数理解析研究所講究録
第190巻 1973年 46-56

大型の一般固有値問題の解法

東大 大型計算機センター 名取 亮

§1. 序

構造物の振動解析などに現われる一般固有値問題

$$Av = \lambda Bv$$

を考える。ここで，

A: n次元，対称，正定値行列，バンド幅 $(2m_A+1)$

B: 〃 〃 〃 〃 $(2m_B+1)$

または，対角，非負定値行列

λ: すべて正

である。計算の目的は

(1) 小さい方から数個のλとvを求めること.

(2) $\mu_B < \lambda < \mu_T$ の範囲にあるλとvを求めること.

のどちらかである。

解法として，K-J. Bathe の thesis[1] にある2つの方法を紹介する。

(1) determinant search method

3角分解（行列式の計算）と inverse iteration を効率よく組合せた方法

(2) subspace iteration algorithm

simultaneous iteration の拡張

A，B のバンド幅が小さいときは（1）の方法，バンド幅が大きいときは（2）の方法が有効である。

§2. determinant search method

2.1 序

determinant search method は以前から知られているが，大型の行列の場合には，この方法だけで固有値を精度よく求めようとすると3角分解の回数が多くなり，バンド幅が非常に小さい時以外は計算量が多く実用にならない。ここで紹介する方法は，determinant search で固有値の近似値を求め，それを出発値として inverse iteration を行なう方法である。

inverse iteration では，3角分解を行なうから，特性多項式

$$p(\mu) = \det(A - \mu B)$$

の値を計算できる。

$$LDL^T = (A - \mu_k B)$$

と分解したとすると，

$$p(\mu_k) = det(LDL^T) = \prod_{i=1}^{n} d_{ii}$$

となる。さらに，"Dの負の要素の個数が μ_k より小さな固有値の個数と一致する" ことを利用できる。

2．2　$A-\mu_k B$ の3角分解

$\mu_k > \lambda_1$（最小固有値）のときは $A-\mu_k B$ は正定値ではないから，コレスキー分解はできないが，LDL^T 分解はできる。これは，ガウスの消去法を用いて行なえる。

$$L_{n-1}^{-1} \cdots L_2^{-1} L_1^{-1} (A-\mu_k B) = U$$

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

$$DL^T = U$$

一般にpivotingを行なわないと不安定になることがあるが，pivotingを行なうとバンド幅が大きくなってしまう。ここでは，固有値の近似値を求めることが目的であるから，もし不安定がおきたら μ_k を少し大きくしてやりなおせばよい。実例では不安定はおこらなかった。

3角分解と行列式の計算に要する演算回数は，

$$\frac{1}{2}nm^2 + \frac{5}{2}nm + 2n \qquad ; \quad m = m_A = m_B$$

$$\frac{1}{2}nm^2 + \frac{3}{2}nm + 2n \qquad ; \quad m = m_A, \quad m_B = 0$$

2．3　補間スキーム，セカント法

μ_{k-1}，μ_k から μ_{k+1} を求めるには

$$\mu_{k+1} = \mu_k + \eta \cdot \frac{p(\mu_k)}{p(\mu_k) - p(\mu_{k-1})} (\mu_{k-1} - \mu_k)$$

を用いる。　$\eta = 1.0$ とすると，よく知られたセカント法になる。これでは速度が遅いので，加速するために $\eta = 2.0$ とし，μ の変化が小さいときには，2倍にしてゆく。　μ が λ をとびこえると p の符号が変るから検出できる。

その他に，ニュートン法も考えられるが $p'(\mu)$ の計算が大変である。

2．4　出発値

セカント法の出発値として λ_1 より小さな μ_1 と μ_2 が必要である。B が正定値なら $\mu_1 = 0.0$ ととればよいが，μ_2 の選び方が問題である。なるべく λ_1 に近い値をとれば反復は少なくてすむ。ここでは，$\mu_1 = 0.0$ に対して inverse iteration を行なう方法を用いる．ある程度収束したときのベクトルを x_k とすると，

$$r_k = (A - \rho(x_k) B) x_k$$

$$\mu_2 < \{ \rho(x_k) - (r_k, B^{-1} r_k)^{\frac{1}{2}} \}$$

但し，

$$\rho(x_k) = \frac{(x_k, A x_k)}{(x_k, B x_k)} \qquad (\text{レーリー商})$$

となる。実際には上式の代りに

$$\mu_2 = (1-0.01)\rho(x_k)$$

としておけばよい。

2.5 inverse iteration

shift を μ とすると，

$$(A-\mu B)\bar{x}_{k+1} = y_k$$

$$LDL^T \bar{x}_{k+1} = y_k$$

$$L^T \bar{x}_{k+1} = D^{-1}L^{-1} y_k \qquad \left(\text{第1回目は右辺} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \text{からはじめる}\right)$$

back-substitution によって $\bar{x}_{k+1}$ が求まる。次に，

$$\bar{y}_{k+1} = B\bar{x}_{k+1}$$

$$\rho^c(\bar{x}_{k+1}) = \frac{(\bar{x}_{k+1}, y_k)}{(\bar{x}_{k+1}, \bar{y}_{k+1})} \qquad (\text{固有値の補正項})$$

$$y_{k+1} = (\bar{y}_{k+1} - \alpha_1 w_1 - \cdots - \alpha_t w_t) / (\bar{x}_{k+1}, \bar{y}_{k+1})^{\frac{1}{2}}$$

ここで，

$$w_i = B v_i$$

$$\alpha_j = (\bar{y}_{k+1}, v_j)$$

これは，すでに求まっている t 個の固有ベクトル $v_1, \cdots, v_t$

との直交化を行なうためである。

演算回数は,

$$4nm + 2nt + 5n \quad ; \quad m = m_A = m_B$$

$$2nm + 2nt + 5n \quad ; \quad m = m_A , \ m_B = 0$$

inverse iteration の収束判定は, 固有値

$$\lambda_i^{(k+1)} = \mu + \rho^c(\bar{x}_{k+1}) \quad , \quad k = 1, 2, \cdots$$

の相対誤差

$$t_i^{(k+1)} = \frac{|\lambda_i^{(k+1)} - \lambda_i^{(k)}|}{\lambda_i^{(k+1)}} \quad , \quad \lambda_i^{(1)} = 0.0$$

が, ある値より小さくなったかどうかで行なう.

§3. subspace iteration algorithm

3.1 序

subspace iteration は標準形の固有値問題に対して, Bauer,[2] Jennings[3], Rutishauser[4] などによって提案された simultaneous iteration を一般固有値問題に拡張したものである。

X_k を p 個のベクトルからなる行列としたとき,

(a) $A X_{k+1} = B X_k$

inverse iteration, X_k の各列は最小固有値に対応するベクトルに収束する.

(b) $A X_{k+1} = B X_k R_{k+1}^{-1}$

R_{k+1} は上3角行列で、X_{k+1} の中のベクトルが B-orthogonal になるようにきめる。収束の rate は $\max(\lambda_{i-1}/\lambda_i, \lambda_i/\lambda_{i+1})$

(c) $A \bar{X}_{k+1} = B X_k$

$A_{k+1} = \bar{X}_{k+1}^T A \bar{X}_{k+1}$

$B_{k+1} = \bar{X}_{k+1}^T B \bar{X}_{k+1}$

$A_{k+1} Q_{k+1} = B_{k+1} Q_{k+1} \Lambda_{k+1}$

$X_{k+1} = \bar{X}_{k+1} Q_{k+1}$

収束の rate は $(\lambda_i/\lambda_{p+1})$

これらはすべて、p 次元部分空間 $\mathcal{E}_k$ による反復で、

$$\mathcal{E}_{k+1} = \{x \mid Ax = By \; ; \; y \in \mathcal{E}_k\}$$

(a)(b)(c) いずれの方法でも，部分空間としては同じものが作り出される。(c) における A_{k+1}, B_{k+1} は A, B の $\mathcal{E}_{k+1}$ への projection である。

演算回数はそれぞれ異なるから，これらを適当に組合せて用いるとよい。

3．2 一般化されたヤコビ法

projected operator に対する固有値問題を解くことを考える。

$$A v = \lambda B v$$

と書くと，A，B は p×p の full matrix である。

特徴 (1) B は非常に ill-condition になることがある。

(2) iteration vector が固有ベクトルに近づいてくると，A，B ともに対角行列に近づく．

(1)から，B をコレスキー分解して標準形に直すことはむつかしい。ここでは (2) を利用して，B を分解せずに直接解くために一般化されたヤコビ法を用いる。

ヤコビ法のときと同様に 2×2 行列を考える。

$$A = \begin{pmatrix} a_{jj} & a_{jk} \\ a_{kj} & a_{kk} \end{pmatrix}, \qquad B = \begin{pmatrix} b_{jj} & b_{jk} \\ b_{kj} & b_{kk} \end{pmatrix}$$

ここで，

$$V = \begin{pmatrix} 1 & \alpha \\ \gamma & 1 \end{pmatrix}$$

として，$V^T A V$ および $V^T B V$ が同時に対角になるように α，γ をきめる。

$$\alpha a_{jj} + (1 + \alpha\gamma) a_{jk} + \gamma a_{kk} = 0$$

$$\alpha b_{jj} + (1 + \alpha\gamma) b_{jk} + \gamma b_{kk} = 0$$

これを解くと，

$$\bar{a}_{kk} = a_{kk} b_{jk} - b_{kk} a_{jk}$$

$$\bar{a}_{jj} = a_{jj} b_{jk} - b_{jj} a_{jk}$$

$$a = a_{jj} b_{kk} - a_{kk} b_{jj}$$

として，

$$\alpha = \bar{a}_{kk} / x \quad , \quad \gamma = - \bar{a}_{jj} / x$$

となる。但し，x は

$$x^2 - a x - \bar{a}_{kk} \bar{a}_{jj} = 0$$

$$x_{1,2} = \frac{a}{2} \pm \sqrt{\frac{a^2 + 4 \bar{a}_{kk} \bar{a}_{jj}}{4}}$$

このうち絶対値の大きい方を用いる．

対角化を行なうときには自由度 i と j の間の coupling を小さくするようにすればよい。coupling の尺度としては，coupling factor，$(a^2_{ij} / a_{ii} a_{jj})^{\frac{1}{2}}$，$(b^2_{ij} / b_{ii} b_{jj})^{\frac{1}{2}}$ を用いる。反復は次のように行なう．

（1） k 回目の反復に対する threshold をきめる。

（2） $i<j$ のすべての (i, j) に対して coupling factor を計算し，どちらかが threshold より大きければ変換を行なう。

（3） 新しい固有値を計算する．（ $\lambda_i = a_{ii} / b_{ii}$ ）

（4） 前の回の固有値と比較して，差が大きければ (1) へ。収束したと判定する前に，すべての coupling factor が小さくなっているかどうかを確かめる。

3．3　初期ベクトルの選び方

X_1 の中のベクトルの選び方が重要である。これらのベクトルがすでに固有ベクトルに近ければ反復回数は少なくてすむ。固有ベクトルの近似値がわかっている場合には，それを使えばよい。（ex. dynamic optimization）　近似ベクトルが不明な場合には，第1回目の右辺，BX_1 の第1列をBの対角要素，第2列以降は b_{ii}/a_{ii} の大きいものから順に i 番目の要素を1とする単位ベクトルにすればよい。

3．4　計算上の注意

（1）部分空間の次元 p の選び方

必要とする固有ベクトルの個数を s とすると，

$$p = \min\{2s, s+8\}$$

（2）収束判定

inverse iteration を行なって，deflation をする場合とちがって有利な点は，固有値，固有ベクトルの精度が悪くても計算の安定性が保たれることである。

相対誤差

$$t_i^{(k+1)} = \frac{|\lambda_i^{(k+1)} - \lambda_i^{(k)}|}{\lambda_i^{(k+1)}} , \quad \lambda_i^{(1)} = 0.0$$

としたとき，$t_i^{(k+1)} < RTOL$ となった固有値は収束した

ものとみなす。s個の固有値がすべて収束するか，最大反復回数 NITEM の後に stop. する。 $RTOL = 10^{-6}$，NITEM = 12 位が適当である。

(3) チェック

小さい方から s 個の固有値がすべて求まったかどうかは，λ_s の少し右に shift して，3角分解し，負の対角要素の個数を数えればチェックできる。§3.3 で述べた出発ベクトルを用いると，途中の固有値がぬけることはほとんどない。

§4. 文献

(1) Klaus-Jürgen Bathe, "Solution Method for Large Generalized Eigenvalue Problems in Structural Engineering," thesis, Univ. of California, Berkeley, 1971

(2) F. L. Bauer, "Das Verfahren der Treppeniteration und verwandte Verfahren zur Lösung algebraischer Eigenwertprobleme", ZAMP, 8, 214–235 (1957)

(3) M. Clint & A. Jennings, "The evaluation of eigenvalues and eigenvectors of real symmetric matrices by simultaneous iteration", Comp. J. 13, 76–80 (1970)

(4) H. Rutishauser, "Computational aspects of F. L. Bauer's simultaneous iteration method," Num. Math. 13, 4–13 (1969)