Lecture Notes from Math 201, Fall 2018

Kyle Ormsby

November 28, 2018

Contents

Linear algebra is pervasive in modern math/science/tech:

- quantum physics
- Google PageRank
- machine learning (PCA, etc.)
- Markov processes
- multivariable differentiation
- multidimensional volume

...

But its origins are elementary:

e.g. Find all (x,y) such that

$$\begin{aligned} 3x+2y &= 5 \\ 2x-y &= 1. \end{aligned}$$

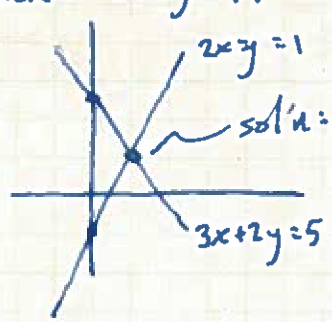
Eliminate variables:

$$\begin{array}{rcl} 3x+2y &=& 5 \\ 4x-2y &=& 2 \\ \hline 7x &=& 7 \end{array} \Rightarrow x=1 \text{ and } 3x+2y=5 \text{ is}$$

$$\begin{aligned} & 3+2y=5 \\ \Rightarrow\ & 2y=2 \\ \Rightarrow\ & y=1 \end{aligned}$$

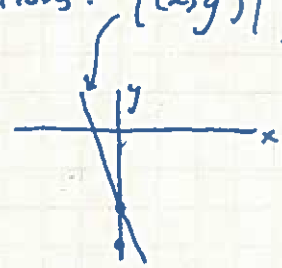
Unique solution: $x=y=1$.

Geometry:

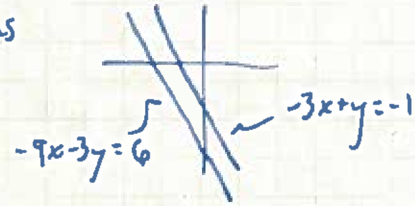


e.g. $-9x-3y=6$
$3x+y=-2$

Solutions: $\{(x,y) \mid y=-2-3x\}$



e.g. $-9x - 3y = 6$ No solutions
$3x + y = -1$



e.g.
$$\begin{aligned} x + 2y + z &= 0 \\ x \quad\;\; + z &= 4 \\ x + y + 2z &= 1 \end{aligned}$$

General idea: Replace a given set of equations with an equivalent set (having the same solution set) but from which solutions are evident.

The following operations do not change the solution set and are called row operations:

① Multiply an equation by a nonzero scalar
② Swap two equations
③ Add a multiple of one row to another

element of the "base field" F (maybe $\mathbb{R}$ or $\mathbb{C}$ or $\mathbb{Z}/5\mathbb{Z}$)

Think Pair Share Why are these operations invertible? Why does this imply solution sets are invariant under row operations?

We will see that row operations are sufficient to solve our problem.

$$\begin{aligned} x + 2y + z &= 0 \\ x \quad\;\; + z &= 4 \\ x + y + 2z &= 1 \end{aligned} \rightsquigarrow \left(\begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 1 & 0 & 1 & 4 \\ 1 & 1 & 2 & 1 \end{array}\right) \xrightarrow[\substack{\text{(eliminate} \\ x \text{ from last two} \\ \text{eq'ns)}}]{\substack{r_2 \to r_2 - r_1 \\ r_3 \to r_3 - r_1}} \left(\begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 0 & -2 & 0 & 4 \\ 0 & -1 & 1 & 1 \end{array}\right)$$

- augmented matrix
- columns correspond to coefficient of x, y, z, and constant value

$$\xrightarrow[\substack{\text{(set coeff of} \\ y \text{ in 2nd eqn} \\ \text{to 1)}}]{r_2 \to -\frac{1}{2} r_2} \left(\begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 0 & 1 & 0 & -2 \\ 0 & -1 & 1 & 1 \end{array}\right) \xrightarrow{r_3 \to r_3 + r_2} \underbrace{\left(\begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & -1 \end{array}\right)}_{\text{echelon form}} \xrightarrow{r_1 \to -2r_2 + r_1} \left(\begin{array}{ccc|c} 1 & 0 & 1 & 4 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & -1 \end{array}\right),$$

$$\xrightarrow{r_1 \to r_1 - r_3} \underbrace{\left(\begin{array}{ccc|c} 1 & 0 & 0 & 5 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & -1 \end{array}\right)}_{\text{reduced echelon form}} \rightsquigarrow \left.\begin{aligned} x &= 5 \\ y &= -2 \\ z &= -1 \end{aligned}\right\}$$

Unique solution! Check that it works.

e.g. $\begin{aligned} x + 2y + z &= 0 \\ x + z &= 4 \\ x + y + z &= 1 \end{aligned}$ i.e. $\left(\begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 1 & 0 & 1 & 4 \\ 1 & 1 & 1 & 1 \end{array}\right) \rightsquigarrow \left(\begin{array}{ccc|c} 1 & 0 & 1 & 4 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 0 & -1 \end{array}\right)$

No solutions as the final row says that $0 = -1$!

e.g. $\left(\begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 1 & 0 & 1 & 4 \\ 1 & 1 & 1 & 2 \end{array}\right) \rightsquigarrow \left(\begin{array}{ccc|c} 1 & 0 & 1 & 4 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 0 & 0 \end{array}\right)$ $\begin{aligned} x + z &= 4 \\ y &= -2 \\ (0 &= 0) \end{aligned}$

Solution set: $\{(x, -2, 4-x) \mid x \in \mathbb{R}\}$, a line in $\mathbb{R}^3$

Today: · Compute reduced echelon form of an augmented matrix.
· Learn how to express an infinite number of solutions in parametric and vector forms.

In a matrix, the <u>leading term</u> of a row is its first nonzero entry. A matrix is in <u>echelon form</u> if each leading term is to the right of the leading term in the row above it (except for the leading term in the first row) and any all 0 rows are at the bottom:

A matrix is in <u>reduced echelon form</u> if it is in echelon form and each leading term is a 1 and is the only nonzero entry in its column.

<u>e.g.</u> $\left(\begin{array}{cccc|c} 1 & & & & a \\ & 1 & & & b \\ & & 1 & & c \\ & & & 1 & d \end{array}\right) \Rightarrow \begin{array}{l} x_1 = a \\ x_2 = b \\ x_3 = c \\ x_4 = d \end{array}$ $\quad \left(\begin{array}{ccccc|c} 1 & 2 & & 3 & & a \\ & & 1 & 4 & & b \\ & & & & 1 & c \\ & & & & & d \end{array}\right) \Rightarrow ?$ (TPS)

<u>TPS</u> · When are there no solutions? — contradictory eq'n
· When is there a unique solution? — no contradiction, every column has a leading term

In reduced echelon form:

<u>e.g.</u>
$$\begin{aligned} 2x_3 + 6x_4 &= 0 \\ x_1 + 2x_2 + x_3 + 3x_4 &= 1 \\ 2x_1 + 4x_2 + 3x_3 + 9x_4 + x_5 &= 5 \end{aligned}$$
has augmented matrix

$$\left(\begin{array}{ccccc|c} 0 & 0 & 2 & 6 & 0 & 0 \\ 1 & 2 & 1 & 3 & 0 & 1 \\ 2 & 4 & 3 & 9 & 1 & 5 \end{array}\right) \xrightarrow{r_1 \leftrightarrow r_2} \left(\begin{array}{ccccc|c} 1 & 2 & 1 & 3 & 0 & 1 \\ 0 & 0 & 2 & 6 & 0 & 0 \\ 2 & 4 & 3 & 9 & 1 & 5 \end{array}\right) \xrightarrow{r_3 \to r_3 - 2r_1} \left(\begin{array}{ccccc|c} 1 & 2 & 1 & 3 & 0 & 1 \\ 0 & 0 & 2 & 6 & 0 & 0 \\ 0 & 0 & 1 & 3 & 1 & 3 \end{array}\right)$$

$$\xrightarrow[r_2 \to \frac{1}{2} r_2]{} \left(\begin{array}{ccccc|c} 1 & 2 & 1 & 3 & 0 & 1 \\ 0 & 0 & 1 & 3 & 0 & 0 \\ 0 & 0 & 1 & 3 & 1 & 3 \end{array}\right) \xrightarrow[r_3 \to r_3 - r_2]{r_1 \to r_1 - r_2} \left(\begin{array}{ccccc|c} 1 & 2 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 3 \end{array}\right)$$

x_2, x_4 are the <u>free variables</u>.

So the original system is equivalent to

$$\begin{aligned} x_1 + 2x_2 \qquad\qquad &= 1 \implies x_1 = 1 - 2x_2 \\ x_3 + 3x_4 &= 0 \implies x_3 = -3x_4 \\ x_5 &= 3 \implies x_5 = 3 \end{aligned}$$

Solution set: $\{(1-2x_2,\ x_2,\ -3x_4,\ x_4,\ 3) \mid x_2, x_4 \in \mathbb{R}\}$

a plane of solutions. This is a <u>parametric</u> description of the sol'ns.

Vector form: $$\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 3 \end{pmatrix} + x_2 \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} 0 \\ 0 \\ -3 \\ 1 \\ 0 \end{pmatrix} \;\middle|\; x_2, x_4 \in \mathbb{R} \right\}$$

<u>e.g.</u> If a system has reduced echelon form

$$\left(\begin{array}{ccccccc|c} 1 & 3 & 0 & 1 & 2 & 0 & 1 & 7 \\ 0 & 0 & 1 & 4 & -1 & 0 & 3 & -2 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

$x_2 \quad x_4 \quad x_5 \quad x_7$ free

then its solution set is

$$\{(7-3x_2-x_4-2x_5-x_7,\ x_2,\ -2-4x_4+x_5-3x_7,\ x_4,\ x_5,\ 3-x_7,\ x_7) \mid x_2, x_4, x_5, x_7 \in \mathbb{R}\}$$

(parametric form)

or $$\left\{ \begin{pmatrix} 7 \\ 0 \\ -2 \\ 0 \\ 0 \\ 3 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} -3 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} -1 \\ 0 \\ -4 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_5 \begin{pmatrix} -2 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_7 \begin{pmatrix} -1 \\ 0 \\ -3 \\ 0 \\ 0 \\ -1 \\ 1 \end{pmatrix} \;\middle|\; x_2, x_4, x_5, x_7 \in \mathbb{R} \right\}$$

(vector form).

Problem Find all parabolas $y = ax^2 + bx + c$ passing through $(1,4)$ & $(3,6)$.

Sol'n To pass through $(1,4)$ we need

$$4 = a + b + c.$$

For $(3,6)$,

$$6 = 9a + 3b + c.$$

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 4 \\ 9 & 3 & 1 & 6 \end{array}\right) \xrightarrow{r_2 \to r_2 - 9r_1} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 4 \\ 0 & -6 & -8 & -30 \end{array}\right) \xrightarrow{r_2 \to \frac{-1}{6} r_2} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 4 \\ 0 & 1 & 4/3 & 5 \end{array}\right)$$

$$\xrightarrow{r_1 \to r_1 - r_2} \left(\begin{array}{ccc|c} 1 & 0 & -1/3 & -1 \\ 0 & 1 & 4/3 & 5 \end{array}\right)$$ in reduced echelon form. Thus

the parabolas in question have

$$(a,b,c) \in \left\{ \left(-1 + \tfrac{1}{3}c,\ 5 - \tfrac{4}{3}c,\ c\right) \mid c \in \mathbb{R} \right\}$$

$$= \left\{ \begin{pmatrix} -1 \\ 5 \\ 0 \end{pmatrix} + c \begin{pmatrix} 1/3 \\ -4/3 \\ 1 \end{pmatrix} \mid c \in \mathbb{R} \right\}$$

TPS Check this.

Vector Spaces

Let F be a field, e.g. $\mathbb{R}, \mathbb{C}, \mathbb{Q}, \mathbb{Z}/2\mathbb{Z}$, etc. (not $\mathbb{Z}$ ...).

Defn A vector space over F (or F-vector space) is a set V together with operations $+ : V \times V \longrightarrow V$ (vector addition)
$\cdot : F \times V \longrightarrow V$ (scalar multiplication)

(Write $v+w$ for $+(v,w)$, λv for $\cdot(\lambda, v)$.) These operations have the following properties for all $x,y,z \in V$, $a,b \in F$:

1. $x+y = y+x$ (commutativity of $+$)
2. $(x+y)+z = x+(y+z)$ (associativity of $+$)
3. $\exists\, 0 \in V$ s.t. $\cancel{w}+0 = w \quad \forall w \in V$
4. $\exists\, -x \in V$ s.t. $x+(-x) = 0$
5. For $1 \in F$, $1 \cdot x = x$
6. $(ab)x = a(bx)$ (associativity of scalar mult)
7. $a(x+y) = ax + ay$
8. $(a+b)x = ax+bx$

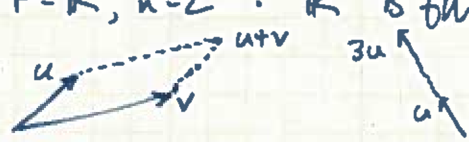
$\left.\right\}$ distributivity (7, 8)

Rmk 2–4 make V a group under $+$. 1 makes this group Abelian. 5–8 say that F acts on V in a manner compatible with $+$. All together, we get a linear structure on V.

e.g. $F^n = \underbrace{F \times \cdots \times F}_{n \text{ factors}} = \{(a_1, \dots, a_n) \mid a_i \in F \text{ for } i = 1, \dots, n\}$

$$(a_1, \dots, a_n) + (b_1, \dots, b_n) := (a_1+b_1, \dots, a_n+b_n)$$
$$c(a_1, \dots, a_n) := (ca_1, \dots, ca_n)$$

sub-eg. (a) $F = \mathbb{R}$, $n=2$: $\mathbb{R}^2$ is the Euclidean plane



(b) $F = \mathbb{Z}/2\mathbb{Z}$, $n=3$: vector space with 8 elts such as $(0,1,0), (0,1,1)$ with $(0,1,0)+(0,1,1) = (0,0,1)$.

(c) $n=1$: $F^1 = F$

(d) $n=0$: $F^0 = \{0\}$, the trivial vector space.

$$0+0=0$$
$$a0=0$$

e.g. $\mathbb{C}$ is an $\mathbb{R}$-vector space :

$$(a+bi)+(c+di) = (a+c)+(b+d)i$$
$$c(a+bi) = ca + cbi$$

e.g. $\mathbb{R}$ is a $\mathbb{Q}$-vector space.

e.g. $M_{m\times n}(F)$ = $m\times n$ matrices with entries in F

$$= \left\{ \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \middle|\ a_{ij} \in F \ \forall i,j \right\}$$

$$(A+B)_{ij} = \underbrace{A_{ij}}_{ij \text{ entry of } A} + B_{ij}$$

$$(cA)_{ij} = c(A_{ij})$$

Q How similar is this to F^{mn}?

e.g. For S any set, let $F^S := \{\text{functions } f: S \to F\}$.

$$(f+g)(s) = f(s)+g(s)$$
$$(cf)(s) = c(f(s))$$

If $S=\{1,\dots,n\}$, F^S is essentially the same as F^n :

$$f: \{1,\dots,n\} \to F \iff (f(1),\dots,f(n)) \in F^n$$

If $S=\mathbb{N}$, get sequences in F.

TPS Put a linear structure on the set of polynomials in a single variable.

Subspaces & Spanning Sets

Defn Let V be an F-vector space, and let $\emptyset \neq S \subseteq V$. A <u>linear combination</u> of vectors in S is a vector

$$v = a_1 u_1 + a_2 u_2 + \cdots + a_n u_n$$

for some $a_1, \ldots, a_n \in F$, $u_1, \ldots, u_n \in S$.

<u>e.g.</u> $S = \{(3,2), (2,-1)\} \subseteq \mathbb{Q}^2$. Is $(-1,4)$ a linear combo of vectors in S? Only if $\exists a, b \in F$ s.t.

$$a(3,2) + b(2,-1) = (-1,4)$$

i.e. $(3a, 2a) + (2b, -b) = (3a+2b,\ 2a-b) = (-1,4)$

i.e. $\left.\begin{aligned} 3a + 2b &= -1 \\ 2a - b &= 4 \end{aligned}\right\}$ system of linear equations!

Performing row ops,

$$\left(\begin{array}{cc|c} 3 & 2 & -1 \\ 2 & -1 & 4 \end{array}\right) \xrightarrow{r_1 \to r_1 - r_2} \left(\begin{array}{cc|c} 1 & 3 & -5 \\ 2 & -1 & 4 \end{array}\right) \xrightarrow{r_2 \to r_2 - 2r_1} \left(\begin{array}{cc|c} 1 & 3 & -5 \\ 0 & -7 & 14 \end{array}\right)$$

$$\xrightarrow{r_2 \to \frac{-1}{7} r_2} \left(\begin{array}{cc|c} 1 & 3 & -5 \\ 0 & 1 & -2 \end{array}\right) \xrightarrow{r_1 \to r_1 - 3r_2} \left(\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & -2 \end{array}\right)$$

so $a = 1$, $b = -2$. Indeed, $1 \cdot (3,2) + (-2)(2,-1) = (-1,4)$ ✓.

Defn V an F-vector space, $\emptyset \neq S \subseteq V$. The <u>span</u> of S, denoted $\operatorname{span}(S)$, is the set of all linear combos of elts of S.

Convention: $\operatorname{span}(\emptyset) = \{0\}$.

<u>e.g.</u> In $\mathbb{R}^2$, $\operatorname{span}(\{(1,1)\}) = \{(a,a) \mid a \in \mathbb{R}\}$

In $\mathbb{R}^3$, $\operatorname{span}(\{(1,0,0), (0,1,0)\}) = \{a(1,0,0) + b(0,1,0) \mid a, b \in \mathbb{R}\}$
$= \{(a, b, 0) \mid a, b \in \mathbb{R}\}$.

Defn A subset $W \subseteq V$ is a (linear or vector) <u>subspace</u> if W is a vector space itself with operations inherited from V.

Prop $W \subseteq V$ is a subspace iff

1. $0 \in W$
2. W is closed under addition ($u, v \in W \Rightarrow u+v \in W$)

e.g. $W = \{(a,0) \mid a \in \mathbb{R}\} \subseteq \mathbb{R}^2$ is a subspace

Pf Letting $a=0$, we see $(0,0) = 0 \in W$. If $(a,0), (b,0) \in W$, then $(a,0)+(b,0) = (a+b, 0) \in W$. If $c \in \mathbb{R}$, $(a,0) \in W$, then $c(a,0) = (ca, 0) \in W$. $\square$

e.g. $\mathbb{R}^{\mathbb{R}}$ = $\mathbb{R}$-vs. of fns $\mathbb{R} \to \mathbb{R}$

$$\begin{aligned} &\mathbb{R}^{\mathbb{R}} \\ &\ \cup\!\shortmid \\ &V = C(\mathbb{R},\mathbb{R}) = \{f:\mathbb{R}\to\mathbb{R} \mid f \text{ is continuous}\} \\ &\ \cup \\ &W = C'(\mathbb{R},\mathbb{R}) = \{f:\mathbb{R}\to\mathbb{R} \mid f \text{ is differentiable}\} \end{aligned}$$

V, W are subspaces of $\mathbb{R}^{\mathbb{R}}$, W is a subspace of V.

e.g. $W = \{(a,b) \mid a,b \in \mathbb{R},\ (a=0 \text{ or } b=0)\}$

= union of two axes

W is not a subspace of $\mathbb{R}^2$: $(0,1)+(1,0) = (1,1) \notin W$.

e.g. $\{0\}$, V are subspaces of V.

Prop If $W_1, W_2 \subseteq V$ are subspaces, then so is $W_1 \cap W_2$.

Pf Have $0 \in W_1$ and $0 \in W_2$, so $0 \in W_1 \cap W_2$. If $u,v \in W_1 \cap W_2$ then $u,v \in W_i$ for $i=1,2$. Hence $u+v \in W_i$ for $i=1,2$. Hence $u+v \in W_1 \cap W_2$. Similarly, for each $\lambda \in F$,

$$\begin{aligned} u \in W_1 \cap W_2 &\Rightarrow u \in W_1 \text{ and } u \in W_2 \\ &\Rightarrow \lambda u \in W_1 \text{ and } \lambda u \in W_2 \\ &\Rightarrow \lambda u \in W_1 \cap W_2. \qquad \square \end{aligned}$$

Prop If S is a subset of a vector space V, then

ⓐ span(S) is a subspace of V,

ⓑ if $W \subseteq V$ is a subspace and $S \subseteq W$, then span$(S) \subseteq W$,

ⓒ every subspace of V is the span of some subset of V.

Pf ⓐ If $S = \emptyset$, span$(S) = \{0\}$ is a subspace. Now suppose $S \neq \emptyset$. For $u \in S$, $0u = 0 \in$ span(S). Now take $x, y \in$ span(S).

Then $x = a_1u_1 + \cdots + a_nu_n$ for some $a_i, b_i \in F$, $u_i, v_i \in S$.
$y = b_1v_1 + \cdots + b_mv_m$

Thus $x+y = a_1u_1 + \cdots + a_nu_n + b_1v_1 + \cdots + b_mv_m$ is also a linear combo of elts of S, so $x+y \in \mathrm{span}(S)$. Finally,

$$cx = (ca_1)u_1 + \cdots + (ca_n)u_n \in \mathrm{span}(S).$$

(b) If $x \in \mathrm{span}(S)$ then $x = a_1u_1 + \cdots + a_nu_n$ for some $a_i \in F$, $u_i \in S$. Since $S \subseteq W$, $u_i \in W$ too. Since W is a subspace, it's closed under add'n & scalar mult, so $x \in W$. Hence $\mathrm{span}(S) \subseteq W$.

(c) $\mathrm{span}(W) = W$. $\square$

<u>Defn</u> We say $S \subseteq V$ <u>generates</u> a subspace W if $\mathrm{span}(S) = W$.

<u>e.g.</u>
- $\{(1,0), (0,1)\}$ generates $\mathbb{R}^2$
- $\{(1,0), (0,1), (3,2)\}$ generates $\mathbb{R}^2$
- $\{1, x, x^2, x^3, \ldots\}$ generates $F[x]$.
- Let $e_1 = (1,0,\ldots,0) \in F^n$

$e_2 = (0,1,0,\ldots,0) \in F^n$

$\vdots$

$e_i = (0,\ldots,0,1,0,\ldots,0) \in F^n$

$\uparrow$ i-th position

$\vdots$

$e_n = (0,\ldots,0,1) \in F^n$.

Then $\{e_1, \ldots, e_n\}$ generates F^n.

<u>e.g.</u> If T is a <u>finite</u> set, let $\chi_t : T \longrightarrow F$ for $t \in T$.
$$s \longmapsto \begin{cases} 1 & s = t \\ 0 & s \neq t \end{cases}$$

Then $\{\chi_t \mid t \in T\}$ generates F^T. (Check!)

<u>TPS</u> What goes wrong if T is infinite?

Linear Independence

Defn A set $S \subseteq V$ is *linearly dependent* if $\exists$ distinct $u_1, \dots, u_n \in S$ and scalars $a_1, \dots, a_n$ not all 0 s.t. $a_1 u_1 + \cdots + a_n u_n = 0$.

e.g. If $0 \in S$, then S is linearly dependent: $\underbrace{1}_{\text{scalar}} \cdot 0 = \underbrace{0}_{\text{vector}}$

e.g. $S = \{(1,-1,0), (-1,0,2), (-5,3,4)\}$.

S is linearly dep iff $\exists a_1, a_2, a_3$, not all 0 s.t.

$$a_1(1,-1,0) + a_2(-1,0,2) + a_3(-5,3,4) = 0 = (0,0,0)$$

$$\left(\begin{array}{ccc|c} 1 & -1 & -5 & 0 \\ -1 & 0 & 3 & 0 \\ 0 & 2 & 4 & 0 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|c} 1 & 0 & -3 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right) \implies \begin{array}{l} a_1 = 3a_3 \\ a_2 = -2a_3 \\ a_3 \text{ arbitrary} \end{array}$$

Taking $a_3 = 1$, we get a nontrivial sol'n w/ $a_1 = 3$, $a_2 = -2$, $a_3 = 1$.

Prop S is lin dep iff $\exists v \in S$ s.t. v is a linear combo of vectors in $S \setminus \{v\}$.

Pf ($\Rightarrow$) Suppose $a_1 u_1 + \cdots + a_n u_n = 0$ w/ $a_i \in F$, $u_i \in S$. WLOG, assume $a_1 \neq 0$. Then $u_1 = \frac{-a_2}{a_1} u_2 - \frac{a_3}{a_1} u_3 - \cdots - \frac{a_n}{a_1} u_n$.

($\Leftarrow$) Say $v = a_1 u_1 + \cdots + a_n u_n$ with $u_i \in S \setminus \{v\}$ and $v \in S$.
Then $a_1 u_1 + \cdots + a_n u_n - v = 0$ so S is lin dep. $\square$

Defn $S \subseteq V$ is *linearly independent* if it is not linearly dependent, i.e. if $a_1 u_1 + \cdots + a_n u_n = 0$ for distinct $u_i \in S$, then $a_1 = \cdots = a_n = 0$.

e.g.
- $\emptyset$ is lin ind.
- $\{u\}$ is lin ind $\forall\, 0 \neq u \in V$.
- $S = \{(1,-1,0), (-1,0,2), (0,1,1)\}$ is lin ind:

$$\left(\begin{array}{ccc|c} 1 & -1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 2 & 1 & 0 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right).$$

Thm $S_1 \subseteq S_2 \subseteq V$. If S_1 is lin ~~ind~~ dep, then S_2 is lin dep. If S_2 is lin ind, then S_1 is lin ind.

Pf Moral exercise. $\square$

Thm If $S \subseteq V$ is lin ind and $v \in V \setminus S$, then $S \cup \{v\}$ is lin dep iff $v \in \text{span}(S)$.

Pf ($\Rightarrow$) If $S \cup \{v\}$ is lin dep then $\exists\, a, a_i \in F$ not all 0 and distinct $u_i \in S$, distinct from v, s.t.

$$av + a_1u_1 + \cdots + a_nu_n = 0.$$

If $a = 0$, we would have $a_1u_1 + \cdots + a_nu_n = 0$, contradicting lin ind of S. Thus $a \neq 0$. Then

$$v = \frac{-a_1}{a}u_1 - \frac{a_2}{a}u_2 - \cdots - \frac{a_n}{a}u_n \in \text{span}(S).$$

($\Leftarrow$) If $v \in \text{span}(S)$, then v is a linear combo of vectors in S. Thus $S \cup \{v\}$ is lin dep. $\square$

e.g. In $F[x]$, $\{1, x, x^2, \ldots, x^n\}$ is lin ind.

★ **Thm** Suppose S is lin ind. Then for $v \in \text{span}(S)$, v can be expressed as a linear combo of vectors in S in a <u>unique</u> way.

Pf Say $v = a_1u_1 + \cdots + a_nu_n = b_1u_1 + \cdots + b_nu_n$ with $a_i, b_i \in F$, $u_i \in S$. (By letting some $a_i, b_i = 0$, we may assume we have the same u_i on both sides.) Then

$$0 = v - v = \sum (a_i - b_i)u_i.$$

By lin ind, $a_i - b_i = 0$ $\forall i$. $\square$

Note This result does not hold if S is lin dep. For example, let $S = \{(1,1), (2,2)\} \subseteq \mathbb{R}^2$. Then

$$\begin{aligned}(3,3) &= (1,1) + (2,2)\\ &= 2(1,1) + \tfrac{1}{2}(2,2)\\ &= 3(1,1) + 0(2,2) \quad \text{etc.}\end{aligned}$$

V an F-v.s. Recall $u_1, \dots, u_n \in V$ linearly independent when
$$a_1u_1 + \cdots + a_nu_n = 0,\ a_i \in F \implies a_1, \dots, a_n = 0.$$
A subset $S \subseteq V$ is lin ind if each of its finite subsets is lin ind.

Thm $S \subseteq V$ lin ind, $v \in \operatorname{span}(S)$. Then v can be expressed uniquely as a linear combo of elts of S.

Prop If $S_1 \subseteq S_2 \subseteq V$ and S_2 is lin ind, then S_1 is lin ind.

Pf Suppose $\sum_{i=1}^{n} a_iu_i = 0$ for some $u_i \in S_1$, $a_i \in F$. Since $S_1 \subseteq S_2$, $u_i \in S_2$ as well. Lin ind of $S_2 \implies a_i = 0\ \forall i$. $\square$

e.g. $V = (\mathbb{Z}/3\mathbb{Z})^3$, a $\mathbb{Z}/3\mathbb{Z}$-vector space.

Note that $|V| = 3^3 = 27$.

Check that $W = \{(x_1, x_2, x_3) \in V \mid x_1 + x_2 + x_3 = 0\} \subseteq V$ is a subspace.

Then $W = \{(-x_2 - x_3, x_2, x_3) \mid x_2, x_3 \in \mathbb{Z}/3\mathbb{Z}\}$ has 9 elements.

Find a lin ind generating set:

Take $v_1 = (2,1,0) \in W$ with $\operatorname{span}\{v_1\} = \{(0,0,0), (2,1,0), (1,2,0)\}$.

Then $v_2 = (1,1,1) \in W \setminus \operatorname{span}\{v_1\}$ so $\{v_1, v_2\}$ is lin ind.

Every element of $\operatorname{span}\{v_1, v_2\}$ has a unique expression of the form $a_1v_1 + a_2v_2$ with $a_1, a_2 \in \mathbb{Z}/3\mathbb{Z}$. Thus $|\operatorname{span}\{v_1, v_2\}| = 9$.

Also $\operatorname{span}\{v_1, v_2\} \subseteq W$, and cardinalities match, so they are equal.

Bases

Defn A subset $B \subseteq V$ is a <u>basis</u> if it is lin ind and spans V.

An <u>ordered basis</u> is a basis whose elements have been listed as a sequence, $B = \{b_1, b_2, \dots\}$.

⚠ The book does not distinguish b/w unordered and ordered bases (its bases are always ordered) but we will!

Prop If B is a basis of V, then every element of V can be expressed uniquely as a linear combo of elements of B.

Pf We have already seen that for B lin ind, every elt of span B has a unique such expression. Since span $B = V$, we are done. $\square$

Defn Let $B = \{v_1, \dots, v_n\}$ be an ordered basis of V. Given $v \in V$, there are unique $a_1, \dots, a_n \in F$ s.t. $v = a_1 v_1 + \cdots + a_n v_n$.
The coordinates of v with respect to B are the components of the vector $(a_1, \dots, a_n) \in F^n$.

e.g. ① $B = \{e_1, e_2, e_3\} \subseteq F^3$, where $e_1 = (1,0,0)$, $e_2 = (0,1,0)$, $e_3 = (0,0,1)$.
The coordinates of (x,y,z) wrt B are x, y, z.

② Let $e_1' = e_3$, $e_2' = e_2$, $e_3' = e_1$ and $B' = \{e_1', e_2', e_3'\}$. Then the coords of (x,y,z) wrt B' are z, y, x.

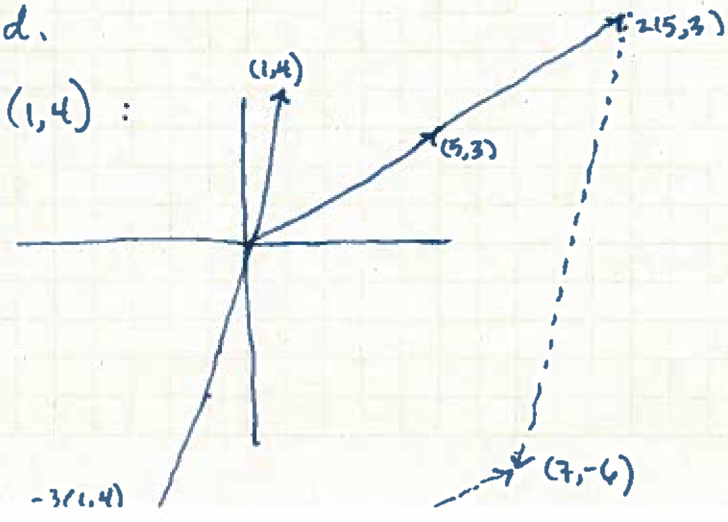
③ $B'' = \{(1,0,0), (1,1,0), (1,1,1)\}$ is an ordered basis of F^3.
Since $(x,y,z) = (x-y)(1,0,0) + (y-z)(1,1,0) + z(1,1,1)$, (x,y,z) has coords $x-y, y-z, z$ wrt B''.

④ $V = M_{2\times 2}(F)$. Then $B = \{M_1, M_2, M_3, M_4\}$ with

$$M_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad M_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad M_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad M_4 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

is an ordered basis of V wrt which the coords of $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ are a, b, c, d.

⑤ $(7,-6) = 2\cdot(5,3) - 3(1,4)$:



Dimension

Defn V is *finite dimensional* if it has a basis with a finite number of elements.

e.g. $\cdot F^n$, $M_{m\times n}(F)$ are fin dim'l

$\cdot F[x]$, $\mathbb{R}^{\mathbb{R}}$ are infinite dim'l.

Thm If V is finite dimensional, then every basis of V contains the same number of elts.

We'll get to the proof....

Defn If V is fin dim'l, the *dimension* of V, denoted $\dim V$ or $\dim_F V$, is the number of elements in any of its bases.

Exchange Lemma Suppose $B=\{v_1,\dots,v_n\}$ is a basis for V, and suppose $w=a_1v_1+\cdots+a_nv_n\in V$ with $a_i\in F$, $a_\ell\neq 0$. Let $B'=(B\setminus\{v_\ell\})\cup\{w\}$. Then B' is also a basis of V.

Pf First show B' is lin ind. WLOG, $\ell=1$. Suppose $bw+b_2v_2+\cdots+b_nv_n=0$. Substituting for w,

$$0=b(a_1v_1+\cdots+a_nv_n)+b_2v_2+\cdots+b_nv_n$$
$$=ba_1v_1+(ba_2+b_2)v_2+\cdots+(ba_n+b_n)v_n.$$

Since the v_i are lin ind, $ba_1=ba_2+b_2=\cdots=ba_n+b_n=0$.

Since $a_1\neq 0$, get $b=0$, so $b_2=\cdots=b_n=0$ as well. Thus B' lin ind.

Now show B' spans V. First note $v_1=\frac{1}{a_1}w-\frac{a_2}{a_1}v_2-\cdots-\frac{a_n}{a_1}v_n$.

Take $v\in V$. Since B spans, $v=c_1v_1+\cdots+c_nv_n$

$$=c_1\left(\frac{1}{a_1}w-\frac{a_2}{a_1}v_2-\cdots-\frac{a_n}{a_1}v_n\right)+c_2v_2+\cdots+c_nv_n$$

$$=\frac{c_1}{a_1}w+\left(c_2-\frac{c_1a_2}{a_1}\right)v_2+\cdots+\left(c_n-\frac{c_1a_n}{a_1}\right)v_n$$

so B' spans V. $\square$

Thm In a finite-dimensional vector space, every basis has the same number of elements.

Pf Let V be a fin dim vector space. Among bases for V, let $B = \{u_1, \dots, u_n\}$ be one of minimal size. Let $C = \{w_1, w_2, \dots\}$ be any other basis. Know $|B| \le |C|$, and want to show $|B| = |C|$.

Let $B_0 = B$ and consider $w_1 \in C$. By the exchange lemma, get a new basis B_1 by swapping w_1 with some u_ℓ. Relabeling if necessary, may assume $\ell = 1$ so $B_1 = \{w_1, u_2, \dots, u_n\}$.

Now consider $w_2 \in C$. Have $w_2 = a_1 w_1 + a_2 u_2 + \cdots + a_n u_n$ since B_1 is a basis. Since w_1, w_2 are lin ind, at least one of $a_2, \dots, a_n$ is nonzero. (Make sure you understand this step!) WLOG, $a_2 \ne 0$, so by exchange lemma, $B_2 = \{w_1, w_2, u_3, \dots, u_n\}$ is a basis. Continuing in this way, eventually get $B_n = \{w_1, \dots, w_n\}$ basis, $\subseteq C$.

In fact, $B_n = C$: if $w_{n+1} \in C \setminus B_n$, then $w_{n+1} = \sum_{i=1}^n d_i w_i$, (b/c B_n basis) but that can't happen b/c C is a basis. Thus $C = B_n$ has n elements. $\square$

Cor Let V be a fin dim vs, $S \subseteq V$ lin ind. Then S can be completed to form a basis of V.

Pf If $V \ne \text{span}(S)$, then for any $v \in V \setminus \text{span}(S)$, $S \cup \{v\}$ is lin ind. Continue until the set spans V. This terminates since o/w we would get an infinite basis. $\square$

Cor V fin dim vs, $V = \text{span}(S)$. Then $\exists T \subseteq S$ which is a basis.

Pf Similar. $\square$

Cor A collection of n vectors in an n-dim'l vector space is lin ind $\iff$ it spans V.

Pf ($\Rightarrow$) Suppose $S \subseteq V$ lin ind, $|S| = n$. We can complete S to a basis B, but if that involves adding any vectors to it, then $|B| > n$ ☠.

($\Leftarrow$) If S spans V, $|S| = n$, then we can shrink S to a basis B, but if that involves removing any vectors, then $|B| < n$ ☠. $\square$

<u>Moral</u> Basis = min'l spanning set

= max'l lin ind set

e.g. (1) $\mathbb{R}^n$ has basis $\{e_1, \dots, e_n\}$

(2) $\{(1,0,0), (1,2,0), (1,2,3)\} \subseteq \mathbb{R}^3$ lin ind $\Rightarrow$ basis.

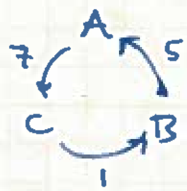
(3) $\mathbb{R}[x]_{\leq 2}$ = const, lin, or quad real polys. Basis $\{1, x, x^2\}$.

Sim, $\{1, 1+2x, 1+2x+3x^2\}$ is a basis

Condorcet's Paradox

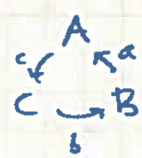
Candidates A, B, C; 29 voters

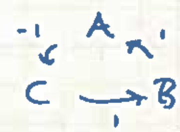
$$\begin{array}{l} A>B>C : 5 \\ A>C>B : 4 \\ B>A>C : 2 \\ B>C>A : 8 \\ C>A>B : 8 \\ C>B>A : 2 \end{array} \implies \begin{array}{ll} A>B : +5 & (17-12) \\ B>C : +1 & (15-14) \\ C>A : +7 & (18-11) \end{array}$$

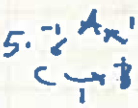
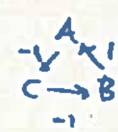
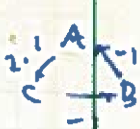
i.e.  a voting paradox or Condorcet cycle.

With head-to-head voting, any outcome can be achieved — the vote scheduler is a dictator!

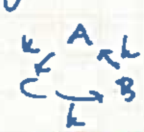
Goal: Use linear algebra to understand how/when such cycles arise.

$V = \Big\{$  $\Big| \; a,b,c \in \mathbb{R} \Big\}$ "=" $\mathbb{R}^3$

An $A>B>C$ voter corresponds to , etc.

The above example amounts to $5 \cdot$  $+ 4 \cdot$  $+ \cdots + 2 \cdot$ 

Call a vector in $\mathbb{R}^3$ purely cyclic if it is of the form (k,k,k), $k \in \mathbb{R}$

Let $C = \{(k,k,k) \mid k \in \mathbb{R}\} = \Big\{$  $\Big| \; k \in \mathbb{R} \Big\}$

Which vectors have no cyclic component? Those perpendicular to C: $(a,b,c) \perp (x,y,z) \iff ax+by+cz=0$.

So $C^{\perp} = \{(a,b,c) \in \mathbb{R}^3 \mid ak+bk+ck=0 \;\; \forall k \in \mathbb{R}\}$

$= \{(a,b,c) \in \mathbb{R}^3 \mid a+b+c=0\}$

$= \{ b(-1,1,0) + c(-1,0,1) \mid b,c \in \mathbb{R}\}$.

This led to the ordered basis $\mathcal{B} = \{(1,1,1), (-1,1,0), (-1,0,1)\}$ of $\mathbb{R}^3$.

~~Ex~~ First coord: cyclic component

$2^{nd}, 3^{rd}$ coords: non-cyclic components.

e.g. $(1,1,-1) = a(1,1,1) + b(-1,1,0) + c(-1,0,1)$

$$\Longleftrightarrow \quad \begin{aligned} a-b-c &= 1 \\ a+b &= 1 \\ a + c &= -1 \end{aligned} \quad \rightsquigarrow \left(\begin{array}{ccc|c} 1 & -1 & -1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & -1 \end{array}\right) \xrightarrow{\text{row ops}} \left(\begin{array}{ccc|c} 1 & 0 & 0 & 1/3 \\ 0 & 1 & 0 & 2/3 \\ 0 & 0 & 1 & -4/3 \end{array}\right)$$

so $(1,1,-1)$ has coords $(1/3, 2/3, -4/3)$ wrt $\mathcal{B}$.

In particular, this "rational preference" (i.e. ordered preference) has a cyclic component!

$C > B > A$ has coords $(-\frac{1}{3}, -\frac{2}{3}, \frac{4}{3})$ wrt $\mathcal{B}$

Call sign of first coord the spin of the rational preference.

pos spin

neg spin

① pos spin:

$$\begin{tikzcd}[column sep=small] & A \arrow[dl, "-1"'] & \\ C \arrow[rr, "1"'] & & B \arrow[ul, "1"'] \end{tikzcd} = \begin{tikzcd}[column sep=small] & A \arrow[dl, "1/3"'] & \\ C \arrow[rr, "1/3"'] & & B \arrow[ul, "1/3"'] \end{tikzcd} + \begin{tikzcd}[column sep=small] & A \arrow[dl, "-4/3"'] & \\ C \arrow[rr, "2/3"'] & & B \arrow[ul, "2/3"'] \end{tikzcd}$$

$A > B > C$

① neg spin:

$$\begin{tikzcd}[column sep=small] & A \arrow[dl, "1"'] & \\ C \arrow[rr, "-1"'] & & B \arrow[ul, "-1"'] \end{tikzcd} = \begin{tikzcd}[column sep=small] & A \arrow[dl, "-1/3"'] & \\ C \arrow[rr, "-1/3"'] & & B \arrow[ul, "-1/3"'] \end{tikzcd} + \begin{tikzcd}[column sep=small] & A \arrow[dl, "4/3"'] & \\ C \arrow[rr, "-2/3"'] & & B \arrow[ul, "-2/3"'] \end{tikzcd}$$

$C > B > A$

② pos spin:

$$\begin{tikzcd}[column sep=small] & A \arrow[dl, "1"'] & \\ C \arrow[rr, "1"'] & & B \arrow[ul, "-1"'] \end{tikzcd} = \begin{tikzcd}[column sep=small] & A \arrow[dl, "1/3"'] & \\ C \arrow[rr, "1/3"'] & & B \arrow[ul, "1/3"'] \end{tikzcd} + \begin{tikzcd}[column sep=small] & A \arrow[dl, "2/3"'] & \\ C \arrow[rr, "2/3"'] & & B \arrow[ul, "-4/3"'] \end{tikzcd}$$

$B > C > A$

② neg spin:

$$\begin{tikzcd}[column sep=small] & A \arrow[dl, "-1"'] & \\ C \arrow[rr, "-1"'] & & B \arrow[ul, "1"'] \end{tikzcd} = \begin{tikzcd}[column sep=small] & A \arrow[dl, "-1/3"'] & \\ C \arrow[rr, "-1/3"'] & & B \arrow[ul, "-1/3"'] \end{tikzcd} + \begin{tikzcd}[column sep=small] & A \arrow[dl, "-2/3"'] & \\ C \arrow[rr, "-2/3"'] & & B \arrow[ul, "4/3"'] \end{tikzcd}$$

$A > C > B$

③ pos spin:

$$\begin{tikzcd}[column sep=small] & A \arrow[dl, "1"'] & \\ C \arrow[rr, "-1"'] & & B \arrow[ul, "1"'] \end{tikzcd} = \begin{tikzcd}[column sep=small] & A \arrow[dl, "1/3"'] & \\ C \arrow[rr, "1/3"'] & & B \arrow[ul, "1/3"'] \end{tikzcd} + \begin{tikzcd}[column sep=small] & A \arrow[dl, "2/3"'] & \\ C \arrow[rr, "-4/3"'] & & B \arrow[ul, "2/3"'] \end{tikzcd}$$

$C > A > B$

③ neg spin:

$$\begin{tikzcd}[column sep=small] & A \arrow[dl, "-1"'] & \\ C \arrow[rr, "1"'] & & B \arrow[ul, "-1"'] \end{tikzcd} = \begin{tikzcd}[column sep=small] & A \arrow[dl, "-1/3"'] & \\ C \arrow[rr, "-1/3"'] & & B \arrow[ul, "-1/3"'] \end{tikzcd} + \begin{tikzcd}[column sep=small] & A \arrow[dl, "-2/3"'] & \\ C \arrow[rr, "4/3"'] & & B \arrow[ul, "-2/3"'] \end{tikzcd}$$

$B > A > C$

Summing row ① contributions from election, get $\begin{tikzcd}[column sep=small] & A \arrow[dl, "-a"'] & \\ C \arrow[rr, "a"'] & & B \arrow[ul, "a"'] \end{tikzcd}$
with $a > 0$ if more on left, $a < 0$ if more on right, $a = 0$ if same left/right.
From row ②, sim get $\begin{tikzcd}[column sep=small] & A \arrow[dl, "b"'] & \\ C \arrow[rr, "b"'] & & B \arrow[ul, "-b"'] \end{tikzcd}$, and from ③ $\begin{tikzcd}[column sep=small] & A \arrow[dl, "c"'] & \\ C \arrow[rr, "-c"'] & & B \arrow[ul, "c"'] \end{tikzcd}$.

The election is then determined by

$$\begin{tikzcd} & A \arrow[dl, "-a+b+c"'] & \\ C \arrow[rr, "a+b-c"'] & & B \arrow[ul, "a-b+c"'] \end{tikzcd}$$

Condorcet cycle when all 3 have same sign.

All positive: $\begin{aligned} -a+b+c &> 0 \\ a-b+c &> 0 \\ a+b-c &> 0 \end{aligned} \implies a,b,c > 0$

check

Thus we have proved the following:

Thm If there is a Condorcet cycle, then $a,b,c>0$ or $a,b,c<0$.

TPS Converse?

Q Given N voters, what fraction of voting profiles result in Condorcet cycles?

Rank of matrices

Defn Let A be an $m \times n$ matrix over F. The row space of A is the subspace of F^n spanned by the rows of A. The column space of A is the subspace of F^m spanned by the columns of A.
The row rank of A is the dimension of its row space. The column rank of A is the dimension of its column space.

Note Row operations = linear combos of rows. So if $A \to B$ via row ops, then $\text{Rowspace}(B) \subseteq \text{Rowspace}(A)$. But we can reverse row ops to get $B \to A$ so the opposite inclusion holds as well. Thus:

Lemma If A, B related by row ops, then they have the same row space. In particular, the reduced echelon form of A has the same row space as A. $\square$

TPS Why are the nonzero rows of a reduced echelon matrix lin ind?

Prop Let A be an $m \times n$ matrix which reduces to E in reduced echelon form. Then the nonzero rows of E for a basis of the row space of A. $\square$

e.g.
$$A = \begin{pmatrix} 1 & 2 & 0 & 4 \\ 3 & 3 & 1 & 0 \\ 7 & 8 & 2 & 4 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & 2/3 & -4 \\ 0 & 1 & -1/3 & 4 \\ 0 & 0 & 0 & 0 \end{pmatrix} \text{ in reduced echelon form}$$

so $\{(1,0,2/3,-4), (0,1,-1/3,4)\}$ is a basis of the row space of A.

Lemma Row ops don't change the column rank of A.

Pf Suppose A is an $m \times n$ matrix with relation
$$c_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + \cdots + c_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} = 0$$
among its columns. This rel'n is equiv to a solution $(c_1, \dots, c_n)$ to the linear system
$$\begin{aligned} c_1 a_{11} + \cdots + c_n a_{1n} &= 0 \\ &\vdots \\ c_1 a_{m1} + \cdots + c_n a_{mn} &= 0 \end{aligned}$$
Row ops don't change sol'ns, so don't change rel'ns among cols. $\square$

We see that rel'ns among columns correspond to rel'ns b/w cols of reduced echelon form of the matrix. The cols containing a pivot form a basis, so the corr cols in A form a basis of its column space!

⚠ Must take the corresponding cols in A, not the cols in E.

e.g. In the previous example, the first 2 cols were pivot cols of E, so $\begin{pmatrix}1\\3\\7\end{pmatrix}, \begin{pmatrix}2\\3\\8\end{pmatrix}$ form a basis of col space of A.

Thm The row rank of a matrix is equal to its column rank.

Pf Let E be the reduced echelon form of a matrix A.
The number of nonzero rows equals the number of pivot columns. □

Defn The rank of a matrix A, denoted $\mathrm{rank}(A)$, is the dimension of its row or column space.

Thm Suppose we have a homogeneous system of linear eq'ns

$$\begin{matrix} a_{11}x_1 + \cdots + a_{1n}x_n = 0 \\ \vdots \qquad\qquad \vdots \quad\ \ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = 0 \end{matrix}$$

Let A be the corresponding matrix. Then the vectors space of solutions has dimension $n - \mathrm{rank}(A)$. (So unique sol'n iff $\mathrm{rank}(A) = n$.).

Pf To solve, we compute REF of A. The number of free variables $=$ # non-pivot columns $= n - \mathrm{rank}(A)$. □

TPS What about a non-homogeneous system?

Linear Transformations

Q How are vector spaces related? **A** By linear transformations.

Defn V, W F-vector spaces. A **linear transformation** from V to W is a function $f: V \to W$ s.t. $\forall v, v' \in V, \lambda \in F$,

$$f(v+v') = f(v) + f(v') \quad \& \quad f(\lambda v) = \lambda f(v).$$

"f preserves addition" "f preserves scalar multn"

"f preserves linear structure"

Note $f: V \to W$ is a lin trans iff $f(v+\lambda v') = f(v) + \lambda f(v')$ $\forall v, v', \lambda$.

Synonyms linear map, homomorphism

e.g. $f: \mathbb{R}^3 \to \mathbb{R}^2$, $(x,y,z) \mapsto (2x+3y, x+y-3z)$ is linear:

$$\begin{aligned} f((x,y,z)+(x',y',z')) &= f(x+x', y+y', z+z') \\ &= (2(x+x')+3(y+y'), x+x'+y+y'-3(z+z')) \\ &= (2x+3y, x+y-3z) + (2x'+3y', x'+y'-3z') \\ &= f(x,y,z) + f(x',y',z'), \end{aligned}$$

$$f(\lambda(x,y,z)) = f(\lambda x, \lambda y, \lambda z) = \cdots = \lambda f(x,y,z).$$

TPS Is $f: \mathbb{R} \to \mathbb{R}$, $x \mapsto x^2$ linear?

Prop If $f: V \to W$ is linear, then $f(0) = 0$ (i.e. $f(0_V) = 0_W$).

Pf Since f is linear, $f(0) = f(0 \cdot 0_V) = 0 \cdot f(0_V) = 0_W$. $\square$

Prop Let V, W be F-vector spaces, $B \subseteq V$ a basis. For each $b \in B$, take $w_b \in W$. Then $\exists!$ linear trans $f: V \to W$ s.t. $f(b) = w_b$ $\forall b \in B$.

Slogan Linear transformations are determined by their action on bases.

Pf Given $v \in V$, have a unique expression $v = a_1 b_1 + \cdots + a_k b_k$ for some $a_i \in F^\times$, $b_i \in B$. Define $f(v) = a_1 w_{b_1} + \cdots + a_k w_{b_k}$. ~~Since B~~ ~~is a basis~~ Well-definition follows from uniqueness of the expression for v. Linearity forces this defn since $f(b) = w_b$. $\square$

Terminology Say f defined on B and extended linearly to V.

(For V, W F-vss, let $\mathcal{L}(V,W) = \mathrm{Hom}(V,W) = \mathrm{Hom}_F(V,W)$ be the set of linear transformations $V \to W$. This forms a vector space via the operations $(f+g)(v) = f(v) + g(v)$, $(\lambda f)(v) = \lambda(f(v))$.)

e.g. $h: \mathbb{R}^2 \to \mathbb{R}^2$ linear with $h(e_1) = (-1,1)$, $h(e_2) = (3,4)$.

Then $h(a,b) = h(ae_1 + be_2) = a h(e_1) + b h(e_2) = (-a, a) + (3b, 4b)$

$= (3b-a, 4b+a)$

e.g. $\pi_i : F^n \longrightarrow F$

$e_j \longmapsto \begin{cases} 1 & \text{if } j = i \\ 0 & \text{if } j \neq i \end{cases}$

TPS What is a formula for $\pi_i(a_1, \ldots, a_n)$?

TPS Is matrix transpose $M_{2\times2}(F) \longrightarrow M_{2\times2}(F)$ linear?

$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \longmapsto \begin{pmatrix} a & c \\ b & d \end{pmatrix}$

Kernel and Image

Lemma Let $f: V \to W$ be a linear trans. For any subspace $U \subseteq V$, $f(U) = \{f(u) \mid u \in U\}$ is a subspace of W.

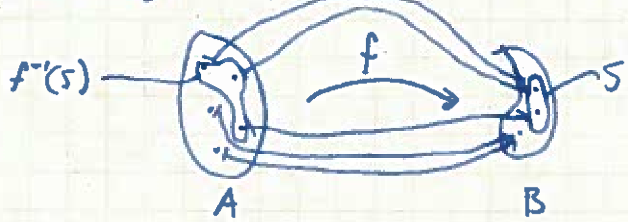
Pf Since $0 \in U$ and $f(0) = 0$, $0 \in f(U)$. Now for $f(u), f(u') \in f(U)$, $f(u) + \lambda f(u') = f(u + \lambda u') \in f(U)$ since U is closed under linear combos. $\square$

Defn The image (or range space) of $f: V \to W$ is
$$\operatorname{im}(f) = \mathcal{R}(f) = f(V) = \{f(v) \mid v \in V\}.$$
The dimn of $\operatorname{im}(f)$ is the rank of f.

e.g. $\frac{d}{dx}: F[x]_{\leq 3} \to F[x]$ has image $F[x]_{\leq 2}$
$$a_0 + a_1 x + a_2 x^2 + a_3 x^3 \longmapsto a_1 + 2a_2 x + 3a_3 x^2$$
and rank 2.

e.g. $h: M_{2\times 2}(F) \longrightarrow F[x]_{\leq 3}$
$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \longmapsto (a+b+2d) + cx^2 + cx^3$$
Any vector in $\operatorname{im}(h)$ has any const term, 0 linear coeff, and equal x^2, x^3 coeff. So $\operatorname{im}(h) = \{r + sx^2 + sx^3 \mid r, s \in F\}$ and $\operatorname{rank}(h) = 2$.

Recall If $f: A \to B$ is a function and $S \subseteq B$, then the preimage of S under f is $f^{-1}(S) := \{a \in A \mid f(a) \in S\}$.



Lemma For any $f: V \to W$ linear and $U \subseteq W$ subspace, $f^{-1}(U)$ is a subspace of V.

Pf $0 \in f^{-1}(U)$ b/c $0 \in U$ and $f(0)=0$. If $v, v' \in f^{-1}(U)$, then $f(v+v') = f(v) + f(v') \in U$ b/c $f(v), f(v') \in U$. Moreover, $f(\lambda v) = \lambda f(v) \in U$ b/c $f(v) \in U$. Thus $v+v', \lambda v \in f^{-1}(U)$ so $f^{-1}(U)$ is a subspace.

Defn The kernel (or null space) of a linear map $f: V \to W$ is the inverse image of $\{0\}$,

$$\ker(f) = \mathcal{N}(f) = f^{-1}(\{0\}) = \{v \in V \mid f(v) = 0\}.$$

Note $\{0\} \subseteq W$ is a subspace so $\ker(f)$ is a subspace of V. The dimension of $\ker(f)$ is the map's nullity.

e.g. $\ker\left(\frac{d}{dx} : F[x] \to F[x]\right) = \{\text{constant polynomials}\}$ so $\frac{d}{dx}$ has nullity 1.

e.g. For $h: M_{2\times 2}(F) \to F[x]_{\le 3}$ as before,

$$\ker(h) = \left\{ \begin{pmatrix} a & b \\ 0 & \frac{-(a+b)}{2} \end{pmatrix} \middle| a, b \in F \right\}$$ so h has nullity 2.

e.g. $\pi_i : \mathbb{R}^n \to \mathbb{R}$ projection onto ith coord has

$$\ker(\pi_i) = \{(a_1, \ldots, a_{i-1}, 0, a_{i+1}, \ldots, a_n) \mid a_j \in \mathbb{R}\}$$

has nullity $n-1$.

Thm If $f: V \to W$ linear, then $\operatorname{rank}(f) + \operatorname{nullity}(f) = \dim V$.

e.g. Check for previous examples.

Pf Let $\{v_1, \ldots, v_k\}$ be a basis for $\ker(f)$. Extend to a basis $\{v_1, \ldots, v_k, v_{k+1}, \ldots, v_n\}$ of V. We show $\{f(v_{k+1}), \ldots, f(v_n)\}$ is a basis of $\operatorname{im}(f)$, and the thm then follows.

Suppose $0 = c_{k+1} f(v_{k+1}) + \cdots + c_n f(v_n)$. Then $0 = f(c_{k+1} v_{k+1} + \cdots + c_n v_n)$ so $c_{k+1} v_{k+1} + \cdots + c_n v_n \in \ker(f)$. Since $\{v_1, \ldots, v_k\}$ is a basis

of $\ker(f)$ and $\{v_1, \dots, v_n\}$ is a basis of V, get $c_{k+1} = \cdots = c_n = 0$.
Thus B is lin ind.

Now suppose $f(v) \in \operatorname{im}(f)$. Have $v = c_1 v_1 + \cdots + c_n v_n$ and
$f(v) = c_1 f(v_1) + \cdots + c_n f(v_n) = c_{k+1} f(v_{k+1}) + \cdots + c_n f(v_n)$ since
$v_1, \dots, v_k \in \ker(f)$. Thus B spans $\operatorname{im}(f)$. $\square$

<u>Recall</u> $f: V \to W$ linear

$\ker(f) = \{v \in V \mid f(v) = 0\}$

$\text{im}(f) = \{f(v) \mid v \in V\}$

$\text{rank}(f) = \dim \text{im}(f)$

$\text{nullity}(f) = \dim \ker(f)$

$\text{rank}(f) + \text{nullity}(f) = \dim V$ (Rank-nullity thm)

<u>Prop</u> Linear $f: V \to W$ is injective iff $\ker(f) = \{0\}$.

<u>Pf</u> ~~If Ex~~ By linearity, $f(0) = 0$, so if f is injective, then $\ker(f) = \{0\}$. Now suppose $\ker(f) = \{0\}$ and that $f(u) = f(v)$. Then

$$0 = f(u) - f(v) = f(u-v) \Rightarrow u - v \in \ker(f) = \{0\}$$
$$\Rightarrow u - v = 0$$
$$\Rightarrow u = v$$

so f is inj. $\square$

<u>Prop</u> Let $S \subseteq V$, $f: V \to W$ linear.

① If S is lin dep, then $f(S) = \{f(s) \mid s \in S\} \subseteq W$ is lin dep.

② If f is injective and S is lin ind, then $f(S) \subseteq W$ is lin ind.

<u>Pf</u> Suppose $\sum a_i s_i = 0$ for some $a_i \in F$, $s_i \in S$. Since f is linear,

$$0 = f(0) = f\left(\sum a_i s_i\right) = \sum a_i f(s_i)$$ so f preserves dependences.

Now suppose f inj, S lin ind. If $0 = \sum a_i f(s_i)$ for some $a_i \in F$, $f(s_i) \in f(S)$, then, by linearity of f, $0 = f(\sum a_i s_i)$, i.e., $\sum a_i s_i \in \ker(f) = \{0\}$. Thus $\sum a_i s_i = 0$ and since S is lin ind, each $a_i = 0$. Thus $f(S)$ lin ind. $\square$

<u>Defn</u> A linear trans $f: V \to W$ is an <u>isomorphism</u> if $\exists$ lin trans $g: W \to V$ s.t. $g \circ f = id_V$ and $f \circ g = id_W$. (Call g the <u>inverse</u> of f, write $g = f^{-1}$.) Write $f: V \cong W$ or $V \cong W$.

Note When a function $f: A \to B$ has an inverse, we know it is bijective, so every linear isomorphism is a bijection. In fact, if $f: V \to W$ is a linear bijection, then its inverse function is also ~~bijective~~ linear (check!), and so f is an isomorphism.

Takeaway: Isomorphism = linear bij'n.

e.g. $M_{2\times 2}(F) \longrightarrow F^4$ is an isomorphism.
$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \longmapsto (a,b,c,d)$$

So is $M_{m\times n}(F) \longrightarrow F^{mn}$ discussed previously.

Prop Linear $f: V \to W$ is an iso iff $\ker(f) = \{0\}$ and $\operatorname{im}(f) = W$.

Pf $\ker(f) = \{0\}$ iff f is inj, $\operatorname{im}(f) = W$ iff f is surj, so both conditions (taken together) are equivalent to f bij, i.e. f an iso. $\square$

Thm Let $\dim V = n < \infty$. Then $V \cong F^n$.

Pf Choose a basis $b_1, \dots, b_n$ of V and let $e_1, \dots, e_n$ be the standard basis of F^n. Define $f: V \to F^n$ by $f(b_i) = e_i$, $1 \le i \le n$, and extending linearly. Then $f(v) = f(\sum a_i b_i) = \sum a_i e_i = (a_1, \dots, a_n) \in F^n$ — this is the map taking v to its coordinates wrt $b_1, \dots, b_n$ which is clearly a bij'n. $\square$

Cor Let V, W be finite dimensional vector spaces. Then V and W are isomorphic iff $\dim V = \dim W$.

Pf First suppose $f: V \cong W$ and let $b_1, \dots, b_n$ be a basis of V. Then $f(b_1), \dots, f(b_n)$ are lin ind (by Prop) and they span W b/c f is surj. Thus $f(b_1), \dots, f(b_n)$ is a basis of W $\Rightarrow \dim W = n = \dim V$.

Now suppose $\dim V = \dim W = n$. By the Thm, there are isos $V \xrightarrow[\cong]{f} F^n \xleftarrow[\cong]{g} W$. Then $g^{-1} \circ f: V \to W$ is an iso. $\square$

Note For $n = 0, 1, 2, \dots$ get only one isomorphism class of n-dim vector space. Choosing an iso $V \to F^n$ is equivalent to choosing a basis of V.

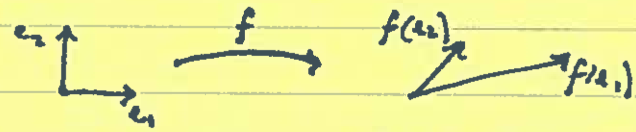
Geometry of Linear Transformations

Goal: Build visual intuition for linear trans

$f: \mathbb{R}^m \to \mathbb{R}^n$, focusing on $m=n=2$.

Recall that $f: V \to W$ linear is specified by its action on a basis of V: Suppose $b_1, \dots, b_n$ form a basis of V. For any $w_1, \dots, w_n \in W$, $\exists!$ lin trans $f: V \to W$ s.t. $f(b_i) = w_i$. (Then $f(\sum a_i b_i) = \sum a_i f(b_i) = \sum a_i w_i$.)

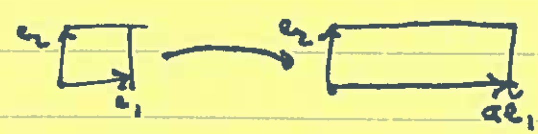
In particular a linear trans $f: \mathbb{R}^2 \to \mathbb{R}^2$ is specified by $f(e_1) = f(1,0)$ & $f(e_2) = f(0,1)$.



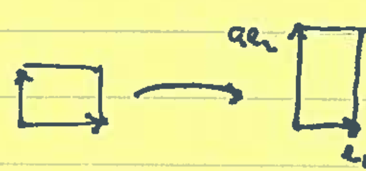
$(a,b) \longmapsto a f(e_1) + b f(e_2)$

Thus it is common to visualize linear trans $f: \mathbb{R}^2 \to \mathbb{R}^2$ by what they do to the unit square $[0,1] \times [0,1] \subseteq \mathbb{R}^2$. Here are some special cases along with their effects on $[0,1]^2$:

Scale
- $e_1 \mapsto a e_1$
  $e_2 \mapsto e_2$

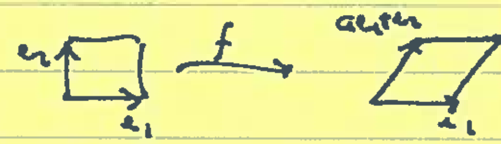


- $e_1 \mapsto e_1$
  $e_2 \mapsto a e_2$



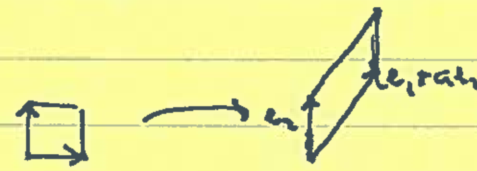
Shear

- $e_1 \mapsto e_1$
- $e_2 \mapsto ae_1 + e_2$



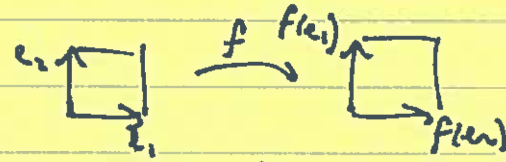
- $e_1 \mapsto e_1 + ae_2$

  $e_2 \mapsto e_2$



Reflect

- $e_1 \mapsto e_2$
- $e_2 \mapsto e_1$

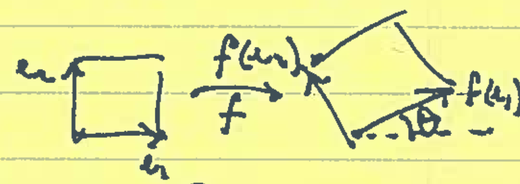


Reflects through the $y=x$ line.

Rotate

- $e_1 \mapsto (\cos\theta, \sin\theta)$

  $e_2 \mapsto (-\sin\theta, \cos\theta)$



Call this map R_θ

Squash

- $e_1 \mapsto e_1$

  $e_2 \mapsto 0$

- 0-map

Fact: Every linear trans $\mathbb{R}^m \to \mathbb{R}^n$ is the composition of a possible squash followed by shears, scales and reflections.

Note: We will prove this when studying matrix inversion.

TPS: Why is $f \circ g$ linear when f, g are linear?

Q: How can we represent R_θ as such a composition?

Special case: $\theta = \pi$. Then $e_1 \mapsto -e_1$, $e_2 \mapsto -e_2$ is clearly the composition of two scales.

Now suppose $\theta \notin k\pi$, $k \in \mathbb{Z}$.

<u>Claim</u> $R_\theta = X_\alpha \circ Y_\beta \circ X_\gamma$ for X_α an x-shear by α
($e_1 \mapsto e_1$, $e_2 \mapsto \alpha e_1 + e_2$)

and Y_β a y-shear by β ($e_1 \mapsto e_1 + \beta e_2$, $e_2 \mapsto e_2$)

with $\alpha = \gamma = -\tan(\theta/2)$, $\beta = \sin\theta$.

Indeed,

$$\begin{aligned} X_\alpha Y_\beta X_\alpha(e_1) &= X_\alpha Y_\beta(e_1) \\ &= X_\alpha(e_1 + \beta e_2) \\ &= e_1 + \beta\alpha e_1 + \beta e_2 \\ &= (1+\beta\alpha)e_1 + \beta e_2 \end{aligned}$$

$$\begin{aligned} X_\alpha Y_\beta X_\alpha(e_2) &= X_\alpha Y_\beta(\alpha e_1 + e_2) \\ &= X_\alpha(\alpha e_1 + \alpha\beta e_2 + e_2) \\ &= X_\alpha(\alpha e_1 + (1+\alpha\beta)e_2) \\ &= \alpha e_1 + (1+\alpha\beta)(\alpha e_1 + e_2) \\ &= (2\alpha + \alpha^2\beta)e_1 + (1+\alpha\beta)e_2 \end{aligned}$$

Now for $\cancel{\alpha = \tan(\theta/2)}$, $\beta = \sin\theta$, have

$$\begin{aligned} 1+\alpha\beta = \cos\theta &\iff 1 + \alpha\sin\theta = \cos\theta \\ &\iff \alpha = \frac{\cos\theta - 1}{\sin\theta} \\ &\iff \alpha = -\tan(\theta/2) \quad \text{(by trigonometry)} \end{aligned}$$

Finally, $2\alpha + \alpha^2\beta = \alpha(1 + (1+\alpha\beta)) = \alpha(1+\cos\theta) = \cos\theta$ (more trig).

<u>TPS</u> Express reflection through $y=-x$ as a comp'n of scale, shear, & reflect transformations.

Matrices

Recall $M_{m\times n}(F)$ = $m\times n$ matrices A w/ entries $A_{ij}\in F$ (entry in i-th row, j-th column)

e.g. $A=\begin{pmatrix}1&2&6\\7&0&-1\end{pmatrix}\in M_{2\times3}(\mathbb{Q})$

has $A_{1,3}=6$.

$M_{m\times n}(F)$ is an F-vs with entry-wise add'n & scalar mult; its dimension is mn w/ basis $\{E(i,j) \mid 1\le i\le m,\ 1\le j\le n\}$

$E(i,j)$: 1 in ij posn, 0 o/w

Now define a product on matrices

$$M_{m\times p}(F)\times M_{p\times n}(F)\longrightarrow M_{m\times n}(F)$$

cols on left = # rows on right

$$(AB)_{ij}=\sum_{k=1}^{p}A_{ik}\cdot B_{kj}$$

A_{ik} steps through i-th row of A; B_{kj} steps through j-th col of B

e.g. $\begin{pmatrix}\boxed{1\ 0\ 2}\\3\ 1\ 4\end{pmatrix}\begin{pmatrix}1&\boxed{0}\\3&\boxed{-1}\\5&\boxed{3}\end{pmatrix}=\begin{pmatrix}11&\boxed{6}\\26&11\end{pmatrix}$ $\qquad (AB)_{12}=\sum_{k=1}^{3}A_{1k}B_{k2}$

$$=1\cdot0+0\cdot(-1)+2\cdot3=6$$

This is the "dot product" of i-th row w/ j-th column where $(a_1,\dots,a_p)\cdot(b_1,\dots,b_p)=\sum_{k=1}^{p}a_kb_k$.

Prop Let $A\in M_{m\times p}(F)$, $B,B'\in M_{p\times n}(F)$, $C\in M_{n\times q}(F)$, $\lambda\in F$, $D,D'\in M_{r\times m}(F)$. Then

(a) $\lambda(AB)=(\lambda A)B=A(\lambda B)$

(b) $A(BC)=(AB)C$

(c) $A(B+B')=(AB)+(AB')$

(d) $(D+D')A=DA+D'A$

Pf 1 of (b)

$$
\begin{aligned}
(A(BC))_{ij} &= \sum_{k=1}^{p} A_{ik} (BC)_{kj} \\
&= \sum_{k=1}^{p} A_{ik} \left(\sum_{l=1}^{n} B_{kl} C_{lj} \right) \\
&= \sum_{k=1}^{p} \sum_{l=1}^{n} A_{ik} (B_{kl} C_{lj}) \\
&= \sum_{l=1}^{n} \sum_{k=1}^{p} A_{ik} (B_{kl} C_{lj}) \\
&= \sum_{l=1}^{n} \sum_{k=1}^{p} (A_{ik} B_{kl}) C_{lj} \\
&= \sum_{l=1}^{n} \left(\sum_{k=1}^{p} A_{ik} B_{kl} \right) C_{lj} \\
&= \sum_{l=1}^{n} (AB)_{il} C_{lj} \\
&= ((AB)C)_{ij} \qquad \square
\end{aligned}
$$

Pf 2 of (b) We will build a dictionary (bij'n) — in fact, linear iso

$$M_{m\times n}(F) \longleftrightarrow \mathcal{L}(F^n, F^m)$$

$$A \longleftrightarrow (x \longmapsto Ax) \qquad \text{(for } x \text{ col vector of length } n\text{)}$$

$$\text{mult} \longleftrightarrow \text{composition}$$

Composition is associative, so matrix mult is as well! $\square$

Matrix mult'n is not commutative! (Even when defined)

TPS How does matrix rank interact w/ scalar mult, add'n?

Identity matrices: The $n\times n$ identity matrix I_n has 1's on diag, 0's elsewhere. Whenever defined, $AI = A$, $IB = B$.

Inverses: $A \in M_{m\times n}(F)$, $B \in M_{n\times m}(F)$. If $AB = I_m$, call A a left inverse for B, B a right inverse for A.

e.g. $A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}$ $B = \begin{pmatrix} 1 & -1 \\ 0 & 0 \\ 0 & 1 \end{pmatrix}$ $\Rightarrow AB = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

$BA = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix}$

so A is a left-inverse for B but B is not a left-inverse for A.

Thm: $A, B \in M_{n\times n}(F)$. TFAE: ① $AB = I_n$, ② $BA = I_n$.

In this case, say A, B invertible, $A^{-1} = B$, $B^{-1} = A$.

TFAE: ③ A is invertible, ④ $\operatorname{rank}(A) = n$, ⑤ the reduced echelon form of A is I_n.

Proof follows from an algorithm for computing inverses.

Calculating the inverse

An example first: Let $A = \begin{pmatrix} 0 & 3 & -1 \\ 1 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix}$. A right inverse to A would satisfy

$$\begin{pmatrix} 0 & 3 & -1 \\ 1 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix}\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

This reduces to 3 problems:

$$A\begin{pmatrix} a \\ d \\ g \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad A\begin{pmatrix} b \\ e \\ h \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad A\begin{pmatrix} c \\ f \\ i \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

$$\updownarrow \qquad\qquad \updownarrow \qquad\qquad \updownarrow$$

$$\left(\begin{array}{ccc|c} 0 & 3 & -1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & -1 & 0 & 0 \end{array}\right) \quad \left(A \,\middle|\, \begin{matrix} 0 \\ 1 \\ 0 \end{matrix}\right) \quad \left(A \,\middle|\, \begin{matrix} 0 \\ 0 \\ 1 \end{matrix}\right)$$

Combine into one "super"-augmented matrix reduction:

$$(A|I) = \left(\begin{array}{ccc|ccc} 0 & 3 & -1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & -1 & 0 & 0 & 0 & 1 \end{array}\right) \xrightarrow{r_1 \leftrightarrow r_2} \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 3 & -1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 1 \end{array}\right)$$

$$\xrightarrow{r_3 \to r_3 - r_1} \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 3 & -1 & 1 & 0 & 0 \\ 0 & -1 & -1 & 0 & -1 & 1 \end{array}\right) \xrightarrow[r_3 \to -r_3]{r_2 \leftrightarrow r_3} \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 & -1 \\ 0 & 3 & -1 & 1 & 0 & 0 \end{array}\right)$$

$$\xrightarrow{r_3 \to r_3 - 3r_2} \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 & -1 \\ 0 & 0 & -4 & 1 & -3 & 3 \end{array}\right) \xrightarrow{r_3 \to -r_3/4} \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 & -1 \\ 0 & 0 & 1 & \frac{-1}{4} & \frac{3}{4} & \frac{-3}{4} \end{array}\right)$$

$$\xrightarrow[r_2 \to r_2 - r_3]{r_1 \to r_1 - r_3} \left(\begin{array}{ccc|c} 1 & 0 & 0 & \\ 0 & 1 & 0 & \underbrace{\begin{matrix} 1/4 & 1/4 & 3/4 \\ 1/4 & 1/4 & -1/4 \\ -1/4 & 3/4 & -3/4 \end{matrix}}_{A^{-1}} \\ 0 & 0 & 1 & \end{array}\right)$$

This easily generalizes to $A \in M_{n\times n}(F)$.

<u>Algorithm for computing A^{-1}</u>:

Perform row ops to $(A|I_n)$ to compute REF of A:

$$(A|I_n) \longrightarrow (\mathrm{REF}(A)|B).$$

Then if $\mathrm{REF}(A) \neq I_n$, then $\mathrm{rank}(A) < n$ and A has no inverse. If $\mathrm{REF}(A) = I_n$, then $\mathrm{rank}(A) = n$, and B = right inverse of A.

Performing the same algorithm on $(A^T|I_n)$ computes left inverse C. (Note $\mathrm{rank}(A^T) = \mathrm{rank}(A)$.) Then $AB = I$ & $CA = I$ so

$C(AB) = CI = C \Rightarrow (CA)B = C$, but $(CA) = I$, so $IB = C$,

i.e. $B = C$, as desired. Thus the algorithm computes the 2-sided inverse $B = A^{-1}$.

Matrices & Linear Transformations

$$M_{m\times n}(F) \longrightarrow \mathcal{L}(F^n, F^m)$$
$$A \longmapsto (f_A : x \mapsto Ax) \quad \text{for } x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Linearity of f_A: $f_A(u+\lambda v) = A(u+\lambda v) = Au + \lambda Av$
$$= f_A(u) + \lambda f_A(v).$$

e.g. $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$ $\quad f_A : F^3 \longrightarrow F^2$
$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \longmapsto \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x+2y+3z \\ 4x+5y+6z \end{pmatrix}$$

Notes ① Coeffs come from corresponding row
② $f_A(e_j) = j$-th column of A.

This gives us an idea on producing an inverse function
$$\mathcal{L}(F^n, F^m) \longrightarrow M_{m\times n}(F)$$
$$f \longmapsto (f(e_1) \cdots f(e_n)) \quad \text{with } f(e_j) \text{ written as a column}$$

Facts (a) $M_{m\times n}(F) \longrightarrow \mathcal{L}(F^n, F^m)$ is linear
(b) and a bij, hence an isomorphism

TPS What does it mean for $A \mapsto f_A$ to be linear?

Q How can we encode a lin trans $f : V \to W$ with a ~~lin trans~~ matrix?
(V, W fin dim)

Idea Choosing a basis for V is equiv to producing an isomorphism $V \xrightarrow{\cong} F^n$. Do this for W as well then use the above assignment.

Suppose $\alpha = \{v_1, \dots, v_n\}$ is an ordered basis of V and $v = c_1 v_1 + \cdots + c_n v_n$ has coords $(c_1, \dots, c_n)$. Get $\phi_\alpha : V \xrightarrow{\cong} F^n$
$$v \longmapsto (c_1, \dots, c_n).$$

Similarly, if $\beta = \{w_1, \dots, w_m\}$ basis of W, get $\phi_\beta : W \xrightarrow{\cong} F^m$.

The $m \times n$ matrix A_α^β representing f wrt these bases is the one making

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ \phi_\alpha \downarrow \cong & & \cong \downarrow \phi_\beta \\ F^n & \xrightarrow[f_{A_\alpha^\beta}]{} & F^m \end{array}$$

commute.

We have

$$\begin{array}{ccc} v_j & \longmapsto & f(v_j) \\ \downarrow & & \\ e_j & \longmapsto & j\text{th column of } A_\alpha^\beta \end{array}$$

so the j-th column of A_α^β must be the β-coords of $f(v_j)$.

I.e. $\boxed{A_\alpha^\beta = (a_{ij}) \text{ where } f(v_j) = a_{1j}w_1 + \cdots + a_{mj}w_m}$.

e.g. $f : \mathbb{R}[x]_{\le 2} \longrightarrow \mathbb{R}[x]_{\le 3}$

$$p \longmapsto xp + p'$$

$\alpha = \{1, x, x^2\}$, $\beta = \{1, x, x^2, x^3\}$

$f(1) = x = 0\cdot 1 + 1\cdot x + 0\cdot x^2 + 0\cdot x^3$

$f(x) = x^2 + 1 = 1\cdot 1 + 0\cdot x + 1\cdot x^2 + 0\cdot x^3$

$f(x^2) = x^3 + 2x = 0\cdot 1 + 2\cdot x + 0\cdot x^2 + 1\cdot x^3$

Thus $A_\alpha^\beta = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ rep's f wrt α, β.

TPS
- What is dim $\mathcal{L}(V, W)$ for dim $V = n$, dim $W = m$?
- How is this related to V^*?

Recall $M_{m\times n}(F) \xrightarrow{\cong} \mathcal{L}(F^n, F^m)$

$$A \longmapsto (f_A : x \mapsto Ax)$$

$$A_f := (f(e_1) \cdots f(e_n)) \longleftarrow f$$

Image For $f: F^n \to F^m$, $im(f) = \{f(x) \mid x \in F^n\} \subseteq F^m$.

$$= span\{f(e_1), \dots, f(e_n)\}$$

Second equality b/c $x = (x_1, \dots, x_n) = x_1 e_1 + \cdots + x_n e_n$
$\longmapsto f(x) = x_1 f(e_1) + \cdots + x_n f(e_n) \in span\{f(e_1), \dots, f(e_n)\}$, giving $im(f) \subseteq span\{f(e_1), \dots, f(e_n)\}$. The other inclusion follows b/c each $f(e_j) \in im(f)$ & $im(f)$ is a subspace.

Prop $im(f) =$ column space of A_f

$rank(f) = rank(A_f)$. $\square$

e.g. $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ -3 & 2 \end{pmatrix}$ then $im(f_A) = span\{(1,0,-3), (1,1,2)\}$

Indeed, $f_A(x,y) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ -3 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x+y \\ y \\ -3x+2y \end{pmatrix} = x\begin{pmatrix} 1 \\ 0 \\ -3 \end{pmatrix} + y\begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$.

Composition

Thm $A \in M_{m\times p}(F)$, $B \in M_{p\times n}(F)$ with associated lin trans $f_A : F^p \to F^m$, $f_B : F^n \to F^p$. Then

$$f_A \circ f_B = f_{AB}.$$

Pf For $x \in F^n$, $(f_A \circ f_B)(x) = f_A(f_B(x)) = A(Bx) = (AB)x$

$$= f_{AB}(x). \quad \square$$

Note Matrix mult'n was invented so that this would happen.

Inverses Suppose $A, B \in M_{n\times n}(F)$, $AB = I_n$.

TPS What ~~$\cdot$~~ is f_{I_n}?

Get $f_A \circ f_B = id_{F^n}$. ~~It follows that f_B is injective, so has kernel $\{0\}$ and thus $rank(f_B) = rank$~~

It follows that f_A is surjective $\Rightarrow rank(f_A) = rank(A) = n$.
In particular, $rank(A) < n \Rightarrow A$ doesn't have a right inverse.
Similarly, f_B injective $\Rightarrow nullity(f_B) = 0 \Rightarrow rank(f_B) = rank(B) = n$
so $rank(B) < n \Rightarrow B$ doesn't have a left inverse.
This completes our argument about matrix inversion! $\square$

Time allowing How do kernels fit into this picture?

Dual Vector Spaces

Defn (1) For V an F-vector space, $V^* := \mathcal{L}(V, F)$ is the dual space of V. Elements of V^* are called linear functionals.

(2) If V is finite dimensional with basis $\{v_1, \ldots, v_n\}$, define $v_i^* \in V^*$ for $i \in \{1, \ldots, n\}$ by its action on $\{v_1, \ldots, v_n\}$:

$$v_i^*(v_j) = \begin{cases} 1 & \text{if } i=j \\ 0 & \text{if } i \neq j. \end{cases}$$

Prop $\{v_1^*, \ldots, v_n^*\}$ is a basis of V^*. In particular, if $\dim V < \infty$, then $\dim V = \dim V^*$.

Pf Since $\dim V < \infty$, $\dim V^* = \dim \mathcal{L}(V, F) = \dim M_{1 \times \dim V}(F)$
$= \dim V$.

Since there are n v_i^*'s, it suffices to show they are linearly independent. If $a_1 v_1^* + \cdots + a_n v_n^* = 0$, then, applying this eqn to v_j, get $a_j = 0$. This holds for all j, so the v_i^* are lin ind. $\square$

Defn $\{v_1^*, \ldots, v_n^*\}$ is the dual basis of $\{v_1, \ldots, v_n\}$.

Double Duals

Defn The dual of V^*, namely V^{**}, is the double dual of V.

Thm There is a natural linear injection $V \longrightarrow V^{**}$. If V is finite dimensional, then this linear transformation is an isomorphism.

Pf Let $v \in V$. Define the evaluation at v map

$$\begin{aligned} ev_v : V^* &\longrightarrow F \\ f &\longmapsto f(v). \end{aligned}$$

Then $ev_v(f+\lambda g) = (f+\lambda g)(v) = f(v) + \lambda g(v) = ev_v(f) + \lambda ev_v(g)$

so ev_v is a linear transformation $V^* \to F$, i.e., $ev_v \in V^{**}$.

We thus get a natural map
$$\varphi: V \longrightarrow V^{**}$$
$$v \longmapsto ev_v$$

and φ is linear: $ev_{v+\lambda w}(f) = f(v+\lambda w) = f(v) + \lambda f(w) = ev_v(f) + \lambda ev_w(f)$

for all $f \in V^*$, $v, w \in V$, $\lambda \in F$. Thus

$$\varphi(v+\lambda w) = ev_{v+\lambda w} = ev_v + \lambda ev_w = \varphi(v) + \lambda\varphi(w).$$

For injectivity, we punt — this requires knowing that V has a basis (containing any specified nonzero v), but we have only proven this for fin dim'l vector spaces. Nevertheless, it's true!

For fin dim'l V, given $v \neq 0 \in V$, $\exists$ basis $B \ni v$.

Define
$$f: V \longrightarrow F$$
$$v \longmapsto 1$$
$$B \setminus \{v\} \longmapsto 0$$
. Then $f \in V^*$ and $ev_v(f) = f(v) = 1$.

Thus $\varphi(v) = ev_v \neq 0 \implies \ker\varphi = \{0\}$

$\implies \varphi$ inj.

By rank-nullity, φ is an isomorphism since

$\dim V = \dim V^* = \dim V^{**}$. $\square$

<u>Dual transformations</u> and <u>transpose matrices</u>

Given $\varphi: V \to W$ linear and $f \in W^*$, we have $f \circ \varphi \in V^*$.

<u>Prop</u> The assignment
$$\varphi^*: W^* \longrightarrow V^*$$
$$f \longmapsto f\circ\varphi$$
is linear.

<u>Pf</u> $\varphi^*(f+\lambda g) = (f+\lambda g)\circ\varphi = f\circ\varphi + \lambda(g\circ\varphi) = \varphi^* f + \lambda\varphi^* g$.

<u>Defn</u>
$$(\,)^T: M_{m\times n}(F) \longrightarrow M_{n\times m}(F)$$
$$(a_{ij}) \longmapsto (a_{ji}) = (a_{ij})^T$$
is the <u>transpose</u>.

<u>Thm</u> Let $\alpha = \{v_1, \dots, v_n\}$, $\beta = \{w_1, \dots, w_m\}$ be ordered basis of V, W, resp.

~~Then~~ Let $A_\alpha^\beta(\varphi)$ denote the matrix of φ wrt α, β.

Let $\alpha^* = \{v_1^*, \dots, v_n^*\}$, $\beta^* = \{w_1^*, \dots, w_m^*\}$ be the dual bases.

Then $A^{\alpha^*}_{\beta^*}(\varphi^*) = A^{\beta}_{\alpha}(\varphi)^T$ for any lin trans $\varphi: V \to W$.

Pf An exercise in (advanced!) bookkeeping:

Let $A^{\beta}_{\alpha}(\varphi) = (a_{ij})$ so that $\varphi(v_j) = \sum_{i=1}^{m} a_{ij} w_i$, $1 \le j \le n$.

Now $\varphi^*(w_k^*)(v_j) = (w_k^* \circ \varphi)(v_j) = w_k^*\left(\sum_{i=1}^{m} a_{ij} w_i\right) = a_{kj}$.

Also $\left(\sum_{i=1}^{n} a_{ki} v_i^*\right)(v_j) = a_{kj}$ for all j. Thus $\varphi^*(w_k^*)$ and $\sum_{i=1}^{n} a_{ki} v_i^*$ agree on a basis $\Longrightarrow \varphi^*(w_k^*) = \sum_{i=1}^{n} a_{ki} v_i^*$.

This says that the k-th column of $A^{\alpha^*}_{\beta^*}(\varphi^*)$ is equal to the k-th row of $A^{\beta}_{\alpha}(\varphi)$ $\forall k$, so

$$A^{\alpha^*}_{\beta^*}(\varphi^*) = A^{\beta}_{\alpha}(\varphi)^T. \qquad \square$$

Recall $V^* = \mathcal{L}(V,F)$

$$\mathcal{L}(V,W) \longrightarrow \mathcal{L}(W^*,V^*)$$
$$\varphi \longmapsto \varphi^* = f\circ\varphi$$

Q How are $\ker\varphi$, $\operatorname{im}\varphi$ related to $\ker\varphi^*$, $\operatorname{im}\varphi^*$?

Defn Let $S \subseteq V$ be a subset of V. The annihilator of S is the subset of V^* defined by $S^\circ = \{f \in V^* \mid f(s) = 0 \ \forall s \in S\}$.

e.g. $V = \mathbb{R}[x]$, $S = \{p \in V \mid p(0) = 0\}$. (so S = multiples of x = const term 0 polynomials)

For $\lambda \in \mathbb{R}$, define $f_\lambda \in V^*$ by $f_\lambda(p) = \lambda p(0)$.

Claim $S^\circ = \{f_\lambda \mid \lambda \in \mathbb{R}\}$.

Indeed, if $p \in S$ then $f_\lambda(p) = \lambda p(0) = \lambda \cdot 0 = 0$ so $f_\lambda \in S^\circ$.

Now suppose $g \in S^\circ$. Restricting g to $\mathbb{R}$ (viewed as const polys) gives a linear trans $g|_{\mathbb{R}} : \mathbb{R} \to \mathbb{R}$. Let $\lambda = g(1)$. Then $g|_{\mathbb{R}}(r) = \lambda r \ \forall r \in \mathbb{R}$. Since $g \in S^\circ$, $g(x^i) = 0$ for $i > 0$. Thus if $p = a_n x^n + \cdots + a_0 \in V$, then $g(p) = a_n g(x^n) + \cdots + a_1 g(x) + a_0 g(1)$
$$= \lambda a_0 = \lambda p(0).$$

Thus $g = f_\lambda$. $\square$

Note $\{f_\lambda \mid \lambda \in \mathbb{R}\}$ is a subspace of V^*.

HW $S^\circ \subseteq V^*$ is a subspace $\forall$ subset $S \subseteq V$.

Lemma Suppose V is fin dim. Let $S \subseteq V$ be a subspace, and $i : S \to V$ be the inclusion map $i(s) = s$. Then $\operatorname{im}(i^*) = S^*$.

Pf HW $\square$

Prop For V fin dim, $S \subseteq V$ subspace,
$$\dim(S) + \dim(S^\circ) = \dim V.$$

Pf $\operatorname{rank}(i^*) + \operatorname{null}(i^*) = \dim V^*$, $\ker(i^*) = S^\circ$, $\dim V = \dim V^*$, so
$$\underbrace{\dim(S^*)}_{=\dim S} + \dim(S^\circ) = \dim V \qquad \square$$

Thm Suppose V, W fin dim, $\varphi \in \mathcal{L}(V,W)$. Then

(a) $\ker(\varphi^*) = \operatorname{im}(\varphi)^\circ$

(b) $\operatorname{null}(\varphi^*) = \operatorname{null}(\varphi) + \dim W - \dim V$.

Pf For (a), note that if $f \in W^*$, then

$$\begin{aligned} f \in \ker(\varphi^*) &\iff f \circ \varphi = 0 \\ &\iff f(\varphi(v)) = 0 \quad \forall v \in V \\ &\iff f(w) = 0 \quad \forall w \in \operatorname{im}(\varphi) \\ &\iff f \in \operatorname{im}(\varphi)^\circ. \end{aligned}$$

For (b), apply the prop to $S = \operatorname{im}(\varphi) \subseteq W$ to obtain

$$\dim \operatorname{im} \varphi + \dim \operatorname{im}(\varphi)^\circ = \dim W.$$

But $\dim \operatorname{im} \varphi = \operatorname{rank}(\varphi)$ & $\dim \operatorname{im}(\varphi)^\circ = \dim \ker \varphi^* = \operatorname{null} \varphi^*$

so $$\operatorname{rank} \varphi + \operatorname{null} \varphi^* = \dim W.$$

By rank-nullity, $\operatorname{rank} \varphi = \dim V - \operatorname{null} \varphi$, so

$$\operatorname{null} \varphi^* = \operatorname{null} \varphi + \dim W - \dim V. \quad \square$$

Cor φ^* is inj $\iff \varphi$ is surjective.

Pf HW $\square$

Thm Suppose V, W fin dim, $\varphi \in \mathcal{L}(V,W)$. Then

(a) $\operatorname{rank} \varphi^* = \operatorname{rank} \varphi$

(b) $\operatorname{im}(\varphi^*) = \ker(\varphi)^\circ$

Pf For (a), apply rank-nullity to φ & φ^*:

$$\begin{aligned} \operatorname{rank} \varphi^* &= \dim W^* - \operatorname{null} \varphi^* \\ \operatorname{rank} \varphi &= \dim V - \operatorname{null} \varphi \end{aligned}$$

$$\Rightarrow \operatorname{rank} \varphi^* - \operatorname{rank} \varphi = \operatorname{null} \varphi + \dim W - \dim V - \operatorname{null} \varphi^* = 0.$$

For (b), suppose $f \in V^*$ is in the image of φ^*, so that $f = \varphi^*(g)$ for some $g \in W^*$. To show $f \in \ker(\varphi)^\circ$, must show $f(v) = 0$ $\forall v \in \ker\varphi$.

For $v \in \ker\varphi$, $f(v) = g(\varphi(v)) = g(0) = 0$ ✓ so $\operatorname{im}\varphi^* \subseteq \ker(\varphi)^\circ$.

Now check dimensions are equal, proving equality:

By the Prop, $\operatorname{null}\varphi + \dim\ker(\varphi)^\circ = \dim V$, so

$$\begin{aligned}\dim\ker(\varphi)^\circ &= \dim V - \operatorname{null}\varphi \\ &= \operatorname{rank}(\varphi) \\ &= \operatorname{rank}\varphi^* \\ &= \dim\operatorname{im}\varphi^*. \qquad \square\end{aligned}$$

__Cor__ φ^* surj iff φ is inj.

__Pf__ HW $\square$

Determinants

Defn The determinant is a multilinear, alternating function of the rows of a square matrix, $\det : M_{n\times n}(F) \to F$, normalized so that its value on the identity matrix is 1.

To explain, for $A \in M_{n\times n}(F)$ with rows $r_1, \dots, r_n \in F^n$, write $\det(r_1, \dots, r_n)$ for $\det A$. Then

① Multilinear: The determinant is a linear fn wrt each row:

$$\det(r_1, \dots, r_{i-1}, r_i + \lambda r_i', r_{i+1}, \dots, r_n)$$
$$= \det(r_1, \dots, r_n) + \lambda \det(r_1, \dots, r_{i-1}, r_i', r_{i+1}, \dots, r_n).$$

② Alternating: The determinant is 0 if two of the rows are equal:

$$\det(r_1, \dots, r_n) = 0 \text{ if } r_i = r_j \text{ for some } i \neq j.$$

③ Normalized: $\det(I_n) = \det(e_1, \dots, e_n) = 1$.

Thm For each $n \geq 0$, $\exists!\ \det : M_{n\times n}(F) \to F$.

For now, assume det exists satisfying ①-③.

Prop [det & row ops] Let $A, B \in M_{n\times n}(F)$.

① If B is obtained from A by swapping two rows, $\det B = -\det A$.
② If B — " — by scaling a row by λ, $\det B = \lambda \det A$.
③ If B — " — by adding a λ scalar mult of one row to another, then $\det B = \det A$

Pf ① In the case of swapping r_1, r_2 in A to get B, compute

$$0 = \det(r_1 + r_2, r_1 + r_2, r_3, \dots, r_n) \quad \text{[alt]}$$
$$= \det(r_1, r_1 + r_2, r_3, \dots, r_n) + \det(r_2, r_1 + r_2, r_3, \dots, r_n) \quad \text{[mult]}$$
$$= \det(r_1, r_1, r_3, \dots, r_n) + \det(r_1, r_2, r_3, \dots, r_n) + \det(r_2, r_1, r_3, \dots, r_n) + \det(r_2, r_2, r_3, \dots)$$
$$= 0 + \det A + \det B + 0 \implies \det B = -\det A$$

② Implied by multilinearity.

③ $$\det(r_1, \lambda r_1 + r_2, r_3, \dots, r_n) = \lambda \cancel{\det(r_1, r_1, \dots, r_n)}^{0} + \det(r_1, r_2, \dots, r_n)$$
$$= \det(r_1, \dots, r_n)$$

e.g.
$$\det\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \det\big((a,b),\ (c,d)\big)$$
$$= \det(ae_1 + be_2,\ ce_1 + de_2)$$
$$= a\det(e_1, ce_1 + de_2) + b\det(e_2, ce_1 + de_2)$$
$$= ac\,\cancelto{0}{\det(e_1,e_1)} + ad\det(e_1,e_2) + bc\det(e_2,e_1) + bd\,\cancelto{0}{\det(e_2,e_2)}$$
$$= ad\det I_2 - bc\det I_2$$
$$= ad - bc.$$

The prop turns Gauss-Jordan reduction into an algorithm for computing det!

e.g.
$$\det\begin{pmatrix} 1 & 2 & -2 \\ 9 & 4 & 0 \\ 2 & 2 & 4 \end{pmatrix} = \det\begin{pmatrix} 1 & 2 & -2 \\ 0 & -14 & 18 \\ 0 & -2 & 8 \end{pmatrix}$$
$$= -\det\begin{pmatrix} 1 & 2 & -2 \\ 0 & -2 & 8 \\ 0 & -14 & 18 \end{pmatrix}$$
$$= 2\det\begin{pmatrix} 1 & 2 & -2 \\ 0 & 1 & -4 \\ 0 & -14 & 18 \end{pmatrix}$$
$$= 2\det\begin{pmatrix} 1 & 2 & -2 \\ 0 & 1 & -4 \\ 0 & 0 & -38 \end{pmatrix}$$
$$= 2(-38)\begin{pmatrix} 1 & 2 & -2 \\ 0 & 1 & -4 \\ 0 & 0 & 1 \end{pmatrix}$$
$$= 2(-38)\det I_3$$
$$= -76.$$

TPS [in groups of 4]

- What is $\det\begin{pmatrix} 4 & 2 & -3 & 8 \\ 0 & 5 & 1 & 3 \\ 0 & 0 & 2 & 6 \\ 0 & 0 & 0 & 3 \end{pmatrix}$?
- What is $\det\begin{pmatrix} \diagdown & \\ 0 & \end{pmatrix}$?

if time

Prop TFAE:

1. $\det A \neq 0$
2. $\operatorname{rank}(A) = n$
3. A invertible.

<u>Recall</u> $\det : M_{n\times n}(F) \to F$ is the unique multilinear, alternating function of the rows of an $n\times n$ matrix, normalized so that $\det(I_n)=1$

Know:
- swapping rows switches sign
- scaling a row scales det
- adding a scaler multiple of one row to another does nothing
- $\det A \neq 0 \iff \operatorname{rank}(A)=n \iff A$ is invertible

To do:
- $\det A^T = \det A$ — Today
- $\det AB = \det A \det B$ — ~~Today~~ HW8F
- row/column expansion — Friday
- permutation expansion $\det A = \sum_{\sigma\in\Sigma_n} \operatorname{sgn}(\sigma) A_{1\sigma(1)}\cdots A_{n\sigma(n)}$ — Wed
- Over $\mathbb{R}$, $|\det A| = \operatorname{vol}(A\cdot[0,1]^n)$ — next Monday
- det exists and is unique — ~~[illegible]~~ Friday

<u>Elementary Matrices</u> An $n\times n$ matrix is called an elementary matrix if it is obtained from I_n through a single row operation.

$$\begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 0 & 1 & & \\ & & & \ddots & & \\ & & 1 & 0 & & \\ & & & & \ddots & \\ & & & & & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & \lambda & & \\ & & & & 1 & \\ & & & & & \ddots \\ & & & & & & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & \ddots & & \\ & & \lambda & & 1 & \\ & & & & & \ddots \\ & & & & & & 1 \end{pmatrix}$$

<u>Fact</u> If E is an $n\times n$ elementary matrix and $A\in M_{n\times k}(F)$, then EA is the matrix obtained from A by performing the row operation associated with E.

<u>Upshot</u> You can perform row ops via mult'n by elementary matrices

<u>e.g.</u> $E=\begin{pmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \leftrightsquigarrow r_2 \to r_2 - 3r_1$

$$\begin{pmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 0 & -1 & 2 \\ 1 & 5 & 6 & 7 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & -6 & -10 & -10 \\ 1 & 5 & 6 & 7 \end{pmatrix}$$

Note $REF(A) = E_\ell \cdots E_2 E_1 A$ for some elementary matrices E_i.

Now look at $\det A^T$ vs $\det A$.

e.g. $\det \begin{pmatrix} a & b \\ c & d \end{pmatrix}^T = \det \begin{pmatrix} a & c \\ b & d \end{pmatrix} = ad - bc$. ✓

Thm $\det AB = \det A \cdot \det B$

Pf HW □

Prop (1) $(AB)^T = B^T A^T$

(2) $(A^T)^{-1} = (A^{-1})^T$ for A invertible.

Pf (1) ✓

(2) The linear trans version says $(f^{-1})^* = (f^*)^{-1}$, which we now prove: For $f: V \to W$ linear iso with inverse $f^{-1}: W \to V$

here $f \circ f^{-1} = id_W \Rightarrow (f \circ f^{-1})^* = id_{W^*}$

$\Rightarrow (f^{-1})^* \circ f^* = id_{W^*}$

Similarly, $f^* \circ (f^{-1})^* = id_{V^*}$, so $(f^{-1})^* = (f^*)^{-1}$. □

Lemma For E elementary, $\det E = \det E^T \neq 0$.

Pf (1) For E given by swapping i, j, $E = E^T$ & $\det E = -1 = \det E^T$.

(2) For E scaling of row by $\lambda \neq 0$, $E = E^T$ & $\lambda = \det E = \det E^T \neq 0$.

(3) For E given by $r_j \to r_j + \lambda r_i$, E^T given by $r_i \to r_i + \lambda r_j$

so $\det E = \det E^T = \det I_n = 1$. □

Thm $\det A = \det A^T$

Pf There are elementary matrices $E_1, \ldots, E_\ell$ s.t.

$$REF(A) = E_\ell \cdots E_1 A \quad (*)$$

$\Rightarrow \det REF(A) = \det E_\ell \cdots \det E_1 \det A$

$\Rightarrow \det A = \det(E_\ell)^{-1} \cdots \det(E_1)^{-1} \det REF(A)$

Taking $\circledast^T$: $\quad REF(A)^T = A^T E_1^T \cdots E_\ell^T$.

Taking det and solving for $\det A^T$ (using $\det E_i = \det E_i^T$):

$$\det A^T = \det(E_1)^{-1} \cdots \det(E_\ell)^{-1} \det REF(A)^T$$
$$= \det(E_\ell)^{-1} \cdots \det(E_1)^{-1} \det REF(A)^T$$

Two cases: (1) $\operatorname{rank} A = n \Longleftrightarrow REF(A) = I_n \Longrightarrow \det REF(A) = 1$

and $REF(A)^T = I_n^T = I_n$ so $\det REF(A)^T = 1$ as well. Thus

$$\det A = \det(E_\ell)^{-1} \cdots \det(E_1)^{-1} = \det A^T.$$

(2) $\operatorname{rank} A < n \Longrightarrow \operatorname{rank} A^T = \operatorname{rank} A < n$

$$\Longrightarrow \det A = \det A^T = 0. \qquad \square$$

<u>Cor</u> det is a multilinear, alternating function of the columns of a square matrix. $\square$

Permutation Expansion of the Determinant

Defn A permutation of a set X is a bijective fn $X \to X$. The set of all permutations of X is called the symmetric group on X.

The symmetric group on $\underline{n} = \{1, \dots, n\}$ is the symmetric group on n letters, denoted Σ_n (or S_n, or $\mathfrak{S}_n$).

Represent $\sigma \in \Sigma_n$ by the $2 \times n$ matrix

$$\underbrace{\begin{pmatrix} 1 & 2 & 3 & \cdots & n \\ \sigma(1) & \sigma(2) & \sigma(3) & \cdots & \sigma(n) \end{pmatrix}}_{\text{rearrangement of } 1, \dots, n}$$

Get $|\Sigma_n| = n!$

e.g. The 6 elements of Σ_3 are

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$$

Defn The permutation matrix corresponding to $\sigma \in \Sigma_n$ is the matrix

$$P_\sigma \in M_{n \times n}(F)$$

with i-th column $e_{\sigma(i)}$.

e.g. $\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \rightsquigarrow P_\sigma = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$

Prop (a) The $\sigma(i)$-th row of P_σ is e_i.

(b) $P_\sigma P_\sigma^T = I_n$ (P_σ is orthogonal)

(c) $\{P_\sigma \mid \sigma \in \Sigma_n\}$ is the set of matrices in $M_{n \times n}(F)$ with exactly one 1 in every row & column, 0's elsewhere.

(d) $P_{\sigma \circ \tau} = P_\sigma P_\tau \ \forall \sigma, \tau \in \Sigma_n$.

Pf (a) The cols of P_σ are $e_{\sigma(1)}, \dots, e_{\sigma(n)}$. If $j = \sigma(i)$, the j-th row of P_σ is $(e_{\sigma(1)j}, \dots, e_{\sigma(i)j}, \dots, e_{\sigma(n)j}) = e_i$.

(b) $(P_\sigma)_{ab} = e_{\sigma(b)a} = \delta_{\sigma(b),a}$. Thus

$$(P_\sigma P_\sigma^T)_{ij} = \sum_{k=1}^n (P_\sigma)_{ik} (P_\sigma^T)_{kj} = \sum_{k=1}^n (P_\sigma)_{ik} (P_\sigma)_{jk}$$

$$= \sum_{k=1}^n \delta_{\sigma(k)i} \delta_{\sigma(k)j} = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$$

$\Rightarrow P_\sigma P_\sigma^T = I_n$.

(c), (d): Moral exc. $\square$

<u>Rmk</u> (1) If $AA^T = I_n$ then $1 = \det A \det A^T = (\det A)^2$ so $\det A = \pm 1$. Thus $\det P_\sigma = \pm 1$ $\forall \sigma \in \Sigma_n$.

(2) Fun to think about permutation matrices as "non-attacking rooks" on an $n \times n$ chessboard.

<u>Defn</u> A <u>transposition</u> in Σ_n is a permutation which interchanges two elts of $\underline{n}$ and fixes all others. Write $(a\,b)$ for the transposition swapping a, b.

<u>Defn</u> The <u>sign</u> of a permutation $\sigma \in \Sigma_n$ is $\operatorname{sgn}(\sigma) = \det(P_\sigma) \in \{\pm 1\}$.

<u>Prop</u> Suppose σ is the composition of k permutations. Then $\operatorname{sgn}(\sigma) = (-1)^k$.

<u>Pf</u> If $k=1$, P_σ obtained from a single row swap so $\det P_\sigma = -1$.

If $\sigma = \tau_1 \circ \cdots \circ \tau_k$ for $k > 1$, τ_i transpositions, then

$P_\sigma = P_{\tau_1} \cdots P_{\tau_k}$ and $\det P_\sigma = \det P_{\tau_1} \cdots \det P_{\tau_k} = (-1)^k$. $\square$

<u>Rmk</u> In Math 332 you'll prove that every elt of Σ_n is a composition of transpositions.

<u>Thm</u> For every $A \in M_{n \times n}(F)$,

$$\det A = \sum_{\sigma \in \Sigma_n} \operatorname{sgn}(\sigma) A_{1\sigma(1)} \cdots A_{n\sigma(n)}$$

Pf The i-th row of A is $A_{i1}e_1+\cdots+A_{in}e_n$ so we seek to compute $\det(A_{11}e_1+\cdots+A_{1n}e_n, \ldots, A_{n1}e_1+\cdots+A_{nn}e_n)$.
Using multilinearity to expand get n^n terms, each of the form $A_{1j_1}A_{2j_2}\cdots A_{nj_n}\det(e_{j_1},\ldots,e_{j_n})$ $\star$
If any $e_{j_i}=e_{j_k}$, get 0, so only permutations

$\sigma=\begin{pmatrix}1&2&\cdots&n\\ j_1&j_2&\cdots&j_n\end{pmatrix}$ contribute. The matrix with rows $e_{j_1},\ldots,e_{j_n}$ is P_σ^T with $\det P_\sigma^T=\det P_\sigma=\operatorname{sgn}(\sigma)$. Thus the contribution of $\star$ is $A_{1\sigma(1)}\cdots A_{n\sigma(n)}\operatorname{sgn}(\sigma)$. $\square$

e.g.

$\begin{pmatrix}1&2&3\\1&2&3\end{pmatrix}$ $\begin{pmatrix}\boxed{a_{11}}&a_{12}&a_{13}\\a_{21}&\boxed{a_{22}}&a_{23}\\a_{31}&a_{32}&\boxed{a_{33}}\end{pmatrix}$ $a_{11}a_{22}a_{33}$

$\begin{pmatrix}1&2&3\\2&1&3\end{pmatrix}$ $\begin{pmatrix}&\square&\\\square&&\\&&\square\end{pmatrix}$ $-a_{12}a_{21}a_{33}$

$\begin{pmatrix}1&2&3\\3&2&1\end{pmatrix}$ $\begin{pmatrix}&&\square\\&\square&\\\square&&\end{pmatrix}$ $-a_{13}a_{22}a_{31}$

$\begin{pmatrix}1&2&3\\1&3&2\end{pmatrix}$ $\begin{pmatrix}\square&&\\&&\square\\&\square&\end{pmatrix}$ $-a_{11}a_{23}a_{32}$

$\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix}$ $\begin{pmatrix}&\square&\\&&\square\\\square&&\end{pmatrix}$ $a_{12}a_{23}a_{31}$

$\begin{pmatrix}1&2&3\\3&1&2\end{pmatrix}$ $\begin{pmatrix}&&\square\\\square&&\\&\square&\end{pmatrix}$ $+\ a_{13}a_{21}a_{32}$

$\overline{\qquad\qquad\qquad}$

det

So far have seen that if $\det : M_{n\times n}(F) \to F$ multilin, alternating in rows with $\det I_n = 1$ exists, then

$$\det A = \sum_{\sigma\in\Sigma_n} \operatorname{sgn}(\sigma) A_{1\sigma(1)} \cdots A_{n\sigma(n)}$$

so existence will imply uniqueness.

<u>Notation</u> For $A \in M_{n\times n}(F)$ and $1 \le i,j \le n$, let $A(i|j)$ be the matrix obtained by deleting the i-th row and j-th column from A.

<u>Lemma</u> Suppose that $n > 1$ and that $D : M_{n-1\times n-1}(F) \to F$ is multilin, alt with $D(I_{n-1}) = 1$. Fix $j \in \{1,\dots,n\}$ and define $d_j : M_{n\times n}(F) \to F$ by $d_j(A) = \sum_{i=1}^{n} (-1)^{i+j} A_{ij} D(A(i|j))$. Then d_j is alt multilin with $d_j(I_n) = 1$.

<u>Pf</u> Direct computation. $\square$

<u>Thm</u> For every $n \geq 1$ $\exists!$ det on $M_{n\times n}(F)$. Moreover, this function satisfies $\det A = \sum_{i=1}^{n} (-1)^{i+j} A_{ij} \det A(i|j)$ for every $j \in \{1,\dots,n\}$ and all $A \in M_{n\times n}(F)$.

<u>Pf</u> Existence follows inductively from the lemma. Uniqueness follows from permutation expansion. Since the determinant is unique, all the fns d_j are equal. $\square$

<u>Rmk</u> This is called cofactor (or Laplace) expansion.
The (i,j) <u>cofactor</u> of A is $(-1)^{i+j} \det A(i|j) =: C_{ij}$.
We get $\det A = \sum_{i=1}^{n} A_{ij} C_{ij} = \sum_{j=1}^{n} A_{ij} C_{ij}$
(use $\det A = \det A^T$.)

e.g. $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}$

Expand along 2nd row:

$$\det A = -2 \det\begin{pmatrix} 2 & 3 \\ 1 & 1 \end{pmatrix} + 0 \cdot \det\begin{pmatrix} 1 & 3 \\ 1 & 1 \end{pmatrix} - 1 \cdot \det\begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}$$

$$= (-2)(-1) - (-1) = 3$$

Along 3rd column:

$$\det A = 3 \det\begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix} - 1 \det\begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix} + 1 \cdot \det\begin{pmatrix} 1 & 2 \\ 2 & 0 \end{pmatrix}$$

$$= 3(2) - 1(-1) + 1(-4) = 3.$$

Rmk If a matrix has many 0's along a row or col, expand along it for quick comp'n:

$$\det\begin{pmatrix} 1 & 3 & 0 \\ 3 & 2 & 3 \\ 1 & 4 & 0 \end{pmatrix} = -3 \det\begin{pmatrix} 1 & 3 \\ 1 & 4 \end{pmatrix} = -3.$$

Defn For $A \in M_{n\times n}(F)$ let $C \in M_{n\times n}(F)$ have $C_{ij} = (-1)^{i+j} \det A(i|j)$ (the matrix of cofactors of A). The adjugate of A is

$$\operatorname{adj}(A) := C^T.$$

Thm $\operatorname{adj}(A) \cdot A = (\det A) I_n$.

Pf Let $B = \operatorname{adj}(A) A$. Then $B_{ii} = \sum_{k=1}^{n} \operatorname{adj}(A)_{ik} A_{kj}$

= expansion of $\det A$ along i-th col

$= \det A$.

For $i \neq j$, remains to show $B_{ij} = 0$. Let M be the matrix obtained by replacing the i-th col of A with A's j-th col. Show $\det M = B_{ij}$, so $B_{ij} = 0$. $\square$

Cor If A is invertible, $A^{-1} = (\det A)^{-1} \operatorname{adj}(A)$. $\square$

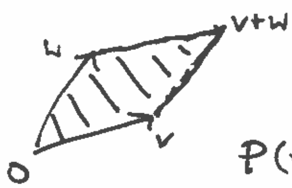
MATH 201: LINEAR ALGEBRA
DETERMINANTS OVER $\mathbb{R}$

Let $v = (x_1, y_1), w = (x_2, y_2) \in \mathbb{R}^2$ be linearly independent vectors. They span the parellelogram

$$P(v, w) = \{av + bw \mid 0 \le a, b \le 1\}.$$

Problem 1. Let M be the matrix with columns v, w so that $M = \begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix}$. Show that $M([0,1]^2) = P(v, w)$ and draw a picture of $P(v, w)$. (Here $[0,1]^2 = [0,1] \times [0,1] = \{(a, b) \mid 0 \le a, b \le 1\}$.)

$M\begin{pmatrix} a \\ b \end{pmatrix} = av + bw$ so $M([0,1]^2) = P(v,w)$.



$P(v,w)$

If the vectors v, w are linearly dependent, it is reasonable to say that the degenerate parallelogram $P(v, w)$ has area 0. This defines a function

$$A : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$$

given by $A(v, w) = \text{area}(P(v, w))$.

Problem 2. Let $k \in \mathbb{R}$. What is $A(kv, w)$? (Be careful with the case $k < 0$.)

For $k \geq 0$, $A(kv,w) = k\,A(v,w)$. In general,

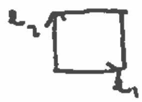
$$A(kv,w) = |k|\,A(v,w).$$

(Multiplying v by k scales the base of the parallelogram by $|k|$.)

Problem 3. Let $k \in \mathbb{R}$. What is $A(v, w + kv)$? (A proof by picture might be appropriate.)

Rotate so that v is on the horizontal axis. Then $P(v,w)$ and $P(v,w+kv)$ have the same base and height. Thus $A(v,w+kv) = A(v,w)$.

Problem 4. What is $A(e_1, e_2)$?



$A(e_1, e_2) = 1.$

Problem 5. The function A nearly has the properties of a determinant function. Explain what proprties it does and does not have in this respect.

The function A is alternating and normalized, but not quite multilinear as scalars pull out as their absolute value.

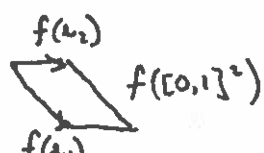
This inspires us to define the *signed area* of $P(v, w)$. For this definition, the order of v and w matters. If v and w are linearly independent, let θ be the angle from v to w, measured counterclockwise. Then $0 < \theta < 2\pi$ and $\theta \neq \pi$. We can then define

$$SA(v, w) = \begin{cases} A(v, w) & \text{if } 0 < \theta < \pi, \\ -A(v, w) & \text{if } \pi < \theta < 2\pi, \\ 0 & \text{if } v, w \text{ linearly dependent.} \end{cases}$$

Problem 6. Prove that $SA(v, w) = \det M$ where M has columns v, w.

SA is alternating, multilinear, and normalized as a function of the columns of a 2×2 matrix. By our $\det M^T = \det M$ theorem, this is equivalent to $SA(v, w) = \det \begin{pmatrix} \vdots & \vdots \\ v & w \\ \vdots & \vdots \end{pmatrix}$.

Problem 7. Consider the linear transformation $f : \mathbb{R}^2 \to \mathbb{R}^2$ given by the matrix $\begin{pmatrix} 1 & 1 \\ -2 & 0 \end{pmatrix}$. Draw a picture of $f([0,1]^2)$. What is its area?



$\det \begin{pmatrix} 1 & 1 \\ -2 & 0 \end{pmatrix} = 0 - (-2) = 2 = \text{area}(f([0,1]^2)).$

Now let $(v_1, \ldots, v_n)$ be an n-tuple of vectors in $\mathbb{R}^n$ (*i.e.*, an element of $(\mathbb{R}^n)^n$). The *parallelepiped* formed by $(v_1, \ldots, v_n)$ is the set

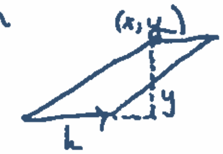
$$P(v_1, \ldots, v_n) = \{t_1 v_1 + \cdots + t_n v_n \mid t_1, \ldots, t_n \in [0, 1]\}.$$

When $n = 2$, this gives the parallelogram $P(v_1, v_2)$. For $n = 3$, we get a solid prism as long as the vectors are linearly independent.

We define the *volume* of a parallelepiped determined by $(v_1, \ldots, v_n)$ as the absolute value of the determinant of the $n \times n$ matrix with columns $v_1, \ldots, v_n$.

Problem 8. Using the properties of the determinant and your intution about how a volume should behave, argue why this definition makes sense. Check it against standard formulas for area and volume when $n = 2$ and $n = 3$.

Here is the $n=2$ check: If $v=(k,0)$, $w=(x,y)$, then $\det\begin{pmatrix} k & x \\ 0 & y \end{pmatrix} = ky - x\cdot 0 = ky$. Geometrically, $P(v,w)$ is



with area xy. For the general case, consider the rotation $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ that takes v to the positive x-axis.

Rotations don't change area, and $\det\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = \cos^2\theta + \sin^2\theta = 1$

so area is preserved.

Problem 9. For $n \times n$ real matrices A, B, interpret the rule $\det(AB) = \det(A)\det(B)$ in terms of volumes.

Since absolute value commutes with products as well, we see that the volume of $(AB)([0,1]^n)$ is the product of the volume of $A[0,1]^n$ and $B[0,1]^n$.

Let V be a fin dim vector space over a field F.

Defn A linear operator on V (or endomorphism of V) is a linear transformation $V \to V$.

Notation · $\mathcal{L}(V) := \mathcal{L}(V,V)$.

· If $f \in \mathcal{L}(V)$ and α is an ordered basis of V, $M_\alpha(f) = M_\alpha^\alpha(f)$ denotes the matrix of f wrt α.

Goal Given $f \in \mathcal{L}(V)$ find a basis α for V s.t. $M_\alpha(f)$ is especially simple.

Suppose, for instance, that $\alpha = \{v_1, \dots, v_n\}$ is a basis for V s.t. $M_\alpha(f) = \mathrm{diag}(c_1, \dots, c_n)$. Then (HW):

- A basis for $\mathrm{im}(f)$ is $\{v_i \mid c_i \neq 0\}$ and $\mathrm{rank}(f) = |\{i \mid c_i \neq 0\}|$
- A basis for $\ker(f)$ is $\{v_i \mid c_i = 0\}$ and $\mathrm{null}(f) = |\{i \mid c_i = 0\}|$.
- $\det(f) = c_1 \cdots c_n$.

We'll address the following:

- Which linear operators on V can be represented by a diagonal matrix?
- If not diagonal, what is the simplest type of matrix by which we can represent a given operator?

Defn A scalar $\lambda \in F$ is an eigenvalue of f if $\exists$ nonzero $v \in V$ s.t. $f(v) = \lambda v$. In that case, v is an eigenvector of f with eigenvalue λ.

e.g. $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2 \in \mathcal{L}(\mathbb{R}^2)$ Then $f(2,3) = (4,6) = 2 \cdot (2,3)$,
$(x,y) \longmapsto (2y - x, 6y - 6x)$

so $v_1 = (2,3)$ is an eigenvector of f with eigenvalue 2. Similarly, $f(1,2) = (3,6) = 3 \cdot (1,2)$, so $v_2 = (1,2)$ is an eigenvector of f with eigenvalue 3.

Consider $\alpha = \{v_1, v_2\}$. Then $M_\alpha(f) = \text{diag}(2,3) = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$.

Prop Let $f \in \mathcal{L}(V)$ and suppose $\alpha = \{v_1, \dots, v_n\}$ is a basis for V consisting of eigenvectors of f. Then $M_\alpha(f) = \text{diag}(\lambda_1, \dots, \lambda_n)$ where λ_i = eigenvalue of v_i for f.

Pf Follows from defn of $M_\alpha(f)$. $\square$

Defn A linear operator $f \in \mathcal{L}(V)$ is diagonalizable if $\exists$ basis of V consisting of eigenvectors of f.

e.g. $f \in \mathcal{L}(\mathbb{R}^2)$ given by $f(x,y) = (-y, x)$. Claim f has no eigenvector. If (a,b) is an eigenvector of f w/ eigenvalue λ, then $(-b, a) = f(a,b) = \lambda(a,b)$ so $-b = \lambda a$, $a = \lambda b$

$$\Longrightarrow \quad b(\lambda^2+1) = a(\lambda^2+1) = 0.$$

Since $(a,b) \neq (0,0)$, get $\lambda^2+1 = 0$ ⚡ for $\lambda \in \mathbb{R}$.

Defn Let $f \in \mathcal{L}(V)$. The characteristic polynomial of f is the polynomial $p_f(x) \in F[x]$ given by $p_f(x) = \det(A - xI)$ where $A = M_\alpha(f)$ for any ordered basis α of V.

Prop Suppose α, β ordered bases of V, let $A = M_\alpha(f)$, $B = M_\beta(f)$. Then $\det(A - xI) = \det(B - xI)$.

Pf $\exists$ invertible P s.t. $A = P^{-1}BP$. Thus $P^{-1}(B - xI)P = P^{-1}BP - P^{-1}xIP = A - xI$ so $B - xI$, $A - xI$ are similar. Finally,

$$\begin{aligned} \det(A - xI) &= \det(P^{-1}(B - xI)P) \\ &= \det P^{-1} \det(B - xI) \det P \\ &= \det(B - xI) \qquad \left(\text{b/c } \det P^{-1} = \frac{1}{\det P}\right). \; \square \end{aligned}$$

e.g. For $f(x,y) = (-y, x)$, $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ and

$$p_f(x) = \det\left(\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} - \begin{pmatrix} x & 0 \\ 0 & x \end{pmatrix}\right) = \det\begin{pmatrix} -x & -1 \\ 1 & -x \end{pmatrix} = x^2+1.$$

Lemma Let $\varphi \in \mathcal{L}(V)$. Then φ is invertible iff $\ker\varphi = \{0\}$.

Pf Rank-nullity. $\square$

Prop Let $f \in \mathcal{L}(V)$, $\lambda \in F$. Then

$$\lambda \text{ is an eigenvalue of } f \iff \lambda \text{ is a root of } p_f(x).$$

Pf Let $\varphi = f - \lambda I \in \mathcal{L}(V)$. Then λ is an eigenvalue of f iff $\ker\varphi \neq \{0\}$ iff φ not invertible iff $p_f(\lambda) = 0$. $\square$

e.g. Let $F = \mathbb{Z}/11\mathbb{Z}$, $f \in \mathcal{L}(F^2)$ given by $A = \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}$.

Then $p_f(x) = \det(A - xI) = \det\begin{pmatrix} 2-x & 1 \\ 1 & 3-x \end{pmatrix} = x^2 - 5x + 5 = (x-6)(x-10)$

so the eigenvalues of f are 6 and 10.

V fin dim vs $/F$, $f \in \mathcal{L}(V)$. When is f diagonalizable?

Defn For $U_1, \dots, U_k$ subspaces of V, say V is the direct sum of $U_1, \dots, U_k$ written $V = U_1 \oplus \cdots \oplus U_k$ if $\forall v \in V$, $\exists! u_i \in U_i$ s.t. $v = u_1 + \cdots + u_k$.

Prop Suppose $V = U_1 \oplus \cdots \oplus U_k$ and let B_i be a basis for U_i. Then

(a) $B_i \cap B_j = \emptyset$ for $i \neq j$,

(b) $B = B_1 \cup \cdots \cup B_k$ is a basis for V,

(c) $\dim V = \dim U_1 + \cdots + \dim U_k$.

Pf HW. $\square$

Lemma Let $\lambda_1, \dots, \lambda_k$ be distinct eigenvalues of $f \in \mathcal{L}(V)$, and suppose u_i is a λ_i eigenvector of f. Then $u_1, \dots, u_k$ are lin ind.

Pf by induction on k. If $k=1$, $u_1 \neq 0$ so $\{u_1\}$ lin ind. Assume now $k>1$ and the result holds for $k-1$. If $c_1u_1 + \cdots + c_ku_k = 0$ for $c_i \in F$, apply $f - \lambda I_k$ to both sides:

$$0 = \sum_{i=1}^{k} (c_i\lambda_i u_i - \lambda_k c_i u_i) = \sum_{i=1}^{k-1} c_i\lambda_i u_i - \lambda_k c_i u_i$$

$$= \sum_{i=1}^{k-1} (c_i\lambda_i - \lambda_k c_i) u_i .$$

By ind'n hypothesis, $c_i\lambda_i - \lambda_k c_i = 0$ for $i<k$.

$$\Rightarrow c_i(\lambda_i - \lambda_k) = 0$$

Since $\lambda_i \neq \lambda_k$, get $c_i = 0$ for $i<k$ so

$$0 = c_1u_1 + \cdots + c_ku_k = c_ku_k$$

$$\Rightarrow c_k = 0 \text{ too.} \quad \square$$

Defn For $\lambda \in F$, the λ-eigenspace of f is the subspace

$$E_\lambda(f) = \ker(f - \lambda I)$$
$$= \{\lambda\text{-eigenvectors of } f\} \cup \{0\}.$$

<u>Lemma</u> Let $\lambda_1, \ldots, \lambda_k$ be distinct eigenvalues of f, and let $U = E_{\lambda_1}(f) + \cdots + E_{\lambda_k}(f)$. Then $U = E_{\lambda_1}(f) \oplus \cdots \oplus E_{\lambda_k}(f)$.

<u>Pf</u> Suffices to show $u_1 + \cdots + u_k = 0$, $u_i \in E_{\lambda_i}(f) \Rightarrow u_i = 0 \ \forall i$.
Let $\{u_{i_1}, \ldots, u_{i_m}\}$ be nonzero u_i's. Then

$u_{i_1} + \cdots + u_{i_m} = 0$ is a trivial lin combo of

lin ind vectors $\lightning$. $\square$

<u>Thm</u> Let $\lambda_1, \ldots, \lambda_k$ be the distinct eigenvalues of f, and let $d_i = \dim E_{\lambda_i}(f)$. TFAE:

(a) f is diagonalizable

(b) $V = E_{\lambda_1}(f) \oplus \cdots \oplus E_{\lambda_k}(f)$

(c) $\dim V = d_1 + \cdots + d_k$.

<u>Pf</u> Let $U_i = E_{\lambda_i}(f)$, $U = U_1 \oplus \cdots \oplus U_k$.

(a) $\Rightarrow$ (b): If f is diagonalizable then every elt of V is a linear combo of eigenvectors of f. Since every eigenvector is in U_i for some i, get $V = U$.

(b) $\Rightarrow$ (c): ✓

(c) $\Rightarrow$ (a): For $i = 1, \ldots, k$, let B_i be a basis of U_i. Then $B = B_1 \cup \cdots \cup B_k$ is a basis of U with $d_1 + \cdots + d_k$ elts so $\dim V = \dim U$. As $U \subseteq V$, get $U = V$. So B is a basis of V, and every elt of B is an eigenvector of f. $\square$

How do we determine $d_1, \ldots, d_k$? Choosing a basis of V gives $V \xrightarrow{\cong} F^n$, $\mathcal{L}(V) \xrightarrow{\cong} M_{n \times n}(F)$

$$f \longmapsto A$$

and $\ker(f - \lambda I) \longleftrightarrow \ker(A - \lambda I)$. So d_i can be computed by reducing $A - \lambda I$, counting non-pivot columns e.g. $f \in \mathcal{L}(\mathbb{R}^2)$, $f(x,y) = (-x+2y, -6x+6y)$

$$\updownarrow e_1, e_2$$

$$A = \begin{pmatrix} -1 & 2 \\ -6 & 6 \end{pmatrix}$$

$$p_f(x) = \det(A - xI) = \det\begin{pmatrix} -1-x & 2 \\ -6 & 6-x \end{pmatrix} = x^2 - 5x + 6 = (x-2)(x-3)$$

$\Longrightarrow$ eigenvalues 2, 3.

$$A - 2I \rightsquigarrow \begin{pmatrix} 1 & -2/3 \\ 0 & 0 \end{pmatrix} \qquad A - 3I \rightsquigarrow \begin{pmatrix} 1 & -1/2 \\ 0 & 0 \end{pmatrix}$$

so $\ker(A-2I)$, $\ker(A-3I)$ are both 1-dim'l, spanned by $(2,3)$, $(1,2)$, resp. Since $1+1=2=\dim \mathbb{R}^2$, f is diagonalizable.

$$M_{\langle (2,3),(1,2) \rangle}(f) = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}.$$

Recall the lin trans $\mathbb{R}^2 \to \mathbb{R}^2$ given by $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ which rotates the plane by $\pi/2$. We have

$$p_A(x) = \det(A - xI) = \det\begin{pmatrix} -x & -1 \\ 1 & -x \end{pmatrix} = x^2 + 1$$

which has no roots in $\mathbb{R}$ and thus A has no eigenvalues.

Now consider the lin trans $\mathbb{C}^2 \to \mathbb{C}^2$ given by A.

Over $\mathbb{C}$, $p_A(x) = (x+i)(x-i)$ and A has eigenvalues $\pm i$.

Prop If $\dim V = n$ and $f \in \mathcal{L}(V)$ has n distinct eigenvalues, then f is diagonalizable.

Pf TPS (know eigenvectors of distinct eigenvalues are lin ind) $\square$

Compute a basis for eigenspace of i:

$$A - iI_2 = \begin{pmatrix} -i & -1 \\ 1 & -i \end{pmatrix} \xrightarrow{r_1 \leftrightarrow r_2} \begin{pmatrix} 1 & -i \\ -i & -1 \end{pmatrix} \xrightarrow{r_2 \to r_2 + ir_1} \begin{pmatrix} 1 & -i \\ 0 & 0 \end{pmatrix}$$

$\Rightarrow \ker(A - iI_2) = \{(iy, y) \mid y \in \mathbb{C}\} = \operatorname{span}\{(i, 1)\}$.

Similarly, $\ker(A + iI_2) = \operatorname{span}\{(-i, 1)\}$.

Check $A\begin{pmatrix} i \\ 1 \end{pmatrix} = i\begin{pmatrix} i \\ 1 \end{pmatrix}$ $\qquad A\begin{pmatrix} -i \\ 1 \end{pmatrix} = -i\begin{pmatrix} i \\ 1 \end{pmatrix}$.

So for $P = \begin{pmatrix} i & -i \\ 1 & 1 \end{pmatrix}$, $P^{-1}AP = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$.

Defn A polynomial $p \in F[x]$ splits over F if $\exists\, c, \lambda_1, \ldots, \lambda_n \in F$ s.t. $p(x) = c(x - \lambda_1)\cdots(x - \lambda_n)$.

Note The λ_i need not be distinct. The number of times a particular value λ occurs among the λ_i is called its algebraic multiplicity.

Fundamental Thm of Algebra Every $p \in \mathbb{C}[x]$ splits over $\mathbb{C}$.

Recall $E_\lambda(f) = \ker(f - \lambda I)$ is the λ-eigenspace of f.

Call $\dim E_\lambda(f)$ the <u>geometric</u> <u>multiplicity</u> of λ.

Call multiplicity of λ as a root of $p_f(x)$ the <u>algebraic</u> <u>multiplicity</u> of λ.

We have already seen that f is diagonalizable iff geometric multiplicities add to $\dim V$.

<u>Prop</u> For $\dim V < \infty$, λ an eigenvalue of $f \in \mathcal{L}(V)$, the geometric multiplicity of λ is $\leq$ alg mult of λ.

<u>Pf</u> Take $v_1, \ldots, v_k$ basis of $E_\lambda(f)$ and extend to $v_1, \ldots, v_n$ basis of V. With respect to this basis, the matrix for f takes the form

$$A = \begin{pmatrix} \lambda I_k & B \\ 0 & C \end{pmatrix} \quad \text{where } B, C \in M_{(n-k)\times(n-k)}(F).$$

Thus

$$p_f(x) = \det\begin{pmatrix} (\lambda - x) I_k & B \\ 0 & C - xI_{n-k} \end{pmatrix}$$

$$= (\lambda - x)^k \det(C - xI_{n-k})$$

$$= (\lambda - x)^k g(x)$$

for some $g \in F[x]$. Thus $k \leq$ alg mult of λ. $\square$

<u>Jordan Form</u>

Suppose $p_f(x)$ splits over F, but f is not diagonalizable.

Does $\exists$ basis α s.t. $M_\alpha(f)$ is still "nice"?

<u>Defn</u> A Jordan <u>block</u> of size k for $\lambda \in F$ is the $k \times k$ matrix $J_k(\lambda)$ w/ λ's on diagonal and 1's on the "superdiagonal"

$$J_k(\lambda) = \begin{pmatrix} \lambda & 1 & & 0 \\ & \lambda & 1 & \\ & & \ddots & \ddots \\ & & & \lambda & 1 \\ 0 & & & & \lambda \end{pmatrix}$$

A matrix is in Jordan form if it is block diagonal with Jordan blocks for various λ along diagonal:

$$\begin{pmatrix} J_{k_1}(\lambda_1) & & & 0 \\ & J_{k_2}(\lambda_2) & & \\ & & \ddots & \\ 0 & & & J_{k_m}(\lambda_m) \end{pmatrix}$$

e.g.

$$\left(\begin{array}{cc|cc|c|ccc} 2 & 1 & & & & & & \\ & 2 & & & & & & \\ \hline & & 2 & 1 & & & & \\ & & & 2 & & & & \\ \hline & & & & 5 & & & \\ \hline & & & & & 4 & 1 & \\ & & & & & & 4 & 1 \\ & & & & & & & 4 \end{array}\right) = \begin{pmatrix} J_2(2) & & & \\ & J_2(2) & & \\ & & J_1(5) & \\ & & & J_3(4) \end{pmatrix}$$

$$=: J_2(2) \oplus J_2(2) \oplus J_1(5) \oplus J_3(4)$$

Thm Let $\dim V < \infty$. Suppose $f \in \mathcal{L}(V)$ and $p_f(x)$ splits over F. Then $\exists$ ord basis α for V s.t. $M_\alpha(f)$ is in Jordan form. The Jordan form is unique up to permutation of the Jordan blocks.

Pf Math 332 via structure thm for fin gen modules over a principal ideal domain. $\square$

TPS · Determine $J_k(\lambda)^m$

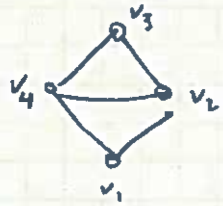
· Use the Jordan form thm to show that every $f \in \mathcal{L}(V)$ with p_f split $/F$ has a matrix rep which is the sum $D + N$ of a diagonal matrix D and nilpotent matrix N (so that $N^r = 0$ for some $r \in \mathbb{N}$).

Walks on Graphs

For $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, $D^l = \mathrm{diag}(\lambda_1^l, \ldots, \lambda_n^l)$.

So if $A = PDP^{-1}$, then (HW) $A^l = PD^lP^{-1}$ is easily computed.

A <u>graph</u> consists of vertices connected by edges:



4 vxs
5 edges

A <u>walk</u> of <u>length</u> l in a graph is a sequence of vxs $u_0, u_1, \ldots, u_l$ in the graph with u_{i-1} connected to u_i by an edge for $i = 1, \ldots, l$.

<u>e.g.</u> v_1v_4, $v_1v_2v_3v_4$ are walks from v_1 to v_4 of length 1 and length 3 above.

<u>Defn</u> Let G be a graph with vxs $v_1, \ldots, v_n$. The <u>adjacency matrix</u> of G is the $n \times n$ matrix $A = A(G)$ defined by

$$A_{ij} = \begin{cases} 1 & \text{if there is an edge connecting } v_i \text{ to } v_j \\ 0 & \text{o/w.} \end{cases}$$

<u>e.g.</u> $A = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}$ for the diamond graph.

<u>Thm.</u> The number of walks of length l from v_i to v_j is $(A^l)_{ij}$.

<u>Pf</u> HW. $\square$

<u>e.g.</u> For $A = A(\text{diamond graph})$, $A^2 = \begin{pmatrix} 2 & 1 & 2 & 1 \\ 1 & 3 & \boxed{1} & 2 \\ 2 & 1 & 2 & 1 \\ 1 & 2 & 1 & 3 \end{pmatrix}$, $A^3 = \begin{pmatrix} 2 & 5 & 2 & 5 \\ 5 & \boxed{4} & 5 & 5 \\ 2 & 5 & 2 & 5 \\ 5 & 5 & 5 & 4 \end{pmatrix}$

so, e.g., 1 path of length 2 from v_2 to v_3

4 paths of length 3 from v_2 to itself.

<u>TPS</u> Verify

Thm If $A \in M_{n\times n}(\mathbb{R})$ is symmetric ($A = A^T$), then A is diagonalizable over $\mathbb{R}$.

Pf Math 202 via spectral thm. $\square$

Note $A(G)$ is symmetric!

Thus $\exists$ diagonal D s.t. $P^{-1}AP = D$ for some P and $A^l = PD^lP^{-1}$.

It follows that the # of walks of length l b/w v_i, v_j is a linear combination of the l-th powers of the eigenvalues of A:

$$c_1\lambda_1^l + \cdots + c_n\lambda_n^l .$$

Defn A walk is <u>closed</u> if it begins and ends at the same vx.

Defn For $A \in M_{n\times n}(F)$, the <u>trace</u> of A is the sum of its diagonal entries, $tr(A) = \sum_{i=1}^{n} A_{ii}$.

Prop For $A = A(G)$, the number of closed walks in G of length l is $tr(A^l)$.

Pf The # of closed walks of length l from v_i to v_i is $(A^l)_{ii}$. Summing over $i = 1, \dots, n$ gives $tr(A^l)$. $\square$

Prop For $A \in M_{n\times n}(F)$ with $p_A(x) = c(x-\lambda_1)\cdots(x-\lambda_n)$,
$tr(A) = \lambda_1 + \cdots + \lambda_n$.

Note True even if the λ_i are in some larger field $K \supseteq F$ (and such a field always exists s.t. p_A splits).

Pf Take P s.t. $P^{-1}AP = J$ is in Jordan form. Then eigenvalues of A are on the diagonal of J, each appearing a # of times equal to its algebraic multiplicity. **Fact** $tr(UV) = tr(VU)$.
Thus $tr(A) = tr(PJP^{-1}) = tr(JPP^{-1}) = tr(J) = \lambda_1 + \cdots + \lambda_n$. $\square$

Cor For $A = A(G) \in M_{n\times n}(\mathbb{R})$, $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ the eigenvalues of A (with multiplicity), Then the # closed walks in G of length l is $\sum_{i=1}^{n} \lambda_i^l$.

Pf. $tr(A^\ell) = \sum$ eigenvalues of $A^\ell = \lambda_1^\ell + \cdots + \lambda_n^\ell$. $\square$

e.g. For the diamond graph,

$$\det(A - xI_4) = x^4 - 5x^2 - 4x = x(x+1)(x^2-x-4)$$

so eigenvalues are $0, -1, \frac{1 \pm \sqrt{17}}{2}$.

Thus the # closed walks in G of length ℓ is

$$w(\ell) = (-1)^\ell + \left(\frac{1+\sqrt{17}}{2}\right)^\ell + \left(\frac{1-\sqrt{17}}{2}\right)^\ell.$$

ℓ	1	2	3	4	5	6
$w(\ell)$	0	10	12	50	100	298

.

Suppose $x_1(t)$ = pop'n of frogs in a pond
$x_2(t)$ = pop'n of flies in a pond

and suppose

$$(*)\quad \begin{aligned} x_1'(t) &= ax_1(t) + bx_2(t) \\ x_2'(t) &= cx_1(t) + dx_2(t) \end{aligned}$$

Let $x(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}$ and $x'(t) = \begin{pmatrix} x_1'(t) \\ x_2'(t) \end{pmatrix}$

Then $(*) \iff (**)\quad x'(t) = \begin{pmatrix} a & b \\ c & d \end{pmatrix} x(t)$

<u>Goal</u> Find $x(t)$ solving $(**)$.

<u>e.g.</u> If $b=c=0$, get $x_1'(t) = ax_1(t)$
$x_2'(t) = dx_1(t)$

so $x_1(t) = k_1 e^{at}$, $x_2(t) = k_2 e^{dt}$, $k_1 = x_1(0)$, $k_2 = x_2(0)$.

Say the system is <u>decoupled</u> since x_1, x_2 don't depend on each other.

Generalize by setting $x(t) = \begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix}$, $x' = Ax$ for some $A \in M_{n\times n}(\mathbb{R})$. If A is diagonalizable over $\mathbb{R}$, then decouple the system as follows: take P s.t.

$$P^{-1}AP = D = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$$

The $x' = Ax$ becomes $x' = PDP^{-1}x$

$$\iff P^{-1}x' = DP^{-1}x$$

Set $y(t) = P^{-1}x(t)$. Then $y'(t) = P^{-1}x'(t)$ and we get the system $y' = Dy$, i.e.

$$\begin{aligned} y_1' &= \lambda_1 y_1 \\ &\vdots \\ y_n' &= \lambda_n y_n \end{aligned}$$

Solutions $y_i(t) = k_i e^{\lambda_i t}$ for $k_i = y_i(0)$, $i = 1, \dots, n$.
Since $x = Py$, this solves the original system with a linear combination of $k_1 e^{\lambda_1 t}, \dots, k_n e^{\lambda_n t}$.

e.g. $x_1' = x_2$ i.e. $x' = Ax$ for $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.
$x_2' = x_1$

Then $P^{-1}AP = D = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ for $P = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$.

Thus $y_1 = k_1 e^t$, $y_2 = k_2 e^{-t}$ and $x = Py$ is

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} k_1 e^t + k_2 e^{-t} \\ k_1 e^t - k_2 e^{-t} \end{pmatrix}$$

Suppose the soln begins at $(1,0)$. Then

$$1 = x_1(0) = k_1 e^0 + k_2 e^0 = k_1 + k_2$$
$$0 = x_2(0) = k_1 - k_2$$

so $k_1 = k_2 = \frac{1}{2}$ and our sol'n is

$$x_1(t) = \tfrac{1}{2}(e^t + e^{-t})$$
$$x_2(t) = \tfrac{1}{2}(e^t - e^{-t})$$

Note May think of $x' = Ax$ specifying "velocity" x' at each $x \in \mathbb{R}^n$. A sol'n is then $\gamma : \mathbb{R} \to \mathbb{R}^n$ s.t. $\gamma' = A\gamma$ i.e. a "flow" through the velocity field.

Alternate solution

Recall $e^t = \sum_{k=0}^{\infty} \frac{1}{k!} t^k$ converges $\forall t \in \mathbb{C}$. Given $A \in M_{n \times n}(\mathbb{C})$

define $e^{At} = \sum_{k=0}^{\infty} \frac{1}{k!} A^k t^k = I_n + At + \frac{1}{2}A^2t^2 + \frac{1}{6}A^3t^3 + \cdots$.

In each entry we get a power series in t that (Thm) converge for all t.

Prop For $A \in M_{n\times n}(\mathbb{R})$, the sol'n to $x' = Ax$ with initial condition $x(0) = p$ is $x = e^{At}p$.

Sketch $(e^{At})' = Ae^{At} \Rightarrow (e^{At}p)' = A(e^{At}p)$.

Computing e^{At}: If $P^{-1}AP = D = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$, then $A^k = P\,\mathrm{diag}(\lambda_1^k, \dots, \lambda_n^k)P^{-1}$. Thus

$$e^{At} = \sum_{k=0}^{\infty} \frac{1}{k!} A^k t^k = \sum_{k=0}^{\infty} \frac{1}{k!} (PD^kP^{-1})t^k$$

$$= P\left(\sum_{k=0}^{\infty} \frac{1}{k!} D^k t^k\right)P^{-1} = Pe^{Dt}P^{-1}$$

and $e^{Dt} = \mathrm{diag}(e^{\lambda_1 t}, \dots, e^{\lambda_n t})$

so $e^{At} = P\,\mathrm{diag}(e^{\lambda_1 t}, \dots, e^{\lambda_n t})P^{-1}$.

e.g. Previously, $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $P = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$, $D = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$

so $e^{At} = Pe^{Dt}P^{-1}$

$$= \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix}\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix}$$

$$= \frac{1}{2}\begin{pmatrix} e^t + e^{-t} & e^t - e^{-t} \\ e^t - e^{-t} & e^t + e^{-t} \end{pmatrix}$$

If $x(0) = (1, 0)$, then $x = e^{At}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} e^t + e^{-t} \\ e^t - e^{-t} \end{pmatrix}$

as we saw earlier!

Inner Products

Goal: Add structure to a vector space that will allow us to define length and angles.

Defn Let V be a vector space over $F = \mathbb{R}$ or $\mathbb{C}$. An inner product on V is a function

$$\langle \, , \rangle : V \times V \longrightarrow F$$
$$(x, y) \longmapsto \langle x, y \rangle$$

s.t. $\forall x, y, z \in V$, $c \in F$,

(1) $\langle x+y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$ and $\langle cx, y \rangle = c\langle x, y \rangle$

(2) $\overline{\langle x, y \rangle} = \langle y, x \rangle$

(3) $\langle x, x \rangle \in \mathbb{R}_{\geq 0}$ and $\langle x, x \rangle = 0$ iff $x = 0$.

Note $F = \mathbb{R}$: non-degenerate positive definite form

$F = \mathbb{C}$: non-degenerate Hermitian form

e.g.

- The ordinary dot product on $\mathbb{R}^n$: $V = \mathbb{R}^n$,
$$\langle (x_1, \dots, x_n), (y_1, \dots, y_n) \rangle = x \cdot y = \sum_{i=1}^n x_i y_i$$
- The ordinary inner product on $\mathbb{C}^n$: $V = \mathbb{C}^n$
$$\langle (x_1, \dots, x_n), (y_1, \dots, y_n) \rangle = x \cdot \overline{y} = \sum_{i=1}^n x_i \overline{y_i} \, .$$
- Let $V = C_{\mathbb{R}}([0,1]) = \{ f : [0,1] \to \mathbb{R} \mid f \text{ cts} \}$,
$$\langle f, g \rangle = \int_0^1 f(t) g(t) \, dt$$

TPS Check pos def.

- $V = \mathbb{R}^2$, $\langle (x_1, x_2), (y_1, y_2) \rangle = 3x_1y_1 + 2x_1y_2 + 2x_2y_1 + 4x_2y_2$

For pos def, $\langle (x_1, x_2), (x_1, x_2) \rangle = 3x_1^2 + 4x_1x_2 + 4x_2^2$

$$= 3\left(x_1^2 + \tfrac{4}{3}x_1x_2 + \tfrac{4}{3}x_2^2\right)$$
$$= 3\left(\left(x_1 + \tfrac{2}{3}x_2\right)^2 - \tfrac{4}{9}x_2^2 + \tfrac{4}{3}x_2^2\right)$$
$$= 3\left(\left(x_1 + \tfrac{2}{3}x_2\right)^2 + \tfrac{8}{9}x_2^2\right) \geq 0$$

with equality iff $x_1 = x_2 = 0$.

- $V = M_{m\times n}(F)$. For $A \in V$, define the <u>conjugate transpose</u> of A by $A^* := \bar{A}^T$ where $\bar{(\ )}$ takes the cpx conjugate of each entry of A. Define

$$\langle A, B\rangle = \operatorname{tr}(B^* A) = \sum_{i=1}^{n} (B^* A)_{ii}$$

<u>Note</u> $m=1$ gives usual inner product

Pos def: exercise.

<u>Prop</u> Let $(V, \langle,\rangle)$ be an inner product space. Then

(1) $\langle x, y+z\rangle = \langle x, y\rangle + \langle x, z\rangle$

(2) $\langle x, cy\rangle = \bar{c}\langle x, y\rangle$

(3) $\langle 0, y\rangle = 0$

(4) if $\langle x, y\rangle = \langle x, z\rangle \ \forall x \in V$ then $y = z$.

<u>Pf</u> (1): $\langle x, y+z\rangle = \overline{\langle y+z, x\rangle} = \overline{\langle y, x\rangle} + \overline{\langle z, x\rangle}$

$= \langle x, y\rangle + \langle x, z\rangle$

(2), (3): exc.

(4): $\langle x, y\rangle = \langle x, z\rangle \ \forall x \implies \langle x, y-z\rangle = 0 \ \forall x$

~~$\Rightarrow y-z=0$~~

In particular, for $x = y-z$ get $\langle y-z, y-z\rangle = 0$, so $y-z=0$. $\square$

Length, distance, components, projections, angles

Defn For $(V, \langle , \rangle)$ an inner product space, the norm (or length) of $x \in V$ is $\|x\| = \sqrt{\langle x, x\rangle} \in \mathbb{R}$. Two vectors are orthogonal (or perpendicular) if $\langle x, y\rangle = 0$. A unit vector x has $\|x\| = 1$ $\iff \langle x, x\rangle = 1$.

e.g. $(\mathbb{R}^n, \cdot)$ has $\|x\| = \sqrt{x_1^2 + \cdots + x_n^2}$

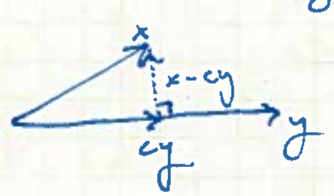
$(\mathbb{C}^n, \cdot)$ has
$$\|z\| = \sqrt{z_1\bar{z}_1 + \cdots + z_n\bar{z}_n} = \sqrt{|z_1|^2 + \cdots + |z_n|^2}$$

Note If $z_j = x_j + iy_j$, $x_j, y_j \in \mathbb{R}$, then
$$\|z\| = \sqrt{x_1^2 + y_1^2 + \cdots + x_n^2 + y_n^2}.$$

Thm (Pythagoras?) Let $(V, \langle , \rangle)$ be an inner product space and suppose $\langle x, y\rangle = 0$. Then $\|x\|^2 + \|y\|^2 = \|x+y\|^2$.

Pf We have $\langle y, x\rangle = \overline{\langle x, y\rangle} = \bar{0} = 0$ as well. Thus
$$\|x+y\|^2 = \langle x+y, x+y\rangle = \langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle = \|x\|^2 + \|y\|^2.$$
$\square$

For $x, y \in V$, $x = (x - cy) + cy$, and cy is in the "direction" of y.



Find c s.t. $\langle x - cy, y\rangle = 0$
$$\iff \langle x, y\rangle - c\langle y, y\rangle = 0$$
$$\iff c = \frac{\langle x, y\rangle}{\langle y, y\rangle} = \frac{\langle x, y\rangle}{\|y\|^2}$$
(as long as $y \neq 0$).

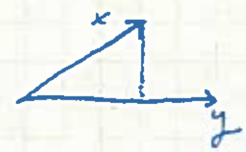
Defn The component of x along y is the scalar
$$c = \frac{\langle x, y\rangle}{\|y\|^2}.$$

The <u>orthogonal projection</u> of x along y is $cy = \dfrac{\langle x,y\rangle}{\|y\|^2}\, y$

<u>e.g.</u> $x=(3,2)$, $y=(5,0) \in \mathbb{R}^2$. Then

$$c = \frac{\langle x,y\rangle}{\|y\|^2} = \frac{(3,2)\cdot(5,0)}{(5,0)\cdot(5,0)} = \frac{15}{25} = \frac{3}{5}$$

and $cy = \frac{3}{5}(5,0) = (3,0)$



<u>Prop.</u> (1) $\|cx\| = |c|\,\|x\|$

(2) $\|x\| = 0$ iff $x = 0$

(3) $|\langle x,y\rangle| \le \|x\|\,\|y\|$ (Cauchy-Schwartz)

(4) $\|x+y\| \le \|x\| + \|y\|$ (triangle)

<u>Pf</u> (1), (2): exc.

(3): If $y=0$, done, so assume $y \ne 0$ and let $c = \dfrac{\langle x,y\rangle}{\|y\|^2}$.

Then $x - cy \perp y$ so, by Pythagoras,

$$\|x-cy\|^2 + \|cy\|^2 = \|x\|^2$$

$$\Longrightarrow \|cy\|^2 \le \|x\|^2$$

$$\Longrightarrow \|x\| \ge \|cy\| = |c|\,\|y\| = \frac{|\langle x,y\rangle|}{\|y\|}.$$

$$\Longrightarrow \|x\|\|y\| \ge |\langle x,y\rangle|.$$

(4): $\|x+y\|^2 = \langle x+y, x+y\rangle = \|x\|^2 + \langle x,y\rangle + \langle y,x\rangle + \|y\|^2$

$= \|x\|^2 + \langle x,y\rangle + \overline{\langle x,y\rangle} + \|y\|^2$

$= \|x\|^2 + 2\,\mathrm{Re}(\langle x,y\rangle) + \|y\|^2$ $\qquad (z+\bar{z} = 2\,\mathrm{Re}(z))$

$\le \|x\|^2 + 2|\langle x,y\rangle| + \|y\|^2$

$\le \|x\|^2 + 2\|x\|\,\|y\| + \|y\|^2$ $\qquad$ (CS)

$= (\|x\| + \|y\|)^2$ $\qquad \square$

<u>Defn.</u> The <u>distance</u> between $x, y \in V$ is

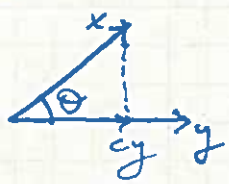
$$d(x,y) := \|x-y\|.$$

<u>Prop</u> (1) $d(x,y) = d(y,x)$

(2) $d(x,y) \geq 0$ with equality iff $x=y$

(3) $d(x,y) \leq d(x,z) + d(z,x)$. $\square$

<u>Angles</u> $(V, \langle , \rangle)$ inner product space over $F = \mathbb{R}$. (not $\mathbb{C}$)



<u>Defn.</u> The <u>angle</u> θ between $x, y \in V$ is

$$\theta = \arccos\left(\frac{\langle x,y\rangle}{\|x\|\,\|y\|}\right).$$

Thus $\langle x,y\rangle = \|x\|\,\|y\| \cos\theta$.

<u>Rmk</u> • By C-S, $|\langle x,y\rangle| \leq \|x\|\,\|y\|$, so

$$-1 \leq \frac{\langle x,y\rangle}{\|x\|\,\|y\|} \leq 1$$ and arccos makes since.

• $\cos\theta = \left\langle \frac{x}{\|x\|}, \frac{y}{\|y\|} \right\rangle$

unit vector in direction of x.

Gram-Schmidt

Let $(V, \langle , \rangle)$ be an inner product space over $F = \mathbb{R}$ or $\mathbb{C}$.

Defn Let $S \subseteq V$. Then S is an <u>orthogonal subset</u> of V if $\langle u, v\rangle = 0$ for all $u \neq v \in S$. If S is an orthogonal subset of V and $\|u\| = 1$ for all $u \in S$, then S is an <u>orthonormal subset</u> of V.

<u>e.g.</u> • The standard basis $e_1, \ldots, e_n$ for F^n is orthonormal with respect to the standard inner product on F^n.

• $\{\frac{1}{\sqrt{2}}(1,1), \frac{1}{\sqrt{2}}(1,-1)\}$ is orthonormal wrt std inner product on $\mathbb{R}^2$

Prop Let $S = \{v_1, \ldots, v_k\}$ be an orthogonal set of nonzero vectors in V, and let $y \in \operatorname{span} S$. Then

$$y = \sum_{j=1}^{k} \frac{\langle y, v_j\rangle}{\langle v_j, v_j\rangle} v_j = \sum_{j=1}^{k} \frac{\langle y, v_j\rangle}{\|v_j\|^2} v_j .$$

Pf Say $y = \sum_{i=1}^{k} a_i v_i$. Then for $j = 1, \ldots, k$

$$\langle y, v_j\rangle = \left\langle \sum_{i=1}^{k} a_i v_i, v_j \right\rangle = \sum_{i=1}^{k} a_i \langle v_i, v_j\rangle = a_j \langle v_j, v_j\rangle .$$

Hence $a_j = \dfrac{\langle y, v_j\rangle}{\langle v_j, v_j\rangle}$. $\square$

Cor Let $S \subseteq V$ be orthonormal, $S = \{v_1, \ldots, v_k\}$, $y \in \operatorname{span} S$. Then $y = \sum_{j=1}^{k} \langle y, v_j\rangle v_j$. $\square$

Cor If $S = \{v_1, \ldots, v_k\} \subseteq V$ is orthogonal, then S is linearly ind.

Pf If $\sum_{i=1}^{k} a_i v_i = 0$ then for $j = 1, \ldots, k$,

$$0 = \langle 0, v_j\rangle = \left\langle \sum_{i=1}^{k} a_i v_i, v_j \right\rangle = a_j \underbrace{\langle v_j, v_j\rangle}_{\neq 0} \Rightarrow a_j = 0 \ \forall j. \ \square$$

<u>e.g.</u> $\mathbb{R}^2$ w/ std inner prod,

$$u = \tfrac{1}{\sqrt{2}}(1,1), \quad v = \tfrac{1}{\sqrt{2}}(1,-1)$$

Then $\beta = \{u, v\}$ is an orthonormal ordered basis for $\mathbb{R}^2$.

<u>Q</u> What are the coords of $y = (4,7)$ wrt β?

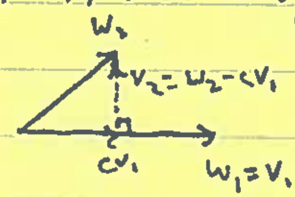
<u>A</u> (TPS)
$$\begin{aligned} y &= \langle y, u\rangle u + \langle y, v\rangle v \\ &= (4,7)\cdot \tfrac{1}{\sqrt{2}}(1,1)\, u + (4,7)\cdot \tfrac{1}{\sqrt{2}}(1,-1)\, v \\ &= \tfrac{11}{\sqrt{2}} u - \tfrac{3}{\sqrt{2}} v \end{aligned}$$

Indeed, $\frac{11}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}}(1,1) - \frac{3}{\sqrt{2}}\frac{1}{\sqrt{2}}(1,-1) = \left(\frac{11}{2}, \frac{11}{2}\right) - \left(\frac{3}{2}, -\frac{3}{2}\right) = (4,7)$ ✓

<u>Gram-Schmidt</u>

Baby case: given $w_1, w_2 \in V$, find orthogonal v_1, v_2 s.t.
$$\operatorname{span}\{w_1, w_2\} = \operatorname{span}\{v_1, v_2\}.$$

Idea: Let $v_1 = w_1$, then "straighten out" w_2 to create v_2:



Here $c = \dfrac{\langle w_2, v_1\rangle}{\|v_1\|^2}$ is the component of w_2 along v_1.

Gram-Schmidt Orthogonalization Algorithm:

Input: $S = \{w_1, \ldots, w_n\}$, a lin ind subset of V.

- Let $v_1 = w_1$.
- For $k = 2, 3, \ldots, n$, define
$$v_k = w_k - \sum_{i=1}^{k-1} \frac{\langle w_k, v_i\rangle}{\|v_i\|^2} v_i.$$

(I.e. subtract orthog proj'ns of w_k along previously found v_j.)

Output: $S' = \{v_1, \ldots, v_n\}$ an orthogonal set with $\operatorname{span} S' = \operatorname{span} S$

<u>or</u> Output: $S'' = \left\{\frac{v_1}{\|v_1\|}, \ldots, \frac{v_n}{\|v_n\|}\right\}$ an orthonormal set with $\operatorname{span} S'' = \operatorname{span} S$.

Pf by induction on n. For $n=1$, ✓. Assume it works for some $n\geq 1$. Then $\operatorname{span}\{v_1,\dots,v_n\} = \operatorname{span}\{w_1,\dots,w_n\}$ and $\{v_1,\dots,v_n\}$ orthogonal. Then

$$v_{n+1} = w_{n+1} - \sum_{i=1}^{n} \frac{\langle w_{n+1}, v_i\rangle}{\|v_i\|^2} v_i .$$

If $v_{n+1}=0$, then $w_{n+1} \in \operatorname{span}\{v_1,\dots,v_n\} = \operatorname{span}\{w_1,\dots,w_n\}$, ↯, so $v_{n+1}\neq 0$. For $j=1,\dots,n$,

$$\langle v_{n+1}, v_j\rangle = \left\langle w_{n+1} - \sum_{i=1}^{n} \frac{\langle w_{n+1}, v_i\rangle}{\|v_i\|^2} v_i, v_j\right\rangle$$

$$= \langle w_{n+1}, v_j\rangle - \sum_{i=1}^{n} \frac{\langle w_{n+1}, v_i\rangle}{\|v_i\|^2}\langle v_i, v_j\rangle$$

$$= \langle w_{n+1}, v_j\rangle - \frac{\langle w_{n+1}, v_j\rangle}{\|v_j\|^2}\langle v_j, v_j\rangle$$

$$= 0.$$

so $\{v_1,\dots,v_{n+1}\}$ orthogonal. It remains to show this set has the correct span. Since $\{v_1,\dots,v_{n+1}\}$ lin ind and

$$\operatorname{span}\{v_1,\dots,v_{n+1}\} \subseteq \operatorname{span}\{v_1,\dots,v_n, w_{n+1}\} \subseteq \operatorname{span}\{w_1,\dots,w_{n+1}\}$$

get equality (both $(n+1)$-dim'l). □

Cor Every nontrivial fin dim'l inner product space has an orthonormal basis. □

e.g. $V=\mathbb{R}_{\leq 1}[x]$, $\langle f,g\rangle = \int_0^1 f(t)g(t)\,dt$.

Apply G-S to $\{1,x\}$ to get orthonormal basis.

Note $\langle 1,x\rangle = \int_0^1 t\,dt = \frac{1}{2}\neq 0$, so not orthogonal.

G-S: $v_1 = 1$.

$$v_2 = x - \frac{\langle x, v_1\rangle}{\|v_1\|^2}v_1 = x - \frac{\langle x,1\rangle}{\|1\|^2}1 = x - \frac{1/2}{\int_0^1 dt} = x-\frac{1}{2}.$$

Have $\|v_1\| = \sqrt{\int_0^1 dt} = 1$

$$\|v_2\| = \sqrt{\langle x - 1/2, x - 1/2 \rangle}$$

$$= \sqrt{\int_0^1 (t - 1/2)^2 \, dt}$$

$$= \sqrt{1/12}$$

So an orthonormal basis for V is $\left\{1, \frac{1}{\sqrt{12}}(x - 1/2)\right\}$.

Orthogonal complements and projections

Defn The (external) direct sum of vector spaces U, W over a field F is the set $U \oplus W := U \times W$ with coordinatewise scalar mult & vector addition:

$$\lambda(u,w) = (\lambda u, \lambda w)$$
$$(u,w) + (u',w') = (u+u', w+w') \qquad u,u' \in U,\ w,w' \in W,\ \lambda \in F$$

Prop Let U, W be subspaces of a vector space V over F s.t. $U + W = V$ and $U \cap W = \{0\}$. Then

$$U \oplus W \xrightarrow{\cong} V$$
$$(u,w) \longmapsto u+w. \qquad \square$$

~~Def~~ Now let $(V, \langle\,,\rangle)$ be an inner product space over $F = \mathbb{R}$ or $\mathbb{C}$.

Defn Let $\emptyset \neq S \subseteq V$. The orthogonal complement of S is

$$S^\perp = \{x \in V \mid \langle x, y\rangle = 0 \ \forall y \in S\}.$$

Exc $S^\perp$ is a subspace of V.

Prop Suppose $\dim V = n$ and $S = \{v_1, \dots, v_k\}$ is an orthonormal subset of V. ① S can be extended to an orthonormal basis $\{v_1, \dots, v_n\}$ of V.

② If $W = \operatorname{span} S$, then $S' = \{v_{k+1}, \dots, v_n\}$ is an orthonormal basis for $W^\perp$.

③ If $W \subseteq V$ is any subspace, then

$$\dim W + \dim W^\perp = \dim V = n.$$

④ If $W \subseteq V$ is any subspace, then $(W^\perp)^\perp = W$.

Pf ① Extend to a basis then apply G-S.

② S' is lin ind as its a subset of a basis. $S' \subseteq W^\perp$ by orthogonality of $\{v_1, \dots, v_n\}$. Thus $\operatorname{span} S' \subseteq W^\perp$.

For $x \in W^\perp$, $x = \sum_{i=1}^{n} \langle x, v_i\rangle v_i = \sum_{i=k+1} \langle x, v_i\rangle v_i \in \operatorname{span} S'$.

③ Choose an orthonormal basis for W then apply ①, ②.

④ Have $(W^\perp)^\perp = \{x \in V \mid \langle x, y \rangle = 0 \ \forall y \in W^\perp\} \supseteq W$.

By ③, $\dim (W^\perp)^\perp = n - \dim W^\perp = \dim W$, so they are equal. □

Prop Let W be a finite dimensional subspace of V. Then $V = W \oplus W^\perp$. I.e. $\forall y \in V \ \exists! u \in W, z \in W^\perp$ s.t. $y = u + z$.

Defn Define u to be the orthogonal projection of y onto W.

If $u_1, \ldots, u_k$ orthonormal basis of W, then

$$u = \sum_{i=1}^{k} \langle y, u_i \rangle u_i .$$

Pf By G-S, $\exists$ orthonormal basis $u_1, \ldots, u_k$ of W. Define $u = \sum_{i=1}^{k} \langle y, u_i \rangle u_i$ and $z = y - u$. Then $u \in W$ and $y = u + z$ for $z = y - u$. Further, for $j = 1, \ldots, k$,

$$\begin{aligned}
\langle z, u_j \rangle &= \langle y - u, u_j \rangle \\
&= \langle y, u_j \rangle - \left\langle \sum_{i=1}^{k} \langle y, u_i \rangle u_i , u_j \right\rangle \\
&= \langle y, u_j \rangle - \sum_{i=1}^{k} \langle y, u_i \rangle \langle u_i, u_j \rangle \\
&= \langle y, u_j \rangle - \langle y, u_j \rangle \langle u_j, u_j \rangle \\
&= 0, \text{ so } z \in W^\perp .
\end{aligned}$$

For uniqueness, suppose $\exists u' \in W, z' \in W^\perp$ s.t. $y = u + z = u' + z'$. Then $u - u' = z - z' \in W \cap W^\perp = \{0\}$. □

Cor The orthogonal projection u of y onto W is the closest vector in W to y: $\|y - u\| \le \|y - w\| \ \forall w \in W$ with equality iff $u = w$.

Pf Write $y = u + z$ with $u \in W, z \in W^\perp$. Let $w \in W$. Then $u - w \in W$, $y - u \in W^\perp$. So $u - w$ and $z = y - u$ are orthogonal.

By Pythagoras,

$$\begin{aligned}\|y-w\|^2 &= \|(u+z)-w\|^2 \\ &= \|(u-w)+z\|^2 \\ &= \|u-w\|^2+\|z\|^2 \\ &\geq \|z\|^2 \\ &= \|y-u\|^2.\end{aligned}$$

Equality iff $\|u-w\|^2=0$, i.e. iff $u=w$. $\square$

e.g. $V=\mathbb{R}^3$ w/ std inner prod. For orthogonal proj'n onto xy-plane, take $\{e_1, e_2\}$ as orthonormal basis. The proj'n of $u=(x,y,z)\in\mathbb{R}^3$ onto this plane is

$$(u\cdot e_1)e_1+(u\cdot e_2)e_2=(x,y,0).$$

The distance of u to the xy-plane is

~~$\|u-(0,0,z$~~

$$\|u-(x,y,0)\| = \|(0,0,z)\| = |z|. \checkmark$$

$(V, \langle\,,\,\rangle)$ an inner product space over $F = \mathbb{R}$ or $\mathbb{C}$.

$W \subseteq V$ fin-dim'l subspace $\Rightarrow$ $V = W \oplus W^\perp$, so every $y \in V$ has a unique expression of the form $y = u + z$, $u \in W$, $z \in W^\perp$. If $\{u_1, \dots, u_k\}$ is an orthonormal basis for W, then

$$u = \sum_{i=1}^{k} \langle y, u_i \rangle u_i$$

and u is the point in W closest to y.

e.g. In $\mathbb{R}^2$, find the line closest to the three points $(0,6)$, $(1,0)$, $(2,0)$. For $y = ax + b$ to pass through the pts, we would need

$$\underset{A}{\begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{pmatrix}} \underset{x}{\begin{pmatrix} a \\ b \end{pmatrix}} = \underset{y}{\begin{pmatrix} 6 \\ 0 \\ 0 \end{pmatrix}}$$

There is no x satisfying this eq'n, so instead we look for $x = (a, b)$ minimizing the error $e := \|y - Ax\|$.

Define $W := \mathrm{im}(A)$. Then to minimize e, we need to compute the projection of $y = (6,0,0)$ onto W. For this, we need an orthonormal basis of W. Begin w/ columns of A & apply Gram-Schmidt: $v_1 = (0,1,2)$

$$v_2 = (1,1,1) - \frac{(1,1,1)\cdot(0,1,2)}{(0,1,2)\cdot(0,1,2)}(0,1,2)$$

$$= (1,1,1) - \tfrac{3}{5}(0,1,2)$$

$$= \left(1, \tfrac{2}{5}, -\tfrac{1}{5}\right).$$

Then $u_1 = \dfrac{v_1}{\|v_1\|} = \dfrac{1}{\sqrt{5}}(0,1,2)$

$u_2 = \dfrac{v_2}{\|v_2\|} = \sqrt{\tfrac{5}{6}}\left(1, \tfrac{2}{5}, -\tfrac{1}{5}\right)$ form an orthonormal basis of W.

The projection of y onto W is

$$u = \langle y, u_1 \rangle u_1 + \langle y, u_2 \rangle u_2$$

$$= (6,0,0)\cdot \tfrac{1}{\sqrt{5}}(0,1,2)\, u_1 + (6,0,0)\cdot \sqrt{\tfrac{5}{6}}\left(1, \tfrac{2}{5}, -\tfrac{1}{5}\right) u_2$$

$$= 6\sqrt{\tfrac{5}{6}}\, u_2$$

$$= 6\sqrt{\tfrac{5}{6}}\sqrt{\tfrac{5}{6}}\left(1, \tfrac{2}{5}, -\tfrac{1}{5}\right)$$

$$= (5,2,-1).$$

Since $(5,2,-1) \in W = \operatorname{im} A$, we can solve

$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 5 \\ 2 \\ -1 \end{pmatrix}$$ and get $a = -3$, $b = 5$.

So the line of best fit is $\boxed{y = -3x + 5.}$

<u>Adjoints</u> (w/o proof)

If $f \in \mathcal{L}(V)$, $\exists!\, f^\dagger \in \mathcal{L}(V)$, the <u>adjoint</u> of f satisfying

$$\langle f(x), y \rangle = \langle x, f^\dagger(y) \rangle$$

$\forall x, y \in V$. If f is represented by a matrix A wrt some ordered basis, then $f^\dagger$ is rep'd by $A^\dagger$ (aka A^*), the conjugate transpose of A: $\langle Ax, y \rangle = \langle x, A^\dagger y \rangle$.

<u>Utility</u>: given $A \in M_{m\times n}(F)$ and $y \in F^m$, want to compute $x \in F^n$ minimizing $\|y - Ax\|$.

<u>Lemma</u> $\operatorname{rank}(A^\dagger A) = \operatorname{rank}(A)$.

<u>Pf</u> Note $A^\dagger A \in M_{n\times n}(F)$. By rank-nullity,

$$\operatorname{rank}(A) = n - \dim(\ker A)$$

$$\operatorname{rank}(A^\dagger A) = n - \dim(\ker A^\dagger A)$$

so it suffices to show $\dim \ker A = \dim \ker A^\dagger A$.

If $x \in \ker A$, then $Ax = 0 \Rightarrow A^+Ax = A^+0 = 0$ so $x \in \ker A^+A$.
Thus $\ker A \subseteq \ker A^+A$.
If $x \in \ker A^+A$, then $A^+Ax = 0 \Rightarrow 0 = \langle x, 0\rangle = \langle x, A^+Ax\rangle = \langle Ax, Ax\rangle$.
By pos def, $Ax = 0$, so $x \in \ker A$, proving the opposite inclusion. $\square$

Cor If $A \in M_{m\times n}(F)$ has rank n, then A^+A is invertible. $\square$

Prop Given $A \in M_{m\times n}(F)$ and $y \in F^m$, there exists $x_0 \in F^n$ such that $\|y - Ax_0\| \leq \|y - Ax\|$ $\forall x \in F^n$. For this x_0, we have $A^+A x_0 = A^+y$. If $\operatorname{rank}(A) = n$, then $x_0 = (A^+A)^{-1}A^+y$.

Pf Want Ax_0 closest to y so looking for the proj'n of y onto $\operatorname{im}(A) \subseteq F^m$. This proves existence. Now want to find $x_0 \in F^n$ s.t. $y = Ax_0 + z$ with $z = y - Ax_0 \in (\operatorname{im} A)^\perp$.

$$\begin{aligned} y - Ax_0 \in (\operatorname{im} A)^\perp &\iff \langle Ax, y - Ax_0\rangle = 0 \quad \forall x \in F^n \\ &\iff \langle x, A^+(y - Ax_0)\rangle = 0 \quad \forall x \in F^n \\ &\iff A^+(y - Ax_0) = 0 \\ &\iff A^+Ax_0 = A^+y. \end{aligned}$$
$\square$

e.g. $A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{pmatrix}$, $y = \begin{pmatrix} 6 \\ 0 \\ 0 \end{pmatrix}$, rank $A = 2$.

$$A^+A = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 5 & 3 \\ 3 & 3 \end{pmatrix}, \quad (A^+A)^{-1} = \frac{1}{6}\begin{pmatrix} 3 & -3 \\ -3 & 5 \end{pmatrix}$$

$$\Rightarrow x_0 = (A^+A)^{-1}A^+y = \frac{1}{6}\begin{pmatrix} 3 & -3 \\ -3 & 5 \end{pmatrix}\begin{pmatrix} 0 & 1 & 2 \\ 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 6 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -3 \\ 5 \end{pmatrix}.$$

Rmk This method avoids computing an orthonormal basis (as before).

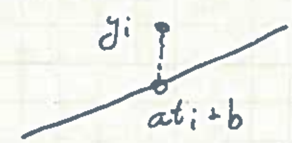
<u>Least Squares</u> Minimizing $\|y - Ax\|$ is called the method of least squares. Imagine that at time t_i we are measuring a quantity $y_i \in F$, $i = 1, \ldots, n$. Want the "best" line $y = ax + b$. Then we want $a, b \in F$ s.t. $y_i = at_i + b$ for each i, i.e.

$$\underset{A}{\begin{pmatrix} t_1 & 1 \\ \vdots & \vdots \\ t_n & 1 \end{pmatrix}} \underset{x}{\begin{pmatrix} a \\ b \end{pmatrix}} = \underset{y}{\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}}$$

Seek to minimize error $\|y - Ax\|$, or equivalently $\|y - Ax\|^2$. But

$$\|y - Ax\|^2 = \sum_{i=1}^{n} (y_i - (at_i + b))^2 .$$

The terms $y_i - (at_i + b)$ are vertical distances



and we are minimizing their squares.

Principal Component Analysis

Suppose we take n measurements of m variables (with real values). Each measurement is a vector in $\mathbb{R}^m$, and our n measurements are then n vectors $x_1, \dots, x_n \in \mathbb{R}^m$ point cloud

The **mean** (or average) of these vectors is

$$\mu := \frac{1}{n}(x_1 + \cdots + x_n).$$

Note The i-th component of μ is just the average value of the i-th variable: $\mu_i = \frac{1}{n}(x_{1i} + \cdots + x_{ni})$.

Q How can we mimic other important statistics?

e.g. For $a_1, \dots, a_n \in \mathbb{R}$, their **variance** is

$$\operatorname{var}(a) = \frac{1}{n-1}\left((a_1 - \mu)^2 + \cdots + (a_n - \mu)^2\right).$$

If we also measure $b_1, \dots, b_n \in \mathbb{R}$, the **covariance** b/w a_i, b_i is

$$\operatorname{cov}(a,b) = \frac{1}{n-1}\left((a_1 - \mu_a)(b_1 - \mu_b) + \cdots + (a_n - \mu_a)(b_n - \mu_b)\right).$$

- Variance measures how much the a_i differ from their mean and its square root is the **standard deviation**.
- Covariance measures how the a_i & b_i depend on each other. E.g. negative cov arises when large a_i predicts small b_i (relative to means).

Going back to multivariate setting, define

$$B = \begin{pmatrix} x_1 - \mu & x_2 - \mu & \cdots & x_n - \mu \end{pmatrix}$$

the $m \times n$ matrix with i-th column $x_i - \mu$. This is the **recentering** of the data so the mean is 0.

Defn The **covariance matrix** $S = \frac{1}{n-1} BB^T$.

Note $S \in M_{m \times m}(\mathbb{R})$ is symmetric.

e.g. $x_1 = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{pmatrix}$ $x_2 = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{pmatrix}$ $x_3 = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{pmatrix}$ $\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \mu_4 \end{pmatrix}$

$$B = \begin{pmatrix} a_1-\mu_1 & b_1-\mu_1 & c_1-\mu_1 \\ a_2-\mu_2 & b_2-\mu_2 & c_2-\mu_2 \\ a_3-\mu_3 & b_3-\mu_3 & c_3-\mu_3 \\ a_4-\mu_4 & b_4-\mu_4 & c_4-\mu_4 \end{pmatrix}$$

Then $S_{11} = \frac{1}{3-1}\left((a_1-\mu_1)^2 + (b_1-\mu_1)^2 + (c_1-\mu_1)^2\right)$ = variance of first variable.

Similarly, S_{ii} = variance of i-th variable.

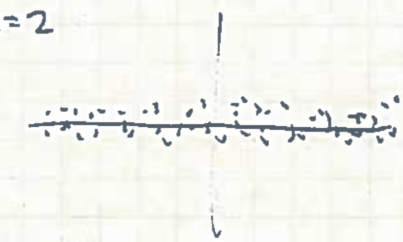
Also $S_{21} = \frac{1}{3-1}\left((a_1-\mu_1)(a_2-\mu_2) + (b_1-\mu_1)(b_2-\mu_2) + (c_1-\mu_1)(c_2-\mu_2)\right)$

= covariance of first and second variables.

Similarly, S_{ij} = cov of i-th & j-th vars.

Defn The total variance is $tr(S) = \Sigma$ var of variables.

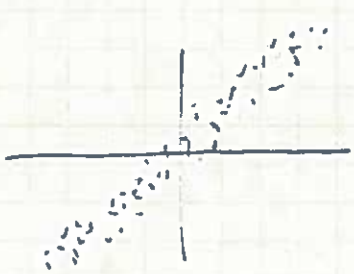
e.g. $m=2$



Observe: S_{11} large, S_{22} small

covariance small

$S = \begin{pmatrix} 95 & 1 \\ 1 & 5 \end{pmatrix}$ (total var 100)



$S = \begin{pmatrix} 50 & 40 \\ 40 & 50 \end{pmatrix}$

Goal Recognize the ~~sim~~ similarity of these data sets with linear algebra.

Spectral Thm If $A \in M_{n\times n}(\mathbb{R})$ and $A = A^T$, then A is orthogonally diagonalizable with real eigenvalues. I.e. $\exists \lambda_1, \dots, \lambda_n \in \mathbb{R}$ and orthogonal nonzero vectors $v_1, \dots, v_n \in \mathbb{R}^n$ s.t. $Av_i = \lambda_i v_i$.

Pf Later.

Note For $B \in M_{m\times n}(\mathbb{R})$, BB^T and B^TB are symmetric real matrices to which the spectral thm applies.

Prop BB^T and B^TB share the same nonzero eigenvalues.

Pf Take v an eigenvector of B^TB with eigenvalue $\lambda \neq 0$, so that

$$B^TBv = \lambda v.$$

Mult on left by B to get

$$BB^T(Bv) = \lambda(Bv).$$

Hence λ is an eigenvalue of BB^T with eigenvector Bv.

(Indeed, $Bv \neq 0$ since $B^TBv = \lambda v \neq 0$.)

Similarly $BB^Tw = \lambda w \Rightarrow B^TB(B^Tw) = \lambda(B^Tw)$

so BB^T, B^TB have the same nonzero eigenvalues. $\square$

TPS What if B is 500×2?

(Find eigenvalues of 2×2 matrix B^TB. These are eigenvalues of BB^T (a 500×500 matrix) and all others are 0.)

Prop The eigenvalues of BB^T and B^TB are all nonnegative.

Pf Take v an eigenvector of B^TB with eigenvalue λ. Then

$$\begin{aligned}\|Bv\|^2 &= (Bv)\cdot(Bv) = (Bv)^T(Bv)\\ &= v^T(B^TB)v\\ &= v^T(\lambda v)\\ &= \lambda v^Tv\\ &= \lambda\|v\|^2.\end{aligned}$$

Since $\|Bv\|^2 \geq 0$ and $\|v\|^2 \neq 0$, must have $\lambda \geq 0$.

$\square$

<u>Recall</u> n measurements of m variables recorded as vectors $x_1, \dots, x_n \in \mathbb{R}^m$ have mean

$$\mu = \frac{1}{n}(x_1 + \cdots + x_n),$$

mean-centered data matrix $B \in M_{m \times n}(\mathbb{R})$ with i-th column $x_i - \mu$ and covariance matrix

$$S = \frac{1}{n-1} BB^T \in M_{m \times m}(\mathbb{R}).$$

S is symmetric, so the spectral theorem (and corollary on matrices BB^T) imply that S has nonnegative real eigenvalues

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m \geq 0.$$

Let $u_1, \dots, u_m$ be the corresponding orthogonal eigen~~values~~vectors.

The vectors $u_1, \dots, u_m$ are the <u>principal components</u> of the data.

<u>Note</u> Total variance $T = Tr(S) = \lambda_1 + \cdots + \lambda_m$.

- The direction (unit vector) u_1 (the first principal direction) accounts for $\frac{\lambda_1}{T}$ of the total variance. The second principal direction u_2 accounts for $\frac{\lambda_2}{T}$ of the total variance. Etc.
- Thus u_1 points in the "most significant" direction of the data set.
- Amongst $u_1^\perp$, u_2 points in the most significant direction. Etc.

<u>Fact</u> The line spanned by u_1 minimizes <u>orthogonal</u> distance from line to point cloud (compare to least squares).

Suppose we are measuring 10 variables ~~and the~~ $T = 100$, $\lambda_1 = 90.5$, $\lambda_2 = 8.9$. Then $\lambda_3, \dots, \lambda_{10} \leq 0.1$ and the data set in $\mathbb{R}^{10}$ has 99.4% of its total variance explained by $\text{span}\{u_1, u_2\}$.