Name: Wei Shen, Gabriel Hanson, Ketian Tu, Ajay Sanjeevan
Computing ID: zyy5hb, gh8sj, kt9sh, zwz3wu

Problem 1

a) Bias
b) Variance
c) B. As λ increases, flexibility of ridge regression fit decreases leading to increased bias
d) B
e) A
f) Reduce the bias from the chosen hypothesis. Reduce the complexity of the problem. Also avoids overfitting.
g) False: overfitting is more likely when hypothesis space is small
h) True
i) False

Problem 2

a) Reject: The training errors are not always accurate. We need to see how the model behaves with the test data and compare the models.
b) Reject: The parameters are tuned on the training set. By tuning parameters on the test set, the test set is converted into a training set.

Problem 3

a) (2) reduces variance at the expense of higher bias.
b) (2) Variance
c) (1) $y=\beta 0$
d) (2) High bias

Problem 4

Yes, it is possible to modify the hypothesis space so that the bayes predictor is included. One would have to allow for irrational numbers as the bayes predictor was irrational.

Problem 5

You want to pick a hypothesis space with a VC dimension that minimizes the sum of the training error and the other term. If the complexity is too low, training error will dominate whereas if the VCdimension is too high, the other term will dominate. In this example, the training error (the empirical risk) is related to the bias while the other term is related to the variance. For the KNN example where K= 1, vcdim = inf. Relating this to the equation in problem 5, while the training error will be small, the test error will be very large, resulting in overfitting.

<u>Problem 6</u>
VC dimension of H is 3:
There exists a set of C of size 3 that is shattered by H
For 4 points the labeling cannot be shattered if +ve and -ve alternate. (The hypothesis can only cross the axis at most 2 times)