An interior-point method for minimizing the sum of piecewise-linear convex functions

Toshihiro Kosaki
The Department of Industrial and Management
Tokyo Institute of Technology
2-12-1 Oh-Okayama, Meguro-ku, Tokyo 152-8552 Japan
kosaki.t.aa@m.titech.ac.jp

(February 26, 2010)

Abstract We consider the problem to minimize the sum of piecewise-linear convex functions under both linear and nonnegative constraints. We convert the piecewise-linear convex problem into a standard form linear programming problem (LP) and apply a primal-dual interior-point method for the LP. From the solution of the converted problem, we can obtain the solution of the original problem.

We establish polynomial convergence of the interior-point method for the converted problem and devise the computaion of the Newton direction.

Keywords: optimization, piecewise-linear convex function, interior-point method

1. Introduction

In this paper, we consider the problem to minimize the sum of piecewise-linear convex functions under both linear constraints and nonnegative constraints. This problem is a generalized one from a piecewise-linear convex problem[1, 3]. The purpose of the paper is to estimate complexity of solving this problem. More precisely speaking, is there a polynomial time method for the problem?

This paper is organized as follows. In section 2, we introduce the conversion from a linear programming problem(LP) with free variables into the standard form LP. The solution of the original problem is obtained from the solution of the converted problem. In section 3, we formulate the problem to minimize the sum of piesewise-linear convex functions. Firstly we convese the problem into an LP with free variable. And secondly we convert the LP into a standard form problem. We apply the interior-point method. We establish polynomial convergence of the algorithm and devise the computation of the Newton direction.

2. The conversion from the LP with free variables into the standard form LP

We consider the following LP with free variables.

$$\begin{aligned} &\min\ c^T x + f^T z \\ &\text{s.t. } Ax + Dz = b \\ &\quad x \geq 0, \end{aligned} \tag{P}$$

where $A \in \Re^{m\times n},\ b \in \Re^m,\ c \in \Re^n,\ f \in \Re^l,\ D \in \Re^{m\times l}$ are constants, $x \in \Re^n$ and $z \in \Re^l$ are variables. Without loss of generality, we can assume that the matrix D is full column. Therefore rank $D = l$. We introduce new variables z_+ and z_-. The standard form is as follow.

$$\begin{aligned} &\min\ c^T x + f^T z_+ - f^T z_- \\ &\text{s.t. } [A\ D\ -D] \begin{bmatrix} x \\ z_+ \\ z_- \end{bmatrix} = b \\ &\quad x \geq 0\ z_+ \geq 0\ z_- \geq 0. \end{aligned} \tag{P1}$$

The dual is as follow.

$$\begin{aligned} &\max\ b^T y \\ &\text{s.t. } A^T y \leq c \\ &\quad D^T y = f. \end{aligned} \tag{D1}$$

The transform $z = z_+ - z_-$ has a drawback that one z can't determine z_+ and z_- and the computation is unstable. The constraints of the dual problem, corresponding to free variables of the primal problem, are equality constraints. We solve the equations in basic variables and substitute it for both the objective function and constarint functions. We obtain a dual problem with only nonbasic variables. We consider the primal problem corresponding to this problem.In case of semidefinite programminng, the conversion was proposed[5]. The advantage is that the size of problem is small. We can obtain the solution of original problem from the solution of the converted problem. We use D_B for a submatrix of D corresponding to the basis, and use the submatrix D_N for the remain. We use subvectors y_B and y_N corresponding to D_B and D_N. And we use, respectively, the submatrices A_B and A_N and subvectors b_B and b_N corresponding to y_B and y_N. We solve equations $D^T y = f$ in basic variables,

$$y_B = D_B^{-T}(f - D_N^T y_N). \tag{1}$$

Substituting y_B into the objective function, we have

$$\begin{aligned} b^T y &= b_B^T D_B^{-T}(f - D_N^T y_N) + b_N^T y_N \\ &= (b_N^T - b_B^T D_B^{-T} D_N^T) y_N + b_B^T D_B^{-T} f. \end{aligned} \tag{2}$$

Substituting y_B into constraint functions, we have

$$A_B^T D_B^{-T}(f - D_N^T y_N) + A_N^T y_N \leq c. \tag{3}$$

The dual problem is as follows.

$$\begin{aligned} &\max\ (b_N^T - b_B^T D_B^{-T} D_N^T) y_N + b_B^T D_B^{-T} f \\ &\text{s.t.}\ \ (A_N^T - A_B^T D_B^{-T} D_N^T) y_N \leq c - A_B^T D_B^{-T} f. \end{aligned} \tag{D2}$$

The primal problem is as follows.

$$\begin{aligned} &\min\ (c - A_B^T D_B^{-T} f)^T x + b_B^T D_B^{-T} f \\ &\text{s.t.}\ \ (A_N - D_N D_B^{-1} A_B) x = b_N - D_N D_B^{-1} b_B \\ &\qquad x \geq 0. \end{aligned} \tag{P2}$$

We use x^* and y^* for the solution of the converted primal and dual problems. The solutions (x, z) and (y_B, y_N) of the original problems are given respectively by $(x^*, D_B^{-1}(b_B - A_B x^*))$ and $(D_B^{-T}(f - D_N^T y^*), y^*)$.

3. Formulation of the problem to minimize the sum of piecewise-linear convex functions

We consider the convex problem to minimize the sum of piecewise-linear convex functions under linear and nonnegative constraints. The sum of piecewise-linear convex function is also convex [2]. The objective function is nonlinear and has some nondifferential points. The formulation is as follows.

$$\begin{aligned} \min_x \sum_{p=1,\dots,q} &\max_{i_p=1,\dots,l_p} \left({c_{i_p}^p}^T x + d_{i_p}^p \right) \\ \text{s.t.}\ \ &Ax = b \\ &x \geq 0, \end{aligned} \tag{P3}$$

where $A \in \Re^{m \times n},\ b \in \Re^m,\ c_{i_p}^p \in \Re^n,\ d_{i_p}^p \in \Re$ are constants, $x \in \Re^n$ is variable. The objective function is the sum of q piecewise-linear convex functions and each piecewise-linear convex function consists of l_p affine funcitions. This problem is written, by introducing both variables t_p $(p = 1, \dots, q)$ and slack variables

$s^p_{i_p}$ $(i_p = 1, \dots, l_p, p = 1, \dots, q)$, as follows

$$
\begin{aligned}
\min\ & \sum_{p=1,\dots,q} t_p \\
\text{s.t. } & c^{pT}_1 x + d^p_1 + s^p_{1_p} = t_p \ p = 1, \dots, q \\
& \vdots \\
& c^p_{i_p} Tx + d^p_{i_p} + s^p_{i_p} = t_p \ p = 1, \dots, q \\
& \vdots \\
& c^p_{l_p}{}^T x + d^p_{l_p} + s^p_{l_p} = t_p \ p = 1, \dots, q \\
& Ax = b \\
& x \geq 0 \\
& s^p_{i_p} \ p = 1, \dots, q \ i_p = 1, \dots, l_p.
\end{aligned}
\tag{P4}
$$

We divide t_p into two nonnegative variables and transpose the terms. And we obtain a standard form LP,

$$
\begin{aligned}
\min\ & \sum_{p=1,\dots,q} t^+_p - \sum_{p=1,\dots,q} t^-_p \\
\text{s.t. } & c^p_{1_p}{}^T x - t^+_p + t^-_p + s^p_{1_p} = -d^p_1 \ p = 1, \dots, q \\
& \vdots \\
& c^p_{i_p}{}^T x - t^+_p + t^-_p + s^p_{i_p} = -d^p_{i_p} \ p = 1, \dots, q \\
& \vdots \\
& c^p_{l_p}{}^T x - t^+_p + t^-_p + s_{l_p{}^p} = -d^p_{l_p} \ p = 1, \dots, q \\
& Ax = b \\
& x \geq 0 \\
& t^+_p \geq 0 \ p = 1, \dots, q \ t^-_p \geq 0 \ p = 1, \dots, q \\
& s^p_{i_p} \geq 0 \ p = 1, \dots, q \ i = 1, \dots, l_p.
\end{aligned}
\tag{P5}
$$

The dual problem is as follows.

$$
\begin{aligned}
\max\ & b^T y - \sum_{p=1,\dots,q} \sum_{i_p=1,\dots,l_p} d^p_{i_p} u^p_{i_p} \\
\text{s.t. } & A^T y + \sum_{p=1,\dots,q} \sum_{i_p=1,\dots,l_p} c^p_{i_p} u^p_{i_p} \leq 0 \\
& \sum_{i_p=1,\dots,l_p} u^p_{i_p} = -1 \ p = 1, \dots, q \\
& u^p_{i_p} \leq 0 \ p = 1, \dots, q \ i_p = 1, \dots, l_p.
\end{aligned}
\tag{D3}
$$

The peculiarity is that the primal (P4) problem has free variables t_p. One of the methods dealing with free variables is that the free variable is represented by the difference of two nonnegative variables. This method has a drawback that variable diverge and numerical difficulties happen. We apply another method, which was introduced in section 2, in order to avoid this difficulty.

We solve one of the constraints, then $u^p_{l_p} = -1 - \sum_{i_p \neq l_p} u^p_{i_p} (\leq 0)$ $(p = 1, \dots, q)$ and we substitute this relation into both objective function and constraint function. The objective function is

$$
\begin{aligned}
& b^T y - \sum_{p=1,\dots,q} \sum_{i_p=1,\dots,l_p} d^p_{i_p} u^p_{i_p} \\
=& b^T y - \sum_{p=1,\dots,q} \sum_{i_p \neq l_p} d^p_{i_p} u^p_{i_p} - \sum_{p=1,\dots,q} \left\{ d^p_{l_p} \left(-1 - \sum_{i_p \neq l_p} u^p_{i_p} \right) \right\} \\
=& b^T y + \sum_{p=1,\dots,q} \sum_{i_p \neq l_p} (d^p_{l_p} - d^p_{i_p}) u^p_{i_p} + \sum_{p=1,\dots,q} d^p_{l_p}.
\end{aligned} \tag{4}
$$

The constraint function is

$$
\begin{aligned}
& A^T y + \sum_{p=1,\dots,q} \sum_{i_p=1,\dots,l_p} c^p_{i_p} u^p_{i_p} \\
=& A^T y + \sum_{p=1,\dots,q} \sum_{i_p \neq l_p} c^p_{i_p} u^p_{i_p} + \sum_{p=1,\dots,q} \left\{ c_{l_p} \left(-1 - \sum_{i_p \neq l_p} u^p_{i_p} \right) \right\} \\
=& A^T y + \sum_{p=1,\dots,q} \sum_{i_p \neq l_p} (c^p_{i_p} - c^p_{l_p}) u^p_{i_p} - \sum_{p=1,\dots,q} c^p_{l_p} (\leq 0).
\end{aligned} \tag{5}
$$

The dual problem is as follows.

$$
\begin{aligned}
\max\ & b^T y + \sum_{p=1,\dots,q} \sum_{i_p \neq l_p} (d^p_{l_p} - d^p_{i_p}) u^p_{i_p} + \sum_{p=1,\dots,q} d^p_{l_p} \\
\text{s.t.}\ & A^T y + \sum_{p=1,\dots,q} \sum_{i_p \neq l_p} (c^p_{i_p} - c^p_{l_p}) u^p_{i_p} \leq \sum_{p=1,\dots,q} c^p_{l_p} \\
& - \sum_{i_p \neq l_p} u^p_{i_p} \leq 1 \ p = 1, \dots, q \\
& u^p_{i_p} \leq 0 \ p = 1, \dots, q, \ i_p = 1, \dots, l_p - 1.
\end{aligned} \tag{D4}
$$

The primal problem is as follow.

$$
\begin{aligned}
\min\ & \left(\sum_{p=1,\dots,q} c^p_{l_p} \right)^T x + \sum_{p=1,\dots,q} s^p_{l_p} + \sum_{p=1,\dots,q} d^p_{l_p} \\
\text{s.t.}\ & (c^p_{1_p} - c^p_{l_p})^T x + s^p_1 - s^p_{l_p} = d^p_{l_p} - d^p_1 \ p = 1, \dots, q \\
& \vdots \\
& (c^p_{i_p} - c^p_{l_p})^T x + s^p_{i_p} - s^p_{l_p} = d^p_{l_p} - d^p_{i_p} \ p = 1, \dots, q \\
& \vdots \\
& (c^p_{l_p-1} - c^p_{l_p})^T x + s^p_{l_p-1} - s^p_{l_p} = d^p_{l_p} - d^p_{l_p-1} \ p = 1, \dots, q \\
& Ax = b \\
& x \geq 0, \\
& s^p_{i_p} \geq 0 \ p = 1, \dots, q \ i_p = 1, \dots, l_p.
\end{aligned} \tag{P6}
$$

We introduce slack variables.

$$
v_x := \sum_{p=1,\dots,q} c^p_{l_p} - A^T y - \sum_{p=1,\dots,q} \sum_{i_p \neq l_p} (c^p_{i_p} - c^p_{l_p}) u^p_{i_p}, \tag{6a}
$$

$$
v_{s^p_{l_p}} := 1 + \sum_{i_p \neq l_p} u^p_{i_p} \ p = 1, \dots, q, \tag{6b}
$$

$$
v_{s^p_{i_p}} := -u^p_{i_p} \ p = 1, \dots, q \ i_p = 1, \dots, l_p - 1. \tag{6c}
$$

The optimality condition is as follow.

$$
\begin{aligned}
&(c_1^p - c_{l_p}^p)^T x + s_1^p - s_{l_p}^p = d_{l_p}^p - d_1^p \ p = 1, \dots, q \\
&\qquad\qquad \vdots \\
&(c_{i_p}^p - c_{l_p}^p)^T x + s_{i_p}^p - s_{l_p}^p = d_{l_p}^p - d_{i_p}^p \ p = 1, \dots, q \\
&\qquad\qquad \vdots \\
&(c_{l_p-1}^p - c_{l_p}^p)^T x + s_{l_p-1^p} - s_{l_p}^p = d_{l_p}^p - d_{l_p-1}^p \ p = 1, \dots, q \\
&Ax = b \\
&x \geq 0 \\
&s_{i_p}^p \geq 0 \ p = 1, \dots, q \ i = 1, \dots, l_p \\
&A^T y + \sum_{p=1,\dots,q} \sum_{i_p \neq l_p} (c_{i_p}^p - c_{l_p}^p) u_{i_p}^p + v_x = \sum_{p=1,\dots,q} c_{l_p}^p \\
&-\sum_{i_p \neq l_p} u_{i_p}^p + v_{s_{l_p}^p} = 1 \ p = 1, \dots, q \\
&u_{i_p}^p + v_{s_{l_p}^p} = 0 \ p = 1, \dots, q, \ i_p = 1, \dots, l_p - 1 \\
&v_x \geq 0 \\
&v_{s_{l_p}^p} \geq 0 \ p = 1, \dots, q \\
&v_{s_{i_p}^p} \geq 0 \ p = 1, \dots, q, \ i_p = 1, \dots, l_p - 1 \\
&x^T v_x = 0 \\
&s_{i_p}^p v_{s_{i_p}^p} = 0 \ p = 1, \dots, q, \ i_p = 1, \dots, l_p.
\end{aligned}
\tag{7}
$$

We use $(x^*, s_1^{p*}, \dots, s_{l_p}^{p\,*})$ for the solution of the converted primal problem (P6). The solution of the original primal problem (P4) is given by

$$
(x, t_p, s_1^p, \dots, s_{l_p}^p) = (x^*, c_{l_p{}^p}^T x^* + s_{l_p}^{p\,*} + d_{l_p}^p, s_{1_p}^{p\,*}, \dots, s_{l_p{}^p}^*). \tag{8}
$$

We use $(y^*, u_1^{p*}, \dots, u_{l_p-1}^{p\ \ *})$ for the solution of the converted dual problem(D4). The solution of the original dual problem (D3) is given by

$$
(y, u_1^p, \dots, u_{l_p-1}^p, u_{l_p}^p) = (y^*, u_1^{p*}, \dots, u_{l_p-1}^{p\ \ *}, -1 - \sum_{i_p \neq l_p} u_{i_p}^{p\,*}). \tag{9}
$$

By the number of the variables in the primal problem (P6), the short step interior-point method (Algorithm SPF[7]) can solve the problem (P6) in $O\left(\sqrt{n + \sum_{p=1}^{q} l_p L}\right)$ iterations. We ascertain this proposition.

We use the following notations.

$$\tilde{x} := \begin{bmatrix} x \\ s_1^1 \\ \vdots \\ s_{i_1}^1 \\ \vdots \\ s_{l_1}^1 \\ \vdots \\ s_1^q \\ \vdots \\ s_{i_q}^q \\ \vdots \\ s_{l_q}^q \end{bmatrix}, \ \tilde{y} := \begin{bmatrix} y \\ u_1^1 \\ \vdots \\ u_{i_1}^1 \\ \vdots \\ u_{l_1-1}^1 \\ \vdots \\ u_1^q \\ \vdots \\ u_{i_q}^q \\ \vdots \\ u_{l_q-1}^q \end{bmatrix}, \ \tilde{s} := \begin{bmatrix} v_x \\ \vdots \\ v_{s_1^1} \\ \vdots \\ v_{s_{i_1}^1} \\ \vdots \\ v_{s_{l_1}^1} \\ \vdots \\ v_{s_1^q} \\ \vdots \\ v_{s_{i_q}^q} \\ \vdots \\ v_{s_{l_q}^q} \end{bmatrix}, \tag{10}$$

$$\tilde{A} := \begin{bmatrix} A & 0 & \dots & 0 & \dots & 0 & 0 & & & & & & & \\ (c_1^1 - c_{l_1}^1)^T & 1 & \dots & 0 & \dots & 0 & -1 & & & & & & & \\ \vdots & & \ddots & & & & \vdots & & & & & & & \\ (c_{i_1}^1 - c_{l_1}^1)^T & 0 & \dots & 1 & \dots & 0 & -1 & & & & & & & \\ \vdots & & & & \ddots & & \vdots & & & & & & & \\ (c_{l_1-1}^1 - c_{l_1}^1)^T & 0 & \dots & 0 & \dots & 1 & -1 & & & & & & & \\ \vdots & & & & & & & \ddots & & & & & & \\ (c_1^q - c_{l_q}^q)^T & & & & & & & & 1 & \dots & 0 & \dots & 0 & -1 \\ \vdots & & & & & & & & & \ddots & & & & \vdots \\ (c_{i_q}^q - c_{l_q}^q)^T & & & & & & & & 0 & \dots & 1 & \dots & 0 & -1 \\ \vdots & & & & & & & & & & & \ddots & & \vdots \\ (c_{l_q-1}^q - c_{l_q}^q)^T & & & & & & & & 0 & \dots & 0 & \dots & 1 & -1 \end{bmatrix}, \tag{11}$$

$$\tilde{b} := \begin{bmatrix} b \\ d_{l_1}^1 - d_1^1 \\ \vdots \\ d_{l_1}^1 - d_{i_1}^1 \\ \vdots \\ d_{l_1}^1 - d_{l_1-1}^1 \\ \vdots \\ d_{l_q}^q - d_1^q \\ \vdots \\ d_{l_q}^q - d_{i_q}^q \\ \vdots \\ d_{l_q}^q - d_{l_q-1}^q \end{bmatrix}, \ \tilde{c} := \begin{bmatrix} \sum_{p=1}^q c_{l_p}^p \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 1 \end{bmatrix}. \tag{12}$$

The feasible interior is written by

$$\mathcal{F}^0 := \left\{ (\tilde{x}, \tilde{y}, \tilde{s}) : \tilde{A}\tilde{x} = \tilde{b},\ \tilde{A}^T\tilde{y} + \tilde{s} = \tilde{c},\ (\tilde{x}, \tilde{s}) > 0 \right\}. \tag{13}$$

The neighborhood is written by

$$\mathcal{N} := \left\{(\tilde{x},\tilde{y},\tilde{s}) \in \mathcal{F}^0 : \|\tilde{X}\tilde{s} - \mu e\|_2 \le 0.4\mu\right\}, \tag{14}$$

where e is a vector whose components are all 1. We use the following algorithm.

step0: Initial point $(\tilde{x}^0,\tilde{y}^0,\tilde{s}^0) \in \mathcal{N}$ is given. Termination criteria μ^*.

$\sigma := 1 - \dfrac{0.4}{\sqrt{n+\sum_{p=1}^q l_p}}$, $\mu^0 := \dfrac{\tilde{x}^{0T}\tilde{s}^0}{n+\sum_{p=1}^q l_p}$, let $k := 0$.

step1: If termination criteria $\mu^k \le \mu^*$ is satisfied, then stop.

step2: Set $\mu^k := \dfrac{\tilde{x}^{k^T}\tilde{s}^k}{n+\sum_{p=1}^q l_p}$. Solve

$$\begin{aligned} \tilde{A}\Delta\tilde{x}^k &= 0 &(15a)\\ \tilde{A}^T\Delta\tilde{y}^k + \Delta\tilde{s}^k &= 0 &(15b)\\ \tilde{S}^k\Delta\tilde{x}^k + \tilde{X}^k\Delta\tilde{s}^k &= \sigma\mu^k e - \tilde{X}^k\tilde{s}^k, &(15c)\end{aligned}$$

and obtain Newton direction $(\Delta\tilde{x}^k,\Delta\tilde{y}^k,\Delta\tilde{s}^k)$.

Let $\left(\tilde{x}^{k+1},\tilde{y}^{k+1},\tilde{s}^{k+1}\right) := \left(\tilde{x}^k,\tilde{y}^k,\tilde{s}^k\right) + \left(\Delta\tilde{x}^k,\Delta\tilde{y}^k,\Delta\tilde{s}^k\right)$.

step3: Set $k := k+1$, go to step1.

3.1. iteration number

We consider at kth iteration. The equality constraints are satisfied by the following relation

$$\tilde{A}\tilde{x}^{k+1} = \tilde{A}\tilde{x}^k = \tilde{b}, \tag{16}$$

$$\tilde{A}^T\tilde{y}^{k+1} + \tilde{s}^{k+1} = \tilde{A}^T\tilde{y}^k + \tilde{s}^k = \tilde{c}. \tag{17}$$

Note that, about dualtiy gap used by optimality criteria, $\Delta\tilde{x}^{k^T}\Delta\tilde{s}^k = 0$, We have the following estimate

$$\begin{aligned}\tilde{x}^{k+1^T}\tilde{s}^{k+1} &= \left(\tilde{x}^k + \Delta\tilde{x}^k\right)^T\left(\tilde{s}^k + \Delta\tilde{s}^k\right)\\ &= \tilde{x}^{k^T}\tilde{s}^k + \tilde{x}^{k^T}\Delta\tilde{s}^k + \tilde{s}^{k^T}\Delta\tilde{x}^k + \Delta\tilde{x}^{k^T}\Delta\tilde{s}^k\\ &= \left(1 - \frac{0.4}{\sqrt{n+\sum_{p=1}^q l_p}}\right)(n + \sum_{p=1}^q l_p)\mu^k.\end{aligned} \tag{18}$$

By the following inequality, the next point generated by the algorithm is also kept in the neighborhood.

$$\begin{aligned}\left\|\tilde{X}^{k+1}\tilde{s}^{k+1} - \mu^{k+1}\right\|_2 &= \left\|\left(\tilde{X}^k + \Delta\tilde{X}^k\right)\left(\tilde{s}^k + \Delta\tilde{s}^k\right) - \mu^{k+1}\right\|_2 &(19)\\ &= \left\|\Delta\tilde{X}^k\Delta\tilde{s}^k\right\|_2 &(20)\\ &= \left\|D^{-1}\Delta\tilde{X}^k D\Delta\tilde{s}^k\right\|_2 \text{ where } D := \tilde{X}^{k^{1/2}}\tilde{S}^{k^{-1/2}} &(21)\\ &\le \frac{\sqrt{2}}{4}\left\|D^{-1}\Delta\tilde{x}^k + D\Delta\tilde{s}^k\right\|_2^2 &(22)\\ &= \frac{\sqrt{2}}{4}\left\|\left(\tilde{X}^k\tilde{S}^k\right)^{-1/2}\left(\sigma\mu^k e - \tilde{X}^k\tilde{s}^k\right)\right\|_2^2 &(23)\\ &\le \frac{\sqrt{2}}{4}\frac{\left\|\tilde{X}^k\tilde{s}^k - \sigma\mu^k e\right\|_2^2}{\min\ \tilde{x}_i^k\tilde{s}_i^k} &(24)\\ &\le \frac{\sqrt{2}}{4}\frac{\left\|\left(\tilde{X}^k\tilde{s}^k - \mu^k e\right) + (1-\sigma)\mu^k e\right\|_2^2}{(1-0.4)\mu^k} &(25)\\ &\le \frac{\sqrt{2}}{4}\frac{0.4^2 + (1-\sigma)^2(n+\sum_{p=1}^q l_p)}{1-0.4} &(26)\\ &\le \frac{32\sqrt{2}}{240}\mu^k &(27)\\ &\le 0.4\mu^{k+1}, &(28)\end{aligned}$$

where $\tilde{X}$ and $\tilde{S}$ are diagonal matrices whose components are $\tilde{x}$ and $\tilde{s}$ respectively. Positivity condition is also satisfied.

Because one iteration can make the duality gap $1-0.4/\sqrt{n+\sum_{p=1}^{q} l_p}$ time, algorithm can obtain optimal solution in $O(\sqrt{n+\sum_{p=1}^{q} l_p}L)$ iterarions, where L is the number of bits expressing the data of the problem. From the solution of the converted problem, we can obtain the solution of the original roblem.

3.2. Newton direction

The time of computing the Newton direction is most part of all computing time. Therefore devising computing the Newton direction is important. For the simplicity, we sonsider the problem about the sum of two piecewise-linear convex functions. The Newton direction is given by the solution of the following system of equations

$$\begin{aligned}
\tilde{A}\tilde{X}^k\tilde{S}^{k^{-1}}\tilde{A}^T\Delta\tilde{y}^k &= -\tilde{A}\tilde{S}^{k^{-1}}\left(\sigma\mu^k e-\tilde{X}^k\tilde{s}^k\right) && \text{(29a)}\\
\Delta\tilde{s}^k &= -\tilde{A}^T\Delta\tilde{y}^k && \text{(29b)}\\
\Delta\tilde{x}^k &= \tilde{S}^{k^{-1}}\left(\sigma\mu^k e-\tilde{X}^k\tilde{s}^k\right)-\tilde{X}^k\tilde{S}^{k^{-1}}\Delta\tilde{s}^k && \text{(29c)}
\end{aligned}$$

We use the following notations.

$$\bar{A} := \begin{bmatrix} A \\ (c^1_1-c^1_{l_1})^T \\ \vdots \\ (c^1_{i_1}-c^1_{l_1})^T \\ \vdots \\ (c^1_{l_1-1}-c^1_{l_1})^T \\ (c^2_1-c^2_{l_2})^T \\ \vdots \\ (c^2_{i_2}-c^2_{l_2})^T \\ \vdots \\ (c^2_{l_2-1}-c^2_{l_2}) \end{bmatrix} \mathrm{diag}(x)\mathrm{diag}(v_x)^{-1} \begin{bmatrix} A \\ (c^1_1-c^1_{l_1})^T \\ \vdots \\ (c^1_{i_1}-c^1_{l_1})^T \\ \vdots \\ (c^1_{l_1-1}-c^1_{l_1})^T \\ (c^2_1-c^2_{l_2})^T \\ \vdots \\ (c^2_{i_2}-c^2_{l_2})^T \\ \vdots \\ (c^2_{l_2-1}-c^2_{l_2}) \end{bmatrix}^T$$

$$+\begin{bmatrix} O &&&&&&&& \\ & s^1_1 v^{-1}_{s^1_1} &&&&&&& \\ && \ddots &&&&&& \\ &&& s_{i^1_1} v^{-1}_{s^1_{i_1}} &&&&& \\ &&&& \ddots &&&& \\ &&&&& s_{l_1-1^1} v^{-1}_{s_{l_1-1^1}} &&& \\ &&&&&& s^2_1 v^{-1}_{s^2_1} && \\ &&&&&&& \ddots & \\ &&&&&&&& s^2_{i_2} v^{-1}_{s^2_{i_2}} \\ &&&&&&&&& \ddots \\ &&&&&&&&&& s^2_{l_2-1} v^{-1}_{s^2_{l_2-1}} \end{bmatrix}, \tag{30}$$

where diag() is a diagonal matrix whose components are elements of the input vector.The coefficient matrices are written by

$$\bar{A}+s^1_{l_1}v^{-1}_{s^1_{l_1}}\begin{bmatrix}0\\e\\0\end{bmatrix}\begin{bmatrix}0\\e\\0\end{bmatrix}^T+s^2_{l_2}v^{-1}_{s^2_{l_2}}\begin{bmatrix}0\\0\\e\end{bmatrix}\begin{bmatrix}0\\0\\e\end{bmatrix}^T. \tag{31}$$

Let A be a nonsingular matrix and B, C, D be matrices of proper size. Note that we have Sherman-Morrison-Woodbury formula [4, 6, 7]

$$(A+BDC)^{-1} = A^{-1}-A^{-1}BD\left(D+DCA^{-1}BD\right)^{-1}DCA^{-1}. \tag{32}$$

we use the notation

$$\hat{A} := \bar{A} + s^1_{l_1} v^{-1}_{s^1_{l_1}} \begin{bmatrix} 0 \\ e \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ e \\ 0 \end{bmatrix}^T . \tag{33}$$

By using the relation

$$\hat{A}^{-1} = \bar{A}^{-1} - \frac{s_{l^1_1} v^{-1}_{s^1_{l_1}}}{1 + s^1_{l_1} v^{-1}_{s^1_{l_1}} \begin{bmatrix} 0 \\ e \\ 0 \end{bmatrix}^T \bar{A}^{-1} \begin{bmatrix} 0 \\ e \\ 0 \end{bmatrix}} \bar{A}^{-1} \begin{bmatrix} 0 \\ e \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ e \\ 0 \end{bmatrix}^T \bar{A}^{-1}, \tag{34}$$

we can obtain $\hat{A}^{-1}$ from $\bar{A}^{-1}$. And by using the relation

$$\left(\tilde{A}\tilde{X}^k \tilde{S}^{k^{-1}} \tilde{A}^T\right)^{-1} = \hat{A}^{-1} - \frac{s^2_{l_2} v^{-1}_{s^2_{l_2}}}{1 + s^2_{l_2} v^{-1}_{s^2_{l_2}} \begin{bmatrix} 0 \\ 0 \\ e \end{bmatrix}^T \hat{A}^{-1} \begin{bmatrix} 0 \\ 0 \\ e \end{bmatrix}} \hat{A}^{-1} \begin{bmatrix} 0 \\ 0 \\ e \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ e \end{bmatrix}^T \hat{A}^{-1}, \tag{35}$$

we can obtain the Newton direction. If $\bar{A}$ has special structure by which it is easy to compute its inverse matrix, then this devise reduces computation time. This method can be applied to the problem to minimize the sum of more than three piecewise-linear convex functions.

References

[1] D. Bertsimas and J. N. Tsitsiklis: *Introduction to Linear Optimization*, Athena Scientific (1997).

[2] S. Boyd and L. Vandenberghe: *Convex Optimization*, Cambridge University Press (2004).

[3] G. B. Dantzig and M. N. Thapa: *Linear Programming 2:Theory and Extensions*, Springer (2003).

[4] M. Iri: *Ippansenkeidaisu* (in Japanese), Iwanami shoten (2003).

[5] K. Kobayashi, K. Nakata and M. Kojima: A Conversion of an SDP Having Free Variables into the Standard Form SDP, Computational Optimization and Applications, 36 (2007), 289-307.

[6] J. Nocedal and S. J. Wright: *Numerical Optimization*, Springer (2006).

[7] S. J. Wright: *Primal-Dual Interior-Point Method*, SIAM Publications (1997).