ADEKVÁTNOSŤ LINEARIZÁCIÍ NELINEÁRNEJ REGRESIE

Klára Hornišová, umerhorn@savba.sk
Ústav merania SAV, Bratislava

Abstrakt. Pre všeobecnú linearizáciu nelineárneho regresného modelu definujeme kritériá jej použiteľnosti podobné kritériám použiteľnosti taylorovskej linearizácie v apriórne danom bode. Oblasti prípustnosti linearizácie sa dajú popísať pomocou zovšeobecnených mier vnútornej a parametrickej krivosti pôvodného modelu.

Kritériá prípustnosti taylorovskej linearizácie. Inferenciu v (regulárnom) nelineárnom regresnom modeli

$$y = \eta(\theta) + \varepsilon \; ; \; \varepsilon \sim N(0, \sigma^2 W), \tag{1}$$

kde $y \in \mathbb{R}^N$ sú merania, $\theta \in \Theta \subseteq \mathbb{R}^p$ je neznámy parameter, $\eta(.) : \Theta \to \mathbb{R}^N$ regresná funkcia, ε náhodné chyby, W známa kladne definitná matica a σ neznámy parameter, možno zjednodušiť, ak ho nahradíme lineárnym modelom. Jeho voľbu možno založiť na apriórnej informácii o skutočnej hodnote $\bar\theta$ neznámeho parametra. Ak sa napríklad vie, že $\bar\theta \in \mathcal{O}(\theta^0)$, kde $\mathcal{O}$ je okolie danej hodnoty $\theta^0 \in \Theta$, model (1) sa zvykne aproximovať svojou taylorovskou linearizáciou v bode θ^0

$$y = \eta(\theta^0) + \partial\eta(\theta^0).(\theta - \theta^0) + \varepsilon =: A\theta + a + \varepsilon =: \tilde\eta(\theta) + \varepsilon \; ; \; \varepsilon \sim N(0, \sigma^2 W). \tag{2}$$

V [3] (pre známe σ) a [2] (pre neznáme σ) sa pri predpoklade zanedbateľnosti tretích derivácií funkcie $\eta(.)$ na $\mathcal{O}$ navrhli kritériá použiteľnosti (2) vzhľadom na skreslenie, ktoré spôsobuje pri rôznych druhoch inferencie.
Namiesto (2) možno použiť všeobecnejšiu linearizáciu

$$y = A\beta(\theta) + a + \varepsilon =: \tilde\eta(\theta) + \varepsilon \; ; \; \varepsilon \sim N(0, \sigma^2 W), \tag{3}$$

kde $A \in \mathbb{R}^{N\times p}$, $h(A) = p$, $a \in \mathbb{R}^N$ a $\beta(.) : \Theta \to \mathbb{R}^p$ je regulárna reparametrizácia.

Ak je dané apriórne rozdelenie $\pi(.)$ na Θ, špeciálne netaylorovské linearizácie, ktoré majú dobrú štatistickú interpretáciu - minimalizujú apriórnu strednú kvadratickú chybu - sú napr. linearizácia vyhladzovaním zo [4]

$$A = Cov_\pi(\eta, \theta) Var_\pi^{-1}\theta, a = E_\pi\eta - AE_\pi\theta, \beta(\theta) = \theta \tag{4}$$

alebo vnútorná linearizácia z [1]

$$A = (u_1, \dots, u_p), a = E_\pi\eta, \beta(\theta) = (A^\top W^{-1} A)^{-1} A^\top W^{-1}(\eta(\theta) - a), \tag{5}$$

kde $u_1, \dots, u_N$ sú ortonormálne vlastné vektory matice $Var_\pi\eta$ porade prislúchajúce jej vlastným hodnotám $\lambda_1 \geq \dots \geq \lambda_N \geq 0$.

Predpokladajme, že $\exists \mathcal{A}_0 \in \mathbb{R}^N$ a $W-$ortogonálny projektor $R_{N\times N}$, $h(R) = p + q < N$, také, že

$$\sup_{\theta\in\Theta} \|(I - R)(\eta(\theta) - \mathcal{A}_0)\|_W \tag{6.I}$$

je zanedbateľné, pričom q je čo najmenšie. Význam majú potom zrejme iba tie linearizácie (3), pre ktoré $a = \mathcal{A}_0$ a $\mathcal{M}(P) \subseteq \mathcal{M}(R)$, kde $P := P_A := A(A^\top W^{-1}A)^{-1}A^\top W^{-1}$. Ďalej budeme namiesto $\mathcal{A}_0$ písať a. Označme $Q := R - P$. Zrejme $h(Q) = q$.

Špeciálne možno predpokladať, že $\exists B \in \mathbb{R}^{N\times m^2}, \mathcal{A}(B) \in \mathbb{R}^{N\times m}, \mathcal{A}_0(B) \in \mathbb{R}^N$ také, že

$$\sup_{\theta\in\Theta} \|\eta(\theta) - (B\beta^{\otimes 2}(\theta) + \mathcal{A}(B)\beta(\theta) + \mathcal{A}_0(B))\|_W \tag{6.II}$$

je zanedbateľné. V tomto prípade nech $\mathcal{A}(\bar 0) = A$, $\mathcal{A}_0(\bar 0) = a$, $P := P_A$, $Q := P_{(I-P)B}$, $R := P + Q$. Modelom (6.II) môže byť okrem taylorovskej kvadratizácie v bode θ^0 napr. kvadratizácia $B = (Cov_\pi(\eta, \theta^{\otimes 2}) - ACov_\pi(\theta, \theta^{\otimes 2}))[Var_\pi\theta^{\otimes 2} - Cov_\pi(\theta^{\otimes 2}, \theta)Var_\pi^+\theta Cov_\pi(\theta, \theta^{\otimes 2})]^+, \mathcal{A} = A - BCov_\pi(\theta^{\otimes 2}, \theta), \mathcal{A}_0 = E_\pi\eta - BE_\pi\theta^{\otimes 2} - \mathcal{A}E_\pi\theta, \beta(\theta) = \theta$, kde $A := Cov_\pi(\eta, \theta)Var_\pi^+\theta$, minimalizujúca apriórnu strednú kvadratickú chybu.
Rozšírením prístupu z [3] a [2] máme:

Definícia 1 . Model (1) sa dá $(\alpha, \delta)-$linearizovať modelom (3) vzhľadom na vystihnutie meraní modelom v množine $\mathcal{O} \subseteq \Theta$, ak

$$\sup_{\bar\theta\in\mathcal{O}} P_{\bar\theta}(T \leq F_{q,N-p-q}(1 - \alpha)) \geq 1 - \alpha - \delta,$$

pre test vnútornej linearity modelu (1) s hladinou významnosti α

$$T := \frac{\|Q(y - a)\|_W^2}{qs^2} \sim F_{q,N-p-q}\left(\frac{\|Q(\eta(\theta) - a)\|_W^2}{\sigma^2}\right) \tag{7.I}$$

alebo

$$T = \frac{\|Q(y - a)\|_W^2}{qs^2} \sim F_{q,N-p-q}\left(\frac{1}{4}\frac{\|QB\beta^{\otimes 2}(\theta)\|_W^2}{\sigma^2}\right), \tag{7.II}$$

kde $s^2 := \frac{\|(I-R)(y-a)\|_W^2}{N-p-q}$.

Veta 1 . *Množina*

$$\mathcal{O} := \{\theta \in \Theta; \frac{\|P(\eta(\theta) - a)\|_W^2}{s^2} \leq \frac{2\sqrt{\gamma_{\max}}}{K_{\text{int,I}}(A)}\}, \tag{8.I}$$

alebo

$$\mathcal{O} := \{\theta \in \Theta; \frac{\|A\beta(\theta)\|_W^2}{s^2} \leq \frac{2\sqrt{\gamma_{\max}}}{K_{\text{int,II}}(A)}\} \tag{8.II}$$

je linearizačnou oblasťou modelu (1) *v zmysle definície 1, kde*

$$\gamma_{\max} := \max\{\gamma \geq 0; P\left(F_{q,N-p-q}(\gamma) \leq F_{q,N-p-q}(0; 1 - \alpha)\right) \geq 1 - \alpha - \delta\},$$

$$K_{\text{int,I}}(A) := \frac{s^2}{\sigma} \sup_{\substack{\theta\in\Theta \\ P(\eta(\theta)-a)\neq\bar 0}} \frac{\|Q(\eta(\theta) - a)\|_W}{\|P(\eta(\theta) - a)\|_W^2} = \frac{s^2}{\sigma} \sup_{\substack{\theta\in\Theta \\ \hat\beta(\theta)\neq\bar 0}} \frac{\|Q(\eta(\theta) - a)\|_W}{\|A\hat\beta(\theta)\|_W^2}, \tag{9.I}$$

kde $\hat\beta(\theta) := (A^\top W^{-1}A)^{-1}A^\top W^{-1}(\eta(\theta) - \mathcal{A}_0)$ je kanonická reparametrizácia,

$$K_{\text{int,II}}(A) := \frac{s^2}{\sigma} \sup_{\substack{\theta\in\Theta \\ \beta(\theta)\neq\bar 0}} \frac{\|(I - P)B\beta^{\otimes 2}(\theta)\|_W}{\|A\beta(\theta)\|_W^2} \tag{9.II}$$

(vnútorná krivosť modelu (1) vzhľadom na linearizáciu $A\beta(\theta) + a$).

Označme $\hat\beta(y, A) = (A^\top W^{-1}A)^{-1}A^\top W^{-1}(y - a)$ ML-odhad parametra β v modeli (3). Prípustnosť odhadu $h^\top\hat\beta(y, A)$ funkcie $h^\top\beta$ v modeli (1) a v množine $\mathcal{O}$ možno opäť posudzovať podľa skreslení pri rôznych druhoch inferencie. Označme $C := (A^\top W^{-1}A)^{-1}A^\top W^{-1}$.

Definícia 2 . Model (1) sa na množine $\mathcal{O}$ volá
c_b-linearizovateľný modelom (3) vzhľadom na výchylku funkcionálu h, ak

$$\sup_{\bar\theta\in\mathcal{O}} \frac{(E_{\bar\theta}[h^\top\hat\beta(y, A) - h^\top\bar\beta])^2}{Var[h^\top\hat\beta(y, A)]} = \sup_{\bar\theta\in\mathcal{O}} \frac{(h^\top[C(\eta(\bar\theta) - a) - \beta(\bar\theta)])^2}{\sigma^2 h^\top CWC^\top h} \leq c_b^2,$$

alebo

$$\int_{\mathcal{O}} \frac{(E_{\bar\theta}[h^\top\hat\beta(y, A) - h^\top\bar\beta])^2}{Var[h^\top\hat\beta(y, A)]} d\pi(\bar\theta) = \frac{\int_{\mathcal{O}}(h^\top[C(\eta(\bar\theta) - a) - \beta(\bar\theta)])^2 d\pi(\bar\theta)}{\sigma^2 h^\top CWC^\top h} \leq c_b^2,$$

kde $\mathcal{C} := \left\{C \in \mathbb{R}^{p\times N}; h(C) = p; \{C(\eta(\theta) - a); \theta \in \mathcal{O}\} \cap \beta(\mathcal{O}) \neq \emptyset\right\}$ alebo $\mathcal{C} := \left\{C \in \mathbb{R}^{p\times N}; h(C) = p; \frac{\pi(\{C(\eta(\theta)-a);\theta\in\mathcal{O}\}\cap\beta(\mathcal{O}))}{\pi(\mathcal{O})} \geq 1 - \varepsilon_0\right\}$.
Minimalizáciou tohoto kritéria pre $C \in \mathcal{C}$ možno nájsť najlepšiu c_b-linearizáciu modelu (1) na $\mathcal{O} \subseteq \Theta$.

Veta 2 . *Nech pre množinu $\mathcal{O}$ platí* $\sup_{\theta\in\mathcal{O}} \frac{(h^\top[C(\eta(\theta)-a)-\beta(\theta)])^2}{s^2 h^\top CWC^\top h} \leq c_b^2$ *alebo* $\sup_{\theta\in\mathcal{O}} \frac{(h^\top CB\beta^{\otimes 2}(\theta))^2}{s^2 h^\top CWC^\top h} \leq c_b^2$. *Potom je model* (1) *vzhľadom na funkcionál h na množine $\mathcal{O}$ c_b-linearizovateľný vzhľadom na výchylku.*

Možno nájsť aj podmnožiny parametrického priestoru, na ktorých je taylorovská linearizácia prípustná pre všetky lineárne funkcionály h.

Veta 3 . *Nech* $\mathcal{O} := \{\theta; \frac{\|A\beta(\theta)\|_W^2}{s^2} \leq M^2\}$. *Potom je model* (1) *pre všetky funkcionály $h^\top\theta$ c_b-linearizovateľný vzhľadom na výchylku, kde* $c_b = \frac{1}{2}M^2 K_{\text{par}}(A)$, *kde*

$$K_{\text{par,II}}(A) := \frac{s^2}{\sigma} \sup_{\substack{\theta\in\Theta \\ \beta(\theta)\neq\bar 0}} \frac{\|PB\beta^{\otimes 2}(\theta)\|_W}{\|A\beta(\theta)\|_W^2}.$$

Apriórna informácia daná množinou $\mathcal{O}$ by však mala byť menšia ako informácia z pozorovaní daná oblasťou spoľahlivosti na hladine $1 - \alpha$

$$\mathcal{S} = \{\theta; \frac{\|A(\hat\beta(y, A) - \beta(\theta))\|_W^2}{s^2} \leq (\frac{1}{2}M^2 K_{\text{par}}(A) + \sqrt{pF_{p,N-p-q}(1 - \alpha)})^2\}.$$

Podmienka $\mathcal{O} \subseteq \mathcal{S}$ je teda ekvivalentná s nerovnosťou

$$\frac{1}{2}M^2 K_{\text{par}}(A) + \sqrt{pF_{p,N-p-q}(1 - \alpha)} \leq M.$$

Veta 4 . *Nech*

$$K_{\text{par}}(A) = \frac{\omega^2}{2\sqrt{pF_{p,N-p-q}(1 - \alpha)}}, \omega^2 \leq 1$$

a

$$1 - \sqrt{1 - \omega^2} \leq MK_{\text{par}}(A) \leq 1 + \sqrt{1 - \omega^2}, \text{ ak } \omega^2 > 0,$$
$$\sqrt{pF_{p,N-p-q}(1 - \alpha)} \leq M, \text{ inak.}$$

Potom $\mathcal{O} \subseteq \mathcal{S}$.

Porovnaním podmienok pre M z vety 4 s podmienkami z viet 1 a 3 dostávame v prípade $K_{\text{par}}(A) > 0$

$$\frac{\omega^2\sqrt{\gamma_{\max}}}{(1 + \sqrt{1 - \omega^2})^2} K_{\text{par}}(A) \leq K_{\text{int}}(A)\sqrt{pF_{p,N-p-q}(1 - \alpha)} \leq \frac{\omega^2\sqrt{\gamma_{\max}}}{(1 - \sqrt{1 - \omega^2})^2} K_{\text{par}}(A)$$

a

$$\frac{1}{\omega^2}(1 - \sqrt{1 - \omega^2})^2 \leq \frac{c_b}{\sqrt{pF_{p,N-p-q}(1 - \alpha)}} \leq \frac{1}{\omega^2}(1 + \sqrt{1 - \omega^2})^2$$

a v prípade $K_{\text{par}}(A) = 0$

$$K_{\text{int}}(A) \leq \frac{2\sqrt{\gamma_{\max}}}{pF_{p,N-p-q}(1 - \alpha)}.$$

Poďakovanie. Na výskum prispela agentúra Vega grantmi č. 2/4026/04 a 1/3016/06.

Literatúra

[1] Hornišová K. (2004). *Intrinsic linearization of nonlinear regression by principal components method* AMUC **LXXIII**, 207–216.
[2] Jenčová A. (2000). *Linearization conditions for regression models with unknown variance parameter* Appl. Math. **45**, 145–160.
[3] Kubáček L. (1995). *On a linearization of regression models* Appl. Math. **40**, 61–78.
[4] Pázman A. (2001). *Linearization of regression models by smoothing* Tatra Mt. Math. Publ. **22**, 13–25.