ML

Q1:-(B) O(n)

Q2:- (B) Logistic Regression

Q3:- (B) Gradient Descent

Q4:-(C) Lasso

Q5:-(A) Stochastic Gradient Descent

Q6:-(A) True

Q7:- (A) scaling cost function by half makes gradient descent converge faster.

Q8:-(B) Correlation

Q9:- (B) It becomes slow when number of features are very large. And (D) it does not make use of dependent variable.

Q10:- (B) Linear Regression will have low bias and high variance and (C) Polynomial with degree 5 will have low bias and high variance.

Q11:- (D) No inference can be made from regression line.

Q12:- We can use batch gradient descent, stochastic gradient descent, or mini-batch gradient descent.
SGD and MBGD would work the best because neither of them need to load the entire dataset into memory in order to take 1 step of gradient descent. Batch would be ok with the caveat that you have enough memory to load all the data.

Q13:-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.