INFO251 – Applied Machine Learning

Lab 6

Suraj R. Nair

Quiz 1 Review

Regression

- When comparing two regression models, the model that produces the higher R2 will provide less biased estimates of the causal impact of the independent variables on the dependent variable:
  - True
  - False

Regression

- When comparing two regression models, the model that produces the higher R2 will provide less biased estimates of the causal impact of the independent variables on the dependent variable:
  - True
  - False

Difference-in-difference

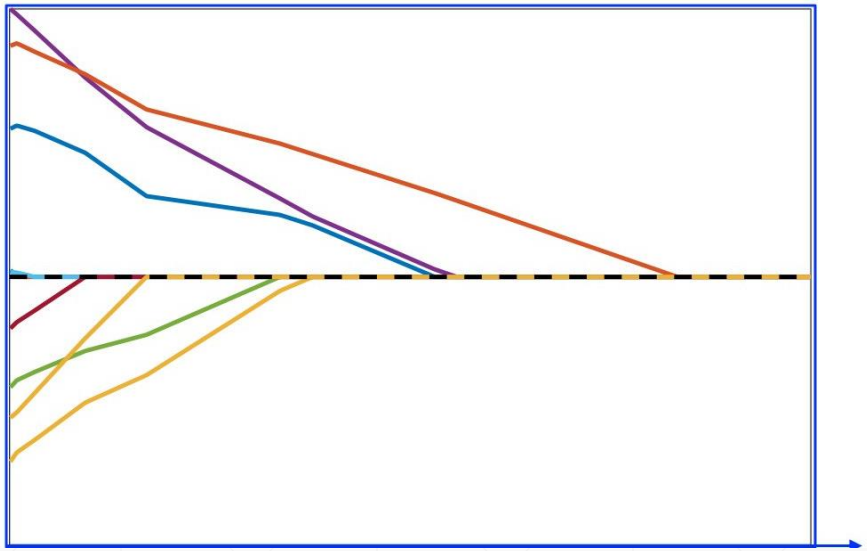
- The key identifying assumption is
  A. Outcomes in the control and treatment group would have been the same in the absence of treatment
  B. Trends in the control and treatment group would have been the same in the absence of treatment
  C. Outcomes pre- and post-treatment would have been the same in the absence of treatment
  D. Outcomes pre-treatment would have been the same in the absence of treatment

Difference-in-difference

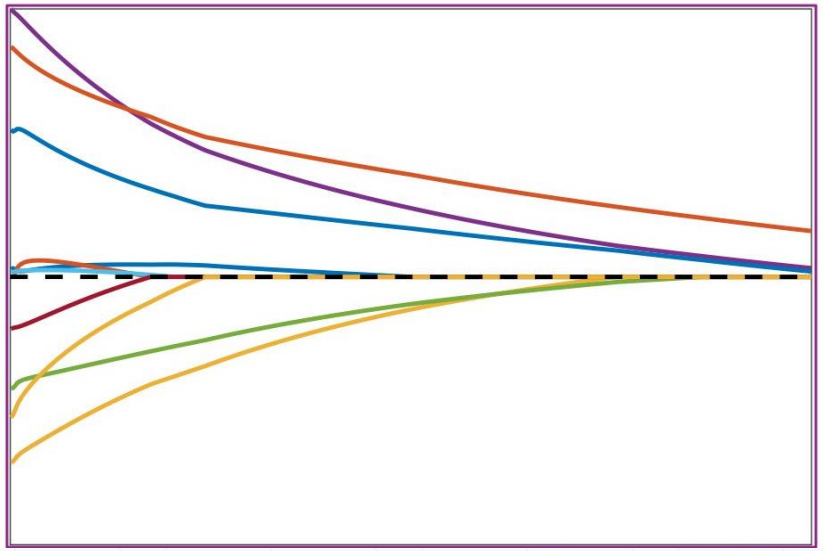
- The key identifying assumption is
  A. Outcomes in the control and treatment group would have been the same in the absence of treatment
  B. Trends in the control and treatment group would have been the same in the absence of treatment
  C. Outcomes pre- and post-treatment would have been the same in the absence of treatment
  D. Outcomes pre-treatment would have been the same in the absence of treatment

Regularization

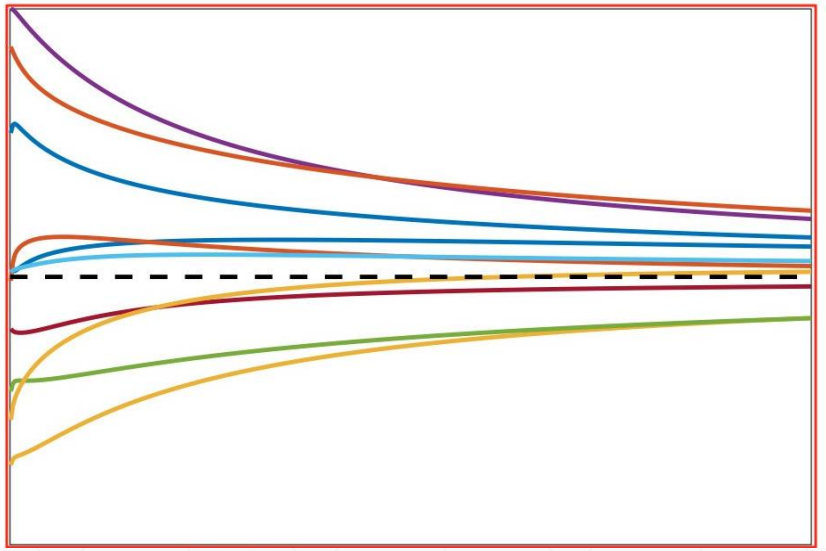
- Match the penalty (Lasso, Ridge, ElasticNet) to the coefficient plot



A



B



C

Regularization

- A – Lasso
- B – ElasticNet
- C – Ridge

Decision Boundaries

- Which of the following algorithms recovers non-linear decision boundaries:
  - K-nearest neighbors (K = 5)
  - SVM
  - Logistic Regression
  - Logistic Regression with lasso regularization

Decision Boundaries

- Which of the following algorithms recovers non-linear decision boundaries:
  - K-nearest neighbors (K = 5)
  - SVM
  - Logistic Regression
  - Logistic Regression with lasso regularization

Gradient Descent

- You are trying to find the parameters for a multivariate linear regression using gradient descent. The algorithm is initialized at some random starting point. However, it is taking very long to converge (> 10,000 iterations). What could be the reason(s)?
  - Step size is too small
  - Step size is too large
  - Data may not have been scaled
  - All of the above

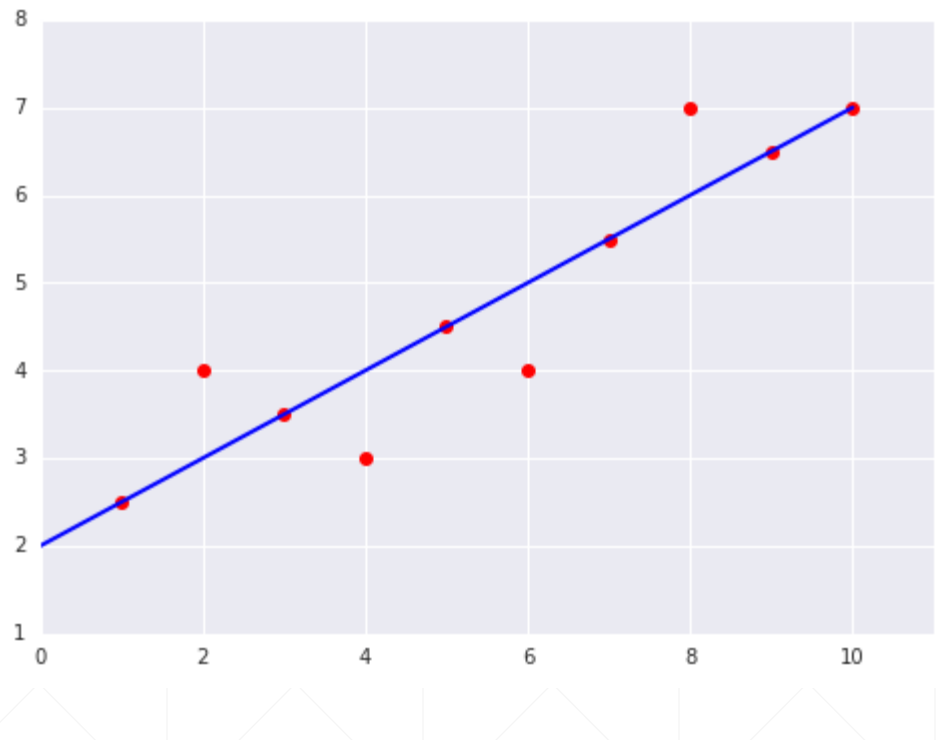
Gradient Descent

- You are trying to find the parameters for a multivariate linear regression using gradient descent. The algorithm is initialized at some random starting point. However, it is taking very long to converge (> 10,000 iterations). What could be the reason(s)?
  - Step size is too small
  - Step size is too large
  - Data may not have been scaled
  - All of the above

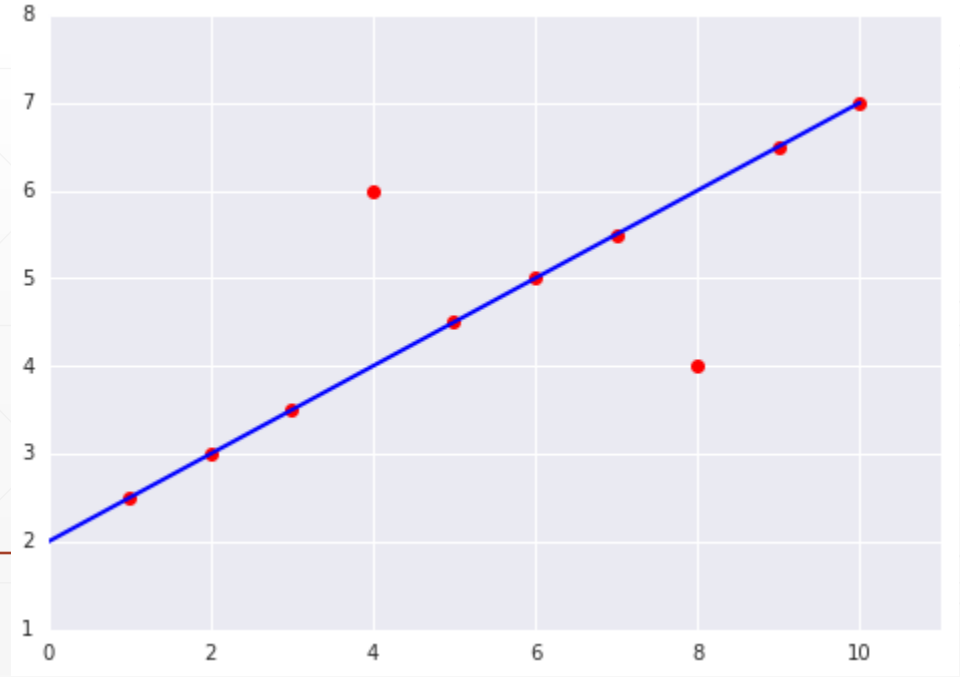
Mean Squared Error

- Suppose you build a linear regression model which predicts y = f(x). Which of these two cases has a higher MSE?
  - A
  - B

A



B

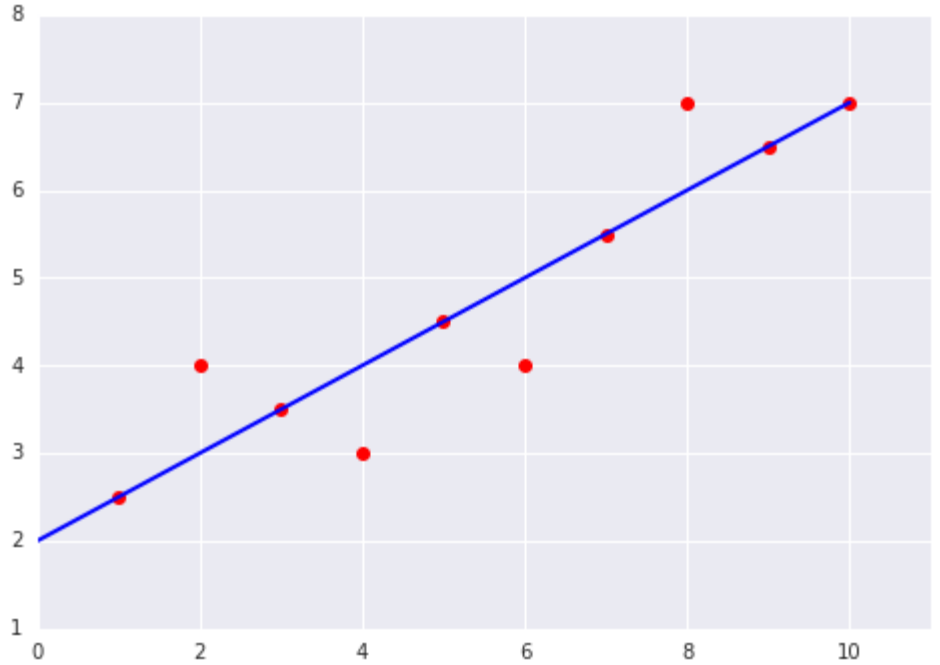


Mean Squared Error

- Suppose you build a linear regression model which predicts y = f(x). Which of these two cases has a higher MSE?
  - A
  - B

  - MSE for A: 4 / 10
  - MSE for B: 8 / 10

A



B