結果の精度を保証するBLASへの依存度が高い行列積アルゴリズムのパフォーマンス

尾崎 克久 * 荻田 武史 †
* 芝浦工業大学 システム理工学部 数理科学科 (ozaki@sic.shibaura-it.ac.jp) / JST, CREST
† 東京女子大学 現代教養学部 数理科学科 / JST, CREST

概要

本報告では，IEEE 754 規格 [1] が定める浮動小数点数を成分とする行列の積を高精度に計算する方法について述べる．浮動小数点数は有限桁であるために，浮動小数点数同士の和や積などの演算結果には誤差が発生する．そのため，演算毎に結果を浮動小数点数に丸める作業が発生し，ここで発生する誤差の蓄積により，結果が正確でない可能性がある．一般的に精度の問題を克服するには，多倍長精度計算の利用は有力である．ただし，事前に問題がどれだけ悪条件かを把握できなければ，適切な演算精度を指定できない．$\mathbb{F}$ を浮動小数点数の集合とし，演算の相対精度を $\mathbf{u}$ とする．$A \in \mathbb{F}^{m\times n}$, $B \in \mathbb{F}^{n\times p}$ に対して

$$|S - AB| \le \mathbf{u}|AB| \tag{1}$$

となる $S \in \mathbb{F}^{m\times p}$ を求めることを目標とする．ここで $|\cdot|$ は各成分に絶対値を取ることを意味し，行列に対する不等式は，成分毎に不等式がすべて成立することを表す．これは，あたかも実数演算を行った結果を最も近い浮動小数点数に丸めた結果と等しい．(1) を自動的に達成する先行研究として，内積単位での高精度計算 [2] が知られている．我々は，行列積から行列の総和への変換法 [3] を適用した後に，[2] の手法を適用し，(1) を満たす S を得る手法を提案する．特徴として，アルゴリズム全体の計算量の中で BLAS(Level 3) を用いた割合が非常に高い．ポスターセッション当日は，これらのパフォーマンスを比較し，また [3] の手法の応用についてもできる限り紹介したい．

参考文献

[1] *IEEE Standard for Floating-Point Arithmetic*, Std 754–2008, 2008.

[2] S. M. Rump, T. Ogita, S. Oishi: Accurate Floating-Point Summation Part II: Sign, K-fold Faithful and Rounding to Nearest, SIAM J. Sci. Comput. 31:2, 1269-1302 (2008).

[3] K. Ozaki, T. Ogita, S. Oishi, S. M. Rump: Error-Free Transformation of Matrix Multiplication by Using Fast Routines of Matrix Multiplication and its Applications, to appear in Numerical Algorithms.