COMP 7180
Quantitative Methods for Data Analytics and Artificial Intelligence

Lecture 1 – Introduction to Linear Algebra: Vectors and Matrices

What is Linear Algebra

- Linear
  - Having to do with lines/planes/etc.
  - For example, x-y = 1, x+y+3z = 7, not *sin*; *log*; x^2, etc.
- Algebra
  - Solving equations involving numbers and symbols
  - From al-jebr (Arabic), meaning reunion of broken parts
  - Abu Ja'far Muhammad ibn Muso al-Khwarizmi, 9th century

What is Linear Algebra

- Linear algebra is the study of vectors, matrices, and linear functions/equations.
- Vectors (columns of numbers) and matrices (2D arrays of numbers) are the language of data.
- **Linear algebra is a branch of mathematics, but the truth of it is that linear algebra is the mathematics of data.**

Vectors in two dimensions

- A two-dimensional vector **v**

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

The first component of **v**

The second component of **v**

- We write **v** as a column. A single letter **v** (in boldface) to denote a vector.

Two basic operations in Linear Algebra

- Vector addition: add vectors

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \qquad \mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \quad \Longrightarrow \quad \mathbf{v} + \mathbf{w} = \begin{bmatrix} v_1 + w_1 \\ v_2 + w_2 \end{bmatrix}$$

  – Subtraction follows the same ideas

$$\mathbf{v} - \mathbf{w} = \begin{bmatrix} v_1 - w_1 \\ v_2 - w_2 \end{bmatrix}$$

- Scalar multiplication: vector can be multiplied by any number (scalar) *c*

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \quad \Longrightarrow \quad c\mathbf{v} = \begin{bmatrix} cv_1 \\ cv_2 \end{bmatrix}$$

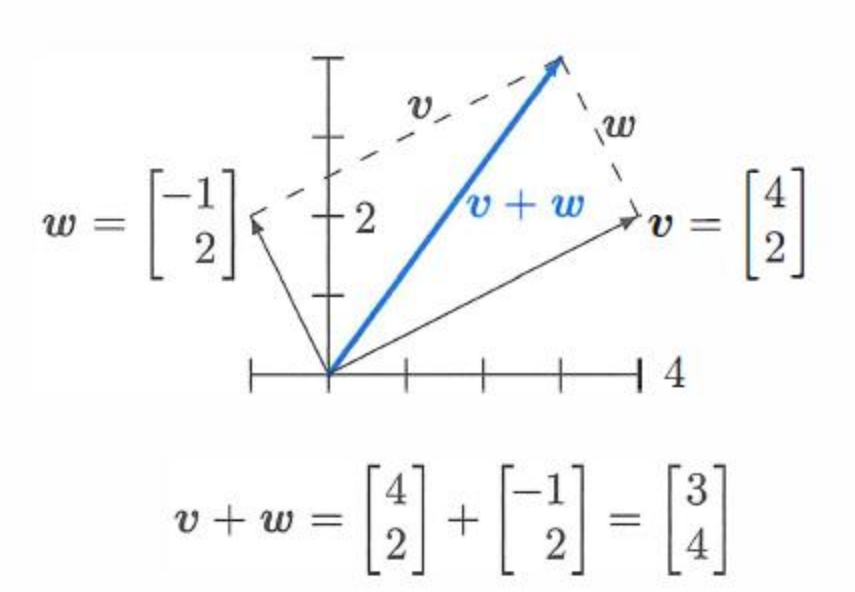
Two basic operations in Linear Algebra

- Linear algebra is built on these two operations
  - Adding vectors: $\mathbf{v} + \mathbf{w}$
  - Multiplying by scalars: $c\mathbf{v}$

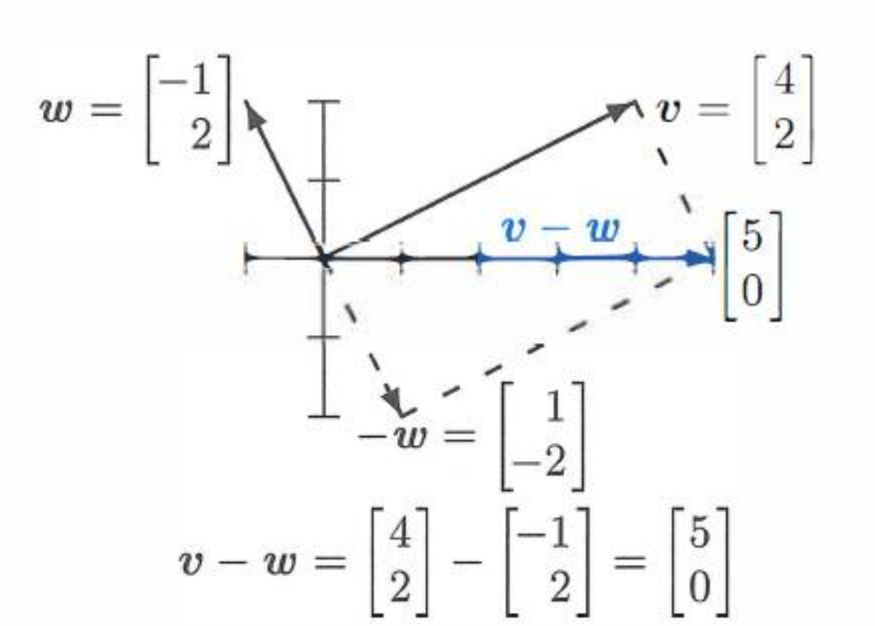
- Linear combination: combine addition with scalar multiplication
  - $\mathrm{c}\mathbf{v} + \mathrm{d}\mathbf{w}$: multiply **v** by *c* and multiply **w** by *d*, then add together
  - $\mathrm{c}\mathbf{v} + \mathrm{d}\mathbf{w}$ is a **"linear combination"** of vector **v** and **w**

Visualize vector addition/subtraction

Vector Addition



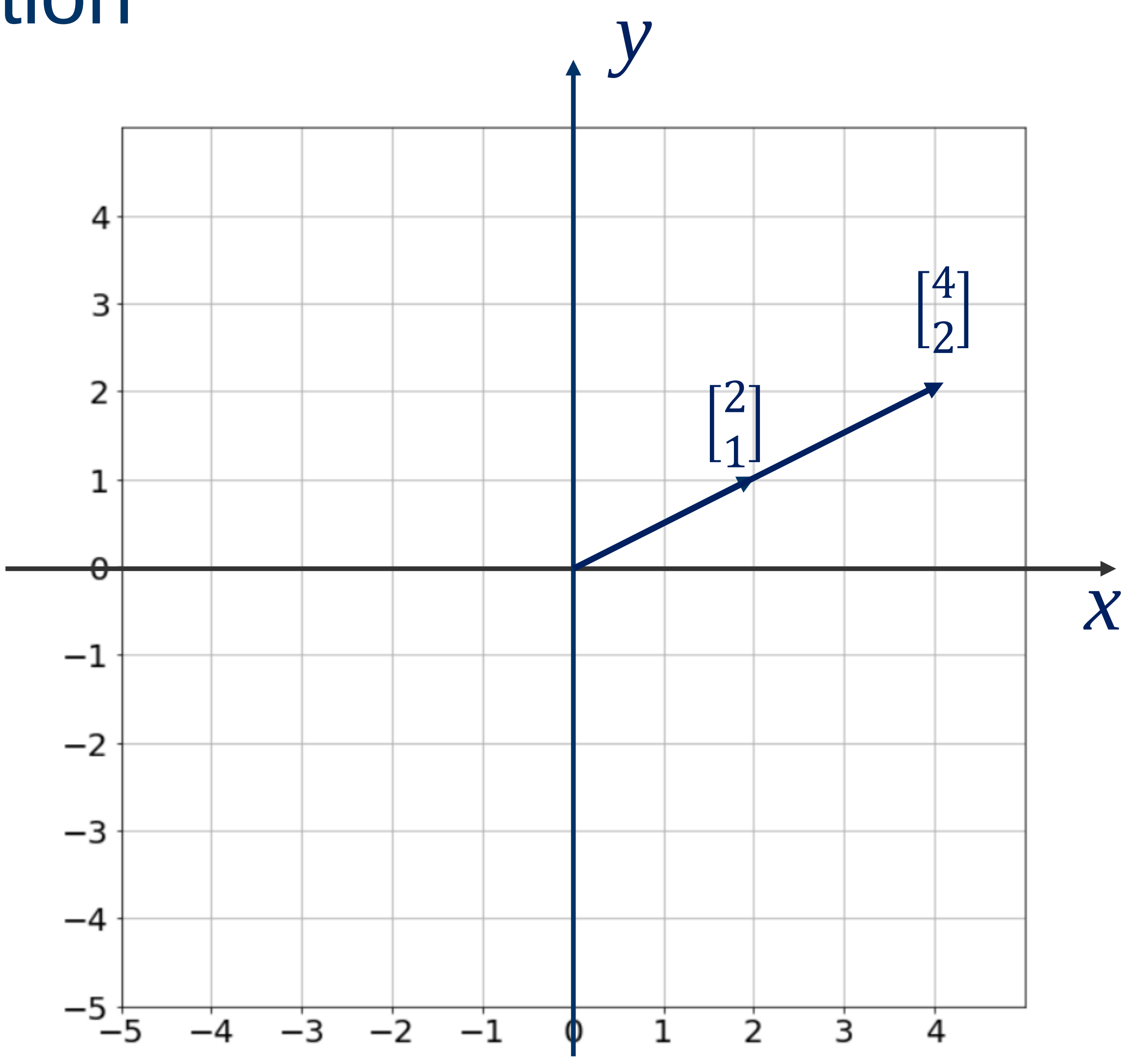
Vector Subtraction



- A vector can be represented by an arrow from the "**origin**" (0, 0)
- $\mathbf{v} + \mathbf{w}$: at the end of **v**, place the start of **w**. Then the third side is $\mathbf{v} + \mathbf{w}$

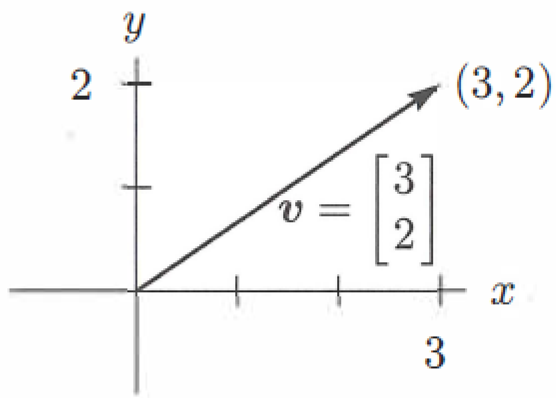
Visualize vector-scalar multiplication

- $\mathbf{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$
- $2\mathbf{v} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$
- If the scalar is positive, vector-scalar multiplication only changes the length (magnitude) of the vector, it does not change the direction of the vector.
- If the scalar is negative, it both changes the length and reverses the direction of the vector.

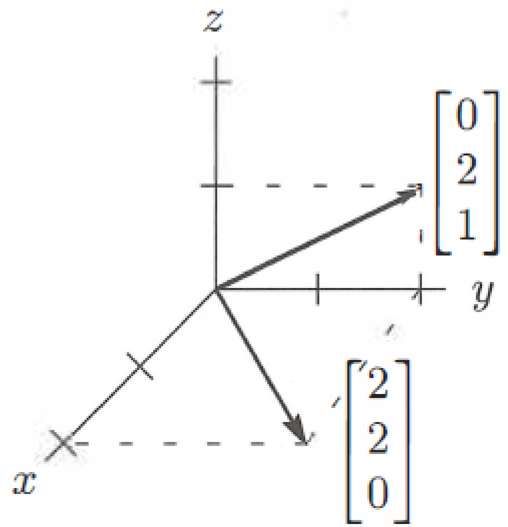


Vectors in Three Dimensions

- A vector with two components corresponds to a point in the *xy* plane.
  - The components of **v** are the coordinates of the point: $x = v_1$, $y = v_2$.



- Now we allow vectors to have three components, then the *xy* plane is replaced by three-dimensional *xyz* space.



Vectors in Three Dimensions

- In three dimensions, **v** + **w** is still found by element-wise addition (the same as in two dimensions).
  - E.g.,

$$\mathbf{v} = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \qquad \mathbf{w} = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} \qquad \Rightarrow \qquad \mathbf{v} + \mathbf{w} = \begin{bmatrix} 3 \\ 4 \\ 3 \end{bmatrix}$$

- Then we can see how to add vectors in 4 or 5 or *n* dimensions. When **w** starts at the end of **v**, the third side is **v** + **w**.

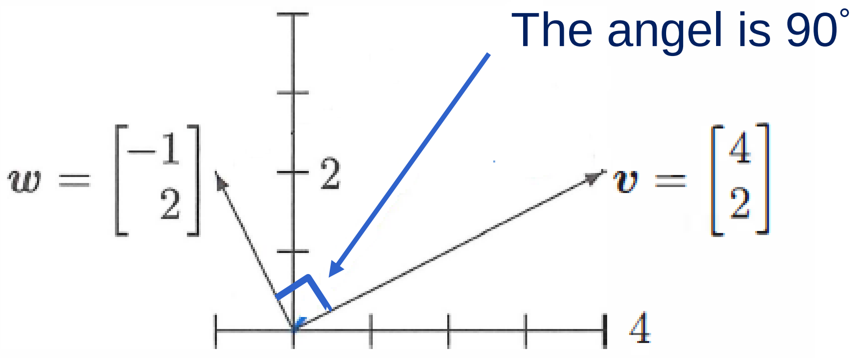
Dot product

- The dot product (or inner product) of $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ and $\mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$ is a scalar

$$\mathbf{v} \cdot \mathbf{w} = v_1 w_1 + v_2 w_2$$

- If the dot product of two vectors is 0, it means that ***these two vectors are perpendicular (orthogonal)***.

$$\begin{bmatrix} 4 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 2 \end{bmatrix} = 4 * (-1) + 2 * 2 = 0$$



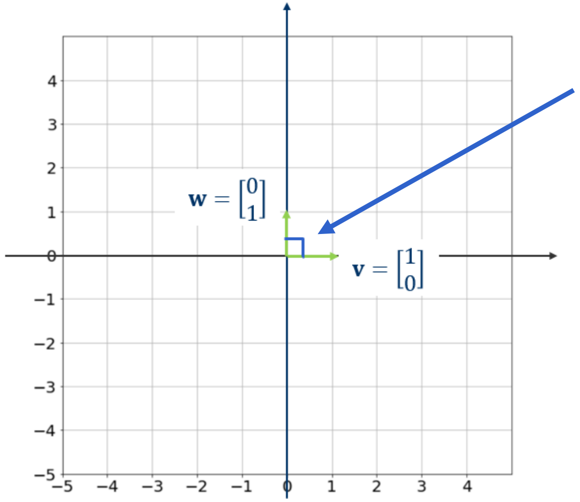
Dot product

- The dot product (or inner product) of $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ and $\mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$ is a scalar

$$\mathbf{v} \cdot \mathbf{w} = v_1 w_1 + v_2 w_2$$

- If the dot product of two vectors is 0, it means that ***these two vectors are perpendicular (orthogonal)***.

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 1 * 0 + 0 * 1 = 0$$



The angel is 90˚

Dot product example

- We have three goods to buy or sell. Their prices are (p_1, p_2, p_3) for each unit – this is the "price vector" **p**. The quantities that we buy or sell are (q_1, q_2, q_3) – positive when we sell, negative when we buy. Selling q_1 units at the price p_1 brings in q_1p_1. The total income (quantities **q** times prices **p**) is the dot product $\mathbf{q} \cdot \mathbf{p}$ in three dimensions:

$$\textbf{Income} = \mathbf{q} \cdot \mathbf{p} = \begin{bmatrix} q_1 \\ q_2 \\ q_3 \end{bmatrix} \cdot \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix} = q_1p_1 + q_2p_2 + q_3p_3$$

  - Total sales equal to total purchases if $\mathbf{q} \cdot \mathbf{p} = 0$, which means that **p** is perpendicular to **q** in three dimensional space.
  - A supermarket with thousands of goods goes quickly into high dimensions.

Length of a vector

- An important case of dot product is the dot product with itself.
  - E.g., **v** = (1, 2, 3)

$$\mathbf{v} \cdot \mathbf{v} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = 1 + 4 + 9 = 14$$

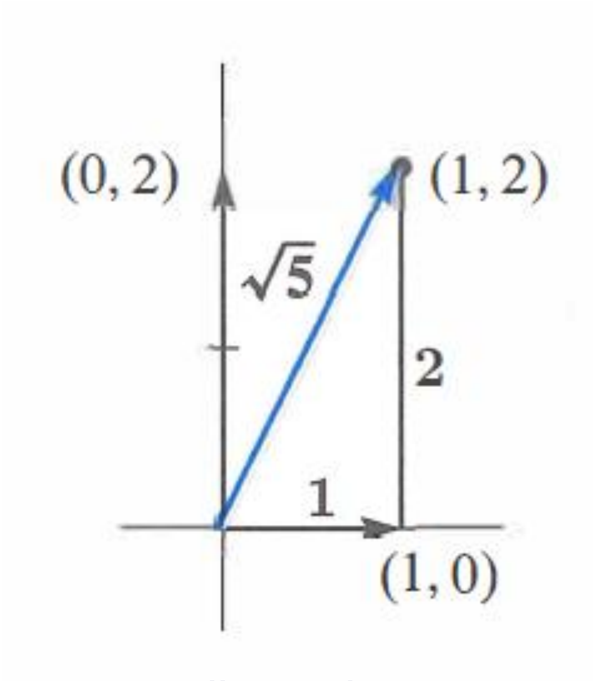
- The dot product $\mathbf{v} \cdot \mathbf{v}$ gives the length of **v** squared: $\|\mathbf{v}\|^2$

- The length $\|\mathbf{v}\|$ of a vector **v** is the square root of $\mathbf{v} \cdot \mathbf{v}$

$$\|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}} = \sqrt{v_1^2 + v_2^2 + \cdots + v_d^2}$$

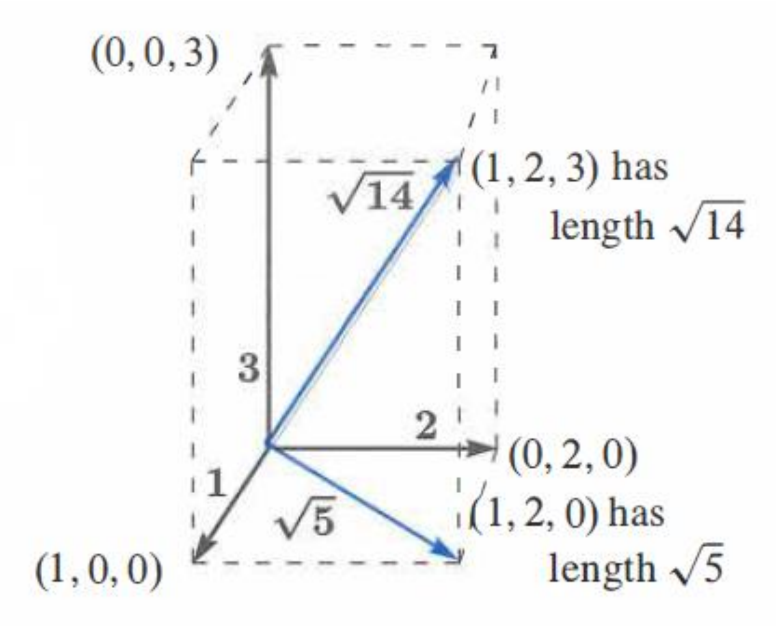
Length of a vector

- $\|\mathbf{v}\|$ is the length of the arrow that represents the vector.



Two-dimensional vector

$$\|\mathbf{v}\| = \sqrt{v_1^2 + v_2^2}$$



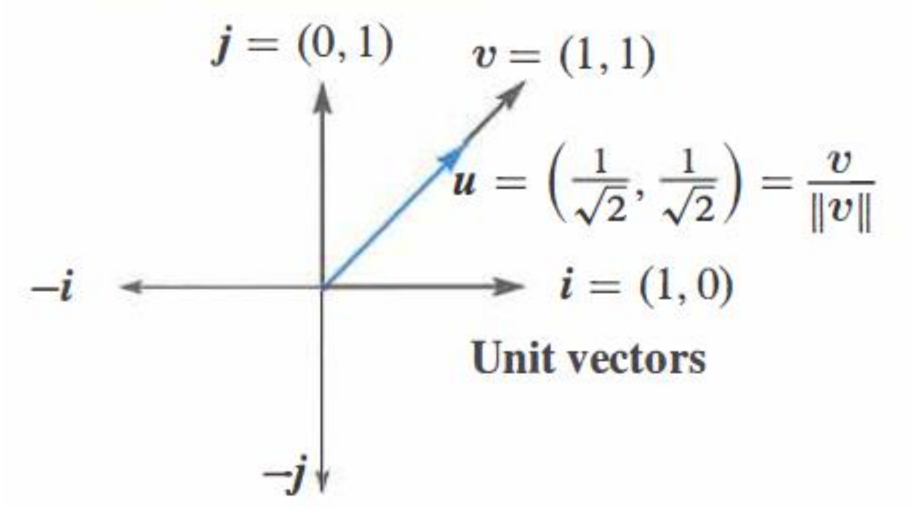
Three-dimensional vector

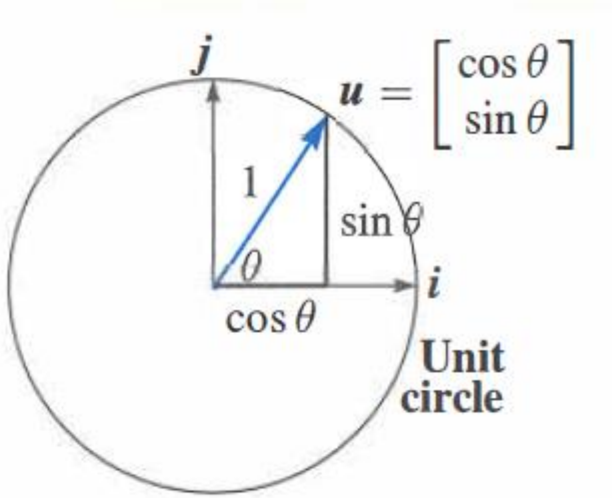
$$\|\mathbf{v}\| = \sqrt{v_1^2 + v_2^2 + v_3^2}$$

Unit Vector

- The word “unit” is always indicating that some measurement equals “one”
- A unit vector $\mathbf{u}$ is a vector whose length equal one: $\mathbf{u} \cdot \mathbf{u} = 1$.
- How to get a unit vector?
  - For any nonzero vector $\mathbf{v}$, we can obtain its unit vector by dividing it by its length $||\mathbf{v}||$
- $\boldsymbol{u} = \frac{\mathbf{v}}{||\mathbf{v}||}$ is a unit vector in the same direction as $\mathbf{v}$.

Unit Vector





The unit vector **u** at angle 45° is obtained by dividing **v** = (1,1) by its length $\|\mathbf{v}\| = \sqrt{2}$.

The **u** = (cosθ, sin θ) is a unit vector at angle θ.

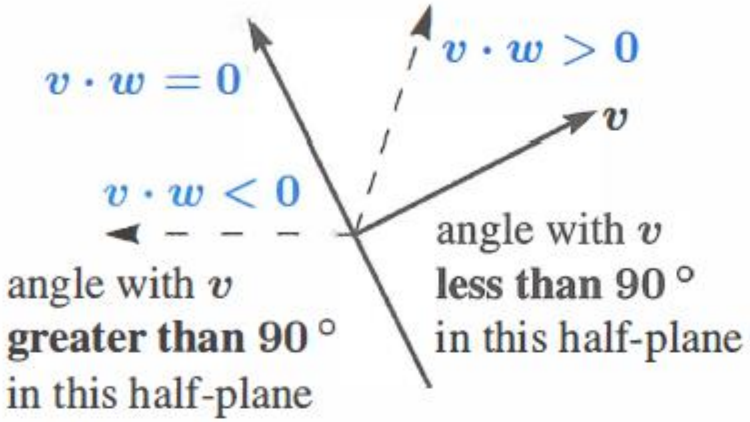
$(\cos\theta)^2 + (\sin\theta)^2 = 1$

The Angle Between Two Vectors

- The dot product $\mathbf{v} \cdot \mathbf{w}$ is zero when the angle between these two vectors **v** and **w** is θ = 90˚. **v** is perpendicular to **w**.

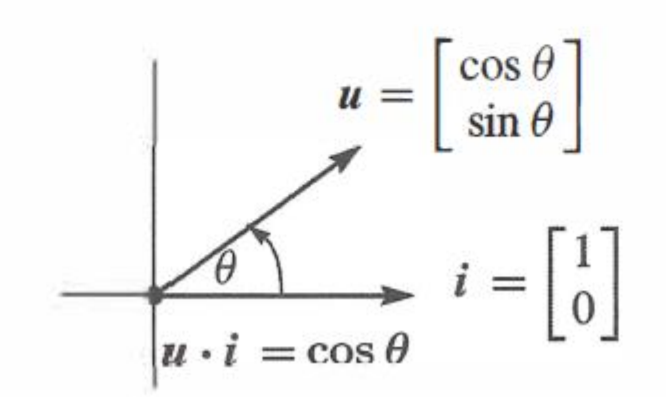
- Zero vector **v** = **0** is perpendicular to every vector **w** because $\mathbf{0} \cdot \mathbf{w}$ is always zero.

- How about $\mathbf{v} \cdot \mathbf{w}$ is not zero?
  - The sign of $\mathbf{v} \cdot \mathbf{w}$ tells whether they are below or above a right angle.
  - The angle is less than 90˚ when $\mathbf{v} \cdot \mathbf{w}$ is positive
  - The angle is above 90˚ when $\mathbf{v} \cdot \mathbf{w}$ is negative
  - The borderline is where vectors are perpendicular to **v**.

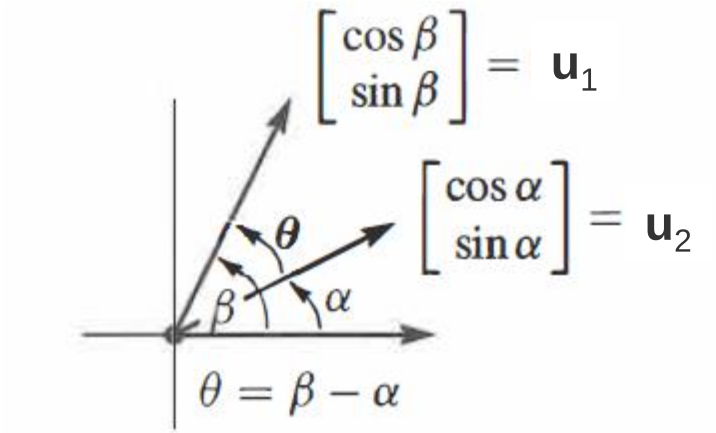


The Angle Between Two Vectors

- The dot product reveals the exact angle θ.
- For two unit vectors $\mathbf{u}_1$, $\mathbf{u}_2$, the dot product $\mathbf{u}_1 \cdot \mathbf{u}_2$ is the cosine of θ. This remains true in *d* dimensions.



$$\mathbf{u} \cdot \boldsymbol{i} = \begin{bmatrix}\cos\theta \\ \sin\theta\end{bmatrix} \cdot \begin{bmatrix}1 \\ 0\end{bmatrix} = \cos\theta$$



$$\mathbf{u}_1 \cdot \mathbf{u}_2 = \begin{bmatrix}\cos\beta \\ \sin\beta\end{bmatrix} \cdot \begin{bmatrix}\cos\alpha \\ \sin\alpha\end{bmatrix} = \cos\beta\cos\alpha + \sin\beta\sin\alpha$$
$$= \cos(\beta - \alpha) = \cos\theta$$

The Angle Between Two Vectors

- How about the exact angle between two non-unit vectors **v** and **w**?
  - Divide the non-unit vectors by their length to get unit vectors $\frac{\mathbf{v}}{\|\mathbf{v}\|}$ and $\frac{\mathbf{w}}{\|\mathbf{w}\|}$
  - Then the dot product of unit vectors $\frac{\mathbf{v}}{\|\mathbf{v}\|}$ and $\frac{\mathbf{w}}{\|\mathbf{w}\|}$ gives $\cos\theta$

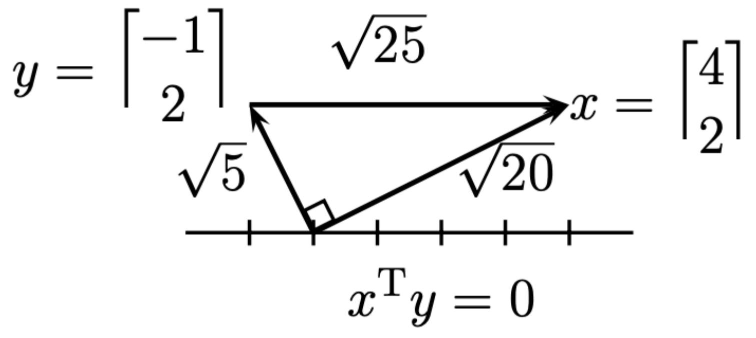
- **Cosine Formula**

  If **v** and **w** are nonzero vectors then $\frac{\mathbf{v}}{\|\mathbf{v}\|} \cdot \frac{\mathbf{w}}{\|\mathbf{w}\|} = \cos\theta$

The Angle Between Two Vectors – Another Point of View

- What is perpendicularity (orthogonality)?
  - Two vectors are perpendicular/orthogonal provided they form a **right triangle**.

Sides of a right triangle $\quad \|x\|^2 + \|y\|^2 = \|x - y\|^2.$



  - Applying the length formula, this test for orthogonality in $\mathbf{R}^n$ becomes:

$$\left(x_1^2 + \cdots + x_n^2\right) + \left(y_1^2 + \cdots + y_n^2\right) = (x_1 - y_1)^2 + \cdots + (x_n - y_n)^2.$$
$$= \left(x_1^2 + \cdots + x_n^2\right) - 2(x_1 y_1 + \cdots + x_n y_n) + \left(y_1^2 + \cdots + y_n^2\right).$$

Orthogonal vectors $\quad x^{\mathrm{T}} y = x_1 y_1 + \cdots + x_n y_n = 0.$

In-Class Exercises

1. What combinations $c\begin{bmatrix}1\\2\end{bmatrix} + d\begin{bmatrix}3\\1\end{bmatrix}$ produce $\begin{bmatrix}14\\8\end{bmatrix}$?

2. Can three vectors **u**, **v**, **w** in the x-y plane have $\mathbf{u} \cdot \mathbf{v} < 0$, $\mathbf{v} \cdot \mathbf{w} < 0$, and $\mathbf{w} \cdot \mathbf{u} < 0$? If Yes, give examples; if no, give the reason.

3. Can more than three vectors in the x-y plane have all negative dot product? Why?

4. Pick any values of x, y, z that satisfy x + y + z = 0. Calculate the angle between **u** = (x, y, z) and **v** = (z, x, y). Try it again with different values of x, y, z that satisfy x + y + z = 0. Any discoveries? And why?

Solution to In-Class Exercises

1. What combinations $c \begin{bmatrix} 1 \\ 2 \end{bmatrix} + d \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ produce $\begin{bmatrix} 14 \\ 8 \end{bmatrix}$?

From $c \begin{bmatrix} 1 \\ 2 \end{bmatrix} + d \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 14 \\ 8 \end{bmatrix}$ we have $\begin{bmatrix} c + 3d = 14 \\ 2c + d = 8 \end{bmatrix}$. Multiplying row 1 by 2, we have $\begin{bmatrix} 2c + 6d = 28 \\ 2c + d = 8 \end{bmatrix}$. Subtracting row 2 from row 1, we have $\begin{bmatrix} 5d = 20 \\ 2c + d = 8 \end{bmatrix}$. So $d = 4$. Then $2c + d = 8$ indicates $2c = 4$. So $c = 2$.

Solution to In-Class Exercises

2. Can three vectors **u**, **v**, **w** in the x-y plane have $\mathbf{v} \cdot \mathbf{w}$ < 0, $\mathbf{v} \cdot \mathbf{w}$ < 0, and $\mathbf{v} \cdot \mathbf{w}$ < 0? If Yes, give examples; if no, give the reason.

Yes. An example is: $\mathrm{u} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \mathrm{v} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}, \mathrm{u} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$.

Solution to In-Class Exercises

3. Can more than three vectors in the x-y plane have all negative dot product? Why?

No. Because negative dot product requires the angle between two vectors to be larger than 90°. However, a circle of the 2-D plane is 360°, which is not possible to be divided into 4 obtuse angles.

Solution to In-Class Exercises

4. Pick any values of x, y, z that satisfy x + y + z = 0. Calculate the angle between **u** = (x, y, z) and **v** = (z, x, y). Try it again with different values of x, y, z that satisfy x + y + z = 0. Any discoveries? And why?

From $x + y + z = 0$, we have $(x + y + z)^2 = 0$. This can be rewritten as:

$(x^2 + y^2 + z^2) + 2(xy + yz + zx) = 0$. This can be further rewritten as:

$-1/2 = (xy + yz + zx) / (x^2 + y^2 + z^2) = (x, y, z) \cdot (z, x, y) / (x^2 + y^2 + z^2) = \frac{\mathbf{u}}{\|\mathbf{u}\|} \cdot \frac{\mathbf{v}}{\|\mathbf{v}\|} = \cos\theta$, where θ is the angle between **u** and **v**. So $\theta = 120°$.

Matrix

Definition: An $m \times n$ matrix, A, is a rectangular array of elements

$$A = \left(a_{ij}\right) = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

m = # of rows

n = # of columns

dimensions = $m \times n$

Matrix Operations

- Matrix addition
  - Matrix can be added if their shapes are the same
  - Matrix addition is like vector addition: element-wise addition

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 2 & 2 \\ 4 & 4 \\ 9 & 9 \end{bmatrix} = \begin{bmatrix} 3 & 4 \\ 7 & 8 \\ 9 & 9 \end{bmatrix}$$

- Matrix multiplied by a scalar *c*
  - Matrix can be multiplied by a scalar
  - Each entry in the matrix will be multiplied by the scalar

$$2\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 6 & 8 \\ 0 & 0 \end{bmatrix}$$

Matrix Operations: Matrix Multiplication

- When can we multiply matrix **A** by matrix **B**?
  - To multiply matrix **A** by matrix **B**, number of columns in **A** must equal to the number of rows in **B**.

$$\mathbf{A}_{m*n}\mathbf{B}_{n*p} = \mathbf{C}_{m*p}$$

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & 4 \\ 1 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & 4 \\ 1 & 1 \\ 1 & 1 \end{bmatrix}$$

What is matrix product **AB**?

- The first way: the dot product way (the usual way to multiply matrices by hand) .
- The product **AB** is filled with dot products: take the dot product of each row of **A** with each column of **B**
- The *ij*-th (i.e., the *i*-th row and the *j*-th column) entry in matrix product **AB** is the dot product between *i*-th row of **A** and *j*-th column of **B**.

$$\begin{bmatrix} * & & & \\ a_{i1} & a_{i2} & \cdots & a_{i5} \\ * & & & \\ * & & & \end{bmatrix}
\begin{bmatrix} * & * & b_{1j} & * & * & * \\ & & b_{2j} & & & \\ & & \vdots & & & \\ & & b_{5j} & & & \end{bmatrix} = \begin{bmatrix} & & * & & & \\ * & * & (AB)_{ij} & * & * & * \\ & & * & & & \\ & & * & & & \end{bmatrix}$$

A is 4 by 5 B is 5 by 6 AB is $(4 \times 5)(5 \times 6) = 4$ by 6

What is matrix product **AB**?

- E.g.1,

$$\begin{bmatrix}1 & 1 & 0\\2 & -1 & 0\\0 & 0 & 1\end{bmatrix}\begin{bmatrix}2 & 2 & 0\\3 & 4 & 1\\0 & 0 & 0\end{bmatrix}=\begin{bmatrix}1*2+1*3+0*0 & 1*2+1*4+0*0 & 1*0+1*1+0*0\\2*2+(-1)*3+0*0 & 2*2+(-1)*4+0*0 & 2*0+(-1)*1+0*0\\0*2+0*3+1*0 & 0*2+0*4+1*0 & 0*0+0*1+1*0\end{bmatrix}=\begin{bmatrix}5 & 6 & 1\\1 & 0 & -1\\0 & 0 & 0\end{bmatrix}$$

- E.g. 2, a column times row $\mathbf{A}_{m*1}\mathbf{B}_{1*p}=\mathbf{C}_{m*p}$

$$\begin{bmatrix}0\\1\\2\end{bmatrix}\begin{bmatrix}1 & 2 & 3\end{bmatrix}=\begin{bmatrix}0*1 & 0*2 & 0*3\\1*1 & 1*2 & 1*3\\2*1 & 2*2 & 2*3\end{bmatrix}=\begin{bmatrix}0 & 0 & 0\\1 & 2 & 3\\2 & 4 & 6\end{bmatrix}$$

  - A column times a row is an “outer” product. The result is a matrix.
  - A row times a column is an “inner” product, that is another name for dot product. The result is a scalar.

What is matrix product **AB**?

- The second way (column picture): Each column of **AB** is a linear combination of the columns of **A**.
  - Matrix **A** times every column of **B**.

$$\mathbf{A}_{m*n}\mathbf{B}_{n*p} = \mathbf{A}[\mathbf{b}_1 \quad \mathbf{b}_2 \quad \mathbf{b}_3 \quad \cdots \quad \mathbf{b}_p] = [\mathbf{A}\mathbf{b}_1 \quad \mathbf{A}\mathbf{b}_2 \quad \mathbf{A}\mathbf{b}_3 \quad \cdots \quad \mathbf{A}\mathbf{b}_p]$$

  - E.g.,

$$\begin{bmatrix} 1 & 1 & 0 \\ 2 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 2 & 2 & 0 \\ 3 & 4 & 1 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 5 & 6 & 1 \\ 1 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$0\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} + 1\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + 0\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$$

$$2\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} + 3\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + 0\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 1 \\ 0 \end{bmatrix}$$

$$2\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} + 4\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + 0\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix}$$

What is matrix product **AB**?

- The third way (row picture): each row of **AB** is a linear combination of rows of **B**.
  - Every row of **A** times matrix **B**

$$\begin{bmatrix} 1 & 1 & 0 \\ 2 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 2 & 2 & 0 \\ 3 & 4 & 1 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} \boxed{5 \quad 6 \quad 1} \\ \boxed{1 \quad 0 \quad -1} \\ \boxed{0 \quad 0 \quad 0} \end{bmatrix}$$

$$\boxed{\begin{aligned} &1[2 \quad 2 \quad 0] + \\ &1[3 \quad 4 \quad 1] + \\ &0[0 \quad 0 \quad 0] = [5 \quad 6 \quad 1] \end{aligned}}$$

$$\boxed{\begin{aligned} &2[2 \quad 2 \quad 0] + \\ &-1[3 \quad 4 \quad 1] + \\ &0[0 \quad 0 \quad 0] = [1 \quad 0 \quad -1] \end{aligned}}$$

$$\boxed{\begin{aligned} &0[2 \quad 2 \quad 0] + \\ &0[3 \quad 4 \quad 1] + \\ &1[0 \quad 0 \quad 0] = [0 \quad 0 \quad 0] \end{aligned}}$$

What is matrix product **AB**?

- The fourth way (columns multiply rows): multiply columns 1 to *n* of **A** by rows 1 to *n* of **B**, then add those matrices together.

$$\begin{bmatrix} | & & | \\ a_1 & \cdots & a_n \\ | & & | \end{bmatrix} \begin{bmatrix} - & b_1 & - \\ & \vdots & \\ - & b_n & - \end{bmatrix} = \begin{bmatrix} a_1 b_1 + \cdots + a_n b_n \end{bmatrix}$$

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

$$= \begin{bmatrix} 2 & 2 & 0 \\ 4 & 4 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 3 & 4 & 1 \\ -3 & -4 & -1 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 5 & 6 & 1 \\ 1 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

The laws for matrix operations

- Addition laws
  - $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$ (commutative law)
  - $c(\mathbf{A} + \mathbf{B}) = c\mathbf{A} + c\mathbf{B}$ (distributive law)
  - $\mathbf{A} + (\mathbf{B} + \mathbf{C}) = (\mathbf{A} + \mathbf{B}) + \mathbf{C}$ (associative law)

- Multiplication laws
  - $\mathbf{AB} \neq \mathbf{BA}$ (the commutative "law" is usually broken)
  - $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$ (distributive law from the left)
  - $(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}$ (distributive law from the right)
  - $\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$ (associative law for $\mathbf{ABC}$)

In-Class Exercise

- Find an examples of 2 by 2 matrices E and F such that EF = 0, although no entries of E or F are zero.

- Solution:

$$E = F = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}$$

Linear combination revisit

- Linear combination of vectors

$$\mathbf{u} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \qquad \mathbf{v} = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \qquad \mathbf{w} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

  – The linear combinations of these three vectors are $x_1\mathbf{u} + x_2\mathbf{v} + x_3\mathbf{w}$

$$x_1\mathbf{u} + x_2\mathbf{v} + x_3\mathbf{w} = x_1 \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} x_1 \\ -x_1 + x_2 \\ -x_2 + x_3 \end{bmatrix}$$

Linear combination revisit

- Linear combination of vectors

$$x_1\mathbf{u} + x_2\mathbf{v} + x_3\mathbf{w} = x_1\begin{bmatrix}1\\-1\\0\end{bmatrix} + x_2\begin{bmatrix}0\\1\\-1\end{bmatrix} + x_3\begin{bmatrix}0\\0\\1\end{bmatrix} = \begin{bmatrix}x_1\\-x_1+x_2\\-x_2+x_3\end{bmatrix}$$

- Represent linear combination of vectors using a matrix
  - Form a matrix **A** where vectors **u**, **v**, **w** are the columns of **A**
  - The linear combination of **u**,**v**,**w** is that matrix **A** multiplies the vector **x** = (x_1, x_2, x_3)

$$\mathbf{Ax} = \begin{array}{c}\begin{bmatrix}1 & 0 & 0\\-1 & 1 & 0\\0 & -1 & 1\end{bmatrix}\\ \begin{array}{ccc}\mathbf{u} & \;\;\mathbf{v} & \;\;\mathbf{w}\end{array}\end{array}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}$$

Linear combination revisit

- Linear combination of vectors

$$x_1\mathbf{u} + x_2\mathbf{v} + x_3\mathbf{w} = x_1\begin{bmatrix}1\\-1\\0\end{bmatrix} + x_2\begin{bmatrix}0\\1\\-1\end{bmatrix} + x_3\begin{bmatrix}0\\0\\1\end{bmatrix} = \begin{bmatrix}x_1\\-x_1+x_2\\-x_2+x_3\end{bmatrix}$$

- Represent linear combination of vectors using a matrix
  - Form a matrix **A** where vectors **u**, **v**, **w** are the columns of **A**
  - The linear combination of **u**,**v**,**w** is that matrix **A** multiplies the vector **x** = (x_1, x_2, x_3)

$$\mathbf{Ax} = \begin{bmatrix}1 & 0 & 0\\-1 & 1 & 0\\0 & -1 & 1\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix} = \begin{bmatrix}x_1\\-x_1+x_2\\-x_2+x_3\end{bmatrix}$$

Two different viewpoints of matrix-vector multiplication

- The usual way to view matrix-vector multiplication (the way you may familiar with)
  - Multiplication a row at a time (**row picture**)

$$\mathbf{Ax} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 * x_1 + 0 * x_2 + 0 * x_3 \\ -1 * x_1 + 1 * x_2 + 0 * x_3 \\ 0 * x_1 - 1 * x_2 + 1 * x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ -x_1 + x_2 \\ -x_2 + x_3 \end{bmatrix}$$

- The new way is to view **Ax** as a linear combination of the columns of **A** (**column picture**). Linear combinations are the key to linear algebra.

$$\mathbf{Ax} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} x_1 \\ -x_1 + x_2 \\ -x_2 + x_3 \end{bmatrix}$$

Linear Equations

- Matrix-vector Multiplication

$$\mathbf{Ax} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \mathbf{b}$$

  – Given any input **x** = (x_1, x_2, x_3), we can compute the output **b**

  – E.g.,

$$\mathbf{x} = \begin{bmatrix} 1 \\ 4 \\ 9 \end{bmatrix} \qquad \mathbf{b} = \mathbf{Ax} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 4 \\ 9 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix}$$

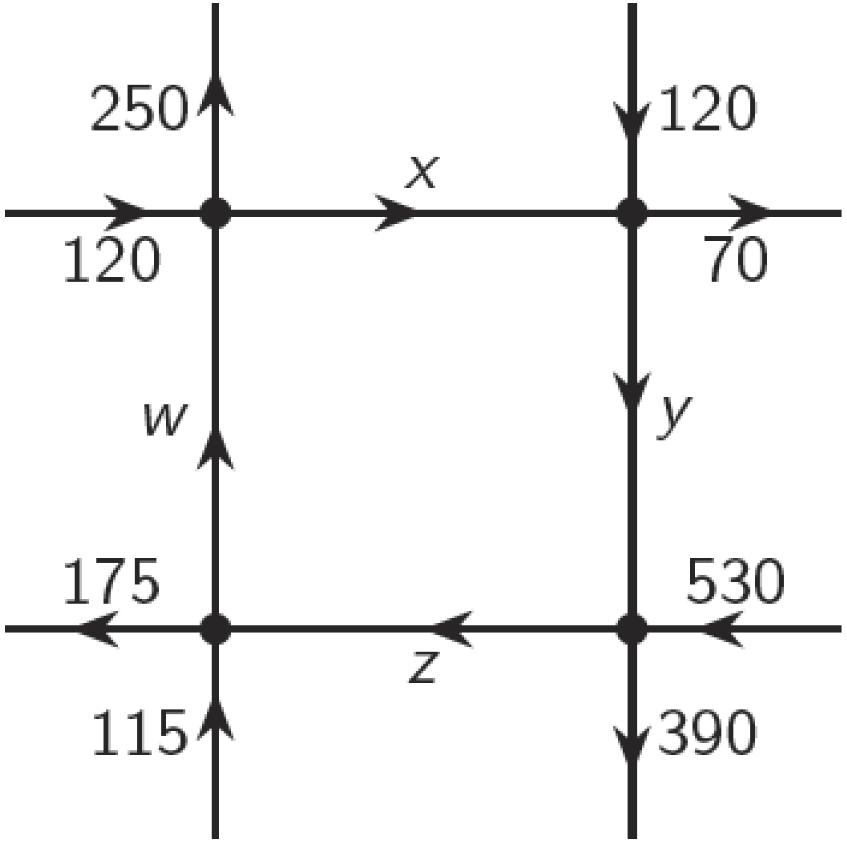
- New question: Given **A**, and **b**, find **x** that satisfy **Ax** = **b**.
  – A system with linear equations.

Application of Linear Equations

- Civil Engineering: How much traffic flows throu the four labeled segments w, x, y, and z?
- System of linear equations:

$$w + 120 = x + 250$$
$$x + 120 = y + 70$$
$$y + 530 = z + 390$$
$$z + 115 = w + 175$$

Traffic flow (cars/hr)



Solving Linear Equations

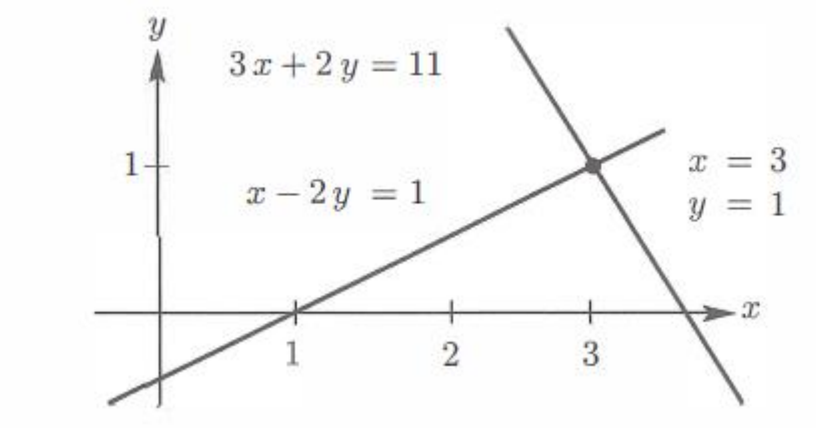
- Solving a system of linear equations is a central problem in linear algebra.

- A small linear system with two equations and two unknown variables.

$$\begin{aligned} x - 2y &= \ \ 1 \\ 3x + 2y &= 11 \end{aligned}$$

Row picture of the linear system

$$\begin{aligned} x - 2y &= \ \ 1 \\ 3x + 2y &= 11 \end{aligned}$$



- Row picture of this linear system
  - The first equation $x - 2y = \ 1$ corresponds to a line in the xy plane.
  - The second equation $3x + 2y = 11$ corresponds to another line in the xy plane.
  - The point (3, 1) where these two lines meet solves both equations.

Column Picture of the linear system

$$\begin{aligned} x - 2y &= \ \ 1 \\ 3x + 2y &= 11 \end{aligned} \quad \Rightarrow \quad x\begin{bmatrix}1\\3\end{bmatrix} + y\begin{bmatrix}-2\\2\end{bmatrix} = \begin{bmatrix}1\\11\end{bmatrix}$$

- This problem is to find the combination of two vector (1,3) and (-2, 2) on the left side that equals to vector (1, 11) on the right side.

- We know x=3 and y = 1 (the same numbers as before) is the solution.

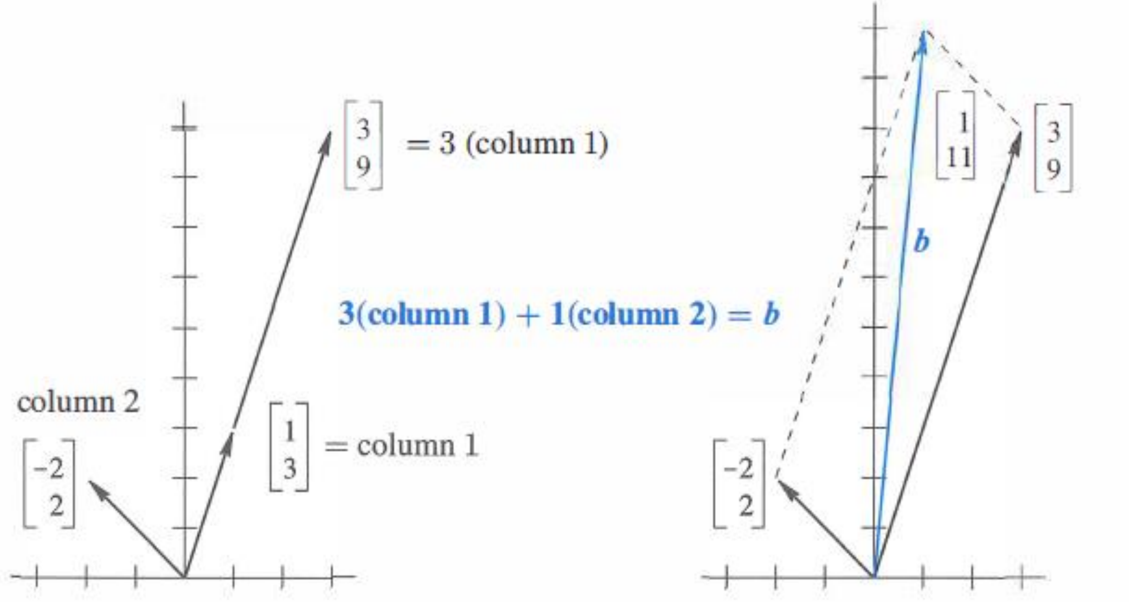
$$3\begin{bmatrix}1\\3\end{bmatrix} + 1\begin{bmatrix}-2\\2\end{bmatrix} = \begin{bmatrix}1\\11\end{bmatrix}$$

Column Picture of the linear system

Vector addition

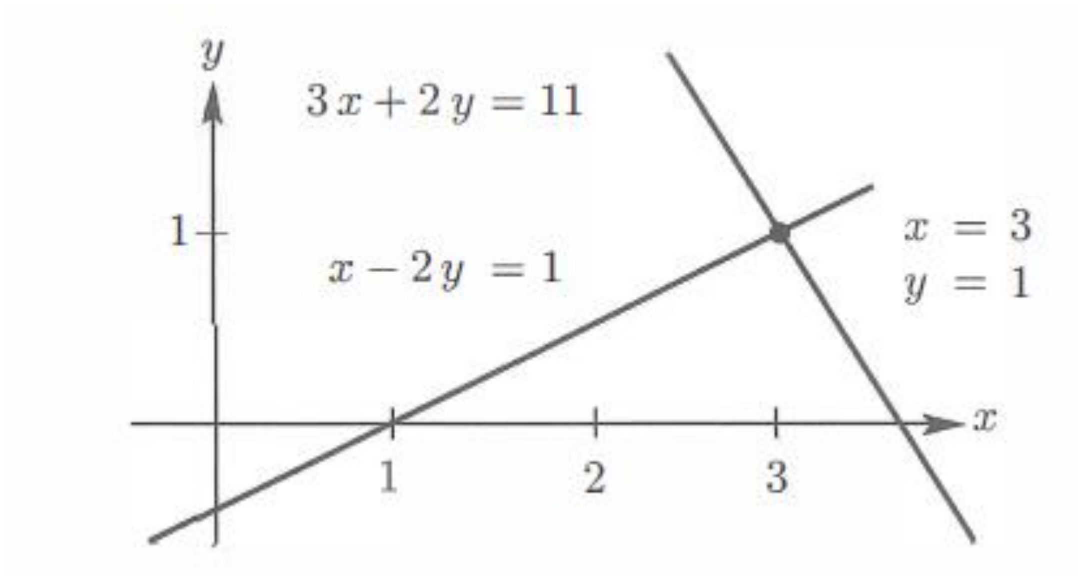
$$\boxed{\boxed{3\begin{bmatrix}1\\3\end{bmatrix}}+\boxed{1\begin{bmatrix}-2\\2\end{bmatrix}}}=\begin{bmatrix}1\\11\end{bmatrix}$$

Scalar multiplication

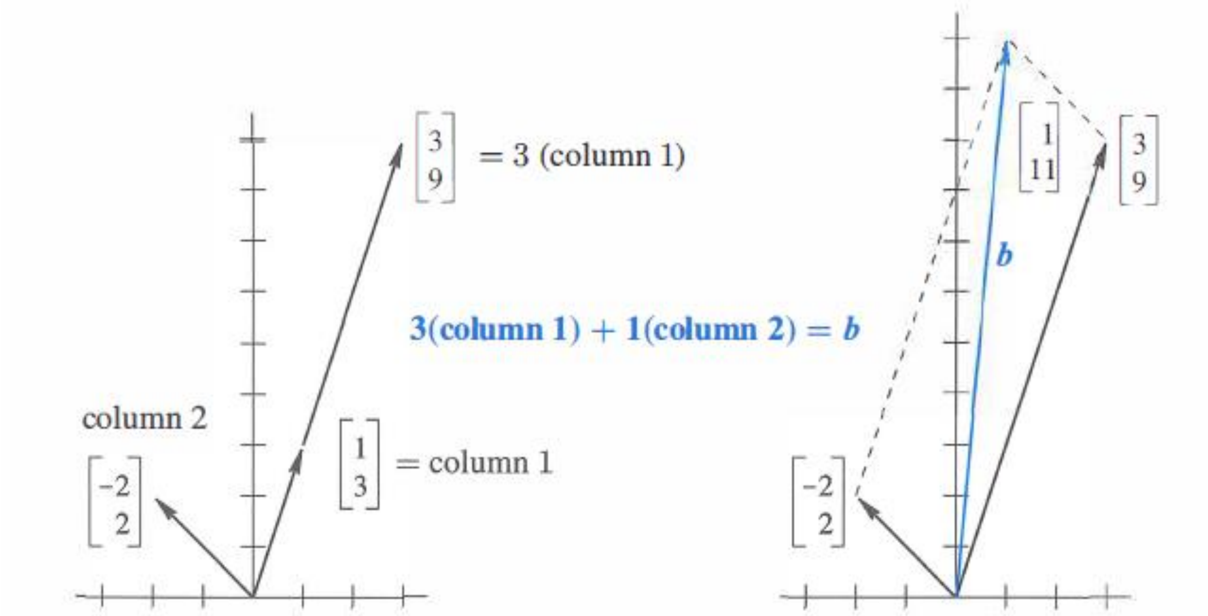


- Linear combination consists of two basic operations: **scalar multiplication** and **vector addition**

Row picture vs. column picture



Row picture



Column picture

The Matrix Form of Linear Equations

- We can represent the linear equations as a matrix problem **Ax** = **b**.

Ax = **b**

$$\begin{aligned} x - 2y &= 1 \\ 3x + 2y &= 11 \end{aligned} \quad \Rightarrow \quad \begin{bmatrix} 1 & -2 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 11 \end{bmatrix}$$

  – The bold capital letter **A** stands for the 2*2 coefficient matrix.
  – The letter bold letter **b** denotes the column vector with two values 1, 11.
  – The unknown variable **x** is also a column vector with two unknown values x an y.

How to systematically solve $\mathbf{Ax} = \mathbf{b}$?

- Suppose $\mathbf{A}$ is a square matrix (n unknown variables with n equations) and $\mathbf{A}$ is invertible. Then, the solution of $\mathbf{Ax} = \mathbf{b}$ is $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$ where $\mathbf{A}^{-1}$ is the inverse of matrix $\mathbf{A}$.

- Why?
  – Let discuss the following concepts before we talk about this solution.
    - Identity matrix
    - Inverse matrix

Identity matrix

- Let's see the following 2*2 matrix and 3*3 matrix

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

- These two matrices $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ are a special type of matrix. They are **identity matrices**. Identity matrix is a square matrix that has 1s on the "main diagonal" and 0s everywhere else. Whatever vector this identity matrix multiplies, that vector is not changed.

- This is like multiply by 1, but for matrices and vectors.

In-Class Exercise

- If $AB = I$ and $BC = I$, prove that $A = C$.

- Solution:

  Using the associative law, we have:

$$A = AI = A(BC) = ABC = (AB)C = IC = C$$

Inverse Matrix

- Suppose **A** is a square matrix. The inverse of A is a matrix B such that BA = I and AB = I. There is at most one such B, and it is denoted by A^{-1} (pronounced “A inverse”):

$$A^{-1}A = I \quad \text{and} \quad AA^{-1} = I$$

- Fundamental property is simple: If you multiply by A and then multiply by A^{-1}, you are back where you started.

Uniqueness of Inverse

- The matrix A cannot have two different inverses.
- Proof: Assume that A has two inverse matrices B and C. According to the property of inverse matrix, we have BA = I and AC = I. Then using associative law, we have

$$B = BI = B(AC) = BAC = (BA)C = IC = C$$

Property of Inverse

- The product AB of invertible matrices is inverted by $B^{-1}A^{-1}$.
- Proof: using associative law to remove parentheses, we have

$$(AB)(B^{-1}A^{-1}) = ABB^{-1}A^{-1} = AIA^{-1} = AA^{-1} = I$$

$$(B^{-1}A^{-1})(AB) = B^{-1}A^{-1}AB = B^{-1}IB = B^{-1}B = I.$$

- A similar rule holds with three or more matrices:

Inverse of ABC $\qquad (ABC)^{-1} = C^{-1}B^{-1}A^{-1}.$

In-Class Exercise

- If A is invertible and AB = AC, prove that B = C.

- Proof: From AB = AC, we have AB - AC = A(B-C) = 0. Since A is invertible, we have
- $(B-C) = A^{-1}A(B-C) = A^{-1}0 = 0 \Rightarrow B = C.$

Systematically solve a system of linear equations

- Solve $\mathbf{Ax} = \mathbf{b}$

- Suppose $\mathbf{A}$ is a square matrix and $\mathbf{A}$ is invertible.

$$\mathbf{Ax} = \mathbf{b} \quad \text{(multiply both sides by } \mathbf{A}^{-1}\text{)}$$

$$\mathbf{A}^{-1}\mathbf{Ax} = \mathbf{A}^{-1}\mathbf{b} \quad (\mathbf{A}^{-1}\mathbf{A} = \mathbf{I})$$

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$$

Questions (We will answer them in the following lectures):

1. Which square matrices are invertible?
2. What if A is not invertible?
3. What if A is not a square matrix?

References and Acknowledgement

- Strang G. Introduction to linear algebra[M]. Wellesley-Cambridge Press, 2022.

The End