Math-Net.Ru

Общероссийский математический портал

И. М. Арбеков, Г. В. Проскурин, О нижней границе вероятности правильной классификации наблюдений, *Теория вероятн. и ее примен.*, 1996, том 41, выпуск 3, 628–633

DOI: <https://doi.org/10.4213/tvp3142>

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением
<http://www.mathnet.ru/rus/agreement>

Параметры загрузки:

IP: 3.85.7.115

26 апреля 2023 г., 18:06:32

КРАТКИЕ СООБЩЕНИЯ

© 1996 г.

АРБЕКОВ И. М.*, ПРОСКУРИН Г. В.*

О НИЖНЕЙ ГРАНИЦЕ ВЕРОЯТНОСТИ
ПРАВИЛЬНОЙ КЛАССИФИКАЦИИ НАБЛЮДЕНИЙ

В работе изучается нижняя граница вероятности правильной классификации двух классов наблюдений, отличающихся сдвигом функции распределения. Для байесовской процедуры распознавания проводится минимизация функционала вероятности по всем распределениям с фиксированной дисперсией. Находятся асимптотически совпадающие при стремящемся к нулю значении параметра сдвига нижняя и верхняя оценки нижней границы вероятности правильной классификации. Полученные результаты обобщаются на многомерный случай.

Ключевые слова и фразы: классификация наблюдений, байесовская процедура.

Рассмотрим случайный вектор (X, ϵ) , где X (наблюдаемая компонента) принадлежит действительной прямой $\mathbf{R}^1$, ϵ (ненаблюдаемая компонента) принимает одно из двух возможных значений 0 или 1. Пусть заданы распределение вероятностей $P\{\epsilon = 0\} = P\{\epsilon = 1\} = \frac{1}{2}$ и условные плотности распределения вероятностей $f(x)$ и $f(x - \mu)$ случайной величины X при условиях $\epsilon = 0$ и $\epsilon = 1$ соответственно.

Процедура классификации состоит в том, чтобы по значению X принять то или иное решение о значении ϵ . Процедура классификации определяется как разбиение $D = (D_0, D_1)$ в $\mathbf{R}^1$ и дает оценку $\epsilon^* = j$ тогда и только тогда, когда $X \in D_j$, $j = 0, 1$. Средняя вероятность правильной классификации $P_D = P\{\epsilon^* = \epsilon\}$ определяется равенством

$$P_D = \frac{1}{2} P\{X \in D_0 \mid \epsilon = 0\} + \frac{1}{2} P\{X \in D_1 \mid \epsilon = 1\}.$$

Оптимальная байесовская процедура $B = (B_0, B_1)$, максимизирующая среднюю вероятность правильной классификации, для данного случая имеет вид [1]

$$B_0 = \{x \mid f(x) \geq f(x - \mu)\}, \quad B_1 = \{x \mid f(x) < f(x - \mu)\}.$$

Вероятность правильной классификации [1]

$$P_B = \frac{1}{2} \int_{\mathbf{R}^1} \max\{f(x), f(x - \mu)\} dx = \frac{1}{2} + \frac{1}{4} \int_{\mathbf{R}^1} |f(x) - f(x - \mu)| dx.$$

В настоящей работе исследуется нижняя граница вероятности P_B по всем распределениям $f(x)$ с фиксированной дисперсией. Полученные результаты обобщаются на многомерный случай.

Теорема 1. Пусть F — класс интегрируемых на $\mathbf{R}^1$ функций $f(x)$, удовлетворяющих условиям

$$f(x) \geq 0, \quad \int_{\mathbf{R}^1} f(x) dx = 1, \quad \int_{\mathbf{R}^1} x f(x) dx = 0, \quad \int_{\mathbf{R}^1} x^2 f(x) dx = 1.$$

*Лаборатории «ТВП», ул. Красикова, 32, 117418 Москва, Россия.

Тогда нижняя грань функционала

$$I(f, \mu) = \int_{\mathbf{R}^1} |f(x) - f(x - \mu)| dx, \quad \mu > 0,$$

по классу функций F удовлетворяет неравенству

$$(1 - 2\mu) \frac{\mu}{\sqrt{3}} \leq \inf_{f \in F} I(f, \mu) \leq \frac{\mu}{\sqrt{3}}.$$

Докажем несколько вспомогательных утверждений, сформулировав их в виде лемм.

Лемма 1. Пусть $\{r_k = k\mu, k = 0, \pm 1, \pm 2, \dots\}$ есть разбиение прямой $\mathbf{R}^1$ на полуинтервалы длины μ : $r_{k+1} - r_k = \mu$; $p_k = \int_{r_k}^{r_{k+1}} f(x) dx$. Тогда для любой $f(x) \in F$ выполняется неравенство $I(f, \mu) \geq 2 \max_k p_k$.

Доказательство. Справедлива следующая цепочка соотношений:

$$\begin{aligned} I(f, \mu) &= \int_{\mathbf{R}^1} |f(x) - f(x - \mu)| dx = \sum_{k=-\infty}^{\infty} \int_{r_k}^{r_{k+1}} |f(x) - f(x - \mu)| dx \\ &\geq \sum_{k=-\infty}^{\infty} \left| \int_{r_k}^{r_{k+1}} f(x) dx - \int_{r_k}^{r_{k+1}} f(x - \mu) dx \right| = \sum_{k=-\infty}^{\infty} |p_{k+1} - p_k|. \end{aligned}$$

Пусть $p_s = \max_k p_k$. Положим

$$m_k = \begin{cases} \max_{u \leq k} p_u, & \text{если } k < s, \\ \max_{u \geq k} p_u, & \text{если } k > s, \\ p_s, & \text{если } k = s. \end{cases}$$

Тогда

$$|p_{k+1} - p_k| \geq |m_{k+1} - m_k| \quad \text{и} \quad \sum_{k=-\infty}^{\infty} |p_{k+1} - p_k| \geq \sum_{k=-\infty}^{\infty} |m_{k+1} - m_k| = 2p_s.$$

Лемма 1 доказана.

Лемма 2. Пусть $f(x) \in F$. Положим

$$f^*(x) = \frac{f(x) + f(-x)}{2} \quad \text{и} \quad p_k^* = \int_{r_k}^{r_{k+1}} f^*(x) dx.$$

Тогда $f^*(x) \in F$ и $\max_k p_k \geq \max_k p_k^*$.

Доказательство. Легко проверить, что функция $f^*(x)$ удовлетворяет условиям теоремы 1. Следовательно, $f^*(x) \in F$. Далее,

$$p_k^* = \frac{p_k + p_{-(k+1)}}{2} \leq \max \{p_k, p_{-(k+1)}\} \quad \text{и} \quad \max_k p_k \geq \max_k p_k^*.$$

Лемма 2 доказана.

Обозначим через l_2 линейное пространство счетных последовательностей вещественных чисел вида

$$l_2 = \left\{ \bar{z} = (z_0, z_1, \dots, z_k, \dots): \sum_{k=0}^{\infty} z_k^2 < \infty \right\}$$

с нормой $\|\bar{z}\| = [\sum_{k=0}^{\infty} z_k^2]^{1/2}$; через $l_2(M_1, M_2)$ обозначим подмножество l_2 вида

$$l_2(M_1, M_2) = \left\{ \bar{z}: z_k \geq 0, \sum_{k=0}^{\infty} z_k = M_1; \sum_{k=0}^{\infty} (k+1)^2 z_k \leq M_2 \right\}.$$

Для $l_2(M_1, M_2)$ легко проверить справедливость следующей леммы 3 [2].

Лемма 3. Множество $l_2(M_1, M_2)$ компактно при любых M_1, M_2 , $0 < M_1 \leq M_2$.

Лемма 4. Пусть последовательность $\bar{u} = \{u_0, u_1, \dots, u_k, \dots\} \in l_2(M_1, M_2)$ такова, что

$$\beta = \inf_{z \in l_2(M_1, M_2)} \max_k z_k = \max_k u_k.$$

Тогда в последовательности $\bar{u}$ лишь конечное число начальных членов отличается от нуля и

$$u_0 = u_1 = \dots = u_{K-1} = \lambda, \quad u_K = \alpha, \quad 0 \leq \alpha < \lambda,$$

для некоторого K . Величина λ удовлетворяет неравенству

$$\lambda \geq \frac{4M_1}{3 + (1 + 48M_2/M_1)^{1/2}}.$$

Доказательство. Нетрудно заметить, что нижняя грань максимального элемента будет достигаться на такой последовательности $\bar{u}$, для которой сумма $\sum_{k=0}^{\infty} (k+1)^2 u_k$ точно равна M_2 . Очевидно также, что если $k > r$ и $u_k > u_r$, то перестановка u_k и u_r сохраняет $\sum_{k=0}^{\infty} u_k$ и уменьшает $\sum_{k=0}^{\infty} (k+1)^2 u_k$. Следовательно, последовательность $\bar{u}$ должна быть невозрастающей: $u_0 \geq u_1 \geq \dots \geq u_k \geq \dots$. Выберем в последовательности $\bar{u}$ члены u_s, u_m, u_r и дадим им приращения $\varepsilon_s, \varepsilon_m, \varepsilon_r$. Для того, чтобы вновь полученная последовательность принадлежала $l_2(M_1, M_2)$ и для нее выполнялось условие леммы, необходимо и достаточно, чтобы величины $\varepsilon_s, \varepsilon_m, \varepsilon_r$ удовлетворяли соотношениям

$$\begin{cases} \varepsilon_s + \varepsilon_m + \varepsilon_r = 0, \\ (s+1)^2 \varepsilon_s + (m+1)^2 \varepsilon_m + (r+1)^2 \varepsilon_r = 0. \end{cases}$$

Отсюда при заданном ε_s получаем, что

$$\begin{aligned} \varepsilon_m &= -\varepsilon_s \frac{(r+1)^2 - (s+1)^2}{(r+1)^2 - (m+1)^2}, \\ \varepsilon_r &= \varepsilon_s \frac{(m+1)^2 - (s+1)^2}{(r+1)^2 - (m+1)^2}. \end{aligned} \quad (1)$$

Пусть

$$u_0 = u_1 = \dots = u_{K-1} > u_K \geq u_{K+1} \geq \dots \geq u_{K+t} \geq \dots$$

и $u_{K+t} > 0$ для некоторых K и t . Если $t \geq 1$, то сделаем следующую операцию. Уменьшим на достаточно малую величину каждое значение u_i , $i = 0, \dots, K-1$, увеличивая в соответствии с (1) значение u_K и уменьшая u_{K+t} . В итоге получим последовательность, максимальный член которой меньше u_0 . Данное противоречие показывает, что $t = 0$ и, следовательно, последовательность, на которой достигается нижняя грань, есть последовательность вида, указанного в заключении леммы 4. Для такой последовательности условия

$$\sum_{k=0}^{\infty} u_k = M_1, \quad \sum_{k=0}^{\infty} (k+1)^2 u_k = M_2$$

приобретают вид

$$K\lambda + \alpha = M_1, \quad \frac{\lambda K(K+1)(2K+1)}{6} + (K+1)^2 \alpha = M_2. \quad (2)$$

Задача сводится к оценке снизу значения λ , удовлетворяющего условиям (2), где $0 \leq \alpha < \lambda$. Из первого равенства находим, что $\lambda \geq M_1/(K+1)$, второе равенство дает

$$M_2 \geq \frac{\lambda K(K+1)(2K+1)}{6} \geq M_1 \frac{2K^2 + K}{6}.$$

Отсюда следует, что

$$K \leq \frac{(1 + 48M_2/M_1)^{1/2} - 1}{4}$$

и

$$\lambda \geq \frac{4M_1}{3 + (1 + 48M_2/M_1)^{1/2}}.$$

Лемма 4 доказана.

Доказательство теоремы 1. Из лемм 1 и 2 получаем, что $I(f, \mu) \geq 2 \max_k p_k^*$. Следовательно,

$$\inf_{f \in F} I(f, \mu) \geq 2 \inf_{f^* \in F^*} \max_k p_k^*, \quad (3)$$

где F^* — множество всех четных функций, удовлетворяющих условиям теоремы 1. Эти условия для четных функций могут быть переписаны в эквивалентном и более удобном для дальнейших выкладок виде:

$$f^*(x) \geq 0, \quad \int_0^\infty f^*(x) dx = \frac{1}{2}, \quad \int_0^\infty x^2 f^*(x) dx = \frac{1}{2}.$$

Каждая из последовательностей $\bar{p}^* = (p_1^*, p_2^*, \dots, p_k^*, \dots)$ в (3) однозначно определяется видом соответствующей функции $f^*(x)$, а совокупность таких последовательностей образует множество $l_2(\frac{1}{2}, M)$, где $M = \frac{1}{2}(1 + 1/\mu)^2$. Этот факт нетрудно проследить, рассмотрев разность

$$\sum_{k=0}^\infty (k+1)^2 \mu^2 p_k^* - \int_0^\infty x^2 f^*(x) dx = \sum_{k=0}^\infty \int_{r_k}^{r_{k+1}} ((k+1)^2 \mu^2 - x^2) f^*(x) dx$$

и заметив, что при $k\mu \leq x \leq (k+1)\mu$ имеет место соотношение

$$|(k+1)^2 \mu^2 - x^2| = |((k+1)\mu - x)((k+1)\mu + x)| \leq \mu(2|x| + \mu).$$

Используя соответствие между последовательностями $\bar{p}^* = (p_1^*, p_2^*, \dots, p_k^*, \dots)$ и функциями $f^*(x)$, находим:

$$2 \inf_{f^*(x) \in F^*} \max_k p_k^* \geq 2 \inf_{\bar{p}^* \in l_2(\frac{1}{2}, M)} \max_k p_k^*. \quad (4)$$

По лемме 4

$$2 \inf_{\bar{p}^* \in l_2(\frac{1}{2}, M)} \max_k p_k^* = 2\lambda, \quad (5)$$

где

$$2\lambda \geq \frac{4}{3 + (1 + 96M)^{1/2}} = \frac{4}{3 + [1 + 48(1 + 1/\mu)^2]^{1/2}}. \quad (6)$$

Заметив, что

$$\left[1 + 48\left(1 + \frac{1}{\mu}\right)^2\right]^{1/2} \leq 1 + 3^{1/2} 4 \left(1 + \frac{1}{\mu}\right),$$

элементарными выкладками легко показать, что

$$\frac{4}{3 + [1 + 48(1 + 1/\mu)^2]^{1/2}} \geq (1 - 2\mu) \frac{\mu}{\sqrt{3}}. \quad (7)$$

Из (3)–(7) получаем оценку

$$\inf_{f \in F} I(f, \mu) \geq (1 - 2\mu) \frac{\mu}{\sqrt{3}}.$$

С другой стороны, для равномерного распределения, заданного на $[-3^{1/2}, 3^{1/2}]$, выполняются условия теоремы 1 и $I(f, \mu) = \mu/\sqrt{3}$. Отсюда следует, что $\inf_{f \in F} I(f, \mu) \leq \mu/\sqrt{3}$. Теорема 1 доказана.

Доказанную выше теорему 1 можно обобщить на многомерный случай. Пусть компонента X принимает значения в d -мерном евклидовом пространстве $\mathbf{R}^d$, $f(\mathbf{x})$, $f(\mathbf{x} - \mathbf{a})$ ($\mathbf{x} \in \mathbf{R}^d$, $\mathbf{a} \in \mathbf{R}^d$) есть условные плотности распределения X при условиях $\varepsilon = 0$ и $\varepsilon = 1$ соответственно. Оптимальная байесовская процедура $B = (B_0, B_1)$ имеет вид

$$B_0 = \{\mathbf{x} \mid f(\mathbf{x}) \geq f(\mathbf{x} - \mathbf{a})\}, \quad B_1 = \{\mathbf{x} \mid f(\mathbf{x}) < f(\mathbf{x} - \mathbf{a})\},$$

вероятность правильной классификации равна

$$P_B = \frac{1}{2} + \frac{1}{4} \int_{\mathbf{R}^d} |f(\mathbf{x}) - f(\mathbf{x} - \mathbf{a})| d\mathbf{x}.$$

Введем обозначения: $\text{cov}(x_i, x_j) = \int_{\mathbf{R}^d} x_i x_j f(\mathbf{x}) d\mathbf{x}$ (ковариация i -й и j -й координат), $Q_f = \|\text{cov}(x_i, x_j)\|$ (матрица ковариаций), $I(f, \mathbf{a}) = \int_{\mathbf{R}^d} |f(\mathbf{x}) - f(\mathbf{x} - \mathbf{a})| d\mathbf{x}$.

Теорема 2. Пусть F^d — класс интегрируемых в $\mathbf{R}^d$ функций $f(\mathbf{x})$, удовлетворяющих условиям

$$f(\mathbf{x}) \geq 0, \quad \int_{\mathbf{R}^d} f(\mathbf{x}) d\mathbf{x} = 1, \quad \int_{\mathbf{R}^d} x_i f(\mathbf{x}) d\mathbf{x} = 0, \quad i = 1, \dots, d, \quad \det Q_f \neq 0.$$

Тогда нижняя грань $I(f, \mathbf{a})$, взятая по всем функциям $f(\mathbf{x}) \in F^d$ и значениям параметра $\mathbf{a}$, для которых $(\mathbf{a}^T Q_f^{-1} \mathbf{a})^{1/2} = \mu > 0$, удовлетворяет неравенству

$$(1 - 2\mu) \frac{\mu}{\sqrt{3}} \leq \inf_{f \in F^d, \mathbf{a}} I(f, \mathbf{a}) \leq \frac{\mu}{\sqrt{3}}.$$

Доказательство. Осуществим отображение $\mathbf{R}^d \rightarrow \mathbf{R}^1$ с помощью преобразования

$$y = \mathbf{x}^T Q_f^{-1} \mathbf{a} (\mathbf{a}^T Q_f^{-1} \mathbf{a})^{-1/2}.$$

Тогда распределение вероятностей с плотностью $f(\mathbf{x})$ индуцирует в $\mathbf{R}^1$ распределение с некоторой плотностью $\psi(y)$, а распределение $f(\mathbf{x} - \mathbf{a})$ индуцирует в $\mathbf{R}^1$ распределение с плотностью $\psi(y - \mu)$. Поскольку величина $\frac{1}{2} + \frac{1}{4} I(f, \mathbf{a})$ представляет собой вероятность правильной классификации для оптимальной процедуры, этот факт влечет за собой выполнение неравенства $I(f, \mathbf{a}) \geq I(\psi, \mu)$. Используя известные соотношения для матрицы ковариаций при линейном преобразовании координат [3], нетрудно показать, что дисперсия случайной величины y равна 1. Отсюда следует, что функция $\psi(y)$ принадлежит классу F интегрируемых функций, удовлетворяющих условиям теоремы 1. Тогда по теореме 1 получаем:

$$\inf_{f \in F^d, \mathbf{a}} I(f, \mathbf{a}) \geq \inf_{\psi \in F} I(\psi, \mu) \geq (1 - 2\mu) \frac{\mu}{\sqrt{3}}.$$

С другой стороны, рассмотрим плотность распределения вероятностей вида

$$\tilde{f}(x_1, x_2, \dots, x_d) = f(x_1) f^*(x_2, \dots, x_d),$$

где

$$f(x_1) = \begin{cases} (2\sqrt{3})^{-1}, & |x_1| \leq 3^{1/2}, \\ 0, & |x_1| > 3^{1/2}, \end{cases}$$

и положим $\mathbf{a} = (\mu, 0, 0, \dots, 0)$. Тогда $(\mathbf{a}^T Q_f^{-1} \mathbf{a})^{1/2} = \mu$ и

$$I(\tilde{f}, \mathbf{a}) = \int_{\mathbf{R}^d} |f(x_1) - f(x_1 - \mu)| f^*(x_2, \dots, x_d) dx = \frac{\mu}{\sqrt{3}}.$$

Отсюда следует, что

$$\inf_{f \in F^d, \mathbf{a}} I(f, \mathbf{a}) \leq \frac{\mu}{\sqrt{3}}.$$

Теорема 2 доказана.

СПИСОК ЛИТЕРАТУРЫ

1. Фукунага К. Введение в статистическую теорию распознавания образов. М.: Наука, 1979, 388 с.
2. Люстерник Л. А., Соболев В. И. Элементы функционального анализа. М.: Наука, 1965, 544 с.
3. Крамер Г. Математические методы статистики. М.: Мир, 1975, 648 с.

Поступила в редакцию
10.VII.1992

Исправленный вариант
3.VIII.1995

© 1996 г.

БЕЖАЕВА З. И.*, ОСЕЛЕДЕЦ В. И.**

О ДИСПЕРСИИ СУММ ДЛЯ ФУНКЦИЙ ОТ СТАЦИОНАРНОГО МАРКОВСКОГО ПРОЦЕССА¹⁾

Для функции от стационарного эргодического марковского процесса изучается вопрос о предельном поведении дисперсии сумм. При некоторых дополнительных предположениях получено необходимое и достаточное условие равенства нулю предельной средней дисперсии.

Ключевые слова и фразы: дисперсия сумм, эргодический марковский процесс, перемешивание, пересечение σ -алгебр.

*МГИЭМ, Б.Вузовский пер., 3/4, Москва, Россия.

**Московский государственный университет им. М. В. Ломоносова, Воробьевы горы, Москва, Россия.

¹⁾Работа выполнена при частичной поддержке Российского фонда фундаментальных исследований (грант № 93-01-00239).