ee364m Exercise Set 2

Due date: January 24, 11:59pm

Question 2.1 (Convex functions on symmetric matrices)**:** Let $\mathcal{S}^n$ denote the $n \times n$ symmetric matrices, so that $X \in \mathcal{S}^n$ has eigenvalue decomposition $X = U \operatorname{diag}(\lambda) U^T$ for an orthogonal $U \in \mathbb{R}^{n \times n}$ and $\lambda \in \mathbb{R}^n$. Von-Neumann's trace inequality states that if $\lambda(X) = (\lambda_1(X), \ldots, \lambda_n(X)) \in \mathbb{R}^n$ denotes the (sorted) eigenvalues of the matrix X, $\lambda_1(X) \geq \cdots \geq \lambda_n(X)$, then

$$\operatorname{tr}(XY) \leq \lambda(X)^T \lambda(Y) \tag{2.1}$$

with equality if and only if there exists an orthogonal U such that $X = U \operatorname{diag}(\lambda(X)) U^T$ and $Y = U \operatorname{diag}(\lambda(Y)) U^T$.

We use inequality (2.1) to characterize all (closed) convex unitarily invariant matrix functions. A function $f : \mathbb{R}^n \to \mathbb{R}$ is permutation symmetric if $f(x) = f(\Pi x)$ for all permutation matrices $\Pi \in \{0, 1\}^{n \times n}$. A function $F : \mathcal{S}^n \to \mathbb{R}$ is *unitarily invariant* if

$$F(X) = F(UXU^T)$$

for all orthogonal matrices $U \in \mathbb{R}^{n \times n}$. Given such a function F, it is immediate that F may depend only on the unsorted eigenvalues of X, that is, there exists a permutation symmetric f such that

$$F(X) = f(\lambda(X)).$$

Here, you show the converse using conjugate duality, and use this to give (sub)gradients of $f_{\mathcal{S}}$.

(a) Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a (permutation) symmetric function, and define is matrix extension

$$f_{\mathcal{S}}(X) := f(\lambda(X)).$$

Use von-Neumann's trace inequality (2.1) to show that

$$(f_{\mathcal{S}})^* = (f^*)_{\mathcal{S}}, \tag{2.2}$$

that is, the convex conjugate of $f_{\mathcal{S}}$ is the matrix extension of the convex conjugate f^* of f.

(b) Conclude that if $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is a closed convex function, then

$$f_{\mathcal{S}} = (f_{\mathcal{S}})^{**}$$

and so $f_{\mathcal{S}}$ is a closed convex unitarily invariant function.

(c) Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be permutation symmetric. Using the Fenchel-Young inequality, show that

$$G \in \partial f_{\mathcal{S}}(X)$$

if and only if there exists an orthogonal $U \in \mathbb{R}^{n \times n}$ such that

$$\lambda(G) \in \partial f(\lambda(X)) \text{ and } G = U \operatorname{diag}(\lambda(G)) U^T \text{ and } X = U \operatorname{diag}(\lambda(X)) U^T.$$

Some additional quite optional questions (for fun)

Question 2.2: If a closed convex f is differentiable at x, then $\partial f(x) = \{\nabla f(x)\}$. Use this and Question 2.1 to show that

$$\nabla \log \det(X) = X^{-1}$$

for $X \succ 0$.

Question 2.3: Using that the Birkhoff polytope

$$\mathcal{P}_n := \left\{ P \in \mathbb{R}_+^{n \times n} \mid P\mathbf{1} = \mathbf{1} \text{ and } P^T\mathbf{1} = \mathbf{1} \right\},$$

that is, the collection of $n \times n$ doubly stochastic matrices, is the convex hull of the $n \times n$ permutation matrices, prove the von-Neumann trace inequality (2.1). *Hint.* The following lemma (which you should attempt to prove) may be useful.

Lemma 2.3.1. *Let $x, y \in \mathbb{R}^n$ satisfy $x_1 \geq \cdots \geq x_n$ and $y_1 \geq \cdots \geq y_n$. Let $\Pi \in \{0,1\}^{n \times n}$ be a permutation matrix. Then $x^T \Pi y \leq x^T y$ with equality if and only if there exists a permutation matrix Q with $Qx = x$ and $Q\Pi y = y$.*