

Non-Assessed Exercise

UNIVERSITY OF MANCHESTER
DEPARTMENT OF COMPUTER SCIENCE

DATA70121: Machine Learning and Statistics I

Lecture 10: Regularised Linear Model

Lecture 10: Regularised Linear Model

Multiple Choice Questions

1. For any given training dataset, the *ordinary least squares* (OLS) algorithm always produces a unique solution to linear regression. True or False?

A. True
B. False

☐

2. A smaller value of the tuning hyperparameter (λ) in LASSO or Ridge regression leads to more complex models. True or False?

A. True
B. False

☐

3. When there are more training examples than input features, the trained Ridge regression model always has a higher RSS than that of the trained OLS model on the training set. True or False?

A. True
B. False

☐

4. Both Ridge and LASSO regressions require feature scaling before model training. True or False?

A. True
B. False

☐

5. To apply a regularised linear model for regression, input features must be *normalised* to a range $[0, 1]$. True or False?

A. True
B. False

☐

6. The Ridge regression penalty term can be negative. True or False?

- A. True
- B. False

☐

7. What is the effect of the regularisation term in the loss function of a regularised linear regression model?

- A. It penalises a large number of features in the model.
- B. It penalises a small number of features in the model.
- C. It penalises large weights in the model.
- D. It penalises small weights in the model.

☐

8. Which form of regularisation is more likely to be useful when dealing with a dataset that has highly correlated features?

- A. L1 regularisation
- B. L2 regularisation
- C. Both L1 and L2 equally
- D. Neither L1 nor L2

☐

9. If the value of the regularisation parameter (λ) is set to zero, what happens to a regularised linear regression model?

- A. The model becomes fully regularised.
- B. The model becomes equivalent to OLS.
- C. The model's weights are all set to zero.
- D. The model's weights are all set to the biggest weight.

☐

10. In the context of LASSO and Ridge regression, what is the role of the regularisation parameter (λ)?

- A. It determines the learning rate of the optimisation algorithm.
- B. It determines the number of features included in the model.
- C. It determines the number of iterations for the optimisation algorithm.
- D. It determines the complexity of the model.

☐

11. Which of the following is NOT a characteristic of Ridge regression?

- A. Ridge regression shrinks coefficients towards zero.
- B. Ridge regression can handle multicollinearity.
- C. Ridge regression sets some coefficients exactly to zero.
- D. Ridge regression uses L2 penalty.

☐

12. Which type of regression model is known to perform feature selection?

- A. Linear regression
- B. Ridge regression
- C. LASSO regression
- D. Polynomial regression

☐

13. Why is LASSO regression sometimes preferred over Ridge regression?

- A. It can handle multicollinearity.
- B. It is less prone to overfitting.
- C. It is more robust to outliers.
- D. It can perform feature selection.

☐

14. In the geometric interpretation of LASSO and Ridge regression, what does the regularisation parameter (λ) represent?

- A. The radius of the constraint region
- B. The number of dimensions in the feature space
- C. The angle between the coefficients vector and the gradient vector
- D. The distance from the origin to the optimal coefficients vector

☐

15. Which of the following statements best describes the shape of the constraint region in Ridge regression?

- A. It is a hypercube.
- B. It is a hypersphere.
- C. It is a hyperplane.
- D. It is a hypercone.

☐

16. Which of the following statements best describes why LASSO regression can yield sparse solutions?

- A. Because the constraint region is a hypersphere.
- B. Because the constraint region is a hypercube.
- C. Because the constraint region intersects the contour of the loss function at an axis.
- D. Because the constraint region is a hyperplane.

☐

17. What can we infer about the coefficients in LASSO and Ridge regression from their geometric interpretation?

- A. In LASSO, some coefficients can become exactly zero.
- B. In Ridge regression, all coefficients are shrunk towards zero but do not become exactly zero.
- C. In Ridge regression, some coefficients can become exactly zero.
- D. In LASSO, all coefficients are shrunk towards zero but do not become exactly zero.

☐

18. Which of the following are reasons to use LASSO or Ridge regression instead of subset selection?

- A. They can perform feature selection and parameter estimation simultaneously because the constraint region is a hypercube.
- B. They are computationally less intensive.
- C. They can handle multicollinearity better.
- D. They can yield interpretable models.
- E. They can handle the case where the number of features is larger than the number of observations.

☐