Линейные процессы: свойства и асимптотические результаты*

Вадим Мармер†

Университет Британской Колумбии, Ванкувер, Канада

Настоящее эссе содержит обзор результатов, касающихся линейных временных рядов и включающих разложение Вольда, свойства спектральных плотностей и операторов сдвига, моделей ARMA, разложение Бевериджа–Нельсона и метод получения асимптотических результатов Филлипса–Соло.

1 Введение

В эконометрической литературе можно выделить два основных подхода к получению асимптотических результатов. Первый подход базируется на предельных теоремах для временных рядов, удовлетворяющих условиям сильного перемешивания. Подробное описание данного подхода дано, например, в Davidson (1994), а также White (2001), причем в последней книге акцент делается на приложениях. Причина популярности концепций перемешивания кроется в том, что это понятие позволяет включить как зависимость, так и неоднородность, так как временные ряды, обладающие свойствами перемешивания, не обязаны быть стационарными. В то же время, этот подход не универсален, так как даже авторегрессионные (AR) процессы первого порядка могут не быть сильно перемешивающими, см. Andrews (1984).

Второй подход, являющийся предметом настоящей обзорной статьи, базируется на линейных процессах $\{X_t : t = 1, 2, \ldots\}$ вида

$$X_t = \sum_{j=1}^{\infty} c_j \varepsilon_{t-j}, \tag{1}$$

где c_j — константы, а $\{\varepsilon_t : t = 1, 2, \ldots\}$ — серийно некоррелированный процесс. Данная модель предоставляет очень интуитивный способ генерации серийно зависимых данных: начиная с серийно некоррелированного процесса $\{\varepsilon_t\}$, можно получить временную зависимость в $\{X_t\}$ путем усреднения случайных величин ε_t.

Асимптотический метод для линейных процессов был разработан в Phillips & Solo (1992) и основан на алгебраическом разложении временного ряда на долгосрочную и переходную компоненты. Разложение было предложено в Beveridge & Nelson (1981) и известно в литературе как разложение Бевериджа–Нельсона (БН). Подход Филлипса и Соло привлекатален тем, что опирается на законы больших чисел (ЗБЧ) и центральные предельные теоремы (ЦПТ) для независимых процессов или мартингал-разностей. Используя линейную структуру процесса и соответствующее БН-разложение, Филлипс и Соло предлагают применить предельные теоремы для независимых процессов или мартингал-разностей к долгосрочной компоненте и показать, что переходной компонентой можно асимптотически пренебречь. Этот метод можно использовать для классических линейных процессов, таких как ARMA-процессы, а также процессов скользящего среднего (MA) бесконечного порядка.[1]

*Цитировать как: Мармер, Вадим (2012) «Линейные процессы: свойства и асимптотические результаты», Квантиль, №10, стр. 33–56. Citation: Marmer, Vadim (2012) "Linear processes: properties and asymptotic results," Quantile, No.10, pp. 33–56.

†Адрес: Department of Economics, University of British Columbia, 997 – 1873 East Mall, Vancouver, BC, V6T 1Z1, Canada. Электронная почта: *vadim.marmer@ubc.ca*

[1] Как в уравнении (1).

Несмотря на свою простоту, метод Филлипса и Соло обладает замечательной общностью в силу знаменитого разложения Вольда, в котором слабостационарные процессы представляются линейными процессами, см., например, Hansen & Sargent (1991). Разложение Вольда часто служит оправданием для моделирования временных рядов как линейных процессов. Если необходимо немного усилить результат разложения Вольда, можно применять метод Филлипса и Соло.

Настоящее эссе имеет далее следующую структуру. В разделе 2 обсуждается разложение Вольда. В разделе 3 описываются некоторые важные инструменты, необходимые для работы с линейными процессами. Их использование иллюстрируется в разделе 4 на примере ARMA-моделей. Наконец, в разделе 5 описывается БН-разложение и асимптотический метод Phillips & Solo (1992).

Помимо цитированных в списке литературы источников, эссе основано также на неопубликованных рукописных лекциях Питера Филлипса.[2]

2 Разложение Вольда

Сфокусируем внимание на слабостационарных процессах, удовлетворяющих следующему определению.

Определение 1 (Слабая стационарность и автоковариационная функция) *Процесс* $\{X_t\}$, *такой что* $\mathbb{E}X_t^2 < \infty$ *для всех* t, *называется слабостационарным, если* $\mathbb{E}X_t = \mu$ *и* $\mathbb{C}(X_t, X_s) = \gamma(|s-t|)$ *для всех* s *и* t, *где* $\gamma(j)$, $j = 0, 1, \ldots$, *называется автоковариационной функцией* $\{X_t\}$.

Слабостационарный серийно некоррелированный процесс с нулевым средним называется белым шумом (БШ). Процессы БШ является строительными блоками в конструкции MA-процессов:

Определение 2 (Белый шум) *Процесс* $\{\varepsilon_t\}$ *называется БШ, если* $\mathbb{E}\varepsilon_t = 0$, $\mathbb{E}\varepsilon_t^2 = \sigma^2 < \infty$ *и* $\mathbb{E}\varepsilon_t\varepsilon_{t-j} = 0$ *для всех* t *и* $j \neq 0$.

Определение 3 (MA) *Процесс* $\{X_t\}$ *называется процессом скользящего среднего порядка* q *(MA(q)), если*

$$X_t = c_0\varepsilon_t + c_1\varepsilon_{t-1} + \ldots + c_q\varepsilon_{t-q}, \tag{2}$$

и $\{\varepsilon_t\}$ *есть БШ.*

Расширение до бесконечного порядка (2), как в уравнении (1), обозначается MA(∞). Разложение Вольда показывает, что любой слабостационарный процесс с нулевым средним и абсолютно суммируемыми автоковариациями может быть представлен в форме MA(∞).

Обозначим за L_2 пространство случайных величин с конечными вторыми моментами. Для $X, Y \in L_2$ определим скалярное произведение

$$\langle X, Y\rangle = \mathbb{E}XY.$$

С таким определением скалярного произведения пространство L_2 становится гильбертовым. Рассмотрим слабостационарный процесс $\{X_t\}$ с нулевым средним, такой что

$$\sum_{j=0}^{\infty} |\gamma(j)| < \infty, \tag{3}$$

[2] Лекции доступны по адресу `http://cowles.econ.yale.edu/korora/phillips/teach/lec-notes.htm`.

где $\gamma(j)$ обозначает автоковариационную функцию

$$\gamma(j) = \mathbb{E}X_t X_{t-j}.$$

Определим $\mathcal{M}_t$ как наименьшее замкнутое подпространство пространства L_2, содержащее все элементы вида

$$\sum_{j=0}^{\infty} c_j X_{t-j}, \text{ такие что } \sum_{j=0}^{\infty} c_j^2 < \infty.$$

Требование $\sum_{j=0}^{\infty} c_j^2 < \infty$ необходимо для того, чтобы элементы $\mathcal{M}_t$ принадлежали L_2. Действительно, в этом случае дисперсию любого элемента из $\mathcal{M}_t$ можно ограничить суммами $\sum_{j=0}^{\infty} c_j^2$ и $\sum_{j=0}^{\infty} |\gamma(j)|$, как показано ниже:

$$\begin{aligned}
\mathbb{V}\left(\sum_{j=0}^{\infty} c_j X_{t-j}\right) &= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} c_i c_j \gamma(i-j) \leq 2\sum_{h=0}^{\infty}\sum_{j=0}^{\infty} c_j c_{j+h}\gamma(h) = 2\sum_{h=0}^{\infty}\gamma(h)\sum_{j=0}^{\infty} c_j c_{j+h} \\
&\leq 2\sum_{h=0}^{\infty}|\gamma(h)|\left|\sum_{j=0}^{\infty} c_j c_{j+h}\right| \leq 2\sum_{h=0}^{\infty}|\gamma(h)|\left(\sum_{j=0}^{\infty} c_j^2\right)^{1/2}\left(\sum_{j=0}^{\infty} c_{j+h}^2\right)^{1/2} \\
&\leq 2\left(\sum_{j=0}^{\infty} c_j^2\right)\sum_{h=0}^{\infty}|\gamma(h)| \qquad (4) \\
&< \infty,
\end{aligned}$$

где последнее неравенство выполнено, если верно неравенство (3). Заметим далее, что последовательность $\mathcal{M}_t$ возрастает:

$$\ldots \subset \mathcal{M}_t \subset \mathcal{M}_{t+1} \subset \ldots.$$

Пусть $P_{\mathcal{M}_t}$ – ортогональная проекция на $\mathcal{M}_t$. Можно записать

$$X_t = \widehat{X}_t + \varepsilon_t, \qquad (5)$$

где

$$\begin{aligned}
\widehat{X}_t &= P_{\mathcal{M}_{t-1}} X_t, \\
\varepsilon_t &= \left(1 - P_{\mathcal{M}_{t-1}}\right) X_t.
\end{aligned}$$

Из известных результатов для гильбертовых пространств следует, что $\widehat{X}_t$ решает задачу наименьших квадратов. Значит, $\widehat{X}_t$ можно интерпретировать как лучший (в среднеквадратическом смысле) линейный предиктор величины X_t на один шаг вперед, а величину ε_t – как ошибку предсказания. Заметим, что $\widehat{X}_t \in \mathcal{M}_{t-1}$, и $\varepsilon_t \in \mathcal{M}_t$, так как

$$\varepsilon_t = X_t - \widehat{X}_t.$$

Более того, в силу свойств ортогональной проекции

$$\varepsilon_t \in \mathcal{M}_{t-1}^{\perp},$$

где

$$\mathcal{M}_t^{\perp} = \{Y \in L_2 : \langle Y, X\rangle = 0 \text{ для всех } X \in \mathcal{M}_t\}.$$

Так как последовательность $\mathcal{M}_t$ возрастает, она содержит все элементы из $\mathcal{M}_{t-h}$, и мы получаем, что $\varepsilon_t \in \mathcal{M}_{t-h}^{\perp}$ для всех $h \geq 1$. Следовательно,

$$\mathbb{E}\varepsilon_t\varepsilon_{t-h} = 0 \text{ для всех } h \geq 1.$$

Более того, дисперсия равна

$$\mathbb{E}\varepsilon_t^2 = \sigma^2 < \infty$$

и не зависит от t, так как процесс $\{X_t\}$ слабостационарен.

Определим

$$\mathcal{E}_t = \left\{ \sum_{j=0}^{\infty} c_j \varepsilon_{t-j} : \sum_{j=0}^{\infty} c_j^2 < \infty \right\},$$

где $\mathcal{E}_t$ является на самом деле замкнутым линейным подпространством пространства L_2. Обозначим за $P_{\mathcal{E}_t}$ ортогональную проекцию на $\mathcal{E}_t$. Имеем:

$$X_t = P_{\mathcal{E}_t} X_t + V_t, \tag{6}$$

где в силу замкнутости и линейности $\mathcal{E}_t$

$$P_{\mathcal{E}_t} X_t = \sum_{j=0}^{\infty} \widehat{c}_j \varepsilon_{t-j} \tag{7}$$

для некоторой последовательности $\{\widehat{c}_j\}$, и

$$V_t = (1 - P_{\mathcal{E}_t}) X_t.$$

Заметим, что $V_t \in \mathcal{M}_t$, и в силу (5)

$$\mathcal{M}_t = \mathcal{M}_{t-1} \oplus \mathcal{S}(\varepsilon_t),$$

где $\mathcal{S}(\varepsilon_t)$ — линейное подпространство, порожденное ε_t, а $\oplus$ обозначает прямую сумму:

$$\mathcal{S}_1 \oplus \mathcal{S}_2 = \{x_1 + x_2 : x_1 \in \mathcal{S}_1, x_2 \in \mathcal{S}_2\}.$$

Так как $P_{\mathcal{E}_t}$ — ортогональная проекция, $\mathbb{E}V_t\varepsilon_t = 0$ в силу (6) и (7), и значит $V_t \notin \mathcal{S}(\varepsilon_t)$. Следовательно, должно быть выполнено включение $V_t \in \mathcal{M}_{t-1}$. Аналогично, так как $\mathcal{M}_{t-1} = \mathcal{M}_{t-2} \oplus \mathcal{S}(\varepsilon_{t-1})$ и $\mathbb{E}V_t\varepsilon_{t-1} = 0$, должны быть выполнены включения $V_t \in \mathcal{M}_{t-1}, \mathcal{M}_{t-2}, \ldots$. Пусть

$$\mathcal{M}_{-\infty} = \cap_{t=-\infty}^{\infty} \mathcal{M}_t.$$

Тогда

$$V_t \in \mathcal{M}_{-\infty} \text{ для всех } t.$$

Таким образом, V_t является элементом любого линейного подпространства $\mathcal{M}_s$, $s \in \mathbb{Z}$, где $\mathbb{Z}$ обозначает множество целых чисел. Весь процесс $\{V_t\}$ можно точно предсказать на основе произвольно далекой истории процесса $\{X_t\}$. Такие процессы называются *детерминистическими.*

Мы вывели разложение Вольда для слабостационарного процесса $\{X_t\}$:

$$X_t = \sum_{j=0}^{\infty} c_j \varepsilon_{t-j} + V_t,$$

где $\{\varepsilon_t\}$ — БШ, а V_t — детерминистический процесс. Такое представление единственно почти наверное.

Можно показать, что

$$c_j = \mathbb{E}X_t\varepsilon_{t-j}/\sigma^2 \tag{8}$$

и

$$c_0 = 1, \tag{9}$$

где $\sigma^2 = \mathbb{V}(\varepsilon_t)$.[3]

Если $V_t = 0$ для всех t, процесс $\{X_t\}$ называется *чисто недетерминистическим.* Разумно предположить, что экономические временные ряды не содержат детерминистических компонент, которые можно бы было точно предсказать на основе очень далеких историй, а значит, являются чисто недетерминистическими.

Возникает желание дать величинам ε структурную экономическую интерпретацию и считать, что они отвечают за фундаментальные шоки переменной X. В этом случае, так как величины ε серийно некоррелированы, эффектом шока после j периодов можно считать частную производную $\partial X_t/\partial\varepsilon_{t-j} = c_j$, которая называется *импульсным откликом.* К сожалению, такая интерпретация инкрементов в разложении Вольда редко бывает оправданной. Заметим, что по построению величины ε являются ошибками *линейного* прогноза на один шаг вперед. Предположим, что фундаментальные шоки являются неожиданными шоками для информационного множества агента. Пусть $\mathcal{F}_t = \sigma\left(X_t, X_{t-1}, \ldots\right)$ — σ-алгебра, порожденная историей X до времени t. Эта σ-алгебра представляет информацию агентов в экономике до момента времени t. Однако, $\widehat{X}_t$ в уравнении (5) должен быть линейным предиктором, и значит, необязательно равен $\mathbb{E}\left(X_t|\mathcal{F}_{t-1}\right)$. Таким образом, истинные фундаментальные шоки могут отличаться от шоков в разложении Вольда. См. обсуждение в главе 4 книги Hansen & Sargent (1991).

3 Инструментарий линейных процессов

Обозначим за $X_1, X_2, \ldots, X_n$ наблюдения за случайной величиной X_t. Тогда ЦПТ формулируется следующим образом:

$$\frac{n^{-1/2}\sum_{t=1}^{n}\left(X_t - \mathbb{E}X_t\right)}{\sqrt{\mathbb{V}\left(n^{-1/2}\sum_{t=1}^{n}X_t\right)}} \to_d N(0,1).$$

Дисперсия $\mathbb{V}\left(n^{-1}\sum_{t=1}^{n}X_t\right)$ представляет особый интерес. С серийно коррелированными данными,

$$\begin{aligned}\mathbb{V}\left(n^{-1/2}\sum_{t=1}^{n}X_t\right) &= n^{-1}\sum_{t=1}^{n}\sum_{s=1}^{n}\mathbb{C}(X_t, X_s)\\ &= n^{-1}\sum_{t=1}^{n}\mathbb{V}(X_t) + 2\sum_{j=1}^{n-1}n^{-1}\sum_{t=j+1}^{n}\mathbb{C}(X_t, X_{t-j}).\end{aligned}$$

Пусть $\gamma(j) = \mathbb{C}(X_t, X_{t-j})$. Предполагая, что $\{X_t\}$ слабостационарен, для больших n дисперсию можно аппроксимировать выражением

$$\omega \equiv \lim_{n\to\infty}\mathbb{V}\left(n^{-1/2}\sum_{t=1}^{n}X_t\right) = \gamma(0) + 2\sum_{j=1}^{\infty}\gamma(j),$$

[3]Запишем $X_t\varepsilon_{t-j} = \left(\sum_{j=0}^{\infty}c_j\varepsilon_{t-j} + V_t\right)\varepsilon_{t-j}$. Из этого следуют уравнения (8) и (9), так как $\{\varepsilon_t\}$ — БШ и $\mathbb{E}V_t\varepsilon_{t-j} = 0$ по построению.

при условии, что предел существует. В этом определении ω обозначает так называемую *долгосрочную дисперсию* процесса $\{X_t\}$.

Определение 4 (Долгосрочная дисперсия) *Для слабостационарного процесса с автоковариационной функцией* $\gamma(j)$, $j = 0, 1, \ldots$, *долгосрочная дисперсия (если существует) задается выражением* $\gamma(0) + 2\sum_{j=1}^{\infty} \gamma(j)$.

Заметим, что необходимым и достаточным условием существования долгосрочной дисперсии является суммируемость

$$\sum_{j=0}^{\infty} |\gamma(j)| < \infty.$$

Ясно, что автоковариационная функция $\gamma(j)$ играет важную роль в эконометрике временных рядов, так как в пределе зависимость по времени отражается долгосрочной дисперсией:

$$n^{-1/2} \sum_{t=1}^{n} (X_t - \mathbb{E}X_t) \to_d N\left(0, \gamma(0) + 2\sum_{j=1}^{\infty} \gamma(j)\right).$$

3.1 Спектральная плотность[4]

Зависимость по времени асимптотически отражается автоковариационной функцией. Удобно агрегировать эту зависимость так называемой функцией спектральной плотности. В то время как автоковариационная функция определена на бесконечной последовательности действительных чисел, спектральную плотность достаточно рассматривать на компактном интервале.

Определение 5 (Спектральная плотность) *Для автоковариационной функции* $\gamma(j)$, $j = 0, 1, \ldots$, *соответствующая спектральная плотность определяется как*

$$f(\lambda) = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \gamma(j) e^{-i\lambda j},$$

где $i = \sqrt{-1}$.

Так как $\gamma(j) = \gamma(-j)$, спектральная плотность действительнозначна.

$$\begin{aligned}
f(\lambda) &= \frac{1}{2\pi}\left(\gamma(0) + \sum_{j=-\infty}^{-1} \gamma(j) e^{-i\lambda j} + \sum_{j=1}^{\infty} \gamma(j) e^{-i\lambda j}\right) \\
&= \frac{1}{2\pi}\left(\gamma(0) + \sum_{j=1}^{\infty} \gamma(-j) e^{i\lambda j} + \sum_{j=1}^{\infty} \gamma(j) e^{-i\lambda j}\right) \\
&= \frac{1}{2\pi}\left(\gamma(0) + \sum_{j=1}^{\infty} \gamma(j)\left(e^{i\lambda j} + e^{-i\lambda j}\right)\right).
\end{aligned}$$

Далее, $e^{i\lambda j} = \cos(\lambda j) + i\sin(\lambda j)$, $e^{-i\lambda j} = \cos(\lambda j) - i\sin(\lambda j)$, и, стало быть,

$$e^{i\lambda j} + e^{-i\lambda j} = 2\cos(\lambda j).$$

[4] Материал этого раздела основан на главе 5 книги Gourieroux & Monfort (1995) и главе 6 книги Hamilton (1994).

Значит,

$$f(\lambda) = \frac{1}{2\pi}\left(\gamma(0) + 2\sum_{j=1}^{\infty}\gamma(j)\cos(\lambda j)\right) = \frac{1}{2\pi}\sum_{j=-\infty}^{\infty}\gamma(j)\cos(\lambda j)\,.$$

Так как $\cos(\lambda j) = \cos(-\lambda j)$, спектральная плотность симметрична относительно нуля. Более того, так как функция cos периодична с периодом 2π, область значений спектральной плотности определяется значениями $f(\lambda)$ в точках $0 \le \lambda \le \pi$.

Автоковариационная функция и спектральная плотность эквивалентны в том смысле, что всегда можно однозначно восстановить автоковариационную функцию по спектральной плотности. Таким образом, они содержат одинаковое количество информации, что демонстрируется следующим результатом.

Теорема 1 *Пусть* $\sum_{h=-\infty}^{\infty}|\gamma(h)| < \infty$. *Тогда* $\gamma(j) = \int_{-\pi}^{\pi} f(\lambda)\, e^{i\lambda j} d\lambda$.

Доказательство. Имеем

$$\int_{-\pi}^{\pi} f(\lambda)\, e^{i\lambda j} d\lambda = \int_{-\pi}^{\pi}\left(\frac{1}{2\pi}\sum_{h=-\infty}^{\infty}\gamma(h)e^{-i\lambda h}\right)e^{i\lambda j}d\lambda = \frac{1}{2\pi}\sum_{h=-\infty}^{\infty}\gamma(h)\int_{-\pi}^{\pi}e^{i\lambda(j-h)}d\lambda,$$

где суммирование и интегрирование можно менять местами, так как $\sum_{h=-\infty}^{\infty}|\gamma(h)| < \infty$. Далее,

$$\int_{-\pi}^{\pi} e^{i\lambda(j-h)}d\lambda = 2\pi \text{ если } j = h.$$

Для $j \neq h$

$$\begin{aligned}
\int_{-\pi}^{\pi} e^{i\lambda(j-h)}d\lambda &= \int_{-\pi}^{\pi}\left(\cos(\lambda(j-h)) + i\sin(\lambda(j-h))\right)d\lambda \\
&= \left.\frac{\sin(\lambda(j-h))}{j-h}\right|_{-\pi}^{\pi} - i\left.\frac{\cos(\lambda(j-h))}{j-h}\right|_{-\pi}^{\pi} \\
&= \frac{\sin(\pi(j-h)) - \sin(-\pi(j-h))}{j-h} - i\frac{\cos(\pi(j-h)) - \cos(-\pi(j-h))}{j-h}.
\end{aligned}$$

Однако, так как функции cos и sin периодичны с периодом 2π,

$$\cos(\pi(j-h)) = \cos(-\pi(j-h) + 2\pi(j-h)) = \cos(-\pi(j-h))\,.$$

Следовательно,

$$\int_{-\pi}^{\pi} e^{i\lambda(j-h)}d\lambda = 0 \text{ если } j \neq h.$$

■

В частности, из теоремы 1 следует, что

$$\gamma(0) = \int_{-\pi}^{\pi} f(\lambda)\, d\lambda.$$

Таким образом, площадь под спектральной плотностью процесса X_t между точками $-\pi$ и π равна дисперсии величины X_t.

Аргумент λ функции $f(\lambda)$ называется частотой. Заметим, что если $\{X_t\}$ слабостационарен с абсолютно суммируемыми автоковариациями, долгосрочная дисперсия определяется значением спектральной плотности в нулевой частоте:

$$\omega_X = \lim_{n\to\infty} \mathbb{V}\left(n^{-1/2}\sum_{t=1}^{n} X_t\right) = \sum_{h=-\infty}^{\infty}\gamma(h) = 2\pi f(0).$$

Далее мы обсудим, как линейные (MA) преобразования слабостационарных процессов влияют на спектральную плотность и долгосрочную дисперсию.

Теорема 2 *Пусть $\{X_t\}$ — слабостационарный процесс с автоковариационной функцией γ_X, такой что $\sum_{j=-\infty}^{\infty}|\gamma_X(j)| < \infty$. Определим $Y_t = \sum_{j=0}^{\infty} c_j X_{t-j}$, где $\sum_{j=0}^{\infty} c_j^2 < \infty$. Тогда процесс $\{Y_t\}$ слабостационарен, а его спектральная плотность $f_Y(\lambda) = \left|\sum_{j=0}^{\infty} c_j e^{-i\lambda j}\right|^2 f_X(\lambda)$, где f_X — спектральная плотность процесса $\{X_t\}$.*

Доказательство. Имеем

$$\begin{aligned}\mathbb{C}(Y_t, Y_{t-h}) &= \mathbb{C}\left(\sum_{j=0}^{\infty} c_j X_{t-j}, \sum_{j=0}^{\infty} c_j X_{t-h-j}\right) = \sum_{j=0}^{\infty}\sum_{k=0}^{\infty} c_j c_k \mathbb{C}(X_{t-j}, X_{t-h-k}) \\ &= \sum_{j=0}^{\infty}\sum_{k=0}^{\infty} c_j c_k \gamma_X(h+k-j).\end{aligned}$$

Таким образом, $\mathbb{C}(Y_t, Y_{t-h})$ не зависит от t. Более того, аналогично выводу неравенства (4),

$$\sum_{j=0}^{\infty}\sum_{k=0}^{\infty} c_j c_k \gamma_X(h+k-j) \le 2\left(\sum_{j=0}^{\infty} c_j^2\right)\sum_{h=0}^{\infty}|\gamma_X(h)| < \infty.$$

Следовательно, $\{Y_t\}$ слабостационарен.

Далее,

$$\begin{aligned}f_Y(\lambda) &= \frac{1}{2\pi}\sum_{h=-\infty}^{\infty}\mathbb{C}(Y_t, Y_{t-h})\, e^{-i\lambda h} = \frac{1}{2\pi}\sum_{h=-\infty}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty} c_j c_k \gamma_X(h+k-j)\, e^{-i\lambda h} \\ &= \sum_{j=0}^{\infty} c_j e^{-i\lambda j}\sum_{k=0}^{\infty} c_k e^{i\lambda k}\frac{1}{2\pi}\sum_{h=-\infty}^{\infty}\gamma_X(h+k-j)\, e^{-i\lambda(h+k-j)} \\ &= \left(\sum_{j=0}^{\infty} c_j e^{-i\lambda j}\right)\left(\sum_{j=0}^{\infty} c_j e^{i\lambda j}\right) f_X(\lambda) = \left|\sum_{j=0}^{\infty} c_j e^{-i\lambda j}\right|^2 f_X(\lambda).\end{aligned}$$

Последнее неравенство верно в силу

$$\sum_j c_j e^{-i\lambda j} = \sum_j c_j(\cos(\lambda j) - i\sin(\lambda j)) = \sum_j c_j\cos(\lambda j) - i\sum_j c_j\sin(\lambda j).$$

Комплексным сопряжением получаем

$$\sum_j c_j\cos(\lambda j) + i\sum_j c_j\sin(\lambda j) = \sum_j c_j(\cos(\lambda j) + i\sin(\lambda j)) = \sum_j c_j e^{i\lambda j},$$

и значит,

$$\left|\sum_{j=0}^{\infty} c_j e^{-i\lambda j}\right|^2 = \left(\sum_{j=0}^{\infty} c_j \cos(\lambda j)\right)^2 + \left(\sum_{j=0}^{\infty} c_j \sin(\lambda j)\right)^2 = \left(\sum_{j=0}^{\infty} c_j e^{-i\lambda j}\right)\left(\sum_{j=0}^{\infty} c_j e^{i\lambda j}\right).$$

■

В данной теореме спектральная плотность в нуле и, стало быть, долгосрочная дисперсия конечны, если $\sum_{j=0}^{\infty} |c_j| < \infty$. Однако абсолютная суммируемость, $\sum_{j=0}^{\infty} |c_j| < \infty$, является более сильным предположением, чем суммируемость квадратов, $\sum_{j=0}^{\infty} c_j^2 < \infty$, как мы сейчас покажем. Пусть $\sum_{j=0}^{\infty} |c_j| < \infty$. Из $\sum_{j=0}^{\infty} |c_j| < \infty$ следует, что $c_j \to 0$ при $j \to \infty$. Значит, последовательность $\{c_j\}$ равномерно ограничена. Наконец, $\sum_{j=0}^{\infty} c_j^2 \le \sup_j |c_j| \sum_{j=0}^{\infty} |c_j| < \infty$.

Предположим, что процесс $\{X_t\}$ слабостационарен и чисто недетерминистичен. Тогда в силу разложения Вольда он имеет MA(∞) представление

$$X_t = \sum_{j=0}^{\infty} a_j \varepsilon_{t-j}, \tag{10}$$

где $\{\varepsilon_t\}$ — БШ, и $\sum_{j=0}^{\infty} a_j^2 < \infty$. Пусть $\mathbb{V}(\varepsilon_t) = \sigma^2$. Так как спектр БШ плоский:

$$f(\lambda) = \frac{\sigma^2}{2\pi} \text{ для всех } \lambda,$$

из теоремы 2 следует, что

$$f_X(\lambda) = \frac{\sigma^2}{2\pi}\left|\sum_{j=0}^{\infty} a_j e^{-i\lambda j}\right|^2,$$

а долгосрочная дисперсия процесса $\{X_t\}$ равна

$$\omega_X = 2\pi f_X(0) = \sigma^2 \left(\sum_{j=0}^{\infty} a_j\right)^2.$$

Если рассматривать уравнение (10) как механизм генерации, то условие $\sum_{j=0}^{\infty} a_j^2 < \infty$ гарантирует, что $\{X_t\}$ слабостационарен. Однако достаточным условием конечности долгосрочной дисперсии является суммируемость $\sum_{j=0}^{\infty} |a_j| < \infty$. Если последнее условие не выполняется, может случиться, что $\sum_{j=-\infty}^{\infty} |\gamma_X(j)| = \infty$. Говорят, что такие процессы обладают *длинной памятью*. Если $\sum_{j=0}^{\infty} a_j^2 < \infty$ выполнено для процесса с длинной памятью, то автоковариационная функция сходится к нулю слишком медленно для того, чтобы долгосрочная дисперсия была конечна.

Определение 6 (Короткая и длинная память) *Слабостационарный процесс с автоковариационной функцией $\gamma(j)$, $j = 0, 1, \ldots$, обладает короткой памятью, если $\sum_{j=0}^{\infty} |\gamma(j)| < \infty$. Длинная память имеет место в случае $\sum_{j=0}^{\infty} |\gamma(j)| = \infty$.*

Пусть $\{Y_t\}$ определен как в теореме 2. Тогда его спектральная плотность и долгосрочная дисперсия равны

$$\begin{aligned} f_Y(\lambda) &= \frac{\sigma^2}{2\pi}\left|\sum_{j=0}^{\infty} a_j e^{-i\lambda j}\right|^2 \left|\sum_{j=0}^{\infty} c_j e^{-i\lambda j}\right|^2, \\ \omega_Y &= \sigma^2 \left(\sum_{j=0}^{\infty} a_j\right)^2 \left(\sum_{j=0}^{\infty} c_j\right)^2. \end{aligned}$$

3.2 Оператор сдвига[5]

Оператор сдвига L отображает процесс $\{X_t\}$ в себя, так что

$$\begin{aligned} LX_t &= X_{t-1}, \\ L^2X_t &= LLX_t = LX_{t-1} = X_{t-2}, \\ &\dots \\ L^hX_t &= X_{t-h}. \end{aligned}$$

Многочлен от оператора сдвига

$$C(L) = \sum_{j=0}^{\infty} c_j L^j$$

преобразует $\{X_t\}$ в другой процесс $\{Y_t\}$, такой что

$$Y_t = C(L)\,X_t = \sum_{j=0}^{\infty} c_j L^j X_t = \sum_{j=0}^{\infty} c_j X_{t-j}.$$

Пусть $A(L) = \sum_{j=0}^{\infty} a_j L^j$ и $B(L) = \sum_{j=0}^{\infty} b_j L^j$. Тогда

$$A(L) + B(L) = \sum_{j=0}^{\infty} (a_j + b_j)\, L^j,$$

и

$$\begin{aligned} A(L)B(L) &= \left(\sum_{j=0}^{\infty} a_j L^j\right)\left(\sum_{j=0}^{\infty} b_j L^j\right) = \sum_{j=0}^{\infty}\sum_{h=0}^{\infty} a_j b_h L^{j+h} \\ &= a_0b_0 + (a_0b_1 + b_0a_1)\,L + (a_1b_1 + a_0b_2 + a_2b_0)\,L^2 + \dots. \end{aligned}$$

Имеем:

$$\begin{aligned} A(L) + B(L) &= B(L) + A(L), \\ A(L)B(L) &= B(L)A(L). \end{aligned}$$

При определенных условиях многочлен от оператора сдвига может быть обращен. Обратный к многочлену $C(L)$ есть еще один многочлен от оператора сдвига, скажем, $B(L)$, такой что

$$C(L)\,B(L) = 1. \tag{11}$$

Можно записать

$$C(L)^{-1} = B(L).$$

Обращение многочленов от оператора сдвига важно по следующей причине. Рассмотрим, например, *авторегрессионный процесс первого порядка* AR(1):

$$X_t = cX_{t-1} + \varepsilon_t. \tag{12}$$

[5]Материал этого раздела основан на главе 5 книги Gourieroux & Monfort (1995) и главе 2 книги Hamilton (1994).

Этот процесс рекурсивно генерируется на основе экзогенного белого шума $\{\varepsilon_t\}$, начального значения X_0 (случайной величины с дисперсией $\mathbb{V}(X_t)$, которую мы рассчитаем позже), и коэффициента c:

$$\begin{aligned} X_1 &= cX_0 + \varepsilon_1, \\ X_2 &= cX_1 + \varepsilon_2, \end{aligned}$$

и.т.д. Таким образом, процесс $\{X_t\}$ является эндогенным решением *разностного уравнения* (12). Это разностное уравнение может быть переписано в виде $X_t - cX_{t-1} = \varepsilon_t$ или

$$\begin{aligned} C(L) X_t &= \varepsilon_t, \\ C(L) &= 1 - cL. \end{aligned}$$

Если $C(L)$ можно обратить, то решение может быть записано как $X_t = C(L)^{-1} \varepsilon_t$. Таким образом, важно вывести условия обратимости многочлена от оператора сдвига и найти способ вычисления коэффициентов обратного многочлена.

Рассмотрим для начала многочлен порядка 1. Без ограничения общности можно положить коэффициент при L^0 единицей $c_0 = 1$:[6]

$$C(L) = 1 - cL.$$

Предположим, что $|c| < 1$. Тогда можно определить многочлен, обратный к $1 - cL$, следующим образом:

$$(1 - cL)^{-1} = 1 + cL + c^2L^2 + \dots. \tag{13}$$

Это определение удовлетворяет (11), так как

$$(1 - cL) \sum_{j=0}^{\infty} c^j L^j = 1.$$

Решение (13) не единственное, удовлетворяющее условию (11). Добавление к нему члена Vc^t, где V — некоторая случайная величина, тоже удовлетворяет ему, так как

$$(1 - cL) Vc^t = Vc^t - VcLc^t = Vc^t - Vcc^{t-1} = 0.$$

Однако если взять $(1 - cL)^{-1} = \sum_{j=0}^{\infty} c^j L^j + Vc^t$, то процесс $(1 - cL)^{-1} \varepsilon_t = \sum_{j=0}^{\infty} c^j \varepsilon_{t-j} + Vc^t\varepsilon_t$ не будет слабостационарным. Поэтому мы накладываем ограничение $V = 0$.

Далее, рассмотрим многочлен от оператора сдвига порядка 2:

$$C(L) = 1 - c_1L - c_2L^2.$$

Мы можем разложить многочлен на множители:

$$1 - c_1L - c_2L^2 = (1 - \lambda_1 L)(1 - \lambda_2 L), \tag{14}$$

где

$$\begin{aligned} c_1 &= \lambda_1 + \lambda_2, \\ c_2 &= -\lambda_1\lambda_2. \end{aligned}$$

[6]Это следует из разложения Вольда. С другой стороны, это можно рассматривать как нормализацию: умножение всех коэффициентов на некоторую константу $c_0 \neq 0$ влияет только на дисперсию процесса.

Другой способ нахождения λ_1 и λ_2 состоит в следующем. Пусть z_1 и z_2 — решения (возможно, комплексные) уравнения

$$1 - c_1 z - c_2 z^2 = 0. \tag{15}$$

Мы можем переписать уравнение (15) в виде

$$0 = 1 - c_1 z - c_2 z^2 = (z_1 - z)(z_2 - z),$$

и, сравнивая последнее уравнение с (14), получаем

$$\lambda_1 = z_1^{-1} \text{ и } \lambda_2 = z_2^{-1}.$$

Многочлен в уравнении (14) можно обратить, если $|\lambda_1| < 1$ и $|\lambda_2| < 1$. Эти условия эквивалентны требованию, чтобы все корни многочлена в уравнении (15) лежали вне единичного круга:

$$|z| > 1.$$

Если это условие выполняется, то

$$\left(1 - c_1 L - c_2 L^2\right)^{-1} = (1 - \lambda_1 L)^{-1} (1 - \lambda_2 L)^{-1} = \left(\sum_{j=0}^{\infty} \lambda_1^j L^j\right) \left(\sum_{j=0}^{\infty} \lambda_2^j L^j\right).$$

С другой стороны, можно записать

$$\frac{1}{1 - \lambda_1 L} \frac{1}{1 - \lambda_2 L} = \frac{1}{\lambda_1 - \lambda_2} \left(\frac{\lambda_1}{1 - \lambda_1 L} - \frac{\lambda_2}{1 - \lambda_2 L}\right).$$

Тогда

$$\left(1 - c_1 L - c_2 L^2\right)^{-1} = \frac{1}{\lambda_1 - \lambda_2} \sum_{j=0}^{\infty} \left(\lambda_1 \lambda_1^j L^j - \lambda_2 \lambda_2^j L^j\right) = \frac{1}{\lambda_1 - \lambda_2} \sum_{j=0}^{\infty} \left(\lambda_1^{j+1} - \lambda_2^{j+1}\right) L^j.$$

Этот результат может быть обобщен на случай многочлена порядка p,

$$C(L) = 1 - c_1 L - \ldots - c_p L^p,$$

так как мы можем разложить его как

$$1 - c_1 L - \ldots - c_p L^p = \prod_{j=1}^{p} (1 - \lambda_j L),$$

где числа λ_j удовлетворяют

$$1 - c_1 z - \ldots - c_p z^p = \prod_{j=1}^{p} (1 - \lambda_j z).$$

Многочлен можно обратить, если корни $1 - c_1 z - \ldots - c_p z^p$ лежат вне единичного круга. Тогда

$$(1 - c_1 L - \ldots - c_p L^p)^{-1} = \prod_{j=1}^{p} (1 - \lambda_j L)^{-1}. \tag{16}$$

3.3 Некоторые обобщения на случай многомерных процессов

До настоящего момента все результаты были сформулированы для одномерных процессов, однако они могут быть обобщены на векторный случай, как показано далее. Предположим, что векторный процесс $\{X_t\}$ слабостационарен, и $\mathbb{E}X_t = 0$ для всех t. Пусть

$$\Gamma(j) = \mathbb{E}X_t X'_{t-j}.$$

Заметим, что для того, чтобы $\{X_t\}$ был слабостационарным, матрица $\Gamma(j)$ не обязана быть симметричной для $j \neq 0$, однако

$$\Gamma(j) = \Gamma(-j)'.$$

Рассмотрим последовательность k-мерных матриц $\{C_j\}$ и определим

$$Y_t = \sum_{j=0}^{\infty} C_j X_{t-j}.$$

Дисперсия величины Y_t равна

$$\begin{aligned}\mathbb{E}Y_t Y_t' &= \mathbb{E}\left(\sum_{i=0}^{\infty} C_i X_{t-i}\right)\left(\sum_{j=0}^{\infty} C_j X_{t-j}\right)' = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} C_i \Gamma(j-i) C_j' \\ &= \sum_{j=0}^{\infty} C_j \Gamma(0) C_j' + \sum_{h=1}^{\infty}\sum_{j=0}^{\infty}\left(C_j \Gamma(h) C'_{j+h} + C_{j+h}\Gamma(h)' C_j'\right).\end{aligned}$$

Пусть $\|A\| = \sqrt{\operatorname{tr}(\mathrm{A}'\mathrm{A})}$. Имеем:

$$\begin{aligned}\left\|\mathbb{E}Y_t Y_t'\right\| &\leq \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\left\|C_i \Gamma(j-i) C_j'\right\| \leq \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\|C_i\|\,\|C_j\|\,\|\Gamma(j-i)\| \\ &\leq 2\sum_{h=0}^{\infty}\|\Gamma(h)\|\left(\sum_{j=0}^{\infty}\|C_j\|^2\right)^2,\end{aligned}$$

где последнее неравенство доказывается так же, как (4). Значит, матрица $\mathbb{E}Y_t Y_t'$ конечна при условиях

$$\sum_{j=0}^{\infty}\|C_j\|^2 < \infty \tag{17}$$

и

$$\sum_{h=-\infty}^{\infty}\|\Gamma(h)\| < \infty. \tag{18}$$

Процесс $\{\varepsilon_t\}$ называется k-мерным БШ, если $\mathbb{E}\varepsilon_t = 0$ и $\mathbb{E}\varepsilon_t\varepsilon_t' = \Sigma$ — положительно определенная матрица для всех t, и $\mathbb{E}\varepsilon_t\varepsilon_s' = 0$ для $t \neq s$. Если $\{X_t\}$ — слабостационарный чисто недетерминистический процесс с нулевым средним, такой что выполнено условие (18), у него имеется, аналогично скалярному случаю, $MA(\infty)$-представление:

$$X_t = \sum_{j=0}^{\infty} C_j \varepsilon_{t-j},$$

где $\{\varepsilon_t\}$ — векторный БШ, матрицы C_j удовлетворяют условию (17), и

$$C_0 = I_k.$$

Снова, как и в скалярном случае, величины ε_t являются ошибками линейного предсказания на один шаг вперед.

Спектральная плотность векторного процесса определяется аналогично скалярному случаю:

$$f(\lambda) = \frac{1}{2\pi}\sum_{j=-\infty}^{\infty}\Gamma(j)\, e^{-i\lambda j} = \frac{1}{2\pi}\left(\Gamma(0) + \sum_{j=1}^{\infty}\left(\Gamma(j)\, e^{-i\lambda j} + \Gamma(j)'\, e^{i\lambda j}\right)\right),$$

Долгосрочная ковариационная матрица снова равна

$$\Omega = 2\pi f(0).$$

Чтобы долгосрочная дисперсия была конечна, достаточно

$$\sum_{j=0}^{\infty}\|C_j\| < \infty.$$

Для векторного БШ $\{\varepsilon_t\}$ спектральная плотность постоянна:

$$f_\varepsilon(\lambda) = \frac{1}{2\pi}\Sigma.$$

Пусть $\{X_t\}$ — слабостационарный процесс со спектральной плотностью f_X. Определим

$$Y_t = \sum_{j=0}^{\infty} C_j X_{t-j}.$$

Тогда спектральная плотность процесса $\{Y_t\}$ равна

$$f_Y(\lambda) = \left(\sum_{j=0}^{\infty} C_j e^{-i\lambda j}\right) f_X(\lambda) \left(\sum_{j=0}^{\infty} C_j' e^{i\lambda j}\right).$$

Пусть $\{C_j\}$ — последовательность матриц порядка k. Многочлен $C(L)$ от оператора сдвига в векторном случае определяется как

$$C(L) = I_k + C_1 L + C_2 L^2 + \ldots,$$

где мы положили $C_0 = I_k$ в соответствии с разложением Вольда. Говорят, что

$$B(L) = C(L)^{-1}$$

если

$$B(L)\, C(L) = I_k.$$

Многочлен $C(L)$ обратим, если корни уравнения

$$\det(C(z)) = 0$$

лежат вне единичного круга. Например, если многочлен $C(L)$ имеет порядок p, то все корни уравнения

$$\det(I + C_1 z + \ldots + C_p z^p) = 0$$

должны быть по модулю больше единицы.

4 Модель ARMA

В этом разделе для скалярного случая иллюстрируются некоторые понятия предыдущего раздела. Пусть $\{\varepsilon_t\}$ — БШ с $\mathbb{V}(\varepsilon_t) = \sigma^2$. Процесс $MA(q)$, скажем $\{X_t\}$, имеет вид:

$$X_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \ldots + \theta_q \varepsilon_{t-q} = \Theta(L)\, \varepsilon_t,$$

где

$$\Theta(L) = 1 + \theta_1 L + \ldots + \theta_q L^q.$$

По теореме 2, процесс $MA(q)$ имеет спектральную плотность

$$f_X(\lambda) = \frac{\sigma^2}{2\pi} \left|\Theta\left(e^{-i\lambda}\right)\right|^2 = \frac{\sigma^2}{2\pi} \left|1 + \theta_1 e^{-i\lambda} + \ldots + \theta_q e^{-i\lambda q}\right|^2,$$

и долгосрочную дисперсию

$$\omega_X = 2\pi f_X(0) = \sigma^2 \left|\Theta(1)\right|^2 = \sigma^2 \left(1 + \theta_1 + \ldots + \theta_q\right)^2.$$

Заметим, что для некоторых значений θ_j, долгосрочная дисперсия может быть равна нулю. Для $0 \leq j \leq q$ эффект шока в период t на X через j периодов равен

$$\theta_j = \frac{\partial X_{t+j}}{\partial \varepsilon_t},$$

и влияние шока нулевой после q периодов.

Определение 7 (AR) *Процесс $\{X_t\}$ называется авторегрессией порядка p, обозначаемой как $AR(p)$, если он имеет вид*

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \ldots + \phi_p X_{t-p} + \varepsilon_t,$$

где $\{\varepsilon_t\}$ есть БШ.

Пусть

$$\Phi(L) = 1 - \phi_1 L - \ldots - \phi_p L^p.$$

Тогда процесс $AR(p)$ может быть переписан в виде

$$\Phi(L)\, X_t = \varepsilon_t.$$

Если корни многочлена $\Phi(z)$ лежат вне единичного круга, $\Phi(L)$ можно обратить, и процесс получает следующее $MA(\infty)$-представление (разложение Вольда):

$$X_t = \Phi(L)^{-1} \varepsilon_t = \Psi(L)\, \varepsilon_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} \tag{19}$$

для некоторого многочлена $\Psi(L)$.

Пусть $p = 1$. Тогда

$$(1 - \phi_1 L)\, X_t = \varepsilon_t,$$

и, в соответствии с уравнением (13),

$$X_t = (1 - \phi_1 L)^{-1} \varepsilon_t = \left(\sum_{j=0}^{\infty} \phi_1^j L^j\right) \varepsilon_t = \sum_{j=0}^{\infty} \phi_1^j \varepsilon_{t-j}.$$

Таким образом, в случае $AR(1)$-процесса коэффициенты многочлена $\Psi(L)$ в соотношении (19) равны $\psi_j = \phi_1^j$. Проверим условие суммируемости квадратов из теоремы 2:

$$\sum_{j=0}^{\infty} \psi_j^2 = \sum_{j=0}^{\infty} \phi_1^{2j} = \frac{1}{1-\phi_1^2} < \infty$$

при условии $|\phi_1| < 1$. Из теоремы 2 следует, что процесс $AR(1)$ слабостационарен, и его спектральная плотность и долгосрочная дисперсия равны

$$\begin{aligned} f_X(\lambda) &= \frac{\sigma^2}{2\pi}\left|\sum_{j=0}^{\infty}\phi_1^j e^{-i\lambda j}\right|^2 = \frac{\sigma^2}{2\pi}\frac{1}{|1-\phi_1 e^{-i\lambda}|^2}, \\ \omega_X &= 2\pi f_X(0) = \frac{\sigma^2}{(1-\phi_1)^2}. \end{aligned}$$

Долгосрочная дисперсия стационарного $AR(1)$-процесса также конечна. Коэффициенты в $MA(\infty)$-представлении стационарного $AR(1)$-процесса абсолютно суммируемы:

$$\sum_{j=0}^{\infty} |\psi_j| = \sum_{j=0}^{\infty} |\phi_1|^j = \frac{1}{1-|\phi_1|} < \infty.$$

Это следует из того факта, что $\psi_j = \phi_1^j \to 0$ при $j \to \infty$ с экспоненциальной скоростью.

В случае $AR(p)$, $p > 1$, в силу (16), $\sum_{j=0}^{\infty} \psi_j^2 < \infty$, если все корни многочлена $\Phi(z)$ лежат вне единичного круга. Тогда $AR(p)$-процесс слабостационарен со спектральной плотностью и долгосрочной дисперсией

$$\begin{aligned} f_X(\lambda) &= \frac{\sigma^2}{2\pi}\frac{1}{|\Phi(e^{-i\lambda})|^2}, \\ \omega_X &= \frac{\sigma^2}{\Phi(1)^2}. \end{aligned}$$

Определение 8 (ARMA) *Процесс $\{X_t\}$ называется $ARMA(p,q)$-процессом, если он имеет вид*

$$\Phi(L) X_t = \Theta(L)\varepsilon_t,$$

где $\{\varepsilon_t\}$ — БШ.

Если корни многочлена $\Phi(z)$ лежат вне единичного круга, то $ARMA(p,q)$-процесс имеет $MA(\infty)$-представление

$$X_t = \Phi(L)^{-1}\Theta(L)\varepsilon_t.$$

Его спектральная плотность и долгосрочная дисперсия равны

$$\begin{aligned} f_X(\lambda) &= \frac{\sigma^2}{2\pi}\frac{\left|\Theta(e^{-i\lambda})\right|^2}{|\Phi(e^{-i\lambda})|^2}, \\ \omega_X &= \sigma^2\frac{|\Theta(1)|^2}{|\Phi(1)|^2}. \end{aligned}$$

Если корни многочлена $\Theta(z)$ лежат вне единичного круга, то процесс $ARMA(p,q)$ имеет $AR(\infty)$-представление

$$\Theta(L)^{-1}\Phi(L) X_t = \varepsilon_t.$$

Таким образом, на практике любой слабостационарный процесс $ARMA\left(p,q\right)$ или $MA\left(\infty\right)$ может быть аппроксимирован моделью $AR\left(m_n\right)$, с порядком m_n, увеличивающимся с n, с меньшей, впрочем, скоростью.

ARMA-процесс с ненулевым средним μ может быть записан в виде

$$\Phi\left(L\right)\left(X_t-\mu\right)=\Theta\left(L\right)\varepsilon_t,$$

или, эквивалентно, как

$$\Phi\left(L\right)X_t=\alpha+\Theta\left(L\right)\varepsilon_t.$$

В последнем уравнении α может рассматриваться как свободный член в модели линейной регрессии, так как уравнение может быть представлено в виде

$$X_t=\alpha+\phi_1X_{t-1}+\ldots+\phi_pX_{t-p}+\Theta\left(L\right)\varepsilon_t.$$

Заметим, что μ и α связаны соотношением

$$\mu=\frac{\alpha}{\Phi\left(1\right)}=\frac{\alpha}{1-\phi_1-\ldots-\phi_p},$$

если многочлен $\Phi(L)$ обратим.

5 Асимптотические результаты

В настоящем разделе мы обсудим ЗБЧ и ЦПТ для линейных процессов $\{X_t\}$ вида

$$X_t=\sum_{j=0}^{\infty}c_j\varepsilon_{t-j}=C\left(L\right)\varepsilon_t, \tag{20}$$

где

$$C\left(L\right)=\sum_{j=0}^{\infty}c_jL^j.$$

Соответствующий метод был разработан в работе Phillips& Solo (1992). Используя подход авторов и предполагая, что $\{\varepsilon_t\}$ — последовательность независимых одинаково распределенных (н.о.р.) случайных величин с нулевым средним и конечной дисперсией, можно вывести ЗБЧ и ЦПТ для серийно коррелированного линейного процесса $\{X_t\}$, опираясь только на ЗБЧ и ЦПТ для н.о.р.-последовательностей и некоторое алгебраическое разложение многочлена $C\left(L\right)$.

Этот метод работает также для более широкого класса последовательностей инноваций $\{\varepsilon_t\}$. Например, вместо предположения о н.о.р. последовательности $\{\varepsilon_t\}$ можно работать со случаем независимости, но неодинаковой распределенности (н.н.о.р.). В этом случае необходимы следующие сильные версии ЗБЧ (СЗБЧ) и ЦПТ.

Лемма 3 (СЗБЧ для н.н.о.р.-последовательностей) *Пусть $\{\varepsilon_t\}$ — последовательность независимых случайных величин, таких что $\sup_t\mathbb{E}\left|\varepsilon_t\right|^{1+\delta}<\infty$ для некоторого $\delta>0$. Тогда $n^{-1}\sum_{t=1}^n\varepsilon_t-n^{-1}\sum_{t=1}^n\mathbb{E}\varepsilon_t\to_{a.s.}0$.*

См. White (2001, следствие 3.9).

Лемма 4 (ЦПТ для н.н.о.р.-последовательностей) *Пусть $\{\varepsilon_t\}$ — последовательность независимых случайных величин, таких что $\mathbb{E}\varepsilon_t=0$ для всех t, $\sup_t\mathbb{E}\left|\varepsilon_t\right|^{2+\delta}<\infty$ для некоторого $\delta>0$, и для всех достаточно больших n выполнено $n^{-1}\sum_{t=1}^n\mathbb{E}\varepsilon_t^2>\delta'>0$. Тогда $n^{-1/2}\sum_{t=1}^n\varepsilon_t/\left(n^{-1}\sum_{t=1}^n\mathbb{E}\varepsilon_t^2\right)^{1/2}\to_d N\left(0,1\right)$.*

См. White (2001, теорема 5.10).

С другой стороны, если нельзя предположить, что инновации $\{\varepsilon_t\}$ независимы, можно опираться на ЗБЧ и ЦПТ для мартингал-разностей.

Определение 9 (Мартингал-разность) *Пусть $\{\mathcal{F}_t\}$ — возрастающая последовательность σ-алгебр. Тогда $\{(u_t, \mathcal{F}_t)\}$ называется мартингалом, если $\mathbb{E}(u_t|\mathcal{F}_{t-1}) = u_{t-1}$ почти наверное. Последовательность $\{(\varepsilon_t, \mathcal{F}_t)\}$ называется мартингал-разностью, если $\mathbb{E}(\varepsilon_t|\mathcal{F}_{t-1}) = 0$ почти наверное.*

Если процесс $\{u_t\}$ задан, σ-алгебра $\mathcal{F}_t$ часто определяется как $\sigma(u_t, u_{t-1}, \ldots)$. Пусть $\{(u_t, \mathcal{F}_t)\}$ — мартингал. Тогда $\{(u_t - u_{t-1}, \mathcal{F}_t)\}$ — мартингал-разность, так как $u_t - u_{t-1} = u_t - \mathbb{E}(u_t|\mathcal{F}_{t-1})$, и, следовательно, $\mathbb{E}(u_t - u_{t-1}|\mathcal{F}_{t-1}) = 0$. Легко видеть, что мартингал-разность $\{\varepsilon_t\}$ серийно некоррелирована: $\mathbb{E}\varepsilon_t\varepsilon_{t-1} = \mathbb{E}(\varepsilon_{t-1}\mathbb{E}(\varepsilon_t|\mathcal{F}_{t-1})) = 0$. В то же время величины ε могут не быть независимыми.

Если $\{\varepsilon_t\}$ — мартингал-разность, следующие СЗБЧ и ЦПТ могут быть полезны.

Лемма 5 (СЗБЧ для мартингал-разностей) *Пусть $\{(\varepsilon_t, \mathcal{F}_t)\}$ — мартингал-разность, такая что $\sup_t \mathbb{E}|\varepsilon_t|^{2+\delta} < \infty$ для некоторого $\delta > 0$. Тогда $n^{-1}\sum_{t=1}^n \varepsilon_t \to_{a.s.} 0$.*

См. White (2001, упражнение 3.77).

Лемма 6 (ЦПТ для мартингал-разностей) *Пусть $\{(\varepsilon_t, \mathcal{F}_t)\}$ — мартингал-разность, такая что $\sup_t \mathbb{E}|\varepsilon_t|^{2+\delta} < \infty$ для некоторого $\delta > 0$. Пусть для всех достаточно больших n выполнено $n^{-1}\sum_{t=1}^n \mathbb{E}\varepsilon_t^2 > \delta' > 0$, и пусть $n^{-1}\sum_{t=1}^n \varepsilon_t^2 - n^{-1}\sum_{t=1}^n \mathbb{E}\varepsilon_t^2 \to_p 0$. Тогда $n^{-1/2}\sum_{t=1}^n \varepsilon_t / \left(n^{-1}\sum_{t=1}^n \mathbb{E}\varepsilon_t^2\right)^{1/2} \to_d N(0,1)$.*

См. White (2001, следствие 5.26).

Лемма 7 (ЦПТ для строго стационарных и эргодичных мартингал-разностей) *Пусть $\{(\varepsilon_t, \mathcal{F}_t)\}$ — строго стационарная и эргодичная мартингал-разность, такая что $\mathbb{E}\varepsilon_t^2 < \infty$. Тогда $n^{-1/2}\sum_{t=1}^n \varepsilon_t \to_d N\left(0, \mathbb{E}\varepsilon_t^2\right)$.*

Заметим, что все рассмотренные альтернативные требования к процессу $\{\varepsilon_t\}$, н.о.р., н.н.о.р. или мартингал-разность, сильнее простого БШ. Таким образом, для использования указанного метода необходимо немного усилить MA(∞)-представление по сравнению с результатом, даваемым разложением Вольда.

5.1 БН-разложение

Обсудим сначала алгебраическое разложение многочлена от оператора сдвига на долгосрочную и переходную компоненты. Такое разложение было предложено в работе Beveridge & Nelson (1981).

Лемма 8 *Пусть $C(L) = \sum_{j=0}^{\infty} c_j L^j$. Тогда*

(a) $C(L) = C(1) - (1-L)\widetilde{C}(L)$, где $\widetilde{C}(L) = \sum_{j=0}^{\infty} \widetilde{c}_j L^j$ с $\widetilde{c}_j = \sum_{h=j+1}^{\infty} c_h$.

(b) Если $\sum_{j=1}^{\infty} j^{1/2}|c_j| < \infty$, то $\sum_{j=0}^{\infty} \widetilde{c}_j^2 < \infty$.

(c) Если $\sum_{j=1}^{\infty} j|c_j| < \infty$, то $\sum_{j=0}^{\infty} |\widetilde{c}_j| < \infty$.

Здесь $C(1)$ — долгосрочная компонента, а $(1-L)\widetilde{C}(L)$ считается переходной компонентой по причинам, которые скоро станут ясны, в частности, из уравнений (22) и (23).

Доказательство. Для доказательства пункта (a), запишем

$$\begin{aligned}\sum_{j=0}^{\infty} c_j L^j &= \sum_{j=0}^{\infty} c_j - \sum_{j=1}^{\infty} c_j + \left(\sum_{j=1}^{\infty} c_j - \sum_{j=2}^{\infty} c_j\right) L + \left(\sum_{j=2}^{\infty} c_j - \sum_{j=3}^{\infty} c_j\right) L^2 \\ &\quad + \ldots + \left(\sum_{j=h}^{\infty} c_j - \sum_{j=h+1}^{\infty} c_j\right) L^h + \ldots .\end{aligned}$$

Переставляя местами слагаемые, получим

$$\begin{aligned}\sum_{j=0}^{\infty} c_j L^j &= \sum_{j=0}^{\infty} c_j - (1-L)\sum_{j=1}^{\infty} c_j - (1-L)\sum_{j=2}^{\infty} c_j L - \ldots - (1-L)\sum_{j=h+1}^{\infty} c_j L^h - \ldots \\ &= \sum_{j=0}^{\infty} c_j - (1-L)\sum_{j=0}^{\infty}\left(\sum_{h=j+1}^{\infty} c_h\right) L^j = C(1) - (1-L)\widetilde{C}(L).\end{aligned}$$

Для доказательства пункта (b),

$$\begin{aligned}\sum_{j=0}^{\infty} \widetilde{c}_j^2 &= \sum_{j=0}^{\infty}\left(\sum_{h=j+1}^{\infty} c_h\right)^2 \le \sum_{j=0}^{\infty}\left(\sum_{h=j+1}^{\infty} |c_h|\right)^2 = \sum_{j=0}^{\infty}\left(\sum_{h=j+1}^{\infty} |c_h|^{1/2} h^{1/4} |c_h|^{1/2} h^{-1/4}\right)^2 \\ &\le \sum_{j=0}^{\infty}\left(\sum_{h=j+1}^{\infty} |c_h| h^{1/2}\right)\left(\sum_{h=j+1}^{\infty} |c_h| h^{-1/2}\right) \le \left(\sum_{h=0}^{\infty} |c_h| h^{1/2}\right)\sum_{j=0}^{\infty}\sum_{h=j+1}^{\infty} |c_h| h^{-1/2}.\end{aligned}$$

Далее, рассмотрим $\sum_{j=0}^{\infty}\sum_{h=j+1}^{\infty} |c_h| h^{-1/2}$. Заметим, что $|c_1|$ входит в сумму только однажды, когда $j=0$, $|c_2|$ появляется в сумме дважды, когда $j=0,1$, и.т.д. Таким образом, $|c_h|$ входит в сумму для $j=0,1,\ldots,h-1$, или h раз. Следовательно,

$$\sum_{j=0}^{\infty}\sum_{h=j+1}^{\infty} |c_h| h^{-1/2} = \sum_{j=0}^{\infty} |c_j| j^{-1/2} j = \sum_{j=0}^{\infty} |c_j| j^{1/2},$$

и

$$\sum_{j=0}^{\infty} \widetilde{c}_j^2 \le \left(\sum_{j=0}^{\infty} |c_j| j^{1/2}\right)^2 .$$

Для доказательства пункта (c),

$$\sum_{j=0}^{\infty} |\widetilde{c}_j| = \sum_{j=0}^{\infty}\left|\sum_{h=j+1}^{\infty} c_h\right| \le \sum_{j=0}^{\infty}\sum_{h=j+1}^{\infty} |c_h| = \sum_{j=0}^{\infty} |c_j| j.$$

■

Заметим, что предположения $\sum_{j=1}^{\infty} j^{1/2}|c_j| < \infty$ и $\sum_{j=1}^{\infty} j|c_j| < \infty$ сильнее, чем конечность долгосрочной дисперсии $\sum_{j=1}^{\infty} |c_j| < \infty$.

Согласно БН-разложению, если процесс $\{X_t\}$ линеен, то

$$X_t = C(L)\varepsilon_t = C(1)\varepsilon_t - (1-L)\widetilde{C}(L)\varepsilon_t = C(1)\varepsilon_t - (\widetilde{\varepsilon}_t - \widetilde{\varepsilon}_{t-1}), \tag{21}$$

где

$$\widetilde{\varepsilon}_t = \widetilde{C}(L)\varepsilon_t.$$

Более того, $\widetilde{\varepsilon}_t$ имеет конечную дисперсию, если $\sum_{j=0}^{\infty} j^{1/2} |c_j| < \infty$. Первое слагаемое в правой части уравнения (21), $C(1)\varepsilon_t$, называется *долгосрочной компонентой*, а $\widetilde{\varepsilon}_t - \widetilde{\varepsilon}_{t-1}$ — *переходной компонентой.*

Аналогичное разложение справедливо в векторном случае. Переходная компонента имеет конечную дисперсию, если $\sum_{j=1}^{\infty} j^{1/2} \|C_j\| < \infty$:

$$\begin{aligned}\sum_{j=0}^{\infty} \left\|\widetilde{C}_j\right\|^2 &= \sum_{j=0}^{\infty} \left\|\sum_{h=j+1}^{\infty} C_h\right\|^2 \le \sum_{j=0}^{\infty} \left(\sum_{h=j+1}^{\infty} \|C_j\|\right)^2 \\ &= \sum_{j=0}^{\infty} \left(\sum_{h=j+1}^{\infty} \|C_h\|^{1/2} h^{1/4} \|C_h\|^{1/2} h^{-1/4}\right)^2 \le \left(\sum_{j=0}^{\infty} j^{1/2} \|C_j\|\right)^2.\end{aligned}$$

Условие пункта (c) теоремы 8 преобразуется в $\sum_{j=1}^{\infty} j \|C_j\| < \infty$ в векторном случае.

5.2 ЗБЧ для линейных процессов

В настоящем разделе мы покажем, как получить ЗБЧ для линейного процесса с помощью БН-разложения, следуя подходу Филлипса и Соло.

Теорема 9 (ЗБЧ для линейных процессов с н.о.р. инновациями) *Пусть* $X_t = C(L)\varepsilon_t$, $\{\varepsilon_t\}$ *— последовательность н.о.р. случайных величин с* $\mathbb{E}|\varepsilon_t| < \infty$ *и* $\mathbb{E}\varepsilon_t = 0$, *и пусть для* $C(L)$ *выполнено* $\sum_{j=1}^{\infty} j|c_j| < \infty$. *Тогда* $n^{-1}\sum_{t=1}^{n} X_t \to_p 0$.

Доказательство. Ключевой элемент доказательства — БН-разложение (21). Сначала заметим, что $\sum_{t=1}^{n} (\widetilde{\varepsilon}_t - \widetilde{\varepsilon}_{t-1})$ является так называемой телескопической суммой:

$$\begin{aligned}\sum_{t=1}^{n} (\widetilde{\varepsilon}_t - \widetilde{\varepsilon}_{t-1}) &= (\tilde{\varepsilon}_1 - \tilde{\varepsilon}_0) + (\tilde{\varepsilon}_2 - \tilde{\varepsilon}_1) + (\tilde{\varepsilon}_3 - \tilde{\varepsilon}_2) + \ldots + (\tilde{\varepsilon}_n - \tilde{\varepsilon}_{n-1}) \\ &= -\widetilde{\varepsilon}_0 + \widetilde{\varepsilon}_n, \end{aligned} \tag{22}$$

так что

$$n^{-1}\sum_{t=1}^{n} X_t = C(1)\, n^{-1}\sum_{t=1}^{n} \varepsilon_t - n^{-1}(\widetilde{\varepsilon}_n - \widetilde{\varepsilon}_0). \tag{23}$$

В слагаемом, соответствующем долгосрочной компоненте, происходит накопление шоков, а в переходной компоненте шоки взаимоуничтожаются.

Так как $\sum_{j=1}^{\infty} j|c_j| < \infty$, мы получаем, что $|C(1)| < \infty$. Следовательно, используя, например, слабый ЗБЧ (СлЗБЧ) для н.о.р.-последовательностей, получаем, что

$$C(1)\, n^{-1}\sum_{t=1}^{n} \varepsilon_t \to_p 0.$$

Осталось показать, что вторым слагаемым (соответствующим переходной компоненте) в уравнении (23) можно асимптотически пренебречь:

$$\mathbb{P}\left\{n^{-1}|\widetilde{\varepsilon}_t| > \delta\right\} \le \frac{\mathbb{E}|\widetilde{\varepsilon}_t|}{n\delta}$$

и

$$\mathbb{E}|\widetilde{\varepsilon}_t| = \mathbb{E}\left|\sum_{j=0}^{\infty} \widetilde{c}_j \varepsilon_{t-j}\right| \le \sum_{j=0}^{\infty} |\widetilde{c}_j|\, \mathbb{E}|\varepsilon_{t-j}| = \mathbb{E}|\varepsilon_0| \sum_{j=0}^{\infty} |\widetilde{c}_j| < \infty \tag{24}$$

при условии $\sum_{j=1}^{\infty} j\,|c_j| < \infty$. Следовательно,

$$n^{-1}\left(\widetilde{\varepsilon}_n - \widetilde{\varepsilon}_0\right) \to_p 0.$$

■

Если же предположить, что $\mathbb{E}\varepsilon_t^2 < \infty$, то можно заменить $\sum_{j=1}^{\infty} j\,|c_j| < \infty$ на $\sum_{j=1}^{\infty} j^{1/2}\,|c_j| < \infty$, так как

$$P\left(n^{-1}\,|\widetilde{\varepsilon}_t| > \delta\right) \le \frac{\mathbb{E}\widetilde{\varepsilon}_t^2}{n^2\delta^2},$$

и $\mathbb{E}\widetilde{\varepsilon}_t^2 < \infty$ при условии $\sum_{j=1}^{\infty} j^{1/2}\,|c_j| < \infty$.

Можно доказать аналогичный ЗБЧ в предположении, что $\{\varepsilon_t\}$ — последовательность н.н.о.р. величин или мартингал-разностей, с помощью соответствующих ЗБЧ (леммы 3 и 5 соответственно). Например, аналогичный результат имеет место, если $\{\varepsilon_t\}$ н.н.о.р., $\sup_t \mathbb{E}\,|\varepsilon_t|^{1+\delta} < \infty$ для некоторого $\delta > 0$ и $\sum_{j=1}^{\infty} j\,|c_j| < \infty$. Если $\{(\varepsilon_t, \mathcal{F}_t)\}$ — мартингал-разность, то аналогичный результат верен при условиях $\sup_t \mathbb{E}\,|\varepsilon_t|^{2+\delta} < \infty$ для некоторого $\delta > 0$ и $\sum_{j=1}^{\infty} j^{1/2}\,|c_j| < \infty$.

До настоящего момента мы предполагали, что $\mathbb{E}X_t = 0$. Можно модифицировать это предположение так, что $X_t = \mu + C(L)\,\varepsilon_t$, где μ — среднее процесса X_t. В этом случае, при тех же условиях имеем $n^{-1}\sum_{t=1}^{n} X_t \to_p \mu$. Например, $AR(1)$-процесс со средним μ определяется как $(1-\phi L)(X_t - \mu) = \varepsilon_t$. Если $|\phi| < 1$, то выборочное среднее процесса X_t сходится по вероятности к μ, если выполнены соответствующим условия на $\{\varepsilon_t\}$.

5.3 ЦПТ для линейных процессов

Далее показано, как базовая ЦПТ для н.о.р. случайных величин может быть обобщена на случай линейных процессов с помощью подхода Филлипса–Соло.

Теорема 10 (ЦПТ для линейных процессов с н.о.р. инновациями) *Пусть $X_t = C(L)\varepsilon_t$, $\{\varepsilon_t\}$ — последовательность н.о.р. случайных величин с $\mathbb{E}\,|\varepsilon_t|^2 = \sigma^2 < \infty$ и $\mathbb{E}\varepsilon_t = 0$, $C(L)$ удовлетворяет условию $\sum_{j=1}^{\infty} j^{1/2}\,|c_j| < \infty$ и $C(1) \neq 0$. Тогда $n^{-1/2}\sum_{t=1}^{n} X_t \to_d N(0, \sigma^2 C(1)^2)$.*

Доказательство. БН-разложение позволяет записать

$$\begin{aligned} n^{-1/2}\sum_{t=1}^{n} X_t &= C(1)\,n^{-1/2}\sum_{t=1}^{n}\varepsilon_t - n^{-1/2}\left(\widetilde{\varepsilon}_n - \widetilde{\varepsilon}_0\right) = C(1)\,n^{-1/2}\sum_{t=1}^{n}\varepsilon_t + o_p(1) \\ &\to_d C(1)\,N\left(0, \sigma^2\right) =^d N(0, \sigma^2 C(1)^2). \end{aligned}$$

Здесь имеет место сходимость по распределению в силу ЦПТ для н.о.р. случайных величин.

■

Использованный подход ясно демонстрирует, почему в случае наличия серийной корреляции асимптотическая дисперсия среднего зависит от долгосрочной дисперсии процесса. Как и ранее, этот подход может быть обобщен на случай, когда $\{\varepsilon_t\}$ — последовательность н.н.о.р. величин или мартингал-разностей.

В векторном случае, пусть $\{\varepsilon_t\}$ — последовательность н.о.р. k-мерных векторов с $\mathbb{E}\varepsilon_t = 0$, и $\mathbb{E}\varepsilon_t\varepsilon_t' = \Sigma$ — конечная матрица. Пусть $X_t = C(L)\,\varepsilon_t$ и $\sum_{j=0}^{\infty} j^{1/2}\,\|C_j\| < \infty$, $C(1) \neq 0$. Так как $n^{-1/2}\sum_{t=1}^{n}\varepsilon_t \to_d N(0, \Sigma)$, получаем, что

$$n^{-1/2}\sum_{t=1}^{n} X_t \to_d N\left(0, C(1)\,\Sigma C(1)'\right).$$

5.4 Сходимость выборочных дисперсий

Рассмотрим простую модель линейной регрессии $Y_t = \beta X_t + U_t$. МНК-оценка β равна $\hat{\beta}_n = \sum_{t=1}^{n} X_t Y_t / \sum_{t=1}^{n} X_t^2 = \beta + \sum_{t=1}^{n} X_t U_t / \sum_{t=1}^{n} X_t^2$. Таким образом, в работе с асимптотическими свойствами МНК-оценки необходимо иметь дело со вторыми выборочными моментами $n^{-1} \sum_{t=1}^{n} X_t^2$. В настоящем разделе мы обсудим, как можно установить сходимость вторых выборочных моментов в случае, когда процесс $\{X_t\}$ линеен.

Предположим, что $\{X_t\}$ — скалярный линейный процесс, удовлетворяющий условиям теоремы 10. Запишем

$$\begin{aligned} X_t^2 &= (C(L)\varepsilon_t)^2 = \left(\sum_{j=0}^{\infty} c_j \varepsilon_{t-j}\right)^2 = \sum_{j=0}^{\infty}\sum_{l=0}^{\infty} c_j c_l \varepsilon_{t-j}\varepsilon_{t-l} \\ &= \sum_{j=0}^{\infty} c_j^2 \varepsilon_{t-j}^2 + 2\sum_{j=0}^{\infty}\sum_{l>j} c_j c_l \varepsilon_{t-j}\varepsilon_{t-l} \\ &= \sum_{j=0}^{\infty} c_j^2 \varepsilon_{t-j}^2 + 2\sum_{h=1}^{\infty}\sum_{j=0}^{\infty} c_j c_{j+h} \varepsilon_{t-j}\varepsilon_{t-j-h} \text{ (замена } l = j+h \text{, так что } h = 1,2,\ldots) \\ &= B_0(L)\varepsilon_t^2 + 2\sum_{h=1}^{\infty} B_h(L)\,\varepsilon_t\varepsilon_{t-h}, \end{aligned}$$

где для $h = 0, 1, \ldots$ выполнено

$$B_h(L) = \sum_{j=0}^{\infty} b_{h,j} L^j = \sum_{j=0}^{\infty} c_j c_{j+h} L^j.$$

Таким образом,

$$\begin{aligned} B_0(L) &= \sum_{j=0}^{\infty} b_{0,j} L^j = \sum_{j=0}^{\infty} c_j^2 L^j, \\ B_1(L) &= \sum_{j=0}^{\infty} b_{1,j} L^j = \sum_{j=0}^{\infty} c_j c_{j+1} L^j, \end{aligned}$$

и т.д. БН-разложение многочлена $B_h(L)$ есть

$$B_h(L) = B_h(1) - (1-L)\widetilde{B}_h(L), \tag{25}$$

где

$$\widetilde{B}_h(L) = \sum_{j=0}^{\infty} \widetilde{b}_{h,j} L^j, \quad \widetilde{b}_{h,j} = \sum_{l=j+1}^{\infty} b_{h,l} = \sum_{l=j+1}^{\infty} c_l c_{l+h}.$$

БН-разложение многочлена $B_h(L)$ верно, если $\sum_{j=1}^{\infty} j^{1/2}|c_j| < \infty$:

$$\begin{aligned} \sum_{j=0}^{\infty} \widetilde{b}_{h,j}^2 &= \sum_{j=0}^{\infty}\left(\sum_{l=j+1}^{\infty} c_l c_{l+h}\right)^2 = \sum_{j=0}^{\infty}\left(\sum_{l=j+1}^{\infty} l^{1/4} c_l c_{l+h} l^{-1/4}\right)^2 \\ &\le \sum_{j=0}^{\infty}\left(\sum_{l=j+1}^{\infty} l^{1/2} c_l^2\right)\left(\sum_{l=j+1}^{\infty} c_{l+h}^2 l^{-1/2}\right) \le \left(\sum_{l=1}^{\infty} l^{1/2} c_l^2\right)\left(\sum_{j=0}^{\infty}\sum_{l=j+1}^{\infty} c_{l+h}^2 l^{-1/2}\right) \\ &= \left(\sum_{l=1}^{\infty} l^{1/2} c_l^2\right)\left(\sum_{l=0}^{\infty} c_{l+h}^2 l^{1/2}\right) \le \left(\sum_{l=1}^{\infty} l^{1/2} c_l^2\right)^2, \end{aligned}$$

и ряд $\sum_{l=1}^{\infty} l^{1/2} c_l^2$ сходится, если сходится $\sum_{l=1}^{\infty} l^{1/2} |c_l|$:

$$\sum_{l=1}^{\infty} l^{1/2} c_l^2 = \sum_{l=1}^{\infty} l^{1/2} |c_l|^2 \leq \sup_j |c_j| \sum_{l=1}^{\infty} l^{1/2} |c_l| < \infty,$$

где $\sup_j |c_j| < \infty$, так как $\sum_{l=1}^{\infty} l^{1/2} |c_l| < \infty$, и значит $|c_l| \to 0$ при $l \to \infty$.

Таким образом, имеем

$$\begin{aligned} X_t^2 &= B_0(L)\varepsilon_t^2 + 2\sum_{h=1}^{\infty} B_h(L)\,\varepsilon_t\varepsilon_{t-h} \\ &= B_0(1)\varepsilon_t^2 + 2\sum_{h=1}^{\infty} B_h(1)\,\varepsilon_t\varepsilon_{t-h} - (1-L)\left(\widetilde{B}_0(L)\,\varepsilon_t^2 + 2\sum_{h=1}^{\infty} \widetilde{B}_h(L)\,\varepsilon_t\varepsilon_{t-h}\right) \\ &= \left(\sum_{j=0}^{\infty} c_j^2\right)\varepsilon_t^2 + u_t - (1-L)\widetilde{v}_t, \end{aligned}$$

где

$$\begin{aligned} u_t &= \varepsilon_t\left(2\sum_{h=1}^{\infty} B_h(1)\,\varepsilon_{t-h}\right), \\ \widetilde{v}_t &= \widetilde{B}_0(L)\,\varepsilon_t^2 + 2\sum_{h=1}^{\infty} \widetilde{B}_h(L)\,\varepsilon_t\varepsilon_{t-h}. \end{aligned}$$

Получаем

$$n^{-1}\sum_{t=1}^{n} X_t^2 = \left(\sum_{j=0}^{\infty} c_j^2\right) n^{-1}\sum_{t=1}^{n}\varepsilon_t^2 + n^{-1}\sum_{t=1}^{n} u_t - n^{-1}\left(\widetilde{v}_n - \widetilde{v}_0\right).$$

Теперь покажем, что

$$n^{-1}\sum_{t=1}^{n} u_t \to_{a.s.} 0.$$

Пусть $\mathcal{F}_t = \sigma(\varepsilon_t, \varepsilon_{t-1}, \ldots)$. Тогда $\{(u_t, \mathcal{F}_t)\}$ — мартингал-разность.

Лемма 11 (СЗБЧ для мартингал-разностей) *Пусть $\{(u_t, \mathcal{F}_t)\}$ — мартингал-разность. Если для некоторого $r \geq 1$ верно $\sum_{t=1}^{\infty} \mathbb{E}|u_t|^{2r}/t^{1+r} < \infty$, то $n^{-1}\sum_{t=1}^{n} u_t \to_{a.s.} 0$.*

См. White (2001, теорема 3.76). Проверим, что $\{u_t\}$ удовлетворяет условию приведенной леммы. Положим $r = 1$. Условие выполняется, если $\sup_t \mathbb{E}u_t^2 < \infty$, так как $\sum_{t=1}^{\infty} t^{-2} < \infty$.

$$\mathbb{E}u_t^2 = 4\sigma^2\mathbb{E}\left(\sum_{h=1}^{\infty} B_h(1)\,\varepsilon_{t-h}\right)^2 = 4\sigma^4\sum_{h=1}^{\infty} B_h(1)^2.$$

Далее,

$$\begin{aligned} \sum_{h=1}^{\infty} B_h(1)^2 &= \sum_{h=1}^{\infty}\left(\sum_{j=0}^{\infty} c_j c_{j+h}\right)^2 \leq \sum_{h=1}^{\infty}\left(\sum_{j=0}^{\infty} c_j^2\right)\sum_{j=0}^{\infty} c_{j+h}^2 \\ &= \left(\sum_{j=0}^{\infty} c_j^2\right)\sum_{h=1}^{\infty}\sum_{j=0}^{\infty} c_{j+h}^2 = \left(\sum_{j=0}^{\infty} c_j^2\right)\sum_{j=1}^{\infty} j c_j^2. \end{aligned}$$

Применяя логику, аналогичную рассуждениям после доказательства теоремы 2 раздела 3.1, получим, что из $\sum_{j=1}^{\infty} j^{1/2} |c_j| < \infty$ следует $\sum_{j=1}^{\infty} j c_j^2 < \infty$. Как и ранее, можно показать, что

$$n^{-1} (\widetilde{v}_n - \widetilde{v}_0) \to_p 0,$$

если только $\sum_{j=1}^{\infty} j^{1/2} |c_j| < \infty$. Наконец, в силу СлЗБЧ для последовательностей н.о.р. величин

$$n^{-1} \sum_{t=1}^{n} \varepsilon_t^2 \to_p \sigma^2.$$

Следовательно,

$$n^{-1} \sum_{t=1}^{n} X_t^2 \to_p \sigma^2 \sum_{j=0}^{\infty} c_j^2 = \mathbb{E} X_t^2.$$

Список литературы

Andrews, D.W.K. (1984). Non-strong mixing autoregressive processes. *Journal of Applied Probability* 21, 930-–934.

Beveridge, S. & C.R. Nelson (1981). A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the ‘business cycle’. *Journal of Monetary Economics* 7, 151—174.

Davidson, J. (1994). *Stochastic Limit Theory*. Oxford University Press: New York.

Gourieroux, C. & A. Monfort (1995). *Statistics and Econometric Models*. Cambridge University Press: Cambridge.

Hamilton, J.D. (1994). *Time Series Analysis*. Princeton University Press: Princeton.

Hansen, L.P. & T.J. Sargent (1991). *Rational Expectations Econometrics*. Westview Press: Boulder.

Phillips, P.C.B. & V. Solo (1992). Asymptotics for linear processes. *Annals of Statistics* 20, 971—1001.

White, H. (2001). *Asymptotic Theory For Econometricians*. Academic Press: San Diego.

Linear processes: properties and asymptotic results

Vadim Marmer

University of British Columbia, Vancouver, Canada

This essay surveys results for linear time series including Wold decomposition, properties of spectral density functions and lag operators, autoregressive moving average models, Beveridge–Nelson decomposition, and Phillips–Solo device for deriving asymptotics.