Linear Regression

26/08/2024

$x_1 \ldots\ldots x_D$	t (target variable)	$y = f_w(x)$ (prediction)
$[x_{11}, \ldots\ldots x_{1D}]$	t_1	y_1
$[x_{21} \ldots\ldots x_{2D}]$	t_2	y_2
$[x.$	$\vdots$	$\vdots$
$[x_{n1} \ldots\ldots x_{nD}]$	t_n	y_n

$$y_i = \left(\sum_{j=1}^{D} W_j x_{ij}\right) + W_0$$

$x \rightarrow \phi \rightarrow y = W$

$x \rightarrow \phi \; ; \; y \rightarrow W^T \phi$

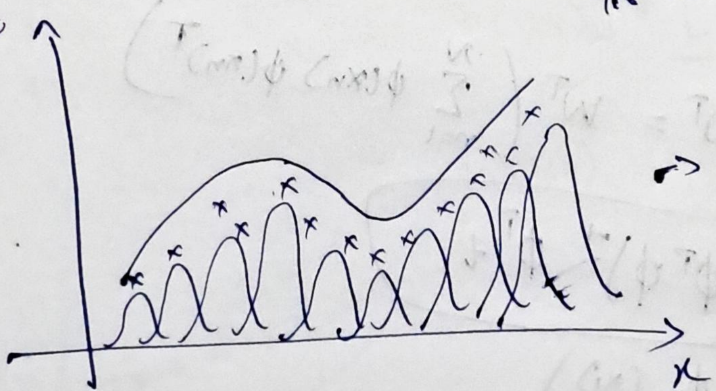
Ex: $x \in \mathbb{R}^1$

$$\phi_i = \begin{bmatrix} 1 \\ x_i \\ x_i^2 \\ \vdots \\ x_i^D \end{bmatrix}$$

$$y_i = \sum_{j=0}^{D} W_D x_i^j = W^T \phi$$

↳ features ϕ

Here features of x is non linear in $\mathbb{R}^1$ but linear in ϕ space



$$\phi = \begin{bmatrix} \vdots \\ e^{-(x-x_i)^2/2\sigma^2} \\ \vdots \end{bmatrix}$$

Radial Basis function

$x^2 = x x^T$

On ~~linear~~ w.

Data matrix

- Measured output has stochastic noise:-
  - Assuming ~~a~~ Gaussian noise: $\frac{1}{\sqrt{2\sigma^2\pi}} e^{-(x-\mu)^2/2\sigma^2}$

$$t = y(x, w) + \varepsilon$$

$$P(t \mid x, w, \beta) = \prod_{n=1}^{N} \mathcal{N}(t_n \mid w^T \phi(x_n), \beta^{-1})$$

$$\ln(P(t \mid x, w, \beta)) = \sum_{n=1}^{N} \ln\left(\frac{1}{\sqrt{2\sigma^2\pi}}\right) + \frac{[t - w^T\phi(x_i)]^2}{-2\sigma^2}$$

{ mean squared error comes from this equation of Expectation maximization due to the Gaussian distribution of errors }

→ Laplace distribution of error results in mean absolute error ~~difference~~

→ ~~outlier~~ more outliers indicates the distribution of errors to be Laplace

Maximizing likelihood:-

$$\nabla \ln(p(t \mid w, \beta)) = \sum_{n=1}^{N} \{t_n - w^T\phi(x_n)\}\,\phi(x_n)^T = 0$$

$$\Rightarrow \sum_{n=1}^{T} t_n \phi(x_n)^T = w^T \left(\sum_{n=1}^{N} \phi(x_n)\,\phi(x_n)^T\right)$$

$$\boxed{w_{ML} = (\phi^T\phi)^{-1}\,\phi^T t}$$

$$\phi = \begin{pmatrix} \phi_0(x_1) & \cdots & \phi_{M-1}(x_1) \\ \vdots & & \vdots \\ \phi_0(x_N) & & \phi_{M-1}(x_N) \end{pmatrix}$$

Bias - variance - Decomposition of linear reg with gaussian err

L = Loss = $(y(x_i) - t_i)^2$

$E(L) = \int\int (y(x_i) - t_i)^2 \, p(x, t) \, dx \, dt$ → minimize this loss

$$\frac{\partial E(L)}{\partial y(x)} = 2\int [y(x) - t] \, p(x,t) \, dt = 0$$

$$\Rightarrow y(x) = \frac{\int t \, p(x,t) \, dt}{p(x)}$$

{assuming a constant value of x}

$$= \int t \, (p(t/x)) \, dt$$

$$\boxed{y(x) = E_t(t/x)}$$

$$(y(x) - t)^2 = [\, y(x) - E[t/x] + E[t/x] - t \,]^2$$

$$= [y(x) - E(t/x)]^2 + [E[t/x] - t]^2$$

→ Non-Reducible error

$$+ \, 2[y(x) - E(t/x)][E(t/x) - t]$$ → 0

{Both error are ass to be uncorrelated}

Expected value is zero.