Chapter 2 Matrix Algebra and Random Vectors

wqcui@ustc

General Statistical Distance

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

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

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

General Statistical Distance

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

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

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

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

Quadratic Form

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

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

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

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

Statistical Distance under Rotated Coordinate System

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

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

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

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

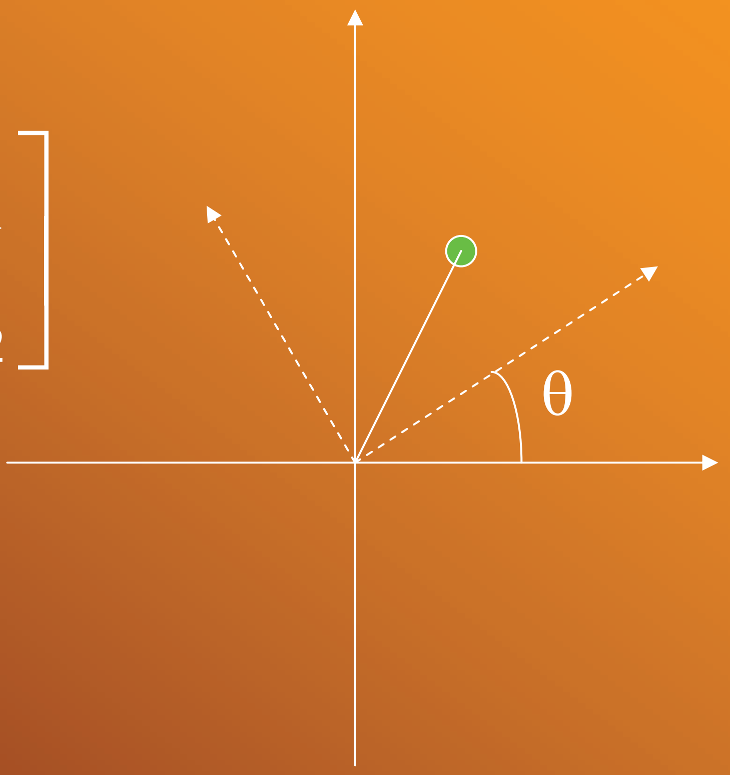
$$d(O,P) = \sqrt{a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2}$$

Rotated Coordinate System

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

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

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

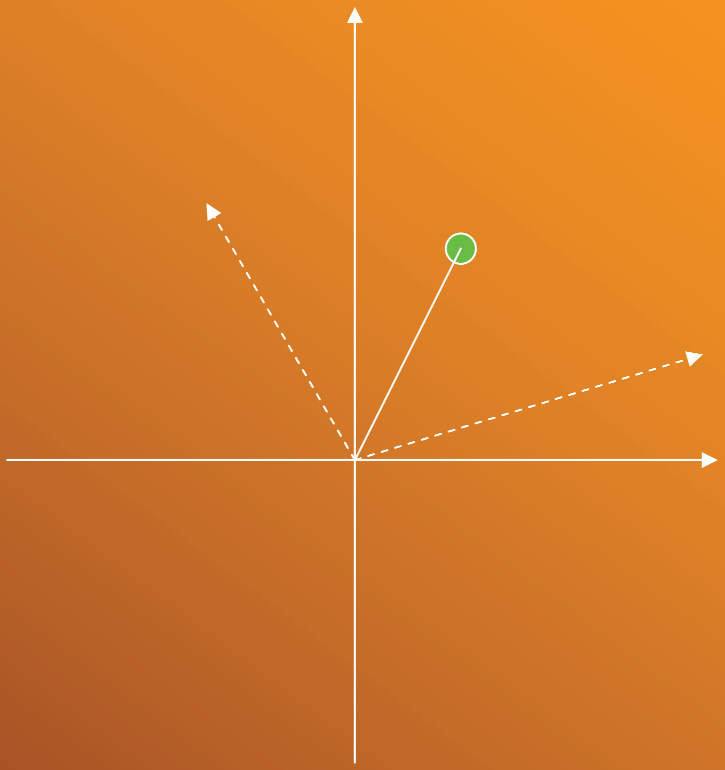


Coordinate Transformation

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

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

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



Quadratic Form in Transformed Coordinate Systems

$$Q(\mathbf{x}) = \mathbf{x}'\mathbf{A}\mathbf{x}$$

$$= (\mathbf{B}^{-1}\mathbf{y})'\mathbf{A}(\mathbf{B}^{-1}\mathbf{y})$$

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

$$= \mathbf{y}'\mathbf{\Lambda}\mathbf{y} = Q(\mathbf{y})$$

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

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

Diagonalized Quadratic Form

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

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

Orthogonal Matrix

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

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

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

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

Diagonalization

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

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

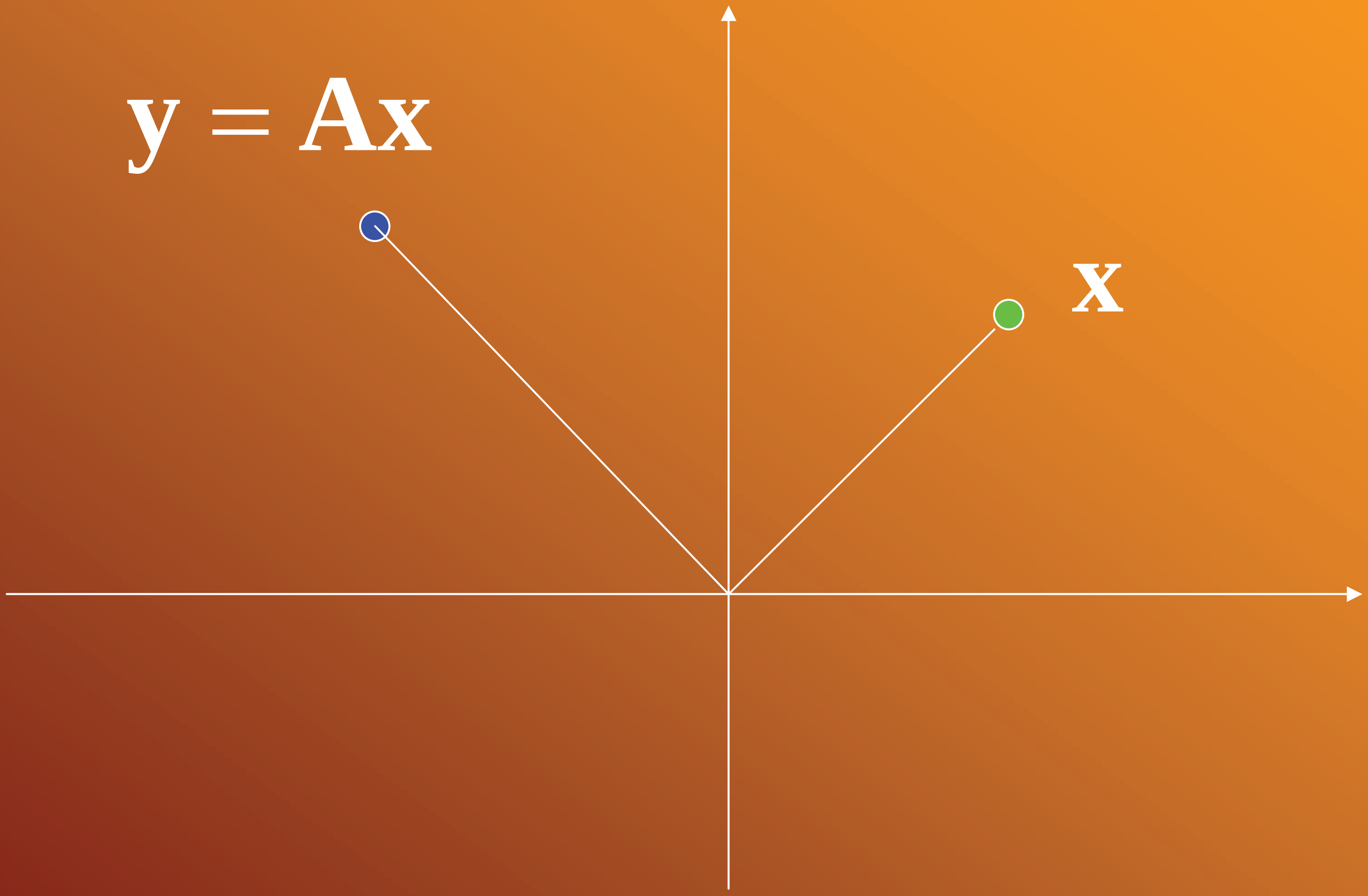
$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{bmatrix} \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_k \end{bmatrix} =$$

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

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

Concept of Mapping

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



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

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

$$\lambda_1 = 1$$

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

$$\lambda_2 = 3$$

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

Spectral Decomposition

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

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

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

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

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 $\mathbf{x}$

Positive Definite Matrix

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

$$= \mathbf{xAx}$$

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

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

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

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

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

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

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

Inverse and Square-Root Matrix

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

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

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

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

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

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

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

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

Random Vectors and Random Matrices

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

Expected Value of a Random Matrix

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

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

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

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 $f(x_i)$

$$\mu_i = E(X_i)$$

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

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

Covariance

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

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

$$= \sigma_{ik}$$

Statistically Independent

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

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

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

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

(the converse is not true in general)

Population Variance-Covariance Matrices

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

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

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

Population Correlation Coefficients

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

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

$$\rho_{ii} = 1$$

Standard Deviation Matrix

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

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

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

Correlation Matrix from Covariance Matrix

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

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

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

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

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

$$\boldsymbol{\Sigma} = E(\mathbf{X}-\boldsymbol{\mu})(\mathbf{X}-\boldsymbol{\mu})' = \left[\begin{array}{c|c} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ --- & --- \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{array}\right]$$

$$= \left[\begin{array}{ccc|ccc} \sigma_{11} & \cdots & \sigma_{1q} & \sigma_{1,q+1} & \cdots & \sigma_{1p} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \sigma_{q1} & \cdots & \sigma_{qq} & \sigma_{q,q+1} & \cdots & \sigma_{qp} \\ -- & -- & -- & -- & -- & -- \\ \sigma_{q+1,1} & \cdots & \sigma_{q+1,q} & \sigma_{q+1,q+1} & \cdots & \sigma_{q+1,p} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \cdots & \sigma_{pq} & \sigma_{p,q+1} & \cdots & \sigma_{pp} \end{array}\right]$$

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

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

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

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

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

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

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

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

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

Linear Combinations of Random Variables

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

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

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

Sample Mean Vector and Covariance Matrix

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

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

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

Partitioning Sample Mean Vector

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

Partitioning Sample Covariance Matrix

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

$$= \left[\begin{array}{ccc|ccc} s_{11} & \cdots & s_{1q} & s_{1,q+1} & \cdots & s_{1p} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ s_{q1} & \cdots & s_{qq} & s_{q,q+1} & \cdots & s_{qp} \\ -- & -- & -- & -- & -- & -- \\ s_{q+1,1} & \cdots & s_{q+1,q} & s_{q+1,q+1} & \cdots & s_{q+1,p} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ s_{p1} & \cdots & s_{pq} & s_{p,q+1} & \cdots & s_{pp} \end{array}\right]$$

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

Cauchy-Schwarz Inequality

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

Proof :

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

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

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

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

Extended Cauchy-Schwarz Inequality

$\mathbf{B}$ positive definite

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

Proof :

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

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

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

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

Maximization Lemma

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

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

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

Proof :

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

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

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

Maximization of Quadratic Forms for Points on the Unit Sphere

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

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

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

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

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

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

Maximization of Quadratic Forms for Points on the Unit Sphere

Proof :

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

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

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

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

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

Maximization of Quadratic Forms for Points on the Unit Sphere

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

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

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

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