ELEG5481 Signal Processing Optimization Techniques
Tutorial Solution 9

Mar. 31, 2013

Q1. (a) Suppose that we are given a set of discrete points $\{x_i\}_{i=1}^{K+L}$. Show that there exists an hyperplane $a^T x = b$ that strictly separates $\{x_i\}_{i=1}^{K}$ from $\{x_i\}_{i=K+1}^{K+L}$, i.e.,

$$a^T x_i < b, \quad i = 1, \ldots, K, \qquad a^T x_i > b, \quad i = K+1, \ldots, K+L,$$

if and only if there exist two parallel hyperplanes $\tilde{a}^T x = \tilde{b} - 1$ and $\tilde{a}^T x = \tilde{b} + 1$ which separate $\{x_i\}_{i=1}^{K}$ from $\{x_i\}_{i=K+1}^{K+L}$ in the sense that

$$\tilde{a}^T x_i \leq \tilde{b} - 1, \quad i = 1, \ldots, K, \qquad \tilde{a}^T x_i \geq \tilde{b} + 1 \quad i = K+1, \ldots, K+L.$$

(b) Suppose that we are given $K + L$ ellipsoids

$$\epsilon_i = \{P_i u + q_i \mid \|u\|_2 \leq 1\}, \quad i = 1, \ldots, K+L,$$

where $P_i \in \mathbf{S}^n$. We are interested in finding a hyperplane that strictly separates $\epsilon_1, \ldots, \epsilon_K$ from $\epsilon_{K+1}, \ldots, \epsilon_{K+L}$, i.e. we want to compute $a \in \mathbf{R}^n$, $b \in \mathbf{R}$ such that

$$a^T x + b < 0 \text{ for } x \in \epsilon_1 \cup \ldots \cup \epsilon_K, \quad a^T x + b > 0 \text{ for } x \in \epsilon_{K+1} \cup \ldots \cup \epsilon_{K+L},$$

or prove that no such hyperplane exists. Express this problem as an SOCP feasibility problem.

Solution:

(a) Suppose that there exists an hyperplane $a^T x = b$ that strictly separate S_1 and S_2, i.e.,

$$a^T x_i < b, \quad i = 1, \ldots, K, \qquad a^T x_i > b \quad i = K+1, \ldots, K+L,$$

Then there exists an $\delta > 0$ such that

$$a^T x_i \leq b - \delta, \quad i = 1, \ldots, K, \qquad a^T x_i \geq b + \delta \quad i = K+1, \ldots, K+L,$$

or equivalently

$$(a/\delta)^T x_i \leq b/\delta - 1, \quad i = 1, \ldots, K, \qquad (a/\delta)^T x_i \geq b/\delta + 1 \quad i = K+1, \ldots, K+L,$$

Then it suffices to take $\tilde{a} = a/\epsilon$ and $\tilde{b} = b/\epsilon$.
Conversely, given $\tilde{a}$ and $\tilde{b}$, it suffices to take $a = \tilde{a}$ and $b = \tilde{b}$.

(b) **Step 1**. We first show that there exists a hyperplane $a^T x = b$ that strictly separates $\epsilon_1, \ldots, \epsilon_K$ from $\epsilon_{K+1}, \ldots, \epsilon_{K+L}$, i.e. we want to compute $a \in \mathbf{R}^n$, $b \in \mathbf{R}$ such that

$$a^T x < b \text{ for } x \in \epsilon_1 \cup \ldots \cup \epsilon_K, \quad a^T x > b \text{ for } x \in \epsilon_{K+1} \cup \ldots \cup \epsilon_{K+L},$$

if and only if there exist $\tilde{a}$ and $\tilde{b}$ such that

$$\tilde{a}^T x \leq \tilde{b} - 1 \text{ for } x \in \epsilon_1 \cup \ldots \cup \epsilon_K, \quad \tilde{a}^T x \geq \tilde{b} + 1 \text{ for } x \in \epsilon_{K+1} \cup \ldots \cup \epsilon_{K+L}.$$

Suppose such a hyperplane $a^T x = b$ exists. Consider the following problem

$$\begin{aligned} v^\star = \sup_{x \in \mathbf{R}^n} \quad & a^T x \\ \text{s.t.} \quad & x \in \epsilon_1 \cup \ldots \cup \epsilon_K, \end{aligned}$$

This problem can be rewritten as

$$v^\star = \max_{i=1,\dots,K} \sup_{x\in\epsilon_i} \quad a^T x$$

An optimal solution of each of the inner problem is $x_i^\star = P_i P_i^T a_i / \|P_i^T a_i\| + q_i$ if $P_i^T a_i \neq 0$, and $x_i^\star = q_i$ if $P_i^T a_i = 0$. Note that this optimal solution is achievable for each inner problems. Therefore, we have

$$v^\star = \max_{i=1,\dots,K} a^T x_i^\star < b$$

The inequality is due to the fact that each $x_i^\star \in \epsilon_i$ and $b > a^T x$ for all $x \in \epsilon_i$. As a result, we have

$$b > \sup_{x\in\epsilon_1\cup\ldots\cup\epsilon_K} a^T x,$$

Similarly, we have

$$b < \inf_{x\in\epsilon_{K+1}\cup\ldots\cup\epsilon_{K+L}} a^T x.$$

This implies that there exists a $\delta > 0$ such that

$$\begin{aligned} b - \delta &\geq \sup_{x\in\epsilon_1\cup\ldots\cup\epsilon_K} a^T x, \\ b + \delta &\leq \inf_{x\in\epsilon_{K+1}\cup\ldots\cup\epsilon_{K+L}} a^T x. \end{aligned}$$

or equivalently

$$a^T x \leq b - \delta \text{ for } x \in \epsilon_1 \cup \ldots \cup \epsilon_K, \quad a^T x \geq b + \delta \text{ for } x \in \epsilon_{K+1} \cup \ldots \cup \epsilon_{K+L},$$

or equivalently

$$\tilde{a}^T x \leq \tilde{b} - 1 \text{ for } x \in \epsilon_1 \cup \ldots \cup \epsilon_K, \quad \tilde{a}^T x \geq \tilde{b} + 1 \text{ for } x \in \epsilon_{K+1} \cup \ldots \cup \epsilon_{K+L},$$

where $\tilde{a} = a/\delta$ and $\tilde{b} = b/\delta$.
Conversely, given $\tilde{a}$ and $\tilde{b}$, it suffices to take $a = \tilde{a}$ and $b = \tilde{b}$.
Step 2. With the previous result, the feasibility problem can be formalized as

$$\begin{aligned} \text{find} \quad & \tilde{a},\ \tilde{b} \\ \text{s.t.} \quad & \tilde{a}^T x_i \leq \tilde{b} - 1,\ \forall\ x_i \in \epsilon_i, i = 1, \dots, K, \\ & \tilde{a}^T x_i \geq \tilde{b} + 1,\ \forall\ x_i \in \epsilon_i, i = K+1, \dots, K+L. \end{aligned} \tag{1}$$

which can be written as

$$\begin{aligned} \text{find} \quad & \tilde{a},\ \tilde{b} \\ \text{s.t.} \quad & \|P_i^T \tilde{a}_i\|_2 + \tilde{a}^T q_i \leq \tilde{b} - 1, \quad i = 1, \dots, K, \\ & -\|P_i^T \tilde{a}_i\|_2 + \tilde{a}^T q_i \geq \tilde{b} + 1, \quad i = K+1, \dots, K+L, \end{aligned}$$

which is an SOCP feasibility problem.

Q2. Show how to convert the conic form SDP to the standard form SDP.
Conic form:

$$\begin{aligned} \min_{x\in\mathbf{R}^n} \quad & c^T x \\ \text{s.t.} \quad & x_1 F_1 + \ldots + x_n F_n + G \preceq 0, \\ & Ax = b, \end{aligned}$$

where $G, F_1, \ldots, F_n \in \mathbf{S}^k$, and $A \in \mathbf{R}^{m\times n}$.
Standard form:

$$\begin{aligned} \min_{X \in \mathbf{S}^n} \quad & \mathbf{tr}(CX) \\ \text{s.t.} \quad & \mathbf{tr}(A_i X) = b_i,\ i = 1, \ldots, p \\ & X \succeq 0, \end{aligned}$$

where $C,\ A_1, \ldots, A_p \in \mathbf{S}^n$.

Solution: Introduce some PSD matrices $Y_0 \in \mathbf{S}_+^k$, and $Y_i \in \mathbf{S}_+^2$ for $i = 1, \ldots, n$. Then we write the conic form SDP as

$$\begin{aligned} \min_{x, \{Y_i\}_{i=0}^n} \quad & c^T x \\ \text{s.t.} \quad & Ax = b, \\ & -Y_0 = x_1 F_1 + \ldots + x_n F_n + G \\ & x_i = [Y_i]_{12},\ i = 1, \ldots, n \\ & Y_i \succeq 0,\ i = 0, \ldots, n. \end{aligned}$$

We now can remove the variable x.

$$\begin{aligned} \min_{\{Y_i\}_{i=0}^n} \quad & \sum_{i=1}^n c_i [Y_i]_{12} \\ \text{s.t.} \quad & \sum_{i=1}^n A_{ji} [Y_i]_{12} = b_j,\ j = 1, \ldots, m \\ & -Y_0 = [Y_1]_{12} F_1 + \ldots + [Y_n]_{12} F_n + G, \\ & Y_i \succeq 0,\ i = 0, \ldots, m. \end{aligned}$$

Introduce a PSD matrices $Y \in \mathbf{S}_+^{k+2n}$, which has the block diagonal structure of $Y = \mathrm{diag}(Y_0, \ldots, Y_n)$. Note that this block diagonal structure can be imposed by the constraint $\mathbf{tr} A_i Y = 0$ for some A_i. Also note that all the equality constraints can be rewrite in this form as well. Therefore we can rewrite the conic form SDP to a standard form SDP.