凸最適化の情報幾何と多項式時間内点法

土谷　隆

(統計数理研究所)

（小原敦美氏（大阪大学）との共同研究）

本研究の目的

- 最近の（連続的）最適化の世界の紹介
- 『計算複雑度』という『情報科学の複雑さ』と『問題の曲率』という『幾何学的複雑さ』を情報幾何を通じて結び付けて論じる。

本講演の目的

- 最近の（連続的）最適化の世界の紹介
- 凸最適化問題に対する情報幾何
- 『計算複雑度』という『情報科学の複雑さ』と『問題の曲率』という『幾何学的複雑さ』を情報幾何を通じて結び付けて論じる。

計算複雑度　　　**モデルの複雑度**

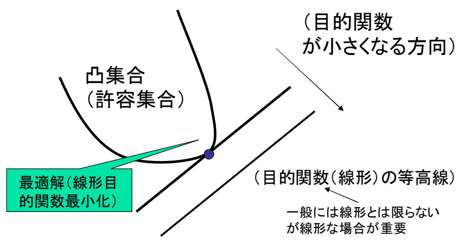
情報幾何

モデリング・数理・アルゴリズム

凸最適化と内点法

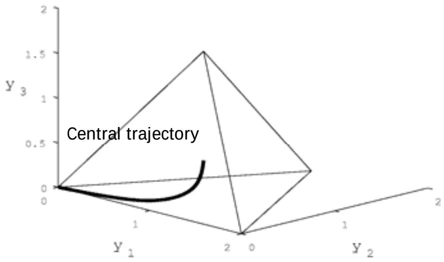
- 凸最適化問題：凸集合上での凸関数最小化問題
  - 線形計画問題：多面体上での線形関数最小化
  - 凸2次計画問題：多面体上での凸2次関数の最小化
- 内点法：凸計画問題を、中心曲線と呼ばれる曲線を離散的に辿りながら解く.

凸最適化問題



許容解集合の次元：数千から数十万次元

Example



内点法

- Karmarkar (1984) : 線形計画に対する(多項式時間)射影変換法
- 小島・水野・吉瀬/田邉(1987):線形計画の主双対内点法 (日本発で世界で使われている数少ないアルゴリズム)
- 現在解ける問題の最大サイズ:

内点法

- Karmarkar (1984) : 線形計画に対する(多項式時間)射影変換法
- 小島・水野・吉瀬/田邉(1987):線形計画の主双対内点法 (日本発で世界で使われている数少ないアルゴリズム)
- 現在解ける問題の最大サイズ:10億変数(資産運用問題)

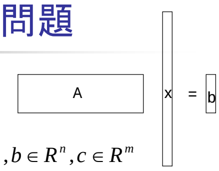
内点法

- Nesterov-Nemirovskii(1994) の Self-concordant 障壁関数による多項式時間凸最適化の一般論
- 凸錐上の線形計画問題:半正定値計画問題、2次錐計画問題
- 対称錐上の線形計画問題に対する多項式時間主双対内点法 --- Euclidean Jordan 代数による一般論 (1995-2000)

登場から20年以上経って …

- Karmarkar 特許は2005年で切れた ☺
- 多項式時間で解けるクラスの凸計画問題が線形計画問題から拡大した.
- 計算機の速度や容量:桁違いに大きくなった.
  (80年代後半のパソコン＝PC9801;
  今のパソコン＝90年代中頃の世界のトップクラスのスーパーコンピュータ)

(古典的)線形計画問題



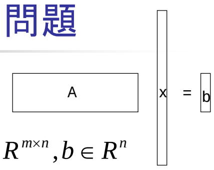
$\min\ c^T x$ (主問題)

$\text{s.t.}\quad Ax = b,\ \ x \geq 0,\ A \in R^{m\times n}, b \in R^n, c \in R^m$

$\max\ b^T y$ (双対問題)

$\text{s.t.}\quad s = c - A^T y,\ \ s \geq 0$

凸錐上の線形計画問題



$\min \ c^T x$ (主問題)

s.t. $Ax = b, \ x \in C, \ A \in R^{m \times n}, b \in R^n$

$\max \ b^T y$ (双対問題)

s.t. $s = c - A^T y, \ s \in C^*,$

C : Convex cone,

C^* : Cone dual to C.

$(C^* = \{s | x^T s \geq 0 \ \forall x \in C\})$

弱双対定理

- 主問題の目的関数値は双対問題の目的関数値以上である。

$$\begin{aligned} c^T x - b^T y &= x^T c - x^T A^T y \\ &= x^T s \geq 0. \end{aligned}$$

等号成立：両問題の最適解が得られている。

対称錐計画問題

- 対称錐: $C = C^*$ （＋等質性）
- 線形計画問題 (C: 第一象限) ローレンツ錐
- 2次錐計画問題 (C: 2次錐の直積)
- 半正定値計画問題 (C: 半正定値対称行列)

これらの問題は内点法で多項式時間で求解可能.
凸2次計画問題など, 目的関数が凸2次の場合もこの枠組みで扱える.

半正定値計画問題（主問題）

min $\mathrm{Tr}[CX]$

s.t. $\mathrm{Tr}[A_i X] = b_i, \ i = 1, \ldots, n$

$X \succeq 0$

X は半正定値対称行列

半正定値計画問題（双対問題）

max $\sum b^T y$

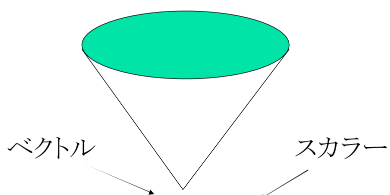
s.t. $Z = C - \sum_{i=1}^{m} A_i y_i,$

$Z \succeq 0,$

Z は半正定値対称行列

2次錐計画問題

2次錐の自画像



$\|x_1\| \leq x_0$

大きな流れ

- 1990年代初頭まで：線形計画問題のみが解ける大規模最適化問題であった。
- 1990年以降 --- 下記の凸計画問題が多項式時間で実用的に解けるようになった。
  - 凸2次計画（多面体上での凸2次関数最小化）
  - 半正定値計画問題（半正定対称行列空間上での線形計画問題）
  - 2次錐計画問題（2次錐の直積上での線形計画問題）

ローレンツ錐

多項式時間アルゴリズム

- 主問題の実行可能解と双対問題の実行可能解の列で、

$$(c^T x^{k+1} - b^T y^{k+1}) \le \left(1 - \frac{0.1}{\sqrt{p}}\right)(c^T x^k - b^T y^k)$$

を満たすものを生成する。(収束速度が p にしか依存しない。p は、C の性質のみから定まるパラメータ)

内点法と実際問題のモデリング

- リニアモーターカー磁気シールド最適設計問題
- データ同化（エルニーニョ予測、解析）
- Support Vector Machine (機械学習、凸2次計画)
- Compressed Sampling, Compressive Sensing (最適化からの視点：線形計画)
- ファイナンス、列車の制御 (凸2次計画)

2次錐計画法による磁気シールド最適設計問題

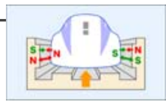
統計数理研究所
土谷　隆
（鉄道技術総合研究所 笹川卓氏との共同研究）

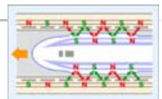
SIAM Journal on Scientific Computing (2003), 統計数理 (2005)

山梨リニア実験線

http://www.rtri.or.jp/

最適磁気シールド設計問題への二次錐計画≒半正定値計画の応用





リニアモーターカー　：　超伝導磁石で浮上，推進

強力な磁場

車体内部をシールドして，内部に磁場がもれないようにする．一方，シールドはできるだけ軽くしたい．

→最適設計問題

図：http://www.pref.yamanashi.jp/

最適化の結果1

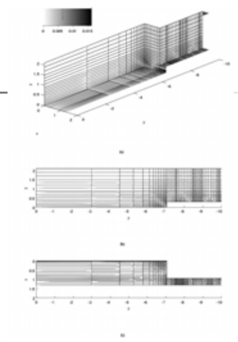
- 問題の大きさ
  主双対変数の数：5007
  自由度（yの次元）：1669
- 時間データ ： 1.8秒
  （精度：10^{-12} ，反復回数：21）

PentiumⅢ 700MHz

最適化の結果2

- 問題の大きさ
  主双対変数の数：約18万
  自由度（yの次元）：約6万
- 時間データ ： 720秒
  （精度：10^{-12}, 反復回数：40回位）

PentiumⅢ 700MHz

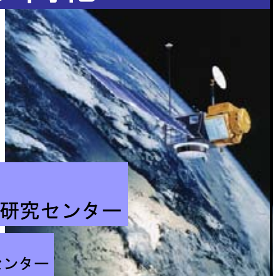


最適化されたシールドの形

データ同化へのグラフィカルモデルの適用（統数研上野助教との共同研究）

- エルニーニョの予測，解析（樋口プロジェクト）**偏微分方程式モデルの実データへのあてはめ**
- アンサンブルカルマンフィルタの尤度を計算するために1981次元の正規分布を推定したい．
- データ数：約300 (人工衛星による海面高度データからトレンドを除去したもの.)
- 半正定値計画問題に対する内点法によりガウシアングラフィカルモデルで推定

アンサンブルカルマンフィルターを用いた大気海洋結合モデルへのデータ同化



上野 玄太
モデリング研究系 / 予測発見戦略研究センター

共同研究者：樋口知之[1]、鍵本崇[2]、広瀬直毅[3]
1. モデリング研究系 / 予測発見戦略研究センター
2. 地球環境フロンティア研究センター
3. 九州大・応用力学研究所

目的: エルニーニョの予測

エルニーニョ

太平洋赤道域の中央部からペルー沿岸にかけて、
海面水温が平年に比べて高くなり、その状態が 1 年程度続く現象

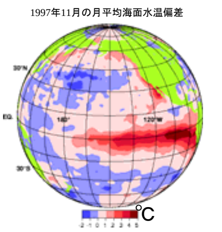
el niño = the boy
el Niño = Jesus Christ
エル･ニーニョ現象 = el fenomeno del Niño

海面水温が平年に比べて低い
状態が続くのは ラニーニャ

la niña = the girl

ほぼ 4 年おきにおきる
↑
これを正確に予測したい

1997年11月の月平均海面水温偏差



http://www.data.kishou.go.jp/climate/elnino/mikata/whatiselnino.html

- 分散共分散行列の逆行列の内メッシュの東西南北近傍に対応する要素だけが非ゼロと仮定
  - 1981×1981の行列
  - 自由度（パラメータ数）は5841
  - 何とかパソコンで求解可能（Pentium IV 2.6GHz, メモリ1GB, Windows XP, Matlab)
- さらに近傍パターンを変えてAICでモデル選択，データ同化モデリング (樋口プロジェクト) に実際に使われている (統計数理研究所のスパコンで28037パラメータのモデルまで探索)

本研究の目的

- 内点法への情報幾何的アプローチ
- 『計算複雑度』という『情報科学の複雑さ』と『問題の曲率』という『幾何学的複雑さ』を結び付けて論じたい。

情報幾何

- 情報幾何学
  - 情報を取り扱うための微分幾何学 (e.g. Amari and Nagaoka (2000))
  - 適用分野
    - 統計学
    - 機械学習
    - Neural Networks
    - 信号処理
    - 制御 etc.,etc.

目的

- 情報幾何の立場から内点法を見直す.
- 情報幾何の立場からの多項式時間予測子・修正子法

凸錐上の線形計画問題（主問題）

(P) min $\langle c, x\rangle$

s.t. $x \in (T + d) \cap \Omega$.

$d \in E(= R^n), \ c \in E^*$

$T : (n - m)$ dim.

linear subspace of E.

Ω : proper closed convex cone.

凸錐上の線形計画問題（双対問題）

(D) min $\langle d, s\rangle$

s.t. $s \in (T^* + c) \cap \Omega^*$.

$T^* = T^\perp$: (m dim.

linear subspace of E^*.)

Ω^* : dual cone of Ω.

双対錐(dual cone)

$$\Omega^* = \{s \mid \langle x, s\rangle \geq 0, \quad \forall x \in \Omega\}$$

第4回応用数学連携フォーラム,
2008/5/30, 東北大学

主問題と双対問題の関係

- 主問題と双対問題： 対称
- 主問題と双対問題： 片一方が解ければもう一つが解ける。

第4回応用数学連携フォーラム,
2008/5/30, 東北大学

凸錐上の線形計画問題

$\mathcal{P} \equiv (T+d) \cap \mathrm{int}(\Omega).$ 主問題の許容領域

$\mathcal{D} \equiv (T^*+c) \cap \mathrm{int}(\Omega^*).$ 双対問題の許容領域

We assume that $\mathcal{P} \neq \emptyset$ and $\mathcal{D} \neq \emptyset$

第4回応用数学連携フォーラム,
2008/5/30, 東北大学

概要

- (直観的ではなく微分幾何学的な意味での)中心曲線の曲率積分に基づいた多項式時間内点法の反復回数の評価.
- 線形計画法の場合
- 情報幾何的な意味

第4回応用数学連携フォーラム,
2008/5/30, 東北大学

正規障壁関数 $\psi(x)$

- Ω 上で定義された凸関数。$x \to \partial\Omega$ で $\psi(x) \to \infty$ となる。
- Self-concordant 障壁関数
- 多項式時間内点法：$\psi(x)$に基づいて構成される。
- $\psi(x)$ の障壁パラメータ p ： 内点法の計算複雑度を定める。

第4回応用数学連携フォーラム,
2008/5/30, 東北大学

中心曲線

- パラメータ t を含む下記の最適化問題の最適解 $x(t)$ が構成する曲線.

$$\min \langle tc, x\rangle + \psi(x), \quad x \in \mathcal{P}.$$

(t: Homotopy parameter) $\quad (0 < t)$

t が∞に近づくと (P) の最適解に収束する。

第4回応用数学連携フォーラム,
2008/5/30, 東北大学

中心曲線

- パラメータ t を含む下記の最適化問題の最適解 $x(t)$ が構成する曲線.

$$\min \ \langle tc,\ x\rangle + \psi(x), \quad x \in \mathcal{P}.$$

(t: Homotopy parameter) $(0 < t)$

この問題を解くにあたっては難しい条件の $x \in \Omega$ を無視できる(障壁関数の利点)。

中心曲線

- この曲線上の点を

第4回応用数学連携フォーラム,
2008/5/30, 東北大学

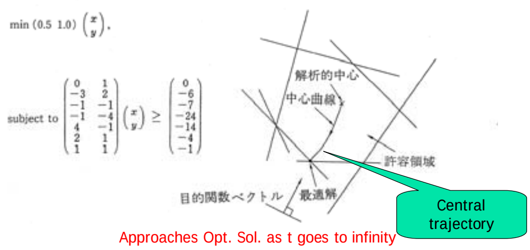
線形計画問題

$$\min \ c^T x, \ \text{s.t.} \ a_i^T x - b_i \geq 0, \ i = 1, \ldots, n$$

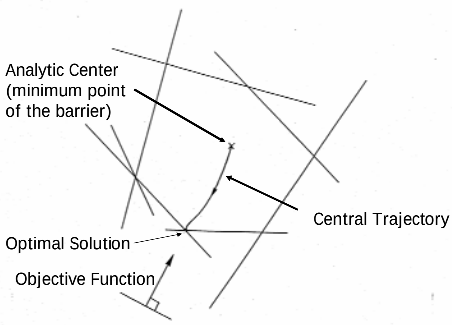
$$\text{Log barrier:} \ -\textstyle\sum_{i=1}^{n} \log(a_i^T x - b_i)$$

第4回応用数学連携フォーラム,
2008/5/30, 東北大学

線形計画問題

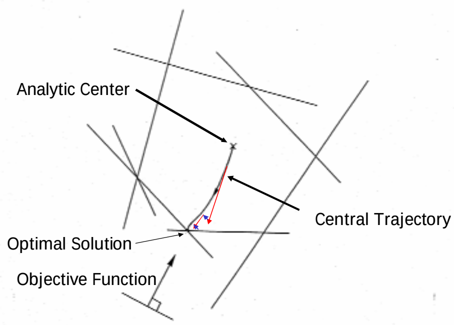


第4回応用数学連携フォーラム,
2008/5/30, 東北大学



第4回応用数学連携フォーラム,
2008/5/30, 東北大学

内点法 ― 中心曲線を近似的に辿る



第4回応用数学連携フォーラム,
2008/5/30, 東北大学

多項式時間内点法

- 中心曲線上の2点 $x(t_1)$ から $x(t_2)$ を辿るのに要する反復回数：

$$O(\sqrt{p}\log(t_2/t_1))$$

- 中心曲線上の目的関数値と最適値の関係：

$$\langle c, x(t)\rangle - \langle c, x_{\mathrm{opt}}\rangle \leq \frac{p}{t}$$

凸錐上の線形計画問題と情報幾何

- Ω に $\psi(x)$ をポテンシャル関数として導入される情報幾何的構造を考える。

ルジャンドル変換

- ルジャンドル変換（勾配写像）

$$s = -\frac{\partial \psi}{\partial x}$$

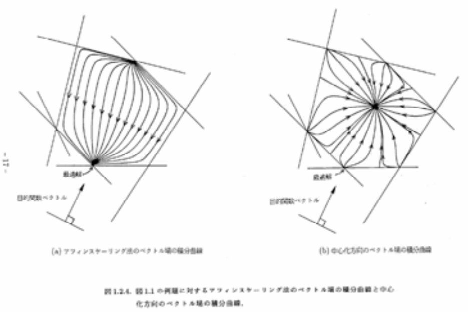
は $\mathrm{int}(\Omega)$ と $\mathrm{int}(\Omega^*)$ の間の大域的写像。

- Ω は s 座標系を用いると Ω^* に見える。

凸錐上の線形計画問題と情報幾何

- リーマン計量: $\psi(x)$ のヘッセ行列
- 互いに双対な2つの接続
  - 接続 ∇ : x 座標系での直線が測地線となる接続。
  - 接続 ∇^*: s 座標系での直線が測地線となる接続。

線形計画の場合 （Log barrier)



これらの曲線は ∇^* 測地線（ ∇ 測地線は上図で任意の直線）

Dikin 楕円体（各点でのリーマン計量による半径1の楕円体）

自己平行部分多様体

- 2つの自己平行部分多様体
  - ∇ 自己平行部分多様体: x 座標系でアフィン空間の開部分集合として書ける多様体。（主問題の許容領域）
  - ∇^*自己平行部分多様体: s 座標系でアフィン空間の開部分集合として書ける多様体。(双対問題の許容領域)

第4回応用数学連携フォーラム, 2008/5/30, 東北大学

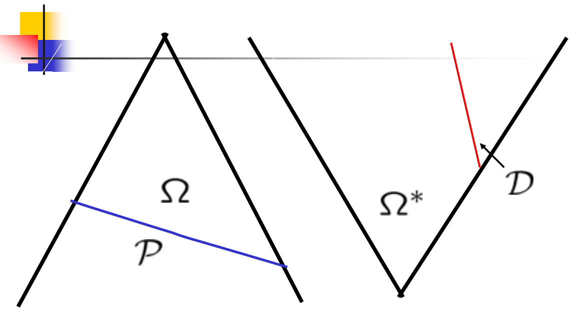
$\mathcal{P}$: Primal feasible region
主問題の許容領域

$\mathcal{D}$: Dual feasible region
双対問題の許容領域

Ω: Primal cone

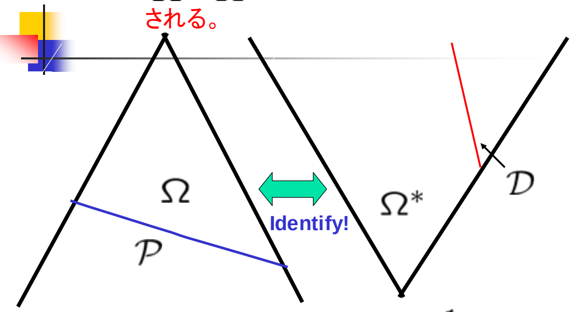
Ω^*: Dual cone of Ω

第4回応用数学連携フォーラム, 2008/5/30, 東北大学



第4回応用数学連携フォーラム, 2008/5/30, 東北大学

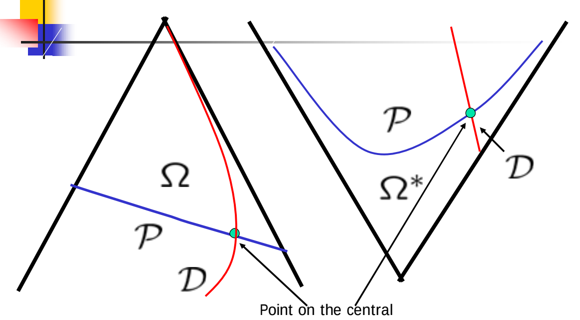
Ω と Ω^*はルジャンドル変換で同一視される。



$$\Omega^* = s(\Omega), \quad \Omega = s^{-1}(\Omega^*)$$

第4回応用数学連携フォーラム, 2008/5/30, 東北大学

$\mathcal{P} \cap \mathcal{D} =$ (t=1 の中心曲線上の点)



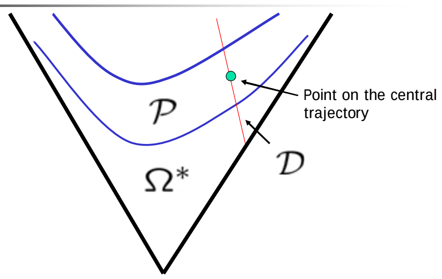
第4回応用数学連携フォーラム, 2008/5/30, 東北大学

曲率と計算複雑度

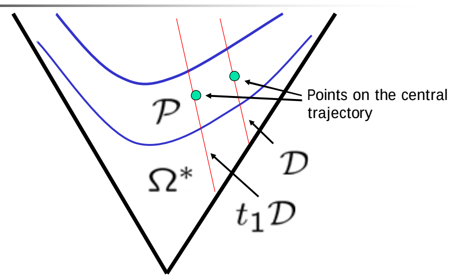
- このような準備の元で、多項式時間内点法の計算複雑度は、中心曲線の埋め込み曲率を用いて精密に表現できる。

第4回応用数学連携フォーラム, 2008/5/30, 東北大学

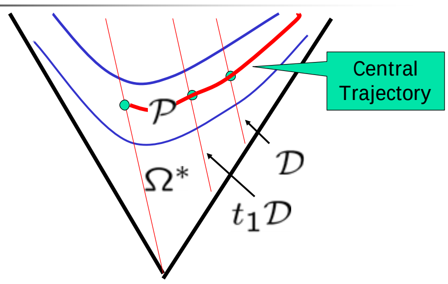
中心曲線



中心曲線



中心曲線



中心曲線

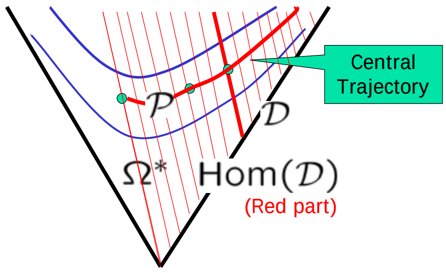
- パラメータ t を含む下記の最適化問題の最適解 $x(t)$ が構成する曲線.

$$\min\ \langle tc,\ x\rangle + \psi(x), \quad x \in \mathcal{P}.$$

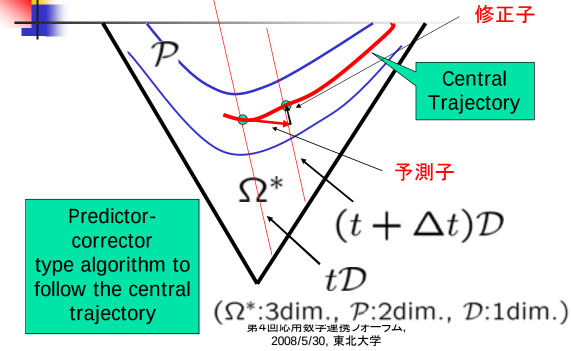
(t: Homotopy parameter) $(0 < t)$

t が∞に近づくと (P) の最適解に収束する。

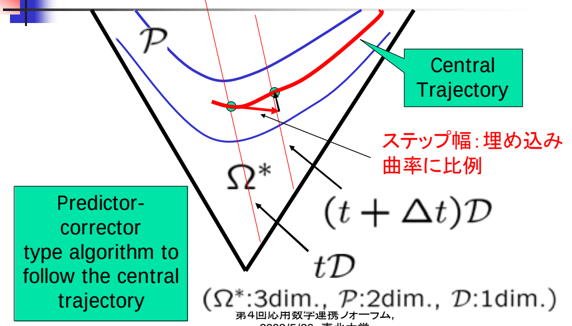
中心曲線＝$\mathcal{P} \cap \mathrm{Hom}(\mathcal{D})$



多項式時間パス追跡法（予測子・修正子法）と埋め込み曲率



多項式時間パス追跡法（予測子・修正子法）と埋め込み曲率



多項式時間内点法

- 中心曲線上の2点$x(t_1)$から$x(t_2)$を辿るのに要する反復回数：

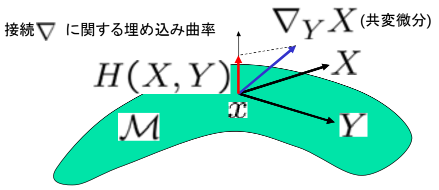
$$O(\sqrt{p}\log(t_2/t_1))$$

- 中心曲線上の目的関数値と最適値の関係：

$$\langle c, x(t)\rangle - \langle c, x_{\mathrm{opt}}\rangle \leq \frac{p}{t}$$

第4回応用数学連携フォーラム,
2008/5/30, 東北大学

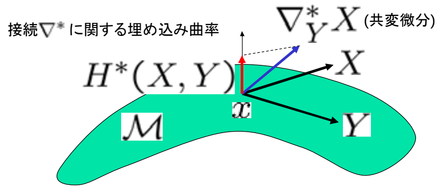
埋め込み曲率



H = 0 on M ⇔ M が∇自己平行多様体

第4回応用数学連携フォーラム,
2008/5/30, 東北大学

埋め込み曲率



H^*= 0 on M ⇔ M が∇*自己平行多様体

第4回応用数学連携フォーラム,
2008/5/30, 東北大学

計算複雑度と曲率

- 曲率積分

$$I_P(t_1,t_2) \equiv \frac{1}{\sqrt{2}}\int_{t_1}^{t_2}\sqrt{\|H^*(\dot{x},\dot{x})\|}\,dt$$

Differential geometric quantity

t_1, t_2: correspond to two points on the central trajectory

第4回応用数学連携フォーラム,
2008/5/30, 東北大学

曲率積分

- Remark

$$\frac{1}{\sqrt{2}}\sqrt{\|H^*(\dot{x},\dot{x})\|} \leq \frac{\sqrt{p}}{t}$$

$$I_P(t_1,t_2) \leq \sqrt{p}\log(t_1/t_2)$$

計算複雑度解析と矛盾のない結果

第4回応用数学連携フォーラム,
2008/5/30, 東北大学

多項式時間内点法

- 中心曲線上の2点 $x(t_1)$ から $x(t_2)$ を辿るのに要する反復回数：

$$O(\sqrt{p}\log(t_2/t_1))$$

- 中心曲線上の目的関数値と最適値の関係：

$$\langle c, x(t)\rangle - \langle c, x_{\mathrm{opt}}\rangle \le \frac{p}{t}$$

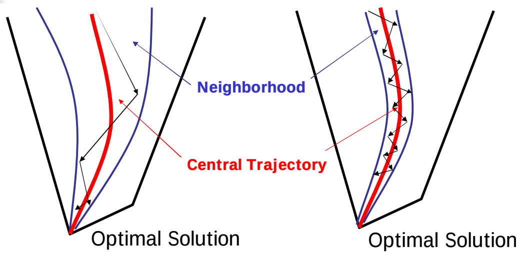
Complexity and Curvature

- 多項式時間内点法の反復回数は、積分を用いて以下のように表現できる。

$$\#\text{ of iterations} \sim \frac{I_P(t_1, t_2)}{\sqrt{\beta}}$$

β: 用いる中心曲線の近傍の大きさ

Path-following Algorithm



β: Large β: Small

- $I_P(t_1, t_2)$ は、$\beta = 1$の場合の反復回数の推定値.
- この積分が大きければ（数百にも上れば）、内点法で効率的に問題を解ける可能性は低くなる。

古典的LPの場合

- 古典的な LP の場合, 中心曲線の全曲率、すなわち I(0, ∞) が存在して、下記の通り評価できる。

LPの係数行列の入力ビット数

I(0,∞) = O($n^{3.5}$*(input size of T))

(independent of c and d; n: # of nonnegative variables).

古典的LPの場合

- 係数行列が0-1行列であるような LP（ネットワーク流問題を含む）の場合,

I(0,∞) = O($n^{4.5}$m)

(n: 非負変数の次元, m: 等式制約の数).

古典的LPの場合

- 情報幾何の立場からは, 本命題は、中心曲線の大域的構造を捉えたものと考えられる。
- アルゴリズムの立場からは、本命題は、LPについて、係数行列にしか計算複雑度が依存しない多項式時間解法が存在することの帰結と考えられる。(係数行列が0－1行列の場合は強多項式解法となる) (Vavasis-Ye, Monteiro-Tsuchiya)

古典的 LP の場合: 計算複雑度と曲率 (主双対内点法)

- 主双対内点法 : 一番良く使われている内点法
主問題と双対問題両方の情報を用いて反復列を作る。
- 主双対内点法についても、中心曲線上の積分を用いた反復回数の表現が知られていた。

$$\#\text{ of iterations} \sim \frac{I_{PD}(t_1,t_2)}{\sqrt{\beta}}$$

β: 用いる中心曲線の近傍の大きさ

古典的 LP の場合: 計算複雑度と曲率 (主双対内点法)

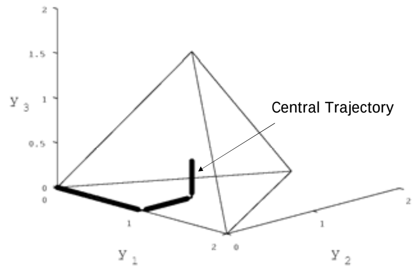
$$I_{PD}(t_1,t_2) \equiv$$

$$\int_{t_1}^{t_2} \sqrt{(\text{primal-dual central trajectory curvature})}\, dt$$

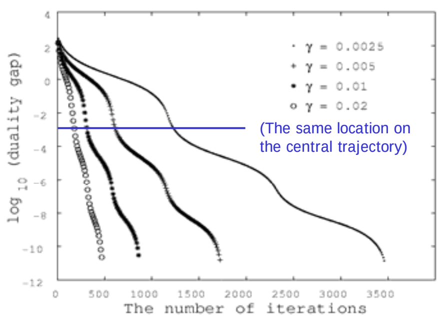
t_1, t_2: correspond to two points on the central trajectory

幾何学的な解釈が分からなかった。

Appendix: Example --- Iteration Complexity and Curvature Integral (primal-dual case)



Summary of the paper

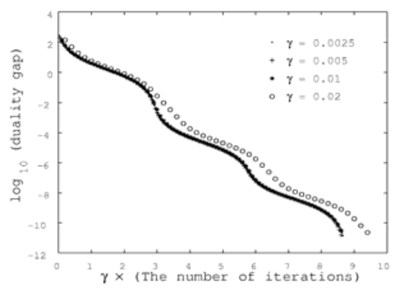


Iterations with different width of neighborhoods

$\sqrt{\beta}$ ($\sqrt{\beta}$ is γ in the figure)

$(\#\text{ of iterations}) \times \gamma \sim I_{PD}(t_1,t_2)$

(用いる近傍の広さには依存しない)



積分が横軸の量として現れている！

ピタゴラス関係と主双対内点法

- 実は、下記のようなピタゴラスの定理が成立する。

(主双対曲率の情報幾何的な意味づけ)

$$(\text{Primal-dual central trajectory curvature})^2 = \frac{\|H^*_{\mathcal{P}}(\dot{x},\dot{x})\|^2}{4} + \frac{\|H_{\mathcal{D}}(\dot{s},\dot{s})\|^2}{4}$$

主内点法と主双対内点法の反復回数の関係

- ピタゴラス関係から, 古典的LPの場合については、主内点法 の方が主双対内点法よりも少ない反復回数で問題をとくことができることがわかる。

$$I_{\mathcal{P}}(t_1, t_2) \le I_{PD}(t_1, t_2)$$

- この関係を半正定値計画に拡張することは興味深い問題である。

文献

- A.Ohara and T.Tsuchiya: An information geometric approach to polynomial-time interior-point algorithms:
complexity bound via curvature integral. Res. Memo No.1055, The Inst. Stat. Math., Dec. 2007 (available at optimization–online)
- 土谷隆、小原敦美: 内点法と情報幾何 --- 計算の複雑さへの微分幾何学的接近法 --- 数学セミナー, 2008年3月号.