

悪条件問題に対するCG法向けIC前処理手法の改善

河合 直聡^{1,a)} 伊田 明弘^{1,b)} 中島 研吾^{1,c)}

概要：不完全コレスキー (IC) 分解は共役勾配法の収束を改善するための手法として広く使用されているが、係数行列の条件数が大きく不定な場合には分解が破綻したり、収束が悪化する場合がある。本研究では分解前の係数行列のブロック化と対角シフトによる正則化に基づく前処理手法を提案し、量子力学アプリケーションから得られる悪条件問題に適用し、収束性を改善できることを示した。

Modified IC Preconditioner of CG method for ill-conditioned problems

MASATOSHI KAWAI^{1,a)} AKIHIRO IDA^{1,b)} KENGO NAKAJIMA^{1,c)}

1. はじめに

固有値解析は、構造解析や量子力学など多くの分野で、解析対象の基本的な特性を調べるために幅広く利用されている。計算機の発展に伴って、要求される解析の精度やモデルが複雑化し、結果として固有値解析の対象は大規模化している。大規模固有値問題を数値的に行う手法は複数提案されている。量子力学などの実問題では、複素平面内の任意の領域内に存在する固有値を計算しなければならない場合がある。これを実現する手法としては、Sakurai-Sugiura法やJacobi-Davidson法など周回積分を利用した手法がある。周回積分は離散的に行われ、離散点毎に大規模な連立一次方程式の求解が必要である。1) 対角成分の絶対値が非対角成分と比較して小さい。あるいは正と負の対角成分が混在している、2) 不正定値である、3) 条件数が大きい、という悪条件な特徴を持つ可能性がある。悪条件な係数行列は計算誤差の増大や破綻を引き起こすため、大規模な問題を高速かつ安定に解ける手法が必要である。

連立一次方程式の解法としては大まかに直接法と反復法に大別されるが、問題が大規模である点から反復法が適している。また、係数行列の性質は悪条件であるため、前処理付きCG法が妥当である。CG法は反復法の一つであ

り、適した前処理との併用により収束性と安定性が向上する。前処理では前処理行列が元の行列に近いほどその効果は高くなるため、不完全コレスキー (IC) 分解に基づく前処理 [1] 適用について考える。IC分解では近似的に完全コレスキー分解を行う方法であり、幅広い分野で用いられている。しかし、IC分解を対角成分の絶対値が非対角成分と比較して小さい係数行列に適用した場合、計算精度の低下や分解破綻が発生する可能性がある。これらの問題が発生することを防ぐために、本研究ではIC分解前の係数行列へのブロック化 [2] と対角シフトによる正則化の適用を提案する。IC前処理付きCG (ICCG) 法への提案手法の適用により、悪条件問題を高速かつ安全に解けることを量子力学アプリケーションの問題 [3][4][5] を対象として確認する。

本稿の構成は次の通りである。2章ではCG法およびIC前処理について述べ、3章では本研究の提案であるブロック化と対角シフトによる正則化について述べる。4章では量子力学アプリケーション問題を対象とした提案手法の評価結果について述べる。

2. 前処理付きCG法

A を大きさが $N \times N$ の係数行列、 x を解ベクトル、 b を右辺ベクトルとして、連立一次方程式

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

をCG法を用いて解く。

CG法はクリロフ部分空間法の写像から得られる探索

¹ 東京大学 情報基盤センター
ITC, University of Tokyo

a) kawai@cc.u-tokyo.ac.jp

b) ida@cc.u-tokyo.ac.jp

c) nakajima@cc.u-tokyo.ac.jp

```

do k = 1, until converge
   $\alpha = \frac{(r^k, p^k)}{(p^k, Ap^k)}$ 
   $x^{k+1} = x^k + \alpha p^k$ 
   $r^{k+1} = r^k - \alpha Ap^k$ 
   $q = P^{-1}r^{k+1}$ 
   $\beta = -\frac{(q, Ap^k)}{(p^k, Ap^k)}$ 
   $p^{k+1} = q + \beta p^k$ 
enddo

```

図 1 前処理付き CG 法のアルゴリズム

ベクトルを利用して近似解を求める手法である。CG 法の収束性は、係数行列 A の最大固有値 λ_{max} と最小固有値 λ_{min} の比 $\kappa = \lambda_{max}/\lambda_{min}$ (条件数) が小さいほど良好となる。一方で、条件数が大きい (悪条件) の場合には計算誤差により探索ベクトルの直交性が維持できなくなり、収束性が悪化する。悪条件な問題を CG 法で解くためには前処理を適用するのが一般的である。前処理を適用した CG 法では前処理行列を P として、以下の方程式を前処理なしの CG 法で解くこと同義である。

$$P^{-1}Ax = P^{-1}b \quad (2)$$

$P = A$ とした場合、係数行列は単位行列となるため、条件数は 1 であり、1 反復で真の解が導かれる。しかし、 P^{-1} を解くことは困難であるため、一般的に $P \approx A$ を満たす前処理行列が選択される。ここで、前処理付き CG 法のアルゴリズムに着目する (図 1)。なお、 x^k, r^k, p^k はそれぞれ k 反復目の近似解、残差、探索ベクトルを表す。本図から分かるように前処理付き CG 法は行列ベクトル積、内積、前処理で構成されている。前処理に必要な計算量やメモリリソースが極端に多い場合には、計算困難となる。従って、前処理は行列ベクトル積と同等のコストであることが望ましい。本研究ではこの要件を満たす前処理手法である IC 分解法の適用について考える。

2.1 不完全コレスキー分解

不完全コレスキー分解は完全コレスキー分解の近似である。完全コレスキー分解では単位上三角行列を U 、対角行列を D として、

$$A = U^T D U \quad (3)$$

と表される形に分解する。この時、 D および U の要素 $D_{i,i}$ および $U_{i,j}$ は、以下の式で算出する。

$$D_{i,i} = A_{i,i} - \sum_{k=1}^{i-1} U_{k,i} D_{i,i} U_{k,i} \quad (4)$$

$$U_{i,j} = \frac{1}{D_{i,i}} \left(A_{i,j} - \sum_{k=1}^{i-1} U_{k,i} D_{i,i} U_{k,j} \right) \quad (5)$$

一方で IC 分解では、単位上三角行列を \bar{U} 、対角行列を \bar{D} として、その各要素 $\bar{U}_{i,j}$ および $\bar{D}_{i,i}$ を、以下で表される。

$$\bar{D}_{i,i} = A_{i,i} - \sum_{k=1}^{i-1} \bar{U}_{k,i} \bar{D}_{i,i} \bar{U}_{k,i} \quad (6)$$

$$\bar{U}_{i,j} = \begin{cases} \frac{1}{\bar{D}_{i,i}} \left(A_{i,j} - \sum_{k=1}^{i-1} \bar{U}_{k,i} \bar{D}_{i,i} \bar{U}_{k,j} \right), & A_{i,j} \neq 0 \\ 0, & A_{i,j} = 0 \end{cases} \quad (7)$$

完全コレスキー分解に対して、IC 分解の差は元の係数行列 A の非ゼロ要素が存在する場所のみ、計算を行う点にある。従って $\bar{U}^T \bar{D} \bar{U} = P_{IC} \neq A$ である。しかし、 \bar{U} と A の上三角領域の非ゼロ要素の分布は一致するため、 $q = P_{IC}^{-1} r^{k+1}$ を求めるために必要な演算量、メモリリソースは行列ベクトル積とほぼ同じである。

先行研究 [6] では本研究で対象とする問題を前処理付き CG 法で解けることを報告している。この先行研究では Carp-CG [7] 法の利用が提案されている。Carp と呼ばれる Kaczmarz [8] 法をベースにした特殊な前処理が適用されており、量子力学 (カーボンナノシートの電気特性解析) の問題を対象にした評価が行われている。また、同問題は Carp のような特殊な前処理が必要であることも示唆されている。従って、本研究でも IC 前処理を適用するだけでなく、さらに正則化を施すことを次節で提案する。

3. 正則化

IC 分解の結果から得られる対角行列 \bar{D} は式 6 から分かるように、減算から得られるため要素 $\bar{D}_{i,i}$ は $A_{i,i}$ と比較して小さくなる。さらに、式 7 から $\bar{U}_{i,j}$ は $\bar{D}_{i,i}$ の除算から算出されるため、 $\bar{U}_{i,j}$ は大きくなり、 $\bar{D}_{i+1,i+1}$ はさらに小さくなる。従って、元の行列 A の対角成分が小さい問題では分解の課程で計算誤差が蓄積し、最悪の場合には分解破綻となる。本研究では正則化を用いて A からより対角成分を大きな A' を導出し、IC 分解を適用する手法を提案する。

3.1 ブロック化

本節では IC 分解前処理に対するブロック化の適用により、収束性とロバスト性の向上が期待出来ることを示す。

ブロック化では分解対象となる係数行列 A の部分行列をそれぞれブロックとして扱い、IC 分解を適用する。適用に際しては全てのブロックが同じ大きさの正方行列である必要がある。そこで、各ブロックのサイズを l とした時の要素数が N' の行列 A' を以下の式で定義する。

$$N' = \lfloor N/l \rfloor + l', \quad (l' = l - N \% l) \quad (8)$$

$$A' = \begin{bmatrix} A & \emptyset \\ \emptyset & I_{l'} \end{bmatrix}, \quad I_{l'} \in \mathbb{C}^{l' \times l'} \quad (9)$$

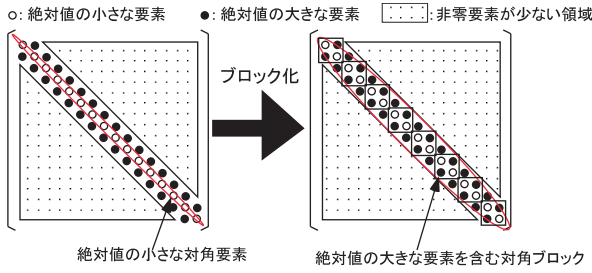


図 2 ブロック IC 分解での対角行列

A' の要素を $m = N'/l$ 個のブロックに分割する．分割したブロック A^b に含まれる要素は A' の各要素 $A'_{i,j}$ として次式で表される．

$$A^b_{i,j} = \begin{bmatrix} A'_{l*(i-1)+1, l*(j-1)+1} & \cdots & A'_{l*(i-1)+1, l*j} \\ A'_{l*(i-1)+2, l*(j-1)+1} & \cdots & A'_{l*(i-1)+2, l*j} \\ \vdots & \ddots & \vdots \\ A'_{l*i, l*(j-1)+1} & \cdots & A'_{l*i, l*j} \end{bmatrix} \quad (10)$$

ブロック化した行列 A^b を IC 分解した結果得られる行列をそれぞれ U^b, D^b とすると，各行列の要素 $U^b_{i,j}$ および $D^b_{i,j}$ とすると，以下の式で表される．

$$\overline{D^b_{i,i}} = A^b_{i,i} - \sum_{k=1}^{i-1} \overline{U^b_{k,i}} \overline{D^b_{i,i}} \overline{U^b_{k,i}} \quad (11)$$

$$\overline{U^b_{i,j}} = \begin{cases} \overline{D^b_{i,i}}^{-1} \left(A^b_{i,j} - \sum_{k=1}^{i-1} \overline{U^b_{k,i}} \overline{D^b_{i,i}} \overline{U^b_{k,j}} \right), & A^b_{i,j} \neq 0 \\ \emptyset, & A^b_{i,j} = 0 \end{cases} \quad (12)$$

対角成分が非対角成分より絶対値が十分に小さい問題を対象とした場合，ブロック化の適用により対角要素が相対的に大きくなる．図 2 に示すように，対角ブロックには元の行列の非対角成分が含まれるためである．結果，ブロック化を適用しない IC 分解と比較して精度低下，分解破綻が起こりにくく，ロバスト性が向上する．

また，図 3 に示すように，元のブロック ($A^b_{i,j}$) が疎な行列の場合でも，除算および乗算の結果 ($\overline{U^b_{i,j}}$) は比較的密な行列となる．このような元の係数行列で 0 の位置に非ゼロ要素が入ることを Fill-in と呼ぶ．ブロック化では Fill-in が発生し，分解結果はより A に近づくため，収束性向上が期待できる．

3.2 対角シフト

本節では係数行列 A' の対角成分の操作によりロバスト性が向上することを示す．

3 で述べたように係数行列の対角成分が小さい場合，分解精度の低下や分解破綻を引き起こす．そこで，より直接的に対象の行列の対角成分に定数 α を足す (対角シフト) 事を提案する．

対角シフトの適用により式 13 は次式で表される．

$$A' = \begin{bmatrix} A + \alpha I_N & 0 \\ 0 & I_{l'} \end{bmatrix}, \quad I_N \in \mathbb{C}^{N \times N} \quad (13)$$

ここで， α を定数とする． A' は対角シフト α により対角成分が大きくなるため，IC 分解による演算精度低下や分解破綻が起こりにくくすることが可能である．ただし， α を極端に大きくした場合， $P \approx A$ を満たさなくなる．問題毎に最適な値を探索する必要があると想定される．

なお，ここで定義した正則化行列 A' は CG 法のアルゴリズム内で前処理にのみ用いるものであり (図 4 の赤字)，それ以外の計算では A (図 4 の青字) を用いる．

4. 評価

本稿で提案した正則化の効果を一般的な ICCG との比較により確認する．

4.1 対象とする問題

本研究で対象とする問題は量子力学の分野から得られる問題である．本稿で取り扱う問題は全部で 3 種類，23 ケースである．以下に問題の特徴を示す．なお，いずれの問題の係数行列も不正定値対象である．

- Kohn-Sham[3]

Kohn-Sham 方程式から原子同士，電子同士の干渉を考慮した電子軌道の解析を目的とした問題である．6 ケースのモデルが存在し，自由度は 57,575 ~ 76,163，非ゼロ要素数は 1 行辺り平均で 20 ~ 24 個である．

- Graehene[4]

炭素分子のから構成される分子 (カーボンナノチューブやフラーレンなど) の電気的特性の解析を目的とした問題である．9 ケースのモデルが存在し，自由度は 128 ~ 1,000,000，非ゼロ要素数は 1 行あたり 13 または 4 個である．

- Spin[5]

電子スピンを分子構造内の相互作用も考慮して解析するための問題である．8 ケースのモデルが存在し，自由度は 252 ~ 2,704,156，非ゼロ要素数は 1 行あたり 6 ~ 12 個である．本問題では対角が 0 の場合が存在するため，IC 分解は A' の該当要素を 10^{-8} で置換した上で実施した．

各問題を解くための右辺ベクトルは 1 ~ 10 の間の乱数とした．また，反復計算では次式を満たした場合に解が収束したと判断した．

$$\frac{\|r^k\|_2}{\|r^0\|_2} \leq 10^{-7} \quad (14)$$

4.2 評価結果

図 5 に正則化を行わなかった場合，ブロックサイズ 4 のブロック化のみを適用した場合， $\alpha = 100$ の対角シフトのみを適用した場合，両方の正則化を適用した場合の

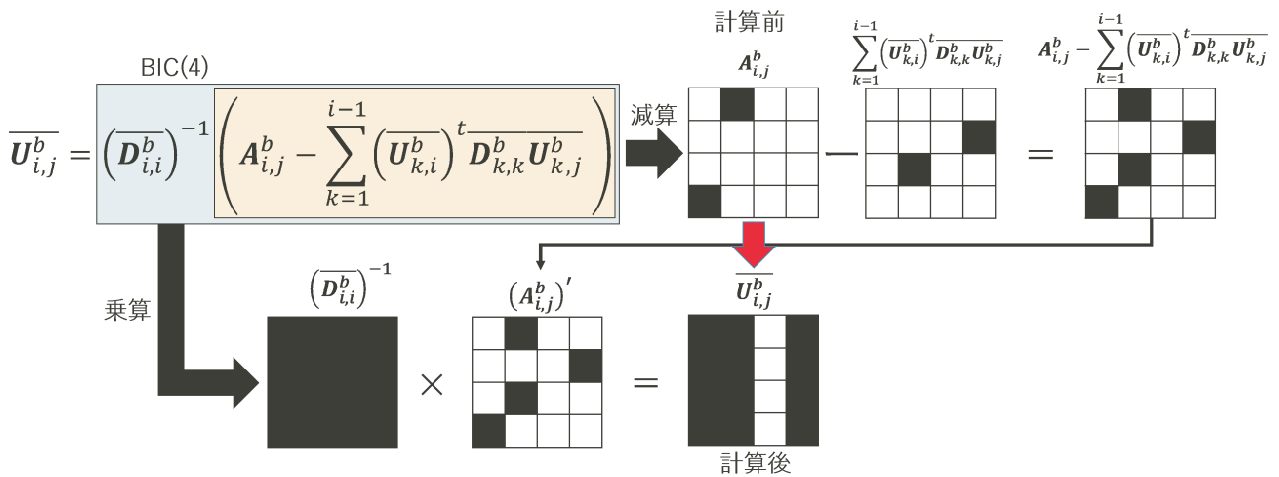


図 3 ブロック IC 分解での fill-in の発生プロセス

```

do k = 1, until converge
  alpha = (r^k, p^k) / (p^k, Ap^k)
  x^{k+1} = x^k + alpha p^k
  r^{k+1} = r^k - alpha Ap^k
  q = (BIC(A'))^{-1} r^{k+1}
  beta = -(q, Ap^k) / (p^k, Ap^k)
  p^{k+1} = q + beta p^k
enddo

```

図 4 正則化を適用した ICCG 法のアルゴリズム
(BIC() はブロック IC 分解演算子を示す.)

結果を示す (それぞれを ICCG, BICCG(4), $\alpha = 100.0$, BICCG(4), $\alpha = 100.0$ と表記). なお青は解けたケース数, 赤は解けなかったケース数, 黒は分解破綻となったケース数を示している. 正則化なしの ICCG では 5 ケースしか解けていないのに対して対角シフトの適用により, 9 ケースの問題が解けている. さらにブロック化の適用で分解破綻を起こしていた問題が改善でき, 最終的に 23 ケース中 15 ケースが解けることを確認した. 解けなかったのはいずれも Spin の問題であった. これは spin の問題の係数行列は対角要素が正または負であるが, 絶対値は非対角要素と比較して小さくなく, IC 分解による計算誤差の増加が少なかったためと推察する.

次に, ブロックサイズが 1 (ICCG), 2, 4, 8, 16, 32 の場合と, シフト量 α が 1.0, 10.0, 100.0 の場合の全ての組み合わせ (18 通り) の収束までの反復回数を図 5 で解けた 15 ケースの問題で評価した. 表 1 はその結果であり, 収束までの反復回数が一番少なかった場合の結果と, その時のブ

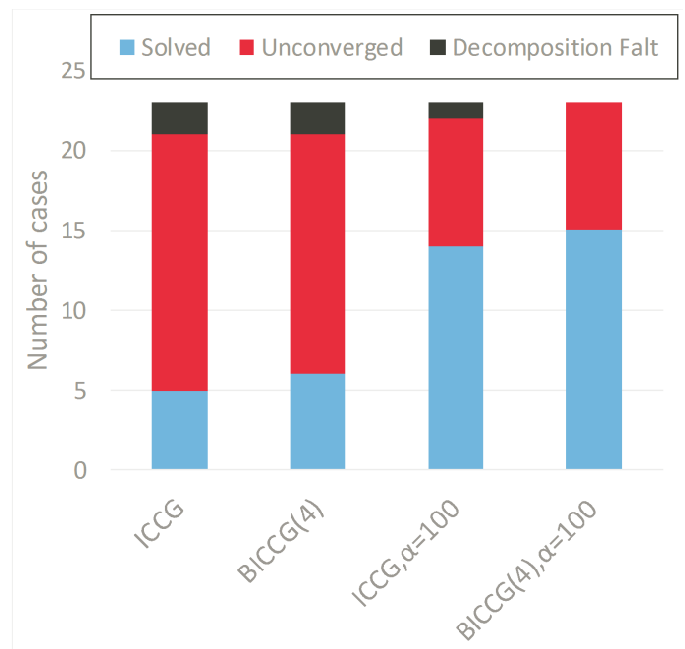


図 5 正則化による効果

ロックサイズ, 対角シフト量を示している. 本結果から, 比較的要素数の少ない問題ではブロックサイズを大きくした BICCG が最も効果的であり, 要素数の大きな問題では対角シフトを適用した方が収束性が良いことがわかる. この傾向は対象の係数行列のバンド幅がいずれも大きい事に起因すると考えられる. 要素数が小さい問題ではブロックサイズを大きくすることにより行列 A の 0 要素に対する Fill-in の割合が大きくなり, $P^{-1} \approx A^{-1}$ となる. 結果, 同条件では対角シフトは逆効果となると推察する. 一方で要素数の大きな問題ではブロック化を適用しても Fill-in の割合が大きくなるため, 計算精度を上げる対角シフトの効果は得られると考えられる.

表 1 反復回数が最小の場合の正則化の条件

問題の種類	問題番号	要素数	反復回数	最適なブロックサイズ	最適な対角シフト量
Kohn-Sham	1	57,575	1795	16	10.0
	2	59,927	1105	16	1.0
	3	62,279	1087	32	1.0
	4	64,631	1356	1	1.0
	5	76,163	788	1	0.0
	6	57,575	686	1	0.0
Graphen	1	1,000	268	16	0.0
	2	10,000	671	16	0.0
	3	100,000	1335	1	100.0
	4	1,000,000	1335	2	100.0
	5	128	13	32	0.0
	6	256	28	16	0.0
	7	8,192	601	1	100.0
	8	32,768	1205	1	100.0
	9	131,072	2382	1	100.0

5. まとめ

本稿では量子力学アプリケーションの固有値問題から得られる連立一次方程式を ICCG 法で解くために、IC 前処理への正則化を提案した。これは対象とする方程式の係数行列が悪条件であり、IC 分解課程での精度低下や分解破綻を抑制するためである。提案した正則化はブロック化と対角シフトの 2 つである。結果、一般の ICCG では 23 ケース中 5 ケースしか解けなかったが、正則化の適用で 15 ケースの問題を解くことができた。

一方で正則化を適用しても spin の問題を解くには至らなかった。またブロック化による収束性改善の効果が想定よりも小さかった。これらは係数行列のバンド幅が大きいことが原因と考えられる。今後はブロック化に Reverse Cuthill Mckee などの Reordering 手法を併用し、さらなる収束性改善に務める。

謝辞 本研究の遂行に関して、貴重なご意見を頂いた先生方（東京大学・埴敏博先生，大島聡史先生，星野哲也先生，北海道大学・岩下武史先生）に感謝の意を表す。また，本研究は JST CREST「ppOpen-HPC」プロジェクト（日本）および SPPEXA「ESSEX」（ドイツ）プロジェクトの支援を受けた研究である。

参考文献

- [1] Saad, Y.: *Iterative Methods for Sparse Linear Systems*, SIAM, Philadelphia, PA, 2nd edition (2003).
- [2] van der Vorst, H. A.: Large tridiagonal and block tridiagonal linear systems on vector and parallel computers, *Parallel Computing*, Vol. 5, No. 1-2, pp. 45-54 (1987).
- [3] Davydov, D., Young, T. D. and Steinmann, P.: On the adaptive finite element analysis of the Kohn-Sham equations: methods, algorithms, and implementation, *International Journal for Numerical Methods in Engineering* (2015).
- [4] Neto, A. C., Guinea, F., Peres, N. M., Novoselov, K. S. and Geim, A. K.: The electronic properties of graphene, *Reviews of modern physics*, Vol. 81, No. 1, p. 109 (2009).
- [5] Thies, J.: Dnnbesetzte Eigenwertlser auf Heterogenen Supercomputern (in German) (2015).
- [6] Galgon, M., Krämer, L., Thies, J., Basermann, A. and Lang, B.: On the parallel iterative solution of linear systems arising in the FEAST algorithm for computing inner eigenvalues, *Parallel Computing*, Vol. 49, pp. 153-163 (2015).
- [7] Gordon, D. and Gordon, R.: CARP-CG: A robust and efficient parallel solver for linear systems, applied to strongly convection dominated PDEs, *Parallel Computing*, Vol. 36, No. 9, pp. 495-515 (2010).
- [8] Haller, R. and Szwarc, R.: Kaczmarz algorithm in Hilbert space, *Studia Math*, Vol. 169, No. 2, pp. 123-132 (2005).