Zbierka úloh z viacrozmerných štatistických analýz

Lenka Filová

25. septembra 2019

1 Opakovanie maticovej algebry

Príklad 1. Dokážte, že každá symetrická idempotentná matica A je pozitívne semidefinitná matica, ktorej všetky vlastné čísla sú rovné buď 0 alebo 1, čiže je ju možné zapísať v tvare

$$A = U\mathrm{diag}(1, \ldots, 1, 0, \ldots, 0)U^T,$$

kde U je ortogonálna matica vlastných vektorov matice A. Dokážte tiež, že stopa symetrickej idempotentnej matice je rovná jej hodnosti.

Príklad 2. Nech A je matica typu $n \times m$ a nech $(A^T A)^-$ je pseudoinverzná matica k $A^T A$. Ukážte, že $(A^T A)^- A^T$ je jednou z pseudoinverzií matice A.

Príklad 3. Ortogonálny projektor na lineárny podpriestor L priestoru $\mathbb{R}^n$ je taká matica P typu $n \times n$, pre ktorú platí:
a) $\forall x \in \mathbb{R}^n: \; Px \in L$
b) $\forall x \in L: \; Px = x$
c) $\forall x, y \in \mathbb{R}^n: \; \langle x, Py \rangle = \langle Px, y \rangle$.
Presvedčte sa, že každý ortogonálny projektor je idempotentná symetrická matica. Dokážte, že existuje jediný ortogonálny projektor na daný lineárny podpriestor L.

Príklad 4. Nech $A_{p\times p}$ je symetrická matica a $a \in \mathbb{R}^p$. Ukážte, že potom

$$\frac{\partial a^T x}{\partial x} = a, \; \frac{\partial x^T A x}{\partial x} = 2Ax, \; \frac{\partial^2 a^T x}{\partial x \partial x^T} = 2A.$$

Príklad 5. Nech

$$A = \begin{pmatrix} 1 & 4 & 2 \\ 4 & 3 & 1 \\ 2 & 1 & 9 \end{pmatrix}.$$

Nájdite gradient a Hessián kvadratickej formy $Q(x) = x^T A x$.

Príklad 6. Nájdite vzorec pre $\det(A + aa^T)$ a $(A + aa^T)^{-1}$.

Príklad 7. Ukážte, že pre ľubovoľné matice $A_{n\times p}$, $B_{p\times n}$ majú súčiny AB a BA rovnaké nenulové vlastné hodnoty.

2 Náhodné vektory

Príklad 1. Nech X, Y sú nezávislé náhodné premenné s konečnou disperziou a nech $X_1 = X$, $X_2 = X + Y$ a $X_3 = X - Y$. Nájdite kovariančnú maticu náhodného vektora $Z = (X_1, X_2, X_3)^T$. Presvedčte sa, že náhodné vektory $(X_1, X_2)^T$, $(X_1, X_3)^T$ a $(X_2, X_3)^T$ majú regulárne kovariančné matice, ale náhodný vektor Z má singulárnu kovariančnú maticu.

Príklad 2. Dokážte, že korelačná matica ľubovoľného náhodného vektora, ktorého komponenty majú konečné a nenulové disperzie, je pozitívne semidefinitná.

Príklad 3. Nech $V = (X, Y, Z)^T$ je náhodný vektor s konečnými a nenulovými disperziami komponent a nech $\rho_{XY}, \rho_{XZ}, \rho_{YZ}$ sú korelačné koeficienty zložiek vektora V. Dokážte, že potom

$$\rho_{XY} + \rho_{XZ} + \rho_{YZ} \geq -3/2,$$

$$2\rho_{XY}\rho_{XZ}\rho_{YZ} + 1 \geq \rho_{XY}^2 + \rho_{XZ}^2 + \rho_{YZ}^2$$

Príklad 4. Náhodný vektor $(X, Y, Z)^T$ má strednú hodnotu $(1, 1, 1)$ a kovariančnú maticu

$$\begin{pmatrix} 2 & 0 & 1 \\ 0 & 3 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

Nájdite strednú hodnotu a kovariančnú maticu náhodného vektora $(X - Z, X - Y + Z)^T$.

Príklad 5. Nech X a Y sú n-rozmerné náhodné vektory s konečnými druhými momentmi. Ukážte, že

$$\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + \mathrm{Cov}(X, Y) + \mathrm{Cov}(Y, X)$$

Príklad 6. Nech X, Y, Z sú n-rozmerné náhodné vektory s konečnými druhými momentmi. Ukážte, že

$$\mathrm{Cov}(X + Y, Z) = \mathrm{Cov}(X, Z) + \mathrm{Cov}(Y, Z)$$

Príklad 7. Nech X a Y sú n-rozmerné náhodné vektory s konečnými druhými momentmi. Ďalej, nech $a \in \mathbb{R}^p$, $b \in \mathbb{R}^q$ sú reálne vektory, A je matica typu $p \times n$ a B je matica typu $q \times n$. Ukážte, že

$$\mathrm{Cov}(a + AX, b + BY) = A\mathrm{Cov}(X, Y)B^T.$$

Príklad 8. Nech X, Y, Z sú náhodné premenné a $a, b \in \mathbb{R}$. Ukážte, že

1. $E(a|Y) = a$,
2. $E(aX) + bZ|Y) = aE(X|Y) + bE(Z|Y)$,
3. Ak $X \geq 0$, potom $E(X|Y) \geq 0$,
4. Ak sú X a Y nezávislé, potom $E(X|Y) = E(X)$.

Príklad 9. Nech X, Y sú náhodné premenné. Nájdite $E(Y|X)$ pre nasledovné situácie:

1. $f(x, y) = x + y$ pre $x, y \in [0, 1]$,
2. $f(x, y) = 6x^2 y$ pre $x, y \in [0, 1]$.

Príklad 10. Nech náhodný vektor $(X, Y)^T$ má rovnomerné rozdelenie na obdĺžniku $S = [a, b] \times [c, d] \subseteq \mathbb{R}^2$. Nájdite $E(Y|X)$.
Príklad 11. Nech X, Y sú nezávislé náhodné premenné s rovnomerným rozdelením na $[0, 1]$ a $Z = X + Y$. Nájdite $E(Z|X)$ a $E(XZ|X)$.

3 Kopuly

Príklad 1. Ukážte, že $\min(u, v)$ a $\max(u + v - 1, 0)$ sú kopule.

Príklad 2. Nech X, Y sú náhodné premenné so združeným rozdelením daným kopulou C. Nájdite rozdelenie $Z = (X, Y)$, ak

1. $X, Y \sim U(0, 1)$, $C(u, v) = \min(u, v)$
2. $X, Y \sim U(0, 2)$, $C(u, v) = uv$
3. $X, Y \sim Exp(1)$, C je Gumbelova kopula s parametrom $\theta = 1$.

Príklad 3. Nájdite limitu Frankovej kopuly

$$C_\theta^F(u, v) = -\frac{1}{\theta} \log \left[1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{e^{-\theta} - 1} \right], \ \theta \in \mathbb{R} \setminus \{0\}$$

Príklad 4. Nech X, Y majú exponenciálne rozdelenie s parametrami $\lambda_x = 0.5$ a $\lambda_y = 0.3$. Nájdite ich združené rozdelenie, ktoré je dané Gumbelovou kopulou s parametrom $\theta = 3$.

Príklad 5. Nájdite príklad dvojrozmerného rozdelenia, ktoré nie je štandardizované normálne, ale jeho marginálne rozdelenia sú štandardizované normálne.

Príklad 6. Nech X a Y sú náhodné premenné so združenou distribučnou funkciou

$$H(x, y) = (1 + e^{-x} + e^{-y})^{-1} \forall x, y \in \mathbb{R}$$

- Nájdite marginálne rozdelenia X a Y
- Nájdite kopulu náhodných premenných X a Y

Príklad 7. Uvažujme funkciu

$$C_\rho(u,v) = \Phi_\rho(\Phi^{-1}(u), \Phi^{-1}(v)),$$

kde Φ_ρ označuje združenú distribučnú funkciu 2-rozmerného normálneho rozdelenia s nulovou strednou hodnotou a korelačným koeficientom ρ. Ukážte, že C_ρ je kopula. Overte, že C_ρ spĺňa v oboch smeroch Sklarovu vetu. Ukážte, že ak $\rho = 0$, potom $C_\rho = \Pi$.

4 Viacrozmerné normálne rozdelenie

Príklad 1. Určte rozdelenie náhodného vektora $Y = AX$, ak $A = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$, kde $X = (X_1, X_2)^T$ má štandardizované dvojrozmerné normálne rozdelenie. Ukážte, že náhodný vektor $Y = (Y_1, Y_2)^T$ má nezávislé zložky.

Príklad 2. Pomocou základných vlastností charakteristickej funkcie dokážte, že náhodný vektor X má rozdelenie $N_p(\mu, \Sigma)$ vtedy a len vtedy, keď pre každý vektor $a \in \mathbb{R}$ platí $a^T X \sim N(a^T\mu, a^T\Sigma a)$.

Príklad 3. Nech $X \sim N_2(\mu, \Sigma)$, kde $\mu = (2,2)^T$ a $\Sigma = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. Dokážte, že ak

$$A = \begin{bmatrix} 1 & -2 \\ -3 & 6 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 4 \\ 1 & 2 \end{bmatrix},$$

tak AX a BX sú nezávislé.

Príklad 4. Dokážte, že stredná hodnota náhodného vektora $X \sim N_p(\mu, \Sigma)$ je μ a jeho kovariančná matica je Σ.

1. Pomocou charakteristickej funkcie náhodného vektora $X \sim N_p(\mu, \Sigma)$
2. Pomocou vety o rozdelení podvektorov z prednášky.

Príklad 5. Nech vektor $(X, Y, Z)^T$ má hustotu

$$f(x,y,z) = \frac{\sqrt{3}}{16\pi^{3/2}} \exp\left(-\frac{1}{8}(2x^2 + 4y^2 - 2y(z+5) + (z+5)^2) \right)$$

pre $x, y, z \in \mathbb{R}$. Nájdite jeho kovariančnú maticu.

Príklad 6. Nech $X = (X_1, X_2)^T \sim N_4(\mu, \Sigma)$, kde X_1, X_2 sú dvojrozmerné podvektory X a

$$\mu = (2, -1, 3, 1)^T,$$

$$\Sigma = \begin{pmatrix} 7 & 3 & -3 & 2 \\ 3 & 6 & 0 & 4 \\ -3 & 0 & 5 & -2 \\ 2 & 4 & -2 & 4 \end{pmatrix}$$

Nájdite rozdelenie $X_1 | X_2 = x_2$

Príklad 7. Nech

$$X \sim N_2 \left(\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \right)$$

$$Y | X = x \sim N_2 \left(\begin{pmatrix} X_1 \\ X_1 + X_2 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right).$$

Nájdite rozdelenie $Y_2 | Y_1 = y_1$, kde $Y = (Y_1, Y_2)^T$.

Príklad 8. Nech $X_1, \ldots, X_n$ sú nezávislé náhodné vektory s rozdelením $X_j \sim N_p(\mu_j, \Sigma)$, $j = 1, \ldots, n$. Ukážte, že potom

$$V_1 = c_1 X_1 + \ldots + c_n X_n \sim N_p(\sum_{j=1}^{n} c_j \mu_j, (\sum_{j=1}^{n} c_j^2) \Sigma).$$

Ďalej, ak $V_2 = b_1 X_1 + \ldots + b_n X_n$, potom $V = (V_1, V_2)^T \sim N_{2p}(\mu_V, \Sigma_V)$, kde

$$\begin{bmatrix} (\sum_{j=1}^{n} c_j^2) \Sigma & (b^T c) \Sigma \\ (b^T c) \Sigma & (\sum_{j=1}^{n} b_j^2) \Sigma. \end{bmatrix}$$

Teda V_1 a V_2 sú nezávislé, ak $\sum_{j=1}^{n} c_j b_j = 0$.

5 Wishartovo rozdelenie a kvadratické formy

Príklad 1. Nech X_1, X_2, X_3 sú nezávislé náhodné premenné s rozdelením $N(0, 1)$. Nájdite rozdelenie kvadratických foriem

$$Y = \frac{2}{3}(X_1^2 + X_2^2 + X_3^2 - X_1 X_2 - X_1 X_3 - X_2 X_3),$$

$$Z = \frac{1}{3}(X_1^2 + X_2^2 + X_3^2 + 2X_1 X_2 + 2X_1 X_3 + 2X_2 X_3).$$

Príklad 2. Nech $X_1, \ldots, X_{20}$ je náhodný výber z $N_6(\mu, \Sigma)$. Nájdite

1. rozdelenie $(X_1 - \mu)' \Sigma^{-1} (X_1 - \mu)$

2. rozdelenie nS

3. rozdelenie $20BSB'$, ak $B = \begin{bmatrix} 1 & -1/2 & -1/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1/2 & -1/2 & 1 \end{bmatrix}$

Príklad 3. Nech $M \sim W_p(\Sigma, n)$. Nájdite rozdelenie náhodných premenných $M_{11}, \ldots, M_{pp}$.

Príklad 4. Nech $M \sim W_p(\Sigma, n)$ a nech $c \in \mathbb{R}^p$ je taký vektor, že $c^T \Sigma c > 0$. Vyjadrite pomocou distribučnej funkcie rozdelenia ξ_n^2 pravdepodobnosť udalosti $c^T M c \in (a, b)$, kde a, b.

Príklad 5. Dokážte, že ak S je výberová kovariančná matica pre náhodný výber $X_1, \ldots, X_n$ z rozdelenia $N_p(\mu, \Sigma)$, t.j.

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

tak platí $nS \sim W_p(\Sigma, n-1)$. Ukážte, že $\frac{n}{n-1} S$ je nevychýlený odhad matice Σ.

Príklad 6. Nech S je výberová kovariančná matica a $\bar{X}$ je výberový priemer pre náhodný výber $X_1, \ldots, X_n$ z rozdelenia $N_p(\mu, \Sigma)$. Dokážte, že potom S a $\bar{X}$ sú nezávislé.

Príklad 7. Overte, že Hottelingovo rozdelenie $T^2(1, m)$ je identické rozdeleniu $F(1, m)$ a rozdeleniu náhodnej premennej T^2, kde $T \sim t_m$.

Príklad 8. Nech $X \sim N_p(\mu, \Sigma)$ a A je matica typu $p \times p$. Nájdite strednú hodnotu kvadratickej formy $X'AX$.

Príklad 9. Nech $\Lambda \sim \Lambda(1, m, n)$. Dokážte, že potom platí

$$\frac{m}{n} \frac{1 - \Lambda}{\Lambda} \sim F(n, m).$$

Príklad 10. Nech $\Lambda \sim \Lambda(p, m, 1)$. Dokážte, že potom platí

$$\frac{1 - \Lambda}{\Lambda} \sim T^2(p, m).$$

6 Metóda maximálnej vierohodnosti a test pomerom vierohodnosti

Príklad 1. Nech $X_1, \ldots, X_n$ je náhodný výber z rovnomerného rozdelenia na intervale $[0, \theta]$. Nájdite odhad parametra θ metódou maximálnej vierohodnosti.

Príklad 2. Nech $X_1, \ldots, X_n$ je náhodný výber z dvojrozmerného rozdelenia s hustotou

$$f(x_1, x_2) = \frac{1}{\theta_1^2\theta_2}\frac{1}{x_2}\exp\left\{-\left(\frac{x_1}{\theta_1 x_2} + \frac{x_2}{\theta_1\theta_2}\right)\right\},\ x_1, x_2 > 0.$$

1. Vypočítajte odhad parametra $\theta = (\theta_1, \theta_2)^T$ metódou maximálnej vierohodnosti.
2. Nájdite Cramer-Raovo dolné ohraničenie a asymptotickú varianciu $\hat{\theta}$.
3. Pomocou testu pomerom vierohodností testujte hypotézu $H_0 : \theta = (\theta_{01}, \theta_{02})^T$.

Príklad 3. Nech $X_1, \ldots, X_n$ je náhodný výber z rozdelenia $N_p(\mu, I_p)$. Nájdite odhad parametra μ metódou maximálnej vierohodnosti.

Príklad 4. Nech $X_1, \ldots, X_n$ je náhodný výber z rozdelenia $N_p(\mu, \Sigma)$, kde matica Σ je neznáma, ale vieme, že je diagonálna, t.j. $\Sigma = diag(\sigma_{11}, \ldots, \sigma_{pp})$. Nájdite odhady parametrov μ a Σ metódou maximálnej vierohodnosti.

Príklad 5. Pomocou Wilksovej vety odvoďte test pre testovanie hypotézy, že kocka je vyvážená, ak ňou hádžeme n krát.

Príklad 6. Nech $X_1, \ldots, X_n$ je náhodný výber z rozdelenia $N_p(\mu, \Sigma)$. Uvažujme test pomerom vierohodnosti pre hypotézu H_0, že zložky náhodného vektora sú združene nezávislé, t.j. že Σ je diagonálna matica. Dokážte, že testovacia štatistika je $-n \ln \det(R)$, kde R je výberová korelačná matica.

7 Lineárne hypotézy o neznámych parametroch

Príklad 1. Uvažujme rozdelenie $N_3(\mu, \Sigma)$. Formulujte hypotézu $H_0 : \mu_1 = \mu_2 = \mu_3$ v tvare $A\mu = a$.

Príklad 2. Nasimulujte výber z normálneho rozdelenia s $\mu = (1, 2)^T$ a $\Sigma = \begin{pmatrix} 1 & 1/2 \\ 1/2 & 2 \end{pmatrix}$ a testujte $H_0 : 2\mu_1 - \mu_2 = 0.2$ pri známej aj pri neznámej Σ. Porovnajte výsledky.

Príklad 3. Nech $X \sim N_2(\mu, \Sigma)$, kde $\Sigma = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$. Z tohto rozdelenia urobíme náhodný výber o rozsahu $n = 6$ s výberovým priemerom $\bar{x} = (1, 1/2)^T$. Na hladine $\alpha = 0.05$ testujte nasledovné hypotézy:

1. $H_0 : \mu = (2, 2/3)^T$
2. $H_0 : \mu_1 + \mu_2 = 7/2$
3. $H_0 : \mu_1 - \mu_2 = 1/2$

4. $H_0 : \mu_1 = 2$

Príklad 4. Testujte hypotézy z predchádzajúceho príkladu pri neznámej Σ a $S = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$.

Príklad 5. Nech $X \sim N_p(\mu, \Sigma)$, kde Σ nepoznáme, ale vieme, že to je diagonálna matica. Odvoďte test pre testovanie hypotézy $H_0 : \mu = \mu_0$.