1.

Question 1

After training a ridge regression model, you find that the training and test set accuracies are 0.98 and 0.54 respectively. Which of the following would be the best choice for the next ridge regression model you train?

1 point

- You are overfitting, the next model trained should have a lower value for alpha
- (•) You are overfitting, the next model trained should have a higher value for alpha
- You are underfitting, the next model trained should have a lower value for alpha
- You are underfitting, the next model trained should have a higher value for alpha

2.

Question 2

After training a Radial Basis Function (RBF) kernel SVM, you decide to increase the influence of each training point and to simplify the decision surface. Which of the following would be the best choice for the next RBF SVM you train?

1 point

- (•) Decrease C and gamma
- Increase C and gamma
- Increase C, decrease gamma
- Decrease C, increase gamma

3.

Question 3

Which of the following is an example of multiclass classification? (Select all that apply)

1 point

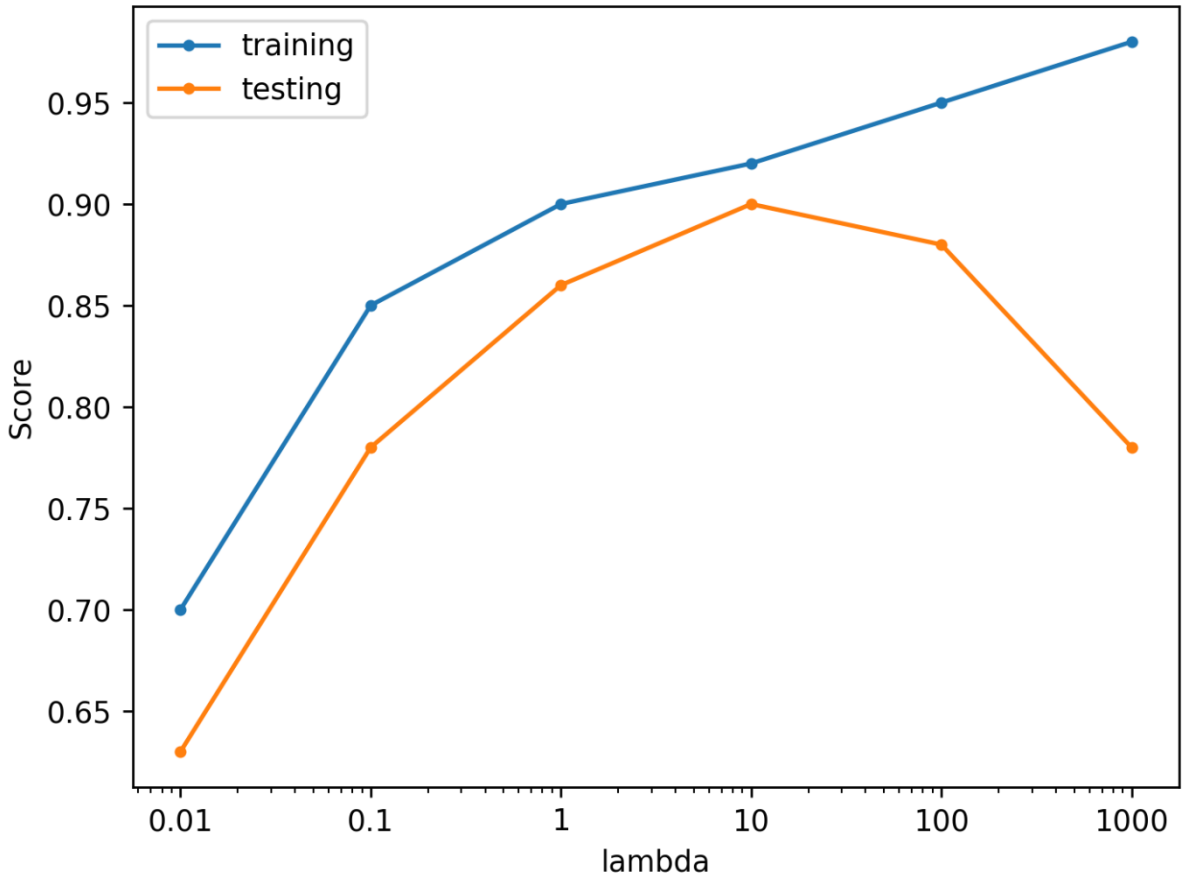
- [x] Classify a piece of fruit so that it is assigned to one of four classes: apples, oranges, bananas, or lemons
- [ ] Predict whether an article is relevant to one or more topics (e.g. sports, politics, finance, science)
- [ ] Predicting both the rating and profit of soon to be released movie
- [ ]

Classify a voice recording as an authorized user or not an authorized user.

4.

Question 4

Looking at the plot below which shows accuracy scores for different values of a regularization parameter lambda, what value of lambda is the best choice for generalization?



1 point

10

5.

Question 5

Suppose you are interested in finding a parsimonious model (the model that accomplishes the desired level of prediction with as few predictor variables as possible) to predict housing prices. Which of the following would be the best choice?

1 point

Logistic Regression

Ridge Regression

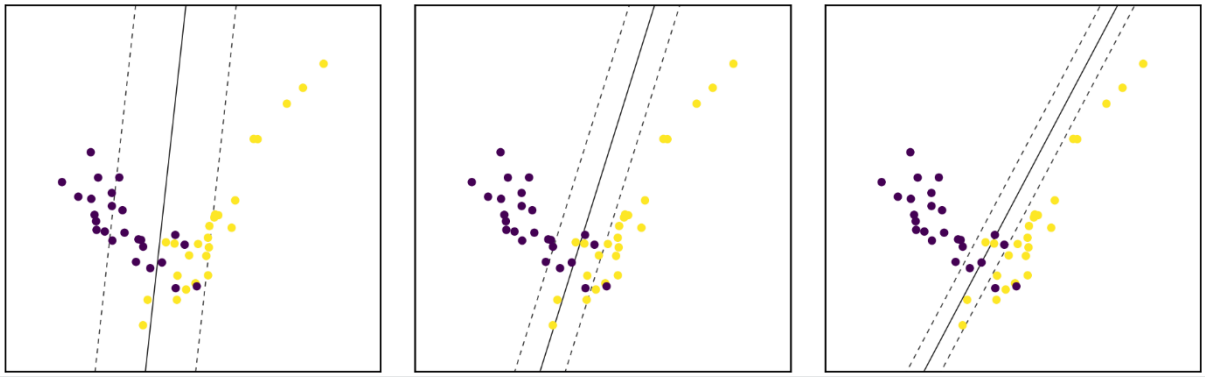
Ordinary Least Squares Regression

Lasso Regression

6.

Question 6

Match the plots of SVM margins below to the values of the C parameter that correspond to them.



1 point

10, 1, 0.1

0.1, 1, 10

10, 0.1, 1

1, 0.1, 10

7.

Question 7

Use Figures A and B below to answer questions 7, 8, 9, and 10.

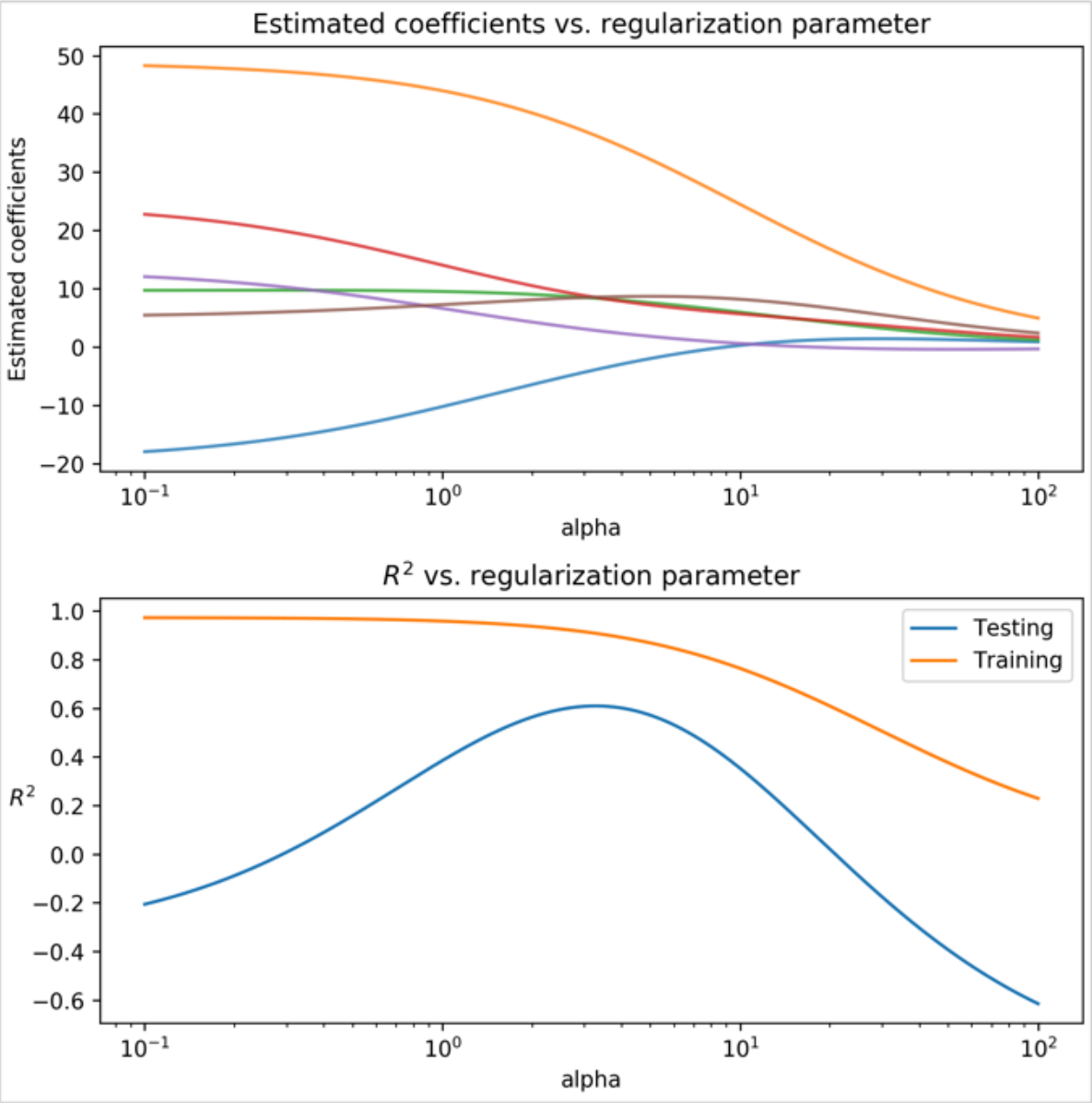

Estimated coefficients vs. regularization parameter
Estimated coefficients
50
40
30
20
10
0
−10
−20
10^{-1}
10^{0}
10^{1}
10^{2}
alpha
R^2 vs. regularization parameter
R^2
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
Testing
Training
10^{-1}
10^{0}
10^{1}
10^{2}
alpha

Figure A

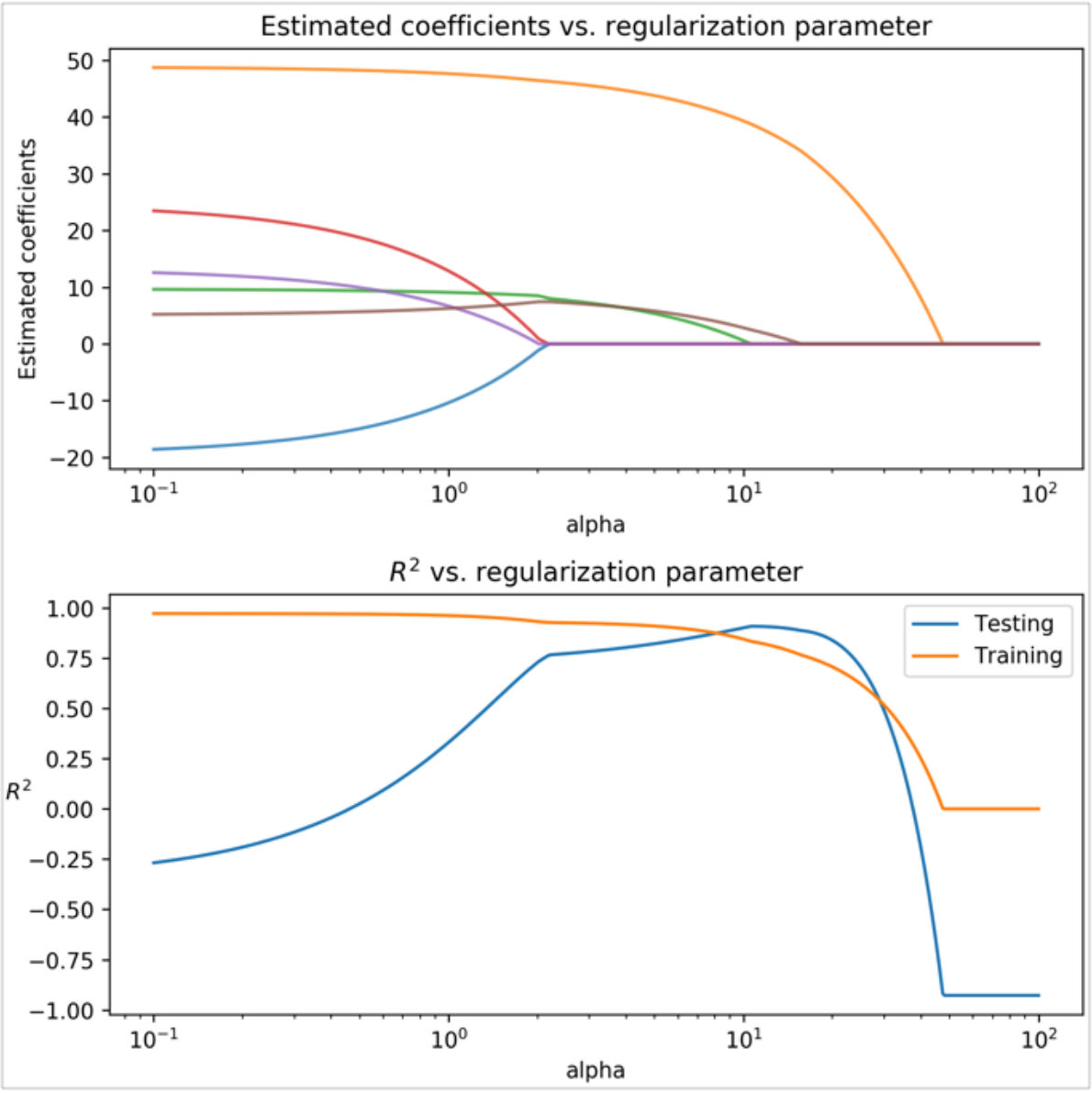


Figure B

Looking at the two figures (Figure A, Figure B), determine which linear model each figure corresponds to:

1 point

- (●) Figure A: Ridge Regression, Figure B: Lasso Regression
- ( ) Figure A: Lasso Regression, Figure B: Ridge Regression
- ( ) Figure A: Ordinary Least Squares Regression, Figure B: Ridge Regression
- ( ) Figure A: Ridge Regression, Figure B: Ordinary Least Squares Regression
- ( )

Figure A: Ordinary Least Squares Regression, Figure B: Lasso Regression

Figure A: Lasso Regression, Figure B: Ordinary Least Squares Regression

8.

Question 8

Looking at Figure A and B, what is a value of alpha that optimizes the R2 score for the Ridge Model? (Note the log scale for alpha.)

1 point

3

9.

Question 9

Looking at Figure A and B, what is a value of alpha that optimizes the R2 score for the Lasso Model? (Note the log scale for alpha.)

1 point

10

10.

Question 10

When running a LinearRegression() model with default parameters on the same data that generated Figures A and B the output coefficients are:

Coef 0	-19.5
Coef 1	48.8
Coef 2	9.7
Coef 3	24.6
Coef 4	13.2
Coef 5	5.1

For the Lasso regression model that maximizes R2 score on the test set, what will be the approximate value of Coeff 0?

1 point

0

11.

Question 11

Which of the following is true of cross-validation? (Select all that apply)

1 point

☑ Helps prevent knowledge about the test set from leaking into the model

☑ Increases generalization ability and computational complexity

- [x] Fits multiple models on different splits of the data
- [ ] Removes need for training and test sets
- [ ] Increases generalization ability and reduces computational complexity

12.

Question 12

Which of the following are true statements about validation curves? (Select all that apply)

1 point

- [x] Validation curves should show both the mean cross-validation accuracy, and the variation about the mean (typically as a shaded region above and below the curve).
- [x] Validation curves show how mean cross-validation accuracy varies a function of a hyperparameter, such as the SVM gamma parameter.
- [x] Validation curves should only be generated using a single train-test split.
- [x] Validation curves are used to compare the accuracy of a model across different cross-validation folds.