87강 보충강의

A : $m\times n$ 행렬 $b \in \mathbb{R}^n$ $x = A^+b \Rightarrow Ax=b$ 최소제곱해

↳ 최소 norm을 가진 해이다.

pf) $x = A^+b \in row(A)$

row(A)에 속함

* A가 full column rank가 아닐 때 $Ax=b$를 만족하는 해는 반드시 존재하고 그 해는 유일하며, 모든 해중에서 가장 작은 norm을 가진다.

pf) $Ax=b$ $x_0 = proj_{row(A)}x_0 + proj_{null(A)}x_0$ 로 유일하게 나타낼수 있다. (∵ 정사영정리)

$b = Ax_0 = A(proj_{row(A)}x_0 + proj_{null(A)}x_0)$ $Ax=0$

$= A\,proj_{row(A)}x_0 + A\,proj_{null(A)}x_0 = A\,\boxed{proj_{row(A)}x_0}$

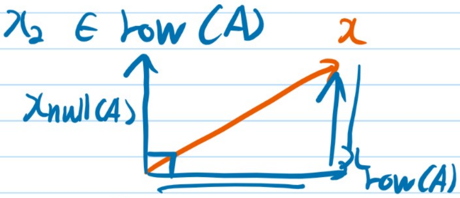
$x_1 = proj_{row(A)}x_0$ 가 $Ax=b$의 해이고, $x_1 \in row(A)$

*유일성 $Ax_2 = b$, $x_2 \in row(A)$

$A(x_1-x_2) = Ax_1 - Ax_2 = b-b = 0$

↳ $Ax=0$ 의 해 $\begin{cases} x_1-x_2 \in null(A) \\ x_1-x_2 \in row(A) \end{cases}$

$x_1 - x_2 = 0 \Rightarrow x_1 = x_2$



* 가장 작은 norm

↳ 모든 것보다도 더 작다.

임의의 해를 x라고 하자 ($Ax=b$의 해) $x = proj_{row(A)}x + x_{null(A)}$

$\|x\| = \sqrt{\|proj_{row(A)}x\|^2 + \|proj_{null(A)}x\|^2}$

$= \sqrt{\|x_1\|^2 + \|proj_{null(A)}x\|^2} \geq \|x_1\|$ ┘

* A^+b가 왜 최소 norm을 가지는가? $rank(A)=k$

pf) $A^+b \in row(A)$

$A^+b = \boxed{V'(\Sigma'^{-1}U'^Tb)} \in row(A)$

$V' = [v_1 \cdots v_k] \rightarrow v_1 \sim v_k$ row(A)의 정규직교기저 ┘

* V'은 row(A)의 정규직교기저이다. A: $m\times n$ 행렬.

pf) $rank(A)=k$ 라고 하면, $V' = [v_1 \cdots v_k]$

$A^TA = VDV^T$ $V = [v_1 \cdots v_k \mid \overbrace{v_{k+1} \cdots v_n}^{n-k\text{개}}] \in null(A^TA)$

$rank(A) = rank(A^TA) \Rightarrow A^TA$의 고유값들 $\lambda_1 \cdots \lambda_k \Rightarrow$ 나머지는 영.

$null(A) = null(A^TA)$

↳ $nullity(A) = n-k$ (차원정리에 의해. $rank(A)=k$)

$\{v_{k+1}, \cdots, v_n\}$은 $null(A)$의 기저 임을 알수있다.

$\{v_1, \cdots, v_k\}$는 $span\{v_{k+1}, \cdots, v_n\}$에 전부 직교하는 집합이다.

↳ $null(A)$

$\{v_1, \cdots, v_k\} \subset null(A)^\perp = row(A)$

$\underbrace{\text{정규직교}}_{}$ (k개) $\Rightarrow$ basis $\Rightarrow$ $row(A)$의 정규직교 기저. ($dim(row(A)) = k$)

* A가 $m \times n$ full column rank

$$A^+ = (A^TA)^{-1}A^T$$

pf) $A^TA = (V'\Sigma'^T \cancel{U'^T})(\cancel{U'}\Sigma' V'^T) = \boxed{\underbrace{V'}_{n\times n}\Sigma'^2\underbrace{V'^T}_{n\times n}}$

A가 full column rank $\Rightarrow$ A^TA가 가역이므로 $\quad V'^{-1} = V'^T$

$(A^TA)^{-1} = (V'\Sigma'^2V'^T)^{-1} = V'\Sigma'^{-2}V'^T$

$(A^TA)^{-1}A^T = (V'\Sigma'^{-2}V'^T)(U'\Sigma'V'^T)^T$

$= (V'\Sigma'^{-\cancel{2}}\cancel{V'^T})(\cancel{V'}\Sigma'^{\cancel{1}}U'^T)$

$= V'\Sigma'^{-1}U'^T = A^+$