Linear Transformations

(Sections 2.6, 7.1)

Recall: matrix-vector multiplication

Given a vector $\vec{v} \in \mathbb{R}^n$ as a column, you can multiply $\vec{v}$ by an $m \times n$ matrix A to get a vector in $\mathbb{R}^m$.

A takes vectors in $\mathbb{R}^n$ to vectors in $\mathbb{R}^m$, and A distributes over addition and scalar multiplication.

Generalize to arbitrary vector spaces!

Linear Transformations

(or Linear maps, Linear operators)

Start with V, W vector spaces.

A linear transformation T is a function from V to W (written "$T: V \to W$") such that

$\forall\ x, y \in V$ and $c \in \mathbb{R}$,

$$\boxed{\begin{aligned} T(x+y) &= T(x) + T(y) \\ T(c \cdot x) &= c \cdot T(x) \end{aligned}}$$

The point: T distributes over addition and scalar multiplication

Example 1: (matrices) Recall that,

since all matrices distribute over addition and scalar multiplication in $\mathbb{R}^n$, every $m \times n$ matrix gives a linear transformation $T_A : \mathbb{R}^n \to \mathbb{R}^m$ defined by

$$\boxed{T_A(x) = A \cdot x}$$

for $x \in \mathbb{R}^n$ and A an $m \times n$ matrix.

The Converse: Every linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ can be represented as a matrix A_T! Think of $\mathbb{R}^n, \mathbb{R}^m$ as column vectors.

Let $\{e_1, e_2, \dots, e_n\}$ be the standard basis vectors of $\mathbb{R}^n$.

1^{st} column of $A_T = T(e_1) \in \mathbb{R}^m$

2^{nd} column of $A_T = T(e_2) \in \mathbb{R}^m$

$\vdots$

n^{th} column of $A_T = T(e_n) \in \mathbb{R}^m$

Show $T(x) = A_T(x)$ for all $x \in \mathbb{R}^n$.

Let $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$

$$= \begin{bmatrix} x_1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ x_2 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \cdots + \begin{bmatrix} 0 \\ \vdots \\ 0 \\ x_n \end{bmatrix}$$

$$= x_1 \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \cdots + x_n \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

$$= x_1 e_1 + x_2 e_2 + \cdots + x_n e_n$$

$$= \sum_{i=1}^{n} x_i e_i$$

$$T(x) = T\left(\sum_{i=1}^{n} x_i e_i\right)$$

$$= T(x_1 e_1 + x_2 e_2 + \cdots + x_n e_n)$$

Since T is a linear transformation

$$= T(x_1 e_1) + T(x_2 e_2) + \cdots + T(x_n e_n)$$

$$= x_1 T(e_1) + x_2 T(e_2) + \cdots + x_n T(e_n)$$

$$= x_1 \cdot (1^{st} \text{ column of } A_T)$$

$$+ x_2 (2^{nd} \text{ column of } A_T)$$

$$+ \vdots$$

$$+ x_n (n^{th} \text{ column of } A_T)$$

But the i^{th} column of A_T can be recovered as $A_T(e_i)$ for $1 \leq i \leq n$!

So $T(x) = x_1 \cdot (1^{st} \text{ column of } A_T)$

$$+ x_2 (2^{nd} \text{ column of } A_T)$$

$$+ \vdots$$

$$+ x_n (n^{th} \text{ column of } A_T)$$

$$= x_1 A_T(e_1)$$

$$+ x_2 A_T(e_2)$$

$$\vdots$$

$$+ x_n A_T(e_n)$$

all matrices are linear transformations

$$= A_T(x_1 e_1) + A_T(x_2 e_2) + \cdots + A_T(x_n e_n)$$

$$= A_T(x_1 e_1 + x_2 e_2 + \cdots + x_n e_n)$$

$$= A_T \cdot x$$

Moral: Linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^m$ are the same as $m \times n$ matrices.

Example 2: (zero and identity maps)

Let V and W be any vector spaces.

Then there always exists one linear transformation from V to W: the zero map!

$$T(x) = 0_W \quad \forall\, x \in V.$$

Let's check this is linear: let $y \in V$, $c \in \mathbb{R}$. Then

$$\begin{aligned} T(x) + T(y) &= 0_W + 0_W \\ &= 0_W \\ &= T(x+y) \end{aligned}$$

$$T(c \cdot x) = 0_W$$

$$= c \cdot T(x)$$

since $T(x) = 0_W$!

T is a linear transformation, then.

From V to V, we also have

the identity transformation I,

$$I(x) = x \quad \forall \; x \in V.$$

Check this is linear: Let $y \in V$, $c \in \mathbb{R}$.

$$I(x+y) = x+y = I(x) + I(y)$$

$$I(c \cdot x) = c \cdot x = c \cdot I(x)$$

I is a linear transformation from V to V.

zero transformation = analog of the zero matrix

identity transformation = " " " identity matrix

Shorter Way to Check that a Function is a Linear Transformation!

If V and W are vector spaces and $T: V \to W$, then T is a linear transformation precisely when

$$\boxed{T(c \cdot x + y) = c \cdot T(x) + T(y)}$$

$$\forall\, x, y \in V,\ c \in \mathbb{R}$$

Example 3: (limits on sequence space)

Let $V = \mathcal{S}$, the vector space of all sequences of real numbers. Let $W \subseteq \mathcal{S}$ be the subspace of all sequences $(a_n)_{n=1}^{\infty} \in \mathcal{S}$ such that $\lim_{n\to\infty} a_n$ exists.

Show W is a subspace

Define $T: W \to \mathbb{R}$,

$$\boxed{T\left((a_n)_{n=1}^{\infty}\right) = \lim_{n\to\infty} a_n}$$

Let's show T is a linear transformation.

From calculus, if $(a_n)_{n=1}^{\infty}$ and $(b_n)_{n=1}^{\infty}$ are sequences and $\lim_{n\to\infty} a_n = L$, $\lim_{n\to\infty} b_n = M$ and $c \in \mathbb{R}$,

$$\lim_{n\to\infty} (a_n + b_n) = L + M$$

$$\lim_{n\to\infty} (c \cdot a_n) = c \cdot L$$

Suppose that $(a_n)_{n=1}^{\infty}, (b_n)_{n=1}^{\infty} \in W$ with limits $L, M \in \mathbb{R}$, respectively.

Let $c \in \mathbb{R}$.

Then

$$T\left(c\,(a_n)_{n=1}^{\infty} + (b_n)_{n=1}^{\infty}\right)$$

$$= T\left((ca_n)_{n=1}^{\infty} + (b_n)_{n=1}^{\infty} \right)$$

$$= T\left((ca_n + b_n)_{n=1}^{\infty} \right) \leftarrow$$

$$= \lim_{n\to\infty} (ca_n + b_n)$$

$$= c \cdot L + M \quad \text{(calculus rules)}$$

$$= c \cdot \lim_{n\to\infty} a_n + \lim_{n\to\infty} b_n$$

$$= c \cdot T\left((a_n)_{n=1}^{\infty}\right) + T\left((b_n)_{n=1}^{\infty}\right) \checkmark$$

So T is a linear transformation!

Example 4: (a non-linear transformation)

Define $f: \mathbb{R}^3 \to \mathbb{R}$,

$$f\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{cases} x+y+z, & y \neq 0 \\ x, & y = 0. \end{cases}$$

Show f is not linear.

Solution: What we need: either two explicit vectors $\begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix}$ and $\begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix}$ such that f does not distribute over their sum OR an explicit vector $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ and

a scalar $c \in \mathbb{R}$ with

$$f\left(c \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) \neq c f\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right)$$

Try to come up with examples that break distributivity over sums:

$$\begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix}$$

$$\begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ 16 \end{bmatrix}$$

$$\begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix}$$

$$\begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ 16 \end{bmatrix}$$

The sum of these vectors is

$$\begin{bmatrix} 3 \\ 0 \\ 19 \end{bmatrix}.$$

$$f\left(\begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix}\right) = 1+1+3 = 5$$

$$f\left(\begin{bmatrix} 2 \\ -1 \\ 16 \end{bmatrix}\right) = 2-1+16 = 17$$

$$f\left(\begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix}\right) + f\left(\begin{bmatrix} 2 \\ -1 \\ 16 \end{bmatrix}\right) = 22$$

But

$$f\left(\begin{bmatrix}1\\1\\3\end{bmatrix}+\begin{bmatrix}2\\-1\\16\end{bmatrix}\right)$$

$$= f\left(\begin{bmatrix}3\\0\\19\end{bmatrix}\right) = 3 \neq 22$$

So f is not linear!

Note: f does distribute over scalar multiplication!

Example 4 : (integrals) Let V denote the vector space of all functions defined on $[0,1]$ with real value outputs:

$$V = \text{all functions } f:[0,1] \to \mathbb{R}.$$

The addition and scalar multiplication is the same as $\mathcal{F}(\mathbb{R})$.

Let $W \subseteq V$ be the subspace consisting of all continuous $f:[0,1] \to \mathbb{R}$.

Recall this means

$$\lim_{x \to a} f(x) = f(a) \quad \forall \quad a \in (0,1),$$

$$\lim_{x \to 1^-} f(x) = f(1), \quad \text{and}$$

$$\lim_{x \to 0^+} f(x) = f(0).$$

Define $T: W \to \mathbb{R}$ by

$$\boxed{T(f) = \int_0^1 f(x)\,dx}.$$

Since f is continuous, we know the integral exists (Calc I).

Show T is a linear transformation.

Let $f, g \in W$ and $c \in \mathbb{R}$. Then

$$T(cf+g) = \int_0^1 ((c\cdot f+g)(x))\,dx$$

(how we add functions in V) $$= \int_0^1 ((c\cdot f)(x) + g(x))\,dx$$

(how we scalar multiply in V) $$= \int_0^1 (c\cdot f(x) + g(x))\,dx$$

(integrals distribute over addition) $$= \int_0^1 c\cdot f(x)\,dx + \int_0^1 g(x)\,dx$$

(integrals pull out constants) $$= c\int_0^1 f(x)\,dx + \int_0^1 g(x)\,dx$$

$$= c\cdot T(f) + T(g)$$ ✓

This shows T is a linear transformation.

Observe that we need W to be a vector space: sums and scalar multiples of continuous functions are continuous. The usual notation for W is $C([0,1])$.

Remark: (derivatives) Let $V = \mathcal{F}(\mathbb{R})$ and let W be the subspace of $\mathcal{F}(\mathbb{R})$ consisting of all differentiable functions:

$f \in W$ if

$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a} \text{ exists}$$

$\forall\, a \in \mathbb{R}$. We call this number $f'(a)$.

If $f, g \in W$, $c \in \mathbb{R}$

from Calc I
$$\begin{cases} (f+g)'(a) = f'(a) + g'(a) \\ (c \cdot f)'(a) = c \cdot f'(a). \end{cases}$$

This shows W is a subspace of $\mathcal{F}(\mathbb{R})$.

Let's try to define a linear transformation T on W by

$$T(f) = f' \quad \forall f \in W.$$

$$T : W \to ?$$

The map T does <u>not</u> send differentiable functions to differentiable functions: let $f(x) = \int_0^x |t|\,dt$.

By the Fundamental Theorem of Calculus,

$f'(x) = |x|$, not differentiable at $x = 0$!

There are examples of functions

$g: \mathbb{R} \to \mathbb{R}$ such that

1) g is continuous $\forall\, a \in \mathbb{R}$

2) g is NOT differentiable for any $a \in \mathbb{R}$.

So setting $f(x) = \int_0^x g(t)\,dt$,

$f'(x) = g(x)$, not differentiable anywhere!

What should T map to?

Example 5: (point evaluation on $\mathcal{F}(\mathbb{R})$)

Fix $c \in \mathbb{R}$ and define

$T_c : \mathcal{F}(\mathbb{R}) \to \mathbb{R}$ by

$$\boxed{T_c(f) = f(c)}$$

So for example:

$$T_2(f) = f(2)$$

$$T_0(f) = f(0)$$

$$T_\pi(f) = f(\pi) \text{ etc.}$$

Let's show T_c is a linear transformation:

take $f, g \in \mathcal{F}(\mathbb{R})$ and $d \in \mathbb{R}$.

Then for all $c \in \mathbb{R}$,

$$T_c(df+g) = (df+g)(c)$$

(addition in $\mathcal{F}(\mathbb{R})$) $\quad = (df)(c) + g(c)$

(scalar multiplication in $\mathcal{F}(\mathbb{R})$) $\quad = d \cdot f(c) + g(c)$

$$= d \cdot T_c(f) + T_c(g)$$

So T is a linear transformation.

Observation: (zero goes to zero)

Let V and W be vector spaces and let $T: V \to W$ be a linear transformation. Then

$$\underline{T(0_V) = 0_W}$$

Why? $T(0_V) = T(0_V + 0_V)$

$T(0_V) = T(0_V) + T(0_V)$
(T distributes over sums)

Every $x \in V$ has an additive inverse. Let $-T(0_V)$ denote the additive inverse of $T(0_V)$.

Starting with

$$T(0_v) = T(0_v) + T(0_v),$$

add $-T(0_v)$ to both sides.

$$\underbrace{\cancel{T(0_v) + (-T(0_v))}}_{0_w} = T(0_v) + \underbrace{\cancel{T(0_v) + (-T(0_v))}}_{0_w}$$

$$0_w = T(0_v) + 0_w = T(0_v) \checkmark$$

This fact is yours to use for the rest of class.

Two Subspaces Associated to a Linear Transformation

Let V and W be vector spaces and let $T: V \to W$ be a linear transformation.

There are two collections of vectors associated to T: the kernel of T, denoted $\ker(T)$, and the range of T, denoted by $\mathrm{Ran}(T)$, where

$$\boxed{\begin{aligned} \ker(T) &= \{ x \in V \mid T(x) = 0_W \} \\ \mathrm{Ran}(T) &= \{ y \in W \mid \exists\, x \in V,\ T(x) = y \} \end{aligned}}$$

Here, the notation

$\{ P \mid Q \}$ means

such that

" the collection of all P such that Q "

Then $\ker(T)$ is a subspace of V

and $\mathrm{Ran}(T)$ is a subspace of W

Why? We'll only show $\ker(T)$ is a subspace. We need to show

1) $0_V \in \ker(T)$

2) If $x, z \in \ker(T)$ and $c \in \mathbb{R}$, then $c \cdot x + z \in \ker(T)$.

1) Since T is a linear transformation, we showed $T(0_V) = 0_W$. This gives us that $0_V \in \ker(T)$.

2) Let $x, z \in \ker(T)$ and $c \in \mathbb{R}$.

Then x any y satisfy the defining property of $\ker(T)$:

$$T(x) = 0_W$$

$$T(z) = 0_W$$

Then

$$T(c \cdot x + z) = c \cdot T(x) + T(z) \quad (T \text{ is linear})$$

$$= c \cdot 0_W + 0_W$$

$$= 0_W + 0_W$$

$$= 0_W \checkmark$$

So $c \cdot x + z \in \ker(T)$ and $\ker(T)$ is a subspace of V!

Subspaces Associated to Matrices

Let A be an $m \times n$ matrix.

We have two subspaces associated to A that don't show up for general linear transformations: the column space of A, denoted $\mathrm{col}(A)$, and the row space of A, denoted $\mathrm{row}(A)$, where

$$\boxed{\begin{array}{l} \mathrm{col}(A) = \text{all linear combinations of the columns of } A \\ \mathrm{row}(A) = \text{'' \quad '' \quad '' \quad '' \quad '' rows '' \quad ''} \end{array}}$$

In fact, since

$$Ae_i = i^{th} \text{ column of } A$$
$$\forall \; 1 \le i \le n,$$

$$\mathrm{col}(A) = \mathrm{Ran}(A).$$

We have $\mathrm{row}(A) = \mathrm{Ran}(A^t)$.

Finally, $\ker(A)$ is sometimes called the nullspace of A, and denoted $\mathrm{null}(A)$. The book sometimes calls $\mathrm{Ran}(A)$ the image of A, denoted $\mathrm{im}(A)$.

Refinement of derivative question:

W is the vector space of differentiable functions on $\mathbb{R}$, $T(f) = f'$ for $f \in W$

What is $\mathrm{Ran}(T)$?