SEEM5020 2026 Spring: Homework 7 Solution

- **Q1. Median Trick for Matrix Data.**

  Solution:

  Correctness when the majority is good. Suppose $\sum_{i=1}^{\ell} Y_i \geq \ell/2 + 1$. We show two things:

  (i) Any "good" $\mathbf{D}_i$ (with $Y_i = 1$) satisfies $\omega_i \geq \ell/2$, so the algorithm finds an output.

  (ii) Any $\mathbf{D}_i$ output by the algorithm satisfies $\|\mathbf{D}_i - \mathbf{AB}\|_F \leq \varepsilon \|\mathbf{A}\|_F \|\mathbf{B}\|_F$.

  For (i): If $Y_i = 1$ and $Y_j = 1$, then by the triangle inequality,

  $$\|\mathbf{D}_i - \mathbf{D}_j\|_F \leq \|\mathbf{D}_i - \mathbf{AB}\|_F + \|\mathbf{D}_j - \mathbf{AB}\|_F < \frac{\varepsilon}{3}\|\mathbf{A}\|_F\|\mathbf{B}\|_F + \frac{\varepsilon}{3}\|\mathbf{A}\|_F\|\mathbf{B}\|_F = \frac{2\varepsilon}{3}\|\mathbf{A}\|_F\|\mathbf{B}\|_F.$$

  Since there are at least $\ell/2 + 1$ good matrices and at least $\ell/2$ of them have index $j \neq i$, we have $\omega_i \geq \ell/2$.

  For (ii): Suppose for contradiction that the algorithm outputs $\mathbf{D}_i$ with $\|\mathbf{D}_i - \mathbf{AB}\|_F \geq \varepsilon\|\mathbf{A}\|_F\|\mathbf{B}\|_F$, yet $\omega_i \geq \ell/2$. Then there exist at least $\ell/2$ indices $j \neq i$ with $\|\mathbf{D}_i - \mathbf{D}_j\|_F \leq \frac{2\varepsilon}{3}\|\mathbf{A}\|_F\|\mathbf{B}\|_F$. For any such j, by the reverse triangle inequality:

  $$\|\mathbf{D}_j - \mathbf{AB}\|_F \geq \|\mathbf{D}_i - \mathbf{AB}\|_F - \|\mathbf{D}_i - \mathbf{D}_j\|_F \geq \varepsilon\|\mathbf{A}\|_F\|\mathbf{B}\|_F - \frac{2\varepsilon}{3}\|\mathbf{A}\|_F\|\mathbf{B}\|_F = \frac{\varepsilon}{3}\|\mathbf{A}\|_F\|\mathbf{B}\|_F.$$

  So $Y_j = 0$ for all such j. Including index i itself (also bad), there are at least $\ell/2 + 1$ bad indices, meaning $\sum Y_i \leq \ell - (\ell/2 + 1) < \ell/2$. This contradicts $\sum Y_i \geq \ell/2 + 1$.

  Probability bound. Each Y_i is an independent Bernoulli random variable with $\mathbb{E}[Y_i] \geq 2/3$. Let $Y = \sum_{i=1}^{\ell} Y_i$ and $\mu = \mathbb{E}[Y] \geq 2\ell/3$. Since ℓ is even and Y is integer-valued, the event $Y < \ell/2 + 1$ is equivalent to $Y \leq \ell/2$.

  Also, since $\ell/2 = (1 - \frac{1}{4}) \cdot (2\ell/3) \leq (1 - \frac{1}{4})\mu$, the Chernoff bound gives

  $$\begin{aligned}
  \mathbb{P}\left[Y < \frac{\ell}{2} + 1\right] &= \mathbb{P}\left[Y \leq \frac{\ell}{2}\right] \\
  &\leq \mathbb{P}\left[Y \leq \left(1 - \frac{1}{4}\right)\mu\right] \\
  &\leq \exp\left(-\frac{\mu(1/4)^2}{2}\right) \\
  &\leq \exp\left(-\frac{(2\ell/3)}{32}\right) = e^{-\ell/48}.
  \end{aligned}$$

  Now set $\ell = 2\lceil 24 \ln(1/\delta) \rceil$. Then $\ell \geq 48 \ln(1/\delta)$, so $e^{-\ell/48} \leq \delta$. Therefore, $\mathbb{P}[Y < \ell/2 + 1] \leq \delta$, and hence with probability at least $1 - \delta$ we have $Y \geq \ell/2 + 1$. By the first part, the algorithm outputs some $\mathbf{D}_i$ satisfying $\|\mathbf{D}_i - \mathbf{AB}\|_F \leq \varepsilon\|\mathbf{A}\|_F\|\mathbf{B}\|_F$.

- **Q2. Variance Bound for Multi-Sample Approximate Matrix Multiplication.**

  Solution:

  Define $\mathbf{D}^{(i)}$ as the matrix obtained in the i-th sample, so $\mathbf{D} = \frac{1}{t}\sum_{i=1}^{t} \mathbf{D}^{(i)}$. Since $\mathbb{E}[\mathbf{D}^{(i)}_{x,y}] = (\mathbf{AB})_{x,y}$ for each i, we have $\mathbb{E}[\mathbf{D}_{x,y}] = (\mathbf{AB})_{x,y}$.

  Now consider:

  $$\mathbb{E}[\|\mathbf{D} - \mathbf{AB}\|_F^2] = \sum_{x,y} \mathbb{E}[(\mathbf{D}_{x,y} - (\mathbf{AB})_{x,y})^2] = \sum_{x,y} \mathrm{Var}[\mathbf{D}_{x,y}].$$

Since $\mathbf{D}_{x,y} = \frac{1}{t}\sum_{i=1}^{t}\mathbf{D}_{x,y}^{(i)}$ and the t samples are independent:

$$\mathrm{Var}[\mathbf{D}_{x,y}] = \frac{1}{t^2}\sum_{i=1}^{t}\mathrm{Var}[\mathbf{D}_{x,y}^{(i)}] = \frac{1}{t}\mathrm{Var}[\mathbf{D}_{x,y}^{(1)}].$$

From the $t = 1$ case, we know:

$$\mathrm{Var}[\mathbf{D}_{x,y}^{(1)}] = \sum_{i=1}^{n}\frac{(\mathbf{A}_{x,i}\mathbf{B}_{i,y})^2}{p_i} - (\mathbf{AB})_{x,y}^2.$$

Summing over all x, y:

$$\begin{aligned}\mathbb{E}[\|\mathbf{D}-\mathbf{AB}\|_F^2] &= \frac{1}{t}\sum_{x,y}\left(\sum_{i=1}^{n}\frac{(\mathbf{A}_{x,i}\mathbf{B}_{i,y})^2}{p_i} - (\mathbf{AB})_{x,y}^2\right)\\ &= \frac{1}{t}\left(\sum_{i=1}^{n}\frac{\|\mathbf{A}_{(*,i)}\|_2^2\|\mathbf{B}_{(i,*)}\|_2^2}{p_i} - \|\mathbf{AB}\|_F^2\right).\end{aligned}$$

Substituting $p_i = \frac{\|\mathbf{A}_{(*,i)}\|_2\|\mathbf{B}_{(i,*)}\|_2}{\sum_{j=1}^{n}\|\mathbf{A}_{(*,j)}\|_2\|\mathbf{B}_{(j,*)}\|_2}$:

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

- **Q3. Optimal Sampling Probabilities.**

  Solution:

  For convenience, denote $c_i = \|\mathbf{A}_{(*,i)}\|_2^2\|\mathbf{B}_{(i,*)}\|_2^2$. We minimize $\sum_{i=1}^{n} c_i/p_i$ subject to $\sum_{i=1}^{n} p_i = 1$. Define the Lagrangian:

  $$L(p_1,\ldots,p_n,\lambda) = \sum_{i=1}^{n}\frac{c_i}{p_i} + \lambda\left(\sum_{i=1}^{n}p_i - 1\right).$$

  Taking partial derivatives and setting them to zero:

  $$\frac{\partial L}{\partial p_i} = -\frac{c_i}{p_i^2} + \lambda = 0 \implies p_i = \sqrt{\frac{c_i}{\lambda}} = \frac{\sqrt{c_i}}{\sqrt{\lambda}}.$$

  Applying the constraint $\sum_{i=1}^{n} p_i = 1$:

  $$\frac{1}{\sqrt{\lambda}}\sum_{i=1}^{n}\sqrt{c_i} = 1 \implies \sqrt{\lambda} = \sum_{i=1}^{n}\sqrt{c_i}.$$

  Therefore:

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

- **Q4. Geometric Median Trick for Vector Estimation.**

  Solution:

  (a) Let $G = \{i : \|\hat{v}_i - v\|_2 \le r\}$ and $B = [s] \setminus G$. Denote $g = |G|$ and $b = |B|$. By assumption, $g \ge 2s/3$ and $b \le s/3$.

Define $F(x) = \sum_{i=1}^{s} \|x - \hat{v}_i\|_2$. Since $\tilde{v}$ is the geometric median, $F(\tilde{v}) \leq F(x)$ for all x.

For any $x \in \mathbb{R}^d$ and $i \in G$, by the triangle inequality and $\|\hat{v}_i - v\|_2 \leq r$:

$$\|x - \hat{v}_i\|_2 - \|v - \hat{v}_i\|_2 \geq \|x - v\|_2 - 2\|\hat{v}_i - v\|_2 \geq \|x - v\|_2 - 2r.$$

For $i \in B$, by the reverse triangle inequality:

$$\|x - \hat{v}_i\|_2 - \|v - \hat{v}_i\|_2 \geq -\|x - v\|_2.$$

Summing over all i:

$$\begin{aligned} F(x) - F(v) &= \sum_{i=1}^{s} (\|x - \hat{v}_i\|_2 - \|v - \hat{v}_i\|_2) \\ &\geq g(\|x - v\|_2 - 2r) - b\|x - v\|_2 = (g - b)\|x - v\|_2 - 2gr. \end{aligned}$$

If $\|x - v\|_2 > \frac{2gr}{g-b}$, then $F(x) > F(v)$, so x cannot be a minimizer. Therefore every minimizer $\tilde{v}$ satisfies $\|\tilde{v} - v\|_2 \leq \frac{2gr}{g-b}$.

Using $g \geq 2s/3$ and $b \leq s/3$:

$$\frac{2gr}{g - b} \leq \frac{2gr}{g/2} = 4r.$$

(b) The variables $Y_1, \ldots, Y_s$ are independent Bernoulli with $\mathbb{E}[Y_i] \geq 3/4$. Let $Y = \sum_{i=1}^{s} Y_i$ and $\mu = \mathbb{E}[Y] \geq 3s/4$.

By part (a), if $Y \geq 2s/3$, then $\|\tilde{v} - v\|_2 \leq 4r$. Therefore:

$$\mathbb{P}[\|\tilde{v} - v\|_2 > 4r] \leq \mathbb{P}[Y < 2s/3].$$

Since $\mu \geq 3s/4$, we have $2s/3 \leq \left(1 - \frac{1}{9}\right)\mu$ (as $\frac{8}{9} \cdot \frac{3s}{4} = \frac{2s}{3}$). By the Chernoff bound $\mathbb{P}[Y < (1 - \delta')\mu] \leq e^{-\mu\delta'^2/2}$ with $\delta' = 1/9$:

$$\mathbb{P}\left[Y < \frac{2s}{3}\right] \leq \mathbb{P}\left[Y < \left(1 - \frac{1}{9}\right)\mu\right] \leq \exp\left(-\frac{\mu(1/9)^2}{2}\right) \leq \exp\left(-\frac{(3s/4)}{162}\right) = e^{-s/216}.$$

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

If $s \geq 216 \ln(1/\delta)$, then $e^{-s/216} \leq \delta$, and therefore $\mathbb{P}[\|\tilde{v} - v\|_2 \leq 4r] \geq 1 - \delta$.