大規模連立一次方程式に対する高並列前処理技術について

今倉 暁
筑波大学 計算科学研究センター

共同研究者
櫻井鉄也(筑波大学)，住吉光介(沼津高専)，松古栄夫(KEK)

本日のトピック

- 大規模連立一次方程式

$$A\boldsymbol{x} = \boldsymbol{b}, \quad A \in \mathbb{R}^{n \times n}, \quad \boldsymbol{x}, \boldsymbol{b} \in \mathbb{R}^n$$

のための**(前処理付き)Krylov部分空間法**の概略について紹介する.

- 高並列性を考慮した前処理として，反復法を用いた**重み付き定常反復型前処理**を導入し，その**パラメータを最適化手法**を提案・有効性の検証を行う.

第1トピック

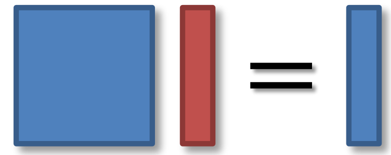
「前処理付きKrylov部分空間法の概略について」

連立一次方程式とその数値解法

連立一次方程式

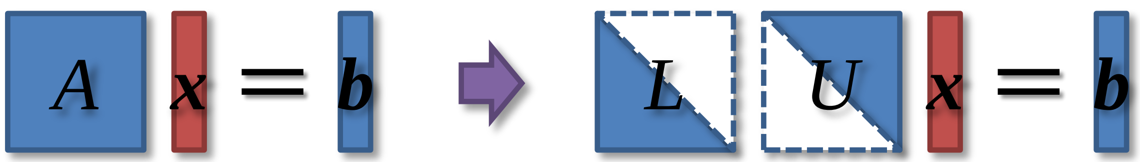
>> 入力 : $A \in \mathbb{R}^{n\times n}, \boldsymbol{b} \in \mathbb{R}^{n}$ 出力 : $\boldsymbol{x} \in \mathbb{R}^{n}$

$$A\boldsymbol{x} = \boldsymbol{b}$$

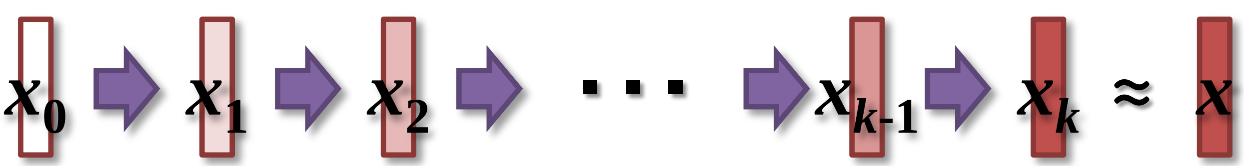


連立一次方程式の数値解法

>> 直接法 : 有限回の演算で解を計算



>> 反復法 : 解に収束する近似解列を逐次的に生成



連立一次方程式とその数値解法

連立一次方程式に対する反復法

>> 定常反復法

ex) Jacobi法, Gauss-Seidel法, SOR法, ...

$$Ax = b \Leftrightarrow x = f(x)$$

$$x_{k+1} = f(x_k) = M^{-1}Nx_k + M^{-1}b \quad (A = M - N)$$

>> Krylov部分空間法

ex) 共役勾配法(CG法), GMRES法, BiCGSTAB法, ...

$$x_k = x_0 + z_k$$

$$z_k \in \mathcal{K}_k(A, r_0) := \mathrm{span}\{r_0, Ar_0, \ldots, A^{k-1}r_0\}$$

>> その他

ex) 幾何的/代数的マルチグリッド法, ...

Krylov部分空間法の概略

Krylov部分空間法

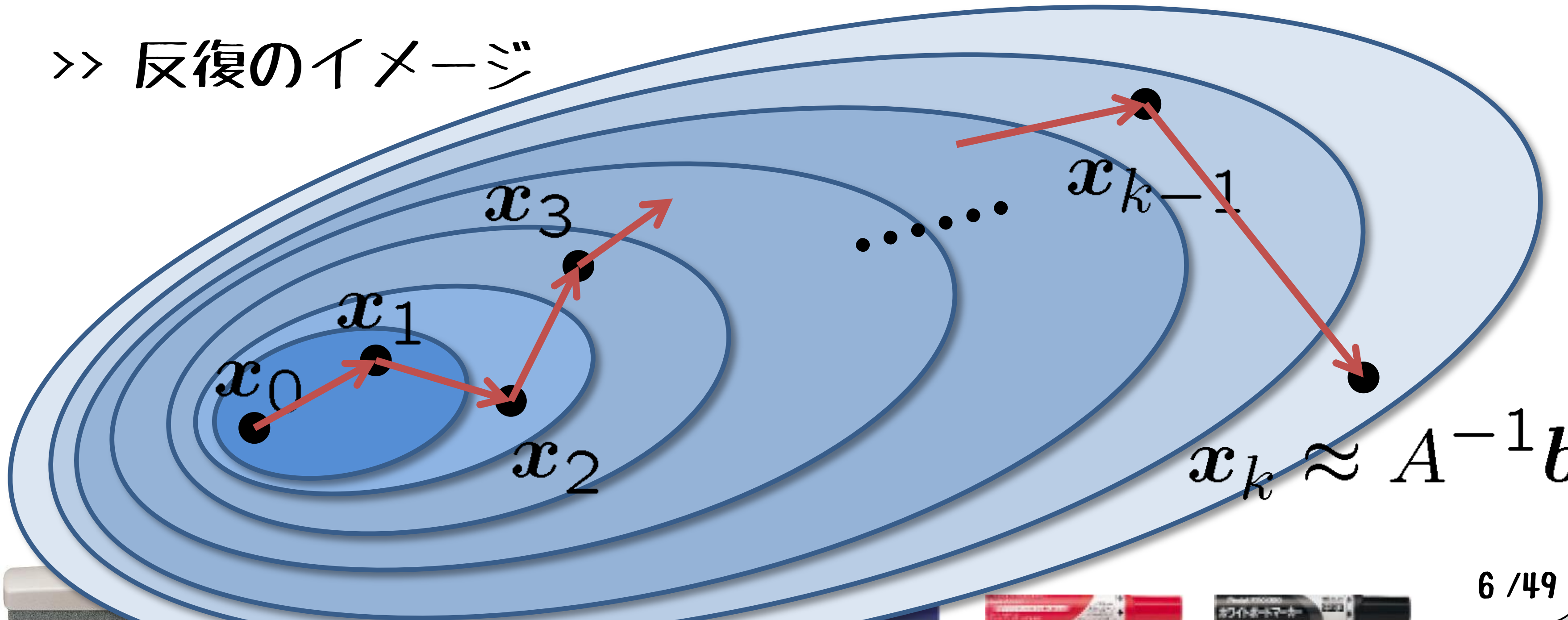
>> 定義：Krylov部分空間に基づく射影法

>> 入力：初期近似解 x_0，初期残差 $r_0 := b - Ax_0$

>> 出力：

$$x_k = x_0 + z_k, \quad z_k \in \mathcal{K}_k(A, r_0)$$
$$:= \mathrm{span}\{r_0, Ar_0, \ldots, A^{k-1}r_0\}$$

>> 反復のイメージ



Krylov部分空間法の概略

- Krylov部分空間法

  >> アルゴリズム

Input

$$x_0, r_0 := b - Ax_0$$

Krylov basis

$$V_k \leftarrow \mathscr{K}_k(A, r_0) := \mathrm{span}\{r_0, Ar_0, \ldots, A^{k-1}r_0\}$$

Approximate solution

$$x_k = x_0 + V_k s_k, \quad s_k \in \mathbb{R}^k$$

Output

$$x_k, r_k = b - Ax_k$$

Krylov部分空間法の概略

Krylov部分空間法

>> アルゴリズム

Input

$$x_0, r_0$$

Krylov basis

Krylov部分空間（の基底ベクトル）を生成

収束するまで反復

Approximate solution

何らかの条件に基づき近似解を構築

Output

$$x_k, r_k$$

Krylov basis

- Lanczos過程
- Arnoldi過程
- 双Lanczos過程

Approximate solution

- Ritz-Galerkin条件
- Petrov-Galerkin条件
- 最小残差条件

Krylov部分空間法の前処理の概略

Krylov部分空間法の前処理

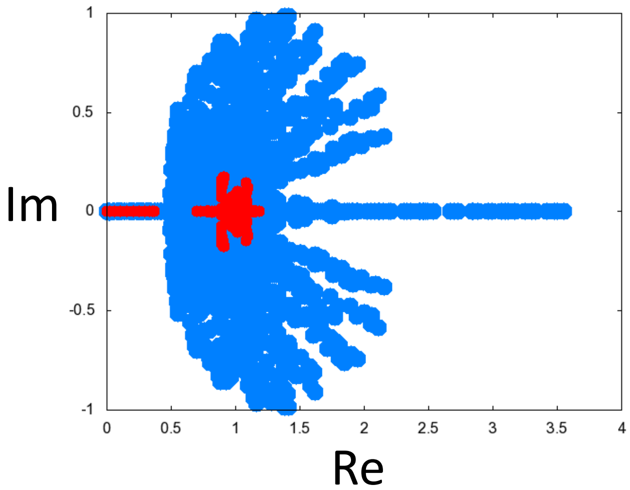
>> Krylov部分空間法の収束性

一般に，係数行列の**固有値分布**に依存

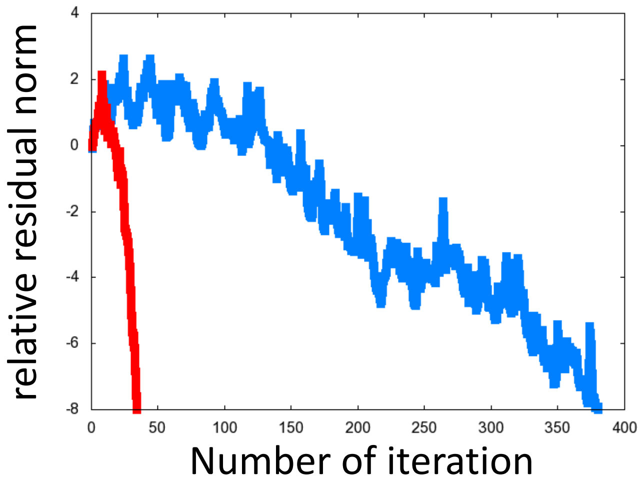
>> 前処理（収束性の改善を図る）

$$Ax = b \quad \Leftrightarrow \quad AK^{-1}y = b, \quad x = K^{-1}y$$

係数行列の固有値分布



収束履歴

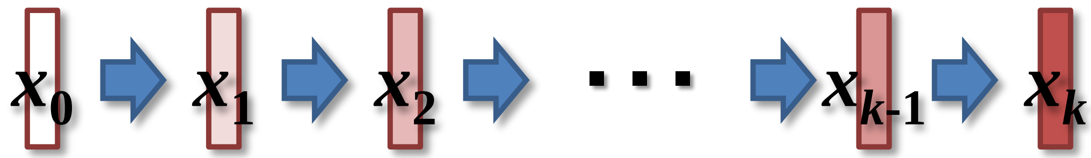


Krylov部分空間法の前処理の概略

Krylov部分空間法に対する前処理の適用

Krylov部分空間法

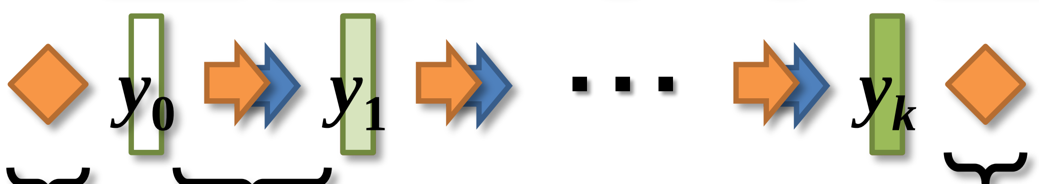
$$A\,x = b$$



1反復の計算（Aの行列ベクトル積 + 内積, AXPY）

前処理付きKrylov部分空間法

$$A\,K^{-1}\,y = b, \quad x = K^{-1}\,y$$



事前準備

1反復の計算

後処理

（Aの行列ベクトル積 + Kの方程式の求解 + 内積, AXPY）

Krylov部分空間法の前処理の概略

理想的な前処理

>> **高速**

-- 各反復でのKの方程式$Kw = v$が高速に解けること

-- 事前準備・後処理のコストが小さいこと

>> **高精度**

-- 収束性を大きく改善出来ること（$AK^{-1} \approx I$）

>> **高並列**（大規模並列化を行う場合）

-- 事前準備・後処理および前処理の適用の際の並列化効率が高いこと

Krylov部分空間法の前処理の概略

前処理の分類

>> 直接法に基づく前処理

-- 各反復でのKの方程式$Kw = v$が容易に解けるよう，
K またはK^{-1}を事前に計算する.

◎ 反復あたりの計算コスト小

▲ 事前準備コスト大，行列の構造が陽的に必要

ex) ILU前処理，近似逆行列前処理，多項式前処理，...

>> 反復法に基づく前処理

-- 各反復で$Kw = v\,(K \approx A)$を解く代わりに，
$Aw = v$ を反復法で近似的に計算する.

◎ 事前準備コスト小or無し，行列の構造は不要な場合も

▲ 多くのパラメータを持つ

ex) 定常反復法，GMRES法，...

第2トピック

「重み付き定常反復型前処理に対するパラメータ最適化手法の提案」

第2トピックの目次

- 反復法に基づく前処理

- 重み付き定常反復型解法
  - >> 定常反復型解法の概略
  - >> 重み付き定常反復型解法の導入

- パラメータ最適化手法
  - >> 重み付き定常反復型解法の収束性解析
  - >> 重み付き定常反復型前処理に対するパラメータ最適化手法の提案

- 数値実験

反復法に基づく前処理

反復法に基づく前処理の概略（再掲）

>> 各反復での計算

-- 各反復で $Kw = v\ (K \approx A)$を解く代わりに，$Aw = v$ を反復法で近似的に計算する.

>> メリット・デメリット

◎ 事前準備コスト小or無し，行列の構造は不要な場合も

▲ 多くのパラメータを持つ

>> 使用される反復法の例

-- 定常反復法: Jacobi法，Gauss-Seidel法，SOR法，...

-- Krylov部分空間法: CGNE法，GMRES法，...

反復法に基づく前処理

- 前処理で用いる理想的な反復法

>> **高速**

-- 反復法として高い収束性を持つこと
(低精度までの収束性が良ければOK)

-- 反復あたりの計算コストが小さいこと

>> **高安定**

-- 残差が大きく振動しないこと
(少ない反復回数で停止するため)

>> **高並列性** (大規模並列化を行う場合)

-- 反復あたりの計算が高い並列化効率を持つこと

第2トピックの目次

- 反復法に基づく前処理
- 重み付き定常反復型解法
  - >> 定常反復型解法の概略
  - >> 重み付き定常反復型解法の導入
- パラメータ最適化手法
  - >> 重み付き定常反復型解法の収束性解析
  - >> 重み付き定常反復型前処理に対するパラメータ最適化手法の提案
- 数値実験

重み付き定常反復型解法

定常反復型解法

>> 基本的アイディア

-- 線形方程式と同値な不動点方程式

$$x = f(x) \Leftrightarrow Ax = b$$

の不動点 $x \left(= A^{-1}b\right)$ を以下の反復により求める.

$$x_{k+1} = f(x_k), \quad k = 0, 1, 2, \ldots$$

-- 一般には, $\boxed{A = M - N}$ とし，ベクトル値関数 f を

$$f(x) := M^{-1}Nx + M^{-1}b$$

と置くことで，反復式は以下のように書かれる.

$$\boxed{x_{k+1} = M^{-1}Nx_k + M^{-1}b, \quad k = 0, 1, 2, \ldots}$$

重み付き定常反復型解法

定常反復型解法

>> 収束性

-- 反復行列 $G := M^{-1}N$ のスペクトル半径 $\rho(G) := \max_{i=1,\ldots,n} |\lambda_i(G)|$ に依存する.

>> 収束定理

-- 定常反復型解法が任意の x_0, b に対して収束するための必要十分条件は，以下の不等式を満たすことである.

$$\rho(G) < 1$$

-- また，その収束性は

$$\lim_{k\to\infty}\left(\max_{x_0\in\mathbb{R}^n}\frac{\|e_k\|}{\|e_0\|}\right)^{1/k} = \rho(G)$$

のように書ける.

重み付き定常反復型解法

理想的な定常反復型解法

>> **高速**

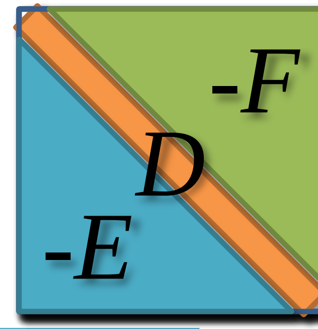
-- 高い収束性を持つこと（$\rho(G)$ が小さいこと）

-- 反復当たりの計算コストが小さいこと

>> **高並列性**（大規模並列化を行う場合）

-- 反復あたりの計算が高い並列化効率を持つこと

重み付き定常反復型解法



定常反復型解法の例

>> 標準的な定常反復型解法（定常反復法）

解法	M	N
Jacobi法	D	$E+F$
Gauss-Seidel法	$D-E$	F
SOR法	$(D-\omega E)/\omega$	$\{(1-\omega)D+\omega F\}/\omega$

>> 標準的でない定常反復型解法

解法	M	N
ILU型	$A \approx \tilde{L}\tilde{U}$	$M-A$
近似逆行列型	$\min \|\|I-AM\|\|_{\mathrm{F}}$	$M-A$

重み付き定常反復型解法

重み付き定常反復型解法の導入

>> 基本的アイディア：重みパラメータを用いて収束を改善

-- 定常反復型解法の反復式

$$x_{k+1} = M^{-1}Nx_k + M^{-1}b$$

を重みパラメータ $\omega \in \mathbb{R}$ を用い,

$$\begin{aligned} x_{k+1} &= \omega\left(M^{-1}Nx_k + M^{-1}b\right) + (1-\omega)x_k \\ &= x_k + \omega M^{-1}(b - Ax_k) \end{aligned}$$

のように変更することで，収束を加速する

>> 定常反復型解法の拡張

-- パラメータ $\omega = 1$ の時，定常反復型解法に帰着

第2トピックの目次

- 反復法に基づく前処理
- 重み付き定常反復型解法
  - >> 定常反復型解法の概略
  - >> 重み付き定常反復型解法の導入
- パラメータ最適化手法
  - >> 重み付き定常反復型解法の収束性解析
  - >> 重み付き定常反復型前処理に対するパラメータ最適化手法の提案
- 数値実験

重み付き定常反復型解法

- 重み付き定常反復型解法の収束性

>> 行列の分割

-- 係数行列を

$$A = M_\omega - N_\omega, \quad \begin{aligned} M_\omega &= M/\omega, \\ N_\omega &= \{N + (1-\omega)A\}/\omega \end{aligned}$$

と分割した定常反復型解法とみることが出来る.

>> 収束性

反復行列 $G_\omega := M_\omega^{-1} N_\omega$ のスペクトル半径 $\rho(G_\omega)$ に依存

重み付き定常反復型解法

重み付き定常反復型解法の収束性

>> 重みパラメータの最適値

Proposition 1: Optimal parameter

$C(\gamma,\rho)$を複素平面上の中心$\gamma \in \mathbb{R}$半径$\rho \in \mathbb{R}$の円で囲まれた領域あるとし，$\gamma^*, \rho^* \in \mathbb{R}$を

$$\gamma^*, \rho^* := \operatorname*{arg\,min}_{\gamma,\rho} \left|\frac{\rho}{\gamma}\right|,$$

$$s.t. \quad \lambda_i(M^{-1}A) \in C(\gamma,\rho), \quad i = 1, 2, \ldots, n$$

と定義する．

この時，以下が成り立つ．

$$\boxed{\operatorname*{arg\,min}_{\omega\in\mathbb{R}} \rho(G_\omega) = \frac{1}{\gamma^*}, \quad \min_{\omega\in\mathbb{R}} \rho(G_\omega) = \left|\frac{\rho^*}{\gamma^*}\right|}$$

重み付き定常反復型解法

重み付き定常反復型解法の収束性

>> 重みパラメータの最適値

Proposition 2: Convergence condition

重み付き定常反復型解法における ω の最適値を $\omega^* := 1/\gamma^*$ とする．この時，

$$\rho(G_{\omega^*}) = \left|\frac{\rho^*}{\gamma^*}\right| < 1$$

となるための，必要十分条件は，

$$\mathcal{R}e\left(\lambda_i(M^{-1}A)\right) > 0, \quad i = 1, 2, \ldots, n$$

または，

$$\mathcal{R}e\left(\lambda_i(M^{-1}A)\right) < 0, \quad i = 1, 2, \ldots, n$$

が満たされることである．

パラメータ最適化手法

重み付き定常反復型前処理に対するパラメータ最適化

>> 重み付き定常反復型前処理の収束性（再掲）

Remark of Propositions 1 and 2

$$\mathcal{R}e\left(\lambda_i(M^{-1}A)\right) > 0, \quad i = 1, 2, \ldots, n$$

または

$$\mathcal{R}e\left(\lambda_i(M^{-1}A)\right) < 0, \quad i = 1, 2, \ldots, n$$

$$\omega^* = \frac{1}{\gamma^*}, \quad \rho(G_{\omega^*}) = \left|\frac{\rho^*}{\gamma^*}\right| < 1$$

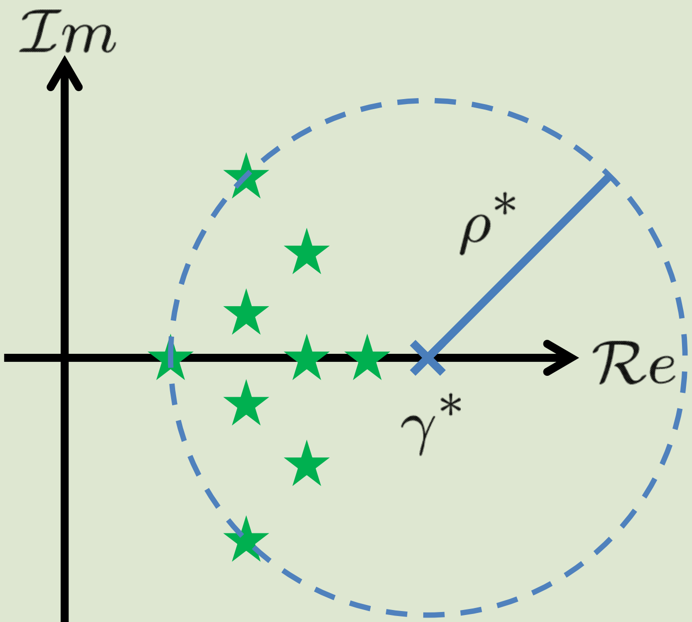
$$\left(\begin{aligned} &\gamma^*, \rho^* := \arg\min_{\gamma,\rho} \left|\frac{\rho}{\gamma}\right|, \\ &s.t. \quad \lambda_i(D^{-1}A) \in C(\gamma, \rho), \quad i = 1, 2, \ldots, n \end{aligned} \right)$$

パラメータ最適化手法

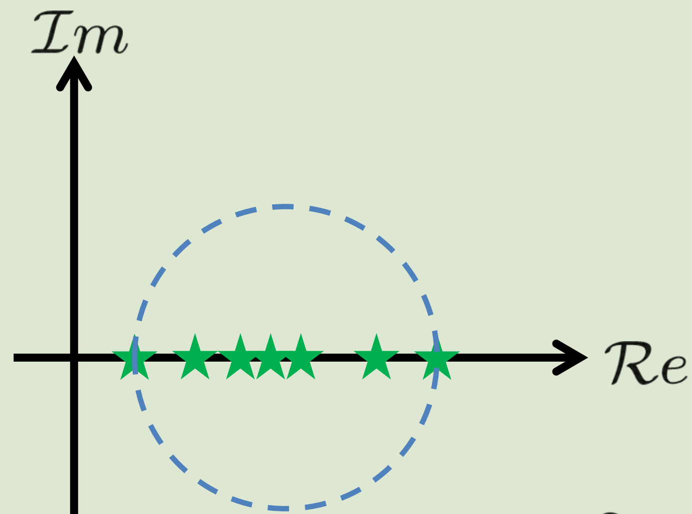
重み付き定常反復型前処理に対するパラメータ最適化

>> 重み付き定常反復型前処理の収束性（再掲）

Rough sketch



$\omega^* = 1/\gamma^*, \rho(G_{\omega^*}) = \rho^*/\gamma^*$



$$\omega^* = \frac{2}{\lambda_{\max} + \lambda_{\min}}$$

$$\rho(G_{\omega^*}) = \frac{\lambda_{\max} - \lambda_{\min}}{\lambda_{\max} + \lambda_{\min}}$$

固有値分布から重みパラメータは最適化可能

パラメータ最適化手法

- 重み付き定常反復型前処理に対するパラメータ最適化

>> 基本的アイディア：オフラインチューニング

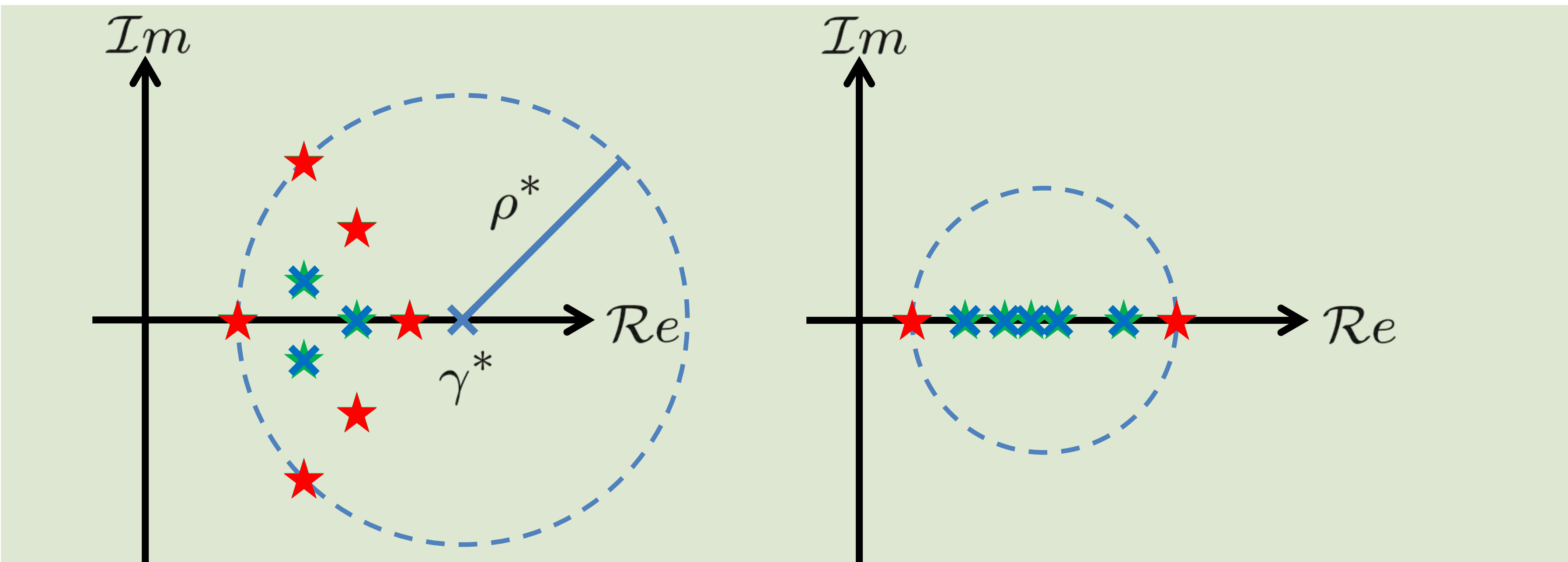
Input	・初期近似解 x_0 ・行列の分割 $A = M - N$
Optimization I	・$M^{-1}A$ の固有値分布を(近似的に)計算
Optimization II	・計算された固有値分布を元に， $\omega^* = 1/\gamma^*, \rho(G_{\omega^*}) = \|\rho^*/\gamma^*\|$ を計算
Preconditioned Krylov	・近似最適パラメータを用いた重み付き定常反復型前処理付きKrylov部分空間法を適用

パラメータ最適化手法

最適化 I

>> 最適化に必要な固有値

Rough sketch



外部固有値のみから重みパラメータは最適化可能

外部固有値の近似値を用いて最適化可能

パラメータ最適化手法

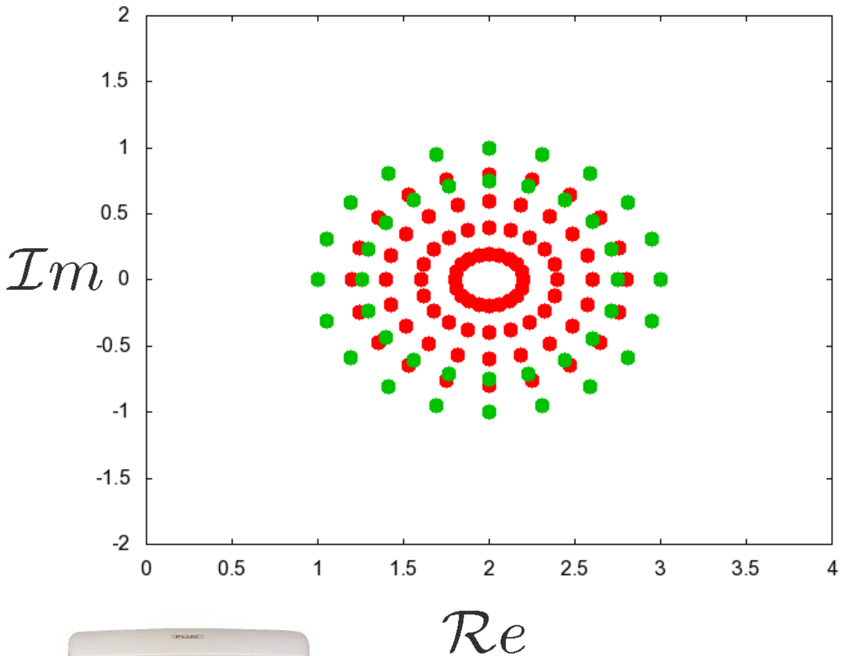
最適化 I

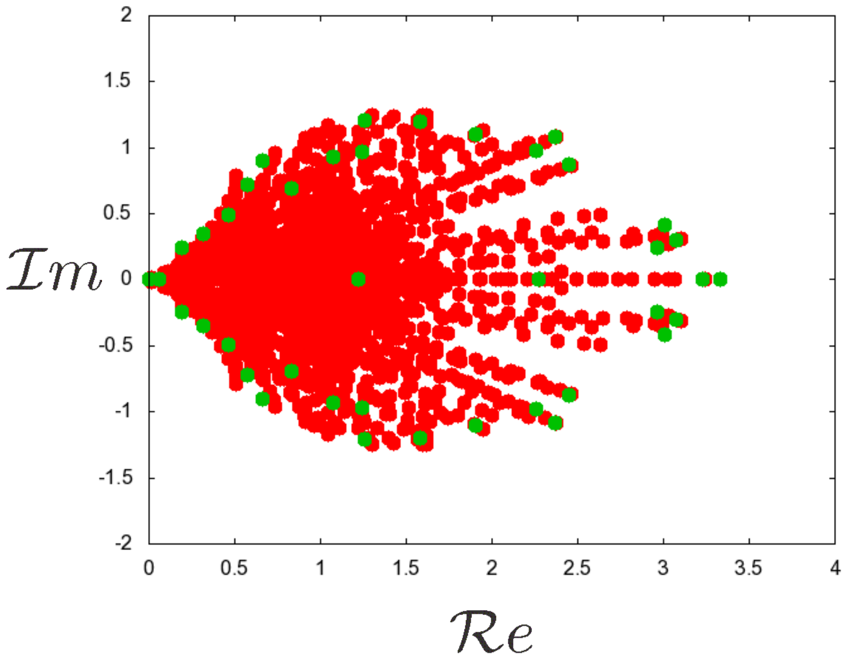
>> 外部固有値の近似

-- 反復法である「Arnoldi法」を利用

>> Arnoldi法で計算される近似固有値

-- 真の固有値，近似固有値 （40反復目）





パラメータ最適化手法

- 重み付き定常反復型前処理に対するパラメータ最適化

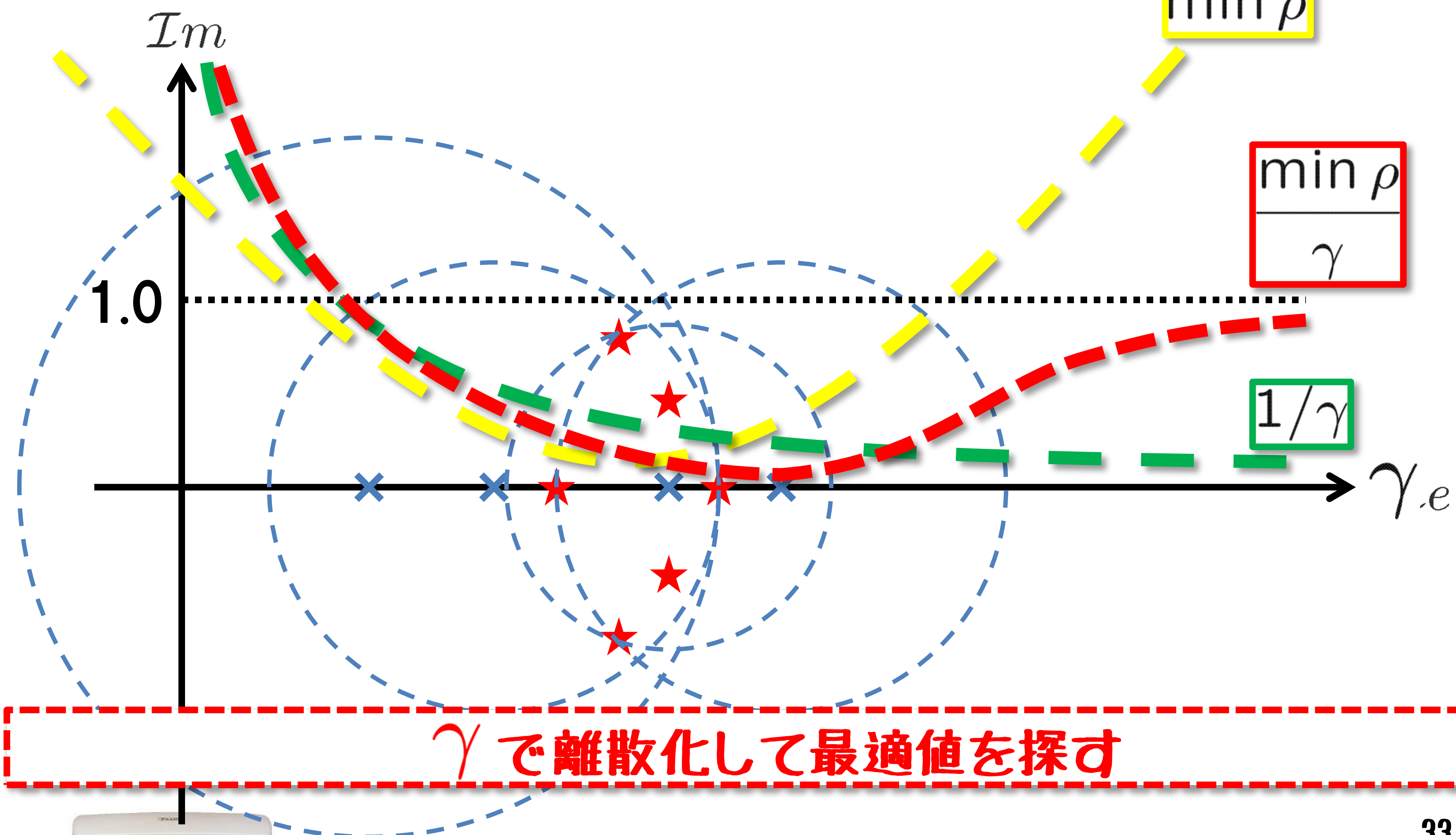
>> 基本的アイディア：オフラインチューニング

Input	・初期近似解 x_0 ・行列の分割 $A = M - N$
Optimization I	・Arnoldi法により $M^{-1}A$ の外部固有値を近似
Optimization II	・計算された固有値分布を元に， $\omega^* = 1/\gamma^*, \rho(G_{\omega^*}) = \lvert\rho^*/\gamma^*\rvert$ を計算
Preconditioned Krylov	・近似最適パラメータを用いた重み付き定常反復型前処理付きKrylov部分空間法を適用

パラメータ最適化手法

最適化 II

>> 最適な円の計算



γ で離散化して最適値を探す

パラメータ最適化手法

- 重み付き定常反復型前処理に対するパラメータ最適化

  >> 基本的アイディア：オフラインチューニング

Input
- 初期近似解 x_0
- 行列の分割 $A = M - N$

Optimization I
- Arnoldi法により$M^{-1}A$の外部固有値を近似

Optimization II
- γ で離散化して最適な円 $C(\gamma^*, \rho^*)$ を書く
  $(\omega^* = 1/\gamma^*, \rho(G_{\omega^*}) = |\rho^*/\gamma^*|)$

Preconditioned Krylov
- 近似最適パラメータを用いた重み付き定常反復型前処理付きKrylov部分空間法を適用

パラメータ最適化手法

- 重み付き定常反復型前処理に対するパラメータ最適化

>> 具体的なアルゴリズム

Optimization I

1: For l = 1, 2, …, lmax
2: Operate lth step of the Arnoldi method and get l approx. extreme eigenvalues
3: Compute $\tilde{\omega}_l$ and $\rho(G_{\tilde{\omega}_l})$ from computed eigenvalues
4: IF $|\tilde{\omega}_l - \tilde{\omega}_{l-1}/\tilde{\omega}_l| \leq \epsilon$ then exit
5: End For
6: Set $\tilde{\omega}_i = \tilde{\omega}_l, \rho(G_{\tilde{\omega}_i}) = \rho(G_{\tilde{\omega}_l})$

>> 実用上のアイディア

-- Arnoldi法の反復はωの相対誤差で停止

-- 予めMの候補を設定し，M毎にωを最適化することで，Mの選択を行うことも可能．

第2トピックの目次

- 反復法に基づく前処理
- 重み付き定常反復型解法
  - >> 定常反復型解法の概略
  - >> 重み付き定常反復型解法の導入
- パラメータ最適化手法
  - >> 重み付き定常反復型解法の収束性解析
  - >> 重み付き定常反復型前処理に対する パラメータ最適化手法の提案
- 数値実験

数値実験

数値実験I

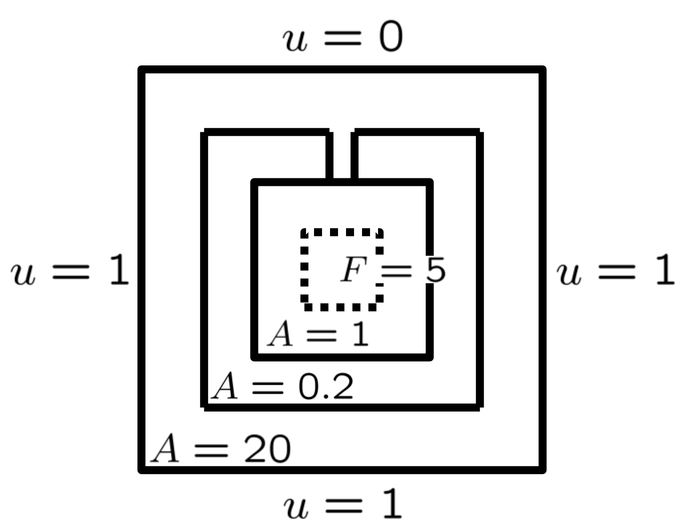
>> 目的：最適化手法の有効性を検証する

>> モデル問題

-- 単位正方領域 $(x, y) \in (0,1) \times (0,1)$ 上の偏微分方程式

$$-(A(x,y)u_x)_x - (A(x,y)u_y)_y + \alpha \exp(2(x^2+y^2))u_x = F(x,y)$$

離散化して得られる2500次の実非対称連立一次方程式



数値実験

数値実験I

>> 比較解法

-- Jacobi法, 重み付きJacobi法,

-- Gauss-Seidel法, 重み付きGauss-Seidel法

>> 各種パラメータ

-- 初期近似解 : $x_0 = [0, 0, \ldots, 0]^\top$

-- Arnoldi法の最小/最大反復回数 : 10回/20回

-- Arnoldi法の停止条件 : 10^-2

>> 計算機環境

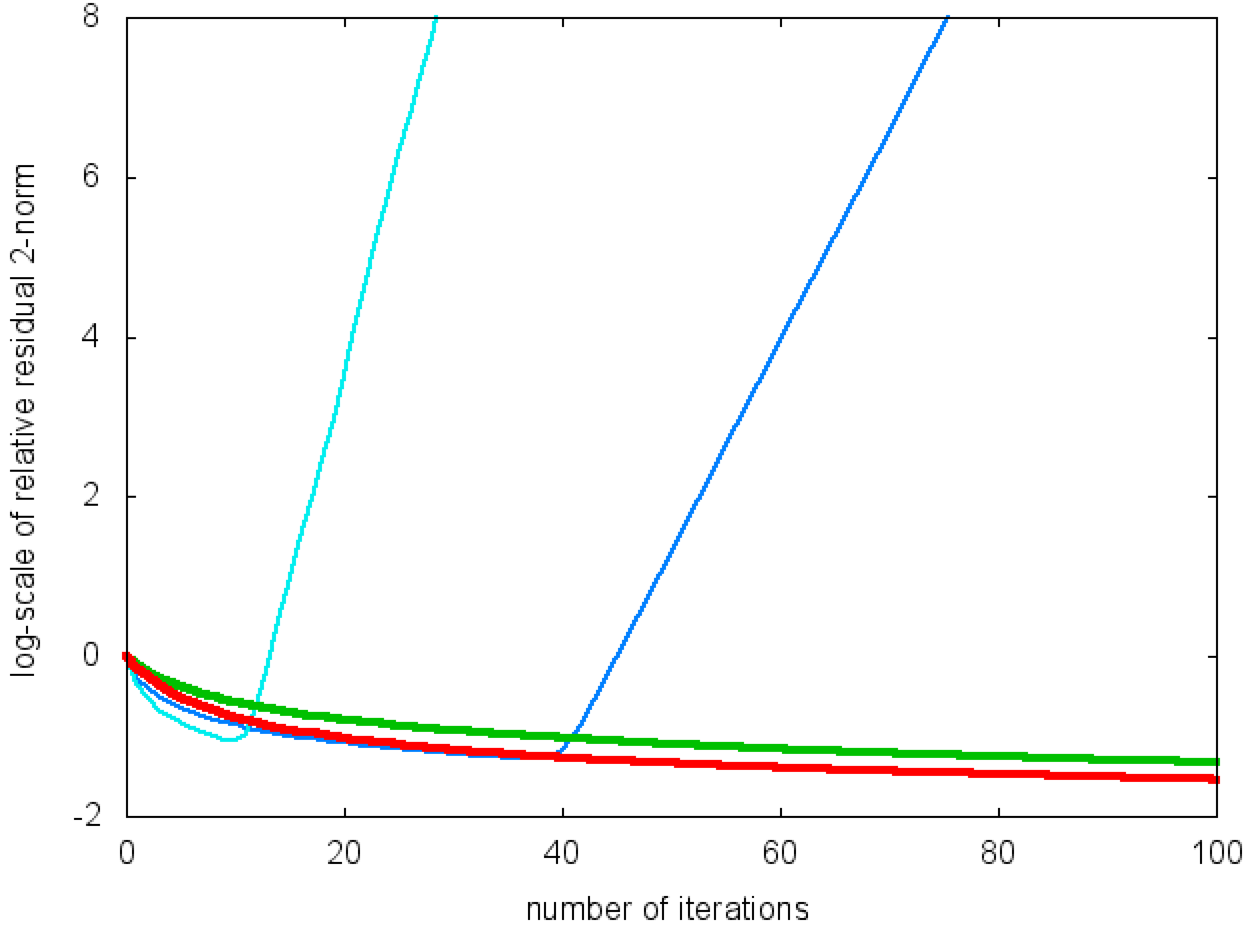
-- OS: CentOS, CPU: Intel Xeon X5550 (2.67GHz),
Memory: 48GB

-- Compiler: GNU Fortran ver.4.1.2,
Compile option: -O3

数値実験

実験結果I

>> 収束履歴（J, GS, W-J, W-GS）



数値実験

数値実験II

>> 目的: 前処理としての性能を評価する

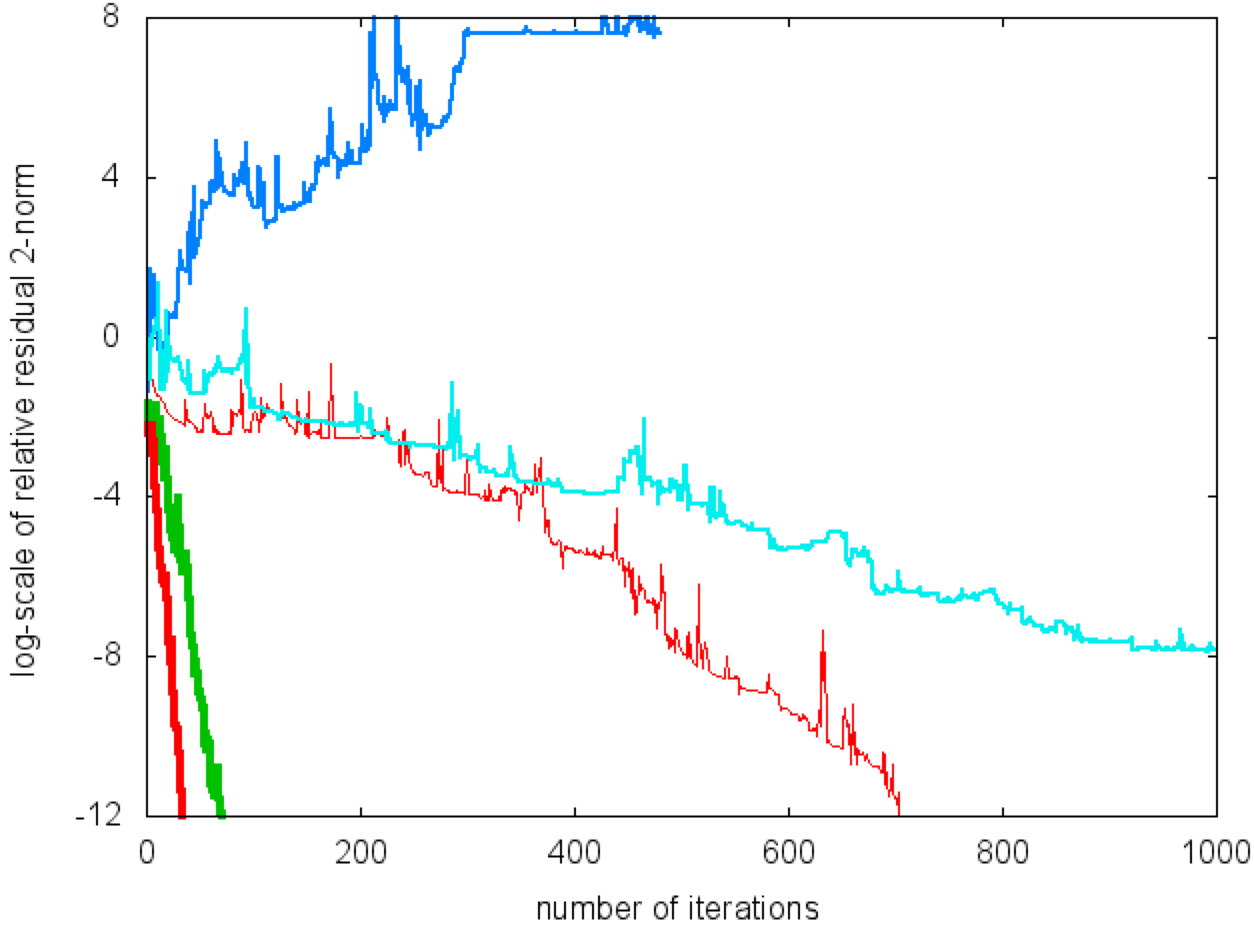
>> 比較前処理 (Krylov部分空間解法: BiCGSTAB法)
-- Jacobi法, 重み付きJacobi法,
-- Gauss-Seidel法, 重み付きGauss-Seidel法

>> 各種パラメータ
-- 初期近似解 : $x_0 = [0, 0, \ldots, 0]^\top$
-- 前処理の反復回数 : 10回
-- Arnoldi法の最小/最大反復回数 : 10回/20回
-- Arnoldi法の停止条件 : 10^-2

数値実験

実験結果II

>> 収束履歴（non, J, GS, W-J, W-GS）



数値実験

数値実験III

>> 目的: 超新星爆発計算における有効性を検証する

>> 3次元ニュートリノ輻射輸送 [K.Sumiyoshi, 2012]

-- ニュートリノ分布（空間3次元 + ニュートリノの運動量3次元）

$$f_\nu := f_\nu(r, \theta, \phi; \epsilon, \theta_\nu, \phi_\nu; t)$$

-- Boltzmann方程式

$$\frac{1}{c}\frac{\partial f_\nu}{\partial t} + \frac{\partial f_\nu}{\partial s} = \left[\frac{1}{c}\frac{\partial f_\nu}{\partial t}\right]_{\text{collision}}$$

-- 3次元球座標系で表現 -> 保存系に変形

$$\frac{1}{c}\frac{\partial f_\nu}{\partial t} + \frac{\mu_\nu}{r^2}\frac{\partial}{\partial r}(r^2 f_\nu) + \frac{1}{r}\frac{\partial}{\partial \mu_\nu}[(1-\mu_\nu^2)f_\nu] + \frac{\sqrt{1-\mu_\nu^2}\cos\phi_\nu}{r\sin\theta}\frac{\partial}{\partial\theta}(\sin\theta f_\nu)$$
$$-\frac{\sqrt{1-\mu_\nu^2}\cos\theta}{r\sin\theta}\frac{\partial}{\partial\phi_\nu}(\sin\phi_\nu f_\nu) + \frac{\sqrt{1-\mu_\nu^2}\sin\phi_\nu}{r\sin\theta}\frac{\partial f_\nu}{\partial\phi} = \left[\frac{1}{c}\frac{\delta f_\nu}{\delta t}\right]_{\text{collision}}$$

数値実験

超新星爆発計算

>> 概略

-- 時間発展：陰解法

-> 時間ステップ毎に大規模線形方程式の求解が必要

$$(N_r, N_\theta, N_\phi, N_\epsilon, N_{\theta_\nu}, N_{\phi_\nu}) = (200, 9, 9, 14, 6, 12)$$

→ n ≃ 1600万, Nnz ≃ 12.8億, Nnz/n ≃ 80

-- 要求される時間発展：実時間で約1秒

-> 時間ステップを大きくとりたい ($\Delta t \simeq 10^{-5}$程度)

-- 時間ステップを大きくすると方程式が解きにくくなる.

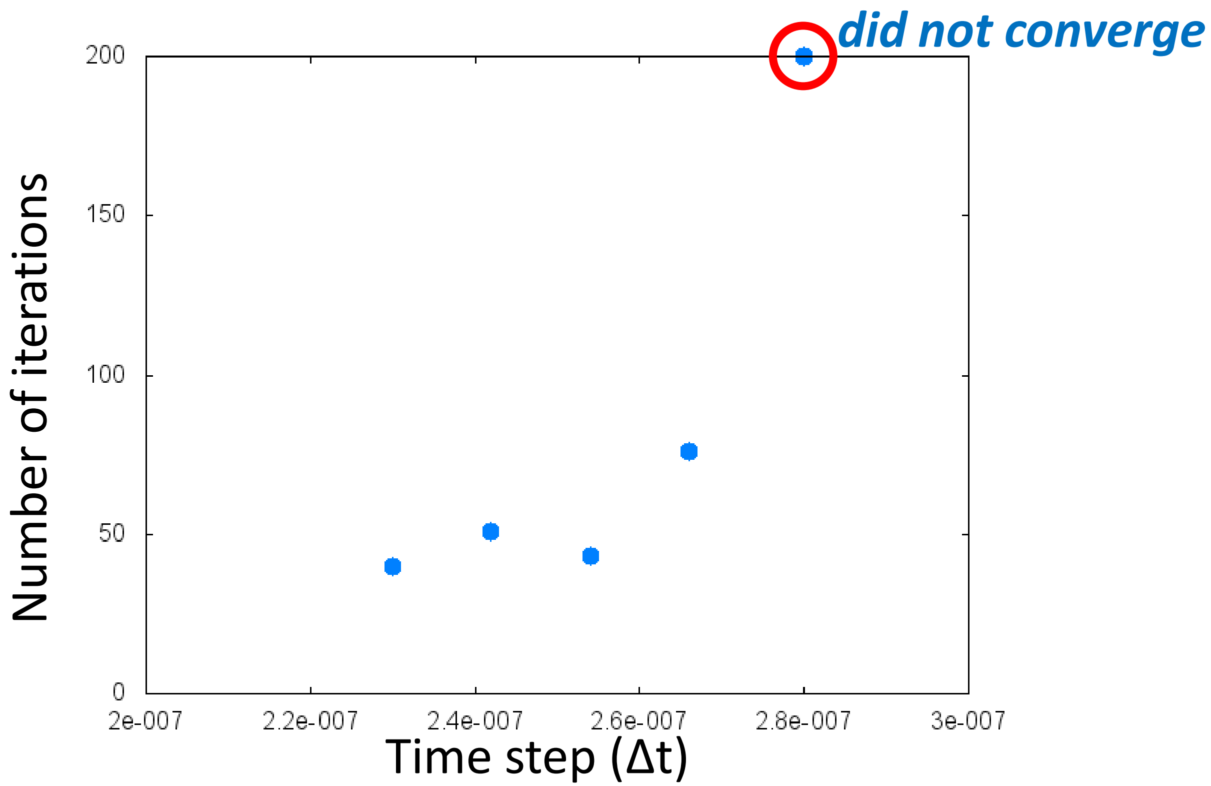
>> 従来手法(高並列性を考慮)

-- 対角スケーリング前処理付きBiCGSTAB法

数値実験

超新星爆発計算

>> 対角スケーリング前処理付きBiCGSTAB法の収束性



数値実験

提案法の概略

>> 重み付きJacobi型前処理

$$x_{k+1} = x_k + \omega D^{-1}(b - Ax_k)$$

-- 重みパラメータ ω を最適化

-- 対角行列 D は事前に準備した候補の中から $\rho(G_{\omega^*})$ を最小化するものを選択

数値実験

提案法の概略

>> 各種パラメータ

-- 前処理の反復回数 : 10回

-- Arnoldi法の最小/最大反復回数 : 10回 / 20回

-- Arnoldi法の停止条件 : 10^-2

-- 対角行列の候補 :

$$d_i = a_{ii}$$

$$d_i = \sum_j |a_{ij}|$$

$$d_i = \sqrt{\sum_j a_{ij}^2}$$

計算環境

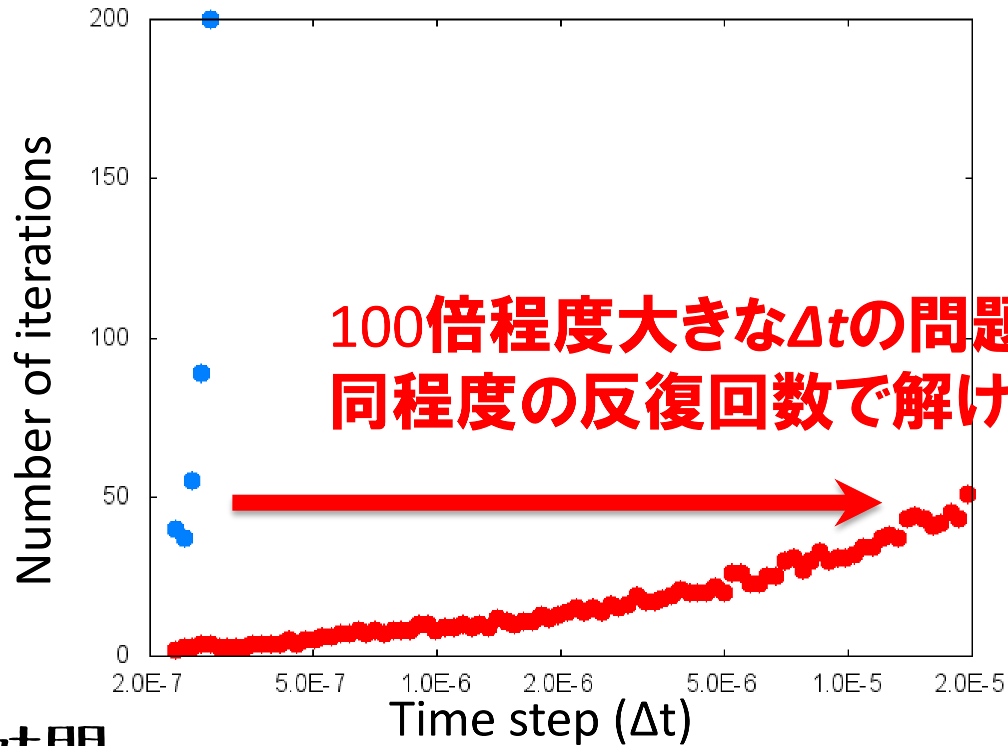
>> KEK SR16000/M1

-- 1 node 32 cores (logical 64 cores)

数値実験

超新星爆発計算

>> 提案する前処理法の収束性



>> 計算時間

-- 対角スケーリング　：0.24秒/反復

-- 提案前処理　　　　：1.75秒/反復（最適化4～7秒）

まとめ・今後の課題

まとめ・今後の課題

まとめ

Conclusion

高並列性を考慮した前処理として，反復法を用いた重み付き定常反復型前処理を導入し，そのパラメータ最適化手法を提案・有効性の検証を行った．

数値実験から，少ない計算コストで重みパラメータを最適化でき，また超新星爆発計算に対して前処理の有効性が確認された．

今後の課題

Future works

・並列化効率の測定
・他の前処理との比較