ELEG5481 Signal Processing Optimization Techniques
Tutorial Solution 11

Apr. 15, 2013

Q1. Consider the following problem.

$$\begin{aligned} \min_{x} \quad & \|x-y\|_2^2 \\ \text{s.t.} \quad & x \geq a,\ c - b^T x \geq 0. \end{aligned}$$

where $y \in \mathbf{R}^n$, $a, b \in \mathbf{R}_+^n$, and $c \in \mathbf{R}_{++}$ are given data. We assume the problem is strictly feasible.

(a) Write down the KKT conditions.

(b) Let $x^\star$ denote the primal optimal solution and $\sigma^\star$ denote the dual optimal solution associated with the constraint $c - b^T x \geq 0$. From the KKT conditions, derive the relationship between $x^\star$ and $\sigma^\star$.

(c) Devise a fast algorithm to find the primal and dual optimal solutions.

Solution:

(a) Let λ and σ denote the dual variables associated with constraints $x \geq a$ and $c - b^T x \geq 0$, respectively. The KKT conditions are

$$x \geq a, \tag{1}$$
$$c - b^T x \geq 0, \tag{2}$$
$$\lambda \geq 0, \tag{3}$$
$$\sigma \geq 0, \tag{4}$$
$$\lambda_i (x_i - a_i) = 0,\ \forall\ i, \tag{5}$$
$$\sigma(c - b^T x) = 0, \tag{6}$$
$$2(x-y) - \lambda + \sigma b = 0. \tag{7}$$

(b) From Eq.(6), we can see that

$$x = (y - \tfrac{\sigma}{2} b) + \tfrac{1}{2}\lambda.$$

Let us investigate the relationship between x and σ.

Case 1: Suppose $y_i - \frac{\sigma}{2} b_i > a_i$. As $\lambda_i \geq 0$, we have $x_i > a_i$. From Eq.(5), we have $\lambda_i = 0$. Thus $x_i = y_i - \frac{\sigma}{2} b_i$.

Case 2: Suppose $y_i - \frac{\sigma}{2} b_i = a_i$. Then we have $x_i = a_i + \frac{1}{2}\lambda_i$. From Eq.(5), we have $\lambda_i^2 = 0$. Therefore $\lambda_i = 0$ and $x_i = y_i - \frac{\sigma}{2} b_i$.

Case 3: Suppose $y_i - \frac{\sigma}{2} b_i < a_i$. Then $0 \leq x_i - a_i = (y_i - \frac{\sigma}{2} b_i) + \frac{1}{2}\lambda_i - a_i < \frac{1}{2}\lambda_i$. From Eq.(5), we have $x_i = a_i$ and $\lambda_i = 2(a_i - y_i - \frac{\sigma}{2} b_i)$.

Summarizing these three cases, we have

$$x_i = \max\{y_i - \tfrac{\sigma}{2} b_i, a_i\}, \quad i = 1, \ldots, n.$$

and $\lambda_i = \max\{2y_i - \sigma b_i, 2a_i\} - (2y_i - \sigma b_i)$.

(c) The KKT conditions reduce to

$$c - b^T x \geq 0, \tag{8}$$
$$\sigma \geq 0, \tag{9}$$
$$\sigma(c - b^T x) = 0, \tag{10}$$
$$x_i = \max\{y_i - \tfrac{\sigma}{2} b_i, a_i\}, \quad i = 1, \ldots, n. \tag{11}$$

To solve this equations, let us first consider the three cases.

Case 1: Suppose that $c - \sum_{i=1}^n b_i \max\{y_i, a_i\} > 0$. We have $\sigma = 0$, as if $\sigma > 0$, then

$$c - b^T x = c - \sum_{i=1}^n b_i \max\{y_i - \tfrac{\sigma}{2} b_i, a_i\} \geq c - \sum_{i=1}^n b_i \max\{y_i, a_i\} > 0,$$

which violate Eq.(10). We also have

$$x_i = \max\{y_i, a_i\}, \quad i = 1, \dots, n.$$

Case 2: Suppose that $c - \sum_{i=1}^n b_i \max\{y_i, a_i\} = 0$. Then it can be easily seen that $\sigma = 0$ and $x_i = \max\{y_i, a_i\}, \quad i = 1, \dots, n$ is a solution.
Case 3: Suppose $c - \sum_{i=1}^n b_i \max\{y_i, a_i\} < 0$. We must have $\sigma > 0$. For otherwise,

$$0 \leq c - b^T x = c - \sum_{i=1}^n b_i \max\{y_i, a_i\} < 0,$$

which is a contradiction. Therefore, we have $c = b^T x$, or

$$c = \sum_{i=1}^n b_i \max\{y_i - \tfrac{\sigma}{2} b_i, a_i\}.$$

Define $f(\sigma) = \sum_{i=1}^n b_i \max\{y_i - \frac{\sigma}{2} b_i, a_i\}$ with domain $\sigma \geq 0$. We want to solve the equation $f(\sigma) = c$ for σ. Define $f_i(\sigma) = b_i \max\{y_i - \frac{\sigma}{2} b_i, a_i\}$. Then we have

$$f_i(\sigma) = \begin{cases} y_i b_i - \frac{\sigma}{2} b_i^2, & \text{if } \sigma \leq \sigma_i \\ b_i a_i, & \text{if } \sigma > \sigma_i, \end{cases}$$

where $\sigma_i = 2(y_i - a_i)/b_i$. Because $f_i(\sigma)$ is decreasing, by definition of $f(x)$, so is $f(x)$. To solve for $f(x) = c$, we can simply use bisection.
A better method exists by observing that $f_i(\sigma)$ is piece-wise, and so is $f(x)$. Without loss of generality, we assume that $\sigma_1 \leq \sigma_2 \dots \leq \sigma_n$. By definition of $f(\sigma)$ we have, when $\sigma_i \leq \sigma \leq \sigma_{i+1}$ (we define $\sigma_0 = 0$ and $\sigma_{n+1} = \infty$),

$$f(\sigma) = \sum_{j=1}^n f_j(\sigma) = \sum_{j=1}^i f_j(\sigma) + \sum_{j=i+1}^n f_j(\sigma) = \sum_{j=1}^i b_j a_j + \sum_{j=i+1}^n (y_j b_j - \tfrac{\sigma}{2} b_j^2) \tag{12}$$

Our next step is to determine the solution $\sigma^\star$ of $f(\sigma) = c$ lies in which region $[\sigma_{i'}, \sigma_{i'+1})$. As $f(\sigma)$ is decreasing and $\sigma_0 \leq \sigma_1 \leq, \dots, \sigma_n$, we can compute $f(\sigma_i)$ for $i = 1, \dots, n$ and find out the region $f(\sigma_{i'}) \leq c < f(\sigma_{i'+1})$. This implies that $\sigma_{i'} \leq \sigma^\star < \sigma_{i'+1}$. Within this region $[\sigma_{i'}, \sigma_{i'+1})$, we have

$$f(\sigma) = \sum_{j=1}^n f_j(\sigma) = \sum_{j=1}^{i'} f_j(\sigma) + \sum_{j=i'+1}^n f_j(\sigma) = \sum_{j=1}^{i'} b_j a_j + \sum_{j=i'+1}^n (y_j b_j - \tfrac{\sigma}{2} b_j^2). \tag{13}$$

The solution $x^\star$ can be easily computed as $f(\sigma)$ is just linear now.