行列スケーリングから
非正曲率空間上の測地的凸最適化へ

平井 広志

東京大学大学院 情報理工学系研究科
数理情報学専攻

JST さきがけ

hirai@mist.i.u-tokyo.ac.jp

JST数学関連3領域連携WS「情報科学と拓く新しい数理科学」

2022年9月12日，札幌

さきがけ数理構造活用：

「新しい凸性に基づくアルゴリズムと最適化理論」

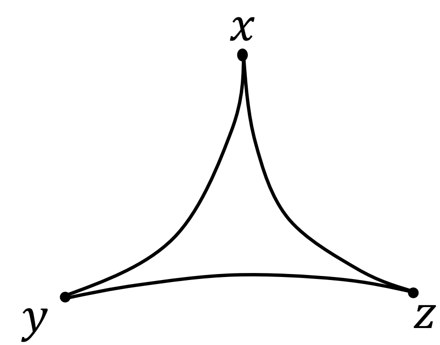
〜 **非正曲率空間の凸性**に基づく新しい連続・離散最適化理論,

および，計算複雑度・アルゴリズム論の展開

やろうとしていること（おおざっぱ）：

- ユークリッド空間 $\mathbb{R}^n$ に埋め込めない・埋め込むことが自然でない空間上の最適化の理論とアルゴリズムの設計
- それを用いて効率的に解ける（多項式時間で解ける）問題クラスを拡げたい

対象とする非正曲率空間：



CAT(0) 空間 $:=$ 測地的距離空間で三角形が痩せている

アダマール空間 $:=$ 完備なCAT(0)空間

アダマール多様体 $:=$ 非正な断面曲率をもつ完備単連結リーマン多様体

$=$ アダマール空間で多様体

CAT(0) 空間では，２点をつなぐ測地線（≃最短路）が一意

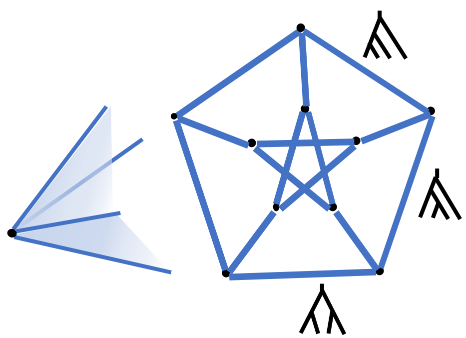
→ 測地線に沿って凸性・凸関数が自然に定義される

→ 最適化理論の舞台としての期待

情報科学・応用数学においても活用の機運が高まっている

例：BHV 系統樹空間（Billera, Holmes, Vogtmann 2001），

多項式時間測地線計算アルゴリズム（Owen 2011）



今日の講演：理論計算機科学の先端的話題における関わり

- 作用素スケーリング & シンボリックランク計算の最近の発展
- 講演者の成果について

（２重確率）行列スケーリング（Sinkhorn 1964）

A: $n \times n$ 非負行列

正対角行列 R, C によって RAC を２重確率行列にできるか？

$RAC\,\mathbf{1} \approx \mathbf{1}, (RAC)^\top \mathbf{1} \approx \mathbf{1}$ としたい

応用

- マルコフ連鎖の推定
- 線形計算の前処理
- 最適輸送，Entropic regularization (Cuturi 2013)

$\vdots$

Sinkhorn のアルゴリズム

- 行正規化：$A \leftarrow RA \quad \text{s.t.} \ (RA)\mathbf{1} = \mathbf{1}; \ R = \mathrm{diag}(\cdots \frac{1}{\sum_k A_{ik}} \cdots)$
- 列正規化：$A \leftarrow \ AC \ \text{s.t.} \ (AC)^{\top}\mathbf{1} = \mathbf{1}; \ C = \ \mathrm{diag}(\cdots \frac{1}{\sum_k A_{kj}} \cdots)$
- これをくりかえす

$$
\begin{array}{c c}
 & \begin{array}{cccc} 4 & 10 & 1 & 1 \end{array} \\
\begin{array}{c} 6 \\ 3 \\ 4 \\ 3 \end{array} & \begin{pmatrix} & 6 & & \\ 1 & & 1 & 1 \\ & 4 & & \\ 3 & & & \end{pmatrix}
\end{array}
\ \xrightarrow{\text{行正規化}} \
\begin{array}{c c}
 & \begin{array}{cccc} 4/3 & 2 & 1/3 & 1/3 \end{array} \\
\begin{array}{c} 1 \\ 1 \\ 1 \\ 1 \end{array} & \begin{pmatrix} & 1 & & \\ 1/3 & & 1/3 & 1/3 \\ & 1 & & \\ 1 & & & \end{pmatrix}
\end{array}
$$

$$
\xrightarrow{\text{列正規化}} \
\begin{array}{c c}
 & \begin{array}{cccc} 1 & 1 & 1 & 1 \end{array} \\
\begin{array}{c} 1/2 \\ 9/4 \\ 1/2 \\ 3/4 \end{array} & \begin{pmatrix} & 1/2 & & \\ 1/4 & & 1 & 1 \\ & 1/2 & & \\ 3/4 & & & \end{pmatrix}
\end{array}
\ \xrightarrow{\text{行正規化}} \ \cdots
$$

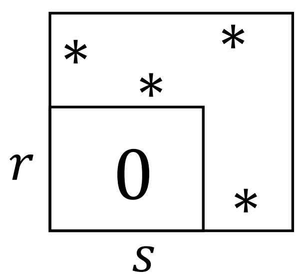
スケーリング可能性の特徴付け

Thm (Sinkhorn-Knopp 1976, Rothblum-Schneider 1989): 以下は同値.

(a) A は（近似的）２重確率スケーリング可能：

$$\forall \epsilon > 0, \exists R, C \ \text{ s.t. } \ \|RAC\mathbf{1} - \mathbf{1}\| < \epsilon, \|(RAC)^{\top}\mathbf{1} - \mathbf{1}\| < \epsilon$$

(b) $\displaystyle \inf_{x,y>0} \log \frac{\left(x^{\top} A y\right)^{n}}{x_1 x_2 \ldots x_n y_1 y_2 \ldots y_n} > -\infty$

(c) $A =$ 

, $r + s > n$ のように行と列を並べ替えれない

(d) Sinkhornアルゴリズムが収束

連続最適化と行列スケーリング

$$\inf \quad \log \frac{(x^\top A y)^n}{x_1 x_2 \ldots x_n y_1 y_2 \ldots y_n} \quad \text{s.t.} \quad x > 0, y > 0$$

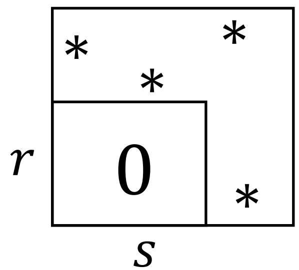
変数変換 $x_i = e^{s_i}, y_j = e^{t_j}$ すると凸最適化

- $R = \mathrm{diag}(\cdots x_i \cdots), \;\; C = \mathrm{diag}(\cdots y_j \cdots)$
- Sinkhornアルゴリズム $\simeq$ 交互最適化

(a) 近似的スケーリング可能性 $\Leftrightarrow$ $\inf_{s,t \in \mathbb{R}^n} \|\nabla f(s,t)\| = 0$

(b) 最適化問題の有界性 $\Leftrightarrow$ $\inf_{s,t \in \mathbb{R}^n} f(s,t) > -\infty$

離散最適化と行列スケーリング

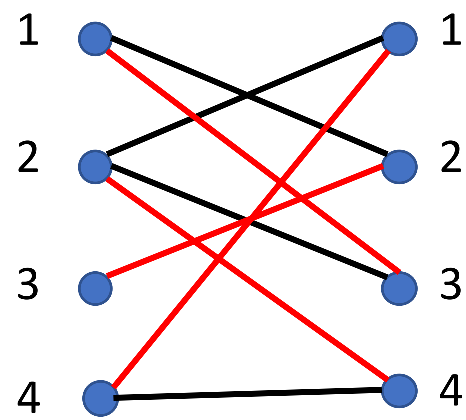
(c) $A=$ 

, $r+s>n$ となるように並べ替えれない

$\Leftrightarrow$ 2 部グラフG_Aに完全マッチングが存在

Hall の結婚定理

$$A=\begin{array}{c|cccc} & 1 & 2 & 3 & 4 \\ \hline 1 & & 6 & 2 & \\ 2 & 1 & & 1 & 1 \\ 3 & & 4 & & \\ 4 & & & & 3 \end{array}$$

$G_A=$ 

Linial, Saromoditsky, Wigderson 2000:

Sinkhorn 反復（多項式回）による完全マッチング存在判定

作用素スケーリング（Gurvits 2004）

$A_1, A_2, \ldots, A_m \in \mathbb{C}^{n \times n}$ がつくる正値作用素 $X \mapsto \sum_k A_k X A_k^\dagger$ が

2 重確率 $\underset{\text{def}}{\Leftrightarrow}$ $\sum_k A_k A_k^\dagger = I$，$\sum_k A_k^\dagger A_k = I$

正則行列 g, hによって，$g(A_1, A_2, \ldots, A_m)h^\dagger$ を2 重確率にできるか？

作用素 Sinkhornアルゴリズム（Gurvitsのアルゴリズム, flip-flop）

- 行正規化: $A_k \leftarrow gA_k$ s.t. $g(\sum_k A_k A_k^\dagger)g^\dagger = I$
- 列正規化: $A_k \leftarrow A_k h^\dagger$ s.t. $h(\sum_k A_k^\dagger A_k)h^\dagger = I$
- これを繰り返す

スケーリング可能性の特徴付け

Thm (Gurvitz 2004): 以下は同値.

(a) $(A_1, A_2, \ldots, A_m)$ は（近似的）2重確率スケーリング可能

(b) $$\inf_{X,Y \succ 0} \log \frac{\left(\mathrm{tr} \sum_k XA_kYA_k^\dagger\right)^n}{\det X \det Y} > -\infty$$

$\{X$: 正定値エルミート$\}^2$上凸でないけど「測地的凸」

(c) $$SA_kT = \begin{array}{c c} & \begin{array}{|cc|}\hline * & * \\ * & \\ \hline \end{array} \\ r & \begin{array}{|c|c|} \hline 0 & * \\ \hline \end{array} \\ & s \end{array}$$ $(\forall k)$, $r + s > n$ となる正則行列 S, T なし.

交互最適化

(d) 作用素 Sinkhorn アルゴリズムが収束.

Garg, Gurvits, Oliveira, Wigderson 2020 (FOCS 2016):

多項式回の Sinkhorn反復 で判定できる.

シンクボリックランク計算（Edmonds 問題）

Edmonds 1967

$A_1, A_2, \dots, A_m \in \mathbb{K}^{n\times n}$, $x_1, x_2, \dots, x_m$ 変数

$$A \coloneqq A_1x_1 + A_2x_2 + \cdots + A_mx_m$$

A の $\mathbb{K}(x_1, x_2, \dots, x_m)$ 上のランクは多項式時間で計算できるか？

- 理論計算機科学の重要な未解決問題
- 組合せ最適化との関連，フレームワークの剛性，行列補完・・・

$$SA_kT = \begin{array}{cc} \begin{array}{c} \\ r \end{array} & \left[\begin{array}{c|c} * \quad * & * \\ \hline 0 & * \end{array}\right] \\ & \begin{array}{c} s \end{array} \end{array} \quad (\forall k),\ r+s>n$$ となる $S,T \in GL(n,\mathbb{K})$ なし．

→ 正則性の必要条件

非可換 Edmonds 問題 Ivanyos, Qiao, Subrahmanyam 2017

$A_1, A_2, \ldots, A_m \in \mathbb{K}^{n\times n}$, $x_1, x_2, \ldots, x_m$ 非可換変数 $x_i x_j \neq x_j x_i$

$$A := A_1 x_1 + A_2 x_2 + \cdots + A_m x_m$$

A の $\mathbb{K}(\langle x_1, x_2, \ldots, x_m \rangle)$ 上のランクは多項式時間で計算できるか？

自由斜体（free skew field）
Amitsur 1966

非可換ランク（nc-rank）

$$SA_kT = \begin{array}{cc} & \\ r & \begin{bmatrix} * & * \\ 0 & * \end{bmatrix} \\ & s \end{array}$$

$(\forall k)$, $r+s>n$ となる $S, T \in GL(n, \mathbb{K})$ なし.

正則性の必要十分条件になる

Cohn 1995, Fortin,Rautenauer 2004

Nc-rank in P

- Garg, Gurvits, Oliveira, Wigderson 2020 (FOCS 2016): $\mathbb{K} = \mathbb{Q}, \mathbb{R}, \mathbb{C}$
  - Sinkhorn 反復（多項式回）によるスケーリング可能性判定
  - アダマール多様体上の測地的凸最適化問題に対する交互最適化

    → さらなる発展（Bürgisser et al. FOCS2019）

- Ivanyos, Qiao, Subrahmanyam 2018 (ITCS2017) : $\mathbb{K}$ 一般
  - Wong sequence ～ 交互道アルゴリズムの代数的一般化

    → 組合せ最適化の代数化，量子情報との関わり，…

これから説明

- Hamada, Hirai 2021 : $\mathbb{K}$ 一般
  - モジュラ束上の劣モジュラ最適化
  - 非多様体的なアダマール空間上の測地的凸最適化

群軌道閉包上の最適化（ Bürgisser et al. FOCS2019 ）

G: 簡約代数群（$\mathbb{C}$上）, $\pi: G \to GL(N, \mathbb{C})$: 有理表現, $v \in \mathbb{C}^N \setminus \{0\}$

$$\text{inf. } \log \|\pi(g)v\| \text{ s.t. } g \in G$$

作用素スケーリング $\simeq$ Left-right action

$G := SL(n, \mathbb{C}) \times SL(n, \mathbb{C})$

$\pi(g, h)(A_1, A_2, \ldots, A_m) := g(A_1, A_2, \ldots, A_m)h^\dagger$

数理科学・情報科学の様々な問題が定式化される：

行列・テンソル正規モデル, テンソルスケーリング,

Quantum marginal, Hornの問題, Null cone メンバシップ, Brascamp-Lieb 不等式,

Geometric Complexity Theory (GCT) との関連, …

G: 簡約代数群（$\mathbb{C}$上），$\pi: G \to GL(N, \mathbb{C})$: 有理表現，$v \in \mathbb{C}^N \setminus \{0\}$

K: 極大コンパクト，$\|\cdot\|$: K不変ノルム

$$\text{inf.} \ \log \|\pi(g)v\| \ \text{ s.t. } \ Kg \in G/K$$

$G/K \simeq G \cap \{$正定値エルミート$\}$: 非正曲率対称空間

→ アダマール多様体

測地線 $t \mapsto g e^{tH} g^{\dagger}$

Thm [Bürgisser et al. FOCS2019]

目的関数 f（Kempf-Ness 関数）は測地的凸

Thm [Kempf-Ness + Hilbert-Mumford] 以下は同値

(a) $\inf_{x \in G/K} \|\nabla f(x)\| = 0$

(b) $\inf_{x \in G/K} f(x) > -\infty \quad (\Leftrightarrow 0 \notin \overline{\pi(G)v}\,)$

(c) No 1-parameter subgroup $t \mapsto e(t)$ s.t. $\lim_{t \to \infty} \pi\big(e(t)\big)v = 0$

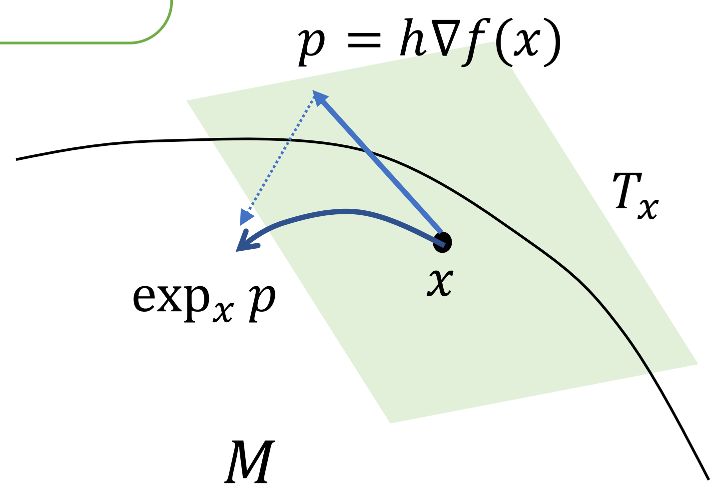
Bürgisser et al. FOCS2019: 最適化の枠組み～リーマン多様体 G/K 上の最適化

１次法（勾配法），２次法（信頼領域 Newton法）

$\mathrm{poly}\left(\frac{1}{\epsilon}, \cdots\right)$ 反復で

$\|\nabla f(x)\| < \epsilon$ なる x を出力

$\mathrm{poly}\left(\log\frac{1}{\epsilon}, \cdots ✸\right)$ 反復で…



定ステップ幅アルゴリズムでは原理的に困難な場合あり

Franks, Reichenbach 2019

→ チャレンジ：内点法のようなアルゴリズム

有界性判定に対する無限遠境界（Boundary at Infinity）からのアプローチ

Kapovich, Leeb, Milson 2009, Hirai 2022:

有界性の判定（条件(c)）は,

無限遠境界錐上の測地的凸最適化として定式化できる.

非多様体的なアダマール空間

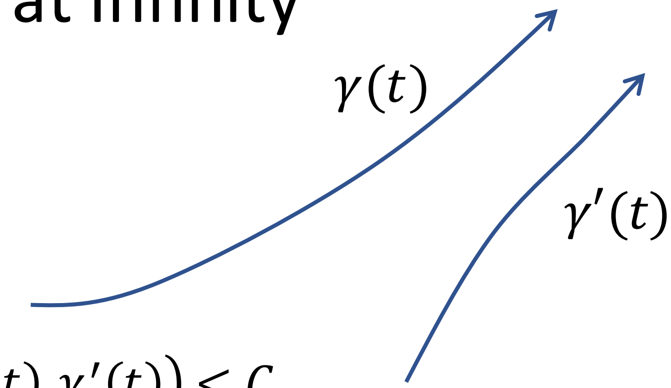
Hamada, Hirai 2021:

非可換ランク計算の多項式時間アルゴリズム

Left-right action のときの

無限遠境界錐上の測地的凸最適化を解いているとみなせる

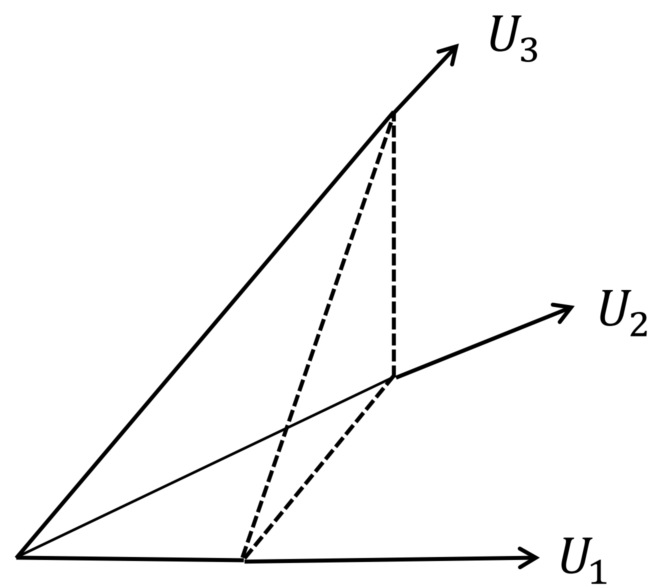
Boundary at Infinity



アダマール多様体の無限遠境界

:= 測地光線の同値類

$$\gamma \sim \gamma' \Leftrightarrow d\big(\gamma(t), \gamma'(t)\big) < C$$



対称空間の無限遠境界（錐）

～（ユークリッド的）ビルディング

という構造が入る

$SL(n,\mathbb{C})/SU(n) = \{X \succ 0, \det X = 1\}$ のとき:

ベクトル空間の列（flag） $U_1 \subset U_2 \subset \cdots \subset U_{n-1} \subset \mathbb{C}^n$ に cone を対応させる.

すべての flag について cone を考えて貼り合わせる.

非可換ランク計算への劣モジュラ最適化からのアプローチ

Thm (Fortin, Reutenauer 2004)

$$\text{nc-rank} \sum_k A_k x_k = 2n - \text{Max.} \quad r + s$$

$$\text{s.t.} \quad SA_kT = \begin{array}{rc} & \left[\begin{array}{c|c} * & * \\ \hline 0 & * \end{array}\right] \\ r & \\ & s \end{array} \quad (\forall k)$$

$$S, T \in GL(n, \mathbb{K})$$

Max. $\dim X + \dim Y$

s.t. $A_k(X, Y) = \{0\}\ (\forall k)$

$X, Y \subseteq \mathbb{K}^n$: ベクトル部分空間

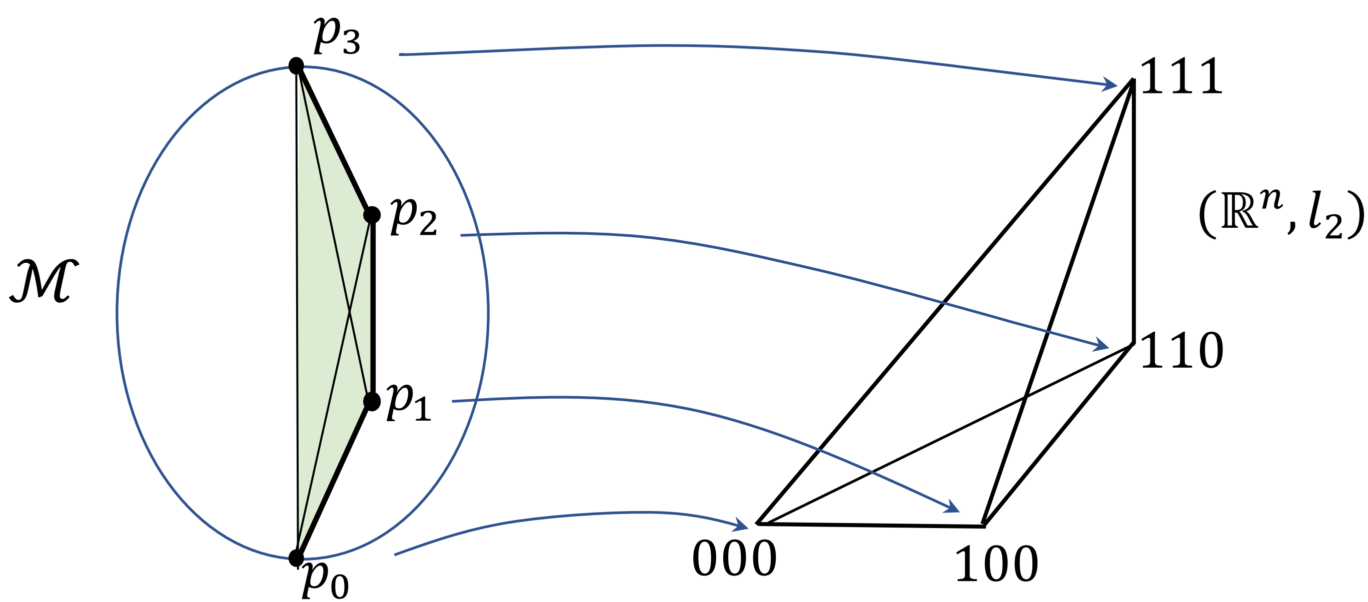
where $A_k(x, y) \coloneqq x^\top A_k y$

ベクトル部分空間のなすモジュラ束 $\mathcal{M}$ 上の劣モジュラ最適化とみなせる

$$f(p) + f(q) \geq f(p \vee q) + f(p \wedge q) \quad (p, q \in \mathcal{M})$$

モジュラ束 $\mathcal{M}$ $\overset{}{\underset{\text{オーソスキーム複体}}{\Longrightarrow}}$ $K(\mathcal{M})$: アダマール空間 (Chalopin, Chepoi, Hirai, Osajda 2020)

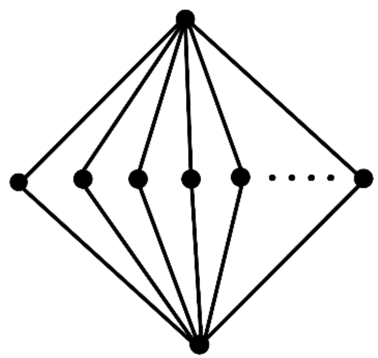
劣モジュラ関数 $f : \mathcal{M} \to \mathbb{R}$ $\underset{\text{Lovász 拡張}}{\Longrightarrow}$ $\bar{f} : K(\mathcal{M}) \to \mathbb{R}$ 測地的凸 (Hirai 2018)



非可換ランク ＝ アダマール空間の測地的凸最適化問題の最適値

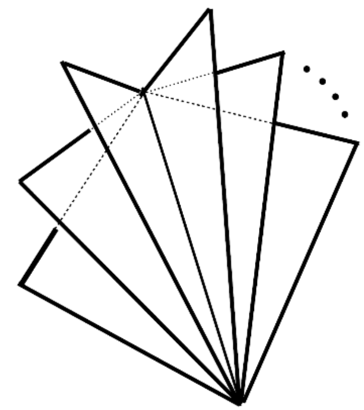
$\cong$

$SL(n, \mathbb{C})/SU(n)$ の無限遠境界上の最適化（$\mathbb{K} = \mathbb{C}$ のとき）



高さ２のモジュラ束 $\mathcal{M}$

→



$K(\mathcal{M})$

Bačák 2014: 分割近接点法

$$\text{Min.} \quad f(x) = \sum_{i=1}^{N} f_i(x) \quad \text{s.t.} \quad x \in K$$

Iterate: $x^{k+1} \longleftarrow \operatorname{argmin}_{x \in K} \ f_{k \bmod N}\ (x) + \frac{1}{\lambda_k} d(x, x^k)^2$

Ohta, Pálfia 2015: f: 強凸 ⇒ 収束レート（sublinear）

Hamada, Hirai 2021: これらを応用する

- $A_k(X,Y) = \{0\}(\forall k)$ → ペナルティ項 $C \sum_k \operatorname{rank} A_k|_{X,Y}$ → Lovász拡張
- 摂動（強凸化），Lipschitz 定数の評価，目的関数の整数性 → 多項式回の反復で十分
- １反復が多項式時間で実行可能 ← 束論的議論
- $\mathbb{K} = \mathbb{Q}$ のときは，ビット複雑度を抑える必要がある
  → p 進付値を利用して，$GF(p)$ 上の問題に帰着させる

まとめ

- 行列スケーリング ～ 作用素スケーリング ～ 群軌道上の最適化,
  シンボリックランク計算，アダマール空間上の凸最適化,
  の関わりを紹介した.

- このように情報科学の先端的話題に非正曲率空間上の凸最適化が
  現れ始めてきており，重要な研究の方向性であると考えている.

- 従来型の離散最適化問題においてもしばしば有用
  → 離散凸解析ビヨンド $\mathbb{Z}^n$ (Hirai 2013～)

おわり

平井広志：「行列スケーリングの数理」（学部 4 年演習資料）

平井広志：「代数的組合せ最適化---Edmonds 問題の最近の発展について」

（2019 年度組合せ最適化セミナー資料）

平井のページからダウンロード可

H. Hirai: Convex analysis on Hadamard spaces and scaling problems, arXiv, 2022.

M. Hamada and H. Hirai:
Computing the nc-rank via discrete convex optimization on CAT(0) spaces,
SIAM Journal on Applied Geometry and Algebra 5 (2021), 455--478.