ODVOZENÍ OBLASTI NECITLIVOSTI PRO PARAMETRY STŘEDNÍ HODNOTY REGULÁRNÍHO SMÍŠENÉHO LINEÁRNÍHO REGRESNÍHO MODELU BEZ PODMÍNEK

Hana Boháčová
Univerzita Pardubice, Fakulta ekonomicko-správní, Ústav matematiky

Abstract: *The aim of this paper is to find an explicit form of an insensitivity region for a linear function of the fixed effects parameters in a regular mixed linear regression model without constraints, to explore a possible graphical representation of this region and its significance for the determination of the estimates in the mentioned model.*

Keywords: *Insensitivity region, mixed linear regression model, fixed effects parameters, maximum likelihood estimator of fixed effect parameters, variance components.*

1. Použité značení

$\boldsymbol{I}$	jednotková matice
$\boldsymbol{A}^+$	Moore-Penroseova pseudoinverze matice $\boldsymbol{A}$ (viz [7])
$\boldsymbol{M_A}$	matice ortogonální projekce na vektorový prostor kolmý k vektorovému prostoru generovanému sloupci matice $\boldsymbol{A}$, ($\boldsymbol{M_A} = \boldsymbol{I}-\boldsymbol{A}\boldsymbol{A}^+$)
$r(\boldsymbol{A})$	hodnost matice $\boldsymbol{A}$
$tr(\boldsymbol{A})$	stopa matice $\boldsymbol{A}$, definuje se pro čtvercové matice jako součet diagonálních prvků
$\boldsymbol{Y} \sim N_n(\boldsymbol{X\beta}, \boldsymbol{\Sigma})$	náhodný vektor $\boldsymbol{Y}$ má n-rozměrné normální rozdělení se střední hodnotou $\boldsymbol{X\beta}$ a varianční maticí $\boldsymbol{\Sigma}$
$Var_{\boldsymbol{\theta 0}}[\hat{b}(\boldsymbol{\theta})]$	varianční matice odhadu $\hat{b}(\boldsymbol{\theta})$ za předpokladu, že skutečná hodnota parametru $\boldsymbol{\theta}$ je $\boldsymbol{\theta_0}$
θ_i	i-tá složka vektoru $\boldsymbol{\theta}$

2. Úvod

Jednou z často používaných metod určování bodových odhadů parametrů je metoda maximální věrohodnosti. Hlavní předností této metody jsou její asymptotické vlastnosti. Maximálně věrohodný odhad parametrů střední hodnoty ve smíšeném lineárním regresním modelu, který bude blíže popsán v následujícím odstavci, je funkcí parametrů varianční matice (v dalším textu označovaných jako varianční komponenty). K určení odhadu je tedy nejprve třeba vhodně zvolit vstupní hodnoty variančních komponent. Otázkou je, jak poznáme, zda daná volba těchto vstupních hodnot byla dobrá a zda vůbec data, která máme k dispozici umožňují určení kvalitních odhadů parametrů střední hodnoty.

3. Maximálně věrohodné odhady parametrů střední hodnoty a variančních komponent ve smíšeném lineárním regresním modelu

Uvažujme smíšený lineární regresní model

$$\boldsymbol{Y} \sim N_n(\boldsymbol{X\beta}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}}). \qquad (1)$$

Data, která máme k dispozici, jsou obsažena v observačním vektoru $\boldsymbol{Y}$, neznámé parametry tvoří vektor $\boldsymbol{\beta}$. Dále předpokládáme, že složky vektoru $\boldsymbol{Y}$ jsou v případě, že data jsou přesná (tedy nedošlo k žádné chybě při jejich získávání), lineárními funkcemi vektoru parametrů $\boldsymbol{\beta}$,

což je v modelu (1) vyjádřeno tím, že střední hodnota vektoru $\boldsymbol{Y}$ je $\boldsymbol{X\beta}$. V praxi většinou k tomuto modelu dospějeme pomocí linearizace. Nechť matice $\boldsymbol{X}$ je plné hodnosti ve sloupcích, $r(\boldsymbol{X}) = k$. Uvažujme model s r variančními komponentami $q_1,...,q_r$, tedy varianční matice vektoru $\boldsymbol{Y}$ je tvaru

$$\text{Var } \boldsymbol{Y} = \boldsymbol{\Sigma_\theta} = \sum_{i=1}^{r} q_i \, \boldsymbol{V}_i \,, \tag{2}$$

kde $\boldsymbol{V_1}, \ldots, \boldsymbol{V_r}$ jsou známé symetrické matice, přičemž musí platit, že varianční matice $\boldsymbol{\Sigma_\theta}$ je pozitivně definitní. Existence více variančních komponent v reálných situacích znamená, že data pocházejí z několika různě přesných zdrojů, například byla měřena několika různými přístroji, matice $\boldsymbol{V}_i$ pak obvykle bývají diagonální.

Naším cílem je najít odhady parametrů střední hodnoty $\boldsymbol{\beta}$ a variančních komponent $q_1,...,q_r$ metodou maximální věrohodnosti. Pro usnadnění zápisu budeme v dalším symbolem $\boldsymbol{\theta}$ značit vektor, jehož složkami jsou varianční komponenty.

V modelu (1) jsou věrohodnostní rovnice pro $\boldsymbol{\beta}$ a $\boldsymbol{\theta}$ (například podle [7]) tvaru

$$[\boldsymbol{X}'(\boldsymbol{\Sigma_\theta})^{-1}\boldsymbol{X}]\boldsymbol{\beta} = \boldsymbol{X}'(\boldsymbol{\Sigma_\theta})^{-1}\boldsymbol{Y}, \tag{3}$$

$$\text{tr}[(\boldsymbol{\Sigma_\theta})^{-1}\boldsymbol{V}_i] = \boldsymbol{Y}'(\boldsymbol{M_X}\boldsymbol{\Sigma_\theta}\boldsymbol{M_X})^{+}\boldsymbol{V}_i(\boldsymbol{M_X}\boldsymbol{\Sigma_\theta}\boldsymbol{M_X})^{+}\boldsymbol{Y}, \; i = 1,...,r\,. \tag{4}$$

Vzhledem k předpokladu pozitivní definitnosti matice $\boldsymbol{\Sigma_\theta}$ je tato varianční matice i regulární. Protože navíc předpokládáme plnou sloupcovou hodnost matice $\boldsymbol{X}$, je regulární i matice $\boldsymbol{X}'(\boldsymbol{\Sigma_\theta})^{-1}\boldsymbol{X}$ a existuje k ní tedy matice inverzní. Proto z rovnice (3) můžeme přímo vyjádřit odhad parametru $\boldsymbol{\beta}$ ve tvaru

$$\hat{b} = [\boldsymbol{X}'(\boldsymbol{\Sigma_\theta})^{-1}\boldsymbol{X}]^{-1}\boldsymbol{X}'(\boldsymbol{\Sigma_\theta})^{-1}\boldsymbol{Y} \tag{5}$$

Explicitní vyjádření odhadu parametru $\boldsymbol{\theta}$, které vznikne úpravou soustavy rovnic (4), je uvedeno v [1].

Ze vztahu (5) je vidět, že odhad parametru $\boldsymbol{\beta}$ je funkcí proměnné $\boldsymbol{\theta}$, jak bylo zmíněno v úvodu. Potřebujeme proto stanovit nějakou vhodnou vstupní hodnotu variančních komponent, kterou budeme značit $\boldsymbol{\theta_0}$. Zároveň potřebujeme umět posoudit, jestli daná volba vstupních variančních komponent umožňuje získání použitelného odhadu parametrů střední hodnoty. To nám umožní právě oblast necitlivosti, jejímž odvozením se budeme zabývat v následujícím odstavci.

Obvykle se nejprve řešením soustavy (4) získá odhad variančních komponent a ten se pak použije jako vstupní hodnota do vztahu (5). Soustava (4) se ale musí řešit iteračně a i zde je nutná nějaká vstupní počáteční hodnota variančních komponent, kterou je třeba stanovit na základě observačního vektoru $\boldsymbol{Y}$. V situaci, kdy nepotřebujeme znát odhady variančních komponent, se proto zdá být snazší použít jako vstupní hodnotu pro (5) rovnou počáteční hodnoty variančních komponent a pomocí oblasti necitlivosti pak rozhodnout, zda tato volba byla vhodná a umožnila nám získat rozumný odhad parametrů střední hodnoty.

4. Oblasti necitlivosti pro parametry střední hodnoty

Počáteční hodnoty variančních komponent vstupující do vztahu (5) mohou podstatně ovlivnit výsledné odhady, jejich volba je proto velmi důležitá. Zkusme si položit otázku, co se stane, změníme-li varianční komponenty $\boldsymbol{\theta_0}$ vstupující do vzorce (5) o $\delta\boldsymbol{\theta}$. Výsledný odhad parametrů střední hodnoty pak můžeme přibližně vyjádřit pomocí diferenciálu následovně

$$\hat{b}\,(\boldsymbol{\theta}_0+\delta\boldsymbol{\theta}) \approx \hat{b}\,(\boldsymbol{\theta}_0)+\frac{\partial \hat{b}(q_0)}{\partial q'}\delta\boldsymbol{\theta}. \tag{6}$$

Protože po úpravě

$$\frac{\partial \hat{b}(q_0)}{\partial q_i} = -\,[\boldsymbol{X}'(\boldsymbol{\Sigma}_{\theta 0})^{-1}\boldsymbol{X}]^{-1}\boldsymbol{X}'(\boldsymbol{\Sigma}_{\theta 0})^{-1}\,\boldsymbol{V}_i(\boldsymbol{\Sigma}_{\theta 0})^{-1}[\boldsymbol{Y}-\boldsymbol{X}\,\hat{b}\,(\boldsymbol{\theta}_0)], \tag{7}$$

(kde $\boldsymbol{\Sigma}_{\theta 0}$ značí matici typu (2) s $\boldsymbol{\theta}_0$ místo $\boldsymbol{\theta}$), můžeme podle (6) psát

$$\hat{b}\,(\boldsymbol{\theta}_0+\delta\boldsymbol{\theta}) \approx \hat{b}\,(\boldsymbol{\theta}_0)-\sum_{i=1}^{r}\,[\boldsymbol{X}'(\boldsymbol{\Sigma}_{\theta 0})^{-1}\boldsymbol{X}]^{-1}\boldsymbol{X}'(\boldsymbol{\Sigma}_{\theta 0})^{-1}\,\boldsymbol{V}_i(\boldsymbol{\Sigma}_{\theta 0})^{-1}[\boldsymbol{Y}-\boldsymbol{X}\,\hat{b}\,(\boldsymbol{\theta}_0)]\delta\theta_i. \tag{8}$$

Z hlediska kvality výsledného odhadu se zdá být rozumné, aby při změně vstupních hodnot variančních komponent nedošlo k přílišnému nárůstu disperze odhadu parametrů střední hodnoty. Budeme tedy hledat takovou množinu vstupních hodnot $\boldsymbol{\theta}_0+\delta\boldsymbol{\theta}$, které nepovedou ke zvýšení směrodatné odchylky odhadu dané lineární kombinace složek vektoru $\boldsymbol{\beta}$ o více než $100\cdot e$ % v porovnání se směrodatnou odchylkou odhadu stejné lineární kombinace vycházejícího ze vstupní hodnoty $\boldsymbol{\theta}_0$. Takovou množinu nazveme oblastí necitlivosti pro lineární funkci $\boldsymbol{h}'\boldsymbol{\beta}$ pro dané $\boldsymbol{\theta}_0$ a dané ε, v dalším textu ji pro stručnost budeme nazývat oblastí necitlivosti pro parametry střední hodnoty. To, že místo směrodatných odchylek odhadů jednotlivých složek parametru $\boldsymbol{\beta}$ zkoumáme směrodatné odchylky lineárních kombinací složek vektoru $\boldsymbol{\beta}$, má své opodstatnění. Zvolíme-li za vektor $\boldsymbol{h}$ i-tý jednotkový vektor, je příslušná lineární kombinace rovna i-té složce vektoru $\boldsymbol{\beta}$, o možnost sledovat směrodatné odchylky složek odhadu jsme tedy nepřišli. Naopak máme navíc možnost věnovat se i lineárním kombinacím určeným jinými vektory než jednotkovými. Podle výše uvedeného hledáme množinu takových $\boldsymbol{\theta}_0+\delta\boldsymbol{\theta}$, pro která platí

$$\{Var_{\theta 0}\,[\boldsymbol{h}'\,\hat{b}\,(\boldsymbol{\theta}_0+\delta\boldsymbol{\theta})]\}^{1/2} \le (1+\varepsilon)\,\{Var_{\theta 0}\,[\boldsymbol{h}'\,\hat{b}\,(\boldsymbol{\theta}_0)]\}^{1/2}\,. \tag{9}$$

Na základě vztahu (5) dostaneme

$$Var_{\theta 0}\,[\,\hat{b}\,(\boldsymbol{\theta}_0)] = [\boldsymbol{X}'(\boldsymbol{\Sigma}_{\theta 0})^{-1}\boldsymbol{X}]^{-1}, \tag{10}$$

tedy

$$Var_{\theta 0}\,[\boldsymbol{h}'\,\hat{b}\,(\boldsymbol{\theta}_0)] = \boldsymbol{h}'\,[\boldsymbol{X}'(\boldsymbol{\Sigma}_{\theta 0})^{-1}\boldsymbol{X}]^{-1}\boldsymbol{h}. \tag{11}$$

Na základě (8) můžeme odvodit následující přibližné vyjádření

$$\begin{aligned} Var_{\theta 0}\,[\,\hat{b}\,(\boldsymbol{\theta}_0+\delta\boldsymbol{\theta})] &\approx Var_{\theta 0}\,[\,\hat{b}\,(\boldsymbol{\theta}_0)]+ \\ +\,[\boldsymbol{X}'(\boldsymbol{\Sigma}_{\theta 0})^{-1}\boldsymbol{X}]^{-1}\boldsymbol{X}'(\boldsymbol{\Sigma}_{\theta 0})^{-1}\boldsymbol{\Sigma}_{\delta\theta}\,(\boldsymbol{M}_X\boldsymbol{\Sigma}_{\theta 0}\boldsymbol{M}_X)^{+}\,\boldsymbol{\Sigma}_{\delta\theta}(\boldsymbol{\Sigma}_{\theta 0})^{-1}\boldsymbol{X}[\boldsymbol{X}'(\boldsymbol{\Sigma}_{\theta 0})^{-1}\boldsymbol{X}]^{-1}. \end{aligned} \tag{12}$$

$\boldsymbol{\Sigma}_{\delta\theta}$ zde značí matici $\sum_{i=1}^{r} dq_i\,\boldsymbol{V}_i$. Výslednou oblast necitlivosti pak označíme $N_{\boldsymbol{h}'\boldsymbol{\beta},\boldsymbol{\theta}}$. Po úpravách ji můžeme zapsat takto:

$$N_{\boldsymbol{h}'\boldsymbol{\beta},\boldsymbol{\theta}0} = \{\boldsymbol{\theta}_0+\delta\boldsymbol{\theta}:\,(\delta\boldsymbol{\theta})'\boldsymbol{W}_{\boldsymbol{h}}\delta\boldsymbol{\theta} \le (2\varepsilon+\varepsilon^2)\boldsymbol{h}'[\boldsymbol{X}'(\boldsymbol{\Sigma}_{\theta 0})^{-1}\boldsymbol{X}]^{-1}\boldsymbol{h}\}, \tag{13}$$

kde $\boldsymbol{W}_{\boldsymbol{h}}$ je matice s prvky

$$\{\boldsymbol{W}_{\boldsymbol{h}}\}_{i,j} = \boldsymbol{h}'[\boldsymbol{X}'(\boldsymbol{\Sigma}_{\theta 0})^{-1}\boldsymbol{X}]^{-1}\boldsymbol{X}'(\boldsymbol{\Sigma}_{\theta 0})^{-1}\boldsymbol{V}_i(\boldsymbol{M}_X\boldsymbol{\Sigma}_{\theta 0}\boldsymbol{M}_X)^{+}\boldsymbol{V}_j(\boldsymbol{\Sigma}_{\theta 0})^{-1}\boldsymbol{X}[\boldsymbol{X}'(\boldsymbol{\Sigma}_{\theta 0})^{-1}\boldsymbol{X}]^{-1}\boldsymbol{h}. \tag{14}$$

Protože e je malé kladné číslo (obvykle se volí $e < 0{,}5$), můžeme většinou v (14) použít $2e$ místo $(2e + e^2)$.

Podrobnější postup odvození oblastí necitlivosti je možné najít např. v [3], [4] nebo [6]. Dá se dokázat, že matice $\boldsymbol{W_h}$ je singulární a vektor vstupních variančních komponent $\boldsymbol{\theta_0}$ je kolmý na prostor generovaný sloupci matice $\boldsymbol{W_h}$. Důkaz je možné najít např. v [6]. Znamená to, že kvadratická forma $(\delta\boldsymbol{\theta})'\boldsymbol{W_h}\delta\boldsymbol{\theta}$ uvnitř množiny (13) určuje singulární kuželosečku. Pro $r = 2$ je tak oblast necitlivosti vlastně pás vymezený dvěma rovnoběžnými přímkami, které jsou navíc rovnoběžné s orientovanou úsečkou spojující počátek soustavy souřadnic s bodem $\boldsymbol{\theta_0}$.

5. Použití oblastí necitlivosti

Oblasti necitlivosti jsou, jak víme, množinami možných vstupních hodnot variančních komponent, které nezpůsobí příliš velký nárůst směrodatné odchylky odhadů. Abychom mohli posoudit kvalitu odhadů parametrů střední hodnoty, respektive rozhodnout o tom, jestli za odhad parametru můžeme považovat už hodnotu určenou ze vztahu (5) dosazením počátečních hodnot variančních komponent, porovnáme oblast necitlivosti pro parametry střední hodnoty s oblastí spolehlivosti pro varianční komponenty. Nevyplývá-li z konkrétní situace jiná potřebná volba vektoru $\boldsymbol{h}$, volíme obvykle za $\boldsymbol{h}$ postupně jednotkové vektory a všechny příslušné oblasti necitlivosti porovnáváme s oblastí spolehlivosti. Odvození oblasti spolehlivosti pro varianční komponenty si vysvětlíme pro případ dvou variančních komponent – tedy pro situaci, kdy $r = 2$, pro větší hodnoty r je postup analogický. Oblast spolehlivosti budeme hledat ve tvaru obdélníku se středem v bodě $\hat{q}(\boldsymbol{\theta_0})$ (tj. odhad variančních komponent získaný při použití počáteční hodnoty $\boldsymbol{\theta_0}$ první iterací soustavy (4)), který pokrývá skutečné hodnoty variančních komponent s danou pravděpodobností 1-α. Podle Čebyševovy nerovnosti (viz např. [1]) platí

$$P\left\{ \left|\hat{q}_1(q_0) - q_1\right| \le k\sqrt{Var_{q0}\left[\hat{q}_1(q_0)\right]} \right\} \ge 1 - \frac{1}{k^2} \tag{15}$$

a podobně

$$P\left\{ \left|\hat{q}_2(q_0) - q_2\right| \le k\sqrt{Var_{q0}\left[\hat{q}_2(q_0)\right]} \right\} \ge 1 - \frac{1}{k^2}. \tag{16}$$

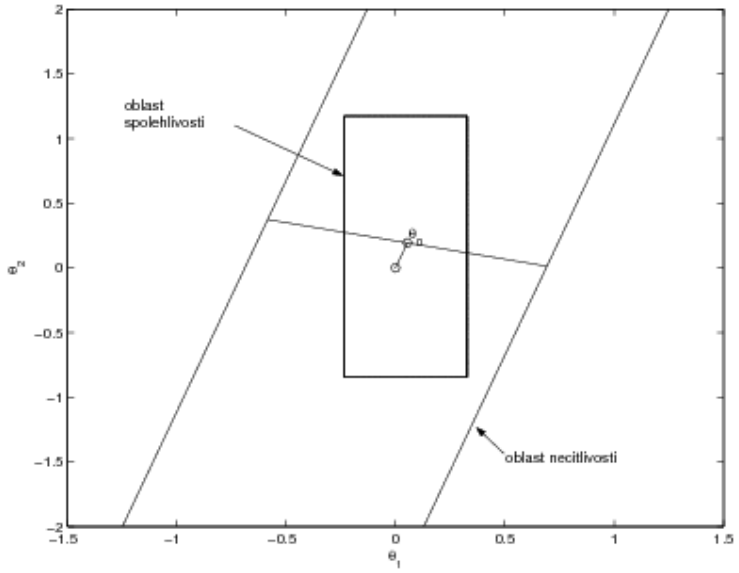
Bonferroniho nerovnost, která je uvedena např. v [5], pak zohledňuje (15) i (16):

$$P\left\{ \left|\hat{q}_1(q_0) - q_1\right| \le k\sqrt{Var_{q0}\left[\hat{q}_1(q_0)\right]} \wedge \left|\hat{q}_2(q_0) - q_2\right| \le k\sqrt{Var_{q0}\left[\hat{q}_2(q_0)\right]} \right\} \ge 1 - \frac{2}{k^2}. \tag{17}$$

Potřebujeme, aby $1 - \frac{2}{k^2} = 1 - a$, tedy $k = \sqrt{\frac{2}{a}}$. Oblast spolehlivosti pro varianční komponenty $\mathrm{E}_{\boldsymbol{\theta},\boldsymbol{\theta 0}}$ je tedy množina

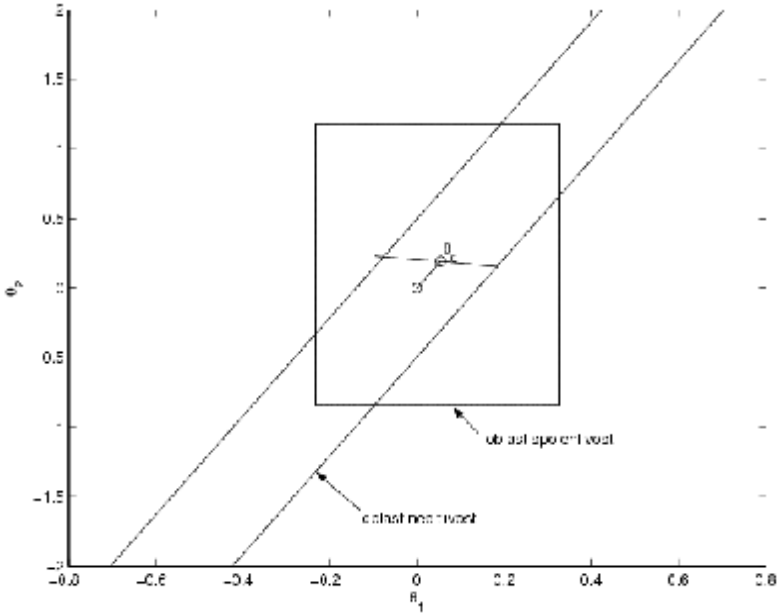
$$\mathrm{E}_{\boldsymbol{\theta},\boldsymbol{\theta 0}} = \left\{ q = \begin{pmatrix} q_1 \\ q_2 \end{pmatrix} : \left|\hat{q}_1(q_0) - q_1\right| \le \sqrt{\frac{2}{a} Var_{q0}\left[\hat{q}_1(q_0)\right]} \wedge \left|\hat{q}_2(q_0) - q_2\right| \le \sqrt{\frac{2}{a} Var_{q0}\left[\hat{q}_2(q_0)\right]} \right\}. \tag{18}$$

Pokud bude v konkrétním případě obdélník představující oblast spolehlivosti uvnitř pásu oblasti necitlivosti, jak je uvedeno na obrázku 1, znamená to, že s $\hat{b}(\boldsymbol{\theta_0})$ můžeme pracovat jako s kvalitním odhadem parametrů střední hodnoty.



Obr. 1: Vzájemná poloha oblasti necitlivosti a oblasti spolehlivosti – kvalitní odhad parametru

Pokud ovšem bude oblast spolehlivosti oblast necitlivosti výrazně přesahovat, můžeme se při určování odhadů velmi snadno dostat do situace, kdy disperze složek určeného odhadu budou příliš velké a takovému odhadu pak nemůžeme přikládat velkou váhu. Vzájemná poloha oblasti necitlivosti a oblasti spolehlivosti odpovídající takové situaci je na obrázku 2.



Obr. 2: Vzájemná poloha oblasti necitlivosti a oblasti spolehlivosti – nekvalitní odhad parametru

6. Závěr

Pokud by při odhadování parametrů střední hodnoty konkrétního modelu nastala vzájemná poloha oblasti necitlivosti a oblasti spolehlivosti podobná té na obrázku 2, je třeba zvolit místo $\boldsymbol{\theta_0}$ jinou vstupní hodnotu variančních komponent. Můžeme třeba $\boldsymbol{\theta_0}$ použít jako počáteční hodnotu do soustavy rovnic (4), tyto rovnice iteračně vyřešit a výsledný odhad variančních komponent pak použít pro získání odhadu parametrů střední hodnoty. Pokud ani toto nepomůže, je třeba zkoušet jiné počáteční hodnoty vstupující do výše zmíněné iterační

procedury. Ukazuje se však, že v některých modelech vůbec není možné volbou vstupních variančních komponent docílit vzájemné polohy jako na obrázku 1.

Použitá literatura:

[1] ANDĚL, J. *Statistické metody*. 2. vydání. Praha: MATFYZPRESS, 2003.300 s. ISBN 80-85863-27-8.

[2] BOHÁČOVÁ, H. Odhad parametrů střední hodnoty a parametrů varianční matice ve smíšeném lineárním modelu s podmínkami typu I a II. *In Scientific papers of the University of Pardubice – Series D*, 2007, s. 5-10. ISSN 1211-555X.

[3] BOHÁČOVÁ, H., HECKENBERGEROVÁ, J. Oblasti necitlivosti pro parametry střední hodnoty ve smíšeném lineárním regresním modelu s podmínkami typu I a s nimi spojené výpočetní problémy. *In Forum Statisticum Slovacum*, 2007, roč. 3, č. 6, s. 31-35. ISSN 1336-7420.

[4] BOHÁČOVÁ, H. Insensitivity region for variance components in general linear model. *In Acta Universitatis Palackianae Olomucensis, Facultas Rerum Naturalium, Mathematica*, 2008, roč. 49, s. 7-22. ISSN 0231-9721

[5] HUMAK, K. M. S. *Statistische Methoden der Modellbildung, Band I Statistische Inferen fűr lineare Parameter*, Berlin: Akademie – Verlag, 1977, 516 s.

[6] KUBÁČEK, L., KUBÁČKOVÁ, L. *Statistika a metrologie.* Olomouc: Univerzita Palackého v Olomouci– vydavatelství, 2000. 307 s. ISBN 80-244-0093-6.

[7] RAO, C. R., KLEFFE, J. *Estimation of Variance Components and Applications*, Amsterdam – New York – Oxford – Tokyo: North-Holland, 1988, 496 s. ISBN 0-444-70023-4

Kontaktní adresa:

Mgr. Hana Boháčová
Univerzita Pardubice
Fakulta ekonomicko-správní
Ústave matematiky
Studentská 84
532 10 Pardubice
Email: Hana.Bohacova @upce.cz