# Lecture 3

## A Short Review of Matrix Algebra

Reading: Zelterman, 2015 Chapter 4; Izenman, 2008 Chapter 3.1-3.3

*DSA 8070 Multivariate Analysis*
September 5 - September 9, 2022

Whitney Huang
Clemson University





Notes

### Agenda

1. Motivation
2. Basic Matrix Concepts
3. Some Useful Matrix Tools/Facts





Notes

### Why Matrix Algebra?

**Data:**

$$\boldsymbol{X} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \cdots & \cdots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix}$$

**Summary Statistics:**

$\bar{X} = \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \vdots \\ \bar{x}_p \end{bmatrix} = \begin{bmatrix} \frac{1}{n}\sum_{i=1}^n x_{i1} \\ \frac{1}{n}\sum_{i=1}^n x_{i2} \\ \vdots \\ \frac{1}{n}\sum_{i=1}^n x_{ip} \end{bmatrix} = \frac{1}{n}\boldsymbol{X}^T\boldsymbol{1}$ is the sample mean vector,

and $\boldsymbol{S} = \begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1p} \\ s_{21} & s_{22} & \cdots & s_{2p} \\ \vdots & \cdots & \cdots & \cdots \\ s_{p1} & s_{p2} & \cdots & s_{pp} \end{bmatrix} = \frac{1}{n-1}\boldsymbol{X}^T(I - \frac{1}{n}\boldsymbol{1}\boldsymbol{1}^T)\boldsymbol{X}$ is the sample covariance matrix. Many matrix algebra techniques will be applied to this matrix in multivariate analysis





Notes

## Covariance Matrices

- Covariance Matrix

$$\boldsymbol{\Sigma} = \underbrace{\begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_{pp} \end{bmatrix}}_{\textbf{population}\text{ covariance matrix}}, \quad \boldsymbol{S} = \underbrace{\begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1p} \\ s_{21} & s_{22} & \cdots & s_{2p} \\ \vdots & \cdots & \cdots & \cdots \\ s_{p1} & s_{p2} & \cdots & s_{pp} \end{bmatrix}}_{\textbf{sample}\text{ covariance matrix}}$$

- Since $\sigma_{jk} = \sigma_{kj}$ (likewise $s_{jk} = s_{kj}$) for all $j \neq k \Rightarrow \boldsymbol{\Sigma}$ and $\boldsymbol{S}$ are symmetric
- $\boldsymbol{\Sigma}$ and $\boldsymbol{S}$ are also non-negative definite





Notes

## Vectors

- A column array of $p$ elements is called a vector of dimension $p$ and is written as

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix}$$

- The transpose of the column vector $\boldsymbol{x}$ is a row vector

$$\boldsymbol{x}^T = [x_1 \quad x_2 \quad \cdots \quad x_p]$$

- $L_{\boldsymbol{x}}^{-1}\boldsymbol{x}$, where $L_{\boldsymbol{x}} = \sqrt{\sum_{j=1}^{p} x_j^2}$, is called a unit vector





Notes

## Matrices

- A matrix $A$ is an array of elements $a_{ij}$ with $n$ rows and $p$ columns:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{bmatrix}$$

- The transpose $A^T$ has $p$ rows and $n$ columns. The $j$-th row of $A^T$ os the $j$-th column of $A$

$$A^T = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1p} & a_{2p} & \cdots & a_{np} \end{bmatrix}$$





Notes

## Identity Matrix and Inverse Matrix

- An identity matrix, denoted by $I$, is a square matrix with $1$'s along the diagonal and $0$'s everywhere else. For example

$$I_{3\times 3} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

- Consider two square matrices $A$ and $B$ with the same dimension. If

$$AB = BA = I,$$

then $B$ is the inverse of $A$, denoted by $A^{-1}$





Notes

## Orthogonal Matrices

- A square matrix $Q$ is orthogonal if

$$QQ^T = Q^TQ = I$$

- If $Q$ is orthogonal, its rows and columns have unit length (i.e., $L_{\boldsymbol{q}_j} = 1$) and are mutually perpendicular (i.e., $\boldsymbol{q}_j^T\boldsymbol{q}_k = 0$ for any $j \neq k$)

- **Example:**

$$Q = \frac{1}{3}\begin{bmatrix} 2 & -2 & 1 \\ 1 & 2 & 2 \\ 2 & 1 & -2 \end{bmatrix}$$





Notes

## Eigenvalues and Eigenvectors

- A square matrix $A$ has an eigenvalue $\lambda$ with corresponding eigenvector $\boldsymbol{x} \neq 0$ if

$$A\boldsymbol{x} = \lambda\boldsymbol{x}.$$

The eigenvalues of $A$ are the solution to $|A - \lambda I| = 0$

- A normalized eigenvector is denoted by $\boldsymbol{e}$ with $\boldsymbol{e}^T\boldsymbol{e} = 1$

- A $p \times p$ matrix $A$ has $p$ pairs of eigenvalues and eigenvectors

$$\lambda_1, \boldsymbol{e}_1 \quad \lambda_2, \boldsymbol{e}_2 \quad \cdots \quad \lambda_p, \boldsymbol{e}_p$$





Notes

## Spectral Decomposition



- Eigenvalues and eigenvectors will play an important role in DSA 8070. For example, principal components are based on the eigenvalues and eigenvectors of sample covariance matrices
- The spectral decomposition of a $p \times p$ symmetric matrix $A$ is $A = \lambda_1 \boldsymbol{e}_1 \boldsymbol{e}_1^T + \lambda_2 \boldsymbol{e}_2 \boldsymbol{e}_2^T + \cdots + \lambda_p \boldsymbol{e}_p \boldsymbol{e}_p^T$. This can be written in the following matrix form:

$$\underbrace{\begin{bmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \cdots & \boldsymbol{e}_p \end{bmatrix}}_{P} \underbrace{\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_p \end{bmatrix}}_{\Lambda} \underbrace{\begin{bmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \cdots & \boldsymbol{e}_p \end{bmatrix}^T}_{P^T}$$



Notes

## Determinant and Trace



- The trace if a $p \times p$ matrix $A$ is the sum of the diagonal elements, i.e., $\text{trace}(A) = \sum_{i=1}^{p} a_{ii}$
- The trace of a square, symmetric matrix $A$ is the sum of the eigenvalues, i.e., $\text{trace}(A) = \sum_{i=1}^{p} a_{ii} = \sum_{i=1}^{p} \lambda_i$
- The determinant of a square, symmetric matrix $A$ is the product of the eigenvalues, i.e., $|A| = \prod_{i=1}^{p} \lambda_i$



Notes

## Positive Definite Matrix



- For a $p \times p$ symmetric matrix $A$ and a vector $\boldsymbol{x} = \begin{bmatrix} x_1 & x_2 & \cdots & x_p \end{bmatrix}^T$ the quantity $\boldsymbol{x}^T A \boldsymbol{x} = \sum_{i=1}^{p} \sum_{j=1}^{p} a_{ij} x_i x_j$ is called a quadratic form
- If $\boldsymbol{x}^T A \boldsymbol{x} \geq 0$ for any vector $\boldsymbol{x}$, both $A$ and the quadratic form are said to be non-negative definite

  $\Rightarrow$ all the eigenvalues of $A$ are non-negative
- If $\boldsymbol{x}^T A \boldsymbol{x} > 0$ for any vector $\boldsymbol{x} \neq \boldsymbol{0}$, both $A$ and the quadratic form are said to be positive definite

  $\Rightarrow$ all the eigenvalues of $A$ are positive



Notes

## Square-Root Matrices

- Spectral decomposition of a positive definite matrix $A$ yields

$$A = \sum_{j=1}^{p} \lambda_j \boldsymbol{e}_j \boldsymbol{e}_j^T = P\Lambda P^T,$$

with $\Lambda_{p\times p} = \mathrm{diag}(\lambda_j)$, all $\lambda_j > 0$, and $P_{p\times p} = \begin{bmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \cdots & \boldsymbol{e}_p \end{bmatrix}$ an orthonormal matrix of eigenvectors. Then

$$A^{-1} = P\Lambda^{-1}P^T = \sum_{j=1}^{p} \frac{1}{\lambda_j} \boldsymbol{e}_j \boldsymbol{e}_j^T$$

- With $\Lambda^{\frac{1}{2}} = \mathrm{diag}(\lambda_j^{\frac{1}{2}})$, a square-root matrix is

$$A^{\frac{1}{2}} = P\Lambda^{\frac{1}{2}}P^T = \sum_{j=1}^{p} \sqrt{\lambda_j}\, \boldsymbol{e}_j \boldsymbol{e}_j^T$$





Notes

## Partitioning Random vectors

- If we partition the $p \times 1$ random vector $\boldsymbol{X}$ into two components $\boldsymbol{X}_1, \boldsymbol{X}_2$ of dimensions $q \times 1$ and $(p-q) \times 1$ respectively, then the mean vector and the variance-covariance matrix need to be partitioned accordingly

- Partitioned mean vector:

$$\mathbb{E}[\boldsymbol{X}] = \mathbb{E}\begin{bmatrix} \boldsymbol{X}_1 \\ \boldsymbol{X}_2 \end{bmatrix} = \begin{bmatrix} \mathbb{E}[\boldsymbol{X}_1] \\ \mathbb{E}[\boldsymbol{X}_2] \end{bmatrix} = \begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix}$$

- Partitioned covariance matrix:

$$\boldsymbol{\Sigma} = \begin{bmatrix} \mathrm{Var}(\boldsymbol{X}_1) & \mathrm{Cov}(\boldsymbol{X}_1, \boldsymbol{X}_2) \\ \mathrm{Cov}(\boldsymbol{X}_2, \boldsymbol{X}_1) & \mathrm{Var}(\boldsymbol{X}_2) \end{bmatrix} = \begin{bmatrix} \underbrace{\boldsymbol{\Sigma}_{11}}_{q\times q} & \underbrace{\boldsymbol{\Sigma}_{12}}_{q\times(p-q)} \\ \underbrace{\boldsymbol{\Sigma}_{21}}_{(p-q)\times q} & \underbrace{\boldsymbol{\Sigma}_{22}}_{(p-q)\times(p-q)} \end{bmatrix}$$

A Short Review of Matrix Algebra
CLEMSON UNIVERSITY
Motivation
Basic Matrix Concepts
Some Useful Matrix Tools/Facts



Notes

Notes