ELEG5481 Signal Processing Optimization Techniques
Tutorial Solution 5

Mar. 4, 2013

Q1. Please point out what are the flaws in the following arguments.

(a) Question: Is the set $C = \{x \mid a_1^T x \leq b_1 \text{ or } a_1^T x \leq b_2\}$ convex?
No. The union of two convex is not convex.

(b) Question: Show that the set $V = \{x \in \mathbf{R}^n \mid \|x - x_0\| \leq \|x - x_i\|, i = 1, \ldots, n\}$ is a polyhedron.
Argument: Because $\|x - x_0\| \leq \|x - x_i\| \Longrightarrow (x_0 - x_i)^T x \geq \frac{1}{2}(x_0^T x_0 - x_i^T x_i)$ for $i = 1, \ldots, n$, we have

$$V = \{x \mid (x_0 - x_i)^T x \geq \tfrac{1}{2}(x_0^T x_0 - x_i^T x_i), i = 1, \ldots, n\}.$$

Therefore V is a polyhedron.

(c) Given a polyhedron $P = \{x \mid Ax \preceq b\}$ with nonempty interior, show how to find $x_0, \ldots, x_K$ such that P is $V(x_0)$.
Argument: Let x_i be a point in the interior of P. Denote the ith row of A as a_i^T. Let x_i be the mirror image of x_0 with respect to the hyperplane $a_i^T x = b$.

(d) Question: Show the following set is convex: $C = \{x \mid B(x, a) \subset S\}$, where $a \geq 0$, S is a convex set, and $B(x, a) = \{y \mid \|y - x\| \leq a\}$.
Argument: Let $x_1, x_2 \in C$, and $0 \leq \theta \leq 1$. We have $\|y - (\theta x_1 + (1 - \theta)x_2)\| \leq a$. (The remaining argument is skipped.)

Solution:
(a) C can be convex in some cases. A correct answer could be "C is not convex generally".
(b) This only shows shat V is a subset of $\{x \mid (x_0 - x_i)^T x \geq \frac{1}{2}(x_0^T x_0 - x_i^T x_i), i = 1, \ldots, n\}$.
(c) You need to verify that $x_0, \ldots, x_K$ are correct.
(d) y is not defined.

Q2. Prove the Hadamard's inequality

$$\mathbf{det} P \leq \prod_{i=1}^{n} P_{ii},$$

where $P \in \mathbf{S}_+^n$.

(a) Step1: (*Cholesky decomposition*) Show that a PD matrix P can be decomposed as $P = R^T R$, where R is an upper triangular matrix with positive diagonal entries.
(b) Step2: Use the Cholesky decomposition to prove that the Hadamard's inequality is true for a PD matrix P.
(c) Step3: Show that the Hadamard's inequality is true for a non-PD symetric matrix P.

Solution:
(a) Show by induction. For $n = 1$, this is obviously true. Suppose this is true for n, we show this is true for $n + 1$. We partition P as

$$P = \begin{bmatrix} P_{11} & P_{12} \\ P_{12}^T & P_{22} \end{bmatrix},$$

where $P_{11} \in \mathbf{R}^{n\times n}$, $P_{12} \in \mathbf{R}^n$ and $P_{22} \in \mathbf{R}$. Since P_{11} is PD, we can decompose $P_{11} = R_{11}^T R_{11}$ where R_{11} is upper triangular with positive diagonals. We claim that the following R defined as

$$R = \begin{bmatrix} R_{11} & R_{11}^{-T} P_{12} \\ 0 & \sqrt{P_{22} - P_{12}^T P_{11}^{-1} P_{12}} \end{bmatrix}$$

is the Cholesky factor we are looking for. We first need to show that R is well defined. As R_{11}^T is lower triangular with positive diagonals, R_{11} is invertible and $R_{11}^{-T} P_{12}$ is well defined. We need $P_{22} - P_{12}^T P_{11}^{-1} P_{12} > 0$ as well. This can be seen from

$$P_{22} - P_{12}^T P_{11}^{-1} P_{12} = \begin{bmatrix} -P_{11}^{-1} P_{12} \\ 1 \end{bmatrix}^T \begin{bmatrix} P_{11} & P_{12} \\ P_{12}^T & P_{22} \end{bmatrix}^T \begin{bmatrix} -P_{11}^{-1} P_{12} \\ 1 \end{bmatrix} > 0$$

We now verify $P = R^T R$.

$$R^T R = \begin{bmatrix} R_{11}^T & 0 \\ P_{12}^T R_{11}^{-1} & \sqrt{P_{22} - P_{12}^T P_{11}^{-1} P_{12}} \end{bmatrix} \begin{bmatrix} R_{11} & R_{11}^{-T} P_{12} \\ 0 & \sqrt{P_{22} - P_{12}^T P_{11}^{-1} P_{12}} \end{bmatrix} = \begin{bmatrix} R_{11}^T R_{11} & P_{12} \\ P_{12}^T & P_{22} \end{bmatrix} = P.$$

(b) Let $P = R^T R$ be the Cholesky decomposition. We write $R = [r_1, \dots, r_n]$. Then we have $r_i^T r_i = P_{ii}$, which implies that $r_{ii}^2 \le P_{ii}$. Since R is an upper triangular matrix, we have $\mathbf{det} R = \prod_{i=1}^n r_{ii}$. Therefore,

$$\mathbf{det} P = (\mathbf{det} R)^2 = (\prod_{i=1}^n r_{ii})^2 \le \prod_{i=1}^n P_{ii}.$$

(c) If P is not PD but PSD, then $\mathbf{det} P = 0$. We also have $\prod_{i=1}^n P_{ii} \ge 0$.

Q3. Show that

$$f(x) = \prod_{k=1}^n x_k^{\alpha_k}, \quad \mathbf{dom}\, f = \mathbf{R}_{++}^n,$$

is concave, where α_k are nonegative numbers with $\sum_k \alpha_k = 1$. Hint:

(a) Show the following result first: For a symmetric matrix $A \in \mathbf{R}^{n\times n}$, if A is diagonally dominant (i.e. $|A_{i,i}| \ge \sum_{j\ne i} |A_{j,i}|$, for $i = 1, \dots, n$) and if $A_{i,i} > 0$ for $i = 1, \dots, n$, then A is positive semidefinite.

(b) Compute the Hessian of $f(x)$ and show that it is negative semidefinite.

Solution: We prove a slightly stronger result. We assume that A is Hermitian. Then for any $x \in \mathbf{R}^n$,

$$\begin{aligned} x^H A x &= \sum_{i=1}^n |x_i|^2 A_{ii} + 2 \sum_{i>j} \mathcal{R}\{A_{ij} x_i^* x_j\} \\ &\ge \sum_{i=1}^n \left(|x_i|^2 \sum_{j\ne i} |A_{ij}| \right) + 2 \sum_{i>j} \mathcal{R}\{A_{ij} x_i^* x_j\} \\ &= \sum_{i>j} (|x_i|^2 |A_{ij}| + |x_j|^2 |A_{ji}|) + 2 \sum_{i>j} \mathcal{R}\{A_{ij} x_i^* x_j\} \\ &\ge \sum_{i>j} |A_{ij}| (|x_i|^2 + |x_j|^2) - 2 \sum_{i>j} |A_{ij}| |x_i| |x_j| \\ &\ge 0. \end{aligned}$$

Now let us show that $f(x)$ is concave by second order condition.

$$\begin{aligned}\frac{\partial f}{\partial x_i} &= \alpha_i x_i^{-1} \prod_{k=1}^{n} x_k^{\alpha_k} \\ \frac{\partial^2 f}{\partial x_i \partial x_j} &= \alpha_i \alpha_j x_i^{-1} x_j^{-1} \prod_{k=1}^{n} x_k^{\alpha_k}, \quad i \neq j \\ \frac{\partial^2 f}{\partial x_i^2} &= \alpha_i(\alpha_i - 1) x_i^{-2} \prod_{k=1}^{n} x_k^{\alpha_k}\end{aligned}$$

If we let $q = x^{-1}$, then

$$\begin{aligned}\nabla^2 f(x) &= \left(\prod_{k=1}^{n} x_k^{\alpha_k}\right)\big(\mathrm{diag}(q)\alpha\alpha^T \mathrm{diag}(q) - \mathrm{diag}(q)\mathrm{diag}(\alpha)\mathrm{diag}(q)\big) \\ &= \left(\prod_{k=1}^{n} x_k^{\alpha_k}\right)\big(\mathrm{diag}(q)(\alpha\alpha^T - \mathrm{diag}(\alpha))\mathrm{diag}(q)\big)\end{aligned}$$

If we have $\alpha\alpha^T - \mathrm{diag}(\alpha) \preceq 0$, then $\nabla^2 f(x) \preceq 0$, and we are done. So let us show that $\mathrm{diag}(\alpha) - \alpha\alpha^T \succeq 0$. But $\mathrm{diag}(\alpha) - \alpha\alpha^T$ is diagonally dominant, as $(\alpha_i - \alpha_i^2) - \sum_{j \neq i} \alpha_i \alpha_j = \alpha_i - \alpha_i(\sum_{j=1}^{n} \alpha_j) = 0$, hence $\mathrm{diag}(\alpha) - \alpha\alpha^T$ is PSD.