MTH 372 Practical Machine Learning

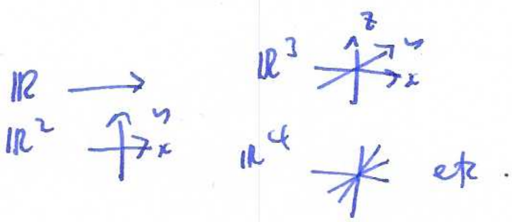
Joseph Maher joseph.maher@csi.cuny.edu
web page: http://www.math.csi.cuny.edu/~maher
office 1S-222 office hours M 12:20 – 2:15 (in 1S-222) W 12:20 – 1:10 (in 1S-214)

— students w/ disability ← make sure I know what to do

(Review of) Linear Algebra

Vector spaces

Example: $\mathbb{R}^n = \{(x_1, x_2, \ldots, x_n) \mid x_i \in \mathbb{R}\}$ e.g. $\mathbb{R}$, $\mathbb{R}^2$, $\mathbb{R}^3$, $\mathbb{R}^4$ etc.



Defn A set V is a vector space over $\mathbb{R}$ if there are operations of addition $+: V \times V \to V$, $(v,w) \mapsto v+w$

and scalar multiplication $\mathbb{R} \times V \to V$, $(a,v) \mapsto a \cdot v$

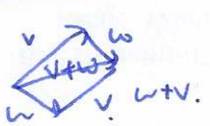
such that:
a) additive ~~inv~~ identity: there is an element $0 \in V$ s.t. $0 + v = v$ for all $v \in V$
b) additive inverse: for all $v \in V$, there is $(-v) \in V$ s.t. $v + (-v) = 0$
c) commutative: for all $v, w \in V$, $v + w = w + v$
d) distributive: for all $a \in \mathbb{R}$ and $v, w \in V$, $a(v+w) = av + aw$
e) associative: for all $a, b \in \mathbb{R}$, $v \in V$ $a(bv) = (ab)v$
f) distributive: for all $a, b \in \mathbb{R}$, $v \in V$ $(a+b)v = av + bv$
g) multiplicative identity: for all $v \in V$, $1v = v$

notation $v \in V$ is a vector. $a \in \mathbb{R}$ is a scalar.

warning: can't multiply vectors! vw doesn't make sense :(.

geometric interpretation: a vector has length and direction.

• addition



• scalar mult:

exception $0 \in V$ has length 0, direction undefined.

warning: 0 ambiguous: 0 scalar or 0 vector?

Examples of vector spaces $\mathbb{R}^n$, set of polynomials of degree d, set of all polynomials, set of all functions $f: \mathbb{R} \to \mathbb{R}$ (last two not finite dimensional) set of sequences $(x_1, x_2, \ldots)$

Def^n A **linear map/transformation** is a function $T: V \to W$ s.t. for all $a, b \in \mathbb{R}$, $v, w \in V$,

~~$T(av + bw) = aT(v) + bT(w)$.~~

~~Equivalently~~, T preserves addition $T(v+w) = T(v) + T(w)$
and scalar multiplication $T(av) = aT(v)$

Examples : P_d = set of polynomials of deg d. $\begin{aligned} \mathcal{E}: P_d &\to \mathbb{R} \\ p &\mapsto p(0) \end{aligned}$ is a linear map. (check!)

Q: what about $\begin{aligned} \alpha: P_d &\to \mathbb{R} \\ p &\mapsto p(2) \end{aligned}$?
evaluation at 2

Q: what about change of coordinates $\begin{aligned} p(x) &\mapsto p(x+1) \\ p(x) &\mapsto p(2x) \end{aligned}$?

Q: what about $p(x) \mapsto p(x^2)$?

Example differentiation : $\begin{aligned} D: P_d &\to P_{d-1} \\ p &\mapsto \frac{dp}{dx} \end{aligned}$ (check!) Q: what about integration?

Q: what about replacing P_d with $C = \{ f: \mathbb{R} \to \mathbb{R} \}$?

Example shift map on sequences $(x_1, x_2, x_3, \ldots) \mapsto (x_2, x_3, \ldots)$

Example geometric transformations of $\mathbb{R}^2$.

- expansion. $v \mapsto 2v$
- rotation 

Q: what maps are not linear?

Example let $H(V, W) = \{$set of linear maps from V to $W\}$.
check: this is a vector space! $S, T: V \to W$ with: $S(v) + T(v) = (S+T)(v)$.
$aS(v) = S(av)$.

Def^n A **matrix** is an array of numbers $A = \underbrace{\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & & & \vdots \\ a_{m1} & \cdots & & a_{mn} \end{bmatrix}}_{n \text{ cols.}} \Big\} m \text{ rows}$

Def^n matrix multiplication $\overset{m \times n}{A} v \leftarrow n \times 1$ $(Av) \leftarrow m \times 1$.

$$\underset{m \times n}{\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}} \underset{n \times 1}{\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}} = \underset{m \times 1.}{\begin{bmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{bmatrix}}$$

key fact matrix multiplication is a linear map. $A(v+w) = Av + Aw$.
$aA(v) = A(av)$.

Examples find matrices for the previous examples.

$T: P_2 \to \mathbb{R}$, $p \mapsto p(0)$. $p \leftrightarrow a_0 + a_1 x + a_2 x^2$ then $T(p) = a_0$. so $T = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}$

$\leftrightarrow \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix}$.

$T: P_2 \to P_2$, $p \mapsto p(2)$. $T(p) = a_0 + a_1 \cdot 2 + a_2 \cdot 4$ so $T = \begin{bmatrix} 1 & 2 & 4 \end{bmatrix}$.

$T: P_2 \to P_2$, $p(x) \mapsto p(x+1)$. $a_0 + a_1 x + a_2 x^2 \mapsto a_0 + a_1(x+1) + a_2(x+1)^2$

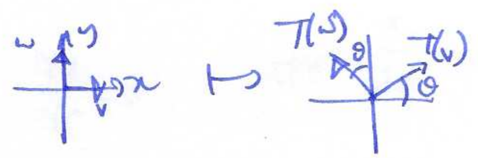
$$\begin{aligned} & a_0 + a_1 x + a_2 x^2 \\ & + a_1 \quad + 2a_2 x \\ & + a_2 \end{aligned}$$

$T = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix}$.

$T: P_3 \to P_2$, $p \mapsto \frac{dp}{dx}$. $a_0 + a_1 x + a_2 x^2 + a_3 x^3 \mapsto a_1 + 2a_2 x + 3a_3 x^2$.

$T = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix}$

Rotation in $\mathbb{R}^2$:



$T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$.

expansion/contraction: $\begin{bmatrix} 2 & 0 \\ 0 & 1/2 \end{bmatrix}$

Q: what does $\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$ do geometrically? (not clear — we will find out later).

Example let M_2 = 2x2 matrices (vector space!).

Q: is $A \mapsto A + A^T$ a linear map? can you find a matrix for it?

Q: what about $A \mapsto A^2$? (no).

Defn A subset $U \subseteq V$ is a subspace if U is a vector space (w/ same operations as V!)

Propn $U \subseteq V$ is a subspace iff it is closed under addition and scalar multiplication.

Proof (feel free to skip) $\Rightarrow$ if U is a subspace then $\underbrace{u+v \in U}_{\text{closed under addition}}$ and $\underbrace{au \in U}_{\text{closed under scalar mult.}}$ ✓.

$\Leftarrow$ defn of vector space $\square$.

Def^{n} let $T: V \to W$ be a linear map.

- the kernel of T $\ker(T) = \{v \in V \mid T(v) = 0\}$.
- the image of T $\mathrm{im}(T) = \{w \in W \mid \exists v \in V \text{ s.t. } T(v) = w\}$.

Prop^{n} ① $\ker(T) \subseteq V$ is a subspace
② $\mathrm{im}(T) \subseteq W$ is a subspace

Proof ① if $v, w \in \ker(T)$ then $T(v+w) = T(v) + T(w) = 0 + 0 = 0$
if $a \in \mathbb{R}$, $v \in \ker(T)$ then $T(av) = aT(v) = a.0 = 0$.

② if $w_1, w_2 \in \mathrm{im}(T)$ then $\exists$ v_1, v_2 s.t. $T(v_1) = w_1$ and $T(v_2) = w_2$
so $T(v_1 + v_2) = T(v_1) + T(v_2) = w_1 + w_2$ so $w_1 + w_2 \in \mathrm{im}(T)$.
if $a \in \mathbb{R}$, $w \in \mathrm{im}(T)$ then $\exists$ $v \in V$ s.t. $T(v) = w$, then $T(av_g) = aT(v_g) = aw$
so $aw \in \mathrm{im}(T)$. $\square$.

Example (aside): $V =$ $f: \mathbb{R} \to \mathbb{R}$ ~~dif~~ smooth (∞-differentiable).
then $T = \frac{d}{dx}: V \to V$ Q: what is $\ker(T)$? $\frac{df}{dx} = 0 \Rightarrow f =$ constant functions.

- consider a linear differential equation $\frac{d^2f}{dx^2} + 4\frac{df}{dx} - 7f = 0$. (*)

then $T = \frac{d^2}{dx^2} + 4\frac{d}{dx} + 7$ is a linear map $T: V \to V$ and solutions of (*) correspond to $\ker(T)$.

Bases, dimension, linear maps as matrices.

Def^{n} A $\boxed{\text{set}}$ of vectors $(v_1, v_2, \dots v_n)$ is a **basis** for V if for any $v \in V$ there are unique scalars $a_1, \dots, a_n$ s.t. $v = a_1v_1 + a_2v_2 + \dots + a_nv_n$.

warning not a set an ordered tuple. i.e. $(v_1, v_2) \neq (v_2, v_1)$.

Def^{n} the dimension of V, $\dim(V)$ is the number of elements in a basis for V.

Thm For any vector space V, all bases of V have the same number of elements.

Def^{n} standard basis of $\mathbb{R}^n$ is $\{\underbrace{(1,0,\dots)}_{e_1}, \underbrace{(0,1,0,\dots)}_{e_2}, \dots, \underbrace{(0,0,\dots 1)}_{e_n}\}$