Teoretické příklady (výběr/cv.5) 30.03.2021

1

$$\frac{\partial |X|}{\partial X} = (\mathrm{Adj}(X))^T \qquad X \text{ čtvercová, nesymetrická}$$

Nechť $X = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \Rightarrow |X| = ad - bc$

$$\frac{\partial |X|}{\partial X} = \frac{\partial}{\partial X}[ad-bc] = \begin{pmatrix} \frac{\partial}{\partial a}(ad-bc) & \frac{\partial}{\partial b}(ad-bc) \\ \frac{\partial}{\partial c}(ad-bc) & \frac{\partial}{\partial d}(ad-bc) \end{pmatrix}$$

$$= \begin{pmatrix} d & -c \\ -b & a \end{pmatrix} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}^T = (\mathrm{Adj}(X))^T$$

2

$$\frac{\partial |X|}{\partial X} = 2\,\mathrm{Adj}(X) - \mathrm{Diag}(\mathrm{Adj}\,X) \qquad X \text{ čtvercová, symetrická}$$

Nechť $X = \begin{pmatrix} a & b \\ b & c \end{pmatrix} \Rightarrow |X| = ac - b^2$

$$\frac{\partial |X|}{\partial X} = \frac{\partial}{\partial X}[ac-b^2] = \begin{pmatrix} \frac{\partial}{\partial a}(ac-b^2) & \frac{\partial}{\partial b}(ac-b^2) \\ \frac{\partial}{\partial b}(ac-b^2) & \frac{\partial}{\partial c}(ac-b^2) \end{pmatrix}$$

$$= \begin{pmatrix} c & -2b \\ -2b & a \end{pmatrix} = \begin{pmatrix} 2c & -2b \\ -2b & 2a \end{pmatrix} - \begin{pmatrix} c & 0 \\ 0 & a \end{pmatrix}$$

$$= 2\,\mathrm{Adj}(X) - \mathrm{Diag}(\mathrm{Adj}(X))$$

3

$$\frac{\partial\, tr(XA)}{\partial X} = \frac{\partial\, tr(AX)}{\partial X} = A^T$$

Recht: $X = \begin{pmatrix} x_1 & x_2 \\ x_3 & x_4 \end{pmatrix}$; $A = \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix}$

$$XA = \begin{pmatrix} a_1x_1 + a_3x_2 & a_2x_1 + a_4x_2 \\ a_1x_3 + a_3x_4 & a_2x_3 + a_4x_4 \end{pmatrix}$$

$$tr(XA) = a_1x_1 + a_2x_3 + a_3x_2 + a_4x_4$$

$$\frac{\partial\, tr(XA)}{\partial X} = \frac{\partial}{X}\Big[\underbrace{a_1x_1 + a_2x_3 + a_3x_2 + a_4x_4}_{(*)}\Big]$$

$$= \begin{pmatrix} \frac{\partial}{\partial x_1}(*) & \frac{\partial}{\partial x_2}(*) \\ \frac{\partial}{\partial x_3}(*) & \frac{\partial}{\partial x_4}(*) \end{pmatrix} = \begin{pmatrix} a_1 & a_3 \\ a_2 & a_4 \end{pmatrix}$$

$$= \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix}^T = A^T$$

5 $$\frac{\partial tr(X^{-1}A)}{\partial X} = \frac{\partial tr(AX^{-1})}{\partial X} = -X^{-1}A^{T}X^{-1}$$

Nach: $X = \begin{pmatrix} x_1 & x_2 \\ x_3 & x_4 \end{pmatrix}$; $A = \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix}$

$$X^{-1} = \frac{1}{det(X)} \cdot Adj(X) = \frac{1}{x_1x_4 - x_2x_3}\begin{pmatrix} x_4 & -x_2 \\ -x_3 & x_1 \end{pmatrix}$$

$$X^{-1}A = \begin{pmatrix} a_1x_4 - a_3x_2 & a_2x_4 - a_4x_2 \\ -a_1x_3 + a_3x_1 & -a_2x_3 + a_4x_1 \end{pmatrix} \cdot \frac{1}{x_1x_4 - x_2x_3}$$

$$tr(X^{-1}A) = \frac{a_1x_4 - a_3x_2 - a_2x_3 + a_4x_1}{x_1x_4 - x_2x_3} = (*)$$

$$\frac{\partial tr(X^{-1}A)}{\partial X} = \frac{\partial}{\partial X}\left[\underbrace{\frac{a_1x_4 - a_2x_3 - a_3x_2 + a_4x_1}{x_1x_4 - x_2x_3}}_{(*)}\right]$$

$$= \begin{pmatrix} \frac{\partial}{\partial x_1}(*) & \frac{\partial}{\partial x_2}(*) \\ \frac{\partial}{\partial x_3}(*) & \frac{\partial}{\partial x_4}(*) \end{pmatrix}$$

$$\frac{\partial}{\partial x_1}\left[\frac{a_1x_4 - a_2x_3 - a_3x_2 + a_4x_1}{x_1x_4 - x_2x_3}\right] =$$

$$= \frac{\boxed{a_4(x_1x_4 - x_2x_3) - (a_1x_4 - a_2x_3 - a_3x_2 + a_4x_1)x_4}}{(x_1x_4 - x_2x_3)^2}$$

$$\| $$

$$- a_1x^4 + a_2x_3x_4 + a_3x_2x_4 + a_4x_1x_4$$

$$\frac{\partial}{\partial x_2}\left[\frac{a_1x_4 - a_2x_3 - a_3x_2 + a_4x_1}{x_1x_4 - x_2x_3}\right] =$$

$$= \frac{-a_2(x_1x_4 - x_1x_3) + (a_1x_4 - a_2x_3 - a_3x_2 + a_4x_1)x_3}{(x_1x_4 - x_2x_3)^2}$$

$$\frac{\partial}{\partial x_3}\left[\frac{a_1x_4 - a_2x_3 - a_3x_2 + a_4x_1}{x_1x_4 - x_2x_3}\right] =$$

$$= \frac{-a_3(x_1x_4 - x_2x_3) + (a_1x_4 - a_2x_3 - a_3x_2 + a_4x_1)\cdot x_2}{(x_1x_4 - x_2x_3)^2}$$

$$\frac{\partial}{\partial x_4}\left[\frac{a_1x_4 - a_2x_3 - a_3x_2 + a_4x_1}{x_1x_4 - x_2x_3}\right] =$$

$$= \frac{a_1(x_1x_4 - x_2x_3) - (a_1x_4 - a_2x_3 - a_3x_2 + a_4x_1)\cdot x_1}{x_1x_4 - x_2x_3}$$

Spočteme teď: $\boxed{-X^{-1} \cdot A \cdot X^{-1}} = (\square)$

$$(\square) = \frac{(-1)}{(x_1x_4 - x_2x_3)^2} \cdot \begin{pmatrix} x_4 & -x_3 \\ -x_2 & x_1 \end{pmatrix} \cdot \begin{pmatrix} a_1 & a_3 \\ a_2 & a_4 \end{pmatrix} \cdot \begin{pmatrix} x_4 & -x_3 \\ -x_2 & x_1 \end{pmatrix}$$

$$= \frac{(-1)}{(x_1x_4 - x_2x_3)^2} \cdot \begin{pmatrix} a_1x_4 - a_2x_3 & a_3x_4 - a_4x_3 \\ -a_1x_2 + a_2x_1 & -a_3x_2 + a_4x_1 \end{pmatrix} \begin{pmatrix} x_4 & -x_3 \\ -x_2 & x_1 \end{pmatrix}$$

$$= \frac{(-1)}{(x_1x_4 - x_2x_3)^2} \begin{pmatrix} a_1x_4^2 - a_2x_3x_4 - a_3x_2x_4 + a_4x_2x_3 & \vdots \\ & \vdots \end{pmatrix}$$

$\longrightarrow$ ostatní prvky se ověří analogicky!

tudíž platí, že

$$\boxed{\frac{\partial \operatorname{tr}(X^{-1}A)}{\partial X} = -X^{-1}A^{T}X^{-1}}$$

Ostatní příklady (t.j. 4, 6, 7, 8) samostatně!

Maximum Likelihood Estimation

3) $f_\theta(x_1, x_2) = \frac{1}{\theta_1 \theta_2} \cdot \exp\left\{-\left(\frac{x_1}{\theta_1} + \frac{x_2}{\theta_2}\right)\right\}$ $\quad x_1 \geq 0$, $x_2 \geq 0$

$\theta_1, \theta_2 \in \mathbb{R}^+$; Data $\mathbb{X} = \{(x_{i1}, x_{i2}) ; i = 1, \dots n\}$

$$L(\mathbb{X}, \theta_1, \theta_2) = \prod_{i=1}^{n} f_\theta(x_{i1}, x_{i2}) =$$

$$= \prod_{i=1}^{n} \frac{1}{\theta_1 \theta_2} \cdot \exp\left(-\left(\frac{x_{i1}}{\theta_1} + \frac{x_{i2}}{\theta_2}\right)\right)$$

$$= (\theta_1 \theta_2)^{-n} \cdot \exp\left\{-\frac{\sum_{i=1}^{n} x_{i1}}{\theta_1} - \frac{\sum_{i=1}^{n} x_{i2}}{\theta_2}\right\}$$

$$\ell(\mathbb{X}, \theta_1, \theta_2) = \log L(\mathbb{X}, \theta_1, \theta_2) =$$

$$-n(\ln \theta_1 + \ln \theta_2) - \frac{\sum_{i=1}^{n} x_{i1}}{\theta_1} - \frac{\sum_{i=1}^{n} x_{i2}}{\theta_2}$$

$$\frac{\partial}{\partial \theta_1} \ell(\mathbb{X}, \theta_1, \theta_2) = -\frac{n}{\theta_1} + \frac{\sum_{i=1}^{n} x_{i1}}{\theta_1^2} := 0$$

$$\frac{\partial}{\partial \theta_2} \ell(\mathbb{X}, \theta_1, \theta_2) = -\frac{n}{\theta_2} + \frac{\sum_{i=1}^{n} x_{i2}}{\theta_2^2} := 0$$

$$\Longrightarrow \quad \hat{\theta}_1 = \frac{1}{n}\sum_{i=1}^{n} X_{i1}$$ (průměr prvých složek)

$$\hat{\theta}_2 = \frac{1}{n}\sum_{i=1}^{n} X_{i2}$$ (průměr druhých složek)

Také víme, že

$$\sqrt{n}(\hat{\theta}_1, \hat{\theta}_2) \overset{as}{\sim} N(\theta; I^{-1}(\theta))$$

Fisherova informace

(asymptotický rozptyl)

(zároveň R-C mez)

$$I_n(\theta) = -E\begin{pmatrix} \frac{\partial^2}{\partial\theta_1^2} \ell(X, \theta_1, \theta_2) & \frac{\partial^2}{\partial\theta_1 \partial\theta_2} \ell(X, \theta_1, \theta_2) \\ \frac{\partial^2}{\partial\theta_1 \partial\theta_2} \ell(X, \theta_1, \theta_2) & \frac{\partial^2}{\partial\theta_2^2} \ell(X, \theta_1, \theta_2) \end{pmatrix}$$

$$= -E\begin{bmatrix} \frac{n}{\theta_1^2} - \frac{2\sum_{i=1}^{n} X_{i1}}{\theta_1^3} & 0 \\ 0 & \frac{n}{\theta_2^2} - \frac{2\sum_{i=1}^{n} X_{i2}}{\theta_2^3} \end{bmatrix} = (\text{II})$$

Hustota $f_\theta(x_1, x_2)$ je faktorizovatelná $\Rightarrow$ X_1 a X_2 jsou nezávislé, zároveň platí, že $X_1 \sim Exp(\theta_1)$ a $X_2 \sim Exp(\theta_2)$

$\Longrightarrow$ $EX_1 = \theta_1$ a $EX_2 = \theta_2$!

$$(\square) = \begin{pmatrix} \frac{2n}{\theta_1^2} - \frac{n}{\theta_1^2} & \emptyset \\ \emptyset & \frac{2n}{\theta_2^2} - \frac{n}{\theta_2^2} \end{pmatrix} =$$

$$= \begin{pmatrix} \frac{n}{\theta_1^2} & \emptyset \\ \emptyset & \frac{n}{\theta_2^2} \end{pmatrix}$$

Ostatné príklady (t.j. 10, 11 a 12) samostatne!

Štatistické testy pomerom vierohodnosti:

13 Data $X_1, \dots X_n \sim N_p(\mu, \Sigma)$

$\mathbb{X} = \{X_1, \dots X_n\}$

$\left.\begin{matrix} H_0: \mu = \mu_0 \\ H_A: \mu \neq \mu_0 \end{matrix}\right\}$ p-rozmerné rozdelenie ($\mu, \mu_0 \in \mathbb{R}^p$)

Obecná vierohodnost pro $N_p(\mu, \Sigma)$:

$$L(\mathbb{X}, \mu, \Sigma) = |2\pi\Sigma|^{-\frac{n}{2}} \cdot \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}(X_i-\mu)\Sigma^{-1}(X_i-\mu)^T\right\}$$

$$\ell(\mathbb{X}, \mu, \Sigma) = -\frac{n}{2}\log(|2\pi\Sigma|) - \frac{1}{2}\sum_{i=1}^{n}(X_i-\mu)\Sigma^{-1}(X_i-\mu)^T$$

Za platnosti nulové hypotézy ($H_0: \mu = \mu_0$) dostáváme: (a chceme odhad pro Σ)

$$\ell_0(X, \Sigma) = -\frac{n}{2}\log(|2\pi\Sigma|) - \frac{1}{2}\sum_{i=1}^{n}(X_i - \mu_0)\Sigma^{-1}(X_i - \mu_0)^T$$

dosadíme, co známe z H_0!

$$= \frac{-np\log 2\pi}{2} - \frac{n}{2}\log|\Sigma| - \frac{1}{2}\sum_{i=1}^{n}(X_i - \mu_0)\Sigma^{-1}(X_i - \mu_0)^T$$

$$= \frac{-np\log 2\pi}{2} - \frac{n}{2}\log|\Sigma| - \frac{1}{2}\left\{tr\left(\Sigma^{-1}\sum_{i=1}^{n}(X_i - \mu_0)(X_i - \mu_0)^T\right)\right\}$$

Var-cov matice

suma přes „n"

Teď to musíme derivovat vzhledem k Σ a položit rovno nule (vektorová nula)!

Využijeme tohle:

$$\boxed{\frac{\partial \log|X|}{\partial X} = (X^T)^{-1}}$$

$$\boxed{\frac{\partial\, tr(X^T A)}{\partial X} = A}$$

použijeme pro $\boxed{X = \Sigma^{-1}}$

Teda:

$$\frac{\partial}{\partial(\Sigma^{-1})} \ell_0(X, \Sigma) = \frac{n}{2}\Sigma - \frac{1}{2}\sum_{i=1}^{n}(X_i - \mu_0)(X_i - \mu_0)^T$$

↳ řešení: $$\hat{\Sigma} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu_0)(X_i - \mu_0)^T$$

A zároveň platí následující:

$$\hat{\Sigma} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu_0)(X_i - \mu_0)^T =$$

$$= \frac{1}{n}\sum_{i=1}^{n}\Big\{(X_i - \bar{X})(X_i - \bar{X})^T + (\bar{X} - \mu_0)(\bar{X} - \mu_0)^T + 2(\bar{X} - \mu_0)(X_i - \bar{X})\Big\}$$

(*) průměr $v \{X_1, \dots X_n\} = X$

$$= \boxed{\frac{1}{n}\sum (X_i - \bar{X})(X_i - \bar{X})^T} + (\bar{X} - \mu_0)(\bar{X} - \mu_0)^T$$

S

↳ co sme chceli ukázat!

Ukážeme ještě, že LR statistika pro nulovou hypotézu $H_0\ \mu = \mu_0$ má následující tvar

$$-2\log\lambda = n\log\left(1 + (\bar{X} - \mu_0)^T S^{-1}(\bar{X} - \mu_0)\right)$$

→ potřebujeme porovnat věrohodnost za nulovej hypotézy (to už máme: $\ell_0(X, \hat{\Sigma})$) a věrohodnost za alternativy, t.j. $\ell(X, \hat{\mu}, \hat{\Sigma})$.

Takže potřebujeme ještě odhady $\hat{\mu}$ a $\hat{\Sigma}$

Jako samostatné cvičení ukázat, že

$$\hat{\mu} = \frac{1}{n}\sum X_i = \bar{X} \quad \text{(výběrový průměr)}$$

$$S = \frac{1}{n}\sum (X_i - \bar{X})(X_i - \bar{X})^T$$

(nemýlit s $\hat{\Sigma}$ za platnosti nulovej hypotézy $\hat{\Sigma}_0 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu_0)(X_i - \mu_0)^T$)

pro rozlišení $\hat{\Sigma}$ a $\hat{\Sigma}_0$

obecný

za H_0

Konečně dostáváme:

$$-2\lambda = 2(\ell(X, \hat{\mu}, S) - \ell_0(X, \hat{\Sigma}_0))$$

$$= 2(\ell(X, \hat{\mu}, S) - \ell_0(X, \underbrace{S + (\bar{X} - \mu_0)(\bar{X} - \mu_0)}_{\text{dosazeno z } (\blacktriangle)}))$$

$$= -n \log|S| + n \log\left|S + (\bar{X}-\mu_0)(\bar{X}-\mu_0)^T\right|$$

$$- \operatorname{tr}\left\{S^{-1} \sum_{i=1}^{n} (X_i - \bar{X})(X_i - \bar{X})^T\right.$$

$$+ \operatorname{tr}\left\{\left(S + (\bar{X}-\mu_0)(\bar{X}-\mu_0)^T\right)^{-1} \sum_{i=1}^{n} (X_i-\mu_0)(X_i-\mu_0)^T\right.$$

$$= n \log\left\{|S|\left(1 + (\bar{X}-\mu_0)^T S^{-1} (\bar{X}-\mu_0)\right)\right\} - n \log|S|$$

$$= n \log\left(1 + (\bar{X}-\mu_0)^T S^{-1} (\bar{X}-\mu_0)\right)$$

15 Pomocou programu R!