Q1-1: The parameters to be estimated in the Linear Regression model $y = \beta_0 + \beta_1 x$ are

1. β_0, β_1
2. y
3. β_0, y
4. β_1, y

Q1-1: The parameters to be estimated in the Linear Regression model $y = \beta_0 + \beta_1 x$ are

1. β_0, β_1 ⬅
2. y
3. β_0, y
4. β_1, y

Q1-2: In the regression model $y = \beta_0 + \beta_1 x$, the change in y for a one unit increase in x is:

1. Will always be the same amount, β_0
2. Will always be the same amount, β_1
3. Will depend on both β_0 and β_1
4. None of above

Q1-2: In the regression model $y = \beta_0 + \beta_1 x$, the change in y for a one unit increase in x is:

1. Will always be the same amount, β_0
2. Will always be the same amount, β_1 ←
3. Will depend on both β_0 and β_1
4. None of above

If $y = \beta_0 + \beta_1 x, x' = x + 1$, then we have $y' = \beta_0 + \beta_1 x' = \beta_0 + \beta_1 (x + 1) = \beta_0 + \beta_1 x + \beta_1 = y + \beta_1$.

Q1-3: Suppose that the value of r^2 for an estimated regression model is exactly zero. Which are true? (Multiple answers)

1. The slope coefficient estimate will be zero
2. The fitted line will be horizontal
3. The fitted line will be vertical
4. The intercept coefficient estimate will be zero

Q1-3: Suppose that the value of r^2 for an estimated regression model is exactly zero. Which are true? (Multiple answers)

1. The slope coefficient estimate will be zero ←
2. The fitted line will be horizontal ←
3. The fitted line will be vertical
4. The intercept coefficient estimate will be zero

If $r^2 = 0$, then the linear function has exactly the same error as that of a constant. So it is just the function $y = \bar{y}$

Q2-1: In general, the Least Squares Regression approach finds the equation: (multiple answers)

1. that includes the best set of predictor variables
2. of the best fitting straight line/hyperplane through a set of points
3. with the lowest r^2, after comparing all possible models
4. that has the smallest sum of squared errors

Q2-1: In general, the Least Squares Regression approach finds the equation: (multiple answers)

1. that includes the best set of predictor variables
2. of the best fitting straight line/hyperplane through a set of points
3. with the lowest r^2, after comparing all possible models
4. that has the smallest sum of squared errors

For 1: there can be noise in the labels, so may not find the best set of predictor variables.
For 3: actually the objective is to get the lowest sum of squared errors, so is to get the highest r^2

Q2-2: Suppose you train two linear regression models on the same dataset, one with 0 regularization, one use large positive λ for regularization. You get the following two vectors of coefficients.

$$\theta_1 = [55, 66, 77, 88]$$
$$\theta_2 = [5, 6, 7, 8]$$

Which linear model has utilized regularization during training?

1. Model 1
2. Model 2
3. Need more information to tell

Q2-2: Suppose you train two linear regression models on the same dataset, one with 0 regularization, one use large positive λ for regularization. You get the following two vectors of coefficients.

$$\theta_1 = [55, 66, 77, 88]$$
$$\theta_2 = [5, 6, 7, 8]$$

Which linear model has utilized regularization during training?

Regularization will penalize the norm of the parameter vector, so it will lead to smaller norm solutions.

1. Model 1
2. Model 2 ⟵
3. Need more information to tell

Q2-3: Consider the regression problem

$$\min_{\beta} ||\boldsymbol{y} - \boldsymbol{X}\beta||^2 + \lambda \, ||\beta||^2$$

Which of the following is appropriate if we want to further penalize the flexibility of the model?

1. Increase λ
2. Decrease λ
3. Set $\lambda = 1$
4. Set $\lambda < 0$

Q2-3: Consider the regression problem

$$\min_{\beta} ||\boldsymbol{y} - \boldsymbol{X}\beta||^2 + \lambda \, ||\beta||^2$$

Which of the following is appropriate if we want to further penalize the flexibility of the model?

1. Increase λ
2. Decrease λ
3. Set $\lambda = 1$
4. Set $\lambda < 0$

Q3-1: Is logistic regression an appropriate substitute for linear regression?

1. Yes
2. No

Q3-1: Is logistic regression an appropriate substitute for linear regression?

1. Yes
2. No ⟵

Logistic regression is for linear classification (though it’s called regression for historical reasons).

Q3-2: Given the training data

$$(x, y): (0, +), (1, -), (2, +), (3, -)$$

Is this true: A logistic regression model can be trained to classify the data points with zero training error?

1. True
2. False

Q3-2: Given the training data

$$(x, y): (0, +), (1, -), (2, +), (3, -)$$

Is this true: A logistic regression model can be trained to classify the data points with zero training error?

1. True
2. False ⬅

The decision boundary between + and – by a logistic regression model must be a linear hyperplane. It is a threshold in 1-dim space. Then it cannot get zero classification errors on the data since the labels are interweaving.

Q3-3: If a dataset is linearly separable, which of the following training methods is more suitable to train a logistic regression classifier?

1. MLE
2. MAP

Q3-3: If a dataset is linearly separable, which of the following training methods is more suitable to train a logistic regression classifier?

1. MLE
2. MAP