Restringovaná (reziduální) metoda maximální věrohodnosti - REML

Jednou z vlastností maximálně věrohodného odhadu varianční složky je, že nebere v úvahu stupně volnosti potřebné při odhadu pevného efektu. Výsledný odhad varianční složky pak není nestranný (až na speciální případy).

Tento problém je odstraněn při použití restringované (reziduální) metody maximální věrohodnosti, jejíž základní myšlenkou je odhad varianční složky založený na reziduích získaných při odhadu pevných efektů metodou nejmenších čtverců.

Věrohodnostní funkce

Uvažujme n rozměrný normálně rozložený vektor y se střední hodnotou $X\beta$ a kovarianční maticí $V(\theta)$, tedy $y \sim N_n(X\beta, V(\theta))$. Předpokládejme, že X je $n \times p$ rozměrná matice s hodností p a V je $n \times n$ rozměrná pozitivně definitní matice.

REML je založeno na lineární kombinaci $A'y$, která neobsahuje žádné pevné efekty, tj. $A'X = 0$, kde A je $n \times (n-p)$ rozměrná matice. Přičemž každá lineární funkce $L'y$ taková, že $L'X = 0$, může být vyjádřena jako funkce $A'y$. Tedy $A'y \sim N_{(n-p)}(0, A'VA)$ a

$$L_{REML} = (2\pi)^{-(n-p)/2} |A'VA|^{-1/2} \, \mathrm{e}^{-\frac{1}{2} y'A(A'VA)^{-1}A'y} \,. \tag{1}$$

Běžně je však REML věrohodnostní funkce uváděná ve tvaru

$$L_{REML} = (Konst.) |V|^{-1/2} |X'V^{-1}X|^{-1/2} \, \mathrm{e}^{-\frac{1}{2}(y-X\hat{\beta})'V^{-1}(y-X\hat{\beta})}, \tag{2}$$

kde $\hat{\beta} = (X'V^{-1}X)^{-1}X'V^{-1}y$. Dále bude ukázáno, že tyto zápisy jsou ekvivalentní.

Pomocná tvrzení a označení

1. Pokud M a N jsou $m \times m$ matice, pak $|MN| = |M||N|$.

2. Uvažujme regulární symetrickou matici

$$M = \begin{pmatrix} M_{11} & M_{12} \\ M_{12}' & M_{22} \end{pmatrix}.$$

Pak

$$|M| = |M_{22}||M_{11} - M_{12}M_{22}^{-1}M_{12}'|.$$

Důkaz. Platí

$$\begin{pmatrix} I & -M_{12}M_{22}^{-1} \\ 0 & I \end{pmatrix} \begin{pmatrix} M_{11} & M_{12} \\ M_{12}' & M_{22} \end{pmatrix} = \begin{pmatrix} M_{11} - M_{12}M_{22}^{-1}M_{12}' & 0 \\ M_{12}' & M_{22} \end{pmatrix},$$

$$\left| \begin{pmatrix} I & -M_{12}M_{22}^{-1} \\ 0 & I \end{pmatrix} \right| = 1 \text{ a } \left| \begin{pmatrix} M_{11} - M_{12}M_{22}^{-1}M_{12}' & 0 \\ M_{12}' & M_{22} \end{pmatrix} \right| = |M_{22}||M_{11} - M_{12}M_{22}^{-1}M_{12}'|,$$

čímž dostáváme tvrzení. □

3.
$$\frac{\partial ln|B(t)|}{\partial t} = \mathrm{Tr}\left[B^{-1}\frac{\partial B(t)}{\partial t}\right]$$

4.
$$\frac{\partial B(t)^{-1}}{\partial t} = -B^{-1}\frac{\partial B(t)}{\partial t}B^{-1}$$

Tvrzení 1. *Uvažujme matice V, X a A výše uvedených vlastností, pak*

$$V^{-1} = V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1} + A(A'VA)^{-1}A'.$$

Důkaz. Platí

$$\begin{aligned}(V^{-1}X,\, A)^{-1}V^{-1}(V^{-1}X,\, A)'^{-1} &= [(V^{-1}X,\, A)'V(V^{-1}X,\, A)]^{-1}\\ &= \begin{pmatrix} X'V^{-1}X & 0\\ 0 & A'VA\end{pmatrix}^{-1}\\ &= \begin{pmatrix} (X'V^{-1}X)^{-1} & 0\\ 0 & (A'VA)^{-1}\end{pmatrix}\end{aligned}$$

Dále

$$\begin{aligned}V^{-1} &= (V^{-1}X,\, A)[(V^{-1}X,\, A)'V(V^{-1}X,\, A)]^{-1}(V^{-1}X,\, A)'\\ &= (V^{-1}X,\, A)\begin{pmatrix} (X'V^{-1}X)^{-1} & 0\\ 0 & (A'VA)^{-1}\end{pmatrix}(V^{-1}X,\, A)'\\ &= V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1} + A(A'VA)^{-1}A'.\end{aligned}$$

□

Tedy

$$\begin{aligned}y'A(A'VA)^{-1}A'y &= y'[V^{-1} - V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1}]y\\ &= y'V^{-1}y - y'V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1}y\\ &= y'V^{-1}y - y'V^{-1}X(X'V^{-1}X)^{-1}(X'V^{-1}X)(X'V^{-1}X)^{-1}X'V^{-1}y\\ &= (y' - y'V^{-1}X(X'V^{-1}X)^{-1}X')V^{-1}(y - X(X'V^{-1}X)^{-1}X'V^{-1}y)\\ &= (y - X(X'V^{-1}X)^{-1}X'V^{-1}y)'V^{-1}(y - X(X'V^{-1}X)^{-1}X'V^{-1}y)\\ &= (y - X\hat{\beta})'V^{-1}(y - X\hat{\beta}), \qquad (3)\end{aligned}$$

čímž jsme ukázali ekvivalenci exponentů výrazů (1) a (2).

Označme dále $W = V^{-1} - V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1} = A(A'VA)^{-1}A'$ a derivaci $V_i = \partial V/\partial\theta_i$. Úpravami obdržíme

$$\begin{aligned}\frac{\partial \ln|A'VA|}{\partial\theta_i} &= \mathrm{Tr}\left[(A'VA)^{-1}\frac{\partial A'VA}{\partial\theta_i}\right] = \mathrm{Tr}\left[(A'VA)^{-1}A'V_iA\right] =\\ &= \mathrm{Tr}\left[A(A'VA)^{-1}A'V_i\right] = \mathrm{Tr}[WV_i] \qquad (4)\end{aligned}$$

a dále

$$\begin{aligned}
\frac{\partial[\ln|V|+\ln|X'V^{-1}X|}{\partial\theta_i} &= \mathrm{Tr}[V^{-1}V_i] - \mathrm{Tr}\left[(XV^{-1}X)^{-1}\frac{\partial(X'V^{-1}X)}{\partial\theta_i}\right] \\
&= \mathrm{Tr}[V^{-1}V_i] - \mathrm{Tr}\left[(X'V^{-1}X)^{-1}X'\frac{\partial V^{-1}}{\partial\theta_i}X\right] \\
&= \mathrm{Tr}[V^{-1}V_i] - \mathrm{Tr}\left[(X'V^{-1}X)^{-1}X'V^{-1}V_iV^{-1}X\right] \\
&= \mathrm{Tr}[V^{-1}V_i] - \mathrm{Tr}\left[V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1}V_i\right] \\
&= \mathrm{Tr}[(V^{-1} - V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1})V_i] \\
&= \mathrm{Tr}[WV_i]. \qquad (5)
\end{aligned}$$

Z výrazů (4) a (5) vyplývá

$$\begin{aligned}
\ln|A'VA| &= (Konst.) + \ln|V| + \ln|X'V^{-1}X| \\
|A'VA| &= (Konst.)|V||X'V^{-1}X|. \qquad (6)
\end{aligned}$$

Celkem jsme tedy pomocí (3) a (6) dokázali ekvivalenci mezi výrazy (1) a (2).

Restringované maximálně věrohodné odhady

Uvažujme dále náhodný vektor $y \sim N(X\beta, V)$ ve tvaru $y = X\beta + \sum_{i=1}^{s+1} Z_i b_i$, kde $\mathrm{cov}(y) = \sum_{i=1}^{s+1} \sigma_i^2 Z_i Z_i' := V$. Potom lineární kombinace $A'y$ bude tvaru $A'y = A'X\beta + \sum_{i=1}^{s+1} A'Z_i b_i$.

ML odhady

Jak je ukázáno například v [2], maximálně věrohodné rovnice jsou

$$\frac{\partial l}{\partial\beta} = X'V^{-1}y - X'V^{-1}X\beta = 0 \qquad (7)$$

$$\frac{\partial l}{\partial\sigma_i^2} = -\frac{1}{2}\,\mathrm{Tr}(V^{-1}Z_iZ_i') + \frac{1}{2}(y - X\beta)'V^{-1}Z_iZ_i'V^{-1}(y - X\beta) = 0, \qquad (8)$$

pro $i = 1, \ldots, s+1$.

REML odhady

Substitucí $y = A'y$, $X = A'X = 0$, $Z = A'Z$ a $V = A'VA$ dostaneme restringované maximálně věrohodné rovnice

$$\mathrm{Tr}[(A'VA)^{-1}A'Z_iZ_i'A] = y'A(A'VA)^{-1}A'Z_iZ_i'A(A'VA)^{-1}A'y, \qquad (9)$$

pro $i = 1, \ldots, s+1$.

Příklad

Uvažujme nejjednodušší případ $y \sim N(X\beta, V = \sigma^2 I)$. Tedy $i = 1$ a $Z_i = I$ a podívejme se, jak se liší ML a REML odhady σ^2.

ML odhady

Dosazením do rovnice (8) a použitím ML odhadu $\hat{\beta} = (X'V^{-1}X)^{-1}X'V^{-1}y$ obdržíme

$$\begin{aligned}
\mathrm{Tr}\left(\frac{1}{\sigma^2}I_{n\times n}\right) &= (y - X\hat{\beta})'\frac{1}{\sigma^2}I\frac{1}{\sigma^2}I(y - X\hat{\beta}),\\
\frac{n}{\sigma^2} &= \frac{(y - X\hat{\beta})'(y - X\hat{\beta})}{(\sigma^2)^2}\\
\hat{\sigma}^2 &= \frac{(y - X\hat{\beta})'(y - X\hat{\beta})}{n}.
\end{aligned}$$

REML odhady

Dosazením do rovnice (9) a využitím rovnosti (3) dostaneme

$$\begin{aligned}
\mathrm{Tr}[\frac{1}{\sigma^2}(A'A)^{-1}A'A] &= y'A\frac{1}{\sigma^2}(A'A)^{-1}A'\frac{1}{\sigma^2}A(A'A)^{-1}A'y,\\
\mathrm{Tr}[\frac{1}{\sigma^2}I_{(n-p)\times(n-p)}] &= \frac{y'A(A'A)^{-1}A'y}{(\sigma^2)^2},\\
\frac{n-p}{\sigma^2} &= \frac{(y - X\hat{\beta})'(y - X\hat{\beta})}{(\sigma^2)^2},\\
\hat{\sigma}^2 &= \frac{(y - X\hat{\beta})'(y - X\hat{\beta})}{n-p}.
\end{aligned}$$

Literatura

[1] LaMotte, L. R. *A Direct Derivation of the REML Likelihood Function.* Statistical Papers, 48, 2007, s. 321–327.

[2] Wimmer, G. *Bioštatistika.*