1. Вспомогательные определения

В данном разделе введем некоторые обозначения, которые будем использовать в дальнейшем.

Определение 1.1. Случайный процесс $\{Y_t : t \in \mathbb{Z}\}$ называют стационарным (в широком смысле), если

1. $\mathsf{E}Y_t \equiv \mathrm{const}$ (среднее постоянно по времени);
2. $\mathsf{cov}(Y_t, Y_{t+h}) = \gamma(h)$ (ковариация зависит только от лага h).

Замечание 1.1. Поскольку $\gamma(0) = \mathsf{cov}(Y_t, Y_t) = \mathsf{D}Y_t$, то дисперсия также не меняется со временем.

Замечание 1.2. Далее под стационарностью будет подразумеваться именно стационарность в широком смысле.

Определение 1.2. Случайный процесс $\{\varepsilon_t\}$ называют белым шумом $\mathrm{WN}(0, \sigma^2)$, если он стационарный, $\mathsf{E}\varepsilon_t = 0$, $\gamma(h) = 0\ \forall h \neq 0$ и $\mathsf{D}\varepsilon_t = \sigma^2$.

Определение 1.3. Спектральной плотностью стационарного процесса называется такая функция $f(\omega)$, что

$$\gamma(h) = 2\int_0^{1/2} e^{2\pi h\omega \mathrm{i}} f(\omega) d\omega.$$

Определение 1.4. Пусть $\{Y_t\}$ — стационарный процесс. Функцию

$$I(\omega) = \frac{1}{n}\left|\sum_{j=1}^{n} Y_j e^{-2\pi\omega j \mathrm{i}}\right|^2$$

называют периодограммой выборки размера n процесса $\{Y_t\}$.

2. Процессы с длинной памятью

Определение 2.1. Говорят, что стационарный процесс $\{Y_t\}$ обладает длинной памятью, если

$$\sum_{h=0}^{H} |\gamma(h)| \to \infty,$$

при $H \to \infty$. Иначе говорят, что $\{Y_t\}$ обладает короткой памятью:

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

Существуют и альтернативные определения процессов с длинной памятью, которые можно найти в [1, Section 3.1]. Там же показано, что они согласованы с определением 2.1.

Пример 2.1. Процессом с короткой памятью является, например, модель ARMA(p, q), поскольку $|\gamma(h)| \leqslant CR^h$, где $C > 0$ и $0 < R < 1$ [2].

Примером стационарного процесса с длинной памятью является дробно интегрированный процесс. Для его определения необходимо ввести понятие дробного интегрирования $(1-L)^d$, где L — оператор сдвига. Например, для $d = 1$ имеем $(1-L)Y_t = Y_t - Y_{t-1}$, для $d = 2$ — $(1-L)^2 Y_t = Y_t - 2Y_{t-1} + Y_{t-2}$, и так далее. Обобщим этот оператор для нецелых d с помощью разложения в ряд Тейлора функции $(1-x)^d$ в нуле:

$$\begin{aligned}(1-x)^d &= 1 - dx - \frac{d(1-d)}{2}x^2 - \frac{d(1-d)(2-d)}{3!}x^3 - \dots \\ &= \sum_{j=0}^{\infty} \pi_j(d) x^j = \sum_{j=0}^{\infty} \binom{d}{j} (-1)^j x^j,\end{aligned}$$

где $\binom{d}{j}$ — биномиальный коэффициент. Коэффиенты $\pi_j(d)$ удовлетворяют соотношению

$$\pi_j(d) = (-1)^j \binom{d}{j} = \frac{j-1-d}{j}\pi_{j-1}(d) = \frac{\Gamma(j-d)}{\Gamma(j+1)\Gamma(-d)}, \tag{1}$$

где $\Gamma(x)$ — гамма функция. Заметим, что второе равенство в формуле (1) верно для любых d, третье же верно только для $d \notin \mathbb{N} \cup \{0\}$, поскольку гамма функция не определена для неположительных целых чисел.

Предположение 2.1. Пусть $\{X_t\}$ представляет собой MA(∞) процесс,

$$X_t = \sum_{j=0}^{\infty} c_j \varepsilon_{t-j}, \quad \{\varepsilon_t\} \sim \mathrm{WN}(0, \sigma^2),$$

который абсолютно суммируемый, $\sum_{j=0}^{\infty} |c_j| < \infty$, и $\sum_{j=0}^{\infty} c_j \neq 0$.

Определение 2.2. Пусть процесс $\{Y_t\}$ определен соотношением

$$Y_t = (1-L)^{-d} X_t = \sum_{k=0}^{\infty} \pi_k(-d) X_{t-k}, \quad d < \frac{1}{2},$$

где $\pi_k(-d)$ из формулы (1), $\{X_t\}$ уловлетворяет предположению 2.1 и существует такое $s > 1 - d$, что

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

Процесс $\{Y_t\}$ называют дробно интегрированным процессом порядка d (FI(d)).

Предложение 2.1. Процесс $\{Y_t\}$ из определения 2.2 является стационарным с $\mathsf{E}Y_t = 0$ при $d < 1/2$. Его спектральная плотность определяется выражением

$$\begin{aligned} f_Y(\omega) &= 4^{-d} \sin^{-2d}(\pi\omega)\, f_X(\omega), \quad \omega > 0 \\ &\sim \omega^{-2d} f_X(0), \quad \omega \to 0, \end{aligned} \tag{2}$$

где $f_X(\omega)$ — спектральная плотность $\{X_t\}$.

Доказательство. См. [3, Proposition 6.1]. □

Замечание 2.1. Из формулы (2) видно, что спектральная плотность дробно интегрированного процесса монотонно убывает (возрастает) тогда и только тогда, когда монотонно убывает (возрастает) спектральная плотность процесса $\{X_t\}$.

Следствие 2.1. В условиях предложения 2.1 при $0 < d < 1/2$

$$\gamma_Y(h) \sim C_{\gamma,d} h^{2d-1}, \quad h \to \infty,$$

где

$$C_{\gamma,d} = f_X(0) \frac{\Gamma(1-2d)}{\Gamma(d)\Gamma(1-d)}.$$

Доказательство. См. [3, Corollary 6.1]. □

Из следствия 2.1 сразу следует, что процесс FI(d) с $d \in (0, 1/2)$ обладает длинной памятью. При $d \leqslant 0$ процесс обладает короткой памятью [3, Section 6.2].

Пример 2.2. Если $\{X_t\}$ является белым шумом WN$(0, \sigma^2)$, то $\{Y_t\}$ называют дробно интегрированным шумом (FIN). Его спектральная плотность имеет вид

$$f_Y(\omega) = \sigma^2 4^{-d} \sin^{-2d}(\pi\omega).$$

Отсюда следует, что дробно интегрированный шум всегда обладает монотонной спектральной плотностью.

Пример 2.3. Стационарный и обратимый ARMA процесс удовлетворяют предположениям о $\{X_t\}$ в определении 2.2 [3, Proposition 3.5]. Процесс $\{Y_t\}$ в таком случае называют дробно интегрированным ARMA процессом или коротко ARFIMA(p, d, q). Его спектральная плотность имеет вид

$$f_Y(\omega) = \sigma^2 4^{-d} \sin^{-2d}(\pi\omega) \frac{\left|\theta(e^{-2\pi\omega \mathrm{i}})\right|^2}{\left|\phi(e^{-2\pi\omega \mathrm{i}})\right|^2},$$

где ϕ и θ — характеристические полиномы AR и MA частей ARMA соответственно.

2.1. Возникновение процессов с длинной памятью

Нас интересуют процессы с монотонной спектральной плотностью, поскольку они довольно распространены в реальном мире. Такими процессами являются процессы со степенной спектральной плотностью $f(\omega) \sim \omega^{-\alpha}$, имеющие большое применение в различных областях, например, в физике, биологии, астрофизике, геофизике и экономике.

Процессы с длинной памятью, являющиеся частным случаем процессов со степенной спектральной плотностью, довольно распространены. Например, в работе [4] обнаружена длинная память в таких среднегодовых гидрологических временных рядах, как количество осадков, температура и данных о речном стоке. В работе [5] на наличие длинной памяти исследовалась скорость ветра в Ирландии, в работе [6] исследовался эффект длинной памяти у сейсмических данных. Помимо геофизики, длинная память встречается также в финансах [7, 8].

3. Оценка параметров

Будем считать, что $\{X_t\}$ из определения 2.2 представляет собой ARMA(p,q) процесс с гауссовским белым шумом $\{\varepsilon_t\}$. Тогда $f_X(\omega) = f_X(\omega; \boldsymbol{\psi}, \sigma)$, где

$$\boldsymbol{\psi} = (\phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q)^{\mathrm{T}}.$$

Поставим задачу оценить вектор параметров $\boldsymbol{\varphi} = (d, \boldsymbol{\psi}, \sigma)^{\mathrm{T}}$.

3.1. Maximum likelihood estimation (MLE)

Пусть $\{Y_t\}$ — стационарный дробно интегрированный процесс. Тогда вектор

$$Y = (Y_1, \ldots, Y_n)^{\mathrm{T}} \sim \mathcal{N}_n(\mathbf{0}, \Sigma_Y),$$

где $\Sigma_Y = (\gamma_Y(|i-j|))_{i,j=1}^n$ — ковариационная матрица Y. Совместная плотность распределения Y равна

$$(2\pi)^{-n/2} |\Sigma_Y|^{-1/2} \exp\left\{ -\frac{1}{2} Y^{\mathrm{T}} \Sigma^{-1} Y \right\}.$$

Рассмотрим логарифм функции правдоподобия. Отбрасывая аддитивные константы, получаем

$$\ell(Y; \boldsymbol{\varphi}) := -\frac{1}{2} \ln |\Sigma_Y| - \frac{1}{2} Y^{\mathrm{T}} \Sigma_Y^{-1} Y.$$

Тогда $\widehat{\boldsymbol{\varphi}}_{\mathrm{ML}} = \operatorname{argmax} \ell(Y; \boldsymbol{\varphi})$.

3.2. Whittle estimation

При больших n вычисление ковариационной матрицы Σ_Y может быть трудоемким. Поэтому вместо логарифма функции правдоподобия можно рассматривать ее оценку (с точностью до константы) [9]:

$$\ell_W(Y, \boldsymbol{\varphi}) := -\frac{1}{m}\sum_{j=1}^{m}\left(\ln f_Y(\omega_j; \boldsymbol{\varphi}) + \frac{I_Y(\omega_j)}{f_Y(\omega_j; \boldsymbol{\varphi})}\right),$$

где $m = \lfloor (n-1)/2 \rfloor$, $\omega_j = j/n$, $j = 1, 2, \ldots, m$. Заметим, что $f_Y(\omega; \boldsymbol{\varphi}) = \sigma^2 g_Y(\omega; d, \boldsymbol{\psi})$. Тогда

$$\ell_W(Y; \boldsymbol{\varphi}) = -\ln\sigma^2 - \frac{1}{m}\sum_{j=1}^{m}\ln g_Y(\omega_j; d, \boldsymbol{\psi}) - \frac{1}{m}\sigma^{-2}\sum_{j=1}^{m}\frac{I_Y(\omega_j)}{g_Y(\omega_j; d, \boldsymbol{\psi})}.$$

Таким образом, решая уравнение $\frac{d}{d\sigma}\ell_W(Y, \boldsymbol{\varphi}) = 0$, получаем

$$\widehat{\sigma}_W^2 = \frac{1}{m}\sum_{j=1}^{m}\frac{I_Y(\omega_j)}{g_Y(\omega_j; d, \boldsymbol{\psi})}.$$

Остальные параметры находятся, подставляя $\widehat{\sigma}_W^2$ в $\ell_W(Y, \boldsymbol{\varphi})$ и максимизируя полученную функцию.

Список литературы

1. Palma Wilfredo. Long-Memory Time Series: Theory and Methods. — Wiley, 2006.
2. Time Series Analysis: Forecasting and Control / Box G., Jenkins G., Reinsel G., and Ljung G. — Fifth ed. — 2016.
3. Hassler Uwe. Time Series Analysis with Long Memory in View. — Wiley, 2018.
4. Hipel Keith W., McLeod Ian. Time series modelling of water resources and environmental systems. — Elsevier, 1994.
5. Haslett John, Raftery Adrian E. Space-Time Modelling with Long-Memory Dependence: Assessing Ireland's Wind Power Resource // Journal of the Royal Statistical Society. Series C (Applied Statistics). — 1989. — Vol. 38, no. 1. — P. 1–50.
6. Long memory effects and forecasting of earthquake and volcano seismic data / Mariani Maria C., Bhuiyan Md Al Masum, Tweneboah Osei K. and Gonzalez-Huizar Hector // Physica A: Statistical Mechanics and its Applications. — 2020. — Vol. 559. — P. 125049.
7. Barkoulas J., Labys W. C., Onochie J. I. Fractional dynamics in international commodity prices // Journal of Futures Markets. — 1997. — Vol. 17. — P. 161–189.
8. Guglielmo Maria Caporale Luis Gil-Alana, Plastun Alex. Long memory and data frequency in financial markets // Journal of Statistical Computation and Simulation. — 2019. — Vol. 89, no. 10. — P. 1763–1779.
9. Whittle P. The Analysis of Multiple Stationary Time Series // Journal of the Royal Statistical Society. Series B (Methodological). — 1953. — P. 125–139.