

УДК 519.226.3

В. ОЛЬМАН

СОХРАНЕНИЕ БАЙЕСОВСКОГО СВОЙСТВА ОТНОСИТЕЛЬНО ДИСКРЕТНЫХ МЕР ПРИ ПРЕОБРАЗОВАНИЯХ

(Представил Н. Алумяэ)

В теории статистического оценивания часто возникает необходимость оценивать не сам неизвестный параметр распределения θ , а некоторую функцию $f(\theta)$ от него. Так, например, в теории надежности обычно требуется оценивать по наблюдениям не интенсивность отказов θ , а вероятность безотказной работы в течение времени T , т. е. величину $e^{-\theta T}$. Естественно, возникает вопрос: для каких преобразований f некоторое заданное статистическое свойство оценки $\hat{\theta}$ параметра θ сохраняется при оценивании $f(\theta)$ с помощью оценки $\hat{f}(\theta)$? Известно, например, что если $\hat{\theta}$ — оценка максимального правдоподобия (ОМП), то для широкого класса функций $\hat{f}(\hat{\theta})$ является ОМП величины $f(\theta)$. В настоящем сообщении исследуется сохранение байесовского свойства относительно дискретных мер и показано, что только линейные преобразования сохраняют байесовость.

Рассмотрим задачу оценивания функции $f(\theta)$, $\theta \in R^p$, $p \geq 1$, по наблюдению $x \in X$ при условии, что плотность $p(x, \theta)$ распределения x относительно некоторой меры $d\mu(x)$ известна с точностью до параметра θ и что $f: R^p \rightarrow R^r$, $r \geq 1$. Неизвестный параметр θ является случайным независимым от x вектором, подчиненным некоторому вероятностному распределению G . При квадратичной потере риск произвольной оценки $\tilde{\theta}(x)$ определяется интегралом

$$\int_{\theta \in R^p} \int_{x \in X} (\tilde{\theta}(x) - \theta)^2 p(x, \theta) d\mu(x) dG(\theta)$$

и минимизирующая этот функционал байесовская оценка $\theta_G(x)$ имеет вид

$$\theta_G(x) = \int_{\theta \in R^p} \theta p(x, \theta) dG(\theta) / \int_{\theta \in R^p} p(x, \theta) dG(\theta).$$

Определение. Назовем функцию f сохраняющей байесовость оценки θ_G , если статистика $f(\theta_G(x))$ является байесовской оценкой $f(\theta)$ относительно распределения G .

В соответствии с определением f сохраняет байесовость θ_G , если для почти всех x по мере $d\mu(x)$ имеет место равенство

$$f(\theta_G(x)) = \int_{\theta \in R^p} f(\theta) p(x, \theta) dG(\theta) / \int_{\theta \in R^p} p(x, \theta) dG(\theta). \quad (1)$$

Пусть распределение G сосредоточено на конечном множестве векторов $\Theta = \{\theta_1, \theta_2, \dots, \theta_n\}$, $\theta_i \in R^p$, с нагрузками p_1, p_2, \dots, p_n , $n \geq 1$. Тогда равенство (1) имеет вид

$$f \left(\frac{\sum_{i=1}^n \theta_i p(x, \theta_i) p_i}{\sum_{i=1}^n p(x, \theta_i) p_i} \right) = \frac{\sum_{i=1}^n f(\theta_i) p(x, \theta_i) p_i}{\sum_{i=1}^n p(x, \theta_i) p_i}.$$

Или, обозначив $q_i(x) = p(x, \theta_i) p_i / \sum_{i=1}^n p_i p(x, \theta_i)$, имеем

$$f \left(\sum_{i=1}^n \theta_i q_i(x) \right) = \sum_{i=1}^n f(\theta_i) q_i(x). \quad (2)$$

Пусть $\theta_G(X) = \Theta_0$ — множество значений статистики θ_G . Произвол в значениях $f(\theta_i)$, $i=1, 2, \dots, n$, удовлетворяющих уравнению (2), определяется пересечением $\Theta_0 \cap \Theta$. В частности, если $\Theta_0 \cap \Theta = \emptyset$, то $f(\theta_i)$, $i=1, 2, \dots, n$, могут быть произвольными, и в точках $\theta_G(x) \neq \theta_j$, $x \in X$, $j=1, 2, \dots, n$, значения $f(\theta_G(x))$ задаются равенством (2). Рассмотрим случай $\Theta_0 \cap \Theta \neq \emptyset$ и перенумеруем векторы $\theta_1, \dots, \theta_n$ так, что

$$\Theta_0 \cap \Theta = \{\theta_1, \theta_2, \dots, \theta_k\}, \quad 1 \leq k \leq n.$$

Введем обозначения: F — матрица $r \times n$, $F = (f(\theta_1) \dots f(\theta_n))$, Q — матрица $p \times n$, $Q = (\theta_1 \dots \theta_n)$, $q^T(x) = (q_1(x), \dots, q_n(x))$, l_j — n -мерный столбец с 1 на j -м месте, и 0 на всех остальных местах, $j=1, 2, \dots, n$, $l = \sum_{j=1}^n l_j$. Пусть

$$X_j = \{x \in X : \theta_G(x) = \theta_j\}, \quad j=1, 2, \dots, k.$$

Тогда в матричном виде

$$Q(q(x) - l_j) = 0, \quad x \in X_j, \quad j=1, 2, \dots, k,$$

и, кроме того, по определению $l^T(q(x) - l_j) = 0$, $1 \leq j \leq n$, и следовательно,

$$\tilde{Q}(q(x) - l_j) = 0, \quad x \in X_j, \quad j=1, 2, \dots, k,$$

где $\tilde{Q} = (Q^T l)^T$.

Образует линейное подпространство L , натянутое на векторы $\{q(x) - l_j, x \in X_j\}_{j=1}^k$. Пусть d — размерность подпространства L , а \tilde{L} — матрица порядка $d \times n$, строки которой являются базисом L , т. е. $L = \mathcal{L}(\tilde{L})$. Очевидно, что $d + v \leq n$, где $v = R(\tilde{Q})$ — ранг матрицы \tilde{Q} . Пусть Q_1 матрица такая, что $\tilde{Q}Q_1^T = 0$, $L_1Q_1^T = 0$ и $R(Q_1) = n - d - v$, т. е. $\mathcal{L}(Q_1) = \mathcal{L}^\perp(\tilde{L}) \cap \mathcal{L}^\perp(\tilde{Q})$, где \mathcal{L}^\perp обозначает ортогональное дополнение к соответствующему подпространству. Но из равенства (2) имеем

$$F(q(x) - l_j) = 0, \quad x \in X_j, \quad j=1, 2, \dots, k,$$

а следовательно, $\mathcal{L}(F) \subset \mathcal{L}^\perp(\tilde{L})$, и получаем, что

$$F = AQ + be^T + BQ_1, \quad (3)$$

где A и B произвольные матрицы порядков $r \times p$ и $r \times R(Q_1)$ соответственно, а b — произвольный r -мерный вектор. Перепишем (3) в виде

$$f(\theta_i) = A\theta_i + b + B\eta_i, \quad i=1, 2, \dots, n, \quad (4)$$

где η_1, \dots, η_n — столбцы матрицы Q_1 , получаем следующий результат.

Теорема. Все решения уравнения (2) относительно функции $f(\cdot)$ имеют вид

$$f(u) = Au + b + B \sum_{i=1}^n \eta_i q_i(x), \quad x \in X,$$

где $u \in \Theta_0$ и $u = \sum_{i=1}^n \theta_i q_i(x)$, $x \in X$.

Следствие. Если $d+v=n$, то

$$f(u) = Au + b, \quad u \in \Theta_0.$$

В качестве примера рассмотрим оценивание функции f от среднего θ в схеме независимых одномерных нормальных наблюдений. Другими словами, в соответствии с обозначениями $p=r=1$, $X=R^h$, $x = (x_1, \dots, x_h)$, $p(x, \theta) = (\sqrt{2\pi}\sigma)^{-h} \exp[-\sum (x_i - \theta)^2/2\sigma^2]$, $\sigma_h^2 > 0$,

$d\mu(x) = dx$. В силу достаточности статистики $\bar{x} = \sum_{i=1}^h x_i/k$ можно принять $k=1$, а кроме того, не умаляя общности, положим $\sigma^2=1$. Пусть $\theta_1 < \theta_2 < \dots < \theta_n$ — точки сосредоточения распределения G , а p_1, \dots, p_n — соответствующие нагрузки. Тогда байесовская оценка $\theta_G(x)$ параметра θ имеет вид, $i=1, 2, \dots, n$,

$$\theta_G(x) = \sum_{i=1}^n \theta_i q_i(x), \quad q_i(x) = \exp[-(x - \theta_i)^2/2] / \sum_{j=1}^n \exp\left[-\frac{(x - \theta_j)^2}{2}\right].$$

Нетрудно убедиться, что θ_G возрастающая по $x \in R^1$ функция, причем область ее значений — открытый интервал (θ_1, θ_n) , и следовательно, $\Theta \cap \Theta_0 = (\theta_2, \dots, \theta_{n-1})$. Обозначим через $x_2 < \dots < x_{n-1}$ решение системы

$$\theta_G(x_j) = \theta_j, \quad j=2, \dots, n-1. \quad (5)$$

Очевидно, что в силу возрастания $\theta_G(x)$ система (5) имеет единственное решение относительно x_2, \dots, x_{n-1} , и следовательно, $X_j = x_j$, $j=2, \dots, n-1$.

Рассмотрим матрицу

$$A = \{q_i(x_j) - \delta_{ij}\}_{i=1, 2, \dots, n}^{j=2, 3, \dots, n-1}$$

и убедимся, что ее ранг равен $n-2$. По теореме Леви—Деспланка [1], квадратная подматрица $\{g_i(x_j) - \delta_{ij}\}$ $i, j=2, \dots, n-1$ невырождена, так как

$\sum_{i=2}^{n-1} q_i(x_j) < 1$, и следовательно, ранг матрицы A есть

$n-2$. Осталось заметить, что матрица

$$\begin{pmatrix} \theta_1 & \dots & \theta_n \\ 1, 1, & \dots, & 1 \end{pmatrix}$$

имеет ранг 2, и значит, выполнено условие следствия к теореме. Таким образом, только линейная функция сохраняет байесовское свойство относительно конечного дискретного априорного распределения при оценивании функции от среднего нормального закона.

ЛИТЕРАТУРА

1. Маркус М., Минк Х. Обзор по теории матриц и матричных неравенств. М., «Наука», 1972.

Институт кибернетики
Академии наук Эстонской ССР

Поступила в редакцию
4/X 1986

HINNANGUTE TEISENDUSE BAYESI OMADUSE SÄILIVUS APRIORORSETE DISKREETSETE MÕOTUDE KORRAL

Statistiliste hinnangute teorias tekib sageli vajadus hinnata mitte tundmatut jao-
tuse parameetrit θ ennast, vaid temast sõltuvat funktsiooni $f(\theta)$. Sellistes olukordades
on kasulik teada, kas parameetri θ hinnangu $\hat{\theta}$ mingi statistiline omadus säilib
 $f(\theta)$ hindamisel statistikuga $f(\hat{\theta})$. Artiklis on iseloomustatud selliseid teisendusi, mis
säilitavad hinnangu Bayesi omaduse lõplikus arvus punktides kontsentreeritud apriorse
diskreetse mõõdu suhtes ruutkaofunktsiooni korral. On näidatud, et ühemõõtmelise nor-
maaljaotuse keskvaartuse hindamisel sellised teisendused ammenduvad lineaarfunktsiooni
klassiga.

V. OLMAN

INVARIANCE OF BAYES PROPERTY SUBJECT TO DISCRETE MEASURE UNDER TRANSFORMATIONS

In many problems arising in the theory of statistical estimation the value necessary
to estimate is not directly the unknown parameter θ of the considered distribution but
some transformation $f(\theta)$ of it. The question considered in the paper may be, in the
most general way, stated as follows: for what kind of transformations f some given
statistical property of the estimator $\hat{\theta}$ of the parameter θ is preserved when
 $f(\theta)$ is estimated by $f(\hat{\theta})$. In this paper a Bayes property of estimator is in question
under quadratic risk subject to a priori discrete measure, concentrated in finite numbers
of points. The characterization of transformations which are invariant in respect to
Bayes property is obtained. It is shown that for normal case these transformations are
only linear.