

並列非線型共役勾配法アルゴリズムとその性能評価

西 田 晃^{†1}

大規模疎行列の固有値を数値的に求める場合、Krylov 部分空間法は最も有力な手法として知られている。Krylov 部分空間法には複数の手法があり、一部の解法は適切な前処理手法と組み合わせることによってより高速に固有値を計算することができるが、前処理の効果、計算精度に与える影響などについては未解明な点が多い。本稿では、固有値解法において必要となる非線型共役勾配法の並列アルゴリズム、前処理手法等について検討するとともに、複数の計算機アーキテクチャ上での性能評価結果について報告する。

Algorithm of Parallel Nonlinear Conjugate Gradient Method and its Performance Evaluation

AKIRA NISHIDA ^{†1}

When we need to compute some eigenvalues of a large sparse matrix numerically, the Krylov subspace method is known to be the optimal choice. Combining appropriate preconditioners with the some Krylov subspace method, you can compute the eigensolutions more efficiently, although there are many unsolved problems on the preconditioners. This study discusses the parallel algorithm of the nonlinear conjugate gradient method, which is required for eigenvalue problems, and their preconditioning, with their performance results on some computer architectures.

1. 背 景

大規模疎行列の固有値を数値的に求める場合、いくつかの解法を考えることができ、冪乗

法、逆反復法、Lanczos 法、Davidson 法、共役勾配法等の Krylov 部分空間法が提案されている。多くの問題では、特定の範囲にある数個の固有値（及び固有ベクトル）を求めればよい。ため、疎行列性を保存するこれらの解法は、微分方程式の離散化に伴う大規模計算などにおいて広く利用されている。

本研究では、平成 14-19 年度科学技術振興機構 CREST 事業の一環として、反復解法ライブラリ Lis^{*1} を開発、配布し、様々な並列計算機上で大規模な線型方程式を解くための環境を提供している。また平成 20 年度には九州大学情報基盤研究開発センターにおいて固有値解法の実装を行い、同年 11 月より疎行列固有値解法に対応した新版を公開している。本稿では、Lis において実装している並列非線型共役勾配法のアルゴリズム、前処理手法等の詳細について検討するとともに、複数の計算機アーキテクチャ上での性能評価結果について報告する。

2. 固有値解法における前処理

本稿では、標準固有値問題

$$Ax = \lambda x \tag{1}$$

について考える。固有値 λ が既知であると仮定すると、これに対応する固有ベクトルは

$$(A - \lambda)x = 0, \quad x \neq 0 \tag{2}$$

を解くことにより求める。すなわち、固有値解法における理想的な前処理行列 T は $A - \lambda$ の逆行列であり、実際には未知の λ を Ritz 値で置き換えた前処理行列を考えることもできる。一般にこのように前処理行列を取ると不定値となるため、 T が対称正定値でなければならない場合には

$$T \approx A^{-1} \tag{3}$$

と取る。特に連立一次方程式 $Ax = f$ の解法が与えられている場合には、前処理の計算も容易である。このように定めた T に関して、

$$w^{(i)} = Tr \tag{4}$$

すなわち

$$T^{-1}w^{(i)} = r \tag{5}$$

^{†1}九州大学情報基盤研究開発センター

Research Institute for Information Technology, Kyushu University

^{*1}<http://www.ssisc.org/lis/>

を解くのが、固有値解法における前処理の位置付けである¹⁾。

3. Krylov 部分空間法

以下では、Krylov 部分空間法の原理とその具体的な解法について述べる。

Krylov 部分空間法に属する最も単純な解法は、冪乗法、及びその拡張である部分空間反復法である。冪乗法では、サイズ n の大規模行列 A の固有値を絶対値最大のものから^{*1} 求めるため、これを次元 $\nu \ll n$ のベクトル部分空間 G_l 上への直交射影に関する固有値問題に近似する。射影を表す行列を π_l とすれば、問題は G_l において

$$\pi_l(Ax_l - \lambda_l x_l) = 0 \quad (6)$$

を満たす近似固有対

$$\lambda_l \in C, \quad 0 \neq x_l \in G_l \quad (7)$$

の計算に帰着される。部分空間 S を生成する r 個の独立なベクトルを

$$U = [u_1, \dots, u_r], \quad 1 \leq r < n \quad (8)$$

とすると、 G_l 内の正規直交基底の選び方によって、

$$S_l = A^l S, \quad l = 1, 2, \dots \quad (9)$$

とすれば $r = 1$ の場合は冪乗法、また $r > 1$ ならば部分空間反復法となる。同様にして、逆反復法は A ではなく、 A^{-1} をベクトルに適用することにより、 A の固有値を絶対値最小のものから求める。

$\{u_1, \dots, u_r\}$ によって生成される Krylov 部分空間を

$$\mathcal{K}_l = \text{lin}(S, AS, \dots, A^{l-1}S) \quad (10)$$

とすると、 A が Hermite 行列ならば Lanczos 法 ($r = 1$)、ブロック Lanczos 法 ($r > 1$) が得られる。これらの方法では、写像が三重対角行列 H_l で表される \mathcal{K}_l の正規直交基底 $\{v_i\}_1^l$ を求め、 H_l の固有対を元の問題の近似解として計算する。ただし、解の精度、また近似固有ベクトルを計算する際には、元の問題に対する逆反復法の適用が必要となる点に注意する必要がある。

Davidson 法では、正規直交基底 $\{v_i\}_1^k$ で張られる部分空間 $\mathcal{K} = \text{span}\{v_1, \dots, v_k\}$ 上で行列 A の Ritz 値 θ_k 及び Ritz ベクトル u_k を計算する際に、残差 $r = Au_k - \theta_k u_k$ に関する

修正方程式

$$M_k t = r, \quad M_k = D_A - \theta_k I \quad (11)$$

を解いて u_k を更新し、 \mathcal{K} の次元を拡張する。 D_A は A の対角成分である。具体的には、 t を \mathcal{K} と直交化して v_{k+1} を得る。 $V_{k+1} = [v_1, \dots, v_{k+1}]$ と置けば、新しい Ritz 対 (θ_{k+1}, u_{k+1}) は行列

$$H_{k+1} = V_{k+1}^* A V_{k+1} \quad (12)$$

の固有対として計算されることになる。 $M_k = I$ の場合には Lanczos 法と同一になるので、Davidson 法は修正を伴う Lanczos 法の一つと考えることができる。なお、

$$M_k^{-1} \approx (A - \theta_k I)^{-1} \quad (13)$$

を残差ベクトル r に対する前処理行列と考えると、この方法では θ_k に対応する近似固有ベクトル u_k の方向の成分を増幅させることになり、対角優位な行列の最大固有値を求める場合にしか顕著な効果が得られないことが分かっている²⁾。このため、Jacobi-Davidson 法では u_k の直交補空間から更新のための成分を取り出すことによってこの問題点を解消している³⁾。Davidson 法も、固有ベクトルが必要な場合には得られた固有値をもとに逆反復法等により計算することになるため、余分な計算量を要する。

実対称行列 A に関する固有値問題

$$Ax = \lambda x \quad (14)$$

の最小固有値、またはこれと同値な問題

$$x = \mu Ax, \quad \mu = 1/\lambda \quad (15)$$

の最大固有値の計算を

Rayleigh 商

$$\mu(x) = \frac{x^T x}{x^T Ax} \quad (16)$$

の極値問題に帰着し、最急勾配方向が

$$\nabla \mu(x) \equiv g(x) = \frac{2(x - \mu Ax)}{x^T Ax} \quad (17)$$

であることから、適当な係数 α_i 、修正方向

$$p_i = -g_i + \beta_{i-1} p_{i-1}, \quad \beta_{i-1} = \frac{g_i^T g_i}{g_{i-1}^T g_{i-1}} \quad (18)$$

などを用いることにより、非線型共役勾配法を導くことができる⁴⁾。これは 1951 年に提案さ

*1 以下単に最大固有値と書く。

れた手法^{5),6)}であるが1980年代以降の研究⁷⁾⁻⁹⁾により、以下に述べる前処理と組み合わせることによって高速に固有値を計算できることが知られている。

前処理付固有値解法のアルゴリズムは、前処理行列 $T \approx A^{-1}$, TA に関する m_k 次多項式 $P_{m_k}(TA)$ を用いて以下のように書くことができる。

- (1) 初期ベクトル $x^{(0)}$ を選択する
- (2) m_k 回の反復により $x^{(k)} = P_{m_k}(TA)x^{(0)}$ を計算する
- (3) $\mu^{(k)} = (x^{(k)}, x^{(k)}) / (x^{(k)}, Ax^{(k)})$ を計算する

前処理付共役勾配法の反復は、適当な初期ベクトル $x^{(0)}$ と対応する修正ベクトル $p^{(0)} = 0$ を用いて、

$$\mu^{(i)} = (x^{(i)}, x^{(i)}) / (x^{(i)}, Ax^{(i)}) \quad (19)$$

$$r = x^{(i)} - \mu^{(i)} Ax^{(i)} \quad (20)$$

$$w^{(i)} = Tr \quad (21)$$

$$x^{(i+1)} = w^{(i)} + \tau^{(i)} x^{(i)} + \gamma^{(i)} p^{(i)} \quad (22)$$

$$p^{(i+1)} = w^{(i)} + \gamma^{(i)} p^{(i)} \quad (23)$$

と書くことができる。ここでは行列束 $x^{(i)} - \mu^{(i)} Ax^{(i)}$ に関する $\text{span}\{w, x^{(i)}, p^{(i)}\}$ 上の Ritz 値, Ritz ベクトルを Rayleigh-Ritz 法を用いて計算し、最大 Ritz 値に対応する Ritz ベクトルを $x^{(i+1)}$ とする。すなわち、係数 $\tau^{(i)}, \gamma^{(i)}$ の値は、 $\text{span}\{w, x^{(i)}, p^{(i)}\}$ 上での局所的な最適解をもとに決められる。これによって、ベクトル間の直交性をもとに各係数を明に計算する必要のある従来の方法と比較して、少ない計算量で更新値を求めることができる。

この方法で用いている前処理は、 $T \approx A^{-1}$ をベクトルに適用している点で、実際には逆反復法に近い意味を持っている。ただし、近似逆行列を用いていることから、精度に関しては問題が生じる可能性があり、この点に注意しなければならない。

4. 非対称問題への適用

非対称問題への適用に関しては、先に挙げた冪乗法系の手法の他、以下の系統の手法が提案されている。

Arnoldi 法・Davidson 法

$\{u_1, \dots, u_r\}$ によって生成される Krylov 部分空間を

$$\mathcal{K}_l = \text{lin}(S, AS, \dots, A^{l-1}S) \quad (24)$$

とすると、 A が非 Hermitian 行列ならば Arnoldi 法、ブロック Arnoldi 法が得られる。これらの方法では、写像が Hessenberg 行列 H_l で表される \mathcal{K}_l の正規直交基底 $\{v_i\}_1^l$ を求め、 H_l の固有対を元の問題の近似解として計算する。

Davidson 法では、正規直交基底 $\{v_i\}_1^k$ で張られる部分空間 $\mathcal{K} = \text{span}\{v_1, \dots, v_k\}$ 上で行列 A の Ritz 値 θ_k 及び Ritz ベクトル u_k を計算する際に、残差 $r = Au_k - \theta_k u_k$ に関する修正方程式

$$M_k t = r, \quad M_k = D_A - \theta_k I \quad (25)$$

を解いて u_k を更新し、 \mathcal{K} の次元を拡張する。 D_A は A の対角成分である。具体的には、 t を \mathcal{K} と直交化して v_{k+1} を得る。 $V_{k+1} = [v_1, \dots, v_{k+1}]$ と置けば、新しい Ritz 対 (θ_{k+1}, u_{k+1}) は行列

$$H_{k+1} = V_{k+1}^* A V_{k+1} \quad (26)$$

の固有対として計算されることになる。 $M_k = I$ の場合には Arnoldi 法と同一になるので、Davidson 法は修正を伴う Arnoldi 法の一つと考えることができる。

共役残差法による方法

行列 A が非対称な場合には、共役勾配法の代わりに Rayleigh 商以外の汎関数に対して共役残差法を適用する方法が提案されており¹⁰⁾、並列化に適している。本研究でもこの手法を Lis に実装している。ここでは、固有値問題の残差を

$$r = \lambda x - Ax, \quad (27)$$

$$\lambda = (Ax, x) / (x, x) \quad (28)$$

で定義する。 x が固有値ベクトルに等しければ、 λ は対応する固有値に一致する。ここでは最小化すべき関数として残差 r の内積

$$F(r) = (r, r) \quad (29)$$

を選び、共役残差法 (Orthomin(1)) を適用する。すなわち、初期値 x_0 から

$$\lambda^{(0)} = (Ax^{(0)}, x^{(0)}) / (x^{(0)}, x^{(0)}), \quad (30)$$

$$r^{(0)} = \lambda^{(0)} x^{(0)} - Ax^{(0)}, \quad (31)$$

$$p^{(0)} = r^{(0)} \quad (32)$$

を求め、以下の反復

$$\alpha^{(i)} = \frac{[(r^{(i)}, Ap^{(i)}) - \lambda^{(i)}(r^{(i)}, p^{(i)})]}{[(Ap^{(i)}, Ap^{(i)}) - 2\lambda^{(i)}(Ap^{(i)}, p^{(i)}) + (\lambda^{(i)})^2(p^{(i)}, p^{(i)})]}, \quad (33)$$

$$x^{(i+1)} = x^{(i)} + \alpha^{(i)}p^{(i)}, \quad (34)$$

$$\lambda^{(i+1)} = (Ax^{(i+1)}, x^{(i+1)})/(x^{(i+1)}, x^{(i+1)}), \quad (35)$$

$$r^{(i+1)} = \lambda^{(i+1)}x^{(i+1)} - Ax^{(i+1)}, \quad (36)$$

$$\beta^{(i)} = -[(Ar^{(i+1)}, Ap^{(i)}) - \lambda^{(i+1)}\{(Ar^{(i+1)}, p^{(i)}) + (Ap^{(i)}, r^{(i+1)})\} + (\lambda^{(i+1)})^2(r^{(i+1)}, p^{(i)})] / [(Ap^{(i)}, Ap^{(i)}) - 2\lambda^{(i+1)}(Ap^{(i)}, p^{(i)}) + (\lambda^{(i+1)})^2(p^{(i)}, p^{(i)})], \quad (37)$$

$$p^{(i+1)} = r^{(i+1)} + \beta^{(i)}p^{(i)} \quad (38)$$

を相対残差

$$\epsilon^{(i)} = \frac{\|\lambda^{(i)}x^{(i)} - Ax^{(i)}\|_2}{\|\lambda^{(i)}x^{(i)}\|_2} \quad (39)$$

が十分小さくなるまで繰り返す。

5. Lis を用いた実装

既存の線型計算ライブラリはいくつかあるが、本研究では、以前より開発を進めている Lis に、非線型共役勾配法の実装を行った。

海外においても、Valencia 工科大学による固有値解法ライブラリ SLEPc (Argonne 米国立研究所の並列反復解法ライブラリ PETSc ^{*1} をベースとして開発されている) ^{*2} や、Lawrence Berkeley 米国立研究所による並列反復解法ライブラリ HyPre ^{*3} などに、疎行列向けの固有値解法が実装されている ¹¹⁾。

*1 <http://www-unix.mcs.anl.gov/petsc/>

*2 <http://www.grycap.upv.es/slep/>

*3 <http://computation.llnl.gov/casc/hypr/>

これらのライブラリでは本研究の実装と同様に、いずれもオブジェクト指向に基づいた設計を行っており、並列化はライブラリのレベルで実現されている。また、すべてのデータは API を用いて処理されている。すなわち、行列、ベクトルデータ等の生成・廃棄、及びこれらのオブジェクトに対する操作は、それぞれの操作を記述する API を呼び出すことにより処理される。各前処理手法はそれぞれ線型解法として実装され、必要に応じて前処理として利用することができるようになっており、これらを組み合わせて評価することができる。

本研究では、Lis 上に共役勾配法を実装し、評価を行った。なお、逆反復法等の内部処理においても、 A^{-1} を反復ベクトルに適用する際に線型方程式を解く必要があり、前処理の利用は有効である。表 1, 2, 3 に対応している固有値解法、線型方程式解法、前処理手法を示す。Lis ではこれらを自由に組み合わせて新たに固有値解法を構成することができる。

表 1 Lis で利用可能な固有値解法

Power Iteration
Inverse Iteration
Approximate Inverse Iteration
Conjugate Gradient
Lanczos Iteration
Subspace Iteration
Conjugate Residual

表 2 Lis で利用可能な線型方程式解法

CG	CR
BiCG	BiCR
CGS	CRS
BiCGSTAB	BiCRSTAB
GPBiCG	GPBiCR
BiCGSafe	BiCRSafe
BiCGSTAB(l)	BiCRSTAB(l)
Jacobi	Gauss-Seidel
SOR	Orthomin(m)
TFQMR	MINRES
GMRES(m)	FGMRES(m)
IDR(s)	

表 3 Lis で利用可能な前処理

Jacobi	SSOR
ILU(k)	ILUT
Crout ILU	I+S
SA-AMG	Hybrid
SAINV	Additive Schwarz
ユーザ定義前処理	

実装した解法は、Lis の固有値計算アルゴリズムとして収録され、要素演算と線型方程式解法、前処理に関するライブラリを必要に応じて使用している。反復解法部は行列等のデータに関する要素演算として定義された Lis の API を用いて記述されており、MPI を用いた並列化はこのレベルで行われている。

Lis に実装されている前処理手法について、主なものを簡単に述べる。

- 対角スケールリング・Jacobi 前処理
これらは前処理としては単純な方法であるが、スケーラビリティに関して良好な性能を示す。
- SSOR・Hybrid 前処理
- 不完全 LU 分解前処理
ILU 前処理のスケーラブルな実装。問題サイズがプロセッサ数に比例する場合に、ほぼ一定の計算時間で処理することができる。ILU(k), ILUT, Crout ILU が実装されている。
- SA-AMG¹²⁾ 前処理
代数的マルチグリッド法の逐次・並列実装。Lis では前処理の一つとして実装され、クリロフ部分空間法の前処理として用いることができる。
- 近似逆行列前処理
- Additive Schwarz 前処理

6. 性能評価

Lis 上に実装した共役勾配法の各計算機上での性能について調べるため、九州大学情報基盤研究開発センターの富士通 PRIMEQUEST580 (1.6GHz Intel デュアルコア Itanium 2 プロセッサ ×32)、日立 SR16000 (4.7GHz IBM デュアルコア POWER6 プロセッサ ×16)

の DDR InfiniBand クラスタ、Xeon 5570 サーバ (2.93 GHz クラッドコアプロセッサ ×2)、及び東北大学サイバーサイエンスセンターの SX-9 1 ノードを用いて性能評価を行った。なお、他の固有値解法の性能については文献¹³⁾を参照されたい。

実験では、1次元 Poisson 問題の係数行列について、最小固有対 (最小固有値と対応する固有ベクトル) を共役勾配法により計算した。相対残差ノルムの閾値を 10^{-5} とし、非零要素数がプロセッサ数に比例するよう問題サイズを取って反復解法部の実行時間を測定した結果を表 4-7 に示す。なお、Lis 上に実装した疎行列ベクトル積に関するベンチマークプログラム spmvtest での評価結果をもとに、スカラー計算機については行列格納形式として CRS を、ベクトル計算機については DIA を使用した。また OpenMP を併用した予備的な評価結果を踏まえ、並列化は性能の良好な MPI のみによる方式で行った。

問題サイズなど計測条件が計算機によって若干異なるが、局所 ILU 前処理はスカラー計算機で効果が見られるのに対し、ベクトル計算機では対角スケールリングが良い性能を示している。これはスカラー性能の違いによるものと思われる。ただしこの問題では、逆反復法、共役勾配法で得られる最小固有値は一致しなかった。実際には、表 3 に示した近似逆反復法 (Approximate Inverse Iteration, すなわち逆反復法において、 A^{-1} の代わりに $T \approx A^{-1}$ を適用する手法) での結果と一致する。これは、共役勾配法における前処理の効果が、逆反復法に類似したものであることによるもので、共役勾配法は逆反復法と比較して短時間で解が求まっているものの、精度に関しては注意が必要である。

7. むすび

本稿では、非線型共役勾配法とその前処理手法について、本研究でこれまでに得られている知見をまとめるとともに、複数のアーキテクチャ上での結果について報告した。

本解法の収束特性等に関しては、未解明な部分も残っている。今後大規模問題を対象として、各解法の特性を明らかにしていきたい。

参考文献

- 1) Bai, Z., Demmel, J., Dongarra, J., Ruhe, A. and vander Vorst, H.(eds.): *Templates for the Solution of Algebraic Eigenvalue Problems*, SIAM (2000).
- 2) Sleijpen, G.L.G. and vander Vorst, H.A.: A Jacobi-Davidson iteration method for linear eigenvalue problems, *SIAM J. Matrix Anal. Appl.*, Vol.17, No.2, pp.401-425 (1996).
- 3) 西田晃, 小柳義夫: 大規模固有値問題のための Jacobi-Davidson 法とその特性につい

表 4 PRIMEQUEST580 上での局所 ILU 前処理, 対角スケーリング付き共役勾配法による最小固有対の計算結果

Problem Size	500,000	1,000,000	2,000,000	4,000,000
# cores	8	16	32	64
time (s) (ILUCG)	1.22	1.74	3.12	6.05
time (s) (SCG)	11.1	11.0	12.3	12.8

表 5 SR16000 上での局所 ILU 前処理, 対角スケーリング付き共役勾配法による最小固有対の計算結果

Problem Size	500,000	1,000,000	2,000,000	4,000,000	8,000,000	16,000,000
# cores	8	16	32	64	128	256
time (s) (ILUCG)	0.156	0.433	0.707	1.19	7.49	8.89
time (s) (SCG)	0.877	1.46	3.55	0.994	2.14	4.37

表 6 Xeon 5570 サーバ上での局所 ILU 前処理, 対角スケーリング付き共役勾配法による最小固有対の計算結果

Problem Size	62,500	125,000	250,000	500,000
# cores	1	2	4	8
time (s) (ILUCG)	0.417	0.195	0.285	0.446
time (s) (SCG)	1.63	1.63	3.02	4.58

表 7 SX-9 上での局所 ILU 前処理, 対角スケーリング付き共役勾配法による最小固有対の計算結果

Problem Size	62,500	125,000	250,000
# cores	1	2	4
time (s) (ILUCG)	12.2	5.08	5.26
time (s) (SCG)	0.914	0.961	0.994

て, 情報処理学会論文誌, Vol.41, No.SIG8, pp.101-106 (2000).

- 4) Fletcher, R. and Reeves, C.M.: Function minimization by conjugate gradients, *Comp. J.*, Vol.7, pp.149-154 (1964).
- 5) Bradbury, W.W. and Fletcher, R.: New Iterative Method for Solution of the Eigenproblem, *Numer. Math.*, Vol.9, pp.259-267 (1966).
- 6) Hestenes, M.R. and Karush, W.: A method of gradients for the calculation of the characteristic roots and vectors of a real symmetric matrix, *J. Res. Natl. Inst. Stand. Technol.*, Vol.47, pp.45-61 (1951).
- 7) Knyazev, A.V.: Preconditioned eigensolvers—an oxymoron?, *Electron. Trans. Numer. Anal.*, Vol.7, pp.104-123 (electronic) (1998). Large scale eigenvalue problems (Argonne, IL, 1997).
- 8) Knyazev, A.V.: Toward the Optimal Preconditioned Eigensolver: Locally Optimal Block Preconditioned Conjugate Gradient Method, *SIAM J. Sci. Comput.*, Vol.23, No.2, pp.517-541 (2001).
- 9) Arbenz, P. and Lehoucq, R.: A comparison of algorithms for modal analysis in the absence of a sparse direct method, Technical Report SAND2003-1028J, Sandia National Laboratories (2003).
- 10) Suetomi, E. and Sekimoto, H.: Conjugate gradient like methods and their application to eigenvalue problems for neutron diffusion equation., *Annals of Nuclear Energy*, Vol.18, No.4, pp.205-227 (1991).
- 11) Dongarra, J.: Freely Available Software for the Solution of Linear Algebra Problems, Technical Report <http://www.netlib.org/utk/people/JackDongarra/lasw.html>, Innovative Computing Laboratory, University of Tennessee (2009).
- 12) Fujii, A., Nishida, A. and Oyanagi, Y.: Evaluation of Parallel Aggregate Creation Orders : Smoothed Aggregation Algebraic Multigrid Method, *High Performance Computational Science A*, pp.99-122 (2005).
- 13) 西田晃: 疎行列固有値解法における前処理の特性について, 情処研報, Vol.2008, No.125, pp.103-108 (2008).