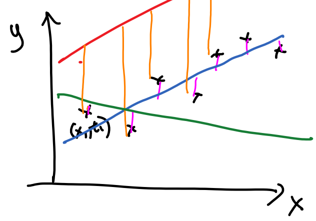
Lecture 1: Linear to Nonlinear Regression

Monday, 9 May 2022 8:58 AM

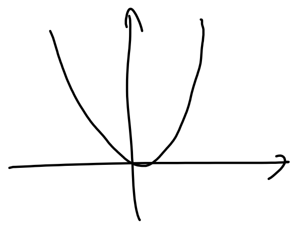


$y = w_0 + w_1 x$

Formal Def^n of Linear Reg Problem

Given $D = \{(\underline{x}_n, \underline{t}_n)\}_{n=1}^{N}$

$x_n = [x_1, x_2, x_3, \ldots, x_D]$ $\quad x \in \mathbb{R}^D$



Want: fit $y(x; w) = \underline{w_0 + w_1 x_1}$ to D

$$\underbrace{\underset{w}{\text{minimize}} \ \frac{1}{2} \sum_{n=1}^{N} \underbrace{(y(x_n, w) - t_n)}_{e_n}^2}_{E(w) \text{ (error function)}}$$

$y = 1 \cdot w_0 + w_1 x_1$

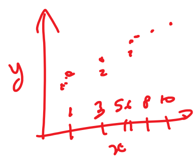
$[1 \quad x]$

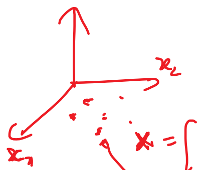
let $x_0 = 1$ then $\boxed{y(x, w) = \sum_{i=0}^{D} w_i x_i = w^T x}$

$$\boxed{\underset{w}{\text{minimize}} \ \frac{1}{2} \sum_{n=1}^{N} (w^T x_n - t_n)^2}$$

"How"

let $\mathbf{X} = \begin{array}{cc} & D \\ N & \begin{bmatrix} — x_1 — \\ — x_2 — \\ \vdots \\ — x_N — \end{bmatrix} \end{array}$ $\quad t = \begin{bmatrix} t_1 \\ t_2 \\ \vdots \\ t_N \end{bmatrix}$





$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

$\mathbf{X} = \begin{bmatrix} 1 \\ 3 \\ 5 \\ 6 \\ 8 \\ 10 \end{bmatrix}$ $\quad \mathbf{X} = \begin{bmatrix} x_{1,1} & x_{1,2} \\ x_{2,1} & x_{2,2} \\ \vdots & \\ x_{N,1} & x_{N,2} \end{bmatrix}$

$$E(w) = \frac{1}{2} (t - Xw)^T (t - Xw)$$

$$= \frac{1}{2} \left[t^T t - t^T X w - (Xw)^T t + (Xw)^T (Xw) \right]$$

$$= \frac{1}{2} \left[t^T t - 2 w^T X^T t + w X^T X w \right]$$

$$= \frac{1}{2} t^T t - w^T X^T t + \frac{1}{2} w X^T X w$$

$$\nabla_w E(w) = 0$$

$$\Rightarrow -X^T t + (X^T X) w = 0$$

$$\Rightarrow (X^T X) w = X^T t$$

$$\Rightarrow \boxed{w_{ML} = \underbrace{\boxed{(X^T X)^{-1} X^T}}_{\text{pseudo inverse}} t}$$

polynomial regression

$$y(x,w) = w_0 + w_1 x_1 + w_2 x_1^2 + w_3 x_1^3 + \dots + w_p x_1^p$$

$$= \sum_{i=0}^{p} w_i x_1^i$$

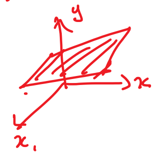
$$x \longrightarrow [1, x_1, x_1^2, x_1^2 \dots, x_1^p]$$

$$\phi(x) = [\underbrace{\phi_0(x)}_{\phi_0(x)=1}, \phi_1(x), \dots, \phi_{M-1}(x)]$$

$$y(x,w) = \sum_{i=0}^{M-1} w_i \phi_i(x) = w^T \phi(x)$$ } M parameters

$$\boxed{y(x,w) = w^T \phi(x)}$$

$x = [x_1, x_2]$



$$X = \begin{array}{cc} & D \\ N & \begin{bmatrix} - x_1 - \\ - x_2 - \\ \vdots \\ - x_N - \end{bmatrix} \end{array}, \quad \Phi = \begin{array}{cc} & M \\ N & \begin{bmatrix} - \phi(x_1) - \\ - \phi(x_2) - \\ \vdots \\ - \phi(x_N) - \end{bmatrix} \end{array}$$

$\phi(x) = [x_1^0 \; x_1^1 \; x_1^2 \dots x_1^p]$

e.g.

$$\phi = \begin{array}{cc} & M \\ & \begin{bmatrix} 1 & x_1^1 & x_1^2 & x_1^3 & \dots & x_1^p \\ 1 & x_2 & x_2 & \dots & & x_2^p \\ \vdots \\ 1 & x_N & x_N^2 & x_N^3 & \dots & x_N^p \end{bmatrix} & N \end{array}$$

W

$$\begin{bmatrix} x_1 & x_2, \\ & \end{bmatrix} \Rightarrow \begin{bmatrix} 1, x_1, x_2, x_1 x_2, x_1^2, x_2^2, \\ \end{bmatrix}$$

---

minimize w

$$\tilde{E}(w) = \underbrace{\frac{1}{2} \sum_{n=1}^{N} (y(x_n, w) - t_n)^2}_{} + \frac{\lambda}{2} \underbrace{\|w\|^2}_{\frac{\lambda}{2} w^T w} \rightarrow w_0^2 + w_1^2 + \dots + w_{m-1}^2$$

$$\begin{bmatrix} x & x & x & x \\ \\ \end{bmatrix}$$

$$\begin{bmatrix} [x, x^2], [x, x] \\ \end{bmatrix}$$

$$\nabla_w \tilde{E}(w) = 0$$

$$-X^T t + (X^T X) w + \lambda w = 0$$

$$-X^T t + (X^T X) w + \lambda I w = 0 \qquad \begin{bmatrix} 1 & & 0 \\ & 1 & \\ 0 & & 1 \end{bmatrix}$$

$$-X^T t + (X^T X + \lambda I) w = 0$$

$$\boxed{w_{LR} = (X^T X + \lambda I)^{-1} X^T t}$$