

フラグメント分子軌道法におけるコレスキー分解を用いた 積分近似の適用

(東京大・生研¹, 国立衛生研², JST-CREST³, NEC ソフト⁴, 立教大・理⁵, 神戸大院・工⁶)

○沖山 佳生¹, 中野 達也^{1,2,3}, 山下 勝美⁴, 望月 祐志^{1,3,5}, 田中 成典^{3,6}

【序論】 フラグメント分子軌道 (FMO) 法 [1] は、様々な機能性タンパク質に対して適用されて来た。計算機環境の発展は対象の大規模化を促す一方で、計算の高精度化も同時に要求している。こうした状況に対応するため、近年電子状態計算の高速化手法として、最も計算コストを要する 2 電子積分の生成と扱いに対して因子分解等の近似が適用されるようになった。次世代の FMO 法が目指すものは動力学を含む多数の構造サンプルに対する計算であり、大規模生体分子系に対してさらなる構造空間の探索を行うためには、積分計算の高速化が必要不可欠である。そこで我々は FMO 法の枠組みでコレスキー分解 [2-7] に基づく積分近似を導入した。この近似法では補助基底として原子軌道 (AO) ペアから直裁に構築されるコレスキー基底を用いる。従来の積分ルーチンがそのまま利用でき、エネルギーやその微分計算の実装が容易であるという特長がある。また、クーロン計量の 2 中心積分に関する閾値でスクリーニングを行うことで、精度コントロールができるため、十分な精度を維持しながら積分を高速化することが可能である (閾値ゼロの極限では厳密解を得る)。この方法を我々は CDAM (Cholesky decomposition with adaptive metric) と呼び、開発版の ABINIT-MPX プログラム [8] へ組み込みを行い、FMO-HF ならびに MP2 計算に対して高速化とその精度検証を行った。

【方法】 本研究における積分近似は、Density fitting (DF) 法や Resolution of the identity (RI) 法と非常に類似した考え方を持っており、補助基底を用いて 4 中心の 2 電子積分を次のように擬似 3 中心積分で展開する： $(pq, rs) \approx \sum_{IJ} (pq|I)G_{IJ}^{-1}(J|rs)$ 。ここでクーロン計量行列 $G_{IJ} = (I|J)$ である。ただし外部から補助基底を導入する代わりに、AO 積から直裁に得られるコレスキー基底を補助基底と考える。コレスキー基底の候補となる AO ペアは、まずクーロン計量の 2 中心積分の値でスクリーニングされる： $(I'|I') = (pq|pq) \geq \delta$ 。さらに、その候補から構築されたクーロン計量行列に対する完全ピボット・コレスキー分解によって線形従属性を取り除くことで、最終的に計算に用いるコレスキー基底が選択される。

$$(pq, rs) \approx \sum_{IJ} (pq|I)G_{IJ}^{-1}(J|rs) = \sum_{IJK} (pq|I)Z_{IK}Z_{JK}(J|rs) = \sum_K L_{K,pq}L_{K,rs} \quad (1)$$

補助基底を含む積分は 4 中心積分であるが、擬似的に 3 中心積分と見なすことができる。フラグメントのモノマー・ダイマー・トリマー (高々千軌道程度) が積分計算の対象である FMO 法の枠組みでは、 $L_{K,pq}$ をメモリ上に分散して確保することが可能であり、ファイル I/O を考慮する必要がない点も重要である。 $L_{K,pq}$ および $X_{I,pi} = \sum_r L_{I,pr}C_{ri}$ を用いて、クーロン積分行列ならびに交換積分行列は次のように近似される [9]。

$$J_{pq} = \sum_{rs} (pq, rs)P_{rs} \approx \sum_{rs,I} L_{I,pq}L_{I,rs}P_{rs}, \quad K_{pq} = \frac{1}{2} \sum_{rs} (pr, qs)P_{rs} \approx \sum_{i,I} X_{I,pi}X_{I,qi} \quad (2)$$

また、MP2 計算に必要な 2 電子積分の AO から MO 表示への変換は $B_{I,ia} = \sum_q X_{I,qi} C_{qa} = \sum_{pq} L_{I,pq} C_{pi} C_{qa}$ を用いて、次のように近似される [10]。

$$(ia, jb) = \sum_{pqrs} C_{pi} C_{qa} C_{rj} C_{sb} (pq, rs) \approx \sum_I B_{I,ia} B_{I,jb} \quad (3)$$

このように行列同士の簡単な積の形に落とすことができ、さらに並列化も容易である。

【結果と考察】 Chignolin、Trp-cage および Crambin に対するテスト計算の結果を表に示す。基底関数は 6-31G、計算機環境は Intel Xeon CPU X5272 (3.40GHz) の 8 コア (メモリ 3.5GB / コア)、BLAS レベル 3 の行列演算には Intel Math Kernel Library を用いた。また、コレスキー基底をスクリーニングするための閾値は $\delta = 1.0 \times 10^{-3}$ である。CDAM 近似を用いた場

	Chignolin		Trp-cage		Crambin	
	conventional	CDAM	conventional	CDAM	conventional	CDAM
電子/原子/フラグメント	572 / 138 / 10		1158 / 304 / 20		2520 / 642 / 46	
基底関数	815		1686		3597	
時間 (秒)						
HF コレスキー基底生成		635		1043		2645
SCF	366	620	860	1129	2351	2128
(monomer + dimer)		59%		76%		110%
MP2 monomer	46	23	67	30	110	45
		203%		221%		247%
dimer	915	105	2235	294	5361	724
		869%		761%		741%
Total	1330	1386	3190	2552	8558	5926
		96%		125%		144%
エネルギー (a.u.)						
HF	-3797.8267957	-3797.8267624	-7436.323404	-7436.323335	-17996.8613272	-17996.8611329
$\Delta E(\text{HF})$		0.0000333		0.0000688		0.0001943
MP2	-7.6556242	-7.6556231	-15.29950838	-15.29950442	-32.6243663	-32.6243525
$\Delta E(\text{MP2})$		0.0000011		0.0000040		0.0000138

合、IFIE 解析を行うのに十分な精度を保ちながら、MP2 計算を大いに加速していることがわかる (石川らによる FMO-RI 法 [11] も参照のこと)。また、我々は式 (2) で表されるクーロン積分および交換積分に対する近似を導入することで HF 計算も加速したいと考えている。しかし、ここではコレスキー基底の生成がボトルネックであり、また Crambin サイズの系に対して SCF 計算が従来よりも加速されてはいるが、まだまだ不十分である。対象となる系や基底関数のサイズの増加に対する CDAM 近似の優位性を活かしつつ、コレスキー基底生成過程の並列化やメモリ分散処理・SCF アルゴリズムの改良と併せて、さらなる計算速度の改善が今後の課題である。

【謝辞】 本研究は RISS および JST-CREST プロジェクトの支援のもとで行われている。

【文献】 [1] D. G. Fedorov, K. Kitaura, *J. Phys. Chem. A* **111** (2007) 6904. [2] F. Aquilante, *et al.*, *J. Chem. Phys.* **127** (2007) 114107. [3] L. Boman, *et al.*, *J. Chem. Phys.* **129** (2008) 134107. [4] I. Røeggen, *et al.*, *J. Chem. Phys.* **128** (2008) 194107. [5] F. Aquilante, *et al.*, *J. Chem. Phys.* **129** (2008) 034106. [6] F. Aquilante, *et al.*, *J. Chem. Phys.* **130** (2009) 154107. [7] T.B. Pedersen, *et al.*, *Theo. Chem. Acc.* in press. [8] Y. Mochizuki, K. Yamashita, *et al.*, *Chem. Phys. Lett.* **457** (2008) 396. [9] H. A. Früchtl, *et al.*, *Int. J. Quantum Chem.* **64** (1996) 63. [10] C. Hättig, *NIC Series* **31** (2006) 245. [11] T. Ishikawa, *et al.*, *Chem. Phys. Lett.* **474** (2009) 195.