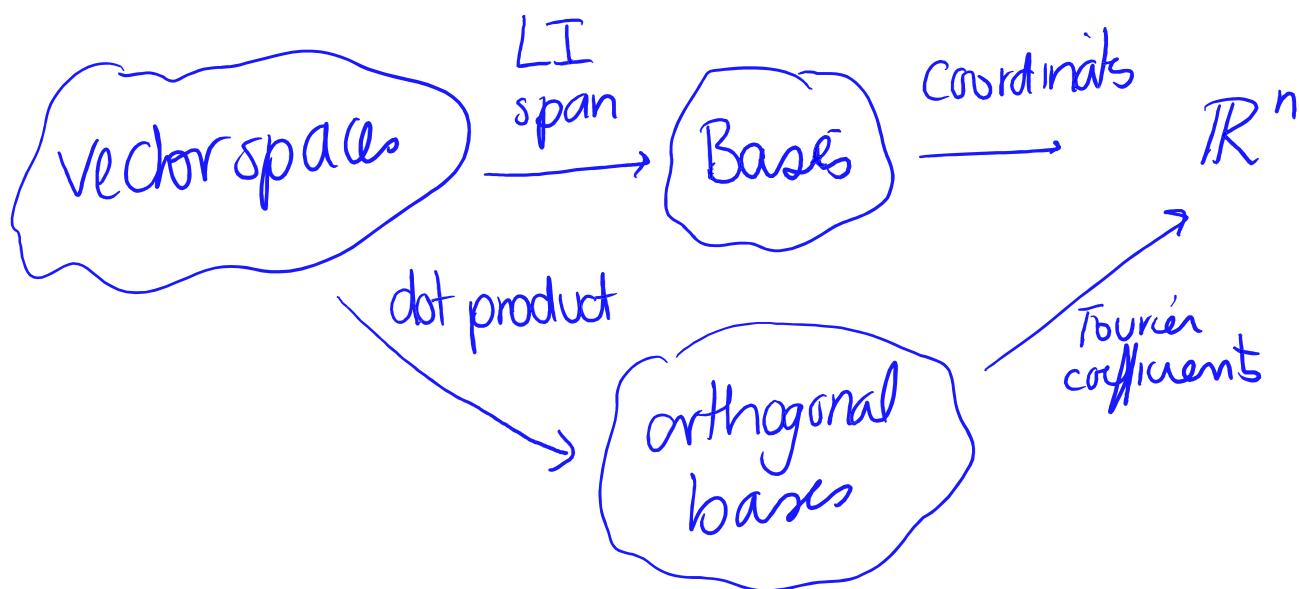


Chapter 24: Linear transformations

So: we have learned about vector spaces
and bases for vector spaces
and matrices.



matrices showed up:

- to solve linear systems
- to solve vector equations
- as algorithms to solve for bases
- as algebraic objects "generalized numbers"
you can add, multiply & invert

then

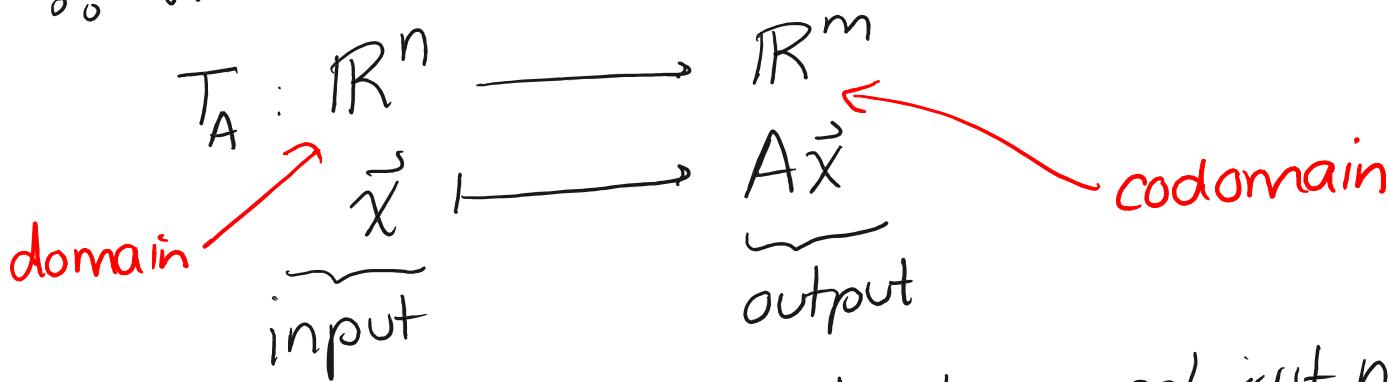
- as functions acting on vectors
(so that eigenvectors are the most special)

Let's elaborate on this last point, which is key.

So let A be an $m \times n$ matrix.

Then if $\vec{x} \in \mathbb{R}^n$, $A\vec{x} \in \mathbb{R}^m$.

∴ we can define a function



Because the inputs and outputs are not just numbers, but rather vectors, we call this a **transformation** rather than function (for historical reasons).

But it is a special transformation:

$$\begin{aligned} & \cdot A(\vec{0}) = \vec{0} \\ & \cdot A(\vec{x} + \vec{y}) = A\vec{x} + A\vec{y} \\ & \cdot A(c\vec{x}) = cA\vec{x} \end{aligned} \quad \left. \right\} \text{A preserves the operations on a vector space}$$

A function or transformation satisfying these three properties is called **linear**.

So matrix multiplication is a linear transformation from \mathbb{R}^n to \mathbb{R}^m .

What does linearity give us?

Say $\{\vec{u}, \vec{v}\}$ is a basis.

Suppose you have calculated

$$A\vec{u} = \vec{w}$$

$$A\vec{v} = \vec{z}$$

then for any $\vec{x} = a\vec{u} + b\vec{v}$ we have

$$\begin{aligned} A(a\vec{u} + b\vec{v}) &= a(A\vec{u}) + b(A\vec{v}) \\ &= a\vec{w} + b\vec{z}. \end{aligned}$$

So if you know the vectors $A\vec{u}$ & $A\vec{v}$ you know everything. No surprises. Everything about your linear transformation is known as soon as you know where it sends a basis.

Def'n: Let U and W be vector spaces. A linear transformation T is a map from U to W satisfying

- 1) for all $\vec{u}, \vec{v} \in U$, $T(\vec{u} + \vec{v}) = T(\vec{u}) + T(\vec{v})$
- 2) for all $\vec{u} \in U$, $c \in \mathbb{R}$, $T(c\vec{u}) = cT(\vec{u})$

Example: if A is an $m \times n$ matrix then

$T(\vec{u}) = A\vec{u}$ is the formula of a linear transformation from $U = \mathbb{R}^n$ to $W = \mathbb{R}^m$.

What other examples do we know?

eg) Let $U = W = F(\mathbb{R})$.

Then $T(f) = \frac{d}{dx}f$ is a linear transformation from $F(\mathbb{R})$ to $F(\mathbb{R})$ because:

$$\bullet \frac{d}{dx}(f+g) = \frac{d}{dx}f + \frac{d}{dx}g$$

$$\bullet \frac{d}{dx}(cf) = c \frac{d}{dx}f \quad \text{if } c \text{ is a real number}$$

eg) Let U be a vector space with basis $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\} = B$ and $W = \mathbb{R}^3$.

Then $T(a\vec{u}_1 + b\vec{u}_2 + c\vec{u}_3) = (a, b, c)$

(the coordinate map relative to this basis)

$$\text{or: } T(\vec{x}) = [\vec{x}]_B$$

is a linear transformation from U to \mathbb{R}^3 because:

$$\bullet T((a\vec{u}_1 + b\vec{u}_2 + c\vec{u}_3) + (d\vec{u}_1 + e\vec{u}_2 + f\vec{u}_3))$$

$$= T((a+d)\vec{u}_1 + (b+e)\vec{u}_2 + (c+f)\vec{u}_3)$$

$$= (a+d, b+e, c+f) = T(a\vec{u}_1 + b\vec{u}_2 + c\vec{u}_3) + T(d\vec{u}_1 + e\vec{u}_2 + f\vec{u}_3)$$

$$\begin{aligned}
 T(e(au_1 + bu_2 + cu_3)) &= T(ea u_1 + eb u_2 + ec u_3) \\
 &= (ea, eb, ec) \\
 &= e(a, b, c) \\
 &= e T(a \bar{u}_1 + b \bar{u}_2 + c \bar{u}_3).
 \end{aligned}$$

eg (non linear)

$$U = \mathbb{R}^2, W = \mathbb{R}^2$$

$$f(x, y) = (x^2 + y^2, x^2 - y)$$

is not a linear transformation because,

for example:

$$f(1, 0) = (1, 1) \quad \text{but } f(2, 0) = (4, 4) \neq 2f(1, 0).$$

eg) (non linear)

$$U = P_2 \quad W = P_4$$

$$T(f(x)) = f(x)^2$$

$$\begin{aligned} \text{eg: } T(x) &= x^2 \\ T(x^2) &= x^4 \end{aligned}$$

is not linear because

$$T(x+x^2) = (x+x^2)^2 = x^2 + 2x^3 + x^4 \neq T(x) + T(x^2).$$

$$\text{eg (linear)} \quad U = P_2 \quad W = P_3$$

$T(p(x)) = xp(x)$ is linear because

$$\begin{aligned}
 T(p(x) + q(x)) &= x(p(x) + q(x)) = xp(x) + xq(x) \\
 &= T(p(x)) + T(q(x)),
 \end{aligned}$$

$$\begin{aligned}
 \text{eg: } T(4+x) &= 4x + x^2 \\
 T(x+x^2) &= x^2 + x^3
 \end{aligned}$$

$$\begin{aligned}
 T(c p(x)) &= x(c p(x)) \\
 &= c(x p(x)) \\
 &= c T(p(x)).
 \end{aligned}$$

e.g) Let $U=W=\mathbb{R}^n$.
Suppose V is a subspace of \mathbb{R}^n .
Then orthogonal projection onto V is a linear transformation (whose image is V).

Why? Recall the formula:

V : orthogonal basis $\{v_1, \dots, v_k\}$

then

$$\text{proj}_V(\vec{u}) = \frac{\vec{u} \cdot \vec{v}_1}{\|\vec{v}_1\|^2} \vec{v}_1 + \dots + \frac{\vec{u} \cdot \vec{v}_k}{\|\vec{v}_k\|^2} \vec{v}_k$$

We recall:

$$(\vec{u} + \vec{w}) \cdot \vec{v}_i = \vec{u} \cdot \vec{v}_i + \vec{w} \cdot \vec{v}_i$$

$$\begin{aligned}
 \Rightarrow \frac{(\vec{u} + \vec{w}) \cdot \vec{v}_i}{\|\vec{v}_i\|^2} v_i &= \frac{\vec{u} \cdot \vec{v}_i + \vec{w} \cdot \vec{v}_i}{\|\vec{v}_i\|^2} v_i \\
 &= \frac{\vec{u} \cdot \vec{v}_i}{\|\vec{v}_i\|^2} \vec{v}_i + \frac{\vec{w} \cdot \vec{v}_i}{\|\vec{v}_i\|^2} \vec{v}_i
 \end{aligned}$$

$$(c \vec{u}) \cdot \vec{v}_i = c(\vec{u} \cdot \vec{v}_i)$$

$$\Rightarrow \frac{(c \vec{u}) \cdot \vec{v}_i}{\|\vec{v}_i\|^2} v_i = c \left(\frac{\vec{u} \cdot \vec{v}_i}{\|\vec{v}_i\|^2} v_i \right)$$

It follows (exercise) that

- $\text{proj}_V(\vec{u} + \vec{w}) = \text{proj}_V(\vec{u}) + \text{proj}_V(\vec{w})$
- $\text{proj}_V(c\vec{u}) = c \text{proj}_V(\vec{u})$.

Big theorem on linear transformations:

- a) Suppose $T: U \rightarrow W$ is a linear transformation and $\{\vec{u}_1, \dots, \vec{u}_n\}$ is a basis for U . Then T is completely determined by $T(\vec{u}_1), \dots, T(\vec{u}_n)$.
- b) Suppose $\{u_1, \dots, u_n\}$ is a basis for a vector space U and $\{w_1, \dots, w_n\}$ is ANY set of vectors in a vector space W . Then there is exactly one linear transformation $T: U \rightarrow W$ satisfying $T(u_1) = w_1, \dots, T(u_n) = w_n$.

Proof:

- a) By "completely determined" we mean: the value of $T(\vec{u})$, for any \vec{u} , is known.
- Let $\vec{u} \in U$. Then $\vec{u} = a_1\vec{u}_1 + \dots + a_n\vec{u}_n$.
By linearity of T , $T(\vec{u}) = a_1T(\vec{u}_1) + \dots + a_nT(\vec{u}_n)$.
So there are no choices left to make
- b) Define T by $T(a_1\vec{u}_1 + \dots + a_n\vec{u}_n) = a_1w_1 + \dots + a_nw_n$.
Since $\{u_1, \dots, u_n\}$ is a basis of U , this gives a unique formula

for $T(U)$. (ex) T is a linear transformation. \square

So: it seems like we only need to know what T does on a basis to fully describe T . How could we do so efficiently?

Thm (the matrix of a linear transformation)

Let $\{u_1, \dots, u_n\}$ be a basis for U & $\{w_1, \dots, w_m\} = B$ a basis for W . If $T: U \rightarrow W$ is a linear transformation then the matrix of T relative to these bases is

$$A = \begin{bmatrix} [T(u_1)]_B & [T(u_2)]_B & \dots & [T(u_n)]_B \end{bmatrix}.$$

Special case: $U = \mathbb{R}^n$, $W = \mathbb{R}^m$.

The standard matrix of a linear transformation

is

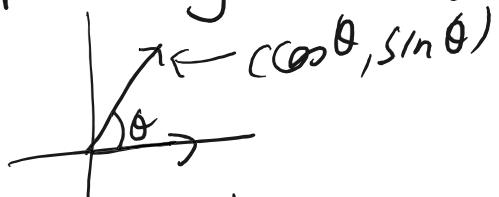
$$A = \begin{bmatrix} T(e_1) & T(e_2) & \dots & T(e_n) \end{bmatrix}$$

$\uparrow \quad \uparrow \quad \uparrow$
column vectors.

example: If T is multiplication by A then the standard matrix of T is A .

eg) If T is rotation of the plane by θ (angle)

then $T\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix}$



$$T\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -\sin \theta \\ \cos \theta \end{pmatrix}$$

$$\therefore A = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$$

g) If T is the state machine of a dynamical system, A is the matrix with transition probabilities.

eg) Consider the subspace

$$U = \text{span} \{ 1, \underbrace{\sin(x), \cos(x)}_{\text{basis for this subspace}} \} \subseteq F(\mathbb{R}).$$

Then

$$\frac{d}{dx} 1 = 0 \Rightarrow [0]_B = (0, 0, 0)$$

$$\frac{d}{dx} \sin(x) = \cos(x) : \text{coordinates } (0, 0, 1)$$

$$\frac{d}{dx} \cos(x) = -\sin(x) : \text{coordinates } (0, -1, 0).$$

\therefore The matrix of $\frac{d}{dx}$ on U is

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}$$

funny: its eigenvalues
are $0, i, -i \dots$

linear transformations are to vector spaces

as

differentiable functions are to Calculus (\mathbb{R})