Recitation 3 Solutions

March 12, 2005

Optimal Routing in a Communication Network

Example 2.1.3, p.198-p.200, text book
Example 5.7, p.453-454, "Data Networks" by Dimitri Bertsekas and Robert Gallager

2.3.3

(a) Since the scaling matrix H^k is positive definite and diagonal, we have that each diagonal term H_i^k of the matrix H^k is positive. Define $y = (H^k)^{1/2}x$, and $h^k(y) = f((H^k)^{-1/2}y)$. Then $\nabla h^k(y) = (H^k)^{-1/2}\nabla f(x^k)$ and the set $X = \{x \mid \alpha_i \leq x_i \leq \beta_i,\ i = 1, \ldots, n\}$ is transformed to $Y^k = \{y \mid \hat{\alpha}_i^k \leq x_i \leq \hat{\beta}_i^k,\ i = 1, \ldots, n\}$ with $\hat{\alpha}_i^k = (H_i^k)^{1/2}\alpha_i,\ \ \hat{\beta}_i^k = (H_i^k)^{1/2}\beta_i$ for all i, k. Through this transformation the gradient projection method in y-space is given by

$$y^{k+1} = y^k + \gamma^k(\bar{y}^k - y^k), \tag{1}$$

$$\bar{y}^k = [y^k - s^k\nabla h^k(y^k)]^+, \tag{2}$$

where $[\cdot]^+$ denotes the projection on the set Y. Hence

$$\bar{y}_i^k = \begin{cases} \hat{\alpha}_i^k & \text{if } y_i^k - s^k\left(\nabla h^k(y^k)\right)_i \leq \hat{\alpha}_i^k, \\ \hat{\beta}_i^k & \text{if } y_i^k - s^k\left(\nabla h^k(y^k)\right)_i \geq \hat{\beta}_i^k, \\ y_i^k - s^k\left(\nabla h^k(y^k)\right)_i & \text{otherwise.} \end{cases}$$

Going back to the x-space we obtain the desired form for the coordinates of the vector $\bar{x}^k$, namely

$$x_i^k = \begin{cases} \alpha_i & \text{if } x_i^k - \frac{s^k}{H_i^k}\cdot\frac{\partial f(x^k)}{\partial x_i} \leq \alpha_i, \\ \beta_i & \text{if } x_i^k - \frac{s^k}{H_i^k}\cdot\frac{\partial f(x^k)}{\partial x_i} \geq \beta_i, \\ x_i^k - \frac{s^k}{H_i^k}\cdot\frac{\partial f(x^k)}{\partial x_i} & \text{otherwise.} \end{cases}$$

(b) Here we have

$$x^{k+1} = x^k + \gamma^k(\bar{x}^k - x^k),$$

$$\bar{x}^k = \arg\min_{x\in X}\left\{\nabla f(x^k)'(x - x^k) + \frac{1}{2s^k}(x - x^k)'H^k(x - x^k)\right\}, \tag{1}$$

where γ^k is the stepsize, H^k is diagonal with $H_i^k > 0$, and X is a simplex, i.e.

$$X = \{x \mid \sum_{i=1}^{n} x_i = r, \quad x \geq 0\}.$$

Since the objective function in (1) is convex, the first order necessary conditions are also sufficient. Thus, $\bar{x^k}$ is optimal in (1) if and only if

$$\bar{x_i^k} > 0 \implies \frac{\partial f(x^k)}{\partial x_j} + \frac{H_j^k}{s^k}(\bar{x}_j - x_j^k) \geq \frac{\partial f(x^k)}{\partial x_i} + \frac{H_i^k}{s^k}(\bar{x}_i - x_i^k), \quad \forall\, j.$$

To find $\bar{x^k}$ algorithmically, we try to find w such that

$$\frac{\partial f(x^k)}{\partial x_j} + \frac{H_j^k}{s^k}(x_j - x_j^k) \geq w, \quad \forall\, j, \tag{2}$$

with equality only if $x_i > 0$. So we start with some small value of w and increase it up to the point where

$$\sum_{j=1}^{n} \max\left\{0, \frac{s^k}{H_j^k}\left(w - \frac{\partial f(x^k)}{\partial x_j} + \frac{H_j^k}{s^k} x_j^k\right)\right\} = r. \tag{3}$$

If $\bar{w}$ is the value obtained (i.e. $\bar{w}$ satisfies both (2) and (3)), then the desired solution $\bar{x^k}$ is given by

$$\bar{x_i^k} = \max\left\{0, \frac{s^k}{H_i^k}\left(\bar{w} - \frac{\partial f(x^k)}{\partial x_i} + \frac{H_i^k}{s^k} x_i^k\right)\right\}, \quad \forall\, i.$$

3.1.6

By introducing the new variables

$$y_i = \ln(x_i), \qquad i = 1, .., n,$$

the original problem is transformed to an equivalent problem

$$\min \sum_{i=1}^{n} \alpha_i e^{y_i}$$

$$\text{subject to } \sum_{i=1}^{n} \alpha_i y_i = 0, \qquad y_i \in,$$

and the Lagrangian function of this problem is

$$L(y, \lambda) = \sum_{i=1}^{n} \alpha_i (e^{y_i} + \lambda y_i).$$

The 1st order necessary conditions are

$$\nabla_y L(y^*, \lambda^*) = 0 \quad \Leftrightarrow \quad e^{y_i^*} + \lambda^* = 0, \qquad i = 1, .., n$$

$$\nabla_\lambda L(y^*, \lambda^*) = 0 \quad \Leftrightarrow \quad \alpha_1 y_1^* + \cdots + \alpha_n y_n^* = 0.$$

The above system possesses a unique solution

$$y^* = 0, \qquad \lambda^* = -1.$$

Since the function $f(y) = \sum_{i=1}^n \alpha_i e^{y_i}$ is continuous and coercive, and the feasible set

$$\{y \in | \sum_{i=1}^n \alpha_i y_i = 0\}$$

is closed, by Weierstrass' theorem, the minimization problem has at least one global minimum y^*. Furthermore, every feasible point is regular [here $h(y) = \sum_{i=1}^n \alpha_i y_i$ and $\nabla h(y)' = (\alpha_1, \dots, \alpha_n)' \neq 0$], so that the 1st order necessary conditions must be satisfied at a global minimum. Because $y^* = 0$ and $\lambda^* = -1$ is the unique pair satisfying the 1st order necessary conditions, the point $y^* = 0$ is the unique global minimum, which in the x-space corresponds to $x^* = (1, \dots, 1)'$ with the minimum cost value $f(x^*) = \alpha_1 + \cdots + \alpha_n = 1$.

Now let $x_1, ..., x_n$ be positive numbers and let

$$x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n} = a > 0$$

for some positive number a. Since $1 + \cdots + n = 1$, we have

$$x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n} = a^{\alpha_1 + \cdots + \alpha_n}.$$

Hence

$$\left(\frac{x_1}{a}\right)^{\alpha_1} \cdots \left(\frac{x_n}{a}\right)^{\alpha_n} = 1$$

and from the above minimization problem we get

$$\sum_{i=1}^n \alpha_i \frac{x_i}{a} \geq 1$$

or

$$\sum_{i=1}^n \alpha_i x_i \geq a = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}.$$