

ML ODHAD μ A Σ

ĎALŠIE POMOCNÉ TVRDENIA

Lema 1*. Platí

$$\frac{\partial \mathbf{m}'\mathbf{x}}{\partial \mathbf{x}} = \mathbf{m},$$

$$\frac{\partial \mathbf{x}'\mathbf{A}\mathbf{x}}{\partial \mathbf{x}} = 2\mathbf{A}\mathbf{x}.$$

Dôkaz. Urobte ako cvičenie.

Nech \mathbf{B} je regulárna $n \times n$ matica, ktorej prvky sú diferencovateľnými funkciami premennej t , čiže $\{\mathbf{B}\}_{i,j} = b_{ij} = b_{ij}(t)$, $i, j = 1, 2, \dots, n$,

$\frac{\partial \mathbf{B}}{\partial t}$ je $n \times n$ matica, ktorej prvky sú $\frac{\partial b_{ij}(t)}{\partial t}$, $i, j = 1, 2, \dots, n$

$\frac{\partial \det \mathbf{B}}{\partial \mathbf{B}}$ je $n \times n$ matica, ktorej prvky sú $\frac{\partial \det \mathbf{B}}{\partial b_{ij}}$, $i, j = 1, 2, \dots, n$,

$$\text{diag} \mathbf{B} = \begin{pmatrix} \{\mathbf{B}\}_{1,1} & 0 & \dots & 0 \\ 0 & \{\mathbf{B}\}_{2,2} & \dots & 0 \\ \vdots & & & \\ 0 & & & \{\mathbf{B}\}_{n,n} \end{pmatrix}.$$

Značenie je rovnaké ako v kapitole 8. "Pomocné tvrdenia" textu "Plánovanie regresného experimentu".

Lema 2*. Pre symetrickú regulárnu $n \times n$ maticu \mathbf{B} a symetrickú maticu \mathbf{C} platí

$$\frac{\partial \text{tr} \mathbf{B}^{-1} \mathbf{C}}{\partial \mathbf{B}} = -2\mathbf{B}^{-1} \mathbf{C} \mathbf{B}^{-1} + \text{diag}(\mathbf{B}^{-1} \mathbf{C} \mathbf{B}^{-1}).$$

Dôkaz. Platí

$$\begin{aligned} \frac{\partial \text{tr} \mathbf{B}^{-1} \mathbf{C}}{\partial b_{11}} &= -\text{tr} \mathbf{C} \mathbf{B}^{-1} \frac{\partial \mathbf{B}}{\partial b_{11}} \mathbf{B}^{-1} = \\ &= -\text{tr} \mathbf{C} \mathbf{B}^{-1} \begin{pmatrix} 1 & 0 & 0 & \dots \\ 0 & 0 & 0 & \\ 0 & 0 & 0 & \\ \vdots & & & \\ 0 & 0 & 0 & \end{pmatrix} \mathbf{B}^{-1} = -\text{tr} \begin{pmatrix} 1 & 0 & 0 & \dots \\ 0 & 0 & 0 & \\ 0 & 0 & 0 & \\ \vdots & & & \\ 0 & 0 & 0 & \end{pmatrix} \mathbf{B}^{-1} \mathbf{C} \mathbf{B}^{-1} = \\ &= -\{\mathbf{B}^{-1} \mathbf{C} \mathbf{B}^{-1}\}_{11}, \\ \frac{\partial \text{tr} \mathbf{B}^{-1} \mathbf{C}}{\partial b_{12}} &= -\text{tr} \mathbf{C} \mathbf{B}^{-1} \frac{\partial \mathbf{B}}{\partial b_{12}} \mathbf{B}^{-1} = \end{aligned}$$

$$\begin{aligned}
&= -\text{tr} \mathbf{C} \mathbf{B}^{-1} \begin{pmatrix} 0 & 1 & 0 & \dots \\ 1 & 0 & 0 & \\ 0 & 0 & 0 & \\ \vdots & & & \\ 0 & 0 & 0 & \end{pmatrix} \mathbf{B}^{-1} = -\text{tr} \begin{pmatrix} 0 & 1 & 0 & \dots \\ 1 & 0 & 0 & \\ 0 & 0 & 0 & \\ \vdots & & & \\ 0 & 0 & 0 & \end{pmatrix} \mathbf{B}^{-1} \mathbf{C} \mathbf{B}^{-1} = \\
&= -\{\mathbf{B}^{-1} \mathbf{C} \mathbf{B}^{-1}\}_{21} - \{\mathbf{B}^{-1} \mathbf{C} \mathbf{B}^{-1}\}_{21} = 2\{\mathbf{B}^{-1} \mathbf{C} \mathbf{B}^{-1}\}_{12}.
\end{aligned}$$

Úplne analogicky dostávame

$$\begin{aligned}
\frac{\partial \text{tr} \mathbf{B}^{-1} \mathbf{C}}{\partial b_{ii}} &= -\text{tr} \mathbf{C} \mathbf{B}^{-1} \frac{\partial \mathbf{B}}{\partial b_{ii}} \mathbf{B}^{-1} = \\
&= \{\mathbf{B}^{-1} \mathbf{C} \mathbf{B}^{-1}\}_{ii}
\end{aligned}$$

a pre $i \neq j$

$$\frac{\partial \text{tr} \mathbf{B}^{-1} \mathbf{C}}{\partial b_{ij}} = -\text{tr} \mathbf{C} \mathbf{B}^{-1} \frac{\partial \mathbf{B}}{\partial b_{ij}} \mathbf{B}^{-1} = -2\{\mathbf{B}^{-1} \mathbf{C} \mathbf{B}^{-1}\}_{ij},$$

teda

$$\frac{\partial \text{tr} \mathbf{B}^{-1} \mathbf{C}}{\partial \mathbf{C}} = -2\mathbf{B}^{-1} \mathbf{C} \mathbf{B}^{-1} + \text{diag}(\mathbf{B}^{-1} \mathbf{C} \mathbf{B}^{-1}). \quad \square$$

Lema 3*. Pre symetrickú $n \times n$ maticu \mathbf{B} platí

$$2\mathbf{B} - \text{diag} \mathbf{B} = \mathbf{0} \Leftrightarrow \mathbf{B} = \mathbf{0}.$$

Dôkaz. Spravte ako cvičenie.

Združenú funkciu hustoty rozdelenia náhodného výberu $\mathbf{X}_{n,p,1} = (\mathbf{X}'_1, \dots, \mathbf{X}'_n)'$ uvažovanú pri danom \mathbf{x} (realizácia $\mathbf{X} \in \mathcal{R}^p$) ako funkciu vektorového parametra $\boldsymbol{\theta} \in \mathcal{R}^q$ nazývame funkciou vierohodnosti

$$L(\mathbf{x}; \boldsymbol{\theta}) = \prod_{i=1}^n f(\mathbf{x}_i; \boldsymbol{\theta}),$$

resp. jej logaritmus, teda

$$l(\mathbf{x}, \boldsymbol{\theta}) = l(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n, \boldsymbol{\theta}) = \ln L(\mathbf{x}; \boldsymbol{\theta}) = \sum_{i=1}^n \ln f(\mathbf{x}_i; \boldsymbol{\theta}).$$

Vierohodnostnými rovnicami rozumieme systém

$$\sum_{i=1}^n \frac{\partial \ln f(\mathbf{x}_i, \boldsymbol{\theta})}{\partial \theta_k} = 0, \quad k = 1, 2, \dots, q.$$

Majme náhodný výber $\mathbf{X} = (\mathbf{X}'_1, \mathbf{X}'_2, \dots, \mathbf{X}'_n)'$, kde $\mathbf{X}_i \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\boldsymbol{\Sigma}$ je regulárna. Potom

$$L(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = |2\pi\boldsymbol{\Sigma}|^{-\frac{n}{2}} e^{-\frac{1}{2} \sum_{i=1}^n (\mathbf{x}_i - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x}_i - \boldsymbol{\mu})},$$

čiže

$$l(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \ln L(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = -\frac{n}{2} \ln |2\pi\boldsymbol{\Sigma}| - \frac{1}{2} \sum_{i=1}^n (\mathbf{x}_i - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x}_i - \boldsymbol{\mu}).$$

Platí

$$\begin{aligned} (\mathbf{x}_i - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x}_i - \boldsymbol{\mu}) &= (\mathbf{x}_i - \bar{\mathbf{x}} + \bar{\mathbf{x}} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x}_i - \bar{\mathbf{x}} + \bar{\mathbf{x}} - \boldsymbol{\mu}) = \\ &= (\mathbf{x}_i - \bar{\mathbf{x}})' \boldsymbol{\Sigma}^{-1} (\mathbf{x}_i - \bar{\mathbf{x}}) + (\bar{\mathbf{x}} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\bar{\mathbf{x}} - \boldsymbol{\mu}) + \\ &\quad + 2(\mathbf{x}_i - \bar{\mathbf{x}})' \boldsymbol{\Sigma}^{-1} (\bar{\mathbf{x}} - \boldsymbol{\mu}), \end{aligned}$$

čiže

$$\begin{aligned} \sum_{i=1}^n (\mathbf{x}_i - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x}_i - \boldsymbol{\mu}) &= \sum_{i=1}^n (\mathbf{x}_i - \bar{\mathbf{x}})' \boldsymbol{\Sigma}^{-1} (\mathbf{x}_i - \bar{\mathbf{x}}) + n(\bar{\mathbf{x}} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\bar{\mathbf{x}} - \boldsymbol{\mu}) + \\ + 2 \sum_{i=1}^n (\mathbf{x}_i - \bar{\mathbf{x}})' \boldsymbol{\Sigma}^{-1} (\bar{\mathbf{x}} - \boldsymbol{\mu}) &= \text{tr} \sum_{i=1}^n (\mathbf{x}_i - \bar{\mathbf{x}})' \boldsymbol{\Sigma}^{-1} (\mathbf{x}_i - \bar{\mathbf{x}}) + n(\bar{\mathbf{x}} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\bar{\mathbf{x}} - \boldsymbol{\mu}) = \\ &= \text{tr} \left[\boldsymbol{\Sigma}^{-1} \left\{ \sum_{i=1}^n (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})' \right\} \right] + n(\bar{\mathbf{x}} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\bar{\mathbf{x}} - \boldsymbol{\mu}) + \end{aligned}$$

$$(A) \quad = n \text{tr} \left\{ \boldsymbol{\Sigma}^{-1} \mathbf{S}^{(real)} \right\} + n(\bar{\mathbf{x}} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\bar{\mathbf{x}} - \boldsymbol{\mu}),$$

lebo

$$\begin{aligned} \mathbf{S}^{(real)} &= \frac{1}{n} \sum_{i=1}^n (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})' \quad \text{a} \quad 2 \sum_{i=1}^n (\mathbf{x}_i - \bar{\mathbf{x}})' \boldsymbol{\Sigma}^{-1} (\bar{\mathbf{x}} - \boldsymbol{\mu}) = \\ &= 2 \left\{ n \frac{1}{n} \sum_{i=1}^n \mathbf{x}_i' \boldsymbol{\Sigma}^{-1} (\bar{\mathbf{x}} - \boldsymbol{\mu}) - n \bar{\mathbf{x}}' \boldsymbol{\Sigma}^{-1} (\bar{\mathbf{x}} - \boldsymbol{\mu}) \right\} = 0. \end{aligned}$$

Dostávame

$$\begin{aligned} (B) \quad l(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) &= -\frac{n}{2} \ln |2\pi\boldsymbol{\Sigma}| - \frac{n}{2} \text{tr} \left\{ \boldsymbol{\Sigma}^{-1} \mathbf{S}^{(real)} \right\} - \frac{n}{2} \text{tr} \left\{ \boldsymbol{\Sigma}^{-1} (\bar{\mathbf{x}} - \boldsymbol{\mu})(\bar{\mathbf{x}} - \boldsymbol{\mu})' \right\} = \\ &= -\frac{np}{2} \ln 2\pi - \frac{n}{2} \ln (\det(\boldsymbol{\Sigma})) - \frac{n}{2} \text{tr} \left\{ \boldsymbol{\Sigma}^{-1} \left[\mathbf{S}^{(real)} - (\bar{\mathbf{x}} - \boldsymbol{\mu})(\bar{\mathbf{x}} - \boldsymbol{\mu})' \right] \right\}. \end{aligned}$$

Teda vierohodnostné rovnice sú

$$\left. \frac{\partial l}{\partial \boldsymbol{\mu}} \right|_{\boldsymbol{\mu}=\hat{\boldsymbol{\mu}}^{(real)}, \boldsymbol{\Sigma}=\hat{\boldsymbol{\Sigma}}^{(real)}} = \mathbf{0},$$

$$\left. \frac{\partial l}{\partial \boldsymbol{\Sigma}} \right|_{\boldsymbol{\mu}=\hat{\boldsymbol{\mu}}^{(real)}, \boldsymbol{\Sigma}=\hat{\boldsymbol{\Sigma}}^{(real)}} = \mathbf{0}.$$

Ak $n \geq p + 1$, pomocou Lemy 1* dostávame z prvého systému vierohodnostných rovníc

$$-2(\hat{\Sigma}^{(real)})^{-1}\bar{\mathbf{x}} + 2(\hat{\Sigma}^{(real)})^{-1}\hat{\boldsymbol{\mu}}^{(real)} = \mathbf{0},$$

čiže

$$\hat{\boldsymbol{\mu}}^{(real)} = \bar{\mathbf{x}},$$

teda

$$\hat{\boldsymbol{\mu}} = \bar{\mathbf{X}}.$$

Ďalej budeme pokračovať bez komplikovaného značenia a využijeme ostatné lemy. Dostávame z druhého systému vierohodnostných rovníc

$$-\frac{n}{2} \frac{\partial}{\partial \Sigma} \{ \ln(\det(\Sigma)) + \text{tr} [\Sigma^{-1}(\mathbf{S} + (\bar{\mathbf{x}} - \boldsymbol{\mu})(\bar{\mathbf{x}} - \boldsymbol{\mu})')] \} = \mathbf{0},$$

$$2 \{ \Sigma^{-1} - \Sigma^{-1}(\mathbf{S} + (\bar{\mathbf{x}} - \boldsymbol{\mu})(\bar{\mathbf{x}} - \boldsymbol{\mu})')\Sigma^{-1} \} - \\ - \text{diag} \{ \Sigma^{-1} - \Sigma^{-1}(\mathbf{S} + (\bar{\mathbf{x}} - \boldsymbol{\mu})(\bar{\mathbf{x}} - \boldsymbol{\mu})')\Sigma^{-1} \} = \mathbf{0},$$

čiže

$$\Sigma^{-1} - \Sigma^{-1}(\mathbf{S} + (\bar{\mathbf{x}} - \boldsymbol{\mu})(\bar{\mathbf{x}} - \boldsymbol{\mu})')\Sigma^{-1} = \mathbf{0}.$$

Výsledne

$$\hat{\Sigma} = \mathbf{S}.$$

Chceme ukázať, že pre každú realizáciu $\mathbf{x}_1, \dots, \mathbf{x}_n$ je

$$l(\mathbf{x}_1, \dots, \mathbf{x}_n; \bar{\mathbf{x}}, \mathbf{S}^{(real)}) - l(\mathbf{x}_1, \dots, \mathbf{x}_n; \hat{\boldsymbol{\mu}}, \hat{\Sigma}) \geq 0.$$

Budeme ešte potrebovať niekoľko pomocných tvrdení.

Lema 4*. Čísla $\lambda_1, \dots, \lambda_p$ sú korene rovnice $|\mathbf{S}^{(real)} - \lambda \Sigma| = 0$ práve vtady, ak sú koreňmi rovnice $|\Sigma^{-\frac{1}{2}} \mathbf{S}^{(real)} \Sigma^{-\frac{1}{2}} - \lambda \mathbf{I}| = 0$.

Dôkaz. vyplýva z implikácií

$$|\mathbf{S}^{(real)} - \lambda \Sigma| = 0 \iff |\Sigma^{\frac{1}{2}} (\Sigma^{-\frac{1}{2}} \mathbf{S}^{(real)} \Sigma^{-\frac{1}{2}} - \lambda \mathbf{I}) \Sigma^{\frac{1}{2}}| = 0$$

$$\iff |\Sigma^{\frac{1}{2}}| |\Sigma^{-\frac{1}{2}} \mathbf{S}^{(real)} \Sigma^{-\frac{1}{2}} - \lambda \mathbf{I}| |\Sigma^{\frac{1}{2}}| = 0 \iff |\Sigma^{-\frac{1}{2}} \mathbf{S}^{(real)} \Sigma^{-\frac{1}{2}} - \lambda \mathbf{I}| = 0,$$

prčom navyše platí (podľa Lemy 8.9 v texte "Plánovanie regresného experimentu"), že

$$\text{tr}(\Sigma^{-\frac{1}{2}} \mathbf{S}^{(real)} \Sigma^{-\frac{1}{2}}) = \sum_{i=1}^p \lambda_i, \quad |\Sigma^{-\frac{1}{2}} \mathbf{S}^{(real)} \Sigma^{-\frac{1}{2}}| = \lambda_1 \dots \lambda_p.$$

Lema 5*. Pre $x > 0$ je $\ln x + 1 - x \leq 0$.

Dôkaz. : Pre funkciu $f(x) = e^{x-1} - x$ platí, že $f(0) = \frac{1}{e}$, $f(1) = 0$. Pre $x > 0$ je minimum tejto funkcie v tom čísle x , pre ktoré $f'(x) = 0$. Teda $f'(x) = e^{x-1} - 1 = 0$, čiže $x = 1$. Pretože $f''(1) = e^{x-1}|_{x=1} = 1 > 0$, v bode $x = 1$ funkcia $f(x)$ nadobúda minimum. A minimálna hodnota funkcie je $f(1) = 0$. Pre $x > 0$ preto $e^{x-1} - x \geq 0$, $e^{x-1} \geq x$, $x - 1 \geq \ln x$ a konečne $\ln x + 1 - x \leq 0$.

Pomocou (A), (B), Lemy 4* a Lemy 5* dostávame pre každé $\boldsymbol{\mu} \in \mathcal{R}^p$ a $\boldsymbol{\Sigma}$ pozitívne definitnú symetrickú $p \times p$ maticu:

$$\begin{aligned}
& l(\mathbf{x}_1, \dots, \mathbf{x}_n; \bar{\mathbf{x}}, \mathbf{S}^{(real)}) - l(\mathbf{x}_1, \dots, \mathbf{x}_n; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \\
& = -\frac{n}{2} \ln(\det \mathbf{S}^{(real)}) - \frac{n}{2} \text{tr} \mathbf{S}^{(real)-1} \mathbf{S}^{(real)} - \frac{n}{2} (\bar{\mathbf{x}} - \bar{\mathbf{x}})' \mathbf{S}^{(real)-1} (\bar{\mathbf{x}} - \bar{\mathbf{x}}) + \\
& \quad + \frac{n}{2} \ln(\det \boldsymbol{\Sigma}) + \frac{n}{2} \text{tr} \boldsymbol{\Sigma}^{-1} \mathbf{S}^{(real)} + \frac{n}{2} (\bar{\mathbf{x}} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\bar{\mathbf{x}} - \boldsymbol{\mu}) \geq \\
& \quad \geq -\frac{n}{2} \ln \frac{\det \mathbf{S}^{(real)}}{\det \boldsymbol{\Sigma}} - \frac{np}{2} + \frac{n}{2} \text{tr} \boldsymbol{\Sigma}^{-1} \mathbf{S}^{(real)} = \\
& \quad = -\frac{n}{2} \ln \frac{\det \mathbf{S}^{(real)}}{\det \boldsymbol{\Sigma}^{\frac{1}{2}} \det \boldsymbol{\Sigma}^{\frac{1}{2}}} - \frac{np}{2} + \frac{n}{2} \text{tr} \boldsymbol{\Sigma}^{-\frac{1}{2}} \mathbf{S}^{(real)} \boldsymbol{\Sigma}^{-\frac{1}{2}} = \\
& \quad = -\frac{n}{2} \ln \det(\boldsymbol{\Sigma}^{-\frac{1}{2}} \mathbf{S}^{(real)} \boldsymbol{\Sigma}^{-\frac{1}{2}}) - \frac{np}{2} + \frac{n}{2} \text{tr} \boldsymbol{\Sigma}^{-\frac{1}{2}} \mathbf{S}^{(real)} \boldsymbol{\Sigma}^{-\frac{1}{2}} = \\
& \quad = -\frac{n}{2} \ln(\lambda_1 \dots \lambda_p) - \frac{np}{2} + \frac{n}{2} (\lambda_1 + \dots + \lambda_p) = \\
& \quad = -\frac{n}{2} \{ \ln \lambda_1 + \dots + \ln \lambda_p + 1 + \dots + 1 - \lambda_1 - \dots - \lambda_p \} = \\
& \quad = -\frac{n}{2} \{ (\ln \lambda_1 + 1 - \lambda_1) + \dots + (\ln \lambda_p + 1 - \lambda_p) \} \geq 0.
\end{aligned}$$