CS/ECE/ME 532

Unit 3 Practice Problems

1. In (a) - (e), let the SVD of a matrix be given as $\boldsymbol{A} = \boldsymbol{U}\Sigma\boldsymbol{V}^T$, with $\Sigma_{1,1} = \sigma_1$ denoting the first singular value.

a) σ_1 is the largest value of $||\boldsymbol{A}\boldsymbol{x}||_2$ for any unit norm vector $\boldsymbol{x}$.
True False

b) σ_1 is the largest value of $||\boldsymbol{x}^T\boldsymbol{A}||_2$ for any unit norm vector $\boldsymbol{x}$.
True False

c) σ_1 is ℓ_2 norm of the vector $\boldsymbol{x}$ that maximizes $||\boldsymbol{x}^T\boldsymbol{A}||_2$.
True False

d) $\sigma_1 = ||\boldsymbol{A}||_{op}$.
True False

e) $\sigma_1^2 = \sum_{i,j} \boldsymbol{A}_{i,j}^2$.
True False

2. You collect eight, four dimensional data points that you store as columns in a matrix $\boldsymbol{X}$:

$$\boldsymbol{X} = \begin{bmatrix}\boldsymbol{x}_1 \ \boldsymbol{x}_2 \ \dots \ \boldsymbol{x}_8\end{bmatrix} \tag{1}$$

You cluster the 8 data points by running the k-means algorithm with $k = 3$, which produces cluster centers, $\boldsymbol{t}_1, \boldsymbol{t}_2, \boldsymbol{t}_3$:

$$\boldsymbol{T} = [\boldsymbol{t}_1 \ \boldsymbol{t}_2 \ \boldsymbol{t}_3] = \begin{bmatrix} 1 & 1 & -2 \\ -4 & -2 & 0 \\ 7 & -6 & 2 \\ 7 & -6 & 9 \end{bmatrix} \tag{2}$$

a) The data points $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3$ are assigned to cluster $\boldsymbol{t}_1$, while $\boldsymbol{x}_4$ is assigned to $\boldsymbol{t}_2$, and the remaining data points are assigned to $\boldsymbol{t}_3$. Specify the cluster assignment matrix $\boldsymbol{W}$, so that $\boldsymbol{X} \approx \boldsymbol{T}\boldsymbol{W}^T$.

b) What is the rank of $\boldsymbol{T}\boldsymbol{W}^T$? Why?

3. You are told that a 3-by-4 matrix $\boldsymbol{X} = \begin{bmatrix} 4 & . & . & . \\ . & -2 & . & . \\ . & . & . & 2 \end{bmatrix}$ has singular-value decomposition $\boldsymbol{X} = \boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^T$ where $\boldsymbol{S} = \begin{bmatrix} 4\sqrt{6} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ and $\boldsymbol{V} = \frac{1}{2}\begin{bmatrix} 1 & -1 & 1 & -1 \\ 1 & 1 & 1 & 1 \\ -1 & -1 & 1 & 1 \\ -1 & 1 & 1 & -1 \end{bmatrix}$. Find the matrix $\boldsymbol{X}$. Be sure to explain how you obtained your answer.

4. Your are given n data points in $\mathbb{R}^7$, and you want to cluster them using k means, $k = 2$, with initial clusters centers $\boldsymbol{t}_1$ and $\boldsymbol{t}_2$. Write pseudo-code for the k means algorithm. Use matrix notation: i.e, approximate $\boldsymbol{X} \approx \boldsymbol{T}_i\boldsymbol{W}_i^T$, where $\boldsymbol{T}_i$ is the matrix of cluster centers on iteration i, and $\boldsymbol{W}_i$ is the cluster assignment matrix.

5. The informative SVD. Consider a data matrix $\boldsymbol{X}$, where the m rows correspond to different training examples and the n columns correspond to different features. Let $\boldsymbol{y}$ be an $m \times 1$ vector with the labels for each example. Suppose the full SVD of $\boldsymbol{X}$ is given by $\boldsymbol{X} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^{\mathsf{T}}$, where:

$$\boldsymbol{U} = \frac{1}{7}\begin{bmatrix} 4 & 2 & 2 & -5 \\ 1 & 4 & 4 & 4 \\ 4 & -5 & 2 & 2 \\ 4 & 2 & -5 & 2 \end{bmatrix} \qquad \boldsymbol{\Sigma} = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad \boldsymbol{V} = \frac{1}{3}\begin{bmatrix} 2 & -2 & 1 \\ 1 & 2 & 2 \\ 2 & 1 & -2 \end{bmatrix}$$

Answer the following questions.

a) For a weight vector $\boldsymbol{w} \in \mathbb{R}^n$, the vector $\boldsymbol{X}\boldsymbol{w}$ is its prediction of the labels. Give a basis for the set of all such prediction vectors.

b) If we restrict the weight vector to satisfy $\|\boldsymbol{w}\|_2 \leq 1$, what is the largest possible prediction $\boldsymbol{X}\boldsymbol{w}$ (as measured in terms of its 2-norm)?

c) Are there weight vectors such that $\boldsymbol{X}\boldsymbol{w} = \boldsymbol{0}$? If so, find a basis for the set of all such vectors.

d) Write an expression for the pseudo-inverse $\boldsymbol{X}^{\dagger}$ satisfying $\boldsymbol{X}^{\dagger}\boldsymbol{X} = \boldsymbol{I}$ (you may leave it in factored form)

e) Suppose $\boldsymbol{y} = \begin{bmatrix} 7 \\ 0 \\ 0 \\ 0 \end{bmatrix}$. Compute the value of $\min_{\boldsymbol{w}} \|\boldsymbol{X}\boldsymbol{w} - \boldsymbol{y}\|_2^2$.