連立一次方程式Ax=bの精度保証

小室 和範
早稲田大学 基幹理工学部 応用数理学科
e-mail : komuro@akane.waseda.jp

1　はじめに

連立一次方程式 $Ax = b$ を最近点への丸めのみを用いて計算する際の誤差評価を事前誤差評価により行い、GPU上での実装を可能にする。$\mathbb{F}$ を浮動点小数点数の集合、係数行列 $A \in \mathbb{F}^{n\times n}$ とする。また、本報告中の計算は、IEEE754規格に基づく倍精度浮動小数点演算を用い [1]、アンダーフロー、オーバーフローは起こらないと仮定する。Oishi-Rumpが文献 [2] において、IEEE754に基づく丸めモードの変更を用い、連立一次方程式 $Ax = b, A \in \mathbb{F}^{n\times n}, b \in \mathbb{F}^n$ における解を精度保証付きで求める方法を提案している。これは、A が区間の場合にも拡張が可能である。連立一次方程式 $Ax = b$ において区間を持った A, b を

$$A =< A_{mid}, A_{rad} >$$
$$b =< b_{mid}, b_{rad} >$$

のように表す。このとき、$A_{mid} \in \mathbb{F}^{n\times n}$ は A の中心,$A_{rad} \in \mathbb{F}^{n\times n}$ は A の半径を表す。同様に $b_{mid} \in \mathbb{F}^n$ は b の中心,$b_{rad} \in \mathbb{F}^n$ は b の半径を表す。$\tilde{x} \in \mathbb{F}^{n\times n}(\tilde{x} = A_{mid}\backslash b_{mid})$ を近似解,$R \in \mathbb{F}(R = \mathrm{inv}(A_{mid}))$ としたとき、$\|RA-I\|_\infty \leq 1$ であれば、A は正則であり、$Ax = b$ における精度保証式は、区間を持った場合も同様に

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

と評価される。本報告書中、$\mathbf{u}, e, \gamma_n, \tilde{\gamma}_n$ はそれぞれ $\mathbf{u} = 2^{-53}, e = (1, \cdots, 1)^T \in \mathbb{F}^n, \gamma_n = \dfrac{n\mathbf{u}}{1 - n\mathbf{u}}, \tilde{\gamma}_n = \mathrm{fl}\left(\dfrac{n\mathbf{u}}{1 - n\mathbf{u}}\right)$ を表す。

2　先行研究

連立一次方程式 $Ax = b$ においては、以下の評価式が成立する。

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

$Ax = b$ における上記した精度保証式を評価するために、$\|RA-I\|_\infty \leq \alpha$ と $\|R(A\tilde{x}-b)\|_\infty \leq \beta$ を丸めの向きを変更する手法で求めるアルゴリズムを以下に記す。本報告中のアルゴリズムはコード記述の簡略化のために Matlab, INTLAB における関数に似たコードで記述する。

Algorithm 2.0 (Oishi and Rump [2]).$A \in \mathbb{F}^{n\times n}$, $b \in \mathbb{F}^{n\times n}$, $\tilde{x}$ を $Ax = b$ の近似解, R を A の近似逆行列としたとき、丸めの向きを変更して $\|RA - I\|_\infty$ の上限 α と $\|R(A\tilde{x} - b)\|_\infty$ の上限 β を求めるアルゴリズム

$$
\begin{aligned}
&\mathrm{function}[\alpha, \beta] = \mathrm{std.rnd}(A, b, \tilde{x}, R)\\
&\mathrm{setround}(-1)\\
&\underline{G} = \mathrm{fl}(RA - I);\ \underline{r} = \mathrm{fl}(A\tilde{x} - b);\\
&\mathrm{setround}(+1)\\
&\overline{G} = \mathrm{fl}(RA - I);\ \overline{r} = \mathrm{fl}(A\tilde{x} - b);\\
&\alpha = \mathrm{fl}(\|\max(|\underline{G}|, |\overline{G}|)\|_\infty);\\
&r_{mid} = (\underline{r} + \overline{r})/2; r_{rad} = r_{mid} - \underline{r};\\
&\mathrm{setround}(-1)\\
&\underline{t} = \mathrm{fl}(R \cdot r_{mid});\\
&\mathrm{setround}(+1)\\
&\overline{t} = \mathrm{fl}(R \cdot r_{mid});\\
&\beta = \mathrm{fl}(\|\max(|\underline{t}|, |\overline{t}|) + |R| \cdot r_{rad}\|_\infty);
\end{aligned}
$$

Ogita-Rump-Oishiが (1) の関係式において，最近点丸めにおける浮動小数点演算の誤差評価を用いて，丸めモードの変更を用いない手法を提案している [3]。$\mathrm{fl}(\cdot)$ は括弧内の計算をすべて最近点への丸めで計算を行ったことを表す。このとき (1) の関係式は最近点丸めにおける計算のみを用いて、以下のように計算される。

$$\|x - \tilde{x}\|_\infty \leq \mathrm{fl}\left(\frac{\beta/(1 - \alpha)}{1 - 3\mathbf{u}}\right). \qquad (2)$$

本報告では、アンダーフローが起こらないと仮定するため、文献 [3] における誤差評価式を用い精度保証式 (2) における α を再評価する。

2.1 $\|RA-I\|_\infty$ の上限 α の再評価

アンダーフローが起こらないと仮定した場合、$\|RA-I\|_\infty$ の上限 α は次のように求められる。

$$\begin{aligned} & \|RA-I\|_\infty \\ \le\ & \|\mathrm{fl}(RA)-I\|_\infty + \|\gamma_n|R||A|\|_\infty \end{aligned}$$

$$\begin{aligned} & \|\gamma_n|R||A|\|_\infty \\ =\ & \|\gamma_n|R|(|A|e)\|_\infty \\ \le\ & \|\gamma_n|R|(1+\mathbf{u})^{n-1}\mathrm{fl}(|A|e)\|_\infty \\ \le\ & \|\gamma_n(1+\mathbf{u})^{2n-1}\mathrm{fl}(|R|(|A|e))\|_\infty \\ \le\ & \mathrm{fl}\left(\frac{\tilde{\gamma}_{n+1}\||R|(|A|e)\|_\infty}{1-2n\mathbf{u}}\right) \\ =:\ & c_1 \end{aligned}$$

$M=\mathrm{fl}(RA)$ とすると、

$$\begin{aligned} & \|\mathrm{fl}(RA)-I\|_\infty \\ =\ & \|M-I\|_\infty \\ \le\ & \|\mathrm{fl}(M-I)+\mathrm{fl}(\mathbf{u}\cdot I)\|_\infty \\ \le\ & \|\mathrm{fl}(M-I)\|_\infty + \|\mathrm{fl}(\mathbf{u}\cdot I)\|_\infty \end{aligned}$$

ここで、

$$\begin{aligned} \|\mathrm{fl}(M-I)\|_\infty \ \le\ & \mathrm{fl}\left(\frac{\|M-I\|_\infty}{1-n\mathbf{u}}\right) \\ =:\ & c_2 \end{aligned}$$

$$\begin{aligned} \|\mathrm{fl}(\mathbf{u}\cdot I)\|_\infty \ =\ & \mathrm{fl}(\|\mathbf{u}\cdot I\|_\infty) \\ =:\ & c_3 \end{aligned}$$

とすれば、$\|RA-I\|_\infty$ の上限 α は c_1, c_2, c_3 により次のようになる。

$$\begin{aligned} & \|RA-I\|_\infty \\ \le\ & c_1+c_2+c_3 \\ \le\ & (1+\mathbf{u})^2\mathrm{fl}(c_1+c_2+c_3) \\ \le\ & \mathrm{fl}\left(\frac{c_1+c_2+c_3}{1-3\mathbf{u}}\right) =: \alpha \end{aligned}$$

Algorithm 2.1 $A \in \mathbb{F}^{n\times n}$, $b \in \mathbb{F}^{n\times n}$, $\tilde{x}$ を $Ax=b$ の近似解, R を A の近似逆行列としたとき、丸めモードを変更しないで $\|RA-I\|_\infty$ の上限 α を求めるアルゴリズム

function$[\alpha]$ = regularity1(A, R)

setround(0);

$c_1 = \mathrm{fl}\left(\frac{\tilde{\gamma}_{n+1}\||R|(|A|e)\|_\infty}{1-2n\mathbf{u}}\right)$;

$M = \mathrm{fl}(RA)$;

$c_2 = \mathrm{fl}\left(\frac{\|M-I\|_\infty}{1-n\mathbf{u}}\right)$;

$c_3 = \mathrm{fl}\left(\frac{\tilde{\gamma}_2\||M|+|I|\|_\infty}{(1-2\mathbf{u})(1-n\mathbf{u})}\right)$;

$\alpha = \mathrm{fl}\left(\frac{c_1+c_2+c_3}{1-3\mathbf{u}}\right)$;

end

2.2 $\|R(A\tilde{x}-b)\|_\infty$ の上限 β の評価

$\|R(A\tilde{x}-b)\|_\infty$ の上限 β は次のように求められる。

$$\begin{aligned} & \|R(A\tilde{x}-b)\|_\infty \\ \le\ & \||R\mathrm{fl}(A\tilde{x}-b)| + |R|\gamma_{n+1}(|A||\tilde{x}|+|b|)\|_\infty \end{aligned}$$

ここで、

$$\begin{aligned} & |R\mathrm{fl}(A\tilde{x}-b)| \\ \le\ & |\mathrm{fl}(R(A\tilde{x}-b)) + \gamma_n|R||\mathrm{fl}(A\tilde{x}-b)|| \end{aligned}$$

となるので、

$$\mathrm{fl}(R(A\tilde{x}-b)) =: c_1$$

$$\begin{aligned} & \gamma_n|R||\mathrm{fl}(A\tilde{x}-b)| \\ \le\ & \gamma_n(1+\mathbf{u})^n\mathrm{fl}(|R||A\tilde{x}-b|) \\ \le\ & \mathrm{fl}\left(\frac{\tilde{\gamma}_{n+1}|R||A\tilde{x}-b|}{1-(n+1)\mathbf{u}}\right) \\ =:\ & c_2 \end{aligned}$$

とする。さらに、

$$\begin{aligned} & |R|\gamma_{n+1}(|A||\tilde{x}|+|b|) \\ \le\ & |R|\gamma_{n+1}(1+\mathbf{u})^{n+1}\mathrm{fl}(|A||\tilde{x}|+|b|) \\ \le\ & \gamma_{n+1}(1+\mathbf{u})^{2n+1}\mathrm{fl}(|R|(|A||\tilde{x}|+|b|)) \\ \le\ & \mathrm{fl}\left(\frac{\tilde{\gamma}_{n+2}|R|(|A||\tilde{x}|+|b|)}{1-(2n+2)\mathbf{u}}\right) \\ =:\ & c_3 \end{aligned}$$

とすれば、$\|R(A\tilde{x}-b)\|_\infty$ の上限 β は c_1, c_2, c_3 により次のようになる。

$$
\begin{aligned}
& \|R(A\tilde{x}-b)\|_\infty \\
\le\ & \||c_1+c_2|+c_3\|_\infty \\
\le\ & \||c_1|+|c_2|+c_3\|_\infty \\
\le\ & \|(1+\mathbf{u})^2\mathrm{fl}(|c_1|+|c_2|+c_3)\|_\infty \\
\le\ & \mathrm{fl}\left(\frac{\||c_1|+|c_2|+c_3\|_\infty}{1-3\mathbf{u}}\right)
\end{aligned}
$$

Algorithm 2.2 $A \in \mathbb{F}^{n\times n}$, $b \in \mathbb{F}^{n\times n}$, $\tilde{x}$ を $Ax=b$ の近似解, R を A の近似逆行列としたとき、丸めモードを変更しないで $\|R(A\tilde{x}-b)\|_\infty$ の上限 β を求めるアルゴリズム

$$
\begin{aligned}
& \mathrm{function}[\beta] = \mathrm{numerator1}(A, b, \tilde{x}, R) \\
& \mathrm{setround}(0); \\
& c_1 = \mathrm{fl}(R(A\tilde{x}-b)); \\
& c_2 = \mathrm{fl}\left(\frac{\tilde{\gamma}_{n+1}|R||A\tilde{x}-b|}{1-(n+1)\mathbf{u}}\right); \\
& c_3 = \mathrm{fl}\left(\frac{\tilde{\gamma}_{n+2}|R|(|A||\tilde{x}|+|b|)}{1-(2n+2)\mathbf{u}}\right); \\
& \beta = \mathrm{fl}\left(\frac{\||c_1|+|c_2|+c_3\|_\infty}{1-3\mathbf{u}}\right); \\
& \mathrm{end}
\end{aligned}
$$

3 本報告における手法

連立一次方程式における精度保証法を A, b が区間の場合に拡張し、事前誤差評価を用いた最近点丸めのみを用いた手法を提案する。A, b が区間を持つとき、$\|RA-I\|_\infty \le \alpha$ と $\|R(A\tilde{x}-b)\|_\infty \le \beta$ を丸め変更する手法で求めるアルゴリズムを以下に記す。

Algorithm 3.0 $A =< A_{mid}, A_{rad} >$,$b =< b_{mid}, b_{rad} >$,$\tilde{x}$ を $Ax=b$ の近似解,R を A の近似逆行列としたとき,丸めを変更して $\|RA-I\|_\infty$ の上限 α と $\|R(A\tilde{x}-B)\|_\infty$ の上限 β を求めるアルゴリズム

$$
\begin{aligned}
& \mathrm{function}[\alpha,\beta] = \mathrm{std.rnd2}(A, b, \tilde{x}, R) \\
& G = \mathrm{fl}(RA-I);\ r = \mathrm{fl}(A\tilde{x}-b); \\
& \mathrm{setround}(+1) \\
& \alpha = \mathrm{fl}(\sup(\|G\|_\infty)); \\
& \mathrm{setround}(-1) \\
& \underline{t} = \mathrm{fl}(R\cdot\mathrm{mid}(r)); \\
& \mathrm{setround}(+1) \\
& \overline{t} = \mathrm{fl}(R\cdot\mathrm{mid}(r)); \\
& \beta = \mathrm{fl}(\|\max(|\underline{t}|, |\overline{t}|) + |R|\cdot r_{rad}\|_\infty);
\end{aligned}
$$

3.1 $\|RA-I\|_\infty$ の評価 (A が区間の場合)

$A =< A_{mid}, A_{rad} >$ のように A が区間の場合

$$
\begin{aligned}
& RA-I \\
\in\ & < R\cdot A_{mid}-I, |R|A_{rad} >
\end{aligned}
$$

となる。このとき、$\|RA-I\|_\infty$ の上限 α は次のように求められる。

$$
\begin{aligned}
& \|RA-I\|_\infty \\
\le\ & \||R\cdot A_{mid}-I|+|R|A_{rad}\|_\infty \\
\le\ & \|R\cdot A_{mid}-I\|_\infty + \||R|A_{rad}\|_\infty
\end{aligned}
$$

ここで、$\|R\cdot A_{mid}-I\|_\infty$ の上限を d_1 とすると、d_1 は区間なしのときと同様に評価できる。さらに、

$$
\begin{aligned}
& \||R|A_{rad}\|_\infty \\
=\ & \||R|(A_{rad}e)\|_\infty \\
\le\ & \||R|(1+\mathbf{u})^{n-1}\mathrm{fl}(A_{rad}e)\|_\infty \\
\le\ & \|(1+\mathbf{u})^{2n-1}\mathrm{fl}(|R|(A_{rad}e))\|_\infty \\
\le\ & \mathrm{fl}\left(\frac{\||R|(A_{rad}e)\|_\infty}{1-2n\mathbf{u}}\right) \\
=:\ & d_2
\end{aligned}
$$

とすれば、$\|RA-I\|_\infty$ の上限 α は d_1, d_2 により次のようになる。

$$
\begin{aligned}
& \|RA-I\|_\infty \\
\le\ & d_1+d_2 \\
\le\ & (1+\mathbf{u})\mathrm{fl}(d_1+d_2) \\
\le\ & \mathrm{fl}\left(\frac{d_1+d_2}{1-2\mathbf{u}}\right) =: \alpha
\end{aligned}
$$

Algorithm 3.1 $A=<A_{mid},A_{rad}>$,$\tilde{x}$ を $Ax=b$ の近似解,R を A の近似逆行列としたとき, 丸めモードを変更しないで $\|RA-I\|_\infty$ の上限 α を求めるアルゴリズム

function$[\alpha]$ = regularity2(A_{mid},A_{rad},R)
setround(0);
$d_1=$ regularity1(A_{mid},R);
$d_2=\mathrm{fl}\left(\frac{\||R|(A_{rad}e)\|_\infty}{1-2n\mathbf{u}}\right)$;
$\alpha=\mathrm{fl}\left(\frac{d_1+d_2}{1-2\mathbf{u}}\right)$;
end

3.2 $\|R(A\tilde{x}-b)\|_\infty$ の評価 (A,b が区間の場合)

A,b が次のような区間の場合

$$
\begin{aligned}
A &= <A_{mid},A_{rad}> \\
b &= <b_{mid},b_{rad}>
\end{aligned}
$$

$$
\begin{aligned}
&R(A\tilde{x}-b) \\
\in\ &<R(A_{mid}\tilde{x}-b_{mid}),|R|(A_{rad}|\tilde{x}|+b_{rad})>
\end{aligned}
$$

となる。このとき、$\|R(A\tilde{x}-b)\|_\infty$ の上限 β は次のように求められる。

$$
\begin{aligned}
&\|R(A\tilde{x}-b)\|_\infty \\
\le\ &\||R(A_{mid}\tilde{x}-b_{mid})|+|R|(A_{rad}|\tilde{x}|+b_{rad})\|_\infty
\end{aligned}
$$

ここで、$|R(A_{mid}\tilde{x}-b_{mid})|$ の部分は区間なしの場合と同様に、

$$
\begin{aligned}
c_1' &= \mathrm{fl}(R(A_{mid}\tilde{x}-b_{mid})) \\
c_2' &= \mathrm{fl}\left(\frac{\tilde{\gamma}_{n+1}|R||A_{mid}\tilde{x}-b_{mid}|}{1-(n+1)\mathbf{u}}\right) \\
c_3' &= \mathrm{fl}\left(\frac{\tilde{\gamma}_{n+2}|R|(|A_{mid}||\tilde{x}|+|b_{mid}|)}{1-(2n+2)\mathbf{u}}\right)
\end{aligned}
$$

とすれば

$$
|R(A_{mid}\tilde{x}-b_{mid})|\le|c_1'+c_2'|+c_3'
$$

である。また、

$$
\begin{aligned}
&|R|(A_{rad}|\tilde{x}|+b_{rad}) \\
\le\ &|R|(1+\mathbf{u})^{n+1}\mathrm{fl}(A_{rad}|\tilde{x}|+b_{rad}) \\
\le\ &(1+\mathbf{u})^{2n+1}\mathrm{fl}(|R|(A_{rad}|\tilde{x}|+b_{rad})) \\
\le\ &\mathrm{fl}\left(\frac{|R|(A_{rad}|\tilde{x}|+b_{rad})}{1-(2n+2)\mathbf{u}}\right) \\
=:\ &c_4
\end{aligned}
$$

より、$\|R(A\tilde{x}-b)\|_\infty$ の上限 β は c_1',c_2',c_3',c_4 を用いて次のように表せる。

$$
\begin{aligned}
&\|R(A\tilde{x}-b)\|_\infty \\
\le\ &\||c_1'+c_2'|+c_3'+c_4\|_\infty \\
\le\ &\||c_1'|+|c_2'|+c_3'+c_4\|_\infty \\
\le\ &\|(1+\mathbf{u})^3\mathrm{fl}(|c_1'|+|c_2'|+c_3'+c_4)\|_\infty \\
\le\ &\mathrm{fl}\left(\frac{\||c_1'|+|c_2'|+c_3'+c_4\|_\infty}{1-4\mathbf{u}}\right)=:\beta
\end{aligned}
$$

Algorithm 3.2 $A=<A_{mid},A_{rad}>$,$b=<b_{mid},b_{rad}>$,$\tilde{x}$ を $Ax=b$ の近似解,R を A の近似逆行列としたとき, 丸めモードを変更しないで $\|R(A\tilde{x}-b)\|_\infty$ の上限 β を求めるアルゴリズム

function$[\beta]$ = numerator2$(A_{mid},A_{rad},b_{mid},b_{rad},\tilde{x},R)$
setround(0);
$c_1'=\mathrm{fl}(R(A_{mid}\tilde{x}-b_{mid}))$;
$c_2'=\mathrm{fl}\left(\frac{\tilde{\gamma}_{n+1}|R||A_{mid}\tilde{x}-b_{mid}|}{1-(n+1)\mathbf{u}}\right)$;
$c_3'=\mathrm{fl}\left(\frac{\tilde{\gamma}_{n+2}|R|(|A_{mid}||\tilde{x}|+|b_{mid}|)}{1-(2n+2)\mathbf{u}}\right)$;
$c_4=\mathrm{fl}\left(\frac{\||R|(A_{rad}|\tilde{x}|+b_{rad})\|_\infty}{1-(2n+2)\mathbf{u}}\right)$;
$\beta=\mathrm{fl}\left(\frac{\||c_1'|+|c_2'|+c_3'+c_4\|_\infty}{1-4\mathbf{u}}\right)$;
end

4 数値実験

それぞれのアルゴリズムを実装し数値実験を行った。各項目において異なる行列で10回計算を行い、その平均値を表に表す。表における"*"は $\alpha\ge1$ となった場合であり，精度保証失敗を表す.

- 計算環境

1. CPU: Intel(R) Xeon(R) CPU X5680 @3.33GHz (ハイパースレッディング有/24コア相当)

2. Memory: 96 GB
3. OS: Yellow dog Linux for CUDA 6.1
4. GPU: NVIDIA TeslaC2070
5. MATLAB (R2012a, Parallel computing tool box), GotoBLAS2 1.13

- テスト行列の作成方法

$$\mathrm{randn}('\mathrm{state}', i)$$
$$A = \mathrm{randn}(n)$$
$$A = \mathrm{gallery}('\mathrm{randsvd}', \mathrm{n}, 10^{\mathrm{s}}, 3, \mathrm{n}, \mathrm{n}, 1)$$
$$A_{mid} = A$$
$$A_{rad} = |A_{mid}| * u$$
$$b_{mid} = A * e$$
$$b_{rad} = A_{rad} * e$$

- 数値実験結果

テスト行列を $A = \mathrm{randn}(n)$ で作成し、行列サイズ n を変えながら各アルゴリズムでの計算結果を比較したものが表 1, 表 2, 表 3 である。また、テスト行列を $n = 1000, A = \mathrm{gallery}(\cdots)$ で作成し、条件数 10^s を変えながら各アルゴリズムでの計算結果を比較したものが表 4, 表 5 である。

表 1.$\|RA - I\|_\infty$ の上限 α の比較

n	std.rnd2	Alg 3.1,3.2	GPU
100	7.02e-12	3.03e-10	3.03e-10
500	9.09e-11	8.86e-09	8.84e-09
1000	1.77e-10	2.63e-08	2.63e-08
2000	4.35e-10	9.43e-08	9.44e-08
5000	1.03e-08	3.99e-06	4.00e-06
10000	1.80e-08	1.06e-05	1.06e-05

表 2.$\|x - \tilde{x}\|_\infty$ の上限 $\beta/(1-\alpha)$ の比較

n	std.rnd2	Alg 3.1,3.2	GPU
100	5.02e-12	3.36e-10	3.36e-10
500	2.98e-11	9.21e-09	9.21e-09
1000	4.39e-11	2.70e-08	2.70e-08
2000	8.07e-11	9.61e-08	9.61e-08
5000	1.34e-09	4.04e-06	4.04e-06
10000	1.84e-09	1.07e-05	1.07e-05

表 3. 計算時間の比較 (秒)

n	std.rnd2	Alg 3.1,3.2	GPU
100	0.0044	0.0022	0.0215
500	0.0563	0.0368	0.0864
1000	0.2405	0.1676	0.1779
2000	1.1197	0.7759	0.6409
5000	14.9551	11.2420	4.6051
10000	105.0689	60.7180	26.0132

表 4.$\|RA - I\|_\infty$ の上限 α の比較

cond(A)	std.rnd2	Alg 3.1,3.2	GPU
10^2	4.78e-12	8.91e-10	9.02e-10
10^4	1.68e-10	4.78e-08	4.85e-08
10^6	9.70e-09	3.34e-06	3.39
10^8	6.76e-07	2.60e-04	2.64e-04
10^{10}	4.96e-05	0.0216	0.0220
10^{12}	0.0041	*	*
10^{14}	*	*	*

* : verification failed ($1 \leq \alpha$)

表 5.$\|x - \tilde{x}\|_\infty$ の上限 $\beta/(1-\alpha)$ の比較

cond(A)	std.rnd2	Alg 3.1,3.2	GPU
10^2	1.58e-12	9.17e-10	9.17e-10
10^4	7.45e-11	4.92e-08	4.92e-08
10^6	5.18e-09	3.44e-06	3.44e-06
10^8	3.94e-07	2.68e-04	2.68e-04
10^{10}	3.28e-05	0.0228	0.0228
10^{12}	0.0028	*	*
10^{14}	*	*	*

* : verification failed ($1 \leq \alpha$)

5 まとめ

本報告では、A, b が区間を持つ場合の $Ax = b$ における最近点丸めのみを用いた精度保証法を提案した。提案手法は、丸め変更を用いる手法より行列積の計算回数が少ないため高速に計算できる。一方、事前誤差評価による過大評価によって精度が 1~2 桁悪くなる。

参考文献

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

[2] S. Oishi, S. M. Rump: Fast Verification Solutions of Matrix Eqiations, Numer. Math., 90:4 (2002), 755-773.

[3] T. Ogita, S.M. Rump, S. Oishi, Verified solution of linear systems without directed rounding, Technical Report, No. 2005-04, Advanced Research Institute for Science and Engineering, Waseda University.