EE 378B Homework 4 Solution

Due to: Samir Khan

Problem 1.

(a) Assuming $\|z\|_2 \leq 1$, we have

$$\begin{aligned}
\|A^T z\|_2^2 &= z^T AA^T z, \\
&= z^T BB^T z + z^T (AA^T - BB^T) z, \\
&\leq z^T \left(\sum_{i=1}^{m} \sigma_i(BB^T) \ell_i \ell_i^T \right) z + \lambda_1(AA^T - BB^T), \\
&= z^T \left(\sum_{i=k+1}^{m} \sigma_i(BB^T) \ell_i \ell_i^T \right) z + \|AA^T - BB^T\|_{\text{op}}, && \text{(because } L_k^T z = 0\text{)} \\
&\leq \sigma_{k+1}(BB^T) + \|AA^T - BB^T\|_{\text{op}}, \\
&\leq \sigma_{k+1}(AA^T) + 2\|AA^T - BB^T\|. && \text{(by Weyl's inequality)}
\end{aligned}$$

Then,

$$\|A - L_k L_k^T A\|_{\text{op}}^2 = \|A^T - A^T L_k L_k^T\|_{\text{op}}^2 = \max_{\|z\|_2 \leq 1} \|A^T (I - L_k L_k^T) z\|_2^2.$$

The matrix $I - L_k L_k^T$ is the orthogonal projection onto $\text{col}(L_k)^\perp$, so $(I - L_k L_k^T)z$ satisfies the hypothesis of (a). Using that bound then gives

$$\|A - L_k L_k^T A\|_{\text{op}}^2 \leq \sigma_{k+1}(AA^T) + 2\|AA^T - BB^T\|_{\text{op}} = \sigma_{k+1}^2(A) + 2\|AA^T - BB^T\|,$$

which is what we wanted sinc $\sigma_{k+1}(A) = \|A - A_k\|_{\text{op}}$.

(b) Based on our work in the previous part, it suffices to show that with probability at least $1 - m^{-2}$,

$$\|AA^T - BB^T\|_{\text{op}} \leq Cn\|A\|_{\text{op}} \|A\|_{1\to 2} \sqrt{\frac{\log m}{\ell}}.$$

To do this, we write

$$B = \sum_{j=1}^{\ell} \frac{1}{\sqrt{\ell p_{i(j)}}} a_{i(j)} e_j^T,$$

which leads to

$$BB^T = \left(\sum_{j=1}^{\ell} \frac{1}{\sqrt{\ell p_{i(j)}}} a_{i(j)} e_j^T \right) \left(\sum_{k=1}^{\ell} \frac{1}{\sqrt{\ell p_{i(k)}}} a_{i(k)} e_k^T \right)^T = \frac{1}{\ell} \sum_{j,k=1}^{\ell} \frac{1}{\sqrt{p_{i(j)} p_{i(k)}}} a_{i(j)} e_j^T e_k a_{i(k)}^T = \frac{1}{\ell} \sum_{j=1}^{\ell} \frac{1}{p_{i(j)}} a_{i(j)} a_{i(j)}^T.$$

Then,

$$\|AA^T - BB^T\|_{\text{op}} = \|AA^T - \frac{1}{\ell} \sum_{j=1}^{\ell} \frac{1}{p_{i(j)}} a_{i(j)} a_{i(j)}^T\|_{\text{op}} = \| \sum_{j=1}^{\ell} \frac{1}{\ell} \left(\frac{1}{p_{i(j)}} a_{i(j)} a_{i(j)}^T - AA^T \right) \|_{\text{op}}.$$

Now, define

$$Z_j = \frac{1}{\ell}\left(\frac{1}{p_{i(j)}} a_{i(j)} a_{i(j)}^T - AA^T\right) = \frac{1}{\ell}\left(n a_{i(j)} a_{i(j)}^T - AA^T\right).$$

Then $\mathbb{E}[Z_j] = 0$ and

$$\|Z_j\|_{\mathrm{op}} \leq \frac{n}{\ell}\|a_{i(j)} a_{i(j)}^T\|_{\mathrm{op}} + \frac{1}{\ell}\|AA^T\|_{\mathrm{op}} \leq \frac{n}{\ell}\|A\|_{1\to 2}^2 + \frac{1}{\ell}\|A\|_{\mathrm{op}}^2.$$

Then,

$$\mathbb{E}[Z_j^2] = \frac{1}{\ell^2}\mathbb{E}\left[\left(n a_{i(j)} a_{i(j)}^T - AA^T\right)^2\right] = \frac{1}{\ell^2}\left(n\sum_{j=1}^n (a_j a_j^T)^2 - (AA^T)^2\right) \prec \frac{n}{\ell^2}\sum_{j=1}^n (a_j a_j^T)^2.$$

Note that

$$\sum_{j=1}^n (a_j a_j^T)^2 = \sum_{j=1}^n a_j a_j^T a_j a_j^T \prec \|A\|_{1\to 2}^2 \sum_{j=1}^n a_j a_j^T = \|A\|_{1\to 2}^2 AA^T.$$

Combining the previous two displays, we conclude that

$$\mathbb{E}[Z_j^2] \prec \frac{n}{\ell^2}\|A\|_{1\to 2}^2 AA^T \implies \|\mathbb{E}[Z_j^2]\|_{\mathrm{op}} \leq \frac{n}{\ell^2}\|A\|_{1\to 2}^2\|A\|_{\mathrm{op}}^2.$$

At this point, we can apply Bernstein's inequality with

$$K = \frac{n}{\ell}\|A\|_{1\to 2}^2 + \frac{1}{\ell}\|A\|_{\mathrm{op}}^2 \quad \text{and} \quad \sigma^2 = \frac{n}{\ell}\|A\|_{1\to 2}^2\|A\|_{\mathrm{op}}^2$$

to conclude that with probability at least $1 - m^{-2}$ we have

$$\|\sum_{j=1}^{\ell} Z_j\|_{\mathrm{op}} \leq C\left(\sigma\sqrt{\log(m)} \vee K\log(m)\right).$$

To obtain the exact result of the problem, we need to show that both $\sigma\sqrt{\log m}$ and $K\log m$ are bounded by

$$Cn\|A\|_{\mathrm{op}}\|A\|_{1\to 2}\sqrt{\frac{\log m}{\ell}}.$$

Clearly we always have

$$\sigma\sqrt{\log m} = \sqrt{n}\|A\|_{1\to 2}\|A\|_{\mathrm{op}}\sqrt{\frac{\log m}{\ell}} < n\|A\|_{1\to 2}\|A\|_{\mathrm{op}}\sqrt{\frac{\log m}{\ell}}.$$

If we additionally assume that $\ell > \log m$, we also have

$$K\log m = \left(n\|A\|_{1\to 2}^2 + \|A\|_{\mathrm{op}}^2\right)\frac{\log m}{\ell} \leq \left(n\|A\|_{1\to 2}^2 + \|A\|_{\mathrm{op}}^2\right)\sqrt{\frac{\log m}{\ell}},$$

and all that is left is to verify that

$$n\|A\|_{1\to 2}^2 + \|A\|_{\mathrm{op}}^2 \leq Cn\|A\|_{1\to 2}\|A\|_{\mathrm{op}}.$$

In particular, we take $C = 2$. It is easy to check that

$$\|A\|_{1\to 2} \leq \|A\|_{\mathrm{op}} \leq \sqrt{n}\|A\|_{1\to 2},$$

and these bounds give

$$n\|A\|_{1\to 2}^2 + \|A\|_{\text{op}}^2 \le 2n\|A\|_{1\to 2}^2 \le 2n\|A\|_{1\to 2}\|A\|_{\text{op}}.$$

Finally, we conclude that if $\ell > \log m$, then we have

$$\|AA^T - BB^T\|_{\text{op}} \le Cn\|A\|_{\text{op}}\|A\|_{1\to 2}\sqrt{\frac{\log m}{\ell}}$$

with probability at least $1 - m^{-2}$.

(c) As before, let

$$Z_j = \frac{1}{\ell}\left(\frac{1}{p_{i(j)}} a_{i(j)} a_{i(j)}^T - AA^T\right),$$

but now we take

$$p_i = \frac{\|a_i\|_2^2}{\|a_1\|_2^2 + \cdots + \|a_n\|_2^2} = \frac{\|a_i\|_2^2}{\|A\|_F^2}.$$

Then,

$$\|Z_j\|_{\text{op}} \le \frac{1}{\ell}\|\frac{1}{p_{i(j)}} a_{i(j)} a_{i(j)}^T\| + \frac{1}{\ell}\|A\|_{\text{op}}^2 \le \frac{2\|A\|_F^2}{\ell},$$

since the operator norm is bounded by the Frobenius norm. Similarly,

$$\mathbb{E}[Z_j^2] = \frac{1}{\ell^2}\left(\sum_{j=1}^n \frac{\|A\|_F^2}{\|a_j\|_2^2}(a_j a_j^T)^2 - (AA^T)^2\right) \prec \frac{\|A\|_F^2}{\ell^2}\sum_{j=1}^n a_j a_j^T = \frac{\|A\|_F^2}{\ell^2} AA^T.$$

Thus

$$\|\mathbb{E}[Z_j^2]\|_{\text{op}} \le \frac{\|A\|_F^2}{\ell^2}\|AA^T\|_{\text{op}} \le \frac{\|A\|_F^4}{\ell^2}.$$

So we can apply Bernstein with

$$K = \frac{2\|A\|_F^2}{\ell} \quad \text{and} \quad \sigma^2 = \frac{\|A\|_F^4}{\ell}$$

and obtain that with probability at least $1 - m^{-2}$ we have

$$\|\sum_{j=1}^{\ell} Z_j\|_{\text{op}} \le C\left(\sigma\sqrt{\log m} \vee K\log m\right).$$

Now

$$\sigma\sqrt{\log m} = \|A\|_F^2\sqrt{\frac{\log m}{\ell}},$$

and if $\ell > \log m$,

$$K\log m = 2\|A\|_F^2\frac{\log m}{\ell} \le 2\|A\|_F^2\sqrt{\frac{\log m}{\ell}},$$

which was the bound we wanted.