



<http://dx.doi.org/10.15407/dopovidi2016.06.025>

УДК 519.7

О. А. Галкін

Київський національний університет ім. Тараса Шевченка

E-mail: galkin.o.a@gmail.com

Асимптотичні властивості Σ -класифікатора для багатокласових задач розпізнавання з нееліптичним розподілом даних

(Представлено членом-кореспондентом НАН України А. В. Анісімовим)

Досліджуються асимптотичні властивості Σ -класифікатора, що не вимагає попередньої інформації про розподіл або форму розділової кривої. Побудовано математичний апарат для розв'язання багатокласових задач розпізнавання з нееліптичним розподілом даних на основі методу мажоритарного голосування. Досліджено процедуру визначення форм розділових кривих Σ -класифікатора по геометричній структурі даних, що лежать в основі Σ -схеми. Визначено умови, при яких Σ -класифікатор є асимптотично еквівалентним правилу Байеса.

Ключові слова: правило Байеса, Σ -класифікатор, асимптотична збіжність.

Постановка задачі. Дана робота присвячена дослідженню результатів узгодженості Σ -класифікатора, який описано в [1], а також визначенню умов, при яких запропоноване правило класифікації буде асимптотично еквівалентним правилу Байеса. Враховуючи той факт, що досліджувані функції глибини є обмеженими, будемо вважати, що вони обмежені. Визначимо $V_\tau(r_1, r_2) = 1$, якщо $r_2 > d_\tau(r_1)$ та $V_\tau(r_1, r_2) = 0$, якщо $r_2 \leq d_\tau(r_1)$ при $r_1, r_2 \in [0, 1]$ та $\tau \in \mathbb{R}^{\Delta_0}$. Вирази $V_\tau(E_H(z), E_U(z))$ та $V_{\bar{\tau}_M}(E_{H_n}(z), E_{U_m}(z))$ відповідають Σ -класифікатору та його емпіричному аналогу, що має таку форму при $\bar{\tau}_M = \arg \min\{\bar{\Psi}_M(\tau)\}$: а) якщо $E_{U_m}(z) > d_{\bar{\tau}_M}(E_{H_n}(z))$, тоді $z \in U$; б) якщо $E_{U_m}(z) < d_{\bar{\tau}_M}(E_{H_n}(z))$, тоді $z \in H$.

Коефіцієнти помилкової класифікації для $V_\tau(E_H(z), E_U(z))$ та $V_{\bar{\tau}_M}(E_{H_n}(z), E_{U_m}(z))$ можна визначити таким чином:

$$\Psi(V_\tau) = p_1 P_H\{x: V_\tau(E_H(x), E_U(x)) = 1\} + p_2 P_U\{x: V_\tau(E_H(x), E_U(x)) = 0\} \quad (1)$$

та

$$\begin{aligned} \Psi_M(\bar{V}_M) = & p_1 P_H \{x: V_{\tau_M}(E_{H_n}(x), E_{U_m}(x)) = 1\} + \\ & + p_2 P_U \{x: V_{\tau_M}(E_{H_n}(x), E_{U_m}(x)) = 0\}, \end{aligned} \quad (2)$$

де $\Psi_M(\bar{V}_M)$ є умовною ймовірністю помилкової класифікації, коли $E_{U_m}(X)$ та $d_{\tau_M}(E_{H_n}(X))$ використовуються для класифікації елемента X .

Властивості Σ -класифікатора. Розглянемо функцію $V_\omega(r_1, r_2) = \omega(r_1)$, де змінну ω будемо використовувати у якості показника скінченного об'єднання інтервалів [2]. Після застосування даної функції до $(E_H(x), E_U(x))$, класифікація точки x відбуватиметься виключно на основі її глибини відносно H . Крім того, пов'язана із функцією ω помилка невірної класифікації може бути визначена таким чином:

$$\Psi(V_\omega) = p_1 P_H \{x: V_\omega(E_H(x), E_U(x)) = 1\} + p_2 P_U \{x: V_\omega(E_H(x), E_U(x)) = 0\}. \quad (3)$$

Далі припустимо, що для $a \in \mathbb{R}$ та $\forall \tau \in \mathbb{R}^{\Delta_0}$

$$P_H \{x: E_U(x) = d_\tau(E_H(x))\} = P_U \{x: E_U(x) = d_\tau(E_H(x))\} = 0, \quad (4)$$

$$P_H \{x: E_H(x) = a\} = P_U \{x: E_H(x) = a\} = 0. \quad (5)$$

Твердження 1 . $\exists V_0 \in G$, що $\Psi(V_0) = \inf_{V \in G} \Psi(V)$ при умові, якщо H та U задовольняють (4) та (5).

Теорема 1. Нехай $E_H(\cdot)$ та $E_U(\cdot)$ задовольняють асимптотичній збіжності $\sup_z |E_{H_n}(z) - E_H(z)| \rightarrow 0$ та $\sup_z |E_{U_m}(z) - E_U(z)| \rightarrow 0$ при $\min(n, m) \rightarrow \infty$, а також є неперервними. Крім того припустимо, що H та U задовольняють (4) та (5). Тоді $\bar{\Psi}_M(V)$ асимптотично сходиться до $\Psi(V)$ при $\min(n, m) \rightarrow \infty$ для $\forall V \in G$.

Доведення. Розглядаючи майже напевну збіжність послідовностей, ми використовуємо теорему Скорохода в сенсі розподільної збіжності для побудови відповідного представлення емпіричних розподілів [3]. Припустимо, що випадкові вибірки $\{Z_n\}$ та $\{X_m\}$ визначені на ймовірнісному просторі (E, η, ε) , а випадковість введених величин залежить від $e \in E$. Таким чином, ми приймаємо $\{Z_n(e)\}$ та $\{X_m(e)\}$ для послідовностей, а також H_n^e та U_m^e для емпіричних розподілів, що базуються на $\{Z_1(e), \dots, Z_n(e)\}$ та $\{X_1(e), \dots, X_m(e)\}$ відповідно. Для спрощення запису введемо позначення $\bar{\Psi}_M^e$ та \bar{V}_M^e .

Можна стверджувати, що \exists множина $E_0 \in \eta$, яка $\varepsilon(E_0) = 1$ та для $\forall e \in E_0$ послідовності функцій розподілу $\{H_n^e\}$ та $\{U_m^e\}$ сходяться до H та U відповідно. Даний висновок слідує з r -вимірної теореми Гливленко–Кантеллі, оскільки $\min(n, m) \rightarrow \infty$. Тоді \exists такий ймовірнісний простір (I, η, ε_I) , що якщо $e \in E_0$, тоді \exists такі послідовності випадкових величин $\{X_n^{e,H}\}_{n \geq 0}$ та $\{X_m^{e,U}\}_{m \geq 0}$, що

а) розподілом $X_n^{e,H} \in H_n^e$, а розподілом $X_m^{e,U} \in U_m^e$ для $\forall n, m = 1, \dots$. Також розподілом $X_0^{e,H} \in H$, а розподілом $X_0^{e,U} \in U$.

б) послідовності $\{X_n^{e,H}\}_{n \geq 1}$ та $\{X_m^{e,U}\}_{m \geq 1}$ сходяться майже напевно до випадкових величин $X_0^{e,H}$ та $X_0^{e,U}$ відповідно.

Наступні результати мають місце при застосуванні доданих послідовностей теореми Скорохода.

Зазначимо, що множина $E^* = E_0 \cap E_1$ має ймовірність 1, оскільки E_1 є множиною одиничної ймовірності, в якій збіжності $\sup_z |E_{H_n}(z) - E_H(z)| \rightarrow 0$ та $\sup_z |E_{U_m}(z) - E_U(z)| \rightarrow 0$ задоволені. Отже ми маємо

$$\begin{aligned} \bar{\Psi}_M^e(V) &= p_1 P_{H_n^e} \{z : V(E_{H_n^e}(z), E_{U_m^e}(z)) = 1\} + p_2 P_{U_m^e} \{z : V(E_{H_n^e}(z), E_{U_m^e}(z)) = 0\} = \\ &= p_1 \varepsilon_1 \{V(E_{H_n^e}(X_n^{e,H}), E_{U_m^e}(X_n^{e,H})) = 1\} + p_2 \varepsilon_1 \{V(E_{H_n^e}(X_m^{e,U}), E_{U_m^e}(X_m^{e,U})) = 0\} = \\ &= W_n^e + Q_m^e. \end{aligned} \quad (6)$$

де $e \in E^*$.

Крім того, має місце така рівність:

$$\begin{aligned} |E_{H_n^e}(X_n^{e,H}) - E_H(X_0^{e,H})| &= |E_{H_n^e}(X_n^{e,H}) - E_H(X_n^{e,H})| + |E_H(X_n^{e,H}) - E_H(X_0^{e,H})| = \\ &= W_{n,1}^e + W_{n,2}^e. \end{aligned} \quad (7)$$

Оскільки точка e розглядається як така, що належить до E_1 , то $\{W_{n,1}^e\}_n \rightarrow 0$. Однак, оскільки $e \in E_0$, $\varepsilon_1 \xrightarrow{\text{ac}} X_0^{e,H}$, де $X_n^{e,H}$ є випадковими величинами, що визначені на I .

Отже, $W_{n,2}^e \varepsilon_1 \xrightarrow{\text{ac}} 0$, оскільки E_H є неперервною функцією. Використовуючи аналогічний підхід до $E_{U_m^e}(X_n^{e,H})$, ми отримуємо

$$(E_{H_n^e}(X_n^{e,H}), E_{U_m^e}(X_n^{e,H})) \xrightarrow{\text{ac}} (E_H(X_0^{e,H}), E_U(X_0^{e,H}))$$

відносно ймовірності ε_1 . У даному випадку межею множини $\{(t, b) : V(t, b) = 0\} \in \{(t, b) : t = d(b)\}$, коли $V = V_d$ (d є многочленом) та множиною вигляду $\{\{a_i\} \times [0, 1] : i = 1, \dots, f+1\}$, коли V є показником об'єднання f інтервалів.

Тому, як слідує з (4) та (5),

$$W_n^e \rightarrow p_1 \varepsilon_1 \{V(E_H(X_0^{e,H}), E_U(X_0^{e,H})) = 1\} = p_1 P_H(z : V(E_H(z), E_U(z)) = 1), \quad (8)$$

оскільки розподілом $X_0^{e,H} \in H$. Отже, результат має місце, оскільки (8) виконується для $\forall e \in E^*$, що є множиною одиничної ймовірності. Теорему доведено.

Далі наведено лему 1, з якої слідує, що $\bar{V}_M \xrightarrow{\text{ac}} V_0$, де \bar{V}_M мінімізує емпіричний коефіцієнт помилкової класифікації, а V_0 — множинний коефіцієнт помилкової класифікації.

Лема 1. *Нехай $\bar{V}_M = \arg \min_{V \in G} \{\bar{\Psi}_M(V)\}$ та $V_0 = \arg \min_{V \in G} \{\Psi(V)\}$. Крім того, нехай асимптотична збіжність розглядається, як*

$$P_H\{x : \bar{V}_M(E_H(x), E_U(x)) \rightarrow V_0(E_H(x), E_U(x))\} = 1 \quad (9)$$

та

$$P_U\{x : \bar{V}_M(E_H(x), E_U(x)) \rightarrow V_0(E_H(x), E_U(x))\} = 1. \quad (10)$$

Тоді, якщо V_0 є унікальним, то $\bar{V}_M \xrightarrow{\text{ac}} V_0$ при $\min(n, m) \rightarrow \infty$.

Доведення. Як слідує з теореми 1, \exists така множина одиничної ймовірності E^* , що $\bar{\Psi}_M^e(V_0) \xrightarrow{\text{ac}} \Psi(V_0)$ для $\forall e \in E^*$. Крім того, E^* містить множину E_0 , в якій виконується асимптотична збіжність $\sup_z |E_{H_n}(z) - E_H(z)| \rightarrow 0$ та $\sup_z |E_{U_m}(z) - E_U(z)| \rightarrow 0$, а також має місце теорема Скорохода [4].

Розглянемо послідовність $\{\bar{V}_M^e\}_M$, де $e \in E^*$ є фіксованою точкою. Для доведення даної леми необхідно показати, що вся послідовність сходиться до E_0 . Дана збіжність буде мати місце, якщо кожна підпослідовність $\{\bar{V}_M^e\}_M$ буде містити додаткову підпослідовність, що сходиться до E_0 . Отже, розглянемо підпослідовність $\{\bar{V}_M^e\}_M$. Зазначимо, що дана підпослідовність містить асимптотично збіжну підпослідовність. Позначимо границю збіжної підпослідовності послідовності $\{\bar{V}_M^e\}_M$, як V^e .

Розглядаючи послідовність $\{\bar{\Psi}_M^e(\bar{V}_M^e)\}_M$, ми маємо

$$\begin{aligned} \bar{\Psi}_M^e(\bar{V}_M^e) &= p_1 \varepsilon_1 \{\bar{V}_M^e(E_{H_n^e}(X_n^{e,H}), E_{U_m^e}(X_n^{e,H})) = 1\} + \\ &+ p_2 \varepsilon_1 \{\bar{V}_M^e(E_{H_n^e}(X_m^{e,U}), E_{U_m^e}(X_m^{e,U})) = 0\}, \end{aligned} \quad (11)$$

як і в доведенні теореми 1.

Аналогічним чином ми отримуємо, що

$$(E_{H_n^e}(X_n^{e,H}), E_{U_m^e}(X_n^{e,H})) \xrightarrow{\text{ac.}} (E_H(X_0^{e,H}), E_U(X_0^{e,H})). \quad (12)$$

Зазначимо, що збіжність $\bar{V}_M^e \rightarrow V^e$ означає, що $\bar{V}_M^e(E_H(z), E_U(z)) = 1$ при умові, якщо $V^e(E_H(z), E_U(z)) = 1$, а $(E_H(z), E_U(z))$ не належить межі множини $\{V^e = 1\}$.

Розглядаючи асимптотичну збіжність відносно ймовірності ε_1 , ми отримуємо такі асимптотичні збіжності:

$$\Lambda_{\{\bar{V}_M^e(E_{H_n^e}(X_n^{e,H}), E_{U_m^e}(X_n^{e,H}))=1\}} \rightarrow \Lambda_{\{\bar{V}^e(E_H(X_0^{e,H}), E_U(X_0^{e,H}))=1\}} \quad (13)$$

та

$$\Lambda_{\{\bar{V}_M^e(E_{H_n^e}(X_n^{e,H}), E_{U_m^e}(X_n^{e,H}))=0\}} \rightarrow \Lambda_{\{\bar{V}^e(E_H(X_0^{e,H}), E_U(X_0^{e,H}))=0\}}, \quad (14)$$

оскільки ймовірність того, що $X_0^{e,H}$ належить межі $\{V^e = 1\}$ дорівнює нулю.

З рівності (11) та теореми Лебега про мажоровану збіжність слідує, що $\bar{\Psi}_M^e(\bar{V}_M^e) \rightarrow \Psi(V^e)$ для точки e , оскільки всі випадкові величини є додатними та обмеженими 1 [5]. Крім того, в даній точці e має місце збіжність $\bar{\Psi}_M^e(V_0) \xrightarrow{\text{ac.}} \Psi(V_0)$, та як $e \in E^*$.

Таким чином, $\Psi(V^e) = \Psi(V_0)$, оскільки

$$\Psi(V_0) = \lim \bar{\Psi}_M^e(V_0) \geq \lim \bar{\Psi}_M^e(\bar{V}_M^e) = \Psi(V^e) \geq \Psi(V_0), \quad (15)$$

що слідує з визначення V_0 та \bar{V}_M^e . Отже, $V^e = V_0$, враховуючи унікальність V_0 . Лему доведено.

Далі наведемо основний результат роботи.

Теорема 2. Припустимо, що $p_1 = p_2$, а функції щільності $h_1(\cdot)$ та $h_2(\cdot)$ з H та U відповідно мають вигляд $h_i(z) = v_i |\Xi_i|^{-1/2} f_i((z - \varepsilon_i)' \Xi_i^{-1} (z - \varepsilon_i))$ з $h_1(\cdot) = h_2(\cdot)$ та $\Xi_1 = \Xi_2$, де $i = 1, 2$. Якщо $\tau_1 = (1, 0, \dots, 0)$, а функція глибини, що використовується в алгоритмі класифікації є глибиною Махаланобіса, напівпросторовою глибиною, симплексною глибиною, або проекційною глибиною, тоді ми маємо, що $\Omega(\Psi_M(\bar{V}_M)) \rightarrow \Psi(V_{\tau_1})$ при $\min(n, m) \rightarrow \infty$.

Доведення. З леми 1 слідує, що $\bar{V}_M \xrightarrow{\text{ac.}} V_{\tau_1}$ при $\min(n, m) \rightarrow \infty$. Крім того, $V_0 = V_{\tau_1}$.

Зазначимо також, що

$$|\Psi_M(\bar{V}_M) - \Psi(V_{\tau_1})| \leq p_1 \int |\Lambda_{\{\bar{V}_M(E_{H_n}(x), E_{U_m}(x))=1\}} - \Lambda_{\{V_{\tau_1}(E_H(x), E_U(x))=1\}}| h_1(x) dx +$$

$$+ p_2 \int |\Lambda_{\{\bar{V}_M(E_{H_n}(x), E_{U_m}(x))=0\}} - \Lambda_{\{V_{\tau_1}(E_H(x), E_U(x))=0\}}| h_2(x) dx. \quad (16)$$

З теореми Лебега про мажоровану збіжність маємо

$$|\Psi_M(\bar{V}_M) - \Psi(V_{\tau_1})| \xrightarrow{a.s.} 0 \quad (17)$$

при $\min(n, m) \rightarrow \infty$. Даний результат має місце при об'єднанні емпіричних функцій глибини з множинними функціями глибини з майже напевною поточною збіжністю. Остаточний результат маємо при повторному застосуванні теореми Лебега про мажоровану збіжність. Теорему доведено.

Враховуючи припущення теореми 2, можна стверджувати, що правило Байєса еквівалентно такій схемі: 1) якщо $E_U(z) > E_H(z)$, тоді z належатиме U ; 2) якщо $E_U(z) < E_H(z)$, тоді z належатиме H . Таким чином, $\Psi(V_{\tau_1})$ відповідає байєсівському ризику. Отже, на основі заданих припущень можна стверджувати, що запропонований Σ -класифікатор еквівалентний правилу Байєса.

Цитована література

1. Галкин О. А. Непараметричный метод классификации для задач с неэллиптическим распределением данных на основе глубиннозависимой Σ -схемы // Вісн. Київ. нац. ун-ту ім. Тараса Шевченка. Сер. фіз.-мат. науки. – 2015. – № 3. – С. 60–65.
2. Lange T., Mosler K., Mozharovskyi P. Fast nonparametric classification based on data depth // Statist. Papers. – 2014. – **55**. – P. 53–67.
3. Cuesta-Albertos J. A., Nieto-Reyes A. The random Tukey depth // Computational Statistics & Data Analysis. – 2008. – **52**. – P. 4980–4987.
4. Li J., Zuo R. Y. New nonparametric tests of multivariate locations and scales using data depth // Statistical Sci. – 2004. – **19**. – P. 687–694.
5. Zuo Y. J. Projection-based depth functions and associated medians // The Annals of Statistics. – 2003. – **31**. – P. 1463–1484.

References

1. Galkin O. A. Bull. of Taras Shevchenko National University of Kiev. Ser. Phys. & Math., 2015, No 3: 60–65 (in Ukrainian).
2. Lange T., Mosler K., Mozharovskyi P. Statist. Papers, 2014, **55**: 53–67.
3. Cuesta-Albertos J. A., Nieto-Reyes A. Computational Statistics & Data Analysis, 2008, **52**: 4980–4987.
4. Li J., Zuo R. Y. Statistical Sci., 2004, **19**: 687–694.
5. Zuo Y. J. The Annals of Statistics, 2003, **31**: 1463–1484.

Надійшло до редакції 14.09.2015

А. А. Галкин

Киевский национальный университет им. Тараса Шевченко

E-mail: galkin.o.a@gmail.com

Асимптотические свойства Σ -классификатора для многоклассовых задач распознавания с неэллиптическим распределением данных

Исследуются асимптотические свойства Σ -классификатора, который не требует предварительной информации о распределении или форме разделительной кривой. Построен математический аппарат для решения многоклассовых задач распознавания с неэллиптическим распределением данных.

ским распределением данных на основе метода мажоритарного голосования. Исследована процедура определения форм разделительных кривых Σ -классификатора по геометрической структуре данных, лежащих в основе Σ -схемы. Определены условия, при которых Σ -классификатор является асимптотически эквивалентным правилу Байеса.

Ключевые слова: правило Байеса, Σ -классификатор, асимптотическая сходимость.

O. A. Galkin

Taras Shevchenko National University of Kiev

E-mail: galkin.o.a@gmail.com

The asymptotic properties of a Σ -classifier for multiclass recognition problems with non-elliptic distribution of data

The asymptotic properties of a Σ -classifier that does not require a priori information about the distribution or shape of a dividing curve are studied. A mathematical apparatus is built to solve multiclass recognition problems with a non-elliptic distribution of data on the basis of the majority voting method. The procedure is studied to determine the shapes of dividing curves of a Σ -classifier by the geometric structure of data that are the basis of the Σ -scheme. The conditions, under which the Σ -classifier is asymptotically equivalent to the Bayes rule, are defined.

Keywords: Bayes rule, Σ -classifier, asymptotic convergence.