# Linear Algebra Basics

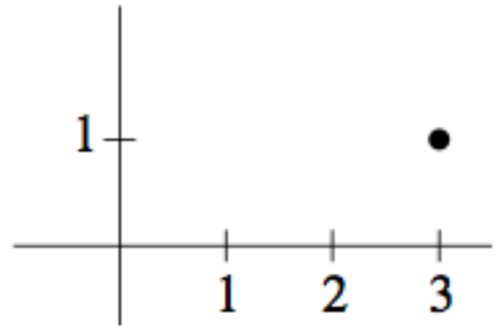
Dr. Naresh Manwani
Assistant Professor, CSE
IIIT-Hyderabad

IIIT, HYDERABAD

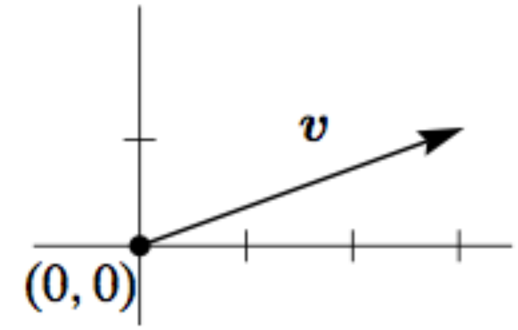
# Vectors

$$v = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

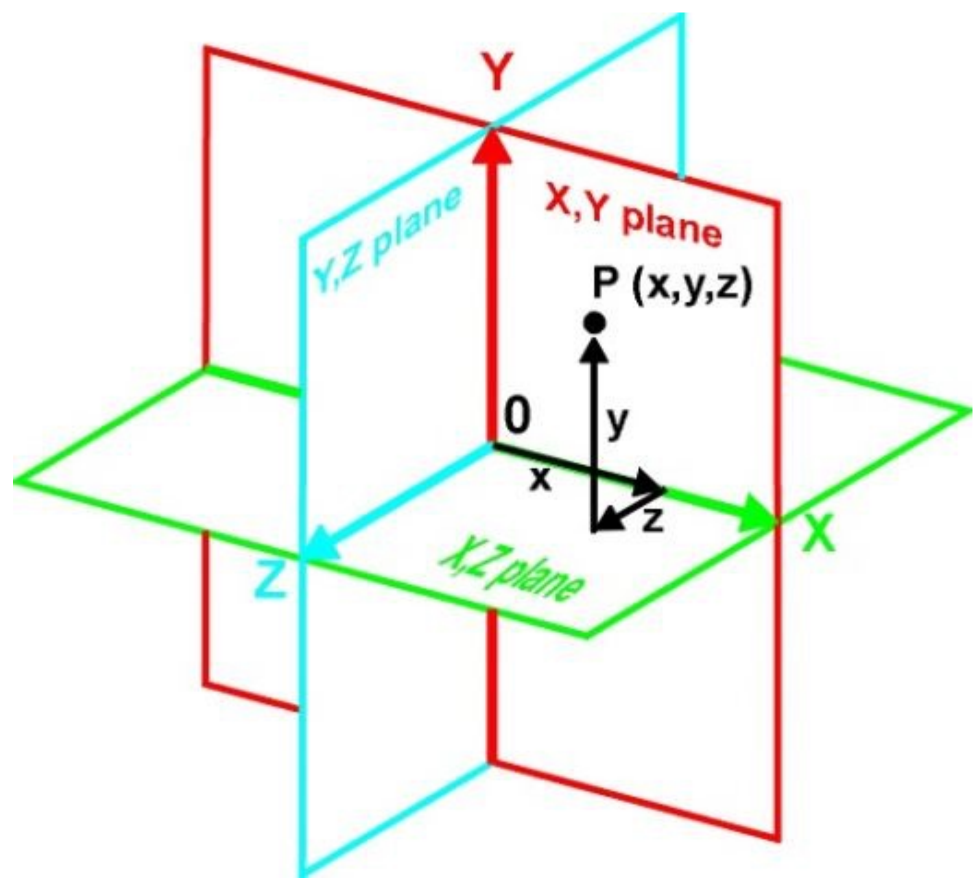
column vector



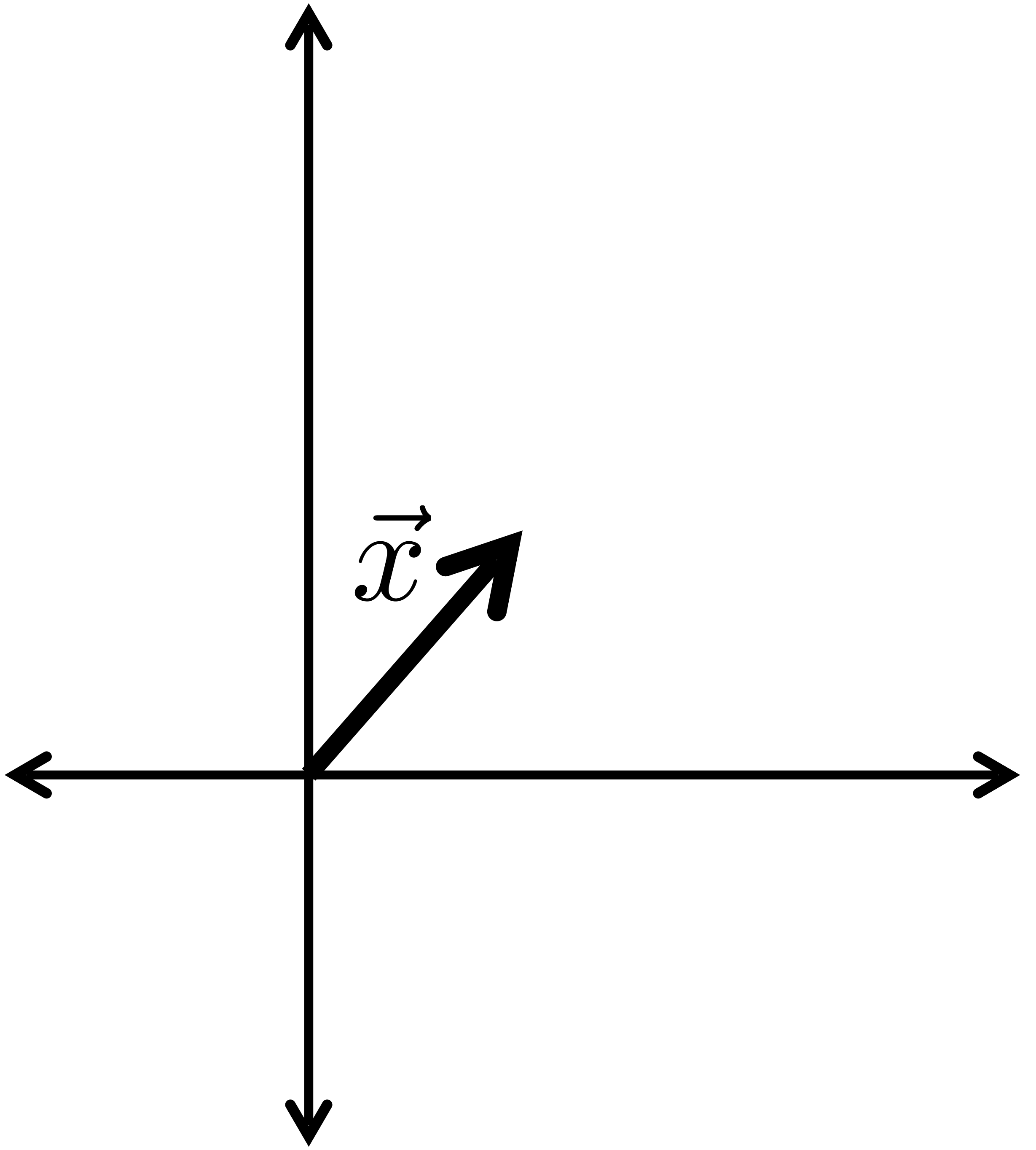

point (3, 1)



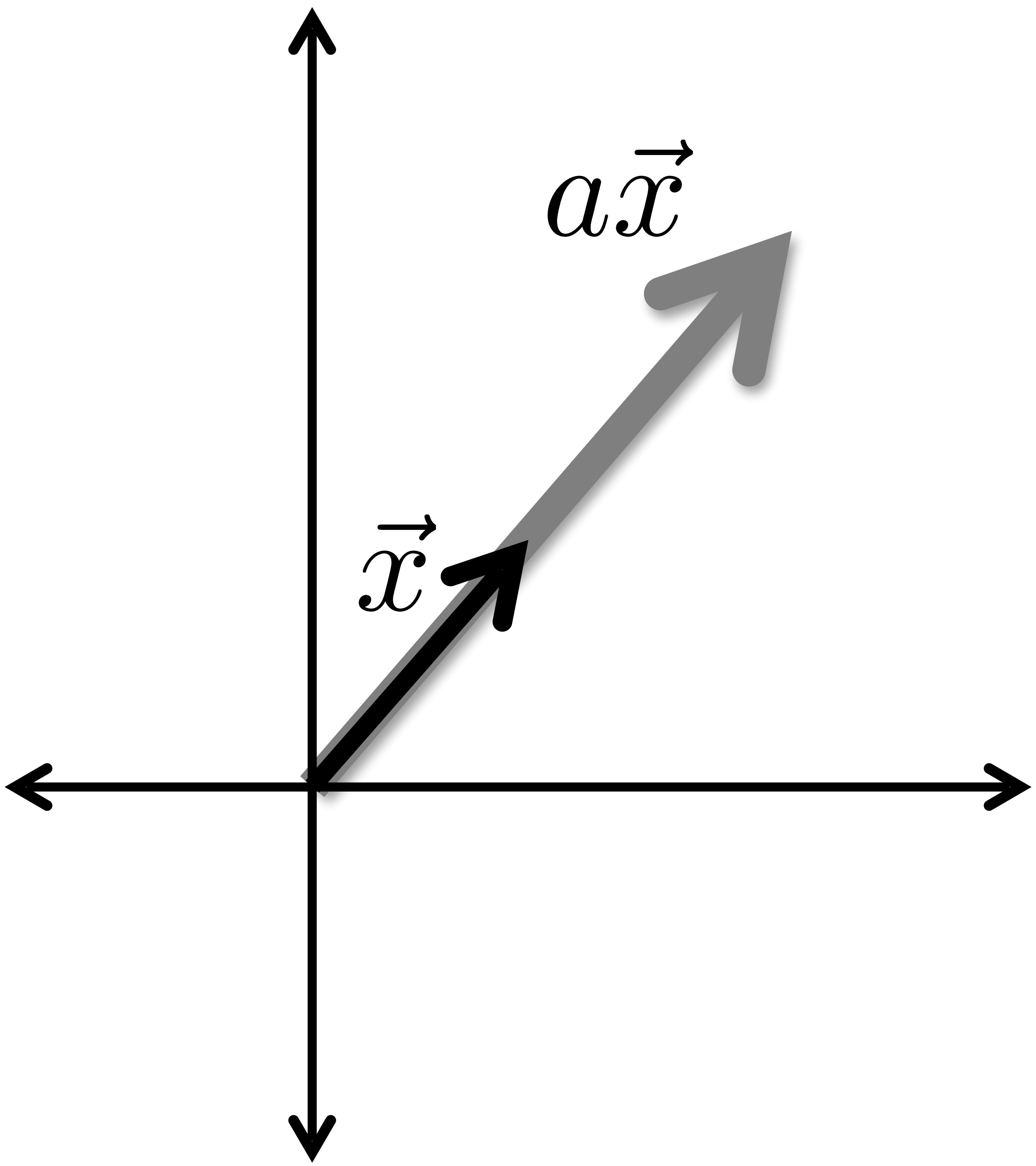

arrow to (3, 1)



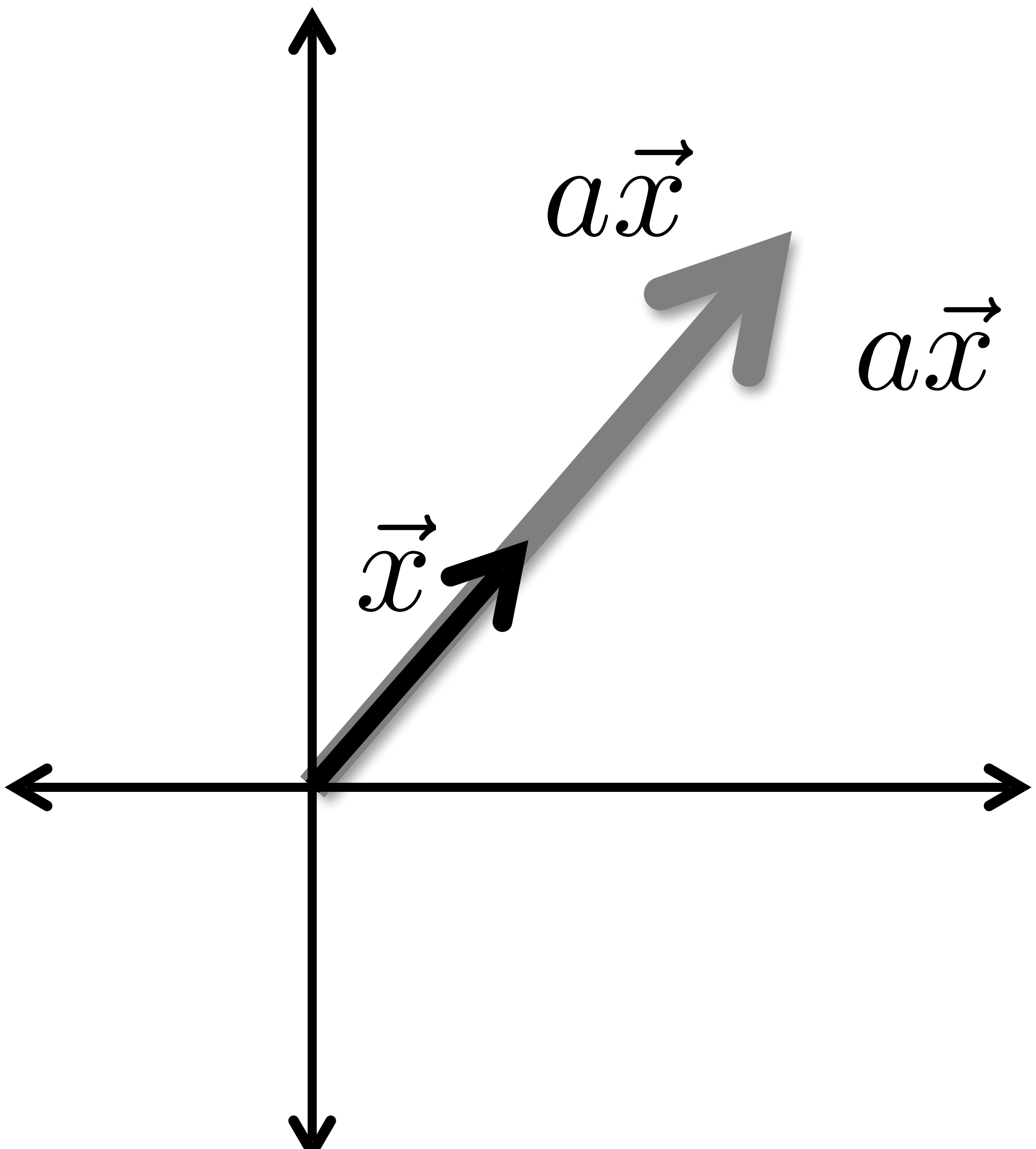
# Scalar times vector



# Scalar times vector



# Scalar times vector




$$a\vec{x} = a \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix} = \begin{pmatrix} ax_1 \\ ax_2 \\ \vdots \\ ax_N \end{pmatrix}$$

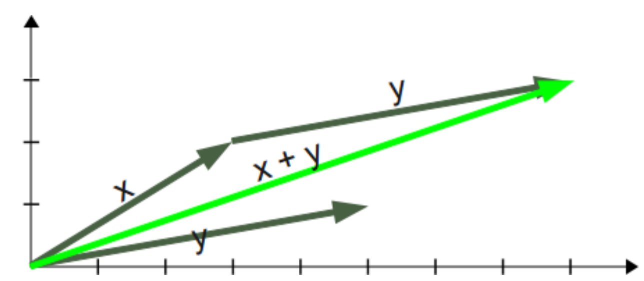
# Vector Addition

Let $\mathbf{x}, \mathbf{y} \in \Re^d$, then

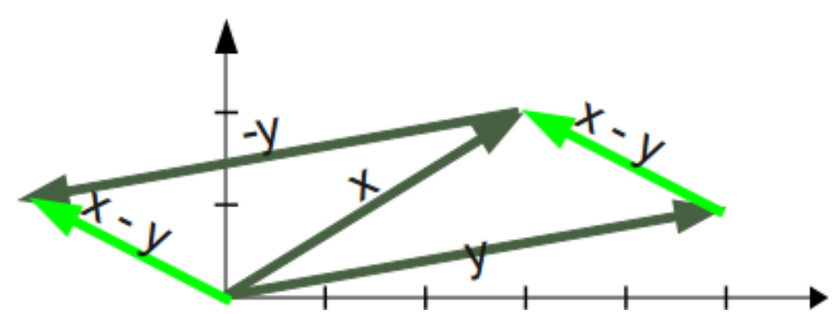
$$\mathbf{x} + \mathbf{y} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_d \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_d + y_d \end{pmatrix}$$

# Example : Vector Addition

- Let $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ and $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$

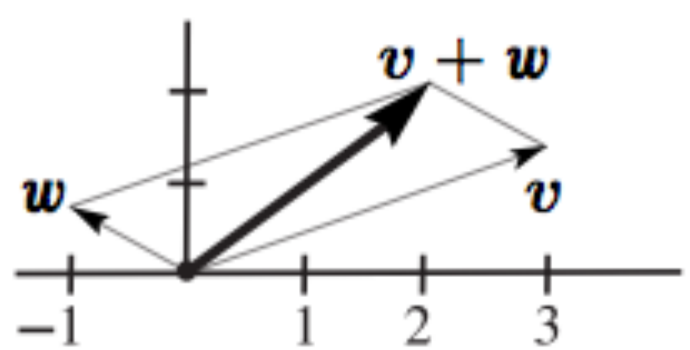



- Then, $\mathbf{x} + \mathbf{y} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \end{pmatrix}$




- Then, $\mathbf{x} - \mathbf{y} = \begin{pmatrix} x_1 - y_1 \\ x_2 - y_2 \end{pmatrix}$

$$\boldsymbol{v} + \boldsymbol{w} = \begin{bmatrix} 3 \\ 1 \end{bmatrix} + \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$



# Scalar Multiplication of a Vector

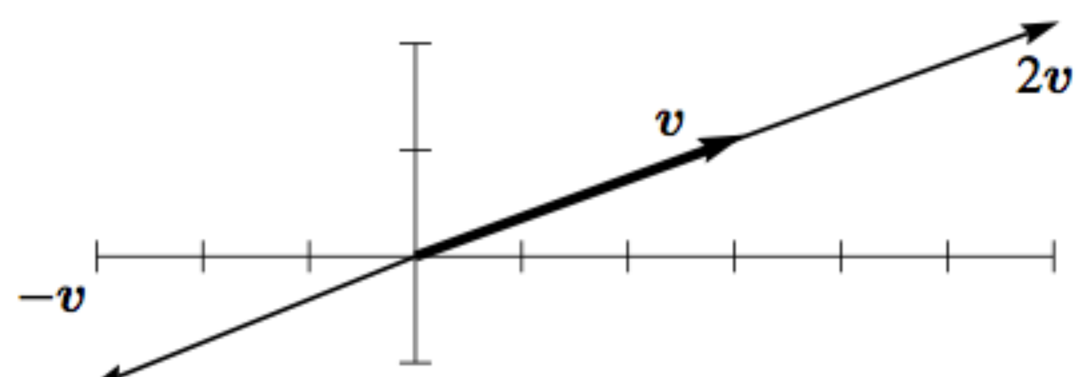
Let $\mathbf{a} \in \Re^d$, then

- When $k > 0$, $\mathbf{a}$ and $k\mathbf{a}$ are in the same direction and $|k\mathbf{a}| = k|\mathbf{a}|$.
- When $k < 0$, $\mathbf{a}$ and $k\mathbf{a}$ are in the opposite direction and $|k\mathbf{a}| = -k|\mathbf{a}|$.

$$\mathbf{v} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$$

$$2\boldsymbol{v} = \begin{bmatrix} 6 \\ 2 \end{bmatrix} \quad -\boldsymbol{v} = \begin{bmatrix} -3 \\ -1 \end{bmatrix}$$

column vectors




arrows to (6, 2) and (−3, −1)

# Linear Combination of Vectors

$$\mathbf{v} = \begin{pmatrix} 3 \\ 1 \end{pmatrix} \text{ and } \mathbf{w} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

| | |
|---|---|
| **Linear combination** $2v + 3w$ | $2\begin{bmatrix} 3 \\ 1 \end{bmatrix} + 3\begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 5 \end{bmatrix}.$ |

# Dot Products

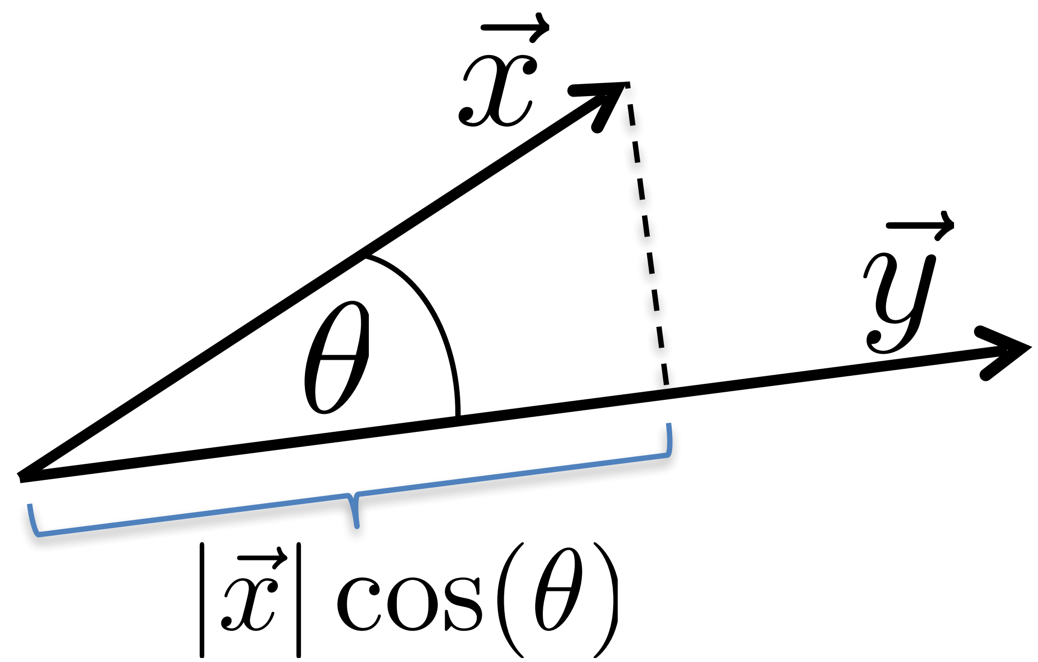
**Definition :** The **dot product** or **inner product** of two vectors $\mathbf{x} \in \Re^d$ and $\mathbf{y} \in \Re^d$ is the number $\mathbf{x}.\mathbf{y}$

$$\mathbf{x}.\mathbf{y} = x_1 y_1 + x_2 y_2 + \ldots x_d y_d$$

**Properties :**

- The dot product $\mathbf{x}.\mathbf{y}$ equals $\mathbf{y}.\mathbf{x}$
- If $\mathbf{x}$ and $\mathbf{y}$ are perpendicular to each other, then $\mathbf{x}.\mathbf{y} = 0$

# Dot product geometric intuition: "Overlap" of 2 vectors




$$\vec{x} \cdot \vec{y} = |\vec{x}||\vec{y}|\cos(\theta)$$

# Matrix times a vector

$$\overrightarrow{y} = \overleftrightarrow{W}\,\overrightarrow{x}$$

# Matrix times a vector

$$\overrightarrow{y} = \overleftrightarrow{W}\,\overrightarrow{x}$$

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_M \end{pmatrix} = \begin{pmatrix} W_{11} & W_{12} & \cdots & W_{1N} \\ W_{21} & W_{22} & \cdots & W_{2N} \\ \vdots & \vdots & & \vdots \\ W_{M1} & W_{M2} & \cdots & W_{MN} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix}$$

# Matrix times a vector

$$\overrightarrow{y} = \overleftrightarrow{W}\,\overrightarrow{x}$$

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_M \end{pmatrix} = \begin{pmatrix} W_{11} & W_{12} & \cdots & W_{1N} \\ W_{21} & W_{22} & \cdots & W_{2N} \\ \vdots & \vdots & & \vdots \\ W_{M1} & W_{M2} & \cdots & W_{MN} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix}$$

M X 1 M X N N X 1

# Matrix times a vector:
# **inner product interpretation**

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_i \\ \vdots \\ y_M \end{pmatrix} = \begin{pmatrix} W_{11} & W_{12} & \cdots & W_{1N} \\ W_{21} & W_{22} & \cdots & W_{2N} \\ \vdots & \vdots & & \vdots \\ W_{i1} & W_{i2} & \cdots & W_{iN} \\ \vdots & \vdots & \ddots & \vdots \\ W_{M1} & W_{M2} & \cdots & W_{MN} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix}$$

- Rule: the $i^{th}$ element of **y** is the dot product of the $i^{th}$ row of W with **x**

# Matrix times a vector:
# **inner product interpretation**

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_i \\ \vdots \\ y_M \end{pmatrix} = \begin{pmatrix} W_{11} & W_{12} & \cdots & W_{1N} \\ W_{21} & W_{22} & \cdots & W_{2N} \\ \vdots & \vdots & & \vdots \\ W_{i1} & W_{i2} & \cdots & W_{iN} \\ \vdots & \vdots & \ddots & \vdots \\ W_{M1} & W_{M2} & \cdots & W_{MN} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix}$$

- Rule: the $i^{th}$ element of **y** is the dot product of the $i^{th}$ row of W with **x**

# Matrix times a vector: **inner product interpretation**

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_i \\ \vdots \\ y_M \end{pmatrix} = \begin{pmatrix} W_{11} & W_{12} & \cdots & W_{1N} \\ W_{21} & W_{22} & \cdots & W_{2N} \\ \vdots & \vdots & & \vdots \\ W_{i1} & W_{i2} & \cdots & W_{iN} \\ \vdots & \vdots & \ddots & \vdots \\ W_{M1} & W_{M2} & \cdots & W_{MN} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix}$$

- Rule: the $i^{th}$ element of **y** is the dot product of the $i^{th}$ row of W with **x**

# Matrix times a vector: **inner product interpretation**

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_i \\ \vdots \\ y_M \end{pmatrix} = \begin{pmatrix} W_{11} & W_{12} & \cdots & W_{1N} \\ W_{21} & W_{22} & \cdots & W_{2N} \\ \vdots & \vdots & & \vdots \\ W_{i1} & W_{i2} & \cdots & W_{iN} \\ \vdots & \vdots & \ddots & \vdots \\ W_{M1} & W_{M2} & \cdots & W_{MN} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix}$$

- Rule: the $i^{th}$ element of **y** is the dot product of the $i^{th}$ row of W with **x**

# Matrix times a vector: **inner product interpretation**

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_i \\ \vdots \\ y_M \end{pmatrix} = \begin{pmatrix} W_{11} & W_{12} & \cdots & W_{1N} \\ W_{21} & W_{22} & \cdots & W_{2N} \\ \vdots & \vdots & & \vdots \\ W_{i1} & W_{i2} & \cdots & W_{iN} \\ \vdots & \vdots & \ddots & \vdots \\ W_{M1} & W_{M2} & \cdots & W_{MN} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix}$$

- Rule: the $i^{th}$ element of **y** is the dot product of the $i^{th}$ row of W with **x**

# Matrix times a vector:
# **outer product interpretation**

$$
\begin{array}{l}
\;\;\overrightarrow{W}^{(1)} \\
\;\;\;\downarrow \\
\begin{pmatrix}
W_{11} & W_{12} & \cdots & W_{1N} \\
W_{21} & W_{22} & \cdots & W_{2N} \\
\vdots & \vdots & & \vdots \\
W_{M1} & W_{M2} & \cdots & W_{MN}
\end{pmatrix}
\begin{pmatrix}
x_1 \\ x_2 \\ \vdots \\ x_N
\end{pmatrix} =
\end{array}
$$

- The product is a weighted sum of the columns of W, weighted by the entries of x

# Matrix times a vector: **outer product interpretation**

$$\begin{array}{c}\overrightarrow{W}^{(1)}\\ \downarrow\end{array}$$

$$\begin{pmatrix} W_{11} & W_{12} & \cdots & W_{1N} \\ W_{21} & W_{22} & \cdots & W_{2N} \\ \vdots & \vdots & & \vdots \\ W_{M1} & W_{M2} & \cdots & W_{MN} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix} = x_1 \overrightarrow{W}^{(1)} +$$

- The product is a weighted sum of the columns of W, weighted by the entries of x

# Matrix times a vector:
# **outer product interpretation**

$$\begin{array}{c} \overrightarrow{W}^{(1)} \\ \downarrow \end{array}$$

$$\begin{pmatrix} W_{11} & W_{12} & \cdots & W_{1N} \\ W_{21} & W_{22} & \cdots & W_{2N} \\ \vdots & \vdots & & \vdots \\ W_{M1} & W_{M2} & \cdots & W_{MN} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix} = x_1 \overrightarrow{W}^{(1)} + x_2 \overrightarrow{W}^{(2)}$$

- The product is a weighted sum of the columns of W, weighted by the entries of x

# Matrix times a vector:
# **outer product interpretation**

$$\begin{array}{l} \;\;\overrightarrow{W}^{(1)} \\ \;\;\;\;\downarrow \\ \begin{pmatrix} W_{11} & W_{12} & \cdots & W_{1N} \\ W_{21} & W_{22} & \cdots & W_{2N} \\ \vdots & \vdots & & \vdots \\ W_{M1} & W_{M2} & \cdots & W_{MN} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix} = x_1 \overrightarrow{W}^{(1)} + x_2 \overrightarrow{W}^{(2)} + \cdots + x_N \overrightarrow{W}^{(N)} \end{array}$$

- The product is a weighted sum of the columns of W, weighted by the entries of x

# Example of the outer product method

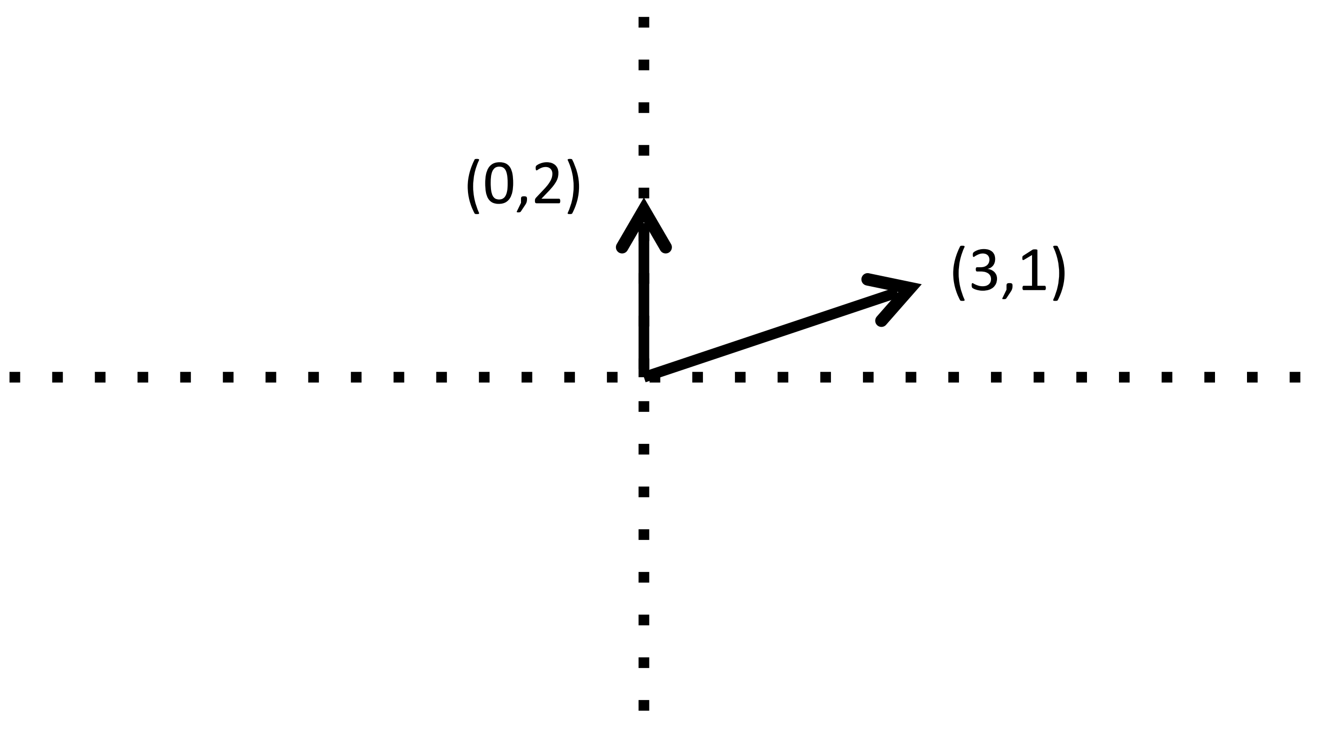
$$\overleftrightarrow{M}$$

$$\begin{pmatrix} 0 & 3 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} =$$

# Example of the outer product method

$\overleftrightarrow{M}$

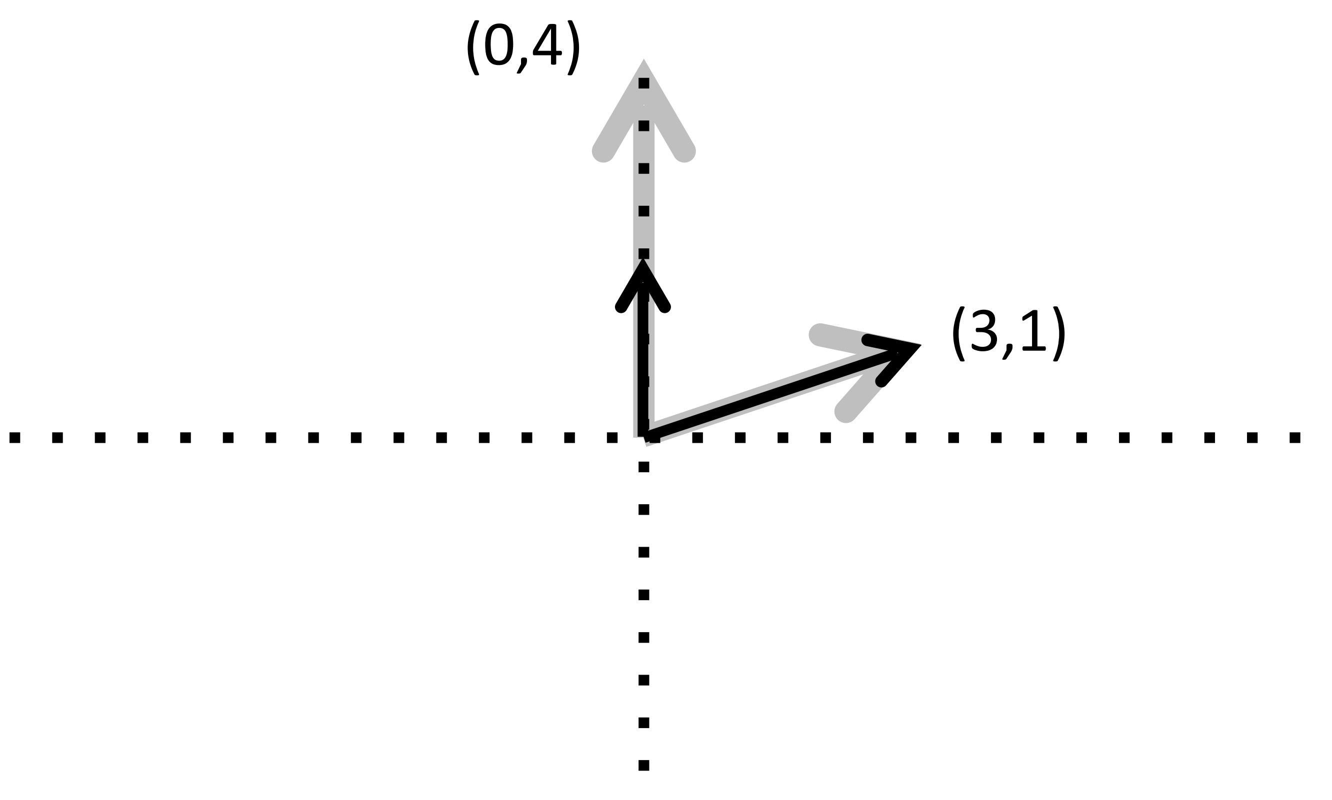
$$\begin{pmatrix} 0 & 3 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = 2 \begin{pmatrix} 0 \\ 2 \end{pmatrix} + 1 \begin{pmatrix} 3 \\ 1 \end{pmatrix}$$



# Example of the outer product method

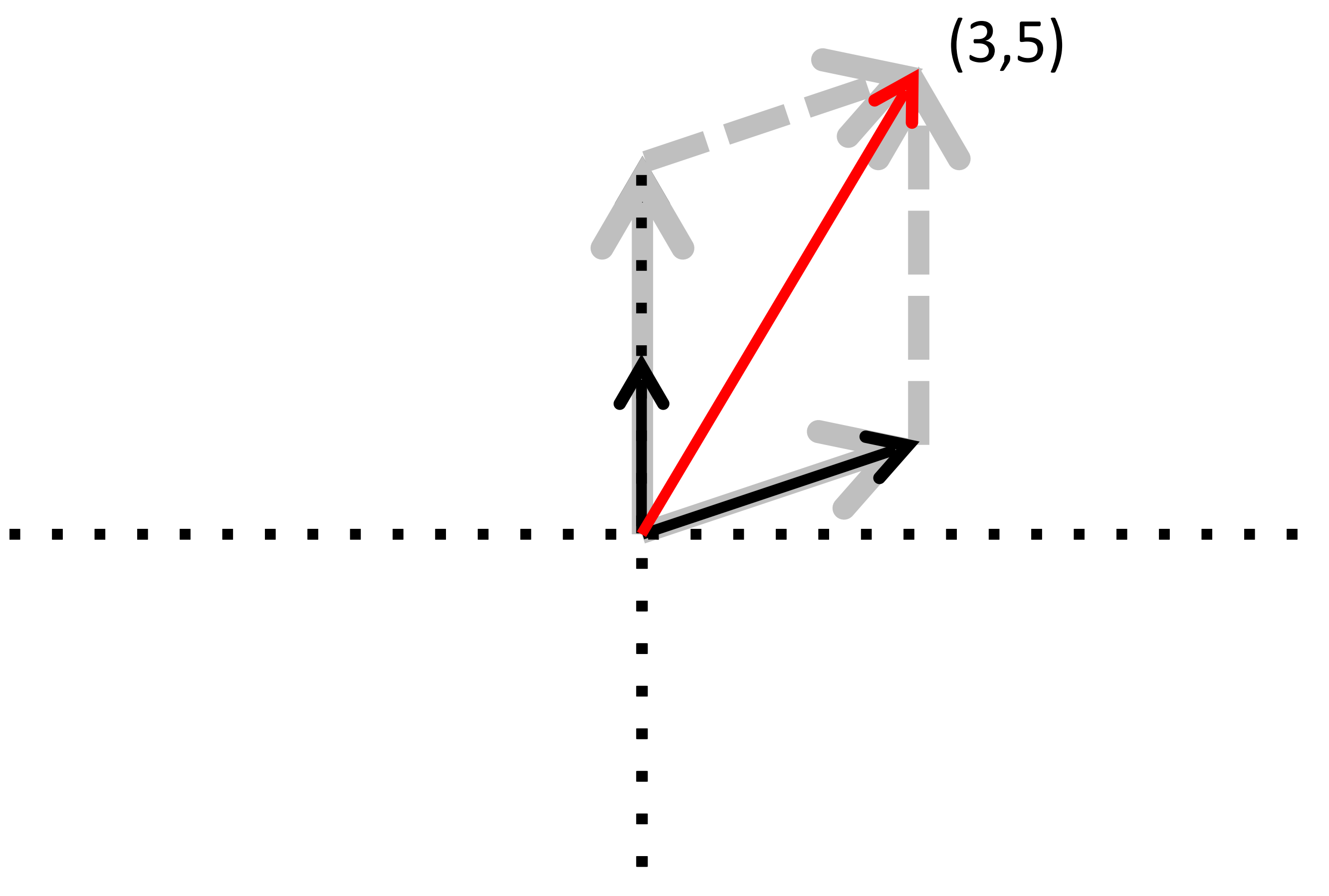
$$\overleftrightarrow{M}$$

$$\begin{pmatrix} 0 & 3 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = 2 \begin{pmatrix} 0 \\ 2 \end{pmatrix} + 1 \begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 4 \end{pmatrix} + \begin{pmatrix} 3 \\ 1 \end{pmatrix}$$



# Example of the outer product method

$$\overbrace{\begin{pmatrix}0 & 3\\ 2 & 1\end{pmatrix}}^{M}\begin{pmatrix}2\\ 1\end{pmatrix} = 2\begin{pmatrix}0\\ 2\end{pmatrix} + 1\begin{pmatrix}3\\ 1\end{pmatrix} = \begin{pmatrix}0\\ 4\end{pmatrix} + \begin{pmatrix}3\\ 1\end{pmatrix} = \begin{pmatrix}3\\ 5\end{pmatrix}$$




- Note: different combinations of the columns of **M** can give you any vector in the plane

(we say the columns of **M** "span" the plane)

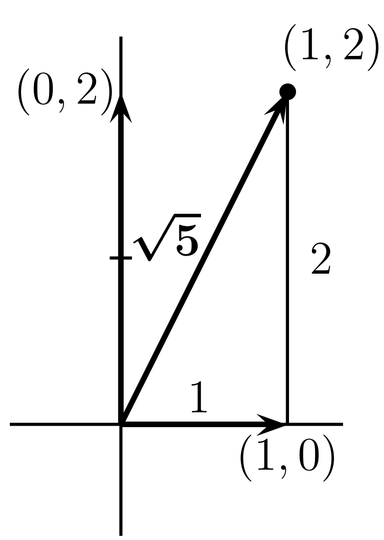
# Norm of a Vector

**Definition :** For $\mathbf{x} \in \Re^d$, the **norm** of $\mathbf{x}$ is defined as

$$||\mathbf{x}|| = (x_1^2 + x_2^2 + \dots x_d^2)^{\frac{1}{2}}$$

**Properties :**

- $||\alpha \mathbf{x}|| = |\alpha| ||\mathbf{x}||$, for any $\alpha \in \Re$
- $||\mathbf{x} + \mathbf{y}|| \leq ||\mathbf{x}|| + ||\mathbf{y}||$ (triangle inequality)
- $||\mathbf{x}|| = 0 \Rightarrow \mathbf{x} = 0$

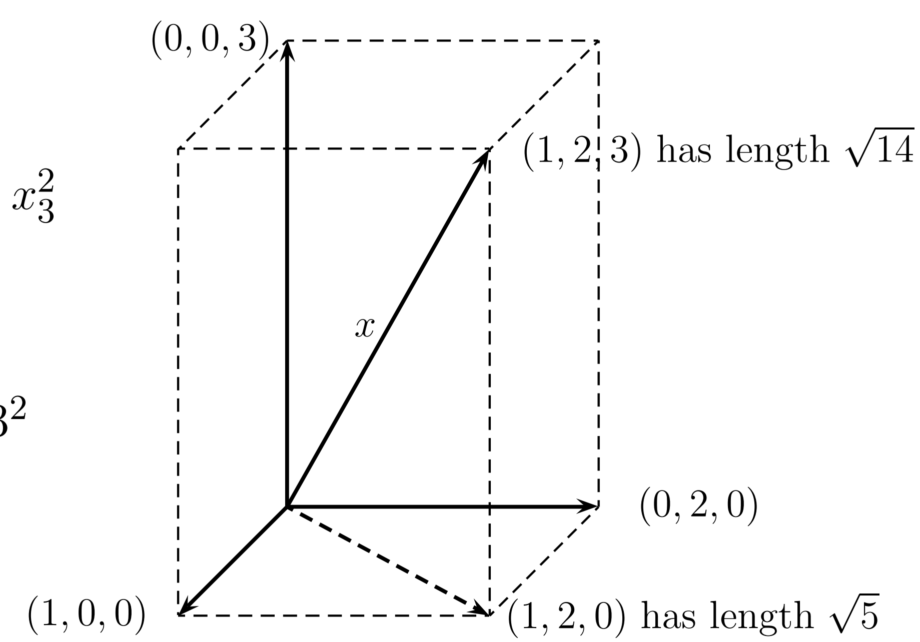


(a)

$$\|x\|^2 = x_1^2 + x_2^2 + x_3^2$$

$$5 = 1^2 + 2^2$$

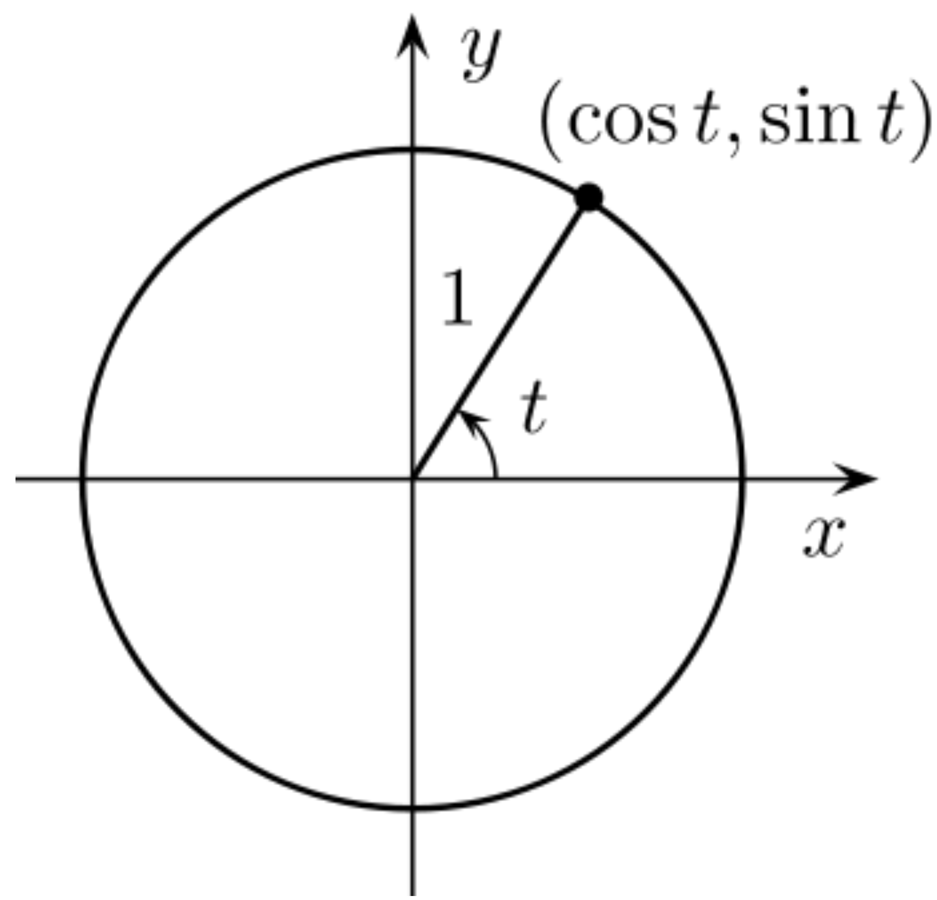
$$14 = 1^2 + 2^2 + 3^2$$



(b)

# Unit Vectors

- A **unit vector u** is a vector whose length equals one. Then $\mathbf{u}.\mathbf{u} = 1$.
- $\mathbf{u} = \frac{\mathbf{v}}{||\mathbf{v}||}$ is a unit vector in the same direction as $\mathbf{v}$.



# Cauchy - Schwarz Inequality

Let $\mathbf{a}, \mathbf{b} \in \Re^d$, then

$$\mathbf{a}.\mathbf{b} \leq ||\mathbf{a}||.||\mathbf{b}||$$

and the equality holds only when either $\mathbf{a}$ or $\mathbf{b}$ is a multiple of the other.

**Proof:**

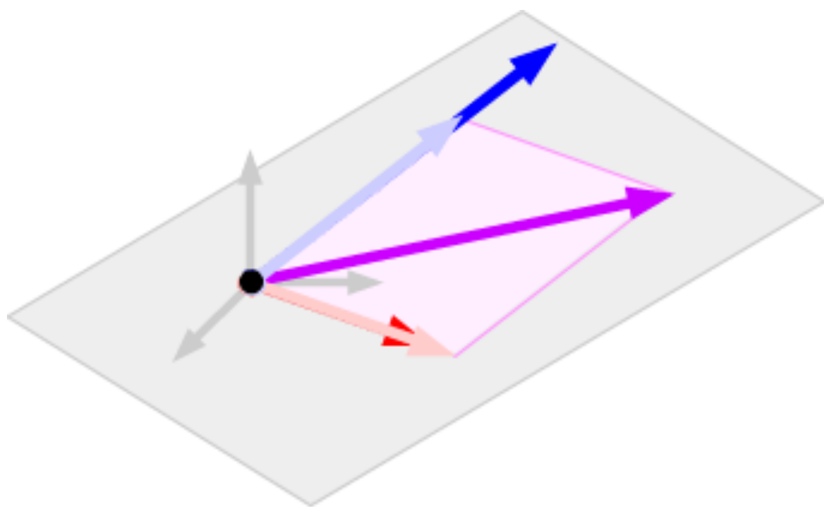
- If $\mathbf{a} = 0$, then equality holds. Thus, lets assume $\mathbf{a}$ is nonzero.
- We also assume that $\mathbf{a}.\mathbf{b} \neq 0$, otherwise the inequality is obviously true (neither $||\mathbf{a}||$ nor $||\mathbf{b}||$ can be negative)
- Let $\mathbf{c} = \mathbf{a} - \frac{\mathbf{a}.\mathbf{b}}{\mathbf{b}.\mathbf{b}}\mathbf{b}$. Then, $\mathbf{c}.\mathbf{b} = 0$ ($\mathbf{c}$ is orthogonal to $\mathbf{b}$).
- Applying Pythagorean Theorem to $\mathbf{a} = \frac{\mathbf{a}.\mathbf{b}}{\mathbf{b}.\mathbf{b}}\mathbf{b} + \mathbf{c}$

$$||\mathbf{a}||^2 = ||\mathbf{c}||^2 + \left|\frac{\mathbf{a}.\mathbf{b}}{\mathbf{b}.\mathbf{b}}\right|^2 ||\mathbf{b}||^2 = ||\mathbf{c}||^2 + \frac{|\mathbf{a}.\mathbf{b}|^2}{||\mathbf{b}||^2} \geq \frac{|\mathbf{a}.\mathbf{b}|^2}{||\mathbf{b}||^2}$$

- Thus, $\mathbf{a}.\mathbf{b} \leq ||\mathbf{a}||.||\mathbf{b}||$
- If $\mathbf{a}.\mathbf{b} = ||\mathbf{a}||.||\mathbf{b}||$, then $||\mathbf{c}||^2 = 0$. Hence, $\mathbf{c} = 0$. Which means $\mathbf{a} = \frac{\mathbf{a}.\mathbf{b}}{\mathbf{b}.\mathbf{b}}\mathbf{b}$

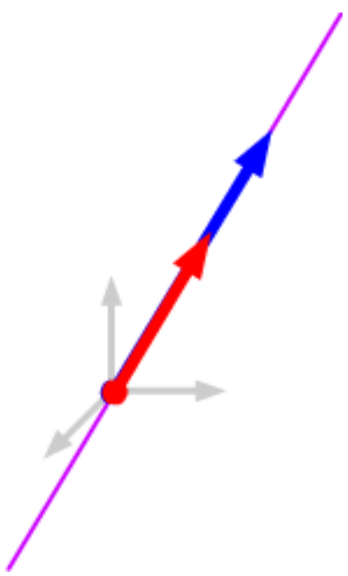
# Spanning Space of Vectors

- **Vectors are not parallel**: there is a single plane through the origin containing them. Any linear combination of those vectors lies in that plane and any vector in that



$$\mathbf{z} = a\mathbf{x} + b\mathbf{y},\ \ a, b \in \Re$$

- **Vectors are parallel**: any linear combination of them must be parallel to both, so as before, the span of two parallel vectors consists of the line through the origin which contains the two vectors.



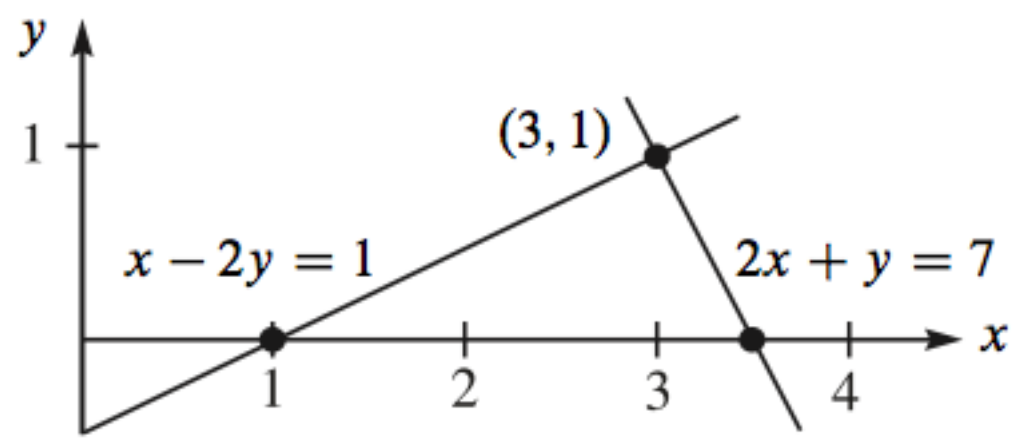
$$\text{if } \mathbf{x} = c\mathbf{y};\quad c \in \Re,$$
$$\text{then } \mathbf{z} = a\mathbf{x} + b\mathbf{y} = (ac + b)\mathbf{y}$$

If there are 3 linearly independent vectors in 3D, then
What will be their spanning space ?

# System of Linear Equations

# Example

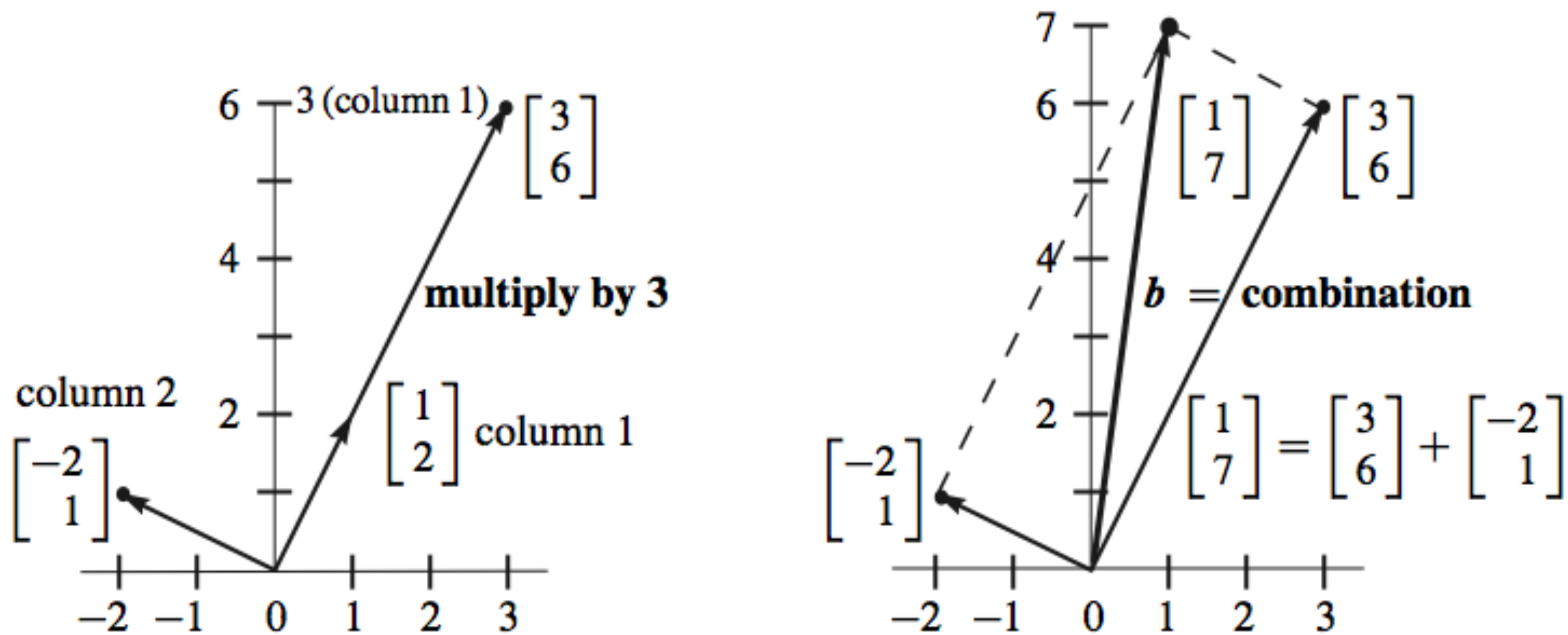
$$\begin{array}{llcr} \textbf{Two equations} & x - 2y &=& 1 \\ \textbf{Two unknowns} & 2x + y &=& 7 \end{array}$$




**ROWS** ***The row picture shows two lines meeting at a single point (the solution).***

# Column Space View

$$x\begin{bmatrix}1\\2\end{bmatrix}+y\begin{bmatrix}-2\\1\end{bmatrix}=\begin{bmatrix}1\\7\end{bmatrix}=\boldsymbol{b}.$$



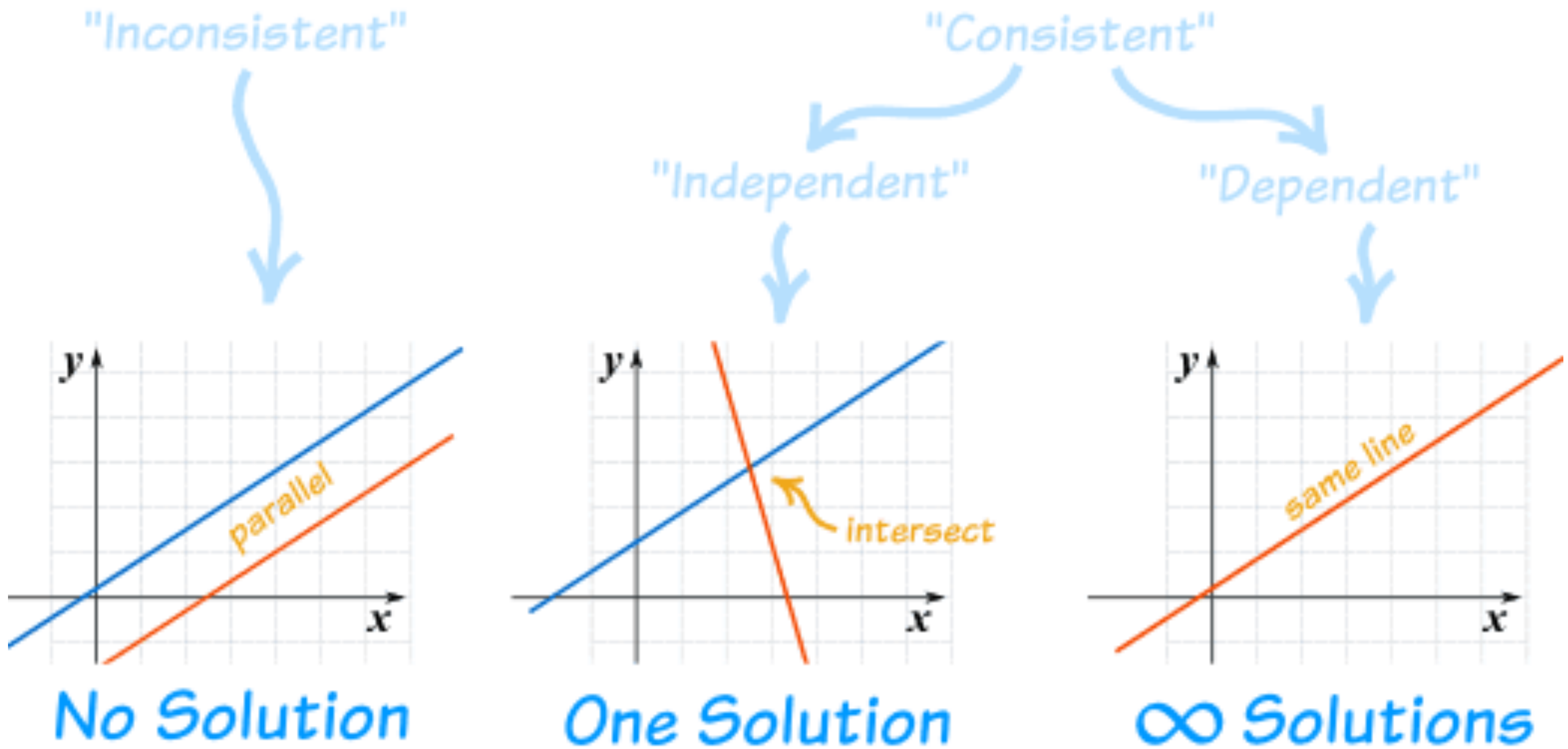

**COLUMNS** *The column picture combines the column vectors on the left side of the equations to produce the vector b on the right side.*

# System of Linear Equations in 2D

$$
\begin{aligned}
A\mathbf{x} &= \quad b \\
\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}
\end{aligned}
$$

$$
\begin{aligned}
a_{11}x + a_{12}y &= b_1 \\
a_{21}x + a_{22}y &= b_2
\end{aligned}
$$

# Row Space View

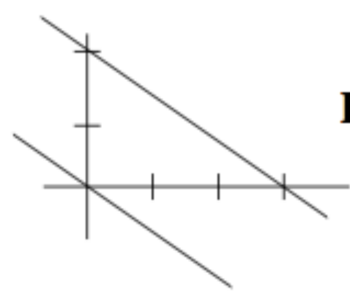



| | |
|---|---|
| No Solution | $[a_{11} \quad a_{12}] = c[a_{21} \quad a_{22}]$ and $b_1 \neq b_2$ |
| Unique Solution | $[a_{11} \quad a_{12}]$ and $[a_{21} \quad a_{22}]$ are linear independent |
| Infinite Solution | $[a_{11} \quad a_{12} \quad b_1] = c[a_{21} \quad a_{22} \quad b_2]$ |

# No Solution

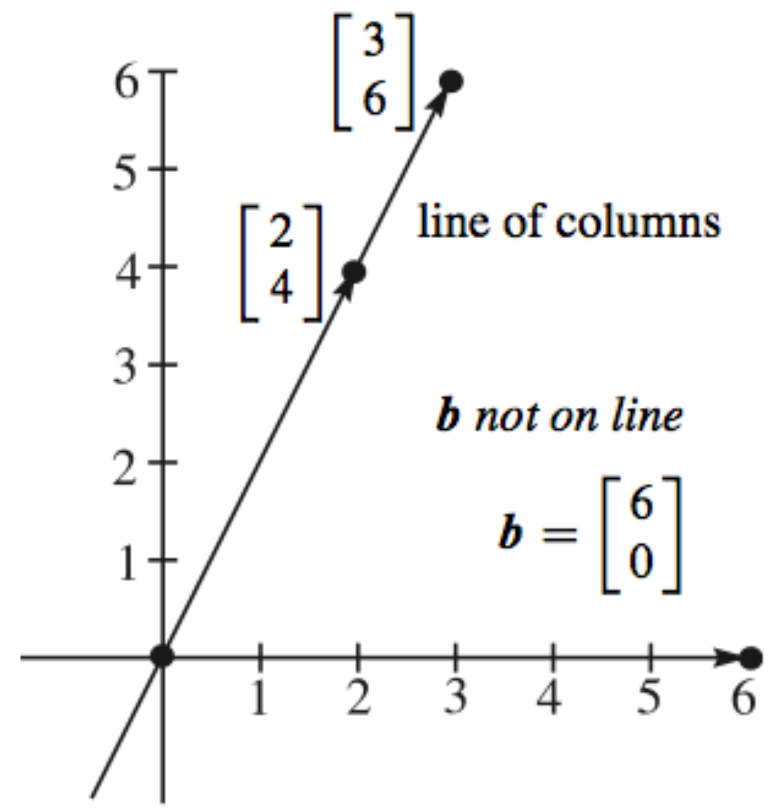
## Row Space View

$$\begin{aligned} 2v_1 - 3v_2 &= 6 \\ 4v_1 - 6v_2 &= 0 \end{aligned}$$



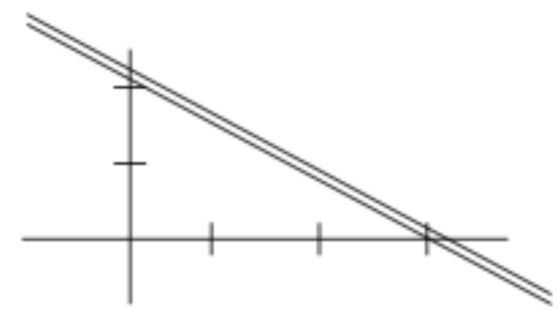
**Parallel lines**
**no solution**

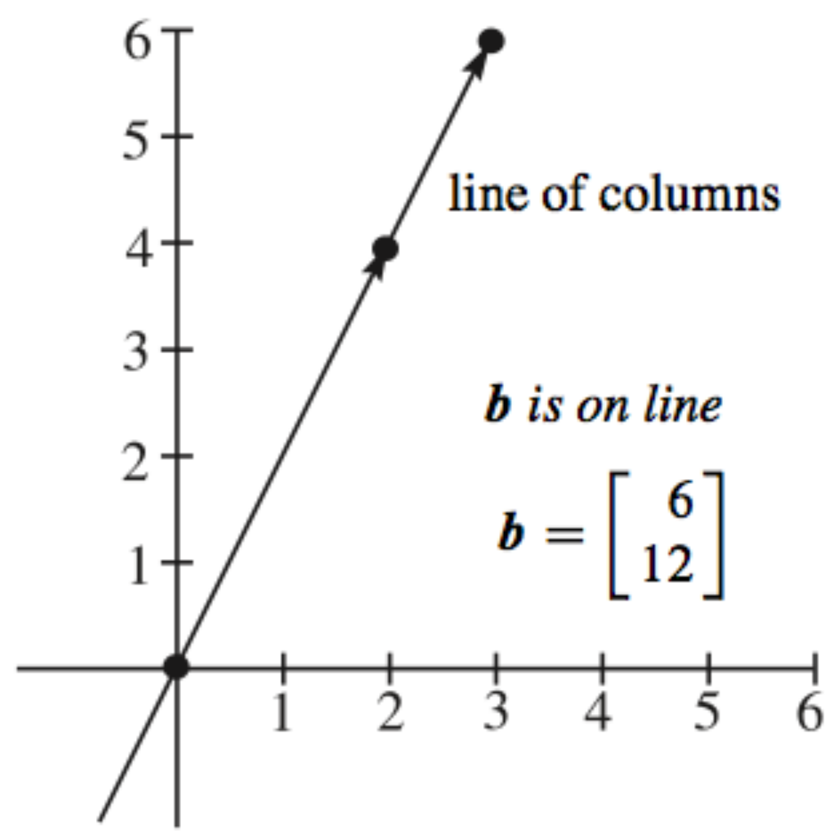
## Column Space View



# Infinite Solutions

$$A = \begin{bmatrix} 2 & 3 \\ 4 & 6 \end{bmatrix} \qquad \begin{aligned} 2v_1 - 3v_2 &= \mathbf{6} \\ 4v_1 - 6v_2 &= \mathbf{12} \end{aligned}$$





# System of Linear Equations in 3D

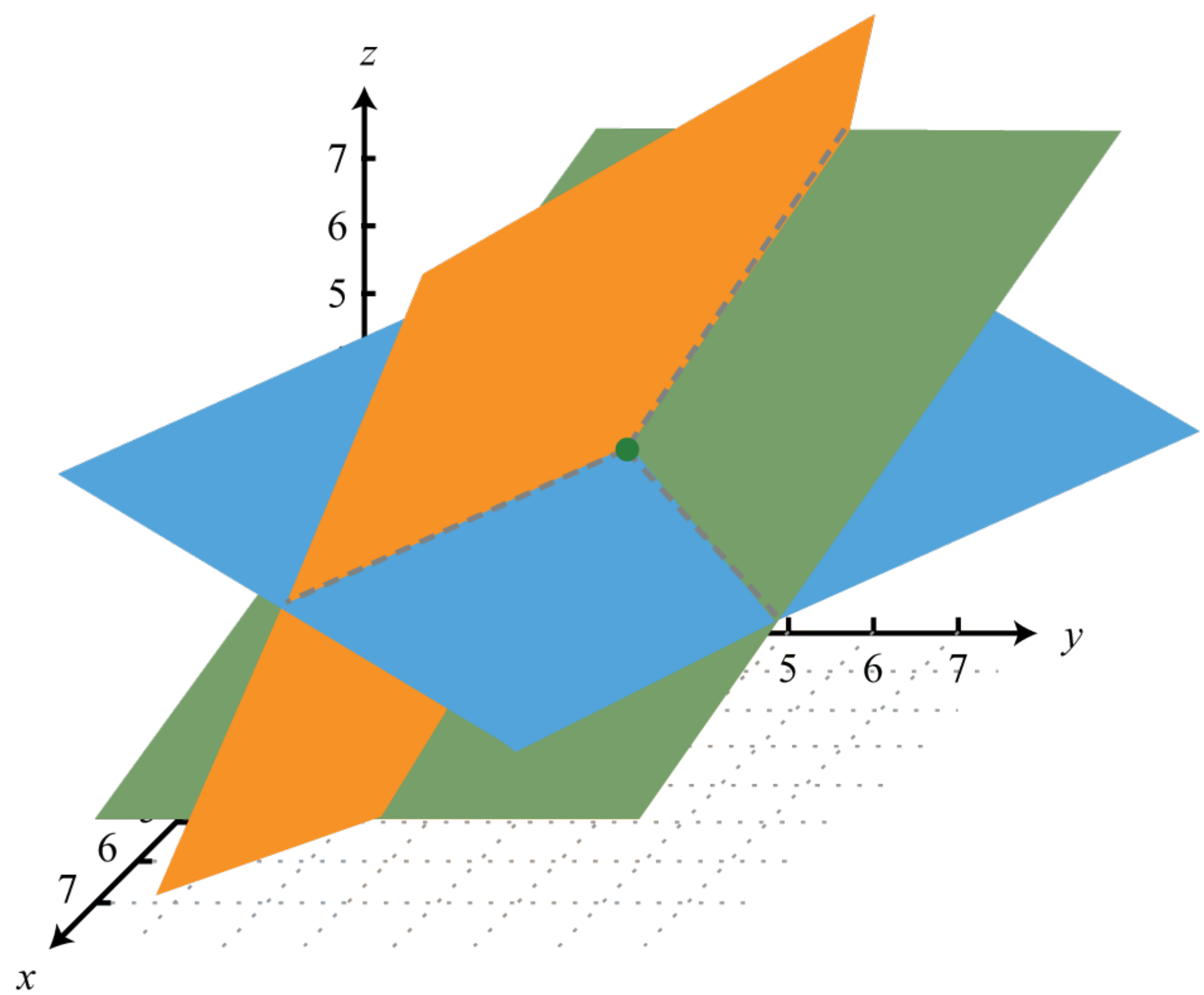
$$A\mathbf{x} = \mathbf{b}$$

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

$$\begin{pmatrix} \mathbf{a_1}^T \\ \mathbf{a_2}^T \\ \mathbf{a_3}^T \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

# Unique Solution

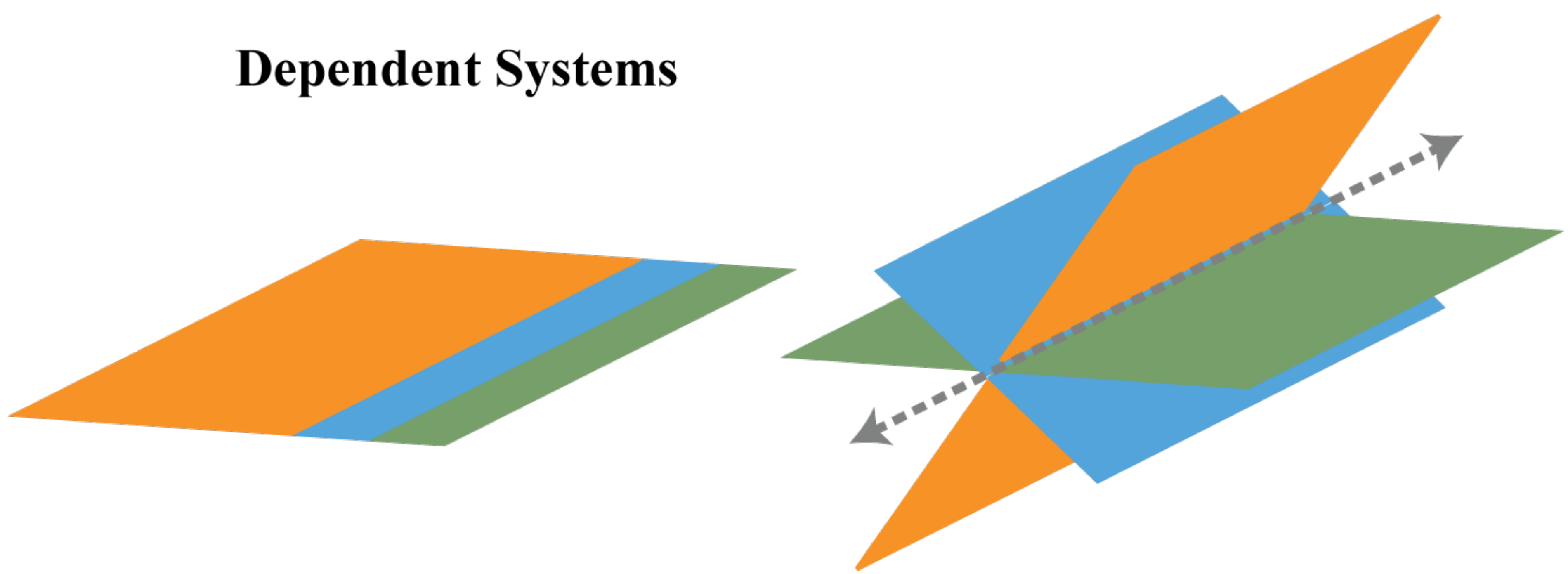
- $\mathbf{a_1}, \mathbf{a_2}$ and $\mathbf{a_3}$ are linearly independent
- $c_1\mathbf{a_1} + c_2\mathbf{a_2} + c_3\mathbf{a_3} = 0$ if and only if $c_1 = c_2 = c_3 = 0$



# Infinite Solutions

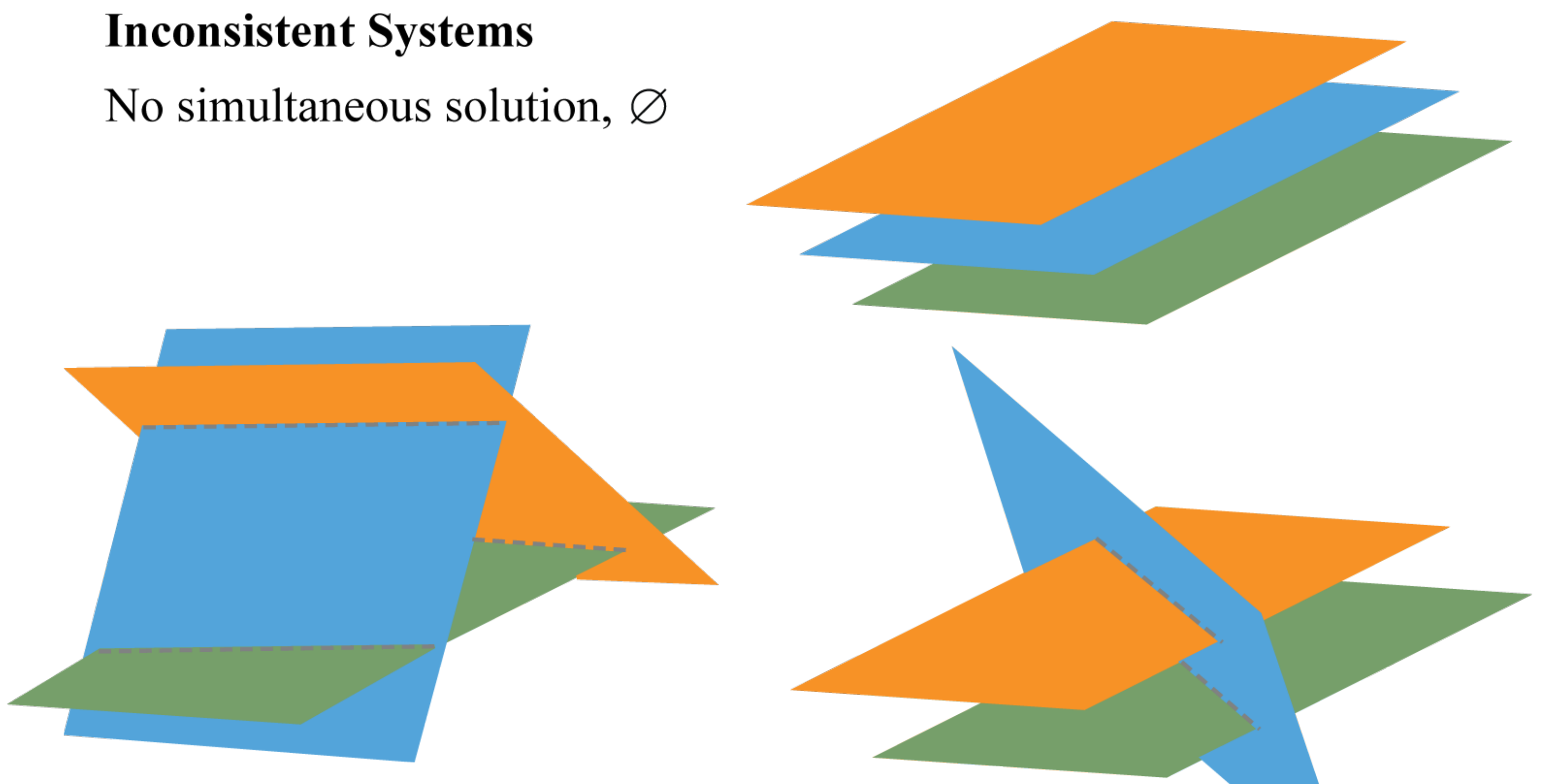
- $\mathbf{a}_1, \mathbf{a}_2$ and $\mathbf{a}_3$ are linearly dependent.
- Thus, there exist scalars $c_1, c_2, c_3$ not all zero such that

$$c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + c_3\mathbf{a}_3 = 0$$



# Inconsistent System

- Case 1: $\mathbf{a}_2 = k_1\mathbf{a}_1$ and $\mathbf{a}_3 = k_2\mathbf{a}_1$ and $b_1 \neq b_2 \neq b_3$
- Case 2: $\mathbf{a}_2 = k_1\mathbf{a}_1$ and $b_1 \neq b_2$. $\mathbf{a}_3$ and $\mathbf{a}_1$ are linearly independent.
- ..
- ..

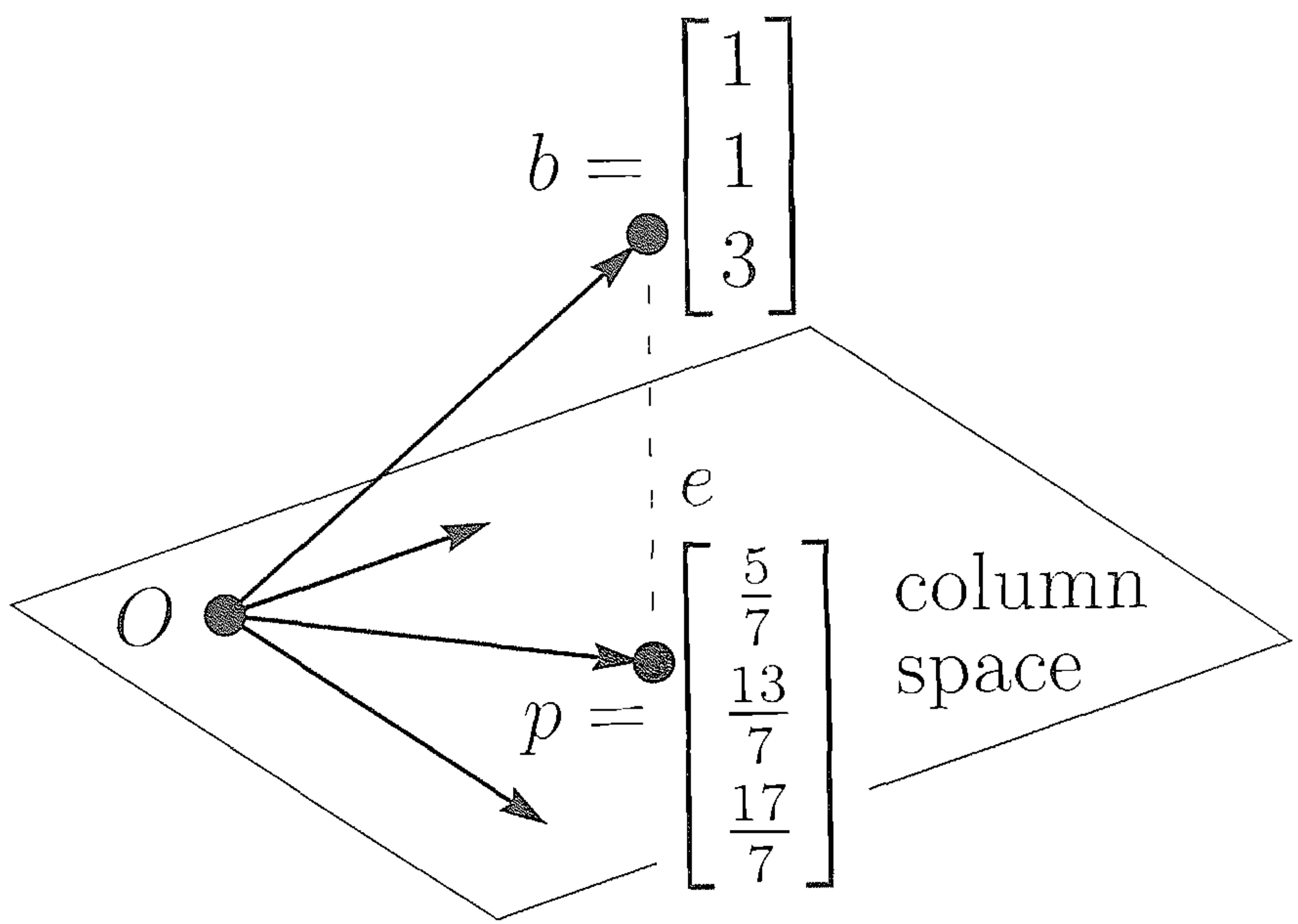


# A general condition for Inconsistency

$$Rank(A) \neq Rank([A \;\; \mathbf{b}])$$

$$\begin{aligned} Rank(A) &= \text{Maximum number of linearly independent column vectors in } A \\ &= \text{Maximum number of linearly independent row vectors in } A \end{aligned}$$

# Column Space View of System of Linear Equations



# Vector Spaces and Subspaces

# Vector Space

Definition : A vector space is a set of vectors together with rules for addition and for multiplication by real numbers.

- Associativity of addition : $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$
- Commutativity of addition : $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$
- Identity element of addition : $\exists \mathbf{0} \in \mathcal{V}$, called the zero vector, such that $\mathbf{v} + \mathbf{0} = \mathbf{v}$, $\forall \mathbf{v} \in \mathcal{V}$
- Inverse element of addition: $\forall \mathbf{v} \in \mathcal{V}$, $\exists -\mathbf{v} \in \mathcal{V}$, called the additive inverse of $\mathbf{v}$, such that $\mathbf{v} + (-\mathbf{v}) = 0$
- Identity element of scalar multiplication: $1\mathbf{v} = \mathbf{v}$, where 1 denotes the multiplicative identity.
- Distributivity of scalar multiplication with respect to vector addition: $a(\mathbf{u} + \mathbf{v}) = a\mathbf{u} + a\mathbf{v}$
- Distributivity of scalar multiplication with respect to field addition: $(a + b)\mathbf{v} = a\mathbf{v} + b\mathbf{v}$

# Subspace

Definition : A subspace of a vector space is a set of vectors (including origin) that satisfies two requirements:

If $\mathbf{v}$ and $\mathbf{w}$ are vectors in the subspace and $c \in \Re$, then

- $\mathbf{v} + \mathbf{w}$ is in the subspace
- $c\mathbf{v}$ is in the subspace

# Subspaces

Consider the vector space of $\Re^d$, then

- 0-dimensional space is the smallest subspace which contains the vector $\mathbf{0}$
- empty set is not allowed to be a subspace
- The whole space $\Re^d$ is the largest subspace of its own

# Column Space

Let $A$ be a $m \times n$ matrix.

**Column space of $A$ :** It consists of all combinations of the columns of matrix $A$. It is denoted by $C(A)$. It forms a subspace of $\Re^m$.

- Suppose $\mathbf{b}$ and $\mathbf{b}'$ lie in the column space, so that $A\mathbf{x} = \mathbf{b}$ for some $\mathbf{x}$ and $A\mathbf{x}' = \mathbf{b}'$ for some $\mathbf{x}'$. Then $A(\mathbf{x} + \mathbf{x}') = \mathbf{b} + \mathbf{b}'$, so that $\mathbf{b} + \mathbf{b}'$ is also a combination of the columns.

- If $\mathbf{b}$ is in the column space $C(A)$, so is any multiple $c\mathbf{b}$. If some combination of columns produces $\mathbf{b}$ (say $A\mathbf{x} = \mathbf{b}$), then multiplying that combination by $c$ will produce $c\mathbf{b}$. In other words, $A(c\mathbf{x}) = c\mathbf{b}$.

# Null Space

**Nullspace :** The nullspace of a matrix $A \in \Re^{m \times n}$ consists of all vectors $\mathbf{x} \in \Re^n$ such that $A\mathbf{x} = \mathbf{0}$. It is denoted by $N(A)$. It is a **subspace** of $\Re^n$, just as the column space was a subspace of $\Re^m$.

- If $A\mathbf{x} = \mathbf{0}$ and $A\mathbf{x}' = \mathbf{0}$, then $A(\mathbf{x} + \mathbf{x}') = \mathbf{0}$.
- If $A\mathbf{x} = \mathbf{0}$ then $A(c\mathbf{x}) = \mathbf{0}$.

**Examples**

- Solution for $\begin{pmatrix} 1 & 0 \\ 5 & 4 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ is $x = 0, y = 0$. Thus the nullspace has only $\mathbf{0}$. Important thing to note here is that the columns of $A$ are independent.
- Let $B = \begin{pmatrix} 1 & 0 & 1 \\ 5 & 4 & 9 \\ 2 & 4 & 6 \end{pmatrix}$. Third column of $B$ is linear combination of the first two columns. The nullspace of $B$ contains the vectors of the form $[c \;\; c \;\; -c]$, $c \in \Re$. The nullspace of $B$ is a line.

# Linear Independence

**Definition :** Suppose $c_1\mathbf{v}_1 + \ldots + c_k\mathbf{v}_k = \mathbf{0}$ only happens when $c_1 = \ldots = c_k = 0$. Then the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ are linearly independent. If any $c$'s are nonzero, then $\mathbf{v}$'s are linearly dependent. One vector is a combination of the others.

**Implications :** If the columns of $A$ are linearly independent, then the nullspace of $A$ contains only zero vector. $A\mathbf{x} = \mathbf{0}$ only when $\mathbf{x} = \mathbf{0}$.

**Example :**

- Columns of $A = \begin{pmatrix} 3 & 4 & 2 \\ 0 & 1 & 5 \\ 0 & 0 & 2 \end{pmatrix}$ are linearly independent.
- Columns of identity matrix are linearly independent.
- A set of $n$ vectors in $\Re^m$ must be linearly dependent.

# Spanning Set

**Definition :** If a vector space $\mathcal{V}$ consists of all linear combinations of vectors $\mathbf{w}_1, \ldots, \mathbf{w}_l$, then these vectors span the space. Thus, $\forall \mathbf{v} \in \mathcal{V}$, $\exists c_1, \ldots, c_l \in \Re$ such that $\mathbf{v} = c_1\mathbf{w}_1 + \ldots + c_l\mathbf{w}_l$.

- It is permitted that different combination of $\mathbf{w}$'s produce the same vector $\mathbf{v}$. Because the spanning set might be excessively large.
- The vectors $\mathbf{w}_1 = [1, 0, 0]$, $\mathbf{w}_2 = [0, 1, 0]$ and $\mathbf{w}_3 = [2, 0, 0]$ span a plane (the x-y plane) in $\Re^3$. The first two vectors also span this plane, whereas $\mathbf{w}_1$ and $\mathbf{w}_3$ span only a line.
- Column vectors $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ of $n \times n$ identity matrix span $\Re^n$. Any vector $\mathbf{b} = [b_1\ b_2\ \ldots\ b_n] \in \Re^n$ can be expressed as

$$\mathbf{b} = b_1\mathbf{e}_1 + b_2\mathbf{e}_2 + \ldots + b_n\mathbf{e}_n$$

# Basis

**Definition :** A basis for vector space $\mathcal{V}$ is a sequence of vectors having two properties at once:

- The vectors are linearly independent (not too many vectors).
- They span the space $\mathcal{V}$ (not too few vectors).

**Implications :**

- Every vector in the space is a linear combination of the basis vectors, because they span.
- It also means that the combination is unique (why?). There is one and only one way to express a vector as a linear combination of the basis vectors.

**Examples :**

- Columns $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ of $n \times n$ identity matrix forms a basis for $\Re^n$.
- There can be infinitely many basis for $\Re^n$.
- Whenever an $n \times n$ matrix $A$ is invertible, its columns are independent and form a basis for $\Re^n$

# Orthonormal Basis

The vectors $q_1, \dots, q_n$ are ***orthonormal*** if

$$q_i^{\mathrm{T}} q_j = \begin{cases} 0 & \text{whenever } i \neq j, \quad \text{giving the orthogonality;} \\ 1 & \text{whenever } i = j, \quad \text{giving the normalization.} \end{cases}$$

**Standard basis**

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \cdots, \quad e_n = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}.$$

There can be infinitely many orthonormal basis !

# Invertible Matrices

**Definition :** Let $A$ be an $n \times n$ matrix. An $n \times n$ matrix $B$ such that $BA = I$ is called a left inverse of $A$; an $n \times n$ matrix $B$ such that $AB = I$ is called a right inverse of $A$. If $AB = BA = I$, then $B$ is called a two-sided inverse of $A$ and $A$ is said to be invertible.

**Properties :**

- If $A$ has left inverse $B$ and right inverse $C$, then $B = C$.
- If $A$ and $B$ are invertible, then $AB$ is invertible and $(AB)^{-1} = B^{-1}A^{-1}$.
- If $A$ is invertible, then the homogeneous system $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.
- If $A$ is invertible, then the system of linear equations $A\mathbf{x} = \mathbf{y}$ has a unique solution for every $\mathbf{y}$.

# Eigenvalues and Eigenvectors

- Given a matrix $A$, a vector $\mathbf{x}$ is called its eigenvector if it does not change its direction after multiplying with $A$.
- $A\mathbf{x} = \lambda\mathbf{x}$
- where $\lambda$ is the eigenvalue corresponding to eigenvector $\mathbf{x}$. $\lambda$ tells whether $\mathbf{x}$ has shrunk or scaled after transformation $A$

# Properties

If $A\mathbf{x} = \lambda\mathbf{x}$, then

- $A^2\mathbf{x} = \lambda^2\mathbf{x}$
- $A^{-1}\mathbf{x} = \lambda^{-1}\mathbf{x}$
- $(A + cI)\mathbf{x} = (\lambda + c)\mathbf{x}$
- $(A - \lambda I)\mathbf{x} = 0$
- $A - \lambda I$ is singular
- $det(A - \lambda I) = 0$

# Finding Eigenvalues and Eigenvectors

**1.** ***Compute the determinant of*** $A - \lambda I$. With $\lambda$ subtracted along the diagonal, this determinant starts with $\lambda^n$ or $-\lambda^n$. It is a polynomial in $\lambda$ of degree $n$.

**2.** ***Find the roots of this polynomial***, by solving $\det(A - \lambda I) = 0$. The $n$ roots are the $n$ eigenvalues of $A$. They make $A - \lambda I$ singular.

**3.** For each eigenvalue $\lambda$, ***solve*** $(A - \lambda I)\boldsymbol{x} = \mathbf{0}$ ***to find an eigenvector*** $\boldsymbol{x}$.

# Example 1

**Example 1** The matrix $A$ has two eigenvalues $\lambda = 1$ and $\lambda = 1/2$. Look at $\det(A-\lambda I)$:

$$A = \begin{bmatrix} .8 & .3 \\ .2 & .7 \end{bmatrix} \quad \det\begin{bmatrix} .8-\lambda & .3 \\ .2 & .7-\lambda \end{bmatrix} = \lambda^2 - \frac{3}{2}\lambda + \frac{1}{2} = (\lambda - 1)\left(\lambda - \frac{1}{2}\right).$$

I factored the quadratic into $\lambda - 1$ times $\lambda - \frac{1}{2}$, to see the two eigenvalues $\boldsymbol{\lambda} = \mathbf{1}$ and $\boldsymbol{\lambda} = \frac{\mathbf{1}}{\mathbf{2}}$. For those numbers, the matrix $A - \lambda I$ becomes *singular* (zero determinant). The eigenvectors $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ are in the nullspaces of $A - I$ and $A - \frac{1}{2}I$.

$(A - I)\boldsymbol{x}_1 = 0$ is $A\boldsymbol{x}_1 = \boldsymbol{x}_1$ and the first eigenvector is $(\mathbf{.6},\ \mathbf{.4})$.

$(A - \frac{1}{2}I)\boldsymbol{x}_2 = 0$ is $A\boldsymbol{x}_2 = \frac{\mathbf{1}}{\mathbf{2}}\boldsymbol{x}_2$ and the second eigenvector is $(\mathbf{1},\ \mathbf{-1})$:

$$\boldsymbol{x}_1 = \begin{bmatrix} .6 \\ .4 \end{bmatrix} \quad \text{and} \quad A\boldsymbol{x}_1 = \begin{bmatrix} .8 & .3 \\ .2 & .7 \end{bmatrix}\begin{bmatrix} .6 \\ .4 \end{bmatrix} = \boldsymbol{x}_1 \quad (A\boldsymbol{x} = \boldsymbol{x} \text{ means that } \lambda_1 = 1)$$

$$\boldsymbol{x}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad \text{and} \quad A\boldsymbol{x}_2 = \begin{bmatrix} .8 & .3 \\ .2 & .7 \end{bmatrix}\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} .5 \\ -.5 \end{bmatrix} \quad (\text{this is } \tfrac{1}{2}\,\boldsymbol{x}_2 \text{ so } \lambda_2 = \tfrac{1}{2}).$$

# Example 1 :Continued

$\lambda = 1$

$Ax_1 = x_1 = \begin{bmatrix} .6 \\ .4 \end{bmatrix}$

$\lambda = .5$

$Ax_2 = \lambda_2 x_2 = \begin{bmatrix} .5 \\ -.5 \end{bmatrix}$

$x_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

$\lambda^2 = 1$

$A^2 x_1 = (1)^2 x_1$

$\lambda^2 = .25$

$A^2 x_2 = (.5)^2 x_2 = \begin{bmatrix} .25 \\ -.25 \end{bmatrix}$

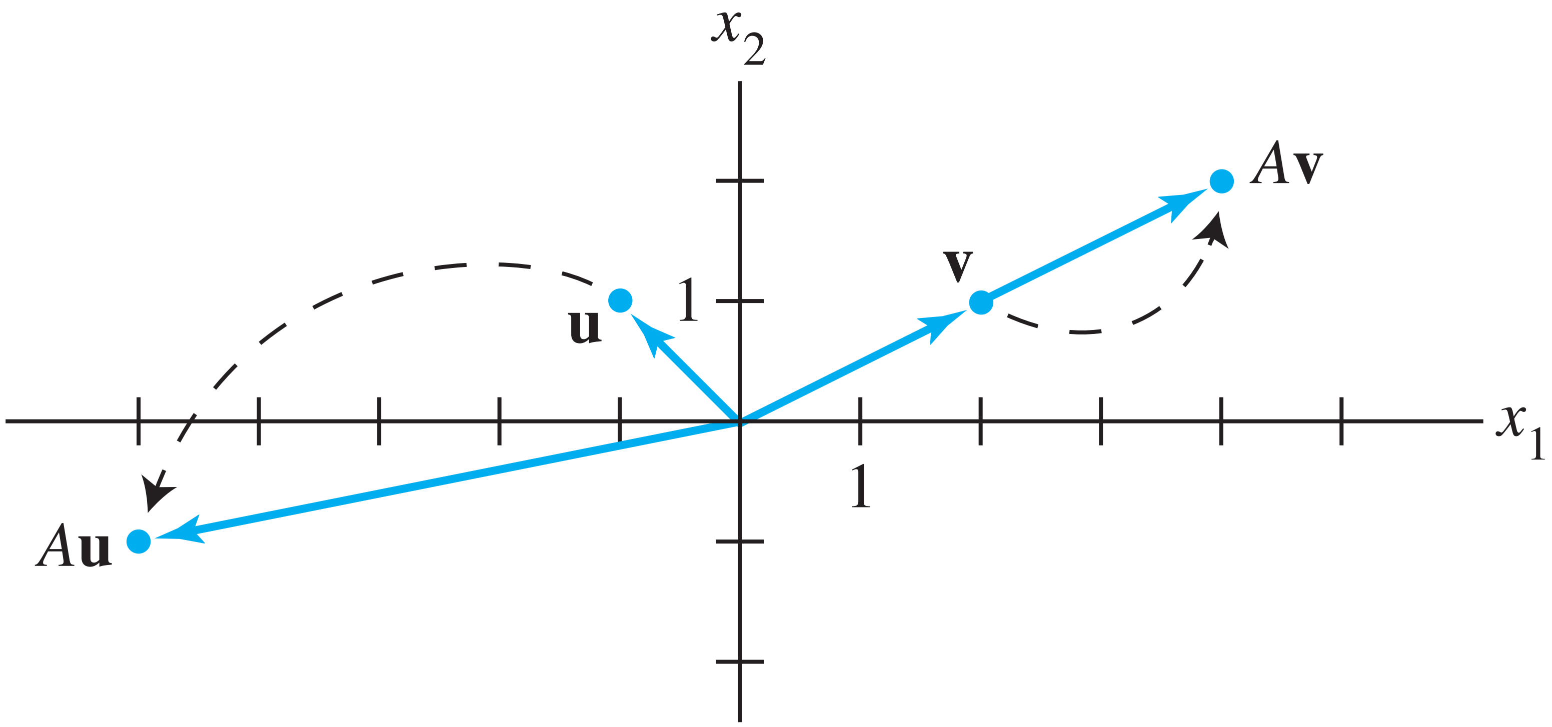
$$A\boldsymbol{x} = \lambda \boldsymbol{x}$$

$$A^2\boldsymbol{x} = \lambda^2 \boldsymbol{x}$$

# Example

Let $A = \begin{pmatrix} 3 & -2 \\ 1 & 0 \end{pmatrix}$, $\mathbf{u} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$, $\mathbf{v} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$

We see that, $A\mathbf{v} = 2\mathbf{v}$



# Diagonalizing a Matrix

**Diagonalization :** Suppose the $n \times n$ matrix $A$ has $n$ linearly independent eigenvectors $\mathbf{x}_1, \ldots, \mathbf{x}_n$. Construct an eigenvector matrix $S$ by putting them as columns of $S$. Then $S^{-1}AS$ is the eigenvalue matrix $\Lambda$.

$$\begin{aligned} AS &= A[\mathbf{x}_1 \ \ \mathbf{x}_2 \ \ \ldots \ \ \mathbf{x}_n] \\ &= [\lambda_1\mathbf{x}_1 \ \ \lambda_2\mathbf{x}_2 \ \ \ldots \ \ \lambda_n\mathbf{x}_n] \\ &= [\mathbf{x}_1 \ \ \mathbf{x}_2 \ \ \ldots \ \ \mathbf{x}_n] \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ldots & 0 \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix} \\ &= S\Lambda \end{aligned}$$

Thus, $S^{-1}AS = \Lambda$ or $A = S\Lambda S^{-1}$

- Matrix $S$ has an inverse, because its columns were assumed to be linearly independent.
- Let $A = \begin{pmatrix} 1 & 5 \\ 0 & 6 \end{pmatrix}$, then $\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 5 \\ 0 & 6 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 6 \end{pmatrix}$
- $A^2 = S\Lambda S^{-1} S\Lambda S^{-1} = S\Lambda^2 S^{-1}$.
- $A^k = S\Lambda^k S^{-1}$
- $\begin{pmatrix} 1 & 5 \\ 0 & 6 \end{pmatrix}^k = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 6^k \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 6^{k-1} \\ 0 & 6^k \end{pmatrix}$

# Diagonalizing a Matrix

- Any matrix that has no repeated eigenvalues, can be diagnonalized
- The eigenvectors is $S$ come in the same order as the eigenvalues in $\Lambda$.
- Some matrices have too few eigenvectors due to repeated eigenvalues. Those matrices can not be diagonalised.
- Examples : $A = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$
- Eigenvectors for $n$ different $\lambda$'s are independent. Then we can diagonalise $A$.

# Symmetric Matrices

- A symmetric matrix has only real eigenvalues.
- The eigenvectors can be chosen orthonormal.
- **Spectral Decomposition :** Every symmetric matrix has the factorization $A = Q\Lambda Q^T$ with real eigenvalues in $\Lambda$ and orthonormal eigenvectors in $S = Q$
- Symmetric Diagonalization : $A = Q\Lambda Q^{-1} = Q\Lambda Q^T$

# Orthogonal Matrices

- A matrix **Q** with orthonormal columns is called orthogonal matrix

**Orthonormal columns**

$$\begin{bmatrix} — & q_1^{\mathrm{T}} & — \\ — & q_2^{\mathrm{T}} & — \\ & \vdots & \\ — & q_n^{\mathrm{T}} & — \end{bmatrix} \begin{bmatrix} | & | & & | \\ q_1 & q_2 & \cdots & q_n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdot & 0 \\ 0 & 1 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & 1 \end{bmatrix} = I.$$

$$Q^{\mathrm{T}}Q = I$$

# Example 1 : Rotation Matrices

$$Q = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}, \qquad Q^{\mathrm{T}} = Q^{-1} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$

# Example 2: Permutation Matrices

$$\text{If} \quad P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \quad \text{then} \quad P^{-1} = P^{\mathrm{T}} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$