Глава 4. Регрессионный анализ для стационарных объясняющих переменных

Прежде, чем переходить к изложению материала этой главы, заметим, что в этой главе мы не будем различать в обозначениях случайные величины и их наблюдаемые значения – в обоих случаях будут использоваться строчные буквы.

4.1. Асимптотическая обоснованность стандартных процедур

В главе 1 мы уже отмечали, что рассмотренные там случаи, в которых можно использовать стандартные процедуры регрессионного анализа несмотря на то, что объясняющие переменные являются стохастическими (ситуации A, A′, B, C), не охватывают наиболее интересных для нас моделей стационарных и нестационарных временных рядов. Это замечание относится и к широко используемым на практике моделям авторегрессии, в том числе и стационарным.

Рассмотрим модель авторегрессии AR(p)

$y_t = \alpha + a_1 y_{t-1} + a_2 y_{t-2} + \ldots + a_p y_{t-p} + \varepsilon_t$,

где ε_t – инновации, образующие процесс белого шума с $D(\varepsilon_t) = \sigma_\varepsilon^2$. Эту модель можно представить в виде линейной модели регрессии

$y_t = x_t^T \theta + \varepsilon_t$,

где

$x_t = (1, y_{t-1}, y_{t-2}, \ldots , y_{t-p})^T$, $\theta = (\alpha , a_1 , a_2 , \ldots , a_p)^T$.

Но мы не можем использовать для нее результаты, полученные в ситуациях A и B. Хотя ε_s и x_t статистически независимы при $s \ge t$, они оказываются зависимыми уже при $s = t - 1$, поскольку ε_{t-1} участвует в формировании случайной величины y_{t-1}, входящей в состав x_t . Это нарушает условие, входящее в определения ситуаций A и B.

Мы не можем также использовать и результаты, полученные в ситуациях A′ и C. Там требовалось, чтобы условное распределение вектора ε при фиксированной матрице X имело вид $N(0, \sigma^2 V)$ с положительно определенной (невырожденной) матрицей V (в ситуации A′ это единичная матрица). Однако при фиксированных значениях $x_{t+1} = (1, y_t , y_{t-1}, \ldots , y_{t-p+1})^T$ и x_t значение ε_t известно с полной определенностью.

Тем не менее, если AR(p) модель стационарна, то положение вполне благополучно:

Ситуация D

- процесс y_t порождается моделью

$$y_t = \alpha + a_1 y_{t-1} + a_2 y_{t-2} + \ldots + a_p y_{t-p} + \varepsilon_t \quad (\varepsilon_t \text{ – инновации});$$

- все корни полинома $1 - a_1 z - a_2 z^2 - \ldots - a_p z^p = 0$ лежат за пределами единичного круга;
- $\varepsilon_t \sim i.i.d.$, $E(\varepsilon_t) = 0$, $D(\varepsilon_t) = \sigma^2 > 0$, $E(\varepsilon_t^4) = \mu_4 < \infty$.

При выполнении перечисленных условий, для оценки наименьших квадратов $\hat{\theta}_n$ вектора коэффициентов $\theta = (\alpha, a_1, a_2, \ldots, a_p)$, полученной по n наблюдениям, выполняется соотношение

$$n^{1/2}(\hat{\theta}_n - \theta) \to N(0, \sigma^2 Q^{-1}),$$

где Q – положительно определенная матрица, элементы которой выражаются в явной форме через математическое ожидание и автокорреляции процесса y_t . Ковариационная матрица $\sigma^2 Q^{-1}$ асимптотического распределения может быть оценена состоятельно посредством $S_n^2(X_n^T X_n / n)^{-1}$, и это означает, что асимптотически обоснованны статистические процедуры, трактующие распределение $\hat{\theta}_n$ как $N(\theta, S_n^2(X_n^T X_n)^{-1})$. (Здесь X_n – матрица значений объясняющих переменных в n наблюдениях.)

Иными словами, и в рассматриваемой ситуации можно пользоваться <u>стандартными</u> методами регрессионного анализа, имея в виду их <u>асимптотическую</u> обоснованность.

Если перейти к процессам, <u>стационарным относительно детерминированного тренда</u>, то следует отметить возникающую здесь особенность, связанную со сходимостью распределения оценок наименьших квадратов к асимптотическому распределению. Мы поясним эту особенность на следующем примере.

Пусть ряд y_t порождается простой моделью временного тренда

$$y_t = \alpha + \beta t + \varepsilon_t,$$

где $\varepsilon_t \sim i.i.d.$, $E(\varepsilon_t) = 0$, $D(\varepsilon_t) = \sigma^2 > 0$, $E(\varepsilon_t^4) = \mu_4 < \infty$. Если и здесь записать модель в стандартной форме

$$y_t = x_t^T \theta + \varepsilon_t, \quad x_t = (1, t)^T, \quad \theta = (\alpha, \beta),$$

то чтобы получить <u>невырожденное</u> асимптотическое распределение оценки наименьших квадратов $\hat{\theta}_n = (\hat{\alpha}_n, \hat{\beta}_n)$, приходится использовать <u>различные</u> нормирующие множители: $(\hat{\alpha}_n - \alpha_n)$ умножается на $T^{1/2}$, а $(\hat{\beta}_n - \beta_n)$ умножается на $T^{3/2}$. Однако это различие компенсируется тем, что аналогичным образом ведут себя и стандартные ошибки для $\hat{\alpha}_n$ и $\hat{\beta}_n$. Как результат, обычные t-статистики имеют асимптотическое $N(0, 1)$ распределение. Иными словами, можно использовать стандартную технику регрессионного анализа, имея в виду ее асимптотическую обоснованность.

Те же самые принципы можно использовать и для исследования процесса авторегрессии произвольного порядка, стационарного относительно детерминированного временного тренда.

<u>Ситуация Е</u>

- процесс y_t порождается моделью

$$y_t = \alpha + \beta t + a_1 y_{t-1} + a_2 y_{t-2} + \ldots + a_p y_{t-p} + \varepsilon_t \quad (\varepsilon_t \text{ – инновации});$$

- все корни полинома $1 - a_1 z - a_2 z^2 - \ldots - a_p z^p = 0$ лежат за пределами единичного круга;
- $\varepsilon_t \sim i.i.d.$, $E(\varepsilon_t) = 0$, $D(\varepsilon_t) = \sigma^2 > 0$, $E(\varepsilon_t^4) = \mu_4 < \infty$.

При выполнении этих предположений обычные t-статистики и статистики qF (где q –количество линейных ограничений на коэффициенты, а F – обычная F-статистика критерия для проверки выполнения этих ограничений) имеют асимптотические $N(0, 1)$ и $\chi^2(q)$ распределения. Можно использовать стандартную технику регрессионного анализа, имея в виду ее асимптотическую обоснованность.

Если не ограничиваться процессами авторегрессии, но оставаться в классе стационарных моделей, то и в этом случае все еще можно надеяться на возможность использования стандартной техники регрессионного анализа, опять имея в виду ее асимптотическую обоснованность.

Рассмотрим линейную модель

$$y = X\theta + \varepsilon \,, \qquad X = X_n \,,$$

или, в эквивалентной форме,

$$y_t = x_t^T \theta + \varepsilon_t \,, \quad t = 1, 2, \ldots, n \,,$$

где

$$x_t = (x_{t1}, x_{t2}, \ldots, x_{tp})^T \; -$$

вектор значений p объясняющих переменных в t-м наблюдении, и пусть $\hat{\theta}_n$ – оценка наименьших квадратов вектора коэффициентов $\theta = (\theta_1, \theta_2, \ldots, \theta_p)^T$, полученная по n наблюдениям. Известно (см., например, [Green (1997)]), что следующие три условия обеспечивают состоятельность и асимптотическую нормальность $\hat{\theta}_n$ при $n \to \infty$:

- $$\operatorname{plim}\left(\frac{1}{n}\sum_{t=1}^{n} x_t \varepsilon_t\right) = 0$$

  [в эквивалентной форме: $\operatorname{plim}\,(n^{-1} X_n^T \varepsilon) = 0$];
- $$\operatorname{plim}\left(\frac{1}{n}\sum_{t=1}^{n} x_t x_t^T\right) = Q$$

  [в эквивалентной форме: $\operatorname{plim}\left(n^{-1} X_n^T X\right) = Q$], где Q – положительно определенная матрица;
- $$\left(\frac{1}{\sqrt{n}}\sum_{t=1}^{n} x_t \varepsilon_t\right) \to N\left(0, \sigma^2 Q\right)$$

  [в эквивалентной форме: $(n^{-1/2} X_n^T \varepsilon) \to N(0, \sigma^2 Q)$].

(Здесь plim – предел по вероятности; стрелка в последнем условии обозначает сходимость по распределению.) Если эти условия выполнены, то при $n \to \infty$, как и в ситуации D,

$$\sqrt{n}\,(\hat{\theta}_n - \theta) \to N(0, \sigma^2 Q^{-1}).$$

В работе [Mann, Wald (1943)] авторы показали следующее ***(теорема Манна-Вальда)***. Если

- $\text{plim}\left(\frac{1}{n}\sum_{t=1}^{n} x_t x_t^T\right) = Q$

[в эквивалентной форме: $\text{plim}\ (n^{-1} X_n^T X) = Q$] , где Q – положительно определенная матрица,

- $\varepsilon_t \sim i.i.d.$, $E(\varepsilon_t) = 0$, $D(\varepsilon_t) = \sigma^2 > 0$, $E|\varepsilon_t|^m < \infty$ для всех $m = 1, 2, \ldots$,
- $E(x_t\varepsilon_t) = 0$, $t = 1, 2, \ldots, n$,

то тогда выполнены также первое и третье условия из предыдущей тройки условий, обеспечивающих состоятельность и асимптотическую нормальность $\hat{\theta}_n$ при $n \to \infty$.

Заметим, что условие $E(x_t\varepsilon_t) = 0$, $t = 1, 2, \ldots, n$, в сочетании с $E(\varepsilon_t) = 0$, означает, что

$Cov(x_{tk}, \varepsilon_t) = 0$ для $k = 1, 2, \ldots, p$,

т.е. означает некоррелированность значений объясняющих переменных с ε_t в совпадающие моменты времени. Условие $E|\varepsilon_t|^m < \infty$ для всех $m = 1, 2, \ldots$, выполняется, в частности, для нормального распределения ε_t.

Цитированные результаты можно объединить теперь вместе.

Ситуация F

Пусть в линейной модели

$y_t = x_t^T\theta + \varepsilon_t$, $t = 1, 2, \ldots, n$,

где

$x_t = (x_{t1}, x_{t2}, \ldots, x_{tK})^T$ –

вектор значений K объясняющих переменных в t-м наблюдении, $\hat{\theta}_n$ – оценка наименьших квадратов вектора коэффициентов $\theta = (\theta_1, \theta_2, \ldots, \theta_K)$, полученная по n наблюдениям. Пусть для этой модели выполнены следующие условия:

- $\text{plim}_{n\to\infty}\left(\frac{1}{n}\sum_{t=1}^{n} x_t x_t^T\right) = Q \quad \left(\text{т.е. } \text{plim}_{n\to\infty}\frac{1}{n}X_n^T X_n = Q\right)$,

  где Q – положительно определенная матрица,
- $\varepsilon_t \sim i.i.d.$, $E(\varepsilon_t) = 0$, $D(\varepsilon_t) = \sigma^2 > 0$, $E|\varepsilon_t|^m < \infty$ для всех $m = 1, 2, \ldots$,
- $Cov(x_{tk}, \varepsilon_t) = 0$ для $k = 1, 2, \ldots, K$.

Тогда при $n \to \infty$

$\sqrt{n}\ (\hat{\theta}_n - \theta) \to N(0, \sigma^2 Q^{-1})$.

Предположим теперь, что x_t – ***стационарный векторный (K-мерный) ряд***, так что

$E(x_t) = \mu = const$, $Cov(x_t) = Q$, $Cov(x_{tk}, x_{t+s,\,l}) = \gamma_{kl}(s)$

при всех t, s для каждой пары $k, l = 1, 2, \ldots, K$. (Здесь $\gamma_{kl}(s)$ – ***кросс-корреляция*** значений k-ой и l-ой компонент векторного ряда x_t, разнесенных на s единиц времени. Если рассматривать s как аргумент, а $\gamma_{kl}(s)$ как функцию от s , то $\gamma_{kl}(s)$ – ***кросс-корреляционная функция*** k-ой и l-ой компонент векторного ряда x_t .) Тогда первое из трех условий, перечисленных в ситуации F, обеспечивает возможность оценивания неизвестной ковариационной матрицы $Cov(x_t) = Q$ простым усреднением доступных наблюдению матриц $x_t x_t^T$ по достаточно длинному интервалу $t = 1, 2, \ldots, n$.

В рамки ситуации F помещается достаточно распространенный класс ***ARX моделей*** :

$$y_t = a_1 y_{t-1} + a_2 y_{t-2} + \ldots + a_p y_{t-p} + z_t^T \beta + \varepsilon_t ,$$

где

$z_t = (z_{t1}, z_{t2}, \ldots, z_{tM})^T$, $\beta = (\beta_1 , \beta_2 , \ldots , \beta_M)^T$.

Подобная модель вписывается в ситуацию F , если положить

$x_t = (y_{t-1}, y_{t-2} , \ldots, y_{t-p}, z_1 , z_2 , \ldots , z_M)^T$,

$\theta = (a_1 , a_2 , \ldots , a_p , \beta_1 , \beta_2 , \ldots , \beta_M)^T$.

Пусть для этой модели выполнены следующие условия:

- z_t – стационарный векторный (M-мерный) ряд;
- $\operatorname*{plim}_{n\to\infty}\left(\frac{1}{n}\sum_{t=1}^{n} z_t z_t^T\right) = Q_Z$,

  где Q_Z – положительно определенная матрица;
- $\varepsilon_t \sim i.i.d., E(\varepsilon_t) = 0, D(\varepsilon_t) = \sigma^2 > 0, E|\varepsilon_t|^m < \infty$ для всех $m = 1, 2, \ldots$;
- $Cov(z_{tm} , \varepsilon_t) = 0$ для $m = 1, 2, \ldots, M$;
- $Cov(y_{t-j} , \varepsilon_t) = 0$ для $j = 1, 2, \ldots, p$;
- все корни уравнения $a(z) = 1 - a_1 z - a_2 z^2 - \ldots - a_p z^p = 0$ лежат вне единичного круга.

Тогда (см. [Green (1993)]) выполнено и первое условие ситуации F, и при $n \to \infty$

$n^{1/2} (\hat{\theta}_n - \theta) \to N(0, \sigma^2 Q^{-1})$.

Последнее из перечисленных условий, касающееся корней уравнения $a(z) = 0$, обеспечивает ***стабильность*** модели ARX. Последнее означает, что по мере продвижения в будущее (т.е. с ростом t) устанавливается определенная ***"долговременная" (long-run)*** связь между переменными y_t , z_{t1}, z_{t2}, …, z_{tM} , по отношению к которой происходят достаточно быстрые осцилляции.

4.2. Динамические модели

Среди различных ARX моделей, в эконометрических исследованиях широкое применение нашли ***динамические модели (модели с авторегрессионно распределенными запаздываниями – ADL)***

$$\begin{aligned} y_t = \alpha_0 + a_1 y_{t-1} + a_2 y_{t-2} + \ldots + a_p y_{t-p} + \\ + (\beta_{10} x_{1,t} + \beta_{11} x_{1,t-1} + \ldots + \beta_{1r} x_{1,t-r}) + \\ + \ldots + \\ + (\beta_{s0} x_{s,t} + \beta_{s1} x_{s,t-1} + \ldots + \beta_{sr} x_{s,t-r}) + \varepsilon_t . \end{aligned}$$

Для такой модели используют обозначение ***ADL(p,r; s)***, где p – глубина запаздываний по переменной y_t , r – глубина запаздываний по переменным $x_{1,t}$, $x_{2,t}$, …, $x_{s,t}$, не являющимся запаздываниями переменной y_t , s – количество таких переменных. При такой форме записи допускается, что некоторые из коэффициентов

β_{ij} равны нулю, так что глубина запаздываний может быть различной для различных переменных $x_{i,\,t}$.

Модель ADL(p,r; s) можно представить в компактном виде

$a(L)\, y_t = \mu + b_1(L)\, x_{1,\,t} + \ldots + b_s(L)\, x_{s,\,t} + \varepsilon_t$,

где

$a(L) = 1 - a_1 L - a_2 L^2 - \ldots - a_p L^p$,

$b_i(L) = \beta_{i\,0} + \beta_{i\,1} L + \ldots + \beta_{i\,r} L^r$, $i = 1, \ldots , s$.

Если выполнено условие стабильности, то y_t представляется в виде

$$y_t = \frac{1}{a(L)}\mu + \frac{1}{a(L)} b_1(L) x_{1,t} + \ldots + \frac{1}{a(L)} b_s(L) x_{s,t} + \frac{1}{a(L)}\varepsilon_t ,$$

или

$$y_t = \frac{1}{a(L)}\mu + c_1(L) x_{1,t} + \ldots + c_s(L) x_{s,t} + \frac{1}{a(L)}\varepsilon_t ,$$

где

$$c_i(L) = \frac{b_i(L)}{a(L)} .$$

Долговременную связь между переменными можно найти, полагая в выражении для y_t

$L = 1$, $\varepsilon_t \equiv 0$.

При этом получаем

$$y_t = \frac{1}{a(1)}\mu + c_1(1) x_{1,t} + \ldots + c_s(1) x_{s,t} ;$$

строго говоря, в последнем выражении указание на момент t следует исключить:

$$y = \frac{1}{a(1)}\mu + c_1(1) x_1 + \ldots + c_s(1) x_s .$$

Коэффициенты $c_1(1)$, … , $c_s(1)$ в последнем сотношении называются ***долгосрочными мультипликаторами (long-run multipliers)***. Поясним это название на примере модели ADL(1, 1; 1), которую запишем в виде

$(1 - \alpha_1 L)\, y_t = \mu + \beta_0\, x_t + \beta_1\, x_{t-1} + \varepsilon_t$.

При $|\alpha_1| < 1$ получаем равносильное представление

$$y_t = \frac{1}{(1-\alpha_1 L)}\mu + \frac{1}{(1-\alpha_1 L)}(\beta_0 x_t + \beta_1 x_{t-1} + \varepsilon_t),$$

т.е.

$y_t = (1 + \alpha_1 + \alpha_1^2 + \ldots)\, \mu + (1 + \alpha_1 L + \alpha_1^2 L^2 + \ldots)(\beta_0\, x_t + \beta_1\, x_{t-1} + \varepsilon_t)$,

из которого последовательно находим:

$\partial y_t / \partial x_t = \beta_0$,

$\partial y_{t+1} / \partial x_t = \partial y_t / \partial x_{t-1} = \beta_1 + \alpha_1 \beta_0$,

$\partial y_{t+2} / \partial x_t = \partial y_t / \partial x_{t-2} = \alpha_1 \beta_1 + \alpha_1^2 \beta_0$,

$\partial y_{t+3} / \partial x_t = \partial y_t / \partial x_{t-3} = \alpha_1^2 \beta_1 + \alpha_1^3 \beta_0$, и т.д. Правые части дают значения ***импульсных мультипликаторов***, показывающих влияние единовременного (импульсного) изменения значения x_t на текущее и последующие значения переменной y_t . Просуммировав полученные выражения, получаем:

$$\partial y_t / \partial x_t + \partial y_t / \partial x_{t-1} + \partial y_t / \partial x_{t-2} + \partial y_t / \partial x_{t-3} + \ldots =$$

$$= \beta_0 (1 + \alpha_1 + \alpha_1^2 + \ldots) + \beta_1(1 + \alpha_1 + \alpha_1^2 + \ldots) =$$

$$= (1 - \alpha_1 L)^{-1}(\beta_0 + \beta_1).$$

Правая часть этого соотношения, как легко заметить, представляет собой долгосрочный мультипликатор рассматриваемой ADL(1, 1; 1). В соответствии с левой частью, этот мультипликатор показывает изменение значения y_t при изменении на единицу текущего и всех предыдущих значений переменной x_t .

Прежде, чем перейти к рассмотрению примера оценивания конкретной ADL модели, следует заметить следующее.

При выполнении условий, обеспечивающих возможность использования стандартной техники регрессионного анализа (имеется в виду ее асимптотическая обоснованность – см. разд. 4.1, ситуация F):

- Обычная t-статистика имеет асимптотическое $N(0,1)$ распределение.
- Если F – обычная F-статистика для проверки гипотезы о выполнении q линейных ограничений на коэффициенты модели, то статистика qF имеет асимптотическое χ^2 распределение с q степенями свободы.
- При умеренном количестве наблюдений параллельно с асимптотическими распределениями для t и qF можно для контроля использовать и точные (стандартные) распределения (распределение Стьюдента для t-статистики, распределение Фишера для F-статистики). Согласованность получаемых при этом результатов подкрепляет уверенность в правильности соответствующих статистических выводов.
- При наличии в правой части запаздывающих значений объясняемой переменной проверку гипотезы об отсутствии автокоррелированности у ряда ε_t следует производить, используя ***критерий Бройша – Годфри***. Критерий Дарбина – Уотсона не годится для этой цели, поскольку в данном случае значения

статистики Дарбина – Уотсона d смещены в направлении значения $d = 2$, так что использование таблиц Дарбина – Уотсона приводит к неоправданно частому неотвержению указанной гипотезы ("презумпция некоррелированности ε_t").

Пример

Рассмотрим модель ADL(3, 2; 1)

$$(1 - 0.5L - 0.1L^2 - 0.05L^3)\, y_t = 0.7 + (0.2 + 0.1\, L + 0.05L^2)\, x_t + \varepsilon_t\, .$$

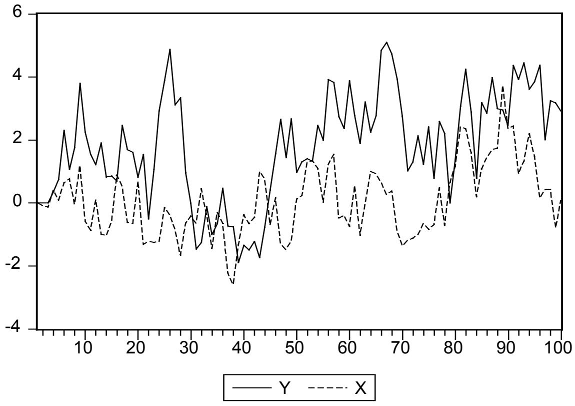
Для нахождения долговременной связи между переменными y и x полагаем $L = 1$ и $\varepsilon_t \equiv 0$:

$$(1 - 0.5 - 0.1 - 0.05)\, y = 0.7 + (0.2 + 0.1 + 0.05)\, x\, ,$$

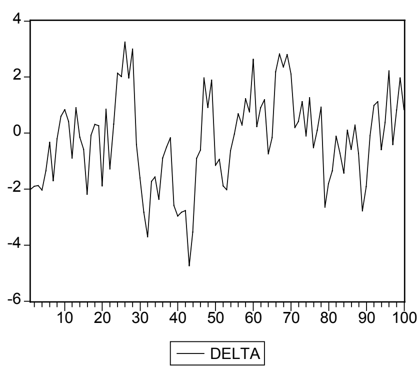
т.е. $0.35\, y = 0.7 + 0.35\, x$, или

$$y = 2 + x\, .$$

На приводимом ниже графике представлены смоделированная реализация ряда $x_t = 0.7x_{t-1} + \varepsilon_{xt}$, $\varepsilon_{xt} \sim i.i.d.\ N(0, 1)$, и соответствующая ей реализация ряда y_t , порождаемого указанной моделью ADL(3, 2; 1), где $\varepsilon_t \sim i.i.d.\ N(0, 1)$, причем ряд ε_t порождается независимо от ряда ε_{xt} . В качестве начальных значений при моделировании были взяты: $x_1 = 0, y_1 = y_2 = y_3 = 0$.



Имея в распоряжении только эти две реализации, мы не знаем, с какой моделью имеем дело. Начнем с оценивания статической модели $y_t = \mu + \beta x_t + \varepsilon_t$ методом наименьших квадратов; в результате получаем оцененную модель $y_t = 1.789 + 0.577x_t + e_t$, где e_t – ряд остатков. График ряда остатков имеет вид:



Здесь обнаруживается явная автокоррелированность ряда остатков, которая подтверждается построенной для него коррелограммой

ACF PAC
F C PAC Q-Stat Prob

.	.		(	4	
|*****	|*****	.696	.696	9.981	.000
.	.		(	7	
|****	|*	.536	.099	9.868	.000
.	*|		-	9	
|***	.	.364	0.081	3.801	.000
.	. |		-	9	
|**	.	.227	0.056	9.279	.000
.	. |		-	1	
|*	.	.130	0.015	01.10	.000
. |	. |		-	1	
.	.	.057	0.020	01.46	.000

и критерием Бройша – Годфри с запаздыванием на один шаг, который дает *P*-значение 0.0000. Это означает, что мы имеем дело не со статической, а с динамической моделью. Поэтому следует прежде всего рассмотреть характер поведения обоих рядов и произвести их идентификацию.

Для ряда x_t коррелограмма имеет вид

ACF	PACF	AC	PAC	Q-Stat	Prob
. |*****	. |*****	0.686	0.686	48.468	0.000
. |***	*| .	0.429	-0.079	67.594	0.000
. |*	*| .	0.193	-0.132	71.527	0.000
. | .	*| .	0.024	-0.066	71.591	0.000
*| .	| .	-0.058	0.003	71.958	0.000
*| .	*| .	-0.140	-0.107	74.090	0.000

По этой коррелограмме ряд x_t идентифицируется как AR(1).

Для ряда y_t коррелограмма имеет вид

ACF	PACF	AC	PAC	Q-Stat	Prob
.	. |******	0.767	0.767	60.58	0.000
|******					
.	. |*	0.629	0.100	101.75	0.000
|*****					
	. | .	0.494	-0.042	127.37	0.000
|****					
.	. | .	0.399	0.019	144.32	0.000
|***					
.	. | .	0.318	-0.003	155.21	0.000
|**					
.	. | .	0.257	0.004	162.38	0.000
|**					

так что и этот ряд идентифицируется как AR(1).

Такой предварительный анализ ограничивает рассмотрение моделью ADL с глубиной запаздываний, равной единице:

$$y_t = \mu + a_1 y_{t-1} + \beta_0 x_t + \beta_1 x_{t-1} + \varepsilon_t .$$

Оценивая такую модель ADL(1, 1; 1), получаем:

Dependent Variable: Y

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.558588	0.157276	3.55163	0.0006
Y(-1)	0.695204	0.066095	10.51828	0.0000
X	0.208971	0.126135	1.65673	0.1009
X(-1)	0.161690	0.132352	1.22166	0.2249

Анализ остатков не выявляет автокоррелированности (P-значение критерия Бройша–Годфри при AR(1) альтернативе равно 0.164), не выявляет значимых отклонений от нормальности распределения ε_t (P-значение критерия Jarque – Bera = 0.267), не обнаруживает гетероскедастичности (P-значение критерия Уайта = 0.159), так что можно, опираясь на приведенные выше факты, использовать асимптотическую теорию статистических выводов и на ее основе использовать результаты, получаемые при применении t- и F-критериев.

При проверке гипотезы H_0: $\beta_0 = \beta_1 = 0$ получаем при использовании F-распределения P-значение 0.0032; при использовании асимптотического распределения $\chi^2(2)$ получаем P-значение 0.0022. В обоих случаях эта гипотеза отвергается. Исключение из правой части модели запаздывающей переменной x_{t-1}, коэффициент при которой статистически незначим и имеет большее P-значение, чем коэффициент при x_t, дает:

Dependent Variable: Y

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.517868	0.154098	3.360648	0.0011
Y(-1)	0.719738	0.063131	11.40064	0.0000
X	0.310343	0.095241	3.258511	0.0015

R-squared	0.637523	Mean dependent var	1.844751
Adjusted R-squared	0.629971	S.D. dependent var	1.709520
S.E. of regression	1.039901	Akaike info criterion	2.945962
Sum squared resid	103.8138	Schwarz criterion	3.024602
Log likelihood	-142.8251	F-statistic	84.42207
Durbin-Watson stat	2.256404	Prob(F-statistic)	0.000000

Здесь все коэффициенты имеют высокую значимость, а остатки вполне удовлетворительны.

Если из предыдущей модели исключить не x_{t-1}, а x_t, то это приводит к оцененной модели

Dependent Variable: Y

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.567821	0.158600	3.580215	0.0005
Y(-1)	0.692523	0.066673	10.38690	0.0000
X(-1)	0.305939	0.100582	3.041698	0.0030

R-squared	0.632818	Mean dependent var	1.844751
Adjusted R-squared	0.625169	S.D. dependent var	1.709520
S.E. of regression	1.046627	Akaike info criterion	2.958857
Sum squared resid	105.1611	Schwarz criterion	3.037497
Log likelihood	-143.4634	F-statistic	82.72550
Durbin-Watson stat	2.221594	Prob(F-statistic)	0.000000

По критерию Шварца чуть более предпочтительной выглядит модель с исключенной x_{t-1} , так что на ней можно и остановиться. Посмотрим, к какому долговременному соотношению приводит такая модель.

Итак, мы останавливаемся на оцененной модели

$(1 - 0.720\, L)\, y_t = 0.518 + 0.310\, x_t + e_t$.

Полагая $L = 1$ и $e_t \equiv 0$, получаем: $0.28\, y_t = 0.518 + 0.310\, x_t$, так что долговременное соотношение оценивается как

$y = 1.839 + 1.107x$.

Это соотношение, конечно, несколько отличается от теоретического. Посмотрим, однако, что дало бы оценивание динамической модели ADL(3, 2; 1). Оценивая такую модель, получаем:

Dependent Variable: Y

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.539617	0.174057	3.100239	0.0026
Y(-1)	0.590293	0.105583	5.590795	0.0000
Y(-2)	0.153936	0.120517	1.277303	0.2048
Y(-3)	-0.031297	0.099814	-0.313555	0.7546
X	0.205570	0.129483	1.587626	0.1159
X(-1)	0.191959	0.153608	1.249666	0.2147
X(-2)	-0.024779	0.138389	-0.179053	0.8583

R-squared	0.643818	Mean dependent var	1.882787
Adjusted R-squared	0.620072	S.D. dependent var	1.706160
S.E. of regression	1.051648	Akaike info criterion	3.008022
Sum squared resid	99.53667	Schwarz criterion	3.193826
Log likelihood	-138.8891	F-statistic	27.11327
Durbin-Watson stat	1.999213	Prob(F-statistic)	0.000000

Если найти долговременное соотношение между y и x на основе такого оцененного уравнения по той же схеме, что и прежде, то получаем:

$y = 1.882 + 1.300\, x$,

и это соотношение отнюдь не ближе к теоретическому, чем то, которое мы получили по редуцированному уравнению. Впрочем, и по критерию Шварца полная оцененная модель хуже редуцированной.

4.3. Векторная авторегрессия

Полулярной моделью связи между временными рядами является ***векторная авторегрессия (VAR – vector autoregression)***.

В своей простейшей форме такая модель связывает два ряда y_{1t} и y_{2t} следующим образом:

$$y_{1t} = \mu_1 + \pi_{11.1}\, y_{1,\, t-1} + \pi_{12.1}\, y_{2,\, t-1} + \varepsilon_{1t}\, ,$$

$$y_{2t} = \mu_2 + \pi_{21.1}\, y_{1,\, t-1} + \pi_{22.1}\, y_{2,\, t-1} + \varepsilon_{2t}\, ,$$

т.е. , в отличие от простого процесса авторегрессии, значение y_{1t} связывается не только с запаздыванием $y_{1,t-1}$, но и с запаздыванием $y_{2,t-1}$ второй переменной y_{2t} . Случайные величины ε_{1t} и ε_{2t} являются ***инновациями*** :

- $Cov(\varepsilon_{jt}\, , \varepsilon_{ls}) = 0$ для $t \ne s$ при любых $j\, , l = 1, 2$;
- $Cov(\varepsilon_{jt}\, , y_{l,\, t-r}) = 0$ для $r \ge 1$ при любых $j\, , l = 1, 2$.

В то же время, для совпадающих моментов времени случайные величины ε_{1t} и ε_{2t} могут быть коррелированными.

Модель векторной авторегрессии для двух рядов допускает включение в правые части уравнений для y_{1t} и y_{2t} и большего количества запаздываний этих переменных. Наибольший порядок запаздываний, включаемых в правую часть, называется ***порядком*** векторной авторегрессии. Если этот порядок равен p , то для такой модели используют обозначение ***VAR(p)***.

В общем случае рассматривается k временных рядов $y_{1t}, y_{2t}, \ldots, y_{kt}$. Модель векторной авторегрессии порядка p предполагает, что связь между этими рядами имеет вид

$$\begin{aligned} y_{1t} = \mu_1 &+ \pi_{11.1}\, y_{1,\,t-1} + \pi_{11.2}\, y_{1,\,t-2} + \ldots + \pi_{11.p}\, y_{1,\,t-p} + \\ &+ \pi_{12.1}\, y_{2,\,t-1} + \pi_{12.2}\, y_{2,\,t-2} + \ldots + \pi_{12.p}\, y_{2,\,t-p} + \\ &+ \ldots + \\ &+ \pi_{1k.1}\, y_{k,\,t-1} + \pi_{1k.2}\, y_{k,\,t-2} + \ldots + \pi_{1k.p}\, y_{k,\,t-p} + \varepsilon_{1t}\,, \end{aligned}$$

$$\begin{aligned} y_{2t} = \mu_2 &+ \pi_{21.1}\, y_{1,\,t-1} + \pi_{21.2}\, y_{1,\,t-2} + \ldots + \pi_{21.p}\, y_{1,\,t-p} + \\ &+ \pi_{22.1}\, y_{2,\,t-1} + \pi_{22.2}\, y_{2,\,t-2} + \ldots + \pi_{22.p}\, y_{2,\,t-p} + \\ &+ \ldots + \\ &+ \pi_{2k.1}\, y_{k,\,t-1} + \pi_{2k.2}\, y_{k,\,t-2} + \ldots + \pi_{2k.p}\, y_{k,\,t-p} + \varepsilon_{2t}\,, \\ &\ \ldots \end{aligned}$$

$$\begin{aligned} y_{kt} = \mu_k &+ \pi_{k1.1}\, y_{1,\,t-1} + \pi_{k1.2}\, y_{1,\,t-2} + \ldots + \pi_{k1.p}\, y_{1,\,t-p} + \\ &+ \pi_{k2.1}\, y_{2,\,t-1} + \pi_{k2.2}\, y_{2,\,t-2} + \ldots + \pi_{k2.p}\, y_{2,\,t-p} + \\ &+ \ldots + \\ &+ \pi_{kk.1}\, y_{k,\,t-1} + \pi_{kk.2}\, y_{k,\,t-2} + \ldots + \pi_{kk.p}\, y_{k,\,t-p} + \varepsilon_{kt}\,, \end{aligned}$$

где $\pi_{ij.r}$ - коэффициент при $y_{j,\,t-r}$ в уравнении для $y_{i\,t}$.

Здесь $\varepsilon_{1t}, \varepsilon_{2t}, \ldots, \varepsilon_{kt}$ – случайные величины, для которых

- $Cov(\varepsilon_{jt}, \varepsilon_{ls}) = 0$ для $t \neq s$ при любых $j, l = 1, \ldots, k$;
- $Cov(\varepsilon_{jt}, y_{l,\,t-r}) = 0$ для $r \geq 1$ при любых $j, l = 1, \ldots, k$;
- $Cov(\varepsilon_{jt}, \varepsilon_{lt})$ могут отличаться от нуля.

Случайные величины $\varepsilon_{1t}, \varepsilon_{2t}, \ldots, \varepsilon_{kt}$ образуют случайный вектор $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t}, \ldots, \varepsilon_{kt})^T$, компоненты которого некоррелированы по времени и не коррелированы с запаздывающими значениями переменных $y_{1t}, y_{2t}, \ldots, y_{kt}$. Этот вектор называют ***вектором инноваций (обновлений) относительно информационного множества***

$$\boldsymbol{Y}_{t-1} = (\, y_{1,\,t-1}, y_{1,\,t-2}, \ldots, y_{1,\,t-p}, \ldots, y_{k,\,t-1}, y_{k,\,t-2}, \ldots, y_{k,\,t-p}) \,.$$

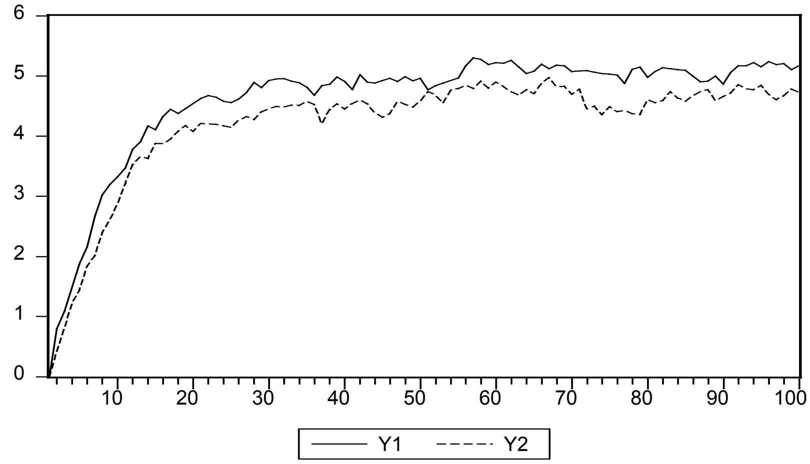
Пример

Рассмотрим следующую модель VAR(1) для двух рядов ($k = 2, p = 1$):

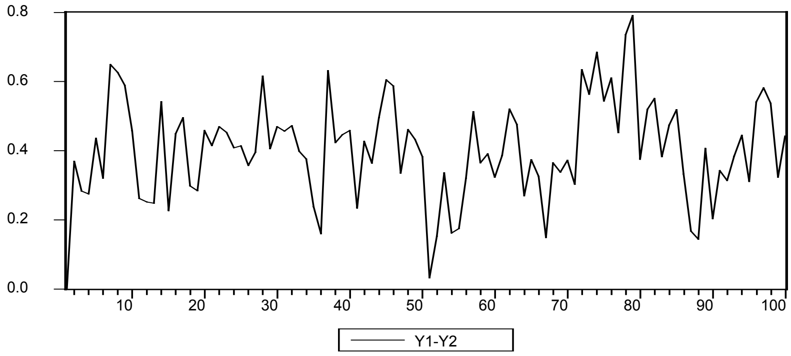
$y_{1t} = 0.6 + 0.7\, y_{1,\,t-1} + 0.2\, y_{2,\,t-1} + \varepsilon_{1t}$,

$y_{2t} = 0.4 + 0.2\, y_{1,\,t-1} + 0.7\, y_{2,\,t-1} + \varepsilon_{2t}$.

Приводимый ниже график иллюстрирует поведение смоделированной пары y_{1t}, y_{2t} порождаемой этой моделью для t = 2, 3, …, 100. В качестве начальных значений были взяты $y_{11} = y_{21} = 0$; ε_{1t} и ε_{2t} моделировались как независимые случайные величины, имеющие одинаковое нормальное распределение $N(0, 0.1^2)$.



Следующий график представляет поведение разности $(y_{1t} - y_{2t})$.



Мы видим, что с течением времени поведение рядов стабилизируется: они осциллируют вокруг установившихся уровней. Второй график показывает, что установившийся уровень для ряда y_{1t} превышает установившийся уровень для ряда y_{2t} приблизительно на 0.4 (среднее арифметическое разности $y_{1t} - y_{2t}$ равно 0.403). Такой характер поведения пары y_{1t}, y_{2t} указывает на стабильность данной модели VAR .

Предсказать стабильный характер поведения реализаций рядов, связанных VAR моделью, можно, анализируя коэффициенты модели. Для этого удобно записать VAR(p) модель для k рядов в более компактной форме

$$y_t = \mu + \Pi_1 y_{t-1} + \Pi_2 y_{t-2} + \ldots + \Pi_p y_{t-p} + \varepsilon_t .$$

Здесь

$$y_t = (y_{1t}, y_{2t}, \ldots, y_{kt})^T, \ \mu = (\mu_1, \mu_2, \ldots, \mu_k)^T, \ \varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t}, \ldots, \varepsilon_{kt})^T ,$$

$\Pi_r = (\pi_{ij.r})$ – матрица размера $k \times k$ коэффициентов при $y_{1,\,t-r}, y_{2,\,t-r}, \ldots, y_{k,\,t-r}$ в k уравнениях.

Последнее представление можно записать как

$$y_t - \Pi_1 y_{t-1} - \Pi_2 y_{t-2} - \ldots - \Pi_p y_{t-p} = \mu + \varepsilon_t ,$$

$$(I_k - \Pi_1 L - \Pi_2 L^2 - \ldots - \Pi_p L^p)\, y_t = \mu + \varepsilon_t ,$$

или

$$A(L)\, y_t = \mu + \varepsilon_t ,$$

где

$$A(L) = I_k - \Pi_1 L - \Pi_2 L^2 - \ldots - \Pi_p L^p .$$

Условие стабильности такой VAR модели состоит в следующем:

- Все k корней уравнения

$$\det(I_k - z\,\Pi_1 - z^2\,\Pi_2 - \ldots - z^p\,\Pi_p) = 0 \qquad (\text{т.е. } \det A(z) = 0)$$

лежат <u>за пределами единичного круга</u> на комплексной плоскости (т.е. модули всех k корней <u>больше единицы</u>).

Если это условие выполняется, то при продвижении вперед по оси времени система постепенно "забывает" о том, при каких начальных значениях $y_1, y_2, \ldots, y_p$ она начала реализовываться. Стабильное состояние системы находится путем приравнивания $L = 1$ и $\varepsilon_t = 0$. При этом получаем

$$A(1)\,y_t = \mu\,,$$

так что стабильное состояние определяется как

$$y_t = A^{-1}(1)\,\mu\,.$$

Пример

Продолжим рассмотрение модели VAR(1) для двух рядов

$$y_{1t} = 0.6 + 0.7\,y_{1,\,t-1} + 0.2\,y_{2,\,t-1} + \varepsilon_{1t}\,,$$

$$y_{2t} = 0.4 + 0.2\,y_{1,\,t-1} + 0.7\,y_{2,\,t-1} + \varepsilon_{2t}\,.$$

В компактной форме эта система имеет вид

$$y_t = \mu + \Pi_1\,y_{t-1} + \varepsilon_t\,,$$

где

$$y_t = \begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix},\quad \mu = \begin{pmatrix} 0.6 \\ 0.4 \end{pmatrix},\quad \Pi_1 = \begin{pmatrix} 0.7 & 0.2 \\ 0.2 & 0.7 \end{pmatrix},\quad \varepsilon_t = \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix},$$

или

$$A(L)\,y_t = \mu + \varepsilon_t\,,$$

где

$$A(L) = I_2 - \Pi_1 L = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 0.7L & 0.2L \\ 0.2L & 0.7L \end{pmatrix} = \begin{pmatrix} 1-0.7L & -0.2L \\ -0.2L & 1-0.7L \end{pmatrix},$$

так что

$$A(1) = \begin{pmatrix} 0.3 & -0.2 \\ -0.2 & 0.3 \end{pmatrix},\quad A^{-1}(1) = \begin{pmatrix} 6 & 4 \\ 4 & 6 \end{pmatrix}.$$

Уравнение $\det A(z) = 0$ принимает здесь вид

$$\det A(z) = \begin{pmatrix} 1-0.7z & -0.2\,z \\ -0.2\,z & 1-0.7z \end{pmatrix} = 0\,,$$

т.е. $(1 - 0.7\,z)^2 - (0.2\,z)^2 = 0$, или $(1 - 0.9\,z)(1 - 0.5\,z) = 0$. Оба корня $z = 1/0.9$ и $z = 1/0.5$ больше 1, т.е. условие стабильности выполняется.

Долгосрочное (стабильное) поведение системы находим по формуле

$$y_t = A^{-1}(1)\mu = \begin{pmatrix} 6 & 4 \\ 4 & 6 \end{pmatrix}\begin{pmatrix} 0.6 \\ 0.4 \end{pmatrix} = \begin{pmatrix} 5.2 \\ 4.8 \end{pmatrix}.$$

Таким образом, стабильное состояние системы определяется здесь как

$y_{1t} = 5.2$, $y_{2t} = 4.8$,

так что стабильное состояние разности $y_{1t} - y_{2t}$ есть

$y_1 - y_2 = 0.4$.

Соответственно, с течением времени, независимо от начальных условий, ряд y_{1t} начинает осциллировать вокруг уровня 5.2, а ряд y_{2t} начинает осциллировать вокруг уровня 4.8; разность $(y_{1t} - y_{2t})$ осциллирует вокруг уровня 0.4 . Именно такое поведение смоделированных реализаций рассматриваемой VAR(1) мы и наблюдали ранее.

Векторные авторегрессии, определенные так, как было указано выше, называют также ***замкнутыми VAR***, отличая тем самым эти модели от ***открытых VAR***, в правые части которых наряду с запаздывающими значениями переменных, находящихся в левых частях уравнений ***(эндогенные переменные)***, входят и некоторые другие переменные и их запаздывания ***(экзогенные переменные)***. Проводя различие между эндогенными и экзогенными переменными, по-существу предполагают, что значения экзогенных переменных формируются вне рассматриваемой системы , а значения эндогенных переменных порождаются в рамках этой системы . Фактически, система в этом случае рассматривается как условная по отношению к экзогенным переменным. Заметим, что в замкнутой VAR экзогенные переменные отсутствуют.

Открытую VAR можно представить в виде

$$A(L)\, y_t = \mu + B(L)\, x_t + \varepsilon_t\,,$$

где

$$A(L) = \mathrm{I} - \Pi_1 L - \Pi_2 L^2 - \ldots - \Pi_p L^p$$

и $B(L)$ – матричные полиномы.

Если все решения уравнения $\det A(z) = 0$ лежат за пределами единичного круга на комплексной плоскости, что необходимо для обеспечения стабильности системы, то тогда справедливо также представление

$$y_t = A^{-1}(L)\, \mu + C(L)\, x_t + A^{-1}(L)\, \varepsilon_t\,,$$

где $C(L) = A^{-1}(L)B(L)$ – ***передаточная функция (transfer function)***. Функция $C(L)$ – матричная функция; она устанавливает влияние единичных изменений в экзогенных переменных на эндогенные переменные.

Долговременную (долгосрочную, стабильную, long-run) связь между экзогенными и эндогенными переменными можно найти, полагая в последнем представлении $L = 1$ и $\varepsilon_t \equiv 0$. При этом получаем:

$$y_t = A^{-1}(1)\, \mu + C(1)\, x_t\,.$$

Матрица $C(1)$ называется ***матрицей долгосрочных мультипликаторов.*** Ее (i , j)-й элемент $c_{ij}(1)$ представляет влияние единичного изменения x_{jt} на y_{it} в долговременном плане (см. интерпретацию *долгосрочных мультипликаторов* в разд. 4.2).

Пример

На базе рассмотренной выше замкнутой модели VAR(1) для двух рядов построим открытую VAR

$$y_{1t} = 0.6 + 0.7\, y_{1,\, t-1} + 0.2\, y_{2,\, t-1} + 0.1\, x_{1,\, t-1} + 0.2\, x_{2,\, t} + \varepsilon_{1t}\,,$$

$$y_{2t} = 0.4 + 0.2\, y_{1,\, t-1} + 0.7\, y_{2,\, t-1} + 0.2\, x_{1,\, t} + 0.4\, x_{2,\, t-1} + \varepsilon_{2t}\,.$$

Здесь μ и матричный полином $A(L)$ – те же, что и ранее, а

$$B(L) = B_0 + B_1 L = \begin{pmatrix} 0 & 0.2 \\ 0.2 & 0 \end{pmatrix} + \begin{pmatrix} 0.1 & 0 \\ 0 & 0.4 \end{pmatrix} L = \begin{pmatrix} 0.1\,L & 0.2 \\ 0.2 & 0.4\,L \end{pmatrix},$$

так что

$$B(1) = B_0 + B_1 = \begin{pmatrix} 0.1 & 0.2 \\ 0.2 & 0.4 \end{pmatrix}.$$

Матрица долгосрочных мультипликаторов равна

$$C(1) = A^{-1}(1) B(1) = \begin{pmatrix} 6 & 4 \\ 4 & 6 \end{pmatrix} \begin{pmatrix} 0.1 & 0.2 \\ 0.2 & 0.4 \end{pmatrix} = \begin{pmatrix} 1.4 & 2.8 \\ 1.6 & 3.2 \end{pmatrix},$$

так что стабильное решение есть

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 5.2 \\ 4.8 \end{pmatrix} + \begin{pmatrix} 1.4 & 2.8 \\ 1.6 & 3.2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix},$$

т.е.

$$y_1 = 5.2 + 1.4 x_1 + 2.8 x_2\,,$$
$$y_2 = 4.8 + 1.6 x_1 + 3.2 x_2\,.$$

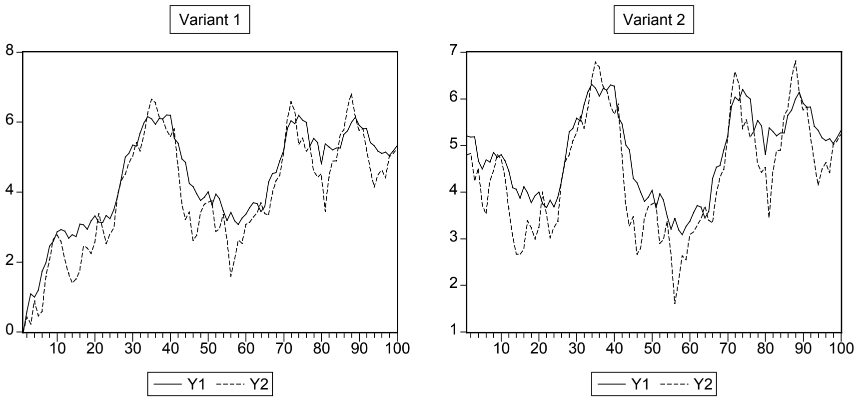
Ниже приведены графики смоделированных реализаций этой открытой системы в случае, когда $x_{1,\,t}$ и $x_{2,\,t}$ – независимые друг от друга AR(1) ряды,

$$x_{1,\,t} = 0.7\, x_{1,\,t-1} + v_{1t}\,, \quad x_{2,\,t} = 0.5\, x_{2,\,t-1} + v_{2t}\,; \quad v_{1t} \text{ и } v_{2t} \sim i.i.d.\ N(0,\,1).$$

В качестве начальных значений при моделировании взяты

- $x_{11} = x_{21} = 0,\ y_{11} = y_{21} = 0$ (вариант 1)
- $x_{11} = x_{21} = 0,\ y_{11} = 5.2,\ y_{21} = 4.8$ (вариант 2).

В результате получаем:



В первом случае, из-за несоответствия начальных значений переменных стабильным соотношениям, системе требуется некоторое время, чтобы выйти на стабильный режим. Во втором случае начальные значения переменных согласованы с долговременными соотношениями между переменными.

Рассмотрим следующую замкнутую VAR(1) для двух переменных:

$$y_{1t} = 0.8\, y_{1,\,t-1} + 0.2\, y_{2,\,t-1} + \varepsilon_{1t}\,,$$
$$y_{2t} = 0.2\, y_{1,\,t-1} + 0.8\, y_{2,\,t-1} + \varepsilon_{2t}\,.$$

Для этой системы

$$\begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} 0.8L & 0.2L \\ 0.2L & 0.8L \end{pmatrix} \begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix},$$

так что

$$A(L) = \begin{pmatrix} 1-0.8L & -0.2L \\ -0.2L & 1-0.8L \end{pmatrix}.$$

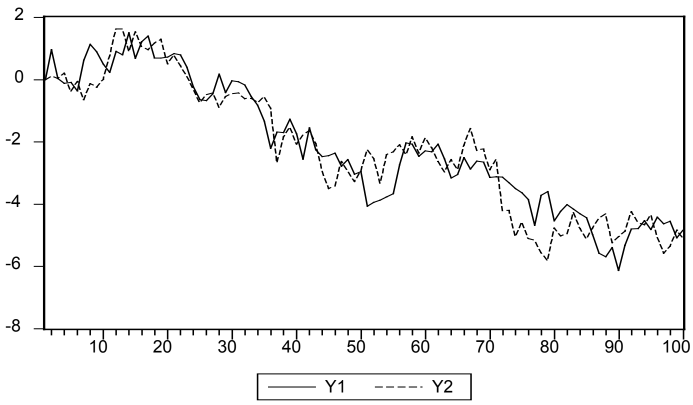
При этом

$$A(1) = \begin{pmatrix} 0.2 & 0.2 \\ 0.2 & 0.2 \end{pmatrix},$$

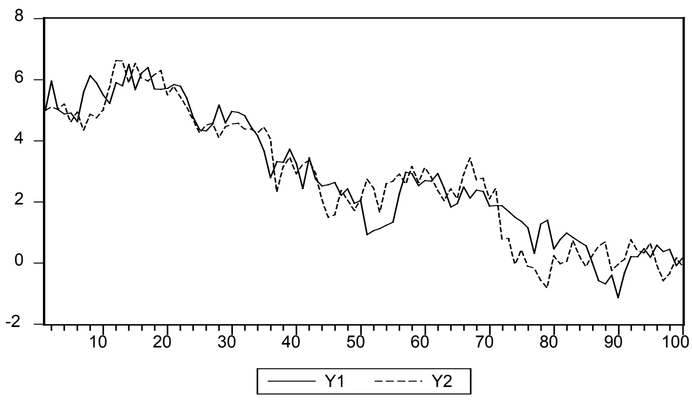
определитель этой матрицы равен нулю, и матрица $A^{-1}(1)$ не определена.

Уравнение $\det A(z) = 0$ имеет здесь вид $(1 - 0.8\,z)^2 - (0.2\,z)^2 = 0$, т.е. $(1 - z)(1 - 0.6\,z) = 0$.

Корни этого уравнения равны (1/0.6) и 1. Наличие корня, равного 1, нарушает условие стабильности системы. Как ведут себя в этом случае реализации системы? Ниже приводится соответствующий график.



Здесь стабилизации системы не наблюдается. Можно предположить, что это происходит из-за неудачного выбора начальных значений $y_{11} = y_{21} = 0$. Перемоделируем реализации, полагая начальные значения приблизительно равными наблюдаемому “конечному” уровню: $y_{11} = y_{21} = 5$. Новые реализации



по-прежнему не стабилизируются, и это отражает фундаментальное отличие рассматриваемой нестабильной модели VAR от стабильной.

4.4. Некоторые частные случаи динамических моделей

Чтобы не загромождать изложение, мы ограничимся здесь рассмотрением моделей, входящих в качестве частных случаев в модель ADL(1,1;1)

$y_t = \mu + a_1 y_{t-1} + \beta_0 x_t + \beta_1 x_{t-1} + \varepsilon_t$.

Эти частные случаи, несмотря на свою простоту, дают схематические представления девяти широко используемых типов моделей.

Различные типы моделей соответствуют различным ограничениям на вектор коэффициентов $\theta = (a_1 , \beta_0 , \beta_1)$. При наличии двух ограничений мы говорим об ***однопараметрической*** модели, а при наличии одного ограничения – о ***двухпараметрической модели***. Полная модель ADL(1,1;1) является ***трехпараметрической***. Ниже мы рассматриваем 9 различных типов моделей.

(1) Статическая регрессия $(a_1 = \beta_1 = 0)$: $y_t = \mu + \beta_0 x_t + \varepsilon_t$.

Здесь на значение y_t влияет только значение x_t в тот же момент времени; предшествующие значения y_{t-1} и x_{t-1} не влияют на y_t .

Такая модель обычно не характерна для данных, получаемых последовательно во времени, поскольку в таких ситуациях, как правило, случайные величины ε_t автокоррелированы.

(2) Процесс авторегрессии $(\beta_0 = \beta_1 = 0)$: $y_t = \mu + a_1 y_{t-1} + \varepsilon_t$.

Здесь значение y_t зависит только от значения y_{t-1} ; значения переменной x_t в моменты t и $(t-1)$ не влияют на y_t .

Подобные ситуации затрудняют экономический анализ и проведение соответствующей экономической политики из-за того, что в этом случае нет “управляющей” переменной, значения которой можно было бы устанавливать самостоятельно с целью управления значениями переменной y_t .

(3) Модель опережающего показателя $(a_1 = \beta_0 = 0)$:

$y_t = \mu + \beta_1 x_{t-1} + \varepsilon_t$.

Такие модели могут использоваться для прогнозирования, если изменения показателя y следуют с запаздыванием за изменениями показателя x с достаочной надежностью. Однако, при отсутствии серьезных теоретических оснований, коэффициент β_1 вовсе не обязан быть постоянным. В последнем случае это может приводить к некачественным прогнозам, особенно в периоды структурных изменений, когда хороший прогноз особенно необходим. Кроме того, не видно каких-то особых причин для исключения из правой части запаздывающих значений переменной y .

(4) Модель скорости роста $(a_1 = 1, \beta_1 = -\beta_0)$

$\Delta y_t = \mu + \beta_0 \Delta x_t + \varepsilon_t$,

($\Delta = 1 - L$, так что $\Delta y_t = y_t - y_{t-1}$, $\Delta x_t = x_t - x_{t-1}$), соответствует модели статической регрессии, но не для рядов в уровнях, а для рядов в разностях (для ***продифференцированных*** данных). Однако переход к рядам разностей оправдан только если исходные ряды имеют стохастический тренд и коинтегрированы. Об этом мы будем подробно говорить в последующих главах. А пока укажем только на то, что при неоправданном переходе к рядам разностей теряется информация о характере долговременной экономической связи между рядами в уровнях.

(5) Модель распределенных запаздываний $(a_1 = 0)$

$y_t = \mu + \beta_0 x_t + \beta_1 x_{t-1} + \varepsilon_t$

не содержит в правой части запаздываний переменной y . Она страдает теми же недостатками, что и статическая регрессия, но к ним еще может добавиться и проблема мультиколлинеарности переменных x_t и x_{t-1} .

(6) Модель частичной корректировки $(\beta_1 = 0)$

$\boldsymbol{y_t = \mu + a_1 y_{t-1} + \beta_0 x_t + \varepsilon_t}$

не содержит в правой части запаздывающих значений переменной x . К такой модели приводят, например, следующие соображения.

Пусть $y^*_t = \alpha + \beta x_t$ – целевой уровень переменной y , а фактически приращение $\Delta y_t = y_t - y_{t-1}$ описывается моделью

$$y_t - y_{t-1} = (1-\lambda)(y^*_t - y_{t-1}) + \varepsilon_t, \quad 0 \le \lambda \le 1,$$

т.е.

$$y_t = (1-\lambda)\, y^*_t + \lambda\, y_{t-1} + \varepsilon_t ,$$

так что, с точностью до случайной ошибки ε_t , текущее значение y_t равно взвешенному среднему целевого y^*_t и предыдущего значения переменной y . (Например, y_t – уровень запасов, x_t – уровень продаж.) Тогда

$$y_t = y_{t-1} + (1-\lambda)(\alpha + \beta x_t - y_{t-1}) + \varepsilon_t = (1-\lambda)\,\alpha + \lambda\, y_{t-1} + (1-\lambda)\,\beta x_t + \varepsilon_t ,$$

или

$$y_t = \mu + a_1 y_{t-1} + \beta_0 x_t + \varepsilon_t ,$$

где

$$\mu = (1-\lambda)\,\alpha , \quad a_1 = \lambda , \ \beta_0 = (1-\lambda)\,\beta .$$

Во многих случаях вывод подобных уравнений приводит к автокоррелированным ошибкам, а игнорирование x_{t-1} часто порождает оценку коэффициента a_1 , существенно отличающуюся от оценки a_1 в полной модели.

(7) Фальстарт, или приведенная форма (β_0 = 0):

$\boldsymbol{y_t = \mu + a_1 y_{t-1} + \beta_1 x_{t-1} + \varepsilon_t}$ **.**

К такой модели можно придти, например, если $x_t = \lambda x_{t-1} + u_t$. Тогда подстановка выражения для x_t в полное уравнение ADL(1,1;1) дает:

$$\begin{aligned} y_t &= \mu + a_1 y_{t-1} + \beta_0 x_t + \beta_1 x_{t-1} + \varepsilon_t \\ &= \mu + a_1 y_{t-1} + (\beta_0 \lambda + \beta_1)\, x_{t-1} + (\varepsilon_t + \beta_0 u_t), \end{aligned}$$

или

$$y_t = \mu + a_1 y_{t-1} + \beta^*_1 x_{t-1} + \varepsilon^*_t .$$

По одному последнему уравнению (***приведенная форма*** исходного уравнения) невозможно восстановить значения β_0 и β_1 , не зная значения λ . Т.е. мы можем оценить коэффициенты приведенной формы, но не коэффициенты ***структурной формы*** (исходного представления ADL(1,1;1)).

(8) Авторегрессионные ошибки ($\beta_1 = - a_1\beta_0$):

$\boldsymbol{y_t = \mu + a_1 y_{t-1} + \beta_0 x_t - a_1\beta_0 x_{t-1} + \varepsilon_t}$ **.**

Запишем это уравнение в виде

$$y_t - a_1 y_{t-1} = (1-a_1)\,\alpha + \beta_0 (x_t - a_1 x_{t-1}) + \varepsilon_t .$$

В последнем уравнении легко узнается известное ***преобразование Кохрейна – Оркатта***, используемое для преодоления проблемы автокоррелированности ошибок в модели парной регрессии

$$y_t = \alpha + \beta_0 x_t + u_t , \quad u_t = a_1 u_{t-1} + \varepsilon_t , \quad |a_1| < 1.$$

(9) Модель коррекции ошибок ($|a_1| < 1, \ \beta_0 + \beta_1 = b(1-a_1), b \ne 0$):

$$\boldsymbol{\Delta y_t = \mu + \beta_0 \Delta x_t - (1-a_1)(y_{t-1} - b\, x_{t-1}) + \varepsilon_t ,}$$

или

$$\boldsymbol{\Delta y_t = \beta_0 \Delta x_t - (1-a_1)(y_{t-1} - a - b\, x_{t-1}) + \varepsilon_t ,}$$

где $a = \mu/(1-a_1)$, $b = (\beta_0 + \beta_1)/(1-a_1)$.

Модели такого вида будут очень часто встречаться у нас при рассмотрении связей между нестационарными временными рядами. В этих случаях такая модель описывает механизм поддержания ***долговременной связи***

$y = a + b\,x$

между переменными y_t и x_t в форме коррекций отклонений

$y_{t-1} - a - b\,x_{t-1}$

от долговременной связи в предыдущий момент времени.

Замечание

Исходную (полную) модельADL(1,1;1)

$y_t = \mu + a_1 y_{t-1} + \beta_0 x_t + \beta_1 x_{t-1} + \varepsilon_t$

всегда можно преобразовать к виду

$y_t - y_{t-1} = \mu - (1 - a_1)\, y_{t-1} + \beta_0 (x_t - x_{t-1}) + (\beta_0 + \beta_1)\, x_{t-1} + \varepsilon_t$.

Если выполнено условие $|a_1| < 1$ (условие стабильности модели), то

$\Delta y_t = \mu + \beta_0 \Delta x_t - (1 - a_1)(\, y_{t-1} - ((\beta_0 + \beta_1)/(1 - a_1))\, x_{t-1}) + \varepsilon_t$,

и при $\beta_0 + \beta_1 \neq 0$ мы получаем модель коррекции ошибок.

Таким образом, модели с $|\, a_1 \,| < 1$ и $\beta_0 + \beta_1 \neq 0$ могут быть представлены в равносильной форме в виде модели коррекции ошибок.

Обратим теперь внимание на следующее. На практике мы имеем дело только со статистическими данными и не можем знать точно, какая именно модель лежала в основе ***процесса порождения данных (data generating process – DGP)***. Мы можем только, привлекая какие-то теоретические положения или результаты ранее проведенных исследований с другими множествами данных, выбрать некоторую ***статистическую модель (statistical model – SM)***, которую, по нашему мнению, можно использовать для описания процесса порождения данных. Выбрав такую модель, мы производим ее оценивание и затем можем по оцененной модели проверять различные гипотезы о ее коэффициентах, строить доверительные интервалы для коэффициентов и прогнозировать значения объясняемых переменных для нового набора объясняющих переменных. Между тем, здесь решающее значение имеет соотношение между истинным процессом порождения данных и выбранной статистической моделью.

Если статистическая модель SM оказывается более полной по сравнению с DGP, то тогда оценивание SM приводит к менее эффективным оценкам. С другой стороны, если процесс порождения данных оказывается полнее, чем выбранная SM, то это приводит к более неприятным последствиям – смещению оценок. Вследствие этого, обычно рекомендуется следовать принципу “от общего к частному”, т.е. первоначально выбирать в качестве статистической модели достаточно полную модель, а затем, производя последовательное тестирование статистической модели, редуцировать исходную статистическую модель к более экономной форме.

Пример

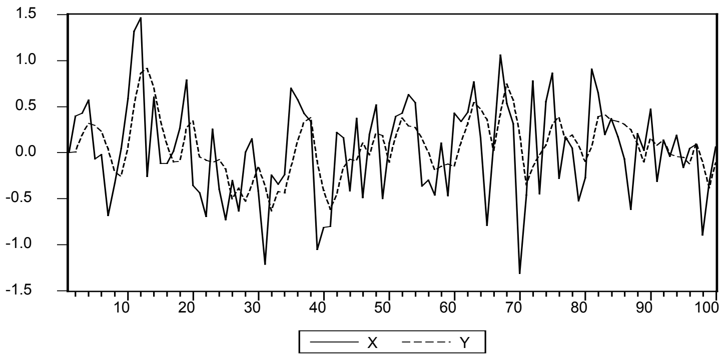
Статистические данные (n = 100) порождены стабильной моделью ADL(1,1,1)

$\boldsymbol{y_t = 0.5\, y_{t-1} + 0.2\, x_t + 0.3\, x_{t-1} + \varepsilon_t}$, $\boldsymbol{\varepsilon_t \sim i.i.d.\ N(0,\, 0.1^2)}$,

$\boldsymbol{x_t = 0.5\, x_{t-1} + v_t}$, $\boldsymbol{v_t \sim i.i.d.\ N(0,\, 0.5^2)}$,

причем ряды ε_t и v_t порождаются независимо друг от друга.

Смоделированные реализации имеют вид



Оценивание по этим реализациям полной модели ADL(1,1;1) в качестве статистической модели дает следующие результаты:

Dependent Variable: Y

Sample(adjusted): 2 100

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.014122	0.009556	1.477773	0.1428
Y(-1)	0.555208	0.034143	16.26107	0.0000
X	0.188567	0.018421	10.23666	0.0000
X(-1)	0.258377	0.020673	12.49808	0.0000

R-squared	0.913395	Mean dependent var	0.062869
Adjusted R-squared	0.910660	S.D. dependent var	0.310554
S.E. of regression	0.092824	Akaike info criterion	-1.876660
Sum squared resid	0.818547	Schwarz criterion	-1.771806

Исключая из правой части статистической модели константу, получаем:

Dependent Variable: Y

Sample(adjusted): 2 100

Variable	Coefficient	Std. Error	t-Statistic	Prob.
Y(-1)	0.565569	0.033621	16.82186	0.0000
X	0.190325	0.018495	10.29043	0.0000
X(-1)	0.256578	0.020764	12.35668	0.0000

R-squared	0.911404	Mean dependent var	0.062869
Adjusted R-squared	0.909558	S.D. dependent var	0.310554
S.E. of regression	0.093394	Akaike info criterion	-1.874134
Sum squared resid	0.837363	Schwarz criterion	-1.795494
Log likelihood	95.76965	Durbin-Watson stat	2.218619

Редуцированная модель признается лучшей по критерию Шварца. Проверка ее на адекватность дает следующие результаты.

- Коррелограмма ряда остатков соответствует процессу белого шума.
- Критерий Бройша – Годфри указывает на отсутствие автокоррелированности у ряда ε_t (P-значение = 0.375 при AR(1) альтернативе и 0.165 при AR(2) альтернативе).
- Критерий Jarque – Bera не обнаруживает значимых отклонений от нормальности (P-значение = 0.689).
- Критерий Уайта не обнаруживает гетероскедастичности (P-значение = 0.285).

Иными словами, применение критериев адекватности к оцененной модели дает удовлетворительные результаты.

Посмотрим теперь, что дает оценивание по тем же данным выбираемых в качестве SM перечисленных ранее 8 редуцированных моделей.

SM_1 *(статическая регрессия):* $y_t = \mu + \beta_0 x_t + \varepsilon_t$.

Оцененная модель:

Dependent Variable: Y

Sample: 1 100

Variable	Coefficient	Std. Error	t-Statistic	Prob.
X	0.271208	0.053356	5.082965	0.0000

R-squared	0.174472	Mean dependent var	0.062241
Adjusted R-squared	0.174472	S.D. dependent var	0.309046
S.E. of regression	0.280794	Akaike info criterion	0.307561
Sum squared resid	7.805700	Schwarz criterion	0.333613
Log likelihood	-14.37805	Durbin-Watson stat	0.839862

В правой части этой статистической модели нет запаздывающих значений объясняемой переменной. Поэтому здесь можно ориентироваться на значения статистики Дарбина – Уотсона. Низкое значение этой статистики указывает на автокоррелированность ряда ε_t , т.е. на неправильную спецификацию выбранной статистической модели.

SM_2 *Процесс авторегрессии:* $y_t = \mu + a_1 y_{t-1} + \varepsilon_t$.

Оцененная модель:

Dependent Variable: Y

Sample(adjusted): 2 100

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.013941	0.020679	0.674149	0.5018
Y(-1)	0.764874	0.065621	11.65594	0.0000

Поскольку в этой статистической модели правая часть содержит запаздывающее значение объясняемой переменной, ориентироваться на статистику Дарбина – Уотсона не следует. Проверку на отсутствие автокоррелированности для ряда ε_t выполняем, используя критерий Бройша – Годфри. При AR(1) альтернативе P-значение этого критерия равно 0.00003, так что гипотеза некоррелированности случайных величин ε_t отвергается. Следовательно, выбранная статистическая модель специфицирована неправильно.

SM_3 *Модель опережающего показателя*: $y_t = \mu + \beta_1 x_{t-1} + \varepsilon_t$.

Оцененная модель:

Dependent Variable: Y

Sample(adjusted): 2 100

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.049508	0.019722	2.510238	0.0137
X(-1)	0.455497	0.037291	12.21457	0.0000

При AR(1) альтернативе P-значение критерия Бройша – Годфри равно 0.0002, гипотеза некоррелированности случайных величин ε_t отвергается. Выбранная статистическая модель специфицирована неправильно.

SM$_4$ ***Модель скорости роста:*** $\Delta y_t = \mu + \beta_0 \Delta x_t + \varepsilon_t$

Оцененная модель

Dependent Variable: D(Y)

Sample(adjusted): 2 100

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.001126	0.021384	-0.052674	0.9581
D(X)	0.040538	0.033362	1.215078	0.2273
Log likelihood	13.74152	F-statistic		1.476415
Durbin-Watson stat	1.574116	Prob(F-statistic)		0.227286

При AR(1) альтернативе P-значение критерия Бройша – Годфри равно 0.029, гипотеза некоррелированности случайных величин ε_t отвергается. Выбранная статистическая модель специфицирована неправильно.

SM$_5$ ***Модель распределенных запаздываний:***

$y_t = \mu + \beta_0 x_t + \beta_1 x_{t-1} + \varepsilon_t$

Оцененная модель

Dependent Variable: Y

Sample(adjusted): 2 100

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.046032	0.018096	2.543741	0.0126
X	0.156214	0.035435	4.408526	0.0000
X(-1)	0.414363	0.035435	11.69370	0.0000

При AR(1) альтернативе P-значение критерия Бройша – Годфри равно 0.00000, гипотеза некоррелированности случайных величин ε_t отвергается. Выбранная статистическая модель специфицирована неправильно.

SM$_6$ ***Модель частичной корректировки:***

$y_t = \mu + a_1 y_{t-1} + \beta_0 x_t + \varepsilon_t$

Оцененная модель

Dependent Variable: Y

Sample(adjusted): 2 100

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.007088	0.015431	0.459354	0.6470
Y(-1)	0.753212	0.048925	15.39514	0.0000
X	0.253493	0.028588	8.867013	0.0000

При AR(1) альтернативе P-значение критерия Бройша – Годфри равно 0.012, гипотеза некоррелированности случайных величин ε_t отвергается. Выбранная статистическая модель специфицирована неправильно.

SM$_7$ ***Приведенная форма***: $y_t = \mu + a_1 y_{t-1} + \beta_1 x_{t-1} + \varepsilon_t$.

Оцененная модель

Dependent Variable: Y
Sample(adjusted): 2 100

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.020438	0.013757	1.485648	0.1406
Y(-1)	0.517457	0.048968	10.56734	0.0000
X(-1)	0.318058	0.028613	11.11579	0.0000

S.E. of regression	0.133909	Akaike info criterion	-1.153476
Sum squared resid	1.721440	Schwarz criterion	-1.074836

Здесь P-значение критерия Бройша – Годфри при AR(1) альтернативе равно 0.499, а при AR(2) альтернативе равно 0.538. Гипотеза некоррелированности случайных величин ε_t не отвергается, и можно перейти к проверке адекватности другими критериями. Критерий Jarque – Bera не обнаруживает значимых отклонений от нормальности (P-значение = 0.937). Критерий Уайта не обнаруживает гетероскедастичности (P-значение = 0.348).

Иными словами, применение критериев адекватности к оцененной модели дает удовлетворительные результаты. Поэтому возможно осуществить редукцию модели, основываясь на статистической незначимости константы в правой части уравнения. Исключение константы из правой части дает:

Dependent Variable: Y

Variable	Coefficient	Std. Error	t-Statistic	Prob.
Y(-1)	0.532005	0.048276	11.02001	0.0000
X(-1)	0.316252	0.028765	10.99445	0.0000

S.E. of regression	0.134740	Akaike info criterion	-1.150947
Sum squared resid	1.761018	Schwarz criterion	-1.098520

Модель без константы предпочтительнее по критерию Шварца.

С точки зрения анализа остатков, последняя модель вполне может быть использована для описания процесса порождения данных. Однако если мы сравним результаты ее оценивания с полученными ранее результатами оценивания модели $y_t = a_1 y_{t-1} + \beta_0 x_t + \beta_1 x_{t-1} + \varepsilon_t$, то обнаружим, что в модели с включением x_t в правую часть значения критериев Акаике (– 1.874) и Шварца (– 1.795) гораздо предпочтительнее.

SM$_8$ *Авторегрессионные ошибки:*

$$y_t = \mu + a_1 y_{t-1} + \beta_0 x_t - a_1\beta_0 x_{t-1} + \varepsilon_t\,.$$

Использование нелинейного (итерационного) метода наименьших квадратов для оценивания параметров этого уравнения дает следующие результаты

Dependent Variable: Y
Sample(adjusted): 2 100
Convergence achieved after 19 iterations
Y=C(1)+C(2)*Y(-1)+C(3)*X-(C(2)*C(3))*X(-1)

	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	0.014489	0.020617	0.702743	0.4839
C(2)	0.749747	0.070184	10.68267	0.0000
C(3)	0.052577	0.036535	1.439066	0.1534

R-squared	0.592630	Mean dependent var	0.062869
Adjusted R-squared	0.584144	S.D. dependent var	0.310554

S.E. of regression	0.200267	Akaike info criterion	-0.348499
Sum squared resid	3.850250	Schwarz criterion	-0.269859
Log likelihood	20.25069	Durbin-Watson stat	1.447077

При AR(1) альтернативе P-значение критерия Бройша – Годфри равно 0.0002, гипотеза некоррелированности случайных величин ε_t отвергается. Выбранная статистическая модель специфицирована неправильно.

Рассмотрим также оценивание SM в форме модели коррекции ошибок (хотя эта модель и не является редуцированной).

SM$_9$ ***Модель коррекции ошибок*:**

$$\Delta y_t = \mu + \beta_0\,\Delta x_t - (1 - a_1)(\,y_{t-1} - b\,x_{t-1}) + \varepsilon_t\,.$$

Оцененная модель (нелинейный метод наименьших квадратов):

Dependent Variable: D(Y)
Sample(adjusted): 2 100
Convergence achieved after 4 iterations
D(Y) =C(1)+C(2)*D(X) + (C(3)-1)*(Y(-1)-C(4)*X(-1))

	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	0.014122	0.009556	1.477773	0.1428
C(2)	0.188567	0.018421	10.23666	0.0000
C(3)	0.555208	0.034143	16.26107	0.0000
C(4)	1.004839	0.078119	12.86299	0.0000

R-squared	0.816395	Mean dependent var	-0.001100
Adjusted R-squared	0.810597	S.D. dependent var	0.213288
S.E. of regression	0.092824	Akaike info criterion	-1.876660
Sum squared resid	0.818547	Schwarz criterion	-1.771806
Log likelihood	96.89465	Durbin-Watson stat	2.248395

P-значение критерия Бройша – Годфри при AR(1) альтернативе равно 0.130, а при AR(2) альтернативе равно 0.318; гипотеза некоррелированности случайных величин ε_t не отвергается. Критерий Jarque – Bera не обнаруживает значимых отклонений от нормальности (P-значение = 0.711). Критерий Уайта не обнаруживает гетероскедастичности (P-значение = 0.380).

Иными словами, применение критериев адекватности к оцененной модели дает удовлетворительные результаты. Опираясь на них, редуцируем модель, исключая из правой части константу; при этом получаем:

Dependent Variable: D(Y)
Convergence achieved after 3 iterations
D(Y) =C(2)*D(X) + (C(3)-1)*(Y(-1)-C(4)*X(-1))

	Coefficient	Std. Error	t-Statistic	Prob.
C(2)	0.190325	0.018495	10.29043	0.0000
C(3)	0.565569	0.033621	16.82186	0.0000
C(4)	1.028710	0.080225	12.82279	0.0000

R-squared	0.812174	Mean dependent var	-0.001100
Adjusted R-squared	0.808261	S.D. dependent var	0.213288
S.E. of regression	0.093394	Akaike info criterion	-1.874134
Sum squared resid	0.837363	Schwarz criterion	-1.795494

Log likelihood	95.76965	Durbin-Watson stat	2.218619

т.е.

$\Delta y_t = 0.190\,\Delta x_t - 0.434(y_{t-1} - 1.029\,x_{t-1}) + e_t$.

Модель без константы предпочтительнее по критерию Шварца.

Уединяя y_t в левой части уравнения, получаем:

$y_t = 0.566\,y_{t-1} + 0.190\,x_t + 0.253\,x_{t-1} + e_t$.

Сравним это уравнение с реально использованным для моделирования DGP: $y_t = 0.5\,y_{t-1} + 0.2\,x_t + 0.3\,x_{t-1} + \varepsilon_t$

и с результатом оценивания соответствующей ему статистической модели:

$y_t = 0.565\,y_{t-1} + 0.190\,x_t + 0.257\,x_{t-1} + \varepsilon_t$.

Найдем долговременное соотношение между переменными y_t и x_t, соответствующее теоретическому процессу порождения данных:

$y_t = 0.5\,y_t + 0.2\,x_t + 0.3\,x_t \;\Rightarrow\; y = x$.

В то же время, долговременное соотношение, получаемое по оцененной SM, соответствующей этому DGP:

$y_t = 0.565\,y_t + 0.190\,x_t + 0.257\,x_t \;\Rightarrow\; y = 1.002\,x$.

Далее, долговременное соотношение, получаемое по оцененной SM_9 (в варианте без константы в правой части):

$y_t = 0.566\,y_t + 0.190\,x_t + 0.253\,x_t \;\Rightarrow\; y = 1.021\,x$.

Наконец, если взять результаты оценивания модели SM_7 (приведенная форма) без включения константы, то для этого случая получаем:

$y_t = 0.532\,y_t + 0.316\,x_t \;\Rightarrow\; y = 0.675\,x$.

Эти результаты указывают на возможность серьезных последствий, проистекающих из неправильной спецификации SM, когда эта спецификация оказывается уже спецификации DGP. Заметим, что в рамках такой SM отнюдь не всегда удается обнаружить статистическими методами “узость” выбранной спецификации. Мы смогли это сделать в рамках оцененных статистических моделей $SM_1 - SM_6$ и SM_8, но не в модели SM_7 .

Рассмотрим теперь обратную ситуацию, когда, напротив, выбранная для оценивания статистическая модель SM оказывается полнее (“шире”) модели DGP, так что модель, соответствующая DGP, является частным случаем статистической модели, выбранной для оценивания.

В качестве DGP будем последовательно брать модели (1) – (8), а в качестве SM – полную модель ADL(1,1;1) без ограничений на коэффициенты:

$y_t = \mu + a_1\,y_{t-1} + \beta_0\,x_t + \beta_1\,x_{t-1} + \varepsilon_t$.

Значения коэффициентов при переменных в моделях (1) – (8) будем брать такими же, как и в исходной модели ADL(1,1;1)

$y_t = 0.5\,y_{t-1} + 0.2\,x_t + 0.3\,x_{t-1} + \varepsilon_t$.

При моделировании DGP во всех случаях берется $\varepsilon_t \sim i.i.d.\ N(0, 0.1^2)$.

DGP_1 : Статическая регрессия

$\boldsymbol{y_t = 0.2\,x_t + \varepsilon_t}$.

Оцененная статистическая модель

Dependent Variable: Y

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.004647	0.010300	-0.451175	0.6529
Y(-1)	0.102848	0.101833	1.009966	0.3151
X	0.186813	0.020222	9.238033	0.0000
X(-1)	0.000201	0.028272	0.007101	0.9943

R-squared	0.507795	Mean dependent var	0.001230
Adjusted R-squared	0.492252	S.D. dependent var	0.143412
S.E. of regression	0.102190	Akaike info criterion	-1.684398
Sum squared resid	0.992068	Schwarz criterion	-1.579545

P-значение критерия Бройша – Годфри при AR(1) альтернативе равно 0.760, а при AR(2) альтернативе равно 0.951, гипотеза некоррелированности случайных величин ε_t не отвергается. Критерий Jarque – Bera не обнаруживает значимых отклонений от нормальности (P-значение = 0.733). Критерий Уайта не обнаруживает гетероскедастичности (P-значение = 0.770).

Иными словами, применение критериев адекватности к оцененной модели дает удовлетворительные результаты. Опираясь на них, можно перейти к проверке гипотез о значениях коэффициентов. При проверке гипотезы о занулении константы и коэффициентов при y_{t-1} и x_{t-1} получаем значение обычной F-статистики, равное F = 0.738, и qF = 2.214. Исходя из F-распределения для статистики F , получаем P-значение 0.532. Использование асимптотического распределения χ^2 (3) для qF приводит к P-значению 0.529. При обоих вариантах гипотеза о занулении трех указанных коэффициентов не отвергается. Тем самым, можно перейти к оцениванию редуцированной модели $y_t = \mu + \beta_0 x_t + \varepsilon_t$, и это дает:

Dependent Variable: Y

Sample: 1 100

Variable	Coefficient	Std. Error	t-Statistic	Prob.
X	0.190067	0.019291	9.852604	0.0000

R-squared	0.494995	Mean dependent var	0.001935
Adjusted R-squared	0.494995	S.D. dependent var	0.142860
S.E. of regression	0.101522	Akaike info criterion	-1.727139
Sum squared resid	1.020359	Schwarz criterion	-1.701087

Редуцированная модель лучше полной и по критерию Акаике и по критерию Шварца. Остатки от оцененной редуцированной модели проходят тесты на нормальность, отсутствие автокоррелированности и гетероскедастичности.

DGP$_2$: *Процесс авторегрессии*

$$y_t = 0.5\, y_{t-1} + \varepsilon_t .$$

Оцененная статистическая модель

Dependent Variable: Y

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.004519	0.010315	-0.438075	0.6623
Y(-1)	0.576756	0.084134	6.855173	0.0000
X	-0.013220	0.020253	-0.652774	0.5155
X(-1)	0.021476	0.020228	1.061719	0.2911

R-squared	0.338422	Mean dependent var	-0.007891

Adjusted R-squared	0.317530	S.D. dependent var	0.123820
S.E. of regression	0.102290	Akaike info criterion	-1.682441
Sum squared resid	0.994011	Schwarz criterion	-1.577588

P-значение критерия Бройша – Годфри при AR(1) альтернативе равно 0.600, а при AR(2) альтернативе равно 0.773, гипотеза некоррелированности случайных величин ε_t не отвергается. Критерий Jarque – Bera не обнаруживает значимых отклонений от нормальности (P-значение = 0.654). Критерий Уайта не обнаруживает гетероскедастичности (P-значение = 0.956).

При проверке гипотезы о занулении константы и коэффициентов при x_t и x_{t-1} получаем значение обычной F-статистики, равное $F = 0.641$, и $qF = 1.283$. Исходя из F-распределения для статистики F, получаем P-значение 0.529. Использование асимптотического распределения $\chi^2(3)$ для qF приводит к P-значению 0.527. При обоих вариантах гипотеза о занулении трех указанных коэффициентов не отвергается. Тем самым, можно перейти к оцениванию редуцированной модели $y_t = a_1 y_{t-1} + \varepsilon_t$, и это дает:

Dependent Variable: Y

Variable	Coefficient	Std. Error	t-Statistic	Prob.
Y(-1)	0.575922	0.082705	6.963585	0.0000

R-squared	0.328274	Mean dependent var	-0.007891
Adjusted R-squared	0.328274	S.D. dependent var	0.123820
S.E. of regression	0.101482	Akaike info criterion	-1.727825
Sum squared resid	1.009258	Schwarz criterion	-1.701612

Редуцированная модель предпочтительнее и по критерию Акаике и по критерию Шварца. Анализ остатков не выявляет значимых отклонений от сделанных предположении в отношении ряда ε_t .

DGP$_3$: ***Модель опережающего показателя***

$$\mathbf{y_t = 0.3\, x_{t-1} + \varepsilon_t}\ .$$

Оцененная модель

Dependent Variable: Y

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.005202	0.010305	-0.504767	0.6149
Y(-1)	0.052373	0.054196	0.966363	0.3363
X	-0.012475	0.020310	-0.614213	0.5405
X(-1)	0.315962	0.020645	15.30433	0.0000

R-squared	0.736662	Mean dependent var	0.004035
Adjusted R-squared	0.728346	S.D. dependent var	0.196154
S.E. of regression	0.102236	Akaike info criterion	-1.683501
Sum squared resid	0.992959	Schwarz criterion	-1.578647

P-значение критерия Бройша – Годфри при AR(1) альтернативе равно 0.614, а при AR(2) альтернативе равно 0.868, гипотеза некоррелированности случайных величин ε_t не отвергается. Критерий Jarque – Bera не обнаруживает значимых отклонений от нормальности (P-значение = 0.740). Критерий Уайта не обнаруживает гетероскедастичности (P-значение = 0.804).

При проверке гипотезы о занулении константы и коэффициентов при x_t и y_{t-1} получаем значение обычной F-статистики, равное $F = 0.577$, и $qF = 1.730$. Исходя из F-распределения для статистики F, получаем P-значение 0.632. Использование асимптотического распределения $\chi^2(3)$ для qF приводит к P-значению 0.630. При обоих вариантах гипотеза о занулении трех указанных коэффициентов не отвергается. Тем самым, можно перейти к оцениванию редуцированной модели $y_t = \beta_1 x_{t-1} + \varepsilon_t$, и это дает:

Dependent Variable: Y

Variable	Coefficient	Std. Error	t-Statistic	Prob.
X(-1)	0.315777	0.019302	16.35987	0.0000

R-squared	0.731866	Mean dependent var	0.004035
Adjusted R-squared	0.731866	S.D. dependent var	0.196154
S.E. of regression	0.101572	Akaike info criterion	-1.726058
Sum squared resid	1.011044	Schwarz criterion	-1.699844

Редуцированная модель предпочтительнее и по критерию Акаике и по критерию Шварца. Анализ остатков не выявляет значимых отклонений от сделанных предположении в отношении ряда ε_t.

DGP$_4$: ***Модель скорости роста***

$\Delta y_t = \mu + \beta_0 \Delta x_t + \varepsilon_t$.

Оцененная модель

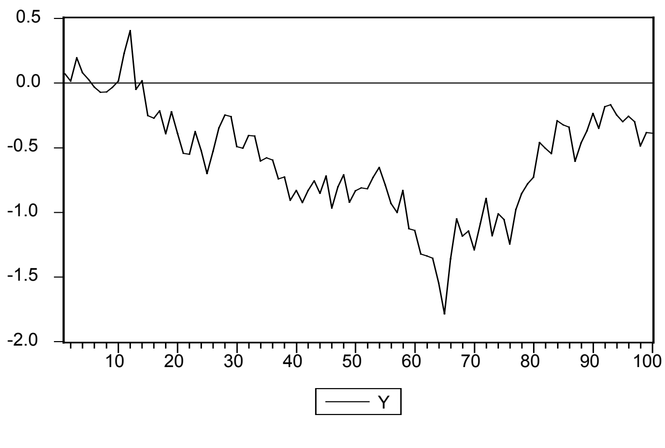
Dependent Variable: Y

Included observations: 99 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.029026	0.017741	-1.636049	0.1051
Y(-1)	0.959750	0.024228	39.61303	0.0000
X	0.184064	0.019993	9.206520	0.0000
X(-1)	-0.173461	0.020340	-8.528162	0.0000

R-squared	0.944972	Mean dependent var	-0.599911
Adjusted R-squared	0.943234	S.D. dependent var	0.425076
S.E. of regression	0.101277	Akaike info criterion	-1.702356
Sum squared resid	0.974412	Schwarz criterion	-1.597502
Log likelihood	88.26661	F-statistic	543.7988
Durbin-Watson stat	1.745026	Prob(F-statistic)	0.000000

Отметим близкое к единице оцененное значение коэффициента при y_{t-1}, что может говорить о том, что в DGP "истинное" значение коэффициента равно $a_1 = 1$. Но тогда нарушается условие стабильности системы. И действительно, график ряда y_t, полученного в результате моделирования, имеет вид

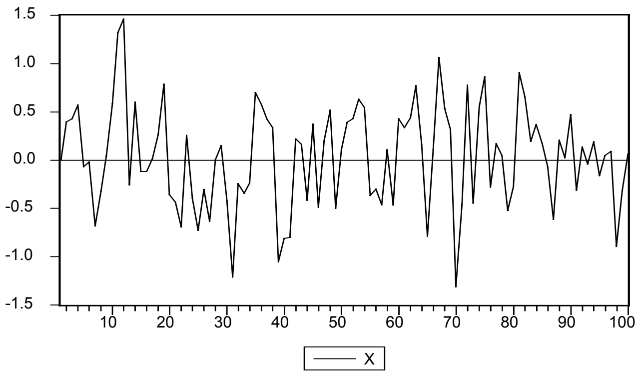


явно указывающий на нестационарность ряда.

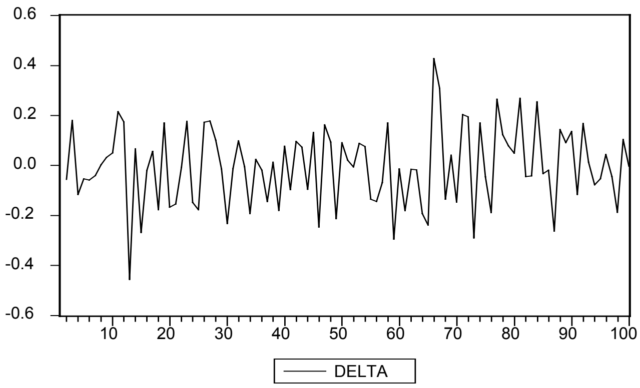
Вопрос о проверке гипотез типа H_0: $a_1 = 1$ требует специального рассмотрения, и мы будем рассматривать его в последующих главах. Сейчас же, исходя из наблюдаемого поведения ряда y_t и близости оцененного значения коэффициента к 1, займемся оцениванием модели

$$\Delta y_t = \mu + \beta_0 x_t + \beta_1 x_{t-1} + \varepsilon_t .$$

Мы можем использовать для нее стандартную (асимптотическую) технику статистических выводов, поскольку реализация ряда x_t имеет вид



указывающий на стационарность этого ряда, и реализация ряда разностей Δy_t имеет вид



говорящий в пользу стационарности ряда Δy_t.

В результате оценивания получаем:

Dependent Variable: D(Y)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.004915	0.010297	-0.477336	0.6342
X	0.185224	0.020163	9.186333	0.0000
X(-1)	-0.179782	0.020163	-8.916410	0.0000

Здесь мы, конечно, обращаем внимание на статистическую незначимость константы, а также на то, что оцененные коэффициенты при переменных x_t и x_{t-1} близки по абсолютной величине и противоположны по знаку. В связи с этим, мы в рамках статистической модели

$$\Delta y_t = \mu + \beta_0 x_t + \beta_1 x_{t-1} + \varepsilon_t$$

проверяем гипотезу H_0: $\mu = 0$, $\beta_0 = -\beta_1$. Используя F-распределение для F-статистики и распределение хи-квадрат $\chi^2(2)$ для статистики $qF = 2F$, получаем в обоих случаях P-значение 0.876. Гипотеза H_0 не отвергается, и можно перейти к оцениванию модели с такими ограничениями, т.е. модели $\Delta y_t = \beta_0 \Delta x_t + \varepsilon_t$. Оцененная модель с ограничениями:

Dependent Variable: D(Y)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(X)	0.182495	0.015882	11.49045	0.0000

P-значение критерия Бройша – Годфри при AR(1) альтернативе равно 0.328, а при AR(2) альтернативе равно 0.605; гипотеза некоррелированности случайных величин ε_t не отвергается. Критерий Jarque – Bera не обнаруживает значимых отклонений от нормальности (*P*-значение = 0.673). Критерий Уайта не обнаруживает гетероскедастичности (*P*-значение = 0.988). Таким образом, и в этом случае мы вышли в результате тестирования на статистическую модель, имеющую ту же спецификацию, что и DGP.

$\mathbf{DGP_5}$ *Модель распределенных запаздываний*

$$y_t = \mu + 0.2\, x_t + 0.3\, x_{t-1} + \varepsilon_t\,.$$

Оцененная модель

Dependent Variable: Y

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.003282	0.010206	-0.321584	0.7485
Y(-1)	0.000523	0.058294	0.008975	0.9929
X	0.181735	0.019977	9.097166	0.0000
X(-1)	0.289502	0.024922	11.61638	0.0000

R-squared	0.804020	Mean dependent var	0.010665
Adjusted R-squared	0.797831	S.D. dependent var	0.225113
S.E. of regression	0.101218	Akaike info criterion	-1.703515
Sum squared resid	0.973283	Schwarz criterion	-1.598661

P-значение критерия Бройша – Годфри при AR(1) альтернативе равно 0.972, а при AR(2) альтернативе равно 0.826; гипотеза некоррелированности случайных величин ε_t не отвергается. Критерий Jarque – Bera не обнаруживает значимых отклонений от нормальности (P-значение = 0.689). Критерий Уайта не обнаруживает гетероскедастичности (P-значение = 0.433).

Проверяем гипотезу H_0: $\mu = 0,\ a_1 = 0$. Используя *F*-распределение для *F*-статистики и распределение хи-квадрат $\chi^2(2)$ для статистики $qF = 2F$, получаем в обоих случаях P-значение 0.950. Гипотеза H_0 не отвергается, и можно перейти к оцениванию модели с $\mu = 0,\ a_1 = 0$, т.е. модели $y_t = \mu + \beta_0\, x_t + \beta_1\, x_{t-1} + \varepsilon_t$. Оцененная модель с ограничениями:

Dependent Variable: Y

Variable	Coefficient	Std. Error	t-Statistic	Prob.
X	0.181461	0.019754	9.185972	0.0000
X(-1)	0.289367	0.019756	14.64736	0.0000

$\mathbf{DGP_6}$ *Модель частичной корректировки*

$$y_t = \mu + 0.5\, y_{t-1} + 0.2\, x_t + \varepsilon_t\,.$$

Оцененная модель:

Dependent Variable: Y

Variable	Coefficient	Std. Error	t-Statistic	Prob.

C	-0.005099	0.010263	-0.496869	0.6204
Y(-1)	0.592041	0.071296	8.304016	0.0000
X	0.188766	0.020280	9.308077	0.0000
X(-1)	0.000973	0.025019	0.038898	0.9691

P-значение критерия Бройша – Годфри при AR(1) альтернативе равно 0.904, а при AR(2) альтернативе равно 0.723; гипотеза некоррелированности случайных величин ε_t не отвергается. Критерий Jarque – Bera не обнаруживает значимых отклонений от нормальности (P-значение = 0.691). Критерий Уайта не обнаруживает гетероскедастичности (P-значение = 0.533).

Проверяем гипотезу H_0: $\mu = 0$, $\beta_1 = 0$. Используя F-распределение для F-статистики и распределение хи-квадрат $\chi^2(2)$ для статистики $qF = 2F$, получаем в обоих случаях P-значение 0.884. Гипотеза H_0 не отвергается, и можно перейти к оцениванию модели с $\mu = 0$, $\beta_1 = 0$, т.е. модели $y_t = \mu + a_1\, y_{t-1} + \beta_0\, x_t + \varepsilon_t$. Оцененная модель с ограничениями:

Dependent Variable: Y

Variable	Coefficient	Std. Error	t-Statistic	Prob.
Y(-1)	0.592666	0.056702	10.45233	0.0000
X	0.188468	0.019202	9.814814	0.0000

R-squared	0.690477	Mean dependent var	0.004719
Adjusted R-squared	0.687286	S.D. dependent var	0.180476
S.E. of regression	0.100924	Akaike info criterion	-1.728911
Sum squared resid	0.988000	Schwarz criterion	-1.676484

DGP$_7$ ***Приведенная форма***

$$y_t = 0.5\, y_{t-1} + 0.3\, x_{t-1} + \varepsilon_t\,.$$

Оцененная модель

Dependent Variable: Y

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.005886	0.010271	-0.573100	0.5679
Y(-1)	0.559497	0.042762	13.08387	0.0000
X	-0.010318	0.020320	-0.507807	0.6128
X(-1)	0.316645	0.020229	15.65291	0.0000

P-значение критерия Бройша – Годфри при AR(1) альтернативе равно 0.701, а при AR(2) альтернативе равно 0.827; гипотеза некоррелированности случайных величин ε_t не отвергается. Критерий Jarque – Bera не обнаруживает значимых отклонений от нормальности (P-значение = 0.740). Критерий Уайта не обнаруживает гетероскедастичности (P-значение = 0.586).

Проверяем гипотезу H_0: $\mu = 0$, $\beta_0 = 0$. Используя F-распределение для F-статистики и распределение хи-квадрат $\chi^2(2)$ для статистики $qF = 2F$, получаем в обоих случаях P-значение 0.734. Гипотеза H_0 не отвергается, и можно перейти к оцениванию модели с $\mu = 0$, $\beta_0 = 0$, т.е. к оцениванию модели $y_t = a_1\, y_{t-1} + \beta_1\, x_{t-1} + \varepsilon_t$. Оцененная модель с ограничениями:

Dependent Variable: Y

Variable	Coefficient	Std. Error	t-Statistic	Prob.
Y(-1)	0.561389	0.041849	13.41452	0.0000

X(-1)	0.313207	0.019269	16.25422	0.0000

DGP$_8$ ***Авторегрессионные ошибки***

$$y_t = 0.5\,y_{t-1} + 0.2\,x_t - 0.1\,x_{t-1} + \varepsilon_t\,.$$

Оцененная модель:

Dependent Variable: Y

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.004519	0.010315	-0.438075	0.6623
Y(-1)	0.576756	0.084134	6.855173	0.0000
X	0.186780	0.020253	9.222532	0.0000
X(-1)	-0.093875	0.025414	-3.693770	0.0004

P-значение критерия Бройша – Годфри при AR(1) альтернативе равно 0.600, а при AR(2) альтернативе равно 0.773; гипотеза некоррелированности случайных величин ε_t не отвергается. Критерий Jarque – Bera не обнаруживает значимых отклонений от нормальности (P-значение = 0.654). Критерий Уайта не обнаруживает гетероскедастичности (P-значение = 0.682).

Исключим из статистической модели статистически незначимую константу:

Dependent Variable: Y

Variable	Coefficient	Std. Error	t-Statistic	Prob.
Y(-1)	0.578309	0.083705	6.908866	0.0000
X	0.186426	0.020151	9.251413	0.0000
X(-1)	-0.094531	0.025263	-3.741827	0.0003

Заметим, что произведение оцененных коэффициентов при y_{t-1} и x_t равно 0.108, т.е. близко по абсолютной величине и противоположно по знаку коэффициенту при x_{t-1}. В связи с этим наблюдением, проверим гипотезу H$_0$: $\beta_1 = -\,a_1\beta_0$. Здесь мы имеем дело с <u>нелинейной</u> гипотезой, и результаты проверки могут зависеть от формы записи этого ограничения на коэффициенты. Поэтому мы проверяем указанную гипотезу в трех формах:

H$_0$: $\beta_1 = -\,a_1\beta_0$; H$_0$: $\beta_0 = -(\beta_1/a_1)$; H$_0$: $a_1 = -(\beta_1/\beta_0)$.

Соответствующие этим формам P-значения $\chi^2(1)$-критериев равны 0.515, 0.514 и 0.506, так что выводы в отношении гипотезы H$_0$ согласуются: эта гипотеза не отвергается. Последнее означает, что можно перейти к оцениванию модели с ограничением на коэффициенты в виде $\beta_1 = -\,a_1\beta_0$, т.е. к модели $y_t = a_1 y_{t-1} + \beta_0 x_t - a_1\beta_0 x_{t-1} + \varepsilon_t$. В итоге получаем оцененную модель

Dependent Variable: Y
Convergence achieved after 3 iterations
Y =C(1)*Y(-1)+C(2)*X-(C(1)*C(2))*X(-1)

	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	0.575812	0.083369	6.906747	0.0000
C(2)	0.182370	0.019110	9.543254	0.0000

которую можно записать в виде

$$y_t = 0.576\,y_{t-1} + 0.182\,x_t - 0.105\,x_{t-1} + \varepsilon_t\,.$$

Таким образом, во всех рассмотренных случаях, когда DGP являлся частным случаем выбранной для оценивания статистической модели, последовательное применение метода редукции модели от общего к частному (с предварительной проверкой SM на адекватность) выводило нас на редуцированные модели, спецификация которых соответствовала спецификации DGP. В то же время, как мы видели перед этим, при движении “от частного к общему” возможны ситуации, когда остановка происходит на модели, существенно отличающейся от DGP, хотя и проходящей стандартные тесты на адекватность имеющимся статистическим данным.

Это еще раз подчеркивает предпочтительность использования при подборе моделей по статистическим данным метода “от общего к частному”, т.е. первоначальному выбору достаточно общей модели, проверки ее на адекватность имеющимся статистическим данным, и, в случае признания выбранной модели адекватной данным, последующей редукции этой модели с использованием стандартных критериев спецификации.

В связи с последними замечаниями, еще раз обратимся к модели линейной регрессии с автокоррелированными ошибками, образующими стационарный процесс авторегрессии первого порядка. В учебной литературе по эконометрике довольно часто делается упор на эту модель как средство преодоления проблемы автокоррелированности ошибок в рамках известных процедур Кохрейна – Оркатта или Прайса – Уинстена. Однако, как ясно из предыдущего изложения, такая модель (в нашей нумерации – модель 8) является всего лишь весьма частным случаем общей динамической модели ADL(1,1;1). В рамках этой общей модели

$$y_t = \mu + a_1 y_{t-1} + \beta_0 x_t + \beta_1 x_{t-1} + \varepsilon_t$$

модель, о которой идет речь, выделяется выполнением соотношения

$$\beta_1 = -a_1\beta_0 .$$

В то же время, при $\beta_0 \neq 0$ общую модель ADL(1,1;1) можно представить в виде

$$(1 - a_1L)y_t = \beta_0 \left(1 + \frac{\beta_1}{\beta_0} L\right) x_t + \varepsilon_t ,$$

или

$$a(L)\, y_t = b(L)\, x_t + \varepsilon_t ,$$

где

$$a(L) = 1 - a_1L ,\ b(L) = \beta_0 \left(1 + \frac{\beta_1}{\beta_0} L\right).$$

(Для простоты мы полагаем $\mu = 0$.)

Если $a_1 = -\dfrac{\beta_1}{\beta_0}$, то модель принимает вид

$$(1 - a_1L)y_t = \beta_0 (1 - a_1L)\, x_t + \varepsilon_t ,$$

так что многочлены $a(L)$ и $b(L)$ имеют ***общий множитель*** $(1 - a_1L)$. Разделив обе части уравнения на этот общий множитель, получаем

$$y_t = \beta_0 x_t + u_t ,$$

где

$$u_t = \frac{\varepsilon_t}{1 - a_1L} ,$$

так что $(1 - a_1L)\, u_t = \varepsilon_t$ и $u_t = a_1 u_{t-1} + \varepsilon_t$.

В связи с наличием общего множителя, модель авторегрессионных ошибок относят к классу моделей, носящему название COMFAC (common factors). Рассматриваемая модель обязана принадлежностью к этому классу именно наличию ограничения $\beta_1 = -a_1\beta_0$. Класс COMFAC является весьма частным случаем моделей с авторегрессионно распределенными запаздываниями. Поэтому применение обычной процедуры проверки автокоррелированности ошибок в модели регрессии $y_t = \beta_0 x_t + u_t$ и коррекции обнаруженной автокоррелированности посредством авторегрессионного преобразования переменных, вообще говоря, некорректно. Правильный порядок действий должен состоять

- В установлении пригодности модели $a(L)\ y_t = b(L)\ x_t + \varepsilon_t$ с помощью различных критериев адекватности; гипотеза о том, что ряд ε_t является гауссовским белым шумом, не должна отвергаться – в противном случае следует говорить о непригодности уже этой общей модели.
- В проверке гипотезы о том, что многочлены $a(L)$ и $b(L)$ имеют общие корни.
- Наконец, в случае подтверждения обеих гипотез следует проверить гипотезу H_0: $a_1 = 0$ (она соответствует модели статической регрессии). Заметим, что отвержение этой гипотезы непосредственно в модели с автокоррелированными ошибками вовсе не доказывает наличия указанных общих множителей.

Однако здесь имеются некоторые сложности.

На первом шаге гипотеза H_0: “ε_t – белый шум” проверяется, в частности, против альтернативы H_A: $\varepsilon_t \sim AR(k)$ с $k \le p$, т.е.

$$\varepsilon_t = \rho_1\,\varepsilon_{t-1} + \ldots + \rho_p\,\varepsilon_{t-p} + v_t\,,$$

где $v_t \sim i.i.d.$ и хотя бы одно $\rho_j \ne 0$. Модель, соответствующая альтернативе, имеет вид

$$y_t = a_1 y_{t-1} + \beta_0 x_t + \beta_1 x_{t-1} + \rho_1\,\varepsilon_{t-1} + \ldots + \rho_p\,\varepsilon_{t-p} + v_t\,,$$

и, фактически, речь идет о проверке гипотезы H_0: $\rho_1^2 = \ldots = \rho_p^2 = 0$ против H_A: $\rho_1^2 + \ldots + \rho_p^2 \ne 0$. Такую проверку можно осуществить, используя стандартный LM критерий Бройша – Годфри. В то же время не рекомендуется использовать для этой цели критерии, основанные на статистиках Бокса – Пирса и Льюнга – Бокса из разд. 3.1 и предназначенные для анализа “сырых” рядов. (См., например, статью [Kwan, Sim (1996)].)

Проблемы возникают и с применением стандартного критерия Вальда для проверки гипотезы H_1: $\beta_1 = -a_1\beta_0$ против альтернативы H_A: $\beta_1 \ne -a_1\beta_0$. Дело в том, что эта гипотеза не является линейной, а в таких случаях результаты применения критерия Вальда зависят от того, в какой форме записано ограничение: $\beta_1 = -a_1\beta_0$, $a_1 = -\beta_1/\beta_0$ или $\beta_0 = -\beta_1/a_1$, что может приводить к противоречивым выводам.

Отметим также проблему, связанную с последовательным использованием нескольких критериев проверки гипотез. В рамках рассмотренной процедуры приходится, по крайней мере, сначала проверять гипотезу H_1 о наличии общих множителей, а затем, если она не отвергается, проверять гипотезу H_2: $a_1 = 0$ о некоррелированности ошибок в статической модели регрессии.

Пусть каждая из этих гипотез проверяется на уровне значимости α, скажем, $\alpha = 0.05$. Какова в такой ситуации вероятность ошибочного отвержения модели статической регрессии? Имеем:

P{ошибочно отвергается хотя бы одна из гипотез H_1, H_2} $\le$

$\le P${ошибочно отвергается H_1} + P{ошибочно отвергается H_2} =

$= \alpha + \alpha = 2\alpha$.

Следовательно, если положить $\alpha = 0.025$, то вероятность отвержения модели статической регрессии в рамках двухступенчатой процедуры не будет превышать 0.05. Заметим, что при этом мы еще не принимали в расчет ошибки, связанные с возможностью неправильной диагностики общей модели.