ML ODHAD $\boldsymbol{\mu}$ A $\boldsymbol{\Sigma}$

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 $\{\mathbf{B}\}_{i,j} = b_{ij} = b_{ij}(t),\ i,j = 1,2,...,n$,

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

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

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

Lema 2. *Platí*

$$\frac{\partial \mathbf{B}^{-1}}{\partial t} = -\mathbf{B}^{-1}\frac{\partial \mathbf{B}}{\partial t}\mathbf{B}^{-1}.$$

Dôkaz. Prvky matice $\mathbf{B}^{-1}$ označme $b_{ij}^{(-1)},\ i,j = 1,2,...,n$. Tiež sú diferencovateľnými funkciami premennej t, čiže $b_{ij}^{(-1)} = b_{ij}^{(-1)}(t), i,j = 1,2,...,n$. Pre $i,j = 1,2,...,n$ je

$$\{\mathbf{B}\mathbf{B}^{-1}\}_{i,j} = \sum_{k=1}^{n} b_{ik}(t) b_{kj}^{(-1)}(t) = \delta_{ij}$$

(δ_{ij} je tzv. Kroneckerovo delta, čiže $\delta_{ij} = 0$ pre $i \neq j$ a $\delta_{ij} = 1$ pre $i = j$.) Preto

$$0 = \frac{\partial}{\partial t}\{\mathbf{B}\mathbf{B}^{-1}\}_{i,j} = \sum_{k=1}^{n} \frac{\partial}{\partial t}\left[b_{ik}(t) b_{kj}^{(-1)}(t)\right] =$$

$$= \sum_{k=1}^{n} \frac{\partial b_{ik}(t)}{\partial t} b_{kj}^{(-1)}(t) + \sum_{k=1}^{n} b_{ik}(t)\frac{\partial b_{kj}^{(-1)}(t)}{\partial t},\quad i,j = 1,2,...,n,$$

čo v maticovom zápise je

$$\frac{\partial \mathbf{B}}{\partial t}\mathbf{B}^{-1} + \mathbf{B}\frac{\partial \mathbf{B}^{-1}}{\partial t} = \mathbf{0},$$

čiže

$$\frac{\partial \mathbf{B}^{-1}}{\partial t} = -\mathbf{B}^{-1}\frac{\partial \mathbf{B}}{\partial t}\mathbf{B}^{-1}. \qquad \square$$

Lema 3. *Nech* $\mathbf{C}$ *je* $n \times n$ *matica konštánt. Platí*

$$\frac{\partial Tr\mathbf{BC}}{\partial t} = Tr\mathbf{C}\frac{\partial \mathbf{B}}{\partial t}.$$

Dôkaz.

$$\frac{\partial Tr\mathbf{BC}}{\partial t} = \frac{\partial}{\partial t}\sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}(t)c_{ji} = \sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial b_{ij}(t)}{\partial t}c_{ji} = Tr\frac{\partial \mathbf{B}}{\partial t}\mathbf{C} = Tr\mathbf{C}\frac{\partial \mathbf{B}}{\partial t}. \qquad \square$$

Z predchádzajúcich dvoch liem priamo dostávame

Dôsledok 4. *Platí*

$$\frac{\partial Tr\mathbf{B}}{\partial t} = Tr\frac{\partial \mathbf{B}}{\partial t}.$$

Dôsledok 5. *Platí*

$$\frac{\partial Tr\mathbf{B}^{-1}(t)\mathbf{C}}{\partial t} = -Tr\mathbf{C}\mathbf{B}^{-1}\frac{\partial \mathbf{B}(t)}{\partial t}\mathbf{B}^{-1}.$$

Lema 6. *Platí*

$$\frac{\partial det\mathbf{B}}{\partial \mathbf{B}} = \begin{cases} (\det \mathbf{B})(\mathbf{B}^{-1})', & \textit{ak je } \mathbf{B} \textit{ nesymetrická} \\ (\det \mathbf{B})(2\mathbf{B}^{-1} - diag\mathbf{B}^{-1}), & \textit{ak je } \mathbf{B} \textit{ symetrická.} \end{cases}$$

Dôkaz. Determinant regulárnej $n \times n$ matice $\mathbf{B}$ sa dá písať ako

$$det\mathbf{B} = \{\mathbf{B}\}_{i,1}B_{i,1} + \{\mathbf{B}\}_{i,2}B_{i,2} + ... + \{\mathbf{B}\}_{i,n}B_{i,n}$$

pre $i \in \{1, 2, ..., n\}$, pričom $B_{s,t}$ je doplnok $(n-1)$–ho stupňa determinantu $det\mathbf{B}$ patriaci k prvku $\{\mathbf{B}\}_{s,t}$ (pozri napr. [Kořínek, V., Základy algebry, Nakladatelství ČSAV, Praha, 1953], str. 270). Preto

$$\frac{\partial det\mathbf{B}}{\partial \{\mathbf{B}\}_{i,j}} = \frac{\partial}{\partial \{\mathbf{B}\}_{i,j}}(\{\mathbf{B}\}_{i,1}B_{i,1} + \{\mathbf{B}\}_{i,2}B_{i,2} + ... + \{\mathbf{B}\}_{i,n}B_{i,n}).$$

Dostávame

$$\frac{\partial det\mathbf{B}}{\partial \{\mathbf{B}\}_{i,j}} = B_{i,j}$$

a

$$\frac{\partial det\mathbf{B}}{\partial\mathbf{B}} = \begin{pmatrix} B_{1,1} & B_{1,2} & \dots & B_{1,n} \\ B_{2,1} & B_{2,2} & \dots & B_{2,n} \\ \dots & & & \\ B_{n,1} & B_{n,2} & \dots & B_{n,n} \end{pmatrix} = (\det\mathbf{B})(\mathbf{B}^{-1})',$$

lebo

$$\mathbf{B}^{-1} = \begin{pmatrix} \frac{B_{1,1}}{\det\mathbf{B}} & \frac{B_{2,1}}{\det\mathbf{B}} & \dots & \frac{B_{n,1}}{\det\mathbf{B}} \\ \frac{B_{1,2}}{\det\mathbf{B}} & \frac{B_{2,2}}{\det\mathbf{B}} & \dots & \frac{B_{n,2}}{\det\mathbf{B}} \\ \vdots & & & \\ \frac{B_{1,n}}{\det\mathbf{B}} & \frac{B_{2,n}}{\det\mathbf{B}} & \dots & \frac{B_{n,n}}{\det\mathbf{B}} \end{pmatrix}$$

(pozri napr. [Kořínek, V., Základy algebry, Nakladatelství ČSAV, Praha, 1953], str. 320). Toto platí o nesymetrickej matici $\mathbf{B}$. V prípade, že $\mathbf{B}$ je symetrická, teda

$$\{\mathbf{B}\}_{r,s} = \{\mathbf{B}(b_{11}, b_{12}, ..., b_{1n}, b_{22}, b_{23}, \dots, b_{2n}, \dots, b_{n-1\ n-1}, b_{n-1\ n}, b_{nn})\}_{r,s} =$$

$$= \begin{cases} b_{rs}, & \text{ak } r \leqq s, \\ b_{sr}, & \text{ak } r > s. \end{cases}$$

Pre symetrickú maticu teda

$$\mathbf{B} = \begin{pmatrix} \{\mathbf{B}\}_{1,1} & \{\mathbf{B}\}_{1,2} & \dots & \{\mathbf{B}\}_{1,n} \\ \{\mathbf{B}\}_{2,1} & \{\mathbf{B}\}_{2,2} & \dots & \{\mathbf{B}\}_{2,n} \\ \vdots & & & \\ \{\mathbf{B}\}_{n,1} & \{\mathbf{B}\}_{n,2} & \dots & \{\mathbf{B}\}_{n,n} \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{12} & b_{22} & \dots & b_{2n} \\ \vdots & & & \\ b_{1n} & b_{2n} & \dots & b_{nn} \end{pmatrix}.$$

Preto

$$\frac{\partial det\mathbf{B}}{\partial b_{ii}} = \sum_{k=1}^{n}\sum_{l=1}^{n} \frac{\partial det\mathbf{B}}{\partial\{\mathbf{B}\}_{k,l}} \frac{\partial\{\mathbf{B}\}_{k,l}}{\partial b_{ii}} = \frac{\partial det\mathbf{B}}{\partial\{\mathbf{B}\}_{i,i}} \frac{\partial\{\mathbf{B}\}_{i,i}}{\partial b_{ii}} =$$

$$= \frac{\partial}{\partial\{\mathbf{B}\}_{i,i}} \left[\{\mathbf{B}\}_{i,1}B_{i,1} + \{\mathbf{B}\}_{i,2}B_{i,2} + ... + \{\mathbf{B}\}_{i,n}B_{i,n}\right].1 = B_{i,i}, \quad i = 1, 2, ..., n.$$

Pre $i < j$ je

$$\frac{\partial det\mathbf{B}}{\partial b_{ij}} = \sum_{k=1}^{n}\sum_{l=1}^{n} \frac{\partial det\mathbf{B}}{\partial\{\mathbf{B}\}_{k,l}} \frac{\partial\{\mathbf{B}\}_{k,l}}{\partial b_{ij}} = \frac{\partial det\mathbf{B}}{\partial\{\mathbf{B}\}_{i,j}} \frac{\partial\{\mathbf{B}\}_{i,j}}{\partial b_{ij}} + \frac{\partial det\mathbf{B}}{\partial\{\mathbf{B}\}_{j,i}} \frac{\partial\{\mathbf{B}\}_{j,i}}{\partial b_{ij}} =$$

$$= \frac{\partial}{\partial\{\mathbf{B}\}_{i,j}} \left[\{\mathbf{B}\}_{i,1}B_{i,1} + \{\mathbf{B}\}_{i,2}B_{i,2} + ... + \{\mathbf{B}\}_{i,n}B_{i,n}\right].1 +$$

$$+ \frac{\partial}{\partial\{\mathbf{B}\}_{j,i}} \left[\{\mathbf{B}\}_{j,1}B_{j,1} + \{\mathbf{B}\}_{j,2}B_{j,2} + ... + \{\mathbf{B}\}_{j,n}B_{j,n}\right].1 =$$

$$= B_{i,j} + B_{j,i}.$$

Úplne rovnako pre $i > j$ dostaneme

$$\frac{\partial det\mathbf{B}}{\partial b_{ij}} = B_{j,i} + B_{i,j},$$

čiže

$$\frac{\partial det\mathbf{B}}{\partial \mathbf{B}} = \begin{pmatrix} \frac{\partial det\mathbf{B}}{\partial b_{11}} & \frac{\partial det\mathbf{B}}{\partial b_{12}} & \dots & \frac{\partial det\mathbf{B}}{\partial b_{1n}} \\ \frac{\partial det\mathbf{B}}{\partial b_{12}} & \frac{\partial det\mathbf{B}}{\partial b_{22}} & \dots & \frac{\partial det\mathbf{B}}{\partial b_{2n}} \\ \vdots \\ \frac{\partial det\mathbf{B}}{\partial b_{1n}} & \frac{\partial det\mathbf{B}}{\partial b_{2n}} & \dots & \frac{\partial det\mathbf{B}}{\partial b_{nn}} \end{pmatrix} = (\det \mathbf{B})(2\mathbf{B}^{-1} - diag\mathbf{B}^{-1}). \qquad \square$$

Lema 7. *Pre symetrickú regulárnu $n \times n$ maticu $\mathbf{B}$ platí*

$$\frac{\partial lndet\mathbf{B}(t)}{\partial t} = Tr\mathbf{B}^{-1}\frac{\partial \mathbf{B}(t)}{\partial t}.$$

Dôkaz. Ak si uvedomíme, že $\mathbf{B}$ aj $\mathbf{B}^{-1}$ sú symetrické matice, teda pre $i > j$ platí $\{\mathbf{B}^{-1}\}_{i,j} = \{\mathbf{B}^{-1}\}_{j,i}$, $\left\{\frac{\partial \mathbf{B}}{\partial t}\right\}_{i,j} = \left\{\frac{\partial \mathbf{B}}{\partial t}\right\}_{j,i}$ a tvrdenie predchádzjúcej lemy, čiže

$$\frac{\partial det\mathbf{B}}{\partial b_{ij}} = \begin{cases} 2\{\mathbf{B}^{-1}\}_{i,j}det\mathbf{B}, & \text{ak } i < j, \\ \{\mathbf{B}^{-1}\}_{i,i}det\mathbf{B}, & \text{ak } i = j, \end{cases}$$

dostávame

$$\frac{\partial lndet\mathbf{B}(t)}{\partial t} = \frac{1}{det\mathbf{B}}\frac{\partial det\mathbf{B}}{\partial t} = \frac{1}{det\mathbf{B}}\sum_{i=1}^{n}\sum_{j=i}^{n}\frac{\partial det\mathbf{B}}{\partial b_{ij}}\frac{\partial b_{ij}}{\partial t} = Tr\mathbf{B}^{-1}\frac{\partial \mathbf{B}}{\partial t}. \qquad \square$$

Lema 8. *Pre symetrickú regulárnu $n \times n$ maticu $\mathbf{B}$ a symetrickú maticu $\mathbf{C}$ platí*

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

Dôkaz. Podľa Dôsledku 5 je

$$\frac{\partial Tr\mathbf{B}^{-1}\mathbf{C}}{\partial b_{11}} = -Tr\mathbf{C}\mathbf{B}^{-1}\frac{\partial \mathbf{B}}{\partial b_{11}}\mathbf{B}^{-1} =$$

$$= -Tr\mathbf{C}\mathbf{B}^{-1}\begin{pmatrix} 1 & 0 & 0 & \dots \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ \vdots \\ 0 & 0 & 0 \end{pmatrix}\mathbf{B}^{-1} = -Tr\begin{pmatrix} 1 & 0 & 0 & \dots \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ \vdots \\ 0 & 0 & 0 \end{pmatrix}\mathbf{B}^{-1}\mathbf{C}\mathbf{B}^{-1} =$$

$$= -\{\mathbf{B}^{-1}\mathbf{C}\mathbf{B}^{-1}\}_{11},$$

$$\frac{\partial Tr\mathbf{B}^{-1}\mathbf{C}}{\partial b_{12}} = -Tr\mathbf{C}\mathbf{B}^{-1}\frac{\partial \mathbf{B}}{\partial b_{12}}\mathbf{B}^{-1} =$$

$$= -Tr\mathbf{C}\mathbf{B}^{-1}\begin{pmatrix} 0 & 1 & 0 & \dots \\ 1 & 0 & 0 & \\ 0 & 0 & 0 & \\ \vdots & & & \\ 0 & 0 & 0 & \end{pmatrix}\mathbf{B}^{-1} = -Tr\begin{pmatrix} 0 & 1 & 0 & \dots \\ 1 & 0 & 0 & \\ 0 & 0 & 0 & \\ \vdots & & & \\ 0 & 0 & 0 & \end{pmatrix}\mathbf{B}^{-1}\mathbf{C}\mathbf{B}^{-1} =$$

$$= -\{\mathbf{B}^{-1}\mathbf{C}\mathbf{B}^{-1}\}_{21} - \{\mathbf{B}^{-1}\mathbf{C}\mathbf{B}^{-1}\}_{21} = 2\{\mathbf{B}^{-1}\mathbf{C}\mathbf{B}^{-1}\}_{12}.$$

Úplne analogicky dostávame

$$\frac{\partial Tr\mathbf{B}^{-1}\mathbf{C}}{\partial b_{ii}} = -Tr\mathbf{C}\mathbf{B}^{-1}\frac{\partial \mathbf{B}}{\partial b_{ii}}\mathbf{B}^{-1} =$$

$$= \{\mathbf{B}^{-1}\mathbf{C}\mathbf{B}^{-1}\}_{ii}$$

a pre $i \neq j$

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

teda

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

Lema 9. *Pre symetrickú $n \times n$ maticu $\mathbf{B}$ platí*

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

Dôkaz. Spravte ako cvičenie.

V skriptičkách Multivariátna analýza 2 v 5. kapitole sme dostali, že logaritmus vierohodnostnej funkcie je

$$l(\mathbf{x};\boldsymbol{\mu},\boldsymbol{\Sigma}) = -\frac{n}{2}\ln|2\pi\boldsymbol{\Sigma}| - \frac{n}{2}Tr\left\{\boldsymbol{\Sigma}^{-1}\mathbf{S}^{(real)}\right\} - \frac{n}{2}Tr\left\{\boldsymbol{\Sigma}^{-1}(\overline{\mathbf{x}}-\boldsymbol{\mu})(\overline{\mathbf{x}}-\boldsymbol{\mu})'\right\} =$$

$$= -\frac{np}{2}\ln 2\pi - \frac{n}{2}\ln(det(\boldsymbol{\Sigma})) - \frac{n}{2}Tr\left\{\boldsymbol{\Sigma}^{-1}\left[\mathbf{S}^{(real)} - (\overline{\mathbf{x}}-\boldsymbol{\mu})(\overline{\mathbf{x}}-\boldsymbol{\mu})'\right]\right\}.$$

Teda vierohodnostné rovnice sú

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

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

Pomocou Lemy 1 dostávame z prvého systému vierohodnostných rovníc

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

čiže

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

teda

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

Ďalej budeme pokračovať bez komplikovaného značenia a využijeme Lemy 7,8 a 9. Dostávame z druhého systému vierohodnostných rovníc

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

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

čiže

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

Výsledne

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