CS 395T , Homework 3.

For the optimization problem:

$$\min_{B,C} \sum_{i=1}^{n}\sum_{j=1}^{n} g_{ij}(a_{ij} - (e_i^T B)(C e_j))^2 + \frac{\mu}{2}(\|B\|_F^2 + \|C\|_F^2).$$

Which is equivalent to

$$\min_{B,C} \|P_\Omega(A - BC^T)\|_F^2 + \frac{\mu}{2}(\|B\|_F^2 + \|C\|_F^2). \quad (*)$$

Where, Ω is the set $g_{ij} = 1$, and P_Ω is the projection operater to the set Ω.

(1) Solving $(*)$ by Alternating Minimization, gives

$$2(P_\Omega(A - BC^T))(-C) + \mu \cdot B = 0.$$

$$2(P_\Omega(A - BC^T)) \cdot (-B) + \mu \cdot C = 0,$$

(2). We perform Taylor expansion for the function defined above:

For each row and column in B, C, we define b_i and c_j

$$\frac{\partial F}{\partial b_i} = (\sum_j 2 g_{ij} c_j c_j^T + \mu I) b_i - \sum_j 2 g_{ij} a_{ij} c_j$$

$$\frac{\partial F}{\partial c_j} = (\sum_i 2 g_{ij} b_i b_i^T + \mu I) c_j - \sum_i 2 g_{ij} a_{ij} b_i.$$

Hence, $$\frac{\partial^2 F}{\partial b_i \partial b_j} = \sum_j 2 g_{ij} c_j c_j^T + \mu I \quad (i = j)$$

$$\frac{\partial^2 F}{\partial b_i \partial c_j} = \cancel{2 g_{ij} c_j c_j^T}$$

$$(-2 g_{ij} a_{ij} + 2 g_{ij} c_j^T b_i) I + b_i c_j^T$$

Problem 3:

the necessary condition & sufficient condition for p^* being an optimal sollution of

$$\min_{p} \quad g^T p + \frac{1}{2} p^T B p \qquad B \text{ is symmetric}$$
$$\text{s.t.} \quad p^T A p \leq d^2. \quad , \qquad A \text{ is psd.}$$

This is the extended formulation of Trust Region Sub-problem, with I being replaced by A. Recall the condition in Trust Region Sub-problem.

$$\begin{cases} (B+\lambda I) p^* = -g \\ \lambda(\Delta - \|p^*\|) = 0 \\ (B+\lambda I) \text{ is psd.} \end{cases}$$

We use the similar approach to prove this for general A.

Assume first that there exists $\lambda \geq 0$ such that

$$\begin{cases} (B+\lambda A) p^* = -g \\ \lambda(d - \|P^*\|_A) = 0 \\ B+\lambda A \succeq 0 . \end{cases}$$

By lemma 3.1 in Lecture notes, this implies p^* is the global minimum of the quadratic function

$$\hat{m}(p) = g^T p + \frac{1}{2} p^T (B+\lambda A) p .$$

Then, we consider $m(p) = g^T p + \frac{1}{2} p^T B p$,

Since $\hat{m}(p) \geq \hat{m}(p^*)$, we have

$$m(p) - m(p^*) \geq \frac{1}{2} \lambda (\|p^*\|_A^2 - \|p\|_A^2)$$

plugging in the fact $\lambda(d^2 - \|p^*\|_A^2) = 0$, we have $m(p) \geq m(p^*)$ for $\|p\|_A \leq d$.

For the converse, we assume that p^* is a global solution of the original optimization.

We consider the Lagrangian function defined by

$$L(p,\lambda) = g^T p + \frac{1}{2} p^T (B + \lambda A) p - \frac{1}{2}\lambda d^2.$$

This has a stationary point at p^*.

By setting $\nabla_p L(p^*, \lambda) = 0$, we have $(B+\lambda A)p^* = -g$.

By setting $\nabla_\lambda L(p, \lambda) = 0$, we have $\|p^*\|_A^2 = d^2$ or $\lambda = 0$.

$$g^T p + \frac{1}{2} p^T (B+\lambda A) p \geq g^T p^* + \frac{1}{2} p^{*T} (B+\lambda A) p^*, \quad \|p\|_A^2 = d^2,$$

If we substitute $p^* = -(B+\lambda A)^{-1} g$, we obtain after some arrangement that

$$\frac{1}{2}(p-p^*)(B+\lambda A)(p-p^*) \geq 0.$$

Since the set $p-p^*$ is dense, we have $B+\lambda A \geq 0$.