ЛЕКЦИОННЫЕ И МЕТОДИЧЕСКИЕ МАТЕРИАЛЫ

Анализ временных рядов[1)]

Канторович Г.Г.

Продолжается публикация курса «Анализ временных рядов». В этом номере рассматривается моделирование сезонности в терминах моделей ARIMA, построение авторегрессионных моделей с распределенными лагами (ADL), модели корреляции ошибками (ECM), различные типы экзогенности переменных, начато рассмотрение многомерных временных рядов.

В следующем выпуске будет продолжено рассмотрение моделей многомерных случайных процессов, коинтеграционные регрессии, тестирование коинтеграции.

Лекция 11

Сезонность

Статистические данные динамики экономических величин часто демонстрируют регулярные или почти регулярные колебания. Это явление носит название сезонности. Так, если рассматривать данные об объеме промышленного производства, то уровень производства в январе имеет большую связь с производством в прошлогоднем январе, чем с производством в предыдущем месяце – декабре. Обычно в России в январе наблюдается сильное падение производства по сравнению с декабрем. В развитых странах в летние месяцы происходит снижение безработицы в связи с сезонными работами. Данные о розничных продажах обычно имеют заметный «всплеск» в декабре, что носит название рождественского эффекта.

Из предыдущих лекций видно, что модели типа ARIMA плохо описывают такое сезонное циклическое поведение данных. Другими словами, неявно предполагается, что сезонность предварительно устранена. В традиционной экономической практике такое «десезонирование» данных получило широкое применение. Одним из распространенных методов является хорошо вам известное из курса эконометрики применение dummy-переменных. Например, для квартальных данных строится регрессия на 4 dummy-переменные, каждая из которых равна 1 в соот-

[1)] Подготовлено при содействии Национального фонда подготовки кадров (НФПК) в рамках программы «Совершенствование преподавания социально-экономических дисциплин в вузах».

Канторович Г.Г. – профессор, к. физ.-мат. н., проректор ГУ–ВШЭ, зав. кафедрой математической экономики и эконометрики.

ветствующем квартале и 0 – в остальных (не забудем, что свободный член в этом случае в регрессию не включается). Остатки этой регрессии и являются данными с устраненной сезонностью. Этот метод предполагает «жесткую» сезонность с неизменной величиной влияния каждого квартала в отдельности. Для месячных данных регрессия строится на 12 dummy-переменных, для полугодовых – на 2 dummy-переменные. Иногда рассматривают временные отрезки длиной в четыре недели, тогда в течение года насчитывается 13 промежутков и, соответственно, 13 dummy-переменных. В дальнейшем будем называть такой подход детерминированной сезонностью.

Предположение о «жесткой» сезонности является слишком ограничительным, поэтому официальная статистика применяет более сложные методы устранения сезонности. Широкое распространение получила программа X-11, разработанная в Американском бюро цензов (US Census Bureau), и ее дальнейшие развития. Разработаны и применяются на практике и другие методы. На рис. 11.1 приведены официальные ежемесячные данные Госкомстата РФ о динамике индекса промышленного производства в России в период с января 1990 г. по апрель 1999 г. до и после устранения сезонности [1]. На графике хорошо видно, что сезонность в этом случае наложена на тренд, т.е. данные демонстрируют одновременно и циклическое поведение, и тенденцию к снижению уровня.

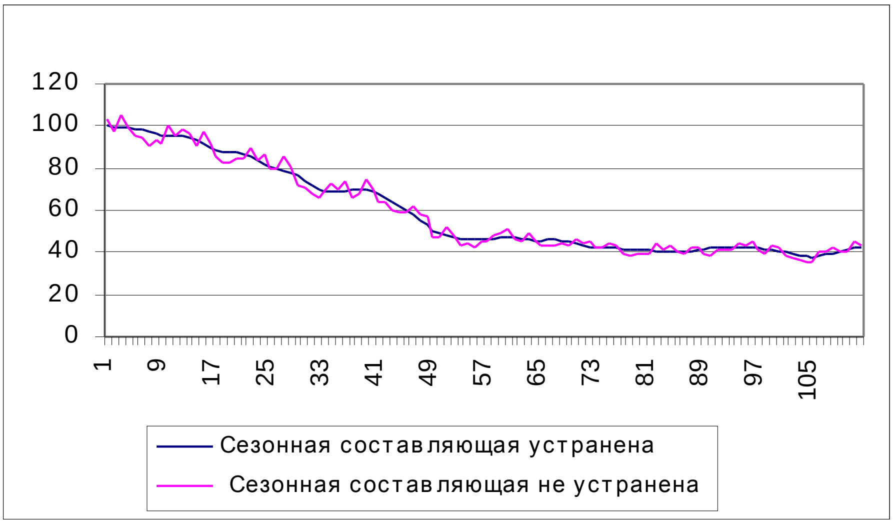

Рис. 11.1. Динамика индекса промышленного производства в России с января 1990 г. по апрель 1999 г., %

Рассмотренные до сих пор модели временной структуры ряда не предназначались для описания именно такого поведения. Мы строили авторегрессионные модели, когда текущее значение ряда было связано с несколькими предыдущими значениями. В принципе, циклическое поведение можно попытаться описать в терминах моделей ARMA, наложив некоторые ограничения на коэффициенты модели. Например, если исследуемые данные являются квартальными, сезонность означает, что данные I квартала текущего года сильнее связаны с данными I квартала предыдущего года, чем с данными предыдущего квартала. Разумно рассматривать модель, в которой явно присутствует связь только с данными одноименного квартала предыдущего года, т.е. модель вида: $x_t = \varphi x_{t-4} + \varepsilon_t$. А для месячных данных разумной будет модель $x_t = \Psi x_{t-12} + \varepsilon_t$. Используя оператор сдвига, получаем для квартальных данных: $(1-\varphi L^4)x_t = \varepsilon_t$, а для месячных данных: $(1-\varphi L^{12})x_t = \varepsilon_t$. Видно, что сезонность – это специальный случай ARMA. По нашей терминологии это ARMA(4,0) или AR(4). Но у него промежуточные коэффициенты равны нулю.

Что будет, если в модели с квартальными данными $\varphi = 1$? Модель принимает вид $x_t = x_{t-4} + \varepsilon_t$. Очевидно, характеристическое уравнение имеет единичный корень. Модель приводится к виду $(1-L^4)x_t = \varepsilon_t$. Я думаю, вы помните, что $(1-L^4)$ раскладывается на 4 сомножителя, и L=1 является корнем. Три остальных корня характеристического уравнения равны соответственно −1 и $\pm i$, где $i = \sqrt{-1}$. Все четыре корня по модулю равны 1. Говорят, что уравнение содержит единичный корень в обычном смысле и, кроме того, три сезонных единичных корня [11]. Аналогично для месячных данных: если $\Psi = 1$, то модель принимает вид $(1-L^{12})x_t = \varepsilon_t$. Разумеется, здесь тоже есть единичный корень в обычном смысле, т.е. и этот ряд относится к типу DS.

Модель, в которой имеется такой четко выраженный сомножитель $(1-\varphi L^4)$ или $(1-\varphi L^{12})$ принято называть *сезонной авторегрессией* (SAR). Аналогично можно рассмотреть сезонную схему для скользящего среднего, т.е. рассмотреть модель типа SMA, которая означает, что в правую часть будет добавлено слагаемое $\beta\varepsilon_{t-4}$. Эта модель представляет собой специальный вид обычного MA(4) или MA(12) с нулевыми промежуточными коэффициентами. Общая модель, описывающая сезонность в подходе Бокса–Дженкинса, носит название SARIMA.

Разработана примерно такая же теория и практика нахождения единичных сезонных корней, как это сделано для обычных единичных корней. Если TS и DS процессы мы называли также процессами с детерминированным и стохастическим трендом соответственно, то для описания сезонности употребляется аналогичная терминология. Если $\varphi = 1$, то присутствует *стохастическая сезонность*. Если $|\varphi| < 1$, то присутствует *детерминированная сезонность*.

Разумеется, в одной модели можно без большого труда сочетать и сезонную составляющую, и обычную ARMA схему. Например, рассмотрим модель типа SARIMA: $(1-L^4)B(L)x_t = \varepsilon_t$, где $B(L) = (1-\beta L)$ (для упрощения записи в модель не включена MA часть). После раскрытия скобок получаем: $(1-L^4-\beta L+\beta L^5)x_t = \varepsilon_t$ или $x_t = \beta x_{t-1} + x_{t-4} - \beta x_{t-5} + \varepsilon_t$. То есть исходная модель приводится к AR(5), но с некоторыми дополнительными ограничениями на коэффициенты. Они выражаются в том, что второй и третий коэффициенты равны нулю, и в том, что коэффициенты первого и пятого слагаемых в сумме равны нулю, т.е. имеют противоположные знаки. Другими словами, это некий частный случай общего AR(5). При оценивании модели для нас важно, что мы уже знаем эти соотношения между коэффициентами и будем оценивать именно такую модель, а не общую, в которой у нас будут лишние переменные и понизится точность оценки наших коэффициентов. Эту модель принято обозначать SAR(4)×AR(1).

Econometric Views позволяет легко строить такие модели, вы просто указываете, что вам надо добавить в модель SAR или SMA соответствующего порядка. Следует отметить, что при задании порядка сезонности нужно учитывать, как описана структура данных в базе EViews. Если данные описаны как нерегулярные, то порядок сезонности задается числом 4, а если вы их уже описали как квартальные, то можно задать SAR(1).

Получается, что если в данных присутствует сезонность, нельзя обойтись моделью ARMA низкого порядка. Допустим, у вас есть ежемесячные данные за 4 или 5 лет. Наличие сезонности означает, что имеется значимая статистическая зависимость от соответствующего месяца предыдущего года. Вы можете сразу построить модель порядка, скажем, 15, т.е. AR(15), но AR(15) приведет к утрате 15 первых точек для построения модели. Коэффициентов определяется очень много, опасность мультиколлинеарности возрастает. При явном использовании схемы SAR возникает другая ситуация, в модели остается только 4 коэффициента. Конечно, некоторое число наблюдений все равно будет потеряно. Но число коэффициентов гораздо меньше, число степеней свободы существенно больше, значит, точность их оценивания возрастает. Кроме того, уменьшается опасность мультиколлинеарности.

Если учет сезонности в терминах подхода Бокса−Дженкинса находит широкое применение на практике, то исследование единичных сезонных корней не столь популярно. Среди работ по этой тематике, кроме уже отмеченной статьи Хиллеберга и других, стоит отметить монографию Франсеса [5].

Авторегрессионные модели с распределенными лагами

До сих пор мы рассматривали только один случайный процесс X_t. Мы исследовали его на стационарность, строили модели процесса, прогнозировали будущие значения процесса. Однако при изучении экономических явлений наибольший интерес представляет взаимозависимость экономических величин. Текущее значение экономической величины будет зависеть не только от ее предыдущих значений, но и от текущего и предыдущих значений других экономических величин. Другими словами, среди регрессоров будут лаговые значения как объясняемой, так и объясняющих величин. Такие модели принято называть авторегрессионными моделями с распределенными лагами, по-английски − ADL (auturegressive distributed lag) models. Другое название таких моделей − ARMAX напоминает об их сходстве с моделями ARIMA. Регрессор X_t как бы заменяет собой белый шум в модели ARIMA. Для случая только одной объясняющей экономической величины общий вид моделей рассматриваемого типа:

$$Y_t = \theta + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + ... + \alpha_p Y_{t-p} + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + ... + \beta_q X_{t-q} + \varepsilon_t, t = 1,2,..T.$$

Случайное возмущение по-прежнему полагается белым шумом. Используя операторные полиномы, получим $\alpha_p(L)Y_t = \theta + \beta_q(L)X_t + \varepsilon_t$. В отличие от модели

ARIMA мы отказываемся от нормирующего условия $\beta_0 \equiv 1$. Мы будем использовать обозначения ADL(p,q) и ARMAX(p,q), чтобы указать количество лагов независимой и зависимой переменных.

Если процесс X_t является стационарным в широком смысле, то, используя его разложение по теореме Вольда, мы получаем представление процесса Y_t в виде модели ARMA(p,∞), и стационарность процесса Y_t полностью определяется авторегрессионной частью исходной модели. Для нестационарного регрессора X_t процесс Y_t является, вообще говоря, также нестационарным. Отдельного рассмотрения заслуживает только случай, когда характеристические уравнения обоих полиномов $\alpha_p(L)$ и $\beta_q(L)$ имеют одинаковые корни, по модулю большие единицы.

Начнем рассмотрение с ситуации, когда регрессор стационарен, и полином $\alpha_p(L)$ не имеет корней вне или на единичной окружности. В этом случае можно переписать модель в виде: $Y_t = \alpha_p(L)^{-1}\theta + \alpha_p(L)^{-1}\beta_q(L)X_t + \alpha_p(L)^{-1}\varepsilon_t$, т.е. стационарная ADL модель представима в виде модели с распределенными лагами без авторегрессионной части. Это означает, что текущее значение величины Y_t зависит от текущего и всех предшествовавших значений величины X_t или что текущее значение X_t влияет на текущее и все последующие значения величины Y_t. Другими словами, текущее значение величины Y_t представляется в виде суммы динамических откликов на изменения величины X_t. Коэффициент влияния значения X_t на величину Y_{t+k} представляет собой частную производную $\frac{\partial Y_{t+k}}{\partial X}$, которую можно выразить как значение производной от операторного выражения:

$$\frac{d^k}{dL^k}\left(\frac{\beta_q(L)}{\alpha_p(L)}\right) \text{ при } L=0.$$

При увеличении параметра k получаем последовательно мгновенный, краткосрочный, среднесрочный и долгосрочный отклики. Для макроэкономических и монетарных моделей, выраженных в уровнях величин, эти отклики принято называть соответствующими мультипликаторами, а для моделей, выраженных в логарифмах величин, – эластичностями. Например, для модели ADL(1,1) отклики выражаются через коэффициенты модели следующим образом:

$$\beta_0, \beta_1 + \alpha_1\beta_0, \alpha_1\beta_1 + \alpha_1{}^2\beta_0 \text{ и т.д.}$$

Долгосрочный отклик определяется как сумма всех промежуточных откликов. Используя стационарность процессов X_t и Y_t, долгосрочный отклик можно найти, просто взяв математические ожидания от обеих частей соотношения. В результате получаем уравнение статического или долгосрочного равновесия.

$$\overline{Y} = \frac{\theta}{\alpha_p(1)} + \frac{\beta_q(1)}{\alpha_p(1)}\overline{X},$$

где через $\overline{Y}$ и $\overline{X}$ обозначены равновесные значения соответствующих величин.

Для удобства анализа долгосрочного и краткосрочного поведения динамического соотношения, выраженного ADL моделью, удобно провести перепараметризацию модели. Начнем с модели ADL(1,1): $Y_t = \theta + \alpha_1 Y_{t-1} + \beta_0 X_t + \beta_1 X_{t-1} + \varepsilon_t$. Заменив Y_t на $Y_{t-1} + \Delta Y_t$, и X_t на $X_{t-1} + \Delta X_t$, получаем:

$$\Delta Y_t = \theta + \beta_0 \Delta X_t - (1-\alpha_1)Y_{t-1} + (\beta_0 + \beta_1)X_{t-1} + \varepsilon_t.$$

Перегруппировка членов дает:

$$\Delta Y_t = \beta_0 \Delta X_t - (1-\alpha_1)\left[Y_{t-1} - \frac{\theta}{1-\alpha_1} - \frac{\beta_0 + \beta_1}{1-\alpha_1}X_{t-1}\right] + \varepsilon_t.$$

Это представление ADL модели называется *моделью коррекции ошибками* (error correction model), сокращенно – ECM[2]. Смысл модели и названия становится ясен, если обратить внимание, что выражение в квадратных скобках может трактоваться как отклонение от долгосрочного равновесия в момент времени t-1. В самом деле, долгосрочное равновесие определяется соотношением

$$\left[\overline{Y} - \frac{\theta}{1-\alpha_1} - \frac{\beta_0 + \beta_1}{1-\alpha_1}\overline{X}\right] = 0,$$

поэтому выражение в квадратных скобках положительно, если значение Y_{t-1} превышает равновесное значение, соответствующее X_{t-1}. Таким образом, текущее (краткосрочное) изменение Y представлено в виде суммы двух слагаемых. Первое из них – это мгновенный отклик на текущее (краткосрочное) изменение X , а второе – поправка на имевшее место в предыдущий момент отклонение от долгосрочного равновесия. При этом, поскольку для стационарности процесса Y_t необходимо выполнение условия $|\alpha_1| < 1$, коэффициент при невязке отрицательный. Это означает, что второе слагаемое «подтягивает» процесс Y_t к долгосрочному соотношению с процессом X_t. Таким образом, модель коррекции ошибками позволяет удобно объединить в рамках одной модели краткосрочную и долгосрочную динамику, а ее коэффициенты имеют содержательную экономическую интерпретацию.

В общем случае модели ADL(p,q) мы должны заменить значения процессов Y_t и X_t в различные моменты времени на их значения в момент t−1 и отклоне-

---

[2] Впервые введена в [14].

ния, например $Y_{t-2} = Y_{t-1} - \Delta Y_t$, $Y_{t-3} = Y_{t-1} - \Delta Y_t - \Delta Y_{t-2}$. В результате получаем ECM представление:

$$\Delta Y_t = \beta_0 \Delta X_t + \sum_{i=1}^{p-1} \delta_i \Delta Y_{t-i} + \sum_{i=1}^{q-1} \gamma_i \Delta X_{t-i} - \alpha_p(1)\left[Y_{t-1} - \frac{\theta}{\alpha_p(1)} - \frac{\beta_q(1)}{\alpha_p(1)} X_{t-1} \right] + \varepsilon_t .$$

Выражение в квадратных скобках по-прежнему представляет корректирующий член, «подправляющий» лаговую структуру отклонениями от долгосрочного равновесия на предыдущем шаге. Это хорошо видно после замены нулями всех отклонений, что соответствует равновесному состоянию. Важно также отметить, что коэффициенты δ_i и γ_i, также как и остальные коэффициенты ECM представления, линейно выражаются через коэффициенты исходной ADL модели, причем это преобразование невырождено.

Наиболее общая ADL модель получается в случае, когда имеется k различных экономических дисциплин в качестве регрессоров. Будем обозначать ее ADL($p,q_1,q_2,...,q_k$), а общее уравнение принимает вид:

$$\alpha_p(L)Y_t = \theta + \beta_{q_1}(L)X_{1t} + \beta_{q_2}(L)X_{2t} + ... + \beta_{q_k}(L)X_{kt} + \varepsilon_t .$$

Преобразование к модели коррекции ошибками проводится в этом случае точно также, как и ранее, но выражение в общем случае становится весьма громоздким.

Оценивание и диагностика ADL моделей близки к моделям ARIMA. Поскольку в модели отсутствует лаговая структура случайного возмущения, основным методом оценивания является метод наименьших квадратов. Мы знаем, что при отсутствии корреляции случайных регрессоров и случайного возмущения МНК дает состоятельные оценки. Разумеется, перед оцениванием модели нужно убедиться, что все переменные являются стационарными.

У нас есть две возможности. Первая: оценить ADL модель, а затем пересчитать параметры и их стандартные ошибки для ECM представления. Вторая: оценить параметры ECM модели непосредственно. Поскольку, как уже упоминалось ранее, коэффициенты ADL и ECM моделей связаны невырожденным линейным преобразованием, можно доказать, что оба пути дают идентичные результаты. Поэтому, выбор представления модели определяется содержательной задачей исследования, а не особенностями процедуры оценивания.

Определение числа лагов для переменных, входящих в модель, неизбежно сопровождается построением ряда моделей и выбором наилучшей из них. Общепринятой стратегией выбора модели в настоящее время является подход от общего к частному (from general to simple), предложенный и развитый Дэвидом Хендри и его соавторами [8]. В соответствии с этим подходом мы начинаем построение с наиболее общей модели как по набору объясняющих переменных, так и по количеству включенных лагов. Построенная методом наименьших квадратов модель подвергается различным тестам: на автокорреляцию остатков, нормальность, гетероскедастичность и т.д. Если модель проходит все диагностические тесты, на следующем шаге исследуется возможность наложения различных ограничений на ее коэффициенты. Такими ограничениями могут быть исключение из модели отдельных переменных или лагов, равенство некоторых коэффициентов между собой и т.п. Обычно проверяемые ограничения порождаются как содержательными теоретическими соображениями, так и численными значениями оценок коэффициентов.

Лекция 12

При оценивании ADL моделей существует еще одна особенность, которую следует иметь в виду. Впрочем, она присутствует при оценивании любой модели со стохастическими регрессорами, но в курсе эконометрики мы оставляли ее за рамками рассмотрения. Рассматривая несколько случайных величин (или процессов) одновременно, мы всегда неявно предполагаем существование их совместного распределения. Для упрощения выкладок и, не теряя общности, ограничимся двумя процессами: Y_t и X_t. Пусть, например, мы строим ADL модель процесса Y_t от одной объясняющей переменной X_t. Совместная плотность распределения $f(Y_t, X_t)$, если она существует, может быть представлена в виде произведения $f(Y_t, X_t) = f(Y_t|X_t)f(X_t)$ условной плотности Y_t при условии X_t и одномерной плотности процесса X_t. Такое разложение называется факторизацией совместного распределения.

В то же время, теоретическая регрессия представляет собой условное математическое ожидание $E(Y_t|X_t) = \int Y_t f(Y_t|X_t)dY_t$, для вычисления которого используется только условное распределение. Фактически для вычисления регрессии не используется часть информации, содержащаяся в совместном распределении. Вопрос о том, в какой степени такой неполный учет информации влияет на оценивание регрессии, представляет очень интересный и важный аспект современного эконометрического исследования.

Рассмотрим два случайных процесса y_t и x_t, определяемые следующей моделью, в которой для упрощения записи математические ожидания положены равными нулю:

$$\begin{aligned} y_t &= \beta x_t + \varepsilon_{1t} \\ x_t &= \alpha_1 x_{t-1} + \alpha_2 y_{t-1} + \varepsilon_{2t}. \end{aligned}$$

Пусть возмущения образуют двумерный нормальный случайный процесс $\vec{\varepsilon}$, каждая компонента которого не имеет автокорреляций (т.е. является белым шумом), но корреляция между одновременными значениями возмущений, вообще говоря, может быть ненулевой. Это означает, что

$$E(\vec{\varepsilon}) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \operatorname{cov}(\vec{\varepsilon}) = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix}, \quad \rho = \frac{\sigma_{12}}{\sqrt{\sigma_{11}\sigma_{22}}}.$$

Двумерное распределение x_t и y_t полностью определяется свойствами вектора $\vec{\varepsilon}$, но у нас нет нужды выписывать распределение в явном виде. Заметим, что процессы x_t и y_t очевидно автокоррелированы. Первое из соотношений является уравнением регрессии, и вопрос состоит в том, в какой степени можно оценивать эту зависимость y_t от x_t, не принимая во внимание второе из соотношений. В ранее применявшихся терминах [10] вопрос звучал так: является ли x_t экзогенным (exogenous) для уравнения регрессии? В соответствии же с современными представлениями [4] этот вопрос недостаточно конкретен: следует заранее определить способ использования модели. В эконометрике мы уже встречались с ситуацией, когда выбор, например, наилучшей модели зависел от того, как мы собираемся ее использовать. Мы различали, в основном, два типа использования: а) для целей прогнозирования и б) для изучения механизма модели, когда нас интересовали численные оценки всех или нескольких параметров. В соответствии с [4] будем различать следующие три цели анализа моделей:

- сделать статистические выводы об одном или нескольких параметрах модели (изучить механизм модели);
- предсказать y_t при заданном x_t, (условный прогноз);
- проверить, является ли уравнение регрессии структурно инвариантным при изменениях одномерного распределения x_t.

Слабая экзогенность (weak exogeneity)

Пусть мы рассматриваем регрессию некой величины y_t на несколько регрессоров. Обозначим через $\vec{\varphi}$ вектор параметров совместной плотности величины y_t и регрессоров, через $\vec{\theta}_1$ – вектор параметров, от которых зависит условная плотность вероятности y_t при заданных регрессорах, а через $\vec{\theta}_2$ – вектор параметров совместной плотности вероятности регрессоров.

Наконец, пусть параметры, относительно которых мы хотим сделать статистические выводы, образуют множество, записанное в виде вектора $\vec{\alpha}$. Мы скажем, что набор регрессоров является слабо экзогенным для параметров $\vec{\alpha}$, если статистические выводы об этих параметрах, сделанные на основании условного распределения, эквивалентны выводам по совместному распределению y_t и регрессоров.

Очевидно, что необходимым условием независимости статистических выводов от информации, содержащейся в условном распределении регрессоров, является возможность выразить $\vec{\alpha}$ в виде функции параметров $\vec{\theta}_1$, т.е. $\vec{\alpha} = \vec{\alpha}(\vec{\theta}_1)$. Кроме того, на компоненты векторов параметров $\vec{\theta}_1$ и $\vec{\theta}_2$ не должны быть наложены совместные ограничения ни в виде равенств, ни в виде неравенств. Это означает, что область допустимых значений параметров $\vec{\theta}_1$ не зависит от того, какие значения принимают параметры из множества $\vec{\theta}_2$, и наоборот. Это свойство по-английски выражается термином variation-free. По-русски можно использовать предложенный Э. Б. Ершовым термин «свободно-варьируемые».

Проиллюстрируем свойство слабой экзогенности на примере вышеприведенной модели. Умножив второе из уравнений на $-\frac{\sigma_{12}}{\sigma_{22}}$ и сложив полученные уравнения, получим:

$$y_t = (\beta + \frac{\sigma_{12}}{\sigma_{22}})x_t - \alpha_1 \frac{\sigma_{12}}{\sigma_{22}} x_{t-1} - \alpha_2 \frac{\sigma_{12}}{\sigma_{22}} y_{t-1} - \varepsilon_{1t} - \frac{\sigma_{12}}{\sigma_{22}} \varepsilon_{2t}.$$

Введя для упрощения записи обозначения:

$$\delta_0 = \beta + \frac{\sigma_{12}}{\sigma_{22}}$$

$$\delta_0 = -\alpha_1 \frac{\sigma_{12}}{\sigma_{22}}$$

$$\delta_0 = -\alpha_2 \frac{\sigma_{12}}{\sigma_{22}}$$

$$u_t = \varepsilon_{1t} - \frac{\sigma_{12}}{\sigma_{22}} \varepsilon_{2t},$$

приходим к следующей модели: $y_t = \delta_0 x_t + \delta_1 x_{t-1} + \delta_2 y_{t-1} + u_t$.

Случайный процесс u_t, как линейная комбинация нормальных процессов, сам является нормальным. Очевидно, что $\mathrm{var}(u_t) = \sigma_{11} - \frac{\sigma_{12}}{\sigma_{22}} = \sigma_{11}(1 - p^2)$ и $\mathrm{cov}(u_t, \varepsilon_{2t}) = \mathrm{cov}(\varepsilon_{1t}, \varepsilon_{2t}) - \frac{\sigma_{12}}{\sigma_{22}} \mathrm{cov}(\varepsilon_{2t}, \varepsilon_{2t}) = 0$ по определению параметров σ_{11}, σ_{12}, σ_{22} и ρ. Таким образом, нормальные случайные процессы u_t и ε_{2t} – некоррелированы и, следовательно, независимы. Полученное уравнение для процесса y_t является другой параметризацией уравнения регрессии $y_t = \beta x_t + \varepsilon_{1t}$.

Введенные ранее векторы параметров конкретизируются для данного примера следующим образом:

$$\vec{\phi} = (\beta, \alpha_1, \alpha_2, \sigma_{11}, \sigma_{12}, \sigma_{22})^T$$

$$\vec{\theta}_1 = (\delta_0, \delta_1, \delta_2, \sigma_{11}^2)^T,\ \vec{\theta}_2 = (\alpha_1, \alpha_2, \sigma_{22})^T.$$

Предположим, что вектор $\vec{\alpha}$ содержит только один параметр β, т.е. нас интересует только коэффициент регрессии y_t на x_t. Параметр β выражается через параметры из множеств $\vec{\theta}_1$ и $\vec{\theta}_2$ следующим образом:

$$\beta = \delta_0 + \frac{\delta_1}{\alpha_1} = \delta_0 + \frac{\delta_2}{\alpha_2}.$$

Очевидно, что β невозможно представить в виде функции от параметров, входящих только в множество $\vec{\theta}_1$. Более того, параметры из множеств $\vec{\theta}_1$ и $\vec{\theta}_2$ связаны очевидным соотношением $\alpha_2\delta_1 = \alpha_1\delta_2$, т.е. не являются свободно варьируемыми. Вывод очевиден: переменная (регрессор) x_t не является слабо экзогенной для параметра β.

Несколько изменим модель. Пусть возмущения ε_{1t} и ε_{2t} будут независимыми ($\sigma_{12} = 0$). Тогда

$$\vec{\phi} = (\beta, \alpha_1, \alpha_2, \sigma_{11}, \sigma_{12}, \sigma_{22})^T$$
$$\vec{\theta}_1 = (\beta, \sigma_{11})^T,\ \vec{\theta}_2 = (\alpha_1, \alpha_2, \sigma_{22})^T.$$

Теперь β является элементом множества $\vec{\theta}_1$ и, конечно, является функцией от самого себя. Также очевидно, что параметры множеств $\vec{\theta}_1$ и $\vec{\theta}_2$ являются свободно варьируемыми. Таким образом, в этом случае x_t является слабо экзогенной переменной для параметра β.

Как еще одну модификацию рассмотрим случай, когда σ_{12} вновь не равно нулю, но мы интересуемся другим параметром $-\delta_0$ из преобразованного уравнения. Очевидно, что параметр δ_0 представим как функция параметров из множества $\vec{\theta}_1$, но условие свободной варьируемости не выполнено. Следовательно, процесс x_t не является слабо экзогенным и для параметра δ_0.

Обратите внимание, что если переменная не является слабо экзогенной для параметра, то это еще не означает, что этот параметр нельзя оценить. Например, попытка оценить параметр β методом наименьших квадратов дает несостоятельную оценку, поскольку x_t очевидно коррелирует с ε_{1t}

$$(\operatorname{cov}(x_t, \varepsilon_{2t}) = \operatorname{cov}(\varepsilon_{1t}, \varepsilon_{2t}) = \sigma_{12}).$$

А вот корреляция переменной x_t с возмущением u_t равна нулю, поэтому оценка параметра δ_0 методом наименьших квадратов состоятельна. Но вот точность оценки (ее эффективность) снижена, так как в модель не инкорпорирована отмеченная связь между параметрами множеств $\vec{\theta}_1$ и $\vec{\theta}_2$.

Сильная экзогенность (strong exogeneity)

В соответствии с [4] переменная x_t называется строго экзогенной для параметра β, если она слабо экзогенна для него, и объясняемая переменная y_t не является причиной по Грэнджеру для переменной x_t. Понятие причинности по Грэнджеру (Granger causality) было введено в работе [6]. Этот термин сегодня расценивается как не очень удачный, прежде всего из-за дословного совпадения с причинно-следственными отношениями в обычном смысле, но он уже прочно укоренился в литературе. Лимер [12] предлагал более удачный термин «предшествование» (precedence), но в практике укоренилась именно причинность по Грэнджеру.

Основной посылкой Грэнджера было то, что будущее не может быть причиной настоящего или прошлого. Поэтому, если событие А произошло после события В, то А определенно не может быть причиной В. Но, как знает каждый, «после того не значит вследствие того». При анализе временных рядов часто хотелось бы знать, предшествует ряд x_t ряду y_t, или y_t предшествует x_t, или они «одновременны». Например, предшествует сжатие денежной массы падению производства в России 1990-х гг., или между ними нет соотношения предшествования (причинности по Грэнджеру).

Понятие причинности по Грэнджеру имеет более широкое применение, чем то, в котором мы будем его использовать. Это, скорее, понятие информационное. Сначала запишем формальное определение. Рассмотрим некоторый процесс Z_t, не важно, многомерный или нет. Представим себе, что мы рассчитываем условное математическое ожидание этого процесса: $E\left(Z_{t+1}|\Omega_t\right)$, где Ω_t – вся возможная информация, которая существует в мире в момент *t*. Потом из всевозможной информации, которая известна к моменту *t*, удаляем информацию о некотором процессе x_t. Если при этом условное математическое ожидание процесса Z_t не изменится, т.е. если $E\left(Z_{t+1}|\Omega_t\right)=E\left(Z_{t+1}|\Omega_t \setminus x_s(s\le t)\right)$, то x_t не является причиной по Гренжеру для Z_t. Если же они не равны между собой, т.е. если изъятие информации об x_t меняет условное математическое ожидание, то x_t является причиной по Грэнджеру для Z_t.

Если x_t – причина по Грэнджеру для Z_t, то это не означает, что между этими процессами есть причинно-следственная связь. Единственный вывод состоит в том, что уж если переменная x_t не является причиной по Грэнджеру для переменной Z_t, то она не является ее причиной и в обычном смысле. Грэнджер [6] предложил метод тестирования причинности по Грэнджеру. Он предложил построить регрессию процесса Z_t на его собственные предыдущие значения и на предыдущие значения процесса x_t: $Z_t=\alpha_0+\sum_{i=1}^{k}\alpha_i Z_{t-i}+\sum_{i=1}^{k}\beta_i x_{t-i}+\varepsilon_t$. А после этого проверить обычную гипотезу о равенстве нулю группы коэффициентов: $\begin{matrix} H_0:\beta_1=...=\beta_k=0 \\ H_1:\beta_1^2+...+\beta_k^2>0 \end{matrix}$. Это обычный F-тест. Как обычно, строится полное уравнение и укороченное, и сравниваются остаточные суммы по F-статистике. Если нулевая гипотеза отвергается, то x_t является причиной по Грэнджеру для Z_t. Если же нулевая гипотеза не отвергается, то прошлое процесса x_t не оказывает влияние на процесс Z_t и не является его причиной по Грэнджеру.

Симс в [16] предложил другой подход: x_t не является причиной по Грэнджеру для y_t, если в регрессии y_t на прошлые, текущие и предыдущие значения x_t коэффициенты при будущих значениях x_t совместно равны нулю. Для проведения теста строим регрессию

$$y_t = \sum_{i=-k}^{m} \beta_i x_{t-1} + \varepsilon_t$$

и проверяем гипотезу H_0: $\beta_{-i} = 0$ ($i = 1, \ldots, k$) против естественной альтернативы. Если нулевая гипотеза отвергается, то знание будущих значений x_t не позволяет улучшать прогноз y_t. Хотя с эконометрической точки зрения тесты Грэнджера и Симса не тождественны, они проверяют одно и то же свойство [3]. Оба теста, впрочем, весьма чувствительны к числу лагов, включенных в тестовое уравнение. Идейно тесты должны включать бесконечное число лагов, т.е. всю предысторию x_t. Но нас ограничивает длина реализации, во-первых, в количестве лагов, которое можно применять, во-вторых, падающее число степеней свободы тоже может повлиять на мощность этого теста. Хорошие рекомендации по выбору числа лагов отсутствуют.

Если группа регрессоров является строго экзогенной для параметра β, то этот параметр можно оценить из уравнения регрессии, используя только информацию об условном распределении. В этом случае также можно прогнозировать процесс y_t, основываясь на прогнозе процесса x_t по его собственным прошлым значениям.

Суперэкзогенность (super exogeneity)

Говорят, что группа регрессоров является суперэкзогенной, если изменение их совместного распределения не меняет условного распределения объясняемой переменной y_t. Это свойство связано с так называемой критикой Лукаса [13]. Она заключается в следующем. Одна из целей эконометрического моделирования состоит в прогнозировании эффекта от изменений экзогенных переменных. Однако при изменении этих переменных экономические агенты видят, что происходят изменения, меняют свое поведение и, тем самым, меняют параметры экономической системы. Поэтому модель с постоянными параметрами, говорил Лукас, не адекватна реальным экономическим системам. Свойство суперэкзогенности выделяет модели экономических систем, к которым критика Лукаса не применима.

В соответствии с уже упоминавшимся подходом Cowles Foundation [10] экзогенность переменных и причинно-следственные отношения между переменными не тестируются, они специфицируются из содержательных априорных соображений. Современный же подход [4] предполагает тестирование экзогенности. По самому смыслу этого свойства для проверки, влияет или нет учет распределения регрессоров на оценки регрессии, использующие только условное распределение, нужно построить обе модели: учитывающую DGP (Data Generating Process) полностью и уравнение регрессии. Но если мы в состоянии построить модель, учитывающую DGP полностью, то необходимость в регрессии просто пропадает. Рассмотренный выше пример с исследованием слабой экзогенности процесса x_t относительно параметра β показывает, что нарушение этого свойства связано с наличием регрессора y_{t-1} во втором из уравнений и с неравенством нулю ковариации σ_{12}. Следовательно, наличие слабой экзогенности является частным случаем общей модели при ограничениях на значения параметров: $\sigma_{12}=0,\ \alpha_2=0$. Поэтому для тестирования слабой экзогенности представляется естественным применение теста множителей Лагранжа, при котором оценивается только модель с ограничениями на параметры. Тест, использующий тест множителей Лагранжа для тестирования слабой экзогенности, был разработан Энглом [4].

В условиях нулевой гипотезы оба уравнения для y_t и для x_t могут быть оценены МНК по отдельности. Обозначим остатки соответствующих регрессий через e_y и e_x, причем в обе регрессии включены свободные члены, если только они не должны там отсутствовать из содержательных соображений. Далее строим регрессию e_y на константу, x_t и $e_x : e_{yt}$ = a + bx_t + ge_{xt} + ε_t.

Нулевая гипотеза принимает вид $H_0 : \gamma = 0$, и при условии справедливости H_0 асимптотическое распределение статистики $(T-1)R^2$ имеет асимптотическое распределение χ_1^2 с одной степенью свободы. Здесь R^2 – множественный коэффициент детерминации последней регрессии, а $(T-1)$ – количество наблюдений при ее оценке.

Асимптотически эквивалентный результат можно получить, используя R^2 регрессии e_x на константу, x_{t-1}, y_{t-1} и e_y. Этот подход требует построения трех регрессий. Можно ограничиться построением только двух, если воспользоваться теоремой Фриша–Вау. В этом варианте теста мы строим регрессию y_t на константу, x_t и e_x, а затем проверяем значимость коэффициента при e_x с помощью статистики Стьюдента. Статистический вывод асимптотически эквивалентен предыдущим модификациям. Последняя версия заключается в регрессии x_t на константу, x_{t-1}, y_{t-1} и e_y и проверке значимости коэффициента при e_y.

Для тестирования сильной экзогенности мы тестируем слабую экзогенность и затем причинность по Грэнджеру, как это описано ранее. Проверка суперэкзогенности проводится в три этапа. Сначала проверяется слабая экзогенность. Затем проверяется стабильность параметров условного распределения из множества $\vec{\theta}_1$. Для этой цели могут использоваться введение dummy-переменных и проверка их значимости, тест Чау и другие тесты.

Если стабильность параметров условного распределения установлена, то на третьем этапе проверяется стабильность параметров из множества $\vec{\theta}_2$. Проверку имеет смысл проводить, используя dummy-переменные. Если параметры из множества $\vec{\theta}_2$ оказываются стабильными, то имеющиеся данные не дают оснований для вывода о суперэкзогенности. Если же нестабильность параметров из $\vec{\theta}_2$ установлена, то значимые dummy-переменные добавляются в качестве дополнительных регрессоров в исследуемую регрессию. Если введенные dummy-переменные оказываются незначимыми в совокупности, то суперэкзогенность установлена.

На другой идее проверки слабой экзогенности основан тест Ву–Хаусмана [7, 18]. При отсутствии слабой экзогенности регрессор x_t коррелирован со случайным возмущением, поэтому МНК оценки коэффициентов регрессии являются смещенными и несостоятельными. Состоятельные оценки дает метод инструментальных переменных (IV), о котором мы говорили в курсе эконометрики. Напомню, что инструментальной переменной (инструментом) для x_t называется любая переменная z_t, некоррелированная со случайным возмущением, но коррелированная с x_t. Тест Ву–Хаусмана сравнивает две оценки: метода наименьших квадратов и метода инструментальных переменных. При нулевой гипотезе слабой экзогенности обе оценки состоятельны и асимптотически эквивалентны. При альтернативной гипотезе оценки не эквивалентны.

В качестве инструмента выбирается оценка x_t методом наименьших квадратов. Проиллюстрируем применение теста Ву–Хаусмана на рассматриваемом примере.

$$y_t = \beta x_t + \varepsilon_{1t}$$
$$x_t = \alpha_1 x_{t-1} + \alpha_2 y_{t-1} + \varepsilon_{2t}$$

с теми же свойствами случайных возмущений. На первом шаге строим регрессию, соответствующую второму уравнению, и обозначим через $\hat{x}_t$ рассчитанные значения x_t. Далее оцениваем регрессию

$$y_t = \beta x_t + \gamma \hat{x}_t + \varepsilon_t$$

и проверяем значимость коэффициента γ. В общем случае нескольких регрессоров строятся инструменты для каждого из них, все инструменты добавляются в регрессию, и проверяется гипотеза о совместной значимости коэффициентов при инструментальных переменных. Если нулевая гипотеза не отвергается, делается вывод о наличии слабой экзогенности.

Впрочем, на практике достаточно часто просто постулируют слабую экзогенность переменных, исходя из содержательных соображений. Напомним, что стационарность переменных и слабая экзогенность регрессоров позволяют получить состоятельные оценки коэффициентов ADL моделей, и обычные процедуры проверки гипотез являются асимптотически верными.

Лекция 13

Многомерные процессы

До сих пор мы рассматривали модели, которые состоят только из одного соотношения, связывающего временные ряды. При этом мы выбирали одну из переменных в качестве эндогенной, а остальные переменные являлись экзогенными. Такое разделение не всегда является естественным, часто приходится рассматривать одновременно несколько соотношений, в которые одни и те же переменные входят и как эндогенные, и как экзогенные. Как видно из прошлой лекции, переменная не всегда может рассматриваться как экзогенная, и мы фактически должны рассматривать модель DGP, состоящую из нескольких уравнений. Это означает моделирование нескольких временных рядов одновременно, другими словами − моделирование многомерного случайного процесса.

Начнем с определений. Рассмотрим вектор $\vec{X}_t = (x_t^1, x_t^2, ..., x_t^k)^T$, каждая компонента которого является временным рядом. Верхним индексом будем обозначать номер компоненты, а нижним по-прежнему − момент времени. Распределение компонент характеризуется семейством совместных плотностей распределения вида: $f_n(x_{t_1}^{i_1}, x_{t_2}^{i_2}, ..., x_{t_n}^{i_n}),\quad n = 1,2,...$. Условием стационарности в узком смысле по-прежнему является независимость от сдвига во времени всего семейства совместных плотностей распределения. Только теперь кроме всевозможных комбинаций значений случайного процесса в различные моменты времени аргументами плотностей вероятности также являются всевозможные комбинации различных компонент в различные моменты времени. Например, для двухмерной плотности получаем из условия стационарности: $f_2(x_t^1, x_t^2) = f_2(x_{t+\tau}^1, x_{t+\tau}^2)$ для любого τ. Совместное распределение компонент для одного и того же момента времени не зависит от времени. Рассмотрим другую функцию распределения, например трехмерную, в которую входят значения первой компоненты в два разных момента времени и второй компоненты в некоторый третий момент времени. Стационарность означает, что $f_3(x_t^1, x_{t+h}^1, x_{t+s}^2) = f_3(x_{t+\tau}^1, x_{t+h+\tau}^1, x_{t+s+\tau}^2)$. Можно сказать, что это свойство инвариантности к сдвигу во времени. То есть, если к каждому моменту времени прибавить величину τ, то функция плотности не изменится. Понятно, что стационарность многомерного процесса влечет за собой стационарность каждой из его компонент.

Как и в одномерном случае, стационарность в узком смысле влечет за собой ряд свойств характеристик случайных процессов. Прежде всего, начнем с математического ожидания. Математическое ожидание для каждой компоненты не зависит от других компонент. Поэтому если многомерный процесс стационарен, математическое ожидание каждой компоненты не зависит от времени. Вектор математических ожиданий $E(\vec{X}_t) = \begin{pmatrix} \mu_1 \\ ... \\ \mu_k \end{pmatrix}$ не зависит от времени.

Теперь рассмотрим моменты второго порядка. Каждая компонента характеризуется дисперсией и автокорреляционной функцией. Если одномерный ряд стационарен, его автокорреляционная и автоковариационная функции зависят только от сдвига τ: $Corr(\tau) = Corr(x_t^i, x_{t+\tau}^i) = \rho_i(\tau)$. Однако теперь можно рассмотреть второй смешанный момент для различных компонент, а также $Corr(x_t^i, x_{t+\tau}^j)$. Такую величину естественно назвать кросс-корреляционной функцией. Если компоненты образуют многомерный стационарный процесс, то кросс-корреляция будет функцией сдвига во времени τ. Обозначим эту функцию $R_{ij}(\tau)$. Довольно очевидно, что $R_{ij}(\tau) = R_{ji}(-\tau)$. При фиксированном значении τ элементы $R_{ij}(\tau)$ образуют матрицу R, зависящую от τ. Значению τ, равному нулю, соответствует корреляционная матрица вектора $\vec{X}_t$.

Как и в одномерном случае, назовем многомерный процесс стационарным в широком смысле (слабо стационарным), если вектор математических ожиданий не зависит от времени, и кросс-корреляции между любыми компонентами зависят только от разности во времени. Можно сказать, что у стационарного многомерного процесса компоненты стационарны и стационарно связаны между собой.

Модель многомерного временного ряда обычно будет задаваться в виде уравнения для каждой из компонент, причем в виде объясняющих переменных будут выступать текущие и предыдущие значения, вообще говоря, всех компонент. Другими словами, модель будет представлена системой одновременных уравнений. Одновременность не дает напрямую использовать такой распространенный метод, как метод наименьших квадратов для оценки параметров многомерных моделей. Эта особенность не имеет специфики временных рядов, а относится именно к оценке одновременных уравнений. Поскольку системы одновременных уравнений не входили еще в данный курс эконометрики, кратко рассмотрим возникающие трудности.

Пусть задана простая Кейнсианская макроэкономическая модель национального дохода, состоящая только из двух следующих уравнений:

$$\begin{cases} Y_t = C_t + Z_t \\ C_t = \alpha + \beta Y_t + u_t \end{cases}$$

где Y_t – национальный доход, C_t – потребление, Z_t – инвестиции и государственные расходы, а u_t – белый шум. Первое из уравнений представляет собой условие равновесия национального дохода, а второе – функцию потребления. Коэффициент β – предельная склонность к потреблению, а коэффициент α – автономное потребление. В соответствии с экономическим смыслом считаем эндогенными переменными Y_t и C_t – ВВП и потребление, а Z_t считаем экзогенной переменной. Эти уравнения и их коэффициенты отражают экономический смысл соотношений, в этом случае говорят, что уравнения заданы в *структурной (structural) форме*. Индекс *t* указывает на то, что переменные являются временными рядами, хотя уравнения и не содержат лаговых составляющих. Пусть собраны данные о потреблении, о расходах и о национальном доходе. Можно ли, отвлекаясь от проблемы экзогенности, оценить функцию потребления просто из второго соотношения методом наименьших квадратов?

Модель представленного типа называется одновременными уравнениями. В данном случае у нас два уравнения, причем эконометрическое у нас только одно. Первое соотношение − это обычное экономическое тождество, мы ничего не оцениваем в этом уравнении. Мы хотим оценить только коэффициенты α и β, но из-за того, что существует дополнительное экономическое тождество, появляется проблема. Если просто подставить C_t в первое соотношение, то получим $Y_t = \alpha + \beta Y_t + Z_t + u_t$. Разрешив полученное уравнение относительно национального дохода, получаем $Y_t = \frac{\alpha}{1-\beta} + \frac{1}{1-\beta} Z_t + \frac{1}{1-\beta} u_t$. Подстановка Y_t во второе уравнение дает так называемую *приведенную (reduced) форму* модели:

$$\begin{cases} C_t = \dfrac{\alpha}{1-\beta} + \dfrac{\beta}{1-\beta} Z_t + \dfrac{1}{1-\beta} u_t \\ Y_t = \dfrac{\alpha}{1-\beta} + \dfrac{1}{1-\beta} Z_t + \dfrac{1}{1-\beta} u_t \end{cases}.$$

Теперь эндогенные переменные выражены через экзогенные переменные и случайные возмущения. Различие между структурной и приведенной формой носит не только математический, но и содержательный характер. Вообще-то нас интересуют коэффициенты структурных уравнений, они имеют экономический смысл. Например, β − это предельная склонность к потреблению. А коэффициенты в приведенной форме уже не несут экономического смысла.

Приведенная форма явно показывает, что переменная Y_t коррелирована со случайным возмущением u_t. Формально в структурное уравнение, определяющее Y_t, не входит никакая случайная компонента. Тем не менее, из-за того, что это одновременные уравнения, Y_t зависит от u_t. Это означает, что метод наименьших квадратов даст смещенную и несостоятельную оценку при оценивании функции потребления. Он приведет к смещению, вызванному совместными уравнениями. Даже асимптотически метод наименьших квадратов, примененный к структурному уравнению, приводит к плохим результатам. Из-за одновременности уравнений ошибка из любого уравнения как бы проникает во все остальные уравнения.

Можно использовать МНК для оценки коэффициентов приведенной формы, если экзогенная переменная Z_t детерминирована или не коррелирует с возмущением u_t. Для нахождения коэффициентов структурной формы применяются два способа.

1. *Косвенный (indirect) метод наименьших квадратов.* Методом наименьших квадратов оцениваем коэффициенты приведенной формы, для чего оцениваем регрессии:

$$\begin{cases} Y_t = \pi_{10} + \pi_{11} \cdot Z_t + \varepsilon_t^1 \\ C_t = \pi_{20} + \pi_{21} \cdot Z_t + \varepsilon_t^2 \end{cases}.$$

Каждая из регрессий оценивается отдельно. Для нахождения параметров структурной формы α и β можно попытаться решить систему уравнений:

$$\frac{\alpha}{1-\beta} = \hat{\pi}_{10}, \quad \frac{\alpha}{1-\beta} = \hat{\pi}_{20}, \quad \frac{\beta}{1-\beta} = \hat{\pi}_{21}, \quad \frac{1}{1-\beta} = \hat{\pi}_{11}.$$

Очевидно, что эта система четырех уравнений с двумя неизвестными, вообще говоря, не имеет решения. В этом и состоит основной недостаток косвенного метода наименьших квадратов. Приведенную систему всегда можно оценить. Предположим, система состоит не из двух, а из 10 уравнений с 10 переменными. Выбираем эндогенные и экзогенные переменные и строим регрессию каждой эндогенной переменной на все экзогенные. После этого основной вопрос заключается в том, можно ли от коэффициентов приведенной формы вернуться однозначно к структурным коэффициентам, которые и имеют содержательный смысл. Говорят, что система идентифицируема, если по коэффициентам приведенной формы можно однозначно вернуться к структурным коэффициентам. Иногда соответствующая система уравнений противоречива, иногда решения будут существовать, но не будут единственными. Существуют системы одновременных уравнений, в которых косвенный МНК приводит к однозначному результату. Идентифицируемость зависит в том числе и от того, сколько экзогенных переменных присутствуют в модели и как они входят в структурные уравнения. В этом, на самом деле, одна из причин того, что в нашем примере мы не можем оценить структурные коэффициенты – у нас одна единственная экзогенная переменная, всего два структурных параметра и четыре – в приведенной системе. С проблемой идентифицируемости вы познакомитесь подробней в дальнейшем курсе эконометрики при изучении темы «системы одновременных уравнений». Кажущийся таким простым косвенный метод наименьших квадратов не всегда может дать оценки структурных параметров.

Для дальнейшего изложения важно понять, что, имея систему одновременных уравнений, нельзя оценивать каждое уравнение в отдельности из-за того, что эндогенные переменные входят и в левую, и в правую часть, и из-за того, что равновесие определяется всеми уравнениями вместе. Случайная ошибка из любого уравнения проникает во все остальные, из-за чего оказываются коррелированны регрессор и случайное возмущение.

2. *Метод инструментальных переменных*. Этот метод известен из курса эконометрики и кратко упоминался в предыдущей лекции. В случае системы одновременных уравнений можно легко предложить набор инструментальных переменных. Они строятся как линейные комбинации экзогенных переменных, подобранные так, чтобы не было коррелированности инструментов и случайных возмущений. Этот метод тоже зависит от того, насколько идентифицируема оцениваемая система уравнений. При отсутствии лаговых переменных двумя относительно широко применяемыми вариантами этого подхода являются так называемые двухшаговый МНК и трехшаговый МНК. Мы не будем на них специально останавливаться. Это отдельная проблема.

Возвращаясь к временным рядам, мы видим, что при оценивании моделей, состоящих из нескольких уравнений, возникают сложности, связанные с наличием среди регрессоров «одновременных» составляющих. Кроме того, надо попытаться разделить переменные на эндогенные и экзогенные. Иногда это разделение экономически оправдано (как в вышерассмотренном примере), иногда не очень. Например, один из подходов к исследованию эффективности рынка – это моделирование временных рядов кросс-курсов различных валют и цен разных финансовых инструментов. Например, рассматривая курсы рубль/доллар, доллар/евро, рубль/евро, трудно сказать, какой из них естественно выбрать эндогенной переменной, а какие – экзогенными.

Векторная авторегрессия

В 1980 г. Симс [15] предложил подход, который позволяет уйти от разделения переменных на экзогенные и эндогенные и избавиться от сложностей, связанных с одновременностью уравнений. Этот подход является к тому же естественным обобщением подхода Бокса–Дженкинса к моделям ARIMA и носит название VAR, или *векторная авторегрессия*. В принципе существует и VARIMA [17], но, как мы увидим дальше, это скорее теоретически, чем практически. Для того, чтобы избежать смещения, связанного с применением МНК непосредственно к каждому уравнению структурной формы, Симс предложил, не производя деления переменных на экзогенные и эндогенные, представить каждую из компонент многомерного случайного процесса как линейную комбинацию от предыдущих значений всех переменных.

Рассмотрим сначала пример двумерного вектора и только одного лага. Пусть, например, X_t^1 – темп роста денежной массы, а X_t^2 – дефлятор ВВП. Запишем систему уравнений указанного вида:

$$\begin{cases} X_t^1 = \alpha_1 + \beta_{11} X_{t-1}^1 + \beta_{12} X_{t-1}^2 + \varepsilon_t^1 \\ X_t^2 = \alpha_2 + \beta_{21} X_{t-1}^1 + \beta_{22} X_{t-1}^2 + \varepsilon_t^2 \end{cases}.$$

Здесь случайные возмущения предполагаются белыми шумами, вообще говоря, коррелированными. Стандартное обозначение этой модели – VAR(1), где 1 – число лагов. Размерность многомерного случайного процесса обычно не указывается.

Модель VAR(p) в матрично-векторных обозначениях имеет вид:

$$\vec{X}_t = \vec{a} + A_1 \vec{X}_{t-1} + A_2 \vec{X}_{t-2} + \ldots + A_p \vec{X}_{t-p} + \vec{\varepsilon}_t,$$

где $\vec{\varepsilon}_t$ – векторный белый шум со следующими свойствами: $E(\vec{\varepsilon}_t) = 0$ для всех t, $\operatorname{cov}(\vec{\varepsilon}_t) = E(\vec{\varepsilon}_t \vec{\varepsilon}_s^{\,T}) = \begin{cases} \Omega, s = t \\ 0, s \neq t \end{cases}$. Модель сразу выписывается в приведенной форме. Одновременные компоненты не входят в правую часть. Для рассмотренного двумерного примера получаем:

$$\begin{pmatrix} X_t^1 \\ X_t^2 \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} + \underbrace{\begin{pmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{pmatrix}}_{A_1} \begin{pmatrix} X_{t-1}^1 \\ X_{t-1}^2 \end{pmatrix} + \begin{pmatrix} \varepsilon_t^1 \\ \varepsilon_t^2 \end{pmatrix}.$$

Модель VAR(p) напоминает модель ARMA, поэтому многие их свойства схожи. Перепишем модель VAR(p), используя оператор лага,:

$$(E - A_1 L - A_2 L^2 - ... - A_p L^p)\vec{x}_t = \vec{\varepsilon}_t.$$

Для упрощения записи мы предполагаем, что нет свободного члена, это не приводит к потери общности, но немного сокращает запись. Для рассматриваемого примера можем записать:

$$\begin{pmatrix} 1-\beta_{11}L & -\beta_{12}L \\ -\beta_{21}L & 1-\beta_{22}L \end{pmatrix} \begin{pmatrix} x_t^1 \\ x_t^2 \end{pmatrix} = \begin{pmatrix} \varepsilon_t^1 \\ \varepsilon_t^2 \end{pmatrix}.$$

Также, как и в одномерном случае, возникает вопрос, можно ли выразить вектор $\vec{x}_t$ через текущие и предыдущие значения вектора случайных возмущений $\vec{\varepsilon}_t$. Очевидно, что если существует обратная матрица, то можно записать:

$$\begin{pmatrix} x_t^1 \\ x_t^2 \end{pmatrix} = \frac{1}{\Delta} \begin{pmatrix} 1-\beta_{22}L & \beta_{12}L \\ \beta_{21}L & 1-\beta_{11}L \end{pmatrix} \begin{pmatrix} \varepsilon_t^1 \\ \varepsilon_t^2 \end{pmatrix}.$$

Если обозначить через I единичную матрицу, то определитель матрицы системы $\Pi = I - A_1$ равен:

$$\Delta = (1-\beta_{11}L)(1-\beta_{22}L) - \beta_{12}\beta_{21}L^2 = 1 - (\beta_{11}+\beta_{22})L + (\beta_{11}\beta_{22} - \beta_{12}\beta_{21})L^2 = (1-\mu_1 L)(1-\mu_2 L),$$

где μ_1 и μ_2 − собственные числа матрицы Π, возможно комплексные.

Если хотя бы одно из собственных чисел равно 0, то матрица системы уравнений вырождена, и $\vec{x}_t$ не выражается через текущие и предыдущие значения векторного белого шума. Вспомнив условия стационарности моделей ARMA(p,q), можно заключить, что при условии $|\mu_1| < 1$ и $|\mu_2| < 1$ каждая из компонент процесса $\vec{x}_t$ выражается в виде бесконечного ряда текущего и предыдущего значений компонент белого шума. Поскольку коэффициенты каждого из разложений будут учитывать в геометрической прогрессии, по теореме Вольда каждая из компонент процесса $\vec{x}_t$ будет стационарной в широком смысле.

Для изучения условий стационарности и свойств моделей VAR(1) удобно использовать сведение матриц A_1 и Π к простейшей форме. Заметим, что матрицы Π и A_1 очевидно имеют одни и те же собственные векторы, а их собственные числа дополняют друг друга до единиц. Обозначим собственные числа матрицы A_1 через $\lambda_1 = 1 - \mu_1$, и $\lambda_2 = 1 - \mu_2$. Поэтому условие, что собственные числа матрицы Π лежат в единичном круге, эквивалентно аналогичному условию для собственных чисел матрицы A_1.

Рассмотрим сначала случай, когда $\lambda_1 \neq \lambda_2$. Из линейной алгебры мы знаем, что собственные векторы, соответствующие различным собственным числам, линейно независимы. Введем невырожденную матрицу C, столбцы которой являются собственными векторами, соответствующими собственным числам λ_1 и λ_2 соответственно. Обозначим через Λ диагональную матрицу diag{λ_1, λ_2}. Тогда очевидно, что $C^{-1}A_1C = \Lambda$ и $A_1 = C\Lambda C^{-1}$. Определим новый вектор переменных соотношением $\vec{y}_t = C^{-1}\vec{x}_t$, отсюда следует также, что $\vec{x}_t = C\vec{y}_t$. Умножая уравнение $\vec{x}_t = A_1\vec{x}_{t-1} + \vec{\varepsilon}_t$ слева на матрицу C^{-1}, получим $\vec{y}_t = \Lambda\vec{y}_{t-1} + \vec{\eta}_t$, где через $\vec{\eta}_t$ обозначен «новый» векторный белый шум.

В новых переменных система уравнений распадается на два отдельных уравнения:

$$y_t^1 = \lambda_1 y_{t-1}^1 + \eta_t^1$$
$$y_t^2 = \lambda_2 y_{t-1}^2 + \eta_t^2,$$

свойства которых можно установить, используя уже известные нам приемы.

Очевидно, что y_t^1 и y_t^2 слабо стационарны при $|\lambda_1| < 1$ и $|\lambda_2| < 1$, и, по крайней мере, одна из компонент вектора $\vec{y}_t$ не является стационарной при нарушении этого условия. Поскольку x_t^1 и x_t^2 являются линейными комбинациями y_t^1 и y_t^2, то условием стационарности векторного процесса $\vec{x}_t$ является нахождение собственных чисел матрицы A_1 внутри единичного круга.

Отдельного рассмотрения заслуживает случай, когда $\lambda_1 = \lambda_2$. В этом случае двух линейно независимых векторов может и не существовать, и матрицу A_1 нельзя привести к диагональной форме. Однако доказано, что тогда существует матрица P, которая приводит матрицу A_1 к так называемой жордановой форме. Мы ограничимся только кратким рассмотрением этого понятия без доказательств. Итак, доказано, что существует невырожденная матрица P, такая что $P^{-1}A_1P = J$ и $A_1 = PJP^{-1}$, где J – жорданова матрица: $J = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$. Определив вектор $\vec{y}_t = P^{-1}\vec{x}_t$ ($\vec{x}_t = P\vec{y}_t$), сведем исходную модель к виду $\vec{y}_t = J\vec{y}_{t-1} + \vec{\eta}_t$, где сейчас векторный белый шум $\vec{\eta}_t = P^{-1}\vec{\varepsilon}_t$. В этом случае система уравнений:

$$y_t^1 = \lambda_1 y_{t-1}^1 + y_{t-1}^2 + \eta_t^1$$
$$y_t^2 = \lambda_2 y_{t-1}^2 + \eta_t^2$$

не распадается на отдельные уравнения, но второе из них может быть решено и исследовано отдельно от первого. Очевидно, что при условии $|\lambda|<1$ второе уравнение представляет собой стационарную модель AR(1), а первое − стационарную модель ADL(1,1). При $|\lambda|\geq 1$ оба уравнения нестационарны. Таким образом, оба рассмотренных случая дают единое условие стационарности: собственные числа матрицы A_1 (или Π) должны лежать строго внутри единичного круга.

Этот результат легко обобщить на случай VAR(1) модели произвольной размерности. Если все собственные числа $\lambda_1, \lambda_2, ..., \lambda_k$ матрицы A_1 различны, то модель распадается на отдельные уравнения $y_t^i = \lambda_i y_{t-1}^i + \eta_t^i \quad (i = 1,2,..,k)$. Каждое из них стационарно тогда и только тогда, когда $|\lambda_i| < 1$. Если же среди собственных чисел $\lambda_1, \lambda_2, ..., \lambda_k$ есть группа равных по величине, то, обозначив их без потери общности номерами 1,2,...,k, приводим модель к тому же виду: $\vec{y}_t = J\vec{y}_{t-1} + \vec{\eta}_t$. Матрица J содержит так называемую жорданову клетку S размера r в левом верхнем углу:

$$S = \begin{pmatrix} \lambda & 1 & . & 0 & 0 \\ 0 & \lambda & . & 0 & 0 \\ . & . & . & . & . \\ 0 & 0 & . & \lambda & 1 \\ 0 & 0 & . & 0 & \lambda \end{pmatrix}$$

На главной диагонали жордановой клетки стоят собственные числа, первая над-диагональ содержит 1, а остальные элементы равны 0. Отдельные уравнения системы принимают вид:

$$\begin{aligned} &y_t^1 = \lambda y_{t-1}^1 + y_{t-1}^2 + \eta_t^1 \\ &. \\ &. \\ &. \\ &y_t^{r-1} = \lambda y_{t-1}^{r-1} + y_{t-1}^r + \eta_t^{r-1} \\ &y_t^r = \lambda y_{t-1}^r + \eta_t^r \\ &y_t^{r+1} = \lambda_{r+1} y_{t-1}^{r+1} + \eta_t^{r+1} \\ &. \\ &. \\ &. \\ &y_t^k = \lambda_k y_{t-1}^k + \eta_t^k . \end{aligned}$$

Преобразованная система ясно показывает, что условие стационарности имеет тот же вид, что и для случая двумерной модели.

Для модели VAR(p) рассмотрение, подобное проведенному выше, невозможно. Для получения условий стационарности можно толковать модель VAR(p) как систему разностных линейных уравнений с постоянными коэффициентами и правой частью $\vec{\varepsilon}_t$. По аналогии со скалярным случаем будем искать общее решение однородной системы $\vec{x}_t - A_1\vec{x}_{t-1} - ... - A_p\vec{x}_{t-p} = 0$ в виде $\vec{x}_t = \vec{C}(\lambda)^t$. Подставив это выражение в уравнение, получим:

$$(\lambda^p I + \lambda^{p-1}A_1 + ... + A_p)\vec{C} = 0.$$

Однородная система линейных уравнений имеет ненулевые решения при λ равных собственным числам матрицы $\lambda^p I + \lambda^{p-1}A_1 + ... + A_p$, т.е. корням характеристического уравнения $\det(\lambda^p I + \lambda^{p-1}A_1 + ... + A_p) = 0$.

Соответствующими нетривиальными решениями системы

$$(\lambda^p I + \lambda^{p-1}A_1 + ... + A_p)\vec{C} = 0$$

являются собственные векторы ее матрицы. Тогда слагаемое, соответствующее собственному числу λ_i, стремится к нулю при неограниченном увеличении *t*. Если среди $p \times k$ собственных чисел есть одинаковые, то нужно заменить в выражении $\vec{x}_t = \vec{C}(\lambda)^t$ постоянный вектор $\vec{C}$ зависящим от времени вектором вида $\vec{C}_0 + \vec{C}_1 t + ... + \vec{C}_{r-1}t^{r-1}$. Во всех случаях условие стационарности выражается компактно: все корни характеристического уравнения должны лежать строго внутри единичного круга.

Если условие стационарности выполнено, и случайное возмущение $\vec{\varepsilon}_t$ обладает свойствами белого шума, то уравнения модели VAR(p) можно оценивать обычным методом наименьших квадратов. Причем каждое уравнение можно оценивать отдельно, т.е. строить регрессию компоненты x_t^i на предыдущие значения всех компонент, включая предыдущие значения самой x_t^i. Во-первых, нет корреляции регрессоров с возмущениями, и нет никаких проблем с состоятельностью оценок. Кроме того, все переменные совершенно равноправны, и поскольку все регрессоры предшествуют объясняемым переменным, не возникает проблем с экзогенностью. Таким образом, оценивание моделей VAR(p) требует только многократного применения метода наименьших квадратов и легко реализуемо в специализированных пакетах. Обычно в компьютерных программах, например в том же Econometric views, задается список переменных, число лагов, и этого достаточно для построения модели.

Одним из больших препятствий при практическом применении VAR моделей является очень большое число подлежащих оцениванию параметров. Рассмотрим простой случай. Как вы знаете, стандартная кейнсианская модель состоит из 7 уравнений. Переменные модели являются макроэкономическими показателями. По каким данным вы будете оценивать такую модель? Ну, не меньше, чем по квартальным. Сколько лагов включать в модель? Конечно, существуют статистические методы оценки числа лагов, но попробуем использовать содержательные знания. Знаменитая Сент-Луисская модель [2] демонстрирует, что в уравнение темпа роста номинального ВВП входят темпы роста денежного агрегата M_1 и темпы роста правительственных расходов на поддержание занятости с запаздыванием вплоть до четырех кварталов. Давайте возьмем модель с 5 лагами, пусть используются квартальные данные. Сколько параметров нам надо оценить, если мы строим VAR модель? 5 лагов – это значит 5 матриц размерности 7 на 7. В каждой 49 параметров, плюс столбец свободных членов, плюс 7 дисперсий случайных составляющих. Итого мы получаем, что надо оценить $49 \cdot 5 + 7 + 7 = 259$ коэффициентов. В каждом из уравнений нужно оценить 37 параметров. Для этого необходимы данные хотя бы лет за 15. Вот эта простая арифметика называется перепараметризацией VAR, т.е. VAR вводит слишком много параметров в модель. Это свойство препятствует широкому применению VAR моделей.

В то же время оказалось, что VAR представление тесно связано с исследованием стационарности и возможностью построения эконометрических моделей с нестационарными регрессорами. Это и будет предметом нашего рассмотрения в последующих лекциях.

* *
*

СПИСОК ЛИТЕРАТУРЫ

1. Индексы интенсивности промышленного производства / Центр экономической конъюнктуры при Правительстве РФ, ежемесячное издание. Также официальный сайт Центра http://www.cea.gov.ru

2. *Carlson K.M.* Does the St. Louis Equation Now Believe in Fiscal Policy // Review., Federal Bank of St. Louis. 1978. V. 60. № 2. P. 17.

3. *Chambarlain G.* The General Equivalence of Granger and Sims Causality // Econometrica. 1982. V. 50. P. 569–582.

4. *Engle R.F.* Wald, Likelihood Ratio and Lagrange Multiplier Tests in Econometrics // Handbook of Econometrics, eds. Z. Griliches and M.D. Intriligator. Chapter 13. North-Holland, 1984.

5. *Franses P.H.* Periodicity and Stochastic Trends in Economic Time Series. Oxford Univercity Press, 1996.

6. *Granger C.V.J.* Investigating Causal Relations by Econometric Methods and Cross-Spectral Methods // Econometrica. 1969. V. 37. P. 424–438.

7. *Hausman A.* Specification Tests in Econometrics // Econometrica. V. 46. P. 1251–1271.

8. *Hendry D.F.* and *Ericsson N.R.* Modeling the Demand for Narrow Money in the United Kingdom and the United States // European Economic Review. 1991. V. 35. P. 833–866.

9. *Hendry D.F.* and *Richard J.-F.* Exogeneity // Econometrica. 1983. V. 51. P. 277–304.

10. *Hood W. C. and Coopmans T. S.* Studies in Econometric Method. Wiley, 1953.

11. *Hylleberg S.*, *Engle R.F.*, *Granger C.W.J.*, Yoo *B.S.* Seasonal Integration and Cointegration // Journal of Econometrics. 1990. № 44. P. 215–238.

12. *Leamer E.E.* Vector Autoregressions for Cansal Inference // K.L. Brunner and A. Meltrer (eds.) Understanding Monetary Regims (Supplement to the Journal of Monetary Economics). 1985. P. 255–304.

13. *Lucas R.E.* Econometric Policy Evaluation: A Critique // K.L. Brunner (eds.) The Phillips Curve and Labor Markets (Supplement to the Journal of Monetary Economics). 1976. P. 19–46.

14. *Sargan J.D.* Wages and Prices in the United Kingdom: A Study in Econometric Methodology // P.E. Hart, G. Mills, K. Whiteker (eds.) Econometric Analysis for National Economic Planning, Butterworth. London, 1964; reprinted in D.F. Hendry and K.F. Wallis (eds.) Econometrics and Quantitative Economics. Oxford: Basil Blackwell, 1984.

15. *Sims C.A.* Macroeconomics and Reality // Econometrica. 1980. V. 48. P. 1–48.

16. *Sims C.A.* Money, Income and Causality // American Economic Review. 1972. V. 62. P. 450–552.

17. *Tsiao G.C.* and *Box G.E.P.* Modelling Multiple Time Series with Applications // Journal of the American Statistical Association. 1981. V. 76. P. 802–816.

18. *Wu D.* Alternative Tests of Independence between Stochastic Regressors and Disturbances // Econometrica. 1973. V. 41. P. 733–750.