Least Squares

Thursday, March 12, 2020 10:47 PM

① Pseudo inverse A^+ : $Ax = b \rightarrow \boxed{\hat{x} = A^+ b}$

$A \rightarrow$ pseudo inverse A^+ $\quad AA^+ \approx I$

$m \times n \qquad n \times m$

If A^{-1} exist, $AA^{-1} = A^{-1}A = I$ then $A^+ = A^{-1}$

in case of rectangular or zero eigenvalues

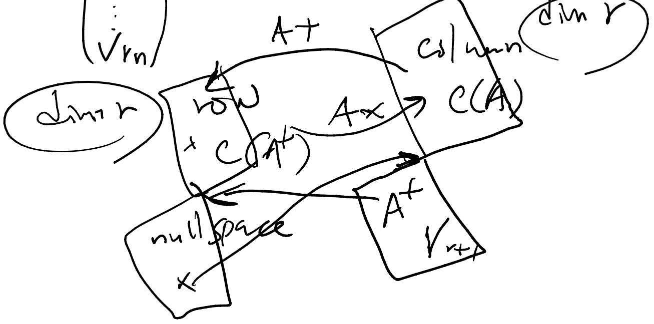
square, but has nullspace other than 0 vector

= columns are ~~dependent~~ $\quad Ax = 0 \rightarrow A^{-1}Ax = 0$

↳ null space

i) $A^+(Ax) = x$ in the row space → Ax brings back to x

ii) $A^+\begin{pmatrix} V_{r+1} \\ \vdots \\ V_{rn} \end{pmatrix} = 0$ in the null space



$A^+A = I$ on top half

$A^+A = 0$ on bottom half

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

$A = U\Sigma V^T$ — if invertible, $A^{-1} = V^{-T}\Sigma^{-1}U^{-1} = V\Sigma^{-1}U^T$

$\boxed{(V^T)^{-1} = (V^{-1})^{-1} = V}$

if Σ is rectangular, $A^+ = V\Sigma^+U^T$ $\qquad \Sigma^+\Sigma = \begin{bmatrix} 1 & & 0 \\ & 1 & \\ 0 & & 0 \end{bmatrix}$

$$\underbrace{\begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & 0 & \\ 0 & & 0 & \\ & & & 0 \end{bmatrix}}_{n} m \rightarrow \underbrace{\begin{bmatrix} 1/\sigma_1 & & 0 & \\ & 1/\sigma_2 & & \\ 0 & & & 0 \end{bmatrix}}_{m} n$$

② Solve $A^TA\hat{x} = A^Tb \rightarrow \hat{x} = (A^TA)^{-1}A^Tb$ $\qquad Ax = b \rightarrow$ when $m = n = r$, A^{-1}

$m \times n$, rank r $\qquad$ ↓ vector of measurements

minimize $\|Ax - b\|_2^2 = (Ax-b)^T(Ax-b)$

$= x^TA^TAx - 2b^TAx + b^Tb$

$\overline{x^TA^Tb}$

미분 → $\boxed{A^TA\hat{x} = A^Tb}$ (normal eqn)

Linear regression

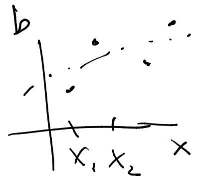
$b_1, b_2, \cdots b_m$

↓

best! fit a straight line $\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_m \end{bmatrix}\begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$

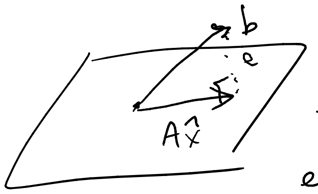
$C + Dx$ $\qquad$ rank 2

unsolvable



Assumptions : A has indep. cols (no nullspace)

A^TA invertible



C(A)

= all possible vector Ax

error = $b - A\hat{x}$

$(Ax)^T(b - A\hat{x}) = x^TA^T(b - A\hat{x}) = 0$

$$\text{error} = b - A\hat{x}$$

$$e \perp C(A) \to (b - A\hat{x}) \perp (Ax) \to (Ax)^T (b - A\hat{x}) = x^T \underbrace{A^T (b - A\hat{x})}_{=0} = 0 \to A^TA\hat{x} = A^Tb$$

$$A\hat{x} = p \text{ (projection } b \text{ onto } C(A))$$
$$= \underbrace{A(A^TA)^{-1}A^T}_{P} \underline{b} \qquad P = A(A^TA)^{-1}A^T, \quad P^2 = A\underline{(A^TA)^{-1}A^T A}(A^TA)^{-1}A^T = P$$

A^TA : attractive symmetry, but <u>size problem</u> $[\text{cond}(A)]^2$, <u>expensive</u> in large problem

applications	A^TA (or A^TCA) ill-posed
mechanical eng.	stiffness matrix
circuit theory	conductance matrix
graph theory	(weighted) graph Laplacian
instead mathematics	Gram matrix (inner product of col. of A)

(3) $A = QR$

good : orthogonal, triangular

$$A^TA\hat{x} = A^Tb$$
$$(QR)^T(QR)\hat{x} = (QR)^Tb$$
$$R^T\underline{Q^TQ}R\hat{x} = R^TQ^Tb$$
$$R\hat{x} = Q^Tb \qquad \text{\underline{safe} to solve and \underline{fast}}$$

(4) minimize $\|Ax - b\|^2 + \underline{\delta^2\|x\|^2}$

regularize : ridge regression

$$\xrightarrow[\text{eqn}]{\text{normal}} (A^TA + \underline{\delta^2 I})x_\delta = A^Tb, \quad \delta \to 0, \; x_\delta \to \hat{x}$$

$$\begin{cases} n \neq k: \sin nx \sin kx = \frac{1}{2}\left[\cos(n-k)x - \cos(n+k)x\right] \\ n = k: (\sin kx)^2 = \frac{1}{2}\left[1 - \cos 2kx\right] \end{cases}$$

$$\int_0^{\pi} S(x) \sin kx \, dx = b_k \frac{\pi}{2}$$