УДК 519.962.22

МІНІМАКСНІ СЕРЕДНЬОКВАДРАТИЧНІ ОЦІНКИ ТРЕНДУ В ЗАДАЧАХ РЕГРЕСІЇ

© Демиденко С.В., Наконечний О.Г.

Київський національний університет ім. Т.Г.Шевченка,
факультет кібернетики
пр-т академіка Глушкова-2, корпус 6, м. Київ 03680, Україна
e-mail: *s.demidenko@gmail.com*

Abstract. This paper describes an approach to the linear minimax estimation of the generalized polinom with unknown partially unbounded parameters. Linear minimax estimations are constructed on the basis of measurements with random noise. We introduce notions of the upper and lower linear minimax estimations. The sufficient conditions for the minimax estimation existence are formulated. Numerical simulation that illustrates these results is presented.

Вступ

Проблеми апроксимації функції тісно зв'язані із задачами оптимізації [1, 2] з одного боку, а також із питаннями оцінки функцій за даними спостережень, що досліджувалися в математичній статистиці [3, 4].

У даній роботі розглядається випадок, коли спостерігається узагальнений поліном із шумом. При цьому кореляційна функція випадкового процесу що моделює шум невідома і належить певній області. При обмеженнях на коефіцієнти розроблено алгоритм побудови гарантованої лінійної оцінки такого полінома.

1. Постановка задачі

Нехай на відрізку $[0, T]$ спостерігається реалізація випадкового процесу $y(t)$, що має вигляд

$$y(t) = \sum_{i=1}^{m} \alpha_i \varphi_i(t) + \xi(t) \tag{1}$$

де $\varphi_i(t)$ $i = \overline{1..m}$ – неперервні на відрізку $[0, T]$ функції, $\xi(t)$ – реалізація неперервного в середньому квадратичному випадкового процесу з нульовим середнім та невідомою кореляційною функцією $R(t, s) \in K$

$$K = \{R : \int_0^T \int_0^T (R(t, s) - R_0(t, s))^2 dtds \le q^2\} \tag{2}$$

$R_0(t, s)$ – відома кореляційна функція неперервна на $[0, T] \times [0, T]$. Нехай також відомо, що коефіцієнти $\alpha_i, i = \overline{1..r}, r \le m$ належать множині G, де

$$G = \{\alpha : |\alpha_i - \widehat{\alpha}_i| \le \beta_i, i = \overline{1..r}\} \tag{3}$$

Позначимо через $\widehat{P(T)}$ лінійну оцінку тренду $P(T) = \sum\limits_{i=1}^{m} \alpha_i \varphi_i(T)$ вигляду $\widehat{P(T)} = \int\limits_0^T u(t)y(t)dt + c$, де $u(t)$ належить простору вимірних за Лебегом інтегрованих з квадратом функцій, c – довільна константа. Позначимо через $\sigma(R, \alpha, u, c)$ середньоквадратичну похибку такої оцінки, тобто

$$\sigma^2(R, \alpha, u, c) = M(P(T) - \widehat{P(T)})^2 \tag{4}$$

Означення 1. Оцінка для якої $\widehat{u}$ та $\widehat{c}$ належать множині $(\widehat{u}, \widehat{c}) \in \arg\min \sigma_1^2(u, c)$, де

$$\sigma_1^2(u, c) = \sup_{G,K} \sigma^2(R, \alpha, u, c)$$

називається мінімаксною середньоквадратичною оцінкою.

2. Умови існування та вигляд мінімаксної оцінки

Твердження 1. Припустимо, що $\varphi_i(t)$, $i = \overline{(r+1)..m}$, лінійно незалежні, тоді існує єдина мінімаксна середньоквадратична оцінка причому

$$\inf_{u,c} \sigma_1^2(u, c) = \inf_{u \in U} \sigma_1^2(u, \widehat{c}) = \sigma_2^2(\widehat{u}) \tag{5}$$

де $\sigma_2^2(u) = \sigma_1^2(u, \sum_{i=1}^r \widehat{\alpha}_i z_i(0))$, $\widehat{u} = \arg\min_{\widehat{u} \in U} \sigma_2^2(u)$, U множина, що визначається із умови

$$U = \{u : z_i(0) = 0, i = \overline{(1+r)..m}\} \tag{6}$$

а $\widehat{c} = \sum\limits_{i=1}^{r} \widehat{\alpha}_i \widehat{z}_i(0)$, $\widehat{z}_i(0) = z_i(0)|_{u=\widehat{u}}$, де $z_i(0) = \varphi_i(T) - \int\limits_0^T u(t)\varphi_i(t)dt$.

Доведення. Зауважимо, що має місце рівність

$$\sigma^2(R, \alpha, u, c) = (\sum_{i=1}^{m} \alpha_i z_i(0) - c)^2 + \int_0^T \int_0^T R(t,s)u(t)u(s)dtds \tag{7}$$

$$\sigma_1^2(u, c) = \sup_G (\sum_{i=1}^{m} \alpha_i z_i(0) - c)^2 + \sup_R \int_0^T \int_0^T R(t,s)u(t)u(s)dtds = I_1(u,c) + I_2(u) \tag{8}$$

Крім того $I_1 = \infty$ при $u \notin U$ і

$$I_1(u, c) = \left(\sum_{i=1}^{r} \beta_i |z_i(0)| + \left| c - \sum_{i=1}^{r} \widehat{\alpha}_i z_i(0) \right| \right)^2 \tag{9}$$

при $u \in U$.

$$I_2(u) = \int_0^T \int_0^T R_0(t,s)u(t)u(s)dtds + q^2 \int_0^T u^2(t)dt \tag{10}$$

Далі зауважимо, що так як функції $\varphi_i(t)$ лінійно незалежні, то множина U є непорожньою опуклою замкненою в $L_2(0,T)$ множиною. Функціонал

$$\sigma_2^2(u) = \left(\sum_{i=1}^{r} \beta_i |z_i(0)|\right)^2 + I_2(u) \tag{11}$$

є слабконапівнеперервним, сильно опуклим функціоналом і значить існує єдина функція $\widehat{u} \in \arg\min \sigma_2^2$, в силу нерівності

$$\sigma_1^2(u,c) \geq \sigma_2^2(u) \geq \min_{u \in U} \sigma_2^2(u) = \sigma_2^2(\widehat{u}) \tag{12}$$

яка перетворюється в рівність при $u = \widehat{u}$, $c = \widehat{c}$, одержимо необхідне. □

Позначимо далі через V множину матриць B, що мають вигляд

$$\begin{aligned} &B = (b_{ij})_{i,j=\overline{1,r}} \\ &b_{ij} = \int_{-1}^{1} .. \int_{-1}^{1} x_i x_j \nu(dx) \end{aligned} \tag{13}$$

$\nu(\cdot)$ – пробігає множину ймовірносних мір, що зосереджені на гіперкубі $[-1..1] \times .. \times [-1..1]$.

Твердження 2. Має місце рівність

$$\min_{u \in U} \sigma_2^2(u) = \max_{B \in V} \min_{u \in U} \sigma_3^2(u,B) \tag{14}$$

де

$$\begin{aligned} &\sigma_3^2(u,B) = \sum_{i,j=1}^{r} \beta_i \beta_j b_{ij} z_i(0) z_j(0) + I_2(u) = (Bz_1(0), z_1(0)) + I_2(u) \\ &z_1(0) = (\beta_1 z_1(0), .., \beta_r z_r(0)) \end{aligned} \tag{15}$$

Доведення. Так як

$$\begin{aligned} &\left(\sum_{i=1}^{r} \beta_i |z_i(0)|\right)^2 = \max_{-1 \leq x_i \leq 1} \left(\sum_{i=1}^{r} \beta_i x_i z_i(0)\right)^2 = \\ &= \max_{\nu(\cdot)} \int_{-1}^{1} .. \int_{-1}^{1} \left(\sum_{i=1}^{r} \beta_i x_i z_i(0)\right)^2 \nu(dx) = \max_{B \in V} (Bz_1(0), z_1(0)) \end{aligned} \tag{16}$$

то в силу теореми про мінімакс одержимо необхідну рівність. □

Твердження 3. Має місце рівність

$$\min_{u \in U} \sigma_3^2(u,B) = \sigma_3^2(\widehat{u},B) \tag{17}$$

Доведення. Нехай $z_i(t)$ є розв'язком рівняння

$$\frac{z_i(t)}{dt} = -\varphi_i(t)u(t), z_i(T) = \varphi_i(T), i = \overline{1,..,m}, z_i(0) = 0, i = \overline{(r+1),..,m} \qquad (18)$$

Введемо також функції $p_i(t)$, що є розв'язками рівнянь

$$\frac{p_i(t)}{dt} = 0, i = \overline{1,..,m} \qquad (19)$$

$$p_i(0) = \sum_{j=1}^{r} b_{ij}\beta_j z_j(0), i = \overline{1,..,r} \qquad (20)$$

Тоді похідна Гато від функціонала $\sigma_3^2(u, B)$ буде мати вигляд

$$\frac{1}{2}\left(\sigma_3^2\right)' = -\sum_{i=1}^{m} \varphi_i(t)p_i(t) + \int\limits_0^T R_0(t,s)u(s)ds + q^2u(t) \qquad (21)$$

Таким чином для $\widehat{u}(s)$ одержимо інтегральне рівняння

$$\int\limits_0^T R_0(t,s)\widehat{u}(s)ds + q^2\widehat{u}(t) = \sum_{i=1}^{m} \varphi_i(t)p_i(t) \qquad (22)$$

Запишемо $\widehat{u}$ у іншому вигляді, для цього зауважимо, що $p_i(t)$ не залежить від t і тому якщо позначити їх через p_i, то одержимо, що

$$\widehat{u}(t) = \sum_{i=1}^{m} p_i\psi_i(t) \qquad (23)$$

де $\psi_i(t)$ є розв'язком інтегрального рівняння

$$\int\limits_0^T R_0(t,s)\psi_i(s)ds + q^2\psi_i(t) = \varphi_i(t) \qquad (24)$$

Тоді $z_i(t) = -\sum\limits_{j=1}^{m} p_i z_{ij}(t) + \varphi_i(T)$, де $z_{ij}(t) = \int\limits_t^T \varphi_i(s)\psi_j(s)ds.$

Так як

$$p_i = p_i(0) = \sum_{j=1}^{r} b_{ij}\beta_j z_j(0), \quad i = \overline{1,..,r},$$

то для $z_j(0)$ одержимо систему рівнянь

$$z_i(0) = -\sum_{j=1}^{m}\sum_{k=1}^{r} b_{ik} z_{ij}(0)\beta_k z_k(0) + \varphi_i(T) \qquad (25)$$

а константи p_i, $i = \overline{(r+1)..m}$ визначаються із системи рівнянь

$$\sum_{i=1}^{m} \int_0^T \varphi_j(t)\psi_i(t)dtp_i = \varphi_j(T), j = \overline{r+1,..,m} \tag{26}$$

□

Наслідок 1. Нехай r=0. Тоді мінімаксна оцінка має вигляд

$$\widehat{\widehat{P}} = \sum_{i=1}^{m} p_i \int_0^T \psi_i(t)y(t)dt \tag{27}$$

Величини p_i визначаються з системи рівнянь

$$\sum_{i=1}^{m} p_i \int_0^T \varphi_j(t)\psi_i(t)dt = \varphi_j(T), j = \overline{1,..,m} \tag{28}$$

Наслідок 2. Припустимо, що $R_0(t,s) = 0$. Тоді $\psi_i(t) = q^{-2}\varphi_i(t)$.

Наслідок 3. Мінімаксна оцінка має вигляд

$$\widehat{\widehat{P}} = \sum_{i=1}^{m} \widehat{p_i} \int_0^T \psi_i(t)y(t)dt + \widehat{c} \tag{29}$$

числа $\widehat{p_i}$ знаходяться з розв'язку задачі на мінімум наступної функції

$$F(p_1,..,p_m) = \left(\sum_{k=1}^{r} \beta_k |\varphi_k(T) - \sum_{i=1}^{m} p_i z_{ki}(0)|\right)^2 + \sum_{i,j=1}^{m} p_i p_j (R_0\psi_i,\psi_j) + q^2 \sum_{i,j=1}^{m} p_i p_j (\psi_i,\psi_j) \tag{30}$$

де

$$(R\psi_i,\psi_j) = \int_0^T .. \int_0^T R_0(t,s)\psi_i\psi_j dtds, (\psi_i,\psi_j) = \int_0^T .. \int_0^T \psi_i\psi_j dtds, \tag{31}$$

при обмеженнях

$$\varphi_i(T) = \sum_{k=1}^{m} p_k \int_0^T \varphi_i(t)\psi_k(t)dt, \quad i = \overline{r+1,..,m} \tag{32}$$

Доведення. Із твердження 3 одержимо наступний вигляд функції $\widehat{u}(t)$ на якій досягається мінімум середньоквадратичної похибки оцінки $\widehat{P}(T)$, $\widehat{u}(t) = \sum_{i=1}^{m} \widehat{p}_i\psi_i(t)$, враховуючи вигляд функціоналу $\sigma^2(u,c)$ одержимо необхідне. □

Приклад 1. Проілюструємо застосування наслідку 3 на прикладі наступної моделі $\varphi_1(t) = \sin(t), \varphi_2(t) = \sin(t^2), \varphi_3(t) = \cos(t), \beta_1 = 2, R_0 = 0, \widehat{\alpha_1} = 0$. Випадковий процес $\xi(t)$ має вигляд $\xi(t) = \xi_1 * sin(20 * t)$, де ξ_1 – випадкова величина, що має нормальний розподіл з параметрами $(0, 0.2)$. Нехай $\alpha = (1\ 1\ 1)^T$ і реалізація випадкового процесу $\xi(t)$ має вигляд $\xi(t) = -0.226 * \sin(20 * t)$. Процедура оцінювання буде проводитися в умовах коли всі зазначені вище параметри фіксовані, а параметри T та q будуть варіюватися в межах $T \in [1, 3]$ та $q \in [0.2, 2.2]$ з кроком 0.05 та 0.4. Відповідний алгоритм було реалізовано в системі Mathematica 6, результати обчислень зображені на наступних графіках.

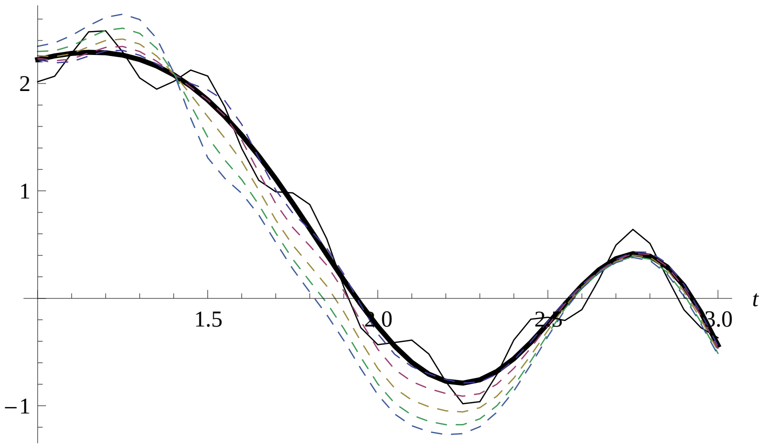

Рис. 1. Реальні значення $P(T)$ (суцільна лінія), реалізація $y(t)$ (суцільна тонка лінія), оцінки $\widehat{P(T)}$ при різних значеннях параметра q (штрихові лінії)

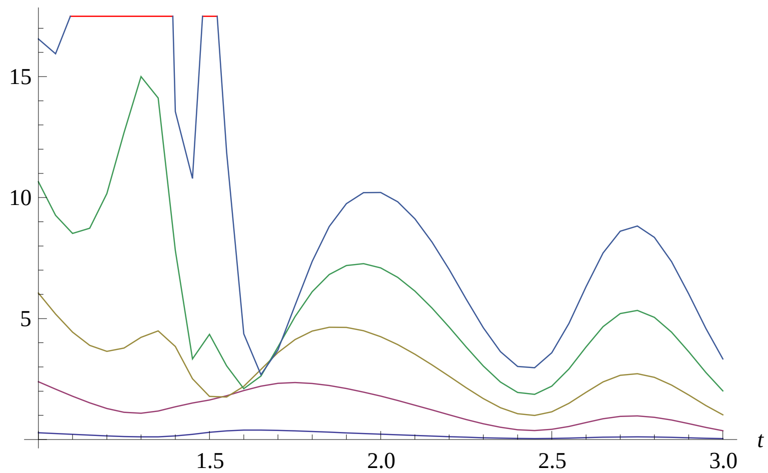

Рис. 2. Мінімаксна середньоквадратична похибка при різних значеннях параметра q

Отже, як видно з малюнку 1, при зменшенні параметра q якість оцінювання покращується, при $q = 0.2$ оцінки розташовані найближче до реальних значень тренду. При цьому найменші значення мінімаксної середньоквадратичної похибки теж досягаються при $q = 0.2$ (мал. 2).

3. Верхня та нижня мінімаксні оцінки

Означення 2. Оцінка $\widehat{\widehat{P}}_+(T)$ та $\widehat{\widehat{P}}_-(T)$ називаються верхньою та нижньою мінімаксною якщо вони мають вигляд

$$\begin{aligned} \widehat{\widehat{P}}_+(T) &= \int_0^T \widehat{u}_+(t)y(t)dt + \widehat{c}_+ \\ \widehat{\widehat{P}}_-(T) &= \int_0^T \widehat{u}_-(t)y(t)dt + \widehat{c}_- \end{aligned} \tag{33}$$

де $(\widehat{u}_+, \widehat{c}_+)$, $(\widehat{u}_-, \widehat{c}_-)$ знаходяться із мінімізації функціоналів $\sigma_+^2(u, c), \sigma_-^2(u, c)$ відповідно і для яких справедливі нерівності

$$\sigma_-^2(u, c) \le \sigma^2(u, c) \le \sigma_+^2(u, c) \tag{34}$$

Твердження 4. Оцінки для яких має місце рівність

$$\begin{aligned} \widehat{u}_+(t) &= \sum_{i=1}^m p_i^+ \psi_i(t), \\ \widehat{u}_-(t) &= \sum_{i=1}^m p_i^- \psi_i(t) \end{aligned} \tag{35}$$

для яких числа $p_i^\pm, i = \overline{1, .., m}$ знаходяться із систем рівнянь

$$\psi_i(T) = \sum_{i=1}^m p_k^\pm + z_{ik}(0) + \delta_{ir} p_i^\pm (\beta_i^\pm)^2 \tag{36}$$

де

$$\delta = \begin{cases} 1, & i \le r \\ 0, & i > r \end{cases}$$

є відповідно нижніми та верхніми мінімаксними оцінками.

Доведення. Введемо множини G_+ та G_-

$$\begin{aligned} G_+ &= \{\alpha : \sum_{i=1}^r \frac{(\alpha_i - \widehat{\alpha}_i)^2}{\beta_i^{+2}} \le 1\} \\ G_- &= \{\alpha : \sum_{i=1}^r \frac{(\alpha_i - \widehat{\alpha}_i)^2}{\beta_i^{-2}} \le 1\} \end{aligned} \tag{37}$$

числа β_i^+ та β_i^- виберемо так, щоб $G_- \subset G \subset G_+$. Тоді очевидно, що

$$\begin{aligned} \sigma_+^2(u,c) &= \sup_{G_+,R} \sigma^2(R,\alpha,u,c) \\ \sigma_-^2(u,c) &= \sup_{G_-,R} \sigma^2(R,\alpha,u,c) \end{aligned} \tag{38}$$

Легко бачити, що

$$\sigma_+^2(u,c) = ((\sum_{i=1}^{r}(\beta_i^+)^2 z_i^2(0))^{\frac{1}{2}} + |c - \sum_{i=1}^{r} \widehat{\alpha}_i z_i(0)|)^2 + (R_0u,u) + q^2(u,u) \tag{39}$$

Звідки $\widehat{c}_+ = \sum_{i=1}^{r} \widehat{\alpha}_i \widehat{z}_i^+(0)$, де $z_i^+(t)$ є розв‘язком рівняння

$$\frac{d\widehat{z}_i^+}{dt} = -\varphi_i(t)\widehat{u}^+(t), \widehat{z}_i^+(T) = \varphi_i(T) \tag{40}$$

де $\widehat{u}^+(t) \in \arg\min \sigma_+^2(u)$, $\sigma_+^2 = \sum_{i=1}^{r}(\beta_i^+)^2 z_i^2(0) + (R_0u,u) + q^2(u,u)$

Звідки як і в твердженні 3 одержимо, що

$$\widehat{u}_+(t) = \sum_{i=1}^{m} p_i^+ \psi_i(t) \tag{41}$$

де p_i^+ знаходиться із системи рівнянь

$$\begin{cases} p_j^+(\beta_j^+)^2 = \varphi_j(T) - \sum_{i=1}^{m} p_i^+ z_{ij}(0), \ j = \overline{1,..,r} \\ \varphi_j(T) = \sum_{i=1}^{m} p_i^+ z_{ij}(0), \ j = \overline{1+r,..,m} \end{cases}$$

Зауваження 1. Нехай множина значень параметрів є G_+, тоді мінімаксна оцінка співпадає із верхньою мінімаксною оцінкою.

□

Висновки

В роботі показано, що при деяких умовах існують єдині мінімаксні середньоквадратичні оцінки узагальнених поліномів. Дані оцінки знайдені у вигляді лінійної комбінації розв'язків інтегральних рівнянь Фредгольма другого роду. Вводяться поняття верхніх і нижніх мінімаксних оцінок та наводиться їх вигляд. При певних умовах на коефіцієнти розроблено алгоритми для обчислення оцінок. В системі Mathematica 6 створено відповідну програмну реалізацію.

Список літератури

1. *Лоран Г.Ж.* Апроксимация и оптимизация. – М.: Мир, 1975. – 496 с.
2. *Тихомиров В.М.* Некоторые вопросы теории приближений. – М.: МГУ, 1980. – 304 с.
3. *Себер Д.* Линейный регрессионный анализ. – М.: Мир, 1980. – 456 с.
4. *Норман Дрейпер, Гарри Смит* Прикладной регрессионный анализ. – 3-е изд. – М.: «Диалектика», 2007. – С. 912.
5. *Полянин А. Д., Манжиров А. В.* Справочник по интегральным уравнениям. – М.: Физматлит, 2003. – 608 с.

Статья поступила в редакцию 13.10.2008