Nanyang Technological University, SPMS, MAS

MH4501 Multivariate Analysis, Semester 2 AY 2022-23

Assignment 1 (Due date: 4:30 PM, 9 February, 2023)

Remark: NTU places very high importance on honesty in academic work submitted by students, and adopts a policy of **ZERO** tolerance on cheating and plagiarism.

1. Let $x_1, \ldots, x_n$ be n samples from a multivariate population. Let

$$\bar{x}_n = \frac{1}{n}\sum_{i=1}^{n} x_i,\ d_{n+1} = x_{n+1} - \bar{x}_n,\ e_n = x_n - \bar{x}_n,\ Q_n = \sum_{i=1}^{n}(x_i - \bar{x}_n)(x_i - \bar{x}_n)^\top.$$

Verify the following recursive formulas:

(a) (when we add one sample,)

$$\bar{x}_{n+1} = \bar{x}_n + \frac{d_{n+1}}{n+1},\ Q_{n+1} = Q_n + \left(1 - \frac{1}{n+1}\right) d_{n+1}d_{n+1}^\top;$$

(b) (when we remove one sample,)

$$\bar{x}_{n-1} = \bar{x}_n - \frac{e_n}{n-1},\ Q_{n-1} = Q_n - \left(1 + \frac{1}{n-1}\right) e_n e_n^\top.$$

2. Let Y be a random vector with mean $\mu = \mathbb{E}[Y]$ and covariance matrix $\Sigma = \mathrm{Cov}(Y)$, and A is a constant matrix. Prove that

$$\mathbb{E}[Y^\top AY] = \mathrm{tr}(A\Sigma) + \mu^\top A\mu.$$

3. (a) Let $B_{p\times p}$ be a symmetric positive-definite matrix. Use the eigenvalue decomposition of B to prove that

i.
$$\max_{x\neq 0} \frac{x^\top Bx}{x^\top x} = \lambda_1 \quad \text{(attained when } x = u_1\text{)}$$

where λ_1 is the largest eigenvalue of B and u_1 is the associated normalized eigenvector.

(b) Find the maximum value of the ratio $\frac{x^\top Ax}{x^\top x}$ for any nonzero vectors $x = (x_1, x_2, x_3)^\top$ if

$$A = \begin{pmatrix} 13 & -4 & 2 \\ -4 & 13 & -2 \\ 2 & -2 & 10 \end{pmatrix}.$$

4. [Example of a non-normal bivariate distribution with normal marginals]
Let X_1 be $N(0,1)$, and let

$$X_2 = \begin{cases} -X_1, & \text{if } -1 \le X_1 \le 1, \\ X_1, & \text{otherwise.} \end{cases}$$

Show each of the following.

(a) X_2 also has an $N(0,1)$ distribution.
Hint: try to prove $\mathbb{P}(X_2 \le x) = \mathbb{P}(X_1 \le x)$ for three cases when $x < -1$; when $-1 \le x \le 1$; and when $x > 1$.

(b) X_1 and X_2 do not have a bivariate normal distribution.
Hint: find a contradiction to the multinormal properties by showing

$$\mathbb{P}(X_1 - X_2 = 0) \neq 0.$$

Note that probability for a continuous random variable at a point is zero.

5. Let X be $N_3(\mu, \Sigma)$ with $\mu = (2, -3, 1)^\top$ and

$$\Sigma = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 2 & 2 \end{pmatrix}.$$

(a) Find the distribution of $3X_1 - 2X_2 + X_3$.

(b) Relabel the variables if necessary, and find a 2×1 vector a such that X_2 and $X_2 - a^\top \begin{pmatrix} X_1 \\ X_3 \end{pmatrix}$ are independent.

[Optional Survey] How do you rate the difficulty of this assignment (1=easy, 5=hard)? Any feedback about lectures, assignments, tutorials, and labs so far?