Linear Algebra (Leon, Steven)

§1.1 Introduction: Systems of Linear Equations

What is linear algebra? The road not taken.

motivation — Let's return to junior high school: solving simultaneous eq'ns.

Ex: $3x + 5y = 6$
$2x + 2y = 4$

Problem was to find $x, y$ satisfying both equations.

Typically: $x + y = 2 \Rightarrow y = 2 - x$
$\Rightarrow 3x + 10 - 5x = 6$
$\Rightarrow -2x = -4;\ x = 2.$
$y = 0.$

the curious child — [Maybe you as a kid start wondering...]
What about $3x + 5y = 6$
$6x + 10y = 10$ ?

What about $3x + 5y = 6$
$6x + 10y = 12$ ?

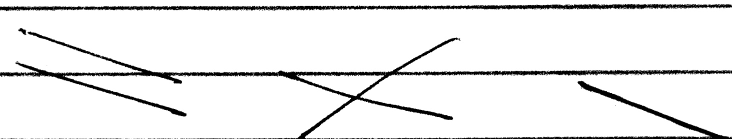
~~In the first case,~~ $\cancel{3x + 5y = 5 \Rightarrow 5y = 5 - 3x;\ 5 = 6.}$ ✗

Second case, $3x + 5y = 6$
$3x + 5y = 6$ $\Bigg\} \Rightarrow ?$ $x + \frac{5}{3}y = 2;$
any $(x,y)$ on this line will do (∞ly many sol'ns)

~~contrary to teaching~~ — Point: Could have no solutions (inconsistent)
one solution
infinitely many } (consistent)

What's really going on?

geometric interpretation — Geometric interpretation.




the curious child again — Another question you might have: What happens in higher dimensions?

Han dynasty problem

Ex: $3x + 2y + z = 39$

$2x + 3y + z = 34$

$x + 2y + 3z = 26$

How does one find the sol'n set for this sort of thing? (define)

Clever(?) observations:

legal moves

1. Changing the order of eq'ns doesn't change the sol'n set
2. Multiplying an eq'n by a nonzero number doesn't change...
3. Adding one eq'n to another in the system " " ...

(Why not? (#3) Well, any sol'n of the original will be a sol'n of the new, so sol'n(new) $\supset$ sol'n(old).)

equivalent systems

Defn: If two systems of linear eq'ns have the same sol'n sets, we say they are equivalent.

Rk: The above operations produce equivalent systems.
So we can [hope to] use them to simplify the system to one whose sol'n set is obvious (and the same as the original)

simplifying to triangular form

$$\begin{aligned} 3x+2y+z&=39\\ 2x+3y+z&=34\\ x+2y+3z&=26 \end{aligned} \sim \begin{aligned} x+2y+3z&=26\\ 2x+3y+z&=34\\ 3x+2y+z&=39 \end{aligned} \sim \begin{aligned} x+2y+3z&=26\\ -y-5z&=-18\\ -4y-8z&=-39 \end{aligned}$$

$$\sim \begin{aligned} x+2y+3z&=26\\ -y-5z&=-18\\ 12z&=33 \end{aligned}$$

; and this system is trivial:

$z = 33/12$;

$-y - 165/12 = -18$

$\Rightarrow y = 51/12$

$x = 117/12$

Terminology: solving by "back substitution."

## More terminology:

1. "m×n linear system" (m eq'ns, n unknowns) ← [overdetermined: m > n; under...]
2. "square system"
3. strictly triangular system (in $k^{th}$ eq'n, first $k-1$ coefficients are zero, $k^{th}$ coefficient $\neq 0$)

~~Ex~~ Observe that any (square) strictly triangular system has a unique soln (via back substitution)

Book keeping

Another mediocre observation: could just use the numbers

$$\left[\begin{array}{ccc|c} 3 & 2 & 1 & 39 \\ 2 & 3 & 1 & 34 \\ 1 & 2 & 3 & 26 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 2 & 3 & 26 \\ 2 & 3 & 1 & 34 \\ 3 & 2 & 1 & 39 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 2 & 3 & 26 \\ 0 & -1 & -5 & -18 \\ 1 & 2 & 3 & 39 \end{array}\right] \sim \ldots$$

$$\sim \left[\begin{array}{ccc|c} 1 & 2 & 3 & 26 \\ 0 & -1 & -5 & -18 \\ 0 & -4 & -8 & -39 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 2 & 3 & 26 \\ 0 & 1 & 5 & 18 \\ 0 & -4 & -8 & -39 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 2 & 3 & 26 \\ 0 & 1 & 5 & 18 \\ 0 & 0 & 12 & 33 \end{array}\right]$$

## More terminology:

1. coefficient matrix, augmented mx. for a linear system
2. elementary row operations
3. pivotal row

§1.2 Row Echelon Form

Do we always get triangular form?

Okay. Let's try to solve the following (homogeneous) system: (Q. What is an obvious sol'n)

$$\left[\begin{array}{cccc|c} 2 & -4 & 2 & -2 & 0 \\ 2 & -4 & 3 & -4 & 0 \\ 4 & -8 & 3 & -2 & 0 \\ 0 & 0 & -1 & 2 & 0 \end{array}\right] \sim \left[\begin{array}{cccc|c} 1 & -2 & 1 & -1 & 0 \\ 2 & -4 & 3 & -4 & 0 \\ 4 & -8 & 3 & -2 & 0 \\ 0 & 0 & -1 & 2 & 0 \end{array}\right]$$

$$\overset{R_2 \to R_2 - 2R_1}{\sim} \left[\begin{array}{cccc|c} 1 & -2 & 1 & -1 & 0 \\ 0 & 0 & 1 & -2 & 0 \\ 4 & -8 & 3 & -2 & 0 \\ 0 & 0 & -1 & 2 & 0 \end{array}\right] \sim \left[\begin{array}{cccc|c} 1 & -2 & 1 & -1 & 0 \\ 0 & \underline{0} & 1 & -2 & 0 \\ 0 & \underline{0} & -1 & 2 & 0 \\ 0 & \underline{0} & -1 & 2 & 0 \end{array}\right] \sim \dots$$

Trouble (?): we can't put this into strict triangular form: all possible "pivot elements" are 0.

not strictly "row echelon form"

We continue anyway: $\sim \dots \sim$ $\left[\begin{array}{cccc|c} 1 & -2 & 1 & -1 & 0 \\ 0 & 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$,

ending up with the [underdetermined] system

$$x_1 - 2x_2 + x_3 - x_4 = 0$$
$$x_3 - 2x_4 = 0.$$

Terminology:

i) row echelon form (lead entry = 1; # of leading 0's increases)
ii) lead variables ($x_1, x_3$)
iii) free variables ($x_2, x_4$)
iv) Gaussian elimination.

Form of sol'n set

(Q) What kind of a solution set do we get from doing this?

$$x_1 + x_3 = 2x_2 + x_4 \qquad x_3 = 2x_4$$

; we see that if we choose any of numbers $(x_2, x_4)$, we can solve for $x_1, x_3$ and get a sol'n.

So we get an entire plane of solutions:

Back substitution: $x_1 + 2x_4 = 2x_2 + x_4$ ; $x_1 = 2x_2 - x_4$
$x_3 = 2x_4$ ; $x_3 = 2x_4$.

So any 4-tuple $(2x_2 - x_4, x_2, 2x_4, x_4)$ will do.

another obvious remark

Rk: We could've continued one step further in the process:

$$\left[\begin{array}{cccc|c} 1 & -2 & 1 & -1 & 0 \\ 0 & 0 & 1 & -2 & 0 \end{array}\right] \sim \left[\begin{array}{cccc|c} 1 & -2 & 0 & 1 & 0 \\ 0 & 0 & 1 & -2 & 0 \end{array}\right]$$

to obtain $x_1 - 2x_2 + x_4 = 0$ ; $x_1 = 2x_2 - x_4$,
$x_3 - 2x_4 = 0$ $\quad x_3 = 2x_2$.

Terminology: i) reduced row echelon form
ii) Gauss-Jordan reduction

Comment on consistent underdetermined systems: never have unique sol'ns!

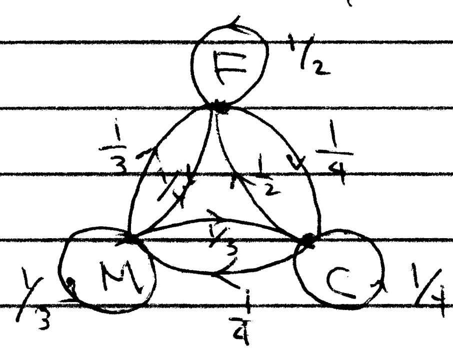
Rk: Underdetermined systems are "usually consistent";
further if consistent ⇒ will have $\infty$-ly many sol'ns
[Why? Because of the existence of free variables]

Rk: Thus any underdetermined homogeneous system has infinitely many solutions
(i.e. never

example

Col. economic example: How to assign relative values to goods that reflect a bartering system?

(Leontief models)




Let $x_1$ = total value of F goods

$x_2$ = M

$x_3$ = C

Then (if fair)

$$x_1 = \frac{1}{2}x_1 + \frac{1}{3}x_2 + \frac{1}{2}x_3$$
$$x_2 = \frac{1}{4}x_1 + \frac{1}{3}x_2 + \frac{1}{4}x_3$$
$$x_3 = \frac{1}{4}x_1 + \frac{1}{3}x_2 + \frac{1}{4}x_3.$$

$\ldots \Rightarrow (5,3,3)$ is a sol'n:

$$\left(\begin{array}{l} x_1 = \frac{5}{3}x_3 \\ x_2 = x_3 \end{array}\right)$$

## §1.3 Basic matrix notation and algebra

notation

What is a matrix? Simply an array of numbers

$$(\alpha_{ij})_{\substack{1\le i\le m\\ 1\le j\le n}} = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & & & \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} \end{bmatrix}$$

usu. denoted by a capital letter, e.g. $A$

Rk: $m$ rows, $n$ columns $\Big\} \Rightarrow$ an "$m \times n$" mx. Notation: $\mathcal{M}_{m,n}(\mathbb{R})$.

Special cases: $m\times 1$ mx called a column vector; $1\times n$ mx (also) called a row vector

Rk: Text writes columns $[\ ]$ and rows $(\ ,\ ,\ )$ (inconsistent)

algebraic operations & relations

Basic ~~notation~~ definitions & algebraic ops.

i) equality: We say $A = B$ if i) $A, B \in \mathcal{M}_{m,n}$ (same size)

ii) $\alpha_{ij} = \beta_{ij}$ $\forall i,j$ (entries equal)

ii) add'n: If $A, B \in \mathcal{M}_{m,n}$, we define $A+B$ to be the $m\times n$ mx with entries $\gamma_{ij} = \alpha_{ij} + \beta_{ij}$.

iii) scalar mult'n: $A \in \mathcal{M}_{m,n}$; $c \in \mathbb{R}$; $\Rightarrow$ we define $cA \in \mathcal{M}_{m,n}$ by $\gamma_{ij} = c\alpha_{ij}$.

iv) zero mx: the mx $0$ of appropriate size with $\gamma_{ij} = 0$

v) additive inverse: $-A := (-1)A$

Rk: Of course, $A + -A = 0$.

funny conventions

Nonstandard but useful conventions:

i) column vectors will be denoted w/ boldface, lowercase letters: $\mathbf{x}, \mathbf{y}, \mathbf{z}$

ii) row vectors by $\vec{x}, \vec{y}, \vec{z}$, etc (arrow above)

iii) If $A = (a_{ij}) = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$,

$\vec{a}_i = [a_{i1} \;\; a_{i2} \;\; \cdots \;\; a_{in}]$ will denote the $i^{th}$ row;

$\mathbf{a}_j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}$ will denote the $j^{th}$ column.

Ex: $\begin{bmatrix} 3 & 2 & 5 \\ -1 & 8 & 4 \end{bmatrix}$ $\quad \vec{a}_2 = [-1 \;\; 8 \;\; 4]$

$\mathbf{a}_2 = \begin{bmatrix} 2 \\ 8 \end{bmatrix}$ ; etc.

Rk: So we ~~may~~ (will) abbreviate $A$ as either

i) $A = [\mathbf{a}_1 \; \mathbf{a}_2 \; \cdots \; \mathbf{a}_n]$ (its column vectors)

ii) $A = \begin{bmatrix} \vec{a}_1 \\ \vec{a}_2 \\ \vdots \\ \vec{a}_m \end{bmatrix}$ (or its row vectors)

(This will come in useful!)

Back to algebraic ops

Most important algebraic operation: Matrix multiplication

Q How should we define it?

The "best" way turns out to be the following:

matrix multiplication

**Defn [matrix multiplication]:**

Given $A = (a_{ik})_{\substack{1 \le i \le m \\ 1 \le k \le n}}$ an $m \times n$ mx.

$B = (b_{kj})_{\substack{1 \le k \le n \\ 1 \le j \le s}}$ an $n \times s$ mx,

The matrix $AB$ is defined to be the $m \times s$ mx. $(c_{ij})$

where $c_{ij} := \vec{a}_i \cdot b_j$

$= [a_{i1}\ a_{i2}\ \cdots\ a_{in}] \cdot \begin{bmatrix} b_{1j} \\ \vdots \\ b_{nj} \end{bmatrix} = \sum_{k=1}^{n} a_{ik} b_{kj}$

← the dot product of the $i$th row of $A$ vs the $j$th column of $B$.

example

**Ex:** $\begin{bmatrix} -3 & 2 & 4 \\ 1 & 0 & 4 \end{bmatrix} \begin{bmatrix} 1 & 1 & 2 \\ 2 & -1 & 0 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & -1 & -6 \\ 1 & 5 & 2 \end{bmatrix}$

$2 \times 3 \qquad 3 \times 3$

Rk: Note $\begin{bmatrix} -3 & 2 & 4 \\ 1 & 0 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$; $\begin{bmatrix} -3 & 2 & 4 \\ 1 & 0 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 5 \end{bmatrix}$, etc.

Important observation

**Key Observation:** (In general, $AB = [Ab_1\ Ab_2\ \cdots\ Ab_s]$)

non-commutativity

Rk: Also notice: $\begin{bmatrix} 1 & 1 & 2 \\ 2 & -1 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} -3 & 2 & 4 \\ 1 & 0 & 4 \end{bmatrix}$ doesn't make sense.
$(3 \times 3) \quad (2 \times 3)$

Even if it does, no guarantee that $AB = BA$:
matrix multiplication is noncommutative

Ex: $\begin{bmatrix} 0 & 2 \\ 3 & 5 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 2 & 5 \end{bmatrix}$

~~Back to~~ linear systems

**Relation with linear systems:**

Suppose we have an $m \times n$ system $\quad a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1$

$$\vdots$$

$$a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m .$$

We can now abbreviate this by letting $A = (a_{ij})$, $\mathbb{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$, $\mathbb{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$;

Then the above system is equivalent to $A\mathbb{x} = \mathbb{b}$:

$$A\mathbb{x} = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \end{bmatrix} \quad (*)$$

For $A\mathbb{x}$ to equal $\mathbb{b}$, the system would have to be satisfied.

Comment

Rk: Thus in some sense our goal is the same one we've had since elementary school: to solve $A\mathbb{x} = \mathbb{b}$ for $\mathbb{x}$.

Q. Can't we just divide by $A$? (Q. When can't we, in $ax = b$?)

alternate viewpoint of linear systems

**Another key observation:** (Alternate viewpoint)

Notice (in $(*)$) that $\quad A\mathbb{x} = x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ \vdots \\ a_{m2} \end{bmatrix} + \dots + x_n \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix}$

$$(= x_1\mathbb{a}_1 + \dots + x_n\mathbb{a}_n)$$

So $\quad A\mathbb{x} = \mathbb{b}$

$$\iff \underline{x_1}\,\mathbb{a}_1 + \underline{x_2}\,\mathbb{a}_2 + \dots + \underline{x_n}\,\mathbb{a}_n = \mathbb{b}$$

i.e. the question is: can we find coeffs $x_1, \dots, x_n$ s.t.

$$\mathbb{b} = x_1\mathbb{a}_1 + \dots + x_n\mathbb{a}_n \,?$$

11

linear combination — **Def'n:** Let $\mathbb{v}_1, \dots, \mathbb{v}_n$ be vectors in $\mathbb{R}^m$.

Any sum of the form $c_1\mathbb{v}_1 + c_2\mathbb{v}_2 + \dots + c_n\mathbb{v}_n$ ; $c_i \in \mathbb{R}$ is called a <u>linear combination</u> of $\mathbb{v}_1, \dots, \mathbb{v}_n$.

span — **Def'n:** The colln. of all linear combinations of $\mathbb{v}_1, \dots, \mathbb{v}_n$ is called the <u>span</u> of $\mathbb{v}_1, \dots, \mathbb{v}_n$.

alt. view — **Rk:** Thus the question "Can we solve $A\mathbb{x} = \mathbb{b}$?"

$\iff$ Is $\mathbb{b}$ in the span of $\mathbb{a}_1, \dots, \mathbb{a}_n$?

(ie, can $\mathbb{b}$ be realized as a linear comb'n of $\mathbb{a}_1, \dots, \mathbb{a}_n$?)

"consistency theorem" — Ie, **Thm:** $A\mathbb{x} = \mathbb{b}$ is consistent $\iff$ $\mathbb{b} \in \text{span}\{\mathbb{a}_1, \dots, \mathbb{a}_n\}$

## §1.4 More on matrix multiplication. (fairly empty section)

Recall: Given $A = (a_{ij})$ an $m \times n$ matrix, $B = (b_{ij})$ an $n \times s$ mx,

defn mult. $AB = (c_{ij})$ ~~was~~ the $m \times s$ matrix

with $c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$;

(Further, we saw $AB = [Ab_1 \ldots Ab_s]$).

Although in general $AB \neq BA$, matrix multiplication does satisfy some familiar algebraic properties:

laws

1. $A + B = B + A$
2. $(A+B)+C = A+(B+C)$
3. $(AB)C = A(BC)$
4. $A(B+C) = AB + AC$
5. $(A+B)C = AC + BC$
6. $(\alpha\beta)A = \alpha(\beta A)$, etc.

elem. proofs

How does one prove them?

Pf (1): WTS $A + B = B + A$. Well, we need show corresponding entries are equal.

Let $c_{ij}$ denote the $(i,j)^{th}$ entry of $A+B$,

$\gamma_{ij}$ - - - - - - - - - - - - $B+A$

Since $c_{ij} = a_{ij} + b_{ij} = b_{ij} + a_{ij} = \gamma_{ij}$, we're done.

Pf (4): Let $D = A(B+C)$ $E = AB + AC$.

By definition, $d_{ij} = \sum_{k=1}^{n} a_{ik}(b_{kj} + c_{kj})$

and $e_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} + \sum_{k=1}^{n} a_{ik} c_{kj}$.

Clearly $d_{ij} = e_{ij}$; so we're done.

Pf (3): Say $A$ is $m \times n$, $B$ is $n \times r$, $C$ is $r \times s$.
Let $D = AB$, $E = BC$.

We want to show $DC = AE$.

Well... the

$(i,j)^{\text{th}}$ entry of $DC$ is $\sum_{l=1}^{r} d_{il} c_{lj}$

$$= \sum_{l=1}^{r} \left[ \sum_{k=1}^{n} a_{ik} b_{kl} \right] c_{lj}.$$

$(i,j)^{\text{th}}$ entry of $AE$ is $\sum_{k=1}^{n} a_{ik} e_{kj}$

$$= \sum_{k=1}^{n} a_{ik} \left[ \sum_{l=1}^{r} b_{kl} c_{lj} \right]$$

These are equal, so we're done.

Rk: Difficulty is mainly in choosing good notation, and bookkeeping

Comment: So $A(AA) = (AA)A$, so we just write $AAA$ or $A^3$. (etc)

Application (Markov chains) Let $x$ = # of married women, $y$ = # unmarried.

Suppose we have a transition matrix $\begin{bmatrix} .7 & .2 \\ .3 & .8 \end{bmatrix} = A$.

embodying the statistics that .7 of married women stay married,
.2 of single women get married, etc.

Then we can model the future ...

(Q) Does this system reach an equilibrium?

I. **Identity mx:** Now that we have multiplication, we'd like a multiplicative identity (analogue of 1), i.e. a mx $I$ s.t. $IA = A = AI$.

~~Q. What mx works for $I$? (Introduce Kronecker $\delta_{ij}$)~~

$A^{-1}$ **Multiplicative Inverses:**

**Defn:** $A$ an $n \times n$ mx. If $\exists\, B \in M_{n,n}(\mathbb{R})$ s.t. $AB = BA = I$, then we say $A$ is invertible (nonsingular) and call $B$ the (!) (multiplicative) inverse of $A$. (denoted $A^{-1}$)

**Ex:** $\begin{bmatrix} 1 & 3 & -2 \\ 2 & 5 & -3 \\ -3 & 2 & -4 \end{bmatrix}$ and $\begin{bmatrix} 14 & -8 & -1 \\ -17 & 10 & 1 \\ -19 & 11 & 1 \end{bmatrix}$ are inverses

Rk: Was the "the" justified? Could there be more than one such $B$?

→ ~~Rk: $A\vec{x} = \vec{b}$.~~ Think: If $B, B'$ both inverses $\Rightarrow B = BAB' = B'$.

Rk: Funny question: could we have $B$ s.t. $AB = I$, $BA \neq I$?

~~Q. Are all mxs. invertible? [Create examples]~~

**Observation:** If $A, B \in M^*_{n,n}(\mathbb{R})$ (i.e. are invertible) then $AB$ is also invertible: $(AB)^{-1} = B^{-1}A^{-1}$.

Pf: Obvious.

[Transposes: Rules (read yourself). Most important $(AB)^T = ?$]

[Digression: Proof by induction]

Exercises in class: 14, 16, 30, 36.

## §1.5 Elementary Matrices

Let's begin with a few simple observations:

equivalent systems

• Lemma: $A\vec{x} = \vec{b}$ an $m \times n$ linear system, $M \in \mathcal{M}_{m,m}$.

Claim: $M$ invertible $\Rightarrow$ $A\vec{x} = \vec{b}$ is equivalent to $MA\vec{x} = M\vec{b}$

(i.e., has the same soln set as)

Pf: Suppose $\vec{v}$ solves $A\vec{x} = \vec{b}$, i.e. $A\vec{v} = \vec{b}$.

Well, certainly $MA\vec{v} = M\vec{b}$, i.e.

any soln of I solves II.

Suppose $\vec{v}$ solves $MA\vec{x} = M\vec{b}$ ... (Duh). Trivial.

Rk: Now it's trivial, but it's not a priori trivial.

elem. mxs

• Second simple observation: "Elementary Row Operations"

~~Let $A = (\alpha_{ij})$.~~ We can switch $R_1$ and $R_2$ by

switching $R_1$ and $R_2$ in $I$ and then left-multiplying:

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{bmatrix} = \begin{bmatrix} \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{bmatrix}.$$

Similarly for the other elementary row operations.

Defn: If the matrix $E$ can be obtained from $I$ via a single elementary row operation, we call $E$ an elementary mx.

elem. mxs invertible

• Remark: Every elementary mx. is invertible.

Pf: Take the elem. mx. that reverses the operation: that's $E^{-1}$!

E.g. $R_3 \to R_3 + 5R_1$ is reversed by $R_3 \to R_3 - 5R_1$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 5 & 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -5 & 0 & 1 \end{bmatrix}$$

A ~ B

**Defn:** We say the mx $A$ is <u>row equivalent</u> to $B$ $(A \sim B)$ if $A$ can be obtained from $B$ by elementary row ops.

Pf: Equivalently: $A \sim B \iff \exists$ elem mxs $E_1, \ldots, E_k$ s.t. $B = E_k E_{k-1} \cdots E_1 A$.

[ Q: If $A \sim B$, is $B \sim A$?
$A \sim B$, $B \sim C \Rightarrow A \sim C$? ]

We can use this definition to obtain a nice (?)

charzn of invertibility

<u>Characterization of Invertibility:</u>

Thm: $A \in M_{n,n}(\mathbb{R})$. Then TFAE:

i) $A$ is invertible
ii) $Ax = \vec{0}$ has a unique soln $(\vec{0})$
iii) $A \sim I$
iv) $A$ is a product of elementary matrices.

Pf: i) $\Rightarrow$ ii) obvious.

ii) $\Rightarrow$ iii) Reduce $A$ to reduced row echelon form.
Suppose $\alpha_{ii} = 0$ for some $i$; then $\vec{a}_n = \vec{0}$
$\Rightarrow$ $A$ has infinitely many solns ✗
Thus $\alpha_{ii} = 1$ $\forall i = 1, \ldots, n$; so the reduced form is $I$

iii) $\Rightarrow$ iv) obvious: $E_k \cdots E_2 E_1 A = I$; so
$A = E_1^{-1} E_2^{-1} \cdots E_k^{-1}$.

iv) $\Rightarrow$ i) obvious.

~~examination~~

Remark: Did we use the $\vec{0}$ in (ii) anywhere? In fact, $\iff$
"$Ax = b$ has a unique soln for any $b$."

Computation of $A^{-1}$

How does one determine an inverse?

Computation of Inverses: Book-keeping 101.

We just noticed that if $A$ is invertible, then $A \sim I$;

i.e., $\exists$ elem. mxs s.t. $E_n \cdots E_1 A = I$.

Well then $E_n \cdots E_1 = A^{-1}$!

This fact suggests the following "method":

Given $A$, take $[A \mid I]$.

Try to change $A$ into $I$ by applying elementary mxs.;

and apply those mxs. to $I$ on the right to keep track.

$\sim \cdots \sim [E_k \cdots E_1 A \mid E_k \cdots E_1]$.

When the LHS becomes $I$, the RHS will be $A^{-1}$!

Ex: $A = \begin{bmatrix} 1 & 3 & -2 \\ 2 & 5 & -3 \\ -3 & 2 & -4 \end{bmatrix}$

Exercises in class: 17, 29, 31

LU factorization (using "Type III" elem. mxs)

LU Factorization*:

Defns: $A \in M_{n,n}(\mathbb{R})$ is called (strict) (upper/lower) triangular if ...;

diagonal if ...

Observation: $[A \in M_{n,n}]$ If $A$ can be reduced to a strict upper triangular form using elem. row ops. of the form $R_i \to R_i + \alpha R_j$,

then $A = LU$ for some lower triangular mx $L$,

upper triangular mx $U$.

"A has a LU factorization" (p.65) (Read yourself)

(§2.2 and its properties.)

§2.1 The determinant

pf. by induction — < Introduce proof by induction $\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$

(a lie)

motivation — "Goal": To associate a number with a matrix that .

To get there, we need a few definitions:

minor matrix — **Defn:** [Minor matrices] $A \in M_{n,n}(\mathbb{R})$.

We define the $(i,j)^{th}$ minor matrix $M_{ij}$ to be the $(n-1)\times(n-1)$ mx obtained by excising $A$'s $i^{th}$ row & $j^{th}$ column

1x1 det — **Defn:** If $A$ is a $1\times1$ mx, we define $\det A = |A| = a_{11}$

nxn det — **Defn:** [Determinant] $A$ an $n\times n$ mx, $n\geq 2$.

We define the determinant $\det(A)$ or $|A|$ of the matrix $A$

by $\boxed{\det(A) := a_{11}A_{11} + a_{12}A_{12} + \dots + a_{1n}A_{1n}}$,

cofactor — where $A_{ij} := (-1)^{i+j}\det(M_{ij})$,

the $(i,j)^{th}$ cofactor of $A$

inductive defn — **Rk:** This is an inductive definition: $n\times n$ dets. are defined in terms of $(n-1)\times(n-1)$ determinants, which are ....

examples — **Ex:** $\det\begin{bmatrix}3&2\\-2&4\end{bmatrix} = 3(-1)^{1+1}\det[4] + 2(-1)^{1+2}\det[-2]$

$= 3\cdot4 - 2\cdot(-2) = 16$. In general, $\begin{vmatrix}a&b\\c&d\end{vmatrix} = ad-bc$

**Ex:** $\begin{vmatrix}1&0&7\\1&3&2\\1&-2&4\end{vmatrix} = 1\cdot\begin{vmatrix}3&2\\-2&4\end{vmatrix} - 0\cdot\begin{vmatrix}1&2\\1&4\end{vmatrix} + 7\cdot\begin{vmatrix}1&3\\1&-2\end{vmatrix}$.

$= 16 - 0 + -35 = -19$.

**Ex:** $4\times4$ etc

**Rk:** "cofactor expansion"/"expansion by minors".

Rk: We could, if we wanted, take cofactor expansions down any row or column. (Do some!)

Useful Thm

Thm: Cofactor expansions down any row or column yield the same value, viz. $\det(A)$.

Pf: Not particularly illuminating (Appendix B, Fraleigh & Beauregard)

application

Ex: Find determinant of

$$A = \begin{bmatrix} 3 & 2 & 0 & 1 & 3 \\ -2 & 4 & 1 & 2 & 1 \\ 0 & -1 & 0 & 1 & -5 \\ -1 & 2 & 0 & -1 & 2 \\ 0 & 0 & 0 & 0 & 2 \end{bmatrix}$$

practice: pf. by induction

Rk: What is the determinant of a diagonal matrix?
(Trivial conseq. of prev. thm, but) (Do as an example of proof by induction)

Claim: $A = (a_{ij})$ an $n \times n$ diagonal mx $\Rightarrow \det(A) = \prod_{i=1}^{n} a_{ii}$

Pf: By induction. If $n=1$, then done.
So assume statement is true for $(n-1)\times(n-1)$ matrices.

Well, $\det(A) = a_{11}A_{11} + \dots + a_{1n}A_{1n}$

$= a_{11}A_{11}$ $\quad (a_{1j} = 0$ for all $j \neq 1)$

$= a_{11}\det(M_{11})$

But $M_{11}$ is an $(n-1)\times(n-1)$ mx which is diagonal;

so by Induction Hypothesis $\det(M_{11}) = \prod_{i=2}^{n} a_{ii}$

$\Longrightarrow \det(A) = a_{11}\prod_{i=2}^{n} a_{ii} = \prod_{i=1}^{n} a_{ii}$ ∎

(Rk: Similarly for triangular mxs)

The "useful thm." can be used to imply various properties of det:

Consequences of useful thm.

Thm: [Properties of the determinant]. $A$ an $n \times n$ mx

i) $\det(A) = \det(A^T)$

ii) Let $A'$ be obtained from $A$ by interchanging 2 rows (cols)

Then $\det(A') = -\det(A)$

Corollary: If $A$ has 2 identical rows (cols), $\det(A) = 0$ (why?)

iii) Let $A'$ be obtained from $A$ by multiplying a row by $\alpha \in \mathbb{R}$

$\Longrightarrow \det(A') = \alpha \det(A)$

practice (iii)

Writing Practice: Pf iii): Take the cofactor expansion of $\det(A')$ along the ith row

$$\det(A') = a'_{i1}A'_{i1} + \dots + a'_{in}A'_{in}.$$

Since $a'_{ij} = \alpha a_{ij}$; $j = 1, \dots, n$, we get

$$\det(A') = \alpha a_{i1}A'_{i1} + \dots + \alpha a_{in}A'_{in}.$$

$$= \alpha a_{i1}A_{i1} + \dots + \alpha a_{in}A_{in} \quad \text{since } A'_{ij} = A_{ij}$$

$$= \alpha \det(A) \blacksquare$$

pf of ii via induction

Pf: The only really interesting one is ii):

Pf: By induction (obvious in case $n=2$, meaningless if $n=1$)

Assume statement true for $(n-1)\times(n-1)$ matrices

Suppose $A'$ is obtained by switching rows $i$ and $r$ in $A$.

Choose a row $k$, where $k \neq i$ or $r$.

By induction hypothesis, $|M_{kj}| = (-1)|M'_{kj}|$.

So $\sum a_{kj}A'_{kj} = \sum -a_{kj}A_{kj}$

(expanded on next page) i.e. $\det(A') = -\det(A)$ $\blacksquare$

Notice that each $M'_{kj}$; $j=1,\dots,n$ is $M_{kj}$ w/ 2 rows switched;

so $|M'_{kj}| = (-1)|M_{kj}|$ by induction hypothesis.

$$\left( \Rightarrow \quad A'_{kj} = (-1)^{k+j}|M'_{kj}| = -A_{kj}. \right)$$

Taking the expansion down the $k$-th row,

$$\det(A') = \sum_{j=1}^{n} a'_{kj} A'_{kj} = \sum_{j=1}^{n} a_{kj}(-1)A_{kj} = -\det(A)$$

In class exercise: 5. Show $\det(\alpha A) = \alpha^n \det(A)$.

<u>Rk</u>: It's not surprising that so many proofs are by induction: the definition is inductive!

## §2.2. Properties of Determinants.

Recall: We showed ($A$ an $n\times n$ mx.)

i) If $A'$ is obtained from $A$ by switching 2 rows $\Rightarrow \det A' = -\det A$

ii) multiplying a row by $\alpha \Rightarrow \det A' = \alpha \det A$.

Q) What will we prove next?

iii) If $B$ is obtained from $A$ by adding a multiple of one row onto another,

$\Longrightarrow \det(B) = \det(A)$.

To do this, we will need the following lemma.

Rk: Recall: how did we show $\det(A)=0$ if two rows were identical?

We expanded down a row different from the identical ones.

Here's what information we get by expanding down one of the rows.

"the wrong cofactor lemma"

Lemma: $A$ an $n\times n$ mx.

Then $a_{i1}A_{j1} + a_{i2}A_{j2} + \dots + a_{in}A_{jn} = \begin{cases} \det(A) & ; i=j \\ 0 & i\neq j \end{cases}$.

Rk: I.e., using the wrong row's cofactors yields ... 0 !!

Or, better: if $i\neq j$, $\langle a_{i1},\dots,a_{in}\rangle \perp \langle A_{j1},\dots,A_{jn}\rangle$

Pf: We've already noted ("useful thm.") that if $i=j$ one gets $\det(A)$.

Now consider

$$C := \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ & & & \\ a_{i1} & a_{i2} & \dots & a_{in} \\ c_{j1} & a_{i2} & \dots & a_{in} \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \begin{array}{l} \\ \\ \leftarrow i^{th} \text{ row} \\ \leftarrow j^{th} \text{ row.} \\ \\ \end{array}$$

i.e. $\vec{c}_j = \vec{a}_i$; all other rows identical ($\vec{c}_k = \vec{a}_k$, $k\neq j$)

If we expand $\det(C)$ down the $j^{th}$ row, we get

$$0 = \det(C) = a_{i1}A_{j1} + \dots + a_{in}A_{jn}. \blacksquare$$

Now we'll use that to show iii):

Pf: Say $B$ is $A$ with $R_k \to R_k + cR_i$,

i.e. $\vec{b}_k = c\vec{a}_i + \vec{a}_k$.

Obviously, $B_{kj} = A_{kj}$ for $j = 1, \dots, n$ since the $(k,j)^{th}$ minors are equal, $j = 1, \dots, n$

Expand $\det(B)$ down the $k^{th}$ row;

$$\det(B) = b_{k1}B_{k1} + \dots + b_{kn}B_{kn}$$
$$= (ca_{i1} + a_{k1})A_{k1} + \dots + (ca_{in} + a_{kn})A_{kn}$$
$$= c(a_{i1}A_{k1} + \dots + a_{in}A_{kn}) + \det(A)$$
$$= \det(A) \text{ by the lemma. } \blacksquare$$

Rk: What are the $\det(E)$ where $E$ are elementary mx's?

$\det(E) = -1, \alpha$, or $1$ depending on type.

So, note; the above result (i–iii) could be rephrased as

Given any elementary mx $E$, $\det(EA) = \det(E)\det(A)$

Corollary: $A$ is invertible $\Leftrightarrow$ $\det(A) \neq 0$

Pf: Let $U = E_k \dots E_1 A$ be the reduced row echelon form of $A$.

$A$ invertible $\Rightarrow \det(U) = \det(I) \neq 0 \Rightarrow \det(A) \neq 0$.

[Exerc. 16] $A$ noninvertible $\Rightarrow \det(U) = 0 \Rightarrow \det(A) = 0$. $\blacksquare$

Thm: $\det(AB) = \det(A)\det(B)$ [read yourself!]

Application: Calculation of determinants. $\begin{vmatrix} 5 & -2 & 1 \\ 3 & 2 & 0 \\ 3 & 3 & 3 \end{vmatrix} = 3 \begin{vmatrix} 5 & -2 & 1 \\ 3 & 2 & 0 \\ 1 & 1 & 1 \end{vmatrix} = -3 \begin{vmatrix} 1 & 1 & 1 \\ 3 & 2 & 0 \\ 5 & -2 & 1 \end{vmatrix}$

$\dots (= 51)$

Rk. Efficiency.

Exerc 3d. $\begin{bmatrix} 2 & 1 & 1 \\ 4 & 3 & 5 \\ -2 & 1 & 2 \end{bmatrix}$ invertible?

## §23 The Adjoint Matrix; Cramer's Rule

**Adjoint mx and inverses:**

Recall the "wrong cofactor lemma": $a_{i1}A_{j1} + \dots + a_{in}A_{jn} = 0$ if $i \neq j$.

This observation provides us an alternate view of the inverse:

adjoint

**Defn:** $A$ an $n \times n$ mx. We define the adjoint of $A$ by

$$\operatorname{adj}(A) := \begin{bmatrix} A_{11} & \cdots & A_{n1} \\ A_{12} & & \vdots \\ \vdots & & \\ A_{1n} & & A_{nn} \end{bmatrix}$$

i.e. the transpose of the "cofactor mx" of $A$.

reln w/ inverse

**Observe:** $A \operatorname{adj}(A) = \det(A) I$

Why? $$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} A_{11} & \cdots & A_{n1} \\ \vdots & & \vdots \\ A_{1n} & \cdots & A_{nn} \end{bmatrix} = \begin{bmatrix} \det A & 0 & \cdots & 0 \\ & & & \\ & & & \end{bmatrix}$$

$$\left( x_{ij} = a_{i1}A_{j1} + a_{i2}A_{j2} + \dots + a_{in}A_{jn} = \begin{cases} \det A & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \right).$$

what is the inverse?

So what is the inverse matrix? $\boxed{A^{-1} = \frac{1}{\det(A)} \operatorname{adj}(A)}$

**Rks** This is no more efficient than our previous method.

Let us give another example of a theoretically interesting observation (but practically impractical method).

## Determinants & Linear Systems: Cramer's Rule

Cramer's Rule

**Thm:** Given the eq'n $A\vec{x} = \vec{b}$, where $A$ is square invertible,

then $\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ where $x_k = \dfrac{\det(B_k)}{\det(A)}$; $k = 1, \ldots, n$

(where $B_k$ is the mx obtained by replacing the $k^{th}$ col. of $A$ w/ $\vec{b}$)

example

Ex: 
$$5x_1 - 2x_2 + x_3 = 1$$
$$3x_1 + 2x_2 = 3$$
$$x_1 + x_2 - x_3 = 0$$

$A = \begin{bmatrix} 5 & -2 & 1 \\ 3 & 2 & 0 \\ 1 & 1 & -1 \end{bmatrix}$, $B_1 = \begin{bmatrix} 1 & -2 & 1 \\ 3 & 2 & 0 \\ 0 & 1 & -1 \end{bmatrix}$

$B_2 = \begin{bmatrix} 5 & 1 & 1 \\ 3 & 3 & 0 \\ 1 & 0 & -1 \end{bmatrix}$

$B_3 = \begin{bmatrix} 5 & -2 & 1 \\ 3 & 2 & 3 \\ 1 & 1 & 0 \end{bmatrix}$

We calculate $|A| = 15$, $|B_1| = -5$
$|B_2| = -15$
$|B_3| = -20$

So $\vec{x} = \begin{bmatrix} \frac{1}{3} \\ 1 \\ \frac{4}{3} \end{bmatrix}$.

"This illustrates the folly of using Cramer's rule to solve linear systems."

proof using adjoint mx

Pf of Cramer's Rule: $A\mathbb{x} = \mathbb{b} \Rightarrow \mathbb{x} = A^{-1}\mathbb{b}$

$$\Rightarrow \mathbb{x} = \frac{1}{\det A} \operatorname{adj}(A)\mathbb{b}.$$

$$\Rightarrow x_k = \frac{1}{\det A}\left(\underbrace{A_{1k}b_1 + A_{2k}b_2 + \ldots + A_{nk}b_n}_{\text{cofactor expansion along } k^{th} \text{ col. of matrix } A \text{ with } \mathbb{b} \text{ as } k^{th} \text{ col}}\right)$$

$$= \frac{1}{\det A} \det(B_k). \blacksquare$$

## §2.2 Properties of Determinants.

Thm: det & elem. row ops.

Recall We showed ($A$ an $n\times n$ mx.)

i) If $\hat{A}$ is obtained from $A$ by switching 2 rows, $\det \hat{A} = -\det A$

ii) If $\hat{A}$ is obtained from $A$ by multiplying a row by $\alpha$, $\det\hat{A} = \alpha \det A$

What will we prove next?

~~iii) If B is obtained from A by adding a multiple of one row to another, then~~ $\cancel{\det(B) = \det(A)}$

comment mx. viewpt.

[Rk. We could rephrase all of the above in terms of elem. mxs:

- that given any elem mx $E$, $\det(EA) = \det(E)\det(A)$

Pf of iii)

Pf of iii): Suppose $\vec{b}_k = c\vec{a}_i + \vec{a}_k$ (i.e. $B$ is $A$ with $R_k \to R_k + cR_i$)

Obviously, $B_{kj} = A_{kj}$ for $j=1,\dots,n$.

(note the $(k,j)^{th}$ minors are all equal, $j=1,\dots,n$)

So if we expand $\det(B)$ along the $k^{th}$ row,

$$\det(B) = b_{k1}B_{k1} + b_{k2}B_{k2} + \dots + b_{kn}B_{kn}$$

$$= (ca_{i1}+a_{k1})A_{k1} + (ca_{i2}+a_{k2})A_{k2} + \dots + (ca_{in}+a_{kn})A_{kn}$$

$$= c\left[a_{i1}A_{k1} + a_{i2}A_{k2} + \dots + a_{in}A_{kn}\right] + \det(A)$$

Now, the term in brackets is the cofactor expansion of $\det(C)$, where $C$ is a matrix with 2 identical rows

($C$ is $A$ with row $k$ replaced by a copy of row $i$)

So $\det(B) = 0 + \det(A)$ ■

Rk: We formalize the last step in the following lemma:

The "wrong cofactor lemma"

<u>Lemma</u>: $A$ an $n\times n$ mx. Then $a_{i1}A_{j1} + a_{i2}A_{j2} + \dots + a_{in}A_{jn} = \begin{cases} \det(A) & ; i=j \\ 0 & ; i\neq j. \end{cases}$

Rk: Ie, using the wrong row's cofactors gives 0.

Rk: Or better: if $i\neq j$, $\langle a_{i1}, \dots, a_{in}\rangle \perp \langle A_{j1}, \dots, A_{jn}\rangle$.

<u>Pf</u>: [First part is from the "useful theorem" about cofactor expansions.]

Consider the matrix $C = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \\ a_{i1} & a_{i2} & \dots & a_{in} \\ a_{i1} & a_{i2} & \dots & a_{in} \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \leftarrow j^{th}$ row.

(ie, $\vec{c}_k = \begin{cases} \vec{a}_k & ; k\neq j \\ \vec{a}_i & ; k=j \end{cases}$)

Expanding $\det C$ along the $j^{th}$ row we get

$$0 = \det C = a_{i1}A_{j1} + a_{i2}A_{j2} + \dots + a_{in}A_{jn}.$$

Consequences of Thm:

<u>Corollary</u>: $A$ is invertible $\Leftrightarrow$ $\det(A)\neq 0$.

<u>Pf</u>: Let $U = E_k \dots E_1 A$ be the reduced row echelon form of $A$

$A$ invertible $\Rightarrow$ $U = I$ $\Rightarrow$ $\det A \neq 0$.

Exerc 16 $\rightarrow$ $A$ noninvertible $\Rightarrow$ $\det(U) = 0$ $\Rightarrow$ $\det A = 0$. ■

[Thm: $\det(AB) = \det(A)\det(B)$ <u>Pf</u>: Read yourself!]

<u>Application</u>: Calculation of determinants

$$\begin{vmatrix} 5 & -2 & 1 \\ 3 & 2 & 0 \\ 1 & 1 & 1 \end{vmatrix} = (-1)\begin{vmatrix} 1 & 1 & 1 \\ 3 & 2 & 0 \\ 5 & -2 & 1 \end{vmatrix} = (-1)\begin{vmatrix} 1 & 1 & 1 \\ 0 & -1 & -3 \\ 5 & -2 & 1 \end{vmatrix} = (-1)\begin{vmatrix} 1 & 1 & 1 \\ 0 & -1 & -3 \\ 0 & -7 & -4 \end{vmatrix}$$

$$= (-1)(-1)\begin{vmatrix} 1 & 1 & 1 \\ 0 & 1 & 3 \\ 0 & -7 & -4 \end{vmatrix} = (-1)^2\begin{vmatrix} 1 & 1 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 17 \end{vmatrix} = 17.$$ Rks: Efficiency.

Exerc. 31.

## §23 The Adjoint Matrix; Cramer's Rule

**Adjoint mx and inverses:**

Recall the "wrong cofactor lemma": $a_{i1}A_{j1} + \dots + a_{in}A_{jn} = 0$ if $i \neq j$.

This observation provides us an alternate view of the inverse:

<u>Defn:</u> $A$ an $n \times n$ mx. We define the <u>adjoint</u> of $A$ by

$$\operatorname{adj}(A) := \begin{bmatrix} A_{11} & \cdots & A_{n1} \\ A_{12} & & \vdots \\ A_{1n} & & A_{nn} \end{bmatrix}$$

i.e. the transpose of the "cofactor mx" of $A$.

<u>Observe:</u> $A \operatorname{adj}(A) = \det(A) I$

why?
$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} A_{11} & \cdots & A_{n1} \\ \vdots & & \vdots \\ A_{1n} & \cdots & A_{nn} \end{bmatrix} = \begin{bmatrix} \det A & 0 & \cdots & 0 \\ & & & \\ & & & \end{bmatrix}$$

$$\left( x_{ij} = a_{i1}A_{j1} + a_{i2}A_{j2} + \dots + a_{in}A_{jn} = \begin{cases} \det A & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \right)$$

So what is the inverse matrix? $\boxed{A^{-1} = \frac{1}{\det(A)} \operatorname{adj}(A).}$

<u>Rk:</u> This is no more efficient than our previous method.

Let us give another example of a theoretically interesting observation (but practically impractical method).

## Determinants & Linear Systems: Cramer's Rule

**Thm:** Given the eq'n $A\vec{x} = \vec{b}$, where $A$ is square invertible,

then $\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ where $x_k = \dfrac{\det(B_k)}{\det(A)}$; $k = 1, \ldots, n$

(where $B_k$ is the mx. obtained by replacing the $k^{th}$ col. of $A$ w/ $\vec{b}$)

**Ex:**

$$\begin{aligned} 5x_1 - 2x_2 + x_3 &= 1 \\ 3x_1 + 2x_2 \quad &= 3 \\ x_1 + x_2 - x_3 &= 0 \end{aligned} \quad ; \quad A = \begin{bmatrix} 5 & -2 & 1 \\ 3 & 2 & 0 \\ 1 & 1 & -1 \end{bmatrix}, \quad B_1 = \begin{bmatrix} 1 & -2 & 1 \\ 3 & 2 & 0 \\ 0 & 1 & -1 \end{bmatrix}$$

$$B_2 = \begin{bmatrix} 5 & 1 & 1 \\ 3 & 3 & 0 \\ 1 & 0 & -1 \end{bmatrix} \quad B_3 = \begin{bmatrix} 5 & -2 & 1 \\ 3 & 2 & 3 \\ 1 & 1 & 0 \end{bmatrix}$$

We calculate $|A| = 15$, $|B_1| = -5$, $|B_2| = -15$, $|B_3| = -20$

So $\vec{x} = \begin{bmatrix} -\frac{1}{3} \\ -1 \\ -\frac{4}{3} \end{bmatrix}$.

"This illustrates the folly of using Cramer's rule to solve linear systems."

Pf of Cramer's Rule: $A\mathbb{x} = \mathbb{b} \Rightarrow \mathbb{x} = A^{-1}\mathbb{b}$

$$\Rightarrow \mathbb{x} = \frac{1}{\det A} \operatorname{adj}(A)\mathbb{b}.$$

$$\Rightarrow x_k = \frac{1}{\det A}\left(A_{1k} b_1 + A_{2k} b_2 + \ldots + A_{nk} b_n\right)$$

$\shortparallel$ cofactor expansion along $k^{th}$ col. of matrix $A$ with $\mathbb{b}$ as $k^{th}$ col.

$$= \frac{1}{\det A} \det(B_k). \blacksquare$$

- The ~~true~~ power of linear algebra comes from the flexibility of the idea. Turns out (?) from the correct viewpoint many constructs are basically "like" matrices

(or, "that wasn't Linear Algebra!")

## §3.1 Vector Spaces

defn of v.s.

<u>Defin:</u> Let $V$ be a set with an <u>addition</u> and a <u>scalar multiplication</u>,

i.e. $\forall \vec{x}, \vec{y} \in V$ $\exists$ a well-defined "$\vec{x}+\vec{y}$" $\in V$;

$\forall \vec{x} \in V$, $\alpha \in \mathbb{R}$ $\exists$ a well-defined "$\alpha\vec{x}$" $\in V$.

If A1 addition is commutative

A2 associative

A3 $\exists$ a (unique) element $\vec{0} \in V$ s.t. $\vec{x}+\vec{0} = \vec{x}$ $\forall \vec{x} \in V$

A4 $\forall \vec{x} \in V$ $\exists$ a (unique) element "$-\vec{x}$" $\in V$ s.t. $\vec{x}+(-\vec{x}) = 0$

S1 $\alpha(\vec{x}+\vec{y}) = \alpha\vec{x}+\alpha\vec{y}$ $\forall \alpha \in \mathbb{R}$, $\vec{x}, \vec{y} \in V$

S2 $(\alpha+\beta)\vec{x} = \alpha\vec{x}+\beta\vec{x}$

S3 $(\alpha\beta)\vec{x} = \alpha(\beta\vec{x})$ $\forall \alpha, \beta \in \mathbb{R}$, $\vec{x} \in V$

S4 $1 \cdot \vec{x} = \vec{x}$ $\forall \vec{x} \in V$

Then we call $V$ a <u>vector space</u> and its elements <u>vectors</u>

remarks (& examples)

<u>Rk:</u> i) "scalar" actually means element of a "field"

ii) The property $\vec{x}, \vec{y} \in V \Rightarrow \vec{x}+\vec{y} \in V$ we call "closed under addition". The other, "closed under scalar mult"

iii) Is $\vec{0}$ unique? Could there be another?

No: $\vec{0}' = \vec{0}' + \vec{0} = \vec{0}$, if both are zeroes

~~minimal~~ set of axioms

<u>Rk:</u> The above set of "axioms" is minimal; other properties are consequences

~~properties~~ which follow

Thm: [Other properties] $V$ a vector space, $\vec{x} \in V$.

i) $0\vec{x} = \vec{0}$

ii) $\vec{x}+\vec{y} = \vec{0} \Rightarrow \vec{y} = -\vec{x}$ (uniqueness of additive inverse)

iii) $(-1)\vec{x} = -\vec{x}$

Pf i) WTS that $0\vec{x} = \vec{0}$.

Well, $\vec{x} = 1\vec{x} = (1+0)\vec{x} = \vec{x} + 0\vec{x}$

$$\Rightarrow \vec{0} = -\vec{x} + \vec{x} = -\vec{x} + (\vec{x} + 0\vec{x})$$
$$= (-\vec{x} + \vec{x}) + 0\vec{x}$$
$$= \vec{0} + 0\vec{x} = 0\vec{x}$$

ii) WTS $\vec{x} + \vec{y} = \vec{0} \Rightarrow \vec{y} = -\vec{x}$.

Well, $-\vec{x} = -\vec{x} + \vec{0}$
$$= -\vec{x} + (\vec{x} + \vec{y})$$
$$= (-\vec{x} + \vec{x}) + \vec{y} = \vec{y}$$

iii) $(-1)\vec{x} = -\vec{x}$:

Well, $\vec{x} + (-1)\vec{x} = (1 + -1)\vec{x} = 0\vec{x} = \vec{0}$ by (i).

$$\Longrightarrow (-1)\vec{x} = -\vec{x} \text{ by (ii)}. \blacksquare$$

examples.

Examples of vector spaces:

i) $\mathbb{R}^n$

ii) $M_{n,m}(\mathbb{R})$ ($\mathbb{R}^{n\times m}$ in the text)

iii) $C[a,b]$ continuous fns. on $[a,b]$ ($C^n[a,b]$)

iv) $P_n$ (poly'ls of degree $< n$)

Other examples?

Warm-up: If $CA = I$, then $AC = I$.
Hint: Show $C$ is invertible.
(since $CA\vec{x} = \vec{b}$ is always consistent)

## §3.2 Subspaces

defn. of subspace

**Defn:** [$V$ a vector space] If $W \subset V$, with the vector add'n and scalar multiplication inherited from $V$, is itself a vector space, we call it a subspace.

examples

**Ex:** $C^n[a,b] \subset C[a,b]$ is a (proper) subspace

**Ex:** Let $S = \{p \in P_n \mid p(0) = 0\}$. $S$ is a subspace.

Q ~~Is $\{p(0)=1\}$ a subspace?~~

**Ex:** $S = \{f \in C^2[a,b] \mid f' + f'' = 0\}$

Q How to check if a subset is a subspace?

check

**Thm:** Any nonempty subset $W \subset V$, a vector space, is a subspace $\Leftrightarrow$ $W$ is closed under vector add'n & scalar multip'lcn

Pf: $\Rightarrow$) trivial

$\Leftarrow$) Say $W$ closed under scalar multiplication.

- Then for any $\vec{w} \in W$, $0\vec{w} = \vec{0}$; so $\exists$ additive identity
- Further, given any $\vec{w} \in W$, $(-1)\vec{w}$ is such that $\vec{w} + (-1)\vec{w} = \vec{0}$; $\exists$ additive inverse

etc., etc. ∎

exercises

Q Is the set of skew-symmetric matrices a subspace?

(20) Q If $U, V$ are subspaces, is $U \cap V$?

(9) Fix $A \in \mathbb{R}^{n\times n}$. Is $B$ s.t. $[A, B] = 0$ a subspace?

**Naturally-occurring subspaces:** (Nullspace & Span)

null space

**Defn:** [$A \in \mathbb{R}^{m\times n}$] We define the null space (aka the kernel) $N(A)$ of $A$ as

$$N(A) := \{\vec{x} \in \mathbb{R}^n \mid A\vec{x} = \vec{0}\}$$

Rk: Observe that $N(A)$ is a subspace.

finding the nullspace of a mx.

example — **Example:** $A = \begin{bmatrix} 1 & -2 & 1 & -1 \\ 2 & -3 & 4 & -3 \\ 3 & -5 & 5 & -4 \\ -1 & 1 & -3 & 2 \end{bmatrix}$ Reduces to $\begin{bmatrix} 1 & 0 & 5 & -3 \\ 0 & 1 & 2 & -1 \end{bmatrix}$

i.e., $x_1 + 5x_3 - 3x_4 = 0$

$x_2 + 2x_3 - x_4 = 0.$

Then a sol'n of $A\vec{x} = \vec{0}$ must be of the form $\begin{bmatrix} -5x_3 + 3x_4 \\ -2x_3 + x_4 \\ x_3 \\ x_4 \end{bmatrix}$ for some $x_3, x_4 \in \mathbb{R}$.

$$= x_3 \begin{bmatrix} -5 \\ -2 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 3 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

I.e. $N(A) = \text{Span}\left\{ \begin{bmatrix} -5 \\ -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ 0 \\ 1 \end{bmatrix} \right\}.$

## Span:

recall span, linear combo^n — **Recall:** Def^n: $[\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n \in V$, vector space$]$ i) Any vector of the form $\alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + \ldots + \alpha_n\vec{v}_n$ is called a linear combination of $\vec{v}_1, \ldots, \vec{v}_n$.

ii) $\text{Span}\{\vec{v}_1, \ldots, \vec{v}_n\} := \{$all linear combo^ns of $\vec{v}_1, \ldots, \vec{v}_n\}$

recall connxn w/ multiplication — **Ex:** (13a) (checking $\vec{b}$ in the span) $u = \begin{bmatrix} -1 \\ 2 \\ 3 \end{bmatrix}$, $v = \begin{bmatrix} 3 \\ 4 \\ 2 \end{bmatrix}$, $b = \begin{bmatrix} 2 \\ 6 \\ 6 \end{bmatrix}$ Is $\vec{b} \in \text{Span}\{\vec{u}, \vec{v}\}$?

Recall $b = \alpha u + \beta v \Leftrightarrow \begin{bmatrix} u & v \\ | & | \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} b \\ | \end{bmatrix}$; so just solve

easy exercise (review) — **Ex:** (19) $[A \in \mathbb{R}^{n\times n}]$ Show TFAE.

i) $N(A) = \{0\}$ || (do this part only)

ii) $A$ is invertible

iii) For each $\vec{b} \in \mathbb{R}^n$, $A\vec{x} = b$ has a unique sol'n.

Thm/Exercise: $[v_1, \ldots, v_n \in V]$ $\text{Span}\{v_1, \ldots, v_n\}$ is a subspace.

Pf: Trivial.

concrete exercises

**Exer.** Does $\operatorname{Span}\left\{\begin{bmatrix}1\\0\\1\end{bmatrix}, \begin{bmatrix}1\\1\\0\end{bmatrix}, \begin{bmatrix}0\\-1\\1\end{bmatrix}\right\} = \mathbb{R}^3$? I.e., "does the set span $\mathbb{R}^3$?"

I.e. "is the set a spanning set for $\mathbb{R}^3$"

The question is, given any $\begin{bmatrix}a\\b\\c\end{bmatrix}$,

is $\begin{bmatrix}a\\b\\c\end{bmatrix} = x_1\begin{bmatrix}1\\0\\1\end{bmatrix} + x_2\begin{bmatrix}1\\1\\0\end{bmatrix} + x_3\begin{bmatrix}0\\1\\1\end{bmatrix}$ for some coeffs. $x_1, x_2, x_3$?

$\iff$ Can we solve $\left[\begin{array}{ccc|c}1&1&0&a\\0&1&-1&b\\1&0&1&c\end{array}\right]$?

__Gaussian elimination:__ $\sim \left[\begin{array}{ccc|c}1&1&0&a\\0&1&-1&b\\0&-1&1&c-a\end{array}\right] \sim \left[\begin{array}{ccc|c}1&1&0&a\\0&1&-1&b\\0&0&0&c-a+b\end{array}\right]$

$\left(\sim \left[\begin{array}{ccc|c}1&0&1&a-b\\0&1&-1&b\\0&0&0&c-a+b\end{array}\right]\right)$

I.e., $x_1 = a-b+x_3$

$x_2 = b+x_3$

$0 = c-a+b$. Note if $c-a+b \neq 0$, __no__ sol'ns;

we have sol'ns only if $c-a+b=0$.

So no, the span is __not__ $\mathbb{R}^3$. (It spans $c = a-b$. $x_3 = x_1 - \ldots$

superfluous elements

(11c) Does $\operatorname{Span}\left\{\begin{bmatrix}-2\\1\end{bmatrix}, \begin{bmatrix}1\\3\end{bmatrix}, \begin{bmatrix}2\\4\end{bmatrix}\right\} = \mathbb{R}^2$?

$\iff$ can we always solve $\left[\begin{array}{ccc|c}-2&1&2&a\\1&3&4&b\end{array}\right]$?

$\sim \left[\begin{array}{ccc|c}1&3&4&b\\-2&1&2&a\end{array}\right] \sim \left[\begin{array}{ccc|c}1&3&4&b\\0&7&10&a+2b\end{array}\right] \sim \left[\begin{array}{ccc|c}1&0&-\frac{2}{7}&-\frac{3a}{7}+\ldots\\0&1&\frac{10}{7}&\frac{a+2b}{7}\end{array}\right]$

Answer is __yes__: in fact we always have infinitely many sol'ns.

In some sense, we have too many vectors in our spanning set.

## §3.3 Linear independence

goal: a coordinate system

Motivation: In some sense, our goal is (again) to create a "coordinate system" for our vector space. It will become evident that to do this, a minimal spanning set is needed (aka a basis)

need a minimal spanning set

Nonexample: $\longleftrightarrow$ (nonunique coordinates)

crucial idea: linear independence (minimality)

The crucial concept is the following:

Defn: [$\vec{v}_1, \ldots, \vec{v}_n \in V$ a vector space]

We say $\{\vec{v}_1, \ldots, \vec{v}_n\}$ are linearly independent if, whenever $c_1\vec{v}_1 + \ldots + c_n\vec{v}_n = \vec{0}$, we know $c_i = 0 \ \forall i$.

(ie, the only way to express the zero vector is the trivial one)

otherwise, we say the $\{v_i\}$ are linearly dependant.

examples

Examples: i) In $P_n$, $\{1, x, x^2, \ldots, x^{n-1}\}$ are linearly independent:

$c_0 + c_1 x + c_2 x^2 + \ldots + c_n x^{n-1} =$ zero poly'l implies all $c_i$ are $0$. (Taylor)

ii) $\{1, 1+x, x, x^2\}$ is linearly dependent:

$1 + (-1)(1+x) + x + 0x^2 \equiv 0$ (ie. is the zero poly'l)

So you see dependence $\Leftrightarrow$ a sort of redundancy;

independence $\Leftrightarrow$ no redundancy.

simple observation about dependence

Observation: $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ linearly dep't $\Leftrightarrow$ one of the $\vec{v}_i$ is a linear comb$^n$ of the others (is in the span of the others)

Pf: $\Rightarrow$) $c_1\vec{v}_1 + \ldots + c_n\vec{v}_n = \vec{0}$, for some $c_1, \ldots, c_n$ not all zero.

Assume, WLOG, $c_1 \neq 0$.

Then $c_1\vec{v}_1 = -c_2\vec{v}_2 + \ldots + -c_n\vec{v}_n$.

$\vec{v}_1 = -\frac{c_2}{c_1}\vec{v}_2 + \ldots + -\frac{c_n}{c_1}\vec{v}_n$.

$\Leftarrow$) Just as easy.

relat'n w/ coordinates' uniqueness

So linear dependence ~ redundancy; independence ~ minimality.

In fact, this minimality ensures a certain uniqueness of coordinates:

**Thm:** [$\vec{v}_1, \dots, \vec{v}_n \in V$, vector space]

$\{\vec{v}_1, \dots, \vec{v}_n\}$ is linearly indep$^t$ $\iff$ ~~every $\vec{v} \in \text{Span}\{\vec{v}_1, \dots, \vec{v}_n\}$~~ is uniquely expressible ~~can be written uniquely~~ as a linear comb$^n$ of the $\vec{v}_i$.

Pf: $\Rightarrow$) Assume $\{\vec{v}_1, \dots, \vec{v}_n\}$ independent, and let $\vec{v} \in \text{Span}\{\vec{v}_i\}$.

If $\vec{v} = c_1\vec{v}_1 + \dots + c_n\vec{v}_n$ and $\vec{v} = d_1\vec{v}_1 + \dots + d_n\vec{v}_n$,

$$\Longrightarrow \quad \vec{0} = (c_1 - d_1)\vec{v}_1 + \dots + (c_n - d_n)\vec{v}_n.$$

By linear independence, $c_i - d_i = 0$; $i = 1, \dots, n$

i.e. $c_i = d_i$; $i = 1, \dots, n$.

$\Leftarrow$) If every $\vec{v} \in \text{Span}\{\vec{v}_i\}$ has a unique expr'$\underline{n}$ as a linear comb$^n$ of the $\vec{v}_1, \dots, \vec{v}_n$,

then certainly $\vec{0}$ does.

Since $\vec{0} = 0\vec{v}_1 + \dots + 0\vec{v}_n$, this must be the only expression yielding $\vec{0}$. ☒

Rk: Thus if $\{\vec{v}_1, \dots, \vec{v}_n\}$ are linearly independent, each $\vec{v}$ in their span is associated with unique coords.

(Q) How does one check linear independence?

~~Example:~~ (2.) ~~Are~~ $\begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ -2 \end{bmatrix} \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix}$ linearly indpt?

$\iff$ Can we uniquely solve $\begin{bmatrix} 2 & 3 & 2 \\ 1 & 2 & 2 \\ -2 & -2 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \stackrel{?}{=} \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$?

Rk: We see this is actually the same thing (or at least closely related to) we studied before:

In general: Let $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n \in \mathbb{R}^n$, $A = [\vec{a}_1\, \vec{a}_2 \ldots \vec{a}_n]$

<u>Claim</u>: $\{\vec{a}_1, \ldots, \vec{a}_n\}$ are linearly indpt $\iff$ A is invertible.

Pf: $\{\vec{a}_1, \ldots, \vec{a}_n\}$ linearly indpt $\overset{?}{\iff}$

$\iff c_1\vec{a}_1 + \ldots + c_n\vec{a}_n = \vec{0}$ forces $c_i = 0\ \forall i$.

$\iff A\vec{c} = \vec{0}$ has only the trivial sol'n.

$\iff A$ is invertible.

Pt: So we have (yet) another way of seeing invertibility.

(Q. If A is invertible, so is $A^T$, and vice-versa. So ...)

<u>Note:</u> Read Example 6 on your own!

<u>Cool example</u>: Linear independence in $C^{(n-1)}[a,b]$: The Wronskian.

<u>Def'n</u>: Given $f_1, f_2, \ldots, f_n \in C^{(n-1)}[a,b]$, we define <u>the Wronskian</u> $W[f_1, \ldots, f_n] : [a,b] \to \mathbb{R}$

$$\text{by } W(x) = \begin{vmatrix} f_1(x) & f_2(x) & \ldots & f_n(x) \\ f_1'(x) & f_2'(x) & \ldots & f_n'(x) \\ \vdots & & & \\ f_1^{(n-1)}(x) & & \ldots & f_n^{(n-1)}(x) \end{vmatrix}$$

<u>Thm</u>: If there exits even <u>one</u> point $x_0 \in [a,b]$ where $W(x_0) \neq 0$, then the $f_i$ are linearly independent

Pf: By contraposition: we show, if the $\{f_1, \ldots, f_n\}$ are linearly dependent then $W[f_1, \ldots, f_n](x) \equiv 0$ for all $x \in [a,b]$.

(Proof trivial)

## §3.4 Basis and Dimension. (pretty empty section)

Motivation: Given a vector space $V$, if we can find a spanning set $\{v_i\}$ that is also linearly independent, then each $v \in V$ can be uniquely expressed as a linear combination of the $\{v_i\}$.

Rk: In fact these are equivalent, as we saw.

basis

Def'n: If i. $\{v_1, \ldots, v_n\}$ span $V$ ii. $\{v_1, \ldots, v_n\}$ are linearly indpt $\Big\} \Rightarrow$ we say $\{v_i\}$ is a basis of $V$.

example

Ex: The null space of $\begin{bmatrix} 1 & 0 & 5 & -3 \\ 0 & 1 & 2 & -1 \end{bmatrix}$ : $\begin{aligned} x_1 + 5x_3 - 3x_4 &= 0 \\ x_2 + 2x_3 - x_4 &= 0 \end{aligned}$

$\longrightarrow$ Sol'n must look like $\begin{bmatrix} -5x_3+3x_4 \\ -2x_3+x_4 \\ x_3 \\ x_4 \end{bmatrix} = x_3 \begin{bmatrix} -5 \\ -2 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 3 \\ 1 \\ 0 \\ 1 \end{bmatrix}$

ie. $N(A) = \text{span}\left\{ \begin{bmatrix} -5 \\ -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ 0 \\ 1 \end{bmatrix} \right\}$. These vectors are linearly indpt, so they form a basis.

dimension

Def'n: $V$ a vector space, $\{v_i\}$ any basis of $V$. We define the dimension of $V$ by $\dim(V) = \#\{v_i\}$.

not yet well-defined

Rk: You should be unhappy with this defn. What's wrong w/ it?

Point: How do we know all bases have the same cardinality?

# indpt ≤ # spanning

Thm: Say $\{v_1, \ldots, v_n\}$ span $V$, and suppose $m > n$.

Claim: Any set of vectors $\{u_1, \ldots, u_m\}$ must be linearly dependent.

Rk: Contrapositive: If $\{u_1, \ldots, u_m\}$ be linearly indpt. $\Rightarrow m \le n$ (# indpt ≤ # spanning)

Commentary: a bit tricky or abstract

Pf: WTS the $\{u_i\}$ are linearly dependent, i.e. $\exists$ coeffs $\hat{c}_1, \dots, \hat{c}_m$ (not all zero) s.t. $\sum_{i=1}^{m} \hat{c}_i u_i = \vec{0}$

Can we create such a combination?

Consider: Each $\vec{u}_i = \sum_{j=1}^{n} a_{ij} \vec{v}_j$; so any comb$^n$

$$\sum_{i=1}^{m} c_i \vec{u}_i = \sum_{i=1}^{m} c_i \sum_{j=1}^{n} a_{ij} \vec{v}_j$$

$$= \sum_{j=1}^{n} \left( \sum_{i=1}^{m} a_{ij} c_i \right) \vec{v}_j$$

Now, the system $\sum_{i=1}^{m} a_{ij} c_j = 0$; $j = 1, 2, \dots, n$

is a homogeneous system with $m$ unknowns, $n$ eq'ns.

$\Rightarrow$ has nontrivial solutions $\blacksquare$

corollary: dim(V) well-defined

Corollary: If $\{v_1, \dots, v_n\}$ and $\{u_1, \dots, u_m\}$ are bases of $V \Rightarrow n = m$

Rk: Thus dim(V) actually is well defined.

trivial facts about bases.

Trivial Theorem: Let $V$ be a vector space, $\dim(V) = n$, $\{v_1, \dots, v_n\}$ a set of vectors in $V$.

Then $\{v_1, \dots, v_n\}$ are linearly indpt $\Leftrightarrow$ $\{v_i\}$ span $V$.

(I.e. any independent set of the right size is a basis; any spanning set of the right size must be indpt.)

Pf: Trivial (do as exercise in class.) (by $*$, say)

Corollary: i) any linearly indpt. set can be extended to a basis

ii) any spanning set can be reduced to a basis

The Exercise are unavoidably awful...

## §3.5 Change of Basis (Change of coordinate systems)

Coordinate systems:

Let $V$ be a vector space and $\mathcal{B} = \{\vec{b}_1, \ldots, \vec{b}_n\}$ a basis of $V$.

Then for each $\vec{v} \in V$, there is a unique expression

$\vec{v} = \alpha_1 \vec{b}_1 + \alpha_2 \vec{b}_2 + \ldots + \alpha_n \vec{b}_n$ of $\vec{v}$ as a linear comb'n

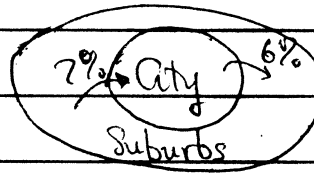
coord vector — We call $(\alpha_1, \alpha_2, \ldots, \alpha_n)^T$ the coordinate vector w.r.t. $\mathcal{B}$.

Rk: In some texts, this is denoted $\vec{v}|_{\mathcal{B}}$

useful to change coord sys. — It will occasionally be elucidating to change coord. systems. from one basis to another

**Example:** Markov chain application.

example (Markov process)




Initial state: 30% of total population in city.

Q. What happens in the long run?

We model this system as follows:

Let $v = \begin{bmatrix} c \\ s \end{bmatrix}$ where $c$ = percent in city "state vector"

$s = 1 - c$ = percent in suburbs $\begin{bmatrix} .3 \\ .7 \end{bmatrix}$ for us

Let $A = \begin{bmatrix} .94 & .02 \\ .06 & .98 \end{bmatrix}$ be the so-called transition mx

Q Think: What's $A\vec{v}$? $\begin{bmatrix} .94c + .02s \\ .06c + .98s \end{bmatrix}$ the new state after 1 year

Q What's the state after n years?

limit — It turns out that $\lim_{n \to \infty} A^n \begin{bmatrix} .3 \\ .7 \end{bmatrix} = \begin{bmatrix} .25 \\ .75 \end{bmatrix}$

i.e., with time, .25 will end up in the city.

steady state vector — Rk: What's $A \begin{bmatrix} .25 \\ .75 \end{bmatrix}$? That is a steady state vector

This result can be clarified using change of basis.

Allow us to magically choose two vectors $\vec{u}_1 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$ $\vec{u}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$.

eigenvectors

Rk: $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$ is a multiple of $\begin{bmatrix} .25 \\ .75 \end{bmatrix}$ and as such is fixed by $A$;

$$A\vec{u}_2 = \dots = .92\vec{u}_2. \quad \text{(eigenvectors)}$$

change to eigenvector basis

If we write $\begin{bmatrix} .3 \\ .7 \end{bmatrix}$ in terms of $\vec{u}_1, \vec{u}_2$ (rather than $\vec{e}_1, \vec{e}_2$ as we implicitly have)

$$\begin{bmatrix} .3 \\ .7 \end{bmatrix} = .25\vec{u}_1 - .05\vec{u}_2,$$

Then the phenomenon becomes clear:

$$A^n\left(.25\vec{u}_1 - .05\vec{u}_2\right) = .25\vec{u}_1 - .05\underline{(.92)^n\vec{u}_2} \to \vec{0}.$$

Insightful comment:

It will be useful to choose (if possible!) a basis composed of vectors for which $A$'s behavior is simple.

change of coords & matrices

Matrices & Change of Coordinates: (From $\mathcal{B}$ to $\mathcal{V}$)

Let $\mathcal{B} = \{b_1, \dots, b_n\}$ and $\mathcal{V} = \{v_1, \dots, v_n\}$ be bases of $\mathbb{R}^n$.

Suppose we know $[x]_{\mathcal{B}} = (c_1\ c_2 \dots c_n)^T$. What is $x_{\mathcal{V}}$, then?

Well, think: $x = c_1 b_1 + c_2 b_2 + \dots + c_n b_n$

$$= [b_1\ b_2 \dots b_n]\, x_{\mathcal{B}}$$

We want $x_{\mathcal{V}}$ so that $x = [v_1\ v_2 \dots v_n]\, x_{\mathcal{V}}$.

So it's easy: $\boxed{x_{\mathcal{V}} = [v_1\ v_2 \dots v_n]^{-1}[b_1 \dots b_n]\, x_{\mathcal{B}}}$

change of coords. example

**Example:** (5) $u_1 = \begin{pmatrix}1\\1\\1\end{pmatrix}$, $u_2 = \begin{pmatrix}1\\2\\2\end{pmatrix}$, $u_3 = \begin{pmatrix}2\\3\\4\end{pmatrix}$

Write $\begin{pmatrix}3\\2\\5\end{pmatrix}$ in terms of $\{u_1, u_2, u_3\}$.

Well, want to solve $\begin{bmatrix}1&1&2\\1&2&3\\1&2&4\end{bmatrix}\begin{bmatrix}a\\b\\c\end{bmatrix} = \begin{bmatrix}3\\2\\5\end{bmatrix}$

Inverting $U$, we get $\begin{bmatrix}a\\b\\c\end{bmatrix} = \begin{bmatrix}2&0&-1\\-1&2&-1\\0&-1&1\end{bmatrix}\begin{bmatrix}3\\2\\5\end{bmatrix} = \begin{bmatrix}1\\-4\\3\end{bmatrix}$

Rechecking, we see $u_1 + -4u_2 + 3u_3 = \begin{pmatrix}3\\2\\5\end{pmatrix}$.

exerc.

**Exerc:** $E = \left\{\begin{pmatrix}5\\3\end{pmatrix}, \begin{pmatrix}3\\2\end{pmatrix}\right\}$, $z = \begin{pmatrix}10\\7\end{pmatrix}$. What's $[z]_E$?

change of basis mx

**Example** (from book): Find change of basis mx

from $B = \left\{\begin{pmatrix}5\\2\end{pmatrix}, \begin{pmatrix}7\\3\end{pmatrix}\right\}$ to $B' = \left\{\begin{pmatrix}3\\2\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix}\right\}$.

Given $[v]_B = \begin{pmatrix}a\\b\end{pmatrix}$, know $\begin{pmatrix}5&7\\2&3\end{pmatrix}\begin{pmatrix}a\\b\end{pmatrix} = \begin{pmatrix}3&1\\2&1\end{pmatrix}\begin{pmatrix}a'\\b'\end{pmatrix}$

$\rightarrow \begin{pmatrix}1&-1\\-2&3\end{pmatrix}\begin{pmatrix}5&7\\2&3\end{pmatrix}\begin{pmatrix}a\\b\end{pmatrix} = \begin{pmatrix}a'\\b'\end{pmatrix}$.

not'n: $[v]_B$

**Notation:** $V$ a vector space, $E = \{v_1, \ldots, v_n\}$ a basis of $V$.

Given any $v \in V$, we define $[v]_E$, the coordinate vector of $v$ w.r.t. $E$,

as $(c_1, \ldots, c_n)^T \in \mathbb{R}^n$

(where $v = c_1 v_1 + \ldots + c_n v_n$)

(is the unique exp'n of $v$ as a linear comb'n of the $v_i$.)

**Example:** $x^2$ in $\{1, 2x, 4x^2-2\}$ is $\left(\frac{1}{2}, 0, \frac{1}{4}\right)^T$.

abstract vector spaces

Change of basis for abstract vector spaces:

Let $V$ be an $n$-dim'l vector space, $\mathcal{B}$ and $\mathcal{B}'$ two bases $\mathcal{B} = \{v_1, \ldots, v_n\}$, $\mathcal{B}' = \{w_1, \ldots, w_n\}$

Suppose
$$v_1 = s_{11} w_1 + s_{21} w_2 + \ldots + s_{n1} w_n$$
$$v_2 = s_{12} w_1 + \ldots + s_{n2} w_n$$
$$\vdots$$
$$v_n = s_{1n} w_1 + \ldots + s_{nn} w_n$$
, let $S = (s_{ij})$.

(Rk: The indices are transposed.)

Q. How do we get from $[v]_{\mathcal{B}}$ to $[v]_{\mathcal{B}'}$?

transition matrix

Claim: $[v]_{\mathcal{B}'} = S[v]_{\mathcal{B}}$ ($S$ is the transition mx.)

Why? Say $[v]_{\mathcal{B}} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, i.e.

$$v = x_1 v_1 + x_2 v_2 + \ldots + x_n v_n.$$
$$= x_1 (s_{11} w_1 + s_{21} w_2 + \ldots + s_{n1} w_n)$$
$$+ x_2 (s_{12} w_1 + \ldots + s_{n2} w_n) + \ldots + x_n (s_{1n} w_1 + \ldots + s_{nn} w_n)$$
$$= \sum_{i=1}^{n} \left( \sum_{j=1}^{n} s_{ij} x_j \right) w_i \text{ ; so if } [v]_{\mathcal{B}'} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

we see $y_i = \sum_{j=1}^{n} s_{ij} x_j$.

i.e. $[v]_{\mathcal{B}'} = S[v]_{\mathcal{B}}$.

example

Exer: Going from $[1, x, x^2]$ to $[1, 2x, 4x^2-2]$

$$1 = 1 + 0 \cdot 2x + 0 \cdot (4x^2-2)$$
$$x = 0 \cdot 1 + \tfrac{1}{2} \cdot 2x + 0 (4x^2-2)$$
$$x^2 = \tfrac{1}{2} \cdot 1 + 0 \cdot 2x + \tfrac{1}{4} (4x^2-2)$$

So $S = \begin{bmatrix} 1 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{4} \end{bmatrix}$

(Do a concrete example)

$\mathcal{B} = \{v_1, \ldots, v_n\}$, $\mathcal{B}' = \{w_1, \ldots, w_n\}$

dumb? observations

**Observations:** $S$ a transition mx $\Rightarrow$ $S$ is invertible

~~Let's think. Given $v \in V$, we have~~

$$[v]_{\mathcal{B}'} = S[v]_{\mathcal{B}}.$$

$$(\Leftrightarrow \quad y_1 w_1 + \ldots + y_n w_n = x_1 v_1 + \ldots + x_n v_n)$$

~~Well, when could $S\mathbf{x} = \mathbf{0}$?~~

~~That would~~ $\Leftrightarrow$ ~~$0 = x_1 v_1 + \ldots + x_n v_n$.~~

$\Leftrightarrow \quad x_1 = \ldots = x_n = 0.$ by linear indpce.

Thus $S$ must be invertible.

**Conversely:** $S$ invertible $\Rightarrow$ it can be used to generate a new basis.

~~For suppose $\{v_1, \ldots, v_n\}$ is a basis of $V$.~~

~~Then $\{S^T v_1, \ldots, S^T v_n\}$ is linearly indpt $\Rightarrow$ still a basis.~~

~~Then $S$ is the matrix transitioning from $\{v_1, \ldots, v_n\}$ to $\{w_i = S^T v_i\}$.~~

## §3.6 Row Space, Column Space: The Rank Theorem

column, row space rank

**Defn:** Let $A \in \mathbb{R}^{m\times n}$, i.e. $A$ be an $m\times n$ matrix.

i) We call span$\{a_1, a_2, \ldots, a_n\}$ the column space of $A$;

ii) We call span$\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m\}$ the row space of $A$, and denote by rank$(A)$ its dimension.

Simple observation: If $A \sim B$ then they have identical row spaces.

finding the rank via reduction

Observe: Thus we can determine a basis for the row space of a mx. via Gaussian elementary

Ex. $\begin{bmatrix} -3 & 1 & 3 & 4 \\ 1 & 2 & -1 & -2 \\ -3 & 8 & 4 & 2 \end{bmatrix} \sim \begin{bmatrix} -3 & 1 & 3 & 4 \\ 1 & 2 & -1 & -2 \\ 0 & 7 & 1 & -2 \end{bmatrix}$

$\sim \begin{bmatrix} 0 & 7 & 0 & -2 \\ 1 & 2 & -1 & -2 \\ 0 & 7 & 1 & -2 \end{bmatrix} \sim \begin{bmatrix} 0 & 7 & 0 & -2 \\ 1 & 2 & -1 & -2 \\ 0 & 0 & 1 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & -1 & -2 \\ 0 & 7 & 0 & -2 \\ 0 & 0 & 1 & 0 \end{bmatrix}$

These mxs. have identical row spaces, viz. Span$\left\{\begin{matrix} (1,2,-1,-2) \\ (0,7,0,-2) \\ (0,0,1,0) \end{matrix}\right\}$.

~~(Pf: Rank(A) = # pivots in reduced row-echelon form)~~

[Some brief comments on linear systems ...]

dumb observation

**Thm:** $A \in \mathbb{R}^{m\times n}$ The column vectors of $A$ are linearly indpt.

$\iff$ $Ax = b$ has at most one sol'n for every $b \in \mathbb{R}^m$

Pf. $\Rightarrow$) Suppose $Ax = b$ had 2 sol'ns, $x_1$ and $x_2$.

Then $A(x_1 - x_2) = 0 \Rightarrow$ columns dependent.

$\Leftarrow$) By hypothesis, $Ax = 0$ has one sol'n.

Corollary: $A \in \mathbb{R}^{n\times n}$ is invertible $\iff$ column vectors form a basis

Observations | Let's return to the example for a moment:

$$A = \begin{bmatrix} -3 & 1 & 3 & 4 \\ 1 & 2 & -1 & -2 \\ -3 & 8 & 4 & 2 \end{bmatrix} \sim \dots \sim \begin{bmatrix} 1 & 0 & 0 & -10/7 \\ 0 & 1 & 0 & -2/7 \\ 0 & 0 & 1 & 0 \end{bmatrix} ; \ \underline{\text{rank}(A) = 3}$$

# of pivots

The null space: is one dim'l

$$x_1 = 10/7\, x_4 \qquad x_2 = 2/7\, x_4 \qquad x_3 = 0 \quad ; \quad \left\{ r \begin{bmatrix} 10/7 \\ 2/7 \\ 0 \\ 1 \end{bmatrix} ; r \in \mathbb{R} \right\}$$

nullity | ie, the nullity is 1.

The rank theorem | Thm. $A \in \mathbb{R}^{m \times n} \Rightarrow \text{rank}(A) + \text{nullity}(A) = n$

Pf: Use the reduced row echelon form exactly as above

[Rk: Proofs by reduced row-echelon form are generally distasteful]

Column space | (Q) How come we don't care about dim(column space)?

dim of = rank. | Thm: dim(column space of $A$) = rank($A$)

Pf: Begin by observing that if $A \sim U$, then $Ax = 0 \Leftrightarrow Ux = 0$.

that is, $\boxed{\text{any dependence between } A\text{'s columns is a dependence between } U\text{'s columns}}$

So let $U$ be the reduced row echelon form of $A$.

Observe that any column w/o pivots is expressible as a linear comb'n of those with pivots; i.e. is in the span of those with pivots.

Thus dim(column space) = # of pivots = rank($A$) ∎

Rk: The columns in $A$ corresponding to the pivot columns in $U$ form a basis of the column space (Ex. above)

(Do example 5!)

example

Example 5: Find the dimension of the subspace spanned by

$$x_1 = \begin{bmatrix} 1 \\ 2 \\ -1 \\ 0 \end{bmatrix}, \quad x_2 = \begin{bmatrix} 2 \\ 5 \\ -3 \\ 2 \end{bmatrix} \quad x_3 = \begin{bmatrix} 2 \\ 4 \\ -2 \\ 0 \end{bmatrix} \quad x_4 = \begin{bmatrix} 3 \\ 8 \\ -5 \\ 4 \end{bmatrix}$$

Well,

$$\text{span}\{x_1, \dots, x_4\} = \begin{bmatrix} 1 & 2 & 2 & 3 \\ 2 & 5 & 4 & 8 \\ -1 & -3 & -2 & -5 \\ 0 & 2 & 0 & 4 \end{bmatrix} \text{'s column space.}$$

$$X \sim A = \begin{bmatrix} 1 & 2 & 2 & 3 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Notice $a_3 = 2a_1$,
$a_4 = a_1 + a_2$.
and $a_1, a_2$ are indpt.

Thus $x_3, x_4 \in \text{Span}\{x_1, x_2\}$ which are indpt.; (the dim^n of the span is 2).

exercises

Exercises:

1a. Find bases for the row space, column space, and null space of $\begin{bmatrix} 1 & 3 & 2 \\ 2 & 1 & 4 \\ 4 & 7 & 8 \end{bmatrix}$.

24. Say $A, B \in \mathbb{R}^{n\times n}$ and $N(A-B) = \mathbb{R}^n$.
Then $A = B$. Proof: $\text{rank}(A-B) = 0$.

10. Let $A \in \mathbb{R}^{m\times n}$ and say $\text{rank}(A) = n$.
Suppose $Ac = Ad$. Does that imply $c = d$?
(Well, $\text{nullity}(A) = 0$; and $A(c-d) = \vec{0}$)
Suppose $\text{rank}(A) < n$. Must $c = d$? No - as long as $c - d \in$ nullspace, ok.

26.

§4.1 Linear Transformations:

Defn. [V, W vector spaces] $T: V \longrightarrow W$ a map from V into W.

If i. $T(\vec{v}_1 + \vec{v}_2) = T(\vec{v}_1) + T(\vec{v}_2)$ $\quad \forall \vec{v}_1, \vec{v}_2 \in V$

linear xform$^n$ ii. $T(\alpha \vec{v}_1) = \alpha T(\vec{v}_1)$ $\quad \forall \vec{v}_1 \in V$

Then we say T is a linear transformation from V to W.

(Rk: Preserves algebraic operations)

examples Examples:

mxs. i) Any matrix $A \in \mathbb{R}^{m \times n}$ can be used to define a linear transform$^n$ from $\mathbb{R}^n$ to $\mathbb{R}^m$ by $T(\vec{v}) := A\vec{v}$.

(Certainly $A(\vec{v}_1 + \vec{v}_2) = A\vec{v}_1 + A\vec{v}_2$ etc.)

proj'n ii) The map $L: \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ is a linear xform$^n$, for
$(x_1, x_2) \mapsto (x_1, 0)$

certainly $L((x_1, x_2) + (y_1, y_2)) = x_1 + y_1 = L(x_1, x_2) + L(y_1, y_2)$

$L(\alpha x_1, \alpha x_2) = \alpha x_1 = \alpha L(x_1, x_2)$

Rk. This is projection onto the first axis, and could be realized with the mx $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$.

rot'n iii) Is 'rotation by angle $\theta$' a linear xform$^n$ on $\mathbb{R}^2$?

(Give a geometric argument)

Rk: In fact, this corresponds to multiplication by $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$

Q Can all linear transform$^n$s be realized as matrices?

diff'n. iv) Is differentiation a linear transform$^n$ on $P_n$?

i.e. i) $D(p_1 + p_2) = Dp_1 + Dp_2$? (What matrix? ??)

ii) $D(\alpha p) = \alpha Dp$?

Basic consequences/properties of linear xform^ns:

Basic props

Thm: $T: V \to W$ a linear xfrm^n. Then
i) $L(\vec{0}_V) = \vec{0}_W$. (more preservation of vector space structure: 0 and additive inverse)
ii) $L(-\vec{v}) = -L(\vec{v})$

Pf: i. $L(\alpha\vec{v}) = \alpha L(\vec{v})$; take $\alpha = 0$.
ii. $L(\vec{v}) + L(-\vec{v}) = L(\vec{v} - \vec{v}) = L(\vec{0}_V) = \vec{0}_W$.
By uniqueness of additive inverses, $L(-\vec{v}) = -L(\vec{v})$.

More basic defns: Kernel and Image.

Defns: $T: V \to W$ a linear xfrm^n.

kernel

i) $\ker(T) := \{\vec{v} \in V \mid T\vec{v} = \vec{0}\}$ (null space)

image (range)

ii) [S a subspace of V] $T(S) := \{\vec{w} \in W \mid \vec{w} = T(\vec{v})$ for some $\vec{v} \in S\}$
is called the image of S (under T)
Rk: More common is to define $\mathrm{Im}(T) = T(V)$;
in this text they call this the range of T.

Simple observations: $\ker(T)$ and $T(S)$ are both subspaces,
i.e. closed under addn & scalar multiplication

examples

Ex: Do examples w/ (ii), (iii), (iv) above.

exercs.

Exercises: 21. $T: V \to W$ is called one-to-one if $T(\vec{v}_1) = T(\vec{v}_2)$ implies $\vec{v}_1 = \vec{v}_2$.
Claim: T is 1-1 $\iff$ $\ker(T) = \{\vec{0}_V\}$.

19. Find kernel & range of $T(p(x)) = xp'(x)$ on $P_3$.

13. Show that if $T_1, T_2$ agree on a basis of V,
then they agree on every vector in V.

## §4.2 Matrix Representations of Linear Transformations

matrix rep'ns of linear xform'n between Euclidean vector spaces.

Rk: mx rep'n of T w.r.t. the standard bases.

<u>Recall</u>: Agreement on basis $\Rightarrow$ Agreement everywhere. (linearity constraining)

<u>Recall</u>: Given any mx $A \in \mathbb{R}^{n \times m}$ we have a corresponding linear transformation $T_A : \mathbb{R}^m \to \mathbb{R}^n$ defined by $v \mapsto Av$.

In fact, the converse is true: Given any linear xform'n $\mathbb{R}^m \to \mathbb{R}^n$, there is a matrix $A$ s.t. $Tv = Av \ \forall v$. (i.e. $T = T_A$.)

<u>Thm</u>: $T : \mathbb{R}^n \to \mathbb{R}^m$ a linear xform'n $\Rightarrow \exists A \in \mathbb{R}^{m \times n}$ s.t. $T = T_A$.

(Rk: We often conflate $A$ and $T_A$)

<u>Pf</u>: Let $\{e_1, \ldots, e_n\}$ denote (as always) the standard basis in $\mathbb{R}^n$.

Let $a_i := Te_i \in \mathbb{R}^m$, and $A = [a_1 \, a_2 \ldots a_m]$

Then $T_A e_i = Ae_i = a_i \ \forall i$, so $T_A$ agrees with $T$ on a basis. ∎

<u>Example</u> (36): $T\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_1 + x_2 \\ x_1 + x_2 + x_3 \end{pmatrix}$. What's $A$? $\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$

Then $\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_1 + x_2 \\ x_1 + x_2 + x_3 \end{pmatrix}$, as desired.

Rk: Thus $L(\mathbb{R}^n, \mathbb{R}^m)$ is basically $\mathbb{R}^{m \times n}$.

Q. What about for abstract vector spaces? Would this work?

For abstract vector spaces

## Matrix rep'ns of linear xform$^{ns}$ between (abstract) vector spaces:

**Thm:** $V, W$ vector spaces, $T: V \to W$ a linear xform$^n$.

$E = \{\vec{v}_1, \ldots, \vec{v}_n\}$; $F = \{\vec{w}_1, \ldots, \vec{w}_m\}$ bases for $V, W$, resp.

defn of mx rep'n

Claim: $\exists$ a matrix $A \in \mathbb{R}^{m\times n}$ s.t. $A[\vec{v}]_E = [T(\vec{v})]_F \; \forall \vec{v} \in V$;

we call $A$ the matrix rep'n of $T$ (w.r.t. bases $E$ and $F$).

Picture:

$$\begin{tikzcd} \vec{v} \quad V \arrow[r, "T"] \arrow[d, "P"', "\cong"] & W \quad \vec{w} \arrow[d, "Q"', "\cong"] \\ {[\vec{v}]_E} \quad \mathbb{R}^n \arrow[r, dashed, "T_A"'] & \mathbb{R}^m \quad {[\vec{w}]_F} \end{tikzcd}$$

The claim is that $\exists\, T_A$ s.t. the diagram commutes, i.e. $QT\vec{v} = T_A P\vec{v} \;\; \forall \vec{v}$.

constructing the mx. rep'n

Pf: Let $A = [\mathbf{a}_1 \ldots \mathbf{a}_n]$ where $\mathbf{a}_i = [Tv_i]_F$.

Notice (composition) $QT\vec{v}_i = [T\vec{v}_i]_F$

and $T_A P\vec{v}_i = T_A e_i = \mathbf{a}_i = [T\vec{v}_i]_F$.

In other words, $QT$ and $T_A P$ agree on $\{\vec{v}_1, \ldots, \vec{v}_n\}$, a basis

$\Rightarrow QT\vec{v} = T_A P\vec{v} \;\; \forall \vec{v}$. ∎

What's really going on?

Rk: What's really going on here?

In fact, $P$ and $Q$ are isomorphisms (and thus invertible).

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[d, "P"', "\cong"] & W \arrow[d, "\cong", "Q"'] \\ \mathbb{R}^n & \mathbb{R}^m \end{tikzcd}$$

Consider $QTP^{-1}: \mathbb{R}^n \to \mathbb{R}^m$.

$A$ is the mx. rep'n of $QTP^{-1}$:

$$\begin{aligned} \mathbf{a}_i &= QTP^{-1}e_i \\ &= QTv_i \\ &= [Tv_i]_F. \end{aligned}$$

(Rk: On book's presentation)

Example 5: $V = P_3$, $W = P_2$; $T: V \to W$ the diff'n op'tor.
What's the matrix rep'n w.r.t. the std bases
$$E = \{1, x, x^2\}, \quad F = \{1, x\} \; ?$$

$$\begin{array}{ccc} V & \xrightarrow{T} & W \\ \downarrow & & \downarrow \\ \mathbb{R}^3 & \dashrightarrow & \mathbb{R}^2 \end{array}$$

We saw $A$ is given by
$$a_i = [Tv_i]_F$$

Well, $Tv_1 = \frac{d}{dx} 1 = 0$ ; $[0]_F = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$
$Tv_2 = \frac{d}{dx} x = 1$ ; $[1]_F = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$
$Tv_3 = \frac{d}{dx} x^2 = 2x$ ; $[2x] = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$.

So $A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$.

Parallel behavior: $\frac{d}{dx}(3x + 4x^2) = 3 + 8x$ $(T\vec{v})$

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 3 \\ 8 \end{bmatrix}.$$

$$T_A [\vec{v}]_E = A[\vec{v}]_E = [T\vec{v}]_F.$$

Exercise: 15. Let $S = \text{Span}\{e^x, xe^x, x^2e^x\}$, $D = \frac{d}{dx}$ on $S \to S$.
Find the mx rep'n of $D$ w.r.t. $\{e^x, xe^x, x^2e^x\}$

16. $T$ a linear operator on $V \to V$, $\ker T \neq \{0\}$.
Show: if $A$ is the mx rep'n, that $A$ is noninvertible.

6

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

Rk: On the book's presentation:

$T: V \to W$ a linear xform$^n$, $E = \{\vec{v}_1, \dots, \vec{v}_n\}$ a basis of $V$,

$F = \{\vec{w}_1, \dots, \vec{w}_m\}$ a basis of $W$.

For any $\vec{v} \in V$, know $\vec{v} = x_1\vec{v}_1 + \dots + x_n\vec{v}_n$.

Claim: $\exists A \in \mathbb{R}^{m \times n}$ ~~that represents T in the sense that~~

$$A\mathbb{x} = \mathbb{y} \Leftrightarrow T\vec{v} = y_1\vec{w}_1 + \dots + y_m\vec{w}_m.$$

But this is the same statement: $A[\vec{v}]_E = [T\vec{v}]_F$.

Notice that we could just as well change the basis for $T: \mathbb{R}^n \to \mathbb{R}^m$:

One particular instance of the theorem.

Thm: $T: \mathbb{R}^n \to \mathbb{R}^m$ a linear xform$^n$,

$E = \{\vec{v}_1, \dots, \vec{v}_n\}$ a basis of $\mathbb{R}^n$, $F = \{\vec{w}_1, \dots, \vec{w}_m\}$, one of $\mathbb{R}^m$

Claim: The matrix rep'n of $T$ w.r.t. $E$ and $F$ is

$$A = [a_1, \dots, a_n] \text{ where } a_i = \cancel{B} W^{-1}T(v_i).$$

$i = 1, \dots, n.$

$$(W = [w_1 \dots w_m])$$

Pf: Just apply the theorem. $a_i = [Tv_i]_F$

$= [w_1 \dots w_m]^{-1} Tv_i$.

dumb comment

Rk: Thus $A = [w_1 \dots w_m]^{-1}[Tv_1\ Tv_2 \dots Tv_n]$, is the mx rep'n of $T$ w.r.t. $E$ & $F$.

~~which one could calculate by reducing:~~

$$[w_1 \dots w_m \mid Tv_1 \dots Tv_n]$$

$$\sim \dots \sim [I \mid A]$$

(See Ex. 6)
(Do 18a)

Rk: The point is mainly that $T$ can have different rep'ns depending on choice of bases.

18a. $E = \{u_1, u_2, u_3\}$, $F = \{b_1, b_2\}$

$$u_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \quad u_2 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \quad u_3 = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} ; \quad b_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad b_2 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}.$$

Say $T(x) = \begin{bmatrix} x_3 \\ x_1 \end{bmatrix}$, a linear xform$^n$ $\mathbb{R}^3 \to \mathbb{R}^2$

What's the mx. rep'n of $T$ w.r.t. $E, F$?

$$A = \begin{bmatrix} b_1 & b_2 \end{bmatrix}^{-1} \begin{bmatrix} Tu_1 & Tu_2 & Tu_3 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 2 \\ -1 & -1 \end{bmatrix}^{-1} \begin{bmatrix} -1 & 1 & 1 \\ 1 & 1 & -1 \end{bmatrix} = \begin{bmatrix} -1 & -2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} -1 & 1 & 1 \\ 1 & 1 & -1 \end{bmatrix}$$

$$= \begin{bmatrix} -1 & -3 & 1 \\ 0 & 2 & 0 \end{bmatrix}.$$

Ex:

$$\begin{tikzcd}
u_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \arrow[r, mapsto, "T"] \arrow[d] & \begin{bmatrix} -1 \\ 1 \end{bmatrix} = Tu_1 \arrow[d] \\
{[u_1]_E} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \arrow[r, mapsto, "T_A"] & \begin{bmatrix} -1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \end{bmatrix} = [Tu_1]_F
\end{tikzcd}$$

$\checkmark$

$\begin{bmatrix} -1 \\ 0 \end{bmatrix} = T_A([u_1]_E)$

and one can check agreement on $u_2, u_3$.

## §4.2 Secret mathematical knowledge

agree on basis $\Rightarrow$ same xformn

(13) <u>Lemma:</u> If $T_1, T_2 : V \to W$ are linear transformations that agree on a basis of $V$, then they agree on every vector of $V$.

<u>Pf:</u> Let $\{\vec{b}_1, \dots, \vec{b}_n\}$ be a basis of $V$; say $T_1(\vec{b}_i) = T_2(\vec{b}_i)\ \forall i$.

Then for any $\vec{v} = \sum_{i=1}^n c_i \vec{b}_i$ in $V$,

$$T_1\vec{v} = T_1\left(\sum_1^n c_i \vec{b}_i\right) = \sum_{i=1}^n c_i T_1 \vec{b}_i \quad \text{by linearity}$$

$$= \sum c_i T_2 \vec{b}_i \quad \text{by hypothesis}$$

$$= T_2\left(\sum c_i \vec{b}_i\right) = T_2\vec{v} \quad \text{by linearity.}$$

1-1. (21) <u>Defn:</u> $T : V \to W$ a linear xformn. is called <u>one-to-one</u> if $T\vec{v}_1 = T\vec{v}_2$ implies $\vec{v}_1 = \vec{v}_2$.

(equiv., if $\vec{v}_1 \neq \vec{v}_2$ implies $T\vec{v}_1 \neq T\vec{v}_2$).

1-1 $\Leftrightarrow$ $\ker T = \{0\}$.

<u>Claim:</u> $T$ is 1-1 $\Leftrightarrow$ $\ker(T) = \{\vec{0}_V\}$.

<u>Pf:</u> $\Rightarrow$) $T(\vec{0}_V) = \vec{0}_W$, so $\vec{0}_V \in \ker(T)$.

And since $T$ is 1-1, if $\vec{v} \in \ker(T)$, i.e. $T\vec{v} = \vec{0}_W$, we know $\vec{v} = \vec{0}_V$. So $\ker(T) = \{\vec{0}_V\}$

$\Leftarrow$) Suppose $\ker(T) = \{\vec{0}_V\}$.

If $T\vec{v}_1 = T\vec{v}_2$, then $\vec{v}_1 - \vec{v}_2 \in \ker(T)$

But that implies $\vec{v}_1 - \vec{v}_2 = \vec{0}$, i.e. $\vec{v}_1 = \vec{v}_2$.

So $T\vec{v}_1 = T\vec{v}_2$ implies $\vec{v}_1 = \vec{v}_2$. ⊠

onto <u>Defn:</u> $T : V \to W$ a linear xformn is called <u>onto</u> if $T(V) = W$,

i.e., the range of $T$ on $V$ is <u>all</u> of $W$.

(i.e., given any $\vec{w} \in W$, $\exists\, \vec{v} \in V$ s.t. $T\vec{v} = \vec{w}$)

<u>Ex:</u> $\mathcal{B}$ a basis of $V$; then $\vec{v} \mapsto [\vec{v}]_\mathcal{B}$ is <u>onto</u> $\mathbb{R}^n$.

exercise involving rank theorem

**Ex**: Let $A \in \mathbb{R}^{n \times n}$, $T_A : \mathbb{R}^n \to \mathbb{R}^n$ given (as usual) by $T_A v = Av$.

Claim: $T_A$ is onto $\Leftrightarrow$ $A$ is invertible.

Pf: $\Leftarrow$) Given $w \in \mathbb{R}^n$, let $v = A^{-1}w$.
Then $T_A v = AA^{-1}w = w$. So $T_A$ is onto.

$\Rightarrow$) $T_A$ is onto $\Leftrightarrow$ dim(columnspace of $A$) $= n$.
ie, rank$(A) = n$.
$\Leftrightarrow$ nullity$(A) = 0$ (Rank Thm.)
$\Leftrightarrow$ $A$ is invertible.

isomorphism

**Defn**: Let $T: V \to W$ be a linear transformation.
If $T$ is 1-1 and onto, we say $T$ is an isomorphism.
(and that $V$ and $W$ are isomorphic) ($V \cong W$)

example

**Ex**: The map $T: V \to \mathbb{R}^n$ from above is an isomorphism.
$v \mapsto [v]_B$

invertible transformations

**FACT**: Thm: $T$ is an isomorphism $\Leftrightarrow$ $T$ is invertible,
ie, $\exists$ a linear xformation $T^{-1} : W \to V$
s.t. $T^{-1} \circ T = Id_V$
$T \circ T^{-1} = Id_W$.

example

**Ex**: $B = \{b_1, \ldots, b_n\}$ a basis of $V$; $T$ defined as above.
Then the map $S: \mathbb{R}^n \to V$
$(\alpha_1, \ldots, \alpha_n) \mapsto \sum \alpha_i b_i$
going from coordinate vector to original vector
is the inverse of $T$.

Point: If $\dim(V) = n$, then $V \cong \mathbb{R}^n$:
any $n$-dim'l vector space is isomorphic to $\mathbb{R}^n$.

## §4.3 Similarity

Rk: In this section, we'll be dealing with linear operators, ie $T: V \to V$, and their mx. rep'ns w.r.t. a basis $\mathcal{B}$ of $V$.

mx rep'n of linear operator w.r.t. a single basis.

Picture:
$$\begin{tikzcd} v \in V \arrow[r, "T"] \arrow[d] & V \ni v \arrow[d] \\ {[v]_{\mathcal{B}} \in \mathbb{R}^n} \arrow[r, dashed, "A_{\mathcal{B}}"] & \mathbb{R}^n \ni [v]_{\mathcal{B}} \end{tikzcd}$$

Let $A_{\mathcal{B}}$ denote the mx. rep'n w.r.t. $\mathcal{B}, \mathcal{B}$.

Notice that if we change $\mathcal{B}$, we change $A_{\mathcal{B}}$.

Q) What's the relation (if any) between $A_{\mathcal{B}}$ and $A_{\mathcal{B}'}$?

We'll need the following notion:

recall transition mx between bases

Recall: Given a vector space $V$, and bases $E = \{\vec{v}_1, \dots, \vec{v}_n\}$
$F = \{\vec{w}_1, \dots, \vec{w}_n\}$,

We found the transition matrix taking $[\vec{v}]_E$ to $[\vec{v}]_F$:

i) Express $\vec{v}_1 = S_{11}\vec{w}_1 + S_{21}\vec{w}_2 + \dots + S_{n1}\vec{w}_n$

$\vec{v}_n = S_{1n}\vec{w}_1 + S_{2n}\vec{w}_2 + \dots + S_{nn}\vec{w}_n$;

ii) Then $S = \begin{bmatrix} S_{11} & S_{12} & \cdots & S_{1n} \\ \vdots & & & \\ S_{n1} & & & S_{nn} \end{bmatrix}$ is the trans'n mx.

rel'n between mx. rep'ns

Thm: $T: V \to V$ a linear operator, $E = \{\vec{v}_1, \dots, \vec{v}_n\}$, $F = \{\vec{w}_1, \dots, \vec{w}_n\}$ bases of $V$.

Let $S$ be the trans'n mx. from $E$ to $F$.

$$(\text{ie, } S[v]_E = [v]_F \ \forall v \in V)$$

Then $$S^{-1} A_F S = A_E.$$

Proof: By construction of $A_E, A_F$, the following diagram commutes:

$$\begin{tikzcd}
{[\vec v]_F} & \mathbb{R}^n \arrow[r, "A_F"] & \mathbb{R}^n \\
\vec v \arrow[u] \arrow[d] & V \arrow[u] \arrow[d] \arrow[r, "T"] & V \arrow[u, "F"'] \arrow[d, "E"] \\
{[\vec v]_E} & \mathbb{R}^n \arrow[r, "A_E"'] & \mathbb{R}^n
\end{tikzcd}$$

Thus
$$\begin{tikzcd}
\mathbb{R}^n \arrow[r, "A_F"] & \mathbb{R}^n \arrow[d, "S^{-1}"] \\
\mathbb{R}^n \arrow[u, "S"] \arrow[r, "A_E"'] & \mathbb{R}^n
\end{tikzcd}$$
commutes,

i.e., $A_E = S^{-1} A_F S$, as desired.

similar

Defn: $A, B \in \mathbb{R}^{n\times n}$. If $\exists S \in \mathbb{R}^{n\times n}$ s.t. $B = S^{-1}AS$

we say $B$ is similar to $A$.

Rk: Notice $B$ similar to $A \Leftrightarrow A$ similar to $B$.

(and $A \sim B$, $B \sim C \Rightarrow A \sim C$. An equivalence relation)

(exercise)

Rk: So above says if $A$ and $B$ are mx rep'ns of $T$, then $A \sim B$.

In fact, the converse is true: if $A$ is a mx rep'n of $T$,

then any $S^{-1}AS = B$ is also a mx rep'n of $T$

example: find the new mx. rep'n.

Example 1: $D$ differentiation on $P_3$.

The matrix representing $D$ w.r.t. $\mathcal{B} = \{1, x, x^2\}$ is $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$

Use the transition mx $S$ from $\{1, 2x, 4x^2-2\}$ to $\{1, x, x^2\}$

to obtain the mx rep'n of $D$ wrt $\mathcal{V}$.

$$\left.\begin{aligned}
\vec v_1 &= \vec b_1 + 0\vec b_2 + 0\vec b_3 \\
\vec v_2 &= 0\vec b_1 + 2\vec b_2 + 0\vec b_3 \\
\vec v_3 &= -2\vec b_1 + 0\vec b_2 + 4\vec b_3
\end{aligned}\right\} \Rightarrow S = \begin{bmatrix} 1 & 0 & -2 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{bmatrix}.$$

$$\left(S^{-1} = \begin{bmatrix} 1 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{4} \end{bmatrix}\right)$$

Then $S^{-1}AS = \begin{bmatrix} 0 & 2 & 0 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{bmatrix}$.

(Check: $D\vec v_1 = 0$, $D\vec v_2 = 2\vec v_1$, $D\vec v_3 = 4\vec v_2$ ✓)

Exerc 5: $L$ the operator on $P_3$ given by $L(p(x)) = xp'(x) + p''(x)$.

Q) What is $L^{100}(3x+5+5x^2)$?

a) Find matrix rep'n $A$ of $L$ w.r.t. $\{1, x, x^2\} = \mathcal{B}$

$$L\vec{b}_1 = L(1) = 0.$$
$$L\vec{b}_2 = L(x) = x + 0 = \vec{b}_2$$
$$L\vec{b}_3 = L(x^2) = x \cdot 2x + 2 = 2\vec{b}_1 + 2\vec{b}_3.$$

$$; A = \begin{bmatrix} 0 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

b) Find mx rep'n $B$ of $L$ w.r.t. $\{1, x, 1+x^2\} = \mathcal{V}$

$$L\vec{v}_1 = L(1) = 0$$
$$L\vec{v}_2 = Lx = x + 0 = \vec{v}_2$$
$$L\vec{v}_3 = L(1+x^2) = x \cdot 2x + 2 = 2x^2 + 2 = 2\vec{v}_3$$

$$; B = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

c) What's the trans'n mx $S$ s.t. $B = S^{-1}AS$?

$$\vec{v}_1 = \vec{b}_1$$
$$\vec{v}_2 = \vec{b}_2 \quad ; \quad S = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix};$$
$$\vec{v}_3 = \vec{b}_1 + \vec{b}_3$$

Calculation yields $B = S^{-1}AS$

Point: $B$ is easier to deal with than $A$. E.g.,

d) If $p(x) = a_0 + a_1 x + a_2(1+x^2)$, what's $L^n(p(x))$?

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

Rk: Notice that $\vec{v}_1, \vec{v}_2, \vec{v}_3$ are all eigenvectors of $L$.

## §5.1 Scalar Product in $\mathbb{R}^n$ (concrete, first instance of "inner product")

Notions that you should already be familiar with from Calc III:

i) dot/scalar product in $\mathbb{R}^n$: $x, y \in \mathbb{R}^{n\times 1}$. Their scalar product is defined as $x^T y$, i.e. $x_1y_1 + \dots + x_ny_n$.

ii) In $\mathbb{R}^2$ or $\mathbb{R}^3$, we have

a. a rel'n with magnitude: $\|x\| = (x^Tx)^{1/2}$

b. and thus a rel'n with distance: $d(x,y) := \|x - y\|$

c. and a geometric interpretation: $x^Ty = \|x\|\|y\|\cos\varphi$. (based on Law of Cosines)

- Recall we define the dir'n of $x$ to be $\frac{x}{\|x\|}$ (normalized vector)

d. Cauchy-Schwarz Inequality: $|x^Ty| \le \|x\|\|y\|$.

e. Def'n: If $x^Ty = 0$, we say $x, y$ are orthogonal.

f. Recall the vector proj'n of $x$ onto $y$:

$$\text{Proj}_y x := \frac{x^Ty}{\|y\|}\frac{y}{\|y\|} = \frac{x^Ty}{y^Ty}\, y.$$

iii) Can extend all of above to $\mathbb{R}^n$

Section contains various applications that you should've seen in Calc III.

- finding eq'n of plane, given point & normal vector.
- cross products in $\mathbb{R}^3$, etc

(not important)

Rk: • Primitive search algorithms based on $\cos\theta$.

(Read applications 2 & 3)

- Correlation mx. a) subtract mean; get deviations b) normalize
  exams & HW? c) take $U^TU$; gives mx. of cosines.

## §5.2 Orthogonal complements & Direct sums in $\mathbb{R}^n$.

orthogonal subspaces (intermediary ~~concept~~)

**Defn:** [$X, Y$ subspaces of $\mathbb{R}^n$] If $x^T y = 0$ for every $x \in X$, $y \in Y$, then we say $X \perp Y$, ie. $X$ and $Y$ are orthogonal (subspaces).

examples

Example: i) (~~trivial example~~) $X = \text{span}\left\{\begin{bmatrix}1\\0\\0\end{bmatrix}\right\}$, $Y = \text{span}\left\{\begin{bmatrix}0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix}\right\}$.

then $X \perp Y$.

ii) (nonexample) $X = \text{span}\left\{\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}\right\}$, $Y = \text{span}\left\{\begin{bmatrix}0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix}\right\}$.

(the xy-plane) (the yz-plane)

iii) (more sophisticated example) $N(A) \perp$ column space of $A^T$, $A \in \mathbb{R}^{m \times n}$.

Say $x \in N(A)$, ie. $Ax = 0$.

ie. $\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ \vdots \\ 0 \end{bmatrix}$

ie., $\vec{a}_i x = 0$ for $i = 1, \ldots, m$.

ie, $x$ is orthogonal to each row vector of $A$

$x$ is orthogonal to any vector in row space $(A)$.

Ie, $N(A) \perp$ row space of $A$ (column space of $A^T$).

comment ~~about~~ range

Remark: The column space of $A^T$ = range of $A^T$, so we're saying $N(A) \perp R(A^T)$.

warning

Q Are we saying that if $y$ is orthogonal to $N(A)$ then $y \in R(A^T)$?

No: ~~Only that if $x \in R(A^T)$ then $x \perp N(A)$.~~

| | |
|---|---|
| orthogonal complement | **Defn:** [$S \subset \mathbb{R}^n$ a subspace] Let $S^\perp = \{ x \in \mathbb{R}^n \mid x^T s = 0 \ \forall s \in S \}$ i.e., the colln. of all vectors orthogonal to every $s \in S$, called the orthogonal complement of $S$ in $\mathbb{R}^n$. |
| | Rk: $S_1 \perp S_2 \Rightarrow S_1 \subset S_2^\perp$ and $S_2 \subset S_1^\perp$. |
| | Trivial observations: ( 0) Of course $S \perp S^\perp$ ) |
| observations | i) $X \perp Y \Rightarrow X \cap Y = \{0\}$ |
| | ii) $Y^\perp$ is a subspace |
| | Pf's: Trivial |
| | Rk: We shall see that, given subspace $S \subset \mathbb{R}^n$, we can decompose $\mathbb{R}^n$ into a "sum" of $S + S^\perp$ |
| "Fund'l Subspaces Thm" (?) | Thm: [$A \in \mathbb{R}^{m \times n}$] $N(A) = R(A^T)^\perp$ (and $N(A^T) = R(A)^\perp$). |
| | ~~Pf: $\subseteq$: $x \in N(A) \Rightarrow x \in R(A^T)$~~, we saw earlier. |
| | $\supseteq$: $x \in R(A^T)^\perp \Rightarrow x \perp$ to any vector in $R(A^T)$ |
| | $\Rightarrow$ in particular, $x \perp$ to the columns of $A^T$ |
| | $\Rightarrow Ax = \vec{0} \Rightarrow x \in N(A)$ |
| | [Apply $N(A) = R(A^T)^\perp$ to $A^T$ to get the second statement] |
| | More facts about orthogonal complements: |
| Properties of orthog. complements | Thm: i) [$S$ subspace of $\mathbb{R}^n$] $\dim(S) + \dim(S^\perp) = n$. |
| | ii) $\{x_1, \ldots, x_r\}$ a basis of $S$, $\{x_{r+1}, \ldots, x_n\}$ a basis of $S^\perp$ $\Rightarrow \{x_1, \ldots, x_r, x_{r+1}, \ldots, x_n\}$ is a basis of $\mathbb{R}^n$ |
| | Pf: i) [Q. What theorem do you think we'll apply?] |
| | Case $S = \{0\}$ is trivial. |
| | Let $\{x_1, \ldots, x_r\}$ be a basis of $S$; let $A = \begin{bmatrix} \vert & & \vert \\ x_1 & \cdots & x_r \\ \vert & & \vert \end{bmatrix}_{(n \times r)}$ |
| | Then $S = R(A)$, and $S^\perp = R(A)^\perp = N(A^T)$. |
| | The Rank Thm $\Rightarrow \underbrace{\dim(N(A^T)) + \dim(R(A^T))}_{(\dim(S^\perp) + \dim(S) = n)} = n$ ∎ |

ii) Claim: $\{x_1, \ldots, x_r ; x_{r+1}, \ldots, x_n\}$ is a basis of $\mathbb{R}^n$.

Pf: STS they are linearly indpt.

Suppose $\sum_{i=1}^{n} c_i x_i = \vec{0}$ in $\mathbb{R}^n$.

Then $\sum_{i=1}^{r} c_i x_i = \sum_{i=r+1}^{n} -c_i x_i = v$

Observe LHS $\in S$, RHS $\in S^\perp$; so $v \in S \cap S^\perp = \{0\}$

$\Rightarrow c_1 = \ldots = c_r = 0 ; c_{r+1} = \ldots = c_n = 0$ ∎

direct sum

Defn: [$U, V$ subspaces of $W$, a vector space]

If every $\vec{w} \in W$ can be uniquely expressed as $\vec{u} + \vec{v}$; $\vec{u} \in U$, $\vec{v} \in V$

then we say $W$ is the direct sum of $U$ and $V$

and write $W = U \oplus V$.

$\mathbb{R}^n = S \oplus S^\perp$ ($S, S^\perp$ form a direct sum)

Example: Given any subspace $S \subset \mathbb{R}^n$, $\mathbb{R}^n = S \oplus S^\perp$.

Pf: Let $\{x_1, \ldots, x_r\}$ be a basis of $S$, $\{x_{r+1}, \ldots, x_n\}$ one of $S^\perp$.

We saw $\{x_1, \ldots, x_n\}$ is a basis of $\mathbb{R}^n$;

given any $w \in \mathbb{R}^n$,

$$w = \underbrace{c_1 x_1 + \ldots + c_r x_r}_{u} + \underbrace{c_{r+1} x_{r+1} + \ldots + c_n x_n}_{v}$$

Is this expression unique?

Suppose $w = u' + v'$; $u' \in S$, $v' \in S^\perp$.

Then $u + v = u' + v' \Rightarrow u - u' = v' - v$.

This object lies in $S \cap S^\perp = \{0\}$.

$\Rightarrow u = u'$, $v = v'$.

So $\mathbb{R}^n = S \oplus S^\perp$ ∎

useful property about orthogonal complementation

**Thm:** [$S$ a subspace of $\mathbb{R}^n$] $(S^\perp)^\perp = S$.

Pf: $\supseteq$: Given any $v \in S$, we know $v \perp$ any $y \in S^\perp$.

$$\Rightarrow v \in (S^\perp)^\perp.$$

$\subseteq$: Let $x \in (S^\perp)^\perp$ i.e., $x$ is $\perp$ to <u>any</u> $y \in S^\perp$.

We know $x = u + v$ for some $u \in S$, $v \in S^\perp$.

Take the scalar product of $x$ with $v$:

$$\underset{\uparrow \text{ since } x \in (S^\perp)^\perp}{0 = v^T x} = v^T(u+v) = \underset{\uparrow \text{ since } v \in S^\perp}{v^T v}$$

$$\Rightarrow \|v\| = 0, \text{ so } v = 0.$$

Thus $x = u$; i.e. $x \in S$. ∎

Final comment: Recall dim (row space) = dim (column space)

$$(\dim(R(A^T)) = \dim(R(A)))$$

observation: $R(A) \cong R(A^T)$ made explicitly clear.

Why? ~~The map~~ $T_A: R(A^T) \longrightarrow R(A)$ is an isomorphism.

Why? Say $A \in \mathbb{R}^{m\times n}$; i.e. $A: \mathbb{R}^n \to \mathbb{R}^m$.

Now, $\mathbb{R}^n = R(A^T) \oplus N(A)$

any $x = u + v$, $u \in R(A^T)$, $v \in N(A)$.

What does $A$ do? $Ax = Au + \cancel{Av}^{\,\vec{0}}$.

~~Observe:~~

i) $A$ maps $R(A^T)$ onto $R(A)$

ii) $A$ is 1-1 on $R(A^T)$: if $u_1, u_2 \in R(A^T)$

and $Au_1 = Au_2 \Rightarrow A(u_1 - u_2) = 0$

$$\Rightarrow u_1 - u_2 \in R(A^T) \cap N(A) = \{0\}.$$

(Rk: Not a deep section)

## §5.4 Scalar products (and norms) for abstract vector spaces.

motivation

**Motivation**: Many familiar phenomena in $\mathbb{R}^n$ (Pythagorean Thm., Cauchy-Schwarz inequality, etc.) are actually not endemic to $\mathbb{R}^n$: They are consequences of simple properties of the dot product....

inner product

**Def'n**: [V a vector space] Let $\langle , \rangle : V \times V \longrightarrow \mathbb{R}$ be a binary operation on V (taking in 2 vectors, returning a real number)

If I. $\langle \vec{v}, \vec{v} \rangle \geq 0 \ \forall \vec{v} \in V$, and $\langle \vec{v}, \vec{v} \rangle = 0 \Leftrightarrow \vec{v} = \vec{0}$,

II. $\langle \vec{x}, \vec{y} \rangle = \langle \vec{y}, \vec{x} \rangle \quad \forall \vec{x}, \vec{y} \in V$

III. $\langle \alpha\vec{x} + \beta\vec{y}, \vec{z} \rangle = \alpha\langle \vec{x}, \vec{z} \rangle + \beta\langle \vec{y}, \vec{z} \rangle \quad \forall \vec{x}, \vec{y}, \vec{z} \in V, \ \alpha, \beta \in \mathbb{R}$

Then we say $\langle , \rangle$ is an <u>inner product on V</u> and that $\{V, \langle , \rangle\}$ is an <u>inner product space</u>.

important examples.

<u>Examples</u>: i. $V = \mathbb{R}^n$; $\langle \mathbf{x}, \mathbf{y} \rangle := \mathbf{x}^T \mathbf{y} = \sum_{i=1}^n x_i y_i$

i.' $V = \mathbb{R}^n$, $\mathbf{w} \in \mathbb{R}^n$ with $w_i > 0 \ \forall i$.

Then $\langle \mathbf{x}, \mathbf{y} \rangle := \sum_{i=1}^n x_i y_i w_i$ the <u>weighted dot product</u> is an inner product.

ii. $A, B \in \mathbb{R}^{m \times n}$ $\quad \langle A, B \rangle := \sum_{i,j=1}^{m,n} a_{ij} b_{ij}$

iii. $V = C[a,b]$; $\langle f, g \rangle := \int_a^b f(x) g(x) \, dx$.

or more general, for $w \in C[a,b]$ s.t. $w(x) > 0$, $\langle f, g \rangle := \int_a^b f(x) g(x) w(x) \, dx$.

iv. $V = P_n$. Fix any $x_1, \ldots, x_n \in \mathbb{R}^n$.

Then $\langle p, q \rangle := \sum_{i=1}^n p(x_i) q(x_i)$ $\left(\text{or } \sum p(x_i) q(x_i) w(x_i) \text{ for } w \text{ a pos. fn.}\right)$

is an inner product.

(checking)

Q. Why is property I satisfied?

related concepts — Once we have an inner product, we can define a notion of orthogonality:

orthogonal — Defn: Let $\vec{u},\vec{v}\in V$, an inner product space. If $\langle \vec{u},\vec{v}\rangle = 0$, we say $\vec{u}$ and $\vec{v}$ are <u>orthogonal</u>.

<u>Ex</u>: In $C[-1,1]$, $\langle 1, x\rangle = \int_{-1}^{1} 1\cdot x\,dx = 0$.

We can also introduce a notion of <u>magnitude</u>:

norm — Defn: $\vec{v}\in V$, an inner product space. We define the <u>norm</u> of $\vec{v}$ by $\|\vec{v}\| := \langle \vec{v},\vec{v}\rangle^{1/2}$

And we can define a notion of projection:

projection — Defn: $\vec{u},\vec{v}\in V$, i.p.s. We define the (vector) <u>projection</u> of $\vec{u}$ onto $\vec{v}$ by $\vec{p} = \dfrac{\langle \vec{u},\vec{v}\rangle}{\langle \vec{v},\vec{v}\rangle}\vec{v}$.

(and the scalar projn as $\dfrac{\langle \vec{u},\vec{v}\rangle}{\|\vec{v}\|}$.)

familiar properties — <u>General Properties of (abstract) inner product spaces</u>: Any such space has a Pythagorean Thm, etc.

Pythagorean Thm — Thm: (Pythagorean) $V$ an inner product space, $\vec{u},\vec{v}\in V$.

If $\vec{u}\perp\vec{v} \Rightarrow \|\vec{u}\|^2 + \|\vec{v}\|^2 = \|\vec{u}+\vec{v}\|^2$.

Rk: If you don't think this is cool, you should get out!

examples — <u>Ex</u>: In $C[-1,1]$, we know $\|1\|^2 + \|x\|^2 = \|1+x\|^2$.

$\left(\Rightarrow \|1+x\|^2 = 2 + \frac{2}{3}\right)$

(Fourier series) — <u>Ex</u>: $V = C[-\pi,\pi]$, $\langle f,g\rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)g(x)\,dx$ (weighted)

Calculus exerc: $\langle \cos x, \sin x\rangle = 0$, $\langle \cos x,\cos x\rangle = \langle \sin x, \sin x\rangle = 1$

$\Rightarrow \|\cos x + \sin x\|^2 = \|\cos x\|^2 + \|\sin x\|^2 = 2$

proof — Proof of Thm: $\|\vec{u}+\vec{v}\|^2 = \langle \vec{u}+\vec{v}, \vec{u}+\vec{v}\rangle = \langle \vec{u},\vec{u}\rangle + 2\cancelto{0}{\langle \vec{u},\vec{v}\rangle} + \langle \vec{v},\vec{v}\rangle$

(Rk: this has nothing to do with angles)

Cauchy-Schwarz

Thm: (Cauchy-Schwarz inequality) $\vec{u}, \vec{v} \in V$ an inner product sp.

$\Rightarrow \quad |\langle \vec{u}, \vec{v} \rangle| \le \|\vec{u}\| \, \|\vec{v}\|$;

Further, equality holds $\Leftrightarrow$ $\vec{u}, \vec{v}$ are linearly dependent.

(Rk: Forgotten part of C-S)

The proof requires two lemmas:

lemmas

<u>Lemmata</u>: Say $\vec{v} \ne \vec{0}$; $\vec{u}, \vec{v} \in V$ an i.p.s., $\vec{p} := \mathrm{Proj}_{\vec{v}} \vec{u}$.

(creating orthogonal pieces)

i) $\vec{u} - \mathrm{Proj}_{\vec{v}} \vec{u} \perp \mathrm{Proj}_{\vec{v}} \vec{u}$:

Pf: $\left\langle \vec{u} - \dfrac{\langle \vec{u}, \vec{v} \rangle}{\langle \vec{v}, \vec{v} \rangle} \vec{v}, \; \dfrac{\langle \vec{u}, \vec{v} \rangle}{\langle \vec{v}, \vec{v} \rangle} \vec{v} \right\rangle = 0$ (trivial calculation)

ii) $\vec{u} = \mathrm{Proj}_{\vec{v}} \vec{u} \iff \vec{u} = \beta \vec{v}$ for some $\beta \in \mathbb{R}$:

Pf: $\Rightarrow$) is true by construction

$\Leftarrow$) $\vec{u} = \beta \vec{v} \Rightarrow \dfrac{\langle \beta \vec{v}, \vec{v} \rangle}{\langle \vec{v}, \vec{v} \rangle} \vec{v} = \beta \vec{v} = \vec{u}$.

proof

<u>Proof of Cauchy-Schwarz:</u> (Slick)

First note that if $\vec{v} = \vec{0}$ then the proof is trivial.

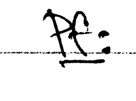
Say $\vec{v} \ne \vec{0}$.

We <u>want to show</u> $\langle \vec{u}, \vec{v} \rangle^2 \le \|\vec{u}\|^2 \|\vec{v}\|^2$

i.e. $\left( \dfrac{\langle \vec{u}, \vec{v} \rangle}{\|\vec{v}\|} \right)^2 \le \|\vec{u}\|^2$

$\leftarrow$ that's $\|\vec{p}\|^2$ the proj'n of $\vec{u}$ on $\vec{v}$.

(nonexistent) Picture




Well, by Pythagorean Thm,

$$\|\vec{p}\|^2 = \|\vec{u}\|^2 - \|\vec{u} - \vec{p}\|^2 \overset{?}{\le} \|u\|^2 ?$$

For sure, and equality holds if $\vec{u} = \vec{p}$,

i.e. $\vec{u}$ is a multiple of $\vec{v}$. ∎

angle

**Def'n:** Given $\vec{u}, \vec{v} \in V$ an inner product space,

we define the <u>angle</u> between $\vec{u}, \vec{v}$ to be $\cos^{-1}\left(\frac{\langle \vec{u}, \vec{v} \rangle}{\|\vec{u}\| \|\vec{v}\|}\right)$

Rk: makes sense by C-S inequality.

norm

<u>Abstract norm:</u>

**Def'n:** $V$ a vector space. If $\exists$ a function $\|\cdot\| : V \to \mathbb{R}$, s.t.

i. $\|\vec{v}\| \geq 0$, with equality $\Leftrightarrow \vec{v} = \vec{0}$,

ii. $\|\alpha \vec{v}\| = |\alpha| \|\vec{v}\|$ for any scalar $\alpha$

iii. $\|\vec{v} + \vec{w}\| \leq \|\vec{v}\| + \|\vec{w}\|$ $\forall \vec{v}, \vec{w} \in V$. ($\Delta \neq$)

Then we call $\| \;\|$ a <u>norm</u> on $V$,

and say $V$ is a <u>normed linear space</u> (vector).

norms

<u>Important examples of norms:</u>

i) On $\mathbb{R}^n$: $\|v\|_1 := \sum_{i=1}^{n} |v_i|$. (the $l^1$ norm)

ii) $\|v\|_p := \left(\sum_{i=1}^{n} |v_i|^p\right)^{1/p}$. ($l^p$)

iii) $\|v\|_\infty := \max\{v_1, \ldots, v_n\}$. ($l^\infty$)

**Def'n:** (distance) Given normed v.s. $V$, $\vec{x}, \vec{y} \in V$,

Then we define $d(\vec{x}, \vec{y}) = \|\vec{x} - \vec{y}\|$.

induced norm

Last important norm: $\{V, \langle , \rangle\}$ an i.p.s

Then $\|\vec{v}\| := (\langle \vec{v}, \vec{v} \rangle)^{1/2}$ is a norm. (Proof?) (use C-S)

[lecture, S.S.]

orthogonal set of vectors.

Thm: orthogonal ⇒ independent.

orthonormal set.

orthogonal ⟶ orthonormal

Example of orthogonal → orthonormal.

§6.1 Eigenvectors & Eigenvalues.

(main defins) Defn: $A \in \mathbb{R}^{n\times n}$, $\lambda \in \mathbb{C}$. If there exists a vector $\vec{v} \neq \vec{0}$ s.t. $A\vec{v} = \lambda\vec{v}$, we say $\lambda$ is an <u>eigenvalue</u> of $A$ and that $\vec{v}$ is an <u>eigenvector</u> belonging to $\lambda$.

simple exercises <u>Exer. 1:</u> If $\lambda$ is an eigenvalue of $A$, the colln. of eigenvectors belonging to $\lambda$ forms a subspace.

(called the <u>eigenspace</u> $E_\lambda$)

Two proofs:

i) Dull: if $\vec{v}_1, \vec{v}_2 \in E_\lambda$ then $A(\vec{v}_1 + \vec{v}_2) =$

$$= A\vec{v}_1 + A\vec{v}_2 = \lambda\vec{v}_1 + \lambda\vec{v}_2 = \lambda(\vec{v}_1 + \vec{v}_2).$$

If $\vec{v} \in E_\lambda$, then $A(\alpha\vec{v}) = \alpha A\vec{v} = \alpha\lambda\vec{v}$.

$= \lambda\alpha\vec{v}$.

ii) Slick: $\vec{v} \in E_\lambda \Leftrightarrow A\vec{v} = \lambda\vec{v}$

$$\Leftrightarrow A\vec{v} - \lambda I\vec{v} = \vec{0}$$

$$\Leftrightarrow (A - \lambda I)\vec{v} = \vec{0}$$

$$\Leftrightarrow \vec{v} \in \text{Nullspace}(A - \lambda I).$$

<u>Exer. 2:</u> $\lambda$ is an eigenvalue of $A \Leftrightarrow \det(A - \lambda I) = 0$.

($\Leftrightarrow$ $A - \lambda I$ is singular)

($\Leftrightarrow$ $A - \lambda I$ has a nontrivial soln)

Pf: Observe: $\lambda$ is an eigenvalue of $A$

$\Leftrightarrow$ there's some nonzero $\vec{v} \in E_\lambda$.

$\Leftrightarrow$ Nullspace $(A - \lambda I)$ is nontrivial.

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

**Remark:** So the above gives us a way of finding eigenvalues (and eigenvectors):

char. poly'l.

**Defn:** $A \in \mathbb{R}^{n \times n}$. We call $\det(A-\lambda I)$ the characteristic polynomial of $A$.

Rk: Notice the roots of $\det(A-\lambda I)$ are the eigenvalues of $A$.

Application: finding eigenvals & eigenspaces

**Ex:** (1b) $A = \begin{bmatrix} 6 & -4 \\ 3 & -1 \end{bmatrix}$ Find evals & espaces. $\det(A-\lambda I) = \det\begin{bmatrix} 6-\lambda & -4 \\ 3 & -1-\lambda \end{bmatrix}$

$$= (6-\lambda)(-1-\lambda)+12 = -6-5\lambda+\lambda^2+12 = \lambda^2-5\lambda+6$$
$$= (\lambda-2)(\lambda-3).$$

$\Longrightarrow \begin{bmatrix} 6 & -4 \\ 3 & -1 \end{bmatrix}$ has eigenvalues 2 and 3.

(exerc)

~~Eigenspaces~~: $E_2$: Want to show $A\vec{v} = 2\vec{v}$, i.e. $[A-2I]\vec{v} = \vec{0}$

Well: $\left[\begin{array}{cc|c} 4 & -4 & 0 \\ 3 & -3 & 0 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & -1 & 0 \\ 0 & 0 & 0 \end{array}\right]$; so $x_1 - x_2 = 0$.

i.e. $E_2 = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix} s;\ s \in \mathbb{R} \right\}$.

$E_3$: $\left[\begin{array}{cc|c} 3 & -4 & 0 \\ 3 & -4 & 0 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & -\frac{4}{3} & 0 \\ 0 & 0 & 0 \end{array}\right]$; $x_1 = \frac{4}{3}x_2$.

$E_3 = \left\{ \begin{pmatrix} \frac{4}{3} \\ 1 \end{pmatrix} r;\ r \in \mathbb{R} \right\}$.

(exerc) **Exerc:** (1a) $\begin{bmatrix} 3 & 2 \\ 4 & 1 \end{bmatrix}$.

More sophisticated exercs: 8. $A \in \mathbb{R}^{n \times n}$ idempotent ($A^2 = A$) show eigenvalues are either 0 or 1

9. $A \in \mathbb{R}^{n \times n}$ nilpotent ($A^k = 0$). Show eigenvalues must be 0.

18. $(A-\lambda I)$ has rank $k \Rightarrow$ what's $\dim(E_\lambda)$?

Some simple (?) observations.

i) If $\lambda \in \mathbb{C}$ is an eigenvalue, then so is $\bar{\lambda}$. ← complex conjugate $(A \in \mathbb{R}^{n\times n})$

Why? $p(\lambda) = \det(A - \lambda I) = 0$

$$\Rightarrow \overline{\det(A-\lambda I)} = 0 \Rightarrow \det(\bar{A} - \bar{\lambda} I) = 0$$
$$\Rightarrow \det(A - \bar{\lambda} I) = 0.$$

ii) If $\lambda, \bar{\lambda} \in \mathbb{C}$ are eigenvalues, $\vec{v} \in E_\lambda \Leftrightarrow \bar{\vec{v}} \in E_{\bar{\lambda}}$.

Pf: Obvious. If $\vec{v}$ is an eigenvector belonging to $\lambda$,

Then $A\bar{v} = \bar{A}\bar{v} = \overline{Av} = \bar{\lambda}\bar{v}$

since $A$ is real

<u>Not-so-simple observations</u>: Product & Sum of Eigenvalues.

<u>Thm</u>: $A \in \mathbb{R}^{n\times n}$, $p(\lambda)$ the characteristic polyl of $A$,
$\{\lambda_1, \ldots, \lambda_n\}$ the roots of $p(\lambda)$ (some may repeat)
(ie, the <u>eigenvalues</u> of $A$)

Then Claim: i. $\lambda_1 \lambda_2 \cdots \lambda_n = \det(A)$

ii. $\sum \lambda_i = \sum_{i=1}^{n} a_{ii}$, the <u>trace</u> of $A$.

<u>Ex</u>: (16) $A = \begin{bmatrix} 6 & -4 \\ 3 & -1 \end{bmatrix}$ ; $\det(A) = -6 - -12 = 6$,
$tr(A) = 5$.

Lo and behold, $\lambda_1 \cdot \lambda_2 = 2 \cdot 3 = 6$,
$\lambda_1 + \lambda_2 = 2 + 3 = 5$.

(not a coincidence)

(Exerc. 16.)

Initial exercise: $(x-3)(x-2)(x-5)(x+7)(x+1) = ?$

Sketch of proof: (Difficult to conceptualize)

I. First notice that $p(\lambda) = \det(A - \lambda I)$, a poly $^{(\text{in } \lambda)}$ of degree $n$.

$\Rightarrow$ it has $n$ roots (some repeated), the eigenvals $\lambda_1, \ldots, \lambda_n$

$$\Rightarrow p(\lambda) = C(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n) \quad (\dagger)$$

(for some $C$)

II. Let's expand $p(\lambda)$ out: $p(\lambda) = b_n\lambda^n + b_{n-1}\lambda^{n-1} + \cdots + b_0$.

(i.e., let $b_k$ be the coeff of $\lambda^k$).

Observe: $C = b_n$ ; $b_{n-1} = C\left(-\sum_{i=1}^{n}\lambda_i\right)$

III. Now, $p(\lambda) = \det \begin{bmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & & a_{2n} \\ \vdots & & & \\ a_{n1} & a_{n2} & \cdots & a_{nn}-\lambda \end{bmatrix}$

Observe: Expanding, we see there is only one term involving more than $n-2$ of the diagonal elements, namely

$$(a_{11}-\lambda)(a_{22}-\lambda)\cdots(a_{nn}-\lambda) \quad (*)$$

(So this is the only term that can contribute to $b_n, b_{n-1}$)

What's the coeff. of $\lambda^n$ in $(*)$? $(-1)^n$. So $C = (-1)^n$

So $p(\lambda) = (-1)^n(\lambda-\lambda_1)\cdots(\lambda-\lambda_n)$ ; $p(0) = \prod \lambda_i$

What's the coeff of $\lambda^{n-1}$ in $(*)$? $(-1)^{n-1}\sum_{i=1}^{n} a_{ii}$.

$$\Rightarrow b_{n-1} = (-1)^{n-1}\sum_{i=1}^{n} a_{ii}$$

$$\| $$

$$(-1)^n(-1)\sum_{i=1}^{n}\lambda_i.$$

Exerc: Thm: $A, B \in \mathbb{R}^{n\times n}$ If $B \sim A \Rightarrow$
A, B have the same char. poly.
( Similar mxs have same eigenvalues.)

Exercises: 26, $B = S^{-1}AS$, $x \in E_{\lambda, B}$;

Show $Sx \in E_{\lambda, A}$.

§6.4. Hermitian matrices.

$\mathbb{C}$ — First recall: • For $z = a+bi \in \mathbb{C}$, we define $\bar{z} = a-bi$, the complex conjugate of $z$.

• The modulus $|z| := \sqrt{\bar{z} z} = \sqrt{a^2+b^2}$.

$\mathbb{C}^n$ — • We can consider vectors in $\mathbb{C}^n$: $(z_1, z_2, \dots, z_n)^T = \mathbb{Z}$.

• We define $\overline{\mathbb{Z}} = (\bar{z}_1, \bar{z}_2, \dots, \bar{z}_n)^T$ and

• $\|\mathbb{Z}\| := (\overline{\mathbb{Z}}^T \mathbb{Z})^{1/2}$

$$= \left( (\bar{z}_1, \dots, \bar{z}_n) \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} \right)^{1/2} = \left( |z_1|^2 + \dots + |z_n|^2 \right)^{1/2}$$

~~Rk: Book notation~~ $\overline{\mathbb{Z}}^T = $ "$\mathbb{Z}^H$" (named.)

inner product — Rk: We can define a (complex) inner product on $\mathbb{C}^n$, namely $\langle \mathbb{Z}, \mathbb{W} \rangle := \mathbb{W}^H \mathbb{Z}$.

(Check this is an inner product.)

Rk: Notice that $\langle \mathbb{Z}, \alpha \mathbb{W} \rangle = \bar{\alpha} \langle \mathbb{Z}, \mathbb{W} \rangle$ (conjugate linear)

Complex matrices: $\mathbb{C}^{m\times n}$

Notation: For $M \in \mathbb{C}^{m\times n}$, say $M = (m_{ij})$,

we define the conjugate $\overline{M}$ by conjugating each entry.

Defn: If $M = \overline{M}^T$, we say $M$ is Hermitian.

(Rk: If $M \in \mathbb{R}^{n\times n}$, what would "Hermitian" mean?)

(Ex: $\begin{bmatrix} 1 & 3+i \\ 3-i & 0 \end{bmatrix}$ is Hermitian)

Exer: $\langle \mathbb{Z}, A\mathbb{W} \rangle = \langle A\mathbb{Z}, \mathbb{W} \rangle$.

Thm: $[A \in \mathbb{C}^{n\times n}]$ If $A$ is Hermitian, then

i) Spec $(A)$ is all real

ii) Eigenvectors corresponding to distinct eigenvalues are orthogonal.

Pf: i) Say $\vec{v} \in E_\lambda$.

Since $A$ is Hermitian, $\langle \vec{v}, A\vec{v} \rangle = \langle A\vec{v}, \vec{v} \rangle$

$$\begin{array}{ccc} \langle \vec{v}, A\vec{v} \rangle & = & \langle A\vec{v}, \vec{v} \rangle \\ \| & & \| \\ \langle \vec{v}, \lambda\vec{v} \rangle & & \langle \lambda\vec{v}, \vec{v} \rangle \\ \| & & \| \\ \bar{\lambda} \langle \vec{v}, \vec{v} \rangle & & \lambda \langle \vec{v}, \vec{v} \rangle. \end{array}$$

So $\bar{\lambda} = \lambda$.

ii) Say $\vec{v}_i \in E_{\lambda_i}$; $i = 1, 2$.

Again, $A$ Hermitian $\Rightarrow$

$$\begin{array}{ccc} \langle A\vec{v}_1, \vec{v}_2 \rangle & = & \langle \vec{v}_1, A\vec{v}_2 \rangle \\ \| & & \| \\ \langle \lambda_1\vec{v}_1, \vec{v}_2 \rangle & & \langle \vec{v}_1, \lambda_2\vec{v}_2 \rangle \\ \| & (\text{note } \lambda_i \in \mathbb{R}) & \| \\ \lambda_1 \langle v_1, v_2 \rangle & & \lambda_2 \langle v_1, v_2 \rangle \end{array}$$

Since $\lambda_1 \neq \lambda_2$, $\langle v_1, v_2 \rangle$ must be $0$.

Corollary: If $A$ is Hermitian & has distinct eigenvalues, then there exists an orthogonal mx diagonalizing $A$.

Pf: $A$ has $n$ distinct eigenvalues $\Rightarrow$ has $n$ orthogonal eigenvectors $\{\vec{v}_1, \dots, \vec{v}_n\}$.

Normalize them: let $\vec{u}_i = \dfrac{\vec{v}_i}{\|\vec{v}_i\|}$; $i = 1, \dots, n$

(they are still orthogonal eigenvectors)

Then $\mathcal{U} = \begin{bmatrix} | & & | \\ u_1 & \cdots & u_n \\ | & & | \end{bmatrix}$ diagonalizes $A$. ∎

Rk: Actually we mean "unitary", not orthogonal.

**Defn:** $[\mathcal{U} \in \mathbb{C}^{n \times n}]$. If the columns of $\mathcal{U}$ form an ON set in $\mathbb{C}^n$ we call $\mathcal{U}$ <u>unitary</u>.

(<u>Rk:</u> Same as orthogonal mx. but for complex matrices)

Rk: Observe that <u>unitary</u> means $\overline{\mathcal{U}}^T \mathcal{U} = I$.

$$(\Rightarrow \mathcal{U}^{-1} = \overline{\mathcal{U}}^T)$$

So the corollary should properly read,

<u>Corollary:</u> If $A$ is Hermitian and has $n$ distinct eigenvalues $\Rightarrow \exists$ a unitary mx. that diagonalizes $A$.

In fact something far more remarkable is true:

<u>The Spectral Theorem:</u> $A$ Hermitian $\Rightarrow \exists$ unitary $\mathcal{U}$ that diagonalizes $A$.

(i.e. $A$ Hermitian $\Rightarrow$ we're <u>guaranteed</u> sufficient eigenvectors)

Recall: Symmetric $\Rightarrow$ orthogonally diagonalizable.

§6.5 Singular Value Decomposition.

SVD

**Defn:** $A \in \mathbb{R}^{m\times n}$ ; $(m \geq n)$. If $\exists\, U \in O(m)$, $V \in O(n)$, and $\Sigma \in \mathbb{R}^{m\times n}$ with nonnegative decreasing diagonal entries $\sigma_i$ (and 0 elsewhere)

s.t. $A = U\Sigma V^T$,

then we call this a <u>singular value decomp'n</u> of $A$, with singular values $\sigma_1, \ldots, \sigma_n$.

construction of SVD

**Thm:** $A \in \mathbb{R}^{m\times n}$. $A$ has a SVD.

**Pf:** We will construct $V$ and $\Sigma$, then find $U$ s.t. $A = U\Sigma V^T$.

$V$

i) Consider: $A^TA$ is symmetric, so all eigenvals are real and $\exists$ an orthogonal diagonalizing mx $V$. (nonunique)

ii) Notice the <u>eigenvalues are all nonnegative</u>: Say $A^TAv = \lambda v$.

$$\lambda\langle v,v\rangle = \langle \lambda v, v\rangle = \langle A^TAv, v\rangle$$
$$= \langle Av, Av\rangle. \; ; \text{ so } \lambda = \frac{\langle Av, Av\rangle}{\langle v,v\rangle} \geq 0.$$

iii) Reorder the columns of $V$ to correspond to eigenvalues of $A^TA$
$$\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n \geq 0.$$

iv) <u>Observe:</u> $\text{rank}(A) = \text{rank}(A^TA) = \#$ of nonzero evals $= r$ say (since $A^TA$ is symmetric)

$\Sigma$

So we have $\lambda_1 \geq \ldots \geq \lambda_r > 0$; $\lambda_{r+1} = \ldots = \lambda_n = 0$.

Let $\sigma_i = (\lambda_i)^{1/2}$, and construct $\Sigma = \left[\begin{array}{ccc|c} \sigma_1 & & & \\ & \ddots & & 0 \\ & & \sigma_r & \\ \hline & 0 & & 0 \end{array}\right]$.

(call it $\Sigma_1$)

v) Let $V_1 = (v_1, \ldots, v_r)$ ; $V_2 = (v_{r+1}, \ldots, v_n)$
(where $V = (v_1, \ldots, v_n)$)

Observe 1) $A^TAv_j = 0$ for $j = r+1, \ldots, n$.

2) dim nullsp. $(A^TA) = n-r$.

Since $v_{r+1}, \ldots, v_n$ are orthogonal, they span nullspace $(A^TA)$.
$= N(A)$.

3) Thus $AV_2 = 0$

(Note: $V = [V_1\ V_2]$ so $VV^T = \ldots$)

4) And $I = VV^T = V_1V_1^T + V_2V_2^T$
$AI = \ldots = AV_1V_1^T$.

$U$ vi) Constructing $U$. s.t. $AV = U\Sigma$.

Consider the first $r$ columns. $Av_j = \sigma_j u_j$ ; $j = 1, \ldots, r$ is necessary

So it's easy to see that $u_j = \frac{1}{\sigma_j} Av_j$ ; $j = 1, \ldots, r$.
(Note $AV_1 = U_1\Sigma_1$)

Observe:

1) The $\{u_j\}$ form an ON set. (easy)

2) $u_j \in R(A)$, which has dim$^n$ $r$; so $\{u_1, \ldots, u_r\}$ are an ONB for $R(A)$

(So that $U$ is invertible). 3) Choose an ONB $\{u_{r+1}, \ldots, u_m\}$ for $N(A^T) = R(A)^\perp$;

Let $U = [U_1\ U_2]$ $U_2 = (u_{r+1}, \ldots, u_m)$

$A = U\Sigma V^T$ Then $U\Sigma V^T = [U_1\ U_2]\begin{bmatrix}\Sigma_1 & 0 \\ 0 & 0\end{bmatrix}\begin{bmatrix}V_1^T \\ V_2^T\end{bmatrix}$

$= U_1\Sigma_1V_1^T = AV_1V_1^T = A$ ∎

Simple observations:

observations 1. $\sigma_1, \ldots, \sigma_n$ are unique; $U$ and $V$ are not.

Constructing $U$ and $V$:

2. $V$ diagonalizes $A^TA$, so $v_j$ are eigenvectors of $A^TA$.
3. $AA^T = U\Sigma\Sigma^T U^T$; so $U$ diagonalizes $AA^T$.

$\Rightarrow$ $u_j$'s are eigenvectors of $AA^T$.

4. Recall: $AV = U\Sigma \Rightarrow \boxed{Av_j = \sigma_j u_j}$; $j = 1, \ldots, n$.

Similarly, $V^TA^T = \Sigma^T U^T$. $U, V$ are orthogonal, so

$$A^T U = V\Sigma^T$$

$$\Rightarrow \boxed{A^T u_j = \sigma_j v_j}; \; j = 1, \ldots, n.$$

Defn: $u_j$ called the left singular vectors of $A$;
$v_j$ called the right singular vectors of $A$.

Steps: 1. Find eigenvals of $A^TA$.

2. Find (orthogonal) diagonalizing mx $V$ for $A^TA$
(and thus all right singular vectors)

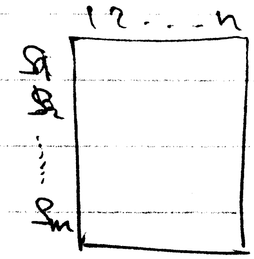
3. Find left singular vectors via $u_i = \frac{1}{\sigma_i} Av_i$; $i = 1, \ldots, r$.

4. Create ONB for $N(A^T)$.

5. Use 3 & 4 to create $U$, then

$$A = U\Sigma V^T.$$

# Principal Component Analysis:




Here's the idea: $m$ individuals, $n$ tests.

Create an $m \times n$ matrix $X$ of <u>deviations from the mean</u>.

<u>Observe</u>: Columns of $X$ will likely be correlated;

however, we hope our "factors" accounting for the scores are <u>uncorrelated</u>.

I.e., we hope to find

i) Orthogonal vectors $\vec{y}_1, \ldots, \vec{y}_r$ that ~~span~~ form a basis for $R(X)$

ii) Further the $\{\vec{y}_i\}$ must be in order of decreasing variance. (i.e. captures the most information)

~~Recall~~

Rk: Observe that $y_i \in R(X)$ implies $y_i = X v_i$ for some $v_i$

We want to maximize variance of $\underline{\mathbf{y_1}}$

<u>FACT</u>:

It suffices to choose $\vec{v}_1$ to be a unit eigenvector of $X^TX$ belonging to the max. eigenvalue $\lambda_1$.

Since eigenvectors of $X^TX$ are the right singular vectors of $X$, we simply choose the right singular vector $\vec{v}_1$ corresponding to the largest singular value.

Then $y_1 = X v_1 = \sigma u_1$ is the first principal component vector

Similarly the other PC vectors.

Point: Doing a SVD of $X$ gives me immediately (orthogonal) vectors corresponding to eigenvalues of $X^T X$ decreasing magnitude, (i.e. of maximal variance in the dataset)

~~Pb: Every vector should be expressible as~~ a linear comb^n of the component vectors!!

There is no reason this shouldn't work!!!

PCA:

Basic notations:

1) Suppose we have n data points for a variable x. $(\text{data set } X_1)$ $(d_1, \dots, d_n)$

2) Let $\bar{x}$ = mean value;

let $\mathbb{x} = (d_1 - \bar{x}, d_2 - \bar{x}, \dots, d_n - \bar{x})$ $= (x_1, \dots, x_n)$

be the vector of deviations from the mean.

3) Defn: Variance $:= \frac{1}{n-1} \sum_{i=1}^{n} x_i^2 = \frac{\mathbb{x} \cdot \mathbb{x}}{n-1}$.

Std dev $:= (\text{variance})^{1/2}$

---

Now, given 2 sets of data, we create

$\mathbb{x}_1$ and $\mathbb{x}_2$ (deviations from mean),

$$\mathrm{cov}(X_1, X_2) := \frac{\mathbb{x}_1 \cdot \mathbb{x}_2}{n-1}$$

~~Res~~ Can form a covariance matrix.

---

Ideas: Form a mtx $X$ of dev. from mean

and factor into $UW$.

columns correspond to hypothetical factors?