

BiCGSTAB(ℓ) 法の有効な使い方について

An Effective Guideline for BiCGSTAB(ℓ) method

藤野清次 (九州大学情報基盤センター)

Seiji FUJINO (Kyushu University)

三浦謙一 (富士通 (株), 九州大学情報基盤センター)

Kenichi MIURA (Fujitsu Ltd., Kyushu University)

1 はじめに

ここでは、連立1次方程式 $Ax = b$ を BiCGSTAB(ℓ) 法 [12] で解くことを考える。ここで、 A は $n \times n$ の正方行列、 b, x は n 次元の右辺ベクトルおよび解ベクトルとする。BiCGSTAB(ℓ) 法は BiCG 法系統の反復解法の一つで、その他に例えば、BiCG 法, CGS 法, BiCGSTAB 法, BiCGSTAB2 法, GPBiCG 法, など多くの解法がある。一般に、BiCG 系統の解法では初期残差 r_0 と初期シャドウ残差 r_0^* が必要になる。前者は、初期値 x_0 を与えれば、 $r_0 = b - Ax_0$ の式から自ずと決まる。しかし、後者の初期シャドウ残差の与え方には任意性が残っている。すなわち、初期シャドウ残差が満たすべき条件としては、内積 $(r_0^*, r_0) \neq 0$ だけである。しかし、この条件を最低限満たすということとその簡便さから、従来は $r_0^* = r_0$ と与えることがよく行なわれてきた。そして、その条件下で様々な反復法の収束性、頑強性、効率などが今まで論じられてきた。

一方、BiCGSTAB(ℓ) 法では計算を始める前に段数 ℓ を陽に与える必要がある。そのため、従来の研究 [8] [9] [11] では収束過程の残差などの情報から、例えば、停滞 (stagnation) や破綻 (break down) などの現象を検出し、段数 ℓ を一時的に増加させることによって停滞や破綻を回避し、それらを通じた後、元の段数に戻るといった適応型の BiCGSTAB(ℓ) 法の研究開発に重点がおかれてきた。元の BiCGSTAB(ℓ) 法の収束性の本質についてはあまり目が向けられてこなかった感がある。

そこで、本研究では、初期シャドウ残差 r_0^* の与え方に任意性が残っていることに着目し、それを適切に与えることによって BiCGSTAB(ℓ) 法の収束特性が向上することを数値実験にて示す。すなわち、 r_0^* に一様乱数を与えた BiCGSTAB(ℓ) 法の有効性を検証する。さらに、収束時の近似解の相対残差の大きさにも注意を向け、単に効率の比較だけでは論じられない BiCGSTAB(ℓ) 法の総合的な収束性について論じる。そして、BiCGSTAB(ℓ) 法の有効な使用指針を示す [4]。

2 BiCGSTAB(ℓ) 法の算法

以下に BiCGSTAB(ℓ) 法の算法を示す。BiCGSTAB(ℓ) 法では、 $k = m\ell + \ell$ ($m = 0, 1, 2, \dots$) に関して残差ベクトルや近似解の更新などが行なわれる。この反復は外部反復と呼ばれる。 ℓ

を段数と呼ぶ。外部反復は二つの内部反復から構成されている。一つは、BiCG法を使うBiCGパートで、もう一つは ℓ 次のMR(Minimal Residual)多項式を用いるパートである。後者では、修正Gram-Schmidtの直交化を利用して ℓ 本のベクトルの直交化が行なわれる。そこでの演算量は $O(\ell^2)$ で見積もられ、BiCGパートの演算量に比べて非常に多くなる。

[Algorithm of BiCGStab(ℓ)]

$k = -\ell$

choose $\mathbf{x}_0, \mathbf{r}_0^*$, compute $\mathbf{r}_0 = \mathbf{b} - A\mathbf{x}_0$

take $\mathbf{u}_{-1} = 0, \mathbf{x}_0 = \mathbf{x}_0, \rho_0 = 1, \alpha = 0, \omega = 1$

repeat until $\|\mathbf{r}_{k+\ell}\|_2$ is small enough

$k = k + \ell$

Put $\hat{\mathbf{u}}_0 = \mathbf{u}_{k-1}, \hat{\mathbf{r}}_0 = \mathbf{r}_k, \hat{\mathbf{x}}_0 = \mathbf{x}_k, \rho_0 = -\omega\rho_0$

For $j = 0, 1, \dots, \ell - 1$ [BiCG part]

$$\rho_1 = (\hat{\mathbf{r}}_j, \mathbf{r}_0^*), \beta = \alpha \frac{\rho_1}{\rho_0}, \rho_0 = \rho_1$$

$$\hat{\mathbf{u}}_i = \hat{\mathbf{r}}_i - \beta \hat{\mathbf{u}}_i \quad (i = 0, 1, \dots, j)$$

$$\hat{\mathbf{u}}_{j+1} = A\hat{\mathbf{u}}_j, \gamma = (\hat{\mathbf{u}}_{j+1}, \mathbf{r}_0^*), \alpha = \frac{\rho_0}{\gamma}$$

$$\hat{\mathbf{r}}_i = \hat{\mathbf{r}}_i - \alpha \hat{\mathbf{u}}_{i+1} \quad (i=0, 1, \dots, j)$$

$$\hat{\mathbf{r}}_{j+1} = A\hat{\mathbf{r}}_j, \hat{\mathbf{x}}_0 = \hat{\mathbf{x}}_0 + \alpha \hat{\mathbf{u}}_0$$

end

For $j = 1, 2, \dots, \ell$ [MR part]

$$\tau_{ij} = \frac{1}{\sigma_i} (\hat{\mathbf{r}}_j, \hat{\mathbf{r}}_i), \hat{\mathbf{r}}_j = \hat{\mathbf{r}}_j - \tau_{ij} \hat{\mathbf{r}}_i \quad (i = 1, 2, \dots, j - 1)$$

$$\sigma_j = (\hat{\mathbf{r}}_j, \hat{\mathbf{r}}_j), \gamma'_j = \frac{1}{\sigma_j} (\hat{\mathbf{r}}_0, \hat{\mathbf{r}}_j)$$

end

$$\gamma_\ell = \gamma'_\ell, \omega = \gamma_\ell$$

$$\gamma_j = \gamma'_j - \sum_{i=j+1}^{\ell} \tau_{ji} \gamma_i \quad (j=\ell-1, \dots, 1)$$

$$\gamma''_j = \gamma_{j+1} + \sum_{i=j+1}^{\ell-1} \tau_{ji} \gamma_{i+1} \quad (j = 1, 2, \dots, \ell - 1)$$

$$\hat{\mathbf{x}}_0 = \hat{\mathbf{x}}_0 + \gamma_1 \hat{\mathbf{r}}_0, \hat{\mathbf{r}}_0 = \hat{\mathbf{r}}_0 - \gamma'_\ell \hat{\mathbf{r}}_\ell, \hat{\mathbf{u}}_0 = \hat{\mathbf{u}}_0 - \gamma_\ell \hat{\mathbf{u}}_\ell$$

$$\hat{\mathbf{u}}_0 = \hat{\mathbf{u}}_0 - \gamma_j \hat{\mathbf{u}}_j, \hat{\mathbf{x}}_0 = \hat{\mathbf{x}}_0 + \gamma''_j \hat{\mathbf{r}}_j, \hat{\mathbf{r}}_0 = \hat{\mathbf{r}}_0 - \gamma'_j \hat{\mathbf{r}}_j \quad (j = 1, 2, \dots, \ell - 1)$$

Put $\mathbf{u}_{k+\ell-1} = \hat{\mathbf{u}}_0, \mathbf{r}_{k+\ell} = \hat{\mathbf{r}}_0, \mathbf{x}_{k+\ell} = \hat{\mathbf{x}}_0$

end.

簡単に演算量について考える。例えば、行列ベクトル積の演算量を1とすると、段数(ℓ)を使って、BiCGSTAB(ℓ)法の平均演算量は $\ell + 5$ と表される[12]。したがって、BiCGSTAB(ℓ)法において $\ell=1$ と $\ell=2$ の演算量の比は $7/6 = \text{約} 1.17$ 、同様に $\ell=1$ と $\ell=8$ との同比は $13/6 = \text{約}$

2.17となる。段数を増やすと演算量は急に増加する。そのため、段数増加という操作は収束性を改善する反面計算時間の増大を招く。

3 初期シャドウ残差 r_0^* について

以下のように固有ベクトルおよび固有値の各個数を表す記号を定める。

k	初期残差 r_0 の固有ベクトルの個数
k^*	初期シャドウ残差 r_0^* の固有ベクトルの個数
K	行列 A の相異なる固有値の個数

BiCG 法系統の反復法では、係数行列 A の性質だけでなく初期残差 r_0 の固有ベクトルの個数と収束性が関係深い。すなわち、残差ベクトル r_n は残差多項式 $R_n(A)$ と初期残差 r_0 との積で表され、初期残差 r_0 も解法の収束性を左右する重要な要因である。また、収束までの反復回数の点から、初期残差 r_0 に含まれる固有値に対応する固有ベクトル成分の個数 k は少ない方が好ましい。したがって、従来のように、 $x_0 = 0$, $r_0 = b$ かつ $r_0^* = r_0$ とおくのは、固有ベクトル成分をこれ以上増加させないという意味で良案の一つである。反復解法の収束性は、固有ベクトル成分の個数： k と k^* の大小関係によって以下のように分類される [1] [3]。

1. $k < K$ のとき。すなわち、初期残差の固有ベクトル成分の個数が少ないとき。
 - (a) $k^* \approx k$ のとき、速い収束性が得られたり、丸め誤差の影響で発散したりする。
 - (b) $k < k^*$ のとき安定に収束する。
2. $k \approx K$ のとき。すなわち、初期残差、初期シャドウ残差の固有ベクトル成分がともに多いとき、反復回数は多いが安定に収束する。

本研究で考察する、初期残差を $r_0 = b$ とおき、初期シャドウ残差 r_0^* に乱数を与える、というのは上記 1.-(b) の場合 (下線部分) にあたる。本研究ではこれらの場合の収束性向上の可能性について調べた [2] [6]。また、一般に、一様乱数にはすべての固有ベクトル成分が含まれる可能性が高いと推測される。

4 テストで用いた乱数について

初期解に代入する乱数には、以下の3種類の一様乱数 (区間は $[0,1]$) を用いた。

- Mersenne Twister 法による乱数 (以下, 'MT' と略す)[7].
- M 系列乱数 ('m 系' と略す)。演算としては、合同法の乗算 (および加算) に対して、排他的論理和をとる論理演算を用いる [5].
- 線形合同法による乱数 ('合同' と略す)。漸化式を使った乱数法である。

4.1 初期値： x_0, x_0^*, r_0, r_0^* の組合せ

初期値： x_0, x_0^*, r_0, r_0^* の組合せを以下に示す。以下では、タイプAとタイプBにおけるBiCGSTAB(ℓ)法の収束性を調べた。

タイプ	場合 の数	初期残差 r_0, r_0^* の与え方	初期解の与え方	
			x_0	x_0^*
A	1	$r_0=r_0^*=b$	$x_0=0$	$x_0^*=0$
B	3	$r_0=b, r_0^*=b-A^T x_0^*$	$x_0=0$	$x_0^*=MT$
			$x_0=0$	$x_0^*=m$ 系
			$x_0=0$	$x_0^*=$ 合同

5 数値実験

数値実験は九州大学情報基盤センターの富士通:スーパーコンピュータ VPP5000 上で倍精度演算で実行した。言語は Fortran90 を使用した。コンパイルは frt コマンドのみとし、特別の最適化オプションは使わなかった。実験は従来のタイプAおよび乱数を使うタイプBの場合について行った。また、タイプBの結果の数字は3種類の乱数のときの平均値である。 n は行列の次元数を表す。表中の'相対残差'とは第 k 回目の反復における $\log_{10}(\|b - Ax_k\|_2 / \|b\|_2)$ を表す。段数 ℓ の最大値は8とした。収束判定は各反復での相対残差2ノルムが $\|r_k\| / \|r_0\| \leq 10^{-12}$ のときとし、最大反復回数は2000回とした。行列に前処理はしなかった。収束の"破綻 (break down)" は以下の σ_k の式で判定した。すなわち、 σ_k の値が計算機イプシロン $\epsilon (=2.2 \times 10^{-16})$ の平方根、すなわち 1.5×10^{-8} 以下のとき残差は"破綻"したと判定した。

$$\sigma_k = (r_k, r_0^*) / (\|r_k\| \cdot \|r_0^*\|) \leq \sqrt{\epsilon} \quad (5.1)$$

5.1 テスト問題

1. [問題 1] 実数 Toeplitz 行列: γ は実数パラメータ。右辺項は厳密解がすべて1となるように定めた [10]。次元数は $n=262144$ である。

$$A := \begin{bmatrix} 2 & 1 & & & & & & 0 \\ 0 & 2 & 1 & & & & & \\ \gamma & 0 & 2 & 1 & & & & \\ & \gamma & 0 & 2 & 1 & & & \\ & & \gamma & 0 & 2 & \ddots & & \\ 0 & & & \ddots & \ddots & \ddots & & \end{bmatrix}$$

2. [問題 2] 正方形領域 $[0,1]^2$ で定義された偏微分方程式から生じた境界値問題 (全周 Dirichlet 境界条件): 実数行列 [6]。 D は実数。次元数 $n=256^2$ 。格子幅 h は $1/257$ 。厳密解を $u(x, y) = 1 + xy$ とし、右辺項 $G(x, y)$ を定めた。

$$-u_{xx} - u_{yy} + Du_x = G(x, y)$$

3. [問題 3] 正方形領域 $[0,1]^2$ で定義された以下の境界値問題 (全周 Dirichlet 境界条件) :
 実数行列 ([6] 一部条件変更). D は実数. 次元数 $n=256^2$. 格子幅 h は $1/257$. 厳密解が
 $u(x, y) = 1 + xy$ となるように右辺項 $G(x, y)$ を定めた.

$$-u_{xx} - u_{yy} + D \left\{ \left(y - \frac{1}{2} \right) u_x + \left(x - \frac{1}{3} \right) \left(x - \frac{2}{3} \right) u_y \right\} - \pi^2 u = G(x, y)$$

5.2 実験結果

表1-表2に問題1の結果を示す. 表2中の下線を付けた数字は収束までの計算時間が最も少なかったものを表す. 以下の表でも同様の意味とする. また, 表中の“-”印は最大反復回数までで収束しなかったものを表す.

表1: Toeplitz 行列のとき, BiCGSTAB(ℓ) 法において計算時間が最も少なかったときの段数 (ℓ), 時間 (秒), 相対残差, 2つのタイプの時間の比

γ	タイプ A			タイプ B			比 (%)
	段数	時間	相対残差	段数	時間	相対残差	
1.4	3	.165	-12.6	2	.153	-12.3	92.7
1.5	3	.184	-12.2	2	.179	-12.3	97.3
1.6	3	.278	-12.2	2	.227	-12.1	81.6
1.7	5	.445	-12.1	2	.270	-12.1	60.6
1.8	-	-	-	2	.357	-12.3	
2.0	-	-	-	2	.677	-12.1	
2.2	-	-	-	3	1.90	-12.1	
2.3	-	-	-	3	5.94	-12.1	

表2: Toeplitz 行列 ($\gamma=1.7$) のときの段数, 反復回数, 時間 (秒), 相対残差, 比

段数	タイプ A			比 (%)	タイプ B				
	回数	時間	相対残差		回数	時間	相対残差	1反復時間	比
1	-	-	-		327	.939	-12.1	2.872	1.0
2	-	-	-		85	.270	-12.1	3.176	1.106
3	破綻	-	-		81	.279	-12.2	3.444	1.199
4	124	.466	-12.1	67.0	83	.312	-12.2	3.759	1.309
5	110	.445	-12.1	75.0	82	.334	-12.3	4.073	1.418
6	132	.580	-12.6	63.9	84	.371	-12.5	4.417	1.538
7	182	.860	-12.3	46.2	84	.397	-12.6	4.726	1.646
8	96	.485	-12.0	83.4	80	.405	-12.1	5.063	1.763

表2のタイプBについては, 反復1回当たりの計算時間 (単位: ミリ秒) と段数 $\ell=1$ のときの同時間を1としたときの比を示す. 表1と表2の結果から以下のことがわかる.

1. より大きな γ の値に対して, 乱数を使うタイプ B の方が従来のタイプ A よりも収束性が頑強である.
2. タイプ B の方が段数も少なく計算時間も短い.
3. タイプ A では, 段数を大きくしても必ずしも反復回数は減らないこと, 逆に増えることもある.
4. タイプ B では反復回数はほぼ一定である.
5. いずれの場合も, 近似解は要求精度と同じ 12 桁の精度を持つ.
6. タイプ A では, 段数 $l=3$ のとき (反復回数が 205 回を超えたとき, 図 1 参照) 残差の破綻が起こった.
7. 段数を増すに従って, 反復 1 回当たりの計算時間がほぼ一定の比率で増えている.

図 1 に $\gamma=1.7$ のときの収束の履歴を示す. 段数 $l=2$ と段数 $l=3$ の二つの場合を示す. いずれの場合もタイプ A とタイプ B のときの残差ノルムの値をプロットした. 図から, タイプ A では収束しなかったものが, タイプ B では $l=2$ と $l=3$ のいずれの場合も収束したことがわかる. 特に, タイプ A で $l=3$ のとき, 反復回数が 205 回を超えたとき残差が急激に増加し (オーダー 10^{-5} からオーダー 10^6 へ), (5.1) 式の “破綻” が検出された. ただし, 図 1 は破綻の後も反復計算を続行させたときの収束の履歴である.

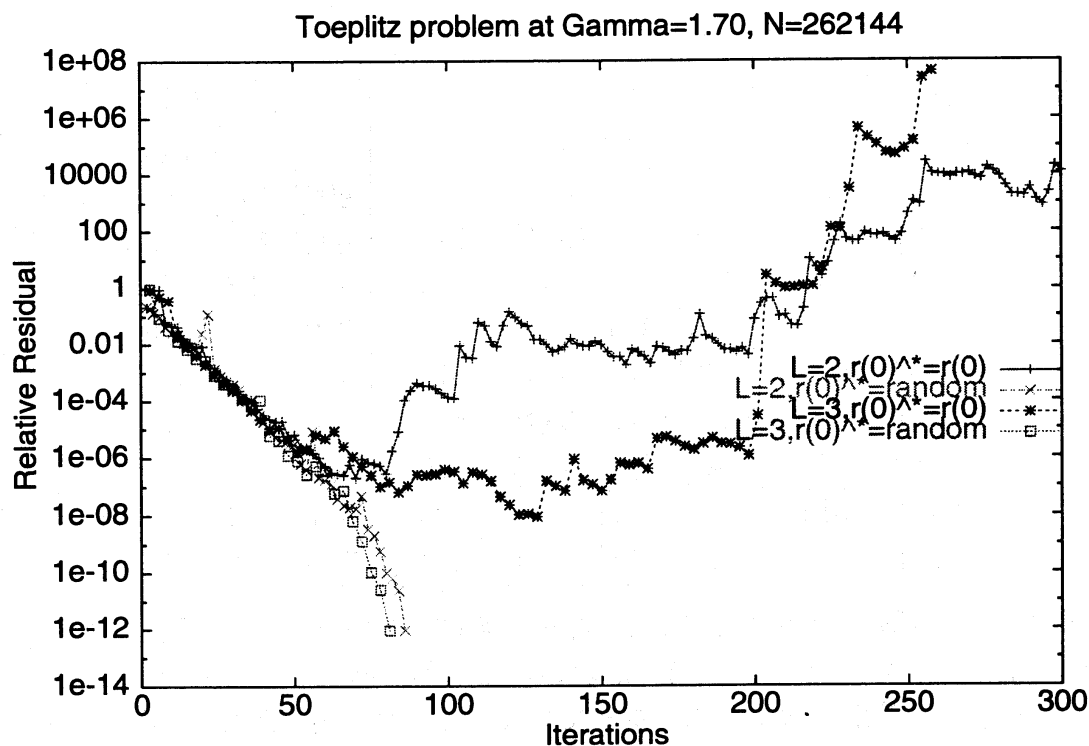


図 1: Toeplitz 問題に対するタイプ A とタイプ B の各収束履歴 ($\gamma=1.7$, 段数 $l=2$ と $l=3$ のとき)

次に、表3-表7に問題2の結果を示す。表3から、(i)タイプAの相対残差が要求精度に比べて悪いこと、一方、タイプBはほぼ要求精度を満たしていること、(ii)計算時間は21.4%～41.0% (反復回数の平均で67.2%に減少、計算時間で65.9%に減少)ほどタイプBの方が少ないことがわかる。

表3: 問題2(Dh=2)のときの段数, 反復回数, 時間(秒), 相対残差, 比

段数	タイプA			比 (%)	タイプB		
	回数	時間	相対残差		回数	時間	相対残差
1	<u>547</u>	<u>.491</u>	-5.4	78.6	447	<u>.386</u>	-12.2
2	582	.529	-8.0	71.6	413	.388	-11.9
3	582	.599	-8.5	67.6	401	.405	-12.2
4	604	.688	-7.8	62.9	<u>387</u>	.433	-11.8
5	660	.770	-6.7	<u>59.0</u>	392	.454	-11.7
6	672	.841	-6.0	61.6	398	.518	-11.6
7	623	.853	-6.5	66.4	418	.567	-10.9
8	648	.954	-5.4	65.2	445	.622	-10.6
平均 (比)	615 (1.0)	.716 (1.0)	-6.8	66.6	413 (.672)	.472 (.659)	-11.6

表4: 問題2(Dh=2とDh=4のとき): 乱数の違いによる反復回数の変化

段数	Dh=2				Dh=4			
	タイプA	タイプB			タイプA	タイプB		
		MT	合同	m系		MT	合同	m系
1	547	448	449	<u>445</u>	-	-	-	<u>919</u>
2	582	414	<u>400</u>	424	558	<u>400</u>	402	406
3	582	396	423	<u>384</u>	576	<u>399</u>	411	402
4	604	384	<u>368</u>	408	524	404	428	<u>384</u>
5	660	390	<u>350</u>	435	530	390	405	<u>380</u>
6	672	426	<u>378</u>	390	576	426	402	<u>396</u>
7	623	427	427	<u>399</u>	588	<u>420</u>	476	469
8	648	440	<u>424</u>	472	656	<u>408</u>	440	448
平均	615	416	402	420	573	407	423	412
比	1.0	.676	.653	.682	1.0	.710	.738	.719

表4の結果から、3種類の乱数による違いは、その差はいずれも小さく、Dh=2ときは“合同”がよく、Dh=4のときは“MT”と“m系”がよいことがわかる。しかし、従来のタイプAに比較すると、いずれも回数が大幅に少なくなっていることがわかる。ただし、表4の平均値は段数 $l=2$ から段数 $l=8$ までの平均である。次に、表5から、Dh=4のときも、タイプAの相対残差が悪いことがわかる。また、計算時間も26%～35%程、タイプBの方が少ないことがわかる。表5の段数が1のときのタイプBの括弧つきの数字は、3種類の乱数のうち収束した'M系'の乱数のときの値を示している。

表5: 問題2(Dh=4)のときの段数, 反復回数, 時間(秒), 相対残差, 比

段数	タイプA			比 (%)	タイプB				
	回数	時間	相対 残差		回数	時間	相対 残差	1反復 時間	比
1	-	-	-	-	(919)	(.767)	(-11.4)	0.835	1.0
2	558	.505	-8.1	72.0	403	.364	-12.0	0.903	1.081
3	576	.561	-7.3	70.1	404	.393	-12.1	0.973	1.165
4	524	.555	-7.8	77.2	405	.429	-11.4	1.059	1.268
5	530	.602	-7.8	73.8	392	.444	-11.8	1.133	1.357
6	576	.703	-7.4	70.7	408	.497	-11.2	1.218	1.459
7	588	.764	-6.8	77.3	455	.591	-11.3	1.299	1.556
8	656	.908	-5.9	65.7	432	.597	-10.9	1.382	1.655
平均 (比)	573 (1.0)	.657 (1.0)	-7.3	72.4	414 (.722)	.474 (.722)	-11.5		

表6から, パラメータ Dh の値を 0.5 から 32 まで変えたとき, いずれもタイプ B の方が従来のタイプ A よりも計算時間(平均 82.5%に減少)も少ないこと, また相対残差の値から要求精度の 12 桁まで精度が得られていることがわかる. 一方, タイプ A の近似解の相対残差の値が非常に悪いことも目立つ. 特に, タイプ A の Dh=2.0 のとき相対残差が-5.4 しか得られず最悪である. また, Dh が 1.0 と 16 の間の値のときも相対残差が最良でも約-9.1 しか得られないなど精度が悪い.

表6: 問題2: Dh を変化させたとき, 最少の計算時間のときの段数, タイプ A と B の時間(秒), 比(%), 相対残差

Dh	段数	時間(秒)		比 (%)	相対残差	
		タイプ A	タイプ B		タイプ A	タイプ B
.5	1	.398	.372	93.5	-10.3	-12.0
1.	1	.455	.386	84.9	-9.1	-11.4
2.	1	.491	.386	78.6	-5.4	-12.2
4.	2	.505	.364	72.0	-8.1	-12.0
8.	2	.463	.357	77.0	-9.1	-12.3
16	2	.415	.364	87.8	-9.1	-12.3
32	2	.460	.386	83.9	-11.1	-12.3
平均		.455	.374	82.5	-8.9	-12.1

次に, 表7から, 収束までの時間が最も短かった(Dh が 2 より小さいとき $\ell=1$ のときが最速, Dh=4 以上のとき $\ell=2$ のときが最速) ときと段数が 8 のときを比べて, 相対残差が大幅に悪くなることなどがわかる. しかし, その度合はタイプ A の方が大きく, 段数をむやみに大きくすると近似解の精度を保持することが難しくなることがわかる. 図2に Dh=4.0, 段数 $\ell=2$ のときの 2 つのタイプの収束履歴を示す.

表 7: 問題 2: 最少の計算時間のときの段数と相対残差 vs. 段数=8 のときの相対残差の比較

Dh	タイプ A の相対残差				タイプ B の相対残差			
	段数=1	段数=2	段数=8	差	段数=1	段数=2	段数=8	差
.5	-10.3	-10.7	-6.5	-3.8	-12.0	-12.3	-9.6	-2.4
1.	-9.1	-7.9	-5.7	-3.4	-11.4	-11.4	-9.2	-2.2
2.	-5.4	-8.0	-5.4	-0.0	-12.2	-11.9	-10.6	-1.6
4.	-	-8.1	-5.9	-2.2	-	-12.0	-10.9	-1.1
8.	-	-9.1	-8.0	-1.1	-	-12.3	-11.8	-0.5
16	-	-9.1	-8.5	-0.6	-	-12.3	-12.0	-0.3
32	-	-11.1	-10.3	-0.8	-	-12.3	-12.2	-0.1
平均	-8.3	-9.1	-7.2		-11.9	-12.1	-10.9	

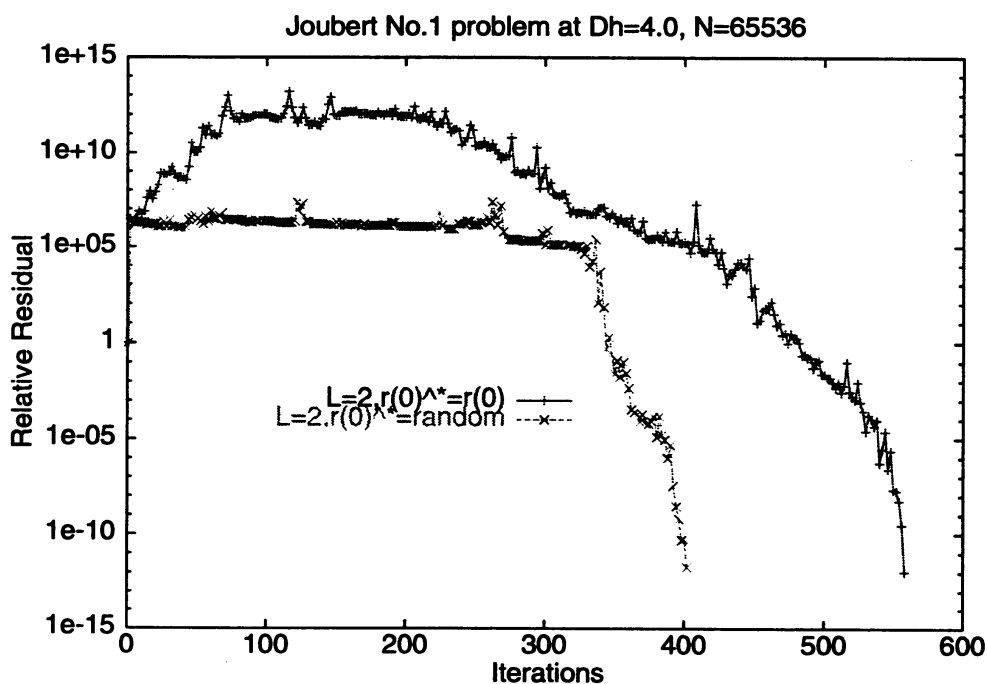


図 2: 問題 2(Dh=4.0) に対するタイプ A とタイプ B の各収束履歴 (段数 L=2 のとき)

表 8-表 9 に問題 3 の結果を各々示す. 表 8 からタイプ A とタイプ B の計算時間の差は 4.4% と小さいことがわかる. しかし, 特に Dh=4 のとき, タイプ A では相対残差が -8.8 しかなく, タイプ B の相対残差との差が非常に大きいことがわかる. 相対残差の平均でも 1.1 桁ほどタイプ B の相対残差がタイプ A よりよいことがわかる. また, タイプ B のとき, 相対残差の値は -12.0 前後の値で非常に安定してよいこともわかる. 表 9 でも, 表 7 と同様に, 従来のタイプ A の方が段数を増やしたときの近似解の相対残差の低下の度合いが大きく近似解の精度が失われやすいことがわかる.

以上の数値実験結果から, 従来のタイプ A の初期値設定法と初期シャドウ残差 r_0^* に一様乱数を与えるタイプ B の BiCGSTAB(ℓ) 法の収束性に関して次の知見が得られた.

表 8: 問題 3 : Dh を変化させたとき, 最少の計算時間のときの段数およびタイプ A と B の各時間 (秒), 比 (%), 相対残差

Dh	段数	時間 (秒)		比 (%)	相対残差	
		タイプ A	タイプ B		タイプ A	タイプ B
.5	1	0.943	0.861	91.3	-10.9	-12.0
1.	1	1.093	1.071	98.0	-11.6	-12.1
2.	1	1.298	1.241	95.6	-10.1	-11.9
4.	1	1.725	1.577	91.4	-8.8	-12.0
8.	2	2.149	2.074	96.5	-10.9	-12.1
16	2	3.384	3.379	99.8	-11.9	-12.2
32	2	5.048	4.936	97.7	-11.6	-12.1
64	3	7.676	7.251	94.4	-12.0	-12.2
平均		2.915	2.799	95.6	-11.0	-12.1

表 9: 問題 3 : 最少の計算時間のときの段数と相対残差 vs. 段数=8 のときの相対残差の比較

Dh	タイプ A の相対残差			タイプ B の相対残差		
	最少のとき	$l=8$ のとき	差	最少のとき	$l=8$ のとき	差
.5	-10.9 ($l=1$)	-9.4	-1.5	-12.0 ($l=1$)	-10.6	-1.4
1.	-11.6 ($l=1$)	-7.5	-4.1	-12.1 ($l=1$)	-9.4	-2.7
2.	-10.1 ($l=1$)	-7.7	-2.4	-11.9 ($l=1$)	-9.5	-2.4
4.	-8.8 ($l=1$)	-5.7	-3.1	-12.0 ($l=1$)	-10.4	-1.6
平均	-10.4	-7.6	-2.8	-12.0	-10.0	-2.0
8.	-10.9 ($l=2$)	-8.9	-2.0	-12.1 ($l=2$)	-11.2	-0.9
16	-11.9 ($l=2$)	-10.0	-1.9	-12.2 ($l=2$)	-11.2	-1.0
32	-11.6 ($l=2$)	-10.9	-0.7	-12.1 ($l=2$)	-11.1	-0.9
64	-12.0 ($l=3$)	-11.9	-0.1	-12.2 ($l=3$)	-11.9	-0.3
平均	-11.6	-10.4	-1.2	-12.2	-11.4	-0.8

1. 段数を大きくすると近似解の相対残差はだんだんと悪くなる. そのため, 段数をむやみに大きな値にするのは危険である. (例えば, 表 3, 表 5, 表 7, 表 9 より)
2. 従来のタイプ A よりもタイプ B が収束性が頑強である. (主として表 1 と表 6 より)
3. タイプ B の近似解の相対残差は, 要求精度と同じオーダ 10^{-12} が保持された. 一方, タイプ A では, 平均でオーダ $10^{-8.9}$ と非常に悪い. このことは計算時間の比較以前の問題である. (表 6 より. 同様の傾向が表 8 でも見られる)
4. 計算時間もタイプ B が従来のタイプ A よりも少ない. (表 1 と表 6 より)

6 まとめ

タイプBのBiCGSTAB(ℓ)法, すなわち, 初期シャドウ残差 r_0^* に一様乱数を使用するBiCG-STAB(ℓ)法を使えば, 計算時間および得られた近似解の相対残差の観点から, 最少の計算時間で妥当な近似解が得られる可能性が高く, そのときの段数の上限も $\ell = 3$ 程度の少ない数で十分である.

謝辞:

本研究を進めるにあたり, 理化学研究所 阿部邦美博士, 東京大学大学院 張紹良助教授, 本学情報基盤センター 南里豪志助教授から有用なご助言を戴きました. 心から謝意を表します.

参考文献

- [1] Abe, K., Fujino, S., Himeno, R., Effect of Initial Residual in Variants of Bi-CG Method, *Trans. of Journal Information*, 5(2002), 17-32.
- [2] Freund, R., Conjugate gradient-type methods for linear systems with complex symmetric coefficient matrices, *SIAM J. Sci. Stat. Comput.*, 13(1992), 425-448.
- [3] 藤野清次, 阿部邦美, BiCG 系統の反復解法に対する効果的な収束改善法について, ハイパフォーマンスコンピューティングと計算科学シンポジウム HPCS2002 講演集, つくば国際会議場, 2002.1, 51-58.
- [4] Fujino, S., On Enhancement of Convergence of BiCGSTAB(L) Method, The Proc. of LATSIS 2002 – Iterative Solvers for Large Linear Systems–, ETH Zurich, 2002.2, p.19.
- [5] 伏見正則, –UP 応用数学選書– 乱数, 東京大学出版会, 2000.
- [6] Joubert, W., Lanczos methods for the solution of nonsymmetric systems of linear equations, *SIAM J. Matrix Anal. Appl.*, 13(1992), 926-943.
- [7] Matsumoto, M., Nishimura, T., Mersenne Twister, *ACM Trans. on Model. Comput. Simul.*, 8(1998), 3-30.
- [8] 宮内努, 伊藤祥司, 張紹良, 名取亮, BiCGSTAB(L) 法における L の動的選択について, 日本応用数理学会論文誌, 11(2001), 49-62.
- [9] 森屋健太郎, 野寺隆, 適応的に ℓ を変化させる BiCGStab(ℓ) 法, 情報処理学会論文誌, 40(1999), 2669-2678.
- [10] Reichel, L., Trefethen, L.N., Eigenvalues and Pseudo-eigenvalues of Toeplitz Matrices, *Lin. Alg. Appl.*, 162-164(1992), 153-185.
- [11] Sleijpen, G., van der Vorst, H.A., An overview of approaches for the stable computation of hybrid BiCG methods, *Appl. Numer. Math.*, 19(1995), 235-254.
- [12] Sleijpen, G., Fokkema, D.R., BiCGStab(ℓ) for linear equations involving unsymmetric matrices with complex spectrum, *Electronic Trans. on Numer. Anal.*, 1(1993), 11-32.