

Linalg

- scalars: real numbers and stuff, \mathbb{R}
- vectors: a list of scalars.
 - we use column vectors: $\begin{bmatrix} 1 \\ \pi \end{bmatrix} \in \mathbb{R}^2$, etc
 - geometric picture: \mathbb{R}^2 is cartesian plane, \mathbb{R}^3 is 3d space, etc
- operations:
 - scalar multiplication: $2 \cdot \begin{bmatrix} 1 \\ \pi \end{bmatrix} = \begin{bmatrix} 2 \\ 2\pi \end{bmatrix}$. geometrically: “stretching/shrinking”
 - vector addition: $\begin{bmatrix} 1 \\ \pi \end{bmatrix} + \begin{bmatrix} 2 \\ -\pi \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \end{bmatrix}$. geometrically: “tail to tip”
- standard vectors (think $\hat{i}, \hat{j}, \hat{k}$ from 3dcalc)
 - \mathbb{R}^2 :
 - $\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$
 - $\vec{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$
 - \mathbb{R}^3 :
 - $\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$
 - $\vec{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$
 - $\vec{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$

- BIG POINT: every vec in \mathbb{R}^n is a linear combination of standard vectors
 - ie, $\begin{bmatrix} 1 \\ \pi \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ \pi \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \pi \begin{bmatrix} 0 \\ 1 \end{bmatrix}$
- future language: $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$ span \mathbb{R}^n
- matrices: an array of scalars:

$$\begin{bmatrix} 1 & 0 & 7 \\ 0 & \pi & \frac{1}{2} \end{bmatrix}$$

- variables to use: A, B, C, U, V

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & \\ \vdots & \vdots & & \\ a_{m1} & & & \end{bmatrix}$$

- A_{i*} means the i th row

Matrix Multiplication

- definition:

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

Figure 1: matrix multiplication. use dot products. see gif.

- matrix mult. can also be viewed row-wise and column-wise (medium post link)
 - let $AB = C$
 - a col of C is a linear combination of cols of A ; weighted by the corresponding col of B
 - a row of C is a linear combination of rows of B , weighted by the corresponding row of A

$A = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix}$ $B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \\ B_{31} & B_{32} \end{bmatrix}$ $AB = C = \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} + A_{13}B_{31} & A_{11}B_{12} + A_{12}B_{22} + A_{13}B_{32} \\ A_{21}B_{11} + A_{22}B_{21} + A_{23}B_{31} & A_{21}B_{12} + A_{22}B_{22} + A_{23}B_{32} \\ A_{31}B_{11} + A_{32}B_{21} + A_{33}B_{31} & A_{31}B_{12} + A_{32}B_{22} + A_{33}B_{32} \end{bmatrix}$

columns of C is linear comb. of columns of A .
 a col of C is A_iB_{*1} .
 Thus, columns of C are linear comb. of A 's weighted by row of B .

rows of C is linear comb. of rows of B .
 a row of C is $A_{*1}B_{11} + A_{*2}B_{21} + A_{*3}B_{31}$.
 Thus, rows of C are linear comb. of B 's weighted by row of A .

first col of C : $\begin{bmatrix} A_{11} \\ A_{21} \\ A_{31} \end{bmatrix} B_{11} + \begin{bmatrix} A_{12} \\ A_{22} \\ A_{32} \end{bmatrix} B_{21} + \begin{bmatrix} A_{13} \\ A_{23} \\ A_{33} \end{bmatrix} B_{31}$
 $= A_{11}B_{11} + A_{21}B_{21} + A_{31}B_{31}$

- matrix-vector mult is just viewing vector as column vector, 1 wide matrix
- matrix mult properties:
 - $m \times n$ matrix times $n \times o$ matrix is a $m \times o$ matrix
 - Associative property for matrices: $A(BC) = (AB)C$
 - Commutative property of scalars: $c(AB) = (cA)B = A(cB)$
 - Left distributive property: $A(B + C) = AB + AC$
 - Right distributive property: $(B + C)A = BA + CA$
 - Identity property: $I_m A = A = AI_n$
- transpose matrix (A^T): just leetcode flip 2d array
- $\vec{x} \cdot \vec{y} = \vec{x}^T \vec{y}$ when you view right half as matrix mult. as scalar
- TODO something about why $(AB)^T = B^T A^T$?
- special matrices
 - $n \times n$: square
 - I_n : the $n \times n$ identity matrix
 - $AI = A$ and $IA = A$ for any A
 - permutation matrices: sudoku the 1s
 - $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} A_{1*} \\ A_{2*} \\ A_{3*} \end{bmatrix} = \begin{bmatrix} A_{2*} \\ A_{3*} \\ A_{1*} \end{bmatrix}$
 - diagonal matrices: entries not on diagonal are 0. usually square.
 - geometrically, they “scale our axes”

- jacobian matrices :3
- ▶ upper triangular: $A_{ij} = 0$ if $i > j$. i.e. entries *strictly* lower than diag must be 0.
- lower triangular: see above.
- swap between lower/upper triang. via transposes
- used to solve systems of equations

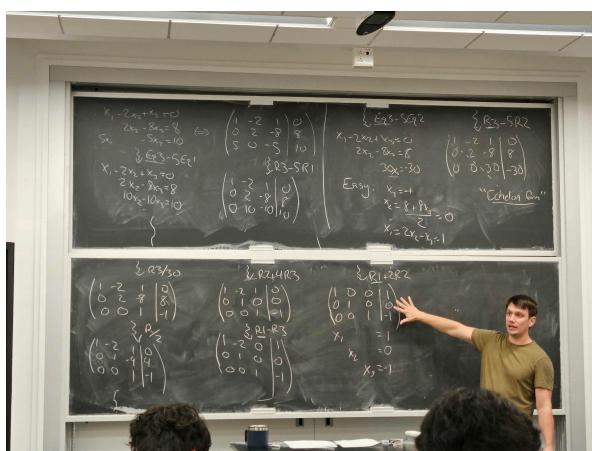
systems of linear equations

- e.g. solve $A\vec{x} = \vec{b}$ where A is a upper triang. matrix

$$A = \begin{bmatrix} 1 & -1 & 2 \\ 0 & 8 & -2 \\ 0 & 0 & 3 \end{bmatrix}, \vec{b} = \begin{bmatrix} 3 \\ 4 \\ 6 \end{bmatrix}$$

- a linear equation is of the form $a_1x_1 + a_2x_2 + \dots + a_nx_n = c$
- review: solving systems in general

get from system of equations matrix to echelon form matrix to reduced echelon form



- elementary row operations
- ▶ elementary matrices E correspond to these.
 - use via EA .
 - these are invertible
- ▶ all reversible
- ▶ scaling rows: `row *= c`

$$E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

Figure 2: scale 3rd row by 3

- ▶ replace rows: `row += c * other_row`

$$E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix}$$

Figure 3: add 3R1 to 3rd row

- ▶ interchange rows
 - permutation matrices: swap the rows of I you want to swap
- two systems are equivalent (have same soln set) if their corresponding augmented matrices are two equivalent
- a linear system is *consistent* if it has solutions (1 or inf) and *inconsistent* otherwise (no solutions)
 - ▶ a system is consistent iff the rightmost column of the augmented matrix is *not* a pivot column
 - ▶ if system consistent, then (inf. many solutions iff free variables exist)
 - ▶ inconsistent class example: we did the solve path and got an echelon form of

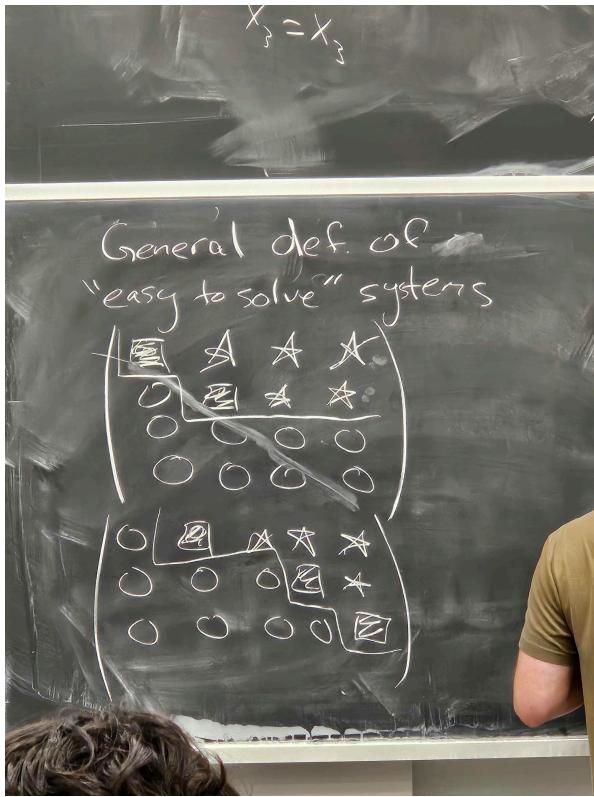
$$\left[\begin{array}{ccc|c} 2 & -3 & 2 & 1 \\ 0 & 1 & -4 & 8 \\ 0 & 0 & 0 & 15 \end{array} \right]$$

but $0 \neq 15$.

- ▶ consistent inf solutions class example: did solve path and got solution set of

$$\left\{ \vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 5 \\ -1 \\ 1 \end{bmatrix} : x_3 \in \mathbb{R} \right\}$$

- Echelon Form (EF) “easy to solve” definition:
 - ▶ all nonzero rows above zero rows
 - ▶ each leading entry (first nonzero entry) of a row is in a col to the right of the leading entry above
 - ▶ all entries below leading entry are zero



- solving “easy to solve” Echelon Form get Reduced EF (REF)
 - subset of Echelon form
 - unique
 - all leading entries are 1s
 - each leading entry is the only nonzero in its column

Last Lecture: REF/EF

$$\left(\begin{array}{ccccc} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array} \right)$$

- A pivot position corresponds to a leading 1 entry in the REF
- A pivot column is a column of A that contains a pivot position.

Gaussian Elimination

- how we get to the reduced echelon form
- eg general solution to:

$$\left[\begin{array}{cccc|c} 1 & -2 & -1 & 3 & 0 \\ -2 & 4 & 5 & -5 & 3 \\ 3 & -6 & -6 & 8 & -3 \end{array} \right]$$

R2 + 2R1, R3 - 3R1, R3 + R2, R2 / 3, R1 + R2

- let non pivot columns be free variables

The image shows a chalkboard with handwritten mathematical work. On the left, there's a matrix in row echelon form with annotations: "Last lecture: REF/EF", "this form is unique", "leading entry = 1st nonzero", "A pivot position is a leading entry in a row", and "A pivot column is a column containing a pivot position". On the right, the matrix is being transformed into reduced row echelon form (REF) through row operations like R1+R2, pivoting, and scaling. Annotations explain the steps and concepts like pivot columns and free variables.

Figure 4: full example of solving a system

- algorithm:
 - ▶ “forward phase” gets to EF
 1. find leftmost nonzero column. (a pivot pos at top.) use row swaps to move a nonzero entry to pivot pos
 2. use row replacement to get zeros below the pivot.
 3. recurse: ignore top row, repeat prev. 2 steps on rest of matrix
 - ▶ “backward phase” get to REF
 4. use rescaling to make all leading entries 1
 5. working left/upward, zero entries above pivot positions

Homogeneous Systems

- a linear system is *homogeneous* iff it has the form $A\vec{x} = \vec{0}$
 - ▶ (non/in homogeneous if that 0 is a nonzero \vec{b})
 - ▶ always consistent: $\vec{x} = \vec{0}$ is always a (trivial) solution
- but do they have nontrivial solutions?
 - ▶ just solve it (g. elim.)
- if a homogeneous and nonhomogeneous system differs only in their $\vec{0}$ vs \vec{b} vector, their solution sets will only differ by a translation
 - ▶ eg you might get

$$\vec{x}_{\text{homogenous}} = x_1 \begin{bmatrix} \frac{1}{5} \\ 0 \\ 1 \end{bmatrix} \text{ and } \vec{x}_{\text{inhomogenous}} = \begin{bmatrix} 0 \\ 8 \\ -2 \end{bmatrix} + x_1 \begin{bmatrix} \frac{1}{5} \\ 0 \\ 1 \end{bmatrix}$$

- ▶ i.e., if $A\vec{x} = \vec{b}$ is consistent and \vec{p} is a solution (i.e. $A\vec{p} = \vec{b}$), then the solution set to $A\vec{x} = \vec{b}$ is $\{\vec{p} + \vec{v} : A\vec{v} = 0\}$
 - (note that for $A\vec{x} = 0$, the soln set is $\{\vec{x} : A\vec{x} = 0\}$)

Matrix inverses

- valid for square matrices

Matrix inverses (only for square matrices)

Solving $A\vec{x} = \vec{b}$

$AB \neq BA$ \Rightarrow Can we divide by A ?
Note: $\vec{x} = \begin{pmatrix} 3 \\ 5 \end{pmatrix}$
Multiplying by $\frac{1}{A}$ (inverses)
 $\left(\frac{1}{A} \right) A = I$

Solve $\begin{pmatrix} 2 & 5 \\ -3 & -7 \end{pmatrix} \vec{x} = \begin{pmatrix} 1 \\ -3 \end{pmatrix}$

$\begin{pmatrix} -7 & -5 \\ 3 & 2 \end{pmatrix} \begin{pmatrix} 2 & 5 \\ -3 & -7 \end{pmatrix} \vec{x} = \begin{pmatrix} 1 & -5 \\ -3 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ -3 \end{pmatrix}$

$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ -3 \end{pmatrix}$

Figure 5: matrix inverses primer/motivating examples

- a matrix A is *invertible* if there is a matrix C s.t. $CA = I$ and $AC = I$
 - ▶ in this case, $C = A^{-1}$
- if A, B invertible, then AB invertible and $(AB)^{-1} = B^{-1}A^{-1}$
 - ▶ imagine $AB\vec{x}$ is B first acting on \vec{x} then A acting on that. to undo, “shoes then socks” stack data structure method, we first A^{-1} to undo the A and then likewise with B , getting $B^{-1}A^{-1}$
- finding inverse not always possible. let $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, trying to find A^{-1} .

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} a+c & b+d \\ a+c & b+d \end{bmatrix}$$

but $0 \neq 1$.

- non-invertible matrices are called *singular*
- if A invertible, $A\vec{x} = \vec{b}$ has a unique solution $\vec{x} = A^{-1}\vec{b}$
- how to try finding A^{-1}
 - ▶ let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and the determinant D be $ad - bc$.
 - ▶ if $D \neq 0$, then $A^{-1} = \frac{1}{D} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$
 - ▶ if $D = 0$, A is singular
- in general for $n \times n$ matrices, find A^{-1} via row reduction
 - ▶ form the matrix $[A | I]$ and put it in REF
 - ▶ if A is invertible, we get $[I | A^{-1}]$
 - ▶ if A is not invertible, then REF of A is not I

Find A^{-1} if $A = \begin{pmatrix} 2 & 5 \\ -3 & -7 \end{pmatrix}$

$\begin{pmatrix} 2 & 5 & | & 1 & 0 \\ -3 & -7 & | & 0 & 1 \end{pmatrix}$

$\xrightarrow{\text{R}_2 + \frac{3}{2}\text{R}_1}$ $\begin{pmatrix} 2 & 5 & | & 1 & 0 \\ 0 & \frac{1}{2} & | & \frac{3}{2} & 1 \end{pmatrix}$

$\xrightarrow{\text{R}_2 \cdot 2}$ $\begin{pmatrix} 2 & 5 & | & 1 & 0 \\ 0 & 1 & | & 3 & 2 \end{pmatrix}$

$\xrightarrow{\text{R}_1 - 5\text{R}_2}$ $\begin{pmatrix} 1 & 0 & | & -7 & -5 \\ 0 & 1 & | & 3 & 2 \end{pmatrix}$

Figure 6: finding matrix inverse example

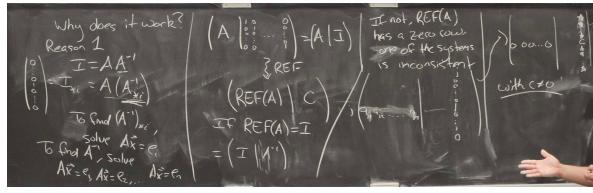


Figure 7: proof of why this method works

- elementary matrices are invertible (noted above).



Figure 8: inverting elementary matrices

- so if REF of A is I , there are E_1, \dots, E_m s.t.

$$(E_m E_{m-1} \dots E_2 E_1) A = I$$

thus, A^{-1} is simply the product of the elementary matrices.

additionally,

$$A = E_1^{-1} \dots E_{m-1}^{-1} E_m^{-1}$$

following from socks/shoes above (link)

Vector spaces

- vector spaces: \mathbb{R}^n
- span: if $\vec{v}_1, \dots, \vec{v}_n \in \mathbb{R}^m$, their span is the set of all linear combinations
 - i.e. $\text{span}\{\vec{v}_1, \dots, \vec{v}_n\} = \{c_1 \vec{v}_1 + \dots + c_n \vec{v}_n : c_1, \dots, c_n \in \mathbb{R}\}$
- consider the question: determine if $\vec{b} \in \text{span}\{\vec{v}_1, \dots, \vec{v}_n\}$.
 - i.e., are there constants such that $c_1 \vec{v}_1 + \dots + c_n \vec{v}_n = \vec{b}$.
 - if we view each \vec{v} as a column vector of a matrix A , we realize we are just asking: solve $A\vec{c} = \vec{b}$
- thus, view columns of a matrix as spanning some linear space (line, plane, hyperplane)
 - $A\vec{x} = \vec{b}$ is consistent iff $\vec{b} \in \text{span}\{\text{the cols of } A\}$
- theorem:
 - the columns of A span \mathbb{R}^n (n being “height” of A)
 - iff $\forall \vec{b} \in \mathbb{R}^n$, $A\vec{x} = \vec{b}$ is consistent
 - iff A has a pivot pos in every row
- theorem:
 - the columns of A are *linearly independent*
 - iff $A\vec{x} = 0$ has only one solution (the trivial $\vec{x} = \vec{0}$)
 - iff A has a pivot pos in every column
 - (intuit: all vecs point in “different directions”)
 - (the column vecs are otherwise *linearly dependent*)
- thrm: if $V = \{\vec{v}_1, \dots, \vec{v}_n\}$ are linearly dependent, $\exists \vec{v}_j \in V, \vec{v}_j \in \text{span}\{V \setminus \vec{v}_j\}$
 - i.e., if vecs are linearly dependent, one is in the span of the others

- thrm: if $\vec{v}_1, \dots, \vec{v}_n \in \mathbb{R}^m$ and $n > m$, then the vecs are linearly dependent.
 - informal proof: form the matrix A from col vecs \vec{v} . A is n wide, m tall. since $n > m$, this means there *must* be some pivot column and therefore inf. solutions.
- let $B = \{\vec{b}_1, \dots, \vec{b}_m\}$. B is a basis for \mathbb{R}^m if $\text{span}\{B\} = \mathbb{R}^m$ and B is linearly independent.
 - all bases of \mathbb{R}^m have the same size (square! think: why? hints above...)
- more generally, let $B = \{\vec{b}_1, \dots, \vec{b}_m\}$. B is a basis for a subspace H if $\text{span}\{B\} = H$ and B is linearly independent.
 - all bases of H have the same size (not necessarily square), calle the dimension of H , $\dim H$
- a subspace is denoted $H \subseteq \mathbb{R}^m$
 - properties:
 1. $0 \in H$
 2. if $\vec{u}, \vec{v} \in H$, then $\vec{u} + \vec{v} \in H$
 3. if $\vec{u} \in H, c \in \mathbb{R}$, then $c\vec{u} \in H$
 - eg plane through \mathbb{R}^3 's origin
 - non-eg a circle in \mathbb{R}^2 . subspaces don't curve? if we escape the subspace by taking two points and adding them it's not actually a subspace.
- subspaces are spans of vectors ($H = \text{span}\{\vec{v}_1, \dots, \vec{v}_k\}$)
- “spanning set theorem”:
 - let $S = \{\vec{v}_1, \dots, \vec{v}_k\}, H = \text{span } S$.
 - if S is linearly dependent (\exists redundancy), then $\exists \vec{v}_j \in S, \vec{v}_j \in \text{span}\{S \setminus \vec{v}_j\}$
 - (this is basically a duplicate of above)
 - note that $\text{span}\{S\} = \text{span}\{S \setminus \vec{v}_j\}$
 - can keep on “removing” vectors. when you are independent, you get B , a basis for H
- let the column space of a matrix, $\text{col}(A)$, be the span of the columns of a matrix A
 - notice: this is a subspace!
- find a basis for $\text{col}(A)$:
 - suppose matrix is in REF
 - take the pivot cols of it!
 - linear dependencies from REF are same as linear dependencies of the original matrix
 - so the basis is the same exact pivot cols
 - notice that if a free variable col can be created from some linear combination of pivot cols in the REF, they can be created with the same linear combination of the same pivot cols of the original matrix!
- *rank* is the dimension of the vector space spanned by a matrix's columns