$A_{m\times n}$

$P_{n\times n} = A(A^TA)^{-1}A^T$ exists if $\boxed{A \text{ is skinny full rank}}$

$= AA^+$

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

$Ax = b$ Not lucky

$b \notin R(A)$

$$\begin{bmatrix} & & \\ & A & \\ & & \end{bmatrix}_{m\times n} \begin{bmatrix} x \end{bmatrix} = \begin{bmatrix} b \end{bmatrix}$$

m eq's

n unknown

$n < m$

$$\boxed{\min_{x\in\mathbb{R}^n} \|Ax-b\|^2}$$

equiv

$$\min_{x\in\mathbb{R}^n} \|Ax-b\|$$

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

$f:\mathbb{R}^n \to \mathbb{R}$

Loss function

$$\nabla f(x) = \begin{bmatrix} \frac{\partial}{\partial x_1} f(x) \\ \frac{\partial}{\partial x_2} f(x) \\ \vdots \\ \frac{\partial}{\partial x_n} f(x) \end{bmatrix} = 2A^T(Ax-b)$$

$$\boxed{\nabla f(x) = 0}$$

$$\boxed{\cancel{2}A^T(Ax-b) = 0}$$

$$\boxed{A^TAx = A^Tb}$$

$Ax = b$

Normal Equations

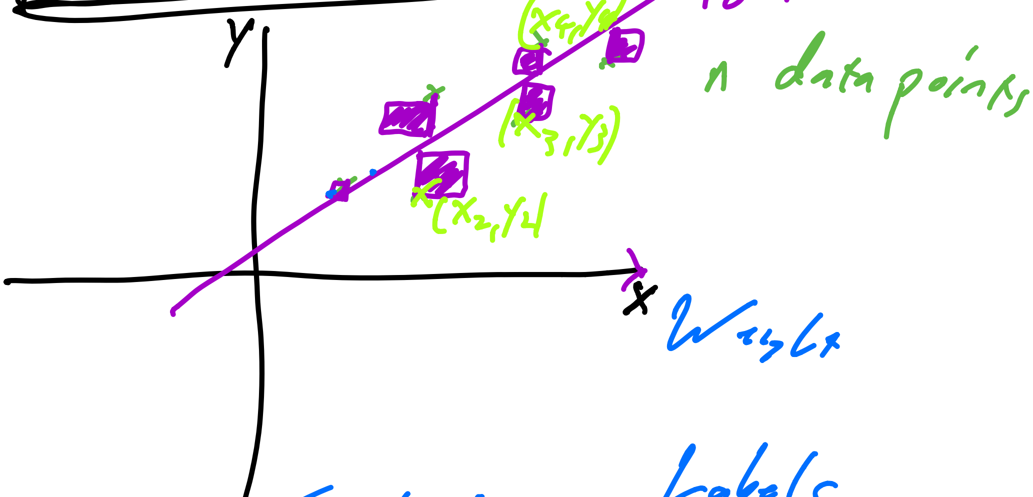
$$\hat{x} = \underbrace{(A^TA)^{-1}A^T}_{A^+} b$$

A^+ Moore Penrose pseudo-inverse

$$A^+ A \hat{x} = A^+ b$$

$$\boxed{x = A^+ b}$$

Least Squares Data Fitting



Features

Labels

$$A = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \\ 1 & x_4 \\ 1 & x_5 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}_{2\times 1} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ \vdots \end{bmatrix}$$

$\beta_0 1 + \beta_1 x_1 \approx y_1$

$\vdots$

$\beta_0 1 + \beta_5 x_5 \approx y_5$

design matrix

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

$n \times p$, $p = 2$

$$\boxed{A\beta = y}$$

$$Ax = b$$

$$\beta = A^{+} y$$

$$= (A^T A)^{-1} A^T y$$

$$A = \begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}_{n \times 2}$$

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

$$= \begin{bmatrix} \sum x_i^2 & \\ & \end{bmatrix}$$

$$A^T A$$