APLIKÁCIA KVADRATICKÝCH ODHADOV V METROLÓGII

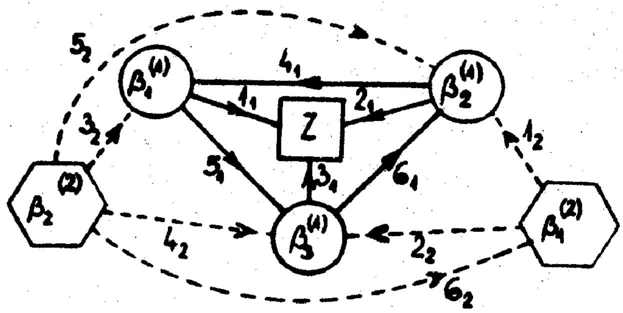
LUBOMÍR KUBÁČEK
Matematický ústav SAV
814 73 Bratislava, Obrancov mieru 49

1. Úvod

Vznik špeciálnej štruktúry (dvojetapovej a zmiešanej) regresného modelu ukazuje nasledovný problém.

Nech Z označuje hodnotu etalónu nejakej fyzikálnej veličiny, ktorá je prijatá konvenciou, t.zn. je bezchybná. S týmto etalónom Z porovnáme etalóny $\beta_1^{(1)}, \beta_2^{(1)}, \beta_3^{(1)}$ (prvá etapa) v súlade s grafom na obr.1.

Obr.1



Orientovaná hrana grafu je priradená náhodnej veličine, realizáciou ktorej vznikne údaj "zmeraný rozdiel medzi hodnotou etalónu na konci hrany a hodnotou etalónu na začiatku hrany". Ak pre stredné hodnoty E(Y) týchto náhodných veličín $Y_i^{(1)}$, i=1,..., 6, platí (t.zn., že v meraní nepôsobili systematické vplyvy):

$$E\begin{pmatrix} Y_1^{(1)} - Z \\ Y_2^{(1)} - Z \\ Y_3^{(1)} - Z \\ Y_4^{(1)} \\ Y_5^{(1)} \\ Y_6^{(1)} \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \\ 1 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} \beta_1^{(1)} \\ \beta_2^{(1)} \\ \beta_3^{(1)} \end{pmatrix}$$

(posledný vzťah zapíšeme v ďalšom $E(Y^{(1)}) = X^{(1)}\beta^{(1)}$) a pre kovariančnú maticu $Var(Y^{(1)})$ náhodného vektora $Y^{(1)}$ platí $Var(Y^{(1)}) = \sigma_1^2 I$ (I označuje identickú maticu), potom najlepší lineárny nevychýlený odhad parametra $\beta^{(1)}$ založený na vektore $Y^{(1)}$ je $\hat{\beta}^{(1)}(Y^{(1)}) = (X^{(1)\prime}X^{(1)})^{-1}X^{(1)\prime}Y^{(1)}$. Nech ďalej etalóny druhej etapy $\beta_1^{(2)}, \beta_2^{(2)}$ boli porovnané s etalónmi prvej etapy tak, že sa opäť zmerali rozdiely hodnôt etalónov podľa grafu na obr.1. Nech (podobne ako v prvej etape) platí

$$E\begin{pmatrix} Y_1^{(2)} \\ Y_2^{(2)} \\ Y_3^{(2)} \\ Y_4^{(2)} \\ Y_5^{(2)} \\ Y_6^{(2)} \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \beta_1^{(1)} \\ \beta_2^{(1)} \\ \beta_3^{(1)} \end{pmatrix} + \begin{pmatrix} -1 & 0 \\ -1 & 0 \\ 0 & -1 \\ 0 & -1 \\ 0 & -1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} \beta_1^{(2)} \\ \beta_2^{(2)} \end{pmatrix}$$

(posledný vzťah zapíšeme v ďalšom $E(Y^{(2)}) = C_{21}\beta^{(1)} + X^{(2)}\beta^{(2)}$) a $Var(Y^{(2)}) = \sigma_2^2 I$, pričom vektory $Y^{(1)}$ a $Y^{(2)}$ sú stochasticky nezávislé.

Podmienkou teraz je, že vektor $Y^{(2)}$ druhej etapy môžeme použiť pre odhad parametrov druhej etapy, ale hodnoty odhadu $\hat{\beta}^{(1)}(Y^{(1)})$ nesmieme meniť. Môžeme teda hľadať odhad $\hat{\beta}^{(2)}(Y^{(1)}, Y^{(2)})$ založený na vektoroch $Y^{(1)}$, $Y^{(2)}$, alebo odhad $\tilde{\beta}^{(2)}(Y^{(2)}, \hat{\beta}^{(1)})$

založený na vektoroch $Y^{(2)}$, $\hat{\beta}^{(1)}(Y^{(1)})$. V druhom prípade na pr. dostávame:

$$E(Y^{(2)} - C_{21}\hat{\beta}^{(1)}(Y^{(1)})) = X^{(2)}\beta^{(2)}$$
$$Var(Y^{(2)} - C_{21}\hat{\beta}^{(1)}) = \sigma_1^2 C_{21}(X^{(1)\prime}X^{(1)})^{-1}C_{21}^{\prime} + \sigma_2^2 I .$$

(Tu je potrebné si všimnúť, že kovariančná matica má štruktúru $\sigma_1^2 H_1 + \sigma_2^2 H_2$, kde H_1 a H_2 sú známe matice, avšak σ_1^2, σ_2^2 známe byť nemusia.) Obidve etapy (súčasné) sú charakterizované modelom

$$(+)\quad E\begin{pmatrix} Y^{(1)} \\ Y^{(2)} \end{pmatrix} = \begin{pmatrix} X^{(1)} & 0 \\ C_{21} & X^{(2)} \end{pmatrix}\begin{pmatrix} \beta^{(1)} \\ \beta^{(2)} \end{pmatrix} ; \quad Var\begin{pmatrix} Y^{(1)} \\ Y^{(2)} \end{pmatrix} = \sigma_1^2\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} + \sigma_2^2\begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix}$$

(opäť kovariančná matica má už spomínanú štruktúru), pričom $\mathcal{M}(C_{21}^{\prime}) \subset \mathcal{M}(X_1^{\prime})$ ($\mathcal{M}(.)$ označuje stĺpcový priestor matice v zátvorkách).

V prípade, že hodnoty σ_1^2, σ_2^2, σ_1^2/σ_2^2 sú a priori neznáme, vzniká problém ako odhadovať parametre druhej etapy $\beta^{(2)}$, prípadne ako odhadovať hodnoty σ_1^2, σ_2^2.

Tu je potrebné poznamenať, že štruktúra (+) regresného modelu vzniká na pr. v geodézii, geofyzike a iných odboroch, takže užitočnosť štúdia štruktúry(+) nie je obmedzená na jedinú aplikačnú oblasť.

V ďalšom je uvedený prehľad základných poznatkov o tejto problematike a niektoré novšie výsledky.

2. Definície a formulácia problému

Definícia 2.1. Regresný model $(Y, X\beta, \Sigma)$ nazveme p-etapovým ak pre maticu plánu X a pre kovariančnú maticu Σ platí:

$$X = \begin{pmatrix} X_1 & 0 & \dots & \dots & \dots & 0 \\ C_{2,1} & X_2 & \dots & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ C_{p,1} & C_{p,2} & \dots & C_{p,p-1} & & X_p \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \Sigma_{1,1} & 0 & & 0 \\ 0 & \Sigma_{2,2} & & 0 \\ \dots & \dots & \dots & \\ 0 & 0 & & \Sigma_{p,p} \end{pmatrix},$$

$\mathcal{M}(C_{i,j}^{\prime}) \subset \mathcal{M}(X_j^{\prime})$, $i=2,\dots,p$, $j=1,\dots,p$, $i \geq j$.

Definícia 2.2. Regresný model nazveme zmiešaným ak pre jeho kovariančnú maticu Σ platí $\Sigma = \sum_{i=1}^{p} \vartheta_i V_i$, $p \geq 2$, pričom V_i sú známe symetrické matice a ϑ_i sú neznáme parametre (variančné komponenty).

Definícia 2.3. Regresný model $(\underline{Y}, \underline{X}\beta, \underline{\Sigma})$ nazveme replikovaným, ak $\underline{X} = \mathbf{1} \otimes X$, $\mathbf{1} = (1,\dots,1)^{\prime}$, $\underline{\Sigma} = I \otimes \Sigma$.

Poznámka 2.1. Z uvedených definícií je zrejmé, akú štruktúru bude mať na pr. replikovaný zmiešaný model, p-etapový zmiešaný replikovaný model a pod.

V ďalšom si všimnime postupy pre odhad parametrov β a ϑ v nereplikovanom modeli a postupy pre odhad parametrov ϑ v replikovanom zmiešanom modeli.

Problémy určenia týchto odhadov v nereplikovaných modeloch sa doteraz študovali najmä v troch základných polohách a to: a) odhad β [1], [4], [12], b) odhad ϑ [5], [6], [11], [13], [15], [16] a c) simultánny odhad β a ϑ [1], [9], [16]; v replikovaných modeloch sa zatiaľ pozornosť sústreďuje na odhad ϑ [2], [5], [6], [15], výnimočne na β [12], [4].

Poznámka 2.2. V [15] je uvedený značný počet citácií prác, ktoré sa zaoberajú problematikou odhadu ϑ.

3. Odhad parametra β

Teoréma 3.1. V 2-etapovom regresnom modeli so známou kovariančnou maticou a s odhadnuteľnými parametrami prvej aj druhej etapy platí:

$$\hat{\beta}^{(2)}(Y^{(1)}, Y^{(2)}) = \hat{\beta}^{(2)}(Y^{(2)}, \hat{\beta}^{(1)}(Y^{(1)})),$$

kde $\hat{\beta}^{(2)}(Y^{(1)}, Y^{(2)})$ je najlepší lineárny odhad založený na vektoroch $Y^{(1)}$ a $Y^{(2)}$ a $\hat{\beta}^{(2)}(Y^{(2)}, \hat{\beta}^{(1)}(Y^{(1)}))$ je najlepší lineárny odhad založený na vektoroch $Y^{(2)}$ a

$\hat{\beta}^{(1)}(Y^{(1)})$ (najlepší lineárny odhad vektora $\beta^{(1)}$ založený na vektore $Y^{(1)}$).

Dôkaz. Pozri teorému 3.1 v [10].

Teoréma 3.2. V p-etapovom ($p > 2$) regresnom modeli so známou kovariančnou maticou a s odhadnuteľnými parametrami β vo všetkých etapách platí:

$$\hat{\beta}^{(i)}(Y^{(1)},\dots,Y^{(i)}) = \hat{\beta}^{(i)}\Big\{Y^{(i)}, \hat{\beta}^{(1)}(Y^{(1)}),\dots, \hat{\beta}^{(i-1)}(Y^{(i-1)}, \hat{\beta}^{(i-2)}\big[Y^{(i-2)}, \hat{\beta}^{(i-3)}(Y^{(i-3)}, \hat{\beta}^{(i-4)}(\dots)),\dots, \hat{\beta}^{(1)}(Y^{(1)})\big]\Big\},$$

$i=1,\dots,p$, práve vtedy, ak $C_{i,j} = 0 \Leftrightarrow i-j \geq 2$, $i=1,\dots,p$, $j=1,\dots,i-1$.

Dôkaz. Pozri teorému 3.2 v [10].

Teoréma 3.3. Nech v replikovanom modeli $(Y = (Y_1', Y_2', \dots, Y_{f+1}')', (1 \otimes X), I \otimes \Sigma)$ je náhodný vektor Y normálne rozdelený. Pri označení

$$S = (1/f)\sum_{i=1}^{f+1}(Y_i - \bar{Y})(Y_i - \bar{Y})', \quad \bar{Y} = [1/(f+1)]\sum_{i=1}^{f+1} Y_i$$

a za predpokladu $f \geq R(\Sigma)$ (hodnosť), platí:

(1) $$\{X[(X')^-_{m(S)}]'\bar{Y}\}^{(p)} \sim N(X\beta, 1 + \tfrac{1}{f}T_2'(Z'SZ)^-T_2\Lambda_{11.2}),$$

kde

$$\mathcal{M}(Z) = Ker(X'), \quad T_2 = Z'\bar{Y},$$

$$\Lambda = \begin{pmatrix} XX^-\Sigma(X^-)'X'/(f+1), & XX^-\Sigma Z/(f+1) \\ Z'\Sigma(X^-)'X'/(f+1), & Z'\Sigma Z/(f+1) \end{pmatrix} = \begin{pmatrix} \Lambda_{11}, & \Lambda_{12} \\ \Lambda_{21}, & \Lambda_{22} \end{pmatrix},$$

$\Lambda_{11.2} = \Lambda_{11} - \Lambda_{12}\Lambda_{22}^-\Lambda_{21}$, $(X')^-_{m(S)}$ je minimum S-seminorm g-inverzia matice X' a (p) označuje $(T_2, \hat{\Lambda}_{22} = Z'SZ/(f+1))$-podmienenosť;

(2) ak označíme $X[(X')^-_{m(S)}]'\bar{Y} = \widehat{X\beta}$, potom $f(\bar{Y} - \widehat{X\beta})'S^-(\bar{Y} - \widehat{X\beta})$ je náhodná premenná s rozdelením pravdepodobnosti $[f(C - \nu_1)/(\nu_1 + f - C + 1)]$. $\cdot F_{f(C-\nu_1),(\nu_1+f-C+1)}$, kde $C = R(\Sigma)$, $\nu_1 = R\langle[\Sigma/(f+1)]\{[\Sigma/(f+1)] + XX'\}^-X\rangle$;

(3) náhodná premenná $\langle(\widehat{X\beta} - X\beta)'\{X[(X')^-_{m(S)}]'S\}^-(\widehat{X\beta} - X\beta)/(1 + \tfrac{1}{f}\hat{v}'S^-\hat{v})\rangle$. $\cdot[f - (C - \nu_1)]/f$ má rozdelenie pravdepodobnosti

$$[f - (C - \nu_1)]\frac{\nu_1}{f - C + 1}F_{\nu_1, f-C+1},$$

pričom $\hat{v} = \bar{Y} - \widehat{X\beta}$;

(4) $$Var(\widehat{X\beta}) = X[(X')^-_{m(S)}]'\Sigma\,\frac{1}{f+1}\,\frac{f-1}{\nu_1 + f - C - 1}.$$

Dôkaz. Pozri lemu 2.1, lemu 2.2, teorému 2.1, teorému 2.2, lemu 2.3 a teorému 2.3 v [4].

Poznámka 3.1. Predchádzajúcu teorému možno využiť pre všetky v praxi sa vyskytujúce regulárne verzie regresného modelu ([3] str.127-138). Na pr. v modeli (5.model v [3] str.126) $E(\xi) = A_{n,k}\Theta$, $\Sigma_\xi = \sigma^2 H_{n,n}/(f+1)$, $b + B_{q,n}\Theta = 0$, kde $R(A) = k \leq n$, $R(H) = n$, $R(B) = q \leq k$, platí

$$\hat{\beta} = \{(A'S^{-1}A)^{-1} - (A'S^{-1}A)^{-1}B'[B(A'S^{-1}A)^{-1}B']^{-1}B(A'S^{-1}A)^{-1}\}A'S^{-1}\xi - (A'S^{-1}A)^{-1}B[B(A'S^{-1}A)^{-1}B']^{-1}b$$

$$Var(\hat{\beta}) = \{(A'\Sigma^{-1}A)^{-1} - (A'\Sigma^{-1}A)^{-1}B'[B(A'\Sigma^{-1}A)^{-1}B']^{-1}B(A'\Sigma^{-1}A)^{-1}\}.$$
$$\cdot\frac{1}{f+1}\,\frac{f-1}{f-(n-k+q)-1}.$$

Poznámka 3.2. Ak v zmiešanom regresnom modeli $(Y, X\beta, \Sigma = \sum_{i=1}^{p}\vartheta_i V_i)$ parametre $\vartheta_1,\dots,\vartheta_p$ ($p \geq 2$) nepoznáme, môžeme na základe vektora Y vo všeobecnosti určiť len ϑ_0-lokálne najlepší odhad vektora $X\beta$ vzťahom

$$\widehat{X\beta} = X\left[(X')^{-}_{m(\sum_{i=1}^{p}\vartheta_{i,0}V_i)}\right]' Y .$$

Pre niektoré funkcionály $f(.): \mathcal{R}^k \to \mathcal{R}^1$, $f(\beta) = f'\beta$, však môže existovať $\underline{\Sigma}$-rovnomerne najlepší odhad (t.zn. rovnomerne najlepší vzhľadom k triede

$$\underline{\Sigma} = \left\{\Sigma:\ \Sigma = \sum_{i=1}^{p}\vartheta_i V_i,\ \vartheta \in \underline{\vartheta}\right\}).$$

<u>Teoréma 3.4.</u> Nech $E(Y) = X\beta$, $Var(Y) = \Sigma \in \underline{\Sigma} = \left\{\Sigma: \Sigma = \sum_{i=1}^{p}\vartheta_i V_i,\ \vartheta \in \underline{\vartheta} \subset \mathcal{R}^p\right\}$, $\beta \in \mathcal{R}^k$. Nech ďalej Σ_0 je matica s vlastnosťou $\forall\{\Sigma \in \underline{\Sigma}\}\ \mathcal{M}(\Sigma) \subset \mathcal{M}(\Sigma_0)\ \&\ \forall\{\Sigma_1 : \Sigma_1$ je pozitívne semidefinitná matica s vlastnosťou $\mathcal{M}(\Sigma) \subset \mathcal{M}(\Sigma_1)$, $\Sigma \in \underline{\Sigma}\}\ \mathcal{M}(\Sigma_0) \subset \mathcal{M}(\Sigma_1)$.

Potom platí: Ak $a \in \mathcal{M}(W)$, kde $W = \Sigma_0 + XX'$, tak $a'Y$ je $\underline{\Sigma}$-rovnomerne najlepší lineárny odhad funkcie $f(\beta) = a'X\beta$, $\beta \in \mathcal{R}^k$, práve vtedy ak

$$a \in \mathcal{M}\left\{W^+X\ Ker\left[X\ W^-\ \sum_{i=1}^{p} V_i M V_i W^- X\right]\right\},$$

kde $M = I - X(X'X)^- X'$. (Tu W^+ je Mooreova-Penroseova inverzia matice W a $(X'X)^-$ je g-inverzia matice $X'X$.)

Dôkaz. Pozri [17].

<u>Poznámka 3.3.</u> Pri odhadovaní funkcie $f(\beta) = f'\beta$, $\beta \in \mathcal{R}^k$, kde $f \in \mathcal{M}(X')$ sa stačí obmedziť na odhady $a'Y$, kde $a \in \mathcal{M}(W)$.

<u>Teoréma 3.5.</u> Nech $E(Y) = X\beta$, $\beta \in \mathcal{R}^k$, $Var(Y) = \Sigma \in \underline{\Sigma} = \left\{\sum_{i=1}^{p}\vartheta_i V_i,\ \vartheta \in \underline{\vartheta}\right\}$. Pre funkciu $f(\beta) = c'\beta$, $\beta \in \mathcal{R}^k$, existuje $\underline{\Sigma}$-rovnomerne najlepší lineárny odhad práve vtedy, ak

$$c \in \mathcal{M}\left\{X'W^-X\ Ker\left[X'W^-\ \sum_{i=1}^{p} V_i M V_i W^- X\right]\right\}$$

(symbol W je definovaný v teoréme 3.4).

Dôkaz. Pozri [17].

4. Odhad parametra ϑ

<u>Teoréma 4.1.</u> V zmiešanom regresnom modeli $E(Y) = X\beta$, $\beta \in \mathcal{R}^k$, $Var(Y) = \Sigma \in \underline{\Sigma} = \left\{\sum_{i=1}^{p}\vartheta_i V_i,\ \vartheta \in \underline{\vartheta}\right\}$ je funkcia $g(\vartheta) = g'\vartheta$, $\vartheta \in \underline{\vartheta}$, nevychýlene odhadnuteľná práve vtedy, ak $g \in \mathcal{M}(K_0)$, kde $\{K_0\}_{i,j} = Tr(V_iV_j - PV_iPV_j)$, $i,j = 1,\dots,p$, $P = X(X'X)^-X'$.

Dôkaz. Pozri v [15].

<u>Definícia 4.1.</u> V zmiešanom regresnom modeli odhad $Y'AY$ nazveme invariantným, ak $\forall\{\delta \in \mathcal{R}^k\}\ (Y - X\delta)'A(Y - X\delta) = Y'AY$.

<u>Lema 4.1.</u> Nech Y je náhodný vektor s danou distribučnou funkciou, $E(Y) \in \mathcal{M}(X)$, kde X je daná matica a $Var(Y) = \Sigma$. Nech ďalej

$$\mathcal{T}_1 = \{Y'AY : A = A',\ X'AX = 0,\ X'A\Sigma = 0\}$$
$$\mathcal{T}_2 = \{Y'AY : A = A',\ X'A = 0\}.$$

Potom pre každú náhodnú premennú $Y'A_1Y \in \mathcal{T}_1$ existuje náhodná premenná $Y'A_2Y \in \mathcal{T}_2$, pre ktorú platí $P\{Y'A_1Y = Y'A_2Y\} = 1$.

Dôkaz. Zrejme $\mathcal{T}_2 \subset \mathcal{T}_1$. Nech $Y'A_1Y \in \mathcal{T}_1$. Potom $Y'MA_1MY \in \mathcal{T}_2$ $(M = I - X(X'X)^-X')$. Označme $\varepsilon = Y - E(Y)$. Pre ε platí: $P\{\varepsilon \in \mathcal{M}(\Sigma)\} = 1$. Pre náhodnú premennú $Y'A_1Y$ teda platí

$$Y'A_1Y = \varepsilon'A_1\varepsilon \quad \text{skoro iste (s.i.)},$$

pretože

$$\varepsilon'A_1E(Y) = 0 \quad \text{s.i.}$$
$$E(Y')A_1\varepsilon = 0 \quad \text{s.i.}$$
$$E(Y')AE(Y) = 0 \ \text{s.i.}.$$

Pre náhodnú premennú $Y'MA_1MY$ platí

$$Y'MA_1MY = \varepsilon' MA_1 M\varepsilon = \varepsilon'(I-P)A_1(I-P)\varepsilon = \varepsilon' A_1 \varepsilon \quad \text{s.i.},$$

pretože $\varepsilon' PA_1\varepsilon = 0$ s.i., $\varepsilon' A_1 P\varepsilon = 0$ s.i., $\varepsilon' PA_1 P\varepsilon = 0$ s.i.. (Tu $P = X(X'X)^- X'$.)

Teda

$$Y'A_1Y \overset{\text{s.i.}}{=} Y'MA_1MY \in \mathcal{T}_2 .$$

Pomocou lemy 4.1 možno dokázať nasledujúcu teorému.

Teoréma 4.2. V zmiešanom regresnom modeli je funkcia $g(\vartheta) = g'\vartheta$, $\vartheta \in \underline{\vartheta}$, nevychýlene a invariantne odhadnuteľná práve vtedy, ak $g \in \mathcal{M}(K_0^{(I)})$, kde $\{K_0^{(I)}\}_{i,j} = \operatorname{Tr}(MV_iMV_j)$, $i,j=1,\dots,p$.

Teoréma 4.3. V zmiešanom regresnom modeli pre ϑ_0-lokálne najlepší nevychýlený a invariantný odhad funkcie $g(\vartheta) = g'\vartheta$, $\vartheta \in \underline{\vartheta}$, kde $g \in \mathcal{M}(K_0^{(I)})$ a $\Sigma_0 = \sum_{i=1}^{p} \vartheta_{i,0} V_i$ je regulárna, v prípade normality vektora Y, je

$$\widehat{g'\vartheta} = Y' \sum_{i=1}^{p} \lambda_i \left[\Sigma_0^{-1} - \Sigma_0^{-1}X(X'\Sigma_0^{-1}X)^- X'\Sigma_0^{-1}\right] V_i \left[\Sigma_0^{-1} - \Sigma_0^{-1}X(X'\Sigma_0^{-1}X)^- X'\Sigma_0^{-1}\right] Y .$$

Koeficienty $\lambda_1,\dots,\lambda_p$ sú riešením systému

$$S_{(M\Sigma_0M)^+}\lambda = g ,$$

kde

$$\left\{S_{(M\Sigma_0M)^+}\right\}_{i,j} = \operatorname{Tr}\left[(M\Sigma_0M)^+ V_i (M\Sigma_0M)^+ V_j\right], \quad i,j=1,\dots,p ,$$

a

$$(M\Sigma_0M)^+ = \Sigma_0^{-1} - \Sigma_0^{-1}X(X'\Sigma_0^{-1}X)^- X'\Sigma_0^{-1} .$$

Dôkaz. Pozri [1].

Teoréma 4.4. V zmiešanom regresnom modeli (β_0, ϑ_0)-lokálne najlepší nevychýlený odhad funkcie $g(\vartheta) = g'\vartheta$, $\vartheta \in \underline{\vartheta}$, kde $g \in \mathcal{M}(K_0)$ a $\Sigma_0 = \sum_{i=1}^{p} \vartheta_{i,0} V_i$ je regulárna, je

$$\widehat{g'\vartheta} = (Y - X\beta_0)' \sum_{i=1}^{p} \lambda_i A_i (Y - X\beta_0) ,$$

$A_i = \Sigma_0^{-1} V_i \Sigma_0^{-1} - \Sigma_0^{-1}X(X'\Sigma_0^{-1}X)^- X'\Sigma_0^{-1} V_i \Sigma_0^{-1} X(X'\Sigma_0^{-1}X)^- X'\Sigma_0^{-1}$, $i=1,\dots,p$ a koeficienty $\lambda_1,\dots,\lambda_p$ sú riešením sústavy

$$\sum_{i=1}^{p} \lambda_i \operatorname{Tr}(A_i V_j) = g_j, \quad j=1,\dots,p.$$

Poznámka 4.1. Pretože $\mathcal{M}(K_0^{(I)}) \subset \mathcal{M}(K_0)$ môže byť nevychýlený odhad funkcie $g(.)$ lepší než invariantný a nevychýlený odhad. Ďalej môže nastať prípad, že nevychýlený odhad existuje, hoci nevychýlený a invariantný odhad neexistuje. Je preto potrebné uvážiť voľbu odhadu v konkrétnej situácii.

Podmienky pre existenciu (β, ϑ)-rovnomerne najlepších nevychýlených odhadov sú pomerne zložité. Ukazuje to nasledujúca teoréma.

Teoréma 4.5. V zmiešanom regresnom modeli $(Y, X\beta, \Sigma = \sum_{i=1}^{p} \vartheta_i V_i)$ označme $\mathcal{V} = \operatorname{span}\{V_1,\dots,V_p\}$; ďalej pre ľubovoľné $V_0 \in \mathcal{V}$, kde V_o je pozitívne definitná definujme:

$$\mathcal{A} = \left\{d \in \mathcal{M}(X) : VV_0^{-1}d \in \mathcal{M}(X),\ V \in \mathcal{V}\right\}$$

a lineárny podpriestor $\mathcal{L}$ priestoru symetrických matíc $\mathcal{S}_n = \{A : A = A',\ A \text{ je typu } n \times n\}$

$$\mathcal{L} = \left\{G : \forall \{V \in \mathcal{V}\} \exists \{G_V \in \mathcal{V}\}\ MVM(MV_0M)^+ G (MV_0M)^+ MVM = MG_VM \ \&\ MG_V V^{-1} X = 0\right\}.$$

Potom $a'Y + Y'AY$, kde $a = V_0^{-1}d$, $d \in \mathcal{A}$ a $A = (MV_0M)^+ G (MV_0M)^+$, $G \in \mathcal{L}$, je (β, ϑ)-rovnomerne najlepší nevychýlený odhad svojej strednej hodnoty

$$E(a'Y + Y'AY) = a'X\beta + \sum_{i=1}^{p} \vartheta_i \operatorname{Tr}(AV_i) .$$

Dôkaz. Pozri [1] teoréma 2.5.

Teoréma 4.6. Nech v replikovanom zmiešanom regresnom modeli $Y = (Y_1',\dots,Y_{k+1}')' \sim N_{(k+1)m}\left[(1 \otimes X)\beta,\ I \otimes \Sigma\right]$ platí: $\Sigma = \sum_{i=1}^{m} \nu_i V_i$, $(\nu_1,\dots,\nu_m)' = \nu \in \nu_* \subset \mathcal{R}^m$, pričom existuje uzavretá guľa vnorená do ν_*, $\beta \in \mathcal{R}^s$ a matica $V(\vartheta) = \sum_{i=1}^{m} \vartheta_i V_i$ je pre uvažovaný bod $\vartheta \in \nu_*$ regulárna. Označme

$$S = (1/k)\sum_{j=1}^{k+1}(Y_j - \bar{Y})(Y_j - \bar{Y})', \ \bar{Y} = [1/(k+1)]\sum_{j=1}^{k+1} Y_j, \ \{M(\vartheta)\}_{i,j} = Tr(V_i[V(\vartheta)]^{-1}V_j[V(\vartheta)]^{-1}),$$

$i,j=1,\ldots,m.$

Potom platí

(1) Pre funkciu $g(.) : \nu_* \to \mathcal{R}^1$, kde $g(\nu) = \lambda'\nu$, $\nu \in \nu_*$, existuje v triede $\mathcal{A} = \{Tr(AS): A \in \mathcal{S}_n\}$ nevychýlený odhad práve vtedy, ak $\lambda \in \mathcal{M}(M(\vartheta)) = \mathcal{M}(K)$, kde $\{K\}_{i,j} = Tr(V_iV_j)$, $i,j=1,\ldots,m.$

(2) Ak v triede $\mathcal{A}$ existuje nevychýlený odhad pre funkciu $g(.)$, potom ϑ-lokálne najlepší odhad v tejto triede je

$$\tau_g(S) = Tr(\sum_{j=1}^{m} \varkappa_j [V(\vartheta)]^{-1}V_j[V(\vartheta)]^{-1}S),$$

kde vektor $\varkappa = (\varkappa_1,\ldots,\varkappa_m)'$ je riešením sústavy

$$M(\vartheta)\varkappa = \lambda.$$

(3) Disperzia odhadu $\tau_g(S)$ dosahuje v bode ϑ dolnú Raovu-Cramérovu hranicu

$$\mathcal{D}[\tau_g(S)] = (2/k)Tr(\sum_{i=1}^{m}\varkappa_i[V(\vartheta)]^{-1}V_i[V(\vartheta)]^{-1}\sum_{j=1}^{m}\varkappa_j[V(\vartheta)]^{-1}V_j[V(\vartheta)]^{-1}V(\vartheta)) =$$

$$= (2/k)^2\varkappa' F(\vartheta)\varkappa = \lambda' F^{-1}(\vartheta)\lambda,$$

kde $\{F(\vartheta)\}_{i,j} = (k/2)Tr([V(\vartheta)]^{-1}V_i[V(\vartheta)]^{-1}V_j) = (k/2)\{M(\vartheta)\}_{i,j}$.

Dôkaz. Pozri v [5].

<u>Teoréma 4.7.</u> Nech v replikovanom zmiešanom regresnom modeli $Y = (Y_1',\ldots,Y_m')' \sim N_{nm}[(1 \otimes X)\beta, \underline{\Sigma} = I \otimes \Sigma = \sum_{i=1}^{p}\vartheta_i(I \otimes V_i)]$ je $V = \sum_{i=1}^{p}\vartheta_{i0}V_i$ regulárna a $S = [1/(m-1)]\sum_{i=1}^{m}(Y_i - \bar{Y})(Y_i - \bar{Y})'$, $\bar{Y} = (1/m)\sum_{j=1}^{m}Y_j$. Potom ϑ_0-lokálne najlepší invariantný a nevychýlený odhad (založený na vektore Y) funkcie $g(\vartheta) = f'\vartheta$, $\vartheta \in \underline{\vartheta}$, kde $g(.)$ je nevychýlene a invariantne odhadnuteľná, má tvar

$$\tau_g(Y) = \sum_{i=1}^{p}\lambda_i\left\{(m-1)Tr(V^{-1}V_iV^{-1}S) + m\bar{Y}'(MVM)^+V_i(MVM)^+\bar{Y}\right\},$$

pričom vektor $\lambda = (\lambda_1,\ldots,\lambda_p)'$ je riešením sústavy

$$[(m-1)S_{V^{-1}} + S_{(MVM)^+}]\lambda = f,$$

$$\{S_{V^{-1}}\}_{i,j} = Tr(V^{-1}V_iV^{-1}V_j), \ i,j=1,\ldots,p, \ \{S_{(MVM)^+}\}_{i,j} = Tr[(MVM)^+V_i(MVM)^+V_j], i,j=1,\ldots p.$$

Dôkaz. Pozri [1].

<u>Poznámka 4.2.</u> Vzťah medzi odhadmi z posledných dvoch teorém je analyzovaný v [7].

L I T E R A T Ú R A

[1] Kleffe,J.: Simultaneous estimation of expectation and covariance matrix in linear models. Math.Oper.Statist.Ser.Statist.9, 1978, 443-478.

[2] Kleffe,J.: C.R.Rao's MINQUE for replicated and multivariate observations. Tech. Rept.Zimm der AdW der DDR 1979, Berlin.

[3] Kubáček,L.: Základy teórie odhadu. Bratislava, Veda 1983.

[4] Kubáček,L.: Regression model with estimated covariance matrix. Math.Slovaca 33, 1983, 395-408.

[5] Kubáček, L.: Repeated regression experiment and estimation of variance components. Math.Slovaca 34, 1984, 103-114.

[6] Kubáček, L.: Estimation of covariance components in a repeated regression experiment. Math.Slovaca 34, 1984, 155-164.

[7] Kubáček, L.: Note to C.R.Rao's MINQUE for replicated observations. Math.Slovaca (v tlači).

[8] Kubáček, L.: Reliability of calculations in the linear estimation and some problems of the analysis of a multistage regression experiment. In:Compstat'84, Proceedings in computational statistics, str.214-224, Wien, Physica Verlag 1984.

[9] Kubáček, L.: Locally best quadratic estimators. Math.Slovaca (v tlači).

[10] Kubáček, L.: Multistage regression model (zaslané do Aplikace matematiky).

[11] Pincus, R.: Estimability of parameters of the covariance matrix and variance components. Math.Oper.Statist.5, 1974, 245-248.

[12] Rao, C.R.: Least squares theory using an estimated dispersion matrix and its application to measurement of signals. In:Proc.Fifth Berkeley Symposium, Vol.I., 1967, 355-372.

[13] Rao, C.R.: Estimation of variance and covariance components. J.Multivariate Anal. 1, 1871, 257-275.

[14] Rao, C.R., Mitra, S.K.: Generalized Inverse of Matrices and Its Applications. N.York, J.Wiley 1972.

[15] Rao, C.R., Kleffe, J.: Estimation of variance components. In:Krisnaiah, P.R.ed. Handbook of Statistics, Vol.I., str.1-40, N.York, North Holland 1980.

[16] Volaufová, J., Kubáček, L.: Locally and uniformly best estimators in replicated regression model. Aplikace matematiky 28, 1983, 386-390.

[17] Zmyšlony, R.: A characterization of best linear unbiased estimators in the general linear model. Institut of Mathematics, Polish Academy of Sciences, 1978, Preprint 159.