

Министерство образования и науки Российской Федерации
Федеральное агентство по образованию
Федеральное государственное образовательное учреждение высшего
профессионального образования «Санкт-Петербургский государственный
университет»
Математико-механический факультет
Кафедра статистического моделирования

Дипломная работа

Бояров Андрей Александрович

«ИССЛЕДОВАНИЕ СТАТИСТИЧЕСКИХ СВОЙСТВ МЕТОДА
МОНТЕ-КАРЛО SSA»

Допущена к защите

Заведующий кафедрой:

д. ф.-м. н., профессор С. М. Ермаков

Научный руководитель:

к. ф.-м. н., доцент Н. Э. Голяндина

Рецензент:

аспирант А. Ю. Шлемов

Санкт-Петербург

2012

Saint Petersburg State University
Faculty of Mathematics and Mechanics
Department of Statistical Modelling

Graduation Thesis

Boiarov Andrei Alexandrovich

INVESTIGATION OF STATISTICAL PROPERTIES OF MONTE CARLO SSA
METHOD

Admitted for defence

Head of Department:

Professor S. M. Ermakov

Scientific Supervisor:

Associate Professor N. E. Golyandina

Reviewer:

PhD student A. Y. Shlemov

Saint Petersburg

2012

Оглавление

Введение	5
Глава 1. Вспомогательные результаты	7
1.1. Метод «Гусеница»-SSA	7
1.2. Красный шум	9
1.3. Оценка параметров красного шума	9
1.3.1. Алгоритмы оценивания параметров красного шума	9
1.3.2. Сравнение результатов работы алгоритмов	11
1.3.3. Реакция оценок на наличие сигнала	14
Глава 2. Метод Монте-Карло SSA	18
2.1. Введение	18
2.2. Одномерный Монте-Карло SSA	18
2.3. Многомерный Монте-Карло SSA	19
2.3.1. Построение статистики критерия	21
2.3.2. Построение критерия	23
2.4. Сравнение методов с помощью оценок мощностей	37
2.4.1. Оценки мощностей	38
2.4.2. Сравнение методов	40
2.4.3. Выводы	46
2.5. Сравнение методов для различных параметров сигнала	47
2.6. Сравнение методов для различных параметров красного шума	49
2.7. Монте-Карло SSA с оценкой параметров красного шума	52
Заключение	55
Список литературы	56
Приложение А. Описание CD	57
Приложение Б. Результаты для эллипсоидов и кубов	58
Б.1. Сравнение методов выбора векторов для доверительных эллипсоидов и кубов	58

Б.2. Сравнение способов построения доверительных областей	61
Б.3. Сравнение методов для различных параметров сигнала	63
Б.4. Сравнение методов для различных параметров красного шума	65
Б.5. Монте Карло SSA с оценкой параметров красного шума	68

Введение

В данной работе рассматривается проблема обнаружения сигнала во временном ряде, являющимся смесью шума и сигнала. Во многих реальных задачах, возникающих, например, в климатологии или в геофизике, представляет интерес обнаружение слабого сигнала в «красном» шуме.

Задача обнаружения сигнала в шуме сводится к построению критерия для проверки гипотезы о том, что ряд является чистым шумом, мощного против альтернативы, заключающейся в наличии в ряде сигнала. Исходно в качестве критерия для решения поставленной задачи был взят метод Монте-Карло SSA, представленный в статье [1], и основанный на базовом методе «Гусеница»-SSA, описанном в [2]. Однако, в [1] формализация построения критерия недостаточна, а его статистические свойства исследованы мало.

Таким образом, задачей, поставленной в данной дипломной работе, является модифицирование метода Монте-Карло SSA так, чтобы он оставался применимым, и исследование статистических свойств построенных критериев. В рамках исследования проводится сравнение модификаций между собой по мощности, исследование влияния изменений различных параметров на результаты работы метода, исследование применения метода в реальных условиях.

Основной идеей метода Монте-Карло SSA является построение проекций скользящих отрезков длины L исходного временного ряда (эти отрезки составляют так называемую траекторную матрицу, возникающую в методе «Гусеница»-SSA) на специальным образом выбранные вектора. Величина проекции и является статистикой критерия. Затем для построения критической области оцениваются распределения построенных проекций с помощью моделирования, т.е. подхода Монте-Карло. В ситуации, когда вектор, на который строятся проекции, один, квантили распределения статистики критерия как раз и задают разбиение на критическую и доверительную области. Однако в многомерной ситуации, т.е. когда векторов много, построение такого разбиения не является однозначным. В описываемом в статье методе в качестве векторов, на которые строились проекции, брались сингулярные вектора, полученные после применения метода «Гусеница»-SSA к исходному временному ряду. Именно это и дало название методу «Монте-Карло SSA». В данной работе предлагается другой выбор векторов для проек-

ций, а также предлагаются разные варианты многомерных доверительных областей.

Сделаем краткий обзор содержания данной работы. В первой главе рассмотрены необходимые для дальнейших исследований результаты. В разделе 1.1 содержится описание метода «Гусеница»-SSA, на котором основан метод Монте-Карло SSA. В разделах 1.2 и 1.3 приведены результаты, касающиеся процесса авторегрессии первого порядка (красного шума): его формальное определение, алгоритмы построения оценок параметров красного шума и их сравнение. Во второй главе рассматривается непосредственно метод Монте-Карло SSA. В разделе 2.1 ставится задача, в разделе 2.2 описывается метод Монте-Карло SSA в одномерном случае. Раздел 2.3 содержит описание метода Монте-Карло SSA в многомерной ситуации, приводятся различные методы построения статистики критерия и самого критерия, исследуется корректность построенных критериев. Раздел 2.4 посвящен сравнению полученных критериев по мощности. Затем приводятся исследования влияния изменения различных параметров на результаты работы критериев: параметров сигнала — в разделе 2.5, параметров шума — в разделе 2.6. Сравнение критериев проводилось в искусственных предположениях, что параметры шума известны. Раздел 2.7 содержит исследование применения Монте-Карло SSA в реальных условиях, когда параметры шума оцениваются по временному ряду. В приложении располагается описание CD диска, на котором находятся программы, написанные на языке программирования «R», которые используются для получения описанных в данной работе результатов.

Глава 1

Вспомогательные результаты

1.1. Метод «Гусеница»-SSA

Рассмотрим базовый метод «Гусеница»-SSA, описанный в [2], на котором основан представленный в данной работе метод Монте-Карло SSA.

Будем рассматривать вещественнозначный временной ряд $F_N = (f_0, \dots, f_{N-1})$ длины $N > 2$. Предполагаем, что ряд ненулевой.

Первый шаг — *вложение*. Процедура вложения переводит исходный временной ряд в последовательность многомерных векторов. Пусть L — некоторое целое число, которое называется *длиной окна*, $1 < L < N$. Процедура вложения образует $K = N - L + 1$ векторов вложения

$$X = (f_{i-1}, \dots, f_{i+L-2})^T, \quad 1 \leq i \leq K.$$

Траекторная матрица ряда F

$$\mathbf{X} = (x_{ij})_{i,j=1}^{L,K} = [X_1 : \dots : X_K] \quad (1.1)$$

состоит из векторов вложения в качестве столбцов. Очевидно, что $x_{ij} = f_{i+j-2}$ и матрица \mathbf{X} имеет одинаковые элементы на «диагоналях» $i + j = \text{const}$. Таким образом, траекторная матрица является *ганкелевой*.

Второй шаг — *сингулярное разложение* траекторной матрицы ряда. Определим $L \times L$ симметричную матрицу:

$$\mathbf{S} = \mathbf{X}\mathbf{X}^T. \quad (1.2)$$

Матрица \mathbf{S} имеет L линейно независимых собственных векторов и L собственных чисел. Обозначим $\lambda_1, \dots, \lambda_L$ *собственные числа* матрицы \mathbf{S} , взятые в неубывающем порядке ($\lambda_1 \geq \dots \geq \lambda_L \geq 0$) и U_1, \dots, U_L — ортонормированную систему *собственных векторов* матрицы S , соответствующую собственным числам.

Пусть $d = \max\{i : \lambda_i > 0\}$. Пусть $V_i = \mathbf{X}^T U_i / \sqrt{\lambda_i}$, $i = 1, \dots, d$, тогда сингулярное разложение матрицы \mathbf{X} может быть записано как

$$\mathbf{X} = \mathbf{X}_1 + \dots + \mathbf{X}_d, \quad (1.3)$$

где $\mathbf{X}_i = \sqrt{\lambda_i} U_i V_i^T$. Набор $(\sqrt{\lambda_i}, U_i, V_i)$ назовем *i-ой собственной тройкой* сингулярного разложения (1.3).

Третий шаг — *группировка собственных троек*. Процедура группировки делит все множество индексов $1, \dots, d$, полученных в результате разложения (1.3), на m непересекающихся подмножеств I_1, \dots, I_m .

Рассмотрим такое подмножество $I = i_1, \dots, i_p$. Результирующая матрица \mathbf{X}_I , соответствующая I , определяется как

$$\mathbf{X}_I = \mathbf{X}_{i_1} + \dots + \mathbf{X}_{i_p}.$$

Вычислив такие матрицы для всех групп $I = I_1, \dots, I_m$, разложение (1.3) можно записать в сгруппированном виде

$$\mathbf{X} = \mathbf{X}_{I_1} + \dots + \mathbf{X}_{I_m}. \quad (1.4)$$

Четвертый шаг — *диагональное усреднение*. На этом шаге алгоритма каждая матрица сгруппированного разложения (1.4) преобразуется в ряд длины N .

Пусть \mathbf{X}_{I_k} — результирующая матрица из (1.4). Тогда ее преобразование соответствует усреднению элементов матрицы вдоль «диагоналей» $i + j = \text{const}$. Если \mathbf{X}_{I_k} является траекторной матрицей некоторого ряда (h_0, \dots, h_{N-1}) , то h_i — есть элемент, лежащий на соответствующей «диагонали». Применяя диагональное усреднение к результирующим матрицам \mathbf{X}_{I_k} , получаем восстановленные ряды $\check{F}_N^{(k)} = (\check{f}_0^{(k)}, \dots, \check{f}_{N-1}^{(k)})$. Исходный ряд $F_N = (f_0, \dots, f_{N-1})$ раскладывается в сумму m рядов

$$f_n = \sum_{k=1}^m \check{f}_n^{(k)}.$$

Отметим, что если рассматривать $L \times L$ матрицу

$$\mathbf{U} = [U_1 : \dots : U_L], \quad (1.5)$$

состоящую из собственных векторов матрицы \mathbf{S} в качестве столбцов, то можно определить матрицу

$$\mathbf{\Lambda} = \mathbf{U}^T \mathbf{S} \mathbf{U}. \quad (1.6)$$

$\mathbf{\Lambda}$ — диагональная матрица, k -ым диагональным элементом которой является k -ое собственное число матрицы \mathbf{S} .

1.2. Красный шум

На практике (например в геофизике и климатологии) для описания реальных данных часто возникает потребность рассматривать модель авторегрессии первого порядка (AR(1)):

$$\xi_i = u + \gamma(\xi_{i-1} - u) + \delta\varepsilon_i, \quad (1.7)$$

где γ, δ — параметры процесса, $|\gamma| < 1, \delta > 0$, u — среднее процесса, ε_i — Гауссовский белый шум, $E\varepsilon_i = 0$, $D\varepsilon_i = 1$. Если ξ_t и ξ_{t+l} — элементы процесса авторегрессии первого порядка, то их автоковариация равна

$$c_l = \text{cov}(\xi_t, \xi_{t+l}) = \frac{\delta^2}{1 - \gamma^2} \gamma^{|l|}. \quad (1.8)$$

Опишем алгоритм моделирования AR(1) процесса (для простоты $u = 0$).

Алгоритм 1. Пусть нужно промоделировать ряд длины N . Исходя из формул (1.7) и (1.8), получаем: $\xi_0 \sim N(0, \sigma^2)$, $\xi_i = \gamma\xi_{i-1} + \delta\varepsilon_i$, где $\varepsilon_i \sim N(0, 1)$, $i = 1, \dots, N - 1$.

В результате такого моделирования получается временной ряд, который имеет название *красный шум*, если $\gamma > 0$.

1.3. Оценка параметров красного шума

В реальных ситуациях редко бывают известны параметры красного шума, поэтому возникает потребность в их оценивании. Рассмотрим алгоритмы построения оценок параметров γ, δ и u , а также их сравнение и поведение в случае наличия в ряде сигнала.

1.3.1. Алгоритмы оценивания параметров красного шума

Первый алгоритм

Пусть $F_N = (f_0, \dots, f_{N-1})$ — временной ряд, являющийся реализацией красного шума с параметрами γ, δ, u . \bar{f} — выборочное среднее F_N , оценка u . Так как f_t и f_{t+l} — элементы красного шума, то автоковариация c_l определяется по формуле (1.8). Оценку для автоковариации можно получить как

$$\hat{c}_l \equiv \frac{1}{N-l} \sum_{i=1}^{N-l} (f_i - \bar{f})(f_{i+l} - \bar{f}). \quad (1.9)$$

Используя формулы (1.8) и (1.9) для оценок двух различных автоковариаций, вычисляем результат для $\hat{\gamma}$ и $\hat{\delta}$. Наиболее точные оценки получаются для $l = 0$ и $l = 1$ в формуле (1.9) [1]. Таким образом

$$\hat{\gamma} = \frac{\hat{c}_1}{\hat{c}_0}, \quad (1.10)$$

$$\hat{\delta} = \sqrt{\frac{\hat{c}_0^2 - \hat{c}_1^2}{\hat{c}_0}}. \quad (1.11)$$

Замечание 1. В силу определения красного шума (1.7) в случае, когда длина ряда N мала, а параметр γ велик, оценка среднего \bar{f} неточна. Поэтому, если известен u , следует в формуле (1.9) использовать его.

Второй алгоритм

Этот алгоритм использует результаты, полученные в предыдущем алгоритме, и основан на утверждениях из статьи [1].

В [1] предлагается оценка автоковариации:

$$\tilde{c}_1 = \hat{c}_1 + \tilde{c}_0 \mu^2, \quad (1.12)$$

где μ^2 задаётся равенством

$$\mu^2 \equiv \frac{1}{N} + \frac{1}{N^2} \sum_{k=1}^{N-1} 2(N-k)\gamma^k = -\frac{1}{N} + \frac{2}{N^2} \left(\frac{N - \gamma^N}{1 - \gamma} - \frac{\gamma(1 - \gamma^{N-1})}{(1 - \gamma)^2} \right). \quad (1.13)$$

Промежуточная оценка $\tilde{\gamma}$ находится как решение уравнения

$$\frac{\hat{c}_1}{\hat{c}_0} = \frac{\tilde{\gamma} - \mu^2(\tilde{\gamma})}{1 - \mu^2(\tilde{\gamma})}. \quad (1.14)$$

Обозначим правую часть (1.14) $g(\tilde{\gamma})$. Решение (1.14) получим с помощью метода Ньютона-Рафсона:

1. $\tilde{\gamma}_1 = \hat{c}_1/\hat{c}_0$.
2. $\tilde{\gamma}_{n+1} = \tilde{\gamma}_n - (g(\tilde{\gamma}_n))/(g'(\tilde{\gamma}_n))$, где $g'(\tilde{\gamma}) = (1 - \mu^2 - (\mu^2)' + (\mu^2)'\tilde{\gamma})/((1 - \mu^2)^2)$,
 $(\mu^2(\tilde{\gamma}))' = 2(N - N\tilde{\gamma} + N\tilde{\gamma}^N - N\tilde{\gamma}^{N+1} - 1 - \tilde{\gamma} + \tilde{\gamma}^N + \tilde{\gamma}^{N+1})/(N^2(1 - \tilde{\gamma})^3)$.

В [1] утверждается, что процедура сходится за 2–5 итерации. Получив таким образом $\tilde{\gamma}$, мы находим \tilde{c}_0 из

$$\tilde{c}_0 = \frac{\hat{c}_0}{1 - \mu^2(\tilde{\gamma})}, \quad (1.15)$$

\tilde{c}_1 вычисляем по формуле (1.12). Оценки $\tilde{\gamma}$ и $\tilde{\delta}$ получаем аналогично выражениям (1.10) и (1.11) соответственно:

$$\tilde{\gamma} = \frac{\tilde{c}_1}{\tilde{c}_0},$$

$$\tilde{\delta} = \sqrt{\frac{\tilde{c}_0^2 - \tilde{c}_1^2}{\tilde{c}_0}}.$$

Третий алгоритм (ARIMA)

Рассмотрим метод, предложенный в статье [3], с помощью которой можно оценить параметры авторегрессии первого порядка.

1.3.2. Сравнение результатов работы алгоритмов

Для результатов оценок γ и δ вычислим $E(\theta - \hat{\theta})^2$, $D(\hat{\theta})$ и $(\theta - E\hat{\theta})^2$, где θ — параметр, $\hat{\theta}$ — его оценка.

Результаты будем считать для разных N — длин временного ряда, разных γ при фиксированном δ . Временной ряд будет являться реализацией процесса авторегрессии первого порядка.

Были проведены эксперименты для $N = 50$ и $N = 500$, при $\delta = 1$, $u = 0$ и 1000 повторов алгоритмов.

Рассмотрим для оценок γ ($N = 50$) квадраты отклонений, дисперсии и квадраты смещений. Из рисунка 1.1 видно, что оценки, полученные по второму алгоритму, почти всюду точнее. Также можно сделать вывод, что с ростом γ убывает $D(\hat{\gamma})$, а $(\gamma - E\hat{\gamma})^2$ возрастает.

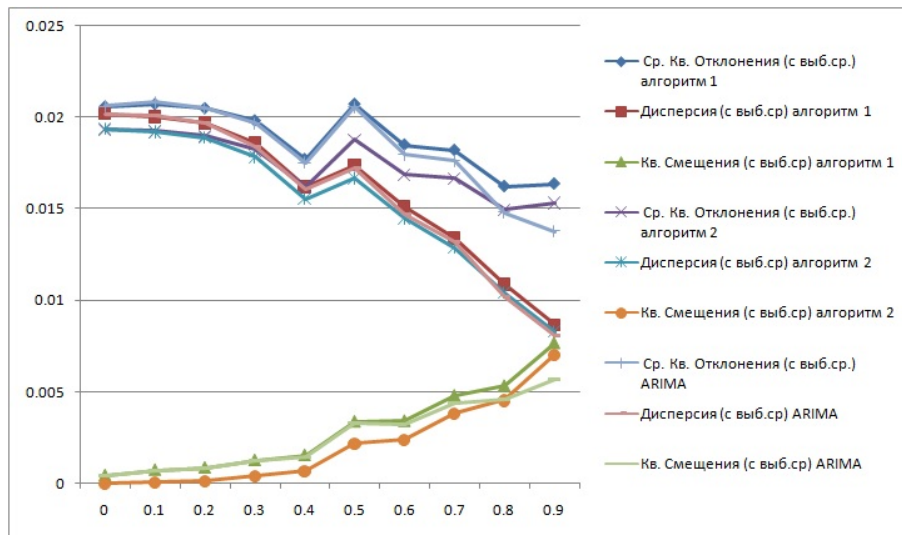


Рис. 1.1. Зависимость средних квадратов отклонений, дисперсий и квадратов смещения оценок γ от γ , $N = 50$.

Про оценки δ при $N = 50$ можно сказать, что они практически одинаковы для всех алгоритмов. Это отображено на рисунке 1.2.

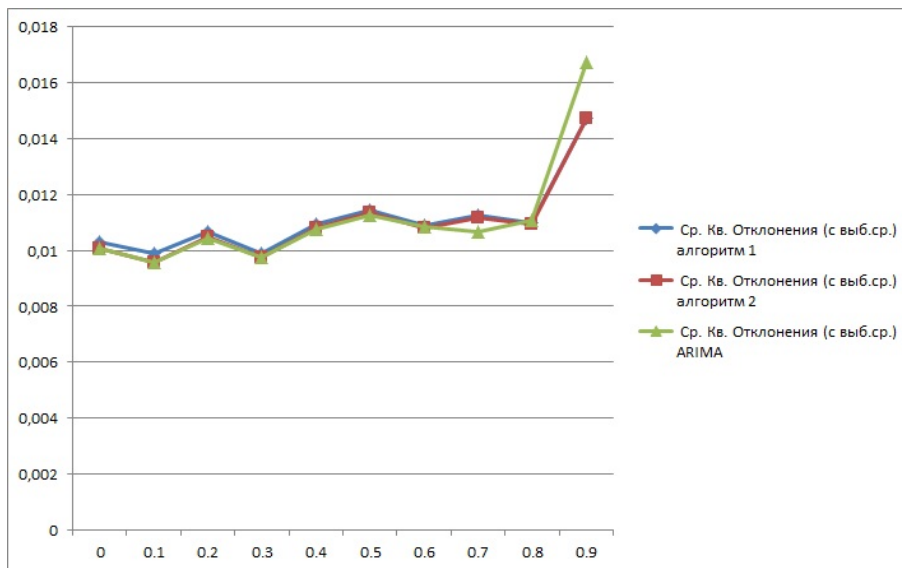


Рис. 1.2. Зависимость средних квадратов отклонений оценок δ от γ , $N = 50$.

При больших длинах временного ряда ($N = 500$) оценки параметра γ , полученные по результатам работы всех трех алгоритмов, практически совпадают. Это хорошо заметно на рисунке 1.3, где изображены графики $E(\gamma - \hat{\gamma})^2$, $D(\hat{\gamma})$ и $(\gamma - E\hat{\gamma})^2$.

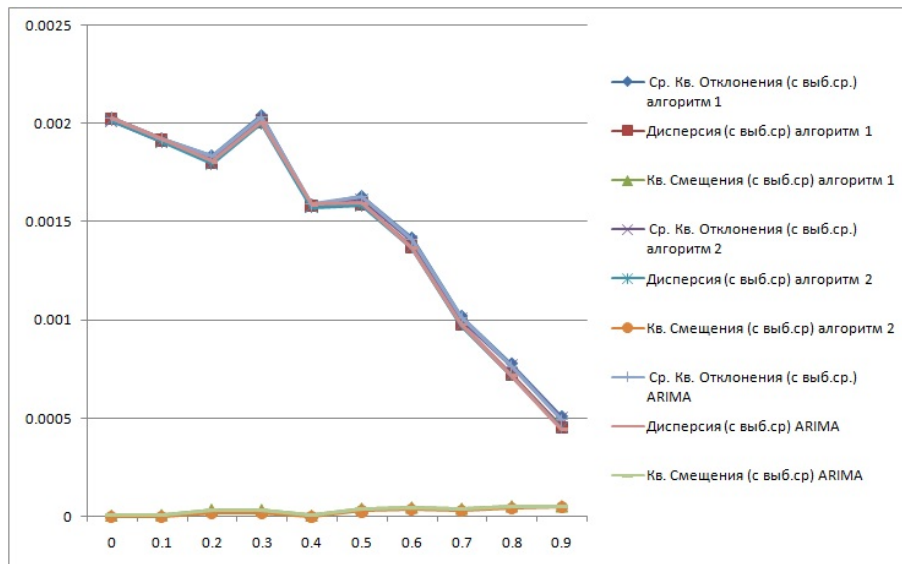


Рис. 1.3. Зависимость средних квадратов отклонений, дисперсий и квадратов смещения оценок γ от γ , $N = 500$.

При больших N оценки δ по всем трем алгоритмам ведут себя почти аналогично случаю малых N , но являются более точными (рисунок 1.4).

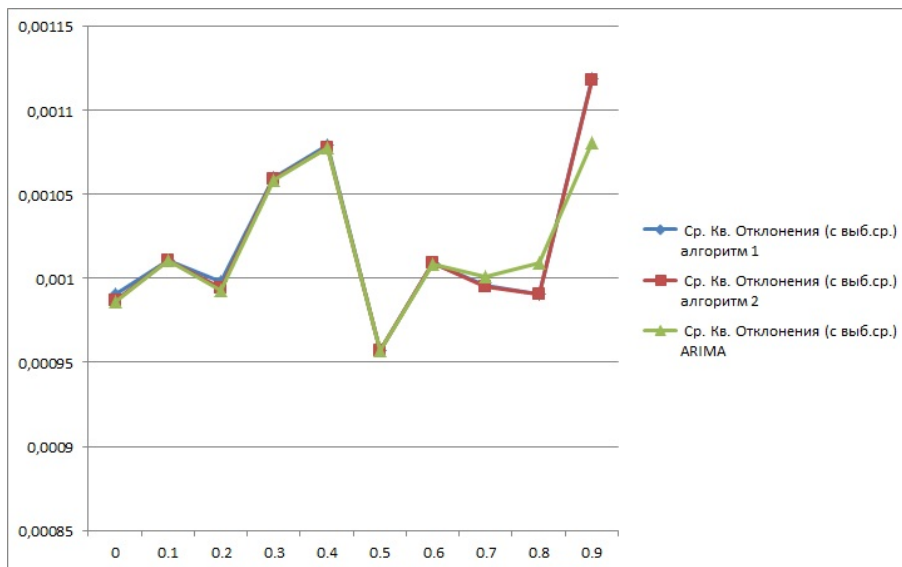


Рис. 1.4. Зависимость средних квадратов отклонений оценок δ от γ , $N = 500$.

В итоге по результатам экспериментов можно сделать вывод, что при небольшой длине ряда нужно применять второй алгоритм, так как он даёт более точную оценку параметра γ , нежели первый и третий алгоритмы.

1.3.3. Реакция оценок на наличие сигнала

Ранее мы рассматривали результаты работы алгоритмов оценки параметров AR(1) процесса применительно к чистому шуму. Теперь установим их реакцию на наличие сигнала. Для этого вычислим для оценок γ и δ с добавлением сигнала и без него $E(\theta - \hat{\theta})^2$, где θ — параметр, $\hat{\theta}$ — его оценка.

Проведем эксперименты для $N = 50$ и $N = 500$, $\delta = 1$, $u = 0$, 1000 повторов, при добавлении сигнала $s_i = \sin(2\pi i/5.5)$, $i = 0, \dots, N - 1$.

При небольшой длине ряда получаем график средних квадратов смещений, изображенный на рисунке 1.5.

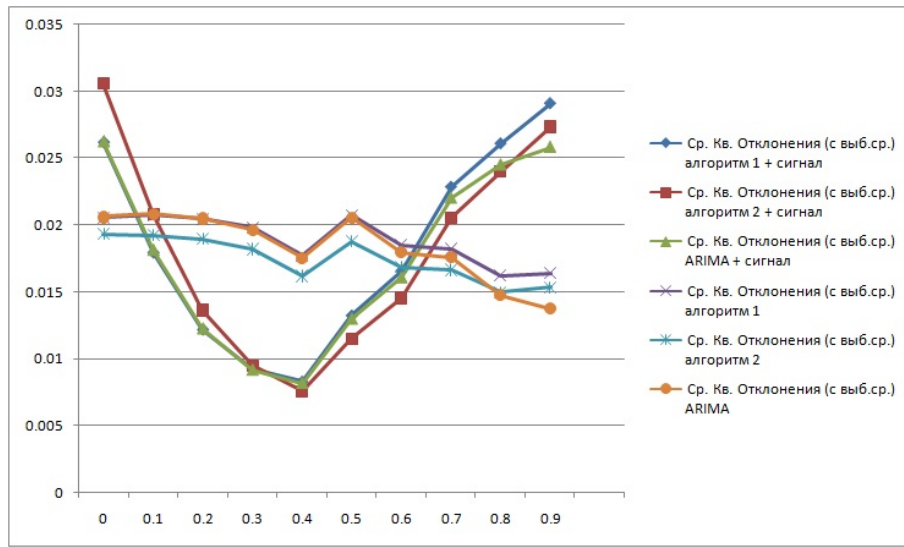


Рис. 1.5. Зависимость средних квадратов отклонений оценок γ от γ , $N = 50$.

Из рисунка 1.5 видно, что при некоторых γ добавление сигнала улучшает оценку, в частности при $\gamma = 0.4$ получается наиболее точная оценка.

Такой эффект объясняется влиянием на результаты работы всех трех алгоритмов сигнала $s_i = \sin(2\pi i/5.5)$, $i = 0, \dots, N - 1$. Это хорошо видно, если применить алгоритмы к чистому сигналу. В этой ситуации оценка параметра γ определяется формулой

$$\hat{\gamma} = \cos(2\pi/T), \quad (1.16)$$

где T — период сигнала (в рассматриваемом случае $T = 5.5$). Поведение функции $\cos(2\pi/T)$ при $T = 1, \dots, N$, где $N = 200$, показано на рисунке 1.6.

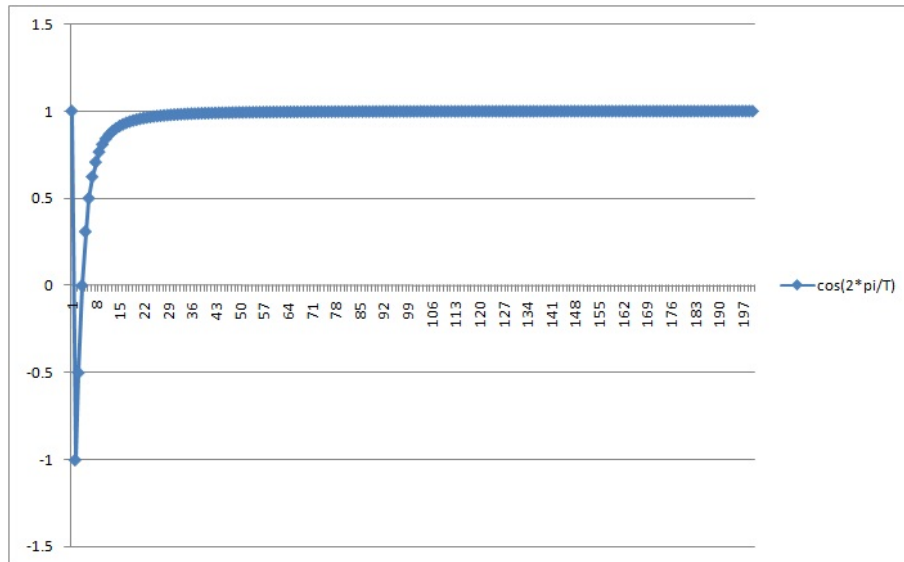


Рис. 1.6. Зависимость $\cos(2\pi/T)$ от T при $T = 1, \dots, 200$.

Рассмотрим результаты работы трех алгоритмов, примененных к чистому синусу при $T = 5.5$, (рисунок 1.7).

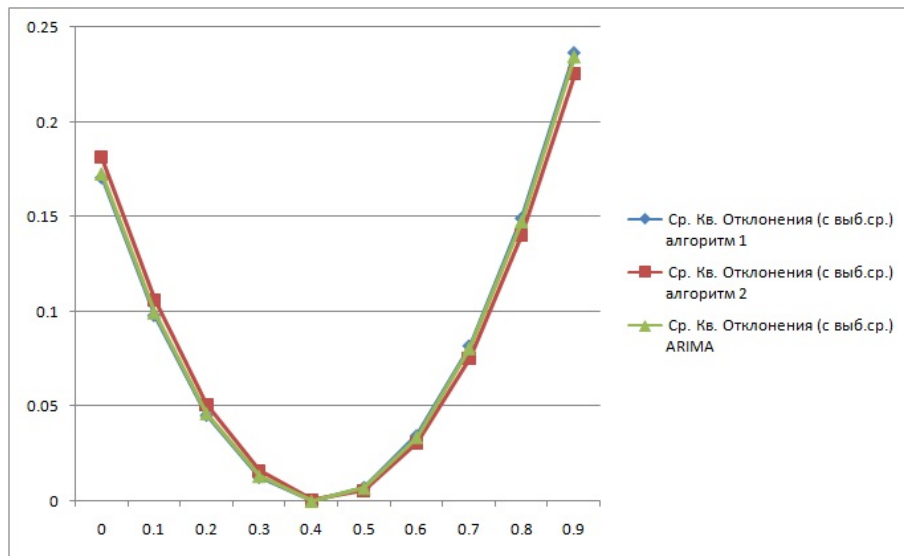


Рис. 1.7. Зависимость средних квадратов отклонений оценок γ от γ , период сигнала 5.5.

Оценки при чистом сигнале, полученные по формуле (1.16), сильно влияют на оценки параметра γ при применении алгоритмов к временному ряду, который состоит из красного шума и сигнала. Рисунок 1.5 построен при добавлении к AR(1) процессу синуса с периодом 5.5. Влияние оценок параметра, полученных на чистом синусе с $T = 5.5$, видно из рисунка 1.7.

Про оценку δ при длине ряда $N = 50$ можно сказать, что при наличии сигнала она

ухудшается относительно оценки без сигнала, что проиллюстрировано на рисунке 1.8.

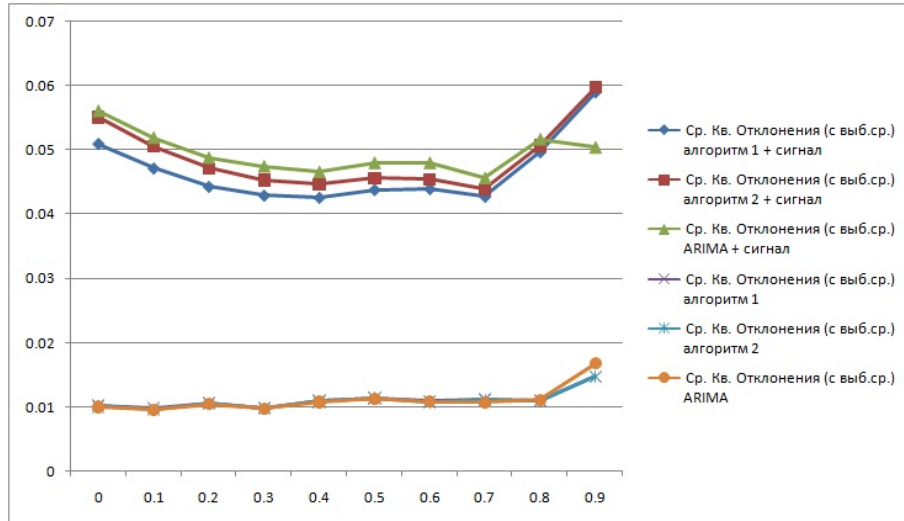


Рис. 1.8. Зависимость средних квадратов отклонений оценок δ от γ , $N = 50$.

Рассмотрим график средних квадратов смещений оценок γ для большой длины ряда.

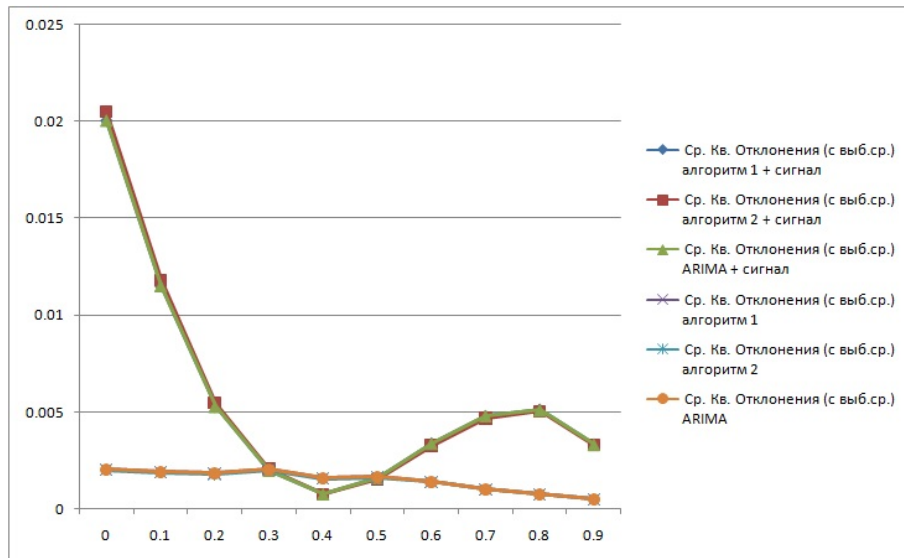


Рис. 1.9. Зависимость средних квадратов отклонений оценок γ от γ , $N = 500$.

Из рисунка 1.9 видно, что для некоторых γ оценка получилась точнее при добавлении сигнала. Этот эффект аналогичен такому же случаю при небольшой длине ряда. Меньшее, чем при $N = 50$, количество более точных результатов с сигналом объясняется большой длиной ряда.

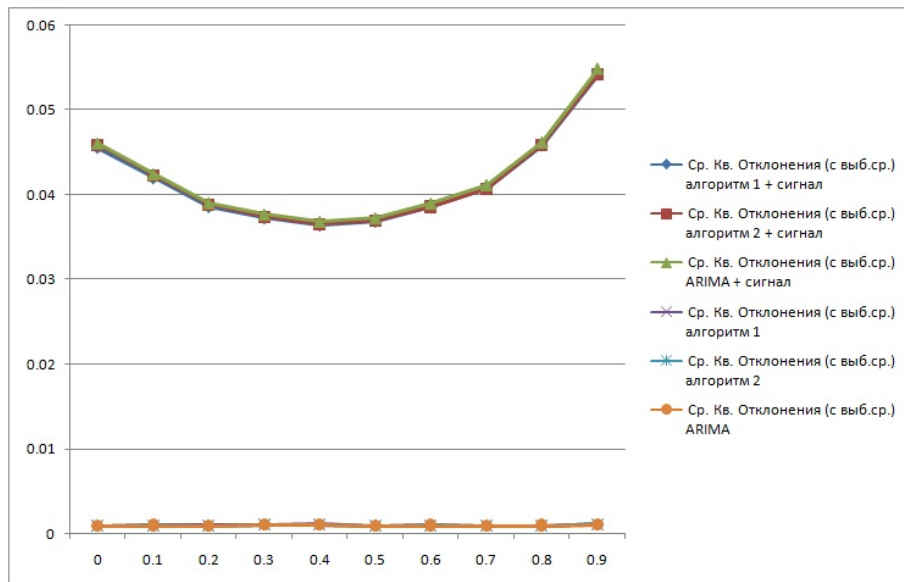


Рис. 1.10. Зависимость средних квадратов отклонений оценок δ от γ , $N = 500$.

При больших N оценка δ становится хуже при добавлении сигнала (рисунок 1.10).

Глава 2

Метод Монте-Карло SSA

2.1. Введение

Пусть $F_N = (f_0, \dots, f_{N-1})$ — исходный ряд. Рассмотрим модель, согласно которой временной ряд имеет структуру $F_N = S_N + R_N$, где $R_N = (r_0, \dots, r_{N-1})$ является красным шумом, а $S_N = (s_0, \dots, s_{N-1})$ — некоторым детерминированным рядом, например $s_i = A \sin(2\pi\omega i + \varphi)$, $i = 0, \dots, N-1$, где A — амплитуда, ω — частота, φ — фаза колебания.

Метод Монте-Карло SSA применяется для проверки гипотезы $H_0 : S_N = 0$ против альтернативы $H_1 : S_N \neq 0$.

Для проверки нулевой гипотезы необходимо выбрать уровень значимости α , построить статистику критерия t и критическую область. На основе попадания t в критическую область принимается решение о том, отвергается или не отвергается гипотеза H_0 .

Пусть после выполнения первых двух шагов базового алгоритма «Гусеница»-SSA получено разложение (1.3) исходного ряда F_N . Тогда, если после применения метода Монте-Карло SSA H_0 отвергается в пользу альтернативной гипотезы, то появляется возможность определить собственные тройки, принадлежащие сигналу с помощью результатов, полученных в ходе работы Монте-Карло SSA. Благодаря этому после процедур группировки и диагонального усреднения можно восстановить сигнал S_N .

2.2. Одномерный Монте-Карло SSA

В этом разделе описан алгоритм Монте-Карло SSA для одного вектора. Входные данные: W — некоторый фиксированный заранее выбранный вектор, α — заданный уровень значимости, $F_N = (f_0, \dots, f_{N-1})$ — временной ряд, L — длина окна, G — число итераций для метода Монте-Карло.

1. Построение статистики критерия.

- а. Используя метод «Гусеница»-SSA, по формуле (1.1) находим траекторную матрицу \mathbf{X} исходного ряда F_N .

- б. Определим статистику критерия как квадрат нормы величины проекции F_N на направление, задаваемое вектором W

$$t = \|\mathbf{X}^T W\|^2. \quad (2.1)$$

2. Построение критерия.

- а. Чтобы построить критическую область, необходимо оценить распределение статистики критерия t , при условии, что H_0 верна. Для этого применим подход Монте-Карло. Промоделируем, используя алгоритм 1, ряд длины N $\xi_N^i = (\xi_0^i, \dots, \xi_{N-1}^i)$. Таким образом, ξ_N^i является реализацией красного шума.
- б. С помощью формулы (1.1) строим траекторную матрицу промоделированного ряда Ξ_i и вычисляем квадрат нормы величины проекции ξ_N^i на направление, задаваемое вектором W , по формуле (2.1) как

$$p_i = \|\Xi_i^T W\|^2. \quad (2.2)$$

- в. Прделав процедуру, описанную в пункте 2б, для всех $i = 1, \dots, G$, где G — заданное число, получаем набор величин (p_1, \dots, p_G) , найденных по формуле (2.2), по которому находятся выборочные $(\alpha/2)$ и $(1 - \alpha/2)$ квантили, доверительная область задается отрезком между этими квантилями.

Замечание 2. Для моделирования красного шума в пункте 2а необходимы его параметры γ, δ . В дальнейших разделах для исследования статистических свойств метода будем считать, что параметры известны. Однако, в реальных задачах параметры $AR(1)$ процесса неизвестны, поэтому в разделе 2.7 рассматривается использование Монте-Карло SSA в случае неизвестных параметров красного шума.

Замечание 3. Описанный выше критерий является точным, если параметры γ, δ известны, иначе — нет.

2.3. Многомерный Монте-Карло SSA

Для того чтобы описанный выше критерий был мощным против альтернативы H_1 , необходимо выбирать вектор W наиболее близким по структуре к сигналу S_N (по

частоте, если S_N — синус). В реальных задачах, когда ничего не известно о сигнале S_N или информация о нем приблизительная, сложно определить один конкретный вектор W для применения к нему метода Монте-Карло SSA. В таких ситуациях имеет смысл рассматривать несколько векторов, таким образом критерий становится многомерным.

Входные данные совпадают с одномерным случаем, только вместо одного вектора W определим набор интересующих нас векторов W_1, \dots, W_H . Пусть матрица

$$\mathbf{W} = [W_1 : \dots : W_H]$$

является матрицей, содержащей эти вектора в качестве столбцов.

Статистика критерия — вектор $\bar{t} = (t_1, \dots, t_H)$, где каждый его элемент t_k определяется аналогично формуле (2.1)

$$t_k = \|\mathbf{X}^T W_k\|^2, \quad k = 1, \dots, H. \quad (2.3)$$

В матричном виде \bar{t} можно записать как вектор диагональных элементов матрицы

$$\mathbf{\Lambda}_F = \mathbf{W}^T \mathbf{X} \mathbf{X}^T \mathbf{W}, \quad (2.4)$$

где \mathbf{X} — траекторная матрица исходного ряда F_N .

Согласно общему алгоритму Монте-Карло SSA для оценки распределения статистики критерия \bar{t} необходимо построить набор проекций на вектора W_1, \dots, W_H . Для каждого вектора W_k найдем выборку проекций $P_k = (p_{k1}, \dots, p_{kG})^T$, где каждый элемент p_{ki} вычисляется по формуле (2.2)

$$p_{ki} = \|\Xi_i^T W_k\|^2, \quad i = 1, \dots, G,$$

Ξ_i — траекторная матрица красного шума $\xi_N^i = (\xi_0^i, \dots, \xi_{N-1}^i)$, промоделированного с помощью алгоритма 1.

В матричном виде P_k записывается как вектор, состоящий из диагональных элементов матрицы

$$\mathbf{\Lambda}_R = \mathbf{W}^T \mathbf{\Xi}_i \mathbf{\Xi}_i^T \mathbf{W}. \quad (2.5)$$

Построив P_k для всех $k = 1, \dots, H$, определим матрицу

$$\mathbf{P} = \begin{pmatrix} P_1^T \\ \vdots \\ P_H^T \end{pmatrix}, \quad (2.6)$$

строками которой являются выборки проекций красного шума на вектора W_1, \dots, W_H . Теперь с помощью матрицы \mathbf{P} оценивается распределение статистики критерия \bar{t} и строятся критическая и доверительная области. Способы построения критерия и статистики критерия оказывает влияние на реальную вероятность ошибки первого рода и на мощность. В данной работе рассмотрено несколько способов и проведено их сравнение.

2.3.1. Построение статистики критерия

При построении статистики критерия \bar{t} и матрицы \mathbf{P} встает вопрос выбора векторов W_1, \dots, W_H . Далее рассматриваются несколько методов определения этой системы.

Собственные вектора красного шума

Данный метод предложен в [1]. Определим W_1, \dots, W_H как собственные вектора автоковариационной матрицы $\text{AR}(1)$ процесса, которая определяется как $\mathbf{S}_R = c_0 \mathbf{\Gamma}$, где c_0 — дисперсия красного шума, $\Gamma_{ij} = \gamma^{|i-j|}$, γ — параметр $\text{AR}(1)$ процесса. Таким образом, система векторов W_1, \dots, W_H зависит от параметра γ .

В дальнейшем будем обозначать данный метод, как «С.в. $\text{AR}(1)$ ».

Замечание 4. В силу замечания 2, при известных параметрах красного шума в методе «С.в. $\text{AR}(1)$ » W_1, \dots, W_H определены однозначно, однако при неизвестном параметре γ , система векторов становится неоднозначной: $W_k = W_k(\hat{\gamma})$, $k = 1, \dots, H$, где $\hat{\gamma}$ — оценка γ .

Синусы и косинусы

В данном методе выбора векторов система W_1, \dots, W_H является однозначно определенной, т.е. при нахождении t_k и p_{ki} по формулам (2.1) и (2.2) соответственно, проекции строятся на одни и те же вектора.

В силу того, что собственные вектора автоковариационной матрицы красного шума при $L \rightarrow \infty$ стремятся к чередующимся синусам и косинусам, частота которых меняется с шагом $\approx 1/(2L)$ (статья [4]), то представляется целесообразным выбрать в качестве W_1, \dots, W_H вектора, являющиеся попеременно синусами и косинусами с шагом по частоте $1/(2L)$.

Для повышения вероятности попадания частоты сигнала S_N в частотную сетку рассмотрим также синусы и косинусы с шагом $1/(4L)$ и $1/(8L)$.

Имеет смысл выбирать W_1, \dots, W_H так, чтобы они образовывали ортогональную систему векторов. Для этого будем рассматривать вектора, которые являются синусами с шагом по частоте $1/L$.

В дальнейшем будем обозначать данный методы в зависимости от шага по частоте как «sin,cos($1/2L$)», «sin,cos($1/4L$)», «sin,cos($1/8L$)», «sin($1/L$)».

Собственные числа красного шума

Этот способ предложен в [1]. В данном методе W_1, \dots, W_H выбираются на каждой итерации алгоритма, таким образом система векторов не является однозначно определенной.

При нахождении статистики критерия t_k вектора W_1, \dots, W_H являются собственными векторами автоковариационной матрицы исходного ряда F_N , определяемой по формуле (1.2), т.е. $W_k = W_k(F_N)$, $k = 1, \dots, H$.

При нахождении p_{ki} для каждого $i = 1, \dots, G$ в качестве W_1, \dots, W_H выбирают собственные вектора автоковариационной матрицы промоделированного согласно нулевой гипотезе ряда ξ_N^i (т.е. красного шума), построенной аналогично (1.2), таким образом, $W_k = W_k(\Xi_i)$, $k = 1, \dots, H$.

В дальнейшем будем обозначать данный метод как «С.ч.АР(1)».

Методы построения статистики критерия

При построении критерия с помощью описанных выше методов выбора системы векторов может возникнуть ситуация, когда сигнал S_N из альтернативной гипотезы имеет частоту, которая не попадает в частотную сетку, которая образована векторами W_1, \dots, W_H . В такой ситуации возникает эффект растекания — в случае верной H_1 статистика критерия не попадает в доверительную область на двух соседних векторах. Для того чтобы избавиться от растекания, будем рассматривать минимумы в соседних парах и суммы в соседних парах. Кроме того, непопадание в доверительную область на паре соседних собственных векторов в методе «Гусеница»-SSA соответствует сигналу (синусу), что актуально при использовании метода выбора векторов, основанного на собственных числах красного шума.

Пусть \bar{t} — общее обозначение для вектора статистики критерия, \mathbf{P} — общее обозначение для матрицы проекций. Вектор, полученный с помощью формулы (2.3), будем обозначать \bar{t}^{sing} , матрицу, полученную по (2.6), — \mathbf{P}^{sing} .

Для минимумов необходимо найти $\bar{t}^{\text{min}} = (t_1^{\text{min}}, \dots, t_{H-1}^{\text{min}})$, где $t_i^{\text{min}} = \min(t_i^{\text{sing}}, t_{i+1}^{\text{sing}})$ для всех $i = 1, \dots, (H-1)$. Теперь непопадание в доверительную область t_i^{min} означает непопадание для пары соседних векторов. Таким же образом для каждого набора проекций на направления, задаваемые векторами W_1, \dots, W_H , будем находить наборы $P_k^{\text{min}} = (p_{k1}^{\text{min}}, \dots, p_{kG}^{\text{min}})^T$, где $p_{kj}^{\text{min}} = \min(p_{kj}^{\text{sing}}, p_{(k+1)j}^{\text{sing}})$ для всех индексов $k = 1, \dots, (H-1)$, $j = 1, \dots, G$.

Аналогично для сумм соседних пар определяется $\bar{t}^{\text{sum}} = (t_1^{\text{sum}}, \dots, t_{H-1}^{\text{sum}})$, где элемент $t_i^{\text{sum}} = (t_i^{\text{sing}} + t_{i+1}^{\text{sing}})$ для всех $i = 1, \dots, (H-1)$, $P_k^{\text{sum}} = (p_{k1}^{\text{sum}}, \dots, p_{kG}^{\text{sum}})^T$, каждый $p_{kj}^{\text{sum}} = (p_{kj}^{\text{sing}} + p_{(k+1)j}^{\text{sing}})$ для всех $k = 1, \dots, (H-1)$, $j = 1, \dots, G$.

Теперь для \bar{t} и \mathbf{P} строятся критические и доверительные области.

2.3.2. Построение критерия

Рассмотрим различные методы построения доверительной области и приведем для них оценки вероятности ошибки первого рода для всех методов выбора векторов и методов построения критерия.

Интерпретация проверки критерия как испытания Бернулли

Рассмотрим выборки вида $\bar{b}_M = (b_1, \dots, b_M)$, где M — число проверок нулевой гипотезы,

$$b_i = \begin{cases} 0, & \text{если на } i\text{-ой итерации } H_0 \text{ не отвергается} \\ 1, & \text{если на } i\text{-ой итерации } H_0 \text{ отвергается.} \end{cases} \quad (2.7)$$

Таким образом, $\bar{b}_M = (b_1, \dots, b_M)$ можно интерпретировать как серию из M испытаний Бернулли с вероятностью успеха r . Для r строится доверительный интервал по методу Уилсона [5]

$$\left\{ \theta \mid Z_{\nu/2} \leq \frac{\hat{r} - \theta}{\sqrt{\theta(1-\theta)/M}} \leq Z_{1-\nu/2} \right\}, \quad (2.8)$$

где \hat{r} — оценка значения вероятности r для выборки \bar{b}_M , $\nu = 0.05$, $Z_{\nu/2}$, $Z_{1-\nu/2}$ — соответствующие квантили стандартного нормального распределения.

Построение критерия

Как было описано в разделе 2.2, доверительная область является отрезком, лежащим между $(\alpha/2)$ и $(1-\alpha/2)$ квантилями, т.е критерий является двусторонним. Однако, в силу того, что мы строим критерий против альтернативы $H_1 : S_N \neq 0$, статистика критерия в случае верности H_1 попадает выше доверительной области, поэтому имеет смысл рассматривать односторонний критерий, когда доверительная область лежит от $-\infty$ до $(1-\alpha)$ квантили. В дальнейшем, чтобы иметь возможность сравнивать по мощности двусторонние и односторонние критерии, будем задавать уровень значимости $\alpha = 0.1$ в первом случае и $\alpha = 0.05$ во втором.

Критерий будем считать построенным, если вероятность ошибки первого рода равна (примерно равна) заданному уровню значимости. Одномерный метод Монте-Карло SSA (раздел 2.2) при известных параметрах красного шума удовлетворяет этому условию.

При проверке условия, что критерий построен, будем рассматривать для вероятности ошибки первого рода $H_0 : \alpha_I = \alpha$, строя доверительную область для такой гипотезы по формуле (2.8), где вместо \hat{r} подставляется α , статистикой критерия является оценка вероятности ошибки первого рода $\hat{\alpha}_I$. В случае уровня значимости $\alpha = 0.1$ доверительная область равна $(0.083, 0.120)$, $\alpha = 0.05$ — $(0.038, 0.065)$.

Пусть для какого-либо метода выбора векторов определены все вектора W_1, \dots, W_Z , где Z в зависимости от метода может быть равным $L/2, L, 2L, 4L$. Будем строить оценки для трех диапазонов соседних номеров векторов, в соответствии с которыми будут выбраны W_1, \dots, W_H из W_1, \dots, W_Z : большой, включающий в себя все номера, средний, включающий половину, и узкий, содержащий только десятую часть всех номеров.

Все оценки будем строить для одних и тех же реализаций красного шума со следующими параметрами: $N = 200$, $\gamma = 0.72$, $\delta = 1$, $u = 0$, $L = 40$, $M = 1000$ — число повторов метода Монте-Карло SSA для построения оценок вероятности ошибки первого рода и мощности, $G = 1000$ — число моделирований красного шума при построении проекций на вектора. Отметим, что в силу выбора таких параметров алгоритма, для метода выбора векторов « $\sin(1/L)$ » не возникает эффект растекания, поэтому при критериях для пар не будем рассматривать этот метод.

Простая доверительная область

Построим для каждого вектора W_k , $k = 1, \dots, H$ доверительную область как в случае одномерного Монте-Карло SSA (раздел 2.2). Тогда нулевая гипотеза отвергается, если хотя бы на одном векторе статистика критерия попадает в критическую область. Критерий будем строить односторонним. Рассмотрим результаты оценивания вероятности ошибки первого рода, жирным шрифтом выделены оценки, которые не попадают в доверительную область гипотезы для соответствующего уровня значимости.

Простая доверительная область ($\alpha = 0.05$)						
Диап.	С.ч.AR(1)	С.век.AR(1)	$\sin, \cos(1/2L)$	$\sin, \cos(1/4L)$	$\sin, \cos(1/8L)$	$\sin(1/L)$
Бол.	0.473	0.781	0.739	0.797	0.819	0.606
Ср.	0.359	0.543	0.503	0.573	0.606	0.380
Узк.	0.236	0.369	0.348	0.373	0.335	0.263
Простая доверительная область для минимумов в соседних парах ($\alpha = 0.05$)						
Диап.	С.ч.AR(1)	С.век.AR(1)	$\sin, \cos(1/2L)$	$\sin, \cos(1/4L)$	$\sin, \cos(1/8L)$	$\sin(1/L)$
Бол.	0.479	0.782	0.771	0.830	0.830	—
Ср.	0.357	0.539	0.533	0.610	0.608	—
Узк.	0.198	0.339	0.324	0.392	0.336	—
Простая доверительная область для сумм в соседних парах ($\alpha = 0.05$)						
Диап.	С.ч.AR(1)	С.век.AR(1)	$\sin, \cos(1/2L)$	$\sin, \cos(1/4L)$	$\sin, \cos(1/8L)$	$\sin(1/L)$
Бол.	0.429	0.732	0.715	0.790	0.806	—
Ср.	0.309	0.477	0.477	0.563	0.582	—
Узк.	0.174	0.305	0.289	0.360	0.316	—

Как видно из приведенных в таблице выше данных, реальная вероятность ошибки первого рода сильно отличается от заданного уровня значимости. Это связано с зависимостью между проекциями на вектора. Если бы они были независимы, то настоящая вероятность ошибки первого рода была бы равной $1 - (1 - \alpha)^d$, где d — число рассматриваемых векторов (длина диапазона номеров), а если бы они все были зависимыми между собой, то настоящая вероятность ошибки первого рода была бы равной α . Но в действительности ситуация не подпадает ни под один из этих крайних случаев, можно

только констатировать тот факт, что между проекциями на некоторые вектора существует ненулевая корреляция. В качестве примеров, ниже приведены корреляционные матрицы для методов выбора векторов: собственные числа красного шума, синусы и косинусы с шагом по частоте $1/(2L)$. Корреляции в остальных методах имеют поведение аналогичное второму примеру.

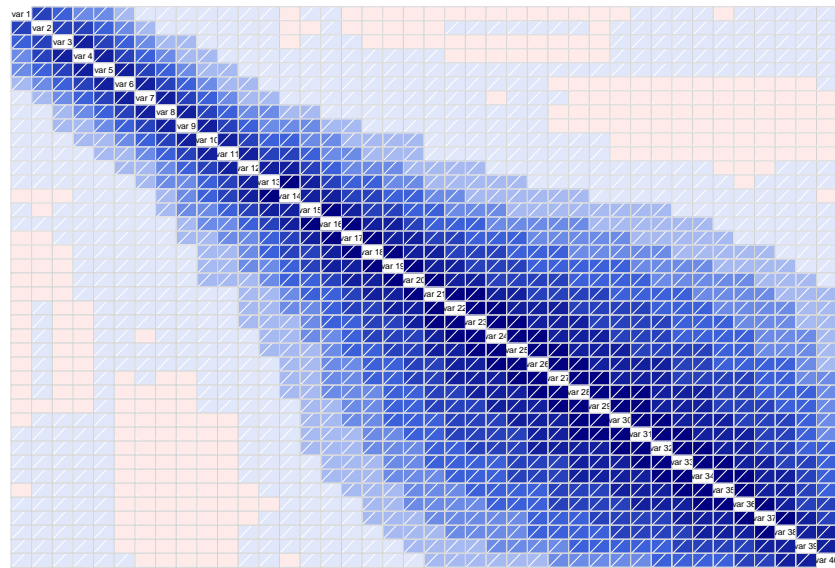


Рис. 2.1. Корреляционная матрица проекций для метода: «С.ч.AR(1)».

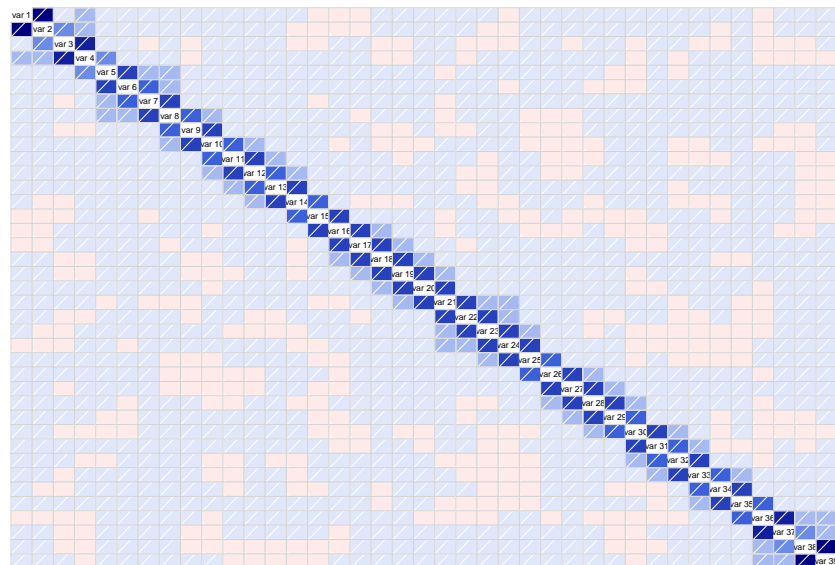


Рис. 2.2. Корреляционная матрица проекций для метода: «sin,cos(1/2L)».

Таким образом необходимо рассматривать другие методы построения доверительной области, учитывающие зависимости между проекциями на вектора.

Теоретический доверительный эллипсоид в нормальной модели

Будем строить доверительную область как доверительный эллипсоид. Для этого рассмотрим утверждение.

Утверждение 1. Пусть $\xi \in \mathbf{R}^m$ имеет многомерное нормальное распределение $\mathcal{N}(\mu, \Sigma)$. Тогда

1. Расстояние Махаланобиса между ξ и μ : $\rho^2(\xi, \mu) = (\xi - \mu)^T(\Sigma)^{-1}(\xi - \mu) \sim \chi^2(m)$.
2. Если $\hat{\Sigma}$ — оценка ковариационной матрицы Σ , то $\rho^2(\xi, \mu) = (\xi - \mu)^T(\hat{\Sigma})^{-1}(\xi - \mu) \sim T_m^2(G - 1)$, где T^2 — распределение Хотеллинга со степенью свободы $(G - 1)$, G — число наблюдений.

Исходя из пункта 1 утверждения 1 необходимо привести к нормальному распределению статистику критерия и матрицу проекций. Например, используем для этого логарифмирование вектора \bar{t} и матрицы \mathbf{P} . В результате получим $\bar{t}^{\log} = (t_1^{\log}, \dots, t_H^{\log})$, где $t_i^{\log} = \log(t_i)$ для всех $i = 1, \dots, H$ и

$$\mathbf{P}^{\log} = \log(\mathbf{P}) = \begin{pmatrix} \log(P_1^T) \\ \vdots \\ \log(P_H^T) \end{pmatrix} = \begin{pmatrix} (P_1^{\log})^T \\ \vdots \\ (P_H^{\log})^T \end{pmatrix}.$$

Теперь используем пункт 2 утверждения.

1. Найдем оценку $\hat{\Sigma}$ ковариационной матрицы, где $\hat{\Sigma}_{ij} = \mathbf{cov}(P_i^{\log}, P_j^{\log})$ для всех $i, j = 1, \dots, H$ и вектор средних $\bar{\mu} = (\mu_1, \dots, \mu_H)$, где μ_k — среднее по P_k^{\log} для всех $k = 1, \dots, H$.
2. Пусть $\gamma = 1 - \alpha$, где α — заданный уровень значимости, $c_\gamma = (T_H^2(G - 1))^{-1}(\gamma)$, где $T_H^2(G - 1)$ — функция распределения Хотеллинга со степенью свободы $(G - 1)$.
3. Вычисляем расстояние Махаланобиса между \bar{t}^{\log} и $\bar{\mu}$, если $\rho^2(\bar{t}^{\log}, \bar{\mu}) < c_\gamma$, то гипотеза не отвергается, иначе — отвергается. Таким образом доверительная область строится как доверительный эллипсоид с центром в $\bar{\mu}$.

В силу того, что с помощью расстояния Махаланобиса можно построить только аналог двустороннего критерия, а не одностороннего, будем задавать уровень значимости $\alpha = 0.1$.

Рассмотрим приведенные ниже результаты оценивания вероятности ошибки первого рода, жирным шрифтом выделены оценки, которые не попадают в доверительную область гипотезы для соответствующего уровня значимости.

Теоретический доверительный эллипсоид ($\alpha = 0.10$)						
Диап.	С.ч.AR(1)	С.век.AR(1)	$\sin, \cos(1/2L)$	$\sin, \cos(1/4L)$	$\sin, \cos(1/8L)$	$\sin(1/L)$
Бол.	0.159	0.127	0.133	0.224	0.266	0.095
Ср.	0.131	0.114	0.107	0.202	0.260	0.104
Узк.	0.161	0.116	0.110	0.183	0.205	0.105
Теоретический доверительный эллипсоид для минимумов в соседних парах ($\alpha = 0.10$)						
Диап.	С.ч.AR(1)	С.век.AR(1)	$\sin, \cos(1/2L)$	$\sin, \cos(1/4L)$	$\sin, \cos(1/8L)$	$\sin(1/L)$
Бол.	0.160	0.160	0.160	0.199	0.275	—
Ср.	0.131	0.143	0.165	0.176	0.259	—
Узк.	0.145	0.147	0.149	0.155	0.206	—
Теоретический доверительный эллипсоид для сумм соседних пар ($\alpha = 0.10$)						
Диап.	С.ч.AR(1)	С.век.AR(1)	$\sin, \cos(1/2L)$	$\sin, \cos(1/4L)$	$\sin, \cos(1/8L)$	$\sin(1/L)$
Бол.	0.163	0.131	0.125	0.238	0.284	—
Ср.	0.131	0.117	0.113	0.212	0.255	—
Узк.	0.152	0.112	0.110	0.186	0.213	—

Как видно из таблицы, большая часть оценок вероятности ошибки первого рода отличаются от заданного уровня значимости. Такое поведение обусловлено тем, что, прологарифмировав \bar{t} и \mathbf{P} , мы в какой-то мере приблизились к нормальности по каждому вектору, что показано на гистограммах выборок проекций на трех векторах для методов выбора векторов: собственные числа красного шума, синусы и косинусы с шагом по частоте $1/(2L)$, но не получили многомерной нормальности.

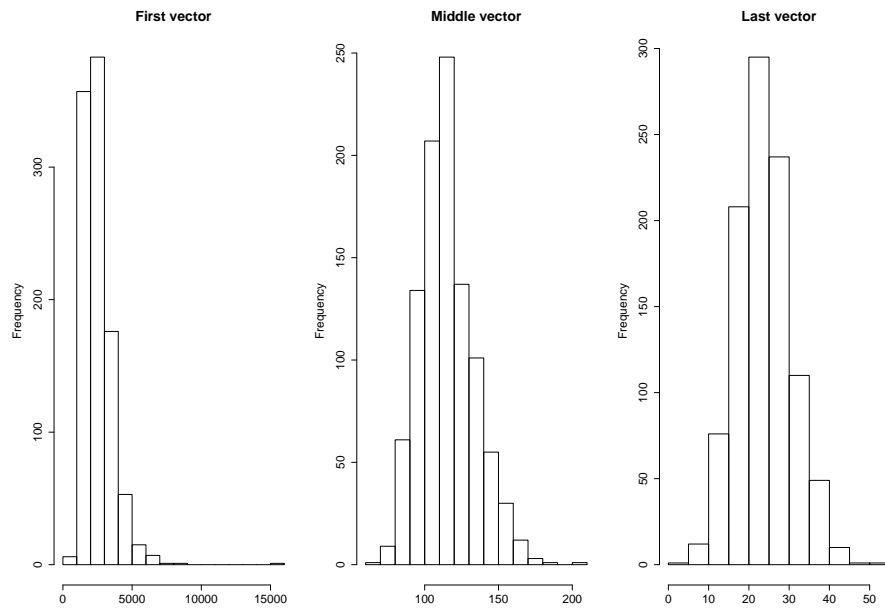


Рис. 2.3. Гистограммы проекций на три вектора до логарифмирования для метода: «С.ч.AR(1)».

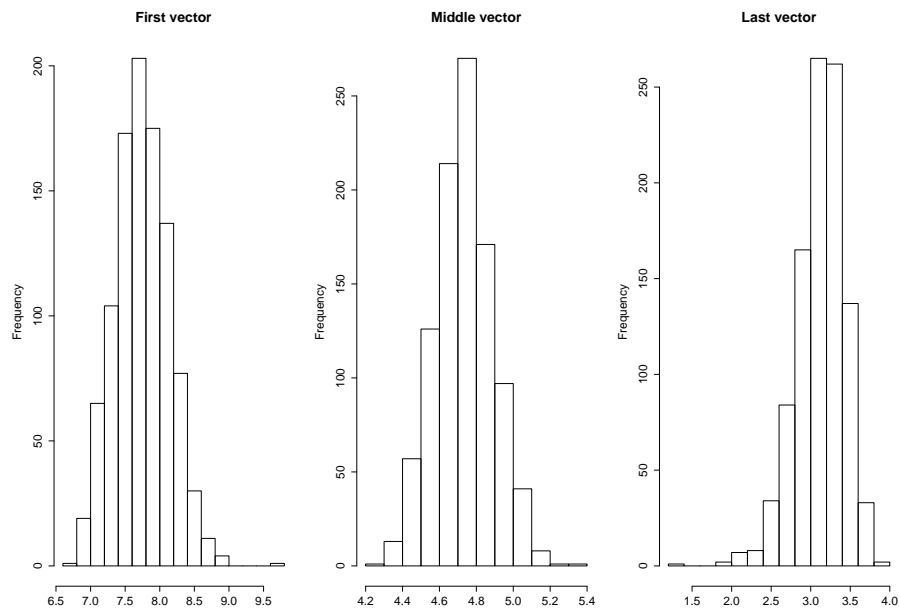


Рис. 2.4. Гистограммы проекций на три вектора после логарифмирования для метода: «С.ч.AR(1)».

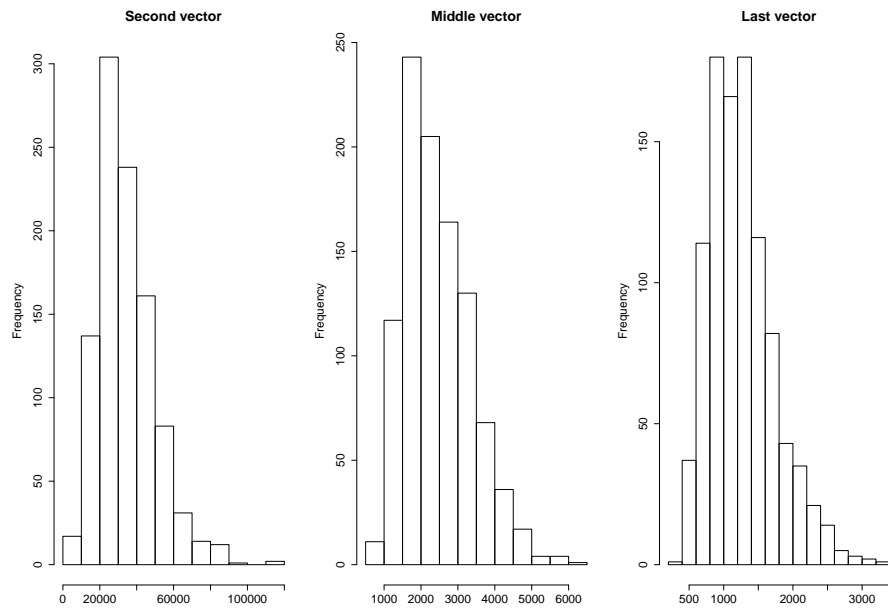


Рис. 2.5. Гистограммы проекций на три вектора до логарифмирования для метода: « $\sin, \cos(1/2L)$ ».

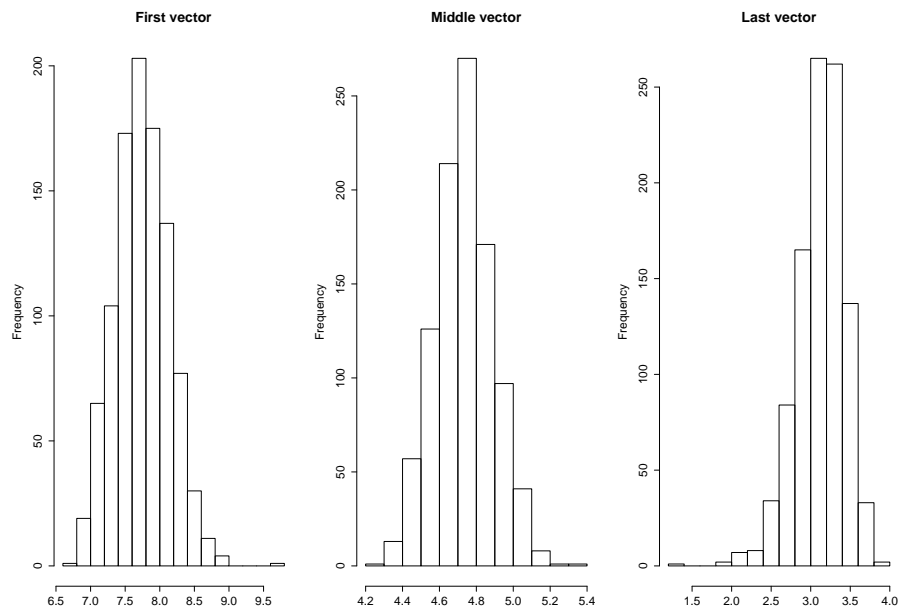


Рис. 2.6. Гистограммы проекций на три вектора после логарифмирования для метода: « $\sin, \cos(1/2L)$ ».

Как видно из этих рисунков, для больших частот распределения прологарифмированных значений статистики критерия видимо отличаются от нормального. Поэтому для больших диапазонов почти везде уровень значимости не равен ошибке первого рода.

При этом величина $\begin{pmatrix} (P_1^{\log})^T \\ \vdots \\ (P_H^{\log})^T \end{pmatrix}$ не имеет многомерной нормальности, что необходи-

мо по пункту 1 утверждения 1. Таким образом, при рассматриваемом методе построения доверительной области реальный уровень значимости в большинстве ситуаций не соответствует теоретическому, следовательно такой метод не имеет смысла использовать в дальнейшем. Можно пойти по пути исследования преобразования, которой дало бы многомерную нормальность, с тем, чтобы заменить им логарифмирование, но мы пойдем по пути построения другой доверительной области, сделав метод непараметрическим. Изменим данный метод, опираясь на выборки проекций.

Замечание 5. Вообще говоря, в статье [1] приведено утверждение, позволяющее аналитически рассчитать распределение проекции красного шума на вектор W , если он имеет синусоидальную форму с периодом T и L/T — целое. Пусть \mathbf{S}_R — автоковариационная матрица красного шума, \mathbf{A} — ее квадратный корень, тогда

$$p_i \sim K \left\| \mathbf{A}^T W \right\|^2 \frac{\chi^2(\nu)}{\nu},$$

где $\nu = 3N/L$ — степень свободы, $K = N - L + 1$.

Доверительные области, построенные по выборке

Будем строить доверительные области, используя полученные согласно методу Монте-Карло проекции промоделированных в соответствии с нулевой гипотезой рядов на вектора. Матрица проекций:

$$\mathbf{P} = \begin{pmatrix} P_1^T \\ \vdots \\ P_H^T \end{pmatrix} = [P^{(1)} : \dots : P^{(G)}] = \begin{pmatrix} p_{11} & \dots & p_{1G} \\ \vdots & \ddots & \vdots \\ p_{H1} & \dots & p_{HG} \end{pmatrix}.$$

Рассмотрим три вида доверительной области: доверительный полукуб, доверительный эллипсоид, доверительный куб. С помощью полукуба строится аналог одностороннего критерия, поэтому уровень значимости в этом случае равен 0.05. С помощью куба, как и с помощью эллипсоида, строится аналог двустороннего критерия, поэтому в обоих этих ситуациях $\alpha = 0.1$.

Для построения будем использовать вектор средних $\bar{\mu} = (\mu_1, \dots, \mu_H)$ и вектор стандартных отклонений $\bar{\sigma} = (\sigma_1, \dots, \sigma_H)$, где σ_k — стандартное отклонение по P_k для всех $k = 1, \dots, H$. Опишем конструирование этих доверительных областей.

Доверительный «полукуб»: находим вектор $\bar{\eta} = (\eta_1, \dots, \eta_G)$, где

$$\eta_j = \max_{1 \leq i \leq H} ((p_{ij} - \mu_i)/\sigma_i), \quad j = 1, \dots, G.$$

Затем вычисляем q : $(1 - \alpha)$ – выборочную квантиль по выборке $\bar{\eta}$. Тогда, если

$$\max_{1 \leq i \leq H} ((t_i - \mu_i)/\sigma_i) < q,$$

то нулевая гипотеза не отвергается, иначе — отвергается.

Доверительный эллипсоид: пусть $P^{(j)}$ — j -ый столбец матрицы \mathbf{P} . Находим вектор $\bar{\eta} = (\eta_1, \dots, \eta_G)$, где

$$\eta_j = \rho^2(P^j, \bar{\mu}), \quad j = 1, \dots, G.$$

Затем вычисляем q : $(1 - \alpha)$ – выборочную квантиль по выборке $\bar{\eta}$. Тогда, если

$$\rho^2(\bar{t}, \bar{\mu}) < q,$$

то нулевая гипотеза не отвергается, иначе — отвергается.

Доверительный «куб»: находим вектор $\bar{\eta} = (\eta_1, \dots, \eta_G)$, где

$$\eta_j = \max_{1 \leq i \leq H} (|p_{ij} - \mu_i|/\sigma_i), \quad j = 1, \dots, G.$$

Затем вычисляем q : $(1 - \alpha)$ – выборочную квантиль по выборке $\bar{\eta}$. Тогда, если

$$\max_{1 \leq i \leq H} (|t_i - \mu_i|/\sigma_i) < q,$$

то нулевая гипотеза не отвергается, иначе — отвергается.

При построении описанных выше доверительных областей будем предварительно логарифмировать вектор \bar{t} и матрицу \mathbf{P} . Это вызвано тем, что такие формы доверительных областей, как куб и полукуб подходят для симметричных распределений, а эллипсоид — для нормальных (см. утверждение 1), поэтому использование \bar{t}^{\log} и \mathbf{P}^{\log} вместо \bar{t} и \mathbf{P} является в данном случае более целесообразным. На рисунке 2.7 изображены рассматриваемые доверительные области в двумерном случае на фоне проекций на вектора.

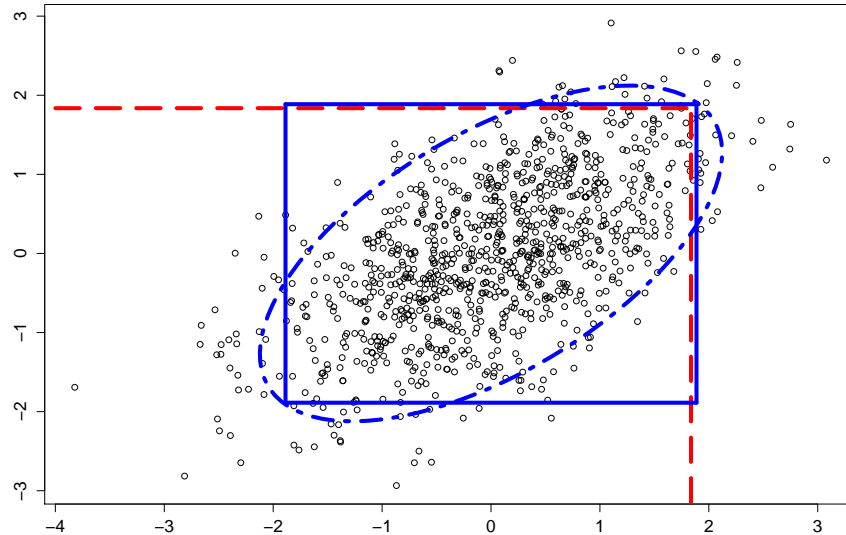


Рис. 2.7. Доверительные полукуб, эллипсоид и куб для двумерного случая.

Заметим, что доверительный эллипсоид, в отличие от доверительных куба и полукуба, не позволяет идентифицировать частоту вектора, на котором произошло попадание в критическую область.

Рассмотрим результаты оценивания вероятности ошибки первого рода для описанных выше методов построения доверительной области, жирным шрифтом выделены оценки, которые не попадают в доверительную область гипотезы для соответствующего уровня значимости.

Оценка вероятности ошибки первого рода						
Доверительный полукуб ($\alpha = 0.05$)						
Диап.	С.ч.AR(1)	С.век.AR(1)	$\sin, \cos(1/2L)$	$\sin, \cos(1/4L)$	$\sin, \cos(1/8L)$	$\sin(1/L)$
Бол.	0.047	0.044	0.050	0.042	0.047	0.044
Ср.	0.048	0.048	0.048	0.044	0.048	0.046
Узк.	0.054	0.054	0.050	0.047	0.057	0.042
Доверительный эллипсоид ($\alpha = 0.10$)						
Диап.	С.ч.AR(1)	С.век.AR(1)	$\sin, \cos(1/2L)$	$\sin, \cos(1/4L)$	$\sin, \cos(1/8L)$	$\sin(1/L)$
Бол.	0.132	0.158	0.158	0.185	0.296	0.115
Ср.	0.113	0.119	0.110	0.130	0.158	0.112
Узк.	0.118	0.124	0.115	0.117	0.117	0.107
Доверительный куб ($\alpha = 0.10$)						
Диап.	С.ч.AR(1)	С.век.AR(1)	$\sin, \cos(1/2L)$	$\sin, \cos(1/4L)$	$\sin, \cos(1/8L)$	$\sin(1/L)$
Бол.	0.103	0.110	0.103	0.104	0.104	0.095
Ср.	0.088	0.108	0.107	0.099	0.101	0.101
Узк.	0.100	0.098	0.102	0.100	0.107	0.097
Доверительный полукуб для минимумов в соседних парах ($\alpha = 0.05$)						
Диап.	С.ч.AR(1)	С.век.AR(1)	$\sin, \cos(1/2L)$	$\sin, \cos(1/4L)$	$\sin, \cos(1/8L)$	$\sin(1/L)$
Бол.	0.047	0.043	0.044	0.044	0.048	—
Ср.	0.049	0.050	0.046	0.045	0.043	—
Узк.	0.051	0.048	0.053	0.049	0.058	—
Доверительный эллипсоид для минимумов в соседних парах ($\alpha = 0.10$)						
Диап.	С.ч.AR(1)	С.век.AR(1)	$\sin, \cos(1/2L)$	$\sin, \cos(1/4L)$	$\sin, \cos(1/8L)$	$\sin(1/L)$
Бол.	0.135	0.144	0.144	0.210	0.325	—
Ср.	0.117	0.120	0.134	0.139	0.174	—
Узк.	0.120	0.115	0.131	0.115	0.133	—
Доверительный куб для минимумов в соседних парах ($\alpha = 0.10$)						
Диап.	С.ч.AR(1)	С.век.AR(1)	$\sin, \cos(1/2L)$	$\sin, \cos(1/4L)$	$\sin, \cos(1/8L)$	$\sin(1/L)$
Бол.	0.098	0.102	0.102	0.107	0.102	—
Ср.	0.090	0.102	0.096	0.098	0.103	—
Узк.	0.094	0.105	0.103	0.102	0.104	—

Доверительный полукуб для сумм соседних пар ($\alpha = 0.05$)						
Диап.	С.ч.AR(1)	С.век.AR(1)	$\sin, \cos(1/2L)$	$\sin, \cos(1/4L)$	$\sin, \cos(1/8L)$	$\sin(1/L)$
Бол.	0.047	0.045	0.051	0.044	0.045	—
Ср.	0.045	0.049	0.045	0.049	0.050	—
Узк.	0.050	0.051	0.049	0.050	0.056	—
Доверительный эллипсоид для сумм соседних пар ($\alpha = 0.10$)						
Диап.	С.ч.AR(1)	С.век.AR(1)	$\sin, \cos(1/2L)$	$\sin, \cos(1/4L)$	$\sin, \cos(1/8L)$	$\sin(1/L)$
Бол.	0.136	0.151	0.155	0.184	0.296	—
Ср.	0.117	0.117	0.118	0.131	0.157	—
Узк.	0.122	0.126	0.120	0.120	0.118	—
Доверительный куб для сумм соседних пар ($\alpha = 0.10$)						
Диап.	С.ч.AR(1)	С.век.AR(1)	$\sin, \cos(1/2L)$	$\sin, \cos(1/4L)$	$\sin, \cos(1/8L)$	$\sin(1/L)$
Бол.	0.102	0.103	0.104	0.100	0.105	—
Ср.	0.083	0.097	0.092	0.093	0.101	—
Узк.	0.096	0.095	0.099	0.097	0.102	—

Как видно из таблиц, приведенных выше, практически во всех ситуациях реальная вероятность ошибки первого рода лежит в соответствующей доверительной области. Этот результат говорит о корректности работы алгоритмов, поэтому для их дальнейшего сравнения по мощностям будем использовать рассматриваемые методы построения доверительной области.

Заметим, что при нахождении оценок было использовано G — число моделированных красного шума при построении проекций на вектора, равное 1000. Как видно из полученных результатов оценивания вероятности ошибки первого рода, этого числа хватило для корректности оценки в большинстве случаев. Однако, в случае векторов, являющихся синусами и косинусами с шагом по частоте $1/(8L)$, значение оценки сильно больше верхней границы доверительной области, соответствующей α . Такой эффект возникает из-за того, что G , равное 1000, является слишком маленьким числом для этого случая, но при $G \rightarrow \infty$ реальная вероятность ошибки первого рода попадет в доверительную область.

Замечание 6. Наблюдения, приведенные выше, показывают недостаток метода выбора векторов: синусы и косинусы с шагом по частоте $1/(8L)$, который заключается

в том, что ему необходимо большее число проекций, нежели для других методов, что делает его более трудоемким.

2.4. Сравнение методов с помощью оценок мощностей

Как было показано ранее, при построении доверительной области по выборке реальные вероятности ошибки первого рода близки к заданному уровню значимости, поэтому такие критерии можно считать построенными. Таким образом, имеет смысл для этих методов оценить мощности и с помощью них сравнить методы между собой. Нашей задачей является определить наиболее мощные методы против альтернативы $H_1 : S_N \neq 0$, где $S_N = (s_0, \dots, s_{N-1})$ — некоторый сигнал. Оценивание было проведено для $s_i = 0.5 \sin(2\pi i/5.5)$, $i = 0, \dots, N - 1$.

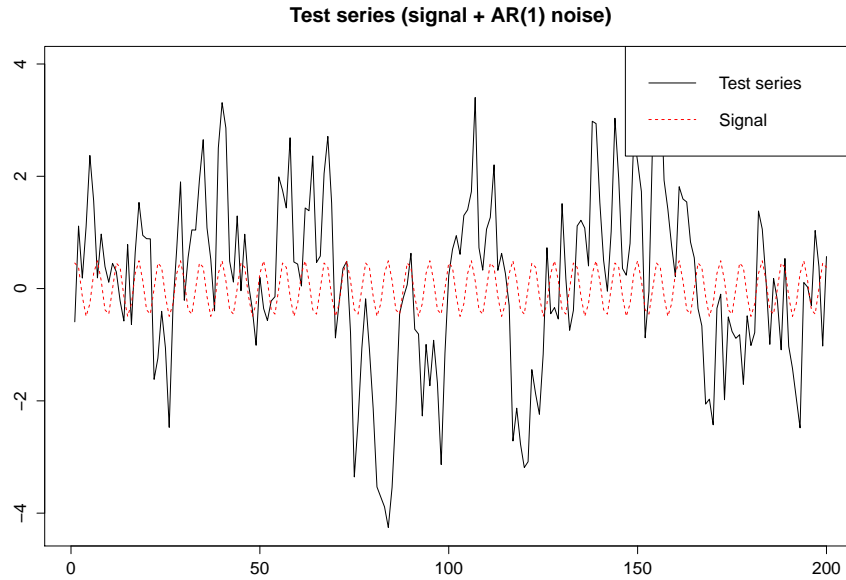


Рис. 2.8. Ряд F_N , для которого были построены оценки мощностей, $s_i = 0.5 \sin(2\pi i/5.5)$.

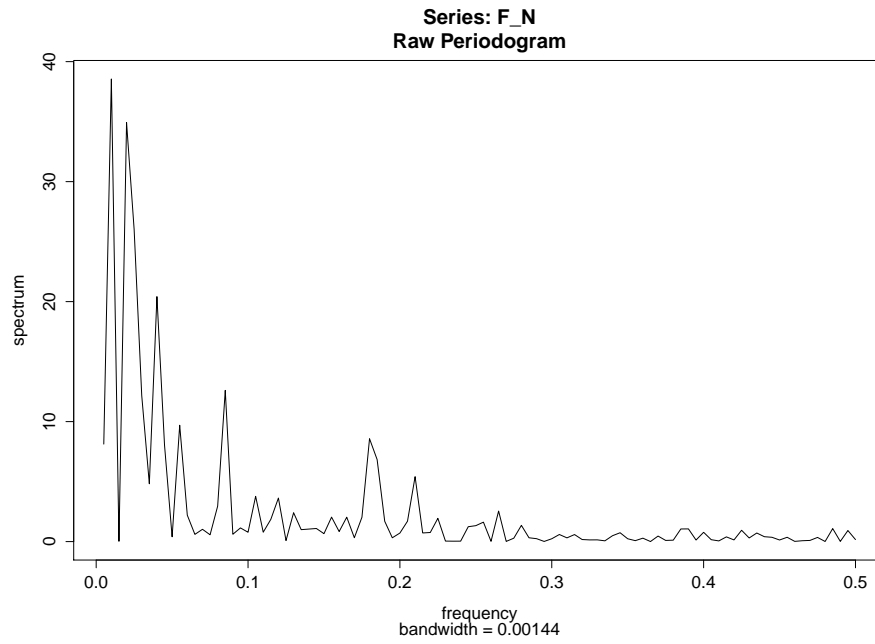


Рис. 2.9. Периодограмма ряда F_N , для которого были построены оценки мощностей.

2.4.1. Оценки мощностей

Оценка мощности, сигнал $s(x) = 0.5 \sin(2\pi x/5.5)$						
Доверительный полукуб ($\alpha = 0.05$)						
Диап.	С.ч.AR(1)	С.век.AR(1)	$\sin, \cos(1/2L)$	$\sin, \cos(1/4L)$	$\sin, \cos(1/8L)$	$\sin(1/L)$
Бол.	0.195	0.641	0.624	0.672	0.667	0.554
Ср.	0.214	0.719	0.709	0.744	0.753	0.635
Узк.	0.274	0.779	0.761	0.788	0.832	0.686
Доверительный эллипсоид ($\alpha = 0.10$)						
Диап.	С.ч.AR(1)	С.век.AR(1)	$\sin, \cos(1/2L)$	$\sin, \cos(1/4L)$	$\sin, \cos(1/8L)$	$\sin(1/L)$
Бол.	0.233	0.519	0.483	0.567	0.615	0.380
Ср.	0.225	0.586	0.577	0.596	0.546	0.484
Узк.	0.290	0.699	0.719	0.682	0.616	0.599
Доверительный куб ($\alpha = 0.10$)						
Диап.	С.ч.AR(1)	С.век.AR(1)	$\sin, \cos(1/2L)$	$\sin, \cos(1/4L)$	$\sin, \cos(1/8L)$	$\sin(1/L)$
Бол.	0.219	0.552	0.574	0.618	0.609	0.506
Ср.	0.249	0.651	0.656	0.691	0.690	0.592
Узк.	0.308	0.724	0.740	0.767	0.800	0.669

Доверительный полукуб для минимумов в соседних парах ($\alpha = 0.05$)						
Диап.	С.ч.AR(1)	С.век.AR(1)	$\sin, \cos(1/2L)$	$\sin, \cos(1/4L)$	$\sin, \cos(1/8L)$	$\sin(1/L)$
Бол.	0.209	0.637	0.691	0.679	0.651	—
Ср.	0.225	0.710	0.745	0.756	0.744	—
Узк.	0.271	0.769	0.790	0.799	0.819	—
Доверительный эллипсоид для минимумов в соседних парах ($\alpha = 0.10$)						
Диап.	С.ч.AR(1)	С.век.AR(1)	$\sin, \cos(1/2L)$	$\sin, \cos(1/4L)$	$\sin, \cos(1/8L)$	$\sin(1/L)$
Бол.	0.232	0.538	0.563	0.607	0.750	—
Ср.	0.221	0.646	0.679	0.617	0.697	—
Узк.	0.305	0.739	0.785	0.691	0.766	—
Доверительный куб для минимумов в соседних парах ($\alpha = 0.10$)						
Диап.	С.ч.AR(1)	С.век.AR(1)	$\sin, \cos(1/2L)$	$\sin, \cos(1/4L)$	$\sin, \cos(1/8L)$	$\sin(1/L)$
Бол.	0.219	0.598	0.672	0.656	0.602	—
Ср.	0.250	0.680	0.740	0.741	0.690	—
Узк.	0.301	0.765	0.790	0.805	0.804	—
Доверительный полукуб для сумм соседних пар ($\alpha = 0.05$)						
Диап.	С.ч.AR(1)	С.век.AR(1)	$\sin, \cos(1/2L)$	$\sin, \cos(1/4L)$	$\sin, \cos(1/8L)$	$\sin(1/L)$
Бол.	0.202	0.668	0.662	0.670	0.662	—
Ср.	0.222	0.743	0.733	0.743	0.750	—
Узк.	0.274	0.798	0.790	0.790	0.831	—
Доверительный эллипсоид для сумм соседних пар ($\alpha = 0.10$)						
Диап.	С.ч.AR(1)	С.век.AR(1)	$\sin, \cos(1/2L)$	$\sin, \cos(1/4L)$	$\sin, \cos(1/8L)$	$\sin(1/L)$
Бол.	0.232	0.518	0.500	0.572	0.616	—
Ср.	0.214	0.615	0.613	0.599	0.548	—
Узк.	0.293	0.730	0.745	0.687	0.624	—
Доверительный куб для сумм соседних пар ($\alpha = 0.10$)						
Диап.	С.ч.AR(1)	С.век.AR(1)	$\sin, \cos(1/2L)$	$\sin, \cos(1/4L)$	$\sin, \cos(1/8L)$	$\sin(1/L)$
Бол.	0.220	0.619	0.655	0.634	0.624	—
Ср.	0.243	0.697	0.721	0.709	0.697	—
Узк.	0.304	0.775	0.780	0.780	0.805	—

Отметим, что диапазоны номеров векторов: большой, средний и узкий брались с таким расчетом, чтобы номера векторов с частотами, близкими к частоте сигнала (т.е. в данном случае к $\omega = 1/5.5$), попадали в эти диапазоны. Для сравнения методов будем рассматривать результаты для среднего диапазона, т.е. половины всех номеров векторов.

Для оценок мощностей были построены доверительные интервалы по формуле (2.8), где вместо α подставляется соответствующая оценка мощности $\hat{\beta}$. Для каждого способа построения доверительной области значения оценок мощностей для методов выбора векторов «С.ч.AR(1)» и «sin(1/L)» являются наименьшими, кроме того доверительные интервалы для них не пересекаются с доверительными интервалами остальных методов и являются самыми маленькими по значениям. Поэтому при дальнейшем сравнении эти два метода рассматриваться не будут. У оставшихся четырех методов доверительные интервалы пересекаются, поэтому необходимо построить критерий для их сравнения.

2.4.2. Сравнение методов

Если рассматривать интерпретацию проверки критерия как испытания Бернулли, то для каждого метода выбора системы векторов при нахождении оценки мощности можно построить выборку вида (2.7). В силу того, что оценки мощностей строятся на одних реализациях ряда F_N для всех методов выбора векторов, эти выборки являются зависимыми. Поэтому для сравнения способов выбора системы векторов проведем двухвыборочные t -тесты в случае зависимых выборок.

В силу того, что для всех способов построения статистики критерия значения оценки мощности для доверительного полукуба значительно больше, чем для других методов построения доверительной области, а также соответствующие доверительные интервалы для $\hat{\beta}$ практически не пересекаются, будем рассматривать только доверительные полукубы. Соответствующие результаты t -тестов, подтверждающие утверждение о том, что метод построения доверительной области с помощью полукуба — лучший, а также результаты t -тестов для доверительных эллипсоидов и кубов приведены в приложении.

Рассмотрим p -значения, полученные в результате t -тестов, жирным шрифтом выделены p -значения, которые меньше 0.1.

Доверительный полукуб, диапазон средний				
Метод Оц.мощности	С.век.AR(1) $\hat{\beta} = 0.719$	$\sin, \cos(1/2L)$ $\hat{\beta} = 0.709$	$\sin, \cos(1/4L)$ $\hat{\beta} = 0.744$	$\sin, \cos(1/8L)$ $\hat{\beta} = 0.753$
С.век.AR(1) $\hat{\beta} = 0.719$	—	0.225	0.001	0.000
$\sin, \cos(1/2L)$ $\hat{\beta} = 0.709$	0.225	—	0.000	0.000
$\sin, \cos(1/4L)$ $\hat{\beta} = 0.744$	0.001	0.000	—	0.106
$\sin, \cos(1/8L)$ $\hat{\beta} = 0.753$	0.000	0.000	0.106	—

Как видно из приведенной выше таблицы, для доверительного полукуба методы выбора системы векторов образуют две группы, внутри которых они не отличаются значимо: первая группа — «С.век.AR(1)» и « $\sin, \cos(1/2L)$ », вторая — « $\sin, \cos(1/4L)$ » и « $\sin, \cos(1/8L)$ ». Между собой группы значимо отличаются, а т.к. во второй группе значения $\hat{\beta}$ больше, то алгоритмы « $\sin, \cos(1/4L)$ » и « $\sin, \cos(1/8L)$ » в данном случае лучше.

Если рассмотреть значения $\hat{\beta}$ для большого и узкого диапазонов, то видно, что найденная тенденция сохраняется, за исключением большого диапазона, где $\hat{\beta}$ у метода « $\sin, \cos(1/4L)$ » незначительно больше, чем у метода « $\sin, \cos(1/8L)$ ». Заметим, что при узком диапазоне значение $\hat{\beta}$ для метода « $\sin, \cos(1/8L)$ » сильно больше, чем для всех остальных.

Доверительный полукуб для минимумов в соседних парах, диапазон средний				
Метод Оц.мощности	С.век.AR(1) $\hat{\beta} = 0.710$	$\sin, \cos(1/2L)$ $\hat{\beta} = 0.745$	$\sin, \cos(1/4L)$ $\hat{\beta} = 0.756$	$\sin, \cos(1/8L)$ $\hat{\beta} = 0.744$
С.век.AR(1) $\hat{\beta} = 0.710$	—	0.000	0.000	0.000
$\sin, \cos(1/2L)$ $\hat{\beta} = 0.745$	0.000	—	0.116	0.896
$\sin, \cos(1/4L)$ $\hat{\beta} = 0.756$	0.000	0.116	—	0.045
$\sin, \cos(1/8L)$ $\hat{\beta} = 0.744$	0.000	0.896	0.045	—

Из результатов, приведенных выше, следует, что для доверительного полукуба при минимумах в соседних парах метод выбора системы векторов «С.век.AR(1)» значительно отличается от трех остальных и его значение $\hat{\beta}$ меньше, чем у других методов. Не отличаются значительно методы « $\sin, \cos(1/2L)$ » и « $\sin, \cos(1/8L)$ », также не отличаются значительно, но в меньшей степени (меньше p -значение), метод « $\sin, \cos(1/2L)$ » и метод « $\sin, \cos(1/4L)$ ». При этом методы « $\sin, \cos(1/4L)$ » и « $\sin, \cos(1/8L)$ » значительно различны, поэтому, в силу того, что у метода « $\sin, \cos(1/4L)$ » значение $\hat{\beta}$ больше, он является в данном случае лучшим.

При рассмотрении значений $\hat{\beta}$ для большого и узкого диапазонов видно, что найденная тенденция сохраняется в том смысле, что методы выбора системы векторов « $\sin, \cos(1/2L)$ », « $\sin, \cos(1/4L)$ » и « $\sin, \cos(1/8L)$ » для всех диапазонов являются лучшими, однако наибольшим значением $\hat{\beta}$ каждый раз обладает новый метод. Заметим, что для этих методов в каждом диапазоне значения $\hat{\beta}$ между собой отличаются не сильно.

Доверительный полукуб для сумм в соседних парах, диапазон средний				
Метод Оц.мощности	С.век.AR(1) $\hat{\beta} = 0.743$	sin,cos(1/2L) $\hat{\beta} = 0.733$	sin,cos(1/4L) $\hat{\beta} = 0.743$	sin,cos(1/8L) $\hat{\beta} = 0.750$
С.век.AR(1) $\hat{\beta} = 0.743$	—	0.123	1.000	0.275
sin,cos(1/2L) $\hat{\beta} = 0.733$	0.123	—	0.096	0.009
sin,cos(1/4L) $\hat{\beta} = 0.743$	1.000	0.096	—	0.209
sin,cos(1/8L) $\hat{\beta} = 0.750$	0.275	0.009	0.209	—

Исходя из результатов, описанных выше, для доверительного полукуба при суммах в соседних парах методы выбора системы векторов «sin,cos(1/4L)» и «sin,cos(1/8L)» не отличаются значимо, при этом метод «sin,cos(1/2L)» отличается значимо от них, а учитывая тот факт, что его значение $\hat{\beta}$ меньше, чем у «sin,cos(1/4L)» и «sin,cos(1/8L)», можно сказать, что эти методы лучше, чем «sin,cos(1/2L)». Метод «С.век.AR(1)» не отличается значимо от остальных трех методов, при этом он совпадает с «sin,cos(1/4L)», поэтому в данном случае лучшими являются методы «С.век.AR(1)», «sin,cos(1/4L)» и «sin,cos(1/8L)». Заметим, что метод «sin,cos(1/8L)» имеет наибольшее значение $\hat{\beta}$.

При рассмотрении значений $\hat{\beta}$ для большого и узкого диапазонов видно, что для большого диапазона все четыре метода имеют близкие значения $\hat{\beta}$ и наибольшим обладает метод «sin,cos(1/4L)». Для узкого диапазона метод «sin,cos(1/8L)» имеет величину $\hat{\beta}$ значительно большую, чем у трех остальных методов, значения $\hat{\beta}$ которых близки. В итоге, можно сказать, что для доверительного полукуба при суммах в соседних парах нельзя явно выделить лучшие методы, хотя в большинстве случаев наибольшим значением $\hat{\beta}$ обладает метод «sin,cos(1/8L)».

Теперь сравним для каждого способа выбора системы векторов различные методы построения статистики критерия.

Сравнение методов для «sin,cos(1/2L)», диапазон средний			
Метод Оц.мощности	обычный $\hat{\beta} = 0.709$	min в парах $\hat{\beta} = 0.745$	суммы в парах $\hat{\beta} = 0.733$
обычный $\hat{\beta} = 0.709$	—	0.000	0.002
min в парах $\hat{\beta} = 0.745$	0.000	—	0.028
суммы в парах $\hat{\beta} = 0.733$	0.002	0.028	—

Из приведенной выше таблицы видно, что все три метода значимо отличаются, а в силу того, что наибольшее значение $\hat{\beta}$ имеет метод выбора векторов «sin,cos(1/2L)» для доверительного полукуба в случае минимумов в соседних парах, он и является лучшим методом.

Сравнение методов для «sin,cos(1/4L)», диапазон средний			
Метод Оц.мощности	обычный $\hat{\beta} = 0.744$	min в парах $\hat{\beta} = 0.756$	суммы в парах $\hat{\beta} = 0.743$
обычный $\hat{\beta} = 0.744$	—	0.010	0.819
min в парах $\hat{\beta} = 0.756$	0.010	—	0.005
суммы в парах $\hat{\beta} = 0.743$	0.819	0.005	—

Как видно из приведенной выше таблицы, метод выбора векторов «sin,cos(1/4L)» для доверительного полукуба в случае минимумов в соседних парах обладает наибольшим значением $\hat{\beta}$ и значимо отличается от двух других методов, которые, в свою очередь, не отличаются значимо. Таким образом, можно сделать вывод, что наилучшим методом выбора системы векторов является метод «sin,cos(1/4L)» для доверительного полукуба в случае минимумов в соседних парах.

Сравнение методов для «sin,cos(1/8L)», диапазон средний			
Метод Оц.мощности	обычный $\hat{\beta} = 0.753$	min в парах $\hat{\beta} = 0.744$	суммы в парах $\hat{\beta} = 0.750$
обычный $\hat{\beta} = 0.753$	—	0.020	0.318
min в парах $\hat{\beta} = 0.744$	0.020	—	0.034
суммы в парах $\hat{\beta} = 0.750$	0.318	0.034	—

По приведенным выше данным видно, что не отличаются значимо методы выбора векторов «sin,cos(1/8L)» для полукуба в случае обычного метода и суммы в парах, при этом они обладают близкими значениями $\hat{\beta}$, которые больше, чем значение $\hat{\beta}$ в случае минимумов в парах. В силу того, что величина $\hat{\beta}$ немного больше для обычного метода, то его и назовем лучшим.

Сравнение методов для «С.век.AR(1)», диапазон средний			
Метод Оц.мощности	обычный $\hat{\beta} = 0.719$	min в парах $\hat{\beta} = 0.710$	суммы в парах $\hat{\beta} = 0.743$
обычный $\hat{\beta} = 0.719$	—	0.393	0.002
min в парах $\hat{\beta} = 0.710$	0.393	—	0.000
суммы в парах $\hat{\beta} = 0.743$	0.002	0.000	—

Как видно из приведенных выше данных, метод выбора векторов «С.век.AR(1)» для доверительного полукуба в случае сумм в соседних парах значимо отличается от двух других методов и имеет наибольшее значение $\hat{\beta}$. Таким образом он является лучшим.

Рассмотрим трудоемкости сравниваемых методов выбора системы векторов. Метод « $\sin, \cos(1/8L)$ » обладает большой трудоемкостью в силу замечания 6 и того, что в нем необходимо построить проекции на наибольшее число векторов: $4L$. Метод выбора векторов « $\sin, \cos(1/4L)$ » обладает средней трудоемкостью, так как проекции в нем строятся на среднее число векторов: $2L$. Отметим, что для корректности этого метода необходимо большее, чем для « $\sin, \cos(1/2L)$ » и «С.век.АР(1)» G — число моделирований красного шума при построении проекций на вектора. Метод « $\sin, \cos(1/2L)$ » и метод «С.век.АР(1)» обладают меньшей трудоемкостью, в силу того, что в них проекции строятся на меньшее число векторов: L . Однако, в методе «С.век.АР(1)» вектора зависят от параметра красного шума γ , поэтому при неизвестном γ на каждой итерации алгоритма строится его оценка и система векторов вычисляется заново.

2.4.3. Выводы

Исходя из приведенных выше сравнений методов построения статистики критерия, способов выбора системы векторов и методов построения доверительной области, можно сделать следующие выводы:

1. Доверительный куб демонстрирует результаты лучше, чем доверительный эллипсоид. Таким образом, для аналога двустороннего критерия лучшим является доверительный куб, а для аналога одностороннего критерия — доверительный полукуб.
2. Метод выбора векторов «С.ч.АР(1)» — самый маломощный, метод « $\sin(1/L)$ » лучше, чем «С.ч.АР(1)», но хуже остальных методов.
3. Методы выбора системы векторов «С.век.АР(1)», « $\sin, \cos(1/2L)$ », « $\sin, \cos(1/4L)$ » и « $\sin, \cos(1/8L)$ » сравнимы по мощности.
4. Результаты для способов построения статистики критерия \bar{t}^{sing} , \bar{t}^{min} и \bar{t}^{sum} также практически одинаковы.
5. Для метода выбора системы векторов « $\sin, \cos(1/8L)$ » необходимо большее, чем для остальных методов, число реализаций красного шума при использовании метода Монте-Карло.

Заметим, что в случае частоты сигнала, попадающей в частотную сетку, мощность метода выбора векторов « $\sin(1/L)$ » возрастает.

Отметим, что в ситуации рассматриваемых параметрах красного шума: $\gamma = 0.72$, $\delta = 1$, параметрах сигнала S_N : $T = 5.5$ — период, $\varphi = 0$ — фаза, $M = 1000$ — числе повторов метода Монте-Карло SSA и $G = 1000$, для которых производилось построение оценок мощности, можно сказать, что наилучшим является метод выбора системы векторов « $\sin, \cos(1/4L)$ » для доверительного полукуба при минимумах в соседних парах.

Замечание 7. В силу сделанных выводов, в дальнейшем будем рассматривать результаты только для доверительных полукубов. Соответствующие результаты для доверительных эллипсоидов и кубов приведены в приложении.

2.5. Сравнение методов для различных параметров сигнала

В предыдущих разделах мы рассматривали альтернативную гипотезу $H_1 : S_N \neq 0$, где $S_N = (s_0, \dots, s_{N-1})$, $s_i = 0.5 \sin(2\pi i/T + \varphi)$, $i = 0, \dots, N-1$ для параметров сигнала: период $T = 5.5$ и фаза $\varphi = 0$. Определим зависимость результатов работы метода от этих параметров. Ясно, что изменение T и φ может оказать влияние только на мощность, но не на вероятность ошибки первого рода.

Период и фаза колебания определяют попадание синуса S_N в частотную сетку, которая зависит от методов выбора системы векторов. Проведем сравнение оценок мощностей для различных T и φ для метода « $\sin, \cos(1/2L)$ ». В этом случае шаг по частотной сетке равен $1/(2L)$. Рассмотрим ситуации, когда период сигнала S_N попадает и не попадает в частотную сетку. В случае попадания в сетку рассмотрим ситуации, когда сигнал S_N является синусом и косинусом (при проекции сигнала на вектор W_k , который является синусом). Также будем рассматривать промежуточные случаи между синусом и косинусом. В итоге, оценим мощности для следующих параметров сигнала:

- $T = 5.5$, $\varphi = 0$ — не попадание в частотную сетку, синус.
- $T = 5.5$, $\varphi = \pi/4$ — не попадание в частотную сетку, промежуточная ситуация.
- $T = 5$, $\varphi = 0$ — попадание в частотную сетку, синус.
- $T = 5$, $\varphi = \pi/2$ — попадание в частотную сетку, косинус.

- $T = 5$, $\varphi = \pi/4$ — попадание в частотную сетку, промежуточная ситуация.

Ниже приведены оценки мощности, построенные при параметрах $M = 1000$ и $G = 1000$.

Оценка мощности для различных сигналов					
Доверительный полукуб ($\alpha = 0.05$)					
Диап.	$T = 5.5$ $\varphi = 0$	$T = 5.5$ $\varphi = \pi/4$	$T = 5$ $\varphi = 0$	$T = 5$ $\varphi = \pi/2$	$T = 5$ $\varphi = \pi/4$
Бол.	0.624	0.634	0.797	0.801	0.821
Ср.	0.709	0.715	0.845	0.847	0.872
Узк.	0.761	0.753	0.885	0.873	0.899
Доверительный полукуб для минимумов в соседних парах ($\alpha = 0.05$)					
Диап.	$T = 5.5$ $\varphi = 0$	$T = 5.5$ $\varphi = \pi/4$	$T = 5$ $\varphi = 0$	$T = 5$ $\varphi = \pi/2$	$T = 5$ $\varphi = \pi/4$
Бол.	0.691	0.703	0.541	0.535	0.547
Ср.	0.745	0.758	0.631	0.635	0.632
Узк.	0.790	0.797	0.695	0.695	0.706
Доверительный полукуб для сумм в соседних парах ($\alpha = 0.05$)					
Диап.	$T = 5.5$ $\varphi = 0$	$T = 5.5$ $\varphi = \pi/4$	$T = 5$ $\varphi = 0$	$T = 5$ $\varphi = \pi/2$	$T = 5$ $\varphi = \pi/4$
Бол.	0.662	0.687	0.715	0.726	0.722
Ср.	0.733	0.737	0.787	0.798	0.793
Узк.	0.790	0.785	0.827	0.824	0.844

Исходя из полученных результатов, можно сделать следующие выводы:

- В силу того, что у метода построения статистики критерия с помощью сумм в соседних парах наименьшие значения мощности больше, чем наименьшие значения для других методов, можно заключить, что этот метод является наиболее устойчивым к изменению периода сигнала.
- В случае непопадания сигнала в частотную сетку, т.е. при $T = 5.5$, лучшие оценки мощности получаются для минимумов в соседних парах.

- В случае попадания сигнала в частотную сетку, т.е. при $T = 5$, лучшие оценки мощности получаются для обычного метода построения статистики критерия.
- Траекторная матрица исходного ряда \mathbf{X} содержит в качестве столбцов вектора с различными фазами (вектора вложения), поэтому изменение фазы φ не оказывает влияния на мощность, что подтверждается полученными результатами.

2.6. Сравнение методов для различных параметров красного шума

Ранее мы получили результаты работы метода Монте-Карло SSA для параметров красного шума $\gamma = 0.72$, $\delta = 1$ и периода сигнала $S_N T = 5.5$. Рассмотрим влияние изменения параметра γ на значения мощности, а также влияние изменения периода T . Для этого приведем оценки мощности для $\gamma = 0.1$, $\gamma = 0.4$, $\gamma = 0.72$ и периодов $T = 5.5$ и $T = 10$. В силу выводов, сделанных в разделе 2.4.3, будем рассматривать только методы выбора векторов «С.век.AR(1)», «sin,cos(1/2L)», «sin,cos(1/4L)» и «sin,cos(1/8L)». Результаты приведены для средних диапазонов при параметрах $M = 1000$ и $G = 1000$.

Оценка мощности, сигнал $\mu(x) = 0.5 \sin(2\pi x/5.5)$				
Доверительный полукуб ($\alpha = 0.05$)				
γ	С.век.AR(1)	sin,cos(1/2L)	sin,cos(1/4L)	sin,cos(1/8L)
0.1	0.794	0.717	0.762	0.774
0.4	0.709	0.629	0.680	0.691
0.72	0.719	0.709	0.744	0.753
Оценка мощности, сигнал $\mu(x) = 0.5 \sin(2\pi x/10)$				
Доверительный полукуб ($\alpha = 0.05$)				
γ	С.век.AR(1)	sin,cos(1/2L)	sin,cos(1/4L)	sin,cos(1/8L)
0.1	0.680	0.703	0.684	0.688
0.4	0.350	0.399	0.372	0.371
0.72	0.219	0.227	0.214	0.219

Оценка мощности, сигнал $\mu(x) = 0.5 \sin(2\pi x/5.5)$				
Доверительный полукуб для минимумов в соседних парах ($\alpha = 0.05$)				
γ	С.век. AR(1)	$\sin, \cos(1/2L)$	$\sin, \cos(1/4L)$	$\sin, \cos(1/8L)$
0.1	0.547	0.761	0.757	0.756
0.4	0.537	0.684	0.688	0.676
0.72	0.710	0.745	0.756	0.744
Оценка мощности, сигнал $\mu(x) = 0.5 \sin(2\pi x/10)$				
Доверительный полукуб для минимумов в соседних парах ($\alpha = 0.05$)				
γ	С.век. AR(1)	$\sin, \cos(1/2L)$	$\sin, \cos(1/4L)$	$\sin, \cos(1/8L)$
0.1	0.641	0.472	0.664	0.689
0.4	0.391	0.228	0.355	0.371
0.72	0.214	0.145	0.194	0.206
Оценка мощности, сигнал $\mu(x) = 0.5 \sin(2\pi x/5.5)$				
Доверительный полукуб для сумм в соседних парах ($\alpha = 0.05$)				
γ	С.век. AR(1)	$\sin, \cos(1/2L)$	$\sin, \cos(1/4L)$	$\sin, \cos(1/8L)$
0.1	0.718	0.752	0.758	0.769
0.4	0.653	0.677	0.677	0.687
0.72	0.743	0.733	0.743	0.750
Оценка мощности, сигнал $\mu(x) = 0.5 \sin(2\pi x/10)$				
Доверительный полукуб для сумм в соседних парах ($\alpha = 0.05$)				
γ	С.век. AR(1)	$\sin, \cos(1/2L)$	$\sin, \cos(1/4L)$	$\sin, \cos(1/8L)$
0.1	0.673	0.640	0.682	0.684
0.4	0.384	0.331	0.370	0.374
0.72	0.215	0.179	0.213	0.217

Как видно из приведенных выше таблиц, при росте периода сигнала S_N значения мощности уменьшаются. В случае, когда период сигнала равен 5.5, наибольшая мощность получается при $\gamma = 0.1$, среднее значение мощности — при $\gamma = 0.72$, наименьшее — при $\gamma = 0.4$. Однако, в случае $T = 5$ самое большое значение мощности для $\gamma = 0.1$, среднее — для $\gamma = 0.4$, самое маленькое — для $\gamma = 0.72$. Отметим, что критерий, построенный с помощью доверительного эллипсоида, является устойчивым по отношению к изменению параметра красного шума γ .

Такая упорядоченность мощностей в зависимости от параметра γ объясняется влиянием красного шума в исходном ряде F_N , который согласно модели имеет структуру $F_N = S_N + R_N$, где S_N — сигнал, R_N — красный шум. Чем больше вклад процесса авторегрессии первого порядка, тем меньше мощность рассматриваемого критерия. Вклад красного шума определяется значением его спектральной плотности. На рисунке 2.10 изображены графики спектральных плотностей в зависимости от частоты для $\gamma = 0.1$, $\gamma = 0.4$ и $\gamma = 0.72$.

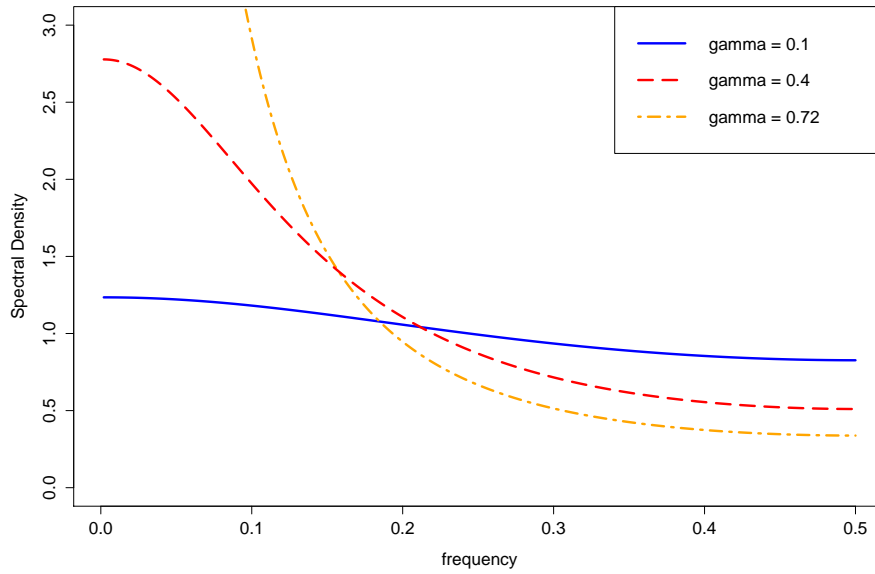


Рис. 2.10. Графики спектральных плотностей AR(1) процесса для $\gamma = 0.1$, $\gamma = 0.4$, $\gamma = 0.72$.

Как видно из рисунка 2.10, для периода сигнала $T = 5.5$, которому соответствует частота $\omega = 0.18$, наибольшее значение спектральной плотности соответствует $\gamma = 0.4$, что показывает наибольший вклад красного шума в этой ситуации и влечет наименьшее значение мощности. Аналогично для $\gamma = 0.72$ значение плотности — среднее, и мощность — средняя, а для $\gamma = 0.1$ плотность — наименьшая, мощность — наибольшая. В случае периода $T = 10$, которому соответствует частота $\omega = 0.1$, для γ , расположенных по убыванию значений спектральной плотности, мощности расположены в возрастающем порядке: $\gamma = 0.72$, $\gamma = 0.4$, $\gamma = 0.1$.

2.7. Монте-Карло SSA с оценкой параметров красного шума

Ранее мы рассмотрели статистические свойства метода Монте-Карло SSA при известных параметрах красного шума. В реальных задачах они неизвестны, поэтому имеет смысл рассмотреть влияние оценок параметров AR(1) процесса на реальную вероятность ошибки первого рода и на мощность. Для оценивание параметров красного шума γ и δ будем использовать второй алгоритм из раздела 1.3.1.

Рассмотрим результаты оценивания вероятности ошибки первого рода, жирным шрифтом выделены оценки, которые не попадают в доверительную область гипотезы для соответствующего уровня значимости.

Оценка вероятности ошибки первого рода при неизвестных параметрах шума						
Доверительный полукуб ($\alpha = 0.05$)						
Диап.	С.ч.AR(1)	С.век.AR(1)	$\sin, \cos(1/2L)$	$\sin, \cos(1/4L)$	$\sin, \cos(1/8L)$	$\sin(1/L)$
Бол.	0.015	0.023	0.027	0.025	0.026	0.015
Ср.	0.021	0.020	0.026	0.031	0.026	0.024
Узк.	0.029	0.024	0.037	0.038	0.043	0.038
Доверительный полукуб для минимумов в соседних парах ($\alpha = 0.05$)						
Диап.	С.ч.AR(1)	С.век.AR(1)	$\sin, \cos(1/2L)$	$\sin, \cos(1/4L)$	$\sin, \cos(1/8L)$	$\sin(1/L)$
Бол.	0.009	0.024	0.022	0.023	0.025	—
Ср.	0.015	0.017	0.032	0.027	0.024	—
Узк.	0.023	0.024	0.031	0.039	0.047	—
Доверительный полукуб для сумм соседних пар ($\alpha = 0.05$)						
Диап.	С.ч.AR(1)	С.век.AR(1)	$\sin, \cos(1/2L)$	$\sin, \cos(1/4L)$	$\sin, \cos(1/8L)$	$\sin(1/L)$
Бол.	0.010	0.020	0.024	0.023	0.026	—
Ср.	0.015	0.022	0.027	0.027	0.024	—
Узк.	0.029	0.020	0.032	0.041	0.046	—

Как видно из приведенных выше данных, реальные вероятности ошибки первого рода уменьшились по сравнению со случаем известных параметров красного шума. Поэтому критерием можно пользоваться, но мощность критерия уменьшится. Самое сильное уменьшение произошло для методов выбора системы векторов «С.ч.AR(1)» и

«С.век.AR(1)». Такой эффект объясняется тем, что эти методы зависят от параметров красного шума. Отметим, что наименьшие изменения произошли для доверительных эллипсоидов, в случае которых сохраняются хорошие значения вероятности ошибки первого рода. Для доверительных полукубов реальная вероятность ошибки первого рода уменьшилась в среднем в 2.4 раза, для доверительных эллипсоидов — в 1.2 раза, для доверительных кубов — в 1.6 раза.

Рассмотрим оценки мощности.

Оценка мощности при неизвестных параметрах шума, сигнал $\mu(x) = 0.5 \sin(2\pi x/5.5)$						
Доверительный полукуб ($\alpha = 0.05$)						
Диап.	С.ч.AR(1)	С.век.AR(1)	$\sin, \cos(1/2L)$	$\sin, \cos(1/4L)$	$\sin, \cos(1/8L)$	$\sin(1/L)$
Бол.	0.049	0.453	0.465	0.548	0.549	0.379
Ср.	0.067	0.560	0.589	0.622	0.636	0.479
Узк.	0.122	0.642	0.640	0.677	0.733	0.546
Доверительный полукуб для минимумов в соседних парах ($\alpha = 0.05$)						
Диап.	С.ч.AR(1)	С.век.AR(1)	$\sin, \cos(1/2L)$	$\sin, \cos(1/4L)$	$\sin, \cos(1/8L)$	$\sin(1/L)$
Бол.	0.047	0.436	0.556	0.557	0.522	—
Ср.	0.067	0.532	0.628	0.636	0.623	—
Узк.	0.115	0.616	0.697	0.702	0.716	—
Доверительный полукуб для сумм соседних пар ($\alpha = 0.05$)						
Диап.	С.ч.AR(1)	С.век.AR(1)	$\sin, \cos(1/2L)$	$\sin, \cos(1/4L)$	$\sin, \cos(1/8L)$	$\sin(1/L)$
Бол.	0.044	0.490	0.526	0.544	0.544	—
Ср.	0.065	0.592	0.618	0.624	0.630	—
Узк.	0.112	0.679	0.667	0.689	0.726	—

Из приведенных выше таблиц видно, что оценки мощности уменьшились, как и вероятности ошибки первого рода, сохранив при этом такое же поведение, как при известных параметрах красного шума. Как и в случае оценок вероятности ошибки первого рода, самое сильное уменьшение произошло для способов выбора векторов «С.ч.AR(1)» и «С.век.AR(1)», и наименьшие изменения произошли для доверительных эллипсоидов. Таким образом, можно сказать, что эллипсоиды наиболее устойчивые. В силу того, что при оценивании мощности оценки параметров красного шума $\hat{\gamma}$ и $\hat{\delta}$ строились по рядам,

которые являлись суммой синуса и красного шума, в отличии от оценок вероятности ошибки первого рода, где $\hat{\gamma}$ и $\hat{\delta}$ строились по чистому красному шуму, мощности уменьшились в среднем в другое число раз, по сравнению с реальными вероятностями ошибки первого рода: для доверительных полукубов — в 1.9 раз, для доверительных эллипсоидов — в 1.1 раз, для доверительных кубов — в 1.4 раза. Отметим, что по сравнению со случаем известных параметров красного шума, в ситуации неизвестных параметров критерий, построенный с помощью эллипсоида, значительно приблизился по мощности к критерию, построенному с помощью куба.

В разделе 1.3.3 было получено, что при периоде сигнала 5.5 оценка для параметра красного шума $\gamma = 0.4$ лучше в случае наличия сигнала в ряде, чем в случае его отсутствия. Такой эффект оказывает влияние на мощность, и в результате при $\gamma = 0.4$ оценка мощности уменьшается в значительно меньшее число раз, чем оценка вероятности ошибки первого рода.

Заключение

Таким образом, в работе были получены следующие результаты.

- Формализована задача из [1], построены корректные модификации Монте-Карло SSA для ее решения.
- Проведено сравнение методов по мощности, в результате которого было выявлено, какие варианты многомерных критериев являются более мощными. Во-первых, это использование «полукуба» в качестве доверительной области. Во-вторых, использование равномерной решетки по частоте для набора векторов, на которые строятся проекции, оказалось не хуже, чем набор собственных векторов красного шума, рассматриваемый в [1].
- Показано, что мощность критерия зависит от попадания частоты сигнала в решетку частот набора векторов. В качестве критерия, имеющего наибольшую минимальную мощность против синусов с разными частотами, является предложенный в данной работе критерий, построенный на суммах проекция на соседние вектора.
- Показано, что оценивание параметров шума по ряду, в целом, уменьшает мощность критерия, но не слишком сильно.
- Продemonстрировано, что свойства критерия не зависят как от фазы выбираемых в качестве проекций векторов, так и от фазы сигнала.

Полученные результаты, во-первых, дают корректно построенные критерии типа Монте-Карло SSA, во-вторых, позволяют выбрать из них более мощные. Тем самым, в частности, появляется возможность сравнить эти критерии с другими, чего не было раньше.

Список литературы

1. Allen M. R., Smith L. A. Monte carlo ssa: detecting irregular oscillations in the presence of coloured noise // Journal of Climate. — 1996. — no. 9. — P. 3373–3404.
2. Голяндина Н. Э. Метод «Гусеница»-SSA: анализ временных рядов: Учеб. пособие. — СПб : BBM, 2004. — 76 с.
3. Gardner G., Harvey A. C., Phillips G. D. A. Algorithm AS154. an algorithm for exact maximum likelihood estimation of autoregressive-moving average models by means of Kalman filtering // Applied Statistics. — 1980. — no. 29. — P. 311–322.
4. Vautard R., Yiou P., Ghil M. Singular-spectrum analysis: A toolkit for short, noisy chaotic signals // Physica D. — 1992. — no. 58. — P. 95–126.
5. Wilson E. B. Probable inference, the law of succession, and statistical inference // Journal of the American Statistical Association. — 1927. — no. 22. — P. 209–212.

Приложение А

Описание CD

К дипломной работе прилагается CD, в котором содержатся следующие материалы.

1. Файл «methods.R» содержит методы, написанные на языке программирования «R», для генерации ряда, построения статистики критерия, проекций на вектора, описанные в разделе 2.2, оценивания параметров красного шума 1.3.1, генерации системы векторов.
2. Файл «Monte-Carlo-SSA-Simple-Ellipsoid.R» содержит программу, написанную на языке программирования «R», вычисляющую оценки вероятности ошибки первого рода или мощности для простого метода построения доверительной области и теоретического эллипсоида.
3. Файл «Monte-Carlo-SSA-SampleQuantile.R» содержит программу, написанную на языке программирования «R», вычисляющую оценки вероятности ошибки первого рода или мощности для доверительных областей, построенных по выборке.
4. Файл «CIs.R» содержит методы, написанные на языке программирования «R», для построения доверительных областей, построенных по выборке: «полукуб», эллипсоид, «куб».
5. Файл «PrintResults.R» содержит вспомогательные методы, написанные на языке программирования «R», предназначенные для вывода информации.

Приложение Б

Результаты для эллипсоидов и кубов

Б.1. Сравнение методов выбора векторов для доверительных эллипсоидов и кубов

Доверительный эллипсоид, диапазон средний				
Метод Оц.мощности	С.век. AR(1) $\hat{\beta} = 0.586$	$\sin, \cos(1/2L)$ $\hat{\beta} = 0.577$	$\sin, \cos(1/4L)$ $\hat{\beta} = 0.596$	$\sin, \cos(1/8L)$ $\hat{\beta} = 0.546$
С.век. AR(1) $\hat{\beta} = 0.586$	—	0.439	0.443	0.004
$\sin, \cos(1/2L)$ $\hat{\beta} = 0.577$	0.439	—	0.146	0.026
$\sin, \cos(1/4L)$ $\hat{\beta} = 0.596$	0.443	0.146	—	0.000
$\sin, \cos(1/8L)$ $\hat{\beta} = 0.546$	0.004	0.026	0.000	—

Доверительный куб, диапазон средний				
Метод Оц.мощности	С.век. AR(1) $\hat{\beta} = 0.651$	$\sin, \cos(1/2L)$ $\hat{\beta} = 0.656$	$\sin, \cos(1/4L)$ $\hat{\beta} = 0.691$	$\sin, \cos(1/8L)$ $\hat{\beta} = 0.690$
С.век. AR(1) $\hat{\beta} = 0.651$	—	0.579	0.000	0.000
$\sin, \cos(1/2L)$ $\hat{\beta} = 0.656$	0.579	—	0.000	0.000
$\sin, \cos(1/4L)$ $\hat{\beta} = 0.691$	0.000	0.000	—	0.884
$\sin, \cos(1/8L)$ $\hat{\beta} = 0.690$	0.000	0.000	0.884	—

Доверительный эллипсоид для минимумов в соседних парах, диапазон средний				
Метод Оц.мощности	С.век.AR(1) $\hat{\beta} = 0.646$	$\sin, \cos(1/2L)$ $\hat{\beta} = 0.679$	$\sin, \cos(1/4L)$ $\hat{\beta} = 0.617$	$\sin, \cos(1/8L)$ $\hat{\beta} = 0.697$
С.век.AR(1) $\hat{\beta} = 0.646$	—	0.005	0.028	0.000
$\sin, \cos(1/2L)$ $\hat{\beta} = 0.679$	0.005	—	0.000	0.144
$\sin, \cos(1/4L)$ $\hat{\beta} = 0.617$	0.028	0.000	—	0.000
$\sin, \cos(1/8L)$ $\hat{\beta} = 0.697$	0.000	0.144	0.000	—

Доверительный куб для минимумов в соседних парах, диапазон средний				
Метод Оц.мощности	С.век.AR(1) $\hat{\beta} = 0.680$	$\sin, \cos(1/2L)$ $\hat{\beta} = 0.740$	$\sin, \cos(1/4L)$ $\hat{\beta} = 0.741$	$\sin, \cos(1/8L)$ $\hat{\beta} = 0.690$
С.век.AR(1) $\hat{\beta} = 0.680$	—	0.000	0.000	0.297
$\sin, \cos(1/2L)$ $\hat{\beta} = 0.740$	0.000	—	0.901	0.000
$\sin, \cos(1/4L)$ $\hat{\beta} = 0.741$	0.000	0.901	—	0.000
$\sin, \cos(1/8L)$ $\hat{\beta} = 0.690$	0.297	0.000	0.000	—

Доверительный эллипсоид для сумм в соседних парах, диапазон средний				
Метод Оц.мощности	С.век.AR(1) $\hat{\beta} = 0.615$	$\sin, \cos(1/2L)$ $\hat{\beta} = 0.613$	$\sin, \cos(1/4L)$ $\hat{\beta} = 0.599$	$\sin, \cos(1/8L)$ $\hat{\beta} = 0.548$
С.век.AR(1) $\hat{\beta} = 0.615$	—	0.861	0.220	0.000
$\sin, \cos(1/2L)$ $\hat{\beta} = 0.613$	0.861	—	0.266	0.000
$\sin, \cos(1/4L)$ $\hat{\beta} = 0.599$	0.220	0.266	—	0.000
$\sin, \cos(1/8L)$ $\hat{\beta} = 0.548$	0.000	0.000	0.000	—

Доверительный куб для сумм в соседних парах, диапазон средний				
Метод Оц.мощности	С.век.AR(1) $\hat{\beta} = 0.697$	$\sin, \cos(1/2L)$ $\hat{\beta} = 0.721$	$\sin, \cos(1/4L)$ $\hat{\beta} = 0.709$	$\sin, \cos(1/8L)$ $\hat{\beta} = 0.697$
С.век.AR(1) $\hat{\beta} = 0.697$	—	0.000	0.070	1.000
$\sin, \cos(1/2L)$ $\hat{\beta} = 0.721$	0.000	—	0.077	0.002
$\sin, \cos(1/4L)$ $\hat{\beta} = 0.709$	0.070	0.077	—	0.090
$\sin, \cos(1/8L)$ $\hat{\beta} = 0.697$	1.000	0.002	0.090	—

Б.2. Сравнение способов построения доверительных областей

Обычный метод построения критерия, диапазон средний			
Метод	$\sin, \cos(1/4L)$ (полукуб)	$\sin, \cos(1/4L)$ (эллипсоид)	$\sin, \cos(1/4L)$ (куб)
Оц.мощности	$\hat{\beta} = 0.744$	$\hat{\beta} = 0.596$	$\hat{\beta} = 0.691$
$\sin, \cos(1/4L)$ (полукуб) $\hat{\beta} = 0.744$	—	0.000	0.000
$\sin, \cos(1/4L)$ (эллипсоид) $\hat{\beta} = 0.596$	0.000	—	0.000
$\sin, \cos(1/4L)$ (куб) $\hat{\beta} = 0.691$	0.000	0.000	—

Минимумы в соседних парах, диапазон средний			
Метод	$\sin, \cos(1/4L)$ (полукуб)	$\sin, \cos(1/8L)$ (эллипсоид)	$\sin, \cos(1/4L)$ (куб)
Оц.мощности	$\hat{\beta} = 0.756$	$\hat{\beta} = 0.697$	$\hat{\beta} = 0.741$
$\sin, \cos(1/4L)$ (полукуб) $\hat{\beta} = 0.756$	—	0.000	0.059
$\sin, \cos(1/8L)$ (эллипсоид) $\hat{\beta} = 0.697$	0.000	—	0.003
$\sin, \cos(1/4L)$ (куб) $\hat{\beta} = 0.741$	0.059	0.003	—

Суммы в соседних парах, диапазон средний			
Метод	$\sin, \cos(1/4L)$ (полукуб) $\hat{\beta} = 0.743$	С.век. AR(1) (эллипсоид) $\hat{\beta} = 0.615$	$\sin, \cos(1/2L)$ (куб) $\hat{\beta} = 0.721$
Оц. мощности			
$\sin, \cos(1/4L)$ (полукуб) $\hat{\beta} = 0.743$	—	0.000	0.007
С.век. AR(1) (эллипсоид) $\hat{\beta} = 0.615$	0.000	—	0.000
$\sin, \cos(1/2L)$ (куб) $\hat{\beta} = 0.721$	0.007	0.000	—

Б.3. Сравнение методов для различных параметров сигнала

Оценка мощности для различных сигналов					
Доверительный эллипсоид ($\alpha = 0.10$)					
Диап.	$T = 5.5$ $\varphi = 0$	$T = 5.5$ $\varphi = \pi/4$	$T = 5$ $\varphi = 0$	$T = 5$ $\varphi = \pi/2$	$T = 5$ $\varphi = \pi/4$
Бол.	0.483	0.488	0.626	0.611	0.634
Ср.	0.577	0.584	0.668	0.664	0.680
Узк.	0.719	0.728	0.772	0.786	0.778
Доверительный куб ($\alpha = 0.10$)					
Диап.	$T = 5.5$ $\varphi = 0$	$T = 5.5$ $\varphi = \pi/4$	$T = 5$ $\varphi = 0$	$T = 5$ $\varphi = \pi/2$	$T = 5$ $\varphi = \pi/4$
Бол.	0.574	0.576	0.760	0.769	0.784
Ср.	0.656	0.666	0.817	0.819	0.844
Узк.	0.740	0.744	0.870	0.864	0.884
Доверительный эллипсоид для минимумов в соседних парах ($\alpha = 0.10$)					
Диап.	$T = 5.5$ $\varphi = 0$	$T = 5.5$ $\varphi = \pi/4$	$T = 5$ $\varphi = 0$	$T = 5$ $\varphi = \pi/2$	$T = 5$ $\varphi = \pi/4$
Бол.	0.563	0.585	0.468	0.481	0.471
Ср.	0.679	0.692	0.579	0.584	0.604
Узк.	0.785	0.784	0.699	0.689	0.698
Доверительный куб для минимумов в соседних парах ($\alpha = 0.10$)					
Диап.	$T = 5.5$ $\varphi = 0$	$T = 5.5$ $\varphi = \pi/4$	$T = 5$ $\varphi = 0$	$T = 5$ $\varphi = \pi/2$	$T = 5$ $\varphi = \pi/4$
Бол.	0.672	0.670	0.496	0.504	0.513
Ср.	0.740	0.741	0.598	0.604	0.617
Узк.	0.790	0.795	0.699	0.694	0.706

Доверительный эллипсоид для сумм в соседних парах ($\alpha = 0.10$)					
Диап.	$T = 5.5$ $\varphi = 0$	$T = 5.5$ $\varphi = \pi/4$	$T = 5$ $\varphi = 0$	$T = 5$ $\varphi = \pi/2$	$T = 5$ $\varphi = \pi/4$
Бол.	0.500	0.528	0.618	0.612	0.630
Ср.	0.613	0.631	0.693	0.683	0.703
Узк.	0.745	0.753	0.793	0.800	0.801
Доверительный куб для сумм в соседних парах ($\alpha = 0.10$)					
Диап.	$T = 5.5$ $\varphi = 0$	$T = 5.5$ $\varphi = \pi/4$	$T = 5$ $\varphi = 0$	$T = 5$ $\varphi = \pi/2$	$T = 5$ $\varphi = \pi/4$
Бол.	0.655	0.663	0.698	0.695	0.707
Ср.	0.721	0.725	0.766	0.767	0.772
Узк.	0.780	0.779	0.821	0.821	0.839

Б.4. Сравнение методов для различных параметров красного шума

Оценка мощности, сигнал $\mu(x) = 0.5 \sin(2\pi x/5.5)$				
Доверительный эллипсоид ($\alpha = 0.05$)				
γ	С.век.AR(1)	$\sin, \cos(1/2L)$	$\sin, \cos(1/4L)$	$\sin, \cos(1/8L)$
0.1	0.600	0.578	0.616	0.573
0.4	0.540	0.503	0.551	0.497
0.72	0.586	0.577	0.596	0.546
Оценка мощности, сигнал $\mu(x) = 0.5 \sin(2\pi x/10)$				
Доверительный эллипсоид ($\alpha = 0.05$)				
γ	С.век.AR(1)	$\sin, \cos(1/2L)$	$\sin, \cos(1/4L)$	$\sin, \cos(1/8L)$
0.1	0.579	0.610	0.474	0.569
0.4	0.319	0.356	0.285	0.359
0.72	0.213	0.228	0.215	0.266
Оценка мощности, сигнал $\mu(x) = 0.5 \sin(2\pi x/5.5)$				
Доверительный куб ($\alpha = 0.05$)				
γ	С.век.AR(1)	$\sin, \cos(1/2L)$	$\sin, \cos(1/4L)$	$\sin, \cos(1/8L)$
0.1	0.738	0.645	0.692	0.703
0.4	0.634	0.567	0.601	0.610
0.72	0.651	0.656	0.691	0.690
Оценка мощности, сигнал $\mu(x) = 0.5 \sin(2\pi x/10)$				
Доверительный куб ($\alpha = 0.05$)				
γ	С.век.AR(1)	$\sin, \cos(1/2L)$	$\sin, \cos(1/4L)$	$\sin, \cos(1/8L)$
0.1	0.617	0.654	0.623	0.615
0.4	0.292	0.338	0.315	0.309
0.72	0.207	0.227	0.207	0.200

Оценка мощности, сигнал $\mu(x) = 0.5 \sin(2\pi x/5.5)$				
Доверительный эллипсоид для минимумов в соседних парах ($\alpha = 0.05$)				
γ	С.век. AR(1)	$\sin, \cos(1/2L)$	$\sin, \cos(1/4L)$	$\sin, \cos(1/8L)$
0.1	0.369	0.566	0.570	0.752
0.4	0.457	0.600	0.517	0.619
0.72	0.646	0.679	0.617	0.697
Оценка мощности, сигнал $\mu(x) = 0.5 \sin(2\pi x/10)$				
Доверительный эллипсоид для минимумов в соседних парах ($\alpha = 0.05$)				
γ	С.век. AR(1)	$\sin, \cos(1/2L)$	$\sin, \cos(1/4L)$	$\sin, \cos(1/8L)$
0.1	0.530	0.400	0.425	0.645
0.4	0.378	0.249	0.283	0.414
0.72	0.229	0.219	0.239	0.308
Оценка мощности, сигнал $\mu(x) = 0.5 \sin(2\pi x/5.5)$				
Доверительный куб для минимумов в соседних парах ($\alpha = 0.05$)				
γ	С.век. AR(1)	$\sin, \cos(1/2L)$	$\sin, \cos(1/4L)$	$\sin, \cos(1/8L)$
0.1	0.530	0.746	0.736	0.720
0.4	0.523	0.670	0.662	0.628
0.72	0.680	0.740	0.741	0.690
Оценка мощности, сигнал $\mu(x) = 0.5 \sin(2\pi x/10)$				
Доверительный куб для минимумов в соседних парах ($\alpha = 0.05$)				
γ	С.век. AR(1)	$\sin, \cos(1/2L)$	$\sin, \cos(1/4L)$	$\sin, \cos(1/8L)$
0.1	0.606	0.458	0.630	0.648
0.4	0.365	0.224	0.331	0.331
0.72	0.231	0.146	0.208	0.205

Оценка мощности, сигнал $\mu(x) = 0.5 \sin(2\pi x/5.5)$				
Доверительный эллипсоид для сумм в соседних парах ($\alpha = 0.05$)				
γ	С.век.AR(1)	$\sin, \cos(1/2L)$	$\sin, \cos(1/4L)$	$\sin, \cos(1/8L)$
0.1	0.575	0.606	0.623	0.573
0.4	0.521	0.533	0.557	0.498
0.72	0.615	0.613	0.599	0.548
Оценка мощности, сигнал $\mu(x) = 0.5 \sin(2\pi x/10)$				
Доверительный эллипсоид для сумм в соседних парах ($\alpha = 0.05$)				
γ	С.век.AR(1)	$\sin, \cos(1/2L)$	$\sin, \cos(1/4L)$	$\sin, \cos(1/8L)$
0.1	0.543	0.505	0.479	0.571
0.4	0.323	0.299	0.286	0.364
0.72	0.219	0.221	0.218	0.258
Оценка мощности, сигнал $\mu(x) = 0.5 \sin(2\pi x/5.5)$				
Доверительный куб для сумм в соседних парах ($\alpha = 0.05$)				
γ	С.век.AR(1)	$\sin, \cos(1/2L)$	$\sin, \cos(1/4L)$	$\sin, \cos(1/8L)$
0.1	0.683	0.728	0.706	0.712
0.4	0.616	0.639	0.616	0.614
0.72	0.697	0.721	0.709	0.697
Оценка мощности, сигнал $\mu(x) = 0.5 \sin(2\pi x/10)$				
Доверительный куб для сумм в соседних парах ($\alpha = 0.05$)				
γ	С.век.AR(1)	$\sin, \cos(1/2L)$	$\sin, \cos(1/4L)$	$\sin, \cos(1/8L)$
0.1	0.653	0.597	0.623	0.625
0.4	0.360	0.304	0.322	0.311
0.72	0.206	0.185	0.210	0.208

Б.5. Монте Карло SSA с оценкой параметров красного шума

Оценка вероятности ошибки первого рода при неизвестных параметрах шума						
Доверительный эллипсоид ($\alpha = 0.10$)						
Диап.	С.ч.AR(1)	С.век.AR(1)	$\sin, \cos(1/2L)$	$\sin, \cos(1/4L)$	$\sin, \cos(1/8L)$	$\sin(1/L)$
Бол.	0.127	0.109	0.129	0.161	0.273	0.077
Ср.	0.108	0.076	0.091	0.120	0.148	0.068
Узк.	0.118	0.079	0.090	0.102	0.112	0.078
Доверительный куб ($\alpha = 0.10$)						
Диап.	С.ч.AR(1)	С.век.AR(1)	$\sin, \cos(1/2L)$	$\sin, \cos(1/4L)$	$\sin, \cos(1/8L)$	$\sin(1/L)$
Бол.	0.045	0.066	0.076	0.073	0.073	0.063
Ср.	0.030	0.038	0.057	0.057	0.070	0.057
Узк.	0.046	0.058	0.077	0.073	0.080	0.076
Доверительный эллипсоид для минимумов в соседних парах ($\alpha = 0.10$)						
Диап.	С.ч.AR(1)	С.век.AR(1)	$\sin, \cos(1/2L)$	$\sin, \cos(1/4L)$	$\sin, \cos(1/8L)$	$\sin(1/L)$
Бол.	0.127	0.111	0.112	0.187	0.298	—
Ср.	0.096	0.074	0.110	0.115	0.160	—
Узк.	0.121	0.082	0.114	0.103	0.114	—
Доверительный куб для минимумов в соседних парах ($\alpha = 0.10$)						
Диап.	С.ч.AR(1)	С.век.AR(1)	$\sin, \cos(1/2L)$	$\sin, \cos(1/4L)$	$\sin, \cos(1/8L)$	$\sin(1/L)$
Бол.	0.041	0.064	0.073	0.076	0.077	—
Ср.	0.023	0.038	0.069	0.058	0.065	—
Узк.	0.040	0.051	0.073	0.082	0.087	—

Доверительный эллипсоид для сумм соседних пар ($\alpha = 0.10$)						
Диап.	С.ч.AR(1)	С.век.AR(1)	$\sin, \cos(1/2L)$	$\sin, \cos(1/4L)$	$\sin, \cos(1/8L)$	$\sin(1/L)$
Бол.	0.126	0.109	0.130	0.169	0.266	—
Ср.	0.109	0.073	0.095	0.113	0.150	—
Узк.	0.116	0.072	0.097	0.106	0.108	—
Доверительный куб для сумм соседних пар ($\alpha = 0.10$)						
Диап.	С.ч.AR(1)	С.век.AR(1)	$\sin, \cos(1/2L)$	$\sin, \cos(1/4L)$	$\sin, \cos(1/8L)$	$\sin(1/L)$
Бол.	0.038	0.053	0.060	0.071	0.076	—
Ср.	0.025	0.034	0.058	0.054	0.067	—
Узк.	0.046	0.055	0.076	0.079	0.080	—
Оценка мощности при неизвестных параметрах шума, сигнал $\mu(x) = 0.5 \sin(2\pi x/5.5)$						
Доверительный эллипсоид ($\alpha = 0.10$)						
Диап.	С.ч.AR(1)	С.век.AR(1)	$\sin, \cos(1/2L)$	$\sin, \cos(1/4L)$	$\sin, \cos(1/8L)$	$\sin(1/L)$
Бол.	0.256	0.444	0.400	0.540	0.601	0.265
Ср.	0.239	0.465	0.463	0.562	0.511	0.290
Узк.	0.318	0.596	0.632	0.638	0.581	0.444
Доверительный куб ($\alpha = 0.10$)						
Диап.	С.ч.AR(1)	С.век.AR(1)	$\sin, \cos(1/2L)$	$\sin, \cos(1/4L)$	$\sin, \cos(1/8L)$	$\sin(1/L)$
Бол.	0.140	0.371	0.429	0.499	0.481	0.347
Ср.	0.100	0.452	0.500	0.571	0.563	0.438
Узк.	0.172	0.573	0.608	0.663	0.699	0.522
Доверительный эллипсоид для минимумов в соседних парах ($\alpha = 0.10$)						
Диап.	С.ч.AR(1)	С.век.AR(1)	$\sin, \cos(1/2L)$	$\sin, \cos(1/4L)$	$\sin, \cos(1/8L)$	$\sin(1/L)$
Бол.	0.246	0.469	0.509	0.574	0.741	—
Ср.	0.233	0.537	0.609	0.548	0.661	—
Узк.	0.316	0.664	0.734	0.646	0.728	—
Доверительный куб для минимумов в соседних парах ($\alpha = 0.10$)						
Диап.	С.ч.AR(1)	С.век.AR(1)	$\sin, \cos(1/2L)$	$\sin, \cos(1/4L)$	$\sin, \cos(1/8L)$	$\sin(1/L)$
Бол.	0.140	0.406	0.537	0.539	0.472	—
Ср.	0.101	0.473	0.607	0.620	0.568	—
Узк.	0.174	0.593	0.689	0.703	0.692	—

Доверительный эллипсоид для сумм соседних пар ($\alpha = 0.10$)						
Диап.	С.ч.AR(1)	С.век.AR(1)	$\sin, \cos(1/2L)$	$\sin, \cos(1/4L)$	$\sin, \cos(1/8L)$	$\sin(1/L)$
Бол.	0.256	0.447	0.417	0.542	0.599	—
Ср.	0.227	0.478	0.504	0.572	0.513	—
Узк.	0.301	0.625	0.653	0.646	0.592	—
Доверительный куб для сумм соседних пар ($\alpha = 0.10$)						
Диап.	С.ч.AR(1)	С.век.AR(1)	$\sin, \cos(1/2L)$	$\sin, \cos(1/4L)$	$\sin, \cos(1/8L)$	$\sin(1/L)$
Бол.	0.137	0.440	0.501	0.514	0.503	—
Ср.	0.096	0.528	0.595	0.585	0.575	—
Узк.	0.160	0.634	0.664	0.671	0.701	—