MATH5473 Homework 1

Lai Yanming

February 20, 2023

5. (a) Let $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$ be unit 2-norm vectors that satisfy $Ax = \sigma y$ with $\sigma = \|A\|_2$. From Theorem 2.5.1 there exists $V_2 \in \mathbb{R}^{n\times(n-1)}$ and $U_2 \in \mathbb{R}^{m\times(m-1)}$ so $V = \begin{bmatrix} x & V_2 \end{bmatrix} \in \mathbb{R}^{n\times n}$ and $U = \begin{bmatrix} y & U_2 \end{bmatrix} \in \mathbb{R}^{m\times m}$ are orthogonal. It is not hard to show that $U^T AV$ has the following structure:

$$U^T AV = \begin{bmatrix} \sigma & w^T \\ 0 & B \end{bmatrix} \equiv A_1.$$

   Since

$$\left\| A_1 \left(\begin{bmatrix} \sigma \\ w \end{bmatrix} \right) \right\|_2^2 \geq \left(\sigma^2 + w^T w\right)^2$$

   we have $\|A_1\|_2^2 \geq \left(\sigma^2 + w^T w\right)$. But $\sigma^2 = \|A\|_2^2 = \|A_1\|_2^2$, and so we must have $w = 0$. An obvious induction argument completes the proof of the theorem.

   (b) Since $U^T A_k V = \operatorname{diag}(\sigma_1, \ldots, \sigma_k, 0, \ldots, 0)$ it follows that $\operatorname{rank}(A_k) = k$ and that $U^T (A - A_k) V = \operatorname{diag}(0, \ldots, 0, \sigma_{k+1}, \ldots, \sigma_p)$ and so $\|A - A_k\|_2 = \sigma_{k+1}$.
   Now suppose $\operatorname{rank}(B) = k$ for some $B \in \mathbb{R}^{m\times n}$. It follows that we can find orthonormal vectors $x_1, \ldots, x_{n-k}$ so $\operatorname{null}(B) = \operatorname{span}\{x_1, \ldots, x_{n-k}\}$. A dimension argument shows that

$$\operatorname{span}\{x_1, \ldots, x_{n-k}\} \cap \operatorname{span}\{v_1, \ldots, v_{k+1}\} \neq \{0\}.$$

   Let z be a unit 2-norm vector in this intersection. Since $Bz = 0$ and

$$Az = \sum_{i=1}^{k+1} \sigma_i \left(v_i^T z\right) u_i$$

   we have

$$\|A - B\|_2^2 \geq \|(A - B)z\|_2^2 = \|Az\|_2^2 = \sum_{i=1}^{k+1} \sigma_i^2 \left(v_i^T z\right)^2 \geq \sigma_{k+1}^2$$

   completing the proof of the theorem.

(d) Suppose that A has SVD $A = U\Sigma V^T$. Denote $\widetilde{A} = QAZ$. Then $\widetilde{A} = QU\Sigma V^T Z = (QU)\Sigma(Z^T V)^T$. Hence $\widetilde{A}$ has the same singular values as A. By definition we conclude that $\|\widetilde{A}\|_p = \|A\|_p$. Hence the Schatten p-norm is unitarily invariant.

(e) We first prove that for $A, B \in \mathbb{R}^{m\times n}$, $\|A - B\| \geq \|\Sigma(A) - \Sigma(B)\|$ for any unitarily invariant norm $\|\cdot\|$. Let $q = \min\{m, n\}$. Use (7.3.7) to identify the singular values of A

$$\sigma_1(A) \geq \cdots \geq \sigma_q(A) \geq 0$$

with the first q nonpositive eigenvalues of the Hermitian matrix

$$\tilde{A} = \begin{bmatrix} 0 & A \\ A^* & 0 \end{bmatrix} \in M_{m+n}$$

of which the $m + n$ ordered eigenvalues are

$$-\sigma_1(A) \leq -\sigma_2(A) \leq \cdots \leq -\sigma_q(A) \leq 0 = \cdots = 0 \leq \sigma_q(A) \leq \cdots \leq \sigma_1(A)$$

and similarly for $\tilde{B}$ and $\tilde{A} - \tilde{B}$. The differences of the ordered eigenvalues of $\tilde{A}$ and $\tilde{B}$ are $\pm[\sigma_1(A) - \sigma_1(B)], \ldots, \pm[\sigma_q(A) - \sigma_q(B)]$ together with 0 ($|m - n|$ times). Although it is not clear how to order this sequence in general, the q smallest elements in an ordering of this sequence are $\{-|\sigma_i(A) - \sigma_i(B)|\}$, and Lemma (7.4.50) applied to $\tilde{A}, \tilde{B}$, and $\tilde{A} - \tilde{B}$ assures us that

$$\sum_{i=1}^{k} -\sigma_i(A - B) \leq \min\left\{\sum_{j=1}^{k} -\left|\sigma_{i_j}(A) - \sigma_{i_j}(B)\right| : 1 \leq i_1 < \cdots < i_k \leq n\right\}$$

for $k = 1, \ldots, q$, which is equivalent to

$$\sum_{i=1}^{k} \sigma_i(A - B) \geq \max\left\{\sum_{j=1}^{k} \left|\sigma_{i_j}(A) - \sigma_{i_j}(B)\right| : 1 \leq i_i < \cdots < i_k \leq n\right\}$$

$k = 1, \ldots, q$. Since $\{|\sigma_i(A) - \sigma_i(B)|\}$ is the set of singular values of $\Sigma(A) - \Sigma(B)$, Corollary (7.4.47) guarantees that $\|A - B\| \geq \|\Sigma(A) - \Sigma(B)\|$ for any unitarily invariant norm $\|\cdot\|$.

If B has rank k, then $\sigma_1(B) \geq \cdots \geq \sigma_k(B) > 0 = \sigma_{k+1}(B) = \cdots = \sigma_q(B)$. Thus,

$$\begin{aligned}
\|A - B\| &\geq \|\Sigma(A) - \Sigma(B)\| \\
&= \|\mathrm{diag}\,(\sigma_1(A) - \sigma_1(B), \ldots, \sigma_k(A) - \sigma_k(B), \sigma_{k+1}(A), \ldots, \sigma_q(A))\| \\
&\geq \|\mathrm{diag}\,(0, \ldots, 0, \sigma_{k+1}(A), \ldots, \sigma_q(A))\|
\end{aligned}$$

where we have used the fact that a unitarily invariant norm on diagonal matrices is a monotone norm because it is a symmetric gauge function of the diagonal entries. On the other hand, supposing SVD of A is $A = U\Sigma V^T$ and $A_k = U\Sigma_k V^T$, we have

$$\begin{aligned}\|A - A_k\| &= \left\|U\Sigma V^T - U\Sigma_k V^T\right\| = \|\Sigma - \Sigma_k\| \\ &= \|\mathrm{diag}\,(0, \ldots, 0, \sigma_{k+1}(A), \ldots, \sigma_q(A))\|\end{aligned}$$

where we use the definition of unitarily invariant norm. Hence

$$\|A - A_k\| = \min_{\mathrm{rank}(B)=k} \|A - B\|$$

(f) For any orthogonal matrix R,

$$\|A - R\|_F^2 = \mathrm{tr}((A-R)^T(A-R)) = \mathrm{tr}(A^TA) + \mathrm{tr}(I) - 2\mathrm{tr}(R^TA)$$

Hence minimizing $\|A - R\|_F$ over orthogonal matrix R is equivalent to maximizing $\mathrm{tr}(R^TA)$. Denote $R^* = UV^T$. We have

$$\begin{aligned}\mathrm{tr}(R^TA) &= \mathrm{tr}(R^TR^*R^{*T}A) = \mathrm{tr}(R^TR^*VU^TU\Sigma V^T) \\ &= \mathrm{tr}(R^TR^*V\Sigma V^T) = \mathrm{tr}(R_1A_1A_1^T)\end{aligned}$$

where $R_1 = R^TR^*$, $A_1 = V\Sigma^{1/2}$. Let a_i be the ith column of A_1. Then

$$\mathrm{tr}(R_1A_1A_1^T) = \mathrm{tr}(A_1^TR_1A_1) = \sum_i a_i^T(R_1)a_i$$

But by the Schwarz inequality,

$$\sum_i a_i^T(R_1)a_i \leq \sum_i \sqrt{(a_i^Ta_i)(a_i^TR_1^TR_1a_i)} = \sum_i a_i^Ta_i = \mathrm{tr}(A_1A_1^T)$$

Thus

$$\mathrm{tr}(R_1A_1A_1^T) \leq \mathrm{tr}(A_1A_1^T)$$

Therefore

$$\mathrm{tr}(R^TA) = \mathrm{tr}(R_1A_1A_1^T) \leq \mathrm{tr}(A_1A_1^T) = \mathrm{tr}(V\Sigma V^T) = \mathrm{tr}(VU^TU\Sigma V^T) = \mathrm{tr}(R^{*T}A)$$

Hence R^* is the maximizer of $\mathrm{tr}(R^TA)$ and minimizer of $\|A - R\|_F$.
For the problem of minimizing $\sum_i \|Rp_i - q_i\|^2$, we have

$$\sum_i \|Rp_i - q_i\|^2 = \sum_i (Rp_i - q_i)^T(Rp_i - q_i) = \sum_i (p_i^Tp_i + q_i^Tq_i - 2q_i^TRp_i)$$

Hence minimizing $\sum_i \|Rp_i - q_i\|^2$ is equivalent to maximizing $\sum_i q_i^TRp_i$. We further derive that

$$\sum_i q_i^TRp_i = \mathrm{tr}\left(\sum_i Rp_iq_i^T\right) = \mathrm{tr}\left(RA_2^T\right)$$

where $A_2 = \sum_i q_ip_i^T$. Assume that SVD of A_2 is $A_2 = U_2\Sigma_2V_2^T$. Then from the above discussion, we know that the maximizer of $\sum_i q_i^TRp_i$, and hence the minimizer of $\sum_i \|Rp_i - q_i\|^2$, is $R^* = U_2V_2^T$.