クラスタリング情報つき半正定値計画問題に対する双対解法

東京工業大学　*ナムチャイシリ　チャールズ　NAMCHAISIRI Charles
05001310　東京工業大学　劉田香　LIU Tianxiang
01704970　東京工業大学　山下真　YAMASHITA Makoto

1.　はじめに

本研究では，以下の対数行列式つき半正定値計画問題を対象とする．

$$
\begin{aligned}
\min_{\boldsymbol{X}\in\mathbb{S}^n}\ & f(\boldsymbol{X}) := \boldsymbol{C}\bullet\boldsymbol{X} - \mu\log\det\boldsymbol{X} + \rho\sum_{i<j}|X_{ij}| \\
& + \lambda\sum_{(i,j)<(s,t)}|X_{ij}-X_{st}| \qquad (\mathcal{P}) \\
\text{s.t. } & \mathcal{A}(\boldsymbol{X}) = \boldsymbol{b}, \boldsymbol{X}\succ\boldsymbol{O}.
\end{aligned}
$$

ここで，$\mathbb{S}^n$ はすべての n 次実対称行列の集合とし，$\boldsymbol{X}\succ\boldsymbol{O}$ は行列 $\boldsymbol{X}$ が正定値であることとする．また，関数 $\mathcal{A}:\mathbb{S}^n\to\mathbb{R}^m$ は $\mathcal{A}(\boldsymbol{X}) = (\boldsymbol{A}_1\bullet\boldsymbol{X}, \boldsymbol{A}_2\bullet\boldsymbol{X}, \dots, \boldsymbol{A}_m\bullet\boldsymbol{X})^T$ と定義し，$\boldsymbol{A}_1, \boldsymbol{A}_2, \dots, \boldsymbol{A}_m \in \mathbb{S}^n$ と $\boldsymbol{b}\in\mathbb{R}^m$ は与えられているとする．$1\le i<j\le n$ と $1\le s<t\le n$ に対して，$(i,j)<(s,t) \leftrightarrow (i<j)\vee(i=j\wedge s<t)$ と定義する．

問題 $(\mathcal{P})$ は,以下の問題の拡張として捉えられる．

$$
\begin{aligned}
\min_{\boldsymbol{X}\in\mathbb{S}^n}\ & \boldsymbol{C}\bullet\boldsymbol{X} - \mu\log\det\boldsymbol{X} + \rho\sum_{i<j}|X_{ij}| \\
\text{s.t. } & \mathcal{A}(\boldsymbol{X}) = \boldsymbol{b}, \boldsymbol{X}\succ\boldsymbol{O}.
\end{aligned} \qquad (1)
$$

問題 (1) は様々な分野に応用を持ち, 統計のグラフィカルモデルとも関係している．また，問題 (1) の解法もいくつか提案されており，特に Nakagaki et al. [1] は Dual Spectral Projected Gradient (DSPG) 手法を提案した. この方法は (1) の双対問題に適用され，数値実験により効率的な求解可能であることが示されている．

Lin et al. [2] では問題 $(\mathcal{P})$ を議論している．目的関数の第 4 項はクラスタリングの情報を取り込むために導入されているが，$\lambda\sum_{(i,j)<(s,t)}|X_{ij}-X_{st}|$ の部分を直接計算するには $\mathcal{O}(n^4)$ 回の計算が必要である．[2] では symmetric Gauss-Seidel based alternating direction method of the multipliers (sGS-ADMM) と proximal augmented Lagrangian method (pALM) を使用した 2 段階アルゴリズムが提案されている．

本研究では，問題 $(\mathcal{P})$ の双対問題に対して，DSPG アルゴリズム [1] を拡張して新たなアルゴリズムを提案する．さらに，生成される点列の目的関数値が問題 $(\mathcal{P})$ の最適値に収束することを証明する．また，提案アルゴリズムでは各反復での計算量が $\mathcal{O}(n^4)$ から $\mathcal{O}(n^2\log n)$ まで削減可能であり，実用的であることを示す．

2.　提案アルゴリズムの概要

$(\mathcal{P})$ の双対問題は，目的関数の定数項を除いて以下の $(\mathcal{D})$ と同値である:

$$
\begin{aligned}
\max_{\boldsymbol{y}\in\mathbb{R}^m, \boldsymbol{W}\in\mathbb{S}^n, \boldsymbol{S}\in\mathbb{S}^n}\ & g(\boldsymbol{y},\boldsymbol{W},\boldsymbol{S}) := \boldsymbol{b}^T\boldsymbol{y} + \mu\log\det\left(\boldsymbol{C} - \mathcal{A}^*(\boldsymbol{y}) + \frac{\boldsymbol{W}}{2} + \frac{\boldsymbol{S}}{2}\right) \\
\text{s.t. } & |\boldsymbol{W}|\le\rho,\ \boldsymbol{S}\in\mathcal{S},\ \boldsymbol{C} - \mathcal{A}^*(\boldsymbol{y}) + \frac{\boldsymbol{W}}{2} + \frac{\boldsymbol{S}}{2}\succ\boldsymbol{O}.
\end{aligned} \qquad (\mathcal{D})
$$

集合 $\mathcal{S}$ は線形写像

$$
[Q^\dagger(\boldsymbol{Z})]_{ij} = \begin{cases} \sum_{(s,t)>(i,j)} Z_{ij,st} - \sum_{(s,t)<(i,j)} Z_{st,ij} & \text{if } i<j \\ [Q^\dagger(\boldsymbol{Z})]_{ji} & \text{if } i>j \\ 0 & \text{if } i=j \end{cases}
$$

と $\mathcal{Z} := \{\boldsymbol{Z}\in\mathbb{S}^{\frac{n(n-1)}{2}} \mid |Z_{ij,st}|\le\lambda\}$ を用いて,

$$
\mathcal{S} = \left\{\boldsymbol{S}\in\mathbb{S}^n \;\middle|\; \boldsymbol{S} = Q^\dagger(\boldsymbol{Z}),\ \text{where } \boldsymbol{Z}\in\mathcal{Z}\right\}
$$

として定義される．ここで，$\boldsymbol{S}$ が $\mathcal{S}$ の要素であるかどうかの判定は $\mathcal{O}(n^4)$ の計算量が必要であることに注意したい．そのため，$\boldsymbol{S}$ の判定の代わりに，$\boldsymbol{S}$ から $\mathcal{S}$ への射影を計算する．$(\boldsymbol{y},\boldsymbol{W},\boldsymbol{S})$ から $\mathcal{M} = \mathbb{R}^m\times\mathcal{W}\times\mathcal{S}$ への射影を以下のように定義する：

$$
\boldsymbol{P}_{\mathcal{M}}(\boldsymbol{y},\boldsymbol{W},\boldsymbol{S}) := (\boldsymbol{y}, [\boldsymbol{W}]_{\le\rho}, [\boldsymbol{S}]_{\mathcal{S}}).
$$

$[[\boldsymbol{W}]_{\le\rho}]_{ij} = \max\{-\rho, \min\{\rho, W_{ij}\}\}$ とし, $[\boldsymbol{S}]_{\mathcal{S}}$ は $\boldsymbol{S}$ から $\mathcal{S}$ への射影とし，以下のように計算する:

$$
[\boldsymbol{S}]_{\mathcal{S}} := \arg\min_{\boldsymbol{U}\in\mathcal{S}}\left\{\sum_{1\le a<b\le n}(U_{ab}-S_{ab})^2\right\}. \qquad (2)
$$

g の勾配ベクトルは以下のように計算できる:

$$\begin{aligned}\nabla g(\boldsymbol{y},\boldsymbol{W},\boldsymbol{S}) &:= (\nabla_{\boldsymbol{y}} g(\boldsymbol{y},\boldsymbol{W},\boldsymbol{S}), \nabla_{\boldsymbol{W}} g(\boldsymbol{y},\boldsymbol{W},\boldsymbol{S}), \nabla_{\boldsymbol{S}} g(\boldsymbol{y},\boldsymbol{W},\boldsymbol{S}))\\ &= \left(\boldsymbol{b}-\mathcal{A}(\boldsymbol{X}(\boldsymbol{y},\boldsymbol{W},\boldsymbol{S})), \frac{1}{2}\boldsymbol{X}(\boldsymbol{y},\boldsymbol{W},\boldsymbol{S}), \frac{1}{2}\boldsymbol{X}(\boldsymbol{y},\boldsymbol{W},\boldsymbol{S})\right).\end{aligned}$$

ここで,

$$\boldsymbol{X}(\boldsymbol{y},\boldsymbol{W},\boldsymbol{S}) := \mu(\boldsymbol{C}+\frac{\boldsymbol{W}}{2}+\frac{\boldsymbol{S}}{2}-\mathcal{A}^*(\boldsymbol{y}))^{-1}$$

である．また，$\boldsymbol{X}^k = \boldsymbol{X}(\boldsymbol{y}^k,\boldsymbol{W}^k,\boldsymbol{S}^k)$ とし、$\mathcal{F}$ は $(\mathcal{D})$ の実行可能領域とする．

提案アルゴリズムの枠組みは以下のように与えられる．

アルゴリズム 1：Novel DSPG

- **Step 0** 初期点 $(\boldsymbol{y}^0,\boldsymbol{W}^0,\boldsymbol{S}^0)\in\mathcal{F}$ を取る．また，反復回数を $k=0$ とする．
- **Step 1** 終了条件を満たすなら,$(\boldsymbol{y}^k,\boldsymbol{W}^k,\boldsymbol{S}^k)$ を最適解として終了する．
- **Step 2** 探索方向を射影により以下のように求める:

$$\begin{aligned}&(\Delta\boldsymbol{y}^k,\Delta\boldsymbol{W}^k,\Delta\boldsymbol{S}^k)\\ :=&P_{\mathcal{M}}((\boldsymbol{y}^k,\boldsymbol{W}^k,\boldsymbol{S}^k)+\nabla g(\boldsymbol{y}^k,\boldsymbol{W}^k,\boldsymbol{S}^k))\\ &-(\boldsymbol{y}^k,\boldsymbol{W}^k,\boldsymbol{S}^k).\end{aligned}$$

- **Step 3** $\boldsymbol{X}^k$ のコレスキー分解と直線探索を使い，

$$(\boldsymbol{y}^k,\boldsymbol{W}^k,\boldsymbol{S}^k)+\nu^k(\Delta\boldsymbol{y}^k,\Delta\boldsymbol{W}^k,\Delta\boldsymbol{S}^k)\in\mathcal{F}$$

  になり、目的関数 g の値が非単調上昇するようなステップ幅 ν^k を求める．
- **Step 4**

$$\begin{aligned}&(\boldsymbol{y}^{k+1},\boldsymbol{W}^{k+1},\boldsymbol{S}^{k+1})\\ :=&(\boldsymbol{y}^k,\boldsymbol{W}^k,\boldsymbol{S}^k)+\nu^k(\Delta\boldsymbol{y}^k,\Delta\boldsymbol{W}^k,\Delta\boldsymbol{S}^k)\end{aligned}$$

  とし，次へ更新する．
- **Step 5** $k\leftarrow k+1$ とし，Step 1 に戻る．

---

3. 収束性解析

提案手法では，以下のように $(\mathcal{D})$ の最適値を得ることができる．双対定理から双対ギャップが 0 となるため，$(\mathcal{P})$ の最適値も同時に得られる．

定理 1. 双対問題 $(\mathcal{D})$ の最適値を g^* とする．提案アルゴリズムはある反復 k で $g(\boldsymbol{y}^k,\boldsymbol{W}^k,\boldsymbol{S}^k)=g^*$ となるか，あるいは

$$\lim_{k\to\infty} g(\boldsymbol{y}^k,\boldsymbol{W}^k,\boldsymbol{S}^k)=g^*$$

である．

この主定理のために本研究では複数の補題に証明を与えているが，以下の補題などは新しい変数 $\boldsymbol{S}$ に関係している．

補題 2. $(\boldsymbol{y}^*,\boldsymbol{W}^*,\boldsymbol{S}^*)$ が $(\mathcal{D})$ の最適解である必要十分条件は,$(\boldsymbol{y}^*,\boldsymbol{W}^*,\boldsymbol{S}^*)$ は $(\mathcal{D})$ の実行可能点であり，かつ，ある $\alpha>0$ が存在して，

$$\boldsymbol{P}_{\mathcal{M}}((\boldsymbol{y}^*,\boldsymbol{W}^*,\boldsymbol{S}^*)+\alpha\nabla g(\boldsymbol{y}^*,\boldsymbol{W}^*,\boldsymbol{S}^*))=(\boldsymbol{y}^*,\boldsymbol{W}^*,\boldsymbol{S}^*)$$

を満たすことである．

この補題により，$(\mathcal{P})$ と $(\mathcal{D})$ の最適解の関係が求められる．$(\boldsymbol{y}^*,\boldsymbol{W}^*,\boldsymbol{S}^*)$ が $(\mathcal{D})$ の最適解であれば，$\boldsymbol{X}^*:=\boldsymbol{X}(\boldsymbol{y}^*,\boldsymbol{W}^*,\boldsymbol{S}^*)$ は $(\mathcal{P})$ の最適解となる．

4. 各反復の計算量

提案アルゴリズムの各反復で最も計算時間を要するのは Step 2 で探索方向を計算するところであり，特に (2) での射影計算である．$\mathcal{S}$ の構成には $\mathcal{O}(n^4)$ の制約が含まれるため，直接計算をしようとすれば $\mathcal{O}(n^4)$ 以上の計算量が必要である．

本研究では，(2) の双対問題に着目した．双対問題に適切な変換を施すと pool-adjacent-violaters アルゴリズム [3] が適用可能となり，計算量は $\mathcal{O}(n^2\log n)$ に抑えられる．主問題にあたる (2) の最適解も，双対問題の最適解から $\mathcal{O}(n^2)$ の計算で算出できる．よって,提案アルゴリズムの各反復の計算量は $\mathcal{O}(n^2\log n)$ であり，[1] の DSPG を直接適用した場合よりも効率的である．

参考文献

[1] Nakagaki, Fukuda, Kim, and Yamashita, "A dual spectral projected gradient method for log-determinant semidefinite problems," *Comput. Optim. Appl.*, **76** (1), pp. 33–68, 2020.

[2] Lin, Sun, Toh, and Wang, "Estimation of sparse Gaussian graphical models with hidden clustering structure," arXiv:2004.08115, 2020.

[3] Barlow, Bartholomew, Bremner and Brunk, *Statistical inference under order restrictions : the theory and applicatin of isotonic regression*, 1972.