Tight Enclosure of Matrix Multiplication with Level 3 BLAS

Katsuhisa Ozaki[1] and Takeshi Ogita[2]

[1] Shibaura Institute of Technology, Department of Mathematical Sciences, Fukasaku 307, Minuma-ku, Saitama-shi, Saitama 337-8570, Japan
ozaki@sic.shibaura-it.ac.jp
[2] Tokyo Woman's Christian University, Division of Mathematical Sciences, 2-6-1 Zempukuji, Suginami-ku, Tokyo 167-8585, Japan

Keywords: Matrix Multiplication, Interval Arithmetic

Introduction

This talk is concerned with enclosure of a product of two matrices. Let $\mathbb{F}$ denote a set of floating-point numbers as defined in IEEE 754 [1]. For $A \in \mathbb{F}^{m\times n}$ and $B \in \mathbb{F}^{n\times p}$, the concern is to obtain an interval matrix $[C]$ which encloses AB, namely, $AB \in [C]$. Notations $\mathtt{fl}(\cdot)$, $\mathtt{fl}_{\triangle}(\cdot)$ and $\mathtt{fl}_{\triangledown}(\cdot)$ mean that all operations in the parenthesis are evaluated by floating-point arithmetic with rounding to nearest, rounding upward and rounding downward, respectively. A well-know method for this problem is to compute $[C] := [\mathtt{fl}_{\triangledown}(AB), \mathtt{fl}_{\triangle}(AB)]$, which involves two matrix products.

Proposed Method

Recently, we developed enclosure methods for AB via three or five floating-point matrix products, which often provide tighter results than the well-known method. First, A and B are split into an unevaluated sum of two matrices as follows

$$A = A^{(1)} + A^{(2)}, \quad B = B^{(1)} + B^{(2)}, \quad A^{(1)}B^{(1)} = \mathtt{fl}(A^{(1)}B^{(1)}) \qquad (1)$$

where $A^{(1)}, A^{(2)} \in \mathbb{F}^{m\times n}$, $B^{(1)}, B^{(2)} \in \mathbb{F}^{n\times p}$ and $\mathtt{fl}(A^{(1)}B^{(1)}) = A^{(1)}B^{(1)}$ means that no rounding error occurs in the evaluation of $A^{(1)}B^{(1)}$.

Then, AB is enclosed by

$$\begin{aligned} AB &= A^{(1)}B^{(1)} + A^{(1)}B^{(2)} + A^{(2)}B \\ &\in \mathtt{fl}(A^{(1)}B^{(1)}) + [\mathtt{fl}_{\nabla}(A^{(1)}B^{(2)}), \mathtt{fl}_{\triangle}(A^{(1)}B^{(2)})] \\ &\quad +[\mathtt{fl}_{\nabla}(A^{(2)}B), \mathtt{fl}_{\triangle}(A^{(2)}B)]. \end{aligned}$$

This method involves five matrix products.

Let a constant β be defined as

$$\beta := \lceil (\log_2 \alpha - \log_2 \mathtt{u})/2 \rceil \,. \tag{2}$$

We define two vectors σ and τ as follows:

$$\sigma_i := 2^{\beta} \cdot 2^{w_i}, \quad \tau_j := 2^{\beta} \cdot 2^{v_j},$$

where

$$\begin{aligned} w_i &:= \lceil \log_2 \max_{1 \le j \le n} |a_{ij}| \rceil \text{ for } \max_{1 \le j \le n} |a_{ij}| \neq 0 \;\;, \;\; w_i = 0 \text{ for } \max_{1 \le j \le n} |a_{ij}| = 0, \\ v_j &:= \lceil \log_2 \max_{1 \le i \le n} |b_{ij}| \rceil \text{ for } \max_{1 \le i \le n} |b_{ij}| \neq 0 \;\;, \;\; v_j = 0 \text{ for } \max_{1 \le i \le n} |b_{ij}| = 0. \end{aligned}$$

If a suitable constant α in (2) can be set, then (1) is satisfied from the following results

$$\begin{aligned} a_{ij}^{(1)} &:= \mathtt{fl}((a_{ij} + \sigma_i) - \sigma_i), & a_{ij}^{(2)} &:= \mathtt{fl}(a_{ij} - a_{ij}^{(1)}), \\ b_{ij}^{(1)} &:= \mathtt{fl}((b_{ij} + \tau_j) - \tau_j), & b_{ij}^{(2)} &:= \mathtt{fl}(b_{ij} - b_{ij}^{(1)}). \end{aligned}$$

In the corresponding original methods in [2] $\alpha := n$ yields (1). However, even if $\alpha < n$, we can prove (1) by a posteriori validation with diagonal scaling. The details of the method and numerical results will be shown in the presentation.

References

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

[2] K. Ozaki, T. Ogita and S. Oishi, Tight and efficient enclosure of matrix multiplication by using optimized BLAS, *Numerical Linear Algebra With Applications* 18(2):237–248, 2011.