Notes on Matrix Analysis

Yang Xu
xuyang1014@pku.edu.cn

December 27, 2023

Abstract

The note is based on [Tao11]. I recommend section 1.3 of this book as a fundamental material since it emphasises the geometric intuition. Studying matrix analysis without learning geometric intuition is like playing games without playing Genshin Impact. This leads to being trapped in a pile of mathematical symbols and overlooking their profound implications.

1 The Geometry of Eigenvalues

Question: Given $0 \preceq \mathbf{A} \preceq \mathbf{B}$, how will you prove that $\lambda_i(\mathbf{A}) \le \lambda_i(\mathbf{B})$ for all i?

1.1 mini-max formula of eigenvalue

Theorem 1.1 (Courant-Fischer min-max theorem)**.** *Let A be an $n \times n$ Hermitian matrix. Then we have*

$$\lambda_i(\mathbf{A}) = \sup_{\dim(V)=i} \underline{f}(V) \overset{\text{def}}{=} \sup_{\dim(V)=i} \inf_{v \in V: |v|=1} v^* \mathbf{A} v \tag{1.1}$$

and

$$\lambda_{n-i+1}(\mathbf{A}) = \inf_{\dim(V)=i} \bar{f}(V) \overset{\text{def}}{=} \inf_{\dim(V)=i} \sup_{v \in V: |v|=1} v^* \mathbf{A} v. \tag{1.2}$$

1.2 Partial Trace

Definition 1.2 (Partial Trace)**.** *Given an $n \times n$ Hermitian matrix $\mathbf{A}$ and an m-dimensional subspace V of $\mathbb{C}^n$, we difine the partial trace*

$$\operatorname{tr}(\mathbf{A} \downharpoonright_V) \overset{\text{def}}{=} \sum_{i=1}^{m} v_i^* \mathbf{A} v_i, \tag{1.3}$$

where $v_1, \dots v_m$ is any orthonormal basis of V.

Theorem 1.3 (Extremal partial trace)**.** *Given an $n \times n$ Hermitian matrix $\mathbf{A}$. Then for any $1 \le k \le n$, one has*

$$\sum_{j=n-k+1}^{n} \lambda_j(\mathbf{A}) \le \operatorname{tr}(\mathbf{A} \downharpoonright_V) \le \sum_{j=1}^{k} \lambda_j(\mathbf{A}). \tag{1.4}$$

It has a direct corollary as follows, since $a_{i_1 i_1} + \cdots + a_{i_k i_k} = \operatorname{tr}(\mathbf{A} \downharpoonright_{\langle e_{i_1}, \dots, e_{i_k} \rangle})$.

Corollary 1.4 (Schur-Horn inequalities)**.**

$$\sum_{j=n-k+1}^{n} \lambda_j(\mathbf{A}) \le a_{i_1 i_1} + \cdots + a_{i_k i_k} \le \sum_{j=1}^{k} \lambda_j(\mathbf{A}). \tag{1.5}$$

The result can also be derived by an application of the Birkhoff's theorem for doubly stochastic matrices, implying that $\operatorname{diag}(\mathbf{A})$ *lies in the permutahedron of* $\lambda_1 \dots \lambda_n$.

Theorem 1.5 (Schur-Horn theorem, 1954)**.** *The orbit of* $\operatorname{diag}(\mathbf{UAU}^*)$ *is exactly the permutahedron of* $\lambda_1 \dots \lambda_n$.

proof sketch. It suffices to consider the $n = 2$ case. □

Since we provide a geometric description for the sum of top-k eigenvalues, we can naturally prove the Ky Fan inequality.

Theorem 1.6 (Ky Fan inequality)**.**

$$\sum_{j=1}^{k} \lambda_j(\mathbf{A}+\mathbf{B}) \leq \sum_{j=1}^{k} \lambda_j(\mathbf{A}) + \sum_{j=1}^{k} \lambda_j(\mathbf{B}) \tag{1.6}$$

Proof. Since

$$\operatorname{tr}(\mathbf{A}+\mathbf{B} \downharpoonleft_V) = \operatorname{tr}(\mathbf{A} \downharpoonleft_V) + \operatorname{tr}(\mathbf{B} \downharpoonleft_V) \leq \sum_{j=1}^{k} \lambda_j(\mathbf{A}) + \sum_{j=1}^{k} \lambda_j(\mathbf{B}) \tag{1.7}$$

holds for any subspace V. Let V maximize the left hand side give the desired result. □

The extremal partial trace provides a geometric description the sum of top-k eigenvalues. Such description can be generalized to the sum of any k eigenvalues

Theorem 1.7 (Wielandt minimax formula)**.** *Let* $1 \leq i_1 < \cdots i_k \leq n$ *be integers. Define a partial flag to be a nested collection* $V_1 \subset \cdots \subset V_k$ *of subspaces of* $\mathbb{C}^n$ *such that* $\dim(V_j) = i_j$ *for all* $1 \leq j \leq k$. *Define the associated* Schubert variety $X(V_1, \dots, V_k)$ *to be the collection of all* k*-dimensional subspaces* W *such that* $\dim(W \cap V_j) \leq j$. *Then for any* $n \times n$ *Hermitian matrix* $\mathbf{A}$,

$$\lambda_{i_1}(\mathbf{A}) + \cdots + \lambda_{i_k}(\mathbf{A}) = \sup_{(V_1,\dots,V_k)\in\mathcal{V}_{i_1,\dots,i_k}} \inf_{W\in X(V_1,\dots,V_k)} \operatorname{tr}(\mathbf{A} \downharpoonleft_W). \tag{1.8}$$

Proof. **Part 1:** We prove that

$$\lambda_{i_1}(\mathbf{A}) + \cdots + \lambda_{i_k}(\mathbf{A}) \geq \inf_{W\in X(V_1,\dots,V_k)} \operatorname{tr}(\mathbf{A} \downharpoonleft_W) \tag{1.9}$$

for $\forall (V_1, \dots, V_k) \in \mathcal{V}_{i_1,\dots,i_k}$. First we simplify the case to $i_1 = 1$. Second we find $W' \in X(V_2, \dots, V_k)$ that minimize $\operatorname{tr}(\mathbf{A} \downharpoonleft_{W'})$ which is at most $\lambda_{i_2}(\mathbf{A}) + \cdots + \lambda_{i_k}(\mathbf{A})$(using induction on k). Thirdly, we check and find the minimal j such that $\dim(W' \cap V_j) < j$, then we add $v \in \mathcal{P}_{W'^{\perp}} V_j$ into W' to be W, which is the construction we require.

Part 2: Let $\mathbf{A}$ be diagonalized and $V_j = \langle e_1, \dots, e_{i_j} \rangle$. Now we prove that

$$\operatorname{tr}(\mathbf{A} \downharpoonleft_W) \geq \lambda_{i_1}(\mathbf{A}) + \cdots + \lambda_{i_k}(\mathbf{A}) \tag{1.10}$$

for $\forall W \in X(V_1, \dots, V_k)$. All we have to do is construct an orthogonal basis $\{v_1, \dots, v_k\}$ of W such that $v_j \in W \cap V_j \cap \operatorname{span}(v_1, \dots, v_{j-1})^{\perp}$, that yields 1.10. □

Now we can establish the Lidskii inequality and the dual Lidskii inequality

Theorem 1.8 (Lidskii inequality)**.**

$$\begin{aligned}
&\lambda_{n-k+1}(\mathbf{B}) + \cdots + \lambda_n(\mathbf{B}) \\
\leq &\lambda_{i_1}(\mathbf{A}+\mathbf{B}) + \cdots + \lambda_{i_k}(\mathbf{A}+\mathbf{B}) - (\lambda_{i_1}(\mathbf{A}) + \cdots + \lambda_{i_k}(\mathbf{A})) \\
\leq &\lambda_1(\mathbf{B}) + \cdots + \lambda_k(\mathbf{B})
\end{aligned} \tag{1.11}$$

2 Eigenvalue Inequalities

Theorem 2.1 (Weyl inequality)**.**

$$\lambda_{i+j-1}(\mathbf{A}+\mathbf{B}) \leq \lambda_i(\mathbf{A}) + \lambda_j(\mathbf{B}) \tag{2.1}$$

for $i, j \geq 1$ *and* $i + j - 1 \leq n$*.*

Theorem 2.2 (Low rank approximation, Eckart-Young-Mirsky theorem)**.** *Assume a Hermitian matrix* $\mathbf{A}$ *with eigen-decomposition*

$$\mathbf{A} = \sum_{i=1}^{n} \lambda_i u_i u_i^\top$$

and let $\mathbf{A}_k$ *be obtained by truncating the eigen-decomposition of* $\mathbf{A}$ *at the k-th term:*

$$\mathbf{A}_k := \sum_{i=1}^{k} \lambda_i u_i u_i^\top.$$

Then we have

$$\|\mathbf{A} - \mathbf{A}_k\| = \min_{\operatorname{rank}(\mathbf{A}') \leq k} \|\mathbf{A} - \mathbf{A}'\|. \tag{2.2}$$

Moreover, a similar statement holds for any unitary-invariant norm (such as Frobenius norm, Schatten norm etc.).

Proof. Using Weyl inequality, we have

$$\lambda_i(\mathbf{A} - \mathbf{A}') \geq \lambda_{i+k}(\mathbf{A}) - \lambda_{k+1}(\mathbf{A}') = \lambda_{i+k}(\mathbf{A}) = \lambda_i(\mathbf{A} - \mathbf{A}_k) \tag{2.3}$$

for $i = 1, \ldots, n-k$. Hence we can observe that $\|\mathbf{A} - \mathbf{A}'\| \geq \|\mathbf{A} - \mathbf{A}_k\|$ for any unitary-invariant norm. □

Theorem 2.3 (Cauchy interlacing law)**.** *For any* $n \times n$ *Hermitian matrix* $\mathbf{A}_n$ *with top left* $(n-1) \times (n-1)$ *minor* $\mathbf{A}_{n-1}$*, then*

$$\lambda_{i+1}(\mathbf{A}_n) \leq \lambda_i(\mathbf{A}_{n-1}) \leq \lambda_i(\mathbf{A}_n) \tag{2.4}$$

Proof. Using the Courant-Fischer min-max theorem, we have that

$$\begin{aligned}
\lambda_i(\mathbf{A}_{n-1}) &= \sup_{\dim(V)=i, V \subset \langle e_1, \ldots e_{n-1} \rangle} \inf_{v \in V : |v|=1} v^* \mathbf{A} v \\
&\leq \sup_{\dim(V)=i} \inf_{v \in V : |v|=1} v^* \mathbf{A} v \\
&= \lambda_i(\mathbf{A}_n).
\end{aligned} \tag{2.5}$$

and similarly for the other side. □