SEEM5020 2026 Spring: Homework 7

This assignment is due at 1:30 pm, March 20, 2026. Please bring your solution and submit it at the beginning of the next class. Total weight of Homework 7: 4%.

- **Q1. [10 Points] Median Trick for Matrix Data.**

  Let $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $\mathbf{B} \in \mathbb{R}^{n\times p}$. Suppose we run ℓ independent trials of the sampling-based approximate matrix multiplication algorithm. In each trial, we draw $t = \lceil \frac{3}{(\varepsilon/3)^2} \rceil$ samples and obtain a matrix $\mathbf{D}_i$ satisfying

  $$\mathbb{P}\left[\|\mathbf{D}_i - \mathbf{AB}\|_F < \frac{\varepsilon}{3}\|\mathbf{A}\|_F\|\mathbf{B}\|_F\right] \geq \frac{2}{3}.$$

  For each $\mathbf{D}_i$, define $\omega_i = |\{j \mid j \neq i \text{ and } \|\mathbf{D}_i - \mathbf{D}_j\|_F \leq \frac{2\varepsilon}{3}\|\mathbf{A}\|_F\|\mathbf{B}\|_F\}|$. The algorithm outputs any $\mathbf{D}_i$ with $\omega_i \geq \ell/2$.

  Prove Theorem 1 (Page 14 of Lecture 7): if $\ell = 2\lceil 24\ln(1/\delta)\rceil$ (slightly modified to make the proof smoother) and $\sum_{i=1}^{\ell} Y_i \geq \ell/2 + 1$ (where $Y_i = \mathbf{1}[\|\mathbf{D}_i - \mathbf{AB}\|_F < \frac{\varepsilon}{3}\|\mathbf{A}\|_F\|\mathbf{B}\|_F]$), then the algorithm outputs $\mathbf{D}_i$ satisfying

  $$\|\mathbf{D}_i - \mathbf{AB}\|_F \leq \varepsilon\|\mathbf{A}\|_F\|\mathbf{B}\|_F,$$

  and this event occurs with probability at least $1 - \delta$.

  Hint: For the first part, show that a "bad" matrix (with error $\geq \varepsilon$) cannot have $\omega_i \geq \ell/2$ when the majority of trials are "good." For the second part, use Chernoff bound to show $\mathbb{P}[\sum Y_i < \ell/2 + 1] \leq \delta$.

- **Q2. [10 Points] Variance Bound for Multi-Sample Approximate Matrix Multiplication.**

  In the approximate matrix multiplication via importance sampling (Pages 13–17 of Lecture 7), we proved that when $t = 1$,

  $$\mathbb{E}[\|\mathbf{D} - \mathbf{AB}\|_F^2] = \left(\sum_{j=1}^{n} \|\mathbf{A}_{(*,j)}\|_2\|\mathbf{B}_{(j,*)}\|_2\right)^2 - \|\mathbf{AB}\|_F^2.$$

  Please extend this to arbitrary $t \geq 1$ and prove that

  $$\mathbb{E}[\|\mathbf{D} - \mathbf{AB}\|_F^2] \leq \frac{1}{t}\left(\sum_{j=1}^{n} \|\mathbf{A}_{(*,j)}\|_2\|\mathbf{B}_{(j,*)}\|_2\right)^2 - \frac{1}{t}\|\mathbf{AB}\|_F^2.$$

  Hint: Define $\mathbf{D}^{(i)}$ as the matrix from the i-th sample, so $\mathbf{D} = \frac{1}{t}\sum_{i=1}^{t} \mathbf{D}^{(i)}$. Use independence of the t samples to decompose the variance.

- **Q3. [10 Points] Optimal Sampling Probabilities.**

  In the approximate matrix multiplication framework (Page 19 of Lecture 7), we showed that

  $$\mathbb{E}[\|\mathbf{D} - \mathbf{AB}\|_F^2] = \sum_{i=1}^{n} \frac{\|\mathbf{A}_{(*,i)}\|_2^2\|\mathbf{B}_{(i,*)}\|_2^2}{p_i} - \|\mathbf{AB}\|_F^2.$$

  To minimize the variance, we solve the constrained optimization problem:

  $$\min_{p_1,\ldots,p_n} \sum_{i=1}^{n} \frac{\|\mathbf{A}_{(*,i)}\|_2^2\|\mathbf{B}_{(i,*)}\|_2^2}{p_i} \quad \text{subject to} \quad \sum_{i=1}^{n} p_i = 1.$$

Using the Lagrange multiplier method, prove that the optimal solution is

$$p_i = \frac{\|\mathbf{A}_{(*,i)}\|_2 \|\mathbf{B}_{(i,*)}\|_2}{\sum_{j=1}^{n} \|\mathbf{A}_{(*,j)}\|_2 \|\mathbf{B}_{(j,*)}\|_2}, \quad i = 1, 2, \ldots, n.$$

- **Q4. [10 Points] Geometric Median Trick for Vector Estimation.**

  Suppose $v \in \mathbb{R}^d$ is an unknown target vector. We run an independent randomized estimator s times and obtain vectors $\hat{v}_1, \hat{v}_2, \ldots, \hat{v}_s \in \mathbb{R}^d$. Assume that for each i,

$$\mathbb{P}[\|\hat{v}_i - v\|_2 \leq r] \geq \frac{3}{4}.$$

  Define the *geometric median* of $\hat{v}_1, \ldots, \hat{v}_s$ to be any minimizer

$$\tilde{v} \in \arg\min_{x \in \mathbb{R}^d} \sum_{i=1}^{s} \|x - \hat{v}_i\|_2.$$

  (a) Prove that if at least $\frac{2s}{3}$ of the vectors $\hat{v}_i$ satisfy $\|\hat{v}_i - v\|_2 \leq r$, then $\|\tilde{v} - v\|_2 \leq 4r$.

  Hint: Let $G = \{i : \|\hat{v}_i - v\|_2 \leq r\}$ and $B = [s] \setminus G$. Compare $F(x) = \sum_{i=1}^{s} \|x - \hat{v}_i\|_2$ with $F(v)$. Use the triangle inequality for indices in G and the reverse triangle inequality for indices in B.

  (b) Let $Y_i = \mathbf{1}[\|\hat{v}_i - v\|_2 \leq r]$. Use the Chernoff bound to show that

$$\mathbb{P}[\|\tilde{v} - v\|_2 > 4r] \leq e^{-s/216}.$$

  Hence, show that by taking $s \geq 216 \ln(1/\delta)$, we obtain $\mathbb{P}[\|\tilde{v} - v\|_2 \leq 4r] \geq 1 - \delta$.

  Remark: The geometric median trick provides a way to boost success probability for vector-valued estimators, analogous to the scalar median trick but applicable in high dimensions.