

EE 290 Mathematics of Data Science - Homework 1

Rules:

1. This assignment is due at 11:59pm on **Tuesday, October 1, 2019**. No late submissions accepted.
2. Please submit the homework solution online via Gradescope. The entry code for gradescope is M4DNZ2. Hand-writing homework is allowed but required to convert to image / pdf file first. You can use any programming language. Submit your code as part of the homework. Tables and figures are preferred to help to illustrate your results. There are several sub-questions for each question. Make sure that it is easy for readers to figure out which sub-questions you are answering.

Problem A) Perturbation bound analysis.

Let A, E be positive definite matrices. Suppose that $A = LL^\top$ and $A + E = (L + G)(L + G)^\top$ are Cholesky factorizations of A and $A + E$, respectively. In this problem, we investigate the perturbation bound for Cholesky decomposition.

1. Prove that for any two matrices $X \in \mathbb{R}^{n \times d}, Y \in \mathbb{R}^{d \times k}$, $\|XY\|_F \leq \|X\|_2 \|Y\|_F$. Use this inequality to justify that $\|A\|_F \geq \|L^{-1}\|_2^{-1} \|L\|_F$ and $\|L\|_F \|L^{-1}\|_2 \leq \|A\|_F \|A^{-1}\|_2$
2. Use the above inequality to show that

$$\frac{\|E\|_F / \|A\|_2}{1 + \sqrt{1 + \|E\|_F / \|A\|_2}} \leq \frac{\|G\|_F}{\|L\|_2}$$

and

$$\frac{\|E\|_F / (\kappa(A) \|A\|_F)}{1 + \sqrt{1 + \|E\|_F / (\kappa(A) \|A\|_F)}} \leq \frac{\|G\|_F}{\|L\|_F},$$

where $\kappa(A) = \|A\|_F \|A^{-1}\|_2$.

Problem B) Orthogonally invariant norms.

Let m, n be given positive integers and $q = \min(m, n)$. For any $A \in \mathbb{R}^{m \times n}$, let $A = V\Sigma W^\top$, where $V \in \mathbb{R}^{m \times m}$ and $W \in \mathbb{R}^{n \times n}$ are orthogonal and $\Sigma \in \mathbb{R}^{m \times n}$ is a non-negative diagonal matrix whose diagonal entries are the nonincreasingly ordered singular values of A : $\sigma_1(A) \geq \dots \geq \sigma_q(A)$. Let $s(A) = [\sigma_1(A), \dots, \sigma_q(A)]^\top$.

We say a norm $\|\cdot\|$ is orthogonally invariant if $\|A\| = \|V\Sigma W^\top\| = \|\Sigma\|$ for any orthogonal matrix V, W . We say a function $g : \mathbb{R}^q \rightarrow \mathbb{R}^+$ is a symmetric gauge function if it is an absolute vector norm such that $g(x) = g(Px)$ for every $x \in \mathbb{R}^q$ and every permutation matrix $P \in \mathbb{R}^{q \times q}$. In this question, we show that the equivalence between orthogonally invariant norms and symmetric gauge functions.

1. Show that any orthogonally invariant norm can be written as a symmetric gauge function of $s(A)$. Let $\|\cdot\|$ be an orthogonally invariant norm on $\mathbb{R}^{m \times n}$. Then for any $A \in \mathbb{R}^{m \times n}$, we have $\|A\| = g(s(A))$ for some symmetric gauge function $g : \mathbb{R}^q \rightarrow \mathbb{R}^+$.
2. Show that any symmetric gauge function of $s(A)$ is an orthogonally invariant norm. Let g be a symmetric gauge function on \mathbb{R}^q . The function $\|\cdot\| : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^+$ defined by $\|A\| = g(s(A))$ is a orthogonally invariant norm on $\mathbb{R}^{m \times n}$. (Hint: you need to show first this definition is a valid norm, and then the norm is orthogonally invariant.)

Problem C) Maximum clique. Read Section 1.2 of <http://www.stat.yale.edu/~yw562/teaching/684/lec01.pdf> and prove the following generalized result. Let ω_n denote the clique number (i.e. the size of the maximum clique) in the Erdős-Rényi graph $G(n, p)$. Show that as $n \rightarrow \infty$, $\frac{\omega_n}{\log(n)}$ converges in probability to a constant. Find the limit.

Problem D) Weyl's inequality.

1. Prove the following Courant-Fischer's variational representation of eigenvalues: Let A be an $n \times n$ real symmetric matrix, with eigenvalues $\lambda_1(A) \geq \dots \geq \lambda_n(A)$. Then for each $i \in [n]$,

$$\lambda_i(A) = \sup_{\dim(V)=i} \inf_{v \in V: \|v\|_2=1} v^\top A v.$$

(Hint: use the eigenvalue decomposition of A .)

2. Prove that: if A, B are both real symmetric, then

$$|\lambda_i(A) - \lambda_i(B)| \leq \|A - B\|_2.$$

Problem E) Experiments for planted clique.

Write code to implement the planted clique detection algorithm (Algorithm 1) in Lecture 3. First generate an Erdős-Rényi graph $G(n, 1/2)$, and then select uniformly at random k points and make it a clique. Conduct experiment under different scenarios as below.

1. Run Algorithm 1 without the last cleaning step. Take n, k pairs as

$$n = 100, k = 10, 20, 30, \dots, 90$$

$$n = 200, k = 10, 30, 50, \dots, 190$$

$$n = 500, k = 10, 30, 50, \dots, 450.$$

For each n, k pair, repeat the process $m = 1000$ times. Denote K_j as the true clique for j -th experiment, and \hat{K}_j as the output of the algorithm. Let $e_{bit} = \frac{1}{m} \sum_{j=1}^m \frac{|K_j \setminus \hat{K}_j|}{K_j}$, $e_{block} = \frac{1}{m} \sum_{j=1}^m \mathbb{1}(K_j \neq \hat{K}_j)$. Plot how e_{bit} and e_{block} changes as k increases. Analyze your results.

2. Run Algorithm 1 with the last cleaning step. Repeat the above process and compare the algorithms with cleaning and without cleaning.
3. Run Algorithm 2 for $n = 500, k = 10, 20, 30, \dots, 100$. Compare the results with Algorithm 1. Record what s works the best for each (n, k) pair. Analyze your results.