PCA

UNDERSTANDING: You run a PCA and the percentage of the variance that is explained by the first two components is 80%. How is that variance calculated?

Answer: The diagonals of the covariance matrix represent the sum of variances of the variables. When you take the sum of the first two diagonals of the covariance matrix, you get the variance that is explained by these two components.

LDA and PCA

TRANSFERRING: What do LDA and PCA have in common and what are their differences?

Answer: Both look for linear combinations of variables and project in a lower dimensional space to explain the data.
LDA: projects in one direction which separates the classes well. It's aim is to minimize the fisher criterion.
PCA: is unsupervised and projects in the direction of the large variance.

SVM/Kernels

UNDERSTANDING: What happens to the performance of a SVM classifier when the diagonal entries of the GRAM matrix are large relative to the off-diagonal entries?

Answer: This problem is called diagonal dominance. The performance of the SVM classifier is poor because the kernel is equivalent to the identity matrix.