

線形方程式の精度保証付き解法 高速アルゴリズム

2007MI137 三輪嵩

指導教員：杉浦洋

1 はじめに

コンピュータでは、高速に数値計算するために、実数を浮動小数点数で近似して計算を行っている。有限桁の小数で近似しているものの例に円周率などの無理数がある。よって、得られる数値計算の結果も近似である。従来はこの数値を検算するのに精度と計算量の点で問題があったが、今では多くの研究により、数値計算による近似計算の検算が、より精度よく、より高速にできるようになった。

本研究では、線形方程式 $Ax + b$ の解を精度保証付きで求める方法を考える。理論上はガウスの消去法を区間演算で実行することによって精度が保証されるが、元数が大きくなるにつれ解の存在区間が過大になり実用性に欠ける。最近、効率的な精度保証付き数値計算の方法が開発され、元数の大きい方程式に対しても、実用的な精度保証が得られるようになってきた。本研究では、その方法を学び、理解を深め、高速な精度保証アルゴリズムの構成を目指す。

2 精度保証付き計算について

2.1 浮動小数点数とは

浮動小数点数システムの規格として、IEEE754 が多くのコンピュータで標準的に用いられている。本研究では IEEE754 に基づく 2 進数浮動小数点数システムを利用する。IEEE754 による浮動小数点数には 5 つのタイプが規定されている。規格化 2 進浮動小数点数、零、非規格化 2 進浮動小数点数、NaN の 4 つである。前 3 つを 2 進浮動小数点数とよぶ。

2 進浮動小数点数とは

$$a = \pm \left(\sum_{n=0}^{\infty} 2^{-k} d_k \right) \times 2^e, d_i \in \{0, 1\}$$

と書ける数である。e は指数という。倍精度では $-1022 \leq e \leq 1023$ となる。倍精度 2 進浮動小数点数全体の集合を \mathbb{F} とする。

3 精度保証のための基本定理

[定理 1] 方程式 $Ax = b$ の近似解 \tilde{x} と A の逆行列の近似行列 A^+ が求められたとき、行列 $R = A^+A - I$ が不等式

$$\|R\|_{\infty} < 1$$

を満たすとき、真の解 x^* が存在し、

$$\|x^* - \tilde{x}\|_{\infty} \leq \frac{\|A^+r\|_{\infty}}{1 - \|R\|_{\infty}} \leq \frac{\|A^+\|_{\infty}\|r\|_{\infty}}{1 - \|R\|_{\infty}} \quad (1)$$

が成り立つ。ここで $r = b - A\tilde{x}$ は残差である。//

式 (1) の右辺を精度保証付きで計算することにより、誤差 $\|x^* - \tilde{x}\|_{\infty}$ の上界が得られる。

4 高速精度保証

式 (1) の右辺の計算で、計算量の大きい部分は逆行列 A^+ の計算と残差 R の計算で、それぞれ計算量は $O(n^3)$ flops である。特に、分母は通常第 1 項の 1 が主要部で、 $\|R\|_{\infty}$ は高い精度を要しない。例えば $\|R\|_{\infty} = 10^{-5}$ のとき、それが 100 倍に計算されたとしても右辺の値はほとんど変化しない。ここで述べる高速精度保証アルゴリズムの基本的なアイデアは、 $\|R\|_{\infty}$ を直接評価せず事前誤差評価で代用し、計算量を $O(n^2)$ flops に削減することである。

以下、行列 A の数値 LU 分解 $A = \tilde{L}\tilde{U}$ と \tilde{L}, \tilde{U} の数値逆行列 $\tilde{L}^+ \cong \tilde{L}^{-1}, \tilde{U}^+ \cong \tilde{U}^{-1}$ がすでに計算されているという前提下で、 $\|R\|_{\infty}$ を $O(n^2)$ flops で計算するアルゴリズムを示す。

$A \in \mathbb{F}^{n \times n}$ の適切な LU 分解法を用いれば、

$$A \cong \tilde{L}\tilde{U} \quad (2)$$

となる。すると次の不等式

$$|A - \tilde{L}\tilde{U}| \leq \gamma_n |\tilde{L}\tilde{U}| \quad (3)$$

が成立する [1]。ここで

$$\gamma_n = \frac{nu}{1-u}$$

であり、 $u \cong 10^{-16}$ は丸め誤差単位である。

下三角行列 \tilde{L} の数値逆行列を \tilde{L}^+ とすると、適切な逆行列計算法を用いれば、

$$|\tilde{L}^+\tilde{L} - I| \leq \gamma_n |\tilde{L}^+||\tilde{L}|. \quad (4)$$

上三角行列 \tilde{U} の数値逆行列を \tilde{U}^+ とすると

$$|\tilde{U}^+\tilde{U} - I| \leq \gamma_n |\tilde{U}^+||\tilde{U}|. \quad (5)$$

が成立する [1]。

定理 1 における A^+ として

$$A^+ = \tilde{U}^+\tilde{L}^+ \cong \tilde{U}^{-1}\tilde{L}^{-1} = (\tilde{L}\tilde{U})^{-1} \cong A^{-1} \quad (6)$$

を採用する。このとき

$$\|R\|_{\infty} = \|A^+A - I\|_{\infty} = \|\tilde{U}^+\tilde{L}^+A - I\|_{\infty}$$

$$= \|\tilde{U}^+\tilde{L}^+A - \tilde{U}^+\tilde{L}^+\tilde{L}\tilde{U}\|_{\infty} + \|\tilde{U}^+\tilde{L}^+\tilde{L}\tilde{U} - I\|_{\infty}$$

最右辺第一項は

$$\|\tilde{U}^+\tilde{L}^+(A - \tilde{L}\tilde{U})\|_{\infty} \leq \gamma_n \|\tilde{U}^+||\tilde{L}^+||\tilde{L}||\tilde{U}|\|_{\infty}$$

最右辺第二項は

$$\begin{aligned} & \| \tilde{U}^+ \tilde{L}^+ \tilde{L} \tilde{U} - I \|_\infty \\ & \leq \gamma_n \| \tilde{U}^+ \| \tilde{L}^+ \| \tilde{L} \| \tilde{U} \|_\infty + \gamma_n \| \tilde{U}^+ \| \tilde{U} \|_\infty \end{aligned}$$

ゆえに

$$\begin{aligned} \| R \|_\infty & \leq 2\gamma_n \| \tilde{U}^+ \| \tilde{L}^+ \| \tilde{L} \| \tilde{U} \|_\infty \\ & \quad + \gamma_n \| \tilde{U}^+ \| \tilde{U} \|_\infty \end{aligned} \quad (7)$$

である。

行列 $A \in \mathbb{F}^{n \times n}$ について, $A \geq 0$ なら, $\|A\|_\infty = \|Ae\|_\infty$ である。

ここで, $e = (1, 1, \dots, 1)^T \in \mathbb{R}^n$ である。このことより, (7) の右辺の各項は

$$\begin{aligned} \| \tilde{U}^+ \| \tilde{L}^+ \| \tilde{L} \| \tilde{U} \|_\infty & = \| \tilde{U}^+ \| \tilde{L}^+ \| \tilde{L} \| \tilde{U} \| e \|_\infty \\ \| \tilde{U}^+ \| \tilde{U} \|_\infty & = \| \tilde{U}^+ \| \tilde{U} \| e \|_\infty \end{aligned} \quad (8)$$

で計算できる。(8)の両式の右辺は, 行列×ベクトル計算のみで, $O(n^2)$ の計算量で計算できる。

5 数値実験

5.1 実験概要

使用した計算機は, 東芝 dynabook MX/25AE である。Mathematica Ver.6.0 上で, $\|R\|_\infty$ の計算を通常の方法で行うプログラムと, 高速化を目的としたプログラムを作成した。そして, n 次元線形方程式 $Ax = b$ を解く実験を行った。係数行列 A には Frank 行列と Hilbert 行列を用い, 真の解 $x^* = (1, 1, \dots, 1)^T$ となるように右辺ベクトル b を与えた。

5.2 実験結果 (Frank 行列)

係数行列を Frank 行列とした。方程式は良条件である。LU 分解法で近似解 \tilde{x} を求め, これを精度保証するために, 誤差ノルムの上界 d を求めた。すなわち, $\|x^* - \tilde{x}\|_\infty \leq d$ が成立すれば, 精度保証は成功である。次数 $n = 2^m$ ($1 \leq m \leq 8$) で双方のアルゴリズムによる実験を行い, 双方とも全ての方程式で近似解の精度保証に成功した。結果を図 1 に示す。

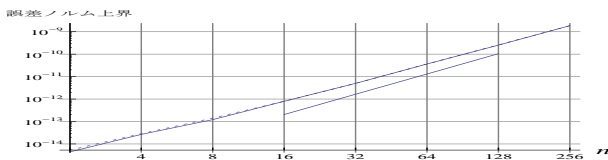


図 1 誤差ノルムの上界

実線は高速, 点線は従来アルゴリズムの結果であるが, 両者はほとんど重なっている。また, 細い短い直線は $d \propto n^3$ を表す。このグラフにより, 高速アルゴリズムは従来アルゴリズムと同程度精密な d を計算していることが分かる。また, ほぼ $d \propto n^3$ が成立していることが分かる。

良条件方程式については, 数値逆行列の残差ノルム $\|R\|_\infty$ は小さいため, 事前誤差評価に基づく荒い評価でも結果に影響を及ぼさないことが分かる。

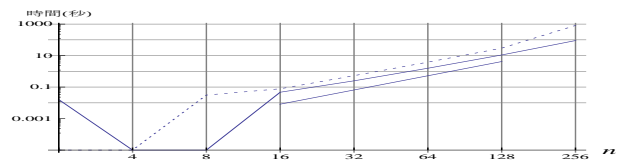


図 2 計算の所要時間

図 2 は所用時間を測定したものである。同様に実線は高速, 点線は従来アルゴリズムの結果である。全ての n で高速アルゴリズムの方が速い。また n が大きいほど高速アルゴリズムと従来アルゴリズムの速度比は大きくなり, $n = 256$ のときは速度比が 10 近くになる。また, 所要時間は $d \propto n^3$ であることが見て取れる。これは, 所要時間がほぼ計算量と比例していることを表す。

5.3 実験結果 (Hilbert 行列)

係数行列を Hilbert 行列とした。方程式は悪条件であり, n が大きくなるにつれ, 急激に条件数が増加し, 数値誤差が増大する。次数 $1 \leq n \leq 12$ で双方のアルゴリズムによる実験を行い, $1 \leq n \leq 11$ では双方とも全ての方程式で近似解の精度保証に成功したが, $n = 12$ のときは, プログラムで $\|R\|_\infty \geq 1$ と判定されたため, 精度保証は出来なかった。 $n = 11$ までの結果を図 3 に示す。

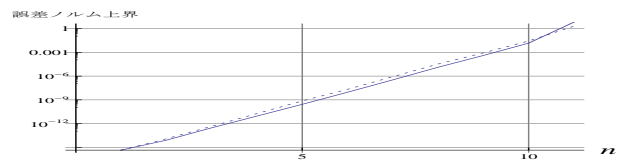


図 3 誤差ノルムの上界

$n = 11$ のときを除き, 高速アルゴリズムの方がやや小さい上界を与えており優秀である。これは式 (1) の分子の計算法の違いによるものと考えられる。 $n = 11$ ではこの関係が逆転する。悪条件が原因で, 数値逆行列の残差ノルム $\|R\|_\infty$ が大きくなり, 高速アルゴリズムの $\|R\|_\infty$ に対する荒い評価が結果に悪影響を及ぼしたのと考えられる。

6 終わりに

本研究では, 計算を高速に行うことができるプログラムの実装に成功した。今後, さらに計算の効率化に関する研究も必要である。

参考文献

- [1] 大石進一:「精度保証付き数値計算」. コロナ社 (2000).
- [2] 杉浦洋:「数値計算の基礎と応用-数値解析学への入門」. サイエンス社 (1997).
- [3] 後藤佳紀:「線形方程式の精度付き解法」. 南山大学数理情報学部数理科学科卒業論文 (2007).