半正定値計画問題における疎性を活用した交互方向乗数法

hikaru

2016 年 6 月 14 日

概要

- 対象とする問題:半正定値計画問題 (semidefinite programming, SDP)
  → 幅広く応用、大規模な問題を解く必要
- 解法:
  主双対内点法 → 多項式時間だが大規模に不向き
  1 次法 → 収束は遅いが一般的に計算量は少ない
- 目的:
  大規模な問題を少ない記憶領域で高速に解きたい

  疎性を活用し 1 次法である交互方向定数法のボトルネックを解消することで達成
  →domain space sparsity & correlative sparsity
- 既存研究
  内点法+疎性の活用

目次

半正定値計画問題

等式標準形の半正定値計画問題

$$
\begin{array}{llll}
(P) & \text{minimize} & \boldsymbol{A}_0 \bullet \boldsymbol{X} \\
 & \text{subject to} & \mathcal{A}(\boldsymbol{X}) = b \\
 & & S^n \ni \boldsymbol{X} \succeq O
\end{array}
\qquad
\begin{array}{lll}
(D) & \text{maximize} & b^T y \\
 & \text{subject to} & \mathcal{A}^*(y) + \boldsymbol{S} = \boldsymbol{A}_0 \\
 & & (\Leftrightarrow \boldsymbol{A}_0 - \mathcal{A}^*(y) \succeq O) \\
 & & S^n \ni \boldsymbol{S} \succeq O, \ \ y \in \mathbb{R}^m
\end{array}
$$

- $b \in \mathbb{R}^m, \ \ \boldsymbol{A}_i \in S^n \ (i = 1, 2, ..., m), \ \ \boldsymbol{A}_0 \in S^n, \ \ S^n$: n 次元対称行列空間
- $\boldsymbol{A} \bullet \boldsymbol{B} := \sum_i \sum_j \boldsymbol{A}_{i,j} \boldsymbol{B}_{i,j} = Tr(\boldsymbol{A}^T \boldsymbol{B})$,
  $\mathcal{A}(\boldsymbol{X}) := (\boldsymbol{A}_1 \bullet \boldsymbol{X}, \boldsymbol{A}_2 \bullet \boldsymbol{X}, ..., \boldsymbol{A}_m \bullet \boldsymbol{X})^T, \ \ \mathcal{A}^*(y) := \sum_{i=1}^m y_i \boldsymbol{A}_i$,
- $\boldsymbol{X} \succeq O \Leftrightarrow \boldsymbol{X}$ の全ての固有値 λ が $\lambda \geq 0 \Leftrightarrow u^T \boldsymbol{X} u \geq 0 \ (\forall u \neq 0)$

交互方向乗数法

$$(D) \quad \text{min} \quad -b^T y \quad \text{s.t.} \quad \mathcal{A}^*(y) + \boldsymbol{S} = \boldsymbol{A}_0,\ \mathcal{S}^n \ni \boldsymbol{S} \succeq O,\ y \in \mathbb{R}^m$$

- 双対問題 (D) に対する拡張ラグランジュ関数

$$\mathcal{L}_\mu(\boldsymbol{X}, y, \boldsymbol{S}) := -b^T y + \boldsymbol{X} \bullet (\mathcal{A}^*(y) + \boldsymbol{S} - \boldsymbol{A}_0) + \frac{1}{2\mu} \|\mathcal{A}^*(y) + \boldsymbol{S} - \boldsymbol{A}_0\|_F^2$$

- 変数を 交互 に更新する反復解法
  (Alternating direction method of multipliers, ADMM)
- k 反復目

$$\begin{aligned}
y^{k+1} &:= \arg\min_{y \in \mathbb{R}^m} \ \mathcal{L}_\mu(\boldsymbol{X}^k, y, \boldsymbol{S}^k) \\
\boldsymbol{S}^{k+1} &:= \arg\min_{S \in \boldsymbol{S}^n} \ \mathcal{L}_\mu(\boldsymbol{X}^k, y^{k+1}, \boldsymbol{S}), \quad \boldsymbol{S} \succeq O \\
\boldsymbol{X}^{k+1} &:= \boldsymbol{X}^k + \frac{\mathcal{A}^*(y^{k+1}) + \boldsymbol{S}^{k+1} - \boldsymbol{A}_0}{\mu}
\end{aligned}$$

更新式の解析解

- y, $\boldsymbol{S}$, $\boldsymbol{X}$ の更新式には解析解が存在

$$y^{k+1} := -(\mathcal{A}\mathcal{A}^*)^{-1}(\mu(\mathcal{A}(\boldsymbol{X}^{\boldsymbol{k}}) - b) + \mathcal{A}(\boldsymbol{S}^k - \boldsymbol{A}_0)), \ ((\mathcal{A}\mathcal{A}^*)_{ij} := \boldsymbol{A}_i \bullet \boldsymbol{A}_j)$$

$\boldsymbol{S}$, $\boldsymbol{X}$ は $\boldsymbol{V}$ を以下とした最適化問題を解くと求められる

$$\boldsymbol{V}^{k+1} := \boldsymbol{A}_0 - \mathcal{A}^*(y^{k+1}) - \mu\boldsymbol{X}^k$$
$$\min_{\boldsymbol{S}\in S^n} \|\boldsymbol{S} - \boldsymbol{V}^{k+1}\|_F^2, \quad \boldsymbol{S} \succeq O$$

$\rightarrow$ $\boldsymbol{V}$ の 固有値分解 を考えると解くことが可能

ADMM の問題点

1. $\mathcal{A}\mathcal{A}^*$ の逆行列計算 (y の更新)
2. $\boldsymbol{V}$ の固有値分解 ($\boldsymbol{X}$, $\boldsymbol{S}$ の更新)

更新式の解析解

- y, $\boldsymbol{S}$, $\boldsymbol{X}$ の更新式には解析解が存在

$$y^{k+1} := -(\mathcal{A}\mathcal{A}^*)^{-1}(\mu(\mathcal{A}(\boldsymbol{X}^{\boldsymbol{k}}) - b) + \mathcal{A}(\boldsymbol{S}^k - \boldsymbol{A}_0)),\ ((\mathcal{A}\mathcal{A}^*)_{ij} := \boldsymbol{A}_i \bullet \boldsymbol{A}_j)$$

$\boldsymbol{S}$, $\boldsymbol{X}$ は $\boldsymbol{V}$ を以下とした最適化問題を解くと求められる

$$\boldsymbol{V}^{k+1} := \boldsymbol{A}_0 - \mathcal{A}^*(y^{k+1}) - \mu \boldsymbol{X}^k$$
$$\min_{\boldsymbol{S} \in S^n} \|\boldsymbol{S} - \boldsymbol{V}^{k+1}\|_F^2, \quad \boldsymbol{S} \succeq O$$

$\rightarrow$ $\boldsymbol{V}$ の 固有値分解 を考えると解くことが可能

ADMM の問題点

1. $\mathcal{A}\mathcal{A}^*$ の逆行列計算 (y の更新) $\rightarrow$ correlative sparsity
2. $\boldsymbol{V}$ の固有値分解 ($\boldsymbol{X}$, $\boldsymbol{S}$ の更新) $\rightarrow$ domain space sparsity

更新式の解析解

- y, $\boldsymbol{S}$, $\boldsymbol{X}$ の更新式には解析解が存在

$$y^{k+1} := -(\mathcal{A}\mathcal{A}^*)^{-1}(\mu(\mathcal{A}(\boldsymbol{X^k}) - b) + \mathcal{A}(\boldsymbol{S}^k - \boldsymbol{A}_0)),\ ((\mathcal{A}\mathcal{A}^*)_{ij} := \boldsymbol{A}_i \bullet \boldsymbol{A}_j)$$

$\boldsymbol{S}$, $\boldsymbol{X}$ は $\boldsymbol{V}$ を以下とした最適化問題を解くと求められる

$$\boldsymbol{V}^{k+1} := \boldsymbol{A}_0 - \mathcal{A}^*(y^{k+1}) - \mu \boldsymbol{X}^k$$
$$\min_{\boldsymbol{S} \in S^n} \|\boldsymbol{S} - \boldsymbol{V}^{k+1}\|_F^2\ , \quad \boldsymbol{S} \succeq O$$

$\rightarrow$ $\boldsymbol{V}$ の 固有値分解 を考えると解くことが可能

ADMM の問題点

1. $\mathcal{A}\mathcal{A}^*$ の逆行列計算 (y の更新) $\rightarrow$ correlative sparsity
2. $\boldsymbol{V}$ の固有値分解 ($\boldsymbol{X}$, $\boldsymbol{S}$ の更新) $\rightarrow$ domain space sparsity

半正定値制約を含む最適化問題

$$\begin{array}{ll} \text{minimize} & f(x, \boldsymbol{X}) \\ \text{subject to} & g(x, \boldsymbol{X}) \in \Omega, \ \ S^n \ni \boldsymbol{X} \succeq O, \ \ x \in \mathbb{R}^l \end{array}$$

$f : \mathbb{R}^l \times S^n \to \mathbb{R}, \ \ g : \mathbb{R}^l \times S^n \to \Omega$

- 等式標準形の SDP の主問題 (P) を含む

domain space sparsity(d-space 疎性)

$N = \{1, 2, ..., n\}, \ \ E := \{(i, j) \in N \times N : i \neq j, \ \ X_{ij} : f, \ g$ の評価に必要 $\}$
$\to \ \ G(N, E)$:domain space sparsity pattern graph

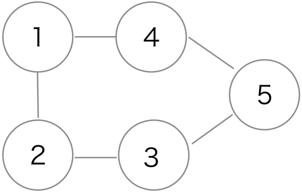
- SDP の例

$$\min \begin{pmatrix} \mathbf{3} & \mathbf{2} & 0 & \mathbf{2} & 0 \\ \mathbf{2} & \mathbf{1} & \mathbf{5} & 0 & 0 \\ 0 & \mathbf{5} & \mathbf{-5} & 0 & \mathbf{2} \\ \mathbf{2} & 0 & 0 & \mathbf{1} & \mathbf{2} \\ 0 & 0 & \mathbf{2} & \mathbf{2} & \mathbf{3} \end{pmatrix} \bullet \boldsymbol{X}, \quad \text{s.t.} \begin{pmatrix} \mathbf{1} & \mathbf{1} & 0 & \mathbf{1} & 0 \\ \mathbf{1} & \mathbf{-2} & \mathbf{3} & 0 & 0 \\ 0 & \mathbf{3} & \mathbf{5} & 0 & \mathbf{2} \\ \mathbf{1} & 0 & 0 & \mathbf{-5} & \mathbf{2} \\ 0 & 0 & \mathbf{2} & \mathbf{2} & \mathbf{1} \end{pmatrix} \bullet \boldsymbol{X} = 1, \ \boldsymbol{X} \succeq O$$

- 疎性があるため下の要素のみで目的関数・等式制約の評価が可能

$$\begin{pmatrix} X_{11} & X_{12} & & X_{14} & \\ X_{21} & X_{22} & X_{23} & & \\ & X_{32} & X_{33} & & X_{35} \\ X_{41} & & & X_{44} & X_{45} \\ & & X_{53} & X_{54} & X_{55} \end{pmatrix}$$

d-space 疎パターングラフ



$\rightarrow$ 特定の要素のみで評価可能

- しかし $\boldsymbol{X} \succeq O$ がある

半正定値行列補完

- 半正定値行列補完

$$\boldsymbol{X} = \begin{pmatrix} 2 & 1 & & 1 & \\ 1 & 9 & 1 & & \\ & 1 & 4 & & 1 \\ 1 & & & 5 & 1 \\ & & 1 & 1 & 4 \end{pmatrix} \xrightarrow{\text{補完}} \overline{\boldsymbol{X}} = \begin{pmatrix} 2 & 1 & 1 & 1 & 2 \\ 1 & 9 & 1 & 0 & 0 \\ 1 & 1 & 4 & 1 & 1 \\ 1 & 0 & 1 & 5 & 1 \\ 2 & 0 & 1 & 1 & 4 \end{pmatrix} \succeq O$$

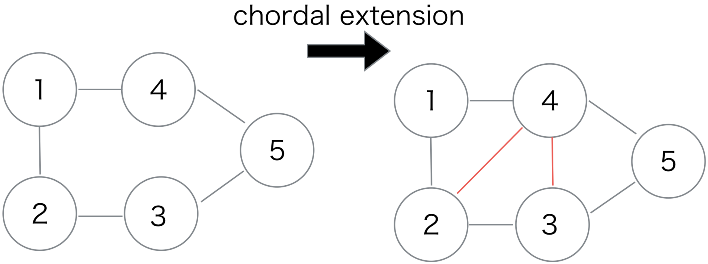
定理:半正定値補完可能な必要十分条件

不完全対称行列 $\boldsymbol{X}$ がある半正定値行列 $\overline{\boldsymbol{X}} \succeq O$ に補完が可能である必要十分条件はコーダルグラフ $G(N, \overline{E})$ の極大クリークを C_k $(k = 1, 2, ..., r)$ とすると $\boldsymbol{X}(C_k) \succeq O$ $(k = 1, 2, ..., r)$

- $G(N, \overline{E})$: $G(N, E)$ のコーダル拡張 (chordal extension)
- コーダルグラフ: 長さが 4 以上のサイクルには chord(弦) をもつグラフ

半正定値制約の分解

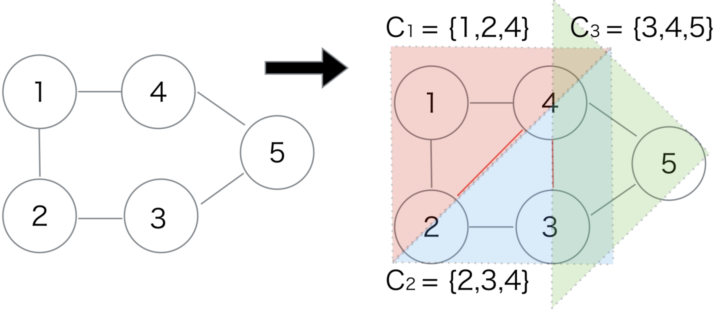
- コーダル拡張 (chordal extension)



$$\boldsymbol{X} \succeq O$$

半正定値制約の分解

- コーダル拡張 (chordal extension)



$$\boldsymbol{X} \succeq O$$

$$\to \begin{pmatrix} X_{11} & X_{12} & X_{14} \\ X_{21} & X_{22} & X_{24} \\ X_{41} & X_{42} & X_{44} \end{pmatrix}, \begin{pmatrix} X_{22} & X_{23} & X_{24} \\ X_{32} & X_{33} & X_{34} \\ X_{42} & X_{43} & X_{44} \end{pmatrix}, \begin{pmatrix} X_{33} & X_{34} & X_{35} \\ X_{43} & X_{44} & X_{45} \\ X_{53} & X_{54} & X_{55} \end{pmatrix} \succeq O$$

domain space conversion

- $G(N, \overline{E})$:d-space 疎パターングラフ $G(N, E)$ のコーダル拡張
- $E \subseteq \overline{E}$ は f, g の評価に必要な要素の位置 (i, j) の集合
- $C_k : G(N, \overline{E})$ の極大クリーク $(k = 1, 2, ..., r)$

domain space conversion(d-space 変換)

$$
\begin{array}{ll}
\text{minimize} & \tilde{f}(x,\ \boldsymbol{X}(C_1),\ \boldsymbol{X}(C_2), ...,\ \boldsymbol{X}(C_r)) \\
\text{subject to} & \tilde{g}(x,\ \boldsymbol{X}(C_1),\ \boldsymbol{X}(C_2), ...,\ \boldsymbol{X}(C_r)) \in \Omega, \\
 & \boldsymbol{X}(C_k) \succeq O \ (k = 1, 2, ..., r),\ x \in \mathbb{R}^l
\end{array}
$$

$\rightarrow$ $\boldsymbol{X}$ の半正定値制約が小さいサイズの行列 $\boldsymbol{X}(C_k)$ の半正定値制約となる

domain space conversion

等式標準形 SDP に d-space 変換

$$
\begin{array}{lll}
(P) & \text{minimize} & \displaystyle\sum_{q=1}^{N} \widehat{\boldsymbol{A}}_0^q \bullet \widehat{\boldsymbol{X}}^q \\
& \text{subject to} & \displaystyle\sum_{q=1}^{N} \widehat{\boldsymbol{A}}_p^q \bullet \widehat{\boldsymbol{X}}^q = \hat{b}_p \ (p = 1, 2, ..., \hat{m}) \\
& & S^{n_q} \ni \widehat{\boldsymbol{X}}^q \succeq O \ (q = 1, 2, ..., N)
\end{array}
$$

- ブロック対角構造にすることで等式標準形 SDP
  $\rightarrow$ ADMM における固有値分解する行列がブロック毎の小さい行列に

$$
\widehat{\boldsymbol{A}}_p = \begin{pmatrix} \widehat{\boldsymbol{A}}_p^1 & O & \dots & O \\ O & \widehat{\boldsymbol{A}}_p^2 & \dots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \dots & \widehat{\boldsymbol{A}}_p^N \end{pmatrix}, \widehat{\boldsymbol{X}} = \begin{pmatrix} \widehat{\boldsymbol{X}}^1 & O & \dots & O \\ O & \widehat{\boldsymbol{X}}^2 & \dots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \dots & \widehat{\boldsymbol{X}}^N \end{pmatrix}
$$

$$(p = 0, 1, 2, ..., \hat{m})$$

更新式の解析解

- y, $\boldsymbol{S}$, $\boldsymbol{X}$ の更新式には解析解が存在

$$y^{k+1} := -(\mathcal{A}\mathcal{A}^*)^{-1}(\mu(\mathcal{A}(\boldsymbol{X}^{\boldsymbol{k}}) - b) + \mathcal{A}(\boldsymbol{S}^k - \boldsymbol{A}_0)), \ ((\mathcal{A}\mathcal{A}^*)_{ij} := \boldsymbol{A}_i \bullet \boldsymbol{A}_j)$$

$\boldsymbol{S}$, $\boldsymbol{X}$ は $\boldsymbol{V}$ を以下とした最適化問題を解くと求められる

$$\boldsymbol{V}^{k+1} := \boldsymbol{A}_0 - \mathcal{A}^*(y^{k+1}) - \mu \boldsymbol{X}^k$$
$$\min_{\boldsymbol{S} \in S^n} \|\boldsymbol{S} - \boldsymbol{V}^{k+1}\|_F^2 , \quad \boldsymbol{S} \succeq O$$

$\rightarrow$ $\boldsymbol{V}$ の固有値分解を考えると解くことが可能

ADMM の問題点

1. $\mathcal{A}\mathcal{A}^*$ の逆行列計算 (y の更新) $\rightarrow$ correlative sparsity
2. $\boldsymbol{V}$ の固有値分解 ($\boldsymbol{X}, \boldsymbol{S}$ の更新) $\rightarrow$ domain space sparsity

$$\text{min} \quad f_0(y) \quad \text{s.t.} \quad F_k(y) \succeq O \ (k = 1, 2, ..., p), \ y \in \mathbb{R}^m$$

→ 等式標準形 SDP の双対問題 (D) を含む ($p = 1$)

correlative sparsity

$$N = \{1, 2, ..., m\}, \ E := \{(i, \ j) : \exists k \text{ s.t. } y_i \text{と } y_j \text{の値が } F_k(y) \text{ の評価に必要 } \}$$
$$\rightarrow G(N, E) : \text{correlative sparsity pattern graph}$$

双対問題 (D) の制約

$$F_1(y) = \boldsymbol{A}_0 - \mathcal{A}^*(y) = \boldsymbol{A}_0 - \sum_{i=1}^{m} y_i \boldsymbol{A}_i \succeq O$$

→ 対角ブロックが 1 つなら $\boldsymbol{A}_i \neq O \ (i = 1, 2, ..., m)$ より F_1 の評価に全ての y_i が必要なため correlative sparsity がない ($G(N, E)$ が完全グラフ)

- 等式標準形 SDP に d-space 変換

  双対問題 (D) の 1 個の制約 $F_1(y) = \boldsymbol{A}_0 - \sum_{i=1}^{m} y_i A_i \succeq O$

  $\rightarrow$ N 個の制約$\hat{F}_k(\hat{y}) = \widehat{\boldsymbol{A}}_0^k - \sum_{i=1}^{\hat{m}} \hat{y}_i \widehat{\boldsymbol{A}}_i^k \succeq O$ $(k = 1, 2, ..., N)$

  $\rightarrow$ 任意の k に対し$\widehat{\boldsymbol{A}}_i^k = O$ または$\widehat{\boldsymbol{A}}_j^k = O$ となる (i, j) が多ければ correlative sparsity が疎になる

  $\rightarrow$ d-space 変換で correlative sparsity が疎になりやすくなる

例)d-space 変換後

$$\widehat{\boldsymbol{A}}_1 = \begin{pmatrix} \widehat{\boldsymbol{A}}_1^1 & & & \\ & O & & \\ & & O & \\ & & & O \end{pmatrix}, \widehat{\boldsymbol{A}}_2 = \begin{pmatrix} O & & & \\ & O & & \\ & & \widehat{\boldsymbol{A}}_2^3 & \\ & & & O \end{pmatrix}, \widehat{\boldsymbol{A}}_3 = \begin{pmatrix} O & & & \\ & O & & \\ & & \widehat{\boldsymbol{A}}_3^3 & \\ & & & \widehat{\boldsymbol{A}}_3^4 \end{pmatrix},$$

$$\widehat{\boldsymbol{A}}_4 = \begin{pmatrix} O & & & \\ & O & & \\ & & O & \\ & & & \widehat{\boldsymbol{A}}_4^4 \end{pmatrix}, \widehat{\boldsymbol{A}}_5 = \begin{pmatrix} \widehat{\boldsymbol{A}}_5^1 & & & \\ & \widehat{\boldsymbol{A}}_5^2 & & \\ & & \widehat{\boldsymbol{A}}_5^3 & \\ & & & \widehat{\boldsymbol{A}}_5^4 \end{pmatrix}, \widehat{\boldsymbol{A}}_6 = \begin{pmatrix} O & & & \\ & \widehat{\boldsymbol{A}}_6^2 & & \\ & & O & \\ & & & O \end{pmatrix}$$

$$\widehat{F}_1(\hat{y}) = \widehat{\boldsymbol{A}}_0^1 - (\hat{y}_1 \widehat{\boldsymbol{A}}_1^1 + \hat{y}_5 \widehat{\boldsymbol{A}}_5^1), \ \widehat{F}_2(\hat{y}) = \widehat{\boldsymbol{A}}_0^2 - (\hat{y}_5 \widehat{\boldsymbol{A}}_5^2 + \hat{y}_6 \widehat{\boldsymbol{A}}_6^2),$$

$$\widehat{F}_3(\hat{y}) = \widehat{\boldsymbol{A}}_0^3 - (\hat{y}_2 \widehat{\boldsymbol{A}}_2^3 + \hat{y}_3 \widehat{\boldsymbol{A}}_3^3 + \hat{y}_5 \widehat{\boldsymbol{A}}_5^3), \ \widehat{F}_4(\hat{y}) = \widehat{\boldsymbol{A}}_0^4 - (\hat{y}_3 \widehat{\boldsymbol{A}}_3^4 + \hat{y}_4 \widehat{\boldsymbol{A}}_4^4 + \hat{y}_5 \widehat{\boldsymbol{A}}_5^4)$$

d-space conversion & correlative sparsity

例)d-space 変換後

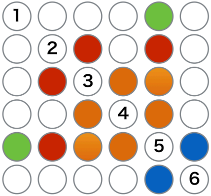
$$\widehat{\boldsymbol{A}}_1 = \begin{pmatrix} \widehat{\boldsymbol{A}}_1^1 & & & \\ & O & & \\ & & O & \\ & & & O \end{pmatrix}, \widehat{\boldsymbol{A}}_2 = \begin{pmatrix} O & & & \\ & O & & \\ & & \widehat{\boldsymbol{A}}_2^3 & \\ & & & O \end{pmatrix}, \widehat{\boldsymbol{A}}_3 = \begin{pmatrix} O & & & \\ & O & & \\ & & \widehat{\boldsymbol{A}}_3^3 & \\ & & & \widehat{\boldsymbol{A}}_3^4 \end{pmatrix},$$

$$\widehat{\boldsymbol{A}}_4 = \begin{pmatrix} O & & & \\ & O & & \\ & & O & \\ & & & \widehat{\boldsymbol{A}}_4^4 \end{pmatrix}, \widehat{\boldsymbol{A}}_5 = \begin{pmatrix} \widehat{\boldsymbol{A}}_5^1 & & & \\ & \widehat{\boldsymbol{A}}_5^2 & & \\ & & \widehat{\boldsymbol{A}}_5^3 & \\ & & & \widehat{\boldsymbol{A}}_5^4 \end{pmatrix}, \widehat{\boldsymbol{A}}_6 = \begin{pmatrix} O & & & \\ & \widehat{\boldsymbol{A}}_6^2 & & \\ & & O & \\ & & & O \end{pmatrix}$$

$$\hat{F}_1(\hat{y}) = \widehat{\boldsymbol{A}}_0^1 - (\hat{y}_1 \widehat{\boldsymbol{A}}_1^1 + \hat{y}_5 \widehat{\boldsymbol{A}}_5^1), \ \hat{F}_2(\hat{y}) = \widehat{\boldsymbol{A}}_0^2 - (\hat{y}_5 \widehat{\boldsymbol{A}}_5^2 + \hat{y}_6 \widehat{\boldsymbol{A}}_6^2),$$

$$\hat{F}_3(\hat{y}) = \widehat{\boldsymbol{A}}_0^3 - (\hat{y}_2 \widehat{\boldsymbol{A}}_2^3 + \hat{y}_3 \widehat{\boldsymbol{A}}_3^3 + \hat{y}_5 \widehat{\boldsymbol{A}}_5^3), \ \hat{F}_4(\hat{y}) = \widehat{\boldsymbol{A}}_0^4 - (\hat{y}_3 \widehat{\boldsymbol{A}}_3^4 + \hat{y}_4 \widehat{\boldsymbol{A}}_4^4 + \hat{y}_5 \widehat{\boldsymbol{A}}_5^4)$$

変換前: 密 ($G(N, E)$ が完全グラフ) $\rightarrow$ 変換後:



correlative sparsity & $\mathcal{A}\mathcal{A}^*$ の疎性

- d-space 変換後の correlative sparsity pattern graph を $G(N, E)$ とすると
  $(i, j) \notin E \Rightarrow$ 全ての k に対し，$\widehat{\boldsymbol{A}}_i^k$, $\widehat{\boldsymbol{A}}_j^k$ の少なくともどちらか一方が O

$$(\widehat{\mathcal{A}\mathcal{A}}^*)_{i,j} = \widehat{\boldsymbol{A}}_i \bullet \widehat{\boldsymbol{A}}_j = \sum_{k=1}^{N} \widehat{\boldsymbol{A}}_i^k \bullet \widehat{\boldsymbol{A}}_j^k = O$$

- $(i, j) \in E$ であっても $\widehat{\boldsymbol{A}}_i \bullet \widehat{\boldsymbol{A}}_j = 0$ になる場合もあるので
  $\widehat{\mathcal{A}\mathcal{A}}^*$ の疎性を $G(N, \widehat{E})$ とすると $\widehat{E} \subseteq E$

  $\rightarrow$ d-space conversion をしても $\widehat{\mathcal{A}\mathcal{A}}^*$ に疎性が現れやすい

y の更新

- y^{k+1} は線形方程式 $\nabla_y \mathcal{L}_\mu(\boldsymbol{X}^k, y, \boldsymbol{S}^k) = 0$
$$\Leftrightarrow (\mathcal{A}\mathcal{A}^*)(y) = -(\mu(\mathcal{A}(\boldsymbol{X}^k) - b) + \mathcal{A}(\boldsymbol{S}^k - \boldsymbol{A}^0))$$
  の解
- d-space 変換をすると correlative sparsity が疎になりやすい
  $\rightarrow$ $\widehat{\mathcal{A}\mathcal{A}^*}$ に疎性が現れやすい
  $\rightarrow$ 行と列を並び替える疎コレスキー分解 (fill-in が少ない) の利用

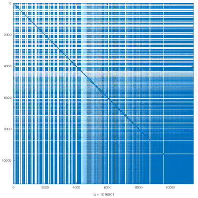


Figure: 並び替え前

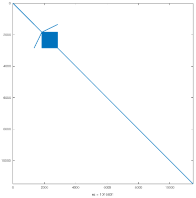


Figure: 並び替え後

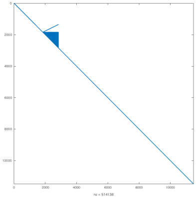


Figure: コレスキー分解

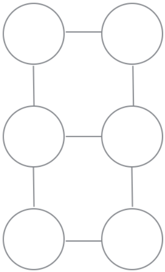
ここまでのまとめ

- 疎性があれば d-space 変換で固有値分解する行列を複数の小さい行列に分解
  $\to$ $\boldsymbol{X}$, $\boldsymbol{S}$ の更新時間の短縮が望める

- d-space 変換で correlative sparsity が疎になりやすくなる
  $\to$ $\widehat{\boldsymbol{A}\boldsymbol{A}^*}$ が疎になりやすくなる
  $\to$ y の更新は $\widehat{\boldsymbol{A}\boldsymbol{A}^*}$ に疎コレスキー分解を利用
  $\to$ メモリを少なく&計算時間短縮が望める

数値実験

- SDP 緩和した格子グラフ (500×10) 上の最大クリーク問題に対して実験を行った



- 実験環境
  - OS:OS X Yosemite 10.10.5
  - CPU:2.7 GHz Intel Core i5
  - メモリ:16 GB 1867 MHz DDR3
  - 内点法:SeDuMi 1.34 (beta)
  - ADMM:MATLAB R2015b で実装
  - 分解:SparseCoLO

- $gap := \frac{|b^T y - \boldsymbol{C} \bullet \boldsymbol{X}|}{1+|b^T y|+|\boldsymbol{C} \bullet \boldsymbol{X}|}$, $pinf := \frac{\|\mathcal{A}(\boldsymbol{X})-b\|_2}{1+\|b\|_2}$, $dinf := \frac{\|\mathcal{A}^*(y)+\boldsymbol{S}-\boldsymbol{C}\|_F}{1+\|\boldsymbol{C}\|_1}$ とし、終了条件は $\delta = \max\{gap, pinf, dinf\} < 10^{-3}$

数値実験

- 最大クリーク... 無向グラフ内の最大のクリークを求める問題

	Memory(GB)	Time(s)	Time/Itr	固有値分解(s)	yの更新(s)	逆行列(s)
ADMM	2.56	over1day	146.6	141.7	0.608	1.7
ADMM - A	12.29	20239	3.4	0.166	2.693	7423
ADMM - B	4.68	1927	0.52	0.208	0.147	-

- ADMM:交互方向乗数法, A:d-space 変換, B:d-space 変換+疎コレスキー分解
- d-space 変換前後のブロック数 N と最大ブロックサイズ Mblk
  - N:1　Mblk:5000　→ N:2095　Mblk:59
- 問題サイズ 行列変数 $(n \times n)$, ベクトル変数 m
  - $n : 5000 \quad m : 9491 \quad \to n : 44058 \quad m : 128493$

数値実験

- 最大クリーク... 無向グラフ内の最大のクリークを求める問題

	Memory(GB)	Time(s)	Time/Itr	固有値分解(s)	yの更新(s)	逆行列(s)
ADMM	2.56	over1day	146.6	141.7	0.608	1.7
ADMM - A	12.29	20239	3.4	0.166	2.693	7423
ADMM - B	4.68	1927	0.52	0.208	0.147	-

- ADMM:交互方向乗数法, A:d-space 変換, B:d-space 変換+疎コレスキー分解
- d-space 変換前後のブロック数 N と最大ブロックサイズ Mblk
  - N:1 Mblk:5000 $\rightarrow$ N:2095 Mblk:59
- 問題サイズ 行列変数 $(n \times n)$, ベクトル変数 m
  - $n : 5000 \quad m : 9491 \rightarrow n : 44058 \quad m : 128493$

数値実験

- 最大クリーク... 無向グラフ内の最大のクリークを求める問題

	Memory(GB)	Time(s)	Time/Itr	固有値分解(s)	yの更新(s)	逆行列(s)
ADMM	2.56	over1day	146.6	141.7	0.608	1.7
ADMM - A	12.29	20239	3.4	0.166	2.693	7423
ADMM - B	4.68	1927	0.52	0.208	0.147	-

- ADMM:交互方向乗数法, A:d-space 変換, B:d-space 変換+疎コレスキー分解
- d-space 変換前後のブロック数 N と最大ブロックサイズ Mblk
  - N:1 Mblk:5000 $\rightarrow$ N:2095 Mblk:59
- 問題サイズ 行列変数 $(n \times n)$, ベクトル変数 m
  - $n : 5000 \quad m : 9491 \rightarrow n : 44058$ m:128493

数値実験

- 最大クリーク... 無向グラフ内の最大のクリークを求める問題

	Memory(GB)	Time(s)	Time/Itr	固有値分解(s)	yの更新(s)	逆行列(s)
ADMM	2.56	over1day	146.6	141.7	0.608	1.7
ADMM - A	12.29	20239	3.4	0.166	2.693	7423
ADMM - B	4.68	1927	0.52	0.208	0.147	-

- ADMM:交互方向乗数法, A:d-space 変換, B:d-space 変換+疎コレスキー分解
- d-space 変換前後のブロック数 N と最大ブロックサイズ Mblk
  - N:1　Mblk:5000　→ N:2095　Mblk:59
- 問題サイズ 行列変数 $(n \times n)$, ベクトル変数 m
  - $n : 5000 \quad m : 9491 \ \rightarrow n : 44058 \quad m : 128493$

数値実験 - 内点法との比較

	Memory(GB)	Time(s)	P-opt	D_opt
内点法	6.96	8243	1.471787	1.471786
ADMM	2.56	over1day	-	-
内点法 - A	11.48	3262	1.472327	1.473374
ADMM - B	4.68	1927	1.474991	1.475355

- ADMM:交互方向乗数法, A:d-space 変換, B:d-space 変換+疎コレスキー分解

数値実験 - 内点法との比較

	Memory(GB)	Time(s)	P-opt	D_opt
内点法	6.96	8243	1.471787	1.471786
ADMM	2.56	over1day	-	-
内点法 - A	11.48	3262	1.472327	1.473374
ADMM - B	4.68	1927	1.474991	1.475355

- ADMM:交互方向乗数法, A:d-space 変換, B:d-space 変換+疎コレスキー分解

- 内点法 :メモリ (多)　計算時間 (多)
- ADMM :メモリ (少)　計算時間 (少)

まとめ

- SDP における 2 つの疎性と、ADMM における疎性の活用を紹介した
  1. domain space sparsity → 行列変数 $\boldsymbol{X}, \boldsymbol{S}$
  2. correlative sparsity → ベクトル変数 y
- ADMM において疎性を活用することで、
  メモリを少なく抑えつつ計算時間の短縮を達成
- 今後の課題
  1. 他の 1 次法との比較
  2. 固有値分解の並列化

参考文献

[1] Z. Wen, D. Godlfarb, and W. Yin. Alternating direction augmented Lagrangian methods for semidefinite programming. Mathematical Programming Computation, 2, 203-230, 2010.

[2] M. Fukuda, M. Kojima, K. Murota, and K. Nakata. Exploiting Sparsity in Semidefinite Programming via Matrix Completion I: General Framework. SIAM Journal on Optimization, 11, 647-674, 2000.

[3] 大規模 SDP 問題を解く研究について http://www.me.titech.ac.jp/~nakata/SDP_large/SDP_large.html

[4] S. Kim, M. Kojima, M. Mevissen, and M. Yamashita. Exploiting sparsity in linear and nonlinear matrix inequalities via positive semidefinite matrix completion, Mathematical Programming, 129, 33-68, 2011.

[5] K. Fujisawa, S. Kim, M. Kojima, Y. Okamoto, and M. Yamashita. User 's Manual for SparseCoLO: Conversion Methods for SPARSE COnic-form Linear Optimization Problems. Department of Mathematical and Computing Sciences Tokyo Institute of Technology, Oh-Okayama, Meguro-ku, Tokyo 152-8552, Japan, Tech. Rep. B-453, 2009.