1. B) always increases
2. C) SSE = SSR – SST
3. C) By its slope
4. A) difference between the actual value and the predicted value
5. B) can be either -1 or 1
6. A) Scatter plot
7. B) f-statistics
8. C) Ridge
9. A) It shows the causal relationship between dependent and independent variables & D) It is a straight line that is the best approximation of the given data set
10. B) Generalizing the test set C) Automatic feature selection
11. A) Normal Equation B) Singular Value Decomposition

12.R-squared evaluates the scatter of the data points around the fitted regression line. It is also called the coefficient of determination, or the coefficient of multiple determination for multiple regression. For the same data set, higher R-squared values represent smaller differences between the observed data and the fitted values. R-squared is always between 0 and 100%:

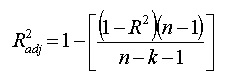
0% represents a model that does not explain any of the variation in the response variable around its mean. The mean of the dependent variable predicts the dependent variable as well as the regression model.

100% represents a model that explains all of the variation in the response variable around its mean.

{\displaystyle R^2 = \frac {\text{Variance explained by the model}}{\text{Total variance}}}

R2 shows how well terms (data points) fit a curve or line. Adjusted R2 also indicates how well terms fit a curve or line, but adjusts for the number of terms in a model. If you add more and more useless variables to a model, adjusted r-squared will decrease. If we add more useful variables, adjusted r-squared will increase.  
Adjusted R2 will always be less than or equal to R2.

You only need R2 when working with samples In other words, R2 isn’t necessary when we have data from an entire population

The formula is:  
[](https://www.statisticshowto.com/wp-content/uploads/2013/09/r-squared-adjusted.jpg)

13. A cost function is a measure of how wrong the model is in terms of its ability to estimate the relationship between X and y. This is typically expressed as a difference or distance between the predicted value and the actual value. The cost function (we may also see this referred to as loss or error.) can be estimated by iteratively running the model to compare estimated predictions against “ground truth” — the known values of y.

The objective of a ML model, therefore, is to find parameters, weights or a structure that minimises the cost function.

14.SST is the maximum sum of squares of errors for the data because the minimum information of Y itself was only used for the baseline model. For the regression model, we square all the differences ③ Ŷ − Ȳ and sum them up, which is called sum of squares due to regression (SSR), ∑(Ŷ − Ȳ)2. SSR is the additional amount of explained variability in Y due to the regression model compared to the baseline model. The difference between SST and SSR is remaining unexplained variability of Y after adopting the regression model, which is called as sum of squares of errors (SSE). SSE can be directly obtained by sum of squares of residual, ∑(Y − Ŷ)2.

15. 1.Mean Squared Error (MSE)

The most common metric for regression tasks is MSE. It has a convex shape. It is the average of the squared difference between the predicted and actual value. Since it is differentiable and has a convex shape, it is easier to optimize.

Image for post

Image for post

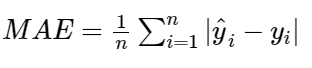
Mean squared error. Image by the author.

MSE penalizes large errors.

2.Mean Absolute Error (MAE)

This is simply the average of the absolute difference between the target value and the value predicted by the model. Not preferred in cases where outliers are prominent.

Image for post



Mean absolute error. Image by the author.

## 3.R-squared or Coefficient of Determination

This metric represents the part of the variance of the dependent variable explained by the independent variables of the model. It measures the strength of the relationship between your model and the dependent variable.

## 4.Root Mean Squared Error (RMSE)

This is the square root of the average of the squared difference of the predicted and actual value.

R-squared error is better than RMSE. This is because R-squared is a relative measure while RMSE is an absolute measure of fit (highly dependent on the variables — not a normalized measure).

Basically, RMSE is just the root of the average of squared residuals.

MAE does not penalize large errors.