CS 229 Machine Learning, spring 2019

Homework 9:

Unsupervised Learning: PCA and SVD

Due Thursday May 9, 11:59pm

Submit by the **blackboard system**

**Question 1 Principal Components Analysis (PCA)**

1. (15pts) Discuss and show the proof of how PCA maximizes the variance of projected data.
2. (15pts) Discuss and show the proof of how PCA minimizes the average projection cost -- defined as the mean squared distance between the data points and their projections.

1. (30pts) Implementation of PCA (code submission is required).

**CASE A**, for students whose project involves high-dim data, e.g., images. Apply PCA before your classification/clustering job. Show whether PCA is helpful on improving your classification/clustering result, e.g., comparing the accuracy with PCA and without PCA.

**CASE B**, for other students. Take the same wine quality data we used in the SVM homework . Apply PCA at first, and then learn SVM from the new representation. Show whether PCA is helpful on improving your classification accuracy (choosing any one of the kernels is ok, and parameters can be set according to the search result in SVM homework).

For both Case A and B,

How the result will be different if different numbers of PCs are selected for the new representation?

**Question 2 Singular Value Decomposition (SVD)**

1. (10pts) Discuss

* How PCA and SVD are related to each other (with proof).
* What’s the difference between PCA and SVD when both of them are used for reduce the dimensionality

1. (30pts) Implementation of SVD (code submission is required).

**CASE A**, for students whose project involves high-dim data, e.g., images. Apply SVD before your classification/clustering job. Show whether SVD is helpful on improving your classification/clustering result, e.g., comparing the accuracy with SVD and without SVD.

**CASE B**, for other students. Take the same wine quality data we used in the SVM homework . Apply SVD at first, and then learn SVM from the reduced dimension space. Show whether SVD is helpful on improving your classification accuracy (choosing any one of the kernels is ok, and parameters can be set according to the search result in SVM homework).