Lecture 20: §3.7 (cont) Matrix representations of linear maps

Recall: A linear transformation $T: \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is a map satisfying:

$\vec{v} \longmapsto T(\vec{v})$

(1) $T(\vec{u}+\vec{v}) = T(\vec{u}) + T(\vec{v})$ for all $\vec{u}, \vec{v}$ in $\mathbb{R}^n$

(2) $T(\alpha\vec{u}) = \alpha\, T(\vec{u})$ for all $\vec{u}$ in $\mathbb{R}^n$, $\alpha \in \mathbb{R}$.

We can view $T(\underline{x}) = A\,\underline{x}$ for some $m \times n$ matrix A.

§1 Null Space & Range

Fix $T: \mathbb{R}^n \longrightarrow \mathbb{R}^m$ linear transformation

We can associate two natural subspaces to T = Null Space & Range

Def 1: The Null Space (or Kernel) of T is the set of vectors $\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ in $\mathbb{R}^n$ mapping to $\underline{0} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$ in $\mathbb{R}^m$ under T, that is

$$\mathcal{N}(T) = \{ \vec{v} \text{ in } \mathbb{R}^n : T(\vec{v}) = \vec{0} \text{ in } \mathbb{R}^n \}$$

Def 2: The Range of T is the image of T, that is

$$\mathcal{R}(T) = \{ \vec{w} \text{ in } \mathbb{R}^m : \text{we can find } \vec{u} \text{ in } \mathbb{R}^n \text{ with } T(\vec{u}) = \vec{w} \}$$

Our interpretation of T as multiplication by A yields:

Theorem 2: • $\mathcal{N}(T) = \mathcal{N}(A)$ (nullspace of A) & $\mathcal{R}(T) = \text{ColSp}(A)$ (= Range of A)

In particular, • $\mathcal{N}(T)$ is a subspace of $\mathbb{R}^n$ with $\dim = \text{null}(A)$

• $\mathcal{R}(T)$ ——— $\mathbb{R}^m$ ——— $= \text{rank}(A)$

We define nullity $(T) = \dim \mathcal{N}(T)$ $\left(= \text{nullity}(A)\right)$

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

Corollary: $\dim \mathcal{N}(T) + \dim \mathcal{R}(T) = n$ $\left(= \#\text{ cols}(A)\right)$

Example: $T: \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ $\quad T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 1 & 0 \\ 8 & -4 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

- $\mathcal{N}(T) = \mathcal{N}\left(\begin{bmatrix} 1 & 0 \\ 8 & -4 \\ 1 & 2 \end{bmatrix}\right)$ $\quad \begin{bmatrix} 1 & 0 \\ 8 & -4 \\ 1 & 2 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 \\ 0 & -4 \\ 0 & 2 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$

  So $x_1 = x_2 = 0$ $\quad \mathcal{N}(T) = \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\} \rightsquigarrow \text{nullity}(T) = 0$

- $\mathcal{R}(T) = ?$ Subspace of $\mathbb{R}^3$ with $\dim = 2 - 0 = 2$

  so it's a plane in $\mathbb{R}^3$.

  $\mathcal{R}(T) = \text{Sp}\left(\begin{bmatrix} 1 \\ 8 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ -4 \\ 2 \end{bmatrix} \right)$ & $\dim = 2$ so we get the equation by computing the normal $\vec{\eta} = \begin{bmatrix} 1 \\ 8 \\ 1 \end{bmatrix} \times \begin{bmatrix} 0 \\ -4 \\ 2 \end{bmatrix}$

  $$\vec{\eta} = \det\left(\begin{bmatrix} i & j & k \\ 1 & 8 & 1 \\ 0 & -4 & 2 \end{bmatrix} \right) = \begin{bmatrix} 20 \\ -2 \\ -4 \end{bmatrix} = 2 \begin{bmatrix} 10 \\ -1 \\ -2 \end{bmatrix}$$

  So $\mathcal{R}(T) = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} : 10x_1 - x_2 - 2x_3 = 0 \right\}$.

§2. Matrix Representations:

Obs: $\boxed{T(\underbrace{\vec{0}}_{\text{in } \mathbb{R}^n}) = \vec{0} \text{ in } \mathbb{R}^m.}$ because $\vec{0} = 0 \cdot \vec{0}$ in both $\mathbb{R}^n$ & $\mathbb{R}^m$ & $\vec{0} = \underset{\text{scalar}}{0}\,\vec{w}$ for any $\vec{w}$ in $\mathbb{R}^m$

so $T(\vec{0}) = T(0 \cdot \vec{0}) \overset{(2)}{=} 0\, \underbrace{T(\vec{0})}_{\vec{w} \text{ in } \mathbb{R}^m} = \vec{0}$ in $\mathbb{R}^m$.

Consequence: $\vec{0}$ in $\mathcal{N}(T)$ & $\vec{0}$ in $\mathcal{R}(T)$ (We knew this because $\mathcal{N}(T)$ & $\mathcal{R}(T)$ are subspaces!)

Matrix Representation Theorem: Any linear transformation $T: \mathbb{R}^n \longrightarrow \mathbb{R}^m$ can be written as $T(\vec{v}) = A \cdot \vec{v}$, where the matrix A of size $m \times n$ has the form $A = [\, T(e_1) \cdots T(e_n) \,]$

canonical basis elements in $\mathbb{R}^n$ (in the given order!)

Alternatively: A = coefficient matrix for the linear expression in each coord. of $T\left(\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}\right)$.

Example: $T: \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ $\quad T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} x_1 + 2x_3 \\ 2x_2 - 5x_3 \end{bmatrix} \rightsquigarrow A = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 2 & -5 \end{bmatrix}$

Here, $T\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $T\left(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$ & $T\left(\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 2 \\ -5 \end{bmatrix}$

Conclusion: T is completely determined by its values on $\{e_1, \ldots, e_n\}$

But, why is the canonical basis better than any other basis for $\mathbb{R}^n$?

Answer: It is not!

same number of vectors as $\dim \mathbb{R}^n = n$

Proposition: Given a basis $\mathcal{B} = \{\vec{v}_1, \ldots, \vec{v}_n\}$ for $\mathbb{R}^n$ & ANY $\{\vec{w}_1, \ldots, \vec{w}_n\}$ in $\mathbb{R}^m$, there is a unique linear transformation $T: \mathbb{R}^n \to \mathbb{R}^m$ with $T(\vec{v}_1) = \vec{w}_1, \ldots, T(\vec{v}_n) = \vec{w}_n$

Example (last time): $T\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 4 \\ 3 \end{bmatrix}$, $T\left(\begin{bmatrix} 1 \\ 2 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 0 \\ 5 \end{bmatrix}$ determine a unique linear transf $T: \mathbb{R}^2 \to \mathbb{R}^3$ because $\left\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix}\right\}$ is a basis for $\mathbb{R}^2$. $\quad$ A: $T(\vec{v}) = \begin{bmatrix} 1 & 0 \\ 8 & -4 \\ 1 & 2 \end{bmatrix} \vec{v}$

This is not true if $\{v_1, \ldots, v_n\}$ is not a basis.

Ex $\quad \vec{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \vec{v}_2 = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$ $\quad$ l.d vectors in $\mathbb{R}^2$

$\vec{w}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad \vec{w}_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

We cannot have a linear transf $T: \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ with $T(\vec{v}_1) = \vec{w}_1$ & $T(\vec{v}_2) = \vec{w}_2$

Why? $\vec{w}_2 = T(\vec{v}_2) = T\Big(\underbrace{2 \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix}}_{= v_1}\Big) \underset{T \text{ linear}}{=} 2\, T\Big(\underbrace{\begin{bmatrix} 1 \\ 0 \end{bmatrix}}_{\vec{v}_1}\Big) = 2 \cdot \vec{w}_1$

but $\vec{w}_2 \neq 2\vec{w}_1$

The problem: explicit linear dependency for $\vec{v}_1, \vec{v}_2$ does NOT hold for $\{\vec{w}_1, \vec{w}_2\}$

Another example: We cannot have $T: \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ linear with $T\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix}$, $T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ & $T\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}$.

Why? $\begin{bmatrix}1\\1\end{bmatrix} = \begin{bmatrix}1\\0\end{bmatrix} + \begin{bmatrix}0\\1\end{bmatrix} \xrightarrow{\text{apply } T} T\left(\begin{bmatrix}1\\1\end{bmatrix}\right) = T\left(\begin{bmatrix}1\\0\end{bmatrix}\right) + T\left(\begin{bmatrix}0\\1\end{bmatrix}\right)$

$$\begin{bmatrix}3\\1\\2\end{bmatrix} = \begin{bmatrix}2\\0\\0\end{bmatrix} + \begin{bmatrix}1\\0\\1\end{bmatrix}$$ This is false!

So T linear with the prescribed assigned values for e_1, e_2 & $\begin{bmatrix}1\\1\end{bmatrix}$ cannot exist.

Proof of Prop: Since $B = \{\vec{v}_1, \dots, \vec{v}_n\}$ is a basis for $\mathbb{R}^n$, any vector $\vec{v}$ in $\mathbb{R}^n$ can be uniquely written as $\boxed{\vec{v} = \alpha_1 \vec{v}_1 + \dots + \alpha_n \vec{v}_n}$ for suitable $\alpha_1, \dots, \alpha_n$

(Recall: $[\vec{v}]_B = \begin{bmatrix}\alpha_1\\ \vdots \\ \alpha_n\end{bmatrix}$ coordinates of $\vec{v}$ with respect to the basis B)

Then, "apply T" to the boxed expression & use the linearity properties

(Remember, we are trying to guess the value of $T(\vec{v})$!!)

$$T(\vec{v}) = T(\alpha_1 \vec{v}_1 + \dots + \alpha_n \vec{v}_n) \underset{\text{Prop (1)}}{=} T(\alpha_1 \vec{v}_1) + \dots + T(\alpha_n \vec{v}_n)$$

$$\underset{\text{Prop (2)}}{=} \alpha_1 \underbrace{T(\vec{v}_1)}_{= \vec{w}_1} + \dots + \alpha_n \underbrace{T(\vec{v}_n)}_{= \vec{w}_n}$$ → (these are the prescribed values for T!)

Conclude: $$T(\vec{v}) = \underbrace{[\vec{w}_1 \ \dots \ \vec{w}_n]}_{\text{matrix of } T(\vec{v}_1), \dots T(\vec{v}_n)} [\vec{v}]_B$$

Can check: This map T is linear $\left(\begin{array}{l} T(\vec{v} + \vec{u}) = T(\vec{v}) + T(\vec{u}) \\ T(\beta \vec{v}) = \beta T(\vec{v}) \end{array} \right)$

The reason: Taking coordinates with respect to a basis is linear!

Given B basis for $\mathbb{R}^n$, $F: \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is linear.

$$\vec{v} \longmapsto [\vec{v}]_B$$

Q: Why are matrix representations useful?

A: Allow for fast compositions!

same number !!

Prop: Given $F: \mathbb{R}^n \longrightarrow \mathbb{R}^{\boxed{m}}$, $G: \mathbb{R}^{\boxed{m}} \longrightarrow \mathbb{R}^s$ two linear transf

then the composition $G \circ F: \mathbb{R}^n \longrightarrow \mathbb{R}^s$ defined by

$\left(\begin{array}{l} G\circ F: \mathbb{R}^n \xrightarrow{F} \mathbb{R}^m \xrightarrow{G} \mathbb{R}^s \\ \text{apply } F \text{ first \& then } G \end{array} \right)$ $\quad \underset{\text{in } \mathbb{R}^n}{\vec{v}} \longmapsto G(\underbrace{F(\vec{v})}_{\text{in } \mathbb{R}^m})$ in $\mathbb{R}^s$

is <u>also</u> a linear transformation.

Furthermore, if $F(\vec{v}) = A\vec{v}$, A of size $m \times n$, then

$G(\vec{w}) = B\vec{w}$, B — $s \times m$

the matrix representing $G\circ F$ is BA (size $s \times n$).

$\boxed{\text{VERY IMPORTANT:}}$ We must multiply B & A in the <u>same order</u> as we compose G & F.

• Before we discuss the proof, let's look at an example.

<u>Example</u>: $F: \mathbb{R}^4 \longrightarrow \mathbb{R}^2$ $\quad F\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}\right) = \begin{bmatrix} x_1 - x_2 \\ x_3 + x_4 \end{bmatrix}$ $\quad A = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$

$G: \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ $\quad G\left(\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}\right) = \begin{bmatrix} y_1 \\ y_1 + 5y_2 \\ -y_1 + y_2 \end{bmatrix}$ $\quad B = \begin{bmatrix} 1 & 0 \\ 1 & 5 \\ -1 & 1 \end{bmatrix}$

So $BA = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 1 & -1 & 5 & 5 \\ -1 & 1 & 1 & 1 \end{bmatrix}$

replace $y_1 = (x_1 - x_2)$, $y_2 = (x_3 + x_4)$ in formula for G

And $G\circ F: \mathbb{R}^4 \longrightarrow \mathbb{R}^3$ is $G\circ F\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}\right) = G\left(\begin{bmatrix} x_1 - x_2 \\ x_3 + x_4 \end{bmatrix}\right) = \begin{bmatrix} (x_1 - x_2) \\ x_1 - x_2 + 5(x_3 + x_4) \\ -(x_1 + x_2) + (x_3 + x_4) \end{bmatrix}$

$= \begin{bmatrix} x_1 - x_2 \\ x_1 - x_2 + 5x_3 + 5x_4 \\ -x_1 + x_2 + x_3 + x_4 \end{bmatrix}$

<u>Check</u>: matrix for $G\circ F$ must have size 3×4 & $= \begin{bmatrix} 1 & -1 & 0 & 0 \\ 1 & -1 & 5 & 5 \\ -1 & 1 & 1 & 1 \end{bmatrix}$, just as the Prop predicted!

<u>Proof</u>: Check 2 linear properties

(1) $G(F(\vec{v} + \vec{u})) \underset{F \text{ linear}}{=} G(\underbrace{F(\vec{v})}_{= \vec{w}_1} + \underbrace{F(\vec{u})}_{\vec{w}_2}) \underset{G \text{ linear}}{=} G(F(\vec{v})) + G(F(\vec{u}))$ ✓

(2) $G(F(\alpha \vec{v})) \underset{F \text{ linear}}{=} G(\alpha \underbrace{F(\vec{v})}_{\vec{w} \text{ in } \mathbb{R}^m}) \underset{G \text{ linear}}{=} \alpha\, G(F(\vec{v}))$ ✓

So we know $G\circ F$ has a matrix representing it. Indeed, the matrix is $\left[\, G\circ F(e_1) \;\cdots\; G\circ F(e_n) \,\right]$

We compute each column vector:

$$G\circ F(e_1) \underset{\text{def of } F}{=} G\left(A\begin{bmatrix}1\\0\\\vdots\\0\end{bmatrix}\right) = G(\underbrace{col_1(A)}_{\vec{w}_1 \text{ in } \mathbb{R}^m}) \underset{\text{def of } G}{=} B\, col_1(A)$$

$$\vdots$$

$$G\circ F(e_n) = G\left(A\begin{bmatrix}0\\\vdots\\0\\1\end{bmatrix}\right) = G(\underbrace{col_n(A)}_{=\vec{w}_n \text{ in } \mathbb{R}^m}) \underset{\text{def of } G}{=} B\, col_n(A)$$

so $\left[G\circ F(e_1) \cdots G\circ F(e_n)\right] = \left[\underbrace{B\,col_1(A)}_{col_1(BA)} \cdots\cdots \underbrace{B\,col_n(A)}_{col_n(BA)}\right] \underset{\text{def of matrix multiplication}}{=} BA$ size $s\times n$

This size is consistent with $G\circ F: \mathbb{R}^n \longrightarrow \mathbb{R}^s$. $\square$