Problem Set 3

CS 7301

Due: 11/29/2021 by 11:59pm

Note: all answers should be accompanied by explanations and code for full credit. Late homeworks will not be accepted.

Problem 1: Missing Entries (50 pts)

Suppose that you are given a symmetric matrix $A \in \mathbb{R}^{n \times n}$ that is missing some entries, e.g., $A_{ij} =?$ for some indices $i, j \in \{1, \ldots, n\}$. To determine which entries are missing, we will use an index matrix $Q \in \{0, 1\}^{n \times n}$ such that $Q_{ij} = 1$ if $A_{ij} =?$ and $Q_{ij} = 0$ otherwise.

1. Explain how to formulate the problem of finding the closest symmetric positive semidefinite matrix to A under the Frobenius norm (over the non-missing entries) as a convex optimization problem.
2. What is the dual of your optimization problem?
3. Write a Python function that takes as input the matrices A and Q, an initial guess for the completion, B, and a number of iterations and returns the the result of applying projected gradient descent, starting at B, for the specified number of iterations.

Problem 2: Matrix Factorizations (50 pts)

1. Consider the following convex function, known as the generalized KL divergence, for two **nonnegative** matrices $A, B \in \mathbb{R}^{m \times n}$.

$$KL(A||B) = \sum_{i}^{m} \sum_{j=1}^{n} \left[A_{ij} \log \left(\frac{A_{ij}}{B_{ij}} \right) - A_{ij} + B_{ij} \right]$$

   Suppose, now that $A \in \mathbb{R}^{m \times n}$ is a nonnegative matrix that we would like to approximate as a product of two nonnegative matrices $C \in \mathbb{R}^{m \times k}, U \in \mathbb{R}^{k \times n}$. Explain how to formulate the problem of finding the closest pair of nonnegative matrices to A under the generalized KL-divergence as a biconvex optimization problem.

2. In Python, write a function that takes as input the matrix A and an integer $k > 0$ and returns a matrix C and U obtained by running block coordinate descent to convergence from a random starting point.
3. Is your block coordinate descent procedure guaranteed to converge to a critical point?