チュートリアル
「精度保証付き数値計算の基礎」

精度保証付き数値計算の概要

荻田 武史
東京女子大学

2018/9/10

数値計算は，解析的に解くことが困難な問題を数値的に解く計算手法であるが，これは実数演算のような厳密な計算ではなく近似計算であり，計算途中で様々な誤差が発生するため，**最終的に得られた結果がどれくらい正しいかは問題依存**である。

数値計算によって得られた結果に対して，数学的に厳密な誤差限界を与える手法が**精度保証付き数値計算**である。

ここでは，いくつかの例を交えながら，精度保証付き数値計算の有用性や必要性について述べる。

計算機援用証明

計算機を用いて数学的な定理を証明することを**計算機援用証明**(computer-assisted proof) と呼ぶ。

精度保証付き数値計算は，**従来の数値計算に数学的な厳密性を付加する**ものであり，計算機援用証明のための新しい強力なツールと成り得る。

以下は，実際に精度保証付き数値計算によって計算機援用証明に成功した代表的な例である。

- ローレンツアトラクターの存在検証 [10]
  （スメイルの第14番目の問題）
- ケプラー予想（球充填問題）の肯定的解決 [3, 4]
  （約400年間の未解決問題）
- Double Bubble予想の肯定的解決 [5]
  （100年間以上の未解決問題）

区間演算

連続した数の集合は閉区間によって表現できる。

区間演算は，実数を区間に，実数演算を区間による演算に，それぞれ置き換えたものである。

何らかの計算を実数演算の代わりに区間演算で実行すると，得られた結果は必ず実数演算による結果を含む区間となる。

例)

$$\pi + \sqrt{2} \in [3.14, 3.15] + [1.41, 1.42] = [4.55, 4.57]$$

残念ながら，実際には，実数演算を単純に区間演算で置き換えただけでは，意味のある結果を得られないことが多いことがわかっている。

区間演算を繰り返すと**区間幅が指数関数的に増大してしまう。**

たとえば，連立一次方程式に対する区間ガウスの消去法は，その典型的な例である。(3章で説明)

このような区間演算の振る舞いについて，丸め誤差解析で著名なJ. H. Wilkinsonは以下のように述べている[11]。

> 区間演算は役に立たないわけではないが，適用可能な状況に至るまでに深刻な制限がある。
>
> 一般に，代数的な計算に対して区間演算を有効な手段とするためには，その使用をできる限り後回しにすることが最良である。

すなわち，通常の数値計算によって**近似解を得た後**に，区間演算によってその近似解の精度を保証する，という考え方が重要である。

丸め誤差の影響

数値計算で用いる浮動小数点演算は有限桁の計算であるため丸め誤差が発生するが，実際にどのような影響があるか，いくつかの例を挙げる。

丸め誤差が，そこまで計算結果に深刻な影響を与えるのか？

Rumpの例題

一般に，数値計算では演算精度が高いほど結果の精度も高くなる傾向がある。

ある演算精度で何らかの計算をして，次にそれよりも高い演算精度で同じ計算をしたときに，双方の結果が近ければ，ある程度は結果の正しさが確認できる？

この経験則は，確かに有効な場合もあるが，残念ながら常に正しいわけではない。

1980年代に，S. M. Rumpは以下の例題を考案した[9]。

$$f(x,y) = 333.75y^6+x^2(11x^2y^2-y^6-121y^4-2)+5.5y^8+\frac{x}{2y}$$

$a = 77617$, $b = 33096$のとき，$f(a,b)$の値を異なる演算精度で計算する。

$$\begin{aligned} \text{単精度:} \quad & f(a,b) \approx 1.172603\cdots \\ \text{倍精度:} \quad & f(a,b) \approx 1.1726039400531\cdots \\ \text{拡張精度:} \quad & f(a,b) \approx 1.172603940053178\cdots \\ \textbf{真の値}\text{:} \quad & f(a,b) = -0.827386\cdots \end{aligned}$$

このように，経験則では対処できない問題もある。

この問題に対して倍精度の区間演算を用いると

$$[-5.91, 4.73] \times 10^{21}$$

という結果を得る。これは，非常に区間幅が大きいため，あまり意味のある結果ではないが，少なくとも真の値を含んでいる。

つまり，**区間演算は間違った答えを決して出さない**ということが重要である。

また，このように区間幅が大きい結果を得たことによって，深刻な丸め誤差が発生していることに気付くことができる場合がある。

2元連立一次方程式

連立一次方程式$Ax = b$の近似解を$\widehat{x}$とすると，その残差は$r := b - A\widehat{x}$と定義される。

もしAが正則で$r = \mathbf{0}$であれば，$\widehat{x}$は真の解である。

rの要素の大きさが小さければ，$\widehat{x}$の精度は良い？

そこで，以下のような例題[7]を考えてみよう。

$$A = \begin{pmatrix} 64919121 & 159018721 \\ 41869520.5 & 102558961 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

このとき，Aは正則で真の解は

$$x = A^{-1}b = \begin{pmatrix} 205117922 \\ -83739041 \end{pmatrix}$$

である。

これに対して，IEEE 754の倍精度浮動小数点演算を用いてガウスの消去法で解を計算すると

$$\widehat{x} = \begin{pmatrix} 106018308.007133 \\ -43281793.0017831 \end{pmatrix}$$

のように1桁も合っていない結果が得られる。

ところが，この$\widehat{x}$に対して残差rを倍精度演算を用いて計算すると，残差の近似は$\widehat{r} = (0,0)^T$ となり，一見$\widehat{x}$は正しい解のように見えてしまう[1]。

残差の計算からでは解が正しいかどうかを判定することができない。

この問題に対して倍精度の区間演算を用いてガウスの消去法で解を計算すると，$-\infty < x_i < \infty$ $(i = 1, 2)$ という区間としては無意味な結果を得るが，これは問題の方程式が解きづらいことを示唆している。

[1]真の残差は$r = (0.616\cdots, 0.085\cdots)^T$である。

IEEE 754–1985浮動小数点演算規格の制定に尽力したW. M. Kahanは，以下のように述べている[6]。

> 浮動小数点演算によって得られた結果と真値に大きな差が生じることは非常に稀であり，常に心配するにはあまりにも稀であるが，だからと言って無視できるほど稀なわけではない。

すなわち，絶対に間違ってはいけないような計算をする場合，丸め誤差を無視してはいけない。

また，丸め誤差だけでなく**打ち切り誤差**や**離散化誤差**も考慮に入れて計算する必要がある。

打ち切り誤差

例えば，$\sin x$ のテイラー展開

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} + \cdots + (-1)^{n-1}\frac{x^{2n-1}}{(2n-1)!} + R_{2n+1}(x) \quad (1)$$

$$R_{2n+1}(x) = (-1)^{n+1}\frac{\cos \xi x}{(2n+1)!}x^{2n+1}, \quad 0 < \xi < 1$$

を用いて，$\sin \frac{\pi}{6}$ の値を計算することを考えよう。

式(1)の右辺を第5項までで打ち切る（$n = 5$）

$$S(x) := x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \frac{x^9}{9!}$$

$x = \frac{\pi}{6}$として$S(\pi/6)$を倍精度浮動小数点演算による区間演算で丸め誤差を考慮しながら計算する。

$\pi/6$も倍精度浮動小数点数では厳密には表現できないため，これを含む区間としてxに代入する。

$$S(\pi/6) \in [0.50000000002027, 0.50000000002029]$$

が得られるが，打ち切り誤差のため$\sin\frac{\pi}{6}$の真値$\frac{1}{2}$を含んでいない。

打ち切り誤差は，剰余項$R_{11}(x) = \dfrac{\cos \xi x}{11!} x^{11}$で与えられる。

ξの値は分からないが,

$$-1 \leq \cos \xi x \leq 1$$

であることを利用して区間演算することによって$R_{11}(\pi/6)$を包含することができるため，これを用いて区間演算を行うと

$$\sin \frac{\pi}{6} \in [0.49999999999, 0.50000000005]$$

のように数学的に厳密に$\sin \frac{\pi}{6}$の真値を包含可能となる。

離散化誤差

例えば，以下の定積分

$$I = \int_{-1}^{1} f(x)dx, \quad f(x) = \frac{2}{x^2+1}$$

を計算する問題を考えよう（Iの真値はπである)。

これを台形公式によって計算する。

$$h = (b-a)/n, \quad x_i = a + ih,\ i = 1, 2, \ldots, n-1$$

$$\int_a^b f(x)dx \approx \frac{h}{2}\left(f(a) + 2\sum_{i=1}^{n-1} f(x_i) + f(b)\right) =: S$$

倍精度浮動小数点演算を用いて$n = 100$で計算すると

$$I \approx 3.1415259869232539$$

が得られる。

丸め誤差を把握するため，区間演算で計算すると

$$S \in [3.1415259869232365, 3.1415259869232673]$$

が得られるが，これは離散化誤差のため真の値を含んでいない（$I = \pi \notin S$）。

台形公式の誤差は，ある$\xi \in [a,b]$を用いると

$$\int_a^b f(x)dx = \frac{h}{2}\left(f(a) + 2\sum_{i=1}^{n-1} f(x_i) + f(b)\right) - \frac{b-a}{12}h^2 f''(\xi)$$

ξの値は分からないが，$f''(x) = \dfrac{4(3x^2-1)}{(x^2+1)^3}$ に対して，x に区間$[-1,1]$を代入して区間演算することによって$f''(\xi)$を包含することができる。

これと台形公式において区間演算を行うと

$$I \in [3.1404, 3.1421]$$

この $f''(x)$ に対する区間演算で見られるように，区間演算は丸め誤差の把握だけでなく，**関数の値域の評価**にも用いられる。

また，微分方程式のような連続な系を数値計算によって解くためには離散化が必要となるが，その際にも離散化誤差が発生する。

極端な場合，それによって「幻影解」が現れることもある。すなわち，離散化された問題の解が，元の問題の解に対する良い近似にならない場合がある。

例として，領域$\Omega = (0, a) \times (0, 1/a)$, $a > 0$の上で，以下のEmden方程式と呼ばれる非線形偏微分方程式の境界値問題の解を考える。

$$-\Delta u = u^2 \text{ in } \Omega, \quad u = 0 \text{ on } \partial\Omega \tag{2}$$

Breuer-Plum-McKennaは，$a = 2.9$の場合に図1のような近似解が得られることを報告した[1]。

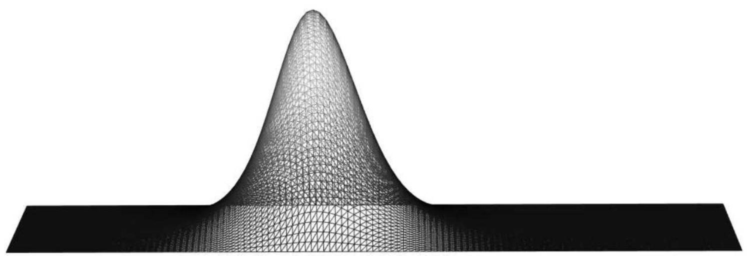

図 1: 幻影解

ところが，この問題は，理論的な解析手法によって，領域の中心についてx方向またはy方向に非対称な解が存在しないということが証明されている[2]。

つまり，図1のような非対称な解は問題(2)の解の良い近似にならない幻影解である。

上記の例は極めて珍しいものであるが，非線形微分方程式の解の存在などを厳密に検討するためには，数値解法によって得られた近似解を検証しなければならないことが分かる。

解の存在証明

不動点定理の成立を数値的に確かめることによって方程式の解の存在保証と存在範囲の保証が可能。

常微分方程式の初期値問題

$$\frac{dx}{dt} = f(x(t), t), \quad x(t_0) = x_0, \quad t \in [t_0, t_1]$$

は，両辺を 0 から t まで積分して

$$x(t) = x_0 + \int_0^t f(x(t), t) dt$$

と変換できる。

写像 $P : C[t_0, t_1] \to C[t_0, t_1]$ を

$$P : x(t) \mapsto x_0 + \int_0^t f(x(t), t)dt$$

とする。

Pが$C[t_0, t_1]$のある集合XをX自身に移すならば，Pのコンパクト性とシャウダーの不動点定理により，Pの不動点がXに存在し，それは元の初期値問題の解の存在を意味する。

これを小さな例題で確かめてみよう。

$$\frac{dx}{dt} = -x^2, \quad x(0) = 1, \quad t \in [0, 0.1]$$

に対して，$P : C[0, 0.1] \to C[0, 0.1]$ を

$$P : x(t) \mapsto 1 + \int_0^t (-x(t)^2) dt$$

とする。ここで，$X \subset C[0, 0.1]$を

$$X = \{x(t) \in C[0, 0.1] \mid 0.8 \le x(t) \le 1.2\}$$

とすると

$$
\begin{aligned}
& 0.8 \le x \le 1.2 \quad \Rightarrow \quad 0.64 \le x^2 \le 1.44 \\
\Rightarrow \quad & -1.44 \le -x^2 \le -0.64 \\
\Rightarrow \quad & 1 + \int_0^t (-1.44)dt \le 1 + \int_0^t (-x^2)dt \le 1 + \int_0^t (-0.64)dt \\
\Rightarrow \quad & 1 - 1.44t \le 1 + \int_0^t (-x^2)dt \le 1 - 0.64t
\end{aligned}
$$

と簡単な計算により，$P(X) \subset X$が分かる（図2を参照）。

これにより，$P(X)$に初期値問題の解が存在することが数学的に保証される。

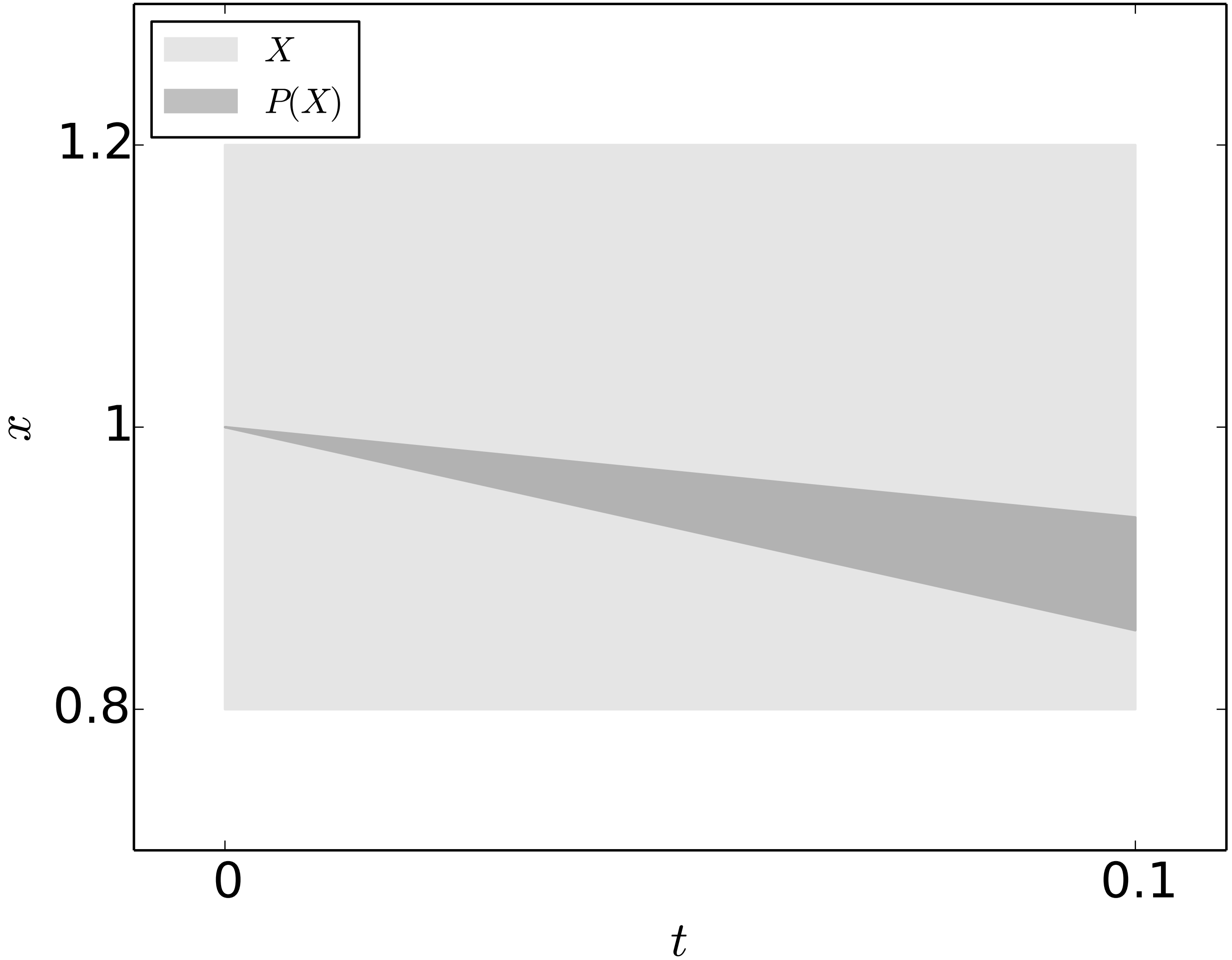

図 2: 不動点定理の成立

最後に，精度保証付き数値計算やその中で活用される数学的な理論を学ぶ上で参考となる和書を列挙しておく。

1. 大石 進一：精度保証付き数値計算, コロナ社, 1999.

2. 大石 進一：非線形解析入門, コロナ社, 1997.

3. 中尾 充宏, 山本 野人：精度保証付き数値計算 – コンピュータによる無限への挑戦, 日本評論社, 1998.

4. 中尾 充宏, 渡部 善隆：実例で学ぶ精度保証付き数値計算 〜 理論と実装 〜, 臨時別冊・数理科学2011年10月 (SGCライブラリ **85**), サイエンス社, 2011.

References

[1] B. Breuer, M. Plum, P. J. McKenna: Inclusions and existence proofs for solutions of a nonlinear boundary value problem by spectral numerical methods, Topics in Numerical Analysis, G. Alefeld, X. Chen (eds.), Springer Vienna, 2001, pp.61–77.

[2] B. Gidas, W.-M. Ni, L. Nirenberg: Symmetry and related properties via the maximum principle, Commun. Math. Phys., **68** (1979), 209–243.

[3] T. C. Hales: A proof of the Kepler conjecture, Ann.

Math., **162** (2005), 1065–1185.

[4] T. C. Hales et al.: A formal proof of the Kepler conjecture, Forum of Mathematics, Pi, **5** (2017), e2, 29 pages.

[5] J. Hass, M. Hutchings, R. Schlafly: The double bubble conjecture, Electron. Res. Announc. Amer. Math. Soc., **1** (1995), 98–102.

[6] W. M. Kahan: The Regrettable Failure of Automated Error Analysis, A Mini-Course prepared for the

conference at MIT on Computers and Mathematics, 1989.

[7] U. W. Kulisch, W. L. Miranker: The Arithmetic of the Digital Computer: A New Approach, SIAM Review, **28** (1986), 1–40.

[8] E. Loh, G. W. Walster: Rump's example revisited, Reliable Computing, **8** (2002), 245–248.

[9] S. M. Rump: Algorithms for verified inclusions: theory and practice, Reliability in Computing: The Role of

Interval Methods in Scientific Computing (R. E. Moore ed.), Academic Press, Boston, 1988, pp.109–126.

[10] W. Tucker: The Lorenz attractor exists, C. R. Acad. Sci. Paris Sér. I Math., **328** (1999), 1197–1202.

[11] J. H. Wilkinson: Modern error analysis, SIAM Review, **13** (1971), 548–568.