Авторегрессионное продолжение ковариационных функций (материалы к специальному курсу)

В.В. Некруткин

кафедра статистического моделирования `http://statmod.ru`, матмех СПбГУ, 2016 г.

Содержание

1 Процессы авторегрессии

1.1 Основные понятия и свойства.

Пусть ξ_n, $n = 0, \pm, \pm 2, \ldots$ — вещественная стационарная в широком смысле последовательность, имеющая нулевое среднее, ковариационную функцию $R_\xi(n)$, спектральную меру m_ξ и ассоциированную стохастическую ортогональную меру μ_ξ.

Если существуют такие натуральные $q \geq 0$ и $p \geq 0$, числа $a_0, \ldots, a_p$, $b_0, \ldots, b_q$, удовлетворяющие условию $a_0 b_0 a_p b_q \neq 0$, а также такой вещественный стандартный белый шум ε_n, что для любого n

$$b_0\xi_n + b_1\xi_{n-1} + \ldots + b_q\xi_{n-q} = a_0\varepsilon_n + a_1\varepsilon_{n-1} + \ldots + a_p\varepsilon_{n-p}, \tag{1.1}$$

то процесс $\xi.$ называется *процессом авторегрессии-скользящего суммирования порядка* (q, p), сокращенно — АРСС(q, p).

Отметим, что среди коэффициентов $a_0, \ldots, a_p$, $b_0, \ldots, b_q$ есть один лишний, но мы будем придерживаться записи (1.1), так как, в зависимости от обстоятельств, иногда будет удобно считать $b_0 = 1$, а иногда — $a_0 = 1$.

При $q = 0$

$$b_0\xi_n = a_0\varepsilon_n + a_1\varepsilon_{n-1} + \ldots + a_p\varepsilon_{n-p},$$

то есть последовательность ξ_n является процессом скользящего суммирования. В случае, когда $p = 0$, ξ_n называется *процессом* (чистой) *авторегрессии* и обозначается АР(q). В дальнейшем нас будет интересовать случай $q > 0$.

Процесс (1.1) называется *реализуемым относительно белого шума* ε (сокращенно — *реализуемым*), если ε_n ортогонально ξ_m при $m < n$.

Поясним последнее определение. Будем считать, что $b_0 = 1$ и перепишем (1.1) в виде

$$\xi_n = -(b_1\xi_{n-1} + \ldots + b_q\xi_{n-q}) + a_0\varepsilon_n + a_1\varepsilon_{n-1} + \ldots + a_p\varepsilon_{n-p}, \tag{1.2}$$

Если теперь заменить в определении белого шума ε_k слово «ортогональность» на «независимость», то выражение (1.2) станет очень похожим на рекуррентную формулу для моделирования последовательности ξ_n.[1] Действительно, ξ_n определяется наборами $(\xi_{n-1}, \ldots, \xi_{n-q})$ и $(\varepsilon_{n-1}, \ldots, \varepsilon_{n-q})$, а также случайной величиной ε_n, которая независима с ε_k при $k < m$.

Поэтому предполагаемый алгоритм должен выглядеть следующим образом: имея на руках реализации случайных величин $(\xi_{n-1}, \ldots, \xi_{n-q})$ и $(\varepsilon_{n-1}, \ldots, \varepsilon_{n-q})$, нужно заново обратиться к генератору псевдослучайных чисел и получить реализацию случайной величины ε_n, независимую от всех предыдущих действий. А потом применяем формулу (1.2) для получения ξ_n.

Но слова «независимо от всех предыдущих действий» и означает, что случайная величина ε_n должна быть независима от всех ξ_k при $k < n$. Меняя обратно «независимость» на «ортогональность», мы и приходим к определению реализуемости.

С понятием реализуемости связано и само название последовательности ξ_n — «процесс авторегрессии-скользящего суммирования». Вторая половина этого термина не требует долгого объяснения: она, конечно, описывает правую часть (1.1).

Чтобы объяснить слово «авторегрессия», рассмотрим реализуемый процесс чистой авторегрессии, который выглядит как

$$\xi_n = -(b_1\xi_{n-1} + \ldots + b_q\xi_{n-q}) + a_0\varepsilon_n.$$

Так как ε_n и $\xi_{n-1}, \ldots, \xi_{n-q}$ ортогональны, то ясно, что коэффициенты $c_k = -b_k$ являются ничем иным, как коэффициентами линейной регрессии ξ_n на $\xi_{n-1}, \ldots, \xi_{n-q}$. Поскольку же все ξ_k вместе составляют единую (стационарную) последовательность, то эта регрессия является «авторегрессией».

[1] Здесь мы не обсуждаем инициализацию этой процедуры.

С коэффициентами $(a_0, \ldots, a_p)$ и $(b_0, \ldots, b_q)$ связывают следующие полиномы комплексной переменной. Введем обозначения

$$Q(z) = b_0 z^q + b_1 z^{q-1} + \ldots + b_{q-1} z + b_q \quad \text{и} \quad Q^*(z) = b_q z^q + b_{q-1} z^{q-1} + \ldots + b_1 z + b_0$$

и аналогичные обозначения P и P^* для полиномов с коэффициентами $(a_0, \ldots, a_p)$. Поскольку $b_0 b_q \neq 0$, то корни полиномов Q, Q^* взаимно обратны. Действительно, так как

$$Q(z) = \sum_{k=0}^{q} b_k z^{q-k} = z^q \sum_{k=0}^{q} b_k z^{-k} = z^q Q^*(z^{-1}), \tag{1.3}$$

то из $Q(z_0) = 0$ следует $Q^*(1/z_0) = 0$ и наоборот. В частности, из (1.3) следует, что полиномы Q и Q^* одновременно имеют корни или не имеют корней на единичной окружности $|z| = 1$.

Предложение 1.1. *Пусть $Q(z) \neq 0$ при $|z| = 1$. Тогда процесс* (1.1) *имеет спектральную плотность f_ξ, выражающуюся формулой*

$$f_\xi(\lambda) = \frac{1}{2\pi} \frac{|P^*(e^{-i\lambda})|^2}{|Q^*(e^{-i\lambda})|^2}, \quad \lambda \in (-\pi, \pi]. \tag{1.4}$$

Доказательство. Обозначим η_n левую (и правую) часть равенства (1.1). Так как $\eta_\cdot$ является линейным преобразованием $\xi_\cdot$, то мы можем стандартным способом вычислить спектральную меру этого процесса. Действительно,

$$\eta_n = \sum_{k=0}^{q} b_k \xi_{n-k} = \sum_{k=0}^{q} b_k \int_{-\pi}^{\pi} e^{i\lambda(n-k)} \mu_\xi(d\lambda) =$$
$$= \int_{-\pi}^{\pi} e^{i\lambda n} \left(\sum_{k=0}^{q} b_k e^{-i\lambda k} \right) \mu_\xi(d\lambda) = \int_{-\pi}^{\pi} e^{i\lambda n} Q^*\big(e^{-i\lambda}\big) \mu_\xi(d\lambda),$$

так что

$$m_\eta(d\lambda) = \big|Q^*\big(e^{-i\lambda}\big)\big|^2 m_\xi(d\lambda).$$

Аналогично из правой части (1.1) получаем, что

$$m_\eta(d\lambda) = \frac{1}{2\pi} \big|P^*\big(e^{-i\lambda}\big)\big|^2 d\lambda.$$

Поэтому

$$\big|Q^*\big(e^{-i\lambda}\big)\big|^2 m_\xi(d\lambda) = \big|P^*\big(e^{-i\lambda}\big)\big|^2 d\lambda$$

и, если $Q^*\big(e^{-i\lambda}\big) \neq 0$ ни при каком λ, то

$$m_\xi(d\lambda) = \frac{1}{2\pi} \frac{|P^*(e^{-i\lambda})|^2}{|Q^*(e^{-i\lambda})|^2} d\lambda.$$

Утверждение доказано. □

В дальнейшем, если не оговорено противное, мы всегда предполагаем, что полином Q не имеет корней на единичной окружности.

Замечание 1.1. Так как $|Q(z)| = |z^q Q^*(1/z)| = |Q^*(1/z)|$ при $|z| = 1$, то спектральная плотность f_ξ может быть представлена также в виде

$$f_\xi(\lambda) = \frac{1}{2\pi} \frac{|P(e^{i\lambda})|^2}{|Q(e^{i\lambda})|^2}.$$

Покажем теперь, что изучение процесса АРСС (1.1) можно свести к изучению процесса чистой авторегрессии с теми же коэффициентами b_k.

Предложение 1.2. *Пусть $\eta_\cdot$ — процесс чистой авторегрессии, имеющий вид*

$$b_0\eta_n + b_1\eta_{n-1} + \ldots + b_q\eta_{n-q} = \varepsilon_n \tag{1.5}$$

и

$$\xi_n = a_0\eta_n + a_1\eta_{n-1} + \ldots + a_p\eta_{n-p}. \tag{1.6}$$

Тогда 1. Процесс $\xi_\cdot$ удовлетворяет (1.1)*. 2. Если процесс $\eta_\cdot$ является реализуемым относительно $\varepsilon_\cdot$, то таким же свойством обладает и $\xi_\cdot$.*

Доказательство. Действительно,

$$\sum_{k=0}^{q} b_k\xi_{n-k} = \sum_{k=0}^{q} b_k \left(\sum_{\ell=0}^{p} a_\ell\, \eta_{n-k-\ell}\right) = \sum_{\ell=0}^{p} a_\ell \left(\sum_{k=0}^{q} b_k\eta_{n-k-\ell}\right) = \sum_{\ell=0}^{p} a_\ell\varepsilon_{n-k}.$$

Второе утверждение очевидным образом следует из (1.5) и (1.6). □

Ввиду Предложения 1.2 дальнейшие рассуждения проводятся для процессов чистой авторегрессии

$$b_0\xi_n + b_1\xi_{n-1} + \ldots + b_q\xi_{n-q} = a_0\varepsilon_n, \tag{1.7}$$

имеющей спектральную плотность

$$f_\xi(\lambda) = \frac{1}{2\pi}\ \frac{|a_0|^2}{|Q^*(e^{-i\lambda})|^2}\ , \quad \lambda \in (-\pi, \pi].$$

Заметим, что эта спектральная плотность не обращается в ноль, поэтому (согласно общему результату теории стационарных в широком смысле последовательностей) процесс ξ_n является линейным преобразованием некоторого белого шума, принадлежащего замыканию в $\mathbb{L}^2(dP)$ всевозможных линейных комбинаций случайных величин ξ_n. Для процессов авторегрессии этот результат можно уточнить.

Предложение 1.3. *Процесс авторегрессии* (1.7) *является линейным преобразованием белого шума ε_n.*

Доказательство. Не умаляя общности, будем считать, что $a_0 = 1$. Согласно (1.7)

$$\varepsilon_n = \int_{(-\pi,\pi]} e^{in\lambda} Q^*\left(e^{-i\lambda}\right)\, \mu_\xi(d\lambda),$$

то есть $\mu_\varepsilon(d\lambda) = Q^*\left(e^{-i\lambda}\right)\, \mu_\xi(d\lambda)$. С другой стороны,

$$\xi_n = \int_{(-\pi,\pi]} e^{in\lambda}\mu_\xi(d\lambda) = \int_{(-\pi,\pi]} e^{in\lambda}\frac{1}{Q^*\left(e^{-i\lambda}\right)}\ Q^*\left(e^{-i\lambda}\right)\mu_\xi(d\lambda) = \int_{(-\pi,\pi]} e^{in\lambda}\frac{1}{Q^*\left(e^{-i\lambda}\right)}\ \mu_\varepsilon(d\lambda).$$

Поскольку функция $1/Q^*\left(e^{-i\lambda}\right)$ принадлежит $\mathbb{L}^2(-\pi,\pi)$, то все доказано. □

Следствие 1.1. Для любого белого шума ε_n и для любых чисел $(b_0, \ldots, b_q)$ таких, что $b_0b_q \neq 0$ и полином $Q*$ не имеет корней на единичной окружности существует процесс авторегрессии ξ_n, удовлетворяющий (1.7).

Доказательство. Снова достаточно рассмотреть случай $a_0 = 1$. Тогда из доказательства Предложения 1.3 следует, что в качестве искомого процесса авторегрессии можно взять

$$\xi_n = \int_{(-\pi,\pi]} e^{in\lambda}\frac{1}{Q^*\left(e^{-i\lambda}\right)}\ \mu_\varepsilon(d\lambda), \tag{1.8}$$

что и доказывает следствие. □

1.2 Реализуемые процессы авторегрессии.

Теорема 1.1. *Пусть $Q^*(z) \neq 0$ при $|z| = 1$. Для того, чтобы процесс авторегрессии* (1.7) *был реализуемым, необходимо и достаточно, чтобы все корни полинома $Q^*(z)$ лежали вне единичного круга.*

Доказательство. Запишем условие реализуемости в явном виде. Скалярно умножив (1.7) на ξ_m при $m < n$, получим, что реализуемость эквивалентна бесконечной системе

$$b_0 R_\xi(n-m) + b_1 R_\xi(n-1-m) + \ldots + b_q R_\xi(n-q-m) = 0, \quad n > m.$$

Обозначая $k = n - m$ получим, что процесс ξ_n реализуем тогда и только тогда, когда при любом $k \geq 1$

$$b_0 R_\xi(k) + b_1 R_\xi(k-1) + \ldots + b_q R_\xi(k-q) = 0. \tag{1.9}$$

Поскольку

$$R_\xi(k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ik\lambda} \frac{|a_0|^2}{|Q^*(e^{-i\lambda})|^2} d\lambda,$$

то (1.9) эквивалентно тому, что при любом $k \geq 1$

$$0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left(\sum_{\ell=0}^{q} b_\ell e^{i(k-\ell)\lambda} \right) \frac{1}{|Q^*(e^{-i\lambda})|^2} \, d\lambda = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ik\lambda} \left(\sum_{\ell=0}^{q} b_\ell e^{-i\ell\lambda} \right) \frac{1}{|Q^*(e^{-i\lambda})|^2} \, d\lambda =$$
$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ik\lambda} \frac{Q^*(e^{-i\lambda})}{|Q^*(e^{-i\lambda})|^2} \, d\lambda = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ik\lambda} \frac{1}{\overline{Q^*(e^{-i\lambda})}} \, d\lambda = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ik\lambda} \frac{1}{Q^*(e^{i\lambda})} \, d\lambda =$$
$$= \frac{1}{2\pi i} \int_{|z=1|} z^{k-1} \frac{1}{Q^*(z)} \, dz.$$

Как хорошо известно из теории функций комплексной переменной, интеграл от мероморфной функции по замкнутому контуру равен нулю, если это все нули этой функции расположены вне контура. Таким образом (так как $k \geq 1$), процесс авторегрессии будет реализуемым, если функция $Q^*(z)$ не имеет корней внутри единичного круга.

Доказательство необходимости для простоты проведем в случае, когда все корни $z_1, \ldots, z_m$ ($m > 0$) полинома $Q^*(z)$, лежащие внутри единичного круга, являются простыми. Поскольку среди чисел z_ℓ нет нуля, то при любом $k \geq 1$

$$\frac{1}{2\pi i} \int_{|z=1|} z^{k-1} \frac{1}{Q^*(z)} \, dz = \sum_{\ell=1}^{m} \mathrm{Res}_{v_\ell} \left(\frac{z^{k-1}}{Q^*(z)} \right) = \sum_{\ell=1}^{m} \frac{v_\ell^{k-1}}{H(v_\ell)} = 0,$$

где $H(z) = dQ^*(z)/dz$. Этого, конечно, быть не может, так как вектора $(1, v_\ell, \ldots, v_\ell^s)^{\mathrm{T}}$ линейно независимы при достаточно больших s. В случае кратных корней доказательство сохраняется, но линейно независимые вектора выглядят более сложно. □

Замечание 1.2. В терминах полинома Q условия Теоремы 1.1 означают, что все корни этого полинома лежат внутри единичного круга.

Следствие 1.2. Пусть $Q^*(z) \neq 0$ при $|z| = 1$. Процесс авторегрессии (1.1) является реализуемым относительно белого шума ε_n тогда и только тогда, когда он может быть представлен в виде реализуемого фильтра белого шума ε_n.

Доказательство. Достаточность условия очевидна. Для доказательства необходимости достаточно рассмотреть случай чистой авторегрессии (1.5). Тогда полином $Q^*(z)$ не имеет корней при $|z| \le 1$ и верно равенство (1.8). Поэтому $\xi_n = \sum_k c_k \varepsilon_{n-k}$, где

$$c_k = \int_{-\pi}^{\pi} e^{-ik\lambda} \frac{1}{Q^*\left(e^{-i\lambda}\right)}\, d\lambda = 0$$

при $k < 0$. □

Предложение 1.4. *Пусть $b_0, \ldots, b_q$ и $a_0, \ldots, a_p$ — некоторые числа, причем $b_0 b_q a_0 a_q \ne 0$. Обозначим $Q^*(z) = \sum_{k=0}^{q} b_k z^k$ и $P^*(z) = \sum_{k=0}^{p} a_k z^k$. Если $Q^*(z) \ne 0$ при $|z| = 1$, то для любого белого шума ε_n существует стационарная последовательность АРСС(q,p), реализуемая относительного ε_n и имеющая спектральную плотность* (1.4).

Доказательство. Во первых, нам достаточно доказывать результат лишь для процесса чистой регрессии, то есть в случае $P^*(z) = 1$. Далее, если все корни полинома $Q^*(z)$ лежат вне единичного круга, то нам достаточно определить ξ_n формулой (1.8).

Рассмотрим случай, когда $Q^*(z) = \left(\prod_{j=1}^{d}(z - z_j)\right) R(z)$, где все нули функции $R(z)$ лежат вне единичного круга, а $|z_j| < 1$ при $j = 1, \ldots, d$. Далее,

$$|Q^*(z)|^2 = Q^*(z)\overline{Q^*(z)} = |R(z)|^2 \prod_{j=1}^{d}(z - z_j)(\overline{z} - \overline{z}_j).$$

Заметим, что в формуле для спектральной плотности полином $Q^*(z)$ нужен только при $|z| = 1$. Поэтому при $|z| = 1$

$$(z - z_j)(\overline{z} - \overline{z}_j) = |z|^2\, |z_j|^2 (1/z_j - 1/z)(1/\overline{z}_j - 1/\overline{z}) = |z_j|^2 (\overline{z} - 1/z_j)(z - 1/\overline{z}_j),$$

и мы приходим к полиному

$$Q_1^*(z) = \prod_{j=1}^{d} z_j \left(\prod_{j=1}^{d}(z - 1/z_j)\right) R(z),$$

который совпадает с $Q^*(z)$ при $|z| = 1$ и все корни которого лежат вне единичного круга. Положив

$$\xi_n = \int_{(-\pi,\pi]} e^{in\lambda} \frac{1}{Q_1^*\left(e^{-i\lambda}\right)}\, \mu_\varepsilon(d\lambda),$$

получим требуемое. □

1.3 Авторегрессия АР(1) и марковские гауссовские стационарные последовательности.

Рассмотрим вещественный реализуемый процесс чистой авторегрессии первого порядка, который можно записать в форме

$$\xi_n = a\xi_{n-1} + b\varepsilon_n, \tag{1.10}$$

где $a, b \in \mathbb{R}$ и $\mathbb{E}(\xi_m \varepsilon_n) = 0$ при $m < n$. Из этих условий сразу же следует, что $R_\xi(n - m) = aR_\xi(n - m - 1)$ при любых $n > m$, то есть $R_\xi(k) = a^k R_\xi(0)$, $k > 0$. Тем самым $\rho_\xi(k) = a^k$ и, следовательно, $|a| \le 1$. Для наглядности записи будем вместо a употреблять букву ρ. Также положим $R_\xi(0) = \sigma^2$ и выразим через ρ и σ число b.

Домножив обе части (1.10) на ξ_n и взяв математическое ожидание, получим, что

$$\sigma^2 = \rho R_\xi(1) + b\mathbb{E}\xi_n\varepsilon_n = \rho^2\sigma^2 + b\mathbb{E}(\xi_n\varepsilon_n).$$

С другой стороны, домножив (1.10) на ε_n, поучим после осреднения, что $\mathbb{E}(\xi_n\varepsilon_n) = b$. Тем самым приходим к равенству $b^2 = \sigma^2(1-\rho^2)$. Выбирая знак "плюс" перед квадратным корнем, приходим к представлению

$$\xi_n = \rho\xi_{n-1} + \sigma\sqrt{1-\rho^2}\,\varepsilon_n, \quad |\rho| \le 1, \quad \sigma > 0, \tag{1.11}$$

которым и будем пользоваться в дальнейшем. Разберем сначала случай $|\rho| = 1$. Если $\rho = 1$, то $\xi_n = \xi_{n-1}$ и спектральная мера m_ξ процесса ξ_n сосредоточена в нуле. При $\rho = -1$, получаем, что $\xi_n = -\xi_{n-1}$ и, следовательно, m_ξ сосредоточена в точке π. В обоих случаях спектральной плотности не существует.

Для проверки условий существования спектральной плотности обратимся к Предложению 1.1. В нашем случае полином Q^* имеет вид $Q^*(z) = 1 - az$, то есть условие Предложения 1.1 приобретает вид $|a| \ne 1$. С учетом того, что мы исследуем реализуемые процессы, приходим к выводу, что необходимое и достаточное условием существования спектральной плотности у процесса авторегрессии (1.11) является неравенство $|\rho| < 1$, и при этом

$$f_\xi(\lambda) = \frac{1}{2\pi}\,\frac{\sigma^2(1-\rho^2)}{|1-\rho e^{i\lambda}|^2} = \frac{1}{2\pi}\,\frac{\sigma^2(1-\rho^2)}{1+\rho^2-2\rho\cos(\lambda)}\,. \tag{1.12}$$

Легко видеть, что спектральная плотность (1.12) монотонно убывает при $\rho > 0$ и монотонно возрастает при $\rho < 0$ на отрезке $[0,\pi]$ (а при $\rho = 0$ последовательность ξ_n просто является белым шумом). При этом

$$f_\xi(0) = \frac{1}{2\pi}\,\frac{\sigma^2(1-\rho^2)}{(1-\rho)^2} = \frac{\sigma^2}{2\pi}\,\frac{1+\rho}{1-\rho} \quad \text{и} \quad f_\xi(\pi) = \frac{\sigma^2}{2\pi}\,\frac{1-\rho}{1+\rho}\,.$$

Пусть $\rho \to 1$. Нетрудно видеть, что тогда $f_\xi(\lambda) \to 0$ при $\lambda > 0$ в то время как $f_\xi(0) \to \infty$. Это означает, что в спектральном разложении процесса (1.11) при $\rho \approx 1$ будут доминировать низкие частоты (а на уровне реализаций процесса — низкочастотные колебания). Поскольку в оптике такая ситуация характерна для красного цвета, то процесс авторегрессии с ρ, близким к $+1$, называют *красным шумом*.

Если $\rho \approx -1$, то ситуация аналогичная, только реализации процесса будут походить на последовательность с периодом 2.

Легко найти явное представление процесса авторегрессии (1.11) как линейного преобразования белого белого шума ε_n. Разлагая функцию $1/Q^*(e^{-i\lambda})$, как это обозначено в Следствии 1.2, получим, что

$$\xi_n = \sigma\sqrt{1-\rho^2}\sum_{k=0}^{\infty}\rho^k\varepsilon_{n-k} = \sqrt{1-\rho^2}\varepsilon_n + \sqrt{1-\rho^2}\sum_{k=1}^{\infty}\rho^k\varepsilon_{n-k} =$$

$$= \sqrt{1-\rho^2}\varepsilon_n + \rho\sqrt{1-\rho^2}\sum_{k=1}^{\infty}\rho^{k-1}\varepsilon_{n-k} = \sqrt{1-\rho^2}\varepsilon_n + \rho\sqrt{1-\rho^2}\sum_{\ell=0}^{\infty}\rho^{\ell}\varepsilon_{n-1-\ell} = \sqrt{1-\rho^2}\varepsilon_n + \rho\xi_{n-1}.$$

Наконец, обратимся к связи между реализуемыми процессами авторегрессии первого порядка и марковскими гауссовскими стационарными последовательностями.

Предложение 1.5. *Пусть ξ_n, $n = 0, 1, \ldots,$ — вещественная гауссовская стационарная последовательность с нулевым средним. Для того, чтобы она была марковской, необходимо и достаточно, ξ_n было реализуемой авторегрессией первого порядка относительно некоторого гауссовского белого шума.*

Доказательство. Пусть последовательность ξ_n марковская. Рассмотрим разложение

$$\xi_n = \mathbb{E}(\xi_n \,|\, \xi_{n-1}, \ldots, \xi_0) + \sigma_n\varepsilon_n,$$

где $\mathbb{E}\varepsilon_n = 0$, $\mathbb{D}\varepsilon_n = 1$ и ε_n ортогональна любой случайной величине вида $\varphi(\xi_0, \ldots, \xi_{n-1})$, имеющей конечный второй момент.

Поскольку вектор $(\xi_0, \ldots, \xi_n)$ — гауссовский, то $\varepsilon_n \in \mathrm{N}(0,1)$, $\mathbb{E}(\xi_n \,|\, \xi_{n-1}, \ldots, \xi_0) = \sum_{j=0}^{n-1} c_j\xi_j$ и случайная величина ε_n не зависит от $(\xi_0, \ldots, \xi_n)$.

Марковость последовательности ξ_n означает, что $\mathbb{P}(\xi_n \in A \,|\, \xi_{n-1}, \ldots, \xi_0) = \mathbb{P}(\xi_n \in A \,|\, \xi_{n-1})$ и поэтому $\mathbb{E}(\xi_n \,|\, \xi_{n-1}, \ldots, \xi_0) = \mathbb{E}(\xi_n \,|\, \xi_{n-1}) = c_{n-1}\xi_{n-1}$, так как вектор (ξ_{n-1}, ξ_n) гауссовский. Наконец, из стационарности ξ_n следует, что числа σ_n и c_n не зависят от n.

Пусть теперь гауссовская стационарная последовательность ξ_n имеет вид

$$\xi_n = a\xi_{n-1} + b\varepsilon_n, \tag{1.13}$$

где ε_n ортогонален ξ_m при $m = 0, \ldots, n-1$ и $\mathbb{E}\varepsilon_n^2 = 1$. Если в рекуррентной формуле (1.13) эту ортогональность можно заменить на независимость, то марковость последовательности становится очевидной по построению. Но, поскольку вектор (ξ_{n-1}, ξ_n) —гауссовский, то $\varepsilon_n \in \mathrm{N}(0,1)$ и ε_n не зависит от ξ_m. Доказательство завершено. □

2 Авторегрессионные продолжения

2.1 Уравнения Юла-Уолкера

Пусть $R(0), \ldots, R(p)$ — некоторые вещественные числа и $R(-k) = R(k)$. При $t \le p+1$ обозначим $\Sigma_t = \left\{r_{k\,l}\right\}_{k,\,l=1}^{t}$ с $r_{k\,l} = R(k-l)$.

Лемма 2.1. *Если матрица Σ_{p+1} неотрицательно определена, а матрица Σ_p положительно определена, то существует единственное решение $A = (a_1, \ldots, a_p)^{\mathrm{T}}$ системы*

$$R(m) + \sum_{k=1}^{p} a_k R(m-k) = 0, \quad m = 1, \ldots, p\ , \tag{2.1}$$

причем это решение удовлетворяет неравенству

$$\sigma^2 \stackrel{\mathrm{def}}{=} R(0) + \sum_{k=1}^{p} a_k R(k) \ge 0. \tag{2.2}$$

Если к тому же матрица Σ_{p+1} положительно определена, то $\sigma^2 > 0$.

Доказательство. Обозначим $B = (R(1), \ldots, R(p))^{\mathrm{T}}$ и $A = (a_1, \ldots, a_p)^{\mathrm{T}}$. Тогда система (2.1) перепишется в виде $B + \Sigma_p A = \mathbf{0}$. Поскольку матрица Σ_p — невырожденная, то единственность вектора $A = -\Sigma_p^{-1} B$ очевидна.

Нам будет удобно при доказательстве (2.2) перейти на вероятностный язык. Возьмем случайный вектор $(\eta_p, \ldots, \eta_0)^{\mathrm{T}}$ с распределением $\mathbb{N}(\mathbf{0}, \Sigma_{p+1})$ и заметим, что по определению $\mathbb{D}\eta_m = R(0)$ и $\mathbb{E}\eta_j\eta_{m-j} = R(m)$.

Рассмотрим условное математическое ожидание $\mathbb{E}(\eta_p \,|\, \overline{\eta}_{p-1})$ с $\overline{\eta}_{p-1} = (\eta_{p-1}, \ldots, \eta_0)^{\mathrm{T}}$, которое, как хорошо известно, является линейной функцией случайных величин $\eta_0, \ldots, \eta_{p-1}$. Запишем этот факт в виде равенства

$$\mathbb{E}(\eta_p \,|\, \eta_0, \ldots, \eta_{p-1}) = \sum_{k=1}^{p} c_k \eta_{p-k} = C^{\mathrm{T}} \overline{\eta}_{p-1}.$$

Поскольку ковариационная матрица вектора $\overline{\eta}_{p-1}$ есть Σ_p, а $\mathbb{E}\eta_p\eta_{m-p} = R(m)$, то $C = \Sigma_p^{-1} B = -A$.

Как хорошо известно, дисперсия случайной величины $e_p \stackrel{\mathrm{def}}{=} \eta_p - C^{\mathrm{T}}\overline{\eta}_{p-1}$ (так называемая «остаточная дисперсия») имеет вид

$$0 \le \mathbb{D}e_p = \mathbb{D}\eta_p - B^{\mathrm{T}}\Sigma_p^{-1}B = R(0) + B^{\mathrm{T}}A,$$

что в точности совпадает с (2.2).

Для окончания доказательства осталось заметить, что, если остаточная дисперсия равна нулю, то $e_p = 0$, η_p является линейной комбинацией $\eta_0, \ldots, \eta_{p-1}$ и, следовательно, матрица Σ_{p+1} вырождена. □

Лемма 2.2. *Пусть $R(n)$ — ненулевая вещественная неотрицательно определенная функция и m — соответствующая спектральная мера. Если m обладает спектральной плотностью f, то для любого $p \ge 1$ матрица $\Sigma_p = \left\{R(l-k)\right\}_{k,l=0}^{p-1}$ является положительно определенной.*

Доказательство. Рассмотрим комплекснозначные $z_0, \ldots, z_{p-1}$. Так как

$$R(n) = \int_{-\pi}^{\pi} e^{in\lambda} f(\lambda) d\lambda,$$

то

$$\sum_{k,\ell=1}^{p} R(k-\ell) z_k \bar{z}_\ell = \int_{-\pi}^{\pi} \left(\sum_{k,\ell=1}^{p} e^{i(k-\ell)\lambda} z_k \bar{z}_\ell \right) f(\lambda) d\lambda = \int_{-\pi}^{\pi} \left| \sum_{k=1}^{p} e^{ik\lambda} z_k \right|^2 f(\lambda) d\lambda.$$

Поскольку функция

$$g(\lambda) = \left| \sum_{k=1}^{p} e^{ik\lambda} z_k \right|^2$$

представляет собой тригонометрический полином, то она (кроме тривиального случай $z_0 = \ldots = z_{p-1} = 0$) имеет на отрезке $(-\pi, \pi)$ лишь конечное число корней. Следовательно,

$$\sum_{k,\ell=1}^{p} R(k-\ell) z_k \bar{z}_\ell = 0$$

лишь в случае $z_0 = \ldots = z_{p-1} = 0$, и утверждение доказано. □

Рассмотрим реализуемый вещественный авторегрессионный процесс порядка $p+1$ (или меньшего порядка, если $a_p = 0$):

$$\xi_n + \sum_{k=1}^{p} a_k \xi_{n-k} = \sigma \varepsilon_n, \tag{2.3}$$

где $\sigma > 0$ и $\varepsilon_1, \ldots, \varepsilon_n, \ldots$ — (нормированный и центрированный) белый шум. Как всегда, предполагается, что $\mathbb{E}\xi_n = 0$. Реализуемость ξ_n (относительно белого шума ε_n) означает, что $\mathbb{E}\xi_n \varepsilon_m = 0$ при $m > n$. Коэффициенты a_m предполагаются вещественными.

Спектральная плотность процесса авторегрессии (2.3) имеет вид

$$f_\xi(\lambda) = \frac{1}{2\pi} \frac{\sigma^2}{\left|1 + \sum_{k=1}^{p} a_k e^{-ik\lambda}\right|^2}, \quad |\lambda| \le \pi.$$

Предложение 2.1 (Уравнения Юла-Уолкера)**.** *Пусть $R_\xi(n)$ — ковариационная функция процесса* (2.3)*. Тогда*
1. При $m \ge 1$

$$R_\xi(m) + \sum_{k=1}^{p} a_k R_\xi(m-k) = 0. \tag{2.4}$$

2. Числа $a_1, \ldots, a_p$ однозначно определяются системой уравнений (2.4) *при $m = 1, \ldots, p$ и удовлетворяют неравенству*

$$R_\xi(0) + \sum_{k=1}^{p} a_k R_\xi(k) = \sigma^2 > 0. \tag{2.5}$$

Доказательство. Первое утверждение уже доказано в Теореме 1.1 о реализуемых процесса авторегрессии. Докажем второе утверждение. Умножим обе части (2.3) на ξ_n и возьмем математическое ожидание. Получим, что

$$R_\xi(0) + \sum_{k=1}^{p} a_k R_\xi(k) = \sigma \mathbb{E}\xi_n \varepsilon_n.$$

С другой стороны, домножая (2.3) ε_n и снова беря математическое ожидание приходим к равенству

$$\mathbb{E}\xi_n \varepsilon_n = -\sum_{l=1}^{p} a_l \mathbb{E}\xi_{n-l}\varepsilon_n + \sigma \mathbb{E}\varepsilon_n^2 = \sigma.$$

Так как $\sigma^2 > 0$, то утверждение доказано. □

Замечание 2.1. 1. Положительность левой части (2.5) непосредственно следует из лемм 2.1, 2.2.
2. Система уравнений (2.4) при $1 \le m \le p$ называется *системой уравнений Юла-Уолкера.*

2.2 Теорема об авторегрессионном продолжении

Предложение 2.2. *Пусть η_n $(n \geq 0)$ — вещественный стационарный в широком смысле процесс с ковариационной функцией $R_\eta(m)$ и ненулевой спектральной плотностью $f_\eta(\lambda)$. Тогда для любого $p \geq 0$ существует реализуемый процесс авторегрессии ξ_n порядка, не большего $p+1$ такой, что $R_\xi(m) = R_\eta(m)$ при $m = 0, \ldots, p$.*

Доказательство. Конечно, можно считать, что $\mathbb{E}\eta_n = 0$. При $p = 0$ результат очевиден: достаточно в качестве ξ_n взять белый шум с дисперсией $R_\xi(0)$. Рассмотрим случай $p > 0$.

Согласно Лемме 2.2, для любого p матрица $\Sigma_{p+1} = \left\{R_\eta(l-k)\right\}_{k,l=1}^{p+1}$ является положительно определенной. Поэтому (см. Лемму 2.1) ковариации $R_\eta(n)$ удовлетворяют равенствам

$$R_\eta(m) + \sum_{k=1}^{p} a_k R_\eta(m-k) = 0, \quad m = 1, \ldots, p\,, \tag{2.6}$$

причем при фиксированном p числа $a_1, \ldots, a_p$ определяются однозначно, а

$$\sigma_p^2 \stackrel{\text{def}}{=} R_\eta(0) + \sum_{k=1}^{p} a_k R_\eta(k) > 0. \tag{2.7}$$

Построим процесс ξ_n следующим образом. При $n = 0, \ldots, p-1$ положим $\xi_n = \eta_n$. При $n \geq p$ определим

$$\xi_n = -\sum_{k=1}^{p} a_k \xi_{n-k} + \sigma_p\, \varepsilon_n, \tag{2.8}$$

где (центрированный и нормированный) белый шум ε_n ортогонален ξ_j при $0 \leq j < n$.

Рассмотрим сначала случай $n = p$, то есть равенство

$$\xi_p = -\sum_{k=1}^{p} a_k \xi_{p-k} + \sigma_p\, \varepsilon_p. \tag{2.9}$$

Домножив обе части (2.9) скалярно на ξ_ℓ при $0 \leq \ell < p$ и воспользовавшись тем, что $\xi_\ell = \eta_\ell$ и $\xi_{p-k} = \eta_{p-k}$, получим равенство

$$\mathbb{E}\xi_p\, \xi_\ell = -\sum_{k=1}^{p} a_k \mathbb{E}\eta_{p-k}\eta_\ell = -\sum_{k=1}^{p} a_k R_\eta(p-\ell-k). \tag{2.10}$$

Используя равенство (2.6) с $m = p - \ell$ получаем, что $\mathbb{E}\xi_p\, \xi_\ell = R_\eta(p-\ell)$, так как $1 \leq p - \ell \leq p$.

Теперь перейдем к случаю $\ell = p$. Тогда, так как $0 \leq p - k < p$, то по уже доказанному

$$\mathbb{E}\xi_p^2 = -\sum_{k=1}^{p} a_k \mathbb{E}\xi_{p-k}\xi_p + \sigma_p \mathbb{E}\xi_p\varepsilon_p = -\sum_{k=1}^{p} a_k R_\eta(k) + \sigma_p \mathbb{E}\xi_p\varepsilon_p,$$

Как следует из (2.9), $\mathbb{E}\eta_p\varepsilon_p = \sigma_p$ и поэтому окончательно получаем, что

$$\mathbb{E}\xi_p^2 = -\sum_{k=1}^{p} a_k \mathbb{E}\xi_{p-k}\xi_p + \sigma_p \mathbb{E}\xi_p\varepsilon_p = . - \sum_{k=1}^{p} a_k R_\eta(k) + \sigma_p^2,$$

то есть, согласно (2.7), $\mathbb{E}\xi_p^2 = R_\eta(0)$.

Далее все по индукции. Предположим, что для любых ℓ, m таких, что $0 \leq \ell \leq m < n$ ковариация $\mathbb{E}\xi_n\xi_\ell$ зависит только от разности $m - \ell$ (обозначим эту ковариацию $R_\xi(m-\ell)$), а при $0 \leq m - \ell \leq p$ эта ковариация совпадает с $R_\eta(m-\ell)$. Нужно доказать, что это же свойство имеет место при $m = n$ и $0 \leq \ell \leq m$.

Обозначим лишь ход рассуждений. Снова рассматриваем (2.7), скалярно умножаем это равенство на ξ_ℓ при $\ell < n$ и приходим к равенству

$$\mathbb{E}\xi_n\xi_\ell = -\sum_{k=1}^{p} a_k \mathbb{E}\xi_{n-k}\xi_\ell.$$

Согласно индуктивному предположению, $\mathbb{E}\xi_{n-k}\xi_\ell$ зависит только от модуля разности $|n-\ell-k|$ и равна $R_\xi(|n-\ell-k|)$, так что $\mathbb{E}\xi_{n-k}\xi_\ell$ действительно зависит только разности $n-\ell$.

Далее, если $n-\ell \le p$, то, как нетрудно видеть, $0 \le |n-k-\ell| < p$ при $1 \le k \le p$ и $0 \le \ell \le n$. Поэтому при этих условиях $\mathbb{E}\xi_{n-k}\xi_\ell = R_\eta(n-k-\ell)$ и мы снова используем равенство (2.6).

Наконец, доказательство того, что $\mathbb{E}\xi_n^2 = R_\eta(0)$, повторяет уже приведенные рассуждения для $n = p$. Утверждение доказано. □

Замечание 2.2. Как доказано в Предложении 2.2, процесс ξ_n является реализуемым процессов авторегрессии. Порядок этой авторегрессии может оказаться меньше p, если $a_p = 0$.

Последовательность $R_\xi(m)$ можно называть *авторегрессионным продолжением* отрезка $R_\eta(0), \ldots, R_\eta(p)$ ковариационной функции процесса η_n. Очевидно, ковариации $R_\xi(m)$ при $m > p$ определяются рекуррентными равенствами Юла-Уолкера (2.4).

Замечание 2.3. Предположим, что процесс η_n Предложения 2.2 является реализуемым процессом авторегрессии порядка d:

$$\eta_n = -\sum_{k=1}^{d} b_k \eta_{n-k} + \sigma\varepsilon'_n.$$

Если в Предложении 2.2 выбрать $p = d$, то авторегрессионным продолжением ряда $R_\eta(0), \ldots, R_\eta(d)$ будет сама ковариационная функция $R_\eta(n)$. Действительно, в этом случае коэффициенты a_k в (2.6) будут равны b_k, а число σ_p^2 в (2.7) равно σ^2. Поэтому последовательность (2.8) можно расценивать как другую (отличающуюся лишь выбором белого шума) реализацию исходного процесса η_n.

Если $p < d$, то это уже не так просто потому, что по определению $b_d \ne 0$, а процесс авторегрессии (2.8) имеет порядок, меньший d.

Рассмотрим случай $p > d$. Согласно (2.4),

$$R_\eta(m) = \sum_{k=1}^{d} b_k R_\eta(m-k)$$

при $m \ge 1$. В то же время выполнено (2.6), причем коэффициенты a_k однозначно определены. Отсюда сразу же следует, что $a_k = b_k$ при $k \le d$ и $a_k = 0$ при $k > d$. Поэтому $\sigma_p^2 = \sigma$ и (как и в случае $p = d$) авторегрессионным продолжением отрезка $R_\eta(0), \ldots, R_\eta(d)$ будет ковариационная функция $R_\eta(n)$.

2.3 Спектральные плотности авторегрессионных продолжений.

Пусть $p \to \infty$. Рассмотрим последовательность $\{R_\xi^{(p)}(m), m \geq 0\}$ авторегрессионных продолжений отрезков $R_\eta(0), \ldots, R_\eta(p-1)$ ковариационной функции процесса η_n. Естественно ожидать, что спектральные плотности $f_\xi^{(p)}(\lambda)$, соответствующие ковариационным функциям $R_\xi^{(p)}$, будут сходиться к спектральной плотности $f_\eta(\lambda)$ исходного процесса η_n.

Легко доказать, что это имеет место в некотором слабом смысле. Для этого нам понадобится следующая лемма.

Лемма 2.3. *Пусть $R_n(k)$ — последовательность ковариационных функций со спектральными мерами m_n, сосредоточенными на $(-\pi, \pi]$. Если для любого k имеет место сходимость $R_n(k) \to R(k)$, где $R(k)$ — ковариационная функция со спектральной мерой m, сосредоточенной в $(-\pi, \pi)$, то $m_n \Rightarrow m$ в том смысле, что $\int_{(-\pi,\pi]} g dm_n \to \int_{(-\pi,\pi]} g dm$ для любой ограниченной непрерывной в $(-\pi, \pi]$ функции g.*

Доказательство. Перейдем от ковариационных функций к корреляционным. А именно, положим $\rho_n(k) = R_n(k)/R_n(0)$ и $\rho(k) = R(k)/R(0)$, а также $\mathcal{P}_n = m_n/R_n(0)$ и $\mathcal{P} = m/R(0)$. Ясно, что $\mathcal{P}_n$ и $\mathcal{P}$ являются распределениями и одновременно — спектральными мерами последовательностей ρ_n и ρ соответственно. Кроме того, сходимость $R_n(k) \to R(k)$ эквивалентна сходимости $\rho_n(k) \to \rho(k)$, а слабая сходимость $m_n \Rightarrow m$ — слабой сходимости $\mathcal{P}_n \Rightarrow \mathcal{P}$. Таким образом, мы имеем

$$\rho_n(k) = \int_{(-\pi,\pi]} e^{ik\lambda} \mathcal{P}_n(d\lambda) \to \rho(k) = \int_{(-\pi,\pi]} e^{ik\lambda} \mathcal{P}(d\lambda).$$

Пусть ξ_n — случайная величина, имеющая распределение $\mathcal{P}_n$, а $\mathcal{L}(\xi) = \mathcal{P}$. Тем самым $\mathbb{E}e^{ik\xi_n} \to \mathbb{E}e^{ik\xi}$. Введем случайные величины $\eta_n = (\xi_n + \pi)/2\pi$ и $\eta = (\xi_n + \pi)/2\pi$. Ясно, что распределение $\mathcal{Q}_n = \mathcal{L}(\xi_n)$ сосредоточено на $(0, 1]$, а распределение $\mathcal{Q} = \mathcal{L}(\xi)$ — на $(0, 1)$. Кроме того,

$$\rho_n(k) = \mathbb{E}e^{ik\xi_n} = \mathbb{E}e^{ik(2\pi\eta_n - \pi)} = e^{-\pi k}\mathbb{E}e^{i2\pi k\eta_n} \to \rho(k) = \mathbb{E}e^{ik\xi} = \mathbb{E}e^{ik(2\pi\eta - \pi)} = e^{-\pi k}\mathbb{E}e^{i2\pi k\eta}$$

Поэтому $\mathbb{E}e^{i2\pi k\eta_n} \to \mathbb{E}e^{i2\pi k\eta}$. Из теоремы о характеризации сходимости на единичном торе ([14, гл. 7 §7]) следует, что $\mathcal{Q}_n \Rightarrow \mathcal{Q}$. Поэтому (так как линейное преобразование непрерывно) $\mathcal{P}_n \Rightarrow \mathcal{P}$ и утверждение доказано. □

Следствие 2.1. В случае, когда у предельной меры m нет нагрузок на всем промежутке $(-\pi, \pi]$, из этого утверждения следует в частности, что для любого отрезка $A = [a, b] \subset (-\pi, \pi]$ имеет место сходимость $m_n(A) \to m(A)$.

Кроме того, легко доказать, что любая непрерывная отделенная от нуля спектральная плотность может быть равномерно приближена спектральной плотностью некоторого процесса авторегрессии.

Предложение 2.3. *Пусть $f(\lambda)$ — спектральная плотность некоторой вещественной стационарной последовательности. Если f непрерывна и отделена от нуля, то существует такой реализуемый процесс авторегрессии, что его спектральная плотность отличается от f сколь угодно мало в равномерной метрике.*

Доказательство. Пусть $0 < m^2 \leq f(\lambda) \leq M^2$. Обозначим $h(\lambda) = 1/\sqrt{f(\lambda)}$. Тогда $1/M \leq h(\lambda) \leq 1/m$. Функция h, очевидно, непрерывна и может быть продолжена на всю вещественную ось по периодичности, причем продолжение тоже будет непрерывной функцией.

По теореме Вейерштрасса о тригонометрических полиномах функцию h можно равномерно приблизить с точностью до сколь угодно малого ε с помощью тригонометрического полинома Q периода 2π. Ясно, что $1/M - \varepsilon \leq Q(\lambda) \leq 1/m + \varepsilon$.

Поскольку всегда можно представить положительный полином Q в виде

$$Q(\lambda) = \sum_{k=0}^{q} c_k e^{-ik\lambda} = \left| \sum_{k=0}^{q} c_k e^{-ik\lambda} \right|,$$

то в обозначениях

$$\widehat{f}(\lambda) = \frac{1}{|Q(\lambda)|^2}$$

получим, что

$$\left|\widehat{f}(\lambda) - f(\lambda)\right| \le \left|\frac{1}{Q(\lambda)} - \frac{1}{h(\lambda)}\right| \left|\frac{1}{Q(\lambda)} + \frac{1}{h(\lambda)}\right| \le \max_{\lambda} \left(1/h(\lambda) + 1/Q(\lambda)\right) \left|\frac{h(\lambda) - Q(\lambda)}{h(\lambda)Q(\lambda)}\right| \le$$
$$\le \frac{2}{1/M - \varepsilon} \, \frac{\varepsilon M}{1/M - \varepsilon}.$$

Тем самым $\widehat{f}$ равномерно аппроксимирует f. □

Предложение 2.3 говорит о том, что непрерывная строго положительная спектральная плотность может быть приближена сколь угодно точно с помощью спектральной плотности некоторого процесса авторегрессии. В то же время остается непонятным, можно ли это сделать с помощью обсуждаемой специальной процедуры.

Лемма 2.3, напротив, рассматривает (в том числе) эту специальную процедуру, но гарантирует лишь слабую сходимость спектральных мер вместо равномерной сходимости плотностей.

На самом деле имеет место следующий результат [11, разд. 3.5].

Теорема 2.1. *Пусть $R_\eta(k)$ — ковариационная функция некоторого вещественного стационарного процесса, имеющего непрерывную ограниченную и отделенную от нуля спектральную плотность $f_\eta(\lambda)$, удовлетворяющую условию Гельдера с показателем γ. Обозначим $f_\xi^{(p)}(\lambda)$ спектральную плотность реализуемого процесса авторегрессии, первые p значений ковариационной функции которого совпадают со значениями R_η. Тогда существует такая положительная постоянная C, зависящая от f_η, что*

$$\left|f_\xi^{(p)}(\lambda) - f_\eta(\lambda)\right| \le Cp^{-\gamma}.$$

Эта теорема объясняет популярность обсуждаемого метода. Но доказательство теоремы довольно сложно и поэтому опускается.

2.4 Авторегрессионные продолжения и система ортогональных полиномов.

Материал этого раздела тесно связан с теорией ортогональных полиномов на окружности, восходящей к [13, гл. XI, XII].

Пусть R и f — соответственно ковариационная функция и спектральная плотность некоторой стационарной последовательности. Для комплекснозначных функций θ_1, θ_2, определенных на единичной окружности $|z| = 1$ и удовлетворяющих условию

$$\int\limits_{-\pi}^{\pi} \left|\theta_k(e^{i\lambda})\right|^2 f(\lambda)d\lambda < \infty, \quad k = 1, 2,$$

определим скалярное произведение равенством

$$\langle \theta_1, \theta_2 \rangle_f = \frac{1}{2\pi} \int\limits_{-\pi}^{\pi} \theta_1(e^{i\lambda})\overline{\theta_2(e^{i\lambda})} f(\lambda)d\lambda. \tag{2.11}$$

Соответствующее гильбертово пространство обозначим $\mathbb{L}^2_f$. Тогда $R(m) = 2\pi \left\langle e^{i\lambda n}, e^{i\lambda(m+n)} \right\rangle_f$ для любого n, то в этом случае $\theta_1(z) = z^n$ и $\theta_2(z) = z^{m+n}$.

Для любого $p \geq 1$ рассмотрим систему

$$R(m) + \sum_{k=1}^{p} a_{p,k} R(m-k) = 0, \quad m = 1, \ldots, p, \tag{2.12}$$

которая, согласно Лемме 2.1, имеет единственное решение $a_{p,1}, \ldots, a_{p,p}$, для которого выполняется

$$\sigma_p^2 \stackrel{\text{def}}{=} R(0) + \sum_{k=1}^{p} a_{p,k} R(k) > 0.$$

Введем полиномы комплексной переменной $Q_p(z)$ равенством

$$Q_p(z) = z^p + a_{p,1} z^{p-1} + \ldots + a_{p,p},$$

дополнительно положив $\sigma_0 = \sqrt{R(0)}$ и $Q_0(z) \equiv 1$.

Лемма 2.4. *Для полиномов Q_p выполнены равенства*

$$\langle Q_p, Q_q \rangle_f = 0, \quad p \neq q; \quad \|Q_p\|_f^2 = \sigma_p^2/2\pi.$$

Доказательство. Равенство $\|Q_0\|_f^2 = \sigma_0^2/2\pi$ выполнено по определению. Перепишем (2.12) в виде

$$0 = \int_{-\pi}^{\pi} \left(e^{im\lambda} + \sum_{k=1}^{p} a_{p,k} e^{i(m-k)\lambda} \right) f(\lambda)d\lambda = \int_{-\pi}^{\pi} e^{-i(p-m)\lambda} \left(e^{ip\lambda} + \sum_{k=1}^{p} a_{p,k} e^{i(p-k)\lambda} \right) f(\lambda)d\lambda =$$
$$= \int_{-\pi}^{\pi} e^{-i(p-m)\lambda}\, Q_p(e^{i\lambda})\, f(\lambda)d\lambda \quad \text{при} \quad 1 \leq m \leq p.$$

Отсюда следует, что при $p \geq 1$ полином Q_p ортогонален в $\mathbb{L}^2_f$ любому полиному степени, меньшей p, и, следовательно, ортогонален Q_q при $0 \leq q < p$.

Далее, при $p \geq 1$

$$2\pi\|Q_p\|_f^2 = \int_{-\pi}^{\pi} Q_p(e^{i\lambda})\overline{Q_p(e^{i\lambda})} f(\lambda)d\lambda = \int_{-\pi}^{\pi} Q_p(e^{i\lambda})\big(e^{-ip\lambda} + \sum_{k=1}^{p} a_{p,k} e^{-i(p-k)\lambda}\big) f(\lambda)d\lambda =$$
$$= \left\langle Q_p(z), z^p \right\rangle_f + \left\langle Q_p(z), M(z) \right\rangle_f,$$

где M — полином степени, меньшей p. Поэтому

$$2\pi\|Q_p\|_f^2 = \int_{-\pi}^{\pi} Q_p(e^{i\lambda}) e^{-ip\lambda} f(\lambda)d\lambda = \int_{-\pi}^{\pi} \left(1 + \sum_{k=1}^{p} a_{p,k} e^{-ik\lambda} \right) f(\lambda)d\lambda = R(0) + \sum_{k=1}^{p} a_{p,k} R(-k) = \sigma_p^2.$$

Лемма доказана. □

Полиномы Q_p называют *ортогональными полиномами Сегё*.

Замечание 2.4. Обозначим $\phi_p = Q_p/\|Q_p\|_f$. Тогда спектральная плотность авторегрессионного продолжения $\xi^{(p)}$ равна

$$f_\xi^{(p)}(\lambda) = \frac{1}{2\pi}\,\frac{\sigma_p^2}{\left|1+\sum_{k=1}^p a_k e^{-ik\lambda}\right|^2} = \frac{1}{2\pi}\,\frac{\sigma_p^2}{\left|1+\sum_{k=1}^p a_k e^{ik\lambda}\right|^2} = \frac{\|Q_p\|_f^2}{\left|e^{-ip\lambda}Q_p(e^{-i\lambda})\right|^2} =$$
$$= \frac{\|Q_p\|_f^2}{|Q_p(e^{-i\lambda})|^2} = \frac{\|Q_p\|_f^2}{\|Q_p\|_f^2\,|\phi_p(e^{-i\lambda})|^2} = \frac{1}{|\phi_p(e^{-i\lambda})|^2}\ .$$

Отметим некоторые свойства полиномов Q_p и ϕ_p (например, [13, гл. XI]).

Предложение 2.4. *1. Пусть ψ_p, $p = 0, 1, \ldots$ — некоторая ортонормированная в $\mathbb{L}_f^2$ система функций, причем ψ_p является полиномом степени p с вещественными коэффициентами. Тогда $\psi_p = \pm\phi_p$.*
2. Если полином[2] P степени $p > 0$ ортогонален в $\mathbb{L}_f^2$ любому полиному степени $0 \le j < p$, то $P(z) = cQ_p(z)$.
3. Если полином P степени $p > 0$ ортогонален в $\mathbb{L}_f^2$ любому моному z^j с $1 < j \le p$, то $P(z) = cz^pQ_p(1/z)$.
4. Полиномы Q_p при $|z| = 1$ удовлетворяют рекуррентным соотношениям

$$Q_{p+1}(z) = zQ_p(z) + a_{p+1,p+1}\,z^p\,Q_p(1/z). \tag{2.13}$$

Доказательство. 1. Очевидно, $\psi_0 = \pm\phi_0$, так это просто одна и та же (с точностью до знака) постоянная. Далее доказательство проводится индукцией по p. Пусть $\psi_p = \pm\phi_p$ при $p < p_0$. Рассмотрим полиномы ψ_{p_0} и ϕ_{p_0}. Для них существуют такие ненулевые вещественные α и β, что $\gamma_{p_0} \stackrel{\text{def}}{=} \alpha\psi_{p_0} - \beta\phi_{p_0}$ является полиномом степени, меньшей, чем p_0.

Далее, полином γ_{p_0} ортогонален любому полиному степени, меньшей p_0, так как и ψ_{p_0} и ϕ_{p_0} обладают этим свойством. Поэтому этот полином ортогонален самому себе, $\gamma_{p_0} \equiv 0$ и $\alpha\psi_{p_0} = \beta\phi_{p_0}$. Наконец, так как $\|\psi_{p_0}\|_f = \|\phi_{p_0}\|_f = 1$, то $|\alpha| = |\beta|$.
2. Это свойство становится очевидным, если рассмотреть разложение $P = \sum_{j=0}^p c_jQ_j$.
3. Для доказательства заметим, что для любых полиномов P_1, P_2 и любого $k \ge 0$ выполняется тождество $\langle P_1(z), P_2(z)\rangle_f = \left\langle z^kP_2(1/z), z^kP_1(1/z)\right\rangle_f$. Действительно,

$$\left\langle z^kP_2(1/z), z^kP_1(1/z)\right\rangle_f = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i\lambda k}P_2(e^{-i\lambda})\,\overline{e^{i\lambda k}P_1(e^{-i\lambda})}f(\lambda)d\lambda =$$
$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} P_1(e^{i\lambda})\overline{P_2(e^{i\lambda})}f(\lambda)d\lambda = \langle P_1(z), P_2(z)\rangle_f\,.$$

Теперь мы видим, что из равенств $\left\langle P(z), z^j\right\rangle_f = 0$ для $j = 1, \ldots, p-1$ следует, что $\left\langle z^pP(1/z), z^{p-j}\right\rangle_f = 0$. Так как $z^pP(1/z)$ — полином степени p, ортогональный z^k для $k = p - j \in \{0, \ldots, p-1\}$, то $z^pP(1/z) = cQ_p(z)$. Отсюда сразу же следует требуемое.
4. По определению, $Q_{p+1}(z)$ ортогонален z^j при $1 \le j \le p$. Кроме того,

$$\left\langle zQ_p(z), z^j\right\rangle_f = \left\langle Q_p(z), z^{j-1}\right\rangle_f = 0$$

при $1 \le j \le p$. Следовательно, полином $Q_{p+1}(z) - zQ_p(z)$ ортогонален z^j при всех $1 \le j \le p$. Применяя утверждение п. 3 видим, что полином $Q_{p+1}(z) - zQ_p(z)$ пропорционален $z^pQ_p(1/z)$. Приравнивая коэффициенты при z^0 и учитывая, что у полинома $zQ_p(z)$ этот коэффициент равен нулю, а у $z^pQ_p(1/z)$ — единице, получаем вид коэффициента пропорциональности: $c = a_{p+1,p+1}$. □

Равенства (2.13) называются *рекуррентными соотношениями Сегё*. Покажем, как эти соотношения можно использовать для рекуррентного по p нахождения коэффициентов авторегрессии $a_{p,k}$.

Приравнивая коэффициенты при одинаковых степенях в правой и левой частях (2.13) и учитывая, что

$$z^pQ_p(1/z) = 1 + \sum_{k=1}^p a_{p,k}z^k = 1 + \sum_{k=1}^p a_{p,p-k+1}z^{p-k+1},$$

[2]Все полиномы в этом разделе имеют вещественные коэффициенты.

получим, что при $1 \le k \le p$

$$a_{p+1,k} = a_{p,k} + a_{p+1,p+1}\, a_{p,p+1-k}, \tag{2.14}$$

в то время как члены с z^{p+1} сокращаются, а при $k = p+1$ мы приходим к тождеству $a_{p+1,p+1} = a_{p+1,p+1}$. Тем самым числа $a_{p+1,k}$ рекуррентно выражаются через $a_{p+1,p+1}$ и $a_{p,j}$, $j = 1, \ldots, p$. Теперь, подставляя (2.14) в уравнение

$$R(p+1) + \sum_{k=1}^{p+1} a_{p+1,k} R(p+1-k) = 0$$

и решая его относительно $a_{p+1,p+1}$, находим все коэффициенты $a_{p+1,k}$. А именно,

$$0 = R(p+1) + a_{p+1,p+1} R(0) + \sum_{k=1}^{p} a_{p+1,k} R(p+1-k) =$$
$$= R(p+1) + a_{p+1,p+1} R(0) + \sum_{k=1} a_{p,k} R(p+1-k) + a_{p+1,p+1} \sum_{k=1} a_{p,p+1-k} R(p+1-k),$$

откуда можно найти $a_{p+1,p+1}$, а потом из (2.14) — все остальные $a_{p+1,k}$.

Этот прием, позволяющий последовательно получать коэффициенты авторегрессии $a_{p,k}$ для увеличивающихся p, в статистике стационарных процессов носит название *алгоритма Левинсона*, см., например, [9, разд. 3.8.1].

2.5 Энтропийная характеризация авторегрессионного продолжения.

Авторегрессионное продолжение отрезка $R_\eta(0),\dots,R_\eta(p)$ ковариационной функции процесса η_n не является единственным неотрицательно определенным продолжением этого отрезка. Действительно, если взять $q>p$, то авторегрессионное продолжение отрезка $R_\eta(0),\dots,R_\eta(q)$ будет автоматически продолжением отрезка $R_\eta(0),\dots,R_\eta(p)$.

Приведем пример других продолжений, следуя [2], где используется несколько более аналитический язык.

Поскольку матрица $\Sigma_{p+1}=\left\{R_\eta(l-k)\right\}_{k,l=0}^p$ положительно определена, то при достаточно малом $\mu>0$ матрица $\Sigma_{p+1}^{(\mu)}=\Sigma_{p+1}-\mu I$ тоже будет положительно определена (здесь I — единичная $(p+1)\times(p+1)$ матрица). Следовательно, вместо (2.6) и (2.7) мы получим

$$R_\eta^{(\mu)}(m)+\sum_{k=1}^p a_k^{(\mu)}R_\eta^{(\mu)}(m-k)=0,\quad m=1,\dots,p$$

и

$$\sigma_p^2(\mu)\stackrel{\text{def}}{=}R_\eta^{(\mu)}(0)+\sum_{k=1}^p a_k^{(\mu)}R_\eta^{(\mu)}(k)>0,$$

где $R_\eta^{(\mu)}(k)=R_\eta(k)$ при $1\le k\le p$ и $R_\eta^{(\mu)}(0)=R_\eta(0)-\mu>0$.

Зададим последовательность $\xi_n^{(\mu)}$ следующим образом: случайные величины $\xi_0^{(\mu)},\dots,\xi_p^{(\mu)}$ имеют среднее ноль и ковариационную матрицу $\Sigma_{p+1}^{(\mu)}$, а при $n>p$ задаются равенством

$$\xi_n^{(\mu)}=-\sum_{k=1}^p a_k^{(\mu)}\xi_{n-k}^{(\mu)}+\sigma_p(\mu)\,\varepsilon_n,$$

где (центрированный и нормированный) белый шум ε_n ортогонален $\xi_j^{(\mu)}$ при $0\le j<n$.

Повторяя рассуждения Предложения 2.2, мы получим, что $\xi_n^{(\mu)}$ является авторегрессионным стационарным процессом, а его ковариационная функция $R_\xi^{(\mu)}(n)$ является продолжением отрезка $R_\eta^{(\mu)}(0),\dots,R_\eta^{(\mu)}(p)$. Иначе говоря, $R_\xi^{(\mu)}(0)=R_\eta(0)-\mu$, а $R_\xi^{(\mu)}(k)=R_\eta(k)$ при $1\le k\le p$.

Если теперь прибавить к $\xi_n^{(\mu)}$ независимый с этими случайными величинами белый шум $\varepsilon_n^{(\mu)}$ с $\mathbb{E}\varepsilon_n^{(\mu)}=0$ и $\mathbb{D}\varepsilon_n^{(\mu)}=\mu$, то мы получим продолжение отрезка $R_\eta(0),\dots,R_\eta(p)$.

Замечание 2.5. Спектральная плотность процесса $\zeta_n^{(\mu)}=\xi_n^{(\mu)}+\varepsilon_n^{(\mu)}$ имеет вид

$$f_\zeta^{(\mu)}(\lambda)=\mu+\frac{1}{2\pi}\,\frac{\sigma_p^2(\mu)}{\left|1+\sum_{k=1}^p a_k^{(\mu)}e^{-ik\lambda}\right|^2}\ ,\quad |\lambda|\le\pi. \tag{2.15}$$

Для описания места авторегрессионного продолжения отрезка $R_\eta^{(\mu)}(0),\dots,R_\eta^{(\mu)}(p)$ среди всех продолжений нам понадобится процитировать два классических, но не простых для доказательства утверждений.

Теорема 2.2 (Колмогоров, см. [3], стр. 622)**.** *Пусть η_n, $n\in\mathbb{Z}$ — стационарный в широком смысле процесс со средним ноль. Для того, чтобы существовал белый шум ε_n такой, что*

$$\eta_n=\sum_{k=-\infty}^n c_{n+k}\varepsilon_k, \tag{2.16}$$

необходимо и достаточно, чтобы существовала спектральная плотность $f_\eta(\lambda)$ процесса η_n такая что

$$S(f)=S_\eta(f)\stackrel{\text{def}}{=}\int_{-\pi}^{\pi}\ln f_\eta(\lambda)d\lambda>-\infty. \tag{2.17}$$

В англоязычной литературе величина $S(f)$ иногда называется *энтропийным интегралом* (entropy integral).

Процесс, удовлетворяющий (2.16), называется *регулярным* относительно белого шума ε_n. Отметим, что реализуемые процессы авторегрессии, как и стационарные процессы со спектральной плотностью (2.15), удовлетворяют неравенству (2.17).

Теорема 2.3 (Колмогоров-Сегё, см. [4], стр. 257)**.** *Пусть η^*_{n+1} — наилучший (в среднеквадратическом) линейный прогноз случайной величины η_{n+1}, построенный по случайным величинам η_k при $-\infty < k \le n$. Если η_n удовлетворяет* (2.16)*, то*

$$\mathbb{E}\big(\eta^*_{n+1} - \eta_{n+1}\big)^2 = 2\pi e^{S_\eta(f)/2\pi}. \tag{2.18}$$

Пусть дана ковариационная функция $R_\eta(n)$. Зафиксировав отрезок $R_\eta(0), \ldots, R_\eta(p)$ этой ковариационной функции, рассмотрим те ее неотрицательно определенные продолжения, для которых выполнено неравенство (2.17). Значит, для них будут выполнены аналоги (2.18).

Начиная с работ Берга [5], [6] (в интернете доступна диссертация Берга [7], в которой изложены результаты этих статей) для характеризации авторегрессионного продолжения используется принцип "максимальной случайности" (поскольку в качестве меры случайности обычно используется энтропия, то этот принцип обычно называют *Принципом Максимальной Энтропии*).

Немного изменив для наших целей фразу из [8] (цитируется по [9, гл. 7.4]), принцип максимальной случайности можно сформулировать таким образом:

если мы делаем выводы на основе неполной информации, то должны опираться[3] *на такую вероятностную модель, которая приводит к максимальной случайности, допускаемой нашей информацией.*

В нашем случае мерой "случайности" процесса можно считать ошибку наилучшего прогноза на один шаг при условии, что все предыдущая траектория процесса известна (чем более "случаен" процесс, тем труднее его прогнозировать).

Если мы приняли принцип "максимальной случайности" в данной постановке для нашей редакции, то пришли с учетом (2.18) к следующей задаче: *найти для фиксированного p такое неотрицательно определенное продолжение отрезка $R_\eta(0), \ldots, R_\eta(p)$, для которого энтропийный интеграл* (2.17) *максимален.*

Оказывается, что решением такой оптимизационной задачи и является авторегрессионное продолжение. Впервые это заметил Берг, хотя в его доказательстве и содержатся некоторые пробелы. Список ссылок на другие подходы и доказательства этого факта можно найти в [10].

С собственно энтропией задачу максимизации функционала (2.17) связывают следующие соображения (эти соображения и приводят к термину "метод максимальной энтропии").

Пусть случайный вектор $\bar\xi = (\xi_1, \ldots, \xi_n)^{\mathrm{T}}$ имеет распределение $\mathcal{P}$ с плотностью $p(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^n$. Тогда (дифференциальной) *энтропией* n-мерного распределения $\mathcal{P}$ называется функционал

$$H_\xi = H(\mathcal{P}) = -\int_{\mathbb{R}^n} \ln\big(p(\mathbf{x})\big) p(\mathbf{x}) d\mathbf{x},$$

где $0\ln(0) = 0$.

Не останавливаясь на причинах, по которым энтропия может служить подходящей мерой случайности (кстати, это считается общепринятым, см. [8]), приведем некоторые ее свойства, необходимые для изложения.

Ясно, что если случайные величины ξ_k независимы с энтропиями H_k, то

$$H_\xi = \sum_{k=1}^n H_k.$$

Непосредственно доказывается следующий простой факт. Если случайный вектор $\bar\xi = (\xi_1, \ldots, \xi_n)^{\mathrm{T}}$ имеет плотность распределения $p(x_1, \ldots, x_n)$ и энтропию H_ξ и если $A : \mathbb{R}^L \mapsto \mathbb{R}^L$ — невырожденная матрица, то для случайного вектора $\bar\eta = A\bar\xi$ энтропия имеет вид

$$H_\eta = H_\xi + \ln(|\det A|). \tag{2.19}$$

Отсюда сразу следует вид энтропии многомерного гауссовского распределения с нулевым средним и невырожденной ковариационной матрицей. Действительно, для одномерной гауссовской случайной величины ζ со средним ноль и положительной дисперсией σ^2 энтропия легко считается и имеет вид

$$H_\zeta = \frac{1}{2}\,\ln(2\pi e) + \frac{1}{2}\,\ln(\sigma^2).$$

[3] далее в [8]: "...на такое вероятностное распределение, которое имеет максимальную энтропию, допускаемую нашей априорной информацией."

Поэтому, если $\xi_1, \ldots, \xi_n$ независимы и $\xi_k \in \mathrm{N}(0, \sigma_k^2)$, то их совместная энтропия равна

$$H_\xi = \frac{n}{2}\,\ln(2\pi e) + \frac{1}{2}\,\ln\left(\prod_{k=1}^{n} \sigma_k^2\right).$$

Отсюда стандартным приведением ковариационной матрицы к диагональному виду с использованием (2.19) сразу получается, что для гауссовского вектора $\bar{\eta}$ с нулевым средним и невырожденной ковариационной матрицей Σ

$$H_\eta = \frac{n}{2}\,\ln(2\pi e) + \frac{1}{2}\,\ln\left(\det\Sigma\right).$$

Рассмотрим теперь бесконечную последовательность $\eta = (\eta_0, \ldots, \eta_n, \ldots)$ такую, что для любого n вектор $(\eta_0, \ldots, \eta_n)$ имеет плотность распределения $p_n(\mathbf{x})$ и энтропию H_n. Если существует предел h последовательности $h_n = H_n/n$ при $n \to \infty$, то число h называется *удельной энтропией* последовательности η.

Пусть η — гауссовская стационарная последовательность со средним ноль и ковариационной функцией $R_\eta(m)$. Обозначим Σ_n ковариационную матрицу случайного вектора $(\eta_0, \ldots, \eta_{n-1})$ и потребуем, чтобы при любом m эта матрица была невырожденной. Тогда автоматически вектор $(\eta_0, \ldots, \eta_{n-1})$ будет иметь плотность распределения и, следовательно, энтропию

$$H_n = \frac{n}{2}\,\ln(2\pi e) + \frac{1}{2}\,\ln\left(\det\Sigma_n\right). \tag{2.20}$$

Оказывается (при необременительных ограничениях), в этом случае удельная энтропия существует и имеет достаточно простой вид, связанный с энтропийным интегралом (2.17). А именно, имеет место следующее утверждение.

Теорема 2.4. *Пусть гауссовская стационарная последовательность η имеет ограниченную сверху и отделенную от нуля спектральную плотность f. Тогда удельная энтропия этой последовательности h существует и*

$$h = \frac{1}{2}\,\ln(2\pi e) + \frac{1}{4\pi}\int_{-\pi}^{\pi} \ln\big(2\pi f(\lambda)\big)d\lambda = 1/2 + \ln(2\pi) + \frac{1}{4\pi}\,S_f, \tag{2.21}$$

где S_f определено в (2.17).

Доказательство. Ограничимся лишь изложением хода доказательства теоремы 2.4. Оно основано на классическом результате Г. Сегё о предельном поведении спектра эрмитовых теплицевых матриц. Приведем это утверждение в удобной для нас форме.

Пусть $R(n)$ — ковариационная функция, имеющая спектральную плотность f. Обозначим $\Sigma_n = \big\{R(k-l)\big\}_{k,l=0}^{n-1}$ и $\tau_{n,0}, \ldots, \tau_{n,n-1}$ — собственные числа матрицы Σ_n. Наконец, введем распределение собственных чисел матрицы Σ_n равенством

$$\mathcal{P}_n = \begin{pmatrix} \tau_{n,0} & \ldots & \tau_{n,n-1} \\ 1/n & \ldots & 1/n \end{pmatrix}.$$

Теорема 2.5 (Сегё, [11], гл. 7)**.** *Если спектральная плотность f ограничена, то распределение $\mathcal{P}_n$ слабо сходится при $n \to \infty$ к распределению $\mathcal{P}$ случайной величины $2\pi f(\alpha)$, где α равномерно распределена на $[-\pi, \pi]$.*

При дополнительном (и упрощающем) предположении $\sum_{m=-\infty}^{\infty} |R(m)| < \infty$ доказательство этого факта изложено в [12, теор. 9 разд. 4.4]. Следует только отметить, что в подобной литературе, не имеющей непосредственного отношения к случайным процессами, ковариации представлены в виде

$$R(m) = \frac{1}{2\pi}\int_0^{2\pi} e^{-im\omega} g(\omega)d\omega,$$

где функция g связана со спектральной плотностью f равенством

$$g(\omega) = \begin{cases} 2\pi f(\omega) & \text{при} \quad 0 \le \omega < \pi, \\ 2\pi f(\omega - 2\pi) & \text{при} \quad \pi \le \omega < 2\pi. \end{cases}$$

Дальнейший ход рассуждений таков. Прежде всего, имеет место следующее утверждение.

Лемма 2.5 ([12], лемма 6 разд. 4.2)**.** *Обозначим* $m_f = \operatorname{essinf} f$ *и* $M_f = \operatorname{esssup} f$*. Тогда для любых* n *и* k

$$2\pi m_f \le \tau_{n,k} \le 2\pi M_f.$$

Тем самым и распределения $\mathcal{P}_n$ и распределение случайной величины $2\pi f(\alpha)$ сосредоточены на одном и том же компакте $[2\pi m_f, 2\pi M_f]$.

Предположим теперь, что спектральная плотность не только ограничена, но и отделена от нуля. Тогда функция $F(x) = \ln x$ непрерывна и ограничена на отрезке $[2\pi m_f, 2\pi M_f]$ и, следовательно, по теореме Сегё

$$\int F d\mathcal{P}_n = \frac{1}{n}\sum_{k=0}^{n-1} \ln(\tau_{n,k}) = \frac{1}{n}\ln\left(\prod_{k=0}^{n-1}\tau_{n,k}\right) = \frac{\ln(\det\Sigma_n)}{n} \to \int F d\mathcal{P} = \frac{1}{2\pi}\int_{-\pi}^{\pi} \ln\big(2\pi f(\lambda)\big) d\lambda.$$

Учитывая (2.17) и (2.20), получаем (2.21). □

Таким образом, задачи максимизации дисперсии наилучшего прогноза (2.18) и удельной энтропии $\lim H_n/n$ для гауссовского стационарного процесса с ограниченной и отделенной от нуля спектральной плотностью оказываются эквивалентными.

Как уже говорилось, решением этих задач при условии, что первые p автоковариаций процесса известны, является авторегрессионное продолжение ковариационной функции процесса. Кстати, технически простое и основанное на теоретико-информационных соображениях доказательство этого факта можно найти в [10].

Список литературы

[1] Т. Андерсон (1976), Статистический анализ временных рядов, М.. Мир.

[2] V.F. Pisarenko (1973), The Retrieval of Harmonics from a Covariance Function, Geophys. J. R. astr. Soc., 33, p. 347 – 366.

[3] А.Н. Ширяев (2004), Вероятность-2, 3-е изд., М., МЦНМО.

[4] А.В. Булинский, А.Н. Ширяев (2003), Теория случайных процессов, М., Физматлит.

[5] J.P. Burg (1967), Maximum entropy spectral analysis, paper presented at the *37-th Annual International Meeting, Soc. of Explor. Geophys.*, Oklahoma City, Okla., Oct. 31, 1967.

[6] J.P. Burg (1968), A new analysis technique for time series data, paper presented at *Advanced Study Institute on Signal Processing*, NATO, Enschede, Netherlands, 1968. (Reprinted in: Childers D. G., ed., Modern spectral analysis. New York: IEEE Press. 1978. p. 42 – 48).

[7] J.P. Burg (1975), Maximum Entropy Spectral Analysis., Ph.D. dissertation, Stanford University, Stanford, CA. May 1975.

[8] Э. Т. Джейнс (1982), О логическом обосновании методов максимальной энтропии, ТИИИЭР, т. 70, N° 9, с. 33 – 51.

[9] С.Л. Марпл-мл. (1990), Цифровой спектральный анализ и его приложения, М.. Мир.

[10] B.S. Choi, Th.M.Cover (1987), A proof of Burg's Theorem, in: C.R. Smith and G.J. Ericson (eds.), Maximum-Entropy and Bayesian Spectral Analysis and Estimation Peoblems, D. Reydel Publishing Company, p. 75 – 84.

[11] У. Гренандер, Г. Сегё (1961), Теплицевы формы и их приложения, М., Изд. иностр. лит.

[12] R. M. Gray (2006), Toeplitz and Circulant Matrices: A Review, Communications and Information Theory Vol. 2, No 3, p. 155 -– 239.

[13] G. Szegő (1975). *Orthogonal polynomials*, 4nd ed., AMS, Providence, Rhode Island.

[14] П. Биллингсли (1977), Сходимость вероятностных мер, М., Наука.