# Verification methods for system of linear equations in rounding to nearest

Yusuke Morikura[1,*] and Shin'ichi Oishi[2,3]

[1]Graduate School of Fundamental Science and Engineering, Waseda University
Building 63, Room 419, Okubo 3-4-1, Shinjuku, Tokyo 169-8555, Japan
[2]Faculty of Science and Engineering, Waseda University
[3]JST, CREST

[*]m.myusuke@suou.waseda.jp

**Keywords:** verified numerical computations, system of linear equations, high performance computation

This talk is concerned with verification methods for systems of linear equations:

$$Ax = b, \quad A \in \mathbb{R}^{n\times n},\ b \in \mathbb{R}^n. \tag{1}$$

We do not assume special structures for $A$. The verification methods for systems of linear equations (1) give existence to the exact solution and error bounds between the exact solution and an approximate solution by using floating point arithmetic defined by the IEEE 754 standard [1]. The aim of this talk is to propose verification methods for systems of linear equations which are portable and high precision, and its implementation for high performance computers. In this talk, we introduce the method by using super computers (Fujitsu FX10 [2]).

Let $I$ be the $n \times n$ identity matrix. Let $\tilde{x}$ be an approximate solution of (1). To obtain an error bound of an approximate solution, the following inequality is often-used [3]. If we can find $R \in \mathbb{R}^{n\times n}$ such that

$$\|RA - I\|_\infty < 1, \tag{2}$$

then $A^{-1}$ exists and

$$\|\tilde{x} - A^{-1}b\|_\infty \quad \leq \quad \frac{\|R(A\tilde{x} - b)\|_\infty}{1 - \|RA - I\|_\infty}. \tag{3}$$

Usually, $R$ is an approximate inverse of $A$. Many verified computations for (2) and (3) use the switches of rounding modes defined by the IEEE 754 standard [1]. However, for recent some of high performance computers like Graphics Processing Units, super computers, etc., to control the rounding modes is difficult. Therefore, we propose verification methods for systems of linear systems in rounding to nearest.

First, we improve the verification method [4] by using error estimates [5]. Since the verification method [4] use a priori error estimates of floating-point arithmetic in rounding to nearest, it dose not need the switches of rounding modes. It can work on a wide range of computational environments. Its a priori error estimates are based on traditional type estimates. We propose a new method by using new error estimates developed by Rump [5].

Next, we introduce the verification method by using super computers. Especially, in case of using super computers, we consider large scale problems. $n$ is up to $10^5$. If $n$ is increase, it is known that the accuracy will be worse. In addition, since our method is based on a priori error estimates, overestimation occurs. Therefore, we obtain a good approximate solution by using iterative refinement based on accurate dot product [6]. By using the accurate algorithm for calculating dot product [6] with its error bound in twice the working precision, we get sharp error bounds of residual $A\tilde{x} - b$. Since we apply the accurate matrix multiplication [7], we improve adaptable range.

The details of our method and numerical experiments will be shown at the presentation.

This research is partially supported by Initiative on Promotion of Supercomputing for Young or Women Researchers, Supercomputing Division, Information Technology Center, The University of Tokyo.

**References:**

[1] ANSI/IEEE 754-1985, IEEE Standard for Binary Floating-Point Arithmetic, New York, 1985.

[2] Fujitsu FX10, Fujitsu, Supercomputer PRIMEHPC FX10, `http://www.fujitsu.com/global/services/solutions/tc/hpc/products/primehpc/`

[3] S. Oishi and S. M. Rump, Fast Verification of Solutions of Matrix Equations, *Numer. Math.*, vol. 90, no. 4, pp. 755–773, 2002.

[4] T. Ogita, S. M. Rump, and S. Oishi, Verified Solutions of Linear Systems without Directed Rounding, *Technical Report 2005-04, Advanced Research Institute for Science and Engineering, Waseda University Tokyo, Japan*, 2005.

[5] S.M. Rump, Error estimation of floating-point summation and dot product, *BIT Numerical Mathematics*, vol. 52, no. 1, pp. 201–220, 2012.

[6] T. Ogita, S. M. Rump and S. Oishi, Accurate sum and dot product, *SIAM Journal on Scientific Computing*, vol. 26, no. 6, pp. 1955–1988, 2005.

[7] K. Ozaki, T. Ogita, S. Oishi and S. M. Rump, Error-Free Transformation of Matrix Multiplication by Using Fast Routines of Matrix Multiplication and its Applications, *Numerical Algorithms*, vol. 59, no. 1, pp. 95–118, 2012.