2011 International Symposium on Nonlinear Theory and its Applications
NOLTA2011, Kobe, Japan, September 4-7, 2011

Accurate matrix multiplication: Improvement of error-free splitting

Katsuhisa Ozaki[†,**], Takeshi Ogita[‡,**] and Shin'ichi Oishi[*,**]

†Department of Mathematical Sciences, College of Systems Engineering and Science, Shibaura Institute of Technology
307 Fukasaku, Minumaku, Saitama-shi, Saitama 337-8570, Japan
‡Division of Mathematical Sciences, Tokyo Woman's Christian University
2-6-1 Zempukuji, Suginami-ku, Tokyo 167-8585, Japan
* Faculty of Science and Engineering, Waseda University
3-4-1 Okubo, Shinjyuku-ku, Tokyo 169-8555, Japan
** CREST, Japan Science and Technology Agency
Email: ozaki@sic.shibaura-it.ac.jp

Abstract—Recently, algorithms for accurate matrix multiplication have been developed by the authors. A main characteristic of the algorithms is to exploit level-3 BLAS routines, which can be highly optimized in terms of execution time. In the algorithms, error-free splitting for floating-point matrices is the key technique. In this paper, the improvement of error-free splitting is focused on. It is shown that the accuracy of matrix products can be improved by the modified error-free splitting, compared to the previous one.

1. Introduction

This paper is concerned with accurate matrix multiplication. Floating-point arithmetic is performed very fast by recent architectures. On the other hand, since a significand of a floating-point number is finite, floating-point arithmetic may cause rounding errors on each arithmetic operation. If rounding errors accumulate, an inaccurate result may be output. To overcome this problem, there are following possibilities:

- multi-precision library [2, 3]
- mixed-precision library [7]
- accurate dot product or summation algorithm [4, 5, 6, 8]

Recently, we have developed new and accurate algorithms for floating-point matrix multiplication [1]. These mainly use functions supported in BLAS (Basic Linear Algebra Subprograms). It is known that the performance of the function for a matrix product in such BLAS is nearly peak, for example GotoBLAS2, Intel Math Kernel Library, ATLAS and so forth. Our algorithms can receive much benefit for optimization and parallelization from the BLAS. When our method requires 3 matrix products for obtaining an approximate result, then accuracy of a result by new error splittings is better than that of a pure-floating result.

For obtaining an accurate result, the key techniques of our algorithms are error-free splittings for floating-point matrices. We develop an another variant of the error-free splittings. As a result, the accuracy of the computed result is improved, compared to the previous algorithm. Finally, numerical results are presented to illustrate the efficiency of the proposed algorithms.

References

[1] K. Ozaki, T. Ogita, S. Oishi: Tight and efficient enclosure of matrix multiplication by using optimized BLAS, Numerical Linear Algebra With Applications, Vol. 18, Issue 2, pp. 237-248.

[2] The MPFR Library: `http://www.mpfr.org/`

[3] exflib - extend precision floating-point arithmetic library:
http://www-an.acs.i.kyoto-u.ac.jp/~fujiwara/exflib/exflib-index.html

[4] T. Ogita, S. M. Rump, S. Oishi: Accurate sum and dot product, SIAM J. Sci. Comput. 26, 1955–1988 (2005).

[5] S.M. Rump: Ultimately Fast Accurate Summation. SIAM J. Sci. Comput. 31:5, 3466-3502 (2009).

[6] S. M. Rump, T. Ogita, S. Oishi: Accurate Floating-Point Summation Part I: Faithful Rounding, SIAM J. Sci. Comput. 31:1, 189-224 (2008).

[7] X. Li, J. Demmel, D. Bailey, G. Henry, Y. Hida, J. Iskandar, W. Kahan, S. Kang, A. Kapur, M. Martin, B. Thompson, T. Tung, D. Yoo, Design, Implementation and Testing of Extended and Mixed Precision BLAS, ACM Transactions on Mathematical Software 28:2, 152-205 (2002).

[8] J. Demmel, Y. Hida: Accurate and Efficient Floating Point Summation, SIAM J. Sci. Comput. 25:4, 1214–1248 (2003).