Linear Algebra

- Orthogonality -

Jinsun Park

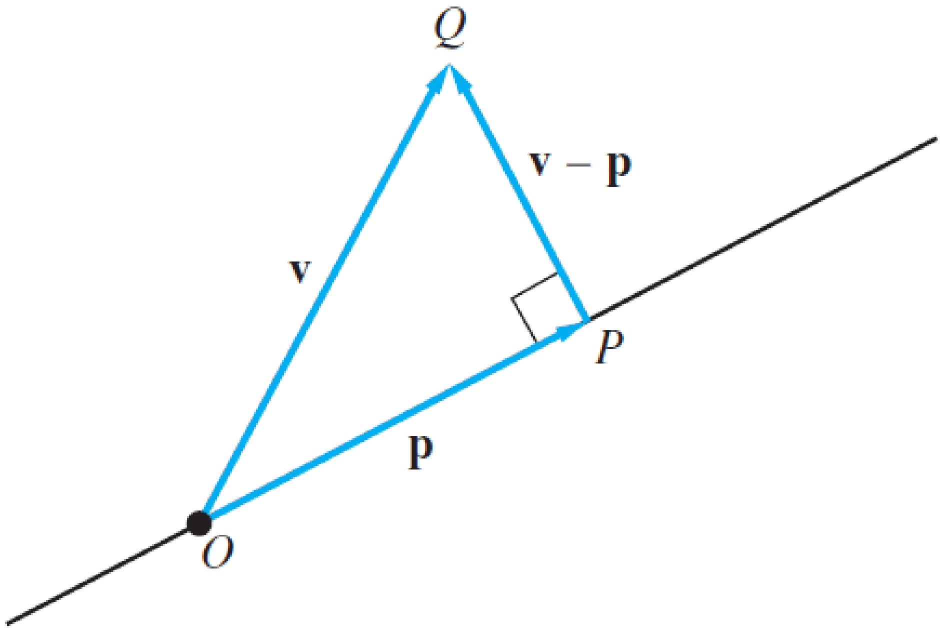
Visual Intelligence and Perception Lab., CSE, PNU

Orthogonality

Orthogonality

- Let l be a line passing through the origin, and let Q be a point not on l. Find the point P on l that is closest to Q

- Condition: QP is perpendicular to OP



- We can think of orthogonality as a generalization of the concept of perpendicularity to any vector space with an inner product

The Scalar Product in $\mathbb{R}^n$

The Scalar Product in $\mathbb{R}^2$ and $\mathbb{R}^3$

- Two vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ in $\mathbb{R}^n$ may be regarded as $n \times 1$ matrices
- The product $\boldsymbol{x}^T\boldsymbol{y}$ is called the *scalar product* of $\boldsymbol{x}$ and $\boldsymbol{y}$

$$\boldsymbol{x}^T\boldsymbol{y} = x_1y_1 + x_2y_2 + \cdots + x_ny_n$$

- Given a vector $\boldsymbol{x}$ in either $\mathbb{R}^2$ or $\mathbb{R}^3$, its *Euclidean length* can be defined in terms of the scalar product

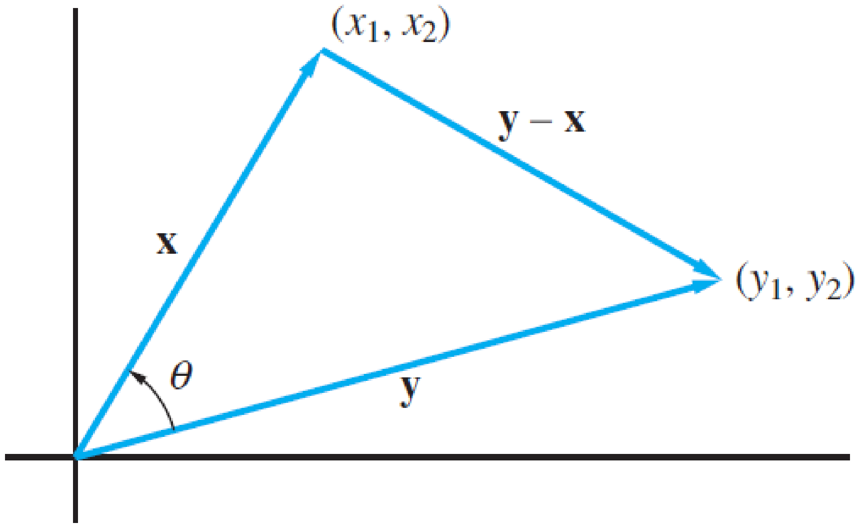
$$\|\boldsymbol{x}\| = (\boldsymbol{x}^T\boldsymbol{x})^{\frac{1}{2}} = \begin{cases} \sqrt{x_1^2 + x_2^2} & \text{if } \boldsymbol{x} \in \mathbb{R}^2 \\ \sqrt{x_1^2 + x_2^2 + x_3^2} & \text{if } \boldsymbol{x} \in \mathbb{R}^3 \end{cases}$$

The Scalar Product in $\mathbb{R}^2$ and $\mathbb{R}^3$

- Given two nonzero vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ as directed line segments starting at the same point, the angle between the two vectors is defined as the angle θ between the line segments

- The distance between the vectors can be measured by the length of the vector joining the terminal point of $\boldsymbol{x}$ to that of $\boldsymbol{y}$

- Let $\boldsymbol{x}$ and $\boldsymbol{y}$ be vectors in either $\mathbb{R}^2$ or $\mathbb{R}^3$. The distance between $\boldsymbol{x}$ and $\boldsymbol{y}$ is defined to be the number $||\boldsymbol{x} - \boldsymbol{y}||$



The Scalar Product in $\mathbb{R}^2$ and $\mathbb{R}^3$

- If $\boldsymbol{x}$ and $\boldsymbol{y}$ are two nonzero vectors in either $\mathbb{R}^2$ or $\mathbb{R}^3$ and θ is the angle between them, then

$$\boldsymbol{x}^T\boldsymbol{y} = \|\boldsymbol{x}\|\|\boldsymbol{y}\|\cos\theta$$

- Proof

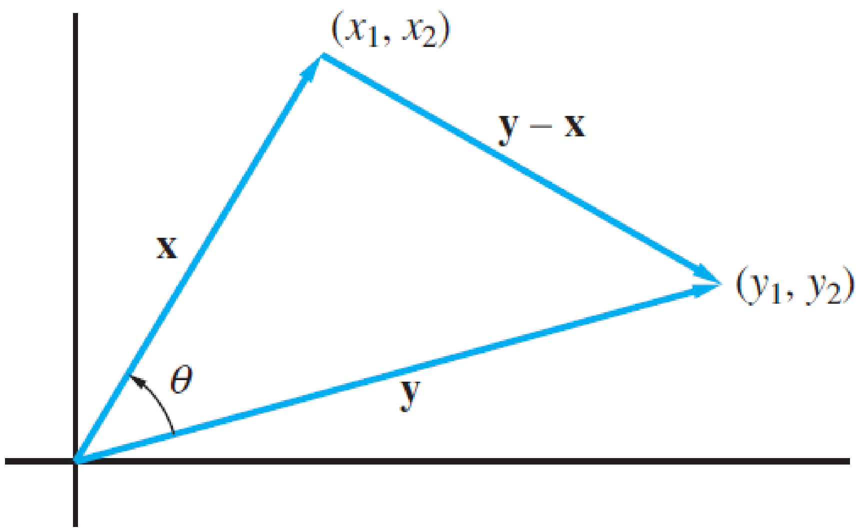
The vectors $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{y}-\mathbf{x}$ may be used to form a triangle as in Figure 5.1.1. By the law of cosines, we have

$$\|\mathbf{y}-\mathbf{x}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - 2\|\mathbf{x}\|\,\|\mathbf{y}\|\cos\theta$$

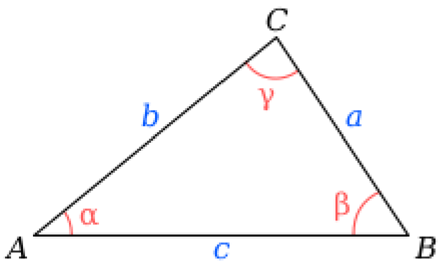
and hence it follows that

$$\begin{aligned}\|\mathbf{x}\|\,\|\mathbf{y}\|\cos\theta &= \tfrac{1}{2}(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - \|\mathbf{y}-\mathbf{x}\|^2)\\ &= \tfrac{1}{2}(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - (\mathbf{y}-\mathbf{x})^T(\mathbf{y}-\mathbf{x}))\\ &= \tfrac{1}{2}(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - (\mathbf{y}^T\mathbf{y} - \mathbf{y}^T\mathbf{x} - \mathbf{x}^T\mathbf{y} + \mathbf{x}^T\mathbf{x}))\\ &= \mathbf{x}^T\mathbf{y}\end{aligned}$$

■



Law of cosines



$$c^2 = a^2 + b^2 - 2ab\cos\gamma,$$
$$a^2 = b^2 + c^2 - 2bc\cos\alpha,$$
$$b^2 = a^2 + c^2 - 2ac\cos\beta.$$

https://en.wikipedia.org/wiki/Law_of_cosines

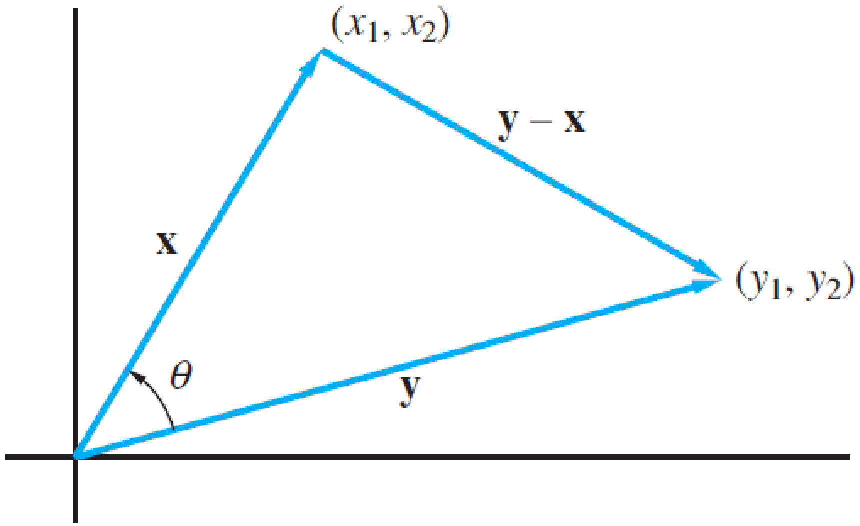
The Scalar Product in $\mathbb{R}^2$ and $\mathbb{R}^3$

- If $\boldsymbol{x}$ and $\boldsymbol{y}$ are nonzero vectors, we can specify their directions by forming unit vectors

$$\boldsymbol{u} = \frac{1}{\|\boldsymbol{x}\|}\boldsymbol{x} \qquad \boldsymbol{v} = \frac{1}{\|\boldsymbol{y}\|}\boldsymbol{y}$$

- If θ is the angle between $\boldsymbol{x}$ and $\boldsymbol{y}$, then

$$\cos\theta = \frac{\boldsymbol{x}^T\boldsymbol{y}}{\|\boldsymbol{x}\|\|\boldsymbol{y}\|} = \boldsymbol{u}^T\boldsymbol{v}$$



Cauchy-Schwarz Inequality

- If $\boldsymbol{x}$ and $\boldsymbol{y}$ are vectors in either $\mathbb{R}^2$ or $\mathbb{R}^3$, then

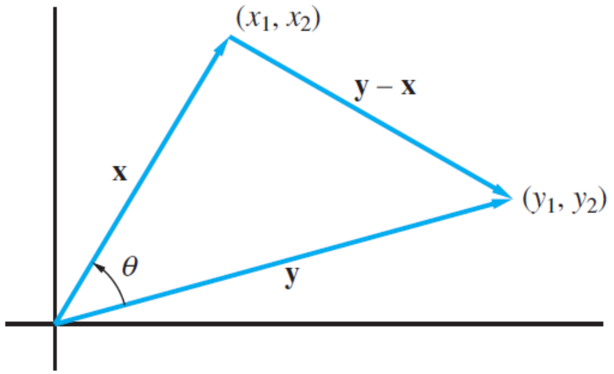
$$|\boldsymbol{x}^T\boldsymbol{y}| \leq \|\boldsymbol{x}\|\|\boldsymbol{y}\|$$

  with equality holding if and only if one of the vectors is $\mathbf{0}$ or one vector is a multiple of the other

- Proof

$$(1)\ \boldsymbol{x}^T\boldsymbol{y} = \|\boldsymbol{x}\|\|\boldsymbol{y}\|\cos\theta$$

The inequality follows from (1). If one of the vectors is $\mathbf{0}$, then both sides of (2) are 0. If both vectors are nonzero, it follows from (1) that equality can hold in (2) if and only if $\cos\theta = \pm 1$. But this would imply that the vectors are either in the same or opposite directions and hence that one vector must be a multiple of the other. ■



- If $\boldsymbol{x}^T\boldsymbol{y} = 0$, either one of the vectors is the zero vector or $\cos\theta = 0$. If $\cos\theta = 0$, the angle between the vectors is a right angle

- The vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ in either $\mathbb{R}^2$ or $\mathbb{R}^3$ are said to be *orthogonal* if $\boldsymbol{x}^T\boldsymbol{y} = 0$
  - The vector $\mathbf{0}$ is orthogonal to every vector in the vector space

Scalar and Vector Projections

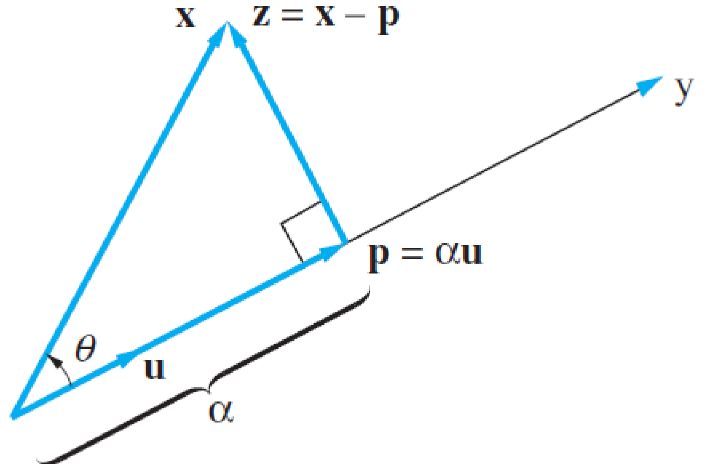
- Let $\boldsymbol{x}$ and $\boldsymbol{y}$ be nonzero vectors in either $\mathbb{R}^2$ or $\mathbb{R}^3$

- We would like to write $\boldsymbol{x}$ as a sum of the form $\boldsymbol{p} + \boldsymbol{z}$
  - $\boldsymbol{p}$ is in the direction of $\boldsymbol{y}$ (and the unit vector $\boldsymbol{u} \parallel \boldsymbol{y}$)
  - $\boldsymbol{z}$ is orthogonal to $\boldsymbol{p}$

- We wish to find α such that $\boldsymbol{p} = \alpha\boldsymbol{u}$ is orthogonal to $\boldsymbol{z} = \boldsymbol{x} - \alpha\boldsymbol{u}$.

- The scalar α must satisfy

$$\alpha = \|\boldsymbol{x}\| \cos\theta = \frac{\|\boldsymbol{x}\|\|\boldsymbol{y}\| \cos\theta}{\|\boldsymbol{y}\|} = \frac{\boldsymbol{x}^T\boldsymbol{y}}{\|\boldsymbol{y}\|}$$



- The scalar α is called the *scalar projection* of $\boldsymbol{x}$ onto $\boldsymbol{y}$, and the vector $\boldsymbol{p}$ is called the *vector projection* of $\boldsymbol{x}$ onto $\boldsymbol{y}$

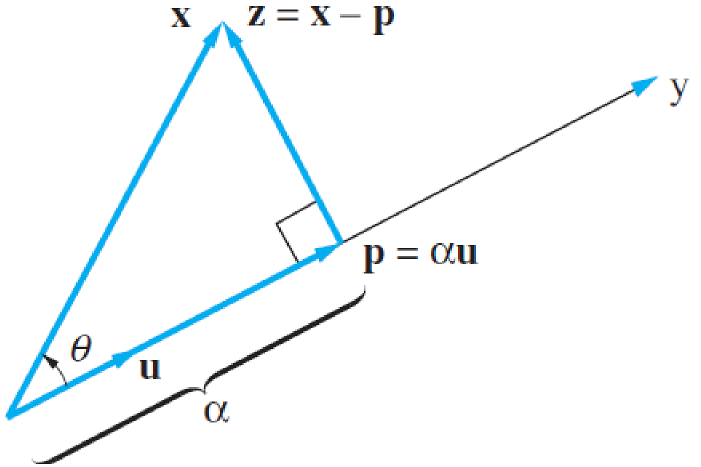
Scalar and Vector Projections

- Scalar projection of $\boldsymbol{x}$ onto $\boldsymbol{y}$

$$\alpha = \frac{\boldsymbol{x}^T\boldsymbol{y}}{\|\boldsymbol{y}\|}$$

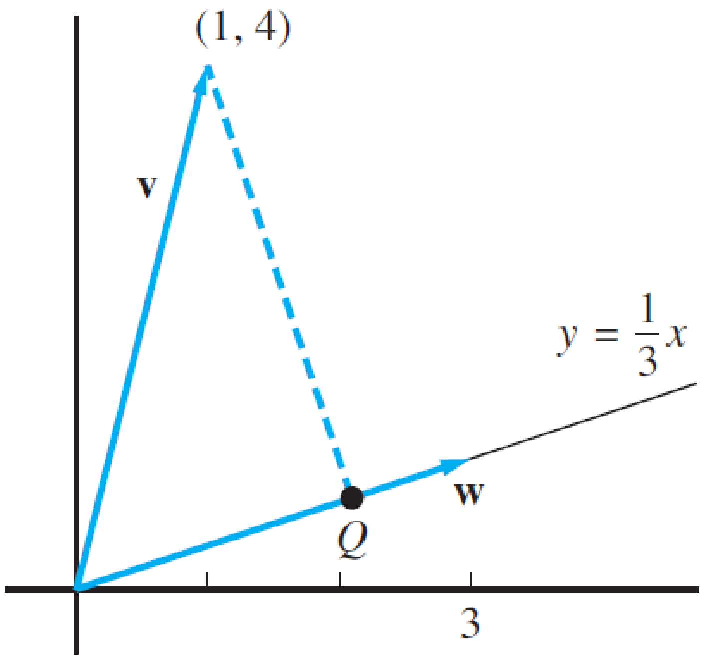
- Vector projection of $\boldsymbol{x}$ onto $\boldsymbol{y}$

$$\boldsymbol{p} = \alpha\boldsymbol{u} = \alpha\frac{1}{\|\boldsymbol{y}\|}\boldsymbol{y} = \frac{\boldsymbol{x}^T\boldsymbol{y}}{\boldsymbol{y}^T\boldsymbol{y}}\boldsymbol{y}$$



Scalar and Vector Projections

Ex 5. Determine the coordinates of Q



- Solution

The vector $\mathbf{w} = (3, 1)^T$ is a vector in the direction of the line $y = \frac{1}{3}x$. Let $\mathbf{v} = (1, 4)^T$. If Q is the desired point, then Q^T is the vector projection of $\mathbf{v}$ onto $\mathbf{w}$.

$$Q^T = \left(\frac{\mathbf{v}^T\mathbf{w}}{\mathbf{w}^T\mathbf{w}}\right)\mathbf{w} = \frac{7}{10}\begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 2.1 \\ 0.7 \end{bmatrix}$$

Thus, $Q = (2.1, 0.7)$ is the closest point. ■

Scalar and Vector Projections

- If $\boldsymbol{N}$ is a nonzero vector and P_0 is a fixed point, the set of points P such that $\overrightarrow{P_0P}$ is orthogonal to $\boldsymbol{N}$ forms a plane π in 3-space that passes through P_0

- The vector $\boldsymbol{N}$ and the plane π are said to be normal to each other

- A point $P = (x, y, z)$ will lie on π if and only if

$$\left(\overrightarrow{P_0P}\right)^T \boldsymbol{N} = 0$$

- If $\boldsymbol{N} = (a, b, c)^T$ and $P_0 = (x_0, y_0, z_0)$, this equation can be written in the form

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0$$

Scalar and Vector Projections

Ex 6. Find the equation of the plane passing through the point $(2, -1, 3)$ and normal to the vector $N = (2, 3, 4)^T$

- Solution

$\overrightarrow{P_0P} = (x - 2, y + 1, z - 3)^T$. The equation is $(\overrightarrow{P_0P})^T \mathbf{N} = 0$, or

$$2(x - 2) + 3(y + 1) + 4(z - 3) = 0$$

Scalar and Vector Projections

Ex 7. Find the equation of the plane that passes through the points

$$P_1 = (1,1,2) \qquad P_2 = (2,3,3) \qquad P_3 = (3,-3,3)$$

- Solution

Let

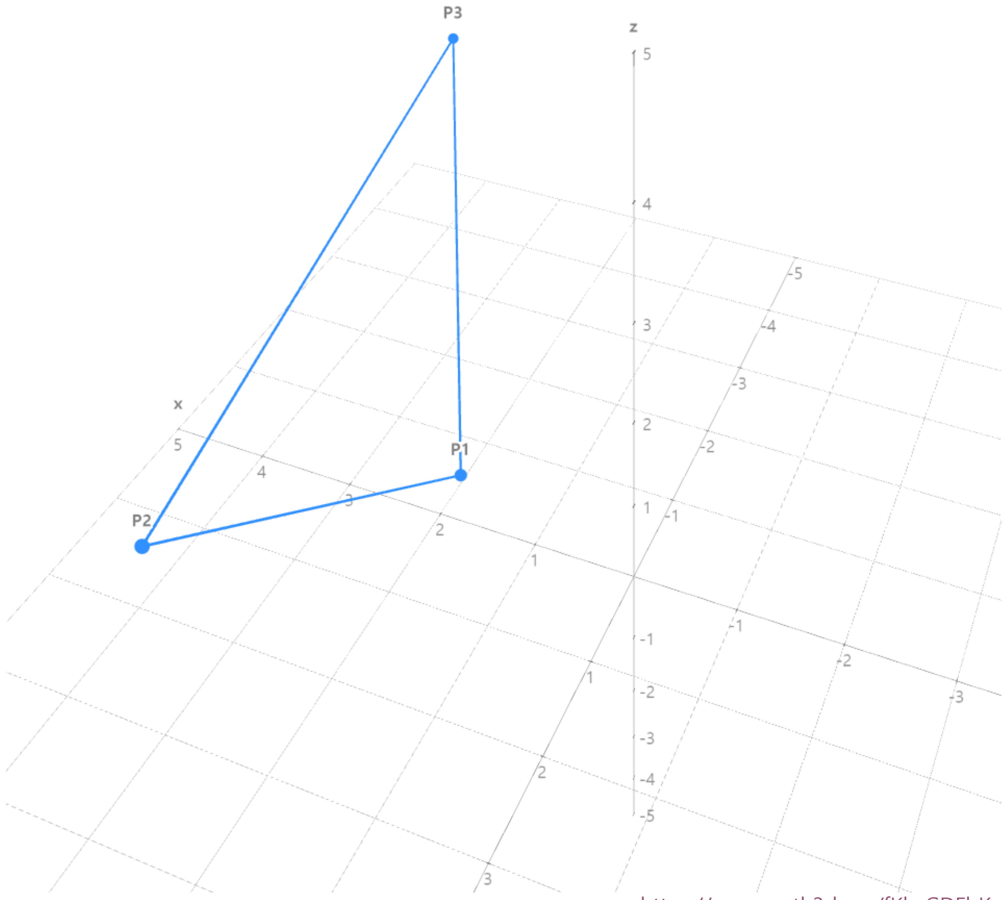
$$\mathbf{x} = \overrightarrow{P_1P_2} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \quad \text{and} \quad \mathbf{y} = \overrightarrow{P_1P_3} = \begin{bmatrix} 2 \\ -4 \\ 1 \end{bmatrix}$$

The normal vector $\mathbf{N}$ must be orthogonal to both $\mathbf{x}$ and $\mathbf{y}$. If we set

$$\mathbf{N} = \mathbf{x} \times \mathbf{y} = \begin{bmatrix} 6 \\ 1 \\ -8 \end{bmatrix}$$

then $\mathbf{N}$ will be a normal vector to the plane that passes through the given points. We can then use any one of the points to determine the equation of the plane. Using the point P_1, we see that the equation of the plane is

$$6(x-1) + (y-1) - 8(z-2) = 0 \qquad \blacksquare$$



https://www.math3d.org/fKbsGDFbK

Scalar and Vector Projections

Ex 8. Find the distance from the point $(2,0,0)$ to the plane $x+2y+2z=0$

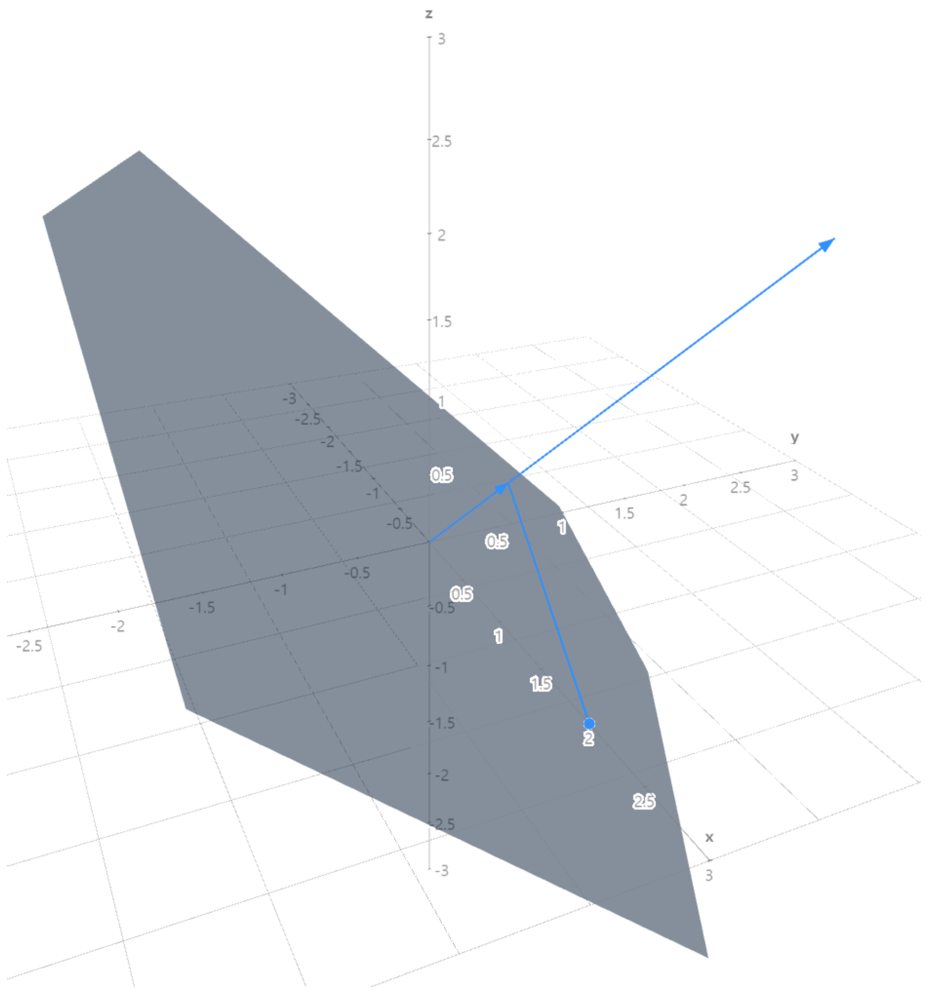
- Solution

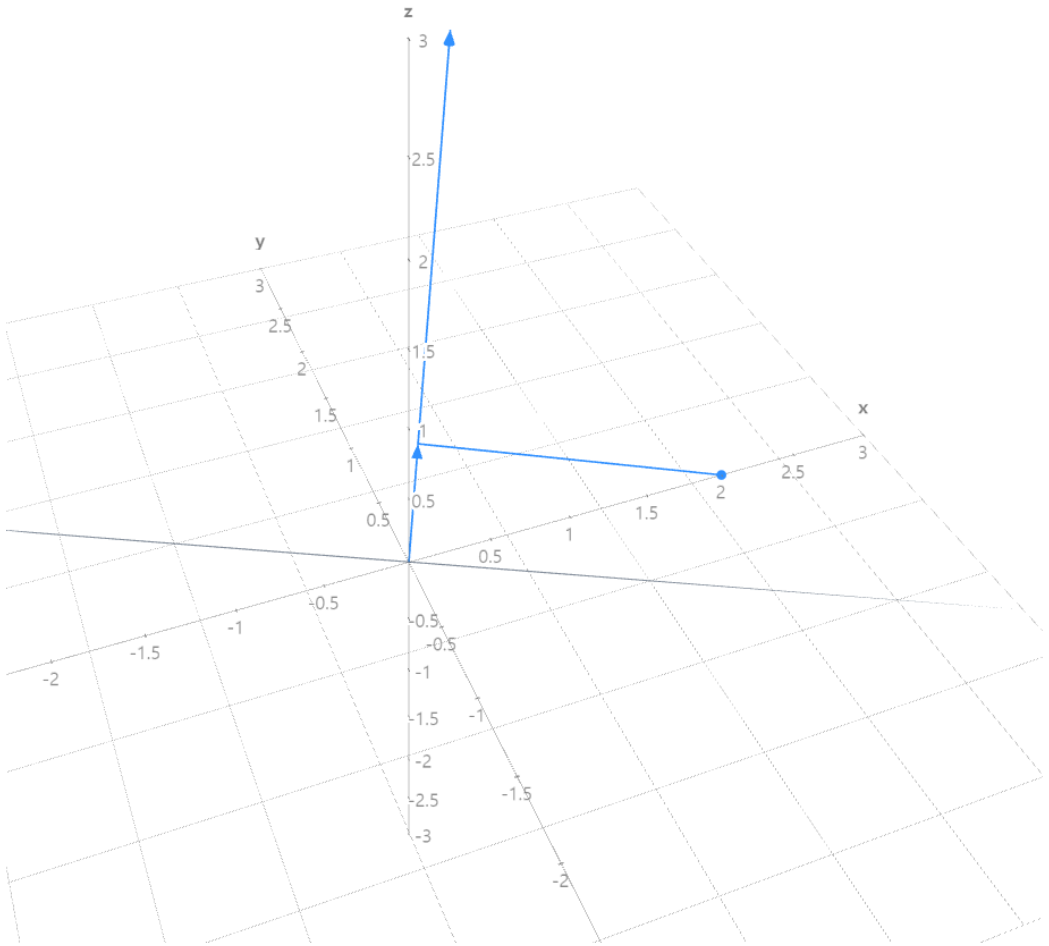
The vector $\mathbf{N} = (1, 2, 2)^T$ is normal to the plane and the plane passes through the origin. Let $\mathbf{v} = (2, 0, 0)^T$. The distance d from $(2, 0, 0)$ to the plane is simply the absolute value of the scalar projection of $\mathbf{v}$ onto $\mathbf{N}$. Thus,

$$d = \frac{|\mathbf{v}^T\mathbf{N}|}{\|\mathbf{N}\|} = \frac{2}{3} \qquad \blacksquare$$

Scalar and Vector Projections

Ex 8. Find the distance from the point $(2, 0, 0)$ to the plane $x + 2y + 2z = 0$





https://www.math3d.org/du87WQD3E

Scalar and Vector Projections

- For any nonzero vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ in $\mathbb{R}^3$

$$\|\boldsymbol{x} \times \boldsymbol{y}\| = \|\boldsymbol{x}\|\|\boldsymbol{y}\| \sin\theta$$

- Proof

$$\sin\theta = \sqrt{1-\cos^2\theta} = \sqrt{1-\frac{(\boldsymbol{x}^T\boldsymbol{y})^2}{\|\boldsymbol{x}\|^2\|\boldsymbol{y}\|^2}} = \frac{\sqrt{\|\boldsymbol{x}\|^2\|\boldsymbol{y}\|^2-(\boldsymbol{x}^T\boldsymbol{y})^2}}{\|\boldsymbol{x}\|\|\boldsymbol{y}\|}$$

$$\|\boldsymbol{x}\|\|\boldsymbol{y}\|\sin\theta = \sqrt{\|\boldsymbol{x}\|^2\|\boldsymbol{y}\|^2-(\boldsymbol{x}^T\boldsymbol{y})^2} = \sqrt{(x_1^2+x_2^2+x_3^2)(y_1^2+y_2^2+y_3^2)-(x_1y_1+x_2y_2+x_3y_3)^2}$$
$$= \sqrt{(x_2y_3-x_3y_2)^2+(x_3y_1-x_1y_3)^2+(x_1y_2-x_2y_1)^2} = \|\boldsymbol{x} \times \boldsymbol{y}\|$$

- If either $\boldsymbol{x}$ or $\boldsymbol{y}$ is the zero vector, then $\boldsymbol{x} \times \boldsymbol{y} = \boldsymbol{0}$ and hence the norm of $\boldsymbol{x} \times \boldsymbol{y}$ will be 0

Orthogonality in $\mathbb{R}^n$

- The definitions that have been given for $\mathbb{R}^2$ and $\mathbb{R}^3$ can all be generalized to $\mathbb{R}^n$

- If $\boldsymbol{x} \in \mathbb{R}^n$, then the Euclidean length of $\boldsymbol{x}$ is defined by

$$\|\boldsymbol{x}\| = (\boldsymbol{x}^T\boldsymbol{x})^{\frac{1}{2}} = (x_1^2 + x_2^2 + \cdots + x_n^2)^{\frac{1}{2}}$$

- If $\boldsymbol{x}$ and $\boldsymbol{y}$ are two vectors in $\mathbb{R}^n$, then the distance between the vectors is $\|\boldsymbol{x} - \boldsymbol{y}\|$

- The Cauchy-Schwarz inequality holds in $\mathbb{R}^n$. Consequently

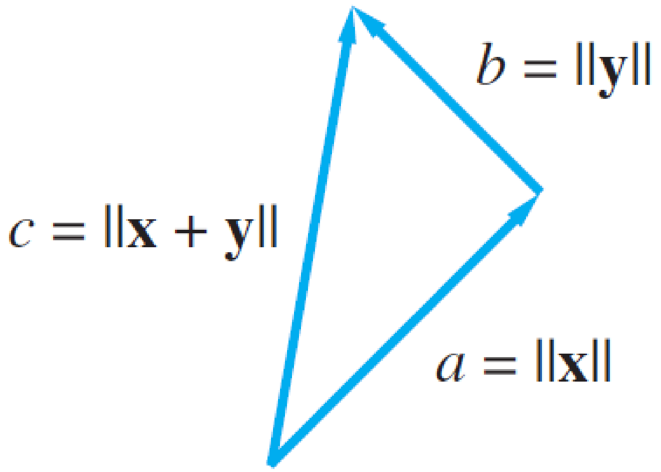
$$-1 \leq \frac{\boldsymbol{x}^T\boldsymbol{y}}{\|\boldsymbol{x}\|\|\boldsymbol{y}\|} \leq 1$$

for any nonzero vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ in $\mathbb{R}^n$

Orthogonality in $\mathbb{R}^n$

- The angle θ between two nonzero vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ in $\mathbb{R}^n$ is given by

$$\cos\theta = \frac{\boldsymbol{x}^T\boldsymbol{y}}{\|\boldsymbol{x}\|\|\boldsymbol{y}\|} \qquad 0 \le \theta \le \pi$$



- The vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ are said to be orthogonal if $\boldsymbol{x}^T\boldsymbol{y} = 0$
  - The symbol $\perp$ is used to indicate orthogonality
  - $\boldsymbol{x} \perp \boldsymbol{y}$

- If $\boldsymbol{x}$ and $\boldsymbol{y}$ are vectors in $\mathbb{R}^n$, then

$$\|\boldsymbol{x}+\boldsymbol{y}\|^2 = (\boldsymbol{x}+\boldsymbol{y})^T(\boldsymbol{x}+\boldsymbol{y}) = \|\boldsymbol{x}\|^2 + 2\boldsymbol{x}^T\boldsymbol{y} + \|\boldsymbol{y}\|^2$$

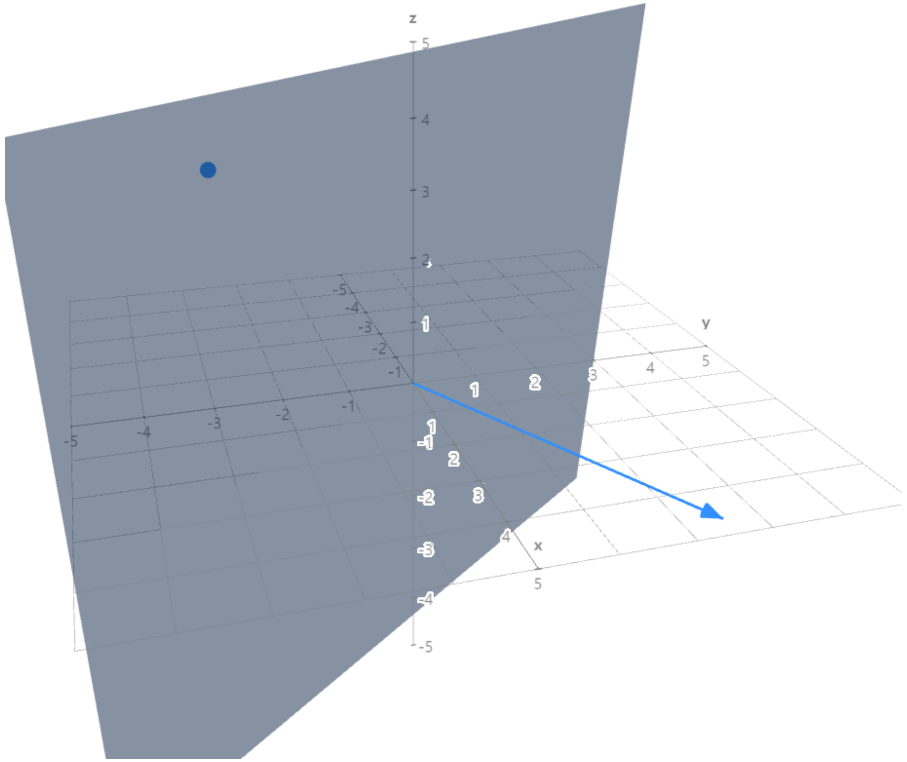
- In the case that $\boldsymbol{x}$ and $\boldsymbol{y}$ are orthogonal, the equation becomes the *Pythagorean law*

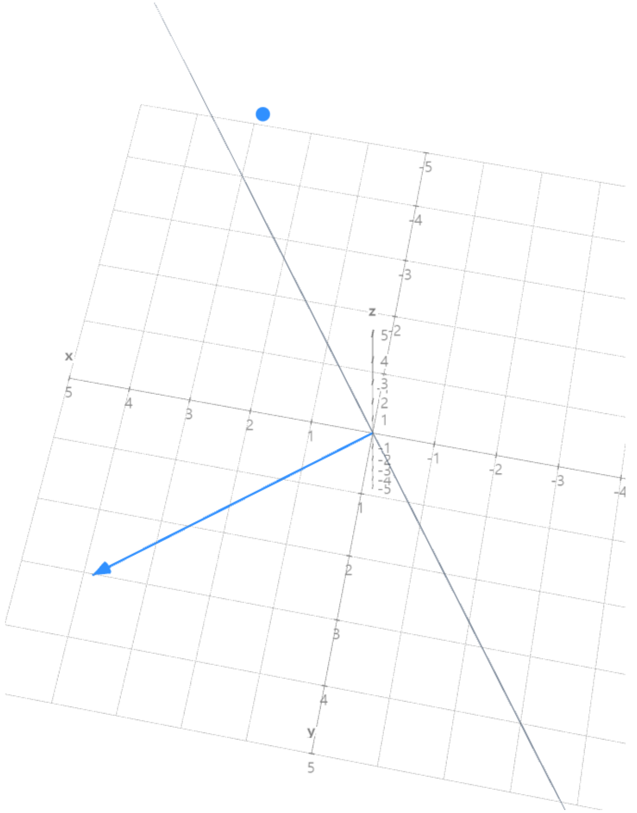
$$\|\boldsymbol{x}+\boldsymbol{y}\|^2 = \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2$$

Exercises

11. Find the distance from the point $(2, -3, 4)$ to the plane

$$8(x-2)+6(y+2)-(z-4)=0$$





https://www.math3d.org/vZUWW8fkA

Orthogonal Subspaces

Orthogonal Subspaces

- Let A be an $m \times n$ matrix and let $\boldsymbol{x} \in N(A)$, the null space of A. Since $A\boldsymbol{x} = 0$,

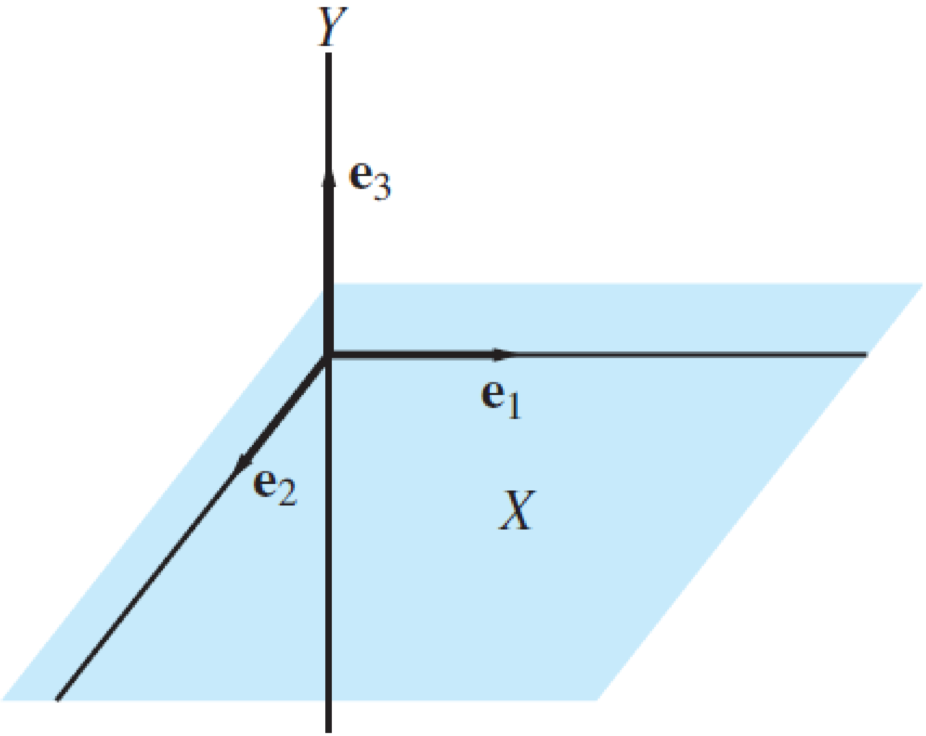
$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = 0 \qquad \text{for } i = 1, \cdots, m$$

- $\boldsymbol{x}$ is orthogonal to the i-th column vector of A^T for $i = 1, \cdots, m$
- Each vector in $N(A)$ is orthogonal to every vector in the column space of A^T

- Two subspaces X and Y of $\mathbb{R}^n$ are said to be orthogonal if $\boldsymbol{x}^T\boldsymbol{y} = 0$ for every $\boldsymbol{x} \in X$ and every $\boldsymbol{y} \in Y$.
- If X and Y are orthogonal, we write $X \perp Y$

Orthogonal Subspaces

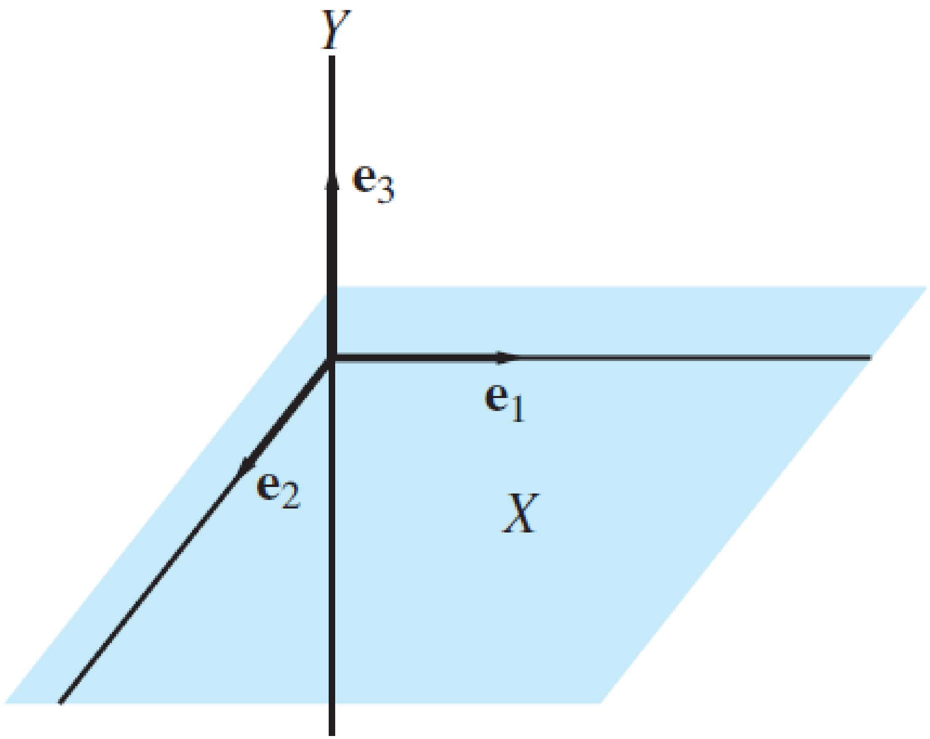
- Are xy-plane and yz-plane orthogonal?



$\boldsymbol{x}_1 = (1, 1, 0)^T$ $\boldsymbol{x}_2 = (0, 1, 1)^T$

Orthogonal Subspaces

Ex 2. Let X be the subspace of $\mathbb{R}^3$ spanned by $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$, and let Y be the subspace spanned by $\boldsymbol{e}_3$



Orthogonal Subspaces

- Let Y be a subspace of $\mathbb{R}^n$. The set of all vectors in $\mathbb{R}^n$ that are orthogonal to every vector in Y will be denoted $Y^\perp$

$$Y^\perp = \{\boldsymbol{x} \in \mathbb{R}^n | \boldsymbol{x}^T \boldsymbol{y} = 0 \quad \text{for every} \quad \boldsymbol{y} \in Y\}$$

- The set $Y^\perp$ is called the *orthogonal complement* of Y (The symbol is sometimes read "Y perp")

- The subspaces $X = Span(\boldsymbol{e}_1)$ and $Y = Span(\boldsymbol{e}_2)$ of $\mathbb{R}^3$ are orthogonal, but they are not orthogonal complements

$$X^\perp = Span(\boldsymbol{e}_2, \boldsymbol{e}_3) \qquad Y^\perp = Span(\boldsymbol{e}_1, \boldsymbol{e}_3)$$

Orthogonal Subspaces

1) If X and Y are orthogonal subspaces of $\mathbb{R}^n$, then $X \cap Y = \{\mathbf{0}\}$

2) If Y is a subspace of $\mathbb{R}^n$, then $Y^\perp$ is also a subspace of $\mathbb{R}^n$

- Proof

Proof of (1) If $\mathbf{x} \in X \cap Y$ and $X \perp Y$, then $\|\mathbf{x}\|^2 = \mathbf{x}^T\mathbf{x} = 0$ and hence $\mathbf{x} = \mathbf{0}$.

Proof of (2) If $\mathbf{x} \in Y^\perp$ and α is a scalar, then, for any $\mathbf{y} \in Y$,

$$(\alpha\mathbf{x})^T\mathbf{y} = \alpha(\mathbf{x}^T\mathbf{y}) = \alpha \cdot 0 = 0$$

Therefore, $\alpha\mathbf{x} \in Y^\perp$. If $\mathbf{x}_1$ and $\mathbf{x}_2$ are elements of $Y^\perp$, then

$$(\mathbf{x}_1 + \mathbf{x}_2)^T\mathbf{y} = \mathbf{x}_1^T\mathbf{y} + \mathbf{x}_2^T\mathbf{y} = 0 + 0 = 0$$

for each $\mathbf{y} \in Y$. Hence, $\mathbf{x}_1 + \mathbf{x}_2 \in Y^\perp$. Therefore, $Y^\perp$ is a subspace of $\mathbb{R}^n$.

Fundamental Subspaces

- Let A be an $m \times n$ matrix. A vector $\boldsymbol{b} \in \mathbb{R}^m$ is in the column space of A if and only if $\boldsymbol{b} = A\boldsymbol{x}$ for some $\boldsymbol{x} \in \mathbb{R}^n$

- If we think of A as a linear transformation mapping $\mathbb{R}^n$ to $\mathbb{R}^m$, then the column space of A is the same as the range of A. Let us denote the range of A by $R(A)$

$$R(A) = \{\boldsymbol{b} \in \mathbb{R}^m | \boldsymbol{b} = A\boldsymbol{x} \quad \text{for some} \quad \boldsymbol{x} \in \mathbb{R}^n\} = \text{The column space of A}$$

- The column space of A^T, $R(A^T)$, is a subspace of $\mathbb{R}^n$

$$R(A^T) = \{\boldsymbol{y} \in \mathbb{R}^n | \boldsymbol{y} = A^T\boldsymbol{x} \quad \text{for some} \quad \boldsymbol{x} \in \mathbb{R}^m\}$$

- $R(A^T)$ is essentially the same as the row space of A. Thus, $\boldsymbol{y} \in R(A^T)$ if and only if $\boldsymbol{y}^T$ is in the row space of A

- For some $\boldsymbol{n} \in N(A)$ and $\boldsymbol{y} \in R(A^T)$, $\boldsymbol{n}^T\boldsymbol{y} = 0$, therefore, $N(A) \perp R(A^T)$

Fundamental Subspaces

- If A is an $m \times n$ matrix, then $N(A) = R(A^T)^{\perp}$ and $N(A^T) = R(A)^{\perp}$

- Proof

On the one hand, we have already seen that $N(A) \perp R(A^T)$, and this implies that $N(A) \subset R(A^T)^{\perp}$. On the other hand, if $\mathbf{x}$ is any vector in $R(A^T)^{\perp}$, then $\mathbf{x}$ is orthogonal to each of the column vectors of A^T and, consequently, $A\mathbf{x} = \mathbf{0}$. Thus, $\mathbf{x}$ must be an element of $N(A)$ and hence $N(A) = R(A^T)^{\perp}$. This proof does not depend on the dimensions of A. In particular, the result will also hold for the matrix $B = A^T$. Consequently,

$$N(A^T) = N(B) = R(B^T)^{\perp} = R(A)^{\perp} \qquad \blacksquare$$

Fundamental Subspaces

- If S is a subspace of $\mathbb{R}^n$, then $\dim S + \dim S^\perp = n$. Furthermore, if $\{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_r\}$ is a basis for S and $\{\boldsymbol{x}_{r+1}, \cdots, \boldsymbol{x}_n\}$ is a basis for $S^\perp$, then $\{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_r, \boldsymbol{x}_{r+1}, \cdots, \boldsymbol{x}_n\}$ is a basis for $\mathbb{R}^n$

- Proof

If $S = \{\mathbf{0}\}$, then $S^\perp = \mathbb{R}^n$ and

$$\dim S + \dim S^\perp = 0 + n = n$$

If $S \neq \{\mathbf{0}\}$, then let $\{\mathbf{x}_1, \ldots, \mathbf{x}_r\}$ be a basis for S and define X to be an $r \times n$ matrix whose ith row is $\mathbf{x}_i^T$ for each i. By construction, the matrix X has rank r and $R(X^T) = S$. By Theorem 5.2.1,

$$S^\perp = R(X^T)^\perp = N(X)$$

It follows from Theorem 3.6.5 that

$$\dim S^\perp = \dim N(X) = n - r$$

To show that $\{\mathbf{x}_1, \ldots, \mathbf{x}_r, \mathbf{x}_{r+1}, \ldots, \mathbf{x}_n\}$ is a basis for $\mathbb{R}^n$, it suffices to show that the n vectors are linearly independent. Suppose that

$$c_1\mathbf{x}_1 + \cdots + c_r\mathbf{x}_r + c_{r+1}\mathbf{x}_{r+1} + \cdots + c_n\mathbf{x}_n = \mathbf{0}$$

Let $\mathbf{y} = c_1\mathbf{x}_1 + \cdots + c_r\mathbf{x}_r$ and $\mathbf{z} = c_{r+1}\mathbf{x}_{r+1} + \cdots + c_n\mathbf{x}_n$. We then have

$$\mathbf{y} + \mathbf{z} = \mathbf{0}$$
$$\mathbf{y} = -\mathbf{z}$$

Thus, $\mathbf{y}$ and $\mathbf{z}$ are both elements of $S \cap S^\perp$. But $S \cap S^\perp = \{\mathbf{0}\}$. Therefore,

$$c_1\mathbf{x}_1 + \cdots + c_r\mathbf{x}_r = \mathbf{0}$$
$$c_{r+1}\mathbf{x}_{r+1} + \cdots + c_n\mathbf{x}_n = \mathbf{0}$$

Since $\mathbf{x}_1, \ldots, \mathbf{x}_r$ are linearly independent,

$$c_1 = c_2 = \cdots = c_r = 0$$

Similarly, $\mathbf{x}_{r+1}, \ldots, \mathbf{x}_n$ are linearly independent and hence

$$c_{r+1} = c_{r+2} = \cdots = c_n = 0$$

So $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ are linearly independent and form a basis for $\mathbb{R}^n$. ■

Fundamental Subspaces

- If U and V are subspaces of a vector space W and each $\boldsymbol{w} \in W$ can be written uniquely as a sum $\boldsymbol{u} + \boldsymbol{v}$, where $\boldsymbol{u} \in U$ and $\boldsymbol{v} \in V$, then we say that W is a *direct sum* of U and V

$$W = U \oplus V$$

- If S is a subspace of $\mathbb{R}^n$, then $\mathbb{R}^n = S \oplus S^\perp$

- Proof

The result is trivial if either $S = \{\mathbf{0}\}$ or $S = \mathbb{R}^n$. In the case where the dimension of S is r, $0 < r < n$, it follows from Theorem 5.2.2 that each vector $\mathbf{x} \in \mathbb{R}^n$ can be represented in the form

$$\mathbf{x} = c_1\mathbf{x}_1 + \cdots + c_r\mathbf{x}_r + c_{r+1}\mathbf{x}_{r+1} + \cdots + c_n\mathbf{x}_n$$

where $\{\mathbf{x}_1, \ldots, \mathbf{x}_r\}$ is a basis for S and $\{\mathbf{x}_{r+1}, \ldots, \mathbf{x}_n\}$ is a basis for $S^\perp$. If we let

$$\mathbf{u} = c_1\mathbf{x}_1 + \cdots + c_r\mathbf{x}_r \quad \text{and} \quad \mathbf{v} = c_{r+1}\mathbf{x}_{r+1} + \cdots + c_n\mathbf{x}_n$$

then $\mathbf{u} \in S$, $\mathbf{v} \in S^\perp$, and $\mathbf{x} = \mathbf{u} + \mathbf{v}$. To show uniqueness, suppose that $\mathbf{x}$ can also be written as a sum $\mathbf{y} + \mathbf{z}$, where $\mathbf{y} \in S$ and $\mathbf{z} \in S^\perp$. Thus,

$$\begin{aligned} \mathbf{u} + \mathbf{v} &= \mathbf{x} = \mathbf{y} + \mathbf{z} \\ \mathbf{u} - \mathbf{y} &= \mathbf{z} - \mathbf{v} \end{aligned}$$

But $\mathbf{u} - \mathbf{y} \in S$ and $\mathbf{z} - \mathbf{v} \in S^\perp$, so each is in $S \cap S^\perp$. Since

$$S \cap S^\perp = \{\mathbf{0}\}$$

it follows that

$$\mathbf{u} = \mathbf{y} \quad \text{and} \quad \mathbf{v} = \mathbf{z}$$

Fundamental Subspaces

- If S is a subspace of $\mathbb{R}^n$, then $(S^{\perp})^{\perp} = S$

- Proof

On the one hand, if $\mathbf{x} \in S$, then $\mathbf{x}$ is orthogonal to each $\mathbf{y}$ in $S^{\perp}$. Therefore, $\mathbf{x} \in (S^{\perp})^{\perp}$ and hence $S \subset (S^{\perp})^{\perp}$. On the other hand, suppose that $\mathbf{z}$ is an arbitrary element of $(S^{\perp})^{\perp}$. By Theorem 5.2.3, we can write $\mathbf{z}$ as a sum $\mathbf{u} + \mathbf{v}$, where $\mathbf{u} \in S$ and $\mathbf{v} \in S^{\perp}$. Since $\mathbf{v} \in S^{\perp}$, it is orthogonal to both $\mathbf{u}$ and $\mathbf{z}$. It then follows that

$$0 = \mathbf{v}^T\mathbf{z} = \mathbf{v}^T\mathbf{u} + \mathbf{v}^T\mathbf{v} = \mathbf{v}^T\mathbf{v}$$

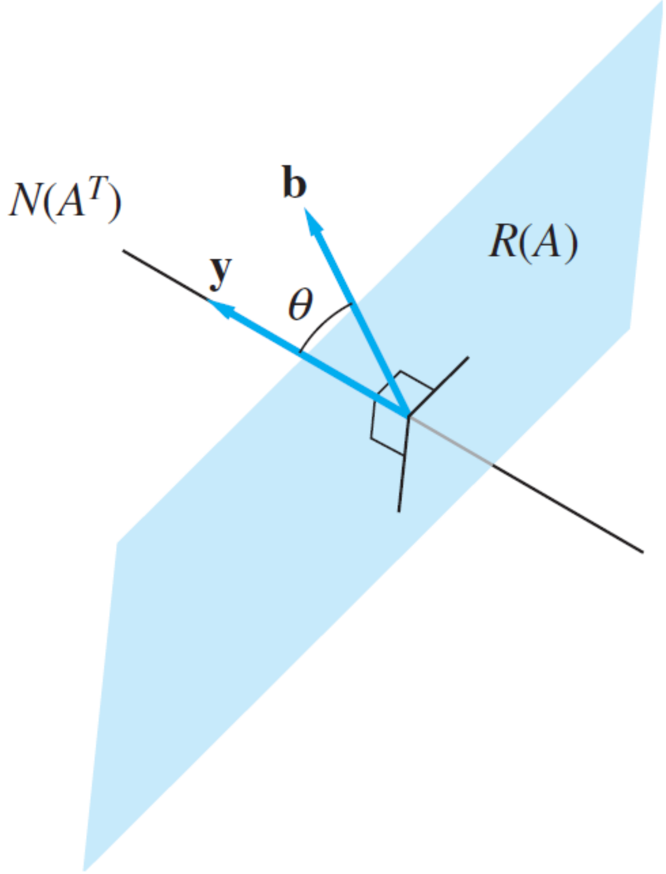
and, consequently, $\mathbf{v} = \mathbf{0}$. Therefore, $\mathbf{z} = \mathbf{u} \in S$ and hence $S = (S^{\perp})^{\perp}$. ■

Fundamental Subspaces

- If A is an $m \times n$ matrix and $\boldsymbol{b} \in \mathbb{R}^m$, then either there is a vector $\boldsymbol{x} \in \mathbb{R}^n$ such that $A\boldsymbol{x} = \boldsymbol{b}$ or there is a vector $\boldsymbol{y} \in \mathbb{R}^m$ such that $A^T\boldsymbol{y} = \boldsymbol{0}$ and $\boldsymbol{y}^T\boldsymbol{b} \neq 0$

- Proof

Corollary 5.2.5 is illustrated in Figure 5.2.2 for the case where $R(A)$ is a two-dimensional subspace of $\mathbb{R}^3$. The angle θ in the figure will be a right angle if and only if $\mathbf{b} \in R(A)$.



Fundamental Subspaces

Ex 4. Let

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

Find the bases for $N(A), R(A^T), N(A^T),$ and $R(A)$

- Solution

We can find bases for $N(A)$ and $R(A^T)$ by transforming A into reduced row echelon form:

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

Since (1, 0, 1) and (0, 1, 1) form a basis for the row space of A, it follows that $(1, 0, 1)^T$ and $(0, 1, 1)^T$ form a basis for $R(A^T)$. If $\mathbf{x} \in N(A)$, it follows from the reduced row echelon form of A that

$$x_1 + x_3 = 0$$
$$x_2 + x_3 = 0$$

Thus,

$$x_1 = x_2 = -x_3$$

Setting $x_3 = \alpha$, we see that $N(A)$ consists of all vectors of the form $\alpha(-1, -1, 1)^T$. Note that $(-1, -1, 1)^T$ is orthogonal to $(1, 0, 1)^T$ and $(0, 1, 1)^T$.

To find bases for $R(A)$ and $N(A^T)$, transform A^T to reduced row echelon form:

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

Thus, $(1, 0, 1)^T$ and $(0, 1, 2)^T$ form a basis for $R(A)$. If $\mathbf{x} \in N(A^T)$, then $x_1 = -x_3$, $x_2 = -2x_3$. Hence, $N(A^T)$ is the subspace of $\mathbb{R}^3$ spanned by $(-1, -2, 1)^T$. Note that $(-1, -2, 1)^T$ is orthogonal to $(1, 0, 1)^T$ and $(0, 1, 2)^T$. ■

Exercises

6. Is it possible for a matrix to have the vector $(1,2,3)$ in its row space and $(2,1,-1)^T$ in its null space?