Multivariate Statistics

Sudipta Das

Assistant Professor,
Department of Data Science,
Ramakrishna Mission Vivekananda University, Kolkata
Slides adapted from Jhonson & Winchern

Outline I

1. Review of Linear Algebra
   - Vectors and Matrix
   - Matrix inequalities and Maximization

Matrix and Random Vectors I

- Vectors

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_p \end{bmatrix} \quad or \quad x' = \left[x_1, x_2, x_3, \ldots, x_p\right]$$

- Eucledian distance form origin, length or 2-norm

$$||x||_2 = \sqrt{x_1^2 + x_2^2 + \ldots + x_p^2}.$$

- Angle between vectors x and y

$$\cos(\theta) = \frac{x'.y}{||x||_2||y||_2}$$

Matrix and Random Vectors II

- Linear dependence of vectors:- A set of vectors $x_1, x_2, \ldots, x_n$ is said to be linearly dependent if there exists constants $c_1, c_2, \ldots, c_k$, not all zero, such that

$$c_1x_1 + c_2x_2 + \ldots + c_nx_n = 0.$$

- Vectors of same dimensions that are not linearly dependent are said to be linearly independent.

Matrix and Random Vectors III

- Matrices

$$A_{n\times p} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1p} \\ x_{21} & x_{22} & \dots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{np} \end{bmatrix}$$

- A square matrix $A_{n\times n}$ is symmetric if $A = A'$.
- Inverse of a square matrix A is A^{-1}, where $|A| \neq 0$ and $AA^{-1} = I$.
- A matrix Q is called orthogonal matrix if

$$Q^{-1} = Q'.$$

- A square matrix A is said to have an eigenvalue λ, with corresponding eigenvector $x \neq 0$, if

$$Ax = \lambda x.$$

Matrix and Random Vectors IV

- Result: Let A be a $n \times n$ square symmetric matrix. Then A has n pairs of eigenvalues and eigenvectors-namely,

$$\lambda_1, e_1; \lambda_2, e_2; \ldots; \lambda_n, e_n.$$

The eigenvectors can be chosen to satisfy $1 = e_1' e_1 = \ldots = e_n' e_n$ and be mutually perpendicular. The eigenvectors are unique unless two or more eigenvalues are equal.

- Result: The spectral decomposition of a $n \times n$ symmetric matrix A is given by

$$A = \lambda_1 e_1 e_1' + \lambda_2 e_2 e_2' \ldots + \lambda_n e_n e_n'.$$

- Example 2.10 (Page 61)

Matrix and Random Vectors V

- A square matrix A is said to be positve definite if

$$x'Ax > 0$$

for all vectors $x \neq 0$.

- Spectral Decomposition of square syemmetric

$$A = P\Lambda P',$$

where

$$P = [e_1 : e_2 : \ldots e_n] = \begin{bmatrix} e_{11} & e_{21} & \ldots & e_{n1} \\ e_{12} & e_{22} & \ldots & e_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ e_{1n} & e_{2n} & \ldots & e_{nn} \end{bmatrix},$$

and

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}.$$

- Thus,
  - Inverse

$$A^{-1} = P\Lambda^{-1}P'.$$

  - Sqaure Root

$$A^{\frac{1}{2}} = P\Lambda^{\frac{1}{2}}P'.$$

  - Factorization

$$A = A^{\frac{1}{2}}A^{\frac{1}{2}}.$$

Matrix inequalities and Maximization I

- Cauchy-Schwarz Inequality: Let b and d any two $p \times 1$ vectors. Then

$$(b'd)^2 \leq (b'b)(d'd)$$

with equality iff $b = cd$ for some constant c.

Matrix inequalities and Maximization II

- Extended Cauchy-Schwarz Inequality: Let b, d be any two $p \times 1$ vectors and B be a positive definite matrix. Then

$$(b'd)^2 \leq (b'Bb)(d'B^{-1}d)$$

with equality iff $b = cB^{-1}d$ for some constant c.

Matrix inequalities and Maximization III

- Maximization Lemma: Let B be positive definite and d be a given vector. Then, for an arbitrary nonzero vector x,

$$\max_{x \neq 0} \frac{(x'd)^2}{x'Bx} = d'B^{-1}d$$

with the maximum attained when $x = cB^{-1}d$ for any constant $c \neq 0$.

Matrix inequalities and Maximization IV

- Maximization of Quadratic Forms for Points on the Unit Sphere: Let $B_{p\times p}$ be a positive definite matrix with eigenvalues $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \ldots \geq \lambda_p \geq 0$ and associated normalized eigenvectors $e_1, e_2, \ldots, e_p$. Then

  1. $$\max_{x\neq 0} \frac{x'Bx}{x'x} = \lambda_1 \text{ attained when } x = e_1$$

     and

  2. $$\min_{x\neq 0} \frac{x'Bx}{x'x} = \lambda_p \text{ attained when } x = e_p$$

  3. Moreover, for $k = 1, \ldots, p-1$

     $$\max_{x\perp e_1,\ldots,e_k} \frac{x'Bx}{x'x} = \lambda_{k+1} \text{ attained when } x = e_{k+1}.$$

Matrix inequalities and Maximization V

- Sketch of proof:
  Let $B = P\Lambda P'$ and $y = P'x$, where $P = [e_1 : e_2 : \ldots : e_p]$ and $\Lambda = diag(\lambda_1, \lambda_2, \ldots, \lambda_p)$.
  1. Thus,

$$\frac{x'Bx}{x'x} = \frac{x'P\Lambda P'x}{x'PP'x} = \frac{y'\Lambda y}{y'y}$$

Hence,

$$\max_{x\neq 0} \frac{x'Bx}{x'x} \Leftrightarrow \max_{y\neq 0} \frac{y'\Lambda y}{y'y}$$

Now,

$$\frac{y'\Lambda y}{y'y} = \frac{\sum_{i=1}^{p} \lambda_i y_i^2}{\sum_{i=1}^{p} y_i^2} \leq \lambda_1 \frac{\sum_{i=1}^{p} y_i^2}{\sum_{i=1}^{p} y_i^2} = \lambda_1.$$

Also, the maximum is attained at $y = [1, 0, \ldots, 0]'$, equivalently at $x = Py = e_1$.

Matrix inequalities and Maximization VI

2. Similarly.
3. Note that
$$\mathbf{x} = P\mathbf{y} = y_1\mathbf{e_1} + \ldots + y_i\mathbf{e_i} + \cdots + y_p\mathbf{e_p}$$
and
$$\mathbf{e_i'x} = y_1\mathbf{e_i'e_1} + \ldots + y_i\mathbf{e_i'e_i} + \cdots + y_p\mathbf{e_i'e_p} = y_i$$
Hence,
$$\mathbf{x}\perp\mathbf{e}_1, \ldots, \mathbf{e}_k \Rightarrow y_i = 0 \forall i \leq k$$

Matrix inequalities and Maximization VII

Thus,

$$\max_{\mathbf{x}\perp\mathbf{e}_1,\ldots,\mathbf{e}_k} \frac{\mathbf{x}'B\mathbf{x}}{\mathbf{x}'\mathbf{x}} \Leftrightarrow \max_{\mathbf{y}:y_i=0\forall i\leq k} \frac{\mathbf{y}'\Lambda\mathbf{y}}{\mathbf{y}'\mathbf{y}}$$

Therefore,

$$\frac{y'\Lambda y}{y'y} = \frac{\sum_{i=1}^{p}\lambda_i y_i^2}{\sum_{i=1}^{p} y_i^2} = \frac{\sum_{i=k+1}^{p}\lambda_i y_i^2}{\sum_{i=k+1}^{p} y_i^2} \leq \lambda_{k+1}\frac{\sum_{i=k+1}^{p} y_i^2}{\sum_{i=k+1}^{p} y_i^2} = \lambda_{k+1}.$$

Also, the maximum is attained at $y = [\underbrace{0,\ldots,0}_{k},1,\ldots,0]'$, equivalently at $x = Py = e_{k+1}$.