UCDAVIS
DEPARTMENT OF STATISTICS

# STA 108 Applied Statistical Methods: Regression Analysis

## Linear Algebra

Shizhe Chen, PhD

Spring 2020

You are expected to read the course notes **before** lectures.

# Linear Algebra

**Why linear algebra?**

Crucial knowledge in the era of data science

Prerequisite for understanding multiple linear regression

**Why now?**

R's syntax uses vectors and matrices

**What to take away from this?**

Look out for the image in the lower right corner, meaning "not required"[1]

[1]... but will be investigated in R

# Basic Definitions

# Vectors and Matrices

An $n$-dimensional **vector** is: $\mathbf{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in \mathbb{R}^n$

An $n \times p$ **matrix** is: $\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{pmatrix} \in \mathbb{R}^{n \times p}$

The $j$th column of a matrix is written as $\mathbf{a}_j$ and the $(i, j)$th element of a matrix is written as $a_{ij}$.

By convention, vectors are column vectors unless specified otherwise.

# Vector and Matrix Tranpose

We will denote the **transpose** of a vector or matrix using the symbol ($'$ or $^{\mathrm{T}}$).

The **transpose** of a vector turns a column vector into a row vector

$$\mathbf{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}_{n \times 1} \qquad \mathbf{a}^{\mathrm{T}} = \begin{pmatrix} a_1 & \cdots & a_n \end{pmatrix}_{1 \times n}$$

The **transpose** of a matrix turns the columns into rows

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{pmatrix}_{n \times p} \qquad \mathbf{A}^{\mathrm{T}} = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1p} & a_{2p} & \cdots & a_{np} \end{pmatrix}_{p \times n}$$

# Vector and Matrix Transpose: Examples

What is the transpose of $\mathbf{a} = \begin{pmatrix} 1 \\ 3 \\ 5 \\ 7 \end{pmatrix}$?

What is the transpose of $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \\ 7 & 8 \end{pmatrix}$?

# Vector and Matrix Transpose: Examples

What is the transpose of $\mathbf{a} = \begin{pmatrix} 1 \\ 3 \\ 5 \\ 7 \end{pmatrix}$?

What is the transpose of $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \\ 7 & 8 \end{pmatrix}$?

Answers:

$$\mathbf{a}^{\mathrm{T}} = \begin{pmatrix} 1 & 3 & 5 & 7 \end{pmatrix}, \qquad \mathbf{A}^{\mathrm{T}} = \begin{pmatrix} 1 & 3 & 5 & 7 \\ 2 & 4 & 6 & 8 \end{pmatrix}$$

# Matrix Transpose: Useful Properties

Here are some useful properties of matrix transpose:

- $(\mathbf{A}^{\mathrm{T}})^{\mathrm{T}} = \mathbf{A}$
- $(\mathbf{A} + \mathbf{B})^{\mathrm{T}} = \mathbf{A}^{\mathrm{T}} + \mathbf{B}^{\mathrm{T}}$ ($\mathbf{A} + \mathbf{B}$ is matrix addition)
- $(\mathbf{AB})^{\mathrm{T}} = \mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}$ ($\mathbf{AB}$ is matrix multiplication)
- $(\mathbf{A}^{-1})^{\mathrm{T}} = (\mathbf{A}^{\mathrm{T}})^{-1}$ ($\mathbf{A}^{-1}$ is matrix inverse)
- $(c\mathbf{A})^{\mathrm{T}} = c\mathbf{A}^{\mathrm{T}}$ ($c\mathbf{A}$ is scalar multiplication)

# Diagonal and Identity Matrices:

A **diagonal matrix** is a square matrix that has zeros in the off-diagonals:

$$\begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & a_n \end{pmatrix}_{n\times n}$$

In this course, we also write $\mathrm{diag}(a_1, \ldots, a_n)$ to denote a diagonal matrix

The **identity matrix** is a special type of diagonal matrix with ones on the diagonal, $\mathbf{I}_n = \mathrm{diag}(1, \ldots, 1)$

# Symmetric Matrix:

A **symmetric matrix** is square and symmetric along the diagonals elements

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pp} \end{pmatrix}$$

with $a_{ij} = a_{ji}$ for all $i \neq j$.

The transpose of a symmetric matrix is by definition itself, i.e., $\mathbf{A} = \mathbf{A}^{\mathrm{T}}$.

Examples: are these matrices symmetric?

$$\begin{pmatrix} 1 & 5 & 8 \\ 5 & 3 & 6 \\ 8 & 6 & 2 \end{pmatrix} \qquad \begin{pmatrix} 1 & 5 & 8 \\ 5 & 3 & 6 \\ 3 & 6 & 2 \end{pmatrix}$$

# Matrix Trace:

The **trace** of a square matrix

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}_{n \times n}$$

is $\mathrm{tr}(\mathbf{A}) = \sum_{i=1}^{n} a_{ii}$, i.e., the sum of the diagonal elements.

Example:

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 5 & 3 & 5 \\ 2 & 6 & 1 \end{pmatrix} \qquad \mathrm{tr}(\mathbf{A}) = 1 + 3 + 1 = 5.$$

# Matrix Trace: Useful Properties

Here are some useful properties of matrix trace:

- $\mathrm{tr}(\mathbf{A}) = \mathrm{tr}(\mathbf{A}^{\mathrm{T}})$
- $\mathrm{tr}(\mathbf{A} + \mathbf{B}) = \mathrm{tr}(\mathbf{A}) + \mathrm{tr}(\mathbf{B})$
- $\mathrm{tr}(\mathbf{AB}) = \mathrm{tr}(\mathbf{BA})$
- $\mathrm{tr}(c\mathbf{A}) = c \cdot \mathrm{tr}(\mathbf{A})$
- If $\mathbf{A}$ is symmetric, $\mathrm{tr}(\mathbf{A}) = \sum_{j=1}^{n} \lambda_j$, where $\lambda_j$ is the $j$th eigenvalue of $\mathbf{A}$.

# Matrix Calculations

# Matrix Equality

For two $n \times p$ matrices $\mathbf{A}$ and $\mathbf{B}$, we say that $\mathbf{A}$ is equal to $\mathbf{B}$ $(\mathbf{A} = \mathbf{B})$ if and only if $a_{ij} = b_{ij}$ for all $i, j$.

**Example:**

$$\text{if } \mathbf{A} = \begin{pmatrix} 2 & 8 & 9 \\ 3 & 2 & 3 \\ 3 & 8 & 4 \end{pmatrix}, \qquad \mathbf{B} = \begin{pmatrix} 2 & 8 & 9 \\ 3 & 2 & 3 \\ 3 & 8 & 4 \end{pmatrix}, \text{ then } \mathbf{A} = \mathbf{B}$$

$$\text{if } \mathbf{A} = \begin{pmatrix} 2 & 8 & 9 \\ 3 & 2 & 3 \\ 3 & 8 & 4 \end{pmatrix}, \qquad \mathbf{B} = \begin{pmatrix} 2 & 8 & 9 \\ 9 & 2 & 3 \\ 1 & 8 & 4 \end{pmatrix}, \text{ then } \mathbf{A} \neq \mathbf{B}$$

# Matrix Addition

Given two matrices $\mathbf{A} = \{a_{ij}\}_{n\times p}$ and $\mathbf{B} = \{b_{ij}\}_{n\times p}$ of the **same dimensions**, the **addition** $\mathbf{A}+\mathbf{B}$ produces $\mathbf{C} = \{c_{ij}\}_{n\times p}$ such that $c_{ij} = a_{ij} + b_{ij}$.

**Example:**

$$\text{Given } \mathbf{A} = \begin{pmatrix} 2 & 8 & 9 \\ 3 & 2 & 3 \end{pmatrix} \quad \text{and} \quad \mathbf{B} = \begin{pmatrix} 1 & 3 & 5 \\ 4 & 7 & 1 \end{pmatrix}, \text{ we have}$$

$$\mathbf{A}+\mathbf{B} = \begin{pmatrix} 2+1 & 8+3 & 9+5 \\ 3+4 & 2+7 & 3+1 \end{pmatrix} = \begin{pmatrix} 3 & 11 & 14 \\ 7 & 9 & 4 \end{pmatrix}$$

# Matrix Subtration

Given two matrices $\mathbf{A} = \{a_{ij}\}_{n \times p}$ and $\mathbf{B} = \{b_{ij}\}_{n \times p}$ of the **same dimensions**, the **subtraction** $\mathbf{A} - \mathbf{B}$ produces $\mathbf{C} = \{c_{ij}\}_{n \times p}$ such that $c_{ij} = a_{ij} - b_{ij}$.

**Example:**

$$\text{Given } \mathbf{A} = \begin{pmatrix} 2 & 8 & 9 \\ 3 & 2 & 3 \end{pmatrix} \quad \text{and} \quad \mathbf{B} = \begin{pmatrix} 1 & 3 & 5 \\ 4 & 7 & 1 \end{pmatrix}, \text{ we have}$$

$$\mathbf{A} - \mathbf{B} = \begin{pmatrix} 2-1 & 8-3 & 9-5 \\ 3-4 & 2-7 & 3-1 \end{pmatrix} = \begin{pmatrix} 1 & 5 & 4 \\ -1 & -5 & 2 \end{pmatrix}$$

# Inner Product between Two Vectors

The **inner product** of two $n$-dimensional vectors $\mathbf{x} = (x_1, \ldots, x_n)^{\mathrm{T}}$ and $\mathbf{y} = (y_1, \ldots, y_n)^{\mathrm{T}}$ is

$$\mathbf{x}^{\mathrm{T}}\mathbf{y} = (x_1 \ x_2 \ \cdots \ x_n) \cdot \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \sum_{i=1}^{n} x_i y_i$$

**Example:**

$$\text{Given } \mathbf{x} = (1 \ 3 \ -4)^{\mathrm{T}} \quad \text{and} \quad \mathbf{y} = (3 \ -2 \ 1)^{\mathrm{T}}, \text{ we have}$$

$$\begin{aligned} \mathbf{x}^T\mathbf{y} &= 1 \cdot 3 + 3 \cdot (-2) + (-4) \cdot 1 \\ &= 3 - 6 - 4 = -7 \end{aligned}$$

**Note:** $\mathbf{x}, \mathbf{y}$ must have the same length.

# Outer Product between Two Vectors

The **outer product** of two vectors $\mathbf{x} = (x_1, \ldots, x_n)^{\mathrm{T}}$ and $\mathbf{y} = (y_m, \ldots, y_m)^{\mathrm{T}}$ is

$$\mathbf{x}\mathbf{y}^{\mathrm{T}} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \cdot (y_1 \; y_2 \; \cdots \; y_m) = \begin{pmatrix} x_1y_1 & x_1y_2 & \cdots & x_1y_m \\ x_2y_1 & x_2y_2 & \cdots & x_2y_m \\ \vdots & \vdots & \ddots & \vdots \\ x_ny_1 & x_ny_2 & \cdots & x_ny_m \end{pmatrix}_{n\times m}$$

**Example:**

$$\text{Given } \mathbf{x} = (1\ 3)^{\mathrm{T}} \quad \text{and} \quad \mathbf{y} = (3\ \ -2\ 1)^{\mathrm{T}}, \text{ we have}$$

$$\mathbf{x}\mathbf{y}^{\mathrm{T}} = \begin{pmatrix} 1\cdot 3 & 1\cdot(-2) & 1\cdot 1 \\ 3\cdot 3 & 3\cdot(-2) & 3\cdot 1 \end{pmatrix} = \begin{pmatrix} 3 & -2 & 1 \\ 9 & -6 & 3 \end{pmatrix}$$

**Note:** $\mathbf{x}, \mathbf{y}$ can have different lengths.

# Matrix-Scalar Product

The **matrix-scalar product** of $\mathbf{A} = \{a_{ij}\}_{n \times p}$ and $b \in \mathbb{R}$ is the matrix $\mathbf{C} = \{c_{ij}\}_{n \times p}$ such that $c_{ij} = ba_{ij}$

$$b \cdot \mathbf{A} = \mathbf{A} \cdot b = \begin{pmatrix} ba_{11} & ba_{12} & \cdots & ba_{1p} \\ ba_{21} & ba_{22} & \cdots & ba_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ ba_{n1} & ba_{n2} & \cdots & ba_{np} \end{pmatrix}_{n \times p}$$

**Example:**

$$\text{Given } \mathbf{A} = \begin{pmatrix} 1 & -3 \\ 4 & 2 \end{pmatrix} \quad \text{and} \quad b = 3, \text{ we have}$$

$$b \cdot \mathbf{A} = \begin{pmatrix} 3 \cdot 1 & 3 \cdot (-3) \\ 3 \cdot 4 & 3 \cdot 2 \end{pmatrix} = \begin{pmatrix} 3 & -9 \\ 12 & 6 \end{pmatrix}$$

# Matrix-Vector Product

The **matrix-vector product** of $\mathbf{A} = \{a_{ij}\}_{n\times p}$ and $\mathbf{x} = (x_1, \ldots, x_p)^{\mathrm{T}}$ is

$$\mathbf{Ax} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{pmatrix}_{n\times p} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix}_{p\times 1} = \begin{pmatrix} \sum_{j=1}^{p} a_{1j}x_j \\ \sum_{j=1}^{p} a_{2j}x_j \\ \vdots \\ \sum_{j=1}^{p} a_{nj}x_j \end{pmatrix}_{n\times 1}$$

**Example:**

$$\text{Given } \mathbf{A} = \begin{pmatrix} 1 & -3 \\ 4 & 2 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad \mathbf{x} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \text{ we have}$$

$$\mathbf{Ax} = \begin{pmatrix} 1 \cdot 3 + (-3) \cdot 1 \\ 4 \cdot 3 + 2 \cdot 1 \\ 0 \cdot 3 + 1 \cdot 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 14 \\ 1 \end{pmatrix}$$

**Note:** The length of $\mathbf{x}$ must match the number of columns of $\mathbf{A}$

# Matrix-Matrix Product

The **matrix-matrix product** of $\mathbf{A} = \{a_{ij}\}_{n\times p}$ and $\mathbf{B} = \{b_{ij}\}_{p\times m}$ is

$$\mathbf{AB} = \begin{pmatrix} a_{11} & \cdots & a_{1p} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{np} \end{pmatrix}_{n\times p} \begin{pmatrix} b_{11} & \cdots & b_{1m} \\ \vdots & \ddots & \vdots \\ b_{p1} & \cdots & b_{pm} \end{pmatrix}_{p\times m}$$

$$= \begin{pmatrix} \sum_{j=1}^{p} a_{1j}b_{j1} & \sum_{j=1}^{p} a_{1j}b_{j2} & \cdots & \sum_{j=1}^{p} a_{1j}b_{jm} \\ \sum_{j=1}^{p} a_{2j}b_{j1} & \sum_{j=1}^{p} a_{2j}b_{j2} & \cdots & \sum_{j=1}^{p} a_{2j}b_{jm} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{j=1}^{p} a_{nj}b_{j1} & \sum_{j=1}^{p} a_{nj}b_{j2} & \cdots & \sum_{j=1}^{p} a_{nj}b_{jm} \end{pmatrix}_{n\times m}$$

**Note:** The number of rows of $\mathbf{B}$ (i.e., p) must match the number of columns of $\mathbf{A}$ (i.e., p)

# Matrix-Matrix Product Example

Given $\mathbf{A} = \begin{pmatrix} 1 & 3 & 4 \\ 4 & 7 & 5 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 1 & 6 \end{pmatrix}$, we have

$$\begin{aligned} \mathbf{AB} &= \begin{pmatrix} 1 & 3 & 4 \\ 4 & 7 & 5 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 1 & 6 \end{pmatrix} \\ &= \begin{pmatrix} 1 \cdot 1 + 3 \cdot 3 + 4 \cdot 1 & 1 \cdot 2 + 3 \cdot 4 + 4 \cdot 6 \\ 4 \cdot 1 + 7 \cdot 3 + 5 \cdot 1 & 4 \cdot 2 + 7 \cdot 4 + 5 \cdot 6 \end{pmatrix} \\ &= \begin{pmatrix} 14 & 38 \\ 30 & 66 \end{pmatrix} \end{aligned}$$

# Typical Mistakes in Matrix-Matrix Product

In general, for $\mathbf{A}$ and $\mathbf{B}$ that have the same dimensions:

- $\mathbf{AB} \neq \mathbf{BA}$
- $\mathbf{AB} = \mathbf{CB}$ generally **DOES NOT** imply that $\mathbf{A} = \mathbf{C}$

# Properties of Matrices

# Matrix Inverse

A square (not necessarily symmetric) matrix $\mathbf{A} = \{a_{ij}\}_{n \times n}$ is **invertible** (or **non-singular**) if there exists a matrix $\mathbf{A}^{-1}$ such that

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}_{n \times n}$$

A matrix $\mathbf{A} = \{a_{ij}\}_{n \times n}$ is **invertible** if and only if it has **full rank**, i.e., $\text{rank}(\mathbf{A}) = n$

If $\mathbf{A}$ and $\mathbf{B}$ are invertible, then $(\mathbf{A}\mathbf{B})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$ (**why?**)

# Matrix Inverse for $2 \times 2$ Case

**Claim:** For a $2 \times 2$ matrix $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, the matrix inverse is $\mathbf{A}^{-1} = \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$

**Proof:** Show $\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I}_{2\times 2}$.

# Matrix Inverse: Example

Given $\mathbf{A} = \begin{pmatrix} 1 & 3 \\ 2 & 1 \end{pmatrix}$, the inverse is $\begin{pmatrix} -1/5 & 3/5 \\ 2/5 & -1/5 \end{pmatrix}$:

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

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

# Example: Multiple Linear Regression[2]

$$Y = \mathbf{X}\beta + \epsilon$$

$$\hat{\beta} = (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}Y$$

[2] Not required before learning multiple linear regression

# Projection Matrix[3]

**Definition**: A **square** matrix $\mathbf{P}$ is a projection matrix if and only if $\mathbf{P}^2 = \mathbf{P}$ (idempotent).

- $(\mathbf{I} - \mathbf{P})(\mathbf{I} - \mathbf{P}) = (\mathbf{I} - \mathbf{P})$ and $\mathbf{PP} = \mathbf{P}$.
- $\mathbf{P}(\mathbf{I} - \mathbf{P}) = \mathbf{0}$.

**Example:**

In $\hat{Y} = \mathbf{X}\hat{\beta} = \mathbf{X}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}Y$, $\mathbf{P} = \mathbf{X}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}$ is a projection matrix ( $\hat{Y} = \mathbf{P}Y$).

- $\mathbf{P}$ is the projection matrix onto $\mathcal{R}(\mathbf{X})$ (column space of $\mathbf{X}$).
- $\mathbf{I} - \mathbf{P}$ is the projection matrix onto $\mathcal{R}(\mathbf{X})^{\perp}$.
- $\mathbf{PX} = \mathbf{X}$
- $(\mathbf{I} - \mathbf{P})\mathbf{X} = \mathbf{0}$

[3] Not required before learning multiple linear regression

# Rank of a Matrix

**Linear independence:** vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$ are linearly independent if $\sum_{i=1}^n c_i \mathbf{a}_i \neq 0$ unless $c_i = 0$ for all $i$.

The **rank** of a matrix $\mathbf{A} = \{a_{ij}\}_{n \times p}$ is a number of linearly independent rows/columns

- **column rank**: of $\mathbf{A}$ is the number of linearly independent columns
- **row rank**: of $\mathbf{A}$ is the number of linearly independent rows

We say that $\mathbf{A} = \{a_{ij}\}_{n \times p}$ is full rank if $\mathrm{rank}(\mathbf{A}) = \min(n, p)$.

- If $n < p$, **full rank** implies **full row rank**, i.e., $\mathrm{rank}(\mathbf{A}) = n$
- If $n > p$, **full rank** implies **full column rank**, i.e., $\mathrm{rank}(\mathbf{A}) = p$

# Examples:

What is the **rank** of

- the matrix $\begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}$?
- the matrix $\begin{pmatrix} 1 & 2 & 1 \\ 1 & 2 & 1 \\ 2 & 3 & 1 \\ 0 & 1 & 1 \end{pmatrix}$?

# Length and Orthogonality

**Length of a vector:** the length of a vector $\mathbf{x}$ is measured by its $\ell_2$ norm, i.e., $\|\mathbf{x}\| = \sqrt{\mathbf{x}^{\mathrm{T}}\mathbf{x}}$

**Orthogonal vectors:** two $n$-dimensional vectors $\mathbf{x}$ and $\mathbf{y}$ are orthogonal if $\mathbf{x}^{\mathrm{T}}\mathbf{y} = 0$.

**Orthogonal matrix:** a matrix $\mathbf{A} = \{a_{ij}\}_{n\times n}$ is orthogonal if its columns are orthogonal with unit norm. If $\mathbf{A}$ is orthogonal, then $\mathbf{A}^{\mathrm{T}}\mathbf{A} = \mathbf{I}_{n\times n}$.

If $\mathbf{A}$ is **square and orthogonal**, then $\mathbf{A}^{\mathrm{T}} = \mathbf{A}^{-1}$.

# Matrix Determinant

The **determinant** of a square matrix $\mathbf{A} = \{a_{ij}\}_{n\times n}$ is a real-valued function from $\mathbb{R}^{n\times n} \to \mathbb{R}$, and is denoted as $|\mathbf{A}|$ or $\det(\mathbf{A})$.

The determinant can be calculated using a recursive formula.

For a $2 \times 2$ matrix $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, $|\mathbf{A}| = ad - bc$.

For a $3 \times 3$ matrix $\mathbf{A} = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$, show that

$$|\mathbf{A}| = aei + bfg + cdh - ceg - bdi - afh$$

# Properties of Matrix Determinants

Here are some useful properties of matrix determinants:

- $|\mathbf{A}| = |\mathbf{A}^{\mathrm{T}}|$
- $|\mathbf{A}^{-1}| = |\mathbf{A}|^{-1}$
- $|\mathbf{AB}| = |\mathbf{A}||\mathbf{B}|$ (if $\mathbf{A}$ and $\mathbf{B}$ are both square matrices)
- $|b\mathbf{A}| = b^n|\mathbf{A}|$ (if $b \in \mathbb{R}$ and $\mathbf{A} = \{a_{ij}\}_{n\times n}$)
- If $\mathbf{A}$ is symmetric, $|\mathbf{A}| = \prod_{j=1}^{n} \lambda_j$, where $\lambda_j$ is the $j$th eigenvalue of $\mathbf{A}$

# Eigenvalue and Eigenvector

**Definition:** Let $\mathbf{A} = \{a_{ij}\}_{n \times n}$. If $\mathbf{A}\mathbf{x} = \lambda \mathbf{x}$ where $\mathbf{x} \neq \mathbf{0}$, then $\lambda$ is an **eigenvalue** of $\mathbf{A}$ and $\mathbf{x}$ is an **eigenvector** of $\mathbf{A}$

We can find the **eigenvalue** and **eigenvector** of a matrix by solving the following eigenvalue problem:

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

For a symmetric matrix $\mathbf{A} = \{a_{ij}\}_{n \times n}$ with eigenvalues $\lambda_1, \ldots, \lambda_n$:

- $\mathrm{rank}(\mathbf{A})$ is the number of non-zero eigenvalues.
- $\mathrm{trace}(\mathbf{A}) = \sum_{j=1}^{n} \lambda_j$.
- $|\mathbf{A}| = \prod_{j=1}^{n} \lambda_j$

# Eigenvalue and Eigenvector: Example

Find the eigenvalues and eigenvectors of $\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$

**Step 1:** Take determinant of $\mathbf{A} - \lambda\mathbf{I}$

$$|\mathbf{A}-\lambda\mathbf{I}| = \begin{vmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{vmatrix} = (2-\lambda)^2-1 = \lambda^2-4\lambda+3 = (\lambda-3)(\lambda-1)$$

**Step 2:** Solve the eigenvalue problem $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$ for both eigenvalues.

$$\mathbf{x}_{\lambda=1} = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \qquad \mathbf{x}_{\lambda=3} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

# Matrix Definiteness

A symmetric matrix $\mathbf{A} = \{a_{ij}\}_{n \times n}$ is **positive definite** if $\mathbf{x}^T \mathbf{A} \mathbf{x} > 0$ for all non-zero $\mathbf{x}$

**An equivalent definition:** A symmetric matrix $\mathbf{A} = \{a_{ij}\}_{n \times n}$ is **positive definite** if all eigenvalues of $\mathbf{A}$ are positive

**Properties of positive definite matrix:**

- All diagonals elements of $\mathbf{A}$ are positive
- $\mathbf{A}$ is invertible, and $\mathbf{A}^{-1}$ is also positive definite
- $\mathrm{trace}(\mathbf{A}) > 0$
- $|\mathbf{A}| > 0$
- If $\mathbf{A}$ is $n \times p$ of rank $p$, then $\mathbf{A}^{\mathrm{T}} \mathbf{A}$ is positive definite

Throughout the course, we write $\mathbf{A} \succ 0$ to indicate positive definiteness

# Matrix Definiteness: Example

Verify that the matrix $\mathbf{A} = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$ is positive definite

**Proof:** show that $\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} > 0$ for all $\mathbf{x} \neq 0$.

# Matrix Definiteness: Example 2

All diagonal elements of a positive definite matrix are positive

**Proof:** Use the property $\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} > 0$ for all $\mathbf{x} \neq 0$.

# Matrix Decompositions

# Eigenvalue (Spectral) Decomposition

**Spectral Theorem:** For any symmetric matrix $\mathbf{A} = \{a_{ij}\}_{n\times n}$, there exists an orthogonal matrix $\mathbf{T}$ such that

$$\mathbf{T}^{\mathrm{T}}\mathbf{A}\mathbf{T} = \mathbf{\Lambda}$$

where $\mathbf{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ is a diagonal matrix with $\lambda_j \in \mathbb{R}$.

**Some Properties of Spectral Decomposition:**

- By convention, $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$
- $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of $\mathbf{A}$ and the columns of $\mathbf{T}$ $(\mathbf{t}_1, \ldots, \mathbf{t}_n)$ are the corresponding eigenvectors
- Note that $\mathbf{T}^{\mathrm{T}}\mathbf{T} = \mathbf{T}\mathbf{T}^{\mathrm{T}} = \mathbf{I}_{n\times n}$
- Related to **Principal Component Analysis**

# Singular Value Decomposition

- Eigenvalue Decomposition works only for **symmetric matrix**
- **Every matrix** has a Singular Value Decomposition (SVD)
- Often, SVD is the best way to think about matrices

# Singular Value Decomposition

The **Singular Value Decomposition** (SVD) decomposes any matrix $\mathbf{A} = \{a_{ij}\}_{n \times p}$ into a product of three matrices:

$$\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{V}^{\mathrm{T}}$$

such that

- $\mathbf{U}$ is an orthogonal $n \times n$ matrix $(\mathbf{U}^{\mathrm{T}}\mathbf{U} = \mathbf{I}_{n \times n})$
- $\mathbf{V}$ is an orthogonal $p \times p$ matrix $(\mathbf{V}^{\mathrm{T}}\mathbf{V} = \mathbf{I}_{p \times p})$
- $\mathbf{D}$ is a diagonal matrix with $d_{ii} > 0$ for all $i \leq \min(n, p)$.
- Columns of $\mathbf{U}$ are **left singular vectors** and columns of $\mathbf{V}$ are **right singular vectors**
- The diagonal elements of $\mathbf{D}$ are the **singular values**

**Note:** The SVD is unique up to signs of columns of $\mathbf{U}$ and $\mathbf{V}$.