HW #2 Due: 4/3/2018

In the following problems, you can use existing tools to find eigenvalues and eigenvectors of matrices. For pedagogical reasons, you need to complete other parts of the programs by yourself, but not calling existing libraries.

1. In the lecture, we mentioned that the term $\lambda \sum_i w_i^2$ is a regularization term.
   i. Use your own words to explain why it is so.
   ii. If we implement two regression programs, one using $\lambda \sum_i w_i^2$ and the other one using $\lambda \sum_i w_i^{10}$, which one do you expect has lower bias? How about variance? Why? Assuming that we use the same λ for both programs.
2. We mentioned that the covariance matrix may be ill-conditioned. Find the (sample) covariance matrices for the three classes of the Iris dataset and compute the condition numbers for the covariance matrices. For simplicity, use the following as the condition number: $\kappa(\mathrm{A}) = \left|\frac{\lambda_{\max}}{\lambda_{\min}}\right|$, where $\lambda_{\max}$ and $\lambda_{\min}$ are the largest and smallest eigenvalues of matrix A.
3. In this problem, you are asked to use the Iris dataset to perform PCA dimensionality reduction before classification. Randomly draw 35 samples in each class to find the vectors $\boldsymbol{w}_{(j)}$ for the largest two principal components. Recall that PCA is unsupervised; therefore, you need to use 35×3 = 105 data points to find the parameters of the PCA. Implement the 3-NN classifier to test the rest 15 samples in each class and record the accuracy. Repeat the drawing and the *k*-NN classification 10 times and compute the average accuracy and variance. For simplicity, use the Euclidean distance in the *k*-NN computation.
4. Following the general steps of problem 2, but use the FA approach for dimensionality reduction. For simplicity, you may assume $\Psi = 0$ and use the LS solutions.
5. Repeat problem 2 by using LDA as the reduction method. Remember to compute the parameters for each class in order to use LDA.