MATH 1553

INTRODUCTION TO LINEAR ALGEBRA.

Fall 2015, Georgia Tech

LECTURE 1.

algebra - arithmetic ($+ - \times \div$) with symbols

from al-jebr (Arabic): reunion of broken parts

eg. $x = 9 - 4x \rightsquigarrow 5x = 9$

9th c. Abu Ja'far Muhammad ibn Musa al-Khwarizmi

(algorithm)

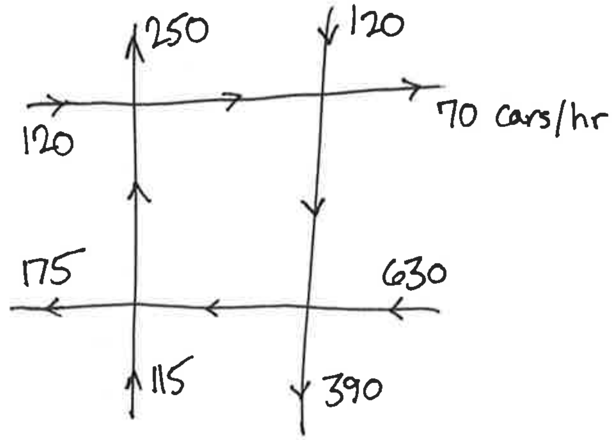
linear - having to do with lines/planes/etc.

$x + y + 3z = 7$ not: $\sin(x)$, $\ln(x)$, x^2, etc.

> Almost every engineering problem, no matter how huge, can be reduced to linear algebra
>
> $Ax = b$ or
>
> $Ax = \lambda x$

examples: load & displacement, finite element analysis, stress & strain, LCR circuits, flow in a network of pipes, computer vision, machine learning, data analysis.

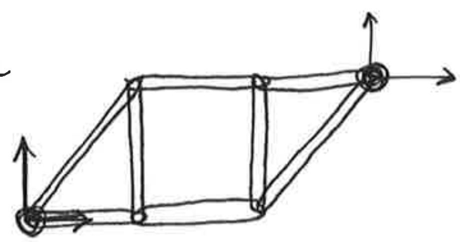
Traffic Flow



How much traffic in the four unlabelled segments?

⇝ system of linear equations

Stress & strain



Find force at each joint given the forces at the two ends.

⇝ system of lin eqns.

Chemistry

$$\_\_\ C_2H_6 + \_\_\ O_2 \longrightarrow \_\_\ CO_2 + \_\_\ H_2O$$

⇝ sys of lin eqns

Genetics

genotypes AA = brown eyes
Aa = brown
aa = blue

A = dominant gene
a = recessive.

if we only breed with AA's, what happens to the population with each generation (offspring gets one gene, randomly, from each parent)

$$AA_{n+1} = AA_n + \tfrac{1}{2} Aa_n$$
$$Aa_{n+1} = \tfrac{1}{2} Aa_n + aa_n$$

$$\begin{pmatrix} 1 & 1/2 & 0 \\ 0 & 1/2 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

Geometry Find the equation of a circle passing through 3 given pts,

say $(1,0)$ $(0,1)$ $(1,1)$.

general form: $a(x^2+y^2)+bx+cy+d=0$

$\rightsquigarrow$ sys of lin eqns

Astronomy Compute the orbit of a planet

Similar: $ax^2+bx+cy^2+dy+e=0$

Kepler's first law: orbit of an asteroid around the sun is an ellipse

Google "the 25 billion dollar eigenvector"

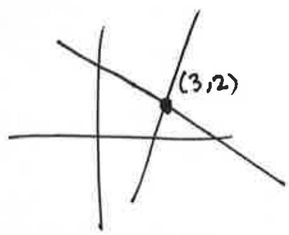
each web page has some importance, which it shares via outgoing links $\rightsquigarrow$ sys of lin eqns

LECTURE 2.

Section 1.1 Systems of Linear Equations

Solution to linear equation is a line, plane, etc.
Solution to a <u>system</u> of linear equations is an <u>intersection</u> of lines, planes, etc.

e.g. $$\begin{aligned} x - 3y &= -3 \\ 2x + y &= 8 \end{aligned}$$ $\rightsquigarrow$



Two vars $\rightsquigarrow$ possibilites are: line, pt, $\emptyset$.

$\boxed{\text{CLICKER}}$ In how many ways can 3 planes intersect in $\mathbb{R}^3$?

Let's solve:
$$\begin{aligned} x + 2y + 3z &= 6 \\ 2x - 3y + 2z &= 14 \\ 3x + y - z &= -2 \end{aligned}$$

Idea: Eliminate all but one x from first col, one y from 2^{nd} col, one z from 3^{rd}.

Tools: ① add a multiple of one eqn to another
② swap rows
③ multiply an eqn by a const.

Why do these not change the solution?

First solve the system in long-hand.
Then rewrite using matrix notation

augmented matrix

$$\begin{pmatrix} 1 & 2 & 3 & 6 \\ 2 & -3 & 2 & 14 \\ 3 & 1 & -1 & -2 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 3 \end{pmatrix}$$

$$\Rightarrow \quad \begin{aligned} x \quad\quad &= 1 \\ y \quad &= -2 \\ z &= 3 \end{aligned}$$

Try this:

$$\begin{aligned} x+2y &= 10 \\ 2x-2y &= -4 \\ 3x+5y &= 20 \end{aligned} \quad \rightsquigarrow \quad \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 4 \\ 0 & 0 & 6 \end{pmatrix}$$

$\rightsquigarrow$ $0=6$. The system is in<u>consistent</u>, meaning no solution:

More examples:

$$\begin{aligned} y-4z &= 8 \\ 2x-3y+2z &= 1 \\ 4x-8y+12z &= 1 \end{aligned} \qquad\qquad \begin{aligned} x+z &= 6 \\ z-3y &= 7 \\ 2x+y+3z &= 15 \end{aligned}$$

Make up (reverse engineer) a linear system for your friend!

Lecture 3

Section 1.2 Row Reduction and Echelon Forms

A matrix is in row echelon form if

1. all zero rows are at the bottom
2. each leading entry of a row is to the right of the leading entry of the row above
3. below a leading entry of a row all entries are 0.

$\square$ = pivot

$$\begin{pmatrix} \square & * & * & * & * \\ 0 & \square & * & * & * \\ 0 & 0 & 0 & \square & * \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \rightsquigarrow \text{easy to solve!}$$

A matrix is in reduced row echelon form if also:

4. the leading entry in each nonzero row is 1
5. each leading 1 is the only nonzero entry in its col.

$$\begin{pmatrix} 1 & 0 & * & 0 & * \\ 0 & 1 & * & 0 & * \\ 0 & 0 & 0 & 1 & * \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \rightsquigarrow \text{easier to solve!}$$

$\boxed{\text{CLICKER}}$ Which are in reduced row echelon form?

$$\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \quad \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 1 & 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 1 & 8 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 17 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

<u>Theorem</u>. *Each matrix is equivalent to one and only one matrix in reduced row echelon form.*

The Row Reduction Algorithm

Step 1. Swap rows (if needed) so in the leftmost nonzero column the top entry (= pivot) is nonzero.

Step 2. Scale the top row so the pivot becomes 1.

Step 3. Use row replacement to create zeros above and below the pivot.

Step 4. Cover up the top row and go back to Step 1.

When there are no nonzero rows, the result is in reduced row echelon form.

Examples. Solve
$$\begin{aligned} x + 2y + 3z &= 9 \\ 2x - y + z &= 8 \\ 3x \quad - z &= 3 \end{aligned}$$

Solve
$$\begin{aligned} 3x + y + 3z &= 2 \\ x \quad + 2z &= -3 \\ 2x + y + z &= 4 \end{aligned}$$

Solve
$$\begin{aligned} a + 2b + d &= 3 \\ c + d - 2e &= 1 \end{aligned}$$

Solutions of Linear Systems

We want to go from reduced row echelon form to the solution of the linear system.

① $\begin{pmatrix} 1 & 0 & 5 & 0 \\ 0 & 1 & 2 & 1 \end{pmatrix} \rightsquigarrow \begin{aligned} x_1 \quad\; + 5x_3 &= 0 \\ x_2 + 2x_3 &= 1 \end{aligned}$

$\Rightarrow$ x_3 can be anything (it is a <u>free variable</u>)

$x_2 = 1 - 2x_3$

$x_1 = -5x_3$

$\Rightarrow$ Solution is a <u>line</u>.

② $\begin{pmatrix} 1 & 0 & 5 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \rightsquigarrow 0 = 1 \Rightarrow$ inconsistent

<u>Theorem</u>. A linear system is consistent if and only if (exactly when) the last column of the augmented matrix does not have a pivot.
If it is consistent, the solution can be a point, line, plane, etc.

$\boxed{\text{CLICKER}}$ A linear system has 4 variables and 3 equations. What are the possible solution sets?

a. nothing b. point c. line d. plane

e. 3-plane f. 4-plane

Lecture 4

Section 1.3 Vector Equations

A <u>vector</u> is a matrix with one row or one column:

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

row vector — column vector

Adding vectors: add component-wise

$$\begin{pmatrix} 1 \\ 2 \end{pmatrix} + \begin{pmatrix} 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 4 \\ 6 \end{pmatrix}$$

Scaling vectors: scale component-wise

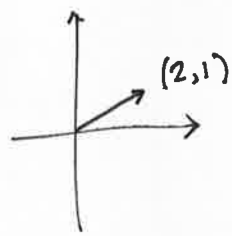
$$7\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 7 \\ 14 \end{pmatrix}$$

So for $u = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ and $v = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$,

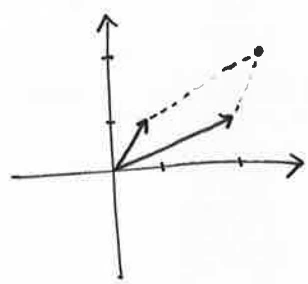
$$2u - v = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$$

Geometry

A length n vector can be drawn as a point or arrow in $\mathbb{R}^n$:



Parallelogram rule for addition:



$$\begin{pmatrix}2\\1\end{pmatrix}+\begin{pmatrix}1\\1\end{pmatrix}=\begin{pmatrix}3\\2\end{pmatrix}$$

Scaling just makes a vector longer or shorter:

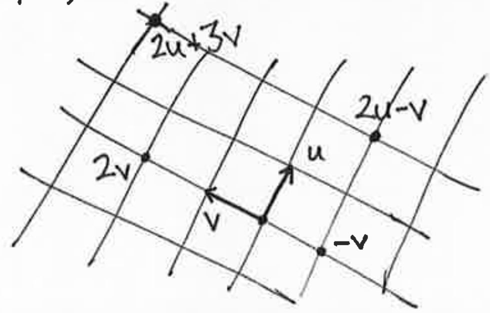
Linear Combinations

A <u>linear combination</u> of the vectors $v_1, \ldots, v_k$ is a vector

$$c_1 v_1 + \cdots + c_k v_k$$

where $c_1, \ldots, c_k$ are real numbers.

picture:



$\boxed{\text{CLICKER}}$ If u is a linear comb. of $v_1, \ldots, v_k$ then there is one way to write u as such (that is, only one choice for $c_1, \ldots, c_k$).

Q. Is $\begin{pmatrix} 8 \\ 16 \\ 3 \end{pmatrix}$ a linear comb of $\begin{pmatrix} 1 \\ 2 \\ 6 \end{pmatrix}$ and $\begin{pmatrix} -1 \\ -2 \\ 1 \end{pmatrix}$?

$\rightsquigarrow$ solve $c_1 \begin{pmatrix} 1 \\ 2 \\ 6 \end{pmatrix} + c_2 \begin{pmatrix} -1 \\ -2 \\ 1 \end{pmatrix} = \begin{pmatrix} 8 \\ 16 \\ 3 \end{pmatrix}$

$\rightsquigarrow$ solve
$$\begin{aligned} c_1 - c_2 &= 8 \\ 2c_1 - 2c_2 &= 16 \\ 6c_1 + c_2 &= 3 \end{aligned}$$

$\rightsquigarrow$ row reduce $\begin{pmatrix} 1 & -1 & 8 \\ 2 & -2 & 16 \\ 6 & 1 & 3 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & 5 \\ 0 & 1 & -3 \\ 0 & 0 & 0 \end{pmatrix}$

$\rightsquigarrow$ $c_1 = 5;\ c_2 = -3.$

Spans

Span $\{v_1, \ldots, v_k\}$ is the set of linear combinations of $v_1, \ldots, v_k$.

We just saw: the question of whether u is in Span $\{v_1, \ldots, v_k\}$ is equivalent to a linear system, solved by row reducing

$$\begin{pmatrix} v_1 & \cdots & v_k & u \end{pmatrix}$$

col. vectors.

$\boxed{\text{CLICKER}}$ What shape can Span $\{v_1, \ldots, v_k\}$ be?
empty, pt, line plane, circle.

Application

Some production costs:

	Materials	Labor	Overhead
Widget	\$1	\$2	\$3
Gadget	\$2	\$3	\$1

Q. What are possible expenditures on materials, labor, and overhead?

A. Span of $(1,2,3)$ and $(2,3,1)$

1.4 The Matrix Equation $Ax = b$

Multiplying matrices

same size!

first, row x col: $\begin{pmatrix} a_1 & \cdots & a_n \end{pmatrix} \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = a_1 b_1 + \cdots + a_n b_n$

next, matrix x col: $\underset{m \times n}{\begin{pmatrix} r_1 \\ \vdots \\ r_m \end{pmatrix}} \underset{n \times 1}{\begin{pmatrix} b \end{pmatrix}} = \underset{m \times 1}{\begin{pmatrix} r_1 b \\ \vdots \\ r_m b \end{pmatrix}}$

example: $\begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 5 \cdot 2 + 6 \cdot 3 \\ 7 \cdot 2 + 8 \cdot 3 \end{pmatrix} = \begin{pmatrix} 28 \\ 38 \end{pmatrix}$

Guess how to do matrix x matrix.

Linear systems vs matrix eqns vs vector eqns

linear system

$$\begin{aligned} x_1 + 2x_2 &= 5 \\ 3x_1 + 4x_2 &= 6 \end{aligned} \quad \longleftrightarrow \quad \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 5 \\ 6 \end{pmatrix}$$

matrix eqn

$$x_1 \begin{pmatrix} 1 \\ 3 \end{pmatrix} + x_2 \begin{pmatrix} 2 \\ 4 \end{pmatrix} = \begin{pmatrix} 5 \\ 6 \end{pmatrix}$$

vector eqn

Q. Say $u = (u_1, u_2, u_3)$, $v = (v_1, v_2, v_3)$, $w = (w_1, w_2, w_3)$
Write $3u - 5v + 7w = 0$ as matrix eqn.

Solutions to linear systems vs spans

Fact. $Ax = b$ has a solution
$\iff$ b in span of columns of A

Why?

Q. Which of the following vectors are in the span of
$(2,3,1,4,0)$, $(3,4,-1,3,5)$, $(1,-1,2,4,3)$?

- $(3,6,-5,-2,-7)$
- $(6,19,-3,4,-12)$

$\boxed{\text{CLICKER}}$ Which of the following true statements can be checked without calculation?

a. $(0,1,2)$ is in span of $(3,3,4)$, $(0,10,20)$, $(0,-1,-2)$
b. $(0,1,2)$ is in span of $(3,3,4)$, $(0,5,7)$, $(0,6,8)$
c. $(0,1,2)$ is in span of $(3,3,4)$, $(0,1,0)$, $(0,0,\sqrt{2})$

Theorem. The following are equivalent:

1. $Ax = b$ has a solution for all b
2. Each b is in span of cols of A
3. The span of cols of A is $\mathbb{R}^m$
4. A has pivot posn in each row.

$A = m \times n$

Why?

Properties of the Matrix Product Ax

- $A(u+v) = Au + Av$
- $A(cu) = cAu$ $\qquad$ c = real num.

$\rightsquigarrow$ say the multiplication is linear.

Check these!

So: If $Au = 0$ and $Av = 0$

then
$$\begin{aligned} A(cu+dv) &= A(cu) + A(dv) \\ &= cAu + dAv \\ &= c\cdot 0 + d\cdot 0 = 0. \end{aligned}$$

that is: each vector in Span$\{u,v\}$ is a soln to $Ax = 0$!

$\boxed{\text{CLICKER}}$ If $b \neq 0$ then the solutions to $Ax = b$ is: always/sometimes/never a span.

Fibonacci Numbers

0, 1, 1, 2, 3, ... in *Liber Abaci* (1202) by Leonardo of Pisa aka Fibonacci and earlier in India

Like a vector in $\mathbb{R}^\infty$ $(f_1, f_2, f_3, \dots)$ satisfying

$$f_1 + f_2 = f_3$$
$$f_2 + f_3 = f_4 \qquad \text{etc.}$$

Are there other solutions in $\mathbb{R}^\infty$?

Can scale: 0, 2, 2, 4, ...

What about: 1, 0, 1, 1, ... can't get this by scaling

Check: Any other soln is a linear combo of these
$\rightsquigarrow$ solutions form a plane in $\mathbb{R}^\infty$

Are there any nice sequences in this span?
arithmetic?
geometric?

If there are nice sequences in the span, can we find formulas for the n^{th} term?
Can we use this to find formula for f_n?

1.5 Solution Sets of Linear Systems

Homogeneous systems

$Ax = 0$

$\rightsquigarrow$ $x = 0$ always a solution: trivial soln.

Fact. $Ax = 0$ has soln $\Leftrightarrow$ there is a free var
$\Leftrightarrow$ there is a col w/o pivot

Examples. $\begin{pmatrix} 1 & 2 & 0 & -1 \\ -2 & -3 & 4 & 5 \\ 2 & 4 & 0 & -2 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & -8 & -7 \\ 0 & 1 & 4 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ two free vars

$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 2 & 1 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$ one free var

$\begin{pmatrix} 1 & 1 & 1 & 1 \end{pmatrix}$ three free vars

Solutions are: planes through origin $\longleftrightarrow$ spans
so if $v_1, \ldots, v_k$ are solns then
Span$\{v_1, \ldots, v_k\}$ are solns.

The dim of the plane is the number of free vars.

$\boxed{\text{CLICKER}}$ A homogeneous system with same/greater/fewer equations than vars can have zero/one/∞ many solns. Which combinations are possible?

Parametric forms

Say free vars for $Ax=0$ are $x_k, \dots, x_n$.
Then solns to $Ax=0$ can be written as

$$x_k v_k + \cdots + x_n v_n \quad (*) \qquad (\text{for some } v_k, \dots, v_n)$$

$\rightsquigarrow$ solutions are $\mathrm{Span}\{v_k, \dots, v_n\}$

$(*)$ is the param. form of the solutions.

Q. Find parametric form for above examples.

Nonhomogeneous systems

$Ax=b, \quad b \neq 0.$

If we solve as before (Sec. 1.4) can find soln in terms of free vars. $x_k, \dots, x_n$
Can then rewrite soln as

$$p + x_k v_k + \cdots + x_n v_n \text{ for some } v_k, \dots, v_n$$

This is the parametric soln.

Examples. Do above examples with

$b = (3, -1, 6)$

$b = (4, 2, 4)$

$b = (9)$

Is there a b making $Ax=b$ inconsistent?

Homogeneous vs. Nonhomogeneous

Key realization: If v, w are solutions to $Ax=b$
then v−w is solution to $Ax=0$ (why?)

↑ associated homog. system.

This means: ① Solutions to $Ax=b$ parallel to solutions to $Ax=0$

② Solutions to $Ax=b$ obtained by taking one solution and adding all possible solns to $Ax=0$.

So by understanding $Ax=0$ we gain understanding of $Ax=b$ for all b. This gives structure to the set of equations $Ax=\square$.

$\boxed{\text{CLICKER}}$ Make a 3×2 matrix A so that $Ax=b$ is consistent exactly when b lies on $x=y=z$ and so solutions are a line of slope 2. (or other slopes).

How does the solution change as we slide b along $x=y=z$?

Q Describe the solution set as b varies in above three examples.

1.6 Applications of Linear Systems

Balancing Chemical Equations

sodium bicarbonate + citric acid $\longrightarrow$ sodium citrate + water + carbon dioxide

$$\underline{x_1}\, NaHCO_3 + \underline{x_2}\, C_6H_8O_7 \longrightarrow \underline{x_3}\, Na_3C_6H_5O_7 + \underline{x_4}\, H_2O + \underline{x_5}\, CO_2$$

$$\begin{aligned}
Na:&\quad x_1 = 3x_3\\
H:&\quad x_1 + 8x_2 = 5x_3 + 2x_4\\
C:&\quad x_1 + 6x_2 = 6x_3 + x_5\\
O:&\quad 3x_1 + 7x_2 = 7x_3 + x_4 + 2x_5
\end{aligned}$$

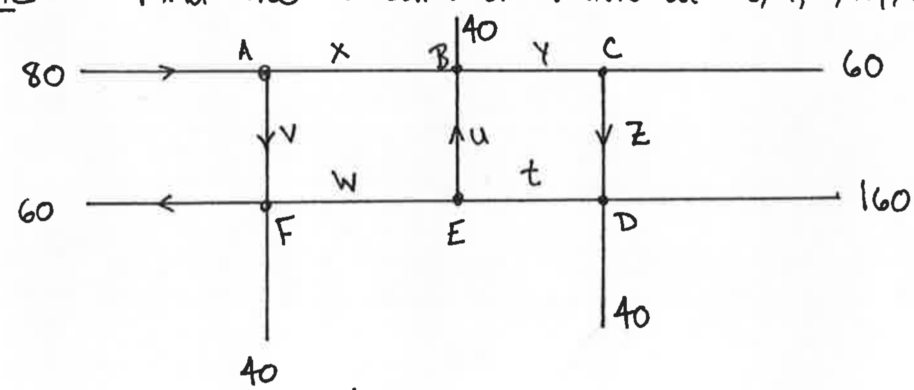
$\rightsquigarrow$ $Ax = 0$ where A is

$$\begin{pmatrix} 1 & 0 & -3 & 0 & 0 \\ 1 & 8 & -5 & -2 & 0 \\ 1 & 6 & -6 & 0 & -1 \\ 3 & 7 & -7 & -1 & -2 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 & -1/3 \\ 0 & 0 & 1 & 0 & -1/3 \\ 0 & 0 & 0 & 1 & -1 \end{pmatrix}$$

$$\rightsquigarrow \begin{aligned} x_1 &= x_5\\ x_2 &= x_5/3\\ x_3 &= x_5/3\\ x_4 &= x_5\\ x_5 &\text{ free} \end{aligned} \qquad \rightsquigarrow (3,1,1,3,3) \text{ or any multiple.}$$

NETWORK FLOW

<u>Traffic</u> Find the amount of traffic at t, u, v, w, x, y, and z.



$$
\begin{array}{c}
\begin{array}{ccccccc|c} x & y & z & u & v & w & t & \end{array} \\
\left(\begin{array}{ccccccc|c}
1 & 0 & 0 & 0 & 1 & 0 & 0 & 80 \\
1 & -1 & 0 & 1 & 0 & 0 & 0 & 40 \\
0 & 1 & -1 & 0 & 0 & 0 & 0 & 60 \\
0 & 0 & 1 & 0 & 0 & 0 & -1 & -120 \\
0 & 0 & 0 & 1 & 0 & 1 & -1 & 0 \\
0 & 0 & 0 & 0 & 1 & 1 & 0 & 100
\end{array}\right)
\end{array}
$$

Conservation law: traffic into an intersection equals traffic out.

$x = w - 20$
$y = t - 60$
$z = t - 120$
$u = t - w$
$v = 100 - w$
t, w free!

Q. What if x, y streets are closed for construction. How should we reroute traffic?

A. $x, y = 0 \Rightarrow z < 0$
$\Rightarrow$ need to reverse z traffic!

<u>Electrical</u> <u>circuits</u>: Similar! Conservation law is Kirchhoff's first law.

ECONOMICS

Leontief's Exchange Model (closed version)

Economy has sectors (manufacturing, communication, etc.)
Each sector has output each year, which it completely distributes among the sectors.
Dollar value of output = price.
Want to find equilibrium prices: income = expenses for each sector.

Distribution of output given by exchange table, for example:

note: cols sum to 1!

	Output from			
	A	B	C	Purchased by
	.3	.3	.3	A
	.4	.1	.5	B
	.3	.6	.2	C

Equilibrium means: $.3p_A + .3p_B + .3p_C = p_A$ etc.

$$\rightsquigarrow \begin{pmatrix} -.7 & .3 & .3 \\ .4 & -.9 & .5 \\ .3 & .6 & -.8 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & -.82 \\ 0 & 1 & -.9 \\ 0 & 0 & 0 \end{pmatrix} \rightsquigarrow \begin{array}{l} p_A = .82p_C \\ p_B = .92p_C \\ p_C \text{ free} \end{array}$$

So C is most expensive. Why does this make sense?

1.7 LINEAR INDEPENDENCE

A set of vectors $v_1, \ldots, v_k$ in $\mathbb{R}^n$ is linearly indep. if

$$x_1 v_1 + \cdots + x_k v_k = 0$$

has only the trivial solution. It is lin. dep. otherwise.

So lin. dep. means there are $c_1, \ldots, c_k$ not all zero so

$$c_1 v_1 + \cdots + c_k v_k = 0.$$

"linear dependence relation"

<u>Fact</u>. The cols of A are linearly indep

$\Leftrightarrow$ $Ax = 0$ has only the trivial soln. why?

Example. Is $\{(1,1,1), (1,-1,2), (3,1,4)\}$ lin indep?

no: $2 \cdot (1,1,1) + (1,-1,2) - (3,1,4) = 0$

$\boxed{\text{CLICKER}}$ For which x are $(1,1,x)$, $(1,x,1)$, $(1,1,x)$ lin dep?

<u>One vector</u> $\{v\}$ lin dep $\Leftrightarrow$ $v = 0$.

<u>Two vectors</u> $\{v, w\}$ lin dep $\Leftrightarrow$ v is a mult of w or vice versa

$\Leftrightarrow$ v, w lie on a line.

<u>More generally</u> $\{v_1, \dots, v_k\}$ lin dep $\Leftrightarrow$ $v_1, \dots, v_k$ lie in a $(k-1)$-plane

$\Leftrightarrow$ some v_i is a linear comb of $v_1, \dots, v_{i-1}$.

Why? If $\{v_1, \dots, v_k\}$ lin dep $\rightsquigarrow$ $c_1 v_1 + \dots + c_k v_k = 0$, c_i not all 0.

Choose largest i so $c_i \neq 0$, move $c_1 v_1 + \dots + c_{i-1} v_{i-1}$ to other side, divide both sides by c_i

$\rightsquigarrow$ have v_i as linear combo of $v_1, \dots, v_{i-1}$.

(Other direction easier.)

Beware! If $v_1 = 0$ this still works. Need to interpret $v_1, \dots, v_{i-1}$ as $\emptyset$ and $\operatorname{Span} \emptyset = 0$.

<u>Example 4</u> $u = (3,1,0)$ $v = (1,6,0)$

Describe $\operatorname{span}\{u, v\}$.

Explain why w in $\operatorname{Span}\{u,v\}$ $\leftrightarrow$ $\{u, v, w\}$ lin dep.

<u>Two More Facts</u>

① Say $v_1, \dots, v_k$ in $\mathbb{R}^n$

If $k > n$ then $\{v_1, \dots, v_k\}$ lin dep.

② If one of $v_1, \dots, v_k$ is 0 then $\{v_1, \dots, v_k\}$ is lin dep.

1.8 Intro to Linear Transformations

$A = m \times n$ matrix

$\rightsquigarrow$ function $T : \mathbb{R}^n \to \mathbb{R}^m$ "matrix transformation"

where $T(v) = Av$

domain: $\mathbb{R}^n$

target/codomain: $\mathbb{R}^m$

image/range: span of cols of A why?

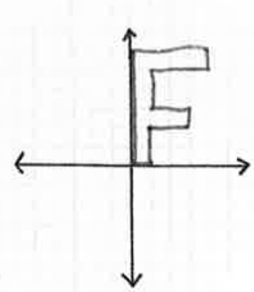
This is a <u>fourth</u> point of view: matrices, vector eqns, linear systems, matrix trans.

Will use it to describe dynamical systems, e.g. genetics example from day 1.

Example. $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}$ $u = (3,4)$ $b = (7,5,7)$

- Find $T(u)$
- Find v so $T(v) = b$. How many such v?
- Find c so there is no v with $T(v) = c$.

$\boxed{\text{CLICKER}}$ What does $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ do to this F?



Geometric examples

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{projection}$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{reflection in plane}$$

Q. How to get reflection in line, point?

$$\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \quad \text{rotation + dilation}$$

Q. How to get just dilation?

$$\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \quad \text{shear}$$

Linear transformations

A function $T : \mathbb{R}^n \to \mathbb{R}^m$ is linear if

(1) $T(u+v) = T(u) + T(v)$ for all u, v in $\mathbb{R}^n$

(2) $T(cv) = cT(v)$ for all v in $\mathbb{R}^n$, c in $\mathbb{R}$

It follows that:

$$T(c_1 v_1 + \cdots + c_k v_k) = c_1 T(v_1) + \cdots + c_k T(v_k)$$

This is the principle of superposition in engineering. (v = signal, $T(v)$ = response).

Fact. Every matrix transformation is linear   why?

Next time: Every linear transf is a matrix transf!

1.9 The Matrix of a Linear Transformation

Last time: every matrix transf. is a linear transf.

<u>Theorem</u>. Every linear transf. is a matrix transf.

This means: for any linear transf $T: \mathbb{R}^n \to \mathbb{R}^m$
there is an $m \times n$ matrix A so $T(v) = Av$
for all v in $\mathbb{R}^n$.

i^{th} spot

why? Take $A = \big(T(e_1) \cdots T(e_n) \big)$ $\quad e_i = (0,\ldots,0,1,0,\ldots,0)$

Check $Ae_i = T(e_i)$

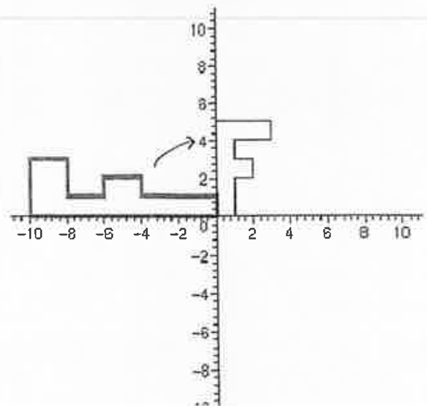
It follows from linearity (= superposition)
that $Av = T(v)$ for all v.

Q. Find the matrix that rotates $\mathbb{R}^2$ by $\pi/4$.
Or θ.

Q. Find a matrix that puts $\mathbb{R}^2$ rigidly onto the yz-plane in $\mathbb{R}^3$.

Q. Find the matrix that reflects $\mathbb{R}^2$ about $y = x$.

$\boxed{\text{CLICKER}}$ Find a matrix that does this:



One-to-one and onto

$T : \mathbb{R}^n \to \mathbb{R}^m$ is one-to-one if each b in $\mathbb{R}^m$ is the image of at most one v in $\mathbb{R}^n$.
It is onto if the image of T is $\mathbb{R}^m$, that is, each b in $\mathbb{R}^m$ is the image of at least one v in $\mathbb{R}^m$.

Q. What can we say about the relative sizes of m & n if T is one-to-one, onto, or both?

Theorem. Say $T : \mathbb{R}^n \to \mathbb{R}^m$ is a lin. transf. corresp. to a matrix A.

- T is onto $\Leftrightarrow$ cols of A span $\mathbb{R}^m$
- T is one-to-one $\Leftrightarrow$ cols of A are lin ind $\Leftrightarrow$ $Ax = 0$ has only 0 solution.

Why?

Q. Do the following matrices give linear tranf's that are one-to-one? onto?

$$\begin{pmatrix} 1 & 0 & 7 \\ 1 & 1 & 2 \\ 2 & 1 & 9 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 2 & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$$

Q. Draw a picture of a one-to-one $\mathbb{R} \to \mathbb{R}^3$
of an onto $\mathbb{R}^3 \to \mathbb{R}$.

CHAPTER 1 IN A NUTSHELL

We want to solve linear systems.

why? engineering, econ, chem, physics, ...

$\rightsquigarrow$ matrices & row echelon form

Solutions to $Ax=0$ is a plane P_0 through 0.

$\dim P_0 = \#$ free vars $= \#$ cols $- \#$ pivots

can write P_0 as a span $\rightsquigarrow$ parametric form.

Solutions to $Ax=b$ is a plane P_b parallel to P_0

If p is one given solution to $Ax=b$

then $P_b = p + P_0$ $\rightsquigarrow$ param form for P_b.

Matrix eqns correspond to vector eqns, so:

$Ax=b$ is consistent $\Leftrightarrow$ b in span of cols of A.

If A is $m \times n$ get $T_A : \mathbb{R}^n \to \mathbb{R}^m$, matrix transformation.

Matrix transformations and linear transformations are the same thing.

$Ax=b$ is consistent for all b $\Leftrightarrow$ cols of A span $\mathbb{R}^m$

$\Leftrightarrow$ A has a pivot in each row

$\Leftrightarrow$ T_A is onto

$Ax=b$ has exactly one solution $\Leftrightarrow$ cols of A are linearly ind.

$\Leftrightarrow$ A has a pivot in each col

$\Leftrightarrow$ T_A is one-to-one.

LIGHT GAME

Consider a 5×5 grid. In each position there is a light, which can be on or off. You start with some lights on and the goal is to turn them all off. When you click on a square, the four lights above, below, left, and right will toggle.

Q. Which starting configurations have a solution?
What if all lights are on at the start?

Hint: Think about mod 2 arithmetic.

What about $n \times n$? Other variations?

Play the game on the course web site!

2 Matrix Algebra

2.1 Matrix Operations

Terminology ~~Basic Operations~~

A = $m \times n$ matrix

a_{ij} or A_{ij} = ij^{th} entry

a_{ii} are <u>diagonal</u> entries

diagonal matrix: all non-diagonal entries are 0.

zero matrix: all entries 0.

Sums and Scalar Multiples

Same as for vectors: component-wise

$\rightsquigarrow$ matrices must be same size to add.

Basic rules:

$$A+B = B+A \qquad r(A+B) = rA + rB$$

$$(A+B)+C = A+(B+C) \qquad (r+s)A = rA + sA$$

$$A+0 = A \qquad (rs)A = r(sA)$$

Matrix Multiplication

$A = m \times n$, $B = n \times p$

$$\underbrace{\begin{pmatrix} r_1 \\ \vdots \\ r_m \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} c_1 & \cdots & c_p \end{pmatrix}}_{B}$$

AB is $m \times p$:

$\rightsquigarrow (AB)_{ij} = \sum_k a_{ik} b_{kj} = r_i \cdot b_j$

$\boxed{\text{CLICKER}}$ Are there $A, B \neq 0$ with $AB = 0$?

Fact. $T_{AB} = T_A \circ T_B$

so: matrix multiplication $\longleftrightarrow$ composition of lin. trans.

why? enough to check $T_{AB}(e_i) = T_A \circ T_B(e_i)$

$$\begin{aligned} T_{AB}(e_i) &= AB(e_i) = i^{\text{th}} \text{ col of } AB \\ &= \begin{pmatrix} r_1 c_i \\ \vdots \\ r_m c_i \end{pmatrix} = Ac_i \\ &= A(Be_i) = A(T_B(e_i)) \\ &= T_A \circ T_B(e_i) \end{aligned}$$

Another View: $AB = A(c_1 \cdots c_p) = (Ac_1 \;\cdots\; Ac_p)$

Properties of Matrix Multiplication

$A(BC) = (AB)C$

$A(B+C) = AB + AC$

$(B+C)A = BA + CA$

$r(AB) = (rA)B = A(rB)$

$I_m A = A = A I_n$

$I_m = m \times m$ identity matrix

$$\begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$$

Most interesting is associativity: multiplication is associative because function composition is! (Or just check with the formula).

Commutativity? Find A, B so $AB \neq BA$.

In general: $AB \neq BA$

$AB = AC$ does <u>not</u> mean $B = C$

$AB = 0$ does <u>not</u> mean A or B is 0.

Powers

$A^k = A \cdots A$ (k times)

Transpose

$A = m \times n \rightsquigarrow A^T = n \times m \quad (A^T)_{ij} = A_{ji}$

Properties: $(A^T)^T = A$

$(A+B)^T = A^T + B^T$

$(rA)^T = rA^T$

and $(AB)^T = B^T A^T$

2.2 THE INVERSE OF A MATRIX

$A = n \times n$ matrix

A is <u>invertible</u> (or nonsingular) if there is a matrix A^{-1} with

$$AA^{-1} = A^{-1}A = I_n$$

A^{-1} = <u>inverse</u> of A.

example. $\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$

Q. Can you guess the inverse of $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$?

Q. Find a matrix that is not invertible.

Fact. Inverses are unique.

why? If B and C are inverses of A then

$$B = BI = BAC = C.$$

$\boxed{\text{CLICKER}}$ Which of the following linear transformations of $\mathbb{R}^3$ correspond to invertible matrices?

- projection to xy-plane
- rotation about z-axis by π.
- reflection through origin
- reflection through xy-plane

The 2x2 case

Fact. Say $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, write $\det(A) = ad - bc$ "determinant"

If $\det(A) \neq 0$ then A is invertible and

$$A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

If $\det(A) = 0$ then A is not invertible.

why? if $\det(A) \neq 0$, just check!

other part harder. show $\det AB = \det A \det B$.

example. $\begin{pmatrix} 3 & 2 \\ 2 & 2 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & -1 \\ -1 & 3/2 \end{pmatrix}$

Solving Linear Systems via Inverses

Fact. If A is invertible, $Ax = b$ has exactly one solution, namely $x = A^{-1} b$ why?

example. Solve

$$\begin{aligned} 2x + 3y + 2z &= 1 \\ x + 3z &= 1 \\ 2x + 2y + 3z &= 1 \end{aligned}$$

using $\begin{pmatrix} 2 & 3 & 2 \\ 1 & 0 & 3 \\ 2 & 2 & 3 \end{pmatrix}^{-1} = \begin{pmatrix} -6 & -5 & 9 \\ 3 & 2 & -4 \\ 2 & 2 & -3 \end{pmatrix}$

Q. Better than old way?

Some Facts A, B invertible $n \times n$ matrices

① A^{-1} is invertible and $(A^{-1})^{-1} = A$

② AB is invertible and $(AB)^{-1} = B^{-1}A^{-1}$

③ A^T is invertible and $(A^T)^{-1} = (A^{-1})^T$

why?

Q. What is $(ABC)^{-1}$?

An Algorithm for finding A^{-1}

$A = n \times n$ matrix.

Row reduce $(A \mid I_n)$

If reduction has the form $(I_n \mid B)$ then

A is invertible and $B = A^{-1}$

Otherwise A is not invertible.

Q ~~example~~. Find $\begin{pmatrix} 1 & 3 & 3 \\ 1 & 4 & 3 \\ 1 & 3 & 4 \end{pmatrix}^{-1}$

Why does this work? First answer: we can think of the algorithm as simultaneously solving

$$Ax_1 = e_1$$
$$Ax_2 = e_2 \quad \text{etc.}$$

But the cols of A^{-1} are $A^{-1}e_i$, which is x_i.

Second answer is more algebraic...

Elementary Matrices

An elementary matrix is one that differs from I_n by one row op.

Fact. If E is an elem matrix for some row op then EA differs from A by same row op.

why? check for each row op.

Fact. Elem. matrices are invertible. why?

Theorem. An $n \times n$ matrix A is invertible iff it is row equiv. to I_n. In this case the seq of row ops taking A to I_n also takes I_n to A^{-1}

This gives a second explanation of the algorithm.
Why is it true? Because:

$$E_k \cdots E_1 A = I \qquad (\text{mult. on right by } A^{-1})$$

$$\rightsquigarrow E_k \cdots E_1 I = A^{-1}$$

CRYPTOGRAPHY

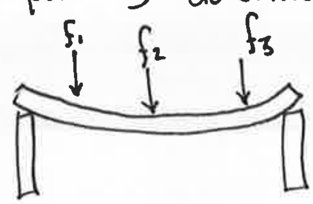
- Encode letters A,…,Z by 1,…,26.
- Choose a matrix A, say $n \times n$.
- Break messages into blocks of size $n \rightsquigarrow$ vectors.
- Apply A to vectors to get encrypted message.

Example: $A = \begin{pmatrix} 2 & 3 & 1 \\ 2 & 0 & 1 \\ 0 & 0 & 6 \end{pmatrix}$ encoded message: $\begin{pmatrix} 112 \\ 52 \\ 36 \end{pmatrix}$

What is the unencoded message?

STRUCTURAL ENGINEERING

Suppose you put 3 downward forces on an elastic beam:



Hooke's law $\rightsquigarrow$ the vertical displacements at those three points y_1, y_2, y_3 are given by a linear transformation: $A\begin{pmatrix} f_1 \\ f_2 \\ f_3 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$

So if you want to achieve a certain displacement, use A^{-1} to find the required forces!

2.3 Characterizations of Invertible Matrices

<u>Invertible Matrix Theorem</u>. $A = n \times n$ matrix, $T_A : \mathbb{R}^n \to \mathbb{R}^n$ the associated linear transf.

TFAE:

(a) A is invertible

(b) A is row equiv to I_n.

(c) A has n pivots

(d) $Ax = 0$ has only 0 soln

(e) cols of A are lin ind.

(f) T_A is one-to-one

(g) $Ax = b$ is consistent for all b in $\mathbb{R}^n$

(h) cols of A span $\mathbb{R}^n$

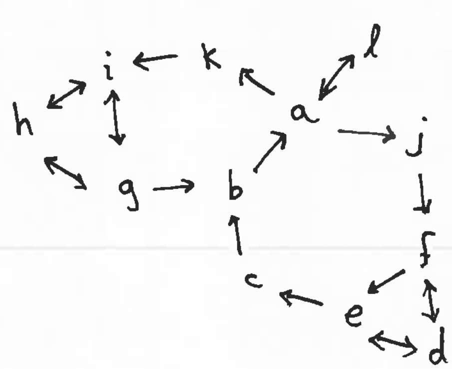
(i) T_A is onto

(j) A has a left inverse

(k) A has a right inverse

(l) A^T is invertible.

why? one possible road map:



So, there are only two kinds of <u>square</u> matrices:

invertible/non singular and non-invertible/singular.

For invertible matrices (a) - (ℓ) hold and for non-invertible matrices the negations of (a)-(ℓ) hold.

Q. State the negations of (a) - (ℓ).

Q. Are the following equivalent? $\boxed{\text{CLICKER}}$

(m) rows of A span $\mathbb{R}^n$

(n) rows of A are lin ind.

(o) $Ax=b$ has exactly one solution for all b in $\mathbb{R}^n$

(p) $\det(A) \neq 0$ where $\det(A)$ is volume of the parallelpiped (= brick) spanned by cols of A

(q) A^2 invertible.

<u>Invertible Functions</u>

A function $f: X \to Y$ is <u>invertible</u> if there is a function $g: Y \to X$ so $f \circ g$ and $g \circ f$ = id.

that is: $g \circ f(x) = x$ for all $x \in X$ <u>and</u>

$f \circ g(y) = y$ for all $y \in Y$

<u>Fact 1</u>. If a function has an inverse, it is unique

~~and~~ ⇝ call it f^{-1}.

<u>Fact 2</u>. Invertible functions are one-to-one and onto.

<u>Fact</u>. $A = n \times n$ matrix, T_A the assoc. lin. transf.

T_A is invertible <u>as a function</u> if and only if A is invertible.

And in this case $(T_A)^{-1} = T_{A^{-1}}$

why? T_A invertible $\overset{\text{Fact 2.}}{\Longrightarrow}$ T_A onto $\overset{\text{IMT}}{\Longrightarrow}$ A inv.

A inv $\Longrightarrow$ $T_A \circ T_{A^{-1}} = T_{AA^{-1}} = T_I = id$

& $T_{A^{-1}} \circ T_A = id$

2.5 MATRIX FACTORIZATIONS

Recall: If we want to solve $Ax=b$ for many b, good to find A^{-1}.

What if A is not invertible? Or not even square?

When solving $Ax=b_1$ and $Ax=b_2$, doing basically the same row ops. Should have a way of not repeating the same steps.

An <u>LU-factorization</u> of $A = m\times n$ matrix is

$$A = LU$$

where $L = m\times m$ lower triang.

U = an echelon form of A

To solve $Ax=b$

$\rightsquigarrow LUx=b$

$\rightsquigarrow$ ① Solve $Ly=b$

② Solve $Ux=y$

this uses <u>only</u> back sub, not elimination.

$$\begin{tikzcd}
x & y & b \\
\mathbb{R}^n \arrow[r, "U"] \arrow[rr, bend right, "A"'] & \mathbb{R}^m \arrow[r, "L"] & \mathbb{R}^m
\end{tikzcd}$$

again: matrix mult. is composition!

example. $A = \begin{pmatrix} 6 & 0 & 2 \\ 24 & 1 & 8 \\ -12 & 1 & -3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ -2 & 1 & 1 \end{pmatrix} \begin{pmatrix} 6 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

L records negatives of row rep. ops.
(be careful to go in order!)

why? row ops are elem matrices

$\rightsquigarrow E_4 E_3 E_2 E_1 A = U$

$\rightsquigarrow A = (E_1^{-1} E_2^{-1} E_3^{-1} E_4^{-1})\, U$

$= LU$

since our E_i in this case are all lower Δ, their product is.

$\boxed{\text{CLICKER}}$ Which are true?

① Every matrix has an LU fact.

② LU fact's are unique.

③ LU is really faster than row red.

example. Solve $Ax = \begin{pmatrix} 4 \\ 19 \\ -6 \end{pmatrix}$ (same A)

① Solve $Ly = \begin{pmatrix} 4 \\ 19 \\ -6 \end{pmatrix}$

$\rightsquigarrow y = \begin{pmatrix} 4 \\ 3 \\ -1 \end{pmatrix}$

② solve $Ux = \begin{pmatrix} 4 \\ 3 \\ -1 \end{pmatrix}$

$\rightsquigarrow x = \begin{pmatrix} 1 \\ 3 \\ -1 \end{pmatrix}$

A non-square example:

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

The above procedure for LU-fact works when we don't need to swap rows in row reduction. What to do in that case?

$$PA = LU \quad \text{or} \quad A = PLU$$

P = permutation matrix.

LU has many applications... next time.

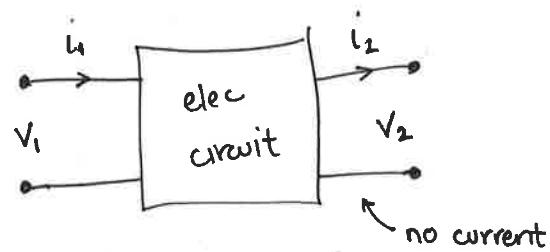
Some extra examples to try:

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & -4 & 6 \\ 3 & -9 & -3 \end{pmatrix} \quad \begin{pmatrix} 4 & 3 \\ 6 & 3 \end{pmatrix} \quad \begin{pmatrix} 1 & 1 & 1 \\ 2 & 3 & 5 \\ 4 & 6 & 8 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

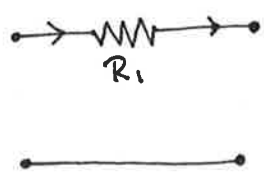
LU-factorizations in Electrical Engineering

In an electrical circuit, current i and voltage v often change by a linear transf. (by Ohm's law & Kirchoff's laws).



so $A \begin{pmatrix} v_1 \\ i_1 \end{pmatrix} = \begin{pmatrix} v_2 \\ i_2 \end{pmatrix}$ for some "transfer matrix" A.

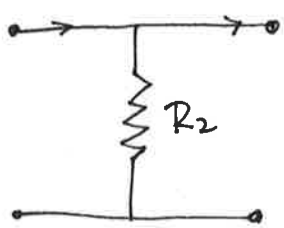
examples:



series circuit

$$A = \begin{pmatrix} 1 & -R_1 \\ 0 & 1 \end{pmatrix}$$

current unchanged, voltage decreases proportional to current



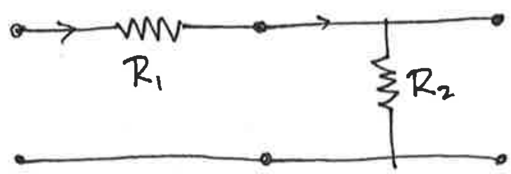
shunt circuit

$$A = \begin{pmatrix} 1 & 0 \\ -1/R_2 & 1 \end{pmatrix}$$

voltage unchanged, current decreased proportional to voltage

If we string these together we get a ladder circuit. The transfer matrix for the ladder circuit is the product of the matrices for the components. (Why does this make sense? Think about function composition!)

So the matrix for



is $$\begin{pmatrix} 1 & 0 \\ -1/R_2 & 1 \end{pmatrix}\begin{pmatrix} 1 & -R_1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -R_1 \\ -1/R_2 & 1+R_1/R_2 \end{pmatrix}$$

Can you reverse engineer a ladder circuit whose transfer matrix is

$$\begin{pmatrix} 1 & -8 \\ -.5 & 5 \end{pmatrix} \;?$$

Use LU!

2.8 Subspaces of $\mathbb{R}^n$.

Subspaces

A subspace of $\mathbb{R}^n$ is a subset V with
① If u, v in V then $u+v$ in V
② If u in V and c in $\mathbb{R}$ then $cu \in V$.

Note: by ②, 0 must be in V.

Example. If v_1, v_2 in $\mathbb{R}^n$ then $\operatorname{span}\{v_1, v_2\}$ is a subspace.

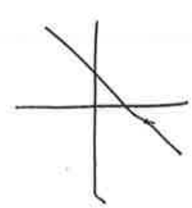
why? linear combos of linear combos are lin. combos

Fact. Subspaces are same as spans

why? just saw why spans are subspaces.
if V is a subspace then $V = \operatorname{span}\{V\}$.

Also recall: spans are same as planes thru 0.

Non-example. This is not a span



what fails?

If $V = \operatorname{span}\{v_1, \ldots, v_k\}$ we say V is the subspace **generated** by $v_1, \ldots, v_k$

$\boxed{\text{CLICKER}}$ Consider $V = \left\{ \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} \in \mathbb{R}^4 : ad - bc = 0 \right\}$.
Is V a subspace of $\mathbb{R}^4$?

Column space and Null space

$A = m \times n$ matrix

The column space of A is the span of the columns in $\mathbb{R}^m$
The null space is the set of solutions to $Ax = 0$ in $\mathbb{R}^n$.

Example. $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{pmatrix}$

col space is $\text{span}\left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}$ = line in $\mathbb{R}^3$
null space is $x + y = 0$ line in $\mathbb{R}^2$

Fact. Null spaces are spans. why?

Bases

V = subspace of $\mathbb{R}^n$
A basis for V is a set of vectors $\{v_1, \ldots, v_k\}$ so
1. $V = \text{span}\{v_1, \ldots, v_k\}$
2. the v_i are lin ind.

Note: $k = \dim V$

Standard basis for $\mathbb{R}^n$: $e_1, \ldots, e_n$

Q. What is a basis for the set of Fibonacci sequences in $\mathbb{R}^\infty$?

Q. Find a basis for the null space of $\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} = A$

A. Solns to $Ax = 0$: $y\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + z\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$

$\rightsquigarrow$ basis: $\left\{ \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \right\}$

check generation & lin ind.

Q. Find a basis for the col space of (same) A.

A. $\left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}$

<u>Fact</u>. In general, the pivot cols of A form a basis for the col space (not the reduced pivot cols!)

why?

In particular: A is invertible iff cols of A form a basis for $\mathbb{R}^n$ $\quad$ $A = n \times n$

2.9 Dimension and Rank

Bases as coordinate systems

V = subspace of $\mathbb{R}^n$

$\mathcal{B} = \{b_1, \ldots, b_k\}$ basis for V

x in V

$\rightsquigarrow$ can write x uniquely as

$$c_1 b_1 + \cdots + c_k b_k \qquad \text{why?}$$

$\rightsquigarrow$ $[x]_{\mathcal{B}} = \begin{pmatrix} c_1 \\ \vdots \\ c_k \end{pmatrix}$ "$\mathcal{B}$-coords of x"

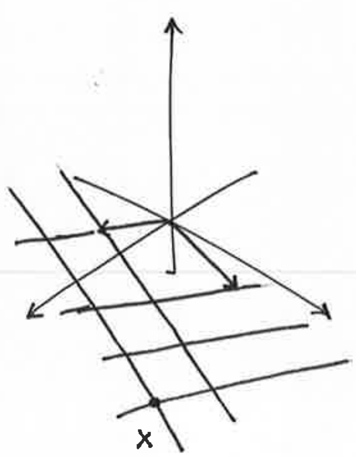
example. $b_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ $b_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$

$V = \text{span}\{b_1, b_2\}$

$\mathcal{B} = \{b_1, b_2\}$ is a basis for V. why?

$x = \begin{pmatrix} 5 \\ 3 \\ 5 \end{pmatrix}$

$\rightsquigarrow [x]_{\mathcal{B}} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$



Dimension

V = subspace of $\mathbb{R}^n$

$\dim V$ = # vectors in a basis for V — why is this well defined?

Note: basis for $\{0\}$ is $\{\}$

$\Rightarrow \dim\{0\} = 0.$

$\boxed{\text{CLICKER}}$ U, V are 2-dim subspaces of $\mathbb{R}^4$. What are the possible dims for $U \cap V$?
(What about $U+V$?)

Rank Theorem

$$\operatorname{rank}(A) = \dim \operatorname{col}(A) = \#\text{ pivot cols}$$

$$\dim \operatorname{Nul}(A) = \#\text{ non-pivot cols}$$

Rank Thm. $A = m \times n$ matrix

$$\operatorname{rank} A + \dim \operatorname{Nul} A = n.$$

example. $\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} = A$ $\quad \operatorname{rank} A = 1, \ \dim \operatorname{Nul} A = 2$

or T_A crushes 2 dim, leaving 1

$\boxed{\text{CLICKER}}$ A, B 3×3. What are possible values of $\operatorname{rank} AB$ if $\operatorname{rk} A = \operatorname{rk} B = 2$?

Two More Theorems

Basis Thm. • V = k-dim subspace of $\mathbb{R}^n$

- Any k lin ind vectors of V form a basis for V
- Any k vectors that span V form a basis for V.

Invertible Matrix Thm (cont)

(m) cols of A form a basis for $\mathbb{R}^n$

(n) $\text{Col } A = \mathbb{R}^n$

(o) $\dim \text{Col } A = n$

(p) $\text{rk } A = n$

(q) $\text{Nul } A = \{0\}$

(r) $\dim \text{Nul } A = 0$

why? (m) $\leftrightarrow$ cols A span $\mathbb{R}^n$

$Ax=b$ consist for all $b \longrightarrow n \longrightarrow o \longrightarrow p \longrightarrow r$

$\longrightarrow q \longrightarrow$ $Ax=0$ has only 0 soln.

CHAPTER 2 IN A NUTSHELL

Still solving $Ax=b$...

<u>Products</u> $AB \neq BA$

$T_{AB} = T_A \circ T_B$

<u>Inverses</u> $AB = BA = I \rightsquigarrow B = A^{-1}$

$T_A^{-1} = T_{A^{-1}}$

$Ax = b \rightsquigarrow x = A^{-1}b$ (easy to solve for many b)

example: flexibility matrix $D \cdot \text{forces} = \text{displacement}$

Find A^{-1} by: $(A|I) \rightsquigarrow (I|A^{-1})$

Another view: $E_k \cdots E_1 A = I$ E_i elementary

$\rightsquigarrow E_k \cdots E_1 = A^{-1}$

<u>Invertible Matrix Thm</u> $A = n \times n$ matrix

TFAE: A invertible

$A \sim I$

A has n pivots

$\text{Nul } A = \{0\}$

T_A 1-1

T_A onto

$Ax=b$ consist. for all b.

$\text{rank } A = n$

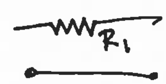
etc...

LU Decompositions Say $A = LU$ L = unit lower Δ
U = echelon form

$\rightsquigarrow$ easy to solve $Ax=b$ for many b
even if A not invertible or $\square$

Step 1. Solve $Ly=b$
Step 2. Solve $Ux=y$

To find L, U: row reduce, record neg. of row ops
notes: must go col by col.
only use lower row replacement

Application: circuits  $\begin{pmatrix} 1 & -R_1 \\ 0 & 1 \end{pmatrix}$ $\begin{pmatrix} V \\ i \end{pmatrix}$

 $\begin{pmatrix} 1 & 0 \\ -1/R_2 & 1 \end{pmatrix}$

Subspaces A subspace of $\mathbb{R}^n$ is a nonempty subset closed under taking linear combos

Subspaces = spans = planes thru 0.

$\mathrm{Col}\,A$ = span of cols of A
$\mathrm{Nul}\,A$ = solns to $Ax=0$
both are subspaces

Basis for a subspace: a lin ind. set that spans
dim of a subspace = # basis elts

Find bases for $\mathrm{Col}\,A$, $\mathrm{Nul}\,A$ by row reducing
$\hookrightarrow$ pivot cols $\searrow$ param. form

$\mathcal{B}$ = basis $\rightsquigarrow$ find $\mathcal{B}$-coords for x, $[x]_{\mathcal{B}}$ by solving $(b_1 \cdots b_k)c = x$.

Rank Thm $A = m \times n$ matrix

$$\text{rank } A + \dim \text{Nul } A = n.$$

Basis Thm $V = k$-dim subsp. of $\mathbb{R}^n$

① Any k lin. ind. vectors form a basis

② Any k spanning vectors form a basis

3.1 Introduction to Determinants

Given a matrix, want a number that tells us if the matrix is invertible or not

$\rightsquigarrow$ volume of the parallelpiped determined by rows will work.

Let's see. When is $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ invertible?

Row reduce: $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \rightsquigarrow \begin{pmatrix} a & b \\ ac & ad \end{pmatrix} \rightsquigarrow \begin{pmatrix} a & b \\ 0 & ad-bc \end{pmatrix}$

So invertible iff $\det = ad - bc$ is nonzero.

Can do same for 3×3 but formula is much more complicated

It turns out there is such a function that works for $n \times n$ matrices. It is called the determinant.

<u>Formula for determinant</u>

Will give a <u>recursive</u> formula. $\det(0 \times 0 \text{ matrix}) = 1$

$A = (a_{ij})$ $n \times n$

$A_{ij} =$ ij-minor of A

$= n \times n$ matrix obtained by deleting i-th row j-th col.

$C_{ij} = (-1)^{i+j} \det A_{ij}$ $=$ ij-cofactor of A.

$\det A = \sum_{j=1}^{n} a_{ij} C_{ij}$ "cofactor expansion"

1x1 case $\det(a_{11}) = a_{11} \cdot C_{11} = a_{11} \cdot \det(0\times 0) = a_{11} \cdot 1 = a_{11}$

2x2 case $\det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11} \cdot \det(a_{22}) - a_{12} \cdot \det(a_{21})$

$= a_{11} a_{22} - a_{21} a_{12}$

3x3 case Write down the general formula...

example. Compute*

$$\det\begin{pmatrix} 5 & 1 & 0 \\ -1 & 3 & 2 \\ 4 & 0 & -1 \end{pmatrix}$$

It turns out that you can expand across any row or col:

$$\det A = \sum_{j=1}^{n} a_{ij} C_{ij} \quad \text{for any } i$$

$$\det A = \sum_{i=1}^{n} a_{ij} C_{ij} \quad \text{for any } j$$

So look for the row/col with the most zeros

example. Compute $\det\begin{pmatrix} 2 & 1 & 0 \\ 1 & 1 & 0 \\ 5 & 9 & 1 \end{pmatrix}$

* A trick for 3x3 matrices: add all products of triples on ↘ diags and subtract those on ↙ diags

e.g. $\det\begin{pmatrix} 5 & 1 & 0 \\ -1 & 3 & 2 \\ 4 & 0 & -1 \end{pmatrix} = 5 \cdot 3 \cdot -1 + 1 \cdot 2 \cdot 4 + 0 \cdot -1 \cdot 0$
$- 0 \cdot 3 \cdot 4 - 1 \cdot -1 \cdot -1 - 5 \cdot 2 \cdot 0$

Triangular matrices

Fact. If A is upper/lower triangular, $\det A$ is product of diag entries

why?

What about "off-triangular"?

e.g. $\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 2 & 3 \\ 0 & 4 & 5 & 6 \\ 7 & 8 & 9 & 10 \end{pmatrix}$

3.2 PROPERTIES OF DETERMINANTS

* We still don't know that our $2n$ formulas for $\det A$ are the same, but for now, lets assume they are.

<u>Effect of row ops</u>

$$\text{If we perform the row op } \left\{\begin{array}{l} \text{row repl.} \\ \text{row swap} \\ \text{row scale by } k \end{array}\right\} \text{ to } A, \det A \text{ changes by factor of } \left\{\begin{array}{c} 1 \\ -1 \\ k \end{array}\right\}$$

why? last one easy using cofactor exp.

What about row replacement? can check directly for 2×2. For 3×3 use cofactor exp. across a row <u>not</u> involved in the row replacement. In each minor we have done (the same) row replacement - but on a 2×2 matrix. So the a_{ij} don't change and the $\det A_{ij}$ don't either, so det stays same!

Can use this to compute det's more quickly.

$$\begin{pmatrix} 0 & 6 & 11 \\ 2 & 7 & 9 \\ 1 & 3 & 4 \end{pmatrix} \xrightarrow[-1]{\text{factor}} \begin{pmatrix} 1 & 3 & 4 \\ 2 & 7 & 9 \\ 0 & 6 & 11 \end{pmatrix} \xrightarrow[1/2]{\text{factor}} \begin{pmatrix} 2 & 6 & 8 \\ 2 & 7 & 9 \\ 0 & 6 & 11 \end{pmatrix}$$

$$\xrightarrow[1]{\text{factor}} \begin{pmatrix} 2 & 6 & 8 \\ 0 & 1 & 1 \\ 0 & 0 & 3 \end{pmatrix} \leftarrow \det = 6$$

so original matrix has det $6 \cdot 1/2 \cdot -1 = -3$.

If we row reduce without row scales:

$$\det A = (-1)^{\#\text{ swaps}} (\text{product of diag entries of REF})$$

And we see:

$$A \text{ invertible} \iff \det A \neq 0.$$

Determinants and Products

Can check that if:

E is an elem. matrix corr. to $\left\{\begin{array}{l}\text{row repl.}\\ \text{row swap}\\ \text{row scale by } k\end{array}\right\}$ then $\det E = \left\{\begin{array}{l}1\\ -1\\ k\end{array}\right.$

We now see that if E is elem. matrix

$$\det EB = \det E \det B$$

From this we can deduce first:

If $A = E_1 \cdots E_k$ then

$$\det A = \det E_1 \cdots \det E_k \qquad \left(\begin{array}{c}\text{apply one at a}\\ \text{time}\end{array}\right)$$

and then for any A, B

$$\det AB = \det A \det B$$

Just break A into $E_1 \cdots E_k$ and apply previous two facts.

Properties of Determinants (review)

1. A invertible $\Leftrightarrow$ $\det A \neq 0$.

2. $\det A$ = vol. of parallelpiped determined by rows (or cols) of A

3. $\det AB = \det A \det B$

4. $$\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij} \quad \text{any } i$$
$$\sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij} \quad \text{any } j$$

5. The elem. matrix corresponding to $\left\{\begin{array}{l}\text{row replacement}\\ \text{row swap}\\ \text{row scale by } k\end{array}\right\}$ has det $\left\{\begin{array}{c}1\\ -1\\ k.\end{array}\right.$

$\rightsquigarrow$ can compute det via row ops.

3.3 CRAMER'S RULE

$A = n \times n$ matrix

$b = n \times 1$

$\rightsquigarrow$ $A_i(b)$ = matrix obtained by replacing ith col of A by b.

<u>Thm</u> (Cramer's rule) A = invertible $n \times n$ matrix, b in $\mathbb{R}^n$.

The solution to $Ax = b$ has

$$x_i = \frac{\det A_i(b)}{\det A} \qquad i = 1, \ldots, n.$$

why? Write $A = (c_1 \cdots c_n)$

Say $Ax = b$

Then
$$\begin{aligned} A \cdot I_i(x) &= A\,(e_1 \cdots x \cdots e_n) \\ &= (Ae_1 \cdots Ax \cdots Ae_n) \\ &= (c_1 \cdots b \cdots c_n) \\ &= A_i(b) \end{aligned}$$

$$\rightsquigarrow \det A \underbrace{\det I_i(x)}_{= x_i} = \det AI_i(x) = \det A_i(b)$$

$$x_i \det A$$

example. Solve using Cramer's rule:

$$x + z = 1$$
$$y + z = 1$$
$$x + y = 4$$

Inverse formula

Thm. If A is invertible, then

$$A^{-1} = \frac{1}{\det A}\begin{pmatrix} C_{11} & \cdots & C_{n1} \\ \vdots & \ddots & \vdots \\ C_{1n} & \cdots & C_{nn} \end{pmatrix} = \frac{1}{\det A}(C_{ji})$$

why? j^{th} col of A^{-1} is an x so $Ax = e_j$

$\rightsquigarrow$ ij-entry of A^{-1} is $\dfrac{\det A_i(e_j)}{\det A}$

But $\det A_i(e_j) = (-1)^{i+j} \det A_{ji} = C_{ji}$.

example. Find $\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}^{-1}$

Determinants, volumes, and linear transformations

Recall $\det A$ = volume...

So T_A takes the unit cube (vol = 1) to a shape of vol = $\det A$.

By linearity, smaller cubes (vol = V) go to ~~cubes~~ shapes of vol = $\det A \cdot V$

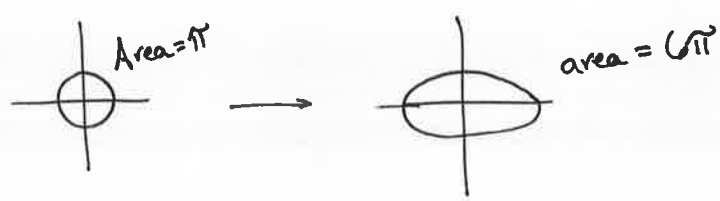
By Calculus, arbitrary shapes of vol V go to shapes of vol $\det A \cdot V$.

example 1. If you shear a sheep

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

area stays same!

example 2. ellipses $\begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}$



Population Growth

First half of course: $Ax = b$

Next up: $Ax = \lambda x$ "eigenvalues and eigenvectors"

Example. In a population of rabbits:

a) half of rabbits survive first year
of those, half survive second year
max life span is 3 yrs.

b) on ave, rabbits produce 0, 6, 8 rabbits in 1^{st}, 2^{nd}, 3^{rd} yrs.

Current age distribution is

$$v_0 = \begin{pmatrix} 24 \\ 24 \\ 20 \end{pmatrix} \quad \begin{array}{r} \text{age } 0 \\ 1 \\ 2 \end{array}$$

$\rightsquigarrow$ Age transition matrix: $\begin{pmatrix} 0 & 6 & 8 \\ 1/2 & 0 & 0 \\ 0 & 1/2 & 0 \end{pmatrix} = A$

After n years, age distr. is

$$A^n v_0$$

Want to find a stable age distribution

$\rightsquigarrow$ $Av_0 = \lambda v_0$ (same ratios)

Can you do with guess & check?

or Try $v_0 = \begin{pmatrix} 32 \\ 8 \\ 2 \end{pmatrix}$

5.1 EIGENVECTORS AND EIGENVALUES

$A = n \times n$ matrix

If $Av = \lambda v$ for some nonzero v in $\mathbb{R}^n$

and λ in $\mathbb{R}$

then v is called an eigenvector for A

and λ is the corresponding eigenvalue.

$\boxed{\text{eigen} = \text{characteristic}}$

examples. $A = \begin{pmatrix} 0 & 6 & 8 \\ 1/2 & 0 & 0 \\ 0 & 1/2 & 0 \end{pmatrix}$ $\quad V = \begin{pmatrix} 32 \\ 8 \\ 2 \end{pmatrix}$ $\quad \lambda = 2$

$A = \begin{pmatrix} 2 & 2 \\ -4 & 8 \end{pmatrix}$ $\quad V = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ $\quad \lambda = 4$

example. Check that $\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ are eigenvectors of $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$

example. Check that $\lambda = 3$ is an eigenvalue of

$$A = \begin{pmatrix} 2 & -4 \\ -1 & -1 \end{pmatrix}$$

How? Want $Av = 3v$

$$Av - 3v = 0$$
$$Av - 3Iv = 0$$
$$(A - 3I)v = 0$$

$\rightsquigarrow$ row reduce $\begin{pmatrix} -1 & -4 \\ -1 & -4 \end{pmatrix}$

This works for any λ. That is, λ is an eigenvalue if and only if

$$(A - \lambda I)v = 0$$

has a nontrivial solution.

Or $A - \lambda I$ is not invertible

Or $\det A - \lambda I = 0$.

The set of all solutions to $(A - \lambda I)v = 0$ is the eigenspace of A corresp. to λ.

These are all vectors in $\mathbb{R}^n$ that get scaled by λ.

example. Find eigenvectors, eigenvalues, and eigenspaces of $\begin{pmatrix} 5 & -6 \\ 3 & -4 \end{pmatrix}$ and draw the picture.

example. Find a basis for the $\lambda = 2$ eigenspace of

$$\begin{pmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{pmatrix}$$

<u>Thm</u>. The eigenvalues of a triangular matrix are the entries on the diagonal.

Why? What is the determinant of $A - \lambda I$?

Zero Eigenvalues

$\lambda = 0$ means $Ax = 0x$ has a nontrivial soln

$\rightsquigarrow$ same as saying A is not invertible.

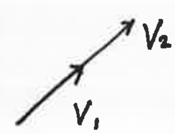
So: A invertible $\iff$ 0 is not an eigenvalue of A.

Distinct Eigenvalues

Thm. If $v_1, \dots, v_k$ are eigenvectors corresp. to distinct eigenvalues $\lambda_1, \dots, \lambda_k$ then $v_1, \dots, v_k$ are lin ind.

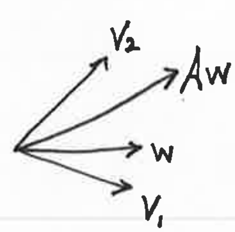
Why? Think about $k=2$.

If v_1, v_2 lin dep, obviously can't have different λ_1, λ_2



both get stretched the same.

And $k=3$:



Say $\lambda_1 = 1$
$\lambda_2 = 2$

$\implies$ no other eigenvectors in v_1v_2-plane

$\implies$ any other eigenvector must be lin ind.

More examples

Find the eigenvalues/eigenvectors without calculation

- $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ $\quad T_A = \mathrm{id}$
- $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ $\quad T_A =$ proj. to x-axis
- $A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ $\quad T_A =$ rot. by θ.
- $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ $\quad T_A =$ reflect about $y = x$
- $A = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$ $\quad T_A =$ stretch in x-dir
- $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ $\quad T_A =$ shear

Non-eigenvectors

What does A do to non-eigenvectors?

example. $\begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}$ almost all vectors get pulled to x-axis = eigenvector with largest eigenvalue.

$\rightsquigarrow$ similar to rabbit populations

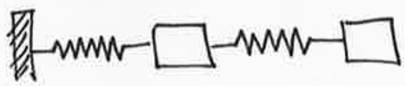
More exercises

$\begin{pmatrix} 2 & 7 \\ -1 & -6 \end{pmatrix}$ $\begin{pmatrix} 0 & 1 \\ -2 & -3 \end{pmatrix}$ $\begin{pmatrix} 10 & -9 \\ 4 & -2 \end{pmatrix}$

Structural Engineering

Tacoma Narrows Bridge – why is this an eigenvalue problem?

example: two masses on springs



This system has a natural frequency (actually two) where the masses move together in harmony. In other words, the position vector $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ changes by scaling only $\rightsquigarrow$ eigenvector

frequency $\longleftrightarrow$ eigenvalue.

if the wind, say, blows at that frequency, motion will get larger $\rightsquigarrow$ bridge collapse.

Preview of Diff Eq

Equation of Motion: $\underbrace{M}_{\text{mass matrix}}\ddot{u} + \underbrace{K}_{\text{stiffness matrix}}u = 0$

Assume $u = v\sin(\omega t)$

$$\rightsquigarrow\ -\omega^2 Mv\sin(\omega t) + Kv\sin(\omega t) = 0$$

$$\rightsquigarrow\ [K - \omega^2 M]v = 0$$

Eigenvectors and Difference Eqns

Say we want to solve $x_{k+1} = Ax_k$

in other words, need a whole seq $x_0, x_1, x_2, \ldots$

with $x_1 = Ax_0$, $x_2 = Ax_1$, etc.

example. $A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} b \\ a+b \end{pmatrix}$

What if we find an eigenvector v of A?

Then we claim we have a soln

$$\begin{array}{cccc} v, & \lambda v, & \lambda^2 v, & \ldots \\ \parallel & \parallel & \parallel & \\ x_0 & x_1 & x_2 & \end{array}$$

why? $Ax_k = A(\lambda^k v) = \lambda^k Av = \lambda^k \lambda v = \lambda^{k+1} v = x_{k+1}$.

What about our A?

$$\det A - \lambda I = \det \begin{pmatrix} -\lambda & 1 \\ 1 & 1-\lambda \end{pmatrix} = \lambda^2 - \lambda - 1$$

$\rightsquigarrow \lambda = \dfrac{1 \pm \sqrt{5}}{2}$ golden ratio

5.2 THE CHARACTERISTIC POLYNOMIAL

Recall: λ is an eigenval of A

$\Leftrightarrow (A-\lambda I)x = 0$ has non-0 soln

$\Leftrightarrow A-\lambda I$ not inv.

$\Leftrightarrow \det A-\lambda I = 0$.

Now: $\det A-\lambda I$ is a polynomial in λ (why?)

called the characteristic polynomial of A

$\rightsquigarrow$ its roots are the eigenvals of A.

e.g. the char poly of $\begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix}$ is $\lambda^2 - 6\lambda + 1$

$\rightsquigarrow$ eigenvals are $\dfrac{6 \pm \sqrt{32}}{2} = 3 \pm 2\sqrt{2}$

For any 2×2: char poly of $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is $\lambda^2 - \underbrace{(a+d)}_{\text{"trace"}}\lambda + \underbrace{(ad-bc)}_{\det}$.

<u>Algebraic Multiplicity</u>: The algebraic multiplicity of an eigenval λ is its multiplicity as a root of the char poly.

e.g. if char poly is $\lambda^4 - \lambda^2$ then the eigenvals are $0, 1, -1$ with alg. mult. $2, 1, 1$.

<u>Fact</u>. An $n\times n$ matrix has exactly n (complex) roots with multiplicity.

$\boxed{\text{CLICKER}}$ Find eigenvals (with alg mult) of the rabbit population matrix

$$\begin{pmatrix} 0 & 6 & 8 \\ 1/2 & 0 & 0 \\ 0 & 1/2 & 0 \end{pmatrix}$$

Similarity

A, B are similar if there is a C so that

$$A = CBC^{-1}$$

The idea is that A, B are essentially doing the same thing, but w.r.t. different bases.

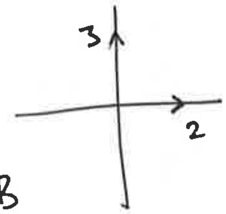
example. $B = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$ $A = \begin{pmatrix} 1 & 2 \\ -1 & 4 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$

Fact. If A, B similar then char poly A = char poly B.

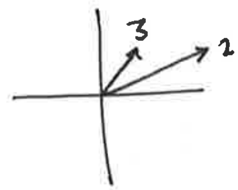
why?
$$\begin{aligned} \det(A - \lambda I) &= \det(CBC^{-1} - \lambda I) \\ &= \det(CBC^{-1} - C\lambda I C^{-1}) \\ &= \det(C(B - \lambda I)C^{-1}) \\ &= \det C \, \det(B - \lambda I) \, \det C^{-1} \\ &= \det(CC^{-1}) \det(B - \lambda I) \\ &= \det(B - \lambda I). \end{aligned}$$

In above example: char poly is $\lambda^2 - \underbrace{5}_{\text{trace}}\lambda + \underbrace{6}_{\text{det}}$.

In the example B has eigenvecs e_1, e_2 for eigenvals $2, 3$
A has eigenvects $\begin{pmatrix}2\\1\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix}$ for eigenvals $2, 3$



B



This makes sense because C^{-1} takes $\begin{pmatrix}2\\1\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix}$ to e_1, e_2 then B stretches e_1, e_2 then C takes e_1, e_2 back to $\begin{pmatrix}2\\1\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix}$.

__example__. Do a similar analysis of

$$B = \begin{pmatrix}1 & 0\\0 & -1\end{pmatrix} \qquad A = \begin{pmatrix}-1 & 0\\0 & 1\end{pmatrix} = \begin{pmatrix}0 & -1\\1 & 0\end{pmatrix}\begin{pmatrix}1 & 0\\0 & -1\end{pmatrix}\begin{pmatrix}0 & 1\\-1 & 0\end{pmatrix}$$

__More Char. Poly Probs__

$$\begin{pmatrix}2 & 3 & 1\\3 & 2 & 4\\0 & 0 & -1\end{pmatrix} \qquad \begin{pmatrix}2 & 1 & 1\\0 & 1 & 0\\1 & 0 & 1\end{pmatrix} \qquad \begin{pmatrix}1 & 1 & 1\\1 & 1 & 1\\1 & 1 & 1\end{pmatrix}$$

$$\begin{pmatrix}2 & 1 & 1\\1 & 1 & 0\\1 & 0 & 1\end{pmatrix}$$

Structural Engineering: Column Buckling

Say we have a column with a compressive force.
How exactly will the column buckle?

Idea: Approximate column by finite # of pts, say
$(0,0), (0,1), \ldots, (0,6)$.

Buckling $\rightsquigarrow$ $(0,0), (x_1,1), (x_2,2), \ldots, (x_5,5), (0,6)$
(roughly).

Engineers $\rightsquigarrow$ difference eqn

$$(x_{i-1} - 2x_i + x_{i+1}) + \lambda x_i = 0$$

Why is this reasonable?

(discrete version of $d^2x/dy^2 + \lambda x = 0$).

λ depends on the force and stiffness matrix.

$\rightsquigarrow$ [redacted] solve this diff. eqn as above. Need eigenvector of

$$\begin{pmatrix} 2 & -1 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & -1 & 2 \end{pmatrix}$$

For most λ, only eigenvector is $0 \Longrightarrow$ no buckling!
Above matrix has three eigenvals

$\lambda=1$ $\lambda=2$ $\lambda=3$

5.3 Diagonalization

We have seen that it is useful to take powers of matrices,
e.g. rabbit populations, diff. eqns.

If A is diagonal, A^k is easy to compute.

e.g. what is $\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}^{10}$?

What if A is not diagonal?

e.g. find $\begin{pmatrix} 1 & 2 \\ -1 & 4 \end{pmatrix}^{10}$. Not easy!

But we saw
$$\begin{pmatrix} 1 & 2 \\ -1 & 4 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$$
$$A \quad = \quad C \quad B \quad C^{-1}$$

So
$$A^{10} = (CBC^{-1})(CBC^{-1})\cdots(CBC^{-1})$$
$$= CB^{10}C^{-1}$$
$$= \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 2^{10} & 0 \\ 0 & 3^{10} \end{pmatrix}\begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} \qquad \text{easy!}$$

Upshot: The diagonalization of A $(A = CBC^{-1})$ is useful!

Diagonalization $\quad A = n\times n$ matrix

A is diagonalizable if it is similar to a diagonal matrix:

$$A = CDC^{-1} \qquad D \text{ diagonal.}$$

Thm. $A = n\times n$ matrix

A is diagonalizable $\iff$ A has n lin. ind. evectors $v_1, \ldots, v_n$

In this case
$$A = \underbrace{(v_1 \cdots v_n)}_{C} \underbrace{\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}}_{D} \underbrace{(v_1 \cdots v_n)^{-1}}_{C^{-1}}$$

λ_i = evalue for v_i.

why? $(v_1 \cdots v_n)^{-1}$ takes each v_i to e_i

D stretches each e_i by λ_i

C takes the e_i back to v_i

so: net effect is stretching each v_i by λ_i.

examples. Diagonalize if possible.

$$\begin{pmatrix} 2 & 6 \\ 0 & -1 \end{pmatrix} \quad \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix} \quad \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \quad \begin{pmatrix} a & b \\ b & a \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 2 & 0 & 0 \\ 1 & 2 & 1 \\ -1 & 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 2 & 4 & 6 \\ 0 & 2 & 2 \\ 0 & 0 & 4 \end{pmatrix}$$

remember: for 3×3 matrices, you often need to guess eigenvalues if you can't factor.

Q. Use your diagonalization of $\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$ to find a formula for the n^{th} Fibonacci number.

$\boxed{\text{CLICKER}}$ Which are true?

① If A is diagonalizable then A^2 is.
② If A is diagonalizable then A^{-1} is.
③ If A^2 is diag. then A is.

<u>Fact</u>. If A has n distinct eigenvals then A is diagonalizable. why?

<u>Non-distinct Eigenvals</u>

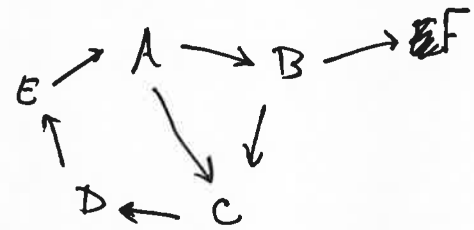
$A = n \times n$ with eigenvals $\lambda_1, \ldots, \lambda_k$
$a_i =$ alg. mult. of λ_i
$d_i =$ dim of λ_i eigenspace

① $d_i \leq a_i$ all i

② A is diagonalizable $\Leftrightarrow \sum d_i = n$
$\Leftrightarrow d_i = a_i$ for all i and char poly has n real roots

③ If A is diagonalizable, the basis vectors for the eigenspaces give a basis for $\mathbb{R}^n$.

Application: Social Networks

Consider a social network  etc.

Want to find communities, say, a group of people so there is a directed path connecting any two.

Make a matrix M ij-entry is # arrows from i to j.

Then the ij-entry of M^2 is # paths of length 2 from i to j.   why?

Similar for M^3, etc.. So ij-entry of

$$M + \cdots + M^k$$

is # paths of length at most k.

$\rightsquigarrow$ look for positive minors.

Leading eigenvalue is a measure of how connected the network is.

APPLICATION : BUSINESS

Say your rental car business has 3 locations
Make a matrix M ij-entry is probability that
a car at location i ends at location j

eg.

$$M = \begin{pmatrix} .3 & .4 & .5 \\ .3 & .4 & .3 \\ .4 & .2 & .2 \end{pmatrix}$$

Note cols add to 1
again

Eigenvector with eigenvalue 1 is steady state $\begin{pmatrix} .3\bar{8} \\ .3\bar{3} \\ .2\bar{7} \end{pmatrix}$
Any other vector gets pulled to this state.
Applying powers of M gives the state
after some number of iterations.

Why is this similar/same as Google?

CHAPTER 3 IN A NUTSHELL

The determinant is a function

$$\det : \{n \times n \text{ matrices}\} \longrightarrow \mathbb{R}$$

It has several formulas:

$$\det A = \sum_{j=1}^{n} a_{ij} C_{ij} \quad \text{any } i \qquad C_{ij} = (-1)^{i+j} \det A_{ij}$$

"cofactor"

$$= \sum_{i=1}^{n} a_{ij} C_{ij} \quad \text{any } j \qquad A_{ij} = ij^{\text{th}} \text{ minor}$$

So: A triangular $\Rightarrow \det A$ = product of diag entries

The determinant behaves nicely under row ops:

$$\text{If we perform} \left\{ \begin{array}{l} \text{row repl.} \\ \text{row swap} \\ \text{row scale by } k \end{array} \right\} \text{det changes by a factor} \left\{ \begin{array}{l} +1 \\ -1 \\ k \end{array} \right.$$

This should be the <u>defn</u> of det!

Consequences:
1. Can compute det by row red.
2. A inv $\Leftrightarrow \det A \neq 0$.
3. $\det AB = \det A \det B$
4. $\det A$ = signed volume of parallelpiped spanned by rows of A
5. the cofactor formula holds.

Cramer's Rule: A invertible $\Rightarrow$ solns to $Ax=b$ has

$$x_i = \frac{\det A_i(b)}{\det A} \qquad A_i(b) = \text{replace } i^{th} \text{ col of } A \text{ by } b.$$

$$\rightsquigarrow A^{-1} = \frac{1}{\det A}(C_{ij})^T$$

Linear transformations: $A = n \times n$ matrix

$S =$ region in $\mathbb{R}^n$ with finite volume

$$\rightsquigarrow \text{vol } T_A(S) = \det A \cdot \text{vol}(S).$$

CHAPTER 5 IN A NUTSHELL

If A is a matrix, v a nonzero vector, λ a number with

$$Av = \lambda v$$

then v is an eigenvector for A with eigenvalue λ.

<u>Facts</u> ① A inv $\iff$ 0 is not an eigenval of A

② Eigenvects for distinct eigenvals are lin ind.

Difference Eqns: If v is an eigenvector for A then

$x_{k+1} = Ax_k$ has soln $x_k = \lambda^k v$

Finding eigenvalues: Solve $\underbrace{\det A - \lambda I}_{\text{characteristic poly.}} = 0$

(when $n > 2$, need to be clever!)

Finding eigenvectors/eigen<u>spaces</u>: Given λ solve

$$(A - \lambda I)x = 0.$$

Why? Finding eigenvalues & eigenspaces gives a concrete description of what A is <u>doing</u> to $\mathbb{R}^n$.

<u>Similarity</u> A, B are similar if $A = CBC^{-1}$ some C

<u>Fact</u>. Similar matrices have <u>same</u> eigenvalues, cor<u>respond</u>ing eigenspaces.

So: similar matrices <u>do</u> same thing to $\mathbb{R}^n$, just with respect to different bases.

Diagonalization A matrix A is **diagonalizable** if it is similar to a diagonal matrix D: $A = CDC^{-1}$.

$\rightsquigarrow$ can take powers $A^n = CD^nC^{-1}$

Thm. A is diagonalizable $\iff$ A has n lin. ind. eigenvectors
$n \times n$

If the eigenvects are $v_1, \dots, v_n$ & **corresp.** evals are $\lambda_1, \dots, \lambda_n$ then $A = CDC^{-1}$ where $C = (v_1 \cdots v_n)$

$D = \text{diag}(\lambda_1, \dots, \lambda_n)$

So: if A has n distinct evals, it is diag'able.

Thm. If λ is an eigenvalue of a matrix:

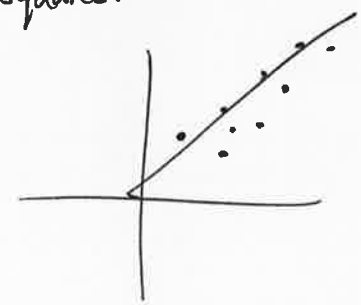
dim of λ-eigenspace $\leq$ alg. mult. of λ

Applications Google, rental cars, bunnies, social networks, column buckling, natural frequencies

6 ORTHOGONALITY

Main goal: solve $Ax = b$ as close as possible
(if no actual solution)
$\rightsquigarrow$ method of least squares.

Applications: Linear regression



Plus: facial recognition, image compression,
etc...

6.1 INNER PRODUCTS

Dot Product $\quad u, v$ in $\mathbb{R}^n$ (col. vectors)

$$u \cdot v = u^T v$$
$$= \sum u_i v_i$$

Some properties: $u \cdot v = v \cdot u$

$(u+v) \cdot w = u \cdot w + v \cdot w$

$(cu) \cdot v = c(u \cdot v)$

$u \cdot u \geq 0 \quad (u \cdot u = 0 \Leftrightarrow u = 0)$.

Length $\quad v$ in $\mathbb{R}^n$

$\|v\| = \sqrt{v \cdot v}$ $\quad$ length (or norm) of v

why? Pythagorean thm!

Fact. $\|cv\| = |c| \cdot \|v\|$

v is a unit vector if $\|v\| = 1$.

Q. Find the unit vector in the direction of $(1,2,3,4)$

Distance $\quad u, v$ in $\mathbb{R}^n$

$\text{dist}(u,v) = \|u-v\| = \|v-u\|$

Q. Compute the dist. b/w $(1,1,1)$ and $(1,4,-3)$.

Orthogonality u, v in $\mathbb{R}^n$

Fact. $u \perp v \iff u \cdot v = 0$

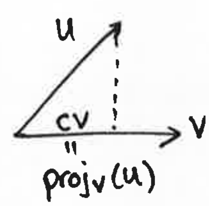
why? $u \perp v \iff \|u\|^2 + \|v\|^2 = \|u - v\|^2$

$$\iff u \cdot u + v \cdot v = u \cdot u - 2u \cdot v + v \cdot v$$

$$\iff u \cdot v = 0.$$

Q. Find a vector $\perp$ to $(1,2,3)$.

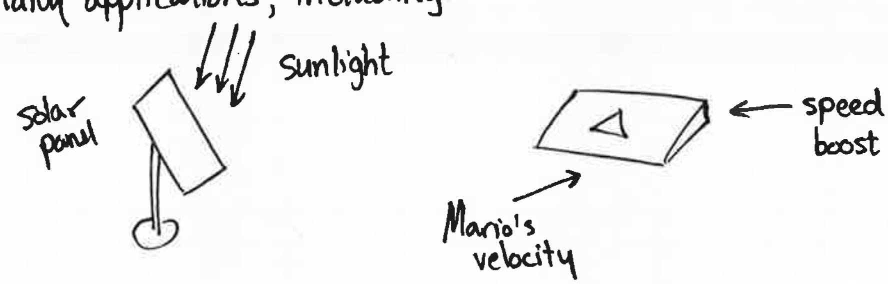
Projections u, v in $\mathbb{R}^n$



Fact. $\mathrm{proj}_v(u) = \frac{u \cdot v}{v \cdot v} v$

why? $(u - cv) \cdot v = 0 \rightsquigarrow c = \frac{u \cdot v}{v \cdot v}$

Many applications, including:

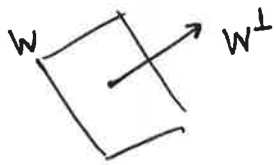


Orthogonal complements

W = subspace of $\mathbb{R}^n$

$W^\perp = \{ v \text{ in } \mathbb{R}^n : \ v \cdot w = 0 \text{ for all } w \text{ in } W \}$

e.g.



<u>Facts</u>. ① $W^\perp$ is a subspace

② $(W^\perp)^\perp = W$

③ $\dim W + \dim W^\perp = n$

④ If $W = \text{span}\{w_1, \ldots, w_k\}$

then $W^\perp = \{ v \text{ in } \mathbb{R}^n : \ v \cdot w_1 = \cdots = v \cdot w_k = 0 \}$

Q. What is $W^\perp$ if $W = \text{span}\{e_1, e_2\}$ in $\mathbb{R}^3$?

What about $W = \text{span}\left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix} \right\}$

<u>Thm</u>. $A = m \times n$ matrix

$(\text{Row } A)^\perp = \text{Nul } A \qquad (\text{or } (\text{Col } A)^\perp = \text{Nul } A^T)$

why? $Ax = 0$ same as $x \perp$ all rows of A!

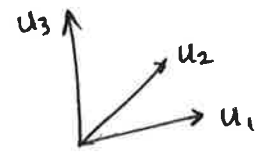
So $\text{span}\left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix} \right\}^\perp = \text{Col}\begin{pmatrix} 1 & -1 \\ 1 & 2 \\ 1 & -1 \end{pmatrix}^\perp = \text{Nul}\begin{pmatrix} 1 & 1 & 1 \\ -1 & 2 & -1 \end{pmatrix} = \text{span}\left\{ \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \right\}$

6.2 Orthogonal Sets

A set of vectors is orthogonal if each pair is

<u>example</u>. $\left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \right\} = B$ check pairwise dot products.

<u>Fact.</u> An orthog. set of nonzero vectors is lin ind.



why? suppose not: $c_1u_1 + c_2u_2 + c_3u_3 = 0$
dot both sides with u_1: $c_1\, u_1 \cdot u_1 = 0 \Rightarrow c_1 = 0$.

<u>Orthogonal bases</u>

orthog. basis = basis that is orthog.

<u>Thm.</u> Say $\{u_1, \ldots, u_k\}$ is an orthog. basis for subspace W of $\mathbb{R}^n$
If y in W then $y = \sum c_i u_i$ where

$$c_i = \frac{y \cdot u_i}{u_i \cdot u_i} = \text{length of proj of } y \text{ to span}\{u_i\}$$

why? $y \cdot u_i = (c_1u_1 + \cdots + c_ku_k) \cdot u_i = c_i\, u_i \cdot u_i$

<u>example</u>. Find B-coords of $(6, 1, -8)$ (B as above)

(Much quicker than solving $Ax = b$!).

Components of a vector

Say $L = \text{span}\{u\}$ in $\mathbb{R}^n$

Given y in $\mathbb{R}^n$ want to decompose it to

$$\underbrace{y = y_L}_{\text{parallel to } L} + \underbrace{y_{L^\perp}}_{\text{orthog. to } L.}$$

How? $y_L = \text{proj}_L\, y = \dfrac{y \cdot u}{u \cdot u} u$

$\rightsquigarrow y_{L^\perp} = y - y_L$

example. Wind vector is $\begin{pmatrix} 2 \\ 3 \end{pmatrix}$ (northwest)

what is the force applied to a train car on a track that has slope 2?

example. $u = (1,1,1)$, $y = (6,1,-8)$.

$$y = -\frac{1}{3}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + -\frac{2}{3}\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} + 7\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$$

from above calculation

$$\rightsquigarrow y = \begin{pmatrix} -1/3 \\ -1/3 \\ -1/3 \end{pmatrix} + \begin{pmatrix} 19/3 \\ 4/3 \\ -23/3 \end{pmatrix}$$

Orthonormal sets & matrices

orthonormal set = orthog. set where each vector is unit.

Q. How to turn an orthog. set to an orthon. one?

Fact. $U = m\times n$ matrix w. orthonormal cols

$$\Rightarrow U^T U = I_n$$

Also: $Ux \cdot Uy = x\cdot y$ any x, y.

In particular: $\|Ux\| = \|x\|$

$x \perp y \Leftrightarrow Ux \perp Uy$

$\{x_1, \dots, x_k\}$ orthon. $\Leftrightarrow \{Ux_1, \dots, Ux_k\}$ is

example $\begin{pmatrix} 1/3 & -2/3 & 2/3 \\ 2/3 & -1/3 & -2/3 \\ 2/3 & 2/3 & 1/3 \end{pmatrix}$

If $A = n\times n$ matrix with orthon. cols
we say A is orthogonal.

orthogonal matrices $\longleftrightarrow$ rotations

why?

6.3 Orthogonal Projection

Last time: $y = y_L + y_{L^\perp}$ $\qquad$ $\dim L = 1$

This time $y = y_W + y_{W^\perp}$ $\qquad$ $\dim W =$ anything.

Recall: $y_L = \frac{y \cdot u}{u \cdot u} u \qquad L = \text{span}\{u\}$

<u>Thm.</u> $W =$ subsp. of $\mathbb{R}^n$

y in $\mathbb{R}^n$

Then can write y uniquely as $y = y_W + y_{W^\perp}$

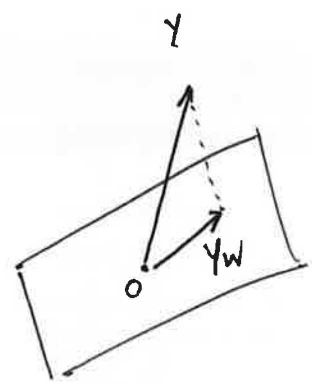
where y_W in W, $y_{W^\perp}$ in $W^\perp$

Moreover if $\{u_1, \dots, u_k\}$ is orthog. basis for W then

$$y_W = \frac{y \cdot u_1}{u_1 \cdot u_1} u_1 + \cdots + \frac{y \cdot u_k}{u_k \cdot u_k} u_k$$

Why does this make sense?

y_W is proj to W of ~~closest to~~ y



<u>example.</u> $y = (1,0,0)$

$W = \text{span}\left\{ \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}$

$\rightsquigarrow$ Find y_W.

Q. Find the matrix A corr. to orthog. proj to W

To do this, project e_1, e_2, e_3 to W. Those are cols of A.

$\boxed{\text{CLICKER}}$ Without calculation, what is A^2?

a. I b. $-I$ c. O d. A e. $-A$

$\boxed{\text{CLICKER}}$ Without calculation, what are eigenvals of A?

a. 0 b. 1 c. -1 d. 2

<u>Special Case 1 of Thm</u>: If $\{u_1, \ldots, u_k\}$ orthonormal then

$$y_W = (y \cdot u_1)u_1 + \cdots + (y \cdot u_k)u_k.$$
$$= UU^T y \quad \text{(why?)}.$$

<u>Special Case 2 of Thm</u>: If $\{u_1, \ldots, u_n\}$ is orthog. basis for $\mathbb{R}^n$

& $W = \text{Span}\{u_1, \ldots, u_k\}$

any $y = c_1u_1 + \cdots + c_nu_n$

Then $y_W = c_1u_1 + \cdots + c_ku_k$.

<u>Special Case 3 of Thm</u> If y in W then $y_W = y$.

Best Approximation

W = subsp. of $\mathbb{R}^n$

y_W = closest pt in W to y

= proj of y to W

= W-part of y

Fact. $\|y - y_W\| < \|y - w\|$ any w in W, $w \neq y_W$

why? $y - w = \underbrace{(y - y_W)}_{\text{in } W^\perp} + \underbrace{(y_W - w)}_{\text{in } W}$

$\rightsquigarrow \|y-w\|^2 = \|y - y_W\|^2 + \|y_W - w\|^2$

$\Rightarrow \|y-w\|^2 > \|y - y_W\|^2$

Q. Find dist. from e_1 to span $\left\{ \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}$

Will use best approx. to find best soln to $Ax = b$ even when there isn't an actual soln.

Image Compression

Say you have an 8×8 greyscale image
$\rightsquigarrow$ 8×8 matrix A.

Look at any row: $x_1 \; x_2 \; \cdots \; x_7 \; x_8$

Replace it with: $\underbrace{\frac{x_1+x_2}{2} \;\; \frac{x_3+x_4}{2} \;\; \cdots \;\; \frac{x_7+x_8}{2}}_{\substack{\text{approximation coeffs} \\ \text{(averages)}}} \;\; \underbrace{\frac{x_1-x_2}{2} \;\; \cdots \;\; \frac{x_7-x_8}{2}}_{\text{detail coeffs}}$

or compute AW_1 where

$$W_1 = \begin{pmatrix} 1/2 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 \\ 1/2 & 0 & 0 & 0 & -1/2 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 0 & 0 & 1/2 & 0 & 0 \\ 0 & 1/2 & 0 & 0 & 0 & -1/2 & 0 & 0 \\ 0 & 0 & 1/2 & 0 & 0 & 0 & 1/2 & 0 \\ 0 & 0 & 1/2 & 0 & 0 & 0 & -1/2 & 0 \\ 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 1/2 \\ 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & -1/2 \end{pmatrix}$$

orthogonal!

Key point: if $x_1 = x_2$ (roughly) get a zero detail coeff (roughly)

Do same to cols: $W_1^T A W_1$
$\rightsquigarrow$ upper-left 4×4 minor records averages over 2×2 grids.

Now repeat the process on upper left 4×4 minor,
then the upper left 2×2 minor:

$$W_3^T W_2^T W_1^T A W_1 W_2 W_3$$

$$= W^T A W \qquad W = W_1 W_2 W_3 \text{ orthogonal!}$$

When you download an image, you first get the 2×2 approx coeffs, then 4×4, etc. until you get the whole image. That's why you see successively finer approximations.

To decode use the fact that $W^T = W^{-1}$
(well, you need to replace W with an ortho<u>normal</u> version).

Orthonormality reduces distortion, e.g. lengths are preserved.

6.4. THE GRAM-SCHMIDT PROCESS

Idea: want to convert bases for subspaces into orthogonal ones, e.g. for purposes of projecting.

example 1. Find an orthog. basis for span$\{x_1, x_2\}$ where $x_1 = (1,1,0)$, $x_2 = (1,1,1)$.

Set $v_1 = x_1$

$v_2 = x_2 - \text{proj}_{\text{span}\{v_1\}}(x_2)$

$= (0,0,1)$ ✓

example 2. Find an orthog. basis for span$\{x_1, x_2, x_3\}$ x_1, x_2 above $x_3 = (3,1,1)$.

Set v_1, v_2 as above, so just need to fix up x_3:

$v_3 = x_3 - \text{proj}_{\text{span}\{v_1, v_2\}}(x_3)$

$= (1,-1,0)$

<u>Thm</u> (Gram-Schmidt process) Say $\{x_1, \ldots, x_k\}$ is a basis for a nonzero subspace W of $\mathbb{R}^n$. Inductively define:

$v_1 = x_1$

$v_i = x_i - \text{proj}_{\text{span}\{v_1, \ldots, v_{i-1}\}}(x_i)$ $\quad 2 \le i \le k$.

Then $\{v_1, \ldots, v_k\}$ is an orthog. basis for W.

Q. Find orthog. basis for span of $(1,1,1,1)$, $(-1,4,4,-1)$, $(4,-2,2,0)$.

QR Factorizations

Thm. $A = m\times n$ matrix with lin ind cols

$\rightsquigarrow$ $A = QR$ where Q has orthonormal cols

R is upper triang with pos. diag entries

Method 1. Q obtained from Gram-Schmidt vectors & normalization.

$R = Q^TA$ This works since:

$$Q^TA = Q^TQR = IR = R$$

For above example $A = \begin{pmatrix} 1 & 1 & 3 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}$ have

$Q = \begin{pmatrix} 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 0 & 1 & 0 \end{pmatrix}$ $R = Q^TA$ need to multiply out.

Method 2. Q obtained from GS

R records the operations used in GS

(like how L records operations in row reduction)

Above example: $Q = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}$ $R = \begin{pmatrix} 1 & \boxed{1} & \boxed{2} \\ 0 & 1 & \boxed{1} \\ 0 & 0 & 1 \end{pmatrix}$ check $QR = A$

The first $\boxed{1}$ comes from: $V_2 = X_2 - V_1$

The other $\boxed{2}$ & $\boxed{1}$ come from: $V_3 = X_3 - 2V_1 - V_2$

This isn't really a proper QR decomp because Q has orthog. cols, not orthonormal. Get a real QR factorization by scaling cols of Q by $1/\sqrt{2}, 1, 1/\sqrt{2}$

& rows of R by $\sqrt{2}, 1, \sqrt{2}$ why does this work?

QR Method for finding eigenvalues

A = $n \times n$ matrix

do $A = Q_1 R_1$ QR factorization

$A_1 = R_1 Q_1$ swap Q and R

$= Q_2 R_2$ and find QR factorization of result

$A_2 = R_2 Q_2$ swap, etc...

The A_k converge to an upper Δ matrix and the diag entries converge to eigenvals of A.

Why? The first thing to note is that each A_k is similar to A (hence same eigenvals). Indeed:

$$A_1 = R_1 A R_1^{-1} \quad \text{so } A_1 \sim A$$
$$A_2 = R_2 A_1 R_2^{-1} \quad \text{so } A_2 \sim A_1 \sim A$$

So all A_k have same eigenvals. Only thing to check is the convergence.

Idea of convergence. Suppose eigenvals are $\lambda_1 > \lambda_2 > \dots > \lambda_n$

You show that $a_{ii} \longrightarrow \lambda_i$

$$\text{and } |a_{ij}| \approx |\lambda_i / \lambda_j|^k \quad i \geq j+1$$

goes to 0

where a_{ij} here are entries of A_k.

6.5 Least Squares Problems

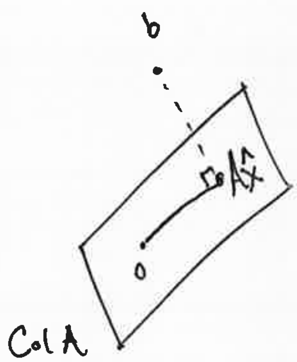
$A = m \times n$. A least squares soln to $Ax = b$ is an $\hat{x}$ in $\mathbb{R}^n$ with

$$\|b - A\hat{x}\| \le \|b - Ax\|$$

for all x in $\mathbb{R}^n$

Thm. The least squares solns to $Ax=b$ are the solns to

$$(A^T A)x = (A^T b)$$



why? By Best Approx. Thm: $A\hat{x} = \text{proj}_{\text{Col}A}(b)$

$\rightsquigarrow b - A\hat{x} \perp \text{Col } A$

$\rightsquigarrow b - A\hat{x} \perp$ each col of A

$\rightsquigarrow A^T(b - A\hat{x}) = 0$

$\rightsquigarrow A^T b - A^T A\hat{x} = 0$

$\rightsquigarrow (A^T A)\hat{x} = A^T b$

examples • $A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{pmatrix}$ $b = \begin{pmatrix} 6 \\ 0 \\ 0 \end{pmatrix}$ $\quad$ $A = \begin{pmatrix} 2 & 0 \\ -1 & 1 \\ 0 & 2 \end{pmatrix}$ $b = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$

Thm. $A = m \times n$. TFAE

a. $Ax = b$ has a unique least sq. soln for all b in $\mathbb{R}^m$

b. cols of A are lin ind

c. $A^T A$ invertible

In this case, l.s.s. is $(A^T A)^{-1} A^T b$

(Why isn't this same as $A^{-1} b$?)

Application: Best-fit lines

Q. Find best-fit line through $(0,6)$, $(1,0)$, $(2,0)$

Need m, b so $y = mx + b$ equal/close to these pts.

$$\rightsquigarrow \begin{aligned} 6 &= m \cdot 0 + b \\ 0 &= m \cdot 1 + b \\ 0 &= m \cdot 2 + b \end{aligned}$$

$$\rightsquigarrow \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} b \\ m \end{pmatrix} = \begin{pmatrix} 6 \\ 0 \\ 0 \end{pmatrix}$$ inconsistent!

Least squares soln: $(5, -3)$ $\rightsquigarrow$ $y = -3x + 5$

$\boxed{\text{CLICKER}}$ What does this line minimize?

a. the sum of the squares of the distances from the data pts to the line

b. ... vertical distances...

c. ... horizontal distances...

d. ... maximal distance...

QR Method for Least Squares

$A = m \times n$, lin ind cols

$\rightsquigarrow A = QR$

Then the least sq. soln to $Ax = b$ is

$$\hat{x} = R^{-1}Q^T b$$

why? $A\hat{x} = QR\hat{x} = QRR^{-1}Q^T b = QQ^T b$

$$= \text{proj}_{\text{Col}\,Q}(b)$$

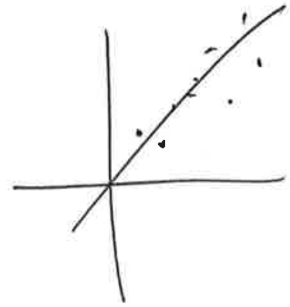
$$= \text{proj}_{\text{Col}(A)}(b)$$

In our example: $\begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 1/\sqrt{3} & -1/\sqrt{2} \\ 1/\sqrt{3} & 0 \\ 1/\sqrt{3} & 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} \sqrt{3} & \sqrt{3} \\ 0 & \sqrt{2} \end{pmatrix}$

This method is nice <u>if</u> you already have $A = QR$.

Chapter 6 in a Nutshell

Big goal: Solve $Ax=b$ as close as possible



Dot products $u \cdot v = 0 \iff u \perp v$

$\iff \|u\| + \|v\| = \|u-v\|$

Projections $\{u_1, \dots, u_k\}$ = orthog. basis for a subsp. W of $\mathbb{R}^n$

$\rightsquigarrow \text{proj}_W(v) = \frac{v \cdot u_1}{u_1 \cdot u_1} u_1 + \dots + \frac{v \cdot u_k}{u_k \cdot u_k} u_k$

If w in W: $\|v - w\| \geq \|v - \text{proj}_W(v)\|$

Orthogonal Matrices U = $n \times n$ matrix is orthogonal if cols are orthonormal.

TFAE: ① U is orthog.

② $U^T U = I$

③ U preserves dot products

If U is $m \times n$ and cols form orthonormal basis for subsp W of $\mathbb{R}^n$ then $UU^T v = \text{proj}_W(v)$.

Gram-Schmidt Process If $\{w_1, \dots, w_k\}$ is any basis for subsp W of $\mathbb{R}^n$ get an orthog. basis by:

$v_1 = u_1$

$v_2 = u_2 - \text{proj}_{\text{span}\{v_1\}}(u_2)$

$\vdots$

$v_k = u_k - \text{proj}_{\text{span}\{v_1, \dots, v_{k-1}\}}(u_k)$

QR Factorization $A = m \times n$ has lin ind cols

$\rightsquigarrow A = QR$ $\quad$ Q has orthonorm. cols

R upper triang & pos. entries on diag,

- Cols of Q are result of Gram-Schmidt + normalization.
- Entries of R record steps of Gram-Schmidt + normalization.

e.g.
$$A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 1 \\ 1 & 6 \end{pmatrix} \underset{\text{Gram-Schmidt}}{=} \begin{pmatrix} 1 & -2 \\ 1 & -1 \\ 1 & -1 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} 1 & +2 \\ 0 & 1 \end{pmatrix} \underset{\text{normalization}}{=} \begin{pmatrix} 1/2 & -2/\sqrt{22} \\ 1/2 & -1/\sqrt{22} \\ 1/2 & -1/\sqrt{22} \\ 1/2 & 4/\sqrt{22} \end{pmatrix} \begin{pmatrix} 2 & 4 \\ 0 & \sqrt{22} \end{pmatrix}$$

Least Squares Solns A l.s.s. of $Ax = b$ is $\hat{x}$ in $\mathbb{R}^n$ s.t.

$\|b - A\hat{x}\| \leq \|b - Ax\|$ all x in $\mathbb{R}^n$.

Method 1 Solve $A^TAx = A^Tb$

Method 2 If $A = QR$ then $\hat{x} = R^{-1}Q^Tb$

7.1 Diagonalization of Symmetric Matrices

A is <u>symmetric</u> if $A = A^T$

<u>Fact</u>. If A is symm then its eigenspaces are orthogonal

why? Say v_1 is a λ_1 evector
v_2 is a λ_2 evector & $\lambda_1 \neq \lambda_2$

Want $v_1 \cdot v_2 = 0$

Have $\lambda_1 v_1 \cdot v_2 = (\lambda_1 v_1)^T v_2 = (Av_1)^T v_2 = v_1^T A v_2$
$= v_1^T(\lambda_2 v_2) = \lambda_2 v_1 \cdot v_2$
$\Rightarrow (\lambda_1 - \lambda_2)\, v_1 \cdot v_2 = 0 \Rightarrow v_1 \cdot v_2 = 0.$

example. diagonalize $\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$

More is true!

A matrix A is <u>orthogonally diagonalizable</u> if $A = CDC^{-1}$
where C is orthogonal $(C^{-1} = C^T)$ and D is diag.

<u>Thm</u>. $A = n \times n$ matrix is orthog. diag'able $\Leftrightarrow$ A is symmetric.

So just by looking, $\begin{pmatrix} 1 & 4 & 5 \\ 4 & 2 & 6 \\ 5 & 6 & 3 \end{pmatrix}$ is diagonalizable!

why? one direction is easy.

Say A is orthog. diag'able : $A = CDC^{-1}$

$$A^T = (C^{-1})^T D^T C^T$$
$$= CDC^{-1} = A$$

for other direction, need more... take the next course!

Spectral Decomposition

Say A is orth. diag'able. We just saw A is symmetric, but more is true.

$$A = CDC^{-1} = CDC^T = (v_1 \cdots v_n)\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}\begin{pmatrix} v_1^T \\ \vdots \\ v_n^T \end{pmatrix}$$

$$= (\lambda_1 v_1 \cdots \lambda_n v_n)\begin{pmatrix} v_1^T \\ \vdots \\ v_n^T \end{pmatrix}$$

$$= \lambda_1 v_1 v_1^T + \cdots + \lambda_n v_n v_n^T \quad \text{why?}$$

Each $\lambda_i v_i v_i^T$ is a symm. matrix and $v_i v_i^T$ is projection to span$\{v_i\}$.

Q. Find spectral decomp. of $\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$.