最小二乗問題の新解法と逆問題への応用

国立情報学研究所
速水 謙
hayami@nii.ac.jp

数学月間懇話会（第9回）
2013年7月22日
@東京大学数理科学研究科

目次

（1）最小二乗法とは？

連立一次方程式の解法

$$\begin{cases} 2x+3y=5 & \cdots(1) \\ 3x-y=2 & \cdots(2) \end{cases}$$

（ガウス）の消去法

↓

解　$x=1, y=1$

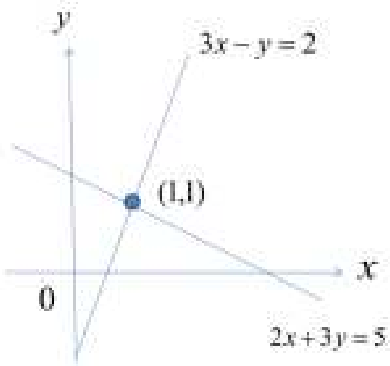
連立一次方程式の幾何学的解釈

$$\begin{cases} 2x+3y=5 & \cdots(1) \\ 3x-y=2 & \cdots(2) \end{cases}$$

（ガウスの）消去法

↓

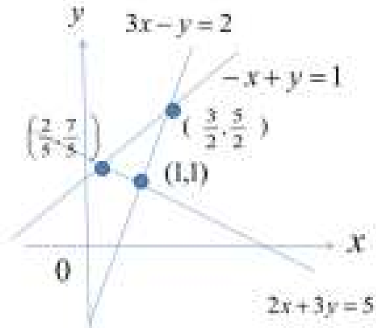
解　$x=1,\ y=1$



未知数の個数よりも式の数が多い場合は？

$$\begin{cases} 2x+3y=5 & \cdots (1) \\ 3x-y=2 & \cdots (2) \\ -x+y=1 & \cdots (3) \end{cases}$$

解は存在しない



最小二乗法

$$\begin{cases} 2x+3y=5 & \cdots (1) \\ 3x-y=2 & \cdots (2) \\ -x+y=1 & \cdots (3) \end{cases}$$

残差の二乗を最小化

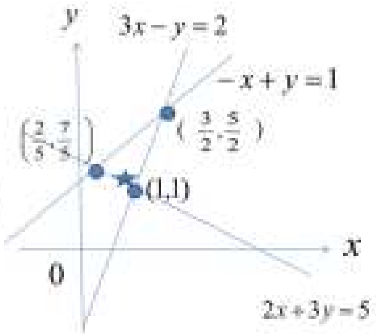
$$(2x+3y-5)^2+(3x-y-2)^2+(-x+y-1)^2$$
$$\Downarrow$$
最小化

最小二乗解

$$\begin{cases} 2x+3y=5 & \cdots(1) \\ 3x-y=2 & \cdots(2) \\ -x+y=1 & \cdots\ (3) \end{cases}$$

最小二乗解★

$$\left(\frac{137}{150},\frac{83}{75}\right) \cong (0.913, 1.11)$$



最小二乗解の求め方

$f(x,y)=(2x+3y-5)^2+(3x-y-2)^2+(-x+y-1)^2$ が最小

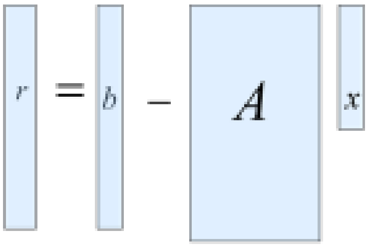
$$\frac{\partial f}{\partial x}=0 \iff 14x+2y=5$$

$$\frac{\partial f}{\partial y}=0 \iff 2x+11y=14$$

最小二乗解

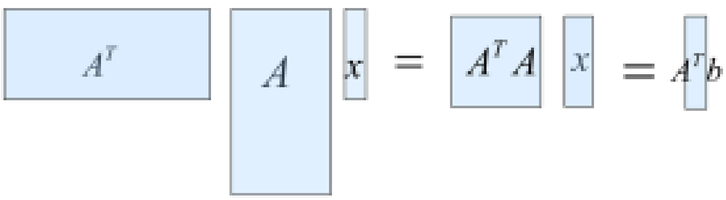
$$(x,y)=\left(\frac{137}{150},\frac{83}{75}\right) \cong (0.913, 1.11)$$

一般には



$\| r \|_2^2 = r^T r$ の最小化（最小二乗問題）

正規方程式（連立一次方程式）



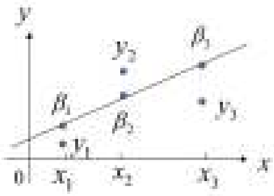
最小二乗解： $x = (A^T A)^{-1} A^T b$

なぜ二乗か？

$\beta_i = ax_i + b$, 観測値 y_i

確率

$$P(0 \le y_i - \beta_i \le \Delta y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y_i-\beta_i)^2}{2\sigma^2}} \Delta y$$

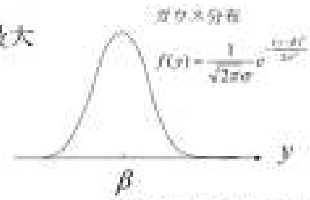


同時確率

$$\left(\frac{\Delta y}{\sqrt{2\pi}\sigma}\right)^3 \prod_{i=1}^{3} e^{-\frac{(y_i-\beta_i)^2}{2\sigma^2}} = \left(\frac{\Delta y}{\sqrt{2\pi}\sigma}\right)^3 e^{-\frac{\sum_{i=1}^{3}(y_i-\beta_i)^2}{2\sigma^2}} \to 最大$$

$$\Updownarrow$$

$$\sum_{i=1}^{3} (y_i-\beta_i)^2 \to 最小$$

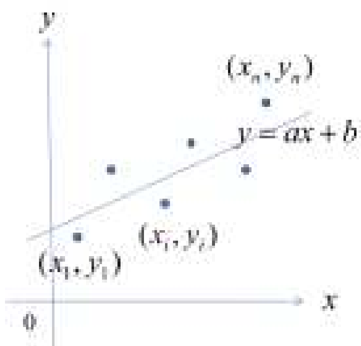


(文献[4] Abdulle, Wanner)

最小二乗法の応用

実験、観測データの解析

$$ax_1 + b = y_1 + \varepsilon_1$$
$$ax_2 + b = y_2 + \varepsilon_2$$
$$\vdots$$
$$ax_n + b = y_n + \varepsilon_n$$



誤差の二乗和を最小にする。

$$\sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} (ax_i + b - y_i)^2 \to 最小$$

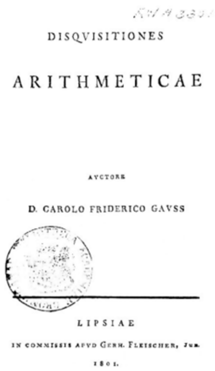
（2）最小二乗法の歴史

Carl Friedrich Gauss
（1777－1855）
ドイツの数学者

「数学は科学の女王であり、
数論は数学の女王である。」

- 「整数論」に関する大著
- 1798年発行
- ガウス24歳

「数論は数学の中でも
最も美しい」


DISQVISITIONES

ARITHMETICAE

AVCTORE

D. CAROLO FRIDERICO GAVSS

LIPSIAE

IN COMMISSIS APVD GERH. FLEISCHER, JUN.

1801.

ガウスのその後
（数理工学、応用数理学者として活躍）

- 天文学：小惑星セレスの軌道の予測
  （最小二乗法の発見と活用）
- 測量学：
  （最小二乗法の活用,微分幾何学の構築）

・電磁気学（ガウスの定理,磁束密度の単位）
・正規分布（ガウス分布,測定の誤差論）
・連立一次方程式の解法
（ガウスの消去法,ガウス・ザイデル法）
・・・

ドイツ旧10マルク紙幣（表）
(1989-2001)

- ガウスの肖像
- 正規（ガウス）分布曲線
- ゲッチンゲン大学の建物

ドイツ旧10マルク紙幣（裏）

- ガウスが発明したヘリオトロープ（測量器具）
- ハノーヴァー地方の三角測量図

ガウスと最小二乗法

- ラプラスが「最小一乗法」を発表（1799）
- ガウスが小惑星Ceresの軌道予測（1801）

・1801年1月1日,イタリアの天文学者Piazziが発見,2月11日まで追跡
・9月にガウスが軌道を計算,予測,12月7日に予測通りに再発見
・その後最小二乗法を用いて軌道を精密に計算

・ルジャンドル,最小二乗法を発表（1805）

・ガウス,最小二乗法の原理を説明（1809）

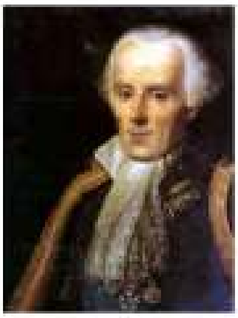
1795年に発見したと主張。「なぜ二乗か？」をガウス分布を用いて説明。

（ガウスの消去法にも言及。）　ルジャンドルの反論

・ガウス,最小二乗法の論文を発表（1823）

ガウス分布に限らず一般の誤差分布に対する最適性を示す。

Pierre-Simon Laplace
(1749-1827)



フランスの数学者
天体力学,確率論,微分方程式論などに貢献
「ラプラス変換」

ガウスによる小惑星Ceresの軌道予測

1801	Longitude	Latitude		Longitude	Latitude
Jan. 1	53° 23' 06.38"	3° 06' 43.16"	23	53° 44' 12.46"	1° 38' 46.78"
2	53° 19' 38.18"	3° 02' 26.46"	28	54° 15' 18.52"	1° 21' 04.92"
3	53° 16' 37.70"	2° 58' 08.04"	30	54° 30' 10.52"	1° 14' 14.24"
4	53° 14' 21.44"	2° 53' 51.98"	31	54° 38' 05.58"	1° 10' 51.02"
10	53° 07' 57.64"	2° 28' 33.64"	Feb. 1	54° 46' 27.14"	1° 07' 34.18"
13	53° 10' 05.60"	2° 16' 46.08"	2	54° 55' 01.52"	1° 04' 18.10"
14	53° 11' 34.20"	2° 12' 34.02"	5	55° 22' 45.20"	0° 54' 34.54"
19	53° 26' 01.98"	1° 55' 37.82"	8	55° 53' 04.32"	0° 45' 08.28"
21	53° 34' 22.68"	1° 46' 13.06"	11	56° 26' 28.20"	0° 35' 55.02"
22	53° 39' 11.58"	1° 42' 28.80"			

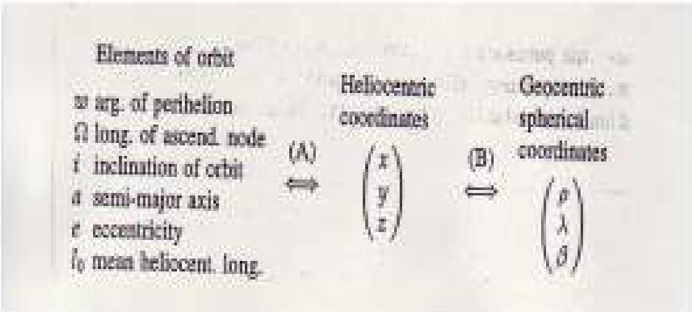
Table 1.1 The observations of Piazzi

Sonnenferne	326° 53' 50"
Ω	81° 1' 44"
Neigung der Bahn	10° 36' 21"
Logarithmus der halben grossen Axe	0.4414902
Excentricität	0.0819603
Epoche: 31 Dec. 1800 mittl. helioc. Länge	77° 34' 29"

Table 1.2 The elements of Ceres (Gauss Dec. 1801)

(文献[4] Abdulle, Wanner)

ガウスの予測法



*Kepler*の法則と非線形方程式の解法

↓

後に非線形最小二乗法(*Gauss* − *Newton*法)を用いる

(文献[4] Abdulle, Wanner)

ケプラーの法則 (Wikipediaより)

ケプラーは、ティコ·ブラーエの観測記録から、太陽に対する火星の運動を推定し、以下のように定式化した。

第1法則（楕円軌道の法則）
感星は、太陽をひとつの焦点とする楕円軌道上を動く。
(1609)

第2法則（面積速度一定の法則）
惑星と太陽とを結ぶ線分が単位時間に描く面積は一定である。
(1609)

第3法則（調和の法則）
惑星の公転周期の2乗は、軌道の長半径の3乗に比例する。
(1619)

Adrien - Marie Legendre
(1752 - 1833)

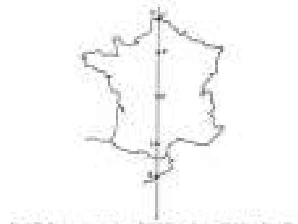
フランスの数学者。（肖像はWikipediaから転載）
統計学、数論、代数学、解析学に貢献

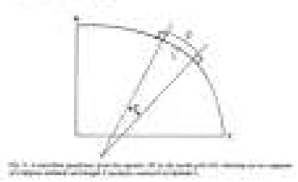
最小二乗法の測地学への応用（ガウス）

子午線上の測地データをもとに地球の楕円度を計算.（1799）

（子午線の四分円の一千万分の一
:1メートルの最初の定義に用いた）

D : *Dunkirk*
P : *Paris*(*Pantheon*)
E : *Evaux*
C : *Carcassonne*
B : *Barcelona*





（文献[2],Stigler）

参考文献

[1] S.G.ギンディキン著,三浦伸夫訳,ガウスが切り開いた道, シュプリンガー・フェアラーク東京,1996.

[2] C.F. Gauss, (A.A. Clarke tr.), *Disquisitiones Aritmeticae*, Yale University Press, 1965. (C.F. ガウス著, 高瀬正仁訳,ガウス整数論,朝倉書店, 1995.)

[3] D. Teets, K. Whitehead, The discovery of Ceres: How Gauss became famous, *Mathematics Magazine,* 72, No.2 (1999), 83-93.

[4] A. Abdulle and G. Wanner, 200 years of least squares method, *Elemente der Mathematik*, 57 (2002), 45-60.

[5] S. Stigler, Gauss and the invention of least squares, *The Annals of Statistics*, 9, No.3 (1981), 465-474.

(3)新しい反復解法

大規模・悪条件最小二乗問題に対する内部反復前処理法の開発

もりくに けいいち

保國惠一(国立情報学研究所)

(総合研究大学院大学 複合科学研究科 情報学専攻 5年一貫制博士課程 2013年3月修了 学位論文の研究)

速水 謙(国立情報学研究所)

行ったこと 1/2

大規模でタチが悪い 最小二乗問題 を効率よく解くために

新しい「前処理法」

を考案した

最小二乗問題とは

工学, 産業, 社会科学に広く現れる重要な基礎問題

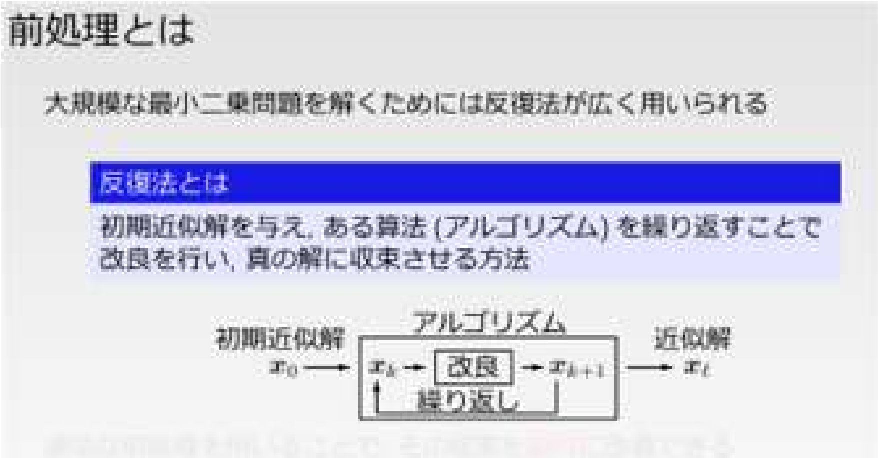
$$\min_{x \in \mathbf{R}^n} \|b - Ax\|_2, \quad A \in \mathbf{R}^{m \times n},\ b \in \mathbf{R}^m$$

方程式の数より未知数の数が多い/少ない場合に
実用的に意味のある答えを推定する問題

具体例
天文補償光学, コンピュータ断層撮影 (CT), 信号処理, 最適化 など

行ったこと 2/2

- 提案法がどんな最小二乗問題でも解けることを証明した
  $\longrightarrow$ 任意の A のサイズおよびランク (階数),
  任意のベクトル b に対して破綻することなく最小二乗解を与える
- ベンチマーク問題 (標準試験問題) に対する数値実験により
  多くの問題に対して従来法よりも優れていることを示した
  $\longrightarrow$ 問題によっては従来法の約 1/40 の計算時間で済む
- 実問題 (画像再構成) に適用し, 有効性を示した
  $\longrightarrow$ 従来法よりも短い時間で誤差をより小さくすることができた



2013.7.22. 最小2乗法_速水 29

反復法とは？

- 直接法
  （ガウスの）消去法のように有限回の計算で
  真の解が求まる方法。

（大規模な問題では計算時間やメモリーが大。）

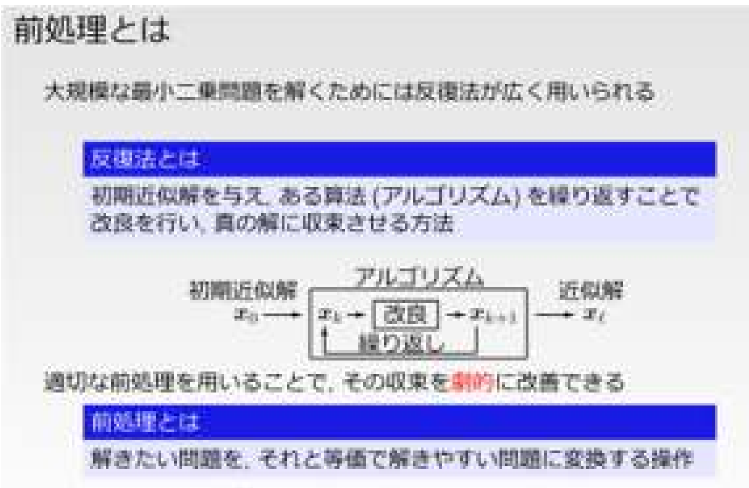
2013.7.22. 最小2乗法_速水 30

反復法

- 簡単な反復法の例（Richardson法）
  連立一次方程式
  $$A\,x = b$$
  に対して適当な初期解x_0を定め
  $$x_{k+1} = x_k + \omega(b - Ax_k) \quad (k = 0,1,2,\cdots)$$
  (ωは適切な定数） により解を改良していく。

（ガウス：「眠っててもできる。」）

・大規模問題に対しても計算量、メモリーが少なくて済む。
・問題は収束性。



前処理とは

$$Ax = b$$

に対して、適切な前処理行列 M を用いて

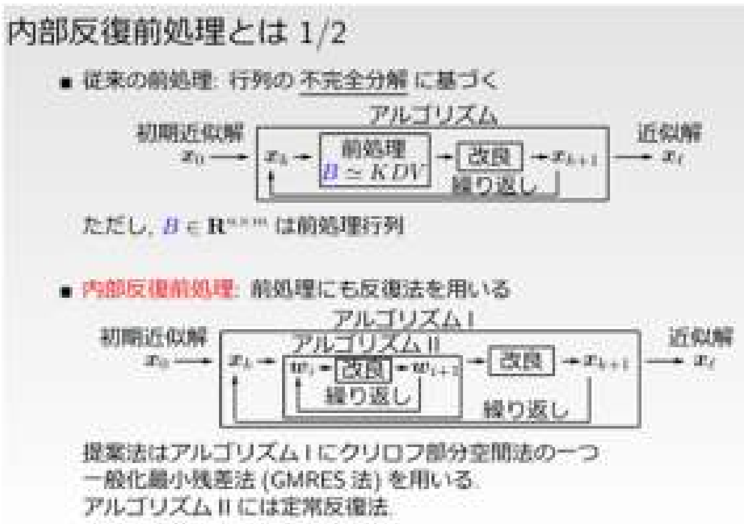
$$MAx = Mb$$

と変形してから、反復法を適用すると速く収束する。

$$M = A^{-1}$$

とおけば一発で解けるが、計算時間がかかる。

$M \sim A^{-1}$ となる安価な Mを求めるのがみそ。



内部反復前処理とは 2/2

元の問題 $\min_{x\in\mathbf{R}^n} \| b - Ax \|_2$

$\Updownarrow \iff \mathcal{R}(B^\top BA) = \mathcal{R}(A) \iff \mathcal{R}(B^\top) = \mathcal{R}(A)$

前処理された問題 $\min_{x\in\mathbf{R}^n} \| Bb - BAx \|_2 = \min_{x\in\mathbf{R}^n} \| Bb - BAx \|_2 \cdots (*)$

提案法は B を陽につくらず，反復法によって陰的に与える

提案法は (*) に対して GMRES 法を用いる

主定理

$\mathcal{R}(B^\top) = \mathcal{R}(A)$ ならば，下記が成り立つ

任意の $b \in \mathbf{R}^m$ および任意の $x_0 \in \mathbf{R}^n$ に対して *BA-GMRES* 法は破綻することなく $\min_{x\in\mathbf{R}^n} \|b - Ax\|_2$ の解を与える

$\Updownarrow$

$\mathcal{N}(A) \cap \mathcal{R}(B) = \{0\}$

従来の問題点 1: 悪条件問題 1/2

従来の手法は悪条件問題に対して計算時間を多く要する

LargeRegFile (回路最適化問題) 2,111,154×801,374, 非ゼロ密度 $2.9 \cdot 10^{-4}$%

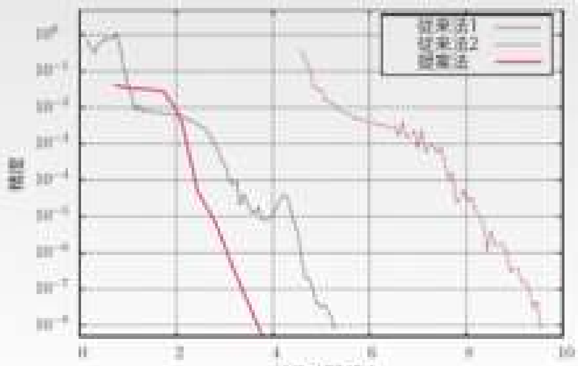


Figure: 計算時間の比較 (大規模悪条件).

従来の問題点 1: 悪条件問題 2/2

従来の手法は悪条件問題に対して計算時間を多く要する

30,000×3,000, 非ゼロ密度 0.1%, $\kappa_2(A) = 1.3 \cdot 10^7$, MATLAB 関数 sprandn

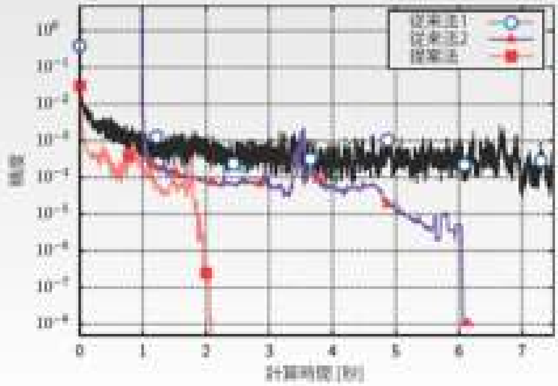


Figure: 計算時間の比較 (中規模悪条件).

従来の問題点 2: 記憶容量

- 従来法は, 前処理行列を陽に構成
  ⟶ 記憶容量を多く要す

$$n\left\{\boxed{A^{\mathsf{T}}A}\right. \simeq \boxed{L}\ \boxed{D}\ \boxed{L^{\mathsf{T}}}$$

  ⟶ 計算時間も要す
- 提案法は, 前処理行列を陽に構成する必要がない
  ⟶ 必要な記憶容量は A の列数のオーダー $\mathcal{O}(n)$

  ⟶ 計算時間も少なく済む

従来の問題点3: ランク落ち 1/2

- 「ランク落ち」という問題に対して有効な従来法はほとんどなかった
  → 「ランク落ち」であると破綻 (ゼロ割り) することも

$$\mathrm{rank}A < \min(m, n) \implies \frac{c}{0}$$

- ランク落ちでも提案法は破綻しないことを証明

主定理

内部反復行列が *semi-convergent* ならば，下記が成り立つ
内部反復前処理付き *BA-GMRES* 法 (提案法) が任意の $b \in \mathbf{R}^m$ および任意の $x_0 \in \mathbf{R}^n$ に対して破綻することなく $\min_{x \in \mathbf{R}^n} \|b - Ax\|_2$ の解を与える

ただし，H が *semi-convergent* である $\overset{\text{def}}{\iff} \lim_{k \to \infty} H^k$ が存在する

従来の問題点3: ランク落ち 2/2

従来法よりも提案法が有効であることを実験的に示した

- size: 21,251×10,144
- 非ゼロ密度 0.1%
- $\kappa_2(A) = 2.9 \cdot 10^6$
- New York Power Authority (NYPA)

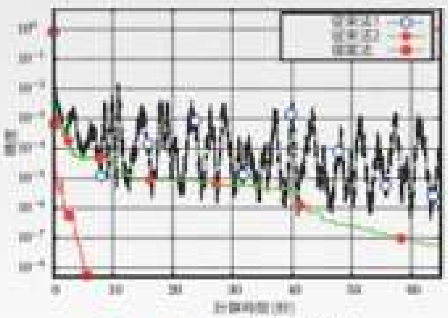
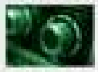


Figure: 計算時間の比較

従来の問題点 4: パラメータチューニング

- 効率よく計算を行うためには，前処理のパラメータの値を適切に設定する必要がある



⟶ 従来法では調整は困難

- 提案法のパラメータの準最適値を自動的に見つける手法を考案した

⟶ 数値実験を通じて，有効性を示した

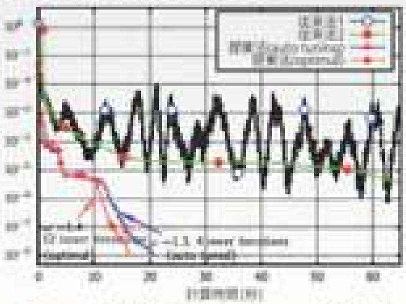


Figure: 自動パラメータチューニング.

参考文献

Morikuni, K., and Hayami, K.,
Inner-iteration Krylov subspace methods for least squares problems,
SIAM Journal on Matrix Analysis and Applications, Vol. 34, No. 1, pp. 1-22, 2013.
http://epubs.siam.org/doi/abs/10.1137/110828472

保國惠一氏がこの論文で SIAM Student Prizeを受賞(2013年7月)
http://siam.org/prizes/sponsored/student_paper.php

(SIAM: Society for Industrial and Applied Mathematics)

関連文献：

Morikuni, K. and Hayami, K.,
Convergence of inner-iteration GMRES methods for least squares problems,
NII Technical Reports, National Institute of Informatics, Tokyo,
NII-2012-005E, pp. 1-9, August 2012.
http://www.nii.ac.jp/TechReports/12-005E.html