

MINIMAXOVÉ ODHADY V MODELU LINEÁRNÍ REGRESE

JAROSLAV MICHÁLEK

1. Úvod

Uvažujme lineární regresní model

$$Y = X\beta + e, \quad (1)$$

kde X je konstantní matice typu $n \times p$ (matice plánu), β je p -rozměrný vektor neznámých parametrů, Y je n -rozměrný náhodný vektor (vektor pozorování) a $e = (e_1, \dots, e_n)'$ je n -rozměrný vektor náhodných chyb.

Budeme předpokládat, že náhodné chyby e_i , $i = 1, 2, \dots, n$ jsou nezávislé, nesystematické a homogenní tj. $E e = O$ a $D e = \text{var } e = E e e' = \sigma^2 I_n$ (přičemž O značí nulový vektor případně nulovou matici – podle kontextu a I_n případně jenom I bude značit jednotkovou matici příslušného řádu, zde řádu n). Dále nechť e_i mají stejnou distribuční funkci F s hustotou f vzhledem k Lebesgueově míře. O distribuční funkci F budeme předpokládat, že patří do třídy distribučních funkcí \mathcal{F} a zatím ji nebudeme blíže specifikovat.

Dále budeme $\mathcal{M}_{n \times p}$ značit množinu reálných matic typu $n \times p$, $\mathcal{M}_n^>$ respektive \mathcal{M}_n^{\geq} množinu pozitivně definitních respektive pozitivně semidefinitních matic reálných čísel typu $n \times n$.

Odhad $\tilde{\beta}$ parametru β , který minimalizuje asymptotickou varianční matici vektoru $\tilde{\beta}$ pro nejméně příznivé rozdělení F_0 ze třídy \mathcal{F} se nazývá (asymptoticky) minimaxový odhad parametru β (viz [7], [12]). Je dobře známo (viz [7], [12]), že asymptoticky minimaxové odhady mají velmi dobré robustní vlastnosti. Např. v [6] je pro třídu \mathcal{F} ϵ -kontaminovaných symetrických rozdělení dokázáno za poměrně obecných podmínek, že rozdělení F , které minimalizuje Fisherovu informaci ve třídě \mathcal{F} , je rovno právě nejméně příznivému rozdělení této třídy a že takto získaný minimaxový odhad parametru β je M -odhadem s generující funkcí $\psi = -f'_0/f_0$. Ve [12] je potom uveden seznam různých tříd rozdělení \mathcal{F} , příslušné nejméně příznivé rozdělení F_0 a generující funkce ψ_0 příslušného M -odhadu, který je zároveň minimaxovým odhadem pro tuto třídu rozdělení minimalizující jeho asymptotickou varianční matici.

Uvedený minimaxový přístup je založen na apriorní informaci, že rozdělení náhodných chyb e_i patří do předem specifikované třídy \mathcal{F} a ztrátová funkce je zastoupena asymptotickou varianční maticí odhadu. Na rozdíl od tohoto přístupu Bunke v [2] volí odlišný minimaxový přístup a ukazuje jeho jisté robustní vlastnosti. Vychází z odhadu $\tilde{\gamma}$ parametrické funkce

$$\gamma = C\beta + c, \quad (2)$$

kde $C \in \mathcal{M}_{s \times p}$ a $c \in \mathbb{R}^s$ ($\mathbb{R} = (-\infty, \infty)$), vzhledem ke ztrátové funkci

$$Z(\vartheta, \tilde{\gamma}) = E_{\vartheta}(\tilde{\gamma} - \gamma)' Z(\tilde{\gamma} - \gamma), \quad (3)$$

kde $Z \in \mathcal{M}_n^{\geq}$ a E_{ϑ} je operátor střední hodnoty počítaný při dané hodnotě $\vartheta = (\beta, \sigma, F)$. Lineární nehomogenní odhad $\tilde{\gamma}$ konstruuje tak, že minimalizuje suprémum ztrátové funkce $\sup Z(\vartheta, \tilde{\gamma})$ přes všechny možné hodnoty ϑ . Uvedený přístup dále modifikuje Läuter v [9], který suprémum ztrátové funkce Z hledá přes specificky volenou třídu možných hodnot parametru β . Tomuto přístupu a jeho rozšíření bude dále věnována pozornost.

2. Minimaxové odhady lineární parametrické funkce při apriorní informaci o parametru β

Budeme uvažovat regresní model (1) s obecnou známou varianční maticí náhodných chyb $R = D\epsilon \in \mathfrak{M}_n^+$. Odhad parametrické funkce γ , která je dána vztahem (2), budeme hledat ve tvaru

$$\hat{\gamma} = UY + u,$$

kde $U \in \mathfrak{M}_{s \times n}$ a $u \in \mathbb{R}^s$, vzhledem ke ztrátové funkci (3), kterou nyní zapíšeme ve tvaru

$$Z(\beta, U, u) = E_{\beta}(\hat{\gamma} - \gamma)'Z(\hat{\gamma} - \gamma). \quad (4)$$

Dále budeme předpokládat, že je dána množina $G \subseteq \mathbb{R}^s$ a apriorní informace o parametru β dané vztahem $\beta \in G$.

Jestliže

$$\inf_{u \in \mathbb{R}^s, U \in \mathfrak{M}_{s \times n}} \sup_{\beta \in G} Z(\beta, U, u) = \sup_{\beta \in G} Z(\beta, \hat{U}, \hat{u}),$$

budeme říkat, že $\hat{\gamma} = \hat{U}Y + \hat{u}$ je minimaxový nehomogenní lineární odhad parametrické funkce $\gamma = C\beta + u$ a zkráceně jej budeme označovat MILE-odhad (minimax inhomogenous linear estimator).

Následující věta (viz [9], [3]) popisuje explicitní vyjádření MILE-odhadu pro apriorní informaci o parametru β tvaru

$$\beta \in G = \{\beta: \beta'Q\beta \leq 1\}, \quad Q \in \mathfrak{M}_s^+. \quad (5)$$

Věta 1. Necht' je dán regresní model (1) se známou varianční maticí náhodných chyb R . Položme

$$S = Q^{-1/2}X'R^{-1}XQ^{-1/2} \\ F = S^{-1}Q^{-1/2}C'ZCQ^{-1/2}S^{-1}.$$

Necht' existuje matice $A \in \mathfrak{M}_s^+$ a konstanta $a_0 > 0$ tak, že platí

- (i) $a_0^{-1/2}(F + A)^{1/2} \geq S^{-1}$
- (ii) $a_0^{1/2}(F + A)^{-1/2}A = SA$
- (iii) $a_0^{-1/2}\text{Tr}(F + A)^{1/2} = 1 + \text{Tr}S^{-1}$.

Potom označíme-li $D_{a_0} = a_0^{-1/2}(F + A)^{1/2} - S^{-1}$, platí, že

$$\hat{\gamma} = \hat{U}Y + \hat{u},$$

kde

$$\hat{U} = CQ^{-1/2}D_{a_0}Q^{-1/2}X'(R + XQ^{-1/2}D_{a_0}Q^{-1/2}X')^{-1} \quad (6) \\ \hat{u} = c,$$

je MILE-odhad parametrické funkce (2) vzhledem ke ztrátové funkci (4) při apriorní informaci o parametru β (5).

Navíc platí, že

$$Z(O, \hat{U}, c) \leq Z(\beta, \hat{U}, c) \leq Z(O, \hat{U}, c) + a_0$$

pro každé $\beta \in G$.

Z vyjádření (6) je patrné, že MILE-odhad závisí na ztrátové funkci (4) prostřednictvím matice Z . Kromě toho je výpočtově značně náročné najít konstantu a_0 a matici A tak, aby vyhovovaly

předpokladům (i), (ii) a (iii). Vlastní výpočet MILE-odhadu na základě (6) nepředstavuje zvláštní těžkosti, s výhodou lze využít např. programového systému MATLAB (viz např. [13]) apod. Dále ukážeme speciální případ, kdy lze snadno najít explicitní vzorce pro výpočet MILE-odhadu.

Předpokládejme, že $C = I$, $Q = \frac{1}{\kappa}I$, $Z = I$ a $\kappa > 0$ je konstanta. Pak $F = \kappa^{-1/2}(X'R^{-1}X)^{-1}$ a snadno lze nahlédnout, že podmínka (i) je splněna pro $a_0 \leq \kappa$ a $A = O$. Je-li matice A nulová, pak podmínka (ii) je splněna triviálně a z podmínky (iii) dostaneme, že

$$a_0 = \kappa(\text{Tr}[(X'R^{-1}X)^{-1}] / (\kappa + \text{Tr}[(X'R^{-1}X)^{-1}]))^2.$$

Odtud po dosazení do (6) ihned dostaneme, že

$$\hat{U} = \frac{\kappa}{\kappa + \text{Tr}[(X'R^{-1}X)^{-1}]} (X'R^{-1}X)^{-1} X'R^{-1}. \quad (7)$$

Porovnáme-li pro $c = O$ získaný MILE-odhad parametrické funkce $\gamma = C\beta = \beta$ získaný za podmínky $\beta'\beta \leq \kappa$ pomocí vzorce (7) s nejlepším nestranným lineárním odhadem $\beta^* = (X'R^{-1}X)^{-1} X'Y$, dostaneme následující větu.

Věta 2. MILE-odhad parametru β v regresním modelu (1) se známou variační maticí R vzhledem ke ztrátové funkci

$$Z(\beta, U, u) = E_{\beta}(UY + u - \beta)'(UY + u - \beta)$$

při apriorní informaci $\beta'\beta \leq \kappa$ lze vyjádřit ve tvaru

$$\hat{\beta} = \frac{\kappa}{\kappa + \text{Tr}[(X'R^{-1}X)^{-1}]} \beta^*.$$

Jiný speciální případ, kdy matice Z určující ztrátovou funkci má hodnotu rovnou jedné, vede k minimaxovým odhadům, jež jsou popsány v [8].

3. Minimaxové odhady skalární parametrické funkce při apriorní informaci o parametrech β a R

V tomto odstavci se budeme zabývat MILE-odhadem skalární parametrické funkce $\gamma = a'\beta$, kterou dostaneme jako speciální případ parametrické funkce (2) volbou $s = 1$, $C' = a \in \mathcal{M}_{p \times 1}$, $c = 0$. MILE-odhady budeme hledat ve tvaru $\hat{\gamma} = U'Y + u$, kde $U \in \mathbb{R}^n$ a $u \in \mathbb{R}$. Dále budeme předpokládat, že jsou dány množiny $G \subseteq \mathbb{R}^p$ a $\mathcal{R} \subseteq \mathcal{M}_n^{\geq}$ a apriorní informace o parametru β tvaru $\beta \in G$ a apriorní informace o neznámé varianční matici R vektoru náhodných chyb e ve tvaru $R \in \mathcal{R}$.

Budeme uvažovat kvadratickou ztrátovou funkci

$$Z((\beta, R), U, u) = E_{(\beta, R)}(\gamma - \hat{\gamma})^2 = E_{(\beta, R)}(a'\beta - U'Y - u)^2 \quad (8)$$

a zavedeme funkci

$$J(U, u) = \sup_{\beta \in G, R \in \mathcal{R}} Z((\beta, R), U, u). \quad (9)$$

Odhad $\hat{\gamma} = \hat{U}'Y + \hat{u}$ nazveme MILE-odhadem parametrické funkce γ vzhledem ke ztrátové funkci (8) při apriorní informaci $\beta \in G$ a $R \in \mathcal{R}$, jestliže

$$J(\hat{U}, \hat{u}) = \inf_{U \in \mathbb{R}^n, u \in \mathbb{R}} J(U, u).$$

Potom $\sigma^2(a, G, \mathcal{R}) = J(\hat{U}, \hat{u})$ je minimaxová střední kvadratická chyba MILE-odhadu při apriorní informaci $\beta \in G$ a $R \in \mathcal{R}$ a $\delta = \sup_{\beta \in G, R \in \mathcal{R}} |E(\hat{\gamma} - \gamma)|$ je minimaxové vychýlení MILE odhadu $\hat{\gamma}$.

Z vlastností funkce $J(U, u)$ lze snadno odvodit, že minimaxový odhad $\hat{\gamma}$ existuje, jestliže jsou splněny následující dvě podmínky:

$$(P1) \quad \{(U, u) : J(U, u) < \infty\} \neq \emptyset$$

(symbol \emptyset značí prázdnou množinu)

$$(P2) \quad \{U : \sup_{R \in \mathcal{R}} U'RU = 0\} = \{O\}.$$

Jestliže $G \neq \mathbb{R}^p$ a platí-li podmínky (P1) a (P2), lze minimaxové vychýlení δ zapsat ve tvaru $\delta = \sup_{\beta \in G} |(a - X'\hat{U})'\beta - \hat{u}|$ a tedy MILE-odhad $\hat{\gamma}$ není nestranný.

Pro výpočet MILE-odhadu je užitečná následující věta (viz [10]).

Věta 3. Nechť G je konvexní množina, $\{U : \sup_{\beta \in G} |(a - X'U)'\beta| < \infty\} \neq \emptyset$ a platí (P2). Pak položíme-li

$$\Phi(U) = \frac{1}{4} [\sup_{\beta \in G} (a - X'U)'\beta - \inf_{\beta \in G} (a - X'U)'\beta]^2 + \sup_{R \in \mathcal{R}} U'RU,$$

dostaneme MILE-odhad parametrické funkce γ ve tvaru $\hat{\gamma} = \hat{U}'Y + \hat{u}$, přičemž

$$\Phi(\hat{U}) = \inf_{U \in \mathbb{R}^n} \Phi(U)$$

a

$$\hat{u} = \frac{1}{2} [\sup_{\beta \in G} (a - X'\hat{U})'\beta + \inf_{\beta \in G} (a - X'\hat{U})'\beta].$$

Zřejmým důsledkem věty 3 je, že za předpokladu symetrie množiny G vzhledem k prvku $\beta \in G$ lze funkci Φ přepsat do tvaru

$$\Phi(U) = \sup_{\beta \in G - \beta} [(a - X'U)'\beta]^2 + \sup_{R \in \mathcal{R}} U'RU$$

a $\hat{u} = \beta'(a - X'\hat{U})$. Pro minimaxovou střední kvadratickou chybu pak dostaneme $\sigma^2(a, G, \mathcal{R}) = \Phi(\hat{U})$.

V případě, že $G = \mathbb{R}^p$, lze z věty 3 odvodit následující dvě tvrzení (viz [10]).

(T1) MILE-odhad $\hat{\gamma}$ existuje, jestliže hodnota matice $(X':a) \in \mathcal{M}_{n \times (p+1)}$ je p .

(T2) Existuje-li MILE-odhad $\hat{\gamma} = \hat{U}'Y + \hat{u}$ a položíme-li $\mathcal{U} = \{U : a = X'U\}$ a $\Phi_1 = \sup_{R \in \mathcal{R}} U'RU$

pak platí $\hat{u} = 0$ a $\Phi_1(\hat{U}) = \inf_{U \in \mathcal{U}} \Phi_1(U)$.

Uvedených výsledků lze využít ke stanovení explicitních výrazů pro MILE-odhady pro některé speciální případy množin G a \mathcal{R} .

4. Speciální případy volby G a \mathcal{R}

V tomto odstavci budeme předpokládat, že množina G je zvolena z následujících tří možností

- (I) $G = \{\beta : \beta'Q_1\beta \leq 1\}$, $Q_1 \in \mathcal{M}_p^>$
- (II) $G = \{\beta : \beta'Q_1\beta \leq 1, L\beta = b\}$, $Q_1 \in \mathcal{M}_p^>$, $L \in \mathcal{M}_{s \times p}$, hodnota L je s a $b \in \mathbb{R}^s$
- (III) $G = \{\beta = (\beta^1, \dots, \beta^p) : |\beta_i| \leq c_i, i = 1, \dots, p\}$, $c = (c_1, \dots, c_p)' \in \mathbb{R}^p$.

Dále budeme uvažovat následující tři možné volby množiny \mathcal{R} :

- (A) $\mathcal{R} = \{R \in \mathcal{M}_n^> : \text{Tr}(Q_2 R) \leq 1\}$, $Q_2 \in \mathcal{M}_n^>$
- (B) $\mathcal{R} = \{R \in \mathcal{M}_n^> : \text{Tr}(R - R_0)^2 \leq q_0\}$, $R_0 \in \mathcal{M}_n^>$, $q_0 > 0$ je konstanta
- (C) $\mathcal{R} = \{R = (r_{ij} \in \mathcal{M}_n^> : |r_{ij} - r_{ij}^{(0)}| \leq b_{ij}, i, j = 1, \dots, n\}$,
 $R_0 = (r_{ij}^{(0)}) \in \mathcal{M}_n^>$, $B = (b_{ij}) \in \mathcal{M}_n^>$.

Potom (viz [10]) lze odvodit explicitní výrazy pro MILE-odhad parametrické funkce γ pro některé kombinace volby množin G a \mathcal{R} . Výsledky jsou uvedeny v následujícím přehledu.

I A: Nechť G je typu I a \mathcal{R} je typu A. Pak

$$\begin{aligned}\hat{U} &= Q_2 X(Q_1 + X'Q_2X)^{-1}a, \quad \hat{u} = 0 \\ \sigma^2(a, G, \mathcal{R}) &= a'(Q_1 + X'Q_2X)^{-1}a \\ \delta^2 &= a'(Q_1 + X'Q_2X)^{-1}Q_1(Q_1 + X'Q_2X)^{-1}a.\end{aligned}$$

I B: Nechť G je typu I a \mathcal{R} je typu B. Pak dostaneme stejné výsledky jako v I A, když Q_2 nahradíme maticí $(R_0 + q_0I)^{-1}$.

II A: Nechť G je typu II a \mathcal{R} je typu A. Pak

$$\begin{aligned}\hat{U} &= Q_2 X(I + \tilde{Q}_1 X'Q_2X)^{-1}\tilde{Q}_1 a \\ \hat{u} &= (a - X'\hat{U})Q_1^{-1}L(LQ_1^{-1}L')^{-1}b \\ \sigma^2(a, G, \mathcal{R}) &= ((I + \tilde{Q}_1 X'Q_2X)^{-1}\tilde{Q}_1 a)'a\end{aligned}$$

přičemž

$$\tilde{Q}_1 = [(1 - b'(LQ_1^{-1}L')^{-1}b)Q_1^{-1}(I - L'(LQ_1^{-1}L')^{-1}LQ_1^{-1})].$$

II B: Nechť G je typu II a \mathcal{R} je typu B. Pak dostaneme stejné výsledky jako ve II A, když za Q_2 dosadíme maticí $(R_0 + q_0I)^{-1}$.

III C: Nechť G je typu III a \mathcal{R} typu C. Potom lze nalézt množiny G_1 a G_2 typu I tak, že $G_1 \subset G \subset G_2$ a množiny \mathcal{R}_1 a \mathcal{R}_2 typu A tak, že $\mathcal{R}_1 \subset \mathcal{R} \subset \mathcal{R}_2$ a aproximaci MILE-odhadu $\hat{\gamma}$ včetně aproximací střední kvadratické chyby $\sigma^2(a, G, \mathcal{R})$ lze najít užitím výsledků I A pro dvojice G_1, \mathcal{R}_1 a G_2, \mathcal{R}_2 . Podrobněji je nastíněná aproximace popsána v [10].

Vhodnou volbou Q_1 a Q_2 v I A a v I B lze získat analogické výsledky jako ve větě 2.

5. Závěr

Použití minimaxových odhadů naráží v mnohých situacích na obtíže spojené zejména s jejich numerickým výpočtem. Vyjádření navrhovaná ve čtvrtém odstavci tuto obtíž pro vybrané specifické apriorní informace v parametrech obcházejí, na základě uvedených výsledků není obtížné pomocí soudobých softwarových systémů (např. MATLAB apod.) získat numerické představy o chování MILE-odhadů. Podrobnější vyšetření jejich robustních vlastností vyžaduje jejich další statistické posouzení. Rovněž se ukazuje, že zobecnění uvedeného postupu na vektorové parametrické funkce tvaru (2) nepředstavuje zvláštní těžkosti. Zajímavá se rovněž jeví srovnání minimaxového přístupu s metodami "ridge regrese".

Zobecnění celé problematiky na nelineární regresní funkce by mohlo být efektivním využitím apriorní informace o neznámých parametrech a mohlo by vést k nalezení vhodných počátečních odhadů pro jiné iterační postupy se známými statistickými vlastnostmi.

Další zajímavé použití minimaxového přístupu v oblasti neparametrické regrese ukazuje práce [5]. Z tohoto hlediska se další použití minimaxového přístupu jeví jako nadějně.

LITERATURA:

- [1] Bublik, B. N., Danilov V. J., Nakonečnyj A. G. (1983):, *Problem of control in linear systems*, UMKVO Kiev (in Russian).
- [2] Bunke, O. (1975), *Last Squares Estimator as Robust Minimax Estimators*, Math. Operationsforsch. Statist., 6 p. 687-688.
- [3] Bunke, H., Bunke O. (1986), *Statistical Inference in Linear Models*, Wiley Series in Probability and Mathematical Statistics, Volume 1, New York.
- [4] Bunke, H., Bunke O. (1989), *Nonlinear Regression, Functional Relations and Robust Methods*, Wiley Series in Probability and Mathematical Statistics Volume 2, New York.
- [5] Heckman, N. E., Woodroove, M. (1991), *Minimax Bayes Estimation in Nonparametric Regression*, The Annals of Statistics. Vol. 19, No 4.p. 2003-2014.
- [6] Huber, P. J. (1969), *Théorie de l'inférence statistique robuste*, Presses de l' Université, Montréal.
- [7] Jurečková, J. (1989), *Robust statistical inference in linear models in [3]*, p. 135-208.

- [8] Kuks, J. (1972), *Minimálnaja ocenka koeficientov regresii*, Izd. Akad. Nauk. Estonskoj SSSR, 21 p. 73-78 (rusky).
- [9] Läuter, H. (1975), *A Minimax Linear Estimator for Linear Parameters Under Restrictions in Form of Inequalities*, *Math. Operationsforsch. Statist.*, 6, p. 689-695.
- [10] Michálek, J., Nakonečnyj, A. G. (1992), *Minimax Estimators for Parameters of Linear Regression Models*, (v tisku).
- [11] Nakonečnyj, A. G. (1985), *Minimálne ocenivanijsje funkcionalov ot rešenij variacionnych uravnenij v Giltelovych prostranstvach*, Kiev KGU (rusky).
- [12] Novovičová, J. (1980), *M-odhady a SA-odhady v lineárním regresním modelu*, *Sborník ROBUST 80* s. 119-130.
- [13] *The Student Edition of MATLAB for MS-DOS Personal Computers*, Prentice Hall, Englewood Cliffs.