Gradient Descent for Linear Regression

$$w_j^{t+1} = w_j^t - \alpha \frac{\partial J(w)}{\partial w_j}$$

$$\Downarrow$$

$$w_j^{t+1} = w_j^t - \alpha \frac{1}{n} \sum_{i=1}^{n} \left(f_w(x^{(i)}) - y^{(i)} \right) x_j^{(i)}$$

$j = 1, \dots, d$

$\tilde{y}^{(i)}$

$$J(w) = \frac{1}{2n} \sum_{i=1}^{n} \left[f_w(x^{(i)}) - y^{(i)} \right]^2$$

Least Mean Square Error

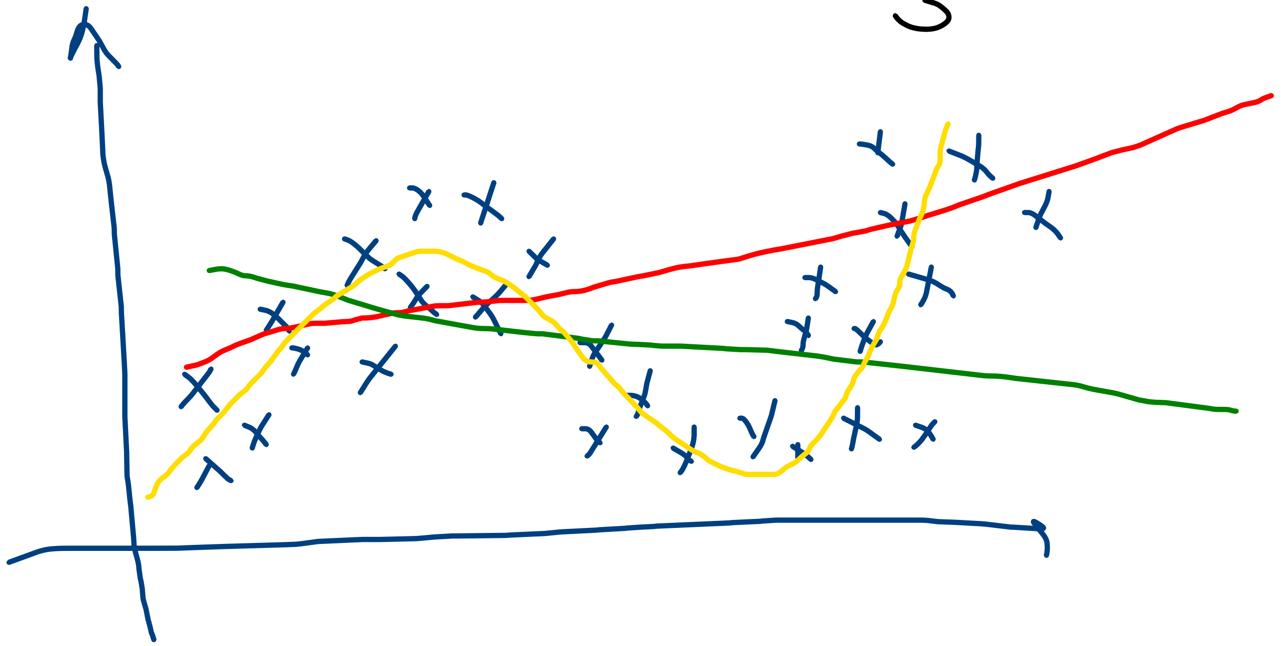
(LMSE)

① epoch → T and $\underbrace{\|w_{new} - w_{old}\|_2}_{\text{Convergence}} < \varepsilon$

LMSE

② $\underline{\underline{\alpha}}$ ⇒ Loss

3



non linear

<u>Linear</u> Reg.

$$f_w(x) = w_0' + w_1' x_1 + w_2' x_2 + \dots + w_d' x_d$$

خط ← 2D أصغر

hyper Plane

Linear basis function models

$$f_w(x) = \sum_{i=0}^{d} w_i x_i \quad , \quad x_0 = 1$$

$$f_w(x) = \sum_{i=0}^{d} w_i \, \phi_i(x)$$ basis function

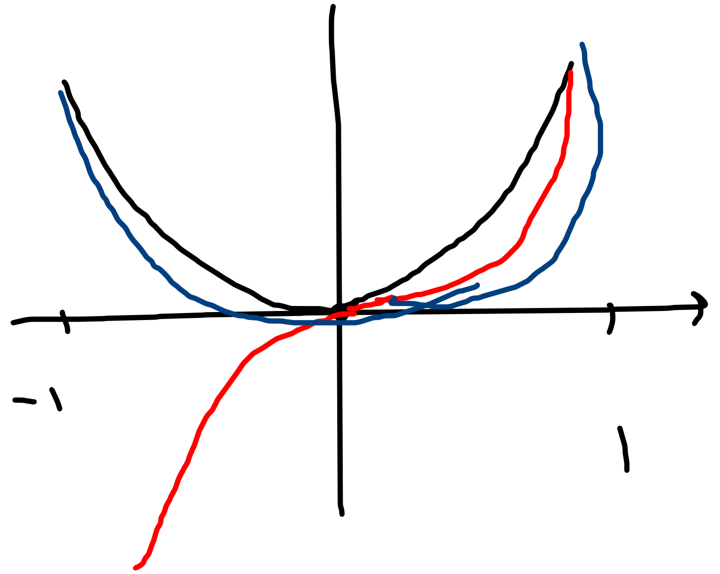
$\phi_0(x) = 1$
$\phi_i(x) = x_i$ $\Rightarrow$ Linear Regression

$i = 1, \dots, d$

1 - Polynomial Basis function

$$\phi_i(x) = x^i$$

$$f_w(x) = \sum_{i=0}^{d} w_i \, \phi_i(x)$$



$$X = \begin{array}{c} \begin{bmatrix} 1 & 5 & -1 \\ 2 & 3 & 0 \\ 3 & -1 & 4 \\ 4 & 2 & 1 \end{bmatrix} \begin{array}{c} x^{(1)} \\ x^{(2)} \\ x^{(3)} \\ x^{(4)} \end{array} \\ \begin{array}{ccc} x_1 & x_2 & x_3 \end{array} \end{array}$$

$$f_w(X) = w_0 + w_1 x_1 + w_2 x_2 + w_3 x_3$$

$$\phi_i(X) = x_i$$

$$X$$

$$\phi_2(\cancel{X}) = X^2$$

$$f_w(x) = \underline{\underline{w_0'}} + \underline{\underline{w_1'}} X + \underline{\underline{w_2'}} X^2$$

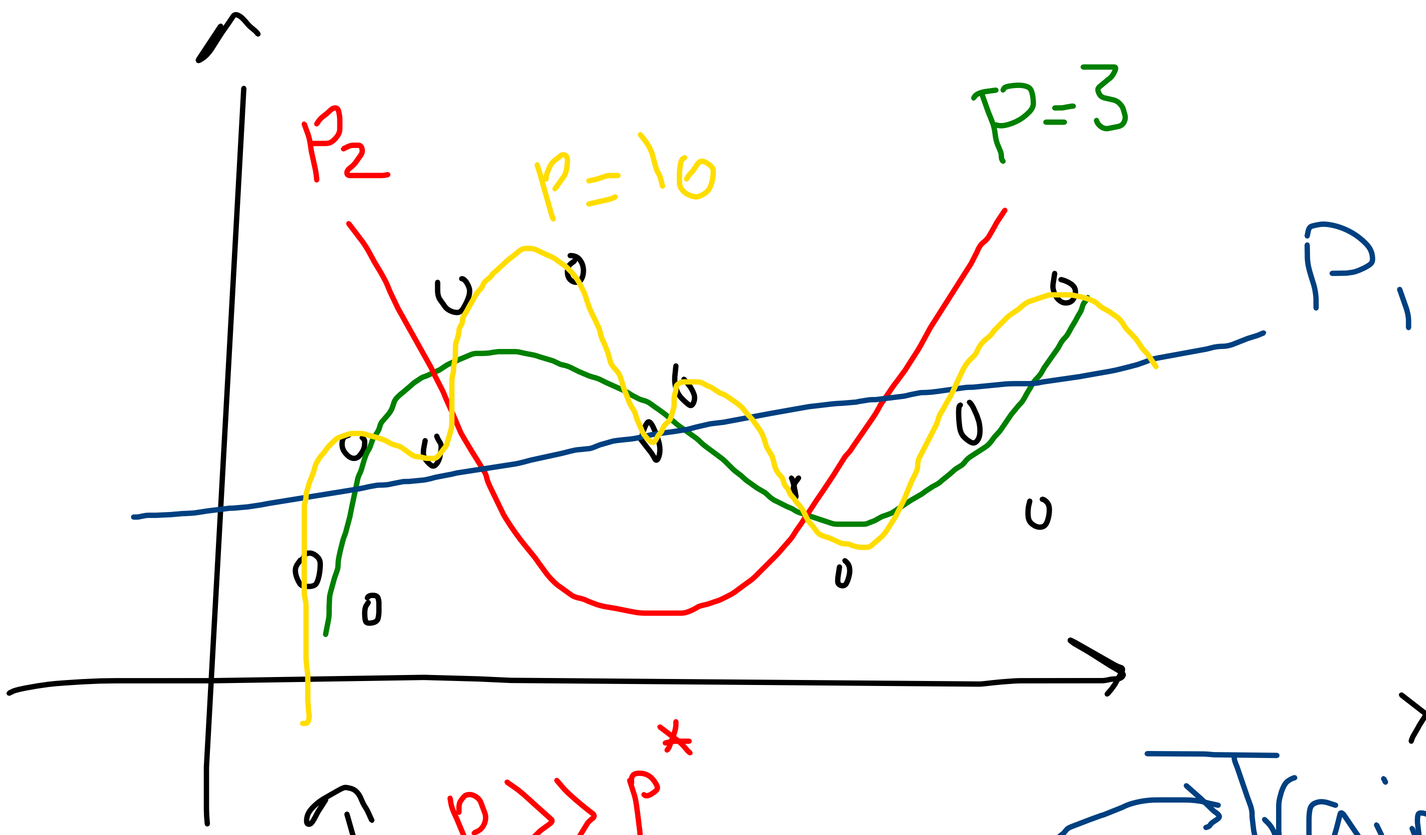
$$\boxed{f_w^{p}(X) = w_0 + w_1 X + w_2 X^2 + \cdots + w_p X^p = \sum_{j=0}^{p} w_j X^p}$$

$$f^{p}_{w}(x) = \sum_{j=0}^{p} w_j x^j$$

$P=2$ 0.001

$P=4$ 0.001

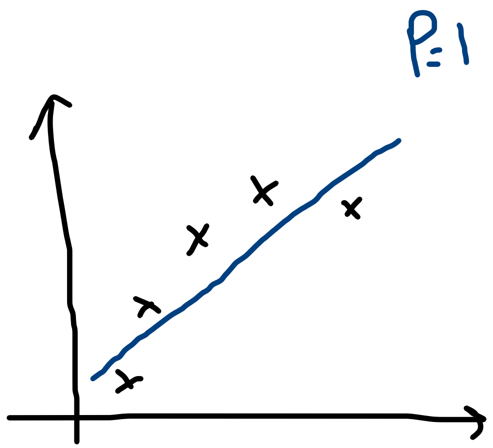
$p \ll p^*$



$p \gg p^*$

Train × آموزش

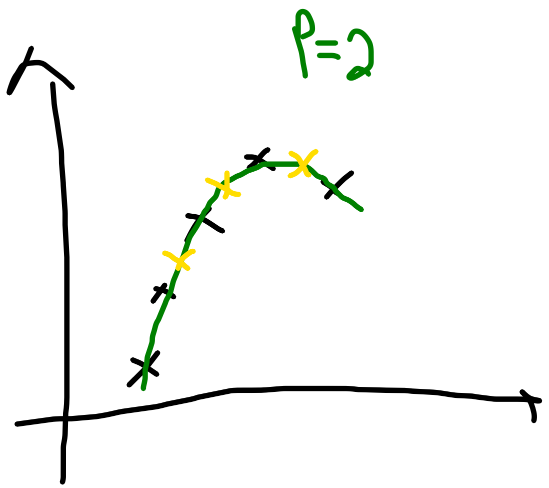
Test × آزمون



Underfitting

$w_0 + w_1 x$

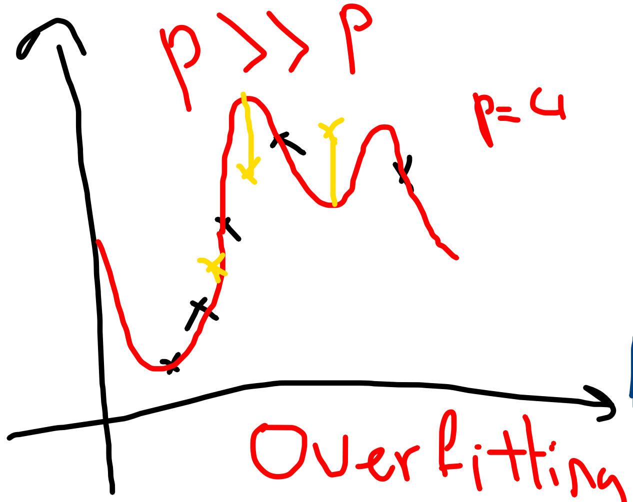
Error >>



Correct Fit

$w_0 + w_1 x + w_2 x^2$

Error ~ 0

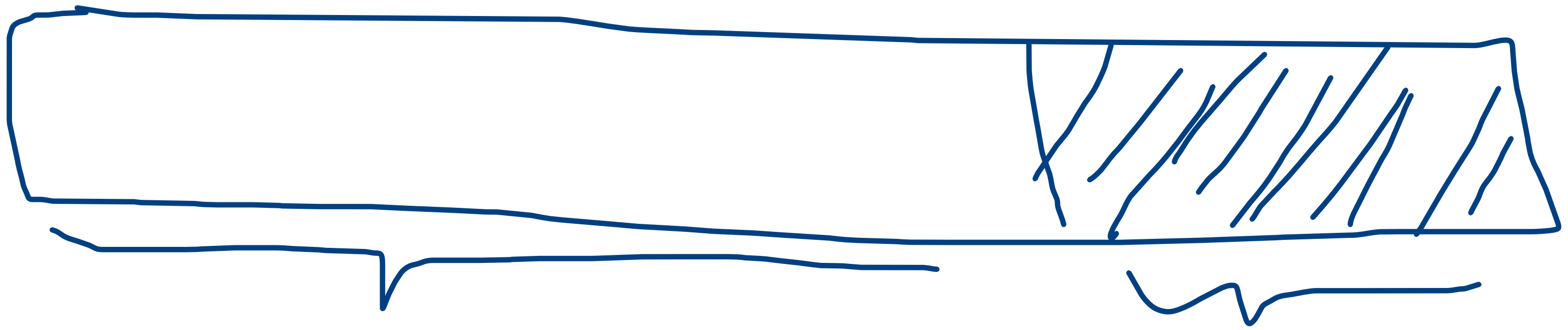


Overfitting

4

$w_0 + w_1 x + w_2 x^2 + w_3 x^3 + w_4 x^4$

Error ~ 0



آموزش ⇐ ساخت مدل

آزمون ⇐ ارزیابی مدل

10

0.01

Underfitting

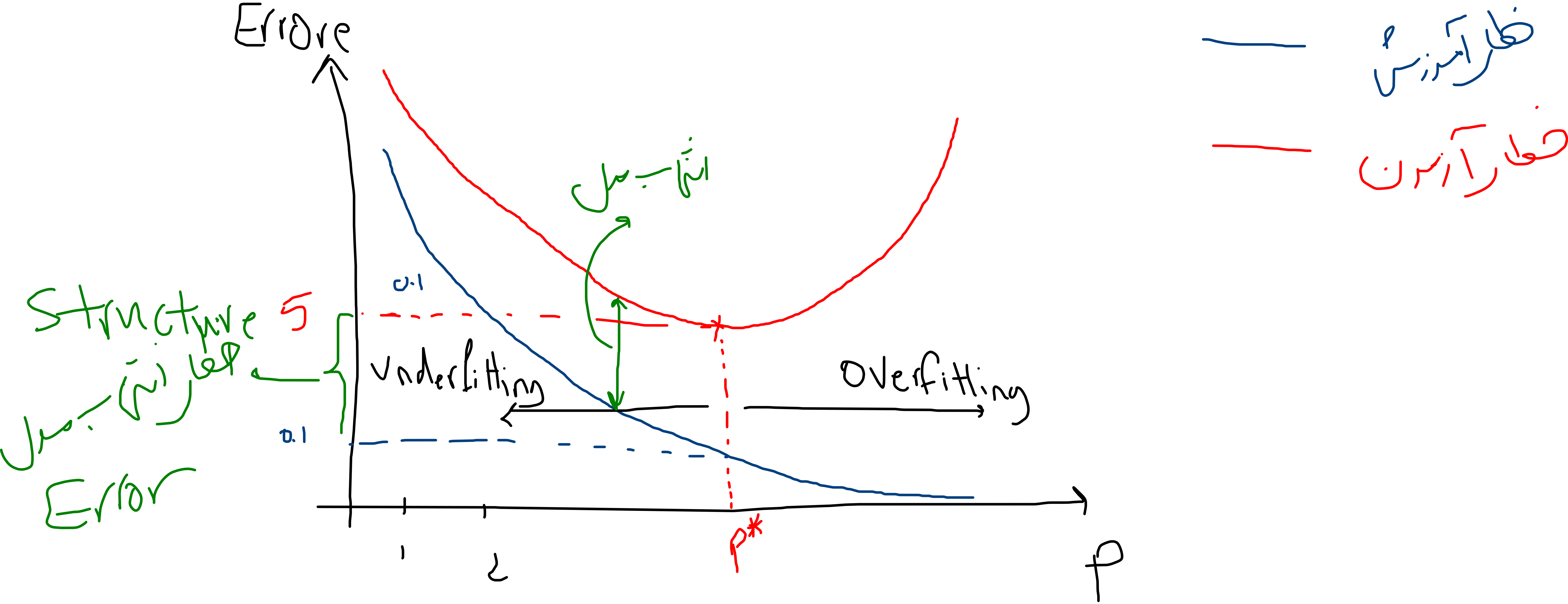
خطای آموزش بالا

Correct fit

خطای آموزش قابل قبول

Overfitting

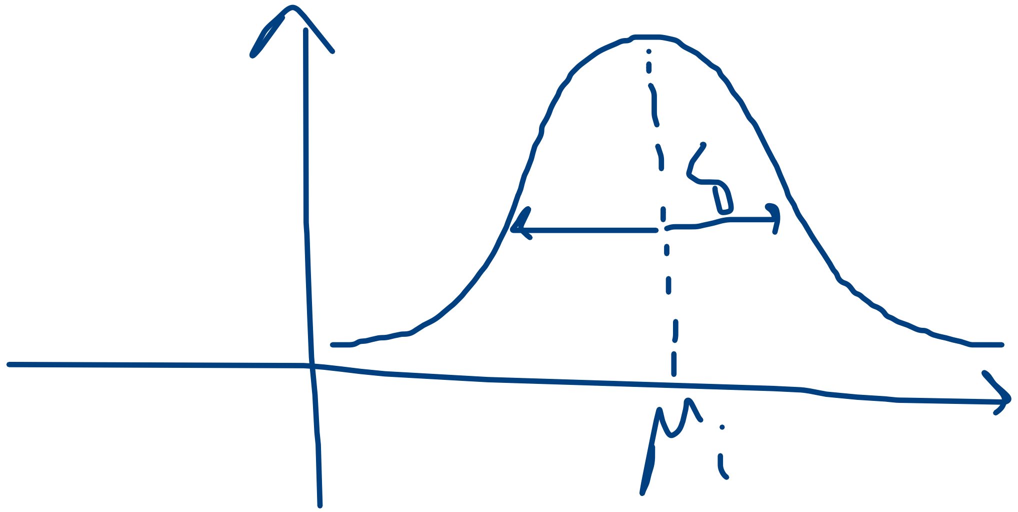
خطای آموزش قابل قبول


Errore
خطای آموزش
خطای آزمون
0.1
0.1
Structure
5
Underfitting
Overfitting
Error
1
2
p*
p

1, 2

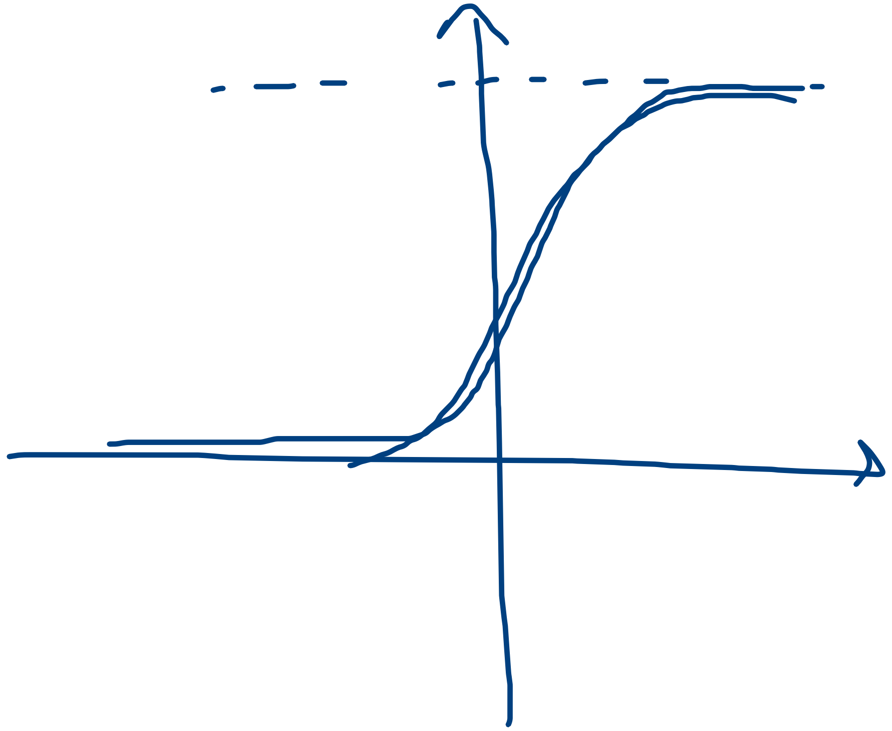
- Gaussian Basis function

$$\phi_j(x) = \exp\left(-\frac{(x-\mu_j)^2}{2s^2}\right)$$

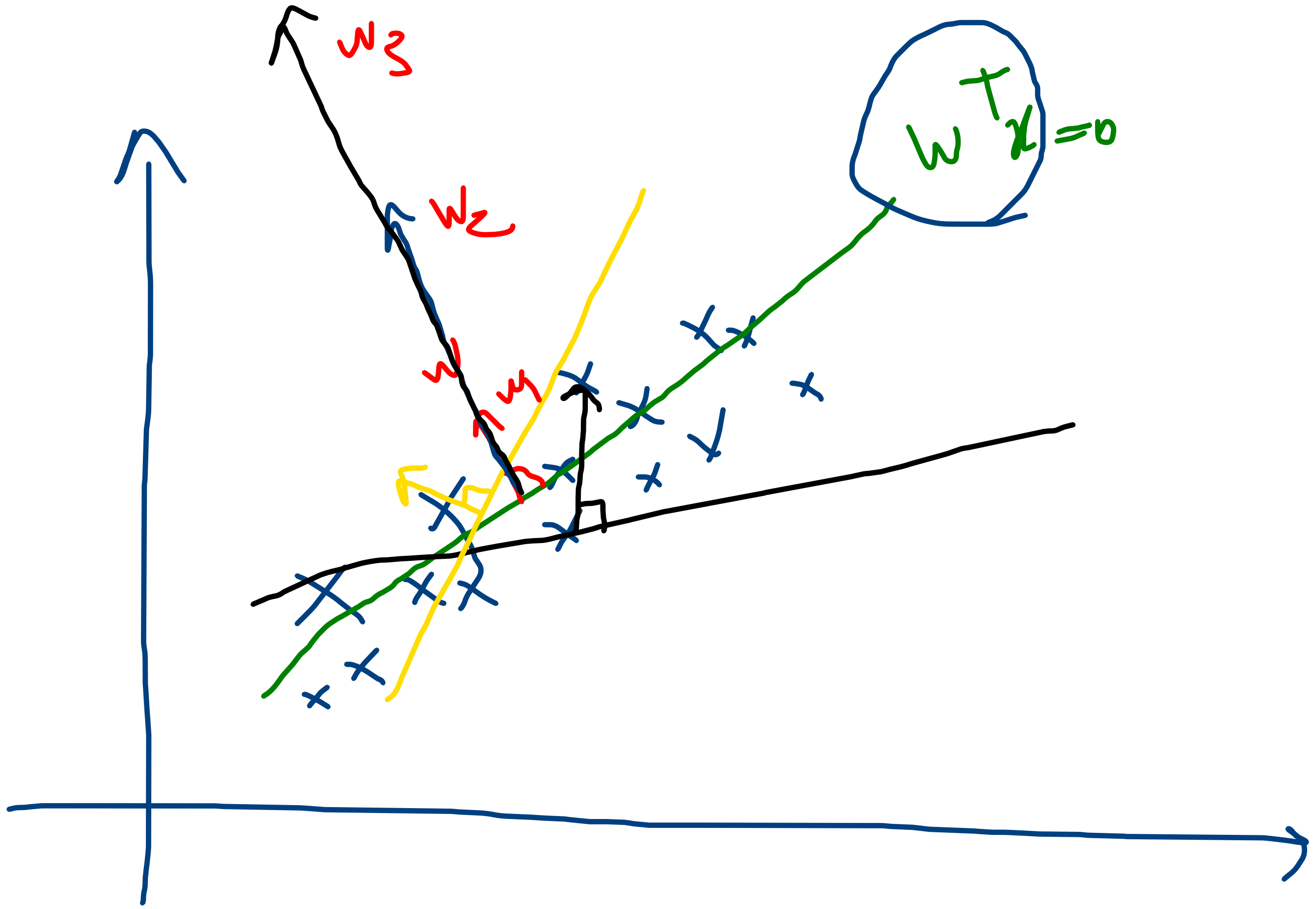


- Sigmoidal Basis function

$$\phi_j(x) = \sigma\left(\frac{x-\mu_j}{s}\right)$$



$$\sigma(x) = \frac{1}{1+e^{-x}}$$



w جهت

w اندازه

$$\|w_1\|_2 < \|w_2\|_2 < \|w_3\|_2$$

$$0 = 5 + 6x \qquad \begin{bmatrix} 5 \\ 6 \end{bmatrix}$$

$$0 = 30 + 36x \qquad \begin{bmatrix} 30 \\ 36 \end{bmatrix}$$

$$0 = 500 + 600x \qquad \begin{bmatrix} 500 \\ 600 \end{bmatrix}$$

$$\underbrace{J(w) = \frac{1}{2n}\sum_{i=1}^{n}\left(w^T x^{(i)} - y^{(i)}\right)^2}_{\text{خطای مدل}} + \underbrace{\lambda \|w\|_2}_{\text{پنالتی / جریمه}}$$

تابع هدف $\Rightarrow$ تابع هزینه

Loss function

پنالتی $\rightarrow$ طول

$$\min_{w^*} J(w)$$

$\lambda \geq 0 \Rightarrow w$ میزان تاثیر طول $\nearrow$ $\lambda \nearrow$

Regularized Linear Regression

$\rightarrow$ Regularization