区分線形凸関数の和の最小化問題に対する内点法

小崎敏寛
東京工業大学

(2009 年 6 月 3 日)

和文概要 区分線形凸関数の和の最小化問題を考える。この問題に対する多項式オーダーの解法を提案する。その手法では、区分線形凸関数の和の最小化問題を標準形の線形計画問題に変換し、主双対内点法を適用する。アルゴリズムの多項式オーダー性を示し、ニュートン方向の計算の工夫を記述する。

キーワード: 最適化, 区分線形凸関数, 内点法

1. はじめに

本稿では、区分線形凸関数の和を等式制約と非負制約の下で最小化する問題を考える。この問題は区分線形計画問題 [1, 3] を一般化した問題である。課題としては、この問題の計算量はどのくらいであるかである。より正確に言うと、多項式オーダーの解法はあるのかということである。

本稿の構成は、以下の通りである。2 節では、自由変数を持つ線形計画問題の標準形への変換を示し、変換後の問題の解から変換前の解が得られることを示す。3 節では、区分線形凸関数の和の最小化問題の定式化を行う。2 節で導入した変換を適用し、標準形に変換する。この問題に対して内点法を適用する。アルゴリズムの多項式オーダー性を示し、ニュートン方向の計算の工夫を記す。

2. 自由変数を持つ線形計画問題の標準形への変換

次の自由変数を持つ線形計画問題を考える。

$$\begin{align*}
\text{min } & c^T x + f^T z \\
\text{s.t. } & Ax + Dz = b \\
& x \geq 0.
\end{align*}$$

(P)

ただし、定数 $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$, $f \in \mathbb{R}^l$, $D \in \mathbb{R}^{m \times l}$, 変数 $x \in \mathbb{R}^n$, $z \in \mathbb{R}^l$. 一般性を失うことなく、行列 $D$ がフル列ランクであると仮定できる。すると $\text{rank} D = l$ である。

標準形は次のようになる。

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

(P1)
双対問題は次のようになる。
\[
\begin{align*}
    \max & \quad b^T y \\
    \text{s.t.} & \quad A^T y \leq c \\
    & \quad D^T y = f.
\end{align*}
\]
(D1)

内点法を適用する前の変換 \( z = z_+ - z_- \) には問題点がある。それは \( z \) を一つ定めるのに \( z_+ , z_- \) は一意に定まらず，計算が不安定になることである。

主問題の自由変数に対応する双対問題の制約は等式制約になる。その等式を基底変数について解いて，目的関数，制約式に代入し非基底変数のみの双対問題を考える。この問題に対する主問題を考える変換法が SDP に対して提案されている [5]．その利点は，問題のサイズが小さくなることである。変換した問題の解から元の問題の解を得ることができる。

行列 \( D \) の基底部分を \( D_B \) 残りの部分を \( D_N \) とする。\( D_T y = f \) を基底変数について解くと，
\[
y_B = D_B^{-T} (f - D_N^T y_N).
\]
(2.1)

目的関数に代入すると，
\[
b^T y = b_B^T D_B^{-T} (f - D_N^T y_N) + b_N^T y_N
\]
\[
= (b_N^T - b_B^T D_B^{-T} D_N^T) y_N + b_B^T D_B^{-T} f.
\]
(2.2)

制約式に代入すると，
\[
A_B^T D_B^{-T} (f - D_N^T y_N) + A_N^T y_N \leq c.
\]
(2.3)

双対問題は次のようになる。
\[
\begin{align*}
    \max & \quad (b_N^T - b_B^T D_B^{-T} D_N^T) y_N + b_B^T D_B^{-T} f \\
    \text{s.t.} & \quad (A_N^T - A_B^T D_B^{-T} D_N^T) y_N \leq c - A_B^T D_B^{-T} f.
\end{align*}
\]
(D2)

主問題は次のようになる。
\[
\begin{align*}
    \min & \quad (c - A_B^T D_B^{-T} f)^T x + b_B^T D_B^{-T} f \\
    \text{s.t.} & \quad (A_N - D_N D_B^{-1} A_B) x = b_N - D_N D_B^{-1} b_B
\end{align*}
\]
(P2)

変換後の主問題の解を \( x^* \) 同時変換後の双対問題の解を \( y^* \) とすると，元の問題の解 \( (x, z) \) はそれぞれ \( (x^*, D_B^{-1} (b_B - A_B x^*)) \)，\( (D_B^{-T} (f - D_N^T y^*), y^*) \) で与えられる。

3. 目的関数が区分線形凸関数の和の最小化問題の定式化

区分線形凸関数の和も凸関数である [2]。目的関数が複数の区分線形凸関数の和である凸計画問題を考える。
\[
\begin{align*}
    \min_x & \quad \sum_{p=1,...,q} \max_{i_p=1,...,l_p} \left(p^T x + d_i^p\right) \\
    \text{s.t.} & \quad Ax = b \\
    & \quad x \geq 0.
\end{align*}
\]
(P)
ただし，定数
\[ A \in \mathbb{R}^{m \times n}, \quad b \in \mathbb{R}^m, \quad c_i^p \in \mathbb{R}^n, \quad d_i^p \in \mathbb{R}, \] 変数は \( x \in \mathbb{R}^n \). この問題は変数 \( t_p (p = 1, \ldots, q) \) を使うと次のように書ける。

\[
\min \sum_{p=1}^{q} t_p \\
\text{s.t. } c_{1}^{pT} x + d_{1}^{p} \leq t_p p = 1, \ldots, q \\
\vdots \\
c_{i}^{pT} x + d_{i}^{p} \leq t_p p = 1, \ldots, q \\
\vdots \\
c_{l_{p}}^{pT} x + d_{l_{p}}^{p} \leq t_p p = 1, \ldots, q \\
Ax = b \\
x \geq 0.
\]

スラック変数を導入し，標準形になおすと次のようになる。

\[
\min \sum_{p=1}^{q} t_p^+ - \sum_{p=1}^{q} t_p^- \\
\text{s.t. } c_{1}^{pT} x - t_p^+ + t_p^- + s_i^p = -d_{1}^{p} p = 1, \ldots, q \\
\vdots \\
c_{i}^{pT} x - t_p^+ + t_p^- + s_i^p = -d_{i}^{p} p = 1, \ldots, q \\
\vdots \\
c_{l_{p}}^{pT} x - t_p^+ + t_p^- + s_i^p = -d_{l_{p}}^{p} p = 1, \ldots, q \\
Ax = b \\
x \geq 0, \\
t_p^+ \geq 0 p = 1, \ldots, q, \quad t_p^- \geq 0 p = 1, \ldots, q, \\
s_i^p \geq 0 p = 1, \ldots, q i = 1, \ldots, l_p.
\]

双対問題は次のようになる。

\[
\max b^T y - \sum_{p=1}^{q} \sum_{i=1}^{l_p} d_i^p u_i^p \\
\text{s.t. } A^T y + \sum_{p=1}^{q} \sum_{i=1}^{l_p} c_i^p u_i^p \leq 0 \\
\sum_{i=1}^{l_p} u_i^p = -1 p = 1, \ldots, q \\
u_i^p \leq 0 p = 1, \ldots, q i = 1, \ldots, l_p.
\]

主問題は自由変数を持つ．単純な変換法として，自由変数を非負変数の差として表す手法がある．この手法は，変数が発散するため数値的不安定になる欠点を持つ．この欠点を回避する手法を提案する．
制約式を解いて，\( u_p^i = -1 - \sum_{i \neq l_p} u_p^i (\leq 0) \) \( p = 1, \ldots, q \) として，目的関数，制約関数に代入する．目的関数は，

\[
v^T y - \sum_{p=1, \ldots, q} \sum_{i \neq l_p} d_p^i u_p^i
\]

\[
= v^T y - \sum_{p=1, \ldots, q} \sum_{i \neq l_p} d_p^i u_p^i - \sum_{p=1, \ldots, q} \left\{ d_p^l \left( -1 - \sum_{i \neq l_p} u_p^i \right) \right\}
\]

\[
= v^T y + \sum_{p=1, \ldots, q} \sum_{i \neq l_p} (d_p^i - d_p^l) u_p^i + \sum_{p=1, \ldots, q} d_p^l,
\]

(3.1)

制約式は，

\[
A^T y + \sum_{p=1, \ldots, q} \sum_{i \neq l_p} c_p^i u_p^i
\]

\[
= A^T y + \sum_{p=1, \ldots, q} \sum_{i \neq l_p} c_p^i u_p^i - \sum_{p=1, \ldots, q} \left\{ c_p^l \left( -1 - \sum_{i \neq l_p} u_p^i \right) \right\}
\]

\[
= A^T y + \sum_{p=1, \ldots, q} \sum_{i \neq l_p} (c_p^i - c_p^l) u_p^i - \sum_{p=1, \ldots, q} c_p^l (\leq 0)
\]

(3.2)

なる．双対問題は次のようになる．

\[
\begin{align*}
\text{max} & \quad b^T y + \sum_{p=1, \ldots, q} \sum_{i \neq l_p} (d_p^i - d_p^l) u_p^i + \sum_{p=1, \ldots, q} d_p^l \\
\text{s.t.} & \quad A^T y + \sum_{p=1, \ldots, q} \sum_{i \neq l_p} (c_p^i - c_p^l) u_p^i \leq \sum_{p=1, \ldots, q} c_p^l \\
& \quad - \sum_{i \neq l_p} u_p^i \leq 1 \quad p = 1, \ldots, q \\
& \quad u_p^i \leq 0 \quad p = 1, \ldots, q, \ i = 1, \ldots, l_p - 1.
\end{align*}
\]

(D)

この問題の主問題は次のようになる．

\[
\begin{align*}
\text{min} & \quad \left( \sum_{p=1, \ldots, q} c_p^l \right)^T x + \sum_{p=1, \ldots, q} s_p^l + \sum_{p=1, \ldots, q} d_p^l \\
\text{s.t.} & \quad (c_1^l - c_l^p)^T x + s_1^p - s_l^p = d_p^l - d_l^p \quad p = 1, \ldots, q \\
& \quad \vdots \\
& \quad (c_q^l - c_l^p)^T x + s_q^p - s_l^p = d_p^l - d_l^p \quad p = 1, \ldots, q \\
& \quad A x = b \\
& \quad x \geq 0, \\
& \quad s_p^i \geq 0 \quad p = 1, \ldots, q, \ i = 1, \ldots, l_p.
\end{align*}
\]

(P)
スラック変数を導入する。

\[
    v_x := \sum_{p=1,\ldots,q} c_{lp}^p - A^T y - \sum_{p=1,\ldots,q} \sum_{i \neq lp} (c_i^p - c_{lp}^p) u_i^p, \quad (3.3)
\]

\[
    v_{sp}^p := 1 + \sum_{i \neq lp} u_i^p, \quad p = 1, \ldots, q \quad (3.4)
\]

\[
    v_i^p := -u_i^p, \quad p = 1, \ldots, q, \quad i = 1, \ldots, l_p - 1. \quad (3.5)
\]

最適条件は次のようになる。

\[
    (c_1 - c_{l_p}^p)^T x + s_1^p - s_{l_p}^p = d_{l_p}^p - d_i^p \quad p = 1, \ldots, q
\]

\[
    \vdots
\]

\[
    (c_{l_p-1} - c_{l_p}^p)^T x + s_{l_p-1}^p - s_{l_p}^p = d_{l_p}^p - d_i^p \quad p = 1, \ldots, q
\]

\[
    Ax = b,
\]

\[
    x \geq 0,
\]

\[
    s_i^p \geq 0 \quad p = 1, \ldots, q, \quad i = 1, \ldots, l_p,
\]

\[
    A^T y + \sum_{p=1,\ldots,q} \sum_{i \neq lp} (c_i^p - c_{lp}^p) u_i^p + v_x = \sum_{p=1,\ldots,q} c_{lp}^p,
\]

\[
    - \sum_{i \neq lp} u_i^p + v_{sp}^p = 1 \quad p = 1, \ldots, q,
\]

\[
    u_i^p + v_{sp}^p = 0 \quad p = 1, \ldots, q, \quad i = 1, \ldots, l_p - 1,
\]

\[
    v_x \geq 0,
\]

\[
    v_{sp}^p \geq 0 \quad p = 1, \ldots, q,
\]

\[
    v_{sp}^p \geq 0 \quad p = 1, \ldots, q, \quad i = 1, \ldots, l_p - 1,
\]

\[
    x^T v_x = 0
\]

\[
    s_i^p v_i^p = 0 \quad p = 1, \ldots, q, \quad i = 1, \ldots, l_p.
\]

変換後の主問題の最適解を \((x^*, s_1^{*p}, \ldots, s_{l_p}^{*p})\) とすると、元の主問題の最適解は、

\[
    (x, t_p, s_1^p, \ldots, s_{l_p}^p) = (x^*, c_{lp}^p x^* + s_{lp}^p + d_{l_p}, s_1^{*p}, \ldots, s_{l_p}^{*p}) \quad (3.7)
\]

と得られる。変換後の双対問題の最適解を \((y^*, u_1^{*p}, \ldots, u_{l_p-1}^{*p})\) とすると、元の双対問題の最適解は、

\[
    (y, u_1^p, \ldots, u_{l_p-1}^p, u_{l_p}^p) = (y^*, u_1^{*p}, \ldots, u_{l_p-1}^{*p}, -1 - \sum_{i \neq lp} u_i^{*p}) \quad (3.8)
\]

と得られる。

問題のサイズからショートステップ法（例えば Algorithm SPF[6]）で \(O \left( \sqrt{n + \sum_{p=1}^{q} L_p L} \right) \)
反復である。以下でこのことを確かめる。
次の表記を使う。

\[
\begin{align*}
\tilde{x} &:= \begin{bmatrix} x \\ s_1^q \\ \vdots \\ s_i^q \\ \vdots \\ s_{i-1}^q \end{bmatrix}, \\
\tilde{y} &:= \begin{bmatrix} y \\ u_1^q \\ \vdots \\ u_i^q \\ \vdots \\ u_{i-1}^q \end{bmatrix}, \\
\tilde{s} &:= \begin{bmatrix} v_x \\ v_{s_1^q} \\ \vdots \\ v_{s_i^q} \\ \vdots \\ v_{s_{i-1}^q} \end{bmatrix}, \\
\tilde{A} &:= \begin{bmatrix} A & 0 & \ldots & 0 & \ldots & 0 & 0 \\
(c_1^q - c_1^q) & 1 & \ldots & 0 & \ldots & 0 & -1 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\
(c_i^q - c_i^q) & 0 & \ldots & 1 & \ldots & 0 & -1 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\
(c_{i-1}^q - c_{i-1}^q) & 0 & \ldots & 0 & \ldots & 1 & -1 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\
(c_q^q - c_q^q) & 1 & \ldots & 0 & \ldots & 0 & -1 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\
(c_{i-1}^q - c_{i-1}^q) & 0 & \ldots & 1 & \ldots & 0 & -1 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\
(c_{q-1}^q - c_{q-1}^q) & 0 & \ldots & 0 & \ldots & 1 & -1 \\
\end{bmatrix},
\end{align*}
\]

(3.9)
実行可能内点を
\[ F_0 := \{ (\tilde{x}, \tilde{y}, \tilde{s}) : \tilde{b} \leq \tilde{A} \tilde{x} = \tilde{b}, \tilde{A}^T \tilde{y} + \tilde{s} = \tilde{c}, (\tilde{x}, \tilde{s}) > 0 \} \] と表す. 近傍を
\[ N := \{ (\tilde{x}, \tilde{y}, \tilde{s}) \in F_0 : \|\tilde{X}\tilde{s} - \mu e\|_2 \leq 0.4 \mu \} \] と表す. 次のアルゴリズムを用いる.

step0: 初期点 $(\tilde{x}^0, \tilde{y}^0, \tilde{s}^0) \in N$ が与えられる. 終了基準 $\mu^*$ を与える.

\[ \sigma := 1 - \frac{0.4}{\sqrt{n + \sum_{p=1}^q l_p}}, \mu^0 := \frac{\tilde{x}^{0T} \tilde{s}^0}{n + \sum_{p=1}^q l_p}, k := 0 \] とする.

step1: もし終了基準 $\mu^k \leq \mu^*$ をみたすならば終了する.

step2: $\mu^k := \frac{\tilde{x}^{kT} \tilde{s}^k}{n + \sum_{p=1}^q l_p}$ として

\[ \tilde{A} \Delta \tilde{x}^k = 0 \] と
\[ \tilde{A}^T \Delta \tilde{y}^k + \Delta \tilde{s}^k = 0 \] と
\[ \tilde{S}^k \Delta \tilde{x}^k + \tilde{X}^k \Delta \tilde{s}^k = \sigma \mu^k e - \tilde{X}^k \tilde{s}^k \] を解きニュートン方向 $(\Delta \tilde{x}^k, \Delta \tilde{y}^k, \Delta \tilde{s}^k)$ を得る.

$(\tilde{x}^{k+1}, \tilde{y}^{k+1}, \tilde{s}^{k+1}) := (\tilde{x}^k, \tilde{y}^k, \tilde{s}^k) + (\Delta \tilde{x}^k, \Delta \tilde{y}^k, \Delta \tilde{s}^k)$ とする.

step3: $k := k + 1$ としてstep1に戻る.

3.1. 反復数

$k$ 反復目を考える. 等式制約は次の関係より成り立つ.

\[ \tilde{A} \tilde{x}^{k+1} = \tilde{A} \tilde{x}^k = \tilde{b}, \]
正値性もなりたつ。

次の不等式が成り立つことよりアルゴリズムで得られる次の点も近傍に入る。

最適性の基準として使う、双対ギャップについて、

\[ \Delta \tilde{x}^T \Delta \tilde{s}^k = 0 \]

ことに注意すると、次の評価を\( \Delta \tilde{x}^k T \Delta \tilde{s}^k \)と表現するビット数として、

\[ \| X_{k+1}^t \tilde{s}_{k+1} - \mu_{k+1} \|_2 = \| ( \tilde{X}_k + \Delta \tilde{X}_k ) ( \tilde{s}_k + \Delta \tilde{s}_k ) - \mu_{k+1} \|_2 \] (3.18)

\[ = \| \Delta \tilde{X}_k \Delta \tilde{s}_k \|_2 \] (3.19)

\[ = \| D^{-1} \Delta \tilde{X}_k D \Delta \tilde{s}_k \|_2 \] ただし \( D := \tilde{X}_k^{1/2} \tilde{s}_k^{-1/2} \) (3.20)

\[ \leq \frac{\sqrt{2}}{4} \| D^{-1} \Delta \tilde{x}_k + D \Delta \tilde{s}_k \|_2 \] (3.21)

\[ = \frac{\sqrt{2}}{4} \| ( \tilde{X}_k \tilde{s}_k )^{-1/2} ( \sigma \mu_k e - \tilde{X}_k \tilde{s}_k ) \|_2 \] (3.22)

\[ \leq \frac{\sqrt{2}}{4} \| \tilde{X}_k \tilde{s}_k^k - \sigma \mu_k e \|_2 \] (3.23)

\[ \leq \frac{\sqrt{2}}{4} \| ( \tilde{X}_k \tilde{s}_k^k - \mu_k e ) + (1 - \sigma) \mu_k e \|_2 \] (3.24)

\[ \leq \frac{\sqrt{2}}{4} \frac{0.4^2 + (1 - \sigma)^2 (n + \sum_{p=1}^q l_p)}{(1 - 0.4)} \] (3.25)

\[ \leq \frac{32 \sqrt{2}}{240} \mu_k \] (3.26)

\[ \leq 0.4 \mu_k^{k+1} \] (3.27)

正値性もなりたつ。

一反復で双対ギャップを \( 1 - 0.4/\sqrt{n + \sum_{p=1}^q l_p} \) にできるので、アルゴリズムは \( L \) を問題を表すビット数として、\( O(\sqrt{n + \sum_{p=1}^q l_p} L) \) 反復で最適解を得ることができる。

以下では、簡単のため二つの区分線形凸関数の和の場合を考える。ニュートン方向は次の方程式系の解として得られる。

\[ \tilde{X} \tilde{S}^{k-1} \tilde{A}^T \tilde{y}^k = -\tilde{A} \tilde{S}^{k-1} ( \sigma \mu_k e - \tilde{X} \tilde{s}_k ) \] (3.28a)

\[ \Delta \tilde{s}_k = -\tilde{A}^T \Delta \tilde{x}_k \] (3.28b)

\[ \Delta \tilde{x}_k = \tilde{S}^{k-1} ( \sigma \mu_k e - \tilde{X} \tilde{s}_k ) - \tilde{X} \tilde{S}^{k-1} \Delta \tilde{s}_k \] (3.28c)
次の表現を使う。

\[
\begin{pmatrix}
A \\
(c_1 - c_l)^T \\
\vdots \\
(c_i - c_l)^T \\
\vdots \\
(c_{i-1} - c_l)^T \\
(f_1 - f_l)^T \\
\vdots \\
(f_j - f_l)^T \\
\vdots \\
(f_{l-1} - f_l)
\end{pmatrix}
\cdot \text{diag}(x) \cdot \text{diag}(v_z)^{-1}
\begin{pmatrix}
A \\
(c_1 - c_l)^T \\
\vdots \\
(c_i - c_l)^T \\
\vdots \\
(c_{i-1} - c_l)^T \\
(f_1 - f_l)^T \\
\vdots \\
(f_j - f_l)^T \\
\vdots \\
(f_{l-1} - f_l)
\end{pmatrix}^T
\]

係数行列は次のように表せる。

\[
\begin{pmatrix}
O \\
\ast_{11}v_{s1}^{-1} \\
\ast_{12}v_{s2}^{-1} \\
\ast_{21}v_{s2}^{-1} \\
\ast_{22}v_{s2}^{-1}
\end{pmatrix}
\]  

\[
\begin{pmatrix}
0 \\
0 \\
0 \\
0 \\
0
\end{pmatrix}
\begin{pmatrix}
e \\
e \\
e \\
e
\end{pmatrix}
\]

\[
\tilde{A} := A + s_1 l_{v_{s1}}^{-1} \begin{bmatrix} 0 \\ 0 \\ e \\ 0 \end{bmatrix} + s_2 l_{v_{s2}}^{-1} \begin{bmatrix} 0 \\ 0 \\ 0 \\ e \end{bmatrix}
\]  

\[
(\hat{A} + BDC)^{-1} = A^{-1} - A^{-1}B(D + DCA^{-1}BD)^{-1}DCA^{-1}
\]

が成り立つことに注意する。

\[
\hat{A} := A + s_1 l_{v_{s1}}^{-1} \begin{bmatrix} 0 \\ 0 \\ e \\ 0 \end{bmatrix}
\]  

\[
(3.29)
\]

\[
(3.30)
\]

\[
(3.31)
\]
とすると，

\[
\hat{A}^{-1} = \tilde{A}^{-1} - \frac{s_{l_1} v_{s_{l_1}}^{-1}}{1 + s_{l_1} v_{s_{l_1}}^{-1}} \tilde{A}^{-1} \left[ \begin{array}{c} 0 \\ \delta \\ 0 \end{array} \right] \left[ \begin{array}{c} 0 \\ e \\ 0 \end{array} \right]^{T} \tilde{A}^{-1}.
\]

（3.33）

の関係を用いて，\(\hat{A}^{-1}\)を求める。そして，

\[
\left( \hat{A}^T \hat{X}^k \hat{S}^{k^{-1}} \hat{A} \right)^{-1} = \hat{A}^{-1} - \frac{s_{l_2} v_{s_{l_2}}^{-1}}{1 + s_{l_2} v_{s_{l_2}}^{-1}} \hat{A}^{-1} \left[ \begin{array}{c} 0 \\ \delta \end{array} \right] \left[ \begin{array}{c} 0 \\ e \end{array} \right]^{T} \hat{A}^{-1}.
\]

（3.34）

の関係を用いて，ニュートン方向の計算を得る。この手法は，複数の区分線形凸関数の和の場合も適用できる。

参考文献