CS395T: Numerical Optimization for Graphics and AI: Homework I

1 Guideline

- Please complete **6** problems out of **14** problems. It is required to choose at least one problem from each section, i.e., Linear Algebra, Probability, Geometry/Topology.
- You are welcome to complete more problems.

2 Linear Algebra

Notations. $A \succeq 0$ means A is positive semidefinite, i.e., A is symmetric and all its eigenvalues are non-negative. $\|A\|$ denotes the spectral norm, i.e., the maximum singular value of A. Given a symmetric matrix X, we use $\lambda_1(X) \geq \cdots \geq \lambda_n(X)$ to denote its eigenvalues in the decreasing order.

Problem 1. The exponential map for a square matrix A is given by

$$\exp(A) := \sum_{i=0}^{\infty} \frac{1}{i!} A^i.$$

Derive an explicit expression for

$$\exp\Big(\begin{bmatrix} 0 & -z & y \\ z & 0 & -x \\ -y & x & 0 \end{bmatrix} \Big).$$

□

Problem 2. Given a 2×2 block matrix

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{12}^T & A_{22} \end{pmatrix}.$$

Suppose $A \succeq 0$. Then

$$\|A\| \leq \|A_{11}\| + \|A_{22}\|.$$

□

Problem 3. Let $\circ$ be the entry-wise product operator. Namely, given two matrices $A = (a_{ij})_{1 \leq i \leq n, 1 \leq j \leq m}, B = (b_{ij})_{1 \leq i \leq n, 1 \leq j \leq m} \in \mathbb{R}^{n \times m}$, $A \circ B = (a_{ij} b_{ij})_{1 \leq i \leq n, 1 \leq j \leq m}$. Show that

$$\|A \circ B\| \leq \|A\| \cdot \|B\|.$$

□

Problem 4. Given a square $X \in \mathbb{R}^{n \times n}$. We define the projection operator $\mathcal{P}_{O(m)}(\cdot) : \mathbb{R}^{n \times n} \rightarrow O(m)$ to the space of orthogonal matrix as follows

$$\mathcal{P}_{O(m)}(X) = UV^T, \quad X = U \Sigma V^T,$$

$X = U\Sigma V^T$ is the singular value decomposition. Given a square matrix $X \in \mathbb{R}^{n\times n}$. Suppose there exists a orthogonal matrix R such that

$$\|X - R\| \le \epsilon \le \frac{1}{3}.$$

Then

$$\|\mathcal{P}_{O(m)}(X) - R\| \le \epsilon + \epsilon^2.$$

□

Problem 5. Let $A, N \in \mathbb{R}^{n\times n}$ be two symmetric matrices. The Wely's inequality tells us that

$$|\lambda_1(A+N) - \lambda_1(A)| \le \|N\|.$$

Here we are looking for a much tight bound between $\lambda_1(A+N)$ and $\lambda_1(A)$. Denote $\boldsymbol{u}$ as the top-eigenvector of A, i.e., $\|\boldsymbol{u}\| = 1$ and $A\boldsymbol{u} = \lambda_1(A)\boldsymbol{u}$. Suppose

$$\lambda_1(A) - \lambda_2(A) \ge \|N\| + |\boldsymbol{u}^T N \boldsymbol{u}|.$$

Then

$$-|\boldsymbol{u}^T N \boldsymbol{u}| \le \lambda_1(A+N) - \lambda_1(A) \le |\boldsymbol{u}^T N \boldsymbol{u}| + \frac{\|N\|^2 - |\boldsymbol{u}^T N \boldsymbol{u}|^2}{\lambda_1(A) - \lambda_2(A)}.$$

□

3 Probability

Problem 6. Four points are chosen on the unit sphere. What is the probability that the origin lies inside the tetrahedron determined by the four points? □

Problem 7. You have $n > 1$ numbers $0, \cdots, n-1$ arranged on a circle. A random walker starts at 0 and at each step moves at random to one of its two nearest neighbors. For each i, compute the probability p_i that, when the walker is at i for the first time, all other points have been previously visited, i.e., that i is the last new point. For example, $p_0 = 0$. □

Problem 8. Let X be a random positive semidefinite matrix, and let A be a fixed positive definite matrix. Then, $\forall A$,

$$Pr[X \succeq A] \le \mathrm{Tr}(E(X)A^{-1}).$$

Here $X \succeq A$ means $X - A$ is positive semidefinite. □

Problem 9. Let $x_i \in \mathbb{R}, 1 \le i \le n$ be independent random variables that satisfies

$$E(x_i) = 0, \quad |x_i| \le 1.$$

Find the **smallest** possible constant c such that

$$Pr\big(|\sum_{i=1}^{n} x_i| \ge c\sqrt{n\log(n)}\big) \le O(\frac{1}{n^2}).$$

□

Problem 10. Suppose we choose a permutation π of the ordered set $N = \{1, 2, \cdots, n\}$ uniformly at random from the space of all permutations of N. Let $L(\pi)$ denote the length of the longest increasing subsequence in permutation π.

- For large n and some positive constant c, prove that $E[L(\pi)] \ge c\sqrt{n}$.
- Derive a upper bound on $E[L(\pi)]$.
- Derive a concentration bound on $L(\pi)$, namely, determine $f_1(n)$ and $f_2(n)$ so that $f_1(n) \le E[L(\pi)] \le f_2(n)$ with high probability.

4 Geometry and Topology

Problem 11. Consider multiple points in an Euclidean space. The maximum pairwise distance is upper bounded by 2. Determine a tight bound on the radius of the enclosing ball of these points. □

Problem 12. We color each edge of a maximally connected planar graph with one of three colors such that each face (triangle) has all three colors in its boundary.

- Show that a 4-coloring of the vertices implies a 3-coloring of the edges.
- Show that a 3-coloring of the edges implies a 4-coloring of the vertices.

Problem 13. Consider orthogonal matrices $R \in O(m), \det(R) = -1$. Collect its diagonal entries $R_{11}, \cdots, R_{mm}$ into a vector in $\mathbb{R}^m$. Prove that the convex hull of these vectors is equivalent to the convex hull of points $(\pm 1, \cdots, \pm 1)$ with a odd number of -1.

Problem 14. We have covered how to estimate the best rigid transformation between a pair of point clouds. Here we study the consistency of such pair-wise transformations among multiple point clouds. Consider n point clouds $\mathcal{P} = \{P_1, \cdots, P_n\}$. Each point cloud consists of m points i.e., $P_i = (\boldsymbol{p}_{i1}, \cdots, \boldsymbol{p}_{im}) \in \mathbb{R}^{l \times m}$, where l is the dimension of the ambient space. We assume that points $\boldsymbol{p}_{ij}, 1 \leq i \leq n$ for each fixed j are in correspondence. With this setup, we denote the optimal rigid transformation from P_i and P_j as $T_{ij} = (R_{ij}, \boldsymbol{t}_{ij})$. As we have learned in class, R_{ij} and $\boldsymbol{t}_{ij}$ admit a close-form solution via singular value decomposition.

Now we consider the consistency of these rigid transformations among multiple point clouds. For each triple of point clouds P_i, P_j, P_k, we say the pair-wise rigid transformations $T_{ij} = (R_{ij}, \boldsymbol{t}_{ij})$,$T_{jk} = (R_{jk}, \boldsymbol{t}_{jk})$ and $T_{ki} = (R_{ki}, \boldsymbol{t}_{ki})$ are consistent if $T_{ki} \circ T_{jk} \circ T_{ij} = Id$ or in other words

$$\begin{aligned} R_{ki} R_{jk} R_{ij} &= I_l \\ R_{ki} R_{jk} \boldsymbol{t}_{ij} + R_{ki} \boldsymbol{t}_{jk} + \boldsymbol{t}_{ki} &= 0 \end{aligned} \tag{1}$$

We say $\mathcal{P}$ is *regular* if the pair-wise transformations are consistent among all triples $1 \leq i \leq j \leq k \leq n$. In general, if you form $\mathcal{P}$ by sampling point clouds randomly, $\mathcal{P}$ is not regular. So this problem is to study under what conditions $\mathcal{P}$ is regular:

- Derive the condition for $l = 2$, $n = 3$ and $m = 3$.
- Derive the sufficient conditions for other configurations of l,m,and n.