

1.4 連立1次方程式

連立1次方程式は

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n = b_n \end{cases} \quad (1.13)$$

と書ける. 式の本数と未知数の数が同じこと, $x_1 \cdot x_2$ のような未知数の積を含む項がないことに注意しよう. 式(1.13)は $n \times n$ の行列 A と n 次元の列ベクトル x, b を用いて

$$Ax = b$$

と書ける. このとき

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}, \quad x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

である.

$Ax = b$ が解を持つ条件は

$$\det(A) \neq 0$$

$$\text{rank}(A) = n$$

である.

この方程式を実際に解くにはガウスの消去法を使う. コンピュータに解かせることを前提にするなら, 単純かつ汎用性のある手順を作り上げることが重要である.

1.4.1 直接法 - ガウスの消去法 -

ガウス (Gauss) の消去法では, まず最初に第1番方程式を利用して x_1 の係数を消去 (elimination) する. それには第1番方程式を何倍かして加え第2番方程式から第 n 番方程式までの x_1 の係数を0にすればよい. この操作は同値な変形であ

名取亮編

数値計算法, オム社, 1998

1.4 連立1次方程式

る. つまり第 i 番方程式用の乗数を α_{i1} とすると

$$\begin{array}{r} \alpha_{i1}(a_{11}x_1 + \dots + a_{1n}x_n) = \alpha_{i1}b_1 \\ +) \quad a_{i1}x_1 + \dots + a_{in}x_n = b_i \\ \hline (\alpha_{i1}a_{11} + a_{i1})x_1 + \dots + (\alpha_{i1}a_{1n} + a_{in})x_n = \alpha_{i1}b_1 + b_i \end{array}$$

である. $\alpha_{i1}a_{11} + a_{i1} = 0$ とするには

$$\alpha_{i1} = -\frac{a_{i1}}{a_{11}} \quad (i = 2, \dots, n)$$

とすればよい. $a_{11} \neq 0$ ならこの操作は実行できて, 方程式ごとに異なった $\alpha_{21}, \alpha_{31}, \dots, \alpha_{n1}$ が求まる. 第1番方程式に α_{i1} を乗じて第 i 番方程式 ($i = 2, \dots, n$) に加える操作を繰り返した結果, 連立方程式(1.13)は同値な連立方程式

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{22}^{(1)}x_2 + \dots + a_{2n}^{(1)}x_n = b_2^{(1)} \\ \vdots \\ a_{n2}^{(1)}x_2 + \dots + a_{nn}^{(1)}x_n = b_n^{(1)} \end{cases}$$

に変換される. ここで $a_{22}^{(1)}, b_2^{(1)}$ などとあるのは

$$\begin{cases} \alpha_{i1} = -\frac{a_{i1}}{a_{11}} \\ a_{ij}^{(1)} = a_{ij} + \alpha_{i1}a_{1j} \quad (i=2, \dots, n) \\ b_i^{(1)} = b_i + \alpha_{i1}b_1 \quad (i=2, \dots, n) \end{cases}$$

によって更新された値である. この操作を第 $(k-1)$ 番方程式を用いた消去まで繰り返すと

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1k}x_k + \dots + a_{1n}x_n = b_1 \\ a_{22}^{(1)}x_2 + \dots + a_{2k}^{(1)}x_k + \dots + a_{2n}^{(1)}x_n = b_2^{(1)} \\ \vdots \\ a_{kk}^{(k-1)}x_k + \dots + a_{kn}^{(k-1)}x_n = b_k^{(k-1)} \\ \vdots \\ a_{nk}^{(k-1)}x_k + \dots + a_{nn}^{(k-1)}x_n = b_n^{(k-1)} \end{cases}$$

となる. 次の第 k 段では「第 k 番方程式を用いて第 $(k+1)$ 番以降の方程式の x_k を消去」する. すなわち $a_{kk}^{(k-1)} \neq 0$ のとき

$$\alpha_{ik} = -\frac{a_{ik}^{(k-1)}}{a_{kk}^{(k-1)}} \quad (i = k+1, \dots, n)$$

を用いて

$$\begin{cases} a_{ij}^{(k)} &= a_{ij}^{(k-1)} + \alpha_{ik} a_{kj}^{(k-1)} \\ b_i^{(k)} &= b_i^{(k-1)} + \alpha_{ik} b_k^{(k-1)} \end{cases}$$

という操作を $i = k+1, \dots, n; j = k+1, \dots, n$ に施せばよい。このようにして第 $(n-1)$ 番方程式による消去までを繰り返すと、連立方程式は

$$\begin{cases} a_{11}x_1 + \dots + a_{1k}x_k + \dots + a_{1n}x_n = b_1 \\ \quad a_{22}^{(1)}x_2 + a_{2k}^{(1)}x_k + \dots + a_{2n}^{(1)}x_n = b_2^{(1)} \\ \quad \quad \quad \quad \quad a_{kk}^{(k-1)}x_k + \dots + a_{kn}^{(k-1)}x_n = b_k^{(k-1)} \\ \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad a_{n-1,n-1}^{(n-2)}x_{n-1} + a_{n-1,n}^{(n-2)}x_n = b_{n-1}^{(n-2)} \\ \quad a_{nn}^{(n-1)}x_n = b_n^{(n-1)} \end{cases}$$

となる。ここまでの過程を前進消去過程 (forward elimination process), または上三角化過程という。いちばん下の第 n 番方程式には、未知数が x_n だけしか含まれていないので、 $a_{nn}^{(n-1)} \neq 0$ なら

$$x_n = \frac{b_n^{(n-1)}}{a_{nn}^{(n-1)}}$$

と解ける。第 n 番方程式から x_n が求まれば、第 $(n-1)$ 番方程式より

$$x_{n-1} = \frac{b_{n-1}^{(n-2)} - a_{n-1,n}^{(n-2)}x_n}{a_{n-1,n-1}^{(n-2)}}$$

が求まる。このように下から順に繰り返せば、 x_k を計算する時点で、 x_{k+1}, \dots, x_n は既知なので

$$x_k = \frac{b_k^{(k-1)} - \sum_{j=k+1}^n a_{kj}^{(k-1)}x_j}{a_{kk}^{(k-1)}}$$

として求められる。この過程を後代入過程 (backward substitution process) という。前進消去が問題なく終了していれば $a_{kk}^{(k-1)} \neq 0$ である。

ガウスの消去法を用いれば、解が存在する多くの連立方程式は解けるが、次のようなとき

$$\begin{cases} 0x_1 + 1x_2 + 2x_3 = 2 \\ 1x_1 + 0x_2 + 3x_3 = 2 \\ 3x_1 + 1x_2 + 0x_3 = -3 \end{cases} \quad (1.14)$$

は a_{11} が 0 となって、解が存在するのにもかかわらず解けない。このような場合でもアルゴリズムが破綻しないよう、部分軸選択 (partial pivoting) を導入する。

部分軸選択とは第 k 段において、第 k 番方程式をそのまま使うのではなく、第 k 番方程式から第 n 番方程式までの x_k の係数 $a_{k,k}^{(k-1)}, \dots, a_{n,k}^{(k-1)}$ から絶対値が最大のものを探し、その係数を持つ方程式と第 k 番方程式を入れ換えてから消去を行う。これによって、単に解けるようになるだけでなく、各段の作業において $|\alpha_{ik}| \leq 1$ が保証され、コンピュータを使用する際の丸め誤差の増幅を防ぐことにも役立っている。消去結果のすべての対角要素がゼロでなければ、この行列は正則である。

例として、式 (1.14) を部分軸選択つきガウスの消去法で解いてみよう。

前進消去過程

$k = 1$ 段:

1 列目 1 行以下で絶対値最大の要素は 3 行目にあるので、第 1 番方程式と第 3 番方程式を交換する。

$$\begin{cases} 3x_1 + 1x_2 + 0x_3 = -3 \\ 1x_1 + 0x_2 + 3x_3 = 2 \\ 0x_1 + 1x_2 + 2x_3 = 2 \end{cases}$$

第 2 番方程式の x_1 を消去するため第 1 番方程式 $\times (\ominus \frac{1}{3})$ を第 2 番方程式に加える。

$$\begin{cases} 3x_1 + 1x_2 + 0x_3 = -3 \\ 0x_1 + \frac{1}{3}x_2 + 1x_3 = \frac{1}{3} \\ 0x_1 + 1x_2 + 2x_3 = 2 \end{cases}$$

$k = 2$ 段:

2 列目 2 行以下で絶対値最大の要素は 3 行目にあるので、第 2 番方程式と第 3 番方程式を交換する。

$$\begin{cases} 3x_1 + 1x_2 + 0x_3 = -3 \\ 0x_1 + 1x_2 + 2x_3 = 2 \\ 0x_1 + \frac{1}{3}x_2 + 1x_3 = \frac{1}{3} \end{cases}$$

第3番方程式の x_2 を消去するため、第2番方程式 $\times (-\frac{1}{3})$ を第3番方程式に加える。

$$\begin{cases} 3x_1 + 1x_2 + 0x_3 = -3 \\ 0x_1 + 1x_2 + 2x_3 = 2 \\ 0x_1 + 0x_2 + \frac{1}{3}x_3 = \frac{1}{3} \end{cases}$$

$k = 3$ 段:

第3番方程式の x_3 の係数はゼロではない。前進消去過程ですべての対角要素 $\neq 0$ となったので、行列 A は正則である。

後退代入過程

$k = 3$ 段: $\frac{1}{3}x_3 = \frac{1}{3}$ より $x_3 = 1$.

$k = 2$ 段: $x_2 + 2x_3 = 2$ を移項した $x_2 = 2 - 2x_3$ に $x_3 = 1$ を代入して $x_2 = 0$.

$k = 1$ 段: $3x_1 + x_2 = -3$ より $x_1 = \frac{-3-x_2}{3}$, $x_2 = 0$ を代入して $x_1 = -1$.

部分軸選択つきガウスの消去法のアルゴリズムをプログラムふうにとまとめたのが例 1.1 である。記号は

- $=$ は右辺の値を左辺に代入
- \rightleftharpoons は値の交換
- for $k = 1, 2, \dots, n$ は k を 1 から n まで変えた繰り返し

を意味する。 ϵ 以下をゼロとみなす。 ir にはゼロとみなされた回数が入る。例 1.1 のアルゴリズムに必要な浮動小数演算の概数を見積もると

前進消去過程	左辺	$\frac{2}{3}n^3$ 回	(*)
	右辺	n^2 回	(**)
後退代入過程	右辺	n^2 回	(***)

<例 1.1> $Ax = b$ を部分軸選択つきガウスの消去法で解く。

[アルゴリズム]

```

ir=0
for k = 1, 2, ..., n
    amax = |akk| ; ip = k
    for i = k + 1, k + 2, ..., n
        if |aik| > amax then
            amax = |aik| ; ip = k
        end if
    end for
    if amax > ε then
        for j = k + 1, k + 2, ..., n
            aip,j  $\rightleftharpoons$  akj
            for i = k + 1, k + 2, ..., n
                aik = -aik/akk
                for j = k + 1, k + 2, ..., n
                    aij = aij + aik · akj (*)
                end for
                bi = bi + aik · bk (**)
            end for
        end for
        ir = ir + 1
    end if
    if ir  $\neq$  0 then
        for k = n, n - 1, ..., 1
            xk = (bk -  $\sum_{j=k+1}^n$  akj · xj) / akk (***)
        end for
    end if

```

である。左辺の前進消去と右辺の前進消去・後退代入は分離可能で、しかも複数の右辺に対しても左辺の前進消去は1回でよい。したがって n 元連立1次方程式を p

回解くための浮動小数演算回数はおおよそ

$$\frac{2}{3}n^3 + p(2n^2) \text{ 回}$$

である。 $p = n$ のとき、すなわち $AX = B$ を満足する行列 X 、あるいは逆行列 ($AX = I$) を求めるとき必要な浮動小数演算回数はおおよそ $\frac{8}{3}n^3$ 回である。式 (1.8) の解を逆行列を使って書けば

$$x = A^{-1}b$$

だが、コンピュータでこの式のとおり逆行列 A^{-1} を作って解こうとすると

$$\frac{8}{3}n^3 + 2n^2 \text{ 回}$$

の浮動小数演算が必要となる。ガウスの消去法で直接計算する場合の $\frac{2}{3}n^3 + 2n^2$ と比べるといかに無駄な計算をしているかがわかる。しかも、逆行列を使うと誤差が入りやすい。このため、逆行列の要素の値が知りたいとき以外、数値計算では逆行列を使わないのが常識である。

ガウスの消去法 k 段の操作

P_k : k 番方程式と ip 番方程式を入れ換える

G_k : k 番方程式を α_{ik} 倍して i 番方程式 ($i = k + 1, \dots, n$) に加える

は次のような行列で表せる。

$$P_k = \begin{matrix} & & k \text{ 列} & & ip \text{ 列} & & \\ \begin{matrix} k \text{ 行} \\ ip \text{ 行} \end{matrix} & \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \ddots & & & \\ & & & & 0 & \dots & 1 & \dots & \dots & \\ & & & & & \ddots & & & & \\ & & & & & & & 1 & & \dots & \\ & & & & & & & & & \ddots & \\ & & & & & & & & & & 1 \end{pmatrix} \end{matrix}$$

$$G_k = \begin{matrix} & & & & k \text{ 列} & & \\ \begin{matrix} 1 \\ \vdots \\ \alpha_{k+1,k} \\ \vdots \\ \alpha_{i,k} \\ \vdots \\ \alpha_{n,k} \end{matrix} & \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \ddots & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \end{matrix}$$

ガウスの消去法の前進消去過程は P_k, G_k を使って

$$G_{n-1}P_{n-1} \dots G_kP_k \dots G_1P_1Ax = G_{n-1}P_{n-1} \dots G_1P_1b$$

と書ける。 $G_{n-1}P_{n-1} \dots G_kP_k \dots G_1P_1A$ は上三角行列 (upper triangular matrix)

$$U = \begin{pmatrix} a_{11} & \dots & \dots & \dots & a_{1n} \\ & a_{22}^{(1)} & \dots & \dots & a_{2n}^{(1)} \\ & & a_{kk}^{(k-1)} & \dots & a_{kn}^{(k-1)} \\ & & & \dots & a_{nn}^{(n-1)} \end{pmatrix}$$

である。

$$G_{n-1}P_{n-1} \dots G_1P_1A = U$$

より

$$(G_{n-1}P_{n-1} \dots G_1P_1)^{-1}U = A$$

と書いて A の LU 分解 (LU decomposition または LU factorization) という事もある。狭義の LU 分解では A を下三角行列 (lower triangular matrix) L と上三角行列 U に分解しておき、方程式 $Ax = b$ の右辺に対する計算を

$$Ly = b$$

$$Ux = y$$

の2段階、すなわち L に対する前進代入と U に対する後退代入に分けて行う。LU 分解とガウスの消去法に本質的な差はない。またガウスの消去法には、扱う行列の特徴、計算環境などによっていろいろな手法がある。

1.3 ノルム, 誤差, 残差, 条件数

$n \times n$ の行列 A と n 次元ベクトル x, b で表された連立 1 次方程式

$$Ax = b \quad (1.8)$$

を考える. ベクトル x を 1 つの数値で代表させたものにノルム (norm) があり, $\|x\|$ と書く. ベクトルのノルムには

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

$$\|x\|_2 = \left(\sum_{i=1}^n x_i^2 \right)^{\frac{1}{2}}$$

$$\|x\|_\infty = \max_i |x_i|$$

などがあり, $\|x\|_1$ を 1-ノルム, $\|x\|_2$ を 2-ノルム (ユークリッドノルム) と呼ぶ. 同様に行列のノルム $\|A\|$ には

$$\|A\|_1 = \sup_{x \neq 0} \frac{\|Ax\|_1}{\|x\|_1}$$

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

$$\|A\|_\infty = \sup_{x \neq 0} \frac{\|Ax\|_\infty}{\|x\|_\infty}$$

などがある. 行列の 1-ノルム $\|A\|_1$ は

$$\|A\|_1 = \max_j \|a_j\|_1 = \max_j \left(\sum_{i=1}^n |a_{ij}| \right)$$

が成り立つので, 簡単に計算できる. 2-ノルム $\|A\|_2$ に対しては

$$\|A\|_2 = (\lambda \max(A^T A))^{\frac{1}{2}}$$

という式が成り立つ. $\lambda \max(A^T A)$ は $A^T A$ の最大固有値であり, A が対称なら

$$(\lambda \max(A^T A))^{\frac{1}{2}} = |\lambda \max(A)|$$

である. 一般に A の絶対値最大の固有値を $\rho(A)$ とすると

$$\rho(A) \leq \|A\|_2, \quad \rho(A) \leq \|A\|_1$$

← 問題あり!

である。 $\rho(A)$ をスペクトル半径という。

ベクトルと行列のノルムは性質

$$\begin{aligned}\|\alpha x\| &= |\alpha| \cdot \|x\| \\ \|x + y\| &\leq \|x\| + \|y\| \\ \|Ax\| &\leq \|A\| \cdot \|x\|\end{aligned}$$

を満足する。また1-ノルムと2-ノルムは

$$\frac{1}{\sqrt{n}} \|A\|_1 \leq \|A\|_2 \leq \sqrt{n} \|A\|_1$$

という関係を満足する。

さて、式(1.8)をコンピュータで解いて得られた数値解を \hat{x} とすると、数を有限桁の浮動小数で表現し、有限桁で計算を実行したことにより一般には $\hat{x} \neq x$ となる。このとき

$$\hat{x} - x$$

を誤差 (error) あるいは誤差ベクトルという。誤差の程度は x によっても変化するため x のノルムで割って、**相対誤差** (relative error) あるいは相対誤差ベクトル

$$\frac{\hat{x} - x}{\|x\|}$$

で評価することもある。誤差の大きさだけが興味の対象なら、スカラー量である相対誤差ノルム

$$\frac{\|\hat{x} - x\|}{\|x\|}$$

で評価すればよい。

実際に解 x がわかっていることは稀なので (わかっているなら計算しない)、別の基準を用いて計算解の正しさを評価することになる。もっともよく使われるのは

$$b - A\hat{x}$$

で**残差** (residual) または残差ベクトルという。

残差に対しても用途に応じて、相対残差、相対残差ノルムを考える。通常、ベクトルに対するノルムは2-ノルムが使われる。

式(1.8)をコンピュータで計算したとき、 A は正しく表されたが b に何らかの誤差 δb があって、その結果、計算解が $x + \delta x$ になったとしよう。式で書くと

$$A(x + \delta x) = b + \delta b \quad (1.9)$$

である。

式(1.8),(1.9)から

$$\delta x = A^{-1}(\delta b)$$

となり、両辺のノルムをとれば

$$\|\delta x\| \leq \|A^{-1}\| \cdot \|\delta b\| \quad (1.10)$$

である。式(1.8)の両辺のノルムをとれば

$$\|A\| \cdot \|x\| \geq \|b\| \quad (1.11)$$

なので、式(1.10),(1.11)を合わせると

$$\frac{\|\delta x\|}{\|x\|} \leq \|A\| \cdot \|A^{-1}\| \frac{\|\delta b\|}{\|b\|} \quad (1.12)$$

となる。 $\|A\| \cdot \|A^{-1}\|$ のことを行列 A の**条件数** (condition number) といい、 $\text{cond}(A)$ と書く。 $AA^{-1} = I$ より

$$\text{cond}(A) = \|A\| \cdot \|A^{-1}\| \geq 1$$

である。

式(1.12)は式(1.9)の右辺に入った誤差の割合 $\frac{\|\delta b\|}{\|b\|}$ が、左辺 $\frac{\|\delta x\|}{\|x\|}$ でどれくらいの倍率になりうるかを示している。式(1.12)はコンピュータ上の数値表現や演算方式に関係していないことに注意しよう。これは純粋に数学的な関係であり、行列 A と右辺 $b, \delta b$ によって決まる。

実際、コンピュータで連立1次方程式を解けば

$$\frac{\|\delta x\|}{\|x\|} = \alpha \frac{\|\delta b\|}{\|b\|} \leq \text{cond}(A) \frac{\|\delta b\|}{\|b\|}$$

を満足する α が求められる。 α は $\text{cond}(A)$ に近くなることもある。 $\text{cond}(A)$ が大きいとき**悪条件方程式**、あるいは方程式の性質がよくないという。悪条件方程式は誤差が入りやすいので、コンピュータで扱うときにはよりいっそうの注意がいる。

定常反復法

- Jacobi法 n 本の方程式に独立に

$$x_i^{(k)} = \left\{ b_i - \sum_{j \neq i} a_{ij} x_j^{(k-1)} \right\} / a_{ii}$$

- Gauss-Seidel法 i より前の $x_l^{(k)}$ に利用する

$$x_i^{(k)} = \left\{ b_i - \sum_{j < i} a_{ij} x_j^{(k)} - \sum_{j > i} a_{ij} x_j^{(k-1)} \right\} / a_{ii}$$

この部分がより新しい値

- Successive Over-Relaxation法 (逐次的過剰緩和法)

$$x_i^{(k)} = \omega \tilde{x}_i^{(k)} + (1-\omega) x_i^{(k-1)}$$

$$\tilde{x}_i^{(k)} = \left\{ b_i - \sum_{j < i} a_{ij} \tilde{x}_j^{(k)} - \sum_{j > i} a_{ij} x_j^{(k-1)} \right\} / a_{ii}$$

Gauss-Seidel法の反復解

$x_i^{(k)}$ に直接更新するのでなく $\omega \in (0, 2)$ により更新量を調節

$0 < \omega < 1$: 過小 $\omega = 1$: Gauss-Seidel法 $1 < \omega < 2$: 過大