ENMG 616 Advanced Optimization & Techniques

Homework Assignment 5

Due Date: December 03, 2023

You need to show the steps followed to get the final answer (Do not just give the final result). The homework should be submitted to Moodle as ***one*** *pdf file. Please insert a copy of your Matlab code in the submitted file.*

1 Hidden Partition Problem

Suppose we are given a set of vertices $\mathcal{V} = \mathcal{S}_1 \cup \mathcal{S}_2$ such that

$$|\mathcal{S}_1| = |\mathcal{S}_2| \text{ and } \mathcal{S}_1 \cap \mathcal{S}_2 = \emptyset.$$

Also assume that

$$\mathbb{P}\{(i,j) \in \mathcal{E}\} = \begin{cases} p & \text{if } \{i,j\} \in \mathcal{S}_1 \text{ or } \{i,j\} \in \mathcal{S}_2 \\ q & \text{otherwise,} \end{cases}$$

where $q < p$. Hence, two vertices from the same group are connected by an edge with probability p higher than the probability of connecting two vertices from different groups. This can model the problem of clustering the network of people on a social media platform into connected groups. Assume we are given an adjacency matrix $\mathbf{A}$ with

$$A_{i,j} = \begin{cases} 1 & \text{if } (i,j) \text{ belong to the same group} \\ 0 & \text{otherwise} \end{cases}$$

We aim to find the hidden partition in the data by solving the maximum likelihood estimate:

$$\left(\widehat{\mathcal{S}}_1, \widehat{\mathcal{S}}_2\right) = \text{argmax}_{\mathcal{S}_1, \mathcal{S}_2} \mathbb{P}\{\mathbf{A} \,|\, \mathcal{S}_1, \mathcal{S}_2\} \qquad \text{s.t. } |\mathcal{S}_1| = |\mathcal{S}_2|.$$

This problem finds the sets that maximize the probability of having the given adjacency matrix. Assuming that different edges are independent, the problem above can be re-formulated as follows:

$$\begin{aligned} \left(\widehat{\mathcal{S}}_1, \widehat{\mathcal{S}}_2\right) &= \text{argmax}_{|\mathcal{S}_1| = |\mathcal{S}_2|} \mathbb{P}\{\mathbf{A} \,|\, \mathcal{S}_1, \mathcal{S}_2\} \\ &= \text{argmax}_{|\mathcal{S}_1| = |\mathcal{S}_2|} \textstyle\prod_{i,j} \mathbb{P}\{e_{i,j} \,|\, \mathcal{S}_1, \mathcal{S}_2\} \\ &= \text{argmax}_{|\mathcal{S}_1| = |\mathcal{S}_2|} \textstyle\sum_{i,j} \log\left(\mathbb{P}\{e_{i,j} \,|\, \mathcal{S}_1, \mathcal{S}_2\}\right). \end{aligned}$$

Here the last equality holds by taking the log of the function. We know that

$$\log \mathbb{P}\{e_{i,j} \,|\, \mathcal{S}_1, \mathcal{S}_2\} = \begin{cases} \log(p) & \text{if } (i,j) \text{ belong to the same partition and } A_{i,j} = 1 \\ \log(1-p) & \text{if } (i,j) \text{ belong to the same partition and } A_{i,j} = 0 \\ \log(q) & \text{if } (i,j) \text{ belong to the different partition and } A_{i,j} = 1 \\ \log(1-q) & \text{if } (i,j) \text{ belong to the different partition and } A_{i,j} = 0. \end{cases}$$

Hence, the optimization problem can written as

$$\begin{aligned}
\left(\widehat{\mathcal{S}}_1, \widehat{\mathcal{S}}_2\right) &= \text{argmax}_{|\mathcal{S}_1|=|\mathcal{S}_2|} \sum_{i,j:\text{ same partition}} \left((1-A_{i,j})\log(1-p) + A_{i,j}\log(p)\right) \\
&+ \sum_{i,j:\text{different partition}} \left((1-A_{i,j})\log(1-q) + A_{i,j}\log(q)\right) \\
&= \text{argmax}_{|\mathcal{S}_1|=|\mathcal{S}_2|} \sum_{i,j:\text{ same partition}} \left(-A_{i,j}\log(1-p) + A_{i,j}\log(p)\right) \\
&+ \sum_{i,j:\text{different partition}} \left(-A_{i,j}\log(1-q) + A_{i,j}\log(q)\right) \\
&= \text{argmax}_{|\mathcal{S}_1|=|\mathcal{S}_2|} \sum_{i,j:\text{ same partition}} A_{i,j}\log\left(\frac{p}{1-p}\right) \\
&+ \sum_{i,j:\text{different partition}} A_{i,j}\log\left(\frac{q}{1-q}\right) \\
&= \text{argmax}_{|\mathcal{S}_1|=|\mathcal{S}_2|} \sum_{i,j:\text{ same partition}} A_{i,j}\log\left(\frac{p}{1-p}\right) - A_{i,j}\log\left(\frac{q}{1-q}\right) \\
&+ \sum_{i,j} A_{i,j}\log\left(\frac{q}{1-q}\right) \\
&= \text{argmax}_{|\mathcal{S}_1|=|\mathcal{S}_2|} \sum_{i,j:\text{ same partition}} A_{i,j}
\end{aligned}$$

1.1 Re-formulating the Optimization Problem

1. Let

$$x_i = \begin{cases} 1 & \text{if } i \in \mathcal{S}_1 \\ -1 & \text{if } i \in \mathcal{S}_2. \end{cases}$$

Show that the problem defined above is equivalent to the following optimization problem

$$\text{argmax}_x \sum_{i,j} A_{i,j}(1 + x_i x_j) \qquad \text{s.t. } \sum_{i=1}^{n} x_i = 0, \quad x_i \in \{-1, +1\}.$$

Allowing for a bit of imbalance in x_i, we pass the constraint to the objective. The problem becomes

$$\text{argmax}_x \mathbf{x}^T \mathbf{A} \mathbf{x} - \lambda \left(\sum_i x_i \right)^2 \qquad \text{s.t.} \qquad x_i \in \{-1, +1\}.$$

By applying a change of variable, $\mathbf{X} = \mathbf{x}\mathbf{x}^T$ and letting $\widetilde{A} = \lambda \mathbf{1}\mathbf{1}^T - \mathbf{A}$, we obtain

$$\begin{aligned} \mathbf{x}^T \mathbf{A} \mathbf{x} - \lambda \left(\textstyle\sum_i x_i \right)^2 &= \mathbf{x}^T \mathbf{A} \mathbf{x} - \lambda \mathbf{x}^T \mathbf{1}\mathbf{1}\mathbf{x} \\ &= \text{Tr}(\mathbf{A}\mathbf{x}\mathbf{x}^T) - \lambda \text{Tr}(\mathbf{1}\mathbf{1}^T \mathbf{x}\mathbf{x}^T) \\ &= -\langle \mathbf{X}, \widetilde{\mathbf{A}} \rangle. \end{aligned}$$

2. Show that the problem can be re-formulated as follows

$$\text{argmin}_{\mathbf{X}} \langle \mathbf{X}, \widetilde{\mathbf{A}} \rangle \qquad \text{s.t. } \mathbf{X} \succeq 0, \quad \text{rank}(\mathbf{X}) \leq 1, \quad \mathbf{X}_{ii} = 1$$

3. Suppose that we want to solve the optimization problem above using ADMM. Explain in few sentences that the above problem can be reformulated as follows:

$$\min_{\mathbf{X}, \mathbf{Z}} \langle \mathbf{X}, \widetilde{\mathbf{A}} \rangle \qquad \text{s.t } \mathbf{X} \succeq 0, \quad Z_{i,i} = 1, \quad \mathbf{Z} = \mathbf{X}.$$

4. Write the ADMM update rules for the problem above.

1.2 Min-Max Problem

Consider the following bi-linear min-max optimization problem

$$\min_{\mathbf{x}} \max_{\mathbf{y}} \mathbf{x}^T \mathbf{A} \mathbf{y}.$$

In this problem, we will generate a random matrix $\mathbf{A}$ and solve the above problem using three different methods.

1. Randomly generate the matrix $\mathbf{A}$ using the following procedure:
   (a) Set $n = 10$
   (b) Generate a random matrix $\mathbf{B} \in \mathbb{R}^{n \times n}$ using the following Matlab code: $\mathbf{B} = \text{randn}(n, n)$.
   (c) Let $\mathbf{A} = \mathbf{B} \times \mathbf{B}'$.
2. Implement EG, OGDA and PP methods to solve the above problem.
3. Run the above three algorithms and plot the following error term

$$\mathcal{E}(\mathbf{x}^k, \mathbf{y}^k) = \|\mathbf{x}^k\|^2 + \|\mathbf{y}^k\|^2$$

   versus iteration number r.
4. Repeat the above experiment when $\mathbf{A}$ is generated as follows:
   (a) Set $n = 10$
   (b) Generate a random matrix $\mathbf{B} \in \mathbb{R}^{n \times n}$ using the following Matlab code: $\mathbf{B} = \text{randn}(n, n)$.
   (c) Let $\mathbf{A} = \mathbf{B} \times \mathbf{B}'$.
   (d) $\mathbf{A} = \mathbf{A} + 10 * eye(n)$.