

MACHINE LEARNING

Q1. Ans. :- (D)  $O(n^3)$

Q2. Ans. :- (C) Polynomial Regression

Q3. Ans. :- (B) Gradient Descent

Q4. Ans. :- (C) Lasso

Q5. Ans. :- (D) All of the above

Q6. Ans. :- (B) True

Q7. Ans. :- (C) it does not matter whether half is there or not.

Q8. Ans. :- (C) Both of them

Q9. Ans. :- (B) It becomes slow when number of features are very large.

Q10. Ans. :- (A) Linear Regression will have high bias and low variance.  
(D) Polynomial with degree 5 will have high bias and low variance.

Q11. Ans. :- (C) It discovers causal relationship.  
(D) No inference can be made from regression line.

Q12. Ans. :- Batch gradient descent, stochastic gradient descent or mini-batch gradient descent. It does not need to load the entire dataset into memory for taking 1st step of gradient descent. Batch gradient descent is used when it has enough memory to load all data. But Normal equations method cannot be used because computational complexity grows very quickly with number of features.

Q13. Ans. :- The normal equations method does not require normalizing the features, so it remains unaffected by features in the training set having very different scales. Feature scaling is required for the various gradient descent algorithms. Feature scaling will help gradient descent converge quicker. The cost function will have the shape of an elongated bowl, so the Gradient Descent Algorithms will take a long time to converge. To solve this you should scale the data before training the model.