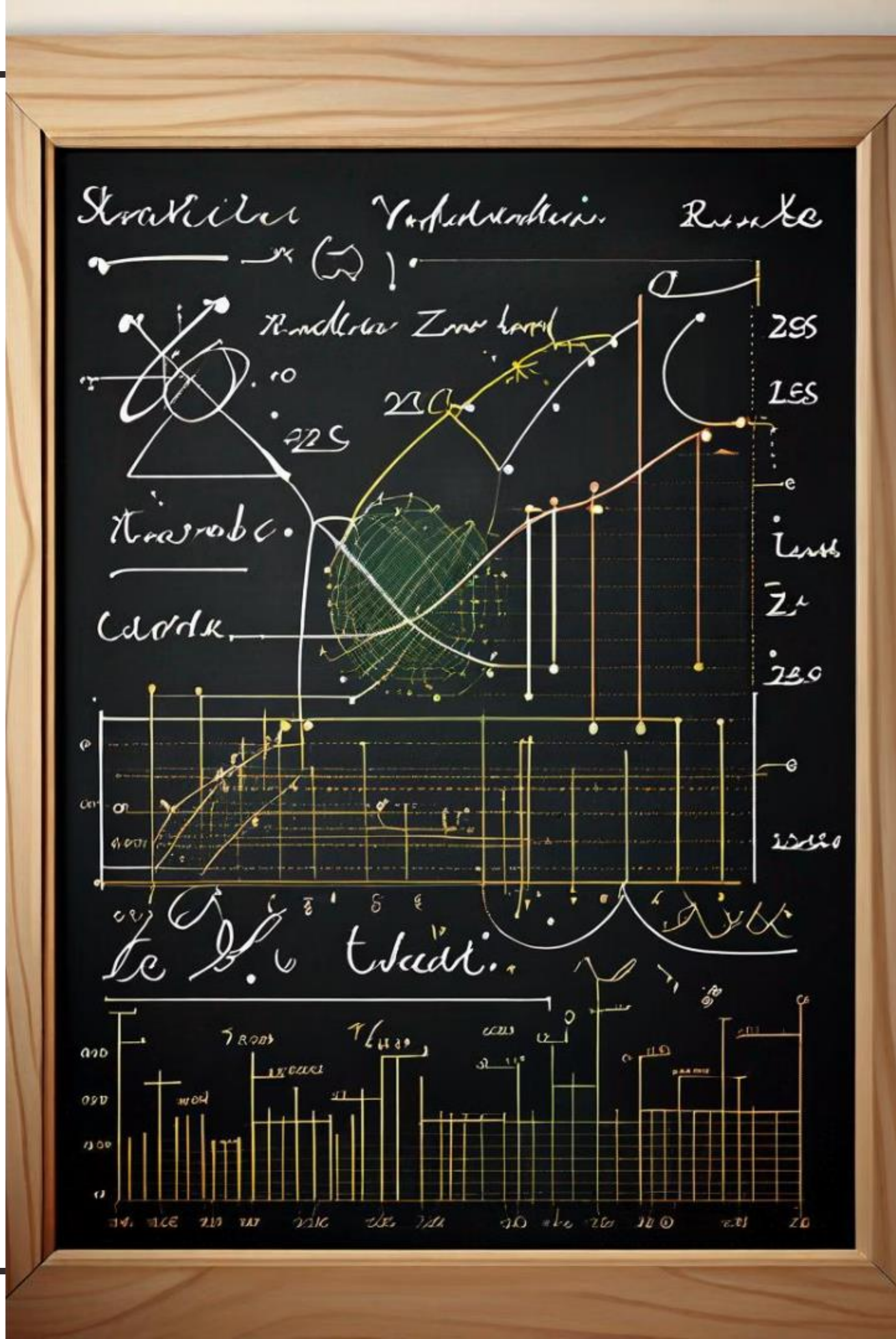
# Linear Algebra, Statistics, and Probability

Abdul Aziz

Assistant Professor, Department of CSE, KUET.

# Introduction to Linear Algebra, Statistics, and Probability Theory

Explore the fundamental mathematical disciplines that underpin modern data analysis and scientific discovery. From the principles of vector spaces to the laws of probability, gain a solid foundation in these essential quantitative fields.



# Vectors and Vector Spaces

### Understanding Vectors

Vectors are mathematical objects that have both magnitude and direction. They are essential in linear algebra and have numerous applications.

### Vector Operations

Vectors can be added, subtracted, and multiplied by scalars, enabling the representation of complex relationships and transformations.

### Vector Spaces

A vector space is a collection of vectors that obey certain rules, allowing for the study of abstract mathematical structures.

# The vector Space $R^n$

**Definition 1.**

Let $(u_1, u_2, \ldots, u_n)$ be a sequence of *n* real numbers. The set of all such sequences is called ***n*-space (or *n*-dimensional. space)** and is denoted $\mathbf{R}^n$.

$u_1$ is the **first component** of $(u_1, u_2, \ldots, u_n)$.

$u_2$ is the **second component** and so on.

***Example 1***

- $\mathbf{R}^2$ is the collection of all sets of two ordered real numbers. For example, (0, 0) , (1, 2) and (-2, -3) are elements of $\mathbf{R}^2$.
- $\mathbf{R}^3$ is the collection of all sets of three ordered real numbers. For example, (0,0, 0) and (-1,3, 4) are elements of $\mathbf{R}^3$.

**Definition 2.**

Let $\mathbf{u} = (u_1, u_2, \ldots, u_n)$ and $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ be two elements of $\mathbf{R}^n$.
We say that **u** and **v** are **equal** if $u_1 = v_1, \ldots, u_n = v_n$.
Thus two elements of $\mathbf{R}^n$ are equal if their **corresponding components** are equal.

**Definition 3.**

Let $\mathbf{u} = (u_1, u_2, \ldots, u_n)$ and $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ be elements of $\mathbf{R}^n$ and let $c$ be a scalar. **Addition** and **scalar multiplication** are performed as follows:

*Addition:* $\mathbf{u} + \mathbf{v} = (u_1 + v_1, \ldots, u_n + v_n)$

*Scalar multiplication :* $c\mathbf{u} = (cu_1, \ldots, cu_n)$

## Example 2

Let $\mathbf{u} = (-1, 4, 3)$ and $\mathbf{v} = (-2, -3, 1)$ be elements of $\mathbf{R}^3$.

Find $\mathbf{u} + \mathbf{v}$ and $3\mathbf{u}$.

**Solution:**

### *Example 3*

In $R^2$ , consider the two elements (4, 1) and (2, 3).

*Find their sum and give a geometrical interpretation of this sum.*

we get $(4, 1) + (2, 3) = (6, 4)$.

The vector (6, 4), the sum, is the diagonal of the parallelogram.

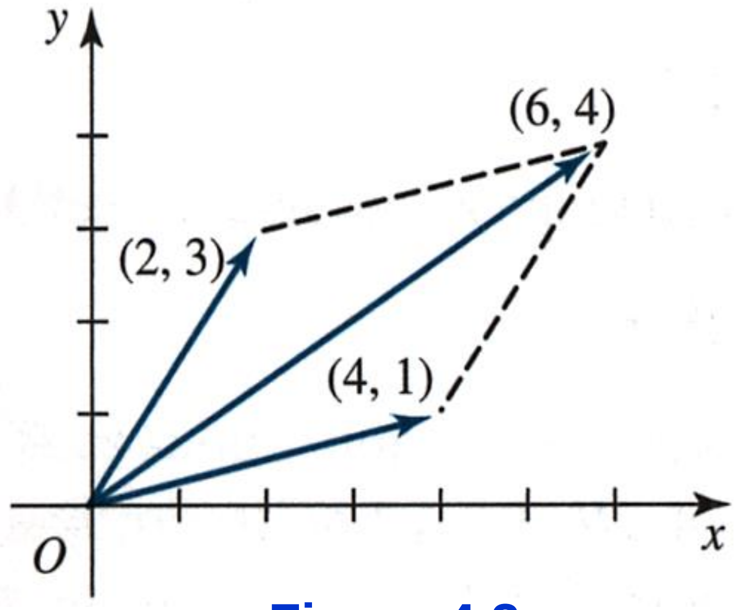



Figure 4.2

# Example 4

Consider the scalar multiple of the vector (3, 2) by 2, we get

$$2(3, 2) = (6, 4)$$

Observe in Figure 4.3 that (6, 4) is a vector in the same direction as (3, 2), and 2 times it in length.

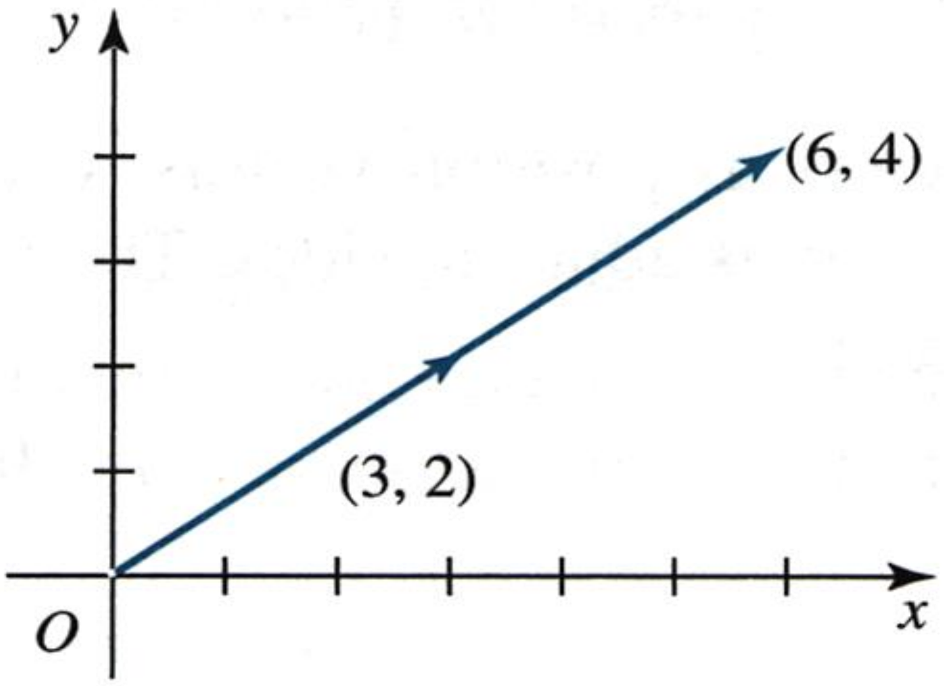


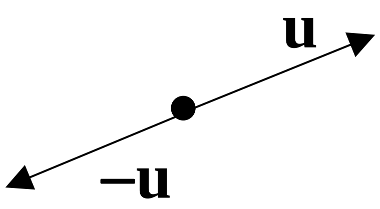

Figure 4.3

# Zero Vector

The vector (0, 0, …, 0), having *n* zero components, is called the **zero vector** of $\mathbf{R}^n$ and is denoted $\mathbf{0}$.

## *Negative Vector*

The vector $(-1)\mathbf{u}$ is writing $-\mathbf{u}$ and is called **the negative of u**. It is a vector having the same length (or magnitude) as $\mathbf{u}$, but lies in the opposite direction to $\mathbf{u}$.




## *Subtraction*

Subtraction is performed on element of $\mathbf{R}^n$ by subtracting corresponding components.

# Theorem 4.1

Let **u**, **v**, and **w** be vectors in $\mathbf{R}^n$ and let $c$ and $d$ be scalars.

(a) $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$

(b) $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$

(c) $\mathbf{u} + \mathbf{0} = \mathbf{0} + \mathbf{u} = \mathbf{u}$

(d) $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$

(e) $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$

(f) $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$

(g) $c(d\mathbf{u}) = (cd)\mathbf{u}$

(h) $1\mathbf{u} = \mathbf{u}$

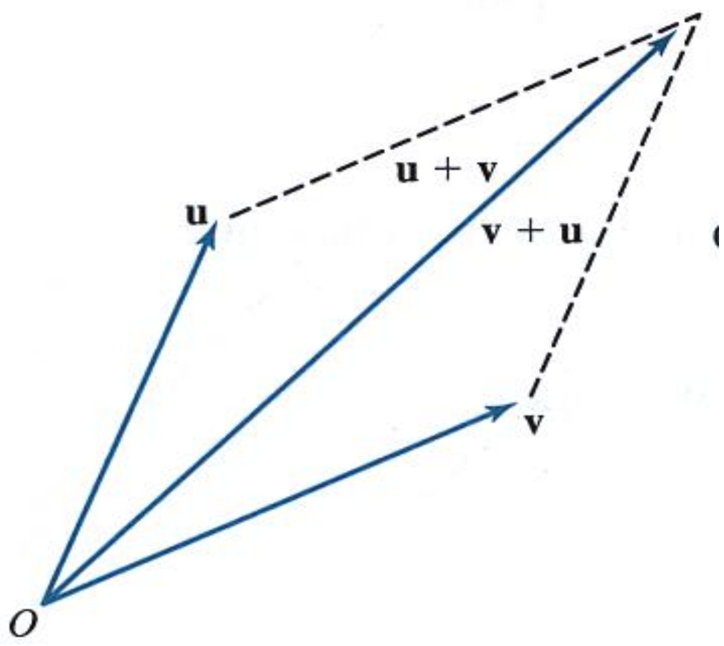



**Figure 4.4**

Commutativity of vector addition

$\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$

# Column Vectors

Row vector: $\mathbf{u} = (u_1, u_2, \ldots, u_n)$

Column vector: $\begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}$

We defined addition and scalar multiplication of column vectors in $\mathbf{R}^n$ in a componentwise manner:

$$\begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} + \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} u_1 + v_1 \\ \vdots \\ u_n + v_n \end{bmatrix} \quad \text{and} \quad c\begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} = \begin{bmatrix} cu_1 \\ \vdots \\ cu_n \end{bmatrix}$$

# Dot Product, Norm, Angle, and Distance

**Definition**

Let $\mathbf{u}=(u_1, u_2, \ldots, u_n)$ and $\mathbf{v}=(v_1, v_2, \ldots, v_n)$ be two vectors in $\mathbf{R}^n$. The **dot product** of **u** and **v** is denoted **u.v** and is defined by

$$\mathbf{u}\cdot\mathbf{v}=u_1v_1+\cdots+u_nv_n$$

The dot product assigns a real number to each pair of vectors.

***Example 1***

Find the dot product of

$$\mathbf{u} = (1, -2, 4) \text{ and } \mathbf{v} = (3, 0, 2)$$

**Solution**

# Properties of the Dot Product

Let **u**, **v**, and **w** be vectors in $\mathbf{R}^n$ and let $c$ be a scalar. Then

1. $\mathbf{u.v} = \mathbf{v.u}$
2. $(\mathbf{u} + \mathbf{v}).\mathbf{w} = \mathbf{u.w} + \mathbf{v.w}$
3. $c\mathbf{u.v} = c(\mathbf{u.v}) = \mathbf{u}.c\mathbf{v}$
4. $\mathbf{u.u} \geq 0$, and $\mathbf{u.u} = 0$ if and only if $\mathbf{u} = \mathbf{0}$

# Norm of a Vector in $\mathbf{R}^n$

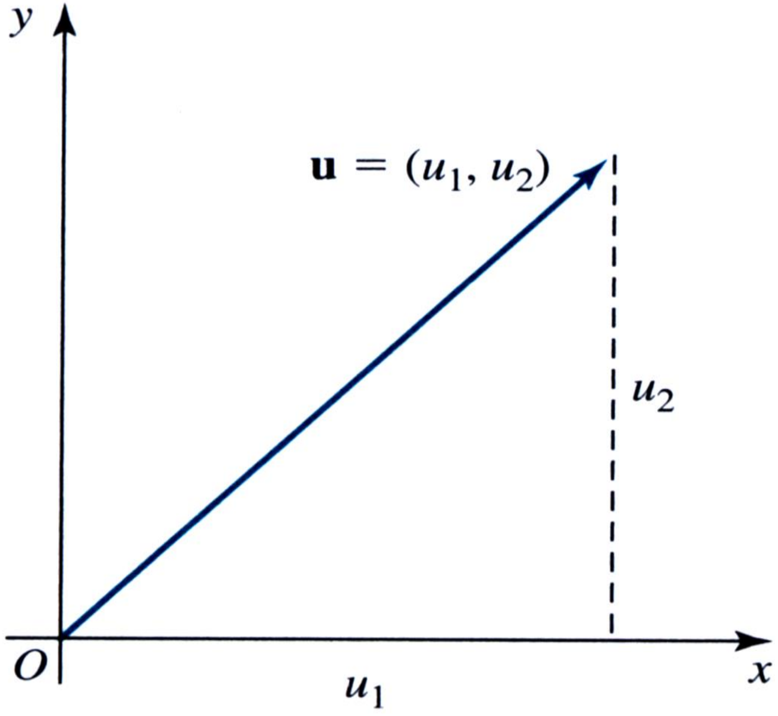



**Figure 4.5** length of $\mathbf{u}$

## Definition

The **norm** (**length** or **magnitude**) of a vector $\mathbf{u} = (u_1, \ldots, u_n)$ in $\mathbf{R}^n$ is denoted $\|\mathbf{u}\|$ and defined by

$$\|\mathbf{u}\| = \sqrt{(u_1)^2 + \cdots + (u_n)^2}$$

*Note:*

The norm of a vector can also be written in terms of the dot product $\|\mathbf{u}\| = \sqrt{\mathbf{u} \cdot \mathbf{u}}$

# Example 2

Find the norm of each of the vectors $\mathbf{u} = (1, 3, 5)$ of $\mathbf{R}^3$ and $\mathbf{v} = (3, 0, 1, 4)$ of $\mathbf{R}^4$.

**Solution**

**Definition**

A **unit vector** is a vector whose norm is 1.

If $\mathbf{v}$ is a nonzero vector, then the vector

$$\boxed{\mathbf{u} = \frac{1}{\|\mathbf{v}\|}\mathbf{v}}$$

is a unit vector in the direction of $\mathbf{v}$.

This procedure of constructing a unit vector in the same direction as a given vector is called **normalizing** the vector.

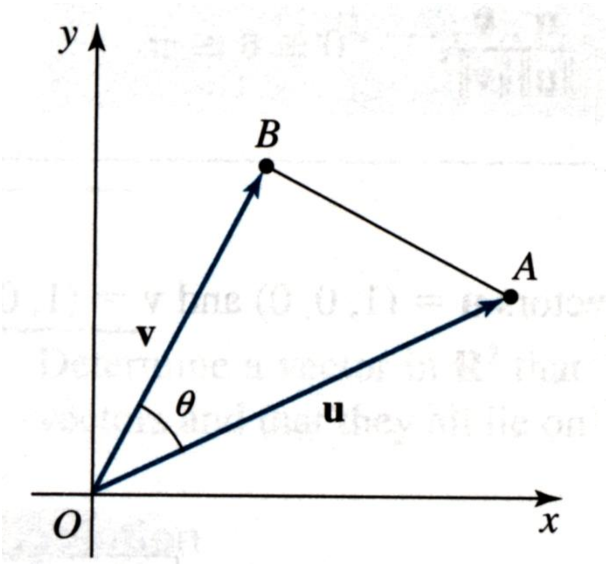
# Example 3

(a) Show that the vector (1, 0) is a unit vector.

(b) Find the norm of the vector (2, –1, 3). Normalize this vector.

**Solution**

# Angle between Vectors ( in $R^2$)




The law of cosines gives:

$$\boxed{\cos\theta = \frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|}}$$

← **Figure**

# Angle between Vectors (in R $^n$)

## Definition

Let $\mathbf{u}$ and $\mathbf{v}$ be two nonzero vectors in $\mathbf{R}^n$.

The **cosine of the angle $\theta$** between these vectors is

$$\cos\theta = \frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|} \qquad 0 \le \theta \le \pi$$

***Example 4***

Determine the angle between the vectors $\mathbf{u} = (1, 0, 0)$ and $\mathbf{v} = (1, 0, 1)$ in $\mathbf{R}^3$.

**Solution**

# Operations with Matrices

**Matrix representations**:

- An uppercase case: *A*, *B*, *C*, …
- A representative element enclosed in brackets: $[a_{ij}]$, $[b_{ij}]$
- A rectangular array of numbers: $\begin{bmatrix} \boxed{\begin{matrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{matrix}} & \boxed{\begin{matrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{matrix}} & \begin{matrix} \cdots \\ \cdots \\ \ddots \\ \cdots \end{matrix} & \boxed{\begin{matrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{matrix}} \end{bmatrix}$

**Vector (column/row matrix):** boldface lowercase

$\mathbf{a}_1$, $\mathbf{a}_2$, …, $\mathbf{a}_n$

# Definitions

- ***Equality of Matrices***
  Two matrices $A = [a_{ij}]$ and $B = [b_{ij}]$ are **equal** if they have the same size ($m \times n$) *and* $a_{ij} = b_{ij}$ for $1 \leq i \leq m$ and $1 \leq j \leq n$.
- ***Matrix Addition***
  If $A = [a_{ij}]$ and $B = [b_{ij}]$ are matrices of size $m \times n$, then their **sum** is the $m \times n$ matrix given by $A+B = [a_{ij} + b_{ij}]$.
  The sum of two matrices of different sizes is undefined.
- ***Scalar Multiplication***
  If $A = [a_{ij}]$ is an $m \times n$ matrix and $c$ is a scalar, then the **scalar multiplication** of $A$ by $c$ is the $m \times n$ matrix given by $cA = [ca_{ij}]$

# Example 1

Consider the four matrices

$$A=\begin{bmatrix}1 & 2\\ 3 & 4\end{bmatrix},\ B=\begin{bmatrix}1\\ 3\end{bmatrix},\ C=\begin{bmatrix}1 & 3\end{bmatrix},\ D=\begin{bmatrix}1 & 2\\ x & 4\end{bmatrix}$$

- Matrices *A* and *B* are **not** equal because they are of different sizes.
- Similarly, *B* and *C* are **not** equal.
- Matrices *A* and *D* are **equal** *if and only if* (iff) $x = 3$

# Subtraction of Matrices

- If $A$ and $B$ are of *the same size*, $A–B$ represents the sum of $A$ and $(–B)$. That is, $A–B = A+(–1)B = [a_{ij} – b_{ij}]$.
- $cA – dB = [ca_{ij} – db_{ij}]$.
- ***Example 3***:

$$A = \begin{bmatrix} 1 & 2 & 4 \\ -3 & 0 & -1 \\ 2 & 1 & 2 \end{bmatrix} \text{ and } B = \begin{bmatrix} 2 & 0 & 0 \\ 1 & -4 & 3 \\ -1 & 3 & 2 \end{bmatrix}$$

$$3A - B = 3\begin{bmatrix} 1 & 2 & 4 \\ -3 & 0 & -1 \\ 2 & 1 & 2 \end{bmatrix} - \begin{bmatrix} 2 & 0 & 0 \\ 1 & -4 & 3 \\ -1 & 3 & 2 \end{bmatrix} = \begin{bmatrix} 3\cdot 1 - 2 & 3\cdot 2 - 0 & 3\cdot 4 - 0 \\ 3\cdot(-3) - 1 & 3\cdot 0 - (-4) & 3\cdot(-1) - 3 \\ 3\cdot 2 - (-1) & 3\cdot 1 - 3 & 3\cdot 2 - 2 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 6 & 12 \\ -10 & 4 & -6 \\ 7 & 0 & 4 \end{bmatrix}$$

# Matrix Multiplication

- If $A = [a_{ij}]$ is an $m \times n$ matrix and $B = [b_{ij}]$ is an $n \times p$ matrix, then the **product** $AB$ is an $m \times p$ matrix $AB = [c_{ij}]$, where

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} = a_{i1} b_{1j} + a_{i2} b_{2j} + \cdots + a_{in} b_{nj}$$

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \boxed{a_{21} \quad a_{22} \quad \cdots \quad a_{2n}} \\ \vdots & \vdots & & \vdots \\ \boxed{a_{i1} \quad a_{i2} \quad \cdots \quad a_{in}} \\ \vdots & \vdots & & \vdots \\ \boxed{a_{m1} \quad a_{m2} \quad \cdots \quad a_{mn}} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1j} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2j} & \cdots & b_{2p} \\ \vdots & \vdots & & \vdots & & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nj} & \cdots & b_{np} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1j} & \cdots & c_{1p} \\ c_{21} & c_{22} & \cdots & c_{2j} & \cdots & c_{2p} \\ \vdots & \vdots & & \vdots & & \vdots \\ c_{i1} & c_{i2} & \cdots & c_{ij} & \cdots & c_{ip} \\ \vdots & \vdots & & \vdots & & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mj} & \cdots & c_{mp} \end{bmatrix}$$

$$A \qquad B \qquad = \qquad AB$$

$$m \times n \qquad n \times p \qquad \boxed{m \times p}$$

# Example 4

Find the product $AB$, where $A = \begin{bmatrix} -1 & 3 \\ 4 & -2 \\ 5 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} -3 & 2 \\ -4 & 1 \end{bmatrix}$

$$\underset{3\times 2}{\begin{bmatrix} -1 & 3 \\ 4 & -2 \\ 5 & 0 \end{bmatrix}} \underset{2\times 2}{\begin{bmatrix} -3 & 2 \\ -4 & 1 \end{bmatrix}} = \underset{3\times 2}{\begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \\ c_{31} & c_{32} \end{bmatrix}} = \begin{bmatrix} -9 & 1 \\ -4 & 6 \\ -15 & 10 \end{bmatrix}$$

$c_{11} = (-1)(-3) + (3)(-4) = -9 \qquad c_{12} = (-1)(2) + (3)(1) = 1$

$c_{21} = (4)(-3) + (-2)(-4) = -4 \qquad c_{22} = (4)(2) + (-2)(1) = 6$

$c_{31} = (5)(-3) + (0)(-4) = -15 \qquad c_{32} = (5)(2) + (0)(1) = 10$

# Example 5

$$(a)\begin{bmatrix}1 & 0 & 3\\ 2 & -1 & -2\end{bmatrix}\begin{bmatrix}-2 & 4 & 2\\ 1 & 0 & 0\\ -1 & 1 & -1\end{bmatrix}=\begin{bmatrix} & & \\ & & \end{bmatrix}_{2\times 3}$$

$$(c)\begin{bmatrix}1 & 2\\ 1 & 1\end{bmatrix}\begin{bmatrix}-1 & 2\\ 1 & -1\end{bmatrix}=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}_{2\times 2}$$

$$(d)\begin{bmatrix}1 & -2 & -3\end{bmatrix}\begin{bmatrix}2\\ -1\\ 1\end{bmatrix}=\begin{bmatrix}1\end{bmatrix}_{1\times 1}$$

$$(e)\begin{bmatrix}2\\ -1\\ 1\end{bmatrix}\begin{bmatrix}1 & -2 & -3\end{bmatrix}=\begin{bmatrix}2 & -4 & -6\\ -1 & 2 & 3\\ 1 & -2 & -3\end{bmatrix}_{3\times 3}$$

$AB \neq BA$

Matrix multiplication is ***not***, in general, **commutative**.

# Systems of Linear Equations

- Matrix Equation: $A\mathbf{x} = \mathbf{b}$

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \Leftrightarrow \begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3 \end{cases}$$

  - $A$: coefficient matrix; **x** and **b**: column matrix (vector)

- ***Example 6***: Solve the matrix equation $A\mathbf{x} = \mathbf{0}$

# Diagonal Matrix & Trace

- A ***square*** matrix

$$A=\begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}$$

  is called a **diagonal matrix** *if* all entries that *not* on the main diagonal are **zero**.

- The **trace** of an $n \times n$ matrix $A$ is the *sum* of the main diagonal entries. That is,

$$Tr(A)=a_{11}+a_{22}+\cdots+a_{nn}$$

# 2.2 Properties of Matrix Operations

## *Theorem 2.1*

### *Properties of Matrix Addition and Scalar Multiplication*

- If $A$, $B$, and $C$ are $m \times n$ matrices and $c$ and $d$ are scalars, then the following properties are *true*.

| | |
|---|---|
| 1. $A+B = B+A$ | Commutative property of addition |
| 2. $A+(B+C) = (A+B)+C$ | Associative property of addition |
| 3. $(cd)A = c(dA)$ | Associative property of multiplication |
| 4. $1A = A$ | Multiplication identity |
| 5. $c(A+B) = cA + cB$ | Distributive property |
| 6. $(c+d)A = cA + dA$ | Distributive property |

# Zero Matrix & Additive Identity

- If $A$ is an $m \times n$ matrix and $O_{mn}$ is the $m \times n$ matrix consisting entirely of **zeros**, then $A + O_{mn} = A$.
- The matrix $O_{mn}$ is called a **zero matrix**, and it serves as the **additive identity** for the set of all $m \times n$ matrices.
- ***Theorem 2.2****: Properties of Zero Matrix*
  If $A$ is an $m \times n$ matrix and $c$ is a scalar, then the following properties are *true*.
  1. $A + O_{mn} = A$.
  2. $A + (-A) = O_{mn}$. $\Rightarrow -A$ is the **additive inverse** of $A$.
  3. If $cA = O_{mn}$, then $c = 0$ **or** $A = O_{mn}$.

# Matrix Equation

| Real Numbers | $m \times n$ Matrices |
| --- | --- |
| $x + a = b$ | $X + A = B$ |
| $x + a + (-a) = b + (-a)$ | $X + A + (-A) = B + (-A)$ |
| $x + 0 = b - a$ | $X + O = B - A$ |
| $x = b - a$ | $X = B - A$ |

- ***Ex. 2***: Solve for *X* in the equation 3*X*+*A* = *B*, where

$$A = \begin{bmatrix} 1 & -2 \\ 0 & 3 \end{bmatrix} \text{ and } B = \begin{bmatrix} -3 & 4 \\ 2 & 1 \end{bmatrix}$$

$$X = \frac{1}{3}(B - A) = \frac{1}{3}\left( \begin{bmatrix} -3 & 4 \\ 2 & 1 \end{bmatrix} - \begin{bmatrix} 1 & -2 \\ 0 & 3 \end{bmatrix} \right) = \frac{1}{3}\begin{bmatrix} -4 & 6 \\ 2 & -2 \end{bmatrix} = \begin{bmatrix} -\frac{4}{3} & 2 \\ \frac{2}{3} & -\frac{2}{3} \end{bmatrix}$$

# Theorem 2.3

- ***Properties of Matrix Multiplication***

If $A$, $B$, and $C$ are matrices (*with sizes such that the given matrix products are defined*) and $c$ is a scalar, then the following properties are *true*.

1. $A(BC) = (AB)C$ — Associative property
2. $A(B+C) = AB + AC$ — Distribution property
3. $(A+B)C = AC + BC$ — Distribution property
4. $c(AB) = (cA)B = A(cB)$

# Noncommutativity

- A ***commutative property*** for matrix multiplication was NOT listed in Theorem 2.3.
- If $A$ is of size $\underline{2\times 3}$ and $B$ is of size $\underline{3\times 3}$, then the product $AB$ is **defined**, but the product $BA$ is not.
- ***Example 4***: Show that $AB$ and $BA$ are not equal for the matrices and

$$A=\begin{bmatrix}1 & 3\\ 2 & -1\end{bmatrix} \qquad B=\begin{bmatrix}2 & -1\\ 0 & 2\end{bmatrix}$$

$$\left.\begin{aligned} AB&=\begin{bmatrix}1 & 3\\ 2 & -1\end{bmatrix}\begin{bmatrix}2 & -1\\ 0 & 2\end{bmatrix}=\begin{bmatrix}2 & 5\\ 4 & -4\end{bmatrix}\\ BA&=\begin{bmatrix}2 & -1\\ 0 & 2\end{bmatrix}\begin{bmatrix}1 & 3\\ 2 & -1\end{bmatrix}=\begin{bmatrix}0 & 7\\ 4 & -2\end{bmatrix}\end{aligned}\right\} AB\neq BA$$

# Cancellation Property

- It does ***NOT*** have a general *cancellation property* for matrix multiplication.
- If $AC = BC$, it is ***NOT*** necessary true that $A = B$.
- ***Example 5***: Show that $AC = BC$.

$$A = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 4 \\ 2 & 3 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & -2 \\ -1 & 2 \end{bmatrix}$$

$$\left.\begin{aligned} AB &= \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 2 & 4 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} -2 & 4 \\ -1 & 2 \end{bmatrix} \\ AC &= \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & -2 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} -2 & 4 \\ -1 & 2 \end{bmatrix} \end{aligned}\right\} AC = BC, \text{ but } A \neq B$$

# Identity Matrix & Theorem 4

- A **square matrix** that has 1's on the main diagonal and 0's elsewhere.
- The identity matrix of order ***n***: $$I_n = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$
- ***Theorem 2.4**: Properties of the Identity Matrix*
  If $A$ is a matrix of size $m \times n$, then the following properties are *true*.
  1. $AI_n = A$. 2. $I_m A = A$.
- If $A$ is a <u>*square matrix*</u> of order $n$, then $AI_n = I_n A = A$.

# Repeated Multiplication

- Repeated multiplication of a *square* matrix:
  For a positive integer $k$, $A^k$ is

$$A^k = \underbrace{AA\cdots A}_{k \text{ factors}}$$

- $A^0 = I_n$, where $A$ is a square matrix of order $n$.
- $A^j A^k = A^{j+k}$
- $(A^j)^k = A^{jk}$

$j$ and $k$ are nonnegative integer.

- ***Example 3***: Find $A^3$ for the matrix $A = \begin{bmatrix} 2 & -1 \\ 3 & 0 \end{bmatrix}$.

$$A^3 = \left( \begin{bmatrix} 2 & -1 \\ 3 & 0 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ 3 & 0 \end{bmatrix} \right) \begin{bmatrix} 2 & -1 \\ 3 & 0 \end{bmatrix} = \begin{bmatrix} 1 & -2 \\ 6 & -3 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ 3 & 0 \end{bmatrix} = \begin{bmatrix} -4 & -1 \\ 3 & -6 \end{bmatrix}$$

# The Transpose of a Matrix

- The **transpose** of a matrix is formed by *writing its columns as rows*.

$$A=\begin{bmatrix} a_{11} & \boxed{a_{12}} & \boxed{a_{13}} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix}_{m\times n} \Leftrightarrow A^T=\begin{bmatrix} a_{11} & a_{21} & a_{31} & \cdots & a_{m1} \\ a_{12} & a_{22} & a_{32} & \cdots & a_{m2} \\ a_{13} & a_{23} & a_{33} & \cdots & a_{m3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & a_{3n} & \cdots & a_{mn} \end{bmatrix}_{n\times m}$$

- A matrix *A* i

$\Rightarrow$ a symme

# Theorem 2.6

### *Properties of Transpose*

- If $A$ and $B$ are matrices (*with sizes such that the given matrix products are defined*) and $c$ is a scalar, then the following properties are *true*.
  1. $(A^T)^T = A$ **Transpose of a transpose**
  2. $(A+B)^T = A^T + B^T$ **Transpose of a sum**
  3. $(cA)^T = c(A^T)$ **Transpose of a scalar multiplication**
  4. $(AB)^T = B^T A^T$ **Transpose of a product**
-
- For *any* matrix $A$, the matrix $AA^T$ is ***symmetric***.

$$(A+B+C)^T = A^T + B^T + C^T \qquad (ABC)^T = C^T B^T A^T$$

$$pf:(AA^T)^T = (A^T)^T A^T = AA^T$$

# Example 9

- Show that are equal.

$$(AB)^T = B^T A^T$$

$$A = \begin{bmatrix} 2 & 1 & -2 \\ -1 & 0 & 3 \\ 0 & -2 & 1 \end{bmatrix} \qquad B = \begin{bmatrix} 3 & 1 \\ 2 & -1 \\ 3 & 0 \end{bmatrix}$$

***Sol:***

$$AB = \begin{bmatrix} 2 & 1 & -2 \\ -1 & 0 & 3 \\ 0 & -2 & 1 \end{bmatrix}\begin{bmatrix} 3 & 1 \\ 2 & -1 \\ 3 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 6 & -1 \\ -1 & 2 \end{bmatrix} \Rightarrow (AB)^T = \begin{bmatrix} 2 & 6 & -1 \\ 1 & -1 & 2 \end{bmatrix}$$

$$B^T A^T = \begin{bmatrix} 3 & 2 & 3 \\ 1 & -1 & 0 \end{bmatrix}\begin{bmatrix} 2 & -1 & 0 \\ 1 & 0 & -2 \\ -2 & 3 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 6 & -1 \\ 1 & -1 & 2 \end{bmatrix}$$

$$(AB)^T = B^T A^T$$

# 2.3 The Inverse of a Matrix

- ***Definition of an Inverse of a Matrix***

  An $n \times n$ matrix $A$ is **invertible** if there exists an $n \times n$ matrix $B$ such that $AB = BA = I_n$
  - $I_n$ is the identity matrix of order $n$.
  - The matrix $B$ is called the (multiplicative) inverse of $A$.
  - A matrix that does **NOT** have an inverse is called **noninvertible**.
- *Nonsquare* matrices do NOT have inverse.
- ***Theorem 2.7: Uniqueness of an Inverse Matrix***

  If $A$ is an **invertible** matrix, then *its inverse is **unique***.

  The inverse of $A$ is denoted by $A^{-1}$.

$$AA^{-1} = A^{-1}A = I$$

# Example 2

Find the inverse of the matrix

$$A = \begin{bmatrix} 1 & 4 \\ -1 & -3 \end{bmatrix}$$

- ***Sol:*** To find the inverse of *A*, try to solve the matrix equation *AX* = *I* for *X*.

$$\begin{bmatrix} 1 & 4 \\ -1 & -3 \end{bmatrix}\begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} x_{11} + 4x_{21} & x_{12} + 4x_{22} \\ -x_{11} - 3x_{21} & -x_{12} - 3x_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\Rightarrow \begin{cases} x_{11} + 4x_{21} = 1 \\ -x_{11} - 3x_{21} = 0 \end{cases} \quad \therefore x_{11} = -3,\ x_{21} = 1$$

$$\begin{cases} x_{12} + 4x_{22} = 0 \\ -x_{12} - 3x_{22} = 1 \end{cases} \quad \therefore x_{12} = -4,\ x_{22} = 1$$

$$A^{-1} = X = \begin{bmatrix} -3 & -4 \\ 1 & 1 \end{bmatrix}$$

Using matrix multiplication to check the result.

$$\begin{bmatrix} 1 & 4 \\ -1 & -3 \end{bmatrix}\begin{bmatrix} -3 & -4 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

# Example 6

$$A=\begin{bmatrix}1 & 1\\ 2 & 4\end{bmatrix}$$

Compute $A^{-2}$ in two different ways and show that the results are equal.

1. $(A^2)^{-1}$:

$$A^2=\begin{bmatrix}1 & 1\\ 2 & 4\end{bmatrix}\begin{bmatrix}1 & 1\\ 2 & 4\end{bmatrix}=\begin{bmatrix}3 & 5\\ 10 & 18\end{bmatrix}\Rightarrow\Delta=(3)(18)-(5)(10)=4$$

$$\therefore (A^2)^{-1}=\frac{1}{4}\begin{bmatrix}18 & -5\\ -10 & 3\end{bmatrix}=\boxed{\begin{bmatrix}\frac{9}{2} & -\frac{5}{4}\\ -\frac{5}{2} & \frac{3}{4}\end{bmatrix}}$$

2. $(A^{-1})^2$:

$$A=\begin{bmatrix}1 & 1\\ 2 & 4\end{bmatrix}\Rightarrow\Delta=(1)(4)-(2)(1)=2$$

the same result

$$\therefore A^{-1}=\frac{1}{2}\begin{bmatrix}4 & -1\\ -2 & 1\end{bmatrix}=\begin{bmatrix}2 & -\frac{1}{2}\\ -1 & \frac{1}{2}\end{bmatrix}$$

$$(A^{-1})^2=\begin{bmatrix}2 & -\frac{1}{2}\\ -1 & \frac{1}{2}\end{bmatrix}\begin{bmatrix}2 & -\frac{1}{2}\\ -1 & \frac{1}{2}\end{bmatrix}=\boxed{\begin{bmatrix}\frac{9}{2} & -\frac{5}{4}\\ -\frac{5}{2} & \frac{3}{4}\end{bmatrix}}$$

# Example 7

Find $(AB)^{-1}$ for the matrices and

using the fact that $A^{-1}$ and $B^{-1}$ are given by

$$A=\begin{bmatrix}1&3&3\\1&4&3\\1&3&4\end{bmatrix}\qquad B=\begin{bmatrix}1&2&3\\1&3&3\\2&4&3\end{bmatrix}$$

$$A^{-1}=\begin{bmatrix}7&-3&-3\\-1&1&0\\-1&0&1\end{bmatrix}\quad B^{-1}=\begin{bmatrix}1&-2&1\\-1&1&0\\\frac{2}{3}&0&-\frac{1}{3}\end{bmatrix}$$

***Sol:***

$$(AB)^{-1}=B^{-1}A^{-1}=\begin{bmatrix}1&-2&1\\-1&1&0\\\frac{2}{3}&0&-\frac{1}{3}\end{bmatrix}\begin{bmatrix}7&-3&-3\\-1&1&0\\-1&0&1\end{bmatrix}=\begin{bmatrix}8&-5&-2\\-8&4&3\\5&-2&-\frac{7}{3}\end{bmatrix}$$

$$AB=\begin{bmatrix}10&23&21\\11&26&24\\12&27&24\end{bmatrix}\Rightarrow\left[\begin{array}{ccc:ccc}10&23&21&1&0&0\\11&26&24&0&1&0\\12&27&24&0&0&1\end{array}\right]\Rightarrow\left[\begin{array}{ccc:ccc}1&0&0&8&-5&-2\\0&1&0&-8&4&3\\0&0&1&5&-2&-\frac{7}{3}\end{array}\right]$$

# Theorem 2.11

- ***Systems of Equations with Unique Solutions***

If $A$ is an **invertible** matrix, then the system of linear equations $A\mathbf{x} = \mathbf{b}$ has a **unique solution** given by
$\mathbf{x} = A^{-1}\mathbf{b}$.

***pf***: $A\mathbf{x} = \mathbf{b} \Rightarrow A^{-1}(A\mathbf{x}) = A^{-1}(\mathbf{b}) \Rightarrow A^{-1}A\mathbf{x} = A^{-1}\mathbf{b} \Rightarrow \mathbf{x} = A^{-1}\mathbf{b}$

- ***Example 8: Use an inverse matrix to solve each system***

$$(a)\ \begin{aligned} 2x+3y+z&=-1\\ 3x+3y+z&=\ \ 1\\ 2x+4y+z&=-1 \end{aligned} \qquad A=\begin{bmatrix}2&3&1\\3&3&1\\2&4&1\end{bmatrix} \Rightarrow A^{-1}=\begin{bmatrix}-1&1&0\\-1&0&1\\6&-2&-3\end{bmatrix}$$

$$(c)\ \begin{aligned} 2x+3y+z&=0\\ 3x+3y+z&=0\\ 2x+4y+z&=0 \end{aligned} \qquad (a)\ \mathbf{x}=A^{-1}\mathbf{b}=\begin{bmatrix}-1&1&0\\-1&0&1\\6&-2&-3\end{bmatrix}\begin{bmatrix}-1\\1\\-2\end{bmatrix}=\begin{bmatrix}2\\-1\\-2\end{bmatrix} \quad (c)\ \mathbf{x}=A^{-1}\mathbf{b}=\begin{bmatrix}0\\0\\0\end{bmatrix}$$

# Eigenvalues and Eigenvectors

# 6.1 Definitions

Definition 1: A nonzero vector **x** is an ***eigenvector*** (or *characteristic vector*) of a square matrix **A** if there exists a scalar **λ** such that **Ax = λx**. Then **λ** is an ***eigenvalue*** (or *characteristic value*) of **A**.

*Note*: The zero vector can not be an eigenvector even though **A0 = λ0**. But **λ = 0** can be an eigenvalue.

Example:

$$Show\ x = \begin{bmatrix} 2 \\ 1 \end{bmatrix} is\ an\ eigenvector\ for\ A = \begin{bmatrix} 2 & -4 \\ 3 & -6 \end{bmatrix}$$

$$Solution: Ax = \begin{bmatrix} 2 & -4 \\ 3 & -6 \end{bmatrix}\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$But\ \ for\ \ \lambda = 0,\ \ \lambda x = 0\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$Thus, x\ is\ an\ eigenvector\ of\ A, and\ \lambda = 0\ is\ an\ eigenvalue.$$

# Geometric interpretation of Eigenvalues and Eigenvectors

An $n \times n$ matrix **A** multiplied by $n \times 1$ vector **x** results in another $n \times 1$ vector **y=Ax**. Thus, **A** can be considered as a transformation matrix.

In general, a matrix acts on a vector by changing both its magnitude and its direction. However, a matrix may act on certain vectors by changing only their magnitude, and leaving their direction unchanged (or possibly reversing it). These vectors are the **eigenvectors** of the matrix.

A matrix acts on an eigenvector by multiplying its magnitude by a factor, which is positive if its direction is unchanged and negative if its direction is reversed. This factor is the **eigenvalue** associated with that eigenvector.

# 6.2 Eigenvalues

Let x be an eigenvector of the matrix **A.** Then there must exist an eigenvalue $\lambda$ such that $$Ax = \lambda x$$ or, equivalently,

$$Ax - \lambda x = 0$$ or

$$(A - \lambda I)x = 0$$

If we define a new matrix $B = A - \lambda I$, then

$$Bx = 0$$

If **B** has an inverse, then $x = B^{-1}0 = 0$. But an eigenvector cannot be zero.

Thus, it follows that x will be an eigenvector of **A** if and only if **B** does not have an inverse, or equivalently $\det(B)=0$, or

$$\det(A - \lambda I) = 0$$

This is called the **characteristic equation** of **A**. Its roots determine the eigenvalues of **A**.

# 6.2 Eigenvalues: examples

Example 1: Find the eigenvalues of $A = \begin{bmatrix} 2 & -12 \\ 1 & -5 \end{bmatrix}$

$$|\lambda I - A| = \begin{vmatrix} \lambda - 2 & 12 \\ -1 & \lambda + 5 \end{vmatrix} = (\lambda - 2)(\lambda + 5) + 12$$

$$= \lambda^2 + 3\lambda + 2 = (\lambda + 1)(\lambda + 2)$$

two eigenvalues: $-1, -2$

*Note:* The roots of the characteristic equation can be repeated. That is, $\lambda_1 = \lambda_2 = \ldots = \lambda_k$. If that happens, the eigenvalue is said to be of multiplicity k.

Example 2: Find the eigenvalues of

$$A = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

$$|\lambda I - A| = \begin{vmatrix} \lambda - 2 & -1 & 0 \\ 0 & \lambda - 2 & 0 \\ 0 & 0 & \lambda - 2 \end{vmatrix} = (\lambda - 2)^3 = 0$$

$\lambda = 2$ is an eigenvector of multiplicity 3.

# 6.3 Eigenvectors

To each distinct eigenvalue of a matrix **A** there will correspond at least one eigenvector which can be found by solving the appropriate set of homogenous equations. If $\boldsymbol{\lambda}_i$ is an eigenvalue then the corresponding eigenvector $\mathbf{x}_i$ is the solution of $(\mathbf{A} - \boldsymbol{\lambda}_i\mathbf{I})\mathbf{x}_i = \mathbf{0}$

Example 1 (cont.):

$$\lambda = -1 : (-1)I - A = \begin{bmatrix} -3 & 12 \\ -1 & 4 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & -4 \\ 0 & 0 \end{bmatrix}$$

$$x_1 - 4x_2 = 0 \Rightarrow x_1 = 4t,\ x_2 = t$$

$$\mathbf{x}_1 = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = t\begin{bmatrix} 4 \\ 1 \end{bmatrix},\ t \neq 0$$

$$\lambda = -2 : (-2)I - A = \begin{bmatrix} -4 & 12 \\ -1 & 3 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & -3 \\ 0 & 0 \end{bmatrix}$$

$$\mathbf{x}_2 = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = s\begin{bmatrix} 3 \\ 1 \end{bmatrix},\ s \neq 0$$

# 6.3 Eigenvectors

Example 2 (cont.): Find the eigenvectors of

$$A = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

Recall that λ = 2 is an eigenvector of multiplicity 3.

Solve the homogeneous linear system represented by

$$(2I - A)\mathbf{x} = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Let $x_1 = s,\ x_3 = t$. The eigenvectors of $\lambda$ = 2 are of the form

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} s \\ 0 \\ t \end{bmatrix} = s\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + t\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix},$$ *s* and *t* not both zero.

# Probability Axioms and Formulas

# Probability Axioms and Formulas

- We have known that a sample space is a set of all outcomes of a random process or experiment and that an event is a subset of a sample space.

**Probability Axioms**

Let $S$ be a sample space, A **probability function** $P$ from the set of all events in $S$ to the set of real numbers satisfies the following three axioms: For all events $A$ and $B$ in $S$,

1. $0 \leq P(A) \leq 1$
2. $P(\emptyset) = 0$ and $P(S) = 1$
3. If $A$ and $B$ are disjoint (that is, if $A \cap B = \emptyset$), then the probability of the union of $A$ and $B$ is

$$P(A \cup B) = P(A) + P(B).$$

# Probability Axioms and Formulas

**Probability of the Complement of an Event**

If $A$ is any event in a sample space $S$, then

$$P(A^c) = 1 - P(A). \qquad 9.8.1$$

**Probability of a General Union of Two Events**

If $S$ is any sample space and $A$ and $B$ are any events in $S$, then

$$P(A \cup B) = P(A) + P(B) - P(A \cap B). \qquad 9.8.2$$

# Expected Value

**• Definition**

Suppose the possible outcomes of an experiment, or random process, are real numbers $a_1, a_2, a_3, \ldots, a_n$, which occur with probabilities $p_1, p_2, p_3, \ldots, p_n$. The **expected value** of the process is

$$\sum_{k=1}^{n} a_k p_k = a_1 p_1 + a_2 p_2 + a_3 p_3 + \cdots + a_n p_n.$$

# Expected Value of a Lottery

- Suppose that 100 people pay $1 each to play a lottery game with a single prize of $98 for exact one winner.

- What is the expected value of a ticket?

# Example 5 – *Expected Value of a Lottery*

- Suppose that 500,000 people pay \$5 each to play a lottery game with the following prizes:

- a grand prize of \$1,000,000, 10 second prizes of \$1,000 each, 1,000 third prizes of \$500 each, and 10,000 fourth prizes of \$10 each.

- What is the expected value of a ticket?

- **Solution:**

- Each of the 500,000 lottery tickets has the same chance as any other of containing a winning lottery number, and so

for all *k* = 1, 2, 3, . . . , 500000.

$$p_k = \frac{1}{500000}$$

# Example 5 – *Solution*

- Let $a_1, a_2, a_3, \ldots, a_{500000}$ be the net gain for an individual ticket, where $a_1 = 999995$ (the net gain for the grand prize ticket, which is one million dollars minus the \$5 cost of the winning ticket),

- $a_2 = a_3 = \cdot \cdot \cdot = a_{11} = 995$ (the net gain for each of the 10 second prize tickets),

- $a_{12} = a_{13} = \cdot \cdot \cdot = a_{1011} = 495$ (the net gain for each of the 1,000 third prize tickets), and

- $a_{1012} = a_{1013} = \cdot \cdot \cdot = a_{11011} = 5$ (the net gain for each of the 10,000 fourth prize tickets).

# Example 5 – *Solution*

- Since the remaining 488,989 tickets just lose \$5, $a_{11012} = a_{11013} = \cdot \cdot \cdot = a_{500000} = -5$.

- The expected value of a ticket is therefore

$$\sum_{k=1}^{500000} a_k p_k = \sum_{k=1}^{500000} \left( a_k \cdot \frac{1}{500000} \right) \qquad \text{because each } p_k = 1/500000$$

$$= \frac{1}{500000} \sum_{k=1}^{500000} a_k \qquad \text{by Theorem 5.1.1(2)}$$

# Example 5 – *Solution*

$$= \frac{1}{500000}(999995 + 10 \cdot 995 + 1000 \cdot 495 + 10000 \cdot 5 + (-5) \cdot 488989)$$

$$= \frac{1}{500000}(999995 + 9950 + 495000 + 50000 - 2444945)$$

$$= -1.78.$$

- In other words, a person who continues to play this lottery for a very long time will probably win some money occasionally but on average will lose \$1.78 per ticket.

# Conditional Probability

- Imagine a couple with two children, each of whom is equally likely to be a boy or a girl. Now suppose you are given the information that one is a boy. What is the probability that the other child is a boy?

- Figure 9.9.1 shows the four equally likely combinations of gender for the children.

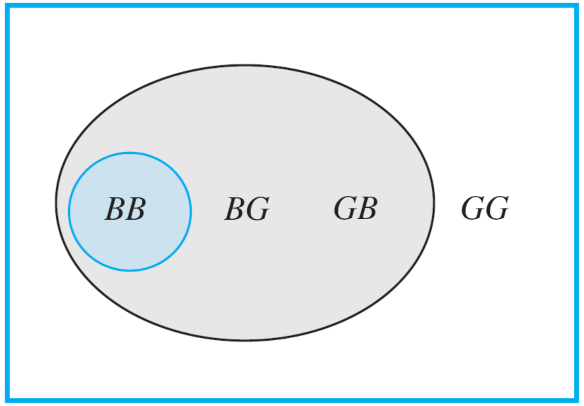



**Figure 9.9.1**

# Conditional Probability

- A generalization of this observation forms the basis for the following definition.

**• Definition**

Let $A$ and $B$ be events in a sample space $S$. If $P(A) \neq 0$, then the **conditional probability of $B$ given $A$**, denoted $P(B \mid A)$, is

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)}. \tag{9.9.1}$$

- Note that when both sides of the formula for conditional probability (formula 9.9.1) are multiplied by $P(A)$, a formula for $P(A \cap B)$ is obtained:

$$P(A \cap B) = P(B \mid A) \cdot P(A). \tag{9.9.2}$$

# Independent Events

- Informally, two events are independent: the probability of one event to happen does not depend on whether the other happens or not.

- Formally, events $A$ and $B$ can be *independent* in the sense that $P(A \mid B) = P(A)$ and $P(B \mid A) = P(B)$.

**• Definition**

If $A$ and $B$ are events in a sample space $S$, then $A$ and $B$ are **independent** if, and only if,

$$P(A \cap B) = P(A) \cdot P(B).$$

# Bayes' Theorem

- Suppose that one urn contains 3 blue and 4 gray balls and a second urn contains 5 blue and 3 gray balls. A ball is selected by choosing one of the urns at random and then picking a ball at random from that urn. If the chosen ball is blue, what is the probability that it came from the first urn?

- This problem can be solved by carefully interpreting all the information that is known and putting it together in just the right way.

- Let $A$ be the event that the chosen ball is blue, $B_1$ the event that the ball came from the first urn, and $B_2$ the event that the ball came from the second urn.

# Bayes' Theorem

- Because 3 of the 7 balls in urn one are blue, and 5 of the 8 balls in urn two are blue,

$$P(A \mid B_1) = \frac{3}{7} \quad \text{and} \quad P(A \mid B_2) = \frac{5}{8}.$$

- And because the urns are equally likely to be chosen,

$$P(B_1) = P(B_2) = \frac{1}{2}.$$

# Bayes' Theorem

- Moreover, by formula (9.9.2),

$$\boxed{P(A \cap B) = P(B \mid A) \cdot P(A).} \qquad 9.9.2$$

$$P(A \cap B_1) = P(A \mid B_1) \cdot P(B_1) = \frac{3}{7} \cdot \frac{1}{2} = \frac{3}{14}, \quad \text{and}$$

$$P(A \cap B_2) = P(A \mid B_2) \cdot P(B_2) = \frac{5}{8} \cdot \frac{1}{2} = \frac{5}{16}.$$

- But *A* is the disjoint union of ($A \cap B_1$) and ($A \cap B_2$), so by probability axiom 3,

$$P(A) = P((A \cap B_1) \cup (A \cap B_2)) = P(A \cap B_1) + P(A \cap B_2) = \frac{3}{14} + \frac{5}{16} = \frac{59}{112}.$$

# Bayes' Theorem

- Finally, by definition of conditional probability,

$$P(B_1 \mid A) = \frac{P(B_1 \cap A)}{P(A)} = \frac{\frac{3}{14}}{\frac{59}{112}} = \frac{336}{826} \cong 40.7\%.$$

- Thus, if the chosen ball is blue, the probability is approximately 40.7% that it came from the first urn.

- The steps used to derive the answer in the previous example can be generalized to prove Bayes' Theorem.

# Bayes' Theorem

- 

**Theorem 9.9.1 Bayes' Theorem**

Suppose that a sample space $S$ is a union of mutually disjoint events $B_1, B_2, B_3, \ldots, B_n$, suppose $A$ is an event in $S$, and suppose $A$ and all the $B_1$ have nonzero probabilities. If $k$ is an integer with $1 \leq k \leq n$, then

$$P(B_k \mid A) = \frac{P(A \mid B_k)P(B_k)}{P(A \mid B_1)P(B_1) + P(A \mid B_2)P(B_2) + \cdots + P(A \mid B_n)P(B_n)}$$

# Example 3 – *Applying Bayes' Theorem*

- Most medical tests occasionally produce incorrect results, called false positives and false negatives.

- When a test is designed to determine whether a patient has a certain disease, a **false positive** result indicates that a patient has the disease when the patient does not have it.

- A **false negative** result indicates that a patient does not have the disease when the patient does have it.

# Basic Statistical Concepts

# Properties of distributions: Center

- There are three main measures of center
  - Mean (M): the arithmetic average
    - Add up all of the scores and divide by the total number
    - Most used measure of center
  - Median (Mdn): the middle score in terms of location
    - The score that cuts off the top 50% of the from the bottom 50%
  - Mode: the most frequent score
    - Good for nominal scales (e.g. eye color)

# The Mean

- The most commonly used measure of center
- The arithmetic average
  - Computing the mean

  – The formula for the population mean is (a parameter):

$$\mu = \frac{\boxed{\sum X}}{\boxed{N}}$$

Divide by the total number in the population

Add up all of the X's

# Spread (Variability)

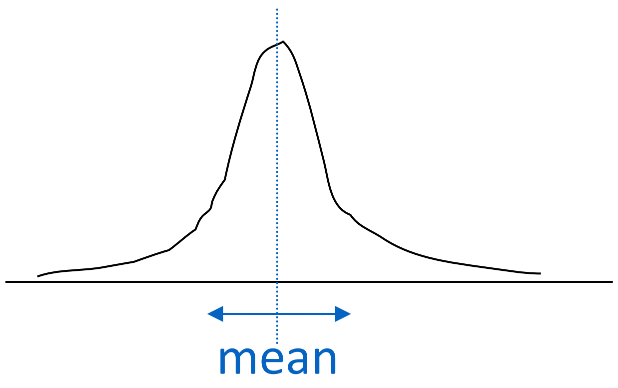
- How similar are the scores?
  - <u>Range</u>: the maximum value - minimum value
    - Only takes two scores from the distribution into account
    - Influenced by extreme values (outliers)
  - <u>Standard deviation (SD)</u>: (essentially) the average amount that the scores in the distribution deviate from the mean
    - Takes all of the scores into account
    - Also influenced by extreme values (but not as much as the range)
  - <u>Variance</u>: standard deviation squared

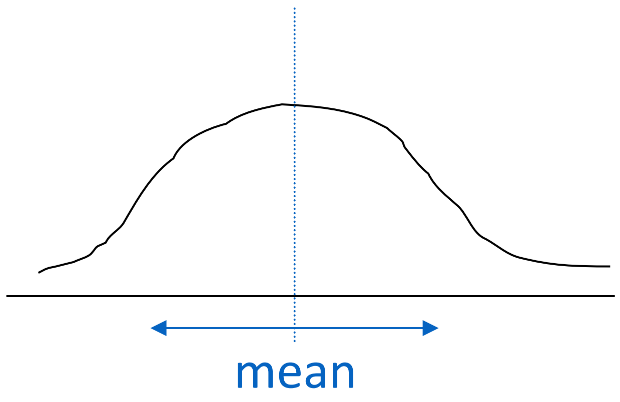
# Variability

- Low variability
  - The scores are fairly similar

High variability

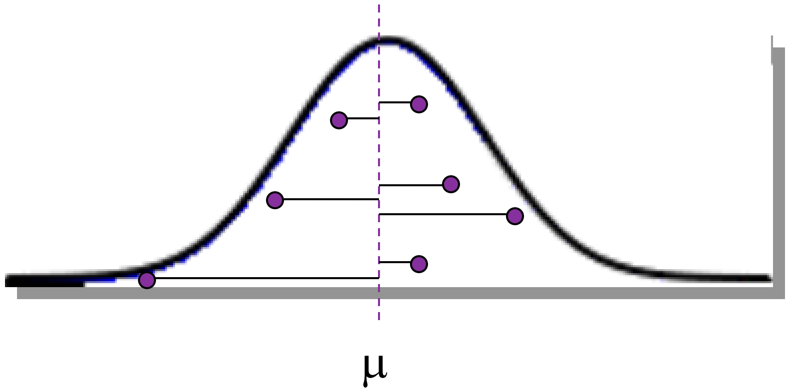
The scores are fairly dissimilar



# Standard deviation

- The **standard deviation** is the most popular and most important measure of variability.
  - The ***standard deviation*** measures how far off all of the individuals in the distribution are from a standard, where that standard is the mean of the distribution.
    - Essentially, the average of the deviations.

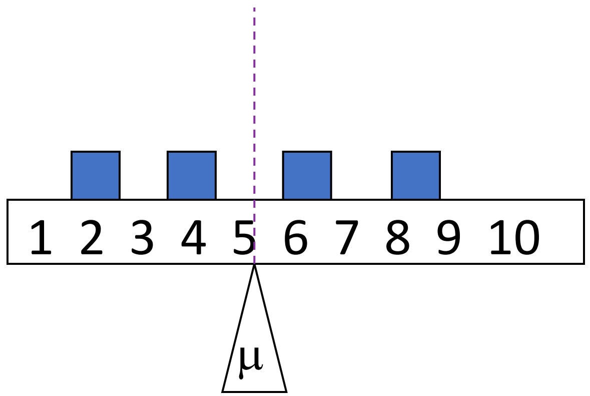


# An Example: Computing the Mean

<u>Our population</u>

2, 4, 6, 8

$$\mu = \frac{\sum X}{N} = \frac{2+4+6+8}{4} = \frac{20}{4} = 5.0$$

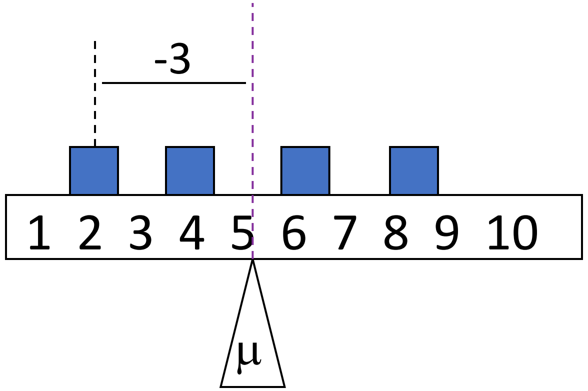


# An Example: Computing Standard Deviation (population)

Our population

2, 4, 6, 8

- ***Step 1:*** To get a measure of the deviation we need to subtract the population mean from every individual in our distribution.




$$\mu = \frac{\Sigma X}{N} = \frac{2+4+6+8}{4} = \frac{20}{4} = 5.0$$

X - μ = deviation scores

2 - 5 = -3

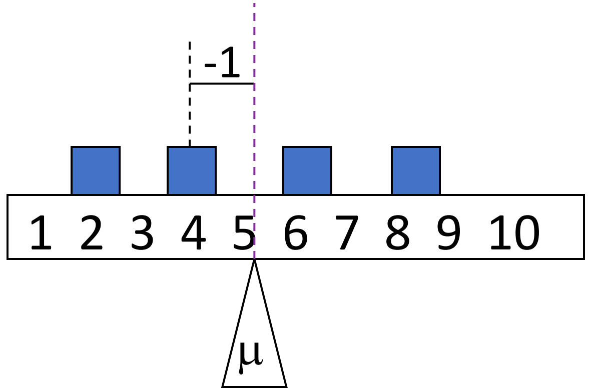
# An Example: Computing Standard Deviation (population)

Our population

2, 4, 6, 8

- ***Step 1:*** To get a measure of the deviation we need to subtract the population mean from every individual in our distribution.

$$\mu = \frac{\Sigma X}{N} = \frac{2+4+6+8}{4} = \frac{20}{4} = 5.0$$




X - μ = deviation scores

2 - 5 = -3

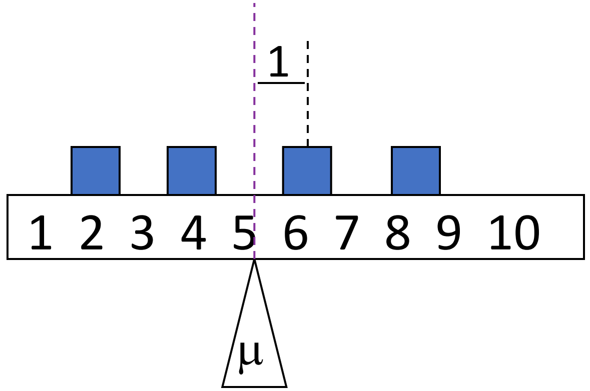
4 - 5 = -1

# An Example: Computing Standard Deviation (population)

<u>Our population</u>

2, 4, 6, 8

- ***Step 1:*** To get a measure of the deviation we need to subtract the population mean from every individual in our distribution.




$$\mu = \frac{\Sigma X}{N} = \frac{2+4+6+8}{4} = \frac{20}{4} = 5.0$$

$X - \mu$ = deviation scores

$2 - 5 = -3$

$6 - 5 = +1$

$4 - 5 = -1$

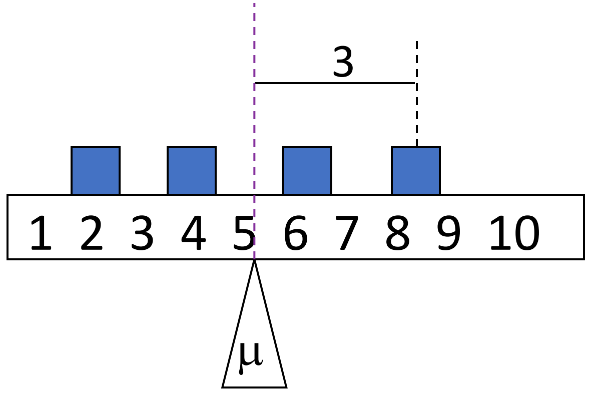
# An Example: Computing Standard Deviation (population)

Our population

2, 4, 6, 8

- ***Step 1:*** To get a measure of the deviation we need to subtract the population mean from every individual in our distribution.

$$\mu = \frac{\Sigma X}{N} = \frac{2+4+6+8}{4} = \frac{20}{4} = 5.0$$




X - μ = deviation scores

2 - 5 = -3

4 - 5 = -1

6 - 5 = +1

8 - 5 = +3

Notice that if you add up all of the deviations they must equal 0.

# An Example: Computing Standard Deviation (population)

- ***Step 2***: So what we have to do is get rid of the negative signs. We do this by squaring the deviations and then taking the square root of the **sum of the squared deviations (SS)**.

$X - \mu$ = deviation scores

$2 - 5 = -3$

$4 - 5 = -1$

$6 - 5 = +1$

$8 - 5 = +3$

$SS = \sum (X - \mu)^2$

$= (-3)^2 + (-1)^2 + (+1)^2 + (+3)^2$

$= 9 + 1 + 1 + 9 = 20$

# An Example: Computing Standard Deviation (population)

- ***Step 3***: Compute ***Variance*** (which is simply the average of the squared deviations (SS))
  - So to get the mean, we need to divide by the number of individuals in the population.

variance = $\sigma^2$ = SS/N

SS = 20, N = 4

$\sigma^2$ = 20/4 = 5.0

# An Example: Computing Standard Deviation (population)

- ***Step 4:*** Compute **Standard Deviation**
  - To get this we need to take the square root of the population variance.

standard deviation = σ = $\sqrt{\sigma^2} = \sqrt{\frac{\sum (X-\mu)^2}{N}}$

$$\sigma = \sqrt{5} = 2.24$$

# An Example: Computing Standard Deviation (population)

- To review:
  - *Step 1*: Compute deviation scores
  - *Step 2*: Compute the SS
  - *Step 3*: Determine the variance
    - Take the average of the squared deviations
    - Divide the SS by the N
  - *Step 4*:  Determine the standard deviation
    - Take the square root of the variance

# An Example: Computing Standard Deviation (SAMPLE)

- To review:
  - *Step 1*: Compute deviation scores
  - *Step 2*: Compute the SS
  - *Step 3*: Determine the variance
    - Take the average of the squared deviations
    - Divide the SS by **(n-1)**
  - *Step 4*: Determine the standard deviation
    - Take the square root of the variance

# Relationships between variables

- Example: Suppose that you notice that the more you study for an exam, the better your score typically is.
  - This suggests that there is a relationship between study time and test performance.
  - We call this relationship a *correlation*.

# Relationships between variables

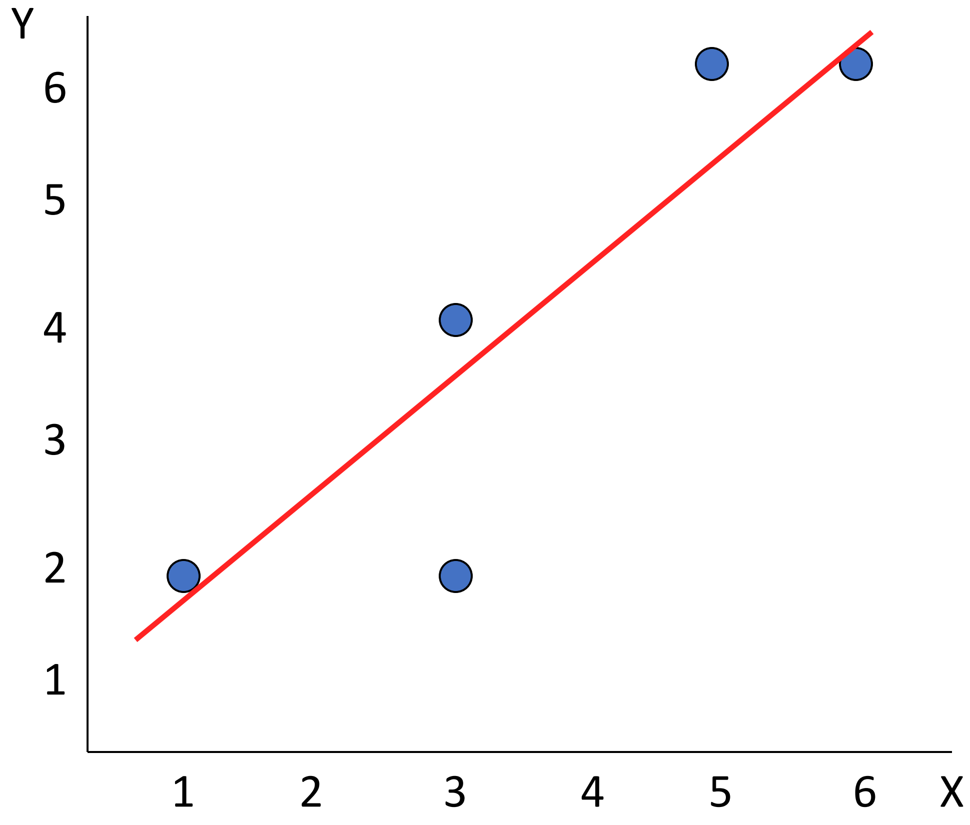
- Properties of a correlation
  - Form (linear or non-linear)
  - Direction (positive or negative)
  - Strength (none, weak, strong, perfect)
- To examine this relationship you should:
  - Make a <u>scatterplot</u>
  - Compute the <u>Correlation Coefficient</u>

# Scatterplot

- Plots one variable against the other
- Useful for “seeing” the relationship
  - Form, Direction, and Strength
- Each point corresponds to a different individual
- Imagine a line through the data points

# Scatterplot

| Hours study X | Exam perf. Y |
|---|---|
| 6 | 6 |
| 1 | 2 |
| 5 | 6 |
| 3 | 4 |
| 3 | 2 |



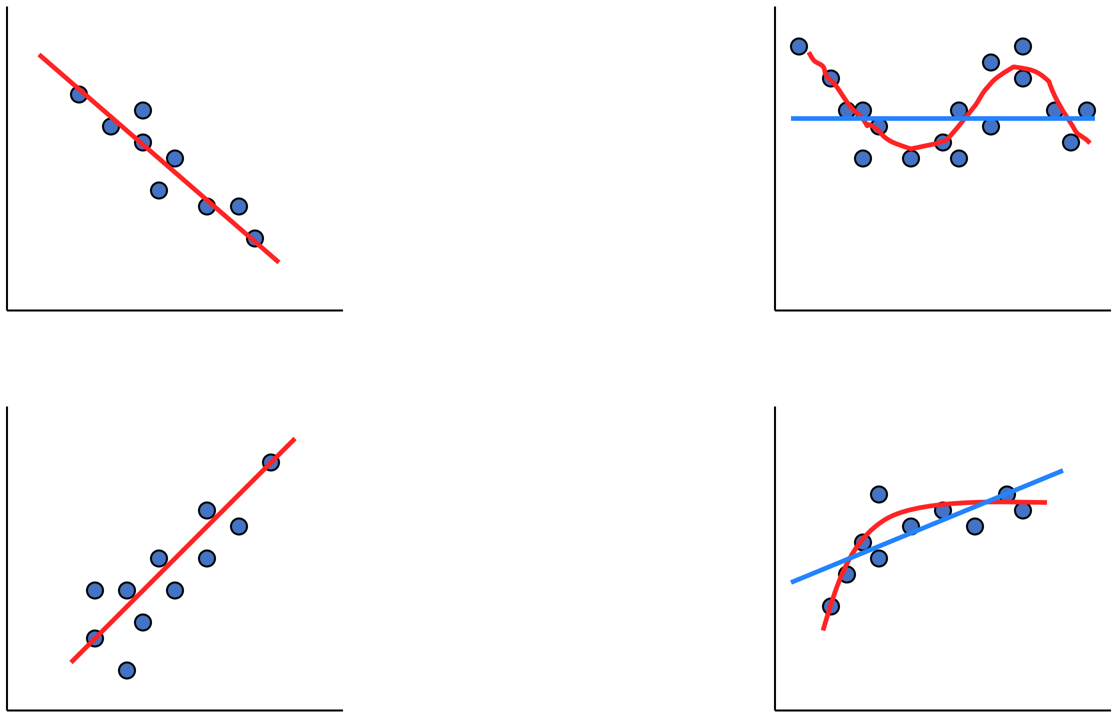
# Correlation Coefficient

- A numerical description of the relationship between two variables
- For relationship between two continuous variables we use Pearson's r
- It basically tells us how much our two variables vary together
    - As X goes up, what does Y typically do
        - X↑, Y↑
        - X↑, Y↓
        - X↑, Y⇳

# Form

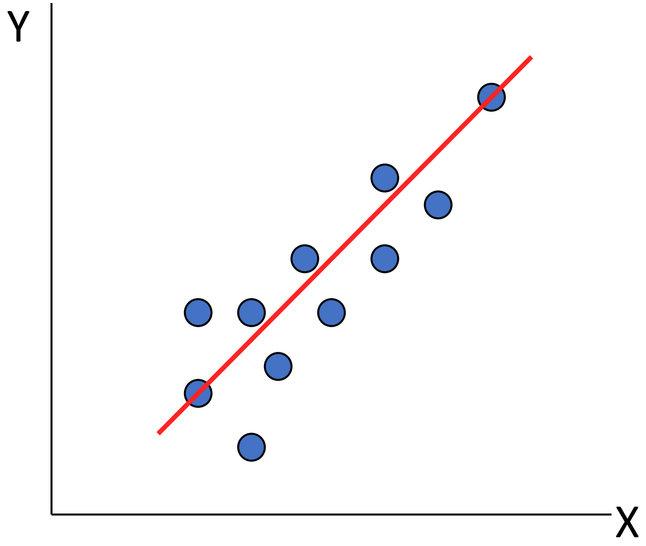
Linear

Non-linear



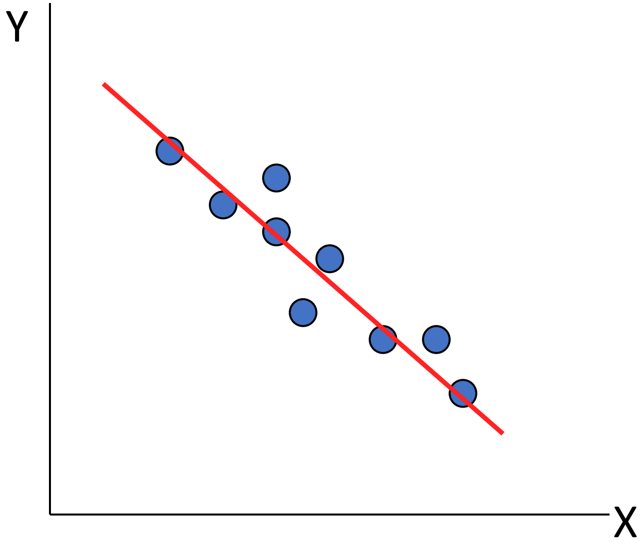
# Direction

## Positive




- As X goes up, Y goes up
- X & Y vary in the same direction
- positive Pearson's r

## Negative

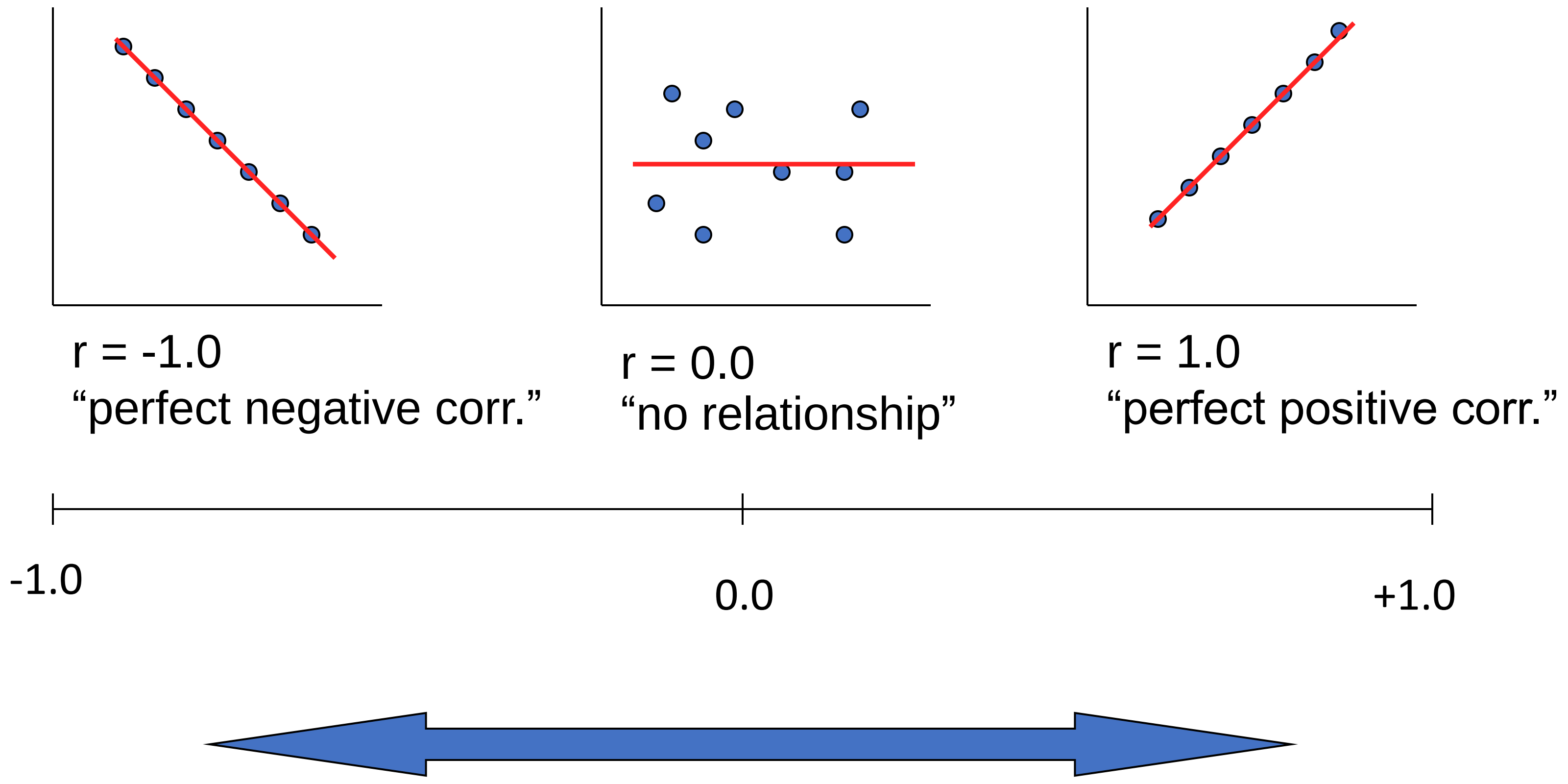



- As X goes up, Y goes down
- X & Y vary in opposite directions
- negative Pearson's r

# Strength

- Zero means “no relationship”.
  - The farther the r is from zero, the stronger the relationship
- The strength of the relationship
  - Spread around the line (note the axis scales)

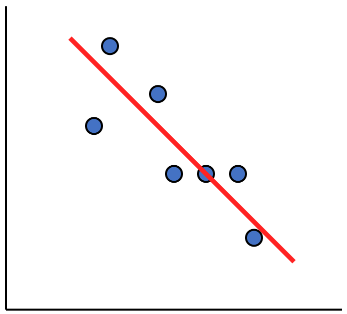
# Strength




The farther from zero, the stronger the relationship

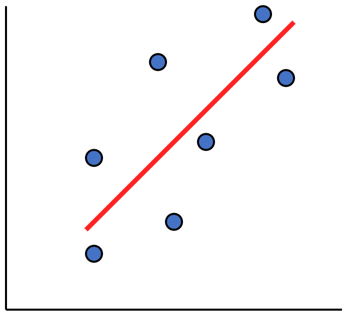
# Strength

Rel A



r = -0.8

Rel B



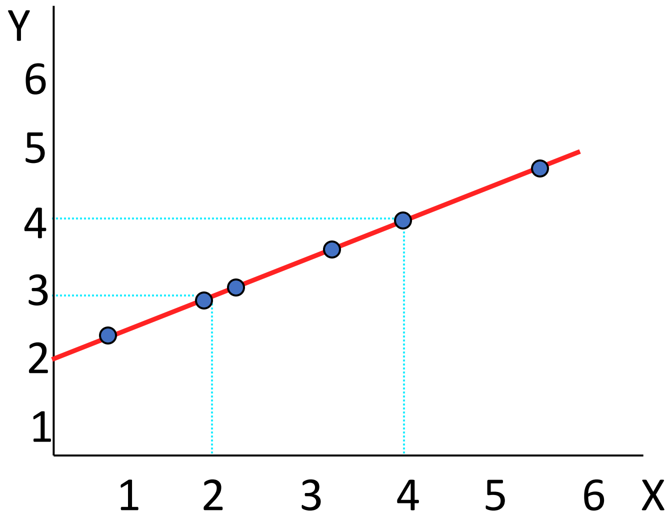
r = 0.5


Which relationship is stronger?

Rel A, -0.8 is stronger than +0.5

# Regression

- Compute the equation for the line that best fits the data points




Y = (X)(slope) + (intercept)

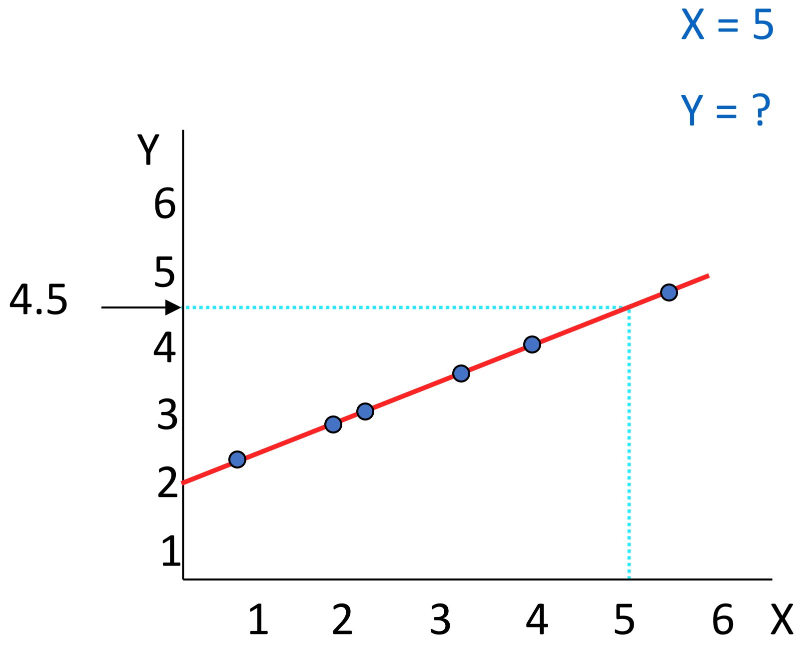
0.5

2.0

$$\frac{\text{Change in Y}}{\text{Change in X}} = \text{slope}$$

# Regression

- Can make specific predictions about Y based on X




$Y = (X)(.5) + (2.0)$

$Y = (5)(.5) + (2.0)$

$Y = 2.5 + 2 = 4.5$