DESIGN EXPERIMENTU PRO REGRESNÍ MODELY S PODMÍNKAMI TYPU I

TUČKOVÁ M.[1,2,*], KUBÁČEK L.[1], TUČEK P.[2]

UNIVERSITAS PALACKIANA OLOMUCENSIS

[1] Katedra matematické analýzy a aplikací matematiky
[2] Katedra geoinformatiky
Přírodovědecká fakulta, Univerzita Palackého v Olomouci
*korespondující autor: michaela.tuckova@upol.cz

Přírodovědecká fakulta

SHRNUTÍ

Poster prezentuje hledání lokálně optimálního návrhu měření pro regresní modely s podmínkami typu I. Zabýváme se kriteriálními funkcemi lokální A-optimality, C-optimality a D-optimality, pro které odvozujeme konkrétní podoby gradientů a následně demonstrujeme iterační proces výpočtu.

REGRESNÍ MODEL S PODMÍNKAMI TYPU I

Uvažujeme regulární nelineární model s nelineárními podmínkami typu I daný vztahem

$$\mathbf{Y} = \phi(\mathbf{x}, \boldsymbol{\theta}) + \boldsymbol{\epsilon}, \qquad \eta_j(\boldsymbol{\theta}) = 0, \quad j = 1, \dots, q,$$

kde $\mathbf{Y}$ je observační vektor, ϕ a η_j jsou známé nelineární funkce, $\mathbf{x} = \{x_1, \dots, x_n\}$ je množina experimentálních bodů, $\boldsymbol{\theta} = (\theta_1, \dots, \theta_k)^T$ je vektor neznámých parametrů a $\boldsymbol{\epsilon}$ je vektor chyb měření se $E(\boldsymbol{\epsilon}) = 0$ a $\mathrm{Var}(\boldsymbol{\epsilon}) = \sigma^2 \cdot \mathbf{I}$. Jestliže označíme

$$\mathbf{F} = \left.\frac{\partial \phi(\mathbf{x}, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}^T}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}^0} \text{ jako } (n \times k) \text{ matici plánu}, \quad \mathbf{B} = \left.\frac{\partial \eta_j(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^T}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}^0} \text{ jako } (q \times k) \text{ matici podmínek},$$

potom můžeme psát model v lineární formě jako

$$\begin{aligned} \mathbf{Y} &\sim \phi(\mathbf{x}, \boldsymbol{\theta}^0) + \left.\frac{\partial \phi(\mathbf{x}, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}^T} \cdot \delta\boldsymbol{\theta}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}^0} + \boldsymbol{\epsilon} &&\Rightarrow \quad \mathbf{Y} - \phi(\mathbf{x}, \boldsymbol{\theta}^0) = \mathbf{F} \cdot \delta\boldsymbol{\theta} + \boldsymbol{\epsilon}, \\ \eta_j(\boldsymbol{\theta}) &\sim \eta_j(\boldsymbol{\theta}^0) + \left.\frac{\partial \eta_j(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^T} \cdot \delta\boldsymbol{\theta}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}^0} = 0 &&\Rightarrow \quad \eta_j(\boldsymbol{\theta}) - \eta_j(\boldsymbol{\theta}^0) = \{\mathbf{B}\}_j \cdot \delta\boldsymbol{\theta} = 0, \quad j = 1, \dots, q, \end{aligned}$$

kde $\delta\boldsymbol{\theta} = \boldsymbol{\theta} - \boldsymbol{\theta}^0$. Jestliže ξ je návrh (pravděpodobnostní míra na množině experimentálních bodů $\mathbf{x}$), potom informační matice pro parametr $\boldsymbol{\theta}$ je tvaru

$$\mathbf{M}(\boldsymbol{\theta}) = N \cdot \sigma^{-2} \cdot \underbrace{\sum_{x_i \in \mathrm{Sp}(\xi)} \left(\frac{\partial \phi(x_i, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right) \cdot \left(\frac{\partial \phi(x_i, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right)^T \cdot \xi(x_i)}_{\mathbf{M}(\boldsymbol{\theta}, \xi)},$$

kde $\mathbf{M}(\boldsymbol{\theta}, \xi)$ je normovaná informační matice, $\mathrm{Sp}(\xi) = \{x_i : \xi(x_i) > 0, x_i \in \mathbf{x}\}$ je nosič návrhu ξ a N je celkový počet měření. Jestliže $\mathbf{M} = \mathbf{M}(\boldsymbol{\theta})$ je regulární matice, potom platí, že varianční matice NLNO $\widehat{\boldsymbol{\theta}}$ je tvaru

$$\mathrm{Var}(\widehat{\boldsymbol{\theta}}) = \mathbf{M}^{-1} - \mathbf{M}^{-1}\mathbf{B}^T \left[\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T\right]^{-1} \mathbf{B}\mathbf{M}^{-1} \qquad \Rightarrow \qquad \mathbf{V}_{\boldsymbol{\theta}}(\xi) = N \cdot \sigma^{-2} \cdot \mathrm{Var}(\widehat{\boldsymbol{\theta}}),$$

kde $\mathbf{V}_{\boldsymbol{\theta}}(\xi)$ je normovaná varianční matice $\widehat{\boldsymbol{\theta}}$ (viz. [2]).
Uvažujeme-li konvexní kriteriální funkci Φ, pak je lokálně optimální návrh ξ^* takový návrh, který splňuje podmínku: $\Phi\left[\mathbf{V}_{\boldsymbol{\theta}}(\xi^*)\right] = \min_{\xi \in \Xi} \Phi\left[\mathbf{V}_{\boldsymbol{\theta}}(\xi)\right]$, kde Ξ je konvexní množina všech návrhů.

KRITÉRIUM LOKÁLNÍ A-OPTIMALITY

Kriteriální funkce lokální A-optimality pro model s podmínkami typu I je dána vztahem

$$\boxed{\Phi\left[\mathbf{V}_{\boldsymbol{\theta}}(\xi)\right] = \mathrm{Tr}\left[\mathbf{V}_{\boldsymbol{\theta}}(\xi)\right] = \mathrm{Tr}\left[N \cdot \sigma^{-2}\left[\mathbf{M}^{-1} - \mathbf{M}^{-1}\mathbf{B}^T(\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\mathbf{B}\mathbf{M}^{-1}\right]\right].}$$

Pro výpočetní algoritmus je třeba určit gradient $\nabla\Phi(\mathbf{V})$ kriteriální funkce, tj.

$$\begin{aligned} \nabla\Phi(\mathbf{V}) = N \cdot \sigma^{-2} \Big[&-\mathbf{M}^{-2} + \mathbf{M}^{-1}\mathbf{B}^T(\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\mathbf{B}\mathbf{M}^{-2} \\ &-\mathbf{M}^{-1}\mathbf{B}^T(\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\mathbf{B}\mathbf{M}^{-2}\mathbf{B}^T(\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\mathbf{B}\mathbf{M}^{-1} + \mathbf{M}^{-2}\mathbf{B}^T(\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\mathbf{B}\mathbf{M}^{-1}\Big]. \end{aligned}$$

KRITÉRIUM LOKÁLNÍ C-OPTIMALITY

K definování této kriteriální funkce je nutná existence známé funkce parametrů $h(\boldsymbol{\theta})$, která je diferencovatelná. Při splnění těchto předpokladů pak pro regulární regresní model platí

$$\mathrm{Var}(h(\widehat{\boldsymbol{\theta}})) = \partial h^T \mathrm{Var}(\widehat{\boldsymbol{\theta}}) \partial h \qquad \Rightarrow \qquad \mathbf{V}^h_{\boldsymbol{\theta}}(\xi) = N \cdot \sigma^{-2} \cdot \mathrm{Var}\left(h(\widehat{\boldsymbol{\theta}})\right).$$

Následně je kritérium lokální C-optimality definováno vztahem

$$\boxed{\Phi\left[\mathbf{V}^h_{\boldsymbol{\theta}}(\xi)\right] = \partial h^T(\boldsymbol{\theta})\mathbf{V}^h_{\boldsymbol{\theta}}(\xi)\partial h(\boldsymbol{\theta}) = N \cdot \sigma^{-2}\left[\partial h^T(\boldsymbol{\theta})\left[\mathbf{M}^{-1} - \mathbf{M}^{-1}\mathbf{B}^T(\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\mathbf{B}\mathbf{M}^{-1}\right]\partial h(\boldsymbol{\theta})\right].}$$

Gradient $\nabla\Phi(\mathbf{V})$ kriteriální funkce je tvaru

$$\begin{aligned} \nabla\Phi(\mathbf{V}) = N \cdot \sigma^{-2}[&- \mathbf{M}^{-1}\partial h \partial h^T \mathbf{M}^{-1} + \mathbf{M}^{-1}\mathbf{B}^T(\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\mathbf{B}\mathbf{M}^{-1}\partial h \partial h^T \mathbf{M}^{-1} \\ &- \mathbf{M}^{-1}\mathbf{B}^T(\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\mathbf{B}\mathbf{M}^{-1}\partial h \partial h^T \mathbf{M}^{-1}\mathbf{B}^T(\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\mathbf{B}\mathbf{M}^{-1} \\ &+ \mathbf{M}^{-1}\partial h \partial h^T \mathbf{M}^{-1}\mathbf{B}^T(\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\mathbf{B}\mathbf{M}^{-1}]. \end{aligned}$$

KRITÉRIUM LOKÁLNÍ D-OPTIMALITY

V regresním modelu s podmínkami nelze apriorně definovat kritérium lokální D-optimality vztahem známým z regresních modelů bez podmínek, tedy jako

$$\Phi\left[\mathbf{V}_{\boldsymbol{\theta}}(\xi)\right] = \ln\det\left[\mathbf{V}_{\boldsymbol{\theta}}(\xi)\right]$$

neboť normovaná varianční matice $\mathbf{V}_{\boldsymbol{\theta}}(\xi)$ je často singulární, přestože je parametr $\boldsymbol{\theta}$ odhadnutelný. Z tohoto důvodu je třeba k problematice lokálně D-optimálního návrhu zvolit odlišný přístup než v případě předchozích kriteriálních funkcí.
Matice podmínek $\mathbf{B}$ je rozdělena do bloků $\mathbf{B} = (\mathbf{B}_1, \mathbf{B}_2)$ tak, aby platilo $r(\mathbf{B}_1) = q$. Pak je možné zapsat podmínky typu I jako

$$\mathbf{B}_1\delta\boldsymbol{\theta}_1 + \mathbf{B}_2\delta\boldsymbol{\theta}_2 = \mathbf{0} \quad \rightarrow \quad \delta\boldsymbol{\theta}_1 = -\mathbf{B}_1^{-1}\mathbf{B}_2\delta\boldsymbol{\theta}_2,$$

čímž je původní model s podmínkami typu I převeden na model bez podmínek:

$$\mathbf{Y} \sim \underbrace{\mathbf{F}\begin{pmatrix} -\mathbf{B}_1^{-1}\mathbf{B}_2 \\ \mathbf{I} \end{pmatrix}}_{\mathbf{F}^*} \delta\boldsymbol{\theta}_2 + \boldsymbol{\epsilon},$$

kde $\mathbf{F}^*$ je nová matice plánu. Normovaná informační matice nového modelu je tvaru

$$\mathbf{M}^*(\boldsymbol{\theta}_2, \xi) = \sum_{x_i \in \mathrm{Sp}(\xi)} \{\mathbf{F}^*\}_i \{\mathbf{F}^*\}_i^T \xi(x_i),$$

kde $\{\mathbf{F}^*\}_i^T$ je $i - t\acute{y}$ řádek nové matice plánu $\mathbf{F}^*$. Tímto postupem je zaručena regulárnost matice $\mathbf{M}^*(\boldsymbol{\theta}_2, \xi)$ (důkaz viz. [1]), takže je možné definovat kriteriální funkci lokální D-optimality stejným způsobem jako je tomu v regresních modelech bez podmínky, tedy

$$\boxed{\Phi\left[\mathbf{M}^*(\boldsymbol{\theta}_2, \xi)\right] = -\ln\left[\det(\mathbf{M}^*(\boldsymbol{\theta}_2, \xi))\right].}$$

Gradient $\nabla\Phi(\mathbf{M}^*)$ kriteriální funkce je vyjádřen jako

$$\nabla\Phi(\mathbf{M}^*) = -\left[\mathbf{M}^*(\boldsymbol{\theta}_2, \xi)\right]^{-1}.$$

K uvedené problematice je možné zvolit alternativní přístup užitím QR rozkladu matice B^T. Popis tohoto přístupu je uveden ve [2].

Příklad

Pro jednoduchost uvažujme lineární regresní model s podmínkou typu I

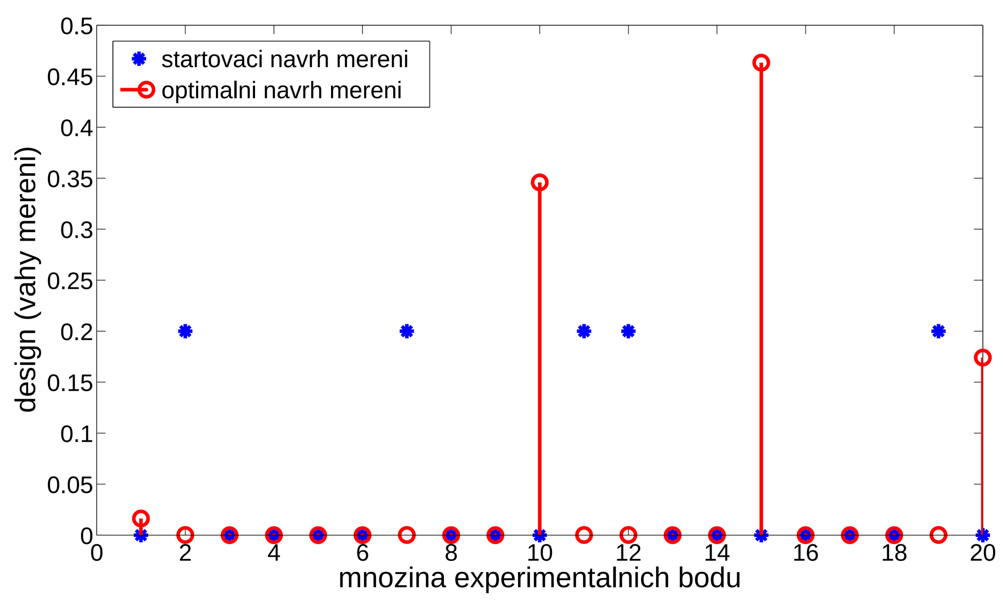
$$\begin{aligned} Y_i &= \theta_1 + \theta_2 \cdot x_i + \epsilon_i, \quad \text{pro } i = 1, \dots, 10 \\ Y_i &= \theta_3 + \theta_4 \cdot x_i + \theta_5 \cdot x_i^2 + \epsilon_i, \quad \text{pro } i = 11, \dots, 20 \\ \theta_1 &+ \theta_2 \cdot x_{10} - (\theta_3 + \theta_4 \cdot x_{10} + \theta_5 \cdot x_{10}^2) = 0. \end{aligned}$$

Dále uvažujme celkový počet měření $N = 20$, množinu experimentálních bodů $\mathbf{x} = (1, \dots, 20)$, matici plánu $\mathbf{F}$ a matici podmínek $\mathbf{B}$ tvaru:

$$\mathbf{F} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ & & \vdots & & \\ 1 & 10 & 0 & 0 & 0 \\ 0 & 0 & 1 & 11 & 11^2 \\ & & \vdots & & \\ 0 & 0 & 1 & 20 & 20^2 \end{pmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 1 & 10 & -1 & -10 & -100 \end{bmatrix}.$$

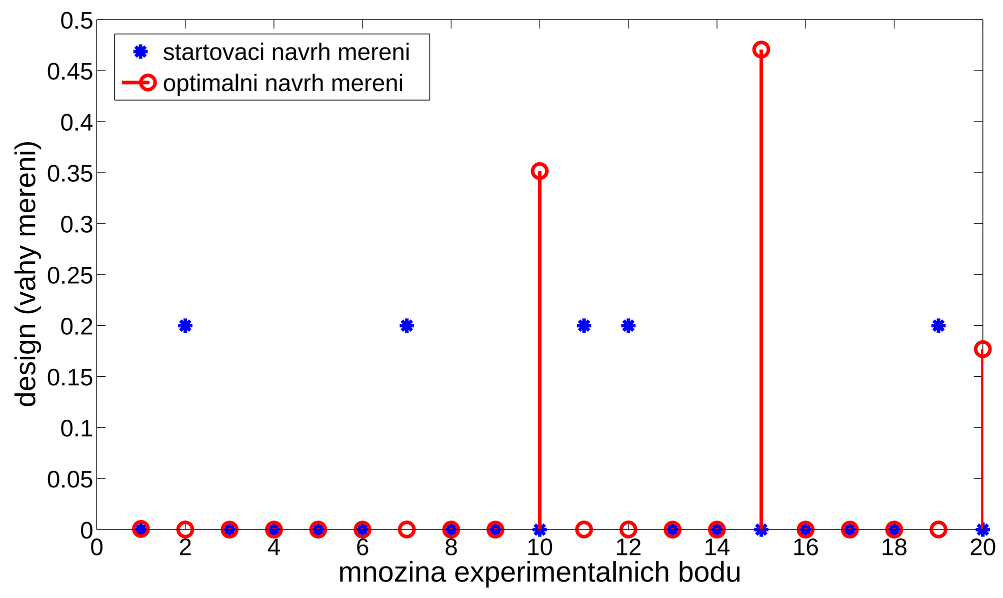
Aby byl parametr $\boldsymbol{\theta}$ odhadnutelný z jakéhokoliv návrhu ξ_j, kde $j = 1, 2, \dots$ v průběhu iteračního procesu, je nutné zajistit odhadnutelnost parametru $\boldsymbol{\theta}$ ze startovacího návrhu ξ_0. To nastane tehdy, bude-li dodržena regulárnost matice $\mathbf{M}(\boldsymbol{\theta}, \xi_0) + \mathbf{B}^T\mathbf{B}$. Tuto podmínku například splňuje startovací návrh rovnoměrně rozdělený do bodů $\mathrm{Sp}(\xi_0) = \{2, 7, 11, 12, 19\}$.

Nejdříve stanovíme lokálně A-optimální návrh měření. Určíme gradient kriteriální funkce, který dosadíme do výpočetního algoritmu založeného na metodě největšího spádu.
Z obrázku 1 je patrné, že nosič lokálně A-optimálního návrhu obsahuje 4 body, kdy relativní počet měření v každém tomto bodě je:
$\xi(x_1) \cdot N = 0.0164 \cdot 20 \doteq 1$,
$\xi(x_{10}) \cdot N = 0.3459 \cdot 20 \doteq 7$,
$\xi(x_{15}) \cdot N = 0.4633 \cdot 20 \doteq 9$,
$\xi(x_{20}) \cdot N = 0.1743 \cdot 20 \doteq 3$.



Obr. 1.: Lokálně A-optimální návrh měření.

Nyní najdeme lokálně C-optimální návrh měření. Matice $\mathbf{F}$ a matice $\mathbf{B}$ mají stejnou strukturu jako v předchozím případě, avšak vyskytuje se zde navíc diferencovatelná funkce parametrů $h(\boldsymbol{\theta})$ ve formě: $0 \cdot \theta_1 + 0 \cdot \theta_2 + 1 \cdot \theta_3 - 2 \cdot \theta_4 - 0.5 \cdot \theta_5$. Potom $\partial h^T = [0, 0, 1, -2, -0.5]$.
Z obrázku 2 je patrné, že nosič lokálně C-optimálního návrhu obsahuje 3 body, kdy relativní počet měření v každém tomto bodě je:
$\xi(x_{10}) \cdot N = 0.3518 \cdot 20 \doteq 7$,
$\xi(x_{15}) \cdot N = 0.4710 \cdot 20 \doteq 9$,
$\xi(x_{20}) \cdot N = 0.1772 \cdot 20 \doteq 4$.



Obr. 2.: Lokálně C-optimální návrh měření.

Podrobný popis výpočetního algoritmu je uveden ve [4]. Pro situaci regresního modelu s podmínkou je však potřebné tento algoritmus upravit jak je naznačeno ve [2].

Plány do budoucna

V budoucnu bychom se rádi zabývali dalšími kritérii optimality v regresních modelech s podmínkami typu I i II. Pozornost chceme věnovat i ověření konvergence jiných výpočetních algoritmů než zde prezentovanému.

Poděkování. Autoři by rádi poděkovali podpoře projektu CZ.1.07/2.2.00/15.0276 spolufinancovaného Evropským sociálním fondem (Operační program vzdělávání pro konkurenceschopnost) a státním rozpočtem České republiky. Dále Studentským projektům PrF_2012_08 a PrF_2012_17 Grantové agentury Univerzity Palackého v Olomouci.

Reference.

[1] Fišerová E., Kubáček L., Kunderová P. (2007). *Linear statistical models. Regularity and Singularities.* Academia, Praha.

[2] Pázman A. (2002). *Optimal design of nonlinar experiments with parameter constraints.* Metrika (2002) 56: 113–130.

[3] Pukelsheim F. (1993). *Optimal design of experiments.* Wiley and Sons, New York.

[4] Wynn H. P. (1972). *Results in the Theory and Construction of D-optimum Experimental Designs.* Journal of the Royal Statistical Society. Series B. Vol 34., No 2. (1972) 133–147