Доклады Академии наук СССР
1974. Том 217, № 3

УДК 519.242.34 *МАТЕМАТИКА*

К. О. ДЖАПАРИДЗЕ, А. М. ЯГЛОМ

ПРИМЕНЕНИЯ МОДИФИЦИРОВАННОГО «МЕТОДА НАКОПЛЕНИЯ» ФИШЕРА К ОЦЕНКЕ ПАРАМЕТРОВ СПЕКТРА СЛУЧАЙНЫХ ПРОЦЕССОВ

(Представлено академиком А. Н. Колмогоровым 28 XII 1973)

Известно, что «уравнения правдоподобия», решением которых является оценка максимального правдоподобия $\theta^* \equiv (\theta_1^*, \ldots, \theta_p^*)$ (здесь и ниже знак $\equiv$ означает «равно по определению») неизвестного значения θ_0 p-мерного параметра θ, входящего в выражение для плотности вероятности случайной величины, часто оказываются очень сложными. В связи с этим Фишер [1] (см. также [2]) разработал приближенный метод накопления (method of scoring), позволяющий, начиная с какой-то более грубой оценки $\tilde{\theta}$, приблизиться к оценке θ^*. Метод Фишера фактически представляет собой небольшую модификацию классического «метода касательной» Ньютона (ср. [3], стр. 154), от которого он отличается лишь тем, что матрица производных от левых частей уравнений правдоподобия по параметрам θ_i здесь заменяется своей состоятельной оценкой $-\mathcal{I}(\theta)$ ($\mathcal{I}$ — информационная матрица Фишера). Позже Лекам [4] отметил, что при широких условиях уже первое приближение «метода накопления» оказывается асимптотически эффективной и асимптотически нормальной оценкой θ_0 (ср. также [1], стр. 722).

Как было указано в [5], аналогичный прием приложим и ко многим другим задачам о нахождении статистических оценок. Пусть $\mathbf{x}_T$ (T пробегает неограниченное множество вещественных значений) — выборочное значение случайной величины $\mathbf{X}_T$ в пространстве $\mathfrak{X}_T$, распределение вероятностей которой при всех T известно с точностью до значения θ_0 параметра $\theta \equiv (\theta_1, \ldots, \theta_p)$. Предположим, что при определенных условиях уравнения $l_j(\theta; \mathbf{x}_T) = 0$, $j = 1, \ldots, p$, где l_j — гладкие функции θ, зависящие от $\mathbf{x}_T$, имеют решение $\theta^* = \theta^*(\mathbf{x}_T)$ такое, что $T^\alpha(\theta^* - \theta_0)$, где $\alpha > 0$, при $T \to \infty$ имеет предельное распределение с ограниченными моментами первых двух порядков. Пусть далее $\tilde{\theta}$ — это T^β-состоятельная оценка θ_0 (т. е. $T^\beta(\tilde{\theta}_j - \theta_{0j})$ ограничено по вероятности при всех $j = 1, 2, \ldots, p$), где $\beta > \alpha/2$, а $\mathcal{D}(\mathbf{x}_T) \equiv \|D_{jk}(\mathbf{x}_T)\|$ — состоятельная оценка матрицы $\mathcal{F}(\theta;\ \mathbf{x}_T) \equiv \|F_{jk}(\theta;\ \mathbf{x}_T)\| \equiv \|\partial l_j(\theta;\ \mathbf{x}_T)/\partial\theta_k\|$ (невырожденной при $\theta = \theta_0$ с вероятностью 1) такая, что $T^{\alpha/2}|D_{jk}(\mathbf{X}_T) - F_{jk}(\tilde{\theta}; \mathbf{X}_T)| \xrightarrow{p} 0$ при $T \to \infty$ для всех j и k ($\xrightarrow{p}$ означает «стремится по вероятности»).

Положим

$$\hat{\theta} = \tilde{\theta} - \mathcal{D}^{-1}(\mathbf{x}_T)\mathbf{l}(\tilde{\theta};\ \mathbf{x}_T), \qquad \mathbf{l} = (l_1, \ldots, l_p); \tag{1}$$

тогда, используя квадратичную сходимость метода Ньютона, нетрудно доказать, что $T^\alpha(\hat{\theta} - \theta_0)$ и $T^\alpha(\theta^* - \theta_0)$ при $T \to \infty$ будут иметь одинаковое предельное распределение (т. е. оценка $\hat{\theta}$ асимптотически будет столь же хорошей, как и θ^*).

Пусть теперь $X(t)$ — стационарный случайный процесс с дискретным (целочисленным) или непрерывным t и $EX(t) = 0$, имеющий спектральную плотность $f(\lambda; \theta)$, зависящую от $\theta \equiv (\theta_1, \ldots, \theta_p)$, а $\mathbf{x}_T \equiv (x(1), \ldots, x(T))$

или $\mathbf{x}_T \equiv (x(t), 0 \leqslant t \leqslant T)$. Рассмотрим уравнения

$$l_j(\theta; \mathbf{x}_T) \equiv -\frac{1}{4\pi}\int\limits_{\Lambda} \frac{f(\lambda;\theta)-I_T(\lambda)}{[f(\lambda;\theta)]^2}\frac{\partial f(\lambda;\theta)}{\partial\theta_j}d\lambda = 0, \quad j=1,\ldots,p, \qquad (2)$$

где

$$I_T(\lambda) \equiv \frac{1}{2\pi T}\left|\sum_{t=1}^{T} x(t)e^{it\lambda}\right|^2 \quad \text{или} \quad \equiv \frac{1}{2\pi T}\left|\int\limits_0^T x(t)e^{it\lambda}dt\right|^2,$$

а $\Lambda \equiv [-\pi, \pi]$ или $\equiv (-\infty, \infty)$ соответственно. Как известно (см. ([6–9])), уравнения (2) при широких условиях регулярности имеют решение $\theta^* \equiv (\theta_1^*, \ldots, \theta_p^*)$, которое для гауссовского $X(t)$ является асимптотически нормальной и асимптотически эффективной оценкой θ_0 такой, что

$$\|\lim_{T\to\infty} TE(\theta_j^*-\theta_{0j})(\theta_k^*-\theta_{0k})\| =$$

$$= \left\|\frac{1}{4\pi}\int\limits_{\Lambda}\frac{\partial f(\lambda;\theta_0)}{\partial\theta_{0j}}\frac{\partial f(\lambda;\theta_0)}{\partial\theta_{0k}}\frac{d\lambda}{[f(\lambda;\theta_0)]^2}\right\|^{-1} \equiv [\mathcal{I}(\theta_0)]^{-1}. \qquad (3)$$

Последнее равенство и асимптотическая нормальность θ^* будут иметь место также и для более общих линейных процессов $X(t)$ с дискретным временем и конечным четвертым семиинвариантом, если только от θ зависит лишь отношение $f(\lambda)/\sigma^2 = \varphi(\lambda; \theta)$, где σ^2 — средний квадрат ошибки наилучшего линейного прогноза процесса $X(t)$ на единицу времени вперед ([6–8]). При этом во всех указанных случаях легко доказать, что при широких условиях регулярности

$$T^{1/4}\left|\frac{\partial l_j(\tilde\theta; \mathbf{X}_T)}{\partial\theta_k} + I_{jk}(\theta_0)\right| \xrightarrow{P} 0, \quad T\to\infty, \qquad (4)$$

где $I_{jk}(\theta_0)$ — это (j, k)-й элемент $\mathcal{I}(\theta_0)$, а $\tilde\theta$ — произвольная $T^{1/4+\varepsilon}$-состоятельная оценка θ_0, $\varepsilon > 0$ (ср., например, ([10]), теорема 1).

Ниже всегда будет предполагаться, что речь идет о процессах $X(t)$, для которых верны приведенные выше результаты об оценках θ^* и что выполняются все нужные для этого (и для справедливости результата, касающегося оценки $\hat\theta$ формулы (1)) условия регулярности; в частности, рассматриваемые состоятельные оценки (обозначаемые знаком ~ сверху) всегда будут считаться $T^{1/4+\varepsilon}$-состоятельными, где $\varepsilon > 0$. Оценки параметров спектра, обладающие теми же асимптотическими свойствами, что и решение θ^* уравнений (1), будут, как это часто делается, называться *асимптотически эффективными*.

1. Воспользуемся формулой $f(\lambda; \hat\theta) = f(\lambda; \tilde\theta) + \sum_{k=1}^{p}[\partial f(\lambda; \theta')/\partial\theta_k'](\hat\theta_k - \tilde\theta_k)$, где $\theta_k' \in [\theta_k, \hat\theta_k]$, и примем, что $D_{jk}(\mathbf{x}_T)$ получается из формулы для $-I_{jk}(\theta_0)$ с помощью замены значения θ_0 на $\tilde\theta$ в функции $f(\lambda; \theta_0)$ в знаменателе и в первой частной производной этой функции под знаком интеграла и на θ' во второй частной производной функции $f(\lambda; \theta_0)$. Тогда (1) можно будет переписать в виде системы p уравнений, отличающихся от (2) лишь тем, что функция $f(\lambda; \theta)$ в числителе теперь заменяется на $f(\lambda; \hat\theta)$, а в той же функции в знаменателе и в ее производной неизвестное θ заменяется на $\tilde\theta$. Таким образом, при отыскании асимптотически эффективных оценок θ_0, в уравнениях (2) можно заменить неизвестное θ в знаменателе и под знаком производной произвольной состоятельной оценкой $\tilde\theta$. Иному доказательству этого факта посвящены работы ([11], [12]); для частного случая дискрет-

ного t и функции $f(\lambda;\theta)$, линейно зависящей от θ, указанный факт составляет основное содержание работы ([13]) *.

2. Пусть $f(\lambda)=\sigma^2|g(e^{i\lambda})|^2/(2\pi)$, $-\pi\leqslant\lambda\leqslant\pi$, где $g(z)=1+\alpha_1 z+\ldots+\alpha_p z^p$ имеет лишь корни z_i с $|z_i|>1$; предположим, что неизвестны значения $\boldsymbol{\alpha}=(\alpha_1,\ldots,\alpha_p)$ (и, быть может, σ^2). Используя снова (1) и (2) и приняв за $\mathscr{D}(\mathbf{x}_T)$ матрицу $-\mathscr{I}(\tilde{\boldsymbol{\alpha}})$, получаем следующую систему уравнений, определяющую асимптотически эффективные оценки $\hat{\alpha}=(\hat{\alpha}_1,\ldots,\hat{\alpha}_p)$:

$$\sum_{k=1}^{p}\int_{-\pi}^{\pi}\frac{\cos(k-j)\lambda\,d\lambda}{f(\lambda;\tilde{\boldsymbol{\alpha}})}[\hat{\alpha}_k-\tilde{\alpha}_k]=\int_{-\pi}^{\pi}\frac{I_T(\lambda)\sum\limits_{k=0}^{p}\tilde{\alpha}_k\cos(k-j)\lambda\,d\lambda}{[f(\lambda;\tilde{\boldsymbol{\alpha}})]^2},\tag{5}$$

где $j=1,\ldots,p$, $\tilde{\boldsymbol{\alpha}}=(\tilde{\alpha}_1,\ldots,\tilde{\alpha}_p)$ — состоятельная оценка $\boldsymbol{\alpha}$, а $\tilde{\alpha}_0=1$. Если же мы примем за $D_{jk}(\mathbf{x}_T)$ линейный функционал от $I_T(\lambda)$, получающийся при подстановке в $-I_{jk}(\boldsymbol{\alpha})$ отношения $I_T(\lambda)/[f(\lambda;\tilde{\boldsymbol{\alpha}})]^2$ вместо $[f(\lambda;\tilde{\boldsymbol{\alpha}})]^{-1}$, то придем к уравнениям

$$\int_{-\pi}^{\pi}\frac{I_T(\lambda)\,[\cos j\lambda+\sum\limits_{k=1}^{p}(2\tilde{\alpha}_k-\hat{\alpha}_k)\cos(j-k)\lambda]\,d\lambda}{[f(\lambda;\tilde{\boldsymbol{\alpha}})]^2}=0,\quad j=1,\ldots,p,\tag{6}$$

предполагавшимся, исходя из других соображений, в ([14], [15]).

3. Пусть теперь $f(\lambda)=\sigma^2|g(e^{i\lambda})|^2/2\pi|h(e^{i\lambda})|^2$, $-\pi\leqslant\lambda\leqslant\pi$, где $g(z)$ — то же, что и выше, $h(z)=1+\beta_1 z+\ldots+\beta_q z^q$, все корни $h(z)$ превосходят по модулю единицу, а неизвестны параметры $\boldsymbol{\alpha}=(\alpha_1,\ldots,\alpha_p)$, $\boldsymbol{\beta}=(\beta_1,\ldots,\beta_q)$ (и, быть может, σ^2). Здесь уравнения (2) удобно записать в виде

$$l_j(\boldsymbol{\alpha},\boldsymbol{\beta};\mathbf{x}_T)=0,\quad j=1,\ldots,p;\qquad l_{p+k}(\boldsymbol{\alpha},\boldsymbol{\beta};\mathbf{x}_T)=0,\quad k=1,\ldots,q,\tag{7}$$

где последние q уравнений (содержащие производные от $f(\lambda)$ по параметрам $\beta_1,\ldots,\beta_q$) — линейные по $\beta_1,\ldots,\beta_q$. Выберем какие-то состоятельные оценки $\tilde{\boldsymbol{\alpha}}=(\tilde{\alpha}_1,\ldots,\tilde{\alpha}_p)$ и определим $\tilde{\boldsymbol{\beta}}=(\tilde{\beta}_1,\ldots,\tilde{\beta}_q)$ как решения последних q уравнений (7), в которых $\boldsymbol{\alpha}$ заменено на $\tilde{\boldsymbol{\alpha}}$. Представим матрицу $\mathscr{I}(\theta)=\mathscr{I}(\boldsymbol{\alpha},\boldsymbol{\beta})$ порядка $p+q$ в виде «клеточной матрицы» $\begin{Vmatrix}\Phi & -\Omega' \\ -\Omega & \Psi\end{Vmatrix}$, где Φ, Ψ, Ω и Ω' — это $(p\times p)$-, $(q\times q)$-, $(q\times p)$- и $(p\times q)$-матрицы, явный вид которых выписан в ([15]), стр. 380 и 392. Так как $\tilde{l}_i=l_i(\tilde{\boldsymbol{\alpha}},\tilde{\boldsymbol{\beta}};\mathbf{x}_T)=0$ при $i=p+1,\ldots,p+k$, то используя (1), (2) и (4), легко найдем, что

$$\hat{\boldsymbol{\alpha}}=\tilde{\boldsymbol{\alpha}}+[\mathscr{E}_p-\tilde{\Phi}^{-1}\tilde{\Omega}'\tilde{\Psi}^{-1}\tilde{\Omega}]^{-1}\tilde{\Phi}^{-1}\tilde{\mathbf{l}}_1\tag{8}$$

(где $\mathscr{E}_p$ — единичная $(p\times p)$-матрица, $\tilde{\Phi}$, $\tilde{\Psi}$, $\tilde{\Omega}$ и $\tilde{\Omega}'$ — состоятельные оценки Φ, Ψ, Ω и Ω', а $\mathbf{l}_1=(l_1,\ldots,l_p)$) будет асимптотически эффективной оценкой $\boldsymbol{\alpha}$. Подставив $\hat{\boldsymbol{\alpha}}$ в последние q уравнений (7) и разрешив получающуюся линейную систему относительно $\beta_1,\ldots,\beta_q$, найдем асимптотически эффективную оценку $\hat{\boldsymbol{\beta}}$ вектора $\boldsymbol{\beta}$. Далее, если определить $\tilde{\Phi}$ как результат подстановки в формулу для Φ вместо $|g(e^{i\lambda})|^{-2}$ комбинации $\tilde{\sigma}^{-2}I_T(\lambda)|\tilde{h}\cdot(e^{i\lambda})|^2|\tilde{g}(e^{i\lambda})|^{-4}$, где $\tilde{h}(z)=1+\tilde{\alpha}_1 z+\ldots+\tilde{\alpha}_p z^p$, $\tilde{g}(z)=1+\tilde{\beta}_1 z+\ldots+\tilde{\beta}_q z^q$, а через $\boldsymbol{\alpha}^{(1)}$ обозначить решение первых p уравнений (7), в которых в произведении $|h(e^{i\lambda})|^2|g(e^{i\lambda})|^{-4}$ под знаком интеграла $h(e^{i\lambda})$ заменено на $\tilde{h}(e^{i\lambda})$, а $g(e^{i\lambda})$ — на $\tilde{g}(e^{i\lambda})$ (так что уравнения уже оказываются линейными

---

* В ([13]), а также и в ([14–16]), вместо интегралов по λ от $-\pi$ до π используются соответствующие римановы суммы, отвечающие разбиению $\Lambda=[-\pi,\pi]$ на T равных частей. Вопрос о законности такой замены специально исследовался в ([13]); здесь мы на нем не будем задерживаться.

по $\alpha_1, \ldots, \alpha_p$), то, как нетрудно проверить, $\widetilde{\Phi}(\widetilde{\alpha}-\widetilde{\alpha}^{(1)})=\widetilde{\mathrm{l}}_1$; при этом (8) обращается в формулу Хеннана ([14, 15])

$$\hat{\alpha}=\widetilde{\alpha}+[\mathscr{E}_p-\widetilde{\Phi}^{-1}\widetilde{\Omega}'\widetilde{\Psi}^{-1}\widetilde{\Omega}]^{-1}(\widetilde{\alpha}-\widetilde{\alpha}^{(1)}) \tag{9}$$

(использовавшего некоторые специальные оценки $\widetilde{\alpha}$ и $\widetilde{\Psi}$).

4. Пусть $f(\lambda)=b(\lambda)/|h(e^{i\lambda})|^2$, где $h(z)$ — то же, что и выше, а $b(\lambda)=\alpha_1 W_1(\lambda)+\ldots+\alpha_p W_p(\lambda)$, $W_k(\lambda)$ — известные функции, а неизвестны параметры $\alpha\equiv(\alpha_1, \ldots, \alpha_p)$, $\beta\equiv(\beta_1, \ldots, \beta_q)$. Записав уравнения (2) в виде (7), мы снова найдем, что последние q из них линейны по $\beta_1, \ldots, \beta_q$ и поэтому опять можем, выбрав оценки $\widetilde{\alpha}\equiv(\widetilde{\alpha}_1, \ldots, \widetilde{\alpha}_p)$, определить затем $\widetilde{\beta}\equiv(\widetilde{\beta}_1, \ldots \ldots \widetilde{\beta}_q)$ как решения этих q уравнений, в которых α заменено на $\widetilde{\alpha}$.

Представим снова $\mathscr{I}(\alpha, \beta)$ в виде «клеточной матрицы» и обозначим ее «клетки» так же, как в примере 3 (явный вид соответствующих матриц Φ, Ψ, Ω и Ω' указан в ([16])); тогда по тем же причинам, что и выше, формула (8) и в этом случае будет задавать асимптотически эффективную оценку $\hat{\alpha}$ параметров α. Примем теперь, что $\widetilde{\Phi}=\Phi(\widetilde{\beta})$ (от α матрица Φ вообще не зависит), и определим $\widetilde{\alpha}^{(1)}$ как решение линейной системы первых p уравнений (7), в которых h заменено на $\widetilde{h}$, а α в знаменателе подынтегрального выражения — на $\widetilde{\alpha}$; тогда опять $\widetilde{\Phi}^{-1}\widetilde{\mathrm{l}}_1=\widetilde{\alpha}-\widetilde{\alpha}^{(1)}$, так что (8) и здесь можно переписать в виде (9). Последний результат представляет собой упрощение и обобщение основного результата работы ([16]).

5. Совершенно аналогичный подход применим и в случае непрерывного времени t, если $f(\lambda)=K|A(i\lambda)|^2/(2\pi|B(i\lambda)|^2)$, $-\infty<\lambda<\infty$, где $A(i\lambda)\equiv(i\lambda)^p+\alpha_1(i\lambda)^{p-1}+\ldots+\alpha_p$, $B(i\lambda)\equiv(i\lambda)^q+\beta_1(i\lambda)^{q-1}+\ldots+\beta_q$, и все корни $A(z)$ и $B(z)$ имеют отрицательные мнимые части. Если неизвестны параметры $\alpha\equiv(\alpha_1, \ldots, \alpha_p)$ и $\beta\equiv(\beta_1, \ldots, \beta_q)$, то записав (2) в виде (7), мы снова получим, что последние q из этих уравнений линейны по $\beta_1, \ldots, \beta_q$. Если, как и выше, $\widetilde{\alpha}$ — произвольная состоятельная оценка α, а $\widetilde{\beta}$ — решение указанных q уравнений, в которых α заменено на $\widetilde{\alpha}$, то равенство (8), где Φ, Ψ, Ω и Ω' — это «клетки» матрицы $\mathscr{I}(\alpha, \beta)$ (явный вид которых легко выписать), снова будет определять асимптотически эффективную оценку $\hat{\alpha}$ вектора α. Выбрав матрицу $\widetilde{\Phi}=\Phi(\widetilde{\alpha}, \widetilde{\beta})$ так, чтобы она представляла собой линейный функционал от периодограммы $I_T(\lambda)$ (это условие однозначно определяет $\widetilde{\Phi}$), мы снова получим, что $\widetilde{\Phi}^{-1}\widetilde{\mathrm{l}}_1=\widetilde{\alpha}-\widetilde{\alpha}^{(1)}$, где $\widetilde{\alpha}^{(1)}$ определяется вполне аналогично тому, как это делалось в примерах 3 и 4. Поэтому и в этом случае равенство (8) можно переписать в виде (9).

Тбилисский государственный университет

Институт физики атмосферы
Академии наук СССР
Москва

Поступило
25 XII 1973

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

[1] *R. A. Fisher* Proc. Cambr. Phil. Soc., v. 22, 700 (1925). [2] *С. Р. Рао*, Линейные статистические методы и их применения, М., 1968, § 5g. [3] *И. С. Березин, Н. П. Жидков*, Методы вычислений, т. 2, М., 1959. [4] *L. Le Cam*, Proc. 3 Berkeley Symp. Math. Stat. Probl., v. 1, 1956, p. 129. [5] *К. О. Джапаридзе*, Теор. вероятн. и ее примен., т. 19, № 2 (1974). [6] *P. Wittle*, Ark. Mat., B. 2, 423 (1953). [7] *P. Whittle*, Bull. Intern. Stat. Inst., v. 39, 105 (1962). [8] *A. M. Walker*, J. Austr. Math. Soc., c. 4, 363 (1964). [9] *К. О. Джапаридзе*, Теор. вероятн. и ее примен., т. 15, 548 (1970). [10] *U. Grenender. M. Rosenblatt*, Ann. Math. Stat., v. 24, 537 (1953). [11] *К. А. Джапаридзе*, Теор. вероятн. и ее примен., т. 16, 563 (1971). [12] *К. О. Джапаридзе*, Там же, т. 19, 120 (1974). [13] *M. L. Clevenson*, Techn. Rep. No. 15, Statistics Dept., Stanford Univ., Stanford, 1970. [14] *E. J. Hannan*, Biometrika, v. 56, 579 (1969). [15] *E. J. Hannan*, Multiple Time Series, N. Y., 1970. [16] *E. Parzen*, Bull. Intern. Stat. Inst., v. 44, Book 2, 1971, p. 315.