Homework 4

1. Suppose we run the k-means algorithm on the n columns a_j of an $m \times n$ matrix A, to obtain k group representatives $b_1, \ldots, b_k$, and an assignment of each column to a group. As we have seen in ECE 133A, the k-means algorithm tries to minimize an objective

$$J^{\text{clust}} = \frac{1}{n} \sum_{j=1}^{n} \|a_j - b_{c_j}\|^2$$

where c_j is the index of the group that column a_j is assigned to.

   (a) Interpret the k-means algorithm as a low-rank factorization method to solve

$$\begin{array}{ll} \text{minimize} & \|A - BC\|_F^2 \\ \text{subject to} & C_{ij} \in \{0, 1\} \quad i = 1, \ldots, k,\ j = 1, \ldots n \\ & C_{1j} + \cdots + C_{kj} = 1,\ j = 1, \ldots, n. \end{array}$$

   The variables in the problem are the $m \times k$ matrix B and the $k \times n$ matrix C. The constraints on C simply mean that every column of C must be a unit vector (a zero-one vector with exactly one element equal to one).

   (b) To improve the quality of the approximation BC computed by the k-means algorithm, one can re-optimize over C by solving

$$\text{minimize} \quad \|A - BC\|_F^2$$

   with a $k \times n$ matrix variable C. In this problem the matrix B is kept fixed, and has as its columns the k group representatives computed by the k-means algorithm. What is the complexity of solving this optimization problem in the variable C?

   (c) The rank-k approximation computed by the k-means algorithm will be different from the optimal rank-k approximation from a truncated SVD of A (which minimizes the difference $\|A - BC\|_F$). What are possible reasons to prefer the k-means factorization?

2. Suppose A, B are $m \times n$ matrices that satisfy

$$AA^T = BB^T.$$

   We show that $B = AX$ for some orthogonal matrix X.

(a) Explain why A and B have SVDs of the form

$$A = U\Sigma V_1^T, \qquad B = U\Sigma V_2^T.$$

These are full SVDs, *i.e.*, with U, V_1, V_2 square and orthogonal.

(b) Show that $A^T B$ has a polar decomposition

$$A^T B = QH \qquad \text{where } Q = V_1 V_2^T \text{ and } H = V_2 \Sigma^T \Sigma V_2^T.$$

(c) show that $B = AQ$.

3. A undirected graph is *complete* if all pairs of vertices are adjacent.

(a) What is the (unweighted) Laplacian L of the complete graph with n vertices?

(b) What are the eigenvalues of L?