Lecture XIV: §5.7 Linear transformation for abstract vector spaces.

§1. Main idea:

We can easily generalize linear transformations $T: \mathbb{R}^n \longrightarrow \mathbb{R}^m$ to linear transformations between abstract vector spaces because we have a way to add & scale multiply vectors in abstract vector spaces.

Our maps $T: V \longrightarrow W$ will be linear if the respect these 2 operations. More precisely:

Def: Fix V, W two (abstract) vector spaces & let $T: V \longrightarrow W$, $\vec{v} \longmapsto T(\vec{v})$ be a map. We say T is a linear transformation if

(1) $T(\underbrace{\vec{v} +_V \vec{u}}_{\text{sum in } V}) = T(\vec{v}) \underbrace{+_W}_{\text{sum in } W} T(\vec{u})$ for all $\vec{u}, \vec{v}$ in V & (2) $T(\underbrace{\alpha \cdot_V \vec{v}}_{\text{scalar mult in } V}) = \alpha \underbrace{\cdot_W}_{\text{in } W} T(\vec{v})$ for $\vec{v}$ in V, α in $\mathbb{R}$

§2 Examples:

$\boxed{\text{Example 0}}$ Take $V = \mathbb{R}^n$, $W = \mathbb{R}^m$. Then $T: \mathbb{R}^n \longrightarrow \mathbb{R}^m$ linear in the usual sense, so it's just multiplication by an $m \times n$ matrix.

$\boxed{\text{Example 1}}$ $T: \underbrace{\mathcal{P}_2}_{\text{polynomials of deg} \leq 2} \longrightarrow \mathbb{R}$ $\qquad T(P_{(x)}) = P_{(1)}$

Claim: T is linear

Why? because of how we defined addition of polynomials

$$T(\underbrace{P_{(x)} + Q_{(x)}}_{\text{new poly}}) = (P+Q)_{(1)} = P_{(1)} + Q_{(1)} = T(P_{(x)}) + T(Q_{(x)})$$

$$T(\underbrace{\alpha P_{(x)}}_{\text{new polynomial}}) = (\alpha P)_{(1)} \underset{\text{because of how we defined scalar mult in } \mathbb{R}^2}{=} \alpha P_{(1)} = \alpha T(P_{(x)}) \quad \checkmark$$

- These statements are true but they don't really help us understand what T is

Better way: write $P_{(x)} = a + bx + cx^2 \rightsquigarrow P_{(1)} = a+b+c$

So we can really view T as the linear map $\tilde{T}: \mathbb{R}^3 \longrightarrow \mathbb{R}$

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} \longmapsto a+b+c$$

Here $\begin{bmatrix}a\\b\\c\end{bmatrix} = [p]_B$ coordinates of p in the basis $\{1, x, x^2\}$ for $\mathcal{P}_2$

- This example generalizes to $T: C_{[-2,3]} \longrightarrow \mathbb{R}$
$$f \longmapsto f(1)$$
(we can replace 1 by any fixed number x_0 in $[-2,3]$)

We call it an evaluation map at $x_0 = 1$. It is linear!

Example 2: Taking coordinates with respect to a fixed basis B for $\mathbb{V}$ where $\dim \mathbb{V}$ is finite

Fix $\dim \mathbb{V} = p$. Then $T: \mathbb{V} \longrightarrow \mathbb{R}^p$ is a linear transf.
$$\vec{v} \longmapsto [\vec{v}]_B$$

Why?
$$\vec{v} = \alpha_1 \vec{v}_1 + \cdots + \alpha_p \vec{v}_p \rightsquigarrow [\vec{v}]_B = \begin{bmatrix}\alpha_1\\ \vdots \\ \alpha_p\end{bmatrix}$$
$$+ \quad \vec{w} = \beta_1 \vec{v}_1 + \cdots + \beta_p \vec{v}_p \rightsquigarrow [\vec{w}]_B = \begin{bmatrix}\beta_1\\ \vdots \\ \beta_p\end{bmatrix}$$
$$\vec{v} + \vec{w} = (\alpha_1+\beta_1)\vec{v}_1 + \cdots + (\alpha_p+\beta_p)\vec{v}_p \overset{(1)}{\rightsquigarrow} \boxed{[\vec{v}+\vec{w}]_B = \begin{bmatrix}\alpha_1+\beta_1\\ \vdots \\ \alpha_p+\beta_p\end{bmatrix} = [\vec{v}]_B + [\vec{w}]_B \checkmark}$$
(sum scalars columnwise)

Also
$$\alpha \vec{w} = (\alpha\beta_1)\vec{v}_1 + \cdots + (\alpha\beta_p)\vec{v}_p \overset{(2)}{\rightsquigarrow} \boxed{[\alpha\vec{w}]_B = \begin{bmatrix}\alpha\beta_1\\ \vdots \\ \alpha\beta_p\end{bmatrix} = \alpha[\vec{w}]_B \checkmark}$$

Obs: We can combine Examples 1 & 2 via composition

$$\begin{tikzcd} T: \mathcal{P}_2 \arrow[r] \arrow[d, "F"'] & \mathbb{R} \arrow[d, equal] \\ \tilde{T}: \mathbb{R}^3 \arrow[r] & \mathbb{R} \end{tikzcd}$$

$$F = \mathcal{P}_2 \longrightarrow \mathbb{R}^3$$
$$p \longmapsto [p]_{\underbrace{\{1,x,x^2\}}_{=B}}$$

Take $B = \{1, x, x^2\}$ standard basis for $\mathcal{P}_2$.
$$T(a+bx+cx^2) = a+b+c$$
$$\tilde{T}\left(\begin{bmatrix}a\\b\\c\end{bmatrix}\right) = a+b+c$$

$\rightsquigarrow$ Translation: For any $p_{(x)}$ in $\mathcal{P}_2$ we have
$$\boxed{T_{(p)} = \tilde{T}([p]_B) = \tilde{T} \circ F_{(p)}}$$

KEY FACT

We will be able to do a similar translation from $T: \mathbb{V} \longrightarrow \mathbb{W}$ with $\dim \mathbb{V} = n$, $\dim \mathbb{W} = m$ to $\tilde{T}: \mathbb{R}^n \longrightarrow \mathbb{R}^m$ by identifying $\mathbb{V}$ with $\mathbb{R}^n$ via coordinates wrt a basis $B_{\mathbb{V}}$ for $\mathbb{V}$

$\mathbb{W}$ with $\mathbb{R}^m$ —— $B_{\mathbb{W}}$ for $\mathbb{W}$

(More on this, next time)

Example 3 $T: C_{[0,1]} \longrightarrow \mathbb{R}$ is a linear transformation

$$f(x) \longmapsto \int_0^1 f(x)\,dx$$

Example 4 $T: C'_{(0,1)} \longrightarrow C_{(0,1)}$

$$f(x) \longmapsto f'(x)$$

(Recall: $C'_{(0,1)}$ = continuous functions $f: (0,1) \to \mathbb{R}$ with $f': (0,1) \to \mathbb{R}$ also continuous)

T is a linear transformation

Example 5 $T: M_{2\times 3} \longrightarrow M_{4\times 3}$ is a linear transform.

$$A \longmapsto \boxed{\begin{bmatrix} 1 & 3 \\ 0 & 4 \\ 0 & 5 \\ 0 & 6 \end{bmatrix}} A$$

replacing this by any other fixed matrix of size 4×2 will have the same effect.

(Left multiplication by a fixed matrix of appropriate size)

Example 6 $T_1: \mathbb{V} \longrightarrow \mathbb{V}$ is linear & $T_2: \mathbb{V} \longrightarrow \mathbb{W}$ is linear

$\vec{v} \longmapsto \vec{v}$ (identity map) $\qquad \vec{v} \longmapsto \mathbb{0}_{\mathbb{W}}$ (zero map)

§3 Basic Properties:

Theorem 1: Assume $\mathbb{V}$ has finite dimension, and basis $B = \{v_1, \dots, v_p\}$

Then, given any set $\{\vec{w}_1, \dots, \vec{w}_p\}$ of vectors in $\mathbb{W}$, we can find a UNIQUE linear transformation $T: \mathbb{V} \longrightarrow \mathbb{W}$ with $T(\vec{v}_1) = \vec{w}_1$,

$\vdots$

$T(\vec{v}_p) = \vec{w}_p$.

Obs: We don't need $\mathbb{W}$ to have finite dimension, only $\mathbb{V}$.

Why is Theorem 1 true? Use coordinates with respect to $\mathcal{B}$ & result when $\mathbb{V}=\mathbb{R}^n$, $\mathbb{W}=\mathbb{R}^m$

$$T(\vec{v}) = T(\alpha_1\vec{v}_1+\cdots+\alpha_p\vec{v}_p) \underset{T \text{ linear!}}{=} \alpha_1 \underbrace{\boxed{T(\vec{v}_1)}}_{=\vec{w}_1}+\cdots+\alpha_p \underbrace{\boxed{T(\vec{v}_p)}}_{=\vec{w}_p}$$

So $T(\vec{v}) = \alpha_1\vec{w}_1+\cdots+\alpha_p\vec{w}_p$ where $[v]_{\mathcal{B}} = \begin{bmatrix}\alpha_1\\ \vdots \\ \alpha_p\end{bmatrix}$

Check: The map T defined by this formula is linear

- Uniqueness follows because we found a unique possibility for defining $T(\vec{v})$ using $\mathcal{B}$ & $\vec{w}_1,\ldots,\vec{w}_p$.

Application: Find a linear transformation $T:\mathcal{P}_3\to\mathcal{P}_2$ with

$T(1)=2+x$, $T(x)=x-x^2$, $T(x^2)=5-10x$ & $T(x^3)=2$.

Solution: $T(a+bx+cx^2+dx^3) \overset{T \text{ linear}}{=} aT(1)+bT(x)+cT(x^2)+dT(x^3)$

$= a(2+x)+b(x-x^2)+c(5-10x)+d\cdot 2$

we have preassigned values for $T(1), T(x), T(x^2)$ & $T(x^3)$

$= (2a+5c+2d)+(a-b-10c)x+(-b)x^2$

Note: If we fix $\mathcal{B}=\{1,x,x^2,x^3\}$ basis for $\mathbb{V}$, $\mathcal{B}_2=\{1,x,x^2\}$ —— $\mathbb{W}$, we can identify $T:\mathcal{P}_3\to\mathcal{P}_2$ with $\tilde{T}:\mathbb{R}^4\to\mathbb{R}^3$

$$\begin{bmatrix}a\\b\\c\\d\end{bmatrix}\longmapsto\begin{bmatrix}2a+5c+2d\\a+b-10c\\-b\end{bmatrix}$$

Here $[a+bx+cx^2+dx^3]_{\mathcal{B}}=\begin{bmatrix}a\\b\\c\\d\end{bmatrix}$

$[T(a+bx+cx^2+dx^3)]_{\mathcal{B}_2}=\begin{bmatrix}2a+5c+2d\\a+b-10c\\-b\end{bmatrix}$

So we get a "commutative diagram" we going from $\mathbb{V}\to\mathbb{R}^3$ in 2 ways must be the same!

way 1

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[d, "{[\ ]_B}"', "\cong"] \arrow[dr, phantom, "\circlearrowleft"] & W \arrow[d, "{[\ ]_{B_2}}", "\cong"'] \\ \mathbb{R}^4 \arrow[r, "\tilde{T}"'] & \mathbb{R}^3 \end{tikzcd}$$

way 2

$\rightsquigarrow$ $[T(p)]_{B_2} = \tilde{T}([p]_B)$

(way 1) (way 2)

Upshot: Studying linear transformations between vectors spaces of finite dimension is as EASY as studying linear maps $T: \mathbb{R}^n \longrightarrow \mathbb{R}^m$. Just pick coordinates on V & W. :)

§4. Null Space & Range of $T: V \longrightarrow W$:

• Next, we want to study the analog of Null Space & Range for $F: \mathbb{R}^n \longrightarrow \mathbb{R}^m$ in this new framework. We define them in the same way!

Def 1: The Null Space of T is $\mathcal{N}(T) = \{\vec{v} \text{ in } V : T(\vec{v}) = \mathbb{O}_W\}$

Def 2: The Range of T is $\mathcal{R}(T) = \{\vec{w} \text{ in } W : \text{we can find } \vec{v} \text{ in } V \text{ with } T(\vec{v}) = \vec{w}\}$

(Range = image of the map T)

• Just as it happen for linear maps $\mathbb{R}^n \longrightarrow \mathbb{R}^m$ these 2 sets are vector subspaces of V & W! More precisely:

Theorem 1: (1) $T(\mathbb{O}_V) = \mathbb{O}_W$

(2) $\mathcal{N}(T)$ is a subspace of V

(3) $\mathcal{R}(T)$ ——————— W

Proof (1) We need to use $\mathbb{O}_V = 0 \cdot \mathbb{O}_V$ & $\mathbb{O}_W = 0 \cdot \vec{w}$ for any $\vec{w}$ in W

$$T(\mathbb{0}_{\mathbb{V}}) = T(0\cdot \mathbb{0}_{\mathbb{V}}) \underset{T \text{ linear}}{=} 0\cdot \underbrace{T(\mathbb{0}_{\mathbb{V}})}_{=:\vec{w}} = \mathbb{0}_{\mathbb{W}}.$$

(2) Need to check $\mathcal{N}(T)$ satisfies the 3 properties for subspaces of $\mathbb{V}$

(S1) $\mathbb{0}_{\mathbb{V}}$ in $\mathcal{N}(T)$ True by (1)

(S2) If $\vec{v},\vec{u}$ in $\mathcal{N}(T)$, then $\vec{v}+\vec{u}$ in $\mathcal{N}(T)$.

Why? $T(\vec{v}) = \mathbb{0}_{\mathbb{W}}$, $T(\vec{u}) = \mathbb{0}_{\mathbb{W}}$ then $T(\vec{v}+\vec{u}) \overset{T\text{ linear}}{=} T(\vec{v}) + T(\vec{u}) = \mathbb{0}_{\mathbb{W}} + \mathbb{0}_{\mathbb{W}} = \mathbb{0}_{\mathbb{W}}$

so $\vec{v}+\vec{u}$ is also in $\mathcal{N}(T)$. ✓

(S3) If $\vec{v}$ in $\mathcal{N}(T)$ & α in $\mathbb{R}$, then $\alpha\cdot\vec{v}$ in $\mathcal{N}(T)$.

Why? $T(\vec{v}) = \mathbb{0}_{\mathbb{W}}$, so $T(\alpha\cdot\vec{v}) = \alpha\cdot T(\vec{v}) = \alpha\cdot \mathbb{0}_{\mathbb{W}} = \mathbb{0}_{\mathbb{W}}$ ✓

(3) Need to check that $\mathcal{R}(T)$ satisfies the 3 properties for subspaces of $\mathbb{W}$.

(S1) $\mathbb{0}_{\mathbb{W}}$ in $\mathcal{R}(T)$ because $\mathbb{0}_{\mathbb{W}} = T(\mathbb{0}_{\mathbb{V}})$.

(S2) If $\vec{w}_1,\vec{w}_2$ in $\mathcal{R}(T)$, then $\vec{w}_1+\vec{w}_2$ in $\mathcal{R}(T)$

Why? We know $T(\vec{v}_1) = \vec{w}_1$ for some $\vec{v}_1$ in $\mathbb{V}$

$T(\vec{v}_2) = \vec{w}_2$ —— $\vec{v}_2$ ——

Then $\vec{w}_1+\vec{w}_2 = T(\vec{v}_1) + T(\vec{v}_2) \underset{T\text{ linear}}{=} T(\underbrace{\vec{v}_1+\vec{v}_2}_{=\vec{v} \text{ in } \mathbb{V}})$ as we wanted! ✓

(S3) If $\vec{w}$ in $\mathcal{R}(T)$ & α in $\mathbb{R}$, then $\alpha\cdot\vec{w}$ in $\mathcal{R}(T)$

Why? We have $\vec{w} = T(\vec{v})$ for some $\vec{v}$ in $\mathbb{V}$.

Then $\alpha\cdot\vec{w} = \alpha\, T(\vec{v}) \underset{T\text{ linear}}{=} T(\underbrace{\alpha\vec{v}}_{\vec{u} \text{ in } \mathbb{V}})$ as we wanted! ✓ $\square$

- Our next result says that we can compute $\mathcal{R}(T)$ very fast when $\mathbb{V}$ has finite dimension

Theorem 2: Fix $T:\mathbb{V}\longrightarrow\mathbb{W}$ linear transformation & assume $\dim\mathbb{V}=p$. Let $\mathcal{B}=\{\vec{v}_1,\ldots,\vec{v}_p\}$ be a basis for $\mathbb{V}$. Then:

(1) $\mathcal{R}(T) = \operatorname{Sp}(T(\vec{v}_1),\ldots,T(\vec{v}_p))$, so $\mathcal{R}(T)$ has dimension $\leq p$.

(2) $\mathcal{N}(T) = \{\mathbb{0}_{\mathbb{V}}\}$ if and only if $\{T(\vec{v}_1),\ldots,T(\vec{v}_p)\}$ is l.i. in $\mathbb{W}$

Proof: (1) Pick $\vec{w}$ in $R(T)$ & fix $\vec{v}$ in $\mathbb{V}$ with $w = T(\vec{v})$

Since $\mathcal{B}$ is a basis for $\mathbb{V}$ we can write $\vec{v} = \alpha_1\vec{v}_1 + \cdots + \alpha_p\vec{v}_p$, where $[\vec{v}]_{\mathcal{B}} = \begin{bmatrix}\alpha_1\\ \vdots \\ \alpha_p\end{bmatrix}$ in $\mathbb{R}^p$.

But then $\vec{w} = T(\vec{v}) = T(\alpha_1\vec{v}_1 + \cdots + \alpha_p\vec{v}_p) = \alpha_1\boxed{T(\vec{v}_1)} + \cdots + \alpha_p\boxed{T(\vec{v}_p)}$

Conclude: $\vec{w}$ in $Sp(T(\vec{v}_1), \ldots, T(\vec{v}_p))$ for any $\vec{w}$ in $R(T)$ $(*)$

- Since $T(\vec{v}_1), \ldots, T(\vec{v}_p)$ are all in $R(T)$ & $R(T)$ is a subspace of $\mathbb{W}$

we have $Sp(T(v_1), \ldots, T(\vec{v}_p))$ lies in $R(T)$

From this & $(*)$ we get $R(T) = Sp(T(\vec{v}_1), \ldots, T(\vec{v}_p))$.

(2) We start by setting up a dependency relation for $\{T(\vec{v}_1), \ldots, T(\vec{v}_p)\}$ is l.i.

$(*)$ $\boxed{0_{\mathbb{W}} = \alpha_1 T(\vec{v}_1) + \cdots + \alpha_p T(\vec{v}_p)} \underset{T\text{ linear}}{=} T(\alpha_1\vec{v}_1 + \cdots + \alpha_p\vec{v}_p)$

This says $\alpha_1\vec{v}_1 + \cdots + \alpha_p\vec{v}_p$ is in $\mathcal{N}(T)$

- If $\mathcal{N}(T) = \{0_{\mathbb{V}}\}$ we get $0_{\mathbb{V}} = \alpha_1\vec{v}_1 + \cdots + \alpha_p\vec{v}_p$

but $\{\vec{v}_1, \ldots, \vec{v}_p\}$ is l.i (it's a basis!) so this forces $\alpha_1 = \cdots = \alpha_p = 0$

Going back to the boxed expression $(*)$, we conclude $\{T(\vec{v}_1), \ldots, T(\vec{v}_p)\}$ is l.i.

- Conversely, assume $\{T(\vec{v}_1), \ldots, T(\vec{v}_p)\}$ is l.i. We want to show $\mathcal{N}(T) = \{0_{\mathbb{V}}\}$

To do so, we pick $\vec{v}$ in $\mathcal{N}(T)$ & write it using the basis $\mathcal{B}$, that is

$(*)$ $\vec{v} = \beta_1\vec{v}_1 + \cdots + \beta_p\vec{v}_p$ for some $\beta_1, \ldots, \beta_p$ in $\mathbb{R}$

Now, we apply T to both sides:

$\underset{\vec{v}\text{ in }\mathcal{N}(T)}{\boxed{\vec{0}} =} T(\vec{v}) = T(\beta_1\vec{v}_1 + \cdots + \beta_p\vec{v}_p) \underset{T\text{ linear}}{=} \boxed{\beta_1 T(\vec{v}_1) + \cdots + \beta_p T(\vec{v}_p)}$

So $\vec{0} = \beta_1 T(\vec{v}_1) + \cdots + \beta_p T(\vec{v}_p)$ forces $\beta_1 = \cdots = \beta_p = 0$

l.i

So replacing in $(*)$ we get $\vec{v} = 0\cdot\vec{v}_1 + 0\cdot\vec{v}_2 + \cdots + 0\cdot\vec{v}_p = 0_{\mathbb{V}}$.

Conclude: The only element in $\mathcal{N}(T)$ is $0_{\mathbb{V}}$. $\square$

Example: $T: \mathcal{P}_2 \longrightarrow \mathbb{R}^2$ $\quad T(P) = \begin{bmatrix} P(1) \\ P'(1)\end{bmatrix}$ for P in $\mathcal{P}_2$ is linear

$\mathcal{N}(T) = \{P_{(x)} = a + bx + cx^2 : P(1) = a + b + c = 0,\ P'(1) = b + 2c = 0\}$ $\rightsquigarrow \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2\end{bmatrix} \longrightarrow \begin{array}{c} \ \ \ \text{indep} \\ \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2\end{bmatrix} \\ \text{REF}\end{array}$ $\quad a = c$, $b = -2c$

$= \{c + (-2c)x + cx^2 : c \text{ in } \mathbb{R}\}$

$= \{c(1 - 2x + x^2) : c \text{ in } \mathbb{R}\} = Sp(1 - 2x + x^2)$ $\quad$ dim $= 1$

$$\mathcal{R}(T) = \left\{ \begin{bmatrix} a+b+c \\ b+2c \end{bmatrix} : a,b,c \text{ in } \mathbb{R} \right\}$$ $\dim = 2 \leq 3 = \dim \mathcal{P}_2$

$$= \left\{ a\begin{bmatrix} 1 \\ 0 \end{bmatrix} + b\begin{bmatrix} 1 \\ 1 \end{bmatrix} + c\begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\} = \mathrm{Sp}\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right)$$

$B = \{1, x, x^2\} \rightsquigarrow T(1) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \; T(x) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \; T(x^2) = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \rightsquigarrow$ as we found

- The next result relates injectivity of T & $\mathcal{N}(T) = \mathbb{0}_V$

Proposition: Fix $T: V \longrightarrow W$ linear transformation.

(1) $T(\vec{v}) = T(\vec{u})$ if and only if $\vec{u} - \vec{v}$ in $\mathcal{N}(T)$

(2) T is injective (meaning $T(\vec{v}) = T(\vec{u})$ forces $\vec{v} = \vec{u}$) if & only if $\mathcal{N}(T) = \{\mathbb{0}_V\}$

Proof: (2) follows easily from (1) because $\vec{u} - \vec{v} = \mathbb{0}_V$ means $\vec{v} = \vec{u}$.

To show (1), we use the linear properties of T:

$T(\vec{v}) = T(\vec{u})$ means $\boxed{\mathbb{0}_W} = T(\vec{u}) - T(\vec{v}) = T(\vec{u}) + (-1)T(\vec{v})$
$= T(\vec{u}) + T((-1)\vec{v}) = T(\vec{u} + (-1)\vec{v}) = \boxed{T(\vec{u} - \vec{v})}$

This means $\vec{u} - \vec{v}$ in $\mathcal{N}(T)$ $\square$

ASIDE OBSERVATION:

Q: Why is this relevant? We can thing of breaking (or "partition") V into different set (that do not intersect), according to their value under T

For each w in W we set $V_w = \{\vec{u} \text{ in } V : T(\vec{u}) = w\}$

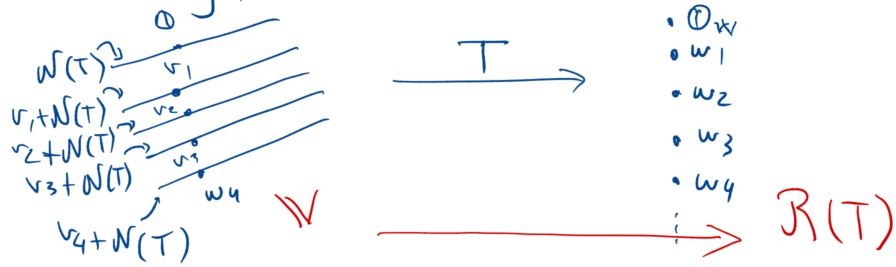
- If w is not in $\mathcal{R}(T)$, then $V_w = \emptyset$ (empty set)
- If w is in $\mathcal{R}(T)$, we know $\vec{w} = T(\vec{v})$ for some $\vec{v}$ in V

Prop (1) says that $V(w) = \vec{v} + \mathcal{N}(T)$

Why? Pick $\vec{u}$ in $V(w)$, then $T(\vec{u}) = T(\vec{v})$ so $(\vec{u} - \vec{v})$ in $\mathcal{N}(T)$

But then $\vec{u} = \vec{v} + \underbrace{(\vec{u} - \vec{v})}_{\text{in } \mathcal{N}(T)}$

Essentially, we can view $\mathbb{V}$ as "translations" of $\mathcal{N}(T)$ by vectors $\vec{v}$.



This is the essence of the <u>rank-nullity theorem</u>, as we will see next.

§5. Rank-Nullity Theorem

<u>Def</u>: Assume $\mathbb{V}$ is finite dimensional, & let $T: \mathbb{V} \to \mathbb{W}$ be a linear transformation.

- $\text{nullity}(T) = \dim \mathcal{N}(T)$ (finite & $\leq \dim \mathbb{V}$)
- $\text{rank}(T) = \dim \mathcal{R}(T)$ ($\leq \dim \mathbb{V}$ by Theorem 2)

<u>Rank-Nullity Theorem</u>: Assume $T: \mathbb{V} \to \mathbb{W}$ linear & $\dim \mathbb{V}$ finite

Then: $\text{rank}(T) + \text{nullity}(T) = \dim \mathbb{V}$.

<u>Consequence</u>: A $m \times n$ matrix $\quad \text{rank}(A) + \text{nullity}(A) = n$ (# cols of A)

<u>Proof of Consequence</u>: A defines a linear map $T: \mathbb{R}^n \longrightarrow \mathbb{R}^m$

$$\vec{v} \longmapsto A\cdot\vec{v}$$

<u>But</u> $\mathcal{N}(A) = \mathcal{N}(T)$

$\mathcal{R}(A) = \mathcal{R}(T)$ (Lecture 23)

Rank-Nullity Thm says: $\underbrace{\dim \mathcal{N}(T)}_{=\text{nullity}(A)} + \underbrace{\dim \mathcal{R}(T)}_{=\text{rank}(A)} = n$

- Before discussing the proof of Rank-Nullity Theorem, we show some examples:
(optional reading, but very insightful)

<u>EXAMPLE 1</u>: $\mathbb{V} = \mathcal{P}_2$, $\mathbb{W} = \mathcal{R}(T) = \mathbb{R}^2$ $\quad T: \mathcal{P}_2 \to \mathbb{R}^2 \quad T(p) = \begin{bmatrix} p(1) \\ p'(1) \end{bmatrix}$

$$\mathbb{V}_{\begin{bmatrix} 1 \\ 0 \end{bmatrix}} = 1 + \mathcal{N}(T) = 1 + \alpha(1 - 2x + x^2) \quad \text{for } \alpha \text{ in } \mathbb{R}$$

$$\mathbb{V}_{\begin{bmatrix} 1 \\ 1 \end{bmatrix}} = x + \mathcal{N}(T) = x + \beta(1 - 2x + x^2) \quad \text{for } \beta \text{ in } \mathbb{R}$$

Q: What about $\mathbb{W}_{\vec{w}}$ for other $\vec{w}$ in $\mathbb{R}^2$?

$$\mathbb{W}_{\vec{w}} = \vec{v} + \mathcal{N}(T)$$, but what is $\vec{v}$?

Use $\left\{\begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}1\\1\end{bmatrix}\right\}$ is a basis for $\mathcal{R}(T)$ Write $\vec{w} = a\begin{bmatrix}1\\0\end{bmatrix} + b\begin{bmatrix}1\\1\end{bmatrix}$

So $\vec{w} = a\,T(1) + b\,T(x) = T(a+bx)$

Answer $\mathbb{W}_{a\left[\begin{smallmatrix}1\\0\end{smallmatrix}\right]+b\left[\begin{smallmatrix}1\\1\end{smallmatrix}\right]} = (a+bx) + \mathcal{N}(T)$

basis for $\mathcal{N}(T)$

Claim: $\{\underbrace{1, x}_{\downarrow T \text{ basis for } \mathcal{R}(T)=\mathbb{R}^2}, \underbrace{\boxed{1-2x+x^2}}_{\downarrow\ 0_{\mathbb{R}^2}}\}$ is a basis for $\mathcal{P}_2$

$$\dim \mathcal{P}_2 = \dim \mathcal{R}(T) + \dim \mathcal{N}(T)$$
$$3 = 1 + 2$$

This is the rank-nullity theorem in a nutshell!

Example 2: $T: M_{2\times 3} \longrightarrow \mathcal{P}_4$

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \longmapsto (a_{12}+a_{23})x^4 + (2a_{22}+3a_{13})x^3 + (a_{11}-a_{23})$$

- T is linear because we have a linear map $\tilde{T}: M_{2\times 3} \longrightarrow \mathbb{R}^5$, $A \longmapsto [T(A)]_B$ after choosing the basis $B = \{1, x, x^2, x^3, x^4\}$ for $\mathcal{P}_4$.

Indeed: $[T(A)]_B = \begin{bmatrix} a_{11}-a_{23} \\ 0 \\ 0 \\ 2a_{22}+3a_{13} \\ a_{12}+a_{23} \end{bmatrix}$ only involves linear expressions in the coefficients of $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$

- $\mathcal{N}(T) = ?$ Need $\begin{cases} a_{12}+a_{23} = 0 \\ 2a_{22}+3a_{13} = 0 \\ a_{11}-a_{23} = 0 \end{cases}$ (This system also characterizes $\mathcal{N}(\tilde{T})$)

Must solve: $\begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 3 & 0 & 2 & 0 \\ 1 & 0 & 0 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} a_{11} \\ a_{12} \\ a_{13} \\ a_{21} \\ a_{22} \\ a_{23} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

$$\begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 3 & 0 & 2 & 0 \\ 1 & 0 & 0 & 0 & 0 & -1 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 3 & 0 & 2 & 0 \end{bmatrix} \longrightarrow \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & \frac{2}{3} & 0 \end{bmatrix}}_{\text{REF}}$$

indep vars

$a_{11} = a_{23}$
$a_{12} = a_{23}$
$a_{13} = -\frac{2}{3} a_{22}$

Any A in $\mathcal{N}(T)$ is $A = \begin{bmatrix} a_{23} & a_{23} & -\frac{2}{3}a_{22} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} = a_{21}\underbrace{\begin{bmatrix} 0&0&0 \\ 1&0&0 \end{bmatrix}} + a_{22}\underbrace{\begin{bmatrix} 0&0&-\frac{2}{3} \\ 0&1&0 \end{bmatrix}} + a_{23}\underbrace{\begin{bmatrix} 1&1&0 \\ 0&0&1 \end{bmatrix}}$

So $\mathcal{N}(T)$ has basis $\left\{ \underset{E_{13}}{\underset{\shortparallel}{\begin{bmatrix} 0&0&0 \\ 1&0&0 \end{bmatrix}}}, \underset{E_{22}-\frac{2}{3}E_{13}}{\underset{\shortparallel}{\begin{bmatrix} 0&0&-\frac{2}{3} \\ 0&1&0 \end{bmatrix}}}, \underset{E_{11}+E_{12}+E_{23}}{\underset{\shortparallel}{\begin{bmatrix} 1&1&0 \\ 0&0&1 \end{bmatrix}}} \right\}$ so $\dim = 3$.

- $\underline{\mathcal{R}(T) = ?}$ We know $\mathcal{B}' = \{E_{11}, E_{12}, E_{13}, E_{21}, E_{22}, E_{23}\}$ basis for $M_{2\times 3}$

ensures $\mathcal{R}(T) = \mathrm{Sp}(T(E_{11}), T(E_{12}), T(E_{13}), T(E_{21}), T(E_{22}), T(E_{23}))$

By Rank-Nullity: $\dim \mathcal{R}(T) = \dim M_{2\times 3} - \dim \mathcal{N}(T) = 6 - 3 = 3$

So we know $\mathcal{R}(T) \neq \mathcal{P}_4$ & a basis will be obtained by finding 3 l.i. vectors among $\{T(E_{11}), \dots, T(E_{23})\}$

$T(E_{11}) = 1$

$T(E_{12}) = x^4$

$T(E_{13}) = 3x^3$ $\quad \rightsquigarrow$ Can pick $\{1, 3x^3, x^4\}$

$T(E_{21}) = 0$

$T(E_{22}) = 2x^3$ $\qquad \mathcal{R}(T) = \mathrm{Sp}(1, 3x^3, x^4)$

$T(E_{23}) = x^4 - 1$ $\qquad = \mathrm{Sp}(1, x^3, x^4)$

We get information about $\mathcal{R}(\tilde{T})$ from this; just take $[\ ]_{\mathcal{B}}$ of these 3 vectors:
($\mathcal{B}$ = standard basis for $\mathcal{P}_4$)

$\mathcal{R}(\tilde{T}) = \mathrm{Sp}([1]_{\mathcal{B}}, [x^3]_{\mathcal{B}}, [x^4]_{\mathcal{B}})$

$$= \mathrm{Sp}\left(\begin{bmatrix} 1\\0\\0\\0\\0 \end{bmatrix}, \begin{bmatrix} 0\\0\\0\\1\\0 \end{bmatrix}, \begin{bmatrix} 0\\0\\0\\0\\1 \end{bmatrix} \right) = \mathrm{Sp}(e_1, e_4, e_5) \text{ in } \mathbb{R}^5$$

This is consistent with the formula we had for $\tilde{T}(A)$ (2^{nd} & 3^{rd} entry were $= 0$).

§6. Proof of the Rank-Nullity Theorem: (optional)

Fix $\dim \mathbb{V} = p$ & $\mathcal{B} = \{\vec{v}_1, \ldots, \vec{v}_p\}$ a basis for $\mathbb{V}$

- We start by recalling a fact: $\operatorname{rank}(T) \leq p$ (by Theorem 2)
- We prove 2 special cases, where $\operatorname{rank}(T) = 0$ & $\operatorname{rank}(T) = p$

(1) Assume $\operatorname{rank}(T) = 0$, meaning $R(T) = \{\mathbb{0}_{\mathbb{W}}\}$. In this case $T(\vec{v}) = \mathbb{0}_{\mathbb{W}}$ for all $\vec{v}$ so $\mathcal{N}(T) = \mathbb{V}$, & $\operatorname{nullity}(T) = \dim \mathbb{V}$

Conclusion: $\operatorname{rank}(T) + \operatorname{nullity}(T) = 0 + \operatorname{nullity}(T) = \dim \mathbb{V}$ ✓

(2) Assume $\operatorname{rank}(T) = p$, meaning $R(T) = \operatorname{Sp}(T(\vec{v}_1), \ldots, T(\vec{v}_p))$ has $\dim = p$

In particular $\{T(\vec{v}_1), \ldots, T(\vec{v}_p)\}$ is l.i., so by Theorem 2, $\mathcal{N}(T) = \{\mathbb{0}_{\mathbb{V}}\}$, nullity $= 0$

Conclusion: $\operatorname{rank}(T) + \operatorname{nullity}(T) = \dim \mathbb{V} + 0 = \dim \mathbb{V}$ ✓

All that remains: is to prove the statement whenever $0 < \operatorname{rank}(T) < p$

Names: $r = \operatorname{rank}(T)$, $d = \operatorname{nullity}(T)$ WANT To show: $r + d = p$

Since $r < p$, we know $\dim \operatorname{Sp}(T(\vec{v}_1), \ldots, T(\vec{v}_p)) = r < p$ so $\{T(\vec{v}_1), \ldots, T(\vec{v}_p)\}$ is NOT l.i. Again, by Theorem 2 this means $\mathcal{N}(T) \neq \{\mathbb{0}_{\mathbb{V}}\}$.

In particular $0 < \operatorname{nullity}(T) < p$ & we can pick a basis for $\mathcal{N}(T)$

Pick $\{\vec{w}_1, \ldots, \vec{w}_r\}$ basis for $\mathcal{R}(T)$

$\{\vec{v}_1, \ldots, \vec{v}_d\}$ —— $\mathcal{N}(T)$

Write $\vec{w}_1 = T(\vec{u}_1)$, ..., $\vec{w}_r = T(\vec{u}_r)$ for some $\vec{u}_1, \ldots, \vec{u}_r$ in $\mathbb{V}$

Claim: $S = \{\vec{u}_1, \ldots, \vec{u}_r, \vec{v}_1, \ldots, \vec{v}_d\}$ is a basis for $\mathbb{V}$

From here we get $p = r + d$. (Total # of elements in any basis for $\mathbb{V}$)

(1) S is l.i.:

$$(*) \quad \mathbb{0}_{\mathbb{V}} = \alpha_1 \vec{u}_1 + \cdots + \alpha_r \vec{u}_r + \beta_1 \vec{v}_1 + \cdots + \beta_d \vec{v}_d \qquad \text{Apply } T$$

$$\downarrow T$$

$$\mathbb{0}_{\mathbb{W}} = \alpha_1 T(\vec{u}_1) + \cdots + \alpha_r T(\vec{u}_r) + \beta_1 \mathbb{0}_{\mathbb{W}} + \cdots + \beta_d \mathbb{0}_{\mathbb{W}}$$

So $0_W = \alpha_1 \vec{w}_1 + \cdots + \alpha_r \vec{w}_r$ But $\{\vec{w}_1, \ldots, \vec{w}_r\}$ is li

so $\boxed{\alpha_1 = \cdots = \alpha_r = 0}$

Now, replace this back in (*) to get an equation only involving the β's.

$0_V = \beta_1 \vec{v}_1 + \cdots + \beta_d \vec{v}_d$ but $\{\vec{v}_1, \ldots, \vec{v}_d\}$ is li so $\boxed{\beta_1 = \cdots = \beta_d = 0}$

Combine the boxed expressions to conclude S is li.

(2) S spans V:

Pick any $\vec{v}$ in V & apply T: Then $T(\vec{v})$ is in $R(T)$, so we can write it as $T(\vec{v}) = \alpha_1 \vec{w}_1 + \cdots + \alpha_r \vec{w}_r$

$$= \alpha_1 T(\vec{u}_1) + \cdots + \alpha_r T(\vec{u}_r) = T(\alpha_1 \vec{u}_1 + \cdots \alpha_r \vec{u}_r)$$

Conclude $\vec{v}$ & $\vec{u} = \alpha_1 \vec{u}_1 + \cdots + \alpha_r \vec{u}_r$ have the same image under T

By Proposition 1, $\vec{v} - \vec{u}$ lies in $N(T)$.

Hence, we get $\vec{v} - \vec{u} = \beta_1 \vec{v}_1 + \cdots + \beta_d \vec{v}_r$

$$\rightsquigarrow \vec{v} = \vec{u} + \beta_1 \vec{v}_1 + \cdots + \beta_d \vec{v}_d$$

$$= \alpha_1 \vec{u}_1 + \cdots + \alpha_r \vec{u}_r + \beta_1 \vec{v}_1 + \cdots + \beta_d \vec{v}_d$$

So $\vec{v}$ lies in $Sp(S)$

Conclusion: S spans V. $\square$

Upshot of the proof: The sets $V_{w_1}, \ldots, V_{w_r}$ describe V

$$V_{w_1} = u_1 + N(T)$$
$$\vdots$$
$$V_{w_r} = u_r + N(T)$$

(see aside Observation from early)

$$\boxed{V_{\alpha_1 \vec{w}_1 + \cdots + \alpha_r \vec{w}_r} = (\alpha_1 \vec{u}_1 + \cdots + \alpha_r \vec{u}_r) + N(T)}$$