МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ,
МОЛОДЁЖИ И СПОРТА УКРАИНЫ
Севастопольский национальный технический университет

В.В. ХОХЛОВ

ИССЛЕДОВАНИЕ И ПРОГНОЗИРОВАНИЕ ЭКОНОМИЧЕСКИХ ПРОЦЕССОВ С ИСПОЛЬЗОВАНИЕМ ЭКСПЛОРАТОРНОГО ФАКТОРНОГО АНАЛИЗА МНОГОМЕРНЫХ ВРЕМЕННЫХ РЯДОВ

Севастополь 2012

УДК 336:519.2
ББК 33.24:В17
Х89

Рецензенты:
д-р экон. наук, проф., заведующий кафедрой Статистики и экономического прогнозирования Харьковского национального экономического университета Е.В. Раевнева;
д-р экон. наук, проф., заведующий кафедрой Экономической кибернетики и экономико-математических методов Санкт-Петербургского государственного университета экономики и финансов С.Г. Светуньков;
д-р экон. наук, профессор кафедры Экономической кибернетики Киевского национального университета им. Т.Г. Шевченко Е.И. Ляшенко.

Научный редактор – канд. экон. наук, доцент, и.о. заведующего кафедрой Учета и аудита Севастопольского национального технического университета Т.М. Одинцова.

Рекомендовано к печати Ученым Советом СевНТУ, протокол № 8 от 29.03.2012 г.

Хохлов В.В.
Х 89 Исследование и прогнозирование экономических процессов с использованием эксплораторного факторного анализа многомерных временных рядов: монография /В.В.Хохлов. – Севастополь: Изд-во СевНТУ, 2012. – 160 с.
ISBN 978–617–612–019–3

В монографии рассмотрены и решены с помощью новых подходов основные проблемы эксплораторного факторного анализа. Предложена авторегрессионно-факторная модель многомерного временного ряда, а также процедура получения прогноза его компонент. Обоснован новый подход к оценке уровня рискованности экономической системы. Разработан метод факторизации симультанных уравнений, с помощью которого получены несмещенные и эффективные оценки параметров системы этих уравнений. Предложенные методы и модели позволили проанализировать состояние экономической системы страны и сделать прогноз основных макроэкономических показателей.

Хохлов В.В.
Х89 Дослідження і прогнозування економічних процесів із використанням експлораторного факторного аналізу багатовимірних часових рядів: монографія /В.В. Хохлов. – Севастополь: Вид-во СевНТУ, 2012. – 160 с.
ISBN 978–617–612–019–3

У монографії розглянуті і вирішені за допомогою нових підходів основні проблеми експлораторного факторного аналізу. Запропонована авторегрессіонно-факторна модель багатомірного тимчасового ряду, а також процедура одержання прогнозу його компонент. Обгрунтовано новий підхід до оцінки рівня ризикованості економічної системи. Розроблено метод факторизації симультанних рівнянь, за допомогою якого отримані незміщені та ефективні оцінки параметрів системи цих рівнянь. Запропоновані методи і моделі дозволили проаналізувати стан економічної системи країни та зробити прогноз основних макроекономічних показників.

УДК 336:519.2
ББК 33.24:В17

ISBN 978–617–612–019–3

СОДЕРЖАНИЕ

ВВЕДЕНИЕ

Мировую экономику с регулярной периодичностью сотрясают разного рода кризисы: финансовые, энергетические, политические, продовольственные и прочие, среди которых наиболее заметным является кризис перепроизводства. У каждого отдельного кризиса всегда есть причина, появившаяся в результате объективных обстоятельств. И, как правило, эти обстоятельства возникают неожиданно, в тот момент, когда что-либо исправить или очень сложно, или просто невозможно. Во многом причиной такой неожиданности являются недоработки экономической науки, не сумевшей вовремя увидеть момент начала формирования негативного явления. Поэтому совершенствование инструментов, как исследования экономических процессов, так и предвидения тенденций их динамики, всегда будет актуальным.

Одна из главных целей экономической науки состоит в том, чтобы изучить объект исследования до такой степени, чтобы опираясь на обретенные знания выявить его основные закономерности развития и на этой основе выработать такие рекомендации, которые будут способствовать поступательному росту экономики и не позволяет ей оказаться в очередном кризисе. В этой связи, наука должна постоянно предлагать новые методы исследования и прогнозирования, а также постоянно совершенствовать существующие научные доктрины.

Одним из таких методов является эксплораторный факторный анализ. В настоящее время он широко используется такими гуманитарными науками, как социология и психология. Этот метод дает возможность раскрыть механизмы взаимодействия различных частей исследуемого объекта. Причем сам объект в социологических и психологических исследованиях – не техническая система, поведение которой можно описать аналитически, он – сверхсложная стохастическая система, поведение которой зачастую не объясняется видимыми причинами. И эксплораторный факторный анализ – фактически, единственный метод, позволяющий результативно исследовать такие объекты.

В исследованиях экономических систем, которые также являются сверхсложными стохастическими системами, эксплораторный факторный анализ практически не используется, и этому имеется несколько причин. Самая главная причина – идентификация решения, которая при традиционных методах факторного анализа допускает неоднозначность: нагрузки факторов при переменных могут быть с помощью методов «вращения» преобразованы к другому виду, а поэтому – факторных решений может быть бесчисленное множество. Другая причина – плохие вычислительные свойства методов оценивания параметров модели, причем методов, которые должны давать, казалось бы, самые надежные результаты. Так, итерационная процедура получения максимально правдоподобных оценок по традиционной методике сходится в исключительно редких случаях. Поэтому на практике используют очень простые методы поиска параметров факторной модели, которые применялись еще в середине прошлого века без привлечения вычислительной техники. В какой степени можно доверять решениям, полученным этими методами, − вопрос открытый, но в исследовании, целью которого является улучшение материальной стороны жизни людей, даже предположение о неточности решения является недопустимым.

Нерешенные или решенные не до конца некоторые проблемы эксплораторного факторного анализа составляют еще одну группу причин, по которым этот метод математической статистики не задействован в исследованиях экономической сферы в должной степени. К таким проблемам можно отнести проблему определения числа факторов, проблему общности исследуемых показателей, проблему вычисления значений факторов и проблему получения однозначного факторного решения. В работе предложено решение этих проблем.

Подход к описанию сверхсложной системы, постулируемый эксплораторным факторным анализом, позволяет сконструировать новый класс моделей исследования многомерных временных рядов. Эти модели позволяют выявить, как авторегрессионные закономерности, так и корреляционные связи между исследуемыми экономическими переменными.

Эксплораторный факторный анализ позволил создать новый подход к решению проблемы оценки уровня рискованности в целом для всей экономической системы. Пространство общих факторов эксплораторного факторного анализа можно рассматривать как пространство экономического риска, на котором может быть задана мера риска. В зависимости от того, в какой части этого пространства находится система, можно судить о степени риска, понимаемого в этом контексте, как степени близости к трудноразрешимым экономическим проблемам.

В работе уточняется понятие «системы одновременных уравнений» и обосновывается более точное ее название на русском языке – «система симультанных уравнений». Показано, что явление симультанности экономических показателей характерно не только для эндогенных переменных, но и для экзогенных. В этом случае обычные методы оценивания коэффициентов симультанных уравнений дают смещенные оценки. Эксплораторный факторный анализ становится основой для метода факторизации симультанных уравнений, дающего несмещенные и более эффективные оценки параметров модели.

Разработанные методы позволили проанализировать поведение стохастических факторов, определяющих развитие экономической системы Украины. Сделан прогноз величин наиболее важных макроэкономических показателей. Построена модель симультанных уравнений экономики страны, на основании которой определены такие значения экзогенных переменных, какие дадут удвоение валового внутреннего продукта, потребления домашних хозяйств и доходов населения, при уменьшении уровня безработицы и инфляции.

1. КОНЦЕПЦИЯ СТОХАСТИЧЕСКОГО ФАКТОРНОГО АНАЛИЗА РАЗВИТИЯ ЭКОНОМИЧЕСКИХ СИСТЕМ

1.1. Методология исследования развития экономических систем

Исследование развития экономических систем тесно связано с его прогнозированием, это, фактически, две стороны познания сути происходящих экономических процессов и явлений. Исследование позволяет вскрыть механизм функционирования экономической системы, на основании которого может быть построен прогностический инструмент. С другой стороны, научное предвидение течения экономического процесса позволяет глубже заглянуть в его существо, что служит целям исследования развития систем.

В исследовании и прогнозировании экономических процессов существует достаточно устоявшаяся терминология, которую приведем ниже.

Исследование научное - процесс выработки новых знаний, один из видов познавательной деятельности. Характеризуется объективностью, воспроизводимостью, доказательностью, точностью; имеет два уровня - эмпирический и теоретический. Наиболее распространенным является деление исследований на фундаментальные и прикладные, количественные и качественные, уникальные и комплексные.

Анализ – общенаучный способ изучения экономических явлений и процессов, экономических систем, сущность которого состоит в разложении исследуемого предмета на составные части и элементы, их отдельное изучение и выяснение связей между ними. Синоним научного исследования [1].

Прогноз – научно обоснованное суждение о возможном состоянии объекта в будущем, о возможных путях и времени его осуществления.

Прогнозирование экономики – процесс разработки экономических прогнозов, которые основаны на изучении закономерностей развития экономических явлений и процессов, выявляет наиболее вероятные и альтернативные пути их развития, и является базисом для выбора и обоснования экономической политики, которая учитывала бы объективные аспекты познанных закономерностей [2].

Экономический процесс – последовательная смена явлений, состояний в развитии экономической системы, в ходе которого возможен ее переход в новое качество, более совершенную форму. Экономический процесс может иметь все формы и виды развития: эволюционный и революционный, прогресс и регресс.

Исследование и прогнозирование экономических процессов приобретают на нынешнем этапе развития экономики и социальных отношений особую значимость и важность. Этот этап характеризуется стремительным развитием научно-технического прогресса, появлением принципиально новых технологий и продуктов, которые коренным образом меняют не только технологии производства, но и образ жизни людей. Яркой чертой настоящего этапа является глубокая глобализация экономики, которая несет не только преимущества международного разделения труда, но и существенные экономические риски. Кризис, появляющийся в одной стране, распространяется далеко за ее пределы, и становится большой проблемой для стран с неокрепшей экономикой. Все это существенным образом сказывается не только на количественные характеристики динамики экономических систем, но и на их качественный характер.

Объектом исследования является макроэкономическая система, содержание которой может быть представлено исходя из различных позиций, связанных с территориальным расположением, видом хозяйственной деятельности, составом элементов и прочее; т.е. макроэкономическая система складывается из отраслей производства товаров и услуг, товарных, финансовых, фондовых рынков, региональных экономических систем. На экономическую систему не могут не оказывать определенных воздействий система государственного регулирования экономических отношений, к которой можно отнести, помимо верховного аппарата законодательной и исполнительной власти, центральный банк, а также фискальные и контролирующие органы.

Таким образом, хозяйственно-экономическая структура на самом высоком уровне представляется собой сверхсложный комплекс взаимоотношений различного вида, в основе которых должны лежать, прежде всего, экономические и финансовые связи, однако на них накладываются и социальные, и политические, и финансовые, и трудовые, и информационные, и даже личностные отношения. Если была бы возможность графически изобразить многообразие связей между всеми экономическими субъектами государственной структуры, то такой рисунок имел бы такие размеры и такую густоту переплетения линий, что даже для поверхностного рассмотрения был бы попросту непригоден, не говоря уже о глубоком анализе имеющих место закономерностей развития экономики. Даже если отбросить многочисленных связи, имеющие стохастический характер, и принять к рассмотрению постоянные, детерминированные отношения, то и в этом случае структурный граф был бы малопригоден для практических целей. Поэтому для рассмотрения таких сверхсложных объектов необходим системный подход, который позволяет не упустить при существенном упрощении представления об объекте наиболее значительные связи и закономерности, а также рассмотреть все определяющие причинно-следственные каналы, рассматривая экономику в целом как единое целое во всем многообразии ее элементов.

Рыночный характер экономических отношений предполагает открытость экономической системы перед другими экономическими системами, в том числе и других государств; и в то же время она является частью более общей общественной системы, представляя собой источник ресурсов, продуктов и услуг для потребления обществом.

Как и любой объект, рассматриваемый с позиции системного анализа, экономическая система может быть охарактеризована с помощью системных объектов: входа, внутреннего процесса, выхода, цели, обратной связи и ограничений.

На вход системы поступают различные ресурсы: природные, трудовые, финансовые, информационные, технологические, научные, продукция и услуги других систем. Выход системы представляет собой совокупность материальных благ и услуг, предназначенных для удовлетворения потребностей населения и других систем.

Окруженная внешней средой, в том числе и обществом, экономическая система не может не быть подверженной определенным воздействиям со стороны этой среды, причем такие воздействия, как природные, так и общественные, являются многообразными. Внешняя природная среда воздействует на экономику посредством обеспеченности ее ресурсами и природными условиями жизни людей, являющихся источником наиболее важного для любой экономической системы ресурса – трудового. Воздействия внешней общественной среды проявляются в мотивах поведения людей, участвующих в производственном процессе, обнаруживаются через морально-этические и идеологические условия, присущие обществу, через национальные и бытовые особенности различных групп населения, и прочее.

Процессы, происходящие в экономической системе, связаны, прежде всего, с ее функционированием. При этом сама системе обладает чрезвычайной сложностью, которая определяется с одной, внешней стороны, – многообразием необходимых для жизненных потребностей продуктов и услуг, с другой, внутренней стороны – продукции, предназначенной для производственных процессов. Фактор совершенствования технических возможностей, в том числе и информационных, как и появление принципиально новых технологий, также оказывает все возрастающее воздействие на экономическую систему, и приобретает решающее значение для ее развития.

Главная часть экономической системы – подсистема производства состоит из различных отраслей, производственных комплексов, промышленных объединений, связанных как горизонтальными, так и вертикальными взаимодействиями. Ее правильная работа не может происходить в отрыве от финансовой подсистемы, также представляющей собой сверхсложный организм, состоящий из субъектов финансовой сферы и финансового рынка.

Функционирование экономической системы не может быть хаотичным и отдаленным от каких-либо целевых установок, причем сам внутренний механизм системы не может выработать такие установки. Отсутствие управления экономикой с внешней стороны, к которой относится и государственная машина, предполагает классическая экономическая теория. Однако лишь классические законы экономики не могут обеспечить бескризисное ведение хозяйства, что показала Великая депрессия 20-30-х годов прошлого столетия. Поэтому корректировка развития экономической системы является объективной необходимостью, что подтверждает финансово-экономический кризис начала двадцать первого века. Корректировка невозможна без формулирования цели управления. Следовательно, цель становится определяющим системным элементом для развития экономики. От сформированной цели в значительной мере зависит то, будет ли экономика расти, куда экономическая система направится для преумножения благ потребления, будут ли эти блага востребованы, удовлетворят ли они вкусы и чаяния населения, восстановится ли социальный мир и воцарится ли спокойствие тех, для кого и существуют экономические системы.

Управление любой системой предусматривает наличие канала обратной связи, по которому проходит объективная, адекватная информация об объекте управления, анализ поступившей информации и выработка регулирующего воздействия, причем в рыночной экономике оно может быть лишь экономическим, но не административно-командным. Схема экономики, как управляемой системы, включающей в себя саму экономическую систему и блок управления, представлен на рисунке 1.1.

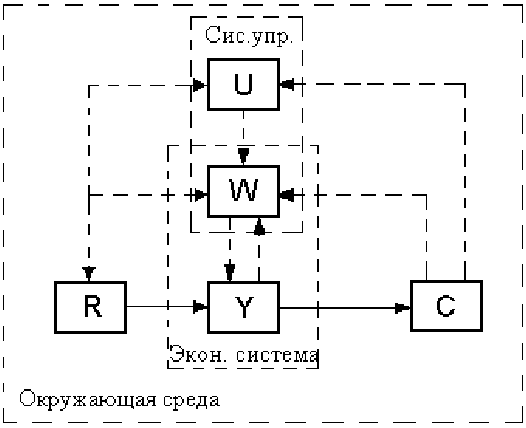

Рис. 1.1 – Экономическая система, как управляемая система

На этом рисунке введены следующие обозначения:

R – внешние ресурсы, поступающие в систему;

Y – производство, включающая в себя подсистему производства, подсистему распределения, подсистему формирования внутренних инвестиций; подсистему развития технологических процессов;

C – конечное потребление;

W – регулятор;

U – блок управления.

Связи RY и YC, отмеченные сплошными линиями обозначают следование материальных ресурсов и продуктов, прерывистые стрелки обозначают информационные потоки, причем CU, CW, YW – каналы обратной связи, UW, WY – управляющие воздействия.

Ресурсы из окружающей среды поступают в производственную подсистему, где производится материальная продукция или услуги. Далее продукт, после того как от него в подсистеме распределения отделен внутренний продукт, предназначенный для нужд производства, направляется на выход экономической системы и служит для конечного потребления, накопления или экспорта в другие экономические системы. В центральном блоке Y, помимо процессов производства и распределения, происходит формирование инвестиционных средств для создания новых производств и модернизации уже существующих. По каналам обратной связи в блоке управления анализируется текущее экономическое положение и принимаются решения о необходимых корректировках, которые материализуются регулятором – блоком W.

Информация, предназначенная для принятия управляющих воздействий, должна быть обработана до такой степени корректно, чтобы она обладала такими свойствами, как достоверность, наглядность, содержательность, обозримость и компактность. В противном случае, когда из всего вороха информационных потоков не выделены золотые зерна истины, принимать ответственные решения невозможно. Поэтому на этапе обработки информации необходимы методы, позволяющие выявить самые существенные закономерности динамики экономической системы.

Целью анализа развития экономических систем является постижение наиболее важных сторон процессов, происходящих в экономической системе, и прогнозирование ее состояния на определенный промежуток времени в будущем. Постановкой, решением задач анализа и прогнозирования развития экономических систем, разработкой методологии экономического моделирования в нашей стране занимаются научные коллективы под руководством В.М. Гейца, В.М. Глушкова, Т.С. Клебановой, Ю.Г. Лысенко, В.И. Елейко, В.Я. Зарубы, А.И. Черняка, В.Н. Соловьева и др.

Экономическая система − совокупность всех экономических процессов, совершающихся в обществе на основе сложившихся в нем отношений собственности и хозяйственного механизма. В любой экономической системе первичную роль играет производство в совокупности с распределением, обменом и потреблением. Во всех экономических системах для производства требуются экономические ресурсы, а результаты хозяйственной деятельности распределяются, обмениваются и потребляются. Характерной чертой любой экономической системы является ее развитие, происходящее вследствие, как внутренних механизмов самоорганизации, так и внешних воздействий окружающей среды, к которой причисляются другие экономические системы и социум. С течением времени меняется состояние системы и характер ее поведения, что чаще всего может быть представлено в виде временного ряда наблюдений или регистраций значений экономических показателей. Это то, что объединяет экономические системы. В то же время в экономических системах есть также элементы, которые отличают их друг от друга, к которым можно отнести: социально-экономические отношения; организационно-правовые формы хозяйственной деятельности; хозяйственный механизм; система стимулов и мотиваций участников; экономические связи между предприятиями и организациями. Все это должно быть отражено при построении модели экономической системы в такой степени, чтобы она отображала наиболее существенные характеристики, но в тоже время не была перегружена. Достичь это можно с помощью адекватных методологических подходов к исследованию сверхсложных систем.

Методология – это логическая организация деятельности человека, состоящая в определении цели и предмета исследования, подходов и ориентиров в его ведении, выборе средств и методов, определяющих наилучший результат. Любая деятельность человека характеризуется методологией. Но в исследовательской деятельности методология играет решающую роль.

Методология любого исследования начинается с выбора, постановки и формулирования его цели. Если объектом исследования является система, предполагающая наличие управления, то в методологическом отношении очень важным оказывается понимать и учитывать класс этой системы. Если это – класс социально-экономических систем, то основополагающим ее элементом является человек, деятельность которого определяет особенности всех процессов функционирования и развития системы. Связи, благодаря которым существует эта система, характеризуют сложные и противоречивые отношения между людьми, основанные на их интересах, ценностях, мотивах и установках. Какими бы совершенными ни были современные технические средства, их роль зависит от интересов человека, мотивов использования и освоения. Система управления не может не строится без учета деятельности человека. Можно исследовать технику, но нельзя ее исследовать в отрыве от человека и всех факторов ее использования в его деятельности.

Для данного объекта исследования предметом исследования является проблема, представляющая собой реальное противоречие, требующее своего разрешения. Функционирование системы управления характеризуется множеством разнообразных проблем, которые выступают как противоречие стратегии и тактики управления, условий рынка и возможностей самой экономической системы. Цель является основой распознавания и выбора проблем в исследовании.

Следующей составляющей в содержании методологии исследования является подход. Подход − это ракурс исследования, исходная позиция, отправная точка, с которой исследование начинается и которая определяет его направленность относительно цели. Подходы могут быть аспектный, системный и концептуальный. Аспектный подход представляет собой выбор одной грани проблемы по принципу актуальности с учетом ограниченности ресурсов, выделенных на исследование. Системный подход отражает более высокий уровень методологии исследования. Он требует максимально возможного учета всех аспектов проблемы в их взаимосвязи и целостности, выделения главного и существенного, определения характера связей между аспектами, свойствами и характеристиками. Концептуальный подход предполагает предварительную разработку концепции исследования, т.е. комплекса ключевых положений, определяющих направленность, архитектонику и преемственность исследования.

Подходы могут быть эмпирическими, прагматическими и научными. Если они в основном опираются на опыт − эмпирические, если − на задачи получения ближайшего результата − прагматические. Наиболее эффективным является научный подход, который характеризуется научной постановкой целей исследования и использованием научного аппарата в его проведении. Методология исследования должна включать также определение и формулирование ориентиров и ограничений. Они позволяют проводить исследование более последовательно и целеустремленно. Ориентиры могут быть гибкими и жесткими, а ограничения явными или неявными.

Главную роль в методологии играют средства и методы исследования, которые можно разделить на три группы: формально-логические, общенаучные и специфические. Формально-логические − это методы интеллектуальной деятельности человека, составляющей основу исследований. Общенаучные методы отражают научный аппарат исследования, определяющий эффективность любого их типа. Специфические − это методы, которые рождаются спецификой систем правления и отражают специфику управленческой деятельности.

Целью методологии прогнозирования является разработка и обоснование программы действий органов управления по формированию и поддержанию действенного хозяйственного механизма в экономической системе, стимулирование воспроизводственной и инновационной деятельности экономических субъектов – элементов экономической системы.

Методология прогнозирования динамических процессов предполагает соблюдение ряда принципов. К общим принципам, не зависящих от конкретного объекта или субъекта прогнозирования, следует отнести следующие:

- принцип системности предполагает сочетание всех элементов прогнозируемой системы;
- принцип согласованности предусматривает соблюдение строгой иерархии, соподчиненности и взаимосвязи всех звеньев системы;
- принцип общности методических подходов − основывается на том положении, что для экономических систем характерны общие закономерности и явления, связанные с особенностями национальной экономики в целом;
- принцип последовательности и непрерывности прогнозно-аналитических расчетов базируется: во-первых, на взаимной увязке оперативных, текущих, среднесрочных и долгосрочных прогнозов; во-вторых, непрерывности прогнозных расчетов, позволяющей субъектам управления своевременно выявлять негативные отклонения от прогнозируемой ситуации и вносить соответствующие коррективы;
- принцип индикативности прогнозов предполагает, что они носят рекомендательный характер;
- принцип вариантности основан на наличии нескольких сценариев (пессимистичного, оптимистичного, оптимального) в построении прогнозных моделей;
- принцип развития производственных отношений предполагает, что при расчетах потенциального роста или замедления в динамике экономической системы следует учитывать качественно новые витки в прогнозируемом периоде.

К специальным принципам, связанным непосредственно с организацией прогнозных работ в конкретной отрасли экономики, следует
отнести следующие:

- принцип индивидуальности означает, что при общности основных тенденций развития каждому экономическому субъекту все же свойственны индивидуальные особенности, которые следует учитывать в прогнозных расчетах;
- принцип сбалансированности основан на адекватном сочетании параметров и показателей прогнозируемых подсистем и системы в целом путем уравновешивания целей роста эффективности и ресурсов для их достижения;
- принцип адаптивности предполагает, что организационно-управленческие системы предприятий под воздействием экономических процессов изменяются, адаптируясь к складывающимся условиям, что, в свою очередь, предполагает соблюдение принципа адекватности;
- принцип адекватности основан на внесении соответствующих изменений в контролируемые государством условия реализации форм и методов ведения хозяйства;
- принцип государственного регулирования и реализации экономической политики предполагает, что прогнозные расчеты должны рассматривать реальные возможности государства воздействовать на развитие отрасли, чтобы не допустить обвальных, неуправляемых процессов в рамках общности целей и задач политики государства, направленной на повышение эффективности национальной экономики.

Необходимость изучения будущего привело к появлению таких понятий, как предвидение, прогноз и прогнозирование.

Суть предвидения состоит в возможности предугадать тенденции будущего развития исследуемого объекта, явления на основе глубокого изучения закономерностей, взаимодействия внутренних и внешних факторов, воздействия случайных событий, рисковых ситуаций, определяющих его динамику с целью выявления их возможных последствий.

Научной основой предвидения, согласно Н. Д. Кондратьеву [3], является сочетание трех подходов, трех ступеней познания: статического, циклично-динамического и генетического. Теория и методология предвидения сформулирована Н. Д. Кондратьевым в его докладах «Перспективы развития сельского хозяйства России» (1924), «Основы перспективного плана развития сельского и лесного хозяйства» (1925) и развита в работах: «Проблема предвидения», «Критические заметки о плане развития народного хозяйства» (1926) и «План и предвидение» (1927) [4].

В теории экономической динамики Н. Д. Кондратьев выделил три типа предвидения социально-экономической действительности [3, с. 105 – 107]. Первый тип предвидения – это предвидение таких событий, которые представляются событиями иррегулярными. Примером таких событий могут служить: конкретные размеры урожая или производства на определенную дату, конкретные размеры экспорта, конкретный уровень цен в определенный момент времени и т. д. Поскольку эти события являются однократными, конкретными, сами по себе они не могут быть включены в ту или иную формулу закона или ряда законов. Для того чтобы предвидеть их точно, нужно располагать почти идеальным знанием хозяйственного положения в исходный момент и почти всей совокупности закономерностей хода окружающих событий. Только тогда можно предсказать интересующие события как результат перекрещивания этих закономерностей. Такой тип вполне конкретного предвидения, если бы он был точным, представлял бы наибольшее практическое значение. Он давал бы наиболее точную ориентировку в возможном ходе событий. Но именно потому, что исследователи обычно не располагают идеальным запасом знаний, такой тип предвидения представляется и наиболее трудным. И на практике, особенно когда такое предвидение делается на более или менее отдаленный срок, в большинстве случаев оно оказывается ошибочным.

Второй тип предвидения имеет место тогда, когда речь идет о предвидении наступления того или иного более или менее регулярно повторяющегося события. Примером такого предвидения могут служить предсказания наступления экономических кризисов, явлений, связанных с сезонными колебаниями конъюнктуры, и т. д. Предсказания такого рода также весьма трудны. Но поскольку идет речь о предсказании событий не в конкретно-количественном выражении, а лишь в форме утверждения вероятного наступления или ненаступления в известный период времени события, возникающего более или менее закономерно и периодически, то предвидение такого типа оказывается в известных пределах более доступным, чем предвидение первого типа. И часто, хотя далеко не всегда, можно по наличию тех или других симптомов за некоторое непродолжительное время предвидеть более или менее точно наступление таких событий.

Третий тип предвидения заключается в предвидении лишь общего развития тех или иных социально-экономических тенденций. Этот тип предвидения не локализует предсказываемых событий точно во времени и не характеризует их в точной количественной форме. В отношении количественной характеристики, где это допускает природа интересующих событий, он может давать лишь приблизительные естественные границы развития тенденций. Примерами такого предвидения могут служить предсказания роста тех или иных отраслей хозяйства, общего повышательного или, наоборот, понижательного движения цен и т.д. Этоттип предвидения, особенно когда речь идет о предвидении в рамках значительного отрезка времени, наиболее доступен при современном уровне социально-экономического знания.

Кондратьев отмечал, что цикличность развития народного хозяйства объективна, ее не следует упускать из рассмотрения и необходимо включать в методику предвидения накладывающиеся друг на друга циклы разной длительности, учитывать их взаимное усиление или ослабевание, а также воздействие факторов неопределенности. При этом может отсутствовать четко выраженная математическая зависимость, вызванная одновременным влиянием множества неравноценных по направлению и силе факторов, порождающих флуктуации. Поэтому в ходе предвидения целесообразно использовать правила [3 , с. 127]:

- необходимо отказаться от количественного выражения тех элементов, предвидеть изменения которых в количественной форме, по крайней мере, на данной стадии знания, вообще невозможно. В таких случаях необходимо ограничиться указанием тенденций соответствующих явлений;
- в тех случаях, когда явление допускает количественное выражение и доступно некоторому количественному предвидению, но лишь на короткий срок, необходимо отказаться от количественного выражения его перспектив на длительные сроки вперед, ограничиваясь на эти длительные сроки лишь характеристикой общих тенденций;
- когда явление допускает известное количественное предвидение на относительно длительный срок, но лишь при условии отказа от дробления его на мелкие составные части, следует при проектировании перспектив на длительное будущее ограничиться рассмотрением явления в суммарном виде.

С учетом перечня возможных факторов строится множество альтернатив будущего развития исследуемой системы, включающее как пессимистичные (с учетом катастрофического риска), так и оптимистичные (с учетом нулевого и отрицательного риска) варианты. Затем заведомо невозможные из них отбрасываются, а наиболее вероятные принимаются как шаги стратегии и тактики принятия решений.

Критерии предсказания будущего развития основываются на анализе действительности и возможностей его стихийного развития, а также наличия объективных предпосылок воздействия на стихийное развитие. Построение реальных перспектив может быть осуществлено, в том случае если проблемы, возникшие в результате случайных событий, могут быть преодолены в любой конкретный промежуток времени. Таким образом, для научного предвидения обоснование перспектив должно быть построено на основе выявленной тенденций случайного хода событий. Поэтому следует опираться на знание выявленных закономерностей в экономической системе ретроспективного периода.

Научное предвидение является важным элементом предсказания будущего, но оно в виде результата дает лишь качественную картинку будущих наиболее возможных событий, и на его основе в большинстве случаев невозможно составить какие-либо количественные расчеты. Такими расчетами занимается научное прогнозирование.

Под прогнозом понимается научно обоснованное суждение о возможном состоянии некоторого объекта в будущем, об альтернативных путях и времени перехода в это состояние [5, с. 11].

Прогноз развития системы невозможен, если отсутствует закономерная связь событий, если исходить из того, что поведение участников социально-экономической жизни непредсказуемо, поскольку возможность прогноза опирается на предпосылку существования всеобщей причинной связи событий [4, с. 124]. Полное знание всех причин и расположение элементов действительности также невозможно, но при этом возможно познание закономерностей развития. Однако учет закономерностей в прогнозировании осложняется тем, что они являются результатом взаимодействия большого числа явлений, и любая закономерность относительна.

Прогнозирование выполняет три основных функции и имеет три стадии [6]:

- предвидение возможных тенденций изменений в будущем той сферы деятельности (объекта, процесса), с которыми нужно иметь дело, выявление закономерностей, тенденций, факторов, обусловливающих эти перемены (исследовательская стадия);
- выявление альтернативных вариантов воздействия на траекторию развития объекта в результате принятия тех или иных решений, оценка последствий реализации этих решений (стадия обоснования управленческих решений);
- оценка хода и последствий выполнения решений, непредвиденных изменений внешней среды, чтобы своевременно скоординировать решения в случае необходимости (стадия слежения и коррекции).

Эти три функции и стадии взаимно переплетены, итеративно повторяются и являются составными элементами управленческой деятельности в любой сфере. Действительно, прежде чем делать выводы о возможном варианте развития экономической системы в будущем, очевидно, следует изучить ее ретроспективу, выявить тенденции замедления и роста, иными словами, провести оценку и анализ объекта исследования. Затем на основе возможных пессимистичного, оптимистичного и оптимального исходов развития объекта в будущем можно построить алгоритм принятия мер управляющего характера с учетом особенностей функционирования и организации каждого подкомплекса. И если предпринятые шаги в общем и хозяйственном менеджменте не оправдали надежд субъектов управления, то необходимо как можно в более короткие сроки посредством механизма обратных связей поменять или откорректировать тактику действий, нацеленную на достижение стратегических задач.

Прежде чем принять то или иное управленческое решение, нужно изучить и предвидеть тенденции изменений в самом объекте социально-экономической науки и окружающей среде, осмыслить возможные последствия – ближайшие и отдаленные – принимаемых решений или совершаемых действий. Без этого результат может оказаться противоположным ожидаемому.

Основное значение прогноза, по словам Н. Д. Кондратьева, лежит в том, что он отвечает настоятельным запросам нашего практического действия в процессе жизненной и социальной борьбы [4, с.117]. «Всюду ставится вопрос об активном вмешательстве в ход событий окружающей социально-экономической среды и вопрос о предвидении хода последующих событий. Вот почему в социально-экономической жизни проблема прогноза имеет особо глубокое практическое значение».

Для прогноза существенны три элемента [4, с. 119]:

1) переход от событий, данных в опыте, к событиям, которые еще не даны в нем;

2) переход к событиям, которые не даны не только потому, что они еще нам не известны, но и потому, что они еще не совершились;

3) переход не произвольный, а научно обоснованный, опирающийся на установленную достаточную для суждения вероятность появления события или событий.

Подводя итог, можно сформулировать основные методологические позиции процедуры прогнозирования экономической системы. Прогноз является непременным условием познания процессов социально-экономической жизни и важнейшим фактором эффективности любых принимаемых решений. Совокупность прогнозов создает научно-аналитическую и многовариантную поисковую базу, поэтому эффективный менеджмент всегда предполагает опережающий, предвидящий подход. Прогнозирование – это самостоятельная, причем первичная функция управления, неразрывно связанная со всеми другими функциями (рисунок 1.1). Цель прогноза – разработка и обоснование программы действий органов управления, направленной на формирование и поддержание действенного хозяйственного механизма в отраслях народного хозяйства, стимулирование воспроизводственной и инновационной деятельности их субъектов, способствующей увеличению производства и реализации продукции и услуг в объемах, удовлетворяющих платежеспособный спрос покупателей и потребности государства.

Достижение цели прогноза предполагает последовательное решение следующих задач:

- проведение анализа влияния различных факторов на динамику экономической системы;
- определение границ неопределенности и риска, связанных с воздействием внутренних и внешних факторов на изменение показателей и критериев развития экономической системы;
- моделирование возможных вариантов развития системы на основе анализа воздействия факторов неопределенности и риска на ее динамику в целом, и адекватного изменения поведения субъектов рынка;
- расчет прогнозов отдельных показателей развития элементов экономической системы;
- обоснование организационно-экономических мер, необходимых для выполнения программ повышения эффективности функционирования системы.

Таким образом, современная концепция методологии прогнозирования предполагает рассмотрение комплекса методов, приемов, подходов, направленных на повышение эффективности и конкурентоспособности хозяйствующих субъектов рынка. Методическая постановка процедуры прогнозирования динамических отклонений и случайных факторов при функционировании экономической системы должна исходить из основных законов в методологии прогнозирования:

- закон системности, в соответствии с которым в прогнозно-аналитических расчетах рассматривается взаимосвязь всей совокупности сложно направленных факторов внешней и внутренней среды;
- закон онтогенеза предполагает, что экономическим системам с элементами управления свойственны определенные циклы, развитие которых можно регулировать и прогнозировать таким образом, чтобы процесс цикличной динамики с учетом случайных факторов носил предсказуемый характер;
- закон композиции, который предполагает, что, с одной стороны, прогноз должен быть детальным, с другой, прогнозно-аналитические расчеты не должны содержать ничего лишнего, рассматривая только основные, принципиальные моменты;
- закон экономической целесообразности, который означает, что прогнозируемые сценарии развития системы должны предполагать эффективное использование имеющихся ресурсов, вовлечение резервов с целью повышения общей эффективности системы.

В настоящий момент существует большое число различных методов и моделей анализа и прогнозирования экономических систем. Если рассматривать только формализованные методы, основанные на моделях, то в качестве инструментария прогнозирования используются: статистические методы [5, 7, 8, 9], детерминированные методы [10, 11], методы эконофизики, вейвлет-анализ, спектральный анализ [12, 13, 14], а также другие методы – метод «Гусеница», генетический алгоритм, самообучающаяся нейронная сеть, метод SSA [15, 16]. Однако ни один из перечисленных методов в полной мере не отвечает указанным принципам и законам методологии прогнозирования. Поэтому создание метода в большей степени удовлетворяющего этой методологии является актуальным.

1.2. Анализ структуры и динамики временных рядов экономической информации.

Исследование и анализ экономической системы в качестве необходимого условия предполагает использование наиболее полной информации о ней. Такая информация представляет собой значения самого широкого спектра переменных величин, характеризующих ее состояние в определенные моменты времени. Как правило – это результирующие показатели за отчетные периоды: итоги дня, недели, месяца, квартала, года и т.д. Совокупность значений отдельного показателя представляют собой временной ряд.

Под временным рядом понимается упорядоченные по времени значения некоторой переменной величины. Причем это понятие существенно отличается от такого широко используемого в статистике понятия, как выборка, представляющая собой некоторый, не зависящий от времени, набор значений случайной величины, отобранных из ее генеральной совокупности. Значения переменной величины из временного ряда естественным образом зависят от фактора времени, и в значительной мере определяются моментом, в который она фиксируется. В противоположность этому, величина элемента выборки, во-первых, никак не связывается с моментом времени, когда она зарегистрирована в ходе какого-либо эксперимента или исследования; и, во-вторых, этим моментом времени, как и любым другим, не может быть определяема. Поэтому повсеместно используемые методы анализа информации и модели реальных процессов, основанные на подходах, базирующихся на обработке выборочных данных, такие как: классический регрессионный анализ, детерминированный факторный анализ, метод наименьших квадратов и многие другие, не подходят к анализу временных рядов. Для исследования объектов, для которых временной фактор является существенным и во многом определяющим, и построения их моделей используются принципиально другие методы, среди которых по частоте применения можно выделить спектральный анализ, стохастический анализ, авторегрессионные и автоковариационные модели и методы.

Временной ряд состоит из двух элементов. Первый – это период времени, в течение которого формировался итог переменной величины (экономического показателя), или в конце которого измерили переменную величину. Второй – само значение переменной величины, называемое уровнем ряда, а именно, ее итог или моментная величина.

Ряды различаются по следующим классификационным признакам:

- форма представления уровня ряда, в соответствии с которой можно различать ряды, составленные из показателей в натуральных единицах измерения (например, в тыс. грн.) и относительных, безразмерных показателей (например, рентабельность);
- характер временного параметра: интервальные (например, количество выданных кредитов за месяц) и моментные ряды (курс акции на момент закрытия торговой сессии);
- расстояние между датами: равноотстоящие (с равными интервалами между соседними парами элементов ряда) и неравноотстоящие, когда интервалы имеют разные длины;
- наличие пропущенных значений: полные и неполные временные ряды;
- характер случайного процесса: стационарные ряды, в которых параметры (главным образом, математическое ожидание и дисперсия) не меняются с течением времени; и нестационарные ряды – у них один из параметров или все параметры меняются с течением времени.

Динамика временного ряда может характеризоваться тремя основными тенденциями, а именно, трендом, циклическими и сезонными колебаниями, при этом последние две называются периодическими составляющими. Тренд представляет собой зависимость случайной величины от времени, имеющая определенный характер – линейный или нелинейный, задаваемая самим случайным процессом. Отсутствие тренда говорит о том, что значения переменной колеблются относительно своего среднего значения, которое все же также можно считать трендом, поскольку на первом этапе анализа временного ряда из него необходимо выделить тренд, в том числе и среднее значение, для исследования закономерностей в отклонениях от линии тренда.

Отходы от тренда могут иметь периодический характер, который часто может быть объяснен очевидной причиной, например зависимость приливов и отливов вкладов населения от времени года: в сезон отпусков и накануне больших праздников вкладчики склоны снимать деньги со счетов для расходования средств на отдых или подарки, а в другие периоды - накапливать.

Циклическая составляющая формируется под воздействием долговременных циклических факторов, которые на протяжении значительного временного периода, измеряемого десятками лет, изменяют социальную, политическую и экономическую атмосферу. Экономические циклы связаны, прежде всего, с переходом на принципиально новые технологии, которые поначалу приводят к всплеску деловой активности, росту производства, а затем эти технологии начинают устаревать, что приводит к спаду деловой активности.

За всю обозримую историю человечество претерпело множество циклических преобразований в экономике: от рабского способа производства в античные времена до подневольного и, фактически, рабского труда в коммунистической системе, которая теоретикам марксизма-ленинизма представлялась вершиной экономического совершенства. Однако, соревнование коммунистического производства в советском варианте и глобализированной мировой экономики оказалось не в пользу первой, и показало всю несостоятельность развития производительных сил, как на основе нерыночных принципов, так и в закрытости от мировых рынков.

Экономические циклы определяются не только сменой социальных формаций, но в большей степени развитием технологической базы производства валового продукта, появлением принципиально новых технологий. Так, третий цикл, в соответствии с теорией Кондратьева [17], длившийся с середины 90-х годов XIX столетия по середину 40-х годов XX века, связан с появлением тяжелого машиностроения, электроэнергетики, неорганической химии, производства стали и электрических двигателей. А четвертый цикл, длившийся с середины 40-х годов до начала 80-х годов прошлого столетия, – связан с революционными достижениями в химической промышленности, в частности нефтепереработки, в массовом производстве автомобилей и других машин. Пятый, еще незавершившийся цикл характеризуется появлением электроники, робототехники, вычислительной, лазерной и телекоммуникационной техники. Однако, эти глобальные экономические волны достаточно сложно учесть при анализе экономических показателей, величины которых рассматриваются на протяжении временного периода, несоизмеримо малого по сравнению с длиной экономического цикла, и, в частности, с длиной кондратьевской волны.

Сезонная составляющая показывает колебания показателя в течение года и является периодически повторяющейся компонентой временного ряда. Свойство сезонности означает, что через примерно равные промежутки времени форма кривой, которая описывает поведение переменной, повторяет свои характерные очертания. После выделения тренда выявить наличие сезонной составляющей достаточно несложно, для чего существует обширный круг критериев [18], к тому же сезонные и, вообще, периодические составляющие выявляются автоматически, если в качестве метода исследования временного ряда используется спектральный анализ.

Для решения практических задач, связанных с прогнозированием поведения одного показателя по значениям одномерного временного ряда, наиболее важным является выявление тренда и сезонной составляющей. Циклическая составляющая, по изложенным выше причинам, не включается в модель.

Исследование тренда содержит два основных этапа: на первом − ряд значений исследуемого показателя проверяется на наличие тренда, на втором − происходит подбор математической зависимости, согласующейся с трендом, и производится выравнивание временного ряда.

При проверке на наличие тренда во временном ряду используются следующие методы и способы [19]:

1. *Графический метод*, когда на графике по оси абсцисс откладывается время, а по оси ординат – уровни ряда. Соединив полученные точки линиями, в большинстве случаев можно выявить тренд визуально.

2. *Метод средних*, согласно которому изучаемый ряд делится на два равных подряда, для каждого из которых определяется средняя величина $\overline{Y}_1$ и $\overline{Y}_2$. И если эти средние величины, в соответствии с формальным статистическим критерием, различаются существенно, то признается наличие тренда.

3. *Метод Кокса и Стюарта*, согласно которому ряд делится на три равные по числу уровней группы и существенное различие выявляется между средними уровнями первой и третьей групп.

4. *Метод Валлиса и Мура*, согласно которому наличие тренда признается в том случае, если ряд не содержит фазы, т.е. перемену знака при определении абсолютного изменения цепным способом.

5. *Метод серий*, согласно которому каждый уровень ряда считается принадлежащим к одному из двух типов, например, один тип составляют уровни меньшие среднего значения элементов ряда, второй тип – больше него. Затем в образовавшейся последовательности типов устанавливается число серий R. Они называются последовательностью уровней одинакового типа, которая граничит с уровнями другого типа. Если в ряду общая тенденция к росту или снижению уровней отсутствует, то число серий является случайной величиной, распределенной приближенно по нормальному закону. Следовательно, если закономерности в изменениях уровней нет, то случайная величина R оказывается в доверительном интервале, определяемом средним значением числа серий $\overline{R}$, среднеквадратическим отклонением и уровнем доверительной вероятности. Тогда с заданной вероятностью тренд имеет место, если установленное число серий ряда не входит в доверительный интервал, и тренд отсутствует, если установленное число серий находится в этом интервале.

Непосредственно сам тренд может быть выделен следующими способами:

1. *Укрупнение интервалов*, когда ряд динамики делят на некоторое достаточно большое число равных интервалов. Если интервальные средние уровни не позволяют увидеть тенденцию, то увеличивают размах интервалов, уменьшая одновременно их число;

2. *Методом скользящей средней*, когда уровни ряда заменяются средними величинами, получаемыми из данного уровня и нескольких симметрично его окружающих уровней. Такие средние называются интервалом сглаживания;

3. *Метод аналитического выравнивания*, когда тренд выражается в виде теоретической функции от времени, и тогда элемент ряда для определенного момента времени выражается посредством значения функции для этого момента и случайной компоненты, в которую входят сезонная составляющая и стохастическое отклонение. Временная функция может быть как линейной, так и нелинейной, причем нелинейность может быть описана полиномом любой степени (чаще не выше третьей), экспонентой и т.д.

Сама функция подбирается так, чтобы минимизировать сумму квадратов отклонений реальных значений от трендовых − теоретических. После того как функция подобрана, оцениваются параметры модели. Если случайные отклонения независимы друг от друга (в этом случае автокорреляционная функция должна быть вырожденной: равняться единице для одного и того же момента времени, и равняться нулю для разных моментов времени, таким образом исходные данные представляют собой выборку), то оценка параметров модели производится с помощью метода наименьших квадратов. Когда же случайные отклонения автокоррелированны, используется обобщенный метод наименьших квадратов.

Аналитический вид сезонной составляющей может быть подобран с использованием ряда Фурье, а также могут быть определены все неизвестные параметры модели, предусматривающей наличие периодической составляющей [20].

При исследовании какого-либо экономического явления чаще всего рассматривается одномерный ряд, представляющий собой наблюдения за одним показателем. Это объясняется наличием достаточно большого числа хорошо разработанных и надежных методов для такого случая, а также простотой вычислительных процедур и наглядностью результатов. Однако полная информация об экономической системе не может быть представлена лишь одним показателем, и чем сложнее система, тем большим числом показателей она характеризуется. В этом случае говорят, что состояние экономической системы описывается многомерным временным рядом. Структура такого ряда является намного более сложной, чем у одномерного ряда. Во-первых, каждая компонента многомерного ряда – переменная – представляет собой одномерный ряд со своим трендом, периодическими составляющими и случайным отклонением. Во-вторых, переменные, наблюдения за которыми составляют многомерный временной ряд, являются взаимозависимыми случайными величинами, природа взаимодействий которых является объективной и особенно характерной для экономических систем.

Таким образом, значение переменной под номером j, $(j = 1, 2, \dots, n)$, которое для момента времени $t, (t = 1, 2, \dots, N)$ обозначим $x_{t,j}$, определяется, с одной стороны, ее значениями в предыдущие моменты времени, что может быть описано авторегрессионными зависимостями; а с другой стороны – поведением и влиянием других переменных. Здесь n и N – соответственно, число переменных, и число наблюдений за ними (длина многомерного ряда).

Существует обширный круг методов исследования многомерных временных рядов, основанных на спектральной теории векторных процессов. Основная идея, объединяющая все эти методы состоит в следующем [21, С.43]. Прежде всего, рассматривается векторный случайный процесс, обладающий свойством стационарности, который сводится к тому, что матрица ковариационных функций обладает свойством инвариантности по отношению к сдвигу временной прямой. Для скалярного процесса это свойство имеет вид:

$$\gamma(s, t) = \gamma(s + \tau, t + \tau),$$

где $\gamma(s, t)$ – ковариационная функция, т.е. ковариация значения некоторой переменной в момент времени t с ее значением в момент времени s;

τ – временной сдвиг.

Для векторного процесса это свойство аналогично:

$$\mathbf{\Gamma}(s, t) = \mathbf{\Gamma}(s + \tau, t + \tau),$$

где $\mathbf{\Gamma}(\cdot)$ – ковариационная матрица, состоящая из парных ковариационных функций между переменными с номерами i и j: $\gamma_{i,j}(s, t)$; $(i = 1, 2, \dots, n; j = 1, 2, \dots, n)$.

При рассмотрении экономических систем в большинстве случаев свойство стационарности не выполняется. К примеру, наличие тренда делает случайный процесс нестационарным. Однако с помощью различных методик нестационарный процесс можно свести к стационарному. К таким методикам, в частности, относится выделение тренда. Поэтому методы классической спектральной теории остаются востребованными при анализе векторных случайных процессов произвольной природы происхождения.

Далее в соответствии со спектральной теорией ставится задача: заменить (аппроксимировать) рассматриваемый стохастический процесс, для которого выполняется свойство $\mathbf{\Gamma}(s, t) = \mathbf{\Gamma}(s + \tau, t + \tau)$, некоторым новым, но для которого ковариационная функция имеет более простую, а именно диагональную, форму, т.е. ковариации для разных моментов времени равны нулю и положительны (не равны нулю) только для одного и того же момента времени. Для этого необходимо подобрать соответствующее преобразование.

В реальных исследованиях ковариационные функции неизвестны, но само преобразование, во-первых, должно диагонализировать ковариационную матрицу, и, во-вторых, должно быть задано a priori. Такое преобразование выполняется методами Фурье, поскольку представление реального векторного процесса в форме ряда Фурье диагонализирует ковариационную матрицу и позволяет, в конечном итоге, составить правдоподобный прогноз поведения системы или процесса, тем самым достигнув конечной цели исследования.

Однако спектральный анализ настроен на изучение процессов, для которых характерны сезонные или любые периодические колебания. Построение спектра позволяет в достаточной мере точно сформулировать основные гипотезы относительно вида модели векторного процесса, основой которой служат повторяющиеся с определенной регулярностью механизмы воспроизведения поведения экономической системы. Таким образом, временные ряды, которые служат исходными данными для анализа, должны иметь достаточную длину. Чтобы выявить ежегодные сезонные составляющие, векторный процесс должен наблюдаться в течение нескольких лет. И, в случае, когда исследуется экономическая система с неустоявшейся динамикой, что характерно для кризисных периодов, подобрать многомерный временной ряд необходимой, для спектрального анализа, длины становится непреодолимо сложно. Это первая причина того, что методы спектрального анализа мало пригодны для изучения экономических процессов в период, отмеченный тенденциями спада, сворачивания производства, массовым банкротством предприятий и элементами стагнации, а также на этапах выхода производства из кризиса. Вторая причина состоит в том, что нестабильной экономической динамике присуще, прежде всего, отсутствие каких-либо значимых периодических колебаний.

Среди распространенных в последнее время методов исследования векторных процессов следует отметить экспертно-аналитический метод, называемый ЖОК [22], предназначенный для оценки результатов влияния переменных на итоговые показатели и друг на друга. Этот метод использует экономико-математическую модель многомерного временного ряда, в которой коэффициенты непосредственного влияния переменных друг на друга и начальные условия задаются экспертами.

В начале исследования в соответствии с этим методом экспертным путем определяется список наиболее существенных переменных, которые необходимо учесть при анализе конкретной ситуации. При этом некоторая часть переменных может носить качественный характер, например, качество продукции, состояние рынка и т.д. После этого определяются необходимые для работы модели уровни переменных, соответствующие начальному состоянию экономического объекта, и для нечисловых переменных проводится оцифровка. Они оцениваются экспертами по некоторой шкале (от -1 до +1). Затем экспертами составляется блок-схема непосредственных влияний существенных переменных друг на друга и оценивается степень непосредственных влияний с помощью той же шкалы. Получается экономико-математическая модель в виде взвешенного ориентированного графа с начальными данными в вершинах. Затем с помощью вычислительной техники просчитываются влияния второго, третьего и т.д. уровней, соответствующие второму, третьему и т.д. моментам времени вплоть до получения стабильного состояния. Результат работы модели − конечные уровни переменных.

Метод ЖОК позволяет просчитать развитие экономической структуры при различных сценариях, а именно: «прогноз», «поиск» и «оптимизация». Сценарий «прогноз» показывает результат при отсутствии управляющих воздействий. Он демонстрирует, как будет развиваться ситуация, если в нее не вмешиваться. В сценариях типа «поиск» анализируются результаты изменений при наличии тех или иных конкретных воздействий на управляющие переменные, и осуществляется эвристический процесс оптимизации, а также анализ поведения системы при тех или иных воздействиях на начальные значения переменных. В сценариях «оптимизация» задаются целевые показатели и условия, которых необходимо добиться, т.е. условия выхода на определенные уровни, например, рентабельность должна быть не менее 40%. С помощью оптимизационных алгоритмов находится наилучшее управление, позволяющее достигнуть цели или максимально к ней приблизиться. Однако найденные компьютером рекомендации могут включать слишком резкие изменения тех или иных начальных параметров, поэтому результаты расчетов скорее указывают на перспективные варианты изменения управляющих параметров, чем непосредственно задают план действий.

Система ЖОК позволяет проследить динамику изменения значений исследуемых переменных вплоть до их стабилизации. При этом факт стабилизации является важным методологическим выводом из экспериментов с моделью ЖОК: «После первоначальных всплесков замкнутая экономическая система стабилизируется, хотя бы и на весьма низком уровне производства и потребления» [22, С. 89].

Методы, основанные на экспертных оценках, безусловно, имеют право на существование. Но мнения экспертов, даже самых эрудированных и подготовленных, являются субъективными, а компилирование нескольких субъективностей для выведения общей оценки экспертов вряд ли может привести к несубъективному результату. Подлинно научный подход к изучению сложных объектов состоит в выявлении объективного образа предмета исследования, что возможно лишь в результате обработки экзогенных данных методами, лишенными какой-либо субъективности. Впрочем, если нет ни таких данных, ни методов, то в качестве единственного источника информации остается мнение экспертов. Однако если объект исследования представлен многомерным временным рядом, в котором по определению содержится полная информация об объекте, то выявить существенные переменные, а также существенные связи между ними можно не прибегая к помощи экспертов, а с помощью объективных и более совершенных методов.

К таким методам можно отнести использование VAR-модели – векторной авторегрессионной модели. Она описывается системой уравнений, число которых равно числу исследуемых переменных. Каждое уравнение представляет собой зависимость данной переменной от значений всех переменных, в том числе и ее самой, в p предшествующих моментах времени (в этом случае говорят, что имеет место авторегрессия порядка p). Модель учитывает влияние, как собственных лаговых значений, так и лаговых значений других переменных. Таким образом, она позволяет установить и аналитическую авторегрессионную зависимость переменных, и влияние других переменных на текущее значение каждой из них.

Однако если исследуются n переменных, и порядок авторегрессии равен p, то число подлежащих оценке коэффициентов равно $(n + p \cdot n^2)$. К примеру, при $n = 5, p = 4$ необходимо найти 105 значений, и это очень большое число [5, С. 293]. Чтобы рассмотреть экономический процесс на основе принципа системности, то потребуется ввести в модель более нескольких десятков переменных и даже при небольшом порядке авторегрессии, скажем $p = 2$, число коэффициетов минимум составит 820. Чтобы получить состоятельные оценки параметров модели, длина ряда должна быть больше такого числа. Этот факт делает метод мало пригодным для практического использования в анализе многомерных временных рядов.

Чтобы обойти эту проблему, VAR-модель трансформируют к виду классического регрессионного уравнения, связывающего данную переменную с лаговыми значениями всех переменных, сведенных в одну матрицу. Далее, применяя n раз метод наименьших квадратов, находят оценки всех коэффициентов исходной модели. Очевидно, такая процедура дает результат, но можно ли ему верить?

Применение метода наименьших квадратов к оценке параметров регрессионного уравнения обусловлено несколькими априорными допущениями. Одно из них состоит в том, что регрессоры должны быть независимыми величинами. Однако в многомерном временном ряде переменным присущи корреляционные связи, и считать столбцы матрицы регрессоров взаимно независимыми будет некорректно. Таким образом, имеет место нарушение одного из предположений классической регрессионной модели. Далее, еще одно предположение регрессионного анализа говорит о том, что независимые переменные и случайные отклонения уравнения не должны коррелировать. Однако, также как и для системы одновременных уравнения, у которой объясняемые и объясняющие переменные в разных уравнениях входят, как в правую часть, так и в левую часть, и поэтому ряд объясняющих переменных, стоящих в правой части уравнений, зависит от случайных ошибок, по отношению к предлагаемому подходу трансформации VAR-модели можно показать, что некоторые регрессоры коррелируют со случайными отклонения. Следовательно, нарушается еще одно предположение регрессионного анализа. А каждое нарушение предположения регрессионного анализа требует корректировки как модели многомерного временного ряда, так и метода оценки ее параметров.

Таким образом, создание модели, учитывающей авторегрессионные зависимости и корреляционные связи между переменными, а также разработка методов получения оценок параметров такой модели, отображающей процессы неустоявшейся динамики, являются актуальными.

1.3. Эксплораторный факторный анализ развития экономических систем

Глубокое исследование экономических процессов, явлений и объектов невозможно без привлечения методов и средств математической статистики, использование которых в экономической сфере объединяется под общим названием – «эконометрия», и эксплораторный факторный анализ является неотъемлемой ее частью.

Эконометрия является математическим аппаратом экономического анализа, с одной стороны. С другой стороны, в силу обширного числа различных методов ею используемых и развиваемых, эконометрия давно уже рассматривается как самостоятельная наука. И эта наука позволяет:

- всесторонне анализировать и оценивать сконструированные с помощью экономической теории и экономических подходов математические модели объекта исследования;
- на основе аппарата теории статистического вывода проверять различные гипотезы об объекте исследования, тем самым устанавливать истинность экономической теории;
- получать экономически обоснованные и статистически достоверные прогнозы развития экономических показателей, что является научной базой принятия решений на различных уровнях: начиная от обоснования хозяйственного решения на предприятии, и заканчивая выбором стратегии экономического развития государства.

Эконометрия является сочетанием двух областей знаний: экономической теории, которая часто использует математические формулировки, и математической статистики. Эконометрия позволяет получить числовое решение практической проблемы, причем в ее основе лежат не любые теоретические допущения, а подлинные факты – статистика наблюдений за реальным объектом. Именно ориентация на достоверность происходящих событий является главным стержнем эконометрического подхода к исследованию экономического явления. И полное соответствие модели, получаемой как итог изучения экономического объекта, к действительности представляет собой главный критерий качества эконометрического исследования.

Все сказанное относится и к эксплораторному факторному анализу

К особенностям использования этого раздела эконометрии следует отнести:

1) большое число экономических показателей, описывающих процесс или объект и являющихся случайными величинами;
2) взаимосвязь или коррелированность этих переменных между собой;
3) неизвестная природа формирования случайности во взаимосвязях между показателями.

Первая особенность определяется системным подходом, характерным для эксплораторного факторного анализа. Согласно этому подходу любой процесс или объект должен быть отображен в описании во всем многообразии как внутренних, так и внешних связей. Безусловно, полная детализация структуры явления приведет ее к бесконечномерности. И все же, на первом этапе исследования ни один аспект в функционировании явления, ни один показатель, характеризующий его особенности, не должен быть проигнорирован. Такой подход ведет к формированию огромного объема информации, обработка которого не мыслима без использования современных эконометрических методов.

Вторая особенность возникает в силу известного постулата, согласно которому все экономические переменные связаны между собой. Причем связь между двумя переменными может быть прямой: скажем, предложение некоторого продукта напрямую влияет на его цену. Но связь между переменными может быть и опосредственной: затраты на рекламу вряд ли непосредственно влияют на предложение продукта, но стимуляция спроса на него, благодаря рекламе, с неизбежностью приведет к росту предложения на этот товар.

Третья особенность определяется случайностью всего происходящего в этом мире, причем о природе происхождения случайности заранее ничего не известно. А именно, трудно наперед сказать, что определило числовое значение экономического показателя в данный момент: то ли его значение в предыдущие моменты времени, то ли влияние на него в данный момент других факторов, а может быть и то и другое; а может быть влияние обстоятельства, не попавшего в поле зрения исследователя. Природа случайности несчетна, трудно даже предусмотреть все аспекты ее проявления.

Эконометрическое исследование традиционно имеет несколько этапов. На первом этапе, исходя из общеэкономических подходов, выделяются все возможные связи между экономическими переменными, а затем на основании статистического материала и, исходя из целей исследования, выбираются наиболее важные из них для дальнейшего включения в модель. Например, рыночный спрос и предложение должны объясняться соотношениями между объемом выпуска, объемом продаж, ценой предлагаемой, ценой запрашиваемой, потребительским доходом.

Включить в модель все переменные, чтобы объяснить поведение каждой из них, невозможно, поскольку всегда останется, хотя бы одна, внешняя переменная, не подлежащая объяснению, и это делает модель неполной или условной. Причем увеличением числа переменных и числа уравнений преодоление условности модели представляется невыполнимым.

На втором этапе выбранные соотношения группируются в модель, математические особенности которой зависят от характера экономических процессов. К примеру, модель спроса и предложения представляет собой систему одновременных уравнений: уравнения спроса, уравнения предложения и уравнения реакции рынка. Это линейные уравнения первой степени, но со стохастическими составляющими. Модели исследования финансового риска предполагают использование аппарата теории временных рядов, и, в частности, рядов Фурье. Динамические модели трудно представить без применения дифференциальных или разностных уравнений.

Эксплораторный факторный анализ позволяет объединить эти два этапа в один, при этом нет необходимости убирать из рассмотрения какие-либо экономические показатели. Факторная модель позволяет рассматривать экономический процесс со всеми регистрируемыми переменными, не смотря на очевидную возможность появления эффекта мультиколлинеарности.

Неуклонное следование требованиям эконометрического подхода приводит к тому, что все эконометрические модели обладают общими свойствами:

- поведение экономически переменных определяется с помощью совместных и одновременных операций с некоторым конечным числом математических соотношений;
- модель, допускающая упрощение действительности, тем не менее, улавливает главные характерные черты изучаемого объекта;
- на основе достигнутого с ее помощью понимания реальной системы всегда можно предсказать ее будущее движение и управлять этим движением.

На третьем этапе статистические данные используются для оценки параметров модели. Причем выбор метода получения оценок также является достаточно ответственным моментом эконометрического исследования. Этот метод должен дать наилучшие для данного фактического материала оценки. Традиционно, к таким статистическим методам относятся метод максимального правдоподобия и метод наименьших квадратов. Хотя специфика статистических данных может сделать невозможным применение одного из них. К примеру, при наличии мультиколлинеарности в исходных данных использование метода наименьших квадратов для получения оценок регрессии приведет к существенным искажениям картины реального объекта и к фатальным ошибкам в принятии решений на основе такой модели.

На последнем этапе осуществляется проверка оцененной модели на адекватность к действительности. На основании ретроспективного прогноза делается вывод о достаточной реалистичности получаемой с ее помощью картины объекта и возможности использования ее в достижении целей исследования, или модель признается неадекватной. В последнем случае эконометрическое исследование должно быть воспроизведено с более ранних этапов, чтобы в конечном итоге получить соответствие модели и реального объекта.

К эконометрическим относится обширный круг методов математической статистики. Первым из них был классический регрессионный анализ. Применение этого метода обширно, но в экономике оно ограничивается достаточно жесткими предположениями о свойствах неизвестных случайных возмущений со стороны внешних факторов, не объясняемых с помощью модели. Так, наличие гетероскедастичности (неоднородности дисперсии ошибок уравнения регрессии) или наличие автокоррелированности ошибок, или наличие сезонности в исходных данных, делает невозможным применение этого метода в исследовании экономических явлений.

Развитие регрессионного анализа как эконометрического метода, идет по пути преодоления противоречий между природой экономического объекта и допустимыми предпосылками использования математики в объяснении естества явлений.

Близко по идеологии к регрессионному анализу расположен другой важный эконометрический метод моделирования экономических процессов – использование систем одновременных уравнений. Однако, в отличие от регрессионных моделей, в такой системе должна быть задана жесткая структура уравнений, отвечающая критериям идентифицируемости, причем одна и та же экономическая переменная может выступать как в роли результирующего показателя, так и в роли объясняющей величины, – в регрессионных уравнениях это невозможно. Формирование системы одновременных уравнений может быть проделано лишь на основании четких экономических представлений о природе изучаемого явления.

Регрессионный анализ и системы одновременных уравнений являются центральными методами эконометрии, однако, не менее важными являются другие разделы эконометрии: теория временных рядов (хотя этот раздел в силу своей важности и объемности рассматривается как самостоятельная наука), дискриминантный анализ, дисперсионный анализ; линейное, стохастическое, нелинейное и дискретное программирование, теория массового обслуживания, теория стохастических процессов, метод Монте-Карло, кластерный анализ и, наконец, факторный анализ.

В отечественной экономической литературе очень часто под факторным анализом понимается раздел экономического анализа конкретных показателей, чаще всего на основании индексной мультипликативной модели. И в этой связи такой факторный анализ не принято считать отдельной научной дисциплиной. В математической статистике под факторным анализом принято понимать нечто иное. А именно, словарь-справочник «Математика и кибернетика в экономике» [23] приводит такую формулировку: «Факторный анализ – направление многомерного анализа, которое исследует структуру матриц ковариаций и корреляций». Поэтому факторный анализ является ветвью математической статистики, и когда с его помощью исследуется экономический объект, факторный анализ становится эконометрическим методом. Для того, чтобы отличать этот статистический метод от других, имеющих такое же название, принято называть его «эксплораторный факторный анализ» − исследовательский факторный анализ. У стохастического факторного анализа существует и другая его разновидность – «конфирматорный факторный анализ», т.е. подтверждающий факторный анализ. Эта ветвь получила широкое распространение в психологических и социальных исследованиях.

Однако полные возможности эксплораторного факторного анализа, как метода познания мира, в должной мере пока еще не реализованы. Попытаемся указать на его потенциал.

Наука является процессом, состоящим из непрерывного цикла построения моделей наблюдаемой действительности, которая вначале кажется необозримой в своем многообразии и не может быть хоть как-нибудь охвачена моделью. В естественных науках наряду с накапливанием наблюдений происходит формирование теоретических положений, которые затем проверяются на новых данных и постепенно совершенствуются.

Математика предоставляет возможность построения неограниченного числа моделей, с помощью которых можно точно описывать действительность. На ранних этапах научного познания реального мира приходилось полагаться на интуицию и гениальность отдельного исследователя, который составлял такую модель и проверял ее соответствие действительным условиям. Однако эти обе функции исследователя, а именно построение модели и проверку ее адекватности, можно систематизировать и формализовать.

Теоретические положения физики, химии и других естественных наук в значительной степени основаны на математических моделях. Для этих наук характерно то, что в условиях одного эксперимента в принципе можно проверить определенную закономерность путем манипулирования величинами в широких пределах, пользуясь их причинной обусловленностью. В науках, занимающихся изучением поведения человека, к которым относится психология, социология, частично медицина и многие смежные области других отраслей знаний, где отсутствует четкая функциональная связь между причиной и следствием, один опыт не играет такой решающей роли. Присущие им закономерности можно обнаружить лишь путем статистической проверки результатов многих опытов, принимая или отклоняя определенные гипотезы. Процедура статистической проверки начинается с формулировки нулевой гипотезы («нет никакого различия») и альтернативной («имеет место различие»). Затем на основании выборки производится проверка нулевой гипотезы относительно альтернативной с помощью соответствующего критерия. Результатом статистической проверки является вывод о том, в скольких случаях на каждые 100 проведенных испытаний в предположении определенной модели отклонения от нее можно считать случайными. Может быть определена вероятность ошибочного принятия альтернативной гипотезы. Вероятность допустить ошибку при принятии нулевой гипотезы тоже имеет вполне определенное значение, называемое уровнем значимости. Таким образом, исследователь на заданном уровне значимости может остановиться на одной из этих гипотез. Имеется большой набор критериев значимости, с помощью которых осуществляется проверка различных гипотез. Выполнив процедуру проверки статистических гипотез, исследователь приходит к выводу о возможности принятия определенной модели. Для этой процедуры важно, на основании каких соображений исследователь определяет вид модели, описывающей данное явление.

В то время как статистическую проверку гипотез можно рассматривать в качестве попытки систематизировать процедуру принятия решений на фоне варьирующих наблюдений, статистическое оценивание можно считать попыткой систематизировать интуицию исследователя. При статистическом оценивании в предположении определенной математической модели речь идет о том, чтобы по возможности точно определить истинное значение параметра генеральной совокупности исходя из выборочных наблюдений. Примерами статистического оценивания являются предсказание исходов выборов и оценка величины брака в промышленном производстве. Если соблюдены предпосылки выборочного метода и принята определенная модель, то можно приступать к построению доверительных границ, определяя интервал, который, например, в 99 из 100 случаев покрывает истинное значение параметра генеральной совокупности.

В попытке систематизировать интуицию эксплораторный факторный анализ представляет собой шаг вперед по сравнению с методом статистического оценивания. Этот анализ является методикой, которая в определенном смысле сама является источником возникновения гипотез.

Рассмотрим сначала специфический характер гипотез, порождаемых факторным анализом. Исходной предпосылкой служит то, что несколько измеряемых переменных сильно коррелируют между собой. Это означает, что либо они взаимно определяют друг друга, либо связь между этими переменными обусловливается какой-то третьей величиной, которую непосредственно измерить нельзя. Модель эксплораторного факторного анализа всегда связана с последним предположением, т.е. измеряемая величина является лишь формой проявления величины, остающейся на заднем плане и не поддающейся непосредственному измерению. Это предположение во многих случаях вполне реально. Возникает задача, можно ли по данным переменным выделить величину, так называемый *фактор,* который объяснил бы наблюдаемые связи. Слово *фактор* используется в другом смысле, чем это принято в отечественной экономической литературе: речь идет о математической величине, получаемой на основе наблюдений. Спрашивается, какова самая простая линейная гипотеза о структуре этой величины, скрывающейся за коррелированными переменными. Сколько нужно гипотетических величин, или факторов, и каких, чтобы по возможности наиболее точно воспроизвести и объяснить наблюдаемые связи между переменными? Каким образом свести большое количество данных к возможно более простой концепции с минимальной потерей информации?

Получить ответы на такие вопросы — сокровенная давняя мечта исследователей. Эксплораторный факторный анализ пытается это сделать. Образно говоря, он заглядывает за кулисы того, что непосредственно измеряется, и стремиться определить истинные функциональные величины, лежащие в основе данного явления. Основная цель эксплораторного факторного анализа состоит в выявлении гипотетических величин, или факторов, по большому числу экспериментальных данных. Факторы должны быть по возможности простыми и достаточно точно описывать и объяснять наблюдаемые величины. Итак, эксплораторный факторный анализ является методом, упорядочивающим кажущуюся хаотичность изучаемого явления и генерирующим новые гипотезы. Число выделяемых факторов должно быть меньше набора исходных величин, структура этих факторов и их взаимосвязь должны быть по возможности более простыми. Возникает вопрос: что же находится за кулисами экспериментальных данных и как может выглядеть система выделенных факторов в самом простом случае?

Исходной предпосылкой анализа является наличие взаимосвязи между несколькими одновременно наблюдаемыми переменными. В качестве количественной меры связи между двумя переменными используется коэффициент корреляции. Он может принимать значения от –1 до +1. При этом, если он приближается к нулю, то это свидетельствует об отсутствии линейной связи, и чем более он близок к +1 или к –1, тем более тесная линейная связь существует между переменными. Все вычисленные коэффициенты корреляции располагаются соответствующим образом в корреляционной матрице. В этой матрице содержится важная информация о взаимоотношениях переменных с учетом влияния помех, причиной которых может являться, например, неоднородность материала. При анализе такой корреляционной матрицы получают гипотетические величины, так называемые факторы, которые находятся в определенных взаимоотношениях с переменными. Примечательно то, что факторный анализ делает возможным выдвижение дифференцированных гипотез о структуре взаимосвязи переменных и факторов, не задаваясь этой структурой заранее и не имея о ней никаких сведений. Эта структура находится по результатам наблюдений. Полученные гипотезы могут быть проверены в ходе дальнейших экспериментов.

Итак, задачей эксплораторного факторного анализа является нахождение простой структуры, которая бы достаточно точно отражала и воспроизводила реальные, существующие в природе зависимости. При этом на реализацию принципа простой структуры оказывают сильное влияние вид и количество экспериментальных данных. Односторонний подход к подбору переменных неизбежно приведет к таким факторам, которые не будут соответствовать изучаемому явлению. Результат эксплораторного факторного анализа определяется постановкой всего исследования. Нельзя надеяться на то, что хаотически собранные данные, подвергнутые эксплораторному факторному анализу по имеющейся стандартной программе, дадут осмысленные результаты. Но не исключено, что конечный результат эксплораторного факторного анализа может оказаться неудовлетворительным также в том случае, когда сбор данных проводился целенаправленно, согласно намеченному плану исследования.

Как в любой прикладной науке, здесь следует обращать внимание на различие между математической моделью и реальным содержанием изучаемого явления. Алгебраически вычислительная сторона метода, в которой речь идет только о решении системы уравнений и точности вычислений, является лишь одним аспектом проблемы. Для эксплораторного факторного анализа характерен также статистический подход, применяемый, например, при проверке гипотезы о числе факторов, подлежащих выделению. Наряду с этим существует еще проблема содержательной интерпретации выделенных факторов, что не имеет места при построении математико-статистической модели. Три вышеназванных аспекта – алгебраически вычислительную сторону, статистический подход и интерпретацию факторов – следует учитывать при проведении эксплораторного факторного анализа и разграничивать их.

Применение эксплораторного факторного анализа особенно эффектно в таких областях, где невозможно манипулировать наблюдаемыми переменными. Исследование поведения человека или его физиологических параметров, а также различных общественных явления нельзя проводить классическим способом, поддерживая на одном и том же уровне определенные величины или вводя различные ограничения. Однако для большинства явлений природы характерна совместная вариация признаков, которая в неизменном окружающем мире дает возможность после точного анализа делать выводы, аналогичные результатам классических экспериментов.

Классическим примером применения эксплораторного факторного анализа является теория интеллектуальных возможностей. Можно предположить, что интеллектуальные возможности состоят из большого числа параметров, характеризующих различные способности человека, такие, как память, внимательность, компетентность и т. д. Этим способностям с помощью психологических тестов пытаются дать количественные оценки и определяют их взаимосвязь. Тесты способностей, как правило, характеризуются положительными корреляциями. Спирмэн [24] предположил, что все возможные совокупности корреляций обусловливаются одним генеральным фактором g, влияющим на все переменные. Этот генеральный фактор совпадает с тем, что обычно в обиходе называют смышленостью. После его психометрической изоляции можно определить, какие методы измерения более всего пригодны для его оценки, т.е. вычисляют нагрузки этого фактора по различным тестам. Обширные исследования показали, что одного генерального фактора недостаточно для описания наблюдений над интеллектуальными способностями. Может быть постулирована более сложная гипотеза, допускающая несколько факторов для адекватного статистического описания интеллекта. Это достигается путем введения новых методов измерения, связанных с решением арифметических задач, с проверкой способностей к абстрактному мышлению, с проверкой словарного запаса и т. д. Другими словами, методы измерения представляют собой набор тестов, применяемый к репрезентативной выборке индивидуумов. Этим самым исследователи получили в руки измерительный инструмент для оценки гипотетической величины интеллекта, расчлененного на отдельные факторы. С помощью факторного анализа было показано, что интеллектуальные возможности человека можно представить в виде эмпирических показателей, а не только в виде слов или понятий. Эти показатели имеют количественные оценки. Понятие интеллекта заменяется рядом факторов, чьи отображения устанавливаются с помощью методов факторного анализа. На основе психологических тестов, выявляющих интеллектуальные возможности человека, можно выделить набор факторов, которые большей частью коррелированны положительно, например факторы пространственного воображения, понимания слов и т.д. С помощью факторного анализа эмпирическим путем проверялись различные психологические теории интеллектуальных возможностей и умственных способностей человека. Такие исследования нашли практическое применение при оказании помощи в выборе профессии, при распределении рабочей силы по имеющимся рабочим местам, в решении проблем воспитания детей в семье и школе, в психотерапии и т. д.

Аналогично исследованию интеллектуальных возможностей с помощью факторного анализа многие ученые занимались изучением индивидуальных особенностей. Среди этих ученых, прежде всего, следует назвать Каттелла [25], который с помощью почти 2000 различных психологических тестов выделял в многочисленных экспериментах от 50 до 60 факторов. Каждый из этих факторов оценивался с помощью набора психологических тестов, примененных к отдельному испытуемому.

Различные теории, которые возникли в результате использования факторного анализа, хорошо согласуются с известными клиническими наблюдениями и делают возможной классификацию людей по их характеру. Например, с помощью набора тестов можно определить, является ли испытуемый неврастеником или обладает бодрым и детерминированным характером, причем опыт показал, что результаты тестов соответствуют заключениям врачей.

Стохастический факторный анализ применяется и в других областях знаний, например, в социологии – для анализа результатов выборов. Но следует отметить, что факторный анализ как научная методика оправдал себя лишь в отдельных направлениях психологии. Аппарат факторного анализа привлекают в первую очередь к исследованию таких явлений, где необходимо выявить наиболее простую структуру по большому числу наблюдаемых переменных. По мере накопления наблюдаемого материала выявить структуру только с помощью логических рассуждений становится труднее и на это затрачивается много времени. Факторный анализ позволяет произвести упорядочение данных, коротко описать взаимоотношения между переменными и получить вспомогательный материал для проверки интуитивных соображении исследователя.

Эксплораторный факторный анализ занял свое особое место среди других эконометрических методов в силу специфики и особенностей факторной модели. Кендэл [26] указывает в качестве исходной позиции при разграничении многомерных методов различия в понятиях «зависимость» и «взаимозависимость» одной или нескольких переменных от остальных. Конкретная постановка вопроса позволяет определить, что является результативной и факторной величиной. Самым простым примером теории, занимающейся исследованием зависимости, является регрессионный анализ. Кендэл сюда же относит дисперсионный и ковариационный анализ. В противоположность этому под взаимозависимостью понимается исследование связей между несколькими переменными без выделения результативных факторных величин. Сюда относится корреляционный и ассоциативный анализ, анализ функциональных связей, компонентный и факторный анализ. Дискриминантный анализ и вычисление канонических корреляций нельзя отнести ни к одному из этих видов анализа.

Так как эксплораторный факторный анализ исходит из коэффициентов корреляции или ковариации, то его следует считать статистическим методом. Однако определение главных компонент корреляционной матрицы является вычислительной абстракцией, которая не распространяется на генеральную совокупность. Но аналогично определение среднего значения и регрессионной прямой тоже является вычислительной процедурой, осуществляемой по определенному алгоритму. Однако такое представление данных через среднее значение и регрессионную прямую является одной из статистических задач.

В многомерном регрессионном анализе задаются целью получить оптимальную оценку зависимой переменной или результирующего показателя исходя из нескольких независимых переменных, которые точнее будет называть регрессорами. Последние могут зависеть друг от друга, т.е. коррелировать, однако в регрессионном анализе одним из основных предположений является то, что ошибки в измерениях регрессоров, как случайных переменных, пренебрежимо малы по сравнению со случайными отклонениями зависимой переменной.

Результирующий показатель оценивается в целях прогнозирования так, чтобы погрешность была по возможности минимальной. В противоположность этому в эксплораторном факторном анализе пытаются определить новые переменные, так называемые факторы, величины которых могут быть оценены для каждого момента времени в будущем. Если в качестве исходных данных используются не временные значения переменных, а величины показателей ряда различных объектов, то величины факторов могут быть найдены для каждого такого объекта. В регрессионном анализе ограничиваются сопоставлением изменений, лежащих на поверхности явлений, в факторном анализе пытаются заглянуть за кулисы явлений, обнаружить основные причины, лежащие в основе этих изменений, и произвести их оценку в каждом отдельном случае. Модель факторного анализа более сложна. В эксплораторном факторном анализе все переменные одинаково важны и все оцениваются, в то время как в регрессионном анализе производится расчленение переменных. По всем переменным определяется минимально возможное число независимых факторов, которые затем могут быть оценены для всех рассматриваемых объектов.

Оба метода различаются постановкой целей и ни в коем случае не противопоставляются друг другу. Более того, использование подходов и модели факторного анализа в регрессионном анализе может решить ряд существенных проблем последнего.

Существенное различие между обоими методами состоит в том, что эксплораторный факторный анализ не довольствуется данными переменными, а ищет те немногие величины, которые воспроизводят переменные, взаимосвязь между ними и объясняют их.

В то время как в регрессионном анализе предполагается различие между исходными величинами и целевой функцией, т.е. производится разгруппировка переменных, в дисперсионном анализе исходят из данных, уже сгруппированных по различным объектам. Оценка расхождения между двумя средними значениями выполняется с помощью t-критерия. В рамках дисперсионного анализа значительно расширяются возможности проверки гипотезы существенности различия между парами средних. В случае однофакторного комплекса проверяется гипотеза, является ли существенным различие между средними значениями отдельных групп и средним значением всей совокупности. Вид гипотезы зависит от постановки задачи и от исходных данных.

Большим достоинством дисперсионного анализа является возможность исследования влияния сразу нескольких факторов. В многофакторном дисперсионном анализе проверяется гипотеза различия между векторами средних значений.

Термин «фактор» общепринят в дисперсионном анализе и означает наблюдаемую величину, что отличается от толкования этого термина в факторном анализе.

Постановка задачи в эксплораторном факторном анализе существенно отличается от ее постановки в дисперсионном анализе. Задачей дисперсионного анализа является исследование влияния одного или нескольких факторных признаков и их взаимодействий на результативный признак и оценка этого влияния. При этом отдельные значения результативного признака увязываются с выбранной моделью. Заранее задаются факторы, число которых ограничено, и их уровни.

В эксплораторном факторном анализе переменные исследуются в отношении их взаимосвязи. При этом заранее не постулируется, что наблюдаемые переменные полностью представляют исследуемую область. Напротив, ставится задача определения количества и вида линейно-независимых друг от друга переменных (факторов), которые бы достаточно точно воспроизводили взаимосвязи наблюдаемых данных. Эксплораторный факторный анализ пытается выявить существенные величины, которые определяют вариацию большого числа переменных. В то время как в дисперсионном анализе имеют дело с одной или несколькими переменными, заданными заранее, и по разработанной методике проверяют их влияние. В эксплораторном факторном анализе пытаются через имеющиеся, ничем не ограниченное число наблюдаемых переменных воспроизвести величины полностью их объясняющие, и оценить эти величины для отдельных объектов.

При конкретных исследованиях имеет смысл оба метода комбинировать. В самом начале, при неопределенной еще структуре исследуемой области, полезно провести факторный анализ, чтобы выявить факторы, вызывающие рассеяние, а потом экспериментировать с ними или, например, провести дисперсионный анализ, чтобы доказать их влияние на третьи величины. Кроме того, можно объединить обе концепции в одном эксперименте и при получении оценок факторов.

Выработка правила, позволяющего приписать некоторый элемент к одной из двух групп, является основной задачей дискриминантного анализа. При такой постановке вопроса должны иметься значения m переменных для n объектов, полученные в результате наблюдений. Кроме того должна быть известна принадлежность каждого объекта к одной из двух групп. К какой из двух этих групп теперь относить новые объекты исходя из имеющейся информации? Ответ на вопрос упирается в подбор такой линейной комбинации переменных, которая бы оптимально разделила обе группы. В геометрических терминах задача сводится к определению положения новой оси координат, удовлетворяющего следующему условию: проекции объектов двух заранее известных групп должны казаться, возможно, более разделенными. Процедура расчетов не ограничивается двумя группами объектов, ее можно распространить на несколько групп в многомерном пространстве.

Существенное отличие от эксплораторного факторного анализа заключается в том, что в дискриминантном анализе определяется разделение точек, причем заранее уже известно, какие точки к какой группе должны относиться. С помощью эксплораторного факторного анализа в принципе такие группы можно найти.

В то время, как дисперсионный и дискриминантный анализ имеют немного общего с эксплораторным факторным анализом, метод главных компонент, или компонентный анализ, тесно связан с факторным анализом. Этот метод позволяет при заданной n-мерной корреляционной матрице найти новую ортогональную n-мерную систему координат и именно так, чтобы максимум полной дисперсии лежал в направлении первой главной оси, а максимум оставшейся дисперсии – в направлении второй главной оси и т.д. Компонентный анализ является важнейшим методом выделения факторов. Однако его следует строго отличать от эксплораторного факторного анализа как такового.

Метод главных компонент заключается в нахождении последовательности ортогональных осей координат, вдоль которых каждый раз в убывающем порядке определяется максимум полной дисперсии. Процедуру можно прекратить в любом месте и, например, выбрать только первые две главные компоненты, которые воспроизводят 80 процентов полной дисперсии. Проекции объектов на главные компонентные оси можно определить точно. Существенное отличие от эксплораторного факторного анализа заключается в том, что диагональные элементы корреляционной матрицы равны единице, а это означает равенство единице значений общностей, и, как следствие, равенство нулю значений характерностей. Таким образом, в модели компонентного анализа характерные факторы отсутствуют, в чем и состоит главное отличие метода главных компонент и эксплораторного факторного анализа.

Единичная дисперсия всех переменных в компонентном анализе рассматривается в совокупности как общая дисперсия. Хотя это приводит к однозначному решению, однако это нереально во всех практических ситуациях. Пожалуй, никогда не проводится анализ переменных, при котором имела бы место гипотеза, что полная дисперсия равна общей, т.е. что она определяется только наблюдаемыми переменными. Лишь в том случае, когда эта гипотеза подтверждается критерием, можно применять модель компонентного анализа.

Главные компоненты можно рассматривать как возможные математические описания измерений полной дисперсии и им нельзя приписать никакого другого эквивалента в действительности. В соответствии с этим, метод главных компонент позволяет определить ортогональную систему координат с определенными сильными свойствами, которой в действительности чрезвычайно редко могут соответствовать какие-либо условия. Это находится в прямом соответствии с тем, что главные компоненты ортогональны и не подлежат дальнейшему вращению.

В противоположность этому модель эксплораторного факторного анализа позволяет отыскать весь комплекс взаимосвязанных дисперсий. При этом вычислительная операция по определению главных компонент может быть использована в качестве одного из этапов, а именно этапа поиска предварительных факторных решений.

В эконометрии, как в других научных дисциплинах, исследователь постоянно сталкивается с проблемой необходимости разработки приемлемой модели, с помощью которой можно объяснять, а, следовательно, и предсказывать различные экономические явления. Одним из главных источников данных, используемых для этих целей, служат временные ряды наблюдений, которые становятся все более доступными для исследователя, и могут относиться к различным временным интервалам (году, кварталу, месяцу …). Главной особенностью временных рядов наблюдений, отличающей их от других источников данных, является свойство зависимости наблюдений, относящихся к различным периодам времени. Для того чтобы временные ряды использовались эффективно, необходимо применять динамические спецификации как для детерминированной, так и для стохастической составляющей частей модели. В эконометрическом анализе временных рядов принято считать, что модель полностью определена, если задана совместная функция плотности распределения вероятности для ее зависимых переменных.

Одна из целей теории временных рядов состоит в получении стохастических моделей для дискретных рядов во временной области, обладающих максимальной простотой и минимальным числом параметров и при этом адекватно описывающих наблюдения.

Получение таких моделей важно по следующим причинам:

- они могут помочь понять природу системы, генерирующей временные ряды;
- их можно использовать для оптимального прогнозирования будущих значений рядов;
- если исследуются два или несколько связанных временных рядов, можно расширить модели так, чтобы они описывали динамические взаимосвязи между рядами и, следовательно, позволяли оценивать передаточные функции;
- они могут быть использованы для выработки стратегии оптимального управления; эта стратегия указывает, каким образом надо изменять регулируемую переменную для того, чтобы минимизировать возмущение некоторой зависимой переменной.

Возможность оптимального прогнозирования, понимание динамических взаимосвязей и оптимальное управление имеют большое практическое значение. Например, прогнозирование оптимального сбыта нужно для планирования коммерческих операций; модели передаточных функций необходимы для улучшения проектирования и контроля заводских процессов; стратегии оптимального управления нужны для регулирования важнейших показателей производственных процессов вручную или при помощи участвующей в процессе вычислительной техники.

В отличие от временных рядов в традиционном стохастическом факторном анализе предполагается независимость наблюдений, относящихся к различным временным периодам, и в этом состоит принципиальное отличие факторного анализа от анализа временных рядов. С другой стороны, в анализе временных рядов до тонкостей рассматриваются варианты корреляции наблюдений, но в стороне остается вопрос о взаимной связи различных переменных многокомпонентного временного ряда. Проблемы построения передаточных функций в практическом аспекте решены лишь для небольшого числа связанных временных рядов, причем связи предполагаются парными.

Модель эксплораторного факторного анализа с некоторым уточнением может быть распространена на временные ряды, т.е. для наблюдений, не являющихся независимыми, причем, ее использование поможет учесть коррелированность компонент временного ряда, и определить причинную структуру взаимосвязей между компонентами.

Стохастический факторный анализ был разработан при проведении психометрических исследований. Он возник из потребности узнать что-либо о структуре взаимосвязей большого числа измеримых переменных. В настоящее время он все еще тесно связан с психологией, однако применение его в эконометрических исследованиях может быть еще более успешным, если хотя бы воспроизвести подходы, применяемые в психометрии, к экономическим объектам, не говоря уже о том, чтобы пойти дальше, развивая методологию эксплораторного факторного анализа с учетом специфики экономических процессов.

Целью метода является выделение из большого числа наблюдаемых переменных наиболее простых показателей (факторов), которые бы описывали данный объект изучения, как можно точнее воспроизводили бы данные, полученные в результате наблюдения, и, в определенном смысле, также объясняли внутренние объективно существующие закономерности. В принципе эксплораторный факторный анализ не ограничивается лишь воспроизведением исходных данных, что, правда, также является его целью. Эксплораторный факторный анализ решает большой объем задач. С его помощью пытаются разработать такую рациональную модель, которая не только воспроизводит исходные данные, но и позволяет получить интерпретируемую систему показателей модели, имеющих разные эквиваленты в действительности. Затем на втором этапе можно эту модель проверить на новых данных.

Первоочередной задачей эксплораторного факторного анализа является не проверка гипотезы, а ее формирование. Этим эксплораторный факторный анализ отличается от других распространенных статистических методов. Он принадлежит к одному из тех методов, которые делают возможным формирование гипотезы на основе большого комплекса данных. При отыскании факторов, наиболее существенно влияющих на исследуемый признак и стоящих за наблюдаемыми переменными, необходимо руководствоваться принципом как можно более простого объяснения природы данного явления. Это стремление к упрощению находит свое выражение в том, что выделяется такое количество факторов, которое минимально необходимо для воспроизводимости наблюдаемых данных. Это дает возможность более надежно интерпретировать каждый фактор. Для проверки соответствия факторной модели некоторому объекту используется конфирматорный факторный анализ.

После построения факторной модели она используется для анализа развития экономических систем. Именно эта модель позволяет рассмотреть как структуру взаимосвязей между самими наблюдаемыми переменными, так и между отдельными элементами экономической системы.

1.4. Концептуальный подход к использованию эксплораторного факторного анализа для исследования динамики экономических систем

Исследование любой системы, в том числе и экономической, ставящей целью раскрыть механизм ее функционирования, и на основании которого сделать предсказуемым ее движение, не может не основываться на принципе системности. Этот принцип предполагает принятие во внимание как можно большего количества связей и внутренних по отношению к системе, и внешних. Весь этот комплекс, если система сверхсложная, может быть описан посредством очень большого числа показателей, по своей природе представляющих собой случайные переменные. Изменение в течение некоторого промежутка времени значения каждой такой переменной, безотносительно к тому, что стало этому причиной, указывает на изменения состояния системы. Такое изменение свидетельствует о динамике системы, которая будет определять траекторию ее движения в некотором фазовом пространстве.

Фазовое пространство представляет собой пространство всех состояний системы, сущность которого состоит в том, что состояние любой, даже очень сложной системы изображается единственной точкой, а эволюция этой системы – перемещением этой точки. К примеру, двумерное фазовое пространство развития динамической системы имеет вид раскручивающейся спирали.

Фазовое пространство механической системы является четномерным и определяется обычными пространственными координатами частиц системы и их импульсами или обобщенными импульсами, т.е. в самом простом случае – произведением массы частицы на ее ускорение вдоль соответствующей пространственной оси. Так фазовое пространство одной материальной точки имеет шесть измерений, три из которых – координаты с ортонормированным базисом, а три других – компоненты импульса. Механическая система из двух материальных точек будет иметь 12 измерений, и т.д. Динамика таких систем описывается уравнениями Гамильтона [59].

Формальное перенесение механистического понятия фазового пространства на экономические системы не будет корректным, поскольку эти системы в силу особенностей их функционирования в большей степени соответствуют динамическим системам, для которых фазовое пространство не обязательно четномерно и динамика на нем задается не обязательно уравнениями Гамильтона [60].

Чтобы определиться с понятием фазового пространства экономической системы, необходимо уточнить на основании чего одно состояние системы можно отличить от другого, в какой точке этого пространства находится система и какими координатами положение этой точки можно определить. Очевидно, что значения переменных могут указывать на то, в каком состоянии находится система. Даже небольшое изменение одной из переменных будет говорить о том, что система поменяла состояние. Поскольку каждая переменная может принимать значения на достаточно большой длины интервале, то число состояний может стремиться к бесконечности. Воспользовавшись агрегированием состояний, такую динамику системы можно описать с помощью цепей Маркова, и движение системы можно изобразить в виде ее перемещения по графу состояний с определенными вероятностями переходов [60]. Однако моделирование поведения системы таким способом не является наглядным с точки зрения того, что целью исследования является прогноз значений показателей, описывающих систему, а знание вероятности перехода в текущий момент времени из данного состояния в те, с которыми оно связано стрелками графа, в качестве конечной цели анализа эволюции системы, отличается от поставленной задачи. К тому же, поскольку система может оказаться в любом из возможных состояний, прогноз положения системы посредством аппарата цепей Маркова через несколько интервалов времени делается малопригодным с практической точки зрения.

В качестве фазового пространства, в котором исследуется движение экономической системы, можно взять множество значений переменных, а сами переменные – в качестве осей координат. Положение системы в таком пространстве будет задаваться единственной точкой, и также в координатах этого пространства можно рассматривать движение системы [61]. Если число переменных больше трех, то графическое изображение движения системы в пространстве более чем трехмерном, невозможно. Однако наглядность динамики системы не является самоцелью.

Другой аспект делает невозможным рассмотрение динамики экономического объекта в системе координат, отождествляемых с наблюдаемыми переменными. Эти переменные, как компоненты многомерного временного ряда, являются взаимно коррелированными. И если их представить в векторном виде, то взаимозависимость указывает на то, что среди них можно выделить целые группы компланарных векторов, т.е. лежащих в разных гиперплоскостях. Поэтому исходные переменные (векторы) не могут рассматриваться в качестве базиса системы координат фазового пространства.

Эти ситуацию можно исправить, выбрав среди векторов, лежащих в одной гиперплоскости, необходимый базис (для двумерной плоскости – это два неколлинеарных вектора), остальные векторы можно выразить посредством базисных векторов. Совокупность всех базисных некомпланарных векторов определит новую системы координат. Таким образом, необходимая и достаточная для описания динамики объекта размерность фазового пространства меньше, чем число исходных переменных. Проблема состоит в том, какие переменные (существенные переменные) необходимо выбрать в качестве базисных векторов, при этом вектора должны быть некомпланарными.

Другая проблема связана с косоугольностью такой системы координат, поскольку исходные наблюдаемые переменные лишь в исключительных случаях могут быть ортогональными, т.е. их попарные скалярные произведения, при векторном представлении переменных, должны быть равны нулю, что будет говорить об их полной независимости. Собственно, косоугольная система координат, кроме неудобства расчета координат точки, связанного с необходимостью учета угла между осями, практически мало отличается от ортогональной. Более того, ортогональная система координат – частный случай косоугольной системы, у которой угол между осями составляет девяносто градусов. Трудность состоит в том, что углы между осями в рассматриваемом случае не будут равными, что предполагается для косоугольной системы координат. Поэтому необходима дополнительная процедура преобразования системы координат с разными углами между осями в систему с одинаковыми и, лучше – прямыми углами.

Таким образом, возникает необходимость преобразования исходных данных, представленных в виде многомерного временного ряда, к такому их представлению, что можно было бы выделить независимые ортогональные векторы, которые будут играть роль базисных векторов фазового пространства для исследования движения экономической системы. При этом преобразование временного ряда не должно привести к потере информации, а лишь к сжатию ее избыточности, устранению дублирования и повторов. Укажем на основные предпосылки к возможности такого преобразования исходных данных.

Если многомерный временной ряд конечной длины записать в виде матрицы, то по отношению к ней можно поставить вопрос о ее ранге. Такая матрица будет неполного ранга в том случае, когда у нее окажутся взаимозависимые столбцы (столбцы – для варианта представления исходных данных следующим образом: столбец соответствует одной переменной, а строки – значениям переменных для разных моментов времени, т.е. число строк должно быть больше числа столбцов, поскольку наблюдений должно быть достаточное число для проведения процедур оценивания). Основное предположение по отношению к рассматриваемому виду временного ряда состоит в том, что переменные связаны некоторыми зависимостями друг с другом, причем эффект мультиколлинеарности не исключается. Поэтому в матрице должны оказаться зависимые столбцы, что будет указывать на то, что ранг матрицы меньше числа столбцов. Тогда исключая зависимые столбцы, а фактически, отбрасывая зависимые переменные, можно добиться того, что усеченная матрица станет матрицей полного ранга.

Далее остается подобрать ортогонализирующее преобразование. Сделать это не сложно, если ковариационную матрицу представить в виде произведения собственных векторов и собственных значений. Умножение усеченной матрицы, составленной из независимых векторов-переменных, на матрицу собственных векторов и будет тем преобразованием, которое даст в качестве результата базисные векторы фазового пространства, в котором уже корректно можно рассматривать движение экономической системы.

Математически описанная процедура выглядит следующим образом. Пусть $\boldsymbol{X}$ – матрица размером $N \cdot n$ исходных данных, представляющих многомерный временной ряд, $\boldsymbol{X}^*$ − усеченная матрица размером $N \cdot m$, в которой оставлены лишь независимые переменные; где N − число наблюдений (длина временного ряда), n − число исходных переменных, m – число отобранных независимых переменных. При этом $\boldsymbol{X}$ – матрица неполного ранга:

$$Rank(\boldsymbol{X}) = m < n;$$

$\boldsymbol{X}_*$– матрица полного ранга:

$$Rank(\boldsymbol{X}_*) = m,$$

т.е. ранг равен числу столбцов матрицы.

Тогда, с точностью до скалярного множителя (зависящего от числа наблюдений), можно указать, что ковариационная матрица $\boldsymbol{C}$ имеет вид:

$$\boldsymbol{C} = \boldsymbol{X}_*^T \boldsymbol{X}_*. \tag{1.1}$$

Поскольку $\boldsymbol{C}$ − симметричная матрица, то она может быть представлена в следующем виде:

$$\boldsymbol{C} = \boldsymbol{Q}\boldsymbol{L}\boldsymbol{Q}^T, \tag{1.2}$$

где $\boldsymbol{Q}$ − матрица собственных векторов ковариационной матрицы,

$\boldsymbol{L}$ − диагональная матрица собственных значений (чисел) ковариационной матрицы,

$(\cdot)^T$ − символ операции транспонирования.

Поскольку матрица собственных векторов обладает свойством ортогональности:

$$\boldsymbol{Q}^T \boldsymbol{Q} = \boldsymbol{I},$$

где $\boldsymbol{I}$ – единичная матрица, то умножение (1.4.2) на $\boldsymbol{Q}$ слева, и на $\boldsymbol{Q}^T$ справа приводит к диагонализации ковариационной матрицы:

$$\boldsymbol{Q}^T \boldsymbol{C} \boldsymbol{Q} = \boldsymbol{L}. \tag{1.3}$$

Подставляя в (1.3) вместо $\boldsymbol{C}$ вычисляемые значения ковариаций переменных (1.1), получаем следующее выражение:

$$\boldsymbol{Q}^T \boldsymbol{X}_*^T \boldsymbol{X}_* \boldsymbol{Q} = \boldsymbol{L}.$$

Следовательно, $\boldsymbol{Q}$ – матрица собственных векторов ковариационной матрицы является линейным преобразованием исходных переменных, приводящим их к ортогонализации. Таким образом, столбцы матрицы $\boldsymbol{W}$, получаемые в результате преобразования,

$$\boldsymbol{W} = \boldsymbol{X}_* \boldsymbol{Q}, \tag{1.4}$$

являются базисными ортогональными векторами фазового пространства:

$$\boldsymbol{W}^T \boldsymbol{W} = \boldsymbol{L}.$$

Для каждого момента времени в такой системе координат можно указать положение экономической системы, и соединив их -мерной кривой становится возможным проследить ее движение. Экстраполяция этой кривой на моменты времени в будущем позволит сделать прогноз движения системы. Однако, графическое изображение движения системы и наглядная интерпретация прогноза возможны лишь в случае, когда число независимых переменных и, соответственно, базисных вектора, два или три. Для пространств большей размерности динамика системы может быть задана лишь аналитически.

Таким образом, показано, что на основе исходных переменных можно построить фазовое пространство с ортогональной системой координат, в котором становится возможным исследовать движение экономической системы. Проблемой остается отбор из множества наблюдаемых исходных переменных независимых. Следуя формальным процедурам особой необходимости в таком отборе нет. Поскольку разложение ковариационной матрицы всех исходных переменных с неполным рангом на собственные векторы и собственные значения возможен, при этом часть собственных значений будет равна нулю, причем в том количестве, которое будет соответствовать числу зависимых переменных. Пусть это число равно k ($k = n - m$), тогда k последних столбцов в исходной матрице будут автоматически отброшены. Это решение проблемы, но вряд ли оно имеет экономический и здравый смысл. При таком формальном подходе среди отброшенных переменных могут оказаться важные для экономической интерпретации компоненты. Поэтому необходимо в рамках данного подхода сконструировать другую процедуру формирования фазового пространства динамики экономической системы.

Если внимательно рассмотреть преобразование (1.4), то можно заметить, что матрица $\boldsymbol{W}$ представляет собой главные компоненты, получаемые в соответствии с методом с тем же названием (метод главных компонент или компонентный анализ). В основе метода лежит разложение ковариационной или корреляционной матрицы наблюдаемых переменных на собственные векторы и собственные значения. Компоненты вычисляются по выражению, аналогичному (1.4). Первая компонента интерпретируется как первая ось новой системы координат, вдоль которой облако точек – наблюдаемых значений вытянуто в наибольшей степени, и объясняет наибольшую совокупную дисперсию исходных данных. Вторая компонента направлена ортогонально первой оси вдоль второго по величине рассеяния исходных точек, и объясняет вторую по величине дисперсию, и так далее.

Число компонент в общем случае равно числу переменных, но если выборочная ковариационная (корреляционная) матрица имеет неполный ранг, то часть собственных значений равно нулю, и соответствующие компоненты отсутствуют. В компонентном анализе происходит переход от наблюдаемых переменных к компонентам, которые интерпретируются по вкладу в общую дисперсию всего множества наблюдаемых величин, причем компоненты, в отличие от исходных переменных являются ортогональными, т.е. независимыми новыми переменными.

Компонентный анализ чаще всего используется как метод преобразования исходных данных к ортогональному виду, причем в некоторых случаях рассматриваются не все компоненты, число которых может быть равно числу наблюдаемых переменных, когда ковариационная матрица полного ранга, а анализируется их «главная» часть – главные компоненты. Малосущественные компоненты, имеющие незначительный вклад в совокупную дисперсию, отбрасываются. Это позволяет понизить число анализируемых переменных без существенной потери информации. Однако в силу недостаточной объективности определения числа главных компонент, в этом анализе в основном рассматриваются все компоненты, каждая из которых представляет собой образ некоторой исходной переменной. Проблема состоит лишь в том, какую переменную отождествить с данной компонентой.

Решение обозначенных проблем приводит к эксплораторному факторному анализу, родственному методу главных компонент. Модель факторного анализа является более детализированной и включает так называемые общие факторы – аналоги главных компонент, а также специфические факторы. Число общих факторов всегда намного меньше числа исходных переменных, но специфических (или характерных) факторов – ровно столько, сколько наблюдаемых переменных. Такая модель с помощью общих факторов позволяет не только выявить основные тенденции в динамике экономической системы, но и посредством характерных факторов учесть специфику каждой переменной. Еще одно существенное отличие этих двух методов: в компонентном анализе компоненты формируются на основании переменных, а в факторном анализе факторы определяют поведение переменных, т.е. имеет место разная причинная направленность во взаимодействии переменных с компонентами и факторами. Именно такая логическая увязка делает пространство общих факторов более подходящим в качестве фазового пространства.

Факторы выступают в качестве экзогенных переменных, задающих поведение экономической системы, они же определяют значения наблюдаемых переменных. Таким образом, независимые латентные переменные – факторы, как некоторые скрытые от непосредственного наблюдения причины, предопределяют динамику экономической системы, и это с одной стороны. С другой стороны, движение системы можно фиксировать в пространстве общих факторов, ортогональными осями которого являются общие факторы. При этом каждому состоянию системы, фактически описываемому совокупностью n значений переменных $\{x_{t1}, x_{t2}, \dots, x_{tn}\}$ в данный момент времени t, будет соответствовать единственная точка -мерного пространства общих факторов с координатами $\{f_{t1}, f_{t2}, \dots, f_{tm}\}$, где f_{tj} − значение j-го фактора в момент времени t, m − число факторов.

Пространство наблюдаемых переменных и пространство общих факторов должно быть связано некоторой операцией отображения каждой точки факторного пространства в пространство исходных переменных. Эксплораторный факторный анализ задает такое отображение посредством модели, связывающей значение наблюдаемой переменной со значениями факторов с точностью до некоторой погрешности, зависящей от характерных факторов. В качестве линейного оператора отображения служит матрица факторных нагрузок.

Поскольку факторы задают поведение системы, совокупность их значений будет определять состояние системы. Поэтому исследование закономерностей изменения факторов будет ключевым при анализе поведения экономической системы. Следует отметить, что модель, на основании которой будет производиться прогноз, для независимых компонент имеет на два порядка меньше параметров, чем у VAR-модели. К тому же векторный прогноз ортогональных составляющих является наиболее корректным [62, с. 38].

В модели, связывающей наблюдаемые переменные и общие факторы, последние являются экзогенными переменными. На основании именно этой модели производится прогноз эндогенных переменных, т.е. переменных, объясняемых модельным уравнением. Этот прогноз, однако, делается на основании будущих – прогнозных значений внешних переменных. Поэтому необходимо уравнение, отображающее динамику факторов, на основании которого можно будет произвести прогноз их значений.

Общие факторы экономической системы являются надпричинами ее закономерностей. Сами факторы – независимы, и не оказывают непосредственного воздействия друг на друга. Особенности и характерные черты каждого фактора определяются им самим. Поскольку если бы динамику факторов определяло что-то стоящее над ними, скажем некоторый надфактор, то его можно было бы выявить так же, как факторный анализ позволяет выявлять латентные переменные, и рассматривать этот надфактор в качестве обычного фактора стоящего над наблюдаемыми показателями системы; при этом размерность факторного пространства стала бы меньше.

Независимость фактора и внутренний механизм формирования закономерностей его динамики делает логичным предположение об авторегрессионных зависимостях значений фактора. В самом деле, если что-то и определило величину фактора в данный момент времени, так это – его же значение в предыдущие моменты времени, поскольку ничто другое повлиять на это не могло. Если восстановить значения факторов в прошедшие моменты времени, а сделать это возможно, поскольку косвенная информация об этом содержится в многомерных временных рядах наблюдаемых переменных, то воссоздав траекторию движения факторов в фазовом пространстве, становится возможным судить о будущем направлении их движения, а значит – делать прогноз для факторов. А на основании значений факторов в прогнозный период строить прогноз для наблюдаемых переменных.

Таким образом, динамика экономической системы определяется двумя видами зависимостей. Первая задается закономерностями движения факторов, определяющих поведение экономической системы, а вторая – отображением факторов на множество значений наблюдаемых переменных. Эти зависимости должны стать основой для математической модели экономической системы.

Выводы

Проведенный анализ современных научных исследований в области методологии предвидения и прогнозирования показал, что исследование и прогнозирование экономических процессов приобретают на нынешнем этапе развития экономики и социальных отношений особую значимость и важность. Этот этап характеризуется стремительным развитием научно-технического прогресса, появлением принципиально новых технологий и продуктов, которые коренным образом меняют не только технологии производства, но и образ жизни людей. Яркой чертой настоящего этапа является глубокая глобализация экономики, которая несет не только преимущества международного разделения труда, но и существенные экономические риски. Кризис, появляющийся в одной стране, распространяется далеко за ее пределы, и становится большой проблемой для стран с неокрепшей экономикой. Все это существенным образом сказывается не только на количественные характеристики динамики экономических систем, но и на их качественный характер.

Современная концепция методологии прогнозирования предполагает рассмотрение комплекса методов, приемов, подходов, направленных на повышение эффективности и конкурентоспособности хозяйствующих субъектов рынка. Однако существующие методы прогнозирования не в достаточной степени соответствуют методической постановке процедуры прогнозирования динамических отклонений и случайных факторов при функционировании экономической системы, которая должна исходить из основных законов в методологии прогнозирования: системности, онтогенеза, композиции и экономической целесообразности.

Анализ наиболее широко используемых в настоящее время методов прогнозирования показал, что создание модели, учитывающей авторегрессионные зависимости и корреляционные связи между переменными, а также разработка методов получения оценок параметров такой модели, отображающей процессы неустоявшейся динамики, являются актуальными.

Для того, чтобы в наиболее полной мере исследовать экономическую систему целесообразно использовать эксплораторный факторный анализ Именно факторная модель позволяет рассмотреть как структуру взаимосвязей между самими наблюдаемыми переменными, так и между отдельными элементами экономической системы.

Динамика экономической системы определяется двумя видами зависимостей. Первая задается закономерностями движения факторов, определяющих поведение экономической системы, а вторая – отображением факторов на множество значений наблюдаемых переменных. Эти зависимости должны стать основой для математической модели экономической системы.

2. МЕТОДЫ ОЦЕНИВАНИЯ ПАРАМЕТРОВ ЭКСПЛОРАТОРНОЙ ФАКТОРНОЙ МОДЕЛИ ЭКОНОМИЧЕСКОГО ПРОЦЕССА.

2.1. Модель эксплораторного факторного анализа экономического процесса

Чтобы численно проиллюстрировать вычислительную сторону факторного анализа, приведем небольшой пример. Пусть в результате наблюдений за пятью показателями получены данные, на основании которых вычислена корреляционная матрица $\boldsymbol{R}$

$$\boldsymbol{R} = \begin{pmatrix} 1 & -0{,}56 & 0{,}48 & -0{,}16 & 0{,}08 \\ -0{,}56 & 1 & -0{,}42 & 0{,}14 & -0{,}07 \\ 0{,}48 & -0{,}42 & 1 & -0{,}12 & 0{,}06 \\ -0{,}16 & 0{,}14 & -0{,}12 & 1 & -0{,}02 \\ 0{,}08 & -0{,}07 & 0{,}06 & -0{,}02 & 1 \end{pmatrix}.$$

Корреляционная матрица является симметричной, т.е. элементы, лежащие выше диагонали, представляют собой зеркальное отражение элементов, лежащих ниже диагонали матрицы. На главной диагонали всегда располагаются единицы, поскольку коэффициент корреляции показателя самого с собой равен единице.

Целью факторного анализа является извлечение на поверхность величины, так называемого фактора, который, по возможности, точнее позволил бы воспроизвести наблюдаемые корреляции с использованием соответствующей процедуры вычислений. Этот фактор и связанная с ним процедура вычислений в начале исследования являются гипотетическими.

Известно, что симметричные матрицы получаются в результате произведения некоторой матрицы на себя же, но транспонированную. Поэтому, логично было бы предположить, что корреляционная матрица может быть получена не только в результате непосредственного вычисления по исходным данным, а именно:

$$\boldsymbol{R} = \frac{1}{N-1}\boldsymbol{Z}^T\boldsymbol{Z},$$

где $\boldsymbol{Z}$ – матрица N наблюдений за показателями, центрированными и нормированными (т.е. из каждого значения вычтено среднее и эта разность разделена на среднеквадратическое отклонение);

$[\,]^T$ –символ транспонирования.

Тот же результат (но как будет видно ниже – почти тот же) может быть получен с помощью перемножения других матриц, в нашем случае:

$$\boldsymbol{R} = \boldsymbol{a}\,\boldsymbol{a}^T,$$

где матрица-столбец $\boldsymbol{a}$ имеет вид:

$$\boldsymbol{a} = \begin{pmatrix} 0{,}8 \\ -0{,}7 \\ 0{,}6 \\ -0{,}2 \\ 0{,}1 \end{pmatrix}$$

В самом деле,

$$\boldsymbol{a}\boldsymbol{a}^T = \begin{pmatrix} 0{,}8 \\ -0{,}7 \\ 0{,}6 \\ -0{,}2 \\ 0{,}1 \end{pmatrix} \begin{pmatrix} 0{,}8 & -0{,}7 & 0{,}6 & -0{,}2 & 0{,}1 \end{pmatrix} =$$

$$= \begin{pmatrix} 0{,}64 & -0{,}56 & 0{,}48 & -0{,}16 & 0{,}08 \\ -0{,}56 & 0{,}49 & -0{,}42 & 0{,}14 & -0{,}7 \\ 0{,}48 & -0{,}42 & 0{,}36 & -0{,}12 & 0{,}06 \\ -0{,}16 & 0{,}14 & -0{,}12 & 0{,}04 & -0{,}02 \\ 0{,}08 & -0{,}07 & 0{,}06 & -0{,}02 & 0{,}01 \end{pmatrix}$$

В последней матрице видно точное воспроизведение всех элементов корреляционной матрицы, за исключением диагональных элементов. Эти значения называются общностями.

Элементы вектора $\boldsymbol{a}$ позволяют получить коэффициенты корреляции между разными переменными, причем их число значительно меньше числа наблюдений. Формально говоря, фактически вся информация о взаимосвязях показателей заключена в пяти числах – элементах вектора $\boldsymbol{a}$. Все это дает основание говорить, что за наблюдаемыми корреляциями стоит некоторая невидимая величина, называемая фактором, которая причинно обуславливает эти корреляции.

В факторном анализе чаще всего применяется линейная модель. Это связано, во-первых, с целью объяснить наблюдаемое явление в терминах более простой теории и, во-вторых, с тем, что отход от линейной модели приводит к математическим трудностям, разрешить которые для ряда практических задач бывает достаточно нелегко.

Обычно в статистическом исследовании рассматривается группа объектов, характеризуемых некоторыми общими для них свойствами. В каждом конкретном случае термину «объект» могут соответствовать самые разные элементы: компании, банки, финансовые бумаги и т.д. Измерения этих свойств (признаков) объектов будем называть значениями их *переменных.* Всюду далее для обозначения общего числа объектов используется буква N для обозначения числа параметров – буква n. Отдельную переменную будем обозначать через X_j. Индекс i используется для обозначения любого объекта или наблюдения; $i = 1,2,3, ..., N$. Значение параметра X_j для объекта i обозначим X_{ij}, обращая внимание на порядок индексов. Частное значение X_{ij} называется *наблюденным (выборочным) значением* и соответствует измерению при произвольном начале координат и произвольных единицах измерения.

В литературе по статистике принято для обозначения статистических характеристик выборки (и гипотетических, или скрытых, параметров) использовать буквы греческого алфавита, а для обозначения наблюденных значений параметров – буквы латинского алфавита.

Среднее значение для N наблюдений параметра X_j определяется следующим образом

$$\overline{X}_j = \frac{1}{N}\sum_{i=1}^{N} X_{ij}\,. \tag{2.1}$$

Математическое ожидание этой переменной будем обозначать греческой буквой μ_j. Наблюденное значение переменной может быть преобразовано к более удобной форме, если зафиксировать начало координат и единицы измерения. Пусть начало координат соответствует среднему выборочному значению, тогда величина

$$x_{ij} = X_{ij} - \overline{X}_j\,. \tag{2.2}$$

называется отклонением переменной от его среднего значения.

Выборочная дисперсия, а именно ее несмещенная оценка, переменной X_j определяется в соответствии с выражением

$$s_j^2 = \frac{1}{N-1}\sum_{i=1}^{N} x_{ij}^2 \,. \tag{2.3}$$

Истинное, но неизвестное значение этой дисперсии обозначается σ_j^2. Теперь приняв в качестве единицы измерения среднеквадратическое отклонение s_j, определим нормированное значение переменной X_j для объекта или наблюдения i:

$$z_{ij} = \frac{x_{ij}}{s_j}. \tag{2.4}$$

Множество значений $z_{ij}, (i = 1,2,...,N)$ называется стандартной формой задания переменной Z_j.Очевидно, что математическое ожидание нормированной переменной равно нулю, а ее дисперсия равна единице.

Для двух произвольных переменных X_j и X_k можно определить их коэффициент *ковариации* в соответствии с выражением

$$s_{jk} = \frac{1}{N-1}\sum_{i=1}^{N} x_{ij}\, x_{ik} \,, \tag{2.5}$$

а соответствующая характеристика истинного распределения обозначается через σ_{jk}. Коэффициент корреляции обозначается через ρ_{jk}, его выборочное значение подсчитывается из выражения

$$r_{jk} = \frac{s_{jk}}{s_j\, s_k}, \tag{2.6}$$

т.е. он равен отношению коэффициента ковариации к произведению среднеквадратических отклонений переменных.

Обычно первым этапом факторного анализа является вычисление коэффициентов корреляции между всеми изучаемыми переменными.

Задача факторного анализа состоит в тоw, чтобы выразить, параметр Z_j в терминах скрытых гипотетических *факторов.* Простейшей моделью для описания одной переменной в терминах нескольких других может служить линейная модель. Однако и в рамках линейной модели в зависимости от целей анализа возможны различные варианты. Рассмотрим две такие цели: 1) выделение максимальной дисперсии; 2) «наилучшую» аппроксимацию выборочных корреляций.

К. Пирсон [27] предложил эвристический метод сжатия большого массива информации с одновременным выделением максимальной дисперсии. Позднее Г. Хотеллинг [28] развил этот метод, создав метел *главных компонент,* или *компонентный анализ.* Модель компонентного анализа имеет вид:

$$Z_j = a_{j1}F_1 + a_{j2}F_2 + ... + a_{jn}F_n,\ (j = 1,2,...,n), \tag{2.7}$$

где каждая из наблюденных переменных линейно зависит от n некоррелированных между собой новых компонент $F_1, F_2,..., F_n$. Важное свойство метода состоит в том, что каждая очередная компонента дает максимально возможный вклад в суммарную дисперсию параметров. Обычно на практике оставляют небольшое число компонент, особенно в том случае, если на их долю приходится достаточно большой процент суммарной дисперсии переменных. Однако для точной аппроксимации корреляций между переменными необходимы все компоненты.

В противоположность подходу, связанному с максимизацией дисперсии, в модели *классического факторного анализа* требуется наилучшим образом аппроксимировать корреляции. Основная модель факторного анализа может быть записана в таком виде:

$$Z_j = a_{j1}F_1 + a_{j2}F_2 + \ldots + a_{jm}F_m + d_jU_j,\ (j = 1,2,\ldots,n). \qquad (2.8)$$

Здесь параметр Z_j линейно зависит от *m общих* факторов и характерного фактора (обычно *m* много меньше *n*). Общие факторы учитывают корреляции между переменными, характерный фактор учитывает оставшуюся (в том числе и связанную с различными погрешностями) дисперсию. Коэффициенты при факторах часто называют *нагрузками.*

Если быть точными, то следовало бы факторы и коэффициенты при них в выражении (2.8) обозначать греческими буквами, так как здесь речь идет о гипотетических характеристиках генеральной совокупности. Однако обозначения (2.8) настолько установились в литературе по факторному анализу, что не стоит менять их. Следует помнить только, что одни и те же символы в (2.7) и (2.8) обозначают разные вещи.

Используя (2.8), можно записать выражение для значения *j*-го параметра у *i*-го объекта:

$$z_{ij} = \sum_{p=1}^{m} a_{jp}F_{ip} + d_jU_{ij},\ (i = 1,2,\ldots,N;\ j = 1,2,\ldots,n). \qquad (2.9)$$

Здесь F_{ip} – значение общего фактора F_p, для объекта или наблюдения под номером *i*, а каждое *m* членов суммы $a_{jp}F_{ip}$ отражает степень влияния соответствующего фактора на параметр Z_j; d_jU_{ij} – остаток, невязка, в выражении для наблюдаемого значения.

Без потери общности можно предположить, что факторы имеют нулевые средние значения и единичные дисперсии (хотя фактически факторы не известны). Кроме того, считают, что характерные факторы независимы как между собой, так и по отношению к общим факторам. В модели (2.8) под факторами F_p, понимают случайные величины с некоторым законом распределения, чаще всего предполагаемым нормальным. Из предположения о том, что факторы – нормально распределенные независимые случайные величины, следует, очевидно, предположение о том, что параметры Z_j имеют многомерное нормальное распределение.

Наибольший теоретический интерес вызывает модель (2.8), и в ней – основная проблема факторного анализа: оценка $n \cdot m$ нагрузок общих факторов. Для решения этой проблемы были предложены различные методы, рассмотренные в следующем разделе. В этих методах для вычисления факторных нагрузок используются матрицы коэффициентов корреляций между переменными.

Поскольку в факторном анализе переменная описывается линейной комбинацией некоторого небольшого числа других переменных (плюс остаток), то может показаться, что факторный анализ очень похож на регрессионный анализ. Однако в регрессионном анализе набор независимых переменных предполагается измеряемым в действительности, тогда как в факторном анализе эти переменные являются гипотетическими и лишь могут быть оценены с помощью измеренной информации при последующем анализе.

В соответствии с моделью (2.8) дисперсия переменной может быть выражена через факторы. Так, подставив в выражение выборочной дисперсии (2.3) значение нормированной величины, определяемой из выражения (2.9), получим:

$$s_j^2=\sum_{l=1}^{N} z_{ij}^2=\sum_{p=1}^{m} a_{jp}^2\left(\sum_{l=1}^{N} F_{ip}^2/(N-1)\right)+d_j^2\sum_{l=1}^{N} U_{ij}^2/(N-1)+$$

$$+2\sum_{p=1}^{m}\sum_{q=1}^{m} a_{jp}a_{jq}\left(\sum_{l=1}^{N} F_{ip}F_{iq}/(N-1)\right)+2d_j\sum_{p=1}^{m} a_{jp}\left(\sum_{l=1}^{N} F_{ip}U_{ij}/(N-1)\right). \tag{2.10}$$

Поскольку дисперсия переменной, заданной в стандартном виде, равна единице и все переменные (включая факторы) предполагаются заданными в стандартном виде, предыдущее выражение можно переписать в таком виде:

$$s_j^2=\sum_{p=1}^{m} a_{jp}^2+d_j^2+2\sum_{p=1}^{m}\sum_{q=1}^{m} a_{jp}a_{jq}r_{F_pF_q}+2d_j\sum_{p=1}^{m} a_{jp}r_{F_pU_j}. \tag{2.11}$$

Характерные факторы всегда не коррелированны с общими факторами; принять, что общие факторы не коррелируют также и между собой, то выражение (2.11) можно упростить:

$$s_j^2=1=a_{j1}^2+a_{j2}^2+\ldots+a_{jm}^2+d_j^2. \tag{2.12}$$

Члены в правой части представляют доли дисперсии параметра Z_j, приходящуюся на соответствующие факторы. Например, a_{j2}^2 есть вклад фактора F_2, в дисперсию параметра Z_j. *Полный вклад* V_p фактора F_p в суммарную дисперсию параметров определяется выражением

$$V_p=\sum_{j=1}^{n} a_{jp}^2,\ (p=1,2,\ldots,m), \tag{2.13}$$

а полный вклад всех общих факторов в суммарную дисперсию переменных равен

$$V=\sum_{p=1}^{m} V_p. \tag{2.14}$$

Отношение $\dfrac{V}{n}$ служит иногда показателем полноты факторизации.

Из выражения для общей дисперсии (2.12) следуют два важных понятия факторного анализа. Первое − *общность* переменной Z_j, которая представляет собой сумму квадратов факторных нагрузок

$$h_j^2=a_{j1}^2+a_{j2}^2+\ldots+a_{jm}^2,\ (j=1,2,\ldots,n), \tag{2.15}$$

и вторoе − *характерность* – вклад характерного фактора в значение дисперсии. Характерность показывает, насколько в общих факторах учтена суммарная дисперсия переменной. Иногда удобно разбивать характерность на две части: ту, что действительно связана со спецификой изучаемых переменных, и ту, что связана с ошибками измерений. Если к данному набору параметров добавить несколько новых, то изучение добавочных коэффициентов корреляции может привести к появлению дополнительных общих факторов. Это означает, что при рассмотрении первоначального набора переменных эти возможные связи данной переменной отражались в той доле характерности, которая связана с действительной спецификой переменной. Эта доля называется *специфичностью* переменной. Оставшаяся доля, связанная с несовершенством измерений, называется *дисперсией* или *ненадежностью* переменной. Дополнение дисперсии ошибки ДО полной иногда называют «надежностью» переменной. В психологии эту характеристику переменной (отличая ее от дисперсии ошибки) измеряют обычно значением коэффициента корреляции между двумя отдельными реализациями одного и того же теста или между близкими вариантами теста, реализованными параллельно. Если две такие реализации переменной обозначить Z_{j*} и Z_j, тогда их коэффициент корреляции r_{j*j} измеряет надежность переменной.

Если характерный фактор на два аналогичных вида, то модель (2.9) приобретает вид

$$Z_j = a_{j1}F_1 + a_{j2}F_2 + ... + a_{jm}F_m + b_jS_j + e_jE_j,\ (j = 1,2,...,n), \qquad (2.16)$$

где S_j и E_j – соответственно факторы специфичности и ошибки, b_j и e_j – коэффициенты при них.

Поскольку факторы специфичности и ошибки между собой не коррелируют, то для характерности и ее составляющих частей справедливо соотношение

$$d_j^2 = b_j^2 + e_j^2. \qquad (2.17)$$

Суммарная дисперсия может быть записана в виде:

$$s_j^2 = h_j^2 + b_j^2 + e_j^2 = 1. \qquad (2.18)$$

Таким образом, суммарная дисперсия состоит из общности (соответствует факторам *F*) и характерности, или иначе, суммарная дисперсия переменной состоит из общности, специфичности (соответствует факторам *S*) и дисперсии ошибки (соответствует факторам *E*).

Методы факторного анализа позволяют получить для каждой переменной значения общности h_j^2 и характерности d_j^2. Далее характерность может быть разбита на свои составляющие: специфичность и дисперсию ошибки. При этом способ разбиения не связан с видом факторного решения. К примеру, если из опыта известно значение надежности переменной Z_j, то значение дисперсии ошибки можно получить из соотношения

$$e_j^2 = 1 - r_{j*j}. \qquad (2.19)$$

Зная дисперсию ошибки, из (1.2.17) можно получить значение специфичности:

$$b_j^2 = d_j^2 - e_j^2, \qquad (2.20)$$

где характерность d_j^2 известна из результатов факторного анализа. Поскольку надежность является дополнением дисперсии ошибки до единицы, то из (1.2.18) следует, что

$$r_{j*j} = h_j^2 + b_j^2, \qquad (2.21)$$

и, следовательно,

$$h_j^2 = r_{j*j} - b_j^2 \le r_{j*j}. \qquad (2.22)$$

Другими словами, общность любой переменной не превышает его надежности и равно ей только в случае нулевой специфичности.

Доля общности от дисперсии, учитываемой общими факторами, измеренная в процентах, обычно называется *полнотой факторизации* C_j переменной Z_j. Этот показатель обычно меньше 100. Равенство ста процентам возможно лишь в случае отсутствия специфичности у переменной, что практически встречается крайне редко.

Рассмотрим некоторые проблемы, связанные с факторизацией реальных данных. Пользуясь моделью (2.8), можно записать выражения для всех *n* переменных в виде

$$Z_1 = a_{11}F_1 + a_{21}F_2 + ... + a_{m1}F_m + d_1U_1,$$
$$Z_2 = a_{12}F_1 + a_{22}F_2 + ... + a_{m2}F_m + d_2U_2,$$
$$...$$
$$Z_n = a_{1n}F_1 + a_{2n}F_2 + ... + a_{mn}F_m + d_nU_n \qquad (2.23)$$

Такая система в факторном анализе называется *факторным отображением*.

В отображении (2.23) общие факторы $F_p\,(p = 1,2,...,m)$ могут быть как коррелированными, так и некоррелированными. В то же время как характерные факторы $U_j\;(j = 1,2,...,m)$ всегда предполагаются не коррелированными как между собой, так и с общими факторами. В уравнении для каждой конкретной переменной некоторые из коэффициентов могут оказаться равными нулю и, следовательно, эта переменная будет выражена через меньшее чем *m* число общих факторов. Число общих факторов, через которые выражается данная переменная, называется его *сложностью*.

Факторный анализ позволяет получить не только отображение, но и значения коэффициентов корреляции между переменными и факторами. Таблица таких коэффициентов корреляции называется *факторной структурой.* Для выполнения полного факторного анализа необходимы и отображение и структура.

Рассмотрим теперь функциональные взаимосвязи между элементами структуры и коэффициентами отображения. Умножив любое из уравнений (2.23) на соответствующие факторы, произведя суммирование по всем наблюдениям *N* и разделив на *N*, получим:

$$r_{Z_jF_1} = a_{j1} + a_{j2}r_{F_1F_2} + ... + a_{jp}r_{F_1F_p} + ... + a_{jm}r_{F_1F_m},$$
$$...$$
$$r_{ZjF_2} = a_{j1}r_{F_pF_1} + a_{j2}r_{F_pF_2} + ... + a_{jp} + ... + a_{jm}r_{F_pF_m}, \qquad (2.24)$$
$$...$$
$$r_{Z_jF_m} = a_{j1}r_{F_mF_1} + a_{j2}r_{F_1F_2} + ... + a_{jp}r_{F_mF_p} + ... + a_{jm}$$

и

$$r_{Z_jU_j} = d_j. \qquad (2.25)$$

Из уравнения (2.25) видно, что коэффициент корреляции переменной с характерным фактором всегда равен коэффициенту при характерном факторе в соответствующем уравнении отображения. Там, где это не будет вызывать недоразумения, будем понимать под факторной структурой таблицу коэффициентов корреляция параметров только с общими факторами, т.е. таблицу значений $r_{Z_jF_p}$.

Иногда уравнениями (2.24) пользуются для оценки элементов структуры, но чаще, когда известны коэффициенты корреляции между переменными и факторами и коэффициенты корреляции между самими факторами. С помощью уравнений (2.24) получают значения коэффициентов отображения. Формально (2.24) можно рассматривать мак *n* систем линейных уравнений *m*-го порядка относительно неизвестных $a_{jp}\;(j = 1,2,...,n;\; p = 1,2,...,m)$ с известными левыми частями. Эти системы можно разрешить относительно неизвестных коэффициентов a_{jp}.

Из (2.24) следует, что элементы $r_{Z_jF_p}$ структуры, вообще говоря, не равны коэффициентам a_{jp} отображения. В случае, когда общие факторы не коррелированны между собой, т.е. $r_{F_pF_q} = 0$, уравнения сворачиваются к виду

$$r_{Z_jF_p} = a_{jp}\;(j = 1,2,...,n;\; p = 1,2,...,m) \qquad (2.26)$$

Таким образом, элементы структуры совпадают с соответствующими коэффициентами отображения только в случае некоррелированных факторов. Это означает, что если анализ ведется в предположении некоррелированности факторов, то полное решение получается непосредственно из факторного отображения, а коэффициенты корреляции между переменными и факторами равны коэффициентам отображения.

Как уже указывалось, для выполнения полного факторного анализа необходимы и структура и отображение. Структура позволяет выявить корреляции между переменными и факторами, которые необходимы для идентификации факторов и их последующей оценки. Отображение в форме уравнений регрессии выражает линейные взаимосвязи переменных и факторов. Оно может также служить для воспроизведения корреляций между переменными при проверке значимости решения. Отображение оказывается также полезным при сравнении разных систем факторов для данного набора параметров.

Рассмотрим вопрос, как совокупность факторов объясняет взаимосвязи между параметрами.

Исходная информация представляет собой выборочные коэффициенты корреляции между параметрами. Как же они соотносятся с коэффициентами корреляции, подсчитанными на основе аппроксимирующей линейной модели? Если коэффициенты корреляции, вычисленные из модели, мало отличаются от выборочных коэффициентов корреляции, то говорят, что модель хорошо описывает экспериментальные данные, в противном случае, модель считается неадекватной. Факторное отображение (2.23) позволяет следующим образом получить коэффициент корреляции между двумя параметрами: перемножим любые два уравнения из (2.23), просуммируем по всем объектам *N* и разделим на *N*. Помня, что факторы заданы в стандартном виде, запишем выражение для коэффициента корреляции между параметрами Z_j и Z_k $(j,k=1,2,...,n)$

$$
\begin{aligned}
r'_{jk} &= a_{j1}a_{k1}+a_{j2}a_{k2}+...+a_{jm}a_{km}+\\
&+(a_{j1}a_{k2}+a_{j2}a_{k1})r_{F_1F_2}+...+(a_{j1}a_{km}+a_{jm}a_{k1})r_{F_1Fm}+\\
&+(a_{j2}a_{k3}+a_{j3}a_{k2})r_{F_2F_3}+...+(a_{j2}a_{km}+a_{jm}a_{k2})r_{F_2F_m}+\\
&+...+a_{j1}d_k r_{F_1U_k}+a_{k1}d_j r_{F_1U_j}+...+a_{jm}d_k r_{F_mU_k}+\\
&+a_{km}d_j r_{F_mUj}+d_j d_k r_{U_jU_k},
\end{aligned}
\tag{2.27}
$$

где вычисленный *коэффициент корреляции* обозначен r'_{jk} в отличие от выборочного коэффициента корреляции r_{jk}. Эти обозначения используются далее повсюду.

Как указывалось выше, характерные факторы предполагаются не коррелированными как между собой, так и с общими факторами, т.е. $r_{F_pU_j}=r_{F_pU_k}=r_{U_jU_k}=0$. Если общие факторы также некоррелированны, то уравнение (2.27) упрощается: $r_{F_pF_q}=0$, и поэтому все члены под первой строкой исчезают. Таким образом, в случае некоррелированных общих факторов коэффициент корреляции между любыми двумя переменными получается из факторного отображения с помощью уравнения

$$r'_{jk}=a_{j1}a_{k1}+a_{j2}a_{k2}+...+a_{jm}a_{km},\ (j\neq k, k=1,2,...n). \tag{2.28}$$

Здесь попросту записана сумма произведений коэффициентов отображения, соответствующих рассматриваемым переменным. Понятно, что коэффициент корреляции параметра с самим собой равен единице. Для вычисленных коэффициентов корреляции это свойство также имеет место благодаря введению характерных факторов. Действительно, положив в (2.27) $k=j$, получим коэффициент корреляции параметра с самим собой и, снова предположив некоррелированность факторов, тотчас увидим, что корреляция параметра с самим собой равна сумме общности и характерности параметра.

Выяснив различие между вычисленными и выборочными коэффициентами корреляции, обратимся теперь к тому, что их связывает. Из-за ошибок, вызванных неточностью эксперимента и представительностью выборки, коэффициенты корреляции, вычисленные из факторного отображения, в более общем виде (2.27) или в случае некоррелированных факторов в виде (2.28), вообще говоря, не равны выборочным коэффициентам корреляции. Это является следствием общенаучного принципа, состоящего в том, что теоретическая модель всегда проще, нежели действительный объект, и потому естественно ожидать несовпадения данных экспериментальных и вычисленных из теории. В частности, в факторном анализе значения коэффициентов корреляции r'_{jk} вычисленных с помощью линейных комбинаций, могут несколько отличаться от выборочных.

После получения факторного отображения необходимо выяснить, несколько адекватно оно описывает корреляции между переменными. Для этого с помощью отображения получают вычисленные корреляции, которые далее вычитаются из выборочных коэффициентов корреляции. Итоговые разности называют *остаточными коэффициентами корреляции*:

$$\bar{r}_{jk} = r_{jk} - r'_{jk}, \tag{2.29}$$

где r_{jk} – выборочный коэффициент корреляции, r'_{jk} – вычисленный коэффициент корреляции, полученный с помощью отображения. В случае некоррелированных общих факторов выражение для остаточных коэффициентов корреляции записывается в виде:

$$\bar{r}_{jk} = r_{jk} - (a_{j1}a_{k1} + a_{j2}a_{k2} + \ldots + a_{jm}a_{km}). \tag{2.30}$$

Теперь, изучая значения и распределение остатков $\bar{r}_{jk}$ можно судить о степени приближения выборочных коэффициентов корреляции вычисленным. Вообще говоря, значения остатков близки к нулю. Поскольку вычисление остатков связано с устранением влияния общих факторов, это означает устранение взаимосвязей между переменными. Можно, следовательно, ожидать, что остатки распределены так же, как некоторая выборка равного объема с нулевой корреляцией. Стандартное отклонение такой выборки определяется формулой

$$\begin{gathered} \sigma_{r=0} = \frac{1}{\sqrt{N-1}}, \\ \sigma_{\bar{r}} \le \frac{1}{\sqrt{N}}, \end{gathered} \tag{2.31}$$

где $\sigma_{\bar{r}}$ есть стандартное отклонение совокупности остатков. Такой показатель качества модели не очень хорош, так как он учитывает только размер выборки, в то время как значения остатков могут также зависеть и от других характеристик выборки, в первую очередь, от числа переменных.

На основании критерия (2.31), зная только размер выборки, можно делать, например, следующие заключения. Если $\sigma_{\bar{r}}$ значительно больше, чем $\frac{1}{\sqrt{N}}$, то какие-то связи между параметрами остались неучтенными и, следовательно, модель нуждается в модификации. Если, наоборот, $\sigma_{\bar{r}}$ значительно меньше $\frac{1}{\sqrt{N}}$ то, возможно, в модель включены несущественные связи между переменными. И, наконец, если стандартное отклонение остатков лишь немного меньше отклонения для выборки с нулевой корреляцией, то факторное решение может считаться приемлемым в смысле критерия (2.31).

В любой науке наблюдаемое явление может быть описано многими, не противоречащими друг другу способами. Выбор конкретного способа интерпретации зависит от цели работы. Эта произвольность или неопределенность, долгое время бывшая предметом исследования ученых, кратко отмечена Мултаном [29]: «любая группа явлений может быть непротиворечиво описана разными путями, вернее, с помощью бесконечно большого числа путей. Независимо от причин, по которым мы выбираем способ интерпретации, мы можем выбрать любой способ, кажущийся нам наиболее целесообразным. Если бы ученые всегда помнили, сколь много приемлемых интерпретаций наблюдаемых явлений может быть выдвинуто, то быть может, они отказались бы от ультимативных попыток поиска единственно верных научных теорий».

Задачи факторного анализа точно так же являются неопределенными в том смысле, что для заданного набора переменных и коэффициентов корреляции между ними, коэффициенты факторного отображения могут быть вычислены неоднозначно. Иначе говоря, может быть найдено бесконечное число ортогональных (независимых) систем факторов, адекватно описывающих выборочные коэффициенты корреляции. Это свойство было известно математикам еще со времен первых исследований в факторном анализе. Впервые оно было формально изложено Хотеллингом [30] в применении к методу главных компонент. Много позже Т. Андерсон и Г. Рубин [31] дали краткое доказательство этого факта в терминах матричной алгебры.

Неопределенность модели, т.е. неоднозначность факторных нагрузок a_{jp}, имеет своей причиной то обстоятельство, что факторное решение, определяя m-мерное пространство, содержащее общие факторы, не определяет базиса в этом пространстве, а следовательно, не определяет положения факторов в нем. С этой неопределенностью приходится сталкиваться дважды: на первой этапе, − при поиске какого-либо решения, удовлетворяющего модели в статистическом смысле, и на втором этапе, − при придании этому решению вида, наиболее удобного с точки зрения интерпретации. Большинство вычислительных процедур факторного анализа дают неоднозначное решение для факторных нагрузок, исключение составляет метод главных компонент. В то же время, любое решение может быть найдено в каноническом виде после того, как факторное решение, достаточно хорошо описывающее выборку, найдено. Система факторов подвергается такому вращению, чтобы полученная в итоге система (столь же хорошо описывающая выборку, что и исходная) оказалась более интерпретируемой с точки зрения специалистов соответствующей области.

Запишем некоторые из введенных ранее понятий в компактном матричном обозначении. Возникающие при этом удобства связаны не только с простотой записи: часто свойства, ускользающие в обычной громоздкой записи, выявляются в матричных обозначениях. Выпишем в матричной форме некоторые фундаментальные уравнения для факторных отображений и структур.

Введем матричное обозначение некоторых основных понятий факторного анализа. Так, совокупность n рассматриваемых нормированных переменных представляется в виде вектора (матрицы строки)

$$\boldsymbol{z} = (Z_1\ Z_2 \dots Z_n); \tag{2.32}$$

совокупность всех N нормированных наблюдаемых значений всех n переменных – в виде матрицы:

$$\boldsymbol{Z} = \begin{pmatrix} z_{11} & z_{12} & \cdots & z_{1n} \\ z_{21} & z_{22} & \cdots & z_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ z_{N1} & z_{N2} & \cdots & z_{Nn} \end{pmatrix}. \tag{2.33}$$

Аналогично представляются факторы:

$$\boldsymbol{f} = \left(F_1\ F_2 \dots F_m\right),\ \boldsymbol{F} = \begin{pmatrix} f_{11} & f_{12} & \dots & f_{1m} \\ f_{21} & f_{22} & \dots & f_{2m} \\ & \dots & \dots & \\ f_{N1} & f_{N2} & \dots & f_{Nm} \end{pmatrix}, \tag{2.34}$$

$$\boldsymbol{u} = \left(U_1\ U_2 \dots U_n\right),\ \boldsymbol{U} = \begin{pmatrix} u_{11} & u_{12} & \dots & u_{1n} \\ u_{21} & u_{22} & \dots & u_{2n} \\ & \dots & \dots & \\ u_{1N} & u_{2N} & \dots & u_{Nn} \end{pmatrix}. \tag{2.35}$$

Коэффициенты при факторах в факторном отображении можно представить в виде матрицы:

$$\boldsymbol{M} = \left(\begin{array}{cccc|cccc} a_{11} & a_{12} & \dots & a_{1m} & d_1 & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & a_{2m} & 0 & d_2 & \dots & 0 \\ & \dots & & & & & \dots & \\ a_{n1} & a_{n2} & \dots & a_{nm} & 0 & 0 & \dots & d_n \end{array}\right) = \left(\boldsymbol{A} \mid \boldsymbol{D}\right), \tag{2.36}$$

где полная матрица $\boldsymbol{M}$ отображения состоит из матрицы $\boldsymbol{A}$ коэффициентов при общих факторах и диагональной матрицы $\boldsymbol{D}$ коэффициентов при характерных факторах.

Пользуясь введенными матрицами, можно переписать выражение для факторного отображения (2.23) в виде:

$$\boldsymbol{Z} = \left(\boldsymbol{f} \mid \boldsymbol{u}\right)\boldsymbol{M}^T = \left(\boldsymbol{f} \mid \boldsymbol{u}\right)\left(\frac{\boldsymbol{A}^T}{\boldsymbol{D}}\right) = \boldsymbol{f}\boldsymbol{A}^T + \boldsymbol{u}\boldsymbol{D}, \tag{2.37}$$

Кроме отображения, факторный анализ изучает также структуру, т.е. таблицу коэффициентов корреляции между переменными и факторами (отображение совпадает со структурой только в случае некоррелированных факторов). Коэффициенты корреляции параметров с характерными факторами совпадают с коэффициентами при характерных факторах в (2.25), и поэтому диагональную матрицу значений этих коэффициентов корреляции естественно обозначить через $\boldsymbol{D}$. Факторную структуру можно изобразить в виде матрицы:

$$\boldsymbol{S} = \begin{pmatrix} s_{11} & s_{12} & \dots & s_{1n} \\ s_{21} & s_{22} & \dots & s_{2n} \\ & & \dots & \\ s_{m1} & s_{m2} & \dots & s_{mn} \end{pmatrix}, \tag{2.38}$$

где через s_{pj} обозначены коэффициенты корреляции переменных с общими факторами: $s_{pj} = r_{F_p Z_j}$, $(p = 1,2,\dots,m;\ j = 1,2,\dots,n)$.

Выразив в матричной форме факторное отображение и структуру, можно рассмотреть некоторые виды взаимосвязи между ними. Запишем факторное отображение (1.37) с учетом *N* объектов:

$$\boldsymbol{Z} = \boldsymbol{F}\boldsymbol{A}^T + \boldsymbol{U}\boldsymbol{D}. \tag{2.39}$$

Умножим обе части этого выражения на транспонированную матрицу факторных значений и разделим на скаляр *(N-1):*

$$\frac{1}{N-1}\boldsymbol{F}^T\boldsymbol{Z} = \left(\frac{1}{N-1}\boldsymbol{F}^T\boldsymbol{F}\right)\boldsymbol{A}^T + \left(\frac{1}{N-1}\boldsymbol{F}^T\boldsymbol{U}\right)\boldsymbol{D}. \tag{2.40}$$

Левая часть полученного выражения есть не что иное, как $\boldsymbol{S}$:

$$\frac{1}{N-1}\boldsymbol{F}^T\boldsymbol{Z}=\boldsymbol{S}, \tag{2.41}$$

и каждый элемент этой матрицы есть коэффициент корреляции между переменными и факторами. Поскольку факторы как общие, так и характерные заданы в стандартном виде, то выражения в круглых скобках правой части (2.40) представляют собой также матрицы корреляций, соответственно, между общими факторами и между общими и характерными факторами. Обозначим матрицу корреляций между общими факторами через $\boldsymbol{R}_{F,F}$, она имеет вид:

$$\boldsymbol{R}_{F,F}=\frac{1}{N-1}\boldsymbol{F}^T\boldsymbol{F}=\begin{pmatrix} 1 & r_{F_1F_2} & \dots & r_{F_1F_m} \\ r_{F_2F_1} & 1 & \dots & r_{F_2F_m} \\ & & \dots & \\ r_{F_mF_1} & r_{F_mF_2} & \dots & r_{F_mF_m} \end{pmatrix}. \tag{2.42}$$

Одним из предположений факторного анализа есть отсутствие корреляции между общими и характерными факторами, т.е. матрица корреляций между общими и характерными факторами равна нулю:

$$\boldsymbol{R}_{F,U}=\frac{1}{N-1}\boldsymbol{F}^T\boldsymbol{U}=\boldsymbol{0}. \tag{2.43}$$

Подставляя в (2.40) выражения (2.41), (2.42) и (2.43), получим

$$\boldsymbol{S}=\boldsymbol{R}_{F,F}\boldsymbol{A}^T. \tag{2.44}$$

Таким образом, получено фундаментальное соотношение, связывающее факторное отображение $\boldsymbol{A}$ с факторной структурой $\boldsymbol{S}$, а именно: матрица факторной структуры равна произведению от умножения матрицы отображения на матрицу коэффициентов корреляции между факторами. Из (2.44) видно, что если факторы некоррелированны (т.е. $\boldsymbol{R}_{F,F}$ есть единичная матрица), то элементы структуры равны соответствующим элементам отображения. Умножив справа обе части (2.44) на обратную матрицу корреляций между общими факторами, получим соответствующее выражение для отображения:

$$\boldsymbol{A}^T=\boldsymbol{R}_{F,F}^{-1}\boldsymbol{S} \quad \text{или} \quad \boldsymbol{A}=\boldsymbol{S}^T\boldsymbol{R}_{F,F}^{-1}. \tag{2.45}$$

Пользуясь соотношением между отображением и структурой, можно теперь установить несколько соотношений для матрицы вычисленных корреляций. По определению, матрица выборочных коэффициентов корреляции равна:

$$\boldsymbol{R}=\frac{1}{N-1}\boldsymbol{Z}^T\boldsymbol{Z}. \tag{2.46}$$

Подставив в это уравнение (2.39), получим:

$$\boldsymbol{R}=\frac{1}{N-1}\left(\boldsymbol{F}\boldsymbol{A}^T+\boldsymbol{U}\boldsymbol{D}\right)^T\left(\boldsymbol{F}\boldsymbol{A}^T+\boldsymbol{U}\boldsymbol{D}\right)=\frac{1}{N-1}\left(\boldsymbol{A}\boldsymbol{F}^T+\boldsymbol{D}\boldsymbol{U}^T\right)\cdot\left(\boldsymbol{F}\boldsymbol{A}^T+\boldsymbol{U}\boldsymbol{D}\right)=$$

$$=\frac{1}{N-1}\left(\boldsymbol{A}\boldsymbol{F}^T\boldsymbol{F}\boldsymbol{A}^T+\boldsymbol{A}\boldsymbol{F}^T\boldsymbol{U}\boldsymbol{D}+\boldsymbol{D}\boldsymbol{U}^T\boldsymbol{F}\boldsymbol{A}^T+\boldsymbol{D}\boldsymbol{U}^T\boldsymbol{U}\boldsymbol{D}\right)=\boldsymbol{A}\boldsymbol{R}_{F,F}\boldsymbol{A}^T+\boldsymbol{D}\boldsymbol{R}_{U,U}\boldsymbol{D}; \tag{2.47}$$

где учтена нулевая корреляция между общими и характерными факторами, $\boldsymbol{R}_{U,U}$ – матрица корреляции между характерными факторами, поскольку они независимы друг от друга и представлены в стандартном виде , то эта матрица равна единичной. Таким образом, выражение (2.47) можно записать в виде:

$$\boldsymbol{R}=\boldsymbol{A}\boldsymbol{R}_{F,F}\boldsymbol{A}^T+\boldsymbol{D}^2. \tag{2.48}$$

Если общие факторы не коррелируют друг с другом, что в большинстве случаев предполагается в факторном анализе экономических процессов, то матрица корреляции общих факторов равна единичной, и выражение (2.48) окончательно приобретает вид:

$$\boldsymbol{R} = \boldsymbol{A}\boldsymbol{A}^{T} + \boldsymbol{D}^{2}\ . \tag{2.49}$$

Уравнение (2.49) в факторном анализе называется «фундаментальной факторной теоремой». Следует отметить, что правая часть (2.49) представляет собой истинные, но неизвестные значения коэффициентов парной корреляции между переменными, которые обозначаются греческой буквой $\boldsymbol{P}$, и, строго говоря, именно эту букву следует употреблять в последнем выражении. Но в факторном анализе настолько «устоялось» употребление обозначений, используемых в (2.49), что вслед за этим также будут употребляться эти обозначения. Но там, где будет принципиально важно подчеркнуть, что речь идет собственно об истинных корреляциях, будет употребляться греческая буква.

Факторный анализ занимается описанием массивов экспериментальных данных (выборочных коэффициентов корреляции) с помощью модели − факторного отображения (2.23) ИЛИ в матричном обозначении (2.39). Как видно из (2.49), имея факторное отображение, можно вычислить коэффициенты корреляции между переменными, пользуясь только коэффициентами при общих факторах. Для того, чтобы матрица вычисленных коэффициентов корреляция адекватно описывала матрицу выборочных коэффициентов корреляции, необходимо, чтобы ее диагональные элементы вычислялись, исходя из части отображения, отвечающей общим факторам. Таким образом, если на главной диагонали матрицы выборочных коэффициентов корреляции стоят числа, приближающие общности, то факторное решение будет включать и характерные факторы. С другой стороны, если на главной диагонали матрицы выборочных коэффициентов корреляции стоят единицы, то, для того чтобы уравнение (2.49) позволило получать единицы в матрице вычисленных коэффициентов корреляции, необходимо, чтобы факторное решение включало только общие факторы. В этой случае не делается предположения о наличии характерных факторов и, следовательно, используется модель компонентного анализа. Если же на диагонали стоят надежности – величины в диапазоне между общностями и единицей, то факторное решение будет включать общие факторы и факторы ошибки, а специфичность будет учтена дисперсией общих факторов. В зависимости от того, какие числа стоят на главной диагонали матрицы выборочных коэффициентов корреляции, общие факторы учитывают разные части дисперсии переменных.

2.2. Оценка общности стохастических факторов и исследуемых переменных экономической системы.

Разработано много способов предварительных оценок общностей. Один Тэрстоун [32] предложил двенадцать различных способов. Но пока все способы остаются без исчерпывающего теоретического обоснования.

Один из самых распространенных способов оценки общности состоит в выборе наибольшего коэффициента корреляции данной переменной с остальными переменными. В соответствии с этим способом в строке корреляционной матрицы данной переменной выбирается наибольший коэффициент парной корреляции вне зависимости от его знака, и он используется в качестве оценки общности. При этом отрицательный коэффициент берется с положительным знаком. Никаких дополнительных вычислений не требуется, что является достоинством данного метода. Именно его простота определила широкое распространение метода в факторном анализе.

Однако наибольший коэффициент корреляции в строке корреляционной матрицы не имеет непосредственной связи с общностью. Коэффициент корреляции является случайной величиной и каждое его возможное значение зависит от случайных обстоятельств, в то время как общность является следствием воздействия на значения переменных и на их корреляции с общими факторами. Поэтому значения общностей и их оценки должны отражать наиболее существенные аспекты факторной структуры. Как уже отмечалось, грубая оценка общности мало влияет на точность факторного решения, но лишь при очень большом числе исследуемых признаков. Когда число переменных невелико (хотя как провести грань между понятиями «велико» и «невелико»), то подобная оценка может существенно исказить факторное решение. Такое исследование провел Иберла [41], которое показало, что даже при незначительном изменении общностей, подставляемых на главную диагональ корреляционной матрицы при оценке параметров факторной модели, значения факторных нагрузок меняются значительно. При этом число переменных было не так уж и мало: оно составляло 12 показателей.

Другим по степени распространенности способом оценки общности является использование коэффициента множественной корреляции, используемого в регрессионном анализе, т.е. меры связи одной переменной со всеми остальными переменными.

Теоретические принципы для этого способа основаны на следующих соображениях. Квадрат коэффициента корреляции между переменной и факторами, определяющими его значения и взаимосвязи с другими переменными, равен общности данной переменной. Такой коэффициент корреляции между переменной и общими факторами равен коэффициенту множественной корреляции, но имеющую другую природу по сравнению с аналогичным понятием, используемым вне факторного анализа. И, если объективно подойти к этой проблеме, то следует отметить, что по определению, коэффициентом множественной корреляции вектора Z с векторами $F_1, F_2, ..., F_m$ называется величина

$$\min_{\alpha_1, \alpha_2, ... \alpha_m} \sqrt{\overline{\left(Z - \sum_j \boldsymbol{\alpha}_j F_j \right)^2}} \, ,$$

где черта сверху означает среднее значение; таким образом, общность переменной равна квадрату коэффициента множественной корреляции между переменной и общими факторами:

$$h_j^2 = R^2_{Z_j \cdot F_1, F_2, ..., F_m} \, . \qquad (2.50)$$

Поскольку в качестве базиса пространства теоретические значения переменных, получаемых по факторному уравнению, также хороши, как и сами факторы [42], то эти переменные можно взять в качестве базиса пространства, в котором можно определить величины коэффициентов множественной корреляции. Поэтому общности можно получить по выражению

$$h_j^2 = R^2_{Z_j \cdot Z_1^*, ..., Z_n^*} \, , \qquad (2.51)$$

где $Z_1^*, ..., Z_n^*$ − теоретические значения переменных, полученных по факторной модели. Кайзер [43] показал, что если *n* векторов $Z_1^*, ..., Z_n^*$ лежат в пространстве общих факторов, то при вычислении квадрата коэффициента множественной корреляции (2.1.2) некоторые из них можно исключить. Исключение переменной Z_j^* из уравнения факторной модели, приводит к выражению

$$h_j^2 = R^2_{Z_j \cdot Z_1^*, ..., Z_{j-1}^*, Z_{j+1}^*, ..., Z_n^*} \, . \qquad (2.52)$$

Таким образом, общность переменной равна квадрату его коэффициента множественной корреляции с *n-1* «общими частями» остальных переменных. Очевидная трудность этого подхода состоит в том, что для вычисления искомой общности, необходимо знать значения остальных *n-1* общностей. Кайзер предложил итеративную процедуру, рассчитанную на использование вычислительных возможностей компьютерной техники, но процедура сходится лишь для некоторых корреляционных матриц, отсюда он заключил, что метод «…не имеет практической ценности для эмпирических матриц»[43, с.10].

Ранее уже отмечалось, что квадрат коэффициента множественной корреляции, вычисленный из уравнения зависимости данной переменной и остальными, может рассматриваться как нижняя граница оценки общности. Сам коэффициент множественной корреляции известен из регрессионного анализа, и он является мерой дисперсии данной переменной, общей с другими переменными исследуемого множества. В то время как общность является мерой дисперсии данной переменной, обусловленной общими для нескольких переменных факторами. Разность между ними является общей дисперсией, которая может быть приписана только данной переменной, поэтому использовать квадрат коэффициента множественной корреляции, вычисляемой по уравнению регрессии, не может быть оправдано с теоретической точки зрения.

Более того, редуцированная корреляционная матрица, т.е. матрица, у которой главная диагональ заменена значениями общности, уже не будет матрицей Грамма. В этом смысле значения квадратов коэффициентов множественной корреляции не являются оценками общности. Тем не менее, несмотря на теоретическую недопустимость, этот способ получил самое широкое распространение в факторном анализе, в силу того, что он давал наиболее точные по сравнению с другими способами оценки общности.

Перейдем к рассмотрению принципиально нового подхода к получению оценок общностей, основанного на свойствах информации, содержащейся в выборочных данных.

Понятие информации интуитивно толкуется как некоторая совокупность сведений, определяющих меру знаний о тех или иных событиях, явлениях, фактах и прочее. Однако такое определение информации ничего не дает для построения количественной теории информации, которая могла быть использована для решения различных задач и не только инженерных, но математических и экономических.

Очевидно, что всякая информация получается в результате некоторого опыта или наблюдения. Можно сказать, что до опыта нельзя однозначно ответить на интересующий нас вопрос, можно лишь высказать ряд предположений. Таким образом, до опыта имеется некоторая неопределенность в соответствующей ситуации, в исходе тех или иных событий. После опыта, т.е. получения информации, можно ответить на вопрос либо однозначно, либо число предположений уменьшится, и, соответственно, уменьшится неопределенность. Таким образом, для оценки количества получаемой информации необходимо найти меру неопределенности той или иной ситуации.

Уменьшение неопределенности в результате опыта можно принять за наиболее общую меру количества получаемой информации. В этом смысле говорят, что информация обратна неопределенности.

В теории информации известно [44], что информация относительно некоторой точки дискретного пространства x_k (точка такого пространства − это элементарное событие), содержащаяся в y_i − точке другого дискретного пространства, сводится к измерению вероятности x_k от ее априорного значения $p(x_k)$ к ее апостериорному значению $p(x_k \mid y_i)$. Количество информации, содержащееся в событии y_i относительно появления события x_k, определяется как

$$I(x_k, y_i) = \log \frac{p(x_k \mid y_i)}{p(x_k)} . \tag{2.53}$$

Основание логарифма, используемого в (2.53), фиксирует величину единицы измерения информации. Чаще всего используют основание 2, в этом случае она называется «бит». Единица информации, измеренная в битах, означает, что вероятность события x_k увеличилась в 2 раза.

Количество информации (2.53) обладает очень важным свойством симметрии по отношению к x_k и y_i, т.е.

$$I(x_k, y_i) = I(y_i, x_k) . \tag{2.54}$$

Таким образом, можно заметить:

$$I(x_k, y_i) = I(y_i, x_k) \le \begin{cases} \log\dfrac{1}{p(x_k)}, \\ \log\dfrac{1}{p(y_i)}. \end{cases} \qquad (2.55)$$

О величинах правой части (2.55) говорят, что они являются количеством собственной информации для x_k и для y_i. Т.е. собственная информация события x_k обозначается и равна

$$I(x_k) = -\log p(x_k). \qquad (2.56)$$

Таким образом, информация в каком-либо событии измеряется логарифмом величины обратной вероятности его появления.

Дисперсия переменной случайной величины также несет информацию. По определению, дисперсией называется математическое ожидание квадрата отклонения случайной величины от своего математического ожидания. Тем самым, в дисперсии заключена информация о мере колеблемости рассматриваемой величины. Общая и характерная части дисперсии переменной так же являются частями информации, и к ним может быть применен информационный подход к оценке общности и характерности.

Корреляционная матрица парных корреляций между всеми переменными рассматриваемой совокупности показателей, содержит полную информацию обо всех признаках и взаимосвязях между ними. При этом истинные значения всех корреляций, составляющих матрицу $\boldsymbol{P}$ ($\boldsymbol{\rho}_{ij}$ − ее элементы), − неизвестны. Факторный анализ исходит из того, что матрица истинных значений корреляций выражается через, опять таки, неизвестные – значения факторных нагрузок (фундаментальная теорема факторного анализа):

$$\boldsymbol{P} = \boldsymbol{A}\boldsymbol{A}^T + \boldsymbol{D}^2. \qquad (2.57)$$

Исключение из матрицы какой-либо переменной не поменяет оставшиеся коэффициенты корреляции, но изменит количество информации, содержащейся в ней. Если определить это информационное изменение, то может быть получен ключ к оценке общности переменной.

Поскольку информация измеряется логарифмом обратной вероятности, определим вероятности соответствующих событий. Какие это события?

Информация о самой переменной может содержаться в событии, заключающемся в том, что истинная матрица $\boldsymbol{P}$ равна выборочной корреляционной матрице без коэффициентов корреляции данной переменной с остальными, т.е. они вычеркнуты из матрицы. Другими словами, реализация этого события говорит о том, что на самом деле присутствие данной переменной в анализе явления необязательно, поскольку она никой полезной информации не несет и на остальные переменные не влияет. Обозначим это событие $A = \{\boldsymbol{P} = \boldsymbol{R}_{-j}\}$.

Другое событие может быть следующим: корреляции между всеми переменными отсутствуют, т.е. истинная корреляционная матрица равна единичной, т.е. такой, у которой на главной диагонали стоят единицы, а остальные элементы − нули. В таком событии заключается информация обо всех корреляциях между переменными, если точно выбрать соответствующую вероятность. Это событие назовем *B*, и его обозначение имеет вид $B = \{\boldsymbol{P} = \boldsymbol{I}\}$.

Вероятности этих событий − условные, поскольку все эти события происходят тогда, когда опыт дает конкретную выборочную корреляционную матрицу $\boldsymbol{R}$. Таким образом, определим вероятности: $p(\boldsymbol{P} = \boldsymbol{R}_{-j} \mid \boldsymbol{R})$ и $p(\boldsymbol{P} = \boldsymbol{I} \mid \boldsymbol{R})$.

Чтобы найти вероятности, нужно знать закон распределения случайных величин. В данном случае такими величинами являются коэффициенты выборочной корреляции.

Известна функция совместного распределения элементов выборочной корреляционной матрицы $\boldsymbol{R}$, знакомая под названием распределения Уишарта. Она имеет вид [15]

$$W(\boldsymbol{P})=\boldsymbol{\gamma}_{N,n}\left|\boldsymbol{P}\right|^{-\frac{N-1}{2}}\left|\boldsymbol{R}\right|^{\frac{N-n-2}{2}}e^{-\frac{N-1}{2}tr\left(\boldsymbol{P}^{-1}\boldsymbol{R}\right)}, \tag{2.58}$$

где N – число наблюдений,

n – число переменных,

tr – операция вычисления следа матрицы (след матрицы равен сумме элементов на главной диагонали),

$\boldsymbol{\gamma}_{N,n}$ – коэффициент распределения Уишарта:

$$\boldsymbol{\gamma}_{N,n}=\left(2^{(N+n+1)/2}\boldsymbol{\pi}^{n(n-1)/4}\prod_{i=1}^{n}\boldsymbol{\Gamma}\left(\frac{N-i+1}{2}\right)\right)^{-1}, \tag{2.59}$$

$\Gamma(\cdot)$ – гамма-функция, определяемая выражениями:

$$\boldsymbol{\Gamma}(x)=(x-1)\boldsymbol{\Gamma}(x-1), \boldsymbol{\Gamma}(1)=1,\ \boldsymbol{\Gamma}(1/2)=\sqrt{\boldsymbol{\pi}}.$$

Зная это распределение, можно найти искомые вероятности.

Условная вероятность события А вычисляется по формуле

$$p(\boldsymbol{P}=\boldsymbol{R}_{-j}\,|\,\boldsymbol{R})=\frac{p(\boldsymbol{R}_{-j}\cap\boldsymbol{R})}{p(\boldsymbol{R})}=\frac{p(\boldsymbol{R}_{-j})}{p(\boldsymbol{R})}, \tag{2.60}$$

а условная вероятность события В – равна

$$p(\boldsymbol{P}=\boldsymbol{I}\,|\,\boldsymbol{R})=\frac{p(\boldsymbol{I}\cap\boldsymbol{R})}{p(\boldsymbol{R})}=\frac{p(\boldsymbol{I})}{p(\boldsymbol{R})}. \tag{2.61}$$

Используя функцию распределения Уишарта, найдем соответствующие вероятности.

$$p(\boldsymbol{R}_{-j})=W(\boldsymbol{R}_{-j})=\boldsymbol{\gamma}_{N,n}\,|\boldsymbol{P}|^{-\frac{N-1}{2}}\,|\boldsymbol{R}_{-j}|^{\frac{N-n-2}{2}}\,e^{-\frac{N-1}{2}tr\boldsymbol{P}^{-1}\boldsymbol{R}_{-j}}\Bigg|_{0}^{\boldsymbol{P}=\boldsymbol{R}_{-j}}, \tag{2.62}$$

После подстановки верхнего, нижнего пределов и небольших преобразований, получаем:

$$p(\boldsymbol{R}_{-j})=\boldsymbol{\gamma}_{N,n}\,|\boldsymbol{R}_{-j}|^{-\frac{n+1}{2}}\,e^{-\frac{N-1}{2}(n-1)}. \tag{2.63}$$

Вероятность того, что истинные корреляции равны выборочным значениям, может быть получена из выражения:

$$p(\boldsymbol{R})=W(\boldsymbol{R})=\boldsymbol{\gamma}_{N,n}\,|\boldsymbol{P}|^{-\frac{N-1}{2}}\,|\boldsymbol{R}|^{\frac{N-n-2}{2}}\,e^{-\frac{N-1}{2}tr\boldsymbol{P}^{-1}\boldsymbol{R}}\Bigg|_{0}^{\boldsymbol{P}=\boldsymbol{R}}, \tag{2.64}$$

в окончательном виде (2.64) становится

$$p(\boldsymbol{R})=\boldsymbol{\gamma}_{N,n}\,|\boldsymbol{R}|^{-\frac{n+1}{2}}\,e^{-\frac{N-1}{2}n}. \tag{2.65}$$

Вероятность отсутствия корреляций между переменными вычисляется на основании выражения и равно:

$$p(\boldsymbol{I})=W(\boldsymbol{I})=\boldsymbol{\gamma}_{N,n}\,|\boldsymbol{P}|^{-\frac{N}{2}}|\boldsymbol{R}|^{\frac{N-n}{2}}\,e^{-\frac{N-1}{2}tr\boldsymbol{P}^{-1}\boldsymbol{R}}\Bigg|_{0}^{\boldsymbol{P}=\boldsymbol{I}}=\boldsymbol{\gamma}_{N,n}\,|\boldsymbol{R}|^{\frac{N-n-2}{2}}\,e^{-\frac{N-1}{2}n} \tag{2.66}$$

Подставляя найденные вероятности в выражения (2.60) и (2.61) получим, соответственно для условной вероятности отсутствия влияния переменной Z_j при условии, что она имеется в корреляционной матрице:

$$p(\boldsymbol{P}=\boldsymbol{R}_{-j}\mid \boldsymbol{R})=|\boldsymbol{R}|^{(n+1)/2}|\boldsymbol{R}_{-j}|^{-(n+1)/2}e^{(N-1)/2}, \tag{2.67}$$

и для условной вероятности отсутствия корреляции между переменными, при условии, что в реальности она имеет место:

$$p(\boldsymbol{P}=\boldsymbol{I}\mid \boldsymbol{R})=|\boldsymbol{R}|^{(N-1)/2}. \tag{2.68}$$

Теперь найдем информацию, содержащуюся в корреляционной матрице о переменной Z_j − она равна логарифму обратной вероятности:

$$\begin{aligned} I(Z_j,\boldsymbol{R})&=-\ln\left(|\boldsymbol{R}|^{(n+1)/2}|\boldsymbol{R}_{-j}|^{-(n+1)/2}e^{(N-1)/2}\right)=\\ &=-\frac{n+1}{2}\ln|\boldsymbol{R}|+\frac{n+1}{2}\ln|\boldsymbol{R}_{-j}|-\frac{N-1}{2}. \end{aligned} \tag{2.69}$$

Информация обо всех переменных в корреляционной матрице имеет вид:

$$I(\boldsymbol{Z},\boldsymbol{R})=-\ln\left(|\boldsymbol{R}|^{(N-1)/2}\right)=-\frac{N-1}{2}\ln|\boldsymbol{R}|. \tag{2.70}$$

Информация о самой переменной, которая получается на основании вычисления вероятности о том, что данная переменная в корреляциях не участвует, измеряет «характерность» переменной. Именно изменение обобщенной величины корреляции от нуля до реальной величины, полученной за счет данной переменной, несет такую информацию, однако она является относительной. Теперь, если эту информацию исключить из информации о корреляциях всех переменных, то относительная доля такого изменения даст измерение той части дисперсии, которая приходится на данную переменную со стороны других, что и составляет общность переменной. Таким образом, значение общности переменной определяется на основании выражения:

$$h_j^2=\frac{I(\boldsymbol{Z},\boldsymbol{R})-I(\boldsymbol{Z}_{-j},\boldsymbol{R})}{I(\boldsymbol{Z},\boldsymbol{R})}=1-\frac{I(\boldsymbol{Z}_{-j},\boldsymbol{R})}{I(\boldsymbol{Z},\boldsymbol{R})}. \tag{2.71}$$

Вычитаемая из единицы величина представляет долю информации о корреляции данной переменной, приходящейся на полную информацию, что численно равно характерности данной переменной. А сумма общности и характерности как раз и равна единице, что подтверждает внутреннюю логику и непротиворечивость данных рассуждений.

После подстановки в (2.71) величин информаций (2.69) и (2.70), проведя небольшие преобразования, окончательно получим значение общности переменной Z_j:

$$h_j^2=1-\frac{(n+1)\left(\ln|\boldsymbol{R}|-\ln|\boldsymbol{R}_{-j}|\right)-N+1}{(N-1)\ln|\boldsymbol{R}|}. \tag{2.72}$$

Теперь остается проверить, на сколько значение (2.72) близко к истинной величине. Пользуясь реальными данными, сделать это непросто, поскольку истинные значения общностей неизвестны. Более того, их значение, равное сумме квадратов коэффициентов корреляций переменной с факторами, не может не зависеть от числа факторов, которое до момента начала исследования экономического явления так же неизвестно. Поэтому проверка может быть только на смоделированном объекте, таком, как, к примеру, в первом разделе.

Будем считать, что для объяснения корреляций между шестью переменными имеется два фактора. Пусть матрица факторных нагрузок имеет вид:

$$\boldsymbol{A}=\begin{pmatrix}0{,}8 & 0{,}1\\ 0{,}7 & 0{,}4\\ 0{,}9 & 0{,}2\\ 0{,}1 & 0{,}9\\ 0{,}3 & 0{,}7\\ 0{,}2 & 0{,}8\end{pmatrix}. \tag{2.73}$$

Тогда редуцированная матрица коэффициентов корреляций, т.е. такая матрица, у которой на главной диагонали вместо единиц стоят значения общностей, имеет вид

$$\widetilde{\boldsymbol{R}}=\boldsymbol{A}\boldsymbol{A}^{T}=\begin{pmatrix}0{,}65 & 0{,}6 & 0{,}74 & 0{,}17 & 0{,}31 & 0{,}24\\ 0{,}6 & 0{,}65 & 0{,}71 & 0{,}43 & 0{,}49 & 0{,}46\\ 0{,}74 & 0{,}71 & 0{,}85 & 0{,}27 & 0{,}41 & 0{,}34\\ 0{,}17 & 0{,}43 & 0{,}27 & 0{,}82 & 0{,}66 & 0{,}74\\ 0{,}31 & 0{,}49 & 0{,}41 & 0{,}66 & 0{,}58 & 0{,}62\\ 0{,}24 & 0{,}46 & 0{,}34 & 0{,}74 & 0{,}62 & 0{,}68\end{pmatrix}, \tag{2.74}$$

и эти значения общностей будем считать, с точки зрения вычислительной процедуры, истинными или эталонными.

Сравним с этими значениями значения общностей, полученных посредством предлагаемого метода, и признаваемого на настоящий момент наиболее точным способом − квадрата коэффициента множественной корреляции.

Оценку квадрата коэффициента множественной корреляции данной переменной с остальными переменными можно определить лишь на основании уравнения зависимости этой переменной с остальными, т.е. регрессионного уравнения. Но известен другой способ, в соответствии с которым он определяется на основании выражения [45]

$$R^2_{Z_j, Z_1 \dots Z_{j-1}, Z_{j+1}, \dots, Z_n}=1-\frac{1}{r^{jj}}, \tag{2.75}$$

где r^{jj} − диагональный элемент обратной матрицы $\boldsymbol{R}^{-1}$, стоящий на j –ом месте. Оценки значений общностей, полученных в соответствии с этим способом обозначим $\hat{h}^2_{R,j}$, т.е. $\hat{h}^2_{R,j}=R^2_{Z_j, Z_1 \dots Z_{j-1}, Z_{j+1}, \dots, Z_n}$. Соответствующие оценки общностей, полученных по матрице (2.74), но у которой на главной диагонали уже стоят единицы, как у выборочной матрицы, имеют вид:

$$\hat{\boldsymbol{h}}^2_{\boldsymbol{R}}=(0{,}565\ 0{,}589\ 0{,}663\ 0{,}625\ 0{,}521\ 0{,}591), \tag{2.76}$$

жирный шрифт означает вектор (матрицу-строку) значений.

Вектор оценок значений общностей, полученных с использованием информационного подхода по выражению (2.72) равен:

$$\hat{\boldsymbol{h}}^2=(0{,}639\ 0{,}635\ 0{,}621\ 0{,}629\ 0{,}646\ 0{,}635). \tag{2.77}$$

«Крышечка» над символом вектора значений означает оценку неизвестных, но истинных величин. При этом сами истинные величины параметров обозначаются без нее. Приведем эти значения общностей для данного примера:

$$\boldsymbol{h}^2=(0{,}65\ 0{,}65\ 0{,}85\ 0{,}82\ 0{,}58\ 0{,}68). \tag{2.78}$$

Для того, чтобы избежать субъективности при сравнении разных оценок, воспользуемся стандартной процедурой анализа соответствий между различными случайными величинами, известной в мире под названием ANOVA. Истинные значения любого неизвестного параметра в математической статистике не являются случайными величинами, таковыми есть их оценки. Истинные общности в факторном анализе также считаются величинами детерминированными. Но в нашем примере, когда элементы факторной структуры воспроизводятся искусственно, «истинные значения» получены в результате вычислительной процедуры, за которой в действительности стояли бы реальные выборочные данные. А, как известно, в выборке не может содержаться вся информация о генеральной совокупности, и в силу случайности выборки, полученные эталонные значения так же наделены природой случайности. Поэтому применение статистического подхода к анализу разных случайных величин на предмет их схожести, который предполагает задаваемую наперед величину ошибки в принятии решения, в данном случае уместен.

ANOVA (ANalysis Of VAriance) − широко распространенный метод выявления различий переменных, получаемых либо разным способом, либо в неодинаковые периоды времени [46]. Итоговым результатом метода служит принятие или отвержение, с определенным заранее уровнем значимости, гипотезы об отсутствии различий у рассматриваемых величин.

Сравним значения оценок общностей − квадратов коэффициентов множественной корреляции − с «истинными значениями». Стандартная таблица ANOVA-анализа с применением однофакторной модели имеет вид, приведенный в таблице 2.1. Напомним, фактор в дисперсионном анализе имеет иное значение, чем в статистическом факторном анализе. В данном случае для ANOVA фактором служит измерение общности.

Таблица 2.1 − Сравнение квадратов коэффициентов множественной корреляции с «истинными значениями»

	Сумма квадратов	Степени свободы	Средние квадраты	F-значение
Фактора	0,040	1	0,0400	5,822
Ошибок	0,068	10	0,0068	
Полная	0,108	11	0,0098	

Итоги анализа различий между оценками общности, полученные с помощью информационного подхода, и значений общности, заданных посредством факторных нагрузок, приведены в таблице 2.2.

Таблица 2.2 − Сравнение «информационных» оценок общности с «истинными значениями»

	Сумма квадратов	Степени свободы	Средние квадраты	F-значение
Фактора	0,015	1	0,015	2,645
Ошибок	0,057	10	0,0057	
Полная	0,072	11	0,0065	

Для проверки гипотезы о тождественности разных групп величин найденные значения F-критерия необходимо сравнить с табличным значением распределения Фишера. Для данных степеней свободы и уровня значимости α (наперед задаваемой ошибки в статистических рассуждениях), традиционно равной 0,05 в эконометрических исследованиях, эта величина имеет значение 4,965. Если вычисленное значение F-критерия больше табличного, то гипотеза об идентичности величин отклоняется с вероятностью ошибки 0,05. Если же F-значение меньше табличного, то гипотеза принимается.

Проверка показала, что у оценок общности, как квадратов коэффициентов множественной корреляции, с эталонными значениями нет ничего общего. И, напротив, предлагаемые оценки общности на основе информационного подхода согласуются с «истинными значениями» общности.

Знание более точных значений оценок общности позволит повысить точность эконометрического исследования с использованием факторного анализа.

2.3. Оценка нагрузок стохастических факторов на переменные экономической системы.

Центральной проблемой факторного анализа является проблема оценки параметров факторной структуры. Именно ее решение позволяет продуктивно использовать факторный анализ для решения различных задач. Более того, без оценок факторных нагрузок факторная модель не представляет собой никакой ценности, поскольку именно количественный, а не качественный, анализ составляет основную идею факторного анализа.

В предыдущем разделе указывались основные методы и способы получения оценок параметров факторной модели. Каждый из них имеет свои достоинства, но и каждому присущи свои недостатки. Общая проблема, связывающая все существующие методы, состоит в отсутствии однозначного решения задачи оценивания. Этот недостаток, однако, имеет свою положительную сторону: используя методы вращения факторного решения, можно получить более интерпретируемое решение. Для получения однозначного решения необходимо сформулировать дополнительные ограничения на параметры факторной структуры, и получить оценки в виде условного экстремума.

Поскольку большинство методов факторного анализа дают оценки параметров в ходе итерационных процедур, сходимость которых в значительной мере зависит от начального приближения, то следует в самом начале исследования получить достаточно хорошее приближение. В качестве начального «хорошего» приближения традиционно используются собственные векторы корреляционной матрицы. Как известно, компонентный анализ в качестве своего решения использует эти векторы. Но в компонентном анализе предполагается наличие такого количества компонент (факторов) сколько переменных рассматривается в анализе. Т.е. на предварительном этапе не учитывается число факторов. Поэтому целесообразно на этапе получения предварительного решения задать число факторов и модифицировать собственные векторы так, чтобы они приблизились к истинным значениям.

Основания для такого подхода достаточно простые: редуцированная матрица получается в результате перемножения матрицы факторных нагрузок:

$$\boldsymbol{A}\boldsymbol{A}^{\boldsymbol{T}} = \widetilde{\boldsymbol{R}} \tag{2.79}$$

где $\widetilde{\boldsymbol{R}}$ – редуцированная матрица, т.е. матрица у которой на главной диагонали стоят значения общностей.

Редуцированная корреляционная матрица может быть разложена на собственные векторы и собственные значения:

$$\widetilde{\boldsymbol{R}} = \boldsymbol{Q}\boldsymbol{L}\boldsymbol{Q}^{\boldsymbol{T}}, \tag{2.80}$$

где $\boldsymbol{Q}$ и $\boldsymbol{L}$ – матрицы, соответственно, собственных векторов и собственных значений редуцированной матрицы.

Последнее выражение может быть подставлено в (2.80)

$$\boldsymbol{A}\boldsymbol{A}^{\boldsymbol{T}} = \boldsymbol{Q}\boldsymbol{L}\boldsymbol{Q}^{\boldsymbol{T}}. \tag{2.81}$$

После несложных преобразований (2.81) преобразуется к виду

$$\boldsymbol{A} = \boldsymbol{L}^{-\boldsymbol{1}}\boldsymbol{Q}^{\boldsymbol{T}}\boldsymbol{A}\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{Q}\boldsymbol{A} \tag{2.82}$$

Выражение (2.82) дает основание для итерационной процедуры поиска значений матрицы факторных нагрузок, которые построены на основании собственных значений редуцированной корреляционной матрицы с учетом информации о числе факторов. Процедура будет построена на том предположении, что значение матрицы $\boldsymbol{A}$, стоящей в левой части выражения (2.82) уточняется вычислительной процедурой действий над матрицами, стоящими в правой части этого выражения. Таким образом, простая итерационная процедура задается уравнением

$$\boldsymbol{A}_{i+1} = \boldsymbol{L}^{-1}\boldsymbol{Q}^{T}\boldsymbol{A}_{i}\boldsymbol{A}_{i}^{T}\boldsymbol{Q}\boldsymbol{A}_{i}, \;\; i = 1,2,\ldots \tag{2.83}$$

где i – номер итерации, которая будет продолжаться, пока матрицы двух соседних итераций не будут отличаться друг от друга на величину наперед задаваемой ошибки.

Практика показывает, однако, что процедуры, подобные (2.83) имеют проблемы со схождением, которое зависит от начального приближения. Итерационная процедура может и не дать сходящийся результат для некоторых исходных данных. Причем теоретические исследования сходимости многомерных операторов известны для небольшого круга матричных итерационных процедур.

Процесс сходимости можно ускорить и, в большинстве случаев, гарантировать, если воспользоваться матричным аналогом метода Ньютона-Канторовича численного решения трансцендентного уравнения. Кратко опишем его.

Пусть необходимо решить матричное уравнение, не имеющее аналитического решения

$$F(\boldsymbol{X}) = 0, \tag{2.84}$$

т.е. в ходе итерационной процедуры найти такое значение матрицы $\boldsymbol{X}$, которое удовлетворяло бы уравнению (2.84). Такая процедура имеет вид

$$\boldsymbol{I}\otimes\boldsymbol{X}_{i+1} = \boldsymbol{I}\otimes\boldsymbol{X}_i - \left(\boldsymbol{I}\otimes\boldsymbol{H}F(\boldsymbol{X}_i)\right)\left(\frac{\partial F(\boldsymbol{X}_i)}{\partial \boldsymbol{X}_i}\right)^{-1}, \tag{2.85}$$

где $\otimes$ - знак произведения Кронекера,

$\boldsymbol{I}$ - единичная матрица,

i – также как и для (2.83) означает номер итерации,

$\boldsymbol{H}$ - идемпотентная матрица, имеющая вид

$$\boldsymbol{H} = (\boldsymbol{X}_i - \boldsymbol{X}_{i-1})\left((\boldsymbol{X}_i - \boldsymbol{X}_{i-1})^T(\boldsymbol{X}_i - \boldsymbol{X}_{i-1})\right)^{-1}(\boldsymbol{X}_i - \boldsymbol{X}_{i-1})^T, \tag{2.86}$$

на первом шаге, когда известна только $\boldsymbol{X}_0$**,** в качестве идемпотентной матрицы берется единичная, поскольку она, также как и идемпотентная, обладает свойством:

$$\boldsymbol{HH} = \boldsymbol{H}.$$

Заметим также, что для обращения производной она должна быть квадратной матрицей, что возможно в случае, когда матричный функционал $F(\boldsymbol{X})$ имеет туже размерность, что и неизвестная матрица $\boldsymbol{X}$.

В нашем случае уравнением, подлежащим решению, служит выражение (2.81), перепишем его в форме (2.86), но неизвестную матрицу будем обозначать $\boldsymbol{A}$:

$$\boldsymbol{A} - \boldsymbol{L}^{-1}\boldsymbol{Q}^T\boldsymbol{A}\boldsymbol{A}^T\boldsymbol{Q}\boldsymbol{A} = \boldsymbol{0}, \tag{2.87}$$

Левая часть (2.87) представляет собой матричный функционал $F(\boldsymbol{A})$, который имеет ту же размерность, что и неизвестная матрица $\boldsymbol{A}$**.** Его производная равна

$$\frac{\partial F(\boldsymbol{A})}{\partial \boldsymbol{A}} = \boldsymbol{I}_{nm} - 2(\boldsymbol{Q}\boldsymbol{L}^{-1}\otimes\boldsymbol{A}^T\boldsymbol{Q}\boldsymbol{A}) - \boldsymbol{Q}\boldsymbol{A}\boldsymbol{A}^T\boldsymbol{Q}\boldsymbol{L}^{-1}\otimes\boldsymbol{I}_m, \tag{2.88}$$

где n – число переменных;

m – число факторов

и нижний индекс единичной матрицы указывает на ее размерность.

Тогда итерационная процедура в соответствии с (2.85) задается выражением:

$$\begin{aligned}\boldsymbol{I}\otimes\boldsymbol{A}_{i+1} = \boldsymbol{I}\otimes\boldsymbol{A}_i - \left(\boldsymbol{I}\otimes\boldsymbol{H}_i\left(\boldsymbol{A}_i - \boldsymbol{L}^{-1}\boldsymbol{Q}^T\boldsymbol{A}_i\boldsymbol{A}_i^T\boldsymbol{Q}\boldsymbol{A}_i\right)\right)\cdot \\ \cdot\left(\boldsymbol{I}_{nm} - 2(\boldsymbol{Q}\boldsymbol{L}^{-1}\otimes\boldsymbol{A}^T\boldsymbol{Q}\boldsymbol{A}) - \boldsymbol{Q}\boldsymbol{A}\boldsymbol{A}^T\boldsymbol{Q}\boldsymbol{L}^{-1}\otimes\boldsymbol{I}_m\right)^{-1},\end{aligned} \tag{2.89}$$

где $\boldsymbol{H}_i = (\boldsymbol{A}_i - \boldsymbol{A}_{i-1})\left((\boldsymbol{A}_i - \boldsymbol{A}_{i-1})^T(\boldsymbol{A}_i - \boldsymbol{A}_{i-1})\right)^{-1}(\boldsymbol{A}_i - \boldsymbol{A}_{i-1})^T$ для $i = 1,2,3,\ldots$ и $\boldsymbol{H}_1 = \boldsymbol{I}$.

Вычислительные эксперименты показывают, что сходимость достигается достаточно быстро, причем решение получается всякий раз, каким бы ни было начальное приближение. Поэтому в качестве начальной матрицы $\boldsymbol{A}_0$ удобно брать часть редуцированной корреляционной матрицы, а именно первые m ее столбцов, или столько же любых других ее столбцов.

После того как получено хорошее приближение факторных нагрузок, их значения можно улучшить, воспользовавшись методом максимального правдоподобия. В математической статистике это наиболее мощный инструмент получения оценок на основании выборочных данных. Суть его состоит в получении таких оценок, которые бы максимизировали функцию правдоподобия. Такой функцией, как правило, служит совместная плотность распределения случайных величин, по которым имеются ряды статистических данных. Как известно [47] совместной плотностью является произведение плотностей распределения отдельных рассматриваемых случайных величин. И в соответствии с методом максимального правдоподобия значение оцениваемого параметра подбирается так, чтобы значение совместной плотности было бы максимальным, а это соответствует максимуму вероятности того, что оценка совпадает с истинным, но неизвестным значением параметра.

Максимум любой функции определяется после приравнивания нулю первой производной, но поскольку дифференцировать произведения, а после чего искать корень уравнения производной нулю, сложно, то часто совместную плотность логарифмируют.

И именно логарифм плотности считают функцией правдоподобия, тем более, что функция и ее логарифм имеют максимум в одной точке.

В факторном анализе неизвестным параметром, предназначенным для оценивания, является матрица факторных нагрузок, которая определяет значения истинной корреляционной матрицы. Плотность совместного распределения коэффициентов корреляций и составляет основу функции правдоподобия. Совместное распределение элементов корреляционной матрицы подчинено закону Уишарта, плотность распределения которого имеет вид (2.58). Логарифмирование и исключение из его результата слагаемых и множителей, не зависящих от матрицы факторных нагрузок дает следующее выражение для функции правдоподобия:

$$g = ln|\boldsymbol{P}| + tr(\boldsymbol{P}^{-1}\boldsymbol{R}), \tag{2.90}$$

где $\boldsymbol{P}$ – истинная корреляционная матрица, имеющая вид:

$$\boldsymbol{P} = \boldsymbol{A}\boldsymbol{A}^T + \boldsymbol{D}^2,$$

$\boldsymbol{D}^2$ – диагональная матрица характерностей каждой переменной;

$\boldsymbol{R}$ – матрица выборочных корреляций.

Теперь найдем производную функции правдоподобия g по неизвестному параметру $\boldsymbol{A}$:

$$\frac{\partial g}{\partial \boldsymbol{A}} = 2vec(\boldsymbol{P}^{-1}\boldsymbol{A}) - 2vec(\boldsymbol{P}^{-1}\boldsymbol{R}\boldsymbol{P}^{-1}\boldsymbol{A}), \tag{2.91}$$

где vec – операция преобразования матрицы в вектор – матрицу столбец, путем взятия строки за строкой.

Затем приравняем производную, т.е. выражение (2.91), нулю, и найдем выражение, дающее основание для итерационной процедуры:

$$\boldsymbol{A} = \boldsymbol{R}\boldsymbol{P}^{-1}\boldsymbol{A}\,. \tag{2.92}$$

Авторы метода максимального правдоподобия в факторном анализе Лоули и Максвелл [36], а вслед за ними и известный эксперт в области факторного анализа – Харман [45] отмечали сложности в сходимости процедур, построенных на выражении (2.92). Ради улучшения сходимости равенство (2.92) преобразовывали к виду, который давал возможность трактовать его как способ вычисления собственных векторов некоторой матрицы, упрощали с целью снизить размерность обращаемой матрицы (это было особенно актуально во времена ограниченных возможностей вычислительной техники) и тому подобное. Более того, исследование сходимости итерационных процедур метода максимального правдоподобия продолжается и по сей день. Однако матричный аналог метода Ньютона-Канторовича (2.95) позволяет получить гарантированно сходящуюся процедуру получения максимально правдоподобной оценки матрицы факторных нагрузок.

Матричным функционалом, нуль которого необходимо найти, является выражение:

$$F(\boldsymbol{A}) = \boldsymbol{A} - \boldsymbol{R}\boldsymbol{P}^{-1}\boldsymbol{A}. \tag{2.93}$$

Для составления итерационной процедуры найдем производную

$$F(\boldsymbol{A})F'(\boldsymbol{A}) = \boldsymbol{I}_{nm} + \boldsymbol{P}^{-1}\boldsymbol{R}\otimes\boldsymbol{A}^T\boldsymbol{P}^{-1}\boldsymbol{A} + \boldsymbol{E}_{nm}(\boldsymbol{A}^T\boldsymbol{P}^{-1}\boldsymbol{R}\otimes\boldsymbol{P}^{-1}\boldsymbol{A}) - \\ -\boldsymbol{P}^{-1}\boldsymbol{R}\otimes\boldsymbol{I}_m. \tag{2.94}$$

где $\boldsymbol{I}_{nm}$ – единичная матрица, размерностью $nm \times nm$;

$\boldsymbol{E}_{nm}$ – перестановочная матрица, возникающая при дифференцировании транспонированной матрицы по исходной, она обладает свойством:

$$\boldsymbol{E}_{nm}^T\boldsymbol{E}_{nm} = \boldsymbol{E}_{nm}\boldsymbol{E}_{nm}^T = \boldsymbol{I}_{nm}.$$

К примеру, пусть матрица $\boldsymbol{A}$ имеющет три строки и два столбца, тогда результатом дифференцирования транспонированной матрицы $\boldsymbol{A}^T$ по исходной будет матрица $\boldsymbol{E}_{3,2}$, которая выглядит следующим образом:

$$\boldsymbol{E}_{3,2} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Итерационная процедура состоит из нескольких этапов. На первом этапе оцениваются общности переменных и, соответственно, − характерности $\boldsymbol{D}^2$. Итерационная процедура (2.83) дает оптимальное начальное приближение $\boldsymbol{A}_0$ для получения максимально правдоподобных оценок матрицы факторных нагрузок. Затем для получения значений идемпотентной матрицы (2.86) на первой итерации определяются значения матриц факторных нагрузок и корреляций:

$$\boldsymbol{A}_1 = \boldsymbol{A}_0 - \boldsymbol{R}\boldsymbol{P}_0^{-1}\boldsymbol{A}_0;$$
$$\boldsymbol{P}_1 = \boldsymbol{A}_1\boldsymbol{A}_1^T + \boldsymbol{D}^2.$$

После этого основная итерационная процедура имеет вид:

$$\boldsymbol{I_n}\otimes\boldsymbol{A}_{i+1} = \boldsymbol{I_n}\otimes\boldsymbol{A}_i - \left(\boldsymbol{I_n}\otimes\boldsymbol{H}_i\left(\boldsymbol{A}_{i-1} - \boldsymbol{R}\boldsymbol{P}_{i-1}^{-1}\boldsymbol{A}_{i-1}\right)\right)\cdot$$
$$\cdot\,(\boldsymbol{I}_{nm} + \boldsymbol{P}_i^{-1}\boldsymbol{R}\otimes\boldsymbol{A}_i^T\boldsymbol{P}_i^{-1}\boldsymbol{A}_i + \boldsymbol{E}_{nm}\left(\boldsymbol{A}_i^T\boldsymbol{P}_i^{-1}\boldsymbol{R}\otimes\boldsymbol{P}_i^{-1}\boldsymbol{A}_i\right) -$$
$$-\boldsymbol{P}_i^{-1}\boldsymbol{R}\otimes\boldsymbol{I}_m)^{-1}, \qquad (2.95)$$

где $\boldsymbol{H}_i = (\boldsymbol{A}_i - \boldsymbol{A}_{i-1})\left((\boldsymbol{A}_i - \boldsymbol{A}_{i-1})^T(\boldsymbol{A}_i - \boldsymbol{A}_{i-1})\right)^{-1}(\boldsymbol{A}_i - \boldsymbol{A}_{i-1})^T$,

$i = 1,2, \ldots$ и процедура выполнятся до тех пор, пока матрицы факторных нагрузок соседних итераций не станут отличаться на величину наперед заданной погрешности.

Выделить матрицу $\boldsymbol{A}_{i+1}$ из кронекеровского произведения несложно, но процедуру (2.95) можно упростить, если воспользоваться свойством умножения произведения Кронекера на вектор:

$$(\boldsymbol{A}\otimes\boldsymbol{B})vec(\boldsymbol{C}) = vec(\boldsymbol{A}\boldsymbol{C}\boldsymbol{B}^T).$$

Введем матрицу свертки $\boldsymbol{S}$, имеющей такую же размерность, как и матрица факторных нагрузок $\boldsymbol{A}$, т.е. $n \times m$. Обе матрицы имеют одинаковый ранг, равный числу факторов m. Матрица свертки – блочная: верхняя ее часть представляет собой диагональную матрицу, а нижняя часть – целиком состоит из нулей. К примеру, матрица факторных нагрузок имеет пять строк и три столбца, тогда матрица свертки имеет вид:

$$\boldsymbol{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Умножим (2.95) на $vec(\boldsymbol{S})$, это выражение свернется в вектор. Затем, преобразовав вектор в обычную матрицу с помощью операции $unvec$, окончательно получим

$$\boldsymbol{A}_{i+1} = \boldsymbol{A}_i - unvec\{\left(\boldsymbol{I_n}\otimes\boldsymbol{H}_i\left(\boldsymbol{A}_{i-1} - \boldsymbol{R}\boldsymbol{P}_{i-1}^{-1}\boldsymbol{A}_{i-1}\right)\right)\cdot$$
$$\cdot\,(\boldsymbol{I}_{nm} + \boldsymbol{P}_i^{-1}\boldsymbol{R}\otimes\boldsymbol{A}_i^T\boldsymbol{P}_i^{-1}\boldsymbol{A}_i + \boldsymbol{E}_{nm}\left(\boldsymbol{A}_i^T\boldsymbol{P}_i^{-1}\boldsymbol{R}\otimes\boldsymbol{P}_i^{-1}\boldsymbol{A}_i\right) -$$
$$-\boldsymbol{P}_i^{-1}\boldsymbol{R}\otimes\boldsymbol{I}_m)^{-1}vec(\boldsymbol{S})\}. \qquad (2.96)$$

Выражение (2.96) определяет сходящуюся итерационную процедуру, дающую максимально правдоподобные оценки коэффициентов матрицы факторных нагрузок. Однако выбирая различные начальные приближения для матрицы факторных нагрузок в начале итерационной процедуры, в результате получаем практически бесчисленное множество различных, но эквивалентных решений. Все они одинаково описывают корреляции исходных переменных, т.е. формальное перемножение матрицы факторных нагрузок на транспонированную для различных эквивалентных решений будет давать равные матрицы.

Единственное решение может быть найдено лишь при одном условии. Исходная для вычисления факторных нагрузок матрица корреляций является симметричной, таким образом, у нее $\frac{n(n+1)}{2}$ различных элемента. Если учесть, что на главной диагонали корреляционной матрицы стоят единицы, а в случае использования редуцированной корреляционной матрицы на главной диагонали стоят уже оцененные значения характерностей, то для определения $n \cdot m$ неизвестных факторных нагрузок имеется $\frac{n(n-1)}{2}$ исходных величин. Несложно увидеть, что условием единственности факторного решения, по алгебраическим соображениям, является условие равенства числа факторов половине числа переменных, уменьшенного на единицу:

$$m = \frac{n-1}{2}. \tag{2.97}$$

Если число переменных – четное, то условие (2.97) автоматически невыполнимо. Как правило, $m > \frac{n-1}{2}$, т.е. число неизвестных больше числа уравнений, что ведет, в случае совместности системы уравнений, к бесчисленному множеству решений. Однако и в случае, когда $m < \frac{n-1}{2}$, также возникают проблемы с однозначностью решения. Но даже тогда, когда выполняется условие (2.97), составить систему с соответствующим количеством уравнений, имеющих единственное решение, крайне сложно и для большого числа переменных – практически невозможно.

Эти рассуждения построены на алгебраическом подходе. Все же следует отметить, что поиск факторных нагрузок по матрице корреляций тождественен получению собственных векторов симметричной матрицы, а это − решение однородного уравнения, которое имеет бесчисленное множество решений. Все это и есть причина бесчисленности различных эквивалентных максимально правдоподобных факторных решений.

Этот факт – безусловно, относится к недостаткам метода, но большое множество решений позволяет выбрать из них такое, которое бы отвечало дополнительным требованиям. Наиболее распространенное одно из них – это предельная интерпретабельность факторного решения.

С наибольшей наглядностью фактор можно интерпретировать, когда матрица корреляций факторов с переменными, т.е. матрица факторных нагрузок, имеет «простой» вид, а именно, когда элементы в столбце принимают только либо большие, либо малые значения. Причем, поскольку переменные и факторы – нормированы, то «большие» значения – это величины, близкие к единице, а малые значения – к нулю. Разные факторы при этом должны «нагружать» разные переменные таким образом, чтобы, по возможности, «нули» и «единицы» для разных столбцов не совпадали. Такое требование ведет к ортогональности разных столбцов, т.е. скалярное произведение двух разных столбцов равно нулю. Следовательно, в итоге данное требование сводится к тому, чтобы произведение транспонированной матрицы факторных нагрузок на саму себя было диагональной матрицей:

$$\boldsymbol{A}^T \boldsymbol{A} = \boldsymbol{J}, \tag{2.98}$$

где $\boldsymbol{J}$ – некоторая диагональная матрица, она в отдельном случае может быть единичной, но поскольку фактор может быть тесно связан с несколькими переменными, и, соответственно, несколько факторных нагрузок могут быть близкими к единице, то матрица $\boldsymbol{J}$ в общем случае – не единичная.

Если настаивать на требовании, чтобы каждый фактор нагружал разные переменные, а факторное решение удовлетворяла бы принципам «простой структуры», то в среднем на каждый фактор должно приходиться $\frac{n}{m}$ переменных, таким образом, матрица $\boldsymbol{J}$ должна иметь вид:

$$\boldsymbol{J} = \frac{n}{m}\boldsymbol{I}.$$

Для того, чтобы выбрать из множества максимально правдоподобных решений такое, которое удовлетворяло бы ограничению (2.98), необходимо функцию правдоподобия (2.90) дополнить данным ограничением, т.е. составить функцию Лагранжа:

$$g = ln|\boldsymbol{P}| + tr(\boldsymbol{P}^{-1}\boldsymbol{R}) + tr\big(\boldsymbol{\Lambda}(\boldsymbol{A}^T\boldsymbol{A} - s\boldsymbol{I})\big),$$

где $\boldsymbol{\Lambda}$ − матрица множителей Лагранжа, размером $m \cdot m$,

$s = \frac{n}{m}$ − множитель, соответствующий принципу простой структуры,

остальные обозначения такие же, как и выше.

Теперь найдем производную функции Лагранжа по матрице факторных нагрузок $\boldsymbol{A}$, затем по матрице множителей Лагранжа $\boldsymbol{\Lambda}$. Эти производные после соответствующих преобразований имеют вид:

$$\frac{\partial g}{\partial \boldsymbol{A}} = 2vec\{\boldsymbol{P}^{-1}\boldsymbol{A} - \boldsymbol{P}^{-1}\boldsymbol{R}\boldsymbol{P}^{-1}\boldsymbol{A} + \boldsymbol{A}\boldsymbol{\Lambda}\};$$

$$\frac{\partial g}{\partial \boldsymbol{\Lambda}} = vec\{\boldsymbol{A}^T\boldsymbol{A} - s\boldsymbol{I}\}.$$

Эти производные приравняем нулю, и поскольку равенство нулю матрицы, преобразованной в вектор, и самой матрицы имеет одинаковый математический смысл, данные выражения дают систему матричных уравнений:

$$\begin{cases} \boldsymbol{P}^{-1}\boldsymbol{A} - \boldsymbol{P}^{-1}\boldsymbol{R}\boldsymbol{P}^{-1}\boldsymbol{A} + \boldsymbol{A}\boldsymbol{\Lambda} = \boldsymbol{0}; \\ \boldsymbol{A}^T\boldsymbol{A} - s\boldsymbol{I} = \boldsymbol{0}. \end{cases} \tag{2.99}$$

Преобразуем первое уравнение системы (2.99) к выражению, на основании которого можно будет выразить матрицу факторных нагрузок, для этого умножим это уравнение на матрицу истинных корреляций $\boldsymbol{P}$ слева, получим:

$$\boldsymbol{A} - \boldsymbol{R}\boldsymbol{P}^{-1}\boldsymbol{A} + \boldsymbol{P}\boldsymbol{A}\boldsymbol{\Lambda} = \boldsymbol{0}. \tag{2.100}$$

Умножим последнее уравнение (2.100) на транспонированную матрицу факторных нагрузок слева, и воспользовавшись вторым уравнением системы (2.99), а именно тем, что $\boldsymbol{A}^T\boldsymbol{A} = s\boldsymbol{I}$, получим значение матрицы коэффициентов Лагранжа:

$$\boldsymbol{\Lambda} = (\boldsymbol{A}^T\boldsymbol{P}\boldsymbol{A})^{-1}(\boldsymbol{A}^T\boldsymbol{R}\boldsymbol{P}^{-1}\boldsymbol{A} - s\boldsymbol{I}). \tag{2.101}$$

Далее, подставив матрицу $\boldsymbol{\Lambda}$ в (2.100), и выразив, после очевидных упрощений, матрицу факторных нагрузок, окончательно получим выражение, дающее основание для итерационной процедуры поиска факторного решения:

$$\boldsymbol{A} = \boldsymbol{R}\boldsymbol{P}^{-1}\boldsymbol{A} - \boldsymbol{P}\boldsymbol{A}(\boldsymbol{A}^T\boldsymbol{P}\boldsymbol{A})^{-1}(\boldsymbol{A}^T\boldsymbol{R}\boldsymbol{P}^{-1}\boldsymbol{A} - s\boldsymbol{I}). \tag{2.102}$$

Для начала итерационной процедуры необходима матрица начального приближения факторных нагрузок $\boldsymbol{A}_0$, которая может быть получена в соответствии с процедурой (2.83). После этого сама итерационная процедура определяется двумя выражениями:

$$\begin{cases} \boldsymbol{P}_i = \boldsymbol{A}_i\boldsymbol{A}_i^T + \boldsymbol{D}_i^2; \\ \boldsymbol{A}_{i+1} = \boldsymbol{R}\boldsymbol{P}_i^{-1}\boldsymbol{A}_i - \boldsymbol{P}_i\boldsymbol{A}_i\left(\boldsymbol{A}_i^T\boldsymbol{P}_i\boldsymbol{A}_i\right)^{-1}\left(\boldsymbol{A}_i^T\boldsymbol{R}\boldsymbol{P}_i^{-1}\boldsymbol{A}_i - s\boldsymbol{I}\right). \end{cases} (i = 0,1,\dots) \tag{2.103}$$

Причем для получения матрицы $\boldsymbol{P}_i$ нет необходимости вычислять значения характерностей $\boldsymbol{D}_i^2$, поскольку в результате сложения на диагонали матрицы теоретических значений корреляции переменных должны стоять единицы, поэтому после получения произведения $\boldsymbol{A}_i\boldsymbol{A}_i^T$ диагональ этой матрицы заменяется единицами.

Практическое использование процедуры (2.103) показало, что процедура сходится достаточно быстро и хорошо, однако для значений факторных нагрузок для четных и нечетных итераций. Т.е. разность между матрицами соседних итераций образует матрицу, элементы которой от итерации к итерации не меняют своих численных значений, после нескольких первоначальных вычислений. Чтобы вывести ход итерации из замкнутого цикла и обеспечить сходимость к одному решению, и вместе с тем удовлетворение условия (2.98), процедура (2.103) должна быть дополнена еще одним выражением. Его содержание состоит в том, то в качестве входной величины для следующей итерации берется матрица, лежащая «посередине» между предыдущей матрицей и матрицей, получаемой в ходе уточнения на данной итерации. Таким образом, итерационная процедура, дающее однозначное решение для данного начального значения матрицы факторных нагрузок, имеет следующий вид:

$$\begin{cases} \boldsymbol{P}_i = \boldsymbol{A}_i\boldsymbol{A}_i^T + \boldsymbol{D}_i^2; \\ \boldsymbol{A}_{i+\frac{1}{2}} = \boldsymbol{R}\boldsymbol{P}_i^{-1}\boldsymbol{A}_i - \boldsymbol{P}_i\boldsymbol{A}_i\left(\boldsymbol{A}_i^T\boldsymbol{P}_i\boldsymbol{A}_i\right)^{-1}\left(\boldsymbol{A}_i^T\boldsymbol{R}\boldsymbol{P}_i^{-1}\boldsymbol{A}_i - s\boldsymbol{I}\right); \\ \boldsymbol{A}_{i+1} = \frac{1}{2}\left(\boldsymbol{A}_{i+\frac{1}{2}} + \boldsymbol{A}_i\right);\ (i = 0,1,\dots). \end{cases} \tag{2.104}$$

Здесь индекс $\frac{1}{2}$ означает промежуточные вычисления. Практическое использование вычислительной процедуры (2.104) показало устойчивую сходимость к конечному результату, причем использование в качестве начального приближения различных «хороших» приближений, задаваемых (2.103), показало идентичность факторных решений. Единственные отличия состояли в том, что факторы менялись местами: первый становился вторым, третий – первым, а второй – третьим, или их следование осуществлялось в другой комбинации. Причем значения факторных нагрузок существенно не отличались друг от друга. Теоретическое исследование итерационной процедуры говорит о том, что требование выполнения условий простой структуры (2.78), приводит к формированию достаточных условий сходимости итерационной процедуры при любом начальном приближении.

Несложно увидеть, что решение, получаемое в ходе процедуры (2.104), удовлетворяет условию (2.98). А именно, пусть в ходе осуществления итераций получено факторное решение, т.е. матрицы факторных нагрузок соседних итераций не отличаются друг от друга на величину наперед заданной погрешности, и индекс номера итерации можно опустить. Тогда, умножив второе уравнение системы вычислительных тождеств (2.104) или выражение (2.102) на $\boldsymbol{A}^T$ слева, получим:

$$\begin{aligned}\boldsymbol{A}^T\boldsymbol{A} &= \boldsymbol{A}^T\boldsymbol{R}\boldsymbol{P}^{-1}\boldsymbol{A} - \boldsymbol{A}^T\boldsymbol{P}\boldsymbol{A}(\boldsymbol{A}^T\boldsymbol{P}\boldsymbol{A})^{-1}(\boldsymbol{A}^T\boldsymbol{R}\boldsymbol{P}^{-1}\boldsymbol{A} - s\boldsymbol{I}) = \\ &= \boldsymbol{A}^T\boldsymbol{R}\boldsymbol{P}^{-1}\boldsymbol{A} - \boldsymbol{I}(\boldsymbol{A}^T\boldsymbol{R}\boldsymbol{P}^{-1}\boldsymbol{A} - s\boldsymbol{I}) = \\ &= \boldsymbol{A}^T\boldsymbol{R}\boldsymbol{P}^{-1}\boldsymbol{A} - \boldsymbol{A}^T\boldsymbol{R}\boldsymbol{P}^{-1}\boldsymbol{A} + s\boldsymbol{I} = s\boldsymbol{I}.\end{aligned}$$

В разделе 2.4 будет показано, что факторное решение (2.104) без дополнительных «вращений», которые в факторном анализе традиционно применяются для облегчения трактовки или улучшения интерпретабельности решения, удовлетворяет требованиям простой структуры, и оно четко может быть истолковано. А к ним относятся и такое: каждая переменная может быть отнесена к одному фактору, и число переменных, которые могут быть отнесены к каждому фактору, распределяется равномерно.

2.4. Оценка значений многомерных факторов модели экономической системы.

В анализе экономических процессов важной стороной является изучение поведения факторов во времени. Знание того, как ведут себя скрытые от непосредственного измерения причины взаимодействий между отдельными сторонами явления, делает завершенным ход исследования, позволяя разобраться практически со всеми нюансами поведения экономического объекта и делая прогноз его поведения наиболее достоверным.

Традиционная интерпретация факторного анализа предполагает, что в качестве исходных данных берется не статистика определенных характеристик одного объекта или процесса для различных моментов времени, а наблюдения этих характеристик, произведенные по различным объектам или индивидуумам. И определение значений факторов осуществляется для соответствующих точек исходной информации. Однако исследование экономического процесса предполагает, что исходные данные собираются как фиксация некоторых параметров процесса в определенные моменты времени, либо они получаются как итоговые результаты за определенный промежуток времени (например, годовой валовый доход).

Исследование значений факторов во временном разрезе в традиционном факторном анализе не проводилось. Тем не менее, такое исследование возможно при определенных предварительных предположениях. Последовательность наблюдений, являющаяся по своей природе выборкой, не должна подпадать под определение «временного ряда», т.е. удовлетворять двум основным условиям:

1) наблюдения в различные моменты времени рассматриваются как случайные величины, являющиеся взаимно независимыми и, в частности, значение, получаемое в некоторый момент времени $t_k, (k = 1,2, \dots n)$, не должно зависеть от того, какие значения были зарегистрированы до этого момента времени;

2) наблюдения, произведенные в разные моменты времени, образуют стационарную последовательность, т.е. закон распределения вероятности $k-$го члена выборки (многомерной случайной величины $\boldsymbol{x}_k = \boldsymbol{x}(t_k)$) должен оставаться одним и тем же при изменении его номера k, в частности, от t_k не должны зависеть основные числовые характеристики случайной величины $\boldsymbol{x}_k$ – ее математическое ожидание и дисперсия.

Если эти два требования не выполняются и исходные данные представляют собой временной ряд, то традиционное статистическое исследование такой последовательности производится с помощью специального раздела математической статистики – анализа временных рядов. Впрочем, аппарат факторного анализа может быть использован для решения задач изучения временных рядов, и на это будет указано в третьем разделе. Однако сейчас сделаем предположение о том, что значения показателей экономического процесса в различные моменты времени представляют собой независимые случайные величины, и их совокупность образует стационарную последовательность, т.е. в статистическом плане исходная информация представляет собой некоторую выборку из генеральной совокупности, вероятностные параметры которой не меняются с течением времени.

Различные подходы к оценке значений факторов исследовал Харман [45]. С того момента достижения в этой области не подвергались никакой ревизии, ни в каком-либо направлении не развивались и сохранили все недостатки, на которые укажем прежде, чем будет предложено кардинальное решение этого вопроса.

Набор из n переменных можно проанализировать либо в терминах только общих факторов, тогда на главной диагонали матрицы корреляций $\boldsymbol{R}$ стоят единицы, либо в терминах совокупности общих и характерных факторов, и тогда на главной диагонали $\boldsymbol{R}$ стоят значения общностей. Первый подход приводит к модели компонентного анализа (модели главных компонент): матрица $\boldsymbol{R}$ – матрица полного ранга, факторное решение дается в терминах n общих факторов и имеет вид

$$\boldsymbol{Z} = \boldsymbol{F}\boldsymbol{A}^T, \tag{2.105}$$

где $\boldsymbol{Z}$ − матрица центрированных и нормированных значений исходных переменных;

$\boldsymbol{F}$ − матрица значений факторов;

$\boldsymbol{A}^T$ − транспонированная матрица факторных нагрузок.

Поскольку $\boldsymbol{A}$ – квадратная невырожденная матрица, то она имеет обратную матрицу. Поэтому искомые значения факторов определяются просто:

$$\boldsymbol{F} = \boldsymbol{Z}\boldsymbol{A}^{-T}. \tag{2.106}$$

Это решение является точным и однозначным, и не связано ни с какими «оценками».

При использовании компонентного анализа к факторной модели матрица факторных нагрузок $\boldsymbol{A}$ является неквадратной, имеющей порядок равный $n \cdot m$. И в этом случае несложно получить значения факторов, умножив (2.105) на $\boldsymbol{A}$ справа:

$$\boldsymbol{Z}\boldsymbol{A} = \boldsymbol{F}\boldsymbol{A}^T\boldsymbol{A}, \tag{2.107}$$

и решив уравнение (2.107) относительно $\boldsymbol{F}$, найдем значения факторов:

$$\boldsymbol{F} = \boldsymbol{Z}\boldsymbol{A}(\boldsymbol{A}^T\boldsymbol{A})^{-\mathbf{1}}. \tag{2.108}$$

Если матрица $\boldsymbol{A}$ является матрицей собственных значений $\boldsymbol{R}$, то $\boldsymbol{A}^T\boldsymbol{A}$ - диагональная матрица собственных значений корреляционной матрицы $\boldsymbol{R}$, и тогда (2.108) преобразуется к виду:

$$\boldsymbol{F} = \boldsymbol{Z}\boldsymbol{A}\boldsymbol{L}^{-\mathbf{1}}, \tag{2.109}$$

где $\boldsymbol{L}$ – диагональная матрица собственных значений $\boldsymbol{R}$.

Величины значений факторов, полученные по (2.109), обладают свойством ортогональности, т.е. матрица $\boldsymbol{F}^T\boldsymbol{F}$ − диагональная:

$$\boldsymbol{F}^T\boldsymbol{F} = \boldsymbol{L}^{-\mathbf{1}}\boldsymbol{A}^T\boldsymbol{Z}^T\boldsymbol{Z}\boldsymbol{A}\boldsymbol{L}^{-\mathbf{1}} = \boldsymbol{L}^{-\mathbf{1}}\boldsymbol{A}^T\boldsymbol{R}\boldsymbol{A}\boldsymbol{L}^{-\mathbf{1}} = \boldsymbol{L}^{-\mathbf{1}}\boldsymbol{L}\boldsymbol{L}^{-\mathbf{1}} = \boldsymbol{L}^{-\mathbf{1}}. \tag{2.110}$$

Диагональный вид матрицы произведений значений факторов $\boldsymbol{F}^T\boldsymbol{F}$ означает, что процедура (2.110) дает ортогональное факторное решение. Однако значения факторов, как и значения преобразованных исходных переменных, должны быть нормированными, т.е. $\boldsymbol{F}^T\boldsymbol{F}$ должна быть равна единичной матрице:

$$\boldsymbol{F}^T\boldsymbol{F} = \boldsymbol{I}, \tag{2.111}$$

а матрица $\boldsymbol{L}^{-\mathbf{1}}$ в (2.110) – обратная матрица собственных значений − не может быть единичной ни для каких исходных данных.

Если, однако, наряду с общими факторами модель включает в себя и характерные факторы, то решение уже нельзя получить столь просто. В этом случае число факторов и общих, и характерных превышает число переменных, и для матрицы факторной структуры не существует обратной. В этом случае традиционным подходом является построение «наилучшего» в смысле наименьших квадратов описания, что подразумевает использование регрессионных методов. Но все же такой подход имеет один существенный недостаток: оценивание с помощью множественной регрессии значений факторов приводит к их коррелированности [48]. С этим утверждением соглашаются вслед и Харман [45, с. 368], и Иберла [41, с. 244]. Тем не менее, можно показать, что полученные с помощью многомерного регрессионного анализа оценки (здесь-таки получаются именно оценки, а не однозначно определенные по формальной вычислительной процедуре значения) факторов при некоторых условиях являются ортогональными.

В указанных литературных источниках получение значений оценок проведено в такой же степени путано, как и приблизительно. Главный недостаток состоит в следующем: в изложении материала отсутствует ссылка на то, оттуда получено главное уравнение, на котором, собственно, и построен этот метод. Не претендуя на новизну, попытаемся раскрыть суть метода, освятив «темные пятна» указанных источников.

Уравнение регрессии, связывающее факторы и переменные можно запасать в виде

$$\boldsymbol{f}_t = \boldsymbol{z}_t \boldsymbol{B}, \tag{2.112}$$

где $\boldsymbol{f}_t$ – вектор значений факторов в момент времени *t*, размером $1 \cdot m$;

$\boldsymbol{z}_t$ – вектор центрированных и нормированных значений переменных в момент времени *t*, размером $1 \cdot n$;

$\boldsymbol{B}$ – матрица неизвестных коэффициентов регрессии, размером $n \cdot m$.

В общем случае линейное уравнение регрессии должно включать и случайную составляющую. Ее наличие не предусматривали «классики», вслед за ними также ею пренебрежем, поскольку природа случайности значений факторов до сих пор не исследовалась.

Умножим (2.112) на $\boldsymbol{z}^T$ слева и возьмем математическое ожидание от всего уравнения:

$$E\{\boldsymbol{z}^T \boldsymbol{f}_t\} = E\{\boldsymbol{z}^T \boldsymbol{z}_t \boldsymbol{B}\}.$$

Матрица коэффициентов корреляции является неизменной во времени матрицей, поэтому она, константа, может быть вынесена за знак математического ожидания:

$$E\{\boldsymbol{z}^T \boldsymbol{f}_t\} = E\{\boldsymbol{z}^T \boldsymbol{z}_t\} \boldsymbol{B}. \tag{2.113}$$

Левая часть (2.113) – математическое ожидание произведения переменных на факторы,– представляет собой корреляции между ними, а этими корреляциями является матрица факторных нагрузок $\boldsymbol{A}$. Математическое ожидание в правой части (2.113) отражает корреляции между переменными, т.е. это математическое ожидание равно матрице корреляций $\boldsymbol{R}$. Таким образом, выражение (2.113) может быть преобразовано к виду:

$$\boldsymbol{A} = \boldsymbol{R}\boldsymbol{B}, \tag{2.114}$$

именно вывод этого уравнения отсутствовал у Хармана.

Теперь, поскольку матрица выборочных корреляций с единицами на главной диагонали – полного ранга, то для нее существует обратная, поэтому из (2.114) могут быть найдены оценки коэффициентов регрессии

$$\boldsymbol{B} = \boldsymbol{R}^{-1} \boldsymbol{A}\ . \tag{2.115}$$

Оценки значений факторов получаются после подставки коэффициентов регрессии (2.115) в уравнение (2.112)

$$\boldsymbol{f}_t = \boldsymbol{z}_t \boldsymbol{R}^{-1} \boldsymbol{A}. \tag{2.116}$$

Последнее уравнение связывает значения факторов в момент времени t со значениями переменных в этот же момент. Для всех моментов времени выражение (2.116) преобразуется к виду:

$$\boldsymbol{F} = \boldsymbol{Z} \boldsymbol{R}^{-1} \boldsymbol{A}, \tag{2.117}$$

где $\boldsymbol{F}$ – матрица значений факторов для всех моментов времени;

$\boldsymbol{Z}$ – матрица центрированных и нормированных значений переменных для всех моментов времени.

Попытаемся проверить ортогональность данной оценки значений факторов, для этого транспонируем матрицу (2.117) и затем умножим ее слева на исходную матрицу:

$$\boldsymbol{F}^T \boldsymbol{F} = \boldsymbol{A}^T \boldsymbol{R}^{-1} \boldsymbol{Z}^T \boldsymbol{Z} \boldsymbol{R}^{-1} \boldsymbol{A}. \tag{2.118}$$

Умножение на множитель $\frac{1}{N-1}$ выражения (2.118) делает произведение $\boldsymbol{Z}^T \boldsymbol{Z}$ матрицей выборочных корреляций $\boldsymbol{R}$, поскольку произведение обратной матрицы на исходную есть матрица единичная, то (2.118) преобразуется к виду:

$$\boldsymbol{F}^T\boldsymbol{F} = \boldsymbol{A}^T\boldsymbol{R}^{-1}\boldsymbol{A}. \qquad (2.119)$$

Если для оценки матрицы факторных нагрузок $\boldsymbol{A}$ используется метод главных факторов – наиболее распространенный до сего момента времени метод, − то $\boldsymbol{A}$ состоит из собственных векторов матрицы выборочных корреляций. И поскольку обратная матрица в (2.119) имеет те же собственные векторы, что и исходная, то (2.119) преобразуется к виду:

$$\boldsymbol{F}^T\boldsymbol{F} = \boldsymbol{L}^{-1}, \qquad (2.120)$$

где $\boldsymbol{L}$ – матрица собственных значений выборочной корреляционной матрицы.

Таким образом, этот метод дает такой же результат, как и компонентный метод, и корреляции между значениями факторов, вычисленных с использованием метода главных факторов, аналогичны (2.120); и, соответственно, замечание о коррелированности факторов при применении методов многомерного регрессионного анализа – не совсем корректно. Тем не менее, применение других методов оценки матрицы факторных нагрузок, кроме метода главных факторов, когда столбцы факторных нагрузок не есть собственные векторы корреляционной матрицы, ведет к нарушению ортогональности факторной структуры. Следовательно, не само использование многомерной регрессии ведет к коррелированности факторов, а к этому приводят результаты использования соответствующего метода оценки факторных нагрузок.

Харман [45] и Иберла [41] подробно рассмотрели вопросы проверки точности оценки значений факторов. Этот момент является достаточно важным, поэтому рассмотрим теорию проблемы в изложении «классиков».

Прежде всего, они исходили из того, что матрица факторных нагрузок $\boldsymbol{A}$ является ортогональной т.е. ее столбцы – собственные векторы матрицы корреляций. Как теперь известно, это ведет к ортогональности значений факторов.

Если известна оценка матрицы коэффициентов регрессии, то подставляя в (2.112) значение переменных для некоторого момента времени, получим оценку значения фактора для данного момента. Ставится задача: насколько точна такая оценка? Мерой качества такой оценки могут служить коэффициент множественной корреляции и коэффициент множественной детерминации [41, с. 251].

При парной линейной связи между лишь двумя переменными коэффициент корреляции характеризует тесноту связи между двумя переменными. Причем чем ближе его значение по абсолютной величине к единице, тем теснее связь между переменными. Соответствующим показателем во множественной регрессии является коэффициент множественной корреляции, или совокупный коэффициент корреляции. Он показывает тесноту связи зависимой переменной, или целевой функции, от переменных, входящих в правую часть уравнения регрессии.

Проверка точности оценки значения фактора осуществляется на основании корреляции между истинным, но неизвестным значением фактора и его оценкой, но поскольку оценка вычисляется по уравнению регрессии, то критерий качества оценки получается на основании коэффициента множественной корреляции между истинным значением фактора и значениями переменных, на основании которых получена оценка.

Для фактора F_j коэффициент множественной корреляции имеет вид [45, с. 373]

$$R_j = \sqrt{a_{j,1}b_{j,1} + a_{j,2}b_{j,2} + \cdots + a_{j,n}b_{j,n}}, \qquad (2.121)$$

где $a_{j,i}$ – элемент матрицы факторных нагрузок $\boldsymbol{A}$,

$b_{j,i}$ − элемент матрицы коэффициентов множественной регрессии $\boldsymbol{B}$.

Квадрат коэффициента множественной корреляции называется коэффициентом множественной детерминации. В соответствии с этим определением для фактора F_j коэффициент множественной детерминации имеет вид

$$R_j^2 = a_{j,1}b_{j,1} + a_{j,2}b_{j,2} + \cdots + a_{j,n}b_{j,n}. \qquad (2.122)$$

Если учесть (2.115), т.е. принять во внимание формулу оценки коэффициентов регрессии, то (2.122) можно представить в следующем виде:

$$R_j^2 = \boldsymbol{a}_j^T\boldsymbol{R}^{-1}\boldsymbol{a}_j, \qquad (2.123)$$

где $\boldsymbol{a}_j$ − j-й столбец матрицы факторных нагрузок.

Коэффициент множественной детерминации состоит из нескольких слагаемых, каждое из которых вносит свой вклад в R_j^2. Каждое из этих слагаемых в свою очередь является произведением коэффициента регрессии и коэффициентом корреляции между j-м фактором и i-й переменной. Если какое-либо слагаемое в (2.122) значительно по величине, то соответствующая переменная вносит большой вклад в определение фактора F_j. По коэффициенту множественной детерминации можно определить, какая часть дисперсии фактора обусловлена переменными, а какая часть обусловлена неучтенными переменными.

Эти же соображения можно отнести и к влиянию фактора на переменные: большое значение коэффициента детерминации говорит о значимости данного фактора в формировании поведения переменных. И contra versa – небольшое значение этого коэффициента свидетельствует о незначимости данного фактора, а в решении проблемы точности оценки значений факторов – о недостаточно хорошем качестве оценки значения фактора.

Коэффициент множественной детерминации, как и коэффициент множественной корреляции, принимает значение не большее, чем единица. Поэтому близость к единице будет свидетельствовать о значимости данного фактора и достаточно хорошей оценке его значения. К тому же коэффициент множественной корреляции позволяет проверить гипотезу о значимости этого коэффициента с заданным уровнем значимости (вероятностью ошибки), что позволяет объективировать процесс проверки меры близости оценки фактора к его истинному значению.

Проверка осуществляется с помощью статистической таблицы «Критические значения коэффициента множественной корреляции». Для данного уровня значимости, заданного числа переменных и наблюдений вычисленный коэффициент множественной корреляции сравнивается с его табличным критическим значением. Если вычисленный коэффициент больше табличного, то его истинное значение отлично от нуля.

Табличное значение задается на основании распределения Фишера по формуле:

$$\rho_{k_1,k_2}^{(\alpha)} = \sqrt{\frac{1}{1+\frac{1}{k_1 \cdot F(k_1,k_2)}}}, \tag{2.124}$$

где $\rho_{k_1,k_2}^{(\alpha)}$ – критическое значение множественного коэффициента корреляции с α − уровнем значимости и k_1, k_2 степенями свободы;

k_1– число переменных;

$k_2 = N - k_1 - 1$, где N − число наблюдений (объем выборки);

$F(k_1, k_2)$ – распределение Фишера с k_1, k_2 степенями свободы и с тем же, что и у критического значения коэффициента корреляции, уровнем значимости.

Если гипотеза о значимости коэффициента множественной корреляции отклоняется, то соответствующий фактор интерпретировать в принципе нельзя, поскольку и вклад переменных в информацию о нем незначителен. Такое положение несложно исправить, уменьшив число факторов на единицу. Правда, проверка на значимость с помощью таблицы критических значений коэффициента множественной корреляции не совсем корректна, так она составлена в предположении независимости переменных, с которыми имеется корреляция некоторого признака [41, с. 252]. Но коррелированность «независимых» переменных ведет к завышению оценки коэффициента множественной корреляции. Это следует из известного в теории вероятностей утверждения о вероятности суммы произвольных, в том числе и зависимых событий, согласно которому вероятность суммы произвольных событий равна сумме вероятностей этих событий, уменьшенной на вероятность произведения этих событий:

$$P(A + B) = P(A) + P(B) - P(AB), \tag{2.125}$$

где $P(\cdot)$ – вероятность события;

$A + B$ − сумма зависимых событий;

AB − произведение событий, т.е. вероятность того, что они одновременно произойдут, а для независимых событий эта вероятность равна произведению вероятностей этих событий:

$$P(AB) = P(A) \cdot P(B), \tag{2.126}$$

но события зависимы, и для них сумма вероятностей все одно уменьшается на вероятность их совместного наступления.

Случайная величина ξ тесно связана с вероятностью события состоящего в том, что ξ примет то или иное значение. И поскольку оценка той или иной случайной величины несет в себе информацию об истинном, но неизвестном ее значении, а информация тесно зависит от вероятности, то отрицательная часть (2.125) ведет к ее снижению. Поясним – почему. Левая часть (2.125) – вероятность суммы событий − представляет собой вероятность того, что фактор связан с несколькими переменными, которые, в общем-то, взаимозависимы. Т.е. вероятность, стоящая в левой части (2.125), связана со значением фактора. В правой части (2.125) – сумме вероятностей событий – каждое слагаемое несет вероятность того, что отдельная переменная формируется фактором, и этот момент отмечен величиной слагаемого. Но общая сумма вероятностей должна быть скорректирована на вероятность их произведения. Так как переменные в определенный момент t проявляются практически все, то такая коррекция должна иметь значительный вес. Поэтому табличные критические значения коэффициента множественной корреляции составлены для независимых переменных, в том числе и потому, что коррекция на вероятность их совместного наступления в практическом плане зависит от специфики самих переменных и в общем случае учтена быть не может.

Завышение оценки значимости фактора может вести лишь к одному результату: в факторном анализе будут появляться факторы, значимость которых подтверждается критерием, но в реальности значимыми они не являются, т.е. соответствующим оценкам значений фактора верить не приходится. Поэтому решение этой проблемы должно лежать в плоскости минимизации числа факторов, минимизации – не как оптимизационной задачи, а выбора минимально необходимого для адекватного анализа числа факторов.

Небольшой промежуточный вывод: оценка значений факторов, основанная на методе множественной регрессии и осуществляемая в соответствии с выражением (2.117), дает ортогональные величины факторов, но они не являются нормированными, т.е. сумма квадратов значений отдельного фактора не равно единице. Это существенный недостаток, поскольку в факторной модели нормированные значения переменных выражаются через значения факторов, и если последние – не нормированы, то и сами переменные нормированными быть не могут. С формальной точки зрения это противоречие преодолено быть не может, поэтому в последнее время появился подход к модели факторного анализа, позволяющий «зрительно» обойти эту проблему [49]. В факторную модель включают центрированные, но не нормированные величины, поэтому в самой модели присутствуют средние значения переменных, отклонения от которых задают факторы:

$$\boldsymbol{z}_t = \bar{\boldsymbol{z}} + \boldsymbol{f}_t\boldsymbol{A}^T + \boldsymbol{u}_t, \tag{2.127}$$

где $\boldsymbol{z}_t$ − в этом контексте, − вектор центрированных переменных в момент времени t, т.е. у значений переменных вычтены средние значения;

$\bar{\boldsymbol{z}}$ – вектор средних значений переменных;

$\boldsymbol{f}_t$ − вектор значений факторов в момент времени t, в этом контексте значения факторов считаются центрированными, но не нормированными;

$\boldsymbol{A}$ − матрица факторных нагрузок;

$\boldsymbol{u}_t$− вектор случайных отклонений в момент времени t, дисперсии которых составляют характерности для переменных.

Идеальным в факторном анализе выглядит ситуация, когда модель соотношения перемененных и факторов дается в натуральных единицах, тогда проблема с центрированием и нормированием выглядела бы надуманной. Но есть один вопрос, на который очень сложно ответить: в каких единицах должен измеряться фактор так, чтобы через него выражались переменные разной размерности? К тому же одна переменная выражается через совокупность факторов. Эта практически не разрешимая проблема как раз и снимается с помощью центрирования и нормирования. Причем одно лишь центрирование проблемы не снимает, поскольку, если из некоторой величины вычесть среднее значение, имеющее ту же единицу измерения, то от этого размерность центрированной переменной ни исчезнет. От размерности освобождаются лишь при делении центрированной величина на среднеквадратическое отклонение.

Поэтому модель (2.127) неправомочна. Значение фактора должно быть безразмерным и также центрированным, как центрированы исходные переменные. Причем вернуться к натуральному измерению переменных не составляет труда, если известны их средние величины и среднеквадратические отклонения, но, опять же, значения факторов должны быть и центрированными (имеющими нулевое математическое ожидание) и нормированными (деленными на собственные среднеквадратические отклонения).

Поставим задачу получить оценки значений факторов для всех моментов времени, математические ожидания которых были бы равны нулю, а дисперсии – единице. Т.е. получить так называемые «стандартные» значения факторов.

Напомним основную модель факторного анализа для всех моментов:

$$\boldsymbol{Z} = \boldsymbol{F}\boldsymbol{A}^T + \boldsymbol{U}\boldsymbol{D}, \tag{2.128}$$

где $\boldsymbol{Z}$ – матрица размером $N \cdot n$ наблюдений переменных, ее значения центрированы и нормированы;

$\boldsymbol{F}$ – матрица размером $N \cdot m$ значений факторов, именно ее значения предстоит оценить;

$\boldsymbol{A}$ – матрица размером $n \cdot m$ факторных нагрузок, будем предполагать, что оценка ее элементов уже получена с помощью метода максимального правдоподобия, изложенного в п.2.2;

$\boldsymbol{U}$ − матрица размером $N \cdot n$ значений характерных факторов.

$\boldsymbol{D}$ – матрица размером $n \cdot n$ нагрузок при характерных факторах

Выше рассматривалась модель (2.39), которая ничем не отличается от (2.128), однако внесем в нее некоторые коррективы. Поскольку поставлена задача получения оценок значения лишь общих факторов, то второе слагаемое – влияние на поведение переменной характерных факторов, − можно рассматривать как некоторые случайные отклонения от основной зависимости, задаваемой общими факторами. В этой связи перепишем модель (2.128) в виде:

$$\boldsymbol{Z} = \boldsymbol{F}\boldsymbol{A}^T + \boldsymbol{V}, \tag{2.129}$$

где $\boldsymbol{V} = \boldsymbol{U}\boldsymbol{D}$ – матрица случайных отклонений, математическое ожидание которой, как несложно увидеть, равно нулю, а дисперсия − $\boldsymbol{D}^2$, т.е. значения характерностей.

Очевидным представляется стремление получить такие оценки значений факторов, которые минимизировали бы ошибки в описании переменных посредством факторной модели. При этом ошибками будем считать отклонения реальных значений переменных от теоретической зависимости, определяемой значениями факторов:

$$\widehat{\boldsymbol{V}} = \boldsymbol{Z} - \boldsymbol{F}\boldsymbol{A}^T, \tag{2.130}$$

где $\widehat{\boldsymbol{V}}$ − ошибки в объяснении переменных посредством факторов. Эти ошибки могут рассматриваться как оценки неизвестных значений характерностей в общей факторной модели.

В теории статистического оценивания широко известен метод наименьших квадратов оценивания параметров регрессии, который позволяет минимизировать сумму квадратов ошибок, т.е. отклонений реальных значений результирующего признака от «идеальной» регрессионной зависимости этого признака, определяемой некоторой группой переменных. В нашем случае выражение (2.129) можно считать уравнением регрессии, в котором значения факторов $\boldsymbol{F}$ играют роль неизвестных параметров, оценки которых подлежат определению, матрица факторных $\boldsymbol{A}$ нагрузок может быть интерпретирована как матрица значений независимых переменных, $\boldsymbol{Z}$ − матрица значений зависимой переменной.

Необходимым условием получения оценок является условие достаточности исходных данных, при котором числа уравнений должно хватить для определения неизвестных. В нашем случае размерность только одной матрицы $\boldsymbol{Z}$, элементы которой составляют исходные данные, равна $N \cdot n$, а число неизвестных – элементов матрицы значений факторов $\boldsymbol{F}$ составляет $N \cdot m$, где m − число факторов − значительно меньше числа переменных n. Даже если учесть, что часть информации «израсходована» на получение оценки факторных нагрузок, число которых равно $n \cdot m$, исходных данных все равно достаточно для получения оценок значений факторов.

Основным соотношением для решения вопроса о достаточности первоначальной информации является следующее: наблюдений за переменными больше числа значений факторов плюс число факторных нагрузок:

$$N \cdot n > N \cdot m + n \cdot m. \tag{2.131}$$

Несложно показать, что даже при максимально возможном числе факторов, составляющем половину от числа переменных, условие (2.131) выполняется. Пусть $m = \frac{n}{2}$, тогда $N \cdot n > N \cdot \frac{n}{2} + n \cdot \frac{n}{2}$; $N > N \cdot \frac{1}{2} + n \cdot c$; $\frac{N}{2} > \frac{n}{2}$; $N > n$; а условие, когда число наблюдений больше числа исследуемых величин, является основополагающим требованием статистических исследований, и выполняется всегда. При этом следует отметить, что в практических исследованиях число интерпретабельных факторов, которые и составляют основу факторной модели конкретного явления, не бывает больше четверти числа переменных.

Таким образом, в качестве целевой функции при получении оценок значений факторов можно выбрать сумму квадратов ошибок (2.130), для матрицы (не для вектора или матрицы-строки) она имеет вид:

$$tr\{(\boldsymbol{Z} - \boldsymbol{F}\boldsymbol{A}^T)^T(\boldsymbol{Z} - \boldsymbol{F}\boldsymbol{A}^T)\}, \tag{2.132}$$

где $tr -$ след матрицы, т.е. сумма диагональных элементов матрицы.

Однако необходимо найти не просто значения $\boldsymbol{F}$, которые минимизировали бы (2.132), но эти значения должны быть нормированными, т.е. обладать свойством (2.131): $\boldsymbol{F}^T\boldsymbol{F} = \boldsymbol{I}$. Для поиска минимума целевой функции при условии наложенных ограничений (2.131) составим функцию Лагранжа:

$$g = tr\{(\boldsymbol{Z} - \boldsymbol{F}\boldsymbol{A}^T)^T(\boldsymbol{Z} - \boldsymbol{F}\boldsymbol{A}^T)\} + tr\{\boldsymbol{\Lambda}(\boldsymbol{F}^T\boldsymbol{F} - \boldsymbol{I})\}, \tag{2.133}$$

где $\boldsymbol{\Lambda} -$ матрица множителей Лагранжа.

Найдем экстремумы функции (2.133), для этого вычислим частные производные g по $\boldsymbol{F}$ и по $\boldsymbol{\Lambda}$:

$$\begin{aligned}\frac{\partial g}{\partial \boldsymbol{F}} &= \frac{\partial}{\partial \boldsymbol{F}}(tr\{(\boldsymbol{Z} - \boldsymbol{F}\boldsymbol{A}^T)^T(\boldsymbol{Z} - \boldsymbol{F}\boldsymbol{A}^T)\} + tr\{\boldsymbol{\Lambda}(\boldsymbol{F}^T\boldsymbol{F} - \boldsymbol{I})\} + tr\{\boldsymbol{\Lambda}(\boldsymbol{F}^T\boldsymbol{F} - \boldsymbol{I})\}) = \\ &= \frac{\partial}{\partial \boldsymbol{F}}(tr\{\boldsymbol{Z}^T\boldsymbol{Z} - \boldsymbol{Z}^T\boldsymbol{F}\boldsymbol{A}^T - \boldsymbol{A}\boldsymbol{F}^T\boldsymbol{Z} + \boldsymbol{A}\boldsymbol{F}^T\boldsymbol{F}\boldsymbol{A}^T\} + tr\{\boldsymbol{\Lambda}\boldsymbol{F}^T\boldsymbol{F} - \boldsymbol{\Lambda}\}) = \\ &= \{\boldsymbol{E}(\boldsymbol{A}^T \otimes \boldsymbol{F}\boldsymbol{A}^T) + (\boldsymbol{F}\boldsymbol{A}^T \otimes \boldsymbol{A}^T) - \boldsymbol{E}(\boldsymbol{A}^T \otimes \boldsymbol{Z}) - (\boldsymbol{Z} \otimes \boldsymbol{A}^T) + \boldsymbol{E}(\boldsymbol{\Lambda}^T \otimes \boldsymbol{F}) + \\ &\quad + (\boldsymbol{F}\boldsymbol{\Lambda}^T \otimes \boldsymbol{I})\}vec\{\boldsymbol{I}\} = vec\{2\boldsymbol{F}\boldsymbol{A}^T\boldsymbol{A} - 2\boldsymbol{Z}\boldsymbol{A} + 2\boldsymbol{F}\boldsymbol{\Lambda}\};\end{aligned}$$

$$\frac{\partial g}{\partial \boldsymbol{\Lambda}} = vec\{\boldsymbol{F}^T\boldsymbol{F} - \boldsymbol{I}\}.$$

Приравняем нулю найденные частные производные, и после очевидных преобразований получим систему матричных уравнений:

$$\begin{cases} -\boldsymbol{Z}\boldsymbol{A} + \boldsymbol{F}(\boldsymbol{A}^T\boldsymbol{A} + \boldsymbol{\Lambda}) = \boldsymbol{0} \\ \boldsymbol{F}^T\boldsymbol{F} = \boldsymbol{I} \end{cases}. \tag{2.134}$$

Из первого уравнения системы (2.134) выразим матрицу $\boldsymbol{F}$, и найденное значение подставим во второе уравнение системы:

$$\boldsymbol{F} = \boldsymbol{Z}\boldsymbol{A}(\boldsymbol{A}^T\boldsymbol{A} + \boldsymbol{\Lambda})^{-1}, \tag{2.135}$$

$$\boldsymbol{F}^T = (\boldsymbol{A}^T\boldsymbol{A} + \boldsymbol{\Lambda})^{-1}\boldsymbol{A}^T\boldsymbol{Z}^T,$$

$$(\boldsymbol{A}^T\boldsymbol{A} + \boldsymbol{\Lambda})^{-1}\boldsymbol{A}^T\boldsymbol{Z}^T\boldsymbol{Z}\boldsymbol{A}(\boldsymbol{A}^T\boldsymbol{A} + \boldsymbol{\Lambda})^{-1} = \boldsymbol{I}.$$

Выразим матрицу, стоящую в середине левой части, умножив слева и справа на обратную матрицу, содержащую множители Лагранжа, получим

$$\boldsymbol{A}^T\boldsymbol{Z}^T\boldsymbol{Z}\boldsymbol{A} = (\boldsymbol{A}^T\boldsymbol{A} + \boldsymbol{\Lambda})^2.$$

Извлечение квадратного корня дает следующее выражение:

$$(\boldsymbol{A}^T\boldsymbol{Z}^T\boldsymbol{Z}\boldsymbol{A})^{\frac{1}{2}} = \boldsymbol{A}^T\boldsymbol{A} + \boldsymbol{\Lambda},$$

откуда получаем матрицу множителей Лагранжа

$$\boldsymbol{\Lambda} = (\boldsymbol{A}^T\boldsymbol{Z}^T\boldsymbol{Z}\boldsymbol{A})^{\frac{1}{2}} - \boldsymbol{A}^T\boldsymbol{A}. \tag{2.136}$$

Подставляя найденную матрицу (2.136) в выражение (2.135), получим оценку значений факторов

$$\widehat{\boldsymbol{F}} = \boldsymbol{Z}\boldsymbol{A}(\boldsymbol{A}^T\boldsymbol{Z}^T\boldsymbol{Z}\boldsymbol{A})^{-\frac{1}{2}}, \tag{2.137}$$

которая должна быть ортогональной, т.е. значения факторов – нормированы, и, соответственно, ее математическое ожидание должно быть равно нулю. Проверим эти свойства.

Сначала установим, чему равно математическое ожидание оценки. Предварительно отметим, что поскольку исходные переменные центрированы и нормированы, то математическое ожидание матрицы таких значений равно нулю:

$$\mathcal{M}\{\boldsymbol{Z}\} = 0,$$

где $\mathcal{M}\{\cdot\}$ – символ математического ожидания. При этом следует отметить, что при выполнении условий центрирования и нормирования, произведение матриц $\boldsymbol{Z}^T\boldsymbol{Z}$ равно корреляционной матрице исходных переменных $\boldsymbol{R}$. Это замечание должно быть сделано, для того, чтобы обойти проблему, связанную с определением математического ожидания обратной случайной величины. Если математическое ожидание некоторой случайной величины равно, скажем a, то есть $\mathcal{M}\{\xi\} = a$, то математическое ожидание обратного значения случайной величины в общем случае не равно обратному значению математического ожидания: $\mathcal{M}\left\{\frac{1}{\xi}\right\} \neq \frac{1}{a}$. И поскольку матрица $\boldsymbol{Z}$ находится под знаком обратной величины в выражении (2.137), то такая проблема имеет место. Решение этой проблемы лежит за пределами данного исследования, поэтому положим то, что с чего начинается факторный анализ – с вычисления корреляционной матрицы. И так как в (2.137) под знаком обратной величины находится не просто матрица $\boldsymbol{Z}$, а ее момент второго порядка $\boldsymbol{Z}^T\boldsymbol{Z}$, то с точностью до скалярных множителей далее будем полагать, что:

$$\boldsymbol{Z}^T\boldsymbol{Z} = \boldsymbol{R},$$

Что касается множителей, то имеется в виду множитель, равный $\frac{1}{N-1}$. На этот множитель можно умножить обе части (2.137), от чего это тождество не потеряет своей справедливости. Причем $\boldsymbol{R}$ − выборочная корреляционная матрица, которая по отношению к оператору взятия математического ожидания может считаться постоянной. Данное обстоятельство должно быть доказано, или доказательно пояснено, причем это возможно, но находится за пределами данного исследования. Поэтому уравнение (2.137) можно переписать в виде

$$\widehat{\boldsymbol{F}} = \boldsymbol{Z}\boldsymbol{A}(\boldsymbol{A}^T\boldsymbol{R}\boldsymbol{A})^{-\frac{1}{2}}. \tag{2.138}$$

А теперь можно найти математическое ожидание оценки факторных нагрузок, учитывая, что постоянную величину можно вынести за знак математического ожидания:

$$\mathcal{M}\{\widehat{\boldsymbol{F}}\} = \mathcal{M}\left\{\boldsymbol{Z}\boldsymbol{A}(\boldsymbol{A}^T\boldsymbol{R}\boldsymbol{A})^{-\frac{1}{2}}\right\} = \mathcal{M}\{\boldsymbol{Z}\}\boldsymbol{A}(\boldsymbol{A}^T\boldsymbol{R}\boldsymbol{A})^{-\frac{1}{2}} = \boldsymbol{0}\cdot\boldsymbol{A}(\boldsymbol{A}^T\boldsymbol{R}\boldsymbol{A})^{-\frac{1}{2}} = \boldsymbol{0}.$$

Таким образом, оно равно нулю, что и требовалось доказать. А так как истинные, но неизвестные значения центрированных и нормированных факторов равны нулю, то доказано утверждение, что данные оценки являются несмещенными.

Покажем теперь, что оценка (2.138) является ортонормированной, для этого найдем произведение $\widehat{\boldsymbol{F}}^T\widehat{\boldsymbol{F}}$:

$$\widehat{\boldsymbol{F}}^T\widehat{\boldsymbol{F}} = (\boldsymbol{A}^T\boldsymbol{Z}^T\boldsymbol{Z}\boldsymbol{A})^{-\frac{1}{2}}\boldsymbol{A}^T\boldsymbol{Z}^T\boldsymbol{Z}\boldsymbol{A}(\boldsymbol{A}^T\boldsymbol{Z}^T\boldsymbol{Z}\boldsymbol{A})^{-\frac{1}{2}} = (\boldsymbol{A}^T\boldsymbol{Z}^T\boldsymbol{Z}\boldsymbol{A})^{-\frac{1}{2}}(\boldsymbol{A}^T\boldsymbol{Z}^T\boldsymbol{Z}\boldsymbol{A})^{\frac{1}{2}} = \boldsymbol{I}.$$

Поскольку это произведение равно единичной матрице, то оценка значений факторов является ортогональной, и это – ставилось как основное требование при получении этой оценки.

Следовательно, оценка значений факторов (2.137) является несмещенной и ортонормированной.

Ортонормированные оценки являются наилучшими оценками среди всех прочих оценок, поскольку ортогональность предполагает взаимную независимость факторов, что отражает основное исходное положение факторного анализа − независимость причин, вызывающих взаимодействия между переменными, и если факторы коррелируют друг с другом, то это означает, что над ними стоит «надфактор», вызывающий такое явление и подлежащий выявлению. Нормированность делает возможным как получение на основании значений факторов таких же нормированных значений переменных, так и использование классической модели факторного анализа, а не различных модификаций типа (2.127), не имеющих физического смысла, строго говоря.

Выводы

Рассмотрены особенности использования модели эксплораторного факторного анализа для исследования экономических систем, обоснована необходимость получения оценок общностей для подстановки в редуцированную матрицу, с тем, чтобы точнее оценить параметры факторной модели экономической системы.

В настоящее время в качестве оценок общностей используются коэффициенты множественной корреляции. В разделе получены оценки общностей стохастических факторов на основе информационного подхода. С точки зрения статистических критериев близости оценки к истинным значениям предложенная оценка является лучше традиционно используемой.

В результате решения оптимизационной задачи поиска максимума функции правдоподобия при наличии ограничений, накладываемых на параметры факторной модели, найдены максимально правдоподобные оценки факторных нагрузок, отвечающие принципу «простой структуры».

Идеальным в эксплораторном факторном анализе выглядит ситуация, когда модель соотношения перемененных и факторов дается в натуральных единицах. Но тогда в каких единицах должен измеряться фактор так, чтобы через него выражались переменные разной размерности? К тому же одна переменная выражается через совокупность факторов. Эта практически не разрешимая проблема снимается с помощью центрирования и нормирования. Т.е. значения стохастических факторов должны удовлетворять условию ортонормированности. Оценка значений факторов, основанная на методе множественной регрессии, дает ортогональные величины факторов, но они не являются нормированными. Поставленная задача получить оценки значений факторов, математические ожидания которых были бы равны нулю, а дисперсии – единице для всех моментов времени, была решена и такие оценки значений факторов получены.

3. ЭКСПЛОРАТОРНЫЙ ФАКТОРНЫЙ АНАЛИЗ МНОГОМЕРНЫХ ВРЕМЕННЫХ РЯДОВ

3.1. Методы и модели анализа многомерных временных рядов.

Большинство статистических методов использует модели, в которых наблюдения в различные моменты времени предполагаются независимыми. Во многих приложениях зависимость между наблюдениями рассматривается как помеха. В планируемых экспериментах вводится рандомизация эксперимента, что позволяет обосновать анализ, проводимый без учета зависимости наблюдений. Однако разнообразные данные в коммерции, экономике, технике и естественных науках поступают в форме временных рядов, в которых наблюдения зависимы и характер этой зависимости интересен сам по себе. Временным рядом называется упорядоченная во времени последовательность наблюдений. Совокупность существующих методов анализа таких рядов зависимых наблюдений называется анализом временных рядов.

Основной чертой, выделяющей этот анализ рядов среди других видов статистического анализа, является существенность порядка, в котором производятся наблюдения. Если в некоторых задачах статистического анализа различных явлений значения переменных для разных моментов времени являются статистически независимыми, или это априорно предполагается, то во временных рядах они зависимы, и характер зависимости может определяться положением наблюдений в последовательности. Природа ряда и структура порождающего ряд процесса предопределяют порядок образования последовательности.

В экономической сфере имеется обширный круг явлений, которые интересно и важно изучать в их развитии и изменении во времени. С течением времени изменяются деловая активность, цены на различные материальные и финансовые активы, значения биржевых индексов, режим протекания того или иного производственного процесса, и тому подобное. Совокупность изменений какой-либо одной характеристики подобного рода в течение некоторого периода времени и представляет временной ряд. Это могут быть ежедневные величины индекса потребительских цен, или поквартальные данные о валовом национальном продукте.

Цели изучения временных рядов могут быть различными, и различные причины, по которым желательно исследовать временные ряды, могут определять выбор методов анализа явления. Различаются пять типов исследования временных рядов [50]:

1) на самом поверхностном уровне берут какой-либо ряд и строят простую математическую систему, которая *описывает* его поведение в сжатом виде;

2) проникая глубже, делают попытку *объяснить* поведение ряда с помощью других переменных и определить соотношение между наблюдениями и некоторыми структурными законами поведения, т.е. для объяснения наблюдений в качестве гипотезы строится модель;

3) результаты анализа, проведенного на первых двух этапах, можно использовать для *прогнозирования* поведения ряда. В случае 1), даже когда ничего не известно об основном механизме, генерирующем ряд, исходят из предположения, что в системе существует достаточная инерция, чтобы гарантировать в будущем такое же поведение, как и в прошлом. Случай 2) характеризуется большим проникновением в область рассматриваемых причинно-следственных связей и можно делать проекции в будущее более уверенно;

4) в случае 2) может потребоваться *контроль* системы с целью выработки сигналов, предупреждающих о грядущих неблагоприятных событиях, причем сигнал может вырабатываться непосредственно либо самой моделью, либо на основании анализа того, что может случиться, если изменятся некоторые параметры модели, и эти изменения произойдут в реальности;

5) для еще более глубокого анализа экономического явления может возникнуть необходимость рассмотреть *совместное развитие* во времени нескольких переменных

или, другими словами, исследуемая переменная может быть вектором наблюдений. И такой тип исследования представляет наибольший интерес.

Характеризуя временной ряд в целом, следует отметить, что в отличие от случайной выборки ряд следует рассматривать как смесь четырех компонент:

1) тренда или долгосрочного движения;
2) более или менее регулярных колебаний относительно тренда;
3) сезонной составляющей;
4) остатка или несистематического случайного эффекта.

Несомненно, ряд удобно представлять в виде суммы этих четырех компонент и одной из целей анализа является разложение ряда на эти составляющие для отдельного изучения. Для рядов экономических данных наличие тренда, вероятнее всего, может объясняться наличием постоянных тенденций, однообразно действующих в одном направлении. Колебания относительно долгосрочного движения, скорее всего, происходят по совокупности объективных причин. И всем регулярным компонентам, в том числе и сезонной составляющей, присуще некоторое возмущение, порождаемое случайными событиями, появление которых предсказать не представляется возможным. Но то, что это имеет место в действительности и что эффекты от различных сил аддитивны, является предположением и имеет характер гипотезы, которую всегда следуют отклонить, если сконструированная модель плохо соответствует исходным данным.

Пусть наблюдаемым временным рядом являются значения некоторой переменной величины $z_1, z_2, \ldots, z_T$, то есть имеется T чисел таких значений, наблюдаемых в равностоящие моменты времени. Эти моменты пронумерованы целыми числами *1, 2, …,T*. Отрицательные значения этих чисел означают соответствующие моменты в прошлом. Достаточно общей математической (статистической или вероятностной) моделью служит уравнение вида

$$z_t = f(t) + u_t,\ t = 1, 2, \ldots, T \tag{3.1}$$

В этой модели наблюдаемый ряд рассматривается как сумма некоторой полностью детерминированной последовательности $f(t)$, которую можно назвать систематической составляющей, и случайной последовательности u_t, подчиняющейся некоторому вероятностному закону. Эти компоненты наблюдаемого ряда ненаблюдаемы, они являются теоретическими величинами. Например, если фиксируется ежедневный объем продаж некоторой компании, то $f(t)$ может представлять планируемую величину реализации продукции, а u_t – особенности спроса данного дня, в эту же составляющую включаются искажения или ошибки в собираемой информации.

Рассмотрение временного параметра в модели согласуется с интуитивным представлением о том, что он собой представляет. Одно из таких представлений состоит в том, что время течет в одном направлении и не может быть обращено вспять. Другое – что события, близкие по времени, должны быть сравнительно сильно связаны, а события, разделенные большим промежутком времени, не должны иметь сильной связи. Можно рассматривать различные варианты математической модели (3.1), в которых влияние времени может сказываться либо только на функции или последовательности $f(t)$, либо только на вероятностном процессе, определяющем случайную составляющую u_t, либо, наконец, на обеих этих компонентах.

Первоначально анализ временных рядов базировался на моделях, в которых влияние временного параметра проявлялось только в систематической составляющей. Эта ситуация называется классической [51], поскольку она в известной мере восходит к тем временам, когда Гаусс и его последователи развивали теорию и метод наименьших квадратов с целью применения их в астрономии и физике. В таких моделях предполагается, что течение времени никак не отражается на случайной составляющей. То есть предполагается, что математическое ожидание случайной составляющей тождественно равно нулю, дисперсия равна некоторой постоянной и что значения u_t в различные моменты времени некоррелированны. Такое определение приводит к тому, что всякую зависимость от времени приходится включать в систематическую составляющую $f(t)$.

Последовательность $f(t)$ может зависеть от некоторых неизвестных коэффициентов и от известных величин, меняющихся со временем. В этом случае ее называют функцией регрессии. Методы статистических выводов для коэффициентов функции регрессии оказываются полезными во многих областях статистики. Своеобразие методов, относящихся именно к временным рядам, состоит в том, что здесь исследуются те модели, в которых упомянутые выше величины, меняющиеся со временем, являются известными функциями t.

Анализ временных рядов имеет важное практическое значение во многих областях деятельности. А одной из важных практических целей является прогнозирование временных рядов. Использование доступных к моменту времени t наблюдений временного ряда для прогнозирования его значения в некоторый момент времени в будущем $t+l$ может явиться основой для планирования в экономике и торговле, планирования выпуска продукции, складского контроля и контроля выпуска, управления и оптимизации промышленных процессов, решения социальных проблем и многих-многих других задач. Для всех этих задач существует необходимость в прогнозе вперед на интервал, называемым временем упреждения и зависящий от конкретной проблемы. К примеру, время упреждения в проблеме складского контроля определено как период, начинающийся с момента передачи заказа на пополнение склада поставщиком и длящийся до тех пор, пока заказ не доставлен на склад.

Предполагается, что наблюдения доступны в дискретные, равноотстоящие моменты времени. Например, в проблеме прогнозирования сбыта продукции сбыт z_t в текущем месяце t и сбыт в предыдущие месяцы могут быть использованы для прогноза сбыта с упреждением $l = 1, 2, \dots, 12$ месяцев.

Обозначим через $\hat{z}_t(l)$ сделанный в момент t прогноз сбыта z_{t+l} в некоторый момент $t+l$ в будущем, т.е. с упреждением l.

Функция $\hat{z}_t(l), l = 1, 2, \dots$, дающая в момент времени t прогнозы для всех будущих времен упреждения, называется прогнозирующей функцией в момент времени t [52]. Цель исследования временного ряда состоит в том, чтобы получить такую прогнозирующую функцию, у которой среднее значение квадрата отклонения $z_{t+l} - \hat{z}_t(l)$ – истинного значения от прогнозируемого значения является наименьшим для каждого упреждения l. Такой прогноз традиционно считается наилучшим.

В дополнение к вычислению наилучшего прогноза необходимо также указать его точность, чтобы, например, можно было бы оценить риск, связанный с решениями, основанными на прогнозировании. Точность прогноза может быть выражена вероятностными пределами по обе стороны от каждого прогнозного значения, т.е. являться следствием построения интервального прогноза. Смысл такого прогнозного интервала состоит в том, что значение временного ряда, которое появится в соответствующее время, с заданной наперед вероятностью будет лежать внутри этого интервала.

Идея использования математических моделей для описания поведения физических объектов является общепризнанной. В частности, иногда удается получить модель, основанную на физических законах, что дает возможность вычислить точное значение какой-либо зависящей от времени величины в любой момент времени. Такая модель является детерминированной. Однако для большинства практических задач получить совершенно точное значение исследуемого признака нельзя из-за случайных возмущений, действующих на объект исследования. Например, остатки на депозитных счетах коммерческого банка на определенный момент, на которые оказывают влияние большое число различных факторов. Для таких объектов нельзя предложить детерминированную модель, допускающую точное вычисление будущего поведения объекта. Такая модель называется вероятностной или стохастической.

Модели временных рядов, необходимые для получения оптимального прогнозирования в действительности являются стохастическими. При этом необходимо различать стохастический процесс, с которым идентифицируется вероятностная модель, и наблюденный временной ряд [52, с.22]. Временной ряд $z_1, z_2, \dots, z_T$ из T последовательных наблюдений рассматривается как выборочная реализация из бесконечной популяции таких временных рядов, генерируемых стохастическим процессом.

Важным и одним из самых ранних классом стохастических моделей для описания временных рядов является класс стационарных моделей. Они основаны на предположении, что основные вероятностные характеристики стохастического процесса являются неизменными с течением времени, т.е. процесс остается в равновесии относительно постоянного среднего уровня. Однако в индустрии, коммерции, экономике и финансовой сфере, где прогнозирование имеет особо важное значение, многие временные ряды лучше описываются нестационарными моделями, к которым, в частности, можно отнести модель экспоненциально взвешенных скользящих средних [53].

Укажем на основные виды классических моделей, используемых в анализе временных рядов [52]. Для этого необходимо ввести некоторые простые операторы.

Оператор сдвига назад B, определяемый как $Bz_t = z_{t-1}$, и тогда, соответственно, возведение оператора в степень m будет означать сдвиг назад на m шагов: $B^m z_t = z_{t-m}$.

Оператор сдвига вперед (обратная по отношению к оператору B) $F = B^{-1}$. Этот оператор задается как $Fz_t = z_{t+1}$, и, соответственно, $F^m z_t = z_{t+m}$.

Разностный оператор со сдвигом назад ∇, который можно выразить через оператор B следующим образом: $\nabla z_t = z_t - z_{t-1} = (1 - B)z_t$.

Оператор суммирования S – обратный оператору ∇, выражаемый как

$$Sz_t = \sum_{j=0}^{\infty} z_{t-j} = z_t + z_{t-1} + z_{t-2} + \cdots = (1 + B + B^m + \cdots)z_t = (1 - B)^{-1}z_t == \nabla^{-1}z_t.$$

Модель линейного фильтра. Эта модель исходит из предположения о том, что временные ряды, в которых последовательные значения сильно зависимы, следует рассматривать как генерируемые последовательностью независимых импульсов a_t. Эти импульсы – реализации случайных величин с фиксированным распределение, которое обычно предполагается нормальным с нулевым средним и дисперсией σ_a^2. Такая последовательность случайных величин $a_t, a_{t+1}, a_{t+2}, \ldots$ называется белым шумом.

Предполагается, что белый шум a_t можно трансформировать в процесс z_t при помощи линейного фильтра $\psi(B)$. Операция линейной фильтрации заключается в вычислении взвешенной суммы предыдущих наблюдений так, что стохастический процесс приобретает вид:

$$z_t = \mu + a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \cdots = \mu + \psi(B)a_t. \qquad (3.2)$$

В общем $\mu -$ параметр, определяющий «уровень» процесса, и

$$\psi(B) = 1 + \psi_1 B + \psi_2 B^2 + \cdots$$

$-$ линейный оператор, преобразующий a_t в z_t и называемый передаточной функцией фильтра.

Последовательность $\psi_1, \psi_2, \ldots$, образованная весами, может быть конечной или бесконечной. Если эта последовательность является сходящейся, фильтр называется устойчивым, а процесс z_t будет стационарным. Параметр μ в этом случае – среднее значение, вокруг которого процесс варьирует. В других случаях z_t нестационарен, и μ не имеет какого-либо иного смысла, кроме как некоторой точки отсчета уровня процесса.

Модели авторегрессии. Так называемая модель авторегрессии является исключительно полезной стохастической моделью для описания некоторых часто встречающихся на практике рядов. В этой модели текущее значение процесса выражается как конечная линейная совокупность предыдущих значений процесса и импульса a_t. Обозначим значения стохастического процесса в равноотстоящие моменты времени $t, t-1, t-2, \ldots$ как $z_t, z_{t-1}, z_{t-2}, \ldots$. Пусть $\tilde{z}_t, \tilde{z}_{t-1}, \tilde{z}_{t-2}, \ldots$ будут отклонениями от μ, например $\tilde{z}_t = z_t - \mu$. Тогда

$$\tilde{z}_t = \phi_1 \tilde{z}_{t-1} + \phi_2 \tilde{z}_{t-2} + \cdots + \phi_p \tilde{z}_{t-p} + a_t \qquad (3.3)$$

называется процессом авторегрессии (*АР*) порядка p. В (3.3) $\phi_1, \phi_2, \ldots, \phi_p$ – некоторые коэффициенты.

Название «авторегрессия» объясняется тем, что линейная модель вида

$$z = \phi_1 x_1 + \phi_2 x_2 + \cdots + \phi_p x_p + a,$$

связывающая зависимую переменную z со множеством, так называемых независимых переменных $x_1, x_2, \dots, x_p$ плюс свободный член a, отождествляемый с ошибкой или случайным отклонением, называется уравнением регрессии; при этом говорят, что z регрессирует на $x_1, x_2, \dots, x_p$. В уравнении (3.3) переменная $\tilde{z}_t$ регрессирует на свои предшествующие значения, поэтому модель авторегрессирующая.

Если определить, что оператор авторегрессии порядка p имеет вид

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p,$$

то модель авторегрессии можно компактно записать в следующем виде:

$$\phi(B)\tilde{z}_t = a_t.$$

Эта модель содержит $p + 2$ неизвестных параметра: $\mu, \phi_1, \phi_2, \dots, \phi_p, \sigma_a^2$, которые на практике следует оценить по наблюдениям. Дополнительный параметр σ_a^2 − дисперсия белого шума.

Нетрудно заметить, что модель авторегрессии является частным видом модели линейного фильтра (3.2). Величину $\tilde{z}_{t-1}$ можно исключить из правой части (3.3) подстановкой из уравнения (3.3), считая ее зависимой переменной:

$$\tilde{z}_{t-1} = \phi_1 \tilde{z}_{t-2} + \phi_2 \tilde{z}_{t-3} + \cdots + \phi_p \tilde{z}_{t-p-1} + a_{t-1}.$$

Аналогичным образом можно исключить $\tilde{z}_{t-2}$ и так далее, получив в результате бесконечный ряд из a. С помощью введенных операторов модель авторегрессии можно записать в следующем виде:

$$\tilde{z}_t = \psi(B) a_t,$$

где

$$\psi(B) = \phi^{-1}(B).$$

Процессы авторегрессии могут быть как стационарными, так и нестационарными. Чтобы процесс был стационарным, нужно выбрать ϕ так, чтобы веса $\psi_1, \psi_2, \dots$, в выражении $\psi(B) = \phi^{-1}(B)$ образовывали сходящийся ряд.

Модели скользящего среднего. Модель авторегрессии (3.3) выражает отклонение стохастического процесса $\tilde{z}_t$ в виде конечной взвешенной суммы p предыдущих отклонений процесса $\tilde{z}_{t-1}, \tilde{z}_{t-2}, \dots, \tilde{z}_{t-p}$ плюс случайный импульс a_t. А также эта модель выражает это же отклонение как бесконечную взвешенную сумму случайных импульсов.

Другим очень важным в описании наблюдаемых временных рядов типом моделей является так называемый процесс скользящего среднего. Пусть $\tilde{z}_t$ линейно зависит от конечного числа q предыдущих случайных импульсов a. Такой стохастический процесс

$$\tilde{z}_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \cdots - \theta_q a_{t-q} \tag{3.4}$$

называется процессом скользящего среднего (CC) порядка q, где $\theta_1, \theta_2, \dots, \theta_q$ коэффициенты модели или веса, которые не обязательно должны быть положительными.

Если определить оператора скользящего среднего порядка q как

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q,$$

то модель скользящего среднего можно компактно представить в следующем виде:

$$\tilde{z}_t = \theta(B) a_t. \tag{3.5}$$

Она содержит $q + 2$ неизвестных параметра $\mu, \theta_1, \theta_2, \dots, \theta_p, \sigma_a^2$, которые на практике должны оцениваться по наблюдениям.

Смешанные модели авторегрессии – скользящего среднего. Для достижения большей гибкости в подгонке моделей к наблюдаемым временным рядам иногда целесообразно объединить в одной модели и авторегрессию, и скользящее среднее. Это приводит к комбинированной модели авторегрессии – скользящего среднего

$$\tilde{z}_t = \phi_1 \tilde{z}_{t-1} + \phi_2 \tilde{z}_{t-2} + \cdots + \phi_p \tilde{z}_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \cdots - \theta_q a_{t-q}, \tag{3.6}$$

или

$$\phi(B)\tilde{z}_t = \theta(B) a_t,$$

в которой имеется $p + q + 2$ неизвестных параметра: $\mu, \phi_1, \phi_2, \dots, \phi_p, \theta_1, \theta_2, \dots, \theta_p, \sigma_a^2$, оцениваемых по наблюдаемым данным.

На практике часто оказывается, что адекватное описание наблюдаемых временных рядов достигается при помощи моделей авторегрессии, скользящего среднего или комбинированной модели, в которых p и q не больше, а часто и меньше 2 [52].

Нестационарные модели. Многие ряды, практически встречающиеся в промышленности или торговле (например, биржевые цены), обнаруживают нестационарный характер и, в частности, не колеблются относительно фиксированного среднего. Тем не менее, их свойства могут быть в некотором смысле однородными. Например, хотя уровень относительно которого происходят флуктуации, может быть разным в разные времена, поведение рядов с учетом различий в уровне оказывается во многом сходным. Такой ряд может быть представлен обобщенным оператором авторегрессии $\varphi(B)$, в котором один или несколько нулей полинома $\varphi(B)$ (т.е. один или несколько корней уравнения $\varphi(B) = 0$) равны единице. В этом случае оператор $\varphi(B)$ имеет вид

$$\varphi(B) = \phi(B)(1 - B)^d,$$

где $\phi(B)$ − оператор стационарной авторегрессии, d − некоторое целое число. При этом обобщенная модель, описывающая однородный нестационарный процесс, имеет вид

$$\varphi(B)z_t = \phi(B)(1 - B)^d z_t = \theta(B)a_t,$$

т.е.

$$\phi(B)w_t = \theta(B)a_t, \qquad (3.7)$$

где

$$w_t = \nabla^{\mathrm{d}} z_t. \qquad (3.8)$$

Таким образом, однородный нестационарный процесс может быть описан моделью, которая требует чтобы -я разность процесса была стационарной. На практике d обычно равно 0, 1 или максимум 2.

Процесс, определенный (3.7) или (3.8), представляет собой эффективную модель для описания стационарных и нестационарных временных рядов и называется процессом авторегрессии – проинтегрированного скользящего среднего (АРПСС) порядка $(p,\ d, q)$. Процесс определен как

$$w_t = \phi_1 w_{t-1} + \cdots + \phi_p w_{t-p} + a_t - \theta_1 a_{t-1} - \cdots - \theta_q a_{t-q} \qquad (3.9)$$

с $w_t = \nabla^{\mathrm{d}} z_t$. Следует отметить, что если заменить w_t на $z_t - \mu$, то при $d = 0$ модель (3.9) будет содержать как частный случай и стационарную комбинированную модель (3.6), а также модели авторегрессии и скользящего среднего.

Если из выражения (3.8) выразить z_t, то получится следующее соотношение:

$$z_t = S^d w_t, \qquad (3.10)$$

где S – оператор суммирования, определенный как

$$Sw_t = \sum_{j=0}^{\infty} w_{t-j} = w_t + w_{t-1} + w_{t-2} + \cdots$$

Таким образом, обобщенный процесс авторегрессии – проинтегрированного скользящего среднего можно получить из белого шума a_t с помощью трех операций фильтрации. Первый фильтр имеет вход a_t передаточную функцию $\theta(B)$ и выход e_t, где

$$e_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \cdots - \theta_q a_{t-q} = \theta(B)a_t. \qquad (3.11)$$

Второй фильтр имеет вход e_t, передаточную функцию $\phi^{-1}(B)$ и выход w_t

$$w_t = \phi_1 w_{t-1} + \cdots + \phi_p w_{t-p} + e_t = \phi^{-1}(B)e_t. \qquad (3.12)$$

И наконец, третий фильтр имеет вход w_t и выход z_t с передаточной функцией S^d.

Для описания сезонных временных рядов используется специальная форма модели (3.9).

Рассмотрим некоторые проблемы прогнозирования с использованием универсальной модели процесса авторегрессии – проинтегрированного скользящего среднего (АРПСС)

$$\varphi(B)z_t = \theta(B)a_t, \qquad (3.13)$$

где оператор $\varphi(B)$ имеет вид следующего выражения: $\varphi(B) = \phi(B)\nabla^{\mathrm{d}}$.

Рассмотрим прогноз значения $z_{t+l}, l \geq 1$ по отношению к текущему моменту времени *t*. О таком прогнозе можно сказать, что он делается в момент времени *t* с упреждением *l*.

Для этого проанализируем три явные формы представления модели. Наблюдение z_{t+l}, генерируемое процессом (3.13), можно выразить следующим образом:

1) непосредственно при помощи разностного уравнения

$$z_{t+l} = \varphi_1 z_{t+l+1} + \cdots + \varphi_{p+d} z_{t+l-p-d} - \theta_1 a_{t+l-1} - \cdots - \theta_q a_{t+l-q} + a_{t+l}; \quad (3.14)$$

2) как бесконечную взвешенную сумму текущего и предшествующих импульсов a_j

$$z_{t+l} = \sum_{j=-\infty}^{t+l} \psi_{t+l-j} a_j = \sum_{j=0}^{\infty} \psi_j a_{t+l-j}, \quad (3.15)$$

где ψ_0 и веса ψ можно найти приравниванием коэффициентов при одинаковых степенях B правой и левой частей уравнения

$$\varphi(B)(1 + \psi_1 B + \psi_2 B^2 + \cdots) = \theta(B). \quad (3.16)$$

3) как бесконечную взвешенную сумму предыдущих наблюдений плюс случайный импульс

$$z_{t+l} = \sum_{j=1}^{\infty} \pi_j z_{t+l-j} + a_{t+l}. \quad (3.17)$$

Если $d \geq 1$, то

$$\tilde{z}_{t+l-1} = \sum_{j=1}^{\infty} \pi_j z_{t+l-j} \quad (3.18)$$

будет взвешенным средним, потому что сумма весов равна единице: $\sum_{j=1}^{\infty} \pi_j = 1$. Веса π можно получить из равенства

$$\varphi(B) = (1 - \pi_1 B - \pi_2 B^2 - \cdots)\theta(B). \quad (3.19)$$

Обозначим прогноз величины z_{t+l} как $\hat{z}_t(l)$, при этом он является линейной функцией текущего и предшествующих наблюдений $z_t, z_{t-1}, z_{t-2}, \ldots$. Поэтому он также будет линейной функцией текущего и предшествующих импульсов $a_t, a_{t-1}, a_{t-2}, \ldots$.

Предположим, что наилучший прогноз существует и имеет вид

$$\hat{z}_t(l) = \psi_l^* a_t + \psi_{l+1}^* a_{t-1} + \psi_{l+2}^* a_{t-2} + \cdots,$$

где веса $\psi_l^*, \psi_{l+1}^*, \ldots$ должны быть определены. Тогда с учетом (3.14) среднеквадратическая ошибка прогноза равна

$$\mathcal{M}\{z_{t+l} - \hat{z}_t(l)\}^2 = \left(1 + \psi_1^2 + \cdots + \psi_l^2\right)\sigma_a^2 + \sum_{j=0}^{\infty} (\psi_{l+j} - \psi_{l+j}^*)^2 \sigma_a^2. \quad (3.20)$$

В последнем выражении $\mathcal{M}\{\cdot\}$ – символ математического ожидания.

Ошибка прогноза может быть минимизирована приравниванием $\psi_{l+j} == \psi_{l+j}^*$. Этот вывод является частным случаем более общих результатов теории прогнозирования, развитой Колмогоровым [54]. В результате имеет место соотношение [52]:

$$z_{t+l} = (a_{t+l} + \psi_1 a_{t+l-1} + \cdots + \psi_{l-1} a_{t+1}) + (\psi_l a_t + \psi_{l+1} a_{t-1} + \cdots), \quad (3.21)$$

$$z_{t+l} = e_t(l) + \hat{z}_t(l), \quad (3.22)$$

где $e_t(l)$ – ошибка прогноза $\hat{z}_t(l)$ с упреждением l.

Отсюда вытекает ряд важных выводов. Обозначим условное математическое ожидание z_{t+l} при условии, что все z до момента t известны через $\mathcal{M}_t\{z_{t+l}\}$. Тогда

$$\hat{z}_t(l) = \psi_l a_t + \psi_{l+1} a_{t-1} + \cdots = \mathcal{M}_t\{z_{t+l}\}. \quad (3.23)$$

Это значит, что прогноз с минимальной среднеквадратической ошибкой в момент t с упреждением l есть условное математическое ожидание z_{t+l} в момент t. Когда $\hat{z}_t(l)$ рассматривается как функция l при фиксированном t она называется прогнозирующей функцией для момента t.

Ошибка прогноза для упреждения l равна

$$e_t(l) = a_{t+l} + \psi_1 a_{t+l-1} + \cdots + \psi_{l-1} a_{t+1}. \quad (3.24)$$

Так как математическое ожидание ошибки равно нулю: $\mathcal{M}_t\{e_t(l)\}$, прогноз будет несмещенный.

Дисперсия ошибки прогноза $V(l)$ равна

$$V(l) = var\{e_t(l)\} = (1 + \psi_1^2 + \cdots + \psi_{l-1}^2)\sigma_a^2. \quad (3.25)$$

Еще один важный вывод состоит в том, что не только $\hat{z}_t(l)$ есть прогноз величины z_{t+l} с минимальной среднеквадратической ошибкой, но и любая линейная комбинация прогнозов$\sum_{l=1}^{L} w_l \hat{z}_t(l)$, где L – число таких прогнозов, w – некоторые веса этих прогнозов, – также является прогнозов с минимальной среднеквадратической ошибкой соответствующей линейной комбинации будущих наблюдений: $\sum_{l=1}^{L} w_l z_{t+l}$.

Например, пусть с помощью выражения (3.23) по ежемесячным данным получен прогноз сбыта товара на один, два и три месяца вперед: $\hat{z}_t(1), \hat{z}_t(2), \hat{z}_t(3)$. Тогда $\hat{z}_t(1) + \hat{z}_t(2) + \hat{z}_t(3)$ – это прогноз с минимальной среднеквадратической ошибкой сбыта товара $z_{t+1} + z_{t+2} + z_{t+3}$ в течение следующего квартала.

Остаточные ошибки как ошибки прогноза на один шаг вперед определяются на основании выражения (3.19), и имеют вид

$$e_t(1) = z_{t+1} - \hat{z}_t(1) = a_{t+1}. \tag{3.26}$$

Следовательно, остаточные ошибки, генерирующие процесс, ошибки, которые вводятся как множество независимых случайных величин или импульсов, тождественны ошибкам прогноза на шаг вперед.

Отсюда следует, что для прогноза с минимальной среднеквадратической ошибкой, ошибки прогноза на шаг вперед должны быть некоррелированными. Это утверждение выглядит вполне логичным, поскольку в противном случае, т.е. если бы они были коррелированными, то ошибка прогноза на шаг вперед a_{t+1} могла быть предсказана на основании известных ошибок прогноза $a_t, a_{t-1}, a_{t-2}, \ldots$ И в этом случае сумма прогноза ошибки и самого прогноза был бы лучшим прогнозом, что противоречит утверждению о единственности лучшего прогноза.

Однако следует отметить, что ошибки прогноза $e_t(l)$ и $e_t(l+j)$, сделанные для разных упреждений с одного и того же момента времени t, все-таки коррелированны, если эти моменты l и j времени значительно отстоят от текущего момента времени t. Одно из следствий этого состоит в том, что часто функции прогноза лежат либо целиком выше, либо целиком ниже фактических значений ряда.

То, что прогноз с минимальной среднеквадратической ошибкой прогноза $\hat{z}_t(l)$ для упреждения l – это условное математическое ожидание $\mathcal{M}_t\{z_{t+l}\}$ случайной величины z_{t+l} в момент времени t, позволяет записать выражения для прогноза любым из трех различных способов, соответствующих трем способам представления модели: (3.14), (3.15) и (3.17). Для $l > 0$ имеется три различных способа выражения прогнозов.

Прогнозы, полученные из разностного уравнения. Переходя в (3.14) к условным математическим ожиданиям в момент времени t, получаем

$$\mathcal{M}_t\{z_{t+l}\} = \hat{z}_t(l) = \varphi_1 \mathcal{M}_t\{z_{t+l-1}\} + \cdots + \varphi_{p+d} \mathcal{M}_t\{z_{t+l-p-d}\} - \theta_1 \mathcal{M}_t\{a_{t+l-1}\} - \cdots - \theta_q \mathcal{M}_t\{a_{t+l-q}\} + a_{t+l}. \tag{3.27}$$

Прогноз в проинтегрированном виде. Пользуясь (3.15), можно получить выражение для прогноза

$$\mathcal{M}_t\{z_{t+l}\} = \hat{z}_t(l) = \psi_1 \mathcal{M}_t\{a_{t+l-1}\} + \cdots + \psi_{l-1} \mathcal{M}_t\{a_{t+1}\} + \psi_t \mathcal{M}_t\{a_t\} + \psi_{l+1} \mathcal{M}_t\{a_{t-1}\} + \cdots + a_{t+1}, \tag{3.28}$$

Прогноз, как взвешенное среднее предшествующих наблюдений и прогнозов, сделанных в тот же момент с меньшими упреждениями. Наконец, переходя в (3.17) к условным математическим ожиданиям, можно получить третью форму выражения прогноза:

$$\mathcal{M}_t\{z_{t+l}\} = \hat{z}_t(l) = \sum_{j=1}^{\infty} \pi_j \mathcal{M}_t\{z_{t+l-j}\} + \mathcal{M}_t\{a_{t+l}\}. \tag{3.29}$$

Нужно отметить, что прогноз с минимальной среднеквадратической ошибкой определен через условное математическое ожидание

$$\mathcal{M}_t\{z_{t+l}\} = \mathcal{M}_t\{z_{t+l} | z_t, z_{t-1}, \ldots\},$$

для которого теоретически требуется знание всех прошлых z до бесконечности. Однако условие обратимости, наложенное на общую модель АРПСС, требует, чтобы веса π образовывали сходящийся ряд. Отсюда следует, что для вычисления прогноза с заданной точностью, начиная с некоторого k, влиянием $z_{t-j}, j > k$, можно пренебречь. На практике веса π обычно затухают довольно быстро так, что какое бы представление модели ни использовалось в вычислениях, лишь небольшая часть ряда $z_t, z_{t-1}, \ldots, z_{t-k}$ необходима для вычисления прогноза с достаточной точностью.

При вычислении условных математических ожиданий, входящих в выражения (3.27), (3.28), (3.29), следует учесть, что если j – неотрицательное числа, то

$$\mathcal{M}_t\{z_{t-j}\} = z_{t-j}, j = 0, 1, 2, \ldots,$$
$$\mathcal{M}_t\{z_{t+j}\} = \hat{z}_t(j), j = 0, 1, 2, \ldots,$$
$$\mathcal{M}_t\{a_{t-j}\} = a_{t-j} = z_{t-j} - \tilde{z}_{t-j-1}(1), j = 0, 1, 2, \ldots,$$
$$\mathcal{M}_t\{a_{t+j}\} = 0, j = 0, 1, 2, \ldots . \tag{3.30}$$

Следовательно, чтобы получить прогноз $\hat{z}_t(l)$, необходимо выразить модель z_{t+l} с помощью любого из трех рассмотренных выше явных представлений и поступать с членами в правой части согласно следующим правилам:

- члены $z_{t-j}, (j = 0, 1, 2, \dots)$, известные уже к моменту t, оставить без изменения;
- члены $z_{t+j}, (j = 0, 1, 2, \dots)$, еще не известные, заменить их прогнозами $\hat{z}_t(j)$ на момент времени t;
- члены $a_{t-j}, (j = 0, 1, 2, \dots)$, уже известные, определить по выражению

$$z_{t-j} - \tilde{z}_{t-j-1}(1);$$

- члены уравнений вида $a_{t+j}, (j = 0, 1, 2, \dots)$, еще не известные, заменить нулями.

Непосредственный выбор модели, по которой осуществляется прогноз, зависит от особенностей поставленной в исследовании задачи. На практике же чаще всего удобнее пользоваться непосредственно представлением модели разностным уравнением (3.14).

Пример [52, с.128]. Ряд значений биржевых цен на акции компании IBM к концу торгов хорошо описываются моделью

$$(1 - 0{,}8B)(1 - B)z_{t+l} = a_{t+l},$$

т.е.

$$(1 - 1{,}8B + 0{,}8B^2)z_{t+l} = a_{t+l},$$

или

$$z_{t+l} = 1{,}8z_{t+l-1} - 0{,}8z_{t+l-2} + a_{t+l}.$$

Прогнозы в момент времени t будут иметь вид

$$\hat{z}_t(1) = 1{,}8z_t - 0{,}8z_{t-1},$$
$$\hat{z}_t(2) = 1{,}8\hat{z}_t(1) - 0{,}8z_t,$$
$$\hat{z}_t(l) = 1{,}8\hat{z}_t(l-1) - 0{,}8\hat{z}_t(l-2), l = 3, 4, 5, \dots .$$

Видно, что прогнозы легко вычисляются рекуррентным способом, начиная с $\hat{z}_t(1), \hat{z}_t(2), \dots$.

Классический анализ временных рядов рассматривает лишь одну случайную величину. Рассмотрение совокупности переменных, одновременно наблюдаемых в соответствующие моменты времени, выходит за рамки классических или традиционных представлений. А такой подход является достаточно актуальным, поскольку на значение переменной оказывает влияние не только время, как некоторый фактор, а также и значения других переменных, от которых данная переменная может зависеть вследствие объективных причинно-следственных связей.

В данном разделе рассмотрен так называемый параметрический подход к анализу временных рядов. Помимо него существует также спектральный анализ временных рядов или непараметрический подход. Представление наблюдаемых значений, как некоторого сигнала в форме ряда Фурье, также может быть использовано для целей прогнозирования, однако, чаще всего исследование гармоник, на которые раскладывается реальный ряд, служит для анализа и объяснения апостериорных значений переменной.

В анализе временных рядов исторически имели место два довольно обособленных направления в исследовании временных рядов: частотный или гармонический подход и направление, связанное с анализом зависимости от времени.

Наиболее полный обзор методов анализа многомерных временных рядов с применением частотного подхода приведен у Д. Бриллинджера [55]. Как уже отмечалось, практическое использование методов многомерных временных рядов для анализа экономических проблем не нашло достаточного распространения по ряду причин. Важным условием применения этих методов является предположение об инвариантности по отношению к временным сдвигам: доля значений анализируемого признака, попавших в некоторый временной интервал, должна быть примерно такой же для другого временного интервала той же длины, но имеющего некоторый временной сдвиг по отношению к первоначальному интервалу.

Типичные физические эксперименты обладают свойством временной инвариантности. Например, для многих практических целей неважно, в какой именно день начата серия измерений силы тяжести; если рассмотреть ряд изменений температуры в течение десятков лет, можно убедиться, что он обладает свойством инвариантности. Однако поведение рядов, имеющих отношение к экономике и связанных с социально-экономическими процессами, не обладают свойством инвариантности по отношению ко времени. Объяснением этому служит достаточно очевидное предположение: люди извлекают уроки из прошлого и соответственно меняют свое поведение. Поэтому ряды, имеющие отношение к человеческой деятельности, вообще говоря, не обладают временной инвариантностью [55, с. 13].

Воспроизведем основную модель и основные результаты в многомерном анализе временных рядов, связанных с частотным подходом. Рассматривается $(r+s)$ -мерный временной стационарный ряд

$$\begin{pmatrix} \boldsymbol{X}(t) \\ \boldsymbol{Y}(t) \end{pmatrix}, \tag{3.31}$$

$t = 0, \pm 1, \pm 2, \ldots$, составленный из -мерного ряда $\boldsymbol{X}(t)$ и s-мерного ряда $\boldsymbol{Y}(t)$.

Обозначим математические ожидания компонент ряда

$$\mathcal{M}\{\, \boldsymbol{X}(t)\} = \boldsymbol{\mu_X},$$
$$\mathcal{M}\{\, \boldsymbol{Y}(t)\} = \boldsymbol{\mu_Y}, \tag{3.32}$$

и ковариации

$$\begin{gathered} \mathcal{M}\{\, (\boldsymbol{X}(t+u) - \boldsymbol{\mu_X})(\boldsymbol{X}(t+u) - \boldsymbol{\mu_X})^T\} = \boldsymbol{\Sigma_{XX}}(u), \\ \mathcal{M}\{\, (\boldsymbol{X}(t+u) - \boldsymbol{\mu_X})(\boldsymbol{Y}(t+u) - \boldsymbol{\mu_Y})^T\} = \boldsymbol{\Sigma_{XY}}(u), \\ \mathcal{M}\{\, (\boldsymbol{Y}(t+u) - \boldsymbol{\mu_Y})(\boldsymbol{Y}(t+u) - \boldsymbol{\mu_Y})^T\} = \boldsymbol{\Sigma_{YY}}(u), \end{gathered} \tag{3.33}$$

$u = 0, \pm 1, \pm 2, \ldots$, а также спектральные плотности второго порядка

$$\begin{gathered} \boldsymbol{f_{XX}}(\lambda) = \frac{1}{2\pi}\sum_{u=-\infty}^{\infty} \boldsymbol{\Sigma_{XX}}(u) e^{-i\lambda u}, \\ \boldsymbol{f_{XY}}(\lambda) = \frac{1}{2\pi}\sum_{u=-\infty}^{\infty} \boldsymbol{\Sigma_{XY}}(u) e^{-i\lambda u}, \\ \boldsymbol{f_{YY}}(\lambda) = \frac{1}{2\pi}\sum_{u=-\infty}^{\infty} \boldsymbol{\Sigma_{YY}}(u) e^{-i\lambda u}, \end{gathered} \tag{3.34}$$

где λ – частота гармоники $(-\infty < \lambda < \infty)$, $i = \sqrt{-1}$ – мнимая единица.

Задача анализа временных рядов состоит в выборе такого -мерного вектора $\boldsymbol{c}$ и такого $s \times r$-фильтра $\boldsymbol{A}\{u\}$, чтобы ряд

$$\boldsymbol{c} + \sum_{u=-\infty}^{\infty} \boldsymbol{A}(t+u)\boldsymbol{X}(u) \tag{3.35}$$

был в некотором смысле близок к $\boldsymbol{Y}(t)$.

Получим оценки $\boldsymbol{c}$ и $\boldsymbol{A}\{u\}$, основанные на конечных выборках значений $\boldsymbol{X}(t)$ и $\boldsymbol{Y}(t)$, $t = 0, 1, \ldots, T-1$.

Пусть ряд (3.31) имеет выборочное среднее значение

$$\begin{pmatrix} \boldsymbol{c_X} \\ \boldsymbol{c_Y} \end{pmatrix}, \tag{3.36}$$

и выборочную ковариационную матрицу

$$\begin{pmatrix} \boldsymbol{C_{XX}} & \boldsymbol{C_{XY}} \\ \boldsymbol{C_{YX}} & \boldsymbol{C_{YY}} \end{pmatrix}. \tag{3.37}$$

Укажем на такой -компонентный вектор $\boldsymbol{c}$ и такую $(s \times r)$ матрицу $\boldsymbol{A}$, которые минимизируют $(s \times s)$ матрицу

$$\mathcal{M}\{\, (\boldsymbol{Y} - \boldsymbol{c} - \boldsymbol{AX})(\boldsymbol{Y} - \boldsymbol{c} - \boldsymbol{AX})^T\}. \tag{3.38}$$

При этом под минимумом по отношению к матрицам понимается упорядочение $A \geq B$, означающее, что матрица $A - B$ неотрицательно определена.

Пусть задан $(r+s)$-мерный вектор

$$\begin{pmatrix} \boldsymbol{X} \\ \boldsymbol{Y} \end{pmatrix}$$

со средним (3.36) и ковариационной матрицей (3.37) и матрица $\boldsymbol{C_{XX}}$ – невырождена, тогда матрицу (3.38) минимизируют следующие величины

$$\boldsymbol{c} = \boldsymbol{c_Y} - \boldsymbol{C_{YX}}\boldsymbol{C_{XX}^{-1}}\boldsymbol{c_X} \tag{3.39}$$

и

$$\boldsymbol{A} = \boldsymbol{C_{YX}}\boldsymbol{C_{XX}^{-1}}. \tag{3.40}$$

Соответствующее минимальное значение равно

$$\boldsymbol{C_{YY}} - \boldsymbol{C_{YX}}\boldsymbol{C_{XX}^{-1}}\boldsymbol{C_{XY}}. \tag{3.41}$$

При этом случайный вектор

$$\boldsymbol{c_Y} + \boldsymbol{C_{YX}}\boldsymbol{C_{XX}^{-1}}(\boldsymbol{X} - \boldsymbol{c_X}) \tag{3.42}$$

называется наилучшим линейным прогнозом $\boldsymbol{Y}$, основанным на $\boldsymbol{X}$.

Величины (3.39) и (3.40) доставляют также минимум детерминанту, следу, диагональным элементам и собственным значениям матрицы (3.38).

Модель многомерного временного ряда (3.35) предполагает наличие детерминированного или стохастического ряда $\boldsymbol{X}(t)$, который, по аналогии с классическим регрессионным анализом, интерпретируется как матрица значений независимых переменных. Тем самым эта модель является обобщением модели классического регрессионного анализа на случай зависимости экзогенных переменных от времени и на случай их стохастичности. Таким образом, этот подход аппроксимации исследуемого ряда с помощью фильтрации другого стационарного ряда и данная модель не предполагают наличия авторегрессии изучаемой совокупности переменных. И при этом также не рассматриваются проблемы корреляции между переменными результирующего ряда $\boldsymbol{Y}(t)$.

Однако существует подход к решению проблем аппроксимации ряда посредством фильтрации не другого, а того же самого ряда. Рассмотрим основные особенности такого подхода.

Пусть имеется векторный ряд $\boldsymbol{X}(t), t = 0, \pm 1, \pm 2, \ldots$ с r компонентами. Его математическое ожидание для всех моментов времени равно

$$\mathcal{M}\{\boldsymbol{X}(t)\} = \boldsymbol{\mu_X}. \tag{3.43}$$

Этот ряд имеет абсолютно суммируемую автоковариационную функцию

$$\mathcal{M}\{(\boldsymbol{X}(t+u) - \boldsymbol{\mu_X})(\boldsymbol{X}(t+u) - \boldsymbol{\mu_X})^T\} = \boldsymbol{\Sigma_{XX}}(u), \tag{3.44}$$

и матрицу спектральной плотности

$$\boldsymbol{f_{XX}}(\lambda) = \frac{1}{2\pi}\sum_{u=-\infty}^{\infty} \boldsymbol{\Sigma_{XX}}(u) e^{-i\lambda u}. \tag{3.45}$$

Чтобы пояснить основную идею этого подхода, воспользуемся аналогией из области передачи данных. Допустим, необходимо передать по каналам связи из одного пункта в другой значения величин $\boldsymbol{X}(t)$, но при этом распоряжении имеется только $l \leq r$ каналов. Предположим, что в процессе передачи ряда данных $\boldsymbol{X}(t)$ по q каналам связи можно описать как фильтрацию, в результате которой получается -компонентный ряд

$$\boldsymbol{\Psi}(t) = \sum_u \boldsymbol{B}(t-u)\boldsymbol{X}(u), \;\; t = 0, \pm 1, \pm 2, \ldots, \tag{3.46}$$

где $\quad \boldsymbol{B}(t-u) - q\,\times$-матричный фильтр. Пусть на выходе каналов принимается ряд, который выступает в качестве оценки для $\boldsymbol{X}(t)$:

$$\boldsymbol{X}^*(t) = \boldsymbol{c} + \sum_u \boldsymbol{D}(t-u)\boldsymbol{\Psi}(u), \tag{3.47}$$

где $\quad \boldsymbol{c}$ – вектор с r компонентами,

$\boldsymbol{D}(t-u) - r \times q$-фильтр.

Необходимо выбрать вектор $\boldsymbol{c}$ и фильтры $\boldsymbol{B}(t-u)$, $\boldsymbol{D}(t-u)$ так, чтобы ряд $\boldsymbol{X}^*(t)$ был близок к ряду $\boldsymbol{X}(t)$.

Эта задача может рассматриваться и как отыскание способа построить такой -мерный ряд $\boldsymbol{\Psi}(t)$, который несет значительную часть информации об исходном ряде $\boldsymbol{X}(t)$. И в этой связи Бриллинджер [55, с. 363] делает такое замечание: «Показатели применяются для того, чтобы измерить изменения величины, не доступной прямому наблюдению, но о которой известно, что она оказывает определенное воздействие на другие величины, непосредственно наблюдаемые. Причем хотя само это воздействие вызывает либо одновременное увеличение, либо уменьшение всех наблюдаемых величин, его эффект маскируется действием многих других причин, по-разному влияющих на отдельные наблюдаемые».

Поэтому введенный ряд $\boldsymbol{\Psi}(t)$ может выступать в роли ряда показателей, описывающих воздействие на $\boldsymbol{X}(t)$ со стороны некоторого скрытого от наблюдателя ряда. При этом предполагается, что ряд $\boldsymbol{\Psi}(t)$ наилучшим среди -мерных рядов образом позволяет восстановить $\boldsymbol{X}(t)$ с помощью линейных операций, инвариантных во времени.

Если ввести ряд $\boldsymbol{\varepsilon}(t)$, описывающий ошибку или искажение передаваемой по каналам связи информации:

$$\boldsymbol{\varepsilon}(t) = \boldsymbol{X}(t) - \boldsymbol{X}^*(t); \tag{3.48}$$

тогда получим

$$\boldsymbol{X}(t) = \boldsymbol{c} + \sum_u \boldsymbol{D}(t-u)\boldsymbol{\Psi}(u) + \boldsymbol{\varepsilon}(t)\,, t = 0, \pm 1, \pm 2, \ldots\,. \tag{3.49}$$

Ряд $\boldsymbol{X}(t)$ оказывается выраженным как профильтрованный вариант ряда $\boldsymbol{\Psi}(t)$, имеющего меньшую размерность, плюс ошибка.

Такие модели находят применение, к примеру, в сейсмологии: пусть $\boldsymbol{\Psi}(u)$ представляет собой ряд импульсов, вызванных q землетрясениями, одновременно происходящих в разных местах. Ряд $\boldsymbol{X}(t)$ представляет собой сигналы, принятые r сейсмографами, и $\boldsymbol{D}(t)$ описывает переходные явления в земной коре, связанные с землетрясениями. Сейсмологов на основании фиксируемой информации о ряде $\boldsymbol{X}(t)$ и априорных знаниях о поведении земной коры во время землетрясений $\boldsymbol{D}(t)$, интересуют свойства ряда $\boldsymbol{\Psi}(u)$.

Укажем на возможность получения оценок параметров модели (3.47), (3.46) на основании конечной выборки наблюдений. Пусть имеется $\boldsymbol{X}$ - случайная матрица с r компонентами и некоторым числом наблюдений, имеющая среднее значение $\boldsymbol{c_X}$ и ковариационную матрицу $\boldsymbol{C_{XX}}$. Поставим задачу одновременной минимизации всех собственных значений симметричной матрицы

$$\mathcal{M}\{\,(\boldsymbol{X}-\boldsymbol{c}-\boldsymbol{DBX})(\boldsymbol{X}-\boldsymbol{c}-\boldsymbol{DBX})^T\} \tag{3.50}$$

за счет выбора вектора $\boldsymbol{c}$ с r компонентами, $(q\times r)$-матрицы $\boldsymbol{B}$ и $(r\times q)$-матрицы $\boldsymbol{D}$**.** Определив соответствующие величины $\boldsymbol{c},\boldsymbol{B},\boldsymbol{D}$, можно убедиться, что они доставляют минимум также монотонным функциям от собственных чисел матрицы (3.50) таким, как след, детерминант и диагональные элементы.

Поскольку любая $r\times r$ матрица $\boldsymbol{A}$ ранга $q\le r$ может быть представлена в виде произведения $\boldsymbol{DB}$, в котором матрица $\boldsymbol{B}$ имеет размер $q\times r$, а матрица $\boldsymbol{C}-r\times q$, то, тем самым, отыскивая $\boldsymbol{B}$ и $\boldsymbol{C}$, будет одновременно найдена матрица $\boldsymbol{A}$ ранга, не превосходящего q, которая минимизирует собственные числа матрицы [55, с. 98]

$$\mathcal{M}\{\,(\boldsymbol{X}-\boldsymbol{c}-\boldsymbol{AX})(\boldsymbol{X}-\boldsymbol{c}-\boldsymbol{AX})^T\}. \tag{3.51}$$

Доказано [55, с. 364], что все собственные числа матрицы (3.50) минимальны, когда параметры модели имеют вид

$$\boldsymbol{B}=\begin{pmatrix}\boldsymbol{q}_1^T\\ \vdots\\ \boldsymbol{q}_q^T\end{pmatrix}, \tag{3.52}$$

$$\boldsymbol{D}=(\boldsymbol{q}_1\quad\cdots\quad\boldsymbol{q}_q)=\boldsymbol{B}^T \tag{3.53}$$

и

$$\boldsymbol{c}=\boldsymbol{c_X}-\boldsymbol{DBc_X}, \tag{3.54}$$

где $\boldsymbol{q}_j^T$ есть -й собственный вектор матрицы $\boldsymbol{C_{XX}}, j=1,2,\dots,r$. Если собственные значения матрицы $\boldsymbol{C_{XX}}$ обозначить как $\lambda_j, j=1,2,\dots,r$, то при выборе значений параметров модели, равных (3.52), (3.53) и (3.54), матрица (3.50) примет наименьшее значение, равное величине

$$\textstyle\sum_{j>q}\lambda_j\,\boldsymbol{q}_j\boldsymbol{q}_j^T. \tag{3.55}$$

Величина

$$\boldsymbol{\psi}_j=\boldsymbol{q}_j^T\boldsymbol{X} \tag{3.56}$$

называется -й главной компонентой матрицы $\boldsymbol{X}, j=1,2,\dots,r$. Таким образом, в модели (3.49) ряд $\boldsymbol{X}(t)$ определяется посредством своих q главных компонент.

Поскольку собственные вектора симметричных матриц – ортогональны, то и главные компоненты, определяемых в анализе многомерных временных рядов как (3.56), также являются ортогональными по отношению друг к другу.

Таким образом, при частотном подходе, делающем упор на исследование спектральных характеристик многомерных временных рядов, наблюдаемый ряд аппроксимируется с помощью своих главных компонент. Следует отметить, что в отличие от классического факторного анализа, в котором главной компонентой матрицы считается ее собственные векторы, в анализе многомерных временных рядов главная компонента – это произведение соответствующего собственного вектора на саму матрицу.

В подходе, связанном с анализом зависимости от времени, распространены те же типы моделей, что и в одномерном случае, а именно: авторегрессия, модель скользящего среднего и смешанная модель. Схемы построения моделей имеют сходную с одномерным вариантом структуру конструкции зависимостей. Так для авторегрессии в многомерном случае имеется r рядов и у каждого из них значение в момент времени t связано не только с его собственными значениями в предыдущие моменты времени, но и с предшествующими значениями других рядов. В каждом из r соотношений имеется остаточный член ε, который может быть разным для различных рядов. В таком случае для описания авторегрессионных рядов порядка k необходимо $p+kp^2$ констант-коэффициентов модели, даже если ε независимы и неавтокоррелированы, что, очевидно, приводит к чрезмерному количеству констант, когда p и k большие.

Для формирования общего вида модели рассмотрим, к примеру, три ряда $x_{1,t}, x_{2,t}, x_{3,t}$. Тогда полностью уравнения можно записать в виде:

$$\begin{cases} \alpha_{01}x_{1,t} + \alpha_{11}x_{1,t-1} + \alpha_{21}x_{2,t-1} + \alpha_{31}x_{3,t-1} = \varepsilon_{1t}; \\ \alpha_{02}x_{2,t} + \alpha_{12}x_{1,t-1} + \alpha_{22}x_{2,t-1} + \alpha_{32}x_{3,t-1} = \varepsilon_{2t}; \\ \alpha_{03}x_{3,t} + \alpha_{13}x_{1,t-1} + \alpha_{23}x_{2,t-1} + \alpha_{33}x_{3,t-1} = \varepsilon_{3t}; \end{cases} \tag{3.57}$$

или, используя оператор сдвига назад,

$$\begin{cases} (\alpha_{01} + \alpha_{11}B)x_{1,t} + \alpha_{21}Bx_{2,t} + \alpha_{31}Bx_{3,t} = \varepsilon_{1t}; \\ \alpha_{12}Bx_{1,t} + (\alpha_{02} + \alpha_{22}B)x_{2,t} + \alpha_{32}Bx_{3,t} = \varepsilon_{2t}; \\ \alpha_{13}Bx_{1,t} + \alpha_{23}Bx_{2,t} + (\alpha_{03} + \alpha_{33}B)x_{3,t} = \varepsilon_{3t}. \end{cases} \tag{3.58}$$

В матричной форме система уравнений (3.58) принимает вид:

$$(\boldsymbol{A_0} + \boldsymbol{A_1}B)\boldsymbol{x}_t = \boldsymbol{\varepsilon}_t, \tag{3.59}$$

где матрицы имеют вид

$$\boldsymbol{A_0} = \begin{pmatrix} \alpha_{01} & 0 & 0 \\ 0 & \alpha_{02} & 0 \\ 0 & 0 & \alpha_{03} \end{pmatrix};$$

$$\boldsymbol{A_1} = \begin{pmatrix} \alpha_{11} & \alpha_{21} & \alpha_{31} \\ \alpha_{12} & \alpha_{22} & \alpha_{32} \\ \alpha_{13} & \alpha_{23} & \alpha_{33} \end{pmatrix};$$

векторы – матрицы столбцы: $\boldsymbol{x}_t$ – вектор наблюдаемых рядов

$$\boldsymbol{x}_t = \begin{pmatrix} x_{1,t} \\ x_{2,t} \\ x_{3,t} \end{pmatrix};$$

$\boldsymbol{\varepsilon}_t$ – вектор случайных отклонений или остатков:

$$\boldsymbol{\varepsilon}_t = \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \\ \varepsilon_{3,t} \end{pmatrix}.$$

Уравнение (3.59) представляет собой модель авторегрессии первого порядка трех рядов исследуемых переменных.

Модель авторегрессии второго порядка, по аналогии с (3.59) можно записать в следующем виде:

$$(\boldsymbol{A_0} + \boldsymbol{A_1}B + \boldsymbol{A_2}B^2)\boldsymbol{x}_t = \boldsymbol{\varepsilon}_t, \tag{3.60}$$

где матрица коэффициентов $\boldsymbol{A_2}$ имеет вид, аналогичный матрице коэффициентов $\boldsymbol{A_1}$, а оператор B^2 производит сдвиг на два шага назад.

И общую авторегрессионную модель произвольного порядка можно записать как

$$\boldsymbol{A}(B)\boldsymbol{x}_t = \boldsymbol{\varepsilon}_t, \tag{3.61}$$

где

$$\boldsymbol{A}(B) = \sum_{m=0}^{p} \boldsymbol{A}_m B^m, \tag{3.62}$$

p – порядок авторегрессии и $\boldsymbol{A}$ являются матрицами коэффициентов размерности $r \times r$.

Формальное решение разностного уравнения (3.61) можно записать в виде (систематические члены предполагаются затухающими)

$$\boldsymbol{x}_t = \boldsymbol{A}^{-1}(B)\boldsymbol{\varepsilon}_t. \tag{3.63}$$

Обратная матрица $\boldsymbol{A}^{-1}$ в общем случае будет равна отношению двух полиномов от B. Разложение единицы, деленной на знаменатель, и умножение на числитель дает бесконечный ряд постоянных коэффициентов, имеющих характер весов при усреднении ε на бесконечном интервале.

Аналогично через матрицы коэффициентов, обозначаемых как $\boldsymbol{M}$, можно записать процесс скользящего среднего, а именно:

$$\boldsymbol{x}_t = \boldsymbol{M}(B)\boldsymbol{\varepsilon}_t, \tag{3.64}$$

где

$$\boldsymbol{M}(B) = \sum_{m=0}^{p} \boldsymbol{M}_m B^m. \tag{3.65}$$

Смешанный процесс многомерного временного ряда имеет вид:

$$\boldsymbol{A}(B)\boldsymbol{x}_t = \boldsymbol{M}(B)\boldsymbol{\varepsilon}_t. \tag{3.66}$$

Предположим, что все ε имеют нулевое математическое ожидание и дисперсию σ^2. Тогда ковариационной матрицей между наблюдениями, расположенных на расстоянии в m шагов друг от друга, будет матрица

$$\boldsymbol{C}_m = \mathcal{M}\{\boldsymbol{x}_t \boldsymbol{x}_{t-m}^T\} \quad (3.67)$$

и после подстановки выражения (3.64) для процесса скользящего среднего в выражение для ковариационной матрицы (3.67) оказывается, что $\boldsymbol{C}_m = \sigma^2 \sum_{m,s} \boldsymbol{M}_{m+s} \boldsymbol{M}_s^T =$ коэфф. при B^m в $\left(\sum \boldsymbol{M}_j B^j\right)\left(\sum \boldsymbol{M}_j^T B^{-j}\right)$, (3.68)

Аналогично для авторегрессионного процесса

$$\boldsymbol{C}_m = \text{коэфф. при } B^m \text{ в } \boldsymbol{A}^{-1}(B)(\boldsymbol{A}^{-1}(B^{-1}))^T, \quad (3.69)$$

а для смешанного процесса

$$\boldsymbol{C}_m = \text{коэфф. при } B^m \text{ в } \boldsymbol{A}^{-1}(B)\boldsymbol{M}(B)\boldsymbol{M}^T(B^{-1})(\boldsymbol{A}^{-1}(B^{-1}))^T. \quad (3.70)$$

Как заметил М. Кендэл [50, с. 145]: «Решить эти уравнения можно, но сделать это нелегко», к тому же эти уравнения могут иметь не единственное решение, что порождает проблему неидентифицируемости. Корень этой проблемы, как отмечал М. Кендэл [50, с. 145], лежит в скользящем среднем, а не в авторегрессионной части модели. При этих обстоятельствах он рекомендует отказаться от компоненты, представленной в модели скользящим средним.

Помимо частотного и временного подходов к анализу многомерных временных рядов, широкое распространение получил, условно говоря, смешанный подход, известный под названием «Фильтр Калмана». Модели фильтрации (а модель авторегрессии, как отмечалось выше, может быть сведена к таковой), служащие для выделения полезной информации из шума, основаны на методах Фурье и в соответствии с этим предполагали стационарность или нечто, достаточно близкое к ней. Подобные методы приводят к описанию фильтра в терминах частотной характеристики, и для того, чтобы получить импульсную переходную функцию, необходим следующий шаг – разложение Фурье. При этом исследуемые признаки должны иметь характер непрерывных величин, что достаточно удобно для построения фильтра с применением аналоговых вычислительных машин. Однако при использовании современной вычислительной техники частотная характеристика не является удобным средством описания фильтра.

Еще два фактора подводят к необходимости модификации модели многомерных временных рядов. Во-первых, на практике спектры имеют не столь общий характер, как это может показаться из научной литературы [56]. Более того, они либо даются исходной физической теорией (кроме констант, подлежащих оценке), либо получаются в результате оценки по прошлым данным. Второй фактор связан с возможностью нестационарности. Одной из причин нестационарности может быть тот факт, что для приближения любой системы к стационарному состоянию требуется определенное время. Если же система наблюдается с момента, когда в ней происходят некоторые переходные процессы, то стационарная модель будет неподходящей. К тому же, системе по ее физической природе может быть присуща существенная нестационарность. Все эти проблемы в значительной мере преодолеваются при использовании методологии фильтрации Калмана.

Рассмотрим динамическую систему, чье положение в момент времени t_k описывается линейным векторным разностным уравнением

$$\boldsymbol{x}_k = \boldsymbol{\Phi}_{k,k-1}\boldsymbol{x}_{k-1} + \boldsymbol{\Gamma}_{k,k-1}\boldsymbol{u}_{k-1} + \boldsymbol{f}_{k-1} + \boldsymbol{w}_{k-1}, \quad (3.71)$$

где $\boldsymbol{x}_k$ – вектор, размерностью n, описывающий состояние системы (многомерный временной ряд);

$\boldsymbol{\Phi}_{k,k-1}$– матрица, размером $n \times n$, переходов системы из состояния в момент времени t_{k-1} в состояние, соответствующее моменту времени t_k;

$\boldsymbol{u}_{k-1}$ – вектор контроля за системой на полуинтервале времени $[t_{k-1}, t_k)$, размерностью p;

$\boldsymbol{\Gamma}_{k,k-1}$ – известная матрица, размером $n \times p$, связанная с контролем состояния системы, благодаря которому вектор контроля изменяется от одного момента времени к другому;

$\boldsymbol{f}_{k-1}$ – вектор, порядка p, силовой функции, представляющей собой известную функцию времени;

$\boldsymbol{w}_{k-1}$ – вектор, порядка p, случайной последовательности (вектор случайных отклонений) с известными свойствами:

$\mathcal{M}\{\boldsymbol{w}_k\} = \boldsymbol{0}$, для всех k;

$$\mathcal{M}\{\boldsymbol{w}_k \boldsymbol{w}_j^T\} = \delta_{kj} \boldsymbol{Q}_k,$$

где δ_{kj} – дельта Кронекера $\left(\delta_{kj} = \begin{cases} 1, & \text{если } k = j \\ 0, & \text{если } k \neq j \end{cases}\right)$; $\boldsymbol{Q}_k$ – неотрицательно определенная матрица ковариаций между случайными отклонениями в момент времени t_k.

Матрица переходов $\boldsymbol{\Phi}_{k,k-1}$ удовлетворяет следующим свойствам:

$$\boldsymbol{\Phi}_{k,k} = \boldsymbol{I}, \text{ для всех } k$$

$$\boldsymbol{\Phi}_{k,j} \boldsymbol{\Phi}_{j,i} = \boldsymbol{\Phi}_{k,i}. \tag{3.72}$$

Данные свойства приводят к тому, что

$$\boldsymbol{\Phi}_{k,j}^T = \boldsymbol{\Phi}_{j,k} \quad . \tag{3.73}$$

Важным моментом в модели фильтра Калмана является начальное состояние системы $\boldsymbol{x}_0$, которое рассматривается как вектор случайной переменной с известными свойствами

$$\mathcal{M}\{\boldsymbol{x}_0\} = \boldsymbol{0}, \ \mathcal{M}\{\boldsymbol{x}_0 \boldsymbol{x}_0^T\} = \boldsymbol{M}_0 \tag{3.74}$$

и

$$\mathcal{M}\{\boldsymbol{w}_k \boldsymbol{x}_0^T\} = 0 \ \text{ для всех } k. \tag{3.75}$$

Далее предполагается, что в каждый момент времени t_k имеется m измерений, составляющих вектор $\boldsymbol{z}_k$, который линейно связан с вектором состояния системы и искажен аддитивным шумом:

$$\boldsymbol{z}_k = \boldsymbol{H}_k \boldsymbol{x}_k + \boldsymbol{v}_k, \tag{3.76}$$

где $\boldsymbol{H}_k$ – известная матрица измерений, размером $m \times n$;

$\boldsymbol{v}_k$ – вектор, порядка m, аддитивной случайной последовательности, обладающий следующими свойствами

$$\mathcal{M}\{\boldsymbol{v}_k\} = \boldsymbol{0}, \text{ для всех } k;$$

$$\mathcal{M}\{\boldsymbol{v}_k \boldsymbol{v}_j^T\} = \delta_{kj} \boldsymbol{R}_k, \tag{3.77}$$

где $\boldsymbol{R}_k$ – неотрицательно определенная матрица ковариации аддитивного шума в момент времени t_k.

Далее предполагается, что случайные последовательности $\boldsymbol{v}_k$ и $\boldsymbol{w}_k$ – некоррелированы. Эти последовательности являются последовательностями «белого шума» с известным свойством:

$$\mathcal{M}\{\boldsymbol{v}_k \boldsymbol{w}_j^T\} = \boldsymbol{0}, \text{для всех } k, j. \tag{3.78}$$

Кроме того, случайная последовательность из уравнения (3.76) также не коррелирует с вектором начального состояния:

$$\mathcal{M}\{\boldsymbol{v}_k \boldsymbol{x}_0^T\} = \boldsymbol{0}, \text{для всех } k. \tag{3.79}$$

После того, как сконструирована модель фильтра Калмана, а она включает два уравнения: (3.71) и (3.76), а также все сформулированные выше предположения относительно основных свойств параметров и случайных величин, можно поставить задачу рекурсивного линейного оценивания.

На основании модели фильтра Калмана определить оценку $\hat{\boldsymbol{x}}_k$ состояния системы в момент времени t_k такую, что она является линейной комбинацией оценки состояния системы в момент времени $\hat{\boldsymbol{x}}_{k-1}$ и измерения данных $\boldsymbol{z}_k$. Оценка должна быть «лучшей» в смысле того, что ожидаемая величина суммы квадратов ошибок при оценивании является минимальной. Таким образом, оценка $\hat{\boldsymbol{x}}_k$ должна быть выбрана так, чтобы

$$\mathcal{M}\{(\hat{\boldsymbol{x}}_k - \boldsymbol{x}_k)^T (\hat{\boldsymbol{x}}_k - \boldsymbol{x}_k)\} = minimum$$

Калман использовал свою модель, главным образом, для решения проблем управления динамическими системами в техническом приложении. Поэтому первое уравнение модели (3.71) имеет силовые компоненты, имеющие отношение к управлению системой. Для решения задач исследования систем, в которых отсутствует силовая составляющая, используется более простая модификация модели.

Пусть динамика системы описывается однородным линейным разностным уравнением

$$\boldsymbol{x}_k = \boldsymbol{\Phi}_{k,k-1} \boldsymbol{x}_{k-1}. \tag{3.80}$$

Непосредственно вектор состояния системы $\boldsymbol{x}_k$ не наблюдается, но имеется вектор измерений $\boldsymbol{z}_k$ (3.76):

$$\boldsymbol{z}_k = \boldsymbol{H}_k \boldsymbol{x}_k + \boldsymbol{v}_k.$$

Таким образом, упрощенная модель фильтра Калмана описывается уравнениями (3.80) и (3.76).

Первое, форма уравнения линейной оценки состояния может быть установлена на основании физических характеристик системы. Состояние системы развивается на основании уравнения (3.80) так, что имея оценку состояния системы $\hat{\boldsymbol{x}}_{k-1}$ в момент времени t_k, можно спрогнозировать оценку для момента времени t_k на основании выражения:

$$\hat{\boldsymbol{x}}_k{}' = \boldsymbol{\Phi}_{k,k-1}\hat{\boldsymbol{x}}_{k-1}, \tag{3.81}$$

если никакая другая информация недоступна. Измерение в момент времени t_k может быть использовано для улучшения оценки. Основываясь на $\hat{\boldsymbol{x}}_k{}'$ и (3.76), можно было бы ожидать, что изменение величин в момент времени t_k был бы равен $\boldsymbol{H}_k\hat{\boldsymbol{x}}_k{}'$. Ошибка при оценивании непосредственно отражается на ошибке ожидаемой величине изменений:

$$\boldsymbol{e}_k = \boldsymbol{z}_k - \boldsymbol{H}_k\boldsymbol{\Phi}_{k,k-1}\hat{\boldsymbol{x}}_{k-1}. \tag{3.82}$$

В соответствии с поставленной задачей рекурсивного оценивания, оценка вектора состояния системы должна быть линейной относительно новых измерений. Определим неизвестную матрицу $\boldsymbol{K}_k$ так, чтобы оценка $\hat{\boldsymbol{x}}_k$ была найдена на основании выражения:

$$\hat{\boldsymbol{x}}_k = \boldsymbol{\Phi}_{k,k-1}\hat{\boldsymbol{x}}_{k-1} + \boldsymbol{K}_k\left(\boldsymbol{z}_k - \boldsymbol{H}_k\boldsymbol{\Phi}_{k,k-1}\hat{\boldsymbol{x}}_{k-1}\right). \tag{3.83}$$

Матрица $\boldsymbol{K}_k$ должна быть определена так, чтобы значение выражения

$$\mathcal{M}\{(\hat{\boldsymbol{x}}_k - \boldsymbol{x}_k)^T(\hat{\boldsymbol{x}}_k - \boldsymbol{x}_k)\}$$

было минимальным, при этом она может рассматриваться как весовая матрица.

Обозначим разность между оценкой состояния системы и ее истинным значением как:

$$\tilde{\boldsymbol{x}}_k = \hat{\boldsymbol{x}}_k - \boldsymbol{x}_k,$$

таким образом,

$$\mathcal{M}\{(\hat{\boldsymbol{x}}_k - \boldsymbol{x}_k)^T(\hat{\boldsymbol{x}}_k - \boldsymbol{x}_k)\} = \mathcal{M}\{\tilde{\boldsymbol{x}}_k{}^T\tilde{\boldsymbol{x}}_k\}. \tag{3.84}$$

Очевидно, что выражение (3.84) может быть переписано следующим образом:

$$\mathcal{M}\{\tilde{\boldsymbol{x}}_k{}^T\tilde{\boldsymbol{x}}_k\} = tr[\mathcal{M}\{\tilde{\boldsymbol{x}}_k\tilde{\boldsymbol{x}}_k{}^T\}],$$

где $tr[\cdot]$ – след матрицы, т.е. оператор взятия суммы диагональных элементов матрицы, стоящей в квадратных скобках. Определим матрицу $\boldsymbol{P}_k$ как

$$\boldsymbol{P}_k = \mathcal{M}\{\tilde{\boldsymbol{x}}_k\tilde{\boldsymbol{x}}_k{}^T\}. \tag{3.85}$$

Теперь получим выражение для $\tilde{\boldsymbol{x}}_k$:

$$\begin{aligned}\tilde{\boldsymbol{x}}_k &= \left(\boldsymbol{\Phi}_{k,k-1}\hat{\boldsymbol{x}}_{k-1} + \boldsymbol{K}_k\left(\boldsymbol{z}_k - \boldsymbol{H}_k\boldsymbol{\Phi}_{k,k-1}\hat{\boldsymbol{x}}_{k-1}\right)\right) = \\ &= \boldsymbol{\Phi}_{k,k-1}\hat{\boldsymbol{x}}_{k-1} - \boldsymbol{K}_k\boldsymbol{H}_k\boldsymbol{\Phi}_{k,k-1}\hat{\boldsymbol{x}}_{k-1} + \boldsymbol{K}_k(\boldsymbol{H}_k\boldsymbol{x}_k + \boldsymbol{v}_k) = \\ &= (\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{H}_k)\boldsymbol{\Phi}_{k,k-1}\tilde{\boldsymbol{x}}_{k-1} + \boldsymbol{K}_k\boldsymbol{v}_k.\end{aligned} \tag{3.86}$$

Выражение (3.86) можно подставить в (3.85) и получить выражение для матрицы $\boldsymbol{P}_k$

$$\boldsymbol{P}_k = \mathcal{M}\Big\{\Big((\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{H}_k)\boldsymbol{\Phi}_{k,k-1}\tilde{\boldsymbol{x}}_{k-1} + \boldsymbol{K}_k\boldsymbol{v}_k\Big)\Big((\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{H}_k)\boldsymbol{\Phi}_{k,k-1}\tilde{\boldsymbol{x}}_{k-1} +$$

$$+\boldsymbol{K}k\boldsymbol{v}kT=\boldsymbol{I}-\boldsymbol{K}k\boldsymbol{H}k\boldsymbol{\Phi}k,k-1\mathcal{M}\boldsymbol{x}k-1\boldsymbol{x}k-1T\boldsymbol{\Phi}k,k-1T\boldsymbol{I}-\boldsymbol{K}k\boldsymbol{H}kT+$$

$$\begin{aligned}&+\boldsymbol{K}_k\mathcal{M}\{\boldsymbol{v}_k\tilde{\boldsymbol{x}}_{k-1}{}^T\}\boldsymbol{\Phi}_{k,k-1}^T(\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{H}_k)^T + \\ &+(\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{H}_k)\boldsymbol{\Phi}_{k,k-1}\mathcal{M}\{\tilde{\boldsymbol{x}}_{k-1}\boldsymbol{v}_k{}^T\}\boldsymbol{K}_k^T + \boldsymbol{K}_k\mathcal{M}\{\boldsymbol{v}_k\boldsymbol{v}_k{}^T\}\boldsymbol{K}_k{}^T.\end{aligned} \tag{3.87}$$

По определению и по аналогии с (3.85), определим матрицу

$$\boldsymbol{P}_{k-1} = \mathcal{M}\{\tilde{\boldsymbol{x}}_{k-1}\tilde{\boldsymbol{x}}_{k-1}{}^T\},$$

а также укажем на матрицу из (3.77):

$$\mathcal{M}\{\boldsymbol{v}_k\boldsymbol{v}_k^T\} = \boldsymbol{R}_k.$$

Рассматривая также следующие свойства введенных величин, а именно

$$\mathcal{M}\{\boldsymbol{v}_k\tilde{\boldsymbol{x}}_{k-1}{}^T\} = \mathcal{M}\{\tilde{\boldsymbol{x}}_{k-1}\boldsymbol{v}_k{}^T\} = \boldsymbol{0},$$

$$\mathcal{M}\{\boldsymbol{v}_k\boldsymbol{v}_{k-1}{}^T\} = \boldsymbol{0},$$

и

$$\mathcal{M}\{\boldsymbol{v}_k\boldsymbol{x}_0{}^T\} = \boldsymbol{0}\ ,$$

выражение (3.87) можно переписать в следующем виде:

$$\boldsymbol{P}_k = (\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{H}_k)\boldsymbol{P'}_k(\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{H}_k)^T + \boldsymbol{K}_k\boldsymbol{R}_k\boldsymbol{K}_k^T, \tag{3.88}$$

где использовано обозначение матрицы

$$\boldsymbol{P'}_k = \boldsymbol{\Phi}_{k,k-1}\boldsymbol{P}_{k-1}\boldsymbol{\Phi}_{k,k-1}^T. \quad (3.89)$$

Раскрывая скобки в (3.88) и после небольших преобразований, получаем

$$\boldsymbol{P}_k = \boldsymbol{P'}_k - \boldsymbol{K}_k\boldsymbol{H}_k\boldsymbol{P'}_k - \boldsymbol{P'}_k\boldsymbol{H}_k^T\boldsymbol{K}_k^T + \boldsymbol{K}_k\left(\boldsymbol{H}_k\boldsymbol{P'}_k\boldsymbol{H}_k^T + \boldsymbol{R}_k\right)\boldsymbol{K}_k^T. \quad (3.90)$$

Матрица $\boldsymbol{P'}_k$ не зависит от $\boldsymbol{K}_k$, поэтому на нее не оказывает эффект выбор матрицы $\boldsymbol{K}_k$. Матрица $\left(\boldsymbol{H}_k\boldsymbol{P'}_k\boldsymbol{H}_k^T + \boldsymbol{R}_k\right)$ является симметричной и неотрицательно-определенной, так что она может быть переписана как результат произведения матрицы $\boldsymbol{S}_k$ на себя транспонированную, т.е. $\boldsymbol{S}_k$– матрица корня квадратного:

$$\boldsymbol{S}_k\boldsymbol{S}_k^T = \boldsymbol{H}_k\boldsymbol{P'}_k\boldsymbol{H}_k^T + \boldsymbol{R}_k. \quad (3.91)$$

Заметим, что последние три члена выражения (3.90) имеют форму квадратичного матричного полинома относительно неизвестной матрицы $\boldsymbol{K}_k$. Подставим (3.91) в выражение (3.90) и предположим существование матрицы $\boldsymbol{A}_k$ такой, что выполняется тождество

$$\boldsymbol{P}_k = \boldsymbol{P'}_k + (\boldsymbol{K}_k\boldsymbol{S}_k - \boldsymbol{A}_k)(\boldsymbol{K}_k\boldsymbol{S}_k - \boldsymbol{A}_k)^T - \boldsymbol{A}_k\boldsymbol{A}_k^T. \quad (3.92)$$

Эта процедура является матричным эквивалентом извлечения корня из квадратичного полинома. Предполагая, что $\boldsymbol{S}_k\boldsymbol{S}_k^T$ – положительно-определенная матрица, а это непосредственно приводит к выражению

$$\boldsymbol{A}_k = \boldsymbol{P'}_k\boldsymbol{H}_k^T(\boldsymbol{S}_k^{-1})^T. \quad (3.93)$$

Когда матрица $\boldsymbol{S}_k$ – сингулярная, то вместо обратной матрицы используется процедура псевдообращения матрицы. В выражении (3.93) только один член – результат произведения – зависит от весовой матрицы $\boldsymbol{K}_k$. Произведение матрицы на себя транспонированную – матрица неотрицательно-определенная, поэтому след матрицы $\boldsymbol{P}_k$ минимизируется посредством выбора

$$\boldsymbol{K}_k\boldsymbol{S}_k = \boldsymbol{P'}_k\boldsymbol{H}_k^T(\boldsymbol{S}_k^{-1})^T,$$

таким образом, оптимальная весовая матрица должна иметь вид:

$$\boldsymbol{K}_k = \boldsymbol{P'}_k\boldsymbol{H}_k^T\left(\boldsymbol{H}_k\boldsymbol{P'}_k\boldsymbol{H}_k^T + \boldsymbol{R}_k\right)^{-1}. \quad (3.94)$$

Подставляя (3.94) и (3.93) в выражение (3.92), получаем окончательное выражение для матрицы $\boldsymbol{P}_k$:

$$\boldsymbol{P}_k = \boldsymbol{P'}_k - \boldsymbol{K}_k\boldsymbol{H}_k\boldsymbol{P'}_k. \quad (3.95)$$

Уравнения (3.83), (3.89), (3.94) и (3.95) составляют фильтр Калмана для модели (3.80) и (3.76).

Расширим модель, включив в нее детерминистическую и случайную силовые составляющие:

$$\boldsymbol{x}_k = \boldsymbol{\Phi}_{k,k-1}\boldsymbol{x}_{k-1} + \boldsymbol{f}_{k-1} + \boldsymbol{w}_{k-1}, \quad (3.96)$$

в этой модели отсутствует компонента контроля.

Предположим, что оценка состояния системы основана только на оценке в момент времени t_{k-1}, т.е. никаких измерений в момент времени t_k не производится. Вектор шума $\boldsymbol{w}_k$ не зависит от состояния в момент t_k и нулевое математическое ожидание, таким образом вполне ожидаемо, что он не оказывает влияние на оценку состояния в момент времени t_k. Также, $\boldsymbol{f}_{k-1}$ – известная векторная функция, которая действует на полуинтервале $[t_{k-1}, t_k)$, поэтому оценка может быть получена на основании выражения

$$\hat{\boldsymbol{x}}'_k = \boldsymbol{\Phi}_{k,k-1}\hat{\boldsymbol{x}}_{k-1} + \boldsymbol{f}_{k-1}. \quad (3.97)$$

Покажем , что эта оценка является наилучшей с точки зрения критерия минимума наименьших квадратов.

В теории метода наименьших квадратов известно утверждение, согласно которому наилучшей оценкой случайного вектора $\boldsymbol{x}$ на основании известных величин $\boldsymbol{z}_1, \boldsymbol{z}_2, \dots, \boldsymbol{z}_q$, если известна плотность их совместного распределения $p\left(\boldsymbol{x}, \boldsymbol{z}_1, \boldsymbol{z}_2, \dots, \boldsymbol{z}_q\right)$, является условное математическое ожидание случайной величины $\boldsymbol{x}$ при наличии заданных значений $\boldsymbol{z}_1, \boldsymbol{z}_2, \dots, \boldsymbol{z}_q$:

$$\hat{\boldsymbol{x}} = \mathcal{M}\{\, \boldsymbol{x} | \boldsymbol{z}_1, \boldsymbol{z}_2, \dots, \boldsymbol{z}_{k-1}\}.$$

Таким образом, наилучшей оценкой неизвестного вектора состояний $\boldsymbol{x}_k$ на основании измерений $\boldsymbol{z}_1, \boldsymbol{z}_2, \dots, \boldsymbol{z}_{k-1}$ является условное математическое ожидание

$$\hat{\boldsymbol{x}}'_k = \mathcal{M}\{\, \boldsymbol{x}_k | \boldsymbol{z}_1, \boldsymbol{z}_2, \dots, \boldsymbol{z}_{k-1}\} =$$

$$= \mathcal{M}\left\{ \left(\boldsymbol{\Phi}_{k,k-1}\boldsymbol{x}_{k-1} + \boldsymbol{f}_{k-1} + \boldsymbol{w}_{k-1}\right) \middle| \boldsymbol{z}_1, \boldsymbol{z}_2, \dots, \boldsymbol{z}_{k-1} \right\} =$$
$$= \boldsymbol{\Phi}_{k,k-1}\mathcal{M}\{\boldsymbol{x}_{k-1}|\boldsymbol{z}_1, \boldsymbol{z}_2, \dots, \boldsymbol{z}_{k-1}\} + \mathcal{M}\{\boldsymbol{f}_{k-1}|\boldsymbol{z}_1, \boldsymbol{z}_2, \dots, \boldsymbol{z}_{k-1}\} +$$
$$+\mathcal{M}\{\boldsymbol{w}_{k-1}|\boldsymbol{z}_1, \boldsymbol{z}_2, \dots, \boldsymbol{z}_{k-1}\}. \tag{3.98}$$

Первое математическое ожидание в (3.98) может быть найдено по определению:

$$\mathcal{M}\{\boldsymbol{x}_{k-1}|\boldsymbol{z}_1, \boldsymbol{z}_2, \dots, \boldsymbol{z}_{k-1}\} = \widehat{\boldsymbol{x}}_{k-1}. \tag{3.99}$$

$\boldsymbol{f}_{k-1}$ – не случайный вектор, поэтому

$$\mathcal{M}\{\boldsymbol{f}_{k-1}|\boldsymbol{z}_1, \boldsymbol{z}_2, \dots, \boldsymbol{z}_{k-1}\} = \boldsymbol{f}_{k-1}. \tag{3.100}$$

В основных предположениях модели фильтра Калмана отмечалось, что вектор шума $\boldsymbol{w}_{k-1}$ не зависит от состояний системы для всех моментов времени, и также не зависит от погрешностей измерений, поэтому:

$$\mathcal{M}\{\boldsymbol{w}_{k-1}|\boldsymbol{z}_1, \boldsymbol{z}_2, \dots, \boldsymbol{z}_{k-1}\} = \boldsymbol{0}. \tag{3.101}$$

Подставляя выражения (3.99), (3.100) и (3.101) в уравнение (3.98), получаем наилучшую оценку состояния системы, основанной на оценке в момент t_{k-1}:

$$\widehat{\boldsymbol{x}}'_k = \boldsymbol{\Phi}_{k,k-1}\widehat{\boldsymbol{x}}_{k-1} + \boldsymbol{f}_{k-1},$$

таким образом, получена оценка (3.97).

Предположим, что произведено измерение в момент времени t_k. Аналогично предположениям, которые были произведены для оценки (3.83), уточненную оценку определим следующим образом:

$$\widehat{\boldsymbol{x}}_k = \widehat{\boldsymbol{x}}'_k + \boldsymbol{K}_k(\boldsymbol{z}_k - \boldsymbol{H}_k\widehat{\boldsymbol{x}}'_k), \tag{3.102}$$

где $\boldsymbol{K}_k$ − неизвестная весовая матрица. Рассмотрим матрицу $\widetilde{\boldsymbol{x}}_k$

$$\widetilde{\boldsymbol{x}}_k = \widehat{\boldsymbol{x}}_k - \boldsymbol{x}_k = \left(\boldsymbol{\Phi}_{k,k-1}\widehat{\boldsymbol{x}}_{k-1} + \boldsymbol{f}_{k-1} + \boldsymbol{K}_k(\boldsymbol{z}_k - \boldsymbol{H}_k\widehat{\boldsymbol{x}}'_k)\right) -$$
$$-\boldsymbol{\Phi}_{k,k-1}\boldsymbol{x}_{k-1} - \boldsymbol{f}_{k-1} - \boldsymbol{w}_{k-1} =$$
$$= \boldsymbol{\Phi}_{k,k-1}\widetilde{\boldsymbol{x}}_{k-1} - \boldsymbol{K}_k\boldsymbol{H}_k\boldsymbol{\Phi}_{k,k-1}\widetilde{\boldsymbol{x}}_{k-1} + \boldsymbol{K}_k\boldsymbol{H}_k\boldsymbol{w}_{k-1} - \boldsymbol{w}_{k-1} + \boldsymbol{K}_k\boldsymbol{v}_k =$$
$$= (\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{H}_k)\left(\boldsymbol{\Phi}_{k,k-1}\widetilde{\boldsymbol{x}}_{k-1} - \boldsymbol{w}_{k-1}\right) + \boldsymbol{K}_k\boldsymbol{v}_k. \tag{3.103}$$

Рассмотрим матрицу:

$$\boldsymbol{P}_k = \mathcal{M}\{\left((\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{H}_k)\left(\boldsymbol{\Phi}_{k,k-1}\widetilde{\boldsymbol{x}}_{k-1} - \boldsymbol{w}_{k-1}\right) + \boldsymbol{K}_k\boldsymbol{v}_k\right) \cdot$$
$$\cdot \left((\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{H}_k)\left(\boldsymbol{\Phi}_{k,k-1}\widetilde{\boldsymbol{x}}_{k-1} - \boldsymbol{w}_{k-1}\right) + \boldsymbol{K}_k\boldsymbol{v}_k\right)^T\}$$

Поскольку математическое ожидание квадратичной формы имеет вид:

$$\mathcal{M}\left\{\left(\boldsymbol{\Phi}_{k,k-1}\widetilde{\boldsymbol{x}}_{k-1} - \boldsymbol{w}_{k-1}\right)\left(\boldsymbol{\Phi}_{k,k-1}\widetilde{\boldsymbol{x}}_{k-1} - \boldsymbol{w}_{k-1}\right)^T\right\} =$$
$$= \boldsymbol{\Phi}_{k,k-1}\boldsymbol{P}_{k-1}\boldsymbol{\Phi}^T_{k,k-1} + \boldsymbol{Q}_{k-1},$$

на основании статистических свойств вектора шума $\boldsymbol{w}_{k-1}$.

Определим матрицу $\boldsymbol{P'}_k$ следующим образом:

$$\boldsymbol{P'}_k = \boldsymbol{\Phi}_{k,k-1}\boldsymbol{P}_{k-1}\boldsymbol{\Phi}^T_{k,k-1} + \boldsymbol{Q}_{k-1}. \tag{3.104}$$

Таким образом, матрица $\boldsymbol{P}_k$ имеет следующий вид:

$$\boldsymbol{P}_k = (\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{H}_k)\boldsymbol{P'}_k(\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{H}_k)^T + \boldsymbol{K}_k\boldsymbol{R}_k\boldsymbol{K}^T_k, \tag{3.105}$$

и форма уравнения (3.105) ничем не отличается от уравнения (3.88), единственное отличие состоит в значении матрицы $\boldsymbol{P'}_k$. Следовательно, оптимальная весовая матрица $\boldsymbol{K}_k$ получается на основании выражения (3.94), а матрица $\boldsymbol{P}_k$ в конечном итоге может быть найдена по выражению (3.95).

Уравнения (3.102), (3.104), (3.94) и (3.95) устанавливают фильтр Калмана для модели с силовыми компонентами, принимающей вид уравнений (3.96) и (3.76). Следует отметить, что единственный эффект случайного шума $\boldsymbol{w}_{k-1}$, действующего в динамической системе, проявляется в увеличении неопределенности при оценивании благодаря аддитивному присутствию в выражении (3.104). С другой стороны, детерминистическая составляющая $\boldsymbol{f}_k$ появляется лишь в уравнении оценки (3.102). Во всем остальном модель фильтра с силовыми компонентами (3.96) ничем не отличается от модели, представленной линейным однородным разностным уравнением (3.80). Поскольку алгоритмы управления, используемые в общей модели фильтра Калмана (3.71), предназначены для технических систем, что с большим трудом может быть применено к экономическим системам, то проблемы оценивания параметров модели с компонентами управления рассматриваться не будут.

3.2. Факторный подход к исследованию экономических процессов с использованием многомерных временных рядов.

В анализе многомерных временных рядов возникает много проблем [50, с.145], некоторые, и в частности, проблема идентифицируемости к настоящему моменту времени не решены до конца. Как отметил М. Кендэл, «проблема идентифицируемости все еще ждет полного и удобного для практики решения» [50, с. 146]. Более того, механистическое обобщение одномерного случая на многомерные временные ряды, которое сводится к тому, что рассматриваются три вида моделей: авторегрессии, скользящего среднего и смешанные, − ведет к нивелированию проблемы разделения временного влияния на значение переменной в данный момент наблюдения и корреляции с другими переменными.

Традиционные подходы анализа многомерных временных рядов предполагают влияние на данную переменную в данный момент времени как ее значений в предшествующие моменты, так и значений других переменных, и в текущий момент, и в предшествующие. Для этого вводится понятие кросс-корреляции между различными переменными, причем разработанный аппарат позволяет рассматривать характеристики тесноты связи среди переменных лишь между их парами. У спектральной плотности кросс-спектра для двух рядов наряду с действительной частью появляется и мнимая часть, и в этой связи возникают еще две характеристики: когерентность – отношение амплитуды кросс-спектра к произведению спектральных плотностей каждого из двух рядов, − и, так называемое, «усиление» − произведение отношения спектральных плотностей двух рядов на когерентность.

«Интерпретация этих величин в анализе временных рядов, как мне кажется, дело более трудное, чем интерпретация кросс-коррелограммы, которая, по крайней мере, показывает силу связи меду двумя рядами при различных опережениях или запаздываниях», − отмечал М. Кендэл [50]. Таким образом, слабая интерпретабельность получаемых результатов, или ее полная невозможность, глубокая запутанность воздействий между переменными для разных моментов времени делают анализ многомерных временных рядов малопривлекательным, и в этой связи – и малораспространенным.

Главной нерешенной проблемой анализа многомерных временных рядов является невозможность отдельного учета влияния временных сдвигов и влияния других переменных. Ответ на вопрос, что привело к данному значению переменной: либо ее значение в предыдущие моменты времени, либо влияние других переменных, − является центральным в многомерном анализе временных рядов как с точки зрения выяснения характера поведения экономической системы, так и для корректного решения проблем прогнозирования. Поэтому в самодостаточной модели многомерного временного ряда должно быть разделение временного аспекта и проблемы взаимодействий между различными переменными, между которыми существует объективная связь.

Пусть имеется несколько переменных, наблюдения для каждой из которых образуют временной ряд. Естественно, что наблюдения произведены в одни и те же моменты времени и их число значительно больше числа переменных. Для построения модели сделаем основное предположение. Значение величины переменной в момент времени t непосредственно не зависит от ее значений в моменты $t-1, t-2, \dots$. Но поведение совокупности переменных, как и в обычном факторном анализе, определяется поведением факторов. Именно факторы подвержены влиянию временных сдвигов. Величина отдельного значения фактора для данного момента времени зависит только от того, какие значения фактор принимал в предшествующие моменты, и не зависит от значений других факторов, поскольку в ортогональной факторной модели факторы являются независимыми друг от друга переменными. Именно поведение каждого фактора можно рассматривать как автокорреляционный процесс.

Построим модель многомерного временного ряда с разделением проблем, связанных с временем и с корреляцией между исследуемыми признаками.

Будем рассматривать значения наблюдаемых переменных, составляющих многомерный временной ряд в виде вектора наблюдений в момент времени t:

$$\boldsymbol{z_t} = (z_{t1} \quad z_{t2} \quad \cdots \quad z_{tn}),$$

где n − число переменных. Также как в обычном факторном анализе будем рассматривать наблюдаемые значения переменных в стандартной форме, т.е. они должны быть центрированными и нормированными. Если речь идет о временных рядах, то строго говоря, у наблюдаемой переменной должен быть удален тренд, который может иметь и нелинейный характер. Однако на этом этапе исследования будем полагать, что вероятностные характеристики наблюдаемой переменной полностью описываются математическим ожиданием и дисперсией в силу основного предположения, что эта переменная непосредственно не зависит от времени.

Обозначим вектор значений факторов, определяющих поведение наблюдаемых переменных в момент времени t, как

$$\boldsymbol{f_t} = (f_{t1} \quad f_{t2} \quad \cdots \quad f_{tm}),$$

где m − число факторов.

Для получения обобщенной многомерной авторегрессионной модели факторов, рассмотрим сначала случай авторегрессии первого порядка:

$$\boldsymbol{f_t} = \boldsymbol{f_{t-1}}\boldsymbol{\Phi_1} + \boldsymbol{a_t}, \tag{3.106}$$

где $\boldsymbol{\Phi_1}$ – матрица перехода значений факторов от момента времени $(t-1)$ к t,
$\boldsymbol{a_t}$ − вектор «белого шума», или импульсов, или случайных отклонений

$$\boldsymbol{a_t} = (a_{t1} \quad a_{t2} \quad \cdots \quad a_{tm}).$$

Пусть, как и для одномерного случая, B − оператор сдвига назад, тогда выражение (3.106) можно записать в следующем виде

$$\boldsymbol{f_t} - \boldsymbol{f_{t-1}}\boldsymbol{\Phi_1} = \boldsymbol{a_t}$$

или

$$\boldsymbol{f_t} - B\boldsymbol{f_t}\boldsymbol{\Phi_1} = \boldsymbol{a_t},$$

и тогда модель (3.106) можно представить следующим образом:

$$\boldsymbol{f_t}(\boldsymbol{I} - B\boldsymbol{\Phi_1}) = \boldsymbol{a_t}. \tag{3.107}$$

В этой модели неизвестными параметрами, подлежащими оценке, являются матрица перехода на один шаг времени вперед $\boldsymbol{\Phi_1}$, а также дисперсии m компонент вектора белого шума.

Если предположить, что переход значений факторов от момента времени t к моменту времени $t-2$ определяется произведением матриц перехода, что логично следует из следующего соотношения:

$$\boldsymbol{f_{t-1}} = \boldsymbol{f_{t-2}}\boldsymbol{\Phi_1} + \boldsymbol{a_{t-1}}, \tag{3.108}$$

где, подставляя (3.108) в (3.106)

$$\boldsymbol{f_t} = (\boldsymbol{f_{t-2}}\boldsymbol{\Phi_1} + \boldsymbol{a_{t-1}})\boldsymbol{\Phi_1} + \boldsymbol{a_t},$$

получаем

$$\boldsymbol{f_t} = \boldsymbol{f_{t-2}}\boldsymbol{\Phi_1}\boldsymbol{\Phi_1} + \boldsymbol{a_{t-1}}\boldsymbol{\Phi_1} + \boldsymbol{a_t} = \boldsymbol{f_{t-2}}\boldsymbol{\Phi_1^2} + \boldsymbol{a_{t-1}}\boldsymbol{\Phi_1} + \boldsymbol{a_t}, \tag{3.109}$$

то модель авторегрессии факторов 2-го порядка имеет вид

$$\boldsymbol{f_t} = \boldsymbol{f_{t-1}}\boldsymbol{\Phi_1} + \boldsymbol{f_{t-2}}\boldsymbol{\Phi_1^2} + \boldsymbol{a_t},$$

или, используя оператор сдвига назад:

$$\boldsymbol{f_t}\left(\boldsymbol{I} - B\boldsymbol{\Phi_1} - B^2\boldsymbol{\Phi_1^2}\right) = \boldsymbol{a_t}. \tag{3.110}$$

Процедура, использованная при получении (3.109), если ее применить несколько раз, ведет к модели фильтра. В модели авторегрессии, в отличие от модели фильтра, вместо значений шума в предыдущие моменты времени используются значения факторов. В модели авторегрессии 2-го порядка, также как в модели первого порядка, такое же число неизвестных параметров.

Обобщая (3.110) на случай авторегрессии произвольного порядка, получаем модель следующего вида

$$\boldsymbol{f_t}\left(\boldsymbol{I} - \sum_{j=1}^{p} B^j\boldsymbol{\Phi_1^j}\right) = \boldsymbol{a_t}, \tag{3.111}$$

где p – порядок авторегрессии. Поскольку матрица перехода из одного момента времени в любой другой выражается через матрицу перехода между соседними моментами времени, то далее ее будем обозначать без индекса, т.е. просто – $\boldsymbol{\Phi}$.

Таким образом, многомерная авторегрессионная модель факторов порядка p имеет следующий вид:

$$\boldsymbol{f_t}\left(\boldsymbol{I} - \sum_{j=1}^{p} B^j \boldsymbol{\Phi}\right) = \boldsymbol{a_t}. \tag{3.112}$$

Значения исследуемых переменных в текущий момент времени определяются значениями факторов в текущий момент:

$$\boldsymbol{z}_t = \boldsymbol{f}_t \boldsymbol{A}^T + \boldsymbol{v}_t, \tag{3.113}$$

где $\boldsymbol{A}$ − матрица факторных нагрузок, значения которых не меняются с течением времени;

$\boldsymbol{v}_t$ – значения произведения характерных факторов на соответствующие нагрузки в момент времени t, рассматриваются как погрешности факторного объяснения наблюдаемых переменных.

Рассмотрим модель факторного анализа временных рядов с авторегрессией первого порядка факторов. Она состоит из двух уравнений: уравнения авторегрессии факторов (3.115) и уравнения зависимости наблюдаемых переменных от факторов (3.113), т.е. определяется системой уравнений

$$\begin{cases} \boldsymbol{f_t} = \boldsymbol{f_{t-1}} \boldsymbol{\Phi} + \boldsymbol{a_t} \\ \boldsymbol{z}_t = \boldsymbol{f}_t \boldsymbol{A}^T + \boldsymbol{v}_t \end{cases}. \tag{3.114}$$

Сделаем для нее основные предположения.

Математическое ожидание случайных отклонений факторной авторегресии равно нулю для всех моментов времени:

$$\mathcal{M}\{\boldsymbol{a_t}\} = 0, t = 1, 2, \dots \tag{3.115}$$

Случайные отклонения факторной авторегрессии имеют постоянную дисперсию σ^2 для любого момента времени:

$$\mathcal{M}\{\boldsymbol{a}_t^T \boldsymbol{a_t}\} = \sigma^2 \boldsymbol{I}, t = 1, 2, \dots \tag{3.116}$$

где $\boldsymbol{I}$ – единичная матрица порядка m.

Случайные отклонения уравнения факторной автрегрессии для разных моментов времени не связаны друг с другом, т.е. между ними отсутствует корреляция:

$$\mathcal{M}\{\boldsymbol{a}_t^T \boldsymbol{a}_{t-s}\} = \boldsymbol{0}, t, s = 1, 2, \dots, \ \ t \neq s. \tag{3.117}$$

Поскольку наблюдаемые величины центрированы и нормированы, т.е. приведены к стандартному виду, то этому виду должны соответствовать и значения факторов. Поэтому очередное предположение следующее: математическое ожидание значений факторов для всех моментов времени равно нулю:

$$\mathcal{M}\{\boldsymbol{f_t}\} = 0, t = 0,1, 2, \dots \tag{3.118}$$

Значения факторов для начального момента времени – ортогональны:

$$\boldsymbol{f}_0^T \boldsymbol{f}_0 = \boldsymbol{I}. \tag{3.119}$$

Факторы и случайные отклонения не коррелируют друг с другом:

$$\mathcal{M}\{\boldsymbol{f}_t^T \boldsymbol{a_t}\} = \boldsymbol{0}, t = 1, 2, \dots \tag{3.120}$$

Матрица перехода $\boldsymbol{\Phi}$ должна иметь общее для подобного рода матриц свойство, а именно: чтобы вычислить по отношению к $\boldsymbol{\Phi}$ обратную матрицу, достаточно ее транспонировать:

$$\boldsymbol{\Phi}^{\mathrm{T}} = \boldsymbol{\Phi}^{-1}. \tag{3.121}$$

Этому свойству удовлетворяют матрицы, составленные из собственных векторов некоторой симметричной матрицы.

Погрешности уравнения зависимости наблюдаемых переменных от факторов имеют нулевое математическое ожидание:

$$\mathcal{M}\{\boldsymbol{v_t}\} = 0, t = 1, 2, \dots \tag{3.122}$$

Эти погрешности не коррелируют ни с наблюдаемыми переменными, ни с факторами, ни со случайными отклонениями факторной авторегрессии:

$$\mathcal{M}\{\boldsymbol{z}_t^T \boldsymbol{v_t}\} = \mathcal{M}\{\boldsymbol{f}_t^T \boldsymbol{v_t}\} = \mathcal{M}\{\boldsymbol{a}_t^T \boldsymbol{v_t}\} = \boldsymbol{0}, t = 1, 2, \dots \tag{3.123}$$

Погрешности второго уравнения системы (3.105+9) имеют постоянный с течением времени момент второго порядка:

$$\mathcal{M}\{\boldsymbol{v}_t^T \boldsymbol{v_t}\} = \boldsymbol{D^2}, t = 1,2, \dots, \tag{3.124}$$

где $\boldsymbol{D^2}$ − диагональная матрица дисперсий погрешностей, которая по аналогии с традиционным факторным анализом может быть названа матрицей значений характерностей.

Погрешности $\boldsymbol{v}$ для разных моментов времени не коррелируют между собой:

$$\mathcal{M}\{\boldsymbol{v}_t^T \boldsymbol{v}_{t-s}\} = \boldsymbol{0}, t, s = 1, 2, \dots, \ \ t \neq s. \tag{3.125}$$

Уравнение (3.106) предполагает, что значения факторов в некоторый момент времени являются случайными величинами. Рассмотрим некоторые свойства этих случайных величин, следующие из принятых предположений. Для этого преобразуем уравнение авторегрессии (3.106) в модель фильтра:

$$\boldsymbol{f_t} = (\boldsymbol{f_{t-2}\Phi} + \boldsymbol{a_{t-1}})\boldsymbol{\Phi} + \boldsymbol{a_t} = \cdots = \boldsymbol{f}_0\boldsymbol{\Phi}^{\mathrm{t}} + \boldsymbol{a}_1\boldsymbol{\Phi}^{\mathrm{t}-1} + \\ +\boldsymbol{a}_2\boldsymbol{\Phi}^{\mathrm{t}-2} + \cdots + \boldsymbol{a}_{t-1}\boldsymbol{\Phi} + \boldsymbol{a}_t = \boldsymbol{f}_0\boldsymbol{\Phi}^{\mathrm{t}} + \sum_{\mathrm{i}=1}^{\mathrm{t}} \boldsymbol{a}_i\boldsymbol{\Phi}^{\mathrm{t-i}}. \tag{3.126}$$

Используя выражение (3.126) и предположение о свойствах случайных величин модели (3.105+9), определим ковариацию значений факторов для момента времени t. Для этого найдем математическое ожидание произведения значений факторов:

$$\mathcal{M}\{\boldsymbol{f}_t^T\boldsymbol{f}_t\} = \left(\boldsymbol{f}_0\boldsymbol{\Phi}^{\mathrm{t}} + \sum_{\mathrm{i}=1}^{\mathrm{t}} \boldsymbol{a}_i\boldsymbol{\Phi}^{\mathrm{t-i}}\right)^T\left(\boldsymbol{f}_0\boldsymbol{\Phi}^{\mathrm{t}} + \sum_{\mathrm{i}=1}^{\mathrm{t}} \boldsymbol{a}_i\boldsymbol{\Phi}^{\mathrm{t-i}}\right) = \\ = \boldsymbol{I} + \sum_{\mathrm{i}=1}^{\mathrm{t}} \sigma^2\boldsymbol{I} = \boldsymbol{I} + t\sigma^2\boldsymbol{I} = (1 + t\sigma^2)\boldsymbol{I}. \tag{3.127}$$

Поскольку матрица ковариаций – диагональная, то значения факторов для любого момента времени – ортогональны, но на диагонали этой матрицы находятся значения большие единицы.

Матрица ковариаций значений факторов для моментов времени, расположенных рядом, имеет вид

$$\mathcal{M}\{\boldsymbol{f}_{t-1}^T\boldsymbol{f}_t\} = \mathcal{M}\{\boldsymbol{f}_{t-1}^T\boldsymbol{f}_{t-1}\boldsymbol{\Phi} + \boldsymbol{f}_{t-1}^T\boldsymbol{a}_t\} = (1 + (t-1)\sigma^2)\boldsymbol{\Phi}. \tag{3.128}$$

Выражение (3.128) показывает, что эта матрица уже не диагональная, поэтому свойство ортогональности значений факторов для разных моментов времени не выполняется, и ковариации между факторами определяются матрицей перехода. В частности, для произвольных моментов времени матрица ковариаций между значениями факторов определяется следующим выражением:

$$\mathcal{M}\{\boldsymbol{f}_{t-s}^T\boldsymbol{f}_t\} = \mathcal{M}\left\{\left(\boldsymbol{f}_0\boldsymbol{\Phi}^{\mathrm{t-s}} + \sum_{\mathrm{i}=1}^{\mathrm{t-s}} \boldsymbol{a}_i\boldsymbol{\Phi}^{\mathrm{t-s-i}}\right)^T\left(\boldsymbol{f}_0\boldsymbol{\Phi}^{\mathrm{t}} + \sum_{\mathrm{i}=1}^{\mathrm{t}} \boldsymbol{a}_i\boldsymbol{\Phi}^{\mathrm{t-i}}\right)\right\} = \\ = \boldsymbol{\Phi}^{\mathrm{s}} + \sum_{\mathrm{i}=1}^{\mathrm{t-s}} \left(\boldsymbol{\Phi}^{\mathrm{t-s-i}}\right)^T\boldsymbol{a}_i^T\boldsymbol{a}_i\boldsymbol{\Phi}^{\mathrm{t-s-i}} = \boldsymbol{\Phi}^{\mathrm{s}} + (\mathrm{t}-\mathrm{s})\sigma^2 + \boldsymbol{\Phi}^{\mathrm{s}} = \\ = (1 + (\mathrm{t}-\mathrm{s})\sigma^2)\boldsymbol{\Phi}^{\mathrm{s}}. \tag{3.129}$$

И в частности, когда $\mathrm{t} = \mathrm{s}$, имеем

$$\mathcal{M}\{\boldsymbol{f}_0^T\boldsymbol{f}_t\} = \boldsymbol{\Phi}^t, \tag{3.130}$$

т.е. ковариация значений факторов в некоторый момент времени со значениями в начальный момент времени не зависит от дисперсии случайных отклонений авторегрессии факторов.

Теперь рассмотрим связь между факторами и значениями наблюдаемых переменных. Естественно, что эта связь определяется уравнением (3.113), поэтому рассмотрим ковариацию между факторами и исследуемыми признаками

$$\mathcal{M}\{\boldsymbol{f}_t^T\boldsymbol{z}_t\} = \mathcal{M}\{\boldsymbol{f}_t^T(\boldsymbol{f}_t\boldsymbol{A}^T + \boldsymbol{v}_t)\} = \mathcal{M}\{\boldsymbol{f}_t^T\boldsymbol{f}_t\boldsymbol{A}^T + \boldsymbol{f}_t^T\boldsymbol{v}_t\} = (1 + t\sigma^2)\boldsymbol{A}^T. \tag{3.131}$$

Таким образом, ковариация между факторами и переменными определяется так же, как и в обычном факторном анализе, матрицей факторных нагрузок, но скорректированной на временной множитель $(1 + t\sigma^2)$. Однако для начального момента времени выполняется одна из теорем классического факторного анализа, а именно:

$$\mathcal{M}\{\boldsymbol{f}_0^T\boldsymbol{z}_0\} = \boldsymbol{A}^T. \tag{3.132}$$

Ковариация между факторами для некоторого момента времени и значениями наблюдаемых переменных для другого момента времени, число шагов между которыми равно s можно определить на основании уравнений (3.131) и (3.129), и она выглядит следующим образом

$$\mathcal{M}\{\boldsymbol{f}_t^T\boldsymbol{z}_{t-s}\} = (1 + (\mathrm{t}-\mathrm{s})\sigma^2)\boldsymbol{\Phi}^{\mathrm{s}}\boldsymbol{A}^T, \tag{3.133}$$

Рассмотрим вопрос о корреляции между наблюдаемыми переменными для некоторого момента времени t, которые согласно догмам факторного анализа объясняются исключительно факторными нагрузками. Для этого воспользуемся моделью (3.105+9). Матрица корреляций между переменными имеет вид:

$$\boldsymbol{R} = \mathcal{M}\{\boldsymbol{z}_t^T\boldsymbol{z}_t\} = \mathcal{M}\{(\boldsymbol{f}_t\boldsymbol{A}^T + \boldsymbol{v}_t)^T(\boldsymbol{f}_t\boldsymbol{A}^T + \boldsymbol{v}_t)\} = \\ = \mathcal{M}\{\boldsymbol{A}\boldsymbol{f}_t^T\boldsymbol{f}_t\boldsymbol{A}^T + \boldsymbol{A}\boldsymbol{f}_t^T\boldsymbol{v}_t + \boldsymbol{v}_t^T\boldsymbol{f}_t\boldsymbol{A}^T + \boldsymbol{v}_t^T\boldsymbol{v}_t\} =$$

$$= \mathcal{M}\{\boldsymbol{A}\boldsymbol{f}_t^T\boldsymbol{f}_t\boldsymbol{A}^T + \boldsymbol{v}_t^T\boldsymbol{v}_t\} = (1 + t\sigma^2)\boldsymbol{A}\boldsymbol{A}^T + \boldsymbol{D}^2. \qquad (3.134)$$

В выражении (3.133) значения на главной диагонали корреляционной матрицы могут и не иметь величину, равную единице из-за наличия временного множителя $(1 + t\sigma^2)$, большего единицы. На самом деле этот вопрос требует дополнительного теоретического исследования. Но, если элементы главной диагонали матрицы $(1 + t\sigma^2)\boldsymbol{A}\boldsymbol{A}^T$ меньше единицы, то значения характерностей $\boldsymbol{D}^2$ будут дополнять эти элементы до единицы, и тогда выражение (3.133) совпадет с основной теоремой факторного анализа. Основным условием для этого является выполнение неравенств

$$aa_{ii} < \frac{1}{1+t\sigma^2}, i = 1,2, \dots n, \qquad (3.135)$$

где aa_{ii} – диагональный элемент матрицы $\boldsymbol{A}\boldsymbol{A}^T$,
n − число наблюдаемых переменных.

Рассмотрим вид основных свойств факторной модели авторегрессии первого порядка факторов в предельном значении, т.е. когда $t \to \infty$. При исследовании асимптотических свойств оценок параметров статистических моделей в этом случае традиционно предполагается, что значение дисперсии погрешностей стремится к нулю, т.е.

$$\lim_{t\to\infty} \sigma^2 = 0.$$

Тогда основные тождества факторной модели авторегрессии первого порядка приобретают следующий смысл:

значения факторов в предельном значении не только ортогональны, но и нормированы:

$$\mathcal{M}\{\boldsymbol{f}_t^T\boldsymbol{f}_t\} = \boldsymbol{I},\ t \to \infty, \qquad (3.136)$$

ковариация между соседними значениями факторов совпадает с матрицей переходов:

$$\mathcal{M}\{\boldsymbol{f}_{t-1}^T\boldsymbol{f}_t\} = \boldsymbol{\Phi},\ t \to \infty, \qquad (3.137)$$

ковариация между значениями факторов, расположенных на расстоянии s шагов друг от друга равна матрице переходов, возведенной в степень s:

$$\mathcal{M}\{\boldsymbol{f}_{t-s}^T\boldsymbol{f}_t\} = \boldsymbol{\Phi}^s, t \to \infty\ , \qquad (3.138)$$

ковариация между значениями факторов и значениями наблюдаемых переменных определяется матрицей факторных нагрузок:

$$\mathcal{M}\{\boldsymbol{f}_t^T\boldsymbol{z}_t\} = \boldsymbol{A}^T, t \to \infty\ , \qquad (3.139)$$

ковариация между значениями факторов для одного момента времени и значениями наблюдаемых переменных для момента времени, отстоящего на s шагов определяется матрицей переходов, возведенной в степень s и матрицей факторных нагрузок

$$\mathcal{M}\{\boldsymbol{f}_t^T\boldsymbol{z}_{t-s}\} = \boldsymbol{\Phi}^s\boldsymbol{A}^T, t \to \infty, \qquad (3.140)$$

для предельного случая корреляция между переменными определяется основной теоремой факторного анализа:

$$\boldsymbol{R} = \boldsymbol{A}\boldsymbol{A}^T + \boldsymbol{D}^2, t \to \infty. \qquad (3.141)$$

Одним из главных постулатов факторной модели авторегрессии первого порядка факторов является утверждение, состоящее в том, что исследуемые переменные предшествующих моментов времени не оказывают воздействия на текущее значение переменных, но ковариация между ними формально может быть вычислена. Для предельного случая она имеет вид

$$\boldsymbol{R}_{t,t-s} = \boldsymbol{A}\boldsymbol{\Phi}^s\boldsymbol{A}^T + \boldsymbol{D}^2, t \to \infty, \qquad (3.142)$$

где $\boldsymbol{R}_{t,t-s}$ – матрица корреляции между наблюдаемыми переменными для моментов времени, расположенных на расстоянии в s шагов.

3.3. Факторная авторегрессионная модель анализа многомерных временных рядов экономической информации.

Рассмотрим экономическую систему, поведение которой описывается рядом показателей. Показатели являются переменными случайными величинами, связанными, или наоборот, не связанными между собой некоторыми неизвестными заранее зависимостями. Поскольку каждый показатель важен для описания системы, то форма зависимости между переменными не будет рассматриваться на этом этапе исследования, целью которого является получение прогноза для всей совокупности показателей.

Сделаем очевидные, но важные предположения для начала исследования: фиксируемые и измеряемые переменные несут полную информацию об экономическом объекте, поскольку будем считать, что регистрируются значения всех возможных характеристик объекта. Избыточность информации, являющаяся вполне логичной в таком случае, не является проблемой, которую необходимо разрешить на данном этапе, и не может привести к искажению выводов.

Существующие методы получения прогнозов основаны на теории Винера-Колмогорова прогнозирования случайных процессов. А эта теория рассматривает процессы, разложимые в ряд Фурье. При этом условие стационарности, или близости к ней, является существенным.

Будем исходить из того, что случайный процесс (управляемый или нет) может быть как стационарным, так и нестационарным. Причем переменные, описывающие экономический процесс, влияют друг на друга, и предсказание поведения отдельного показателя без учета влияния на него других показателей, будет вести к заведомо ложным выводам.

Итак, переменные взаимодействуют между собой, обуславливая тем самым свое будущее поведение. За этим взаимодействием, или взаимным влиянием стоят некие причины – неизвестные, нерегистрируемые и ненаблюдаемые факторы. Переменные в таком случае можно рассматривать как косвенное измерение этих факторов.

Более того, взаимодействие между переменными обусловлено тем, что они отражают разные стороны одного фактора, − и это с одной стороны, а с другой стороны – факторы являются тем скрытым механизмом, который заставляет взаимодействовать или коррелировать разные переменные величины, отражающие различные стороны экономического объекта.

Факторы, или причины, между собой не связаны. И это предположение делается на основании следующего постулата: если бы такая связь имела бы место, то это означало бы, что за взаимодействием таких факторов стоит одна общая причина, объясняющая поведение связанных факторов. И, чтобы установить закономерности динамических изменений исследуемого процесса, достаточно учесть фактор, стоящий над факторами. В факторном анализе такое предположение является основополагающим: факторы – это ортогональные, т.е. независимые надпеременные.

Таким образом, будем исходить их того, что факторы являются независимыми величинами, имеющими неслучайную природу. Т.е. их поведение некоторым образом, но пока неизвестным, предопределено.

Значения факторов определяют состояние экономической системы, и изменение факторов ведет к тому, что система переходит из состояния в состояние с течением времени. Сам переход обусловлен временным фактором, однако, у такого перехода должны быть общие для разных моментов времени закономерности. Если закономерности переходов отсутствуют, то динамика системы имеет характер броуновского движения. Для экономической системы такое движение возможно, однако более общим будет предложение о некотором «разумном» движении, имеющем хоть какую-нибудь закономерность.

Будем фиксировать переходы в определенные моменты времени, рассматривая значения экономических показателей за соответствующий промежуток времени (год, квартал, месяц, etc).

Если закономерности переходов меняются с течением времени, то экономический процесс является нестационарным. Если закономерности переходов остаются неизменными со временем, то экономический процесс – стационарный.

Будем рассматривать фактор, как причину движения экономической системы, абстрагируясь от его места по отношению к самой системе по следующим соображениям. Характер процессов внутри системы может быть обусловлен либо строго внутренними механизмами взаимодействия отдельных ее элементов, либо только внешними «толчками», либо сочетанием тех и других причин. Очевидно, внешние факторы являются причинами внутренних изменений (обратное трудно представить). Поэтому говорить о «внутренних» факторах, не вступая в противоречие с предположением о независимости факторов, будет некорректно. И фактор, имея, безусловно, внешнее происхождение, может оказаться энергетикой структурных внутренних сдвигов, т.е. частью самой экономической системы.

Таким образом, факторы обуславливают поведение экономической системы и, соответственно, определяют значения показателей, с помощью которых система контролируется. Поэтому, чтобы прогнозировать поведение системы с минимальной ошибкой, следует прогнозировать значения факторов, а значения экономических показателей восстанавливать по оцененным величинам факторов.

Именно динамику факторов, как совокупности независимых друг от друга величин, каждую из которых можно рассматривать как компоненту вектора, следует трактовать как векторный процесс, подпадающий под общую теорию Винера-Колмогорова прогнозирования случайных процессов [56, с.145].

Эта теория рассматривает, в основном, прогноз одномерной величины на основании ее спектральных характеристик. Многомерные величины трактуются как вектор с независимыми координатами, наличие корреляции между которыми приведет к искажению не только результатов прогноза, - но и к искажению идентификации самого процесса.

Математическая модель

Рассмотрим дискретный вариант изменения в экономической системе: факторы меняют свои значения в определенные моменты времени $t = 0,1,2,...$

Динамика изменения значений факторов описывается уравнением переходов

$$\boldsymbol{f}_t = \boldsymbol{f}_{t-1}\boldsymbol{\Phi}_{t-1,t}, \tag{3.143}$$

где $\boldsymbol{f}_t = (f_{t1}, f_{t2}, \dots, f_{tm})$ – вектор значений факторов в момент времени t,

$m-$ число факторов;

$\boldsymbol{\Phi}_{t-1,t}$ – матрица переходов значений факторов от момента времени $t-1$ к моменту t, которая имеет вид:

$$\boldsymbol{\Phi}_{t-1,t} = \begin{pmatrix} \varphi_{11} & \cdots & \varphi_{1m} \\ \vdots & \ddots & \vdots \\ \varphi_{m1} & \cdots & \varphi_{mm} \end{pmatrix}$$

Значения изменяемых переменных задаются уравнением факторной структуры

$$\boldsymbol{z}_t = \boldsymbol{f}_t\boldsymbol{A}^T + \boldsymbol{v}_t, \tag{3.144}$$

где $\boldsymbol{z}_t = (z_{t1}, z_{t2}, \dots, z_{tm})$ – вектор стандартизированных значений переменных – экономических показателей в момент времени t; $n-$ число изменяемых показателей; $(\cdot)^T$ знак транспонирования матрицы;

$\boldsymbol{A}-$ матрица факторных нагрузок, имеющая вид

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix}$$

$\boldsymbol{v}_t = (v_{t1}, v_{t2}, \dots, v_{tm})$ – вектор случайных возмущений уравнения факторной структуры.

Стандартизированное значение экономического показателя означает, что из его реального, наблюденного значения вычли среднее значение и эту разность разделили на среднеквадратическое отклонение:

$$z_{tk} = \frac{x_{tk} - \bar{x}_k}{\sigma_k},$$

где $x_{tk}-$ зафиксированное, реальное значение экономического показателя;

$\bar{x}_k-$ среднее значение k-го показателя;

σ_k- его среднеквадратическое отклонение.

По своей математической сути уравнения (3.143) и (3.144) представляют собой фильтр Калмана, причем (3.143) – уравнение переходов, (3.144) – уравнение измерений. Однако Калман предполагал, что все матрицы известны и необходимо было оценить значение переданной, но неизвестной информации $\boldsymbol{f}_t$ по принятому с наложением шумов сигналу $\boldsymbol{z}_t$. В нашем случае, ни одна матрица, кроме значений наблюдаемых переменных $\boldsymbol{z}_t$, не известна, неизвестными являются также значения факторов. Предстоит оценить матрицу переходов, матрицу факторных нагрузок и найти значения факторов. Такой постановки задачи у Калмана нет, и соответственно воспользоваться ходом его решения данной задачи не представляется возможным. Поэтому предварительно исследуем свойства случайных отклонений, а также факторов.

Случайные отклонения уравнения переходов и уравнения факторной структуры должны удовлетворять следующим условиям:

- математическое ожидание случайного отклонения равно нулю

$$\mathcal{M}\{\boldsymbol{v}_t\} = 0, (t = 0, 1, 2, \ldots, N),$$

здесь $\mathcal{M}\{\cdot\}$ – символ математического ожидания, и N – число всех наблюдений;

- случайные отклонения не коррелируют с переменными и факторами

$$\begin{matrix} \mathcal{M}\{\boldsymbol{v}_t^T\boldsymbol{f}_t\} = 0; \\ \mathcal{M}\{\boldsymbol{v}_t^T\boldsymbol{f}_{t-1}\} = 0; \end{matrix} (t = 1, 2, \ldots, N);$$

- момент второго порядка случайного отклонения уравнения факторной структуры для каждого момента времени равен

$$\mathcal{M}\{\boldsymbol{v}_t^T\boldsymbol{v}_t\} = \boldsymbol{D}^2,$$

где $\boldsymbol{D}^2$ – диагональная матрица, называемая в факторном анализе матрицей значений характерностей; она представляет собой матрицу дисперсий, так называемых, характерных факторов.

- момент второго порядка случайных отклонений уравнения факторной структуры для разных моментов времени равен нулю:

$$\mathcal{M}\{\boldsymbol{v}_{t-1}^T\boldsymbol{v}_t\} = 0$$

Теперь отметим основное предположение для факторов:

факторы в данный момент времени являются ортогональными, и это справедливо для всех моментов времени

$$\boldsymbol{f}_t^T\boldsymbol{f}_t = \boldsymbol{I}; (t = 0, 1, 2, \ldots, N),$$

здесь $\boldsymbol{I}$ – единичная матрица размером $m \times m$.

Следующие утверждения вытекают из выше сделанных предположений.

1) Парные корреляции между переменными образуют матрицу $\boldsymbol{R}$, которая для всех моментов времени имеет вид

$$\begin{gathered} \boldsymbol{R} = \mathcal{M}\{\boldsymbol{z}_t^T\boldsymbol{z}_t\} = \mathcal{M}\{(\boldsymbol{f}_t\boldsymbol{A}^T + \boldsymbol{v}_t)^T(\boldsymbol{f}_t\boldsymbol{A}^T + \boldsymbol{v}_t)\} = \\ = \boldsymbol{A}\boldsymbol{f}_t^T\boldsymbol{f}_t\boldsymbol{A}^T + \boldsymbol{A}\boldsymbol{f}_t^T\mathcal{M}\{\boldsymbol{v}_t\} + \mathcal{M}\{\boldsymbol{v}_t^T\}\boldsymbol{f}_t\boldsymbol{A}^T + \mathcal{M}\{\boldsymbol{v}_t^T\boldsymbol{v}_t\} = \\ = \boldsymbol{A}\boldsymbol{A}^T + \boldsymbol{D}^2. \end{gathered}$$

Уравнение

$$\boldsymbol{R} = \boldsymbol{A}\boldsymbol{A}^T + \boldsymbol{D}^2$$

называется «фундаментальной теоремой факторного анализа».

2) Корреляция между переменными и факторами имеет вид:

$$\mathcal{M}\{\boldsymbol{f}_t^T\boldsymbol{z}_t\} = \mathcal{M}\{\boldsymbol{f}_t^T\boldsymbol{f}_t\boldsymbol{A}^T + \boldsymbol{f}_t^T\boldsymbol{v}_t\} = \boldsymbol{A}^T,$$

здесь учтена ортогональность факторов и равенство нулю математического ожидания случайных отклонений. Таким образом, корреляции между переменными и факторами определяют матрицу факторных нагрузок.

3) Переходы значений факторов определяют переходы значений экономических показателей.

Покажем это. Обозначим матрицу переходов значений показателей как $\boldsymbol{M}_{t-1,t}$, т.е.

$$\boldsymbol{M}_{t-1,t} = \mathcal{M}\{\boldsymbol{z}_{t-1}^T\boldsymbol{z}_t\}$$

тогда, поскольку

$$\boldsymbol{z}_{t-1} = \boldsymbol{f}_{t-1}\boldsymbol{A}^T + \boldsymbol{v}_{t-1},$$

то

$$\begin{gathered} \boldsymbol{M}_{t-1,t} = \mathcal{M}\{(\boldsymbol{f}_{t-1}\boldsymbol{A}^T + \boldsymbol{v}_{t-1})^T(\boldsymbol{f}_t\boldsymbol{A}^T + \boldsymbol{v}_t)\} = \\ = \boldsymbol{A}\boldsymbol{f}_{t-1}^T\boldsymbol{f}_t\boldsymbol{A}^T + \mathcal{M}\{\boldsymbol{v}_{t-1}^T\}\boldsymbol{f}_t\boldsymbol{A}^T + \boldsymbol{A}\boldsymbol{f}_{t-1}^T\mathcal{M}\{\boldsymbol{v}_t\} + \mathcal{M}\{\boldsymbol{v}_{t-1}^T\boldsymbol{v}_t\} \end{gathered}$$

Учитывая предположение, что случайные отклонения для разных моментов времени независимы друг от друга, и имея в виду, что

$$\boldsymbol{\Phi}_{t-1,t} = \boldsymbol{f}_{t-1}^T\boldsymbol{f}_t,$$

получаем зависимость между переходами показателей и переходами факторов

$$\boldsymbol{M}_{t-1,t} = \boldsymbol{A}\boldsymbol{\Phi}_{t-1,t}\boldsymbol{A}^T.$$

4) Чтобы найти обратную для матрицы переходов факторов, достаточно ее транспонировать.

Убедимся в этом. Умножим (3.142+1) слева на $\boldsymbol{f}_t^T$, тогда получим :

$$\boldsymbol{f}_t^T\boldsymbol{f}_t = \boldsymbol{f}_{t-1}^T\boldsymbol{f}_{t-1}\boldsymbol{\Phi}_{t-1,t}. \quad (3.145)$$

Учитывая ортогональность факторов, а также то, что произведение факторов для разных моментов времени задают матрицу переходов:

$$\boldsymbol{f}_t^T \boldsymbol{f}_{t-1} = \boldsymbol{\Phi}_{t,t-1},$$

выражение (3.145) преобразуется к виду

$$\boldsymbol{I} = \boldsymbol{\Phi}_{t,t-1} \boldsymbol{\Phi}_{t-1,t},$$

из которого следует

$$\boldsymbol{\Phi}_{t,t-1} = \boldsymbol{\Phi}_{t,t-1}^{-1}.$$

С другой стороны, транспонируем (3.142+1), после чего умножим справа на $\boldsymbol{f}_{t-1}$, получим:

$$\boldsymbol{f}_t^T \boldsymbol{f}_{t-1} = \boldsymbol{\Phi}_{t-1,t}^T \boldsymbol{f}_{t-1}^T \boldsymbol{f}_{t-1},$$
$$\boldsymbol{\Phi}_{t,t-1} = \boldsymbol{\Phi}_{t-1,t}^T.$$

После чего и получается, что

$$\boldsymbol{\Phi}_{t-1,t}^T = \boldsymbol{\Phi}_{t-1,t}^{-1}. \tag{3.146}$$

Таким же свойством обладают матрицы собственных векторов, соответствующих собственным значениям симметрической матрицы.

Покажем, что случайные отклонения уравнения факторной структуры для разных моментов времени являются независимыми:

$$\mathcal{M}\{\boldsymbol{v}_{t-1}^T \boldsymbol{v}_t\} = \mathcal{M}\{(\boldsymbol{z}_{t-1} - \boldsymbol{f}_{t-1}\boldsymbol{A}^T)^T (\boldsymbol{z}_t - \boldsymbol{f}_t \boldsymbol{A}^T)\} =$$
$$= \mathcal{M}\{\boldsymbol{z}_{t-1}^T \boldsymbol{z}_t - \boldsymbol{A}\boldsymbol{f}_{t-1}^T \boldsymbol{z}_t - \boldsymbol{z}_{t-1}^T \boldsymbol{f}_t \boldsymbol{A}^T + \boldsymbol{A}\boldsymbol{f}_{t-1}^T \boldsymbol{f}_t \boldsymbol{A}^T\}. \tag{3.147}$$

Рассмотрим отдельные слагаемые последнего выражения. Математическое ожидание первого слагаемого имеет вид

$$\mathcal{M}\{\boldsymbol{z}_{t-1}^T \boldsymbol{z}_t\} = \boldsymbol{M}_{t-1,t} = \boldsymbol{A}\boldsymbol{\Phi}_{t-1,t}\boldsymbol{A}^T.$$

Пользуясь свойством (3.146), легко заметить, что из уравнения (3.143) следует

$$\boldsymbol{f}_{t-1} = \boldsymbol{f}_t \boldsymbol{\Phi}_{t-1,t}^T,$$

или, если последнее выражение транспонировать, –

$$\boldsymbol{f}_{t-1}^T = \boldsymbol{\Phi}_{t-1,t} \boldsymbol{f}_t^T,$$

и тогда математическое ожидание двух сомножителей второго слагаемого выражения (3.147) приобретает вид:

$$\mathcal{M}\{\boldsymbol{f}_{t-1}^T \boldsymbol{z}_t\} = \mathcal{M}\{\boldsymbol{\Phi}_{t-1,t} \boldsymbol{f}_t^T \boldsymbol{z}_t\} = \boldsymbol{\Phi}_{t-1,t}\boldsymbol{A}^T.$$

Аналогично несложно заметить, что математическое ожидание двух сомножителей третьего слагаемого равно:

$$\mathcal{M}\{\boldsymbol{z}_{t-1}^T \boldsymbol{f}_t\} = \boldsymbol{A}\boldsymbol{\Phi}_{t-1,t}.$$

Относительно внутренних сомножителей четвертого слагаемого можно сказать, что их произведение равно матрице переходов факторов $\boldsymbol{\Phi}_{t-1,t}$.

В итоге, выражение (3.147) приобретает вид

$$\mathcal{M}\{\boldsymbol{v}_{t-1}^T \boldsymbol{v}_t\} = \boldsymbol{A}\boldsymbol{\Phi}_{t-1,t}\boldsymbol{A}^T - \boldsymbol{A}\boldsymbol{\Phi}_{t-1,t}\boldsymbol{A}^T - \boldsymbol{A}\boldsymbol{\Phi}_{t-1,t}\boldsymbol{A}^T + \boldsymbol{A}\boldsymbol{\Phi}_{t-1,t}\boldsymbol{A}^T,$$

т.е. действительно случайные отклонения для разных моментов времени являются независимыми. Изначально это предполагалось, на этом предположении были выведены рассмотренные свойства, и данные свойства не противоречат предположениям.

Найдем оценку матрицы переходов факторов.

Найдем оценку $\boldsymbol{\Phi}_{t-1,t}$ как решение экстремальной задачи поиска минимума суммы квадратов отклонений матрицы переходов значений переменных от матрицы переходов, задаваемых переходами значений факторов, т.е. в качестве основного выражения, формирующего критерий экстремальной задачи, является выражение:

$$\boldsymbol{M}_{t-1,t} = \boldsymbol{A}\boldsymbol{\Phi}_{t-1,t}\boldsymbol{A}^T,$$

и отличие левой и правой частей этого выражения, возведенное в квадрат, будет являться основным критерием. Следовательно, будем искать минимум функции

$$g_1(\boldsymbol{\Phi}) = tr\{(\boldsymbol{M} - \boldsymbol{A}\boldsymbol{\Phi}\boldsymbol{A}^T)^T (\boldsymbol{M} - \boldsymbol{A}\boldsymbol{\Phi}\boldsymbol{A}^T)\},$$

где tr – «след матрицы», т.е. сумма диагональных элементов матрицы;

$\boldsymbol{M}$ и $\boldsymbol{\Phi}$ – обозначают матрицы, соответственно, $\boldsymbol{M}_{t-1,t}$ и $\boldsymbol{\Phi}_{t-1,t}$, т.е. опущены нижние переходные индексы, поскольку другие переходы не рассматриваются, кроме как из предыдущего в последующее состояние.

Кроме того, оценки должны быть такими, чтобы удовлетворялось условие (3.146) – обратная матрица переходов может быть получена путем транспонирования.

Таким образом, имеет место задача поиска условного экстремума: найти минимум функции g_1 при ограничении (3.146). Для ее решения составим функцию Лагранжа

$$g(\boldsymbol{\Phi}, \boldsymbol{\Lambda}) = g_1(\boldsymbol{\Phi}) + tr\{\boldsymbol{\Lambda}(\boldsymbol{\Phi}^T\boldsymbol{\Phi} - \boldsymbol{I})\},$$

где $\boldsymbol{\Lambda}$ - матрица коэффициентов Лагранжа. Окончательно, функция Лагранжа имеет вид

$$g(\boldsymbol{\Phi}, \boldsymbol{\Lambda}) = tr\{(\boldsymbol{M} - \boldsymbol{A\Phi A}^T)^T(\boldsymbol{M} - \boldsymbol{A\Phi A}^T)\} + tr\{\boldsymbol{\Lambda}(\boldsymbol{\Phi}^T\boldsymbol{\Phi} - \boldsymbol{I})\}. \quad (3.148)$$

Для того, чтобы найти условный экстремум, необходимо найти частные производные функции Лагранжа по ее аргументам $\boldsymbol{\Phi}, \boldsymbol{\Lambda}$.

Частная производная функции g по матрице факторных переходов имеет вид:

$$\begin{aligned}\frac{\partial g}{\partial \boldsymbol{\Phi}} &= \frac{\partial g}{\partial \boldsymbol{\Phi}}(tr\{\boldsymbol{M}^T\boldsymbol{M} - \boldsymbol{M}^T\boldsymbol{A\Phi A}^T - \boldsymbol{A\Phi A}^T\boldsymbol{M} + \boldsymbol{A\Phi}^T\boldsymbol{A}^T\boldsymbol{A\Phi A}^T\}) + tr\{\boldsymbol{\Lambda\Phi}^T\boldsymbol{\Phi} - \boldsymbol{\Lambda}\} = \\ &= (-\boldsymbol{A}^T\boldsymbol{M}\otimes\boldsymbol{A}^T - \boldsymbol{E}(\boldsymbol{A}^T\otimes\boldsymbol{A}^T\boldsymbol{M}) + \boldsymbol{E}(\boldsymbol{A}^T\otimes\boldsymbol{A}^T\boldsymbol{A\Phi A}^T) + \boldsymbol{A}^T\boldsymbol{A\Phi A}^T\otimes\boldsymbol{A}^T)vec\{\boldsymbol{I}\} + \\ &+(\boldsymbol{E}(\boldsymbol{\Lambda}^T\otimes\boldsymbol{\Phi}) + \boldsymbol{\Phi\Lambda}^T\otimes\boldsymbol{I})vec\{\boldsymbol{I}\} = vec\{-\boldsymbol{A}^T\boldsymbol{MA} - (\boldsymbol{AM}^T\boldsymbol{A})^T + (\boldsymbol{A}^T\boldsymbol{A\Phi}^T\boldsymbol{A}^T\boldsymbol{A})^T + \\ &+\boldsymbol{A}^T\boldsymbol{A\Phi A}^T\boldsymbol{A}\} + vec\{\boldsymbol{\Phi\Lambda} + \boldsymbol{\Phi\Lambda}^T\} = vec\{-2\boldsymbol{A}^T\boldsymbol{MA} + 2\boldsymbol{A}^T\boldsymbol{A\Phi A}^T\boldsymbol{A} + 2\boldsymbol{\Phi\Lambda}\}, \quad (3.149)\end{aligned}$$

где $\boldsymbol{E}$ − перестановочная матрица,

$\otimes$ − символ произведения Кронекера,

vec − операция «векторизации» матрицы, ее результатом является вектор-столбец, формируемый из исходной матрицы взятием строки за строкой.

В заключительном равенстве выражения (3.149) было учтено, что матрица коэффициентов Лагранжа – симметричная, т.е.

$$\boldsymbol{\Lambda}^T = \boldsymbol{\Lambda},$$

поскольку уравнение

$$\boldsymbol{\Phi}^T\boldsymbol{\Phi} = \boldsymbol{I},$$

представляющее собой условие, является симметричной матрицей, и, следовательно, матрица коэффициентов Лагранжа должна быть также симметричной.

Найдем частные производные по $\boldsymbol{\Lambda}$

$$\frac{\partial g}{\partial \boldsymbol{\Lambda}} = (I\otimes(\boldsymbol{\Phi}^T\boldsymbol{\Phi} - \boldsymbol{I})vec\{\boldsymbol{I}\} = vec\{\boldsymbol{\Phi}^T\boldsymbol{\Phi} - \boldsymbol{I}\}.$$

Приравнивая частные производные нулю, и переходя от векторной к обычной матричной форме, получим систему уравнений

$$\begin{cases}-\boldsymbol{A}^T\boldsymbol{MA} + \boldsymbol{A}^T\boldsymbol{A\Phi A}^T\boldsymbol{A} + \boldsymbol{\Phi\Lambda} = \boldsymbol{0}, \\ \boldsymbol{\Phi}^T\boldsymbol{\Phi} - \boldsymbol{I} = \boldsymbol{0}.\end{cases} \quad (3.150)$$

Выразим $\boldsymbol{\Phi}$ из первого уравнения системы (3.142+8), пользуясь свойствами векторного представления матриц:

$$(\boldsymbol{A}^T\boldsymbol{A}\otimes\boldsymbol{A}^T\boldsymbol{A})vec\{\boldsymbol{\Phi}\} + (\boldsymbol{I}\otimes\boldsymbol{\Lambda})vec\{\boldsymbol{\Phi}\} = vec\{\boldsymbol{A}^T\boldsymbol{MA}\},$$

$$(\boldsymbol{A}^T\boldsymbol{A}\otimes\boldsymbol{A}^T\boldsymbol{A} + \boldsymbol{I}\otimes\boldsymbol{\Lambda})vec\{\boldsymbol{\Phi}\} = vec\{\boldsymbol{A}^T\boldsymbol{MA}\},$$

$$vec\{\boldsymbol{\Phi}\} = (\boldsymbol{A}^T\boldsymbol{A}\otimes\boldsymbol{A}^T\boldsymbol{A} + \boldsymbol{I}\otimes\boldsymbol{\Lambda})^{-1}vec\{\boldsymbol{A}^T\boldsymbol{MA}\}.$$

Произведем небольшие преобразования, пользуясь, опять же, свойствами векторного представления матриц:

$$(\boldsymbol{\Phi}\otimes\boldsymbol{I})vec\{\boldsymbol{I}\} = (\boldsymbol{A}^T\boldsymbol{A}\otimes\boldsymbol{A}^T\boldsymbol{A} + \boldsymbol{I}\otimes\boldsymbol{\Lambda})^{-1}(\boldsymbol{A}^T\boldsymbol{MA}\otimes\boldsymbol{I})vec\{\boldsymbol{I}\}.$$

Перенесем правую часть последнего выражения налево и вынесем $vec\{\boldsymbol{I}\}$ за скобки:

$$\left(\boldsymbol{\Phi}\otimes\boldsymbol{I} - (\boldsymbol{A}^T\boldsymbol{A}\otimes\boldsymbol{A}^T\boldsymbol{A} + \boldsymbol{I}\otimes\boldsymbol{\Lambda})^{-1}(\boldsymbol{A}^T\boldsymbol{MA}\otimes\boldsymbol{I})\right)vec\{\boldsymbol{I}\} = \boldsymbol{0},$$

откуда

$$\boldsymbol{\Phi}\otimes\boldsymbol{I} = (\boldsymbol{A}^T\boldsymbol{A}\otimes\boldsymbol{A}^T\boldsymbol{A} + \boldsymbol{I}\otimes\boldsymbol{\Lambda})^{-1}(\boldsymbol{A}^T\boldsymbol{MA}\otimes\boldsymbol{I}).$$

Для последнего выражения найдем обратную матрицу

$$\boldsymbol{\Phi}^{-1}\otimes\boldsymbol{I} = ((\boldsymbol{A}^T\boldsymbol{MA})^{-1}\otimes\boldsymbol{I})(\boldsymbol{A}^T\boldsymbol{A}\otimes\boldsymbol{A}^T\boldsymbol{A} + \boldsymbol{I}\otimes\boldsymbol{\Lambda}),$$

$$\boldsymbol{\Phi}^{-1}\otimes\boldsymbol{I} = (\boldsymbol{A}^T\boldsymbol{MA})^{-1}\boldsymbol{A}^T\boldsymbol{A}\otimes\boldsymbol{A}^T\boldsymbol{A} + (\boldsymbol{A}^T\boldsymbol{MA})^{-1}\otimes\boldsymbol{\Lambda},$$

или после умножения обеих частей уравнения на $vec\{\boldsymbol{I}\}$, и сворачивания в вектор, а затем, приравнивая аргументы оператора vec, получаем

$$\boldsymbol{\Phi}^{-1} = (\boldsymbol{A}^T\boldsymbol{MA})^{-1}\boldsymbol{A}^T\boldsymbol{AA}^T\boldsymbol{A} + (\boldsymbol{A}^T\boldsymbol{MA})^{-1}\boldsymbol{\Lambda}.$$

Вынесем общий множитель правой части за скобки, получим

$$\boldsymbol{\Phi}^{-1} = (\boldsymbol{A}^T\boldsymbol{MA})^{-1}(\boldsymbol{A}^T\boldsymbol{AA}^T\boldsymbol{A} + \boldsymbol{\Lambda}). \quad (3.151)$$

Обращение выражения (3.151) дает матрицу факторных переходов, вычисленную с точностью до матрицы множителей Лагранжа:

$$\boldsymbol{\Phi} = (\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{A}^T\boldsymbol{A} + \boldsymbol{\Lambda})^{-1}\boldsymbol{A}^T\boldsymbol{M}\boldsymbol{A}. \tag{3.152}$$

Если транспонировать эту матрицу, то получим

$$\boldsymbol{\Phi}^T = \boldsymbol{A}^T\boldsymbol{M}^T\boldsymbol{A}(\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{A}^T\boldsymbol{A} + \boldsymbol{\Lambda})^{-1}. \tag{3.153}$$

Выражения (3.152) и (3.153) подставим в (3.150) и найдем матрицу $\boldsymbol{\Lambda}$

$$\boldsymbol{A}^T\boldsymbol{M}^T\boldsymbol{A}(\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{A}^T\boldsymbol{A} + \boldsymbol{\Lambda})^{-2}\boldsymbol{A}^T\boldsymbol{M}\boldsymbol{A} = \boldsymbol{I}. \tag{3.154}$$

Из уравнения (3.154) выразим обратную матрицу:

$$(\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{A}^T\boldsymbol{A} + \boldsymbol{\Lambda})^{-2} = (\boldsymbol{A}^T\boldsymbol{M}^T\boldsymbol{A})^{-1}(\boldsymbol{A}^T\boldsymbol{M}\boldsymbol{A})^{-1}. \tag{3.155}$$

Объединим обратные матрицы правой части выражения (3.155):

$$(\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{A}^T\boldsymbol{A} + \boldsymbol{\Lambda})^{-2} = (\boldsymbol{A}^T\boldsymbol{M}\boldsymbol{A}\boldsymbol{A}^T\boldsymbol{M}^T\boldsymbol{A})^{-1}, \tag{3.156}$$

и найдем обратную матрицу для выражения (3.156):

$$(\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{A}^T\boldsymbol{A} + \boldsymbol{\Lambda})^{2} = \boldsymbol{A}^T\boldsymbol{M}\boldsymbol{A}\boldsymbol{A}^T\boldsymbol{M}^T\boldsymbol{A},$$

откуда получим значение матрицы коэффициентов Лагранжа:

$$\boldsymbol{\Lambda} = (\boldsymbol{A}^T\boldsymbol{M}\boldsymbol{A}\boldsymbol{A}^T\boldsymbol{M}^T\boldsymbol{A})^{1/2} - \boldsymbol{A}^T\boldsymbol{A}\boldsymbol{A}^T\boldsymbol{A}, \tag{3.157}$$

Найденное значение, т.е. выражение (3.157), подставим в уравнение (3.152) и окончательно получим

$$\boldsymbol{\Phi} = (\boldsymbol{A}^T\boldsymbol{M}\boldsymbol{A}\boldsymbol{A}^T\boldsymbol{M}^T\boldsymbol{A})^{1/2}\boldsymbol{A}^T\boldsymbol{M}\boldsymbol{A}. \tag{3.158}$$

Следует отметить, что матрицы $\boldsymbol{A}^T\boldsymbol{M}\boldsymbol{A}$ и $\boldsymbol{A}^T\boldsymbol{M}^T\boldsymbol{A}$ не являются симметричными, поскольку матрица переходов $\boldsymbol{M}$ наблюдаемых переменных таковой не является. Однако произведение этих матриц есть матрица симметричная и для нее найти корень квадратный – несложно: достаточно извлечь квадратный корень из элементов диагональной матрицы собственных значений симметричной матрицы в ее разложении на произведение собственных векторов и собственных значений.

3.4. Эксплораторный факторный анализ в прогнозировании многомерных временных рядов.

В отличие от обычной регрессионной зависимости, линейной или нелинейной, лаговое соотношение является наиболее приемлемым инструментом для формулирования прогнозного утверждения. Поскольку во многих ситуациях, когда рассматривается экономический процесс, описываемый несколькими переменными, связанными некоторым причинным механизмом, есть основания полагать, что причины предшествуют следствиям. На практике одни события через некоторое время порождают другие, те в свою очередь становятся причинами следующих событий, и так создается цепь причинности. Так, повышение цен ведет к повышению социальных выплат, но повышение выплат в социальной сфере также может привести к повышению некоторых цен. Таким образом, повышение цен в некоторый период времени зависит от повышения цен в предшествующий момент времени, и это пример лаговой, авторегрессионной зависимости.

Когда исследуется система с сильными внутренними взаимосвязями, особенно характерными для экономики, одно уравнение для моделирования динамических процессов будет явно недостаточно. Чтобы построить достоверный прогноз необходимо объяснить два аспекта: первый − динамическое развитие системы, второй – взаимодействие различных переменных.

Поэтому, для решения вопросов, связанных с прогнозированием многомерных временных рядов используется модель, содержащая два уравнения, первое – уравнение переходов факторов, определяющих динамику системы:

$$\boldsymbol{f}_t = \boldsymbol{f}_{t-1}\boldsymbol{\Phi}_{t-1,t}, \tag{3.159}$$

и второе – уравнение определения значений экономических показателей на основе значений факторов, объясняющих взаимосвязи переменных:

$$\boldsymbol{z}_t = \boldsymbol{f}_t\boldsymbol{A}^T + \boldsymbol{v}_t. \tag{3.160}$$

Пусть на основании N наблюдений за переменными, представляющие собой матрицу $\boldsymbol{Z}$ – уже центрированную и нормированную матрицу реальных наблюдений, − найдены оценки факторных нагрузок $\boldsymbol{A}$, диагональной матрицы характерностей $\boldsymbol{D}^2$, матрицы переходов факторов $\boldsymbol{\Phi}$, и значений факторов для реальных наблюдений $\boldsymbol{F}$, которые определяются на основании уравнения

$$\boldsymbol{F} = \boldsymbol{ZA}(\boldsymbol{A}^T\boldsymbol{Z}^T\boldsymbol{ZA})^{-1/2}. \tag{3.161}$$

Тогда прогнозные оценки значений факторов для моментов в будущем $t = 1, 2, \dots, N_p$ ($N_p -$ глубина прогноза) могут быть получены на основании уравнения (3.159), оценки значений показателей – на основании уравнения (3.160). Однако оценка значений факторов может быть уточнена на основании полученного прогнозного значения вектора переменных $\hat{\boldsymbol{z}}_t$ в момент времени t в будущем и оценки значения факторов, получаемой по уравнению (3.161) с добавленным в него, в качестве последней строки, вектором значений переменных $\hat{\boldsymbol{z}}_t$. Таким образом, для получения наиболее достоверной оценки значений факторов $\hat{\boldsymbol{f}}_t$ может быть использована техника фильтра Калмана.

Основные прогнозные уравнения имеют следующий вид

$$\hat{\boldsymbol{f}}_t = \hat{\boldsymbol{f}}_{t-1}\boldsymbol{\Phi} + (\hat{\boldsymbol{z}}_{t-1} - \tilde{\boldsymbol{f}}_{t-1}\boldsymbol{A}^T)\boldsymbol{K}_{t-1} \tag{3.162}$$

и

$$\hat{\boldsymbol{z}}_t = \hat{\boldsymbol{f}}_t\boldsymbol{A}^T. \tag{3.163}$$

В последних двух уравнениях $\tilde{\boldsymbol{f}}_t-$ вектор уточняющих значений факторов, $\boldsymbol{K}_t -$ весовая матрица, $t-1$ означает предшествующий момент времени. Алгоритм получения оценки прогнозных значений экономических показателей состоит из следующей последовательности вычислений. Прежде всего, задаются начальные значения матриц и векторов.

Начальные значения оценок переменных – это N-я строка наблюдений матрицы $\boldsymbol{Z}$:

$$\hat{\boldsymbol{z}}_0 = \boldsymbol{z}_N;$$

а вся матрица $- \boldsymbol{Z}_0 = \boldsymbol{Z}$; (напомним, что заглавной буквой обозначается матрица, т.е. таблица чисел, а прописной – вектор, или матрица-строка);
начальные значения уточняющих факторов – это N-я строка матрицы значений факторов $\boldsymbol{F}$:

$$\tilde{\boldsymbol{f}}_0 = \boldsymbol{f}_N;$$

начальное значение весовой матрицы:

$$\boldsymbol{K}_0 = (\boldsymbol{AA}^T + \boldsymbol{D}^2)^{-1}\boldsymbol{A};$$

начальное значение вспомогательной матрицы $\boldsymbol{P}$ – единичная матрица $\boldsymbol{I}$:

$$\boldsymbol{P}_0 = \boldsymbol{I}.$$

Далее, для моментов времени $t = 1, 2, \dots, N_p$ рекурсивно производятся следующие вычисления:

$$\begin{cases} \hat{\boldsymbol{f}}_t = \hat{\boldsymbol{f}}_{t-1}\boldsymbol{\Phi} + \left(\hat{\boldsymbol{z}}_{t-1} - \tilde{\boldsymbol{f}}_{t-1}\boldsymbol{A}^T\right)\boldsymbol{K}_{t-1}; \\ \hat{\boldsymbol{z}}_t = \hat{\boldsymbol{f}}_t\boldsymbol{A}^T; \\ \boldsymbol{P'}_t = \boldsymbol{\Phi P}_{t-1}\boldsymbol{\Phi}^T; \\ \boldsymbol{K}_t = (\boldsymbol{AP'}_t\boldsymbol{A}^T + \boldsymbol{D}^2)^{-1}\boldsymbol{AP'}_t; \\ \boldsymbol{P}_t = \boldsymbol{P'}_t - \boldsymbol{P'}_t\boldsymbol{A}^T\boldsymbol{K}_t; \\ \boldsymbol{Z}_t = \begin{pmatrix} \boldsymbol{Z}_{t-1} \\ \hat{\boldsymbol{z}}_t \end{pmatrix}; \\ \tilde{\boldsymbol{f}}_t = \hat{\boldsymbol{z}}_t\boldsymbol{A}(\boldsymbol{A}^T\boldsymbol{Z}_t^T\boldsymbol{Z}_t\boldsymbol{A})^{-1/2}. \end{cases} \tag{3.164}$$

В блочной матрице $\boldsymbol{Z}_{N_p}$, получаемой на последнем этапе вычислений, первые N строк – значения переменных, полученные в результате наблюдений, последние N_p строк – это прогнозные значения переменных. Эти значения – центрированные и нормированные. Чтобы получить реальные значения экономических переменных, необходимо преобразовать центрировано-нормированные значения по следующей формуле:

$$\hat{x}_{ti} = \bar{x}_i + \hat{z}_{ti}\sigma_i, \tag{3.165}$$

где $\hat{x}_{ti}$ − прогнозное значение i-й экономической переменной в реальной единицах измерения для -го прогнозного момента времени;

$\bar{x}_i$ – среднее значение i-й экономической переменной в реальных единицах измерения;

$\hat{z}_{ti}$ − прогнозное значение i-й экономической переменной в безразмерных единицах измерения (центрированное и нормированное) для прогнозного момента времени t;

σ_i – среднеквадратичное отклонение i-й переменной.

Если поведение рассматриваемого экономического объекта не является стационарным, а также, если исходные данные имеют распределение вероятностей отличное от нормального, то в (3.165) среднее значение $\bar{x}_i$ следует заменить на математическое ожидание.

Выводы

Рассмотрение методов анализа многомерных временных рядов показал, что традиционные подходы предполагают влияние на данную переменную в данный момент времени как ее значений в предшествующие моменты, так и значений других переменных, и в текущий момент, и в предшествующие. Для этого вводится понятие кросс-корреляции между различными переменными. У спектральной плотности кросс-спектра для двух рядов наряду с действительной частью появляется и мнимая часть, и в этой связи возникают еще две характеристики: когерентность и «усиление». Слабая интерпретабельность получаемых при таком анализе результатов, или ее полная невозможность, делают анализ многомерных временных рядов малопривлекательным и малораспространенным.

Использование эксплораторного факторного анализа позволяет одновременно рассмотреть и авторегрессионные закономерности, и корреляции между исследуемыми переменными.

Предложена модель, состоящая из двух матричных уравнений: динамики факторов и факторной структуры, − которая позволяет преодолеть основные недостатки традиционных методов анализа многомерных временных рядов.

Получена оценка матрицы переходов уравнения динамики стохастических факторов. Обоснована процедура прогноза компонент многомерного временного ряда.

4. ИССЛЕДОВАНИЕ ЭКОНОМИЧЕСКОГО РИСКА С ПОМОЩЬЮ ЭКСПЛОРАТОРНОГО ФАКТОРНОГО АНАЛИЗА

4.1. Риски как экономическая категория

Процесс принятия решений в экономике на всех уровнях управления происходит в условиях постоянно присутствующей неопределенности состояния внешней и внутренней среды, которая обуславливает частичную или полную неопределенность конечных результатов экономической деятельности. В экономике под неопределенностью принято понимать неполноту или неточность информации об условиях хозяйственной деятельности, в том числе о связанных с ней затратах и полученных результатах [63, с. XIII]. Основными причинами неопределенности являются три обстоятельства, носящих и объективный, и субъективный характер. Первое – случайность происходящих событий, имеющих как непосредственное отношение к экономической системе, так и непрямое, но некоторым косвенным образом влияющее на нее. Второе – противодействие со стороны конкурентов и внешней среды, под которой подразумевается система отношений, как на уровне государства, так и на уровне социума, а также противодействия внутри самой экономической системы вплоть до уровня межличностных отношений. Третье обстоятельство состоит в незнании экономических законов, рыночных правил поведения, в игнорировании рекомендаций экономической науки. Неопределенность также объясняется тем, что экономические проблемы главным образом сводятся к решению частных задач выбора решения из множества допустимых и возможных решений, при этом субъект принятия решения зачастую не располагает полной и объективной информацией, а также критерием для принятия оптимального решения.

В современной экономической теории неопределенность трактуется как «двойник» или «предвестник» риска. Основное различие между риском и неопределенностью состоит в том, что риск напрямую связан с последствиями принятия решения в экономической сфере, и риск может иметь место и в том случае, когда неопределенность как такова отсутствует. В остальном, между риском и неопределенностью существует полная тождественность. Именно такой подход характерен для неоклассической школы экономической теории [64]. При этом неопределенность, как и риск, можно разделить по возможности отклонения реального результата от ожидаемого. Если отклонение может вызвать как положительные, так и в отрицательные исходы, то такая неопределенность и риск являются «спекулятивными», если же отклонения могут быть только в меньшую сторону, то в этом случае риск и неопределенность – «чистые».

В современной экономической теории под риском, в основном, понимается возможность потери части своих ресурсов, недополучения доходов или появления дополнительных расходов в результате осуществления предпринимательской деятельности [65, с.18]. Это определение соответствует понятию чистой неопределенности.

Экономический риск является измеримой величиной, в отличие от неопределенности, на чем настаивают авторы одних из последних литературных источников [65]. Однако существуют меры неопределенности – энтропии, которые приводят к определенным информационным мерам, таким, как например, мера Шеннона [23]. К настоящему моменту работы, посвященные применение этих мер к экономическим объектам и явлениям, отсутствуют, причиной чего может служить увлеченность исследователей качественным – не количественным, анализом проблем неопределенности.

Основным измерителем риска является вероятность события, каким-либо образом связанного с риском. Существует определение, согласно которому экономический риск – это измеримая вероятность неполучения прибыли, или потери стоимости некоторых, в том числе финансовых активов, недополучения доходов от инвестиционных проектов и т.д. [63, с. XVI]. Однако не для каждого вида риска можно установить вероятность как меру риска, дающую возможность оценить величину вероятных потерь или непредусмотренных прибылей в случае спекулятивного риска, поскольку вероятностный подход предполагает измерение вероятности некоторого события. Так для странового вида риска сложно выявить само неблагоприятное событие, влекущее риск, если, конечно, отбросить из рассмотрения практически невозможные события, например, ликвидация государства. Поэтому при определении риска чаще всего прибегают к более многозначному понятию «возможности».

При решении практических вопросов наибольшее внимание уделяется не вероятности неблагоприятного события как таковой, а «подверженности» риску, оцениваемой с позиции стоимости и выражаемой с помощью таких показателей как максимальная сумма потерь в результате конкретного фактора риска, средняя за определенный период величина возможных убытков по конкретному виду операций, среднеквадратическое отклонение выручки от реализации и т.п. Подверженность риску, однако, нельзя рассчитать без использования величины вероятности наступления неблагоприятного события. И это – с одной стороны, с другой же стороны необходимо знать масштаб возможного ущерба в результате такого события с точки зрения чувствительности к его последствиям.

Риск пронизывает все стороны экономической жизни. Он сопутствует всем этапам и стадиям жизненного цикла любой экономической системы. Игнорирование или пренебрежение риском всегда ведет к негативным последствиям не только для экономического субъекта, напрямую страдающего от него, но и для связанных с этим субъектом других экономических систем. А в силу глобальности, переплетенности и интегрированности мировой экономики крах какого-либо более-менее значимой экономической фигуры может привести к эффекту «домино», что подтверждает финансовый кризис, разразившийся в последние годы. В силу чего риск становится экономической категорией, без учета которой невозможно исследование поведения любой экономической системы.

Проблема анализа и учета риска с целью минимизации возможных потерь существует во всех секторах экономики – от промышленности и сельского хозяйства до торговли и туризма. Однако все отрасли связаны между собой посредством финансовой сферы, поэтому именно финансовые риски играют заметную роль в формировании атмосферы риска экономической деятельности. При этом главными угрозами для благополучия экономического субъекта являются рыночные, операционные, кредитные риски, а также риски ликвидности и риски событий [63., с. XV].

Рыночный риск с позиции «чистой» неопределенности представляет собой возможность отрицательного изменения стоимости активов в результате колебаний конъюнктуры рынка, изменений процентных ставок, курсов валют, цен акций, облигаций, товарных и финансовых контрактов. Разновидностями рыночного риска являются, в частности, валютный и процентный риски, которые иногда рассматриваются как самостоятельные виды риска.

Кредитный риск или риск контрагента – возможность потерь в результате неспособности контрагентов исполнять свои обязательства, в частности по выплате процентов и основной суммы долга, в соответствии со сроками и условиями кредитного договора. К кредитному риску относятся также риск дефолта и риск досрочного погашения. Из определения следует, что этот вид риска имеет непосредственное отношение к финансовым институтам, имеющим полномочия на осуществление операций по выдаче кредитов. Но непосредственным источником кредитного риска является финансовое положение заемщиков, поэтому уровень кредитного риска, оцененный по всем кредитно-финансовым учреждениям, характеризует также уровень рискованности предпринимательской сферы в целом.

Как кредитный, так и рыночный риск имеет отношение к одним тем же экономическим субъектам, поэтому в силу этой общности эти риски не могут не влиять друг на друга, и, соответственно, являются взаимозависимыми. Однако кредитный риск, в отличие от рыночного риска, является ассиметричным по своей природе. Это означает, что потенциальный выигрыш при операциях кредитования не может превысить некоторую относительную положительную доходность, верхняя граница которой определена кредитным договором, и ни один заемщик не заплатит кредитору больше договорной суммы. Но при этом потенциальный убыток кредитора может колебаться в значительно более широком диапазоне: от нуля до суммы основного долга, причем к возможным потерям в последнем случае могут добавиться судебные издержки, недополученная прибыль, штрафы и пени при просрочке возврата привлеченных для кредитования средств, а также расходы, связанные с дополнительным привлечением средств для исполнения текущих обязательств кредитора.

Таким образом, несмотря на взаимосвязанность рыночного и кредитного рисков, измерение одного не влечет за собой возможности хотя бы косвенной оценки другого.

Другие виды рисков также имеют собственную специфику, которая не позволяет ввести комплексное измерение уровня рискованности, и не дает возможности по оценке риска одного вида судить о величине риска другого вида. Так риск ликвидности делится на два подвида: риск рыночной ликвидности и риск балансовой ликвидности. Первый означает возможность потерь, вызванных неосуществимостью реализовать актив в необходимом количестве и за достаточно короткий промежуток времени. Второй – возможность возникновения дефицита высоколиквидных средств для выполнения обязательств перед контрагентами. Хотя риск ликвидности и зависит от рыночного риска, но однозначным образом не определяется им, поскольку отсутствие высоко ликвидных активов для исполнения текущих обязательств мало связано с конъюнктурой рынка. Конечно, дефицит у конкретного экономического субъекта может определяться ситуаций на рынке, но прямой зависимости для этих видов риска нет.

Операционный риск состоит в возможности непредвиденных потерь вследствие технических ошибок при проведении операций различного характера, умышленных и неумышленных действий персонала, аварийных случаев, сбоев оборудования, несанкционированного доступа к базам данных и т.п. Операционные риски также определяются возможностью убытков, обусловленных неправильным выбором методов оценки риска и управлением им. Этот вид риска имеет ярко специфический характер и увязать его с другими видами риска достаточно сложно.

Риск события – это возможность непредвиденных потерь вследствие форс-мажорных обстоятельств, изменений в законодательстве, действий государственных органов. К рискам события также относятся часто трактуемые как отдельные виды риска юридические и налоговые риски, риск репутации и риск действий фискальных и регулирующих органов. Этот вид риска, как и в значительной мере операционный риск наиболее трудно поддается формализации и количественной оценке. Объясняется это тем, что эти риски во многом обусловлены так называемым «человеческим фактором» [63,с.XVI].

Указанные риски имеют разную значимость для разных экономических субъектов из-за различной специфики предпринимательской деятельности. У кредитно-финансовых учреждений наибольшие потери происходят вследствие кредитных и рыночных рисков, у субъектов, занимающихся клиринговыми операциями, на передний план выходят операционные риски и риски контрагента. Свою особую специфику имеют предприятия промышленности, торговли, сферы услуг. Все они подвержены систематическим рискам, обусловленным отраслевой принадлежностью и особенностями производственного процесса. Эти риски, называемые техногенными или производственными, имеют некоторую коррелированность с рыночным, кредитным, операционным, риском ликвидности и другими видами риска, но эти связь трудно выделить в явной форме, чтобы по уровню величины одного риска судить об уровнях других рисков. Тем более сложно на основании этих показателей прийти к некоторой комплексной оценке уровня риска в целом.

4.2. Оценки экономических рисков

При оценке риска необходимо учитывать две стороны, характеризующие природу возникновения риска. Одна – это объективное функционирование экономической системы, происходящее под воздействием внешних факторов. Другая сторона определяется субъективными обстоятельствами, сопутствующими выбору стратегий, принятию решений в риск-менеджменте, поскольку определение понятия риск включает в себя следующую сторону: «риск – это возможность несоответствия характеристик экономического состояния объекта значениям, ожидаемым лицами, принимающими решения под действием рыночных факторов» [63, с. 204]. Поэтому измерители риска традиционно делятся на основании двух аспектов риска: первый – изменчивость или волатильность экономических индикаторов, частота событий или их вероятность; второй – чувствительность или зависимость стратегии предпринимательской деятельности, критериев принятия решений к их последствиям. Соответственно измерители риска делятся на две основные категории: 1) вероятностные или статистические показатели; 2) величины, характеризующие чувствительность.

Такое деление является достаточно условным, поскольку вероятностные оценки риска достаточно часто могут иметь информацию о чувствительности. Например, выбор стратегии поведения на рынке, основанной на тезисе − чем выше риск, тем большим становится выигрыш, в значительной степени коррелирует с таким положением теории вероятностей, в котором утверждается, что чем больше интервал, чем больше вероятность попадания туда значения случайной величины. С другой стороны показатели чувствительности могут рассматриваться как вероятностные оценки, особенно в случае, когда такой показатель представляет собой случайную величину, характеристики которой несут больше информации, чем непосредственно сам показатель. Деление на эти две категории становится условным в большей степени, если учесть, что оценкам экономического риска присуще свойство субъективности, поскольку любая оценка становится измерителем риска лишь в процессе ее интерпретации.

Если рассматривать риск с позиции вероятностной категории, то в этом случае наиболее обоснованно характеризовать и измерять его как вероятность возникновения определенного уровня потерь. Таким образом, при обстоятельной всесторонней оценке риска следовало бы устанавливать для каждого абсолютного или относительного значения величины возможных потерь соответствующую вероятность возникновения такой величины. Построение таблицы, содержащей значения возможных потерь, приходящихся на определенный интервал, с соответствующими значениями вероятностей – для случая дискретной случайной величины, или кривой вероятностей потерь для непрерывного случая, является исходной стадией оценки риска. Но применительно к анализу экономической системы сверхсложного типа это чаще всего чрезвычайно сложная задача. Поэтому на практике ограничиваются упрощенными подходами, оценивая риск по одному или нескольким главным показателям или критериям, представляющим обобщенные характеристики, наиболее важные с точки зрения о приемлемости риска. С этой целью выделяются определенные области или зоны риска в зависимости от величины потерь.

Область, в которой потери не ожидаются, считаемся безрисковой областью, ей соответствует отсутствие потерь. Под зоной допустимого риска обычно понимается область, в пределах которой данный вид предпринимательской деятельности сохраняет свою экономическую целесообразность, т.е. потери есть, но они могут быть компенсированы из различных источников, например, за счет резервных фондов или страхования. Границы зоны допустимого риска соответствуют уровню потерь, равному расчетной прибыли от предпринимательской деятельности. Следующая область − зона критического риска. Это область, характеризуемая возможностью потерь в размере свыше величины ожидаемой прибыли и вплоть до величины полной расчетной, ожидаемой выручки от предпринимательства. Зона критического риска характеризуется опасностью потерь, которые заведомо превышают ожидаемую прибыль и могут привести к невозмещаемой утере всех средств, вложенных предпринимателем в дело. В последнем случае экономически субъект не только не получает от сделки никакого дохода, но и несет убытки в сумме всех бесплодных затрат.

Помимо критического рассматривается еще более устрашающий – катастрофический риск. Зона катастрофического риска представляет собой область потерь, которые по своей величине превосходят критический уровень и в пределе могут достигать величины, равной имущественному состоянию предпринимателя. Катастрофический риск способен привести к краху, банкротству, полному крушению предприятия, его закрытию и распродаже имущества. Вне зависимости от имущественного или денежного ущерба к категории катастрофического относится риск, связанный с прямой опасностью для жизни людей или с возникновением экологических катастроф. Потери, превышающие имущественное состояние предпринимателя, не рассматриваются, так как их невозможно взыскать.

Вероятности определенных уровней потерь являются важными показателями, позволяющими высказать суждение об ожидаемом риске и его приемлемости. Построенную кривую распределения вероятностей потерь дохода принято назвать кривой риска. Например, если вероятность катастрофической потери выражается значением вероятности, равно, скажем 0,2; свидетельствующей об ощутимой угрозе потери всего состояния, то рассудительный осмотрительный бизнесмен заведомо откажется от такого дела и не пойдет на подобный риск. Поэтому в соответствии с общепринятым подходом, если при оценке риска предпринимательской деятельности удается построить не всю кривую вероятностей риска, а только установить характерные точки – вероятность нулевых потерь, наиболее вероятный уровень риска и вероятности допустимой, критической, катастрофической потери, − задачу оценки считают успешно решенной. Значения этих показателей считается достаточно, чтобы в большинстве случаев идти с «открытыми глазами на обоснованный риск» [66, с. 38].

К числу традиционных способов оценки риска относятся статистический, экспертный, расчетно-аналитический.

Статистический способ состоит в том, что изучается статистика потерь, имевших место в аналогичных видах предпринимательской деятельности, устанавливается частота появления определенных уровней потерь. Если объем такой статистической информации достаточен, то частоту возникновения данного уровня потерь можно в первом приближении приравнять к вероятности их возникновения и на этой основе построить кривую вероятностей потерь, которая и есть искомая кривая риска. При этом необходимо учитывать одно обстоятельство. Определяя частоту возникновения некоторого уровня потерь путем деления числа соответствующих случаев на их общее число, следует включать в общее число случаев и те предпринимательские сделки, в которых потерь не было, а имел место выигрыш, т.е. превышение расчетной прибыли. В противном случае показатели вероятностей потерь и угрозы риска окажутся завышенными.

Экспертный способ, известный под названием метода экспертных оценок, применительно к экономическому риску реализуется путем обработки мнений опытных экспертов или специалистов. Наиболее желательно, чтобы они дали свои оценки вероятностей возникновения определенных уровней потерь, по которым затем можно было бы найти средние значения экспертных оценок и с их помощью построить кривую распределения вероятностей.

При этом также, как и в статистическом методе, ограничиваются получением экспертных оценок вероятностей возникновения определенного уровня потерь в четырех характерных точках, т.е. экспертным образом устанавливаются показатели наиболее вероятных, допустимых, критических и катастрофических потерь, имея в виду, как их уровни, так и вероятности. По этим четырем характерным точкам предпринимается попытка воспроизвести ориентировочно всю кривую распределения вероятностей потерь. Естественно, при небольшом, как правило, массиве экспертных оценок объем информации для расчета частот недостаточен, и кривая вероятностей такого графика является слишком приблизительной. Достоинство этого способа состоит в одном, что все же определенное представление о риске и характеризующих его показателях можно получить, а это уже значительно лучше, чем не знать ничего. Хотя любая комбинация субъективных мнений вряд ли может считаться объективной.

Расчетно-аналитические методы построения кривой распределения вероятностей потерь и оценки на этой основе показателей экономического риска базируются на теоретических представлениях. Однако, прикладная теория риска хорошо разработана только применительно к страховому и игровому риску. Элементы теории игр, в принципе, применимы ко всем видам экономического риска, но следует констатировать, что прикладные математические методы оценочных расчетов производственного, коммерческого, финансового риска на основе теории игр пока не созданы.

Помимо указанных методов определения степени риска часто используют следующие способы его оценки:

- в ряде случаев меру риска, как степень ожидаемой неудачи при неуспехе в процессе достижения цели, определяют через соотношение вероятности неуспеха и степени неблагоприятных последствий, которые могут наступить в этом случае. Однако при таком подходе сложности удваиваются, поскольку необходимо оценить и вероятность неблагоприятного исхода, и степень ущерба при его наступлении;

- степень риска иногда определяют как величину, равную произведению ожидаемого ущерба на вероятность того, что ущерб произойдет. В связи с установлением взаимосвязи между величиной риска выбираемого решения, а также возможным ущербом, наносимым этим решением, и очевидностью, с которой ущерб причиняется, предполагается, что наилучшим является решение с минимальным значением определяемой величины произведения. Вероятность технического и коммерческого успеха, т.е. учет риска и оценка его степени, определяют в зависимости от характера продукции, которую предполагают получить в результате реализации;

- в ряде случаев для определения степени риска и выбора оптимальных решений применяют методику «дерево решений», Она предполагает графическое построение различных вариантов, которые могут быть приняты, По «ветвям дерева» соотносят субъективные и объективные оценки данных событий (экспертные оценки, размеры потерь и доходов и т.д.). Следуя вдоль построенных «ветвей дерева», используя специальные методики расчета вероятностей, оценивают каждый вариант пути. Риск определяют как сумму ущерба, нанесенного вследствие неверного решения, и расходов, связанных с реализацией данного решения.

Все эти методы предполагают использование лишь одного экономического показателя в качестве носителя информации о риске и никак не учитывают причины или внешние факторы, влияющие на уровень рискованности. Поэтому вероятностные подходы к оценке риска не могут дать в качестве результата некоторое обобщенное представление того, что может быть с экономической системой с течением времени, поскольку эти оценки являются статическими. Плотность распределения вероятностей случайной величины не является функцией времени, и потому не может быть использована в качестве инструмента прогноза уровня вероятности на некоторый момент времени в будущем.

Вторая категория методов оценки риска относится к измерению чувствительности критериев принятия решений к его результатам. Показатели чувствительности разработаны в достаточно полном объеме для измерения уровня финансового риска. Измерителей чувствительности по отношению к движению финансовых переменных существует обширное число, обозначаемое, главным образом, греческими буква, отчего их часто называют «греческими показателями». На рынке инструментов с фиксированным доходом чувствительность к движению процентных ставок измеряется дюрацией. На рынке акций чувствительность к фактору рынка в целом (например, фондовому индексу) называется систематическим риском и обозначается коэффициентом бета. На рынке производных инструментов чувствительность к изменению цены базового актива измеряется коэффициентом дельта, определяющим величину изменения стоимости данного финансового инструмента при малом изменении базового ценового фактора – ставки процента или курса валюты.

Показатели, по своему математическому определению являющиеся производными второго порядка, называются выпуклостью на рынке инструментов с фиксированным доходом и коэффициентом гамма на рынке производных инструментов. Выпуклость измеряет изменчивость дюрации по мере изменения процентной ставки. Аналогично гамма измеряет изменения дельты при изменении цены базового актива. Оба показателя измеряют чувствительность второго порядка (или квадратичную чувствительность) к изменениям финансовых переменных.

Существует множество иных показателей риска, применяемых по отношению к производным инструментам: вега – измеряет изменение стоимости финансового инструмента при изменении волатильности базового ценового фактора; тета – изменение стоимости финансового инструмента при изменении срока, оставшегося до его исполнения; ро – показатель изменения стоимости процентных опционов при изменении процентной ставки; лямбда – изменение стоимости опционов на акции при изменении величины дивидендов. Все эти показатели – производные второго порядка. Используются показатели – производные третьего порядка, отслеживающие изменение вл времени показателей − производных второго порядка. Это «скорость», «очарование», «цвет».

Данные показатели служат для измерения риска в достаточно специфической области фондового рынка и с очень большой степенью натяжки могут рассматриваться в качестве интегрированных измерителей. Лишь показатель бетта несет некоторую информацию об уровне рискованности в экономической системе.

Экономический анализ деятельности предприятий и эффективности инвестиционных проектов на основе бухгалтерской и управленческой отчетности~~)~~, использует следующие показатели, которые фактически являются измерителями риска: запасы, разрывы, коэффициенты ликвидности, финансовой устойчивости, плечо финансового рычага, ~~и~~ плечо производственного рычага, коэффициенты эластичности различных экономических показателей по соответствующим факторам и т.~~–~~д. Все эти измерители в той или иной степени характеризуют чувствительность ~~(~~или порог чувствительности~~)~~ критериев эффективности экономической деятельности к изменениям внутрихозяйственной и внешней, ~~(~~рыночной~~)~~ конъюнктуры [63]. ~~Эти показатели определяют приоритет изучаемых и контролируемых параметров, они помогают вскрыть взаимосвязи и логические зависимости между факторами рисков.~~

Другими показателями риска являются ~~различные косвенные показатели~~: рейтинги ценных бумаг, рейтинги заемщиков, рынков, государств; премии за риск, содержащиеся в доходности различных активов; котировки производных финансовых инструментов; параметры дефицита ~~(длина и продолжительность очереди, объем запасов)~~ и т.д.

Наиболее распространенные подходы к оценке экономического риска основаны на вычислении вероятности появления неблагоприятных событий. Однако вероятность – это теоретическое значение частоты события, и ее нахождение возможно лишь на основе знания закона распределения случайной величины, на основании которого определяется риск. В условиях неопределенности череды событий экономической реальности говорить о стабильности закона распределения вероятностей некоторого показателя достаточно сложно; более того, брать за основу поведение лишь одной переменной величины означает значительно уменьшить достоверность выводов и рекомендаций для принятия решений для управления экономической системой. Поэтому актуальной представляется задача поиска альтернативных подходов для оценки риска. ~~Все эти показатели статистическими, экспертными либо рыночными оценками характеризуют риски активов.~~

4.3. Оценка риска с помощью эксплораторного стохастического факторного анализа

Практическое использование эксплораторного факторного анализа показывает, что на любую экономическую систему влияют разнообразные стохастические факторы. Среди них можно выделить факторы, учитывающие макроэкономические тенденции, государственную политику, финансовые аспекты, рыночные и другие отношения, а также внутрисистемные особенности, и, в частности, так называемый, человеческий фактор. Каждый из них имеет непосредственное отношение к уровню рискованности, проявляемой в многочисленных сферах экономической жизни. Каждый стохастический фактор может способствовать стабильности развития экономической системы, а может явиться причиной сбоев в ее поведении. Поэтому количественное измерение фактора может ответить на вопросы: какое воздействие в данный момент времени данный фактор оказывает на экономическую систему – позитивное или негативное, и какой вклад этот фактор делает в общий уровень экономического риска.

В соответствии с общей моделью эксплораторного стохастического факторного анализа, непосредственно наблюдаемые экономические переменные (показатели) связаны прямой зависимостью со стохастическими факторами в момент времени t:

$$\boldsymbol{z}_t = \boldsymbol{f}_t \boldsymbol{A}^T + \boldsymbol{u}_t\,, \tag{4.1}$$

поэтому отрицательное значение фактора ведет к уменьшению значения переменной. Напомним, в модели (4.1) переменные $\boldsymbol{z}_t$ – нормированные величины, т.е. разность наблюдаемого значения переменной и ее математического ожидания для момента времени t отнесена к ее среднеквадратическому отклонению; значения факторов также являются нормированными величинами и находятся в промежутке от минус единица до плюс единица, что является следствием свойства ортогональности стохастических факторов:

$$\boldsymbol{f}_t^T \boldsymbol{f}_t = \boldsymbol{I};\ \forall t\,, \tag{4.2}$$

где $\boldsymbol{I}$ – единичная матрица.

Таким образом, нулевые значения стохастических факторов не приводят к изменениям переменных, отрицательные значения – к уменьшению наблюдаемых показателей, а положительные – к росту экономической системы. Интересным представляется случай, когда часть стохастических факторов положительна, а другая часть – отрицательна. Тогда положительные значения факторов должны компенсировать отрицательные, что согласуется со здравым смыслом. Это с одной стороны, а с другой стороны – стохастические факторы являются ортогональными, независимыми величинами, что формально делает предыдущее утверждение несправедливым, поскольку взаимное воздействие у независимых величин отсутствует и, строго говоря, значение одной ортогональной составляющей не может привести к изменению величины другой ортогональной составляющей. Однако на отдельный экономический показатель оказывают влияние практически все стохастические факторы, именно их общая линейная (для зависимости (4.1)) комбинация определяет значение переменной, поэтому, действительно, в совокупном факторном воздействии положительные факторные значения компенсируют отрицательные.

Следовательно, для измерения уровня риска конкретной экономической системы необходимо учитывать значения всех стохастических факторов, определяющих поведение этой системы, и мерой ее совокупного риска будет служить некоторая функция нескольких аргументов, представляющих собой числовые величины этих факторов, а также одним из аргументов является временной интервал или момент времени t. В общем виде мера риска может быть в следующей форме:

$$Mr_t = \boldsymbol{\varphi}(\boldsymbol{f}_t, t); \tag{4.3}$$

где Mr_t – мера риска;

$\boldsymbol{\varphi}$ – функция многих переменных;

$\boldsymbol{f}_t = (f_{t,1}\quad f_{t,2}\quad \dots\quad f_{t,m})$ – вектор значений стохастических факторов;

m – число стохастических факторов, описывающих данную экономическую систему.

Функция $\boldsymbol{\varphi}$ должна принимать значения, которые удобно было бы интерпретировать так, как например, коэффициент парной корреляции или вероятность некоторого события, т.е. нулевое значение соответствует отсутствию каких-либо рискованных явлений и уровень риска равен нулю, близость к единице должно говорить о том, что уровень риска катастрофический. При этом в основе меры риска должна лежать аддитивная функция значений стохастических факторов. В качестве такой функции может выступить сумма значений факторов, если предположить, что все стохастические факторы в равной мере являются важными для формирования динамики экономической системы:

$$sf_t = \sum_{j=1}^{m} f_{t,j}\,; \tag{4.4}$$

где sf_t – сумма значений стохастических факторов для момента времени t;

$f_{t,j}$ – значение j-го стохастического фактора в момент времени t.

Если для развития экономической системы некоторые факторы являются более предпочтительными перед другими, то функцию (4.4) можно легко адаптировать посредством ввода весовых коэффициентов.

Поскольку для значений факторов выполняется условие (4.2), то квадрат значения фактора не может превышать значения, равного единице:

$$f_{t,j}^2 \le 1;\ \ \forall j.$$

Таким образом, для значения каждого фактора должно выполняться условие

$$-1 \le f_{t,j} \le +1;$$

для всех моментов времени $\forall t$, и максимальная сумма значений стохастических факторов не может превышать числа слагаемых, т.е. числа факторов:

$$\max\{sf_t\} \le m,$$

а минимальное значение функции (4.4) должно быть больше, чем минус число факторов

$$\min\{sf_t\} \ge -m.$$

Функция φ должна принимать минимальное значение, когда сумма sf_t максимальна, и минимальное в противном случае. То есть функция φ определена на интервале $[-m; m]$ для ее аргумента – суммы значений стохастических факторов. Тогда, с учетом условия, что

$$0 \le Mr_t \le 1,$$

мера риска имеет вид:

$$Mr_t = \frac{m - sf_t}{2m};$$

или

$$Mr_t = \frac{m - \sum_{j=1}^{m} f_{t,j}}{2m}. \qquad (4.5)$$

Проверим свойства функции (4.5). В случае, когда все стохастические факторы принимают наименьшее значение минус единица, то их сумма равна $-m$, и мера риска становится равной

$$Mr_t = \frac{m-(-m)}{2m} = \frac{2m}{2m} = 1,$$

тем самым показывая, что уровень риска имеет максимальное значение. Тогда же, когда факторы в наибольшей степени содействуют развитию экономической системы, сумма их значений – максимальна и равна m, величина меры риска имеет вид

$$Mr_t = \frac{m-m}{2m} = \frac{0}{2m} = 0,$$

что соответствует минимальному уровню риска.

Интересным является момент, когда все стохастические факторы не оказывают никакого воздействия на системы, т.е. имеют нулевые значения, или, когда их сумма равна нулю, т.е. негативное влияние одних факторов компенсируется позитивным воздействием других. В этом случае мера риска равна

$$Mr_t = \frac{m-0}{2m} = \frac{m}{2m} = 0{,}5,$$

что соответствует серединному уровню риска, но не нулевому. Тем самым, отсутствие активности в механизме развития системы, или равенство нулю равнодействующей всех причин развития системы не приводит к отсутствию риска как такового. Таким образом, данная мера риска в большей степени согласуется со здравым смыслом, чем традиционные, устоявшиеся представления о механизмах и оценках риска. К примеру, широко известен тезис о том, что риск не возникает в случае закрытой валютной позиции [66], т.е. тогда, когда требования соответствуют обязательствам, покупки и продажи валюты отсутствуют или в точности совпадают и по размерам, и по времени, что практически не является возможным. При закрытой валютной позиции изменение курса валюты – главный источник валютного риска, – не оказывает никакого влияния на убытки или доходы банка или предприятия, ведущего внешнеэкономическую деятельность. Однако в этом случае риск все-таки имеет место, и может рассматриваться как риск недополучения доходов от неторговых валютных операций, отсутствующих при закрытой валютной позиции, поскольку для получения дохода, необходимо открыть валютную позицию. Недополученный доход можно рассматривать как непрямые или косвенные убытки хозяйствующего субъекта, что является проявлением риска.

С точки зрения меры риска (3.147) в этом случае фактор активности в проведении валютных операций равен нулю, но величина $Mr_t = 0{,}5$, т.е. риск имеет место.

Если каждый стохастический фактор представить в виде оси m-мерной ортогональной системы координат, то пространство, положение в котором будет определяться посредством этой системы координат, можно рассматривать как пространство риска. Например, на рисунке 4.1 изображено пространство риска, когда поведение экономической системы определяется двумя факторами. В этом случае имеется четыре области риска, задаваемые квадрантами двумерного пространства. Первый квадрант соответствует безрисковой зоне, третий квадрант – зоне максимального риска, второй и третий указывают на наличие проявлений риска со стороны, соответственно, первого и второго факторов.

Близкое к единице значение фактора указывает на максимальное удаление от проявлений риска, а значение фактора, равное минус единица – на максимальную близость к риску. Поэтому точка A(1;1) соответствует положению экономической системы с максимальным удалением от риска, а точка

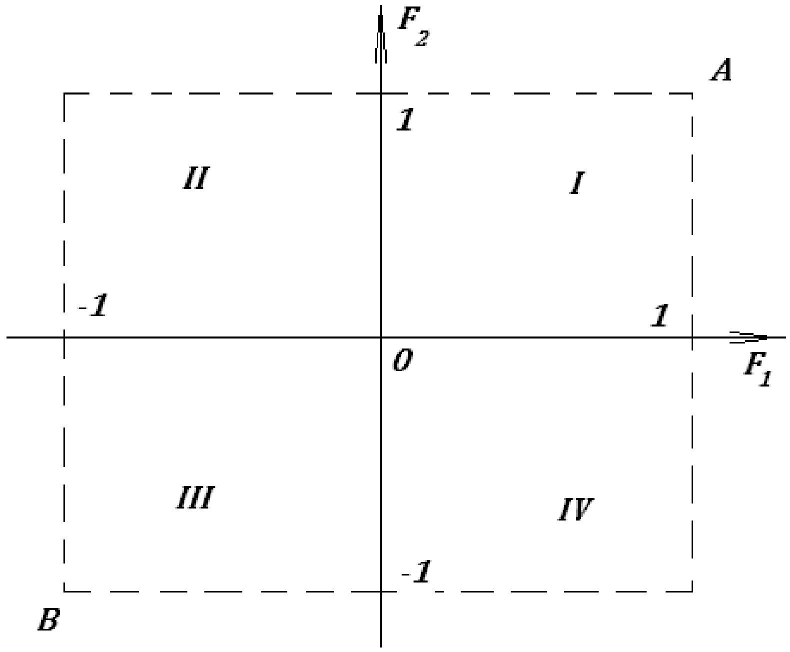

Рис. 4.1 – Пространство риска в случае двух факторов

B(-1;-1) – точка катастрофического риска. Координатами точки в факторном пространстве служат значения факторов, и по этим координатам можно судить в какой зоне риска находится экономическая система, если рассматривать это пространство как пространство риска.

4.4. Прогноз состояния экономической системы с позиций риска

Значения стохастических факторов для данного момента времени однозначно указывают на общий уровень риска с помощью меры риска (4.5), а также по величинам факторов, как по координатам пространства риска, можно судить в какой зоне риска находится экономический субъект. Рассмотрение поведения значений стохастических факторов может указать на «рисковую» динамику экономической системы. Годограф движения системы по пространству риска может служить основой для прогноза состояния экономической системы с позиций риска. Однако графический прогноз возможен лишь для обозримого числа факторов: одного, двух или трех. Если число стохастических факторов больше трех, то визуальное представление годографа невозможно. Поэтому необходим аналитический метод предсказания будущего поведения значений стохастических факторов.

Значения экономических показателей (случайных переменных) в текущий момент времени не может не зависеть от их значений в предшествующие моменты времени. В таких временных зависимостях прослеживается как природа объективных экономических законов, так и внутренних, присущих лишь данной экономической системе, особенностей. Лаговые зависимости являются отличительной чертой экономических и финансовых процессов, и именно они, в основном, определяют характерные черты динамики экономических систем. Таким образом, для экономических систем является характерным и определяющим авторегрессионный процесс.

Но поскольку сами экономические переменные являются лишь измеряемыми и фиксируемыми характеристиками системы, а за этими переменными стоят стохастические факторы, которые и определяют ее поведение, то именно они задают авторегрессионный процесс в экономической системе.

Пусть нам известны значения стохастических факторов эксплораторной факторной модели для $(N+p-1)$ моментов времени ($N-$ «текущих» моментов времени, и p – предшествующих, включая некоторый «нулевой» момент времени). Рассмотрим многомерную факторную авторегрессионную модель порядка p:

$$\boldsymbol{f}_t = \boldsymbol{f}_{t-1}\psi_1 + \boldsymbol{f}_{t-2}\psi_2 + \dots + \boldsymbol{f}_{t-p}\psi_p + \boldsymbol{u}_t; \tag{4.6}$$

где $\boldsymbol{f}_t = (f_{t,1} \quad f_{t,2} \quad \dots \quad f_{t,m})$ – значения стохастических факторов для момента времени t; m – число факторов;

аналогично $\boldsymbol{f}_{t-k} = (f_{t-k,1} \quad f_{t-k,2} \quad \dots \quad f_{t-k,m})$ – значения стохастических факторов для момента времени $t-k; (k = 1,2,\dots,p)$;

$\boldsymbol{\psi}_k = (\boldsymbol{\psi}_1 \quad \boldsymbol{\psi}_2 \quad \dots \quad \boldsymbol{\psi}_m)$ – значения коэффициентов многомерной авторегрессии;

$\boldsymbol{u}_t = (u_{t1} \quad u_{t2} \quad \dots \quad u_{tm})$ – вектор случайных отклонений с нулевым математическим ожиданием и постоянными дисперсиями.

Тогда уравнение (4.6) может быть представлено в виде:

$$\boldsymbol{f}_t = (\boldsymbol{f}_{t-1} \quad \boldsymbol{f}_{t-2} \quad \dots \quad \boldsymbol{f}_{t-p}) \begin{pmatrix} \psi_1 \\ \psi_2 \\ \dots \\ \psi_p \end{pmatrix} + \boldsymbol{u}_t. \tag{4.7}$$

Введем обозначения, обозначим матрицу значений факторов для моментов времени от момента времени 1 до текущего t

$$\boldsymbol{F}_t = \begin{pmatrix} \boldsymbol{f}_1 \\ \boldsymbol{f}_2 \\ \dots \\ \boldsymbol{f}_t \end{pmatrix};$$

матрица значений факторов с временными сдвигами на k шагов от текущего момента времени

$$\boldsymbol{F}_{t-k} = \begin{pmatrix} \boldsymbol{f}_{1-k} \\ \boldsymbol{f}_{2-k} \\ \dots \\ \boldsymbol{f}_{t-k} \end{pmatrix}.$$

Также введем блочную матрицу лаговых значений факторов

$$\Xi_{t-p} = \left(\boldsymbol{F}_{t-1} \mid \boldsymbol{F}_{t-2} \mid \dots \mid \boldsymbol{F}_{t-p} \right),$$

и блочную матрицу коэффициентов авторегрессии

$$\Theta = \begin{pmatrix} \psi_1 \\ \hdashline \psi_2 \\ \hdashline \dots \\ \hdashline \psi_p \end{pmatrix}.$$

Тогда авторегрессионная зависимость (4.7) для произвольного момента времени примет вид:

$$\boldsymbol{F}_t = \Xi_{t-p}\Theta + \boldsymbol{U}_t, \tag{4.8}$$

где $\boldsymbol{U}_t = \begin{pmatrix} \boldsymbol{u}_1 \\ \boldsymbol{u}_2 \\ \dots \\ \boldsymbol{u}_e \end{pmatrix}$ - матрица случайных отклонений.

Ставится задача получить оценку матрицы коэффициентов авторегрессионной модели (4.8) так, чтобы минимизировать сумму квадратов ошибок, но при этом восстанавливаемые по модели значения факторов должны быть ортогональны. Предположим, что имеются значения факторов для $N+p$ моментов времени: $-p+1; -p+2; \dots; -1; 0; 1; 2; \dots; t; \dots; N$. Обозначим матрицы

$$\boldsymbol{F} = \boldsymbol{F}_N;\ \Xi_{-p} = \Xi_{N-p}.$$

Тогда задача оценивания может быть сформулирована следующим образом: найти Θ так, чтобы минимизировать $tr\,(\boldsymbol{F} - \Xi_{-p}\Theta)^T(\boldsymbol{F} - \Xi_{-p}\Theta)$ при условии, что $\boldsymbol{F}^T\boldsymbol{F} = \boldsymbol{I}$ или $(\Xi_{-p}\Theta)^T\Xi_{-p}\Theta = \boldsymbol{I}$. Т.е. найти в классе ортогональных оценок такие, которые минимизировали бы функцию потерь. Для этого составим матричную функцию Лагранжа:

$$g = tr\,(\boldsymbol{F} - \Xi_{-p}\Theta)^T(\boldsymbol{F} - \Xi_{-p}\Theta) + tr\,\Lambda\left((\Xi_{-p}\Theta)^T\Xi_{-p}\Theta - \boldsymbol{I}\right). \tag{4.9}$$

Производная функции Лагранжа по матрице Θ равна

$$\frac{\partial g}{\partial\Theta} = -2(\boldsymbol{F} - \Xi_{-p}\Theta)^T\Xi_{-p} + 2\Xi_{-p}^T\Xi_{-p}\Theta\Lambda;$$

а производная функции Лагранжа по матрице коэффициентов Лагранжа Λ имеет вид

$$\frac{\partial g}{\partial\Lambda} = (\Xi_{-p}\Theta)^T\Xi_{-p}\Theta - \boldsymbol{I}.$$

Таким образом, искомая оценка должна удовлетворять системе матричных уравнений, получаемой от приравнивания нулю матричных производных:

$$\begin{cases} (\boldsymbol{F} - \Xi_{-p}\boldsymbol{\Theta})^T \Xi_{-p} + \Xi_{-p}^T\Xi_{-p}\boldsymbol{\Theta}\boldsymbol{\Lambda} = 0; \\ \quad (\Xi_{-p}\boldsymbol{\Theta})^T \Xi_{-p}\boldsymbol{\Theta} - \boldsymbol{I} = 0. \end{cases} \tag{4.10}$$

Преобразуя уравнения системы (4.10) в векторную форму, и используя свойства произведения векторов на кронекеровские произведения матриц, из первого уравнения системы получим значение матрицы авторегрессионных коэффициентов, выраженных через матрицу коэффициентов Лагранжа:

$$\boldsymbol{\Theta} = \left(\Xi_{-p}^T\Xi_{-p}\right)^{-1}\Xi_{-p}^T\boldsymbol{F}\left(\boldsymbol{I} + \boldsymbol{\Lambda}\right)^{-1}, \tag{4.11}$$

которое после подстановки во второе уравнение системы позволяет найти значение матрицы коэффициентов Лагранжа

$$\boldsymbol{\Lambda} = \left(\boldsymbol{F}^T\Xi_{-p}\left(\Xi_{-p}^T\Xi_{-p}\right)^{-1}\Xi_{-p}^T\boldsymbol{F}\right)^{1/2} - \boldsymbol{I}.$$

Подставляя найденное значение в (4.5+6), окончательно получим оценки коэффициентов авторегрессии:

$$\hat{\boldsymbol{\Theta}} = \left(\Xi_{-p}^T\Xi_{-p}\right)^{-1}\Xi_{-p}^T\boldsymbol{F}\left(\boldsymbol{F}^T\Xi_{-p}\left(\Xi_{-p}^T\Xi_{-p}\right)^{-1}\Xi_{-p}^T\boldsymbol{F}\right)^{-\frac{1}{2}}. \tag{4.12}$$

Полученные оценки позволяют на основании модели (4.6) и значений факторов в p предшествующие моменты времени сделать прогноз динамики факторов.

Прогноз значений стохастических факторов позволит определить, в какой зоне риска будет находиться экономическая система в прогнозный момент времени, и по значениям факторов восстановить какими будут непосредственно наблюдаемые экономические показатели.

Выводы

Рассматривая риск, как экономическую категорию, проведенный анализ измерителей уровня риска показал, что существующие к настоящему моменту времени системы показателей мало приспособлены к тому, чтобы оценить экономический риск в целом с позиции комплексной оценки риска.

Проанализированные методы измерения экономического риска предполагают использование лишь одного экономического показателя в качестве носителя информации о риске и никак не учитывают причины или внешние факторы, влияющие на уровень рискованности. Поэтому актуальным является создание метода, основанного на исследовании информации по как можно наиболее широкому кругу экономических переменных, с целью получения комплексной оценки экономического риска.

Предложено рассматривать m-мерное факторное в качестве пространства риска, и предложена, определенная на этом пространстве комплексная оценка риска.

Для прогноза уровня рискованности экономической системы разработана автрегрессионная модель произвольного порядка для стохастических факторов, и получены оценки параметров этой модели. Прогноз значений стохастических факторов на основании этой модели позволит определить, в какой зоне риска будет находиться экономическая система в прогнозный момент времени.

5. ИССЛЕДОВАНИЕ ЭКОНОМИЧЕСКИХ СИСТЕМ С ПОМОЩЬЮ МОДЕЛЕЙ СИМУЛЬТАННЫХ УРАВНЕНИЙ

5.1. Природа моделей симультанных уравнений

Количественный анализ влияния одних переменных на другие с целью определения такого факта, как, например, что произойдет с числовым значением результирующего показателя (объемом продаж в отрасли), если величина другой, влияющей нее переменной (учетная ставка центрального банка) увеличится, или уменьшиться на заданное количество единиц (процентов), − проводится методами регрессионного анализа. В отдельном уравнении регрессии допускается наличие лишь одного результирующего показателя, поведение которого должно быть объяснено регрессионной моделью, другие переменные, входящие в уравнение, востребованы для объяснения отображаемых моделью закономерностей. Если для объясняющих переменных пополнен ряд условий, называемых предположениями классической регрессионной модели, одно из которых – независимость или некоррелированность объясняющих переменных и стохастических отклонений уравнения, то параметры модели – значения коэффициентов регрессии и величины дисперсий стохастических возмущений можно оценить обычным методом наименьших квадратов (Ordinary Least Square − OLS).

Таким образом, объясняемые переменные, обозначаемые латинской буквой *Y*, стоящие в левой части регрессионных уравнений, рассматриваемых в согласованные или одни в и те же моменты времени (т.е. одновременно), являются по своей природе зависимыми от стоящих в другой части уравнений переменных. В регрессионном анализе последние называются независимыми, и таковыми они должны быть в реальности. Эти переменные традиционно обозначаются латинской буквой *X*. Они, прежде всего, не должны быть стохастическими, поскольку эти переменные считаются задаваемыми экзогенно; не должны коррелировать друг с другом и, как было отмечено выше – со случайными погрешностями уравнений. Эти свойства независимых уравнений, следующие из предположений классической регрессионной модели, нарушаются или искажаются, то обычный метод наименьших квадратов, широко используемый в регрессионном анализе, не может быть использован для оценивания параметров совокупной системы регрессионных уравнений.

В регрессионных уравнениях, которые можно рассматривать в совокупности в виде системы, причинно-следственная связь носит односторонний характер от *X* к *Y*. Однако во многих ситуациях рассмотрение только таких однонаправленных взаимоотношений между переменными не может объяснить механизм взаимодействий в экономической системе. Это появляется в том случае, когда *Y* обусловливается иксами, но некоторые из них, в свою очередь определяются игреком. Такое сочетание причин и следствий приводит к разнонаправленным или ***симультанным*** связям между зависимыми и независимыми переменными регрессионных уравнений, при этом употребление таких названий, основанных на понятии зависимости, по отношению к этим переменным не может носить корректный характер. Поэтому совокупность уравнений, для переменных которых имеют место симультанные связи, носит название модели симультанных уравнений [67,с.717].

В русскоязычных переводах английское название таких моделей «simultaneous-equation models» часто переводится как «система одновременных уравнений». Действительно, один из вариантов перевода слова «simultaneous» имеет значение «одновременный». Но и система обычных регрессионных уравнений подразумевает, что их совокупность рассматривается для одного и того же момента времени, т.е. одновременно. Поэтому термин «одновременный» по отношению к взаимозависимым уравнениям системы не отражает ее специфики и не устанавливает принципиальной границы между двумя классами систем эконометрических уравнений. Другое значение слова «simultaneous» − «удовлетворяющие одним и тем же переменным уравнений» [68, p. 1393]. Подобрать лаконичный аналог, нагруженный тем же смыслом, в русском языке этому английскому слову непросто, поэтому наиболее точным термином, характеризующим взаимозависимые, и в тоже самое время − одновременные уравнения, является англицизм – «симультанные», тем более, что это слово уже прижилось в русском языке [69].

Таким образом, система симультанных уравнений отличается от системы регрессионных уравнений наличием взаимозависимых переменных. В регрессионных уравнениях в правую – объясняющую часть, входят только независимые переменные. Объясняемые или зависимые переменные не могут входить в эту часть уравнения, такая переменная должна находиться лишь в левой части уравнения. В системе симультанных уравнений объясняемые переменные могут появиться как в правой части уравнения, запись которого соответствует выражению определенной экономической закономерности (экономическая форма системы одновременных уравнений), так и в его левой части. В этом случае зависимая – по терминологии регрессионного анализа – переменная является также объясняемой для переменной, в правую часть уравнения которой она входит. Поэтому два термина, употребляемых в регрессионном анализе «зависимая» и «независимая» переменная, заменяются на более корректные − «эндогенная» и «экзогенная» переменная.

Эндогенная – это внутрисистемная переменная, поведение которой объясняется моделью симультанных уравнений. Для каждой такой переменной в модели имеется свое уравнение. Модель симультанных уравнений и служит для исследования их поведения. Экзогенная переменная – внесистемная переменная, которую часто называют предопределенной [70, с.20] переменной. Такая переменная является по своей природе нестохастической, или ее можно считать таковой. В прикладном анализе, основанном на моделях симультанных уравнений, экзогенные переменные играют ту же роль, что и независимые переменные в регрессионном анализе, т.е. главным образом для того, чтобы ответить на вопрос: какие величины будут принимать эндогенные переменные, если совокупность экзогенных переменных будет находиться в определенном множестве числовых значений. Прогноз значений эндогенных переменных возможет лишь на основании прогнозного или задаваемого значения экзогенных переменных.

Наличие симультанных связей в системе уравнений приводит к понятию взаимозависимых уравнений. Два уравнения с минимум двумя эндогенными переменными называются взаимозависимыми, если в одном уравнении первая эндогенная переменная объясняется второй эндогенной переменной, а в уравнении для второй эндогенной переменной она объясняется первой. Покажем, что наличие взаимозависимых уравнений приводит к нарушению основного предположения классической регрессионной модели и невозможности использования метода OLS. Рассмотрим гипотетическую модель симультанных уравнений [67, с.718]

$$Y_1 = \beta_{10} + \beta_{12}Y_2 + \gamma_1 X_1 + u_1; \quad (5.1)$$

$$Y_2 = \beta_{20} + \beta_{21}Y_1 + \gamma_2 X_1 + u_2; \quad (5.2)$$

где Y_1 и Y_2 – взаимно зависимые, или эндогенные переменные, X_1 − экзогенная переменная; u_1, u_2 – случайные отклонения уравнений. Поскольку переменные Y_1 и Y_2 являются стохастическими, а X_1 – предопределенная или детерминированная переменная, то несложно показать, что эндогенные переменные, входящие в объясняющую, правую часть уравнений (5.1) и (5.2), взаимосвязаны со случайными переменными соответствующих уравнений, т.е. соответствующие ковариации отличны от нуля. Но более опасным для решения проблем оценивания является другое утверждение: случайные отклонения зависят от эндогенных переменных. Из этого следует, что случайное отклонение не является стохастическим процессом типа «белый шум», и в стохастических возмущениях появляются составляющие, математическое ожидание которых не может быть равно нулю. Т.е. в процедуре моделирования, которая предполагает упрощение действительных закономерностей таким образом, чтобы отбрасываемые из рассмотрения связи в своей совокупности уравновешивали друг друга, давая в среднем нулевой эффект, происходит существенный сбой. И в случайной компоненте, часто называемой случайной ошибкой, не имеющей значительного влияния на рассматриваемые закономерности, спрятана существенная информация, игнорирование которой может привести к ложным заключениям.

Чтобы показать взаимозависимость симультанных переменных и стохастических отклонений, подставим Y_1 из (5.1) в (5.2). После несложных преобразований получим:

$$Y_2 = \frac{\beta_{20}+\beta_{21}\beta_{10}}{1-\beta_{12}\beta_{21}} + \frac{\beta_{21}\gamma_1+\gamma_2}{1-\beta_{12}\beta_{21}} X_1 + \frac{\beta_{21}}{1-\beta_{12}\beta_{21}} u_1 + u_2. \quad (5.3)$$

Поскольку математическое ожидание Y_2 имеет вид

$$\mathcal{E}\{Y_2\} = \frac{\beta_{20}+\beta_{21}\beta_{10}}{1-\beta_{12}\beta_{21}} + \frac{\beta_{21}\gamma_1+\gamma_2}{1-\beta_{12}\beta_{21}} X_1, \quad (5.4)$$

то отклонение случайной переменной Y_2 от своего математического ожидания равно

$$Y_2 - \mathcal{E}\{Y_2\} = \frac{\beta_{21}}{1-\beta_{12}\beta_{21}} u_1 + u_2. \quad (5.5)$$

Относительно случайных отклонений двух уравнений будем полагать следующие свойства:

$$\mathcal{E}\{u_1\} = \mathcal{E}\{u_2\} = 0;\ \mathcal{D}\{u_1\} = \sigma_1^2; \mathcal{D}\{u_2\} = \sigma_2^2;\ \ \mathcal{E}\{u_1 u_2\} = 0.$$

Поэтому

$$u_1 - \mathcal{E}\{u_1\} = u_1; \quad (5.6)$$

тогда ковариация между второй эндогенной переменной и случайным отклонением первого уравнения может быть записана в следующем виде (выражение ниже получено с учетом (5.5) и (5.6))

$$cov\{Y_2; u_1\} = \mathcal{E}\{(Y_2 - \mathcal{E}\{Y_2\})(u_1 - \mathcal{E}\{u_1\})\} =$$
$$= \mathcal{E}\left\{\left(\frac{\beta_{21}}{1-\beta_{12}\beta_{21}} u_1 + u_2\right) u_1\right\} = \frac{\beta_{21}}{1-\beta_{12}\beta_{21}} \sigma_1^2; \quad (5.7)$$

т.е. ковариация отлична от нуля и переменная, входящая в правую часть уравнения коррелирует со случайным отклонением этого уравнения. Аналогично подобное можно утверждать и о втором уравнении:

$$cov\{Y_1; u_2\} = \frac{\beta_{12}}{1-\beta_{12}\beta_{21}} \sigma_2^2. \quad (5.8)$$

Выражения (5.7) и (5.8) показывают, что нарушается основное предположение классической регрессионной модели о том, что случайные отклонения являются независимыми или по крайней мере некоррелированными с объясняющими переменными. Поэтому метод обычных наименьших квадратов оценки неизвестных параметров системы симультанных уравнений является мало пригодным в данном случае. Более того, найденные с помощью метода OLS оценки коэффициентов регрессии либо не будут иметь экономический смысл, либо будут искажать результаты экономического анализа, поскольку эти оценки являются к тому и смещенными [67, с.726].

Таким образом, в противоположность совокупности регрессионных уравнений, имеющих однонаправленную причинно-следственную связь от независимых переменных к результирующим показателям, в экономическом анализе существуют уравнения с разнонаправленными по объясняющим эффектам, симультанными переменными. Такие переменные являются взаимозависимыми или эндогенными переменными.

Уникальное свойство систем симультанных уравнений состоит в том, что эндогенная переменная – регрессант в «своем» уравнении, − может появиться в другом в качестве регрессора, объясняющей перемен ой. И как следствие, эндогенные переменные становятся коррелированными со случайными возмущениями уравнений. При этом их число в такой системе должно быть равно числу эндогенных переменных.

Обычный метод наименьших квадратов не может быть использован для получения оценок параметров модели симультанных уравнений. Таким образом, возникает потребность в других методах оценивания, способных обойти проблему коррелированности переменных и ошибок. Но перед этим необходимо решить проблему идентифицируемости.

5.2. Проблема идентифицируемости и оценивания параметров модели симультанных уравнений

Проблема идентифицируемости является одной из центральных в моделях симультанных уравнений. Под идентифицирумостью модели симультанных уравнений понимается возможность полного, точного и однозначного получения численных значений параметров – коэффициентов при переменных и дисперсий случайных отклонений уравнений.

Запишем общую (поведенческую) форму системы симультанных k уравнений:

$y_{t1} = \beta_{21}y_{t2} + \beta_{31}y_{t3} + \cdots + \beta_{k1}y_{tk} + \gamma_{11}x_{t1} + \gamma_{21}x_{t2} + .. + \gamma_{n1}x_{tn} + u_{t1};$
$y_{t2} = \beta_{12}y_{t1} + \beta_{32}y_{t3} + \cdots + \beta_{k2}y_{tk} + \gamma_{12}x_{t1} + \gamma_{22}x_{t2} + .. + \gamma_{n2}x_{tn} + u_{t2};$
……………………………………………………………….….;
$y_{tk} = \beta_{1k}y_{t1} + \beta_{k2}y_{t2} + \cdots + \beta_{k-1,k}y_{t,k-1} + \gamma_{1k}x_{t1} + \gamma_{2k}x_{t2} + .. + \gamma_{nk}x_{tn} + +u_{tk};$

(5.9)

где y_{ti} − значение -й эндогенной переменной Y_i $(i = 1,2, \dots, k)$ в момент времени t; $(t = 1,2, \dots T)$; k − число эндогенных переменных, T – число наблюдений;

x_{tl} − значение -й экзогенной переменной X_l $(l = 1,2, \dots, n)$ в момент времени t; $(t = 1,2, \dots T)$;

β_{ij} – коэффициент при эндогенной переменной Y_i в -ом уравнении системы;

γ_{lj} – коэффициент при экзогенной переменной X_l в -ом уравнении системы;

u_{tj} – случайное отклонение j-ого уравнения системы в в момент времени t.

Следует отметить, что для определенной системы уравнений (5.9) не каждая переменная, как эндогенная, так и экзогенная, должна входить в каждое уравнение. Это определяется экономическим смыслом самой системы и критериями идентификации уравнений системы.

Систему уравнений (5.9) можно представить в матричной форме. Введем обозначения:

k – число эндогенных переменных и, соответственно, число уравнений в модели симультанных уравнений;

n − число экзогенных переменных;

$\boldsymbol{y}_t = (y_{t1} \quad y_{t2} \quad \cdots \quad y_{tk})$ – вектор значений эндогенных переменных;
$\boldsymbol{x}_t = (x_{t1} \quad x_{t2} \quad \cdots \quad x_{tn})$ – вектор значений экзогенных переменных;
$\boldsymbol{u}_t = (u_{t1} \quad u_{t2} \quad \cdots \quad u_{tk})$ – вектор случайных отклонений уравнений;

$\mathbf{B} = \begin{pmatrix} 1 & -\beta_{12} & \cdots & -\beta_{1k} \\ -\beta_{21} & 1 & \cdots & -\beta_{2k} \\ \vdots & \vdots & \vdots & \vdots \\ -\beta_{k1} & -\beta_{k2} & \cdots & 1 \end{pmatrix}$ – матрица коэффициентов при эндогенных переменных; на ее главной диагонали стоят единицы, что отображает одну из особенностей системы симультанных уравнений: для каждой эндогенной переменной имеется свое объясняющее уравнение, в котором эта переменная стоит в левой части уравнения;

$-\mathbf{\Gamma} = \begin{pmatrix} \gamma_{11} & \gamma_{12} & \cdots & \gamma_{1k} \\ \gamma_{21} & \gamma_{22} & \cdots & \gamma_{2k} \\ \vdots & \vdots & \vdots & \vdots \\ \gamma_{n1} & \gamma_{n2} & \cdots & \gamma_{nk} \end{pmatrix}$ – матрица коэффициентов при экзогенных переменных;

$\boldsymbol{I} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$ – единичная матрица (с единицами на главной диагонали, нулями на других позициях).

Знак «минус» перед коэффициентами эндогенных переменных и перед матрицей коэффициентов экзогенных переменных потребуется для структурной формы системы уравнений. Заглавными нежирными буквами с одним индексом обозначаются названия переменных, индекс соответствует номеру переменной; строчными нежирными буквами с двумя индексами обозначаются числовые значения соответствующих переменных, первый индекс соответствует моменту времени, второй – номеру переменной; жирные заглавные буквы – это матрицы (таблицы) числовых значений коэффициентов, при этом для обозначения самих коэффициентов используются строчные греческие буквы с двумя индексами, первый из которых – номер переменной, второй – номер уравнения, где появляется этот коэффициент; строчные жирные буквы – векторы (матрицы-строки) числовых значений для момента времени, указанного индексом при векторе.

Тогда система (5.9) приобретает вид

$$\boldsymbol{y}_t = \boldsymbol{y}_t(\boldsymbol{I} - \mathbf{B}) - \boldsymbol{x}_t\mathbf{\Gamma} + \boldsymbol{u}_t. \tag{5.10}$$

Уравнение (5.10) называется поведенческой (экономической) формой модели симультанных уравнений.

В матричном уравнении (5.10), которое теперь будем использовать вместо развернутой записи системы (5.9), имеются некоторые особенности. К предопределенным переменным $\boldsymbol{x}_t$ относятся как текущие (настоящего времени) и лаговые значения внесистемных переменных, так и лаговые значения эндогенных переменных. Для настоящего момента времени эндогенная переменная является стохастической величиной, но ее лаговое значение, отстающее во времени на один или более шагов времени, становится величиной уже известной, а потому и предопределенной, поэтому такая категория величин переходит в группу экзогенных переменных.

Если в правой части (5.10) оставить лишь стохастическую составляющую, а переменные со своими коэффициентами перенести в левую часть, то получим:

$$\boldsymbol{y}_t - \boldsymbol{y}_t(\boldsymbol{I} - \mathbf{B}) + \boldsymbol{x}_t\mathbf{\Gamma} = \boldsymbol{u}_t;$$

$$\boldsymbol{y}_t - \boldsymbol{y}_t + \boldsymbol{y}_t\mathbf{B} + \boldsymbol{x}_t\mathbf{\Gamma} = \boldsymbol{u}_t;$$

или окончательно получаем:

$$\boldsymbol{y}_t\mathbf{B} + \boldsymbol{x}_t\mathbf{\Gamma} = \boldsymbol{u}_t. \tag{5.11}$$

Выражение (5.11) называется структурной формой модели симультанных уравнений.

Если в (5.11) проделать следующие преобразования

$$\boldsymbol{y}_t\mathbf{B} = -\boldsymbol{x}_t\mathbf{\Gamma} + \boldsymbol{u}_t;$$

$$\boldsymbol{y}_t\mathbf{B}\mathbf{B}^{-1} = -\boldsymbol{x}_t\mathbf{\Gamma}\mathbf{B}^{-1} + \boldsymbol{u}_t\mathbf{B}^{-1};$$

$$\boldsymbol{y}_t = -\boldsymbol{x}_t\mathbf{\Gamma}\mathbf{B}^{-1} + \boldsymbol{u}_t\mathbf{B}^{-1};$$

то, выражая вектор эндогенных переменных в явной форме через вектор экзогенных переменных, получаем следующее уравнение:

$$\boldsymbol{y}_t = \boldsymbol{x}_t\mathbf{\Pi} + \boldsymbol{w}, \tag{5.12}$$

где

$$\mathbf{\Pi} = -\mathbf{\Gamma}\mathbf{B}^{-1} \tag{5.13}$$

и $\boldsymbol{w} = \boldsymbol{u}_t\mathbf{B}^{-1}$.

Выражение (5.12) называется приведенной формой (reduced-form) модели симультанных уравнений, которая выражает эндогенные переменные исключительно через экзогенные переменные и стохастические отклонения. Поскольку в правой части (5.12) имеются только эти составляющие, при этом предполагается, что предопределенные переменные не коррелируют со случайными ошибками, то для получения оценок параметров уравнения приведенной формы можно воспользоваться обычным методом наименьших квадратов, т.е. получить OLS-оценки матрицы $\mathbf{\Pi}$, как матрицы коэффициентов регрессии. Коэффициенты этой матрицы достаточно важны для решения задач экономического анализа, но в этом случае модель сводится к традиционной схеме регрессионного анализа, и при этом не учитываются основные зависимости между переменными, определяемые (5.10). Поэтому матрица $\mathbf{\Pi}$ имеет вспомогательный характер, и в некоторых случаях ею можно воспользоваться для восстановления значений матриц коэффициентов структурной формы.

Проблема идентифицируемости начинается с ответа на вопрос: существуют ли однозначные оценки матрицы коэффициентов приведенной формы. Поскольку приведенная форма (5.12) представляет собой модель классического регрессионного анализа, то ответ на этот вопрос достаточно прост: для того, чтобы существовала OLS-оценка матрицы $\mathbf{\Pi}$:

$$\widehat{\mathbf{\Pi}} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{Y}, \tag{5.14}$$

где $\boldsymbol{X}$ – матрица значений экзогенных переменных для всех периодов времени;

$\boldsymbol{Y}$ − матрица значений эндогенных переменных для всех периодов времени;

необходимо и достаточно выполнение матричных действий в (5.14), а для этого матрица $\boldsymbol{X}^T\boldsymbol{X}$ должна иметь обратную. Получение обратной матрицы возможно лишь в одном случае, когда матрица $\boldsymbol{X}^T\boldsymbol{X}$ и, соответственно, $\boldsymbol{X}$ − полного ранга, что возможно лишь тогда, когда переменные этой матрицы независимы, т.е. у этой матрицы отсутствует эффект мультиколлинерности.

Таким образом, приведенная форма симультанных уравнений идентифицируема тогда, и только тогда, когда экзогенные переменные являются независимыми величинами.

После того, как установлена возможность однозначного определения коэффициентов приведенной формы системы уравнений, эта проблема идентифицируемости сводится к ответу на вопрос: может ли быть получены числовые оценки значений параметров отдельного уравнения структурной формы по оцененным коэффициентам приведенной формы. Если это можно сделать, то частное уравнение системы является идентифицируемым; в противном случае, этот уравнение – неидентифицируемо или недоидентифицируемо [67, с.739]. Если выражение (5.13) рассматривать как матричное уравнение, в котором неизвестными величинами являются матрицы коэффициентов эндо и экзогенных переменных, а известной величиной – оцененная матрица коэффициентов приведенной формы, то в одном уравнении оказываются два неизвестных. Как известно для случая превышения количество неизвестных над числом уравнений существует бесчисленное множество решений, если уравнения совместимы. Именно в этой связи возникает проблема однозначности – идентифицируемости системы симультанных уравнений. Но формальное рассмотрение вопроса однозначности решения системы уравнений сводится к рассмотрению каждого конкретного уравнения. Если число неизвестных величин, входящих в него, позволяет на основании статистических данных определить эти неизвестные, то проблема идентифицируемости разрешима. Таким образом, однозначное определение неизвестных величин сводится к правильному соотношению числа переменных и числа уравнений системы, и условия идентифицируемости выглядят вполне достижимыми [67, с. 748]:

1) В модели из k симультанных уравнений для того, чтобы уравнение было идентифицировано, из него должно быть исключено, по крайней мере, $k-1$ переменных (эндогенных и экзогенных), появляющихся в модели:

$$k+n-k_i-n_i \geq k-1;$$

где k_i – число эндогенных переменных -го уравнения;

n_i – число экзогенных переменных -го уравнения.

При этом если исключено точно $k-1$, то уравнение «просто идентифицирумо», если исключено более чем $k-1$, то оно «переидетифицируемо».

2) В модели из k симультанных уравнений для того, чтобы уравнение было идентифицировано, число предопределенных переменных, исключенных из уравнения, должно быть не меньше, чем число эндогенных в этом уравнении, уменьшенное на единицу:

$$n-n_i \geq k-1.$$

Эти два условия носят название «порядковых условий идентифицируемости». Но эти условия являются необходимыми, но не достаточными. Выполнение порядковых условий еще не гарантирует идентифицируемости отдельного уравнения. Необходимым условием является «ранговое условие идентифицируемости»:

В модели k симультанных уравнений с k эндогенными переменными отдельное уравнение идентифицируемо тогда и только тогда, когда, по крайней мере, один ненулевой детерминант меньшего, чем k порядка может быть составлен из коэффициентов при эндо и экзогенных переменных, исключенных из данного уравнения, но включенных в другие уравнения МСУ.

Проиллюстрируем выполнение разных условий идентифицируемости на следующем примере [67, с. 751]. Пусть уравнение (5.8+3) имеет вид:

$$(\boldsymbol{y}_t \;\vdots\; \boldsymbol{x}_t) \cdot \begin{pmatrix} \boldsymbol{B} \\ \hline \boldsymbol{\Gamma} \end{pmatrix} = \boldsymbol{u}; \qquad \begin{pmatrix} y_{t1} \\ y_{t2} \\ y_{t3} \\ y_{t4} \\ \hline x_{t1} \\ x_{t2} \\ x_{t3} \end{pmatrix}^T \cdot \begin{pmatrix} 1 & 0 & \beta_{13} & \beta_{14} \\ \beta_{21} & 1 & 0 & \beta_{24} \\ \beta_{31} & \beta_{32} & 1 & \beta_{34} \\ 0 & 0 & 0 & 1 \\ \hline \gamma_{11} & \gamma_{12} & \gamma_{13} & 0 \\ 0 & \gamma_{22} & \gamma_{23} & 0 \\ 0 & 0 & 0 & \gamma_{34} \end{pmatrix} = \begin{pmatrix} u_{t1} \\ u_{t2} \\ u_{t3} \\ u_{t4} \end{pmatrix}^T.$$

Из первого уравнения исключены переменные Y_4, X_2, X_3 (нулевые элементы первого столбца блочной матрицы $\begin{pmatrix} \boldsymbol{B} \\ \hdashline \boldsymbol{\Gamma} \end{pmatrix}$, соответствующего первому уравнению МСУ). Если из коэффициентов при этих переменных из трех других уравнений можно составить отличный от нуля определитель, то первое уравнение - идентифицируемо, в противном случае – нет. Из этих коэффициентов можно составить единственный определитель, который имеет вид:

$$\begin{vmatrix} 0 & 0 & 1 \\ \gamma_{22} & \gamma_{23} & 0 \\ 0 & 0 & \gamma_{34} \end{vmatrix},$$

и величина которого равна нулю. Следовательно, первое уравнение по ранговому условию идентифицируемости не может иметь однозначных значений коэффициентов. В этом случае, вид этого уравнения МСУ должен быть пересмотрен, для однозначного восстановления матриц $\boldsymbol{B}$ и $\boldsymbol{\Gamma}$ по матрице $\mathbf{\Pi}$.

Таким образом, решение вопроса об идентифицируемости уравнений МСУ с позиции определения возможности однозначного восстановления матриц структурной формы симультанных уравнений по матрице приведенной формы приводит к противоречивым выводам при проверки порядковых и ранговых условий. При этом корректировка структуры уравнения МСУ в случае его неиндентифицируемости ведет к изменению первоначальных теоретических положений построения симультанной модели. Поэтому возникает необходимость иного метода оценивания и других критериев идентифицируемости, позволяющих сохранить первоначальную структуру уравнений.

Для оценивания систем одновременных уравнений в настоящее время имеется большое число методов, которые можно разделить на две группы: с ограниченной информацией и с полной информацией. Первую группу составляют методы, оценивающие параметры каждого отдельного уравнения системы: двухшаговый метод наименьших квадратов [71], метод максимального правдоподобия с ограниченной информацией или метод наименьшего дисперсионного соотношения [72], метод комиссии Коулса [73]. Вторую группу составляют методы, предназначенные для оценивания всей системы в целом. Это – трехшаговый метод наименьших квадратов [71], метод неподвижной точки и метод максимального правдоподобия с полной информацией [67].

Не смотря на то, что для наилучшей передачи природы симультанных уравнений методы второй группы предпочтительней, они используются значительно реже, чем методы с неполной информацией. И этому есть несколько причин [67, с.763]. Первая состоит в обременительной вычислительной процедуре, не смотря даже на современные возможности компьютерной техники. Вторая заключается в том, что методы полной информации ведут к решениям в высшей степени нелинейным по параметрам, и поэтому часто трудно интерпретируемым. Третья относится к случаю, когда имеется ошибка в спецификации (неправильная форма уравнения или исключены нужные переменные) даже в одном уравнении, тогда эта ошибка трансформируется на всю систему в целом. И как результат, методы полной информации очень чувствительны к ошибкам спецификации.

Несмотря на свои особенности, все эти методы объединяются в одном: эндогенные переменные в объясняющей части каждого уравнения заменяются значениями, воспроизводимыми по их регрессии на все экзогенные переменные, т.е. вместо всех эндогенных переменных правой части уравнений подставляются их теоретизированные значения, которые, как предполагается, не должны коррелировать со стохастическими отклонениями. Действительно, предопределенная величина никак не может быть связана со случайной ошибкой в объяснении результирующего показателя. После такой подстановки оценки коэффициентов уравнений находятся на основании обычного метода наименьших квадратов. Двух и трех шаговые процедуры оценивания на первом этапе производят замену реальных значений эндогенных переменных на их регрессионные величины. Затем второй этап двухшаговой процедуры служит для вычисления оценок параметров, а трехшаговой – для нахождения ковариационной матрицы погрешностей системы уравнений. Третий шаг дает окончательную оценку в трехшаговом методе.

Однако замена эндогенных, стохастических переменных линейной или нелинейной комбинацией предопределенных – детерминированных величин не может не вызвать сомнений, тем более, что в теоретическом плане такая замена пока еще не обоснована. Более того, можно показать, что регрессионная подстановка, осуществляемая на основании выражения (5.12), дает величину стохастическую, которая все-таки коррелирует со случайными отклонениями уравнений структурной формы [67, с. 772].

Еще одно обстоятельство делает сомнительным пригодность этого инструмента очищения переменных от корреляции с ошибками, если допустить, что экзогенная переменная является такой же стохастической величиной, как и эндогенная, со своими корреляциями с другими экзогенными переменными. Что, в конечном счете, ведет к корреляции экзогенных переменных со случайными отклонениями. Покажем это на системе из двух симультанных уравнений

$$\begin{cases} Y_1 = \beta_{12}Y_2 + \gamma_{11}X_1 + \gamma_{12}X_2 + u_1; \\ Y_2 = \beta_{21}Y_1 + \gamma_{21}X_1 + \gamma_{22}X_2 + u_2. \end{cases} \tag{5.15}$$

Эндогенные переменные уравнения (5.15) являются взаимозависимыми, как это было показано в п. 5.1. Предположим, что экзогенные – внесистемные переменные коррелируют друг с другом, их наличие одновременно в двух уравнениях делает эти переменные взаимозависимыми. Если пренебречь причинной зависимостью экзогенных переменных от эндогенных, но имея в виду влияние иксов друг на друга и, соответственно, взаимную причинность, то (5.15) может быть записано в следующей форме:

$$X_1 = -\frac{\gamma_{12}}{\gamma_{11}}X_2 - \frac{1}{\gamma_{11}}u_1; \tag{5.16}$$

$$X_2 = -\frac{\gamma_{21}}{\gamma_{22}}X_1 - \frac{1}{\gamma_{22}}u_2. \tag{5.17}$$

Если (5.16) подставить в (5.17), получим:

$$X_2 = \frac{\gamma_{21}}{\gamma_{22}}\frac{\gamma_{12}}{\gamma_{11}}X_2 + \frac{\gamma_{21}}{\gamma_{22}\gamma_{11}}u_1 - \frac{1}{\gamma_{22}}u_2;$$

$$X_2 = \frac{\gamma_{21}}{\gamma_{11}\gamma_{22}-\gamma_{21}\gamma_{12}}u_1 - \frac{\gamma_{11}}{\gamma_{11}\gamma_{22}-\gamma_{21}\gamma_{12}}u_2. \tag{5.18}$$

Тогда взятие математического ожидания от результата умножения обеих частей (5.18) на u_1 даст коэффициент ковариации между X_2 и u_1, входящими в правую часть (5.16):

$$cov\{X_2, u_1\} = \mathcal{E}\{X_2 \cdot u_1\} = \mathcal{E}\left\{\frac{\gamma_{21}}{\gamma_{11}\gamma_{22}-\gamma_{21}\gamma_{12}}u_1u_1 - \frac{\gamma_{11}}{\gamma_{11}\gamma_{22}-\gamma_{21}\gamma_{12}}u_2u_1\right\} =$$
$$= \frac{\gamma_{21}}{\gamma_{11}\gamma_{22}-\gamma_{21}\gamma_{12}}\sigma_1^2.$$

Аналогично устанавливается корреляция между экзогенной переменной и случайным отклонением (5.17):

$$cov\{X_1, u_2\} = \frac{\gamma_{12}}{\gamma_{11}\gamma_{22}-\gamma_{21}\gamma_{12}}\sigma_2^2.$$

σ_1^2 и σ_2^2 – дисперсии случайных отклонений двух уравнений МСУ.

Поскольку коэффициенты ковариаций отличны от нуля, то вместе с эндогенными переменными системы уравнений (5.15) экзогенные также коррелируют со случайными отклонениями. Поэтому использовать регрессию эндогенных переменных на экзогенные, которые тоже коррелируют со случайными отклонениями, для фильтрации такой корреляции, также плохо, как и прямое использование обычного метода наименьших квадратов для оценки коэффициентов системы симультанных уравнений.

Вследствие чего возникает необходимость в новом подходе для оценки параметров системы одновременных уравнений, позволяющем исключить влияние стохастических отклонений уравнений на все переменные как эндо, так и экзогенные.

Таким образом, использование существующих к настоящему моменту времени методов идентификации системы симультанных уравнений приводит к необходимости пересмотра и коррекции теоретических предпосылок, используемых при составлении функциональных соотношений между показателями экономической системы. А в случае стохастичности экзогенных переменных общепринятые методы оценивания коэффициентов системы симультанных уравнений не дадут достоверных решений.

Однако можно обойтись без корректировки исходных теоретических положений для моделирования с помощью системы симультанных уравнений в случае ее неидентифицируемости и сформулировать иные критерии идентифицируемости, если применить метод оценивания, основанный не на использовании регрессии на экзогенные переменные, а на фильтрации корреляции переменных со случайными отклонениями посредством использования факторов, вызывающих такие корреляции.

5.3. Формирование модели симультанных уравнений с помощью стохастического факторного анализа

Стохастический факторный анализ, как отмечалось выше, может быть использован для широкого спектра направлений эконометрических исследований, и, в частности, в качестве инструмента формирования основных гипотез относительно объекта изучения, характерной чертой которого является многообразие связей между его различными сторонами, отображаемыми большим набором показателей – переменных величин. После того, как поставлена цель исследования, именно с формулировки предварительных положений, основанных на известных экономических законах, начинается эконометрическое изучение сложного объекта. Формализовать этот процесс можно лишь перечислением основных этапов и моментов, перечень которых начинается с необходимости решить назревшую проблему, и заканчивается построением модели объекта, предназначенной для испытания возможных вариантов развития процесса и принятия единственно правильной, оптимальной и целесообразной стратегии действий для преодоления изначальной проблемы. Сама формализация процесса исследования представляет интерес ни сколько для процедуры принятия решения, а столько с точки зрения выработки научных подходов к математическому моделированию. Но стохастический факторный анализ позволяет выработать формальные подходы к созданию конструкции модели, основанной на учете многообразных функциональных связей объекта посредством системы математических соотношений в виде уравнений.

При исследовании экономических систем структура модели симультанных уравнений подбирается в соответствии с предварительными теоретическими положениями, основанными на природе объекта, базовыми экономическими законами, которым этот объект подчиняется, целью исследования, имеющейся в наличии статистической информации или возможностью ее получения, и прочими моментами, отражающими специфику моделируемого явления. На этапе создания структуры модели возникает задача решить, поведение каких переменных необходимо объяснить, какие переменные влияют на результирующие показатели, и сколько соотношений потребуется для того, чтобы в достаточной мере описать особенности поведения экономического объекта. То есть, провести, так называемую, спецификацию модели, к которой относится процесс разделения регистрируемых показателей на эндогенные и экзогенные переменные.

Процесс такой спецификации зависит как от цели исследования, в которой должно найти отражение задачи: поведение каких переменных должно быть объяснено, − и, соответственно, такие переменные становятся эндогенными; так и от предварительной информации об объекте исследования. Но помимо априорного разделения переменных, проводимых субъективно, т.е. самим исследователем, существует объективный тест на экзогенность [67, с. 756]. В соответствии с этим тестом в отдельном уравнении МСУ проверяется значимость коэффициента регрессии как у предварительно названных эндогенных переменных, так и у включенных в это же уравнение восстановленных по приведенной форме этих же переменных. Если коэффициенты при восстановленных переменных значимы, то переменные можно считать эндогенными, если эти коэффициенты в соответствии со статистическим тестом – нулевые, то переменные должны трактоваться как экзогенные. Например, в уравнение для эндогенной переменной Y_1 входят, в качестве объясняющих, две другие предварительно названные эндогенные переменные Y_2 и Y_3, и экзогенная переменная X_1:

$$Y_1 = \beta_2 Y_2 + \beta_3 Y_3 + \gamma_1 X_1 + u_1. \tag{5.19}$$

По приведенной форме МСУ можно найти значения Y_2 и Y_3, выраженные исключительно через предопределенные переменные (через все иксы). Обозначим эти оценочные значения $\hat{Y}_2$ и $\hat{Y}_3$. Подставим эти значения в (5.19):

$$Y_1 = \beta_2 Y_2 + \beta_3 Y_3 + \gamma_1 X_1 + \delta_1 \hat{Y}_2 + \delta_2 \hat{Y}_3 + u_1. \qquad (5.20)$$

Теперь, рассматривая (5.20) как обычное уравнение регрессии, оцениваются все регрессионные коэффициенты обычным методом наименьших квадратов. После чего проверяется гипотеза о равенстве нулю коэффициентов δ_1 и δ_2. Если гипотеза отклоняется, то переменные Y_2 и Y_3 могут считаться эндогенными, если гипотеза принимается, то эти переменные – экзогенные.

Данный тест может быть использован лишь в случае полной предопределенности априорно названных экзогенных переменных. Поэтому в случае стохастичности экзогенных переменных проведение такого теста не может дать результат, соответствующий действительности. Это первая причина, по которой практическое использование этого теста на экзогенность вызывает сомнения. Вторая причина заключается в сомнительной логике вынесения решения по результатам проверки гипотезы. Значимость коэффициентов при оценочных значениях эндогенных переменных показывает лишь то, что только сами предопределенные переменные имеют смысловую ценность для уравнения СМУ, поскольку оценочные значения эндогенных показателей представляют собой некоторую комбинацию экзогенных переменных. Наоборот, значимость коэффициентов β_2 и β_3 по сравнению со значимостью δ_1 и δ_2 может указывать на самостоятельную, отдельную от экзогенных величин, важность переменных Y_2 и Y_3 для анализа экономической системы, что указывало бы на эндогенность Y_2 и Y_3. Близость к нулевым значениям β_2 и β_3 могло бы указывать на то, что соответствующие переменные не должны входить в данное уравнение, что, впрочем, никак не решает вопрос о экзогенности. Третья причина состоит в том, что применение обычного метода наименьших квадратов в случае стохастичности переменных, как отмечалось выше, дает смещенные, несостоятельные и неэффективные оценки коэффициентов регрессии, и делать на их основе какие-либо заключения сомнительно.

Таким образом, решение вопроса о выборе эндогенных переменных и отнесении другого набора переменных к экзогенным должно быть основано на цели исследования и априорной информации, но метод получения оценок параметров модели симультанных уравнений должен учитывать специфическую роль каждой переменной, входящей в систему уравнений.

Если считать, что все переменные и эндо, и экзогенные переменные являются стохастическими, проведение факторного анализа позволяет определить меру близости каждой переменной к каждому стохастическому фактору. При этом, если связать данный фактор с эндогенной переменной, то по факторной структуре можно определить какие переменные входят в группу, формируемую этим фактором, и, соответственно, оказывают влияние на объясняемый показатель экономической системы. Таким образом, если число факторов положить равным числу эндогенных переменных, то структура стохастического факторного анализа может указать на структуру модели симультанных уравнений. Эта структура может быть сформирована по степени близость переменной к данному фактору.

Факторную группу должны составить переменные со значимыми коэффициентами факторных нагрузок. В нее войдут как эндогенные, так и экзогенные стохастические величины, при этом та из эндогенных переменных, которая имеет наибольшую нагрузку данного фактора, и будет переменной, объясняемой другими, попавшими в данную факторную группу. Стохастический фактор является причиной, вызывающей корреляцию между переменными, поэтому переменные, с большими нагрузками данного фактора, оказываются тесно связанными между собой; т.е. между ними существуют взаимосвязи, которые можно описать аналитически. Та из эндогенных переменных из факторной группы, которая имеет наибольшую факторную нагрузку, в большей же степени и будет определяться данным фактором, а, следовательно, влиянием других переменных, отождествляемых с данным фактором.

Пусть в результате предварительного анализа выделены k эндогенных переменных и n экзогенных. Не принимая во внимание формальные процедуры определения числа факторов, в качестве такого числа устанавливается число эндогенных переменных k. Тогда матрица факторных нагрузок, оцененная с помощью разработанных методов, имеет вид, представленный в таблице 5.1

Таблица 5.1 – Факторные нагрузки эндо и экзогенных переменных

Переменная	Фактор			
	F_1	F_2	...	F_k
Y_1	a_{11}	a_{12}	...	a_{1k}
Y_2	a_{21}	a_{22}	...	a_{2k}
...	...	...	...	...
Y_k	a_{k1}	a_{k2}	...	a_{kk}
X_1	$a_{(k+1)1}$	$a_{(k+1)2}$	...	$a_{(k+1)k}$
X_2	$a_{(k+2)1}$	$a_{(k+2)2}$	...	$a_{(k+2)k}$
...	...	...	...	...
X_n	$a_{(k+n)1}$	$a_{(k+n)2}$	...	$a_{(k+n)k}$

Таким образом, формальная процедура составления структуры симультанных уравнений выглядит следующим образом. В каждой строке, соответствующей определенной эндогенной переменной (т.е. в верхней части таблицы 5.1), определяется наибольший по абсолютному значению элемент факторной нагрузки. Столбец, содержащий этот элемент, укажет на фактор, с которым в наибольшей степени связана данная переменная. Таким образом, каждая эндогенная переменная будет отнесена к определенному фактору. Поскольку процедура оценивания факторных нагрузок отвечает принципу простой структуры, то переменные будут «разложены» по всем факторам. Однако, изредка, но может оказаться так, что к отдельному фактору не будет отнесена ни одна из переменных. Тогда в соответствующем столбце должен быть найден наибольший по абсолютному значению коэффициент факторной нагрузки, и из соответствующей ему строки выбрана эндогенная переменная, имеющая с этим «пропущенным» фактором наибольшую корреляцию. Тем самым, эта переменная перейдет от фактора, от которого у нее все же большая нагрузка, к фактору с меньшей нагрузкой, но это позволит не исключать данную эндогенную переменную из спецификации. Последнее действие в выборе структуры симультанных уравнений необходимо для соблюдения постулата о том, что все эндогенные переменные важны для экономического анализа системы.

После того, как каждая эндогенная переменная связана с каждым фактором (а число факторов должно быть равно числу эндогенных переменных), в каждом столбце определяются значимые коэффициенты факторных нагрузок. Эти коэффициенты укажут на переменные, которые должны войти в симультанное уравнение в качестве объясняющих. Среди этих переменных уже будут как эндогенные, так и экзогенные. Таким образом, каждая эндогенная переменная связывается с одним стохастическим фактором, наибольшие нагрузки на переменные которого обозначат те из них, что должны войти в данное симультанное уравнение.

Стохастический фактор в данном случае играет роль некоторой инструментальной величины, посредством которой определяется структура соотношений между наблюдаемыми переменными. В общем случае число таких факторов, равное числу симультанных уравнений в системе, больше, чем необходимо для объяснения корреляций между экономическими показателями. Т.е. эти корреляции объясняются посредством факторов с некоторым избытком, что вполне допустимо и в случае обычного варианта использования стохастического факторного анализа, в котором число факторов определяется как число собственных значений корреляционной матрицы, больших единицы. Для случая, когда число факторов больше, чем это число, корреляции между переменными будут просто воспроизведены с большей точностью, при этом свойство ортогональности стохастических факторов нарушаться не будет.

Увязка числа факторов с числом симультанных уравнений имеет следующую логику. Стохастический фактор представляет собой проявление некоторого латентного обстоятельства причинного свойства. Совокупность факторов определяет все возможные структурно-логические и причинно-следственные связи данной системы. Они существуют объективно и не подвержены взаимодействию друг с другом. Если некоторым образом воспроизвести эти факторы, оценить их значения на протяжении определенного временного интервала, то можно с предельно допустимой точностью воссоздать закономерности и тенденции в развитии исследуемой системы. Таким образом, стохастические факторы объективно отображают разные стороны экономического явления или системы.

Эндогенная переменная, в свою очередь, как некоторый результирующий показатель, также служит для описания с максимально возможной достоверностью одной из существенных граней объекта изучения. Выбор, как самих показателей, так и их числа всегда определяется исследователем, поэтому он – субъективен, как бы ни была объективна причина и целесообразность экономического исследования. Эта субъективность в определенной степени преодолевается увеличением числа исследуемых и моделируемых признаков, однако размерность модели не может быть равной бесконечности или очень большой величине, в противном случае практическая ценность модели сводится к нулю. Преодоление субъективности в исследовании является одной из сторон процесса познания, и чем меньше человеческого фактора в формировании образа или модели объекта исследования, тем, очевидно, ближе к истине результат. В этой связи отождествление количества субъективно задаваемых эндогенных переменных и, соответственно, числа уравнений в МСУ, с числом объективно существующих факторов – является некоторым компромиссом в формировании структуры МСУ.

Таким образом, каждый фактор, нагружающий в большей степени «свои» переменные, определяет ими состав симультанного уравнения. Безусловно, одна и та же переменная может входить в группы разных факторов, но именно в этом проявляется природа симультанности.

Качество структуры модели симультанных уравнений может быть оценена с точки зрения распределения максимальных факторных нагрузок между переменными. Т.е., если факторная нагрузка данной эндогенной переменной является максимальной как в строке, так и в столбце той части матрицы факторных нагрузок, которая относится лишь к эндогенным переменным, то такое распределение факторных нагрузок является идеальным. Структура МСУ построенная по такой матрице факторных нагрузок будет наилучшим образом отображать специфику исследуемой экономической системы. Чем больше отклонения от такого распределения, тем менее достоверна модель при таком числе эндогенных переменных.

Неравномерность распределения эндогенных переменных по факторам может свидетельствовать об избыточном их числе. В этом случае возникает вопрос об оптимальном числе симультанных уравнений и, соответственно, числе эндогенных переменных. Ответ на этот вопрос так же может быть получен с помощью стохастического факторного анализа. Если рассмотреть матрицу динамических корреляций между всеми переменными, то число ее собственных значений укажет на количество минимально необходимых факторов. Что, в свою очередь определит минимальное число симультанных уравнений и, соответственно, − необходимых эндогенных переменных. В этом случае, из списка предварительно отмеченных эндогенных переменных следует выбрать те, на которые стохастические факторы имеют наибольшую нагрузку, при этом каждый фактор будет связан только с одной эндогенной переменной. Остальные переменные этого ряда следует либо исключить из модели как несущественные, либо перевести их в экзогенные переменные. Таким образом, может быть сформулирован тест на эндогенность: та из переменных, которая не может считаться предопределенной и имеет от данного стохастического фактора максимальную нагрузку, является эндогенной, объясняемой симультанным уравнением, в которое в качестве объясняющих войдут переменные со значимыми нагрузками от этого же фактора.

5.4. Оценка параметров модели симультанных уравнений

Перейдем к вопросу получения такой оценки параметров модели симультанных уравнений, которая имела бы лучшие статистические свойства по сравнению с оценками, как обычного метода наименьших квадратов, так и двух, трех шаговых м.н.к., а также метода неподвижной точки и прочих методов оценивания коэффициентов уравнений, рассматриваемых в системе одновременно.

Основной недостаток методов, построенных на минимизации суммы квадратов отклонений реальных наблюдений от некоторой теоретической зависимости, состоит в том, что минимизируется сумма квадратов ошибок, которые коррелируют со значениями переменных, входящих в выражение для этих ошибок. Случайное отклонение – один из элементов правой части регрессионного уравнения, − зависит как от всех его элементов: коэффициентов регрессии, значений зависимой и «независимых» переменных. Но связь зависимой переменной со случайным отклонением определяется видом самого уравнения, а по отношению к переменным правой части уравнения существует один из главных постулатов метода наименьших квадратов, который состоит в том, что эти независимые переменные уже предопределены, детерминированы и не зависят от случая. Поэтому ошибка в минимизируемой целевой функции определяется лишь коэффициентами. Тогда коэффициенты регрессии подбираются так, чтобы сумма квадратов ошибок была наименьшей. Но в случае, когда объясняемые переменные коррелируют с ошибкой, целевая функция м.н.к. становится зависимой и от значений самих переменных, тогда м.н.к.-оценка коэффициентов регрессии делается условной, т.е. справедливой лишь для данного набора значений переменных. Это утверждение справедливо и для многошаговых методов, поскольку они предполагают замену симультанных уравнений предопределенными, которые таковыми не являются. Такую модель для исследования экономической системы использовать уже нельзя, поскольку такой анализ предполагает исследование чувствительности результирующих показателей на всевозможные значения объясняемых величин.

Выше указана логическая сторона непригодности методов, основанных на принципе минимизации суммы квадратов ошибок. Существует еще статистическая сторона этой проблемы. Нарушение основного предположения м.н.к. о независимости предопределенных переменных и ошибок ведет к тому, что м.н.к.-оценки в этом случае теряют свои наилучшие свойства: несмещённости, эффективности и состоятельности. Тем самым метод наименьших квадратов, в том числе двух и трехшаговый, становится грубым, очень приблизительным методом, годящийся лишь для поверхностных прикидок, которые, впрочем, могут оказаться ошибочными.

Докажем смещенность оценок параметров модели симультанных уравнений двухшагового метода наименьших квадратов в случае стохастичности эндогенных переменных.

Теорема К1. Если в i-ом симультанном уравнении (i=1, ..., k)

$$\boldsymbol{y}_i = \left(\boldsymbol{Y}_i \,\vdots\, \boldsymbol{X}_i\right)\begin{pmatrix}\boldsymbol{\beta}_i \\ \hdashline \boldsymbol{\gamma}_i\end{pmatrix} + \boldsymbol{u}_i, \qquad (5.21)$$

где $\boldsymbol{y}_i$ − вектор значений эндогенной переменной i-го уравнения;

$\boldsymbol{Y}_i$ − матрица значений объясняющих эндогенных переменных i-го уравнения;

$\boldsymbol{X}_i$ − матрица значений объясняющих экзогенных переменных i-го уравнения;

$\begin{pmatrix}\boldsymbol{\beta}_i \\ \hdashline \boldsymbol{\gamma}_i\end{pmatrix}$ − вектор истинных значений коэффициентов при объясняющих переменных;

$\boldsymbol{u}_i$ − вектор случайных отклонений i-го уравнения;

$\mathrm{E}\{\boldsymbol{u}_i\} = 0; \quad \mathrm{E}\{\boldsymbol{u}_i \boldsymbol{u}_i^T\} = \boldsymbol{\sigma}^2;$

объясняющие переменные коррелируют со случайными отклонениями:

$$\mathrm{E}\{\boldsymbol{Y}_i^T \boldsymbol{u}_i\} = \boldsymbol{\omega}_{Yu_i} \neq 0; \qquad (5.22)$$

$$\mathrm{E}\{\boldsymbol{X}_i^T \boldsymbol{u}_i\} = \boldsymbol{\omega}_{Xu_i} \neq 0; \tag{5.23}$$

где $\boldsymbol{\omega}_{(.)i}^2$ − вектор ковариаций соответствующего множества переменных со случайными отклонениями;

то оценка коэффициентов уравнения (5.21), полученная с помощью двушагового метода наименьших квадратов (2LS) – (5.24), в котором $\hat{\boldsymbol{Y}}_i = \boldsymbol{X}_i(\boldsymbol{X}_i^T\boldsymbol{X}_i)^{-1}\boldsymbol{X}_i^T\boldsymbol{Y}_i$,:

$$\begin{pmatrix} \hat{\boldsymbol{\beta}}_i \\ \hat{\boldsymbol{\gamma}}_i \end{pmatrix}_{2LS} = \begin{pmatrix} \hat{\boldsymbol{Y}}_i^T\boldsymbol{X}_i & \boldsymbol{Y}_i^T\boldsymbol{X}_i \\ \boldsymbol{X}_i^T\boldsymbol{Y}_i & \boldsymbol{X}_i^T\boldsymbol{X}_i \end{pmatrix}^{-1} \begin{pmatrix} \hat{\boldsymbol{Y}}_i^T\boldsymbol{y}_i \\ \boldsymbol{X}_i^T\boldsymbol{y}_i \end{pmatrix} \tag{5.24}$$

является смещенной.

Доказательство:

Несмещенность означает равенство математического ожидания оценки параметра истинному значению параметра. Найдем математическое ожидание оценки коэффициентов двушаговым методом наименьших квадратов:

$$\mathrm{E}\left\{\begin{pmatrix} \hat{\boldsymbol{\beta}}_i \\ \hat{\boldsymbol{\gamma}}_i \end{pmatrix}_{2LS}\right\} = \mathrm{E}\left\{\begin{pmatrix} \hat{\boldsymbol{Y}}_i^T\boldsymbol{X}_i & \boldsymbol{Y}_i^T\boldsymbol{X}_i \\ \boldsymbol{X}_i^T\boldsymbol{Y}_i & \boldsymbol{X}_i^T\boldsymbol{X}_i \end{pmatrix}^{-1} \begin{pmatrix} \hat{\boldsymbol{Y}}_i^T \\ \boldsymbol{X}_i^T \end{pmatrix}\boldsymbol{y}_i\right\} =$$

$$= \mathrm{E}\left\{\begin{pmatrix} \hat{\boldsymbol{Y}}_i^T\boldsymbol{X}_i & \boldsymbol{Y}_i^T\boldsymbol{X}_i \\ \boldsymbol{X}_i^T\boldsymbol{Y}_i & \boldsymbol{X}_i^T\boldsymbol{X}_i \end{pmatrix}^{-1} \begin{pmatrix} \hat{\boldsymbol{Y}}_i^T \\ \boldsymbol{X}_i^T \end{pmatrix}\left(\begin{pmatrix} \boldsymbol{Y}_i & \boldsymbol{X}_i \end{pmatrix}\begin{pmatrix} \boldsymbol{\beta}_i \\ \boldsymbol{\gamma}_i \end{pmatrix} + \boldsymbol{u}_i\right)\right\} =$$

$$= \begin{pmatrix} \boldsymbol{\beta}_i \\ \boldsymbol{\gamma}_i \end{pmatrix} + \mathrm{E}\left\{\begin{pmatrix} \hat{\boldsymbol{Y}}_i^T\boldsymbol{X}_i & \boldsymbol{Y}_i^T\boldsymbol{X}_i \\ \boldsymbol{X}_i^T\boldsymbol{Y}_i & \boldsymbol{X}_i^T\boldsymbol{X}_i \end{pmatrix}^{-1} \begin{pmatrix} \hat{\boldsymbol{Y}}_i^T \\ \boldsymbol{X}_i^T \end{pmatrix}\boldsymbol{u}_i\right\} =$$

$$= \begin{pmatrix} \boldsymbol{\beta}_i \\ \boldsymbol{\gamma}_i \end{pmatrix} + \mathrm{E}\left\{\begin{pmatrix} \hat{\boldsymbol{Y}}_i^T\boldsymbol{X}_i & \boldsymbol{Y}_i^T\boldsymbol{X}_i \\ \boldsymbol{X}_i^T\boldsymbol{Y}_i & \boldsymbol{X}_i^T\boldsymbol{X}_i \end{pmatrix}^{-1} \begin{pmatrix} \boldsymbol{Y}_i^T\boldsymbol{X}_i(\boldsymbol{X}_i^T\boldsymbol{X}_i)^{-1}\boldsymbol{X}_i^T\boldsymbol{u}_i \\ \boldsymbol{X}_i^T\boldsymbol{u}_i \end{pmatrix}\right\} =$$

$$= \begin{pmatrix} \boldsymbol{\beta}_i \\ \boldsymbol{\gamma}_i \end{pmatrix} + \begin{pmatrix} \boldsymbol{Y}_i^T\boldsymbol{X}_i(\boldsymbol{X}_i^T\boldsymbol{X}_i)^{-1}\sigma_{X_i}^2 \\ \omega_{Xu_i} \end{pmatrix}.$$

Откуда видно, что математическое ожидание оценки имеет смещение. ■

Опишем собственный метод получения оценок параметров модели симультанных уравнений. В его основе лежит разложение симультанных переменных на стохастические факторы, которое можно назвать факторизацией, а сам метод − методом факторизации симультанных уравнений (МФСУ). Представление переменных уравнения (5.21), для которых справедливы ковариации (5.22) и (5.23), в виде некоторой комбинации инструментальных факторов производится на основании стохастического факторного анализа с помощью модели:

$$\begin{pmatrix} \boldsymbol{Y}_i & \boldsymbol{X}_i \end{pmatrix} = \boldsymbol{F}\cdot\begin{pmatrix} \boldsymbol{A}_{Y_i}^T & \boldsymbol{A}_{X_i}^T \end{pmatrix} + \boldsymbol{\Sigma}_i; \tag{5.25}$$

где $\boldsymbol{F}$ – матрица значений стохастических факторов;

$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{A}_{Y_i} \\ \boldsymbol{A}_{X_i} \end{pmatrix}$ − матрица факторных нагрузок (ее блоки соответствуют объясняющим эндогенным и экзогенным переменным);

$\boldsymbol{\Sigma}_i$ − матрица случайных отклонений факторной модели для переменных $\begin{pmatrix} \boldsymbol{Y}_i & \boldsymbol{X}_i \end{pmatrix}$;($\mathrm{E}\{\boldsymbol{\Sigma}_i\} = \boldsymbol{0}$; $\mathrm{E}\{\boldsymbol{\Sigma}_i^T\boldsymbol{\Sigma}_i\} = \boldsymbol{D}_i^2$ – диагональная матрица);

при этом стохастические факторы удовлетворяют условиям

$$\mathrm{E}\{\boldsymbol{F}^T\boldsymbol{u}_i\} = 0 \text{ и } \mathrm{E}\{\boldsymbol{F}^T\boldsymbol{\Sigma}_i\} = 0. \tag{5.26}$$

В этом случае уравнение (5.21) преобразуется к виду:

$$\boldsymbol{y}_i = \left(\boldsymbol{F}\left(\boldsymbol{A}_{Y_i}^T \;\vdots\; \boldsymbol{A}_{X_i}^T\right) + \boldsymbol{\Sigma}_i\right)\begin{pmatrix}\boldsymbol{\beta}_i \\ \hdashline \boldsymbol{\gamma}_i\end{pmatrix} + \boldsymbol{u}_i =$$

$$= \boldsymbol{F}\left(\boldsymbol{A}_{Y_i}^T \;\vdots\; \boldsymbol{A}_{X_i}^T\right)\begin{pmatrix}\boldsymbol{\beta}_i \\ \hdashline \boldsymbol{\gamma}_i\end{pmatrix} + \boldsymbol{\Sigma}_i\begin{pmatrix}\boldsymbol{\beta}_i \\ \hdashline \boldsymbol{\gamma}_i\end{pmatrix} + \boldsymbol{u}_i = \boldsymbol{F}\left(\boldsymbol{A}_{Y_i}^T \;\vdots\; \boldsymbol{A}_{X_i}^T\right)\begin{pmatrix}\boldsymbol{\beta}_i \\ \hdashline \boldsymbol{\gamma}_i\end{pmatrix} + \boldsymbol{w}_i; \tag{5.27}$$

где $\boldsymbol{w}_i = \boldsymbol{\Sigma}_i\begin{pmatrix}\boldsymbol{\beta}_i \\ \hdashline \boldsymbol{\gamma}_i\end{pmatrix} + \boldsymbol{u}_i$ вектор стохастических отклонений *i*-го уравнения.

В конечном итоге, следующее уравнение

$$\boldsymbol{y}_i = \left(\boldsymbol{F}\boldsymbol{A}_{Y_i}^T \;\vdots\; \boldsymbol{F}\boldsymbol{A}_{X_i}^T\right)\begin{pmatrix}\boldsymbol{\beta}_i \\ \hdashline \boldsymbol{\gamma}_i\end{pmatrix} + \boldsymbol{w}_i \tag{5.28}$$

является обычным уравнением регрессии, в котором регрессоры в силу условий (5.26) не коррелируют со случайными отклонениями. Поэтому вектор коэффициентов $\begin{pmatrix}\boldsymbol{\beta}_i \\ \hdashline \boldsymbol{\gamma}_i\end{pmatrix}$ может быть оценен обычным методом наименьших квадратов. Таким образом, оценка коэффициентов уравнения методом факторизации симультанных переменных имеет вид:

$$\begin{pmatrix}\hat{\boldsymbol{\beta}}_i \\ \hdashline \hat{\boldsymbol{\gamma}}_i\end{pmatrix} = \begin{pmatrix}\boldsymbol{A}_{Y_i}\boldsymbol{A}_{Y_i}^T & \boldsymbol{A}_{Y_i}\boldsymbol{A}_{X_i}^T \\ \boldsymbol{A}_{X_i}\boldsymbol{A}_{Y_i}^T & \boldsymbol{A}_{X_i}\boldsymbol{A}_{X_i}^T\end{pmatrix}^{-1}\begin{pmatrix}\boldsymbol{A}_{Y_i}\boldsymbol{F}^T\boldsymbol{y}_i \\ \boldsymbol{A}_{X_i}\boldsymbol{F}^T\boldsymbol{y}_i\end{pmatrix}. \tag{5.29}$$

В (5.29) учтена ортогональность стохастических факторов.

Докажем несмещенность оценки (5.29)

Теорема К2. Оценка (5.29) коэффициентов симультанного уравнения (5.21), полученная с помощью метода факторизации симультанных переменных, является несмещенной.

Доказательство. Найдем математическое ожидание оценки (5.29)

$$\mathrm{E}\left\{\begin{pmatrix}\hat{\boldsymbol{\beta}}_i \\ \hdashline \hat{\boldsymbol{\gamma}}_i\end{pmatrix}\right\} = \mathrm{E}\left\{\begin{pmatrix}\boldsymbol{A}_{Y_i}\boldsymbol{A}_{Y_i}^T & \boldsymbol{A}_{Y_i}\boldsymbol{A}_{X_i}^T \\ \boldsymbol{A}_{X_i}\boldsymbol{A}_{Y_i}^T & \boldsymbol{A}_{X_i}\boldsymbol{A}_{X_i}^T\end{pmatrix}^{-1}\begin{pmatrix}\boldsymbol{A}_{Y_i}\boldsymbol{F}^T \\ \boldsymbol{A}_{X_i}\boldsymbol{F}^T\end{pmatrix}\boldsymbol{y}_i\right\} =$$

$$= \mathrm{E}\left\{\begin{pmatrix}\boldsymbol{A}_{Y_i}\boldsymbol{A}_{Y_i}^T & \boldsymbol{A}_{Y_i}\boldsymbol{A}_{X_i}^T \\ \boldsymbol{A}_{X_i}\boldsymbol{A}_{Y_i}^T & \boldsymbol{A}_{X_i}\boldsymbol{A}_{X_i}^T\end{pmatrix}^{-1}\begin{pmatrix}\boldsymbol{A}_{Y_i}\boldsymbol{F}^T \\ \boldsymbol{A}_{X_i}\boldsymbol{F}^T\end{pmatrix}\left(\left(\boldsymbol{Y}_i \;\vdots\; \boldsymbol{X}_i\right)\begin{pmatrix}\boldsymbol{\beta}_i \\ \hdashline \boldsymbol{\gamma}_i\end{pmatrix} + \boldsymbol{w}_i\right)\right\} =$$

$$= \mathrm{E}\left\{\begin{pmatrix}\boldsymbol{A}_{Y_i}\boldsymbol{A}_{Y_i}^T & \boldsymbol{A}_{Y_i}\boldsymbol{A}_{X_i}^T \\ \boldsymbol{A}_{X_i}\boldsymbol{A}_{Y_i}^T & \boldsymbol{A}_{X_i}\boldsymbol{A}_{X_i}^T\end{pmatrix}^{-1}\begin{pmatrix}\boldsymbol{A}_{Y_i}\boldsymbol{F}^T \\ \boldsymbol{A}_{X_i}\boldsymbol{F}^T\end{pmatrix}\left(\left(\boldsymbol{F}\boldsymbol{A}_{Y_i}^T \;\vdots\; \boldsymbol{F}\boldsymbol{A}_{X_i}^T\right)\begin{pmatrix}\boldsymbol{\beta}_i \\ \hdashline \boldsymbol{\gamma}_i\end{pmatrix} + \boldsymbol{w}_i\right)\right\} =$$

$$= \begin{pmatrix}\boldsymbol{\beta}_i \\ \hdashline \boldsymbol{\gamma}_i\end{pmatrix} + \mathrm{E}\left\{\begin{pmatrix}\boldsymbol{A}_{Y_i}\boldsymbol{A}_{Y_i}^T & \boldsymbol{A}_{Y_i}\boldsymbol{A}_{X_i}^T \\ \boldsymbol{A}_{X_i}\boldsymbol{A}_{Y_i}^T & \boldsymbol{A}_{X_i}\boldsymbol{A}_{X_i}^T\end{pmatrix}^{-1}\begin{pmatrix}\boldsymbol{A}_{Y_i}\boldsymbol{F}^T\boldsymbol{w}_i \\ \boldsymbol{A}_{X_i}\boldsymbol{F}^T\boldsymbol{w}_i\end{pmatrix}\right\} = .$$

$$= \begin{pmatrix}\boldsymbol{\beta}_i \\ \hdashline \boldsymbol{\gamma}_i\end{pmatrix} + \begin{pmatrix}\boldsymbol{A}_{Y_i}\boldsymbol{A}_{Y_i}^T & \boldsymbol{A}_{Y_i}\boldsymbol{A}_{X_i}^T \\ \boldsymbol{A}_{X_i}\boldsymbol{A}_{Y_i}^T & \boldsymbol{A}_{X_i}\boldsymbol{A}_{X_i}^T\end{pmatrix}^{-1}\begin{pmatrix}\boldsymbol{A}_{Y_i}\boldsymbol{F}^T\,\mathrm{E}\{\boldsymbol{w}_i\} \\ \boldsymbol{A}_{X_i}\boldsymbol{F}^T\,\mathrm{E}\{\boldsymbol{w}_i\}\end{pmatrix} = \begin{pmatrix}\boldsymbol{\beta}_i \\ \hdashline \boldsymbol{\gamma}_i\end{pmatrix} \blacksquare$$

Таким образом, для случая симультанности экзогенных переменных по критерию несмещенности оценка, полученная по методу факторизации симультанных уравнений, лучше, чем оценка двухшагового метода наименьших квадратов. Покажем эффективность оценки МФСУ перед оценкой 2LS, когда экзогенные переменные стохастичны. Для этого докажем лемму.

Лемма К1. Если экзогенные переменные системы симультанных уравнений являются стохастическими, то

$$\mathrm{E}\{\boldsymbol{X}^T\boldsymbol{u}_i\boldsymbol{u}_i^T\boldsymbol{X}\} = \boldsymbol{\sigma}^2\boldsymbol{A}_X\boldsymbol{A}_X^T + \boldsymbol{\sigma}_{Xu_i}^2 . \qquad (5.30)$$

Доказательство. Стохастичность переменных предполагает их взаимосвязь, коррелированность. Корреляция между переменными объясняется посредством наличия стохастических общих факторов, т.е. с помощью эксплораторной факторной модели:

$$\boldsymbol{X} = \boldsymbol{F}\boldsymbol{A}_X^T + \boldsymbol{\Sigma}_X ,$$

где обозначения приведены в теореме К1. Тогда

$$\mathrm{E}\{\boldsymbol{X}^T\boldsymbol{u}_i\} = \mathrm{E}\left\{(\boldsymbol{F}\boldsymbol{A}_X^T + \boldsymbol{\Sigma}_X)^T\boldsymbol{u}_i\right\} = \mathrm{E}\{\boldsymbol{A}_X\boldsymbol{F}^T\boldsymbol{u}_i + \boldsymbol{\Sigma}_X^T\boldsymbol{u}_i\} \neq 0. \qquad (5.31)$$

Поскольку в (5.20+11) $\boldsymbol{A}_X\,\mathrm{E}\{\boldsymbol{F}^T\boldsymbol{u}_i\} = 0$, то тогда математическое ожидание второго слагаемого не равно нулю. Результат взятия математического ожидания – ковариационную матрицу ошибок факторного и симультанного уравнений, обозначим

$$\mathrm{E}\{\boldsymbol{\Sigma}_X^T\boldsymbol{u}_i\} \equiv \boldsymbol{\sigma}_{Xu_i} .$$

Тогда

$$\mathrm{E}\{\boldsymbol{X}^T\boldsymbol{u}_i\boldsymbol{u}_i^T\boldsymbol{X}\} = \mathrm{E}\{(\boldsymbol{F}\boldsymbol{A}_X^T + \boldsymbol{\Sigma}_X)^T\boldsymbol{u}_i\boldsymbol{u}_i^T(\boldsymbol{F}\boldsymbol{A}_X^T + \boldsymbol{\Sigma}_X)\} =$$

$$= \mathrm{E}\{\boldsymbol{A}_X\boldsymbol{F}^T\boldsymbol{u}_i\boldsymbol{u}_i^T\boldsymbol{F}\boldsymbol{A}_X^T + \boldsymbol{\Sigma}_X^T\boldsymbol{u}_i\boldsymbol{u}_i^T\boldsymbol{F}\boldsymbol{A}_X^T + \boldsymbol{A}_X\boldsymbol{F}^T\boldsymbol{u}_i\boldsymbol{u}_i^T\boldsymbol{\Sigma}_X + \boldsymbol{\Sigma}_X^T\boldsymbol{u}_i\boldsymbol{u}_i^T\boldsymbol{\Sigma}_X\} =$$

$$= \boldsymbol{A}_X\boldsymbol{F}^T\,\mathrm{E}\{\boldsymbol{u}_i\boldsymbol{u}_i^T\}\boldsymbol{F}\boldsymbol{A}_X^T + 0 + 0 + \mathrm{E}\{\boldsymbol{\Sigma}_X^T\boldsymbol{u}_i\}\mathrm{E}\{\boldsymbol{u}_i^T\boldsymbol{\Sigma}_X\} = \boldsymbol{\sigma}^2\boldsymbol{A}_X\boldsymbol{A}_X^T + \boldsymbol{\sigma}_{Xu_i}\boldsymbol{\sigma}_{Xu_i}^T .$$

Аналогично,

$$\mathrm{E}\{\boldsymbol{X}_i^T\boldsymbol{u}_i\boldsymbol{u}_i^T\boldsymbol{X}_i\} = \boldsymbol{\sigma}^2\boldsymbol{A}_{X_i}\boldsymbol{A}_{X_i}^T + \boldsymbol{\sigma}_{X_iu_i}\boldsymbol{\sigma}_{X_iu_i}^T$$

При выводе были использованы свойства квадратичных форм [74]. ■

Из леммы следует, что математическое ожидание произведения стохастических экзогенных переменных на случайные отклонения симультанного уравнения представляет собой матрицу ковариаций между отклонениями факторного и симультанного уравнений. Однако $\boldsymbol{\omega}_{Xu_i} \neq \boldsymbol{\sigma}_{Xu_i}$,

поскольку $\boldsymbol{\omega}_{Xu_i}$ – ковариация случайных отклонений со стохастическими экзогенными переменными *i*-го уравнения, вектор которой имеет размерность $n_i{\cdot}1$, где n_i – число этих переменных, а $\boldsymbol{\sigma}_{Xu_i}$ – ковариация случайных отклонений i-го уравнения со всеми экзогенными переменными с размерностью n·1.

Теорема К3. В случае стохастичности экзогенных переменных оценка коэффициентов симультанного уравнения, полученная с помощью МФСУ (5.20+9), более эффективна, чем оценка этих коэффициентов двухшаговым методом наименьших квадратов (5.20+4).

Доказательство. Ковариационная матрица 2LS-оценки имеет вид

$$\operatorname{cov}\left\{\begin{pmatrix}\hat{\boldsymbol{\beta}}_i\\ \boldsymbol{\gamma}_i\end{pmatrix}_{2LS}\right\} = E\left\{\left(\begin{pmatrix}\hat{\boldsymbol{\beta}}_i\\ \boldsymbol{\gamma}_i\end{pmatrix}_{2LS} - \begin{pmatrix}\boldsymbol{\beta}_i\\ \boldsymbol{\gamma}_i\end{pmatrix}\right)\left(\begin{pmatrix}\hat{\boldsymbol{\beta}}_i\\ \boldsymbol{\gamma}_i\end{pmatrix}_{2LS} - \begin{pmatrix}\boldsymbol{\beta}_i\\ \boldsymbol{\gamma}_i\end{pmatrix}\right)^T\right\} = [i = 1,2,\ldots,k] =$$

$$=E\left\{\left(\begin{pmatrix}\hat{Y}_i^T\hat{Y}_i & \hat{Y}_i^T X_i\\ X_i^T\hat{Y}_i & X_i^T X_i\end{pmatrix}^{-1}\begin{pmatrix}\hat{Y}_i^T\\ X_i^T\end{pmatrix}y^T-\begin{pmatrix}\beta\\ \gamma\end{pmatrix}\right)\left(\begin{pmatrix}\hat{Y}_i^T\hat{Y}_i & \hat{Y}_i^T X_i\\ X_i^T\hat{Y}_i & X_i^T X_i\end{pmatrix}^{-1}\begin{pmatrix}\hat{Y}_i^T\\ X_i^T\end{pmatrix}y^T-\begin{pmatrix}\beta\\ \gamma\end{pmatrix}\right)^T\right\}=$$

$$=E\left\{\left(\begin{pmatrix}\hat{Y}_i^T\hat{Y}_i & \hat{Y}_i^T X_i\\ X_i^T\hat{Y}_i & X_i^T X_i\end{pmatrix}^{-1}\begin{pmatrix}\hat{Y}_i^T\\ X_i^T\end{pmatrix}\left(\left(\hat{Y}_i \;\vdots\; X_i\right)\begin{pmatrix}\beta\\ \gamma\end{pmatrix}+u_i\right)-\begin{pmatrix}\beta\\ \gamma\end{pmatrix}\right)(\ldots)^T\right\}=$$

$$=E\left\{\left(\begin{pmatrix}\beta\\ \gamma\end{pmatrix}+\begin{pmatrix}\hat{Y}_i^T\hat{Y}_i & \hat{Y}_i^T X_i\\ X_i^T\hat{Y}_i & X_i^T X_i\end{pmatrix}^{-1}\begin{pmatrix}\hat{Y}_i^T\\ X_i^T\end{pmatrix}u_i-\begin{pmatrix}\beta\\ \gamma\end{pmatrix}\right)(\ldots)^T\right\}=$$

$$=E\left\{\left(\begin{pmatrix}\hat{Y}_i^T\hat{Y}_i & \hat{Y}_i^T X_i\\ X_i^T\hat{Y}_i & X_i^T X_i\end{pmatrix}^{-1}\begin{pmatrix}\hat{Y}_i^T\\ X_i^T\end{pmatrix}u_i\right)\left(\begin{pmatrix}\hat{Y}_i^T\hat{Y}_i & \hat{Y}_i^T X_i\\ X_i^T\hat{Y}_i & X_i^T X_i\end{pmatrix}^{-1}\begin{pmatrix}\hat{Y}_i^T\\ X_i^T\end{pmatrix}u_i\right)^T\right\}=$$

$$=E\left\{\begin{pmatrix}\hat{Y}_i^T\hat{Y}_i & \hat{Y}_i^T X_i\\ X_i^T\hat{Y}_i & X_i^T X_i\end{pmatrix}^{-1}\begin{pmatrix}\hat{Y}_i^T\\ X_i^T\end{pmatrix}u_i u_i^T\left(\hat{Y}_i \;\vdots\; X_i\right)\begin{pmatrix}\hat{Y}_i^T\hat{Y}_i & \hat{Y}_i^T X_i\\ X_i^T\hat{Y}_i & X_i^T X_i\end{pmatrix}^{-1}\right\}=$$

$$=E\left\{W^{-1}\begin{pmatrix}Y_i^T X(X^TX)^{-1}X^T u_i u_i^T X(X^TX)^{-1}X^T Y_i & Y_i^T X(X^TX)^{-1}X^T u_i u_i^T X_i\\ X_i^T u_i u_i^T X(X^TX)^{-1}X^T Y_i & X_i^T u_i u_i^T X_i\end{pmatrix}W^{-1}\right\}=$$

$$=W^{-1}\left(\begin{array}{c:c}Y_i^T X(X^TX)^{-1}(\sigma^2 A_X A_X^T+\sigma_{Xu_i}\sigma_{Xu_i}^T)(X^TX)^{-1}X^T Y_i & Y_i^T X(X^TX)^{-1}\sigma_{Xu_i}\omega_{Xu_i}^T\\ \hdashline \omega_{Xu_i}\sigma_{Xu_i}^T(X^TX)^{-1}X^T Y_i & \sigma^2 A_{X_i}A_{X_i}^T+\sigma_{X_iu_i}\sigma_{X_iu_i}^T\end{array}\right)W^{-1};(\forall i);$$

(5.32)

где $$W=\left(\begin{array}{c:c}Y_i^T X(X^TX)^{-1}X^T Y_i & Y_i^T X(X^TX)^{-1}X^T X_i\\ \hdashline X_i^T X(X^TX)^{-1}X^T Y_i & X_i^T X_i\end{array}\right). \qquad (5.33)$$

Найдем ковариационную матрицу оценки МФСУ коэффициентов симультанного уравнения (FSE).

$$\operatorname{cov}\left\{\begin{pmatrix}\hat{\beta}_i\\ \gamma_i\end{pmatrix}_{FSE}\right\}=E\left\{\left(\begin{pmatrix}\hat{\beta}_i\\ \gamma_i\end{pmatrix}_{FSE}-\begin{pmatrix}\beta_i\\ \gamma_i\end{pmatrix}\right)\left(\begin{pmatrix}\hat{\beta}_i\\ \gamma_i\end{pmatrix}_{FSE}-\begin{pmatrix}\beta_i\\ \gamma_i\end{pmatrix}\right)^T\right\}=[i=1,2,\ldots,k]=$$

$$=\mathrm{E}\left\{\left(\begin{pmatrix}A_{Y_i}A_{Y_i}^T & A_{Y_i}A_{X_i}^T\\ A_{X_i}A_{Y_i}^T & A_{X_i}A_{X_i}^T\end{pmatrix}^{-1}\begin{pmatrix}A_{Y_i}\\ A_{X_i}\end{pmatrix}F^T y_i-\begin{pmatrix}\beta_i\\ \gamma_i\end{pmatrix}\right)\left(\begin{pmatrix}A_{Y_i}A_{Y_i}^T & A_{Y_i}A_{X_i}^T\\ A_{X_i}A_{Y_i}^T & A_{X_i}A_{X_i}^T\end{pmatrix}^{-1}\begin{pmatrix}A_{Y_i}\\ A_{X_i}\end{pmatrix}F^T y_i-\begin{pmatrix}\beta_i\\ \gamma_i\end{pmatrix}\right)^T\right\}=$$

$$=\mathrm{E}\left\{\left(\begin{pmatrix}A_{Y_i}A_{Y_i}^T & A_{Y_i}A_{X_i}^T\\ A_{X_i}A_{Y_i}^T & A_{X_i}A_{X_i}^T\end{pmatrix}^{-1}\begin{pmatrix}A_{Y_i}\\ A_{X_i}\end{pmatrix}F^T\left(F\left(A_{Y_i}^T \;\vdots\; A_{X_i}^T\right)\begin{pmatrix}\beta_i\\ \gamma_i\end{pmatrix}+u_i\right)-\begin{pmatrix}\beta_i\\ \gamma_i\end{pmatrix}\right)(\ldots)^T\right\}=$$

$$=\mathrm{E}\left\{\left(\begin{pmatrix}\beta_i\\ \gamma_i\end{pmatrix}+\begin{pmatrix}A_{Y_i}A_{Y_i}^T & A_{Y_i}A_{X_i}^T\\ A_{X_i}A_{Y_i}^T & A_{X_i}A_{X_i}^T\end{pmatrix}^{-1}\begin{pmatrix}A_{Y_i}\\ A_{X_i}\end{pmatrix}F^T u_i-\begin{pmatrix}\beta_i\\ \gamma_i\end{pmatrix}\right)(\ldots)^T\right\}=$$

$$=\mathrm{E}\left\{\begin{pmatrix}A_{Y_i}A_{Y_i}^T & A_{Y_i}A_{X_i}^T\\ A_{X_i}A_{Y_i}^T & A_{X_i}A_{X_i}^T\end{pmatrix}^{-1}\begin{pmatrix}A_{Y_i}\\ A_{X_i}\end{pmatrix}F^T u_i u_i^T F\left(A_{Y_i}^T \;\vdots\; A_{X_i}^T\right)\begin{pmatrix}A_{Y_i}A_{Y_i}^T & A_{Y_i}A_{X_i}^T\\ A_{X_i}A_{Y_i}^T & A_{X_i}A_{X_i}^T\end{pmatrix}^{-1}\right\}=$$

$$=\sigma^2\left(\begin{array}{c|c} A_{Y_i}A_{Y_i}^T & A_{Y_i}A_{X_i}^T \\ \hline A_{X_i}A_{Y_i}^T & A_{X_i}A_{X_i}^T \end{array}\right)^{-1}.$$

Покажем положительную определенность матрицы Q, равной разности ковариационной матрицы оценки двухшагового метода наименьших квадратов (2LS) и оценки МФСУ (FSE).

$$Q=\operatorname{cov}\left\{\begin{pmatrix}\hat{\beta}_i \\ \gamma_i\end{pmatrix}_{2LS}\right\}-\operatorname{cov}\left\{\begin{pmatrix}\hat{\beta}_i \\ \gamma_i\end{pmatrix}_{FSE}\right\}=$$

$$=W^{-1}\left(\begin{array}{c|c} Y_i^T X(X^T X)^{-1}(\sigma^2 A_X A_X^T+\sigma_{Xu_i}\sigma_{Xu_i}^T)(X^T X)^{-1}X^T Y_i & Y_i^T X(X^T X)^{-1}\sigma_{Xu_i}\omega_{Xu_i}^T \\ \hline \omega_{Xu_i}\sigma_{Xu_i}^T(X^T X)^{-1}X^T Y_i & \sigma^2 A_{X_i}A_{X_i}^T+\sigma_{X_i u_i}\sigma_{X_i u_i}^T \end{array}\right)W^{-1}-$$

$$-\sigma^2\left(\begin{array}{c|c} A_{Y_i}A_{Y_i}^T & A_{Y_i}A_{X_i}^T \\ \hline A_{X_i}A_{Y_i}^T & A_{X_i}A_{X_i}^T \end{array}\right)^{-1}.$$

Поскольку относительно переменных модели симультанных уравнений известно, что они – стохастические и взаимозависимые, то, отбрасывая в (5.20+5) случайную составляющую, и выражая переменные через значения факторов, получим

$$Q=W^{-1}\left(\begin{array}{c|c} A_{Y_i}A_X^T(A_X A_X^T)^{-1}A_X A_{Y_i}^T(\sigma^2 A_X A_X^T+\sigma_{Xu_i}\sigma_{Xu_i}^T)(A_X A_X^T)^{-1}A_X A_{Y_i i}^T & A_{Y_i}A_X^T(A_X A_X^T)^{-1}\sigma_{Xu_i}\omega_{Xu_i}^T \\ \hline \omega_{Xu_i}\sigma_{Xu_i}^T(A_X A_X^T)^{-1}A_X A_{Y_i i}^T & \sigma^2 A_{X_i}A_{X_i}^T+\sigma_{X_i u_i}\sigma_{X_i u_i}^T \end{array}\right)W^{-1}-$$

$$-\sigma^2\left(\begin{array}{c|c} A_{Y_i}A_{Y_i}^T & A_{Y_i}A_{X_i}^T \\ \hline A_{X_i}A_{Y_i}^T & A_{X_i}A_{X_i}^T \end{array}\right)^{-1};$$

при этом матрица W принимает вид

$$W=\left(\begin{array}{c|c} A_{Y_i}A_X^T(A_X A_X^T)^{-1}A_X A_{Y_i}^T & A_{Y_i}A_X^T(A_X A_X^T)^{-1}A_X A_{X_i}^T \\ \hline A_{X_i}A_X^T(A_X A_X^T)^{-1}A_X A_{Y_i}^T & A_{X_i}A_{X_i}^T \end{array}\right).$$

Поскольку обратная к матрице ковариации оценки МФСУ

$$V=\left(\begin{array}{c|c} A_{Y_i}A_{Y_i}^T & A_{Y_i}A_{X_i}^T \\ \hline A_{X_i}A_{Y_i}^T & A_{X_i}A_{X_i}^T \end{array}\right)$$

с точностью до скалярного множителя и идемпотентной матрицы $A_X^T(A_X A_X^T)^{-1}A_X$ по своей структуре соответствует матрице W, то будем считать эти матрицы примерно тождественными $V\cong W$. Тогда

$$Q=W^{-1}(S-\sigma^2 W)W^{-1},$$

где

$$S=\left(\begin{array}{c|c} A_{Y_i}A_X^T(A_X A_X^T)^{-1}A_X A_{Y_i}^T(\sigma^2 A_X A_X^T+\sigma_{Xu_i}\sigma_{Xu_i}^T)(A_X A_X^T)^{-1}A_X A_{Y_i i}^T & A_{Y_i}A_X^T(A_X A_X^T)^{-1}\sigma_{Xu_i}\omega_{Xu_i}^T \\ \hline \omega_{Xu_i}\sigma_{Xu_i}^T(A_X A_X^T)^{-1}A_X A_{Y_i i}^T & \sigma^2 A_{X_i}A_{X_i}^T+\sigma_{X_i u_i}\sigma_{X_i u_i}^T \end{array}\right).$$

Поскольку матрица W^{-1}, как симметричная, положительно определена, то достаточно показать положительную определенность матрицы разности матриц $(S-\sigma^2 W)$. Поскольку эти матрицы имеют блочную структуру одного порядка, то рассмотрим результат разности матриц по блокам.

Рассмотрим диагональные блоки. Разность верхних левых блоков имеет вид

$$A_{Y_i}A_X^T(A_X A_X^T)^{-1}A_X A_{Y_i}^T(\sigma^2 A_X A_X^T+\sigma_{Xu_i}\sigma_{Xu_i}^T)(A_X A_X^T)^{-1}A_X A_{Y_i i}^T-$$

$$-\sigma^2 A_{Y_i}A_X^T(A_X A_X^T)^{-1}A_X A_{Y_i}^T=\sigma^2 A_{Y_i}A_X^T(A_X A_X^T)^{-1}A_X A_{Y_i}^T+$$

$$+A_{Y_i}A_X^T(A_XA_X^T)^{-1}A_XA_{Y_i}^T\sigma_{Xu_i}\sigma_{Xu_i}^T(A_XA_X^T)^{-1}A_XA_{Y_i\,i}^T-\sigma^2A_{Y_i}A_X^T(A_XA_X^T)^{-1}A_XA_{Y_i}^T=$$

$$=A_{Y_i}A_X^T(A_XA_X^T)^{-1}A_XA_{Y_i}^T\sigma_{Xu_i}\sigma_{Xu_i}^T(A_XA_X^T)^{-1}A_XA_{Y_i\,i}^T,$$

и поскольку это – симметричная матрица, то она положительно определена.

Разность правых нижних блоков равна

$$\sigma^2A_{X_i}A_{X_i}^T+\sigma_{X_iu_i}\sigma_{X_iu_i}^T-\sigma^2A_{X_i}A_{X_i}^T=\sigma_{X_iu_i}\sigma_{X_iu_i}^T,$$

и также представляет собой симметричную положительно определенную матрицу. Поскольку внедиагональные блоки матрицы S симметричны и уменьшаются на одну и туже симметричную величину, то результат разности матриц $(S-\sigma^2W)$ есть симметричная положительно определенная матрица. Поэтому

$$Q=\mathrm{cov}\left\{\begin{pmatrix}\hat{\beta}_i\\ \gamma_i\end{pmatrix}_{2LS}\right\}-\mathrm{cov}\left\{\begin{pmatrix}\hat{\beta}_i\\ \gamma_i\end{pmatrix}_{FSE}\right\}>0;\Rightarrow\ \mathrm{cov}\left\{\begin{pmatrix}\hat{\beta}_i\\ \gamma_i\end{pmatrix}_{2LS}\right\}>\mathrm{cov}\left\{\begin{pmatrix}\hat{\beta}_i\\ \gamma_i\end{pmatrix}_{FSE}\right\},\ (\forall i);$$

следовательно, оценка коэффициентов уравнения, полученная с помощью метода факторизации симультанных уравнений эффективна по отношению к оценке двухшагового метода наименьших квадратов. ■

Аналогично можно показать большую эффективность предложенного метода по сравнению с трехшаговым методом наименьших квадратов – второго по частоте использования метода.

Таким образом, в случае взаимозависимости переменных уравнений системы целесообразно использовать метод факторизации симультанных уравнений.

Выводы

Исследование природы симультанности переменных системы эконометрических уравнений показало, что в такой системе взаимосвязанными могут быть не только эндогенные переменные, но и экзогенные.

Все существующие к настоящему моменту времени методы оценивания параметров системы симультанных уравнений этот факт игнорируют, что приводит к смещенности оценок, получаемых с помощью этих методов.

Разработан метод факторизации системы симультанных переменных, который дает несмещенные оценки ее параметров, причем эти оценки являются более эффективными по сравнению с наиболее распространяемыми методами оценивания коэффициентов уравнений системы.

6. ЭКСПЛОРАТОРНЫЙ ФАКТОРНЫЙ АНАЛИЗ ВРЕМЕННЫХ РЯДОВ В ПРОГНОЗИРОВАНИИ РАЗВИТИЯ ЭКОНОМИЧЕСКОЙ СИСТЕМЫ.

6.1. Анализ состояния экономической системы с помощью эксплораторного факторного анализа

Целью прикладного исследования является, во-первых, анализ состояния экономической системы Украины, динамики уровня общей рискованности этой системы; во-вторых, расчет прогнозных значений основных макроэкономических показателей, выявление наиболее узких мест в развитии системы; в-третьих, построение системы симультанных уравнений для анализа возможных путей выведения экономической системы на путь поступательного развития.

Система макроэкономических показателей обширна, только в одном разделе свода статистической информации Украинского государственного комитета по статистике «Национальные счета» представлено свыше сотни экономических показателей. И поскольку наиболее интересным представляется рассмотрение периода «новейшей истории», т.е начиная с 2005 г. и завершая настоящим моментом, то ограниченность выборки определяет ограниченность числа рассматриваемых показателей. В основу анализа положены статистические данные Укргоскомстата [75] и НБУ [76], которые, в основном, являются квартальными итогами развития экономики Украины, поэтому объем выборки составляет число периодов планирования (семь лет), умноженное на четыре (квартала), т.е. – двадцать восемь. Поэтому целесообразно взять в рассмотрение такое число экономических показателей, которое не превышало бы половины объема выборки.

Выбор самих показателей определился их социальной значимостью. Естественно, как и любой другой выбор, он – субъективен, и, может быть, рассмотрения заслуживает какой-либо другой набор показателей. Но будем считать, что он задан экзогенно на основе работ ведущих экономистов. Возьмем в рассмотрение следующие макроэкономические показатели:

X_1 – валовой внутренний продукт (млн. грн.);
X_2 – учетная ставка НБУ (%);
X_3 – конечные потребительские затраты домашних хозяйств (млн. грн.);
X_4 – государственные расходы (млн. грн.);
X_5 – экспорт (млн. долларов США.);
X_6 – импорт (млн. долларов США.);
X_7 – имеющиеся доходы населения (млн. грн.);
X_8 – капитальные инвестиции (млн. грн.);
X_9 – инфляция (%);
X_{10} – безработица (тыс. чел.);
X_{11} – затраты на выполнение научных работ (млн. грн.);
X_{12} – внешний долг (млн. долларов США.).

Данные показатели имеют парные корреляции, приведенные в таблице 6.1

Таблица 6.1– Парные корреляции экономических показателей

	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}
X_1	1,00	0,05	0,86	0,46	0,70	0,76	0,93	0,56	-0,21	-0,72	0,35	0,49
X_2	0,05	1,00	0,17	0,04	-0,21	-0,16	0,07	0,20	0,49	0,07	0,64	0,53
X_3	0,86	0,17	1,00	0,58	0,60	0,78	0,87	0,74	-0,01	-0,64	0,31	0,56
X_4	0,46	0,04	0,58	1,00	0,27	0,42	0,43	0,58	-0,01	-0,34	0,25	0,14
X_5	0,70	-0,21	0,60	0,27	1,00	0,89	0,62	0,09	0,01	-0,42	-0,15	0,10
X_6	0,76	-0,16	0,78	0,42	0,89	1,00	0,63	0,33	-0,11	-0,63	0,00	0,35
X_7	0,93	0,07	0,87	0,43	0,62	0,63	1,00	0,63	-0,10	-0,61	0,24	0,38
X_8	0,56	0,20	0,74	0,58	0,09	0,33	0,63	1,00	-0,25	-0,56	0,28	0,51
X_9	-0,21	0,49	-0,01	-0,01	0,01	-0,11	-0,10	-0,25	1,00	0,65	0,17	-0,04
X_{10}	-0,72	0,07	-0,64	-0,34	-0,42	-0,63	-0,61	-0,56	0,65	1,00	-0,23	-0,49
X_{11}	0,35	0,64	0,31	0,25	-0,15	0,00	0,24	0,28	0,17	-0,23	1,00	0,59
X_{12}	0,49	0,53	0,56	0,14	0,10	0,35	0,38	0,51	-0,04	-0,49	0,59	1,00

Парные корреляции показателей за период с первого квартала 2005 г. по третий квартал 2011 г. вычислены относительно их одномерных временных трендов, что дает более точный результат по сравнению с их вычислением относительно средних величин.

Собственные величины корреляционной матрицы имеют вид (5,68 2,37 1,46 0,97 0,52 0,46 0,25 0,13 0,08 0,01 0,04 0,02); т.е. четыре значения больше либо равны единице (значение 0,97 практически равно единице). Таким образом, число общих стохастических факторов эксплораторной модели равно четырем. Нагрузки этих факторов на переменные приведены в таблице 6.2.

Таблица 6.2 – Факторные нагрузки на переменные за период с I 2005 по II 2011г.

	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}
F_3	0,639	-0,115	0,661	0,345	0,961	0,928	0,532	0,139	0,106	-0,4	-0,1	0,197
F_2	-0,057	0,562	0,287	0,213	-0,158	-0,101	0,094	0,213	0,806	0,4	0,3	0,212
F_1	0,708	0,11	0,665	0,362	0,081	0,238	0,805	0,774	-0,376	-0,6	0,4	0,439
F_4	-0,083	0,052	0,197	0,265	-0,218	0,183	-0,247	0,345	-0,282	-0,3	0,2	0,407

Для интерпретации факторного решения в каждом столбце таблицы 6.2 выделим наибольшее значение нагрузки (затененный цвет ячейки), тогда в каждой строке таблицы можно увидеть переменные, которые данный фактор нагружает в наибольшей степени. Фактор F_1 нагружает самое большое число показателей – восемь из двенадцати, поэтому он, практически, – генеральный фактор, поэтому его можно назвать социально-экономическим фактором (поэтому ему присвоен первый индекс). F_2 имеет наибольшее влияние на учетную ставку НБУ и инфляцию, т.е. он может рассматриваться как процентный фактор, или фактор процентного риска. Фактор F_3 имеет наибольшие нагрузки на X_5 и X_6 – экспорт и импорт, таким образом его можно интерпретировать как внешнеэкономический фактор. На четвертый фактор не выпало ни одной переменной с максимальным, по модулю, значением нагрузки в столбце, вследствие чего выделим наибольшие значения в строке этого фактора (голубой цвет ячейки). Эти значения приходятся на капитальные инвестиции и внешний долг. Если бы внешние заимствования шли бы на модернизацию экономики, а не на покрытие дефицита государственного, то этот фактор можно было бы назвать фактором развития. Но в настоящей реальности – это фактор обязательств.

На протяжении семи лет – с 2005 г. по 2011 г. в экономике Украины доминировали четыре стохастических фактора, поведение которых на протяжении этого периода представлено на рисунке 6.1.

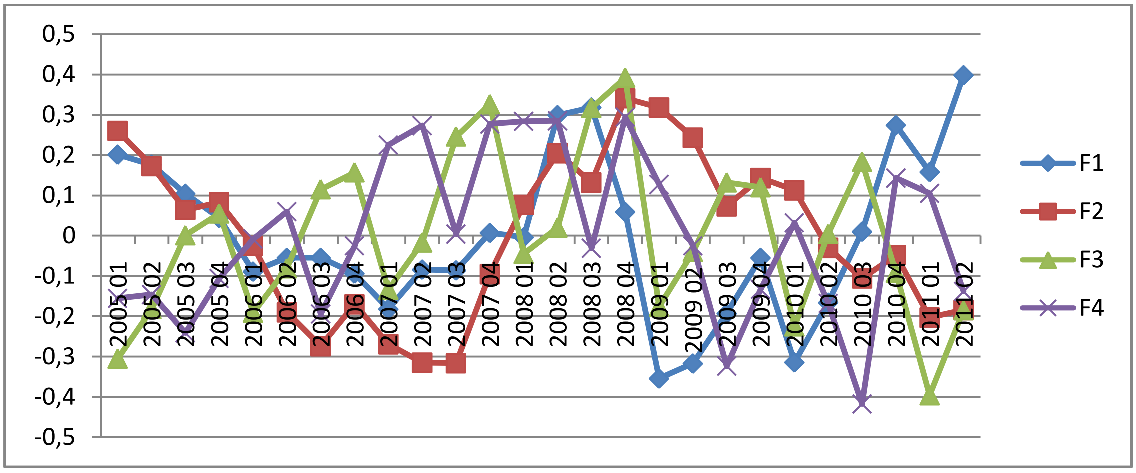

Рис. 6.1 – Динамика значений факторов 2005-11 гг.

График третьего фактора – социально-экономического – с начала рассматриваемого периода имеет «пилообразный» характер с общей тенденцией к росту до конца 2008 г. При этом отчетливо видна закономерность: три шага вперед, и один, но большой шаг назад, причем назад – к почти прежнему состоянию. На участке до 2009 г. было четыре резких отката этого фактора назад. Таким образом, поведение этого фактора посылало кризисные «весточки» в начале 2006, 2007 и 2008 гг. Все это сформировало почву для явного кризисного падения, которое приходится на конец 2008 г. Вместе с резким падением социально-экономического фактора происходит такое же падение внешнеэкономического и фактора обязательств, что сделало кризис чувствительным во всех элементах экономической системы. С конца 2008 г. намечается падение процентного фактора. С начала 2009 г. доминирующую роль приобретает внешнеэкономический фактор, общеэкономический фактор на этом временном интервале приобрел тенденцию к снижению.

На рисунке 6.2 изображен уровень риска на протяжении рассматриваемого периода.

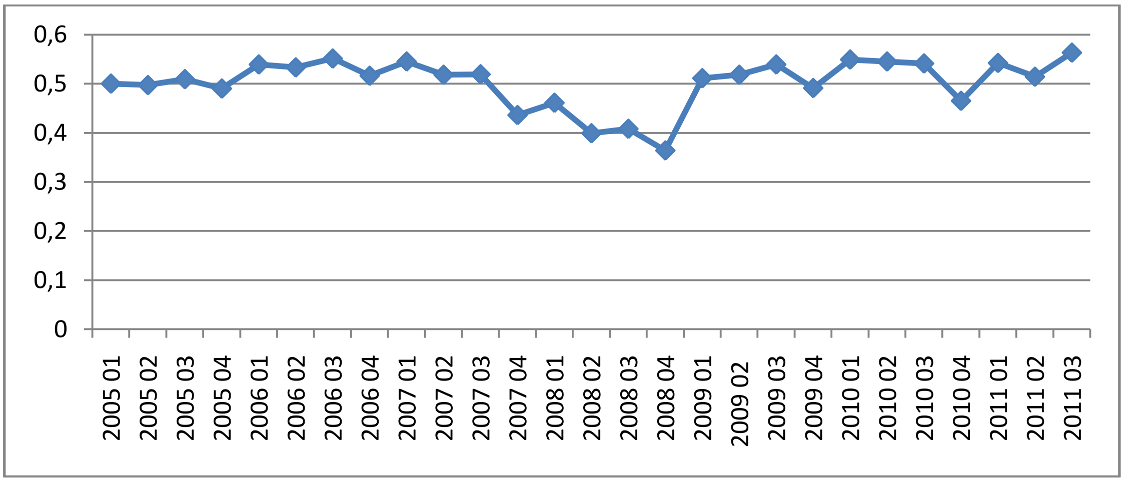

Рис. 6.2. – Уровень риска экономики Украины на протяжении с 2005 по 2011 гг.

Из профиля риска следует, что периоду новейшей экономической истории страны характерен достаточно высокий уровень риска, т.е. страна «привыкла» жить в постоянных кризисных условиях. Наименьшее значение уровня риска приходится на последний квартал 2008 г., с которым связывают начало заметного проявления кризиса в Украине. Между тем, произошедший в тот период откат в худшую сторону является лишь возвратом к обычному состоянию экономики.

Рассмотрим состояние факторов на разных интервалах рассматриваемого периода. Число факторов выбиралось равным числу собственных значений корреляционной матрицы, больших или равных единице. Факторные нагрузки на рассматриваемые экономические показатели на первом интервале 2005-7 гг. (временной интервал I) имели значения, приведенные в таблице 6.3.

Таблица 6.3 – Нагрузки факторов, период 2005-2007 гг.

	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}
F_1^I	0,727	-0,064	0,456	-0,036	0,822	0,62	0,65	0,11	0,314	-0,4	0,7	0,601
F_2^I	-0,057	1	0,03	0,048	-0,525	-0,188	0,023	0,091	-0,108	0,0	0,0	-0,374
F_3^I	0,631	0,059	0,887	0,907	0,145	0,675	0,631	0,961	0,636	-0,2	0,2	0,239

Выделенные ячейки таблицы (максимальные значения в столбцах) позволяют интерпретировать факторы следующим образом. Первый фактор F_1^I так же, как и для всего периода, можно считать социально-экономическим, но он нагружал несколько иные показатели. На начальном интервале рассматриваемого периода этот фактор оказывал доминирующее воздействие на валовой внутренний продукт, экспорт, доходы населения, безработицу, научные затраты и внешний долг, т.е. стимулировал, главным образом, сферу материального производства. Социальная составляющая этого генерального фактора до некоторой степени перешла к фактору F_3^I, который можно рассматривать как фактор потребления. Фактор F_2^I связан лишь с одной переменной – учетной ставкой НБУ, т.е. это – процентный фактор с доминированием процентной политики нашего центрального банка.

Таким образом, на временном интервале I развитие экономической системы определяется акцентами на производственную сферу, потребление и характеризуется заметной ролью НБУ.

Временной интервал II − 2006-8 гг. имел следующие факторные нагрузки (таблица 6.4).

Таблица 6.4 − Нагрузки факторов, период 2006-2008 гг.

	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}
F_1^{II}	0,96	-0,012	0,329	-0,037	0,843	0,669	0,728	0,055	-0,202	-0,8	-0,1	0,519
F_2^{II}	0,114	1	0,266	0,049	0,467	0,584	0,102	-0,274	0,864	0,3	0,7	-0,063
F_3^{II}	0,195	0,035	0,907	0,741	-0,247	-0,025	0,56	0,846	0,285	-0,1	0,0	-0,228

Экономический смысл факторов этого временного интервала остался таким же, как и для первого, однако имеются некоторые отличия по сравнению с предыдущим интервалом. Показатель «импорт» перешел в социально-экономический фактор, а инфляция и научные затраты вошли в сферу доминирования процентного фактора.

На втором временном интервале развитие экономической системы определялось тем же факторами, что и на первом интервале. Однако существенное отличие состоит в том, что в решении инфляционных проблем ведущую роль занял центробанк.

Нагрузки факторов на третьем временном интервале с 2007 по 2009 гг. принимали следующие значения (таблица 6.5).

Таблица 6.5 – Нагрузки факторов, период 2007-2009 гг.

	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}
F_3^{III}	0,449	0,101	0,719	0,756	0,147	0,356	0,635	0,832	-0,039	-0,6	0,5	0,565
F_2^{III}	0,182	0,669	0,107	-0,219	0,204	0,219	0,078	-0,361	0,622	-0,2	0,7	0,715
F_1^{III}	0,777	0,289	0,666	0,304	0,907	0,909	0,621	0,381	-0,048	-0,7	0,4	0,419

F_3^{III} – фактор потребления;

F_2^{III} – процентный фактор;

F_1^{III} – социально-экономический фактор, таким образом, факторы на этом временном интервале – те же, что и на трех предшествующих интервалах. Существенным отличием распределения показателей по факторам является то, что внешний долг от генерального фактора перешел в процентный фактор.

Факторные нагрузки на переменные на четвертом временном интервале – 2009-2010 гг. приведены в таблице 6.6.

Таблица 6.6 – Нагрузки факторов, период 2008-2010 гг.

	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}
F_1^{IV}	0,829	-0,031	0,626	0,428	0,953	0,86	0,626	0,255	-0,691	-0,8	0,5	0,525
F_2^{IV}	-0,076	0,948	-0,042	0,111	-0,134	-0,292	0,101	0,375	0,328	0,2	0,0	0,246
F_3^{IV}	0,554	-0,264	0,656	0,149	0,263	0,319	0,75	0,552	-0,244	-0,3	0,1	0,576

Факторы в матрице факторных нагрузок на этом интервале меняют свой порядок, но остаются с неизменной экономической интерпретацией в сравнении с предшествующими временными интервалами. Выделим переход под непосредственное влияние первого фактора показателя «государственные расходы».

На четвертом временном интервале проявились тенденции такие же, как и для всего периода рассмотрения статистических данных; а именно: функционирование экономической системы определяется четырьмя факторами. Факторные нагрузки – таблице 6.7.

Таблица 6.7 – Нагрузки факторов, период 2009-2011 гг.

	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}
F_1^{V}	0,646	-0,297	0,571	0,71	0,447	0,598	0,382	0,325	-0,442	-0,6	0,6	-0,051
F_2^{V}	-0,383	0,813	-0,249	0,127	0,039	-0,118	-0,395	0,074	-0,192	0,5	-0,3	-0,125
F_3^{V}	0,554	-0,235	0,773	-0,012	0,871	0,676	0,697	0,638	-0,235	-0,2	-0,6	0,558
F_4^{V}	0,357	0,143	0,003	0,006	0,186	-0,29	0,437	-0,012	-0,665	-0,6	0,4	-0,561

Однако в отличие от предшествующих интервалов нагруженность факторами переменных на этом временном интервале приобретает иной характер, и к тому же появляется новый фактор, одновременно нагружающий только две переменные: инфляцию и безработицу, т.е. те показатели, которые в наиболее ярком свете характеризуют кризисные явления. Таким образом, с 2009 г. выявляется фактор кризиса.

Если ВВП, учетную ставку НБУ и конечное потребление домашних хозяйств считать определяющими для факторов социально-экономического, процентного и потребления, то они, при наличии фактора кризиса, имеют несколько иную направленность. Нагрузка первого фактора на государственные расходы возрастает вдвое, так же растет нагрузка на научные затраты, а это указывает на то, что в момент обострения кризиса особую роль приобретает политика государственных расходов и наука. В этом же факторе присутствует безработица, что свидетельствует о важности решения ее проблемы на этом этапе. При чем с одинаковым весом показатель безработицы вошел и в первый фактор, и в четвертый. Процентный фактор «покинула» инфляция, что указывает на слабую возможность центрального банка регулировать инфляционные процессы.

Фактор потребления по числу нагружаемых переменных приблизился к генеральному. Проблемы потребления влияют и на экспорт, и на импорт, а также – на инвестиции и внешний долг. Среди переменных этого также затраты на науку, с таким же весом, как и у первого фактора. Т.е. роль науки при наличии фактора кризиса становится более заметной.

Рассмотренные временные интервалы позволяют судить о том, какова динамика изменений структуры модели эксплораторного факторного анализа на относительно непродолжительном интервале времени. Хотя методы эксплораторного факторного анализа позволяют обходить проблему мультиколлинеарности благодаря тому, что отображение пространства реальных наблюдаемых величин в факторное пространство, имеющее, во-первых, значительно меньшую размерность; а, во-вторых, ортогональный координатный базис; минимальная ширина интервала должна определяться объемом выборочных данных, приходящегося на него. Для двенадцати показателей двенадцать точек наблюдения – это тот минимум, который должна иметь выборка. Динамика факторной структуры указывает на стабильность самих стохастических факторов, лишь отдельные переменные переходили от одного фактора к другому на разных временных интервалах. Но самых характерным для кризисных IV и V интервалов становится схождение двух показателей – инфляции и безработицы к одному фактору: на I, II и III интервалах эти показатели относились к разным факторам. Более того, на V интервале эти два показателя образовали новый стохастический фактор, который не имел место на всех других интервалах. Рассмотрим еще два более продолжительных интервала: до условного начала кризиса (январь 2005 г. – январь 2008 г., интервал VI) и после его начала (с февраля 2008г.– интервал VII). Факторные нагрузки для этих периодов представлены в таблицах 6.8 и 6.9.

Таблица 6.8 – Нагрузки факторов, период 2005-2008 гг.

	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}
F_3^{VI}	0,578	-0,059	0,577	0,702	-0,08	0,136	0,681	0,784	-0,149	-0,6	-0,4	-0,375
F_2^{VI}	-0,269	0,995	0,181	-0,205	-0,305	0,195	-0,167	-0,284	0,637	0,4	0,5	-0,048
F_1^{VI}	0,635	0,078	0,793	0,372	0,746	0,955	0,525	0,39	0,663	0,1	0,8	0,922

Таблица 6.9 – Нагрузки факторов, период 2008-2011 гг.

	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}
F_3^{VII}	0,387	-0,2	0,545	0,005	0,651	0,634	0,428	0,373	-0,257	-0,4	-0,8	0,71
F_2^{VII}	-0,561	0,364	-0,212	-0,116	-0,652	-0,555	-0,294	0,279	0,634	0,6	-0,3	-0,16
F_1^{VII}	0,667	-0,323	0,806	0,662	0,375	0,48	0,746	0,759	-0,477	-0,6	0,5	0,351

Первый и второй стохастические факторы имеют ту же экономическую интерпретацию для обоих интервалов, но с некоторыми нюансами. Так на VI-м интервале экспорт, импорт и внешний долг входят в социально-экономический фактор, а на интервале VII – они переходят во внешнеэкономический фактор. Т.е. третьи факторы у этих интервалов различны. На предкризисном VI интервале – этот фактор имеет преимущественную нагрузку на инвестиции, государственные расходы и доходы населения. А поскольку доходы населения, как и государственные расходы, могут представлять собой ресурсную базу для внутренних инвестиций, то это – инвестиционный фактор, или фактор инвестиционного потенциала. На VII интервале с начала кризиса третий фактор становится внешнеэкономическим.

К процентному фактору на VII интервале одновременно перешли инфляция и безработица. Соединение этих показателей в один отмечается для самых кризисных моментов в рассматриваемом периоде. Нагрузка одного фактора говорит об одной причине динамики этих факторов. Причем знаки нагрузок у этих показателей и у показателя «учетная ставка» одинакова, т.е. у них имеет место однонаправленная тенденция развития. Величины нагрузок процентного фактора у безработицы и инфляции больше, чем у ставки процента, таким образом, два показателя, являющихся наиболее яркими характеристиками кризиса, доминируют над регуляторной политикой центрального банка на VII временном интервале.

На этом интервале безработица с такой же факторной нагрузкой, но со знаком минус, входит в первый фактор, оказывая негативное воздействие на социально-экономические показатели. Однако это является характерным для всех рассмотренных временных периодов.

Можно сделать вывод, что, индикатором углубления кризиса является переход показателя безработицы в тот фактор, который в наибольшей степени нагружает инфляцию. Схождение этих показателей отмечается для интервалов 2008-10 гг., 2008-11 гг., 2009-11 гг. Таким образом, экономический кризис начался в 2008 г. и продолжался до 2011 г.

6.2. Прогноз значений стохастических факторов и экономических показателей

Анализ динамики стохастических факторов, графическое изображение которых представлено на рис. 6.1, показывает определенную цикличность их поведения. Социально-экономический фактор имеет явно обозначенную тенденцию: пик его значений приходится на середину года, а низшая точка – на начало года, т.е. периодичность значений этого фактора составляет четыре квартала. Такую же периодичность имеет фактор обязательств с пиком во втором квартале и с наинизшей точкой спада в третьем.

Внешнеэкономический фактор на протяжении рассматриваемого периода на разных временных участках имел различную периодичность. С начала 2005 г. до начала 2007 г. наблюдалось снижение его значений, после этого следует рост вплоть до третьего квартала 2008 г. И с первого квартала 2009 г. у этого фактора начинает прослеживаться годовая периодичность с отрицательными пиками, случающимися в начале года.

Процентный фактор имеет четырехлетнюю периодичность. Пики его значений приходятся на начало 2005 и 2009 гг., а низшие точки – на первую половину 2007 и 2011 гг.

Таким образом, для факторной модели, построенной на основании ежеквартальной статистики, наиболее подходящий порядок авторегрессии равен четырем.

На основании разработанной модели авторегрессии стохастических факторов сделан прогноз их значений. Динамика значений факторов на период от второго квартала 2011 г. до второго квартала 2013 г. приведена на рисунке 6.3.

Тенденции в поведении социально-экономического фактора, выявленные на основании имеющейся статистики, сохранились и для прогноза. Низшие точки значений фактора приходятся на начало 2012 – 2013 г. Высшая точка 2012 года, приходящаяся на его середину, выходит в положительную область, в отличие от 2011 г., когда она находилась в отрицательной. С 2012 г. у этого фактора намечается тенденция к росту, и, в целом, в социально-экономическом аспекте ситуация в двенадцатом году более благоприятная по сравнению с одиннадцатым годом.

С третьего квартала 2011 г. фактор обязательств также приобретает тенденцию к росту, сохраняя при этом четырехквартальную периодичность.

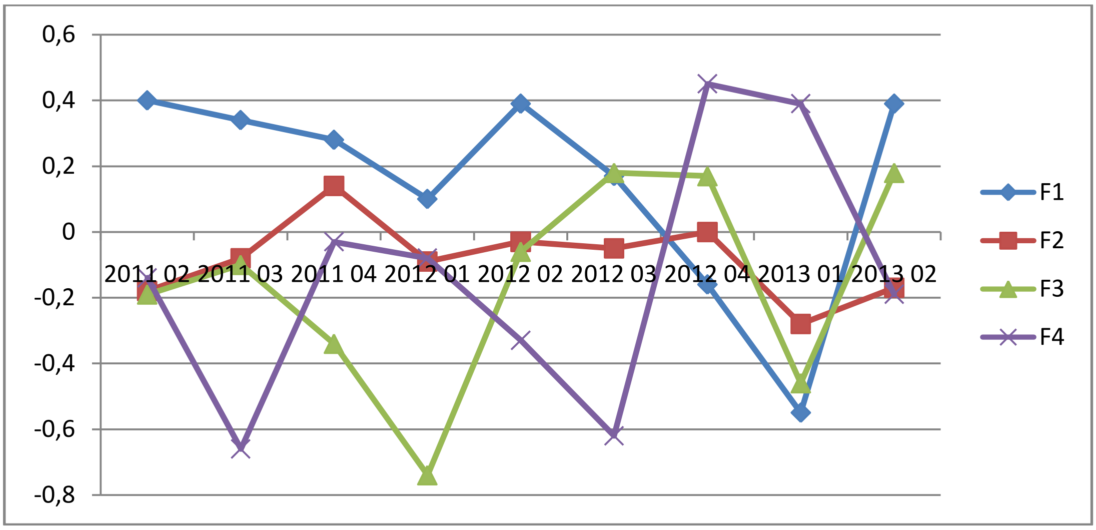

Рис. 6.3 – Прогнозные значения стохастических факторов

Если до 2011 г. динамика внешнеэкономического фактора имела характер: «три шага вперед – один назад», то в последующие годы она меняет такую особенность на противоположную: «один шаг вперед – три шага назад» с общей тенденцией к снижению уровня фактора.

Прогнозные значения процентного фактора, что особенно заметно на рисунке 6.3, имеют наименьшие колебания относительно своей средней величины. Это говорит о том, что процентная составляющая в развитии экономической системы страны будет наиболее стабильной. Однако среднее значение этого фактора находится в отрицательной области, что указывает на недостаточную ожидаемую роль регулятора финансовых ринков.

Можно сделать вывод, что в целом, стохастические факторы не будут благоприятствовать развитию экономики Украины в последующие два года.

Общее поведение факторов, позволяет оценить в целом уровень рискованности в экономической системе. Значение меры риска (4.5) указано на рисунке 6.4.

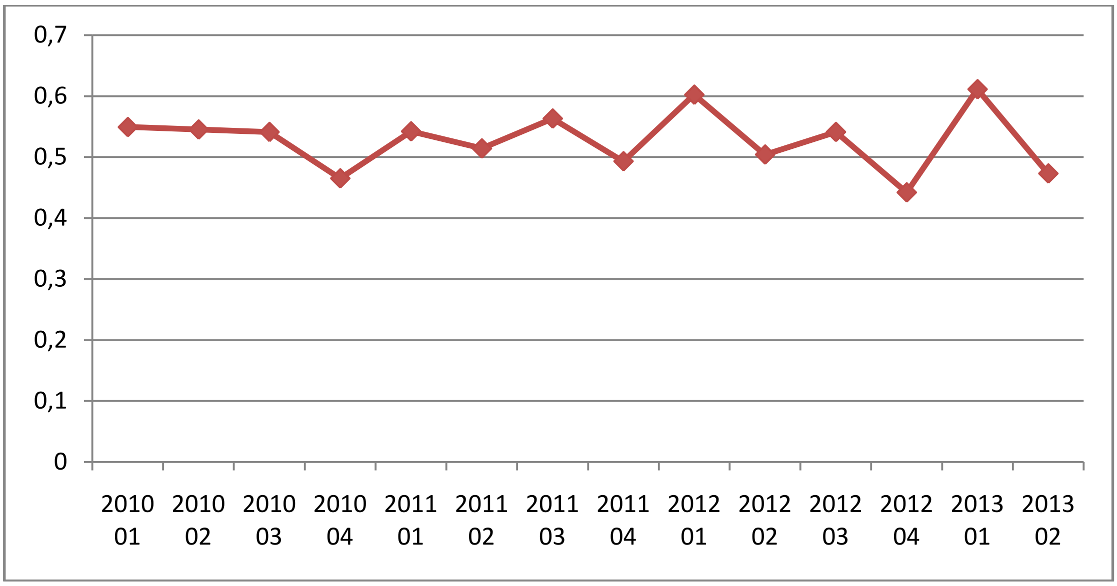

Рис. 6.4 – Реальные и прогнозные значения меры риска

Из этого рисунка видно, что уровень риска остается стабильно высоким и не опускается ниже значения в 0,45 единиц, но, в основном, значение показателя лежит выше отметки в 0,5 единиц.

Перейдем к прогнозу экономических показателей. Результаты прогноза приведены в таблице 6.10.

Таблица 6.10 – Прогноз макроэкономических показателей

Год квартал	ВВП (млн. грн.)	Учетная ставка НБУ (%)	Кон. потрб. ДХ (млн. грн.)	Гос. расх. (млн. грн.)	Экспорт (млн.USD)	Импорт (млн.USD)
2011 03	327 000	9,65	208 000	62 700	165 000	180 000
2011 04	335 000	9,89	214 000	64 800	168 000	183 000
2012 01	340 000	10,11	220 000	66 500	169 000	185 000
2012 02	342 000	10,26	223 000	67 700	171 000	186 000
2012 03	342 000	10,26	226 000	68 500	172 000	187 000
2012 04	344 000	10,12	228 000	69 100	176 000	189 000
2013 01	349 000	9,84	231 000	69 800	181 000	193 000
2013 02	359 000	9,52	235 000	71 000	188 000	200 000

Таблица 6.10 (продолжение) – Прогноз макроэкономических показателей

Год квартал	Дох.насел. (млн.грн.)	Инвестиции (млн.грн.)	Инфляция (% к 12 2004)	Безра-ботица (тыс.чел.)	Затраты на выпол-нение на-учных работ (млн.грн.)	Долг (млн. USD)
2011 03	238 000	60 500	233	1 810	2 630	134 000
2011 04	245 000	65 000	240	1 830	2 710	138 000
2012 01	250 000	66 600	246	1 870	2 780	142 000
2012 02	254 000	64 500	252	1 930	2 840	146 000
2012 03	255 000	59 000	258	2 000	2 880	148 000
2012 04	258 000	51 700	264	2 060	2 910	150 000
2013 01	262 000	44 900	269	2 090	2 930	153 000
2013 02	269 000	41 100	273	2 090	2 950	155 000

Наиболее характерной тенденцией, судя по прогнозу, является то, что инвестиции сокращаются, а внешний долг растет. Причем снижение уровня поступления инвестиций к середине 2013 г. по отношению к началу 2012 г. ожидается на 38%, а для аналогичного периода внешний долг вырастет на 9%. Для того, чтобы экономика развивалась поступательно, необходима обратная тенденция: внешний долг должен сокращаться, а инвестиции – расти.

Прогнозное значение учетной ставки для всего прогнозного периода остается достаточно стабильным на уровне 9,8%; однако в 2013 г. наметится тенденция к снижению ее уровня.

Номинальный валовой внутренний продукт будет расти, и его рост к концу прогнозного периода составит 5,5%. Будут расти и другие макроэкономические показатели. Конечное потребление домашних хозяйств вырастет на 6,8%. Таким же ожидается рост государственных расходов, который составит те же 6,8%. При этом доля государственных расходов в валовом национальном продукте в 2013 г. по сравнению с 2011 г. практические не изменится: во втором квартале 2011 г. она составляла 19,0%; а в том же квартале 2013 г. она составит 19,7%. Доходы населения возрастут на 7,6%; затраты на выполнение научных работ – на 6,1%.

Негативная тенденция сохранится и для уровня безработицы: в течение прогнозного периода в один год она вырастет на 11,8%.

Сохранится и преобладание импорта товаров над экспортом. Причем экспорт увеличится на 11,2%; а импорт – на 8,1%. Таким образом, наметится тенденция к сокращению отрицательного сальдо во внешнеэкономической деятельности.

Однако все позитивные тенденции, на которые указывает прогноз, нивелируются ожидаемым ростом инфляции. Во втором квартале 2013 г. по сравнению с началом 2012 г. относительный рост инфляции составит 11,0%. Если принять во внимание, что уровень инфляции в начале 2012 г. составляет 246% (относительно декабря 2004 г.), а к концу прогнозного периода ее значение равно 273% (относительно декабря 2004 г.), то фактический рост инфляции составит 27%.

Таким образом, в реальном исчислении все макроэкономические показатели, выражаемые в национальной денежной единице, будут снижаться, поскольку их процентный рост значительно ниже этой отметки. И это при том, что внешний долг, измеряемый в долларах США, будет расти.

6.3. Построение модели симультанных уравнений экономической системы Украины.

Основные макроэкономические показатели, которые определяют состояние экономической системы, и которые должны быть объяснены моделью симультанных уравнений, т.е. эндогенные переменные, есть следующие:

Y_1 – валовой внутренний продукт (млн. грн.);
Y_2 – конечное потребление домашних хозяйств (млн. грн.);
Y_3 – экспорт (млн. долларов США);
Y_4 – доходы населения (млн. грн.);
Y_5 – инфляция (%, к декабрю 2004 г.);
Y_6 – безработица (тыс. чел.).

В модели предполагается, что внешними, экзогенными переменными, величины которых, либо задаются, либо достаточно жестко контролируются, являются следующие:

X_1 – учетная ставка Национального банка Украины (%);
X_2 – государственные расходы (млн. грн.);
X_3 – импорт (млн. долларов США);
X_4 – инвестиции (млн. грн.);
X_5 – затраты на выполнение научных работ (млн. грн.);
X_6 – внешний долг (млн. долларов США).

Для раскрытия структуры модели симультанных уравнений оценивается матрица нагрузок инструментальных факторов, в предположении, что число таких факторов равно числу эндогенных переменных, и числу уравнений в системе, т.е. шести.

Таблица 6.11 – Факторные нагрузки на эндогенные и экзогенные переменные.

	Эндогенные переменные						Экзогенные переменные					
	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6	X_1	X_2	X_3	X_4	X_5	X_6
Y_3	**0,427**	0,271	**0,757**	0,261	0,148	-0,281	0,163	-0,034	**0,652**	-0,323	0,074	0,197
Y_1	**0,726**	**0,55**	**0,293**	**0,844**	-0,295	**-0,547**	0,116	0,283	0,267	**0,686**	0,223	0,293
Y_6	0,264	**0,381**	-0,09	0,094	**-0,495**	**-0,679**	0,161	0,227	0,259	**0,393**	**0,385**	**0,578**
Y_4	-0,046	0,069	0,125	**-0,143**	-0,088	-0,09	0,018	**0,288**	0,244	**0,426**	-0,168	0,128
Y_5	-0,011	0,257	-0,192	0,083	**0,715**	0,276	**0,903**	0,15	-0,142	0,24	**0,522**	0,414
Y_2	**-0,389**	**-0,645**	**-0,483**	**-0,441**	-0,194	0,106	**0,362**	**-0,452**	**-0,583**	-0,319	0,167	-0,016

Для выявления структуры модели симультанных уравнений в каждом столбце эндогенной переменной определено максимальное значение нагрузки инструментального фактора (затененная ячейка), таким образом, каждый фактор воспроизводит соответствующую эндогенную переменную. Все переменные «разложились» по своим инструментальным факторам, кроме Y_4, и один из инструментальных факторов остался «незадействованным» в таком распределении. Этому фактору «припишем» переменную Y_4. Далее, в каждой строке отмечаем значимые факторные нагрузки (жирный шрифт), соответствующие переменные войдут в уравнение модели для соответствующей эндогенной переменной. Структура модели симультанных уравнений имеет вид:

$$Y_1 \Leftarrow Y_2, Y_3, Y_4, X_4;$$
$$Y_2 \Leftarrow Y_1, Y_3, Y_4, X_1, X_2, X_3;$$
$$Y_3 \Leftarrow Y_1, X_3;$$
$$Y_4 \Leftarrow X_2, X_4;$$
$$Y_5 \Leftarrow X_1, X_5;$$
$$Y_6 \Leftarrow Y_2, Y_5, Y_4, X_4, X_5, X_6.$$

В уравнениях, построенных на основании этой структуры зависимостей, следует учесть наличие свободного члена, который трактуется как одних из элементов списка экзогенных переменных.

Оценка коэффициентов системы методом факторизации симультанных уравнений (5.29) дает следующие количественные зависимости:

$$Y_1 = -0{,}288Y_2 + 0{,}638Y_3 + 1{,}063Y_4 + 9{,}2865Y_6 + 0{,}0{,}218X_4 + U_1;$$
$$Y_2 = 0{,}215Y_1 - 0{,}421Y_3 + 0{,}446Y_4 + 12{,}856X_1 + 0{,}232X_2 + 0{,}486X_3 + U_2;$$
$$Y_3 = 0{,}117Y_1 + 0{,}703X_3 + U_3;$$
$$Y_4 = 9439{,}7 + 3{,}479X_2 - 0{,}032X_4 + U_4;$$
$$Y_5 = -1{,}317X_1 + 0{,}092X_5 + U_5;$$
$$Y_6 = -0{,}014Y_2 + 21{,}721Y_5 + 529{,}71 + 0{,}002X_4 - 0{,}151X_5 - 4{,}205X_6 + U_6.$$

По коэффициентам модели симультанных уравнений можно судить о характере зависимостей между эндогенными переменными. Но для выработки рекомендаций, какие значения должны принимать экзогенные переменные, чтобы достичь определенных значений макроэкономических показателей, необходимо найти коэффициенты приведенной матрицы, т.е. параметры модели, связывающей все эндогенные переменные с экзогенными. Оценки этих величин, полученные с помощью метода факторизации симультанных уравнений, приведены в таблице 6.12.

Таблица 6.12 – Оценки метода факторизации симультанных уравнений элементов приведенной матрицы

	X_0	X_1	X_2	X_3	X_4	X_5	X_6
Y_1	13250	-272,38	2,964	0,37	0,211	17,315	-39,243
Y_2	6406,52	-32,362	2,274	0,251	0,021	2,874	-6,515
Y_3	1553,69	-31,939	0,348	0,746	0,025	2,03	-4,602
Y_4	9439,66	0	3,479	0	-0,032	0	0
Y_5	0	-1,317	0	0	0	0,092	0
Y_6	438,351	-28,142	-0,032	-0,004	0,002	1,814	-4,112

В таблице X_0 – свободный член.

Из таблицы видно, что учетная ставка НБУ отрицательно сказывается на всех эндогенных переменных, кроме доходов населения. Увеличение процента кредитования коммерческих банков центральным банком страны ведет к сокращению и валового внутреннего продукта, и конечного потребления домашних хозяйств, и экспорта, однако рост учетной ставки НБУ ведет к снижению инфляции и безработицы. Государственные расходы положительно воздействуют на ВВП, потребление ДХ, экспорт и доходы населения, но не оказывают никого влияния на инфляцию, и ведут к снижению безработицы. Импорт, не смотря на отрицательное сальдо внешнеторгового баланса страны, положительно влияет на главные макроэкономические показатели, к тому же при этом знак коэффициента при Y_6 указывает на снижение уровня безработицы при увеличении импорта. Инвестиции так же оказывают положительное влияние на эндогенные переменные, за исключением доходов населения. Затраты на выполненные научные работы стимулируют рост значений всех макроэкономических показателей, в том числе и инфляции, и безработицы. Относительно последней экзогенной переменной можно сделать однозначное заключение: внешний долг оказывает негативное воздействие и на валовой внутренний продукт, и на конечное потребление домашних хозяйств, и на экспорт, и на доходы населения. Причем удельный вес коэффициента при этом показателе самый большой по сравнению с коэффициентами при других экзогенных переменных, выраженных в денежной форме, что указывает на особую значимость этого показателя для формирования оптимальных значений эндогенных переменных.

Оценки приведенной матрицы системы симультанных уравнений с помощью обычного двухшагового метода наименьших квадратов дал следующие результаты (таблица 6.13).

Таблица 6.13 – Оценки двухшагового метода наименьших квадратов элементов приведенной матрицы

	X_0	X_1	X_2	X_3	X_4	X_5	X_6
Y_1	15425,2	-222,39	4,113	0,003	0,136	16,187	-1,215
Y_2	3387,32	1320,42	2,128	0,372	-0,07	-5,061	0,38
Y_3	1808,75	-26,077	0,482	0,703	0,016	1,898	-0,143
Y_4	9439,66	0	3,479	0	-0,032	0	0
Y_5	0	-1,317	0	0	0	0,092	0
Y_6	481,406	-47,433	-0,03	-0,005	0,003	2,078	-0,156

Сравнение элементов приведенной матрицы, оцененных методом факторизации и двухшаговым методом, говорит о тождественности основных закономерностей. Однако имеются принципиальные отличия экономического характера этих методов. Так двухшаговый метод дает положительное значение коэффициента при переменной «внешний долг» в уравнении для переменной «конечное потребление домашних хозяйств», а это означает, что с ростом внешнего долга будет расти это потребление.

Метод факторизации дает противоположное значение знаку этого коэффициента, что в большей степени согласуется с экономической логикой, поскольку домашние хозяйства формируют свой доход, идущий на потребление, в самой последней степени за счет средств внешнего заимствования. Более того, обслуживание обязательств по внешнему долгу в наибольшей степени ложится на домашние хозяйства. Поэтому рост долга как прямым, так и косвенным образом может вести лишь к снижению потребления.

Имеется еще одно принципиальное отличие в знаках коэффициентов приведенной матрицы, и оно касается опять же конечного потребления домашних хозяйств. Двухшаговый метод наименьших квадратов показывает, что с ростом учетной ставки центрального банка растет и потребление, причем вклад процентной ставки в значение потребления – самый весомый, поскольку величина коэффициента равна 1320,4 млн. грн. на 1% (государственные расходы дают прирост потребления в сумме 2,128 млн. грн. на единицу расходов). Метод факторизации указывает на обратную тенденцию во взаимосвязи между учетной ставкой и конечным потреблением домашних хозяйств.

Существенные отличия между двумя методами касаются, практически, одной эндогенной переменной – конечного потребления домашних хозяйств. Остальные зависимости в основных чертах согласуются между собой. Но, если проигнорировать симультанную природу взаимосвязей между эндогенными переменными, а также и между экзогенными, и воспользоваться обычным методом наименьших квадратов для получения оценок коэффициентов при переменных, то уравнения регрессии приобретают такой экономический смысл, который лишь в малой степени согласуется с реальным представлением о характере макроэкономических закономерностей. Коэффициенты регрессии, полученные с помощью OLS-метода приведены в таблице 6.14.

Таблица 6.14 – Оценки обычного метода наименьших квадратов коэффициентов регрессии

	X_0	X_1	X_2	X_3	X_4	X_5	X_6
Y_1	460,049	-2789	-0,27	1,04	0,29	60,926	17,986
Y_2	579,995	-678,86	0,534	0,36	0,074	8,273	656,694
Y_3	8980,26	147,879	0,272	0,799	-0,169	-4,367	105,019
Y_4	-519,139	-2280,4	0,342	0,456	0,137	24,976	593,473
Y_5	50,892	0,802	0,001	0	0	-0,005	1,07
Y_6	1708,5	14,896	0,018	-0,007	-0,006	-0,345	10,568

Из таблицы видно, что главной переменной, задающей положительные значения зависимых переменных (в симультанной модели – это эндогенные переменные), является внешний долг. Он «положительно влияет» на все без исключения переменные, стоящие в левой части уравнений. Уравнение доходов населения имеет отрицательный свободный член, таким образом, существует такое значение каждой из независимых переменных, при котором доходы будут равны нулю, при условии, что остальные переменные правой части уравнения фиксированы на определенном уровне. Так, если при фиксации на значении одна единица переменных X_2, X_3, X_4, X_5, X_6, учетную ставку принять равной 0,044%, доходы населения будут равны нулю, а при больших этой величины значениях ставки процента доходы станут отрицательными.

Из этого следует вывод, что игнорирование симультанной природы экономических переменных ведет к существенному искажению реальной картины зависимостей между анализируемыми показателями. Достоверные заключения для принятия решений можно делать лишь на основании моделей, в максимальной степени «очищенных» от эффекта симультанности.

Проанализируем влияние экзогенных переменных, переменных значения которых можно формировать сознательными реальными действиями, на результирующие показатели.

Первый – валовой внутренний продукт. Уменьшение Национальным банком Украины учетной ставки на один процент приведет к росту ВВП на 272,38 млн. грн. Рост государственных расходов на 1 млн. грн. даст прирост ВВП на 2,964 млн. грн. Рост импорта на 1 млн. долларов США увеличит внутренний продукт на 0,37 млн. грн. Привлечение инвестиций в сумме один млн. грн. ведет к возрастанию ВВП на 0,211 млн. грн. Затраты на науку оказываются самыми рентабельными: вложение в разработку научных работ в размере одного млн. грн. оборачивается семнадцатикратным приростом валового внутреннего продукта Украины. Также существенна роль внешнего долга: сокращение обязательств перед внешними кредиторами в размере одного миллиона американских долларов увеличивает ВВП на 39, 243 млн. грн. Следовательно, приоритетными направлениями для решения задачи роста внутреннего продукта являются научные исследования и сокращение внешнего долга.

Второй – конечное потребление домашних хозяйств, показатель характеризующий уровень жизни граждан страны. Снижение учетной ставки НБУ ведет к росту потребления – один процент соответствует 32,362 млн. грн. Рост государственных расходов положительно сказывается на благосостоянии домашних хозяйств – одна единица прироста расходов дает 2,3 единиц прироста потребления. Не в какой степени существенно, но положительно влияют на конечное потребление ДХ импорт и инвестиции. Увеличение расходов на науку на одну единицу денежных вложений ведет к почти троекратному приросту благосостояния населения. Сокращение внешнего долга на один млн. USD даст прирост конечного потребления на 6,515 млн. грн. Таким образом, как и для показателя ВВП, для потребления домашних хозяйств самими приоритетными направления являются наука и внешний долг.

Третий – экспорт. Так же как и предыдущие показатели зависит от снижения уровня учетной ставки центрального банка: один процент дает прирост экспорта на 31,939 млн. грн. Государственные расходы, импорт и инвестиции в случае неотрицательных приростов дают рост экспорту. Затраты на науку имеют существенный мультипликатор по отношению к экспорту: 1 млн. грн. ведут к двукратному увеличению вывозимой за пределы страны продукции и оказываемых там услуг. Сокращение внешнего долга на один млн. долларов США дают рост экспорта на 4,602 млн. грн.

Четвертый – доходы населения. Они главным образом определяются постоянной составляющей – свободным членом, и зависят от роста государственных расходов. Учетная ставка, импорт, затраты на научные исследования и внешний долг не оказывают влияния на уровень доходов населения. Инвестиции в своем росте ведут хоть и незначительному, но все же снижению доходов.

Пятый – инфляция. Не имеет постоянной составляющей: свободный член равен нулю. И главным образом зависит от политики Национального банка Украины. Рост затрат на выполнение научных работ ведет к некоторому росту инфляции: увеличение объема научных исследований на один млн. грн. дает прирост инфляции в размере 0,092%, эта величина вряд ли может считаться существенной.

Шестой – безработица. Имеет «несжимаемый остаток» в размере 438351 человека. Учетная ставка НБУ, государственные расходы, импорт и внешний долг при своем увеличении снижают уровень безработицы. Однако увеличение расходов на науку в сумме один млн. грн. ведет к росту безработицы почти на две тысячи человек. Это имеет свое объяснение: внедрение научных разработок в сферу материального производства, как правило, ведет к высвобождению трудовых ресурсов. На уровень безработицы оказывает существенное влияние величина валового внутреннего продукта: рост этого эндогенного показателя ведет к снижению безработицы. Поэтому именно с ростом экономической системы автоматически будет решаться и проблема занятости населения.

Таким образом, наиболее приоритетными направления практической политики государственных исполнительных органов являются стимулирование научных исследований и сокращение внешнего долга.

Матрица коэффициентов приведенной факторной модели позволяет ответить на вопрос, какими должны быть значения экзогенных переменных при заданном наборе величин эндогенных показателей. Поставим задачу: какие значения должны иметь учетная ставка НБУ, государственные расходы, импорт, инвестиции, затраты на выполнение научных работ и внешний долг, чтобы удвоились величины ВВП, конечного потребления домашних хозяйств и доходы населения, инфляция осталась без изменений, а безработица – отсутствовала (таблица 6.15).

Таблица 6.15 – Набор величин эндогенных переменных

	Эндогенные переменные					
	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6
Допущение	600000	400000	200000	500000	220	0

Поставленная задача решается с помощью пакета Mathcad, фрагмент файла с программой, в которой оценивались параметры модели эксплораторного факторного анализа, а также матрицы коэффициентов модели симультанных уравнений, приведен на рисунке 6.5. На нем у вектора значений экзогенных переменных: первый элемент – свободный член (равный 0 для начального приближения, и -14 – в решении), обозначения остальных элементов – как в таблице 6.11.

Начальные значения экзогенных переменных

$$x := (0\ \ 7.75\ \ 60202\ \ 190000\ \ 45934\ \ 2125\ \ 123)$$

Заданные значения эндогенных переменных

$$Y := (600000\ \ 400000\ \ 200000\ \ 500000\ \ 220\ \ 0)$$

Программа

Given

$$Y = x \cdot \Pi$$

$$X := \mathrm{Find}(x)$$

Решение

$$X = (-14\ \ 0\ \ 183441\ \ 184499\ \ 257184\ \ 2382\ \ -1915)$$

Рис. 6.5 – Фрагмент Mathcad-файла с поиском значений экзогенных переменных (П –матрица приведенной модели симультанных уравнений0

Таким образом, чтобы произошло удвоение основных макроэкономических показателей, при этом инфляция и уровень безработицы имели социально нейтральные значения, учетная ставка НБУ должна быть нулевой, государственные расходы должны возрасти в три раза и составить 183441 млн. грн., импорт, практически должен остаться без изменений по сравнению с концом 2011 г., инвестиции необходимо увеличить в пять с половиной раз и довести до уровня 257 млрд. грн., затраты на научные разработки достаточно увеличить до двух с половиной миллиардов гривень, но внешний долг должен принять отрицательное значение, что в практическом плане означает его полное погашение. Удвоение валового внутреннего продукта, потребления домашних хозяйств и доходов населения, при неизменном уровне инфляции и нулевой безработицы произойдет по достижению расчетных величин учетной ставки НБУ, государственных расходов, импорта, инвестиций, затрат на выполнение научные работ и внешнего долга. А это зависит от практической деятельности соответствующих государственных структур.

Сравнение данных расчета с прогнозными значениями выглядит просто удручающе. Прогноз говорит о том, что экономика будет развиваться в направлении противоположном от того, который вывел бы экономику страны из стагнации. Вместо сокращения величины внешнего долга будет происходить его повышение. Инвестиции должны стремительно расти, причем расти на порядки, но вместо этого они будут сокращаться. Процент роста объемов средств на выполнение научных работ в два раза меньше, чем рост инфляции, что означает фактическое сокращение финансирования научных разработок.

Прогноз, построенный с помощью эксплораторного факторного анализа, показывает, что в экономической системе страны к настоящему моменту времени факторы, обеспечивающие регулирование поступательного развития, не проявились.

Выводы

Динамика социально-экономического фактора указывала на то, что элементы кризисных явлений проявлялись еще в 2006 г. и 2007 г. В конце 2008 г. происходит резкое падение этого фактора, что указывает на начало экономического кризиса. Но если в 2006 г. и 2007 г. инфляция и безработица обусловливались разными факторами, то с конца 2008 г. эти показатели стали определяться одним, что становится характерным моментом для кризисных периодов.

Прогноз значений стохастических факторов указывает на то, что с 2012 г. у социально-экономического фактора наметится тенденция к росту. Такая же тенденция имеет место и у фактора обязательств. Внешнеэкономический фактор, имея ежегодные всплески, приходящиеся на второй квартал каждого года, приобретет тренд на снижение своих величин. Прогнозные значения процентного фактора имеют наименьшие колебания относительно своей средней величины. Это говорит о том, что процентная составляющая в развитии экономической системы страны будет наиболее стабильной. Однако среднее значение этого фактора находится в отрицательной области, что указывает на недостаточную ожидаемую роль регулятора финансовых ринков. В целом, стохастические факторы не будут благоприятствовать развитию экономики Украины в последующие два года.

Прогноз макроэкономических показателей открыл не очень радостную картину. Инвестиции будут сокращаться, а внешний долг продолжит свой рост. Существенным ожидается рост инфляции, что приведет к нивелированию роста номинального валового внутреннего продукта. В прогнозный период сохранится превышение импорта над экспортом, но наметится тенденция к сокращению отрицательного сальдо во внешнеэкономической деятельности.

Построение модели симультанных уравнений и проведенный на ее основе анализ показал, что наиболее приоритетными направления изменения экономической ситуации к лучшему являются стимулирование научных исследований и сокращение внешнего долга.

Для удвоения валового внутреннего продукта, потребления домашних хозяйств и доходов населения, при неизменном уровне инфляции и нулевой безработицы необходимо, чтобы учетная ставка НБУ была нулевой, государственные расходы должны возрасти в три раза, инвестиции необходимо увеличить в пять с половиной раз, и внешний долг должен быть полностью погашен.

ЗАКЛЮЧЕНИЕ

Проведенный анализ современных научных исследований в области методологии предвидения и прогнозирования показал, что исследование и прогнозирование экономических процессов приобретают на нынешнем этапе развития экономики и социальных отношений особую значимость и важность. Этот этап характеризуется стремительным развитием научно-технического прогресса, появлением принципиально новых технологий и продуктов, которые коренным образом меняют не только технологии производства, но и образ жизни людей. Яркой чертой настоящего этапа является глубокая глобализация экономики, которая несет не только преимущества международного разделения труда, но и существенные экономические риски. Кризис, появляющийся в одной стране, распространяется далеко за ее пределы, и становится большой проблемой для стран с неокрепшей экономикой. Все это существенным образом сказывается не только на количественные характеристики динамики экономических систем, но и на их качественный характер.

Современная концепция методологии прогнозирования предполагает рассмотрение комплекса методов, приемов, подходов, направленных на повышение эффективности и конкурентоспособности хозяйствующих субъектов рынка. Однако существующие методы прогнозирования в недостаточной степени соответствуют методической постановке процедуры прогнозирования динамических отклонений и случайных факторов при функционировании экономической системы, которая должна исходить из основных законов в методологии прогнозирования: системности, онтогенеза, композиции и экономической целесообразности.

Анализ наиболее широко используемых в настоящее время методов прогнозирования показал, что создание модели, учитывающей авторегрессионные зависимости и корреляционные связи между переменными, а также разработка методов получения оценок параметров такой модели, отображающей процессы неустоявшейся динамики, являются актуальными.

Для того, чтобы в наиболее полной мере исследовать экономическую систему целесообразно использовать эксплораторный факторный анализ Именно факторная модель позволяет рассмотреть как структуру взаимосвязей между самими наблюдаемыми переменными, так и между отдельными элементами экономической системы.

Динамика экономической системы определяется двумя видами зависимостей. Первая задается закономерностями движения факторов, определяющих поведение экономической системы, а вторая – отображением факторов на множество значений наблюдаемых переменных. Эти зависимости должны стать основой для математической модели экономической системы.

Рассмотрены особенности использования модели эксплораторного факторного анализа для исследования экономических систем, обоснована необходимость получения оценок общностей для подстановки в редуцированную матрицу, с тем, чтобы точнее оценить параметры факторной модели экономической системы.

В настоящее время в качестве оценок общностей используются коэффициенты множественной корреляции. В работе получены оценки общностей стохастических факторов на основе информационного подхода. С точки зрения статистических критериев близости оценки к истинным значениям предложенная оценка является лучше традиционно используемой.

В результате решения оптимизационной задачи поиска максимума функции правдоподобия при наличии ограничений, накладываемых на параметры факторной модели, найдены максимально правдоподобные оценки факторных нагрузок, отвечающие принципу «простой структуры».

Идеальным в эксплораторном факторном анализе выглядит ситуация, когда модель соотношения перемененных и факторов дается в натуральных единицах. Но тогда в каких единицах должен измеряться фактор так, чтобы через него выражались переменные разной размерности? К тому же одна переменная выражается через совокупность факторов. Эта практически не разрешимая проблема снимается с помощью центрирования и нормирования наблюдаемых переменных, а следовательно, значения стохастических факторов должны удовлетворять условию ортонормированности. Оценка значений факторов, основанная на методе множественной регрессии, дает ортогональные величины факторов, но они не являются нормированными. Поставленная задача получить оценки значений факторов, математические ожидания которых были бы равны нулю, а дисперсии – единице для всех моментов времени, была решена и такие оценки значений факторов получены.

Рассмотрение методов анализа многомерных временных рядов показал, что традиционные подходы предполагают влияние на данную переменную в данный момент времени как ее значений в предшествующие моменты, так и значений других переменных, и в текущий момент, и в предшествующие. Для этого вводится понятие кросс-корреляции между различными переменными. У спектральной плотности кросс-спектра для двух рядов наряду с действительной частью появляется и мнимая часть, и в этой связи возникают еще две характеристики: когерентность и «усиление». Слабая интерпретабельность получаемых при таком анализе результатов, или ее полная невозможность, делают анализ многомерных временных рядов малопривлекательным и малораспространенным.

Использование эксплораторного факторного анализа позволяет одновременно рассмотреть и авторегрессионные закономерности, и корреляции между исследуемыми переменными.

Предложена модель, состоящая из двух матричных уравнений: динамики факторов и факторной структуры, − которая позволяет преодолеть основные недостатки традиционных методов анализа многомерных временных рядов.

Получена оценка матрицы переходов уравнения динамики стохастических факторов. Обоснована процедура прогноза компонент многомерного временного ряда.

Рассматривая риск, как экономическую категорию, проведенный анализ измерителей уровня риска показал, что существующие к настоящему моменту времени системы показателей мало приспособлены к тому, чтобы оценить экономический риск в целом с позиции комплексной оценки риска.

Проанализированные методы измерения экономического риска предполагают использование лишь одного экономического показателя в качестве носителя информации о риске и никак не учитывают причины или внешние факторы, влияющие на уровень рискованности. Поэтому актуальным является создание метода, основанного на исследовании информации по как можно наиболее широкому кругу экономических переменных, с целью получения комплексной оценки экономического риска.

Предложено рассматривать m-мерное факторное в качестве пространства риска, и предложена, определенная на этом пространстве, комплексная оценка риска.

Для прогноза уровня рискованности экономической системы разработана автрегрессионная модель произвольного порядка для стохастических факторов, и получены оценки параметров этой модели. Прогноз значений стохастических факторов на основании этой модели позволит определить, в какой зоне риска будет находиться экономическая система в прогнозный момент времени.

Исследование природы симультанности переменных системы эконометрических уравнений показало, что в такой системе взаимосвязанными могут быть не только эндогенные переменные, но и экзогенные.

Все существующие к настоящему моменту времени методы оценивания параметров системы симультанных уравнений этот факт игнорируют, что приводит к смещенности оценок, получаемых с помощью этих методов.

Разработан метод факторизации системы симультанных переменных, который дает несмещенные оценки ее параметров, причем эти оценки являются более эффективными по сравнению с наиболее распространяемыми методами оценивания коэффициентов уравнений системы.

Эксплораторный факторный анализ позволил проанализировать основные тенденции влияния стохастических факторов на экономику страны. Динамика социально-экономического фактора указала на то, что элементы кризисных явлений проявлялись еще в 2006 и 2007 гг. В конце 2008 г. происходит резкое падение этого фактора, что свидетельствует о начале экономического кризиса. Но если в 2006 и 2007 гг. инфляция и безработица обусловливались разными факторами, то с конца 2008 г. эти показатели стали определяться одним, что становится характерным моментом для кризисных периодов.

Прогноз значений стохастических факторов указывает на то, что с 2012г. у социально-экономического фактора наметится тенденция к росту. Такая же тенденция имеет место и у фактора обязательств. Внешнеэкономический фактор, имея ежегодные всплески, приходящиеся на второй квартал каждого года, приобретет тренд на снижение своих величин. Прогнозные значения процентного фактора имеют наименьшие колебания относительно своей средней величины. Это говорит о том, что процентная составляющая в развитии экономической системы страны будет наиболее стабильной. Однако среднее значение этого фактора находится в отрицательной области, что означает недостаточную ожидаемую роль регулятора финансовых ринков. В целом, стохастические факторы не будут благоприятствовать развитию экономики Украины в последующие два года.

Прогноз макроэкономических показателей открыл не очень радостную картину. Инвестиции будут сокращаться, а внешний долг продолжит свой рост. Существенным ожидается рост инфляции, что приведет к нивелированию роста номинального валового внутреннего продукта. В прогнозный период сохранится превышение импорта над экспортом, но наметится тенденция к сокращению отрицательного сальдо во внешнеэкономической деятельности.

Построение модели симультанных уравнений и проведенный на ее основе анализ показал, что наиболее приоритетными направления изменения экономической ситуации к лучшему являются стимулирование научных исследований и сокращение внешнего долга.

Для удвоения валового внутреннего продукта, потребления домашних хозяйств и доходов населения, при неизменном уровне инфляции и нулевой безработицы необходимо, чтобы учетная ставка НБУ была нулевой, государственные расходы должны возрасти в три раза, инвестиции необходимо увеличить в пять с половиной раз, и внешний долг должен быть полностью погашен.

Эксплораторный факторный анализ позволил сконструировать модель многомерного временного ряда, учитывающую как авторегрессионные зависимости, так и взаимные корреляции между экономическими переменными; сформировать новый подход к комплексному измерению уровня рискованности всей экономической системы; обосновать метод факторизации системы симультанных уравнений, дающий несмещенные и более эффективные, по сравнению с наиболее широко используемыми методами, оценки ее параметров.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Энциклопедический словарь / Научно-редакционный совет: А.М. Прохоров (пред.). – М.: «Советская Энциклопедия», 1981. – 1600 с.
2. Економічна енциклопедія: у трьох томах. Т. 3 / Редкол.: С.В. Мочерний (відп. ред.). – К.: ВЦ «Академія», 2002. – 952 с.
3. Кондратьев Н.Д. Проблемы экономической динамики / Н.Д. Кондратьев. – М.: Экономика, 1989. − 526 с.
4. Кондратьев Н. Д. Избранные сочинения / Н.Д. Кондратьев. – М.: Экономика, 1993. – 246 с.
5. Моделі і методи соціально-економічного прогнозування: підруч. / В.М. Геєць, Т.С. Клебанова, О.І. Черняк, [та ін.]. – Х.: ВД «Інжек», 2008. – 396 с.
6. Яковец Ю.В. Прогнозирование циклов и кризисов / Ю.В. Яковец. − М.: МФК, 2000. – 104 с.
7. Боровиков В.П. Прогнозирование в системе Statistica в среде Windows: Основы теории и интенсивная практика на компьютере: [уч. пособ.]/В.П. Боровиков, Г.И. Ивченко. – М.: Финансы и статистика, 2006. – 368 с.
8. Костіна Н.І. Фінансове прогнозування: методи та моделі: [навч. посіб.]/ Н.І. Костіна, А.А. Алексеев, О.В. Василик. – К.: Тов. «Знання», КОО, 1997. – 184 с.
9. Холден К. Економічне прогнозування / К. Холден, Д. Піл, Дж. Томпсон. – К.: Інформтехніка – ЕМЦ, 1996. – 216 с.
10. Сергеева Л.Н. Моделирование поведения экономических систем методами нелинейной динамики (теории хаоса) / Л.Н. Сергеева. – Запорожье: ЗГУ, 2002. – 227 с.
11. Сергеева Л.Н. Нелинейная экономика: модели и методы: [монография]/ Л.Н. Сергеева. – Запорожье: Полиграф, 2003. – 218с.
12. Вязьмин С.А. Применение вейвлет-анализа в анализе и прогнозировании финансовых рынков / С.А. Вязьмин, В.С. Киреев // Научная сессия МИФИ-2004. – Том 13: Экономика и управление. – С. 60-70. Соловьев
13. Мантенья Р. Введение в эконофизику: Корреляции и сложность в финансах / Р. Мантенья, Г. Стенли. – М.: Мир, 2009. – 192 с.
14. Солов'їв В.М. Моделювання критичних явищ на світовому валютному ринку в умовах глобальної кризи / В.М. Солов'їв // Вісник УБС НБУ. – 2009. − №1 (4). – С. 191-196.
15. Лысенко Ю.Г. Нейронные сети и генетические алгоритмы: [уч. пособ.] / Ю.Г. Лысенко, Н.Н. Иванов, Ю.А. Минц. – Донецк: ООО «Юго-Восток Лтд», 2003. – 265 с.
16. Черняк О.І. Анализ та прогноз динаміки ВВП України за допомогою методу SSA / О.І. Черняк, М.Я. Кудіненко // Економіка і прогнозування. – 2002. − № 4. – С. 134-147.
17. Коротаев А.В., Цирель С.В. Кондратьевские волны в мировой экономической динамике / А.В. Коротаев, С.В. Цирель; Ред. Д.А. Халтурина, А.В. Коротаев // Системный мониторинг. Глобальное и региональное развитие. – М.: Либроком/URSS, 2009. ISBN 978-5-397-00917-1. – С. 189−229.
18. Айвазян С.А. Прикладная статистика. Исследование зависимостей / С.А. Айвазян, И.С. Енюков, Л.Д. Мешалкин. – М.: Финансы и статистика, 1985. – 488 с.
19. Общая теория статистики. Статистическая методология в изучении коммерческой деятельности / под ред. А.А. Спирина, О.Э. Башиной. – М.: Финансы и статистика, 1994. – 296 с.
20. Хохлов В.В. Прогнозирование ресурсов банковских учреждений. / В.В.Хохлов, Е.Л.Гринько. − Вестник НБУ. − № 8(148), 2008.− С.112-118.
21. Хеннан Э. Многомерные временные ряды / Э. Хеннан. – М.: Мир, 1974. – 576 с.
22. Жихарев В.Н. Новый эконометрический метод "ЖОК" оценки результатов взаимовлияний факторов в инженерном менеджменте / В.Н. Жихарев, В.Г. Кольцов, А.И. Орлов; под общей редакцией канд. экон. наук. В.А. Панкова //

Проблемы технологии, управления и экономики: сб. – Ч.1. Краматорск: Донбасская государственная машиностроительная академия, 1999. − С.87-89.
23. Математика и кибернетика в экономике: Словарь-справочник / М.: Экономика, 1971.–223 с.
24. Spearman C. General intelligence objectively determined and measured / C. Spearman.– Am. J. Psychoter, 1904. № 15. – p.201-293.
25. Cattell R. Handbook of multivariate experimental psychology /R. Cattell.– Chicago, 1966.–428p.
26. Kendall M. Factor analysis. Part 1. Factor analysis as a statistical technique/ M. Kendall // J. Roy. Statist. Soc., 1950.– № 12.–p. 60-73.
27. Pearson K. On lines and planes of closest fit to systems of points in space / K. Pearson.– Philosophical Magazine., 1901.– № 6.– p. 559-572.
28. Hotelling H. Analysis of a complex of statistical variables into principal components/ H. Hotelling.–Journal of Social Psychology, 1933.–№ 24.– p. 417-441.
29. Moulton F. The velocity of light / F. Moulton. – Scientific Monthly, 1939.– № 48.– p.481-484.
30. Hotelling H. Analysis of a complex of statistical variables into principal components/ H. Hotelling.–Journal of Social Psychology, 1933.–№ 24.– p. 417-441.
31. Anderson T. Statistical inference in factor analysis /T. Anderson, H. Rubin.– Proc. Of the Third Symp. On Math. Statistics and Probability, 1956.–№ 5.– p. 111-150.
32. Thurstone L. Multiple factor analysis / L. Thurstone. – Psychological Review, 1931.– № 38.– p. 406-427.
33. Tryon R. Cluster analysis / R. Tryon.– Ann. Arbor: Edward Bros., 1939.– P. 122.
34. Хедли Дж. Линейная алгебра / Дж. Хедли.– М.: Высш. шк., 1966.– 290 с.
35. Юл Дж. Теория статистики/Дж. Юл, М. Кендэлл.–М.: Госстатиздат, 1960.– 348 с.
36. Лоули Д. Факторный анализ как статистический метод / Д. Лоули, А. Максвелл.– М.: Мир, 1964.– 224 с.
37. Хохлов В.В. Задачи и методы факторного анализа при моделировании сложных систем/ В.В. Хохлов // Рукопись.– Деп. в УкрНИИНТИ 25.02.93.– Севастополь, 1993.– № 273-Ук93.– 24 с.
38. Dwyer P. The contribution of an orthogonal multiple factor solution to multiple correlation / P. Dwyer.–Psychometrika, 1939.–№ 4.– P. 163-171.
39. Guilford J. Psychometric methods / J. Guilford. − N.Y., 1954.− P.216.
40. Spearman C. Theory of general factor / C. Spearman.− British Journal of Psychology, 1946.− № 36.− P.117-131.
41. Иберла К. Факторный анализ / К. Иберла. − М.: Статистика, 1980.−398с.
42. Kaiser H. A note on Tryon-Kaiser solution for the communalities / H. Kaiser.−Psychometrika, 1989.−№ 24.−P. 269-271.
43. Kaiser H. Solution for the communalities / H. Kaiser.− Research report, № 5,1996, Berkley, U. of California.− 11p.
44. Лапа В.Г. Математические основы кибернетики / В.Г. Лапа.− К.: Вища школа, 1974.−458 с.
45. Харман Г. Современный факторный анализ / Г. Харман. – М.: Статистика, 1972.− 486 с.
46. Бикел П. Математическая статистика/ П. Бикел, К. Доксам. − М.: Финансы и статистика, 1983.– Вып. 2.− 254с.
47. Айвазян С.А. Прикладная статистика: Основы моделирования и первичная обработка данных / С.А. Айвазян, И.С. Енюков, Л.Д. Мешалкин. – М.: Финансы и статистика, 1983.− 471 с.
48. Heermann E.F. Univocal or orthogonal estimators of orthogonal factors / E.F. Heermann. – Psychometrics, 1963, № 28.– P.161-172.
49. Ansary A. Heterogeneous factor analysis model: a Bayesian approach /A. Ansary, K. Jedidi, L. Dube.− Psychometrika, 2002.−Vol. 67, № 1.− P. 49-78.
50. Кендел. М. Временные ряды / М. Кендел. − М.: Финансы и статистика, 1981.− 199 с.

51. Андерсон Т. Статистический анализ временных рядов /Т. Андерсон. – М.: Мир, 1976. – 756 с.
52. Бокс Дж. Анализ временных рядов: прогноз и управление /Дж. Бокс, Г. Дженкинс.−М.: Мир, 1974. −Вып.1.−408 с.
53. Brown R. Smoothing, forecasting and prediction of discrete time-series.−New Jersey: Prentice-Hall, 1982.−246 p.
54. Колмогоров А.Н. Интерполяция и экстраполяция стационарных случайных последовательностей / А.Н. Колмогоров // Известия АН СССР. Сер. Математика.– 1941. Т. 5. – № 3.
55. Бриллинджер Д. Временные ряды: обработка данных и теория / Д. Бриллинджер. – М.: Мир, 1980. – 536 с.
56. Хеннан Э. Многомерные временные ряды /Э. Хеннан. – М.: Мир, 1974. – 575 с.
57. Kalman R.E. New results in linear filtering and prediction theory / R.E. Kalman. – Journal of Basic Engineering, 1961. – 83D.
58. Sorenson H.W. Kalman filtering techniques / H.W. Sorenson.− Advances in Control Systems Theory and Applications, 1966.− vol. 3.− pp. 219-292.
59. Широков Ю.М. Квантовая и классическая механика в представлении фазового пространства / Ю.М. Широков. – ЭЧАЯ, 1979. − №10 (1). – С. 5-50.
60. Королюк В.С. Процессы марковского восстановления в задачах надежности систем / В.С. Королюк, А.Ф. Турбин. – К.: Наук. думка, 1982. – 236с.
61. Кузнецов Д., Ройлих Д. Квантовый шум при отображении фазового пространства / Д. Кузнецов, Д. Ройлих. – Оптика и Спектроскопия, 1997. – № 82 (1).
62. К. Холден. Економічне прогнозування: вступ / К. Холден, Д.А. Піл, Дж.Л. Томпсон. – К.: Інформтехніка – ЕМЦ, 1996 – 216 с.
63. Энциклопедия финансового риск-менеджмента / Под ред. А.А. Лобанова, А.В. Чугунова. – М.: Альпина Паблишер, 2003. – 786 с.
64. Downward P. Risk, uncertainty and inference in post-Keynesian economy / P. Downward. – Paper presented at the INEM-ROPE conference, Univ. of New Hampshire, June 15-17, 1998.
65. Рогов М.А. Риск-менеджмент / М.А. Рогов. – М.: Финансы и статистика, 2001. – 248 с.
66. Балабанов И.Т. Валютный рынок и валютные операции / И.Т. Балабанов. – М.: Финансы и статистика, 1994. – 182 с.
67. Gujarati D. Basic Econometrics / D. Gujarati. − NY: McGraw−Hill, 2004. −1002 p.
68. Longman Dictionary of the English Language. − London: Longman House, 1984. – 1876 p.
69. Википедия – свободная библиотека. – Статья «Симультанный» [Электронный ресурс]. – Режим доступа: http//ru.wikipedia.org/wiki.
70. Davidson J. Econometric Theory / J. Davidson. − Blackwell Publishers: Oxford, U.K. 2000. – 468p.
71. Айвазян С.А. Прикладная статистика: Исследование зависимостей / С.А. Айвазян, И.С. Енюков, Л.Д. Мешалкин. – М.: Финансы и статистика, 1985. – 487 с.
72. Джонстон Д. Эконометрические методы / Д. Джонстон – М.: Статистика, 1980. – 446 с.
73. Маленво Э. Статистические методы в эконометрии / Э. Маленво – М.: Статистика, 1975, вып. 2. – 325 с.
74. Rao C.R. Least squares theory using an estimated dispersion matrix and its application to measurement of signals / C.R. Rao. – Berkeley: The fifth Berkeley symposium, 1968.
75. Макроекономічні показники. Статистичні данні УДКС [Електронний ресурс]. − Режим доступу: http://www.ukrstat.gov.ua.
76. Фінансові ринки. Статистичні данні НБУ [Електронний ресурс]. − Режим доступу: http://www.bank.gov.ua/Statist/sfs.htm.
77. Хохлов В.В. Прогнозирование финансового состояния предприятия на основе многомерного факторного анализа временных рядов / В.В. Хохлов, Е.И. Пискун.− Бизнес Информ.− № 2(1), 2009.− С. 82-87.

78. Хохлов В.В. Оценка развития предприятий, входящих в крупномасштабную экономическую систему / В.В. Хохлов, Е.И. Пискун. − Бизнес Информ.− № 4(2), 2010.− С. 78-81.
79. Хохлов В.В. Факторная оценка функционирования структурного подразделения сложной производственной системы / В.В.Хохлов, Е.И.Пискун. − Бизнес Информ.− № 5(2), 2010.− С. 36-40.
80. Хохлов В.В. Анализ и прогнозирование результативности деятельности крупномасштабных экономических систем и экономики страны с помощью многомерной факторной авторегрессионной модели / В.В. Хохлов, Е.И. Пискун. − Культура народов Причерноморья. Экономические науки. – 2010. − № 178. – С. 190-193.
81. Хохлов В.В. Факторно-кореляційний аналіз депозитів домашніх господарств. − Вісник НБУ. − №7 (185), 2011. – С.33-37.
82. Хохлов В.В. Оценка влияния инновационной деятельности на экономику Украины методом факторизации симультанных переменных / В.В. Хохлов, Е.И. Пискун. − Бизнес Информ.− № 5(1), 2011.− С. 62-65.

Монография
Хохлов Владимир Владимирович

ИССЛЕДОВАНИЕ И ПРОГНОЗИРОВАНИЕ ЭКОНОМИЧЕСКИХ ПРОЦЕССОВ С ИСПОЛЬЗОВАНИЕМ ЭКСПЛОРАТОРНОГО ФАКТОРНОГО АНАЛИЗА МНОГОМЕРНЫХ ВРЕМЕННЫХ РЯДОВ

Монографія
Хохлов Володимир Володимирович

ДОСЛІДЖЕННЯ І ПРОГНОЗУВАННЯ ЕКОНОМІЧНИХ ПРОЦЕСІВ ІЗ ВИКОРИСТАННЯМ ЕКСПЛОРАТОРНОГО ФАКТОРНОГО АНАЛІЗУ БАГАТОВИМІРНИХ ЧАСОВИХ РЯДІВ

(російською мовою)

Коректор Л.П.Светлих
Нормоконтролер І.О.Черевкова
Комп'ютерне складання та верстання В.В.Хохлов

Підп. до друку 25.04.2012 р. Формат 89 x 124/16. Ум. друк. арк. 20,4. Тираж 300 пр. Зам. № _
Видавець та виготовлювач — Севастопольський національний технічний університет
Адреса: вул. Університетська, 33, м. Севастополь, 99053
тел. (0692) 435-210; 435-019. E-mail: root@sevgtu.sebastopol.ua
Свідоцтво суб’єкта видавничої справи ДК № 1272 від 17.03.2003 р.