Lecture 17

(17.0) Quick review of material so far. Let $N \geq 1$ be a positive integer.

- A subset $V \subset \mathbb{R}^N$ is a subspace if (c1) $\vec{0} \in V$, (c2) For $\vec{u}, \vec{v} \in V$, we have $\vec{u}+\vec{v} \in V$, (c3) for $\vec{u} \in V$ and $c \in \mathbb{R}$; we have $c.\vec{u} \in V$.

- Two main examples of subspaces: null space and range of a matrix.

A : (m×n) matrix

$$\mathcal{N}(A) = \{\vec{x} \in \mathbb{R}^n : A\vec{x} = \vec{0}\} \subset \mathbb{R}^n.$$
(null space)

$$\mathcal{R}(A) = \{\vec{y} \in \mathbb{R}^m : A\vec{x} = \vec{y} \text{ for some } \vec{x} \in \mathbb{R}^n\} \subset \mathbb{R}^m.$$
(range)

$$= \{\vec{b} \in \mathbb{R}^m : [A|\vec{b}] \text{ is consistent}\}$$

- Range of a matrix A is also called the column span of A. In general if $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p$ are vectors in $\mathbb{R}^N$, we define

$$\text{Span of } \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\} = \left\{a_1\vec{v}_1 + a_2\vec{v}_2 + \ldots + a_p\vec{v}_p \text{ where } a_1, a_2, \ldots, a_p \text{ are arbitrary real numbers}\right\}$$

$\mathcal{R}(A) = \text{Span}(\text{column vectors in } A)$.

- Basis of a subspace $B = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\} \subset V$ (a subspace of $\mathbb{R}^N$) is a basis of V if • Span(B) = V & • B is linearly independent.

- Dimension of V dim(V) = number of vectors in a basis of V.

$\boxed{\text{Convention}}$ If $V = \{\vec{0}\}$, its only basis is the empty set and its dimension is zero (= # elements in the empty set).

- If A is an $m\times n$ matrix, then <u>$\dim(\mathcal{N}(A))$ is called nullity of A</u> ; and <u>$\dim(\mathcal{R}(A))$ is called rank of A</u>.

<u>Rank- Nullity Theorem</u>

$$\boxed{\text{Nullity of } A + \text{Rank of } A = n \quad (= \text{Number of Columns of } A)}$$

$$\boxed{\text{Rank}(A) = \text{Rank}(A^T)} \longrightarrow \boxed{\begin{array}{l}\text{Rank}(A) \leqslant m \text{ and} \\ \text{Rank}(A) \leqslant n\end{array}}$$

(17.1) How to compute a basis of $\mathcal{N}(A)$ and $\mathcal{R}(A)$.

Let A be an $m\times n$ matrix. Last time we discussed how we compute a basis of $R(A)$. Let us talk about $\mathcal{N}(A)$ (<u>easier</u>).

$$\mathcal{N}(A) = \{\vec{x} \in \mathbb{R}^n \text{ such that } A\vec{x} = \vec{0}\}$$

- Solve the homogeneous system $A\vec{x} = \vec{0}$ and write the set of solutions in vector form.
- If there are l free parameters, the vectors which feature (l of them) in your answer <u>form a basis of $\mathcal{N}(A)$</u>.

$$\boxed{\dim \mathcal{N}(A) = \#\text{ free parameters in the system } A\vec{x} = \vec{0} = \text{Nullity}(A)}$$

As we saw in Method II of the last lecture

$$\boxed{\dim \mathcal{R}(A) = n - \#\text{ free parameters} = \text{Rank}(A)} \leftarrow \text{Rank-Nullity Theorem}$$

Example. $A = \begin{bmatrix} 1 & 0 & 1 & -1 \\ 2 & 1 & 3 & 2 \\ 0 & 2 & 2 & 8 \end{bmatrix}$. Find a basis of $\mathcal{N}(A)$ and a basis of $\mathcal{R}(A)$.

$$\mathcal{N}(A) = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \text{ such that } A \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \right\} \subset \mathbb{R}^4.$$

Put A in (reduced) echelon form to solve the hgs system $A\vec{x} = \vec{0}$.

$$\begin{bmatrix} 1 & 0 & 1 & -1 \\ 2 & 1 & 3 & 2 \\ 0 & 2 & 2 & 8 \end{bmatrix} \xrightarrow[R_2 \to R_2 - 2R_1]{} \begin{bmatrix} 1 & 0 & 1 & -1 \\ 0 & 1 & 1 & 4 \\ 0 & 2 & 2 & 8 \end{bmatrix} \xrightarrow[R_3 \to R_3 - 2R_2]{} \begin{bmatrix} 1 & 0 & 1 & -1 \\ 0 & 1 & 1 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

So $x_1 + x_3 - x_4 = 0$ $\qquad x_1 = -x_3 + x_4$

$x_2 + x_3 + 4x_4 = 0$ $\qquad x_2 = -x_3 - 4x_4$

Free parameters: x_3, x_4

$$\text{Solution set} = \begin{bmatrix} -x_3 + x_4 \\ -x_3 - 4x_4 \\ x_3 \\ x_4 \end{bmatrix} = x_3 \begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 1 \\ -4 \\ 0 \\ 1 \end{bmatrix} \qquad x_3, x_4 \in \mathbb{R} \text{ arbitrary}$$

Basis of $\mathcal{N}(A) = \left\{ \begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ -4 \\ 0 \\ 1 \end{bmatrix} \right\}$. Nullity$(A) = 2$.

Method II for finding a basis of $\mathcal{R}(A) = \text{Span}\left\{ \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \\ 8 \end{bmatrix} \right\}$

says: Basis of $\mathcal{R}(A) = \left\{ \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} \right\}$

Columns of A where $RE(A)$ (reduced echelon form) has pivotal 1's

Rank$(A) = 2$

Rank + Nullity $= 4 = \#$ Columns of A ✓

(17.2) Orthogonal and orthonormal bases.

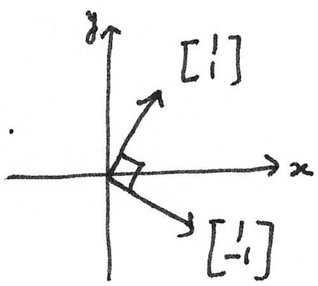
Recall the formula for the dot product: $\vec{u}, \vec{v}$ two vectors in $\mathbb{R}^n$ (i.e. $n \times 1$ matrices)

$\underset{(1\times n)}{\vec{u}^T}\ \underset{(n\times 1)}{\vec{v}} \in \mathbb{R}$ is called the dot product of $\vec{u}$ and $\vec{v}$.

If $\vec{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$ and $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$ then $\vec{u}^T\vec{v} = u_1v_1 + u_2v_2 + \dots + u_nv_n$
$= \vec{v}^T\vec{u}$.

We say $\vec{u} \perp \vec{v}$ ($\vec{u}$ and $\vec{v}$ are orthogonal to each other) if
$\boxed{\vec{u}^T\vec{v} = 0}$.

e.g. • $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ are orthogonal to each other.



• $\vec{0} \perp \vec{v}$ for any vector $\vec{v}$.

Let $V \subset \mathbb{R}^n$ be a subspace, $V \neq \{\vec{0}\}$, and let
$B = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p\}$ be a basis of V.

We say B is an orthogonal basis of V if $\vec{v}_i \perp \vec{v}_j$ for every $i \neq j$.

That is, any two vectors from B are orthogonal to each other.

We say B is an orthonormal basis of V if it is an orthogonal basis and $v_i^T v_i = 1$ for every $i = 1, 2, \dots, p$.

Recall: $\|\vec{v}\|^2 = \vec{v}^T\vec{v}$. So every vector $\vec{v}_i$ in an orthonormal basis must have length 1.

(17.3) Main use of orthogonality is that it implies linear independence.

Theorem. Let $S = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p\} \subset \mathbb{R}^n$. Assume $\vec{0}$ is <u>NOT</u> in S and any two vectors from S are orthogonal to each other. (i.e., $\vec{v}_i^{\,T} v_j = 0$ for every $i \neq j$).

Then S is linearly independent.

The idea is that if $\boxed{a_1\vec{v}_1 + a_2\vec{v}_2 + \dots + a_p\vec{v}_p = \vec{0}}$, then multiplying this equation on the left by $\vec{v}_1^{\,T}$ gives:

$$\underbrace{a_1(\vec{v}_1^{\,T}\vec{v}_1)}_{=\|\vec{v}_1\|^2} + a_2\underbrace{(\vec{v}_1^{\,T}\vec{v}_2)}_{=0} + \dots + a_p\underbrace{(\vec{v}_1^{\,T}\vec{v}_p)}_{=0} = 0$$

That is, $a_1 \cdot \|\vec{v}_1\|^2 = 0$. As $\|\vec{v}_1\|^2 \neq 0$ (remember: $\vec{0}$ is <u>NOT</u> in S) we get $\boxed{a_1 = 0}$. Similarly we can get $\boxed{a_2 = a_3 = \dots = a_p = 0}$.

So, $a_1\vec{v}_1 + a_2\vec{v}_2 + \dots + a_p\vec{v}_p = \vec{0}$ implies $a_1 = a_2 = a_3 = \dots = a_p = 0$, which means precisely that $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p\}$ is <u>linearly independent</u>.

(17.4) Example. $\vec{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$, $\vec{v}_2 = \begin{bmatrix} 3 \\ -1 \\ -1 \end{bmatrix}$, $\vec{v}_3 = \begin{bmatrix} 1 \\ -4 \\ 7 \end{bmatrix}$.

$\vec{v}_1^{\,T}\vec{v}_2 = 1(3) + 2(-1) + 1(-1) = 0$

$\vec{v}_1^{\,T}\vec{v}_3 = 1(1) + 2(-4) + 1(7) = 0$

$\vec{v}_2^{\,T}\vec{v}_3 = 3(1) + (-1)(-4) + (-1)(7) = 3 + 4 - 7 = 0$.

So $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ is an <u>orthogonal set</u>, hence linearly independent.

This set also spans $\mathbb{R}^3$. (Why?) Hence, a basis of $\mathbb{R}^3$.

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

write $\vec{b} = x_1 \vec{v}_1 + x_2 \vec{v}_2 + x_3 \vec{v}_3$. Here is a quick way of figuring out x_1, x_2, x_3. Take dot product with $\vec{v}_1$:

$$\underbrace{\vec{v}_1^{\,T} \vec{b}}_{b_1 + 2b_2 + b_3} = x_1 \underbrace{(\vec{v}_1^{\,T} \vec{v}_1)}_{= \underbrace{\|\vec{v}_1\|^2}_{1^2+2^2+1^2}} + x_2 \underbrace{(\vec{v}_1^{\,T} \vec{v}_2)}_{=0} + x_3 \underbrace{(\vec{v}_1^{\,T} \vec{v}_3)}_{=0}$$

$$b_1 + 2b_2 + b_3 = x_1 (1^2 + 2^2 + 1^2) \equiv \boxed{x_1 = \frac{b_1 + 2b_2 + b_3}{6}}.$$

— similarly you can get x_2 and x_3 by taking dot product with $\vec{v}_2$, $\vec{v}_3$.

This set $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ is <u>NOT</u> orthonormal basis of $\mathbb{R}^3$.

but we can scale these vectors to have length 1 and get an orthonormal basis.

$$\vec{u}_1 = \frac{1}{\|\vec{v}_1\|} \cdot \vec{v}_1 = \frac{1}{\sqrt{6}} \vec{v}_1$$

$$\vec{u}_2 = \frac{1}{\|\vec{v}_2\|} \vec{v}_2 = \frac{1}{\sqrt{11}} \vec{v}_2 \quad \left(\|\vec{v}_2\|^2 = 3^2 + (-1)^2 + (-1)^2 = 11 \right)$$

$$\vec{u}_3 = \frac{1}{\|\vec{v}_3\|} \vec{v}_3 = \frac{1}{\sqrt{66}} \vec{v}_3 \quad \left(\|\vec{v}_3\|^2 = 1^2 + (-4)^2 + 7^2 = 1 + 16 + 49 = 66 \right)$$

orthonormal basis of $\mathbb{R}^3$.

(17.5) Last example highlights that if $B = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}$ is an orthogonal basis of a subspace $V \subset \mathbb{R}^n$, then the problem of writing an arbitrary vector $\vec{w} \in V$, as a linear combination of $\vec{v}_1, \ldots, \vec{v}_p$ is very easy:

$$\vec{w} = x_1\vec{v}_1 + x_2\vec{v}_2 + \ldots + x_p\vec{v}_p \xrightarrow[\text{with } \vec{v}_i]{\text{dot prod.}} \vec{v}_i^{\,T}\vec{w} = x_i\left(\vec{v}_i^{\,T}\vec{v}_i\right)$$

Problem: find $x_1, \ldots, x_p$ so that

$$\boxed{x_i = \frac{\vec{v}_i^{\,T}\vec{w}}{\vec{v}_i^{\,T}\vec{v}_i}} \quad i = 1, 2, \ldots, p.$$

Solution

(17.6) Constructing an orthogonal basis. (Gram-Schmidt)

Input: A basis $\{\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_p\}$ of a subspace $V \subset \mathbb{R}^n$.

Output: An orthogonal basis $\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_p\}$ of V.

Procedure: (Gram-Schmidt)

$$\vec{u}_1 = \vec{w}_1.$$

$$\vec{u}_2 = \vec{w}_2 - \mathrm{Proj}_{\vec{u}_1}(\vec{w}_2) = \vec{w}_2 - \underbrace{\left(\frac{\vec{u}_1^{\,T}\vec{w}_2}{\vec{u}_1^{\,T}\vec{u}_1}\right)\vec{u}_1}_{\text{formula for projection.}}$$

$$\vec{u}_3 = \vec{w}_3 - \mathrm{Proj}_{\vec{u}_1}(\vec{w}_3) - \mathrm{Proj}_{\vec{u}_2}(\vec{w}_3)$$

$$= \vec{w}_3 - \left(\frac{\vec{u}_1^{\,T}\vec{w}_3}{\vec{u}_1^{\,T}\vec{u}_1}\right)\vec{u}_1 - \left(\frac{\vec{u}_2^{\,T}\vec{w}_3}{\vec{u}_2^{\,T}\vec{u}_2}\right)\vec{u}_2$$

... continue this way (it is ~~a recursive~~ an iterative procedure)

Assume we have built up to $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_K$.

$$\vec{u}_{K+1} = \vec{w}_{K+1} - \left(\frac{\vec{u}_1^{\,T}\vec{w}_{K+1}}{\vec{u}_1^{\,T}\vec{u}_1}\right)\vec{u}_1 - \left(\frac{\vec{u}_2^{\,T}\vec{w}_{K+1}}{\vec{u}_2^{\,T}\vec{u}_2}\right)\vec{u}_2$$

$$- \cdots - \left(\frac{\vec{u}_K^{\,T}\vec{w}_{K+1}}{\vec{u}_K^{\,T}\vec{u}_K}\right)\vec{u}_K$$

Example. $V = \text{Span}\left\{\underbrace{\begin{bmatrix}1\\1\\2\end{bmatrix}, \begin{bmatrix}0\\2\\4\end{bmatrix}}\right\} \subset \mathbb{R}^3.$ this set is linearly independent

$\vec{w}_1 = \begin{bmatrix}1\\1\\2\end{bmatrix}$ $\vec{w}_2 = \begin{bmatrix}0\\2\\4\end{bmatrix}$. Let us construct an <u>orthonormal basis</u>.

(Gram-Schmidt) $\vec{u}_1 = \vec{w}_1 = \begin{bmatrix}1\\1\\2\end{bmatrix}$ ✓

$$\vec{u}_2 = \vec{w}_2 - \text{Proj}_{\vec{u}_1}(\vec{w}_2) = \vec{w}_2 - \left(\frac{\vec{u}_1^{\,T}\vec{w}_2}{\vec{u}_1^{\,T}\vec{u}_1}\right)\vec{u}_1$$

$$= \begin{bmatrix}0\\2\\4\end{bmatrix} - \left(\frac{0+2+8}{1^2+1^2+2^2}\right)\begin{bmatrix}1\\1\\2\end{bmatrix} = \begin{bmatrix}0\\2\\4\end{bmatrix} - \frac{10}{6}\begin{bmatrix}1\\1\\2\end{bmatrix}$$

So $\vec{u}_2 = \begin{bmatrix}-\frac{5}{3}\\ 2-\frac{5}{3}\\ 4-\frac{10}{3}\end{bmatrix} = \begin{bmatrix}-5/3\\1/3\\2/3\end{bmatrix}$. Check: $\vec{u}_1^{\,T}\vec{u}_2 = 0$:

$$1\left(\frac{-5}{3}\right) + 1\left(\frac{1}{3}\right) + 2\left(\frac{2}{3}\right) = \frac{-5+1+4}{3} = 0 \checkmark.$$

$\left\{\begin{bmatrix}1\\1\\2\end{bmatrix}, \begin{bmatrix}-5/3\\1/3\\2/3\end{bmatrix}\right\}$: <u>orthogonal</u> bases of V.

To make it orthonormal, divide by $\|\cdot\|$: $\left\{\frac{1}{\sqrt{6}}\begin{bmatrix}1\\1\\2\end{bmatrix}, \frac{1}{\sqrt{30}}\begin{bmatrix}-5\\1\\2\end{bmatrix}\right\}$

<u>orthonormal basis of V</u>.