

特別研究報告書

分離可能な凸計画問題に対する交互方向乗数法と
二次錐計画問題への応用

指導教員 福嶋雅夫 教授

京都大学工学部情報学科
数理工学コース
平成16年4月入学
平成20年3月卒業

大坪 亮太

平成20年1月31日提出

分離可能な凸計画問題に対する交互方向乗数法と 二次錐計画問題への応用

大坪 亮太

摘要

目的関数と制約関数がいくつかの独立した変数を含む関数の和で表される制約付き最適化問題を、分離可能な最適化問題という。制約付き最適化問題の代表的な数値解法の一つとして、乗数法が知られている。しかし、乗数法を分離可能な最適化問題にそのまま適用すると、各反復に現れる部分問題は各変数について分離可能とならないので、もとの問題の分離可能性を活かせない。そこで、分離可能構造を十分に利用できるように乗数法を改良した方法として、交互方向乗数法が提案されている。

本報告書では、閉凸集合を制約に持つ最適化問題に対して、適当な仮定の下で、交互方向乗数法によって生成された点列が、最適解に収束することを証明する。特に、二次錐を制約条件に含む、ある種の二次錐計画問題に対して、交互方向乗数法の各反復に現れる部分問題の解が、陽に求まることを示す。この性質を利用し、交互方向乗数法で問題を解いた場合と、問題の分離可能性を利用せずに解いた場合とを、数値実験により比較し、交互方向乗数法の有効性を検証する。

目次

1	序論	1
2	交互方向乗数法	2
2.1	凸計画問題に対する交互方向乗数法	2
2.2	凸計画問題に対する交互方向乗数法の収束性	3
2.3	分離可能な凸計画問題に対する交互方向乗数法	7
3	二次錐計画問題への応用	8
3.1	交互方向乗数法の二次錐計画問題への応用	8
4	数値実験	9
4.1	パラメータ c の影響	10
4.2	問題の大きさの影響	13
4.3	既存のソルバーとの比較	14
4.3.1	目的関数が二次関数の場合	14
4.3.2	目的関数が一次関数の場合	17
5	結論	19
A	ラグランジュ乗数定理	21
B	分離可能な凸計画問題に対する交互方向乗数法の導出	22

1 序論

以下の形の分離可能な凸計画問題を考える．

$$\begin{aligned} & \text{minimize} && \sum_{i=1}^m F_i(x_i) \\ & \text{subject to} && \sum_{i=1}^m A_i x_i = b \\ & && x_i \in K_i, \quad i = 1, \dots, m \end{aligned} \tag{1.1}$$

ここで， $F_i : \mathbb{R}^{n_i} \rightarrow \mathbb{R}$ は凸関数， $x = (x_1, \dots, x_m)^\top$ であり， x_i は n_i 次元の部分ベクトルである．また， A_i は $r \times n_i$ 行列で， b は r 次元のベクトル， K_i は空でない閉凸集合である．

制約付き最適化問題の代表的な数値解法の一つとして，乗数法 (The method of multipliers) [2, 7, 8] が知られている．乗数法は，1969 年に Hestenes と Powell によって提案された数値解法である．等式制約付き問題のラグランジュ関数にペナルティ項を付加した拡張ラグランジュ関数が局所的に凸になることから，その拡張関数を無制約最小化することで問題の解を求める数値解法のことである．乗数法の利点の一つは，通常のペナルティ法のようにペナルティパラメータを無限に大きくする必要がないことである．

しかし，乗数法を分離可能な最適化問題にそのまま適用すると，各反復に現れる部分問題は各変数について分離可能とならないので，もとの問題の分離可能性を活かせない．そこで，分離可能構造を十分に利用できるように乗数法を改良した方法として，交互方向乗数法 (The alternating direction method of multipliers) [2, 3] が提案されている．Bertsekas と Tsitsiklis [2] は，分離可能な問題が，交互方向乗数法で効率的に扱えることを示している．また，多面体制約を持つ最適化問題に対して，交互方向乗数法により生成された点列の任意の集積点は，ある仮定の下で，最適解となることが示されている [2]．

一方，二次錐計画問題 (Second-order cone programming problem) とは，二次錐の直積に含まれるという制約の下で目的関数を最小化する最適化問題のことである．線形計画問題や凸二次計画問題，二次制約付き凸二次計画問題は，二次錐計画問題に定式化できる．二次錐計画問題は，工学や金融，ロバスト最適化，組み合わせ最適化などの分野に広く応用できる問題のクラスである [1]．

本報告書では，閉凸集合を制約に持つ最適化問題に対して，適当な仮定の下で，交互方向乗数法によって生成された点列が，最適解に収束することを証明する．特に，二次錐計画問題

$$\begin{aligned} & \text{minimize} && \sum_{i=1}^m \left(\frac{1}{2} \alpha_i \|x_i\|^2 + \gamma_i^\top x_i \right) \\ & \text{subject to} && \sum_{i=1}^m x_i = b \\ & && x_i \in K_i, \quad i = 1, \dots, m \end{aligned} \tag{1.2}$$

(ただし， $\alpha_i \geq 0$ ， γ_i ， b は r 次元ベクトル， K_i は二次錐) に対して，交互方向乗数法の各反復に現れる部分問題の解が，陽に求まることを示す．この性質を利用し，交互方向乗数法で問題を解いた場合と，問題の分離可能性を利用せずに解いた場合とを，数値実験により比較する．

以下，本報告書の構成を示す．2.1 節で凸計画問題に対する交互方向乗数法を紹介し，2.2 節でこの数値解法の収束性を証明する．2.3 節で分離可能な凸計画問題に交互方向乗数法を適用し，このとき各反復において分解法が適用できる形になっていることを示す．3 節では交互方向乗数法を二次錐計画問題に適用し，特に問題 (1.2) の場合，各反復に現れる部分問題の解が陽に求まることを

示す．4節で数値実験の結果を報告する．最後に5節で結論を述べる．また，付録Aに2.1節の交互方向乗数法の証明に用いたラグランジュ乗数定理の証明を述べる．付録Bには，交互方向乗数法を分離可能な問題に適用する過程の計算を載せる．

2 交互方向乗数法

本節では，凸計画問題に対する交互方向乗数法を紹介し，この方法の収束性を証明する．また，これを分離可能な凸計画問題に適用する．

2.1 凸計画問題に対する交互方向乗数法

まず，制約付き最適化問題の数値解法の一つである交互方向乗数法について説明する．交互方向乗数法は次の形の最適化問題について定義される．

$$\begin{aligned} & \text{minimize} && G_1(x) + G_2(Bx) \\ & \text{subject to} && x \in C_1, Bx \in C_2 \end{aligned} \quad (2.1)$$

ここで， $G_1: \mathbb{R}^n \rightarrow \mathbb{R}$ で， $G_2: \mathbb{R}^l \rightarrow \mathbb{R}$ は凸関数とする． B は $l \times n$ 行列で， $C_1 \subset \mathbb{R}^n$ と $C_2 \subset \mathbb{R}^l$ は空でない閉凸集合とする．変数 $z \in \mathbb{R}^l$ を用いると，問題 (2.1) は次の等価な問題に変形できる．

$$\begin{aligned} & \text{minimize} && G_1(x) + G_2(z) \\ & \text{subject to} && x \in C_1, z \in C_2, Bx = z \end{aligned} \quad (2.2)$$

この問題の拡張ラグランジュ関数は， $p \in \mathbb{R}^l$ を等式制約に対するラグランジュ乗数， c を正の数とすると，

$$L_c(x, z, p) = G_1(x) + G_2(z) + p^\top (Bx - z) + \frac{c}{2} \|Bx - z\|^2$$

と書ける．このとき，交互方向乗数法の各反復は次式で与えられる．

$$x^{(k+1)} = \operatorname{argmin}_{x \in C_1} \left\{ G_1(x) + p^{(k)\top} Bx + \frac{c}{2} \|Bx - z^{(k)}\|^2 \right\} \quad (2.3)$$

$$z^{(k+1)} = \operatorname{argmin}_{z \in C_2} \left\{ G_2(z) - p^{(k)\top} z + \frac{c}{2} \|Bx^{(k+1)} - z\|^2 \right\} \quad (2.4)$$

$$p^{(k+1)} = p^{(k)} + c(Bx^{(k+1)} - z^{(k+1)}) \quad (2.5)$$

ただし， c は正のパラメータで $p^{(0)}$ と $z^{(0)}$ は任意に選ばれた初期ベクトルである．ここで，式 (2.3) は拡張ラグランジュ関数を $x \in C_1$ という制約の下で最小化する x を与える式であり，式 (2.4) は拡張ラグランジュ関数を $z \in C_2$ という制約の下で最小化する z を与える式である．式 (2.5) はラグランジュ乗数の更新式である．このとき，交互方向乗数法のアルゴリズムを以下に示す．

交互方向乗数法のアルゴリズム

Step 1 初期値 $p^{(0)}, z^{(0)}$ を決める． $k = 0$ とする．

Step 2 終了条件を満たせば，終了する．終了条件を満たさなければ，step3へ．

Step 3 部分問題 (2.3) を解き， $x^{(k+1)}$ を求める．部分問題 (2.4) を解き， $z^{(k+1)}$ を求める．式 (2.5) から $p^{(k+1)}$ を求める． $k := k + 1$ として step2へ．

次節では，交互方向乗数法により生成された点列が最適解に収束することを示す．

2.2 凸計画問題に対する交互方向乗数法の収束性

まず、式 (2.3) の $x^{(k+1)}$ が一意に決まることと、式 (2.4) の $z^{(k+1)}$ が一意に決まることを示すために、次の命題が成り立つことを示す。

命題 1 $F: \mathfrak{R}^n \rightarrow \mathfrak{R}$ を凸関数、 X を空でない閉凸集合とする。このとき、任意の $y \in \mathfrak{R}^n$ と $c > 0$ に対して、 $F(x) + (c/2)\|x - y\|^2$ を集合 X 上で x について最小化する点 $x(y, c)$ が一意に存在する。

証明 任意の $c > 0$ と $y \in \mathfrak{R}^n$ に対して、レベル集合

$$\left\{ x \in X \mid F(x) + \frac{c}{2}\|x - y\|^2 \leq \alpha \right\}, \quad \forall \alpha \in \mathfrak{R} \quad (2.6)$$

が有界であることを示す。これを示せば、 $F(x) + (c/2)\|x - y\|^2$ の x に関する狭義凸性より、最小点 $x(y, c)$ が一意に決まることがいえる。

いま、ある $c > 0$ 、 $y \in \mathfrak{R}^n$ に対してレベル集合 (2.6) が有界でないと仮定すると、

$$\|x^k - y\| \rightarrow \infty, \quad F(x^k) + \frac{c}{2}\|x^k - y\|^2 \leq \alpha, \quad \forall k \quad (2.7)$$

をみたく点列 $\{x^k\}$ が存在する。 $\beta_k = \|x^k - y\|$ とし、一般性を失うことなく任意の k に対して $\beta_k \geq 1$ と仮定する。 $z^k = (x^k - y)/\beta_k$ とおき、 $\hat{F}(x) = F(x + y)$ で凸関数 \hat{F} を定義する。すると、式 (2.7) より、

$$\hat{F}(\beta_k z^k) + \frac{c}{2}(\beta_k)^2 = \hat{F}(x^k - y) + \frac{c}{2}\|x^k - y\|^2 \leq \alpha, \quad \forall k \quad (2.8)$$

が成り立つ。また、 \hat{F} の凸性より、

$$\min_{\|z\|=1} \hat{F}(z) \leq \hat{F}(z^k) \leq \frac{1}{\beta_k} \hat{F}(\beta_k z^k) + \left(1 - \frac{1}{\beta_k}\right) \hat{F}(0)$$

となり、これより次式を得る。

$$(1 - \beta_k) \hat{F}(0) + \beta_k \min_{\|z\|=1} \hat{F}(z) \leq \hat{F}(\beta_k z^k)$$

この式と、式 (2.8) より、

$$\hat{F}(0) + \beta_k \left(\min_{\|z\|=1} \hat{F}(z) - \hat{F}(0) \right) + \frac{c}{2}(\beta_k)^2 \leq \alpha, \quad \forall k \quad (2.9)$$

が成立する。しかし、 $c > 0$ かつ $\beta_k \rightarrow \infty$ であるから式 (2.9) は矛盾である。ゆえに、レベル集合 (2.6) が有界であることが示せた。 ■

この命題 1 により、(2.4) 式の $z^{(k+1)}$ が一意に定まることがわかる。ここで、次の仮定をおく。

仮定 1 問題 (2.1) の最適解集合は空ではなく、行列 $B^\top B$ は正則である。さらに、 $x \in \text{int } C_1$ かつ $Bx \in \text{int } C_2$ を満たす x が存在する¹。

$B^\top B$ が正則であれば、 $B^\top B$ は正定値となる。これより (2.3) 式の右辺の関数が x に関して狭義凸であることが言え、命題 1 と同様に、 $x^{(k+1)}$ も一意に定まる。

次に凸計画問題 (2.1) に対する交互方向乗数法 (2.3)(2.4) (2.5) の収束性を証明する。

¹ C_1 または C_2 が凸多面体の場合、内部という条件は必要なく、凸多面体上の点が存在すればよい。

命題 2 仮定 1 の下で, (2.3)–(2.5) 式のアゴリズムで生成された $\{x^{(k)}, z^{(k)}, p^{(k)}\}$ は有界で, $\{x^{(k)}\}$ は元の問題 (2.1) の最適解に収束し, $\{p^{(k)}\}$ は次の双対問題の最適解 p^* に収束する.

$$\begin{aligned} & \text{maximize} && H_1(p) + H_2(p) \\ & \text{subject to} && p \in \mathfrak{R}^l \end{aligned} \quad (2.10)$$

ただし, 任意の $p \in \mathfrak{R}^l$ に対して, $H_1(p), H_2(p)$ は次式で定義される.

$$H_1(p) = \inf_{x \in C_1} \{G_1(x) + p^\top Bx\}, \quad H_2(p) = \inf_{z \in C_2} \{G_2(z) - p^\top z\} \quad (2.11)$$

以下の補助定理 1 は命題 2 の証明に使用する.

補助定理 1 $J_1: \mathfrak{R}^n \rightarrow \mathfrak{R}$ と $J_2: \mathfrak{R}^n \rightarrow \mathfrak{R}$ は凸関数, Y は \mathfrak{R}^n 上の閉凸集合で, その内部は空ではなく, J_2 は連続的の微分可能とする. そのとき, $y^* = \operatorname{argmin}_{y \in Y} \{J_1(y) + J_2(y)\}$ とすると,

$$y^* = \operatorname{argmin}_{y \in Y} \{J_1(y) + \nabla J_2(y^*)^\top y\} \quad (2.12)$$

が成り立つ.

証明 まず,

$$\begin{aligned} (y^*, z^*) &= \operatorname{argmin}_{y \in Y, z \in \mathfrak{R}^n, z=y} \{J_1(y) + J_2(z)\} \\ z^* &= y^* \end{aligned}$$

であることに注意する. ラグランジュ乗数定理 (付録参照) により, 次式を満足する $\lambda \in \mathfrak{R}^n$ が存在する.

$$y^* = \operatorname{argmin}_{y \in Y} \{J_1(y) + \lambda^\top y\} \quad (2.13)$$

$$y^* = \operatorname{argmin}_{z \in \mathfrak{R}^n} \{J_2(z) - \lambda^\top z\} \quad (2.14)$$

式 (2.14) から, $\lambda = \nabla J_2(y^*)$ となり, 式 (2.13) から式 (2.12) が得られる. ■

命題 2 の証明 補助定理 1 より, $Y = C_1$, $J_1(x) = G_1(x)$, $J_2(x) = p^{(k)\top} Bx + (c/2)\|Bx - z^{(k)}\|^2$ とおくと, 式 (2.3) より次式を得る.

$$\begin{aligned} & G_1(x^{(k+1)}) + [p^{(k)} + c(Bx^{(k+1)} - z^{(k)})]^\top Bx^{(k+1)} \\ & \leq G_1(x) + [p^{(k)} + c(Bx^{(k+1)} - z^{(k)})]^\top Bx, \quad \forall x \in C_1 \end{aligned} \quad (2.15)$$

同様に, $Y = C_2$, $J_1(z) = G_2(z)$, $J_2(z) = -p^{(k)\top} z + (c/2)\|Bx^{(k+1)} - z\|^2$ とおくと, 式 (2.4) より次式を得る.

$$\begin{aligned} & G_2(z^{(k+1)}) - [p^{(k)} + c(Bx^{(k+1)} - z^{(k+1)})]^\top z^{(k+1)} \\ & \leq G_2(z) - [p^{(k)} + c(Bx^{(k+1)} - z^{(k+1)})]^\top z, \quad \forall z \in C_2 \end{aligned} \quad (2.16)$$

問題 (2.2) の最適解を (x^*, z^*) , p^* を双対問題 (2.10) の最適解と仮定する. 式 (2.15) で $x = x^*$, 式 (2.16) で $z = z^*$ とし, ラグランジュ乗数の更新式 (2.5) を使うことで, 次式が成立する.

$$\begin{aligned} & G_1(x^{(k+1)}) + p^{(k+1)\top} Bx^{(k+1)} + c(z^{(k+1)} - z^{(k)})^\top Bx^{(k+1)} \\ & \leq G_1(x^*) + p^{(k+1)\top} Bx^* + c(z^{(k+1)} - z^{(k)})^\top Bx^* \\ & G_2(z^{(k+1)}) - p^{(k+1)\top} z^{(k+1)} \leq G_2(z^*) - p^{(k+1)\top} z^* \end{aligned}$$

この二つの関係式の両辺を加え， $Bx^* = z^*$ という関係から次式を得る．

$$\begin{aligned} G_1(x^{(k+1)}) + G_2(z^{(k+1)}) + p^{(k+1)\top} (Bx^{(k+1)} - z^{(k+1)}) + c(z^{(k+1)} - z^{(k)})^\top B(x^{(k+1)} - x^*) \\ \leq G_1(x^*) + G_2(z^*) \end{aligned} \quad (2.17)$$

一方，鞍点定理 [4] より，

$$G_1(x^*) + G_2(z^*) \leq G_1(x^{(k+1)}) + G_2(z^{(k+1)}) + p^{*\top} (Bx^{(k+1)} - z^{(k+1)}), \quad \forall k \quad (2.18)$$

が成立する．式 (2.17) に (2.18) を加えると，次式が成立する．

$$(p^{(k+1)} - p^*)^\top (Bx^{(k+1)} - z^{(k+1)}) + c(z^{(k+1)} - z^{(k)})^\top B(x^{(k+1)} - x^*) \leq 0 \quad (2.19)$$

ここで任意の k に対して次のように定義する．

$$\bar{x}^{(k)} = x^{(k)} - x^*, \quad \bar{z}^{(k)} = z^{(k)} - z^*, \quad \bar{p}^{(k)} = p^{(k)} - p^* \quad (2.20)$$

$Bx^* = z^*$ より，乗数の更新公式 (2.5) は，

$$\bar{p}^{(k+1)} = \bar{p}^{(k)} + c(B\bar{x}^{(k+1)} - \bar{z}^{(k+1)}) \quad (2.21)$$

と書け，また，

$$\bar{p}^{(k+1)} = \bar{p}^{(k)} + c(Bx^{(k+1)} - z^{(k+1)}) \quad (2.22)$$

と書ける．式 (2.19) に (2.20), (2.22) を用いると，

$$\frac{1}{c} \bar{p}^{(k+1)\top} (\bar{p}^{(k+1)} - \bar{p}^{(k)}) + c(\bar{z}^{(k+1)} - \bar{z}^{(k)})^\top B\bar{x}^{(k+1)} \leq 0$$

となり，この式と式 (2.21) より，

$$\frac{1}{c} \bar{p}^{(k+1)\top} (\bar{p}^{(k+1)} - \bar{p}^{(k)}) + c(\bar{z}^{(k+1)} - \bar{z}^{(k)})^\top \bar{z}^{(k+1)} + (\bar{z}^{(k+1)} - \bar{z}^{(k)})^\top (\bar{p}^{(k+1)} - \bar{p}^{(k)}) \leq 0 \quad (2.23)$$

が成立する．

この関係式の三つの項はそれぞれ次のようになる．

$$\bar{p}^{(k+1)\top} (\bar{p}^{(k+1)} - \bar{p}^{(k)}) = \frac{1}{2} \|\bar{p}^{(k+1)} - \bar{p}^{(k)}\|^2 + \frac{1}{2} \|\bar{p}^{(k+1)}\|^2 - \frac{1}{2} \|\bar{p}^{(k)}\|^2 \quad (2.24)$$

$$(\bar{z}^{(k+1)} - \bar{z}^{(k)})^\top \bar{z}^{(k+1)} = \frac{1}{2} \|\bar{z}^{(k+1)} - \bar{z}^{(k)}\|^2 + \frac{1}{2} \|\bar{z}^{(k+1)}\|^2 - \frac{1}{2} \|\bar{z}^{(k)}\|^2 \quad (2.25)$$

式 (2.23) の三番目の項を計算するために，関係式 (2.16) で $z = z^{(k)}$ とすると，

$$G_2(z^{(k+1)}) - p^{(k+1)\top} z^{(k+1)} \leq G_2(z^{(k)}) - p^{(k+1)\top} z^{(k)} \quad (2.26)$$

となる．いま，補助定理 1 で $Y = C_2$, $J_1(z) = G_2(z)$, $J_2(z) = -p^{(k-1)\top} z + (c/2) \|Bx^{(k)} - z\|^2$ とすると，式 (2.4) より，

$$\begin{aligned} G_2(z^{(k)}) - [p^{(k-1)} + c(Bx^{(k)} - z^{(k)})]^\top z^{(k)} \\ \leq G_2(z) - [p^{(k-1)} + c(Bx^{(k)} - z^{(k)})]^\top z, \quad \forall z \in C_2 \end{aligned}$$

が成立する．この式で $z = z^{(k+1)}$ とし，式 (2.5) を用いると，

$$G_2(z^{(k)}) - p^{(k)\top} z^{(k)} \leq G_2(z^{(k+1)}) - p^{(k)\top} z^{(k+1)} \quad (2.27)$$

が成立する．式 (2.26) に式 (2.27) を加えることで， $0 \leq (z^{(k+1)} - z^{(k)})^\top (p^{(k+1)} - p^{(k)})$ を得る．これは，次と等価である．

$$0 \leq (\bar{z}^{(k+1)} - \bar{z}^{(k)})^\top (\bar{p}^{(k+1)} - \bar{p}^{(k)}) \quad (2.28)$$

不等式 (2.23) に式 (2.24), (2.25), (2.28) を用いると，次式を得る．

$$\|\bar{p}^{(k+1)} - \bar{p}^{(k)}\|^2 + c^2 \|\bar{z}^{(k+1)} - \bar{z}^{(k)}\|^2 \leq (\|\bar{p}^{(k)}\|^2 + c^2 \|\bar{z}^{(k)}\|^2) - (\|\bar{p}^{(k+1)}\|^2 + c^2 \|\bar{z}^{(k+1)}\|^2) \quad (2.29)$$

この式から，点列 $\{\|\bar{p}^{(k)}\|^2 + c^2 \|\bar{z}^{(k)}\|^2\}$ は非増加かつ下に有界で，ある値に収束する．ゆえに，この式の右辺は， $k \rightarrow \infty$ において 0 となることより，

$$\bar{p}^{(k+1)} - \bar{p}^{(k)} \rightarrow 0, \quad \bar{z}^{(k+1)} - \bar{z}^{(k)} \rightarrow 0 \quad (2.30)$$

である．ここで，式 (2.17), (2.18), (2.22) より， $G_1(x^{(k+1)}) + G_2(z^{(k+1)})$ を上と下からおさえ， $k \rightarrow \infty$ とすることで，

$$\lim_{k \rightarrow \infty} [G_1(x^{(k+1)}) + G_2(z^{(k+1)})] = G_1(x^*) + G_2(z^*) = \min_{x \in C_1, z \in C_2, Bx=z} \{G_1(x) + G_2(z)\} \quad (2.31)$$

が成立する．さらに， $\{(x^{(k)}, z^{(k)})\}$ の任意の集積点 (\tilde{x}, \tilde{z}) について， $B\tilde{x} = \tilde{z}$ となるので， \tilde{x} は問題 (2.1) の最適解である．

式 (2.29), (2.20) から， $\{p^{(k)}\}$ と $\{z^{(k)}\}$ は有界で，式 (2.5), (2.30) から， $\|Bx^{(k)} - z^{(k)}\| \rightarrow 0$ となる．また，仮定 1 より $x^{(k)} = (B^\top B)^{-1} B^\top z^{(k)}$ と書け， $\{z^{(k)}\}$ が有界なことから， $\{x^{(k)}\}$ も有界である．ここで，収束する部分列 $\{(x^{(k)}, z^{(k)}, p^{(k)}) \mid k \in K\}$ を考え， $(\tilde{x}, \tilde{z}, \tilde{p})$ をその極限とする．このとき，先ほど示したように， \tilde{x} は問題 (2.1) の最適解である．次に， \tilde{p} が双対問題 (2.10) の最適解であることを示すため， $\hat{p}^{(k+1)} = p^{(k)} + c(Bx^{(k+1)} - z^{(k)})$ と定義する．定義 (2.11) と，式 (2.15), (2.16) より，

$$H_1(\hat{p}^{(k+1)}) = G_1(x^{(k+1)}) + (\hat{p}^{(k+1)})^\top Bx^{(k+1)} \leq G_1(x) + (\hat{p}^{(k+1)})^\top Bx, \quad \forall x \in C_1 \quad (2.32)$$

$$H_2(p^{(k+1)}) = G_2(z^{(k+1)}) - p^{(k+1)\top} z^{(k+1)} \leq G_2(z) + p^{(k+1)\top} z, \quad \forall z \in C_2 \quad (2.33)$$

が成立する．これらの関係式の極限をとると，

$$\limsup_{k \rightarrow \infty, k \in K} H_1(\hat{p}^{(k+1)}) \leq G_1(x) + \tilde{p}^\top Bx, \quad \forall x \in C_1$$

$$\limsup_{k \rightarrow \infty, k \in K} H_2(p^{(k+1)}) \leq G_2(z) - \tilde{p}^\top z, \quad \forall z \in C_2$$

となることから，

$$\limsup_{k \rightarrow \infty, k \in K} H_1(\hat{p}^{(k+1)}) \leq H_1(\tilde{p}), \quad \limsup_{k \rightarrow \infty, k \in K} H_2(p^{(k+1)}) \leq H_2(\tilde{p}) \quad (2.34)$$

が成立する．一方で，式 (2.32) と (2.33) を加え， $B\tilde{x} = \tilde{z}$ を使うと，次式が成立する．

$$\begin{aligned} \lim_{k \rightarrow \infty, k \in K} [H_1(\hat{p}^{(k+1)}) + H_2(p^{(k+1)})] &= G_1(\tilde{x}) + G_2(\tilde{z}) \\ &= \min_{x \in C_1, z \in C_2, Bx=z} \{G_1(x) + G_2(z)\} \end{aligned} \quad (2.35)$$

式 (2.34) と (2.35) と弱双対定理 [4] から， \tilde{p} は双対問題 (2.10) の最適解であることがわかる．

最後に, $\{z^{(k)}, p^{(k)}\}$ の極限が一意に決まることを示す. 式 (2.29) より,

$$\|p^{(k)} - p^*\|^2 + c^2 \|z^{(k)} - z^*\|^2 \quad (2.36)$$

は, 主問題 (2.2) の任意の最適解 (x^*, z^*) と双対問題 (2.10) の任意の最適解 p^* に対して, 非増加な正数列になる. ここで, $\{(z^{(k)}, p^{(k)})\}$ のある極限 $(\tilde{z}_1, \tilde{p}_1)$ を考える. 点列 (2.36) は非増加かつ下に有界より, ある極限に収束する. この極限を $\alpha \geq 0$ とすると,

$$\lim_{k \rightarrow \infty} \left\{ \|p^{(k)} - \tilde{p}_1\|^2 + c^2 \|z^{(k)} - \tilde{z}_1\|^2 \right\} = \alpha$$

ここで, 点列 $\{(z^{(k)}, p^{(k)})\}$ のなかで, $(\tilde{z}_1, \tilde{p}_1)$ に収束する部分列を考えると, $\alpha = 0$ となる. つまり,

$$\lim_{k \rightarrow \infty} p^{(k)} = \tilde{p}_1, \quad \lim_{k \rightarrow \infty} z^{(k)} = \tilde{z}_1 \quad (2.37)$$

である. 点列 $\{(z^{(k)}, p^{(k)})\}$ の任意の極限に対して同様のことが成立するので, $\{(z^{(k)}, p^{(k)})\}$ は収束する. さらに, 仮定 1 より $B^\top B$ が正則なので, $x^{(k)} = (B^\top B)^{-1} B^\top z^{(k)}$ と書け, $\{x^{(k)}\}$ の極限も一意に決まる. \blacksquare

2.3 分離可能な凸計画問題に対する交互方向乗数法

前節で収束性を証明した交互方向乗数法を分離可能な凸計画問題 (1.1) に適用する. 行列 A_1, \dots, A_m を横に並べた行列を $A \in \mathbb{R}^{r \times n}$ とし, A^\top の第 j 列ベクトルを a_j , ベクトル b の第 j 成分を b_j , 部分ベクトル x_1, \dots, x_m を縦に並べたベクトルを x とすると, 問題 (1.1) は以下の形に変形できる.

$$\begin{aligned} & \text{minimize} && \sum_{i=1}^m F_i(x_i) \\ & \text{subject to} && a_j^\top x = b_j, \quad j = 1, \dots, r \\ & && x_i \in K_i, \quad i = 1, \dots, m \end{aligned} \quad (2.38)$$

このとき, 問題 (2.1) において,

$$\begin{aligned} G_1(x) &= \sum_{i=1}^m F_i(x_i), \quad C_1 = K_1 \times K_2 \times \dots \times K_m \\ G_2(z) &= 0, \quad C_2 = \left\{ z \mid \sum_{i=1}^m z_{ji} = b_j, \quad j = 1, \dots, r \right\} \end{aligned}$$

とし, また, a_{ji} を, a_j の部分ベクトル, B を x から $a_{ji}^\top x_i$, $j = 1, \dots, r$, $i = 1, \dots, m$ を成分に持つベクトルへの写像を表す行列とすることで, 問題 (2.38) が得られる. つまり, 問題 (2.1) は問題 (2.38) を含む問題であることがわかる. このことから問題 (2.38), つまり問題 (1.1) に対する交互方向乗数法は, 上の式を式 (2.3)(2.4)(2.5) に代入することにより与えられる. 計算は付録に載せる.

$$x_i^{(k+1)} = \operatorname{argmin}_{x_i \in K_i} \left\{ F_i(x_i) + \sum_{j=1}^r \left\{ \lambda_j^{(k)} a_{ji}^\top x_i + \frac{c}{2} (a_{ji}^\top (x_i - x_i^{(k)}) + w_j^{(k)})^2 \right\} \right\}, \quad i = 1, \dots, m \quad (2.39)$$

$$\lambda_j^{(k+1)} = \lambda_j^{(k)} + c w_j^{(k+1)}, \quad j = 1, \dots, r \quad (2.40)$$

$$w_j^{(k)} = \frac{1}{m} (a_j^\top x^{(k)} - b_j), \quad j = 1, \dots, r \quad (2.41)$$

ここで, 式 (2.39) は, 各部分ベクトル x_i に分解されており, i について独立な問題なので, これらを並列に計算できる.

3 二次錐計画問題への応用

本節では，前節で述べた分離可能な凸計画問題に対する交互方向乗数法を，ある種の二次錐計画問題に適用する．

3.1 交互方向乗数法の二次錐計画問題への応用

まず二次錐を定義する． n 次元ベクトル x の第一成分を x_0 とし，残りの成分からなるベクトルを \bar{x} とする．このとき， n 次元の二次錐は

$$K = \{x \in \mathfrak{R}^n \mid x_0 \geq \|\bar{x}\|\}$$

で定義される．特に，制約に二次錐を含む最適化問題を二次錐計画問題という．

問題 (1.1) で，任意の i について， $A_i = I$ (r 次元の単位行列)， K_i は r 次元の二次錐， $F_i(x_i) = \frac{1}{2}\alpha_i\|x_i\|^2 + \gamma_i^\top x_i$ ($\alpha_i \geq 0$) とすると，以下の形の二次錐計画問題となる．

$$\begin{aligned} & \text{minimize} && \sum_{i=1}^m \left(\frac{1}{2}\alpha_i\|x_i\|^2 + \gamma_i^\top x_i \right) \\ & \text{subject to} && \sum_{i=1}^m x_i = b \\ & && x_i \in K_i, \quad i = 1, \dots, m \end{aligned} \tag{3.1}$$

このとき，交互方向乗数法の各反復に現れる部分問題 (2.39) は，

$$\begin{aligned} x_i^{(k+1)} &= \operatorname{argmin}_{x_i \in K_i} \sum_{j=1}^r \left\{ \frac{1}{2}\alpha_i x_{ij}^2 + \gamma_{ij} x_{ij} + \lambda_j^{(k)} x_{ij} + \frac{c}{2}(x_{ij} - x_{ij}^{(k)} + w_j^{(k)})^2 \right\} \\ &= \operatorname{argmin}_{x_i \in K_i} \sum_{j=1}^r \left\{ \frac{\alpha_i + c}{2} \left(x_{ij} + \frac{1}{\alpha_i + c}(\gamma_{ij} + \lambda_j^{(k)} - c(x_{ij}^{(k)} - w_j^{(k)})) \right)^2 \right. \\ & \quad \left. + \frac{c}{2}(x_{ij}^{(k)} - w_j^{(k)})^2 - \frac{1}{\alpha_i + c} \left\{ \frac{1}{\alpha_i + c}(\gamma_{ij} + \lambda_j^{(k)} + c(w_j^{(k)} - x_{ij}^{(k)})) \right\}^2 \right\} \end{aligned}$$

となる．ここで，

$$\nu_{ij}^{(k)} = -\frac{1}{\alpha_i + c}(\gamma_{ij} + \lambda_j^{(k)} + c(w_j^{(k)} - x_{ij}^{(k)})) \tag{3.2}$$

とおくと， $\alpha_i + c \geq 0$ より，

$$\begin{aligned} x_i^{(k+1)} &= \operatorname{argmin}_{x_i \in K_i} \sum_{j=1}^r (x_{ij} - \nu_{ij}^{(k)})^2 \\ &= \operatorname{argmin}_{x_i \in K_i} \|x_i - \nu_i^{(k)}\|^2, \quad i = 1, \dots, m \end{aligned} \tag{3.3}$$

ただし， $\nu_i^{(k)}$ は $-(\gamma_{ij} + \lambda_j^{(k)} + c(w_j^{(k)} - x_{ij}^{(k)}))/(\alpha_i + c)$ を第 j 成分に持つ r 次元の列ベクトルとする．

部分問題 (3.3) は，二次錐 K_i に含まれる点のなかで，点 $\nu_i^{(k)}$ との距離が最小になる点が解となる．つまり，点 $\nu_i^{(k)}$ を二次錐 K_i に射影した点が解となる．この解は次のように陽に求めることができる [5]．

$\nu_i^{(k)}$ の第一成分を $\nu_{i1}^{(k)}$, 他の成分からなるベクトルを $\bar{\nu}_i^{(k)}$ とすると ,

$$\nu_i^{(k)} = \begin{bmatrix} \nu_{i1}^{(k)} \\ \bar{\nu}_i^{(k)} \end{bmatrix}$$

と書け , このとき部分問題 (3.3) の解は , $i = 1, \dots, m$ に対して ,

$$x_i^{(k+1)} = \begin{cases} \max\{0, \frac{1}{2}(\nu_{i1}^{(k)} + \|\bar{\nu}_i^{(k)}\|)\} \times \begin{bmatrix} 1 \\ \bar{\nu}_i^{(k)}/\|\bar{\nu}_i^{(k)}\| \end{bmatrix} + \max\{0, \frac{1}{2}(\nu_{i1}^{(k)} - \|\bar{\nu}_i^{(k)}\|)\} \times \begin{bmatrix} 1 \\ -\bar{\nu}_i^{(k)}/\|\bar{\nu}_i^{(k)}\| \end{bmatrix} & (\|\bar{\nu}_i^{(k)}\| \neq 0) \\ \max\{0, \nu_{i1}^{(k)}\} \times \begin{bmatrix} 1 \\ 0 \end{bmatrix} & (\|\bar{\nu}_i^{(k)}\| = 0) \end{cases} \quad (3.4)$$

と陽に求まる .

以上より問題 (3.1) に対する交互方向乗数法は , (3.2) 式 , (3.4) 式および

$$\lambda_j^{(k+1)} = \lambda_j^{(k)} + c w_j^{(k+1)}, \quad j = 1, \dots, r \quad (3.5)$$

$$w^{(k)} = \frac{1}{m}(Ax^{(k)} - b), \quad A = [I \ I \ \dots \ I] \in \mathbb{R}^{r \times rm} \quad (3.6)$$

となる . ただし , I は r 次元の単位行列とする .

4 数値実験

本節では , 前節で示した交互方向乗数法 (3.2), (3.4)–(3.6) を (3.1) の形の二次錐計画問題に適用した .

まず , 以下の二次錐計画問題について数値計算を行った結果を述べる .

$$\begin{aligned} & \text{minimize} && \sum_{i=1}^m \left(\frac{1}{2} \alpha_i \|x_i\|^2 + \gamma_i^\top x_i \right) \\ & \text{subject to} && \sum_{i=1}^m x_i = b \\ & && x_i \in K_i, \quad i = 1, \dots, m \end{aligned} \quad (4.1)$$

ただし , γ_i, b は r 次元ベクトル , K_i は r 次元の二次錐とする . γ_i の各成分と α_i は $(0, 1)$ の一様乱数となるように選ぶ . 問題が実行可能になるように , r 次元の二次錐に含まれる点を m 個生成し , その和を b とした . ここで , r 次元の二次錐 $K_i = \{x_i \in \mathbb{R}^r \mid x_{i1} \geq \|\bar{x}_i\|\}$ に含まれる点を生成するため , まず各成分が $(0, 1)$ の一様乱数からなる $r-1$ 次元ベクトル \bar{x}_i を定め , $x_{i1} = 2\|\bar{x}_i\|$ とした .

数値実験では , 終了条件として

$$\left\| \sum_{i=1}^m x_i - b \right\|_\infty \leq 10^{-5} \quad (4.2)$$

を用いた . ただし , $\|\cdot\|_\infty$ は l_∞ -ノルムを表す .

また , 求めた解の精度を確かめるために , 問題 (4.1) の KKT(Karush-Kuhn-Tucker) 条件を考える . この問題の KKT 条件は以下の形に書ける .

$$x - \text{Proj}_K[x - (Mx + \gamma + A^\top \lambda)] = 0 \quad (4.3)$$

$$Ax - b = 0 \quad (4.4)$$

ただし, $K = K_1 \times \dots \times K_m$ であり, r 次の単位行列を I とすると, M は $\alpha_i I (i = 1, \dots, n)$ をブロック対角に並べた行列で, A は I を m 個横に並べた行列である. λ は等式制約に対するラグランジュ乗数を表す. また, $\text{Proj}_K[\cdot]$ は二次錐 K への射影を表す関数とする. 求めた解と真の解とのずれを測るために, 終了条件 (4.2) を満たして計算が終了した時点で得られている解に対して, 次の e_1, e_2 の値を求める.

$$e_1 = \|x - \text{Proj}_K[x - (Mx + \gamma + A^T \lambda)]\|_\infty$$

$$e_2 = \|Ax - b\|_\infty$$

なお, 今回の実験では, 3.2GHz の CPU と 4GB のメモリを持つ計算機上でを行い, アルゴリズムは MATLAB 7 を用いて実装した.

4.1 パラメータ c の影響

まず, パラメータ c の値が, 交互方向乗数法の性能に与える影響を調べる. 実験において, $r = 10$, $m = 10$ の問題を 4 つランダムに生成した. 表 1 は, パラメータ c の値に対して終了条件 (4.2) を満たすまでに要した反復回数を表す. この表から, パラメータ c の値が, アルゴリズムの収束の速さに大きな影響を与えることがわかる. 特に, パラメータ c を, 適当な値に設定することが重要であり, その値よりも小さくても大きくても収束が遅くなる.

表 1: パラメータ c と反復回数の関係

	$c = 0.01$	$c = 0.1$	$c = 1$	$c = 10$
Problem 1	575	59	137	726
Problem 2	632	66	55	255
Problem 3	693	75	62	321
Problem 4	646	64	156	667

次の図 1-4 は, 4 つのそれぞれの問題に対して, 反復回数と $e_1 + e_2$ の値の関係を各パラメータの値ごとに図示したものである.

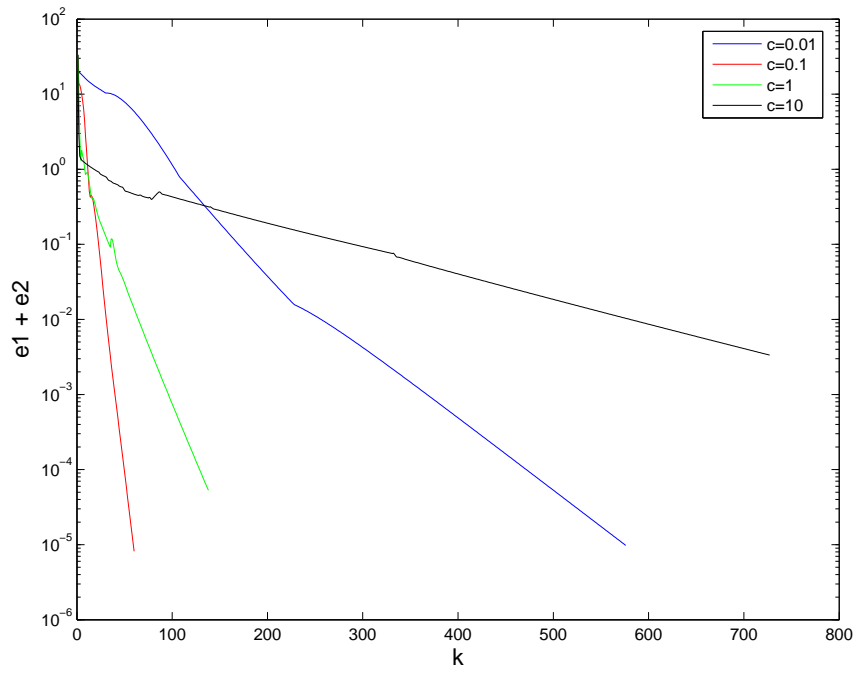


図 1: Problem 1 に対する反復回数と $e_1 + e_2$ の値

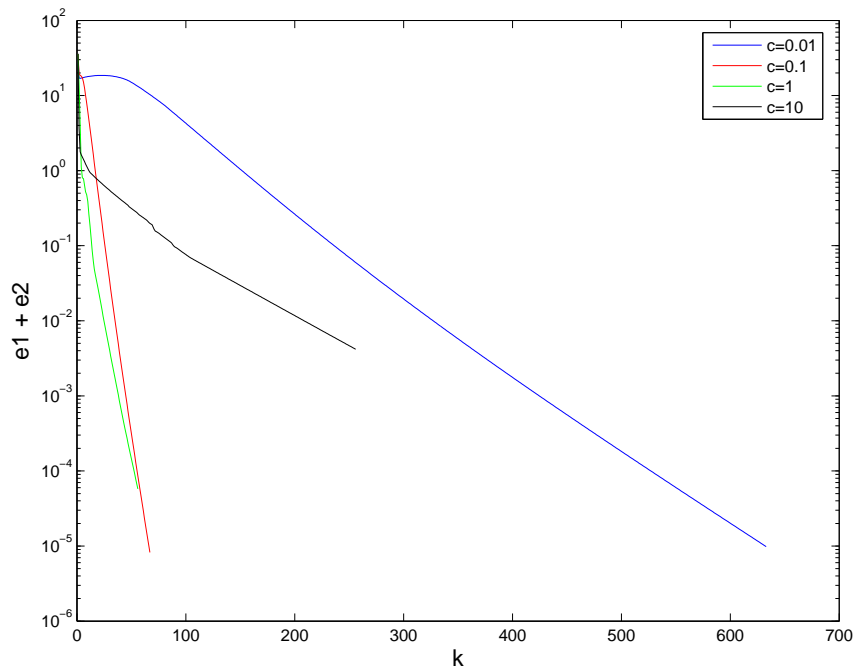


図 2: Problem 2 に対する反復回数と $e_1 + e_2$ の値

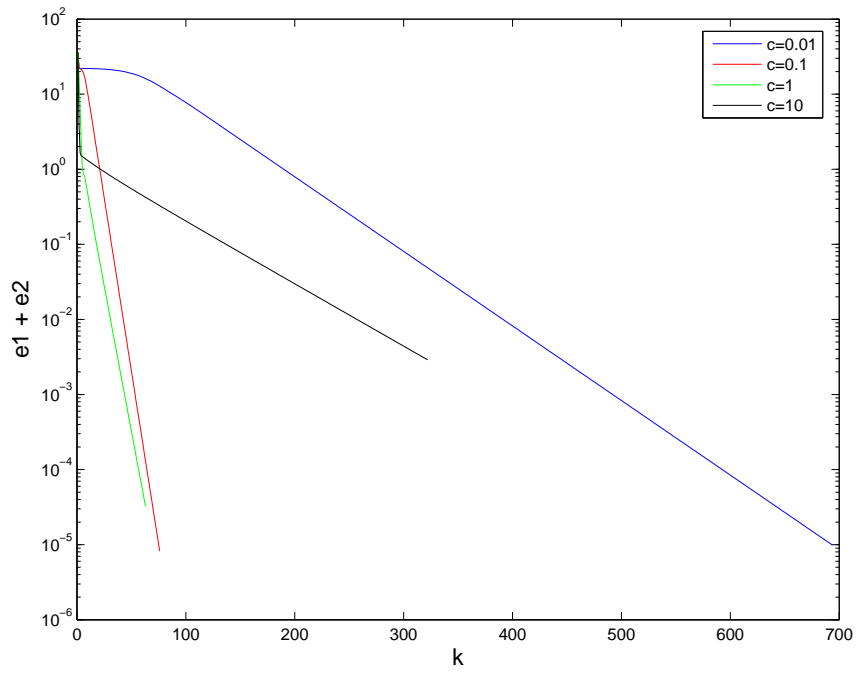


図 3: Problem 3 に対する反復回数と $e_1 + e_2$ の値

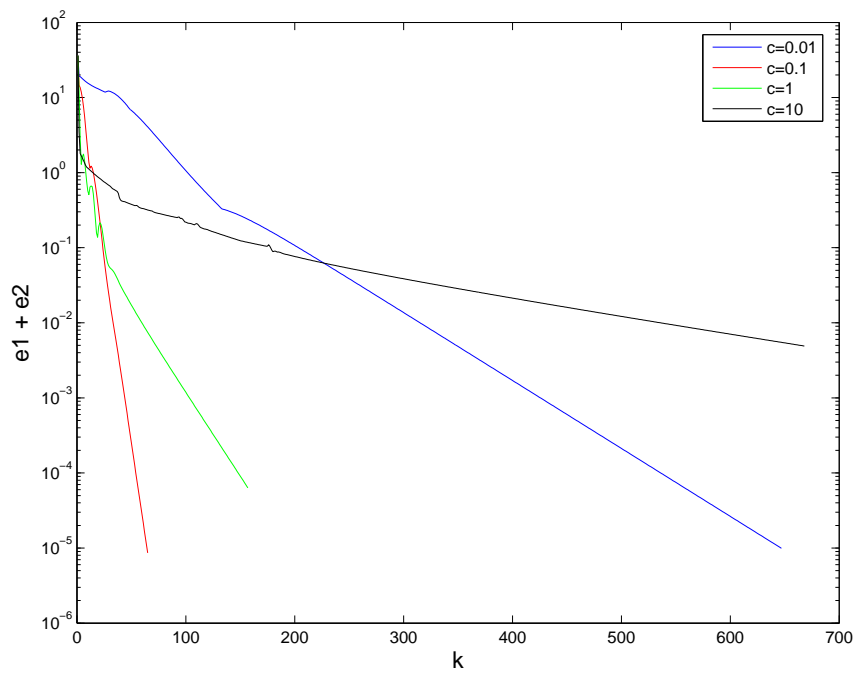


図 4: Problem 4 に対する反復回数と $e_1 + e_2$ の値

4.2 問題の大きさの影響

問題のサイズが交互方向乗数法に与える影響を調べる。いま問題の大きさを決める要素は二つある。それは、ブロックの数 m と各部分ベクトルの次元 r である。これらの影響を調べるため次の実験を行った。今回の実験では並列計算機を用いずに一台のマシんで計算した。

まず、 m を 10 に固定して r の値を変えることにより、各部分ベクトルの次元 r が交互方向乗数法の反復回数にどのような影響を与えるのか調べた。図 5 は、それぞれの r の値について問題を 10 個生成し、各問題に対して終了条件 (4.2) を満たすまでに要した反復回数とその平均回数を表した図である。パラメータ c は 0.3 とした。

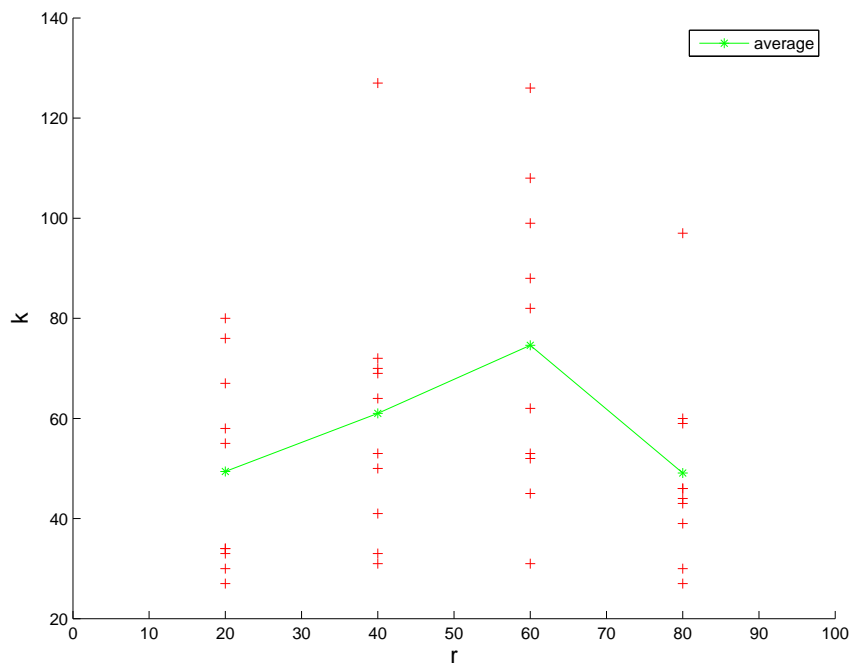


図 5: ブロック数 m を固定し、部分ベクトルの次元 r を変化させたときの反復回数

次に、 r を 10 に固定して m の値を変えることにより、ブロックの数 m が交互方向乗数法の反復回数にどのような影響を与えるのか調べた。図 6 は、それぞれの m の値について問題を 10 個生成し、各問題に対して終了条件 (4.2) を満たすまでに要した反復回数とその平均回数を表した図である。パラメータ c は 0.1 とした。

図 5 より、部分ベクトルの次元 r は反復回数にあまり影響がないことがわかる。一方、図 6 より、ブロック数 m を大きくすると反復回数が増えることがわかる。

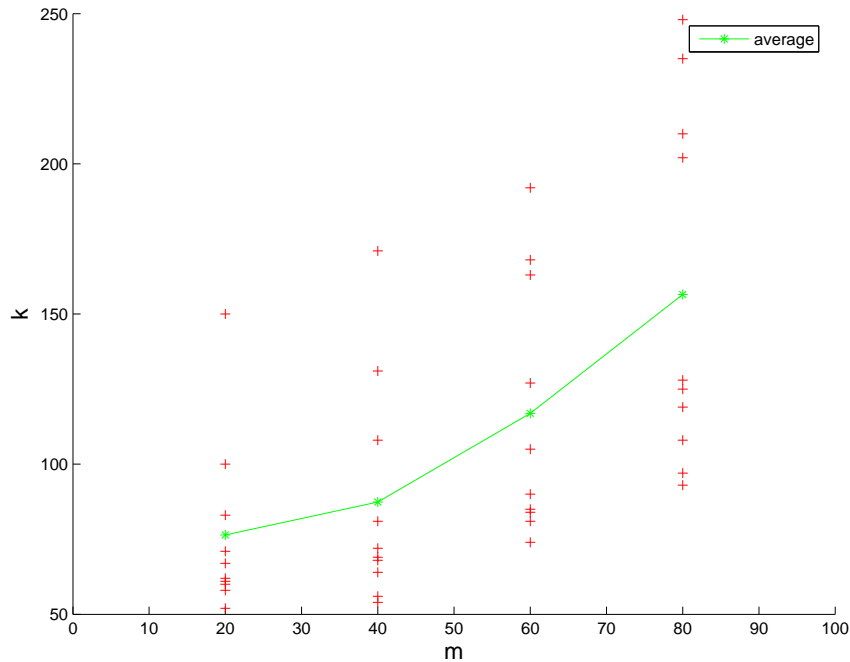


図 6: 部分ベクトルの次元 r を固定し, ブロック数 m を変化させたときの反復回数

4.3 既存のソルバーとの比較

4.3.1 目的関数が二次関数の場合

交互方向乗数法 (ADMM) を用いて問題を解いた場合と, 内点法に基づく既存のソルバー (SDPT3) [6] を用いて問題を解いた場合に対して, それぞれの問題のサイズで計算時間を比較する.

ここで, 実際には並列計算機を用いて実験を行っていないが, データの転送に時間がかからない仮想的な並列計算を行う場合を想定する. このとき, 交互方向乗数法の各反復で (3.4) 式は, m 台のマシンで並列に計算できる. 各反復に要する時間は, 並列化を行わないと, $i = 1, \dots, m$ に対する計算時間の和となる. しかし, m 台のマシンで並列化した場合, 各反復において $i = 1, \dots, m$ のなかの計算時間の最大値が, その反復における計算時間となる. この各反復における計算時間の最大値の和を, 仮想的に並列化できた場合の計算時間とみなし, これとの比較も行う.

まず, m の値を固定して, r の値を変えて問題を 10 個生成し実験した. 表 2 は $m = 10$, 表 3 は $m = 50$ とした場合の平均の計算時間, 反復回数, 得られた解に対する e_1, e_2 の値を表す. 交互方向乗数法のパラメータ c は $m = 10$ のとき 0.3, $m = 50$ のとき 0.2 とした.

表 2: 平均計算時間, 平均反復回数および解の精度 ($m = 10$)

	$r = 10$	$r = 50$	$r = 100$
SDPT3(秒)	0.3508	1.4631	0.7630
ADMM(並列化なし)(秒)	0.0771	0.0814	0.1011
ADMM(仮想的並列化)(秒)	0.0141	0.0149	0.0187
ADMM の反復回数	55.4	49.0	55.4
SDPT3 の e_1	6.1137e-06	1.6037e-04	2.6784e-04
SDPT3 の e_2	1.1753e-10	1.5683e-10	6.0023e-06
ADMM の e_1	3.3348e-06	3.0724e-06	4.9179e-06
ADMM の e_2	6.9902e-06	8.2848e-06	6.2840e-06

表 3: 平均計算時間, 平均反復回数および解の精度 ($m = 50$)

	$r = 10$	$r = 50$	$r = 100$
SDPT3(秒)	0.9945	9.7702	39.1483
ADMM(並列化なし)(秒)	1.0355	0.8938	1.4267
ADMM(仮想的並列化)(秒)	0.1083	0.0913	0.1466
ADMM の反復回数	167.5	120.0	174.1
SDPT3 の e_1	1.4317e-05	1.9576e-04	3.2797e-04
SDPT3 の e_2	9.5472e-12	1.0458e-06	1.5530e-06
ADMM の e_1	1.0570e-06	1.8477e-06	1.3809e-06
ADMM の e_2	9.6338e-06	9.4699e-06	9.6144e-06

表 2 において, $r = 50$ から $r = 100$ と変えたときに SDPT3 による計算時間が減少している. これは, 大規模な問題になると, SDPT3 が行列のスパース性を活用するためである. このとき, e_2 の値は増加していることから解の精度は悪くなる.

次に, r の値を固定して, いくつかの m の値に対して問題を 10 個生成し実験を行った. 表 4 は $r = 10$, 表 5 は $r = 50$ とした場合の平均の計算時間, 反復回数, e_1, e_2 の値を表す. パラメータ c はともに 0.1 とした.

表 4: 平均計算時間, 平均反復回数および解の精度 ($r = 10$)

	$m = 10$	$m = 50$	$m = 100$
SDPT3(秒)	0.3372	1.0075	2.2541
ADMM(並列化なし)(秒)	0.0971	0.7051	1.7608
ADMM(仮想的並列化)(秒)	0.0177	0.0736	0.1628
ADMM の反復回数	69.1	114.0	144.2
SDPT3 の e_1	1.0565e-05	1.0088e-05	2.7655e-05
SDPT3 の e_2	9.0652e-12	1.6011e-11	7.5770e-10
ADMM の e_1	5.7217e-07	2.8345e-07	3.2715e-07
ADMM の e_2	8.2670e-06	8.7527e-06	9.0861e-06

表 5: 平均計算時間, 平均反復回数および解の精度 ($r = 50$)

	$m = 10$	$m = 50$	$m = 100$
SDPT3(秒)	1.4873	10.0850	25.7969
ADMM(並列化なし)(秒)	0.1298	0.7981	2.6950
ADMM(仮想的並列化)(秒)	0.0235	0.0807	0.2392
ADMM の反復回数	78.3	106.8	182.3
SDPT3 の e_1	1.5791e-04	9.4953e-05	3.1167e-04
SDPT3 の e_2	3.0222e-10	3.7198e-07	4.8912e-06
ADMM の e_1	6.5450e-07	5.4277e-07	3.1261e-07
ADMM の e_2	7.4626e-06	8.0816e-06	8.6436e-06

これらの結果から, 4.2 節で確かめたように, 交互方向乗数法は各部分ベクトルの次元 r の影響は小さい一方で, ブロックの数 m の影響が大きいことがわかる. 交互方向乗数法で求めた解の精度を SDPT3 で求めた解の精度と比較すると, e_1 は交互方向乗数法の方が小さいので, KKT 条件の (4.3) 式をより良い精度で満たしていることがわかる. また, e_2 に関しては, SDPT3 の方が小さく, KKT 条件 (4.4) 式については SDPT3 の方が良い解を求めている. これは, SDPT3 が内点法に基づいた方法で, 各反復で制約を満たす実行可能解を求めているので, 制約をどれほど破るかの指標となる e_2 の値は小さくなる. 一方, 交互方向乗数法により生成された点列は収束するにつれ制約を満たす性質があるため, e_2 の値が SDPT3 よりも大きい.

なお, 問題 (4.1) に関する実験では, 目的関数が二次関数であるため, SDPT3 を用いて解く際には, 目的関数が一次関数となるよう, 次の問題に変換して解く必要がある.

$$\begin{aligned}
 & \text{minimize} && \gamma^\top x + u \\
 & \text{subject to} && Ax = b, z = \begin{bmatrix} u + 1 \\ u - 1 \\ \sqrt{2}M^{\frac{1}{2}}x \end{bmatrix} \\
 & && x \in K, z \in K^{r_{m+2}}, u \geq 0
 \end{aligned}$$

ここで γ は γ_i を縦に並べたベクトルで， K^{rm+2} は $rm + 2$ 次元の二次錐を表す．もともと変数が x だけの rm 次元の問題が，この変形により変数が (x, z, u) に増え， $2rm + 3$ 次元の問題になる．そのため， r や m の値が大きくなると，問題の規模が大きくなるので，SDPT3 では計算時間がかかり，さらに次元を大きくするとメモリー不足のため，計算ができなくなる．一方，交互方向乗数法では， rm 次元の問題を， m 個の r 次元の部分問題に分解して解くので，計算に要するメモリーが少なく済み，SDPT3 では扱えないような大規模な問題も解くことができる．特に，問題 (4.1) に対する交互方向乗数法は，各部分問題の解が陽に求まるので計算時間が短くなる．

4.3.2 目的関数が一次関数の場合

つぎに，SDPT3 を適用する際に問題を変換する必要のない，目的関数が一次関数の場合を考える．

$$\begin{aligned} & \text{minimize} && \sum_{i=1}^m \gamma_i^\top x_i \\ & \text{subject to} && \sum_{i=1}^m x_i = b \\ & && x_i \in K_i, \quad i = 1, \dots, m \end{aligned} \quad (4.5)$$

この問題を SDPT3 で解いた場合と交互方向乗数法で解いた場合を，問題のサイズを変えて比較した．ここで，問題 (4.5) の KKT 条件は，

$$x - \text{Proj}_K[x - (\gamma + A^\top \lambda)] = 0 \quad (4.6)$$

$$Ax - b = 0 \quad (4.7)$$

と書けるので，求めた解の精度を測る指標として，次の e_3, e_4 を求める．

$$e_3 = \|x - \text{Proj}_K[x - (\gamma + A^\top \lambda)]\|_\infty$$

$$e_4 = \|Ax - b\|_\infty$$

まず m の値を固定して， r の値を変えて問題を 10 個生成し実験した．

表 6 は $m = 10$ ，表 7 は $m = 100$ とした場合の平均の計算時間，反復回数，得られた解に対する e_3, e_4 の値を表す．また，交互方向乗数法のパラメータ c は $m = 10$ のとき 0.1， $m = 100$ のとき 0.025 とした．

表 6: 平均計算時間，平均反復回数および解の精度 ($m = 10$)

	$r = 10$	$r = 100$	$r = 500$	$r = 1000$	$r = 3000$
SDPT3(秒)	0.1053	0.3282	4.5122	1.6198	9.4800
ADMM(並列化なし)(秒)	0.3201	0.2032	0.2524	0.3230	0.6349
ADMM(仮想的並列化)(秒)	0.0577	0.0375	0.0494	0.0656	0.1456
ADMM の反復回数	230.0	112.0	91.8	82.4	71.5
SDPT3 の e_3	6.3439e-06	6.0842e-06	2.1678e-05	4.1375e-05	5.4895e-05
SDPT3 の e_4	2.9577e-11	4.5464e-09	1.0378e-08	7.8038e-04	1.1607e-03
ADMM の e_3	6.6300e-07	8.7621e-07	1.4337e-06	1.8534e-06	2.1448e-06
ADMM の e_4	9.2056e-06	9.5731e-06	9.1432e-06	9.2726e-06	9.3678e-06

表 7: 平均計算時間, 平均反復回数および解の精度 ($m = 100$)

	$r = 10$	$r = 100$	$r = 500$	$r = 1000$
SDPT3(秒)	0.1621	0.9449	12.3050	23.7330
ADMM(並列化なし)(秒)	17.1590	7.2679	8.6154	12.6092
ADMM(仮想的並列化)(秒)	1.5881	0.6582	0.8235	1.2937
ADMM の反復回数	1413.5	450.1	341.0	341.9
SDPT3 の e_3	5.1953e-06	3.4421e-05	6.1539e-05	6.0333e-04
SDPT3 の e_4	1.4493e-09	3.6586e-08	2.4092e-08	9.0408e-08
ADMM の e_3	6.4195e-08	4.2444e-08	5.1732e-08	5.7375e-08
ADMM の e_4	9.6809e-06	9.6924e-06	9.7469e-06	9.7948e-06

表 2 の場合と同様, 表 6 で $r = 500$ から $r = 1000$ と変えたときに, SDPT3 の計算時間が減少している. これは, 大規模な問題になると, SDPT3 が行列のスパース性を活用するためである. このとき, e_4 の値が増加していることから解の精度は悪くなっていることがわかる.

次に, r の値を固定して, m の値を変えて問題を 10 個生成し実験を行った. 表 8 は $r = 10$, 表 9 は $r = 1000$ とした実験において使用した交互方向乗数法のパラメータの値, 平均の計算時間, 反復回数, 得られた解に対する e_3, e_4 の値を表す.

表 8: 平均計算時間, 平均反復回数および解の精度 ($r = 10$)

	$m = 10$	$m = 50$	$m = 100$
パラメータ c	0.1	0.03	0.025
SDPT3(秒)	0.1125	0.1326	0.1649
ADMM(並列化なし)(秒)	0.1838	3.2223	13.3421
ADMM(仮想的並列化)(秒)	0.0332	0.3351	1.2320
ADMM の反復回数	131.9	518.9	1086.4
SDPT3 の e_3	3.2301e-06	4.7173e-06	4.6733e-06
SDPT3 の e_4	5.9911e-11	4.7351e-10	2.5611e-09
ADMM の e_3	4.9942e-07	1.0393e-07	6.7318e-08
ADMM の e_4	8.5981e-06	7.9533e-06	9.1107e-06

表 9: 平均計算時間, 平均反復回数および解の精度 ($r = 1000$)

	$m = 10$	$m = 50$	$m = 100$
パラメータ c	0.1	0.03	0.025
SDPT3(秒)	1.5867	13.0744	25.8614
ADMM(並列化なし)(秒)	0.3029	4.9393	12.8094
ADMM(仮想的並列化)(秒)	0.0616	0.6206	1.5291
ADMM の反復回数	77.4	262.5	340.3
SDPT3 の e_3	3.7115e-05	5.6419e-04	6.7354e-04
SDPT3 の e_4	1.0259e-03	5.8998e-08	9.8704e-08
ADMM の e_3	1.6022e-06	1.2616e-07	6.9034e-08
ADMM の e_4	9.3277e-06	9.7501e-06	9.7313e-06

目的関数が一次関数の場合は, SDPT3 でも交互方向乗数法と同程度の大きさの問題が解ける. 目的関数が二次関数のときと同様, 交互方向乗数法は各部分ベクトルの次元 r の影響は小さいが, ブロックの数 m の影響を大きく受ける. このことから, r が小さく m が大きいときには SDPT3 で解く方が効率が良い. 逆に, r が相対的に大きいときには問題 (3.1) に対して交互方向乗数法の方が速く解を計算できる. また, m の値が大きいときでも, 交互方向乗数法は m 台のマシンで並列計算可能なので, 並列計算機を利用すれば, 計算時間を短縮することができる. 計算される解の精度については, e_3 は SDPT3 の方が小さく, e_4 は交互方向乗数法の方が小さい. このことから KKT 条件 (4.6) 式に関しては交互方向乗数法の方が, KKT 条件 (4.7) 式については SDPT3 の方が精度の良い解を計算している.

5 結論

本報告書では, 閉凸集合を制約に持つ最適化問題に対し, 適当な仮定をおくことで, 交互方向乗数法によって生成された点列が最適解に収束することを証明した. さらに交互方向乗数法をある種の二次錐計画問題に適用したとき, このアルゴリズムの中に現れる部分問題の解が陽に求まることを示した. この結果を用いて数値実験を行い, 既存のソルバーと比較することで交互方向乗数法の有効性を検証した.

謝辞

日頃から御教授下さり, 本報告書の作成にあたっては細部に至るまで様々な御指摘と適切な御指導を賜った福嶋雅夫教授に深く感謝の意を表します. また, 日頃からお世話になっている山下信雄准教授, 林俊介助教, 福嶋研究室の皆様にも厚く御礼申し上げます.

参考文献

- [1] F. ALIZADEH AND D. GOLDFARB, *Second-order cone programming*, *Mathematical Programming*, **95** (2003), pp. 3–51.

- [2] D. P. BERTSEKAS AND J. N. TSITSIKLIS: *Parallel and Distributed Computation: Numerical Methods*, Prentice Hall, 1989.
- [3] M. FUKUSHIMA, *Application of the alternating direction method of multipliers to separable convex programming problems*, Computational Optimization and Applications, **1** (2001), pp. 93–111.
- [4] 福島 雅夫, 非線形最適化の基礎, 朝倉書店, 2001.
- [5] M. FUKUSHIMA, Z.-Q. LUO AND P. TSENG, *Smoothing functions for second-order-cone complementarity problems*, SIAM Journal on Optimization, **12** (2002), pp. 436–460.
- [6] K. C. TOH, R. H. TÜTÜNCÜ AND M. J. TODD, *SDPT3 version 3.02 - a matlab software for semidefinite-quadratic-linear programming*, 2002, <http://www.math.nus.edu.sg/~mattohkc/sdpt3.html>.
- [7] 矢部 博, 工学基礎 最適化とその応用, 数理工学社, 2006.
- [8] 山下 浩 今野 浩, 非線形計画法, 日科技連出版社, 1978.

A ラグランジュ乗数定理

次の最適化問題を考える．

$$\begin{aligned} & \text{minimize} && F(x) \\ & \text{subject to} && a_i^\top x = b_i, \quad i = 1, \dots, m \\ & && c_j^\top x \leq d_j, \quad j = 1, \dots, r \\ & && x \in K \end{aligned} \tag{A.1}$$

ここで, $F: \mathbb{R}^n \rightarrow \mathbb{R}$ は凸関数, a_i と c_j は \mathbb{R}^n 上のベクトルで, b_i と d_j はスカラーである. K は空でない閉凸集合とする.

この問題のラグランジュ関数は以下で与えられる.

$$L(x, p, u) = F(x) + \sum_{i=1}^m p_i (a_i^\top x - b_i) + \sum_{j=1}^r u_j (c_j^\top x - d_j) \tag{A.2}$$

ここで, p と u はそれぞれ (p_1, \dots, p_m) と (u_1, \dots, u_r) で与えられるベクトルである.

ラグランジュ乗数定理

$$\{x \in \mathbb{R}^n \mid a_i^\top x = b_i, i = 1, \dots, m\} \cap \{x \in \mathbb{R}^n \mid c_j^\top x \leq d_j, j = 1, \dots, r\} \cap \text{int } K \neq \emptyset \tag{A.3}$$

とする (Slater 制約想定). このとき, ベクトル x^* が問題 (A.1) の最適解であることの必要十分条件は, x^* が実行可能で, 次式を満たすベクトル $p^* = (p_1^*, \dots, p_m^*)$ と $u^* \geq 0$ となる $u^* = (u_1^*, \dots, u_r^*)$ が存在することである.

$$F(x^*) = L(x^*, p^*, u^*) = \min_{x \in K} L(x, p^*, u^*) \tag{A.4}$$

$$u_j^* = 0, \quad \forall j \text{ such that } c_j^\top x^* < d_j \tag{A.5}$$

証明 まず十分性を証明する. x^* は実行可能とする. 式 (A.4), (A.5) をみたま $u^* \geq 0$ と p^* が存在するとき, 問題 (A.1) のすべての実行可能解 x に対して,

$$F(x^*) \leq L(x, p^*, u^*) = F(x) + \sum_{i=1}^m p_i^* (a_i^\top x - b_i) + \sum_{j=1}^r u_j^* (c_j^\top x - d_j) \leq F(x) \tag{A.6}$$

となるので, x^* は問題 (A.1) の最適解である.

つぎに必要な性を証明する. x^* は最適解とする. いま, 制約想定 (A.3) が成立することより, 次式を満足するラグランジュ乗数 $u^* \geq 0$ と p^* が存在する.

$$\begin{aligned} 0 & \in \partial F(x^*) + \sum_{i=1}^m p_i^* a_i + \sum_{j=1}^r u_j^* c_j + N_K(x^*) \\ u_j^* & \geq 0, \quad c_j^\top x^* \leq d_j, \quad u_j^* (c_j^\top x^* - d_j) = 0 \quad (j = 1, \dots, r) \\ a_i^\top x^* & = b_i, \quad (i = 1, \dots, m) \end{aligned} \tag{A.7}$$

ただし, $\partial F(x^*)$ は凸関数 F の x^* における劣微分, $N_K(x^*)$ は集合 K の x^* における法線錐である. これより (A.5) が成立. ここで次の問題を考える.

$$\begin{aligned} & \text{minimize} && L(x, p^*, u^*) \\ & \text{subject to} && x \in K \end{aligned}$$

この問題は凸計画問題なので，KKT 条件をみたす点が最適解となる．式 (A.7) より， x^* がこの問題の最適解である．ゆえに，次式が成立する．

$$L(x^*, p^*, u^*) = \min_{x \in K} L(x, p^*, u^*)$$

x^* は問題 (A.1) の最適解であることと，(A.2)，(A.5) より，

$$F(x^*) = L(x^*, p^*, u^*)$$

となり式 (A.4) が成立する． ■

B 分離可能な凸計画問題に対する交互方向乗数法の導出

次のようにおくことで，分離可能な凸計画問題 (1.1) に対する交互方向乗数法 (2.3), (2.4), (2.5) が式 (2.39), (2.40), (2.41) で表せることを示す．

$$G_1(x) = \sum_{i=1}^m F_i(x_i), \quad C_1 = K_1 \times K_2 \times \dots \times K_m,$$

$$G_2(z) = 0, \quad C_2 = \left\{ z \mid \sum_{i=1}^m z_{ji} = b_j, \quad j = 1, \dots, r \right\}$$

ここで， B は x から $a_{ji}^\top x_i$, $j = 1, \dots, r$, $i = 1, \dots, m$ を成分に持つベクトルへの写像を表す行列である．ただし， $a_{ji} \in \mathbb{R}^{n_i}$ は， a_j の部分ベクトルである．また， $x \in \mathbb{R}^n$, $Bx \in \mathbb{R}^m$, $B \in \mathbb{R}^{m \times n}$, $z, p \in \mathbb{R}^m$ である．式 (2.3) より

$$x^{(k+1)} = \operatorname{argmin}_{x \in K_1 \times \dots \times K_m} \left\{ \sum_{i=1}^m F_i(x_i) + \sum_{j=1}^r \sum_{i=1}^m p_{ji}^{(k)} a_{ji}^\top x_i + \frac{c}{2} \sum_{j=1}^r \sum_{i=1}^m (a_{ji}^\top x_i - z_{ji}^{(k)})^2 \right\}$$

となり，各 i について分解できて，次式が成立する．

$$x_i^{(k+1)} = \operatorname{argmin}_{x_i \in K_i} \left\{ F_i(x_i) + \sum_{j=1}^r \left\{ p_{ji}^{(k)} a_{ji}^\top x_i + \frac{c}{2} (a_{ji}^\top x_i - z_{ji}^{(k)})^2 \right\} \right\}, \quad i = 1, \dots, m \quad (\text{B.1})$$

式 (2.4) より

$$z^{(k+1)} = \operatorname{argmin}_{z \in \{z \mid \sum_{i=1}^m z_{ji} = b_j, \quad j=1, \dots, r\}} \left\{ \sum_{j=1}^r \sum_{i=1}^m (-p_{ji}^{(k)} z_{ji}) + \frac{c}{2} \sum_{j=1}^r \sum_{i=1}^m (a_{ji}^\top x_i^{(k+1)} - z_{ji})^2 \right\}$$

となり，各 j について分解できて，次式が成り立つ．

$$z_j^{(k+1)} = \operatorname{argmin}_{z_j \in \{z_j \mid \sum_{i=1}^m z_{ji} = b_j\}} \left\{ \sum_{i=1}^m \left((-p_{ji}^{(k)} z_{ji}) + \frac{c}{2} (a_{ji}^\top x_i^{(k+1)} - z_{ji})^2 \right) \right\}, \quad j = 1, \dots, r \quad (\text{B.2})$$

また，式 (2.5) より，

$$p_{ji}^{(k+1)} = p_{ji}^{(k)} + c(a_{ji}^\top x_i^{(k+1)} - z_{ji}^{(k+1)}), \quad i = 1, \dots, m, \quad j = 1, \dots, r \quad (\text{B.3})$$

が成立する．

ここで、式 (B.2) は凸計画問題なので、KKT 条件から $z_{ji}^{(k+1)}$ を求めることができる。等式制約 $\sum_{i=1}^m z_{ji} = b_j$ に対するラグランジュ乗数を $\lambda_j^{(k+1)}$ とすると、任意の i, j に対して、

$$z_{ji}^{(k+1)} = a_{ji}^\top x_i^{(k+1)} + \frac{1}{c}(p_{ji}^{(k)} - \lambda_j^{(k+1)}) \quad (\text{B.4})$$

となる。式 (B.4) と (B.3) を比較することで、任意の i, j に対して、次式が成立する。

$$\lambda_j^{(k+1)} = p_{ij}^{(k+1)}$$

この式より、式 (B.3) は、任意の i, j に対して、

$$\lambda_j^{(k+1)} = \lambda_j^{(k)} + c(a_{ji}^\top x_i^{(k+1)} - z_{ji}^{(k+1)})$$

と書け、 i について足し合わせて m で割ると、任意の $j = 1, \dots, r$ に対して、

$$\begin{aligned} \lambda_j^{(k+1)} &= \lambda_j^{(k)} + \frac{c}{m} \left(\sum_{i=1}^m a_{ji}^\top x_i^{(k+1)} - \sum_{i=1}^m z_{ji}^{(k+1)} \right) \\ &= \lambda_j^{(k)} + \frac{c}{m} (a_j^\top x^{(k+1)} - b_j) \end{aligned} \quad (\text{B.5})$$

となる。式 (B.4) の $p_{ij}^{(k)}$ を $\lambda_j^{(k)}$ とすることで、任意の i, j に対して、

$$z_{ji}^{(k+1)} = a_{ji}^\top x_i^{(k+1)} + \frac{1}{c}(\lambda_j^{(k)} - \lambda_j^{(k+1)})$$

が成立する。この式と、式 (B.5) より、 $\lambda_j^{(k)}$, $\lambda_j^{(k+1)}$ を消去すると、

$$z_{ji}^{(k+1)} = a_{ji}^\top x_i^{(k+1)} - \frac{1}{m}(a_j^\top x^{(k+1)} - b_j)$$

を得る。つまり、

$$z_{ji}^{(k)} = a_{ji}^\top x_i^{(k)} - \frac{1}{m}(a_j^\top x^{(k)} - b_j) \quad (\text{B.6})$$

と表せ、 $z_{ji}^{(k)}$ を x によって表すことができた。これを式 (B.1) に代入すると、

$$x_i^{(k+1)} = \operatorname{argmin}_{x_i \in K_i} \left\{ F_i(x_i) + \sum_{j=1}^r \left\{ p_{ji}^{(k)} a_{ji}^\top x_i + \frac{c}{2} \left(a_{ji}^\top (x_i - x_i^{(k)}) + \frac{1}{m}(a_j^\top x^{(k)} - b_j) \right)^2 \right\} \right\}, \quad (\text{B.7})$$

$i = 1, \dots, m$

となる。ここで

$$w_j^{(k)} = \frac{1}{m}(a_j^\top x^{(k)} - b_j), \quad j = 1, \dots, r \quad (\text{B.8})$$

とおくと、式 (B.5) は

$$\lambda_j^{(k+1)} = \lambda_j^{(k)} + c w_j^{(k+1)}, \quad j = 1, \dots, r \quad (\text{B.9})$$

となり、式 (B.7) は

$$x_i^{(k+1)} = \operatorname{argmin}_{x_i \in K_i} \left\{ F_i(x_i) + \sum_{j=1}^r \left\{ p_{ji}^{(k)} a_{ji}^\top x_i + \frac{c}{2} \left(a_{ji}^\top (x_i - x_i^{(k)}) + w_j^{(k)} \right)^2 \right\} \right\}, \quad i = 1, \dots, m \quad (\text{B.10})$$

と書ける。以上より、問題 (1.1) に対する交互方向乗数法が導出できた。