「数値計算の基礎１」講義ノート
明治大学　2014 年度秋学期　水 4 限

明治大学理工学部数学科　宮部賢志

2015 年 1 月 15 日

前書き

　この講義ノートは，各回の授業のうちで最低限理解して欲しいことをまとめたものです．説明が不十分な所やユーモアが足りないところがあると思います．授業の復習などに利用して下さい．

1　第 1 回　不正確な計算と実数の表現 (9 月 24 日)

1.1　オリエンテーション

この授業は明治大学理工学部数学科 2 年生向けの数値計算の講義です．休講情報は Oh-o!Meiji を参照して下さい．

毎回の授業の進め方は以下のとおり．最初に講義を 60 分程度行います．その後，レポート課題を出しますので，残りの時間を使って他の履修者および TA とその課題について議論して下さい．終わらなければ宿題として課題を完成させて下さい．次回の授業の初めに提出し，TA がその授業中に採点し，授業終了時までに返却します．

成績は期末試験 70%，レポート 30% で評価します．レポートとは毎回の授業で課されるレポートのことです．

情報の知識についてのアンケートを行いました．Oh-o!Mieji を参照して下さい．

1.2　不正確な計算と実数の表現

2008 年 8 月 28 日のニュース記事に「グーグルの電卓機能が計算ミス」というのがあった．(URL: `http://japan.cnet.com/news/media/20379457/`) 授業日の 2014 年 9 月 24 日現在においてもこの問題は修正されていない．これはグーグルはなぜ修正しないのだろうか？この問題を議論するには実数の計算機での表現についての基礎知識を身につける必要がある．

半角英数字は 1byte，すなわち 8bit で表現されるので，$2^8 = 256$ 種類の文字が表現できる．日本語の漢字，ひらがな，カタカナを表現するには不十分で，普通 2byte で 2^{16} 種類を表現できるようにしている．

実数の表現としてよく使われるのが，**浮動小数点表示**である．浮動小数点表示では，実数 x を

$$x = (-1)^b (1.f_1 f_2 \cdots f_k)_\beta \beta^e$$

の形で表現する．ここで，$(-1)^b$ は符号部，$(1.f_1 f_2 \cdots f_k)_\beta$ は仮数部，β^e は指数部と呼ばれる．β は基数であり，普通は $\beta = 2$ とする．e は整数で，$(1.f_1 f_2 \cdots f_k)_\beta$ は β 進法での表現である．

ここでは特に国際規格 IEEE754 の倍精度について説明しよう．実数を 64bit で表現する：

$$b_0 b_1 b_2 \cdots b_{61} b_{62} b_{63}$$

最初の 1bit の b_0 を符号に，次の 11bit の $b_1 b_2 \cdots b_{11}$ を指数 e に，最後の 52bit を仮数部に割り当てる．指数 $1 \le e \le 2^{11} - 2$ となるように取り，

$$x = (-1)^{b_0} (1.b_{12} b_{13} \cdots b_{63})_2 2^{e-1023}$$

によって対応を与える．

この表現の 2 進法での有効数字は約 52 桁なので 10 進法では約 15 桁ほどとなる．最初に挙げたグーグルの例の数字は 16 桁であることに注意しよう．

問題 1.1. 適当に好きな数字を選んで，それを IEEE754 倍精度表示で表現してください．

問題 1.2. Google 電卓の実数の表現を根拠と共に推定してください．

問題 1.3. 前問を踏まえて，Google 電卓が正しくない計算結果を返す式を一つ与え，どうしてそのようなことが起こるのか説明してください．

2　第 2 回　誤差の種類 (10 月 1 日)

2.1　Excel における計算誤差

「普通の生活で 15 桁の数を扱うことはめったにないし，プログラマーになる予定もないので，このような計算誤差については考える必要はない」と思う人もいるかもしれない．しかし，実数の表現についての理解がなければ，次の例を理解するのは難しいだろう．

Excel において，「57% と 52% の差は 5% 以上か？」と Excel に聞くと NO という答えが返ってくる．「57 と 52 の差は 5 以上か」と聞くと YES という答えが返ってくる．これはどうしてだろうか？ % というのは 100 で割ることを意味する．3% と書いた時には，0.03 を意味する．0.57 と 0.52 が浮動小数点表示されることを考えれば，この現象が起こる理由が理解できるだろう．if の条件判断によって実数を扱う場合は気をつけなければならない．

2.2　桁落ち

10 進で有効数字が 7 桁の場合を考えよう．$x = 7.654321$ と $y = 7.654312$ の差は，$x - y = 0.000009$ となり有効数字が 1 桁になる．このような現象を**桁落ち**と言う．

この桁落ちはアルゴリズムを工夫することで防げる場合がある．

例 2.1.

$$x(\sqrt{x^2+1} - x)$$

という式を考える．有効数字 7 桁で $x = 10^4$ の場合は，$x^2 = 10^8$ なので，$+1$ は無視され，答えは 0 となる．ところがこの式を，

$$x(\sqrt{x^2+1} - x) = \frac{x(\sqrt{x^2+1} - x)(\sqrt{x^2+1} + x)}{\sqrt{x^2+1} + x} = \frac{x}{\sqrt{x^2+1} + x}$$

と変形してから計算を行えば，情報落ちが起こったとしても，正しい値に近い答えが出る．

例 2.2. 2 次方程式 $ax^2 + bx + c = 0$ の実数解を求めることを考える．解の公式

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

をそのまま使うと，b の絶対値が a, c の絶対値に比べて大きい時にけた落ちが起こる．そこで，まず桁落ちが起こらない方の解を求めてから，もう一つの解は，解と係数の関係を使って求めると良い．

2.3　情報落ち

例 2.3. 極端な場合であるが，有効数字 2 桁の場合を考えよう．$a = 5.2$ と $b = 5.7$ の平均を計算する．まず，$a + b = 5.2 + 5.7 = 10.9$ であるが，有効数字 2 桁なので，丸められて 10 となる．これを 2 で割ると平均は 5.0 となり，a, b のどちらよりも小さくなる．

このような場合は，$\frac{a+b}{2} = a + \frac{b-a}{2}$ と変形してから計算すると良い．$b - a = 0.5$ であり，$\frac{b-a}{2} = 0.25$ である．よって平均は $5.2 + 0.25 = 5.45$ から 5.4 と計算される．

一般に大きな数と小さな数を足すときには小さな数が無視されることがある．例えば, 有効数字 7 桁で $10^8 + 1 = 10^8$ となる．このような現象を**情報落ち**と呼ぶ.

例 2.4. n 個のデータの分散は，平均を m とすると，

$$v = \frac{\sum_{k=1}^{n}(x_k - m)^2}{n} = \frac{\sum_{k=1}^{n} x_k^2}{n} - m^2$$

であることを知っている．

確率の計算をする場合，後者の式を使うほうが計算量が少ないことが多い．しかし，後者の式は情報落ちが起こりやすい．そのため，数値計算では前者の定義式を使うことが多い．

問題 2.5. 2 次方程式の解の計算において，

(1)「桁落ちが起こらない方の解」はどうやって判定したら良いだろうか？
(2) 解と係数の関係には，和と積の 2 つあるが，どちらを使うのが良いだろうか？

問題 2.6. 指数関数 e^x を計算するときには，テイラー展開

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$$

を利用する．計算機では無限級数の計算はできないため，ある項までで打ち切ることになる．

(1) 第何項まで計算すれば良いかは，実数 x の大きさや実数の表現とはどういう関係にあるだろうか？
(2) 和を計算するときに，左から足していく場合と右から足していく場合で違いがあるだろうか？あるとすればどちらの方法で計算するのが良いだろうか？

3　第 3 回　誤差の評価と計算ステップ数の評価 (10 月 8 日)

桁落ちの復習から始める．sin の和積の公式

$$\sin x - \sin y = 2\cos\frac{x+y}{2}\sin\frac{x-y}{2}$$

において，$x = 1 + 2^{-45}$, $y = 1$ とする．Google 電卓で計算すると，左辺は $1.5321078 \times 10^{-14}$ であり，右辺は $1.5356315 \times 10^{-14}$ 正確な値は Mathematica によると，$1.535631514196\cdots \times 10^{-14}$. なぜこのようなことが起こるのか？ $\sin x$ も $\sin y$ も共に約 $0.8 = 8.0 \times 10^{-1}$ くらいの数である．それに対し，$\sin x - \sin y$ の差は 10^{-14} 程度であるから，13 桁のずれがある．有効数字が 16 桁だとすれば，最終の有効数字は 3 桁程度になる．これが桁落ちである．このように誤差は正しい計算ができなくても見積もることができる．

3.1　誤差の評価

真の値 x, 近似値 x' に対し，$\Delta x = x' - x$ を誤差という．

$$|\Delta x|$$

を絶対誤差と呼び,

$$\frac{|\Delta x|}{|x|}$$

を相対誤差と呼ぶ．10 進有効数字は

$$-\log_{10}\frac{|\Delta x|}{|x|},\ (x \neq 0)$$

と書ける．

倍精度表示の場合，52 桁プラス，近い方に丸めることにより，$2^{-53} \approx 1.1 \times 10^{-16}$ の相対誤差が起こる．$\pi = 3.141592\cdots$ を倍精度表示した場合の絶対誤差は,

$$1.1 \times 10^{-16} \times 3.1415 < 3.5 \times 2^{-16}$$

と見積もることができる．

様々な演算に関して誤差がどのように蓄積されるか見てみよう．まず，加減算の場合,

$$z = x \pm y$$

に対して,

$$z' = x' \pm y' = (x \pm y) + (\Delta x \pm \Delta y)$$

なので,

$$\Delta z = z' - z = \Delta x \pm \Delta y$$

これより，絶対誤差に関しては安定していることが分かる：

$$|\Delta z| \leq |\Delta x| + |\Delta y|$$

しかし，相対誤差は大きくなることがある：

$$\frac{|\Delta z|}{|z|} \leq \left|\frac{x}{z}\right|\frac{|\Delta x|}{|x|} + \left|\frac{y}{z}\right|\frac{|\Delta y|}{|y|}$$

次に乗算の場合,

$$z' = x' \cdot y' = (x + \Delta x)(y + \Delta y) = xy + y\Delta x + x\Delta y$$

より，相対誤差に関しては安定している：

$$\frac{|\Delta z|}{|z|} \leq \frac{|\Delta x|}{|x|} + \frac{|\Delta y|}{|y|}$$

一方，絶対誤差は次のように見積もることができる：

$$|\Delta z| = |y||\Delta x| + |x||\Delta y|$$

一般に n 変数関数 $y = f(x_1, \cdots, x_n)$ に対しては，

$$\Delta y = \sum_{k=1}^{n} \frac{\partial f(x)}{\partial x_i} \Delta x_i$$

で誤差を見積もることができる．

3.2 計算ステップ数の評価

x^{33} という式を考えてみよう．

$$x^{33} = (\cdots(x \times x) \times x) \cdots x) \times x$$

のように計算すると，32 回の乗算が必要になる．ところが，$x^2 = x \times x$, $x^4 = x^2 \times x^2$, $x^8 = x^4 \times x^4$, $x^{16} = x^8 \times x^8$, $x^{32} = x^{16} \times x^{16}$, $x^{33} = x^{32} \times x$ とすれば，6 回の乗算で済む．このようにアルゴリズムを工夫することで，計算ステップ数を減らし，実効時間を短くすることができる場合がある．

$$S = 1 + \frac{1}{1!} + \frac{1}{2!} + \cdots + \frac{1}{n!}$$

の計算を考えてみよう．そのまま計算すれば，k 項目は $k-1$ 回の乗算と 1 回の除算，そしてそれらを合計するための n 回の加算が必要となる．ところが，$k+1$ 項目は k 項目を k で割れば良いので，n 回の除算と n 回の加算で十分である．

問題 3.1. 桁落ちという現象について，相対誤差という観点から説明してください．

問題 3.2. 底辺が $a \times a$ の正方形，側面の勾配が θ の四角錐の体積は $V = \frac{a^3}{6} \tan\theta$ である．V の相対誤差を 2^{-n} に抑えたいとすれば，a, θ の相対誤差はどの程度にする必要があるか．

問題 3.3. 多項式

$$P(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$$

の計算を考える．n 回の乗算と n 回の加算で計算するようなアルゴリズムを考えてみよう．

4　第 4 回　行列の四則演算 (10 月 15 日)

4.1　様々なソフトウェア

数学系の様々なソフトウェアがあるが，大雑把に数式処理，数値解析，統計処理と分かれる．数式処理のソフトウェアとしては，Mathematica, Maple, Maxima など．数値解析では，Matlab, Octave など．統計処理では，S-Plus, R など．目的に応じて使い分けることになる．C 言語で行列計算をするプログラムを書くよりも，インターネットを探して行列計算のライブラリを探してくるほうが早いし，目的に応じて言語を選んだほうが早い．しかし，ここでは行列の数値計算法を学ぶ目的で，C 言語のような言語で連立方程式を解くプログラムを書くことを想定して話をする．

4.2　行列の計算

「一次連立方程式を解く」という問題を考えよう．解き方の本質的な部分は中学校でも習った．しかし，そのアルゴリズムを統一的に整理し，その誤差の解析を行うには，行列表現を考えたほうがすっきりする．

n 次元ベクトル

$$\boldsymbol{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

は，計算機の中では，配列 (数の組) として表現される．同様に (n, n) 型の行列

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

は，n^2 個の数の組として表現される．行列の和，差，積は，定義されるような型であることを確認した後，成分ごとに計算すればよい．

2 つの n 次元実ベクトルの和は，n 回の実数の加算の計算が必要である．$(1, n)$ 型の実行列と n 次元実ベクトルの積は，n 回の乗算と $n-1$ 回の加算の計算が必要であるから，(n, n) 型の実行列と n 次元実ベクトルの積は，n^2 回の乗算と $n(n-1)$ 回の加算の計算が必要である．

計算機では常に誤差が起こる．行列計算でどのような誤差を起こるのかを見よう．ベクトル $\boldsymbol{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ のノルムを

$$||b||_2 = \sqrt{b_1^2 + \cdots + b_n^2}$$

で定義する．$\boldsymbol{b}$ の近似値 $\boldsymbol{b}' = \boldsymbol{b} + \Delta b$ を考える．単純のため，行列 A に誤差はないとしよう．この時，ベクトル $\boldsymbol{c} = A\boldsymbol{b}$ とその近似値 $\boldsymbol{c}' = A\boldsymbol{b}'$ の誤差 $\Delta\boldsymbol{c}$ は，

$$\Delta\boldsymbol{c} = A\boldsymbol{b}' - A\boldsymbol{b} = A(\boldsymbol{b} + \Delta b) - A\boldsymbol{b} = A\Delta\boldsymbol{b}$$

と評価できる．ノルムをとって，

$$||\Delta\boldsymbol{c}||_2 = ||A\Delta\boldsymbol{b}||_2$$

である．この時，誤差 $||\Delta b||_2$ が小さければ，$||\Delta\boldsymbol{c}||_2$ が小さいと言えるだろうか？

定義 4.1. 行列 A のノルムを，

$$||A||_2 = \sup_{\boldsymbol{x} \neq 0} \frac{||A\boldsymbol{x}||_2}{||\boldsymbol{x}||_2}$$

によって定義する．以下では $||\cdot||$ とも書く．

この時，$||\Delta \boldsymbol{c}||_2 \leq ||A||_2||\Delta \boldsymbol{x}||_2$ となるので，$||A||_2$ は誤差の拡大率の上限を表していることが分かる．
A の固有値を λ_i, その固有ベクトルを x_i とすると，

$$||Ax_i|| = ||\lambda_i x_i|| = |\lambda_i|||x_i||$$

より，

$$||A|| \geq \max_i ||\lambda_i||$$

この右辺を $\rho(A)$ と書き，A のスペクトル半径という．

定理 4.2. $B = A^*A$ とすると，

$$||A||_2 = \sqrt{\rho(B)}$$

である．特に A がエルミート行列ならば，

$$||A||_2 = \rho(A)$$

証明. B はエルミート行列なので，対角化するユニタリ行列 P をとる:

$$D = P^*BP = \begin{pmatrix} \mu_1 & & 0 \\ & \mu_2 & \\ & & \ddots \\ 0 & & & \mu_n \end{pmatrix}$$

$x = Py$ とおくと，$x^* = y^*P^*$ であり，

$$||Ax||_2 = (Ax)^*(Ax) = x^*A^*Ax = y^*P^*BPy = y^*Dy = \sqrt{\sum_i \mu_i |y_i|^2}$$

同様にして，

$$||x||_2 = \sqrt{\sum_i \mu |y_i|^2}$$

μ_i のうち μ_M が最大とすれば，y として第 M 成分のみが 1 で他は 0 である単位ベクトルをとれば，

$$||A||_2 = \sqrt{\mu_M} = \sqrt{\rho(A)}$$

A がエルミート行列であれば，A の固有値 λ_i の固有ベクトルを z_i として，

$$Bz_i = A^*Az_i = A(\lambda_i z_i) = \lambda^2 z_i$$

より，$\lambda^2 = \mu_i$ となる．これより，$||A||_2 = \max_i |\lambda_i| = \rho(A)$. □

定義 4.3. A の条件数 $\kappa(A)$ を

$$\kappa(A) = ||A|| \cdot ||A^{-1}||$$

で定義する．

定理 4.4. 連立方程式 $A\boldsymbol{x} = \boldsymbol{b}$ において，誤差 $\Delta \boldsymbol{b}$ による $\boldsymbol{x}$ の誤差 $\Delta \boldsymbol{x}$ は，

$$\frac{||\Delta \boldsymbol{x}||_2}{||\boldsymbol{x}||_2} \leq \kappa(A)||\frac{||\Delta \boldsymbol{b}||_2}{||\boldsymbol{b}||_2}$$

と評価できる．

証明. $\Delta \boldsymbol{x} = A^{-1}\Delta \boldsymbol{b}$ より，

$$||\Delta \boldsymbol{x}|| \leq ||A^{-1}||||\Delta \boldsymbol{b}||$$

これに $||\boldsymbol{b}|| \leq ||A||||\boldsymbol{x}||$ を乗じて得る． □

条件数が大きい行列は，数値計算で良い結果が出ないことがある．

問題 4.5. $||\cdot||_2$ が以下のノルムの性質を満たしていることを示せ.

(1) (独立性) $||A||_2 = 0 \iff A = O$
(2) (斉次性) $||aA||_2 = |a| \cdot ||A||_2$
(3) (劣加法性) $||A + B||_2 \leq ||A||_2 + ||B||_2$

また,

$$||AB||_2 \leq ||A||_2||B||_2$$

となることを示せ.

問題 4.6. (1) 条件数が大きな行列とはどんな行列だろうか．例として対角行列の場合で考えてみよう．
(2) (i, j) の要素が $\dfrac{1}{i+j-1}$ で定義されるヒルベルト行列は条件数が大きくなる行列として有名である．この行列のどんな特徴が，条件数を大きくさせるのか，考えてみよう．

5　第 5 回　ガウスの消去法 (10 月 22 日)

5.1　アルゴリズム

アルゴリズムとは，問題を解くための手順である．例えば，足し算には筆算やそろばんなどのアルゴリズムがある．まず「アルゴリズムを覚えて問題を解けるようになる」ことが重要である．「数値計算を計算機に行わせる」ためには，アルゴリズムを教えてやらなければならない．また，様々な問題に対処するためには，アルゴリズムが修正できるようにならなければならず，なぜそのアルゴリズムで問題が解けるのかを理解する必要がある．すなわち「アルゴリズムを作る」必要があり，メタな視点が必要となる．

5.2　ガウスの消去法

ガウスの消去法について復習しておこう．以下の連立一次方程式について考える．

$$\begin{aligned} 6x + 5y + 4z &= 8 \\ 12x + 13y + 10z &= 16 \\ 18x + 21y + 17z &= 27 \end{aligned}$$

この方程式を拡大行列を使って次のように表現する．

$$\left(\begin{array}{ccc|c} 6 & 5 & 4 & 8 \\ 12 & 13 & 10 & 16 \\ 18 & 21 & 17 & 27 \end{array}\right)$$

この行列に基本変形を施してまず

$$\left(\begin{array}{ccc|c} * & * & * & * \\ 0 & * & * & * \\ 0 & 0 & * & * \end{array}\right)$$

の形にすることを**前進消去**という．対角成分を 1 にする必要はない．この形にしてから，x_3, x_2, x_1 と順に代入して値を求めることを**後退代入**という．詳しくは適当な線形代数の教科書を参照して欲しい．

さてこの方法を計算機にやらせるにはどうすればよいだろうか．行列 A とベクトル b が変数として与えられている時に，基本変形で行列を「書き換える」という操作は，計算機において「変数に値を代入して書き換える」という操作に相当する．

$$\left(\begin{array}{ccc|c} a_{11} & a_{12} & a_{13} & b_1 \\ a_{21} & a_{22} & a_{23} & b_2 \\ a_{31} & a_{32} & a_{33} & b_3 \end{array}\right)$$

と変数が与えられている状況で，a_{21} の部分が 0 となるように基本変形をするためには，

$$\begin{aligned} a_{21} :=& 0 \quad (= a_{21} - a_{11} \cdot \frac{a_{21}}{a_{11}}) \\ a_{22} :=& a_{22} - a_{12} \cdot \frac{a_{21}}{a_{11}} \\ a_{23} :=& a_{23} - a_{13} \cdot \frac{a_{21}}{a_{11}} \\ b_2 :=& b_2 - b_1 \cdot \frac{a_{21}}{a_{11}} \end{aligned}$$

と変数を書き換えればよい．これをまとめて，3 を n に置き換えて，

$$\begin{aligned} &i = 1 \cdots n \\ &\quad a_{2j} := a_{2j} - a_{1j} \cdot \frac{a_{21}}{a_{11}} \\ &\quad b_2 := b_2 - b_1 \cdot \frac{a_{21}}{a_{11}} \end{aligned}$$

この操作を $i = 2 \cdots n$ 行に対して行い，さらにその操作を $k = 1 \cdots k$ 列に対して行う．まとめるとアルゴリズムは次のように書ける．

$$
\begin{aligned}
k &= 1 \cdots n-1 \\
\quad i &= k+1 \cdots n \\
\quad\quad j &= k+1 \cdots n \\
\quad\quad\quad \alpha &= \frac{a_{ik}}{a_{kk}} \\
\quad\quad\quad a_{ij} &= a_{ij} - \alpha \cdot a_{kj} \\
\quad\quad\quad b_i &= b_i - \alpha \cdot b_k
\end{aligned}
$$

その後，後退代入としては，

$$
\begin{aligned}
x_n &:= \frac{b_n}{a_{nn}} \\
k &= n-1 \cdots 1 \\
\quad x_k &= \frac{b_k - \sum_{j=k+1}^{n} a_{kj} x_j}{a_{kk}}
\end{aligned}
$$

とすればよい．

5.3 部分ピボット選択

前節のガウスの消去法には問題がある．計算の途中で除算が出てくるためである．この割る数が 0 であれば，計算機はエラーを返す．0 ではなかったとしても，非常に小さい数であれば，誤差が大きくなる．そのような例を挙げよう．

有効数字 4 桁で，

$$
\begin{aligned}
-0.001x_1 + 6x_2 &= 6.001 \\
3x_1 + 5x_2 &= 2
\end{aligned}
$$

の方程式を解いてみよう．もちろん正確な解は $(x_1, x_2) = (-1, 1)$ である．第 2 式の x_1 を消去すると，$\alpha = \frac{3}{-0.001} = -3000$ より，

$$
\begin{aligned}
&5 - 6 \cdot (-3000) = 1.801 \times 10^4, \\
&2 - 6.0001 \cdot (-3000) = 1.800 \times 10^4, \\
&x_2 = 9.994 \times 10^{-1}, \\
&x_1 = \frac{6.001 - 6.000 \times 9.994 \times 10^{-1}}{-1.000 \times 10^{-3}} = \frac{6.001 - 5.996}{-1.000 \times 10^{-3}} = -5.000
\end{aligned}
$$

そこでこの問題を解決するため，k 列での前進消去の際に，k 列の k 行目以降で**絶対値の最も大きな要素**がある行と k 行目を入れ替える．これを**部分ピボット選択**と言う．ピボットとは軸のこと．部分と言われているのは，探す要素が行列の一部分だからである．

問題 5.1. ガウスの消去法の乗除算，加減算の回数を数えてみよう．

問題 5.2. 部分ピボット選択付きのガウスの消去法を使って，

$$
\begin{aligned}
6x_1 + 5x_2 + 4x_3 &= 8 \\
12x_1 + 13x_2 + 10x_3 &= 16 \\
18x_1 + 21x_2 + 17x_3 &= 27
\end{aligned}
$$

の解を求めよ．

6　第 6 回　LU 分解 (10 月 29 日)

$\boldsymbol{b}$ を何度も変えて計算する場合，同じような計算をすることになる．そのため，逆行列を求めたくなる．しかし，数値計算では逆行列はほとんど使われない．それよりも LU 分解を使う方が効率的であるためである．

6.1　LU 分解

行列 A の LU 分解とは，下三角行列 L と上三角行列 U に対して，

$$A = LU$$

となるようにすることである．このように分解できれば，

$$L(Ux) = b$$

は，後退代入で求まる．

ガウスの消去法において，ピボット選択がない場合，LU 分解は簡単に求まる．$Ax = b$ をガウスの消去法で解く場合，左から下三角行列 L_i をかけて，

$$L_{n-1}L_{n-2}\cdots L_1 Ax = L_{n-1}L_{n-2}\cdots L_1 b$$

として，

$$U = L_{n-1}L_{n-2}\cdots L_1 A$$

が上三角行列となるようにする．よって，

$$L = L_1^{-1}L_2^{-1}\cdots L_{n-1}^{-1}$$

とすれば，

$$A = LU$$

と LU 分解できる．

6.2　ピボット選択がある場合

ピボット選択がある場合には，置換行列 P_i と変換行列 G_i を使って，

$$Ux = G_{n-1}P_{n-1}G_{n-2}P_{n-2}\cdots G_1 P_1 b$$

と書ける．ここで U は上三角行列である．この右辺を

$$G_{n-1}P_{n-1}G_{n-2}P_{n-2}\cdots G_1 P_2^{-1}P_3^{-1}\cdots P_n^{-1}Pb$$

と見よう．ここで $P = \prod_{k=1}^{n-1} P_k$ である．これは，漸化式で，

$$H_1 = G_1,\ H_i = G_i P_i H_{i-1} P_i^{-1},\ L^{-1} = H_{n-1}$$

とすれば，各 H_i は下三角行列となり，

$$LUx = Pb$$

という形で LU 分解ができる．

7　第 7 回　2 分法とニュートン法 (11 月 5 日)

これまでは，連立一次方程式の解法を学んできた．これからしばらくは非線形方程式の解法について学ぶ．

7.1　2 分法

$\sqrt{2}$ の近似値を求めよう．通常は次のような方法を使うのではないだろうか．$1^2 = 1 < 2$, $2^2 = 4 > 2$ より，$1 < \sqrt{2} < 2$. $1.4^2 = 1.96 < 2$, $1.5^2 = 2.25 > 2$ より，$1.4 < \sqrt{2} < 1.5$. これを繰り返して $\sqrt{2}$ の近似値を求める．

この方法は一体何をしているのか，分析してみよう．$f(x) = x^2 - 2$ とおくと，$\sqrt{2}$ は $f(x) = 0$ の唯一の正の解である．$f(1) = -1 < 0$, $f(2) = 2 > 0$ であり，$f(x)$ は連続であるから，求める解は $(1, 2)$ の中にある．次に，$(1, 2)$ を 10 等分し，$f(1.4) < 0$, $f(1.5) > 0$ であるから，求める解は $(1.4, 1.5)$ の中にある．

この手法を用いて一般の非線形方程式を解いてみよう．連続な実関数 $f(x)$ に対して，$f(a) < 0$, $f(b) > 0$ であるとする．この時，中間値の定理から，$f(x) = 0$ の解が (a, b) の中に少なくとも 1 つ存在する．計算機に行わせるには 10 等分よりも 2 等分の方が都合が良い．$c = \frac{a+b}{2}$ とする．もし $f(c) = 0$ であれば，c が求める解．実際には誤差があるため，真に 0 に等しくなることはめったにない．$f(c) < 0$ であれば，a を c で置き換え，$f(c) > 0$ であれば，b を c で置き換えて，同じことを繰り返す．n 回繰り返すと，a, b の距離は 2^{-n} 倍される．その距離が十分小さくなったところで打ち切って a または b を解とする．これが *2* **分法**と呼ばれるアルゴリズムである．

例 7.1. 2 分法により $\sqrt{2}$ を求める R 言語でのアルゴリズム．

```
f <- function (x) { x^2 -2 }
e <- 10^(-14)
a <- 1
b <- 2
while (b-a > e){
  c <- (a+b)/2
  if (f(c)<0){ a <- c} else { b<- c}
}
a
```

問題 7.2. 2 分法において，解が複数ある場合，どうなるだろうか？ 2 分法で何回繰り返せば良いかは予め分かるだろうか？

7.2　ニュートン法

$\sqrt{2}$ は $f(x) = x^2 - 2$ の唯一の正の解であった．関数 $f(x)$ が微分できる場合には，次のような方法で計算すると収束が速いことがある．

初期値 x_0 から始める．$(x_n, f(x_n))$ での接線は $y = f'(x_n)(x - x_n) + f(x_n)$ であり，この接線と $y = 0$ との交点の x 座標を x_{n+1} とする．すなわち $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$ である．適当な ϵ に対して，$\frac{|x_{n+1} - x_n|}{|x_n|} < \epsilon$ を満たせば，（すなわち x_n がほとんど動かなくなったら）終了する．（収束判定条件については来週詳しく学ぶ.）

$f(x) = x^2 - 2$ の場合，$f'(x) = 2x$ であるから，漸化式は

$$x_{n+1} = x_n - \frac{x^2 - 2}{2x}$$

であるから，$x_0 = 2$ から始めると，1.5, 1.416667, 1.414216, 1.414214 と求まる．

例 7.3. ニュートン法により $\sqrt{2}$ を求める R 言語でのアルゴリズム

```
f <- function (x) { x - (x^2-2)/(2*x)}
```

```
e <- 10^(-15)
a <- 2
b <- 3
n <- 1
while(abs((a-b)/b)>=e){
  b <- a
  a <- f(b)
  n <- n + 1
}
a
```

ニュートン法の長所は収束が速いことである．ニュートン法で $f(x)=0$ の解 α に収束したとする．$g(x)=\frac{f(x)}{f'(x)}$ とおいて，$g'(x)=1-\frac{f(x)f''(x)}{(f'(x))^2}$ より，$g(\alpha)=0, g'(\alpha)=1$. $g(x)$ を $x=\alpha$ でテーラー展開して，

$$g(x)=(x-\alpha)+O((x-\alpha)^2)$$

これより，

$$x_{n+1}-\alpha=x_n-g(x_n)-\alpha=O((x-\alpha)^2)$$

すなわち 2 次の収束をする．2 分法は 1 次の収束であることに注意しよう．

ニュートン法の短所は初期値を間違えると収束しないことがある．例えば，$f(x)=x^3-3x^2+x+3$ で，$x_0=1$ を初期値としてみよ．

7.3　2 変数のニュートン法

$f(x,y)=0, g(x,y)=0$ の連立方程式を考える．$f(x,y)$ の (a,b) での接平面は，

$$z-f(a,b)=f_x(a,b)(x-a)+f_y(a,b)(y-b)$$

であり，$g(x,y)$ の方は，

$$z-g(a,b)=g_x(a,b)(x-a)+g_y(a,b)(y-b)$$

である．平面 $z=0$ と連立した解を次の点とする．

$y=x^2$ と $y=x+2$ の交点をニュートン法で求めよう．上記の方法で漸化式を求めると，

$$x_{n+1}=x_n+\frac{-x_n^2+x_n+2}{2x_n-1},\quad y_{n+1}=y_n+\frac{-x_n^2+x_n+2}{2x_n-1}+x-y+2$$

となる．$(x_0,y_0)=(0,0)$ から始めると，$(-2,0), (-1.2,0.8), (-1.011765,0.9882353), (-1.000046,0.9999542), (-1,1)$ と求まる．$(x_0,y_0)=(3,5)$ から始めると，$(2.2,4.2), (2.011765,4.011765), (2.000046,4.000046), (2,4)$ と求まる．

例 7.4. ニュートン法により $y=x^2$ と $y=x+2$ の交点を求めるための R 言語でのプログラム

```
f <- function (z){
  x <- z[1]
  y <- z[2]
  t <- (-x^2+x+2)/(2*x-1)
  return(c(x+t,t+x+2))
}
norm <- function (z){  max(abs(z[1]),abs(z[2])) }
e <- 10^(-15)
a <- c(0,0)
```

```
b <- c(1,1)
n <- 0
while(norm(a-b)/norm(b)>=e){
  b <- a
  a <- f(b)
  n <- n + 1
}
a
```

問題 7.5. $y = x^2$ と $y = x + 2$ の交点をニュートン法で求めるための漸化式を求めよ.

8　第 8 回　反復法 (11 月 26 日)

漸化式を作ってより良い近似値を求めていく方法を，反復法もしくは反復計算と呼ぶ．反復法において重要なのが「いつ反復計算をやめるべきか」という収束判定基準の設定である．

8.1　様々な収束判定基準

定義 8.1. 実数の数列 $\{x_n\}$ が α に収束するとは，任意の $\epsilon > 0$ に対して，ある自然数 N が定まり，$n > N$ ならば $|x_n - \alpha| < \epsilon$ を満たすことである．

$\{x_n\}$ が与えられても α を計算することはできない．そこで，何らかの方法で，収束したと見なす必要がある．

もともとの問題は「$f(x) = 0$ を満たす x を計算機を用いた反復計算により求める」ということだった．そこで，$f(x)$ が小さいかどうかで判定するという方法が考えられる．これを**残差基準の判定**という．

定義 8.2. 収束判定用の定数 $\epsilon > 0$ を決めておき，

$$|f(x_n)| < \epsilon$$

ならば収束したとみなして計算を打ち切る．

x_n の変化が小さいかどうかで判定する方法を**誤差基準の判定**という．

定義 8.3. 収束判定用の定数 $\epsilon > 0$ を決めておき，

$$|x_n - x_{n-1}| < \epsilon$$

ならば収束したとみなして計算を打ち切る．もしくは，相対的な修正量

$$\frac{|x_n - x_{n-1}|}{|x_{n-1}|} < \epsilon$$

を使うこともある．

この場合，非常に近い数の引き算を行うため，桁落ちが起こりやすい．この判定部分だけ精度をあげて桁落ちを回避するという方法がよく用いられる．

このうちのどれを使うべきか．唯一の正解はない．問題による．

ϵ の選び方にも注意が必要である．「ϵ を小さくすれば良い収束が得られる」わけではない．小さくし過ぎると収束しないこともある．マシンイプシロンとは

$$1 + \epsilon > 1$$

を満たす最小の浮動小数点数である．IEEE754 の倍精度の場合 2^{-52} である．一般にこれより小さい数にするのは危険である．

問題 8.4. 収束判定を誤る例を考えてみよう．その場合にはアルゴリズムや収束判定にどのような修正を加えたら良いだろうか．

8.2　様々なノルム

これまでは 1 変数の場合を考えてきた．多変数の場合には，差の代わりにノルムを考える．

以前にベクトル $\boldsymbol{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ に対して，

$$||\boldsymbol{x}||_2 = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

という *2* **ノルム**を紹介した．これはユークリッドノルムや l_2 ノルムとも呼ばれ，数学で様々な場面で表れる他，様々な良い性質を持つ．しかし，2 乗やルートなど重い計算が必要となる．0 に近いかどうかを調べるなどの特定の目的のためには，わざわざこのノルムを計算する必要はない．

他によく使われるノルムとして，

$$||\boldsymbol{x}||_1 = \sum_{i=1}^{n} |x_i|$$

という 1 ノルム，絶対ノルム，l_1 ノルムや，

$$||\boldsymbol{x}||_\infty = \max_{1\leq i\leq n} |x_i|$$

という最大ノルム，l_∞ ノルムがある．

以前述べた行列ノルム

$$||A|| = \sup_{\boldsymbol{x}\neq\boldsymbol{0}} \frac{||A\boldsymbol{x}||}{||x||}$$

はそれぞれのノルムから誘導される．1 ノルムや最大ノルムの場合には，

$$||A||_1 = \max_{1\leq j\leq n} \sum_{i=1}^{n} |a_{ij}|,$$
$$||A||_\infty = \max_{1\leq i\leq n} \sum_{j=1}^{n} |a_{ij}|$$

のように具体的に計算できる．

$n = 2$ の場合に証明してみよう．$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ と $\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \neq \boldsymbol{0}$ に対して，$m = \max_{j=1,2}(|a_{1j}| + |a_{2j}|)$ とおくと，

$$\begin{aligned} ||Ax||_1 =& || \begin{pmatrix} a_{11}x_1 + a_{1n}x_n \\ a_{n1}x_1 + a_{nn}x_n \end{pmatrix} ||_1 \\ =& |a_{11}x_1 + a_{12}x_2| + |a_{21}x_1 + a_{22}x_2| \\ \leq& (|a_{11}| + |a_{21}|)|x_1| + (|a_{12}| + |a_{22}|)|x_2| \\ \leq& m|x_1| + m|x_2| = m||\boldsymbol{x}||_1 \end{aligned}$$

逆に，$\sum_{i=1}^{n} a_{ij}$ が最大となるのが k であった時，$\boldsymbol{x}$ として，k 成分は 1，それ以外の成分は 0 となるベクトルを考えれば，$||\boldsymbol{x}||_1 = 1$, Ax は A の k 列のベクトルなので，$||Ax||_1 = \sum_{i=1}^{n} |a_{ik}|$．これより，$||A||_1 \geq \max_{1\leq j\leq n} \sum_{i=1}^{n} |a_{ij}|$．

また，$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ と $\boldsymbol{x} = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \neq \boldsymbol{0}$ に対して，

$$\begin{aligned} ||Ax||_\infty =& \max_{i=1,2}(|a_{i1}x_1 + a_{i2}x_2|) \\ \leq& \max_{i=1,2}(|a_{i1}| + |a_{i2}|) \max_{k=1,2} |x_k|, \end{aligned}$$

すなわち，$||A||_\infty \leq \max_{i=1,2} \sum_{j=1}^{2} |a_{ij}|$．

逆に，$A = (a_{ij})$ に対して，k 行目で $\sum_{j=1}^{n} |a_{ij}|$ が最大となったとする．$a_{kj}x_j = |a_{kj}|$ となるように x_j を定める．すると $||\boldsymbol{x}||_\infty = 1$ であり，$||Ax||_\infty \geq \sum_{j=1}^{n} a_{kj}x_j = \sum_{j=1}^{n} |a_{kj}|$．

問題 8.5. 3 つのノルムの間には以下の不等式が成り立つことを示せ．

$$||\boldsymbol{x}||_2 \leq ||\boldsymbol{x}||_1 \leq \sqrt{n}||\boldsymbol{x}||_2$$
$$||\boldsymbol{x}||_\infty \leq ||\boldsymbol{x}||_2 \leq \sqrt{n}||\boldsymbol{x}||_\infty$$
$$||\boldsymbol{x}||_\infty \leq ||\boldsymbol{x}||_1 \leq n||\boldsymbol{x}||_\infty$$

9　第 9 回　反復法と縮小写像の原理 (12 月 3 日)

反復法は $x_{n+1} = f(x_n)$ によってより良い近似値を求めていく方法である．どんな時に収束すると言えるだろうか．

定義 9.1. 集合 $X \subset \mathbb{R}^n$ と写像 $f : X \to X$ を考える．f がノルム $||\cdot||$ に対して**縮小写像**であるとは，ある $L \in [0, 1)$ が存在して，すべての $x, y \in X$ に関して，

$$||f(x) - f(y)|| \leq L||x - y||$$

を満たすことを言う．

任意の縮小写像はリプシッツ連続である．特に連続である．

定理 9.2 (縮小写像の原理)**.** $X \subset \mathbb{R}^n$ を閉集合，$f : X \to X$ を縮小写像とする．

(1) f は X の中に唯一の不動点 *($f(x') = x'$ を満たす x')* を持つ．
(2) 任意の $x \in X$ に対して，$\lim_{k\to\infty} f^k(x) = x'$.

証明. $x_k = f^k(x)$ とすると，$\{x_k\}$ はコーシー列である．$\mathbb{R}^n$ の完備性と，X が閉集合であることから，$\{x_k\}$ の極限 x' は $x' \in X$. f は連続であるから $f(x') = \lim_{k\to\infty} f(x_k) = \lim_{k\to\infty} x_{k+1} = x'$. x'' も不動点であるとすると，$||x' - x''|| = ||f(x') - f(x'')|| \leq L||x' - x''||$. $0 \leq L < 1$ であることから，$||x' - x''|| = 0$，すなわち $x' = x''$. 任意の $x \in X$ に関して $x_k = f^k(x)$ で定まる $\{x_k\}$ が唯一の不動点に収束するから，$\lim_{k\to\infty} f(x) = x'$. □

問題 9.3. 微分可能な関数 $f : \mathbb{R} \to \mathbb{R}$ に対して，$|f'(x)| < q < 1$ となる q が存在すれば，f は縮小写像であることを示せ.

この方法を使って連立一次方程式を解いてみよう.

$A\boldsymbol{x} = \boldsymbol{b}$ が与えられた時，$A = L + D + U$ と分解する．ここで，D は対角行列，L は下三角行列，U は上三角行列である．すると，

$$\begin{aligned}(L + D + U)\boldsymbol{x} &= \boldsymbol{b}\\ D\boldsymbol{x} &= -(L + U)\boldsymbol{x} + \boldsymbol{b}\\ \boldsymbol{x} &= -D^{-1}(L + U)\boldsymbol{x} + D^{-1}\boldsymbol{b}\end{aligned}$$

が成り立つ．ここで，$M = -D^{-1}(L + U)$, $\boldsymbol{c} = D^{-1}\boldsymbol{b}$ とおいて，

$$\boldsymbol{x_{n+1}} = M\boldsymbol{x_n} + \boldsymbol{c} = f(\boldsymbol{x_n})$$

という漸化式を考えよう．これを**ヤコビ法**という．もし f が縮小写像ならば，$\{x_n\}$ は唯一の不動点，すなわち連立一次方程式の解に収束する．f が縮小写像かどうかは行列 M のノルムによって決まる.

$$||f(\boldsymbol{x}) - f(\boldsymbol{y})|| = ||M(\boldsymbol{x} - \boldsymbol{y})|| \leq ||M|| \cdot ||\boldsymbol{x} - \boldsymbol{y}||.$$

ここでは計算しやすいように，$||M||_\infty < 1$ という条件を考えよう．これが収束のための十分条件である．$A = (a_{ij})$, $M = (m_{ij})$ とすると，

$$m_{ii} = 0, \quad m_{ij} = -\frac{a_{ij}}{a_{ii}}$$

であるから，最大値ノルムは定義により，

$$||M||_\infty = \max_i \sum_{j=1}^{n} |m_{ij}| = \max_i \sum_{j=1, j\neq i}^{n} \frac{|a_{ij}|}{|a_{ii}|}$$

である.

例 9.4. ヤコビ法による $\begin{pmatrix} 3 & 1 & 1 \\ 1 & 4 & 2 \\ 3 & 1 & 5 \end{pmatrix} \boldsymbol{x} = \begin{pmatrix} 10 \\ 20 \\ 30 \end{pmatrix}$ を解く R 言語でのプログラム

```
options(digits=15)
#A <- matrix(c(3,1,1,1,4,2,3,1,5), nrow = 3, ncol = 3, byrow = TRUE)
#b = c(10,20,30)
M <- matrix(c(0,-1/3,-1/3,-1/4,0,-2/4,-3/5,-1/5,0),3,3,byrow = TRUE)
d <- c(10/3,20/4,30/5)

f <- function (x){ M %*% x + d }
norm <- function (z){ max(abs(z[1]),abs(z[2])) }
e <- 10^(-15)
x <- c(0,0,0)
y <- c(1,1,1)
n <- 0
while(norm(x-y)/norm(y)>=e){
  y <- x
  x <- f(x)
  n <- n + 1
}
x
```

問題 9.5. ヤコビ法の反復回数を見積りなさい.

10　第 10 回　数値積分とシンプソンの公式 (12 月 10 日)

10.1　区分求積法

連続関数 $f : \mathbb{R} \to \mathbb{R}$ に対して定積分 $\int_a^b f(x)\ dx$ を計算したい．一般に連続関数の不定積分を求めることは難しく，初等関数で表現できないことも多い．そこで，区分求積法の考え方を使って，近似することを考える．

定積分は以下のように表現できる．

$$\int_a^b f(x)\ dx = \lim_{|\Delta|\to 0} \sum_{i=1}^{n-1} f(\zeta_i)\delta x_i$$

ここで，Δ は区間 $[a,b]$ の分割であり，$a = x_1 \leq x_2 \leq \cdots \leq x_n = b$, $x_i < \zeta_i < x_{i+1}$, $\delta x_i = |x_{i+1} - x_i|$, $|\Delta| = \max_i \delta x_i$. そこで，$|\Delta|$ が十分小さな分割によって，定積分は近似できる．

10.2　台形公式

誤差を小さくするため，各区間の近似を長方形ではなく，台形で近似すると，

$$\int_a^b f(x)\ dx \approx \frac{1}{2}\sum_{i=1}^{n-1}(f(x_i) + f(x_{i+1}))(x_{i+1} - x_i)$$

特に n 等分すると，$x_{i+1} - x_i = \frac{a-b}{n}$ である．

10.3　ラグランジュの補間法

2 点あればそれを通る 1 次式はただ 1 つに決まる．同様に 3 点あればそれを通る 2 次式はただ 1 つに決まる．$(x_0, f_0), (x_1, f_1), (x_2, f_2)$ を通る 2 次式は

$$P(x) = \sum_{i=0}^{2} f_i l_i(x)$$

と書ける．ここで，

$$l_0(x) = \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)},$$
$$l_1(x) = \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)},$$
$$l_2(x) = \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)}$$

である．この多項式は，クロネッカーのデルタを使って，

$$l_i(x_k) = \delta_{ik} = \begin{cases} 1 & (i = k) \\ 0 & (i \neq k) \end{cases}$$

という性質を持っていることに注意しよう．

問題 10.1. $n + 1$ 点を通る n 次多項式を求めるラグランジュの補間法を求めよ．

10.4　シンプソンの公式

区間を $2n$ 分割してそれぞれの 2 区間で 2 次式で近似した積分値 n 個を加える，という方法で数値積分を行うのがシンプソンの公式である．

区間 $[a,b]$ を x_0 から x_{2n} により $2n$ 分割する．各区間の 2 次式での近似は

$$y = \frac{(x-x_{2i-1})(x-x_{2i})}{(x_{2i-2}-x_{2i-1})(x_{2i-2}-x_{2i})} f_{2i-2} + \frac{(x-x_{2i-2})(x-x_{2i})}{(x_{2i-1}-x_{2i-2})(x_{2i-1}-x_{2i})} f_{2i-1} + \frac{(x-x_{2i-2})(x-x_{2i-1})}{(x_{2i}-x_{2i-2})(x_{2i}-x_{2i-1})} f_{2i}$$

ここで $y_k = f(x_k)$ である．これを x_{2i-2} から x_{2i} まで積分すると，

$$\int_{x_{2i-2}}^{x_{2i}} f(x)\ dx \approx \frac{h}{3}(y_{2i-2} + 4y_{2i-1} + y_{2i}).$$

ここで，$h = \frac{b-a}{2n}$．よって，

$$\int_{x_0}^{x_{2i}} f(x)\ dx \approx \frac{h}{3}\left(y_0 + 4\sum_{i=1}^{n} y_{2i-1} + 2\sum_{i=1}^{n} y_{2i} + y_{2n}\right).$$

問題 10.2. $\displaystyle\int_0^1 \frac{4dx}{1+x^2} = \pi$ である．この定積分を台形公式やシンプソンの公式を使って近似することで，π の近似値を求めよ．電卓などを使って良い．

11　第 11 回　ガウス積分 (12 月 17 日)

ガウス積分は $n+1$ 個の分点を上手く選んで，$2n+1$ 次の多項式で近似し，積分値の精度を向上させる方法である．

11.1　エルミート補間

相異なる x_0 から x_n に対して，関数値 f_0 から f_n と，微係数 f_0' から f_n' が一致するような，$2n+1$ 次の多項式を求めよう．これを**エルミート補間**という．

その多項式を

$$p_{2n+1}(x) = \sum_{k=0}^{n} f_k h_k(x) + \sum_{k=0}^{n} f_k' g_k(x)$$

とおくと，

$$p_{2n+1}'(x) = \sum_{k=0}^{n} f_k h_k'(x) + \sum_{k=0}^{n} f_k' g_k'(x)$$

である．更なる条件として，

$$h_k(x_i) = \delta_{ki},\ h_k'(x_i) = 0,\ g_k(x_i) = 0,\ g_k'(x_i) = \delta_{ki}$$

を課すと，

$$p_{2n+1}(x_i) = \delta_{ik} f_k,\ p_{2n+1}'(x_i) = \delta_{ik} f_k'$$

となる．この条件を満たす h_k, g_k はラグランジュの補間式を使って，

$$\begin{aligned} h_k =& \{l_k(x)\}^2 \{1 - 2(x - x_k) l_k'(x_k)\}, \\ g_k =& (x - x_k)\{l_k(x)\}^2 \end{aligned}$$

と書ける．

11.2　ルジャンドル多項式

ルジャンドル多項式は，

$$\begin{aligned} P_0(x) =& 1, \\ P_1(x) =& x, \\ P_n(x) =& \frac{2n-1}{n} x P_{n-1}(x) - \frac{n-1}{n} P_{n-2}(x) \end{aligned}$$

によって定義される．

ルジャンドル多項式は次のような性質を持つ．次数の異なるルジャンドル多項式は相関がない．つまり，異なる i, j に対して，

$$\int_{-1}^{1} P_i(x) P_j(x) dx = 0$$

よって，n 次のルジャンドル多項式 $P_n(x)$ は $n-1$ 次の一般的な多項式 $Q(x)$ と相関がない．

$$\int_{-1}^{1} P_n(x) Q(x) dx = 0$$

ルジャンドル方程式の根は，すべて実数で，-1 と 1 の間にある．

11.3 ガウス積分

まず，変数変換により積分区間を $[-1,1]$ に変更する．

$$\int_a^b f(x)dx = \frac{b-a}{2}\int_{-1}^{1} f\left(\frac{a+b}{2}+\frac{b-a}{2}t\right)dt$$

次に，$f(x)$ をエルミートの補間法により，

$$f(x) \approx \sum_{i=0}^{n} y_i h_i(x) + \sum_{i=0}^{n} y_i' g_i(x)$$

と近似する．すると，

$$\int_{-1}^{1} f(x)dx = \sum_{i=0}^{n} \alpha_i y_i + \sum_{i=0}^{n} \beta_i y_i'$$

ここで，

$$\alpha_i = \int_{-1}^{1} h_i(x)dx = \int_{-1}^{1} \{l_i(x)\}^2\{1-2(x-x_i)l_i'(x_i)\}dx,$$
$$\beta_i = \int_{-1}^{1} g_i(x)dx = \int_{-1}^{1} (x-x_i)\{l_i(x)\}^2 dx.$$

ここで，分点として $n+1$ 次のルジャンドル方程式の根 $n+1$ 個を取ってみよう．すると，$(x-x_i)l_i(x)$ は $n+1$ 次式でその根は，ルジャンドル方程式の $n+1$ 個の根なので，$P_{n+1}(x)$ の定数倍である．よって，$\beta_i = 0$ である．

実際に $\int_0^1 \frac{4dx}{1+x^2} = \pi$ を用いて，円周率を小数点以下 10 桁まで求めるのに必要な分割の数は，台形公式の場合 43082 であるのに対し，シンプソンの公式は $14 \times 2 = 28$ で，ガウス積分の場合には 8 である．

問題 11.1. 以下の式で定義されるルジャンドル多項式 $P_n(x)$ について次のことを示せ．

$$P_0(x) = 1,\ P_1(x) = x,\ P_n(x) = \frac{2n-1}{n}xP_{n-1}(x) - \frac{n-1}{n}P_{n-2}(x).$$

直交性，すなわち，相異なる i, j に対して $\int_{-1}^{1} P_i(x)P_j(x) = 0$ となることは使って良い．

(1) $P_n(x)$ は n 次式である．
(2) 任意の多項式は $\sum_{k=0}^{n} b_k P_k(x)$ の形で表される．
(3) 任意の $n < m$ に対して，$\int_{-1}^{1} P_m(x)x^n dx = 0$ を示せ．
(4) ある多項式が $Q(x) = \sum_{k=0}^{n} b_k P_k(x)$ と表されるならば，$\int_{-1}^{1} Q(x)dx = 2b_0$.

12　第 12 回　区間演算 (1 月 14 日)

$$a = [\underline{a}, \overline{a}],\ b = [\underline{b}, \overline{b}],\ \underline{a} > 0,\ \underline{b} > 0$$

とすると,

$$\begin{aligned} a + b &= [\underline{a} + \underline{b}, \overline{a} + \overline{b}], \\ a - b &= [\underline{a} - \overline{b}, \overline{a} - \underline{b}], \\ a \cdot b &= [\underline{ab}, \overline{a}\overline{b}], \\ a/b &= [\underline{a}/\overline{b}, \overline{a}/\underline{b}]. \end{aligned}$$

この時には丸めモードに気をつける必要がある．$\underline{a} < 0$ の時には，$a \cdot b$ は $\min, \max$ を使って書くと楽.

$x = [-0.2, 0.2]$ のとき,

$$y = (x + 1)(x - 1)(x + 3)$$

の真の解は,

$$y \subseteq [2.688, 3.079201435679]$$

である．通常の区間演算を行うと,

$$y \subseteq [1.792, 4.608]$$

となって精度が悪い．一方，式を

$$y = x^3 - 3x^2 - x + 3$$

と展開してから計算すると，結果は

$$y \subseteq [2.656, 3.224]$$

と精度が良くなる．[*1]

$x = [-1, 2],\ y = [1, 2], z = [-3, 2]$ とすると，$xy + xz = [-8, 6],\ x(y + z) = [-4, 2]$ となる．劣分配則 (subdistributive law) が成立するが，一致しないことがある.

$$x \cdot (y + z) \subseteq x \cdot y + x \cdot z$$

問題 12.1. 区間演算は中心と半径を情報として持つことでも行うことができる．その場合の四則演算を書け．単純のため，乗算・除算の場合には，正の値であることを仮定して良い.

[*1] この例は山本野人先生によるものらしい.

演習問題ヒント

問 *4.6.* 対角行列の場合，対角成分がそのまま固有値となる．固有値の絶対値のうち最大のものを M，最小のものを m とすると，条件数は $\frac{M}{m}$ である．すなわち成分の絶対値の最大と最小の比が大きい行列は条件数が大きい．

(n,n) 型のヒルベルト行列では，n が大きくなると各行がほとんど” 並行” になり，行列式が 0 に近づき，固有値のうちの 1 つが 0 に近づく．そのため条件数が大きくなる． □

問 *5.1.* (n,n) 型の行列に対する前進消去，後退代入についての回数を調べよう．(k,k) 成分 $(k=1\sim n-1)$ をピボットとして，i 行 $(i=k+1\sim n)$ に対して，次の操作を行う．$\alpha=\frac{a_{ik}}{a_{kk}}$ において除算 1 回，$a_{ij}=a_{ij}-\alpha\cdot a_{kj}$ において乗算 1 回，減算 1 回，これが $j=k\sim n$. $b_i=b_i-\alpha b_k$ において乗算 1 回，減算 1 回．よって，乗除算の回数は，

$$\sum_{k=1}^{n-1}\sum_{i=k+1}^{n}(1+(n-k+1)+1)$$

で，加減算の回数は，

$$\sum_{k=1}^{n-1}\sum_{i=k1}^{n}((n-k+1)+1)$$

あとはこれを計算すれば良い．重要なのは，オーダーで，どちらも $\frac{n^3}{3}$ くらいになるはずである．

後退代入についても同様に調べると，$\frac{n^2}{2}$ くらいになる． □

問 *7.2.* 解が複数ある場合は，そのうちのどれか 1 つに収束する．どの解に収束するかは，a,b の初期値に依存する．

最初の区間の幅を d とすれば，n 回繰り返した後にの幅は $2^{-n}d$ となる．収束条件を幅が $\leq e$ としていたならば，これより繰り返しの回数が求まる． □

問 *7.5.* $f(x,y)=x^2-y$, $g(x,y)=x-y+2$ とおく．(a,b) での接平面と $z=0$ を連立させて，

$$\begin{aligned}2a(x-a)-(y-b)&=-(a^2-b)\\(x-a)-(y-b)&=-(a-b+2)\end{aligned}$$

これを x,y について解くと，

$$x=a+\frac{-a^2+a+2}{2a-1},\quad y=a+\frac{-a^2+a+2}{2a-1}+2$$

が導かれる．これより漸化式は，

$$x_{n+1}=x_n+\frac{-x_n^2+x_n+2}{2x_n-1},\quad y_{n+1}=x_n+\frac{-x_n^2+x_n+2}{2x_n-1}+2$$

と求まる． □

問 *8.4.* 多くの x について，$f(x)$ がほとんど 0 に近いような $f(x)$ に対しては，収束判定を誤ることがある．そのため，初期値を変えたり，2 項だけではなく 3 項以上動かないなどして判定を厳しくすると，正しい解が求まることがある． □

問 *8.5.* $||\boldsymbol{x}||_1\leq\sqrt{n}||\boldsymbol{x}||_2$ 以外は素直に示すことができる．また，この式は次のように示せる．コーシーシュワルツの不等式

$$\left(\sum_{i=1}^{n}x_iy_i\right)^2\leq\left(\sum_{i=1}^{n}x_i^2\right)\left(\sum_{i=1}^{n}y_i^2\right)^2$$

において，x_i を $|x_i|$ に置き換え，$y_i=1$ とおくと，

$$\left(\sum_{i=1}^{n}|x_i|\right)^2\leq n\cdot\left(\sum_{i=1}^{n}x_i^2\right).$$

ルートをとって，$||\boldsymbol{x}||_1\leq\sqrt{n}||\boldsymbol{x}||_2$. □

問 *9.3.* 任意の $x, y \in \mathbb{R}$ に対して，平均値の定理より，

$$\frac{f(x)-f(y)}{x-y} = f'(\theta)$$

を満たす θ, $x < \theta < y$, が存在する．よって，

$$|f(x)-f(y)| = |f'(\theta)||x-y| < q|x-y|.$$

すなわち f は縮小写像.

注意：多変数にも拡張できる．余力があればやってみよう． □

問 *11.1.* (1) 帰納法による．$n = 0, 1$ の場合は明らか．$P_{n-1}(x)$ が $n-1$ 次式，$P_{n-2}(x)$ が $n-2$ 次式であれば，漸化式より $P_n(x)$ は n 次式である．

(2) 帰納法による．0 次の多項式，すなわち定数は，$b_0 P_0(x)$ の形で表される．n 次の多項式が題意のような形で表されたとすると，$n+1$ 次の多項式 $Q(x)$ は，$Q(x) = b_{n+1}P_{n+1}(x) + R(x)$ と書ける．ここで R はたかだか n 次の多項式．仮定から $R(x)$ も題意のような形で表されるので，$Q(x)$ も題意のような形で表される．特に n 次多項式は $\sum_{k=0}^{n} b_k P_k(x)$ と書ける．

(3) x^n は n 次多項式なので，$x^n = \sum_{k=0}^{n} b_k P_k(x)$ と書ける．よって，

$$\int_{-1}^{1} P_m(x) \sum_{k=0}^{n} b_k P_k(x) dx = \sum_{k=0}^{n} b_k \int_{-1}^{1} P_m(x) P_k(x) dx = 0.$$

(4) $n \geq 1$ について，$\int_{-1}^{1} P_n(x)dx = \int_{-1}^{1} P_n(x)P_0(x)dx = 0$ である．よって，

$$\int_{-1}^{1} Q(x)dx = \int_{-1}^{1} \sum_{k=0}^{n} b_k P_k(x) dx = \int_{-1}^{1} b_0 P_0(x) dx = 2b_0.$$

□

13　今後の予定

授業の準備や理解度等に従って，適宜修正する予定である．

(1) 9 月 24 日 (水) 第 1 回　不正確な計算と実数の表現
(2) 10 月 1 日 (水) 第 2 回　誤差の種類
(3) 10 月 8 日 (水) 第 3 回　誤差の評価と計算ステップ数の評価
(4) 10 月 15 日 (水) 第 4 回　行列の四則演算
(5) 10 月 22 日 (水) 第 5 回　ガウスの消去法
(6) 10 月 29 日 (水) 第 6 回　 LU 分解
(7) 11 月 5 日 (水) 第 7 回　中間テスト，２分法とニュートン法
(8) 11 月 12 日 (水) 休講　→　補講
(9) 11 月 19 日 (水) 休講
(10) 11 月 26 日 (水) 第 8 回　反復法
(11) 12 月 3 日 (水) 第 9 回　反復法の収束と縮小写像の原理
(12) 12 月 10 日 (水) 第 10 回　数値積分とシンプソンの公式
(13) 12 月 17 日 (水) 第 11 回　ガウス積分
(14) 12 月 24 日 (水) 休講
(15) 12 月 29 日 (水) 休講
(16) 1 月 7 日 (水) 休講　→　補講
(17) 1 月 14 日 (水) 第 12 回　中間テスト，区間演算
(18) 1 月 21 日 (水) 補講　総復習
(19) 1 月 22 日 (木) 補講　総復習

中間テストは成績評価に加えない．