Linear algebra for AI and ML

(September - 29)

Stephen Boyd (2018)
(Applied LA & least squares)

G. Strang
Linear algebra and learning from data (2020)

(MIT - OCW) (2019)

Linear algebra for AI ML

- Basics of LA
  - Vectors / Matrices / Vector spaces / basis / dimension / subspaces
  - Linear transformation / affine transformation
  - (Linear functional) ← inner product, norm (ℓ_2 / spectral)
  - Matrix representation of Linear transformation
  - Gram orthogonalization / orthogonal matrices
  - K-mean clustering
- Solving systems of linear equations

$$Ax = b \qquad A \in \mathbb{R}^{m \times n}, \quad b \in \mathbb{R}^m$$

To find $x \in \mathbb{R}^n$

→ Existence and uniqueness of solution.

$$\left\{ \begin{array}{l} (x\text{-linear combination of columns of } A) \\ \langle \text{rows of } A, x \rangle = b_i \end{array} \right\}$$

$ax = b$; $a, b \in \mathbb{R}$ $\quad x = a^{-1}b = b/a$

→ Left inverse / right inverse / inverse of matrix.

For square matrices $m \times m$

→ Methods to compute the solution.

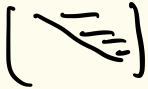
→ Gaussian elimination $\equiv$ LU decomposition of A.

→ In case, A is symmetric, $U = L^T$

$\underbrace{A = LL^T}$: cholesky decomposition of A.

→ QR decomposition (Gram-Schmidt orthogonalization)

$A = QR$



orthogonal

$Ax = b$

$\underbrace{Rx = Q^T b}$

- rotator matrices / reflector matrices

Sensitivity analysis:

$Ax = b$

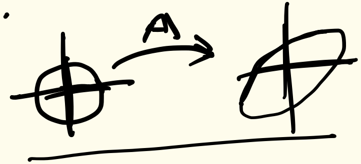
A and b are <u>NOT</u> known exactly.

→ numerical inaccuracies in storing data (1/3).

→ measurement inaccuracies.

$$\underbrace{(A+\Delta A)}\ \hat{x} = \underbrace{(b+\Delta b)} \qquad\qquad Ax = b$$

maxmag(A), minmag(A), condition number.



$$\frac{\|\delta x\|_2}{\|x\|_2} \le \underbrace{k_2(A)}\left[\frac{\|\delta b\|_2}{\|b\|_2} + \frac{\|\delta A\|_2}{\|A\|_2}\right]$$

we will visit again.

$k(A)\uparrow \Rightarrow \dfrac{\|\delta x\|}{\|x\|}$ may be large.

$k(A)\uparrow \Rightarrow$ A is "close" to being singular/non-invertible.

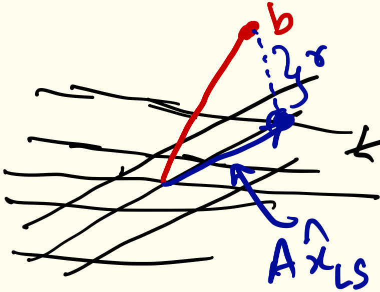
LS problems:

$Ax=b$ (non-square) (A has linearly indep. columns)

→ overdetermined

→ Existence of soln!

$$\min_{x\in\mathbb{R}^n} \|b - Ax\|_2^2 \ : \ x_{LS} = \underbrace{(A^T A)^{-1} A^T}\, b$$



columnspace(A)/
Estimation space

$r \perp A\hat{x}_{LS}$

$\Downarrow$

$(A^T A)\, x_{LS} = A^T b$

<u>normal eqn.</u>

Application of LS problems:

i) LS data fitting

ii) LS classification (boolean/multiclass)

iii) multi-objective LS problem

$J_1, \ldots, J_r$

iv) constrained LS problem

splines, (minimum norm solution)

$\lambda \uparrow \infty$

(KKT matrix)

$y = f(x)$

$f_1, f_2, \ldots, f_p$

$f \approx \underbrace{\alpha_1 f_1 + \alpha_2 f_2 + \ldots + \alpha_p f_p}$

$\boxed{\text{Pending: Sensitivity of LS problem!}}$