Санкт-Петербургский государственный университет

Прикладная математика и информатика

Отчет по учебной практике 1 (проектно-технологическая) (семестр 1)

Метод Monte Carlo SSA для процессов с длинной памятью

Выполнил:

Потешкин Егор Павлович

группа 24.М22-мм

Научный руководитель:

д. ф.-м. н., профессор

Голяндина Нина Эдуардовна

Кафедра Статистического Моделирования

Санкт-Петербург

2025

Оглавление

Введение

Метод Singular Spectrum Analysis (SSA) [1, 2] является мощным инструментом для анализа временных рядов. Он позволяет разложить ряд на интерпретируемые компоненты, такие как тренд, периодические колебания и шум, что значительно упрощает процесс анализа. Метод Monte Carlo SSA [3], в свою очередь, решает задачу обнаружения сигнала в шуме, проверяя соответствующую гипотезу.

Для наиболее распространенного варианта метода Monte Carlo SSA необходимо, чтобы спектральная плотность шума была строго монотонной. Такое ограничение связано с методом SSA и понятием строгой разделимости компонент, без которой оценить доминирующую частоту значимой компоненты не представляется возможным.

В большинстве работ, посвященных методу Monte Carlo SSA, в качестве шума используется модель красного шума — процесса авторегрессии первого порядка с положительным коэффициентом. Такая модель шума обладает строго монотонной спектральной плотностью, однако она плохо описывает временные ряды, обладающие длинной памятью, то есть ряды, автоковариционная функция которых убывает медленней, чем экспоненциальное затухание. Процессы с длинной памятью довольно распространены в реальном мире, например, в работе [4] обнаружена длинная память в таких среднегодовых гидрологических временных рядах, как количество осадков, температура и данных о речном стоке. В работе [5] на наличие длинной памяти исследовалась скорость ветра в Ирландии, в работе [6] исследовался эффект длинной памяти у сейсмических данных. Помимо геофизики, длинная память встречается также в финансах [7, 8].

Помимо проблемы выбора модели шума, существует проблема оценки параметров рассматриваемой модели. В реальных задачах редко встречается ситуация, когда параметры известны, поэтому параметры необходимо наилучшим образом оценить на основе исходного временного ряда. Неправильно оцененные параметры модели могут значительно повлиять на результат Monte Carlo SSA.

В данной работе была расширена применимость метода Monte Carlo SSA на временные ряды с длинной памятью. Также, поскольку истинные параметры модели для конкретного временного ряда неизвестны почти всегда, было произведено численное сравнение различных методов оценки параметров.

Глава 1

Теория случайных процессов

1.1. Вспомогательные определения

Для начала введем некоторые обозначения, которые будем использовать в дальнейшем.

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

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

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

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

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

Определение 1.3. Моделью $\mathrm{ARMA}(p, q)$, где $p,\ q \in \mathbb{N} \cup \{0\}$ называют случайный процесс $\{X_t\}$, удовлетворяющий соотношению

$$X_t = \varepsilon_t + \sum_{i=1}^{p} \phi_i X_{t-i} + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i},$$

где $\{\varepsilon_t\} \sim \mathrm{WN}(0, \sigma^2)$.

Замечание 1.3. Модель $\mathrm{ARMA}(p, q)$ является стационарным и обратимым процессом, если корни характеристических полиномов

$$\Phi(L) = 1 - \sum_{i=1}^{p} \phi_i L^i, \quad \Theta(L) = 1 + \sum_{i=1}^{q} \theta_i L^i$$

лежат вне единичной окружности $\{z : |z| = 1\}$ [9, Section 3.4.1].

Определение 1.4. Процесс $\{X_t\}$ называют красным шумом с параметрами ϕ и σ^2, если $\{X_t\}$ — стационарная модель ARMA(p, q) с $p = 1$, $q = 0$ и $\phi = \phi_1 \in (0, 1)$.

Определение 1.5. Спектральной плотностью стационарного процесса называется такая функция $f(\omega)$, что

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

Определение 1.6. Пусть $\{Y_t\}$ — стационарный процесс. Функцию

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

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

Замечание 1.4. Для любой фиксированной частоты ω_0

$$\mathsf{E}\left(I(\omega_0)\right) \to f(\omega_0), \quad n \to \infty;$$
$$\mathsf{D}\left(I(\omega_0)\right) \to f^2(\omega_0) \neq 0, \quad n \to \infty.$$

Таким образом периодограмма является асимптотически несмещенной, но несостоятельной, оценкой спектральной плотности [10, Раздел 4.5].

1.2. Процессы с длинной памятью

Определение 1.7. Говорят, что стационарный процесс $\{Y_t\}$ обладает длинной памятью, если

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

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

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

Существуют и альтернативные определения процессов с длинной памятью, которые можно найти в [11, Раздел 3.1]. Там же показано, что они согласованы с определением 1.7.

Пример 1.1. Процессом с короткой памятью является, например, стационарная модель ARMA(p, q), поскольку $|\gamma(h)| \leqslant CR^h$, где $C > 0$ и $0 < R < 1$ [9, Section 10.4].

Введем понятие дробного интегрирования $(1-L)^d$, где L — оператор сдвига. Например, для $d=1$ имеем $(1-L)Y_t = Y_t - Y_{t-1}$, для $d=2$ — $(1-L)^2 Y_t = Y_t - 2Y_{t-1} + Y_{t-2}$, и так далее. Обобщим этот оператор для нецелых d с помощью разложения в ряд Тейлора функции $(1-x)^d$ в нуле:

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

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

$$\pi_j(d) = (-1)^j \binom{d}{j} = \frac{j-1-d}{j}\pi_{j-1}(d) = \frac{\Gamma(j-d)}{\Gamma(j+1)\Gamma(-d)}, \tag{1.1}$$

где $\Gamma(x)$ — гамма функция. Заметим, что второе равенство в формуле (1.1) верно для любых d, третье же верно только для $d \notin \mathbb{N} \cup \{0\}$, поскольку гамма функция не определена для неположительных целых чисел.

Определение 1.8. Пусть процесс $\{Y_t\}$ определен соотношением

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

где $\pi_k(-d)$ из формулы (1.1), $\{X_t\}$ — стационарная и обратимая модель ARMA(p,d). Процесс $\{Y_t\}$ называют дробно интегрированной моделью ARMA или ARFIMA(p,d,q).

Предложение 1.1. Процесс $\{Y_t\}$ из определения 1.8 является стационарным процессом с нулевым средним. Его спектральная плотность определяется выражением

$$\begin{aligned}f_Y(\omega) &= 4^{-d} \sin^{-2d}(\pi\omega)\, f_X(\omega) \\ &= 4^{-d} \sin^{-2d}(\pi\omega)\, \sigma^2 \frac{\left|\Theta(e^{-2\pi\omega \mathrm{i}})\right|^2}{\left|\Phi(e^{-2\pi\omega \mathrm{i}})\right|^2}, \quad \omega > 0 \\ &\sim \omega^{-2d} \sigma^2 \frac{|\Theta(1)|^2}{|\Phi(1)|^2}, \quad \omega \to 0,\end{aligned} \tag{1.2}$$

где $\Phi(L)$, $\Theta(L)$ — характеристические полиномы процесса $\{X_t\}$.

Доказательство. См. [10, Proposition 6.1]. □

Замечание 1.5. Из формулы (1.2) видно, что монотонность спектральной плотности процесса $\{Y_t\}$ зависит от поведения спектральной плотности процесса $\{X_t\}$.

Следствие 1.1. В условиях предложения 1.1 при $0 < d < 1/2$

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

где

$$C_{\gamma,d} = \sigma^2 \frac{|\Theta(1)|^2}{|\Phi(1)|^2} \frac{\Gamma(1-2d)}{\Gamma(d)\Gamma(1-d)}.$$

Доказательство. См. [10, Corollary 6.1]. □

Замечание 1.6. Из следствия 1.1 сразу следует, что ARFIMA(p, d, q) с $d \in (0, 1/2)$ обладает длинной памятью.

1.3. Оценка параметров

Пусть $Y_t = (1-L)^{-d} X_t$, $d < 1/2$. Будем считать, что $\{X_t\}$ представляет собой модель ARMA(p, q) с нормально распределенным белым шумом $\{\varepsilon_t\}$. Тогда представим его спектральную плотность в параметрическом виде: $f_X(\omega) = f_X(\omega; \boldsymbol{\psi}, \sigma)$, где

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

Поставим задачу оценить параметры $\boldsymbol{\varphi}^{\mathrm{T}} = \left(d, \boldsymbol{\psi}^{\mathrm{T}}\right)$ и σ^2.

1.3.1. Maximum likelihood estimation (MLE)

Поскольку $\{\varepsilon_t\}$ — гауссовский белый шум, вектор

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

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

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

Рассмотрим логарифм функции правдоподобия. Отбрасывая аддитивные константы, получаем

$$\ell(\boldsymbol{\varphi}, \sigma^2) = -\frac{1}{2} \ln |\boldsymbol{\Sigma}_n| - \frac{1}{2} Y^{\mathrm{T}} \boldsymbol{\Sigma}_n^{-1} Y.$$

Положим $\boldsymbol{\Gamma}_n = \boldsymbol{\Sigma}_n / \sigma^2$ и, максимизируя ℓ по σ^2, получаем

$$\ell_c(\boldsymbol{\varphi}) = -\frac{n}{2} \ln\left(S(\boldsymbol{\varphi})/n\right) - \frac{1}{2} \ln g_n(\boldsymbol{\varphi}), \tag{1.3}$$

где $S(\boldsymbol{\varphi}) = \boldsymbol{Y}^{\mathrm{T}} \boldsymbol{\Gamma}_n \boldsymbol{Y}$, $g_n(\boldsymbol{\varphi}) = |\boldsymbol{\Gamma}_n|$. Тогда

$$\widehat{\boldsymbol{\varphi}}_{\mathrm{ML}} = \operatorname*{argmax}_{\boldsymbol{\varphi}} \ell_c(\boldsymbol{\varphi}), \quad \widehat{\sigma}^2_{\mathrm{ML}} = S(\widehat{\boldsymbol{\varphi}}_{\mathrm{ML}}).$$

Замечание 1.7. В случае ненулевого матожидания $\mathsf{E}Y_t = \mu$, для получения $\widehat{\boldsymbol{\varphi}}_{\mathrm{ML}}$ и $\widehat{\sigma}^2_{\mathrm{ML}}$ вместо $\boldsymbol{Y}$ рассматривается $\boldsymbol{Z} = \boldsymbol{Y} - \mu$.

Замечание 1.8. Для вычисления ℓ_c можно использовать алгоритм Левинсона-Дурбина, имеющий временную трудоемкость $O(n^2)$ [12].

1.3.2. Whittle estimation

Метод максимального правдоподобия применим, когда известно матожидание μ. При неизвестном μ обычно используют его оценку $\overline{\boldsymbol{Y}}$, однако, помимо этого, существует проблема вычислительной сложности метода при больших n.

Обе эти проблемы можно решить, используя оценку Уиттла (Whittle): вместо логарифма функции правдоподобия рассматривается ее оценка (с точностью до константы) [13]. Пусть $f(\omega; \boldsymbol{\varphi}, \sigma^2)$ — спектральная плотность $\{Y_t\}$, $I(\omega)$ — периодограмма $\boldsymbol{Y}$, тогда

$$\ell_W(\boldsymbol{\varphi}, \sigma^2) = -\frac{1}{m}\sum_{j=1}^{m}\left(\ln f(\omega_j; \boldsymbol{\varphi}, \sigma^2) + \frac{I(\omega_j)}{f(\omega_j; \boldsymbol{\varphi}, \sigma^2)}\right),$$

где $m = \lfloor (n-1)/2 \rfloor$, $\omega_j = j/n$, $j = 1, 2, \ldots, m$. Заметим, что $f(\omega; \boldsymbol{\varphi}, \sigma^2) = \sigma^2 g(\omega; \boldsymbol{\varphi})$. Тогда, максимизируя ℓ_W по σ^2, получаем

$$\widehat{\boldsymbol{\varphi}}_W = \operatorname*{argmax}_{\boldsymbol{\varphi}} Q(\boldsymbol{\varphi}), \quad \widehat{\sigma}^2_W = \frac{1}{m}\sum_{j=1}^{m}\frac{I(\omega_j)}{g(\omega_j; \widehat{\boldsymbol{\varphi}}_W)},$$

где

$$Q(\boldsymbol{\varphi}) = -\ln\frac{1}{m}\sum_{j=1}^{m}\frac{I(\omega_j)}{g(\omega_j; \boldsymbol{\varphi})} - \frac{1}{m}\sum_{j=1}^{m}\ln g(\omega_j; \boldsymbol{\varphi}).$$

Замечание 1.9. Такой метод оценки параметров можно использовать при неизвестном среднем, поскольку при ее вычислении не используется значение периодограммы в нуле.

Замечание 1.10. Периодограмму временного ряда можно вычислить за $O(n \log n)$ с помощью быстрого преобразования Фурье, что делает этот метод значительно быстрее MLE для больших n.

1.3.3. Численное сравнение методов оценки параметров

Сравним качество оценок параметров следующих моделей:

1. $d = q = 0$, $p = 1$ — модель AR(1);

2. $p = q = 0$ — модель ARFIMA$(0, \mathrm{d}, 0)$.

3. $p = 1$, $q = 0$ — модель ARFIMA$(1, d, 0)$.

Для оценки параметров этих моделей на языке R [14] были реализованы функции arfima_mle и arfima_whittle, которые соответствуют методам MLE и Whittle соответственно. Для реализации MLE были использованы функции tacvfARFIMA из пакета arfima [15] и DLLoglikelihood из пакета ltsa [12], вычисляющие автоковариационную функцию модели ARFIMA и функцию ℓ_c (1.3) соответственно.

Помимо функций arfima_mle и arfima_whittle, будем использовать:

1. Функцию arima из пакета stats, соответствующую MLE модели ARMA;

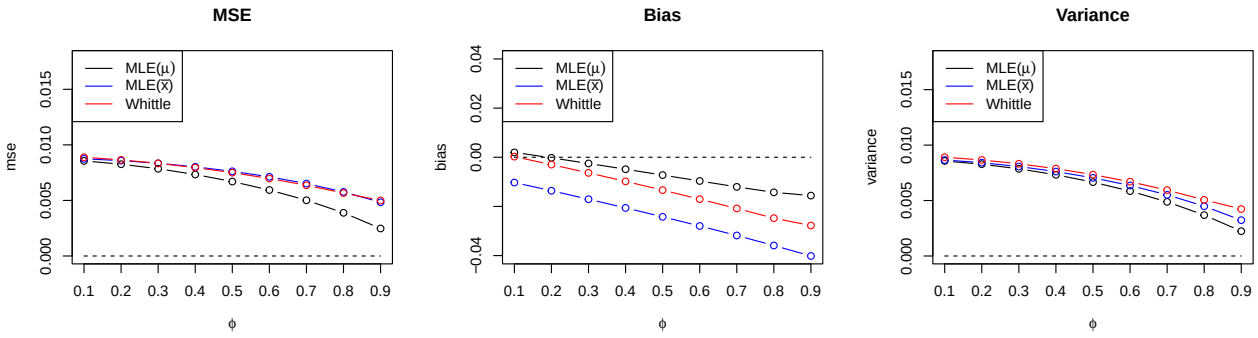
2. Функцию fracdiff из пакета fracdiff [16], соответствующую аппроксимации MLE модели ARFIMA, описанной в работе [5] (обозначим ее за H&R).

Замечание 1.11. Помимо вышеперечисленных функций, в пакете arfima есть функция arfima, которая вычисляет MLE модели ARFIMA. Но данная реализация MLE в некоторых случаях дает оценки хуже, чем arfima_mle. Сравнение оценок обоих реализаций можно найти в разделе A.1.

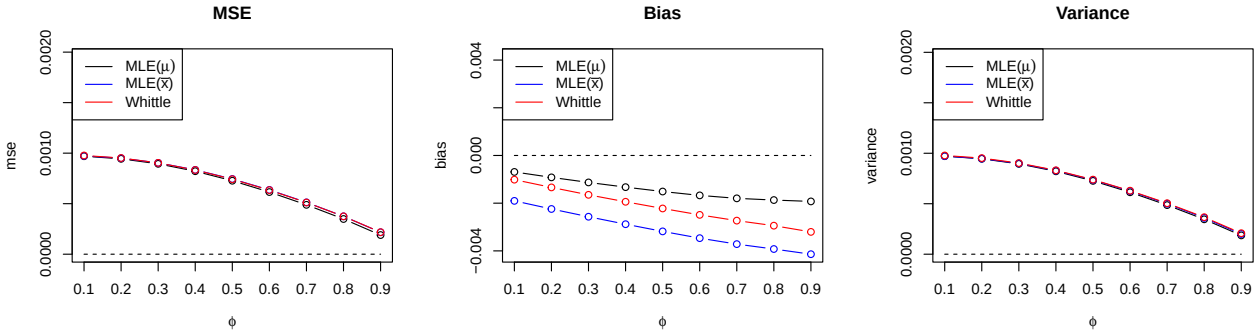
Поскольку для реальных временных рядов матожидание μ неизвестно, будем рассматривать MLE с известным средним и с его оценкой — выборочным средним (будем обозначать их MLE(μ) и MLE($\bar{x}$) соответственно). Не умаляя общности, пусть $\mu = 0$.

На рис. 1.1 и 1.2 изображены среднеквадратичное отклонение, смещение и дисперсия оценок параметров ϕ и d моделей AR(1) и ARFIMA$(0, d, 0)$. Отметим, что все оценки имеют в большинстве своем отрицательное смещение и отличаются между собой в основном степенью смещения. Как и ожидалось, оценка параметров методом максимального правдоподобия с известным средним дает оценку с наименьшим MSE. С другой стороны, если использовать вместо известного среднего выборочное среднее, оценки становятся сильно смещенными. Whittle же, в свою очередь, дает менее смещенную оценку, чем MLE($\bar{x}$), а в случае оценки d имеет смещение даже меньше, чем у MLE(μ). Однако оценки Whittle обладают наибольшей дисперсией среди всех рассмотренных методов, но разница не такая значительная, как в случае смещения.

В таблицах 1.1 и 1.2 представлены значения среднеквадратичной ошибки и смещения оценок параметров d и ϕ модели ARFIMA$(1, d)$. Заметим, что в оценках присутствует смещение, для $\phi = 0.1$ и $\phi = 0.5$ оценки d имеют отрицательное смещение, а



(*а*) $n = 100$

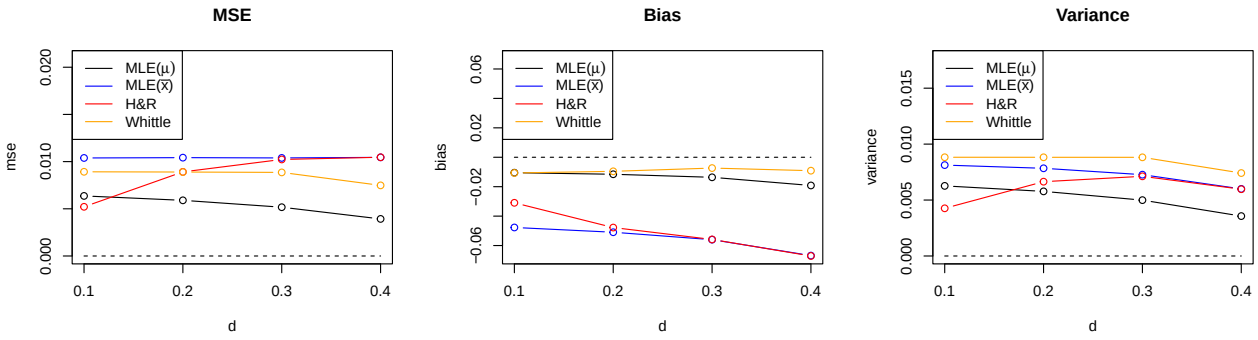


(*б*) $n = 1000$

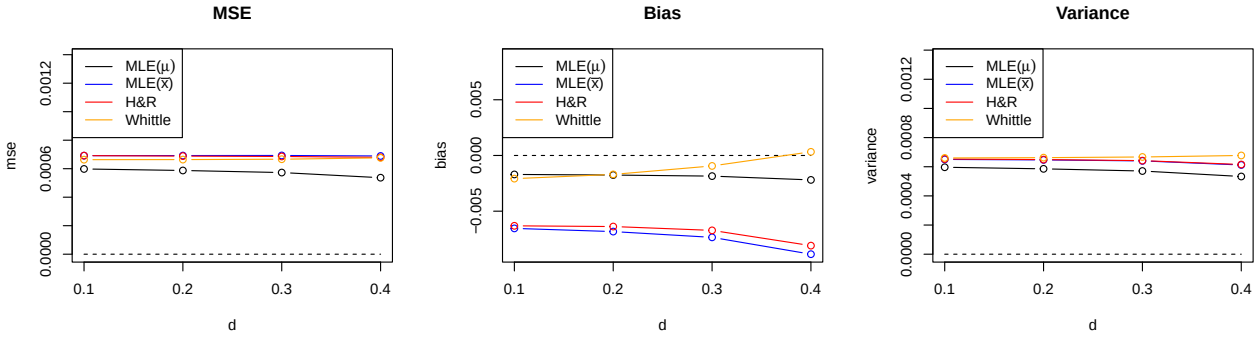
Рис. 1.1. Среднеквадратичное отклонение, смещение и дисперсия оценок параметра ϕ модели AR(1) (500 повторений)

оценки ϕ, наоборот, положительное. Также в таблицах синим цветом выделена лучшая по строке оценка d, а красным — лучшая оценка ϕ. Видно, что в случае коротких рядов ($n = 100$) метод H&R в большинстве случаев дает оценки с наименьшим MSE, однако наименьшее смещение имеют оценки MLE(μ). В случае же длинных рядов ($n = 1000$) наименьшую среднеквадратичную ошибку и смещение дает MLE(μ). Отметим, что даже для длинных рядов оценки MLE($\bar{x}$) имеют, в большинстве случаев, наибольшее смещение и MSE. Оценки H&R, хотя и дают наименьшее после MLE(μ) MSE, также сильно смещены. Оценки методом Whittle выглядят наиболее привлекательными, поскольку имеют наименьшее после MLE(μ) смещение и имеют MSE меньше, чем MLE($\bar{x}$).

Подведем итоги численного сравнения. Если для рассматриваемого ряда известно его матожидание μ (что, конечно, редкость на практике), наилучшим методом является MLE(μ). Если же среднее неизвестно, для коротких рядов оценивать параметры моделей AR(1) и ARFIMA$(0, d, 0)$ следует методом Whittle, а параметры модели ARFIMA$(1, d, 0)$ — методом H&R. В случае длинных рядов параметры рассматриваемых моделей следует оценивать методом Whittle.



(*а*) $n = 100$



(*б*) $n = 1000$

Рис. 1.2. Среднеквадратичное отклонение, смещение и дисперсия оценок параметра d модели ARFIMA$(0, d, 0)$ (500 повторений)

1.3.4. Сходимость оценок к истинным значениям

Известно [10, Theorem 8.1], что

$$\sqrt{n}(\widehat{\boldsymbol{\varphi}}_{\mathrm{ML}} - \boldsymbol{\varphi}_0) \xrightarrow{d} \mathcal{N}_{k+1}(0, \mathcal{I}^{-1}(\boldsymbol{\varphi}_0)), \tag{1.4}$$

где $\boldsymbol{\varphi}_0$ — истинный вектор параметров, $\mathcal{I}(\boldsymbol{\varphi})$ — информационная матрица Фишера. Также известно [10, Proposition 8.3], что вектор $\widehat{\boldsymbol{\varphi}}_W$ имеет такое же асимптотическое распределение, что и $\widehat{\boldsymbol{\varphi}}_{\mathrm{ML}}$.

Покажем, что методы MLE(μ) и Whittle реализованы корректно, посмотрев на дисперсию оценок для $n = 10000$. В таблице 1.3 представлены оценки дисперсий $\hat{d}$ и $\hat{\phi}$, в скобках указана теоретическая дисперсия. Как видим, для разных значений параметров значение дисперсий оценок близки к теоретическим.

Таблица 1.1. Смещение и среднеквадратичное отклонение оценок параметров d и ϕ модели ARFIMA$(1, d, 0)$ ($n = 100$, 500 повторений)

		MSE								Bias							
		MLE(μ)		MLE($\bar{x}$)		H&R		Whittle		MLE(μ)		MLE($\bar{x}$)		H&R		Whittle	
d	ϕ	$\hat{d}$	$\hat{\phi}$	$\hat{d}$	$\hat{\phi}$	$\hat{d}$	$\hat{\phi}$	$\hat{d}$	$\hat{\phi}$	$\hat{d}$	$\hat{\phi}$	$\hat{d}$	$\hat{\phi}$	$\hat{d}$	$\hat{\phi}$	$\hat{d}$	$\hat{\phi}$
0.1	0.1	0.049	0.056	0.119	0.114	0.009	0.018	0.069	0.067	-0.077	0.066	-0.229	0.199	-0.054	0.035	-0.086	0.068
0.2	0.1	0.047	0.055	0.151	0.141	0.025	0.032	0.077	0.073	-0.078	0.067	-0.265	0.232	-0.119	0.099	-0.094	0.074
0.3	0.1	0.041	0.049	0.183	0.165	0.049	0.055	0.084	0.081	-0.076	0.066	-0.301	0.266	-0.179	0.161	-0.109	0.09
0.4	0.1	0.029	0.038	0.211	0.187	0.081	0.089	0.179	0.194	-0.072	0.065	-0.34	0.305	-0.243	0.23	-0.26	0.241
0.1	0.5	0.045	0.041	0.086	0.053	0.010	0.015	0.057	0.054	-0.071	0.034	-0.222	0.151	-0.07	0.034	-0.066	0.024
0.2	0.5	0.042	0.038	0.092	0.055	0.031	0.025	0.074	0.058	-0.081	0.046	-0.244	0.171	-0.154	0.108	-0.153	0.107
0.3	0.5	0.040	0.036	0.1	0.059	0.063	0.043	0.098	0.062	-0.093	0.060	-0.267	0.192	-0.232	0.174	-0.209	0.161
0.4	0.5	0.037	0.033	0.115	0.067	0.104	0.065	0.104	0.066	-0.103	0.073	-0.304	0.226	-0.306	0.235	-0.228	0.177
0.1	0.9	0.029	0.029	0.014	0.01	0.007	0.007	0.034	0.025	0.075	-0.089	0.01	-0.049	0.001	-0.043	0.049	-0.069
0.2	0.9	0.019	0.018	0.011	0.006	0.009	0.004	0.026	0.019	0.046	-0.065	-0.011	-0.035	-0.037	-0.026	0.02	-0.056
0.3	0.9	0.012	0.01	0.009	0.004	0.014	0.003	0.022	0.015	0.016	-0.043	-0.033	-0.023	-0.076	-0.011	-0.024	-0.039
0.4	0.9	0.008	0.006	0.009	0.002	0.025	0.002	0.028	0.01	-0.016	-0.024	-0.061	-0.008	-0.121	0.003	-0.095	-0.016

Таблица 1.2. Смещение и среднеквадратичное отклонение оценок параметров d и ϕ модели ARFIMA$(1, d, 0)$ ($n = 1000$, 500 повторений)

		MSE · 100								Bias · 100							
		MLE(μ)		MLE($\bar{x}$)		H&R		Whittle		MLE(μ)		MLE($\bar{x}$)		H&R		Whittle	
d	ϕ	$\hat{d}$	$\hat{\phi}$	$\hat{d}$	$\hat{\phi}$	$\hat{d}$	$\hat{\phi}$	$\hat{d}$	$\hat{\phi}$	$\hat{d}$	$\hat{\phi}$	$\hat{d}$	$\hat{\phi}$	$\hat{d}$	$\hat{\phi}$	$\hat{d}$	$\hat{\phi}$
0.1	0.1	0.186	0.290	0.27	0.37	0.224	0.317	0.232	0.33	-0.581	0.448	-1.941	1.732	-1.729	1.508	-0.687	0.519
0.2	0.1	0.181	0.287	0.271	0.371	0.267	0.366	0.232	0.329	-0.599	0.465	-2.026	1.824	-1.981	1.773	-0.615	0.468
0.3	0.1	0.174	0.282	0.272	0.372	0.268	0.367	0.232	0.328	-0.639	0.505	-2.177	1.987	-2.176	1.975	-0.476	0.372
0.4	0.1	0.156	0.267	0.272	0.373	0.274	0.371	0.232	0.325	-0.795	0.666	-2.606	2.443	-2.725	2.542	-0.263	0.233
0.1	0.1	0.761	0.8	1.429	1.3	0.504	0.552	1.104	1.05	-2.047	1.588	-5.86	5.102	-3.256	2.741	-2.555	1.904
0.2	0.1	0.710	0.759	1.432	1.302	0.978	0.969	1.213	1.125	-2.018	1.571	-6.08	5.337	-5.157	4.556	-2.721	2.065
0.3	0.1	0.617	0.675	1.462	1.323	1.246	1.175	1.23	1.15	-1.984	1.560	-6.506	5.78	-6.129	5.463	-2.578	1.948
0.4	0.1	0.473	0.539	1.499	1.353	1.507	1.354	1.294	1.18	-2.226	1.861	-7.514	6.838	-7.622	6.905	-2.695	2.099
0.1	0.9	0.338	0.155	0.288	0.122	0.259	0.097	0.387	0.193	0.623	-0.774	-0.095	-0.56	-0.176	-0.504	0.583	-0.92
0.2	0.9	0.273	0.106	0.233	0.077	0.239	0.077	0.326	0.128	0.42	-0.611	-0.388	-0.355	-0.627	-0.303	0.041	-0.652
0.3	0.9	0.241	0.093	0.217	0.068	0.268	0.075	0.326	0.097	0.287	-0.53	-0.667	-0.204	-1.395	-0.029	-1.003	-0.285
0.4	0.9	0.173	0.067	0.182	0.051	0.381	0.076	0.545	0.09	-0.129	-0.295	-1.4	0.178	-2.602	0.359	-3.357	0.389

Метод	$n\mathsf{D}\hat{d}$ (1.83167)	$n\mathsf{D}\hat{\phi}$ (2.98284)
MLE(μ)	1.72733	2.92604
Whittle	1.78350	2.94929

(a) $d = 0.2$, $\varphi = 0.1$

Метод	$n\mathsf{D}\hat{d}$ (1.83167)	$n\mathsf{D}\hat{\phi}$ (2.98284)
MLE(μ)	1.69348	2.90374
Whittle	1.84265	2.96828

(b) $d = 0.4$, $\varphi = 0.1$

Метод	$n\mathsf{D}\hat{d}$ (4.91219)	$n\mathsf{D}\hat{\phi}$ (6.06018)
MLE(μ)	5.00869	6.305
Whittle	5.08405	6.35985

(c) $d = 0.2$, $\varphi = 0.5$

Метод	$n\mathsf{D}\hat{d}$ (4.91219)	$n\mathsf{D}\hat{\phi}$ (6.06018)
MLE(μ)	4.75975	6.06711
Whittle	5.28104	6.4936

(d) $d = 0.4$, $\varphi = 0.5$

Метод	$n\mathsf{D}\hat{d}$ (2.49203)	$n\mathsf{D}\hat{\phi}$ (0.77885)
MLE(μ)	2.42011	0.78209
Whittle	2.44394	0.77549

(e) $d = 0.2$, $\varphi = 0.9$

Метод	$n\mathsf{D}\hat{d}$ (2.49203)	$n\mathsf{D}\hat{\phi}$ (0.77885)
MLE(μ)	2.26718	0.749318
Whittle	2.57117	0.77311

(f) $d = 0.4$, $\varphi = 0.9$

Таблица 1.3. Дисперсия оценок $\hat{d}$ и $\hat{\phi}$, $n = 10000$, 100 повторений

Глава 2

Метод Monte Carlo SSA

2.1. Проверка статистических гипотез

Рассмотрим некоторый критерий со статистикой T. Введем обозначения.

Определение 2.1. Ошибка первого рода — вероятность отвергнуть нулевую гипотезу, если она верна: $\alpha_I(\alpha) = \mathsf{P}_{H_0}(T \in A_{\text{крит}}(\alpha))$.

Определение 2.2. Если $\alpha_I = \alpha$, то говорят, что критерий точный при уровне значимости α, иначе говорят, что критерий неточный. При $\alpha_I < \alpha$ критерий является консервативным, а при $\alpha_I > \alpha$ — радикальным.

Определение 2.3. Мощность критерия против альтернативы H_1 — вероятность отвергнуть нулевую гипотезу, если верна альтернативная: $\beta(\alpha) = \mathsf{P}_{H_1}(T \in A_{\text{крит}}(\alpha))$.

2.1.1. Поправка неточных критериев

Зафиксируем некоторый неточный (консервативный или радикальный) критерий и уровень значимости α^*. Пусть дана зависимость ошибки первого рода от уровня значимости $\alpha_I(\alpha) = \mathsf{P}_{H_0}(\text{p} < \alpha)$. Тогда критерий с формальным уровнем значимости $\widetilde{\alpha}^* = \alpha_I^{-1}(\alpha^*)$ является точным: ошибка первого рода $\alpha_I(\widetilde{\alpha}^*) = \alpha^*$.

Если зависимость $\alpha_I(\alpha)$ неизвестна, она оценивается с помощью моделирования. Приведем алгоритм поправки в этом случае. Помимо критерия и уровня значимости, зафиксируем количество выборок M для оценки $\alpha_I(\alpha)$ и их объем N.

Алгоритм 1. Поправка уровня значимости по зависимости $\alpha_I(\alpha)$ [17]

1. Моделируется M выборок объема N при верной H_0.
2. По моделированным данным строится оценка зависимость ошибки первого рода от уровня значимости $\alpha_I(\alpha)$.
3. Рассчитывается формальный уровень значимости: $\widetilde{\alpha}^* = \alpha_I^{-1}(\alpha^*)$. Критерий с таким уровнем значимости является асимптотически точным при $M \to \infty$.

2.1.2. Сравнение критериев

Точные критерии, проверяющие одну и ту же гипотезу, можно использовать и сравнивать по мощности: чем больше мощность, тем лучше. Если критерий является консервативным, использовать и сравнивать его с другими критерии по мощности также можно, учитывая, что его мощность будет занижена. Радикальный же критерий, без поправки, введенной в разделе 2.1.1, нельзя использовать и сравнивать по мощности с другими критериями. Поэтому введем понятие ROC-кривой, соответствующее мощности критерия, к которому была применена поправка.

Определение 2.4. ROC-кривая — это кривая, задаваемая параметрически

$$\begin{cases} x = \alpha_I(\alpha) \\ y = \beta(\alpha) \end{cases}, \quad \alpha \in [0,1]$$

Замечание 2.1. С помощью ROC-кривых можно сравнивать по мощности неточные (в частности, радикальные) критерии. Отметим, что для точного критерия ROC-кривая совпадает с графиком мощности, так как $\alpha_I(\alpha) = \alpha$.

2.2. Monte Carlo SSA

Метод Monte Carlo SSA (MC-SSA) тесно связан с методом SSA (Singular Spectrum Analysis), состоящим из четырех этапов: *вложения*, *разложения*, *группировки* и *диагонального усреднения*. Поэтому опишем сначала его.

2.2.1. Метод SSA

Пусть $\mathsf{X} = (x_1, \ldots, x_N)$ — временной ряд длины N. Зафиксируем длину окна L, $1 < L < N$. Рассмотрим $K = N - L + 1$ векторов вложения $X_i = (x_i, \ldots, x_{i+L-1})$ и составим из столбцов X_i так называемую траекторную матрицу:

$$\mathbf{X} = [X_1 : \ldots : X_K].$$

Далее траекторная матрица $\mathbf{X}$ разбивается в сумму матриц единичного ранга. В базовом SSA используются собственные векторы матрицы $\mathbf{X}\mathbf{X}^{\mathrm{T}}$, в Toeplitz SSA используются собственные векторы матрицы $\mathbf{T}$ с элементами

$$t_{ij} = \frac{1}{N - |i-j|} \sum_{n=1}^{N-|i-j|} x_n x_{n+|i-j|}, \quad i, j \leqslant L. \tag{2.1}$$

Обозначим за $P_1, \ldots, P_L$ собственные векторы матрицы $\mathbf{X}\mathbf{X}^{\mathrm{T}}$ либо матрицы $\mathbf{T}$. Тогда получаем следующее разложение:

$$\mathbf{X} = \sum_{i=1}^{L} \sigma_i P_i Q_i^{\mathrm{T}} = \mathbf{X}_1 + \ldots + \mathbf{X}_L,$$

где $S_i = \mathbf{X}^{\mathrm{T}} P_i$, $Q_i = S_i / \|S_i\|$, $\sigma_i = \|S_i\|$.

После этого полученные матрицы группируются и каждая из группированных матриц преобразовывается обратно во временной ряд. Таким образом, результатом SSA является разложение временного ряда.

2.2.2. Постановка задачи

Рассмотрим задачу поиска сигнала (неслучайной составляющей) во временном ряде. Модель выглядит следующим образом:

$$\mathsf{X} = \mathsf{S} + \boldsymbol{\xi},$$

где S — сигнал, $\boldsymbol{\xi}$ — стационарный процесс с нулевым средним. Тогда нулевая гипотеза $H_0 : \mathsf{S} = 0$ (отсутствие сигнала, ряд состоит из чистого шума) и альтернатива $H_1 : \mathsf{S} \neq 0$ (ряд содержит сигнал, например, периодическую составляющую).

2.2.3. Множественный тест

Зафиксируем длину окна L и модель шума $\boldsymbol{\xi}$. Пусть $\mathsf{R}_1, \ldots, \mathsf{R}_G$ — реализации $\boldsymbol{\xi}$, которые в дальнейшем будем называть суррогатными. Обозначим за $\mathbf{X}$ и $\boldsymbol{\Xi}_i$, $i = 1, \ldots, G$, траекторные матрицы ряда X и каждой суррогатной реализации соответственно. Рассмотрим H проекционных векторов $W_1, \ldots, W_H$, каждый из которых соответствует некоторой частоте ω_k, $\|W_k\| = 1$, $k = 1, \ldots, H$.

Алгоритм 2. Multiple MC-SSA [18]

1. Для $k = 1, \ldots, H$ вычисляется статистика $\widehat{p}_k = \|\mathbf{X}^{\mathrm{T}} W_k\|^2$, выборка $P_k = \{p_{ki}\}_{i=1}^{G}$ с элементами $p_{ki} = \|\boldsymbol{\Xi}_i^{\mathrm{T}} W_k\|^2$, ее среднее μ_k и стандартное отклонение σ_k.

2. Вычисляется $\eta = (\eta_1, \ldots, \eta_G)$, где

$$\eta_i = \max_{1 \leqslant k \leqslant H} (p_{ki} - \mu_k)/\sigma_k, \quad i = 1, \ldots, G.$$

3. Находится q как выборочный $(1 - \alpha)$-квантиль η, где α — уровень значимости.

4. Нулевая гипотеза не отвергается, если

$$t = \max_{1 \leqslant k \leqslant H} (\widehat{p}_k - \mu_k)/\sigma_k < q.$$

5. Если H_0 отвергнута, вклад W_k (и соответствующей частоты) значим, если $\widehat{p}_k$ превосходит $\mu_k + q\sigma_k$. Таким образом, $[0, \mu_k + q\sigma_k]$ считаются скорректированными интервалами прогнозирования.

2.2.4. Ограничение на модель шума

Для модели шума $\boldsymbol{\xi}$ важно, чтобы спектральная плотность процесса была строго монотонной. Это связано с тем, что в таком случае собственные векторы автоковариационной матрицы стационарного процесса ведут себя как синусоиды с равностоящими частотами, а соответствующие им собственные числа примерно равны значению спектральной плотности в этих частотах. Для процессов с короткой памятью это верно, поскольку теплицеву симметричную матрицу можно аппроксимировать циркулянтной матрицей [19], собственные векторы которой равны

$$v_j = \frac{1}{\sqrt{n}} \left(1, \omega^j, \omega^{2j}, \ldots, \omega^{(n-1)j}\right), \quad j = 0, 1 \ldots, n-1,$$

а соответствующие им собственные числа равны

$$\lambda_j = \sum_{k=0}^{n-1} c_k \omega^j,$$

где $\omega = \exp\{2\pi \mathrm{i}/n\}$. Тогда при строгой монотонности $f_{\boldsymbol{\xi}}(\omega)$ вклад собственных векторов будет попарно различным, делая их сильно разделимыми [2, Раздел 1.5.4]. Если же опустить требование строгой монотонности, компоненты могут смешаться, что сделает невозможным определение доминирующей частоты значимого вектора.

Для процессов с длинной памятью покажем правдивость этого факта, проведя численный эксперимент. Рассмотрим модель ARFIMA$(0, d, 0)$ с $d = 0.4$ и пусть размер автоковариационной матрицы $\boldsymbol{\Sigma}_n$ равен $n = 100$. Частоту векторов будем оценивать с помощью метода ESPRIT [20, Раздел 3.1].

На рис. 2.1 представлены первые 10 собственных векторов матрицы $\boldsymbol{\Sigma}_n$, на рис. 2.2 по оси Ox отложена частота, которой соответствует собственный вектор, а по оси Oy отложено соответствующее вектору собственное число, дополнительно синей линией проведена спектральная плотность процесса. Как видим, действительно, собственные

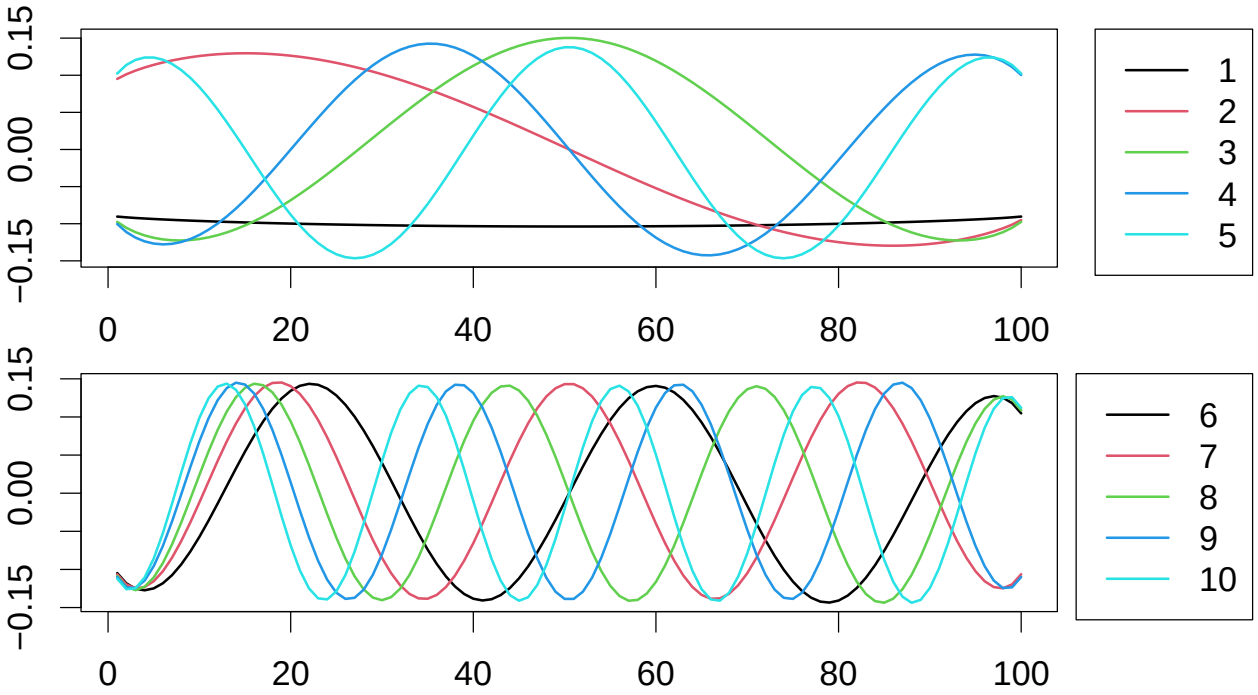


Рис. 2.1. Собственные векторы автоковариационной матрицы модели ARFIMA$(0, d, 0)$

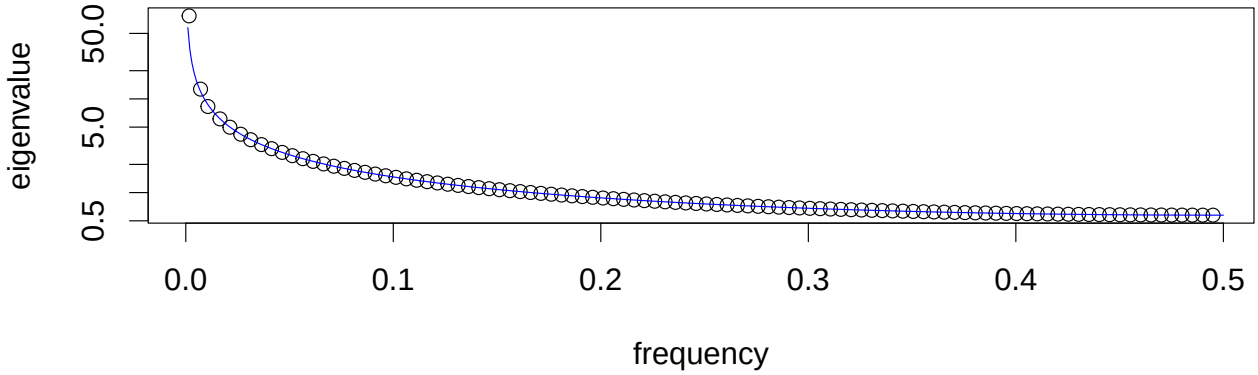


Рис. 2.2. Собственные числа автоковариационной матрицы модели ARFIMA$(0, d, 0)$

векторы ведут себя как периодики, собственные числа хорошо приближаются значением спектральной плотности в соответствующей частоте, и по рис. 2.3 разница между частотами примерно равна $1/(2n) = 0.005$.

2.2.5. Используемый вариант MC-SSA

В разделе 2.2.3 предполагалось, что векторы $W_1, \ldots, W_H$ фиксированные и не зависят от исходного ряда. Такой критерий MC-SSA является точным, то есть ошибка первого рода равна заданному уровню значимости. В этой работе будут рассматриваться векторы W_k, порожденные рядом X, при этом по-прежнему при вычислении p_{ki} используются те же W_k, что и при вычислении $\widehat{p}_k$.

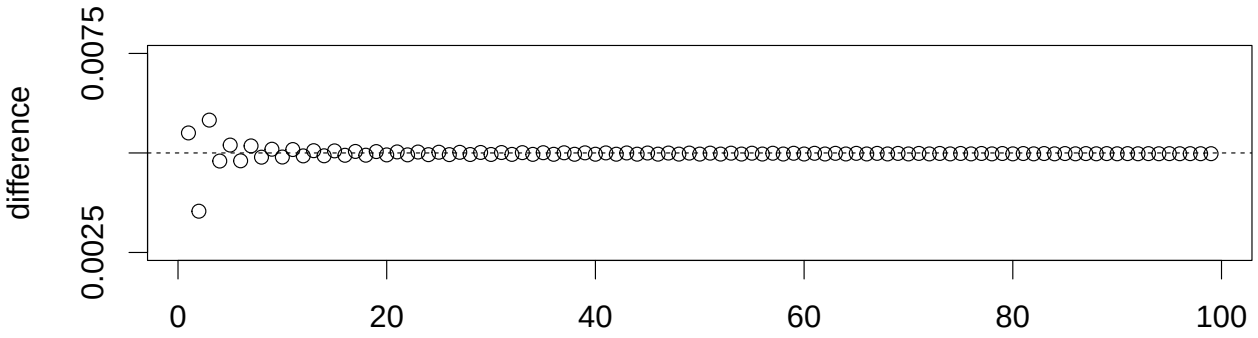


Рис. 2.3. Разница частот между ближайшими собственными векторами

Поскольку в этом варианте векторы W_k не заданы заранее, а порождены исходным рядом, критерий MC-SSA становится, вообще говоря, радикальным. Бороться с этой проблемой позволяет метод эмпирической поправки критерия, описанный в разделе 2.1.1.

В качестве $W_1, \ldots, W_H$ будем брать собственные векторы матрицы $\mathbf{X}\mathbf{X}^{\mathrm{T}}$ или $\mathbf{T}$ (см. формулу (2.1)). Такой способ выбора векторов для проекции самый распространенный, поскольку, если есть значимые векторы, можно восстановить сигнал с помощью SSA на их основе. Будем под MC-SSA подразумевать именно этот вариант критерия. Варианты критерия будут определяться конкретным разложением траекторной матрицы. Заметим, что обычно используется сингулярное разложение.

2.2.6. Сравнение MC-SSA по мощности при разных моделях шума

Пусть $\boldsymbol{\xi}$ — красный шум, а $\boldsymbol{\eta}$ — модель ARFIMA$(0, d, 0)$. Будем считать дисперсию белого шума одинаковой для обоих процессов и равной σ^2. Дисперсии $\boldsymbol{\xi}$ и $\boldsymbol{\eta}$ соответственно равны

$$\mathsf{D}\boldsymbol{\xi} = \frac{\sigma^2}{1-\phi^2}, \quad \mathsf{D}\boldsymbol{\eta} = \sigma^2 \frac{\Gamma(1-2d)}{\Gamma(1-d)^2}.$$

Тогда дисперсии процессов равны тогда и только тогда, когда

$$\phi = \pm\sqrt{1 - \frac{\Gamma(1-d)^2}{\Gamma(1-2d)}}.$$

Пусть $d = 0.4$. Тогда при $\phi \approx 0.719$ процессы $\boldsymbol{\xi}$ и $\boldsymbol{\eta}$ имеют одинаковую дисперсию. На рис. 2.4 изображены спектральные плотности процессов. На нем видно, что процесс $\boldsymbol{\eta}$ имеет меньшее значение плотности для всех значений $\omega \in (0, 0.2)$, за исключением близких к нулю.

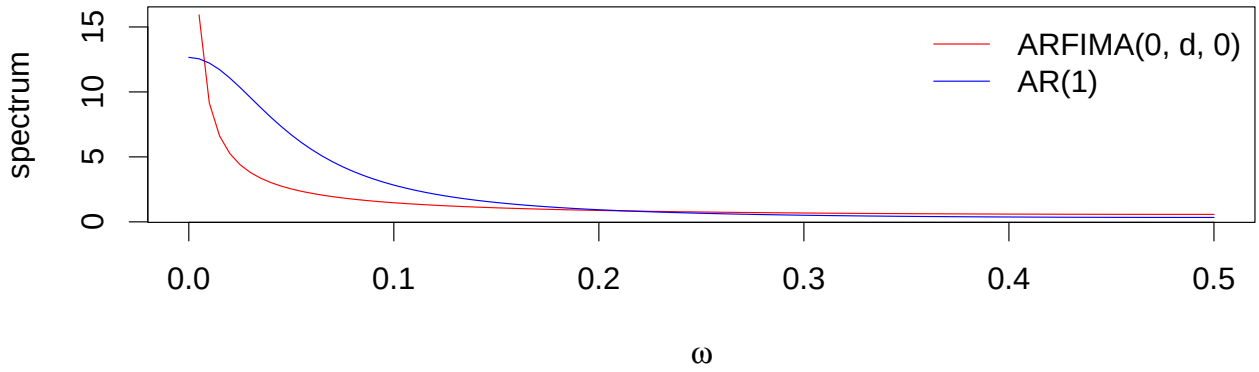
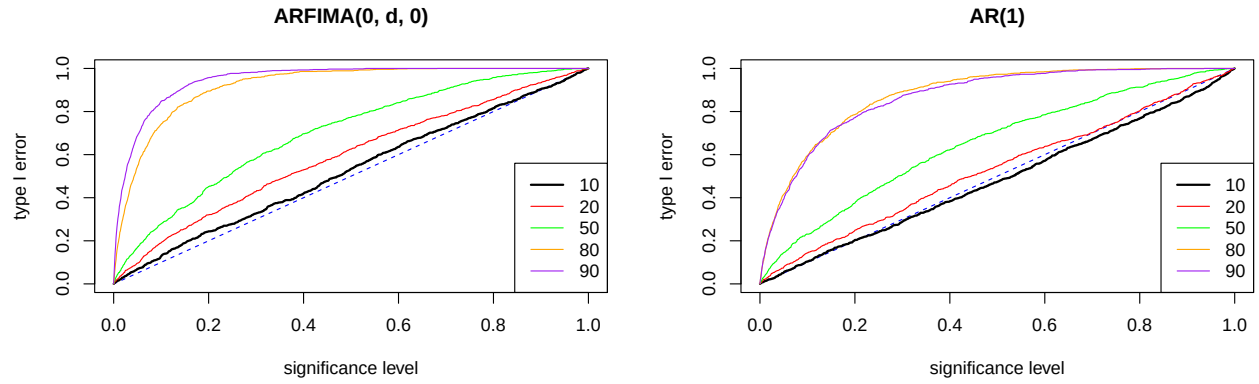


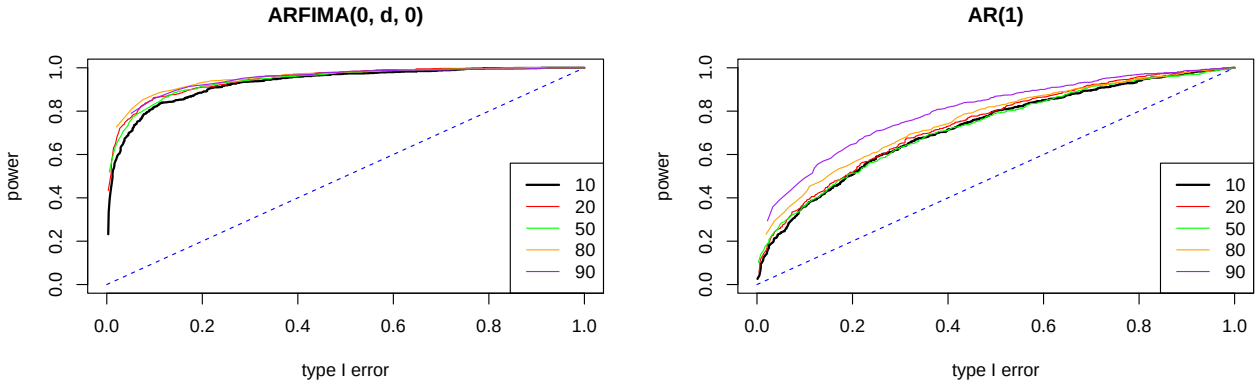
Рис. 2.4. Спектральная плотность процессов с одинаковой дисперсией

Предполагается, что если рассмотреть в качестве альтернативы сигнал с частотой $\omega : f_{\boldsymbol{\eta}}(\omega) < f_{\boldsymbol{\xi}}(\omega)$, то мощность критерия MC-SSA против этой альтернативы при модели шума $\boldsymbol{\eta}$ больше, чем при модели шума $\boldsymbol{\xi}$. Убедимся в этом. Пусть длина ряда $N = 100$, $\sigma^2 = 1$ и

$$\mathsf{S} = \{A\cos(2\pi n\omega)\}_{n=1}^{N}, \quad A = 1,\ \omega = 0.075.$$



(*а*) Ошибка первого рода



(*б*) ROC-кривая

Рис. 2.5. Сравнение мошностей MC-SSA

На рис. 2.5, *а* изображены график ошибок первого рода критериев MC-SSA для

Таблица 2.1. Оценка параметров модели ARFIMA$(0, d, 0)$ ряда Nile Minima

Метод	$\hat{d}$	$\hat{\sigma}^2$
MLE($\bar{x}$)	0.39264	0.48939
H&R	0.39327	0.48934
Whittle	0.40547	0.49026

разных длин окна L. По нему видно, что рассматриваемые критерии являются радикальными, поэтому сравнивать их по мощности будем с помощью ROC-кривых, которые являются графиками мощности критериев, к которым была применена поправка из раздела 2.1.1. По рис. 2.5, *б* видно, что, действительно, мощность критерия против данной альтернативы при модели шума ARFIMA$(0, d, 0)$ больше, чем при модели шума AR(1) с такой же дисперсией.

2.3. Применение MC-SSA на реальных временных рядах с длинной памятью

Рассмотрим несколько примеров реальных временных рядов с длинной памятью и применим к ним MC-SSA. Оценивать параметры будем теми же методами, что и в разделе 1.3.3.

2.3.1. Nile Minima

На рис. 2.6, *а* изображен ежегодный минимальный уровень воды реки Нил за период с 622 по 1284 год (663 наблюдения), данные были взяты из [21]. Нерегулярные циклы или тенденции в этом временном ряду, обусловленные длинной памятью, впервые были обнаружены и обсуждены Хёрстом, британским инженером, который работал гидрологом на реке Нил. Подтверждает присутствие длинной памяти график медленно угасающей автокорреляционной функции на рис. 2.6, *б*.

Оценим параметры модели ARFIMA$(0, d, 0)$. В таблице 2.1 представлены оценки параметров d и σ^2. Поскольку истинное среднее неизвестно и оценка d по Whittle дает наименьшее смещение (см. рис. 1.2), в качестве нулевой гипотезы MC-SSA выберем модель ARFIMA$(0, d, 0)$ с $d = 0.40547$ и $\sigma^2 = 0.48971$, на рис. 2.7 изображена периодограмма ряда вместе с оцененной спектральной плотностью.

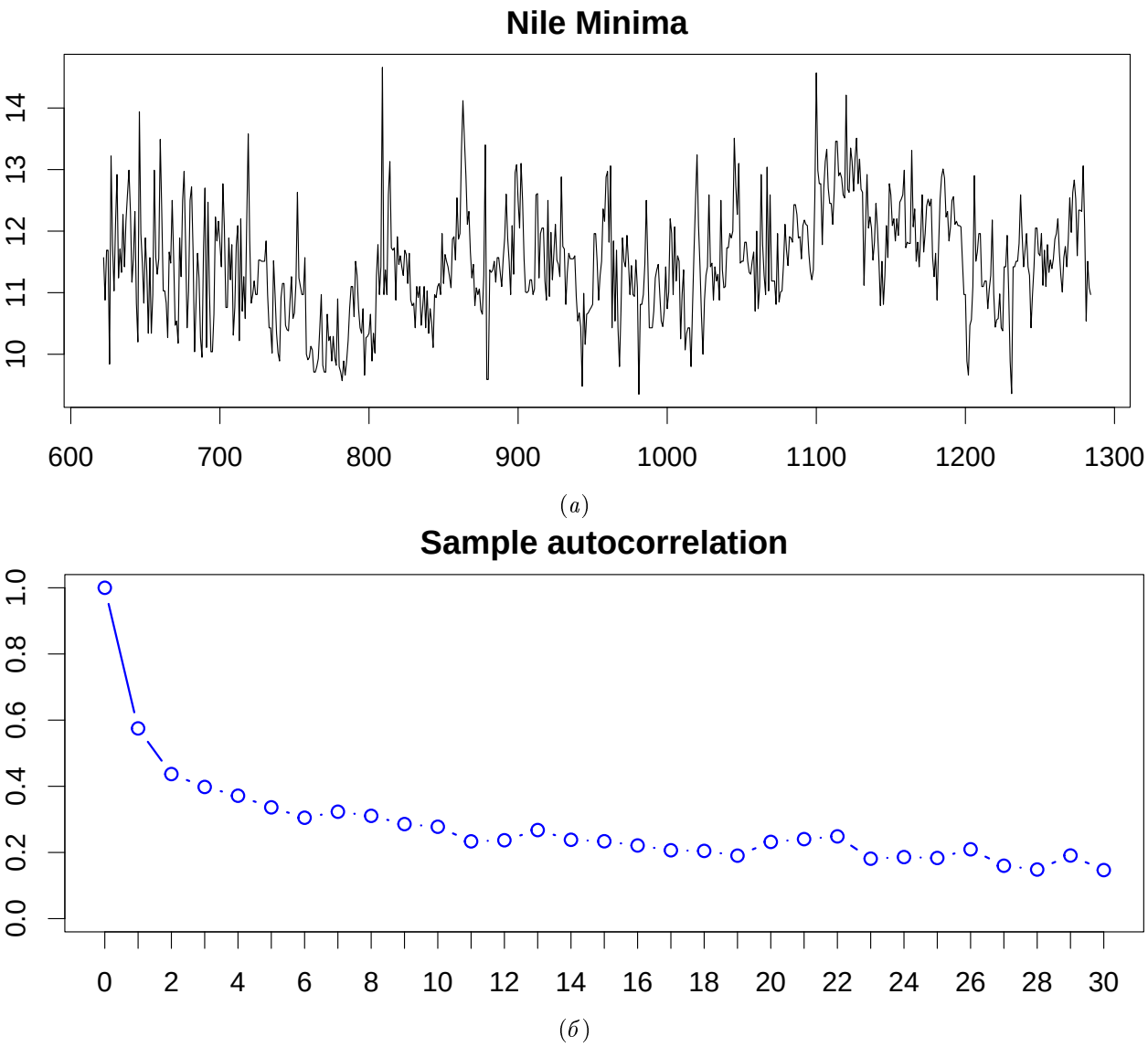


Рис. 2.6. Ежегодный минимальный уровень воды реки Нил

Применим MC-SSA с длиной окна $L = 330 \approx N/2$. На рис. 2.8 изображены 95%-ные доверительные интервалы статистик $\widehat{p}_k$, $k = 1, \ldots, L$ (см. алгоритм 2). Ни одна из статистик не является значимой, это означает, что нет оснований полагать, что в этом временном ряде присутствует неслучайный сигнал.

2.3.2. Ireland Wind

На рис. 2.9 изображены среднесуточные данные о скорости ветра (в узлах) за период с 1961 по 1978 год (6574 дней) на станции Roche's Point в Республике Ирландия [5].

В таблице 2.2 представлены оценки параметров. Полученные оценки примерно одинаковые, но поскольку Whittle дает менее смещенную оценку (см. рис. 1.2 и таблицу 1.2),

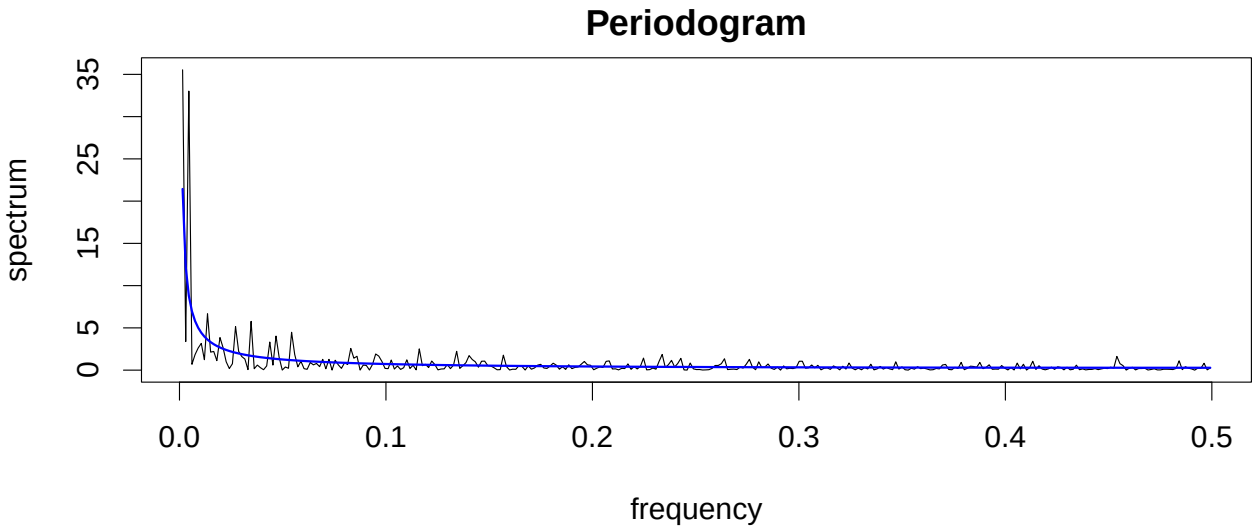


Рис. 2.7. Периодограмма ряда Nile Minima

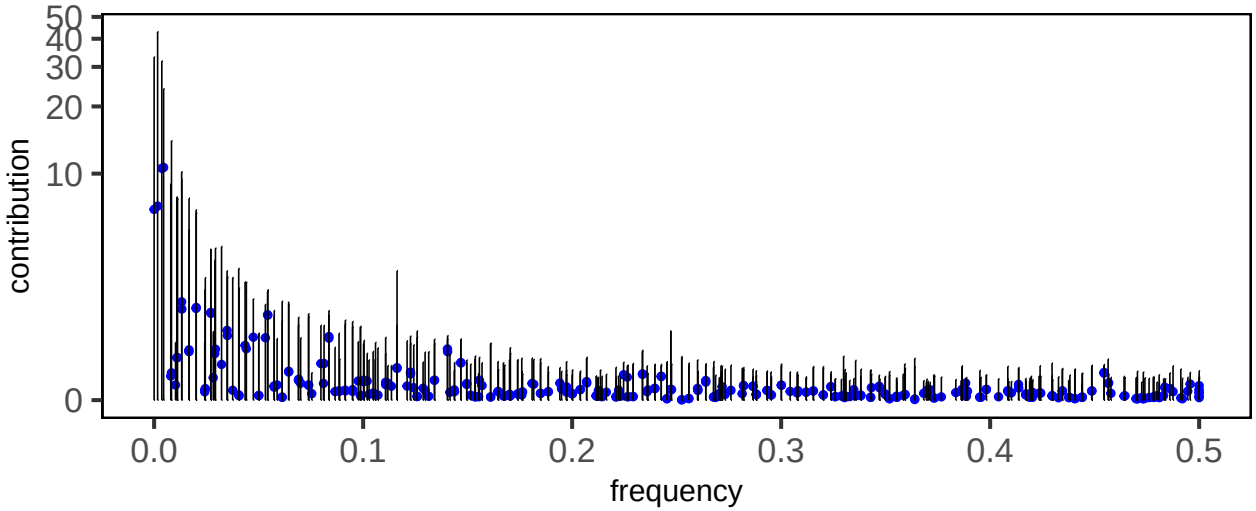


Рис. 2.8. Результат работы MC-SSA для ряда Nile Minima

будем использовать именно ее.

Поскольку ряд достаточно длинный, чтобы не делать поправку, рассмотрим в качестве векторов для проекции косинусы с равностоящими частотами с шагом $1/(2L)$. Также выберем длину окна не слишком большой, чтобы метод MC-SSA считался за адекватное время, скажем, $L = 365$. На рис. 2.10 представлен результат работы MC-SSA для обоих моделей, уровень значимости, как и в прошлом примере, равен $\alpha = 0.05$. Для модели ARFIMA$(0, d, 0)$ значимых векторов два, их периоды равны $365/34 \approx 10.74$ и $365/60 \approx 6.08$ дней соответственно, что сложно как-то интерпретировать. Однако стоит заметить, что вклады значимых векторов не слишком превосходят верхние границы соответствующих предсказательных интервалов, поэтому, скорее всего, векторы значи-

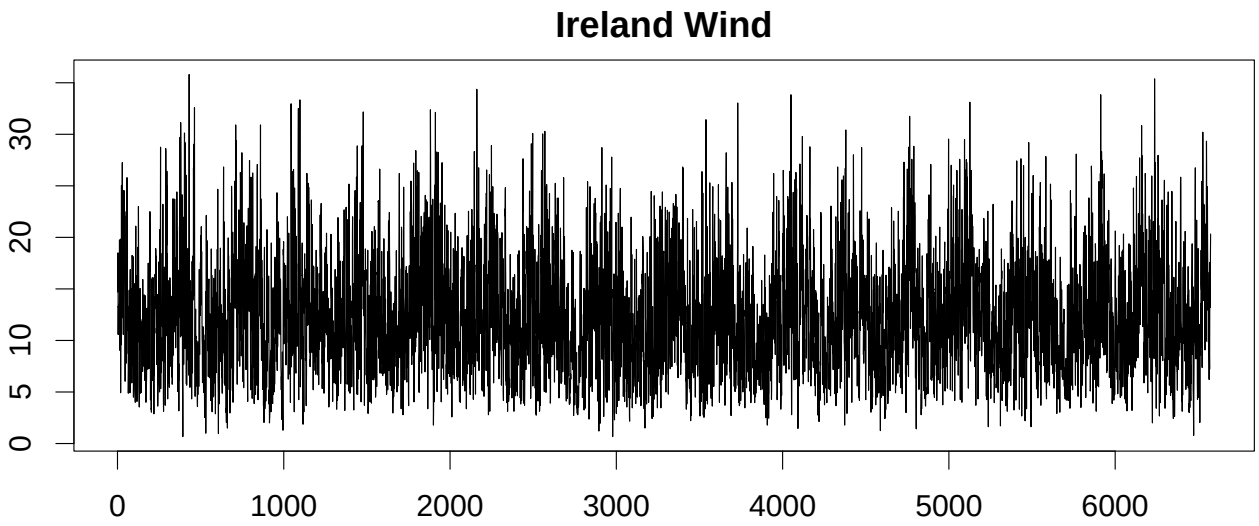
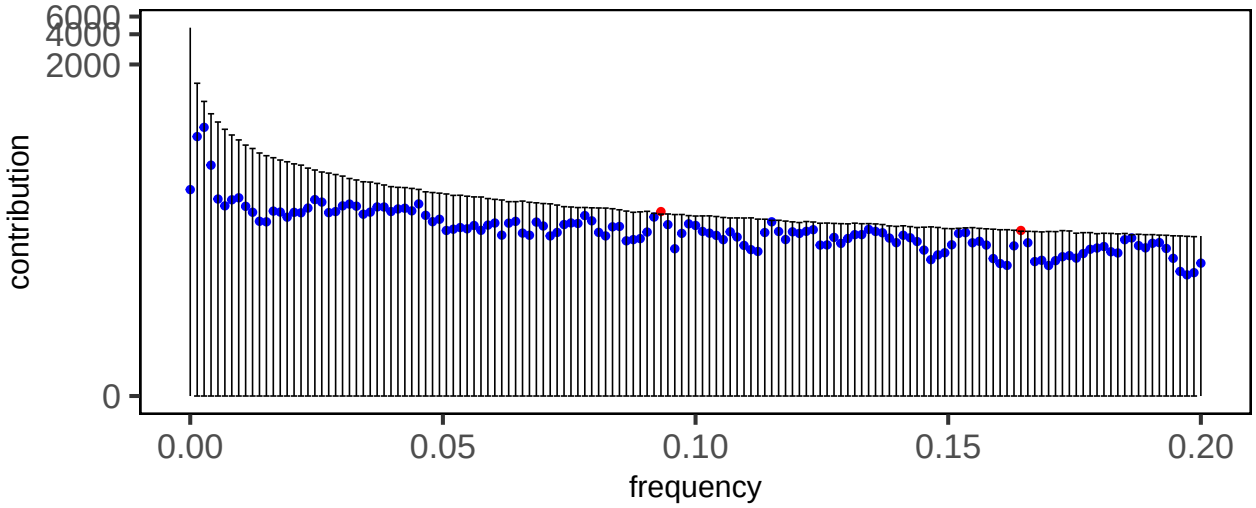


Рис. 2.9. Среднесуточные данные о скорости ветра в Республике Ирландия

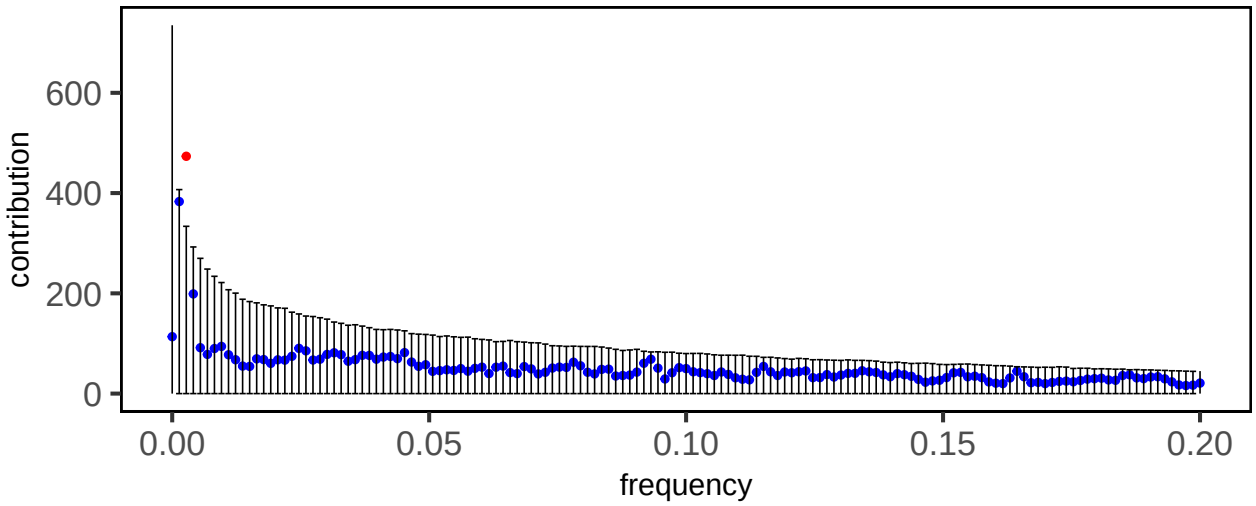
Таблица 2.2. Оценка параметров ряда Ireland Wind

	ARFIMA$(0, d, 0)$		ARFIMA$(1, d, 0)$		
Метод	$\hat{d}$	$\hat{\sigma}^2$	$\hat{d}$	$\hat{\phi}$	$\hat{\sigma}^2$
MLE$(\bar{x})$	0.37117	24.39916	0.17306	0.28403	23.7581
H&R	0.36891	24.38116	0.17245	0.28309	23.73458
Whittle	0.37287	24.40285	0.17598	0.28105	23.75983

мы случайно (количество случайно значимых векторов в среднем равно $\alpha \cdot L = 18.25$). В случае модели ARFIMA$(1, d, 0)$ значим всего один вектор с периодом 365 дней, что интерпретируется как наличие во временном ряде годовой периодичности.



(a) ARFIMA$(0, d, 0)$



($б$) ARFIMA$(1, d, 0)$

Рис. 2.10. Результат работы MC-SSA для ряда Ireland Wind

Глава 3

Метод autoMCSSA

3.1. Сравнение способов задания проекционных векторов

Сравним два способа задания векторов для проекции W_k критерия MC-SSA:

1. Собственные векторы матрицы $\mathbf{X}\mathbf{X}^{\mathrm{T}}$ или $\mathbf{T}$.
2. Косинусы с равноотстоящими частотами $k/(2L)$, $k = 1, \ldots, L$.

Для краткости будем называть соответсвующие им критерии MC-SSA «ev» и «cos» соответственно.

Введем понятие соотношения сигнал-шум. Обычно под ним понимают соотношение

$$\frac{\sum\limits_{n=1}^{N} s_n^2/N}{\mathsf{D}\boldsymbol{\xi}}, \tag{3.1}$$

где $\mathsf{S} = (s_1, \ldots, s_N)$ — сигнал, $\boldsymbol{\xi}$ — шум. Но, поскольку это определение не учитывает поведение спектральной плотности шума, обобщим его.

Определение 3.1. Будем называть

$$\mathrm{SNR}(\mathsf{S}, \boldsymbol{\xi}) = \frac{1}{N} \sum_{j=0}^{N-1} \frac{I_{\mathsf{S}}(j/N)}{f_{\boldsymbol{\xi}}(j/N)} \tag{3.2}$$

соотношением сигнал-шум, где $I_{\mathsf{S}}(\omega)$ — периодограмма сигнала, $f_{\boldsymbol{\xi}}(\omega)$ — спектральная плотность шума.

Замечание 3.1. Для белого шума с плотностью $f(\omega) = \sigma^2$ формулы (3.1) и (3.2) совпадают.

Будем рассматривать ряды длины $N = 100$ и $L \in \{10, 20, 50, 80, 90\}$. В качестве альтернативы рассмотрим

$$\mathsf{S} = \{Ae^{an}\cos(2\pi\omega n)\}_{n=1}^{N}.$$

Нас интересуют два конктетных случая:

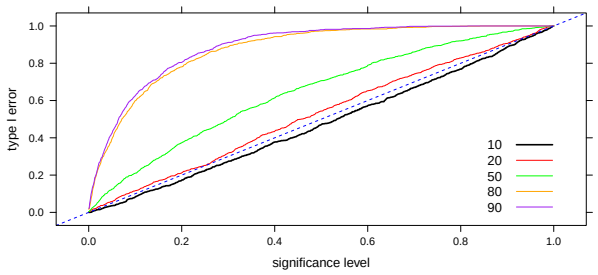
1. $a = 0$ — гармонический ряд;

2. $a \neq 0$ — экспоненциально-модулированный гармонический ряд.

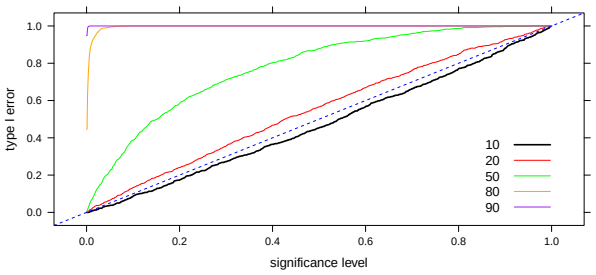
Помимо этого, будем рассматривать два возможных случая ω: когда рассматриваемые L (а значит и $2L$) делятся на $1/\omega$ и когда не делятся. Таким образом, рассмотрим 4 альтернативы:

1. $a = 0$, L делятся на $1/\omega$;
2. $a = 0$, L не делятся на $1/\omega$;
3. $a \neq 0$, L делятся на $1/\omega$;
4. $a \neq 0$, L не делятся на $1/\omega$.

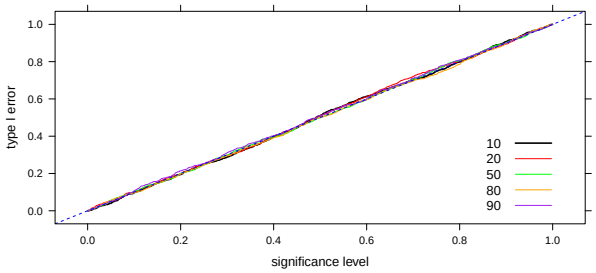
Пусть $\omega_1 = \omega_3 = 0.1$, $\omega_2 = \omega_4 = 0.085$, $A_1 = 0.9$, $A_3 = 0.02$, $a_3 = a_4 = 0.05$, A_2 и A_4 выбираются таким образом, чтобы $\mathrm{SNR}(\mathsf{S}_1, \boldsymbol{\xi}) \approx \mathrm{SNR}(\mathsf{S}_2, \boldsymbol{\xi})$ и $\mathrm{SNR}(\mathsf{S}_3, \boldsymbol{\xi}) \approx \mathrm{SNR}(\mathsf{S}_4, \boldsymbol{\xi})$. В качестве шума рассмотрим модель AR(1) с $\phi = 0.7$ и $\sigma^2 = 1$.



($а$) Проекция на собственные векторы $\mathbf{T}$



($б$) Проекция на собственные векторы $\mathbf{X}\mathbf{X}^{\mathrm{T}}$



($в$) Проекция на косинусы

Рис. 3.1. Ошибка первого рода

На рис. 3.1 изображены графики ошибок первого рода критериев «ev» и «cos» для рассматриваемых длин окна. Как видно по графикам, критерий «ev» радикальный для всех $L > 10$, а «cos», в свою очередь, является точным критерием для любой длины окна.

$a = 0$, $L\omega$ — целые	$L = 10$	$L = 20$	$L = 50$	$L = 80$	$L = 90$
Проекция на собственные векторы $\mathbf{T}$	0.424	0.484	0.525	0.582	0.646
Проекция на косинусы	0.543	0.546	0.546	0.614	0.646
$a = 0$, $L\omega$ — нецелые	$L = 10$	$L = 20$	$L = 50$	$L = 80$	$L = 90$
Проекция на собственные векторы $\mathbf{T}$	0.349	0.432	0.425	0.497	0.537
Проекция на косинусы	0.475	0.431	0.403	0.505	0.488
$a \neq 0$, $L\omega$ — целые	$L = 10$	$L = 20$	$L = 50$	$L = 80$	$L = 90$
Проекция на собственные векторы $\mathbf{X}\mathbf{X}^{\mathrm{T}}$	0.329	0.303	0.307	–	–
Проекция на косинусы	0.480	0.347	0.163	0.228	0.299
$a \neq 0$, $L\omega$ — нецелые	$L = 10$	$L = 20$	$L = 50$	$L = 80$	$L = 90$
Проекция на собственные векторы $\mathbf{X}\mathbf{X}^{\mathrm{T}}$	0.298	0.306	0.282	–	–
Проекция на косинусы	0.469	0.302	0.157	0.209	0.249

Таблица 3.1. Мощность поправленных критериев при уровне значимости $\alpha = 0.1$

В таблице 3.1 представлены значения мощности критериев после поправки, описанной в разделе 2.1.1 для рассмотренных L при уровне значимости $\alpha^* = 0.1$. Прочерки в таблице указывают на то, что критерий при данной длине окна слишком радикальный и построить поправку невозможно. Стоит отметить, что при $a \neq 0$ ряд становится нестационарным, поэтому приходится использовать собственные векторы $\mathbf{X}\mathbf{X}^{\mathrm{T}}$, которые дают для больших L слишком радикальный критерий и тем самым имеют потенциально меньшую мощность, чем при использовании собственных векторов $\mathbf{T}$.

По таблице видно, что в оптимальной длиной окна при постоянной амплитуде сигнала ($a \neq 0$) является $L = 90$. В случае же модуляции оптимальной является $L = 10$. Связано такое поведение мощности с тем, что при $a \neq 0$ частота ω, соответвующая сигналу, растекается по всему спектру частот и чем больше разрешение спектра (в данном случае L), тем сильнее это растекание. Если сравнивать по оптимальным L, то критерий «cos» дает наиболее мощный критерий в общем случае, когда модуляция непостоянная.

Учитывая огромную трудоемкость критерия «ev», и результат его численного сравнения с критерием «cos», рекомендуется использовать косинусы в качестве проекционных векторов, поскольку критерий не является радикальным (следовательно не требует

поправки, что заметно уменьшает трудоемкость MC-SSA) и по мощности не уступает, а в некоторых случаях даже превосходит вариант с векторами, порожденными исходным рядом. Далее будем рассматривать именно такой способ выборка векторов W_k.

3.2. SSA с проекцией

Базовый вариант SSA использует адаптивный базис для оценки подпространства сигнала, но существует возможность зафиксировать некоторые компоненты разложения. Пусть $\mathbf{D} \in \mathbb{R}^{L \times m}$ — матрица, проекцию на столбцы которой мы хотим зафиксировать в разложении $\mathbf{X}$. Тогда SSA с проекцией отличается от базового алгоритма только шагом разложения:

1. В случае, если столбцы матрицы $\mathbf{D}$ не ортонормированны, $\mathbf{D}$ приводится к нужному виду путем ортогонализации Грамма-Шмидта.
2. Вычисляется матрица $\mathbf{C} = \mathbf{D}\mathbf{D}^{\mathrm{T}}\mathbf{X}$.
3. Вычисляется матрица $\mathbf{X}^{\star} = \mathbf{X} - \mathbf{C}$.
4. Матрица $\mathbf{X}^{\star}$ раскладывается в сумму матриц ранга 1.

Замечание 3.2. Таким же образом можно определить матрицу-проектор на подпространство строк матрицы $\mathbf{X}$.

3.3. Автоматическая группировка в SSA

Для ряда X длины N и $0 \leqslant \omega_1 \leqslant \omega_2 \leqslant 0.5$ определим меру, следуя [22]:

$$T(\mathsf{X}; \omega_1, \omega_2) = \frac{1}{\|\mathsf{X}\|^2} \sum_{k:\omega_1 \leqslant k/N \leqslant \omega_2} I_N(k/N), \tag{3.3}$$

где I_N — периодограмма ряда X. Величину $T(\mathsf{X}, \omega_1, \omega_2)$ можно рассматривать как долю вклада частот, содержащегося в интервале $[\omega_1, \omega_2]$.

Будем выделять сигнал S по частоте $\hat{\omega}$ следующим образом:

1. Применить SSA с некоторой длиной окна L.
2. Выбрать первые r компонент разложения, у которых мера T на интервале $[\hat{\omega} - \delta, \hat{\omega} + \delta]$, $\delta > 0$, больше некоторого порога $T_0 \in [0, 1]$, где $r = 2$, если $\hat{\omega} \in (0, 0.5)$, иначе $r = 1$.

3.4. Метод autoMCSSA

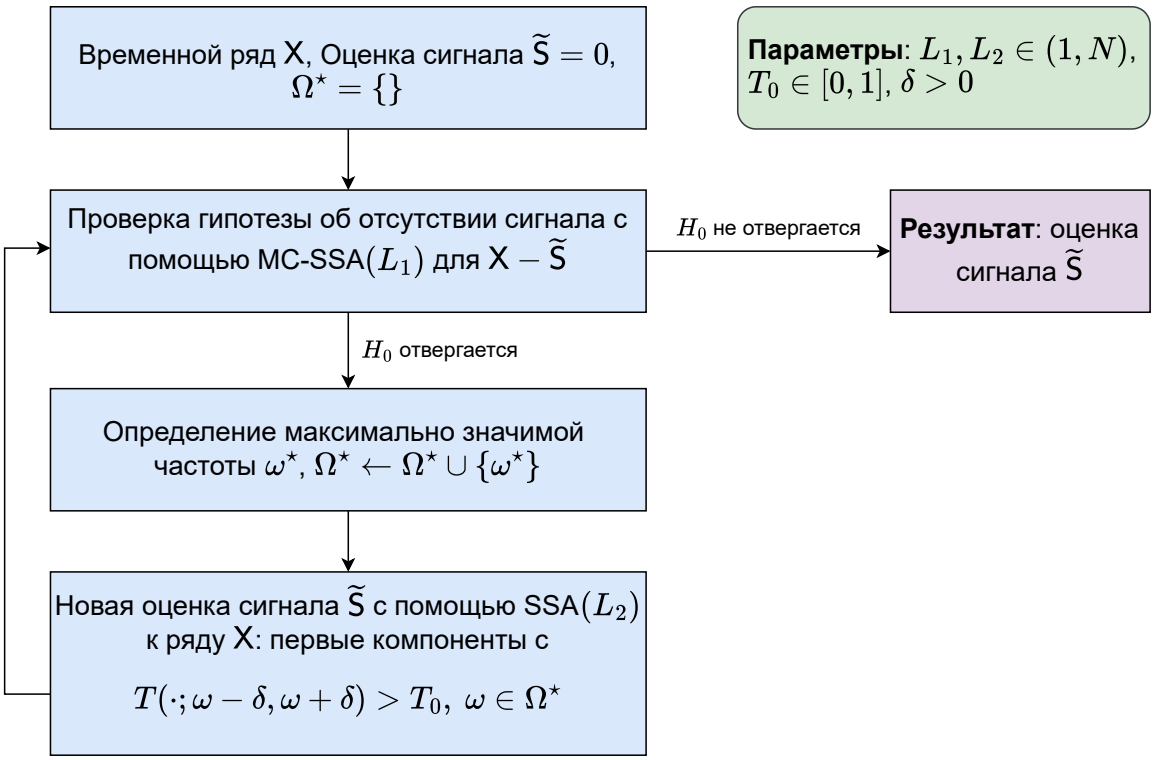


Рис. 3.2. Алгоритм autoMCSSA

На рис. 3.2 изображена блок-схема алгоритма autoMCSSA. Поскольку при оценке неизвестных параметрах шума метод MC-SSA может обнаружить не все частоты, принадлежащие сингалу, предлагается оценивать сигнал последовательно, применяя критерий MC-SSA к остатку ряда до тех пор, пока гипотеза $H_0 : \mathsf{S} = 0$ не перестанет отвергаться. Если на очередной итерации алгоритма гипотеза отвергается, определяется максимально значимая частота $\omega^\star$ и вычисляется новая оценка сигнала $\widetilde{\mathsf{S}}$ с помощью подхода, описанного разделе 3.3. Как только гипотеза перестает отвергаться, алгоритм завершает свою работу и тогда $\widetilde{\mathsf{S}}$ — итоговая оценка сигнала.

Оценивать частоту $\omega^\star$ будем с помощью MC-SSA:

1. Найти индекс наиболее значимой частоты, т.е. $k = \mathrm{argmax}_i(\widehat{p}_i - c_i)$, где c_i — верхняя граница доверительного интервала для $\widehat{p}_i$;

2. Вычислить значение $\hat{\omega}$ как взвешенное среднее частот $\omega_{k-1}, \omega_k, \omega_{k+1}$ с весами $w_i = \max(0, \widehat{p}_i - c_i)$;

Такой способ оценки позволяет получить более точную оценку ω в случае, когда она не попадает в решетку $k/(2L)$.

3.4.1. Пример работы алгоритма

Рассмотрим пример работы autoMCSSA. Пусть $\mathsf{X} = \mathsf{S} + \boldsymbol{\xi}$, где $\boldsymbol{\xi}$ — модель AR(1) с параметрами $\phi = 0.7$ и $\sigma^2 = 1$, $N = 200$, $\mathsf{S} = (s_1, \ldots, s_N)$,

$$s_n = 0.075\, e^{0.02n} \cos(2\pi n/8) + 2\cos(2\pi n/4) + 0.2 \cdot (-1)^n.$$

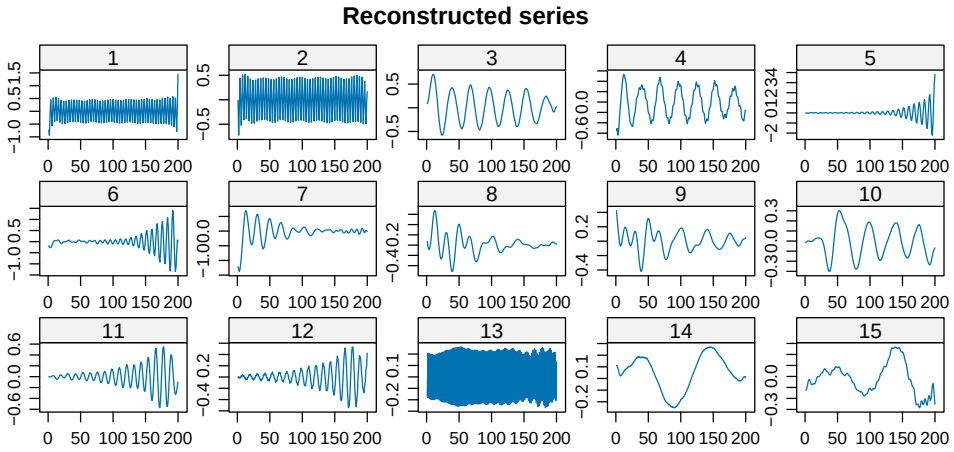


Рис. 3.3. Элементарные восстановленные компоненты

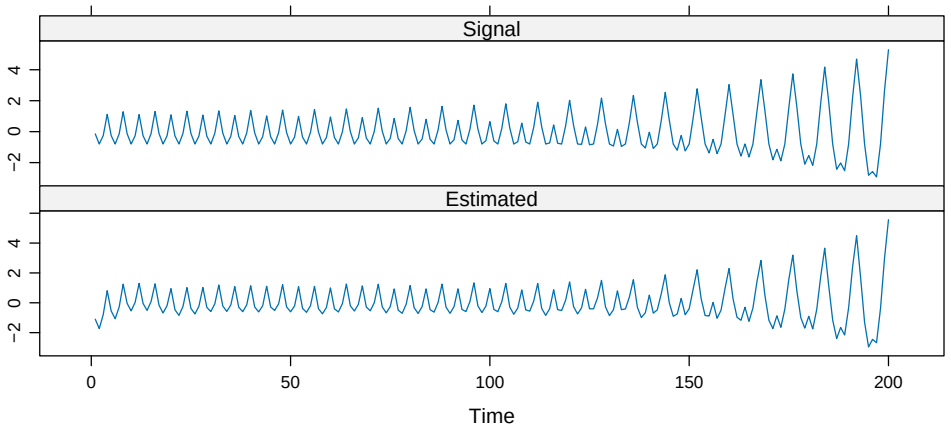


Рис. 3.4. Результат autoMCSSA ($L_1 = 50$, $L_2 = 100$, $\delta = 1/80$, $T_0 = 0.5$)

На рис. 3.3 представлены первые 15 элементарных восстановленных с помощью SSA компонент. Сигналу соответствуют компоненты с индексами 1, 2, 5, 6 и 13. Не

видя формулы, по которой этот сигнал задается, сказать наверняка, какие компоненты неслучайные, проблематично, поскольку компоненты 3, 4 и 11, 12 похожи на пары гармоник. Мы применили алгоритм autoMCSSA к этому ряду и получили, что разработанный метод правильно идентифицировал компоненты, соответствующие сигналу, на рис. 3.4 представлены истинная форма сигнала S и его оценка методом autoMCSSA.

3.4.2. Подходы к выделению сигнала

При описании алгоритма autoMCSSA для выделения сигнала S использовался базовый вариант SSA. Можно обобщить этот подход на SSA с проекцией, зафиксировав подходящий базис.

В данном разделе будем считать, что сигнал представляет из себя экспоненциально-модулированную гармонику:

$$\mathsf{S} = \{Ae^{an}\cos(2\pi\omega n)\}_{n=1}^{N},$$

где $\omega \in (0, 0.5)$. Пусть $\hat{\omega}$ — оценка ω. Обозначим

$$D_1 = \begin{pmatrix} \cos(2\pi\hat{\omega}1) \\ \dots \\ \cos(2\pi\hat{\omega}L) \end{pmatrix},\ D_2 = \begin{pmatrix} \sin(2\pi\hat{\omega}1) \\ \dots \\ \sin(2\pi\hat{\omega}L) \end{pmatrix} \in \mathbb{R}^L.$$

Рассмотрим следующие варианты выделения сигнала S по частоте $\hat{\omega}$:

1. «adaptive»: применить SSA и выбрать первые две компоненты разложения, у которых мера T на интервале $[\hat{\omega}-\delta, \hat{\omega}+\delta]$, $\delta > 0$, больше некоторого порога $T_0 \in [0, 1]$;

2. «semi-adaptive»: применить SSA с проекцией с $\mathbf{D} = D_1 \in \mathbb{R}^{L\times 1}$ и выбрать, помимо компоненты, соответствующей вектору D_1, первую компоненту разложения, у которой мера T на интервале $[\hat{\omega}-\delta, \hat{\omega}+\delta]$, $\delta > 0$, больше некоторого порога $T_0 \in [0, 1]$;

3. «fixed»: применить SSA с проекцией с $\mathbf{D} = [D_1 : D_2] \in \mathbb{R}^{L\times 2}$ и выбрать компоненты разложения, соответсвующие векторам D_1, D_2.

Заметим, что вариант «adaptive» соответствует методу, который использовался в алгоритма autoMCSSA.

Численное сравнение подходов

Проведем численный эксперимент с целью понять, какой из предложенных способов восстановления сигнала наиболее точен. Пусть $N = 99$, процесс $\boldsymbol{\xi}$ — модель AR(1) с параметрами $\phi = 0.7$, $\sigma^2 \in \{0.2, 0.4, 0.6, 0.8, 1\}$. Для SSA $L = 50$, для MC-SSA $L = \widetilde{L} = 40$ (выводы устойчивы к выбору длины окна). В вариантах «adaptive» и «semi-adaptive» $\delta = 0.025$ и $T_0 = 0.5$. Рассмотрим два типа сигнала S, один из которых является частным случаем другого:

1. $a = 0$, $A = 1$ — гармоника с постоянной амплитудой.

2. $a = 0.05$, $A = 0.025$ — экспоненциально-модулированная гармоника.

Возьмем $\omega = 0.115$. Заметим, что при таком выборе частоты сигнала $\widetilde{L}\omega$ не целое, а значит ω не попадает в решетку $k/(2\widetilde{L})$. Оценивать частоту ω будем также, как в разделе 3.4.

На рис. 3.5 изображена зависимость MSE восстановления сигнала от дисперсии белого шума σ^2. По графикам видно, что в случае постоянной амлитуды ($a = 0$) выигрывает вариант «fixed», однако в случае непостоянной амплитуды фиксированный базис оказывается наихудшим. Полуадаптивный базис, являясь неким компромиссом между адаптивным и фиксированным базисами, оказывается вторым по точности в случае $a = 0$ и сравнимым с адаптивным в рассмотренном случае $a \neq 0$. При увеличении $|a|$, начиная с какого-то момента, фиксированная половина базиса ухудшает восстановление сигнала.

Теперь посмотрим, как будут изменяться ошибки при уменьшении/увеличении δ для фиксированного $\widetilde{L}$. На рис. 3.6 δ уменьшена, а на рис. 3.7 увеличена в два раза ($\delta = 0.0125$ и 0.05 соответственно). Из этих графиков видно, что слишком маленькое δ приводит к ухудшению точности адаптивного и полуадаптивного вариантов в случае $a \neq 0$. Связано это с тем, что частота экспоненциально-модулированной гармоники, в отличие от гармоники с постоянной амплитудой, всегда растекается по спектру и чем больше абсолютное значение показателя экспоненты a, тем сильнее это растекание. Увеличение δ в два раза не привело к значительному изменению точности методов, однако, если и дальше увеличивать δ, ошибки, как и в случае слишком маленького δ, опять возрастут.

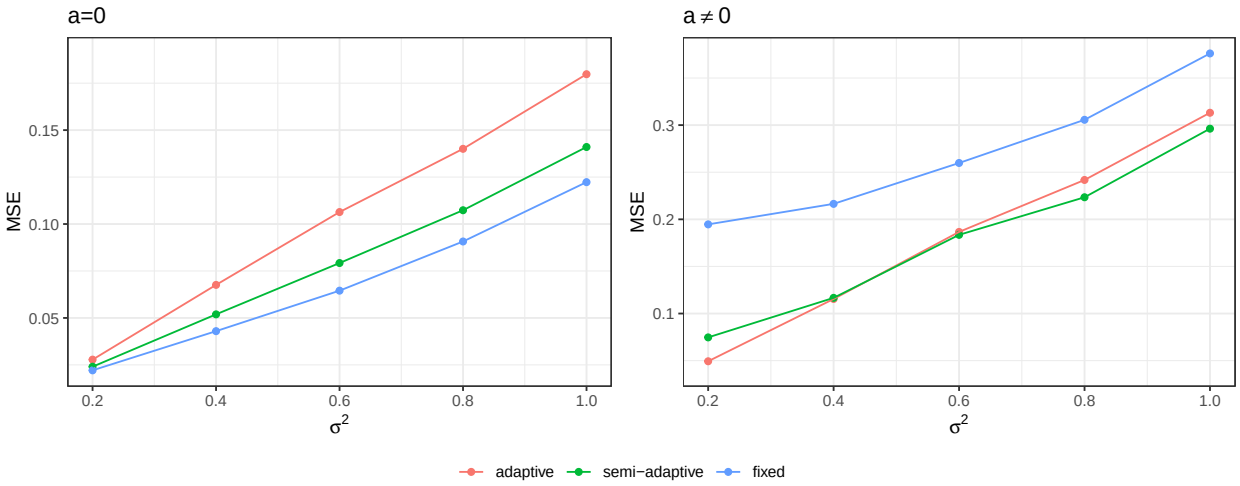


Рис. 3.5. MSE восстановления сигнала ($\delta = 0.025$)

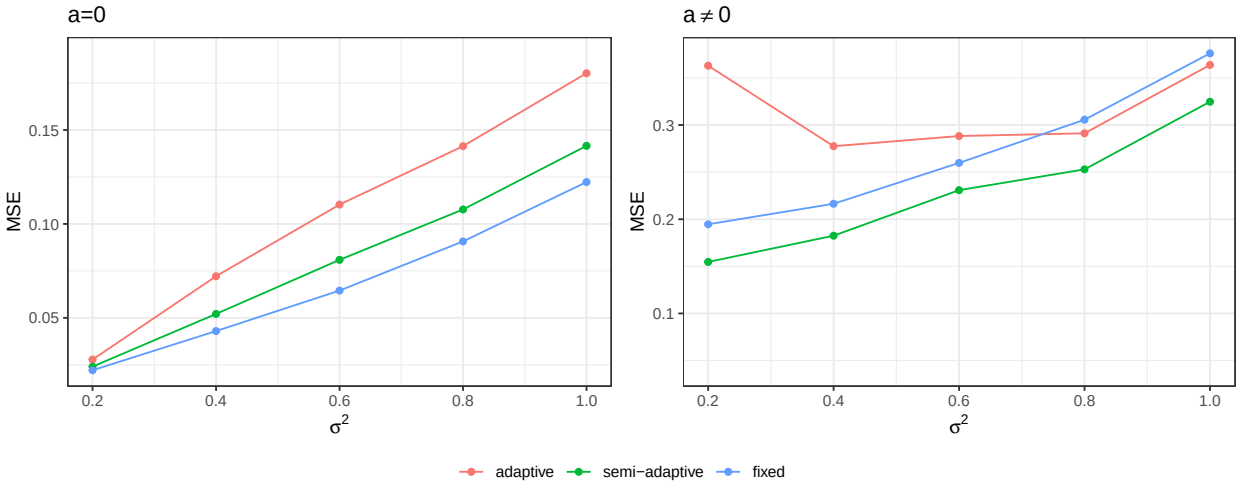


Рис. 3.6. MSE восстановления сигнала ($\delta = 0.0125$)

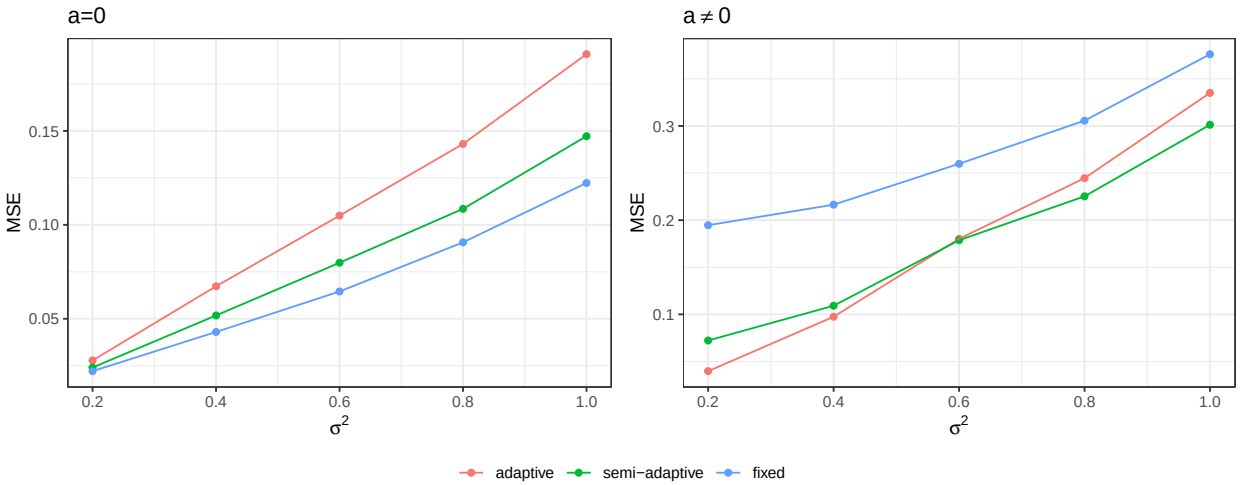


Рис. 3.7. MSE восстановления сигнала ($\delta = 0.05$)

Заключение

В ходе данной работы был реализован метод Monte Carlo SSA, когда в качестве модели шума рассматривается процесс с длинной памятью, а также его численное сравнение с Monte Carlo SSA с моделью красного шума, обладающей такой же дисперсией. Было получено, что если в качестве альтернативы рассмотреть сигнал с некоторой частотой, бóльшую мощность против этой альтернативы дает та модель шума, спектральная плотность которой в этой частоте наименьшая.

Помимо этого, было проведено численное сравнение различных методов оценки параметров модели ARFIMA(p, d, q). Для этого были реализованы методы максимального правдоподобия и Whittle, а также были взяты функции из пакетов языка программирования R. Получено, что при известном среднем метод максимального правдоподобия является наилучшим методом, дающим наименьшую среднеквадратичную ошибку и смещение параметров. Если же среднее неизвестно, наиболее предпочтительным методом является Whittle. Отметим, что реализованный метод максимального правдоподобия оказался лучше, чем в пакете arfima, а качественной реализации метода Whittle на момент написания работы обнаружено не было.

Список литературы

1. Broomhead D. S., King G. P. Extracting qualitative dynamics from experimental data // Physica D: Nonlinear Phenomena. — 1986. — Vol. 20, no. 2–3. — P. 217–236.
2. Golyandina N., Nekrutkin V., Zhigljavsky A. Analysis of Time Series Structure. — Chapman and Hall/CRC, 2001. — ISBN: 9780367801687.
3. Allen M. R., Smith L. A. Monte Carlo SSA: Detecting irregular oscillations in the Presence of Colored Noise // Journal of Climate. — 1996. — Vol. 9, no. 12. — P. 3373–3404.
4. Hipel Keith W., McLeod Ian. Time series modelling of water resources and environmental systems. — Elsevier, 1994.
5. Haslett John, Raftery Adrian E. Space-Time Modelling with Long-Memory Dependence: Assessing Ireland's Wind Power Resource // Journal of the Royal Statistical Society. Series C (Applied Statistics). — 1989. — Vol. 38, no. 1. — P. 1–50.
6. Long memory effects and forecasting of earthquake and volcano seismic data / Mariani Maria C., Bhuiyan Md Al Masum, Tweneboah Osei K. and Gonzalez-Huizar Hector // Physica A: Statistical Mechanics and its Applications. — 2020. — Vol. 559. — P. 125049.
7. Barkoulas J., Labys W. C., Onochie J. I. Fractional dynamics in international commodity prices // Journal of Futures Markets. — 1997. — Vol. 17. — P. 161–189.
8. Guglielmo Maria Caporale Luis Gil-Alana, Plastun Alex. Long memory and data frequency in financial markets // Journal of Statistical Computation and Simulation. — 2019. — Vol. 89, no. 10. — P. 1763–1779.
9. Time Series Analysis: Forecasting and Control / Box G., Jenkins G., Reinsel G., and Ljung G. — Fifth ed. — 2016.
10. Hassler Uwe. Time Series Analysis with Long Memory in View. — Wiley, 2018.
11. Palma Wilfredo. Long-Memory Time Series: Theory and Methods. — Wiley, 2006.
12. McLeod A. I., Yu Hao, Krougly Zinovi. Algorithms for Linear Time Series Analysis: With R Package // Journal of Statistical Software. — 2007. — Vol. 23, no. 5. — Access mode: `https://www.jstatsoft.org/v23/i05/`.
13. Whittle P. The Analysis of Multiple Stationary Time Series // Journal of the Royal Statistical Society. Series B (Methodological). — 1953. — P. 125–139.
14. Team R Core. — R: A Language and Environment for Statistical Computing. — R

Foundation for Statistical Computing, Vienna, Austria, 2024. — Access mode: `https://www.R-project.org/`.

15. Veenstra J.Q. — arfima: Fractional ARIMA (and Other Long Memory) Time Series Modeling : 2012.
16. Maechler Martin. — fracdiff: Fractionally Differenced ARIMA aka ARFIMA(P,d,q) Models : 1999.
17. Ларин Е. С. Метод SSA для проверки гипотезы о существовании сигнала во временном ряде : квалификационная работа магистра ; СПбГУ. — 2022.
18. Golyandina N. Detection of signals by Monte Carlo singular spectrum analysis: multiple testing // Statistics and Its Interface. — 2023. — Vol. 16, no. 1. — P. 147–157.
19. Gray Robert M. Toeplitz and Circulant Matrices: A Review // Foundations and Trends® in Communications and Information Theory. — 2005. — Vol. 2, no. 3. — P. 155–239.
20. Golyandina N., Korobeynikov A., Zhigljavsky A. Singular Spectrum Analysis with R. Use R! — Springer Berlin Heidelberg, 2018. — ISBN: 9783662573808.
21. Beran J. Statistics for Long-Memory Processes. — Chapman & Hall/CRC, 1994.
22. Alexandrov Th. A Method of Trend Extraction Using Singular Spectrum Analysis // RevStat. — 2009. — Vol. 7, no. 1. — P. 1–22.

Приложение А

Графики

А.1. Сравнение arfima_mle и arfima

На рис. А.1, А.2 и А.3 представлены среднеквадратичное отклонение, смещение и дисперсия оценок параметров ϕ и d модели ARFIMA$(1, d, 0)$. По ним, видно, что при $\phi = 0.1$ на рис. А.1 оценки функцией arfima имеют скачок в $d = 0.4$, что может говорить о некоторой вычислительной неустойчивости функции для больших d. Функция arfima_mle не только не имеет подобной проблемы, но и дает более точные оценки, например, при $\phi = 0.5$ на рис. А.2.

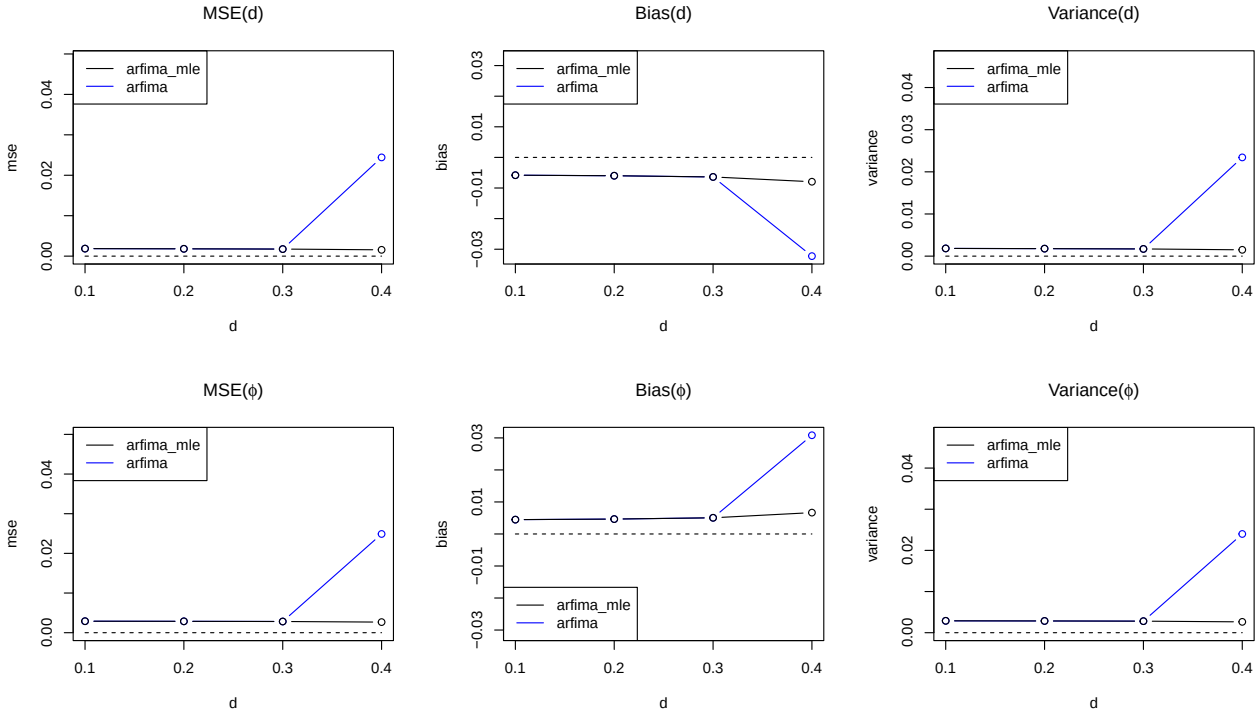


Рис. А.1. Сравнение arfima_mle и arfima при $\phi = 0.1$ ($n = 1000$, 500 повторений)

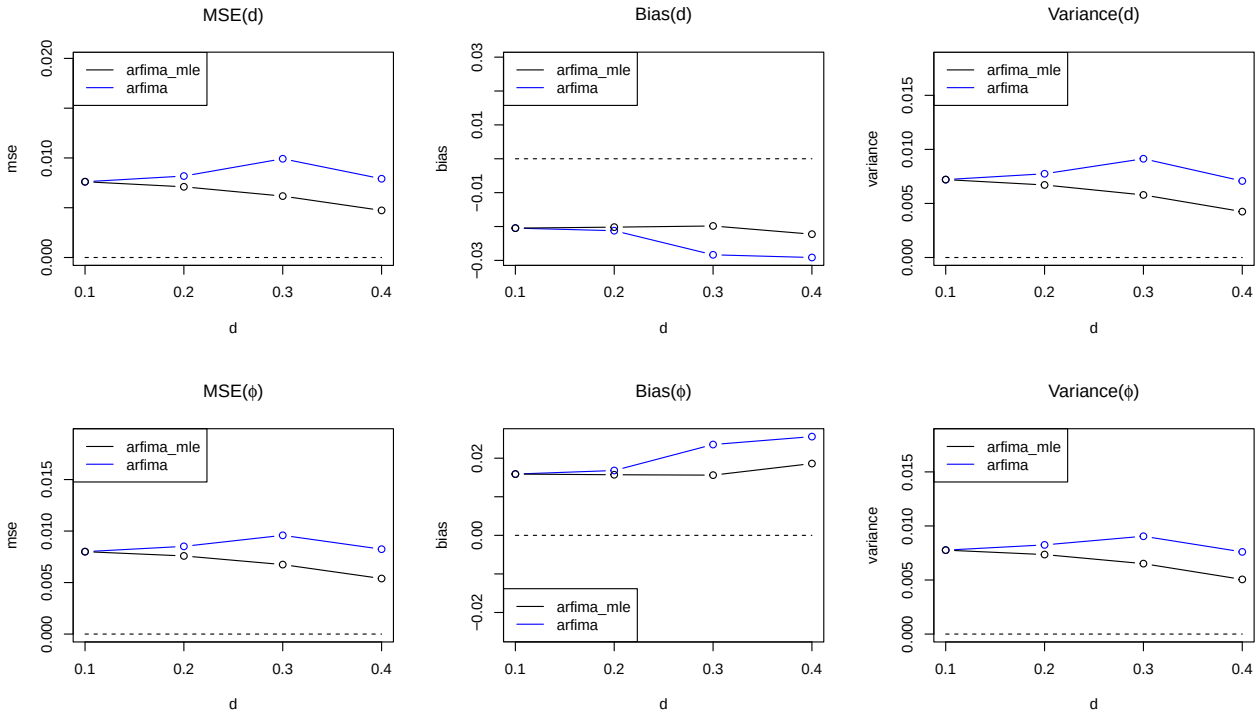


Рис. A.2. Сравнение arfima_mle и arfima при $\phi = 0.5$ ($n = 1000$, 500 повторений)

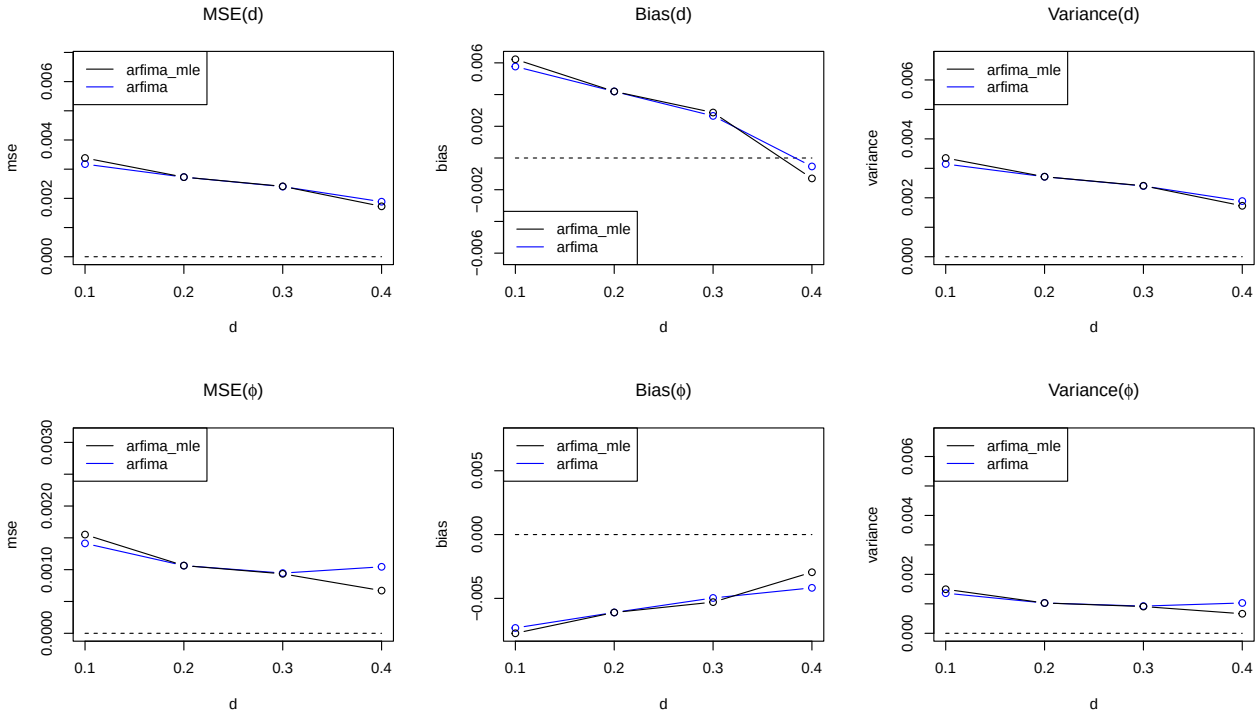


Рис. A.3. Сравнение arfima_mle и arfima при $\phi = 0.9$ ($n = 1000$, 500 повторений)