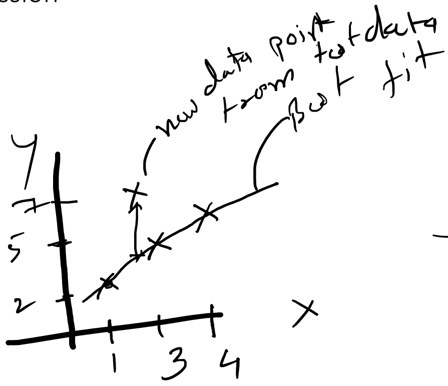
Ridge and Lasso Regression

27 June 2025 19:13




| x | y |
|---|---|
| 1 | 2 |
| 3 | 5 |
| 4 | 7 |

$(y-\hat{y})$

$(2-2)$
$+(5-5)$
$+(7-7)$
$= 0+0+0 = 0$

Acc = 100%.

## Overfitting

training acc = 95%
test acc = 70%

$\left.\begin{array}{l}\text{training acc} = 95\% \\ \text{test acc} = 70\%\end{array}\right\}$ $\boxed{\text{Low Bias} \\ \text{high Variance}}$

test set

| x | y | $\hat{y}$ |
|---|---|---|
| 5 | 6 | 7 |
| 6 | 7 | 9 |

$= (6-7)+(7-9)$
$= 1 + 2$
$= 3$

test Acc = 70%.

## Underfitting

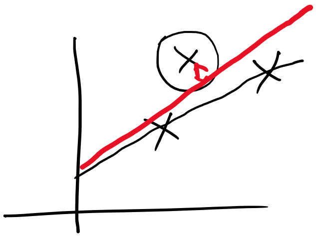
training acc = 60%
test acc = 61%

} high Bias
high Variance

## Generalize Model

training Acc = 90%.
test ACC = 89%.

} Low Bias
Low Variance



Training = 95%
test = 90%

Train = 100%
test = 78%

## Ridge Regression: (L2 Regularization)

to overcome the problem of overfitting

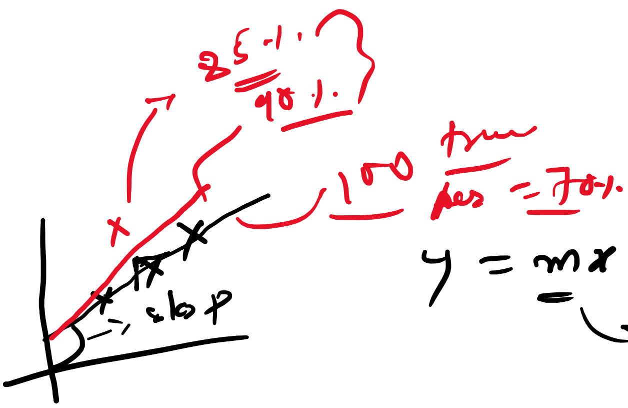
$y = mx + c$ | $h_\theta(x) = \theta_0 + \theta_1 x$

Cost function $J(\theta)$

$$= \frac{1}{m}\sum_{i=1}^{m}(y - \hat{y})^2$$

$$0$$

$$+ \lambda(\text{slope})^2$$




$y = mx + c$

$\to 0.1$ $\lambda = 1$

$$J(\theta) = \frac{1}{m}\sum_{i=1}^{m}(y - \hat{y}) + \underbrace{\lambda(m^2)}_{\text{penalty}}$$

$$0 + 1 \times (0.1)^2$$

$$0 + 0.01$$

$$= 0.01 = 20$$

# Lasso Regression ($L_1$ Regularization)

① to solve the overfitting
② feature selection

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 \; --- \; \beta_n x_n$$

0.0001

**Cost function**

$$J(\theta) = \frac{1}{m} \sum_{i=1}^{m} (y - \hat{y}) + \lambda \, |\text{slope}|$$

or

$$+ \lambda \, |m|$$

$L_1$ - Norms