

大型行列の固有値問題*

名 取 亮**

1. はじめに

数年前までは、行列の固有値問題といえばヤコビ法とべき乗法というのが常識ようになっていた。大抵の計算センターのライブラリーには、この2つの方法によるプログラムが必ずはいっている。ヤコビ法は実対称行列の固有値をすべて計算したいときに用いられ、べき乗法は非対称行列の場合または実対称行列でも数個の固有値だけが必要な場合に用いられる。

物理的な問題に現われる行列や、統計計算における相関行列などは、特別な場合を除いてほとんど実対称行列であるので、ここでは実対称行列に話を限ることにする。

ヤコビ法^{1)~4)}は、面内回転によって絶対値最大の非対角要素をゼロにするような変換をくりかえし行なって、しだいに対角行列に近づけていく方法であるが、一度ゼロになった要素がその後の変換によってゼロでなくなるために反復的に行なわなければならない。 $n \times n$ の行列の場合 $N = n(n-1)/2$ とすると、大体5N回の交換が必要となり、計算量は $10n^3$ 回の乗算と $5n(n-1)/2$ 回の平方根の計算を要する。そのうえヤコビ法は行列が疎であっても、その性質を有効に利用することができない。実際、10~20次元くらいまでの比較的小さい行列に対しては、ヤコビ法は縮退などの悪条件にも強いので有力な方法であるが、次元が大きくなると計算時間が非常にかかる。HITAC 5020 Eで150次元の行列の場合、約1時間を要する。したがって、大型の行列に対してはヤコビ法を用いるよりはハウスホルダー法などによって行列をいったん三重対角行列に変換する方法がすぐれている。これらの方法は、有限回で完結するので、ヤコビ法にくらべてずっと計算量が少なくすむ。この場合、三重対角行列の固有値、固有ベクトルを求めることが必要となるが、これは二分法、逆反復法、QR法などによって比較的簡単に計算することができる。特に大型行列になると、三重対角化に要する時間にくらべて、三重対角行

列の固有値、固有ベクトルの計算のための時間は無視できるほど小さくなる。これらの方法によって400次元くらいまでならHITAC 5020 Eで1時間以内で解けるようになると思われる。もちろんこの場合は、コアメモリーだけでは足りないので、磁気ドラム、磁気ディスクなどの補助メモリーが必要となる。今後電子計算機の性能があがり、記憶容量も大きくなると、もっと大きな行列の固有値問題も解くことができるようになるであろう。これまで述べてきたのは、主として統計計算における相関行列などのように密な行列の場合であるが、有限要素法による構造解析などに現われる行列は行列の要素の大部分がゼロで、いわゆる疎な行列(Sparse matrix)であるので、この性質を有効に利用して計算量および記憶容量を減らすことができる方法を考えなければならない。そうすれば数千次元の行列の固有値問題も解くことが可能となる。記憶容量についていえば、行列をそのままの形で記憶するのではなく、非ゼロ要素だけを記憶しておくようにすればよい。また、この種の問題ではすべての固有値が必要なわけではなく、大抵は大きいほうあるいは小さいほうから数個ないし十数個の固有値だけを計算すればよいのである。大型の疎な行列の固有値問題には、普通べき乗法が用いられているが、べき乗法の欠点は1回の計算で1個の固有値しか計算できないので計算量が、必要な固有値の個数に比例して多くなることである。また、誤差の蓄積のため、あまりたくさん固有値は計算できない。

ランチョス法による三重対角化は、疎であることを利用できるのも魅力的である。しかし、誤差の蓄積を防ぐために再直化が必要になり、そのために計算量が多くなってしまうのが欠点であった。再直化をしないランチョス法によって途中まで三重対角化することによって数個ないし十数個の固有値を精度よく計算する方法を4.で紹介する。

2.では、密な大型行列の固有値問題に対して、行列を三重対角化するための方法を、3.では三重対角化された行列の固有値、固有ベクトルを計算する方法について述べる。

* Eigenvalue Problem of a Large Matrix, by Makoto NATORI
(The Computer Centre, University of Tokyo)

** 東京大学大型計算機センター

2. 行列を三重対角にする方法

行列をいきなり対角化するのではなく、いったん三重対角行列に変換する方法がいくつかある。ここでは、ギヴンス法とハウスホルダー法について説明する。

2.1 ギヴンス法^{5),6)}

ギヴンス法は、ヤコビ法と同様に面内回転をくりかえすことによって行列を三重対角に変換する方法である。すなわち、 (p, q) を $(2, 3), (2, 4), \dots, (2, n), (3, 4), \dots, (3, n), \dots, (n-1, n)$ の順に動かして $(1, 3), (1, 4), \dots, (2, 4), \dots, (2, n), \dots, (n-2, n)$ の要素を次にゼロにしていく。一度ゼロになった要素は、後の変換でもゼロのままであるので $(n-1)(n-2)/2$ 回で三重対角に変換できる。 r 回の変換をした後の行列 $A^{(r)}$ に (p, q) 平面内の回転 R をほどこして $a_{p-1, q}^{(r+1)}$ がゼロになるように変換するためには、

$$\tan \theta = a_{p-1, q}^{(r)} / a_{p-1, p}^{(r)} \quad (1)$$

のように回転角をえらばよい。

この方法では、 $4n^3/3$ 回の乗算と、 $(n-1)(n-2)/2$ 回の平方根の計算によって三重対角にすることができる。

2.2 ハウスホルダー法^{7)~11)}

ギヴンス法は非三重対角要素を1回の変換で1つずつゼロにしていくやり方であったが、ハウスホルダー法では1行1列ずつまとめてゼロにしていくので $(n-2)$ 回の変換で三重対角に変換することができる。 r 回の変換をした後の行列を $A^{(r)}$ とすると

$$A^{(r+1)} = P^{(r)} A^{(r)} P^{(r)} \quad (2)$$

によって $A^{(r+1)}$ が計算される。ここで変換行列 $P^{(r)}$ は、

$$P^{(r)} = I - \mathbf{u}^{(r)}(\mathbf{u}^{(r)})^T / h_r \quad (3)$$

と書ける。ここで

$$(\mathbf{u}^{(r)})^T = (0, \dots, 0, a_{r+1, r}^{(r)} + \operatorname{sgn}(a_{r+1, r}^{(r)})s_r, a_{r+2, r}^{(r)}, \dots, a_{nr}^{(r)}) \quad (4)$$

$$s_r^2 = \sum_{i=r+1}^n (a_{ir}^{(r)})^2 \quad (5)$$

$$h_r = s_r^2 + |a_{r+1, r}^{(r)}| s_r \quad (6)$$

である。実際には行列の対称性を利用して次のように計算したほうがよい。

$$\begin{aligned} A^{(r+1)} &= P^{(r)} A^{(r)} P^{(r)} \\ &= A^{(r)} - \mathbf{u}^{(r)}(\mathbf{u}^{(r)})^T A^{(r)} / h_r - A \mathbf{u}^{(r)}(\mathbf{u}^{(r)})^T \\ &\quad \div h_r + \mathbf{u}^{(r)}(\mathbf{u}^{(r)})^T A^{(r)} \mathbf{u}^{(r)}(\mathbf{u}^{(r)})^T / h_r^2 \end{aligned} \quad (7)$$

ここで

$$\mathbf{p}^{(r)} = A^{(r)} \mathbf{u}^{(r)} / h_r \quad (8)$$

$$\alpha_r = (\mathbf{u}^{(r)})^T \mathbf{p}^{(r)} \quad (9)$$

とすると

$$\begin{aligned} A^{(r+1)} &= A - \mathbf{u}^{(r)}(\mathbf{p}^{(r)})^T - \mathbf{p}^{(r)}(\mathbf{u}^{(r)})^T \\ &\quad + \alpha_r \mathbf{u}^{(r)}(\mathbf{u}^{(r)})^T / h_r \end{aligned} \quad (10)$$

となる。さらに

$$\mathbf{q}^{(r)} = \mathbf{p}^{(r)} - \frac{1}{2} \alpha_r \mathbf{u}^{(r)} / h_r \quad (11)$$

を用いると

$$A^{(r+1)} = A^{(r)} - \mathbf{u}^{(r)}(\mathbf{q}^{(r)})^T - \mathbf{q}^{(r)}(\mathbf{u}^{(r)})^T \quad (12)$$

となる。これから、 $A^{(r)}$ が対称行列なら $A^{(r+1)}$ も対称となることがわかる。

ハウスホルダー法は、 $2n^3/3$ 回の乗算と $(n-2)$ 回の平方根の計算を必要とする。したがって、ギヴンス法の約半分の計算量で三重対角化が行なえるので、密な行列の場合ハウスホルダー法が最もすぐれた方法といえる。

ハウスホルダー法で三重対角化し、その三重対角行列から 3.3 で述べる QR 法を用いて固有値、固有ベクトルを求めた場合と、ヤコビ法で計算した場合との計算時間を比較してみると、第 1 表のようになる¹²⁾。三重対角行列の固有値、固有ベクトルを求める方法には QR 法以外のものもあるが、計算時間はどれを用いても大差はない。ヤコビ法では、収束判定に用いる ϵ の値によって計算時間がかわるが、ここでは

$$\epsilon = \|A\|_{\infty} \times 1.5 \times 10^{-7} \quad (13)$$

とした。ただし、

$$\|A\|_{\infty} = \max_i \left(\sum_{j=1}^n |a_{ij}| \right) \quad (14)$$

である。第 1 表を見ればわかるように、8 次元以上になるとハウスホルダー法のほうが時間が短くなる。

100 次元以上になると、ハウスホルダー法はヤコビ法の $1/16 \sim 1/18$ の時間で計算できることがわかる。HITAC 5020 E の場合、コアメモリー内だけで計算でき

第 1 表 ヤコビ法とハウスホルダー法の計算時間の比較 (単位: 秒, HITAC 5020 E による)

次元数	ヤコビ法	ハウスホルダー法
4	0.004	0.014
5	0.027	0.033
8	0.114	0.092
10	0.271	0.060
50	98.560	8.849
100	1065.917	60.308
140	2367.461	150.876

る最大の次元数は約 140 次元であるが、そのときヤコビ法では 40 分かかるものが、ハウスホルダー法ではわずか 2 分 30 秒でできる。

3. 三重対角行列の固有値と固有ベクトル

3.1 二分法^{5), 6), 13), 14)}

ギブンス法やハウスホルダー法で得られた三重対角行列の固有値を求めるための非常に精度のよい方法に二分法がある。三重対角行列を C として

$$c_{ii} = \alpha_i \quad (15)$$

$$c_{i, i+1} = c_{i+1, i} = \beta_i \quad (16)$$

とすると

$$C = \begin{pmatrix} \alpha_1 & \beta_1 & & & 0 \\ & \beta_1 & \alpha_2 & \beta_2 & \\ & & \ddots & \ddots & \ddots \\ 0 & & & \beta_{n-1} & \alpha_n \end{pmatrix} \quad (17)$$

となる。 $\beta_1, \beta_2, \dots, \beta_{n-1}$ の中にゼロに等しいものがあると、 C はいくつかの三重対角行列に分かれるので C の固有値は、それぞれの行列の固有値問題を解くことによって得られるから、 $\beta_i \neq 0 (i=1, 2, \dots, n-1)$ と仮定しても一般性を失わない。固有値 λ は

$$\begin{vmatrix} \alpha_1 - \lambda & \beta_1 & & & \\ & \beta_1 & \alpha_2 - \lambda & \beta_2 & \\ & & \ddots & \ddots & \ddots \\ & & & \beta_{n-1} & \alpha_n - \lambda \end{vmatrix} = 0 \quad (18)$$

の根として求められる。この行列式の左上から順にとった主小行列式を $f_i(\lambda)$ とすると

$$\left. \begin{aligned} f_0(\lambda) &= 1 \\ f_1(\lambda) &= \alpha_1 - \lambda \\ f_i(\lambda) &= (\alpha_i - \lambda)f_{i-1}(\lambda) - \beta_{i-1}^2 f_{i-2}(\lambda) \end{aligned} \right\} \quad (19)$$

となる。この関数列 $\{f_i(\lambda)\}$ の性質として

$$\left. \begin{aligned} f_i(\lambda_0) &= 0 \text{ のとき} \\ f_{i+1}(\lambda_0)f_{i-1}(\lambda_0) &< 0 \quad (i=1, 2, \dots, n-1) \end{aligned} \right\} \quad (20)$$

となる。なぜなら(19)式から

$$f_{i+1}(\lambda_0) = (\alpha_{i+1} - \lambda_0)f_i(\lambda_0) - \beta_i^2 f_{i-1}(\lambda_0) \quad (21)$$

であるから、 $f_i(\lambda_0) = 0$ とすると

$$f_{i+1}(\lambda_0)f_{i-1}(\lambda_0) = -\beta_i^2 \{f_{i-1}(\lambda_0)\}^2 \leq 0$$

ここで、等号が成り立つことがあるかどうかしらべる。 $\beta_i \neq 0$ であるから、 $f_{i-1}(\lambda_0)$ がゼロになるかどうかしらべればよい。もし $f_{i-1}(\lambda_0) = 0$ になったとすると(21)から $f_{i+1}(\lambda_0) = 0$ となり、すべての $f_i(\lambda_0)$ がゼ

ロにならなければならない。ところが $f_0(\lambda_0) = 1$ であるからこのようなことはない。したがって

$$f_{i+1}(\lambda_0)f_{i-1}(\lambda_0) < 0$$

が成り立つ。この性質によって関数列 $\{f_i(\lambda)\}$ が Sturm 列になっていることがわかり、次の定理が成り立つ。「ある λ_0 に対して $f_i(\lambda_0) (i=0, 1, \dots, n)$ を計算して、隣り合う項の符号が一致する回数を $S(\lambda_0)$ とすると、 λ_0 に等しいかあるいは大きい C の固有値の個数は $S(\lambda_0)$ に等しい。ただし、 $f_i(\lambda_0) = 0$ のときは f_i の符号は $f_{i-1}(\lambda_0)$ の符号と同じであるとみなす。」

この定理を用いると、 C の固有値のうち大きいほうから k 番目の固有値 λ_k を求めるときには、 $S(a) \geq k$, $S(b) < k$ となる a, b をえらべば、 λ_k は区間 (a, b) の間にあることがわかる。 (a, b) の中点 $c = (a+b)/2$ に対して $S(c)$ を計算し、 $S(c) \geq k$ ならば λ_k は区間 (c, b) の中、そうでなければ区間 (a, c) の中にある。これをくりかえしていけば、 p 回で λ_k の存在範囲を $2^{-p}(b-a)$ の幅にまでせまることができる。 C の固有値は、すべて $|\lambda| \leq \|C\|$ であることが知られているから

$$x = \max_{i=2}^{n-1} (|\alpha_i| + |\beta_i|, (|\beta_{i-1}| + |\alpha_i| + |\beta_i|), (|\beta_{n-1}| + |\alpha_n|)) \quad (22)$$

とすれば、すべての固有値は $(-x, x)$ の区間の中にあるから、この区間から出発すれば任意の固有値を求めることができる。

$n=3$ の例を考える。行列 C を

$$C = \begin{pmatrix} 4 & 1 & 0 \\ 1 & 3 & 1 \\ 0 & 1 & 2 \end{pmatrix}$$

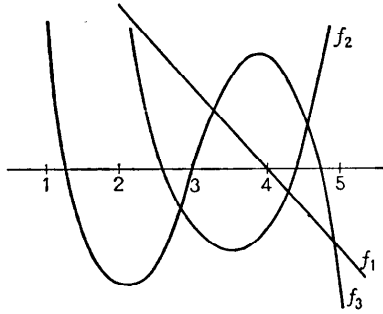
とすると

$$\left. \begin{aligned} f_0(\lambda) &= 1 \\ f_1(\lambda) &= 4 - \lambda \\ f_2(\lambda) &= (3 - \lambda)f_1(\lambda) - 1 \\ f_3(\lambda) &= (2 - \lambda)f_2(\lambda) - f_1(\lambda) \end{aligned} \right\}$$

となり、 $\lambda=1, 2, 4, 5$ に対して符号をしらべると第 2 表のようになり、 $(1, 2)$, $(2, 4)$, $(4, 5)$ の間にそれぞれ 1

第 2 表 $f_0 \sim f_3$ の符号の変化

λ	1	2	4	5
f_0	+	+	+	+
f_1	+	+	0	-
f_2	+	+	-	+
f_3	+	-	+	-
S	3	2	1	0



第1図 $f_1 \sim f_3$ の概形

つつ固有値が存在することがわかる。各関数の概形は第1図のようになる。

4.2 逆反復法^{15)~17)}

二分法などで固有値を計算したあとで、固有ベクトルを求めるためのすぐれた方法として逆反復法がある。これは λ_k の近似値 $\bar{\lambda}_k$ がわかっているときに、

$$(C - \bar{\lambda}_k I) \mathbf{u}^{(r+1)} = \mathbf{u}^{(r)} \tag{23}$$

によって、 $\mathbf{u}^{(r+1)}$ を $\mathbf{u}^{(r)}$ を右辺とする連立方程式を解いて求める。初期ベクトル $\mathbf{u}^{(0)}$ は任意にえらばれる。 $\mathbf{u}^{(0)}$ を C の固有ベクトル \mathbf{x}_i で展開して

$$\mathbf{u}^{(0)} = \sum_{i=1}^n a_i \mathbf{x}_i$$

とすると

$$\begin{aligned} \mathbf{u}^{(r)} &= (C - \bar{\lambda}_k I)^{-r} \mathbf{u}^{(0)} \\ &= \sum_{i=1}^n \frac{a_i}{(\lambda_i - \bar{\lambda}_k)^r} \mathbf{x}_i \\ &= \frac{1}{(\lambda_k - \bar{\lambda}_k)^r} \left[a_k \mathbf{x}_k + \sum_{i \neq k} \left(\frac{\lambda_k - \bar{\lambda}_k}{\lambda_i - \bar{\lambda}_k} \right)^r a_i \mathbf{x}_i \right] \end{aligned} \tag{24}$$

ここで

$$\left| \frac{\lambda_k - \bar{\lambda}_k}{\lambda_i - \bar{\lambda}_k} \right| \ll 1$$

であるから、大抵の場合2回の反復で \mathbf{x}_k の非常によい近似値が得られる。

$\bar{\lambda}_k$ が λ_k に近ければ $(C - \bar{\lambda}_k I)$ はほとんど特異になるが、このことは結果の精度には影響がない。解かれた $\mathbf{u}^{(r)}$ は方程式の解としては精度がおちるが、ほとんど正確に正しい解の定数倍になっていることが示されている¹⁸⁾。固有ベクトルの場合には定数倍は問題にならない。

$(C - \lambda I)$ を係数行列とする連立方程式を一連の右辺に対して解くためには

$$C - \lambda I = LR \tag{25}$$

と分解する。ここで L は左下三角行列、 R は右上三

角行列で、 C が三重対角行列の場合には

$$L = \begin{bmatrix} * & & & 0 \\ * & \dots & & \\ 0 & & ** & * \\ & & & \dots \\ 0 & & & & * \end{bmatrix}$$

$$R = \begin{bmatrix} * & * & * & 0 \\ & \dots & & \\ & & \dots & * \\ & & & \dots & * \\ 0 & & & & * \end{bmatrix}$$

となることがわかる。 LR 分解はガウスの消去法と同じである。 LR 分解ができると方程式は

$$LR \mathbf{u}^{(r+1)} = \mathbf{u}^{(r)} \tag{26}$$

の形にかけて、前向き代入と後向き代入で解くことができる。計算の安定性のために軸交換を用いれば、置換行列 P によって

$$P(C - \lambda I) = LR \tag{27}$$

と分解することになり

$$LR \mathbf{u}^{(r+1)} = P \mathbf{u}^{(r)} \tag{28}$$

を解くことになる。 $P \mathbf{u}^{(r)}$ は $\mathbf{u}^{(r)}$ の要素の順序をかえたものである。

実際に計算する場合には(28)から

$$R \mathbf{u}^{(1)} = L^{-1} P \mathbf{u}^{(0)} \tag{29}$$

で、 $\mathbf{u}^{(0)}$ は任意にえらべるのだから

$$L^{-1} P \mathbf{u}^{(0)} = (1, 1, \dots, 1)^T \tag{30}$$

となるような $\mathbf{u}^{(0)}$ をえらんだとすると、第1回の反復では前向き代入は省略できて、後向き代入だけで $\mathbf{u}^{(1)}$ を求めることができる。第2回の反復

$$LR \mathbf{u}^{(2)} = P \mathbf{u}^{(1)} \tag{31}$$

で得られる $\mathbf{u}^{(2)}$ は普通非常に精度のよい固有値になっている。したがって、わずかに1回半の反復だけで固有ベクトルの計算ができるので、固有値が求まったあとで逆反復法を用いるのは、固有ベクトルを計算するには非常に経済的な方法である。

3.3 QR法^{19),20)}

QR法は、行列 A にユニタリ変換をほどこすことによって、対角に固有値 λ_i をもつ右上三角行列に近づける方法である。すなわち、 r 回の変換の後の行列を $A^{(r)}$ とすると

$$A^{(r)} = Q^{(r)} R^{(r)} \tag{32}$$

$$A^{(r+1)} = R^{(r)} Q^{(r)} \tag{33}$$

ここで $Q^{(r)}$ はユニタリ行列、 $R^{(r)}$ は右上三角行列である。

$$|\lambda_1| > |\lambda_2| > \dots > |\lambda_n|$$

とすると、 $r \rightarrow \infty$ としたとき

$$a_{nn}^{(r)} = \lambda_n + O(\rho_n^r) \tag{34}$$

$$a_{n,n-1}^{(r)} = 0(\rho_n) \quad (35)$$

となる。ただし

$$\rho_n = |\lambda_n / \lambda_{n-1}| \quad (36)$$

である。この収束を速めるために原点の移動を行なう。

$$A^{(r)} - s^{(r)}I = Q^{(r)}R^{(r)} \quad (37)$$

$$R^{(r)}Q^{(r)} + s^{(r)}I = A^{(r+1)} \quad (38)$$

これによって収束率は $|(\lambda_n - s^{(r)}) / (\lambda_{n-1} - s^{(r)})|$ となり、 $s^{(r)}$ を λ_n に近く選べば、収束を速めることができる。 r が大きくなれば $a_{nn}^{(r)}$ が λ_n に近づくと $a_{nn}^{(r)}$ が $s^{(r)}$ のよい値となる。

このようにして $a_{n,n-1}^{(r)}$ が無視できるほど小さくなれば、 $a_{nn}^{(r)}$ が固有値となり残りの固有値は $A^{(r)}$ の n 行 n 列をとり、のぞいた $(n-1)$ 次元の行列の固有値を計算して得られる。したがって、固有値が1個見つかるたびに行列の次数を1つずつ小さくすることができる。特に A が三重対角になっている場合には QR 法に要する計算量は非常に少なくなる。

4. 大型の疎な行列の固有値問題

4.1 べき乗法²¹⁾

行列が疎である場合には、その性質をできるだけ有効に使うために、行列自身は変形せず、もとの行列とベクトルとの積の計算をくりかえし行なうことによって固有値、固有ベクトルを求めるべき乗法が有利である。べき乗法は、初期ベクトル $u^{(0)}$ を任意にえらんで

$$u^{(r+1)} = Au^{(r)} \quad (39)$$

によって次々にベクトルを計算していく方法である。各ベクトルを規格化しておけば、 $r \rightarrow \infty$ で $u^{(r)} \rightarrow x_1$ となる。ここで x_1 は絶対値最大の固有値 λ_1 に対応する固有ベクトルである。この場合行列 A をそのまま記憶しておく必要はなく、あるベクトルが与えられたときに、行列 A とそのベクトルをかけて新しいベクトルを作るサブルーチンを作っておけばよい。このサブルーチンでは、行列 A の性質を有効に使って計算量と記憶容量を減らすようなくふうしておくのである。

λ_1 と x_1 が求まったとき、次に大きい固有値 λ_2 とそれに対応する固有ベクトル x_2 を求めるためには

$$A' = A - \lambda_1 x_1 x_1^T \quad (40)$$

を計算し、 A' に対してべき乗法を適用すればよいのであるが、そうすると A' はもはや疎ではなくなるのでべき乗法の利点が失われてしまう。もとの行列 A をそのまま用いるためには、初期ベクトルとして x_1 に

直交するようなベクトルを選んで、べき乗法を適用することが考えられる。そうすれば今度は x_2 に収束するはずである。次には x_1 と x_2 の両方に直交するような初期ベクトルを選ぶ。以下同じようにして、次々に固有値、固有ベクトルが計算される。ここで注意しなければならないのは、先にいくに従って計算誤差が集積してくるために、 x_1, x_2, \dots, x_{i-1} に直交する初期ベクトルから出発して計算していくと、 x_i に収束するようにみえるが、さらに精度をあげようとして計算を進めるうちに x_1, x_2, \dots, x_{i-1} のうちのどれかと同じものに収束してしまうことがある。このようなときには、 x_i の精度は、ある程度以上にあげることはできないのである。したがって、先にいくほど収束の条件をゆるめてやらないと同じベクトルが出てきてしまうことがある。

4.2 ランチョス法と縮小された三重対角行列

行列を三重対角にする方法のうち、ギブンス法とハウスホルダー法についてはすでに述べたが、これらはいずれも行列 A を逐次変換していく方法であった。それに対して A を変換していくのではなく、ベクトルに A をかける演算をくりかえすことによって三重対角行列の要素を計算するのがランチョス法²²⁾である。 A をそのまま用いて計算する点で、これはべき乗法に類似しているが、べき乗法が反復的であったのに対してランチョス法では n 回で三重対角化が完了する。三重対角行列を(17)式のようなものとする、ランチョス法の手続きは次のようになる。まず任意の規格化された初期ベクトル e_1 から始める

$$\left. \begin{aligned} Ac_1 &= \alpha_1 c_1 + \beta_1 c_2 \\ Ac_2 &= \beta_1 c_1 + \alpha_2 c_2 + \beta_2 c_3 \\ &\dots \\ Ac_i &= \beta_{i-1} c_{i-1} + \alpha_i c_i + \beta_i c_{i+1} \\ &\dots \\ Ac_n &= \beta_{n-1} c_{n-1} + \alpha_n c_n \end{aligned} \right\} \quad (41)$$

ここで

$$\alpha_i = c_i^T A c_i \quad (42)$$

であり、 β_i は c_{i+1} の規格化条件

$$\|c_{i+1}\| = 1 \quad (43)$$

から定められる。初期ベクトルとしては

$$c_1^T = (1, 0, \dots, 0)$$

とするのが普通である。計算に誤差が含まれなければ、

$$c_i^T c_j = 0 \quad (i \neq j) \quad (44)$$

となり、 c_i は互に直交する。したがって、 n 個の独立

なベクトル c_1, c_2, \dots, c_n が求まったところで計算が完了する。しかし、計算誤差の蓄積によって基底ベクトル c_i の間の直交性がくずれてくるので、普通ランチョス法では各回で再直交化を行なっている。そのため計算量がふえてしまうのが欠点であった。

二分法で述べたように、主小行列式 $f_m(\lambda)$ は Sturm 列をなしているの、その性質から $f_m(\lambda)=0$ の根の間には必ず1つ $f_{m-1}(\lambda)=0$ の根が存在する。三重対角行列を m 次で打ち切ったものを考えると、その固有値は $f_m(\lambda)=0$ の根となる。これを大きい順に $\eta_1^{(m)}, \eta_2^{(m)}, \dots, \eta_m^{(m)}$ とすると上述の性質から

$$\eta_{i+1}^{(m)} < \eta_i^{(m-1)} < \eta_i^{(m)} \quad (45)$$

なる不等式が成り立つ。行列 C の固有値、すなわち $f_n(\lambda)=0$ の根を大きい順に $\lambda_1, \lambda_2, \dots, \lambda_n$ とすると、

$$\left. \begin{aligned} \eta_1^{(1)} < \eta_1^{(2)} < \dots < \eta_1^{(m)} < \dots < \lambda_1 \\ \eta_2^{(2)} < \dots < \eta_2^{(m)} < \dots < \lambda_2 \\ \dots \end{aligned} \right\} (46)$$

$$\left. \begin{aligned} \lambda_n < \dots < \eta_m^{(m)} < \dots < \eta_2^{(2)} < \eta_1^{(1)} \\ \lambda_{n-1} < \dots < \eta_{m-1}^{(m)} < \dots < \eta_1^{(2)} \\ \dots \end{aligned} \right\} (47)$$

となって、 C の固有値のうち大きい方から数個と小さい方から数個の固有値は、適当な m で打ち切った三重対角行列の固有値で近似できることが期待される。実際、物理の問題に対して再直交化を行なうランチョス法によって三重対角化した行列を途中で打ち切って計算した固有値がよい近似を与えることが示されている²³⁾。

A の固有ベクトル x を基底ベクトル c_i で展開すると、

$$x = \sum_{i=1}^n u_i c_i \quad (48)$$

これを $Ax = \lambda x$ に代入して(41)を用いると、

$$\beta_{i-1} u_{i-1} + \alpha_i u_i + \beta_i u_{i+1} = \lambda u_i \quad (49)$$

となり

$$u = (u_1, u_2, \dots, u_n)^T \quad (50)$$

は三重対角行列 C の固有ベクトルになっていることがわかる。もし $i > m$ のときに u_i が無視できるほど小さくなっていけば、 x は c_1, c_2, \dots, c_m だけで表わすことができるから三重対角行列を m 次で打ち切っても正しい固有値、固有ベクトルが求まる。ここで注意すべきことは(41)の関係をただただ、 c_i の直交性は直接用いていないことである。したがって各回に再直交化の手続きをしなくても(48)の展開が可能であれば正しい結果が得られるはずである。再直交化をし

なくてすむならば、特に次元の大きい疎な行列の場合には A とベクトルとの積の計算をくふうすれば、ランチョス法に必要な計算量を大幅に減らすことができる。しかも三重対角化を適当なところで打ち切ることによって、その後の固有値、固有ベクトルの計算は次元の小さい行列に対して行えばよく、それによっても計算量と記憶容量が節約できる²⁴⁾。

例として、へりを固定した長方形の膜の振動の問題を考えることにする。

$$-\Delta u = \lambda u \quad (51)$$

ここで、 30×40 の格子に切って、1,200個の内点に対して差分方程式をたてる。第2図のように格子点を番号をつけると

$$4u_i - u_{i+1} - u_{i-1} - u_{i+40} - u_{i-40} = \lambda u_i \quad (52)$$

となる。この1,200次元の行列を再直交化をしないランチョス法で300次元まで三重対角化し、その後で50, 100, 150, 200, 250, 300の各次元で打ち切った三重対角行列に対して固有値を二分法で大きいほうから32個ずつ計算した。結果は第3表に示すとおり、かなり小さな次元で打ち切っても固有値は精度よく求められることがわかる。計算時間は HITAC 5020 E で行なった場合、300次元までの三重対角化に52.3秒、それぞれの次元の行列の固有値を求めるための時間(第3表)も合わせる約96秒であった(計算は単精度、10進7けた)。この例のように三重対角行列を打ち切る次元数をしだいに大きくしていきながら、固有値の収束の状況をみていくことによって、必要な個
第3表 縮小された三重対角行列による固有値の計算 (HITAC 5020 E による)

打ち切った次元数	5けた以上正確な固有値の個数	時間(sec)
50	0	2.8
100	3	5.0
150	5	6.9
200	11	8.5
250	14	9.9
300	14	10.6

1	2	3	4	38	39	40
41	42	43	44	78	79	80
						1200

第2図 格子点の番号

数の固有値を必要な精度で求めることができる。打ち切る次元数を大きくしていくと、本来縮退していないはずの固有値が複数個出てくることがあるが、これは e_i の直交性が失われてくるためではない。

第3表にあるように、32個ずつ二分法で固有値を求めるための時間はわずかである。しかも、いったん三重対角化したあとであるから問題の行列によって時間がかわることはない。それに対して、ランチョス法による三重対角化に要する時間は、問題の行列によって大幅にかわる可能性がある。三重対角化を打ち切る次元を m とし、問題の行列とベクトルとの積を計算するための時間を t_0 とすると、三重対角化に要する時間は大体 mt_0 できる。この例では、1行の中でゼロでない要素の数は高々5個で、要素の位置も規則的であるので行列とベクトルの積を非常に簡単に計算できたのである。

6. おわりに

大型行列の固有値問題を中心にして、固有値問題のいろいろな解法について述べてきた。ここでは実対称行列に話を限ったが、これらの方法の中には実非対称行列、複素エルミート行列、複素一般行列に対しても適用できるものが少なくない。それらの問題については、多くのすぐれた解説書^{25)~28)}があるのでそれを参照していただきたい。また、これらの方法による FORTRAN プログラムが出ている本^{29), 30)}もあるので、プログラム・テクニックのこまかい点についてはそれらにゆずる。

参考文献

- 1) C. G. J. Jacobi: "Über ein leichtes Verfahren die in der Theorie der Säkularstörungen vorkommenden Gleichungen numerisch aufzulösen", *J. reine und angew. Math.* **30** (1846) pp. 51~94.
- 2) A. Schönhage, "Zur Konvergenz des Jacobi-Verfahrens," *Num. Math.* **3** (1961) pp. 374~380.
- 3) A. Schönhage, "Zur quadratischen Konvergenz des Jacobi-Verfahrens," *Num. Math.* **6** (1965) pp. 410~412.
- 4) D. A. Pope and C. Tompkins, "Maximizing functions of rotations—experiments concerning speed of diagonalization of symmetric matrices using Jacobi's method", *J. ACM* **4** (1957) pp. 459~466.
- 5) W. Givens, "A method of computing eigenvalues and eigenvectors suggested by classical results on symmetric matrices", *Nat. Bur. Stand. Appl. Math. Ser.* **29** (1953) pp. 117~112.
- 6) W. Givens, "Numerical computation of the characteristic values of a real symmetric matrix", *Oak Ridge Nat. Lab. Report, ORNL-1574* (1954).
- 7) A. S. Householder and F. L. Bauer, "On certain methods for expanding the characteristic polynomial", *Num. Math.* **1** (1959) pp. 29~37.
- 8) F. L. Bauer, "Sequential Reduction to Tridiagonal Form", *J. Soc. Indust. Appl. Math.* **7** (1959) pp. 107~113.
- 9) J. H. Wilkinson, "Householder's Method for the Solution of the Algebraic Eigenproblem", *Comput. J.* **3** (1960) pp. 23~27.
- 10) J. H. Wilkinson, "Householder's Method for Symmetric Matrices", *Num. Math.* **4** (1962) pp. 354~361.
- 11) J. M. Ortega, "An Error Analysis of Householder's Method for the symmetric Eigenvalue Problem," *Num. Math.* **5** (1963) pp. 211~225.
- 12) 名取, 栃木, 国井: "ヤコビ法とハウスホルダー法の比較", *東京大学大型計算機センター・ニュース*, Vol. 2, No. 9 (1970) pp. 39~45.
- 13) J. M. Ortega, "On Sturm Sequences for Tridiagonal Matrices", *J. ACM*, **7**(1960) pp. 260~263.
- 14) J. H. Wilkinson, "Calculation of the Eigenvalues of a Symmetric Tridiagonal Matrix by the Method of Bisection", *Num. Math.* **4** (1962) pp. 362~367.
- 15) H. Wielandt, "Bestimmung höherer Eigenwerte durch gebrochene Iteration", *Ber. B 44/J/37 der Aerodynamischen Versuchsanstalt, Göttingen*, 1944.
- 16) J. H. Wilkinson, "The Calculation of the Eigenvectors of Codiagonal Matrices", *Comput. J.*, **1** (1958) pp. 148~152.
- 17) J. H. Wilkinson, "Calculation of the Eigenvectors of a Symmetric Tridiagonal Matrix by Inverse Iteration", *Num. Math.* **4** (1962) pp. 368~376.
- 18) J. H. Wilkinson, "Rounding Errors in Algebraic Processes Notes on Applied Science", *Prentice-Hall, New Jersey*, 1963.
- 19) J. G. F. Francis, "The QR transformation, Parts I and II," *Comput. J.* **4** (1961/62) pp. 265~271 and 332~345.
- 20) J. M. Ortega and H. F. Kaiser, "On the LL^T and QR Methods for Symmetric Tridiagonal Matrices", *Comput. J.* **6** (1963) pp. 99~101.

- 21) J. H. Wilkinson, "The calculation of the latent roots and vectors of matrices on the pilot model of ACE", Proc. Camb. Phil. Soc. **50** (1954) pp. 536~566.
 - 22) C. Lanczos, "An iteration method for the solution of the eigenvalue problem of linear differential and integral operators", J. Res. Nat. Bur Stand. **45** (1950) pp. 255~282.
 - 23) T. Sebe and J. Nachamkin, "Variational Buildup of Nuclear Shell Model Bases", Annal. Phys. **51** (1969) pp. 100~123.
 - 24) 高橋, 名取, "大次元の sparse matrix の固有値問題の解法", 情報処理学会, Dec. 1970.
 - 25) J. H. Wilkinson, "The Algebraic Eigenvalue Problem", Oxford university Press, London, 1965.
 - 26) A. S. Householder, "The Theory of Matrices in Numerical Analysis", Blaisdell, New York, 1964.
 - 27) 新谷, "数値計算 I" 朝倉書店, 1967.
 - 28) J. ウォルシュ編, 高須監訳, "数値解析概論", 日本評論社, 1970.
 - 29) A. Ralston and H. Wilf (ed.) "Mathematical Methods for Digital Computers," Vol. 1 and Vol. 2, John Wiley and Sons, 1960 and 1967.
 - 30) 磯田, 大野監修, "FORTRAN による数値計算ハンドブック", オーム社, 1971.
-