АСТРОФИЗИКА

ТОМ 40 МАЙ, 1997 ВЫПУСК 2

УДК: 52-355:51

ЗАКОН ПОДОБИЯ ПРИ СПЕКТРАЛЬНОМ ОЦЕНИВАНИИ ВРЕМЕННО'ГО РЯДА. II.*

В.Ю.ТЕРЕБИЖ

Поступила 9 октября 1996

Во второй части работы (см. [1]) приводятся результаты вспомогательного характера для процесса авторегрессии первого порядка. Дана формальная постановка задачи оценивания спектральной плотности временно'го ряда как обратной задачи математической физики.

3. *Процесс авторегрессии первого порядка.* Для иллюстрации приводимых далее общих соотношений случай частотно-ограниченного белого шума слишком прост, поэтому рассмотрим более сложный пример, играющий вместе с тем важную практическую роль.

В случае гауссова авторегрессионного процесса первого порядка (см. [2,3]) с дисперсией σ^2, обозначаемого обычно посредством AR-1, последовательные его значения $\xi_k \equiv \xi(k \cdot \delta t)$ связаны линейным соотношением

$$\xi_{k+1} = \rho \cdot \xi_k + \varepsilon_{k+1}, \quad k = 0,1\ldots, \tag{3.1}$$

где ρ - коэффициент корреляции между ними (мы считаем $|\rho| < 1$), а (ε_k) - некоррелированные гауссовы случайные величины с диперсией $\sigma_\varepsilon^2 = (1-\rho^2)\sigma^2$. Типичная реализация процесса AR-1 при ρ=0.7 и σ^2=1 представлена на рис. 1.1.

Спектральная плотность процесса AR-1 равна

$$g_0(\nu) = \frac{(1-\rho^2)\sigma^2 \delta t}{1+\rho^2 - 2\rho \cdot \cos(2\pi\nu\delta t)}, \quad |\nu| \le \nu_c. \tag{3.2}$$

Как это обычно делается, перейдем к безразмерной частоте f и введем соответствующую спектральную плотность $s(f)$:

$$f \equiv \nu \cdot \delta t, \quad s(f) \equiv g_0(\nu)/\delta t, \quad -1/2 \le f \le 1/2. \tag{3.3}$$

---

* *Нумерация параграфов в этой и последующих трех статьях этой серии является продолжением нумерации первой статьи, напечатанной в предыдущем выпуске журнала.*

Обозначая еще $b \equiv \sigma_\varepsilon^2 = \left(1-\rho^2\right)\sigma^2 > 0$, получаем:

$$s(f) = \frac{b}{1+\rho^2 - 2\rho\cdot\cos(2\pi f)}, \quad |f| \le 1/2. \tag{3.4}$$

Подстановка (3.2) в (2.11) дает следующее простое выражение для ковариационных коэффициентов:

$$r_n = r_0\cdot\rho^n, \quad r_0 = \sigma^2, \quad n = 0,1\ldots, \tag{3.5}$$

и конечно, $r_{-n}=r_n$. Таким образом, абсолютная величина r_n экспоненциально убывает с ростом запаздывания (процесс AR-1 служит моделью *марковского ряда*). Согласно (2.18) и (3.5), ковариационная матрица процесса AR-1 имеет вид:

$$R = \frac{b}{1-\rho^2}\cdot\begin{bmatrix} 1 & \rho & \rho^2 & \ldots & \rho^{N-1} \\ \rho & 1 & \rho & \ldots & \rho^{N-2} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \rho^{N-1} & & \rho^{N-2} & \ldots & 1 \end{bmatrix}, \tag{3.6}$$

где мы приняли во внимание, что $r_0=b/(1-\rho^2)$. Поскольку в данном случае определитель $D \equiv \det(R) = b^N/\left(1-\rho^2\right)$, входящая в (2.17) обратная матрица R^{-1} равна

$$R^{-1} = b^{-1}\cdot\begin{bmatrix} 1 & -\rho & 0 & \ldots & 0 & 0 & 0 \\ -\rho & 1+\rho^2 & -\rho & \ldots & 0 & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & -\rho & 1+\rho^2 & -\rho \\ 0 & 0 & 0 & \ldots & 0 & -\rho & 1 \end{bmatrix}. \tag{3.7}$$

Таким образом, плотность распределения (2.17) N - мерного вектора ξ можно записать в виде:

$$\varphi(y|b,\rho) = \sqrt{\frac{1-\rho^2}{(2\pi b)^N}}\cdot\exp\left\{-\frac{1}{2b}\left[\left(1-\rho^2\right)y_0^2 + \sum_{k=0}^{N-2}\left(y_{k+1} - \rho\, y_k\right)^2\right]\right\}. \tag{3.8}$$

Модель AR-1 определяется значениями всего двух параметров: b и ρ. По определению, при фиксированной реализации y временно́го ряда *оценки максимального правдоподобия* этих параметров $\hat{b}(y)$ и $\hat{\rho}(y)$ находятся как решения уравнений:

$$\frac{\partial}{\partial b}\left[\ln\varphi(y|b,\rho)\right] = 0, \quad \frac{\partial}{\partial \rho}\left[\ln\varphi(y|b,\rho)\right] = 0. \tag{3.9}$$

Нетрудно проверить, что в случае $N{>>}1$ мы имеем из (3.8) и (3.9):

$$\hat{\rho}(y) \cong \sum_{k=0}^{N-2} y_k y_{k+1} \Big/ \sum_{k=0}^{N-2} y_k^2, \tag{3.10}$$

$$\hat{b}(y) \cong \frac{1}{N} \sum_{k=0}^{N-2} \left(y_{k+1} - \hat{\rho}\, y_k\right)^2. \tag{3.11}$$

Смысл выражений в правых частях двух последних формул вполне понятен: $\hat{\rho}$ оценивается как выборочный коэффициент корреляции последовательных отсчетов, а $\hat{b}$ - как выборочная дисперсия шума ε в модели (3.1).

Для выяснения качества произвольных статистических оценок b и ρ следует найти *матрицу Фишера* [4]. В данном случае это матрица $I(b,\rho)$ второго порядка с элементами:

$$I_{11} \equiv \left< -\frac{\partial^2}{\partial b^2} \ln\varphi(\xi|b,\rho) \right> = \frac{N}{2b^2}, \tag{3.12}$$

$$I_{12} = I_{21} \equiv \left< -\frac{\partial^2}{\partial b\, \partial\rho} \ln\varphi(\xi|b,\rho) \right> = \frac{\rho}{b(1-\rho)^2}, \tag{3.13}$$

$$I_{22} \equiv \left< -\frac{\partial^2}{\partial\rho^2} \ln\varphi(\xi|b,\rho) \right> = \frac{N-1-(N-3)\rho^2}{\left(1-\rho^2\right)^2}. \tag{3.14}$$

Приведенные здесь точные выражения следуют из (3.8); в асимптотическом случае, когда $N>>1$, мы получаем из них для не слишком близких к 1 значений ρ:

$$I \cong N \cdot \begin{bmatrix} 1/\left(2b^2\right) & 0 \\ 0 & 1/\left(1-\rho^2\right) \end{bmatrix}, \quad I^{-1} \cong \frac{1}{N} \cdot \begin{bmatrix} 2b^2 & 0 \\ 0 & 1-\rho^2 \end{bmatrix}. \tag{3.15}$$

Согласно *неравенству информации* (см. [4]), дисперсии произвольных несмещенных оценок b^* и ρ^* искомых параметров ограничены снизу значениями

$$\mathrm{Var}\left(b^*\right) \ge \left(I^{-1}\right)_{11}, \quad \mathrm{Var}\left(\rho^*\right) \ge \left(I^{-1}\right)_{22}. \tag{3.16}$$

Подставляя сюда из (3.15) диагональные элементы обратной матрицы Фишера, получаем следующие нижние границы стандартных отклонений:

$$\sigma\left(b^*\right) \ge b\sqrt{2/N}, \quad \sigma\left(\rho^*\right) \ge \sqrt{\left(1-\rho^2\right)/N}. \tag{3.17}$$

Как видно, по мере увеличения объема выборки точность оценивания параметров b и ρ неограниченно возрастает, то есть оценки, близкие к границе неравенства информации, являются *состоятельными*. Что касается возможной связи между оценками b^* и ρ^*, то из (3.15) следует, что они асимптотически некоррелированы.

Таким образом, если нам задан достаточно протяженный временно'й ряд y и, кроме того, *известно, что он является реализацией процесса* AR-1, то можно с надлежащей точностью оценить значения параметров b и ρ, а после подстановки полученных значений в (3.4) мы найдем и оценку спектральной плотности. Разумеется, последняя представляет собой гладкую функцию.

Иной результат мы получим в том случае, когда информация о типе изучаемого процесса отсутствует. При этом вместо двух параметров приходится оценивать значения гораздо более обширной совокупности, а именно, отсчетов спектральной плотности при различных значениях частоты. Согласно принятой терминологии, здесь производится *непараметрическое оценивание*, хотя в действительности оцениваются тоже значения параметров, только их количество велико. Как уже говорилось во Введении, статистики типа периодограммы Шустера оказываются несостоятельными.

Приведенный пример наглядно демонстрирует важность априорной информации о характере искомого объекта. В этой связи стоит заметить, что метод максимума энотропии Берга [5] по сути дела вводит предположение о том, что процесс относится к классу авторегрессионных AR-p, и лишь его порядок p остается неизвестным и подлежит оцениванию. Именно на этом предположении и на достаточной гибкости описания с помощью процессов AR-p основана та эффективность метода максимума энтропии, которую он показыват во многих реальных ситуациях. Вместе с тем, если исходный процесс далек от авторегрессионного (а к таковым можно отнести уже простую суперпозицию чисто периодической функции и случайного шума), то выводы метода максимума энтропии будут далеки от действительности. Это лишний раз свидетельствует о том, что задача анализа данных состоит не только в использовании *всей* имеющейся априорной информации относительно объекта, но и в использовании *только* этой информации.

4. *Постановка обратной задачи спектрального оценивания.* Напомним общую постановку задачи оценивания, вообще говоря, многомерного параметра a по данным наблюдений [4,6]. Пусть y - известная нам реализация многомерной случайной величины ξ, плотность распределения которой $\varphi(\cdot|a)$ зависит от параметра a. Тогда в качестве *оценки* параметра a может рассматриваться всякая функция $u(y)$, - как говорят, *выборочная статистика*. Следуя Фишеру, среди всех таких статистик нужно разыскать оценку с наименьшим среднеквадратическим отклонением от истинного значения a; эта оценка называется *эффективной.*

При спектральном оценивании в нашем распоряжении имеется конечная выборка значений временного ряда $y \equiv [y_0, y_1, \dots y_{N-1}]'$, являющаяся реализацией гауссовой случайной величины $\xi \equiv [\xi_0, \xi_1, \dots \xi_{N-1}]'$ с плотностью (2.17). Перепишем (2.11) с учетом обозначений (3.3) в виде

$$r_n = \int_{-1/2}^{1/2} \exp(i \cdot 2\pi nf)\, s(f)\, df, \quad n = 0,1\dots, N-1, \tag{4.1}$$

и перейдем в $\varphi(\cdot|r)$ от совокупности коэффициентов ковариации $r \equiv [r_0, r_1, \dots, r_{N-1}]'$ к спектральной плотности процесса. Теперь в качестве неизвестного вектора выступает совокупность значений $s \equiv [s_1, \dots, s_F]'$ спектральной плотности $s(f)$ на выбранной подходящим образом сетке частот, а плотность распределения $\varphi(\cdot|r)$ может рассматриваться как заданная функция вектора s. Таким образом, мы имеем здесь стандартную задачу многомерного параметрического оценивания s по выборке y. Важно подчеркнуть, что для нахождения строгих теоретических ограничений на оценки спектральной плотности достаточно соотношений (4.1). Именно на них и будет базироваться последующий анализ.

Вместе с тем, при нахождении явного выражения для оценки плотности логически строгий подход, связывающий s^* непосредственно с выборкой y, приводит к сложным нелинейным уравнениям. Упростить задачу позволяет следующее соображение: по сути дела мы сталкиваемся здесь с вложенными одна в другую обратными задачами, из которых внешняя заключается в нахождении эффективной оценки $c(y)$ совокупности r, а внутренняя - в оценивании спектральной плотности по случайной выборке коэффициентов ковариации. Поскольку качество решения связано прежде всего с неустойчивостью обратной задачи, можно разбить решение на два этапа: сначала строится подходящая оценка выборочных коэффициентов ковариации $c \equiv [c_0, c_1, \dots, c_{N-1}]'$, а затем обращается (4.1).

Уже без учета статистического характера обратной задачи нетрудно убедиться в неоднозначности ее решения. В самом деле, соотношения (4.1), рассматриваемые как частный случай известной *проблемы моментов*, не позволяют полностью восстановить $s(f)$ на основании лишь конечной совокупности r (так, мы могли бы существенно варьировать высокочастотный хвост $s(f)$, сохраняя практически неизменной совокупность первых коэффициентов ковариации). Для однозначного восстановления требуется обширная априорная информация относительно $s(f)$. Неустойчивость задачи становится тем более очевидной, если мы примем во внимание, что точные значения коэффициентов ковариации r_0, r_1,..., r_{N-1} на опыте неизвестны; в нашем распоряжении имеется лишь случайная выборка y, из элементов которой могут быть построены различные оценки

$c_0(y)$, $c_1(y),\dots,c_{N-1}(y)$ истинных коэффициентов ковариации. Характер обратной задачи (4.1) свидетельствует о том, что неточности последних весьма существенно сказываются на качестве оценивания спектральной плотности.

5. *Энтропия*. Согласно Шеннону [7], *информация* $J(\xi|r)$, связанная со случайной величиной ξ, определяется следующим образом:

$$J(\xi|r) = -\ln\varphi(\xi|r), \tag{5.1}$$

а ее среднее значение есть *энтропия* $\mathcal{E}(r)$:

$$\mathcal{E}(r) = \langle J(\xi|r) \rangle. \tag{5.2}$$

Подставляя в (5.1) выражение (2.17), находим для данного случая:

$$J(\xi|r) = \frac{1}{2}\left[N\ln(2\pi) + \ln D(r) + \xi' R^{-1}\xi\right], \tag{5.3}$$

где $D(r) = \det(R)$. Как известно (см. [8], теорема 3.3.3), если x - N- мерный гауссов случайный вектор с нулевым средним и ковариационой матрицей R, то скалярная случайная величина $\zeta_N \equiv \xi' R^{-1}\xi$ подчиняется распределению χ^2_n. Отсюда следует, что среднее значение и дисперсия ζ_N равны

$$\langle \xi' R^{-1}\xi \rangle = N, \quad \mathrm{Var}\left(\xi' R^{-1}\xi\right) = 2N. \tag{5.4}$$

Усредняя (5.3) с учетом (5.4), получаем для энтропии:

$$\mathcal{E}(r) = \frac{1}{2}\left[N\ln(2\pi e) + \ln D(r)\right]. \tag{5.5}$$

Со времени работы Хотеллинга [9] известно, что возможность компактизации информации, связанной с многомерной случайной величиной, зависит прежде всего от вида спектра ее ковариационной матрицы. Пусть $\alpha \equiv \left[\alpha_0,\alpha_1,\dots\alpha_{N-1}\right]'$ - вектор собственных чисел матрицы R. Поскольку ковариационная матрица предполагается положительно определенной, мы имеем $\alpha_j>0$ для всех j, а определитель (см., например, [10])

$$D = \prod_{j=0}^{N-1}\alpha_j. \tag{5.6}$$

Из двух последних формул следует:

$$\mathcal{E}(r) = \frac{1}{2}\sum_{j=0}^{N-1}\ln\left(2\pi e\alpha_j\right). \tag{5.7}$$

Основной вклад в энтропию вносят наибольшие собственные числа ковариационной матрицы, однако даже в типичном для практики случае, когда диапазон значений α_i весьма широк, неравноценность вклада смягчается логарифмической зависимостью. Далее предполагается для удобства, что совокупность $\{\alpha_i\}$ пронумерована в порядке убывния собственных чисел.

Крымская лаборатория Гос. астрономического института
им. П.К. Штернберга.

SIMILARITY LAW IN A TIME SERIES SPECTRAL ESTIMATION. II

V.Yu.TEREBIZH

In the second part of investigation (see [1]) we give auxiliary results concerning the autoregression process of first order. A strict formulaion of the spectral estimation problem as an inverse problem of mathematical physics is considered.

ЛИТЕРАТУРА

1. *В.Ю.Теребиж,* Астрофизика, 40, 139, 1997 (часть I данной серии).
2. *T.W.Anderson,* The Statistical Analysis of Time Series, Wiley, New York, 1971; Т.Андерсон, Статистический анализ временных рядов, Мир, М., 1976.
3. *В.Ю.Теребиж,* Анализ временны'х рядов в астрофизике, Наука, М., 1992.
4. *M.G.Kendall, A.Stuart,* The Advanced Theory of Statistics, Vols. 2, 3, Griffin, London, 1969; *М.Кендалл, А.Стьюарт,* Статистические выводы и связи, Наука, М., 1973; Многомерный статистический анализ и временные ряды, Наука, М., 1976.
5. *J.P.Burg,* Paper presented at the 37-th Ann. Int. Meeting, Oklahoma City, 1967.
6. *А.А.Боровков,* Математическая статистика, Наука, М., 1984.
7. *C.Shannon,* C.Bell Syst. Techn. J., 27, 379, 623, 1948.

8. *T.W.Anderson,* An Introduction to Multivariate Statistical Analysis, Wiley, New York, 1957; *Т.Андерсон*, Введение в многомерный статистический анализ, ФМ, М., 1963.
9. *H.Hotelling,* Journ. Educ. Psych., 24, 417, 498, 1933.
10. *R.Bellman,* Intruduction to Matrix Analysis, McGrow-Hill, New York, 1960; *Р.Беллман*, Введение в теорию матриц, Наука, М., 1969.