Статистика стационарных случайных последовательностей

В.В. Некруткин

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

Материалы к специальному курсу, 2014 г.

Содержание

1 Вещественные линейные стационарные последовательности

Рассмотрим вещественную стационарную в широком смысле последовательность $\{x_n\}_{n\geq 1}$. По определению $\mathbb{E}x_n^2 < \infty$. Стандартными характеристиками этой последовательности являются

- среднее значение $\mathbb{E}x_n = a$;
- ковариационная функция $R(m) = \mathbb{E}x_n x_{m+n} - a^2$, причем $R(-m) = R(m)$, а $R(0) = \mathbb{D}x_n^2$;
- и спектральная мера $m(d\lambda)$, сосредоточенная на $(-\pi, \pi]$.

Спектральную меру можно описывать ее спектральной функцией $F(\lambda) = m((-\pi, \lambda])$. Если спектральная мера абсолютно непрерывна относительно меры Лебега, то ее обычно задают спектральной плотностью $f(\lambda)$. В случае вещественной стационарной последовательности спектральную плотность можно считать четной функцией.

Соотношение между ковариационной функцией и спектральной мерой хорошо известно:

$$R(m) = \int_{(-\pi,\pi]} e^{im\lambda} m(d\lambda),$$

причем в случае существования спектральной плотности

$$R(m) = \int_{-\pi}^{\pi} e^{im\lambda} f(\lambda) d\lambda = \int_{-\pi}^{\pi} \cos(\lambda) f(\lambda) d\lambda = 2\int_{0}^{\pi} \cos(\lambda) f(\lambda) d\lambda,$$

то есть $R(m) \to 0$ при $m \to \infty$ по теореме Римана-Лебега.

Приведем один из вариантов обратного соотношения.

Предложение 1.1. *Если* $\sum_m |R(m)| < \infty$, *то*

$$f(\lambda) = \frac{1}{2\pi} \sum_m e^{-im\lambda} R(m) = \frac{1}{2\pi} \sum_m e^{im\lambda} R(m) = \frac{1}{2\pi} \left(R(0) + 2\sum_{m=1}^{\infty} \cos(m\lambda) R(m) \right), \tag{1.1}$$

причем f *непрерывна и ограничена.*

Доказательство. Обозначим

$$g(\lambda) = \frac{1}{2\pi} \sum_m e^{-im\lambda} R(m).$$

Ряд в правой части последнего равенства сходится, причем

$$\begin{aligned} \left|g(\lambda_1) - g(\lambda_2)\right| &\leq \frac{1}{2\pi} \sum_m \left|e^{-im\lambda_1} - e^{-im\lambda_2}\right| |R(m)| \leq \\ &\leq \frac{1}{2\pi} \sum_{|m|\leq m_0} \left|1 - e^{-im(\lambda_2-\lambda_1)}\right| |R(m)| + \frac{1}{\pi} \sum_{m>m_0} |R(m)|. \end{aligned}$$

Следовательно, функция g непрерывна. Далее,

$$\int_{-\pi}^{\pi} e^{ik\lambda} g(\lambda) d\lambda = \sum_m R(m) \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i(m+k)\lambda} d\lambda = R(-k) = R(k),$$

что и заканчивает доказательство. □

Следствие 1.1. 1. Из равенства (1.1) сразу же следует, что

$$R(0) + 2\sum_{m=1}^{\infty} R(m) = 2\pi f(0). \tag{1.2}$$

2. Рассуждениями, аналогичными приведенным в Предложении 1.1, можно доказать, что для любого целого $k \ge 1$ из того, что $\sum_m |m|^k \, |R(m)| < \infty$, то следует, что f k раз непрерывно дифференцируема. Таким образом, гладкость спектральной плотности оказывается тесно связанной со скоростью убывания ковариационной функции.

В теории комплекснозначных слабо стационарных последовательностей хорошо известен следующий факт. Если x_m, $m = 0, \pm 1, \pm 2, \ldots$ — слабо стационарная последовательность с нулевым средним и спектральной плотностью f, то существует такой стандартный белый шум ε_m, $m = 0, \pm 1, \pm 2, \ldots$, что

$$x_m = \sum_{j=-\infty}^{\infty} c_{m+j}\, \varepsilon_j = \sum_{\ell=-\infty}^{\infty} c_\ell\, \varepsilon_{\ell-m} \tag{1.3}$$

и $\sum_j |c_j|^2 < \infty$.

Ковариационная функция процесса (1.3) имеет вид

$$R(n) = \mathbb{E} x_0 x_n = \mathbb{E}\left(\sum_j c_j \varepsilon_j\right)\left(\sum_l c_{l+n}\, \varepsilon_l\right) = \sum_{j,l} \mathbb{E} c_j c_{l+n}\, \varepsilon_j \varepsilon_l = \sum_j c_j c_{j+n}. \tag{1.4}$$

Спектральная плотность процесса (1.3) равна

$$f(\lambda) = \frac{1}{2\pi}\left|\sum_j c_j e^{ij\lambda}\right|^2. \tag{1.5}$$

В частности,

$$f(0) = \frac{1}{2\pi}\Big|\sum_j c_j\Big|^2.$$

Замечание 1.1. Меняя порядок нумерации коэффициентов c_j (и компонент белого шума ε_j) можно получить формально другую форму представления линейного процесса. Например, полагая $s = -j$ и $\varepsilon'_n = \varepsilon_{-n}$, мы вместо (1.3) придем к

$$x_m = \sum_{s=-\infty}^{\infty} c_{m-s}\, \varepsilon'_s.$$

Соответственно изменятся и формулы (1.4), (1.5). В зависимости от решаемой задачи мы можем применять разные формы записи одного и того же процесса (1.3).

Мы будем дополнительно требовать, чтобы случайные величины ε_n были вещественны, независимы и одинаково распределены. Тогда последовательность, задаваемая (1.3), будет называться *линейной*. Конечно, линейная последовательность является сильно стационарной. Кроме того, нам удобно будет считать, что и коэффициенты c_j тоже вещественны.[1]

[1] Поскольку здесь происходит переход от стационарных в широком смысле процессов к сильно стационарным процессам, то сходимости в среднеквадратическом ряда (1.3) оказывается недостаточно. Из [5, гл. IV §2 теор. 1] следует, что при описанных условиях ряд (1.3) будет сходиться почти всюду.

Легко написать достаточные условия гладкости спектральной плотности для линейного процесса. Действительно, если $\sum_j |c_j| < \infty$, то

$$\sum_n |R(n)| = \sum_n \Big| \sum_j c_j c_{j+n} \Big| \leq \sum_j |c_j| \sum_n |c_{j+n}| = \left(\sum_n |c_n| \right)^2 < \infty.$$

Следовательно, в этом случае спектральная плотность (1.5) будет непрерывной.

Более того, имеет место следующее утверждение.

Лемма 1.1. *Пусть* $\beta \geq 0$. $\sum_j |j|^\beta |c_j| < \infty$, *то* $\sum_j |j|^\beta |R(j)| < \infty$.

Доказательство. Действительно, так как $|k+j|^\beta \leq 2^\beta \max\big(|k|^\beta, |j|^\beta\big) \leq 2^\beta \big(|k|^\beta + |j|^\beta\big)$, то

$$\sum_j |j|^\beta |R(j)| \leq \sum_j |j|^\beta \sum_k |c_k|\,|c_{k+j}| = \sum_k |c_k| \sum_j |j|^\beta |c_{k+j}| = \sum_k |c_k| \sum_m |m-k|^\beta |c_m| \leq$$
$$\leq 2^\beta \sum_k |c_k| \sum_m \big(|m|^\beta + |k|^\beta\big) |c_m| = 2^{\beta+1} \left(\sum_k |k|^\beta |c_k| \right) \left(\sum_k |c_k| \right) < \infty.$$

Утверждение доказано. □

2 Оценивание среднего

2.1 Асимптотическая дисперсия выборочного среднего стационарной последовательности

Стандартная форма ЦПТ для некоторой последовательности $\{x_n\}_{n\ge1}$ с одинаковыми средними, равными a, и конечными вторыми моментами может быть записана как

$$\mathcal{L}\left(\frac{S_n - na}{B_n}\right) \Longrightarrow \mathrm{N}(0,1), \tag{2.1}$$

где $S_n = x_1 + \ldots + x_n$, $B_n^2 = \mathbb{D}S_n$ и $n \to \infty$.

Если при этом $B_n^2 \sim n\sigma^2$, где $\sigma^2 = \mathrm{const} > 0$, то (2.1) можно переписать в виде

$$\mathcal{L}\left(\frac{S_n - na}{\sigma\sqrt{n}}\right) \Longrightarrow \mathrm{N}(0,1),$$

или, вводя обозначение $\overline{x}_n = S_n/n$, в виде

$$\mathcal{L}\left(\sqrt{n}\ \frac{\overline{x}_n - a}{\sigma}\right) \Longrightarrow \mathrm{N}(0,1),$$

что эквивалентно

$$\mathcal{L}\big(\sqrt{n}\ (\overline{x}_n - a)\big) \Longrightarrow \mathrm{N}(0,\sigma^2). \tag{2.2}$$

Поскольку $\mathbb{D}\overline{x}_n = B_n^2/n \to \sigma^2$, то число σ^2 можно называть *коэффициентом асимптотической дисперсии* выборочного среднего. Для краткости иногда в этом случае употребляют термин *асимптотическая дисперсия.*

Заметим, что (2.2) следует из (2.1) и в том вырожденном случае, когда $B_n^2/n \to \sigma^2 = 0$, если понимать под $\mathrm{N}(b,0)$ распределение, сосредоточенное в точке b.

Вернемся теперь к стационарной последовательности $\{x_n\}_{n\ge1}$. Если у нее спектральная мера m не имеет нагрузки в нуле (например, если у последовательности существует спектральная плотность), то по ЗБЧ в $L^2(d\mathbb{P})$

$$\overline{x}_n - a \to \mu(\{0\}) = 0$$

в $L^2(d\mathbb{P})$ и, следовательно, $\overline{x}_n$ стремится к a по вероятности. Таким образом, при этом условии $\overline{x}_n$ является не только несмещенной, но и состоятельной оценкой среднего a.

Найдем теперь выражение для дисперсии $\overline{x}_n$. Как обычно, для этого мы можем рассматривать только случай $a = 0$.

Предложение 2.1. *Дисперсия $\mathbb{D}\overline{x}_n$ может быть представлена в виде*

$$\mathbb{D}\bar{x}_n = \frac{1}{n}\sum_{j=-(n-1)}^{n-1} \big(1 - |j|/n\big) R(j), \tag{2.3}$$

а также в виде

$$\mathbb{D}\bar{x}_n = \int_{(-\pi,\pi]} \left(\frac{\sin(\lambda n/2)}{n\sin(\lambda/2)}\right)^2 m(d\lambda).$$

Доказательство. Действительно,

$$\mathbb{D}\bar{x}_n = \frac{1}{n^2}\,\mathbb{E}(x_1+\ldots+x_n)^2 = \frac{1}{n^2}\sum_{k,l=1}^{n}\mathbb{E}x_k x_l = \frac{1}{n^2}\sum_{k,l=1}^{n}R(k-l) = \frac{1}{n}\sum_{j=-(n-1)}^{n-1}\big(1-|j|/n\big)R(j).$$

Кроме того,

$$\mathbb{D}\bar{x}_n = \frac{1}{n^2}\int_{(-\pi,\pi]}\sum_{k,l=1}^{n}e^{i(k-l)\lambda}m(d\lambda) = \int_{(-\pi,\pi]}\left|\frac{1}{n}\sum_{k=1}^{n}e^{ik\lambda}\right|^2 m(d\lambda) = \int_{(-\pi,\pi]}\left(\frac{\sin(\lambda n/2)}{n\sin(\lambda/2)}\right)^2 m(d\lambda),$$

что и заканчивает доказательство. □

Перейдем теперь к предельное поведение дисперсии выборочного среднего. Сначала докажем следующую элементарную лемму.

Лемма 2.1. *Если ряд $\sum_{k\geq 1}a_k$ сходится абсолютно, то*

$$\lim_{n\to\infty}\sum_{n=1}^{n-1}(1-k/n)a_k = \sum_{k\geq 1}a_k.$$

Доказательство. Найдем n_0 такое, что $\sum_{k\geq n_0}|a_k|<\varepsilon$. Тогда при $n>n_0$

$$\left|\sum_{k\geq 1}a_k - \sum_{k=0}^{n-1}(1-k/n)a_k\right| \leq \sum_{k\geq n}|a_k| + \frac{1}{n}\sum_{k=1}^{n_0-1}k|a_k| + \sum_{k=n_0}^{n-1}(k/n)\,|a_k| \leq$$
$$\leq \sum_{k\geq n}|a_k| + \frac{1}{n}\sum_{k=1}^{n_0-1}k|a_k| + \sum_{k\geq n_0}|a_k|.$$

Первое слагаемое стремится к нулю при $n\to\infty$, последнее по условию меньше ε, а среднее стремится к нулю при $n\to\infty$ и фиксированном ε. Утверждение доказано. □

Теперь не представляет труда найти коэффициент асимптотической дисперсии для выборочного среднего.

Предложение 2.2. *Если $\sum_m|R(m)|<\infty$, то при $n\to\infty$*

$$n\mathbb{D}\,\overline{x}_n \to R(0) + 2\sum_{m\geq 1}R(m) = 2\pi f(0). \tag{2.4}$$

Доказательство. Утверждение вытекает из равенства (2.3), Леммы 2.1 и Следствия 1.1. □

Замечание 2.1. Таким образом, если $f(0)\neq 0$, то $\mathbb{D}\overline{x}_n \sim 2\pi f(0)/n$ и имеет порядок $O(1/n)$ при $n\to\infty$. Если же $f(0)=0$, то дисперсия $\overline{x}_n$ стремится к нулю быстрее, чем n^{-1}.

2.2 ЦПТ для линейной последовательности

2.2.1 Лемма о слабой сходимости

Лемма 2.2. *Пусть $\xi_n = \eta_{kn} + \beta_{kn}$ при $n \geq n_0(k)$, где η_{kn} и β_{kn} принимают значения в $\mathbb{R}$. Предположим, что выполнены следующие условия.*

1. *Последовательность β_{kn} равномерно по n сходится по вероятности к нулю при $k \to \infty$ в том смысле, что $\sup_n \mathbb{P}(|\beta_{kn}| > \varepsilon) \to 0$ при $k \to \infty$.*
2. *Для любого k $\mathcal{L}(\eta_{kn}) \Rightarrow \mathcal{P}_k$ при $n \to \infty$.*
3. *$\mathcal{P}_k \Rightarrow \mathcal{P}$ при $k \to \infty$.*

Тогда $\mathcal{L}(\xi_n) \Rightarrow \mathcal{P}$ при $n \to \infty$.

Доказательство. Возьмем $f \in \mathbf{C}_{BL}$, $0 \leq f \leq M$. Тогда

$$\begin{gathered}\mathbb{E}f(\xi_n) = \mathbb{E}f(\eta_{kn} + \beta_{kn}) = \mathbb{E}\big(f(\eta_{kn} + \beta_{kn}),\, |\beta_{kn}| > \varepsilon\big) + \mathbb{E}\big(f(\eta_{kn} + \beta_{kn}), |\beta_{kn}| \leq \varepsilon\big) = \\ = \mathbb{E}\big(f(\eta_{kn} + \beta_{kn}), |\beta_{kn}| > \varepsilon\big) + \mathbb{E}\big(f(\eta_{kn} + \beta_{kn}) - f(\eta_{kn}), |\beta_{kn}| \leq \varepsilon\big) + \mathbb{E}\big(f(\eta_{kn}), |\beta_{kn}| \leq \varepsilon\big) = \\ = J_1(k,n) + J_2(k,n) + J_3(k,n).\end{gathered}$$

Далее для любого $\varepsilon > 0$

$$\sup_n J_1(k,n) = \sup_n \mathbb{E}(f(\eta_{kn} + \beta_{kn}), |\beta_{kn}| > \varepsilon) \leq M \sup_n \mathbb{P}(|\beta_{kn}| > \varepsilon) \to 0$$

при $k \to 0$. Значит, выбором k можно сделать J_1 сколь угодно малым равномерно по n. Далее,

$$|J_2(k,n)| \leq L\mathbb{E}\big(|\beta_{kn}|,\, |\beta_{kn}| \leq \varepsilon\big) \leq L\varepsilon.$$

То есть выбором ε слагаемое $J_2(k,n)$ можно сделать сколь угодно малым равномерно по k, n. Затем,

$$J_3(k,n) = \mathbb{E}(f(\eta_{kn}), |\beta_{kn}| \leq \varepsilon) = \mathbb{E}f(\eta_{kn}) - \mathbb{E}(f(\eta_{kn}), |\beta_{kn}| > \varepsilon) = I_1(k,n) + I_2(k,n).$$

Слагаемое $I_2(k,n)$ будет равномерно по n мало при большом k. Что касается $I_1(k,n)$, то

$$I_1(k,n) - \int f d\mathcal{P} = \left(\mathbb{E}f(\eta_{kn}) - \int f d\mathcal{P}_k\right) + \left(\int f d\mathcal{P}_k - \int f d\mathcal{P}\right) = I_3(k,n) + I_4(k).$$

Выбором k делаем I_4 малым, а затем при этом фиксированном k выбираем такое n, что мало I_3. Тем самым утверждение доказано. □

2.2.2 Конечно-зависимые случайные величины

Теоремы Леви (одномерная и многомерная) говорят о слабой сходимости к нормальному распределению суммы (центрированной и нормированной) независимых одинаково распределенных случайных величин. Теорема Линдеберга осуществляет отход от одинаковой распределенности, но сохраняет условие независимости. Что делать, если случайные величины зависимы?

Простым отклонением от независимости является конечно-зависимость.

Определение 2.1. Случайные величины $x_1, \ldots, x_n, \ldots$ называются *конечно-зависимыми*, если существует такое $N \geq 0$, что для любого $l \geq 1$, любых $k_1, \ldots, k_l$ и и любых $j_1, \ldots, j_l$ таких, что $\min j_i - \max k_i > N$ случайные вектора $(x_{k_1}, \ldots, x_{k_l})$ и $(x_{j_1}, \ldots, x_{j_l})$ независимы.

Если число M — минимальное из всех таких N, то величины называются M-зависимыми. Допуская некоторую вольность, мы будем также говорить о $(\leq M)$-зависимых случайных величинах.

Приведем несколько примеров конечно-зависимых случайных величин.

Пример 2.1.

1. Обычные независимые случайные величины — ноль-зависимые.
2. Пусть $y_1, \ldots, y_n, \ldots$ — независимые случайные величины. Обозначим $x_i = y_i y_{j+1}$. Тогда x_i являются (≤ 1)-зависимыми.
3. Положим в условиях предыдущего примера

$$x_k = \sum_{\ell=1}^{M} c_\ell \, y_{k+\ell}.$$

Такие случайные величины конечно-зависимы. Если при этом $\mathbb{E}y_i = \mathrm{const}_1$ и $\mathbb{D}y_i = \mathrm{const}_2 < \infty$, то x_i — стационарная в широком смысле линейная последовательность. Если y_i одинаково распределены, то она является стационарной в узком смысле.

Прежде чем доказывать предельную теорему, остановимся на следующем простом факте.

Лемма 2.3. *Пусть случайные величины $\xi_1, \ldots, \xi_L$ имеют одинаковые дисперсии, равные σ^2. Тогда $\mathbb{D}(\xi_1 + \ldots + \xi_L) \leq L^2\sigma^2$.*

Доказательство. Действительно,

$$\mathbb{D}(\xi_1 + \ldots + \xi_L) = \sum_{i=1}^{L} \mathbb{D}(\xi_i) + 2 \sum_{1 \leq i < j \leq L} \mathrm{Cov}\,(\xi_i, \xi_j) =$$
$$= L\sigma^2 + 2\sigma^2 \sum_{1 \leq i < j \leq L} \rho\,(x_i, x_j) \leq L\sigma^2 + L(L-1)\sigma^2 = L^2\sigma^2,$$

что и требовалось доказать. □

Теорема 2.1. (ЦПТ для конечно-зависимых стационарных последовательностей)

Пусть $x_1, \ldots, x_n, \ldots$ — сильно стационарная $(\leq M)$-зависимая последовательность со средним a, ковариационной функцией $R(m)$ и спектральной плотностью f. Тогда

$$\mathcal{L}\big(\sqrt{n}\,(\overline{x}_n - a)\big) \Rightarrow \mathrm{N}(0, \sigma^2),$$

где

$$\sigma^2 = \sum_{l=-\infty}^{\infty} R(l) = R(0) + 2\sum_{l=1}^{M} R(l) = 2\pi f(0). \tag{2.5}$$

Доказательство. Прежде всего заметим, что $R(l) = 0$ при $l > M$ и положим $a = 0$.

Зафиксируем $k > 2M$ и положим $N = \lfloor n/k \rfloor$. Таким образом, $n = kN + r$, где $0 \leq r < k$. Определим следующие случайные величины.

$$\eta'_{kn} = \Big(\big(x_1 + \ldots + x_{k-M}\big) + \big(x_{k+1} + \ldots + x_{2k-M}\big) + \ldots + \big(x_{k(N-1)+1} + \ldots + x_{kN-M}\big)\Big)/\sqrt{n} \tag{2.6}$$

(в каждой из N сумм по $k - M$ слагаемых, суммы независимы и одинаково распределены),

$$\beta_{kn} = \Big(\big(x_{k-M+1} + \ldots + x_k\big) + \big(x_{2k-M+1} + \ldots + x_{2k}\big) + \ldots + \big(x_{kN-M+1} + \ldots + x_{kN}\big)\Big)/\sqrt{n}$$

(в каждой из N сумм по M слагаемых, суммы независимы и одинаково распределены),

$$\eta''_{kn} = (x_{kN+1} + \ldots + x_n)/\sqrt{n}$$

(оставшиеся иксы, их число не больше $r < k$).

Заметим, что $\sqrt{n}\,\overline{x}_n = \eta'_{kn} + \beta_{kn} + \eta''_{kn}$. Рассмотрим сумму (2.6), считая, что k фиксировано, а $n \to \infty$. Обозначим

$$B_k^2 \stackrel{\mathrm{def}}{=} \mathbb{D}(x_1 + \ldots + x_{k-M}).$$

Если $n \to \infty$, то $N \to \infty$ и поэтому по теореме П. Леви

$$\mathcal{L}(\sqrt{n/N}\eta'_{kn}) \Rightarrow \mathrm{N}\big(0, \mathbb{D}(x_1 + \ldots + x_{k-M})\big) = \mathrm{N}(0, B_k^2).$$

Так как одновременно $n/N \to k$, то $\mathcal{L}(\sqrt{k}\,\eta'_{kn}) \Rightarrow \mathrm{N}(0, B_k^2)$, и следовательно,

$$\mathcal{L}(\eta'_{kn}) \Rightarrow \mathcal{P}_k = \mathrm{N}(0, \sigma_k^2),$$

где $\sigma_k^2 = B_k^2/k$. Далее, согласно Предложению 2.2, при $k \to \infty$

$$\sigma_k^2 = \frac{k-M}{k}\,\frac{\mathbb{D}(x_1 + \ldots + x_{k-M})}{k-M} = \left(\frac{k-M}{k}\right)(k-M)\,\mathbb{D}(\overline{x}_{k-M}) \to \sigma^2,$$

где σ^2 определено в (2.5). Поэтому $\mathcal{P}_k \Rightarrow \mathcal{P} = \mathrm{N}(0, \sigma^2)$.

Обозначая $r = n - kN$, получим из Леммы 2.3, что

$$\mathbb{D}\eta''_{kn} = \frac{1}{n}\,\mathbb{D}(x_{kN+1} + \ldots + x_n) \le \frac{r^2}{n}\,R(0) \le \frac{k^2}{n}\,R(0) \to 0$$

при фиксированном k и $n \to \infty$. Поэтому $\eta''_{kn} \xrightarrow{\mathbb{P}} 0$ и

$$\mathcal{L}(\eta'_{kn} + \eta''_{kn}) \Rightarrow \mathcal{P}_k = \mathrm{N}(0, \sigma_k^2).$$

Наконец, $N/n \le 1/k$ и

$$\sup_n \mathbb{D}\beta_{kn} = \sup_n \frac{N}{n}\,\mathbb{D}(x_1 + \ldots + x_M) \le \sup_n \frac{M^2 N}{n}\,R(0) \le \frac{M^2}{k}\,R(0) \to 0$$

при $k \to \infty$. Применяя Лемму 2.2, мы видим, что утверждение доказано. □

2.2.3 ЦПТ для линейного процесса

Рассмотрим теперь линейный процесс (1.3).

Теорема 2.2. *Пусть*

$$x_n = \sum_j c_{n+j}\varepsilon_j = \sum_j c_j \varepsilon_{j-n},$$

где $\sum_j |c_j| < \infty$ и ε_j — независимые одинаково распределенные случайные величины с нулевым средним и дисперсией 1. Тогда при $n \to \infty$

$$\mathcal{L}(\sqrt{n}\,\bar{x}_n) \Longrightarrow \mathrm{N}(0, 2\pi f(0)).$$

Доказательство. Выберем $k > 1$ и положим

$$x_{kn} = \sum_{|j| \le k} c_j \varepsilon_{j-n}, \quad \text{и} \quad u_{kn} = \sum_{|j| > k} c_j \varepsilon_{j-n}.$$

Тогда

$$\sqrt{n}\,\bar{x}_n = \frac{1}{\sqrt{n}} \sum_{\ell=1}^{n} x_{k\ell} + \frac{1}{\sqrt{n}} \sum_{\ell=1}^{n} u_{k\ell} = \eta_{kn} + \beta_{kn}.$$

Заметим, что при фиксированном k последовательность x_{kn} удовлетворяет Теореме 2.1 с $M = 2k$. Поэтому при $n \to \infty$

$$\mathcal{L}(\eta_{kn}) \Longrightarrow \mathcal{P}_k = \mathrm{N}(0, \sigma_k^2)$$

с

$$\sigma_k^2 = \sum_{\ell=-\infty}^{\infty} R_k(\ell),$$

где

$$R_k(\ell) = \sum_j c_j^{(k)} c_{j+\ell}^{(k)}, \quad \text{а} \quad c_j^{(k)} = \begin{cases} c_j & \text{при} \ \ |j| \le k, \\ 0 & \text{при} \ \ |j| > k. \end{cases}$$

Докажем теперь, что при $k \to \infty$

$$\sigma_k^2 \to \sigma^2 = \sum_{\ell=-\infty}^{\infty} R(\ell).$$

Так как последняя сумма равна $2\pi f(0)$, то отсюда будет следовать, что $\mathcal{P}_k \Rightarrow \mathcal{P} = \mathrm{N}(0, 2\pi f(0))$.

Действительно, так как $c_j^{(k)} \to c_j$ при $k \to \infty$ и $|c_j^{(k)} c_{j+\ell}^{(k)}| \le |c_j|\,|c_{j+\ell}|$, причем

$$\sum_j |c_j|\,|c_{j+\ell}| \le \sup_\ell |c_\ell| \left(\sum_j |c_j| \right) < \infty,$$

то $R_k(\ell) \to R(\ell)$. Аналогичным образом $|R_k(\ell)| \le \sum_j |c_j|\,|c_{j+\ell}|$ и

$$\sum_\ell \sum_j |c_j|\,|c_{j+\ell}| = \left(\sum_j |c_j| \right)^2 < \infty.$$

Следовательно, $\sigma_k^2 \to \sigma^2$.

Теперь, чтобы использовать Лемму 2.2, достаточно показать, что β_{kn} равномерно по n сходится по вероятности к нулю при $k \to \infty$. Действительно,

$$\begin{aligned}
\mathbb{D}\beta_{kn} &= \frac{1}{n}\,\mathbb{E}\left(\sum_{\ell=1}^{n} \sum_{|j|>k} c_j \varepsilon_{j-\ell} \right)^2 = \frac{1}{n} \sum_{\ell,m=1}^{n} \mathbb{E}\left(\sum_{|\alpha|>k} c_\alpha \varepsilon_{\alpha-\ell} \right) \left(\sum_{|\beta|>k} c_\beta \varepsilon_{\beta-m} \right) = \\
&= \frac{1}{n} \sum_{\ell,m=1}^{n} \sum_{|\alpha|,\,|\beta|>k} c_\alpha c_\beta\, \mathbb{E}\varepsilon_{\alpha-\ell}\varepsilon_{\beta-m} = \frac{1}{n} \sum_{\ell,m=1}^{n} \sum_{|\alpha|>k,\,|\alpha-\ell+m|>k} c_\alpha c_{\alpha-\ell+m} \le \\
&\le \frac{1}{n} \sum_{\ell=1}^{n} \sum_{m=1}^{\infty} \sum_{|\alpha|>k,\,|\alpha-\ell+m|>k} |c_\alpha|\,|c_{\alpha-\ell+m}| \le \frac{1}{n} \sum_{\ell=1}^{n} \sum_{|\alpha|>k} \sum_{m:\,|\alpha-\ell+m|>k} |c_\alpha|\,|c_{\alpha-\ell+m}| = \\
&= \frac{1}{n} \sum_{\ell=1}^{n} \sum_{|\alpha|>k} |c_\alpha| \sum_{|\beta|>k} |c_\beta| = \left(\sum_{|\alpha|>k} |c_\alpha| \right)^2 \to 0
\end{aligned}$$

при $k \to \infty$ равномерно по n. Утверждение доказано. $\square$

Замечание 2.2. Из утверждения Теоремы 2.2 следует, что если нам удастся построить состоятельную оценку $\widehat{f}_n(0)$ числа $f(0)$, то мы стандартным образом получим возможность строить асимптотические доверительные интервалы для среднего стационарного процесса.

2.3 О сохранении ЦПТ при линейном преобразовании стационарной последовательности

Теорема 2.3. *Пусть $x_1, \ldots, x_n, \ldots$ — слабо стационарная последовательность с нулевым средним и конечными вторыми моментами. При $j \geq 1$ рассмотрим вещественные числа c_j, удовлетворяющие условию $\sum_j |c_j| < \infty$. Обозначим*

$$y_n = \sum_{j\geq 1} c_j x_{n+j}, \quad n \geq 1.$$

Если $\sigma_n^2 \stackrel{\text{def}}{=} \mathbb{D}(x_1 + \ldots + x_n) \to \infty$ и

$$\mathcal{L}\big((x_1 + \ldots + x_n)/\sigma_n\big) \Longrightarrow \mathrm{N}(0,1),$$

то

$$\mathcal{L}\big((y_1 + \ldots + y_n)/\sigma_n\big) \Longrightarrow \mathrm{N}(0,G^2),$$

где $G = \sum_{j\geq 1} c_j$.

Доказательство. Всегда можно считать, что существует такая ортогональная стохастическая мера μ, что $x_n = \int_{(-\pi,\pi]} e^{in\lambda}\mu(d\lambda)$. Тогда $y_n = \int_{(-\pi,\pi]} e^{in\lambda}g(\lambda)\mu(d\lambda)$, где $g(\lambda) = \sum_{l\geq 1} c_l e^{il\lambda}$ является непрерывной на $[-\pi,\pi]$ функцией. Отметим, что $G = g(0)$.

Предположим сначала, что $c_l = 0$ при $l > N$. Тогда

$$\frac{1}{\sigma_n}(y_1 + \ldots + y_n) = \frac{1}{\sigma_n}\sum_{j=1}^{n}\sum_{\ell=1}^{N} c_\ell x_{j+\ell} = \frac{1}{\sigma_n}\sum_{\ell=1}^{N} c_\ell \sum_{j=1}^{n} x_{j+\ell} =$$
$$= \left(\sum_{\ell=1}^{N} c_\ell\right)\frac{x_1+\ldots+x_n}{\sigma_n} + \frac{1}{\sigma_n}\sum_{\ell=1}^{N} c_\ell\left(\sum_{j=1}^{n} x_{j+\ell} - \sum_{j=1}^{n} x_j\right) = J_1(n) + J_2(n).$$

Ясно, что $\mathcal{L}\big(J_1(n)\big) \Rightarrow \mathrm{N}(0,G^2)$. Докажем, что при фиксированном N последовательность $J_2(n)$ сходится по вероятности к нулю. Действительно, так как при $n > N$ в разности

$$I_2(n,\ell) \stackrel{\text{def}}{=} \sum_{j=1}^{n} x_{j+\ell} - \sum_{j=1}^{n} x_j$$

не более $2N$ различных слагаемых вида $\pm\, x_m$, то $\mathbb{D}I_2(n,\ell) \leq 4N^2\mathbb{D}x_1$ и $\mathbb{D}\big(I_2(n,\ell)/\sigma_n\big) \to 0$ из-за условия $\sigma_n^2 \to \infty$. Следовательно, $J_2(n) \xrightarrow{\mathbb{P}} 0$.

Перейдем к общему случаю. Определим $g_N(\lambda) = \sum_{\ell\leq N} c_\ell\, e^{i\ell\lambda}$ и положим

$$y_n^{(N)} = \int_{(-\pi,\pi]} e^{in\lambda}g_N(\lambda)\mu(d\lambda) = \sum_{1\leq j\leq N} c_j x_{n+j}.$$

Таким образом, $y_n = y_{nN} + z_{nN}$, где

$$z_{nN} = \int_{(-\pi,\pi]} e^{in\lambda}\big(g(\lambda) - g_N(\lambda)\big)\mu(d\lambda) = \sum_{j>N} c_j x_{n+j}.$$

Ясно, что $g_N(\lambda)$ сходится к $g(\lambda)$ равномерно по λ.

Как уже доказано, $\mathcal{L}\big((y_{1N}+\ldots+y_{nN})/\sigma_n\big) \Rightarrow \mathcal{P}_N = \mathrm{N}(0, g_N^2(0))$ при фиксированном N и $n \to \infty$. Одновременно $\mathcal{P}_N \Rightarrow \mathcal{P} = \mathrm{N}(0, g^2(0))$.

Теперь, чтобы сослаться на Лемму 2.2, нужно доказать, что z_{nN} при $N \to \infty$ сходится по вероятности к нулю равномерно по n. Но

$$\mathbb{E}z_{nN}^2 = \int_{(-\pi,\pi]} \left|g(\lambda) - g_N(\lambda)\right|^2 m(d\lambda),$$

где m — спектральная мера последовательности x_n. Поскольку эта мера конечна, а g_N сходится к g равномерно, то утверждение доказано. $\square$

Замечание 2.3. Подчеркнем, что в Теореме 2.3 не требуется, чтобы последовательность $\{x_n\}_{n\geq 1}$ была сильно стационарной, достаточно, чтобы она была слабо стационарной и чтобы для выполнялась ЦПТ. Например, если предположить, что случайные величины x_n независимы, то для них достаточно выполнения условия Линдеберга.

3 Оценивание ковариационной функции

3.1 Различные оценки. Смещение

Теперь наша задача — построить и изучить свойства оценки значения $R(k)$ по отрезку траектории $x_1, \ldots, x_n$ стационарного процесса.

Существует несколько стандартных вариантов таких оценок. Во-первых, в теории часто удобно считать, что математическое ожидание $\mathbb{E}x_1 = a$ известно. В этом случае при $0 \le k < n$ можно использовать статистику

$$\widehat{R}_n^{(0)}(k) = \frac{1}{n-k} \sum_{j=1}^{n-k} (x_j - a)(x_{j+k} - a). \tag{3.1}$$

Если a неизвестно (это практическая ситуация), то аналогом оценки $\widehat{R}_n^{(0)}(k)$ является

$$\widehat{R}_n^{(1)}(k) = \frac{1}{n-k} \sum_{j=1}^{n-k} (x_j - \bar{x}_n)(x_{j+k} - \bar{x}_n), \tag{3.2}$$

где $\bar{x}_n = (x_1 + \ldots + x_n)/n$. Оценки (3.1) и (3.2) можно назвать *эмпирическими.*

Иногда удобно использовать несколько другие оценки. А именно, дальше мы будем иметь дело и со статистиками

$$\widetilde{R}_n^{(0)}(k) = \frac{1}{n} \sum_{j=1}^{n-k} (x_j - a)(x_{j+k} - a) = \frac{n-k}{n}\, \widehat{R}_n^{(0)}(k) \tag{3.3}$$

и

$$\widetilde{R}_n^{(1)}(k) = \frac{1}{n} \sum_{j=1}^{n-k} (x_j - \bar{x}_n)(x_{j+k} - \bar{x}_n) = \frac{n-k}{n}\, \widehat{R}_n^{(1)}(k), \tag{3.4}$$

отличающимися от (3.1) и (3.2) только множителем $1/n$ вместо $1/(n-k)$ перед знаком суммы.

Рассмотрим вопросы смещения этих оценок. Ясно, что $\mathbb{E}\widehat{R}_n^{(0)}(k) = R(k)$, так что оценка $\widehat{R}_n^{(0)}(k)$ несмещенная. Соответственно, $\mathbb{E}\widetilde{R}_n^{(0)}(k) = (1 - k/n)R(k)$ и при $k/n \to 0$ есть асимптотическая несмещенность.

Перейдем теперь к статистикам $\widehat{R}_n^{(1)}(k)$ и $\widetilde{R}_n^{(1)}(k)$. Здесь, как всегда, можно положить $a = 0$, так как от вычитания любой константы из всех x_j величина $\widehat{R}_n^{(1)}(k)$ не меняется.

Предложение 3.1. *Пусть* $\Delta^2 \stackrel{\mathrm{def}}{=} \sum_j |R(j)| < \infty$ *и* $a = 0$. *Тогда*
1. $\mathbb{E}\widehat{R}_n^{(1)}(k) = R(k) + O(1/n)$ *равномерно по* $0 \le k < n$.
2. Если $k/n \to 0$ *при* $n \to \infty$, *то*

$$\lim_{n\to\infty} n \left(\mathbb{E}\widehat{R}_n^{(1)}(k) - R(k) \right) = -2\pi f(0).$$

3. $\mathbb{E}\big|\widehat{R}_n^{(1)}(k) - \widehat{R}_n^{(0)}(k)\big| = O(1/n)$ *при любом фиксированном* k.

Доказательство. 1. Распишем $\widehat{R}_n^{(1)}(k)$.

$$\widehat{R}_n^{(1)}(k) = \frac{1}{n-k} \sum_{j=1}^{n-k} (x_j - \bar{x}_n)(x_{j+k} - \bar{x}_n) =$$

$$= \widehat{R}_n^{(0)}(k) - \frac{1}{n-k} \left(\sum_{j=1}^{n-k} x_j \right) \bar{x}_n - \frac{1}{n-k} \left(\sum_{j=1}^{n-k} x_{j+k} \right) \bar{x}_n + \bar{x}_n^2 \tag{3.5}$$

Согласно Предложению 2.2, $\mathbb{E}\bar{x}_n^2 = O(1/n)$. Далее,

$$\mathbb{E}x_j\bar{x}_n = \frac{1}{n}\sum_{m=1}^{n}\mathbb{E}x_jx_m = \frac{1}{n}\sum_{m=1}^{n}R(j-m), \quad \left|\mathbb{E}x_j\bar{x}_n\right| \le \frac{1}{n}\sum_{m=-\infty}^{\infty}|R(m)| = \Delta^2/n,$$

и поэтому математическое ожидание второго и третьего слагаемого в правой части (3.5) оценивается сверху через Δ^2/n.
Второе и третье утверждение докажем одновременно. Поскольку

$$\frac{1}{n-k}\left(\sum_{j=1}^{n-k}x_j\right)\bar{x}_n = \frac{n}{n-k}\left(\bar{x}_n - (x_{n+k-1}+\ldots+x_n)/n\right)\bar{x}_n =$$
$$= \bar{x}_n^2 + \frac{k}{n-k}\,\bar{x}_n^2 - \frac{x_{n-k+1}+\ldots+x_n}{n-k}\,\bar{x}_n$$

и аналогично

$$\frac{1}{n-k}\left(\sum_{j=1}^{n-k}x_{j+k}\right)\bar{x}_n = \frac{n}{n-k}\left(\bar{x}_n - (x_1+\ldots+x_k)/n\right)\bar{x}_n =$$
$$= \bar{x}_n^2 + \frac{k}{n-k}\,\bar{x}_n^2 - \frac{x_1+\ldots+x_k}{n-k}\,\bar{x}_n,$$

то (3.5) получаем, что

$$n\left(\widehat{R}_n^{(1)}(k) - \widehat{R}_n^{(0)}(k)\right) = -n\bar{x}_n^2 + n\frac{2k}{n-k}\,\bar{x}_n^2 - n\left(\frac{x_{n-k+1}+\ldots+x_n}{n-k}\,\bar{x}_n\right) - n\left(\frac{x_1+\ldots+x_k}{n-k}\,\bar{x}_n\right). \quad (3.6)$$

Рассмотрим четыре слагаемых в правой части (3.6). Прежде всего, из Предложения 2.2 следует, что $n\mathbb{E}\bar{x}_n^2 \to 2\pi f(0)$. Если $k = o(n)$, то $k/(n-k) \sim k/n$ и $nk/(n-k)\mathbb{E}\bar{x}_n^2 \sim 2\pi f(0)k/n$.

Остальные два слагаемых в правой части (3.5) рассматриваются одинаково. Возьмем второе из них. Поскольку величины $j\mathbb{E}\bar{x}_j^2$ ограничены по j, то при $k = o(n)$

$$n\left|\mathbb{E}\left(\frac{x_1+\ldots+x_k}{n-k}\,\bar{x}_n\right)\right| \le \frac{nk}{n-k}\,\mathbb{E}\left(\left|\frac{x_1+\ldots+x_k}{k}\right|\left|\bar{x}_n\right|\right) \le$$
$$\le \frac{\sqrt{kn}}{n-k}\sqrt{k\mathbb{E}\bar{x}_k^2\, n\mathbb{E}\bar{x}_n^2} = O(\sqrt{k/n}).$$

Поэтому при $k = o(n)$

$$n\left(\mathbb{E}\widehat{R}_n^{(1)}(k) - R(k)\right) = \mathbb{E}n\left(\widehat{R}_n^{(1)}(k) - \widehat{R}_n^{(0)}(k)\right) \sim -2\pi f(0) + O(k/n) + O(\sqrt{k/n}) \to -2\pi f(0)$$

и при фиксированном k

$$\limsup_n n\mathbb{E}\left|\widehat{R}_n^{(1)}(k) - \widehat{R}_n^{(0)}(k)\right| \le 2\pi f(0).$$

Утверждение доказано. □

Замечание 3.1. Из доказанного сразу же следует, что при $k = o(n)$

$$\mathbb{E}\widetilde{R}_n^{(1)}(k) = \left(\frac{n-k}{n}\right)\mathbb{E}\widehat{R}_n^{(1)}(k) \sim \left(\frac{n-k}{n}\right)\left(R(k) - 2\pi f(0)/n\right) \sim \left(1 - \frac{k}{n}\right)R(k) - \frac{2\pi f(0)}{n}. \quad (3.7)$$

3.2 Состоятельность оценок ковариационной функции. Асимптотические дисперсии

3.2.1 Общий взгляд

Пусть $x_1, \ldots, x_n, \ldots$ — стационарный в узком смысле процесс с нулевым средним и ковариационной функцией $R_x(m)$. Рассмотрим оценку $\widehat{R}_n^{(0)}(k)$ значения ковариационной функции R_x (то есть предполагаем, что среднее известно. Эта оценка является несмещенной. Другие оценки будут рассмотрены позже). Итак,

$$\widehat{R}_n^{(0)}(k) = \frac{1}{n-k} \sum_{j=1}^{n-k} x_j x_{k+j}.$$

Зафиксируем k. Тогда последовательность $y_j = y_j^{(k)} = x_j x_{j+k}$ является сильно стационарной, а ее среднее равно $R_x(k)$. Более того,

$$\bar{y}_{n-k} = \frac{1}{n-k} \big(y_1 + \ldots + y_{n-k}\big) = \widehat{R}_n^{(0)}(k).$$

Поэтому,[2] если конечен второй момент $\mathbb{E}y_k^2 = \mathbb{E}x_0^2 x_k^2$, то общий ответ известен из Предложения 2.2: в обозначениях

$$R_y(m) = R_y^{(k)}(m) = \mathbb{E}y_1 y_{m+1} - R_x^2(k) = \mathbb{E}x_0 x_k x_m x_{m+k} - R_x^2(k)$$

и при выполнении условия $\sum_j |R_y(j)| < \infty$, вопрос о состоятельности и асимптотической дисперсии оценки $\widehat{R}_n^{(0)}(k)$ решается очень просто:

$$n\mathbb{D}\widehat{R}_n^{(0)}(k) \to R_y^{(k)}(0) + 2\sum_{m\geq 1} R_y^{(k)}(m). \tag{3.8}$$

Нужно ответить на несколько вопросов.

1. При каких разумных условиях у процесса y_j существуют конечные вторые моменты? (Здесь ответ ясен: достаточно потребовать, чтобы $\mathbb{E}x_1^4 < \infty$.)

2. Для каких разумных моделей выражение в правой части (3.8) может быть компактно записано?

3. При каких условиях для $\widehat{R}_n^{(0)}(k)$ выполняется не только закон больших чисел (это следует из (3.8)), но и Центральная Предельная Теорема? Более точно, обозначим Δ^2 правую часть (3.8). При каких условиях $\mathcal{L}\Big(\sqrt{n}\big(\widehat{R}_n^{(0)}(k) - R_x(k)\big)\Big) \Rightarrow \mathrm{N}(0, \Delta^2)$?

4. Что происходит со статистиками $\widehat{R}_n^{(1)}(k)$, $\widetilde{R}_n^{(0)}(k)$ и $\widetilde{R}_n^{(1)}(k)$?

Будем решать эти задачи последовательно.

3.2.2 Предельные дисперсии для линейного процесса

Теорема 3.1. ([1], Теор. 8.4.2, с. 518, Следств. 8.4.1, с. 521)
Пусть x_m имеет вид (1.3), где ε_j — одинаково распределенные независимые случайные величины, имеющие нулевое среднее, единичную дисперсию и конечный четвертый момент. Обозначим $R_x(m)$ ковариационную функцию процесса x_m. Если $\sum_j |c_j| < \infty$, то при $n \to \infty$

$$n\mathbb{D}\widehat{R}_n^{(0)}(k) \to \sum_l \Big(R_x(k+l)R_x(k-l) + R_x^2(l)\Big) + (\mathbb{E}\varepsilon^4 - 3)R_x^2(k). \tag{3.9}$$

[2] Здесь нам удобнее считать, что ряд x_m начинается с нулевого момента времени.

Доказательство. Согласно (3.8), нам достаточно доказать, что в обозначениях $y_j = x_j x_{j+k}$ ряд $\sum_j R_y^{(k)}(j)$ абсолютно сходится и что

$$\sum_j R_y^{(k)}(j) = \sum_\ell \Big(R_x(k+\ell)R_x(k-\ell) + R_x^2(\ell)\Big) + (\mathbb{E}\varepsilon^4 - 3)R_x^2(k).$$

Поскольку оба доказательства проводятся практически одинаково, остановимся только на втором.

Найдем выражение для $R_y^{(k)}(j)$. По определению,

$$R_y^{(k)}(j) = \mathbb{E}y_0y_j - R_x^2(k) = \mathbb{E}x_0x_kx_jx_{j+k} - R_x^2(k) =$$
$$= \mathbb{E}\left(\sum_{i_1} c_{i_1}\varepsilon_{i_1}\right)\left(\sum_{i_2} c_{i_2+k}\varepsilon_{i_2}\right)\left(\sum_{i_3} c_{i_3+j}\varepsilon_{i_3}\right)\left(\sum_{i_4} c_{i_4+j+k}\varepsilon_{i_4}\right) - R_x^2(k) =$$
$$= \sum_{i_1i_2i_3i_4} c_{i_1}\, c_{i_2+k}\, c_{i_3+j}\, c_{i_4+k+j}\, \mathbb{E}\varepsilon_{i_1}\varepsilon_{i_2}\varepsilon_{i_3}\varepsilon_{i_4} - R_x^2(k) =$$
$$= \sum_{i_1=i_2=i_3=i_4} \square + \sum_{i_1=i_2,\, i_3=i_4\, i_1\neq i_3} \square + \sum_{i_1=i_3,\, i_2=i_4\, i_1\neq i_2} \square + \sum_{i_1=i_4,\, i_2=i_3\, i_1\neq i_2} \square - R_x^2(k) =$$
$$= \mathbb{E}\varepsilon^4 I_1(j) + \big(I_2(j) - R_x^2(k)\big) + I_3(j) + I_4(j),$$

где значок $\square$ заменяет выражение $c_{i_1}\, c_{i_2+k}\, c_{i_3+j}\, c_{i_4+k+j}\, \mathbb{E}\varepsilon_{i_1}\varepsilon_{i_2}\varepsilon_{i_3}\varepsilon_{i_4}$. Докажем, что

$$\sum_j I_1(j) = R_x^2(k), \tag{3.10}$$
$$\sum_j \big(I_2(j) - R_x^2(k)\big) = -R_x^2(k),$$
$$\sum_j I_3(j) = \sum_l R_x^2(l) - R_x^2(k), \tag{3.11}$$
$$\sum_j I_4(j) = \sum_l R_x(k+l)R_x(k-l) - R_x^2(k).$$

Поскольку

$$I_1(j) = \sum_i c_ic_{i+k}c_{i+j}c_{i+j+k}$$

и ряд $\sum_m |c_m|$ сходится, то с учетом равенства (1.4)

$$\sum_j I_1(j) = \sum_i c_ic_{i+k}\sum_j c_{i+j}c_{i+j+k} = \sum_i c_ic_{i+k}\sum_l c_lc_{l+k} = R_x^2(k).$$

Тем самым (3.10) доказано. Далее, так как $\mathbb{E}\varepsilon^2 = 1$, то

$$\sum_j \big(I_2(j) - R_x^2(k)\big) = \sum_j \left(\sum_{i\neq m} c_ic_{i+k}c_{j+m}c_{j+m+k} - \sum_i c_ic_{i+k}\sum_m c_{j+m}c_{j+m+k}\right) =$$
$$= -\sum_j\sum_i c_ic_{i+k}c_{j+i}c_{j+i+k} = -\sum_j I_1(j) = -R_x^2(k).$$

Перейдем к I_3. Поскольку

$$I_3(j) = \sum_{i\neq m} c_ic_{m+k}c_{i+j}c_{m+j+k} = \sum_{i,m} c_ic_{m+k}c_{i+j}c_{m+j+k} - \sum_i c_ic_{i+k}c_{i+j}c_{i+j+k} =$$
$$= \sum_{i,m} c_ic_{m+k}c_{i+j}c_{m+j+k} - I_1(j),$$

то

$$\sum_j I_3(j) = \sum_j \sum_i c_i c_{i+j} \left(\sum_m c_{m+k} c_{m+j+k} \right) - \sum_j I_1(j) =$$
$$= \sum_j R_x(j) \left(\sum_i c_i c_{i+j} \right) - R_x^2(k),$$

и мы приходим к выражению (3.11). Наконец,

$$I_4(j) = \sum_{i \neq m} c_i c_{m+k} c_{m+j} c_{i+j+k}.$$

Поэтому

$$\sum_j I_4(j) = \sum_j \sum_{i,m} c_i c_{m+k} c_{m+j} c_{i+j+k} - \sum_j \sum_i c_i c_{i+k} c_{i+j} c_{i+j+k} \tag{3.12}$$

Как уже говорилось, второй член в (3.12) раевн $-R_x^2(k)$. Рассмотрим первую сумму.

$$\sum_{i,j,m} c_i c_{m+k} c_{m+j} c_{i+j+k} = \sum_{i,j} c_i c_{i+j+k} \sum_m c_{m+k} c_{m+j} = \sum_{i,j} c_i c_{i+j+k} R_x(k-j) =$$
$$= \sum_j R_x(k-j) \sum_i c_i c_{i+j+k} = \sum_j R_x(k-j) R_x(k+j).$$

Предложение доказано. $\square$

Замечание 3.2. Абсолютно аналогично доказывается более общее утверждение. А именно, в предположениях Предложения 3.1 для любых фиксированных k, m

$$n \operatorname{Cov}\Big(\widehat{R}_n^{(0)}(k), \widehat{R}_n^{(0)}(m)\Big) \to \sum_l \big(R_x(k+l)R_x(m-l) + R_x(k+l)R_x(m+l)\big) +$$
$$+ \big(\mathbb{E}\varepsilon^4 - 3\big) R_x(k) R_x(m) \tag{3.13}$$

при $n \to \infty$.

3.2.3 Другой вид предельных ковариаций

Выражение в правой части (3.13) (и, в частности, в правой части (3.9)) является довольно громоздким. Оказывается, его можно значительно упростить.

Предложение 3.2. *Если $\sum_j |R_x(j)| < \infty$, то*

$$\sum_l \Big(R_x(k+l)R_x(m-l) + R_x(k+l)R_x(m+l)\Big) = 4\pi \int_{-\pi}^{\pi} \cos(\lambda k)\cos(\lambda m) f^2(\lambda) d\lambda, \tag{3.14}$$

где $f(\lambda)$ — спектральная плотность процесса x_n.

Доказательство. Как уже говорилось, в предположении, что $\sum_j |R_x(j)| < \infty$ спектральная плотность f непрерывна и ограничена и поэтому интеграл в правой части (3.14) конечен. Преобразуем это выражение. Так как

$$f(\lambda) = \frac{1}{2\pi} \sum_j R(j) e^{-i\lambda j},$$

где $R(j) = R_x(j)$, то

$$f^2(\lambda) = \frac{1}{4\pi^2} \sum_{j,p=1}^{\infty} R(j)R(p)e^{-(p+j)\lambda},$$

$$\int_{-\pi}^{\pi} f^2(\lambda)e^{il\lambda}d\lambda = \frac{1}{4\pi} \sum_{j,p} R(j)R(p) \int_{-\pi}^{\pi} e^{i\lambda(l-p-j)}d\lambda = \frac{1}{2\pi} \sum_{j+p=l} R(j)R(p),$$

и

$$\cos(\lambda k)\,\cos(\lambda m) = \frac{1}{4}\left(e^{i(k+m)\lambda} + e^{i(k-m)\lambda} + e^{i(-k+m)\lambda} + e^{-i(k+m)\lambda}\right),$$

то

$$4\pi \int_{-\pi}^{\pi} f^2(\lambda)\cos(\lambda k)\cos(\lambda m)d\lambda =$$

$$= \frac{1}{2}\left(\sum_{p+j=k+m} R(j)R(p) + \sum_{p+j=k-m} R(j)R(p) + \sum_{p+j=-k+m} R(j)R(p) + \sum_{p+j=-k-m} R(j)R(p)\right) =$$

$$= \sum_{p+j=k+m} R(j)R(p) + \sum_{p+j=k-m} R(j)R(p) = \sum_{j} R(j)R(k+m-j) + \sum_{j} R(j)R(k-m-j). \quad (3.15)$$

Делая в обоих слагаемых правой части (3.15) замену $j \mapsto l = j-k$, получаем требуемое. $\square$

Пример 3.1. Посмотрим, как выглядит правая часть (3.13) в случае, когда $x_n = \sigma\varepsilon_n$, то есть когда исходный ряд — белый шум с дисперсией $\sigma^2 > 0$. При этом будем пользоваться представлением (3.14). Прежде всего, $f(\lambda) = \sigma^2/2\pi$. Поэтому при $k, m \geq 0$

$$4\pi \int_{-\pi}^{\pi} \cos(\lambda k)\cos(\lambda m)f^2(\lambda)d\lambda = \frac{\sigma^4}{4\pi}\int_{-\pi}^{\pi} \left(e^{i(k+m)\lambda} + e^{-i(k+m)\lambda} + e^{i(-k+m)\lambda} + e^{i(k-m)\lambda}\right)d\lambda =$$

$$= \frac{\sigma^4}{4\pi}\begin{cases} 8\pi & \text{при} \quad k = m = 0, \\ 4\pi & \text{при} \quad k = m \neq 0, \\ 0 & \text{при} \quad k \neq m; \end{cases} = \begin{cases} 2\sigma^4 & \text{при} \quad k = m = 0, \\ \sigma^4 & \text{при} \quad k = m \neq 0, \\ 0 & \text{при} \quad k \neq m. \end{cases}$$

Таким образом, для белого шума

$$n\,\mathrm{Cov}\left(\widehat{R}_n^{(0)}(k), \widehat{R}_n^{(0)}(m)\right) \to \begin{cases} \sigma^4\left(\mathbb{E}\varepsilon^4 - 1\right) & \text{при} \quad k = m = 0, \\ \sigma^4 & \text{при} \quad k = m \neq 0, \\ 0 & \text{при} \quad k \neq m. \end{cases}$$

Если же последовательность ε_j является еще и гауссовской, то $\mathbb{E}\varepsilon^4 = 3$ и

$$n\,\mathrm{Cov}\left(\widehat{R}_n^{(0)}(k), \widehat{R}_n^{(0)}(m)\right) \to \begin{cases} 2\sigma^4 & \text{при} \quad k = m = 0, \\ \sigma^4 & \text{при} \quad k = m \neq 0, \\ 0 & \text{при} \quad k \neq m. \end{cases}$$

3.3 О центральной предельной теореме

3.3.1 ЦПТ для оценок ковариационной функции

Имеет место следующее утверждение. Обозначим при $K \geq 1$

$$w_{km} = \sum_l \big(R_x(k+l)R_x(m-l) + R_x(k+l)R_x(m+l)\big) + \big(\mathbb{E}\varepsilon^4 - 3\big)\, R_x(k)R_x(m), \tag{3.16}$$

$W_K = \{w_{km}\}_{k,m=0}^{K-1}$ и $\xi_j^{(n)} = \sqrt{n}\,\Big(\widehat{R}_n^{(0)}(j) - R_x(j)\Big)$.

Теорема 3.2. *В условиях Теоремы 3.1 для любого $K \geq 1$ при $n \to \infty$*

$$\mathcal{L}\left(\left(\xi_0^{(n)}, \ldots, \xi_{K-1}^{(n)}\right)^{\mathrm{T}}\right) \Longrightarrow \mathrm{N}(0, W_K). \tag{3.17}$$

Доказательство. Мы не будем приводить доказательство этого предложения. Отметим только, что оно может быть проведено тем же способом, что и Центральная Предельная Теорема 2.2 для среднего: сначала происходит «урезание» линейного процесса x_n до конечной линейной комбинации независимых случайных величин, затем применяется ЦПТ для конечно-зависимых случайных величин и, наконец, используется Лемма 2.2. □

Обсудим теперь результат Предложения 3.2.

Замечание 3.3. 1. Из (3.17) следует одномерная ЦПТ вида

$$\mathcal{L}\big(\sqrt{n}(\widehat{R}_n^{(0)}(k) - R(k))\big) \Longrightarrow \mathrm{N}(0, w_{kk}). \tag{3.18}$$

2. Следует обратить внимание, что ковариации предельного гауссовского распределения $\mathrm{N}(0, \Sigma_K)$ совпадают с предельными ковариациями в (3.13).
3. В статистике предельные теоремы типа (3.18) используются для построения доверительных интервалов, покрывающих оцениваемый параметр. Обсудим (3.18) с этой точки зрения. Для построения асимптотического доверительного интервала значения ковариационной функции $R_x(k)$ достаточно иметь состоятельную оценку $\widehat{w}_{kk}$ дисперсии w_{kk}. Тогда (в облегчающем жизнь предположении, что $\widehat{w}_{kk} > 0$) этот доверительный интервал уровня $1-\gamma$ выглядит как

$$\left(\widehat{R}_n^{(0)}(k) - x_\gamma \sqrt{\widehat{w}_{kk}}/\sqrt{n},\; \widehat{R}_n^{(0)}(k) + x_\gamma \sqrt{\widehat{w}_{kk}}/\sqrt{n}\right),$$

где x_γ — решение уравнения $\Phi(x_\gamma) - \Phi(x_\gamma) = 1 - \gamma$, а Φ — функция распределения стандартного нормального закона.

Согласно (3.9) и (3.14),

$$\begin{aligned} w_{kk} &= \sum_l \Big(R_x(k+l)R_x(k-l) + R_x^2(l)\Big) + (\mathbb{E}\varepsilon^4 - 3)R_x^2(k) = \\ &= 4\pi \int_{-\pi}^{\pi} \cos(\lambda k)\cos(\lambda k) f^2(\lambda) d\lambda + (\mathbb{E}\varepsilon^4 - 3)R_x^2(k). \end{aligned} \tag{3.19}$$

Вторая (содержащая спектральную плотность) формула в правой части (3.19) представляется более удобной для построения состоятельной оценки дисперсии w_{kk}. Действительно, пусть $\widehat{f}_n(\lambda)$ — такая (измеримая по λ) оценка спектральной плотности, что $\sup_\lambda \left|\widehat{f}_n(\lambda) - f(\lambda)\right| \xrightarrow{\mathbb{P}} 0$. Тогда

$$\begin{aligned} \mathbb{P}\left(\left|\int_{-\pi}^{\pi} \widehat{f}_n(\lambda)d\lambda - \int_{-\pi}^{\pi} f(\lambda)d\lambda\right| > \varepsilon\right) &\leq \mathbb{P}\left(\int_{-\pi}^{\pi} \left|\widehat{f}_n(\lambda) - f(\lambda)\right| d\lambda > \varepsilon\right) \leq \\ &\leq \mathbb{P}\left(2\pi \sup_\lambda \left|\widehat{f}_n(\lambda) - f(\lambda)\right| > \varepsilon\right) \to 0. \end{aligned}$$

Однако слагаемое, содержащее четвертый момент $\mathbb{E}\varepsilon^4$ белого шума ε_n, представляется гораздо более сложным для оценивания, так как этот белый шум непосредственно не наблюдается.

Стандартным выходом из этого положения является предположение о гауссовости белого шума ε_n и, следовательно, о гауссовости самого процесса x_n. В этом случае $\mathbb{E}\varepsilon^4 = 3$ и предельные ковариации w_{km} выражаются только через ковариации самого процесса (в более сжатой форме — через его спектральную плотность).

Такое предположение, однако, является произвольным и часто нарушается на практике. Более практичным методом борьбы с этой сложностью является переход от ковариационной к корреляционной функции.

3.3.2 ЦПТ для оценок корреляционной функции

Для перехода от ЦПТ для выборочных ковариаций к ЦПТ для выборочных корреляций нам понадобятся следующее простое утверждение.

Лемма 3.1. *Рассмотрим последовательность случайных векторов $\bar{\xi}_n \in \mathbb{R}^k$ таких, что $\mathcal{L}(\bar{\xi}_n) \Rightarrow \mathrm{N}(0,\Sigma)$ и матрицу $A : \mathbb{R}^k \mapsto \mathbb{R}^m$. Тогда $\mathcal{L}(A\bar{\xi}_n) \Rightarrow \mathrm{N}(0, A\Sigma A^{\mathrm{T}})$.*

Доказательство. Обозначим φ_n характеристическую функцию вектора $\bar{\xi}_n$. По условию, $\varphi_n(\bar{t}) \to e^{-(\Sigma\bar{t},\bar{t})/2}$ для любого $\bar{t} \in \mathbb{R}^k$. Отсюда, поскольку характеристическая функция вектора $A\bar{\xi}_n$ в точке $\bar{s} \in \mathbb{R}^m$ имеет вид $\varphi_n(A^{\mathrm{T}}\bar{s})$, то

$$\varphi_n(A^{\mathrm{T}}\bar{s}) \to e^{-(\Sigma A^{\mathrm{T}}\bar{s}, A^{\mathrm{T}}\bar{s})/2} = e^{-(A\Sigma A^{\mathrm{T}}\bar{s},\bar{s})/2},$$

что и доказывает утверждение. □

Обозначим теперь

$$\widehat{\rho}_n^{(0)}(j) = \frac{\widehat{R}_n^{(0)}(j)}{\widehat{R}_n^{(0)}(0)}$$

и положим $\eta_j^{(n)} = \sqrt{n}\big(\widehat{\rho}_n^{(0)}(j) - \rho(j)\big)$, где $\rho(j) = R(j)/R(0)$.

Теорема 3.3. *В условиях Теоремы 3.1 для любого $k \geq 1$*

$$\mathcal{L}\left(\left(\eta_1^{(n)}, \ldots, \eta_k^{(n)}\right)^{\mathrm{T}}\right) \Longrightarrow \mathrm{N}(0, U_k),$$

где элементы матрицы U_k имеют вид

$$\begin{aligned} u_{j\ell} &= \frac{4\pi}{R^2(0)} \int_{-\pi}^{\pi} \big(\cos(j\lambda) - \rho(j)\big)\big(\cos(l\lambda) - \rho(\ell)\big) f^2(\lambda) d\lambda = \\ &= \sum_r \Big(\rho(r+j) + \rho(r-\ell) - 2\rho(j)\rho(r)\Big)\Big(\rho(r+\ell) + \rho(r-j) - 2\rho(\ell)\rho(r)\Big). \end{aligned} \tag{3.20}$$

Доказательство. Для простоты докажем это утверждение в его одномерном варианте.[3] Иначе говоря, докажем, что для любого k

$$\mathcal{L}\left(\sqrt{n}\big(\widehat{\rho}_n^{(0)}(k) - \rho(k)\big)\right) \Longrightarrow \mathrm{N}(0, u_{kk}). \tag{3.21}$$

[3]Обобщение на многомерный случай не представляет труда. Иногда формулу (3.20) называют формулой Бартлетта.

Заметим, что

$$\sqrt{n}\big(\widehat{\rho}_n^{(0)}(k) - \rho(k)\big) = \sqrt{n}\,\frac{\widehat{R}_n^{(0)}(k) - \rho(k)\widehat{R}_n^{(0)}(0)}{\widehat{R}_n^{(0)}(0)}$$

и рассмотрим сначала предельное распределение последовательности случайных величин

$$\sqrt{n}\left(\widehat{R}_n^{(0)}(k) - \rho(k)\widehat{R}_n^{(0)}(0)\right).$$

Согласно Теореме 3.2, $\mathcal{L}\left(\sqrt{n}\big(\widehat{R}_n^{(0)}(k) - R(k), \widehat{R}_n^{(0)}(0) - R(0)\big)^{\mathrm{T}}\right) \Longrightarrow \mathbb{N}(0, W_{k0})$, где

$$W_{k0} = \begin{pmatrix} w_{kk} & w_{k0} \\ w_{0k} & w_{00} \end{pmatrix},$$

а числа w_{il} определены в (3.16). Тем самым, так как

$$\widehat{R}_n^{(0)}(k) - \rho(k)\widehat{R}_n^{(0)}(0) = \widehat{R}_n^{(0)}(k) - \rho(k)\widehat{R}_n^{(0)}(0) - \big(R(k) - \rho(k)R(0)\big) =$$
$$= \widehat{R}_n^{(0)}(k) - R(k) - \rho(k)\big(\widehat{R}_n^{(0)}(0) - R(0)\big) = Z^{\mathrm{T}}\left(\widehat{R}_n^{(0)}(k) - R(k), \widehat{R}_n^{(0)}(0) - R(0)\right)^{\mathrm{T}},$$

где $Z = (1, -\rho(k))^{\mathrm{T}}$, то, согласно Лемме 3.1

$$\mathcal{L}\left(\sqrt{n}\left(\widehat{R}_n^{(0)}(k) - \rho(k)\widehat{R}_n^{(0)}(0)\right)\right) \Longrightarrow \mathbb{N}(0, Z^{\mathrm{T}} W_{k0} Z),$$

Легко видеть, что $Z^{\mathrm{T}} W_{k0} Z = w_{00}\rho^2(k) - 2w_{0k}\rho(k) + w_{kk}$. Отметим, что согласно Предложению 3.2

$$w_{00} = 4\pi \int_{-\pi}^{\pi} f^2(\lambda)d\lambda + \kappa_4 R^2(0),$$
$$w_{0k} = 4\pi \int_{-\pi}^{\pi} \cos(\lambda k) f^2(\lambda)d\lambda + \kappa_4 R(0)R(k),$$
$$w_{kk} = 4\pi \int_{-\pi}^{\pi} \cos^2(\lambda k) f^2(\lambda)d\lambda + \kappa_4 R^2(k),$$

где $\kappa_4 = \mathbb{E}\varepsilon^4 - 3$. Поэтому

$$w_{00}\rho^2(k) - 2w_{0k}\rho(k) + w_{kk} = 4\pi \int_{-\pi}^{\pi} \left(\rho^2(k) - 2\cos(\lambda k)\rho(k) + \cos^2(\lambda k)\right) f^2(\lambda)d\lambda +$$
$$+\kappa_4\left(R^2(0)\rho^2(k) - 2R(0)R(k)\rho(k) + R^2(k)\right) = 4\pi \int_{-\pi}^{\pi} \big(\cos(k\lambda) - \rho(k)\big)^2 f^2(\lambda)d\lambda.$$

Поскольку по Теореме 3.1 $\widehat{R}_n^{(0)}(0)$ сходится по вероятности к $R(0)$, то (3.21) доказано, причем дисперсия w_{kk} представлена в интегральной форме. Переход от представления w_{kk} в интегральной форме к представлению в виде суммы произведений значений корреляционной функции проводится так же как в Предложении 3.2. □

Пример 3.2. Если исходный процесс — белый шум, то, как нетрудно видеть, $u_{j\ell} = \delta_{j-\ell}$, то есть оценки корреляционной функции белого шума в пределе являются независимыми и имеют стандартное нормальное распределение. На этом факте основан один из популярных критериев проверки того, что стационарная последовательность является белым шумом.

Пример 3.3. Рассмотрим процесс скользящего суммирования

$$x_n = \varepsilon_n + \sum_{j=1}^{p} a_j \varepsilon_{n-j}.$$

Для такого процесса $\rho(\ell) = 0$ при $|\ell| > p$ и, как нетрудно видеть,

$$u_{kk} = 1 + 2\sum_{j=1}^{p} \rho^2(j)$$

при $k > p$.

Пример 3.4. Используя равенство

$$u_{kk} = \sum_{r} \Big(\rho(r+k) + \rho(r-k) - 2\rho(k)\rho(r)\Big)^2,$$

можно легко сосчитать предельные дисперсии оценок ковариационной функции реализуемого процесса авторегрессии первого порядка

$$x_n = \rho x_{n-1} + \sigma\sqrt{1-\rho^2}\varepsilon_n$$

с $\rho(\ell) = \rho^{|\ell|}$. В этом случае

$$u_{kk} = \frac{(1-\rho^{2k})(1+\rho^2)}{1-\rho^2} - 2k\rho^{2k}.$$

В частности, $u_{11} = 1-\rho^2$, $u_{22} = (1+3\rho^2)(1-\rho^2)$ и $u_{kk} \to (1+\rho^2)/(1-\rho^2)$ при $k \to \infty$.

Замечание 3.4. 1. Учитывая п. 3 Замечания 3.3 можно сделать вывод, что имея подходящую (например, равномерно по λ стремящуюся по вероятности к $f(\lambda)$) оценку спектральной плотности, мы можем построить асимптотический доверительный интервал для значения корреляционной функции.
2. Как доказано в [1], утверждение Теоремы 3.3 остается верным и без предположения о конечности четвертого момента белого шума, но зато с более ограничительными условиями на гладкость спектральной плотности. А именно, имеет место следующее утверждение, приводимое здесь без доказательства.

Теорема 3.4. *([1], теор. 8.4.6).*
Пусть x_m имеет вид (1.3), где ε_j — независимые одинаково распределенные случайные величины, имеющие нулевое среднее и единичную дисперсию. Потребуем, чтобы $\sum_j |c_j| < \infty$ и $\sum_j |j| c_j^2 < \infty$. Тогда имеет место утверждение Теоремы 3.3.

Конечно, с чисто математической точки зрения отказ от требования конечности четвертого момента — это серьезное ослабление условий теоремы. С точки зрения практики, тем не менее, это не так важно. Нужны серьезные основания, чтобы считать, что, например, предположение о конечности четвертого момента для повторной выборки не выполнено.

3.3.3 Другие оценки ковариационной и корреляционной функций

Как уже говорилось, оценка $\widehat{R}_n^{(0)}(k)$ значения ковариационной функции $R(k)$ имеет скорее теоретическое, чем практическое значение. На практике вместо $\widehat{R}_n^{(0)}(k)$ используются оценки $\widehat{R}_n^{(1)}(k)$ и $\widetilde{R}_n^{(1)}(k)$, определенные в (3.2) и (3.4). Нам понадобится также оценка $\widetilde{R}_n^{(0)}(k)$ введенная в (3.3).

Предложение 3.3. *Если в теоремах 3.2, 3.3 и 3.4 заменить* $\widehat{R}_n^{(0)}(k)$ *на* $\widehat{R}_n^{(1)}(k)$, $\widetilde{R}_n^{(1)}(k)$ *или* $\widetilde{R}_n^{(0)}(k)$, *то утверждения этих теорем не изменятся.*

Доказательство. Будем доказывать этот факт для статистики $\widehat{R}_n^{(1)}(k)$ и Теоремы 3.2, остальные случаи рассматриваются аналогично.

Согласно третьему пункту Предложения 3.1, $\mathbb{E}\left|\widehat{R}_n^{(1)}(k) - \widehat{R}_n^{(0)}(k)\right| = O(1/n)$. Поэтому

$$\mathbb{P}\left(\sqrt{n}\left|\widehat{R}_n^{(1)}(k) - \widehat{R}_n^{(0)}(k)\right| > \varepsilon\right) \le \frac{\sqrt{n}\,\mathbb{E}\left|\widehat{R}_n^{(1)}(k) - \widehat{R}_n^{(0)}(k)\right|}{\varepsilon} = O(1/\sqrt{n}).$$

Иными словами, $\sqrt{n}\left(\widehat{R}_n^{(1)}(k) - \widehat{R}_n^{(0)}(k)\right) \xrightarrow{\mathbb{P}} 0$ при $n \to \infty$.

Поэтому последовательности распределений случайных векторов

$$\sqrt{n}\,\left(\widehat{R}_n^{(0)}(0) - R_x(0), \ldots, \widehat{R}_n^{(0)}(k) - R_x(k)\right)$$

и

$$\sqrt{n}\,\left(\widehat{R}_n^{(1)}(0) - R_x(0), \ldots, \widehat{R}_n^{(1)}(k) - R_x(k)\right)$$

имеют одинаковый слабый предел (если он, конечно, существует). Согласно Теореме 3.2 этот предел существует и равен $\mathrm{N}(0, W_k)$. □

4 Периодограммы и их свойства

4.1 Суммы мнимых экспонент и ядро Фейера

Пусть $\beta \in \mathbb{R}$ и $n \geq 1$. Обозначим

$$\varphi_n(\beta) = \frac{1}{\sqrt{2\pi n}} \sum_{k=1}^{n} e^{i\beta k} = \frac{e^{i\beta}}{\sqrt{2\pi n}} \begin{cases} n & \text{при} \ \ \beta = 2\pi\ell,\ \ell \in \mathbb{Z}, \\ \dfrac{e^{in\beta}-1}{e^{i\beta}-1} & \text{иначе.} \end{cases} =$$

$$= \frac{e^{i\beta}}{\sqrt{2\pi n}} \begin{cases} n & \text{при} \ \ \beta = 2\pi\ell,\ \ell \in \mathbb{Z}, \\ e^{i(n-1)\beta/2}\ \dfrac{\sin(n\beta/2)}{\sin(\beta/2)} & \text{иначе.} \end{cases}$$

Конечно, $\varphi_n(-\beta) = \overline{\varphi_n(\beta)}$. Кроме того, $\varphi_n(\beta + 2\pi) = \varphi_n(\beta)$ и

$$\int_{\lambda-\pi}^{\lambda+\pi} \varphi_n(\beta) d\beta = \frac{1}{\sqrt{2\pi n}} \sum_{k=1}^{n} \int_{\lambda-\pi}^{\lambda+\pi} e^{i\beta k} d\beta = 0$$

для любого $\lambda \in \mathbb{R}$.

Ядром Фейера называется функция $\Phi_n(\beta) = |\varphi_n(\beta)|^2$. Иными словами, ядро Фейера задается формулой

$$\Phi_n(\beta) = \frac{1}{2\pi n} \left| \sum_{k=1}^{n} e^{i\beta k} \right|^2 = \frac{1}{2\pi} \begin{cases} n & \text{при} \ \ \beta = 2\pi\ell,\ \ell \in \mathbb{Z}, \\ \dfrac{\sin^2\left(\beta n/2\right)}{n \sin^2\left(\beta/2\right)} & \text{иначе.} \end{cases}$$

Ясно, что Φ_n — неотрицательная непрерывная четная 2π-периодическая функция. Ядро Фейера обладает следующими свойствами.

Лемма 4.1. *1. Для любого λ*

$$\int_{-\pi+\lambda}^{\pi+\lambda} \Phi_n(\beta) d\beta = \int_{-\pi}^{\pi} \Phi_n(\lambda + \alpha) d\alpha = 1.$$

2. Для любого $\pi > \beta_0 > 0$ при $n \to \infty$ имеет место сходимость

$$\sup_{\pi \geq |\beta| \geq \beta_0} \Phi_n(\beta) \to 0.$$

3. При $\lambda \in [0, \pi]$ обозначим $\mathcal{P}_n^{(\lambda)}$ распределение, соответствующее плотности, равной функции Φ_n, суженной на $[-\pi + \lambda, \pi + \lambda]$. Если $\lambda \in [0, \pi)$, то $\mathcal{P}_n^{(\lambda)} \Rightarrow \delta_0$ при $n \to \infty$. При $\lambda = \pi$ имеет место сходимость $\mathcal{P}_n^{(\lambda)} \Rightarrow \delta_0/2 + \delta_{2\pi}/2$.

Доказательство. 1. Так как

$$\int_{-\pi+\lambda}^{\pi+\lambda} \Phi_n(\beta) d\beta = \frac{1}{2\pi n} \sum_{k,l=1}^{n} \int_{-\pi+\lambda}^{\pi+\lambda} e^{i\beta(k-l)} d\beta,$$

утверждение очевидно.

2. Отметим, что на промежутке $[\beta_0, \pi]$ функция $1/\sin(\beta/2)$ не возрастает. Поэтому на этом промежутке

$$\Phi_n(\beta) \leq \frac{1}{2\pi n \sin^2(\beta/2)} \leq \frac{1}{2\pi n \sin^2(\beta_0/2)}$$

и второе утверждение тоже доказано.
3. Рассмотрим сначала случай $\lambda = 0$. Пусть $g \in \mathrm{BL}(\mathbb{R})$, $|g| \le M$ и $|g(\beta_1) - g(\beta_2)| \le L\,|\beta_1 - \beta_2|$. Тогда

$$\left|\int_{-\pi}^{\pi} g(\beta)\Phi_n(\beta)d\beta - g(0)\right| = \left|\int_{-\pi}^{\pi} \big(g(\beta) - g(0)\big)\Phi_n(\beta)d\beta\right| \le$$
$$\le \int_{|\beta|\le\beta_0} \big|g(\beta) - g(0)\big|\Phi_n(\beta)d\beta + \int_{\pi>|\beta|>\beta_0} \big|g(\beta) - g(0)\big|\Phi_n(\beta)d\beta \le$$
$$\le L\int_{|\beta|\le\beta_0} |\beta|\Phi_n(\beta)d\beta + 2M\int_{\pi>|\beta|>\beta_0} \Phi_n(\beta)d\beta \le L\beta_0 + \frac{2M(\pi-\beta_0)}{n\sin^2(\beta_0/2)}\ ,$$

что и доказывает утверждение для $\lambda = 0$. Для $\lambda \in (0,\pi)$ доказательство аналогично: нужно рассматривать выражение

$$\int_{\lambda-\pi}^{\lambda+\pi} \big(g(\beta) - g(0)\big)\Phi_n(\beta)d\beta$$

и окружать интервалом длины $2\beta_0$ точку 0, где $\beta_0 < |\lambda - \pi|$.

Для случая $\lambda = \pi$ нужно рассматривать выражение

$$B_n \stackrel{\text{def}}{=} \left|\int_0^{2\pi} g(\beta)\Phi_n(\beta)d\beta - \big(g(0) + g(2\pi)\big)/2\right|.$$

Так как

$$\int_{\pi}^{2\pi} \Phi_n(\beta)d\beta = \int_0^{\pi} \Phi_n(\beta)d\beta = 1/2$$

и, более того,

$$\int_{\pi}^{2\pi} g(\beta)\Phi_n(\beta)d\beta = \int_0^{\pi} g(2\pi-\beta)\Phi_n(2\pi-\beta)d\beta = \int_0^{\pi} g(2\pi-\beta)\Phi_n(\beta)d\beta,$$

то

$$B_n \le \left|\int_0^{\pi} \big(g(\beta) - g(0)\big)\Phi_n(\beta)d\beta\right| + \left|\int_{\pi}^{2\pi} \big(g(\beta) - g(2\pi)\big)\Phi_n(\beta)d\beta\right| \le$$
$$\le \int_0^{\pi} |g(\beta) - g(0)|\Phi_n(\beta)d\beta + \int_0^{\pi} |g(2\pi-\beta) - g(2\pi)|\Phi_n(\beta)d\beta.$$

Повторяя предыдущие рассуждения, получаем, что эта сумма стремится к нулю. □

Следствие 4.1. Пусть g — непрерывная ограниченная функция и $\lambda \in [0,\pi]$. Тогда при $n \to \infty$

$$\int_{-\pi}^{\pi} \Phi_n(\lambda+\alpha)g(\alpha)d\alpha = \int_{-\pi+\lambda}^{\pi+\lambda} \Phi_n(\beta)g(\beta-\lambda)d\beta \to \begin{cases} g(-\lambda) & \text{при}\ \ \lambda \in [0,\pi), \\ g(\pi)/2 + g(-\pi) & \text{при}\ \ \lambda = \pi. \end{cases}$$

Доказательство. Это утверждение непосредственно следует из третьего пункта Леммы 4.1. □

4.2 Периодограммы и выборочные оценки спектральной плотности

Начнем со следующего определения.

Определение 4.1. Пусть $\mathbf{a} = (a_1, \dots, a_n)$ — некоторые вещественные числа. Тогда функция

$$I_{\mathbf{a}}(\lambda) = \frac{1}{2\pi n} \left| \sum_{m=1}^{n} a_m e^{-i\lambda m} \right|^2$$

называется *периодограммой* последовательности $\mathbf{a}$.

Замечание 4.1. 1. Иногда в определении периодограммы используют отличный от $(2\pi n)^{-1}$ нормирующий множитель.
2. Если положить $a_k = 1$ при всех k, то окажется, что $I_{\mathbf{a}}(\lambda) = \Phi_n(\lambda)$, где $\Phi_n(\lambda)$ — ядро Фейера.

Периодограмма любой последовательности $(a_1, \dots, a_n)$ обладает следующими свойствами.

Лемма 4.2. *1.*

$$I_{\mathbf{a}}(\lambda) = \frac{1}{2\pi} \left(\frac{1}{n} \sum_{k=1}^{n} a_k^2 + 2 \sum_{m=1}^{n-1} \frac{1}{n} \sum_{\ell=1}^{n-m} a_\ell \, a_{\ell+m} \cos(m\lambda) \right). \tag{4.1}$$

2. $I_{\mathbf{a}}$ является неотрицательной четной ограниченной непрерывной функцией, удовлетворяющей условию периодичности $I_{\mathbf{a}}(\lambda + 2\pi) = I_{\mathbf{a}}(\lambda)$.
3. Обозначим $I_{\mathbf{a}}^*$ периодограмму последовательности $\mathbf{a}^* = (a_1 - \bar{a}, \dots, a_n - \bar{a})$, где $\bar{a} = \sum_{j=1}^{n} a_j/n$. Пусть $\lambda = 2\pi k/n$, где $k/n \notin \mathbb{Z}$. Тогда $I_{\mathbf{a}}^*(\lambda) = I_{\mathbf{a}}(\lambda)$.

Доказательство. Первое утверждение доказывается непосредственно:

$$\frac{1}{2\pi n} \left| \sum_{m=1}^{n} a_m e^{-i\lambda m} \right|^2 = \frac{1}{2\pi n} \sum_{k,\ell=1}^{n} a_k a_\ell e^{-i(k-\ell)\lambda} = \frac{1}{2\pi n} \left(\sum_{k=1}^{n} a_k^2 + \sum_{k \neq \ell} a_k a_\ell e^{-i(k-\ell)\lambda} \right) =$$

$$= \frac{1}{2\pi n} \left(\sum_{k=1}^{n} a_k^2 + 2 \sum_{1 \le \ell < k \le n} a_\ell a_k \cos{(k-\ell)\lambda} \right) = \frac{1}{2\pi n} \left(\sum_{k=1}^{n} a_k^2 + 2 \sum_{\ell=1}^{n-1} \sum_{k-\ell=1}^{n-\ell} a_\ell a_{\ell+(k-\ell)} \cos{(k-\ell)\lambda} \right) =$$

$$= \frac{1}{2\pi n} \left(\sum_{k=1}^{n} a_k^2 + 2 \sum_{\ell=1}^{n-1} \sum_{m=1}^{n-\ell} a_\ell a_{\ell+m} \cos(m\lambda) \right) = \frac{1}{2\pi n} \left(\sum_{k=1}^{n} a_k^2 + 2 \sum_{m=1}^{n-1} \sum_{\ell=1}^{n-m} a_\ell a_{\ell+m} \cos(m\lambda) \right).$$

Второе не нуждается в специальном доказательстве. Докажем третье утверждение. При $\lambda \neq 2\pi\ell$ с $\ell \in \mathbb{Z}$

$$\sum_{m=1}^{n} e^{-i\lambda m} = \frac{1 - e^{-i\lambda n}}{1 - e^{-i\lambda}} \, e^{-i\lambda}.$$

Последнее выражение превращается в ноль, если дополнительно потребовать, чтобы $\lambda = 2\pi k/n$, $k \in \mathbb{Z}$. Поэтому в условиях Леммы

$$\sum_{m=1}^{n} (a_m - \bar{a}) e^{-i\lambda m} = \sum_{m=1}^{n} a_m e^{-i\lambda m} - \bar{a} \sum_{m=1}^{n} e^{-i\lambda m} = \sum_{m=1}^{n} a_m e^{-i\lambda m}.$$

Лемма доказана. □

Замечание 4.2. Если взять $\lambda = 0$, то окажется, что $I_{\mathbf{a}}(0) = \left(\sum_{m=1}^{n} a_n \right)^2 / (2\pi n)$, в то время как $I_{\mathbf{a}}^*(0) = 0$.

Перейдем теперь к оцениванию спектральной плотности. Согласно Предложению 1.1, при выполнении условия $\sum_m |R(m)| < \infty$

$$f(\lambda) = \frac{1}{2\pi} \sum_m e^{-im\lambda} R(m) = \frac{1}{2\pi} \sum_m \cos(\lambda m) R(m). \tag{4.2}$$

Естественно считать, что выборочным аналогом (4.2), построенным по реализации $x_1, \ldots, x_n$ стационарного процесса, является случайная функция

$$\widehat{f}(\lambda) = \frac{1}{2\pi} \sum_{m=-(n-1)}^{n-1} e^{-im\lambda} \widehat{R}(m) = \frac{1}{2\pi} \sum_{m=-(n-1)}^{n-1} \cos(\lambda m) \widehat{R}(m) = \frac{1}{2\pi} \left(\widehat{R}(0) + 2 \sum_{m=1}^{n-1} \cos(m\lambda) \widehat{R}(m) \right),$$

где $\widehat{R}(m)$ — подходящая оценка ковариационной функции. Для наших целей мы будем использовать два варианта оценок $\widehat{R}(m)$ — оценки (3.3) и (3.4). Иначе говоря (в предположении, что $\mathbb{E}x_1 = 0$) мы будем полагать

$$\widehat{R}(m) = \widetilde{R}_n^{(0)}(m) = \frac{1}{n} \sum_{j=1}^{n-m} x_j \, x_{j+m}$$

и

$$\widehat{R}(m) = \widetilde{R}_n^{(1)}(m) = \frac{1}{n} \sum_{j=1}^{n-m} (x_j - \bar{x}_n)(x_{j+m} - \bar{x}_n).$$

В первом случае вместо $\widehat{f}(\lambda)$ мы будем писать $I_n^{(0)}(\lambda)$, а во втором — $I_n^{(1)}(\lambda)$. Конечно, оценка $I_n^{(0)}(\lambda)$ более удобна для теоретических рассмотрений, в то время как $I_n^{(1)}(\lambda)$ — для практических применений.

Согласно равенству (4.1), $I_n^{(0)}(\lambda)$ является периодограммой последовательности $x_1, \ldots, x_n$, а $I_n^{(1)}(\lambda)$ — периодограммой последовательности $x_1 - \bar{x}, \ldots, x_n - \bar{x}_n$. Таким образом, мы получили два представления одних и тех же оценок спектральной плотности — через оценки ковариационных функций и через периодограммы.

4.3 Среднее значение периодограммы

Предложение 4.1. *Пусть $\lambda \in [0,\pi]$. Тогда*
1.

$$\mathbb{E}I_n^{(0)}(\lambda) = \frac{1}{2\pi}\sum_{r=-n+1}^{n-1}\left(1-\frac{|r|}{n}\right)R(r)\cos(\lambda r) = \frac{1}{2\pi}\left(R(0)+2\sum_{r=1}^{n-1}\left(1-\frac{r}{n}\right)R(r)\cos(\lambda r)\right). \tag{4.3}$$

2.

$$\mathbb{E}I_n^{(0)}(\lambda) = \int_{-\pi}^{\pi}\Phi_n(\lambda-\alpha)f(\alpha)d\alpha = \int_{-\pi+\lambda}^{\pi+\lambda}\Phi_n(\beta)f(\lambda-\beta)d\beta.$$

3. Если спектральная плотность f непрерывна в точке λ, то

$$\mathbb{E}I_n^{(0)}(\lambda) \to f(\lambda).$$

4. Если $\sum_r |r|\,|R(r)| < \infty$, то

$$n\left(\mathbb{E}I_n^{(0)}(\lambda) - f(\lambda)\right) \to -\frac{1}{\pi}\sum_{r=1}^{\infty} rR(r)\cos(\lambda r). \tag{4.4}$$

Доказательство. 1. Равенство следует из того, что, по определению

$$I_n^{(0)}(\lambda) = \frac{1}{2\pi}\sum_{r=-n+1}^{n-1}\cos(r\lambda)\widetilde{R}_n^{(0)}(r),$$

а $\mathbb{E}\widetilde{R}_n^{(0)}(r) = (1-|r|/n)R(r)$.
2.

$$2\pi n\mathbb{E}I_n^{(0)}(\lambda) = \mathbb{E}\left|\sum_{j=1}^{n}x_je^{-i\lambda j}\right|^2 = \sum_{k,l=1}^{n}\mathbb{E}x_kx_le^{-i(k-l)\lambda} = \sum_{k,l=1}^{n}R(k-l)e^{-i(k-l)\lambda} =$$

$$= \sum_{k,l=1}^{n}\int_{-\pi}^{\pi}e^{i(k-l)\alpha}f(\alpha)d\alpha\, e^{-i(k-l)\lambda} = \int_{-\pi}^{\pi}\sum_{k,l=1}^{n}e^{i(k-l)(\alpha-\lambda)}f(\alpha)d\alpha = \int_{-\pi}^{\pi}\left|\sum_{j=1}^{n}e^{-i(\lambda-\alpha)j}\right|^2 f(\alpha)d\alpha.$$

Так как $\Phi_n(\beta) = |\varphi_n(\beta)|^2$, то все доказано.
3. Обозначим $g_\lambda(\beta) = f(\lambda-\beta)$. Согласно Следствию 4.1, при $\lambda \in [0,\pi)$

$$\int_{-\pi+\lambda}^{\pi+\lambda}\Phi_n(\beta)g_\lambda(\beta)d\beta \to g_\lambda(0) = f(\lambda).$$

В то же время при $\lambda = \pi$

$$\int_0^{2\pi}\Phi_n(\beta)g_\pi(\beta)d\beta \to \big(g_\lambda(0)+g_\lambda(2\pi)\big)/2 = \big(f(\pi)+f(-\pi)\big)/2 = f(\pi)$$

так как $f(\pi) = f(-\pi)$.
4. Так как

$$f(\lambda) = \frac{1}{2\pi}\left(R(0)+2\sum_{r\geq 1}R(r)\cos(\lambda r)\right),$$

то из (4.3) следует, что

$$n\left(\mathbb{E}I_n^{(0)}(\lambda)-f(\lambda)\right)=-\frac{1}{\pi}\left(n\sum_{r=n}^{\infty}R(r)\cos(\lambda r)+\sum_{r=1}^{n-1}rR(r)\cos(\lambda r)\right)$$

Первая сумма стремится к нулю, а вторая — к правой части (4.4). □

О смещении оценки $I_n^{(1)}$. Имеет место следующее утверждение ([1, теор. 8.3.6 гл. 8 §8.3]).

Теорема 4.1. *Если спектральная плотность непрерывна в нуле и точке λ, то*

$$\lim n\big(\mathbb{E}I_n^{(1)}(\lambda)-\mathbb{E}I_n^{(0)}(\lambda)\big)\to 0.$$

Таким образом, асимптотические свойства несмещенности обеих оценок совпадают для непрерывных спектральных плотностей.

4.4 Предельное распределение периодограмм

4.4.1 Предельные ковариации периодограммы

Начнем с одной теоремы, которую приведем без доказательства. Пусть x_n — линейный процесс (1.3), где ε_j — независимые случайные величины, имеющие нулевое среднее, единичную дисперсию и конечный четвертый момент. Кроме того, пусть $\sum_j |c_j| < \infty$,

Теорема 4.2. *При $\lambda \in [0, \pi]$ и $n \to \infty$*

$$\mathbb{E}\big(I_n^{(0)}(\lambda) - f(\lambda)\big)^2 \to \begin{cases} 2f^2(\lambda) & \textit{при} \ \ \lambda = 0 \ \ \textit{или} \ \ \lambda = \pi, \\ f^2(\lambda) & \textit{при} \ \ \lambda \in (0, \pi). \end{cases}$$

При этом, если $\lambda_1, \lambda_2 \in [0, \pi]$ и $\lambda_1 \neq \lambda_2$, то при $n \to \infty$

$$\mathbb{E}\big(I_n^{(0)}(\lambda_1) - f(\lambda_1)\big)\big(I_n^{(0)}(\lambda_2) - f(\lambda_2)\big) \to 0.$$

Из теоремы 4.2 следует, что среднеквадратическое отклонение периодограммы, вообще говоря, не стремится к нулю для линейного процесса. Хотя из этого утверждения и не следует несостоятельность периодограммы как оценки спектральной плотности, такая несостоятельность становится весьма правдоподобной.

4.4.2 Результаты и обсуждения

Цель настоящего раздела — доказать следующую теорему.

Теорема 4.3. *Пусть x_n — линейный процесс, записанный в виде*

$$x_n = \sum_m c_m \, \varepsilon_{n-m},$$

где $\sum_m |c_m| < \infty$, и ε_n — последовательность независимых одинаково распределенных случайных величин такая, что $\mathbb{E}\varepsilon_n = 0$, $\mathbb{E}\varepsilon_n^2 = 1$. Тогда при $\lambda \in [0, \pi]$ конечномерные распределения процесса $I_n^{(0)}(\lambda)$ слабо сходятся к конечномерным распределениям случайного процесса $I(\lambda)$ такого, что
1. Если $\lambda \in (0, \pi)$, то случайная величина $\mathcal{L}\big(I(\lambda)\big) = \mathcal{L}\big(f(\lambda)\xi\big)$, где $\xi \in \mathrm{Exp}(1)$.
2. Если $\lambda = 0$ или π, то $\mathcal{L}\big(I(\lambda)\big) = \mathcal{L}\big(f(\lambda)\xi\big)$, где $\xi \in \chi^2(1)$.
3. Значения $I(\lambda)$ при различных λ независимы.

Замечание 4.3. 1. Результат Теоремы 4.3 в некотором смысле может служить объяснением результата Теоремы 4.2. Действительно,

$$\mathbb{E}\big(f(\lambda)\xi - f(\lambda)\big)^2 = f^2(\lambda)\,\mathbb{E}\big(\xi - 1\big)^2 = \begin{cases} f^2(\lambda) & \text{при} \ \ \xi \in \mathrm{Exp}(1), \\ 2f^2(\lambda) & \text{при} \ \ \xi \in \chi^2(1), \end{cases}$$

так как $\mathbb{E}(\xi - 1)^2 = 2$ при $\xi \in \chi^2(1)$, в то время как $\mathbb{E}\big(\xi - 1\big)^2 = 1$ для случая $\xi \in \mathrm{Exp}(1)$.
2. Из Теоремы 4.3 следует несостоятельность периодограммы как оценки спектральной плотности в тех точках λ, где $f(\lambda) > 0$. Действительно, при $n \to \infty$

$$\mathbb{P}\big(\big|I_n^{(0)}(\lambda) - f(\lambda)\big| > \delta\big) \to \mathbb{P}\big(\big|f(\lambda)\xi - f(\lambda)\big| > \delta\big) = \mathbb{P}\big(\big|\xi - 1\big| > \delta/f(\lambda)\big) > 0,$$

где $\xi \in \mathrm{Exp}(1)$ при $\lambda \in (0, \pi)$ и $\xi \in \chi^2(1)$ иначе.

Результат Теоремы 4.3 мы выведем из следующего утверждения, имеющего самостоятельный интерес. Обозначим

$$A_n(\lambda) = \frac{1}{n} \sum_{k=1}^{n} x_k \cos(k\lambda), \quad B_n(\lambda) = \frac{1}{n} \sum_{k=1}^{n} x_k \sin(k\lambda).$$

Ясно, что $\mathbb{E}\, A_n(\lambda) = \mathbb{E}\, B_n(\lambda) = 0$.

Теорема 4.4. *В условиях Теоремы 4.3 конечномерные распределения процесса $\sqrt{n}(A_n(\lambda), B_n(\lambda))$, где $\lambda \in [0,\pi]$, слабо сходятся к конечномерным распределениям двумерного гауссовского процесса $(Z_1(\lambda), Z_2(\lambda))$ такого, что*
1. Случайные процессы Z_1 и Z_2 независимы.
2. При $\lambda \in [0,\pi]$ процессы $Z_1(\lambda)$ и $Z_2(\lambda)$ являются процессами с независимыми значениями, причем $Z_2(0) = Z_2(\pi) = 0$.
3. $\mathbb{E}Z_1(\lambda) = \mathbb{E}Z_2(\lambda) = 0$ при всех λ.
4. $\mathbb{D}Z_1(\lambda) = \mathbb{D}Z_2(\lambda) = \pi f(\lambda)$ при $\lambda \in (0,\pi)$.
5. $\mathbb{D}Z_1(\lambda) = 2\pi f(\lambda)$ при $\lambda = 0$ или π.

Покажем, как из Теоремы 4.4 следует результат Теоремы 4.3.

Следствие 4.2. *Если для процессов $(A_n(\lambda), B_n(\lambda))$ имеет место предельный результат Теоремы 4.4, то для процессов $I_n^{(0)}(\lambda)$ имеет место предельный результат Теоремы 4.3.*

Доказательство. Заметим, что

$$I_n^{(0)}(\lambda) = \frac{1}{2\pi n}\left|\sum_{k=1}^{n} x_k e^{-ik\lambda}\right|^2 = \frac{n}{2\pi}\big(A_n^2(\lambda) + B_n^2(\lambda)\big) = \frac{1}{2\pi}\Big(\big(\sqrt{n}A_n(\lambda)\big)^2 + \big(\sqrt{n}B_n(\lambda)\big)^2\Big). \tag{4.5}$$

Отсюда сразу же следует слабая сходимость конечномерных распределений процесса $I_n^{(0)}(\lambda)$ к конечномерным распределениям процесса $I(\lambda) = \big(Z_1^2(\lambda) + Z_2^2(\lambda)\big)/2\pi$. Конечно, этот процесс имеет независимые значения.

Если $\lambda \in (0,\pi)$, то

$$\mathcal{L}\big(I(\lambda)\big) = \mathcal{L}\left(\frac{1}{2\pi}\big(\pi f(\lambda)\xi_1^2 + \pi f(\lambda)\xi_2^2\big)\right) = \mathcal{L}\Big(f(\lambda)\big(\xi_1^2 + \xi_2^2\big)/2\Big),$$

где $\xi_1, \xi_2 \in \mathrm{N}(0,1)$ и независимы. Так как $\big(\xi_1^2 + \xi_2^2\big)/2 \in \mathrm{Exp}(1)$, то первое утверждение Теоремы 4.3 доказано.

Пусть теперь $\lambda = 0$ или $\lambda = \pi$. Так как $Z_2(\lambda) = 0$ при этих λ, то

$$\mathcal{L}\big(I(\lambda)\big) = \mathcal{L}\left(\frac{1}{2\pi}\,2\pi f(\lambda)\xi_1^2\right) = \mathcal{L}\Big(f(\lambda)\xi_1^2\Big),$$

что и требовалось доказать. □

4.4.3 Доказательство Теоремы 4.4

Начнем с доказательства нескольких лемм.

1. Вспомогательные утверждения.

Лемма 4.3. *1. Пусть $\omega, \omega_1, \omega_2 \in (0, 1/2)$.*
Обозначим $\cos_j(\gamma) = \cos(2\pi\omega j + \gamma)$, $\sin_j(\gamma) = \sin(2\pi\omega j + \gamma)$, $\cos_{j\,1}(\gamma) = \cos(2\pi\omega_1 j + \gamma)$, $\cos_{j\,2}(\gamma) = \cos(2\pi\omega_2 j + \gamma)$, $\sin_{j\,1}(\gamma) = \sin(2\pi\omega_1 j + \gamma)$ и $\sin_{j\,2}(\gamma) = \sin(2\pi\omega_2 j + \gamma)$.

Тогда

$$\sum_{j=0}^{n-1} \cos_j(\gamma) = \frac{\sin(\pi\omega n)}{\sin(\pi\omega)}\,\cos(\pi\omega(n-1) + \gamma),$$

$$\sum_{j=0}^{n-1} \sin_j(\gamma) = \frac{\sin(\pi\omega n)}{\sin(\pi\omega)}\,\sin(\pi\omega(n-1) + \gamma),$$

$$\sum_{j=0}^{n-1} \cos_{j\,1}(\gamma_1)\cos_{j\,2}(\gamma_2) = \frac{\sin(\pi(\omega_1+\omega_2)n)}{2\sin(\pi(\omega_1+\omega_2))}\,\cos\big(\pi(\omega_1+\omega_2)(n-1)+\gamma_1+\gamma_2\big)$$
$$+\frac{1}{2}\begin{cases}\dfrac{\sin(\pi(\omega_1-\omega_2)n)}{\sin(\pi(\omega_1-\omega_2))}\,\cos\big(\pi(\omega_1-\omega_2)(n-1)+\gamma_1-\gamma_2\big) & \text{при } \omega_1\neq\omega_2,\\ n\cos(\gamma_1-\gamma_2) & \text{при } \omega_1=\omega_2,\end{cases}$$

$$\sum_{j=0}^{n-1} \sin_{j\,1}(\gamma_1)\sin_{j\,2}(\gamma_2) = -\frac{\sin(\pi(\omega_1+\omega_2)n)}{2\sin(\pi(\omega_1+\omega_2))}\,\cos\big(\pi(\omega_1+\omega_2)(n-1)+\gamma_1+\gamma_2\big)$$
$$+\frac{1}{2}\begin{cases}\dfrac{\sin(\pi(\omega_1-\omega_2)n)}{\sin(\pi(\omega_1-\omega_2))}\,\cos\big(\pi(\omega_1-\omega_2)(n-1)+\gamma_1-\gamma_2\big) & \text{при } \omega_1\neq\omega_2,\\ n\cos(\gamma_1-\gamma_2) & \text{при } \omega_1=\omega_2,\end{cases}$$

и

$$\sum_{j=0}^{n-1} \cos_{j\,1}(\gamma_1)\sin_{j\,2}(\gamma_2) = \frac{\sin(\pi(\omega_1+\omega_2)n)}{2\sin(\pi(\omega_1+\omega_2))}\,\sin\big(\pi(\omega_1+\omega_2)(n-1)+\gamma_1+\gamma_2\big)$$
$$-\frac{1}{2}\begin{cases}\dfrac{\sin(\pi(\omega_1-\omega_2)n)}{\sin(\pi(\omega_1-\omega_2))}\,\sin\big(\pi(\omega_1-\omega_2)(n-1)+\gamma_1-\gamma_2\big) & \text{при } \omega_1\neq\omega_2,\\ n\sin(\gamma_1-\gamma_2) & \text{при } \omega_1=\omega_2,\end{cases}$$

Доказательство. Утверждение леммы легко доказывается представлением синуса и косинуса в виде суммы мнимых экспонент. □

Следствие 4.3. 1. Для $\lambda_1, \lambda_2 \in [0,\pi]$ при $n \to \infty$

$$\lim \frac{1}{n}\sum_{\ell=1}^{n} \sin(\lambda_1\ell)\sin(\lambda_2\ell) = \begin{cases}0 & \text{при } \lambda_1=\lambda_2=0 \text{ или } \pi,\\ 1/2 & \text{при } \lambda_1=\lambda_2\in(0,\pi),\\ 0 & \text{при } \lambda_1\neq\lambda_2;\end{cases}$$

$$\lim \frac{1}{n}\sum_{\ell=1}^{n} \cos(\lambda_1\ell)\cos(\lambda_2\ell) = \begin{cases}1 & \text{при } \lambda_1=\lambda_2=0 \text{ или } \pi,\\ 1/2 & \text{при } \lambda_1=\lambda_2\in(0,\pi),\\ 0 & \text{при } \lambda_1\neq\lambda_2,\end{cases}$$

а также

$$\lim \frac{1}{n}\sum_{\ell=1}^{n} \cos(\lambda_1\ell)\sin(\lambda_2\ell) = 0.$$

2. Если $\lambda_1 = 2\pi k_1/n$, $\lambda_2 = 2\pi k_2/n$ и $\lambda_1, \lambda_2 \in [0,\pi]$, то

$$\frac{1}{n}\sum_{\ell=1}^{n} \sin(\lambda_1\ell)\sin(\lambda_2\ell) = \begin{cases}0 & \text{при } \lambda_1=\lambda_2=0 \text{ или } \pi,\\ 1/2 & \text{при } \lambda_1=\lambda_2\in(0,\pi),\\ 0 & \text{при } \lambda_1\neq\lambda_2;\end{cases}$$

$$\frac{1}{n}\sum_{\ell=1}^{n} \cos(\lambda_1\ell)\cos(\lambda_2\ell) = \begin{cases}1 & \text{при } \lambda_1=\lambda_2=0 \text{ или } \pi,\\ 1/2 & \text{при } \lambda_1=\lambda_2\in(0,\pi),\\ 0 & \text{при } \lambda_1\neq\lambda_2,\end{cases}$$

а также

$$\frac{1}{n}\sum_{\ell=1}^{n} \cos(\lambda_1\ell)\sin(\lambda_2\ell) = 0.$$

Доказательство. 1. Если $\lambda_1, \lambda_2 = 0$ или π, то утверждение становится очевидным. Для рассмотрения случая $\lambda_1, \lambda_2 \in (0, \pi)$ достаточно перейти в к обозначениям $\lambda_1 = 2\pi\omega_1$ и $\lambda_2 = 2\pi\omega_2$, а также положить $\gamma_1 = \gamma_2 = 0$.

2. Если $\lambda_1 = 2\pi k_1/n$ и $\lambda_2 = 2\pi k_2/n$, то

$$\sin\big(\pi(\omega_1 \pm \omega_2)n\big) = \sin\big((\lambda_1 \pm \lambda_2)n/2\big) = \sin\big(\pi(k_1 \pm k_2)\big) = 0,$$

откуда немедленно вытекает требуемое. □

Лемма 4.4. *Пусть $\varepsilon_1, \ldots, \varepsilon_n, \ldots$ — независимые одинаково распределенные случайные величины с $\mathbb{E}\varepsilon_n = 0$ и $\mathbb{D}\varepsilon_n = 1$. Рассмотрим последовательность a_n и положим $\xi_n = a_n\varepsilon_n$, а также $B_n^2 = a_1^2 + \ldots + a_n^2$. Если $|a_n| \le c$ и $n/B_n^2 = O(1)$, то последовательность ξ_n подчиняется Центральной Предельной Теореме, то есть имеет место сходимость*

$$\mathcal{L}\big((\xi_1 + \ldots + \xi_n)/B_n\big) \Rightarrow \mathrm{N}(0, 1). \tag{4.6}$$

Если дополнительно $B_n^2 \sim \sigma^2 n$, то (4.6) *приобретет вид*

$$\mathcal{L}\big((\xi_1 + \ldots + \xi_n)/\sqrt{n}\big) \Rightarrow \mathrm{N}(0, \sigma^2). \tag{4.7}$$

Доказательство. Докажем, что последовательность ξ_n удовлетворяет условию Линдеберга, то есть что для любого $\theta > 0$ при $n \to \infty$

$$\frac{1}{B_n^2}\sum_{\ell=1}^{n} \mathbb{E}(\xi_\ell^2,\, |\xi_\ell| > \theta B_n) \to 0. \tag{4.8}$$

Так как $B_n \to \infty$, то

$$\mathbb{E}(\xi_\ell^2,\, |\xi_\ell| > \theta B_n) \le c^2\mathbb{E}(\varepsilon_\ell^2,\, c|\varepsilon_\ell| > \theta B_n) = c^2\mathbb{E}(\varepsilon^2,\, |\varepsilon| > \theta B_n/c) \to 0.$$

Поэтому

$$\frac{1}{B_n^2}\sum_{\ell=1}^{n} \mathbb{E}\left(\xi_\ell^2,\, |\xi_\ell| > \theta B_n\right) \le \frac{n}{B_n^2}\; c^2\mathbb{E}\left(\varepsilon^2,\, |\varepsilon| > \theta B_n/c\right) \to 0,$$

что и требовалось доказать. □

Лемма 4.5. *Пусть $\overline{\xi}_n = (\xi_{1,n}, \ldots, \xi_{d,n})^{\mathrm{T}}$ — последовательность d-мерных случайных векторов. Для того, чтобы последовательность $\mathcal{L}(\overline{\xi}_n)$ слабо сходилась к d-мерному гауссовскому распределению $\mathrm{N}(\overline{a}, \Sigma)$, необходимо и достаточно, чтобы для любых $\overline{t} = (t_1, \ldots, t_d)^{\mathrm{T}}$ последовательность $\mathcal{L}\big(\sum_{k=1}^{d} \xi_{k,n}t_k\big)$ слабо сходилась к одномерному нормальному распределению $\mathrm{N}\Big((\overline{a}, \overline{t}), (\Sigma\overline{t}, \overline{t})\Big)$.*

Доказательство. 1. Необходимость. Обозначим $\overline{t} = (t_1, \ldots, t_d)^{\mathrm{T}}$. Если $\mathcal{L}(\overline{\xi}_n) \Rightarrow \mathrm{N}(\overline{a}, \Sigma)$, то по свойствам многомерного гауссовского распределения

$$\mathcal{L}\Big(\sum_{k=1}^{d} \xi_{k,n}t_n\Big) \Rightarrow \mathrm{N}\Big((\overline{a}, \overline{t}), (\Sigma\overline{t}, \overline{t})\Big).$$

2. Достаточность. Из условий следует, что для любого $\tau \in \mathbb{R}$

$$\mathbb{E}e^{i\tau(\overline{\xi}_n, \overline{t})} \to e^{\tau(\overline{a}, \overline{t})}\, e^{-\tau^2(\Sigma\overline{t}, \overline{t})/2}\ .$$

Полагая $\tau = 1$, получаем нужный результат. □

2. Случай гауссовского белого шума.
Теперь убедимся, что Теорема 4.4 верна в случае, когда x_n является гауссовским белым шумом.

Предложение 4.2. *Пусть $x_n = \varepsilon_n$ — гауссовский белый шум с нулевым средним и единичной дисперсией. Тогда конечномерные распределения двумерного процесса $\sqrt{n}(A_n(\lambda), B_n(\lambda))$, $\lambda \in [0, \pi]$, слабо сходятся к конечномерным распределениям двумерного гауссовского процесса $(Z_1(\lambda), Z_2(\lambda))$ такого, что*
1. Z_1 и Z_2 — независимы.
2. При $\lambda \in [0, \pi]$ процессы $Z_1(\lambda)$ и $Z_2(\lambda)$ имеют независимые значения, причем $Z_2(0) = Z_2(\pi) = 0$.
3. $\mathbb{E}Z_1(\lambda) = \mathbb{E}Z_2(\lambda) = 0$ при всех λ.
4. $\mathbb{D}Z_1(\lambda) = \mathbb{D}Z_2(\lambda) = 1/2$ при $\lambda \in (0, \pi)$, $\mathbb{D}Z_1(\lambda) = 1$ при $\lambda \in \{0, \pi\}$.

Доказательство. Прежде всего заметим, что в нашем случае $f(\lambda) = 1/(2\pi)$, так что Предложение 4.2 действительно является частным случаем Теоремы 4.4.

Пусть $0 = \lambda_0 < \lambda_1 < \ldots < \lambda_m = \pi$. Случайный вектор

$$\sqrt{n}\big(A_n(\lambda_0), B_n(\lambda_0), \ldots, A_n(\lambda_j), B_n(\lambda_j), \ldots, A_n(\lambda_m), B_n(\lambda_m)\big), \tag{4.9}$$

очевидно, имеет гауссовское распределение с нулевым средним, ковариационная матрица которого определяется величинами вида $n\mathbb{E}\big(A_n(\lambda_j)A_n(\lambda_k)\big)$, $n\mathbb{E}\big(B_n(\lambda_j)B_n(\lambda_k)\big)$ и $n\mathbb{E}\big(A_n(\lambda_j)B_n(\lambda_k)\big)$. Поскольку x_n — белый шум, то пределы этих величин определены в первом пункте Следствия 4.3. Действительно,

$$n\mathbb{E}\big(A_n(\lambda_j)A_n(\lambda_k)\big) = \frac{1}{n}\ \mathbb{E}\left(\sum_{m=1}^{n} \varepsilon_m \cos(m\lambda_j)\right)\left(\sum_{\ell=1}^{n} \varepsilon_\ell \cos(\ell\lambda_k)\right) = \frac{1}{n}\ \sum_{m=1}^{n} \cos(m\lambda_j)\cos(m\lambda_k).$$

Аналогично

$$n\mathbb{E}\big(B_n(\lambda_j)B_n(\lambda_k)\big) = \frac{1}{n}\ \sum_{m=1}^{n} \sin(m\lambda_j)\sin(m\lambda_k) \quad \text{и} \quad n\mathbb{E}\big(A_n(\lambda_j)B_n(\lambda_k)\big) = \frac{1}{n}\ \sum_{m=1}^{n} \cos(m\lambda_j)\sin(m\lambda_k).$$

Используя Следствие 4.3, приходим к нужному результату. □

Замечание 4.4. Если все рассматриваемые λ_j имеют вид $\lambda_j = 2\pi k_j/n$, то, согласно второму пункту Следствия 4.3, при любом достаточно большом n

$$n\mathbb{E}\big(A_n(\lambda_j)A_n(\lambda_k)\big) = \frac{1}{n}\ \sum_{m=1}^{n} \cos(m\lambda_j)\cos(m\lambda_k) = \begin{cases} 1 & \text{при } \lambda_j = \lambda_k = 0 \text{ или } \pi, \\ 1/2 & \text{при } \lambda_j = \lambda_k \in (0, \pi), \\ 0 & \text{при } \lambda_j \neq \lambda_k. \end{cases}$$

Аналогично

$$n\mathbb{E}\big(B_n(\lambda_j)B_n(\lambda_k)\big) = \begin{cases} 0 & \text{при } \lambda_j = \lambda_k = 0 \text{ или } \pi, \\ 1/2 & \text{при } \lambda_j = \lambda_k \in (0, \pi), \\ 0 & \text{при } \lambda_j \neq \lambda_k \end{cases}$$

и $n\mathbb{E}\big(A_n(\lambda_j)B_n(\lambda_k)\big) = 0$.

Поэтому случайный вектор (4.9) имеет[4] гауссовское распределение с нулевым средним, причем компоненты этого вектора независимы, $B_n(\lambda_0) = B_n(\lambda_m) = 0$, $\mathbb{D}A_n(\lambda_0) = \mathbb{D}A_n(\lambda_m) = 1$, а дисперсии всех остальных компонент вектора равны 1/2.

[4] без всякого предельного перехода

При этом конкретный выбор $\lambda_1, \ldots, \lambda_{m-1}$ не важен, лишь бы они имели вид $\lambda_j = 2\pi k_j/n \in (0, \pi)$ при различных k_j. В частности, если n будет меняться, то числа k_j тоже могут меняться в зависимости от n.

Кроме того, и число m может зависеть от n. В частности, если положить $k_j = j$ при $1 \le j < m$, то можно при нечетном n взять $m - 1 = \lfloor n/2 \rfloor$. Тогда $\lambda_{m-1} = 2\pi(m-1)/n = 2\pi\lfloor n/2\rfloor/n < \pi$. При четном n нужно брать $m = n/2$.

3. Случай произвольного белого шума.
Следующий шаг — доказательство Теоремы 4.4 для произвольного белого шума.

Предложение 4.3. *Утверждение Предложения 4.2 остается верным, если $x_n = \varepsilon_n$, где ε_n — независимые одинаково распределенные случайные величины с $\mathbb{E}x_n = 0$ и $\mathbb{D}x_n = 1$.*

Доказательство. Рассмотрим сначала одномерный случай на примере последовательности $A_n(\lambda)$. Согласно Следствию 4.3 и Лемме 4.4, последовательность $\xi_\ell = \cos(\ell\lambda)\varepsilon_\ell$ удовлетворяет условию Линдеберга с $B_n^2 \sim \sigma^2$, где $\sigma^2 = 1$ при $\lambda = 0$ или π и $\sigma^2 = 1/2$ при $\lambda \in (0, \pi)$.

Отсюда сразу же следует требуемая нормальность (с параметрами 0 и σ^2) предельного распределения последовательности $\sqrt{n}A_n(\lambda)$. Аналогичные рассуждения верны и для $\sqrt{n}B_n(\lambda)$.

Перейдем теперь к общему случаю. Согласно Лемме 4.5, для того, чтобы доказать слабую сходимость конечномерных распределений двумерного процесса $\sqrt{n}\big(A_n(\lambda), B_n(\lambda)\big)$ к нужному многомерному гауссовском распределению, достаточно доказать для произвольного набора частот $0 = \lambda_0 < \lambda_1 < \ldots < \lambda_{m-1} < \lambda_m = \pi$ и для произвольных вещественных чисел $u_0, \ldots, u_m$, $v_0, \ldots, v_m$ предельную нормальность последовательности случайных величин

$$\eta_n = \sqrt{n}\big(A_n(\lambda_0)u_0 + B_n(\lambda_0)v_0 + \ldots + A_n(\lambda_j)u_j + B_n(\lambda_j)v_j + \ldots + A_n(\lambda_m)u_m + B_n(\lambda_m)v_m\big) = \\ = \frac{1}{\sqrt{n}}\sum_{\ell=1}^{n}\left(\sum_{k=0}^{m}\Big(\cos(\ell\lambda_k)u_k + \sin(\ell\lambda_k)v_k\Big)\right)\varepsilon_\ell,$$

а также проверить, что среднее и дисперсия предельного нормального распределения такие, как нужно.[5]

Конечно, проверка предельной нормальности сводится к проверке условия Линдеберга для последовательности случайных величин

$$\xi_\ell = a_\ell\,\varepsilon_\ell = \left(\sum_{k=0}^{m}\big(\cos(\ell\lambda_k)u_k + \sin(\ell\lambda_k)v_k\big)\right)\varepsilon_\ell.$$

Очевидно, $|a_\ell| \le c = \sum_{k=0}^{m}(|u_k| + |v_k|)$. Найдем теперь предельное поведение отношения B_n^2/n, ограничивая для простоты случаем $u_0 = v_0 = u_m = v_m = 0$, то есть рассматривая только $\lambda \in (0, \pi)$.

Тогда согласно Следствию 4.3

$$B_n^2/n = \frac{1}{n}\sum_{\ell=1}^{n}\left(\sum_{k=0}^{m}\big(\cos(\ell\lambda_k)u_k + \sin(\ell\lambda_k)v_k\big)\right)^2 = \sum_{k=1}^{m-1}u_k^2\,\frac{1}{n}\sum_{\ell=1}^{n}\cos^2(\ell\lambda_k) + \\ + \sum_{k=1}^{m-1}v_k^2\,\frac{1}{n}\sum_{\ell=1}^{n}\sin^2(\ell\lambda_k) + 2\sum_{p,r=1}^{m-1}u_p v_r\,\frac{1}{n}\sum_{\ell=1}^{n}\cos(\ell\lambda_p)\sin(\ell\lambda_r) + \sum_{p\ne r}^{m-1}u_p v_r\,\frac{1}{n}\sum_{\ell=1}^{n}\cos(\ell\lambda_p)\cos(\ell\lambda_r) + \\ + \sum_{p\ne r}^{m-1}u_p v_r\,\frac{1}{n}\sum_{\ell=1}^{n}\sin(\ell\lambda_p)\sin(\ell\lambda_r) \to \Delta^2 = \sum_{k=1}^{m-1}(u_k^2 + v_k^2)/2 \qquad (4.10)$$

[5] Со средними здесь все в порядке.

при $n \to \infty$. Согласно теореме Линдеберга это означает, что $\mathcal{L}(\eta_n) \Longrightarrow \mathrm{N}(0, \Delta^2)$. Поскольку переменные u_k, v_k произвольны, то по Лемме 4.5 случайный вектор

$$\sqrt{n}\Big(A_n(\lambda_1), B_n(\lambda_1), \ldots, A_n(\lambda_j), B_n(\lambda_j), \ldots, A_n(\lambda_{m-1}), B_n(\lambda_{m-1})\Big)^{\mathrm{T}}$$

имеет в пределе гауссовское распределение с нулевым средним и диагональной ковариационной матрицей, на диагонали которой стоят дисперсии, равные 1/2. Ровно это (для $\lambda \in (0,\pi)$) и утверждается в доказываемом варианте Теоремы 4.4.

Случай, когда $u_k \neq 0$ и/или $v_k \neq 0$ при $k=0$ и/или $k=m$ абсолютно аналогичен. □

Замечание 4.5. Снова обсудим ситуацию, когда вместо фиксированных λ_j берутся, вообще говоря, зависящие от n частоты вида $\lambda_j = 2\pi k_j/n$. Нетрудно видеть, что все предыдущие рассуждения (и, следовательно, результат) остаются в силе (например, вместо сходимости (4.10) мы получим просто равенство $B_n^2/n = \sum_{k=1}^{m-1}(u_k^2+v_k^2)/2$).

4. Общий случай

Общий случай Теоремы 4.4 будет следовать из следующего предложения.

Предложение 4.4. *Рассмотрим линейный процесс x_n, записанный в виде*

$$x_n = \sum_m c_m\, \varepsilon_{n-m},$$

где $\sum_m |c_m| < \infty$, и ε_n — последовательность независимых одинаково распределенных случайных величин такая, что $\mathbb{E}\varepsilon_n = 0$, $\mathbb{E}\varepsilon_n^2 = 1$. Обозначим

$$J_{n,x}(\lambda) = \frac{1}{\sqrt{n}} \sum_{\ell=1}^{n} x_\ell\, e^{-i\ell\lambda} \quad \text{и} \quad J_{n,\varepsilon}(\lambda) = \frac{1}{\sqrt{n}} \sum_{\ell=1}^{n} \varepsilon_\ell\, e^{-i\ell\lambda}.$$

Тогда

$$J_{n,x}(\lambda) = \left(\sum_m c_m e^{-im\lambda}\right) J_{n,\varepsilon}(\lambda) + V_n(\lambda) \tag{4.11}$$

где $|V_n(\lambda)|$ равномерно по λ стремится по вероятности к нулю при $n \to \infty$.

Доказательство. Положим при $m = 1, \ldots, n$

$$U_{n,m}(\lambda) = -\sum_{\ell=1}^{n} \varepsilon_\ell\, e^{-i\ell\lambda} + \sum_{\ell=1-m}^{n-m} \varepsilon_\ell\, e^{-i\ell\lambda} \tag{4.12}$$

и

$$V_n(\lambda) = \frac{1}{\sqrt{n}} \sum_m c_m e^{-im\lambda}\, U_{n,m}(\lambda).$$

Тогда, делая замену $\ell = j - m$, получим, что

$$\begin{aligned}
J_{n,x}(\lambda) &= \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \left(\sum_m c_m\, \varepsilon_{j-m}\right) e^{-ij\lambda} = \frac{1}{\sqrt{n}} \sum_m c_m \sum_{j=1}^{n} \varepsilon_{j-m} e^{-ij\lambda} = \\
&= \frac{1}{\sqrt{n}} \sum_m c_m \sum_{\ell=1-m}^{n-m} \varepsilon_\ell\, e^{-i(m+l)\lambda} = \sum_m c_m e^{-im\lambda} \frac{1}{\sqrt{n}} \sum_{\ell=1-m}^{n-m} \varepsilon_\ell\, e^{-\ell\lambda} = \\
&= \left(\sum_m c_m e^{-im\lambda}\right)\left(\frac{1}{\sqrt{n}} \sum_{\ell=1}^{n} \varepsilon_\ell\, e^{-i\ell\lambda}\right) + \frac{1}{\sqrt{n}} \sum_m c_m e^{-im\lambda}\, U_{n,m}(\lambda) = \\
&= \left(\sum_m c_m e^{-im\lambda}\right) J_{n,\varepsilon}(\lambda) + V_n(\lambda).
\end{aligned} \tag{4.13}$$

Нетрудно видеть, что после сокращения подобных членов в сумме (4.12) останется ровно $2|m|$ независимых слагаемых при $|m| < n$ и ровно $2n$ независимых слагаемых при $|m| \geq n$. Поэтому

$$\mathbb{E}\big|U_{n,m}(\lambda)\big|^2 \leq 2\min(|m|, n).$$

Далее, зафиксировав некоторое n_0 и взяв $n > n_0$, получим, что

$$\mathbb{E}\big|V_n(\lambda)\big| \leq \frac{1}{\sqrt{n}} \sum_m |c_m| \, \mathbb{E}\big|U_{n,m}(\lambda)\big| \leq \frac{1}{\sqrt{n}} \sum_m |c_m| \left(\mathbb{E}\big|U_{n,m}(\lambda)\big|^2\right)^{1/2} \leq$$
$$\leq \frac{\sqrt{2}}{\sqrt{n}} \sum_m |c_m| \left(\min(|m|, n)\right)^{1/2} \leq \frac{\sqrt{2}}{\sqrt{n}} \sum_{|m|<n_0} |c_m| \, |m|^{1/2} + \sqrt{2} \sum_{|m|\geq n_0} |c_m|.$$

Поэтому при фиксированном n_0

$$\limsup_n \mathbb{E}\big|V_n(\lambda)\big| \leq \sqrt{2} \sum_{|m|\geq n_0} |c_m|.$$

Так как n_0 произвольно, то $\mathbb{E}\big|V_n(\lambda)\big| \to 0$ равномерно по λ. Следовательно, $V_n(\lambda)$ стремится к нулю по вероятности равномерно по λ. □

Замечание 4.6. Поскольку сходимость $V_n(\lambda)$ по вероятности к нулю является равномерной по λ, то

$$J_{n,x}(\lambda_j) = \left(\sum_m c_m e^{-im\lambda_j}\right) J_{n,\varepsilon}(\lambda_j) + V_n(j) \tag{4.14}$$

где $\lambda_j = 2\pi j/n \in [0, \pi]$, а $V_n(j)$ сходится по вероятности к нулю равномерно по j.

Замечание 4.7. Из предложения 4.4 можно сразу получить результат Теоремы 4.3, если заметить, что $I_n^{(0)}(n) = |J_n(\lambda)|^2/2\pi$ и $\big|\sum_m c_m e^{-im\lambda}\big|^2 = 2\pi f(\lambda)$. Поэтому

$$I_{n,x}^{(0)}(\lambda) = f(\lambda) I_{n,\varepsilon}^{(0)}(\lambda) + W_n(\lambda), \tag{4.15}$$

где $W_n(\lambda)$ — некоторый остаток. Предельное поведение $I_{n,\varepsilon}^{(0)}(\lambda)$ нам уже известно, так как Теорема 4.3 уже доказана для белого шума. В частности, при различных λ случайные величины $I_{n,\varepsilon}^{(0)}(\lambda)$ асимптотически независимы, а при $\lambda \in (0, \pi)$ распределение $I_{n,\varepsilon}^{(0)}(\lambda)$ слабо сходится к $\mathrm{Exp}(1)$.

Кроме того, можно доказать (это фактически уже сделано при доказательстве Предложения 4.4), что $\sup_\lambda \mathbb{E}|W_n(\lambda)| \to 0$ и, следовательно, $W_n(\lambda) \to 0$ по вероятности равномерно по λ.

Отсюда сразу же будет следовать общий результат Теоремы 4.3. Более того, так как $\mathbb{E}|W_n(\lambda)| \to 0$ равномерно по λ, то, используя Замечание 4.5, мы получим следующий результат.

Теорема 4.5. *Пусть в условиях Теоремы 4.3 спектральная плотность f является положительной. Тогда любого $r > 1$ и любых $k_j = k_j(n)$ таких, что $1 \leq k_1 < k_2 < \ldots < k_r < n/2$ предельное распределение случайного вектора*

$$\left(I_n^{(0)}\big(\lambda_{k_1}^{(n)}\big)/f\big(\lambda_{k_r}^{(n)}\big), \ldots, I_n^{(0)}\big(\lambda_{k_1}^{(n)}\big)/f\big(\lambda_{k_r}^{(n)}\big)\right)$$

совпадает с совместным распределением r независимых случайных величин, имеющих распределение $\mathrm{EXP}(1)$.

Замечание 4.8. Утверждение Теоремы 4.5 сохранится, если вместо периодограммы «теоретической» $I_n^{(0)}\big(\lambda\big)$ рассматривать «реальную» периодограмму $I_n^{(1)}\big(\lambda\big)$.

Наш план, однако включает доказательство Теоремы 4.4, имеющей самостоятельный интерес.

Следствие 4.4. (*Доказательство Теоремы 4.4*)
Если для комплекснозначного процесса $J_{n,x}(\lambda)$ имеет место результат Предложения 4.4, то для процессов

$$A_n(\lambda) = A_{n,x}(\lambda) = \frac{1}{n}\sum_{k=1}^{n} x_k \cos(k\lambda), \quad B_n(\lambda) = B_{n,x}(\lambda) = \frac{1}{n}\sum_{k=1}^{n} x_k \sin(k\lambda)$$

имеет место результат Теоремы 4.4.

Доказательство. Обозначим

$$A_{n,\varepsilon}(\lambda) = \frac{1}{n}\sum_{k=1}^{n} \varepsilon_k \cos(k\lambda), \quad B_{n,\varepsilon}(\lambda) = \frac{1}{n}\sum_{k=1}^{n} \varepsilon_k \sin(k\lambda).$$

Поскольку $\mathrm{Re}J_{n,x}(\lambda) = \sqrt{n}A_{n,x}(\lambda)$, $\mathrm{Im}J_{n,x}(\lambda) = -\sqrt{n}B_{n,x}(\lambda)$ и аналогичные равенства имеют место для $\mathrm{Re}J_{n,\varepsilon}(\lambda)$ и $\mathrm{Im}J_{n,\varepsilon}(\lambda)$, то равенство (4.11) переписывается в виде

$$\sqrt{n}A_{n,x}(\lambda) = \left(\sum_m c_m \cos(\lambda m)\right)\sqrt{n}A_{n,\varepsilon}(\lambda) - \left(\sum_m c_m \sin(\lambda m)\right)\sqrt{n}B_{n,\varepsilon}(\lambda) + V_n^{(1)}(\lambda),$$
$$\sqrt{n}B_{n,x}(\lambda) = \left(\sum_m c_m \sin(\lambda m)\right)\sqrt{n}A_{n,\varepsilon}(\lambda) + \left(\sum_m c_m \cos(\lambda m)\right)\sqrt{n}B_{n,\varepsilon}(\lambda) + V_n^{(2)}(\lambda),$$

где $V_n^{(1)}(\lambda)$ и $V_n^{(2)}(\lambda)$ сходятся по вероятности к нулю равномерно по λ.

Используя результат Предложения 4.3, мы сразу же видим, что конечномерные распределения двумерных процессов $\sqrt{n}\big(A_{n,x}(\lambda), B_{n,x}(\lambda)\big)$ слабо сходятся к конечномерным распределениям двумерного процесса $\big(Z_{1,x}(\lambda), Z_{2,x}(\lambda)\big)$, где

$$Z_{1,x}(\lambda) = \left(\sum_m c_m \cos(\lambda m)\right) Z_{1,\varepsilon}(\lambda) - \left(\sum_m c_m \sin(\lambda m)\right) Z_{2,\varepsilon}(\lambda),$$
$$Z_{2,x}(\lambda) = \left(\sum_m c_m \sin(\lambda m)\right) Z_{1,\varepsilon}(\lambda) + \left(\sum_m c_m \cos(\lambda m)\right) Z_{2,\varepsilon}(\lambda),$$

а двумерный гауссовский процесс $\big(Z_{1,\varepsilon}(\lambda), Z_{2,\varepsilon}(\lambda)\big)$ обладает следующими свойствами:
1. Процессы $Z_{1,\varepsilon}$ и $Z_{2,\varepsilon}$ — независимы;
2. При $\lambda \in [0,\pi]$ процессы $Z_{1,\varepsilon}$ и $Z_{2,\varepsilon}$ имеют независимые значения, причем $Z_{2,\varepsilon}(0) = Z_{2,\varepsilon}(\pi) = 0$;
3. $\mathbb{E}Z_{1,\varepsilon}(\lambda) = \mathbb{E}Z_{2,\varepsilon}(\lambda) = 0$ при всех λ;
4. $\mathbb{D}Z_{1,\varepsilon}(\lambda) = \mathbb{D}Z_{2,\varepsilon}(\lambda) = 1/2$ при $\lambda \in (0,\pi)$, $\mathbb{D}Z_{1,\varepsilon}(\lambda) = 1$ при $\lambda \in \{0,\pi\}$.

Отсюда, конечно, сразу же выводятся все нужные свойства конечномерных распределений процесса $\big(Z_{1,x}(\lambda), Z_{2,x}(\lambda)\big)$. А именно,
1. Этот процесс является гауссовским, так как линейное преобразование гауссовских векторов является гауссовским вектором.
2. Процессы $Z_{1,x}(\lambda)$ и $Z_{2,x}(\lambda)$ имеют независимые значения.
3. $Z_{2,x}(0) = Z_{2,x}(\pi) = 0$.
4. $\mathbb{E}Z_{1,x}(\lambda) = \mathbb{E}Z_{2,x}(\lambda) = 0$.
5. При $\lambda \in (0,\pi)$

$$\mathbb{D}Z_{1,x}(\lambda) = \mathbb{D}Z_{2,x}(\lambda) = \frac{1}{2}\left(\sum_m c_m \cos(\lambda m)\right)^2 + \frac{1}{2}\left(\sum_m c_m \sin(\lambda m)\right)^2 = \frac{1}{2}\left|\sum_m c_m e^{-im\lambda}\right|^2 = \pi f(\lambda).$$

6. При $\lambda = 0$ или $\lambda = \pi$

$$\mathbb{D}Z_{1,x}(\lambda) = \left(\sum_m c_m \cos(\lambda m)\right)^2 = \left|\sum_m c_m e^{-im\lambda}\right|^2 = 2\pi f(\lambda).$$

7. При $\lambda_1 \neq \lambda_2$ случайные величины $Z_{1,x}(\lambda_1)$ и $Z_{2,x}(\lambda_2)$ некоррелированны и, следовательно, независимы.
8. Для любого $\lambda \in (0, \pi)$

$$\mathbb{E}Z_{1,x}(\lambda)Z_{2,x}(\lambda) = \left(\sum_m c_m \cos(\lambda m)\right)\left(\sum_m c_m \sin(\lambda m)\right) - \left(\sum_m c_m \sin(\lambda m)\right)\left(\sum_m c_m \cos(\lambda m)\right) = 0.$$

Следовательно, случайные процессы $Z_{1,x}$ и $Z_{2,x}$ независимы. Тем самым утверждение доказано. $\square$

Замечание 4.9. Предположим, что спектральная плотность положительна. Тогда при $\lambda \in (0, \pi)$ случайный процесс $I_{n,x}^{(0)}(\lambda)/f(\lambda)$ асимптотически имеет независимые значения, показательно распределенные с параметром 1. Этот факт может служить основанием для проверки гипотезы о том, что $f(\lambda) = \text{const}$.

Замечание 4.10. Результаты Теоремы 4.4 сохраняются, если рассмотреть вместо процессов $I_n^{(0)}(\lambda)$ процессы $I_n^{(1)}(\lambda)$ при $\lambda \in (0, \pi]$.

4.5 Дискретизация периодограммы

В этом разделе мы рассмотрим не всю периодограмму $I_n^{(0)}$ как функцию, заданную на отрезке $[0,\pi]$, а ее значения в некоторых специально выбранных точках $\lambda_k^{(n)}$, число которых будет увеличиваться с ростом n. А именно, положим $\lambda_k^{(n)} = 2\pi k/n$ для тех целых k, при которых $\lambda_k^{(n)} \in [0,\pi]$. Ясно, что $k = 0,\ldots n/2$ при четном n и $k = 0,\ldots(n-1)/2$ при нечетном n. В общем случае $k = 0,\ldots,\lfloor n/2\rfloor$, причем $\lambda_k^{(n)} \in (0,\pi)$ при $k = 1,\ldots,\lceil (n-1)/2\rceil$.

Выбор именно таких точек дискретизации периодограммы обычно связывают с возможностью применения в этом случае быстрого преобразования Фурье, то есть с вычислительными выгодами. Кроме того, $I_n^{(0)}(\lambda_k^{(n)}) = I_n^{(1)}(\lambda_k^{(n)})$ при $k \neq 0$. Теорема 4.5 показывает, что для этого есть и более глубокие причины.

Рассмотрим теперь так называемую *дискретную периодограмму.*

Для каждого $\lambda \in [0,\pi]$ обозначим $\alpha_n(\lambda)$ число вида $2\pi k/n$, ближайшее к λ. В случае, когда таких чисел 2, будем в качестве $\alpha_n(\lambda)$ брать меньшее из них. Положим $\alpha_n(\lambda) = \alpha_n(-\lambda)$ при $\lambda \in [-\pi,0]$ и продолжим функцию α_n по 2π-периодичности на всю прямую. Тогда функция

$$DI_{n,\mathbf{a}}(\lambda) = I_{n,\mathbf{a}}(\alpha_n(\lambda))$$

называется *дискретной периодограммой* последовательности $\mathbf{a} = (a_1,\ldots,a_n)$. Согласно Лемме 4.2, $DI_{n,\mathbf{b}}(\lambda) = DI_{n,\mathbf{a}}(\lambda)$, если $\lambda \in (0,\pi]$ и $\mathbf{b} = (a_1-\bar a,\ldots,a_n-\bar a)$ с $\bar a = (a_1+\ldots+a_n)/n$.

Теорема 4.6. *Утверждение Теоремы 4.4 (и, следовательно, утверждение Теоремы 4.3) сохраняется, если вместо периодограммы $I_n^{(0)}$ рассматривать дискретную периодограмму $DI_n^{(0)}$.*

Доказательство. Доказательство почти дословно повторяет доказательство Теоремы 4.4 и поэтому не приводится. □

4.5.1 Оценивание линейных функционалов от спектральной плотности

Рассмотрим теперь свойства периодограммы в точках $\lambda_k^{(n)} = 2\pi k/n$, $k = 0,\ldots,\lfloor n/2\rfloor$ для линейного процесса.

Теорема 4.7. *Пусть x_n — линейный процесс*

$$x_m = \sum_{j=-\infty}^{\infty} c_{m+j}\varepsilon_j$$

причем $\sum_\ell |\ell||c_\ell| < \infty$, ε_n независимы, $\mathbb{E}\varepsilon_n = 0$, $\mathbb{E}\varepsilon_n^2 = 1$ и $\mathbb{E}\varepsilon_n^4 = m_4 < \infty$. Тогда

1. $\mathbb{E}I_n^{(0)}\big(\lambda_j^{(n)}\big) = f\big(\lambda_j^{(n)}\big) + O(1/n)$ равномерно по j.

2. Если $\lambda_j^{(n)}, \lambda_k^{(n)} \in [0,\pi]$, то

$$\mathrm{Cov}\left(I_n^{(0)}\big(\lambda_j^{(n)}\big), I_n^{(0)}\big(\lambda_k^{(n)}\big)\right) = \begin{cases} 2f^2\big(\lambda_j^{(n)}\big) + O\big(1/n\big) & \text{при } \lambda_j^{(n)} = \lambda_k^{(n)} \in \{0,\pi\}, \\ f^2\big(\lambda_j^{(n)}\big) + O\big(1/n\big) & \text{при } \lambda_j^{(n)} = \lambda_k^{(n)} \neq 0,\pi\,, \\ O(1/n) & \text{при } \lambda_j^{(n)} \neq \lambda_k^{(n)}, \end{cases}$$

причем остаточные члены $O(1/n)$ равномерны по j,k.

Доказательство. Согласно Лемме 1.1, из условия $\sum_\ell |\ell||c_\ell| < \infty$ следует, что $\sum_s |s||R(s)| < \infty$. Поэтому при доказательстве первого утверждения мы можем сослаться на четвертое утверждение Предложения 4.1. Второе утверждение оставим без доказательства. □

На основании Теоремы 4.7 можно строить состоятельные оценки линейных функционалов от спектральной плотности.

Лемма 4.6. *Пусть $h : [a, b] \mapsto \mathbb{R}$ является функцией ограниченной вариации.*[6] *Рассмотрим точки $a = x_0 < x_1 < \ldots < x_p = b$ и обозначим $\Delta_i = x_i - x_{i-1}$. Кроме того, пусть $y_i \in [x_{i-1}, x_i]$. Тогда*

$$\left| \int_a^b h(x)dx - \sum_{i=1}^{p} h(y_i)\Delta_j \right| \leq \max_i \Delta_i \operatorname{Var}(h).$$

Доказательство. Зафиксируем некоторое $\varepsilon > 0$. Обозначим $h_i^* = \sup_{x\in[x_{i-1},x_i]} h(x)$ и выберем точку $z_i^* \in [x_{i-1}, x_i]$ такую, что $h_i^* - h(z_i^*) < \varepsilon$. Аналогично, пусть $h_i^{**} = \inf_{x\in[x_{i-1},x_i]} h(x)$ и $t_i^* \in [x_{i-1}, x_i]$ с $h(t_i^*) - h_i^{**} < \varepsilon$. Тогда

$$\begin{aligned} \left| \int_a^b h(x)dx - \sum_{i=1}^{p} h(y_i)\Delta_j \right| &= \left| \sum_i \int_{x_{i-1}}^{x_i} \big(h(x) - h(y_i)\big)dx \right| \leq \sum_i \int_{x_{i-1}}^{x_i} \big|h(x) - h(y_i)\big|dx \leq \\ \leq \sum_i \int_{x_{i-1}}^{x_i} \max\big(h_i^* - h(y_i), h(y_i) - h_i^{**}\big)dx &= \sum_i \max\big(h_i^* - h(y_i), h(y_i) - h_i^{**}\big)\Delta_i \leq \\ \leq (\max_i \Delta_i) \sum_i \max\big(h(z_i^*) - h(y_i), h(y_i) - h(t_i^*)\big) &+ \varepsilon(b - a). \end{aligned} \tag{4.16}$$

Поскольку каждое слагаемое под знаком суммы в правой части (4.16) представляет собой модуль разности значений функции h, вычисленных в двух точках интервала $[x_{i-1}, x_i]$, то вся сумма не превосходит $\operatorname{Var}(h)$. Устремляя ε к нулю, получаем требуемое. $\square$

Для любых функций g и h, определенных на $[0, \pi]$, и любого $n \geq 2$ введем обозначение

$$S_n\langle g, h\rangle = \frac{2\pi}{n} \sum_{0\leq j\leq \lfloor n/2\rfloor} h\big(\lambda_j^{(n)}\big) g\big(\lambda_j^{(n)}\big).$$

Теорема 4.8. *Пусть $g : [0, \pi] \mapsto \mathbb{R}$ является функцией ограниченной вариации. Обозначим*

$$\langle g, f\rangle = \int_0^\pi f(\lambda)g(\lambda)d\lambda,$$

Тогда в условиях Теоремы 4.7

$$\mathbb{E}S_n\langle g, I_n^{(0)}\rangle = \langle g, f\rangle + O(1/n) \tag{4.17}$$

и

$$n\mathbb{E}\Big(S_n\langle g, I_n^{(0)}\rangle - \langle g, f\rangle\Big)^2 = O(1). \tag{4.18}$$

Доказательство. Для простоты будем рассматривать случай нечетного n. Тогда $S_n\langle g, f\rangle$ является ничем иным, как суммой Римана с постоянным шагом $2\pi/n$ для интеграла $\langle g, f\rangle$. Согласно Следствию 1.1 и Лемме 1.1 спектральная плотность f обладает непрерывной производной. Поэтому она является функцией ограниченной вариации. Легко показать, что произведение функций ограниченной вариации — тоже функция ограниченной вариации. Поэтому из Леммы 4.6 следует, что

$$|S_n\langle g, f\rangle - \langle g, f\rangle| = O(1/n).$$

[6] По определению это означает, что $\operatorname{Var}(h) \stackrel{\text{def}}{=} \sup \sum_{i=0}^{p-1} |h(z_{i+1}) - h(z_i)| < \infty$, где супремум берется по всевозможным разбиениям отрезка $[a, b]$ с узлами разбиения $a = z_0 < z_1 < \ldots < z_p = b$ с всевозможными $p \geq 2$.

С другой стороны, так как $\mathbb{E}S_n\langle g, I_n^{(0)}\rangle = S_n\langle g, \mathbb{E}I_n^{(0)}\rangle$, то из первого утверждения Теоремы 4.7 сразу же следует (4.17). Далее,

$$n\mathbb{E}\left(S_n\langle g, I_n^{(0)}\rangle - \langle g, f\rangle\right)^2 = n\mathbb{D}\left(S_n\langle g, I_n^{(0)}\rangle\right) + n\left(\mathbb{E}S_n\langle g, I_n^{(0)}\rangle - \langle g, f\rangle\right)^2 = n\mathbb{D}\left(S_n\langle g, I_n^{(0)}\rangle\right) + O(1/n).$$

Поэтому вместо среднеквадратического отклонения нам достаточно рассматривать дисперсию. Обозначая для краткости $\lambda_\ell = \lambda_\ell^{(n)}$, получаем согласно второму утверждению Теоремы 4.7, что

$$n\mathbb{D}\left(S_n\langle g, I_n^{(0)}\rangle\right) = \frac{4\pi^2}{n}\sum_{j,k} g(\lambda_j)g(\lambda_k)\mathrm{Cov}\left(I_n^{(0)}(\lambda_j), I_n^{(0)}(\lambda_k)\right) =$$

$$= \frac{4\pi^2}{n}\sum_{0\le k<n/2} g^2(\lambda_k)\mathbb{D}\left(I_n^{(0)}(\lambda_k)\right) + \frac{4\pi^2}{n}\sum_{j\ne k} g(\lambda_j)g(\lambda_k)\mathrm{Cov}\left(I_n^{(0)}(\lambda_j), I_n^{(0)}(\lambda_k)\right) =$$

$$= \frac{4\pi^2}{n}\sum_{0\le k<n/2}\left(g^2(\lambda_k)f^2(\lambda_k) + O(1/n)\right) + \frac{4\pi^2}{n}g^2(0)f^2(0) + O(1/n) + O(1/n^2)\sum_{j\ne k} g(\lambda_j)g(\lambda_k).$$

Так как

$$\frac{2\pi}{n}\sum_{0\le k<n/2} g^2(\lambda_k)f^2(\lambda_k) \to \int_0^\pi g^2(\lambda)f^2(\lambda)d\lambda,$$

а $\sum_{j\ne k} g(\lambda_j)g(\lambda_k) = O(n^2)$, получаем требуемое. □

Замечание 4.11. 1. Аналогичным образом доказывается, что для функций ограниченной вариации g_1, g_2

$$n\mathrm{Cov}\left(S_n\langle g_1, I_n^{(0)}\rangle, S_n\langle g_2, I_n^{(0)}\rangle\right) = O(1).$$

2. Более аккуратный анализ показывает, что

$$n\mathbb{D}\left(S_n\langle g, I_n^{(0)}\rangle\right) \to \sigma^2(g, f) = \int_0^\pi g^2(\lambda)f^2(\lambda)d\lambda + \Xi_1(g, f),$$

и

$$n\mathrm{Cov}\left(S_n\langle g_1, I_n^{(0)}\rangle, S_n\langle g_2, I_n^{(0)}\rangle\right) \to \sigma(g_1, g_2, f) = \int_0^\pi g_1(\lambda)g_2(\lambda)f^2(\lambda)d\lambda + \Xi_2(g, f)$$

где постоянные $\Xi_1(g, f), \Xi_2(g, f)$ обращаются в нуль при $\mathbb{E}\varepsilon_k = 3$.
3. Можно также доказать вариант центральной предельной теоремы для случайных векторов

$$\sqrt{n}\left(S_n\langle g_1, I_n^{(0)}\rangle - \langle g_1, f\rangle,\ S_n\langle g_2, I_n^{(0)}\rangle - \langle g_2, f\rangle\rangle\right):$$

совместное распределение этих векторов слабо сходится к двумерному гауссовскому распределению с нулевым средним и ковариационной матрицей, определенной дисперсиями $\sigma^2(g_1, f), \sigma^2(g_2, f)$ и ковариацией $\sigma(g_1, g_2, f)$.

Таким образом, хотя периодограмма не является состоятельной оценкой спектральной плотности, она позволяет строить состоятельные оценки интегралов от этой плотности. В частности, для грубой оценки поведения спектральной плотности можно использовать «периодограммную гистограмму», аналогичную обычной гистограмме в случае повторной выборки.

4.5.2 Уточнение Теоремы 4.2

Можно задать вопрос: какова в предыдущих рассуждениях роль используемой дискретизации периодограммы? Один из ответов уже приводился — это возможности вычислений с помощью быстрого преобразования Фурье. Более глубокие соображения содержатся в следующем предложении, приводимом без доказательства.

Предложение 4.5. *Пусть в условиях Теоремы 4.7 число* $\varepsilon > 0$ *произвольно. Тогда равномерно по* $\varepsilon < \lambda \neq \mu < \pi - \varepsilon$

$$\mathrm{Cov}\left(I_n^{(0)}(\lambda), I_n^{(0)}(\mu)\right) = \left(\frac{\sin\left(n(\mu-\lambda)\right)/2}{n\sin(\mu-\lambda)/2}\right)^2 f(\mu)f(\lambda) + O(1/n). \tag{4.19}$$

Обсудим результат этого Предложения. Конечно, он интересен только при близких λ и μ. Поэтому зафиксируем λ, положим $\mu = \lambda + 2\alpha_n$ и будем считать, что $\alpha_n \to 0$ при $n \to \infty$. Тогда ввиду гладкости функции f (4.19) превратится в

$$\begin{aligned}\mathrm{Cov}\left(I_n^{(0)}(\lambda), I_n^{(0)}(\mu)\right) = \left(\frac{\sin(n\alpha_n)}{n\sin(\alpha_n)}\right)^2 f(\lambda+2\alpha_n)f(\lambda) + O(1/n) \sim \\ \sim \left(\frac{\sin(n\alpha_n)}{n\sin(\alpha_n)}\right)^2 f^2(\lambda) + O(1/n).\end{aligned} \tag{4.20}$$

Рассмотрим теперь различные поведения последовательности α_n, считая, что $f(\lambda) > 0$.

1. Если $n\alpha_n \to 0$, то

   $$\mathrm{Cov}\left(I_n^{(0)}(\lambda), I_n^{(0)}(\mu)\right) \to f^2(\lambda).$$

   Таким образом, если λ и μ очень близки, то ковариация периодограмм в этих точках в пределе совпадает с дисперсией, как если бы $\mu = \lambda$.

2. Если $n\alpha_n \to \infty$, то первый член (4.5.2) имеет порядок убывания $O(1/(n\alpha_n)^2)$, то есть ковариация в близких точках стремится к нулю, но не быстрее, чем для фиксированных различных λ, μ, где достигается скорость сходимости $O(1/n)$.

3. Если $n\alpha_n \to c > 0$, то все зависит от величины c. При $c \neq \pi k$

   $$\mathrm{Cov}\left(I_n^{(0)}(\lambda), I_n^{(0)}(\mu)\right) \to \frac{\sin^2 c}{c^2}\ f^2(\lambda).$$

   Если же $c = \pi k$, то есть $\mu - \lambda \sim 2\pi k/n$, то ковариация убывает как $O(1/n)$.

5 Оценивание спектральной плотности: окна сглаживания

Итак, мы убедились, что периодограмма $I_n^{(0)}(\lambda)$ не является состоятельной оценкой спектральной плотности $f(\lambda)$, хотя при разумных условиях она асимптотически несмещена. То же самое имеет место и для периодограммы $I_n^{(1)}(\lambda)$. Попробуем понять, в чем здесь дело. Как уже указывалось, $I_n^{(0)}(\lambda)$ может быть представлена как в виде

$$I_n^{(0)}(\lambda) = \frac{1}{2\pi}\left(\widetilde{R}_n^{(0)}(0) + 2\sum_{m=1}^{n-1}\cos(m\lambda)\widetilde{R}_n^{(0)}(m)\right), \tag{5.1}$$

так и в виде

$$I_n^{(0)}(\lambda) = \frac{1}{2\pi n}\left|\sum_{k=1}^{n} x_k e^{-ik\lambda}\right|^2. \tag{5.2}$$

«Недостатки» представления (5.1) понятны, так как дисперсия $\widetilde{R}_n(m)$ при m, близких к n, должна быть гораздо больше, чем при маленьких m, ввиду небольшого числа осредняемых слагаемых вида $x_j x_{j+m}$ Например, при $m = n-1$ оценка $\widetilde{R}_n(n-1)$ состоит только из одного слагаемого $x_1 x_n/n$.

Естественной идеей исправления этой ситуации является введение функции $k_n(m)$, которая бы убывала при росте m, и рассмотрении оценки

$$\widetilde{f}_n(\lambda) = \frac{1}{2\pi}\left(k_n(0)\widetilde{R}_n^{(0)}(0) + 2\sum_{m=1}^{n-1}\cos(m\lambda)k_n(m)\widetilde{R}_n^{(0)}(m)\right)$$

вместо (5.1). Естественно, эта функция должна быть подобрана таким образом, чтобы сохранялась асимптотическая несмещенность оценки.

Представление (5.2) наводит на другие мысли. Поскольку, как уже известно, значения $I_n^{(0)}(\lambda)$ при различных λ асимптотически независимы, а (ввиду непрерывности f) при близких λ еще и приблизительно одинаково распределены, то, осреднив $I_n^{(0)}$ в некоторой окрестности интересующего нас значения значения λ с подходящими весами, можно надеяться на получение состоятельной оценки спектральной плотности в этой точке λ.

Оказывается, эти две идеи при надлежащей реализации просто совпадают.

5.1 Окна сглаживания

Возьмем четную суммируемую функцию $w(\lambda)$, определенную на $\mathbb{R}$ и такую, что $\int_{\mathbb{R}} w(\lambda)d\lambda = 1$. Кроме того, пусть $L_n \to +\infty$ и $w_n(\lambda) = L_n w(L_n\lambda)$. Тогда $\int_{\mathbb{R}} w_n(\lambda)d\lambda = 1$.

Положим

$$\widehat{f}_n^{(0)}(\lambda) = \int_{\mathbb{R}} w_n(\lambda-\alpha)I_n^{(0)}(\alpha)d\alpha \quad \text{и} \quad \widehat{f}_n^{(1)}(\lambda) = \int_{\mathbb{R}} w_n(\lambda-\alpha)I_n^{(1)}(\alpha)d\alpha.$$

Случайные величины такого вида и будут служить оценками спектральной плотности. Отметим сразу, что в случае, когда функция w неотрицательна, эти оценки тоже будут неотрицательными.

Функция w будет называться *частотным окном сглаживания.* Пусть

$$k(x) = \int_{\mathbb{R}} e^{ix\lambda}w(\lambda)d\lambda.$$

Это — прямое преобразование Фурье функции w. Очевидно, $k(0) = 1$, $k(-x) = k(x)$ и $k(x) \to 0$ при $x \to \infty$. Если k — суммируемая функция, то

$$w(\lambda) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{-i\lambda x}k(x)dx.$$

(обратное преобразование Фурье). Если $\int_{\mathbb{R}} w^2(\lambda)d\lambda < \infty$, то

$$\int_{\mathbb{R}} w^2(\lambda)d\lambda = 2\pi \int_{\mathbb{R}} k^2(x)dx$$

(равенство Парсеваля).

Лемма 5.1.

$$\widehat{f}_n^{(0)}(\lambda) = \frac{1}{2\pi} \sum_{r=-n+1}^{n-1} k(r/L_n)\widetilde{R}_n^{(0)}(r)e^{-ir\lambda} \tag{5.3}$$

и

$$\widehat{f}_n^{(1)}(\lambda) = \frac{1}{2\pi} \sum_{r=-n+1}^{n-1} k(r/L_n)\widetilde{R}_n^{(1)}(r)e^{-ir\lambda}. \tag{5.4}$$

Доказательство.

$$\widehat{f}_n^{(0)}(\lambda) = \frac{1}{2\pi} \int_{\mathbb{R}} w_n(\lambda-\alpha)\left(\sum_{r=-n+1}^{n-1} \widetilde{R}_n^{(0)}(r)e^{-ir\alpha}\right)d\alpha = \frac{1}{2\pi} \sum_{r=-n+1}^{n-1} \widetilde{R}_n^{(0)}(r) \int_{\mathbb{R}} w_n(\lambda-\alpha)e^{-ir\alpha}d\alpha.$$

$$\begin{aligned} \int_{\mathbb{R}} L_n w\big(L_n(\lambda-\alpha)\big)e^{-ir\alpha}d\alpha = \int_{\mathbb{R}} L_n w(L_n\nu)e^{-ir\lambda}e^{ir\nu}d\nu = \\ = e^{-ir\lambda}\int_{\mathbb{R}} w(\beta)e^{ir\beta/L_n}d\beta = e^{-ir\lambda}k(r/L_n). \end{aligned}$$

Для $\widehat{f}_n^{(1)}(\lambda)$ так же. □

Функция k называется *ковариационным окном сглаживания.*

5.2 Ковариационные окна сглаживания. Смещение

Рассмотрим ограниченную четную функцию k, сосредоточенную на $[-1,1]$ и равную единице в нуле. Кроме того, будем считать, что $L_n \le n-1$. Тогда (5.3) превратится в

$$\widehat{f}_n^{(0)}(\lambda) = \frac{1}{2\pi} \sum_{r=-L_n}^{L_n} k(r/L_n)\widetilde{R}_n^{(0)}(r)e^{-ir\lambda} = \frac{1}{2\pi} \sum_{r=-L_n}^{L_n} k(r/L_n)\widetilde{R}_n^{(0)}(r)\cos(r\lambda). \tag{5.5}$$

Предложение 5.1. *Предположим, что* $1-k(x) \sim \kappa|x|^q$, *где* $\kappa > 0$ *и* $q > 0$. *Пусть, кроме того,* $\sum_r |r|^p\,|R(r)| < \infty$ *при некотором* $p > 0$ *и* $L_n \to \infty$.

1. Если $q \le p$ *и либо* $p \ge 1$ *и* $L_n^q/n \to 0$, *либо* $p < 1$ *и* $L_n^{q-p+1}/n \to 0$, *то*

$$L_n^q\big(\mathbb{E}\widehat{f}_n^{(0)}(\lambda) - f(\lambda)\big) \to -\frac{\kappa}{2\pi} \sum_{r=-\infty}^{\infty} |r|^q R(r)\cos(r\lambda) = -\frac{\kappa}{\pi} \sum_{r=1}^{\infty} r^q R(r)\cos(r\lambda). \tag{5.6}$$

2. Если $p < q$ *и либо* $p \ge 1$ *и* $L_n^p/n \to 0$, *либо* $p < 1$ *и* $L_n/n \to 0$, *то*

$$L_n^p\left(\mathbb{E}\widehat{f}_n^{(0)}(\lambda) - f(\lambda)\right) \to 0. \tag{5.7}$$

Доказательство. Так как $p > 0$, то

$$f(\lambda) = \frac{1}{2\pi} \sum_{r=-\infty}^{\infty} R(r)\cos(r\lambda).$$

Поскольку $\mathbb{E}\widetilde{R}_n^{(0)}(r) = (1 - |r|/n)R(r)$, то

$$\mathbb{E}\widehat{f}_n^{(0)}(\lambda) = \frac{1}{2\pi} \sum_{|r|\le L_n} R(r)(1-|r|/n)k(r/L_n)\cos(\lambda r) =$$

$$= \frac{1}{2\pi} \sum_{|r|\le L_n} R(r)k(r/L_n)\cos(\lambda r) - \frac{1}{2\pi n} \sum_{|r|\le L_n} R(r)\,|r|\,k(r/L_n)\cos(\lambda r).$$

1. Рассмотрим случай $q \le p$. Тогда

$$L_n^q \mathbb{E}\big(\widehat{f}_n^{(0)}(\lambda) - f(\lambda)\big) = L_n^q \frac{1}{2\pi} \sum_{|r|\le L_n} R(r)\big(k(r/L_n) - 1\big)\cos(\lambda r) -$$

$$-L_n^q \frac{1}{2\pi} \sum_{|r|>L_n} R(r)\cos(\lambda r) - L_n^q \frac{1}{2\pi n} \sum_{|r|\le L_n} R(r)\,|r|\,k(r/L_n)\cos(\lambda r) = J_1 + J_2 + J_3. \tag{5.8}$$

Так как $\sum_r |r|^q\, |R(r)| < \infty$, то $J_2 \to 0$. Рассмотрим J_3. Если $p \ge 1$, то, поскольку $|k(x)| < c$,

$$|J_3| \le C\frac{L_n^q}{n} \sum_{|r|\le L_n} |R(r)|\; |r| \le C\frac{L_n^q}{n} \sum_{r=-\infty}^{\infty} |R(r)|\,|r| \to 0$$

ввиду условия $L_n^q/n \to 0$. Если $p < 1$, то

$$|J_3| \le C\frac{L_n^q}{n} \sum_{|r|\le L_n} |R(r)|\; |r| \le C\frac{L_n^{q-p+1}}{n} \sum_{|r|\le L_n} |R(r)|\,|r|^p,$$

что сходится к нулю из-за условия $L_n^{q-p+1}/n \to 0$.

Займемся первым слагаемым. Ясно, что

$$J_1 = L_n^q \frac{1}{2\pi} \sum_{|r|\le L_n} R(r)\Big(k(r/L_n) - 1 + \kappa\,|r/L_n|^q\Big)\cos(\lambda r) - L_n^q \frac{1}{2\pi} \sum_{|r|\le L_n} \kappa R(r)\,|r/L_n|^q\cos(\lambda r) = I_1 - I_2.$$

Конечно,

$$I_2 \to \frac{\kappa}{2\pi} \sum_{r=-\infty}^{\infty} R(r)|r|^q\cos(\lambda r).$$

Далее, $k(x) - 1 + \kappa|x|^q = |x|^q\varphi(x)$, где $\varphi(x) \to 0$ при $x \to 0$ и φ — ограниченная по модулю функция. Следовательно, так как $p \ge q$, то ряд $\sum_r |R(r)|\; |r|^q$ сходится и утверждение следует из Теоремы Лебега о мажорируемой сходимости.

2. Перейдем к случаю $p < q$. Аналогично (5.8),

$$L_n^p \mathbb{E}\big(\widehat{f}_n^{(0)}(\lambda) - f(\lambda)\big) = L_n^p \frac{1}{2\pi} \sum_{|r|\le L_n} R(r)\big(k(r/L_n) - 1\big)\cos(\lambda r) -$$

$$-L_n^p \frac{1}{2\pi} \sum_{|r|>L_n} R(r)\cos(\lambda r) - L_n^p \frac{1}{2\pi n} \sum_{|r|\le L_n} R(r)\,|r|\,k(r/L_n)\cos(\lambda r) = \Psi_1 + \Psi_2 + \Psi_3.$$

Сходимость к нулю Ψ_2 доказывается точно так же, как сходимость J_2 в (5.8). Нетрудно видеть, что $|\Psi_1| = O(L_n^{p-q})$.

Займемся слагаемым Ψ_3. Если $p \geq 1$, то ряд $\sum_r |R(r)|\,|r|$ сходится и $\Psi_3 = O(L_n^p/n)$. Если же $p < 1$, то для некоторой постоянной $C > 0$

$$|\Psi_3| \leq C\frac{L_n^p}{n}\sum_{|r|\leq L_n}|R(r)|\,|r| = C\frac{L_n^p}{n}\sum_{|r|\leq L_n}|R(r)|\,|r|^p\,|r|^{1-p} \leq C\frac{L_n^p}{n}\,L_n^{1-p}\sum_{|r|\leq L_n}|R(r)|\,|r|^p = O(L_n/n).$$

Утверждение доказано. □

Согласно Предложению 5.1, условия асимптотической несмещенности и скорость сходимости $\mathbb{E}\widehat{f}_n^{(0)}(\lambda)$ к $f(\lambda)$ зависят как от от выбора L_n так и от параметров p, q. Обсудим роль этих параметров.

Параметр p описывает (гарантированную) скорость сходимости ковариационной функции к нулю (или, что тоже самое, гарантированную степень гладкости спектральной плотности). Таким образом, p является характеристикой наблюдаемого процесса. Что касается параметра q, то он в нашей власти.

Если считать, что число p нам известно, то, как следует из Предложения 5.1, при $p \geq 1$ смещение ведет себя следующим образом в зависимости от q:

- при росте q от нуля до p включительно, смещение имеет порядок $O(L_n^{-q})$, а условие на L_n приобретает вид $L_n^q = o(n)$;
- при $q > p$ смещение имеет порядок $o(L_n^{-p})$, а L_n должно удовлетворять условию $L_n^p = o(n)$.

Если же $p < 1$, то

- при росте q от нуля до p включительно, смещение имеет порядок $O(L_n^{-q})$, а условие на L_n приобретает вид $L_n^{q-p+1} = o(n)$;
- при $q > p$ смещение имеет порядок $o(L_n^{-p})$, а L_n должно удовлетворять условию $L_n = o(n)$.

Эти соображения показывают, что при выбор $q > p$ приводит к меньшему асимптотическому смещению, чем выбор $q \in [0, p]$. В то же время для многих стационарных последовательностей $\sum_r |r|^p\,|R|(r) < \infty$ для любого $p > 0$. Например, это имеет место для процессов авторегресии и авторегресии-скользящего суммирования, где $|R(r)| \leq ce^{-\alpha|r|}$ при некоторых положительных c, α.

Отсюда появляется идея, что для таких последовательностей нужно выбирать ядро k с $q = +\infty$. Наиболее естественным в этом смысле является *прямоугольное окно*[7] $k(x) \equiv 1$ при $|x| \leq 1$ (напомним, что $k(x) = 0$ при $|x| > 1$). Докажем для этого окна следующий факт

Предложение 5.2. *Пусть* $|R(r)| \leq ce^{-\alpha|r|}$, $L_n < n$ *и* $L_n/\ln n \to \infty$. *Рассмотрим оценку* (5.5) *с* $k(x) \equiv 1$ *при* $|x| \leq 1$. *Тогда*

$$n\big(\mathbb{E}\widehat{f}_n^{(0)}(\lambda) - f(\lambda)\big) \to -\frac{1}{2\pi}\sum_r R(r)\,|r|\,\cos(\lambda r). \tag{5.9}$$

Доказательство. Действуя так же, как в Предложении 5.1, получим, что

$$n\mathbb{E}\big(\widehat{f}_n^{(0)}(\lambda) - f(\lambda)\big) = -\frac{n}{2\pi}\sum_{|r|>L_n}R(r)\cos(\lambda r) - \frac{1}{2\pi}\sum_{|r|\leq L_n}R(r)\,|r|\,\cos(\lambda r). \tag{5.10}$$

[7]иногда его называют также окном Даниэля.

Второе слагаемое в правой части (5.10), очевидно, стремится к правой части (5.9). Что касается первого слагаемого, то, поскольку

$$\sum_{|r|>L_n} |R(r)| \le c_1 e^{-\alpha L_n}$$

при некотором $c_1 > 0$, а $L_n = \beta_n \ln n$ с $\beta_n \to \infty$, то оно стремится к нулю. □

Замечание 5.1. Прокомментируем условия Предложения 5.1 считая, что $p = q \ge 1$ — целое.
1. Из условия $\sum_r |R(r)|\, |r|^q < \infty$ следует, что спектральная плотность f является q раз непрерывно дифференцируемой. Более того, при четном q формула (5.6) приобретает вид

$$L_n^q\big(\mathbb{E}\widehat{f}_n^{(0)}(\lambda) - f(\lambda)\big) \to (-1)^{q/2+1}\kappa\, f^{(q)}(\lambda).$$

В частности, при $q = 2$

$$L_n^2\big(\mathbb{E}\widehat{f}_n^{(0)}(\lambda) - f(\lambda)\big) \to \kappa\, f^{(2)}(\lambda).$$

Это означает, что вблизи локального максимума смещение отрицательно, а вблизи локального минимума — положительно. Иначе говоря, происходит (в смысле смещения) сглаживание локальных экстремумов.
2. Перейдем от ковариационного окна k к частотному окну w:

$$w(\lambda) = \frac{1}{2\pi} \int_{-1}^{1} e^{-i\lambda x} k(x) dx.$$

Если полученная функция w суммируема, то

$$k(x) = \int_{-\infty}^{\infty} e^{ix\lambda} w(\lambda) d\lambda,$$

и условие $k(0) = 1$ соответствует условию $\int w(\lambda) d\lambda = 1$.
3. В случае $q = 1$ получается, что $1 - k(x) \sim \kappa |x|$, то есть функция k не является дифференцируемой в нуле. Простейший пример этого — окно $k(x) = 1 - |x|$ при $|x| \le 1$ (треугольная плотность), тогда

$$w(\lambda) = 2\, \frac{1 - \cos(\lambda)}{\lambda^2} \ge 0.$$

Последнее неравенство, как уже говорилось, обеспечивает неотрицательность оценки $\widehat{f}_n^{(0)}$. Функция $k(x) = 1 - |x|$ называется *модифицированным окном Бартлетта.*
4. Пусть $\int \lambda^2 |w(\lambda)| d\lambda < \infty$. Тогда k — дважды непрерывно дифференцируемая функция, причем $k'(0) = 0$ и

$$k''(0) = -\int_{-\infty}^{\infty} \lambda^2 w(\lambda) d\lambda.$$

Если последний интеграл не равен нулю, то $q = 2$ и $\theta = \int \lambda^2 w(\lambda) d\lambda$.

Рассмотрим теперь случай оценки $f_n^{(1)}(\lambda)$.

Предложение 5.3. *1. Пусть в условиях Предложения 5.1 $q \le p$, $p \ge 1$ и $L_n^{q+1}/n \to 0$. Тогда имеет место сходимость* (5.6) *с заменой в левой части этого равенства $f_n^{(0)}(\lambda)$ на $f_n^{(1)}(\lambda)$.*

2. Пусть при некотором $\alpha > 0$ выполнено неравенство $|R(r)| \le c\, e^{-\alpha |r|}$. Если $L_n \to \infty$ и $L_n/n \to 0$, то для прямоугольного окна $k(x) = \mathbb{I}_{[-1,1]}(x)$

$$\mathbb{E}\widehat{f}_n^{(1)}(\lambda) - f(\lambda) = O(L_n/n).$$

Доказательство. 1. Поскольку $\mathbb{E}\widetilde{R}_n^{(1)}(k) = (1-k/n)\mathbb{E}\widehat{R}_n^{(1)}(k)$ и согласно (3.7), при $k = o(n)$

$$\mathbb{E}\widehat{R}_n^{(1)}(k) - R(k) \sim 2\pi f(0)/n,$$

то

$$\mathbb{E}\widetilde{R}_n^{(1)}(k) \sim \left(1 - \frac{k}{n}\right) R(k) - \frac{2\pi f(0)}{n}\,. \tag{5.11}$$

Поэтому равенство (5.8) превращается в

$$L_n^q \mathbb{E}\big(\widehat{f}_n^{(1)}(\lambda) - f(\lambda)\big) = J_1 + J_2 + J_3 + J_4, \tag{5.12}$$

где J_1, J_2, J_3 такие же, как в (5.8), а

$$J_4 \sim -L_n^q \frac{f(0)}{n}\,(2L_n + 1) = O(L_n^{q+1}/n).$$

Поскольку пределы J_i при $1 \le j \le 3$ такие же, как в Предложении 5.1, а $J_4 \to 0$, то первое утверждение доказано.

2. Ввиду (5.11)

$$\mathbb{E}\widehat{f}_n^{(1)}(\lambda) - f(\lambda) = -\frac{1}{2\pi}\sum_{|r|>L_n} R(r)\cos(\lambda r) - \frac{1}{2\pi n}\sum_{|r|\le L_n} R(r)\,|r|\,\cos(\lambda r) + \frac{1}{2\pi}(2L_n - 1)O(1/n).$$

Первое слагаемое экспоненциально быстро стремится к нулю, второе имеет порядок $O(n^{-1})$, а третье — $O(L_n/n)$. $\square$

5.3 Предельные дисперсии и ковариации

Теорема 5.1. *Пусть четная функция k задана на $[-1,1]$ и непрерывна на этом отрезке. Предположим, что $L_n \to \infty$ и $L_n = o(n)$. Если стационарный процесс x_n является линейным с коэффициентами c_j, удовлетворяющими условию $\sum_j |c_j| < \infty$ и белым шумом ε_j — н.о.р. с.в. с $\mathbb{E}\varepsilon_j = 0$, $\mathbb{E}\varepsilon_j^2 = 1$ и $\mathbb{E}\varepsilon_j^4 < \infty$, то на отрезке частот $[0,\pi]$*

$$\frac{n}{L_n}\,\mathrm{Cov}\big(\widehat{f}_n^{(1)}(\lambda_1), \widehat{f}_n^{(1)}(\lambda_2)\big) \to \begin{cases} 2f^2(\lambda)\int_{-1}^{1} k^2(x)dx & \textit{при}\ \ \lambda_1 = \lambda_2 = 0, \pi, \\ f^2(\lambda)\int_{-1}^{1} k^2(x)dx & \textit{при}\ \ \lambda_1 = \lambda_2 \in (0,\pi), \\ 0 & \textit{при}\ \ \lambda_1 \ne \lambda_2. \end{cases} \tag{5.13}$$

Следствие 5.1. Для того, чтобы в условиях первого пункта Предложения 5.3 и Теоремы 5.1 среднеквадратического отклонение $\widehat{f}_n^{(1)}(\lambda)$ от $f(\lambda)$ имело оптимальный порядок сходимости к нулю, нужно, чтобы $L_n^{2q+1} \asymp n$. При этом $\mathbb{E}\big(\widehat{f}_n^{(1)}(\lambda) - f(\lambda)\big)^2 \asymp n^{-2q/(2q+1)}$.

Доказательство. Рассмотрим случай $\lambda \in (0,\pi)$. Поскольку

$$\mathbb{E}\big(\widehat{f}_n^{(1)}(\lambda) - f(\lambda)\big)^2 = \mathbb{D}\widehat{f}_n^{(1)}(\lambda) + \big(\mathbb{E}\widehat{f}_n^{(1)}(\lambda) - f(\lambda)\big)^2,$$

то среднеквадратическое отклонение будет стремиться к нулю, если $L_n = o(n)$ и $L_n^{q+1} = o(n)$, при этом

$$\mathbb{E}\big(\widehat{f}_n^{(1)}(\lambda) - f(\lambda)\big)^2 = \frac{L_n}{n}\, f^2(\lambda)\theta^2 + \frac{\kappa^2}{L_n^{2q}}\, C_q^2(\lambda), \tag{5.14}$$

где $\theta^2 = \int_{-1}^{1} k^2(x)dx$ и

$$C_q(\lambda) = -\frac{1}{\pi}\sum_{r=1}^{\infty} |r|^q R(r)\cos(r\lambda).$$

Отметим еще раз, что при $q = 2$

$$C_2(\lambda) = \frac{1}{2\pi} f^{(2)}(\lambda).$$

Если $L_n/n \asymp L_n^{-2q}$, то порядки обоих слагаемых совпадают и являются оптимальными. $\square$

5.3.1 Некоторые стандартные ковариационные окна

Для $\lambda \in (0, \pi)$ при $L_n = \lfloor n^{1/(2q+1)} \rfloor$ равенство (5.14) принимает вид

$$\mathbb{E}\big(\widehat{f}_n^{(1)}(\lambda) - f(\lambda)\big)^2 = \frac{1}{n^{2q/(2q+1)}}\big(f^2(\lambda)\theta^2 + \kappa^2 C_q^2(\lambda)\big) + o(n^{-2q/(2q+1)}).$$

Поскольку $f^2(\lambda)$ и $C_q(\lambda)$ заранее неизвестны, то предельными характеристиками среднеквадратического отклонения являются величины q, κ и $\theta^2 = \int_{-1}^{1} k^2(x)dx$. При этом q отвечает за скорость сходимости, κ — за смещение, а θ^2 — за разброс.

Для прямоугольного окна в условиях второго пункта Предложения 5.3 и Теоремы 5.1

$$\mathbb{E}\big(\widehat{f}_n^{(1)}(\lambda) - f(\lambda)\big)^2 \sim \frac{2L_n}{n} f^2(\lambda),$$

так как дисперсия и смещение в этом случае имеют одинаковый порядок, а $\theta^2 = 2$.

Приведем сводку классических ковариационных окон вместе с этими характеристиками.

Название	$k(x)$, $\lvert x\rvert \le 1$	q	κ	θ^2	sign	L_n	Order
Прямоугольное окно	1	$-\infty$		2	$\pm$	$o(n)$	L_n/n
Модиф. окно Бартлетта	$1-\lvert x\rvert$	1	1	2/3	+	$n^{1/3}$	$n^{-1/3}$
Усеченное окно Даниэля	$\sin(\pi x)/\pi x$	2	$\pi^2/6$	≈ 0.9	$\pm$	$n^{1/5}$	$n^{-2/5}$
Окно Блэкмена-Тьюки (БТ)	$1-2a+2a\cos(\pi x)$	2	$\pi^2 a$	$2(1-4a+6a^2)$	$\pm$	$n^{1/5}$	$n^{-2/5}$
Окно Ханнинга	БТ с $a=0.25$	2	$\pi^2/4$	0.75	$\pm$	$n^{1/5}$	$n^{-2/5}$
Окно Хемминга	БТ с $a=0.23$	2	$0.23\pi^2$	≈ 0.79	$\pm$	$n^{1/5}$	$n^{-2/5}$
Первое окно Парзена	$1-x^2$	2	1	≈ 1.07	$\pm$	$n^{1/5}$	$n^{-2/5}$
Второе окно Парзена	$\begin{cases} 1-6x^2-6\lvert x\rvert^3 & \lvert x\rvert \le 1/2 \\ 2(1-\lvert x\rvert^3) & \lvert x\rvert > 1/2 \end{cases}$	2	6	≈ 0.54	+	$n^{1/5}$	$n^{-2/5}$

Таблица 1: Ковариационные окна, условия на L_n и порядок убывания $\sqrt{\mathbb{E}\big(\widehat{f}_n^{(1)}(\lambda) - f(\lambda)\big)^2}$

Прокомментируем эту таблицу.

Замечание 5.2. Прямоугольное окно (для быстро убывающей ковариационной функции и при выборе $L_n \sim \sqrt{n}$) обеспечивает лучшую сходимость к нулю среднеквадратического отклонения.

Для модифицированной оценки Бартлетта $q = 1$. Следовательно, среднеквадратическое отклонение имеет вид $O(n^{-1/3})$, этому соответствует выбор $L_n \asymp n^{1/3}$. Для остальных окон $q = 2$, среднекв. отклонение равно $O(n^{-2/5})$ и $L_n \asymp n^{1/5}$. Тем самым при $q = 2$ скорость сходимости немного выше, чем при $q = 1$.

Возникает вопрос: почему бы не взять большое q (ведь при $q \to \infty$ скорость сходимости среднеквадратического отклонения стремится к $1/\sqrt{n}$, а задать $k(x) = 1 - |x|^q$ ничего не стоит)? И второй вопрос — зачем придумали так много всяческих окон для $q = 2$?

Ответ на эти вопросы основывается на допредельных соотношениях. Пусть, скажем, $n = 10^3$. Тогда для $q = 1$ главный член среднеквадратической погрешности имеет порядок 10^{-1}, а соответствующее $L_n \approx 10^1$. Для $q = 2$ эти же порядки имеют вид $10^{-6/5} \approx 0.063$, а $L_n \approx 10^{3/5} \approx 4$. Если же взять $q > 2$, то L_n еще уменьшится.

Но все наши рассуждения были основаны на том, что L_n большое ($L_n \to \infty$)! Поэтому при конечных n брать q слишком большим нельзя.

Выбор $q = 2$ обусловлен еще одной важной (особенно при не очень больших n) причиной. Как указывалось в разделе 5.1, оценка $\widehat{f}_n^{(1)}(\lambda)$, кроме ковариационного представления (5.4) имеет и периодограммное представление

$$\widehat{f}_n^{(1)}(\lambda) = \int_{\mathbb{R}} w_n(\lambda - \alpha) I_n^{(1)}(\alpha) d\alpha,$$

где $w_n(\lambda) = L_n w(L_n\lambda)$ и

$$w(\lambda) = \frac{1}{2\pi}\int_{-1}^{1} e^{-i\lambda x} k(x)dx.$$

Из этого представления становится очевидным, что оценка $\widehat{f}_n^{(1)}(\lambda)$ является гарантированно неотрицательной, если $w(\lambda) \geq 0$. А неотрицательность является весьма желательным свойством, особенно при относительно небольших n (при больших n это не так существенно, поскольку для положительной спектральной плотности $\widehat{f}_n^{(1)}(\lambda) \xrightarrow{\mathbb{P}} f(\lambda) > 0$).

Как уже говорилось, случаю целого $q > 2$ соответствует равенство

$$\int_{\mathbb{R}} \lambda^2 w(\lambda)d\lambda = 0,$$

что невозможно при положительном w.

Для прямоугольного окна оценка периодограммы, вообще говоря, является знакопеременной. Для окна Бартлетта $w \geq 0$ (что и обуславливает его жизнеспособность несмотря на небольшую скорость сходимости). Из всех примеров Таблицы 5.3.1, относящихся к $q = 2$, только второе окно Парзена обеспечивает неотрицательность $w(\lambda)$. Для остальных окон это свойство не выполняется. Обилие различных окон с $q = 2$ объясняется попытками уменьшить влияние этих отрицательных значений $w(\lambda)$.

5.4 Предельные теоремы и доверительные интервалы

Теорема 5.2. *При $\lambda \in [0, \pi]$ в условиях Предложения 5.3 и Теоремы 5.1, конечномерные распределения случайного процесса*

$$\xi_n(\lambda) = \sqrt{\frac{n}{L_n}}\left(\widehat{f}_n^{(1)}(\lambda) - \mathbb{E}\widehat{f}_n^{(1)}(\lambda)\right)$$

слабо сходятся к конечномерным распределениям гауссовского процесса с нулевым средним, независимыми значениями и дисперсиями, определенными правой частью равенства (5.13).

Теоремы 5.2 для построения доверительных интервалов, конечно, недостаточно, так как в этой теореме рассматривается разность $\widehat{f}_n^{(1)}(\lambda) - \mathbb{E}\widehat{f}_n^{(1)}(\lambda)$, а не разность $\widehat{f}_n^{(1)}(\lambda) - f(\lambda)$.

Замечание 5.3. Рассмотрим случайный процесс

$$\eta_n(\lambda) = \sqrt{\frac{n}{L_n}}\left(\widehat{f}_n^{(1)}(\lambda) - f(\lambda)\right) = \sqrt{\frac{n}{L_n}}\left(\widehat{f}_n^{(1)}(\lambda) - \mathbb{E}\widehat{f}_n^{(1)}(\lambda)\right) + \sqrt{\frac{n}{L_n}}\big(\mathbb{E}\widehat{f}_n^{(1)}(\lambda) - f(\lambda)\big). \tag{5.15}$$

При $\lambda \in (0, \pi)$ первое слагаемое в правой части (5.15) имеет предельное распределение $\mathrm{N}(0, f^2(\lambda)\theta^2)$, где $\theta = \int_{-1}^{1} k^2(x)dx$. Если

$$\sqrt{\frac{n}{L_n}}\big(\mathbb{E}\widehat{f}_n^{(1)}(\lambda) - f(\lambda)\big) \to 0, \tag{5.16}$$

то этот же слабый предел имеет и случайная величина $\eta(\lambda)$.

Смещение $\mathbb{E}\widehat{f}_n^{(1)}(\lambda) - f(\lambda)$ в случае прямоугольного окна и быстро убывающей ковариационной функции имеет порядок L_n/n, поэтому это условие выполнено.

В условиях первого пункта Предложения 5.3 это, вообще говоря, уже не так, поскольку

$$L_n^q\big(\mathbb{E}\widehat{f}_n^{(0)}(\lambda) - f(\lambda)\big) \to -\frac{\kappa}{\pi}\sum_{r=1}^{\infty} r^q R(r)\cos(r\lambda)$$

и поэтому

$$\sqrt{\frac{n}{L_n}}\big(\mathbb{E}\widehat{f}_n^{(1)}(\lambda) - f(\lambda)\big) \sim -\sqrt{\frac{L_n^{2q+1}}{n}}\ \frac{\kappa}{\pi}\sum_{r=1}^{\infty} r^q R(r)\cos(r\lambda).$$

Тем самым для сохранения предельного распределения $\mathrm{N}(0, f^2(\lambda)\theta^2)$ здесь нужно потребовать, чтобы $L_n^{2q+1}/n \to 0$. При относительно небольших n это, конечно, сильное ограничение. Например, при $q = 2$ это означает, что должно выполняться соотношение $L_n = o(n^{1/5})$.

Пусть теперь $\mathcal{L}(\eta_n(\lambda)) \Rightarrow \mathrm{N}(0, f^2(\lambda)\theta^2)$ и предположим, что $f(\lambda) > 0$. Тогда

$$\mathcal{L}\left(\frac{1}{\theta}\ \sqrt{\frac{n}{L_n}}\left(\frac{\widehat{f}_n^{(1)}(\lambda)}{f(\lambda)} - 1\right)\right) \Rightarrow \mathrm{N}(0, 1),$$

$$\mathrm{P}\left(1 - x_\gamma\theta\sqrt{L_n/n} < \widehat{f}_n^{(1)}(\lambda)/f(\lambda) < 1 + x_\gamma\theta\sqrt{L_n/n}\right) \to 1 - \gamma$$

(здесь, как обычно, x_γ является решением уравнения $\Phi(x) - \Phi(-x) = 1 - \gamma$, где $\Phi(x)$ — функция распределения стандартного нормального закона), и в предположении, что $\widehat{f}_n^{(1)}(\lambda) > 0$, получаем асмптотический доверительный интервал надежности $1 - \gamma$ для $f(\lambda)$:

$$\mathrm{P}\left(\frac{\widehat{f}_n^{(1)}(\lambda)}{1 + x_\gamma\theta\sqrt{L_n/n}} < f(\lambda) < \frac{\widehat{f}_n^{(1)}(\lambda)}{1 - x_\gamma\theta\sqrt{L_n/n}}\right) \to 1 - \gamma.$$

Этот доверительный интервал может, тем не менее, оказаться не очень удобным, если, например, $x_\gamma\theta\sqrt{L_n/n} > 1$ при относительно небольших n.

Альтернативное построение можно описать следующим образом. Так как

$$\mathcal{L}\left(\sqrt{\frac{n}{L_n}}\left(\widehat{f}_n^{(1)}(\lambda) - f(\lambda)\right)\right) \Rightarrow \mathrm{N}(0, \theta^2 f^2(\lambda)),$$

то (считая, что $f(\lambda)$ и $\widehat{f}_n^{(1)}(\lambda)$ положительны), получаем, что

$$\mathcal{L}\left(\sqrt{\frac{n}{L_n}}\left(\ln\left(f(\lambda)\right) - \ln\left(\widehat{f}_n^{(1)}(\lambda)\right)\right)\right) \Rightarrow \mathrm{N}(0, \theta^2),$$

откуда легко выводится, что

$$\mathrm{P}\left(\widehat{f}_n^{(1)}(\lambda)\, e^{-\theta x_\gamma L_n/n} < f(\lambda) < \widehat{f}_n^{(1)}(\lambda)\, e^{-\theta x_\gamma L_n/n}\right) \to 1 - \gamma.$$

Такой доверительный интервал является более стандартным.

6 Оценивание спектральной плотности: метод MEM

6.1 Авторегрессионное продолжение ковариационных функций

6.1.1 Процессы авторегрессии

Основные понятия и свойства. Пусть ξ_n, $n = 0, \pm 1, \pm 2, \ldots$ — стационарная в широком смысле последовательность, имеющая нулевое среднее, ковариационную функцию $R_\xi(n)$, спектральную меру m_ξ и ассоциированную стохастическую 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 белый шум ε_n, что для любого n

$$b_0 \xi_n + b_1 \xi_{n-1} + \ldots + b_q \xi_{n-q} = a_0 \varepsilon_n + a_1 \varepsilon_{n-1} + \ldots + a_p \varepsilon_{n-p}, \tag{6.1}$$

то процесс $\xi_.$ называется *процессом авторегрессии-скользящего суммирования порядка* (q,p), сокращенно — АРСС(q,p).[8]

Процесс (6.1) называется реализуемым относительно белого шума ε, если ε_n ортогонально ξ_m при $m < n$.

В случае, когда $p = 0$, ξ_n называется процессом (чистой) авторегрессии и обозначается АР(q).

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

$$Q(z) = \sum_{k=0}^{q} b_k z^{q-k} = z^q \sum_{k=0}^{q} b_k z^{-k} = z^q Q^*(z^{-1}), \tag{6.2}$$

то из $Q(z_0) = 0$ следует $Q^*(1/z_0) = 0$ и наоборот. В частности, из (6.2) следует, что полиномы Q и Q^* одновременно имеют корни или не имеют корней на единичной окружности $|z| = 1$.

В дальнейшем мы будем считать, что коэффициенты $(a_0, \ldots, a_p)$ и $(b_0, \ldots, b_q)$, а также белый шум ε_n вещественны.

Предложение 6.1. *Пусть $Q(z) \neq 0$ при $|z| = 1$. Тогда процесс* (6.1) *имеет спектральную плотность f_ξ, выражающуюся формулой*

$$f_\xi(\lambda) = \frac{1}{2\pi} \, \frac{|P^*(e^{-i\lambda})|^2}{|Q^*(e^{-i\lambda})|^2}\,, \quad \lambda \in (-\pi, \pi]. \tag{6.3}$$

Доказательство. Обозначим η_n левую (и правую) часть равенства (6.1). Так как $\eta_{\boldsymbol{.}}$ является линейным преобразованием $\xi_{\boldsymbol{.}}$, то мы можем стандартным способом вычислить спектральную меру этого процесса. Действительно,

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

[8]При $q = 0$ (6.1) оказывается процессом скользящего суммирования, слово «авторегрессия» подразумевает, что $q \geq 1$.

так что

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

Аналогично из правой части (6.1) получаем, что

$$m_\eta(d\lambda) = \left|P^*\!\left(e^{-i\lambda}\right)\right|^2 m_\epsilon(d\lambda) = \frac{1}{2\pi}\,\left|P^*\!\left(e^{-i\lambda}\right)\right|^2 d\lambda.$$

Поэтому

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

и, если $Q^*\!\left(e^{-i\lambda}\right) \neq 0$ 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 6.1. 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 (6.1) можно свести к изучению процесса чистой авторегрессии с теми же коэффициентами b_k.

Предложение 6.2. *Пусть $\eta_{\bullet}$ — процесс чистой авторегрессии, имеющий вид*

$$b_0\eta_n + b_1\eta_{n-1} + \ldots + b_q\eta_{n-q} = \varepsilon_n \tag{6.4}$$

и

$$\xi_n = a_0\eta_n + a_1\eta_{n-1} + \ldots + a_p\eta_{n-p}. \tag{6.5}$$

Тогда 1. Процесс $\xi_{\bullet}$ удовлетворяет (6.1). *2. Если процесс $\eta_{\bullet}$ является реализуемым относительно $\varepsilon_{\bullet}$, то таким же свойством обладает и $\xi_{\bullet}$.*

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 (6.4) и (6.5). □

Ввиду Предложения 6.2 дальнейшие рассуждения проводятся для процессов чистой авторегрессии

$$b_0\xi_n + b_1\xi_{n-1} + \ldots + b_q\xi_{n-q} = a_0\varepsilon_n, \tag{6.6}$$

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 6.3. *Процесс авторегрессии* (6.6) *является линейным преобразованием белого шума* ε_n.

Доказательство. Не умаляя общности, будем считать, что $a_0 = 1$. Согласно (6.6)

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

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

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

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

Следствие 6.1. Для любого белого шума ε_n и для любых чисел $(b_0, \ldots, b_q)$ таких, что $b_0 b_q \neq 0$ и полином Q^* не имеет корней на единичной окружности, существует процесс авторегрессии ξ_n, удовлетворяющий (6.6).

Доказательство. Снова достаточно рассмотреть случай $a_0 = 1$. Тогда из доказательства Предложения 6.3 следует, что в качестве искомого процесса авторегрессии можно взять

$$\xi_n = \int_{(-\pi,\pi]} e^{in\lambda} \frac{1}{Q^*\big(e^{-i\lambda}\big)}\ \mu_\varepsilon(d\lambda), \tag{6.7}$$

что и доказывает следствие. □

Реализуемые процессы авторегрессии.

Теорема 6.1. *Пусть* $Q^*(z) \neq 0$ *при* $|z| = 1$. *Для того, чтобы процесс авторегрессии* (6.6) *был реализуемым, необходимо и достаточно, чтобы все корни полинома* $Q^*(z)$ *лежали вне единичного круга.*

Доказательство. Запишем условие реализуемости в явном виде. Скалярно умножив (6.6) 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 $k \geq 1$

$$b_0 R_\xi(k) + b_1 R_\xi(k-1) + \ldots + b_q R_\xi(k-q) = 0. \tag{6.8}$$

Поскольку

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

а коэффициенты полиномов вещественны, то (6.8) 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 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 от противного 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, обозначив H производную функции Q^*, получим, что

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

при любом $k \geq 1$. Этого, конечно, быть не может, так как вектора $(1, z_\ell, \ldots, z_\ell^s)^{\mathrm{T}}$ линейно независимы при достаточно больших s. В случае кратных корней доказательство сохраняется, но линейно независимые вектора выглядят более сложно. □

Замечание 6.2. В терминах полинома Q условия Теоремы 6.1 означают, что все корни этого полинома лежат внутри единичного круга.

Следствие 6.2. Пусть $Q^*(z) \neq 0$ при $|z| = 1$. Процесс авторегрессии (6.1) 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 (6.4). Тогда полином $Q^*(z)$ не имеет корней при $|z| \leq 1$ и верно равенство (6.7). Поэтому $\xi_n = \sum_k c_k\varepsilon_{n-k}$, где

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

и

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

при $k < 0$. □

Предложение 6.4. *Пусть $b_0, \ldots, b_q$ и $a_0, \ldots, a_p$ — некоторые числа, причем $b_0 b_q a_0 a_q \neq 0$. Обозначим $Q^*(z) = \sum_{k=0}^{q} b_k z^k$ и $P^*(z) = \sum_{k=0}^{p} a_k z^k$. Если $Q^*(z) \neq 0$ при $|z| = 1$, то для любого белого шума ε_n существует стационарная последовательность АРСС(q,p), реализуемая относительно ε_n и имеющая спектральную плотность* (6.3).

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 (6.7).

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. □

Авторегрессия первого порядка и марковские гауссовские стационарные последовательности. Рассмотрим вещественный реализуемый процесс чистой авторегрессии первого порядка, который можно записать в форме

$$\xi_n = a\xi_{n-1} + b\varepsilon_n, \tag{6.9}$$

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 (6.9) 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 (6.9) на ε_n, получим 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

$$\xi_n = \rho\xi_{n-1} + \sigma\sqrt{1-\rho^2}\,\varepsilon_n, \quad |\rho| \le 1, \quad \sigma > 0, \tag{6.10}$$

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 6.1. В нашем случае полином Q^* имеет вид $Q^*(z) = 1 - \rho z$, то есть условие Предложения 6.1 приобретает вид $|\rho| \ne 1$. С учетом того, что мы исследуем реализуемые процессы, приходим к выводу, что необходимое и достаточное условием существования спектральной плотности у процесса авторегрессии (6.10) является неравенство $|\rho| < 1$, и при этом

$$f_\xi(\lambda) = \frac{1}{2\pi}\,\frac{\sigma^2(1-\rho^2)}{|1-\rho e^{i\lambda}|^2} = \frac{1}{2\pi}\,\frac{\sigma^2(1-\rho^2)}{1+\rho^2-2\rho\cos(\lambda)}\,. \tag{6.11}$$

Легко видеть, что спектральная плотность (6.11) 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 (6.10) 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 (6.10) как линейного преобразования белого белого шума ε_n. Разлагая функцию $1/Q^*(e^{-i\lambda})$, как это обозначено в Следствии 6.2, получим, что

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

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

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

Предложение 6.5. *Пусть ξ_n, $n = 0, 1, \ldots,$ — вещественная гауссовская стационарная последовательность с нулевым средним. Для того, чтобы она была марковской, необходимо и достаточно, чтобы ξ_n был 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

$$\xi_n = a\xi_{n-1} + b\varepsilon_n, \tag{6.12}$$

где ε_n ортогонален ξ_m при $m = 0, \ldots, n-1$ и $\mathbb{E}\varepsilon_n^2 = 1$. Если в рекуррентной формуле (6.12) 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. $\square$

6.1.2 Уравнения Юла-Уолкера

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

Лемма 6.1. *Если матрица Σ_{p+1} неотрицательно определена, а матрица Σ_p положительно определена, то существует единственное решение $A=(a_p,\ldots,a_1)^{\mathrm{T}}$ системы*

$$R(m)+\sum_{k=1}^{p}a_kR(m-k)=0,\quad m=1,\ldots,p\ , \tag{6.13}$$

причем это решение удовлетворяет неравенству

$$\sigma^2\stackrel{\mathrm{def}}{=}R(0)+\sum_{k=1}^{p}a_kR(k)\geq 0. \tag{6.14}$$

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 обозначим $A=(a_1,\ldots,a_p)^{\mathrm{T}}$. Тогда система (6.13) перепишется в виде $B+\Sigma_pA=\mathbf{0}$. Поскольку матрица Σ_p — невырожденная, то единственность вектора A очевидна.

Нам будет удобно при изучении (6.13) и (6.14) перейти на вероятностный язык. Возьмем случайный вектор $(\eta_0,\ldots,\eta_p)^{\mathrm{T}}$ с распределением $\mathbb{N}(\mathbf{0},\Sigma_{p+1})$ и условное математическое ожидание $\mathbb{E}(\eta_p\,|\,\overline{\eta}_{p-1})$ с $\overline{\eta}_{p-1}=(\eta_0,\ldots,\eta_{p-1})^{\mathrm{T}}$, которое, как хорошо известно, является линейной функцией случайных величин $\eta_0,\ldots,\eta_{p-1}$. Запишем этот факт в виде равенства

$$\mathbb{E}(\eta_p\,|\,\eta_0,\ldots,\eta_{p-1})=\sum_{k=1}^{p}c_k\eta_{p-k}.$$

Тогда разность

$$e_p\stackrel{\mathrm{def}}{=}\eta_p-\sum_{k=1}^{p}c_k\eta_{p-k} \tag{6.15}$$

будет независима со случайными величинами η_{p-m} для $m=1,\ldots p$. Домножив правую часть (6.15) на η_{p-m} и взяв математическое ожидание, мы получим в точности систему (6.13) с $c_k=-a_k$, так как $\mathbb{E}\eta_{p-k}\eta_{p-m}=R(m-k)$. Таким образом, $e_p=\eta_p+A^{\mathrm{T}}\overline{\eta}_{p-1}$ и

$$0\leq\mathbb{D}\eta_p=R(0)-\mathbb{E}A^{\mathrm{T}}\overline{\eta}_{p-1}\overline{\eta}_{p-1}^{\mathrm{T}}A=R(0)-A^{\mathrm{T}}\Sigma_pA=R(0)+B^{\mathrm{T}}\Sigma_p^{-1}\Sigma_pA=R(0)+B^{\mathrm{T}}A,$$

так как $A=-\Sigma_p^{-1}B$. Тем самым (6.14) доказано. Если матрица Σ_{p+1} невырождена, то «остаточная дисперсия» $\mathbb{D}\eta_p$ должна быть строго положительна. □

Лемма 6.2. *Пусть $R(n)$ — ненулевая вещественная неотрицательно определенная функция и m — соответствующая спектральная мера. Если m обладает спектральной плотностью f, то для любого $p\geq 1$ 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

$$\sum_{k,l=0}^{p-1}R(k-l)z_k\bar{z}_l=\int_{-\pi}^{\pi}\left(\sum_{k,l=0}^{p-1}e^{i(k-l)\lambda}z_k\bar{z}_l\right)f(\lambda)d\lambda=\int_{-\pi}^{\pi}\left|\sum_{k=0}^{p-1}e^{ik\lambda}z_k\right|^2f(\lambda)d\lambda.$$

Поскольку функция

$$g(\lambda) = \left| \sum_{k=0}^{p-1} e^{ik\lambda} z_k \right|^2$$

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,

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

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

Рассмотрим реализуемый вещественный авторегрессионный процесс порядка p:

$$\xi_n + \sum_{k=1}^{p} a_k \xi_{n-k} = \sigma \varepsilon_n, \tag{6.16}$$

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 $m > n$. Число p определяется условием $a_p \neq 0$. Коэффициенты a_m предполагаются вещественными.

Спектральная плотность процесса авторегрессии (6.16) 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 6.6 (Уравнения Юла-Уолкера)**.** *Пусть $R_\xi(n)$ — ковариационная функция процесса* (6.16)*. Тогда*
1. При $m \ge 1$

$$R_\xi(m) + \sum_{k=1}^{p} a_k R_\xi(m-k) = 0. \tag{6.17}$$

2. Числа $a_1, \ldots, a_p$ однозначно определяются системой уравнений (6.17) *при $m = 1, \ldots, p$ и удовлетворяют неравенству*

$$R_\xi(0) + \sum_{l=1}^{p} a_l R_\xi(l) = \sigma^2 > 0. \tag{6.18}$$

Доказательство. Первое утверждение уже доказано в Теореме 6.1. Докажем второе утверждение. Умножим обе части (6.16) на ξ_n и возьмем математическое ожидание. Получим, что

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

С другой стороны, домножая (6.16) 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 6.3. 1. Положительность левой части (6.18) непосредственно следует из лемм 6.1, 6.2.
2. Система уравнений (6.17) при $1 \le m \le p$ называется *системой Юла-Уолкера.*

6.1.3 Авторегрессионное продолжение стационарных последовательностей

Теорема об авторегрессионном продолжении.

Предложение 6.7. *Пусть η_n $(n \geq 0)$ — вещественный стационарный в широком смысле процесс с ковариационной функцией $R_\eta(m)$ и спектральной плотностью $f_\eta(\lambda)$. Тогда для любого $p \geq 0$ существует реализуемый процесс авторегрессии порядка p такой, что $R_\xi(m) = R_\eta(m)$ при $m = 0, \ldots, p-1$.*

Доказательство. Конечно, можно считать, что $\mathbb{E}\eta_n = 0$.

Согласно Лемме 6.2, для любого p матрица $\Sigma_{p+1} = \left\{R_\eta(l-k)\right\}_{k,l=1}^{p+1}$ является положительно определенной. Поэтому (см. Лемму 6.1) ковариации $R_\eta(n)$ удовлетворяют равенствам

$$R_\eta(m) + \sum_{k=1}^{p} a_k R_\eta(m-k) = 0, \quad m = 1, \ldots, p, \tag{6.19}$$

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

$$\sigma_p^2 \stackrel{\text{def}}{=} R_\eta(0) + \sum_{k=1}^{p} a_k R_\eta(k) > 0. \tag{6.20}$$

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

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

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

Нам достаточно показать, что для любых n, l таких что $\max(n,l) \geq p$, математическое ожидание $\mathbb{E}\xi_n\xi_l$

a) зависит только от модуля разности $|n-l|$ и

b) при $|n-l| < p$ совпадает с $R_\eta(|n-l|)$.

Будем доказывать этот факт индукцией по $s = \max(n,l) \geq p$. Иначе говоря, индукционное предположение A_s будет состоять в том, что условие *a*) выполняется для всех (n,l) с $\max(n,l) < s$, а условие *b*) — для всех (n,l) с $\max(n,l) < s$ и $|n-l| < p$. Отметим, что при $\max(n,l) < p$ оба эти условия выполняются автоматически.

Ясно, что при $n \geq p$ и $l < n$

$$\mathbb{E}\xi_n\xi_l = -\sum_{k=1}^{p} a_k \mathbb{E}\xi_{n-k}\xi_l. \tag{6.21}$$

С другой стороны, при $n \geq p$

$$\mathbb{E}\xi_n^2 = -\sum_{k=1}^{p} a_k \mathbb{E}\xi_n\xi_{n-k} + \sigma_p^2 \tag{6.22}$$

(этот факт доказывается точно такими же рассуждениями, как равенство (6.18) в Предложении 6.6). Кроме того, по построению $\mathbb{E}\xi_n\xi_l = \mathbb{E}\eta_n\eta_l = R_\eta(|n-l|)$ при $0 \leq n, l < p$.

Возьмем $n = p$ и $0 \leq l < n$. Так как $p - k \in \{0, \ldots, p-1\}$ при $1 \leq k \leq p$, то в этом случае $\mathbb{E}\xi_{p-k}\xi_l = \mathbb{E}\eta_{p-k}\eta_l = R_\eta(p-l-k)$ и формула (6.21) показывает, что $\mathbb{E}\xi_p\xi_l$ зависит только от $n-l$, причем согласно (6.19)

$$\mathbb{E}\xi_p\xi_l = -\sum_{k=1}^{p} a_k R_\eta(p-l-k) = R_\eta(p-l), \tag{6.23}$$

так как $1 \leq p-l \leq p$.

Аналогичным образом из (6.22), (6.23) и (6.20) вытекает, что

$$\mathbb{E}\xi_p^2 = -\sum_{k=1}^{p} a_k \mathbb{E}\xi_p\xi_{p-k} + \sigma_p^2 = -\sum_{k=1}^{p} a_k R_\eta(p-k) + \sigma_p^2 = R_\eta(0). \tag{6.24}$$

Следовательно, база индукции при $s = p$ проверена. Так как $\max(n,l) > \max(n-k,l)$ при $n > p$, $n > l$ и $1 \le k \le p$, то индукционный переход для доказательства того, что $\mathbb{E}\xi_n\xi_l$ зависит только от $|n-l|$, мгновенно следует для $n \ne l$ из равенства (6.21).

При $n = l \ge p$ применение аналогичных рассуждений к равенству (6.22) показывает, что $\mathbb{E}\xi_n^2$ не зависит от $n \ge p$. Следовательно (см. (6.24)), $\mathbb{E}\xi_n^2 = R_\eta(0)$.

Таким образом, осталось доказать, что $\mathbb{E}\xi_n\xi_{n-m} = R_\eta(m)$ при $n \ge p$ и $1 \le m \le p-1$. Применяя равенство (6.21), мы видим, что

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

Поскольку $\max(n-m, n-k) < n$ и $|(n-k)-(n-m)| = |m-k| \le p-1$, то применение индукционного предположения дает

$$\mathbb{E}\xi_n\xi_{n-m} = -\sum_{k=1}^{p} a_k R_\eta(m-k) = R_\eta(m).$$

Утверждение доказано. □

Замечание 6.4. Как доказано в Предложении 6.7, процесс ξ_n является реализуемым процессом 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 (6.17).

Замечание 6.5. Предположим, что процесс η_n Предложения 6.7 является реализуемым процессом авторегрессии порядка d:

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

Если в Предложении 6.7 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 (6.19) будут равны b_k, а число σ_p^2 в (6.20) равно σ^2. Поэтому последовательность (2.6) 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 (2.6) имеет порядок, меньший d.

Рассмотрим случай $p > d$. Согласно (6.17),

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

при $m \ge 1$. В то же время выполнено (6.19), 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 $R_\eta(n)$.

Спектральные плотности авторегрессионных продолжений. Пусть $p \to \infty$. Рассмотрим последовательность $\{R_\xi^{(p)}(m), m \ge 0\}$ 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 6.3. *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 $\mathcal{Q}_n = \mathcal{L}(\eta_n)$ сосредоточено на $(0,1]$, а распределение $\mathcal{Q} = \mathcal{L}(\eta)$ — на $(0,1)$. Кроме того,

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

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 следует, что $\mathcal{Q}_n \Rightarrow \mathcal{Q}$. Поэтому (так как линейное преобразование непрерывно) $\mathcal{P}_n \Rightarrow \mathcal{P}$ и утверждение доказано. □

Следствие 6.3. 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 6.8. *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 $0 < m^2 \le f(\lambda) \le M^2$. Обозначим $h(\lambda) = 1/\sqrt{f(\lambda)}$. Тогда $1/M \le h(\lambda) \le 1/m$. 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 $1/M - \varepsilon \le Q(\lambda) \le 1/m + \varepsilon$.

Поскольку всегда можно представить 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 6.8 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 6.3, 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 [13, разд. 3.5].

Теорема 6.2. *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.

Авторегрессионные продолжения и система ортогональных полиномов. Материал этого раздела тесно связан с теорией ортогональных полиномов на окружности, восходящей к [16, 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

$$\langle \theta_1, \theta_2 \rangle_f = \frac{1}{2\pi} \int\limits_{-\pi}^{\pi} \theta_1(e^{i\lambda}) \overline{\theta_2(e^{i\lambda})} f(\lambda) d\lambda. \tag{6.25}$$

Соответствующее гильбертово пространство обозначим $\mathbb{L}_f^2$. 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 $p \ge 1$ рассмотрим систему

$$R(m) + \sum_{k=1}^{p} a_{p,k} R(m-k) = 0, \quad m = 1, \ldots, p, \tag{6.26}$$

которая, согласно Лемме 6.1, 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 6.4. *Для полиномов Q_p выполнены равенства*

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

Доказательство. Равенство $\|Q_0\|_f^2 = \sigma_0^2/2\pi$ выполнено по определению. Перепишем (6.26) в виде

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

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

Далее, при $p \geq 1$

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

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

Лемма доказана. □

Полиномы Q_p называют полиномами Сеге.

Замечание 6.6. Обозначим $\phi_p = Q_p/\|Q_p\|_f$. Тогда спектральная плотность авторегрессионного продолжения $\xi^{(p)}$ равна

$$f_\xi^{(p)}(\lambda) = \frac{1}{2\pi}\frac{\sigma_p^2}{\left|1+\sum_{k=1}^{p} a_k e^{-ik\lambda}\right|^2} = \frac{\|Q_p\|_f^2}{|Q_p(e^{-i\lambda})|^2} = \frac{\|Q_p\|_f^2}{\|Q_p\|_f^2\,|\phi_p(e^{-i\lambda})|^2} = \frac{1}{|\phi_p(e^{-i\lambda})|^2}\ . \tag{6.27}$$

Отметим некоторые свойства полиномов Q_p и ϕ_p (например,[16, гл. 11]).

Предложение 6.9. *1. Пусть ψ_p, $p = 0, 1, \ldots$ — некоторая ортонормированная в $\mathbb{L}_f^2$ система функций, причем ψ_p является полиномом степени p. Тогда $\psi_p = \pm\phi_p$.*
2. Если полином P степени p ортогонален в $\mathbb{L}_f^2$ любому полиному степени, меньшей p, то $P(z) = cQ_p(z)$.
3. Если полином P степени p ортогонален в $\mathbb{L}_f^2$ любому моному z^j при j от 1 до p, то $P(z) = cz^pQ_p(1/z)$.
4. Полиномы Q_p при $|z| = 1$ удовлетворяют рекуррентным соотношениям

$$Q_{p+1}(z) = zQ_p(z) + a_{p+1,p+1}\, z^pQ_p(1/z). \tag{6.28}$$

Доказательство. 1. Очевидно, $\psi_0 = \phi_0$. Пусть $\psi_p = \phi_p$ при $p < p_0$. Тогда существуют такие ненулевые α, β, что $\gamma_{p_0} \stackrel{\text{def}}{=} \alpha\psi_{p_0} - \beta\phi_{p_0}$ является полиномом степени, меньшей, чем p_0. Поскольку полином γ_{p_0} ортогонален ϕ_p при $p < p_0$, то $\alpha\psi_{p_0} = \beta\phi_{p_0}$. Наконец, так как $\|\psi_{p_0}\|_f = \|\phi_{p_0}\|_f$, то $|\alpha| = |\beta|$.
2. Это свойство становится очевидным, если рассмотреть разложение $P = \sum_{k=0}^{p} c_kQ_k$.
3. Для доказательства заметим, что для любых полиномов P_1, P_2 и любого $k \geq 1$ 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 равенства $\left\langle P(z), z^j\right\rangle_f = 0$ следует, что $\left\langle z^pP(1/z), z^{p-j}\right\rangle_f = 0$ для $j = 1, \ldots, p-1$. Так как $z^pP(1/z)$ — полином степени p, ортогональный z^k для $k = p-j = 0, \ldots, p-1$, то $z^pP(1/z) = cQ_p(z)$. Отсюда сразу же следует требуемое.
4. По определению, $Q_{p+1}(z)$ ортогонален z^j при $1 \leq j \leq p$. 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 тех же j. Следовательно, полином $Q_{p+1}(z) - zQ_p(z)$ ортогонален z^j при всех $1 \leq j \leq p$. Применяя утверждение п. 3 видим, что полином $Q_{p+1}(z) - zQ_p(z)$ пропорционален $z^pQ_p(1/z)$. Приравнивая коэффициенты при степенях этих полиномов, получаем вид постоянной c. □

Равенства (6.28) 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 (6.28) и учитывая, что

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

получим, что при $1 \le k \le p$

$$a_{p+1,k} = a_{p,k} + a_{p+1,p+1}\, a_{p,p+1-k}, \tag{6.29}$$

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 (6.29) 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 $a_{p+1,k}$. Этот прием в статистике стационарных процессов носит название *алгоритма Левинсона*, см., например, [11, разд. 3.8.1].

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

Приведем пример других продолжений, следуя [4], где используется несколько более аналитический язык.

Поскольку матрица $\Sigma_{p+1} = \big\{R_\eta(l-k)\big\}_{k,l=0}^{p}$ 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 (6.19) и (6.20) мы получим

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

и

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

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 $\xi_0^{(\mu)},\ldots,\xi_p^{(\mu)}$ 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 6.7, мы получим, что $\xi_n^{(\mu)}$ является авторегрессионным стационарным процессом, а его ковариационная функция $R_\xi^{(\mu)}(n)$ является продолжением отрезка $R_\eta^{(\mu)}(0),\ldots,R_\eta^{(\mu)}(p)$. 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 $R_\eta(0),\ldots,R_\eta(p)$.

Замечание 6.7. Спектральная плотность процесса $\zeta_n^{(\mu)} = \xi_n^{(\mu)} + \varepsilon_n^{(\mu)}$ имеет вид

$$f_\zeta^{(\mu)}(\lambda) = \mu + \frac{1}{2\pi}\,\frac{\sigma_p^2(\mu)}{\big|1 + \sum_{k=1}^{p} a_k^{(\mu)} e^{-ik\lambda}\big|^2}\ , \quad |\lambda| \le \pi. \tag{6.30}$$

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

Теорема 6.3 (Колмогоров, см. [5], 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*

$$\eta_n = \sum_{k=-\infty}^{n} c_{n+k}\varepsilon_k, \tag{6.31}$$

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

$$S(f) = S_\eta(f) \stackrel{\text{def}}{=} \int_{-\pi}^{\pi} \ln f_\eta(\lambda) d\lambda > -\infty. \tag{6.32}$$

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

Процесс, удовлетворяющий (6.31), называется *регулярным* относительно белого шума ε_n. Отметим, что реализуемые процессы авторегрессии, как и стационарные процессы со спектральной плотностью (6.30), удовлетворяют неравенству (6.32).

Теорема 6.4 (Колмогоров-Сегё, см. [6], 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* (6.31)*, то*

$$\mathbb{E}\big(\eta^*_{n+1} - \eta_{n+1}\big)^2 = 2\pi e^{S_\eta(f)/2\pi}. \tag{6.33}$$

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 (6.32). Значит, для них будут выполнены аналоги (6.33).

Начиная с работ Берга [7], [8] (в интернете доступна диссертация Берга [9], 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 [10] (цитируется по [11, гл. 7.4]), принцип максимальной случайности можно сформулировать таким образом:

если мы делаем выводы на основе неполной информации, то должны опираться[9] *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 (6.33) к следующей задаче: *найти для фиксированного p такое неотрицательно определенное продолжение отрезка $R_\eta(0), \ldots, R_\eta(p)$, для которого энтропийный интеграл* (6.32) *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 [12].

С собственно энтропией задачу максимизации функционала (6.32) 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. [10]), приведем некоторые ее свойства, необходимые для изложения.

[9] далее в [10]: "…на такое вероятностное распределение, которое имеет максимальную энтропию, допускаемую нашей априорной информацией."

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

$$H_\eta = H_\xi + \ln(|\det A|). \tag{6.34}$$

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

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

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 (6.34) 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

$$H_n = \frac{n}{2}\,\ln(2\pi e) + \frac{1}{2}\,\ln\left(\det\Sigma_n\right). \tag{6.35}$$

Оказывается (при необременительных ограничениях), в этом случае удельная энтропия существует и имеет достаточно простой вид, связанный с энтропийным интегралом (6.32). А именно, имеет место следующее утверждение.

Теорема 6.5. *Пусть гауссовская стационарная последовательность η имеет ограниченную сверху и отделенную от нуля спектральную плотность f. Тогда удельная энтропия этой последовательности h существует и*

$$h = \frac{1}{2}\,\ln(2\pi e) + \frac{1}{4\pi}\int_{-\pi}^{\pi} \ln\bigl(2\pi f(\lambda)\bigr)d\lambda = 1/2 + \ln(2\pi) + \frac{1}{4\pi}\,S_f, \tag{6.36}$$

где S_f определено в (6.32).

Доказательство. Ограничимся лишь изложением хода доказательства теоремы 6.5. 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 $\Sigma_n = \bigl\{R(k-l)\bigr\}_{k,l=0}^{n-1}$ 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 6.6 (Сегё, [13], 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 [14, 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

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

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

Лемма 6.5 ([14], 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 (6.32) и (6.35), получаем (6.36). □

Таким образом, задачи максимизации дисперсии наилучшего прогноза (6.33) 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 ковариаций процесса известны, является авторегрессионное продолжение ковариационной функции процесса. Кстати, технически простое и основанное на теоретико-информационных соображениях доказательство этого факта можно найти в [12].

6.2 Метод MEM

Пусть $x_0, \ldots, x_{N-1}$ — отрезок реализации вещественного стационарного процесса с нулевым средним и ковариационной функцией $R(m)$, обладающей спектральной плотностью f. Общая схема применения рассуждений предыдущего раздела к авторегрессионному статистическому оцениванию спектральной плотности f вполне прозрачна и состоит из нескольких шагов:

1. Выбор порядка p аппроксимирующей авторегрессии. Эта авторегрессия имеет вид
$$\xi_n = -\sum_{k=1}^{p} a_k \xi_{n-k} + \sigma\varepsilon_n.$$
2. Построение оценок $\widehat{R}(m)$ ковариационной функции $R(m)$ при $0 \le m \le p$;
3. Нахождение коэффициентов $\widehat{a}_k$ из выборочного аналога системы Юла-Уолкера
$$\widehat{R}(m) + \sum_{k=1}^{p} \widehat{a}_k \widehat{R}(m-k) = 0, \quad m = 1, \ldots, p\,; \tag{6.37}$$
4. Вычисление коэффициента $\widehat{\sigma}^2$ по формуле
$$\widehat{\sigma}^2 = \widehat{R}(0) + \sum_{k=1}^{p} \widehat{a}_k \widehat{R}(k); \tag{6.38}$$
5. Наконец, получение оценки спектральной плотности f:
$$\widehat{f}(\lambda) = \frac{1}{2\pi} \frac{\widehat{\sigma}^2}{\left|1 + \sum_{k=1}^{p} \widehat{a}_k e^{-ik\lambda}\right|^2}\ , \quad |\lambda| \le \pi. \tag{6.39}$$

   В случае, когда нас интересует только форма оцениваемой спектральной плотности (например, расположение ее острых пиков на оси λ), можно использовать более простую статистику
$$\widetilde{f}(\lambda) = \frac{1}{2\pi} \frac{1}{\left|1 + \sum_{k=1}^{p} \widehat{a}_k e^{-ik\lambda}\right|^2}\ , \quad |\lambda| \le \pi. \tag{6.40}$$

6.2.1 Случай фиксированного p

Будем считать, что число p зафиксировано и обсудим статистические свойства этой процедуры.

Согласно Лемме 6.2 матрица $\Sigma_{p+1} = \left\{R(k-l)\right\}_{l,k=0}^{p}$ является неотрицательно определенной и, следовательно, система (6.13) имеет единственное решение, а число σ^2, определенное в (6.14), положительно. Поэтому, если $\widehat{R}(m)$ при $0 \le m \le p$ являются состоятельными оценками $R(m)$, то при $N \to \infty$ матрица $\widehat{\Sigma}_{p+1} = \left\{\widehat{R}(k-l)\right\}_{l,k=0}^{p}$ будет положительно определена с вероятностью, стремящейся к 1.

Следовательно, при больших N с большой вероятностью система (6.37) будет иметь единственное решение, а число $\widehat{\sigma}^2$, определенное в (6.38), будет положительно.

Более того, как нетрудно видеть, при этих условиях $\widehat{a}_k \xrightarrow{\mathbb{P}} a_k$ и $\widehat{\sigma}^2 \xrightarrow{\mathbb{P}} \sigma^2$. Это значит, что при фиксированном p, $n \to \infty$ и состоятельных оценках $\widehat{R}(m)$, $m = 0, \ldots, p$, описанная процедура будет состоятельно оценивать спектральную плотность
$$f(\lambda) = \frac{1}{2\pi} \frac{\sigma^2}{\left|1 + \sum_{k=1}^{p} a_k e^{-ik\lambda}\right|^2}\ , \quad |\lambda| \le \pi,$$
авторегрессионного процесса, обсуждавшегося в предыдущем разделе.

Однако одной состоятельности оценок $\widehat{R}(m)$ не достаточно как для того, чтобы система (6.37) имела единственное решение, так и для того, чтобы правая часть (6.38) была положительна при фиксированном N.

Наиболее популярной оценкой ковариации $R(m)$ по реализации $x_0, \ldots, x_{N-1}$ является

$$\widehat{R}_1(m) = \widehat{R}_1(-m) = \frac{1}{N-m} \sum_{k=0}^{N-m-1} (x_k - \bar{x})(x_{k+m} - \bar{x}). \tag{6.41}$$

Однако матрица $\widehat{\Sigma}^{(1)}_{p+1}$, построенная по таким оценкам ковариационной функции, не будет, вообще говоря, неотрицательно определенной. Тем не менее оценки $\widehat{R}_1(m)$ легко исправить (несколько ухудшив), обеспечив неотрицательную определенность нужной нам ковариационной матрицы.

Докажем сначала следующую лемму.

Лемма 6.6. *Пусть $z_0, \ldots, z_{N-1}$ — вещественные числа и $1 \le L < N$. Обозначим*

$$u_m = \sum_{k=0}^{N-m-1} z_k z_{k+m}$$

и введем Теплицеву $L \times L$ матрицу $\mathbf{S}$ с элементами $s_{ij} = u_{|i-j|}$, где $i, j = 0, \ldots, L-1$. Тогда матрица $\mathbf{S}$ будет неотрицательно определена.

Доказательство. Обозначим

$$z_k^* = \begin{cases} z_k & \text{при } \ 0 \le k \le N-1, \\ 0 & \text{иначе} \end{cases}$$

и рассмотрим $L \times (N+L-1)$ матрицу $\mathbf{X}$ со строками

$$X_k = (z^*_{-k}, z^*_{-k+1}, \ldots, z^*_{N+L-2-k}), \quad k = 0, \ldots, L-1.$$

Например, при $N=4$ и $L=3$ эта матрица имеет вид

$$\mathbf{X} = \begin{pmatrix} z_0 & z_1 & z_2 & z_3 & 0 & 0 \\ 0 & z_0 & z_1 & z_2 & z_3 & 0 \\ 0 & 0 & z_0 & z_1 & z_2 & z_3 \end{pmatrix}.$$

Пусть $0 \le i \le j \le L-1$. Если показать, что $X_i X_j^{\mathrm{T}} = u_{j-i}$, то все будет доказано, так как получится, что $\mathbf{S} = \mathbf{X}\mathbf{X}^{\mathrm{T}}$.

Заметим, что $z^*_{-j+l} = 0$ при $l < j$ и $l \ge N+j$. В остальных случаях $z^*_{-j+l} = z_{l-j}$. Так как $j \ge i$, то $z^*_{-i+l} = z_{l-i}$ при $l \ge j$ и $l \le N+i-1$. Следовательно,

$$X_i X_j^{\mathrm{T}} = \sum_{l=0}^{N-L-2} z^*_{-i+l} z^*_{-j+l} = \sum_{l=j}^{N+i-1} z_{-i+l} z_{-j+l} = \sum_{m=0}^{N+i-j-1} z_m z_{m+j-i}.$$

Утверждение доказано. □

Если теперь в Лемме 6.6 положить $L = p+1$ и $z_k = x_k - \bar{x}$, то мы мгновенно получим, что в обозначениях

$$\widehat{R}_2(m) = \widehat{R}_2(-m) = \frac{1}{N} \sum_{k=0}^{N-m-1} (x_k - \bar{x})(x_{k+m} - \bar{x}) = \frac{N-m}{N} \widehat{R}_1(m), \tag{6.42}$$

матрица $\widehat{\Sigma}^{(2)}_{p+1} = \left\{\widehat{R}_2(l-k)\right\}^p_{l,k=0}$ будет неотрицательно определенной. На практике, конечно, неотрицательная определенность перейдет в положительную определенность.

Обсудим статистические свойства оценок $\widehat{R}_1(m)$ и $\widehat{R}_2(m)$. Хорошо известно (например, [1, разд. 8.3.2]), что при уже обсуждавшемся условии $\sum_{k=-\infty}^{\infty} |R(k)| < \infty$ для фиксированного m

$$\mathbb{E}\widehat{R}_1(m) - R(m) \sim -\frac{1}{N} \sum_{k=-\infty}^{\infty} R(m) = -2\pi f(0)/N.$$

Соответственно,

$$\mathbb{E}\widehat{R}_2(m) - R(m) = \frac{N-m}{N}\Big(\mathbb{E}\widehat{R}_1(m) - R(m)\Big) + \frac{m}{N}\ R(m) \sim \big(-2\pi f(0) + m\big)/N.$$

Видно, что оценки $\widehat{R}_1(m)$ и $\widehat{R}_2(m)$ асимптотически несмещенны, но мало пригодны для оценивания $a_1, \ldots, a_p$ и σ^2 при относительно небольших N.

Поэтому дальше мы будем рассматривать только случай больших N. Для построения авторегрессионных оценок спектральной плотности в случае малых N используются другие методы оценивания авторегрессионных коэффициентов $a_1, \ldots, a_p$. С обзором и сравнением таких методов можно ознакомиться в [11, гл. 8]. Отметим, что в примере раздела 8.6 этой книги для иллюстрации применения некоторых из этих алгоритмов (см. рис. 8.5) выбран процесс авторегрессии порядка 4 при очень небольшом значении $N = 40$.

Рассмотрим предельные распределения оценок $\widehat{R}_1(m)$ и $\widehat{R}_2(m)$. Кроме того, поскольку в уравнениях (6.37) можно заменить оценки ковариационной функции $\widehat{R}(m)$ на оценки корреляционной функции $\widehat{r}(m) = \widehat{R}(m)/\widehat{R}(0)$, то нас могут интересовать предельные распределения оценок $\widehat{r}_1(m)$ и $\widehat{r}_2(m)$.

Имеет место следующее утверждение ([1, теор. 8.4.2, следств. 8.4.1, разд. 8.4.5]). Обозначим

$$\Delta_1(m) = \widehat{R}_1(m) - R(m), \quad \Delta_2(m) = \widehat{R}_2(m) - R(m).$$

Аналогично, положим

$$\delta_1(m) = \widehat{r}_1(m) - r(m), \quad \delta_2(m) = \widehat{r}_2(m) - r(m).$$

Теорема 6.7. *Предположим, что процесс x_n представим в виде*

$$x_n = a + \sum_{k=-\infty}^{\infty} c_k \varepsilon_{n-k}\,,$$

где ε_k — независимые одинаково распределенные случайные величины с $\mathbb{E}\varepsilon_n = 0$, $\mathbb{E}\varepsilon_n^2 = 1$ и $\mathbb{E}\varepsilon_n^4 < \infty$. Если $\sum_{k=-\infty}^{\infty} |c_k| < \infty$, то при фиксированном p и $N \to \infty$

1.

$$\mathcal{L}\Big(\sqrt{N}\big(\Delta_1(0), \ldots, \Delta_1(p)\big)^{\mathrm{T}}\Big) \sim \mathcal{L}\Big(\sqrt{N}\big(\Delta_2(0), \ldots, \Delta_2(p)\big)^{\mathrm{T}}\Big) \Longrightarrow \mathrm{N}(\mathbf{0}, W_{p+1}),$$

где $W_{p+1} = \big\{w_{l,k}\big\}_{l,k=0}^{p}$ и

$$w_{l,k} = 4\pi \int_{-\pi}^{\pi} \cos(\lambda l)\cos(\lambda k) f^2(\lambda) d\lambda + \left(\frac{\mathbb{E}\varepsilon_n^4}{s^4} - 3\right)\ R(l)R(k). \tag{6.43}$$

2.

$$\mathcal{L}\Big(\sqrt{N}\big(\delta_1(0), \ldots, \delta_1(p)\big)^{\mathrm{T}}\Big) \sim \mathcal{L}\Big(\sqrt{N}\big(\delta_2(0), \ldots, \delta_2(p)\big)^{\mathrm{T}}\Big) \Longrightarrow \mathrm{N}(\mathbf{0}, H_{p+1}),$$

где $H_{p+1} = \big\{h_{l,k}\big\}_{l,k=0}^{p}$ и

$$h_{l,k} = \frac{4\pi}{R(0)} \int_{-\pi}^{\pi} \big(\cos(\lambda l) - r_l\big)\big(\cos(\lambda k) - r_k\big) f^2(\lambda) d\lambda. \tag{6.44}$$

Замечание 6.8. Как видно из (6.43), предельные ковариации выборочных ковариаций оценок $R_1(m)$ (и $R_2(m)$) зависят не только от спектральной плотности f, но и от распределения белого шума ε_n. Если этот шум гауссовский, то $\mathbb{E}\varepsilon_n^4 = 3s^4$ и эта зависимость пропадает. В то же время для корреляций, как показывает (6.44), такие неприятности отсутствуют.

Используя результат (и метод доказательства) Теоремы 6.7, можно доказать следующую теорему (см. [15]).

Теорема 6.8. *Предположим, что процесс x_n представим в виде*

$$x_n = a + \sum_{k=0}^{\infty} c_k \varepsilon_{n-k}\,,$$

где ε_k — независимые одинаково распределенные случайные величины с $\mathbb{E}\varepsilon_n = 0$, $\mathbb{E}\varepsilon_n^2 = 1$ и $\mathbb{E}\varepsilon_n^4 < \infty$. Обозначим $R(m)$ ковариационную функцию процесса x_n и пусть при некотором фиксированном p числа a_k и σ^2 определятся равенствами (6.13) и (6.14). Кроме того, положим

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

Пусть теперь $\widehat{R}(m)$ — оценка (6.42) ковариационной функции, построенная при $m=0,\ldots,p$ по отрезку $x_0,\ldots,x_{N-1}$ ряда x_n. Определим $\widehat{a}_1,\ldots,\widehat{a}_p$, $\widehat{\sigma}^2$ и $\widehat{f}_p(\lambda)$ равенствами (6.37), (6.38) и (6.39). Обозначим $\Delta\widehat{A} = (\widehat{a}_1 - a_1,\ldots,\widehat{a}_p - a_p)^{\mathrm{T}}$.

Если $\sum_{k=0}^{\infty}|c_k|<\infty$, то при фиксированном p и $N\to\infty$
1. случайный вектор $\sqrt{N}\big(\widehat{\sigma}^2-\sigma^2,\widehat{a}_1-a_1,\ldots,\widehat{a}_p-a_p\big)^{\mathrm{T}}$ в пределе имеет $(p+1)$-мерное гауссовское распределение;
2. случайные величины $\sqrt{N}(\widehat{\sigma}^2-\sigma^2)$ и $\sqrt{N}(\widehat{a}_k-a_k)$ асимптотически некоррелированы (и, следовательно, асимптотически независимы);
3. $\mathcal{L}\big((\widehat{\sigma}^2-\sigma^2)/\sigma^2\big)\Rightarrow \mathrm{N}(0,\mathbb{E}\varepsilon_n^4-1)$;
4. $\mathcal{L}\big(\sqrt{N}\Delta\widehat{A}\big)\Rightarrow \mathrm{N}(\mathbf{0},\sigma^2\,\Sigma_p^{-1})$, где $\Sigma_p=\big\{R(l-m)\big\}_{l,m=0}^{p-1}$;
5. при $\lambda\in[0,\pi]$ конечномерные распределения процесса

$$\sqrt{N}\ \frac{\widehat{f}_p(\lambda)-f_p(\lambda)}{f_p(\lambda)} \tag{6.45}$$

слабо сходятся к конечномерным распределениям гауссовского случайного процесса с нулевым средним и ковариационной функцией

$$R_f(\lambda,\mu)=\frac{1}{4\pi^2}\left(\mathbb{E}\varepsilon_n^4-1\right)+\frac{64\pi^2}{\sigma^2}f_p(\lambda)f_p(\mu)\,C_\mu^{\mathrm{T}}\Sigma_p^{-1}C_\lambda^{\mathrm{T}}, \tag{6.46}$$

где C_λ — p-мерный вектор с координатами

$$\cos(\lambda j)+\sum_{k=1}^{p}a_k\cos\big(\lambda(k-j)\big),\quad j=1,\ldots,p.$$

Замечание 6.9. 1. Если использовать в качестве оценки ковариационной функции статистику (6.41) вместо (6.42), то формально результат Теоремы 6.8 не изменится, однако при фиксированном N сами оценки (6.37) и (6.39) будут существовать с вероятностью, вообще говоря, меньшей 1.
2. Если случайные величины ε_n имеют нормальное распределение, то $\mathbb{E}\varepsilon_n^4=3$, и формулы (6.45), (6.46) позволяют строить асимптотические доверительные интервалы для значений спектральной плотности только через оценки ковариационной функции.
в системе (6.37) можно заменить оценки ковариаций на оценки корреляций.

Список литературы

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

[2] Д. Бриллинджер (1980). Временные ряды. Обрпботка данных и теория. М., Мир.

[3] K. S. Riedel and A. Sidorenko, "Minimum bias multiple taper spectral estimation," I.E.E.E. Trans. on Signal Processing, vol. 43, pp. 188-195, Jan. 1995.

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

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

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

[7] 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.

[8] 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).

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

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

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

[12] 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.

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

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

[15] H. Akaike (1969), Power spectrum estimation through autoregressive model fitting, Ann. Inst. Statist. Math., 21, 407 – 419.

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

[17] Y. Fan, Q. Yao (2005). *Nonlinear Tillle Series. Nonparametric and Parametric Methods*, Springer, New York etc.