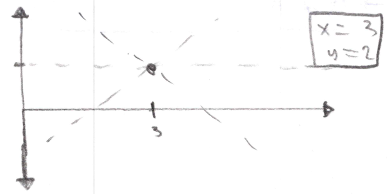
Lecture 1

old-way of solving linear system

system
$$\begin{cases} 2x - 3y = 0 \\ x + y = 5 \\ 2y = 5 \end{cases}$$
① get y=2
② get x value

Geometric Interpretation



* solution to systems of equations equivalent to intersection of lines dictated by this governing eqns

Matrix-Form of Equations

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

General Form:

$$\boxed{Av = b}$$

where

$\boxed{A}$ = (m x n) Matrix

$\boxed{v}$ = (n x 1) vector

$\boxed{b}$ = (m x 1) vector

vector v & Linear combos

$\boxed{v}$ is a point in n-dimensional space

Linear Combination

Fact: Almost any two n x 1 vectors $\boxed{v}$ and $\boxed{w}$ trace out a 2-Dimensional plane in n-dimensional space

→ via their Linear combos

* Linear combo Notation
* of Systems of Equations

$$x\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} + y\begin{bmatrix} -3 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 5 \\ 4 \end{bmatrix}$$

→ Such a solution Only exists if $\begin{bmatrix} 0 \\ 5 \\ 4 \end{bmatrix}$ lies in plane spanned by vectors.

18.06 Cheat Sheet #1

Fact: Equation

$$Av := \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}\begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$$

⇕ (Equivalent to)

$$v_1\begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix} + \dots + v_n\begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$$

→ linear combo of columns of $\boxed{A}$ equal to $\boxed{B}$

DOT Product & Properties

$$v \cdot w = v_1 w_1 + \dots + v_n w_n$$

& $\boxed{v \cdot w = \|v\| \|w\| \cos\theta}$

Lecture 2

Gaussian Elimination

- Basic Method for Solving systems of Eqns

$\boxed{A}$ is coefficient Matrix

Gaussian Elimination

System:
$$\begin{bmatrix} 1 & -1 & 2 \\ -2 & 2 & -3 \\ -3 & -1 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ -3 \end{bmatrix}$$

① form Augmented Matrix $[A \mid b]$

ex.
$$\left[\begin{array}{ccc|c} \boxed{1} & -1 & 2 & 1 \\ \boxed{-2} & 2 & -3 & -1 \\ \boxed{-3} & -1 & 2 & -3 \end{array}\right]$$
□ = Pivots (left most non zero row value)

② Perform Row Eliminations

→ Multiples of Row i to row j $(i < j)$

3 valid Row Operations:

1. Row Addition
2. Row Exchange
3. Scalar Multiplication

⇓
$$\left[\begin{array}{ccc|c} \boxed{1} & -1 & 2 & 1 \\ 0 & \boxed{1} & -2 & 0 \\ 0 & 0 & \boxed{1} & 1 \end{array}\right]$$
( ) denotes REF

Variant: Gauss-Jordan Elimination

- All pivots are $\boxed{1}$
- All values above Pivot are $\boxed{0}$

→ output is RREF

○ Use back substitution to solve

Singularity

○ If REF form has a row of all zeroes than matrix is Singular & might not have solution

REF system

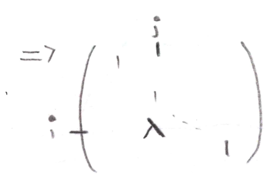
$\boxed{U}$ is REF form of $\boxed{A}$

$$\boxed{Uv = c}$$

Row Operations in Matrix Form

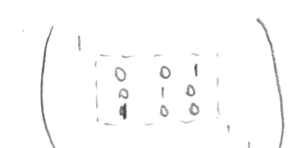
① Elimination Matrix

Adding Multiple of i^{th} row to j $\quad E_{ij}^{(\lambda)}$



② Permutation Matrix

→ swapping rows i & j $\quad P_{ij}$ (swap 2 & 4)



③ Diagonal Matrix

→ Multiplying Row by scalar $\quad D_i^{(\lambda)}$

LU Factorization

- $\boxed{L}$ is sequential Left Product of transformation Matrices (inverted!!)

$$\boxed{\begin{array}{l|l} A = LU & Ux = c \\ b = Lc & \end{array}}$$

Fact: for some matrix $\boxed{L}$ system $Av = b$ equivalent to $Ux = c$

Lecture 3

Matrix Multiplication Mechanics

$$C_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

Rules

- Associativity

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

- Distributivity

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

- $AI = IA$ (unit element)
- $A^P = A \cdot A \cdot$ (p-times)

Inverse Matrix

- Matrix $\boxed{A}$ has inverse A^{-1} that satisfies $\boxed{AA^{-1} = A^{-1}A = I}$

→ Singular Matrices have no inverse

CHECKING INVERTIBILITY

1. Begin w/ following Augmented Matrix

$$[A \mid I]$$

$$\Downarrow$$

$$[I \mid B]$$

if can get to this invertible

Inverse of Matrix Product

$$(AB)^{-1} = B^{-1}A^{-1}$$

Lecture 4

* LU Factorization

Fact: Each elimination Matrix is Lower-triangular

and

$$\left(E_{ij}^{(\lambda)}\right)^{-1} = \left(E_{ij}^{(-\lambda)}\right)$$

$\boxed{A = LU}$ U is upper triangular

Key Point → Gaussian Elimination on square matrix equivalent to product of lower & upper triangular Matrices

Lecture # 5

- A permutation Matrix contains a single $\boxed{1}$ on each row and column with zeroes elsewhere

⟹ # of permutation matrice = $\boxed{n!}$

$$\boxed{PA = LU}$$

(holds for Gaussian elimination)

P = permutation Matrix

U = (upper-triangle) RREF

L = (Lower-triangular) Matrices

Transpose $\boxed{\&}$ Properties

- Switches Rows & columns

↳ $\boxed{\&}$ Dimensions

Properties:

1. $(A+B)^T = B^T + A^T$
2. $(A^T)^T = A$
3. $(AB)^T = B^T A^T$
4. $(A^{-1})^T = (A^T)^{-1}$

Symmetric Matrices

- Symmetric if $\boxed{S = S^T}$

→ ** For any Matrix $\boxed{A}$

→ $A^T A$ is $\boxed{\text{symmetric}}$

LU Factorization of $\boxed{S}$

$$S = LDL^T$$

Lecture # 6

Vector Spaces

- Lines/Planes/Spaces exist in Vector spaces

if $\boxed{V}$ is a vector space

$\boxed{S}$ is a subspace if

- for any $v, w \in S$, $v + w \in S$
- for any $v, \alpha \cdot v \in S$

** Column Space

- column space of an m x n matrix A is a subspace

$$C(A) \subset R^m$$

Fact: A system of Equations only has solutions if & only if

$$\boxed{b \in C(A)}$$

Lecture # 7

Null Space

- consider subspace $\boxed{S} \subset R^3$ spanned by

$$\underset{v}{\begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}} \ \boxed{\&} \ \underset{w}{\begin{bmatrix} 5 \\ -2 \\ -3 \end{bmatrix}}$$

→ 2 vectors not on same line so they constitute a plane

if $A = [1\ 1\ 1]$

$A v = 0$

$\boxed{\&}$

$A w = 0$

(v & w constitute null space of $\boxed{A}$)

Null Space

- The Null Space of an (mxn) matrix $\boxed{A}$ is the subspace $S \subset \mathbb{R}^n$ consisting of vectors $\boxed{v}$ s.t $A v = 0$

Solving For Null Space

Note $N(A) = N(R)$

Step 1 Get Matrix $\boxed{A}$ into RREF Notation via Gauss-Jordan Elimination

Step 2 segregate Pivot variables & free variables

$$\boxed{\#\text{ of free vars} = \#\text{ of non pivot cols}}$$

Step 3 solve for spanning vectors of Null-Space

Lecture # 8

Complete solution for $\boxed{v}$

- Complete solution for any m x n matrix $\boxed{A}$

→ Assume $\boxed{b}$ exists in C(A)

Fact: General solution given by

$$\boxed{V_{gen} = V_{spec} + W_{general}}$$

Procedure for getting whole soln

1. Setup augmented Matrix

$$\left[\, A \mid b \,\right]$$

2. Apply Gauss Jordan elimination to achieve RREF on left side

3. Back substitute for specific solution & General Sol

Note: Put zero rows on Bottom to tell on appearance if system has sols

Matrix Rank: # of pivot columns

$$\boxed{r \leq \min(m,n)}$$

Lecture # 9

Linear Independence

- vectors $v_1, \dots v_n$ are linearly independent

- A basis is a collection of independent vectors which span subspace $\boxed{V}$

→ A basis is not unique to a vector space

Dimension

- A Dimesion of b is the # of vectors making up subspace $\boxed{V}$

** Testing for Linear Indeependance

1. Put vectors in Matrix
2. Convert to REF to find Pivot columns
3. Those columns in $\boxed{A}$ will form basis of subspace

Note: (For Square matrices only)

if $\boxed{r=n}$, $\boxed{A \text{ invertible}}$

Lecture # 10

- To an (mxn) Matrix A we defined subspaces

$\boxed{1.}$ $C(A) \subset \mathbb{R}^m$

$\boxed{2.}$ $N(A) \subset \mathbb{R}^n$

two new to consider

$\boxed{3}$ "Row Space" $C(A^T) \subset \mathbb{R}^n$

$\boxed{4.}$ "Left Nullspace" $N(A^T) \subset \mathbb{R}^m$

- These are call $\boxed{4}$ fundamental subspaces of $\boxed{A}$

Fact: If $\boxed{r}$ is the rank of Matrix $\boxed{A}$

$\boxed{1.}$ Dim $C(A) = r$

$\boxed{2.}$ Dim $N(A) = n - r$

$\boxed{3.}$ Dim $C(A^T) = r$

$\boxed{4.}$ Dim $N(A^T) = m - r$

Complementary subspaces

- Row Space & Null space complementary

- Left Null & Column complementary

Calculating $C(A^T)$ & $N(A^T)$

$\boxed{1.}$ $C(A^T)$ ⇝ Run Gauss Jordan Procedure on Transpose of $\boxed{A}$

$\boxed{2.}$ $N(A^T)$

$\boxed{A.}$ Recall $KA = R$

↳ where $\boxed{K}$ is series of transformation Matrices

B. Solve system

$wR = 0$ for w

C. $\boxed{V = wK}$

Lecture # 11

- $\boxed{2}$ vectors $v, w \in \mathbb{R}^n$ are orthagonal when their dot product is $\boxed{0}$

$$v \perp w \Rightarrow v \cdot w = 0$$

- $\boxed{8}$ if orthagonal Pythagorean theorem holds

$$\|v + w\|^2 = \|v\|^2 + \|w\|^2$$

→ But this idea can be generalized to subspaces

orthagonality of subspaces

- $\boxed{2}$ Subspaces $\boxed{V}$ & $\boxed{W}$ are orthagonal if $v \perp w$ for any $v \in V$ and $w \in \boxed{W}$

→ 2 subspaces can only be orthagonal if $\dim(V) + \dim(W) \leq \boxed{n}$ otherwise the two subspaces would intersect

→ ex. 2-planes in 3D space cant be orthagonal

recall $\begin{cases} \boxed{C(A) \perp N(A^T)} \\ \boxed{N(A) \perp C(A^T)} \end{cases}$

Proof:

- take arbitrary $\boxed{a \in C(A)}$
- take $v \in N(A^T)$

then

$a = Aw$ $\boxed{8}$ $A^T v = 0$

$$\boxed{v \cdot a = 0}$$

orthagonal compliment

- subspace $V^\perp$ s.t in consists of all vectors in $\mathbb{R}^n$ perpindicular to $\boxed{V}$

$N(A^T)$	orthagonal compliment to $C(A)$
$C(A^T)$	orthagonal compliment to $N(A)$

example: $A = \begin{bmatrix} 1 & 0 & 3 & -1 \\ 0 & 1 & -2 & 1 \end{bmatrix}$

$C(A^T)$ spanned by $\begin{pmatrix} 1 \\ 0 \\ 3 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -2 \\ 1 \end{pmatrix}$

$N(A)$ spanned by $\begin{pmatrix} -3 \\ 2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \end{pmatrix}$

- important takeaway; any vector in $\mathbb{R}^n$ can be written as linear combo of rows of $\boxed{A}$ & vector that kills $\boxed{A}$

Lecture # 12

Projections

- Given vector $b \in \mathbb{R}^n$ and linear subspace $V \subset \mathbb{R}^n$, the orthagonal Projection of b onto V is

$$\boxed{p \in V}$$

→ p is closest vector to b contained within V

$\boxed{\text{example}}$

$b = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$, $a = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$, p must be multiple of a

$p = \lambda a$ s.t $(b - \lambda a) \perp a$

$\Downarrow$ Projection Matrix (P_A)

$$p = a \frac{a^T b}{a^T a} \Rightarrow P_a = \frac{a a^T}{a^T a}$$

- For any subspace $V \subset \mathbb{R}^n$ there exists matrix $\boxed{P_V}$ $(n \times n)$ s.t

$$\boxed{\text{orthagonal projection of } b \text{ onto } V \Big\} V = P_V b}$$

Special Case: $V = C(A)$

- so given $b \in \mathbb{R}^n$

→ $\boxed{p \in C(A)}$ & $(b-p) \perp C(A)$

$\downarrow$ $\downarrow$

$p = Av$ $A^T(b-p) = 0$

$$\boxed{p = A(A^TA)^{-1}A^T b}$$

& $\boxed{\text{Projection Matrix} = A(A^TA)^{-1}A^T}$

→ if columns of $\boxed{A}$ are independent, $S = A^TA$ is invertible (no collinearity allowed)

Fact: $\boxed{P_V P_V = P_V}$ (times itself equals self)

Lecture 13

$\boxed{\text{Least Squares Problem}}$

system of Interest $\Big\}$ $\boxed{Av = b}$

→ cant solve b/c # equations greater than # of variables

Goal: minimize error vector

$\|b - Av\| \geq$ (distance from b to subspace C(A)) $= \|b - p\|$

2 steps

$\boxed{1.}$ Find orthagonal projection p of b onto subspace $C(A)$

$\boxed{2.}$ Solve $\boxed{Av = p}$ (which is possible b/c $p \in C(A)$)

Lecture 14

- Checking orthagonality
  - $\boxed{K}$ basis vectors of $\boxed{V}$
  - $\boxed{L}$ basis vectors in $\boxed{W}$

$$\boxed{V \perp W = 0}$$

- vectors orthagonal if

$$v_i \cdot v_j = 0$$

- vectors orthonormal if

orthagonal & $v_i\, v_i^T = \boxed{1}$

- Test for orthornormality

$$Q = [a_1 | a_2 | \cdot\, | a_n]$$

if $\boxed{Q^T Q = I_n}$ (vectors are orthonormal)

- Any subspace has orthonormal subspace

Graham-Schmidt Process

$$V_1 \rightsquigarrow \boxed{a_1} = \frac{V_1}{\|V_1\|}$$

$$V_2 \rightsquigarrow W_2 = V_2 - proj_{a_1} V_2$$

$$\boxed{a_2 = \frac{W_2}{\|W_2\|}}$$

→ so on so forth

Graham-Schmidt Factorization

$$Q = A \underbrace{D_1^{u_1}}_{\text{first step}} \underbrace{E_{12}^{\lambda_2} D_2^{u_2}}_{2^{nd}\text{ step}} \ldots\ldots$$

$\boxed{A = QR}$ where $R = (\ldots)^{-1}$

→ upper triangular

Lecture 15

A $\boxed{\text{linear transformation}}$ is a function: $\phi\; \mathbb{R}^n \to \mathbb{R}^m$

s.t

① $\phi(v + v') = \phi(v) + \phi(v')$

② $\phi(\alpha \cdot v) = \alpha \cdot \phi(v)$

Decomposing into base vector

$$\phi(v) = \phi(v_1 e_1 + \ldots v_n e_n)$$

$$= \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}\, v_j\, e_i$$

→ completely describes ϕ for any v

* There is a one-to-one correspondance between linear transformations & matrices

$$\phi(v) = A v$$

examples: $\mathbb{R}^2 \to \mathbb{R}^2$

① $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ reflection of line $y = x$

② $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ projection onto x-axis

③ $\begin{bmatrix} \lambda & 0 \\ 0 & u \end{bmatrix}$ scaling by factor $\boxed{\lambda}$ in x-direction, & u in y-direction

④ $\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$ rotation by θ about origin

*compound transformations

given by Matrix Multiplication

Lecture 16

- we said before $\phi: \mathbb{R}^m \to \mathbb{R}^n$ corresponded to Matrix $\boxed{A}$ $(m \times n)$ but not strictly true

say that $\phi: \mathbb{R}^2 \to \mathbb{R}^2$ represents rotation by $30°$

$$A = \frac{1}{2}\begin{bmatrix} \sqrt{3} & -1 \\ 1 & \sqrt{3} \end{bmatrix}$$ (representation in xy coordinate system)

Decomposed

$$\phi(x \cdot e_1 + y \cdot e_2) = \left(\frac{x\sqrt{3}}{2} x - \frac{y}{2}\right) \cdot e_1 + \left(\frac{x}{2} + \frac{y\sqrt{3}}{2}\right) \cdot e_2$$

→ However now we want to change reference frame

↓

get some new A'

Relation:

*** $A' = V^{-1} A V$

where,

$\boxed{V = \text{Change of base Matrix}}$

for $v_1 = 2e_1$ & $e_2 = \frac{v_1}{2} - v_2$

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

General Linear Transform Case

Consider General:

$$\phi: \mathbb{R}^n \to \mathbb{R}^m$$

$$\phi(x_1 e_1 + \dots x_n e_n) = (a_{11}x_1 + \dots a_{1n}x_n)e_1 + \dots$$

* It makes sense to change basis of $\boxed{\mathbb{R}^n}$ independently of $\boxed{\mathbb{R}^m}$

$e_1 \dots e_n \rightsquigarrow v_1 \dots v_n$ bases of $\mathbb{R}^n$

$e_1 \dots e_m \rightsquigarrow v_1 \dots v_m$ bases of $\mathbb{R}^m$

$$\boxed{A' = W^{-1} A V}$$

where

$$W = [e_1 | \dots | e_m], \quad V = [e_1 | \dots | e_n]$$

Lecture #17

The Determinant

- Determinant of square Matrix $\boxed{A}$ is the number

$$\det[A]$$

Definition: The factor w/ which a linear transformation ϕ scales the volumes of regions in n-dimensional space

examples:

① Reflection $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $\det(\,) = \boxed{-1}$

↳ preserves magnitude but switches handedness

② Projection onto x-axis

$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $\det = 0$, line has no area

$$\det\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \boxed{ad - bc}$$

$$\det(AB) = \det(A) \cdot \det(B)$$

effect of Row operations on det

① adding multiple of one row to another (unchanged)

② exchanging two rows multiplies determinent negative one (−1)

③ Multiplying a row by $\boxed{\lambda}$ multiplies det by $\boxed{\lambda}$

Important consequence for square matrix

$$\det(A) = \pm(\text{product of pivots})$$

↳ # of Row Exchanges

* if rows/columns of $\boxed{A}$ are linearly dependent,

$$\boxed{\det(A) = 0}$$

Lecture #18

- determinents are linear functions of each row independently

①
$$\det\begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ a_1+b_1 & a_2+b_2 & \cdots & a_n+b_n \\ x_{n1} & x_{n2} & \cdots & x_{nn} \end{pmatrix} = \det\begin{pmatrix} '' & & '' \\ a_1 & \cdots & \\ '' & & '' \end{pmatrix} + \det\begin{pmatrix} '' & & '' \\ b_1 & b_2 & \\ '' & & '' \end{pmatrix}$$

→ very important: Left Matrix is NOT SUM of Matrices on the right, merely same elements save for one row

Also,

②
$$\det\begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ \lambda a_1 & \lambda a_2 & \cdots & \lambda a_n \\ x_{n1} & x_{n2} & \cdots & x_{nn} \end{pmatrix} \to \lambda \cdot \det\begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ a_1 & a_2 & \cdots & a_n \\ x_{n1} & x_{n2} & \cdots & x_{nn} \end{pmatrix}$$

↳ Applied to all n rows as scalar $\quad \det('' \,'') = (\lambda)^n \cdot \det(A)$

Implementing ① on 2x2 Matrix

$$\det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \det\begin{pmatrix} a_{11} & 0 \\ a_{21} & a_{21} \end{pmatrix} + \det\begin{pmatrix} 0 & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

$$= \det\begin{pmatrix} a_{11} & 0 \\ a_{21} & 0 \end{pmatrix} + \det\begin{pmatrix} a_{11} & 0 \\ 0 & a_{22} \end{pmatrix} + \det\begin{pmatrix} 0 & a_{12} \\ a_{21} & 0 \end{pmatrix} + \det\begin{pmatrix} 0 & a_{12} \\ 0 & a_{22} \end{pmatrix}$$

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

More General Formula:

(3x3 case)

$$\det\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \det\begin{pmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{pmatrix} + D\begin{pmatrix} 0 & a_{12} & 0 \\ 0 & 0 & a_{23} \\ a_{31} & 0 & 0 \end{pmatrix}$$

$$+ D\begin{pmatrix} 0 & 0 & a_{13} \\ a_{21} & 0 & 0 \\ 0 & a_{32} & 0 \end{pmatrix} + D\begin{pmatrix} a_{11} & 0 & 0 \\ 0 & 0 & a_{23} \\ 0 & a_{32} & 0 \end{pmatrix} + D(\quad) + D(\quad)$$

* There are as many surviving determinents as possible permutations

Method of Cofactor Expansion

$$\det(A) = a_{i1}C_{i1} + \dots a_{in}C_{in}$$

↳ C_{ij} is the (i,j)-cofactor

↳ where, $C_{ij} = (-1)^{i+j} \operatorname{Det}(M_{ij})$

* This method works along any row or column

example,

$$\det\begin{pmatrix} 7 & 0 & 3 & -1 \\ 0 & 4 & 3 & 0 \\ 2 & 0 & 0 & 1 \\ 0 & 2 & 0 & 1 \end{pmatrix}$$

$$\det(A) = 2 \cdot (-1)^{3+1} \det\begin{bmatrix} 0 & 3 & -1 \\ 4 & 3 & 0 \\ 2 & 0 & 1 \end{bmatrix} + 1 \cdot (-1)^{3+4} \det\begin{pmatrix} 7 & 0 & 3 \\ 0 & 4 & 3 \\ 0 & 2 & 0 \end{pmatrix}$$

Lecture 19

→ Cramer's Rule

- method to solve nxn systems
- unique solution dictated by

$$\boxed{\nu = A^{-1} b}$$

- Take Matrix $\boxed{X}$ where element $x_{ij} = C_{ji}$

$$\det[A] = a_{i1}x_{1i} + \dots + a_{in}x_{ni}$$

Therefore, $$\boxed{A^{-1} = \frac{X}{\det A}}$$

or, $$A^{-1}_{ij} = \frac{C_{ji}}{\det(A)}$$

⇓ substituting into $A\nu = b$ system

$$\boxed{\nu_i = \frac{C_{1i} \cdot b_1 + \dots + C_{ni} \cdot b_n}{\det[A]}}$$

cofactor expansion formula!

Cramers Rule

$$\nu_1 = \frac{\det B_1}{\det A} \dots \nu_n = \frac{\det B_n}{\det A}$$

↳ where $\boxed{B}$ is Matrix A with i^{th} column Replaced w/ $\boxed{b}$

Cross Products

$$\text{cross-Product } (v \times w) = \det\begin{pmatrix} i & v_1 & w_1 \\ j & v_2 & w_2 \\ K & v_3 & w_3 \end{pmatrix}$$

→ explains why $(v \times w) = -(w \times v)$

Lecture 20

→ Goal is to "simplify" Matrices

- Matrix is Diagonalizable

- A matrix is **diagonalizable** if

$$\boxed{A = V \cdot \text{diag}_{d_1, \dots d_n} \cdot V^{-1}}$$

→ Almost all matrices are diagonalizable

Eigenvectors & Eigen values

- Given an nxn matrix $\boxed{A}$ a vector $v \in \mathbb{R}^n$ is a eigenvector if

$$\boxed{Av = \lambda v}$$

→ for some value λ, an eigenvalue

⇓ (connection to diagonal Matrices)

if $$A = V \cdot \text{Diag} \cdot V^{-1}$$

↳ all d_i's are eigenvalues and eigenvectors are columns of $\boxed{V}$

To Solve for $\boxed{\lambda}$ & $\boxed{\nu}$:

solve $$\det(A - \lambda I) = 0$$

characteristic polynomial $\{ \; p(\lambda) = \det(A - \lambda I)$

Facts

1. Sum of eigenvalues of Matrix $\boxed{A}$ is equal to Trace of $\boxed{A}$
2. Product of eigenvalues equal to $\det(A)$

example: (solving for $\boxed{\lambda}$)

$$A = \begin{pmatrix} 2 & 1 \\ 5 & 2 \end{pmatrix}$$

then, $p(\lambda) = \det(A - \lambda I)$

$$= (2-\lambda)(2-\lambda) - 5 \cdot 1 = \lambda^2 - 4\lambda - 1$$

$$\boxed{\lambda_1 = 2 + \sqrt{5}}$$
$$\boxed{\lambda_2 = 2 - \sqrt{5}}$$

for 2x2 case

$$p(\lambda) = \lambda^2 + \text{Tr}(A)\lambda + \det(A)$$

(solving for $\boxed{v}$), Knowing $\boxed{\lambda}$

- compute $(A - \lambda I)$ and put in $\boxed{\text{REF form}}$

solve system $$\begin{bmatrix} 1 & -1/\sqrt{5} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = 0$$

$$\boxed{v_1 = \frac{v_2}{\sqrt{5}}}$$ or any multiple

Lecture 21

Previously: $p(\lambda) = \det(A - \lambda I) = 0$

→ at most $\boxed{n}$ eigenvalues, but there can be less

- An nxn matrix has $\boxed{n}$ eigenvalues, (but they can be repeated, or complex)

$$\boxed{\text{if } \lambda_1 = \lambda_2, \; v_1 \text{ cannot be linear to } v_2}$$

Key: An nxn matrix w/ $\boxed{n}$ distinct eigenvalues is Diagonalizable

Lecture 20

- Goal: Make Matrices as simple as possible

- One of the simplest forms are <u>Diagonal Matrices</u>

$$\operatorname{diag}_{d_1,\dots,d_n} = \begin{bmatrix} d_1 & 0 & & \boxed{0} \\ 0 & d_2 & & \\ & & \ddots & \\ \boxed{0} & & & d_n \end{bmatrix}$$

- some linear Transformation $R_n \to R_n$ corresponds to unit vectors scaling by factor of $\boxed{d}$

→ $\det(\operatorname{diag}_{d_1,\dots,d_n}) = \boxed{d_1 * \dots * d_n}$

Diagonalizable Matrices

$$\boxed{A = V \operatorname{diag}_{d_1 \dots d_n} V^{-1}}$$

- $\boxed{A}$ matrix is diagonalizable if conjugate to a diagonal Matrix

⇒ Almost all nxn Matrices are diagonalizable

→ If you know A is diagonalizable but don't know specifics (can use method of eigenvalues & eigenvectors)

Eigenvectors & Eigenvalues

$$\boxed{A v = \lambda v}$$

- $\boxed{\lambda}$ is eigenvalue
- $\boxed{v}$ eigenvector corresponding to eigenvalue

* $\boxed{\lambda}$ Form values of Diagonal
* $\boxed{v}$ form columns of Matrix V

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

characteristic Polynomial $\{ \; P(\lambda) = \det(A - \lambda I)$

$\begin{cases} \circ \text{ sum of eigenvalues} = Tr(A) \\ \circ \text{ product of eigenvalues} = Det(A) \end{cases}$

Lecture 21

- Characteristic Polynomial has degree $\boxed{n}$ & at most n roots

→ some roots could be <u>repeated</u>

- $\boxed{\text{Algebraic Multiplicity}}$

* can be no linear relation between <u>eigenvectors</u> of <u>different eigenvalues</u>

implication for Diagonalizability

- Matrix $\boxed{A}$ with <u>n distinct eigenvalues</u> is diagonalizable

→ NOTE: ① V is not unique b/c eigenvector can multiply by a constant

② V is not invertible if rank < n.

Extension: Decomposing A to a power

if
$$A = V \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} V^{-1}$$
$$\Downarrow$$
$$A^K = V \begin{bmatrix} \lambda_1^K & \cdots & 0 \\ \vdots & \ddots & \\ 0 & \cdots & \lambda_n^K \end{bmatrix} V^{-1}$$

Lecture 22

- what if eigenvalues are <u>repeated</u>?

$$P(\lambda) = (d_1 - \lambda)^{r_1} (d_2 - \lambda)^{r_2} \dots$$

where $\boxed{r_i}$ is algebraic Multiplicity

→ $\boxed{r_1 + r_2 + \dots + r_s = n}$

Geometric Multiplicity

- Dimension of corresponding space of Eigenvectors

$$\boxed{\text{Geometric Multiplicity} < \text{Algebraic Multiplicity}}$$

example:

$A = \begin{bmatrix} d & 0 \\ 0 & d \end{bmatrix} \qquad B = \begin{bmatrix} d & 1 \\ 0 & d \end{bmatrix}$

→ Both have $\boxed{P(\lambda) = (\lambda - d)^2}$

- both have algebraic multiplicity of 2

- BUT $\boxed{A}$ has geometric Multiplicity of $\boxed{2}$ (any vector can be eigen vector)

- $\boxed{B}$ has geometric multiplicity of $\boxed{1}$ $\begin{pmatrix} x \\ 0 \end{pmatrix}$ is eigenvector

* Matrix is Diagonalizable if and only if

$\boxed{\text{Geometric Multiplicity} = \text{Algebraic Multiplicity}}$

↳ for all eigen values

Jordan Normal form

- Form Decomposition takes when $\boxed{\text{Alg Mult} > \text{Geom Mult}}$ for one or more of eigenvalues

<u>Jordan-Normal Form</u>

$$A = V \begin{bmatrix} J_1 & 0 & 0 \\ 0 & J_2 & 0 \\ 0 & 0 & J_t \end{bmatrix} V^{-1}$$

- each Jordan Block $\boxed{J_i}$ of the form

$$\begin{bmatrix} \lambda & 1 & & & \boxed{0} \\ & \lambda & 1 & & \\ & & \lambda & 1 & \\ \boxed{0} & & & \lambda & 1 \\ & & & & \lambda \end{bmatrix}$$

→ $\boxed{\lambda}$ among eigen values of a

- Total # of d_i that appear on Diagonal $= \boxed{r_i}$ (algebraic Multiplicity)

○ Fill in missing vector in $\boxed{V}$ w/ $N(A-\lambda I)^{\boxed{r_i}}$

example (2 × 2)

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

$$p(\lambda) = (2-\lambda)^2$$

$\boxed{\lambda = 2}$ $r_2 = 2$

$$N(A-\lambda I) = N\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = 0$$

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$N\left(\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\right)$$

check :

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

$$= \begin{pmatrix} 2 & 3 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} \checkmark$$

Lecture 23

○ Using Diagonalization to solve Differential Eqns

$$\dot u(t) = \text{any expression involving } u(t)$$

→ Linear Differential Equations of the form

$$\boxed{\dot u(t) = \lambda u(t)}$$

solving nth order Diff Eq

Form :

$$u^{(n)}(t) = a_{n-1}u^{(n-1)}(t) + \dots a_2 u^{(2)}(t) + a_1 \dot u(t) + a_0 u(t)$$

Notation :

$u_1(t) = u(t)$, $u_2(t) = \dot u(t)$, $u_n(t) = u^{(n-1)}(t)$

System

$$\begin{bmatrix} \dot u_1(t) \\ \dot u_2(t) \\ \vdots \\ \dot u_{n-1}(t) \\ \dot u_n(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 & & & \\ 0 & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & & 1 \\ a_0 & a_1 & a_2 & \cdots & a_{n-1} \end{bmatrix}\begin{bmatrix} u_1(t) \\ \vdots \\ u_n(t) \end{bmatrix}$$

$$\boxed{\dot u(t) = A\, u(t)} \quad \text{where } u(t) = \begin{bmatrix} u_1(t) \\ \vdots \\ u_n(t) \end{bmatrix}$$

when $\boxed{A}$ is a diagonal Matrix

$$u(t) = \begin{bmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{bmatrix}\begin{bmatrix} u_1(t) \\ \vdots \\ u_n(t) \end{bmatrix}$$

$$\boxed{\begin{bmatrix} u_1(t) \\ \vdots \\ u_n(t) \end{bmatrix} = \begin{bmatrix} c_1 e^{d_1 t} \\ \vdots \\ c_n e^{d_n t} \end{bmatrix}}$$

when $\boxed{A}$ not diagonal, but diagonalizable

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

$$\dot u(t) = V\begin{bmatrix} d_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & d_n \end{bmatrix}V^{-1}u(t)$$

$$V^{-1}\dot u(t) = DV^{-1}u(t)$$

$$\dot w(t) = D\,w(t)$$

$\boxed{\text{so}}$,

$$u(t) = V\begin{bmatrix} e^{d_1 t} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & e^{d_n t} \end{bmatrix}V^{-1}a$$

where $a = Vc$

$$\boxed{u(t) = V\begin{bmatrix} c_1 e^{d_1 t} \\ \vdots \\ c_n e^{d_n t} \end{bmatrix}}$$

Lecture 24

○ Take $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ whose

$$p(\lambda) = \boxed{\lambda^2 + 1}$$

→ imaginary #'s come into play

$\boxed{i}$ is such that $\boxed{i^2 = -1}$

or $\boxed{i = \sqrt{-1}}$

complex #'s

$$z = a + bi$$

→ $Re(z) = a$ (Real-Part)

→ $Im(z) = b$ (imaginary-part)

→ $\bar z = a - bi$ (complex-conjugate)

⇒ Can now solve for complex Eigenvalues & eigenvectors

Polar Coordinates

$$z = a + bi \rightsquigarrow (r, \theta)$$

where, $r = \sqrt{a^2 + b^2}$

$$\theta = \arccos\left(\frac{a}{\sqrt{a^2+b^2}}\right)$$

(magnitude & angle in complex plane)

• $\boxed{\theta}$ is called the argument

$$z = r(\cos(\theta) + i\sin(\theta)) = \boxed{re^{i\theta}} \quad (\text{Polar Form})$$

Properties of Polar form

$z^n = r^n e^{in\theta}$ | exponent

$zz' = rr' e^{i(\theta+\theta')}$ Product

Lecture 25

• An n×n symmetric matrix $\boxed{S}$ has $\boxed{n}$ real eigenvalues and $\boxed{n}$ orthonormal eigenvectors

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

$$\boxed{p(\lambda) = \lambda^2 - \lambda(a+c) + ac - b^2}$$

⇓ (Diagonal Decomp)

$$\boxed{S = Q \begin{bmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \lambda_n \end{bmatrix} Q^{-1} \text{ or } Q^T}$$

↳ b/c $Q^{-1} = Q^T$

Matrix "Energy"

• Positive-Definite → all eigenvalues/pivots are positive

• Positive-Semidefinite → all eigenvalues/pivots ≥ 0

$$\boxed{v^T S v \geq 0}$$

• positive-~~semi~~definite

• positive-semidefinite

Lecture 26

• Singular Value Decomposition

→ Rectangular Matrix Represents array of Pixels

→ each element represents color

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

⇓ (can be stored w/ less memory as)

$$A = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 2 & 2 & 2 & 2 & 3 & 3 \end{bmatrix}$$

Goal: Write A as a sum of products between column & row vectors

$$\boxed{A = u_1 \sigma_1 v_1^T + \dots + u_r \sigma_r v_r^T}$$

• u_i is column vector

• v_i^T is row vector

• $\sigma_1, \dots, \sigma_n$ are $\boxed{\text{singular values}}$

How to Obtain Decomposition

• m×m matrix AA^T & n×n matrix A^TA are both symmetric & positive-semidefinite

→ They both have same set of positive eigenvalues $\boxed{\sigma_1^2, \dots, \sigma_r^2}$

The vectors

Left singular vectors: $u_1, \dots, u_r, \dots, u_m \in \mathbb{R}_m$

Right singular vectors: $v_1, \dots, v_r, \dots, v_n \in \mathbb{R}_n$

→ s.t u's & v's are orthonormal

$AA^T u_i = \sigma_i^2 u_i$ & $A^T A v_i = \sigma_i^2 v_i$

↓

$\boxed{A v_i = \sigma_i u_i}$ (A) $\boxed{A^T u_i = \sigma_i v_i}$ (B)

• This holds for all $\boxed{i}$, $i > r$

compact form $\boxed{A = U \Sigma V^T}$

where,

$$U = [u_1 | \dots | u_m]$$

$$\Sigma = \begin{bmatrix} \sigma_1 & 0 & 0 & 0 & 0 \\ 0 & \ddots & & & 0 \\ 0 & & \sigma_r & & 0 \\ 0 & 0 & 0 & 0 & \ddots \end{bmatrix}$$

$$V = [v_1 \dots v_n]$$

Executing Procedure

• Solve for positive squared eigenvalues

• solve eigenvectors for whichever had fewer

• use relations (A) + (B) to solve for other basis

⇒ use Gram Schmidt procedure for any missing basis vectors

SVD (Big-Picture)

• SVD writes any matrix as a sum of rank $\boxed{1}$ Matrices

ex. $A = \begin{bmatrix} 0 & 5 & 0 \\ -3 & 0 & 0 \end{bmatrix}$

↓

$$\begin{bmatrix} 0 & 0 & 0 \\ -3 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 5 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

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

Lecture 27

- SVD tells us how $\boxed{A}$ acts on arbitrary vectors v

$$v = c_1 v_1 + \ldots + c_n v_n$$

$$\rightarrow \boxed{c_i = v_i^T v}$$

so

$$\frac{\|Av\|}{\|v\|} \leq \sigma_1$$

Key: σ_1 is max of $\frac{|Av|}{\|v\|}$, achieved when $v = v_1$

- The Norm is largest singular value of a Matrix

σ_2 is max of $\frac{\|Av\|}{\|v\|}$, when $v \perp v_1$

$\hookrightarrow$ achieved when $\boxed{v = v_2}$

Pseudo-Inverses

- If "pseudo-Inverse" matrix to $m \times n$ matrix $\boxed{A}$

- $n \times m$ "A^{-1}" satisfies

$$"A^{-1}"(u_i) = \frac{v_i}{\sigma_i}$$

Pseudo-Inverse of an $m \times n$ matrix $A = V \Sigma V^T$

$$\Rightarrow \boxed{A^+ = V \Sigma^+ U^T}$$

where, $\Sigma^+ = \begin{bmatrix} \frac{1}{\sigma_1} & & & \boxed{0} \\ & \frac{1}{\sigma_2} & & \\ & & \ddots & \\ \boxed{0} & & \frac{1}{\sigma_r} & 0 \end{bmatrix}$

Note: Pseudo-Invers the same as inverse, when the latter exists

Properties of Pseudo-Inverse

$$A^+A = V\Sigma^+U^T U \Sigma V^T = V\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} V^T$$

$$AA^+ = \text{" " ""} = U\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} U^T$$

Application to Least Squares

- Closest Av is to b achieved for $\boxed{v^+ = A^+ b}$

Polar Decomposition ($m = n$)

- Polar Decomposition of $n \times n$ matrix $\boxed{A}$ is

$$A = QS$$

$Q \rightarrow$ orthogonal

$S \rightarrow$ Positive semi-Definite

$$A = U\Sigma V^T = \underbrace{(UV^T)}_{Q} \cdot \underbrace{V \Sigma V^T}_{S}$$

Lecture 28

Applications of Linear Algebra } Probability

- Probability studies likelihood of future events

- situation with $\boxed{n}$ possible outcomes arise each with probability $\boxed{p}$ respectively

$\boxed{1.}$ $p_1, \ldots, p_n > 0$

$\boxed{2.}$ $p_1 + \ldots + p_n = 1$

- The mean is the probability-weighted sum of all outcomes

$$\mu = p_1 x_1 + \ldots + p_n x_n$$

- The variance is

$$\Sigma = \sum_{n=1}^{n} p_i (x_i - \text{mean})^2$$

- standard Deviation $\sigma = \sqrt{\Sigma}$

Continuous Probability

- Mean also called Expected Value

- We could also study a quantity that takes infinitely many (continuous) values

- In this case probability is a function

$$\boxed{p(x): \mathbb{R} \rightarrow \mathbb{R}_{\geq 0}}$$

$\rightarrow$ called a Probability Distribution

$$P(x \in [a,b]) = \int_a^b p(x)\,dx$$

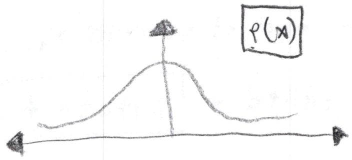
mean: $\int_{-\infty}^{\infty} p(x)\, x \, dx$

var: $\int_{-\infty}^{\infty} p(x)(x-\mu)^2 \, dx$

Uniform Prob. Distribution } $p(x) = \begin{cases} \frac{1}{c} & \text{if } x \in [0,c] \\ 0, & \text{otherwise} \end{cases}$

Normal or Gaussian Distribution } $p(x) = \frac{e^{\frac{-x^2}{2}}}{\sqrt{2\pi}}$

More on Normal Distribution



$\boxed{p(x)}$

$$\mu = \frac{1}{2\pi}\int_{-\infty}^{\infty} x e^{-\frac{x^2}{2}}\, dx = \boxed{0}$$

$$\sigma^2 = 1$$

Normal Probability Distribution of arbitrary mean & var

$$p(x) = \frac{e^{-\frac{(x-\mu)^2}{2\Sigma}}}{\sqrt{2\pi}\;\Sigma}$$

Random Variable

- A quantity $\boxed{X}$ that takes values in $\mathbb{R}$

$\boxed{1.}$ Discrete $(x_1,\dots x_n),(p_1,\dots,p_n)$

$\boxed{2.}$ continuous, $p(x)$

Properties

- $E[X] = \mu$
- $E[(X-\mu)^2] = \Sigma$

Lecture 29

- Covariance of $\boxed{2}$ Random Variables X & Y

$$\Sigma_{XY} = E\big[(X-E[X])(Y-E[Y])\big]$$

↳ Note: Σ_{XX} is just variance

- we need to know the joint probability as well

p_{ij} is the probability $(X,Y)=(x_i,y_j)$

Example: X & Y are coin-flips

- if flips x,y are independent

$$\boxed{p_{HT}=p_{HH}=p_{TT}=p_{TH}} = 0.25$$

$$\mu = \sum_{i,j} p_{ij}\, x_i$$

$$\sigma^2 = \sum_{i,j} p_{ij}(x_i-\mu)(y_j-\nu)$$

$$\nu = \sum_{i,j} y_j$$

Joint PDF: Continuous Case

$$p(x,y): \mathbb{R}^2 \to \mathbb{R}_{\ge 0}$$

→ $P(X\in[a,b],\, Y\in[c,d]) = \int_a^b\int_c^d p(x,y)\,dy\,dx$

$$\boxed{\text{covariance}} = \int_{-\infty}^{+\infty}\int_{-\infty}^{\infty} p(x,y)(x-\mu)(y-\nu)\,dx\,dy$$

Special Case: Independence

$$p(x,y) = a(x)\,b(y)$$

(For any RV discrete or continuous)

$$|\Sigma_{XY}| \le \sqrt{\Sigma_{XX}\;\Sigma_{YY}}$$

covariance Matrix of 2 RV

$$K = \begin{bmatrix} \Sigma_{XX} & \Sigma_{XY} \\ \Sigma_{YX} & \Sigma_{YY} \end{bmatrix}$$

- B/C $\boxed{\Sigma_{XY}=\Sigma_{YX}}$ is symmetric

→ Gaurenteed, Positive-semidefinite

→ Positive-Definite if X & Y not perfectly correlated

Covariance Matrix of Discrete RV

$$K = \sum_{i,j=1}^{K} p_{ij} \begin{bmatrix} x_i-\mu \\ y_j-\nu \end{bmatrix}\begin{bmatrix} x_i-\mu \\ y_j-\nu \end{bmatrix}^T$$

- For $\boxed{n}$ Random Variables

$$K = \begin{bmatrix} \Sigma_{x_1x_1} & \cdots & \Sigma_{x_1x_n} \\ \vdots & \ddots & \vdots \\ \Sigma_{x_nx_1} & \cdots & \Sigma_{x_nx_n} \end{bmatrix}$$

Random vector (vector of RV)

$$X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

- Let p_1, p_2 be the probability that $\boxed{X}$ takes certain vector values

from before

$$p_{ij} = \left(\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} x_i \\ y_j \end{bmatrix}\right)$$

Mean of Random Vector

$$\mu = \sum_i p_i x_i$$

In general, $K = E\big[(X-\mu)(X-\mu)^T\big]$

Linear combo of RV still an RV

$X' = c^T X$ for $c = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$

is a random variable

- $E[(X'-\mu)^2] = \boxed{c^T K c}$

IMPORTANT: Covariance Matrix is positive definite unless $X_1, \dots, X_n$ are Dependent example:

• RV $\boxed{x}$ & $\boxed{Y}$, what is mean & variance of

$$Z = x + y$$

→ mean of $\boxed{Z}$

$$= \begin{bmatrix}1 & 1\end{bmatrix}\begin{bmatrix}\mu \\ \nu\end{bmatrix} = \boxed{\mu + \nu}$$

→ variance of $\boxed{Z}$

$$= \begin{bmatrix}1 & 1\end{bmatrix}\begin{bmatrix}\Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{xy} & \Sigma_{yy}\end{bmatrix}\begin{bmatrix}1 \\ 1\end{bmatrix}$$

Diagonalizing the Covariance Matrix

$$\boxed{K = QDQ^T}$$

• $\boxed{D}$ Diagonal Matrix w/ non-negative eigenvalues

Principal component Analysis

Define: $\begin{bmatrix}Y_1 \\ \vdots \\ Y_n\end{bmatrix} = Y = Q^T X$; of mean $Q^T \mu$

→ $Y_1 \to Y_n$ are all certain linear combos of RV that make up $\boxed{X}$

Big Idea

$$E\left[(Y-\nu)(Y-\nu)^T\right] = \begin{bmatrix}\Sigma_r & & \boxed{0} \\ & \ddots & \\ \boxed{0} & & \Sigma_n\end{bmatrix}$$

$Y_1 \to Y_n$ are all ~~linear~~ mutually independent

→ Practical Definition of its Diagonalization

Lecture 31

∘ Formalism for continuous Distributions

$$K = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} p(x_1, \dots, x_n)\begin{bmatrix}x_1-\mu_1 \\ \vdots \\ x_n-\mu_n\end{bmatrix}\begin{bmatrix}x_1-\mu_1 \\ \vdots \\ x_n-\mu_n\end{bmatrix}^T dx \dots$$

Vector of $\boxed{n}$ Normal Distributions

$$p(y_1, \dots y_n) = \frac{e^{* \; -\frac{(y_1-\nu_1)^2}{2\Sigma_1} - \cdots - \frac{(y_n-\nu_n)^2}{2\Sigma_n}}}{\sqrt{(2\pi)\,\Sigma_1 * \cdots \Sigma_n}}$$

$$* = -\frac{1}{2}\begin{bmatrix}y_1-\nu_1 \\ \vdots \\ y_n-\nu_n\end{bmatrix}^T D^{-1}\begin{bmatrix}y_1-\nu_1 \\ \vdots \\ y_n-\nu_n\end{bmatrix}$$

→ where $D = \begin{bmatrix}\Sigma_1 & & \boxed{0} \\ & \ddots & \\ \boxed{0} & & \Sigma_n\end{bmatrix}$

Least-Squares Application

• A is a Matrix

• $b_1, \dots b_m$ are RV's with variances $\Sigma_1, \dots, \Sigma_m$

Minimize error in

$$Av \sim b = \begin{bmatrix}b_1 \\ \vdots \\ b_m\end{bmatrix}$$

Error term in experiment:

$$\varepsilon = \sum_{i=1}^{n} \frac{(b - Av)_i^2}{\Sigma_i}$$

Lecture 32

• we collect samples $x_1, \dots, x_n$

↳ called a $\boxed{\text{Dataset}}$

Linear Algebra Terminology

$$o = \begin{bmatrix}1 \\ \vdots \\ 1\end{bmatrix} \text{ and } x = \begin{bmatrix}x_1 \\ \vdots \\ x_n\end{bmatrix}$$

$$\boxed{\mu = \frac{o^T x}{o^T o}}$$

$$\Sigma = \frac{\|x - \mu o\|^2}{n-1} = \boxed{\frac{\|Px\|^2}{n-1}}$$

where $\boxed{P = I - \frac{oo^T}{o^To}}$ is the Projection Matrix onto the orthagonal compliment to $\boxed{o}$, an $(n-1)$ dimensional vector space

$$\boxed{P = \begin{bmatrix}\frac{n-1}{n} & & -\frac{1}{n} \\ & \ddots & \\ -\frac{1}{n} & & \frac{n-1}{n}\end{bmatrix}}$$

↳ $P = QDQ^T$

where, $D = \begin{bmatrix}1 & & \boxed{0} \\ & \ddots & \\ \boxed{0} & & 0\end{bmatrix}$

∘ Assume $\boxed{2}$ sets of data $x_1, \dots x_n$; $y_1, \dots y_n$

$$\Sigma_{xy} = \frac{1}{n-1}\left[(x_1-\mu)(y_1-\nu) + \dots (x_n-\mu)\dots\right]$$

$$\boxed{\Sigma_{xy} = \frac{x^T P y}{n-1}}$$

Lecture # 21

Recall we solve for characteristic polynomial: $p(\lambda)$

→ has degree $\boxed{n}$ & at most n roots

- for $\lambda_1 \neq \lambda_2$, v_1 & v_2 are linearly independent

→ $\boxed{n}$ independent eigenvectors form a basis of $\mathbb{R}^n$

Key point: nxn matrix $\boxed{A}$ w/ n distinct eigenvalues is diagonalizable

$$\boxed{A = V\begin{pmatrix}\lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n\end{pmatrix}V^{-1}}$$

↳ where V are eigen vectors

Application: Lucas Numbers

what are they?

$L_0 = 2$, $L_1 = 1$, $L_{n+1} = L_n + L_{n-1}$

consider vector: $a_n = \begin{bmatrix} L_{n+1} \\ L_n \end{bmatrix}$

$a_0 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ $\boxed{\&}$ $a_n = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} a_{n-1}$

for all $n \geq 2$ $\quad a_n = A^n a_0$

Very important Computational Point

*if $A = V\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}V^{-1}$ then

$$\boxed{A^k = V\begin{pmatrix}\lambda_1^k & & \\ & \ddots & \\ & & \lambda_n^k\end{pmatrix}V^{-1}}$$

→ k^{th} power quicker calculated via diagonalization

Exam #3 Review

Note: eigenvalues of Diagonal Matrix are entries on diagonal

* if $\boxed{A = VBV^{-1}}$ matrices

$\boxed{A}$ & $\boxed{B}$ have same ~~Eige~~ eigenvalues & characteristic polynomial

→ there different eigenvectors related by change of base matrix $\boxed{V}$

Lecture 22

→ Now what if eigenvalues are repeated

w/ repeat values $\quad p(\lambda) = (d_1 - \lambda_1)^{r_1}(d_2 - \lambda)^{r_2}\ldots$

$\boxed{r_i}$ known as Algebraic Multiplicity

$$\sum r_i = n$$

Geometric Multiplicity $= \dim N(A - d_i I)$

* $\boxed{\text{geometric Multiplicity} \leq \text{Algebraic Multiplicity}}$

- A matrix is diagonalizable only if $\boxed{\text{geo} = \text{alg}}$ for all eigenvalues

Jordan-Normal form

- if geometric Multiplicities smaller than algebraic

↳ Matrix decomposes in Jordan-Normal

$$A = V\begin{bmatrix} J_1 & & \\ & \ddots & \\ & & J_t \end{bmatrix}V^{-1}$$

where $J_i = \begin{pmatrix}\lambda & 1 & 0 \\ & \ddots & 1 \\ 0 & & \lambda\end{pmatrix}$

$\boxed{\text{where}}$

- dimensions of J_i dictated by algebraic multiplicity of $\boxed{\lambda_i}$

2 things to find out

$\boxed{1.}$ size of Jordan Blocks

$\boxed{2.}$ vector $\boxed{V}$

- vector $\boxed{V}$

* look at solving

$$N[(A - d_1 I)^{r_1}], \ldots,$$

$$N[(A - d_s I_s)^{r_s}]$$

$\boxed{\text{Numerical Example}}$ (of algebraic multiplicity 1 & 2)

$$A = \begin{pmatrix} 10 & -3 & 7 \\ 27 & -10 & 18 \\ 5 & -3 & 3 \end{pmatrix}$$

step 1: characteristic Polynomial

$p(\lambda) = -(\lambda + 1)(\lambda - 2)^2$

so, $\boxed{\begin{matrix}\lambda = 1, \text{ with } r_1 = 1 \\ \lambda = 2, \text{ with } r_2 = 2\end{matrix}}$

$v_1 = N(A + I)$

$v_2 = N(A - 2I)$

$v_3 = N((A - 2I)^2)$

Lecture 23

- Diagonalization can solve differential eqns

differential eqn: $\dot{w}(t) = a(w(t))$

$\boxed{\text{linear differential eqn}}$

$\dot{u}(t) = \lambda u(t)$

consider general case

$u^{(n)}(t) = a_{n-1} u^{(n-1)}(t) + \ldots a_0 u(t)$

Transforming into system

$u_1(t) = u(t), u_2(t) = \dot{u}(t) \ldots u_n(t) = u^{(n-1)}(t)$

so

$$\begin{pmatrix} \dot{u}_1(t) \\ \dot{u}_2(t) \\ \vdots \\ u_n(t) \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ & & \vdots & & \\ a_0 & a_1 & a_2 & & a_{n-1} \end{pmatrix} \begin{pmatrix} u_1(t) \\ u_2(t) \\ u_3(t) \end{pmatrix}$$

Form 1

$$\boxed{\dot{u}(t) = A u(t)}$$

(method of Diagonalization)

$\dot{u}(t) = V(D)V^{-1} u(t)$

$\boxed{\text{call } w(t) = V^{-1}(u(t))}$

→ solve fo diagonal matrix $\boxed{d}$ & plug back to $u(t)$

Lecture 24

- what happens when you encounter non-Real Eigenvalues

imaginary #'s $i^2 = -1$

Comelex #'s $\boxed{a + bi}$

$Re(z) = a$ (Real part of z)

$Im(z) = b$ (imaginary part of z)

$\bar{z} = (a - bi)$ (conjugate of z)

Now, say we have a matrix $\boxed{A}$ whose eigenvalues are

$\lambda_1 = i$
$\lambda_2 = -i$

Then Just solve for complex eigenvectors

Polar Form

$r = \sqrt{a^2+b^2}$ & $\cos(\theta) = \frac{a}{\sqrt{a^2+b^2}}$

$$\boxed{z = a + bi = r^{i\theta}}$$

Lecture 25

- an nxn matrix $\boxed{S}$ has n real eigenvalues and n orthonormal vectors

$$\boxed{S = Q \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} Q^T}$$

b/c for orthonormal vectors $Q^{-1} = Q^T$

$$S = \lambda_1 q_1 q_1^T + \lambda_n q_n q_n^T$$

complex eigenvalues come in pairs

ex. 2x2 matrix

$$p(\lambda) = \frac{Tr \pm \sqrt{Tr^2 - 4 \cdot det}}{2}$$

- A matrix with all eigenvalues/pivots positive is called positive definite
- if $\lambda \geq$, Positive semi-definite

* if $\boxed{S}$ is positive definite or (positive semidefinite)

$v^T S v > 0$	$(v^T S v \geq 0)$

- Matrices of form $A^T A$ are positive semi-definite

→ A 2x2 matrix is positive semidefinite if

$Tr(S)$ & $Det(S) > 0$

$\boxed{\text{Energy of 2x2}}$

$$(x\ y) \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = ax^2 + 2bxy + cy^2$$

◦ Assume we have $\boxed{m}$ datasets $\underline{n \times m \text{ matrix}}$

$$X = \underbrace{\begin{bmatrix} x_1 & y_1 & \cdots \\ \vdots & \vdots & \\ x_n & y_n & \end{bmatrix}}_{\boxed{m}}$$

$$PX = \begin{bmatrix} x_1 - \mu & y_1 - \nu & \cdots \\ \vdots & \vdots & \\ x_n - \mu & y_n - \nu & \end{bmatrix}$$

Covariance Matrix Formula

$$\boxed{K = \frac{X^T P X}{n-1}}$$

Diagonalization of Discrete Datasets

Dataset $\boxed{Y = XQ}$, has

Diagonal covariance, $\boxed{\&}$

mutually independent

Lecture 26

- Singular Value Decomposition (SVD) of a matrix very useful in computations

ex. $\boxed{\text{image Processing}}$

- each pixel represented by rectangular Matrix of integer values
- if all colors random, can't really simplify $\boxed{A}$

→ ** But sometimes images are similar

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

ex. flag of 3 vertical stripes

⇓ can decompose to form

$$A = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 & 2 & 2 & 2 & 3 & 3 & 3 \end{bmatrix}$$

$\boxed{\text{Form of SVD}}$

$$A = u_1 \sigma v_1^T + \dots u_r \sigma_r v_r^T$$

→ $\boxed{r}$ = rank of Matrix $\boxed{A}$

- u_i = column vectors
- v_i^T = row vector
- $\boxed{\sigma_i}$ = singular values

Fact: $m \times m$ matrix A^TA & $n \times n$ matrix AA^T are both symmetric $\boxed{\&}$ positive-semidefinite

→ we know: eigenvectors are orthonormal

so if

left singular vectors $\{ u_1, \dots u_r, \dots u_m \in \mathbb{R}^m$

right singular vectors $\{ v_1, \dots v_u, \dots v_n \in \mathbb{R}^n$

$\boxed{AA^T u_i = \sigma_i^2 u_i}$ or $\boxed{A^TA v_i = \sigma_i^2 v_i}$

How the two are related

$\boxed{A v_i = \sigma_i u_i}$ $\boxed{A^T u_i = \sigma_i v_i}$

holds even for $i \geq r$!! dimensions allowing

↳ where σ_i are singular values

$\boxed{\text{compact form}}$

$$A = U \Sigma V^T$$

where,

$$U = [u_1 | \dots | u_m]$$

$$\Sigma = \begin{bmatrix} \sigma_1 & & & \\ & \sigma_- & & \\ & & \sigma_r & \\ & & & 0 \end{bmatrix}$$

$$V = [v_1 | \dots | v_n]$$

$\boxed{\text{example}}$

$$A = \begin{bmatrix} 3 & 2 \\ 2 & 3 \\ 2 & -2 \end{bmatrix}$$

$$A^TA = \begin{bmatrix} 17 & 8 \\ 8 & 17 \end{bmatrix}$$

- calculate singular values $\boxed{\&}$ eigenvectors

→ use relation above

Lecture 27

$$\sigma_1 = \text{Max} \left(\frac{\|Av\|}{\|v\|} \right), \text{ achieved for } v = v_1$$