代数的組合せ最適化 part II

作用素スケーリングとCapacity

1. 導入：行列スケーリングによるパーマネントの近似計算
2. 作用素スケーリングとCapacity
3. さらなる発展

Part IIでは,

$A = \sum_{i=1}^k A_i x_i$, $A_i: \mathbb{C}$ 上 $n \times n$ 行列, の nc-正則性判定 を扱う.

nc-rank $A = n$ かどうか

行列スケーリング Sinkhorn 1964

$B = (b_{ij})$: $n \times n$ 非負行列, ゼロ列, ゼロ行なし

$$R(B) := \begin{pmatrix} (\sum_j b_{1j})^{-1} & & & \\ & (\sum_j b_{2j})^{-1} & & \\ & & \ddots & \\ & & & (\sum_j b_{nj})^{-1} \end{pmatrix} \quad \begin{array}{l} \text{行スケーリング} \\ B \leftarrow R(B) \cdot B \end{array}$$

$$C(B) := \begin{pmatrix} (\sum_i b_{i1})^{-1} & & & \\ & (\sum_i b_{i2})^{-1} & & \\ & & \ddots & \\ & & & (\sum_i b_{in})^{-1} \end{pmatrix} \quad \begin{array}{l} \text{列スケーリング} \\ B \leftarrow B \cdot C(B) \end{array}$$

Question: 行スケーリングと列スケーリングを交互に繰り返すことで, B は 2 重確率行列に収束するか?

行列スケーリング

$$B \leftarrow R(B) \cdot B,$$

$$B \leftarrow B \cdot C(B); \text{ repeat}$$

Thm (Linial, Samorodnitsky, Wigderson 2000)

行列スケーリングで, B が2重確率行列に収束

$\Leftrightarrow G_B$ に完全マッチングが存在

ここで, $G_B := ([n], [n]; E)$, $ij \in E \Leftrightarrow b_{ij} > 0$

注: 2重確率行列に「なる」わけではない, いくらでも近づくという意味

行列スケーリングで完全マッチングの存在が判定できる! ?

→ (多項式時間で) 出来る (Linial, Samorodnitsky, Wigderson 2000)

パーマネントの計算

- #P完全 (Valiant 1979)
- FPRAS (Jerrum, Sinclair, Vigoda 2004)
- ここでは, Linial, Samorodnitsky, Wigderson 2000の行列スケーリングによる(決定的) e^n 近似アルゴリズムの背景にある考え方を紹介する.
- それが, 作用素スケーリングによるnc-正則性判定 (Gurvits 2004, Garg et al. 2015) につながる.

パーマネント (permanent)

$B = (b_{ij}): n \times n$ 非負行列, ゼロ列, ゼロ行なし

パーマネント $\text{Perm } B := \sum_{\sigma} \prod_{i=1}^n b_{i\sigma(i)}$

$\det B$ から符号を除いたもの

Obs. $\text{Perm } B > 0 \Leftrightarrow G_B$ に完全マッチングが存在

Obs. $G = (U, V; E), |U| = |V| = n$

$$B = (b_{ij}), b_{ij} := \begin{cases} 1 & \text{if } ij \in E \\ 0 & \text{otherwise} \end{cases}$$

$\rightarrow \text{Perm } B = |G \text{ の完全マッチング} |$

古典的 Capacity Gurvits 2008

$B = (b_{ij}): n \times n$ 非負行列, ゼロ列, ゼロ行なし

$$p_B(x_1, x_2, \dots, x_n) := \prod_{i=1}^n (Bx)_i = \prod_{i=1}^n \sum_{j=1}^n b_{ij} x_j$$

Obs. $p_B = (\text{Perm } B)x_1 x_2 \cdots x_n + (\text{その他の項})$

Def (Capacity)

$$\text{Cap}(B) := \inf \left\{ \frac{p_B(x_1, x_2, \dots, x_n)}{x_1 x_2 \cdots x_n} \mid x_1 > 0, x_2 > 0, \dots, x_n > 0 \right\}$$

Obs. $\text{Cap}(B) \geq \text{Perm } B$

Thm (一般化van der Waerden不等式; Gurvits 2008)

$$\text{Perm } B \geq \frac{n!}{n^n} \text{Cap}(B)$$

$\approx e^{-n}$

$$\rightarrow \text{Perm } B \leq \text{Cap}(B) \leq e^n \text{Perm } B$$

Cor: $\text{Cap}(B)$ は $\text{Perm } B$ の e^n 近似

$$\text{Cap}(B) > 0 \Leftrightarrow \text{Perm } B > 0 \Leftrightarrow G_B \text{に完全マッチングあり}$$

数え上げ問題 ($\text{Perm } B$) が近似的に最適化問題 ($\text{Cap}(B)$) になった!

$$\text{Cap}(B) := \inf. \frac{\prod_{i=1}^n (Bx)_i}{x_1 x_2 \cdots x_n}$$

$$\text{s.t. } x_1 > 0, x_2 > 0, \dots, x_n > 0.$$

行列スケールリングは Cap の最小化アルゴリズムとみなせる

Def (対称Capacity)

$$\begin{aligned} \text{Cap}^{\text{sym}}(B) &:= \inf. y^T Bx \\ \text{s. t. } & y_1 y_2 \cdots y_n = x_1 x_2 \cdots x_n = 1, \\ & y_1, y_2, \dots, y_n, x_1, \dots, x_n > 0. \end{aligned}$$

Lem: $\text{Cap}^{\text{sym}}(B) = n \sqrt[n]{\text{Cap}(B)}$

$$\because \text{Cap}^{\text{sym}}(B) = \inf_x \inf_y y^T Bx = n \inf_x \sqrt[n]{\prod_i (Bx)_i} = n \sqrt[n]{\text{Cap}(B)}$$

y に関する
凸計画

$$y^* = \sqrt[n]{\prod_i (Bx)_i} \begin{pmatrix} (Bx)_1^{-1} \\ (Bx)_2^{-1} \\ \vdots \\ (Bx)_n^{-1} \end{pmatrix}$$

$\text{Cap}^{\text{sym}}(B) = \inf_{x,y} y^T B x$ に対する交互最適化

$$(x, y) = (\mathbf{1}, \mathbf{1})$$

y を固定して x で最適化

$$\rightarrow x = \sqrt[n]{\det C(B)} C(B) \mathbf{1}$$

repeat $x = \mathbf{1}$ となるように B を規格化: $B \leftarrow \sqrt[n]{\det C(B)} B \cdot C(B)$

x を固定して y で最適化

$$\rightarrow y = \sqrt[n]{\det R(B)} R(B) \mathbf{1}$$

$y = \mathbf{1}$ となるように B を規格化: $B \leftarrow \sqrt[n]{\det R(B)} R(B) \cdot B$

～ B に対する行列スケーリング

行列スケーリングによる完全マッチングの存在判定

Linial, Samorodnitsky, Wigderson 2000

$ds(B) := \|A^T \mathbf{1} - \mathbf{1}\|^2 + \|A\mathbf{1} - \mathbf{1}\|^2$ --- 2重確率行列への近さ

- $ds(B) < \frac{1}{n} \Rightarrow \text{Cap } B > 0$ (完全マッチング存在)
- 行 (列) スケーリングすると $\text{Cap } B \leq 1$
- 初回以降のスケーリングで, $\text{Cap } B$ 非減少
 $ds(B) > \varepsilon \Rightarrow \text{Cap } B$ は $(1 + \Omega(\varepsilon))$ 倍される.
- B : 整数, 完マ存在 $\Rightarrow \text{Cap } B \geq n^{-n}$ (exp. lower bound)

完マッチング存在なら, $\text{poly}(\text{size}(B), n)$ 反復後に $ds(B) < \frac{1}{n}$
 $\therefore (1 + \Omega(\varepsilon))^N \text{Cap } B \leq 1$ より

※ Cap は Perm に置き換えてもよい

行列スケーリングから作用素スケーリングへ

Obs. B が非負行列 $\Leftrightarrow Bz \geq 0$ ($\forall z \geq 0$)

Def. 正值作用素 $T: \mathbb{C}^{n \times n} \rightarrow \mathbb{C}^{n \times n}$

--- 非負行列の一般化

$$\Leftrightarrow T(X) \succcurlyeq 0 \quad (\forall X \succcurlyeq 0)$$

半正定値

Def. 完全正值作用素 $T: \mathbb{C}^{n \times n} \rightarrow \mathbb{C}^{n \times n}$

$$\Leftrightarrow \exists A_1, A_2, \dots, A_k: T(X) = \sum_i A_i X A_i^\dagger$$

複素共役

Gurvits 2004:

- $A = \sum_i A_i x_i \iff T_A(X) = \sum_i A_i X A_i^\dagger$
- 行列スケーリングと Capacity を完全正值作用素へ拡張

Edmonds問題（正則性判定）に対する Gurvitz 2004のアプローチ

- $A = \sum_i A_i x_i \rightarrow T_A(X) = \sum_i A_i X A_i^\dagger$
- 量子パーマメント QPerm を定義
 - $\text{QPerm } T_A =$ 多項式 $\det \sum_i A_i x_i$ の係数ベクトルのノルム
 - $\text{QPerm } T_A > 0 \Leftrightarrow A:$ 正則
- Capacity $\text{Cap}(T_A) := \inf \left\{ \frac{\det \sum_i A_i X A_i^\dagger}{\det X} \mid X \succ 0 \right\}$
- 作用素スケールリング（ $\sim \text{Cap}(T_A)$ の最小化）

$$A:\text{正則} \leftrightarrow \text{QPerm } T_A > 0 \stackrel{?}{\not\leftrightarrow} \text{Cap}(T_A) > 0 \leftrightarrow A:\text{nc-正則}$$

Garg et al. 2015

Rank nondecreasing 性 Gurvits 2004

$$A = \sum_i A_i x_i, \quad T_A(X) = \sum_i A_i X A_i^\dagger$$

Def T_A : rank nondecreasing

$$\Leftrightarrow \text{rank } T_A(X) \geq \text{rank } X \quad (\forall X \succcurlyeq 0)$$

Lem (本質的にGurvits 2004)

$$T_A: \text{rank nondecreasing} \Leftrightarrow A: \text{nc-正則}$$

2重確率作用素

$$A = \sum_i A_i x_i, T(X) = \sum_i A_i X A_i^\dagger$$

$$\text{Def. } T^*(X) := \sum_i A_i^\dagger X A_i$$

$$\text{Def. } T \text{ が 2 重確率} \Leftrightarrow T(I) = I, T^*(I) = I$$

Lem (Gurvits 2004)

$$T(I) = I, \quad \|T^*(I) - I\|^2 < \frac{1}{n} \quad \Rightarrow \quad T: \text{rank nondecreasing} \quad (\Leftrightarrow \text{nc-正則})$$

Lemの証明

$$X \succcurlyeq 0, X = (U \ V) \begin{pmatrix} \lambda_1 & & & \\ & \ddots & & \\ & & \lambda_r & \\ & & & 0 \end{pmatrix} (U \ V)^\dagger, (U \ V)(U \ V)^\dagger = I$$

$$\begin{aligned} \text{rank } T_A(X) &= \text{rank } \sum_i \lambda_i A_i U U^\dagger A_i^\dagger && \begin{array}{l} \text{半正定値} \\ \sum_i Q_i x = 0 \Leftrightarrow \sum_i x^T Q_i x = 0 \\ \Leftrightarrow (\forall i) Q_i x = 0 \end{array} \\ &= \text{rank } \sum_i A_i U U^\dagger A_i^\dagger \end{aligned}$$

$$= \text{rank } T_A(U U^\dagger)$$

$$\geq \text{tr } T_A(U U^\dagger) \quad T(U U^\dagger) \preccurlyeq T(U U^\dagger) + T(V V^\dagger) = T(I) = I$$

$$= \text{tr } T_A^*(I) U U^\dagger \quad = \sum_i \text{tr } A_i U U^\dagger A_i^\dagger = \sum_i \text{tr } A_i^\dagger A_i U U^\dagger$$

$$T^*(I) = I + \Delta$$

$$= r + \langle \Delta, U U^\dagger \rangle$$

Cauchy-Schwarz

$$\geq r - \|\Delta\| \|U U^\dagger\| > r - \sqrt{\frac{r}{n}}$$

$$A = \sum_i A_i x_i, \quad T(X) = \sum_i A_i X A_i^\dagger, \quad T^*(X) := \sum_i A_i^\dagger X A_i$$

行スケールリング $R(T)(X) := T(I)^{-\frac{1}{2}} T(X) T(I)^{-\frac{1}{2}}$

列スケールリング $C(T)(X) := T(T^*(I)^{-1/2} X T^*(I)^{-1/2})$

Obs. $R(T)(I) = C(T)^*(I) = I \quad \because C(T)^*(X) = \sum_i T(I)^{-\frac{1}{2}} A_i X A_i^\dagger T(I)^{-\frac{1}{2}}$

作用素スケールリング

$$T \leftarrow R(T) \quad \Leftrightarrow \quad A_i \leftarrow T^*(I)^{-1/2} A_i$$

$$T \leftarrow C(T) \quad \Leftrightarrow \quad A_i \leftarrow A_i T^*(I)^{-1/2}$$

repeat

作用素スケーリングはCapを最小化している

Def (Capacity)

$$\text{Cap}(T_A) := \inf \left\{ \frac{\det \sum_i A_i X A_i^\dagger}{\det X} \mid X \succ 0 \right\}$$

Def (対称capacity)

$$\text{Cap}^{\text{sym}}(T_A) := \inf \{ \text{tr} \sum_i A_i X A_i^\dagger Y \mid \det X = \det Y = 1, X, Y \succ 0 \}$$

$$\text{Lem: } \text{Cap}^{\text{sym}}(T_A) = n \sqrt[n]{\text{Cap}(T_A)}$$

Rem: 作用素スケーリング $\sim \text{Cap}^{\text{sym}}$ に対する交互最適化

作用素スケーリングによるnc-正則性判定

Garg, Gurvits, Oliveira, Wigderson 2015

$$A = \sum_i A_i x_i, \quad T(X) = \sum_i A_i X A_i^\dagger, \quad T^*(X) := \sum_i A_i^\dagger X A_i$$

$ds(T) := \|T(I) - I\|^2 + \|T^*(I) - I\|^2$ --- 2重確率作用素への近さ

- $ds(T) < \frac{1}{n} \Rightarrow A$:rank nondecreasing (nc-正則)

- 行 (列) スケーリングすると $\text{Cap}(T) \leq 1$

- 初回以降のスケーリングで, $\text{Cap } T$ 非減少

$ds(T) \geq \varepsilon \Rightarrow \text{Cap } T$ は $(1 + \Omega(\varepsilon))$ 倍される.

- A_i :整数, A :nc-正則 $\Rightarrow \text{Cap } T \geq \overset{?}{(\text{exp. lower bound})}$

Thm (Garg, Gurvits, Oliveira, Wigderson 2015)

$$A = \sum_i A_i x_i, \quad A_i \in \mathbb{Z}^{n \times n}, \quad A: \text{nc-正則}$$

$$\Rightarrow \text{Cap}(T_A) \geq \frac{1}{n^{2n}}$$

→ 作用素スケールリングで $\text{poly}(\text{size}(A), n)$ 反復後に $\text{ds}(T) < \frac{1}{n}$

Thmの証明

$$A:\text{nc-正則} \Rightarrow \exists d, \exists D_i \in \mathbb{K}^{d \times d}, \det(\sum_i D_i \otimes A_i) \neq 0$$

$$\forall X \succcurlyeq 0, \det X = 1, \quad C_i := T_A(X)^{-1/2} A_i X^{1/2} \rightarrow \sum_i C_i C_i^\dagger = I$$

$$\begin{aligned} |\det \sum_i D_i \otimes C_i| &= |\det \sum_i D_i \otimes A_i| \det T_A(X)^{-d/2} \det X^{d/2} \\ &\geq \det T_A(X)^{-d/2} \end{aligned}$$

$$\text{Cap}(T_A) \geq \det T_A(X) \geq \left(\det(\sum_i D_i \otimes C_i) (\sum_i D_i \otimes C_i)^\dagger \right)^{-1/d}$$

固有値に関する
相加相乗平均

$$\geq \left(\frac{nd}{\text{tr}(\sum_i D_i \otimes C_i) (\sum_i D_i \otimes C_i)^\dagger} \right)^n$$

Cauchy-Schwarz

$$\geq \left(\frac{nd}{\sum_i \|D_i\|^2 \sum_i \|C_i\|^2} \right)^n = \left(\frac{d}{\sum_i \|D_i\|^2} \right)^n$$

いま示したこと:

$$D_i \in \mathbb{K}^{d \times d}, \det(\sum_i D_i \otimes A_i) \neq 0$$

$$\Rightarrow \text{Cap}(T_A) \geq \left(\frac{d}{\sum_i \|D_i\|^2} \right)^n$$

実は, $\sum_i \|D_i\|^2 \leq n^2 d$ となる $D_i \in \mathbb{Z}^{d \times d}$ が選べる

Alon's Combinatorial Nullstellensatz (Alon 1999)

$p \in \mathbb{C}[x_1, x_2, \dots, x_N]$: 次数 h , 各 x_i の次数 $\leq h_i$

$\rightarrow \exists (a_1, a_2, \dots, a_N) \in \mathbb{Z}_+^N: \sum_i a_i \leq h, a_i \leq h_i,$

$$p(a_1, a_2, \dots, a_N) \neq 0$$

さらなる展開・拡がり

- Capacityの正体

$$\text{Cap}^{\text{sym}}(T_A) = \inf. \text{tr} \sum_i A_i X A_i^\dagger Y \quad \text{s.t.} \quad \det X = \det Y = 1, X, Y > 0$$

$$= \inf. \sum_i \|U A_i V\|^2 \quad \text{s.t.} \quad U, V \in \text{SL}_n(\mathbb{C})$$

～ 群 $\text{SL}_n(\mathbb{C}) \times \text{SL}_n(\mathbb{C})$ の軌道の閉包上の最小ノルム点問題

作用素スケーリングは、最小ノルム $< \varepsilon$ かどうかを $\text{poly}(\text{size}(A), n, 1/\varepsilon)$ で判定

- Geometric Complexity Theory (Mulmuley, Sohoni 2001 ～)からの問題意識

～ 2つの軌道の閉包が交わるか判定できるか？

～ $\text{poly}(\text{size}(A), n, 1/\varepsilon)$ を $\text{poly}(\text{size}(A), n, \log 1/\varepsilon)$ にできるか？

Cap $\sim$ 非正曲率リーマン多様体上の測地的凸最適化

$$\log \text{Cap}(T_A) := \inf. \log \det \sum_i A_i X A_i^\dagger - \log \det X$$

$$\text{s.t. } X > 0$$

目的関数は凸でない

多様体: $P = \{X > 0\}$

リーマン計量: $S, T \in X$ の接空間 ($= \{\text{エルミート行列}\}$)

$$\langle S, T \rangle_X := \text{tr } X^{-1} S X^{-1} T^\dagger$$

Fact: P は非正曲率リーマン多様体, 測地線(最短路)一意

Fact: 目的関数は P 上, 測地的凸

Allen-Zhu, Garg, Li, Oliveria, Wigderson 2017

P 上で Box constraint Newton 法 $\rightarrow \text{poly}(\text{size}(A), n, \log 1/\varepsilon)$ アルゴリズム

さらに拡がりを見せしており、ついていけない。

キーワード：Brascamp-Lieb不等式,

moment polytope, tensor scaling, quantum marginal,

noncommutative optimization,....

FOCS2019にも論文がでている

part II まとめ

- 行列スケーリング, (古典的) Capacity
- 作用素スケーリング
- Capacity, 対称Capacity