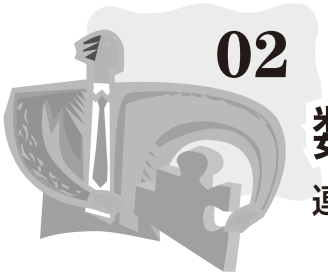

02

数値計算における数学

連立一次方程式の求解法を題材として

基

岩下武史（北海道大学）

計算機で扱う数学

計算科学等の数値シミュレーションは多くの場合，数学や物理学に基づいている．線形代数や微積分といった多くの応用分野で共通的に用いられる数学もあるが，個々の分野において異なった数学的知識が求められるのが普通である．ここで，これらの個別の数学について述べることは，与えられた紙面から見て現実的ではない．そこで，本稿では数値シミュレーションで「数学」を扱う際に共通的に注意すべき点について述べることとする．

数値シミュレーション（数値計算）では，多くの場合（倍精度）浮動小数点数により，さまざまな量が表現される（**図 -1**）．ここで注意しなければならないのは，数値計算で扱われる“数”は誤差を含んでいるという点である．たとえば，倍数度浮動小数点数は 64 桁の 2 進数（64 ビット）で 1 つの数を表すが，π や $\sqrt{2}$ といった無理数はもとより多くの数が近似的に表され，その値には誤差が含まれる．これらの誤差の影響はしばしば計算の進行に伴って拡大し，最終的にシミュレーションの精度に多大な影響を及ぼすことがある．そこで，本稿では多くのシミュレーションにおいて主たる計算核となっている連立一次方程式の求解部を題材に，計算誤差に関連したいくつかの話題について解説する．なお，本稿では $n \times n$ の正則な実行列を係数とする実連立一次方程式を扱うものとする．

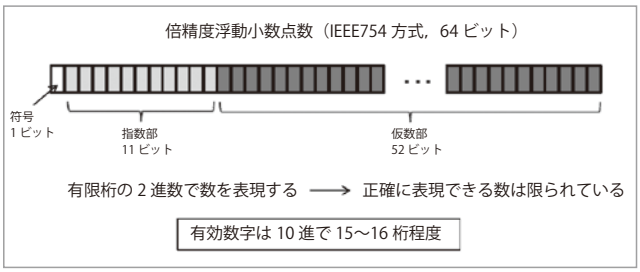

図 -1　倍精度浮動小数点数による数の表現

直接法と反復法

計算機による連立一次方程式の求解法には，直接法と反復法の 2 種類のカテゴリがある．直接法としてはガウスの消去法や LU 分解法等が知られ，行列の変換や代入計算を通じて解ベクトルの要素の各々を直接的に求めていく．一方，初期的に与えられる近似解に対して，ある手順を繰り返し適用することで解を更新し，最終的に十分な精度を持った近似解を得る方法が反復法である．たとえば，ヤコビ法やガウス＝ザイデル法，共役勾配法等が反復法としてよく知られている．さて，ここで「反復法では近似解を求めるに過ぎないが，直接法では正しい解を求めることができる」という考えを持つ人がいるが，これは誤った考えである．先に述べたように，計算機上での計算には誤差が含まれるためである．

以下の n 元連立一次方程式について考える．

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \tag{1}$$

ここで，本来の $\boldsymbol{b}$ に代わり，計算機上では誤差の影響で $\boldsymbol{b}+\delta\boldsymbol{b}$ が右辺ベクトルとなったとする．この場合，

$$\boldsymbol{A}\left(\boldsymbol{x}+\delta\boldsymbol{x}\right) = \boldsymbol{b}+\delta\boldsymbol{b} \tag{2}$$

と書くことができる．式（1）（2）より

$$\delta\boldsymbol{x} = \boldsymbol{A}^{-1}\delta\boldsymbol{b} \tag{3}$$

を得る．式（3）の両辺を $\|\boldsymbol{b}\|$ で割り，さらにノルムの性質（$\|\boldsymbol{A}\boldsymbol{x}\| \leq \|\boldsymbol{A}\|\,\|\boldsymbol{x}\|$）を使って整理すると

$$\frac{\|\delta \boldsymbol{x}\|}{\|\boldsymbol{x}\|} \leq \|\boldsymbol{A}\| \|\boldsymbol{A}^{-1}\| \frac{\|\delta \boldsymbol{b}\|}{\|\boldsymbol{b}\|} \tag{4}$$

を得る．ここで$\kappa(\boldsymbol{A}) = \|\boldsymbol{A}\| \|\boldsymbol{A}^{-1}\|$は条件数と呼ばれ，式（4）は右辺ベクトルの相対誤差が小さくても条件数が大きい場合には得られた解に大きな（相対）誤差が含まれる可能性があることを意味している．すなわち，直接法を用いた場合においても，数値や計算の誤差によって通常“正しい解”は求められず，条件数が大きい場合には解ベクトルに大きな誤差が含まれる場合があることになる．条件数が大きい問題は悪条件な問題と呼ばれ，精度の高い解を得ることは解法によらず困難となる[1]．

残差による近似解の判定

数値シミュレーションで連立一次方程式を解く場合，解ベクトルがどんな値になるかは事前には分からない．では，得られた解が“十分な精度を持つ”かどうかどのように判定すればよいのだろうか．あらかじめ分かっているのは係数行列と右辺ベクトルであるから，通常，以下の残差ベクトル$\boldsymbol{r}$を用いて判定する[2]．

$$\boldsymbol{r} = \boldsymbol{b} - \boldsymbol{A}\tilde{\boldsymbol{x}} \tag{5}$$

ここで，$\tilde{\boldsymbol{x}}$は計算によって得られた（近似）解である．確かに残差ベクトルが0ベクトルとなれば，得られた解は正しい解であると言える．しかし，計算誤差の問題があるため，現実的に残差ベクトルが0ベクトルとなることはほとんどない．そのため，反復法では，残差ベクトルのノルムが右辺ベクトルのノルムと比べて十分に小さい場合に（十分な精度を持つ）近似解が得られたとする場合が多い．しかしながら，残差ベクトルが十分に小さかったとしても近似解に大きな誤差が含まれる場合がある．

正しい解$\boldsymbol{x}$と近似解$\tilde{\boldsymbol{x}}$の間の誤差$\boldsymbol{e}$を

$$\boldsymbol{e} = \boldsymbol{x} - \tilde{\boldsymbol{x}} \tag{6}$$

のように表すものとする．このとき，式（1）（5）（6）より

$$\boldsymbol{A}\boldsymbol{e} = \boldsymbol{r} \tag{7}$$

が成り立つ．ここで，誤差ベクトル$\boldsymbol{e}$が$\boldsymbol{A}$の固有ベクトルであり，かつその固有値の絶対値が非常に小さいと仮定する．このような場合，$\boldsymbol{e}$のノルムがかなり大きな値を持っていたとしても$\boldsymbol{r}$のノルムは小さい値となる．つまり，このようなケースでは残差ベクトルのノルムを（十分に）小さくしたとしても，依然として大きな誤差が解ベクトルに含まれる可能性が残ることとなる．もう少し詳細に見てみよう．

式（7）に対して，式（4）の導出手順と同様に

$$\frac{\|\boldsymbol{e}\|}{\|\boldsymbol{x}\|} \leq \kappa(\boldsymbol{A}) \frac{\|\boldsymbol{r}\|}{\|\boldsymbol{b}\|} \tag{8}$$

を得る．式（8）は，相対残差（残差ベクトルと右辺ベクトルのノルム比）が小さかったとしても係数行列の条件数が非常に大きい場合，大きな誤差が生じ得ること意味している．つまり，悪条件な問題では相対残差が十分に小さい近似解が得られたとしても安心できないということとなる．

では，残差ベクトルのノルムのほかによい指標がなく，相対残差を使わなければならないとするとどのような基準とすればよいのだろうか？　残念ながら，この問いに対する一般的な解答はない．それぞれの応用分野において，過去の経験や実験結果との照合等によって適宜決められているのが現状である．ただし，留意すべき事項として，連立一次方程式の求解においてより重要なことは残差ベクトルよりも誤差ベクトルのノルムを小さくすることにあることを挙げることができる．複数の解法を比較する場合には残差ベクトルのノルムの変化だけではなく，あらかじめ解が分かっている人工的な問題に対して誤差ベクトルの振舞いをチェックすることも時に重要となる．また，実応用分野では，得られた解が実験結果や物理現象をよく模擬しているかどうか十分に吟味する必要がある．

共役勾配法

連立一次方程式の求解法としてよく利用されている反復法に共役勾配（CG：Conjugate Gradient）法がある[3]．CG法は正値対称な係数行列を持つ連立一次方程式に有効な手法で有限要素解析を始めとした多くの解析で利用されている．また，こうした現状を踏まえて計算機やスーパーコンピュータの性能測定のためのベンチマークとしても活用されている．

CG法の長所は反復ごとに算出される残差ベクトルが互いに直交するという点にある．この性質から，CG法は最大でn回の反復数で解を導出することができる．つまり，有限の計算量で求解を行うことができるという点からは，直接法と同様の性質を持つと言うこともできる．しかしながらこの性質は誤差のない計算，つまり「紙と鉛筆」による計算の場合であり，数値計算では事情が異なってくる．

CG法の特徴である残差ベクトルの直交性を確保するために，CG法の反復手順には内積演算が含まれる．しかしながら，一般に内積演算の結果は数値誤差の影響を受けやすい．たとえば，第1番目の要素の絶対値が非常に大きく，それ以外の要素の絶対値が小さい，非常に長いN次元ベクトル$\boldsymbol{f}$と$\boldsymbol{g}$の内積を考える．$\boldsymbol{f}$および$\boldsymbol{g}$のi番目の要素をそれぞれf_i，g_iと表すと，通常の内積演算の手順では，f_1とg_1の積をまず求め，次にf_2とg_2の積をこれに加算し，順次f_iとg_iの積を計算し，加算する．この場合，f_1とg_1の積，すなわち$f_1 \cdot g_1$の絶対値が$f_2 \cdot g_2$の絶対値と比べて非常に大きいと，数値計算上は$f_1 \cdot g_1$に$f_2 \cdot g_2$を足しても結果がまったく変わらないということが生じ得る．このような現象が続いた場合，以降の計算は内積演算の結果に反映されないため，内積は$f_1 \cdot g_1$と計算されることになる．しかしながら，$\boldsymbol{f}$と$\boldsymbol{g}$の要素数が非常に多い場合，いわゆる「ちりも積もれば山となる」で計算結果に反映されなかった$f_2 \cdot g_2 + f_3 \cdot g_3 + \cdots$の（絶対）値が大きなものとなり，計算結果に大きな誤差が含まれるということが生じ得る．このような内積演算や他の計算部分による誤差の影響は，しばしば反復が進むにつれて拡大し，CG法における残差ベクトルの直交性を失わせ，近似解が収束しない（解が得られない）という現象を引き起こす要因となる．

前処理

前章でCG法において，内積演算等に起因する計算誤差の影響で求解のプロセスが破綻する場合があることを述べた．このような現象は他の反復法においても見られるもので，実応用上重要な問題である．本問題の対処法として広く用いられているのが前処理と呼ばれる技術である．前処理は反復法の収束性を向上させる手法で，主に求解時間の短縮を目的として利用されている．しかし，反復ごとに増大する計算誤差の影響が深刻化する前に求解プロセスを完了するという点から言えば，反復法における計算誤差の対処法の1つと捉えることもできる．

前処理とは連立一次方程式（1）を解く代わりに

$$\boldsymbol{M}_1^{-1}\boldsymbol{A}\boldsymbol{M}_2^{-1}\left(\boldsymbol{M}_2\boldsymbol{x}\right) = \boldsymbol{M}_1^{-1}\boldsymbol{b} \qquad (9)$$

を解く方法である．$\boldsymbol{M}_1$および$\boldsymbol{M}_2$（前処理行列）をうまく選ぶことにより，反復法の収束性を向上させることができる．あいまいな表現であるが，なんらかの意味で前処理後の係数行列$\tilde{\boldsymbol{A}} = \boldsymbol{M}_1^{-1}\boldsymbol{A}\boldsymbol{M}_2^{-1}$が単位行列に近い場合，反復法の収束性を改善することができる．つまり$\boldsymbol{M}_1\boldsymbol{M}_2 \simeq \boldsymbol{A}$となるような$\boldsymbol{M}_1$，$\boldsymbol{M}_2$を選ぶことができれば，$\tilde{\boldsymbol{A}} \simeq \boldsymbol{I}$（$\boldsymbol{I}$：単位行列）となり，前処理としての効果が期待できる．

CG法の場合，前処理後の係数行列$\tilde{\boldsymbol{A}}$の対称性を確保するために，$\boldsymbol{M}_2^{-1} = (\boldsymbol{M}_1^{-1})^T = (\boldsymbol{M}_1^T)^{-1}$，すなわち，$\boldsymbol{M}_2 = \boldsymbol{M}_1^T$とするのが一般的である．CG法の前処理に関する手順は各反復中で，

$$\boldsymbol{M}\boldsymbol{z} = \boldsymbol{r} \qquad (10)$$

を解くことにより与えられる．ここで，$\boldsymbol{M} = \boldsymbol{M}_1\boldsymbol{M}_1^T$である．したがって，前処理行列には，式（10）が式（1）よりも簡単に（より計算量が少なく）解けることが求められる．また，$\boldsymbol{M}_1$の導出に多大な計算量を必

要としないことも要求される．前処理に関する計算量が十分に小さいことや前処理行列が係数行列の良い近似であるという性質はCG法に限らず，ほかの反復法の前処理手法に対しても求められる性質である．

実応用分野でよく用いられている前処理手法として，定常反復法前処理や不完全分解前処理，マルチグリッド前処理等を挙げることができる．これらの前処理技術を利用することにより，収束性を大幅に向上し，求解時間を短縮できる例が数多く報告されている．また，前処理を用いない場合，数値計算上十分な精度の近似解を求めることができないといった例もしばしば見られる．こうした背景から，応用分野の中には，前処理を用いることが“常識”とされ，前処理技術の活用が必須となっている分野が存在している．

並列計算と高精度計算

現在，数値シミュレーションの計算環境としてマルチコアプロセッサやクラスタ等，並列計算機が広く普及している．並列計算を行う場合，計算誤差により逐次計算とは異なる結果が生じる場合がある．たとえば，内積演算を2つのスレッドで**図-2**のように並列計算する場合を考える．各々のスレッドで部分和を計算し，その結果を加算して最終的な演算結果とする場合，数学上は逐次計算と同じ結果となるはずであるが，数値計算上では必ずしもそうではない．つまり，誤差のない計算の場合には同一の結果が得られるような並列処理を行った場合でも，数値計算上は異なった結果となり得る．反復法による連立一次方程式の求解の例では，対象とする問題や使用する解法に依存して，並列計算を行った場合に求解までの反復数が5～10%程度変動することが観測されている．このように並列計算では，「数値計算」と「紙と鉛筆による計算」の違いにより注意する必要がある．

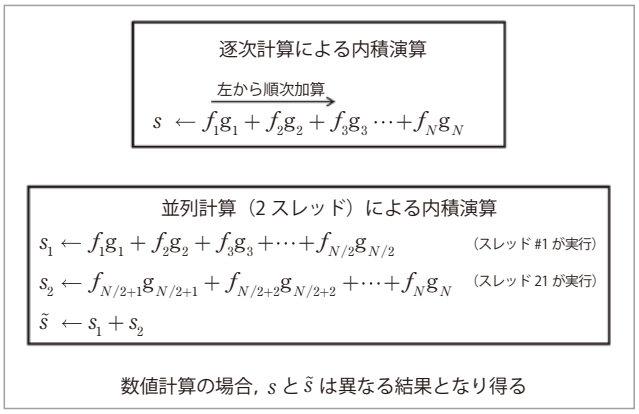

図-2　内積演算の並列処理

近年では，スパコン等の大規模な並列計算機によるシミュレーションが広く行われ，解析規模の拡大も著しい．このような大規模計算では，これまでに述べてきたような計算誤差の問題がより顕著となりやすい．そのため，従来の倍精度浮動小数点数による精度を越える計算精度が要求される場合が生じている．こうした背景のもとで，4倍精度演算や多倍長計算，精度保証計算といった高精度な解析を実現するための計算法に関する研究が近年活発に行われている．

参考文献
1) Demmel, J. : Applied Numerical Linear Algebra, SIAM, Philadelphia, PA (1997).
2) Barrett, R., et al. : Templates for the Solution of Linear Systems, Building Blocks for Iterative Methods, SIAM, Philadelphia, PA (1994).
3) 杉原正顯，室田一雄：線形計算の数理，岩波書店 (2009).

(2015年2月26日受付)

岩下武史（正会員）■ iwashita@iic.hokudai.ac.jp
1998年京都大学大学院工学研究科博士後期課程修了．博士（工学）．京都大学大型計算機センター，同大学学術情報メディアセンターを経て，2014年より北海道大学情報基盤センター教授．