Lecture 11

Last time

- Best low-rank approximation
- Perturbation theory for eigenvalues
- Distances between subspaces

Today

- Distances continued.
- Davis-Kahan $\sin\Theta$ Theorem

Distances and angles between subspaces

Last time we ended with 3 notions of distances between subspaces

1) Distance with optimal rotation

$$\mathrm{dist}_{|||\cdot|||}(U, U^*) = \min_{R \in O(r)} |||UR - U^*|||.$$

e.g., Frobenius, operator

2) Distance between projections

$$|||UU^T - U^*(U^*)^T|||$$

This matrix projects onto $\mathcal{U}$.

3) Pricipal angles

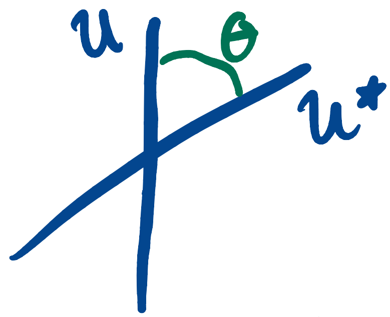
Let $\sigma_1 \geq \ldots \geq \sigma_r \geq 0$ be the singular values of $U^T U^*$. Let $\theta_i = \arccos \sigma_i$ and let

$$\Theta = \begin{bmatrix} \theta_1 & & \\ & \ddots & \\ & & \theta_r \end{bmatrix}, \quad \sin\Theta = \begin{bmatrix} \sin\theta_1 & & \\ & \ddots & \\ & & \sin\theta_r \end{bmatrix}.$$

We measure

$$|||\sin\Theta|||.$$

Interpretation: The angle between a couple of 1D subspaces is clearly defined



What's the angle between a couple of 2D subspaces?

The idea is to define a set of angles instead of just one angle.

The first principle angle

$$\theta_1 = \min_{\substack{u \in \mathcal{U} \cap \mathbb{S}^{d-1} \\ v \in \mathcal{U}^* \cap \mathbb{S}^{d-1}}} \measuredangle(u,v) \qquad (\alpha)$$

$$= \min_{\substack{u \in \mathcal{U} \cap \mathbb{S}^{d-1} \\ v \in \mathcal{U}^* \cap \mathbb{S}^{d-1}}} \arccos(\langle u,v \rangle)$$

Let u_1 and v_1 be the minimizers in (α),
Note that they minimize $\arccos(\langle u, v\rangle)$ iff they maximize $\langle u, v\rangle$

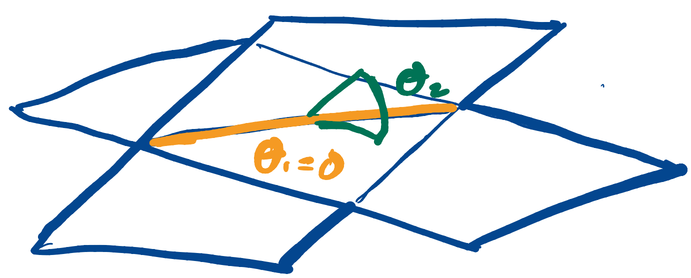

$\cos\theta$.

Exercise: Convince yourself that u_1, v_1 are the top singular vectors of $U^T U^\star$.

Then, we can inductively define

$$\theta_2 = \min_{\substack{u \in (\mathcal{U} \setminus \mathrm{span}(u_1)) \cap \mathbb{S}^{d-1} \\ v \in (\mathcal{U}^\star \setminus \mathrm{span}\{v_1\}) \cap \mathbb{S}^{d-1}}} \arccos(\langle u, v\rangle)$$

Pictorially



Comparison between different "distances"

For our purposes any of these three distances are the "same."

Lemma 4 (Lemmas 2.5 & 2.6 in Chen, Chi, Fan & Ma 2021) Let $U \in \mathbb{R}^{n\times r}$ and $U^\star \in \mathbb{R}^{n\times r}$. Then,

$$\|UU^T - U^\star U^{\star T}\|_{op} = \|\sin\Theta\|_{op} = \|U_\perp^T U^\star\|_{op} = \|U^T U_\perp^\star\|_{op}$$

$$\frac{1}{\sqrt{2}}\|UU^T - U^*U^{*T}\| = \|\sin\Theta\|_F = \|U_\perp^T U^*\|_F = \|U^T U_\perp^*\|_F$$

and, further

$$\|UU^T - U^*U^{*T}\|_{op} \le \min_{R\in O(r)} \|UR - R^*\|_{op} \le \sqrt{2}\|UU^T - U^*U^{*T}\|_{op}$$

$$\frac{1}{\sqrt{2}}\|UU^T - U^*U^{*T}\|_F \le \min_{R\in O(r)} \|UR - R^*\|_F \le \|UU^T - U^*U^{*T}\|_F.$$

$\dashv$

Thus, controlling any would be fine.

Davis-Kahan $\sin\Theta$ theorem

Going back to perturbation analysis for eigenvectors, suppose we have $M = M^* + E \in S^n$ with eigendecompositions:

$$M^* = \sum_{i=1}^{n} \lambda_i^* u_i^* u_i^{*T} = [U^* \ U_\perp^*] \begin{bmatrix} \Lambda^* & \\ & \Lambda_\perp^* \end{bmatrix} \begin{bmatrix} U^{*T} \\ U_\perp^{*T} \end{bmatrix},$$

$(\heartsuit)$

$$M = \sum_{i=1}^{n} \lambda_i u_i u_i^T = [U \ U_\perp] \begin{bmatrix} \Lambda & \\ & \Lambda_\perp \end{bmatrix} \begin{bmatrix} U^T \\ U_\perp^T \end{bmatrix},$$

where $U^* = [u_1^*, \dots, u_r^*]$, $U_\perp^* = [u_{r+1}^*, \dots, u_n^*]$, $\Lambda^* = \mathrm{diag}(\lambda_1^*, \dots, \lambda_r^*)$, and $\Lambda_\perp^* = \mathrm{diag}(\lambda_{r+1}^*, \dots, \lambda_n^*)$

The matrices U, $U_\perp$, Λ and $\Lambda_\perp$ are defined analogously.

Warning: Unlike for eigenvalues, eigenvectors are stable only if their associated eigenvalues are sufficiently far apart.

Example: Consider

$$M^* = \begin{bmatrix} 1+\varepsilon & \\ & 1-\varepsilon \end{bmatrix}, \quad E = \begin{bmatrix} -\varepsilon & \varepsilon \\ \varepsilon & \varepsilon \end{bmatrix}, \quad M = \begin{bmatrix} 1 & \varepsilon \\ \varepsilon & 1 \end{bmatrix}$$

for $\varepsilon \in (0,1)$. Then, a routine computation yields that the leading eigenvectors of M^* and M are

$$u_1^* = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and} \quad u_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Thus, regardless of ε, we have

$$\|u_1 u_1^T - u_1^* u_1^{*T}\|_{op} = \frac{1}{2}\left(\sqrt{3} + \sqrt{2} - 1\right) \approx 1.07$$

$$\|u_1 u_1^T - u_1^* u_1^{*T}\|_F = \sqrt{3 - \sqrt{2}} \approx 1.26$$

The issue here is that M^* is too close to having a 2D eigenspace associated with the top eigenvalue and a small perturbation can move eigenvectors "far away." Thus, we need some control on $\lambda_1(M^*) - \lambda_2(M^*)$. $\dashv$

The following is a seminal result due to Davis and Kahan.

Theorem (Davis-Kahan $\sin\Theta$) Consider M and M^* as in (B). Suppose $\exists$ scalars $a \le b$ and $\Delta > 0$ s.t. any of the following two conditions hold

1) $\{\lambda_1^*, \ldots, \lambda_r^*\} \subseteq [a, b]$, and
$\{\lambda_{r+1}, \ldots, \lambda_n\} \subseteq (-\infty, a-\Delta] \cup [b+\Delta, \infty)$.

2) $\{\lambda_1^*, \ldots, \lambda_r^*\} \subseteq (-\infty, a-\Delta] \cup [b+\Delta, \infty)$, and
$\{\lambda_{r+1}, \ldots, \lambda_n\} \subseteq [a, b]$.

Then,

$$\mathrm{dist}_{op}(U, U^*) \le \sqrt{2}\, \|\sin\Theta\|_{op} \le \sqrt{2}\, \frac{\|E\|_{op}}{\Delta};$$

$$\mathrm{dist}_F(U, U^*) \le \sqrt{2}\, \|\sin\Theta\|_F \le \sqrt{2r}\, \frac{\|E\|_{op}}{\Delta}.$$

$\dashv$

This statement is a bit inconvenient because the eigenvalues of $M = M^* + E$ depend implicitly on both M^* and E.

The following corollary requires more explicit control on $\|E\|_{op}$.

Corollary (☺): Consider M and M^* as in (B). Assume

$$|\lambda_1^*| \geq \dots \geq |\lambda_n^*| \quad \text{and}$$
$$|\lambda_1| \geq \dots \geq |\lambda_n|.$$

Further, suppose

$$\|E\|_{op} < \left(1 - 1/\sqrt{2}\right)\underbrace{\left(|\lambda_r^*| - |\lambda_{r+1}^*|\right)}_{\Delta :=}.$$

Then,

$$\text{dist}_{op}(U, U^*) \leq \sqrt{2}\,\|\sin \Theta\|_{op} \leq \frac{2\|E\|_{op}}{\Delta}$$

$$\text{dist}_F(U, U^*) \leq \sqrt{2}\,\|\sin \Theta\|_F \leq \frac{2\sqrt{r}\,\|E\|_{op}}{\Delta}.$$

+