DESIGN EXPERIMENTU PRO REGRESNÍ MODELY S PODMÍNKAMI TYPU I

M. Tučková[1,2], L. Kubáček[1] a P. Tuček[2]

[1]Katedra matematické analýzy a aplikací matematiky
Univerzita Palackého v Olomouci

[2]Katedra geoinformatiky
Univerzita Palackého v Olomouci

Robust 2012
9. – 14. září 2012, Němčičky

Obsah prezentace

Motivace

Přestože dnes existuje mnoho odvětví optimálního navrhování experimentů, zůstává problematika optimálního navrhování experimentů v regresních modelech s podmínkami do jisté míry ”zapomenutým” tématem.

Podrobněji se touto problematikou zabýval pouze prof. Andrej Pázman ve své publikaci:

Optimal design of nonlinear experiments with paramater constraints. Metrika (2002) 56: 113-130

Popis modelu

Nechť je dán regulární nelineární model

$$\mathbf{Y} = \phi(\mathbf{x}, \boldsymbol{\theta}) + \boldsymbol{\epsilon}$$

s nelineárními podmínkami typu I

$$\eta_j(\boldsymbol{\theta}) = 0, \qquad j = 1, \dots, q,$$

kde

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

Linearizace modelu

Užitím Taylorova rozvoje se zanedbáním členů 2. a vyšších řádů:

$$\begin{aligned} \mathbf{Y} &\sim \phi(\mathbf{x}, \boldsymbol{\theta}^0) + \underbrace{\frac{\partial \phi(\mathbf{x}, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}^T}}_{\mathbf{F}} \cdot \delta\boldsymbol{\theta} \Bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}^0} + \boldsymbol{\epsilon} \\ \mathbf{Y} &= \mathbf{F} \cdot \delta\boldsymbol{\theta} + \boldsymbol{\epsilon}, \end{aligned}$$

$$\begin{aligned} \eta_j(\boldsymbol{\theta}) &\sim \eta_j(\boldsymbol{\theta}^0) + \underbrace{\frac{\partial \eta_j(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^T}}_{\mathbf{B}} \cdot \delta\boldsymbol{\theta} \Bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}^0} = 0 \\ \eta_j(\boldsymbol{\theta}) &= \mathbf{B} \cdot \delta\boldsymbol{\theta} = 0. \end{aligned}$$

Informační matice

Jestliže ξ je návrh (pravděpodobnostní míra na množině experimentálních bodů $\mathbf{x}$) potom informační matici definujeme jako

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

kde

$$\begin{array}{rcl} \mathrm{Sp}(\xi) = \{x_i : \xi(x_i) > 0, x_i \in \mathbf{x}\} & \ldots & \text{je nosič návrhu,} \\ \sigma^2 & \ldots & \text{je chyba měření} \\ N & \ldots & \text{je celkový počet měření.} \end{array}$$

V případě regulárnosti matice $\mathbf{M} = \mathbf{M}(\theta)$ je varianční matice NLNO $\widehat{\boldsymbol{\theta}}$

$$\mathrm{Var}(\widehat{\boldsymbol{\theta}}) = \mathbf{M}^{-1} - \mathbf{M}^{-1}\mathbf{B}^T \left[\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T \right]^{-1} \mathbf{B}\mathbf{M}^{-1}.$$

Konvexní kriteriální funkce

Protože kriteriální funkci Φ uvažujeme jako konvexní, musí lokálně optimální návrh ξ^* splňovat 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.

$$\Downarrow$$

Definování kriteriální funkce jako funkce normované varianční matice

$$\mathbf{V}_{\boldsymbol{\theta}}(\xi) = N \cdot \sigma^{-2} \cdot \mathrm{Var}(\widehat{\boldsymbol{\theta}}).$$

Kritérium lokální D-optimality

V modelu s podmínkami nelze kritérium lokální D-optimality definovat vztahem

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

protože matice $\mathbf{V}_{\boldsymbol{\theta}}(\xi)$ je často singulární, i když je $\boldsymbol{\theta}$ odhadnutelný.

$$\Downarrow$$

Matice $\mathbf{B}_{(q\times k)} = (\mathbf{B}_1, \mathbf{B}_2)$ je rozdělena do bloků tak, aby platilo

$$r(\mathbf{B}_1) = q \qquad \text{resp.} \qquad r(\mathbf{B}_2) = q.$$

Potom lze podmínku typu I psát je tvaru

$$\mathbf{B}\delta\boldsymbol{\theta} = \mathbf{0} \qquad \rightarrow \qquad \mathbf{B}_1\delta\boldsymbol{\theta}_1 + \mathbf{B}_2\delta\boldsymbol{\theta}_2 = \mathbf{0}.$$

Následně převedeme původní model s podmínkami typu I na model bez podmínek.

Kritérium lokální D-optimality

$$\delta\boldsymbol{\theta}_1 = -\mathbf{B}_1^{-1}\mathbf{B}_2\delta\boldsymbol{\theta}_2 \qquad\qquad \delta\boldsymbol{\theta}_2 = -\mathbf{B}_2^{-1}\mathbf{B}_1\delta\boldsymbol{\theta}_1$$

$$\downarrow \qquad\qquad\qquad\qquad \downarrow$$

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

Díky regulárnosti $\mathbf{M}^*$ definujeme kritérium lokální D-optimality jako

$$\begin{aligned} \Phi\left[\mathbf{M}^*(\boldsymbol{\theta}_2,\xi)\right] &= -\ln\left[\det(\mathbf{M}^*(\boldsymbol{\theta}_2,\xi))\right], \\ \text{resp.} \quad \Phi\left[\mathbf{M}^*(\boldsymbol{\theta}_1,\xi)\right] &= -\ln\left[\det(\mathbf{M}^*(\boldsymbol{\theta}_1,\xi))\right]. \end{aligned}$$

Příklad

Uvažujme lineární regresní model

$$\begin{aligned} Y_i &= \theta_1 + \theta_2 \cdot x_i + \epsilon_i, && \text{pro } x_i = i = 1, \ldots, 10 \\ Y_i &= \theta_3 + \theta_4 \cdot x_i + \theta_5 \cdot x_i^2 + \epsilon_i, && \text{pro } x_i = i = 11, \ldots, 20 \end{aligned}$$

s podmínkou typu I

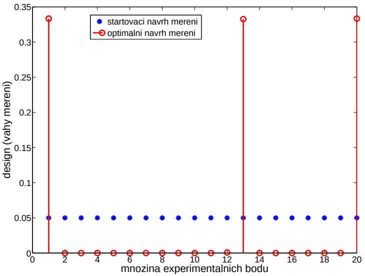
$$\begin{aligned} \theta_1 + \theta_2 \cdot x_{10} - (\theta_3 + \theta_4 \cdot x_{10} + \theta_5 \cdot x_{10}^2) &= 0, \\ \theta_2 - \theta_4 - 2 \cdot \theta_5 \cdot x_{10} &= 0. \end{aligned}$$

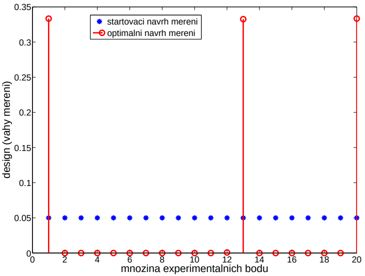
Potom

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

Příklad

$$\mathbf{B}_1 = \begin{bmatrix} 1 & 10 \\ 0 & 1 \end{bmatrix}, \mathbf{B}_2 = \begin{bmatrix} -1 & -10 & -100 \\ 0 & -1 & -20 \end{bmatrix} \qquad \mathbf{B}_1 = \begin{bmatrix} 1 & 10 & -1 \\ 0 & 1 & 0 \end{bmatrix}, \mathbf{B}_2 = \begin{bmatrix} -10 & -100 \\ -1 & -20 \end{bmatrix}$$





Nosič lokálně D-optimálního návrhu:

$$\xi^*(x_1) = 0.3333 \qquad \xi^*(x_{13}) = 0.3333 \qquad \xi^*(x_{20}) = 0.3334$$

Literatura

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