Polynomial Regression

04 February 2022 20:47

$Y \rightarrow x^2 + x$ than $Y \rightarrow x$ S.L R $\Rightarrow$ $\boxed{Y = \theta_0 + \theta_1 \cdot x}$

x is not linearly related y

Polynomial Regression Input / [Independent]

$$Y = \theta_0 + \theta_1 x + \theta_2 x^2 \leftarrow (2) \quad x^2 \leftarrow x$$

Target (y) / dependent

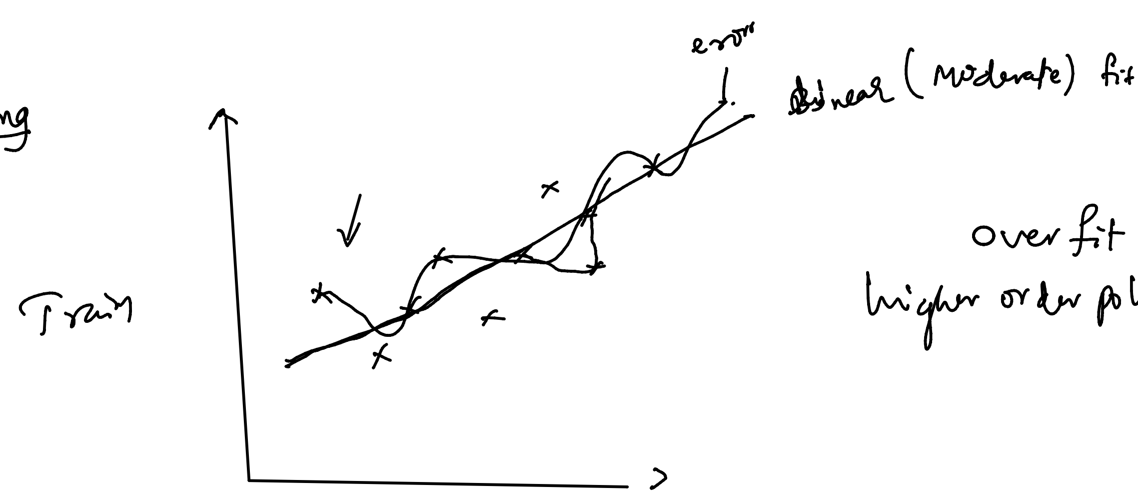
$$Y = \theta_0 + \theta_1 x + \theta_2 x^2 + \theta_3 x^3 \leftarrow (3)$$

$R^2 - 0.951$ order = 5 $\boxed{0.017}$

$R^2 - 0.968$ Order = 30 25 additional features

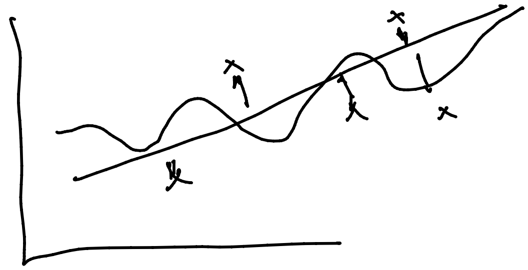
Generalization :-

overfitting



over fit
higher order polynomial

Test



expensive cars (Blue, Black)

Mid (Silver, white, x)

lower (Red, metiz grey)

Cars — Max power, Years, O.R, ... 'Color'

$$y = \theta_0 + \theta_1 \text{max power} + \theta_2 \cdots + \theta_k \text{Col r} + \theta_{k+1} \text{Color}$$

Multicollinearity.

$x_1 - \theta_1 \qquad x_1 \quad \frac{\partial J}{\partial \theta_1}$, x_2 is constant

$x_2 - \theta_2 \qquad x_2 \quad \frac{\partial J}{\partial \theta_2}$, x_1 const -

$\theta_1 \rightarrow Y$

$\theta_2 \rightarrow Y$

Bias vs Variance Tradeoff $\boxed{y = \theta_1 + \theta_2 x_1 + \theta_2 x_2}$ $\boxed{x_3, x_4 \rightarrow Y}$

Bias ← underfitted model

high Variance ← overfitted Model ← Variance high

		Test 1	Test 2	Test 3	
high variance	M_1 (overfitted)	→ [0.85	0.91	0.78]	high variance
high bias	M_2 (Underfitted).	[0.72	0.71	0.73]	

high bias M_2 (Underfitted). [0.72 0.71

0.86] ← Variance less

Moder bias, Varian M_3 (moderate fitted) [0.84 0.83

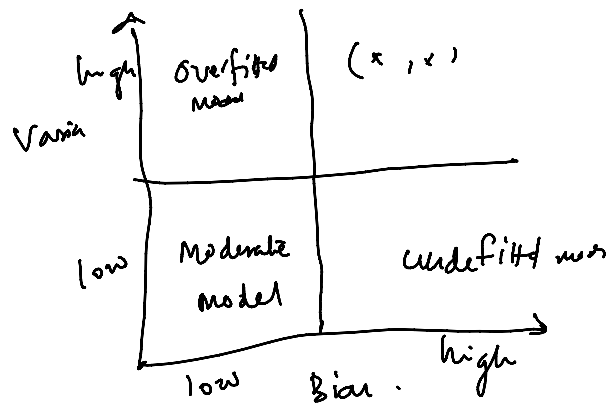
performance

underfitting (overfitted)

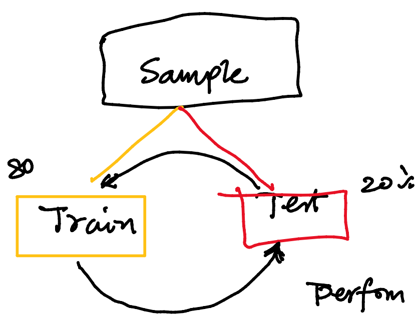
high bias, high variance x [ ]

Wrong sample undefitted Right sample (variance)

Low Bias, Low Variance

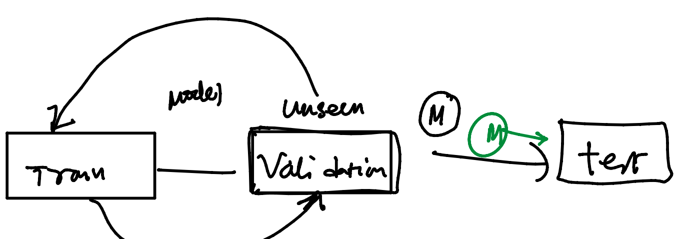


Validation



Perform well on the test dataset

$\boxed{\text{Train}}$ → Completed, Then only use $\boxed{\text{Test}}$



d=2, d=3, d=4, d=5

Sample

70% Train

10% Val

20% Test

Smaller (2000)

2000

Train 1600

400 Test ✓

Cross Validation : ✓ K-Fold , LOOC(X)

(K - No. of folds)

1600 Train

K = 4

400 F_1 | 400 F_2 | 400 F_3 | 400 F_4

	Train		Validation
(1200) M_1	F_1 F_2 F_3	R^2 0.81 ✓	F_4
M_2	F_2 F_3 F_4	R^2 0.82 ✓	F_1
M_3	F_3 F_4 F_1	R^2 0.79 ✓	F_2
M_4	F_4 F_1 F_2	R^2 0.78	F_3

(Variance) (Bias)

R^2 0.75 ✓

R^2 0.74

R^2 0.73

R^2 0.72

Variance

Variance less

Overfitting ✓

(Regularization) ✓

Penalising the coefficients.

$\boxed{\theta_2 = 2.3}$

over fitting ↑

θ - Model

2.3 ← 1.8

Over fitting $2.3 \leftarrow 1.8$ θ - Mode

Penalising the coefficients.

$$y = \theta_0 + \theta_1 x_1 + \underline{\theta_2} x_2 + \cdots \theta_m x_m$$

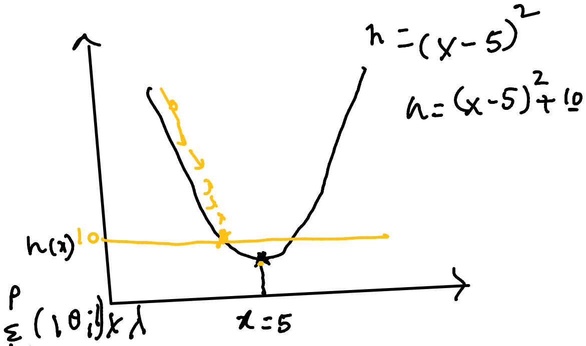
→ θ_2 is unnecessary in the model (x_2 has limited effect y)

→ Magnitude of a θ is higher than the required effect

$x \leftarrow \theta$

y

Shrinkage Parameter (λ)



LASSO Cost function)

$$\frac{1}{2m} \sum_{i=1}^{m} (h_\theta(x) - y)^2 + \frac{1}{2} \sum_{i=1}^{p} (|\theta_i|) \times \lambda$$

$+ \text{sum}(\theta)$

Unknowns, $\theta_0, \theta_1 \cdots \theta_p$

Ridge

$$\frac{1}{2m} \sum_{i=1}^{m} (h_\theta(x) - y)^2 + \frac{1}{2} \sum_{i=1}^{p} (|\theta_i|^2) \times \lambda$$

elastic Net

$$\frac{1}{2m} \sum_{i=1}^{m} (h_\theta(x) - y)^2 + \alpha \frac{1}{2} \sum_{i=1}^{p} (|\theta_i|) + \beta \frac{1}{2} \sum_{i=1}^{p} ((\theta_i|^2)$$

Splines :-