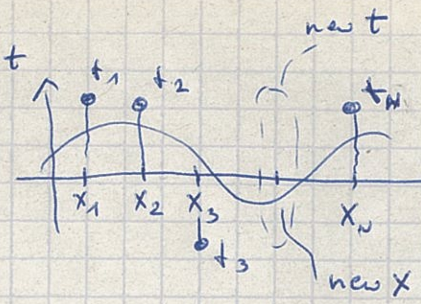
LINEAR REGRESSION (BISHOP CH 3)



DATA, OBSERVATION: $\{x_n, t_n\}_{n=1}^{N}$

MODEL OF DATA: $\underset{\text{new } t}{t} = y(\underset{\text{new } x}{x}, \underline{w})$

POLYNOMIAL MODEL (CURVE FITTING)

$$t = w_0 \cdot 1 + w_1 \cdot x + \dots + w_M x^M = \sum_{j=0}^{M} x^j w_j = (1\ x\ x^2 \cdots x^M)\begin{pmatrix} w_0 \\ w_1 \\ \vdots \\ w_M \end{pmatrix}$$

$\{\phi_j(x)\}$ X-BASED FEATURES

X-BASED **BASIS FUNCTIONS**

HERE: $\phi_j(x) = x^j$

$$= (\phi_0(x)\ \phi_1(x) \cdots \phi_M(x))\begin{pmatrix} w_0 \\ w_1 \\ \vdots \\ w_M \end{pmatrix} = \underline{\phi(x)}^T \cdot \underline{w}$$

CHOICE OF $\underline{w}$?

— IF DATA=TRAINING DATA: t_n REQUIRED FOR x_n

THEN: $t_n \overset{!}{=} y(x_n, \underline{w}^*)$

$$\begin{bmatrix} t_1 = \\ t_2 = \\ \vdots \\ t_N = \end{bmatrix} \cdots \begin{bmatrix} \underline{\phi}^T(x_1) \\ \underline{\phi}^T(x_2) \\ \vdots \\ \underline{\phi}^T(x_N) \end{bmatrix} \underline{w}$$

$$\underline{t} = \underline{\underline{\Phi}} \cdot \underline{w}$$

IF $|\underline{\underline{\Phi}}| \neq 0 \rightarrow \boxed{\underline{w}^* = \underline{\underline{\Phi}}^{-1} \cdot \underline{t}}$

ANALYTIC SOLUTION

$\downarrow$

IF NOT

IF $|\underline{\underline{\Phi}}^T \underline{\underline{\Phi}}| \neq 0 \rightarrow \boxed{\underline{w}^* = \underbrace{(\underline{\underline{\Phi}}^T \underline{\underline{\Phi}})^{-1} \underline{\underline{\Phi}}^T}_{\text{MP-PSEUDO INVERSE}} \underline{t}}$

LEAST-SQUARES (LS) APPROACH (BISHOP 3.1)

ASSUME $y(x_n, \underline{w}) = \hat{t}_n = t_n + e_n$

WHAT CHOICE OF $\underline{w}$ MINIMIZES ERROR OVER ALL DATA?

WHAT IS THE ERROR? $E(\underline{w}) = \frac{1}{2} \sum_{n=1}^{N} e_n^2$ WHY THIS?