Lecture XIX : §3.5 Rank-Nullity theorem
§3.6 Orthonormal bases

Last Time:
- W subspace of $\mathbb{R}^n$ $\rightsquigarrow$ $\dim(W)$ = size of any basis for W
- $W = Sp(\vec{v}_1, \ldots, \vec{v}_m)$ $\rightsquigarrow$ Basis $= \{v_i : x_i$ is a dependent variable of $A\vec{x} = \vec{0}\}$ for $A = [\vec{v}_1 \cdots \vec{v}_m]$ (METHOD 3)

§1. Rank and Nullity:

Fix a matrix A of size $n \times m$

Def: $\text{nullity}(A) = \dim \mathcal{N}(A)$

$\text{rank}(A) = \dim \mathcal{R}(A)$

Q How to compute these numbers?

(1) For Nullity: Solve the system via Gauss Jordan

$\rightsquigarrow$ $\text{nullity}(A) = \#$ indep variables of $A\vec{x} = \vec{0}$.

(2) For Rank: Method 3 says $\#$ of dep. variables = size of a basis for $\mathcal{R}(A)$.

Consequence: Rank-Nullity Theorem

$$\underset{(\#\text{ indep vars})}{\text{nullity}(A)} + \underset{(\#\text{ dep vars})}{\text{rank}(A)} = \underset{(\text{total }\#\text{ of vars})}{\#\text{ columns of } A}.$$

Example: $A = \begin{bmatrix} 1 & 2 & 3 & -1 \\ 2 & 3 & 5 & -1 \\ 1 & 5 & 6 & -4 \end{bmatrix} \xrightarrow{\text{GAUSS-JORDAN}} \text{REF}(A) = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ x_1, x_2 dep

- Basis for Range(A) $= \left\{ \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix} \right\}$ $\rightsquigarrow$ $\text{rank}(A) = 2$.
- To get a basis for $\mathcal{N}(A)$, we write the general form of a solution:

$x_1 + x_3 + x_4 = 0 \rightsquigarrow x_1 = -x_3 - x_4$

$x_2 + x_3 - x_4 = 0 \rightsquigarrow x_2 = -x_3 + x_4$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = x_3 \begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -1 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

li
(complementary location of 0's)

$\rightsquigarrow$ Basis for $\mathcal{N}(A) = \left\{ \begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \\ 1 \end{bmatrix} \right\}$ so nullity $(A) = 2$

Rank - Nullity : $2 + 2 = 4$ ✓

Q What about the old definition of rank? (It looks different!)

. $\text{Oldrank}(A) = \text{Oldrank}(REF(A)) = \#$ non-zero rows of $REF(A)$
$= \text{Rank}(A)$

BUT...

- $\#$ non-zero rows of $REF(A)$ = size of a basis of $\underbrace{\text{RowSp}(REF(A))}_{= \text{RowSp}(A)}$
$= \dim \text{RowSp}(A)$
$= \dim R(A^T) = \text{rank}(A^T)$

Consequence: $\text{Rank}(A) = \text{Rank}(A^T)$, ie Column & row space have the same dimension!

Example above: $A = \begin{bmatrix} 1 & 2 & 3 & -1 \\ 2 & 3 & 5 & -1 \\ 1 & 5 & 6 & -4 \end{bmatrix}$ $\quad \dim \text{RowSp}(A) = 2 = \dim \text{ColSp}(A^T) = \text{Rank}(A^T)$ by earlier calculation

$A^T = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 3 & 5 \\ 3 & 5 & 6 \\ -1 & -1 & -4 \end{bmatrix} \xrightarrow{G-J} \begin{bmatrix} 1 & 0 & 7 \\ 0 & 1 & -3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = REF(A^T)$ $\quad \dim \text{RowSp}(A^T) = 2 = \dim \text{ColSp}(A)$
$= \text{Rank}(A)$

Example: $A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 0 & 1 \end{bmatrix}$ nullity = ?

$\begin{bmatrix} 1 & 2 & 3 \\ 1 & 0 & 1 \end{bmatrix} \xrightarrow[R_1 \leftrightarrow R_2]{} \begin{bmatrix} 1 & 0 & 1 \\ 1 & 2 & 3 \end{bmatrix} \xrightarrow[R_2 \to R_2 - R_1]{} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 2 \end{bmatrix} \xrightarrow[R_2 \to R_2/2]{} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$ x_1, x_2 dep

$\mathcal{N}(A) = \text{Sp}\left(\begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} \right)$ so nullity $= 1$ $\rightsquigarrow$ rank$(A) = 3 - 1 = 2$.

. rank$(A) = \dim R(A)$ & $R(A)$ is a subspace of $\mathbb{R}^2$, also of dim 2.

Conclude: $\mathcal{R}(A) = \mathbb{R}^2$ (same dimension & $\mathcal{R}(A) \subseteq \mathbb{R}^2$)

- $\operatorname{rank}(A^T) = \operatorname{rank}(A) = 2$ so $\operatorname{RowSp}(A)$ has dim 2 as well.

§2. Two applications:

Theorem 1: Fix an $m\times n$ matrix A & $\vec{b}$ in $\mathbb{R}^m$. The system $A\vec{x} = \vec{b}$ is consistent if, and only if, $\operatorname{Rank}(A) = \operatorname{Rank}(A|b)$.

Why? Consistent means $\vec{b}$ is in $\mathcal{R}(A) = \operatorname{Col}\operatorname{Sp}(A)$ ($x_1 \operatorname{Col}_1 A + \cdots + x_n \operatorname{Col}_n A = \vec{b}$ for some numbers $x_1, \ldots, x_n$). This happens if and only if $\operatorname{ColSp}(A) = \operatorname{ColSp}(A|b)$.
The ranks are the dimension of each of these spaces.

Consequence: An $n\times n$ matrix is nonsingular if and only if $\operatorname{rank}(A) = n$.

Why? Nonsingular means $\operatorname{nullity}(A) = 0$ so $\operatorname{rank}(A) = \#\operatorname{cols}(A) - \operatorname{null}(A)$
$= n - 0 = n$.

§3. Orthogonal & orthonormal bases:

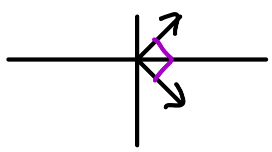
Recall the formula for the dot product:

if $\vec{u} = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}$, $\vec{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$ are vectors in $\mathbb{R}^n$ ($n\times 1$ matrices) then

$$\vec{u}\cdot\vec{v} = \vec{u}^T\vec{v} = u_1v_1 + u_2v_2 + \cdots + u_nv_n \quad (\text{number})$$

- We say $\vec{u}$ & $\vec{v}$ are orthogonal ($\vec{u} \perp \vec{v}$) if $\vec{u}\cdot\vec{v} = 0$

Eg:
- $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ & $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ are orthogonal



- $\vec{0} \perp \vec{v}$ for any $\vec{v}$

- Fix a subspace W of $\mathbb{R}^n$ with $W \neq \{\vec{0}\}$ & let $\mathcal{B} = \{\vec{v}_1, \ldots, \vec{v}_p\}$ be a basis for W.

Def: • We say B is an <u>orthogonal basis</u> for W if $\vec{v}_i \perp \vec{v}_j$ for every $i \neq j$.
(that is, any 2 vectors in B are orthogonal to each other)

• We say B is an <u>orthonormal basis</u> for W if it is an orthogonal basis and $v_i \cdot v_i = 1$ for every $i = 1, \ldots, p$.
(This means $\|v_i\| = \sqrt{1} = 1$ for every i)

<u>Main use of orthogonality</u>: It implies linear independence.

<u>Theorem</u>: Fix a set $S = \{\vec{v}_1, \ldots, \vec{v}_p\}$ in $\mathbb{R}^n$ and assume $\vec{0}$ is <u>NOT</u> in S. If all vectors of S are orthogonal to each other, then S is linearly independent.

<u>Why?</u> Write a dependency relation:

$$\vec{0} = a_1 \vec{v}_1 + a_2 \vec{v}_2 + \cdots + a_p \vec{v}_p$$

& take dot product with $\vec{v}_1$. We can distribute & get

$$0 = \vec{0} \cdot \vec{v} = a_1 \underbrace{\vec{v}_1 \cdot \vec{v}_1}_{\substack{= \|\vec{v}_1\|^2 \neq 0 \\ \text{because } \vec{v}_1 \neq \vec{0}}} + a_2 \underbrace{\vec{v}_2 \cdot \vec{v}_1}_{=0} + \cdots + a_p \underbrace{\vec{v}_p \cdot \vec{v}_1}_{=0}$$

<u>Conclude</u> $0 = a_1 \underbrace{\|\vec{v}_1\|^2}_{\neq 0}$ so $a_1 = 0$.

Same idea shows $a_2 = a_3 = \cdots = a_p = 0$. This says S is l.i.
$(\cdot \vec{v}_2, \cdot \vec{v}_3, \ldots, \cdot \vec{v}_p)$

<u>Example:</u> $\vec{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$, $\vec{v}_2 = \begin{bmatrix} 3 \\ -1 \\ -1 \end{bmatrix}$, $\vec{v}_3 = \begin{bmatrix} 1 \\ -4 \\ 7 \end{bmatrix}$

$\vec{v}_1 \cdot \vec{v}_2 = 1 \cdot 3 - 2 - 1 = 0$

$\vec{v}_1 \cdot \vec{v}_3 = 1 - 8 + 7 = 0$

$\vec{v}_2 \cdot \vec{v}_3 = 3 + 4 - 7 = 0$

$S = \{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ is an orthogonal set hence it's linearly independent

This set has 3 l.i vectors in $\mathbb{R}^3$, so it's a basis for $\mathbb{R}^3$.

- If $\vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$ is an arbitrary vector in $\mathbb{R}^3$, we can write.

$$\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = x_1 \vec{v}_1 + x_2 \vec{v}_2 + x_3 \vec{v}_3 .$$

- Here is a quick way to find x_1, x_2, x_3:

  Take dot product with $\vec{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$

$$b_1 + 2b_2 + b_3 = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} = x_1 \underbrace{\vec{v}_1 \cdot \vec{v}_1}_{= \|\vec{v}_1\|^2 = 1+4+1=6} + x_2 \underbrace{\vec{v}_2 \cdot \vec{v}_1}_{=0} + x_3 \underbrace{\vec{v}_3 \cdot \vec{v}_1}_{=0}$$

Conclude: $x_1 = \dfrac{b_1 + 2b_2 + b_3}{6}$

Similarly using $\cdot \vec{v}_2$ & $\cdot \vec{v}_3$ gives x_2 & x_3:

$$x_2 = \frac{3b_1 - b_2 - b_3}{11} \quad \& \quad x_3 = \frac{b_1 - 4b_2 + 7b_3}{66}$$

- The set $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ is not an orthonormal basis for $\mathbb{R}^3$, but we can scale these vectors to have length 1 & get an orthonormal basis.

$$\vec{u}_1 = \frac{\vec{v}_1}{\|\vec{v}_1\|} = \frac{\vec{v}_1}{\sqrt{6}}, \quad \vec{u}_2 = \frac{\vec{v}_2}{\|\vec{v}_2\|} = \frac{\vec{v}_2}{\sqrt{11}}, \quad \vec{u}_3 = \frac{\vec{v}_3}{\|\vec{v}_3\|} = \frac{\vec{v}_3}{\sqrt{66}}$$

$\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ is an orthonormal basis for $\mathbb{R}^3$.

§3 Coordinates with respect to orthogonal bases.

The last example highlights that if $B = \{\vec{v}_1, \dots, \vec{v}_p\}$ is an orthogonal basis for a subspace W of $\mathbb{R}^n$, then the problem of writing an arbitrary vector $\vec{w}$ of W as a linear comb of $\vec{v}_1, \dots, \vec{v}_p$ is very easy!

Theorem: If $B = \{\vec{v}_1, \dots, \vec{v}_p\}$ is an orthogonal basis for a subspace W of $\mathbb{R}^n$, a $\vec{w}$ is a vector in W, then:

$$\vec{w} = x_1 \vec{v}_1 + \dots + x_p \vec{v}_p$$

where $x_1 = \dfrac{\vec{v}_1 \cdot \vec{w}}{\vec{v}_1 \cdot \vec{v}_1}$, $x_2 = \dfrac{\vec{v}_2 \cdot \vec{w}}{\vec{v}_2 \cdot \vec{v}_2}$, ..., $x_p = \dfrac{\vec{v}_p \cdot \vec{w}}{\vec{v}_p \cdot \vec{v}_p}$.

In short $[\vec{w}]_B = \begin{bmatrix} \vec{v}_1 \cdot \vec{w} / \|\vec{v}_1\|^2 \\ \vec{v}_2 \cdot \vec{w} / \|\vec{v}_2\|^2 \\ \vdots \\ \vec{v}_p \cdot \vec{w} / \|\vec{v}_p\|^2 \end{bmatrix}$ (vector in $\mathbb{R}^p$)

Q: How to find an orthogonal basis? **A** Gram-Schmidt Algorithm (next time!)