数理解析研究所講究録 1297 巻 2002 年 224-233

錐上の多項式制約を持つ最適化問題に対する緩和手法

東京工業大学大学院 情報理工学研究科　脇 隼人 (Hayato Waki)[1]
Department of Mathematical and Computing Sciences, Tokyo Institute of Technology

東京工業大学大学院 情報理工学研究科　小島 政和 (Masakazu Kojima)[2]
Department of Mathematical and Computing Sciences, Tokyo Institute of Technology

Department of Mathematics, Ewha Women's University　Sunyoung Kim[3]

1　はじめに

数理最適化分野には, 実用的に非常に重要であるが, 計算コストが莫大になり, 時として大域的最適解が得られない数理計画問題が数多く存在する. そういう問題に対して, より解きやすい数理計画問題に変形して (同値な変形でなくて良い), 大域的最適値を見積もる手法を, 緩和と呼び, 変形された問題を緩和問題, 緩和問題の最適値, 最適解をそれぞれ, 緩和値, 緩和解と呼ぶ. 緩和問題に望まれる条件として, 緩和値と大域的最適値と緩和値が近く, 簡単に緩和問題を解くことができる, というのがあげられるが, これら両方を満たすのは非常に難しい.

本論文で提案する緩和手法は, 錐上の多項式制約を持つ最適化問題に対する緩和である. この緩和手法は, 一般的であり柔軟性に富んでいて, また, 既存の緩和手法の拡張になっている.

本論文の構成は次の様になっている. 2 節で, 我々の緩和手法を提案する. 3 節では, 我々の緩和手法が, Lasserre[5] の緩和手法の拡張になっていることと, SOCP 緩和について説明する. 4 節では, 我々の緩和手法が, 今までに提案されている緩和理論の拡張になっていることを示す. 最後に 5 節で, 簡単に今後の課題を述べる. また, 本論文は, [3] に基づいているが, 4.2 節では, [4] の 6.2 節の定理 6.4 が [1] の定理 2.1 の拡張になっているということに対して証明を付け加えた.

2　錐上の多項式制約を持つ最適化問題に対する緩和

次の問題を考える:

$$\min g_0(\boldsymbol{x}) \text{ subject to } \boldsymbol{g}(\boldsymbol{x}) = (g_1(\boldsymbol{x}), \dots, g_m(\boldsymbol{x}))^T \in \mathcal{K}. \tag{1}$$

ただし, $g_0(\boldsymbol{x}), g_1(\boldsymbol{x}), \dots, g_m(\boldsymbol{x})$ を n 個の実変数からなる多項式とし, $\mathcal{K}$ を $\mathbb{R}^m$ 上の閉凸錐とする. 以下では, この問題 (1) に対する緩和問題を考える.

説明のために記号を定義する. $\boldsymbol{a} = (a_1, \dots, a_n) \in \mathbb{Z}^n_+$, $\boldsymbol{x} = (x_1, \dots, x_n)$ と書いて, $\boldsymbol{x}^{\boldsymbol{a}}$ を $x_1^{a_1} \cdots x_n^{a_n}$ と定める. また, $\mathcal{A}$ を $\mathbb{Z}^n_+$ の有限部分集合とし, 多項式 $g(\boldsymbol{x})$ の $\boldsymbol{x}^{\boldsymbol{a}}$ における係数を

---

[1] e-mail: waki9@is.titech.ac.jp
[2] e-mail: kojima@is.titech.ac.jp
[3] e-mail: skim@math.ewha.ac.kr

$c(\boldsymbol{a})$ とおく. これにより, 多項式 $g(\boldsymbol{x})$ を,

$$g(\boldsymbol{x}) = \sum_{\boldsymbol{a}\in\mathcal{A}} c(\boldsymbol{a})\boldsymbol{x}^{\boldsymbol{a}}$$

と書くことができる.

問題 (1)の緩和問題を構成するために,線型化を行う.具体的には,$\boldsymbol{x}^{\boldsymbol{a}}$ を $y_{\boldsymbol{a}}$ に置き換える. $g(\boldsymbol{x})$ からできる線形関数を $G((y_{\boldsymbol{a}}\colon \boldsymbol{a}\in\mathcal{A}))$ とおくと,

$$G((y_{\boldsymbol{a}}\colon \boldsymbol{a}\in\mathcal{A})) = \sum_{\boldsymbol{a}\in\mathcal{A}} c(\boldsymbol{a})y_{\boldsymbol{a}}$$

となる. 問題 (1) にこの操作を施すと,

$$\min G_0((y_{\boldsymbol{a}}\colon \boldsymbol{a}\in\mathcal{A})) \text{ subject to } \boldsymbol{G}((y_{\boldsymbol{a}}\colon \boldsymbol{a}\in\mathcal{A})) \in \mathcal{K} \tag{2}$$

となる. ただし, $\boldsymbol{G}((y_{\boldsymbol{a}}\colon \boldsymbol{a}\in\mathcal{A})) = (G_1((y_{\boldsymbol{a}}\colon \boldsymbol{a}\in\mathcal{A})),\ldots,G_m((y_{\boldsymbol{a}}\colon \boldsymbol{a}\in\mathcal{A})))^T$ である. また, $\mathcal{A}$ を全ての多項式 $g_0(\boldsymbol{x}),\ldots,g_m(\boldsymbol{x})$ の各単項式 $\boldsymbol{x}^{\boldsymbol{a}}$ が持つ $\boldsymbol{a}$ の集合とする.

この問題 (2) について考えると, $G_0, \boldsymbol{G}$ は変数を $y_{\boldsymbol{a}}$ とする線形関数であるから, (2) は錐 $\mathcal{K}$ 上の線形計画問題とみなすことができる. したがって, もし, $\mathcal{K}$ が, $\mathbb{R}^m$ の非負領域であれば, 線形計画問題に対する主双対内点法を用いて緩和問題 (2) を解くことができる.

ここでは, 簡単のため, $\mathcal{K}$ を $\mathbb{R}^m$ における閉凸錐としたが, 別に, $\mathbb{R}^m$ 上にとらわれる必要はない. 例えば, 問題 (1) が,

$$\min x_1 \text{ subject to } \begin{pmatrix} 1 & x_1^2+2x_2^2 \\ x_1^2+2x_2^2 & x_1^2x_2 \end{pmatrix} \in \mathcal{S}_+^2, \boldsymbol{x} = (x_1,x_2)^T \in \mathbb{R}^2.$$

という型をしていても, 上で説明したとおりに線型化を行えば,

$$\min y_{20} \text{ subject to } \begin{pmatrix} 1 & y_{20}+2y_{02} \\ y_{20}+2y_{02} & y_{21} \end{pmatrix} \in \mathcal{S}_+^2.$$

となり, この問題は 2 次の半正定値計画問題となる. ここで, $\mathcal{S}_+^2$ とは, 2 次の半正定値行列の作る錐を表している. また, 考えている問題 (1) の錐 $\mathcal{K}$ が錐 $\mathcal{K}_i$ の直積であっても, その各々の錐 $\mathcal{K}_i$ で主双対内点法を適用できるなら, 錐 $\mathcal{K}$ 上の多項式最適化問題に対して上で述べた緩和を施し, 錐上の線型計画問題を構成して緩和値を得ることができる. つまり, 与えられた問題 (1) が主双対内点法が議論できる空間と錐 $\mathcal{K}$ であれば, 上で述べた緩和問題を構成することができる.

3 特徴

この節では, 前節で提案した緩和に対する理論的特徴について述べる.

3.1 Lasserre の SDP 緩和

本論文で提案した緩和方法は, Lasserre の SDP 緩和 [5] の拡張になっている. この項では, そのことについて述べることにする.

まず, Lasserre の SDP 緩和について説明する. Lasserre は, 次の問題に対して SDP 緩和を施すことを提案した：

$$\min g_0(\boldsymbol{x}) \text{ subject to } \boldsymbol{g}(\boldsymbol{x}) \in \mathbb{R}^m_+. \tag{3}$$

これは, 問題 (1) の $\mathcal{K}$ を $\mathbb{R}^m_+$ と置き換えたものである. Lasserre は (3) に (加えても制約領域が変わらないと言う意味で) 冗長な制約を加える. それは,

$$\boldsymbol{u}(\boldsymbol{x}, N_i)\boldsymbol{u}(\boldsymbol{x}, N_i)^T \succeq \boldsymbol{O} \quad (\forall i = 1, \ldots, m, m+1) \tag{4}$$

という制約式である. この制約式は, 各多項式 $g_i(\boldsymbol{x})$ を次数 w_i とし, $\tilde{w}_i = \lceil w_i/2 \rceil$ と定め, $N_{m+1} \geq \lceil m/2 \rceil, N_{m+1} = \max_i \tilde{w}_i$ となるようにおく. さらに, $N_i = N_{m+1} - \tilde{w}_i \ (i = 1 \ldots, m)$ とおく. この様に定めた N_i に対する $\boldsymbol{u}(\boldsymbol{x}, N_i)$ とは, n 個の変数で構成される N_i 次以下の全ての単項式 $\boldsymbol{x}^{\boldsymbol{a}}$ と 1 を並べた列ベクトルである. 具体的に書くと

$$\boldsymbol{u}(\boldsymbol{x}, N_i) = \left(1, x_1, \ldots, x_n, x_1^2, x_1 x_2, \ldots, x_n^2, \ldots, x_1^{N_i}, \ldots, x_n^{N_i}\right)^T$$

である. (4) の対称行列ははランク 1 で, $\boldsymbol{x}$ がどんな値を取っても半正定値対称行列になる. したがって, この制約式 (4) を加えても実行可能領域は変化しない. 問題 (3) は,

$$\min g_0(\boldsymbol{x}) \text{ subject to } \begin{cases} g_i(\boldsymbol{x}) \geq 0, \boldsymbol{u}(\boldsymbol{x}, N_i)\boldsymbol{u}(\boldsymbol{x}, N_i)^T \succeq \boldsymbol{O} \quad (i = 1, \ldots, m), \\ \boldsymbol{u}(\boldsymbol{x}, N_{m+1})\boldsymbol{u}(\boldsymbol{x}, N_{m+1})^T \succeq \boldsymbol{O} \end{cases} \tag{5}$$

という同値な問題になる. 次に, 掛け合わせを行う.

$$\boldsymbol{M}(g_i \boldsymbol{x}) = g_i(\boldsymbol{x})\boldsymbol{u}(\boldsymbol{x}, N_i)\boldsymbol{u}(\boldsymbol{x}, N_i)^T$$

とおくと,

$$\boldsymbol{M}(g_i \boldsymbol{x}) \succeq \boldsymbol{O} \quad \Leftrightarrow \quad g_i(\boldsymbol{x}) \geq 0 \text{ かつ } \boldsymbol{u}(\boldsymbol{x}, N_i)\boldsymbol{u}(\boldsymbol{x}, N_i)^T \succeq \boldsymbol{O}$$

が言えるので, (5) は,

$$\min g_0(\boldsymbol{x}) \text{ subject to } \boldsymbol{M}(\boldsymbol{x}) \succeq \boldsymbol{O},\ \boldsymbol{M}(g_i \boldsymbol{x}) \succeq \boldsymbol{O} \quad (i = 1, \ldots, m) \tag{6}$$

という同値な問題に変形できる. ただし, $\boldsymbol{M}(\boldsymbol{x}) = \boldsymbol{u}(\boldsymbol{x}, N_{m+1})\boldsymbol{u}(\boldsymbol{x}, N_{m+1})^T$ とおく. この問題 (6) に対して, $\boldsymbol{x}^{\boldsymbol{a}}$ を $y_{\boldsymbol{a}}$ と置き換える. これにより, 緩和問題

$$\min g_0(\boldsymbol{x}) \text{ subject to } \boldsymbol{M}(y_{\boldsymbol{a}}) \succeq \boldsymbol{O},\ \boldsymbol{M}(g_i y_{\boldsymbol{a}}) \succeq \boldsymbol{O} \quad (i = 1, \ldots, m) \tag{7}$$

が構成できる.

以上が, Lasserre の提案した SDP 緩和である. Lasserre の SDP 緩和は非常に強力なもので, 弱い仮定の下で, N_{m+1} を大きく取れば, 最適値を得ることができる. 問題 (6) を問題 (1) とみ

なせば我々の緩和は, Lasserre の提案した SDP 緩和の一部拡張になっていることが分かる. ここで「一部」と書いたのは, 我々の緩和方法は, 問題 (6) に対してであり, 問題 (6) の構成方法, つまり, 冗長な制約式や, 制約式同士の掛け合わせについてはまったく言及していない. そこで, 次の項では, 冗長な制約式について述べることにする.

3.2 冗長な制約式

次の 2 つのベクトル値関数 $\boldsymbol{f}_1, \boldsymbol{f}_2 : \mathbb{R}^n \to \mathbb{R}^\ell$ について考える. これらの関数に Cauchy-Schwartz の不等式を適用すると,

$$\left(\boldsymbol{f}_1(\boldsymbol{x})^T \boldsymbol{f}_2(\boldsymbol{x})\right)^2 \leq (\boldsymbol{f}_1(\boldsymbol{x})^T \boldsymbol{f}_1(\boldsymbol{x}))(\boldsymbol{f}_2(\boldsymbol{x})^T \boldsymbol{f}_2(\boldsymbol{x}))$$

が全ての $\boldsymbol{x} \in \mathbb{R}^n$ で成立する. この不等式は,

$$\left(\boldsymbol{f}_1(\boldsymbol{x})^T \boldsymbol{f}_1(\boldsymbol{x}) - \boldsymbol{f}_2(\boldsymbol{x})^T \boldsymbol{f}_2(\boldsymbol{x})\right)^2 + \left(2\boldsymbol{f}_1(\boldsymbol{x})^T \boldsymbol{f}_2(\boldsymbol{x})\right)^2 \leq \left(\boldsymbol{f}_1(\boldsymbol{x})^T \boldsymbol{f}_1(\boldsymbol{x}) + \boldsymbol{f}_2(\boldsymbol{x})^T \boldsymbol{f}_2(\boldsymbol{x})\right)^2$$

と書き換えることができ, これは,

$$\begin{pmatrix} \boldsymbol{f}_1(\boldsymbol{x})^T \boldsymbol{f}_1(\boldsymbol{x}) + \boldsymbol{f}_2(\boldsymbol{x})^T \boldsymbol{f}_2(\boldsymbol{x}) \\ \boldsymbol{f}_1(\boldsymbol{x})^T \boldsymbol{f}_1(\boldsymbol{x}) - \boldsymbol{f}_2(\boldsymbol{x})^T \boldsymbol{f}_2(\boldsymbol{x}) \\ 2\boldsymbol{f}_1(\boldsymbol{x})^T \boldsymbol{f}_2(\boldsymbol{x}) \end{pmatrix} \in \mathcal{N}_3 \tag{8}$$

を意味している. ただし, $\mathcal{N}_3$ は

$$\mathcal{N}_3 = \{(x_0, \boldsymbol{x}_1) \in \mathbb{R} \times \mathbb{R}^2 \mid \|\boldsymbol{x}_1\|_2 \leq x_0\}$$

とし, 3 次の二次錐とする. もちろん (8) は全ての $\boldsymbol{x} \in \mathbb{R}^n$ で成立するので, (8) は冗長な制約式である. この制約式を問題 (1) に加え, 掛け合わせを行い線型化を施せば, 緩和問題として二次錐計画問題を構成することができる. これもまた主双対内点法を適用することができるので, 緩和値を得ることができる.

もう一つ, 二次錐に関する冗長な制約式について述べる. $\boldsymbol{f} : \mathbb{R}^n \to \mathbb{R}^\ell$ とすると, 次のことが言える：

$$\boldsymbol{f}(\boldsymbol{x})\boldsymbol{f}(\boldsymbol{x})^T - \boldsymbol{f}(\boldsymbol{x})\boldsymbol{f}(\boldsymbol{x})^T = \boldsymbol{O} \in \mathcal{S}^\ell_+.$$

これも, 全ての $\boldsymbol{x} \in \mathbb{R}^n$ で成立する. これに対して, ℓ 次半正定値対称行列 $\boldsymbol{C}$ と内積を取る. 半正定値対称行列同士の内積は必ず非負になるので,

$$\boldsymbol{C} \bullet \left(\boldsymbol{f}(\boldsymbol{x})\boldsymbol{f}(\boldsymbol{x})^T - \boldsymbol{f}(\boldsymbol{x})\boldsymbol{f}(\boldsymbol{x})^T\right) \geq 0$$

が成り立つ. ここで, $\boldsymbol{A} \bullet \boldsymbol{B} = \sum_{i=1}^p \sum_{j=1}^p A_{ij} B_{ij} \quad (\boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{p \times p})$ である. 今, $\boldsymbol{C}$ が $\boldsymbol{L} \in \mathbb{R}^{p \times \ell}$ を用いて $\boldsymbol{C} = \boldsymbol{L}^T \boldsymbol{L}$ と記述できたとする (コレスキー分解を用いれば $\boldsymbol{L}$ を求めることができる). すると,

$$\left(\boldsymbol{L}\boldsymbol{f}(\boldsymbol{x})\right)^T \left(\boldsymbol{L}\boldsymbol{f}(\boldsymbol{x})\right) \leq \left(\boldsymbol{f}(\boldsymbol{x})^T \boldsymbol{C} \boldsymbol{f}(\boldsymbol{x})\right)$$

が全ての $\boldsymbol{x} \in \mathbb{R}^n$ で成立する. これも同様に二次錐を用いて

$$\begin{pmatrix} 1 + \boldsymbol{C} \bullet \boldsymbol{f}(\boldsymbol{x})\boldsymbol{f}(\boldsymbol{x})^T \\ 1 - \boldsymbol{C} \bullet \boldsymbol{f}(\boldsymbol{x})\boldsymbol{f}(\boldsymbol{x})^T \\ 2\boldsymbol{L}\boldsymbol{f}(\boldsymbol{x}) \end{pmatrix} \in \mathcal{N}_{p+2}$$

と書くことができる. これを用いても, やはり緩和問題は二次錐計画問題になる.

このようにして, 与えられて問題を二次錐計画問題に緩和して緩和値を得る方法を, SOCP 緩和と呼んでいる. 最後に, このようにして構成される SOCP 緩和と [2] で提案されている SOCP 緩和の違いを述べる. [2] の SOCP 緩和は, 一度, 与えられて問題を SDP 緩和してから生じる $\boldsymbol{X} \succeq \boldsymbol{O}$ という制約に対してさらに緩和することで構成される. 一方, 我々の手法では, 与えられた問題の中に二次錐制約を表現しているものあれば, それと掛け合わせて二次錐計画問題を構成できるし, もしなくても, 上で述べた冗長な制約を加えれば, やはり二次錐計画問題を構成できる.

4 拡張

本論文で提案した緩和手法は, 錐上の多項式制約を持つ最適化問題という非常に幅の広い問題に対して, 冗長な制約式を加えて線型化するという操作によって緩和の枠組みが拡げることができた. この節では, この緩和手法を用いて, 提案されている 2 つの緩和理論を拡張する.

4.1 最適性の十分条件

この節では, 問題 (1) に対する最適性の十分条件を示し, それが, [5] で提案されている問題 (3) に対する最適性の十分条件の拡張になっていることを示す. 以下に, [5] の最適性の十分条件に関する定理をあげておく:

定理 4.1 ([5] の問題 (3) に対する最適性の十分条件)
ζ^* を問題 (3) の最適値とする.

$$g_0(\boldsymbol{x}) - \zeta^* \;=\; q(\boldsymbol{x}) + \sum_{i=1}^{m} g_i(\boldsymbol{x})t_i(\boldsymbol{x}) \quad (\forall \boldsymbol{x} \in \mathbb{R}^n) \tag{9}$$

が成り立つと仮定する. ここで, $q(\boldsymbol{x})$ が次数 $2N_{m+1}$ 以下で, $t_i(\boldsymbol{x})$ が次数 $2N_{m+1} - w_i$ 以下で, 多項式の二乗和で書けるとする. この時, 問題 (3) の最適値 ζ^*, 問題 (10) と問題 (11) の最適値が一致する.

問題 (1) の $g_j(\boldsymbol{x})$ を

$$g_j(\boldsymbol{x}) = \gamma_j + \boldsymbol{d}_j^T\boldsymbol{x} + \sum_{\boldsymbol{a} \in \mathcal{A}^N} c_j(\boldsymbol{a})\boldsymbol{x}^{\boldsymbol{a}} \quad (j = 0, 1, \ldots, m)$$

と書くことにする．ただし，$\mathcal{A}^N$ は，非線型な単項式の指数のベクトル $\boldsymbol{a}$ を集めたもので，具体的に書くと $\mathcal{A}^N = \{\boldsymbol{a} \in \mathcal{A} \mid \sum_{i=1}^n a_i \geq 2\}$ である．また，$\gamma_0 = 0$ とする．これにより，この $g_j(\boldsymbol{x})$ を線型化した $G_j(\boldsymbol{x}, (y_{\boldsymbol{a}}: \boldsymbol{a} \in \mathcal{A}^N))$ は，

$$G_j(\boldsymbol{x}, (y_{\boldsymbol{a}}: \boldsymbol{a} \in \mathcal{A}^N)) = \gamma_j + \boldsymbol{d}_j^T \boldsymbol{x} + \sum_{\boldsymbol{a} \in \mathcal{A}^N} c_j(\boldsymbol{a}) y_{\boldsymbol{a}}$$

とかける．また，同様に $\boldsymbol{G}(\boldsymbol{x}, (y_{\boldsymbol{a}}: \boldsymbol{a} \in \mathcal{A}^N))$ も

$$\boldsymbol{G}(\boldsymbol{x}, (y_{\boldsymbol{a}}: \boldsymbol{a} \in \mathcal{A}^N)) = \left(G_1(\boldsymbol{x}, (y_{\boldsymbol{a}}: \boldsymbol{a} \in \mathcal{A}^N)), \ldots, G_m(\boldsymbol{x}, (y_{\boldsymbol{a}}: \boldsymbol{a} \in \mathcal{A}^N))\right)^T$$

と定める．$G_j(\boldsymbol{x}, (y_{\boldsymbol{a}}: \boldsymbol{a} \in \mathcal{A}^N))$ において，以下の説明のため線型化しても変わらない $\boldsymbol{x}$ はそのままを残した．

問題 (1) は

$$\min g_0(\boldsymbol{x}) \text{ subject to } \boldsymbol{g}(\boldsymbol{x}) \in \mathcal{K}$$

となり，この問題の緩和問題は，

$$\min G_0(\boldsymbol{x}, (y_{\boldsymbol{a}}: \boldsymbol{a} \in \mathcal{A})) \text{ subject to } \boldsymbol{G}(\boldsymbol{x}, (y_{\boldsymbol{a}}: \boldsymbol{a} \in \mathcal{A})) \in \mathcal{K} \tag{10}$$

となる．問題 (10) は錐上の線型計画問題なので，双対問題を考えることができる．双対問題は，

$$\max \sum_{j=1}^m \gamma_j v_j \text{ subject to } \boldsymbol{v} \in \mathcal{G} \tag{11}$$

となる．ただし，

$$\mathcal{G} = \left\{ \boldsymbol{v} \in \mathcal{K}^* \,\middle|\, \sum_{j=1}^m \boldsymbol{d}_j v_j = \boldsymbol{d}_0, \sum_{j=1}^m c_j(\boldsymbol{a}) v_j = c_0(\boldsymbol{a}) \quad (\boldsymbol{a} \in \mathcal{A}^N) \right\}$$

であり，$\mathcal{K}^*$ は $\mathcal{K}$ の双対錐である．ここで，問題 (1) に対する Lagrange 関数 $L(\boldsymbol{x}, \boldsymbol{v})$ を

$$L(\boldsymbol{x}, \boldsymbol{v}) = g_0(\boldsymbol{x}) - \langle \boldsymbol{v}, \boldsymbol{g}(\boldsymbol{x}) \rangle \quad (\boldsymbol{v} \in \mathcal{K}^*)$$

と定める．ここで，$\langle \cdot, \cdot \rangle$ は，内積である．この時，次のことが言える：

補題 4.1 ([3])

1. $(\boldsymbol{v}, \zeta) \in \mathcal{K}^* \times \mathbb{R}$ が

$$L(\boldsymbol{x}, \boldsymbol{v}) = \zeta \quad (\forall \boldsymbol{x} \in \mathbb{R}^n) \tag{12}$$

   をみたすことと，$\boldsymbol{v} \in \mathcal{G}$, $\zeta = \sum_{j=1}^m \gamma_j v_j$ は同値である．

2. $(\boldsymbol{v}, \zeta) \in \mathcal{K}^* \times \mathbb{R}$ が (12) を満たすと仮定する．この時，問題 (10) の目的関数は ζ で下から押さえられる．

以上のことから, 最適性の十分条件を導くことができる.

定理 4.2 ([3])
$\zeta^* > -\infty$ を問題 (1) の最適値とする. さらに, $(\boldsymbol{v}^*, \zeta^*) \in \mathcal{K}^* \times \mathbb{R}$ は (12) を満たすと仮定する. この時, $\boldsymbol{v}^*$ は問題 (11) の最適解で, ζ^*, (10) の最適値と (11) の最適値は全て一致する.

定理 (4.2) が定理 (4.1) の拡張であることを示すためには, 問題 (3) に対して, 補題 (4.1) を適用したときに, 恒等式 (12) が (9) に一致していれば良い. 以下で, それを示す.

問題 (1) を次のようにおけば, 問題 (3) を得る :

$$\begin{array}{lllllll} g_i(\boldsymbol{x}) & = & M(g_i\boldsymbol{x}), & & g_{m+1}(\boldsymbol{x}) & = & M(\boldsymbol{x}) \\ \ell_i & = & \boldsymbol{u}(\boldsymbol{x}, N_i) \text{ の次元}, & & \ell_{m+1} & = & \boldsymbol{u}(\boldsymbol{x}, N_{m+1}) \text{ の次元} \\ \mathcal{K}_i & = & \mathcal{S}_+^{\ell_i} \ (i = 1, \ldots, m+1), & & \mathcal{K} & = & \mathcal{K}_1 \times \cdots \times \mathcal{K}_{m+1} \end{array}$$

したがって, $\mathcal{K}^* = \mathcal{K}$ であるから, Lagrange 関数 $L(\boldsymbol{x}, \boldsymbol{V}_1, \ldots, \boldsymbol{V}_{m+1})$ は,

$$\begin{aligned} L(\boldsymbol{x}, \boldsymbol{V}_1, \ldots, \boldsymbol{V}_{m+1}) &= g_0(\boldsymbol{x}) - \sum_{i=1}^{m} \boldsymbol{V}_i \bullet \boldsymbol{M}(g_i\boldsymbol{x}) \\ &\quad - \boldsymbol{V}_{m+1} \bullet \boldsymbol{M}(\boldsymbol{x}) \quad (\forall \boldsymbol{x} \in \mathbb{R}^n, \boldsymbol{V}_i \in \mathcal{S}_+^{\ell_i}) \end{aligned} \tag{13}$$

となる. ここで,

$$L(\boldsymbol{x}, \boldsymbol{V}_1^*, \ldots, \boldsymbol{V}_{m+1}^*) = \zeta \quad (\forall \boldsymbol{x} \in \mathbb{R}^n),$$

を満たす $(\boldsymbol{V}_1^*, \ldots, \boldsymbol{V}_{m+1}^*, \zeta^*) \in \mathcal{S}^{\ell_1} \times \cdots \times \mathcal{S}_+^{m+1} \times \mathbb{R}$ が存在すると仮定する. この時, $\boldsymbol{V}_i^*$ は

$$\boldsymbol{V}_i^* = \sum_{k=1}^{\ell_i} \lambda_{ik} \boldsymbol{w}_{ik} \boldsymbol{w}_{ik}^T \quad (i = 1, \ldots, m+1)$$

と書ける. ここで, $\boldsymbol{w}_{ik}$, は $\boldsymbol{V}_i^*$ の固有ベクトルに, λ_{ik} は $\boldsymbol{V}_i^*$ の固有値に各々対応している.

したがって,

$$\begin{aligned} L(\boldsymbol{x}, \boldsymbol{V}_1^*, \ldots, \boldsymbol{V}_{m+1}^*) &= g_0(\boldsymbol{x}) - \sum_{i=1}^{m} \left(\sum_{k=1}^{\ell_i} \lambda_{ik} \boldsymbol{w}_{ik} \boldsymbol{w}_{ik}^T \right) \bullet g_i(\boldsymbol{x}) \boldsymbol{u}(\boldsymbol{x}, N_i) \boldsymbol{u}(\boldsymbol{x}, N_i)^T \\ &\quad - \left(\sum_{k=1}^{\ell_{m+1}} \lambda_{m+1k} \boldsymbol{w}_{m+1k} \boldsymbol{w}_{m+1k}^T \right) \bullet \boldsymbol{u}(\boldsymbol{x}, N_{m+1}) \boldsymbol{u}(\boldsymbol{x}, N_{m+1})^T \\ &= g_0(\boldsymbol{x}) - \sum_{i=1}^{m} \left(\sum_{k=1}^{\ell_i} g_i(\boldsymbol{x}) \left(\sqrt{\lambda_{ik}} \boldsymbol{w}_{ik}^T \boldsymbol{u}(\boldsymbol{x}, N_i) \right)^2 \right) \\ &\quad - \sum_{k=1}^{\ell_{m+1}} \left(\sqrt{\lambda_{m+1k}} \boldsymbol{w}_{m+1k}^T \boldsymbol{u}(\boldsymbol{x}, N_{m+1}) \right)^2 \end{aligned}$$

ここで,

$$\begin{aligned} t_i(\boldsymbol{x}) &= \sum_{k=1}^{\ell_i} g_i(\boldsymbol{x}) \left(\sqrt{\lambda_{ik}} \boldsymbol{w}_{ik}^T \boldsymbol{u}(\boldsymbol{x}, N_i) \right)^2 \quad (i = 1, \ldots, m) \\ q(\boldsymbol{x}) &= \sum_{k=1}^{\ell_{m+1}} \left(\sqrt{\lambda_{m+1k}} \boldsymbol{w}_{m+1k}^T \boldsymbol{u}(\boldsymbol{x}, N_{m+1}) \right)^2 \end{aligned}$$

とおくと, (9) になっている.

4.2 緩和問題の実行可能領域

$\widehat{\mathcal{F}}$ を次のように定める：

$$\widehat{\mathcal{F}} = \left\{ \boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{G}\left(\boldsymbol{x}, (y_{\boldsymbol{a}} : \boldsymbol{a} \in \mathcal{A}^N)\right) \in \mathcal{K} \text{ for some } (y_{\boldsymbol{a}} : \boldsymbol{a} \in \mathcal{A}^N) \right\}$$

これは, (10) の実行可能領域を $\boldsymbol{x}$ の空間に射影したものであると考えることができる. $\widehat{\mathcal{F}}$ が凸集合で, この $\widehat{\mathcal{F}}$ は問題 (1) の実行可能領域 $\mathcal{F}$ とすると, $\mathcal{F} \subset \widehat{\mathcal{F}}$ になっていることが分かる. この項では, $\widehat{\mathcal{F}}$ の特徴付けを行い, Fujie-Kojima[1] で提案されている定理 (2.1) を拡張する. 尚, 目的関数は線型関数 $\boldsymbol{c}_0^T\boldsymbol{x}$ とする. もし, 与えられた問題の目的関数が線型でなければ, 変数 x_{n+1} を加えて、目的関数を x_{n+1} にして, $f_0(\boldsymbol{x}) - x_{n+1} \in \mathbb{R}_+$ という制約を加えれば良い. 説明のため記号を定義する. $\mathcal{L}^*$ を

$$\begin{aligned}
\mathcal{L}^* &= \left\{ \boldsymbol{v} \in \mathcal{K}^* \,\middle|\, \boldsymbol{v}^T\boldsymbol{g}(\boldsymbol{x}) \text{ が}\boldsymbol{x}\text{に関して 2 次以上の項を持たない} \right\} \\
&= \left\{ \boldsymbol{v} \in \mathcal{K}^* \,\middle|\, \sum_{j=1}^{m} v_j c_j(\boldsymbol{a}) = 0 \ (\boldsymbol{a} \in \mathcal{A}^N) \right\}
\end{aligned}$$

とすると, 次の集合 $\widetilde{\mathcal{F}}$ が定義できる：

$$\widetilde{\mathcal{F}} = \left\{ \boldsymbol{x} \in \mathbb{R}^n \mid \langle \boldsymbol{v}, \boldsymbol{g}(\boldsymbol{x}) \rangle \geq 0 \ \forall \boldsymbol{v} \in \mathcal{L}^* \right\}$$

この $\widetilde{\mathcal{F}}$ は無限本の線型不等式で構成されている. この $\widehat{\mathcal{F}}$ と $\widetilde{\mathcal{F}}$ の関係を示す.

定理 4.3

1. $\widehat{\mathcal{F}} \subset \widetilde{\mathcal{F}}$
2. $\boldsymbol{g}(\bar{\boldsymbol{x}})$ が $\mathcal{K}$ の内点になるような $\bar{\boldsymbol{x}} \in \mathbb{R}^n$ が存在すると仮定する. この時, $\widehat{\mathcal{F}}$ の閉包と $\widetilde{\mathcal{F}}$ は一致する.

この定理より, $\widehat{\mathcal{F}}$ と $\widetilde{\mathcal{F}}$ はほとんど同じ領域を表していることが分かる. したがって, $\widehat{\mathcal{F}}$ とは幾何的には, $g_j(\boldsymbol{x})$ の非負結合で表される線型不等式が構成する領域だと考えることができる.

最後に, 定理 (4.3) が [1] の拡張であることを示す. [1] では, 次の非凸 2 次計画問題に対する SDP 緩和を提案している：

$$\min \boldsymbol{c}^T\boldsymbol{x} \text{ subject to } p_j(\boldsymbol{x}) = \boldsymbol{x}^T\boldsymbol{Q}_j\boldsymbol{x} + \boldsymbol{q}_j^T\boldsymbol{x} + \gamma_j \leq 0, \ (j = 1, \ldots, m) \tag{14}$$

ここで $\boldsymbol{Q}_j$ は n 次の対称行列, $\boldsymbol{c}, \boldsymbol{q}_j \in \mathbb{R}^n, \gamma_j \in \mathbb{R}$ である. [1] では, 上の非凸 2 次計画問題に対して $\widehat{\mathcal{F}}_{FK}$, $\widetilde{\mathcal{F}}_{FK}$ を次のように定めている：

$$\begin{aligned}
\widehat{\mathcal{F}}_{FK} &= \left\{ \boldsymbol{x} \in \mathbb{R}^n \,\middle|\, \begin{pmatrix} \gamma_j & \boldsymbol{q}_j/2 \\ \boldsymbol{q}_j/2 & \boldsymbol{Q}_j \end{pmatrix} \bullet \begin{pmatrix} 1 & \boldsymbol{x} \\ \boldsymbol{x} & \boldsymbol{X} \end{pmatrix} \leq 0 \text{ for some } \boldsymbol{X} \in \mathcal{S}_+^n \right\} \\
\widetilde{\mathcal{F}}_{FK} &= \left\{ \boldsymbol{x} \in \mathbb{R}^n \,\middle|\, -\sum_{j=1}^{m} s_j p_j(\boldsymbol{x}) \geq 0 \ \forall \boldsymbol{s} \in \mathbb{R}_+^m \text{ such that } \sum_{j=1}^{m} s_j \boldsymbol{Q}_j \succeq \boldsymbol{O} \right\}
\end{aligned}$$

問題 (15) に対して, 我々の定義した $\widehat{\mathcal{F}},\widetilde{\mathcal{F}}$ を定めたとき, $\widehat{\mathcal{F}}_{FK},\widetilde{\mathcal{F}}_{FK}$ が一致していれば, 定理 (4.3) が [1] の定理 (2.1) の拡張であることが言える. 問題 (15) に対して $\widehat{\mathcal{F}}$ を作ると, $\widehat{\mathcal{F}}_{FK}$ と一致することがわかる. $\widetilde{\mathcal{F}}$ は,

$$\widetilde{\mathcal{F}} = \left\{ \boldsymbol{x} \in \mathbb{R}^n \;\middle|\; -\sum_{j=1}^{m} s_j p_j(\boldsymbol{x}) + \begin{pmatrix} 1 & \boldsymbol{x} \\ \boldsymbol{x} & \boldsymbol{x}\boldsymbol{x}^T \end{pmatrix} \bullet \boldsymbol{V} \geq 0 \ \forall (\boldsymbol{s}, \boldsymbol{V}) \in \mathcal{L}^* \right\}$$

となる. 次の命題より, 定理 (4.3) が [1] の拡張であることがわかる :

命題 4.1 $\widetilde{\mathcal{F}} = \widetilde{\mathcal{F}}_{FK}$.

証明 $(\widetilde{\mathcal{F}}_{FK} \subset \widetilde{\mathcal{F}})$. $\bar{\boldsymbol{x}} \in \widetilde{\mathcal{F}}_{FK}$ ならば,

$$-\sum_{j=1}^{m} s_j p_j(\bar{\boldsymbol{x}}) \geq 0 \ \forall \boldsymbol{s} \in \mathbb{R}^m_+ \text{ such that } \sum s_j \boldsymbol{Q}_j \succeq \boldsymbol{O}$$

が成り立つ. この $\boldsymbol{s}$ の集合を Λ とおく. $(\Lambda := \{\boldsymbol{s} \in \mathbb{R}^m_+ \mid \sum s_j \boldsymbol{Q}_j \succeq \boldsymbol{O}\})$. 今, $\mathcal{L}^*$ の元 $(\boldsymbol{s}, \boldsymbol{U})$ について考える. この時,

$$\boldsymbol{U} = \begin{pmatrix} w & \boldsymbol{u}^T \\ \boldsymbol{u} & \boldsymbol{W} \end{pmatrix}$$

とおくと, $-\sum s_j \boldsymbol{Q}_j + \boldsymbol{W} = \boldsymbol{O}$ であり $\boldsymbol{U} \in \mathcal{S}^{n+1}_+$ より, $\boldsymbol{W} \succeq \boldsymbol{O}$ である. したがって, $\boldsymbol{W} = \sum s_j \boldsymbol{Q}_j \succeq \boldsymbol{O}$ である. このことより, $\mathcal{L}^*$ の全ての元 $(\boldsymbol{s}, \boldsymbol{U})$ は必ず, $\boldsymbol{s} \in \Lambda$ となる. 以上より,

$$-\sum_{j=1}^{m} s_j p_j(\bar{\boldsymbol{x}}) \geq 0 \ (\boldsymbol{v} \in \Lambda)$$

より,

$$-\sum_{j=1}^{m} s_j p_j(\bar{\boldsymbol{x}}) + \begin{pmatrix} 1 & \bar{\boldsymbol{x}}^T \\ \bar{\boldsymbol{x}} & \bar{\boldsymbol{x}}\bar{\boldsymbol{x}}^T \end{pmatrix} (1, \bar{\boldsymbol{x}}^T) \bullet \boldsymbol{U} \geq 0 \ (\forall (\boldsymbol{s}, \boldsymbol{U}) \in \mathcal{L}^*)$$

となる. よって, $\bar{\boldsymbol{x}} \in \widetilde{\mathcal{F}}$ である.

$(\widetilde{\mathcal{F}} \subset \widetilde{\mathcal{F}}_{FK})$. このことを示すために $\bar{\boldsymbol{x}} \notin \widetilde{\mathcal{F}}_{FK}$ ならば $\bar{\boldsymbol{x}} \notin \widetilde{\mathcal{F}}$ ということを示す. $\bar{\boldsymbol{x}} \notin \widetilde{\mathcal{F}}_{FK}$ より, $-\sum_{j=1} \bar{s}_j p_j(\bar{\boldsymbol{x}}) < 0$ かつ, $\sum \bar{s}_j \boldsymbol{Q}_j \succeq \boldsymbol{O}$ となる, $\bar{s} \in \mathbb{R}^n_+$ が存在する. $\boldsymbol{V}$ を

$$\boldsymbol{V} = \begin{pmatrix} \bar{\boldsymbol{x}}^T \left(\sum \bar{s}_j \boldsymbol{Q}_j\right) \bar{\boldsymbol{x}} & -\left(\sum \bar{s}_j \boldsymbol{Q}_j\right) \bar{\boldsymbol{x}} \\ -\left(\sum \bar{s}_j \boldsymbol{Q}_j\right) \bar{\boldsymbol{x}} & \sum \bar{s}_j \boldsymbol{Q}_j \end{pmatrix}$$

とおくと, $\boldsymbol{V} \succeq \boldsymbol{O}$ となる. この $(\bar{s}, \boldsymbol{V})$ を用いると,

$$-\sum \bar{s}_j p_j(\boldsymbol{x}) + \begin{pmatrix} 1 & \boldsymbol{x}^T \\ \boldsymbol{x} & \boldsymbol{x}\boldsymbol{x}^T \end{pmatrix} \bullet \boldsymbol{V} = -\left(\sum \bar{s}_j (\boldsymbol{q}_j + 2\boldsymbol{Q}_j \bar{\boldsymbol{x}})\right)^T \boldsymbol{x} - \sum \bar{s}_j r_j + \bar{\boldsymbol{x}}^T \left(\sum \bar{s}_j \boldsymbol{Q}_j\right) \bar{\boldsymbol{x}}$$

となり, 2 次項が出ない. したがって, $(\bar{s}, V) \in \mathcal{L}^*$ であるが,

$$\begin{aligned} -\sum \bar{s}_j p_j(\bar{x}) + \begin{pmatrix} 1 & \bar{x}^T \\ \bar{x} & \bar{x}\bar{x}^T \end{pmatrix} \bullet V &= -\left(\sum \bar{s}_j (q_j + 2Q_j \bar{x})\right)^T \bar{x} \\ &\quad - \sum \bar{s}_j r_j + \bar{x}^T \left(\sum \bar{s}_j Q_j\right) \bar{x} \\ &= -\sum_{j=1}^{m} \bar{s}_j p_j(\bar{x}) < 0 \end{aligned}$$

となり, $\bar{x} \notin \widetilde{\mathcal{F}}$ である. したがって, $\widetilde{\mathcal{F}} \subset \widetilde{\mathcal{F}}_{FK}$ が言えた. □

5 おわりに

本論文では, 錐上の多項式制約を持つ最適化問題の緩和手法について提案した. この手法は, 非常に一般的なものである. 一方で, 理論的にも実用的にも解決しなければいけないことがある.

理論的には, 緩和値と冗長な制約式, 制約式同士の掛け合わせの関係があげられる. どのような冗長な制約式を導入して, どのような掛け合わせを行うと大域的最適値を得られるのか, もしくは, 大域的最適値に近い値が得られるのか, 非常に重要な問題点だと言える.

実用的にも, 同様のことを考えなければならない. また, SOCP 緩和がどれくらい効果を発揮するかも興味深いところである.

参考文献

[1] T. Fujie and M. Kojima. Semidefinite programming relaxation for nonconvex quadratic programs. *Journal of Global Optimization*, Vol. 10, pp. 367–380, 1997.

[2] S. Kim and M. Kojima. Second order cone programming relaxation of nonconvex quadratic optimization problems. *Optimization Methods and Software*, Vol. 15, pp. 201–224, 2001.

[3] M. Kojima, S. Kim, and H. Waki. A general framework for convex relaxation of polynomial optimization problems over cones. Technical Report B-380, Department of mathmatical and Computing Science Tokyo Institute of Technology, 2002.

[4] M. Kojima and L. Tuncel. Cones of matrices and successive convex relaxations of nonconvex sets. *SIAM Journal on Optimization*, Vol. 10, pp. 750–778, 2000.

[5] J. B. Lasserre. Global optimization with ploynomials and the problem of moments. *SIAM Journal on Optimization*, Vol. 11, pp. 796–817, 2001.