Data Science & Artificial Intelligence

Machine Learning

Regression

DDP - 02

Discussion Notes

By- Astha Ma’am

[MCQ]

#Q. Consider the linear regression model $Y = X\beta + \varepsilon$ with $\varepsilon \sim N(0n, \sigma\varepsilon\hat{}2\ \ Inn)$. This model (without intercept) is fitted to data using the ridge regression estimator $\hat{\beta}(\lambda) = \arg\min\beta\ \|Y - X\beta\|\ 2\ 2 + \lambda\|\beta\|\ 2\ 2$ with $\lambda > 0$.

The data are:

$$X\hat{}\top = (-1\ 1\ 1\ -1) \text{ and } Y\hat{}\top = (-1.5\ 2.9\ -3.5\ 0.7\ )$$

What is the maximum likelihood/ordinary least squares estimator of the regression parameter for $\lambda = 0$?

A [-0.3, 0.05]

B [-0.5, 0.1]

C [0.1, -0.2]

D [0.05, -0.3]

$\beta \propto (x\hat{}T\ X)\hat{}-1\ \ X\hat{}T$

$x\hat{}T X = \begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix}$

$(x\hat{}T X)^{-1} = \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix}$

[MCQ]

#Q. Suppose you are training a Ridge Regression model for a particular task and notice the following training error and validation RSS

Train: 57

Validation: 32,714

Would your next to try a Ridge model with a larger or smaller λ

A Larger

B Smaller

C λ does not have an effect here

D Neither larger nor smaller

approval mistake

[MCQ]

#Q. How does ridge regression help in dealing with overfitting in a dataset with a large number of predictors?

A. By increasing the number of observations

B. By introducing a penalty term that shrinks the coefficients towards zero

C. By removing outliers from the dataset

D. By reducing the number of predictors

Penalizes the size of coefficient

[MCQ]

#Q. What role does the Regularization parameter (λ) play in controlling the bias-variance trade-off in a ridge regression model?

A. λ controls the number of predictors in the model

B. λ controls the degree of multicollinearity among predictors

C. λ balances the trade-off between bias and variance

D. λ has no effect on the model's performance

In edge regressor

λ

[MCQ]

#Q. What are the ridge regression coefficients for a dataset with predictors X1, X2, X3, and response variable Y, using a regularization parameter (λ) value of 0.5?

A It depends on the number of observations

B They are calculated using the formula: β^ridge = (X^T X + λ I)^-1 X^T Y

C Ridge regression does not provide coefficients

D They are the same as OLS regression coefficients

[MCQ]

#Q. Increasing the regularizing coefficient value for a ridge regressor will

i. Increase or maintain model bias.
ii. Decrease model bias.
iii. Increase or maintain model variance.
iv. Decrease model variance

A. i & ii

B. i & iv

C. ii & iii

D. ii & iv

* overfitting –

[MCQ]

#Q. Using the data X=[-3,5,4] and Y =[-10,20,20], assuming a ridge penalty λ = 50, what ratio versus the Maximum Likelihood Estimate (MLE) estimate wmle do you think the ridge regression L2 estimate estimate wridge estimate will be?

A. 2

B. 1

C. 0.6

D. 0.5

$$x^T x = \begin{bmatrix} (-3)^2 & -3\times 5 & -3\times 4 \\ -3\times 5 & 5^2 & 5\times 4 \\ -3\times 4 & 5\times 4 & 4^2 \end{bmatrix}$$

$$X^T y = \begin{bmatrix} -3 \\ 5 \\ 4 \end{bmatrix}\begin{bmatrix} -10 & 20 & 20 \end{bmatrix} \begin{bmatrix} 30 \\ 80 \\ 80 \end{bmatrix}$$

$$(X^T X + \lambda I)^{-1} X^T Y$$

$$= 0.5$$

[MCQ]

#Q. As the regularization parameter increases in Ridge regression, do the regression coefficients decrease?

A True

B False

λ increase loss function

[MCQ]

#Q. Which of the following statements are true?

Statement 1: Modifying the cost function can be done by incorporating a penalty equal to the square of the coefficients' magnitudes.

Statement 2: Ridge and Lasso regression are among the basic methods used to mitigate model complexity and counter overfitting issues, which can arise in simple linear regression.

A Statement 1 is incorrect, and statement 2 is correct.

B Statement 1 is correct, and statement 2 is incorrect.

C Both statements 1 and 2 are correct.

D Both statements 1 and 2 are incorrect.

THANK - YOU