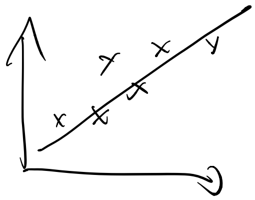
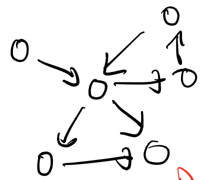
Examples $y = X\beta$

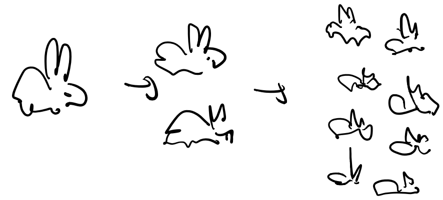


Regression
e.g. lm, glm etc



Pagerank

Graphs & networks
e.g. Markov Models



Difference Equations
e.g. $x_{n+1} = f(x_n)$

$\boxed{y = X\beta}$

$m > n$

$$\begin{aligned} y_1 &= \beta_0 + \beta_1 x_{11}^3 + \beta_2 x_{12} \\ y_2 &= \beta_0 + \beta_1 x_{21}^3 + \beta_2 x_{22} \\ y_3 &= \beta_0 + \beta_1 x_{31}^3 + \beta_2 x_{32} \end{aligned}$$

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} \\ 1 & x_{21} & x_{22} \\ 1 & x_{31} & x_{32} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix}$$

intercepts

βs are unknowns

x's are just numbers

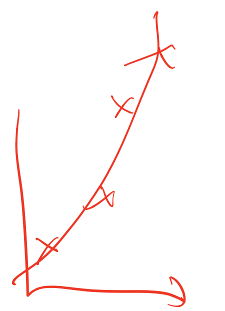
→ Illustration ~ Quadratic fit

$$X = \begin{array}{c} \begin{array}{ccc} x^0 & x^1 & x^2 \end{array} \\ \begin{bmatrix} 1 & 3 & 9 \\ 1 & 4 & 16 \\ 1 & 5 & 25 \\ 1 & 6 & 36 \end{bmatrix} \end{array}$$

$y = \beta_0 + \beta_1 x_1 + \beta_2 x^2$

want $\begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix}$

x	y
	10
4	17
5	29
6	33



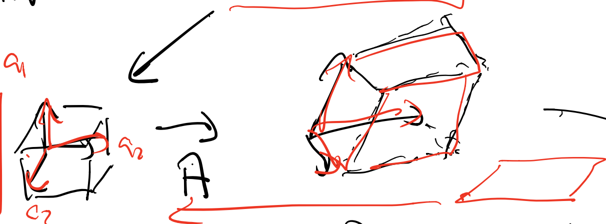
Full rank problems ($m=n=r$)

\# linearly independent rows = \# linearly independent columns

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

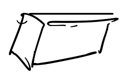
$$\frac{\mathrm{Vol}(\text{After})}{\mathrm{Vol}(\text{Before})} = |\det(A)|, \qquad \mathrm{Vol}(\text{Before}) = 1$$

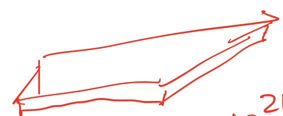
$$A = \begin{bmatrix} | & | & | \\ a_1 & a_2 & a_3 \\ | & | & | \end{bmatrix}$$



Unit cube $\quad\quad$ Parallelepiped

$|\det(A)| = \dfrac{\mathrm{Vol}(\;\cdot\;)}{\mathrm{Vol}(\;\cdot\;) = 1}$





10^{21}

$\left.\dfrac{\sigma_{max}}{\sigma_{min}}\right]$ CN

Condition \# → Almost singular (More with SVD)

$Ax = b$ $\quad m = n.$ $\quad x = A^{-1}b$

(a) $x =$ inv(A) @ b (numerically problematic)

np.linalg $\quad$ scipy.linalg

(b) $x =$ solve(A, b) $\quad$ Do this

Triangle solve

(c) $A = LU \Rightarrow L\underbrace{Ux}_{y} = b$ $\quad\quad Ly = b \leftarrow$

$Ux = y$

LDU

Similar for cholesky (symmetric A) $\quad$ P, L, U $\quad$ Please try in Jupyter.

Overdetermined problems ($n > m$)

Suppose you want $1 for nickels (5¢), dimes (10¢), & quarters (25¢)

N D Q

Many solutions

$4Q \rightarrow \begin{bmatrix} 4 \\ 0 \\ 0 \end{bmatrix} = \$1$

$10D \rightarrow \begin{bmatrix} 0 \\ 10 \\ 0 \end{bmatrix}$ etc

$20N$

$2Q + 5D$

$3Q - 5N$ (Integer valued problem)

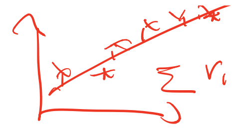
* Choose smallest vector

↓

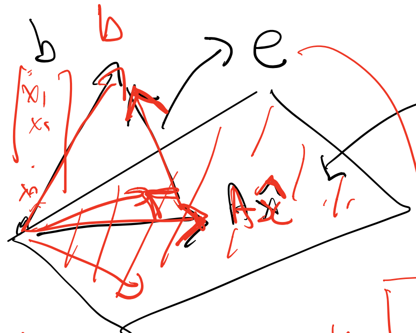
more $\boxed{\text{SVD}}$

Undetermined problems ($m > n$)

No solutions



$\sum r_i = (y_i - x_i)^2$



$$\begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$C(A) = \{\text{All vectors } Ax\}$

$= \text{All linear combinations of columns of } A$

$x_1 a_1 + x_2 a_2 + \dots + x_n a_n = y$

$\boxed{b - A\hat{x}}$

$$\text{Min } \|e\|^2 = \|b - A\hat{x}\|^2$$

$\|b - A\hat{x}\|^2$

$\hat{x}$ $\quad \|A\hat{x} - b\|^2$

Mnemonic

$$Ax = b$$

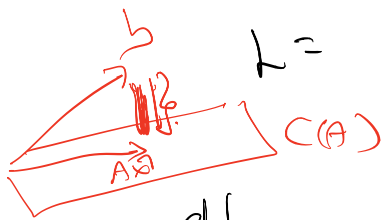
$$(A^TA)^{-1}\, A^TA\hat{x} = A^T b \qquad (A^TA)^{-1}$$

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

$\downarrow$

Normal equations

Calculus (use x for $\hat{x}$)



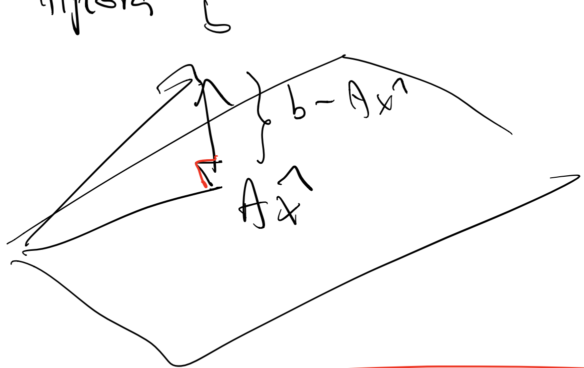
$$L = \|b - Ax\|^2 = (b - Ax)^T (b - Ax)$$

$$\left.\frac{dL}{dx} = 0 \right\} \text{solve}$$

Review vector calculus

$$\boxed{\hat{x} = (A^T A)^{-1} A^T b}$$

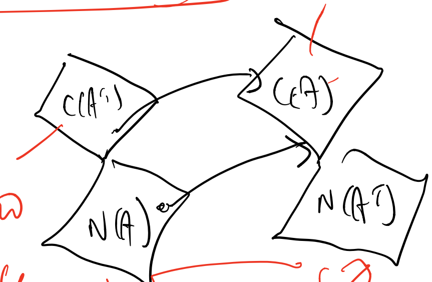
Linear Algebra



Shortest distance

$\boxed{b - A\hat{x} \text{ is orthogonal to } A\hat{x}}$

Column space



Row space

$C(A^T)$

$N(A)$

$N(A^T)$

$\Rightarrow \; b - A\hat{x} \in N(A^T)$

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

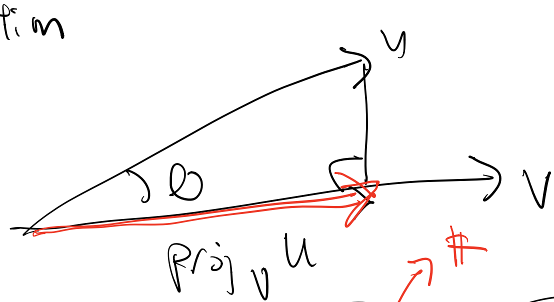
$\Rightarrow \; \boxed{A^T b = A^T A\hat{x}}$

Definition of null space

$\Rightarrow \; \boxed{\hat{x} = (A^T A)^{-1} A^T b}$

gradient descent

Projection



$$\text{proj}_v u = \frac{v^T u}{v^T v} v$$

$$\underbrace{\text{proj}_v u = \frac{v^T u}{v^T v} v}$$

Just a number

#

$$= v \frac{v^T u}{v^T u}$$

$$= \boxed{\frac{v v^T}{v^T v}} u$$

matrix

#

Projection matrix P

$$A(A^TA)^{-1}A^T b$$

$$A\hat{x} \quad \boxed{A(A^TA)^{-1}A^T b}$$

$P^2 = P$ (idempotent)

$I - P$ projects onto orthogonal subspace

Also a projection matrix

$(I-P)b$

$m > n$

least squares

normal equations → Projection

Eigenvector & Eigenvalue

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

So if we can solve problems in new basis given by eigenvectors, matrix multiplication is just SCALING.

Solving

$$Av = \lambda v$$

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

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

Characteristic polynomial
Solve for λ
Plug in λ to find v
Review.

Eigenvalue of P = projection matrix

$$Px = \lambda x$$

Defn of eigenvalue equation.

$$P^2 x = \lambda P x$$

$$Px = \lambda P x$$

$$\lambda x = \lambda^2 x$$

$$\lambda^2 x - \lambda x = 0$$

$$(\lambda^2 - \lambda)x = 0$$

Eigenvalues = 0 or 1

Change of Basis

Let S be standard basis vectors $\left\{ \begin{matrix} | \\ e_1 \\ | \end{matrix}, \begin{matrix} | \\ e_2 \\ | \end{matrix}, \begin{matrix} | \\ e_3 \\ | \end{matrix} \right\}$

Let S' be new basis vectors $\left\{ \begin{matrix} | \\ v_1 \\ | \end{matrix}, \begin{matrix} | \\ v_2 \\ | \end{matrix}, \begin{matrix} | \\ v_3 \\ | \end{matrix} \right\}$

same vector

$$x_S = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \overbrace{a_1 e_1 + a_2 e_2 + a_3 e_3}^{\text{linear combination}}$$

$$x_{S'} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \underline{b_1 v_1 + b_2 v_2 + b_3 v_3}$$

$$\begin{tikzcd} x_S \arrow[r, "P^{-1}"] \arrow[d, "A"'] & x_{S'} \arrow[d, "B"] \\ y_S & y_{S'} \arrow[l, "P"] \end{tikzcd}$$

change of basis matrix

Let $P = \begin{bmatrix} | & | & | \\ v_1 & v_2 & v_3 \\ | & | & | \end{bmatrix}$

- Trace
- Determinant
- Eigenvalues

Same linear map under different bases

$$\boxed{A = PBP^{-1}}$$

A & B are <u>similar</u> matrices

Use of spectral decomposition

"Spectral theorem"

if A is symmetric $N \times N$

$\rightarrow$ N eigenvalues (all real)

$\rightarrow$ N eigenvectors (independent)

$\rightarrow$ Eigenvectors are orthogonal. $v_i^T v_j = 0$ $i \neq j$

$$\boxed{A = V \Lambda V^{-1}}$$

$\frac{v_i}{\|v_i\|}$

eigenvector matrix

diagonal matrix of eigenvalues

If we make V an orthogonal matrix, then there is a nice geometric interpretation

$$\boxed{A = \text{rotate} \rightarrow \text{scale} \rightarrow \text{rotate}}$$

More with SVD

A^{-1} $\begin{bmatrix} \frac{1}{\lambda_1} & & \\ & \frac{1}{\lambda_2} & \\ & & \ddots \end{bmatrix}$

$V \Lambda^{-1} V^{-1}$

$A = V \Lambda V^{-1}$

$A^k = V \Lambda^k V^{-1}$

$A^2 = V \Lambda \cancel{V^{-1} V} \Lambda V^{-1}$

$V \Lambda^2 V^{-1}$ $\begin{bmatrix} \lambda_1^{100} & & \\ & \lambda_2^{100} & \\ 0 & & \lambda_3^{100} \end{bmatrix}^{100}$

Orthogonal Matrix

$$Q = \left\{ v_1, v_2, \ldots, v_n \right\} \qquad \|v_i\| = 1 \qquad v_i^T v_j = 0 \quad i \neq j$$

$$QQ^T = I \qquad Q^TQ = QQ^T = I$$

if Q is $n \times n$ $\quad Q^T = Q^{-1}$ $\qquad Q^T = Q^{-1}$

How to find orthogonal vectors

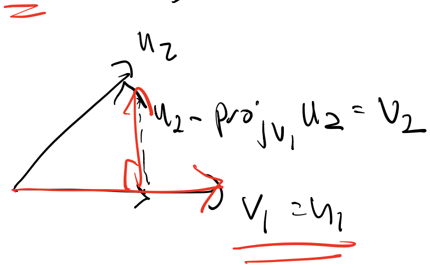
Start w/ non-orthogonal vectors $\{u_1, u_2, \ldots, u_n\}$

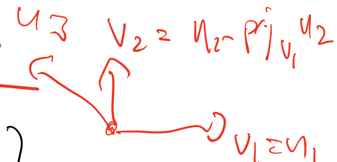
Gram-Schmidt

$$v_1 = u_1$$

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

$$v_3 = u_3 - \text{proj}_{v_2} u_3 - \text{proj}_{v_1} u_3$$





$$Q = \left\{ \frac{v_1}{\|v_1\|}, \frac{v_2}{\|v_2\|}, \frac{v_3}{\|v_3\|}, \ldots \right\}$$

QR Decomposition

$A = QR \rightarrow$ More numerical stable than LU

Eigenvalues of Q

Q is a rotation or reflection $\rightarrow$ All eigenvalues are ± 1

Stable because NO scaling. ie Condition # = 1

PCA

least squares

SVD → NSM

Best Subspaces

Low rank approximation

Geometry

R – S – R'