Matrix Algebra and Random Vectors

Shyh-Kang Jeng
Department of Electrical Engineering/
Graduate Institute of Communication/
Graduate Institute of Networking and Multimedia

1

Outline

- Eigenvalues and Eigenvectors
- Positive Definite Matrices
- A Square-Root Matrix
- Random Vectors and Matrices
- Mean Vectors and Covariance Matrices
- Matrix Inequalities and Maximization

2

Outline

- Eigenvalues and Eigenvectors
- Positive Definite Matrices
- A Square-Root Matrix
- Random Vectors and Matrices
- Mean Vectors and Covariance Matrices
- Matrix Inequalities and Maximization

3

Questions

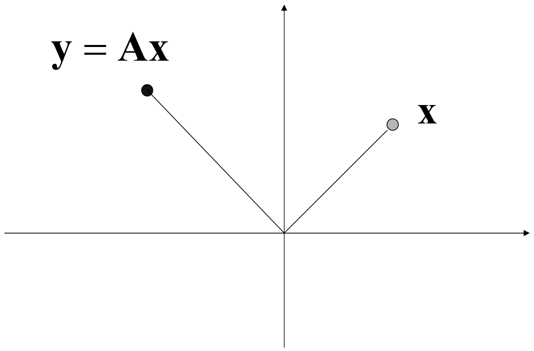
- What are the eigenvalues and eigenvectors?
- How to compute eigenvalues and eigenvectors?
- What is a quadratic form?
- How to express the statistical distance as a quadratic form?

4

Questions

- How to simplify a quadratic form?
- What is an orthogonal matrix?
- How to diagonalize a square matrix?
- What is the spectral decomposition?

Concept of Mapping



Eigenvalues

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$$

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = 0$$

$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 1 & 3 \end{bmatrix}$$

$$|\mathbf{A} - \lambda\mathbf{I}| = \begin{vmatrix} 1-\lambda & 0 \\ 1 & 3-\lambda \end{vmatrix} = (1-\lambda)(3-\lambda) = 0$$

$$\lambda_1 = 1, \quad \lambda_2 = 3$$

Eigenvectors

$$(\mathbf{A} - \lambda_i\mathbf{I})\mathbf{x}_i = 0$$

$$\lambda_1 = 1$$

$$\begin{bmatrix} 0 & 0 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \mathbf{x}_1 = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

$$\lambda_2 = 3$$

$$\begin{bmatrix} -2 & 0 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \mathbf{x}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Quadratic Form

$$Q(\mathbf{x}) = \mathbf{x}'\mathbf{A}\mathbf{x} = \sum_{i=1}^{k}\sum_{j=1}^{k} a_{ij}x_i x_j$$

$$Q(x) = \begin{bmatrix} x_1 & x_2 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1^2 + 2x_1x_2 + x_2^2$$

$$Q(x) = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}\begin{bmatrix} 1 & 3 & 0 \\ 3 & -1 & -2 \\ 0 & -2 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$= x_1^2 + 6x_1x_2 - x_2^2 - 4x_2x_3 + 2x_3^2$$

9

General Statistical Distance

$$P(x_1, x_2, \cdots, x_p), \quad O(0,0,\cdots,0), \quad Q(y_1, y_2, \cdots, y_p)$$

$$d(O,P) = \sqrt{\begin{array}{l}[a_{11}x_1^2 + a_{22}x_2^2 + \cdots + a_{pp}x_p^2 + \\ 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + \cdots + 2a_{p-1,p}x_{p-1}x_p]\end{array}}$$

$$d(P,Q) = \sqrt{\begin{array}{l}[a_{11}(x_1-y_1)^2 + a_{22}(x_2-y_2)^2 + \cdots + \\ a_{pp}(x_p-y_p)^2 + \\ 2a_{12}(x_1-y_1)(x_2-y_2) + 2a_{13}(x_1-y_1)(x_3-y_3) \\ + \cdots + 2a_{p-1,p}(x_{p-1}-y_{p-1})(x_p-y_p)]\end{array}}$$

10

General Statistical Distance

$$d^2(O,P) = [a_{11}x_1^2 + a_{22}x_2^2 + \cdots + a_{pp}x_p^2 +$$

$$2a_{12}x_1x_2 + 2a_{13}x_1x_3 + \cdots + 2a_{p-1,p}x_{p-1}x_p]$$

$$= \begin{bmatrix} x_1 & x_2 & \cdots & x_p \end{bmatrix}\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{12} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1p} & a_{2p} & \cdots & a_{pp} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix}$$

$$= \mathbf{x}'\mathbf{A}\mathbf{x}$$

11

Statistical Distance under Rotated Coordinate System

$$O(0,0), \quad P(\tilde{x}_1, \tilde{x}_2)$$

$$d(O,P) = \sqrt{\frac{\tilde{x}_1^2}{\tilde{s}_{11}} + \frac{\tilde{x}_2^2}{\tilde{s}_{22}}}$$

$$\tilde{x}_1 = x_1\cos\theta + x_2\sin\theta$$

$$\tilde{x}_2 = -x_1\sin\theta + x_2\cos\theta$$

$$d(O,P) = \sqrt{a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2}$$

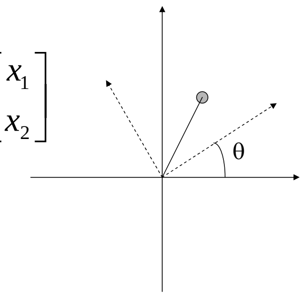
12

Rotated Coordinate System

$$\tilde{x}_1 = x_1 \cos\theta + x_2 \sin\theta$$

$$\tilde{x}_2 = -x_1 \sin\theta + x_2 \cos\theta$$

$$\begin{bmatrix} \tilde{x}_1 \\ \tilde{x}_2 \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$



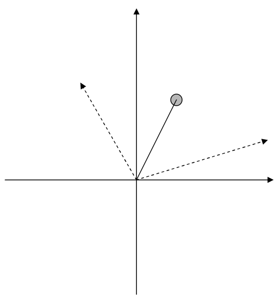
13

Coordinate Transformation

$$\mathbf{y} = \mathbf{B}\mathbf{x}$$

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_k \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1k} \\ b_{21} & b_{22} & \cdots & b_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ b_{k1} & b_{k2} & \cdots & b_{kk} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix}$$

$$\mathbf{y} = \begin{bmatrix} \mathbf{b}_1' \\ \mathbf{b}_2' \\ \vdots \\ \mathbf{b}_k' \end{bmatrix} \mathbf{x}$$



14

Quadratic Form in Transformed Coordinate Systems

$$\begin{aligned} Q(\mathbf{x}) &= \mathbf{x}'\mathbf{A}\mathbf{x} \\ &= (\mathbf{B}^{-1}\mathbf{y})'\mathbf{A}(\mathbf{B}^{-1}\mathbf{y}) \\ &= \mathbf{y}'\mathbf{B}^{-1'}\mathbf{A}\mathbf{B}^{-1}\mathbf{y} \\ &= \mathbf{y}'\mathbf{\Lambda}\mathbf{y} = Q(\mathbf{y}) \end{aligned}$$

$$\mathbf{B}^{-1'}\mathbf{A}\mathbf{B}^{-1} = \mathbf{\Lambda}$$

$$\mathbf{A}\mathbf{B}^{-1} = \mathbf{B}'\mathbf{\Lambda}$$

15

Diagonalized Quadratic Form

$$\mathbf{\Lambda} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_k \end{bmatrix}$$

$$Q(\mathbf{y}) = \mathbf{y}'\mathbf{\Lambda}\mathbf{y} = \sum_{i=1}^{k} \lambda_i y_i^2$$

16

Orthogonal Matrix

$$\mathbf{AA'} = \mathbf{A'A} = \mathbf{I}$$

$$\mathbf{A}^{-1} = \mathbf{A'}$$

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}^{-1} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

17

Diagonalization

$$\mathbf{B}^{-1} = \mathbf{B'}$$

$$\mathbf{AB'} = \mathbf{B'\Lambda}$$

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{bmatrix} [\mathbf{b}_1 \quad \mathbf{b}_2 \quad \cdots \quad \mathbf{b}_k] =$$

$$[\mathbf{b}_1 \quad \mathbf{b}_2 \quad \cdots \quad \mathbf{b}_k] \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_k \end{bmatrix}$$

$$\mathbf{Ab}_i = \lambda_i \mathbf{b}_i$$

18

Spectral Decomposition

$$\mathbf{AB'} = \mathbf{B'\Lambda}$$

$$\mathbf{A} = \mathbf{B'\Lambda B}$$

$$\mathbf{A} = [\mathbf{e}_1 \quad \mathbf{e}_2 \quad \ldots \quad \mathbf{e}_k] \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_k \end{bmatrix} \begin{bmatrix} \mathbf{e}_1' \\ \mathbf{e}_2' \\ \vdots \\ \mathbf{e}_k' \end{bmatrix}$$

$$= \sum_{i=1}^{k} \lambda_i \mathbf{e}_i \mathbf{e}_i'$$

19

Outline

- Eigenvalues and Eigenvectors
- Positive Definite Matrices
- A Square-Root Matrix
- Random Vectors and Matrices
- Mean Vectors and Covariance Matrices
- Matrix Inequalities and Maximization

20

Question

- What is a positive definite matrix?
- Why to learn the concept of positive definite matrix?

21

Positive Definite Matrix

- Matrix $\mathbf{A}$ is non-negative definite if

$$\mathbf{x}'\mathbf{A}\mathbf{x} \geq 0$$

for all $\mathbf{x}' = [x_1, x_2, \ldots, x_k]$

- Matrix A is positive definite if

$$\mathbf{x}'\mathbf{A}\mathbf{x} > 0$$

for all non-zero $\mathbf{x}$

22

Positive Definite Matrix

$$3x_1^2 + 2x_2^2 - 2\sqrt{2}x_1x_2 = [x_1 \quad x_2]\begin{bmatrix} 3 & -\sqrt{2} \\ -\sqrt{2} & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$= \mathbf{xAx}$$

$$\lambda_1 = 4, \lambda_2 = 1$$

$$\mathbf{A} = 4\mathbf{e}_1\mathbf{e}_1^{'} + \mathbf{e}_2\mathbf{e}_2^{'}$$

$$\mathbf{x}'\mathbf{A}\mathbf{x} = 4\mathbf{x}'\mathbf{e}_1\mathbf{e}_1^{'}\mathbf{x} + \mathbf{x}'\mathbf{e}_2\mathbf{e}_2^{'}\mathbf{x} = 4y_1^2 + y_2^2 \geq 0$$

$$y_1 = \mathbf{x}'\mathbf{e}_1 = \mathbf{e}_1^{'}\mathbf{x},\ y_2 = \mathbf{x}'\mathbf{e}_2 = \mathbf{e}_2^{'}\mathbf{x}$$

$$\mathbf{y} = \mathbf{Ex}, \mathbf{E} \text{ orthogonal and } \mathbf{E}^{-1} = \mathbf{E}'$$

$$\mathbf{x}'\mathbf{A}\mathbf{x} > 0 \text{ for all } \mathbf{x}' \neq \mathbf{0}$$

$$\therefore \mathbf{A} \text{ is positive definite}$$

23

Outline

- Eigenvalues and Eigenvectors
- Positive Definite Matrices
- A Square-Root Matrix
- Random Vectors and Matrices
- Mean Vectors and Covariance Matrices
- Matrix Inequalities and Maximization

24

Question

- How to find the square-root of a symmetric square matrix?
- Why to learn the square-root matrix?

25

Inverse and Square-Root Matrix

$$\mathbf{A} = \sum_{i=1}^{k} \lambda_i \mathbf{e}_i \mathbf{e}_i' = \mathbf{P\Lambda P}',$$

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

$$\mathbf{A}^{n} = (\mathbf{P\Lambda P}')(\mathbf{P\Lambda P}')\cdots(\mathbf{P\Lambda P}') = \mathbf{P\Lambda}^{n}\,\mathbf{P}'$$

$$\lambda_i > 0, i = 1, \cdots, k$$

$$\mathbf{A}^{1/2} = \sum_{i=1}^{k} \sqrt{\lambda_i}\mathbf{e}_i \mathbf{e}_i' = \mathbf{P\Lambda}^{1/2}\mathbf{P}'$$

$$\mathbf{A}^{1/2}\mathbf{A}^{1/2} = \mathbf{A}$$

$$\left(\mathbf{A}^{1/2}\right)' = \mathbf{A}^{1/2}$$

$$\mathbf{A}^{-1/2} = \left(\mathbf{A}^{1/2}\right)^{-1}$$

26

Outline

- Eigenvalues and Eigenvectors
- Positive Definite Matrices
- A Square-Root Matrix
- Random Vectors and Matrices
- Mean Vectors and Covariance Matrices
- Matrix Inequalities and Maximization

27

Question

- What is a random vector?
- What is a random matrix?
- How to find the expectation of a random matrix?

28

Random Vectors and Random Matrices

- Random vector
  - Vector whose elements are random variables
- Random matrix
  - Matrix whose elements are random variables

29

Expected Value of a Random Matrix

$$E(\mathbf{X}) = \begin{bmatrix} E(X_{11}) & E(X_{12}) & \cdots & E(X_{1p}) \\ E(X_{21}) & E(X_{22}) & \cdots & E(X_{2p}) \\ \vdots & \vdots & \ddots & \vdots \\ E(X_{p1}) & E(X_{p2}) & \cdots & E(X_{pp}) \end{bmatrix}$$

$$E(X_{ij}) = \begin{cases} \int_{-\infty}^{\infty} x_{ij} f_{ij}(x_{ij}) dx_{ij} \\ \sum_{all\, x_{ij}} x_{ij} p_{ij}(x_{ij}) \end{cases}$$

$$E(\mathbf{AXB}) = \mathbf{A}E(\mathbf{X})\mathbf{B}$$

30

Outline

- Eigenvalues and Eigenvectors
- Positive Definite Matrices
- A Square-Root Matrix
- Random Vectors and Matrices
- Mean Vectors and Covariance Matrices
- Matrix Inequalities and Maximization

31

Questions

- What is the population?
- How to compute population mean vectors and population covariance matrices?
- What is the statistical independence?
- If cov(x1, x2) = 0, will x1 and x2 be statistical independent?

32

Questions

- How to compute the population correlation coefficient matrix?
- How to partition the covariance matrix?
- How to compute population mean vector and covariance matrix of a linear combination of random vectors?

33

Questions

- What are the sample mean vectors and the sample covariance matrices?
- How to partition a sample mean vector and a sample covariance matrix?

34

Population Mean Vectors

Random vector $\mathbf{X} = \begin{bmatrix} X_1 & X_2 & \cdots & X_p \end{bmatrix}'$

Joint probability density function

$f(\mathbf{x}) = f(x_1, x_2, \cdots, x_p)$

Marginal probability distribution $f(x_i)$

$\mu_i = E(X_i)$

$\sigma_i^2 = E(X_i - \mu_i)^2$

$\boldsymbol{\mu} = E(\mathbf{X}) = \begin{bmatrix} \mu_1 & \mu_2 & \cdots & \mu_p \end{bmatrix}$

35

Covariance

$$\mathrm{Cov}(X_i, X_k) = E(X_i - \mu_i)(X_k - \mu_k)$$

$$= \begin{cases} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_i - \mu_i)(x_k - \mu_k) f_{ik}(x_i, x_k) dx_i dx_k \\ \sum_{all\ x_i} \sum_{all\ x_k} (x_i - \mu_i)(x_k - \mu_k) p_{ik}(x_i, x_k) \end{cases}$$

$$= \sigma_{ik}$$

36

Statistically Independent

$$P[X_i \le x_i \text{ and } X_k \le x_k] = P[X_i \le x_i]P[X_k \le x_k]$$

$$f_{ik}(x_i, x_k) = f_i(x_i)\, f_k(x_k)$$

$$f_{12\cdots p}(x_1, x_2, \cdots, x_p) = f_1(x_1)\, f_2(x_2) \cdots f_p(x_p)$$

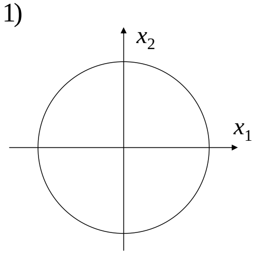
$\mathrm{Cov}(X_i, X_k) = 0$ if X_i, X_k are independent

(the converse is not true in general)

Joint Probability Density Function

$$f(x_1, x_2) = \begin{cases} f(\theta) = \dfrac{1}{2\pi}, x_1^2 + x_2^2 = 1, \theta = \mathrm{atan}\,2(x_2, x_1) \\ 0, otherwise \end{cases}$$

$$= \frac{1}{2\pi}\delta(x_1^2 + x_2^2 - 1)$$



Marginal Probability Density Function

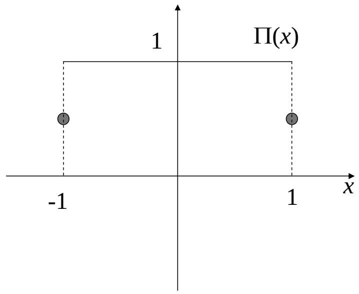
$$f_1(x_1) = \int_{-\infty}^{\infty} f(x_1, x_2)dx_2 = \int_{-\infty}^{\infty} \frac{1}{2\pi}\left\{\delta(x_2 - \sqrt{1 - x_1^2}) + \delta(x_2 + \sqrt{1 - x_1^2})\right\}dx_2$$

$$= \begin{cases} \dfrac{1}{\pi}, |x_1| < 1 \\ \dfrac{1}{2\pi}, |x_1| = 1 \\ 0, otherwise \end{cases}$$

$$= \frac{1}{\pi}\Pi(x_1)$$

$$f_2(x_2) = \frac{1}{\pi}\Pi(x_2)$$

$$f(x_1, x_2) \ne f_1(x_1)\, f_2(x_2)$$



Marginal Means and Covariance

$$\mu_1 = \int_{-\infty}^{\infty} x_1 f_1(x_1)dx_1 = \frac{1}{\pi}\int_{-1}^{1} x_1 dx_1 = 0$$

$$\mu_2 = 0$$

$$\mathrm{Cov}(x_1, x_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x_1 - \mu_1)(x_2 - \mu_2)\, f(x_1, x_2)dx_1 dx_2$$

$$= \int_{-\pi}^{\pi}\int_{0}^{\infty}\rho\cos\theta\rho\sin\theta f(\theta)\delta(\rho - 1)\rho d\rho d\theta$$

$$= \int_{-\pi}^{\pi}\cos\theta\sin\theta\frac{1}{2\pi}d\theta = 0$$

Population Variance-Covariance Matrices

$$\boldsymbol{\Sigma} = E(\mathbf{X}-\boldsymbol{\mu})(\mathbf{X}-\boldsymbol{\mu})'$$

$$= E\left(\begin{bmatrix} X_1-\mu_1 \\ X_2-\mu_2 \\ \vdots \\ X_p-\mu_p \end{bmatrix}\begin{bmatrix} X_1-\mu_1 & X_2-\mu_2 & \cdots & X_p-\mu_p \end{bmatrix}\right)$$

$$= \mathrm{Cov}(\mathbf{X}) = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_{pp} \end{bmatrix}$$

41

Population Correlation Coefficients

$$\boldsymbol{\rho} = \begin{bmatrix} \rho_{11} & \rho_{12} & \cdots & \rho_{1p} \\ \rho_{21} & \rho_{22} & \cdots & \rho_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{p1} & \rho_{p2} & \cdots & \rho_{pp} \end{bmatrix}$$

$$\rho_{ik} = \frac{\sigma_{ik}}{\sqrt{\sigma_{ii}}\sqrt{\sigma_{kk}}}$$

$$\rho_{ii} = 1$$

42

Standard Deviation Matrix

$$\mathbf{V}^{1/2} = \begin{bmatrix} \sqrt{\sigma_{11}} & 0 & \cdots & 0 \\ 0 & \sqrt{\sigma_{22}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{\sigma_{pp}} \end{bmatrix}$$

$$\boldsymbol{\Sigma} = \mathbf{V}^{1/2}\boldsymbol{\rho}\mathbf{V}^{1/2}$$

$$\boldsymbol{\rho} = \mathbf{V}^{-1/2}\boldsymbol{\Sigma}\mathbf{V}^{-1/2}$$

43

Correlation Matrix from Covariance Matrix

$$\boldsymbol{\Sigma} = \begin{bmatrix} 4 & 1 & 2 \\ 1 & 9 & -3 \\ 2 & -3 & 25 \end{bmatrix} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{bmatrix}$$

$$\mathbf{V}^{1/2} = \begin{bmatrix} \sqrt{\sigma_{11}} & 0 & 0 \\ 0 & \sqrt{\sigma_{22}} & 0 \\ 0 & 0 & \sqrt{\sigma_{33}} \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix}$$

$$\mathbf{V}^{-1/2} = \begin{bmatrix} 1/2 & 0 & 0 \\ 0 & 1/3 & 0 \\ 0 & 0 & 1/5 \end{bmatrix}$$

$$\boldsymbol{\rho} = \mathbf{V}^{-1/2}\boldsymbol{\Sigma}\mathbf{V}^{-1/2} = \begin{bmatrix} 1 & 1/6 & 1/5 \\ 1/6 & 1 & -1/5 \\ 1/5 & -1/5 & 1 \end{bmatrix}$$

44

Partitioning Covariance Matrix

$$\mathbf{X} = \begin{bmatrix} X_1 \\ \vdots \\ X_q \\ --- \\ X_{q+1} \\ \vdots \\ X_p \end{bmatrix} = \begin{bmatrix} \mathbf{X}^{(1)} \\ --- \\ \mathbf{X}^{(2)} \end{bmatrix}, \boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\mu}^{(1)} \\ --- \\ \boldsymbol{\mu}^{(2)} \end{bmatrix}$$

$$(\mathbf{X}-\boldsymbol{\mu})(\mathbf{X}-\boldsymbol{\mu})' =$$

$$\begin{bmatrix} (\mathbf{X}^{(1)}-\boldsymbol{\mu}^{(1)})(\mathbf{X}^{(1)}-\boldsymbol{\mu}^{(1)})' & (\mathbf{X}^{(1)}-\boldsymbol{\mu}^{(1)})(\mathbf{X}^{(2)}-\boldsymbol{\mu}^{(2)})' \\ (\mathbf{X}^{(2)}-\boldsymbol{\mu}^{(2)})(\mathbf{X}^{(1)}-\boldsymbol{\mu}^{(1)})' & (\mathbf{X}^{(2)}-\boldsymbol{\mu}^{(2)})(\mathbf{X}^{(2)}-\boldsymbol{\mu}^{(2)})' \end{bmatrix}$$

Partitioning Covariance Matrix

$$\boldsymbol{\Sigma} = E(\mathbf{X}-\boldsymbol{\mu})(\mathbf{X}-\boldsymbol{\mu})' = \begin{bmatrix} \boldsymbol{\Sigma}_{11} & | & \boldsymbol{\Sigma}_{12} \\ --- & | & --- \\ \boldsymbol{\Sigma}_{21} & | & \boldsymbol{\Sigma}_{22} \end{bmatrix}$$

$$= \begin{bmatrix} \sigma_{11} & \cdots & \sigma_{1q} & | & \sigma_{1,q+1} & \cdots & \sigma_{1p} \\ \vdots & \ddots & \vdots & | & \vdots & \ddots & \vdots \\ \sigma_{q1} & \cdots & \sigma_{qq} & | & \sigma_{q,q+1} & \cdots & \sigma_{qp} \\ -- & -- & -- & | & -- & -- & -- \\ \sigma_{q+1,1} & \cdots & \sigma_{q+1,q} & | & \sigma_{q+1,q+1} & \cdots & \sigma_{q+1,p} \\ \vdots & \ddots & \vdots & | & \vdots & \ddots & \vdots \\ \sigma_{p1} & \cdots & \sigma_{pq} & | & \sigma_{p,q+1} & \cdots & \sigma_{pp} \end{bmatrix}$$

$$\boldsymbol{\Sigma}_{21} = \boldsymbol{\Sigma}_{12}^{'}$$

Linear Combinations of Random Variables

Linear combination $\mathbf{c}'\mathbf{X} = c_1 X_1 + \cdots + c_p X_p$

has

$$\text{mean} = E(\mathbf{c}'\mathbf{X}) = \mathbf{c}'\boldsymbol{\mu}$$

$$\text{variance} = \text{Var}(\mathbf{c}'\mathbf{X}) = \mathbf{c}'\boldsymbol{\Sigma}\mathbf{c}$$

where $\boldsymbol{\mu} = E(\mathbf{X})$ and $\boldsymbol{\Sigma} = \text{Cov}(\mathbf{X})$

Example of Linear Combinations of Random Variables

$$E(aX_1 + bX_2) = aE(X_1) + bE(X_2) = a\mu_1 + b\mu_2$$

$$\text{Var}(aX_1 + bX_2) = E[(aX_1 + bX_2) - (a\mu_1 + b\mu_2)]^2$$

$$= E[a(X_1 - \mu_1) + b(X_2 - \mu_2)]^2$$

$$= a^2\sigma_{11} + b^2\sigma_{22} + 2ab\sigma_{12}$$

$$\mathbf{c}' = [a, b], \mathbf{X}' = [X_1, X_2]$$

$$E(\mathbf{c}'\mathbf{X}) = \mathbf{c}'\boldsymbol{\mu}$$

$$\mathbf{Var}(\mathbf{c}'\mathbf{X}) = [a\ b]\begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} = \mathbf{c}'\boldsymbol{\Sigma}\mathbf{c}$$

Linear Combinations of Random Variables

$$\mathbf{Z} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1p} \\ c_{21} & c_{22} & \cdots & c_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ c_{p1} & c_{p2} & \cdots & c_{pp} \end{bmatrix} \mathbf{X} = \mathbf{CX}$$

$$\boldsymbol{\mu}_{\mathbf{Z}} = E(\mathbf{Z}) = E(\mathbf{CX}) = \mathbf{C}\boldsymbol{\mu}_{\mathbf{X}}$$

$$\boldsymbol{\Sigma}_{\mathbf{Z}} = \mathrm{Cov}(\mathbf{CX}) = \mathbf{C}\boldsymbol{\Sigma}_{\mathbf{X}}\mathbf{C}'$$

49

Sample Mean Vector and Covariance Matrix

$$\bar{\mathbf{x}}' = [\bar{x}_1, \bar{x}_2, \cdots \bar{x}_p]$$

$$\mathbf{S}_n = \begin{bmatrix} s_{11} & \cdots & s_{1p} \\ \vdots & \ddots & \vdots \\ s_{1p} & \cdots & s_{pp} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{n}\sum_{j=1}^{n}\left(x_{j1} - \bar{x}_1\right)^2 & \cdots & \frac{1}{n}\sum_{j=1}^{n}\left(x_{j1} - \bar{x}_1\right)\left(x_{jp} - \bar{x}_p\right) \\ \vdots & \ddots & \vdots \\ \frac{1}{n}\sum_{j=1}^{n}\left(x_{j1} - \bar{x}_1\right)\left(x_{jp} - \bar{x}_p\right) & \cdots & \frac{1}{n}\sum_{j=1}^{n}\left(x_{jp} - \bar{x}_p\right)^2 \end{bmatrix}$$

50

Partitioning Sample Mean Vector

$$\bar{\mathbf{x}} = \begin{bmatrix} \bar{x}_1 \\ \vdots \\ \bar{x}_q \\ --- \\ \bar{x}_{q+1} \\ \vdots \\ \bar{x}_p \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{x}}^{(1)} \\ --- \\ \bar{\mathbf{x}}^{(2)} \end{bmatrix}$$

51

Partitioning Sample Covariance Matrix

$$\mathbf{S}_n = \begin{bmatrix} \mathbf{S}_{11} & | & \mathbf{S}_{12} \\ --- & | & --- \\ \mathbf{S}_{21} & | & \mathbf{S}_{22} \end{bmatrix}$$

$$= \begin{bmatrix} s_{11} & \cdots & s_{1q} & | & s_{1,q+1} & \cdots & s_{1p} \\ \vdots & \ddots & \vdots & | & \vdots & \ddots & \vdots \\ s_{q1} & \cdots & s_{qq} & | & s_{q,q+1} & \cdots & s_{qp} \\ -- & -- & -- & | & -- & -- & -- \\ s_{q+1,1} & \cdots & s_{q+1,q} & | & s_{q+1,q+1} & \cdots & s_{q+1,p} \\ \vdots & \ddots & \vdots & | & \vdots & \ddots & \vdots \\ s_{p1} & \cdots & s_{pq} & | & s_{p,q+1} & \cdots & s_{pp} \end{bmatrix}$$

$$\mathbf{S}_{21} = \mathbf{S}_{12}'$$

52

Outline

- Eigenvalues and Eigenvectors
- Positive Definite Matrices
- A Square-Root Matrix
- Random Vectors and Matrices
- Mean Vectors and Covariance Matrices
- Matrix Inequalities and Maximization

53

Questions

- What is the Cauchy-Schwarz inequality?
- How to prove the Cauchy-Schwarz inequality?
- What is the extended Cauchy-Schwarz inequality?
- How to prove the extended Cauchy-Schwarz inequality?

54

Questions

- What is the maximization lemma?
- How to maximize a quadratic form for points on the unit sphere?
- What is the Rayleigh's quotient?
- How to maximize and minimize the Rayleigh's quotient?

55

Cauchy-Schwarz Inequality

$(\mathbf{b}'\mathbf{d})^2 \leq (\mathbf{b}'\mathbf{b})(\mathbf{d}'\mathbf{d})$, with equality when $\mathbf{b} = c\mathbf{d}$

Proof :

$\mathbf{b}\text{-}x\mathbf{d} \neq 0$

$$0 < (\mathbf{b} - x\mathbf{d})'(\mathbf{b} - x\mathbf{d}) = \mathbf{b}'\mathbf{b} - 2x(\mathbf{b}'\mathbf{d}) + x^2(\mathbf{d}'\mathbf{d})$$

$$0 < \mathbf{b}'\mathbf{b} - \frac{(\mathbf{b}'\mathbf{d})^2}{\mathbf{d}'\mathbf{d}} + (\mathbf{d}'\mathbf{d})\left(x - \frac{\mathbf{b}'\mathbf{d}}{\mathbf{d}'\mathbf{d}}\right)^2$$

$$0 < \mathbf{b}'\mathbf{b} - \frac{(\mathbf{b}'\mathbf{d})^2}{\mathbf{d}'\mathbf{d}}$$

56

Extended Cauchy-Schwarz Inequality

$\mathbf{B}$ positive definite

$(\mathbf{b'd})^2 \le (\mathbf{b'Bb})(\mathbf{d'B}^{-1}\mathbf{d})$ with equality when $\mathbf{b} = c\mathbf{B}^{-1}\mathbf{d}$

Proof :

$$\mathbf{b'd} = \mathbf{bId} = \mathbf{b'B}^{1/2}\mathbf{B}^{-1/2}\mathbf{d} = (\mathbf{B}^{1/2'}\mathbf{b})'(\mathbf{B}^{-1/2}\mathbf{d})$$

$$((\mathbf{B}^{1/2}{}'\mathbf{b})'(\mathbf{B}^{-1/2}\mathbf{d}))^2 \le (\mathbf{b'B}^{1/2}\mathbf{B}^{1/2'}\mathbf{b})(\mathbf{d'B}^{-1/2}{}'\mathbf{B}^{-1/2}\mathbf{d})$$

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

$$\mathbf{B}^{1/2} = \sum_{i=1}^{p}\sqrt{\lambda_i}\mathbf{e}_i\mathbf{e}_i'$$

57

Maximization of Quadratic Forms for Points on the Unit Sphere

$\mathbf{B}$ positive definite matrix with eigen values

$\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_p > 0$ and associated eigenvectors

$\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_p \Rightarrow$

$$\max_{\mathbf{x}\ne 0}\frac{\mathbf{x'Bx}}{\mathbf{x'x}} = \lambda_1 \quad (\text{attained when } \mathbf{x} = \mathbf{e}_1)$$

$$\min_{\mathbf{x}\ne 0}\frac{\mathbf{x'Bx}}{\mathbf{x'x}} = \lambda_p \quad (\text{attained when } \mathbf{x} = \mathbf{e}_p)$$

$$\max_{\mathbf{x}\perp \mathbf{e}_1,\cdots,\mathbf{e}_k}\frac{\mathbf{x'Bx}}{\mathbf{x'x}} = \lambda_{k+1} \quad (\text{attained when } \mathbf{x} = \mathbf{e}_{k+1})$$

59

Maximization Lemma

$\mathbf{B}$ positive definite matrix, $\mathbf{d}$ given vector

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

maximum attained when $\mathbf{x} = c\mathbf{B}^{-1}\mathbf{d}$ for $c \ne 0$

Proof :

$$(\mathbf{x'd})^2 \le (\mathbf{x'Bx})(\mathbf{d'B}^{-1}\mathbf{d})$$

$$\mathbf{x'Bx} > 0$$

$$\frac{(\mathbf{x'd})^2}{\mathbf{x'Bx}} \le \mathbf{d'B}^{-1}\mathbf{d}$$

58

Maximization of Quadratic Forms for Points on the Unit Sphere

Proof :

$$\mathbf{B}^{1/2} = \mathbf{P\Lambda}^{1/2}\mathbf{P'}, \quad \mathbf{y} = \mathbf{P'x}$$

$$\frac{\mathbf{x'Bx}}{\mathbf{x'x}} = \frac{\mathbf{x'B}^{1/2}\mathbf{B}^{1/2}\mathbf{x}}{\mathbf{x'PP'x}} = \frac{\mathbf{x'P\Lambda}^{1/2}\mathbf{P'P\Lambda}^{1/2}\mathbf{P'x}}{\mathbf{y'y}}$$

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

$$\mathbf{x} = \mathbf{e}_1, \mathbf{y'} = (\mathbf{Pe})' = [1\,0\cdots 0], \frac{\mathbf{y'\Lambda y}}{\mathbf{y'y}} = \lambda_1 = \frac{\mathbf{e}_1'\mathbf{Be}_1}{\mathbf{e}_1'\mathbf{e}_1}$$

60

Maximization of Quadratic Forms for Points on the Unit Sphere

$$\mathbf{x} = \mathbf{P}\mathbf{y} = y_1\mathbf{e}_1 + y_2\mathbf{e}_2 + \cdots + y_p\mathbf{e}_p$$

$\mathbf{x} \perp \mathbf{e}_1, \cdots, \mathbf{e}_k$ implies

$$0 = \mathbf{e}_i'\mathbf{x} = y_1\mathbf{e}_i'\mathbf{e}_1 + y_2\mathbf{e}_i'\mathbf{e}_2 + \cdots + y_p\mathbf{e}_i'\mathbf{e}_p = y_i, i \le k$$

$$\frac{\mathbf{x}'\mathbf{B}\mathbf{x}}{\mathbf{x}'\mathbf{x}} = \frac{\sum_{i=k+1}^{p}\lambda_i y_i^2}{\sum_{i=k+1}^{p} y_i^2} \le \lambda_{k+1}$$

Taking $y_{k+1} = 1, y_{k+2} = \cdots = y_p = 0$ gives the asserted maximum

61

Maximization of Quadratic Forms for Points on the Unit Sphere

$$\mathbf{x} = \frac{\mathbf{x}_0}{\sqrt{\mathbf{x}_0'\mathbf{x}_0}} \Rightarrow \mathbf{x}'\mathbf{x} = 1, \mathbf{x}'\mathbf{B}\mathbf{x} = \frac{\mathbf{x}_0'\mathbf{B}\mathbf{x}_0}{\mathbf{x}_0'\mathbf{x}_0}$$

$$\max_{\mathbf{x}'\mathbf{x}=1} \mathbf{x}'\mathbf{B}\mathbf{x} = \max_{\mathbf{x}_0 \ne 0} \frac{\mathbf{x}_0'\mathbf{B}\mathbf{x}_0}{\mathbf{x}_0'\mathbf{x}_0} = \lambda_1$$

$$\min_{\mathbf{x}'\mathbf{x}=1} \mathbf{x}'\mathbf{B}\mathbf{x} = \min_{\mathbf{x}_0 \ne 0} \frac{\mathbf{x}_0'\mathbf{B}\mathbf{x}_0}{\mathbf{x}_0'\mathbf{x}_0} = \lambda_p$$

$$\max_{\mathbf{x}'\mathbf{x}=1, \mathbf{x}\perp\mathbf{e}_1,\cdots,\mathbf{e}_k} \mathbf{x}'\mathbf{B}\mathbf{x} = \max_{\mathbf{x}_0 \ne 0, \mathbf{x}_0\perp\mathbf{e}_1,\cdots,\mathbf{e}_k} \frac{\mathbf{x}_0'\mathbf{B}\mathbf{x}_0}{\mathbf{x}_0'\mathbf{x}_0} = \lambda_{k+1}$$

62

Rayleigh's Quotient

$$f(\mathbf{x}) = \frac{\mathbf{x}'\mathbf{B}\mathbf{x}}{\mathbf{x}'\mathbf{x}}$$

$$\mathbf{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix}'$$

$$\mathbf{B} = \begin{bmatrix} \frac{1}{a^2} & 0 \\ 0 & \frac{1}{b^2} \end{bmatrix}, \quad a > b, \quad \lambda_1 = \frac{1}{b^2}, \quad \lambda_2 = \frac{1}{a^2}$$

63

Rayleigh's Quotient in Polar Coordinates

$$\mathbf{x}'\mathbf{B}\mathbf{x} = \frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} = \rho^2\left(\frac{b^2\cos^2\theta + a^2\sin^2\theta}{a^2b^2}\right)$$

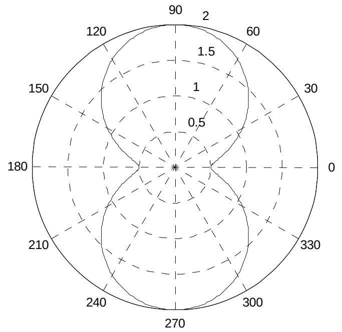
$$\mathbf{x}'\mathbf{x} = x_1^2 + x_2^2 = \rho^2$$

$$\frac{\mathbf{x}'\mathbf{B}\mathbf{x}}{\mathbf{x}'\mathbf{x}} = \frac{b^2\cos^2\theta + a^2\sin^2\theta}{a^2b^2}$$

$$= \left(\frac{1}{2a^2} + \frac{1}{2b^2}\right) + \left(\frac{1}{2a^2} - \frac{1}{2b^2}\right)\cos 2\theta$$

64

Square Root of Rayleigh's Quotient



$a = 2, \; b = 0.5$

Maximum and Minimum of Rayleigh's Quotient

$$a = 2, \quad b = 0.5$$

$$\max_{\mathbf{x} \neq 0} \frac{\mathbf{x}'\mathbf{B}\mathbf{x}}{\mathbf{x}'\mathbf{x}} = \lambda_1 = \frac{1}{b^2} = 4$$

$$\min_{\mathbf{x} \neq 0} \frac{\mathbf{x}'\mathbf{B}\mathbf{x}}{\mathbf{x}'\mathbf{x}} = \lambda_2 = \frac{1}{a^2} = 0.25$$