ECE562, Fall 2020: Problem Set #5

Due Dec 9, 2020

1. **Majorization and Schur-convexity**

   Many nonconvex optimization problems that involve matrix-valued variables (e.g., those arising in MIMO communications) can be transformed into simple problems with scalar variables by relying on *majorization theory*. Majorization makes explicit the notion of *spread* of vector elements, i.e., defines in a specific way the notion of the elements of a vector $\mathbf{x}$ being 'less spread out' than the elements of a vector $\mathbf{y}$. An interesting instance of majorization theory is the *Lorenz curve* in economics, which characterizes the distribution of total wealth in a population.

   For $\mathbf{x} = [x_1, x_2, \ldots, x_n] \in \mathbb{R}^n$ (row vector) let $x_{[1]} \geq \cdots \geq x_{[n]}$ and $x_{(1)} \leq \cdots \leq x_{(n)}$ be the components of $\mathbf{x}$ in decreasing and increasing orders, respectively. For $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ we say that $\mathbf{y}$ *majorizes* $\mathbf{x}$, denoted by $\mathbf{x} \prec \mathbf{y}$, if $\sum_{i=1}^{k} x_{[i]} \leq \sum_{i=1}^{k} y_{[i]}, 1 \leq k < n$ and $\sum_{i=1}^{n} x_{[i]} = \sum_{i=1}^{n} y_{[i]}$ or equivalently, if $\sum_{i=1}^{k} x_{(i)} \geq \sum_{i=1}^{k} y_{(i)}, 1 \leq k < n$ and $\sum_{i=1}^{n} x_{(i)} = \sum_{i=1}^{n} y_{(i)}$. Alternatively, $\mathbf{x} \prec \mathbf{y}$ if and only if $\mathbf{x} = \mathbf{y}\mathbf{P}$ for some doubly stochastic matrix $\mathbf{P}$ [Hardy, Littlewood, Pólya, 1929]. Intuitively, $\mathbf{x} \prec \mathbf{y}$ implies that $\mathbf{x}$ is more *mixed* than $\mathbf{y}$. $\prec$ is a preordering on $\mathbb{R}^n$ and a proper partial ordering on $\mathcal{D} = \{\mathbf{x} : x_1 \geq x_2 \geq \cdots \geq x_n\}$. Functions that preserve the ordering of majorization are said to be *Schur-convex*. More precisely, a real-valued function g defined on a set $\mathcal{C} \subseteq \mathbb{R}^n$ is said to be *Schur-convex* on $\mathcal{C}$ if $\mathbf{x} \prec \mathbf{y}$ on $\mathcal{C} \Rightarrow g(\mathbf{x}) \leq g(\mathbf{y})$ and *Schur-concave* on $\mathcal{C}$ if $\mathbf{x} \prec \mathbf{y}$ on $\mathcal{C} \Rightarrow g(\mathbf{x}) \geq g(\mathbf{y})$. Clearly, g is Schur-concave if and only if $-g$ is Schur-convex.

   (a) Show that if $\mathbf{P}_1$ and $\mathbf{P}_2$ are doubly stochastic, then so is $\mathbf{P} = \mathbf{P}_1\mathbf{P}_2$.

   (b) Show that if $\mathbf{x}\mathbf{P} \prec \mathbf{x}, \forall \mathbf{x} \in \mathbb{R}^n$, then $\mathbf{P}$ is doubly stochastic.

   (c) For $\mathbf{x} \in \mathbb{R}^n$, let $d(\mathbf{x}) = x_{[1]} - x_{[n]}$. If $\mathbf{P}$ is a doubly stochastic matrix such that $P_{ij} \geq \alpha, \forall (i,j)$, prove that $d(\mathbf{x}\mathbf{P}) \leq (1-2\alpha)d(\mathbf{x})$.

   (d) Let $\mathbf{P}_1, \mathbf{P}_2, \ldots$ be a sequence of $n \times n$ doubly stochastic matrices and define the sequence $\mathbf{x}_j = \mathbf{x}_{j-1}\mathbf{P}_j$ for some initial vector $\mathbf{x}_0 \in \mathbb{R}^n$. Clearly, $\mathbf{x}_0 \succ \mathbf{x}_1 \succ \cdots$. Moreover, $\mathbf{x}_j \succ \bar{x}\mathbf{1}, \forall j \geq 0$, where $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$ and $\mathbf{1}$ is the all-ones vector. Show that if all elements of every $\mathbf{P}_j$ are $\geq \alpha$, then $\lim_{j\to\infty} \mathbf{x}_j = \bar{x}\mathbf{1}$.

   (e) Let $\mathbf{A}$ be an $n \times n$ Hermitian matrix with vector of diagonal entries $\mathrm{diag}(\mathbf{A})$. Suppose that the vector of the corresponding eigenvalues is denoted by $\boldsymbol{\lambda}$. Show that $\mathrm{diag}(\mathbf{A}) \prec \boldsymbol{\lambda}$.

   (f) Let $\mathcal{I} \subset \mathbb{R}$ be an interval and $g : \mathcal{I} \to \mathbb{R}$ be a convex function. Prove that $f(\mathbf{x}) = \sum_{i=1}^{2} g(x_i)$ is Schur-convex on $\mathcal{I}^2$.