Leave-one-out Analysis: An Illustrative Example

Hao Yan

May 26, 2022

1 Introduction

Consider the following toy example

$$M = M^\star + E = U^\star \Lambda^\star U^{\star\top} + E \in \mathbb{R}^{n\times n}.$$

Assume that $M^\star$ is a rank-r matrix, and E is symmetric, with $E_{ij} \overset{i.i.d.}{\sim} \mathcal{N}(0, \sigma^2)$. We first show how to establish ℓ_2 bound for the subspace, using $\operatorname{dist}(U, U^\star) := \min_{R\in\mathcal{O}^{r\times r}} \|UR - U^\star\|$.

By **truncated matrix Bernstein inequality**, we have with probability at least $1 - O(n^{-8})$,

$$\|E\| \le 5\sigma\sqrt{n}. \tag{1}$$

Now by **Davis-Kahan** $\sin\Theta$ **Theorem**, which is

Theorem 1.1. *If* $\|E\| \le (1 - 1/\sqrt{2})\left(|\lambda_r^\star| - \left|\lambda_{r+1}^\star\right|\right)$, *then*

$$\operatorname{dist}(U, U^\star) \le \frac{2\,\|EU^\star\|}{|\lambda_r^\star| - \left|\lambda_{r+1}^\star\right|} \le \frac{2\|E\|}{|\lambda_r^\star| - \left|\lambda_{r+1}^\star\right|}.$$

Thus, if we assume that $\sigma\sqrt{n} \le \frac{1-1/\sqrt{2}}{5}|\lambda_r^\star|$, then one has with high probability,

$$\operatorname{dist}(U, U^\star) \le \frac{2\|E\|}{|\lambda_r^\star|} \le \frac{10\sigma\sqrt{n}}{\lambda_r^\star}.$$

From **Weyl's inequality**, the bound (1) also implies that

$$|\lambda_i| \le \|E\| \le 5\sigma\sqrt{n}, \quad \text{for all } i \ge r+1,$$

and

$$|\lambda_i - \lambda_i^\star| \le \|E\| \le 5\sigma\sqrt{n}, \quad \text{for all } i \le r.$$

2 ℓ_∞ Performance Guarantees

Now we focus on the case $r = 1$. Our goal is to obtain bound for

$$\operatorname{dist}_\infty(u, u^\star) := \min\{\|u - u^\star\|_\infty, \|u + u^\star\|_\infty\}.$$

Let us first construct an auxiliary matrix

$$M^{(l)} := \lambda^\star u^\star u^{\star\top} + E^{(l)},$$

where the noise matrix $E^{(l)}$ is generated according to

$$E_{i,j}^{(l)} := \begin{cases} E_{i,j}, & \text{if } i \ne l \text{ and } j \ne l, \\ 0, & \text{else.} \end{cases}$$

Assume that $C\sigma\sqrt{n} < \lambda^\star$, for sufficiently large C. For simplicity, we also assume that

$$\begin{aligned} \|u - u^\star\|_2 &= \mathrm{dist}(u, u^\star), \\ \|u^{(l)} - u^\star\|_2 &= \mathrm{dist}(u^{(l)}, u^\star), \quad 1 \le l \le n \end{aligned}$$

otherwise we simply change the sign of u or $u^{(l)}$. We claim that

$$\|u - u^{(l)}\|_2 = \mathrm{dist}(u, u^{(l)}).$$

To see this, we have

$$\|u - u^{(l)}\|_2 \le \|u - u^\star\|_2 + \|u^{(l)} - u^\star\|_2 \le \frac{20\sigma\sqrt{n}}{\lambda^\star} < 1.$$

And it follows that

$$\|u + u^{(l)}\|_2^2 = 2\|u\|_2^2 + 2\|u^{(l)}\|_2^2 - \|u - u^{(l)}\|_2^2 > 1 > \|u - u^{(l)}\|_2^2.$$

After the previous preparations, we state the ideas here. Since $M^{(l)}$ is very close to M, we expect that u and $u^{(l)}$ are also quite close. By the previous result, this can be shown by establishing a bound for $\mathrm{dist}(u, u^{(l)})$. Now observe that

$$u_l^{(l)} = u^{(l)\top} u^{(l)} u_l^{(l)\top} = \frac{1}{\lambda^{(l)}} u^{(l)\top} M_{\cdot l}^{(l)} = \frac{1}{\lambda^{(l)}} u^{(l)\top} M_{\cdot l}^\star = \frac{\lambda^\star}{\lambda^{(l)}} u^{(l)\top} u^\star u_l^\star,$$

if $\lambda^{(l)} \approx \lambda^\star$ and $u^{(l)\top} u^\star \approx 1$, we have $u_l^{(l)} \approx u_l^\star$. This is not hard to achieve as we already have

$$\frac{|\lambda^{(l)} - \lambda^\star|}{\lambda^\star} \le \|E\|/\lambda^\star \le \frac{5\sigma\sqrt{n}}{\lambda^\star},$$

and

$$\|u^{(l)} - u^\star\|_2 \le \frac{10\sigma\sqrt{n}}{\lambda^\star}.$$

So we proceed as follows.

Step 1. A bound for $\mathbf{dist(u, u^{(l)})}$. This kind of bound is usually established by Davis-Kahan $\sin\Theta$ Theorem. For the moment, assume the condition for Davis-Kahan holds. Since

$$\mathrm{dist}(u, u^{(l)}) \le \frac{2\|(M - M^{(l)})u^{(l)}\|_2}{\lambda^{(l)} - \max_{j\ge 2} |\lambda_j(M^{(l)})|},$$

and

$$\lambda^{(l)} - \max_{j\ge 2} |\lambda_j(M^{(l)})| \ge (\lambda^\star - 5\sigma\sqrt{n}) - 5\sigma\sqrt{n} \ge \lambda^\star/2,$$

we have

$$\mathrm{dist}(u, u^{(l)}) \le \frac{4\|(M - M^{(l)})u^{(l)}\|_2}{\lambda^\star}.$$

Now observe that

$$\begin{aligned} \|(M - M^{(l)})u^{(l)}\|_2 &= \|(E - E^{(l)})u^{(l)}\|_2 \\ &= \|e_l E_{l\cdot} u^{(l)} + u_l^{(l)}(E_{\cdot l} - E_{ll} e_l)\|_2, \end{aligned}$$

where e_l is the l-th standard basis vector. By construction, $u^{(l)}$ is **independent** of $E_{l\cdot}$, which is the key observation in leave-one-out analysis. We have

$$E_{l\cdot} u^{(l)} \sim \mathcal{N}(0, \sigma^2 \|u^{(l)}\|_2^2) = \mathcal{N}(0, \sigma^2)$$

conditioned on $u^{(l)}$. Hence, with probability at least $1 - n^{-10}$,

$$|E_{l\cdot} u^{(l)}| \le 5\sigma\sqrt{\log n}, \quad 1 \le l \le n.$$

In addition, $\|E_{\cdot l} - E_{ll} e_l\|_2 \le \|E_{\cdot l}\|_2 = \|E e_l\|_2 \le \|E\|_2 \|e_l\|_2 = \|E\|_2 \le 5\sigma\sqrt{n}$. Consequently,

$$\begin{aligned}
\|(M - M^{(l)})u^{(l)}\|_2 &\le 5\sigma\sqrt{\log n} + 5\sigma\sqrt{n}|u_l^{(l)}| \\
&\le 5\sigma\sqrt{\log n} + 5\sigma\sqrt{n}\left(\|u\|_\infty + \|u - u^{(l)}\|_\infty\right) \\
&\le 5\sigma\sqrt{\log n} + 5\sigma\sqrt{n}\left(\|u\|_\infty + \|u - u^{(l)}\|_2\right).
\end{aligned}$$

Thus,

$$\operatorname{dist}(u, u^{(l)}) \le \frac{20\sigma\sqrt{\log n} + 20\sigma\sqrt{n}\left(\|u\|_\infty + \|u - u^{(l)}\|_2\right)}{\lambda^\star},$$

which implies that

$$\|u - u^{(l)}\|_2 \le \frac{40\sigma\sqrt{\log n} + 40\sigma\sqrt{n}\|u\|_\infty}{\lambda^\star}, \quad 1 \le l \le n.$$

Now we verify the condition for Davis-Kahan, which is

$$\|M - M^{(l)}\| \le (1 - 1/\sqrt{2})\left(\lambda^{(l)} - \max_{j\ge 2}\left|\lambda_j\left(M^{(l)}\right)\right|\right).$$

It follows that

$$\begin{aligned}
\|M - M^{(l)}\| &\le \|M - M^\star\| + \|M^{(l)} - M^\star\| \\
&= \|E\|_2 + \|E^{(l)}\|_2 \\
&\le 10\sigma\sqrt{n} \le c\lambda^\star \le (1 - 1/\sqrt{2})\left(\lambda^{(l)} - \max_{j\ge 2}\left|\lambda_j\left(M^{(l)}\right)\right|\right).
\end{aligned}$$

Step 2. Show that $\mathbf{u}_{\mathbf{l}}^{(\mathbf{1})} \approx \mathbf{u}_{\mathbf{l}}^\star$. It follows that

$$\begin{aligned}
|u_l^{(l)} - u_l^\star| &= \left|\frac{\lambda^\star}{\lambda^{(l)}} u^{(l)\top} u^\star u_l^\star - u^{\star\top} u^\star u_l^\star\right| \\
&\le \left|u_l^\star\left(\frac{\lambda^\star - \lambda^{(l)}}{\lambda^{(l)}} u^{\star\top} u^{(l)}\right)\right| + \left|u_l^\star u^{\star\top}(u^{(l)} - u^\star)\right| \\
&\le |u_l^\star|\frac{10\sigma\sqrt{n}}{\lambda^\star} + |u_l^\star|\frac{10\sigma\sqrt{n}}{\lambda^\star} \\
&\le \frac{20\sigma\sqrt{n}}{\lambda^\star}\|u^\star\|_\infty.
\end{aligned}$$

Step 3. Now it is time to put everything together.

$$\begin{aligned}
\|u - u^\star\|_\infty = \max_l |u_l - u_l^\star| &\le \max_l\left\{|u_l^{(l)} - u_l^\star| + \|u - u^{(l)}\|_2\right\} \\
&\le \frac{20\sigma\sqrt{n}}{\lambda^\star}\|u^\star\|_\infty + \frac{40\sigma\sqrt{\log n} + 40\sigma\sqrt{n}\,\|u\|_\infty}{\lambda^\star} \\
&\le \frac{40\sigma\sqrt{\log n} + 60\sigma\sqrt{n}\,\|u^\star\|_\infty}{\lambda^\star} + \frac{1}{2}\|u - u^\star\|_\infty.
\end{aligned}$$

Thus,

$$\|u - u^\star\|_\infty \le \frac{80\sigma\sqrt{\log n} + 120\sigma\sqrt{n}\,\|u^\star\|_\infty}{\lambda^\star}.$$

This gives an entry level control.