# CSC311 A3 Non-programming

David Pham: 1002597825

Dec 8 2020

**Question 1(e).**
Suppose we have that

$$x^{(i)} - \mu = M^{(i)}$$

then we can say from the empirical covariance matrix, we have,

$$\frac{1}{N}\sum_{i=1}^{N}(x^{(i)} - \mu)(x^{(i)} - \mu)^T$$

Using what we suggested above:

$$= \frac{1}{N}\sum_{i=1}^{N} M^{(i)} M^{(i)T}$$

Using definition of matrix

$$= \frac{1}{N}\sum_{i=1}^{N} M_{ij} M_{ij}^T$$

Using definition of transpose

$$= \frac{1}{N}\sum_{i=1}^{N} M_{ij} M_{ji}$$

$$= \frac{1}{N}\sum_{i=1}^{N} M_{ji} M_{ij}$$

$$= \frac{1}{N}[M^T M]_{jj}$$

This proves that this is a vectorized form of the empirical covariance formula and that it is symmetric

**Question 2(c).**
Overfitting occurs when the regularization parameter is $< 10^{-4}$. In the graph, you can see the training accuracy vs. validation accuracy to drastically spread out. Underfitting occurs when the regularization parameter is $> 10^{-2}$. You can see if the figure that the training vs validation accuracy are too close together as the regularization parameter gets greater.

**Question 2(e).**
Overfitting occurs when K is $> 10$. In the graph, you can see the training accuracy vs. validation accuracy to drastically spread out. Underfitting occurs in when K $< 5$. You can see if the figure that the training vs validation accuracy are too close together as K gets greater.