# Regression

Root mean-square error (RMSE) – aka norm or Euclidean norm (Euclidean distance b/w hypothesis and label):

Mean absolute error (MAE) – aka norm or Manhattan norm (Manhattan distance, i.e. distance as if you were traversing city blocks):

Choosing LSE as the cost function means that you will minimize both LSE and RMSE through optimization. However, keep in mind that LSE is the cost function while RMSE is the performance metric.

You can choose any order of norm as your performance metric.

gives the number of nonzero elements in the vector, and gives the maximum absolute value in the vector (all other terms go to zero).

Higher means more sensitivity to outliers (large errors have more weight). When outliers are exponentially rare (like when samples are Gaussian-distributed), then RMSE gives you a more accurate picture of your prediction performance. This directly follows from generalized linear models and maximum likelihood estimate.