CS 189: Introduction to Machine Learning - Discussion 12

1. SVD Warmup

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

2. Derivation of PCA

In this question we will derive PCA. PCA aims to find the direction of maximum variance among a dataset. You want the line such that projecting your data onto this line will retain the maximum amount of information. Thus, the optimization problem is

$$\max_{u:\|u\|_2=1} \frac{1}{n}\sum_{i=1}^{n}\left(u^T x_i - u^T \hat{x}\right)^2$$

where n is the number of data points and $\hat{x}$ is the sample average of the data points.

(a) Show that this optimization problem can be massaged into this format

$$\max_{u:\|u\|_2=1} u^T \Sigma u$$

where $\Sigma = \frac{1}{n}\sum_{i=1}^{n}(x_i - \hat{x})(x_i - \hat{x})^T$.

(b) Show that the maximizer for this problem is equal to v_1, where v_1 is the eigenvector corresponding to the largest eigenvalue λ_1. Also show that optimal value of this problem is equal to λ_1.

3. Deriving the second principal component

   (a) Let $J(\mathbf{v_2}, \mathbf{z_2}) = \frac{1}{n}\sum_{i=1}^{n}(\mathbf{x_i} - z_{i1}\mathbf{v_1} - z_{i2}\mathbf{v_2})^T(\mathbf{x_i} - z_{i1}\mathbf{v_1} - z_{i2}\mathbf{v_2})$ given the constraints $\mathbf{v_1^T v_2} = 0$ and $\mathbf{v_2^T v_2} = 1$. Show that $\frac{\partial J}{\partial \mathbf{z_2}} = 0$ yields $z_{i2} = \mathbf{v_2}^T\mathbf{x_i}$.

   (b) We have shown that $z_{i2} = \mathbf{v_2}^T\mathbf{x_i}$ so that the second principal encoding is gotten by projecting onto the second principal direction. Show that the value of $\mathbf{v_2}$ that minimizes J is given by the eigenvector of $\mathbf{C} = \frac{1}{n}\sum_{i=1}^{n}(\mathbf{x_i x_i^T})$ with the second largest eigenvalue. Assumed we have already proved the v_1 is the eigenvector of C with the largest eigenvalue.