

共役勾配法の未解決問題

慶應義塾大学理工学部 野寺 隆 (Takashi NODERA)

1 はじめに

1952年に Hestenes と Stiefel [1] によって提案された共役勾配法 (CG 法ともいう) は, 大型の連立 1 次方程式

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

の反復解法として確固とした地位を築いてきた。その算法は, 単純でエレガントであるので, 様々な科学技術計算の分野で利用されている。

共役勾配法の算法は, 数学的に非常にエレガントで美しい漸化式の形をしているのであるが, 丸め誤差の影響を非常に受けやすいという欠点を持っている。この算法の欠点が災いして, 1955 年後半から 1960 年代を通して余り利用されなくなった。しかし, 1971 年, Reid [3] によって大型の連立 1 次方程式の近似解法として有効で有ることが示されたことで, 再び脚光をあびてきた。ただし, この時点では算法の改良が行われたわけではなく, 共役勾配法の局所的な直交性に着目した解析により, 算法のすべてに渡って直交性が保持される必要がなく, 局所的な直交性が保持できればかなりよい近似解が構成できる ([4] を参照)。

その後, 行列の前処理の研究が進み, 共役勾配法と前処理の組み合わせを用いることにより, 残差ノルムの収束性を向上させることがわかり, 第 3 次の共役勾配法のブームが起った。行列の前処理として Meijerink et al. [5] によって提案された不完全コレスキー分解を併用した共役勾配法は ICCG 法と呼ばれ, 様々な分野で用いられるようになった。また, 従来の直交性による有限回の反復で収束することよりも, Krylov 部分空間における直交する行列多項式を用いた収束性の解析が主流となった。

1952 年から現在まで, 共役勾配法に関連する数百編の論文が出版されている。今日に至っても共役勾配法の根本的な算法の持つ欠点は, 提案された時点から依然として解決されているとは言えない (Greenbaum [2])。

しかし, 共役勾配法はいくつかの未解決な欠点を持つものの大規模で対称な正定値行列を係数とする連立 1 次方程式の近似解を得るための算法として, これ以上の算法はないと言われている。

非対称行列を係数とする連立1次方程式の解法として共役勾配法の算法はどんな位置に
 いるのであろうか。非対称行列系の算法として共役勾配法をそのまま用いることは、でき
 ないことは言うまでもない。基本的に正規方程式を作り、それに対称行列系の共役勾配法
 の算法を適用することになるが、元の行列の条件数を自乗したものを解くことになるので、
 丸め誤差の影響を受け近似解の精度があまりあがらない場合もある。現在では、このよう
 な手法を用いることは、主流ではない。そこで、過去40年の間、非対称行列系に適用可能な
 共役勾配法に似た算法の導出の研究が行われてきた。1984年、Faber と Manteuffel [6] に
 よって非対称行列系の反復解法として共役勾配法と同等の直交性を保持する算法は存在し
 ないことが証明された。この事実により、非対称系の算法の開発は現在も続行中と言える
 が決定的な算法は今だ現れていない。

現在までに、提案された主な非対称行列系の反復解法をまとめると表1のようになる。
 1986年にSaad と Schultz [8] ににより提案されたGMRES法は、我が国ではあまり利用
 されてとはいえないが、欧米では標準的な算法として用いられている。その理由の1つは、
 GMRES法は真の残差を常に最小にする性質を持つからである。わが国では、どちらかと言
 うと1992年、van der Vorst [12] ににより提案されたBiCGSTAB法が主流であるように思わ
 れる。この算法がよく用いられている理由の1つは、算法の明解さにあるように思われる。

本稿では、新しい算法ではないが、旧来から知られているテクニックを利用することで、い
 ままで提案されてきた算法の残差の収束性を向上させる技法について報告することにする。

2 一般共役勾配法

対称行列系に対する共役勾配法の算法が提案されて以来、過去40数年にわたり非対称行
 列系の近似解を求める算法の研究が行なわれてきたのは言うまでもない。

一般共役勾配法 (Generalized Conjugate Gradient Method, GCG法ともいう) と呼ばれ
 る算法が過去十数年にわたっていくつか提案されている ([11,10,9])。本稿では、Weiss [9]
 に従って、一般共役勾配法の算法を次のように記述することができる。

[一般共役勾配法]

- (1) 任意の初期値 x_0 を選び、初期残差 $r_0 = Ax_0 - b$ を計算する。ただし、行列 P は右側
 からの前処理行列とする。
- (2) 次の反復を繰り返す。 ($k = 1, 2, 3, \dots$)
 以下の条件 (i) または (ii) のどちらかを満足する

$$x_{k+1} = x_0 + K_k(PA, Pr_0),$$

表 1 非対称行列系の反復解法

正規方程式に基づく算法	
算法	論文の著者 及び 発表年
CGN (CGNR)	Hestenes and Stiefel '52
CGNE	Craig '55
LSQR	Paige and Saunders '82
Orthogonalization 法	
算法	論文の著者 及び 発表年
GCG 法	Concus and Golub '76, Widlund '78
—	Axelsson '79, '80
ORTHOMIN	Vinsome '76
ORTHORES	Young and Yea '80
ORTHODIR	Young and Yea '80
FOM	Saad '81
GCR	Elman '82, Eisenstat et al. '83
GMRES	Saad and Schultz '86
Biorthogonalization 法	
算法	論文の著者 及び 発表年
BIOMIN = BCG	Lanczos '52, Flecher '76
BIORES = BO	Lanczos '50, Jea and Young '83
BIODIR	Jea and Young '83
BIOMIN ² =CGS	Sonneveld '89
BIORES ²	Gutknecht '90
BIODIR ²	Gutknecht '90
BiCGSTAB	Van der Vorst '90
QMR	Freund '90
他の方法	
算法	論文の著者 及び 発表年
USYMLQ	Saunders, Simon, and Yip '88
USYMQR	Saunders, Simon, and Yip '88

$$r_{k+1} = Ax_{k+1} - b$$

となる x_{k+1} と r_{k+1} を $\|r_{k+1}\|_Z$ が収束するまで計算する。

(i)

$$r_{k+1} Z r_{k+1-i} = 0, \quad \text{for } i = 1, \dots, \sigma, \quad (2)$$

ただし, Z は任意の正則行列とする。

(ii)

$$\|r_{k+1}\|_Z = \min_{\mu_{k+1-\sigma}, \dots, \mu_k} \|r_{k+1-\sigma} + \sum_{i=1}^{\sigma} \mu_{k+1-i} A P_{k+1-i}\|_Z, \quad (3)$$

ただし, Z は対称で正定値行列とする。

この算法に現れる σ はオーダと呼ばれ, σ の選択方法により一般共役勾配法は計算方法に違いが生じる。一般共役勾配法は以下の 4 つの算法に分類できる。

- もし $\sigma = k + 1$ ならば, 新しい残差ベクトルの計算に今までに計算したすべてのベクトルを使用することになるので, この方法を厳密版 (exact version) と呼んでいる。
- もし $\sigma = k \bmod \sigma_{res} + 1$ ならば, σ_{res} 回の反復の後で周期的に再出発 (リスタートともいう) する再出発版 (restarted version) と呼ぶ算法になる。
- もし σ_{max} を固定し $\sigma = \min(k + 1, \sigma_{max})$ とすると, この方法は σ_{max} 個の残差のみを利用する打切版 (truncated version) と呼ぶ算法になる。
- 前述の打切版をある時点で再出発させる算法を結合版 (combined version) と呼ぶ。

ただし, σ は反復回数 k に従属するものであり, 実際は σ_k とする所を σ と記述している。また, もし $\sigma_{res} = \sigma_{max} = 1$ ならば, $\sigma = 1$ となり打切版と再出発版は同じものになる。

任意のベクトル $y \in R^n$ のノルム $\|y\|$ は, $\|y\|_Z = \sqrt{y^T Z y}$ で定義し, 行列 Z はポジティブリアル (positive real) であるものとする。ただし, ポジティブリアルとは, Z の対称部分が正定値行列である。通常, Z は単位行列がユニタリ行列とするのだが, 単位行列とした場合, ユークリッドノルム $\|y\|$ を最小化することになる。もちろん, この Z の選択によって様々な算法に変化することは言うまでもない。

残差ベクトル r_{k+1} は Krylov 空間 $K_{k+1}(AP, r_0)$ において構成されるので, 一般共役勾配法は Krylov 部分空間法の 1 つである。残差ベクトルに関して, 以下の等式が成立する。

$$r_{k+1} = \prod_1^{k+1} (AP) r_0 = \sum_{i=1}^{k+1} \nu_{i,k} (AP)^i r_0 + r_0 \quad (4)$$

$$e_{k+1} = \prod_1^{k+1} (AP) e_0 = \sum_{i=1}^{k+1} \nu_{i,k} (AP)^i e_0 + e_0 \quad (5)$$

ただし, $e_{k+1} = x_{k+1} - \bar{x}$ は誤差ベクトル (\bar{x} は真の解) であり, \prod_1^{k+1} は $k+1$ 次の定数係数を持つ多項式であり, $\prod_1^{k+1}(0) = 1$ を満足する。

もし, 行列 Z が対称で, 正定値行列であるならば, (2) 式は次式と同値になる。

$$\|\bar{r}_{k+1}\|_Z = \min \quad (6)$$

ただし, \bar{r}_{k+1} は疑似残差ベクトルと呼ばれる。(6) 式を満たす算法は, Z ノルムに従って疑似残差を最小にする一般共役勾配法の 1 つである。すなわち, (2) 式が成立すれば, この算法を疑似残差法 (pseudo residual method, PRES 法) と呼んでいる ([9])。

ここで、もし (3) 式を満たすならば、その算法は真の残差を最小にするので、共役残差 (conjugate residual) タイプの算法か、MR (minimal residual) 法となる ([7])。

次に前処理行列 P に関して、 $P = I$ ならば基本算法と呼ばれるものになり、 $P \neq I$ ならば前処理付きの算法となる。

3 疑似残差法

前述の一般共役勾配法において、各反復で $\|\bar{r}_{k+1}\|_Z$ を最小にするのが疑似残差法である。

[疑似残差法 (PRES)]

- (1) 任意の初期値 x_0 を選び、初期残差 $r_0 = Ax_0 - b$ を計算する。ただし、 Z は任意の対称正定値行列とする。
- (2) 以下の反復を $\|r_{k+1}\|_Z$ が収束するまで繰り返す。 ($k = 1, 2, 3, \dots$)

$$d_k = Pr_k, \quad (P: \text{右側からの前処理行列}) \quad (7)$$

$$\alpha_{i,k} = -\frac{r_{k+1-i}^T Z A d_k}{r_{k+1-i}^T Z r_{k+1-i}} \quad \text{for } 1 \leq i \leq \sigma_k, \quad (8)$$

$$\bar{r}_{k+1} = A d_k + \sum_{i=1}^{\sigma_k} \alpha_{i,k} r_{k+1-i} \quad (\bar{r}_{k+1}: \text{疑似残差}) \quad (9)$$

$$\phi_k = 1 / \sum_{i=1}^{\sigma_k} \alpha_{i,k} \quad (10)$$

$$r_{k+1} = \phi_k \bar{r}_{k+1} \quad (11)$$

$$x_{k+1} = \phi_k (d_k + \sum_{i=1}^{\sigma_k} \alpha_{i,k} x_{k+1-i}) \quad (12)$$

疑似残差法は、右側からの前処理行列 P 、任意な正定値行列 Z 、オーダー σ_k の選により、一般共役勾配法の場合に述べた幾つかの算法のバリエーションを考えることができる。もっとも簡単な算法は、行列 Z を単位行列 I と置くことで、ORTHORES 法が得られる。この算法においても、表 2 に示すようなバリエーションがある。また、 $Z = A^T$ と置くと、ATPRES と呼ばれる算法が得られる。この他、行列 Z をいろいろ変えることにより、様々な疑似残差法の算法を導出できる。

ORTHORES 法は基本的に疑似残差を最小とするので、元の残差を最小にはしない。数値実験から、真の残差ベクトルのノルムの収束性を観察すると振動することがわかる。

表 2 前処理なしの疑似残差法

算 法	σ_k
ORTHORES-E	$k+1$
ORTHORES-R(σ_{res})	$(k \bmod \sigma_{\text{res}}) + 1$
ORTHORES-T(σ_{max})	$\min(k + 1, \sigma_{\text{max}})$
ORTHORES-C($\sigma_{\text{max}}, \sigma_{\text{res}}$)	—

表 3 に非対称行列系の算法を用いた数値実験の結果をまとめた。使用した問題の詳細は、文献 [13] を参照してほしい。また、すべての計算は、富士通の並列計算機 AP1000 を用いて行なっている。

表 3 より、BICGSTAB 法や GMRES(5) 法はかなりよい収束性を示している。しかし、残差ノルムの収束性にムラが有り、発散や停滞が起る場合もある。当然、ORTHORES-T(5) 法や ORTHORES-R(5) 法でも残差ベクトルの発散や振動が起る場合もある。しかし、ORTHORES-T(5) 法をある時点 (例えば、50 回の反復の後) で周期的に再出発させる ORTHORES-C(5,50) 法は、ここで用いた例題の全てに対して残差ノルムが収束していることがわかる。

4 再出発

前述の ORTHORES-C(5,50) 法のように打切版を適当な時点で残差を再計算して、再出発 (リスタート) させると、残差ノルムの収束性を格段に向上させることができる (図 1 を参照)。しかし、誤った再出発を行うと、収束を遅くする場合もないわけではない。また、残差ノルムが収束状態に入っている場合には、再出発を行っても残差ノルムの収束性を向上させる効果が現れないことも多い。使用する算法を再出発させる時期が重要なポイントとなるが、周期的に再出発を行なってもかなりの好成绩で残差ノルムを効果的に収束させることができる。

5 おわりに

非対称行列系の近似解を求める一般共役勾配法の問題点について考えてきた。疑似残差法は、真の残差ノルムを最小にはしないが、誤差ノルムを最小にする利点を持っている。し

表 3 非対称行列系の算法による数値実験の結果

算法	a	b	反復回数	時間 (sec)	最終残差ノルム $\ r_k\ $
ORTHORES-T(5)	3	5	oscillate	—	4.4×10^{-4}
	3	50	diverge	—	3.9×10^{-4}
	3	500	536	66	1.6×10^{-14}
	3	5000	487	60	3.5×10^{-14}
ORTHORES-R(5)	3	5	oscillate	—	4.1×10^{-2}
	3	50	5288	477	1.3×10^{-13}
	3	500	544	49	4.8×10^{-15}
	3	5000	1417	127	7.0×10^{-14}
ORTHORES-C(5,50)	3	5	3092	381	1.6×10^{-13}
	3	50	1219	152	5.0×10^{-14}
	3	500	491	59	2.8×10^{-14}
	3	5000	606	75	2.7×10^{-14}
BiCGSTAB	3	5	843	99	4.7×10^{-13}
	3	50	521	61	2.8×10^{-12}
	3	500	570	67	4.1×10^{-7}
	3	5000	diverge	—	1.0×10^{29}
GMRES(5)	3	5	stagnate	—	9.2×10^{-13}
	3	50	594	280	2.2×10^{-13}
	3	500	113	53	2.4×10^{-14}
	3	5000	175	59.7	6.9×10^{-14}

かし、現実の計算では誤差ノルムの評価は不可能である。

メッセージパッシング型の並列計算機を用いる場合には、従来型の前処理はある特殊な例を除いてあまり意味をなさない。そこで近年、前処理なしのシンプルな算法だけを利用して計算を行なうことが多い。その場合、算法の収束性を向上させる手法の一つとして、算法の再出発が重要な意味を持つようになった。疑似残差法に関して、適応的なリスタートを行なう技法の研究も行なわれている ([15])。

参考文献

- [1] M. R. Hestenes and E. Stefel, "Methods of conjugate gradients for solving linear systems," J. Res. Nat. Bur. Standards, Vol. 49 (1952), pp. 409-435.
- [2] A. Greenbaum, "The Lanczos and conjugate gradient algorithms in finite precision arithmetic," Proc. of the Cornelius Lanczos International Centenary Conference, (1994), pp. 49-60.

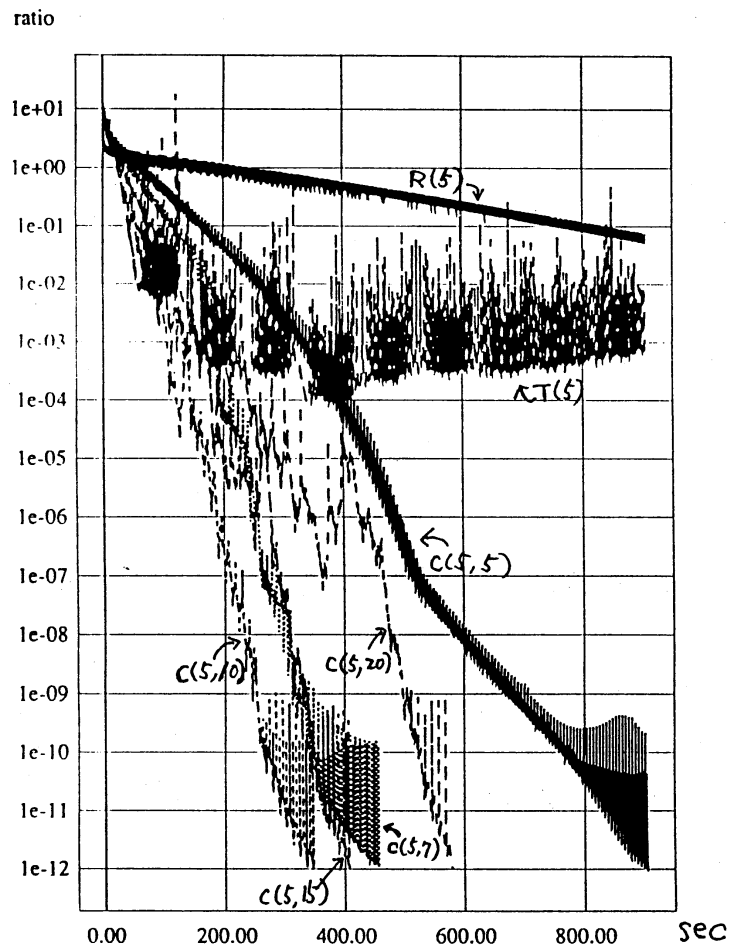


図1 残差ノルムの収束状況 v.s. 収束時間 (sec)

- [3] J. K. Reid, "On the method of conjugate gradients for the solution of large sparse systems of linear equations," (ed. J. K. Reid) Academic Press, (1971), pp. 231-253.
- [4] T. Nodera, "Data flow analysis of orthogonal properties on the conjugate gradient and Lanczos algorithm," Numer. Math. and Appl. (eds. R. Vichnevetsky, J. Vignes), Elsevier Science Publishers B. V., North-Holland (1986), pp. 95-103.
- [5] J. A. Meijerink and H. A. Van der Vorst, "Iterative solution method for linear systems of which the coefficient matrix is a symmetric M-matrix," Math. Comp. Vol. 31 (1977), pp. 148-162.
- [6] V. Faber and T. A. Manteuffel, "Necessary and sufficient conditions for the existence of a conjugate gradient method," SIAM J. Numer. Anal., Vol. 21 (1984), pp. 352-362.
- [7] S. C. Eisenstat, H. C. Elman and M. H. Schultz, "Variational iterative methods for nonsymmetric systems of linear equations," SIAM J. Numer. Anal., Vol. 20 (1983), pp. 345-357.

- [8] Y. Saad and M. H. Schultz, *GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems*,” SIAM J. Sci. Stat. Comput., Vol. 7 (1986), pp. 856-869.
- [9] R. Weiss, *Properties of generalized conjugate gradient methods*, Numer. Lin. Alg. with Appl. Vol. 1(1) (1994), pp. 45-63.
- [10] L. A. Hagemman and D. M. Young, “*Applied iterative methods*,” Academic Press (1981).
- [11] Y. Saad, “*Krylov subspace methods for solving large unsymmetric linear system*,” Math. Comp. 37, (1981), pp. 105-126.
- [12] H. A. van der Vorst, “*Bi-CGSTAB: A fast and smoothly converging variant of Bi-CG for the solution of nonsymmetric linear systems*,” SIAM J. Sci. Stat. Comput., Vol. 13 (1992), pp. 631-644.
- [13] 稲津, 野寺, 非対称行列系に対する疑似残差法と前処理について, 情報研報 Vol. 59 (1995), pp. 27-34.
- [14] 稲津, 野口, 野寺, BiCGStab 系のいろいろなアルゴリズム, 第 2 回 NEC・HPC 研究会論文集 (1995) pp. 22-23.
- [15] 稲津, 疑似残差法の並列化と適応的なリスタート技法, 慶應義塾大学理工学研究科数理科学専攻 修士論文 (1996).