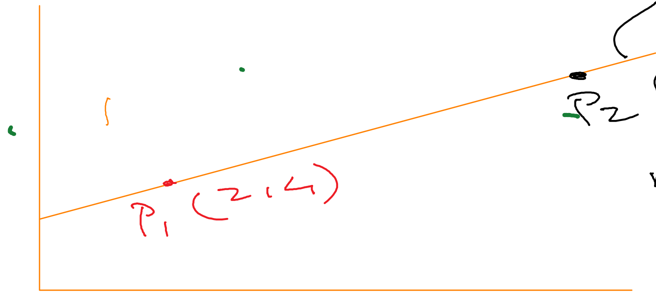
Regularization

Sunday, March 10, 2024 7:51 PM

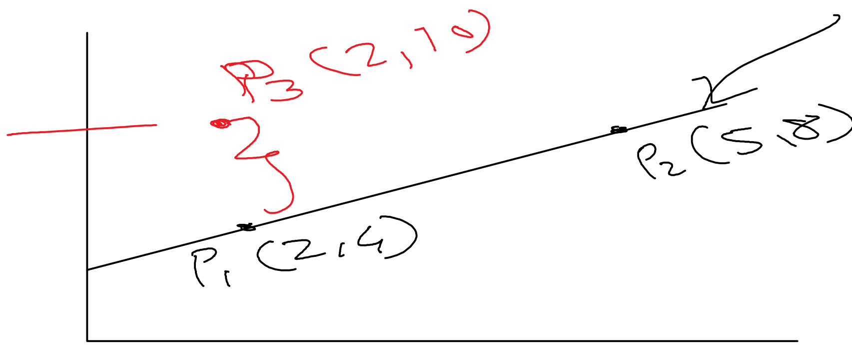
BFL MSE = 0



$m = \frac{8-4}{5-2} = \frac{4}{3} = 1.3$

MSE ≠ 0

MSE = 0



overfitting

Training Accuracy / MSE

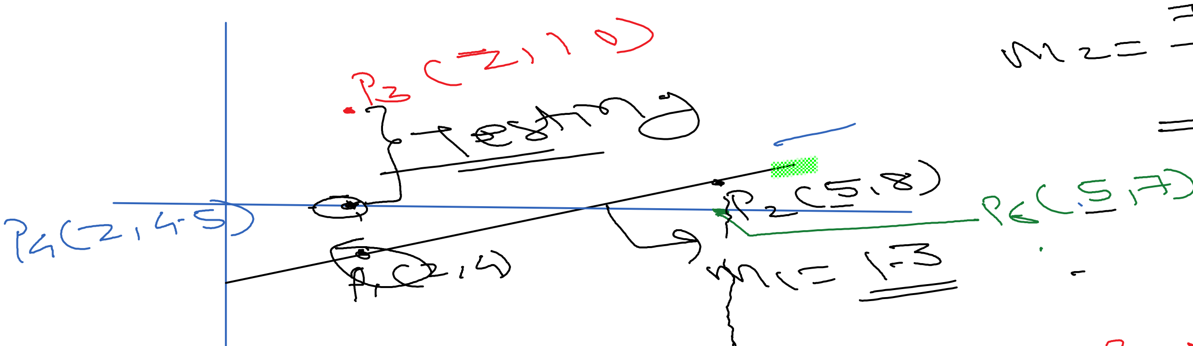
100% 0

Testing Accuracy ⇒ 80%

MSE ≠ 0

Low Bias & High variance

Regularization



$m_2 = \frac{7-4.5}{5-2} = 0.8$

$m_1 = 1.3$

After Regulariz

conclusion

MSE[0] on Training ⇒ MSE[40] ie MSE ≠ 0

MSE[100] on testing ⇒ MSE[60]

After Regulariz

Types
1. L_1 Lasso
2. L_2 Ridge

Ridge Regression

$CF = SSE + \lambda \times slope^2$

$= SSE + \lambda \times [m_1^2 + m_2^2 + m_3^2 + m_4^2 + \dots]$

Case-I

$S.S.E = 0$

λ – Regularization parameter

$\boxed{\lambda = 1}$ Default

$\boxed{m_1 = 1.3}$

$CF = 0 + 1 \times (1.3)^2$

$\boxed{CF = 1.69}$ ⇒ Training

Case-II After Regularization

$m_2 = 0.8$, $SSE \neq 0$, $SSE = 0.5$

$CF = 0.5 + 1 \cdot (0.8)^2$

$\boxed{CF = 1.14}$

$\lambda (slope)^2$

C.F = ...

$$CF = SSE + \alpha \cdot (slope)^2$$

$\alpha = 1$ default
(Regularization parameter)

Range $\alpha = 0.01$ to ∞

Lasso (L1)

$$C.F = SSE + \alpha \cdot |slope|$$

$$= SSE + \alpha\left[|m_1| + |m_2| + |m_3| + --\right]$$

(1) use only in linear model

(2) Lasso ⇒ only 5 feature
incase small set feature

Ridge ⇒ Incase of large set feature

(3) feature selection ⇒ Lasso is Best.

$$y = mx + c$$

$$\rightarrow y = m_1x_1 + m_2x_2 + \cdots + c$$

$(m_1 = 0)$

$$C.F = SSE + \alpha \times \left[|0.4| + |0| + |1|\right]$$

$\hookrightarrow m_2$

C-F= 58E + C1xL

(10800)

$\rightarrow$ m2