1 Deriving the second principal component

1. Recall that PCA tries to minimize the reconstruction error between the data points and the projections of the data points onto the principle componenets. We have derived the first principle component in lecture and last week's recitation. This week we will derive the second principle component. Let $J(\mathbf{v_2}) = \frac{1}{n}\sum_{i=1}^{n} ||\mathbf{x^{(i)}} - z_1^{(i)}\mathbf{v_1} - z_2^{(i)}\mathbf{v_2}||_2^2$ given the constraints $\mathbf{v_1}^T\mathbf{v_2} = 0$ and $\mathbf{v_2}^T\mathbf{v_2} = 1$. Here, n is the number of data points, $\mathbf{v_1}, \mathbf{v_2}$ are the first and the second principle component, and $\mathbf{z^{(i)}}$ denotes the principle encoding of the ith data point $\mathbf{x^{(i)}}$. Recall that we've defined $z_1^{(i)} = \mathbf{v_1}^T\mathbf{x^{(i)}}$. Define $z_2^{(i)}$, which is the second principle encoding of $\mathbf{x^{(i)}}$.

2. Show that the value of $\mathbf{v_2}$ that minimizes J is given by the eigenvector of $\mathbf{C} = \frac{1}{n}\sum_{i=1}^{n}(\mathbf{x^{(i)}}\mathbf{x^{(i)}}^T)$ with the second largest eigenvalue. Assumed we have already proved the $\mathbf{v_1}$ is the eigenvector of $\mathbf{C}$ with the largest eigenvalue.

2 SVD

(a) Find the SVD of $X = \begin{bmatrix} 4 & 4 \\ 3 & -3 \end{bmatrix}$

(b) How does SVD relate to PCA?

(c) How does SVD relate to Matrix Factorization?