Section 5.3 : Diagonalization

Chapter 5 : Eigenvalues and Eigenvectors

Math 1554 Linear Algebra

Motivation: it can be useful to take large powers of matrices, for example A^k, for large k.

But: multiplying two $n \times n$ matrices requires roughly n^3 computations. Is there a more efficient way to compute A^k?

Topics and Objectives

Topics

1. Diagonal, similar, and diagonalizable matrices
2. Diagonalizing matrices

Learning Objectives

For the topics covered in this section, students are expected to be able to do the following.

1. Determine whether a matrix can be diagonalized, and if possible diagonalize a square matrix.
2. Apply diagonalization to compute matrix powers.

Note $\det(CD) = \det(D \cdot C) \Rightarrow \det(\underbrace{P \cdot (B-\lambda I)} \cdot \underbrace{P^{-1}})$

$\det(C)\cdot\det(D) = \det(D)\cdot\det(C)$

$= \det(\underbrace{P^{-1} \cdot P}_{I} \cdot (B-\lambda I))$

$= \det(B-\lambda I).$

Thm If A and B are similar (i.e. $A = P \cdot B \cdot P^{-1}$)

$$\phi_A(\lambda) = \det(A - \lambda I) = \det(B - \lambda I) = \phi_B(\lambda)$$

$$= \det(P \cdot B P^{-1} - \lambda I) = \det(P \cdot (B - \lambda I) P^{-1})$$

$I = P \cdot P^{-1}$

Similar Matrices

Definition

Two $n \times n$ matrices A and B are **similar** if there is a matrix P ($n \times n$) so that $A = PBP^{-1}$.

Theorem

If A and B similar, then they have the same characteristic polynomial.

If time permits, we will explain or prove this theorem in lecture. Note:

- Our textbook introduces similar matrices in Section 5.2, but doesn't have exercises on this concept until 5.3.
- Two matrices, A and B, do not need to be similar to have the same eigenvalues. For example,

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \text{ and } \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Not similar

eigenvalue is 0.

similar $\Rightarrow$ the same eigenvalues

the same eigenvalues $\not\Rightarrow$ similar.

Diagonal Matrices

A matrix is **diagonal** if the only non-zero elements, if any, are on the main diagonal.

The following are all diagonal matrices.

$$\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, \quad \begin{bmatrix} 2 \end{bmatrix}, \quad I_n, \quad \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

We'll only be working with diagonal square matrices in this course.

Powers of Diagonal Matrices

If A is diagonal, then A^k is easy to compute. For example,

$$A = \begin{pmatrix} 3 & 0 \\ 0 & 0.5 \end{pmatrix}$$

· diagonal

· square on diagonal

$$A^2 = \begin{pmatrix} 3 & 0 \\ 0 & 0.5 \end{pmatrix} \begin{pmatrix} 3 & 0 \\ 0 & 0.5 \end{pmatrix} = \begin{pmatrix} 3^2 & 0 \\ 0 & (0.5)^2 \end{pmatrix}$$

$$A^k = \begin{pmatrix} 3^k & 0 \\ 0 & (0.5)^k \end{pmatrix}$$

But what if A is not diagonal?

Diagonalization

Suppose $A \in \mathbb{R}^{n\times n}$. We say that A is **diagonalizable** if it is similar to a diagonal matrix, D. That is, we can write

$$A = PDP^{-1} \quad \text{for some invertible } P \in \mathbb{R}^{n\times n}$$

① Why A and D are similar? (D^k is easy)

$A^k = ?$

$A^2 = (P \cdot D \cdot \boxed{P^{-1}) \cdot (P} \cdot D \cdot P^{-1}) = P \cdot \boxed{D \cdot I \cdot D} \cdot P^{-1} = P \cdot D^2 \cdot P^{-1}$

$A^3 = P \cdot D^3 \cdot P^{-1}$

$\vdots$

$A^k = P \cdot D^k \cdot P^{-1}$

coefficient

$A \cdot \vec{x} =$ lin. combi. of Columns in A

③ Need to find P. How?

$A = P \cdot D \cdot P^{-1}$

$A \cdot P = P \cdot D$

$$A \cdot [\vec{v}_1 \; \vec{v}_2 \; \cdots \; \vec{v}_n] = [\vec{v}_1 \; \cdots \; \vec{v}_n] \begin{bmatrix} a_1 & & 0 \\ & a_2 & \\ 0 & & \ddots \; a_n \end{bmatrix}$$

$$= [a_1\vec{v}_1 \quad a_2\vec{v}_2 \quad \cdots \quad a_n\vec{v}_n]$$

$A \cdot \vec{v}_1 = a_1\vec{v}_1, \quad A\vec{v}_2 = a_2\vec{v}_2, \quad \cdots, \quad A\vec{v}_n = a_n\vec{v}_n$

Diagonalization

Theorem

If A is diagonalizable $\Leftrightarrow$ A has n linearly independent eigenvectors.

Note: the symbol $\Leftrightarrow$ means " if and only if ".

Also note that $A = PDP^{-1}$ if and only if

$$A = \overset{P}{\overset{\shortparallel}{[\vec{v}_1 \ \vec{v}_2 \cdots \vec{v}_n]}} \overset{=D}{\begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}} [\vec{v}_1 \ \vec{v}_2 \cdots \vec{v}_n]^{-1}$$

where $\vec{v}_1, \ldots, \vec{v}_n$ are linearly independent eigenvectors, and $\lambda_1, \ldots, \lambda_n$ are the corresponding eigenvalues (**in order**).

Example 1

Diagonalize if possible.

$$A = \begin{pmatrix} 2 & 6 \\ 0 & -1 \end{pmatrix}$$

① Eigenvalues : $\underbrace{\lambda = 2, -1}$ because A is upper triangular.

② Eigenvectors

(i) $\lambda = 2$ $\quad E_2 = \text{Null}(A - 2I)$

$$A - 2I = \begin{pmatrix} 0 & 6 \\ 0 & 3 \end{pmatrix} \longrightarrow \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{Solution}: c \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

(ii) $\lambda = -1$ $\quad E_{-1} = \text{Null}(A + I)$

$$A + I = \begin{pmatrix} 3 & 6 \\ 0 & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix} \quad \text{Solution}: c \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

$$\vec{v}_2 = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

③ $P = \begin{bmatrix} \vec{v}_1 & \vec{v}_2 \end{bmatrix} = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}$ invertible $\Rightarrow$ A is diagonalizable

$D = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}$ Check: $A = P \cdot D \cdot P^{-1}$

Example 2

Diagonalize if possible.

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

① Eigenvalue $\lambda = 3$

② $E_3 = \text{Null}(A - 3I) = \text{Null}\underbrace{\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}}_{\dim = 1 = \# \text{ of free var.}}$

the only eigenspace

$\Downarrow$

$P = [\vec{v}_1 \quad \vec{v}_2]$ Not invertible.

Not Diagonalizable.

Thm $\lambda_1, \lambda_2, \dots, \lambda_n$ all distinct eigenvalues

$\vec{v}_1, \vec{v}_2, \ \ , \vec{v}_n$ corresponding eigenvectors

$\Rightarrow \{\vec{v}_1, \dots, \vec{v}_n\}$ linearly indep.

Pf: $n=2$: WANT $\{\vec{v}_1, \vec{v}_2\}$ lin. indep.

$a_1\vec{v}_1 + a_2\vec{v}_2 = 0 \Rightarrow a_1 = a_2 = 0$

Suppose $a_1 \neq 0$.

$A(a_1 v_1 + a_2 v_2) = 0$

$a_1 \underline{Av_1} + a_2 \underline{Av_2} = 0$

$$\begin{cases} a_1 \lambda_1 v_1 + a_2 \lambda_2 v_2 = 0 \\ a_1 \lambda_2 v_1 + a_2 \lambda_2 v_2 = 0 \end{cases} \Rightarrow a_1 \underbrace{(\lambda_1 - \lambda_2)}_{\neq 0} \vec{v}_1 = 0$$

$\vec{v}_1 \neq 0$

$\emptyset$

$a_1 = 0$

Distinct Eigenvalues

> **Theorem**
>
> If A is $n \times n$ and has n distinct eigenvalues, then A is diagonalizable.

Why does this theorem hold?

Is it necessary for an $n \times n$ matrix to have n distinct eigenvalues for it to be diagonalizable?

10/16/23

$A \in \mathbb{R}^{n \times n}$ is <u>diagonalizable</u>

$$P = [\overbrace{v_1, v_2, \cdots, v_n}^{\text{eigenvectors}}] \qquad D = \begin{bmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ 0 & & \ddots & \\ & & & \lambda_n \end{bmatrix} \ \text{eigenvalues}$$

definition

$\Longleftrightarrow$ There is an invertible matrix P, a diagonal D such that

$$A = PDP^{-1}$$

$\Longleftrightarrow$ We can find n linearly indep. eigenvectors.

$\Uparrow$

If n distinct eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_n$

$\{v_1, \cdots, v_n\}$: linearly indep.

Non-Distinct Eigenvalues

Theorem. Suppose

- A is $n \times n$
- A has distinct eigenvalues $\lambda_1, \ldots, \lambda_k$, $k \le n$
- a_i = algebraic multiplicity of λ_i
- d_i = dimension of λ_i eigenspace ("geometric multiplicity")

$\overbrace{a_1 \quad \cdots \quad a_k}^{\text{alg. multi.}}$

$\Rightarrow$ Char. Poly. $= \phi_A(\lambda)$

$= (-1)^n (\lambda-\lambda_1)^{a_1}(\lambda-\lambda_2)^{a_2}\cdots(\lambda-\lambda_k)^{a_k}$

$a_1 + \cdots + a_k = n$

$E_{\lambda_i} = \dim(\underbrace{\text{Null}(A-\lambda_i I)}))$

Then

1. $d_i \le a_i$ for all i
2. A is diagonalizable $\Leftrightarrow \Sigma d_i = n \Leftrightarrow d_i = a_i$ for all i
3. A is diagonalizable $\Leftrightarrow$ the eigenvectors, for all eigenvalues, together form a basis for $\mathbb{R}^n$.

$d_1 + d_2 + \cdots + d_k$ = max # of lin. indep. eigenvectors.

- A is diagonalizable $\Leftrightarrow$ $\underset{\le a_1}{d_1} + \cdots + \underset{\le a_k}{d_k} = n$

$\Leftrightarrow$ $a_1 = d_1$, $a_2 = d_2$, $\cdots$, $a_k = d_k$

$\Leftrightarrow$ n lin. indep. eigenvectors.

$\Leftrightarrow$ eigenvectors form a basis for $\mathbb{R}^n$.

Example 3

The eigenvalues of A are $\lambda = 3, 1$. If possible, construct P and D such that $AP = PD$.

$$A = \begin{pmatrix} 7 & 4 & 16 \\ 2 & 5 & 8 \\ -2 & -2 & -5 \end{pmatrix} \qquad D = \begin{bmatrix} 3 & & \\ & 3 & \\ & & 1 \end{bmatrix}$$

① $\lambda = 3$: $\quad E_3 = \text{Null}(A - 3I)$

$$A - 3I = \begin{bmatrix} 4 & 4 & 16 \\ 2 & 2 & 8 \\ -2 & -2 & -8 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 1 & 4 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{RREF.}$$

$x_1 + x_2 + 4x_3 = 0$

$x_1 = -x_2 - 4x_3$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -x_2 - 4x_3 \\ x_2 \\ x_3 \end{bmatrix} = x_2 \underbrace{\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}}_{\vec{v}_1} + x_3 \underbrace{\begin{bmatrix} -4 \\ 0 \\ 1 \end{bmatrix}}_{\vec{v}_2}$$

free var. $\Rightarrow$ geo. multi $= 2$

alg. multi $= 2$

② $\lambda = 1$: $\quad 1 \le$ geo. multi. $\le$ alg. multi. $= 1 \quad \Rightarrow$ A is diagonalizable

$E_1 = \text{Null}(A - I)$

$$A - I = \begin{bmatrix} 6 & 4 & 16 \\ 2 & 4 & 8 \\ -2 & -2 & -6 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 1 & 3 \\ 1 & 2 & 4 \\ 3 & 2 & 8 \end{bmatrix}$$

$$\longrightarrow \begin{bmatrix} 1 & 1 & 3 \\ 0 & 1 & 1 \\ \cancel{0} & \cancel{-1} & \cancel{-1} \end{bmatrix}$$

$$\begin{matrix} 0 & 0 & 0 \end{matrix}$$

$$\longrightarrow \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \quad \underline{\underline{\text{RREF}}}$$

$$\begin{cases} x_1 + 2x_3 = 0 \\ x_2 + x_3 = 0 \end{cases} \qquad \begin{aligned} x_1 &= -2x_3 \\ x_2 &= -x_3 \end{aligned}$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -2x_3 \\ -x_3 \\ x_3 \end{bmatrix} = x_3 \begin{bmatrix} -2 \\ -1 \\ 1 \end{bmatrix} \coloneqq v_3$$

$$P = \begin{bmatrix} -1 & -4 & -2 \\ 1 & 0 & -1 \\ 0 & 1 & 1 \end{bmatrix} \qquad D = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Recall $\mathcal{B} = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ for $\mathbb{R}^n$.

$\vec{x} \in \mathbb{R}^n$, $[\vec{x}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} \Leftrightarrow \vec{x} = c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_n\vec{v}_n$

Basis of Eigenvectors

Express the vector $\vec{x}_0 = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$ as a linear combination of the vectors $\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\vec{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ and find the coordinates of $\vec{x}_0$ in the basis $\mathcal{B} = \{\vec{v}_1, \vec{v}_0\}$.

P

$\vec{x}_0 = c_1\vec{v}_1 + c_2\vec{v}_2$

$\begin{bmatrix} 4 \\ 5 \end{bmatrix} = c_1\begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_2\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$

$\left[\begin{array}{cc|c} 1 & 1 & 4 \\ 1 & -1 & 5 \end{array}\right] \to \cdots$

$c_1 = 4.5 = \frac{9}{2}$ $c_2 = -0.5 = -\frac{1}{2}$

$[\vec{x}_0]_{\mathcal{B}} = \begin{bmatrix} 9/2 \\ -1/2 \end{bmatrix}$

Let $P = [\vec{v}_1\ \vec{v}_2]$ and $D = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, and find $[A^k\vec{x}_0]_{\mathcal{B}}$ where $A = PDP^{-1}$, for $k = 1, 2, \ldots$.

$\Rightarrow$ eigenvalues $=$ $\begin{array}{cc} v_1 & v_2 \\ 1 & -1 \end{array}$

$[A^k\vec{x}_0]_{\mathcal{B}} =$??

$A^k\vec{x}_0 = A^k\left(\frac{9}{2}\cdot v_1 - \frac{1}{2}v_2\right)$

$= \frac{9}{2}A^k v_1 - \frac{1}{2}A^k v_2$

$= \frac{9}{2}\cdot 1^k \cdot v_1 - \frac{1}{2}\cdot(-1)^k \cdot v_2$

$[A^k\vec{x}_0]_{\mathcal{B}} = \begin{bmatrix} \frac{9}{2}\cdot(1)^k \\ -\frac{1}{2}\cdot(-1)^k \end{bmatrix}$

$$[\vec{x}]_B = P^{-1} \cdot \vec{x}$$

$$[A^k \vec{x}_0]_B = P^{-1} \cdot A^k \vec{x}_0$$

$$= \cancel{P^{-1}} \cdot \cancel{P} \cdot D^k \cdot P^{-1} \cdot \vec{x}_0$$

$$= D^k \cdot \underbrace{\left(P^{-1} \cdot \vec{x}_0 \right)}$$

$$= D^k \cdot [x]_B$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}^k \cdot \begin{bmatrix} \frac{9}{2} \\ -\frac{1}{2} \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & (-1)^k \end{bmatrix} \begin{bmatrix} \frac{9}{2} \\ -\frac{1}{2} \end{bmatrix}$$

Basis of Eigenvectors - part 2

Let $\vec{x}_0 = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$, $\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\vec{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ as before.

Again define $P = [\vec{v}_1 \ \vec{v}_2]$ but this time let $D = \begin{bmatrix} 1 & 0 \\ 0 & -1/2 \end{bmatrix}$, and now find $[A^k \vec{x}_0]_{\mathcal{B}}$ where $A = PDP^{-1}$, for $k = 1, 2, \ldots$.

$[A^k \vec{x}_0]_{\mathcal{B}} =$

Basis of Eigenvectors - part 3

Let $\vec{x}_0 = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$, $\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\vec{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ as before.

Again define $P = [\vec{v}_1 \ \vec{v}_2]$ but this time let $D = \begin{bmatrix} 2 & 0 \\ 0 & 3/2 \end{bmatrix}$, and now find $[A^k \vec{x}_0]_{\mathcal{B}}$ where $A = PDP^{-1}$, for $k = 1, 2, \ldots$.

$[A^k \vec{x}_0]_{\mathcal{B}} =$

Additional Example (if time permits)

Note that

$$\vec{x}_k = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \vec{x}_{k-1}, \quad \vec{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad k = 1, 2, 3, \ldots$$

generates a well-known sequence of numbers.

Use a diagonalization to find a matrix equation that gives the n^{th} number in this sequence.

Chapter 5 : Eigenvalues and Eigenvectors

5.5 : Complex Eigenvalues

Topics and Objectives

Topics

1. Complex numbers: addition, multiplication, complex conjugate
2. Complex eigenvalues and eigenvectors.
3. Eigenvalue theorems

Learning Objectives

1. Use eigenvalues to determine identify the rotation and dilation of a linear transform.
2. Rotation dilation matrices.
3. Find complex eigenvalues and eigenvectors of a real matrix.
4. Apply theorems to characterize matrices with complex eigenvalues.

Motivating Question
What are the eigenvalues of a rotation matrix?

Imaginary Numbers

Recall: When calculating roots of polynomials, we can encounter square roots of negative numbers. For example:

$$x^2 + 1 = 0$$

$x^2 = -1$

The roots of this equation are:

$x = \pm\sqrt{-1}$

We usually write $\sqrt{-1}$ as i (for "imaginary").

Addition and Multiplication

The imaginary (or complex) numbers are denoted by $\mathbb{C}$, where

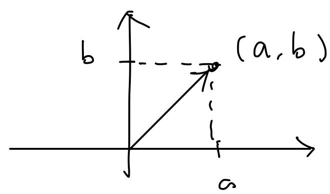
$$\mathbb{C} = \{a + bi \mid a, b \text{ in } \mathbb{R}\}$$

$i = \sqrt{-1}$

$i^2 = -1$.

We can identify $\mathbb{C}$ with $\mathbb{R}^2$: $\quad a + bi \leftrightarrow (a, b)$

$a+bi \longmapsto$



Same as vector addition.

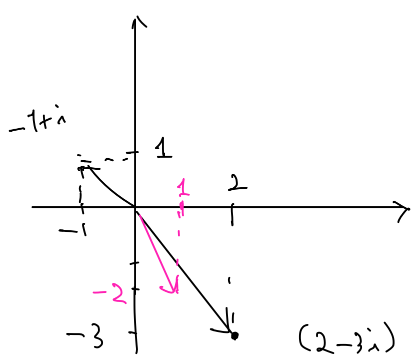
We can add and multiply complex numbers as follows:

$$(2 - 3i) + (-1 + i) = (2 + (-1)) + (-3 + 1)\,i = 1 - 2i$$

$$(2 - 3i)(-1 + i) = 2\cdot(-1) + 2\cdot i + (-3i)\cdot(-1) + (-3i)\cdot i$$

$$= -2 + 2i + 3i + 3 = \boxed{1 + 5i}$$

$-3i^2$, $i^2 = -1$



Complex Conjugate, Absolute Value, Polar Form

We can **conjugate** complex numbers: $\overline{a+bi} =$ $a - bi$

Ex $\overline{(5-2i)} = 5+2i$

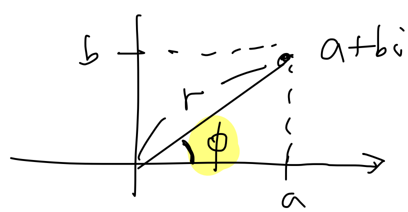
The **absolute value** of a complex number: $|a+bi| = \sqrt{(a+bi)\cdot\overline{(a+bi)}} = \sqrt{a^2+b^2}$

$(a+bi)\cdot\overline{(a+bi)} = (a+bi)\cdot(a-bi)$

$= a^2 - (bi)^2 = a^2 - b^2\cdot i^2 = a^2+b^2$

We can write complex numbers in **polar form**: $a+ib = r(\cos\phi + i\,\sin\phi)$

$z = a+bi \longmapsto$



$r^2 = a^2+b^2$

$r = \sqrt{a^2+b^2} = |a+bi|$

Note For complex z, $z\cdot\bar{z} \geq 0$

$$\begin{cases} r\cdot\cos\phi = a \\ r\cdot\sin\phi = b \end{cases}$$

$z = a+bi = r\cos\phi + r\cdot\sin\phi\, i$

$= r\cdot(\cos\phi + i\sin\phi)$

$$\vec{\bar{x}} = \overline{\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}} = \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \vdots \\ \bar{x}_n \end{bmatrix}$$

In general,

$$\overline{A \cdot \vec{v}} = \overline{A} \cdot \overline{\vec{v}} \underset{\uparrow}{=} A \cdot \overline{\vec{v}}$$

A is real

Complex Conjugate Properties

If x and y are complex numbers, $\vec{v} \in \mathbb{C}^n$, it can be shown that:

- $\overline{(x+y)} = \overline{x} + \overline{y}$
- $\overline{A\vec{v}} = A\overline{\vec{v}}$ ← suppose A is a real matrix (Every entry is real)
- $\text{Im}(x\overline{x}) = 0.$

Example True or false: if x and y are complex numbers, then

$x = a + bi$

$y = c + di$

$\overline{(xy)} = \overline{x}\ \overline{y}$ Yes

$\overline{x \cdot y} = \overline{(ac - bd) + (ad + bc)i}$

$\overline{x} \cdot \overline{y} = (ac - bd) - (ad + bc)i$

Im : Imaginary Part

Re : Real Part

$\text{Im}(5 + 2i) = 2$

$\text{Re}(3 + 4i) = 3$

$x = a + bi$

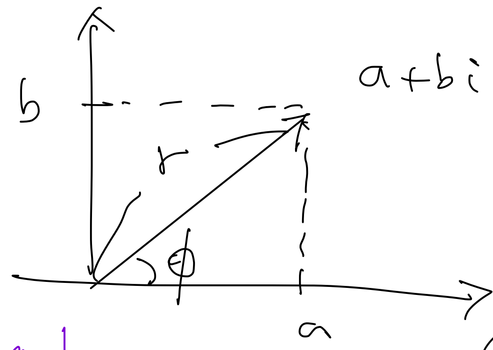
$\text{Im}(x \cdot \overline{x}) = \text{Im}(\underbrace{a^2 + b^2}_{\text{has no imaginary part}}) = 0.$

$$\underset{\sim}{z} = a + bi \Rightarrow$$

$$= r \cdot (\cos\phi + i \sin\phi)$$

$$= r \cdot e^{i\phi}$$

Euler's Formula.



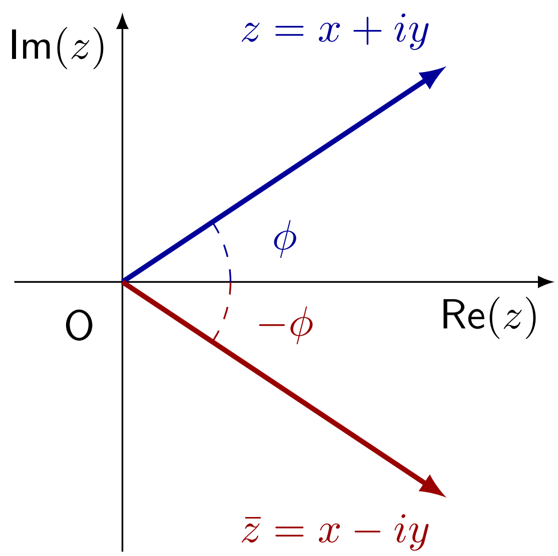
$a+bi$

$$r = |z| = \sqrt{a^2+b^2} = \sqrt{z \cdot \bar{z}}$$

$$\begin{cases} r\cos\phi = a \\ r\sin\phi = b \end{cases}$$

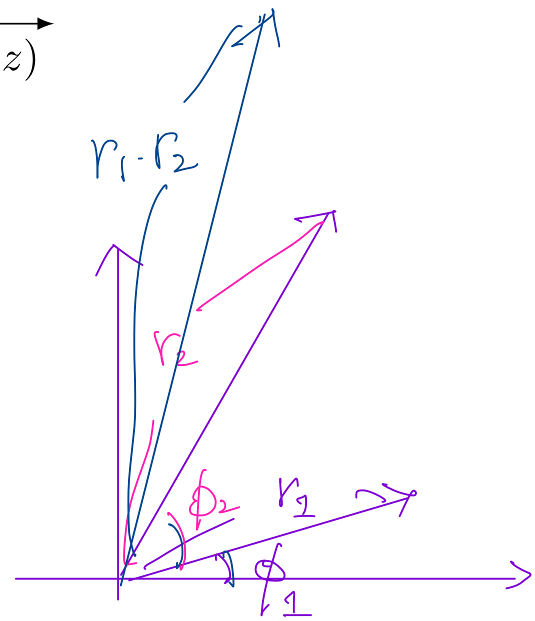
Polar Form and the Complex Conjugate

Conjugation reflects points across the real axis.

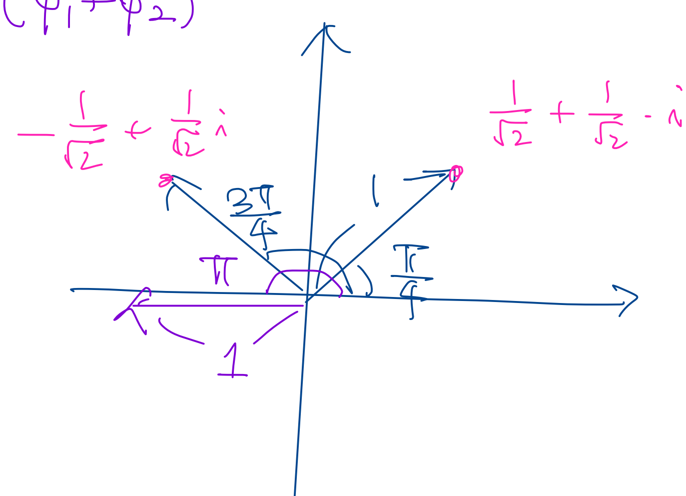


$$z = r_1 \cdot e^{i\phi_1}$$

$$w = r_2 \cdot e^{i\phi_2}$$

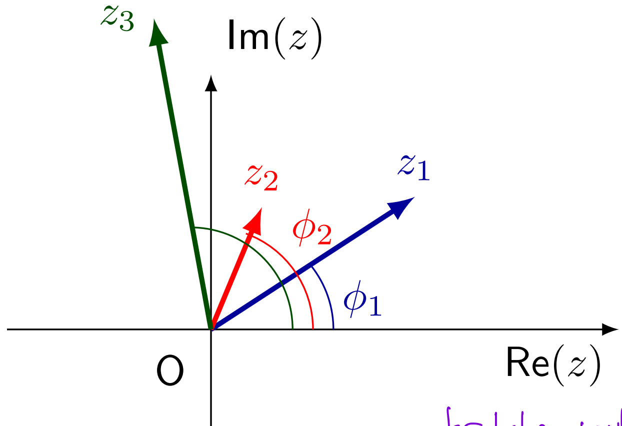


$$z \cdot w = r_1 \cdot r_2 \cdot e^{i(\phi_1+\phi_2)}$$



Euler's Formula

Suppose z_1 has angle ϕ_1, and z_2 has angle ϕ_2.



absolute values of z, w.
"

The product $z_1 z_2$ has angle $\phi_1 + \phi_2$ and modulus $|z|\,|w|$. Easy to remember using Euler's formula.

$$\boxed{z = |z|\,\mathrm{e}^{i\phi}}$$

The product $z_1 z_2$ is:

$$z_3 = z_1 z_2 = (|z_1|\,\mathrm{e}^{i\phi_1})(|z_2|e^{i\phi_2}) = |z_1|\,|z_2|\,\mathrm{e}^{i(\phi_1+\phi_2)}$$

Complex Numbers and Polynomials

$a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 = 0$

Theorem: Fundamental Theorem of Algebra

Every polynomial of degree n has exactly n complex roots, counting multiplicity.

$r_1, \ldots, r_n \in \mathbb{C}$.

$\Rightarrow a_n(x - r_1)(x - r_2) \cdots (x - r_n) = 0$

coefficients $a_n, \ldots, a_0 \in \mathbb{R}$

Theorem

1. If $\lambda \in \mathbb{C}$ is a root of a real polynomial $p(x)$, then the conjugate $\overline{\lambda}$ is also a root of $p(x)$.
2. If λ is an eigenvalue of real matrix A with eigenvector $\vec{v}$, then $\overline{\lambda}$ is an eigenvalue of A with eigenvector $\overline{\vec{v}}$.

Ex: real poly. $p(x) = 0$

one root is $2+i$

$\Rightarrow \overline{2+i} = 2-i$ is also a root.

Example

Four of the eigenvalues of a real 7×7 matrix are $-2, 4+i, -4-i$, and i.
What are the other eigenvalues?

$4-i \quad -4+i \quad -i$

7 distinct eigenvalues

$\Downarrow$

A is diagonalizable

Example

The matrix that rotates vectors by $\phi = \pi/4$ radians about the origin, and then scales (or dilates) vectors by $r = \sqrt{2}$, is

$$A = \begin{bmatrix} r & 0 \\ 0 & r \end{bmatrix} \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

What are the eigenvalues of A? Find an eigenvector for each eigenvalue.

Char. Poly. $= \lambda^2 - (1+1)\lambda + (1\cdot 1 - (-1)\cdot 1)$

$= \lambda^2 - 2\lambda + 2 = 0$

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

$\lambda - 1 = \pm i$

$\lambda = 1 \pm i$

(1) $\lambda_1 = 1 + i$: Null $(A - (1+i)I)$

$A - (1+i)I = \begin{bmatrix} 1-(1+i) & -1 \\ 1 & 1-(1+i) \end{bmatrix} = \begin{bmatrix} -i & -1 \\ 1 & -i \end{bmatrix}$

$\longrightarrow \begin{bmatrix} 1 & -i \\ 0 & 0 \end{bmatrix} \Longrightarrow \begin{bmatrix} i \\ 1 \end{bmatrix} = \vec{v}_1$

② $\lambda_2 = 1 - i$ $\vec{v}_2 = \overline{\begin{bmatrix} i \\ 1 \end{bmatrix}} = \begin{bmatrix} -i \\ 1 \end{bmatrix}$

Example

The matrix in the previous example is a special case of this matrix:

$$C = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

Calculate the eigenvalues of C and express them in polar form.

10/20/23

$A \in \mathbb{R}^{n\times n}$ $\quad \lambda \in \mathbb{C}$ Eigenvalue , $\vec{v} \in \mathbb{C}^n$ Eigenvector

$\Rightarrow$ $\overline{\lambda}$ Eigenvalue $\quad \overline{\vec{v}}$ Eigenvector.

Example

Find the complex eigenvalues and an associated complex eigenvector for each eigenvalue for the matrix.

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

① Char. Eqn : $\lambda^2 - (1+3)\lambda + (1\cdot 3 - (-2)\cdot 1) = 0$

$$\lambda^2 - 4\lambda + 5 = 0$$

$$(\lambda - 2)^2 = \underbrace{\lambda^2 - 4\lambda + 4}_{(\ \ )^2} = -1$$

$\lambda - 2 = \pm i$ $\quad \therefore \lambda = 2+i$, $2-i$

② Eigenvectors : $\lambda_1 = 2+i$

$$A - (2+i)I = \begin{bmatrix} 1-(2+i) & -2 \\ 1 & 3-(2+i) \end{bmatrix} = \begin{bmatrix} -1-i & -2 \\ 1 & 1-i \end{bmatrix}$$

$$\longrightarrow \begin{bmatrix} 1 & 1-i \\ 0 & 0 \end{bmatrix} \qquad \vec{v_1} = \begin{bmatrix} i-1 \\ 1 \end{bmatrix}$$

For $\lambda_2 = 2-i = \overline{\lambda_1}$, $\vec{v_2} = \overline{\vec{v_1}} = \begin{bmatrix} -i-1 \\ 1 \end{bmatrix}$

"Angle"

Section 6.1 : Inner Product, Length, and Orthogonality

Chapter 6: Orthogonality and Least Squares

Math 1554 Linear Algebra

Topics and Objectives

Topics

1. Dot product of vectors
2. Magnitude of vectors, and distances in $\mathbb{R}^n$
3. Orthogonal vectors and complements
4. Angles between vectors

Learning Objectives

1. Compute (a) dot product of two vectors, (b) length (or magnitude) of a vector, (c) distance between two points in $\mathbb{R}^n$, and (d) angles between vectors.
2. Apply theorems related to orthogonal complements, and their relationships to Row and Null space, to characterize vectors and linear systems.

Motivating Question

For a matrix A, which vectors are orthogonal to all the rows of A? To the columns of A?

Inner Product

The Dot Product

(Vector) · (Vector) = number.

The dot product between two vectors, $\vec{u}$ and $\vec{v}$ in $\mathbb{R}^n$, is defined as

n×1 n×1

$$\vec{u} \cdot \vec{v} = \vec{u}^T \vec{v} = \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \underbrace{u_1 v_1 + u_2 v_2 + \cdots + u_n v_n}. \in \mathbb{R}$$

1×n n×1

Example 1: For what values of k is $\vec{u} \cdot \vec{v} = 0$?

$$\vec{u} = \begin{pmatrix} -1 \\ 3 \\ k \\ 2 \end{pmatrix}, \qquad \vec{v} = \begin{pmatrix} 4 \\ 2 \\ 1 \\ -3 \end{pmatrix}$$

$$\vec{u} \cdot \vec{v} = u^T \cdot v = \begin{bmatrix} -1 & 3 & k & 2 \end{bmatrix} \begin{bmatrix} 4 \\ 2 \\ 1 \\ -3 \end{bmatrix}$$

$$= (-1)\cdot 4 + 3 \cdot 2 + k \cdot 1 + 2 \cdot (-3)$$

$$= -4 \cancel{+6} + k \cancel{-6} = k - 4. \quad = 0$$

$$\Rightarrow k = 4.$$

Properties of the Dot Product

The dot product is a special form of matrix multiplication, so it inherits linear properties.

Theorem (Basic Identities of Dot Product)

Let $\vec{u}, \vec{v}, \vec{w}$ be three vectors in $\mathbb{R}^n$, and $c \in \mathbb{R}$.

1. (Symmetry) $\vec{u} \cdot \vec{w} =$ $\vec{w} \cdot \vec{u}$
2. (Linear in each vector) $(\vec{v} + \vec{w}) \cdot \vec{u} =$ $\vec{v} \cdot \vec{u} + \vec{w} \cdot \vec{u}$
3. (Scalars) $(c\vec{u}) \cdot \vec{w} =$ $\vec{u} \cdot (c\vec{w}) = c \cdot (\vec{u} \cdot \vec{w})$
4. (Positivity) $\vec{u} \cdot \vec{u} \geq 0$, and the dot product equals __________

$$\vec{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \qquad \vec{u} \cdot \vec{u} = [u_1 \cdots u_n] \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}$$

$$= u_1^2 + u_2^2 + \cdots + u_n^2 \geq 0$$

if $u_1, \cdots, u_n \geq 0$

<u>Note</u> For complex vector

$$\overline{\vec{v}} \cdot \vec{v} \geq 0$$

The Length of a Vector

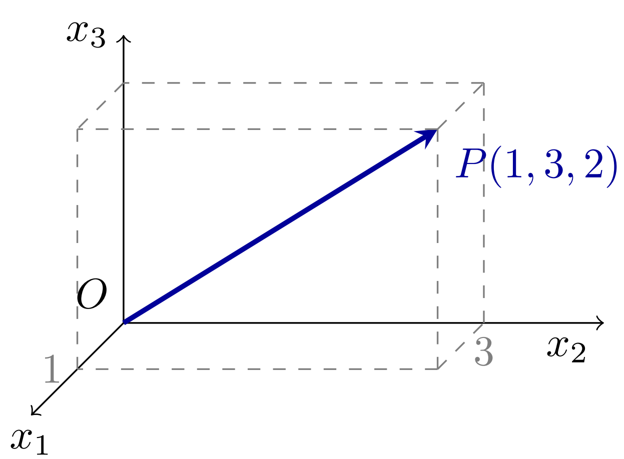
Definition

The **length** of a vector $\vec{u} \in \mathbb{R}^n$ is

$$\|\vec{u}\| = \sqrt{\vec{u} \cdot \vec{u}} = \sqrt{u_1^2 + u_2^2 + \cdots + u_n^2}$$

Example: the length of the vector $\overrightarrow{OP}$ is

$$\sqrt{1^2 + 3^2 + 2^2} = \sqrt{14}$$



Example

Let $\vec{u}, \vec{v}$ be two vectors in $\mathbb{R}^n$ with $\|\vec{u}\| = 5$, $\|\vec{v}\| = \sqrt{3}$, and $\vec{u} \cdot \vec{v} = -1$. Compute the value of $\|\vec{u} + \vec{v}\|$.

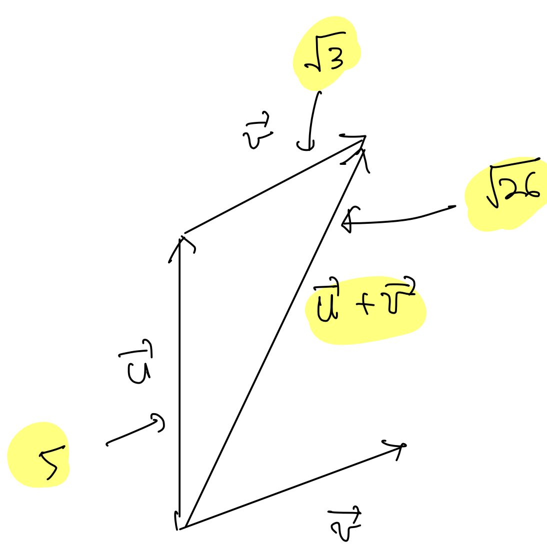
$$\|\vec{u}+\vec{v}\|^2 = \left(\sqrt{(\vec{u}+\vec{v})\cdot(\vec{u}+\vec{v})}\right)^2 = (\vec{u}+\vec{v})\cdot(\vec{u}+\vec{v})$$

$$= \underbrace{\vec{u}\cdot\vec{u}}_{=\|\vec{u}\|^2 = 25} + \underbrace{\vec{u}\cdot\vec{v}}_{=-1} + \underbrace{\vec{v}\cdot\vec{u}}_{=-1} + \underbrace{\vec{v}\cdot\vec{v}}_{=\|\vec{v}\|^2 = (\sqrt{3})^2 = 3}$$

determines angle

$$= 25 - 1 - 1 + 3 = 26$$

$$\|\vec{u}+\vec{v}\| = \sqrt{26}$$



Length of Vectors and Unit Vectors

Note: for any vector $\vec{v}$ and scalar c, the length of $c\vec{v}$ is

$$||c\vec{v}|| = |c|\,||\vec{v}||$$

Definition

If $\vec{v} \in \mathbb{R}^n$ has length one, we say that it is a **unit vector.**

For example, each of the following vectors are unit vectors.

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \vec{y} = \frac{1}{\sqrt{5}}\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad \vec{v} = \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}$$

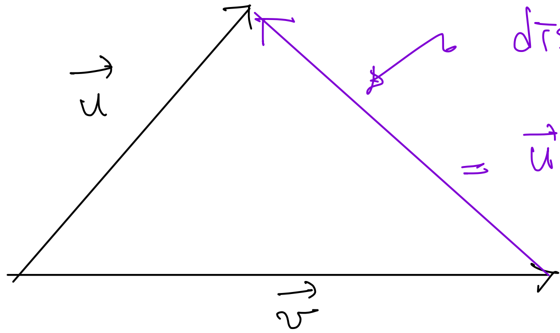
$$\vec{v} = \begin{bmatrix} 1 \\ 3 \end{bmatrix} \qquad ||\vec{v}|| = \sqrt{1^2 + 3^2} = \sqrt{10}$$

$$\frac{1}{\sqrt{10}} ||\vec{v}|| = 1$$

$$\left|\left| \frac{1}{\sqrt{10}} \vec{v} \right|\right| = 1$$

normalization

$$\frac{1}{\sqrt{10}}\vec{v} = \begin{bmatrix} \frac{1}{\sqrt{10}} \\ \frac{3}{\sqrt{10}} \end{bmatrix} \leftarrow \text{unit vector.}$$



distance $= \|\vec{u} - \vec{v}\|$

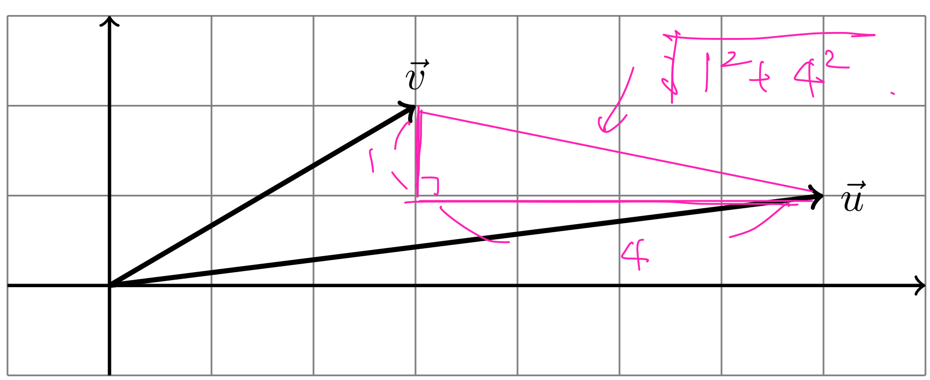
Distance in $\mathbb{R}^n$

Definition

For $\vec{u}, \vec{v} \in \mathbb{R}^n$, the **distance** between $\vec{u}$ and $\vec{v}$ is given by the formula

$$\boxed{\|\vec{u} - \vec{v}\|}$$

Example: Compute the distance from $\vec{u} = \begin{pmatrix} 7 \\ 1 \end{pmatrix}$ and $\vec{v} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$.



$$\text{distance} = \|\vec{u} - \vec{v}\| = \left\| \begin{bmatrix} 7 \\ 1 \end{bmatrix} - \begin{bmatrix} 3 \\ 2 \end{bmatrix} \right\|$$

$$= \left\| \begin{bmatrix} 4 \\ -1 \end{bmatrix} \right\| = \sqrt{4^2 + (-1)^2} = \sqrt{17}$$

$n = 2$

$\vec{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$

$$|\vec{u} \cdot \vec{v}| = |u_1 v_1 + u_2 v_2|$$
$$\leq \sqrt{u_1^2 + u_2^2} \cdot \sqrt{v_1^2 + v_2^2}.$$
$$= \|\vec{u}\| \cdot \|\vec{v}\|$$

The Cauchy-Schwarz Inequality

Theorem: Cauchy-Bunyakovsky–Schwarz Inequality

For all $\vec{u}$ and $\vec{v}$ in $\mathbb{R}^n$,

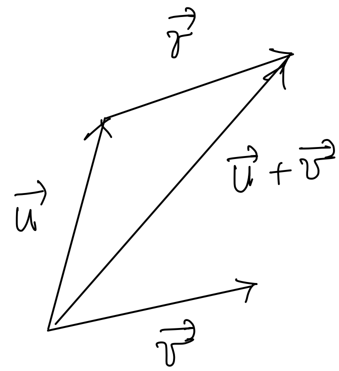
$$|\vec{u} \cdot \vec{v}| \leq \|\vec{u}\|\|\vec{v}\|.$$

Equality holds *if and only if* $\vec{v} = \alpha\vec{u}$ for $\alpha = \dfrac{\vec{u} \cdot \vec{v}}{\vec{u} \cdot \vec{u}}$.

Proof: Assume $\vec{u} \neq 0$, otherwise there is nothing to prove.

Set $\alpha = \dfrac{\vec{u} \cdot \vec{v}}{\vec{u} \cdot \vec{u}}$. Observe that $\vec{u} \cdot (\alpha\vec{u} - \vec{v}) = 0$. So

$$\begin{aligned} 0 \leq \|\alpha\vec{u} - \vec{v}\|^2 &= (\alpha\vec{u} - \vec{v}) \cdot (\alpha\vec{u} - \vec{v}) \\ &= \alpha\vec{u} \cdot (\alpha\vec{u} - \vec{v}) - \vec{v} \cdot (\alpha\vec{u} - \vec{v}) \\ &= -\vec{v} \cdot (\alpha\vec{u} - \vec{v}) \\ &= \frac{\|\vec{u}\|^2\|\vec{v}\|^2 - |\vec{u} \cdot \vec{v}|^2}{\|\vec{u}\|^2} \end{aligned}$$



The Triangle Inequality

Theorem: Triangle Inequality

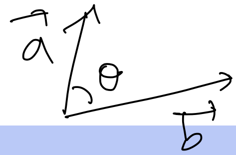
For all $\vec{u}$ and $\vec{v}$ in $\mathbb{R}^n$,

$$\|\vec{u}+\vec{v}\| \leq \|\vec{u}\| + \|\vec{v}\|.$$

Proof:

$$\begin{aligned}
\|\vec{u}+\vec{v}\|^2 &= (\vec{u}+\vec{v})\cdot(\vec{u}+\vec{v}) \\
&= \|\vec{u}\|^2 + \|\vec{v}\|^2 + 2\vec{u}\cdot\vec{v} \\
&\leq \|\vec{u}\|^2 + \|\vec{v}\|^2 + 2\|\vec{u}\|\|\vec{v}\| \\
&= (\|\vec{u}\| + \|\vec{v}\|)^2
\end{aligned}$$

Cauchy – Schwartz.



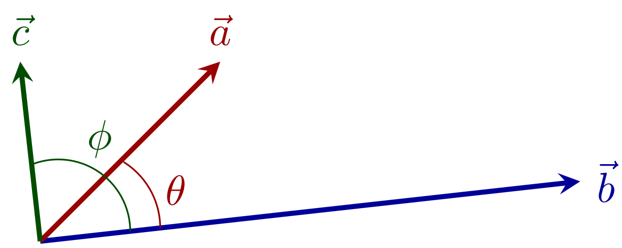
Angles

Theorem

$\vec{a} \cdot \vec{b} = \|\vec{a}\| \|\vec{b}\| \cos\theta$. Thus, if $\vec{a} \cdot \vec{b} = 0$, then:

- $\vec{a}$ and/or $\vec{b}$ are zero vectors, or
- $\vec{a}$ and $\vec{b}$ are perpendicular. $\Leftarrow \cos\theta = 0 \Leftarrow \theta = \frac{\pi}{2}, \frac{3\pi}{2}, \ldots$

For example, consider the vectors below.



$$\cos\theta = \frac{\vec{a} \cdot \vec{b}}{\|\vec{a}\| \cdot \|\vec{b}\|}$$

$$= \left(\frac{\vec{a}}{\|\vec{a}\|}\right) \cdot \left(\frac{\vec{b}}{\|\vec{b}\|}\right)$$

normalized vector of $\vec{a}$

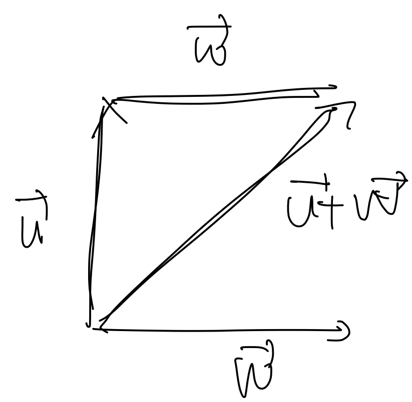
Orthogonality

Definition (Orthogonal Vectors)

Two vectors $\vec{u}$ and $\vec{w}$ are **orthogonal** if $\vec{u} \cdot \vec{w} = 0$. This is equivalent to:

$$\|\vec{u} + \vec{w}\|^2 = \|u\|^2 + \underbrace{2 \cdot \vec{u} \cdot \vec{w}}_{=0} + \|\vec{w}\|^2 = \|u\|^2 + \|\vec{w}\|^2$$

Note: The zero vector in $\mathbb{R}^n$ is orthogonal to every vector in $\mathbb{R}^n$. But we usually only mean non-zero vectors.



10/23/23

$\vec{u}, \vec{v} \in \mathbb{R}^n$

$\vec{u} \cdot \vec{v} = u^T \cdot v = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n$

$\|\vec{u}\| = \sqrt{\vec{u} \cdot \vec{u}}$

Distance between $\vec{u}, \vec{v}$ = $\|\vec{u} - \vec{v}\|$

C-S : $|\vec{u} \cdot \vec{v}| \leq \|\vec{u}\| \cdot \|\vec{v}\|$

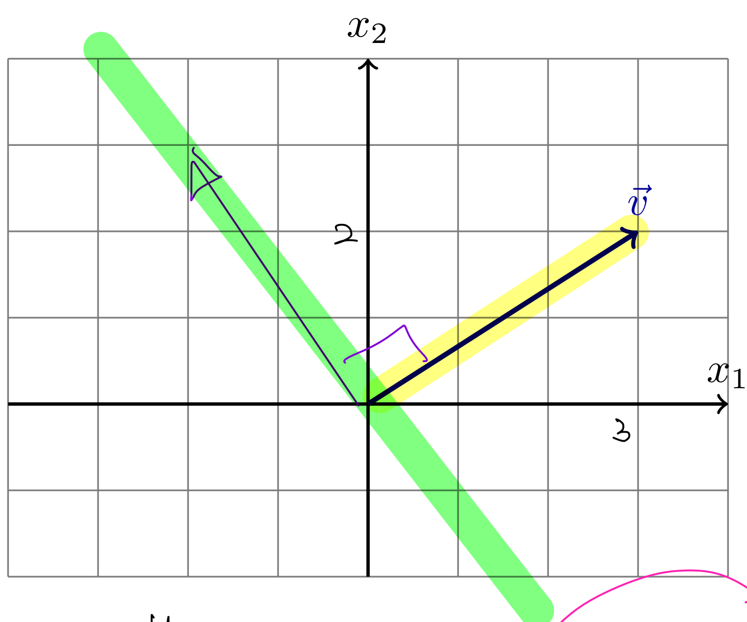
Triangle : $\|\vec{u} + \vec{v}\| \leq \|\vec{u}\| + \|\vec{v}\|$

Angle : $\cos\theta = \dfrac{\vec{u} \cdot \vec{v}}{\|\vec{u}\| \, \|\vec{v}\|}$

$\vec{u}, \vec{v}$ * orthogonal if $\vec{u} \cdot \vec{v} = 0$

Example

Sketch the subspace spanned by the set of all vectors $\vec{u}$ that are orthogonal to $\vec{v} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$.



$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ 2 \end{bmatrix} = 3u_1 + 2u_2 = 0$$

$$\vec{v} \cdot \vec{u} = v^T \vec{u}$$

$$\left\{ \vec{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \in \mathbb{R}^2 : \vec{u} \cdot \vec{v} = 0 \right\}$$

$$= \left\{ \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} : 3u_1 + 2u_2 = 0 \right\} \leftarrow ?$$

$$= \left\{ \begin{bmatrix} -2 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ -3 \end{bmatrix}, \ldots \right\}$$

$$= \left\{ c \begin{bmatrix} -2 \\ 3 \end{bmatrix} : c \in \mathbb{R} \right\} = \text{Null}\left(\underbrace{\begin{bmatrix} 3 & 2 \end{bmatrix}}_{} \right), \quad v^T = \begin{bmatrix} 3 & 2 \end{bmatrix}$$

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

Orthogonal Compliments

Definitions

Let W be a subspace of $\mathbb{R}^n$. Vector $\vec{z} \in \mathbb{R}^n$ is **orthogonal** to W if $\vec{z}$ is orthogonal to every vector in W.

The set of all vectors orthogonal to W is a subspace, the **orthogonal compliment** of W, or $W^\perp$ or 'W perp.'

$$W^\perp = \{\vec{z} \in \mathbb{R}^n \ : \ \vec{z} \cdot \vec{w} = 0 \text{ for all } \vec{w} \in W\}$$

Precious Example.

$$W = \text{Span}\left\{ \begin{bmatrix} 3 \\ 2 \end{bmatrix} \leftarrow \vec{v} \right\}$$

$$W^\perp = \text{Span}\left\{ \begin{bmatrix} -2 \\ 3 \end{bmatrix} \right\}$$

$$= \text{Null}\left(\begin{bmatrix} 3 & 2 \end{bmatrix} \right) = \text{Null}\left(v^T \right)$$

In general, W has a basis $\mathcal{B}$

$$\mathcal{B} = \{ \vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k \}$$

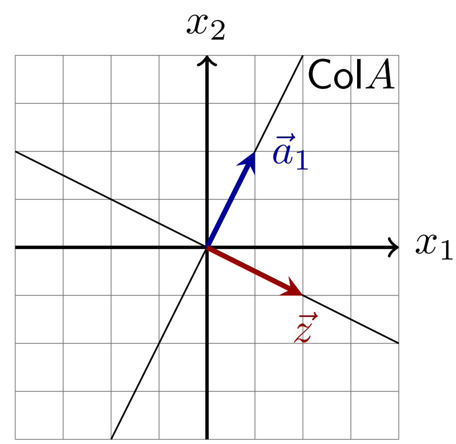
$$W^\perp = \{ \vec{z} : \vec{z} \cdot \vec{w} = 0 \text{ for all } \vec{w} \in W \}$$

$$= \{ \vec{z} : \vec{z} \cdot \vec{v}_1 = 0, \ \vec{z} \cdot \vec{v}_2 = 0, \ldots, \vec{z} \cdot \vec{v}_k = 0 \}$$

$$= \text{Null}\left(\begin{bmatrix} - v_1^T - \\ - v_2^T - \\ \vdots \\ - v_k^T - \end{bmatrix} \right)$$

Example

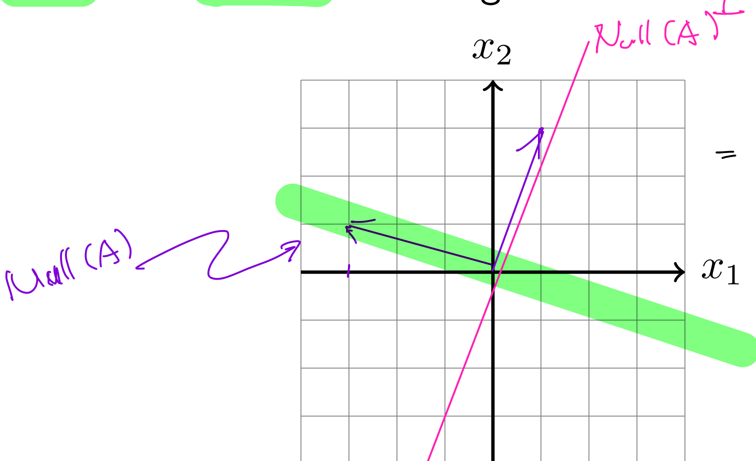
Example: suppose $A = \begin{pmatrix} 1 & 3 \\ 2 & 6 \end{pmatrix}$.



- Col$A$ is the span of $\vec{a}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$
- Col$A^\perp$ is the span of $\vec{z} = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$ $\quad a_1^T$

$\{ \vec{z} : \vec{z} \cdot \vec{a}_1 = 0 \} = \text{Null}\left(\begin{bmatrix} 1 & 2 \end{bmatrix} \right)$

Sketch Null$A$ and Null$A^\perp$ on the grid below.



$\text{Null}(A) = \{ \vec{u} : A \cdot \vec{u} = 0 \}$

$= \left\{ \vec{u} : \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\}$

$= \left\{ \vec{u} : \begin{bmatrix} 1 & 3 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = 0 \right\}$

$u_1 + 3u_2 = 0$

$= \left\{ c \begin{bmatrix} -3 \\ 1 \end{bmatrix} : c \in \mathbb{R} \right\}$

$\text{Null}(A)^\perp = \{ \vec{z} : \vec{z} \cdot \vec{u} = 0 \quad \forall u \in \text{Null}(A) \}$

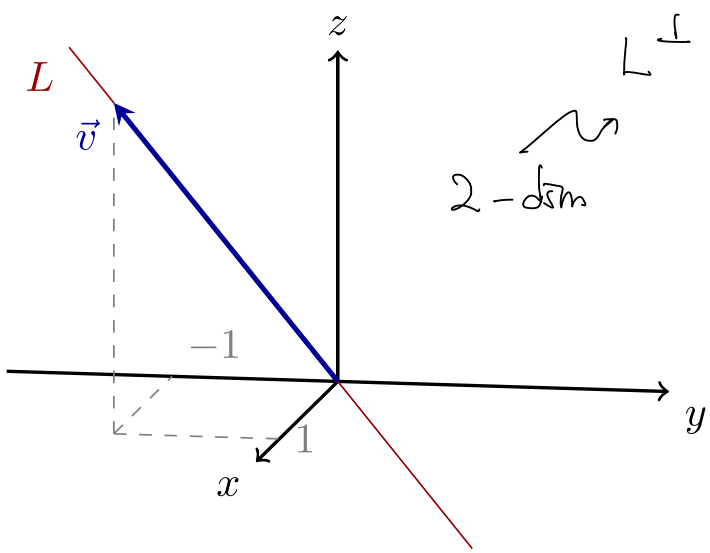
$= \left\{ \vec{z} : \vec{z} \cdot \begin{bmatrix} -3 \\ 1 \end{bmatrix} = 0 \right\} = \{ \vec{z} : -3z_1 + z_2 = 0 \}$

$= \left\{ c \begin{bmatrix} 1 \\ 3 \end{bmatrix} : c \in \mathbb{R} \right\}$

$$L = \text{Span}\left\{\begin{bmatrix}1\\-1\\2\end{bmatrix}\right\} = \left\{c\begin{bmatrix}1\\-1\\2\end{bmatrix} : c \in \mathbb{R}\right\}$$

Example

Line L is a subspace of $\mathbb{R}^3$ spanned by $\vec{v} = \begin{pmatrix}1\\-1\\2\end{pmatrix}$. Then the space $L^\perp$ is a plane. Construct an equation of the plane $L^\perp$.



2-dim

$$L^\perp = \left\{\vec{z} = \begin{bmatrix}z_1\\z_2\\z_3\end{bmatrix} : \vec{z}\cdot\begin{bmatrix}1\\-1\\2\end{bmatrix} = 0\right\}$$

$$= \text{Null}\left(\begin{bmatrix}1 & -1 & 2\end{bmatrix}\right)$$

v^T

$$= \{\vec{z} : z_1 - z_2 + 2z_3 = 0\}$$

Can also visualise line and plane with CalcPlot3D: web.monroecc.edu/calcNSF

Row$A$

$\text{Col}(A) = \text{Row}(A^T)$

$\text{Row}(A) = \text{Col}(A^T)$

Definition

Row$A$ is the space spanned by the rows of matrix A.

We can show that

- $\dim(\text{Row}(A)) = \dim(\text{Col}(A))$
- a basis for Row$A$ is the pivot rows of A

Note that $\text{Row}(A) = \text{Col}(A^T)$, but in general Row$A$ and Col$A$ are not related to each other

$$A = \begin{bmatrix} - & \vec{v}_1 & - \\ & \vdots & \\ - & \vec{v}_m & - \end{bmatrix} \xrightarrow{\text{Row operations}} A'$$

$\text{Row}(A) = \text{Row}(A')$

RREF

Basis for Row(A)

$\Rightarrow \dim(\text{Row}(A)) = \#$ of pivot

$= \dim(\text{Col}(A))$

Dimension Thm : $\dim(\text{Null}(A)) + \dim(\text{Col}(A)) = n$

$$\dim(\text{Col}(A)) = \dim(\text{Row}(A))$$

Example 3

Describe the Null(A) in terms of an orthogonal subspace.

$$= \begin{bmatrix} \vec{v}_1 \cdot \vec{x} \\ \vdots \\ \vec{v}_n \cdot \vec{x} \end{bmatrix}$$

A vector $\vec{x}$ is in $\text{Null}\,A$ if and only if

$$A\vec{x} = \begin{bmatrix} \vec{v}_1 \\ \\ \vec{v}_n \end{bmatrix} \cdot \vec{x} = \vec{0}$$

1. $A\vec{x} = 0$

2. This means that $\vec{x}$ is $\boxed{\text{orthogonal}}$ to each row of A.

3. $\text{Row}\,A$ is $\boxed{\text{Orthogonal}}$ to $\text{Null}\,A$. $\quad (\text{Row}(A))^{\perp} = \text{Null}(A)$

4. The dimension of $\text{Row}\,A$ plus the dimension of $\text{Null}\,A$ equals $\boxed{n}$ $\quad A \in \mathbb{R}^{m \times n}$

Fact : W, $W^{\perp}$ subspaces in $\mathbb{R}^n$

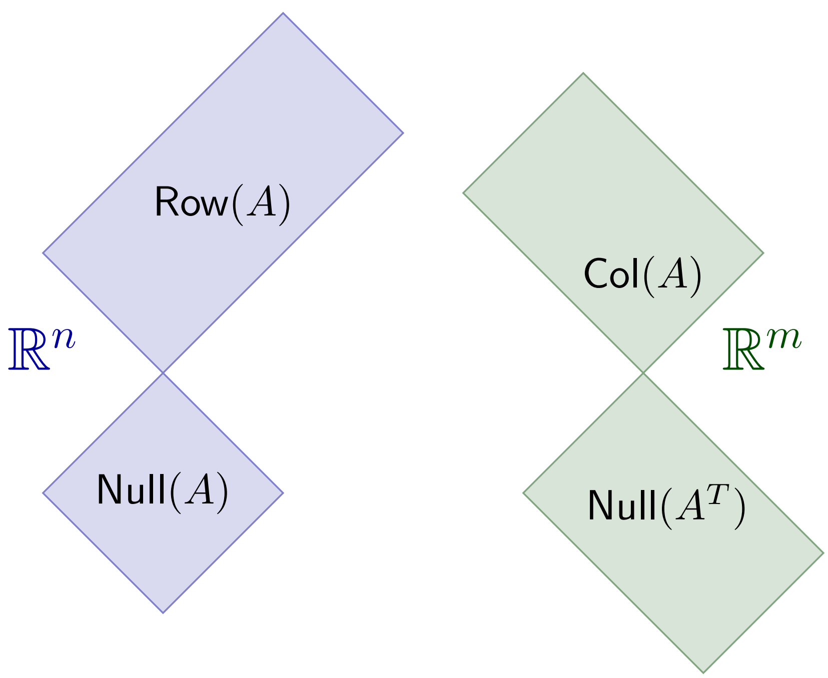
$$\dim(W) + \dim(W^{\perp}) = n$$

- $\mathrm{Row}(A)^{\perp} = \mathrm{Null}(A) = \mathrm{Col}(A^T)^{\perp}$
- $(\mathrm{Row}(A^T))^{\perp} = \mathrm{Col}(A)^{\perp} = \mathrm{Null}(A^T)$

Theorem (The Four Subspaces)

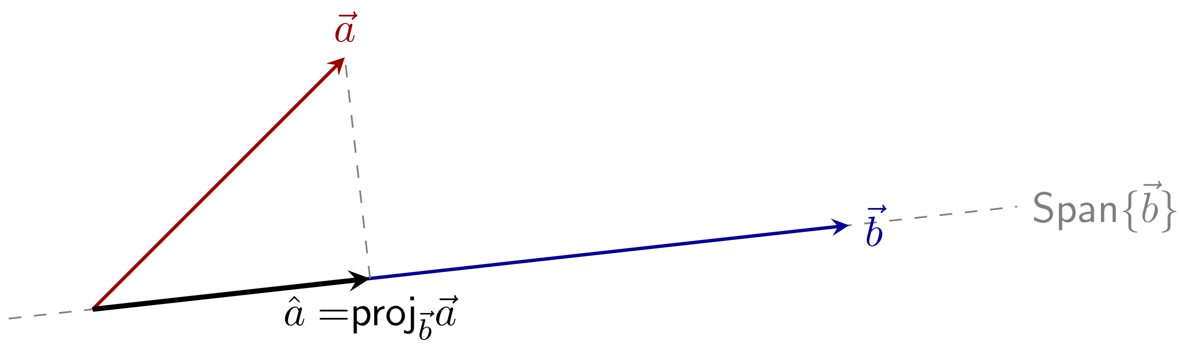
For any $A \in \mathbb{R}^{m\times n}$, the orthogonal complement of $\mathrm{Row}\, A$ is $\mathrm{Null}\, A$, and the orthogonal complement of $\mathrm{Col}\, A$ is $\mathrm{Null}\, A^T$.

The idea behind this theorem is described in the diagram below.



Looking Ahead - Projections

Suppose we want to find the closed vector in Span$\{\vec{b}\}$ to $\vec{a}$.



- Later in this Chapter, we will make connections between dot products and **projections**.
- Projections are also used throughout multivariable calculus courses.

Section 6.2 : Orthogonal Sets

Chapter 6 : Orthogonality and Least Squares

Math 1554 Linear Algebra

Topics and Objectives

Topics

1. Orthogonal Sets of Vectors
2. Orthogonal Bases and Projections.

Learning Objectives

1. Apply the concepts of orthogonality to
   a) compute orthogonal projections and distances,
   b) express a vector as a linear combination of orthogonal vectors,
   c) characterize bases for subspaces of $\mathbb{R}^n$, and
   d) construct orthonormal bases.

Motivating Question

What are the special properties of this basis for $\mathbb{R}^3$?

$$\begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix} / \sqrt{11}, \quad \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} / \sqrt{6}, \quad \begin{bmatrix} -1 \\ -4 \\ 7 \end{bmatrix} / \sqrt{66}$$

Orthogonal Vector Sets

Definition

A set of vectors $\{\vec{u}_1, \ldots, \vec{u}_p\}$ are an **orthogonal set** of vectors if for each $j \neq k$, $\vec{u}_j \perp \vec{u}_k$.

$\vec{u}_j \cdot \vec{u}_k = 0$.

Example: Fill in the missing entries to make $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ an orthogonal set of vectors.

$$\vec{u}_1 = \begin{bmatrix} 4 \\ 0 \\ 1 \end{bmatrix}, \quad \vec{u}_2 = \begin{bmatrix} -2 \\ 0 \\ a=8 \end{bmatrix}, \quad \vec{u}_3 = \begin{bmatrix} 0 \\ \underline{b} \\ c=0 \end{bmatrix}$$

$0 = \vec{u}_1 \cdot \vec{u}_2 = 4 \cdot (-2) + 0 \cdot 0 + 1 \cdot a$, $a = 8$.

$0 = \vec{u}_2 \cdot \vec{u}_3 = (-2) \cdot 0 + 0 \cdot b + 8 \cdot c$, $c = 0$

$0 = \vec{u}_1 \cdot \vec{u}_3$ for any b.

$\{\vec{u}_1, \ldots, \vec{u}_p\}$ lin. indep $\Leftrightarrow$ $c_1\vec{u}_1 + \cdots + c_p\vec{u}_p = 0$ implies $c_1 = c_2 = \cdots = c_p = 0$

Suppose $c_1\vec{u}_1 + \cdots + c_p\vec{u}_p = \vec{0}$

$\Rightarrow \|c_1\vec{u}_1 + \cdots + c_p\vec{u}_p\|^2 = 0 = \underbrace{\overbrace{c_1^2}^{=0}}_{\geq 0}\underbrace{\|u_1\|^2}_{\neq 0} + \underbrace{\overbrace{c_2^2}^{=0}}_{\geq 0}\underbrace{\|\vec{u}_2\|^2}_{\neq 0} + \cdots + \underbrace{\overbrace{c_p^2}^{=0}}_{\geq 0}\underbrace{\|\vec{u}_p\|^2}_{\neq 0}$

Linear Independence

$\Rightarrow c_1 = \cdots = c_p = 0$

Theorem (Linear Independence for Orthogonal Sets)

Let $\{\vec{u}_1, \ldots, \vec{u}_p\}$ be an orthogonal set of vectors. Then, for scalars $c_1, \ldots, c_p$,

$$\left\|c_1\vec{u}_1 + \cdots + c_p\vec{u}_p\right\|^2 = c_1^2\|\vec{u}_1\|^2 + \cdots + c_p^2\|\vec{u}_p\|^2.$$

In particular, if all the vectors $\vec{u}_r$ are non-zero, the set of vectors $\{\vec{u}_1, \ldots, \vec{u}_p\}$ are linearly independent.

$\|c_1\vec{u}_1 + \cdots + c_p\vec{u}_p\|^2 = (c_1\cdot\vec{u}_1 + \cdots + c_p\vec{u}_p)\cdot(c_1\vec{u}_1 + \cdots + c_p\vec{u}_p)$

$= c_1\cdot\vec{u}_1 \cdot c_1\vec{u}_1 + \overbrace{c_1\vec{u}_1\ c_2\vec{u}_2}^{=0} + \cdots + \overbrace{c_1\vec{u}_1 \cdot c_p\cdot\vec{u}_p}^{=0}$

$+ \overbrace{c_2\vec{u}_2 \cdot c_1\vec{u}_1}^{=0} + c_2\vec{u}_2\cdot c_2\vec{u}_2 + \cdots + \overbrace{c_2\vec{u}_2\cdot c_p\vec{u}_p}^{=0}$

$\vdots$

$+ \overbrace{c_p\vec{u}_p\cdot c_1\vec{u}_1}^{=0} + \underbrace{c_p\vec{u}_p\cdot c_2\vec{u}_2}_{=0} + \cdots + c_p\vec{u}_p\cdot c_p\cdot\vec{u}_p$

$= c_1^2\cdot\vec{u}_1\cdot\vec{u}_1 + c_2^2\ \vec{u}_2\cdot\vec{u}_2 + \cdots + c_p^2\cdot\vec{u}_p\cdot\vec{u}_p$

$= c_1^2\ \|\vec{u}_1\|^2 + \cdots + c_p^2\cdot\|\vec{u}_p\|^2$

Recall W : a subspace

$\mathcal{B} = \{\vec{u}_1, \dots, \vec{u}_p\}$ is a basis for W

if $\begin{cases} \mathcal{B} \text{ lin. indep} \\ W = \operatorname{Span} \mathcal{B} \end{cases}$

Orthogonal Bases

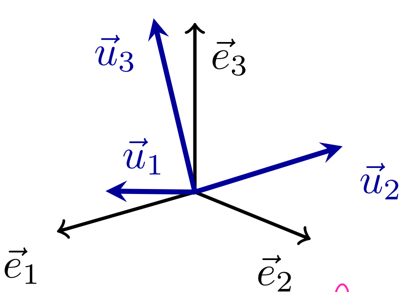
Theorem (Expansion in Orthogonal Basis)

Let $\{\vec{u}_1, \dots, \vec{u}_p\}$ be an orthogonal basis for a subspace W of $\mathbb{R}^n$. Then, for any vector $\vec{w} \in W$,

$$\vec{w} = c_1\vec{u}_1 + \cdots + c_p\vec{u}_p.$$

Above, the scalars are $c_q = \dfrac{\vec{w} \cdot \vec{u}_q}{\vec{u}_q \cdot \vec{u}_q}$.

For example, any vector $\vec{w} \in \mathbb{R}^3$ can be written as a linear combination of $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$, or some other orthogonal basis $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$.



If $\mathcal{B}$ is a orthogonal basis,

For any $\vec{w} \in W$,

$$\vec{w} = c_1\vec{u}_1 + \cdots + c_p\vec{u}_p \quad \text{uniquely}$$

$c_1 = \dfrac{\vec{w} \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1}$, ... $c_p = \dfrac{\vec{w} \cdot \vec{u}_p}{\vec{u}_p \cdot \vec{u}_p}$

$$\vec{u}_1 \cdot \vec{w} = \left(c_1 \vec{u}_1 + \cdots + c_p \vec{u}_p\right) \cdot \vec{u}_1$$

$$\vec{w} \cdot \vec{u}_1 = c_1 \cdot \underbrace{\vec{u}_1 \cdot \vec{u}_1}$$

$$c_1 = \frac{\vec{w} \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1}$$

Example

$$\vec{x} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \vec{u} = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}, \quad \vec{v} = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \quad \vec{s} = \begin{pmatrix} 3 \\ -4 \\ 1 \end{pmatrix}$$

Let W be the subspace of $\mathbb{R}^3$ that is orthogonal to $\vec{x}$.

a) Check that an orthogonal basis for W is given by $\vec{u}$ and $\vec{v}$.

b) Compute the expansion of $\vec{s}$ in basis W.

a) $W = (\mathrm{Span}\{\vec{x}\})^{\perp} \leftarrow$ 2-dim.

① $\vec{u}, \vec{v} \in W$ ($\because$ $\vec{u} \cdot \vec{x} = 0$ $\vec{v} \cdot \vec{x} = 0$)

② $\{\vec{u}, \vec{v}\}$ lin. indep. ($\because$ $\vec{u} \cdot \vec{v} = 0$)

b) $\vec{s} \in W$ because $\vec{x} \cdot \vec{s} = 3 - 4 + 1 = 0$.

$$\vec{s} = c_1 \cdot \vec{u} + c_2 \vec{v}$$

$\{\vec{u}, \vec{v}\}$ orthogonal basis

$$c_1 = \frac{\vec{s} \cdot \vec{u}}{\vec{u} \cdot \vec{u}} = \frac{1 \cdot 3 + (-2)\cdot(-4) + 1 \cdot 1}{1^2 + 2^2 + 1^2} = \frac{12}{6} = 2$$

$$c_2 = \frac{\vec{s} \cdot \vec{v}}{\vec{v} \cdot \vec{v}} = \frac{(-1)\cdot 3 + 0 \cdot(-4) + 1 \cdot 1}{(-1)^2 + 0^2 + 1^2} = \frac{-2}{2} = -1.$$

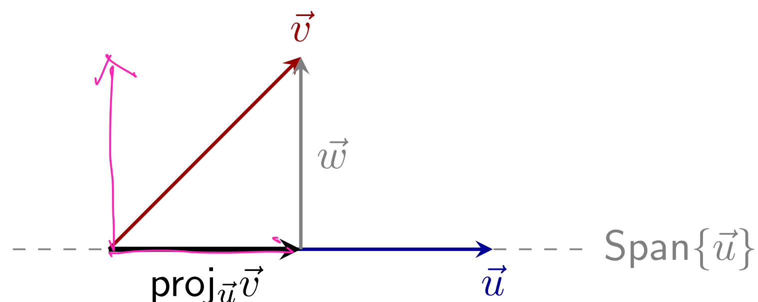
Projections

Let $\vec{u}$ be a non-zero vector, and let $\vec{v}$ be some other vector. The **orthogonal projection of $\vec{v}$ onto the direction of $\vec{u}$** is the vector in the span of $\vec{u}$ that is closest to $\vec{v}$.

$$\boxed{\text{proj}_{\vec{u}}\vec{v} = \frac{\vec{v}\cdot\vec{u}}{\vec{u}\cdot\vec{u}}\vec{u}.}$$

The vector $\vec{w} = \vec{v} - \text{proj}_{\vec{u}}\vec{v}$ is orthogonal to $\vec{u}$, so that

$$\vec{v} = \text{proj}_{\vec{u}}\vec{v} + \vec{w}$$
$$\|\vec{v}\|^2 = \|\text{proj}_{\vec{u}}\vec{v}\|^2 + \|\vec{w}\|^2$$



$\in$ Span$(\vec{u})$

scalar

$$\text{proj}_{\vec{u}}(\vec{v}) = c \cdot \vec{u} = \left(\frac{\vec{u}\cdot\vec{v}}{\vec{u}\cdot\vec{u}}\right)\cdot\vec{u}$$

$$\vec{v} = \underset{\substack{\parallel \\ c\cdot\vec{u}}}{\text{proj}_{\vec{u}}(\vec{v})} + \underbrace{\vec{y}}^{?}$$

orthogonal to $\vec{u}$

$$\vec{u}\cdot\vec{v} = c\cdot\vec{u}\cdot\vec{u} + \cancel{\vec{y}\cdot\vec{u}} \to 0$$

$$c = \frac{\vec{u}\cdot\vec{v}}{\vec{u}\cdot\vec{u}}$$

Example

Let L be spanned by $\vec{u} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$.

$L = \text{Span}\{\vec{u}\}$

$= \text{Span}\{2\cdot\vec{u}\}$

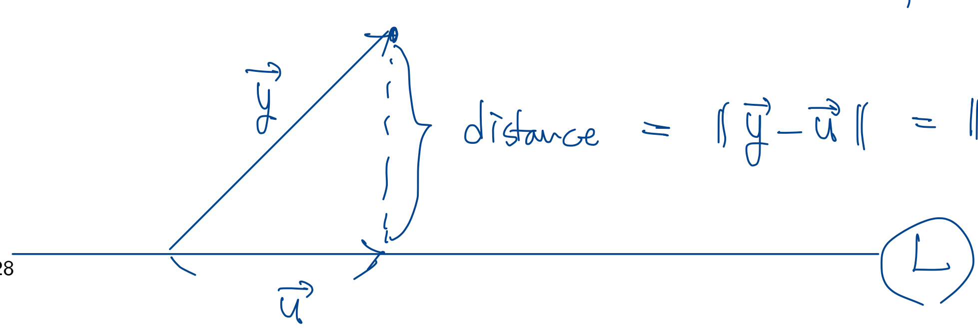
1. Calculate the projection of $\vec{y} = (-3, 5, 6, -4)$ onto line L.
2. How close is $\vec{y}$ to the line L? distance between $\vec{y}$, L

① $\text{proj}_L(\vec{y}) = \text{proj}_{\vec{u}}(\vec{y}) = \left(\dfrac{\vec{u}\cdot\vec{y}}{\vec{u}\cdot\vec{u}}\right)\cdot\vec{u}$

$\text{proj}_{2\vec{u}}(\vec{y})$

$$= \frac{\overbrace{-3+5+6-4}^{=4}}{\underbrace{1^2+1^2+1^2+1^2}_{=4}}\cdot\vec{u} = \vec{u}.$$

②



distance $= \|\vec{y}-\vec{u}\| = \|\vec{y}-\text{proj}_L(\vec{y})\|$

$$= \left\| \begin{bmatrix} -4 \\ 4 \\ 5 \\ -5 \end{bmatrix} \right\| = \sqrt{16+16+25+25}$$

$$= \sqrt{82}$$

Definition

$\|u_1\| = \cdots = \|u_p\| = 1$

Definition (Orthonormal Basis)

An **orthonormal basis** for a subspace W is an orthogonal basis $\{\vec{u}_1, \ldots, \vec{u}_p\}$ in which every vector $\vec{u}_q$ has unit length. In this case, for each $\vec{w} \in W$,

$$\vec{w} = (\vec{w} \cdot \vec{u}_1)\vec{u}_1 + \cdots + (\vec{w} \cdot \vec{u}_p)\vec{u}_p$$

$$\|\vec{w}\| = \sqrt{(\vec{w} \cdot \vec{u}_1)^2 + \cdots + (\vec{w} \cdot \vec{u}_p)^2}$$

$$c_1 = \frac{\vec{w} \cdot \vec{u}_1}{\underbrace{\vec{u}_1 \cdot \vec{u}_1}_{=1}} = \vec{w} \cdot \vec{u}_1$$

Example

The subspace W is a subspace of $\mathbb{R}^3$ perpendicular to $x = (1, 1, 1)$.
Calculate the missing coefficients in the orthonormal basis for W.

$$u = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \qquad \underbrace{v = \frac{1}{\sqrt{6}}}_{} \begin{bmatrix} a \\ b \\ c \end{bmatrix} \begin{matrix} =1 \\ =-2 \\ =1 \end{matrix}$$

① $a+b+c=0$

② $a-c=0$

$u, v \in W \qquad \vec{u} \cdot \vec{x} = \vec{v} \cdot \vec{x} = 0$

$\vec{u} \cdot \vec{v} = 0$

$\|\vec{u}\| = \|\vec{v}\| = 1$

10/27/23

Recall

$\text{proj}_{\vec{u}}(\vec{v}) = \left(\frac{\vec{v}\cdot\vec{u}}{\vec{u}\cdot\vec{u}}\right)\cdot\vec{u}$

$\mathcal{B}$ orthogonal basis $\{\vec{u}_1, \ldots, \vec{u}_p\}$ for W

$\vec{w} \in W \Rightarrow \vec{w} = c_1\vec{u}_1 + \cdots + c_p\vec{u}_p$

$= \frac{\vec{w}\cdot\vec{u}_1}{\vec{u}_1\cdot\vec{u}_1}\vec{u}_1 + \cdots + \frac{\vec{w}\cdot\vec{u}_p}{\vec{u}_p\cdot\vec{u}_p}\cdot\vec{u}_p$

$\mathcal{B}$ orthonormal basis if $\|\vec{u}_1\| = \cdots = \|\vec{u}_p\| = 1$

Orthogonal Matrices

An **orthogonal matrix** is a square ($n \times n$) matrix whose columns are orthonormal.

Theorem

An $m \times n$ matrix U has orthonormal columns if and only if $U^TU = I_n$.

Can U have orthonormal columns if $n > m$?

$\underset{n\times m}{U^T}\cdot\underset{m\times n}{U} \in \mathbb{R}^{n\times n}$

$$\begin{bmatrix} u_1^T \\ \vdots \\ u_n^T \end{bmatrix}\begin{bmatrix} \vec{u}_1 & \cdots & \vec{u}_n \end{bmatrix} = \begin{bmatrix} \underset{=1}{u_1\cdot u_1} & \underset{=0}{u_1\cdot u_2} & \cdots & \underset{=0}{u_1\cdot u_n} \\ \underset{=0}{u_2\cdot u_1} & \underset{=1}{u_2\cdot u_2} & \cdots & \\ & & \ddots & \\ & & & \underset{=1}{u_n\cdot u_n} \end{bmatrix}$$

$= I_n$

Theorem

Theorem (Mapping Properties of Orthogonal Matrices)

Assume $m \times m$ matrix U has orthonormal columns. Then

1. (Preserves length) $\|U\vec{x}\| = \boxed{\|\vec{x}\|}$
2. (Preserves angles) $(U\vec{x}) \cdot (U\vec{y}) = \boxed{\vec{x} \cdot \vec{y}}$
3. (Preserves orthogonality) $\vec{x} \cdot \vec{y} = 0 \iff (U\vec{x}) \cdot (U\vec{y}) = 0$

$U = [\vec{u}_1 \ \cdots \ \vec{u}_m] \qquad \{\vec{u}_1, \ldots, \vec{u}_m\}$ orthonormal

$\| U \vec{x} \|$

$$\|U\vec{x}\|^2 = \left\| U \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} \right\|^2 = \| x_1 \cdot \vec{u}_1 + x_2 \cdot \vec{u}_2 + \cdots + x_m \vec{u}_m \|^2 \quad \text{orthogonal}$$

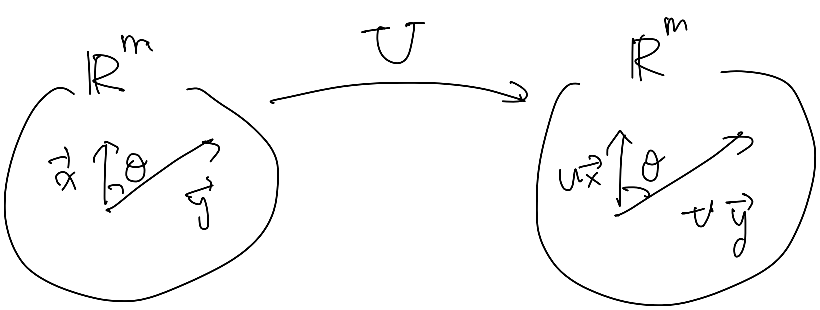
$$= x_1^2 \|\vec{u}_1\|^2 + x_2^2 \|u_2\|^2 + \cdots + x_m^2 \|\vec{u}_m\|^2 \qquad (\|\vec{u}_i\|^2 = 1)$$

$$= x_1^2 + x_2^2 + \cdots + x_m^2 = \|\vec{x}\|^2$$

$$\|U\vec{x}\|^2 = (U\vec{x}) \cdot (U\vec{x}) = (U \cdot \vec{x})^T \cdot (U \cdot \vec{x})$$

$$= x^T \cdot \underbrace{U^T \cdot U}_{= I_m} \cdot \vec{x} = x^T \cdot x$$

$$= \|\vec{x}\|^2$$



Example

$\|Ux\| = \|x\|$

Compute the length of the vector below. U

$$\left\| \underbrace{\begin{bmatrix} 1/2 & 2/\sqrt{14} \\ 1/2 & 1/\sqrt{14} \\ 1/2 & -3/\sqrt{14} \\ 1/2 & 0 \end{bmatrix}}_{4\times 2} \underbrace{\begin{bmatrix} \sqrt{2} \\ -3 \end{bmatrix}}_{2\times 1} \right\| = \left\| \begin{bmatrix} \sqrt{2} \\ -3 \end{bmatrix} \right\| = \sqrt{11}.$$

$\vec{x} \in \mathbb{R}^4$

has orthonormal columns

$$\frac{1}{2}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \cdot \frac{1}{\sqrt{14}}\begin{bmatrix} 2 \\ 1 \\ -3 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

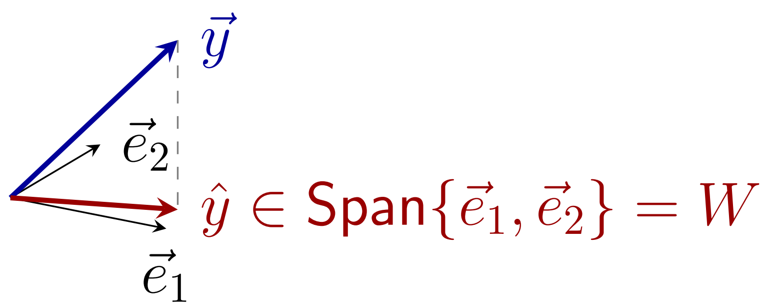
$$\left\| \frac{1}{2}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \right\|^2 = \frac{1}{4} + \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = 1$$

$$\left\| \frac{1}{\sqrt{14}}\begin{bmatrix} 2 \\ 1 \\ -3 \\ 0 \end{bmatrix} \right\|^2 = \frac{1}{14}\left(2^2 + 1^2 + (-3)^2\right) = 1$$

Section 6.3 : Orthogonal Projections

Chapter 6 : Orthogonality and Least Squares

Math 1554 Linear Algebra



Vectors $\vec{e}_1$ and $\vec{e}_2$ form an orthonormal basis for subspace W.
Vector $\vec{y}$ is not in W.
The orthogonal projection of $\vec{y}$ onto $W =$Span$\{\vec{e}_1, \vec{e}_2\}$ is $\hat{y}$.

Topics and Objectives

Topics

1. Orthogonal projections and their basic properties
2. Best approximations

Learning Objectives

1. Apply concepts of orthogonality and projections to
   a) compute orthogonal projections and distances,
   b) express a vector as a linear combination of orthogonal vectors,
   c) construct vector approximations using projections,
   d) characterize bases for subspaces of $\mathbb{R}^n$, and
   e) construct orthonormal bases.

Motivating Question For the matrix A and vector $\vec{b}$, which vector $\widehat{b}$ in column space of A, is closest to $\vec{b}$?

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 0 \\ -4 & -2 \end{bmatrix}, \qquad \vec{b} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Example 1

basis
orthogonal
unit lengths

Let $\vec{u}_1, \ldots, \vec{u}_5$ be an orthonormal basis for $\mathbb{R}^5$. Let $W = \text{Span}\{\vec{u}_1, \vec{u}_2\}$.
For a vector $\vec{y} \in \mathbb{R}^5$, write $\vec{y} = \widehat{y} + w^{\perp}$, where $\widehat{y} \in W$ and $w^{\perp} \in W^{\perp}$.

$$\vec{y} = c_1 \vec{u}_1 + c_2 \vec{u}_2 + c_3 \vec{u}_3 + c_4 \vec{u}_4 + c_5 \vec{u}_5$$

$$= \underbrace{(\vec{y} \cdot \vec{u}_1) \cdot \vec{u}_1 + (\vec{y} \cdot \vec{u}_2) \cdot \vec{u}_2}_{\in W} +$$

$$\text{proj}_W(\vec{y}) = \widehat{y}$$

$$\underbrace{(\vec{y} \cdot \vec{u}_3) \cdot \vec{u}_3 + (\vec{y} \cdot \vec{u}_4) \vec{u}_4 + (\vec{y} \cdot \vec{u}_5) \cdot \vec{u}_5}_{= w^{\perp}}$$

$\Uparrow$

$W^{\perp}$

u_5
u_4
$u_3 \in W^{\perp}$ because

$\Leftrightarrow \quad u_3 \cdot \vec{w} = 0 \qquad \vec{w} \in W$

$\Leftrightarrow \quad u_3 \cdot \vec{u}_1 = 0 \quad , \quad \vec{u}_3 \cdot \vec{u}_2 = 0$

Orthogonal Decomposition Theorem

Theorem

Let W be a subspace of $\mathbb{R}^n$. Then, each vector $\vec{y} \in \mathbb{R}^n$ has the **unique** decomposition

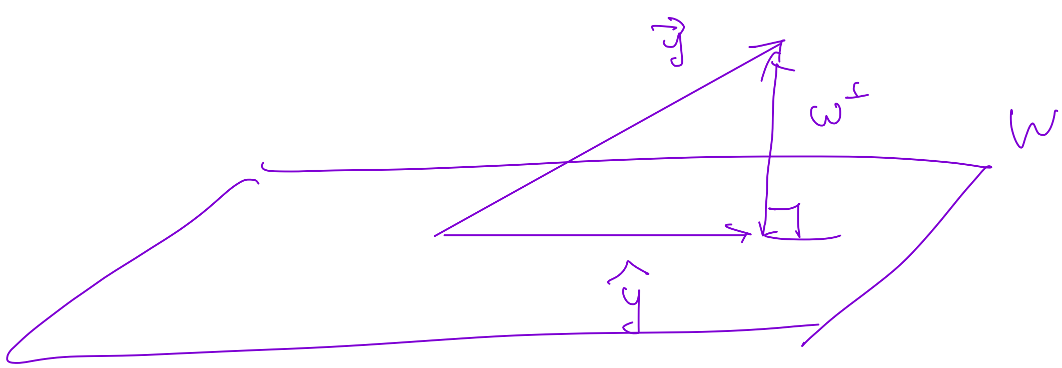
$$\vec{y} = \widehat{y} + w^{\perp}, \quad \widehat{y} \in W, \quad w^{\perp} \in W^{\perp}.$$

And, if $\vec{u}_1, \ldots, \vec{u}_p$ is any orthogonal basis for W,

$$\hat{y} = \frac{\vec{y} \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1}\vec{u}_1 + \cdots + \frac{\vec{y} \cdot \vec{u}_p}{\vec{u}_p \cdot \vec{u}_p}\vec{u}_p.$$

We say that $\widehat{y}$ is the **orthogonal projection of $\vec{y}$ onto W.**

If time permits, we will explain some of this theorem on the next slide.



Explanation (if time permits)

We can write

$$\widehat{y} =$$

Then, $w^{\perp} = \vec{y} - \widehat{y}$ is in $W^{\perp}$ because

Example 2a

$$\vec{y} = \begin{pmatrix} 4 \\ 0 \\ 3 \end{pmatrix}, \quad \vec{u}_1 = \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix}, \quad \vec{u}_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

$\vec{u}_1 \cdot \vec{u}_2 = 0$

Construct the decomposition $\vec{y} = \widehat{y} + w^{\perp}$, where $\widehat{y}$ is the orthogonal projection of $\vec{y}$ onto the subspace $W = \text{Span}\{\vec{u}_1, \vec{u}_2\}$.

orthogonal (Yes)
othonormal (No)

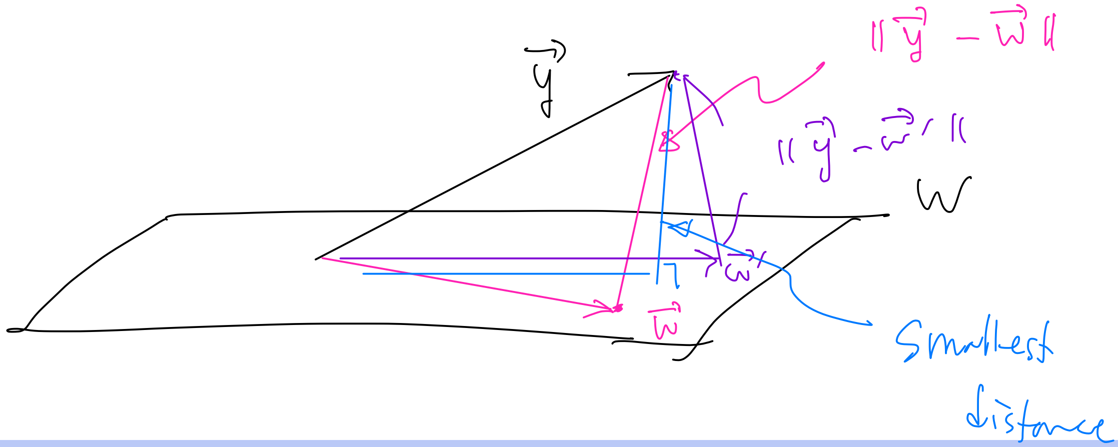
$$\hat{y} = \text{proj}_W(\vec{y})$$

$$= \frac{\vec{y} \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1} u_1 + \frac{\vec{y} \cdot \vec{u}_2}{\vec{u}_2 \cdot \vec{u}_2} u_2$$

$\{u_1, u_2\}$ orthogonal basis for W

$$= \frac{4 \cdot 2 + 0 \cdot 2 + 3 \cdot 0}{2^2 + 2^2} \vec{u}_1 + \frac{3 \cdot 1}{1^2} \vec{u}_2$$

$$= \vec{u}_1 + 3 \cdot \vec{u}_2 = \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix}$$



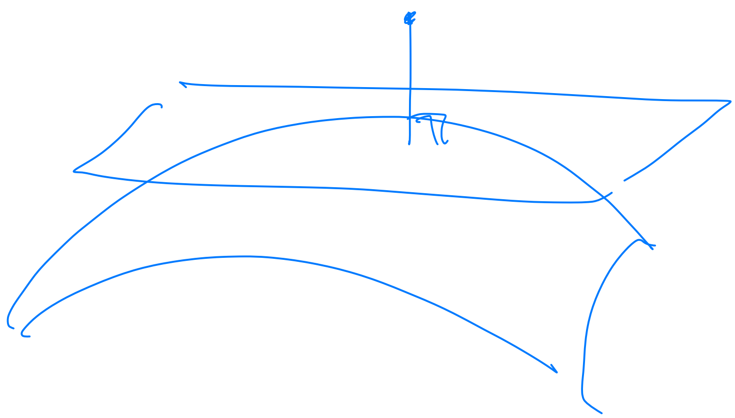
Best Approximation Theorem

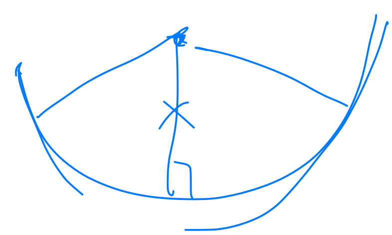
Theorem

Let W be a subspace of $\mathbb{R}^n$, $\vec{y} \in \mathbb{R}^n$, and $\widehat{y}$ is the orthogonal projection of $\vec{y}$ onto W. Then for **any** $\vec{w} \neq \hat{y} \in W$, we have

$$\|\vec{y} - \widehat{y}\| < \|\vec{y} - \vec{w}\|$$

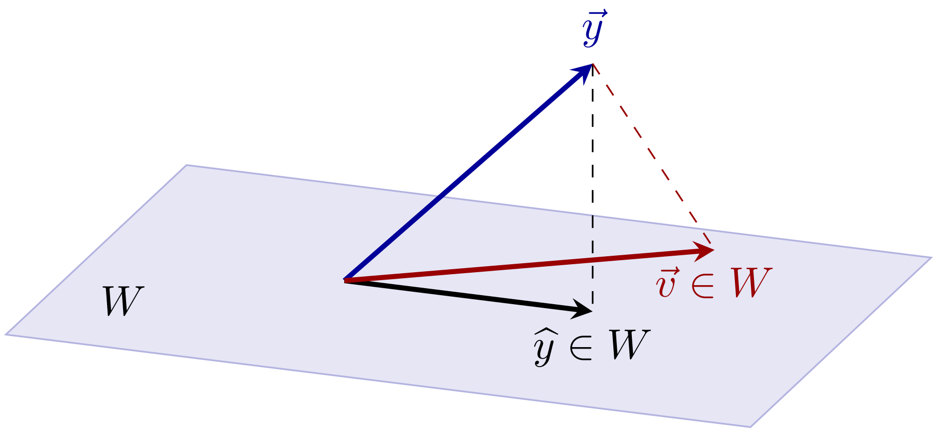
That is, $\widehat{y}$ is the unique vector in W that is closest to $\vec{y}$.





Proof (if time permits)

The orthogonal projection of $\vec{y}$ onto W is the closest point in W to $\vec{y}$.



Example 2b

$$\vec{y} = \begin{pmatrix} 4 \\ 0 \\ 3 \end{pmatrix}, \quad \vec{u}_1 = \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix}, \quad \vec{u}_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

What is the distance between $\vec{y}$ and subspace $W = \text{Span}\{\vec{u}_1, \vec{u}_2\}$? Note that these vectors are the same vectors that we used in Example 2a.

$$\hat{y} = \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix} \longleftarrow \text{minimizes} \quad \| \vec{y} - \underset{1}{c_1} \vec{u}_1 - \underset{3}{c_2} \vec{u}_2 \|$$

$$\text{distance}(\vec{y}, W) = \| \vec{y} - \hat{y} \|$$

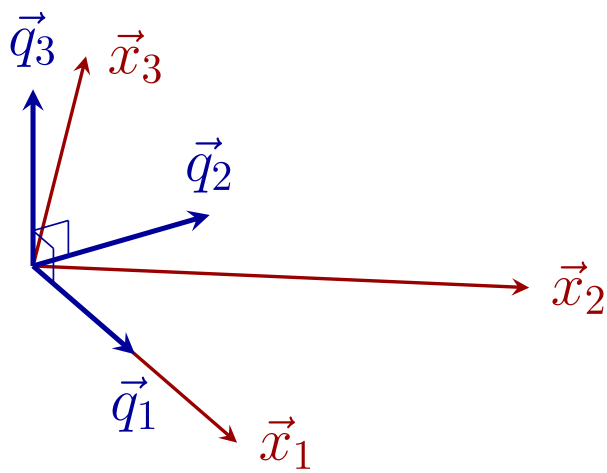
$$= \left\| \begin{bmatrix} 4 \\ 0 \\ 3 \end{bmatrix} - \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix} \right\| = \left\| \begin{bmatrix} 2 \\ -2 \\ 0 \end{bmatrix} \right\|$$

$$= \sqrt{8}.$$

Section 6.4 : The Gram-Schmidt Process

Chapter 6 : Orthogonality and Least Squares

Math 1554 Linear Algebra



Vectors $\vec{x}_1, \vec{x}_2, \vec{x}_3$ are given linearly independent vectors. We wish to construct an orthonormal basis $\{\vec{q}_1, \vec{q}_2, \vec{q}_3\}$ for the space that they span.

Topics and Objectives

Topics

1. Gram Schmidt Process
2. The QR decomposition of matrices and its properties

Learning Objectives

1. Apply the iterative Gram Schmidt Process, and the QR decomposition, to construct an orthogonal basis.
2. Compute the QR factorization of a matrix.

Motivating Question The vectors below span a subspace W of $\mathbb{R}^4$. Identify an orthogonal basis for W.

$$\vec{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad \vec{x}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad \vec{x}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}.$$

Example

The vectors below span a subspace W of $\mathbb{R}^4$. Construct an orthogonal basis for W.

$$\vec{v}_1 = \vec{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad \vec{x}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad \vec{x}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}.$$

$W = \mathrm{Span}(\{\vec{x}_1, \vec{x}_2, \vec{x}_3\})$ Orthogonal basis ?

$\vec{v}_1 = \vec{x}_1$

Find $\vec{v}_2$ s.t. ① $\vec{v}_2 \perp \vec{v}_1$ i.e. $\vec{v}_1 \cdot \vec{v}_2 = 0$

② $\mathrm{Span}\{\vec{x}_1, \vec{x}_2\} = \mathrm{Span}\{\vec{v}_1, \vec{v}_2\}$

lin. Combi. of $\vec{x}_1, \vec{x}_2$

$0 = \vec{v}_1 \cdot \vec{v}_2 = (\vec{x}_2 - c \cdot \underset{\vec{v}_1}{\vec{x}_1}) \cdot \vec{v}_1$

$0 = \vec{x}_2 \cdot \vec{v}_1 - c \cdot \vec{v}_1 \cdot \vec{v}_1 \Rightarrow c = \dfrac{\vec{x}_2 \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1} = \dfrac{3}{4}$

$\vec{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} - \frac{3}{4}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \frac{1}{4}\begin{bmatrix} -3 \\ 1 \\ 1 \\ 1 \end{bmatrix}$

Find $\vec{v}_3 = \vec{x}_3 - c_1 \vec{v}_1 - c_2 \vec{v}_2 \Rightarrow \mathrm{Span}\{x_1, x_2, x_3\} = \mathrm{Span}\{v_1, v_2, v_3\} = W$

<u>Need</u>: $\vec{v}_1 \cdot \vec{v}_3 = 0$, $\vec{v}_2 \cdot \vec{v}_3 = 0$

$0 = \vec{v}_1 \cdot \vec{v}_3 = (\vec{x}_3 - c_1 \vec{v}_1 - c_2 \vec{v}_2) \cdot \vec{v}_1 = \vec{x}_3 \cdot \vec{v}_1 - c_1 \vec{v}_1 \cdot \vec{v}_1$

$\therefore \; c_1 = \dfrac{\vec{x}_3 \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1} = \dfrac{1}{2}$

$$0 = \vec{v}_2 \cdot \vec{v}_3 = (\vec{x}_3 - c_1\vec{v}_1 - c_2\vec{v}_2)\cdot \vec{v}_2$$

$$= \vec{x}_3 \cdot \vec{v}_2 - \underline{c_2} \cdot \vec{v}_2 \cdot \vec{v}_2$$

$$c_2 = \frac{\vec{x}_3 \cdot \vec{v}_2}{\vec{v}_2 \cdot \vec{v}_2} = \frac{1/2}{(\frac{1}{4})^2 \cdot 12} = \frac{1}{\cancel{2}} \cdot \frac{\cancel{16}^{\,8}}{12} = \frac{2}{3}.$$

$$v_3 = \underbrace{\begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} - \frac{1}{2}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}}_{} = v_1 \quad - \frac{\cancel{2}}{3} \cdot \frac{1}{\cancel{4}_{\,2}} \begin{bmatrix} -3 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

$$= \frac{1}{2}\begin{bmatrix} -1 \\ -1 \\ 1 \\ 1 \end{bmatrix} - \frac{1}{6}\begin{bmatrix} -3 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \frac{1}{6}\left(\begin{bmatrix} -3 \\ -3 \\ 3 \\ 3 \end{bmatrix} - \begin{bmatrix} -3 \\ 1 \\ 1 \\ 1 \end{bmatrix} \right)$$

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

The Gram-Schmidt Process

Given a basis $\{\vec{x}_1, \ldots, \vec{x}_p\}$ for a subspace W of $\mathbb{R}^n$, iteratively define

lin. indep.

$(\mathrm{Span}\{v_1\})^{\perp}$ $=$ $W^{\perp}$ part of x_2

$$\vec{v}_1 = \vec{x}_1$$

$$\vec{v}_2 = \vec{x}_2 - \frac{\vec{x}_2 \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1}\vec{v}_1 \qquad = \mathrm{proj}_{\vec{v}_1}(\vec{x}_2)$$

$$\vec{v}_3 = \vec{x}_3 - \left(\frac{\vec{x}_3 \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1}\vec{v}_1 + \frac{\vec{x}_3 \cdot \vec{v}_2}{\vec{v}_2 \cdot \vec{v}_2}\vec{v}_2\right)$$

$\mathrm{proj}_{\mathrm{Span}\{v_1, v_2\}}(\vec{x}_3)$

$\mathrm{Span}\{v_1, v_2\} = \mathrm{Span}\{x_1, x_2\}$

$$\vdots$$

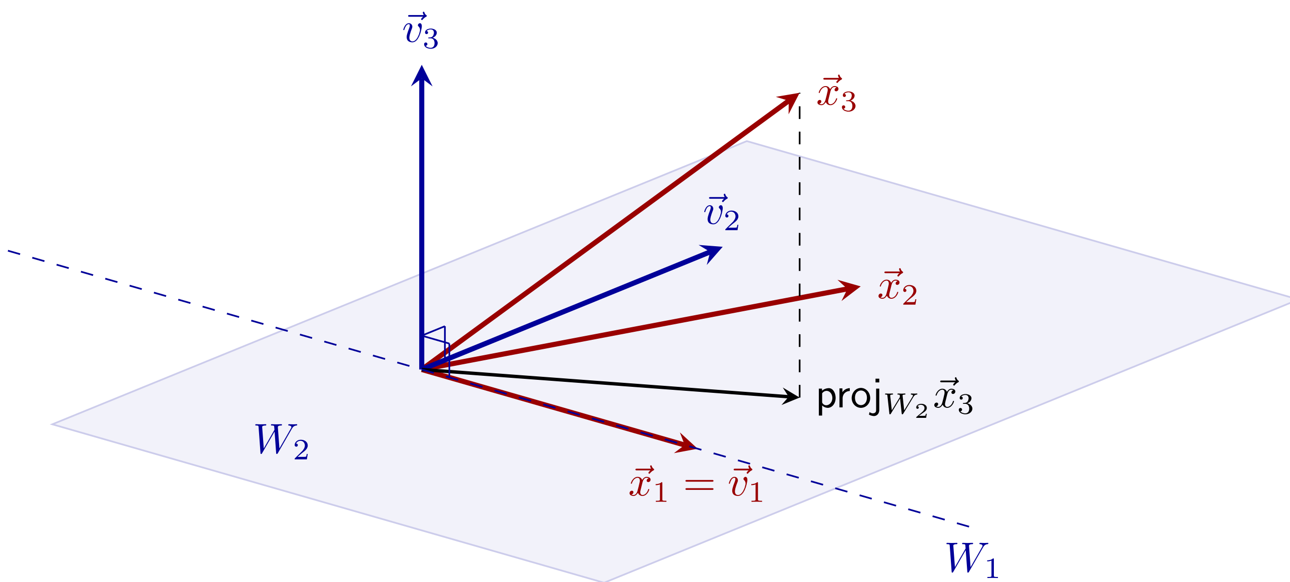
$$\vec{v}_p = \vec{x}_p - \frac{\vec{x}_p \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1}\vec{v}_1 - \cdots - \frac{\vec{x}_p \cdot \vec{v}_{p-1}}{\vec{v}_{p-1} \cdot \vec{v}_{p-1}}\vec{v}_{p-1}$$

Then, $\{\vec{v}_1, \ldots, \vec{v}_p\}$ is an orthogonal basis for W.

Proof

Geometric Interpretation

Suppose $\vec{x}_1, \vec{x}_2, \vec{x}_3$ are linearly independent vectors in $\mathbb{R}^3$. We wish to construct an orthogonal basis for the space that they span.



We construct vectors $\vec{v}_1, \vec{v}_2, \vec{v}_3$, which form our **orthogonal** basis.
$W_1 = \text{Span}\{\vec{v}_1\}$, $W_2 = \text{Span}\{\vec{v}_1, \vec{v}_2\}$.

Lin. Indep $\xrightarrow{G-S}$ Orthogonal $\Longrightarrow$ Orthonormal

$\{\vec{x}_1, \ldots, \vec{x}_p\}$ $\{\vec{v}_1, \ldots, \vec{v}_p\}$ $\{\vec{u}_1, \ldots, \vec{u}_p\}$

Orthonormal Bases

$\vec{u}_1 = \frac{\vec{v}_1}{\|\vec{v}_1\|}, \ldots, \vec{u}_p = \frac{\vec{v}_p}{\|\vec{v}_p\|}$

orthogonal basis with length 1.

||

Definition

A set of vectors form an **orthonormal basis** if the vectors are mutually orthogonal and have unit length. ONB.

Example

The two vectors below form an orthogonal basis for a subspace W. Obtain an orthonormal basis for W.

$$\vec{v}_1 = \begin{bmatrix} 3 \\ 2 \\ 0 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} -2 \\ 3 \\ 1 \end{bmatrix}.$$

$$\vec{u}_1 = \underbrace{\frac{1}{\sqrt{3^2+2^2}}}_{= \frac{1}{\sqrt{13}}} \begin{bmatrix} 3 \\ 2 \\ 0 \end{bmatrix}, \quad \vec{u}_2 = \underbrace{\frac{1}{\sqrt{(-2)^2+3^2+1^2}}}_{= \frac{1}{\sqrt{14}}} \begin{bmatrix} -2 \\ 3 \\ 1 \end{bmatrix}$$

$A = [\vec{x}_1 \ \cdots \ \vec{x}_p] \xrightarrow[\text{normalization}]{G-S} Q = [\vec{u}_1 \ \cdots \ \vec{u}_p]$

$\Downarrow$

$A = QR$

QR Factorization

Theorem

Any $m \times n$ matrix A with linearly independent columns has the **QR factorization**

$$A = QR$$

where

1. Q is $m \times n$, its columns are an orthonormal basis for $\operatorname{Col} A$.
2. R is $n \times n$, upper triangular, with positive entries on its diagonal, and the length of the j^{th} column of R is equal to the length of the j^{th} column of A.

In the interest of time:

- we will not consider the case where A has linearly dependent columns
- students are not expected to know the conditions for which A has a QR factorization

Proof

$\{\vec{x}_1, \dots, \vec{x}_p\}$ lin. indep.

$$\vec{v}_1 = \vec{x}_1 \longrightarrow \vec{u}_1 = \frac{1}{\sqrt{\vec{x}_1 \cdot \vec{x}_1}}\vec{x}_1 = \frac{1}{\|\vec{v}_1\|}\vec{v}_1$$

$$\vec{v}_2 = \vec{x}_2 - (\vec{x}_2 \cdot \vec{u}_1)\cdot \vec{u}_1 \longrightarrow \vec{u}_2 = \frac{1}{\|\vec{v}_2\|}\vec{v}_2$$

$$\vec{v}_3 = \vec{x}_3 - (\vec{x}_3 \cdot \vec{u}_1)\vec{u}_1 - (\vec{x}_3 \cdot \vec{u}_2)\vec{u}_2 \longrightarrow \vec{u}_3 = \frac{1}{\|\vec{v}_3\|}\vec{v}_3$$

$$\vdots$$

$$\vec{x}_1 = \|\vec{x}_1\| \cdot \vec{u}_1$$

$$\vec{x}_2 = \|\vec{v}_2\| \cdot \vec{u}_2 + (\vec{x}_2 \cdot \vec{u}_1)\vec{u}_1$$

$$\vec{x}_3 = \|\vec{v}_3\| \cdot \vec{u}_3 + (\vec{x}_3 \cdot \vec{u}_1)\cdot\vec{u}_1 + (\vec{x}_3 \cdot \vec{u}_2)\cdot\vec{u}_2$$

$$\vdots$$

$$A = \begin{bmatrix} \vec{x}_1 & \vec{x}_2 & \cdots & \vec{x}_p \end{bmatrix} = \begin{bmatrix} \vec{u}_1 & \cdots & \vec{u}_p \end{bmatrix} \begin{bmatrix} \|x_1\| & x_2\cdot u_1 & \\ 0 & \|v_2\| & \\ \vdots & 0 & \\ & \vdots & \\ 0 & 0 & 0 \end{bmatrix}$$

$$\vec{x}_1 = \begin{bmatrix} \vec{u}_1 & \cdots & \vec{u}_p \end{bmatrix} \begin{bmatrix} \|x_1\| \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$\vec{x}_2 = \underbrace{\begin{bmatrix} \vec{u}_1 & \cdots & \vec{u}_p \end{bmatrix}}_{Q} \begin{bmatrix} \vec{x}_2 \cdot \vec{u}_1 \\ \|\vec{v}_2\| \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Example

Construct the QR decomposition for $A = \begin{bmatrix} 3 & -2 \\ 2 & 3 \\ 0 & 1 \end{bmatrix}$.

11/1/23

$\{\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_p\}$ lin. indep.

$\downarrow$ Gram-Schmidt

$W_1 = \text{Span}\{\vec{x}_1\} = \text{Span}\{\vec{v}_1\}$

$\{\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_p\}$ orthogonal set + $W_2 = \text{Span}\{\vec{x}_1, \vec{x}_2\} = \text{Span}\{\vec{v}_1, \vec{v}_2\}$

$\vdots \qquad \vdots$

$W_p = \text{Span}\{x_1, \cdots, x_p\} = \text{Span}\{v_1, \cdots, v_p\}$

How to find $\{v_1, \cdots, v_p\}$

$v_1 = x_1$

$v_2 = x_2 - \text{proj}_{W_1}(x_2) = x_2 - \left(\frac{x_2 \cdot v_1}{v_1 \cdot v_1}\right) \cdot v_1$

$v_3 = x_3 - \text{proj}_{W_2}(x_3) = x_3 - \left(\frac{x_3 \cdot v_1}{v_1 \cdot v_1}\right) v_1 - \left(\frac{x_3 \cdot v_2}{v_2 \cdot v_2}\right) v_2$

$\vdots$

$\{\vec{u}_1, \cdots, \vec{u}_p\}$ orthonormal

$\vec{u}_1 = \frac{\vec{v}_1}{\|\vec{v}_1\|}, \quad \vec{u}_2 = \frac{\vec{v}_2}{\|\vec{v}_2\|}, \quad \cdots, \quad \vec{u}_p = \frac{\vec{v}_p}{\|\vec{v}_p\|}$

$\Rightarrow$

$x_1 \in \text{Span}\{u_1\}$

$x_1 = (x_1 \cdot u_1) \cdot u_1 + 0 \cdot u_2 + 0 \cdot u_3 + \cdots + 0 \cdot u_p$

$x_2 = (x_2 \cdot u_1)\, u_1 + (x_2 \cdot u_2)\, u_2$

$x_2 \in \text{Span}\{u_1, u_2\}$

$x_3 = (x_3 \cdot u_1)\, u_1 + (x_3 \cdot u_2)\, u_2 + (x_3 \cdot u_3)\, u_3$

$\vdots \qquad \vdots$

$x_p = (x_p \cdot u_1)\, u_1 + (x_p \cdot u_2)\, u_2 + \cdots + (x_p \cdot u_p)\, u_p$

$$\Rightarrow \underbrace{[\, x_1 \;\; x_2 \;\; \cdots \;\; x_p \,]}_{A} = \underbrace{[\, u_1 \;\; \cdots \;\; u_p \,]}_{Q} \underbrace{\begin{bmatrix} x_1 \cdot u_1 & x_2 \cdot u_1 & & x_p \cdot u_1 \\ 0 & x_2 \cdot u_2 & & x_p \cdot u_2 \\ 0 & 0 & \cdots & \vdots \\ \vdots & \vdots & & \vdots \\ 0 & 0 & & x_p \cdot u_p \end{bmatrix}}_{R}$$

Q has orthonormal columns

$R \leftarrow$ upper triangular.

7. (4 points) **Show all work for problems on this page.**
Let $\mathcal{B} = \{\vec{x}_1, \vec{x}_2, \vec{x}_3\}$ be a basis for a subspace W of $\mathbb{R}^4$, where

$$\vec{x}_1 = \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}, \qquad \vec{x}_2 = \begin{pmatrix} -2 \\ 1 \\ 1 \\ 2 \end{pmatrix}, \qquad \vec{x}_3 = \begin{pmatrix} 0 \\ 2 \\ -1 \\ -1 \end{pmatrix}.$$

(a) Apply the Gram-Schmidt process to the set of vectors $\{\vec{x}_1, \vec{x}_2, \vec{x}_3\}$ to find an orthogonal basis $\mathcal{H} = \{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ for W. *Clearly show all steps of the Gram-Schmidt process.*

$$v_1 = x_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}$$

$$v_2 = x_2 - \left(\frac{x_2 \cdot v_1}{v_1 \cdot v_1}\right) \cdot v_1 = \begin{bmatrix} -2 \\ 1 \\ 1 \\ 2 \end{bmatrix} - \frac{(-4)}{4} \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}$$

$\mathcal{H} =$ $\boxed{}$

$$= \begin{bmatrix} -1 \\ 0 \\ 2 \\ 1 \end{bmatrix} \qquad (\quad) x_1 - (\quad) x_2$$

$$v_3 = x_3 - \left(\frac{x_3 \cdot v_1}{v_1 \cdot v_1}\right) v_1 - \left(\frac{x_3 \cdot v_2}{v_2 \cdot v_2}\right) v_2 = \begin{bmatrix} 0 \\ 2 \\ -1 \\ -1 \end{bmatrix} - \underbrace{\frac{(-2)}{4}}_{\frac{1}{2}} \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix} - \underbrace{\frac{(-3)}{6}}_{\frac{1}{2}} \begin{bmatrix} -1 \\ 0 \\ 2 \\ 1 \end{bmatrix}$$

$$= \frac{1}{2}\left(\begin{bmatrix} 0 \\ 4 \\ -2 \\ -2 \end{bmatrix} + \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix} + \begin{bmatrix} -1 \\ 0 \\ 2 \\ 1 \end{bmatrix} \right)$$

$$= \frac{1}{2} \begin{bmatrix} 0 \\ 3 \\ 1 \\ -2 \end{bmatrix}.$$

(b) In the space below, **check** that the vectors in the basis $\mathcal{H}$ form an orthogonal set.

$$v_3 = x_3 - c_1 x_1 - c_2 x_2$$

$$v_3 = x_3 - c_1 v_1 - c_2 v_2$$

$$0 = \vec{v}_1 \cdot \vec{v}_3 = v_1 \cdot (x_3 - c_1 v_1 - \underline{c_2 v_2})$$

$$0 = x_3 \cdot v_1 - c_1 \cdot v_1 \cdot v_1 \Rightarrow c_1 = \frac{x_3 \cdot v_1}{v_1 \cdot v_1}$$

Midterm 3. Your initials: ______________

You do not need to justify your reasoning for questions on this page.

(c) (2 points) The standard matrix of a linear transformation $T : \mathbb{R}^3 \to \mathbb{R}^3$ has orthonormal columns. Which one of the following statements is **false**?

Choose only one.

- ◯ $\|T(\vec{x})\| = \|\vec{x}\|$ for all $\vec{x}$ in $\mathbb{R}^3$.
- ◯ If two non-zero vectors $\vec{x}$ and $\vec{y}$ in $\mathbb{R}^3$ are scalar multiples of each other, then $\|T(\vec{x}+\vec{y})\|^2 = \|T(\vec{x})\|^2 + \|T(\vec{y})\|^2$.
- ◯ If $\mathcal{P}$ is a parallelpiped in $\mathbb{R}^3$, then the volume of $T(\mathcal{P})$ is equal to the volume of $\mathcal{P}$.
- ◯ T is one-to-one.

$Q^T \cdot Q = I$

$R = \begin{bmatrix} x_1 \cdot u_1 & x_2 \cdot u_1 \\ 0 & x_2 \cdot u_2 \end{bmatrix}$

$\underset{m\times n}{A} = \underset{m\times n}{Q} \cdot \underset{n\times n}{R}$ upper triangular.

$m = 4$

$n = 2$

$\Rightarrow \quad Q^T \cdot A = Q^T \cdot Q \cdot R = IR = R$.

2. (2 points) Suppose that, in the QR factorization of A, we have Q as given below. Find R.

$$A = \begin{array}{cc} x_1 & x_2 \\ \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} \end{array} \qquad Q = \frac{1}{2} \begin{array}{cc} u_1 & u_2 \\ \begin{bmatrix} 1 & 1/\sqrt{3} \\ 1 & 1/\sqrt{3} \\ 1 & -\sqrt{3} \\ 1 & 1/\sqrt{3} \end{bmatrix} \end{array}$$

$u_2 = \frac{1}{2} \cdot \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ -3 \\ 1 \end{bmatrix}$

$-\sqrt{3} = \frac{1}{\sqrt{3}} \cdot (-3)$

Note: Please fill in the blanks and do not place values in front of the matrix for this problem.

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

$x_1 \cdot u_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \cdot \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = 2$

$x_2 \cdot u_1 = \begin{bmatrix} 1 \\ 1 \\ -1 \\ 1 \end{bmatrix} \cdot \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = 1$

$x_2 \cdot u_2 = \begin{bmatrix} 1 \\ 1 \\ -1 \\ 1 \end{bmatrix} \cdot \frac{1}{2\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ -3 \\ 1 \end{bmatrix}$

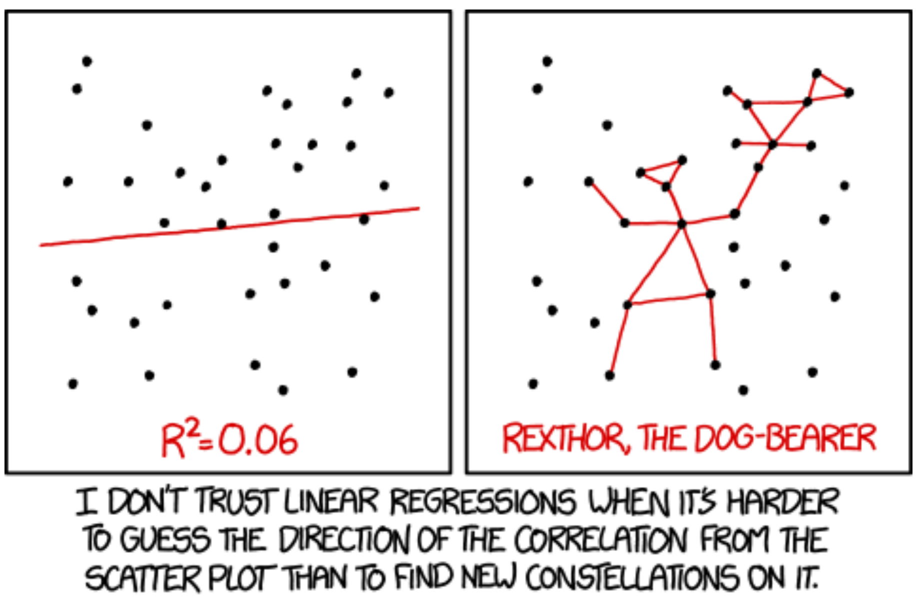
$= \frac{1}{\cancel{2}\sqrt{3}} \cdot \cancel{6} = \sqrt{3}$

$3 = \sqrt{3} \cdot \sqrt{3}$

Section 6.5 : Least-Squares Problems

Chapter 6 : Orthogonality and Least Squares

Math 1554 Linear Algebra



https://xkcd.com/1725

Topics and Objectives

Topics

1. Least Squares Problems
2. Different methods to solve Least Squares Problems

Learning Objectives

1. Compute general solutions, and least squares errors, to least squares problems using the normal equations and the QR decomposition.

Motivating Question A series of measurements are corrupted by random errors. How can the dominant trend be extracted from the measurements with random error?

Inconsistent Systems

Suppose we want to construct a line of the form

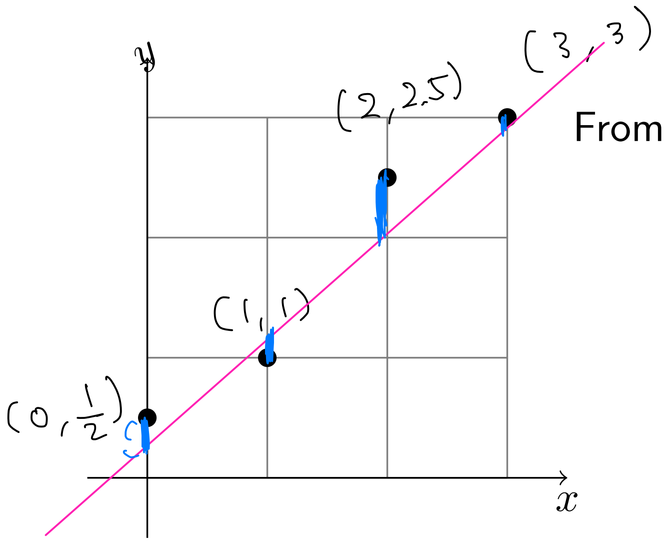
$$y = mx + b$$

Find m, b

best possible

that best fits the data below.

$(4, \underline{?})$



From the data, we can construct the system:

$$\begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} b \\ m \end{bmatrix} = \begin{bmatrix} 0.5 \\ 1 \\ 2.5 \\ 3 \end{bmatrix}$$

Can we 'solve' this inconsistent system?

$$\left.\begin{aligned} 0.5 = \tfrac{1}{2} &= m \cdot 0 + b \\ 1 &= m \cdot 1 + b \\ 2.5 &= m \cdot 2 + b \\ 3 &= m \cdot 3 + b \end{aligned}\right\}$$

The Least Squares Solution to a Linear System

Definition: Least Squares Solution

Let A be a $m \times n$ matrix. A **least squares solution to** $A\vec{x} = \vec{b}$ is the solution $\widehat{x}$ for which

$$\min_{\vec{x} \in \mathbb{R}^n} \|\vec{b} - A\vec{x}\| = \|\vec{b} - A\widehat{x}\| \leq \|\vec{b} - A\vec{x}\|$$

for all $\vec{x} \in \mathbb{R}^n$.

$A\vec{x} = \vec{b}$ If consistant

given

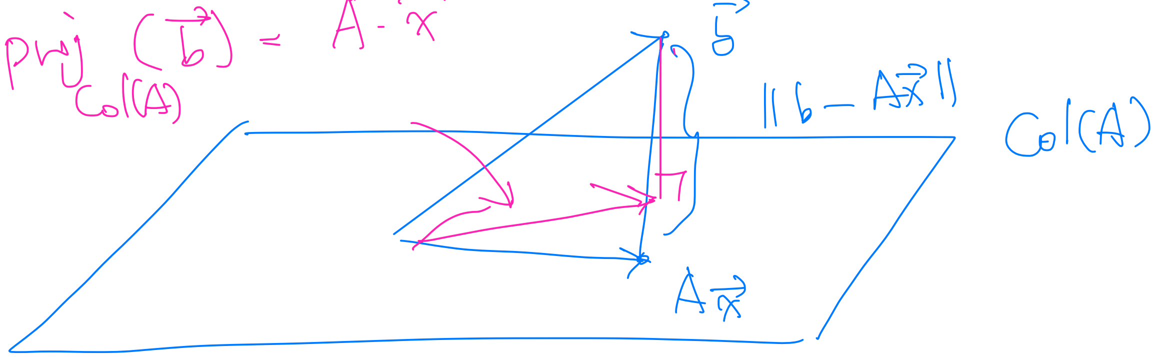
$$\Rightarrow \min_{\vec{x}} \|\vec{b} - A\vec{x}\| = 0$$

change

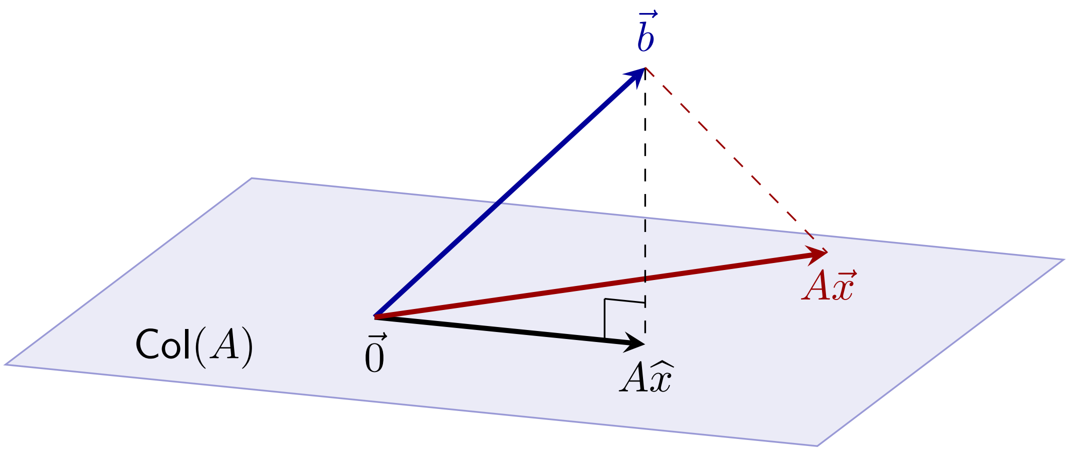
$\|\vec{b} - A\vec{x}\|$ as error

$A\vec{x} = \vec{b}$ Consistent $\iff$ $b \in \mathrm{Col}(A)$

$$\mathrm{proj}_{\mathrm{Col}(A)}(\vec{b}) = A \cdot \hat{x}$$



A Geometric Interpretation



The vector $\vec{b}$ is closer to $A\hat{x}$ than to $A\vec{x}$ for all other $\vec{x} \in$ Col$A$.

1. If $\vec{b} \in \mathrm{Col}\, A$, then $\widehat{x}$ is ...
2. Seek $\widehat{x}$ so that $A\widehat{x}$ is as close to $\vec{b}$ as possible. That is, $\widehat{x}$ should solve $A\widehat{x} = \widehat{b}$ where $\widehat{b}$ is ...

The Normal Equations

Theorem (Normal Equations for Least Squares)

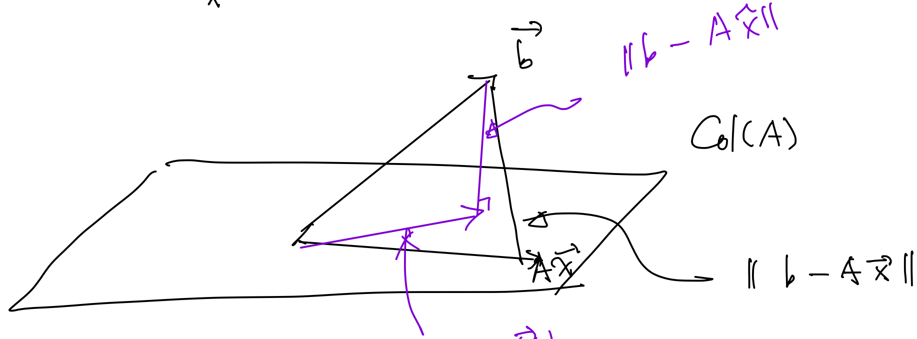
The least squares solutions to $A\vec{x} = \vec{b}$ coincide with the solutions to

$$\underbrace{A^T A\vec{x} = A^T\vec{b}}_{\text{Normal Equations}}$$

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

- $\hat{x}$ is a least-squares solution if

$$\|\vec{b} - A\hat{x}\| = \min_{\vec{x}} \|\vec{b} - A\vec{x}\|$$

- 

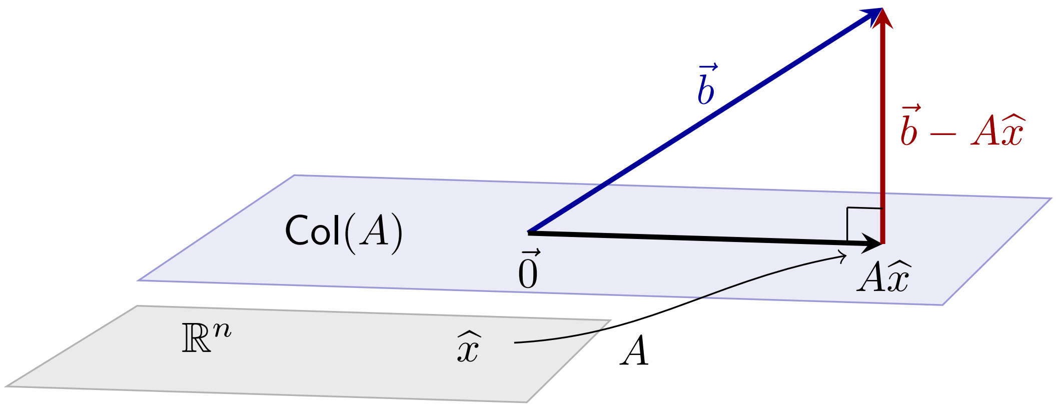
$A\hat{x} = proj_{Col(A)}(\vec{b})$

$\vec{b} - A\hat{x} \perp Col(A)$ $\quad$ $\vec{b} - A\hat{x} \in Col(A)^{\perp} = Null(A^T)$

$A^T \cdot (b - A\hat{x}) = 0$ $\quad$ always consistent

$A^T \cdot b - A^T A\hat{x} = 0 \Rightarrow A^T A\hat{x} = A^T \cdot \vec{b}$

Derivation



The least-squares solution $\hat{x}$ is in $\mathbb{R}^n$.

1. $\widehat{x}$ is the least squares solution, is equivalent to $\vec{b} - A\widehat{x}$ is orthogonal to $\boxed{}\, A$.
2. A vector $\vec{v}$ is in $\mathrm{Null}\, A^T$ if and only if $\boxed{}\, \vec{v} = \vec{0}$.
3. So we obtain the Normal Equations:

• B is symmetric if $B = B^T$

Example

Compute the least squares solution to $A\vec{x} = \vec{b}$, where

$$A = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix}, \qquad \vec{b} = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix}$$

$$\underbrace{A^T . A}\, x = \underbrace{A^T b}$$

Solution:

$$A^T A = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix} = \underbrace{\begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix}}$$

$$T = \begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix}$$ Symmetric

$$A^T \vec{b} = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix} = \begin{bmatrix} 19 \\ 11 \end{bmatrix}$$

$$\begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 19 \\ 11 \end{bmatrix}$$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \frac{1}{84} \begin{bmatrix} 5 & -1 \\ -1 & 17 \end{bmatrix} \begin{bmatrix} 19 \\ 11 \end{bmatrix}$$

$$= \frac{1}{84} \begin{bmatrix} 84 \\ 2 \cdot 84 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$ //

The normal equations $A^T A\vec{x} = A^T\vec{b}$ become:

$\underbrace{A\vec{x}}_{\text{lin. combi. of columns in } A} = \vec{b}$ consistent $\iff$ $b \in \mathrm{Col}(A)$

Note Why $A^TA x = A^Tb$ is consistent?

$A^T b \in \mathrm{Col}(A^T \cdot A) \iff A^TA x = A^Tb$ is consistent

$= \mathrm{Null}(A^T \cdot A)^\perp$

$n \times n$ $\quad m \times n$

$\iff \vec{x} \cdot (A^T \cdot b) = 0 \qquad \vec{x} \in \mathrm{Null}(A^T \cdot A) = \mathrm{Null}(A)$ later

$\iff (x^T \cdot A^T) \cdot b = (A \cdot x)^T \cdot b = 0$ if $\vec{x} \in \mathrm{Null}(A)$

$A \cdot x = 0$

Theorem

Theorem (Unique Solutions for Least Squares)

Let A be any $m \times n$ matrix. These statements are equivalent.

1. The equation $A\vec{x} = \vec{b}$ has a unique least-squares solution for each $\vec{b} \in \mathbb{R}^m$.
2. The columns of A are linearly independent. $\iff$ T is 1-1 $\iff$ $\mathrm{Null}(A) = \{0\}$ $= \mathrm{Null}(A^T \cdot A)$
3. The matrix A^TA is invertible. $\Updownarrow$ $A^T \cdot A$ has lin. indep columns

And, if these statements hold, the least square solution is

$$\boxed{\widehat{x} = (A^TA)^{-1}A^T\vec{b}.}$$

Useful heuristic: A^TA plays the role of 'length-squared' of the matrix A. (See the sections on symmetric matrices and singular value decomposition.)

$A^TA \, x = A^T \cdot b$

$A = Q R$ ← uppertriangular.

↑ has orthonormal Columns

Example

Compute the least squares solution to $A\vec{x} = \vec{b}$, where

$$A = \begin{bmatrix} 1 & -6 \\ 1 & -2 \\ 1 & 1 \\ 1 & 7 \end{bmatrix}, \qquad \vec{b} = \begin{bmatrix} -1 \\ 2 \\ 1 \\ 6 \end{bmatrix}$$

Hint: the columns of A are orthogonal.

$A^T \cdot A x = A^T \cdot b$

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ -6 & -2 & 1 & 7 \end{bmatrix} \begin{bmatrix} 1 & -6 \\ 1 & -2 \\ 1 & 1 \\ 1 & 7 \end{bmatrix} = \begin{bmatrix} 4 & 0 \\ 0 & 90 \end{bmatrix}$$

$$A^T \cdot b = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -6 & -2 & 1 & 7 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \\ 1 \\ 6 \end{bmatrix} = \begin{bmatrix} 8 \\ 45 \end{bmatrix}$$

$$\begin{bmatrix} 4 & 0 \\ 0 & 90 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 8 \\ 45 \end{bmatrix}$$

$x_1 = 2, \quad x_2 = \frac{1}{2}$ //

Theorem (Least Squares and QR)

Let $m \times n$ matrix A have a QR decomposition. Then for each $\vec{b} \in \mathbb{R}^m$ the equation $A\vec{x} = \vec{b}$ has the unique least squares solution

$$R\widehat{x} = Q^T\vec{b}.$$

(Remember, R is upper triangular, so the equation above is solved by back-substitution.)

Example 3. Compute the least squares solution to $A\vec{x} = \vec{b}$, where

$$A = \begin{bmatrix} 1 & 3 & 5 \\ 1 & 1 & 0 \\ 1 & 1 & 2 \\ 1 & 3 & 3 \end{bmatrix}, \qquad \vec{b} = \begin{bmatrix} 3 \\ 5 \\ 7 \\ -3 \end{bmatrix}$$

Solution. The QR decomposition of A is

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

$$\underline{A^T A \vec{x} = A^T b}.$$

$$R \cdot \hat{x} = Q^T \cdot b$$

$$Q^T\vec{b} = \tfrac{1}{2}\begin{bmatrix}1 & 1 & 1 & 1\\ 1 & -1 & -1 & 1\\ 1 & -1 & 1 & -1\end{bmatrix}\begin{bmatrix}3\\5\\7\\-3\end{bmatrix} = \begin{bmatrix}6\\-6\\4\end{bmatrix}$$

And then we solve by backwards substitution $R\vec{x} = Q^T\vec{b}$

$$\underbrace{\begin{bmatrix}2 & 4 & 5\\ 0 & 2 & 3\\ 0 & 0 & 2\end{bmatrix}}_{R}\begin{bmatrix}x_1\\ \cancel{x_2}\ -6\\ \cancel{x_3}\end{bmatrix} = \begin{bmatrix}6\\-6\\4\end{bmatrix}$$

2

$2x_3 = 4 \Rightarrow x_3 = 2$

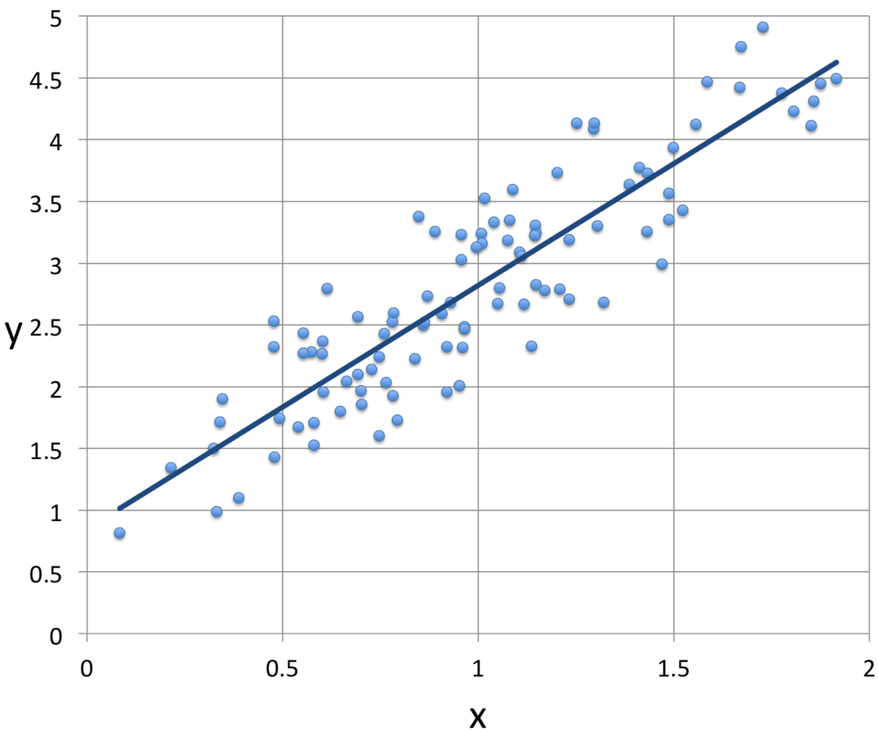
$2\cdot x_2 + 3\cdot 2 = -6 \Rightarrow x_2 = -6$

$2\cdot x_1 + 4\cdot(-6) + 5\cdot 2 = 6$

$2x_1 - 24 + 10 = 6 \Rightarrow x_1 = 10.$

$$\hat{x} = \begin{bmatrix}10\\-6\\2\end{bmatrix}$$

Chapter 6 : Orthogonality and Least Squares
6.6 : Applications to Linear Models



Topics and Objectives

Topics

1. Least Squares Lines
2. Linear and more complicated models

Learning Objectives

For the topics covered in this section, students are expected to be able to do the following.

1. Apply least-squares and multiple regression to construct a linear model from a set of data points.
2. Apply least-squares to fit polynomials and other curves to data.

Motivating Question

Compute the equation of the line $y = \beta_0 + \beta_1 x$ that best fits the data

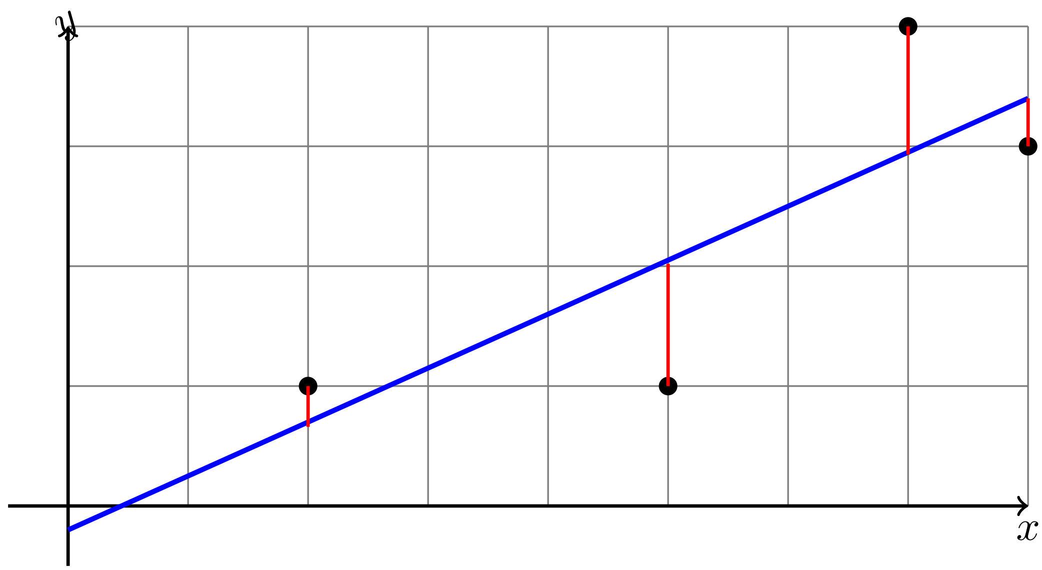
x	2	5	7	8
y	1	1	4	3

The Least Squares Line

Graph below gives an approximate linear relationship between x and y.

1. Black circles are data.
2. Blue line is the **least squares** line.
3. Lengths of red lines are the ____________.

The least squares line minimizes the sum of squares of the ____________.



Model

Example 1 Compute the least squares line $y = \beta_0 + \beta_1 x$ that best fits the data

x	2	5	7	8
y	1	1	4	3

← DATA. $x = 9$, $y = ?$

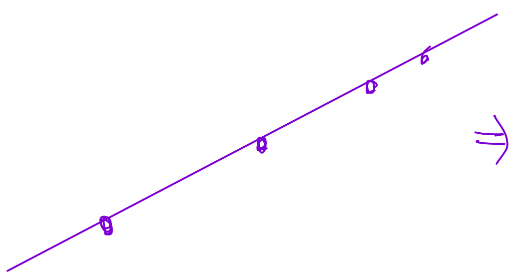
We want to solve

$$\begin{bmatrix} 1 & 2 \\ 1 & 5 \\ 1 & 7 \\ 1 & 8 \end{bmatrix} \underbrace{\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}}_{\vec\beta} = \underbrace{\begin{bmatrix} 1 \\ 1 \\ 4 \\ 3 \end{bmatrix}}_{\vec y}$$

unknown

$$\left\{\begin{aligned} 1 &= 1\cdot\beta_0 + \beta_1\cdot 2 \\ 1 &= 1\cdot\beta_0 + \beta_1\cdot 5 \\ 4 &= 1\cdot\beta_0 + \beta_1\cdot 7 \\ 3 &= 1\cdot\beta_0 + \beta_1\cdot 8 \end{aligned}\right.$$

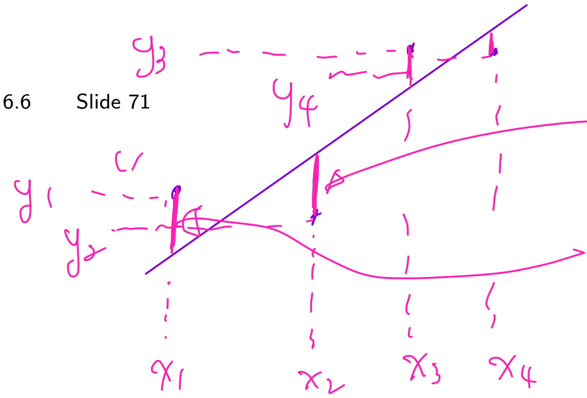
This is a least-squares problem : $X\vec\beta = \vec y$.



variable

$\Rightarrow$ $X\vec\beta = \vec y$ is consistent

Given



$y_2 - (\beta_0 + \beta_1 x_2)$

$y_1 - (\beta_0 + \beta_1 x_1)$

least-square solution

↓

$\hat\beta$

min

↓

$(y_1 - (\beta_0 + \beta_1 x_1))^2 + (y_2 - (\beta_0 + \beta_1 x_2))^2 + \cdots = \|\vec y - X\vec\beta\|^2$

$X\vec{\beta} = \vec{y}$

Normal Eqn : $X^T \cdot X \cdot \vec{\beta} = X^T \cdot \vec{y}$

The normal equations are

square symmetric.

$$X^TX = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 5 & 7 & 8 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 5 \\ 1 & 7 \\ 1 & 8 \end{bmatrix} = \begin{bmatrix} 4 & 22 \\ 22 & 142 \end{bmatrix}$$

$$X^T\vec{y} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 5 & 7 & 8 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 4 \\ 3 \end{bmatrix} = \begin{bmatrix} 9 \\ 59 \end{bmatrix}$$

So the least-squares solution is given by

$$\begin{bmatrix} 4 & 22 \\ 22 & 142 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 9 \\ 59 \end{bmatrix}$$

$$y = \beta_0 + \beta_1 x = \frac{-5}{21} + \frac{19}{42}x$$

As we may have guessed, β_0 is negative, and β_1 is positive.

Least Squares Fitting for Other Curves

We can consider least squares fitting for the form

$$y = c_0 + c_1 f_1(x) + c_2 f_2(x) + \cdots + c_k f_k(x).$$

If functions f_i are known, this is a linear problem in the c_i variables.

Example
Consider the data in the table below.

x	-1	0	0	1
y	2	1	0	6

Determine the coefficients c_1 and c_2 for the curve $y = c_1 x + c_2 x^2$ that best fits the data.

linear

$$\left\{\begin{aligned} 2 &= c_1(-1) + c_2(-1)^2 = -c_1 + c_2 \\ 1 &= c_1 \cdot 0 + c_2 \cdot 0^2 = 0 \cdot c_1 + 0 \cdot c_2 \\ 0 &= c_1 \cdot 0 + c_2 \cdot 0^2 = 0 \cdot c_1 + 0 \cdot c_2 \\ 6 &= c_1 \cdot 1 + c_2 \cdot 1^2 = c_1 + c_2 \end{aligned}\right.$$

$$\underbrace{\begin{bmatrix} 2 \\ 1 \\ 0 \\ 6 \end{bmatrix}}_{=\vec{y}} = \underbrace{\begin{bmatrix} -1 & 1 \\ 0 & 0 \\ 0 & 0 \\ 1 & 1 \end{bmatrix}}_{=X} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \longrightarrow = \vec{\beta}$$

$$X^T \cdot X \vec{\beta} = X^T \cdot \vec{y}$$

WolframAlpha and Mathematica Syntax

Least squares problems can be computed with WolframAlpha, Mathematica, and many other software.

WolframAlpha

$$\texttt{linear fit } \{\{x_1, y_1\}, \{x_2, y_2\}, \ldots, \{x_n, y_n\}\}$$

Mathematica

$$\texttt{LeastSquares}[\{\{x_1, x_1, y_1\}, \{x_2, x_2, y_2\}, \ldots, \{x_n, x_n, y_n\}\}]$$

Almost any spreadsheet program does this as a function as well.

A has lin. indep col.

$\Leftrightarrow$ $\mathrm{Null}(A) = \{0\}$

$\mathrm{Null}(A^T \cdot A)$ $\because$

$A\vec{x} = 0 \Rightarrow A^T A x = 0$

$\vec{x} \in \mathrm{Null}(A) \subseteq \mathrm{Null}(A^T A)$

$\supseteq$

$A^T \cdot A \cdot x = 0 \Rightarrow Ax = ?$

$\Downarrow$

$\|Ax\|^2 = (Ax) \cdot (Ax) = (Ax)^T (Ax) = x^T \cdot A^T \cdot A \cdot x = 0 \Rightarrow Ax = 0.$

$\Leftrightarrow$ $A^T \cdot A$ has lin. indep. columns $\Leftrightarrow$ $A^T A$ invertible

A has lin. indep. $\Leftrightarrow$ A^TA invertible

$\Leftrightarrow$ l.s.s is unique

Midterm 3. Your initials: _______________

8. (8 points) ***Show work*** *on this page with work under the problem, and* ***your answer in the box.***

In this problem, you will use the least-squares method to find the values α and β which best fit the curve

$$y = \alpha \cdot \frac{1}{1+x^2} + \beta$$

to the data points $(-1, 1), (0, -1), (1, 0)$ using the parameters α and β.

(i) What is the augmented matrix for the linear system of equations associated to this least squares problem?

$$1 = \alpha \cdot \frac{1}{1+(-1)^2} + \beta = \frac{1}{2}\alpha + \beta$$

$$-1 = \alpha \frac{1}{1+0^2} + \beta = \alpha + \beta \Rightarrow \underbrace{\begin{bmatrix} \frac{1}{2} & 1 \\ 1 & 1 \\ \frac{1}{2} & 1 \end{bmatrix}}_{=X} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \underbrace{\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}}_{=\vec{y}}$$

$$0 = \alpha \frac{1}{1+1^2} + \beta = \frac{1}{2}\alpha + \beta$$

$$\boxed{\left[\begin{array}{cc|c} \frac{1}{2} & 1 & 1 \\ 1 & 1 & -1 \\ \frac{1}{2} & 1 & 0 \end{array}\right]}$$

(ii) What is the augmented matrix for the normal equations for this system.

$$X^T X \hat{\beta} = X^T \cdot \vec{y}$$

$$X^T \cdot X = \begin{bmatrix} \frac{1}{2} & 1 & \frac{1}{2} \\ 1 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{2} & 1 \\ 1 & 1 \\ \frac{1}{2} & 1 \end{bmatrix} = \begin{bmatrix} \frac{3}{2} & 2 \\ 2 & 3 \end{bmatrix}$$

$$\boxed{\left[\begin{array}{cc|c} \frac{3}{2} & 2 & -\frac{1}{2} \\ 2 & 3 & 0 \end{array}\right]}$$

$$X^T \vec{y} = \begin{bmatrix} \frac{1}{2} & 1 & \frac{1}{2} \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} \\ 0 \end{bmatrix}$$

$$\left[\begin{array}{cc|c} 3 & 4 & -1 \\ 2 & 3 & 0 \end{array}\right]$$

$$\hat{\beta} = (X^T \cdot X)^{-1} \cdot X^T \vec{y} = \frac{1}{\frac{9}{2} - 4} \begin{bmatrix} 3 & -2 \\ -2 & \frac{3}{2} \end{bmatrix} \begin{bmatrix} -\frac{1}{2} \\ 0 \end{bmatrix}$$

(iii) Find a least-squares solution to the linear system from (i) to determine the parameters α and β of the best fitting curve.

$$= \begin{bmatrix} -3 \\ 2 \end{bmatrix}$$

$\alpha =$ -3 $\qquad \beta =$ 2

Midterm 3 Make-up. Your initials: ______________

You do not need to justify your reasoning for questions on this page.

1. (a) (6 points) Suppose A is a real $m \times n$ matrix and $\vec{b} \in \mathbb{R}^m$ unless otherwise stated. Select **true** if the statement is true for all choices of A and $\vec{b}$. Otherwise, select **false**.

true	false	
○	○	For any line $L \in \mathbb{R}^2$ passing through the origin, the matrix corresponding to the transformation that reflects across the line L must always be diagonalizable.
○	○	If A and B are $n \times n$ orthogonal matrices, then AB is also $n \times n$ and orthogonal.
○	○	If A is the reduced row echelon form (RREF) of B and A is diagonalizable, then B is diagonalizable.
○	○	If $\vec{b} \in \text{Col}(A)$, then the least squares solution to the linear system $A\vec{x} = \vec{b}$ is unique.
●	○	For any rectangular $m \times n$ matrix A, $(\text{Nul}A)^\perp = \text{Row}(A^T A.)$
○	○	If the distance of $\vec{w}$ from $\vec{v}$ is equal to the distance of $\vec{w}$ from $-\vec{v}$, then $\vec{w} \cdot \vec{v} = 0$.

$\text{Null}(A^T \cdot A)^\perp$

$\text{Null}(A) = \text{Null}(A^T \cdot A)$.

(b) (2 points) Indicate whether the following situations are possible or impossible.

possible	impossible	
○	○	A diagonal matrix A that is similar to $\begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}$.
○	○	An orthogonal matrix A such that $\lvert \det A \rvert \neq 1$.

8. (4 points) **Show all work for problems on this page.** If $A = QR = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 2 & 2 \\ 0 & 1 \end{pmatrix}$, determine the least-squares solution to $A\hat{x} = \begin{pmatrix} \sqrt{2} \\ 2\sqrt{2} \end{pmatrix}$. You do not need to determine A.

$\hat{x}=$ $\boxed{}$