**Karan Agarwalla – 180050045**

**Problem 1**

**Problem 1.1**

=

Now,

Now we proved in class that is a valid kernel for . Also, if K1 and K2 are valid kernels then K1K2 and aK1+bK2 are valid kernels for a>=0, b>=0. As is a valid kernel, then by first of the results is a valid kernel. Then is valid kernel as it is a positive linear combination of valid kernels.

Now

**Problem 2**

**Problem 2.1**

**(a)**

Hence sigmoid probability in case of binary logistic regression is a special case of multi-class regression with weight vector w = w1 – w­0. Also y(i)in case of logistic regression is simply y1(i) and y0(i) = 1 - y1(i). Hence, simplifying the expression,

which is the cross-entropy loss function used to train binary logistic regression.

**(b)**

(follows from (1))

Let’s define , i.e., probability of kth label on ith example

Hence,

where Pj is vector of size N\*1 whose entries are Pij­ for each example i. Similarly, Yj is vector of size N\*1 whose entries are for each example

Hence,

where P and Y are matrices of size N\*K where K is number of classes and is matrix of size N\*F where F is number of features.

**Problem 2.2**

**(b)**

Test accuracy obtained = 86.32%

Test accuracy for model M = 84.18%

Accuracy is not a good metric when the number of test examples for a particular class dominate the dataset (here class 0). Since in such a case the accuracy of that particular class dominates the evaluation of the model. Here in model M we always predict 0, hence we achieve quite high accuracy score. Hence we need a metric that specifies how well the model does on each class taking into account both false negatives and false positives.

(c)

F1 score obtained for test set: 0.301

F1 score for model M: 0

F1 is a good evaluation metric since it strikes a balance between precision and recall. It takes into account misclassifications for each class and is not dependent on the majority class (here 0). F scores evaluate how well the model does on each class.

(e)

Test accuracy for logistic regression: 84.435%

Test accuracy for perceptron: 78.28%

Hence logistic regression achieves higher accuracy. This can be attributed to the smooth differentiable objective function in case of logistic regression against a step function in perceptron. Hence it is better able to converge to minima and is softer than perceptron, hence finding a better decision boundary.