

ЗАКОН ПОДОБИЯ ПРИ СПЕКТРАЛЬНОМ ОЦЕНИВАНИИ ВРЕМЕННОГО РЯДА. III

В.Ю.ТЕРЕБИЖ

Поступила 9 октября 1996

Принята к печати 10 декабря 1996

Продолжение статей [1,2]. Найдено точное представление основной величины, характеризующей обратную задачу - матрицы Фишера для коэффициентов ковариации. Применение неравенства информации (Рао-Крамера) позволило найти нижнюю границу дисперсии произвольной несмещенной оценки ковариационных коэффициентов. Выводы иллюстрируются примером процесса авторегрессии первого порядка.

6. *Оценки коэффициентов ковариации.* Наиболее употребительными статистическими оценками параметров являются оценки наименьших квадратов и максимального правдоподобия. Как известно, для гауссовых случайных величин эти две оценки совпадают. Оценка максимального правдоподобия $\hat{r}(y)$ (ML-оценка) вектора r есть то значение искомого (векторного) параметра, при котором вероятность получить наблюдаемое значение реализации максимальна, то есть $\hat{r}(y)$ доставляет максимум плотности распределения $\phi(y|r)$, рассматриваемой в качестве функции r при фиксированной реализации y . Необходимые условия экстремума задают $\hat{r}(y)$ как решение системы уравнений

$$\frac{\partial}{\partial r_k} \ln \phi(y|r) = 0, \quad k = 0, 1, \dots, N-1. \quad (6.1)$$

Из (5.3) следует:

$$-\frac{\partial}{\partial r_k} \ln \phi(y|r) = \frac{1}{2} \left[\frac{\partial \ln D}{\partial r_k} + y' \frac{\partial R^{-1}}{\partial r_k} y \right]. \quad (6.2)$$

При помощи формулы (A3) из *Приложения А* первое слагаемое в квадратных скобках (6.2) можно записать следующим образом:

$$\frac{\partial \ln D}{\partial r_k} = \text{tr} \left(R^{-1} \frac{\partial R}{\partial r_k} \right), \quad (6.3)$$

где $\text{tr}(A)$ означает след матрицы A . Представим далее ковариационную матрицу R , заданную соотношением (2.18), как суперпозицию определенных в *Приложении В* матриц Z_k :

$$R = r_0 Z_0 + 2 \sum_{k=1}^{N-1} r_k Z_k. \quad (6.4)$$

Отсюда следует:

$$\frac{\partial R}{\partial r_0} = Z_0 = E_N, \quad \frac{\partial R}{\partial r_k} = 2 Z_k, \quad k = 1, 2, \dots, N-1, \quad (6.5)$$

где E_N - единичная $N \times N$ -матрица. Таким образом,

$$\frac{\partial \ln D}{\partial r_0} = \text{tr}(R^{-1}), \quad \frac{\partial \ln D}{\partial r_k} = 2 \text{tr}(R^{-1} Z_k), \quad k = 1, 2, \dots, N-1. \quad (6.6)$$

Что касается второго слагаемого в (6.2), то, согласно (A1),

$$\frac{\partial R^{-1}}{\partial r_k} = -R^{-1} \frac{\partial R}{\partial r_k} R^{-1}, \quad k = 0, 1, \dots, N-1. \quad (6.7)$$

Здесь вновь следует учесть (6.5).

Подставляя все полученные выше выражения для производных в (6.2), мы приходим к следующей системе уравнений, служащей для нахождения ML-оценки $\hat{R} = R(\hat{r})$ ковариационной матрицы:

$$\text{tr}(\hat{R}^{-1} Z_k) = y' \hat{R}^{-1} Z_k \hat{R}^{-1} y, \quad k = 0, 1, \dots, N-1. \quad (6.8)$$

Для придания (6.8) более компактного вида учтем, что правая часть этой системы равна, как и всякий скаляр, собственному следу. Принимая во внимание также известное свойство $\text{tr}(AB) = \text{tr}(BA)$, находим:

$$\text{tr}[(\hat{R}^{-1} S - E_N) \hat{R}^{-1} Z_k] = 0 \quad k = 0, 1, \dots, N-1. \quad (6.9)$$

Здесь все выборочные данные сосредоточены в матрице $S \equiv yy'$, так что ML-оценка ковариационной матрицы зависит имеено от S . Доумножая (6.9) на r_k , суммируя все полученные уравнения и принимая во внимание разложение (6.4), получаем также:

$$\frac{1}{N} \text{tr}(\hat{R}^{-1} S) = 1. \quad (6.10)$$

На основании равенств $y' R^{-1} y = \text{tr}(y' R^{-1} y) = \text{tr}(R^{-1} yy') = \text{tr}(R^{-1} S)$ последнему соотношению можно придать эквивалентную форму:

$$\frac{1}{N} y' \hat{R}^{-1} y = 1. \quad (6.10a)$$

Найдем еще значение информации в точке максимального правдоподобия. Из (5.3) и (6.10) следует:

$$J(y|\hat{r}) = \frac{1}{2} [N \ln(2\pi e) + \ln D(\hat{r})]. \quad (6.11)$$

Правая часть здесь совпадает со значением энтропии (5.5) в точке $r = \hat{r}$, так что

$$J(y|\hat{r}) = \mathcal{E}(\hat{r}). \quad (6.12)$$

В обратных задачах оценки максимального правдоподобия и наименьших квадратов (близкие к так называемому *инверсному* решению) часто оказываются неудовлетворительными: они требуют слишком тесного с вероятностной точки зрения соответствия данных и модели (см. [3,4]). Строго говоря, ввиду случайности реализации y выборочная информация $J(y|r)$ также представляет собой случайную величину, и потому на опыте мы вправе ожидать реализации не минимального ее значения $J_{\min}(y) = J(y|\hat{r})$, а только близкого к среднему $\langle J(y|r) \rangle = \mathcal{E}(r)$. Поэтому разумно ввести, следуя [5], совокупность *оценок средней информации* $\bar{r}(y)$, определенных условием

$$J(y|\bar{r}) = \mathcal{E}(\bar{r}). \quad (6.13)$$

Подчеркнем, что это соотношение задает, вообще говоря, не одну, а множество оценок. Всякая *приемлемая* (feasible) в статистическом смысле оценка r обязана быть близкой к слою (6.13), однако это условие, конечно, не является достаточным для нахождения *эффективной* оценки (подробное обсуждение можно найти в [3,4]). В случае гауссова процесса определение (6.13) конкретизируется с помощью (5.3) и (5.5) следующим образом:

$$\frac{1}{N} y' \bar{R}^{-1} y = 1. \quad (6.13a)$$

Как видно из (6.12) и (6.13), особенность максимально правдоподобной оценки $\hat{r}(y)$ коэффициентов ковариации временного ряда заключается в том, что она одновременно является и оценкой *средней информации* (среднего правдоподобия). Таким образом, отвергнуть $\hat{r}(y)$ можно было бы только с привлечением более сильных по сравнению с условием (6.13) требований. В отличие от обычной для обратных задач ситуации, здесь сохраняется интерес к непосредственному инверсному оцениванию $\hat{r}(y)$.

В этой связи отметим, что соотношения (6.10) и (6.11) были ранее найдены Бергом и др. [6], причем указанные авторы полагают, что ML-оценка должна находиться из (6.10). В нашей интерпретации (6.10) есть интегральное следствие системы уравнений (6.9), само по себе

задающее множество оценок средней информации, включая и ML-оценку. Для нахождения последней должны использоваться уравнения (6.8) или (6.9).

К сожалению, найти аналитическое решение этих уравнений пока не удастся. Дальше возможно продвинуться в исследовании решений уравнения (6.13а). Заметим прежде всего, что в качестве возможных оценок $\bar{R}$ рассматриваются не произвольные матрицы, но только матрицы Лорана со структурой (2.18), которые определяются значениями N элементов. Известно (см. [7], §3.7), что при $N \gg 1$ спектр матриц Лорана аппроксимируется спектром симметрических циркуляционных матриц. Последние имеют весьма простое спектральное представление, что позволяет существенно упростить (6.13а), а в частных случаях и решить это уравнение. Общее же решение, по-видимому, является сложной функцией выборки y .

Сказанное выше заставляет более внимательно изучить два известных представления для оценок коэффициентов ковариации:

$$c_n = \frac{1}{N-n} \sum_{k=0}^{N-1-n} y_k y_{k+n}, \quad n = 0, 1, \dots, N-1; \quad (6.14)$$

и

$$\tilde{c}_n = \frac{1}{N} \sum_{k=0}^{N-1-n} y_k y_{k+n}, \quad n = 0, 1, \dots, N-1. \quad (6.15)$$

Ввиду (2.10) среднее значение

$$\langle c_n \rangle = \frac{1}{N-n} \sum_{k=0}^{N-1-n} \langle \xi_k \xi_{k+n} \rangle = \frac{1}{N-n} \cdot (N-n) r_n = r_n, \quad (6.16)$$

то есть c - несмещенная оценка r . Поскольку $\tilde{c} = (1 - n/N) \cdot c$, оценка (6.15) при всяком конечном значении n обладает смещением, лишь асимптотически стремящимся к нулю.

Качество $c(y)$ и $\tilde{c}(y)$, рассматриваемых в качестве оценок вектора r , проверялось следующим образом. Для нескольких значений $N \leq 64$ мы моделировали совокупность реализаций временного ряда y для процесса AR-1 при различных значениях ρ (см. §3), затем для каждой реализации согласно (6.14) и (6.15) вычислялись оценки $c(y)$, $\tilde{c}(y)$ и, наконец, вычислялась левая часть (6.10). Ее разброс и систематическое смещение от теоретического значения 1 могут рассматриваться как мера качества рассматриваемой оценки r . Результаты моделирования показывают, что обе оценки $c(y)$ и $\tilde{c}(y)$ удовлетворяют (6.10) с неплохой точностью, причем смещенная оценка $\tilde{c}(y)$ в этом смысле явно предпочтительнее

несмещенной $c(y)$. Поскольку для нас этот вопрос не играет заметной роли, мы будем далее чаще пользоваться представлением (6.14).

Пусть случайный вектор $\eta(y)$ с компонентами

$$\eta_n = c_n - r_n, \quad n = 0, 1, \dots, N-1 \quad (6.17)$$

описывает расхождение между выборочными несмещенными оценками коэффициентов ковариации и их истинными значениями, так что $\langle \eta \rangle = 0$. Пусть далее

$$Q = \langle (c-r) \cdot (c-r)' \rangle = \langle \eta \cdot \eta' \rangle \quad (6.18)$$

- ковариационная матрица c , совпадающая с таковой для вектора η . Точное представление матрицы Q для произвольных значений N можно найти в теореме 8.2.6 [8]. В асимптотической области $N-m \gg 1$, $N-n \gg 1$ элементы Q_{mn} можно выразить либо через спектральную плотность процесса $g_0(v)$:

$$Q_{mn} \approx \frac{2}{\sqrt{(N-m)(N-n)}} \int_{-1/2}^{1/2} \cos(2\pi v m) \cos(2\pi v n) g_0^2(v) dv, \quad (6.19)$$

либо через ковариационную функцию $\{r_n\}$:

$$Q_{mn} \approx \frac{1}{\sqrt{(N-m)(N-n)}} \sum_{k=-\infty}^{\infty} r_k (r_{m+n-k} + r_{m-n-k}). \quad (6.20)$$

Как показывают расчеты, в обычных на практике условиях последние два представления имеют приемлемую точность уже при значениях $N-m$ и $N-n$ порядка 10. Для частотно-ограниченного белого шума мы имеем из (2.15) и (6.19):

$$Q_{00} \approx 2/N, \quad Q_{nn} \approx 1/(N-n), \quad n = 1, 2, \dots, N-1, \quad (6.21)$$

а недиагональные элементы можно считать равными нулю.

Что касается плотности распределения случайного вектора η , то при $N \gg 1$ она стремится к гауссовой плотности с нулевым средним и ковариационной матрицей Q ([8], теорема 8.4.2).

Задача оценивания спектральной плотности $s(f)$ по выборочным значениям ковариационной функции ставится на основании (6.17) и (4.1) следующим образом:

$$c_n = \int_{-1/2}^{1/2} \exp(i \cdot 2\pi n f) s(f) df + \eta_n, \quad n = 0, 1, \dots, N-1. \quad (6.22)$$

Эти уравнения описывают типичную обратную задачу восстановления

оригинала $\mathcal{J}(f)$ на основании наблюдаемой реализации s и гауссовой ошибки η , характеризующейся ковариационной матрицей Q .

7. *Матрица Фишера для ковариационных коэффициентов.* По определению (см. [9]), матрица Фишера задается усреднением вторых производных от логарифма плотности распределения по искомым параметрам. С учетом соотношения (5.1) можно написать:

$$I_{ik}(r) = \langle \frac{\partial^2}{\partial r_i \partial r_k} J(\xi|r) \rangle, \quad i, k = 0, 1, \dots, N-1. \quad (7.1)$$

Подстановка сюда выражения (5.3) дает:

$$2 I_{ik}(r) = \frac{\partial^2 \ln D}{\partial r_i \partial r_k} + \left\langle \xi' \frac{\partial^2 R^{-1}}{\partial r_i \partial r_k} \xi \right\rangle. \quad (7.2)$$

Согласно формуле (A4) Приложения А,

$$\frac{\partial^2 \ln D}{\partial r_i \partial r_k} = \text{tr} \left(R^{-1} \frac{\partial^2 R}{\partial r_i \partial r_k} - R^{-1} \frac{\partial R}{\partial r_i} R^{-1} \frac{\partial R}{\partial r_k} \right). \quad (7.3)$$

Поскольку выражение в угловых скобках (7.2) есть скаляр, можно написать:

$$\begin{aligned} \langle \xi' \frac{\partial^2 R^{-1}}{\partial r_i \partial r_k} \xi \rangle &= \langle \text{tr} \left(\xi' \frac{\partial^2 R^{-1}}{\partial r_i \partial r_k} \xi \right) \rangle = \langle \text{tr} \left(\xi \xi' \frac{\partial^2 R^{-1}}{\partial r_i \partial r_k} \right) \rangle = \text{tr} \left(R \frac{\partial^2 R^{-1}}{\partial r_i \partial r_k} \right) = \\ &= \text{tr} \left(\frac{\partial R}{\partial r_i} R^{-1} \frac{\partial R}{\partial r_k} R^{-1} + \frac{\partial R}{\partial r_k} R^{-1} \frac{\partial R}{\partial r_i} R^{-1} - \frac{\partial^2 R}{\partial r_i \partial r_k} R^{-1} \right), \end{aligned} \quad (7.4)$$

где мы воспользовались также (A2). Из последних трех формул следует:

$$I_{ik}(r) = \frac{1}{2} \text{tr} \left(\frac{\partial R}{\partial r_i} R^{-1} \frac{\partial R}{\partial r_k} R^{-1} \right), \quad i, k = 0, 1, \dots, N-1. \quad (7.5)$$

Полученное представление матрицы Фишера справедливо для произвольной ковариационной матрицы R . В рассматриваемом здесь случае R есть лорановская матрица вида (6.4). Подстановка этого выражения в (7.5) дает искомое точное представление матрицы Фишера для случая гауссова процесса:

$$\begin{cases} I_{00}(r) = \frac{1}{2} \text{tr}(R^{-2}), \\ I_{0k}(r) = I_{k0}(r) = \text{tr}(R^{-1} Z_k R^{-1}), & k = 1, \dots, N-1, \\ I_{ik}(r) = 2 \text{tr}(Z_i R^{-1} Z_k R^{-1}), & i, k = 1, \dots, N-1, \end{cases} \quad (7.6)$$

где семейство матриц $\{Z_k\}$ определено в *Приложении В*.

Рассмотрим в качестве простого примера частотно-ограниченный белый шум (2.15) с единичной дисперсией $\sigma^2 = r_0 = 1$. При этом все старшие коэффициенты ковариации равны нулю, спектральная плотность $s(f) \equiv 1$, а ковариационная матрица R в (2.18) и матрица R^{-1} , обратная к ней, совпадают с единичной матрицей E_N . Обращаясь к формулам (7.6), нетрудно найти, что матрица Фишера $I(r)$ диагональна, причем элементы на главной диагонали равны

$$I_{\infty}(r) = N/2, \quad I_{nn}(r) = N - n, \quad 1 \leq n \leq N-1. \quad (7.7)$$

Важно отметить, что в данном случае $I(r)$ в точности равна обратной матрице Q , определенной в (6.18). В общем случае это равенство не выполняется, однако при $N \gg 1$ можно считать

$$I(r) \approx Q^{-1}. \quad (7.8)$$

Так, численные оценки для случая процесса AR-1 показывают, что соотношение (7.8) выполняется с неплохой точностью уже при $N \approx 10$, если только модуль коэффициента корреляции процесса ρ не слишком близок к 1.

В качестве примера на рис. 7.1 показан вычисленный по формулам (7.6) профиль матрицы $I(r)$ для процесса AR-1. Для удобства построения графика всем элементам матрицы было придано постоянное положительное смещение. По мере возрастания объема выборки N матрица Фишера все более приближается к чисто диагональной матрице, что свидетельствует об асимптотической некоррелированности оценок

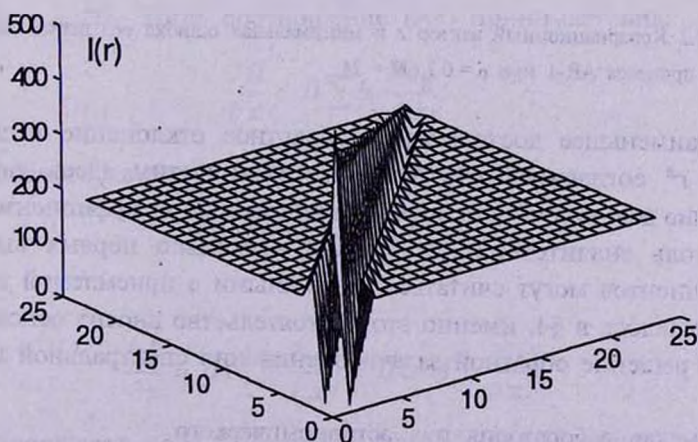


Рис. 7.1 Матрица Фишера $I(r)$ для процесса AR-1 при $\rho = 0.7$ и $N = 24$; всем элементам матрицы придано положительное смещение 180.702.

коэффициентов ковариации.

Как известно [9], знание матрицы Фишера какой-либо величины позволяет указать нижнюю границу для дисперсии наилучшей несмещенной оценки этой величины. Пусть в данном случае $r^*(y)$ означает произвольную несмещенную оценку вектора ковариаций r . Тогда ее дисперсия

$$\text{Var}(r^*) \geq \text{diag}[I^{-1}(r)], \quad (7.9)$$

где $\text{diag}(C)$ означает вектор с элементами, находящимися на главной диагонали квадратной матрицы C .

На рис. 7.2 приведены точные значения компонентов вектора r для тех же условий, которые были описаны выше в связи с рис. 7.1, а

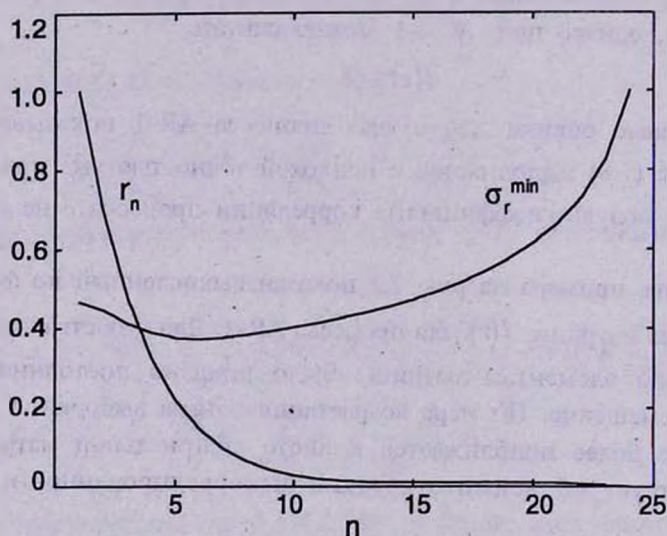


Рис. 7.2 Ковариационный вектор r и минимальная ошибка его несмещенной оценки $\sigma_r^{\min}$ для процесса AR-1 при $\rho = 0.7$, $N = 24$;

также наименьшее достижимое стандартное отклонение несмещенной оценки r^* согласно (7.9). Мы отчетливо видим здесь характерное положение дел: неточность определения старших коэффициентов ковариации столь значительна, что только несколько первых выборочных коэффициентов могут считаться найденными с приемлемой точностью. Как отмечалось в §4, именно это обстоятельство вносит основные трудности в решение обратной задачи оценивания спектральной плотности.

Приложение А

Производные от обратной матрицы и определителя. Пусть A - произвольная неособенная матрица размерами $N \times N$, элементы которой $a_y(x, y)$ являются функциями скаляров x и y . Обозначим через D определитель матрицы A и через A^{-1} - матрицу, обратную A . Тогда

$$\frac{\partial A^{-1}}{\partial x} = -A^{-1} \cdot \frac{\partial A}{\partial x} \cdot A^{-1}, \quad (A1)$$

$$\frac{\partial^2 A^{-1}}{\partial x \partial y} = A^{-1} \cdot \left[\frac{\partial A}{\partial x} \cdot A^{-1} \cdot \frac{\partial A}{\partial y} + \frac{\partial A}{\partial y} \cdot A^{-1} \cdot \frac{\partial A}{\partial x} - \frac{\partial^2 A}{\partial x \partial y} \right] \cdot A^{-1}, \quad (A2)$$

$$\frac{\partial \ln D}{\partial x} = \text{tr} \left[A^{-1} \cdot \frac{\partial A}{\partial x} \right], \quad (A3)$$

$$\frac{\partial^2 \ln D}{\partial x \partial y} = \text{tr} \left[A^{-1} \cdot \frac{\partial^2 A}{\partial x \partial y} - A^{-1} \cdot \frac{\partial A}{\partial x} \cdot A^{-1} \cdot \frac{\partial A}{\partial y} \right], \quad (A4)$$

где $\text{tr}(C)$ - след матрицы C .

Формула (A1) следует из определения обратной матрицы: $A \cdot A^{-1} = E_N$. Повторное дифференцирование (A1) приводит к (A2).

Далее, принимая во внимание определение алгебраического дополнения A_y элемента a_y , мы можем написать:

$$\frac{\partial D}{\partial x} = \sum_{i,j} \frac{\partial D}{\partial a_y} \frac{\partial a_y}{\partial x} = \sum_{i,j} A_y \frac{\partial a_y}{\partial x}. \quad (A5)$$

Обозначим для упрощения записи $A^{-1} \equiv B$, так что элементы матрицы B суть $b_{ij} = A_{ij}/D$. Тогда соотношение (A5) принимает вид:

$$\frac{\partial D}{\partial x} = D \sum_{i,j} b_{ij} \frac{\partial a_y}{\partial x}. \quad (A6)$$

Но если некоторая матрица $C = B \cdot A$, то ее след

$$\text{tr}(C) \equiv \sum_i c_{ii} = \sum_{i,j} b_{ij} a_{ji}. \quad (A7)$$

Поэтому

$$\frac{\partial D}{\partial x} = D \text{tr} \left(B \cdot \frac{\partial A}{\partial x} \right) = D \text{tr} \left(A^{-1} \cdot \frac{\partial A}{\partial x} \right), \quad (A8)$$

что и утверждает (A3). Наконец, формула (A4) получается путем дифференцирования (A3).

Приложение В

Z-матрицы. Пусть все ненулевые элементы $N \times N$ -матрицы T расположены над ее главной диагональю и равны 1. Определим семейство симметрических матриц $\{Z_k\}$ соотношениями:

$$Z_k \equiv \frac{1}{2}(T^k + T^{*k}), \quad k = 0, 1, \dots \quad (B1)$$

В частности,

$$Z_0 = E_N, \quad Z_1 = \frac{1}{2} \begin{bmatrix} 010 \dots 00 \\ 101 \dots 00 \\ \dots \dots \dots \\ 000 \dots 01 \\ 000 \dots 10 \end{bmatrix} \quad (B2)$$

В матрице Z_k , $k \geq 1$ единичные элементы расположены на диагоналях, параллельных главной диагонали и сдвинутых от нее на $\pm k$ шагов.

Из теоремы 6.5.5 [8] следует, что собственные числа Z_1 равны

$$\lambda_j = \cos\left(\frac{\pi j}{N+1}\right), \quad j = 1, 2, \dots, N, \quad (B3)$$

а элементы соответствующих ортонормированных собственных векторов суть

$$V_{kj} = \left(\frac{2}{N+1}\right)^{1/2} \sin\left(\frac{\pi kj}{N+1}\right), \quad k, j = 1, 2, \dots, N. \quad (B4)$$

Представляет значительный интерес спектральное разложение других матриц рассматриваемого семейства.

SIMILARITY LAW IN A TIME SERIES SPECTRAL ESTIMATION. III

V.Yu.TEREBIZH

Continuation of [1,2]. An exact analytic representation of Fisher matrix for covariation coefficients is given. An application of the information (Rao-Cramer) inequality allows us to find a lower boundary for variance of any unbiased estimate of covariation coefficients. The results are illustrated by an example of autoregression process of first order.

ЛИТЕРАТУРА

1. В.Ю.Теребиж, *Астрофизика*, **40**, 139, 1997 (часть I данной серии).
2. В.Ю.Теребиж, *Астрофизика*, **40**, 273, 1997 (часть II данной серии).
3. В.Ю.Теребиж, *Успехи физ. наук*, **165**, 143, 1995.
4. V.Yu.Terebizh, *Int. Journ. of Imaging Systems and Technology*, **6**, 358, 1995.
5. V.Yu.Terebizh, V.V.Biryukov, *Astron. and Astrophys. Transact.*, **6**, 37, 1994.
6. J.P.Burg, D.G.Luenberger, D.L.Wenger, *Tr. IEEE*, **70**, 963, 1982.
7. D.R.Brillinger, *Time Series. Data Analysis and Theory*. Holt, Rinchart and Winston, New York, 1975; Д.Бриллинджер, *Временные ряды. Обработка данных и теория*, Мир, М., 1980.
8. T.W.Anderson, *The Statistical Analysis of Time Series*, Wiley, New York, 1971; Т.Андерсон, *Статистический анализ временных рядов*, Мир, М., 1976.
9. А.А.Боровков, *Математическая статистика*, Наука, М., 1984.