EE 150 - Applications of Convex Optimization in Signal Processing and Communications
Dr. Andre Tkacenko, JPL
Third Term 2011-2012

Homework Set #4

Due on **Thursday, May 3** in class.

1. (10 points) (Adapted from CO-BV, Exercise 3.49) (*Log-concave functions:*) Show that the following functions are log-concave.

(a) *Logistic function:*

$$f(x) = \frac{e^x}{1+e^x}\,,\ \text{dom}(f(x)) = \mathbb{R}\,.$$

(b) *Harmonic mean:*

$$f(\mathbf{x}) = \frac{1}{\frac{1}{x_1} + \cdots + \frac{1}{x_n}}\,,\ \text{dom}(f(\mathbf{x})) = \mathbb{R}^n_{++}\,.$$

(c) *Product over sum:*

$$f(\mathbf{x}) = \frac{\prod_{k=1}^{n} x_k}{\sum_{k=1}^{n} x_k}\,,\ \text{dom}(f(\mathbf{x})) = \mathbb{R}^n_{++}\,.$$

(d) *Determinant over trace:*

$$f(\mathbf{X}) = \frac{\det(\mathbf{X})}{\text{tr}(\mathbf{X})}\,,\ \text{dom}(f(\mathbf{X})) = \mathbb{S}^n_{++}\,.$$

Hint: The *Cauchy-Schwarz inequality* may be useful here, which states that for any inner product $\langle \cdot, \cdot \rangle$, we have

$$|\langle \mathbf{x}, \mathbf{y} \rangle|^2 \leq \langle \mathbf{x}, \mathbf{x} \rangle \cdot \langle \mathbf{y}, \mathbf{y} \rangle\ ,$$

with equality if and only if $\mathbf{x}$ and $\mathbf{y}$ are linearly dependent, which is equivalent to saying that either $\mathbf{y} = C\mathbf{x}$ or $\mathbf{x} = C\mathbf{y}$ for some $C \in \mathbb{F}$.

2. (10 points) (Adapted from CO-BV, Exercise 4.7) (*Convex-affine/convex-concave fractional functions:*)

(a) *Convex-affine fractional functions:* Consider a problem of the following form:

$$\begin{array}{ll} \text{minimize} & \dfrac{f_0(\mathbf{x})}{\mathbf{c}^T\mathbf{x} + d} \\ \text{subject to} & f_k(\mathbf{x}) \leq 0\,,\ k = 1, \ldots, m \\ & \mathbf{A}\mathbf{x} = \mathbf{b} \end{array}\ ,$$

where $f_0(\mathbf{x})\,, \mathbf{f}_1(\mathbf{x})\,, \ldots, f_m(\mathbf{x})$ are convex, and the domain of the objective function is defined as $\left\{\mathbf{x} \in \text{dom}(f_0(\mathbf{x})) : \mathbf{c}^T\mathbf{x} + d > 0\right\}$.

i. Show that this is a quasiconvex optimization problem.

ii. Show that the problem is equivalent to

$$\begin{array}{ll} \text{minimize} & g_0(\mathbf{y},t) \\ \text{subject to} & g_k(\mathbf{y},t) \leq 0\,,\ k=1,\ldots,m \\ & \mathbf{A}\mathbf{y} = \mathbf{b}t \\ & \mathbf{c}^T\mathbf{y} + td = 1 \end{array},$$

where $g_k(\mathbf{x},t)$ is the perspective of $f_k(\mathbf{x})$, i.e., $g_k(\mathbf{x},t) = tf_k(\mathbf{x}/t)$. The variables are $\mathbf{y} \in \mathbb{R}^n$ and $t \in \mathbb{R}$. Show that this problem is convex.

(b) *Convex-concave fractional functions:* Analogous to the problem analyzed above, consider a problem of the form:

$$\begin{array}{ll} \text{minimize} & \dfrac{f_0(\mathbf{x})}{h(\mathbf{x})} \\ \text{subject to} & f_k(\mathbf{x}) \leq 0\,,\ k=1,\ldots,m \\ & \mathbf{A}\mathbf{x} = \mathbf{b} \end{array},$$

where $f_0(\mathbf{x})\,,\mathbf{f}_1(\mathbf{x})\,,\ldots,f_m(\mathbf{x})$ are convex, $h(\mathbf{x})$ is concave, and the domain of the objective function is defined as $\{\mathbf{x} \in \text{dom}(f_0(\mathbf{x})) \cap \text{dom}(h(\mathbf{x})) : h(\mathbf{x}) > 0\}$ and $f_0(\mathbf{x}) \geq 0$ everywhere.

i. Show that this is a quasiconvex optimization problem.

ii. Show that the problem is equivalent to

$$\begin{array}{ll} \text{minimize} & g_0(\mathbf{y},t) \\ \text{subject to} & g_k(\mathbf{y},t) \leq 0\,,\ k=1,\ldots,m \\ & \mathbf{A}\mathbf{y} = \mathbf{b}t \\ & \widetilde{h}(\mathbf{y},t) \leq -1 \end{array},$$

where $g_k(\mathbf{x},t)$ is the perspective of $f_k(\mathbf{x})$ and $\widetilde{h}(\mathbf{x},t)$ is the perspective of $-h(\mathbf{x})$. Show that this problem is convex.

iii. As an example, apply the technique derived in the previous part to the unconstrained problem with

$$f_0(\mathbf{x}) = \frac{1}{m}\text{tr}(\mathbf{F}(\mathbf{x}))\ ,\ h(\mathbf{x}) = (\det(\mathbf{F}(\mathbf{x})))^{\frac{1}{m}}\ ,$$

with $\text{dom}\left(\frac{f_0(\mathbf{x})}{h(\mathbf{x})}\right) = \{\mathbf{x} : \mathbf{F}(\mathbf{x}) \succ \mathbf{0}\}$, where $\mathbf{F}(\mathbf{x}) = \mathbf{F}_0 + x_1\mathbf{F}_1 + \cdots + x_n\mathbf{F}_n$ for given $\mathbf{F}_k \in \mathbb{S}^m$. In this problem, we minimize the ratio of the arithmetic mean over the geometric mean of the eigenvalues of an affine matrix function $\mathbf{F}(\mathbf{x})$.

3. (10 points) (Adapted from CO-BV, Exercise 4.24) (*Complex ℓ_1- ℓ_2-, and ℓ_∞-norm approximation:*) Consider the problem

$$\text{minimize} \quad ||\mathbf{A}\mathbf{x} - \mathbf{b}||_p\ ,$$

where $\mathbf{A} \in \mathbb{C}^{m\times n}$, $\mathbf{b} \in \mathbb{C}^m$, and the variable is $\mathbf{x} \in \mathbb{C}^n$. Recall that the (real or complex) ℓ_p-norm is defined by

$$||\mathbf{y}||_p \triangleq \left(\sum_{k=1}^{m} |y_k|^p\right)^{\frac{1}{p}}$$

for $p \geq 1$, and $||\mathbf{y}||_\infty = \max_{k=1,\ldots,m} |y_k|$. For $p = 1$, 2, and ∞, express the complex ℓ_p-norm approximation problem as a *second-order cone program* (SOCP) or *quadratically constrained quadratic program* (QCQP) with real variables and data.

4. (10 points) (Adapted from CO-AE, Exercise 3.11) (*Using Schur complements to express matrix-based optimization problems as semidefinite programs:*) Formulate each of the following optimization problems as a semidefinite program (SDP). The variable is $\mathbf{x} \in \mathbb{R}^n$ and the function $\mathbf{F}(\mathbf{x})$ is defined as

$$\mathbf{F}(\mathbf{x}) = \mathbf{F}_0 + x_1\mathbf{F}_1 + \cdots + x_n\mathbf{F}_n\,,$$

with $\mathbf{F}_k \in \mathbb{S}^m$. The domain of $f(\mathbf{x})$ in each subproblem is $\mathrm{dom}(f(\mathbf{x})) = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{F}(\mathbf{x}) \succ \mathbf{0}\}$.

(a) Minimize $f(\mathbf{x}) = \mathbf{c}^T\left(\mathbf{F}(\mathbf{x})\right)^{-1}\mathbf{c}$, where $\mathbf{c} \in \mathbb{R}^m$.

(b) Minimize $f(\mathbf{x}) = \max_{k=1,\ldots,K} \mathbf{c}_k^T\left(\mathbf{F}(\mathbf{x})\right)^{-1}\mathbf{c}_k$, where $\mathbf{c}_k \in \mathbb{R}^m$ for $k = 1, \ldots, K$.

(c) Minimize $f(\mathbf{x}) = \sup_{||\mathbf{c}||_2 \leq 1} \mathbf{c}^T\left(\mathbf{F}(\mathbf{x})\right)^{-1}\mathbf{c}$.

(d) Minimize $f(\mathbf{x}) = E\left[\mathbf{c}^T\left(\mathbf{F}(\mathbf{x})\right)^{-1}\mathbf{c}\right]$, where $\mathbf{c}$ is a random vector with mean $\overline{\mathbf{c}} = E[\mathbf{c}]$ and covariance $\mathbf{S} = E\left[(\mathbf{c} - \overline{\mathbf{c}})(\mathbf{c} - \overline{\mathbf{c}})^T\right]$. Assume here that the covariance matrix has the following representation: $\mathbf{S} = \sum_{k=1}^{m} \mathbf{s}_k\mathbf{s}_k^T$, where $\mathbf{s}_k \in \mathbb{R}^m$ for $k = 1, \ldots, m$.

***5.** (30 points) (Adapted from CO-AE, Exercise 2.19) (*Majorization and symmetric functions of eigenvalues:*) Here, we use $\mathbf{x}_{[k]}$ to denote the k-th largest element of a vector $\mathbf{x} \in \mathbb{R}^n$. Thus, $\mathbf{x}_{[1]}, \mathbf{x}_{[2]}, \ldots, \mathbf{x}_{[n]}$ are the elements of $\mathbf{x}$ sorted in decreasing order. We say that a vector $\mathbf{y} \in \mathbb{R}^n$ *majorizes* a vector $\mathbf{x} \in \mathbb{R}^n$ if

$$\begin{aligned}
\mathbf{y}_{[1]} &\geq \mathbf{x}_{[1]}\,, \\
\mathbf{y}_{[1]} + \mathbf{y}_{[2]} &\geq \mathbf{x}_{[1]} + \mathbf{x}_{[2]}\,, \\
\mathbf{y}_{[1]} + \mathbf{y}_{[2]} + \mathbf{y}_{[3]} &\geq \mathbf{x}_{[1]} + \mathbf{x}_{[2]} + \mathbf{x}_{[3]}\,, \\
&\vdots \\
\mathbf{y}_{[1]} + \mathbf{y}_{[2]} + \cdots + \mathbf{y}_{[n-1]} &\geq \mathbf{x}_{[1]} + \mathbf{x}_{[2]} + \cdots + \mathbf{x}_{[n-1]}\,, \\
\mathbf{y}_{[1]} + \mathbf{y}_{[2]} + \cdots + \mathbf{y}_{[n]} &= \mathbf{x}_{[1]} + \mathbf{x}_{[2]} + \cdots + \mathbf{x}_{[n]}\,.
\end{aligned}$$

In other words, the descending-ordered partial sums of $\mathbf{y}$ are greater than or equal to those of $\mathbf{x}$ and the sum of the components of $\mathbf{y}$ and $\mathbf{x}$ are equal.

(a) It can be shown that $\mathbf{y}$ majorizes $\mathbf{x}$ if and only if there exists a *doubly stochastic matrix* $\mathbf{P}$ such that $\mathbf{x} = \mathbf{P}\mathbf{y}$. A doubly stochastic matrix is one with nonnegative entries whose rows and columns add up to unity:

$$P_{k,\ell} \geq 0\,,\ k, \ell = 1, \ldots, n\,,\ \mathbf{P}\mathbf{1} = \mathbf{1}\,,\ \mathbf{1}^T\mathbf{P} = \mathbf{1}^T\,.$$

Use this characterization to show the following: if $f : \mathbb{R} \to \mathbb{R}$ is a convex function and $\mathbf{y}$ majorizes $\mathbf{x}$, then

$$\sum_{k=1}^{n} f(y_k) \geq \sum_{k=1}^{n} f(x_k)\,.$$

(b) We use the notation $\lambda_k(\mathbf{X})$ to denote the k-th largest eigenvalue of a matrix $\mathbf{X} \in \mathbb{S}^n$, so that $\lambda_1(\mathbf{X}), \ldots, \lambda_n(\mathbf{X})$ are the eigenvalues of $\mathbf{X}$ sorted in decreasing order. Let r be any integer in the set $\{1, \ldots, n\}$. Show that

$$\lambda_1(\mathbf{X}) + \cdots + \lambda_r(\mathbf{X}) = \sup \{\operatorname{tr}(\mathbf{XZ}) : \mathbf{Z} \in \mathbb{S}^n\,,\ \mathbf{0} \preceq \mathbf{Z} \preceq \mathbf{I}\,,\ \operatorname{tr}(\mathbf{Z}) = r\}\ . \tag{1}$$

What does this tell us about the convexity properties of the function $g_r(\mathbf{X}) = \lambda_1(\mathbf{X}) + \cdots + \lambda_r(\mathbf{X})$ (i.e., the sum of the largest r eigenvalues of $\mathbf{X}$)?

Hint: Show that the right-hand side of (1) is equal to the left-hand side by using an eigenvalue decomposition of $\mathbf{X}$ to reduce the maximization in (1) to a simple linear program whose solution is the sum of the r largest eigenvalues of $\mathbf{X}$.

(c) Let $\mathbf{X} = \theta \mathbf{U} + (1-\theta)\,\mathbf{V}$ be a convex combination of two matrices $\mathbf{U}, \mathbf{V} \in \mathbb{S}^n$. Use the results of the previous part to show that the vector

$$\mathbf{a} \triangleq \theta \begin{bmatrix} \lambda_1(\mathbf{U}) \\ \vdots \\ \lambda_n(\mathbf{U}) \end{bmatrix} + (1-\theta) \begin{bmatrix} \lambda_1(\mathbf{V}) \\ \vdots \\ \lambda_n(\mathbf{V}) \end{bmatrix}$$

majorizes the vector $\mathbf{b} \triangleq \begin{bmatrix} \lambda_1(\mathbf{X}) & \cdots & \lambda_n(\mathbf{X}) \end{bmatrix}^T$.

(d) Combine the results of parts (a) and (c) to show that if $f : \mathbb{R} \to \mathbb{R}$ is convex, then the function $h : \mathbb{S}^n \to \mathbb{R}$ defined as

$$h(\mathbf{X}) = \sum_{k=1}^{n} f(\lambda_k(\mathbf{X}))$$

is convex.

For example, by taking $f(x) = x\log(x)$, we can conclude that the function $h(\mathbf{X}) = \sum_{k=1}^{n} \lambda_k(\mathbf{X}) \log(\lambda_k(\mathbf{X}))$ is convex on $\mathbb{S}^n_{++}$. This function arises in quantum information theory where it is known as the (negative) Von Neumann entropy. When $\mathbf{X}$ is diagonal, i.e., $\mathbf{X} = \operatorname{diag}(\mathbf{x})$, it reduces to the negative Shannon entropy $\sum_k x_k \log x_k$.

Reading assignments:

1. Read through Chapter 3 and begin Chapter 4 of CO-BV.

Reminders:

Late homework policy for EE 150: Late homeworks will not be accepted. There will be no exceptions to this other than institute established emergency reasons, in which case a signed letter is required from an authorized official.

NCT Problems: Remember that problems with an asterisk, such as ***7** are no collaboration type

(NCT) problems.

Texts: The abbreviation CO-BV corresponds to the textbook "*Convex Optimization*" by Stephen Boyd and Lieven Vandenberghe. In addition, CO-AE refers to the *Additional Exercises for Convex Optimization*, also by Boyd and Vandenberghe. Finally, CVX corresponds to the `cvx` *Users' Guide* by Michael Grant and Stephen Boyd.
