УДК 519.7

Галкін О. А., к.ф.-м.н., асистент

Дослідження непараметричних класифікаторів даних на основі концепції глибинних околів

Київський національний університет імені Тараса Шевченка, 03680, м. Київ, пр-т. Академіка Глушкова 4д,
e-mail: galkin.o.a@gmail.com

O. A. Galkin, candidate of physical and mathematical sciences, assistant

Research of nonparametric data classifiers based on the depth neighborhoods concept

Taras Shevchenko National University of Kyiv, 03680, Kyiv, Glushkova st., 4d,
e-mail: galkin.o.a@gmail.com

Стаття присвячена дослідженню непараметричних глибинних класифікаторів на основі концепції глибинних околів. Розроблено процедуру симетризації функцій глибини на основі методу k-найближчих сусідів, що забезпечує центрально-зовнішнє впорядкування для визначення найближчих сусідів. Побудова симетризації асимптотично гарантує унікальність найглибшої точки, що вирішує проблему опуклої області з нескінченною множиною найглибших точок. Отримані глибинні оцінки k_m найближчих сусідів є ефективними завдяки використанню несферичних та нееліпсоїдальних околів, форма яких визначається локальною геометрією вибірки даних.
Ключові слова: непараметричний класифікатор, симетризація даних, функція глибини

The article is devoted to research nonparametric depth classifiers based on the depth neighborhoods concept. The symmetrization procedure of depth functions is developed on the basis of k-nearest neighbors, which provides centrally external ordering to determine the nearest neighbors. Construction of symmetrization asymptotically guarantees uniqueness of the deepest point, which solves the problem of a convex domain with an infinite set of the deepest points. The resulting depth based estimates of k_m nearest neighbors are effective through the use of non-spherical and non-ellipsoidal neighborhoods, whose form is determined by the local geometry of the sample data. Data depth in the proposed approach is considered relatively empirical distribution associated with sample obtained by adding to the initial data elements of their reflections. Consistency properties obtained in the classification problem can also be applied to the regression problems and density estimates. It is established that the use of spherical namely ellipsoid neighborhoods in density estimates problems can lead to fundamental finite-sample properties. Using of depth neighborhoods is of interest in regression problems where conditional mean of function is estimated on the basis of mutually independent groups of random vector.
Key Words: nonparametric classifier, data symmetrization, depth function

Статтю представив д.ф.-м.н., проф. Анісімов А.В.

Вступ

В рамках дослідження статистичних функцій глибини важливу роль відіграє поняття області глибини (порядку φ), що визначається як множина $C_\varphi(P)=\{z\in \mathrm{R}^r : E(z,P)\geq \varphi\}$ для $\forall\ \varphi>0$. Оскільки області глибини є вкладеними, їх індексація проводиться за ймовірнісною схемою, тобто для $\forall\ \gamma\in[0,1)$ величина $R^\gamma(P)$ вказує на найменше значення $R_\varphi(P)$, що має ймовірність $\geq\gamma$. Тому, для глибинних рівнів та ймовірнісного наповнення використовуються верхні та нижні індекси для областей глибини. Вибіркові форми верхніх глибин можна отримати шляхом заміни P на відповідний емпіричний розподіл $P^{(m)}$ для r-вимірних даних $Z_1,\dots,Z_m$. Можна впорядкувати Z_i таким чином, що

$$E(Z_{(1)}, P^{(m)}) \ge E(Z_{(2)}, P^{(m)}) \ge ... \ge E(Z_{(m)}, P^{(m)}), \quad (1)$$

оскільки вибіркова форма глибини забезпечує центрально-зовнішнє впорядкування елементів даних відносно відповідної найглибшої точки $\bar{\lambda}^{(m)}$. Тому, у глибинному сенсі $Z_{(1)}$ є елементом даних, найближчим до $\bar{\lambda}^{(m)}$, $Z_{(2)}$ - другим елементом даних, найближчим до $\bar{\lambda}^{(m)}$,…, а $Z_{(m)}$ - елементом даних, найвіддаленішим від $\bar{\lambda}^{(m)}$.

Використовуючи впорядкування (1), статистична функція глибини може застосовуватися для визначення сусідів найглибшої точки $\bar{\lambda}^{(m)}$. Однак, для реалізації класифікатора на основі методу *k*-найближчих сусідів необхідно визначити сусідів довільної точки $z \in R^r$.

У даній роботі запропоновано підхід, де глибина даних розглядається відносно емпіричного розподілу $P_Z^{(m)}$, що пов'язаний з вибіркою, отриманою шляхом додавання до початкових елементів даних $Z_1, Z_2, ..., Z_m$ їх відображень $2z - Z_1, ..., 2z - Z_m$ відносно z. Зазначимо, що z є асимптотично унікальною найглибшою точкою відносно $P_Z^{(m)}$. Тому, відповідна побудова симетризації призводить до z-зовнішнього впорядкування такої форми:

$$E(Z_{Z,(1)}, P_Z^{(m)}) \ge E(Z_{Z,(2)}, P_Z^{(m)}) \ge ... \ge E(Z_{Z,(m)}, P_Z^{(m)}), \quad (2)$$

елементи даних якого не є впорядкованими, а використовуються лише для визначення порядку.

Отже, глибинні околи, тобто вибіркові області глибини $C_{Z,\varphi}^{(m)} = C_\varphi(P_Z^{(m)})$ відіграють важливу роль, де $C_Z^{\gamma(m)}$ означає найменшу область глибини $C_{Z,\varphi}^{(m)}$, що містить щонайменше частину γ елементів даних $Z_1, Z_2, ..., Z_m$. Тому, для $\gamma = k/m$, $C_Z^{\gamma(m)}$ є найменшим глибинним околом, що містить k всіх Z_i.

Виклад основного матеріалу. Ідея запропонованого підходу полягає в побудові глибинних класифікаторів на основі методу k-найближчих сусідів шляхом заміни евклідових околів на відповідно визначені глибинні околи. Тобто, запропонований підхід k-найближчих сусідів буде класифікувати елемент z в множину 1 тоді і тільки тоді, коли множина 1 містить більшу кількість елементів даних, ніж множина 2 в найменшому глибинному околі z, що містить k елементів даних - тобто, в $C_Z^{\gamma(m)}$, де $\gamma = k/m$.

Отже, запропонований глибинний класифікатор визначається, як

$$\bar{n}_E^{(m)}(z) = \Lambda\left[\sum_{i=1}^{m} \Lambda\left[X_i = 1\right] D_i^{\gamma(m)}(z) > \sum_{i=1}^{m} \Lambda\left[X_i = 0\right] D_i^{\gamma(m)}(z)\right], \quad (3)$$

де $D_i^{\gamma(m)}(z) = \frac{1}{N_z^{\gamma(m)}} \Lambda\left[Z_i \in C_z^{\gamma(m)}\right]$,

а $N_z^{\gamma(m)} = \sum_{l=1}^{m} \Lambda\left[Z_l \in C_z^{\gamma(m)}\right]$ визначає число елементів даних в глибинному околі $C_z^{\gamma(m)}$. Зазначимо, що запропонований класифікатор (3) отриманий з використанням глибинної оцінки $\bar{\theta}_E^{(m)}(z)$ умовного математичного сподівання $\theta(z)$, оскільки

$$\bar{n}_E^{(m)}(z) = \Lambda\left[\bar{\theta}_E^{(m)}(z) > 1/2\right], \quad (4)$$

де $\bar{\theta}_E^{(m)}(z) = \sum_{i=1}^{m} \Lambda\left[X_i = 1\right] D_i^{\gamma(m)}(z)$.

Відзначимо, що в одновимірному випадку $\bar{n}_E^{(m)}$ зводиться до евклідового класифікатора k-найближчих сусідів незалежно від статистичної функції глибини E. Крім того, запропонований класифікатор є афінно-інваріантним, тобто, якщо $Z_1, ..., Z_m$ та z підлягають загальній афінній трансформації, результат класифікації залишиться незмінним [1]. У даному випадку метод визначення афінно-інваріантного класифікатора k-найближчих сусідів полягає у застосуванні класичного методу k-найближчих сусідів на нормалізованих даних $\bar{\Xi}^{-1/2} Z_i$, $i = 1, ..., m$. Для $\forall$ оберненої $r \times r$ матриці G та $\forall r$-вектора v, оцінка $\bar{\Xi}$ є афінно-інваріантною оцінкою, що має такий вигляд :

$$\bar{\Xi}\left(GZ_1 + v, ..., GZ_m + v\right) \propto G\bar{\Xi}\left(Z_1, ..., Z_m\right)G'. \quad (5)$$

Такий трансформаційний підхід призводить до околів, що не використовують геометрію розподілу в околі точки z та є еліпсоїдами з z-незалежною орієнтацією та формою.

Запропонований глибинний класифікатор на основі методу k-найближчих сусідів є узгодженим при відповідних умовах. Для цього статистична глибинна функція W повиннна задовольняти таким властивостям:

(a) Властивість унікальної максимізації в центрі симетрії. Тобто, $E(\hat{\lambda}, P) > E(z, P)$ для всіх $z \ne \hat{\lambda}$, якщо P є симетричним відносно $\hat{\lambda}$ та має щільність, що є додатньою в $\hat{\lambda}$.

(б) Властивість узгодженості. Тобто, величина $\sup_{z\in V}\left|E(z,P^{(m)})-E(z,P)\right|=o(1)$ майже напевно при $m\to\infty$ для $\forall$ обмеженої r-вимірної борелівської множини V. У даному випадку $P^{(m)}$ є емпіричним розподілом, пов'язаним з m випадковими векторами, що є незалежними однаково розподіленими випадковими величини P.
(в) Властивість неперервності. Тобто, $z\mapsto E(z,P)$ є неперервним в околі $\hat{\lambda}$, якщо розподіл P є симетричним відносно $\hat{\lambda}$ та має щільність, що є додатньою в $\hat{\lambda}$.

Теорема 1. Нехай P є ймовірнісною мірою, симетричною в околі $\hat{\lambda}$ та має щільність, що є додатньою в $\hat{\lambda}$. Також припустимо, що функція глибини E задовольняє властивостям максимального значення в центрі, монотонності відносно найглибшої точки, неперервності та унікальної максимізації в центрі симетрії. Тоді, для $\forall\ \varphi<\varphi_*$, $\exists\,\rho>0$, що $V_{\hat{\lambda}}(\rho)\subset C_{\varphi}(P)$ та для $\forall\ \tau>0$, $\exists\ \varphi<\varphi_*=\max_{z\in R^r}E(z,P)$, що

$$C_{\varphi}(P)\subset V_{\hat{\lambda}}(\tau):=\{z\in R^r:\|z-\hat{\lambda}\|\le\tau\};$$

Доведення. Властивість унікальної максимізації в центрі симетрії означає, що для $\forall\ \varphi\in[\varphi_a,\varphi_*)$ відображення d_φ приймає значення в R_0^+. Таким чином, $\exists\ t_0(\varphi)\in W^{r-1}$, що $d_\varphi(t)\ge d_\varphi(t_0(\varphi))=\rho_\varphi>0$. Тобто, для $\forall\ \varphi\in[\varphi_a,\varphi_*)$, ми маємо $V_{\hat{\lambda}}(\rho_\varphi)\subset C_\varphi(P)$, що доводить результат для цих значень φ. Тому, як слідує з властивості монотонності відносно найглибшої точки, вкладення $C_\varphi(P)$ визначає результат для $\forall\ \varphi<\varphi_*$.

Далі зафіксуємо таке $a>0$, що $z\mapsto E(z,P)$ є неперервною функцією на $V_{\hat{\lambda}}(a)$ та зауважимо, що існування φ_* слідує з властивості максимального значення в центрі, а існування a гарантується властивістю неперервності [2,3]. Остання властивість означає, що $z\mapsto E(z,P)$ досягає мінімуму в $V_{\hat{\lambda}}(a)$, а властивість унікальної максимізації в центрі симетрії припускає, що це мінімальне значення $\varphi_a<\varphi_*$. Використовуючи властивість неперервності, ми отримуємо, що для кожного $\forall\ \varphi\in[\varphi_a,\varphi_*]$,

$$d_\varphi:W^{r-1}\to R^+\ t\mapsto\sup\{d\in R^+:\hat{\lambda}+dt\in C_\varphi(P)\}\quad(6)$$

є неперервною функцією, що поточково сходиться до $d_{\varphi_*}(t)\equiv0$ при $\varphi\to\varphi_*$. Дана збіжність є рівномірною, тобто $\sup_{t\in W^{r-1}}\left|d_\varphi(t)\right|=o(1)$ при $\varphi\to\varphi_*$, оскільки W^{r-1} є компактом. Теорему доведено.

Теорема 2. Нехай P є ймовірнісною мірою, симетричною в околі $\hat{\lambda}$ та має щільність, що є додатньою в $\hat{\lambda}$. Також припустимо, що функція глибини E задовольняє властивостям максимального значення в центрі, монотонності відносно найглибшої точки, неперервності, узгодженості та унікальної максимізації в центрі симетрії. Нехай $N_{\hat{\lambda}}^{\gamma_m(m)}$ є кількістю глибинних найближчих сусідів в $C_{\hat{\lambda}}^{\gamma_m}(P^{(m)})$, де $\gamma_m=k_m/m$ грунтується на послідовності k_m, $P^{(m)}$ означає емпіричний розподіл $Z_1,...,Z_m$, де $Z_1,...,Z_m$ є незалежними та однаково розподіленими випадковими величинами P, а $Z_{\hat{\lambda},(i)}$ - $i-$й глибинний найближчий сусід $\hat{\lambda}$. Тоді, для $\forall\ \tau>0$, $\exists\, m=m(\tau)$, що $\sum_{i=1}^{N_{\hat{\lambda}}^{\gamma_m(m)}}\Lambda[\|Z_{\hat{\lambda},(i)}-\hat{\lambda}\|>\tau]=0$ майже напевно для $\forall\ m\ge m(\tau)$.

Доведення. Як слідує з теореми 1, $\exists\ \varphi<\varphi_*$, що $C_\varphi(P)\subset V_{\hat{\lambda}}(\tau)$ для фіксованого $\tau>0$. Також візьмемо такі $\bar{\varphi}$ та $\mu>0$ що $\varphi<\bar{\varphi}-\mu<\bar{\varphi}+\mu<\varphi_*$. Оскільки $P_{\hat{\lambda}}^{(m)}\to P_{\hat{\lambda}}=P$ при $m\to\infty$, де $P_{\hat{\lambda}}^{(m)}$ та $P_{\hat{\lambda}}$ є $\hat{\lambda}-$симетризованими версіями $P^{(m)}$ та P, відповідно, можна стверджувати, що $\exists$ таке ціле m_0, що

$$C_{\bar{\varphi}+\mu}(P)\subset C_{\bar{\varphi}}(P_{\hat{\lambda}}^{(m)})\subset C_{\bar{\varphi}-\mu}(P)\subset C_\varphi(P)\qquad(7)$$

майже напевно для всіх $\forall\ m\ge m_0$.

Крім того, $\exists\ \rho>0$, що $V_{\hat{\lambda}}(\rho)\subset C_{\bar{\varphi}+\mu}(P)$, тому для $\forall\ m\ge m_0$, має місце таке вкладення:

$$V_{\hat{\lambda}}(\rho)\subset C_{\bar{\varphi}}(P_{\hat{\lambda}}^{(m)})\subset V_{\hat{\lambda}}(\tau)\qquad(8)$$

майже напевно.

Отже, виходячи з неперервності, $Z\approx P$ має щільність, що є додатньою в околі $\hat{\lambda}$ [4]. Тому, $M_m/m\to P[V_{\hat{\lambda}}(\rho)]=P[V_{\hat{\lambda}}(\rho)]>0$ при $m\to\infty$ для $M_m=\sum_{i=1}^m\Lambda[Z_i\in V_{\hat{\lambda}}(\rho)]$. Крім того, для $\forall\ m\ge\bar{m}_0(\ge m_0)$, $\sum_{i=1}^m\Lambda[Z_i\in C_{\bar{\varphi}}(P_{\hat{\lambda}}^{(m)})]\ge M_m\ge k_m$ майже напевно, оскільки $k_m=o(m)$ при $m\to\infty$.

В результаті,

$$C_{\hat{\lambda}}^{\gamma_m}(P^{(m)})=C^{\gamma_m}(P_{\hat{\lambda}}^{(m)})\subset C_{\bar{\varphi}}(P_{\hat{\lambda}}^{(m)})\subset V_{\hat{\lambda}}(\tau)$$

майже напевно для таких значень m, де $\gamma_m = k_m / m$. Отже, $\max_{i=1,...,N_{\lambda}^{\gamma_m(m)}} \|Z_{\lambda,(i)} - \lambda\| \le \tau$ майже напевно для великих m, звідки слідує результат. Теорему доведено.

Зауважимо, що величина

$$\sum_{i=1}^{N_{\lambda}^{\gamma_m(m)}} \Lambda[\|Z_{\lambda,(i)} - \lambda\| > \tau] = 0 \quad (9)$$

є завжди визначеною на відміну від $Z_{\lambda,(i)}$.

Теорема 3. Нехай k_m є послідовністю таких додатніх чисел, що $k_m \to \infty$ та $k_m = o(m)$ при $m \to \infty$. Крім того, нехай E є функцією глибини, що задовольняє властивостям максимальності в центрі, монотонності по відношенню до найглибшої точки, а також властивостям неперервності, унікальної максимізації в центрі симетрії та узгодженості. Припустимо, що $Z \mid [X = l]$ має функцію щільності h_l, множина точок розриву якої має міру Лебега нуль для $l = 0,1$. Тоді, якщо $\sum_m$ є сигма-алгеброю, пов'язаною з (Z_i, X_i), $i = 1,...,m$, має місце узгодженість глибинного класифікатора k_m-найближчих сусідів $n_E^{(m)}$ в (3), тобто:

$$P\left[n_E^{(m)}(Z) \ne X \mid \sum_m\right] - P\left[n_B(Z) \ne X\right] = o_P(1), \quad (10)$$

при $m \to \infty$.

Доведення. Припустимо, що

$$\mathrm{Supp}_+(h) = \{z \in \mathrm{R}^r : h(z) > 0\},$$

а також визначимо $S(h_l)$ для множини точок неперервності h_l, $l = 0,1$, де

$$M = \mathrm{Supp}_+(h) \cap S(h_0) \cap S(h_1). \quad (11)$$

Зазначимо, що Z має функцію щільності $z \mapsto h(z) = p_0 h_0(z) + p_1 h_1(z)$, що слідує з теореми Байєса [5].

Отже, при $P[Z \in M] = 1$ ми маємо, що

$$P\left[Z \in \mathrm{R}^r \setminus M\right] \le P\left[Z \in \mathrm{R}^r \setminus \mathrm{Supp}_+(h)\right] + \\ + \sum_{l \in \{0,1\}} P\left[Z \in \mathrm{R}^r \setminus S(h_l)\right] = \int_{\mathrm{R}^r \setminus \mathrm{Supp}_+(h)} h(z)dz = 0, \quad (12)$$

оскільки $\mathrm{R}^r \setminus S(h_l)$ має нульову міру Лебега, де $l = 0,1$.

Далі припустимо, що $z \in M$, де

$$z \mapsto \theta(z) = p_1 h_1(z) / (p_0 h_0(z) + p_1 h_1(z))$$

є неперервною функцією на M. Оцінку $\bar{\theta}_E^{(m)}$ з (3) можна записати у такій формі:

$$\bar{\theta}_E^{(m)} = \sum_{i=1}^{m} X_i D_i^{\gamma(m)}(z) = \frac{1}{N_z^{\gamma(m)}} \sum_{i=1}^{N_z^{\gamma(m)}} X_{z,(i)}, \quad (13)$$

припускаючи, що $X_{z,(i)} = X_{l(z)}$, де $l(z)$ є таким, що $Z_{z,(i)} = Z_{l(z)}$.

Отже,

$$Q^{(m)}(z) := \Omega[\bar{\theta}_E^{(m)}(z) - \theta(z))^2] \le 2Q_1^{(m)}(z) + 2Q_2^{(m)}(z), \quad (14)$$

де

$$Q_1^{(m)}(z) = \Omega\left[\left|\frac{1}{N_z^{\gamma(m)}} \sum_{i=1}^{N_z^{\gamma(m)}} (X_{z,(i)} - \theta(Z_{z,(i)}))\right|^2\right] \quad (15)$$

та

$$Q_2^{(m)}(z) = \Omega\left[\left|\frac{1}{N_z^{\gamma(m)}} \sum_{i=1}^{N_z^{\gamma(m)}} (\theta(Z_{z,(i)}) - \theta(z))\right|^2\right]. \quad (16)$$

Зауважимо, що величину γ_m замінено на γ для спрощення запису. Крім того, $X_{z,(i)} - \theta(Z_{z,(i)})$, $i = 1,...,m$ є нульовими середніми значеннями взаємно незалежних випадкових величин, на що вказує $\sum_Z^{(m)}$ для сигма-алгебри, породженої Z_i, $i = 1,...,m$ [6].

В результаті, використовуючи той факт, що $N_z^{\gamma(m)} \ge k_m$ майже напевно, ми отримуємо таку рівність:

$$Q_1^{(m)}(z) = \\ = \Omega\left[\frac{1}{(N_z^{\gamma(m)})^2} \sum_{i,l=1}^{N_z^{\gamma(m)}} \Omega\left[(X_{z,(i)} - \theta(Z_{z,(i)}))(X_{z,(l)} - \theta(Z_{z,(l)})) \mid \sum_Z^{(m)}\right]\right] = \\ = \Omega\left[\frac{1}{(N_z^{\gamma(m)})^2} \sum_{i=1}^{N_z^{\gamma(m)}} \Omega\left[(X_{z,(i)} - \theta(Z_{z,(i)}))^2 \mid \sum_Z^{(m)}\right]\right] \le \\ \le \Omega\left[\frac{4}{N_z^{\gamma(m)}}\right] \le \frac{4}{k_m} = o(1), \quad (17)$$

при $m \to \infty$.

Далі, використовуючи нерівність Коші-Шварца, має місце така нерівність:

$$Q_2^{(m)}(z) \le \Omega\left[\frac{1}{N_z^{\gamma(m)}} \sum_{i=1}^{N_z^{\gamma(m)}} (\theta(Z_{z,(i)}) - \theta(z))^2\right] = \\ = \Omega\left[\frac{1}{N_z^{\gamma(m)}} \sum_{i=1}^{N_z^{\gamma(m)}} (\theta(Z_{z,(i)}) - \theta(z))^2 \Lambda[\|Z_{z,(i)} - z\| \le \tau]\right] + \\ + \Omega\left[\frac{1}{N_z^{\gamma(m)}} \sum_{i=1}^{N_z^{\gamma(m)}} (\theta(Z_{z,(i)}) - \theta(z))^2 \Lambda[\|Z_{z,(i)} - z\| > \tau]\right] \le \\ \le \sup_{z' \in V_z(\tau)} |\theta(z') - \theta(z)|^2 + 4\Omega\left[\frac{1}{N_z^{\gamma(m)}} \sum_{i=1}^{N_z^{\gamma(m)}} \Lambda[\|Z_{z,(i)} - z\| > \tau]\right] := \\ := \bar{Q}_2(z;\tau) + \bar{Q}_2^{(m)}(z;\tau) \quad (18)$$

для $\forall\ \tau > 0$.

Тому, для $\forall\ \mu > 0$ можна вибрати таке $\tau = \tau(\mu) > 0$, що $\bar{Q}_2(z; \tau(\mu)) < \mu$. Даний вивід слідує з неперервності θ в z. Крім того очевидно, що $Q_2^{(m)}(z;\tau(\mu))=0$ для великих m, а оскільки $Q_2^{(m)}(z)$ є $o(1)$, те ж саме має місце і для $Q^{(m)}(z)$.

Отже, при $m\to\infty$,

$$\begin{aligned}&\Omega\left[\left|P[\overline{n}_E^{(m)}(Z)\neq X\mid\Sigma_m]-J_*\right|\right]=\\&=\Omega\left[P[\overline{n}_E^{(m)}(Z)\neq X\mid\Sigma_m]-J_*\right]=\\&P[\overline{n}_E^{(m)}(Z)\neq X]-J_*\leq\\&\leq 2\left(\Omega\left[\overline{\theta}_E^{(m)}(Z)-\theta(Z))^2\right]\right)^{1/2}=o(1),\end{aligned}\quad (19)$$

оскільки $P[\overline{n}_E^{(m)}(Z)\neq X\mid\Sigma_m]\geq J_*$ майже напевно. Теорему доведено.

Висновки. Запропонований метод глибинних околів може бути ефективно застосований не лише для задач класифікації, а й для задач оцінки загальної щільності h незалежних та однаково розподілених випадкових r-векторів Z_i, де $i=1,...,m$. Крім того, використання глибинних околів представляє інтерес в регресійних задачах, де умовне середнє значення функції $z\mapsto n(z)=\Omega[X\mid Z=z]$ оцінюється на основі взаємно незалежних групп (Z_i,X_i) випадкового вектора (Z,X) зі значеннями в $\mathrm{R}^r\times\mathrm{R}$, де $i=1,...,m$. Дослідження властивостей даних глибинних методів виходить за рамки даної роботи, однак, досить ймовірним є те, що властивості узгодженості, які отримані для задач класифікації також поширюються на задачі регресії та оцінки щільності. В задачах оцінки щільності використання несферичних, а саме еліпсоїдних околів може привести до фундаментальних скінченно-вибіркових властивостей. У даному випадку глибинні оцінки k_m найближчих сусідів є досить ефективними, оскільки вони використовують несферичні та нееліпсоїдальні околи, форма яких визначається локальною геометрією вибірки.

Список використаних джерел

1. *Holmes C.C.* A probabilistic nearest neighbor method for statistical pattern recognition / C.C. Holmes, N.M. Adams // Journal of the Royal Statistical Society. – 2002. – 64. – P. 295-306.
2. *Godtliebsen F.* Significance in scale space for bivariate density estimation / F. Godtliebsen, J.S. Marron, P. Chaudhuri // Journal of Computational and Graphical Statistics. – 2002. – 11. – P. 1-22.
3. *Mosler K.* Multivariate Dispersions, Central Regions and Depth / K. Mosler. – New York: Springer-Verlag, 2002, P. 1–291.
4. *Silverman B.W.* Density estimation for Statistics and Data Analysis / B.W. Silverman. – London: Chapman and Hall, 1986. 1–175.
5. *Hoeffding W.* Probability inequalitites for sums of bounded random variables // Journal of the American Statistical Association. – 1963. – 58. – P. 14–27.
6. *Zuo Y.* Structural properties and convergence results for contours of sample statistical depth functions / Y. Zuo, R. Serfling // The Annals of Statistics. – 2000. – 28. – P. 484–497.

References

1. HOLMES, C.C. and ADAMS N.M. (2002) A probabilistic nearest neighbor method for statistical pattern recognition. *Journal of the Royal Statistical Society*. 64. p. 295-306.
2. GODTLIEBSEN, F., MARRON, J.S., CHAUDHURI P. (2002) Significance in scale space for bivariate density estimation. *Journal of Computational and Graphical Statistics*. 11. p. 1-22.
3. MOSLER, K. (2002) Multivariate Dispersions, Central Regions and Depth, New York: Springer-Verlag. p. 1- 291.
4. SILVERMAN, B.W. (1986) Density estimation for Statistics and Data Analysis, London: Chapman and Hall. p. 1-175.
5. HOEFFDING, W. (1963) Probability inequalitites for sums of bounded random variables. *Journal of the American Statistical Association*. 58. p. 14-27.
6. ZUO, Y., SERFLING R. (2000) Structural properties and convergence results for contours of sample statistical depth functions. *The Annals of Statistics*. 28. p. 484-497.

Надійшла до редколегії 16.11.15