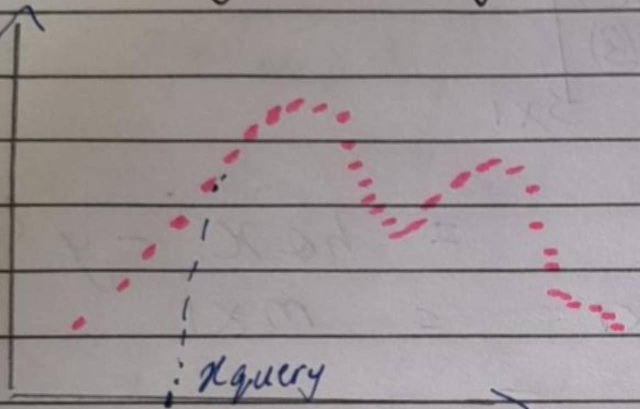
Locally Weighted Regression :-



This is obviously not a straight line, here we will have a query point xquery for which we will determine parameters using pts in its vicinity.

Neighbours of query point have more weight in determining parameters.

linear reg →

$h_\theta(x) = \theta^T x$

loss = $\sum_{i=0}^{m} (y^i - h_\theta x^i)^2$

weighted loss

$$J(\theta) = \sum_{i=0}^{m} \underbrace{w^i}_{\text{weight term}} (y^i - h_\theta(x^i))^2$$

$$\underset{(0,1)}{w^i} = e^{-\left(\frac{x^i - x}{2\tau^2}\right)}$$

x^i = any pt

x = query point

τ = bandwidth parameter

dist ↑ τ ↓

~~Non~~

r every query pt we are gonna learn what's the right set of thetas

$$loss = \sum_{i=1}^{m} w^i [h_\theta x^i - y^i]^2$$

scalar value = w scales loss.
different for every i

lets see how we can scale a vector

$$\underset{3\times3}{\begin{bmatrix} K_1 & 0 & 0 \\ 0 & K_2 & 0 \\ 0 & 0 & K_3 \end{bmatrix}} \underset{3\times1}{\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}} = \underset{3\times1}{\begin{bmatrix} K_1 \\ K_2(2) \\ K_3(3) \end{bmatrix}} \rightarrow \text{scaled vector}$$

$$\underset{m\times n}{X} \quad \underset{n\times1}{\theta} \quad - \quad \underset{m\times1}{y} \quad = h_\theta x^i - y^i = m\times1$$

$$\underset{m\times m}{\begin{bmatrix} w_1 & 0 & 0 \\ 0 & w_2 & 0 \\ 0 & 0 & w_3 \end{bmatrix}} \underset{\substack{m\times1 \\ X\theta - y}}{\begin{bmatrix} y_{pred} \\ \vdots \end{bmatrix}} = \underset{m\times1}{\begin{bmatrix} w_1 y_{pred} \\ w_2 y_{pred} \\ \vdots \end{bmatrix}}$$

weight matrix

$$loss = \sum_{i=1}^{m} w^i [h_\theta(x^i) - y^i]^2 \qquad w = \begin{bmatrix} w_1 & 0 & 0 \\ 0 & w_2 & 0 \\ 0 & 0 & w_3 \\ & \vdots & \end{bmatrix}$$

$$= (X\theta - y)^T w (X\theta - y)$$

Find $\nabla_\theta J(\theta) = 0$

$$\frac{\partial}{\partial\theta} J(\theta) = 0 \quad \text{to minimize loss}$$

$$= (\theta^T X^T - y^T)(WX\theta - Wy)$$

$$= \frac{\partial}{\partial \theta}\left(\theta^T X^T W X \theta - \theta^T X^T W y - y^T W X \theta + \cancel{y^T W y}\right) = 0$$

$$= \frac{\partial}{\partial \theta}\left(\theta^T X^T W X \theta - 2\theta X^T W y\right)$$

$$= 2X^T W X \theta - 2X^T W y \quad = 0 \text{ for minima}$$

$$\theta = \frac{X^T W y}{X^T W X}$$

$$\boxed{\theta = (X^T W X)^{-1} X^T W y}$$

closed form solution for locally weighted regression

as $\tau \uparrow$ you move towards linear regression