1. Some Linear Algebra

Def: A vector space consists of a set V (vectors) along with a set F (scalars, usually $\mathbb{R}$ or $\mathbb{C}$) that is:

(i) Closed under addition: $v, w \in V \Rightarrow v+w \in V$

(ii) Closed under scalar multiplication: $v \in V,\ c \in F \Rightarrow cv \in V$.

Remark: We can deduce some easy consequences:

- $\vec{0} \in V$ (take $c=0$)
- $v \in V \Rightarrow -v \in V$ (take $c=-1$)

If $F=\mathbb{R}$, we say that V is a "real vector space," or an "$\mathbb{R}$-vector space," or a "vector space over $\mathbb{R}$."

A complex vector space is defined similarly (i.e., if $F=\mathbb{C}$)

* Unless specified otherwise, we will assume by default that $F=\mathbb{R}$.

Examples:

(i) $V=\mathbb{R}^n=\{(x_1,\dots,x_n) \mid x_i \in \mathbb{R}\}$ is a vector space.

Proof: "+": $(x_1,\dots,x_n)+(y_1,\dots,y_n)=(x_1+y_1,\dots,x_n+y_n) \in \mathbb{R}^n$ ✓

"$\cdot$": $c(x_1,\dots,x_n)=(cx_1,\dots,cx_n) \in \mathbb{R}^n$

(ii) $V = \mathbb{C}^n = \{ (z_1, \dots, z_n) \mid z_i \in \mathbb{C} \}$

(iii) $V = \mathrm{Poly}_n = \{ a_n x^n + \dots + a_1 x + a_0 \mid a_i \in \mathbb{R} \}$
(polynomials of degree $\leq n$).

(iv) $V = \mathbb{R}[x] = \{ a_0 + a_1 x + a_2 x^2 + \dots + a_k x^k \mid a_i \in \mathbb{R},\ k \in \mathbb{N} \}$.
(polynomials of arbitrary degree)

(v) $V = \mathbb{R}[[x]] = \{ a_0 + a_1 x + a_2 x^2 + \dots \mid a_i \in \mathbb{R} \}$
(formal power series)

(vi) $V = \mathcal{C}^1$ = (once) differentiable functions; $f'(x)$ continuous too

(vii) $V = \mathcal{C}^\infty$ = infinitely differentiable functions; $f^{(k)}(x)$ continuous

(viii) $V = \mathrm{Per}_{2\pi}$ = piecewise continuous functions with $f(x) = f(x + 2\pi)$,
i.e., period $T = \frac{2\pi}{n}$ for some $n \in \mathbb{N}$.

Non-examples:

(i) Polynomials with degree $= n$ $\quad [(x^n + 1) + (2 - x^n) = \underline{3}]$

(ii) The upper half-plane in $\mathbb{R}^2$ $\quad [-1 \cdot (0,1) = (0,-1)]$.

Def: If V is a vector space (over F), then a subspace is a subset $W \subseteq V$ that is also a vector space (over F).

Examples:

(i) Let $V = \{(x,y,z) \mid x,y,z \in \mathbb{R}\} \cong \mathbb{R}^3$

Let $W = \{(x,y,0) \mid x,y \in \mathbb{R}\} \cong \mathbb{R}^2$.

Then W is a subspace of V.

(ii) $\mathrm{Poly}_n \subseteq \mathbb{R}[x] \subseteq \mathbb{R}[[x]]$.

Poly_n is a subspace of both $\mathbb{R}[x]$ and $\mathbb{R}[[x]]$

$\mathbb{R}[x]$ is a subspace of $\mathbb{R}[[x]]$.

(iii) $\mathcal{C}^\infty$ is a subspace of $\mathcal{C}^1$.

Also, note that $\mathcal{C}^1 \supseteq \mathcal{C}^2 \supseteq \mathcal{C}^3 \supseteq \cdots \supseteq \mathcal{C}^\infty$.

Remark: Subspaces in $\mathbb{R}^n$ "look like" hyperplanes through the origin (why?).

Non-examples:

(i) Unit circle in $\mathbb{R}^2$ ($\subseteq \mathbb{R}^2$)

(ii) Polynomials of degree $= n$ ($\subseteq \mathrm{Poly}_n$)

(iii) Upper half-plane ($\subseteq \mathbb{R}^2$)

(iv) The plane $\{(x,y,1) : x,y \in \mathbb{R}\}$ ($\subseteq \mathbb{R}^3$)

(v) Piecewise continuous functions with period exactly 2π ($\subseteq \mathrm{Per}_{2\pi}$).

Def: A set $S \subseteq V$ is linearly independent if none of the vectors in S can be expressed as a linear combination of the others. If S is not linearly independent, then it is linearly dependent.

Remark: $S \subseteq V$ is linearly independent iff for any $v_1, \dots, v_n \in S$,

$$a_1 v_1 + \dots + a_n v_n = 0 \implies a_1 = a_2 = \dots = a_n = 0.$$

Example: Let $V = \mathbb{R}^3$, $S \subseteq V$.

(i) The set $S = \{v_1\}$ is linearly independent iff $v_1 \neq 0$.

(ii) The set $S = \{v_1, v_2\}$ is linearly independent iff v_1 & v_2 don't lie on the same line.

(iii) The set $S = \{v_1, v_2, v_3\}$ is linearly independent iff v_1, v_2, v_3 don't lie on the same plane.

(iv) The set $S = \{v_1, v_2, v_3, v_4\}$ is never linearly independent in $\mathbb{R}^3$.

Another example: Let $V = \mathrm{Poly}_3$

(i) $S = \{1, x, x^2\}$ is linearly independent

(ii) $S = \{1, x, x^2, x^3\}$ is linearly independent

(iii) $S = \{1, x, x^2, 1 + 3x - 4x^2\}$ is linearly dependent

(iv) $S = \{1, x, x^2, x^3 + 1\}$ is linearly independent.

<u>One more example</u>: let $V = \mathcal{C}^1$.

(i) $S = \{\cos t, \sin t\}$ is linearly independent.

<u>Reason</u>: If $C_1 \cos t + C_2 \sin t = 0$ (the zero function), then $C_1 = C_2 = 0$.

(ii) $S = \{e^{2t}, e^{3t}\}$ is linearly independent.

<u>Reason</u>: If $C_1 e^{2t} + C_2 e^{3t} = 0$, then $C_1 = C_2 = 0$.

(iii) $S = \{e^{2t}, e^{-2t}, \cosh 2t\}$ is linearly dependent.

<u>Reason</u>: $\cosh 2t = \frac{1}{2} e^{2t} + \frac{1}{2} e^{-2t}$

or equivalently: $\frac{1}{2} e^{2t} + \frac{1}{2} e^{-2t} - 1 \cosh 2t = 0$.

That is, $C_1 e^{2t} + C_2 e^{-2t} + C_3 \cosh 2t = 0 \not\Rightarrow C_1 = C_2 = C_3 = 0$.

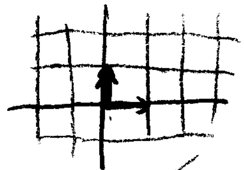
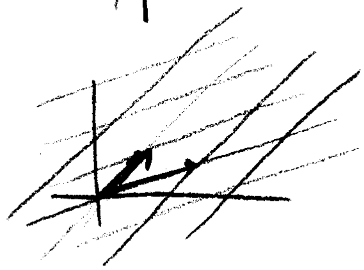
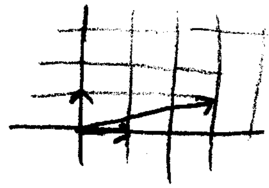
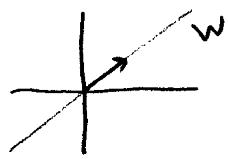
<u>Key concepts</u>: (spanning set vs. basis)

<u>Def</u>: A subset $S \subseteq V$ <u>spans</u> V if every $v \in V$ can be written as $v = a_1 v_1 + \cdots + a_n v_n$ where $v_i \in V$, $a_i \in F$.

Moreover, if S is a basis for V (i.e, if this is the <u>unique</u> way to write v, then S is a <u>basis</u> for V.

Roughly speaking:
- "S <u>spans</u> V" means "S generates all of V"
- "S is a basis for V" means, "S is a <u>minimal</u> set that generates V.

Examples: Let $V = \mathbb{R}^2$.

		Spans $\mathbb{R}^2$?	Basis for $\mathbb{R}^2$?
(i) $S = \{(1,0), (0,1)\}$		Y	Y
(ii) $S = \{(3,1), (1,1)\}$		Y	Y
(iii) $S = \{(1,0), (0,1), (3,1)\}$ unnecessary!		Y	N
(iv) $S = \{(1,1)\}$		N	N

However, S spans a 1-dimensional subspace (a line) of $\mathbb{R}^2$.

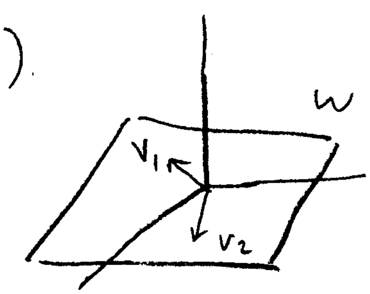
Theorem: Let $S \subseteq V$. The following are equivalent:

(i) S is a basis for V

(ii) S is a minimal spanning set of V

(iii) S is a maximal linearly independent set in V.

Example: Let $V = \mathbb{R}^3$, $W \subseteq V$ any plane (through $\vec{0}$)



Intuition: We need two vectors (not colinear) to generate.

In fact $S = \{v_1, v_2\}$ is a basis for W iff v_1 & v_2 are not colinear.

Go up a dimension (find a basis for V).

$S = \{v_1, v_2, v_3\}$ is a basis for U iff they are not coplanar.

It should be clear how this generalizes to higher dimensions.

Def: The dimension of a vector space is the number of vectors in any basis.

Example: $\dim(\mathbb{R}^n) = n$: Basis: $\{\hat{e}_1, \dots, \hat{e}_n\}$

$\dim(Poly_n) = n+1$: Basis: $\{1, x, \dots, x^n\}$

$\dim(\mathbb{R}[x]) = \infty$: Basis: $\{1, x, x^2, \dots\}$

$\dim(Per_{2\pi}) = \infty$: Basis: $\{1, \cos x, \cos 2x, \dots\} \cup \{\sin x, \sin 2x, \dots\}$

Remark: Any subset $S \subseteq V$ spans a subspace W of V. Denote this subspace by $\langle S \rangle$, or by $\mathrm{Span}(S)$.

We may ask: Is S a basis for W?
If not, then S is not a minimal spanning set, so we can remove some $v_1 \in S$ to get $S' := S \setminus \{v_1\}$, a smaller set that spans W.

We ask again: Is S' a basis for W?
If not, then we can remove some $v_2 \in S'$ to get $S'' := S' \setminus \{v_2\}$, a smaller set that spans W.

If $|S| < \infty$, then eventually this process will terminate, and we'll be left with $\mathcal{B} := S^{(k)}$, a basis for W.

Example: $S = \{(1,0,0), (0,1,0), (1,1,0), (3,1,0)\} \in \mathbb{R}^3$.

Span(S) is a plane W. We can remove $(1,1,0)$ and $(3,1,1)$ to get a basis $\mathcal{B} = \{(1,0,0), (0,1,0)\}$ of W.

This means that
$$W = \{ c_1(1,0,0) + c_2(0,1,0) \mid c_1, c_2 \in \mathbb{R}\}$$
$$= \{ (c_1, c_2, 0) \mid c_1, c_2 \in \mathbb{R} \}.$$

Note that $\{(1,0,0), (3,1,0)\}$ is also a basis for W.

This means that
$$W = \{ c_1(1,0,0) + c_2(3,1,0) \mid c_1, c_2 \in \mathbb{R}\}$$
$$= \{ (c_1 + 3c_2, c_2, 0) \mid c_1, c_2 \in \mathbb{R} \}.$$

Linear maps.

Def: A linear map (or linear operator) is a function $T: V \to W$ between vector spaces V, W satisfying
$$T(ax+by) = aT(x) + bT(y), \quad \text{for all } x, y \in V, \; a, b \in F.$$

Def: The kernel of a linear map T, denoted $\ker(T)$ (also called the nullspace) is the set of vectors such that $T(v) = 0$. That is, $\ker(T) = \{ v \in V \mid T(v) = 0 \}$.

Def: The *image* (or *range*) of T is the set $T(V)$,

i.e., $\mathrm{im}(T) := \{ w \in W \mid T(v) = w \text{ for some } v \in V \}$.

Examples:

(i) Let $V = W = \mathcal{C}^\infty$. $T = \frac{d}{dx}$ is a linear operator.

$T : f(x) \longmapsto f'(x)$. Check: $(af + bg)' = af' + bg'$

Note that $\ker(T) = \{\text{constant functions}\}$.

(ii) Let $V = W = \mathcal{C}^\infty$. $T = \int_0^1$ is a linear operator

$T : f(x) \longmapsto \int_0^1 f(x)$. Check: $\int_0^1 (af + bg) = a\int_0^1 f + b\int_0^1 g$.

(iii) The Laplace transform $\mathcal{L}$ is a linear operator.

$\mathcal{L} : f(t) \longmapsto \int_0^\infty f(t) e^{-st}\, dt$. Check: $\mathcal{L}(af + bg) = a\mathcal{L}(f) + b\mathcal{L}(g)$.

(iv) Any 2×2 matrix is a linear map $A : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$

$$\begin{array}{cc} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} & \\ \uparrow \quad \uparrow & \\ A\vec{e}_1 \quad A\vec{e}_2 & \end{array} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{bmatrix}$$

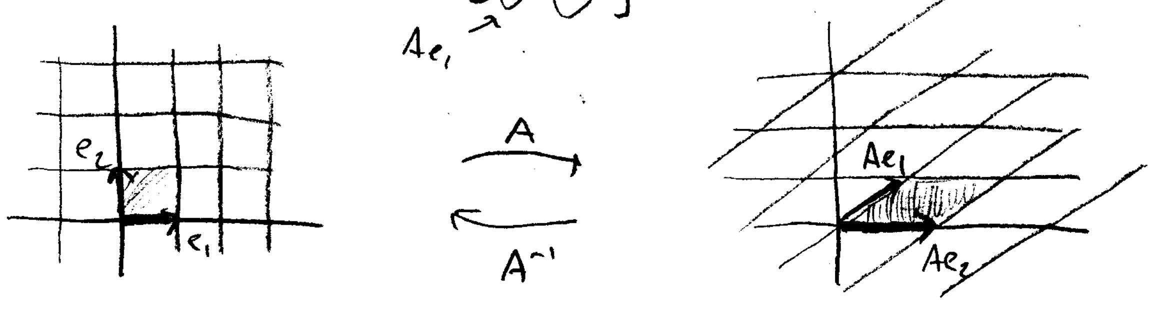
input in $\mathbb{R}^2$

output in $\mathbb{R}^2$

For example, let $A = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}$

Ae_2

Ae_1



Facts: • $|\det A|$ = scaling factor (= area of parallelogram)

negative denotes reflection

• A is <u>invertible</u> iff $\det A \neq 0$. (why?)

• In general, an $n \times m$ matrix is a linear map $A: \mathbb{R}^m \to \mathbb{R}^n$.

$$\begin{bmatrix} & A & \end{bmatrix}_{n\times m} \begin{bmatrix} \vec{x} \end{bmatrix}_{m\times 1} = \begin{bmatrix} \vec{y} \end{bmatrix}_{n\times 1}$$

• $\dim(\operatorname{im}(A)) + \dim(\ker A) = m$

i.e., every "dimension" either gets collapsed or persists.

<u>Connections between linear operators and ODE's</u> (preview):

<u>Examples</u>. Let $V = \mathcal{C}^\infty$

(i) $T = \frac{d^2}{dt^2}$ is a linear operator. $f \xmapsto{\frac{d^2}{dt^2}} f''$

$\ker(T) = \{ f(t) \mid f''(t) = 0 \} = \{ c_1 t + c_2 \mid c_1, c_2 \in \mathbb{R} \}$.

(ii) $T = \frac{d^2}{dt^2} + k^2$ is a linear operator: $y \longmapsto y'' + k^2 y$

$\ker(T) = \{ y(t) \mid y'' + k^2 y = 0 \} = \{ C_1 \cos kt + C_2 \sin kt \mid C_1, C_2 \in \mathbb{R} \}$.

(iii) $T = \frac{d}{dt} + t^2$ is a linear operator.

Check: $T(ay_1+by_2) = \left(\frac{d}{dt}+t^2\right)(ay_1+by_2) =$

$$= a\frac{dy_1}{dt} + at^2y_1 + b\frac{dy_2}{dt} + bt^2y_2$$

$$= a(y_1'+t^2y_1) + b(y_2'+t^2y_2) = aT(y_1) + bT(y_2).$$

$\ker T = \{ y(t) \mid y' + t^2 y = 0 \}$

Big idea: The kernel (or nullspace) of these linear differential operators are solutions to a linear, homogeneous, ODE.

Since $\ker(T)$ is a vector space, the set of solutions (i.e., the general solution) is a vector space as well.

We'll revisit this more later.

Inner products & orthogonality.

Def. Let V be an $\mathbb{R}$-vector space. A function $\langle -,- \rangle : V \times V \to \mathbb{R}$ is a (real) inner product if it satisfies (for all $u,v,w \in V$, $c \in \mathbb{R}$):

$$\left.\begin{array}{l} 1.\ \langle u+v, w\rangle = \langle u,v\rangle + \langle v,w\rangle \\ 2.\ \langle cv, w\rangle = c\langle v,w\rangle \end{array}\right\} \text{"bilinear"}$$

3. $\langle v,w\rangle = \langle w,v\rangle$
4. $\langle v,v\rangle \geq 0$ with equality iff $v=0$.

Remark: If V is a $\mathbb{C}$-vector space, then condition 3 is usually $\langle v,w\rangle = \overline{\langle w,v\rangle}$ (complex conjugation)

Think of an "inner product" as an abstract notion of "dot product."

Defining an inner product gives rise to a geometry, i.e., notions of length, angle, and projection.

- <u>length</u>: $\|v\| := \sqrt{\langle v, v\rangle}$
- <u>angle</u>: $\measuredangle(v, w) = \theta$, where $\cos\theta = \dfrac{\langle v, w\rangle}{\|v\| \cdot \|w\|}$
- <u>projection</u>: If $\|w\| = 1$, then we can project v onto w by defining $\text{proj}_w(v) = \langle v, w\rangle w$. This is the length, or magnitude, of v in the w-direction.

<u>Def</u>: Two vectors $v, w \in V$ are <u>orthogonal</u> if $\langle v, w\rangle = 0$. A set $\{v_1, \ldots, v_n\} \subseteq V$ is <u>orthonormal</u> if $\langle v_i, v_j\rangle = 0$ for all $i \neq j$, and $\|v_i\| = 1$ for all i.

<u>Remarks</u>:

- "orthogonal" is the abstract notion of "perpendicular."
- "orthonormal" means "perpendicular, <u>and</u> unit length"

An equivalent definition is:

* The set $\{v_1, \ldots, v_n\}$ is <u>orthonormal</u> if $\langle v_i, v_j\rangle = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$

Given a vector space V with an inner product, it is usually desirable to have an <u>orthonormal basis</u> of V.

Examples:

(i) Let $V=\mathbb{R}^n$. $\langle v,w\rangle = v\cdot w = \sum_{i=1}^{n} v_i w_i$ is an inner product.

The set $\{e_1,\dots,e_n\}$ is an orthonormal basis of $\mathbb{R}^n$.

Thus, we can write each $v\in\mathbb{R}^n$ uniquely as

$$v = a_1e_1+\cdots+a_ne_n = (a_1,\dots,a_n), \quad \text{where } a_i = \text{proj}_{e_i}(v) = v\cdot e_i$$

(ii) Let $V=\text{Per}_{2\pi}$, $F=\mathbb{C}$, $\langle f,g\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\overline{g(x)}\,dx$ is an inner product.

The set $\{\dots, e^{-2ix}, e^{-ix}, 1, e^{ix}, e^{2ix},\dots\}$ is an orthonormal basis of $\text{Per}_{2\pi}$, w.r.t. this inner product

Thus, we can write each $f(x)\in\text{Per}_{2\pi}$ uniquely as

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{inx} = c_0 + \sum_{n=1}^{\infty} c_n e^{inx} + c_{-n}e^{-inx}, \quad \text{where}$$

$$c_n = \|\text{proj}_{e^{inx}}(f(x))\| = \langle f(x), e^{inx}\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}\,dx$$

(iii) Let $V=\text{Per}_{2\pi}$, but define $\langle f,g\rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)g(x)\,dx$.

The set $\{\frac{1}{\sqrt{2}}, \cos x, \cos 2x,\dots\}\cup\{\sin x, \sin 2x,\dots\}$ is an orthonormal basis of $\text{Per}_{2\pi}$, w.r.t. this inner product.

Thus, we can write each $f(x)\in\text{Per}_{2\pi}$ uniquely as

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n\cos nx + b_n\sin nx, \quad \text{where}$$

$$a_n = \|\text{proj}_{\cos nx}(f(x))\| = \langle f(x), \cos nx\rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos nx\,dx$$

$$b_n = \|\text{proj}_{\sin nx}(f(x))\| = \langle f(x), \sin nx\rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin nx\,dx.$$

<u>Important remark</u>: Sometimes we have an orthogonal (but <u>not</u> orthonormal) basis, $v_1, \ldots, v_n$. There is still a simple way to decompose a vector $v \in V$ into this basis.

In this case, $\left\{ \frac{v_1}{\|v_1\|}, \ldots, \frac{v_n}{\|v_n\|} \right\}$ is an orthonormal basis, so

$$v = a_1 \frac{v_1}{\|v_1\|} + \cdots + a_n \frac{v_n}{\|v_n\|}, \qquad a_i = \left\langle v, \frac{v_i}{\|v_i\|} \right\rangle = \frac{1}{\|v_i\|} \langle v, v_i \rangle = \frac{\langle v, v_i \rangle}{\langle v_i, v_i \rangle^{1/2}}$$

$$v = \frac{a_1}{\|v_1\|} v_1 + \cdots + \frac{a_n}{\|v_n\|} v_n$$

$$v = c_1 v_1 + \cdots + c_n v_n, \qquad c_i = \frac{a_i}{\|v_i\|} = \frac{\langle v, v_i \rangle}{\langle v_i, v_i \rangle} = \frac{\langle v, v_i \rangle}{\|v_i\|^2} \quad (*)$$