

適応的なリスタートを用いた非定常反復法の収束性について

野 寺 隆 † 津野 直人 †

一般共役残差法は、大型で疎な非対称行列系の連立一次方程式の非定常反復解法の一つである。ORTHOMIN(k)法は、この算法のクラスに属するもので、その収束性は記憶する方向ベクトルの数 k に異存する。本稿では適応的なリスタート技法を ORTHOMIN(k) に適用することで、記憶する方向ベクトルの数を増やすことなく、残差ノルムの収束性を格段に向上させる手法について、分散メモリ型の並列計算機 AP1000 を使った数値実験と共に述べる。

Convergence Properties for an Adaptive Restarting Procedure of Nonstationary Iterative Methods

TAKASHI NODERA † and TSUNO NAOTO †

For solving a large and sparse non-symmetric linear systems of equations, generalized conjugate residual method (GCR) is known as one of the nonstationary iterative solvers. One of the main variant of GCR methods is ORTHOMIN(k) algorithm. The convergence of this algorithm is strongly depends on k , which is a number of direction vector. In this paper, applying the adaptive restarting procedure to this algorithm, we are greatly able to improve the convergence of the residual norm. Numerical comparisons of the new algorithms and the original one are given for using a distributed memory parallel machine AP1000.

1. はじめに

科学技術計算に現れる問題には様々なものがあり、偏微分方程式の境界値問題の数値解法も重要な位置をしめている。特に、自己随伴でない楕円型の偏微分方程式の境界値問題では、方程式を有限差分法や有限要素法で離散近似すると、大型で疎(スパース)な係数行列を持つ連立一次方程式

$$Ax = b \quad (1)$$

が得られ、これを効率よく解く算法がひつようとなる。ただし、行列 A は $n \times n$ の非対称な正則行列である。通常、この式(1)の近似解を求める算法は、(1)直接法と(2)反復法に分けることができる。直接法の算法は、行列 A の LU 分解をベースとしており、係数行列が大規模で疎である場合には、直接法は近似解を得るために必要となる計算時間や、さらに算法を実装し計算を行うための記憶領域の面でコスト高となる。そこで算法の実装にかかる記憶領域がすくなくてすみ、なおかつ収束の速い効果的な反復法が求められることになる。

自己随伴な楕円型の偏微分方程式の境界値問題の離散化から得られる対称行列を係数とする連立一次方

式の近似解を求めるためには、共役勾配法¹⁾に代表される有効な算法が存在する。しかし、現在のところ、非対称行列系の算法には、これぞお勧めといわれるものが存在しないのも事実である。その理由の一つに、非対称行列系を解く算法には、ある特殊なケースを除いて対称行列問題の解法として有効な共役勾配法と同等な漸化式で最適な収束性を持つ算法が構成できないことをあげることができる⁵⁾。しかし、共役勾配法とまったく同等な性質を持つ算法ではないが⁸⁾、一般共役勾配法(Generalized Conjugate Gradient Method, GCG法)といわれるクラスに属する算法を構成することができる。通常、これらの算法は、クリロフ部分空間において近似解を構成するので、クリロフ部分空間法と呼ばれることもある。

クリロフ部分空間法には様々な算法が存在するのだが、一般によく知られているものは GMRES 法や ORTHOMIN(k)法、ORTHODIR法、ORTHORES法などや、BiCGStab法などの双共役勾配法(Bi-Conjugate Gradient Method, BCG法)のクラスに属するものである。

本稿では、クリロフ部分空間法の中で一般共役残差法(generalized conjugate gradient method)を取り上げる。一般共役残差法は、毎回の反復で残差ノルムを最小にする算法であり、BCG法やCGS法などの残差ノルムのように激しく振動することはあり得ない

† 慶應義塾大学理工学部
Faculty of Science and Technology, Keio University

解法である。しかし、算法の実装に関していえば、毎回の反復で生成されたすべての方向ベクトルを記憶する必要がある。この算法を実行すると反復が進むに従い記憶領域が増大することになり、実用的な算法とはいえない。そこで、適当にベクトルの数を（例えば、 k 本と）決めてやり、そのベクトルのみを使って算法を続行する ORTHOMIN(k) などが使われることになる。この算法は収束に関していえば、共役勾配法のように最適性をもつものではないが、SOR 法のようにパラメータを決定する必要がなく、最初に方向ベクトルの数を一度決めれば記憶領域も少なくすみ、速く収束してくれば計算コストは非常に安価となる。さらに、この算法は残差ノルムを常に最小にするという性質を持っているが⁴⁾、その収束性に関していえば、近似解を構成するために必要な方向ベクトルを切り捨てて反復を行うので、問題によっては速い収束が得られないことがある。

近年我々が提案した適応的なリスタート技法¹⁶⁾を ORTHOMIN(k) の算法に適用すると、算法の収束性を格段に向上させることが可能となるので、偏微分方程式の離散化から得られた連立一次方程式 (1) の数値実験とともにその適応的なリスタートの有効性について報告する。

2. 一般共役残差法

一般共役残差法は、非対称行列系を解くための反復法である。この算法はクリロフ部分空間 (Krylov subspace method) と呼ばれるクラスに分類されるものであり、各反復によって計算される近似解の列 $\{\mathbf{x}_i\}$ とそれに対応する残差ベクトルの列 $\{\mathbf{r}_i\}$ が

$$\begin{aligned} \mathbf{x}_i &\in \mathbf{x}_0 + \mathcal{K}_i, \\ A^T \mathbf{r}_i &\perp \mathcal{K}_i \end{aligned}$$

という条件を満たすように構成されていく。ここで \mathcal{K}_i は i 次のクリロフ部分空間を表し、

$$\begin{aligned} \mathcal{K}_i &= \mathcal{K}(A, \mathbf{r}_0) \\ &= \text{span}\{\mathbf{r}_0, A\mathbf{r}_0, \dots, A^{i-1}\mathbf{r}_0\} \end{aligned}$$

で定義される。クリロフ部分空間法では、計算による誤差の影響を受けなければ、最大 n 回の反復で真の解を計算するということが保証されているが、そのためにはそれまでの反復で計算された方向ベクトルをすべて記憶しておく必要がある。

2.1 ORTHOMIN(k) 法

前述の GCR 法に対し ORTHOMIN(k) 法は、GCR 法ではすべて記憶しておくべき方向ベクトルを直前の k 本のみに関し、その他の部分では GCR 法とまったく同じように反復を進めていくといった方法である。ORTHOMIN(k) 法の算法を図 1 に示す。GCR 法で必要とされている方向ベクトルを打切りながら反復を進めていくので、最大 n 回で収束するという保証はなくなるが、現実的なメモリ容量や計算時間で実行す

```

1. Choose  $\mathbf{x}_0$ .
2. Compute  $\mathbf{r}_0 = \mathbf{b} - A\mathbf{x}_0$ .
3. Set  $\mathbf{p}_0 = \mathbf{r}_0$ .
4. for  $i = 1, 2, \dots$ 
    4.1  $\mathbf{q}_{i-1} = A\mathbf{p}_{i-1}$ 
    4.2  $\psi_{i-1} = (\mathbf{q}_{i-1}, \mathbf{q}_{i-1})$ 
    4.3  $\alpha_{i-1} = (\mathbf{r}_{i-1}, \mathbf{q}_{i-1}) / \psi_{i-1}$ 
    4.4  $\mathbf{x}_i = \mathbf{x}_{i-1} + \alpha_{i-1}\mathbf{p}_{i-1}$ 
    4.5  $\mathbf{r}_i = \mathbf{r}_{i-1} - \alpha_{i-1}\mathbf{q}_{i-1}$ 
    4.6 for  $j = i - k, i - k + 1, \dots, i - 1$ 
        4.6.1  $\beta_i^{(j)} = -\left\{ (A\mathbf{r}_i, \mathbf{q}_j) \right. \\ \left. + \sum_{h=i-k}^{j-1} \beta_i^{(h)} (\mathbf{q}_h, \mathbf{q}_j) \right\} / \psi_j$ 
    endfor
    4.7  $\mathbf{p}_i = \mathbf{r}_i + \sum_{h=i-k}^{i-1} \beta_i^{(h)} \mathbf{p}_h$ 
    4.8 If converge, escape the loop.
endifor

```

図 1 ORTHOMIN(k) 法の算法

Fig. 1 The Algorithm of ORTHOMIN(k)

ることができ、十分な結果が得られることが多い。また、残差ノルムが単調減少するという GCR 法の性質もそのまま受け継ぐことになる。

2.2 ORTHOMIN-R(k, r) 法

ORTHOMIN(k) 法のバリエーションとして、反復を r 回実行したあとにリスタートを行なう算法が考えられる。このリスタート版を以下 ORTHOMIN-R(k, r) 法と記述することにする。なお、特別な場合として $k = r = 0$ のときに、この算法は最小残差法 (minimal residual method, MR 法) とも呼ばれている⁴⁾。

3. 適応的なリスタート技法

ORTHOMIN-R(k, r) 法は、 r 回目の反復を行なうたびに強制的にリスタートを実行する。そのため、ORTHOMIN(k) 法がよい収束性を示しているときにも強制的なリスタートが悪影響を及ぼし、かえって収束が遅くなってしまうことがある。また問題によっては収束していたものが収束しなくなってしまうようなこともある。このような強制的なリスタートによる悪影響を解消するため、ORTHOMIN(k) 法の収束が減速したときに限りリスタートを行なう適応的なリスタート技法 (adaptive restarted procedure) を導入する。

適応的なリスタート技法は、稲津・野寺¹⁶⁾により ORTHRES 法に導入されたものである。以下、この技法を ORTHOMIN 法に適用した算法のことを ORTHOMIN(k)-AR 法と記述することにする。

残差ノルムの収束性を判断するときに、行列とベク

トルの積や内積といった大きな計算を必要とすると、仮に反復回数が少なくなるとしても計算時間は増えてしまい、満足な結果は得られない。適応的なリスタート技法では、残差ノルムの収束性を判断するときに、余分な計算によるオーバーヘッドや使用するメモリ領域を少なくするため、算法中で使われるスカラの値に注目する。ORTHOMIN(k)法では数回の反復を通して記憶されるスカラは ψ である。いくつかの数値実験を行なった結果、残差ノルムと ψ の間に、例えば図2に示されるような関係がみられる。(この実験に関しては実験1を参照のこと。)このことから次の“予想”を与えることができる。

予想1 ORTHOMIN(k)法は、算法中のスカラ ψ が反復を通じて若干の振動を含めて停滞したふる舞いを示すと、そのときの残差ノルムも同様に停滞したふる舞いを示す。

この予想をもとに、ORTHOMIN(k)法の反復を k 回行なったあとに残差ノルムの収束性を調べ、必要であればリスタートを行ない、リスタートの必要がなければそのまま ORTHOMIN(k)法の反復を k 回続ける、といった算法を ORTHOMIN(k)-AR法として定義する。残差ノルムの収束性の判断、すなわち ψ が停滞しているかどうかは、 ψ の変動係数

$$\frac{\sigma_{n-1}[\psi]}{\bar{x}[\psi]}$$

をもとにする。ここで σ は標準偏差、 \bar{x} は算術平均を表し、記憶してあるすべての ψ から値を計算する。ORTHOMIN(k)-AR法はこの変動係数が与えられた安定基準 ϵ よりも小さいときに ψ が停滞しているとする。すなわち、この条件を満たすと残差ノルムが停滞しているとみなしてリスタートを実行する。

適応的なリスタート技法が必要とされる計算は ψ の変動係数にかかるものだけで、非常に低コストで実行できる。分散型メモリをもつ並列計算機に実装することを考えても、通信を必要としないためボトルネックにはならない。また、余分なメモリ領域は一切必要としない。

以上のような適応的なリスタート技法を導入することにより、ORTHOMIN-R(k, r)法で見られた強制的なリスタートの及ぼす悪影響が解消され、種々の問題に対してより速い収束がみられることが期待される。

4. 数 値 例

矩形領域 $\Omega = [0, 1] \times [0, 1]$ 上で、2次元の楕円型偏微分方程式の境界値問題

$$-u_{xx} - u_{yy} + Du_x = G(x, y) \quad (2)$$

$$u(x, y)|_{\partial\Omega} = 1 + xy$$

に対して解を $u(x, y) = 1 + xy$ と設定し、右辺 $G(x, y)$ を決定する⁹⁾。また、 D は実数の定数であるとする。この問題を5点中心差分により離散化し、連立一次方

程式(1)を決定し、以下に示す数値実験を行なった。ここで、離散化によるメッシュ幅を h で表すものとする。

各算法の初期近似解は零ベクトルとし、収束条件は相対残差ノルムで

$$\frac{\|r_i\|}{\|r_0\|} \leq 1.0 \times 10^{-12} \quad (3)$$

とした。すべての計算は分散メモリ型の並列計算機 AP1000 (富士通)で64個のプロセッサを用い、倍精度計算で実行した。

実験1 [ORTHOMIN(k)法と ψ のふる舞い]

- メッシュサイズ: 256×256
- D の値: $Dh = 2^{-1}$ となるように決定する。
- k の値: $k = 5$
- 最大反復回数: 3000 回

このような条件の元で、ORTHOMIN(k)法の残差ノルムとその算法中のスカラ ψ のふる舞いを調べることとする。

実験1の結果を図2に示す。 D の値を変えた他の問題に対しても図2と同じような、 ψ が停滞しているときには残差ノルムも停滞する、といったふる舞いがみられた。この結果より先の命題を与えている。

次に、適応的なリスタート技法を導入した算法と、その他の算法を様々な問題に対して適用し、結果を比較することとする。

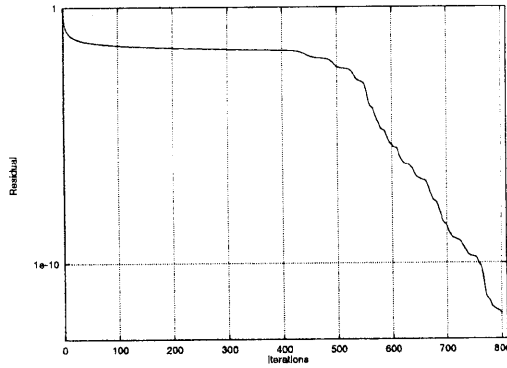
実験2 [各算法の比較]

- メッシュサイズ: 256×256
- D の値: Dh がある決められた値となるように決定する。
- 算法:
 - (1) ORTHOMIN(5)
 - (2) ORTHOMIN-R(5,50)
 - (3) ORTHOMIN-R(5,100)
 - (4) ORTHOMIN(5)-AR
- 最大反復回数: 3000 回

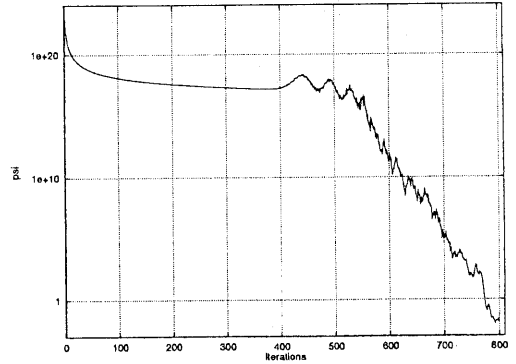
上記の条件で各算法を用いて問題を解き、その結果を比較することとする。なお、適応的なリスタート技法の安定基準である ϵ の値は $\epsilon = 4.0 \times 10^{-3}$ とする。

安定基準の値は、実験1のような ψ のふる舞いをもとにいくつかの点で変動係数を計算し、その値を参考にして定めた。実験2の結果を表1、表2、図3に示す。表1では各算法を実行時間で比較し、表2では ORTHOMIN(k)-AR法で適応的なリスタートの発生した回数を示している。なお、表1の中で“—”は最大反復回数内で収束収束基準が満たされなかったことを表す。

これらの結果より、すべての問題に対して ORTHOMIN(k)-AR法は ORTHOMIN(k)法よりも速く収束していることがわかる。また ORTHOMIN-R(k, r)法と比べてみると、期待通りに ORTHOMIN(k)-AR法



(a) 残差ノルムのふる舞い



(b) ψのふる舞い

図2 残差ノルムとψ
Fig. 2 The residual norm and ψ

表1 各算法の実行時間 (sec)
Table 1 The executive time (sec) of various algorithms

算法	Dh の値							
	2 ⁻²	2 ⁻¹	2 ⁰	2 ¹	2 ²	2 ³	2 ⁴	2 ⁵
ORTHOMIN(5)	512.74	216.30	231.18	261.57	363.27	528.43	—	—
ORTHOMIN-R(5,50)	300.49	219.98	202.91	202.72	211.23	225.16	227.79	244.10
ORTHOMIN-R(5,100)	315.43	211.56	206.06	216.51	239.97	241.06	230.92	301.00
ORTHOMIN(5)-AR	351.52	207.57	212.56	208.20	215.16	211.71	229.67	248.20

表2 適応的なリスタート技法によるリスタートの回数
Table 2 The restarting counts for using the adaptive restart technique

	Dh の値							
	2 ⁻²	2 ⁻¹	2 ⁰	2 ¹	2 ²	2 ³	2 ⁴	2 ⁵
リスタート回数	54	48	99	25	26	15	12	16

で収束が速くなっている例もみられるが、ほぼ同程度の速さとなっており、逆に遅くなっている例もみられる。これは、ORTHOMIN-R(k, r)法の強制的なリスタートの影響が悪い方向ではなく良い方向に働いているためと考えられる。しかし、強制的なリスタートの影響が良い方向に働くか、悪い方向に働くか、というのは問題に左右されどちらになるかをあらかじめ知ることは不可能である。種々の問題に対する安定性を考えるとORTHOMIN(k)-AR法の方が安定しているといえる。

なお、この例の誤差ノルムの収束性を図4に示した。

5. おわりに

一般共役残差法に適応的なリスタート技法を用い

て残差ノルムの収束性について調べてきた。我々の数値実験によれば、適応的なリスタート技法を用いたORTHOMIN(k)法はかなりの好成績をあげていることがわかる。

参考文献

- 1) Hestenes M. R. and Stiefel E.: *Methods of conjugate gradients for solving linear systems*, J. Res. Nat. Bur. Standards, Vol. 49, pp. 409-435 (1952).
- 2) Vinsome, P. K. W.: *Orthomin, an iterative method for solving sparse sets of simultaneous linear equations*, In Proceedings of the Fourth Symposium on Reservoir Simulation, Society of petroleum Engineers of AIME, pp. 149-159

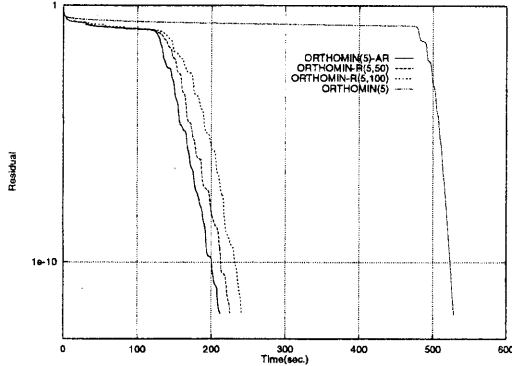


図3 $Dh = 2^3$ の場合の残差ノルムの収束
Fig. 3 The convergence behavior of residual norm for $Dh = 2^3$

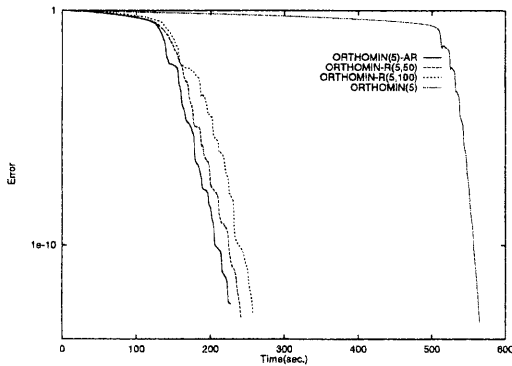


図4 $Dh = 2^3$ の場合の誤差ノルムの収束
Fig. 4 The convergence behavior of error norm for $Dh = 2^3$

(1976).

- 3) Young, D. M. and Jea K. C.: *Generalized conjugate gradient acceleration of nonsymmetrizable iterative methods*, Linear Algebra and Appl., Vol. 34, pp. 159-194 (1980).
- 4) Eisenstat, H. C., Elman, H. C and Schultz, M. H: *Variational iterative methods for non-symmetric systems of linear equation*, SIAM J. Numer. Anal. Vol. 20, pp. 345-357 (1983).
- 5) Faber, V. and Manteuffel, T. A.: *Necessary and Sufficient conditions for existense of a conjugate gradient method*, SIAM Numer. Anal. Vol. 21, pp. 352-362 (1984).
- 6) Saad, Y. and Schultz, M. H.: *GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems*, SIAM J.Sci.Stat. Comp., Vol. 7, No. 3, pp. 856-869 (1986).
- 7) 名取, 野寺: 大型疎行列計算における反復解法, 情報処理, Vol. 28, No. 11, pp. 1452-1459 (1987).
- 8) Van der Vorst H. A.: *Bi-CGSTAB: A fast*

and smoothly converging variant of Bi-CG for the solution of non-symmetric linear systems, SIAM J. Sci. Stat. Comput., Vol. 13, pp. 631-644 (1992).

- 9) Joubert W.: *Lanczos Methods for the solution of non-symmetric systems of Linear Equations*, SIAM. J. Matrix Anal., Vol. 13, No. 3, pp. 926-943 (1992).
- 10) Sleijpen G. L. G. and Fokkema D. R.: *BiCG-STAB(ℓ) for linear equations involving un-symmetric matrices with complex spectrum*, ETNA, Vol. 1, pp. 11-32 (1993).
- 11) R. Weiss: *Properties of generalized conjugate gradient Methods*, Numerical Linear Algebra with Applications, Vol. 1, No. 1, pp. 45-63 (1994).
- 12) 野寺: オペレータ係数法について, 京都大学数理解析研究所講究録 No. 880 (1994).
- 13) Bruaset A. M.: A survey of preconditioned iterative methods, *Pitman Research Notes in Math* (1995).
- 14) Schönauer, W. and Weiss, R.: *An engineering approach to generalized conjugate gradient methods and beyond*, Applied Numerical Mathematics, Vol.19, No.3, pp.175-206 (1985).
- 15) 野寺, 稲津: リスタートを用いた疑似残差法の収束性について, 情報処理学会論文誌, Vol. 37, No. 6, pp. 1237-1240 (1996).
- 16) 稲津, 野寺: 適応的なリスタートを用いた疑似残差法, 情報処理学会論文誌 9月号掲載予定 (1996).