CS395T: Numerical Optimization for Graphics and AI: Homework III

1 Guideline

- Please complete **3** problems out of **6** problems, and please complete at least one problem in the theory session.
- You are welcome to complete more problems.

2 Programming

Each problem in this section counts as two.
Problem 1 and Problem 2. We are interested in solving the following low-rank matrix problem. Given a sparse observation pattern $G = (g_{ij})_{1\leq i,j\leq n} \in \{0,1\}^{n\times n}$ and a data matrix $A = (a_{ij})_{1\leq i,j\leq n} \in \mathbb{R}^{n\times n}$, our goal is to recover a low-rank matrix pair $B \in \mathbb{R}^{n\times r}, C \in \mathbb{R}^{r\times n}$ by minimizing

$$\underset{B,C}{\text{minimize}} \ \sum_{i=1}^{n}\sum_{j=1}^{n} g_{ij}(a_{ij} - (\boldsymbol{e}_i^T B)(C\boldsymbol{e}_j))^2 + \frac{\mu}{2}(\|B\|_{\mathcal{F}}^2 + \|C\|_{\mathcal{F}}^2),$$

where μ is a small constant, which makes the optimization problem non-degenerate. Implement a trust-region method for solving this problem, and compare it against alternating minimization of B and C.

3 Theory

Problem 3. Derive necessary and sufficient conditions for $\boldsymbol{p}^\star$ being an optimal solution to the following optimization problem:

$$\begin{aligned} \underset{\boldsymbol{p}}{\text{minimize}} \quad & \boldsymbol{g}^T\boldsymbol{p} + \frac{1}{2}\boldsymbol{p}^T B\boldsymbol{p} \\ \text{subject to} \quad & \boldsymbol{p}^T A\boldsymbol{p} \leq d^2, \end{aligned} \tag{1}$$

where B is symmetric and A is positive semidefintie (not necessarily positive definite).

Problem 4. Derive a close-form solution for the 2D trust-region problem under the L1-norm:

$$\begin{aligned} \underset{x,y\in\mathbb{R}}{\text{minimize}} \quad & \begin{pmatrix} x & y \end{pmatrix}\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} x & y \end{pmatrix}\begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} \\ \text{subject to} \quad & |x| + |y| \leq d. \end{aligned} \tag{2}$$

Your solution should be as a close-form of $b_1, b_2, a_{11}, a_{12}, a_{22}, d$.

Problem 5. Derive an optimally condition for the following trust region sub-problem for optimizing rotations:

$$\begin{aligned} \underset{R \in SO(m)}{\text{minimize}} \quad & \text{Trace}(C^T R) \\ \text{subject to} \quad & \|R_0 - R\|_{\mathcal{F}}^2 \le d^2. \end{aligned} \tag{3}$$

Here $R_0 \in SO(m)$ and $C \in \mathbb{R}^{m \times m}$ is a constant matrix.

Problem 6. We consider solving the following "sub-problem" using conjugate gradient descent:

$$\begin{aligned} \underset{\boldsymbol{p}}{\text{minimize}} \quad & \phi(\boldsymbol{p}) := \boldsymbol{g}^T \boldsymbol{p} + \frac{1}{2} \boldsymbol{p}^T B \boldsymbol{p} \\ \text{subject to} \quad & \boldsymbol{p}^T C \boldsymbol{p} \le d^2. \end{aligned} \tag{4}$$

Here B is a real symmetric matrix and C is positive definite matrix.

The algorithm we consider is described as follows:

- **Step I:** Set $\boldsymbol{p}_0 = 0, \boldsymbol{r}_0 = -\boldsymbol{g}$. Solve $C\hat{\boldsymbol{r}_0} = \boldsymbol{r}_0$. Set $\boldsymbol{d}_0 = \hat{\boldsymbol{r}}_0, i = 0$.
- **Step II:** Compute $\gamma_i = \boldsymbol{d}_i^T B \boldsymbol{d}_i$. If $\gamma_i > 0$ then continue with Step III. Otherwise compute $\tau > 0$ so that $(\boldsymbol{p}_i + \tau \boldsymbol{d}_i)^T C (\boldsymbol{p}_i + \tau \boldsymbol{d}_i) = d^2$, set $\boldsymbol{p} = \boldsymbol{p}_i + \tau \boldsymbol{d}_i$ and terminate.
- **Step III:** Compute $\alpha_i = \frac{\boldsymbol{r}_i^T \hat{\boldsymbol{r}}_i}{\gamma_i}$, $\boldsymbol{p}_{i+1} = \boldsymbol{p}_i + \alpha_i \boldsymbol{d}_i$. If $\boldsymbol{p}_{i+1}^T C \boldsymbol{p}_{i+1} < d^2$ than Continue with step IV. Otherwise compute $\tau > 0$ so that $(\boldsymbol{p}_i + \tau \boldsymbol{d}_i)^T C (\boldsymbol{p}_i + \tau \boldsymbol{d}_i) = d^2$, set $\boldsymbol{p} = \boldsymbol{p}_i + \tau \boldsymbol{d}_i$ and terminate.
- **Step IV:** Compute $\boldsymbol{r}_{i+1} = \boldsymbol{r}_i - \alpha_i B \boldsymbol{d}_i$. If $\frac{\boldsymbol{r}_{i+1}^T C \boldsymbol{r}_{i+1}}{\boldsymbol{g}^T C \boldsymbol{g}} \le \epsilon^2$ then set $\boldsymbol{p} = \boldsymbol{p}_{i+1}$ and terminate. Otherwise continue with Step V.
- **Step V:** Solve $C\hat{\boldsymbol{r}_{i+1}} = \boldsymbol{r}_{i+1}$. Compute $\beta_i = \frac{\boldsymbol{r}_{i+1}^T \hat{\boldsymbol{r}}_{i+1}}{\boldsymbol{r}_i^T \hat{\boldsymbol{r}}_i}$ and $\boldsymbol{d}_{i+1} = \hat{\boldsymbol{r}}_{i+1} + \beta_i \boldsymbol{d}_i$. Set $i := i + 1$ and continue with step II.

Let $\boldsymbol{p}_j, j = 0, \ddots, i$ be the iterates generated by the algorithm described above.

- Show that $\phi(\boldsymbol{p}_j)$ is strictly decreasing and $\phi(\boldsymbol{p}) \le \phi(\boldsymbol{p}_i)$.
- Show that $\boldsymbol{p}_j^T C \boldsymbol{p}_j$ is strictly increasing for $j = 0, \ddots, i$ and $\boldsymbol{p}^T C \boldsymbol{p} \ge \boldsymbol{p}_j^T C \boldsymbol{p}_j$.

Problem 1 and Problem 2. We are interested in solving the following low-rank matrix problem. Given a sparse observation pattern $G = (g_{ij})_{1\le i,j\le n} \in \{0,1\}^{n\times n}$ and a data matrix $A = (a_{ij})_{1\le i,j\le n} \in \mathbb{R}^{n\times n}$, our goal is to recover a low-rank matrix pair $B \in \mathbb{R}^{n\times r}, C \in \mathbb{R}^{r\times n}$ by minimizing

$$\underset{B,C}{\text{minimize}} \sum_{i=1}^{n}\sum_{j=1}^{n} g_{ij}(a_{ij} - (\boldsymbol{e}_i^T B)(C\boldsymbol{e}_j))^2 + \frac{\mu}{2}(\|B\|_{\mathcal{F}}^2 + \|C\|_{\mathcal{F}}^2),$$

where μ is a small constant, which makes the optimization problem non-degenerate. Implement a trust-region method for solving this problem, and compare it against alternating minimization of B and C.

Solution

a. Alternating Minimization

$$\min_{B,C} \sum_{i,j} g_{ij}(a_{ij} - \boldsymbol{b}_i^T\boldsymbol{c}_j)^2 + \frac{\mu}{2}(\sum_i \|\boldsymbol{b}_i\|_2^2 + \sum_j \|\boldsymbol{c}_j\|_2^2)$$

$$\frac{\partial F}{\partial \boldsymbol{b}_i} = \sum_j 2g_{ij}(\boldsymbol{c}_j^T\boldsymbol{b}_i - a_{ij})\boldsymbol{c}_j + \mu\boldsymbol{b}_i = [\mu I + \sum_j 2g_{ij}\boldsymbol{c}_j\boldsymbol{c}_j^T]\boldsymbol{b}_i - (\sum_j 2g_{ij}a_{ij}\boldsymbol{c}_j)$$

$$\frac{\partial F}{\partial \boldsymbol{c}_j} = [\mu I + \sum_j 2g_{ij}\boldsymbol{b}_i\boldsymbol{b}_i^T]\boldsymbol{c}_j - (\sum_j 2g_{ij}a_{ij}\boldsymbol{b}_i)$$

For proper μ, we can easily solve $\frac{\partial F}{\partial \boldsymbol{b}_i} = \mathbf{0}$ and $\frac{\partial F}{\partial \boldsymbol{c}_j} = \mathbf{0}$.

b. Trust-Region

$$\frac{\partial^2 F}{\partial \boldsymbol{b}_i \partial \boldsymbol{b}_j} = \begin{cases} 0 & \text{if } i \neq j \\ \mu I + 2\sum_k g_{ik}\boldsymbol{c}_k\boldsymbol{c}_k^T & \text{else} \end{cases}$$

$$\frac{\partial^2 F}{\partial \boldsymbol{c}_i \partial \boldsymbol{c}_j} = \begin{cases} 0 & \text{if } i \neq j \\ \mu I + 2\sum_k g_{ik}\boldsymbol{b}_k\boldsymbol{b}_k^T & \text{else} \end{cases}$$

$$\frac{\partial^2 F}{\partial \boldsymbol{b}_i \partial \boldsymbol{c}_j} = 2g_{ij}[(\boldsymbol{c}_j^T\boldsymbol{b}_i - a_{ij})I + \boldsymbol{b}_i\boldsymbol{c}_j^T]$$

Problem 3. Derive necessary and sufficient conditions for $\boldsymbol{p}^\star$ being an optimal solution to the following optimization problem:

$$\begin{aligned} \underset{\boldsymbol{p}}{\text{minimize}} \quad & \boldsymbol{g}^T\boldsymbol{p} + \frac{1}{2}\boldsymbol{p}^T B\boldsymbol{p} \\ \text{subject to} \quad & \boldsymbol{p}^T A\boldsymbol{p} \le d^2, \end{aligned} \tag{5}$$

where B is symmetric and A is positive semidefintie (not necessarily positive definite).

Solution

We claim a similar theorem compared to Theorem 3.1 in Lecture 7.

Theorem 3.1. *Vector $\boldsymbol{p}^\star$ is an optimal solution iff $\boldsymbol{p}^\star$ is feasible and there is a scalar $\lambda \ge 0$ such that the following conditions are satisfied:*

$$(B + \lambda A)\boldsymbol{p}^\star = -\boldsymbol{g} \tag{6}$$

$$\lambda(d - \|\boldsymbol{p}^\star\|_A) = 0 \tag{7}$$

$$B + \lambda A \succeq 0 \tag{8}$$

where $\|\boldsymbol{p}\|_A = \sqrt{\boldsymbol{p}^T A\boldsymbol{p}}$.

Proof. Recall from Lecture 7 the conclusions we have for the unconstrained problem. We will use the following lemma without proving it again.

Lemma 3.1. *Let $m(\boldsymbol{p})$ be the quadratic function defined by*

$$m(\boldsymbol{p}) = \boldsymbol{g}^T\boldsymbol{p} + \frac{1}{2}\boldsymbol{p}^T B\boldsymbol{p}$$

where B is symmetric. Then the following statements are true.

- *m attains a minimum iff $B \succeq 0$ and $\exists \boldsymbol{v}, s.t.\ \boldsymbol{g} = B\boldsymbol{v}$. If $B \succeq 0$, then every $\boldsymbol{p}$ satisfying $B\boldsymbol{p} = -\boldsymbol{g}$ is a global minimizer of m.*
- *m has a unique minimizer iff $B \succ 0$.*

Now we prove that the conditions we derived are necessary and sufficient. Define

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

- **(Sufficiency)** If $\exists \lambda \geq 0, \boldsymbol{p}^\star$ such that (6) (7) (8) are satisfied, then from (6) (8)

  $$\nabla \hat{m}(\boldsymbol{p}^\star) = \boldsymbol{g} + (B+\lambda A)\boldsymbol{p}^\star = \boldsymbol{0}$$

  and $\hat{m}(\boldsymbol{p})$ is convex. This implies $\boldsymbol{p}^\star$ is a global minimizer of $\hat{m}(\boldsymbol{p})$.

  For any $\boldsymbol{p}$ where $\|\boldsymbol{p}\|_A \leq d$,

  $$\begin{aligned} m(\boldsymbol{p}) - m(\boldsymbol{p}^\star) &= \hat{m}(\boldsymbol{p}) - \hat{m}(\boldsymbol{p}^\star) - \frac{1}{2}\lambda(\|\boldsymbol{p}\|_A^2 - \|\boldsymbol{p}^\star\|_A^2) \\ &\geq -\frac{1}{2}\lambda(\|\boldsymbol{p}\|_A^2 - \|\boldsymbol{p}^\star\|_A^2) \\ &\geq -\frac{1}{2}\lambda(d^2 - \|\boldsymbol{p}^\star\|_A^2) \qquad (\|\boldsymbol{p}\|_A \leq d) \\ &= 0 \qquad \text{(Use (7))} \end{aligned}$$

  This means $\boldsymbol{p}^\star$ is a global minimizer of the original problem.

- **(Necessity)** If $\boldsymbol{p}^\star$ is a global optimizer of the original problem, then it must also be a stationary point of the following min-max problem.

  $$\min_{\boldsymbol{p}} \max_{\lambda \geq 0} \mathcal{L}(\boldsymbol{p}, \lambda) := \boldsymbol{g}^T\boldsymbol{p} + \frac{1}{2}\boldsymbol{p}^T(B+\lambda A)\boldsymbol{p} - \frac{1}{2}\lambda d^2$$

  Let $(\boldsymbol{p}^\star, \lambda^\star)$ be a stationary point, it must satisfy

  $$\begin{aligned} &\frac{\partial \mathcal{L}}{\partial \boldsymbol{p}}(\boldsymbol{p}^\star) = \boldsymbol{g} + (B+\lambda^\star A)\boldsymbol{p}^\star = \boldsymbol{0} \\ &\frac{\partial \mathcal{L}}{\partial \lambda}(\lambda^\star) = \frac{1}{2}(\|\boldsymbol{p}^\star\|_A^2 - d^2) = 0 \qquad \text{or} \qquad \{\lambda^\star = 0, \|\boldsymbol{p}^\star\|_A < d\} \end{aligned}$$

  which means $(\boldsymbol{p}^\star, \lambda^\star)$ satisfies (6) and (7).

  If $\lambda^\star = 0$ and $\|\boldsymbol{p}^\star\|_A < d$, then $\boldsymbol{p}^\star$ is an unconstrained minimizer of $m(\boldsymbol{p})$ and consequently

  $$\nabla m(\boldsymbol{p}^\star) = B\boldsymbol{p}^\star + \boldsymbol{g} = \boldsymbol{0}, \qquad \nabla^2 m(\boldsymbol{p}^\star) = B \succeq 0$$

  which means (8) holds for $\lambda^\star$.

  Now assume $\|\boldsymbol{p}^\star\|_A = d$, then $\boldsymbol{p}^\star$ is a minimizer of the following constrained problem

  $$\min_{\boldsymbol{p}} \ \boldsymbol{g}^T\boldsymbol{p} + \frac{1}{2}\boldsymbol{p}^T(B+\lambda^\star A)\boldsymbol{p}, \qquad s.t.\ \|\boldsymbol{p}\|_A = d$$

that means for any $\boldsymbol{p}$ where $\|\boldsymbol{p}\|_A = d$

$$\boldsymbol{g}^T\boldsymbol{p} + \frac{1}{2}\boldsymbol{p}^T(B + \lambda^\star A)\boldsymbol{p} \geq \boldsymbol{g}^T\boldsymbol{p}^\star + \frac{1}{2}(\boldsymbol{p}^\star)^T(B + \lambda^\star A)\boldsymbol{p}^\star$$

use (6), $\|\boldsymbol{p}\|_A = d$ and rearrange, we get

$$\begin{aligned}
&\frac{1}{2}(\boldsymbol{p}^\star)^T(B + \lambda^\star A)\boldsymbol{p}^\star - (\boldsymbol{p}^\star)^T(B + \lambda^\star A)^T\boldsymbol{p} + \frac{1}{2}\boldsymbol{p}^T(B + \lambda^\star A)\boldsymbol{p} \geq 0 \\
\Rightarrow &\frac{1}{2}(\boldsymbol{p} - \boldsymbol{p}^\star)^T(B + \lambda^\star A)(\boldsymbol{p} - \boldsymbol{p}^\star) \geq 0
\end{aligned}$$

One can verify that any direction $\boldsymbol{v}$ where $\boldsymbol{v}^T A\boldsymbol{p}^\star \neq 0$ can be represented as $\boldsymbol{p} - \boldsymbol{p}^\star$ for some $\boldsymbol{p}$ that $\|\boldsymbol{p}\|_A = d$. i.e. assume

$$\boldsymbol{p}^\star - \boldsymbol{p} = \lambda\boldsymbol{v}, \lambda \neq 0, \|\boldsymbol{v}\|_A \neq 0$$

where $\|\boldsymbol{v}\|_A \neq 0$ is a necessary condition for $\boldsymbol{v}^T A\boldsymbol{p}^\star \neq 0$.

$$\begin{aligned}
&\|\boldsymbol{p}\|_A^2 = d^2 \\
\Leftrightarrow &\|\boldsymbol{p}^\star - \lambda\boldsymbol{v}\|_A^2 = d^2 \\
\Leftrightarrow &\|\boldsymbol{p}^\star\|_A^2 + \|\lambda\boldsymbol{v}\|_A^2 - 2\lambda\boldsymbol{v}^T A\boldsymbol{p}^\star = d^2 \\
\Leftrightarrow &\lambda^2\boldsymbol{v}^T A\boldsymbol{v} = 2\lambda\boldsymbol{v}^T A\boldsymbol{p}^\star \\
\Leftrightarrow &\lambda = \frac{2\lambda\boldsymbol{v}^T A\boldsymbol{p}^\star}{\boldsymbol{v}^T A\boldsymbol{v}}
\end{aligned}$$

Then for those $\boldsymbol{v}$ where $\boldsymbol{v}^T A\boldsymbol{p}^\star = 0$, we can use continuity to show that $\boldsymbol{v}^T(B+\lambda A)\boldsymbol{v} \geq 0$. This implies $B + \lambda A \succeq 0$, which completes our proof.

$\square$

Problem 4. Derive a close-form solution for the 2D trust-region problem under the L1-norm:

$$\begin{aligned}
\underset{x,y\in\mathbb{R}}{\text{minimize}} \quad & \begin{pmatrix} x & y \end{pmatrix}\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} x & y \end{pmatrix}\begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} \\
\text{subject to} \quad & |x| + |y| \leq d.
\end{aligned} \tag{9}$$

Your solution should be as a close-form of $b_1, b_2, a_{11}, a_{12}, a_{22}, d$.

Solution

We rewrite this problem as

$$\min_{\boldsymbol{x}} \boldsymbol{b}^T\boldsymbol{x} + \frac{1}{2}\boldsymbol{x}^T A\boldsymbol{x}, \quad s.t. \|\boldsymbol{x}\|_1 \leq d$$

a. If $A \succ 0$, first compute $\hat{\boldsymbol{x}} = -A^{-1}\boldsymbol{b}$, if $\|\hat{\boldsymbol{x}}\|_1 \leq d$, then we found the optimal solution. Otherwise, we have that $\|\boldsymbol{x}^\star\|_1 = d$. Without loss of generality, we assume $-A^{-1}\boldsymbol{b}$ has non-negative coordinates (if not, we can change A and $\boldsymbol{b}$ accordingly without hurting the condition that $A \succ 0$). In this case, the optimal solution lies on segment $S = \{\boldsymbol{x} | x_1 + x_2 = d, x_1, x_2 > 0\}$.

   Suppose we constrain our search domain on line $L = \{\boldsymbol{x} | x_1 + x_2 = d\}$. Then we know the gradient of objective function must be perpendicular to line L, i.e. our optimal solution $\boldsymbol{x}^\star$ must satisfy

   $$\exists\alpha, \quad A\boldsymbol{x}^\star + b = A(\boldsymbol{x}^\star - \hat{\boldsymbol{x}}) = \alpha\mathbf{1}, \quad x_1^\star + x_2^\star = d$$

   This means $\boldsymbol{x}^\star = \hat{\boldsymbol{x}} + \alpha A^{-1}\mathbf{1}$ and $\alpha = \frac{d - \hat{x}_1 - \hat{x}_2}{\mathbf{1}^T A^{-1}\mathbf{1}}$. If our computed $\boldsymbol{x}^\star$ does not lie on segment S, we can truncate $\boldsymbol{x}^\star$ to the nearest end point, i.e. $(0, d)$ or $(d, 0)$.

b. If $A \prec 0$, the minimum is always achieved on one of the end points $\{(0, d), (d, 0), (0, -d), (-d, 0)\}$. Just compute all of them and find the minimum.

c. If A has exactly one eigenvalue that equals to zero, then eigen-decomposition $A = \lambda \boldsymbol{v}\boldsymbol{v}^T + 0\boldsymbol{u}\boldsymbol{u}^T$ where $\|\boldsymbol{v}\|_2 = \|\boldsymbol{u}\|_2 = 1$ and $\lambda \neq 0$.

   Decompose $\boldsymbol{x} = \alpha_v \boldsymbol{v} + \alpha_u \boldsymbol{u}$ and $\boldsymbol{b} = \beta_v \boldsymbol{v} + \beta_u \boldsymbol{u}$. If $\beta_u = 0$,

   c.1 for $\lambda > 0$, we check if there exists α_u such that $\boldsymbol{x} = \frac{-\beta_v}{\lambda}\boldsymbol{v} + \alpha_u \boldsymbol{u}$ lies in the interior of the l_1 ball. If it is the case, we have found our optimal solution. Otherwise, go to the next sub-case.

   c.2 If $\lambda < 0$, or $\lambda > 0$ but optimal point does not lie in interior. Optimal solution is achieved on one of the end points, we can just enumerate them.

   For the rest of this case we assume $\beta_u \neq 0$.

   Note that

$$\nabla_x f = A\boldsymbol{x} + \boldsymbol{b} = (\lambda \alpha_v + \beta_v)\boldsymbol{v} + \beta_u \boldsymbol{u}$$

   Since gradient is always non-zero, we must end up on boundary. Now solve for potential optimal solutions on the segments. Let $\boldsymbol{d}_1 = (1, -1)^T$ and $\boldsymbol{d}_2 = (1, 1)^T$. Equation

$$\nabla_x f^T \boldsymbol{d}_1 = 0$$

   or

$$\nabla_x f^T \boldsymbol{d}_2 = 0$$

   could give solutions for α_v. Then we can easily verify if there exists α_u such that $\boldsymbol{x}$ lies on segments. We also enumerate all four end points take whatever minimum we got. Then we claim that the optimal solution must be examined by the procedure above.

d. If A has two eigenvalues $\lambda_1 > 0 > \lambda_2$, then eigen-decomposition gives $A = \lambda_1 \boldsymbol{v}_1 \boldsymbol{v}_1^T + \lambda_2 \boldsymbol{v}_2 \boldsymbol{v}_2^T$. Let $\boldsymbol{x} = \alpha_1 \boldsymbol{v}_1 + \alpha_2 \boldsymbol{v}_2$ and $\boldsymbol{b} = \beta_1 \boldsymbol{v}_1 + \beta_2 \boldsymbol{v}_2$. The objective function becomes

$$\frac{1}{2}(\lambda_1 \alpha_1^2 + \lambda_2 \alpha_2^2) + (\alpha_1 \beta_1 + \alpha_2 \beta_2)$$

   Since hessian is not positive semi-definite, we must end up on the boundary.

$$\nabla_x f = \begin{pmatrix} \lambda_1 \alpha_1 + \beta_1 \\ \lambda_2 \alpha_2 + \beta_2 \end{pmatrix}$$

   again we can solve equations like

$$\left\{ \begin{array}{c} \nabla_x f^T \boldsymbol{d}_1 = 0 \\ x_1 - x_2 = \pm d \end{array} \right. \quad \text{or} \left\{ \begin{array}{c} \nabla_x f^T \boldsymbol{d}_2 = 0 \\ x_1 + x_2 = \pm d \end{array} \right. \qquad \boldsymbol{d}_1 = (1, -1)^T, \boldsymbol{d}_2 = (1, -1)^T$$

   and then we also examine end points and take the minimum.

Problem 5. Derive an optimally condition for the following trust region sub-problem for optimizing rotations:

$$\begin{aligned} \underset{R \in SO(m)}{\text{minimize}} \quad & \text{Trace}(C^T R) \\ \text{subject to} \quad & \|R_0 - R\|_{\mathcal{F}}^2 \leq d^2. \end{aligned} \tag{10}$$

Here $R_0 \in SO(m)$ and $C \in \mathbb{R}^{m \times m}$ is a constant matrix.

Solution

$$\begin{aligned} & \text{Tr}((R_0 - R)^T (R_0 - R)) \leq d^2 \\ \Leftrightarrow & \text{Tr}(-2R_0^T R) + 2m \leq d^2 \\ \Leftrightarrow & \text{Tr}(R_0^T R) \geq m - \frac{d^2}{2} \end{aligned}$$

Take $\bar{R} = R_0^T R \in SO(m)$ and $\bar{C} = -R_0^T C$, the problem is equivalent to

$$\max_{\bar{R}\in SO(m)} \quad \mathrm{Tr}(\bar{C}^T \bar{R})$$
$$\text{subject to } \mathrm{Tr}(\bar{R}) \geq m - \frac{d^2}{2}$$

Or

$$\max_{\bar{R}\in SO(m)} \min_{\lambda \geq 0} \quad \mathcal{L}(\bar{R}, \lambda) := \mathrm{Tr}((\bar{C} + \lambda I)^T \bar{R}) - \lambda(m - \frac{d^2}{2})$$

If we write down svd decomposition $(\bar{C} + \lambda I)^T = U_\lambda \Sigma V_\lambda^T$ and assume Σ_{mm} is the smallest singular value.

From lecture notes, we know the optimal solution of

$$\max_{\bar{R}\in SO(m)} \mathrm{Tr}(U_\lambda \Sigma V_\lambda^T \bar{R})$$

is $\bar{R} = V_\lambda D_\lambda U_\lambda^T$ where $D_\lambda = diag(1, 1, \ldots, 1, det(U_\lambda V_\lambda^T))$.

If $\bar{R}^\star$ is an optimal solution, there must exists $\lambda^\star$ such that $(\bar{R}^\star, \lambda^\star)$ is a stationary point of the max-min problem, i.e.

1. $\bar{R}^\star = V_{\lambda^\star} D_{\lambda^\star} U_{\lambda^\star}^T$
2. either $\begin{cases} \lambda^\star = 0 \\ \mathrm{Tr}(\bar{R}^\star) \geq m - \frac{d^2}{2} \end{cases}$ or $\begin{cases} \lambda^\star > 0 \\ \mathrm{Tr}(\bar{R}^\star) = m - \frac{d^2}{2} \end{cases}$

Now suppose the above conditions holds for some $\bar{R}^\star \in SO(m), \lambda^\star \geq 0$, we do case analysis on $\lambda^\star$

- If $\lambda^\star = 0$ and $\mathrm{Tr}(\bar{R}^\star) \geq m - \frac{d^2}{2}$, we immediately know that $\bar{R}^\star$ is an optimal solution to the original problem.
- If $\lambda^\star > 0$ and $\mathrm{Tr}(\bar{R}^\star) = m - \frac{d^2}{2}$, for any $\bar{R} \in SO(m)$ where $\mathrm{Tr}(\bar{R}) \geq m - \frac{d^2}{2}$, we have

$$\begin{aligned}
\mathrm{Tr}(\bar{C}^T \bar{R}) &\leq \mathrm{Tr}((\bar{C} + \lambda^\star I)^T \bar{R}) - \lambda^\star(m - \frac{d^2}{2}) \\
&= \mathrm{Tr}((\bar{C} + \lambda^\star I)^T \bar{R}) - \mathrm{Tr}(\lambda^\star I^T \bar{R}^\star) \\
&\leq \mathrm{Tr}((\bar{C} + \lambda^\star I)^T \bar{R}^\star) - \mathrm{Tr}(\lambda^\star I^T \bar{R}^\star) \qquad (\ \bar{R}^\star \text{ maximizes } \mathrm{Tr}((\bar{C} + \lambda^\star I)^T \bar{R})\ ) \\
&= \mathrm{Tr}(\bar{C}^T \bar{R}^\star)
\end{aligned}$$

  which means $\bar{R}^\star$ is an optimal solution to our original problem.

Problem 6. We consider solving the following "sub-problem" using conjugate gradient descent:

$$\begin{aligned}
&\underset{\boldsymbol{p}}{\text{minimize}} \quad \phi(\boldsymbol{p}) := \boldsymbol{g}^T \boldsymbol{p} + \frac{1}{2} \boldsymbol{p}^T B \boldsymbol{p} \\
&\text{subject to} \quad \boldsymbol{p}^T C \boldsymbol{p} \leq d^2.
\end{aligned} \tag{11}$$

Here B is a real symmetric matrix and C is positive definite matrix.

The algorithm we consider is described as follows:

- **Step I:** Set $\boldsymbol{p}_0 = 0, \boldsymbol{r}_0 = -\boldsymbol{g}$. Solve $C\hat{\boldsymbol{r}}_0 = \boldsymbol{r}_0$. Set $\boldsymbol{d}_0 = \hat{\boldsymbol{r}}_0, i = 0$.
- **Step II:** Compute $\gamma_i = \boldsymbol{d}_i^T B \boldsymbol{d}_i$. If $\gamma_i > 0$ then continue with Step III. Otherwise compute $\tau > 0$ so that $(\boldsymbol{p}_i + \tau \boldsymbol{d}_i)^T C (\boldsymbol{p}_i + \tau \boldsymbol{d}_i) = d^2$, set $\boldsymbol{p} = \boldsymbol{p}_i + \tau \boldsymbol{d}_i$ and terminate.
- **Step III:** Compute $\alpha_i = \frac{\boldsymbol{r}_i^T \hat{\boldsymbol{r}}_i}{\gamma_i}$, $\boldsymbol{p}_{i+1} = \boldsymbol{p}_i + \alpha_i \boldsymbol{d}_i$. If $\boldsymbol{p}_{i+1}^T C \boldsymbol{p}_{i+1} < d^2$ than Continue with step IV. Otherwise compute $\tau > 0$ so that $(\boldsymbol{p}_i + \tau \boldsymbol{d}_i)^T C (\boldsymbol{p}_i + \tau \boldsymbol{d}_i) = d^2$, set $\boldsymbol{p} = \boldsymbol{p}_i + \tau \boldsymbol{d}_i$ and terminate.

- **Step IV:** Compute $\boldsymbol{r}_{i+1} = \boldsymbol{r}_i - \alpha_i B\boldsymbol{d}_i$. If $\frac{\boldsymbol{r}_{i+1}^T C \boldsymbol{r}_{i+1}}{\boldsymbol{g}^T C \boldsymbol{g}} \le \epsilon^2$ then set $\boldsymbol{p} = \boldsymbol{p}_{i+1}$ and terminate. Otherwise continue with Step V.

- **Step V:** Solve $C\hat{\boldsymbol{r}_{i+1}} = \boldsymbol{r}_{i+1}$. Compute $\beta_i = \frac{\boldsymbol{r}_{i+1}^T \hat{\boldsymbol{r}}_{i+1}}{\boldsymbol{r}_i^T \hat{\boldsymbol{r}}_i}$ and $\boldsymbol{d}_{i+1} = \hat{\boldsymbol{r}}_{i+1} + \beta_i \boldsymbol{d}_i$. Set $i := i+1$ and continue with step II.

Let $\boldsymbol{p}_j, j = 0, \iddots, i$ be the iterates generated by the algorithm described above.

- Show that $\phi(\boldsymbol{p}_j)$ is strictly decreasing and $\phi(\boldsymbol{p}) \le \phi(\boldsymbol{p}_i)$.

- Show that $\boldsymbol{p}_j^T C \boldsymbol{p}_j$ is strictly increasing for $j = 0, \iddots, i$ and $\boldsymbol{p}^T C \boldsymbol{p} \ge \boldsymbol{p}_j^T C \boldsymbol{p}_j$.

Solution

For understanding the algorithm, we prove the following claim by induction

$$\forall i \ge t \ge 0, \quad \boldsymbol{r}_i = -(B\boldsymbol{p}_i + \boldsymbol{g}), \quad \boldsymbol{d}_t^T \boldsymbol{r}_{i+1} = 0, \quad \boldsymbol{r}_{i+1}^T C^{-1} \boldsymbol{r}_t = 0, \quad \boldsymbol{d}_{i+1}^T B \boldsymbol{d}_t = 0 \tag{12}$$

Proof.

- For $i = t = 0$, the first equation holds immediately. By definition,

$$\begin{aligned} \boldsymbol{r}_1^T C^{-1} \boldsymbol{r}_0 = \boldsymbol{d}_0^T \boldsymbol{r}_1 &= \boldsymbol{d}_0^T (\boldsymbol{r}_0 - \alpha_0 B \boldsymbol{d}_0) \\ &= \boldsymbol{d}_0^T (\boldsymbol{r}_0 - \frac{\boldsymbol{r}_0^T \hat{\boldsymbol{r}}_0}{\boldsymbol{d}_0^T B \boldsymbol{d}_0} B\boldsymbol{d}_0) \\ &= (\boldsymbol{d}_0 - \hat{\boldsymbol{r}}_0)^T \boldsymbol{r}_0 \\ &= 0 \end{aligned}$$

$$\begin{aligned} \boldsymbol{d}_1^T B \boldsymbol{d}_0 &= (\hat{\boldsymbol{r}}_1 + \beta_0 \boldsymbol{d}_0)^T B \boldsymbol{d}_0 \\ &= (\hat{\boldsymbol{r}}_1 + \beta_0 \boldsymbol{d}_0)^T (\boldsymbol{r}_0 - \boldsymbol{r}_1)/\alpha_0 \\ &= \frac{1}{\alpha_0}(-\hat{\boldsymbol{r}}_1^T \boldsymbol{r}_1 + \beta_0 \boldsymbol{d}_0^T \boldsymbol{r}_0) \\ &= 0 \end{aligned}$$

- Assume for $t \le i \le k$, the claim holds, now for $t \le i = k+1$, we have

$$\boldsymbol{r}_{k+1} = \boldsymbol{r}_k - \alpha_k B \boldsymbol{d}_k = -(B\boldsymbol{p}_k + \boldsymbol{g} + \alpha_k B \boldsymbol{d}_k) = -(B\boldsymbol{p}_{k+1} + \boldsymbol{g})$$

and

$$\begin{aligned} \boldsymbol{d}_t^T \boldsymbol{r}_{k+2} &= \boldsymbol{d}_t^T (\boldsymbol{r}_{k+1} - \alpha_{k+1} B \boldsymbol{d}_{k+1}) \\ &= \begin{cases} \boldsymbol{r}_{k+1}^T (\boldsymbol{d}_{k+1} - \hat{\boldsymbol{r}}_{k+1}) & \text{if } t = k+1 \\ 0 & \text{if } t \le k \end{cases} \\ &= 0 \end{aligned}$$

and

$$\begin{aligned} \boldsymbol{r}_{k+2}^T C^{-1} \boldsymbol{r}_t &= \boldsymbol{r}_{k+2}^T \hat{\boldsymbol{r}}_t \\ &= (\boldsymbol{r}_{k+1} - \alpha_{k+1} B \boldsymbol{d}_{k+1})^T \hat{\boldsymbol{r}}_t \\ &= (\boldsymbol{r}_{k+1} - \alpha_{k+1} B \boldsymbol{d}_{k+1})^T (\boldsymbol{d}_t - \beta_{t-1} \boldsymbol{d}_{t-1}) \\ &= \boldsymbol{r}_{k+1}^T \boldsymbol{d}_t - \alpha_{k+1} \boldsymbol{d}_{k+1}^T B \boldsymbol{d}_t \\ &= 0 \end{aligned}$$

This implies,

$$
\begin{aligned}
\beta_{k+1} &= \frac{\hat{\boldsymbol{r}}_{k+2}^T \boldsymbol{r}_{k+2}}{\hat{\boldsymbol{r}}_{k+1}^T \boldsymbol{r}_{k+1}} \\
&= \frac{\hat{\boldsymbol{r}}_{k+2}^T (\boldsymbol{r}_{k+2} - \boldsymbol{r}_{k+1})}{\hat{\boldsymbol{r}}_{k+1}^T (\boldsymbol{r}_{k+1} - \boldsymbol{r}_{k+2})} \\
&= -\frac{\hat{\boldsymbol{r}}_{k+2}^T \alpha_{k+1} B \boldsymbol{d}_{k+1}}{\hat{\boldsymbol{r}}_{k+1}^T \alpha_{k+1} B \boldsymbol{d}_{k+1}} \\
&= -\frac{\hat{\boldsymbol{r}}_{k+2}^T B \boldsymbol{d}_{k+1}}{\hat{\boldsymbol{r}}_{k+1}^T B \boldsymbol{d}_{k+1}} \\
&= -\frac{\hat{\boldsymbol{r}}_{k+2}^T B \boldsymbol{d}_{k+1}}{(\boldsymbol{d}_{k+1} - \beta_k \boldsymbol{d}_k)^T B \boldsymbol{d}_{k+1}} \\
&= -\frac{\hat{\boldsymbol{r}}_{k+2}^T B \boldsymbol{d}_{k+1}}{\boldsymbol{d}_{k+1}^T B \boldsymbol{d}_{k+1}}
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\boldsymbol{d}_{k+2}^T B \boldsymbol{d}_t &= (\hat{\boldsymbol{r}}_{k+2} + \beta_{k+1} \boldsymbol{d}_{k+1})^T B \boldsymbol{d}_t \\
&= (\hat{\boldsymbol{r}}_{k+2} + \beta_{k+1} \boldsymbol{d}_{k+1})^T B \boldsymbol{d}_t \\
&= 0
\end{aligned}
$$

which ends the proof of our claim.

$\square$

We now begin to prove both parts

a. For the first part, note that given (12), step III can be understood as line search along direction $\boldsymbol{d}_i$, where α_i is the minimizer. i.e.

$$
\alpha_i = \arg\min_{\tau} \phi(\boldsymbol{p}_i + \tau \boldsymbol{d}_i) = \frac{1}{2}\tau^2 \gamma_i - \tau \boldsymbol{r}_i^T \hat{\boldsymbol{r}}_i
$$

Since we have $\gamma_i > 0$ in step III, this function is strongly convex. Then by step IV and II, we have

$$
\begin{aligned}
&\boldsymbol{r}_i^T \hat{\boldsymbol{r}}_i = 0 \\
\Leftrightarrow &\boldsymbol{r}_i^T C^{-1} \boldsymbol{r}_i = 0 \\
\Leftrightarrow &\boldsymbol{r}_i = \boldsymbol{0}
\end{aligned}
$$

where the last equality won't happen in step III because

0. For $i = 0$, step II checks $\gamma_0 = \boldsymbol{r}_0^T C^{-T} B C^{-1} \boldsymbol{r} > 0$
1. For $i > 0$, step IV checks $\frac{\boldsymbol{r}^T C \boldsymbol{r}}{\boldsymbol{g}^T C \boldsymbol{g}} > \epsilon^2 > 0$

which implies $\alpha_i > 0$ and $\phi(\boldsymbol{p}_{i+1}) < \phi(\boldsymbol{p}_i)$, i.e. $\phi(\boldsymbol{p}_j)$ is strictly decreasing.

b. For the second part, it's easy to see that $\boldsymbol{p}^T C \boldsymbol{p} \geq \boldsymbol{p}_j^T C \boldsymbol{p}_j$ since all $\boldsymbol{p}_j$ are interior points and $\boldsymbol{p}$ either lies on the boundary (step II, III) or equals to the last $\boldsymbol{p}_j$ (step IV).

   It remains to check $\boldsymbol{p}_j^T C \boldsymbol{p}_j$ is strictly increasing.

$$
\begin{aligned}
\boldsymbol{p}_{i+1}^T C \boldsymbol{p}_{i+1} - \boldsymbol{p}_i^T C \boldsymbol{p}_i &= \alpha_i^2 \boldsymbol{d}_i^T C \boldsymbol{d}_i + 2\alpha_i \boldsymbol{d}_i^T C \boldsymbol{p}_i \\
&> 2\alpha_i \boldsymbol{d}_i^T C \boldsymbol{p}_i \\
&= 2\alpha_i (\hat{\boldsymbol{r}}_i + \beta_{i-1} \boldsymbol{d}_{i-1})^T C (\boldsymbol{p}_{i-1} + \alpha_{i-1} \boldsymbol{d}_{i-1}) \\
&= 2\alpha_i [\beta_{i-1} \boldsymbol{d}_{i-1}^T C \boldsymbol{p}_{i-1} + \beta_{i-1} \boldsymbol{d}_{i-1}^T C \alpha_{i-1} \boldsymbol{d}_{i-1}] \\
&> 2\alpha_i \beta_{i-1} \boldsymbol{d}_{i-1}^T C \boldsymbol{p}_{i-1} \\
&> \dots
\end{aligned}
$$

Since $\boldsymbol{d}_i^T C \boldsymbol{p}_i > \beta_{i-1} \boldsymbol{d}_{i-1}^T C \boldsymbol{p}_{i-1}$ and $\boldsymbol{d}_0^T C \boldsymbol{p}_0 = 0$, we have a proof by induction.