

1 point

1.

From the section "Effect of L1 penalty": Consider the simple model with 2 features trained on the entire sales dataset.

Which of the following values of $l1_penalty$ would not set $w[1]$ zero, but would set $w[2]$ to zero, if we were to take a coordinate gradient step in that coordinate? (Select all that apply)

☐

1.4e8

☒

1.64e8

☒

1.73e8

☐

1.9e8

☐

2.3e8

1 point

2.

Refer to the same model as the previous question.

Which of the following values of $l1_penalty$ would set both $w[1]$ and $w[2]$ to zero, if we were to take a coordinate gradient step in that coordinate? (Select all that apply)

☐

1.4e8

☐

1.64e8

☐

1.73e8

☒

1.9e8

☒

2.3e8

1 point

3.

From the section "Cyclical coordinate descent": Using the simple model (with 2 features), we run our implementation of LASSO coordinate descent on the normalized sales dataset. We apply an L1 penalty of $1e7$ and tolerance of 1.0.

Which of the following ranges contains the RSS of the learned model on the normalized dataset?

☐

Between $8e13$ and $2e14$

☐

Between $2e14$ and $5e14$

☐

Between $5e14$ and $8e14$

☐

Between $8e14$ and $1e15$

☒

Between $1e15$ and $3e15$

1 point

4.

Refer to the same model as the previous question.

Which of the following features were assigned a zero weight at convergence?

☐ constant

☐ sqft_living

☒

bedrooms

1 point

5.

In the section "Evaluating LASSO fit with more features", we split the data into training and test sets and learn weights with varying degree of L1 penalties. The model now has 13 features.

In the model trained with $l1_penalty=1e7$, which of the following features has non-zero weight? (Select all that apply)

☒ constant

☒ sqft_living

☐ grade

☒ waterfront

☐ sqft_basement

1 point

6.

This question refers to the same model as the previous question.

In the model trained with $l1_penalty=1e8$, which of the following features has non-zero weight? (Select all that apply)

☒ constant

☐ sqft_living

☐ grade

☐ waterfront

☐ sqft_basement

1 point

7.

This question refers to the same model as the previous question.

In the model trained with $l1_penalty=1e4$, which of the following features has non-zero weight? (Select all that apply)

☒ constant

☒ sqft_living

☒ grade

☒ waterfront

☒ sqft_basement

1 point

8.

In the section "Evaluating each of the learned models on the test data", we evaluate three models on the test data. The three models were trained with same set of features but different L1 penalties.

Which of the three models gives the lowest RSS on the TEST data?

☒

The model trained with $1e4$

☐

The model trained with $1e7$

☐

The model trained with $1e8$