

計算物理学 2 第 7 回：逆行列・固有値問題

ver. 2018/5/30

1 逆行列

前回は n 元の連立一次方程式の解法について消去法や反復法などの方法を解説しました。与えられた行列 $A = \{a_{ij}\}$ と列ベクトル $\mathbf{b} = \{b_i\}$ を用いて

$$A\mathbf{x} = \mathbf{b} \quad (1)$$

を解き、列ベクトル $\mathbf{x} = \{x_i\}$ を求める問題でした。連立一次方程式を解くプログラムを少し拡張することで行列 A の逆行列 A^{-1} を求めることができます。 A の逆行列は

$$AA^{-1} = I \quad (2)$$

を満たします。 $n \times n$ 行列 A^{-1} を n 個の列ベクトルを並べたものとし、 $A^{-1} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ と書き、単位行列も同様に $I = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ と書くと、式 (2) は

$$A\mathbf{x}_i = \mathbf{e}_i, \quad (i = 1, 2, \dots, n) \quad (3)$$

と n 組の n 元連立一次方程式に分解できます。右辺の \mathbf{b} の値を $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ の n 種類用意したこれらの方程式をすべて解くことで A の逆行列を求めることができます。

1.1 LU 分解

逆行列を求めるためには連立一次方程式を解くいずれの方法を用いてもよいのですが、これらの n 組の連立一次方程式では A の値は同一であるため、消去法では全く同じ計算を行列 A に対して行ってしまうことになり効率がよくありません。ここで説明する LU 分解は同じ行列 A に対して複数の列ベクトル \mathbf{b} に対して連立方程式を解く場合に効率的な方法です。

ガウスの消去法ではまず前進消去を行いました。1 行目の式を用いて変数 x_1 を $2 - n$ 番目の式から消去し、次に 2 行目の式を用いて変数 x_2 を $3 - n$ 番目の式から消去し、と繰り返すことで行列 A を上三角行列の形に変換しました。このときの連立方程式の形は

$$\begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & \cdots & a_{1n}^{(1)} \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} & \cdots & a_{2n}^{(2)} \\ 0 & 0 & a_{33}^{(3)} & \cdots & a_{3n}^{(3)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn}^{(n)} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1^{(1)} \\ b_2^{(2)} \\ b_3^{(3)} \\ \vdots \\ b_n^{(n)} \end{pmatrix}, \quad (4)$$

となっています。このときの右辺の値 $b_i^{(i)}$ はもともとの列ベクトルの値 \mathbf{b} を用いて

$$b_1^{(1)} = b_1 \quad (5)$$

$$b_2^{(2)} = b_2 - \frac{a_{21}^{(1)}}{a_{11}^{(1)}} b_1 \quad (6)$$

$$b_3^{(3)} = b_3^{(2)} - \frac{a_{32}^{(2)}}{a_{22}^{(2)}} b_2^{(2)} = b_3 - \frac{a_{31}^{(1)}}{a_{11}^{(1)}} b_1^{(1)} - \frac{a_{32}^{(2)}}{a_{22}^{(2)}} b_2^{(2)}, \quad (7)$$

\vdots

$$\begin{aligned} b_k^{(k)} &= b_k^{(k-1)} - \frac{a_{k,k-1}^{(k-1)}}{a_{k-1,k-1}^{(k-1)}} b_{k-1}^{(k-1)} \\ &= b_k^{(k-2)} - \frac{a_{k,k-2}^{(k-2)}}{a_{k-2,k-2}^{(k-2)}} b_{k-2}^{(k-2)} - \frac{a_{k,k-1}^{(k-1)}}{a_{k-1,k-1}^{(k-1)}} b_{k-1}^{(k-1)} \\ &= b_k - \sum_{m=1}^{k-1} \frac{a_{km}^{(m)}}{a_{mm}^{(m)}} b_m^{(m)} \end{aligned} \quad (8)$$

行列の形に書くと

$$\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ l_{21} & 1 & 0 & \cdots & 0 \\ l_{31} & l_{32} & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ l_{n1} & l_{n2} & l_{n3} & \cdots & 1 \end{pmatrix} \begin{pmatrix} b_1^{(1)} \\ b_2^{(2)} \\ b_3^{(3)} \\ \vdots \\ b_n^{(n)} \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_n \end{pmatrix} \quad (9)$$

ここで $l_{ij} = a_{ij}^{(j)}/a_{jj}^{(j)}$ です。この式では左辺の行列は下三角行列となっています。この式の $(b_1^{(1)}, b_2^{(2)}, \dots, b_n^{(n)})$ に (4) 式の左辺を代入すると

$$\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ l_{21} & 1 & 0 & \cdots & 0 \\ l_{31} & l_{32} & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ l_{n1} & l_{n2} & l_{n3} & \cdots & 1 \end{pmatrix} \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & \cdots & a_{1n}^{(1)} \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} & \cdots & a_{2n}^{(2)} \\ 0 & 0 & a_{33}^{(3)} & \cdots & a_{3n}^{(3)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn}^{(n)} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_n \end{pmatrix} \quad (10)$$

となります。これは行列 A が下三角行列

$$L = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ l_{21} & 1 & 0 & \cdots & 0 \\ l_{31} & l_{32} & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ l_{n1} & l_{n2} & l_{n3} & \cdots & 1 \end{pmatrix} \quad (11)$$

と上三角行列

$$U = \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & \cdots & a_{1n}^{(1)} \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} & \cdots & a_{2n}^{(2)} \\ 0 & 0 & a_{33}^{(3)} & \cdots & a_{3n}^{(3)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn}^{(n)} \end{pmatrix} \quad (12)$$

に分解できていることになります。これを LU 分解と言います。 $(A = LU)$ 。LU 分解は行列 A に対して行われる操作で列ベクトル \mathbf{b} に依りません。 A の逆行列を求めるときはまず LU 分解を行います。(9) 式は $(b_1^{(1)}, b_2^{(2)}, \dots, b_n^{(n)})$ に関する連立一次方程式だと見ることができます。 L は上三角行列なので、前進代入 (forward substitution) によって、一行目から順番に解を求めることができます。 \mathbf{b} の値を変えながら $b_i^{(i)}$ を求め、それを用いて行列 U を使って後方代入を行って \mathbf{x} を決定します。これは三角行列の逆行列を用いて

$$A^{-1} = U^{-1}L^{-1} \quad (13)$$

を計算していることと同じですが、三角行列の逆行列の計算量は少ないため、直接 A から A^{-1} を計算するよりも効率よくなります。

LU 分解のもう一つの利点は行列式の計算が簡単になることです。三角行列の行列式は対角要素の積になりますので、 A の行列式は

$$\det A = \det L \det U = \prod_{i=1}^n a_{ii}^{(i)} \quad (14)$$

で与えられます。

1.2 修正 Cholesky 分解

LU 分解を少し発展させた例として、修正 Cholesky 分解があります。これは実正平方行列 A を

$$A = LDL^T \quad (15)$$

の形に分解します。ここで L は対角要素が 1 の下三角行列、 D は対角行列で以下のように与えられます。

$$L = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ l_{21} & 1 & 0 & \cdots & 0 \\ l_{31} & l_{32} & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ l_{n1} & l_{n2} & l_{n3} & \cdots & 1 \end{pmatrix}, \quad D = \begin{pmatrix} d_1 & 0 & 0 & \cdots & 0 \\ 0 & d_2 & 0 & \cdots & 0 \\ 0 & 0 & d_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & d_n \end{pmatrix}, \quad (16)$$

DL^T を計算すると

$$DL^T = \begin{pmatrix} d_1 & d_1 l_{21} & d_1 l_{31} & \cdots & d_1 l_{n1} \\ 0 & d_2 & d_2 l_{32} & \cdots & d_2 l_{n2} \\ 0 & 0 & d_3 & \cdots & d_3 l_{n3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & d_n \end{pmatrix}, \quad (17)$$

と上三角行列になることから、修正 Cholesky 分解の L と LU 分解の L は同じもの、 D は $DL^T = U$ の関係にあることがわかります。 d_j と l_{ij} の値は

$$d_1 = a_{11}^{(1)}, \quad (18)$$

$$l_{i1} = \frac{a_{i1}^{(1)}}{d_1} \quad (19)$$

$$d_2 = a_{22}^{(2)} = a_{22} - d_1 l_{21}^2 \quad (20)$$

$$l_{i2} = \frac{a_{i2}^{(2)}}{d_2} = \frac{a_{i2} - d_1 l_{i1} l_{21}}{d_2} \quad (21)$$

\vdots

$$d_j = a_{jj} - \sum_{k=1}^{j-1} d_k l_{jk}^2, \quad (22)$$

$$l_{ij} = \frac{a_{ij} - \sum_{k=1}^{j-1} d_k l_{ik} l_{jk}}{d_j} \quad (23)$$

として求められます。

2 固有値問題

古典力学から量子力学まで物理学の多くの問題は固有値問題に帰着するため、固有値問題を数値的に取り扱うことはとても重要です。講義の後半で取り扱う微分方程式も行列の形に書き直せば固有値問題として取り扱うことができる場合があります。

固有値問題とは n 次元正方行列 A について

$$A\mathbf{u}_i = \lambda_i \mathbf{u}_i \quad (24)$$

となる固有値 $\{\lambda_i\}$ と固有ベクトル $\{\mathbf{u}_i\}$ を求める問題です。

成分で書くと

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} u_{1i} \\ u_{2i} \\ u_{3i} \\ \vdots \\ u_{ni} \end{pmatrix} = \lambda_i \begin{pmatrix} u_{1i} \\ u_{2i} \\ u_{3i} \\ \vdots \\ u_{ni} \end{pmatrix} \quad (25)$$

となります。

固有値は一般的には特性方程式を解くことで得られます。これは

$$\det(A - \lambda I) = \begin{vmatrix} a_{11} - \lambda & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} - \lambda & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} - \lambda \end{vmatrix} = 0 \quad (26)$$

と単位行列 I を用いて書かれます。特性方程式は λ に関して n 次方程式となり、 n 個の解がありますが次元 n が大きい場合に直接解くのは大変です。数値計算では行列 A を対角化することで固有値と固有ベクトルを求めます。

以下では A が実対称行列とします。 n 次の正方実対称行列 A が対角化可能である場合、 A は

$$A = UDU^{-1} \quad (27)$$

と直交行列 U と対角行列 D を用いて分解できます。成分で書くと

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} = \begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ u_{21} & u_{22} & \cdots & u_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ u_{n1} & u_{n2} & \cdots & u_{nn} \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \begin{pmatrix} u_{11} & u_{21} & \cdots & u_{n1} \\ u_{12} & u_{22} & \cdots & u_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ u_{1n} & u_{2n} & \cdots & u_{nn} \end{pmatrix} \quad (28)$$

右辺の U^{-1} を左辺に持ってきて $AU = UD$ として一列ずつ比較をすると D の対角成分が固有値 λ_i に対応し、 $U = (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n)$ となっていることがわかります。つまり固有値問題は行列 A を対角化することによって解くことができます。

2.1 Jacobi 法

(27) 式より対角化とは行列 A に対して $U^{-1}AU$ が対角行列 D となるような直交行列 U を求める問題に帰着します。Jacobi 法では反復法によってこの直交行列 U を求めます。

まずは行列 $A = A_0$ を適当な U_0 という直交行列を用いて

$$A_1 = U_0^{-1}A_0U_0 \quad (29)$$

と変換します。 U_0 は最終的に求めたい解 U ではないため、 A_1 は対角行列ではありませんが、 A_1 が A_0 よりも対角行列に近くなるような U_0 を使います。続いて同様に U_1 という直交行列を用いて A_1 を変換し、

$$A_2 = U_1^{-1}A_1U_1 \quad (30)$$

により A_2 を求め、 A_2 が A_1 よりも対角行列に近くなるように決めます。この操作を繰り返し、一般に直交行列 U_k を用いて行列 A_k を変換し、

$$A_{k+1} = U_k^{-1}A_kU_k \quad (31)$$

が A_k よりも対角行列に近くなるようにします。最終的に A_m が十分対角行列とみなせるようになれば、

$$A_m = U_{m-1}^{-1}A_{m-1}U_{m-1} = (U_{m-1}^{-1}U_{m-2}^{-1})A_{m-2}(U_{m-2}U_{m-1}) = (U_0U_1 \cdots U_{m-1})^{-1}A(U_0U_1 \cdots U_{m-1}) \quad (32)$$

となり、対角化が完了します。ここで対角行列 D は A_m で与えられ、直交行列 U は行列積 $U_0U_1 \cdots U_{m-1}$ で与えられます。

それでは対角行列に近づけるにはどのように直交行列 U_k を各ステップで選べばよいか、という問題になりますが、各ステップで A_k の非対角行列要素 $a_{ij}(i \neq j)$ のうち、絶対値が最大のものをゼロにするように選び

ます。ゼロにしたい行列要素を a_{pq} とします。 $(A$ は直交変換の間つねに対称行列なので $a_{pq} = a_{qp}$ です) このとき、 U_k として Givens 回転 (Jacobi 回転) を用います。これは

$$U_k = \begin{pmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & \cos \theta & 0 & \cdots & 0 & \sin \theta & \\ & & & 0 & 1 & & & 0 & \\ & & & \vdots & & \ddots & & \vdots & \\ & & & 0 & & & 1 & 0 & \\ & & & -\sin \theta & 0 & \cdots & 0 & \cos \theta & \\ & & & & & & & & 1 & \\ & & & & & & & & & \ddots & \\ & & & & & & & & & & 1 \end{pmatrix} \quad (33)$$

の形をしており、 p, q 行や列以外では単位行列、 p, q 行、 p, q 列で二次元回転となっています。式で書くところの行列要素 u_{ij} は

$$u_{pp} = u_{qq} = \cos \theta \quad (34)$$

$$u_{pq} = -u_{qp} = \sin \theta \quad (35)$$

$$u_{ii} = 1 \quad (i \neq p \text{ and } i \neq q) \quad (36)$$

でその他の行列要素はゼロです。 U_k は二次元回転行列なので直交行列です。この U_k を式 (31) に代入すると、 A_{k+1} の行列要素 $a_{ij}^{(k+1)}$ は A_k の行列要素 $a_{ij}^{(k)}$ を用いて

$$a_{pp}^{(k+1)} = a_{pp}^{(k)} \cos^2 \theta + a_{qq}^{(k)} \sin^2 \theta - 2a_{pq}^{(k)} \sin \theta \cos \theta, \quad (37)$$

$$a_{qq}^{(k+1)} = a_{pp}^{(k)} \sin^2 \theta + a_{qq}^{(k)} \cos^2 \theta + 2a_{pq}^{(k)} \sin \theta \cos \theta, \quad (38)$$

$$a_{pq}^{(k+1)} = a_{qp}^{(k+1)} = a_{pq}^{(k)} (\cos^2 \theta - \sin^2 \theta) + (a_{pp}^{(k)} - a_{qq}^{(k)}) \sin \theta \cos \theta \quad (39)$$

$$a_{pj}^{(k+1)} = a_{pj}^{(k)} \cos \theta - a_{qj}^{(k)} \sin \theta, \quad (40)$$

$$a_{qj}^{(k+1)} = a_{pj}^{(k)} \sin \theta + a_{qj}^{(k)} \cos \theta, \quad (41)$$

$$a_{ip}^{(k+1)} = a_{ip}^{(k)} \cos \theta - a_{iq}^{(k)} \sin \theta, \quad (42)$$

$$a_{iq}^{(k+1)} = a_{ip}^{(k)} \sin \theta + a_{iq}^{(k)} \cos \theta, \quad (43)$$

と書けます (p, q いずれも含まない行列要素は不変)。いまこの変換によって A_{k+1} の最大の非対角要素 $a_{pq}^{(k+1)}$ をゼロにしたいので、角度 θ は (39) 式より

$$\tan 2\theta = \frac{-2a_{pq}^{(k)}}{a_{pp}^{(k)} - a_{qq}^{(k)}} \quad (44)$$

とすれば $a_{pq}^{(k+1)} = a_{qp}^{(k+1)} = 0$ とできます。この Givens 回転によって最大の非対角要素はゼロとなりますが、その他の行列要素の値も変動しますので中には値が大きくなる非対角行列要素も存在します。しかし、非対角行列要素の二乗和

$$S_k = \sum_{i \neq j} |a_{ij}^{(k)}|^2 \quad (45)$$

を考えると、これは Givens 回転によって

$$S_{k+1} = S_k - 2|a_{pq}^{(k)}|^2 \quad (46)$$

となり、単調減少します。 S_k は正の値なのでこの操作を繰り返すことで S_k は必ずゼロに収束させることができます。

Jacobi 法の基本的なアルゴリズムは以上の通りですが、以下では実装において計算を簡単にする処方を紹介します。Givens 回転で θ の値を (44) 式から求める必要がありますが、 θ の値そのものを求める必要はなく、実際に計算で必要なのは $\sin \theta$ と $\cos \theta$ の値です。 $t = \tan \theta, \alpha = \cot 2\theta$ ((44) 式より α は既知の値です) とすると t に関する 2 次方程式

$$t^2 + 2\alpha t - 1 = 0 \quad (47)$$

で得られるため、この絶対値の小さい方を用いて (計算が安定します)

$$\cos \theta = \frac{1}{\sqrt{1+t^2}}, \quad \sin \theta = t \cos \theta \quad (48)$$

となります。ただし、 t を求める際は二次方程式の解の公式の表式をそのまま使うと

$$t = -\alpha \pm \sqrt{\alpha^2 + 1} \quad (49)$$

となり、この式は α が非常に大きいときには絶対値の小さい方の解で 2 つの大きい値を使った引き算が行われるため桁落ちが起こります。絶対値が小さい方の解は

$$t = \frac{\operatorname{sgn}(\alpha)}{|\alpha| + \sqrt{\alpha^2 + 1}} \quad (50)$$

と書くことができ、($\operatorname{sgn}(\alpha)$ は α の符号です) この表式では引き算がないため桁落ちが防げます。

また、(37)-(43) 式の計算では桁落ちを防ぐために $\beta = \sin \theta / (1 + \cos \theta)$ を導入し、また $a_{pq}^{(k+1)} = 0$ を用いると

$$a_{pp}^{(k+1)} = a_{pp}^{(k)} - ta_{pq}^{(k)}, \quad (51)$$

$$a_{qq}^{(k+1)} = a_{qq}^{(k)} + ta_{pq}^{(k)}, \quad (52)$$

$$a_{ip}^{(k+1)} = a_{ip}^{(k)} - \sin \theta (a_{iq}^{(k)} + \beta a_{ip}^{(k)}), \quad (53)$$

$$a_{iq}^{(k+1)} = a_{iq}^{(k)} + \sin \theta (a_{ip}^{(k)} - \beta a_{iq}^{(k)}), \quad (54)$$

$$a_{pj}^{(k+1)} = a_{jp}^{(k+1)}, \quad (55)$$

$$a_{qj}^{(k+1)} = a_{jq}^{(k+1)}, \quad (56)$$

として計算します。その他の行列要素の値は変わりません。

直交行列 U は

$$U = U_0 U_1 U_2 \cdots U_{m-1} \quad (57)$$

で与えられますが、計算中すべての U_k を保持しておく必要はなく、初期値を単位行列として、それに順番に U_0, U_1, \dots を反復操作中に右からかけてゆくことで直交行列 U を作ります。いま $k+1$ 回目の反復が終わり

U_k が求まったとします。このとき $U^{(k-1)} = U_0 U_1 \cdots U_{k-1}$ に対して $U^{(k)} = U' U_k$ を計算しますが、 U_k は Givens 回転ですので、新しい行列の p 列目と q 列目

$$u_{ip}^{(k)} = u_{ip}^{(k-1)} - \sin \theta (u_{iq}^{(k-1)} + \beta u_{ip}^{(k-1)}), \quad (58)$$

$$u_{iq}^{(k)} = u_{iq}^{(k-1)} + \sin \theta (u_{ip}^{(k-1)} - \beta u_{iq}^{(k-1)}), \quad (59)$$

のみの値が変わります。Givens 回転行列全体の行列積を計算する必要はありません。

2.2 Householder 法

3 重対角行列 (tridiagonal matrix) とは対角成分とその隣の成分のみが値を持つ行列です。以下の形をしています。

$$T = \begin{pmatrix} a_1 & b_1 & 0 & & \cdots & 0 \\ b_1 & a_2 & b_2 & 0 & & \\ 0 & b_2 & a_3 & b_3 & \ddots & \vdots \\ & 0 & b_3 & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & a_{n-1} & b_{n-1} \\ 0 & & \cdots & 0 & b_{n-1} & a_n \end{pmatrix} \quad (60)$$

3 重対角行列の固有値は特性方程式を解くことで数値的に求めることができます (後日取り上げる予定です)。そのため、固有値問題を解く際に、行列を変換して 3 重対角行列の形にすることも有効です。ここでは一般の実対称行列を 3 重対角行列に直交変換する Householder 法について説明します。

行列 A を三重対角化する直交変換 R を求めます。

$$T = R^T A R \quad (61)$$

Householder 法では以下の形の直交変換 R を考えます。($R^T = R$)

$$R = I - \mathbf{w} \mathbf{w}^T \quad (62)$$

ここで I は単位行列、 \mathbf{w} は長さが $\sqrt{2}$ の n 次元ベクトルです。この形の行列を複数回用いることで行列 $A_0 = A = \{a_{ij}^{(0)}\}$ を三重対角行列に変換します。まずはベクトル

$$\mathbf{u}_1 = \begin{pmatrix} 0 \\ a_{21}^{(0)} + s_1 \\ a_{31}^{(0)} \\ \vdots \\ a_{n1}^{(0)} \end{pmatrix} \quad (63)$$

を考えます。 s_1 は

$$s_1 = \operatorname{sgn}(a_{21}^{(0)}) \sqrt{\sum_{i=2}^n (a_{i1}^{(0)})^2} \quad (64)$$

とし、桁落ちをふせぐため $a_{21}^{(0)}$ と同符号にとります。ここで \mathbf{u}_1 を用いて

$$\mathbf{w}_1 = \frac{\mathbf{u}_1}{\sqrt{s_1(a_{21}^{(0)} + s_1)}} \quad (65)$$

とすると、 \mathbf{w}_1 は長さが $\sqrt{2}$ のベクトルになるため、これを用いて直交行列 $R_1 = I - \mathbf{w}_1 \mathbf{w}_1^T$ が作れます。 R_1 を用いて A_0 を変換すると

$$A_1 = R_1 A_0 R_1 = \begin{pmatrix} a_{11}^{(0)} & -s_1 & 0 & \cdots & 0 \\ -s_1 & a_{22}^{(1)} & a_{23}^{(1)} & \cdots & a_{2n}^{(1)} \\ 0 & a_{32}^{(1)} & a_{33}^{(1)} & \cdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2}^{(1)} & a_{n3}^{(1)} & \cdots & a_{nn}^{(1)} \end{pmatrix} \quad (66)$$

となり、一行目と一列目の三重対角化が完了します。続いてベクトル

$$\mathbf{u}_2 = \begin{pmatrix} 0 \\ 0 \\ a_{32}^{(1)} + s_2 \\ a_{42}^{(1)} \\ \vdots \\ a_{n2}^{(1)} \end{pmatrix} \quad (67)$$

を用いて ($s_2 = \text{sgn}(a_{32}^{(1)}) \sqrt{\sum_{i=3}^n (a_{i2}^{(1)})^2}$ です) 同じようにベクトル \mathbf{u}_2 による直交行列 R_2 を作ることで、2 行目と 2 列目についても三重対角化ができます。

一般化すると、行列 A_{k-1} は $k-1$ 行目、 $k-1$ 列目までの三重対角化が完了しており、 k 行目、 k 列目以降の行列要素は $a_{ij}^{(k-1)}$ で与えられているとします。このとき、ベクトル

$$\mathbf{u}_k = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ a_{k+1,k}^{(k-1)} + s_k \\ a_{k+2,k}^{(k-1)} \\ \vdots \\ a_{nk}^{(k-1)} \end{pmatrix} \quad (68)$$

から長さ $\sqrt{2}$ のベクトル

$$\mathbf{w}_k = \frac{\mathbf{u}_k}{\sqrt{s_k(a_{k+1,k}^{(k-1)} + s_k)}} \quad (69)$$

を作り、これを用いて直交変換

$$R_k = I - \mathbf{w}_k \mathbf{w}_k^T \quad (70)$$

を作り、行列 A_{k-1} を変換すると

$$A_k = R_k A_{k-1} R_k = \begin{pmatrix} a_{11}^{(0)} & -s_1 & 0 & \cdots & & & & & 0 \\ -s_1 & a_{22}^{(1)} & -s_2 & 0 & \cdots & & & & \\ 0 & -s_2 & a_{33}^{(2)} & -s_3 & 0 & \cdots & & & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & & & & \\ & & & -s_{k-1} & a_{k,k}^{(k-1)} & -s_k & 0 & \cdots & 0 \\ & & & 0 & -s_k & a_{k+1,k+1}^{(k)} & a_{k+1,k+2}^{(k)} & \cdots & a_{k+1,n}^{(k)} \\ \vdots & & & \vdots & 0 & a_{k+2,k+1}^{(k)} & a_{k+2,k+2}^{(k)} & \cdots & a_{k+2,n}^{(k)} \\ & & & & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & & & \cdots & 0 & a_{n,k+1}^{(k)} & a_{n,k+2}^{(k)} & \cdots & a_{nn}^{(k)} \end{pmatrix} \quad (71)$$

となり、 k 行目、 k 列目までは三重対角化が完了した形となります。これを $k = n - 1$ まで繰り返すことで三重対角化が完了します。

行列 A_k の計算は $\mathbf{p} = A_{k-1} \mathbf{w}_k$ および

$$\mathbf{q} = \mathbf{p} - \frac{1}{2}(\mathbf{w}_k \cdot \mathbf{p})\mathbf{w}_k \quad (72)$$

を用いて

$$\begin{aligned} A_k &= (I - \mathbf{w}_k \mathbf{w}_k^T) A_{k-1} (I - \mathbf{w}_k \mathbf{w}_k^T) \\ &= A_{k-1} - \mathbf{w}_k (\mathbf{w}_k^T A_{k-1}) - (A_{k-1} \mathbf{w}_k) \mathbf{w}_k^T + \mathbf{w}_k (\mathbf{w}_k^T A_{k-1} \mathbf{w}_k) \mathbf{w}_k^T \\ &= A_{k-1} - \mathbf{w}_k \mathbf{q}^T - \mathbf{q} \mathbf{w}_k^T \end{aligned} \quad (73)$$

とできます。

直交変換 R は

$$R = R_1 R_2 \cdots R_{n-1} \quad (74)$$

で与えられます。

3 演習問題

10×10 行列

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ 1 & 2 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 \\ 1 & 2 & 3 & 4 & 4 & 4 & 4 & 4 & 4 & 4 \\ 1 & 2 & 3 & 4 & 5 & 5 & 5 & 5 & 5 & 5 \\ 1 & 2 & 3 & 4 & 5 & 6 & 6 & 6 & 6 & 6 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 7 & 7 & 7 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 8 & 8 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 9 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \end{pmatrix} \quad (75)$$

を考える。

- (26) この行列を LU 分解せよ。
- (27) この行列を修正 Cholesky 分解せよ。
- (28) この行列を Jacobi 法で対角化せよ。
- (29) この行列を Householder 法で三重対角化せよ。