

COMS 4721 – Machine Learning for Data Science

Daniel Kost-Stephenson – dpk2124

April 19th, 2016

Question 2

Part a)

This does not indicate a problem with their software because the described scenario could occur if one of the top eigenvalues they select are the same and if another eigenvalue is the same (not necessarily in the one they select). If two eigenvalues are the same, there exists a two dimensional subspace in which the variance in every possible direction is the same. Thus, the three eigenvectors obtained by Alice and Bob could be different, but still represent directions of maximum variance in the three dimensional subspace. This means that any rotation of the eigenvectors in the two dimensional subspace (where the eigenvalues are equal) would still yield directions of maximum variance (in the subspace) and would thus mean the software is not incorrect because it could select any two dimensions within that subspace. For example, the eigenvectors are ambiguous in the case of a two dimensional circle with equal variance in every direction.

Part b)

This indicates a problem with their software because the orthonormal projection of a point in a higher dimensional space into a lower dimensional space should be equivalent regardless of the direction of their respective eigenvectors. Thus, Alice and Bob should get the same result when they project points into the three dimensional space spanned by $\{\lambda_1, \lambda_2, \lambda_3\}$ because the first three eigenvalues are equal. The fact that Alice and Bob did not get the same result indicates a problem with their software.

Part c)

The answer from b) does not extend to four dimensions: an orthonormal projection of a point in a higher dimensional space into a lower dimensional space should only be equivalent if the dimensionality of the projected space is equal to the number of equal eigenvectors obtained by Alice and Bob. Since there are four equal eigenvectors and we are projecting the data into a three dimensional space, Alice and Bob could select different eigenvalues to project into, thus there is no guarantee that they will obtain the same result. Therefore, this does not indicate a problem with their current software.