EE378B Inference, Estimation, and Information Processing

Variational principles and eigenvalues inequalities

Andrea Montanari *Lecture 1-2 - Due on 1/19/2021 4PM*

Homework should be submitted via Gradescope, by Monday afternoon: the code will be communicated by an announcement on Canvas.

For getting credit for the class, you are required to present solutions of some of these homeworks during the first 15 minutes of class starting on 1/20. Please, sign up for (at least) one slot, and be sure that your explanation lasts 15 minutes (or less). For these presentations, you are free to choose whatever format you prefer (slides, typed notes, handwriting, ...).

This week, the presentations will be:

- Monday 1/18: Holiday.
- Wednesday 1/20: Problem 2, points (a), (b), (c).

Problem #1: Operator norm

Remember that the operator norm of $\boldsymbol{M} \in \mathbb{R}^{m\times n}$ is defined as

$$\|\boldsymbol{M}\|_{\rm op} := \max\big\{\|\boldsymbol{M}\boldsymbol{x}\|_2 : \|\boldsymbol{x}\|_2 = 1\big\} \tag{1}$$

where, for a vector $\boldsymbol{x} \in \mathbb{R}^n$, $\|\boldsymbol{x}\|_2 \equiv (\sum_{i=1}^n x_i^2)^{1/2}$. The objective of these exercise is to derive various properties of the operator norm from this definition.

(a) Prove that the operator norm admits the following two alternative characterizations

$$\|\boldsymbol{M}\|_{\rm op} = \max\big\{\langle \boldsymbol{x}, \boldsymbol{M}\boldsymbol{y}\rangle \ : \ \|\boldsymbol{x}\|_2 = \|\boldsymbol{y}\|_2 = 1\big\} \tag{2}$$

$$= \sigma_1(\boldsymbol{M})\,, \tag{3}$$

where $\sigma_1(\boldsymbol{M}) \ge \sigma_2(\boldsymbol{M}) \ge \cdots \ge \sigma_n(\boldsymbol{M})$ are the singular values of $\boldsymbol{M}$.

(b) Prove that $\|\boldsymbol{M}\|_{\rm op} = \|\boldsymbol{M}^{\sf T}\|_{\rm op}$ and $\|\boldsymbol{A}\boldsymbol{B}\|_{\rm op} \le \|\boldsymbol{A}\|_{\rm op}\|\boldsymbol{B}\|_{\rm op}$.

(c) Prove that $\|\cdot\|_{\rm op}$ is indeed a norm, i.e. that it enjoys the following properties: (i) $\|a\boldsymbol{M}\|_{\rm op} = |a|\,\|\boldsymbol{M}\|_{\rm op}$ for any scalar $a \in \mathbb{R}$; (ii) $\|\boldsymbol{A}+\boldsymbol{B}\|_{\rm op} \le \|\boldsymbol{A}\|_{\rm op} + \|\boldsymbol{B}\|_{\rm op}$; (iii) if $\|\boldsymbol{M}\|_{\rm op} = 0$, then $\boldsymbol{M} = 0$.

Problem #2: Wielandt's variational principle

The objective of this problem is to prove a generalization of the variational characterization of eigenvalues of a symmetric matric. Let $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ be a symmetric matrix with eigenvalues $\lambda_1(\boldsymbol{A}) \ge \cdots \ge \lambda_n(\boldsymbol{A})$. A partial flag in $\mathbb{R}^n$ with signature $(i_1, \dots, i_k)$ is a sequence of subspaces $V_1 \subset V_2 \subseteq \cdots \subseteq V_k$ such that, for all $1 \le j \le k$, $\dim(V_j) = i_j$. We denote by $\mathcal{F}(i_1, \dots, i_k)$ the set of flags with that signature, and by $\mathcal{S}_k(V_1, \dots, V_k)$ the set of subspaces W of dimension k such that $\dim(W \cap V_j) = j$ for each $j \in \{1, \dots, k\}$.

$$\lambda_{i_1}(\boldsymbol{A}) + \cdots + \lambda_{i_k}(\boldsymbol{A}) = \sup_{(V_1,\dots,V_k)\in\mathcal{F}(i_1,\dots,i_k)} \ \inf_{W\in\mathcal{S}_k(V_1,\dots,V_k)} {\rm tr}(\boldsymbol{A}|_W)\,. \tag{4}$$

For a subspace W, $\dim(W) = k$, $\mathrm{tr}(\boldsymbol{A}|_W)$ is defined by choosing an orthonormal basis $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k$ of W and setting

$$\mathrm{tr}(\boldsymbol{A}|_W) := \sum_{j=1}^{k} \langle \boldsymbol{u}_j, \boldsymbol{A}\boldsymbol{u}_j \rangle \,. \tag{5}$$

In the following we will denote by $\mathcal{R}(\boldsymbol{A}; i_1, \dots, i_k)$ the right-hand side of Eq. (4).

(a) Prove that the value of $\mathrm{tr}(\boldsymbol{A}|_W)$ does not depend on the choice of the orthonormal basis in W.

(b) Let $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n$ be the eigenvectors of $\boldsymbol{A}$ corresponding to $\lambda_1(\boldsymbol{A}), \dots, \lambda_n(\boldsymbol{A})$. Set $V_j = \mathrm{span}(\boldsymbol{v}_1, \dots, \boldsymbol{v}_{i_j})$. Prove that, for any $W \in \mathcal{S}_k(V_1, \dots, V_k)$, $\mathrm{tr}(\boldsymbol{A}|_W) \geq \lambda_{i_1}(\boldsymbol{A}) + \mathrm{tr}(\boldsymbol{A}|_{W'})$, where $\dim(W') = k-1$ and $\dim(W' \cap V_j) = j-1$ for $j \geq 2$ (i.e. $W' \in \mathcal{S}_{k-1}(V_2, \dots, V_k)$).

(c) Continuing from the previous point, deduce, by induction over k, that for any W, $\mathrm{tr}(\boldsymbol{A}|_W) \geq \lambda_{i_1}(\boldsymbol{A}) + \cdots + \lambda_{i_k}(\boldsymbol{A})$, and therefore $\mathcal{R}(\boldsymbol{A}; i_1, \dots, i_k) \geq \lambda_{i_1}(\boldsymbol{A}) + \cdots + \lambda_{i_k}(\boldsymbol{A})$.

(d) Fix $(V_1, \dots, V_k) \in \mathcal{F}(i_1, \dots, i_k)$. Prove that there exists $W \in \mathcal{S}(V_1, \dots, V_k)$ such that $\mathrm{tr}(\boldsymbol{A}|_W) \leq \lambda_{i_1}(\boldsymbol{A}) + \mathrm{tr}(\boldsymbol{A}|_{W'})$, where $\dim(W') = k-1$ and $W' \in \mathcal{S}_{k-1}(V_2, \dots, V_k)$.
[Hint: Let $U_1 := \mathrm{span}(\boldsymbol{v}_{i_1}, \dots . \boldsymbol{v}_n)$. Then $\dim(U_1 \cap V_1) \geq 1$. Choose $\boldsymbol{u} \in U_1 \cap V_1$ of unit norm.]

(e) **[Optional!]** Continuing from the previous point, deduce that there exists W such that $\mathrm{tr}(\boldsymbol{A}|_W) \leq \lambda_{i_1}(\boldsymbol{A}) + \cdots + \lambda_{i_k}(\boldsymbol{A})$.

Problem #3: Lidskii's inequality

Use Eq. (4) to prove Lidskii's inequality: for two symmetric matrices $\boldsymbol{A}$, $\boldsymbol{B} \in \mathbb{R}^{n \times n}$,

$$\lambda_{i_1}(\boldsymbol{A}+\boldsymbol{B}) + \cdots + \lambda_{i_k}(\boldsymbol{A}+\boldsymbol{B}) \leq \lambda_{i_1}(\boldsymbol{A}) + \cdots + \lambda_{i_k}(\boldsymbol{A}) + \lambda_1(\boldsymbol{B}) + \cdots + \lambda_k(\boldsymbol{B}) \,.$$

[Hint: Apply Eq. (4) to $\boldsymbol{A}+\boldsymbol{B}$ and set $V_1, \dots, V_k$ achieving equality for this matrix.]