HW 3b

Problem 1 (2= 1+1 points) We know that if $M_{D\times D}$ is real and symmetric matrix then M can be factored as

$$M = U\Lambda U^T$$

Where $U = [u_1, u_2, \cdots, u_D]$ are the eigenvectors of M. Eigen vectors u_i are usually normalized in that case U is orthonormal i.e U^{-1} =U^T and

$$\Lambda = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_D \end{bmatrix}$$

contains eignevalues along the diagonal.
Let $X_c = U_{N\times D}\Sigma_{D\times D}V^T_{D\times D}$ be reduced SVD of our centrlize matrix of size $X_{C_{N\times D}}$ where $N \geq D$.

1. Prove that V contains the eigen vector of $X_c^T X_C$.
2. Prove $\Sigma^2 = \Sigma\,\Sigma$ contains eigen values of $X_c^T X_c$ along the diagonal.

Similarly U contains the eigen vectors of $X_c X_c^T$ but you need not to prove this part.

Problem 2 (2 points) Let say we have C classes and each point $x_i \in R^d$. Let $n_k =$ number of sample in class k, hence $n = \sum_{k=1}^{C} n_k$ total samples. Class k mean $\mu_k = \frac{\sum_{i:y_i=k} x_i}{n_i}$ and over all mean is $\mu = \frac{\sum_i x_i}{n}$. Also, within class scatter matrix is $S_w = \sum_{k=1}^{C} S_k$ where $S_k = \sum_{i:y_i=k}((x_i - \mu_i)(x_i - \mu_i)^T)$, between class scatter matrix is $S_b = \sum_{k=i}^{k=C} n_k(\mu_i - \mu)(\mu_i - \mu)^T$ and total scatter matrix is $S_T = \sum_i (x_i - \mu)(x_i - \mu)^T$. Prove that

$$S_T = S_w + S_b$$