X

Week 2: Assignment 2

The due date for submitting this assignment has passed.

Due on 2023-08-09, 23:59 IST.

Assignment submitted on 2023-08-09, 12:44 IST

1) The parameters obtained in linear regression ***1 point***

- can take any value in the real space
- are strictly integers
- always lie in the range [0,1]
- can take only non-zero values

Yes, the answer is correct.
Score: 1

Accepted Answers:
can take any value in the real space

2) Suppose that we have N independent variables $(X_1, X_2, \ldots X_n)$ and the dependent variable is Y . Now imagine that you are applying linear regression by fitting the best fit line using the least square error on this data. You found that the correlation coefficient for one of its variables (Say X_1) with Y is -0.005. ***1 point***

- Regressing Y on X_1 mostly does not explain away Y .
- Regressing Y on X_1 explains away Y .
- The given data is insufficient to determine if regressing Y on X_1 explains away Y or not.

Yes, the answer is correct.
Score: 1

Accepted Answers:
Regressing Y on X_1 mostly does not explain away Y .

3) Which of the following is a limitation of subset selection methods in regression? ***1 point***

- (•) They tend to produce biased estimates of the regression coefficients.
- ( ) They cannot handle datasets with missing values.
- ( ) They are computationally expensive for large datasets.
- ( ) They assume a linear relationship between the independent and dependent variables.
- ( ) They are not suitable for datasets with categorical predictors.

No, the answer is incorrect.
Score: 0
Accepted Answers:
They are computationally expensive for large datasets.

4) The relation between studying time (in hours) and grade on the final examination (0-100) in a random sample of students in the Introduction to Machine Learning Class was found to be:Grade = 30.5 + 15.2 (h) ***1 point***
How will a student's grade be affected if she studies for four hours?

- ( ) It will go down by 30.4 points.
- ( ) It will go down by 30.4 points.
- (•) It will go up by 60.8 points.
- ( ) The grade will remain unchanged.
- ( ) It cannot be determined from the information given

Yes, the answer is correct.
Score: 1
Accepted Answers:
It will go up by 60.8 points.

5) Which of the statements is/are True? ***1 point***

- [x] Ridge has sparsity constraint, and it will drive coefficients with low values to 0.
- [x] Lasso has a closed form solution for the optimization problem, but this is not the case for Ridge.
- [x] Ridge regression does not reduce the number of variables since it never leads a coefficient to zero but only minimizes it.
- [x] If there are two or more highly collinear variables, Lasso will select one of them randomly

No, the answer is incorrect.
Score: 0
Accepted Answers:
Ridge regression does not reduce the number of variables since it never leads a coefficient to zero but only minimizes it.
If there are two or more highly collinear variables, Lasso will select one of them randomly

Week 3 ()
Week 4 ()
Week 5 ()
Week 6 ()
Week 7 ()
Week 8 ()
Week 9 ()
Text Transcripts ()
Download Videos ()
Books ()
Problem Solving Session - July 2023 ()

6) Find the mean of squared error for the given predictions: ***1 point***

Y	f(x)
1	2
2	3
4	5
8	9
16	15
32	31

Hint: Find the squared error for each prediction and take the mean of that.

- (●) 1
- ( ) 2
- ( ) 1.5
- ( ) 0

Yes, the answer is correct.
Score: 1

Accepted Answers:
1

7) Consider the following statements: ***1 point***

Statement A: In Forward stepwise selection, in each step, that variable is chosen which has the maximum correlation with the residual, then the residual is regressed on that variable, and it is added to the predictor.
Statement B: In Forward stagewise selection, the variables are added one by one to the previously selected variables to produce the best fit till then

- (●) Both the statements are True.
- ( ) Statement A is True, and Statement B is False
- ( ) Statement A is False and Statement B is True
- ( ) Both the statements are False.

No, the answer is incorrect.
Score: 0

Accepted Answers:
Both the statements are False.

8) The linear regression model $y = a_0 + a_1x_1 + a_2x_2 + \ldots + a_px_p$ is to be fitted to a set of N training data points having p attributes each. Let X be $N \times (p+1)$ vectors of input values (augmented by 1's), Y be $N \times 1$ vector of target values, and θ be $(p+1) \times 1$ vector of parameter values $(a_0, a_1, a_2, \ldots, a_p)$. If the sum squared error is minimized for obtaining the optimal regression model, which of the following equation holds? ***1 point***

- ( ) $X^TX = XY$
- ( ) $X\theta = X^TY$
- ( ) $X^TX\theta = Y$
- (●) $X^TX\theta = X^TY$

Yes, the answer is correct.

Score: 1

Accepted Answers:

$X^T X\theta = X^T Y$

9) Which of the following statements is true regarding Partial Least Squares (PLS) regression? ***1 point***

- (●) PLS is a dimensionality reduction technique that maximizes the covariance between the predictors and the dependent variable.
- ( ) PLS is only applicable when there is no multicollinearity among the independent variables.
- ( ) PLS can handle situations where the number of predictors is larger than the number of observations.
- ( ) PLS estimates the regression coefficients by minimizing the residual sum of squares.
- ( ) PLS is based on the assumption of normally distributed residuals.
- ( ) All of the above.
- ( ) None of the above.

Yes, the answer is correct.
Score: 1

Accepted Answers:

PLS is a dimensionality reduction technique that maximizes the covariance between the predictors and the dependent variable.

10) Which of the following statements about principal components in Principal Component Regression (PCR) is true? ***1 point***

- ( ) Principal components are calculated based on the correlation matrix of the original predictors.
- ( ) The first principal component explains the largest proportion of the variation in the dependent variable.
- (●) Principal components are linear combinations of the original predictors that are uncorrelated with each other.
- ( ) PCR selects the principal components with the highest p-values for inclusion in the regression model.
- ( ) PCR always results in a lower model complexity compared to ordinary least squares regression.

Yes, the answer is correct.
Score: 1

Accepted Answers:

Principal components are linear combinations of the original predictors that are uncorrelated with each other.