1 Guideline

• Please complete 3 problems out of 8 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

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

where $\mu$ 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 $p^*$ being a optimal solution to the following optimization problem:

$$
\min_p \quad g^T p + \frac{1}{2} p^T B p \\
\text{subject to} \quad p^T A p \leq d^2,
$$

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

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

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

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:

\[
\begin{align*}
\text{minimize} & \quad \text{Trace}(C^T R) \\
\text{subject to} & \quad \|R_0 - R\|_F^2 \leq d^2.
\end{align*}
\]

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:

\[
\begin{align*}
\text{minimize} & \quad \phi(p) := g^T p + \frac{1}{2} p^T B p \\
\text{subject to} & \quad p^T C p \leq d^2.
\end{align*}
\]

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

The algorithm we consider is described as follows:

- **Step I:** Set \(p_0 = 0, r_0 = -g\). Solve \(C \hat{r}_0 = r_0\). Set \(d_0 = \hat{r}_0, i = 0\).

- **Step II:** Compute \(\gamma_i = d_i^T B d_i\). If \(\gamma_i > 0\) then continue with Step III. Otherwise compute \(\tau > 0\) so that \((p_i + \tau d_i)^T C (p_i + \tau d_i) = d^2\), set \(p = p_i + \tau d_i\) and terminate.

- **Step III:** Compute \(\alpha_i = \frac{r_i^T r_i}{\gamma_i} \), \(p_{i+1} = p_i + \alpha_i d_i\). If \(p_{i+1}^T C p_{i+1} < d^2\) then Continue with step IV. Otherwise compute \(\tau > 0\) so that \((p_i + \tau d_i)^T C (p_i + \tau d_i) = d^2\), set \(p = p_i + \tau d_i\) and terminate.

- **Step IV:** Compute \(r_{i+1} = r_i - \alpha_i B d_i\). If \(\frac{r_{i+1}^T C r_{i+1}}{g^T C g} \leq \epsilon^2\) then set \(p = p_{i+1}\) and terminate. Otherwise continue with Step V.

- **Step V:** Solve \(C \hat{r}_{i+1} = r_{i+1}\). Compute \(\beta_i = \frac{r_{i+1}^T \hat{r}_{i+1}}{r_i^T r_i}\) and \(d_{i+1} = \hat{r}_{i+1} + \beta_i d_i\). Set \(i := i + 1\) and continue with step II.

Let \(p_j, j = 0, \cdots, i\) be the iterates generated by the algorithm described above.

- Show that \(\phi(p_j)\) is strictly decreasing and \(\phi(p) \leq \phi(p_i)\).

- Show that \(p_j^T C p_j\) is strictly increasing for \(j = 0, \cdots, i\) and \(p_i^T C p_i \geq p_j^T C p_j\).