1) Assume that for a certain linear regression problem involving 4 features, the following weight vectors produce an equal amount of mean square error:

w_1 = [2, 2, 3, 1]

w_2 = [1, 1, 3, 1]

w_3 = [3, 2, 4, 1]

w_4 = [1, 2, 1, 1]

Find $||W||$: $||W_1|| = 4+4+9+1 = 18$, $||W_2|| = 1+1+9+1 = 12$, $||W_3|| = 9+4+16+1 = 30$, $||W_4|| = 1+4+1+1 = 7$

Choose the lowest

Which of the weight vector is likely to be chosen by ridge regression?

- ○ w_1
- ○ w_2
- ○ w_3
- ● w_4

2) Assuming that in the constrained version of ridge regression optimization problem, following are the weight vectors to be considered, along with the mean squared error (MSE) produced by each:

w_1 = [2, 2, 3, 1], MSE = 3

w_2 = [1, 1, 3, 1], MSE = 5

w_3 = [3, 2, 4, 1], MSE = 8

w_4 = [1, 2, 1, 1], MSE = 9

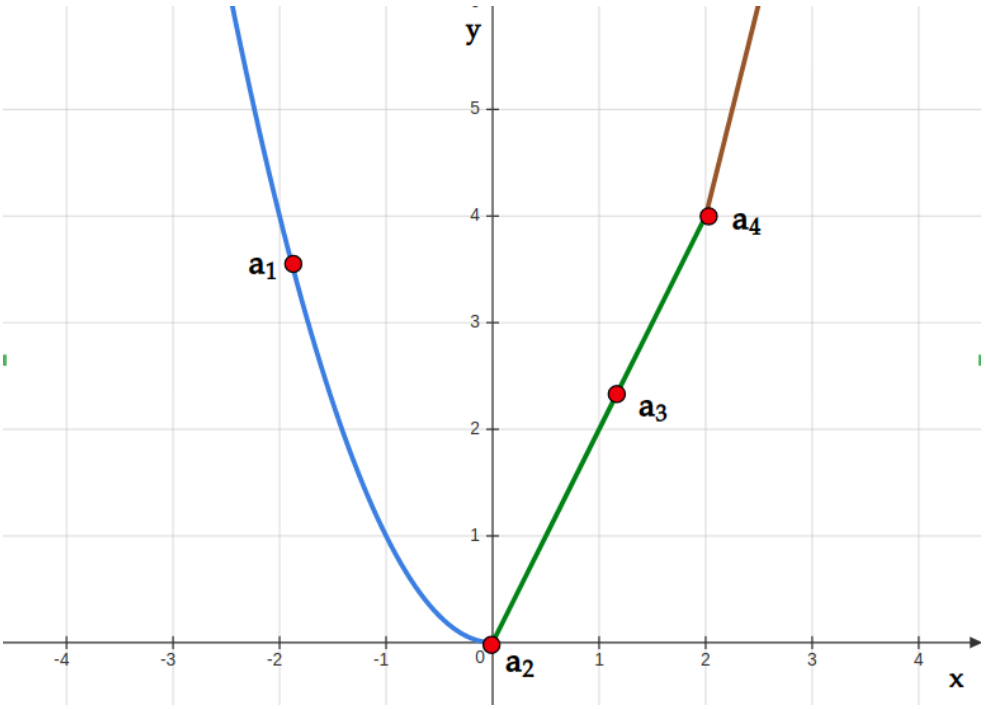
If the value of θ is 13, which of the following weight vectors will be selected as the final weight vector by ridge regression?

Note: θ is as per lectures. That is, $|||w||^2 \leq \theta$

- ○ w_1
- ● w_2
- ○ w_3
- ○ w_4

3) Consider the following piece-wise function as shown in the image:

$$y(x) = \begin{cases} x^2 & x \leq 0 \\ 2x & 0 \leq x \leq 2 \\ 4x - 4 & 2 \leq x \end{cases}$$



How many sub-gradients are possible at points a_1, a_2, a_3 and a_4?

- $a_1 : Many, a_2 : One, a_3 : Many, a_4 : One$
- $a_1 : One, a_2 : Many, a_3 : Many, a_4 : One$
- $a_1 : One, a_2 : Many, a_3 : One, a_4 : Many$
- $a_1 : Many, a_2 : One, a_3 : One, a_4 : Many$

4) For a data set with 1000 data points and 50 features, 10-fold cross-validation will perform validation of how many models?

- [x] 10
- [ ] 50
- [ ] 1000
- [ ] 500

5) For a data set with 1000 data points and 50 features, assume that you keep 80% of the data for training and remaining 20% of the data for validation during k-fold cross-validation. How many models will be validated during cross-validation?

- [ ] 80
- [ ] 20
- [x] 5
- [ ] 4

6) For a data set with 1000 data points and 50 features, how many models will be trained during Leave-One-Out cross-validation?

- [x] 1000
- [ ] 50
- [ ] 5000
- [ ] 20

7) The mean squared error of $\hat{w}_{ML}$ will be small if

- [ ] The eigenvalues of XX^T are small.
- [ ] The eigenvalues of $(XX^T)^{-1}$ are large.
- [x] The eigenvalues of XX^T are large.
- [x] The eigenvalues of $(XX^T)^{-1}$ are small.

8) The eigenvalues of a 3×3 matrix A are 2, 5 and 1. What will be the eigenvalues of the matrix A^{-1}

- [ ] 4, 25, 1
- [ ] 2, 5, 1
- [x] 0.5, 0.2, 1
- [ ] 0.6, 0.9, 0.1

Find $\frac{1}{\lambda}$ to get eigen values of A^{-1}

$\frac{1}{2}, \frac{1}{5}, \frac{1}{1}$