一般化固有値問題

桂田 祐史

2003 年 1 月 28 日,2004 年 12 月 30 日

この文書は以下から入手可能。

http://nalab.mind.meiji.ac.jp/~mk/labo/text/generalized-eigenvalue-problem.pdf

標準固有値問題の計算法については

http://nalab.mind.meiji.ac.jp/~mk/labo/text/eigenvalues.pdf
http://nalab.mind.meiji.ac.jp/~mk/labo/text/eigenvalues-add.pdf

1 序

(行列束とか簡単なことを書きたい…暇が取れるか？)
A, $B \in M(n;\mathbf{C})$ とする。$\lambda \in \mathbf{C}$, $x \in \mathbf{C}^n \setminus \{0\}$ が

$$Ax = \lambda Bx, \quad x \neq 0 \tag{1}$$

をみたすとき、λ を (A,B) の一般化固有値、x を λ に属する一般化固有ベクトルとよぶことにする (あまり普通の言い方ではない？)。

(A,B) の一般化固有値全体を $\sigma(A,B)$ と書こう:

$$\sigma(A,B) := \{\lambda \in \mathbf{C}; \exists y \in \mathbf{C}^n \setminus \{0\} \quad \text{s.t.} \quad Ay = \lambda By\}.$$

正直に言って、あまりテキストがないので詳しいことが分からない。Golub and Van Loan [4] くらいか。特に B が正則でもないような場合は私にとって五里霧中。

2 標準固有値問題への帰着

2.1 問題点

一般固有値問題をそれと同値な標準固有値問題に変換する方法は、以下紹介するように色々ある。

標準固有値問題の解法として、Householder 法のように行列の成分が直接必要なアルゴリズムを利用する場合には、行列の疎性が保存されるかどうかが重要なポイントになる。

これに対して、冪乗法や Lanczos 法を利用する場合には、行列の疎性は必ずしも必要ではない。つまり掛け算

$$x \mapsto Px$$

を効率的に行うのに必ずしも P の疎性は必要ないことに注意しよう。

$$P = P_1 P_2 \cdots P_r$$

と分解できたとして、各 i につき、P_i または P_i^{-1} の一方が疎であれば良いわけである [1]。

2.2 B が正則な場合に使える、ある素朴な方法

B が正則な場合、$\widetilde{A} = B^{-1}A$ とおくと、

$$\widetilde{A}x = \lambda x$$

となる。

命題 2.1 (B が正則な場合) A, B は n 次正方行列、B は正則とするとき、

$$\widetilde{A} := B^{-1}A$$

とおくとき、次の (1), (2) が成り立つ。

(1) 任意の $\lambda \in \mathbf{C}$, $x \in \mathbf{C}^n$ について

$$Ax = \lambda Bx \quad \Longleftrightarrow \quad \widetilde{A}x = \lambda x.$$

(2) 任意の $\lambda \in \mathbf{C}$ について

$$\begin{aligned} \lambda \text{ が } (A,B) \text{ の一般化固有値} \quad &\Longleftrightarrow \quad \lambda \text{ が } \widetilde{A} \text{ の固有値} \\ &\Longleftrightarrow \quad \det(\lambda I - \widetilde{A}) = 0. \end{aligned}$$

ただし I は n 次単位行列である。

そこで (A,B) の一般化固有値 λ の (代数的) 重複度を

$$\det(\lambda I - B^{-1}A) = 0$$

の重複度によって定義する [2]。これから (A,B) の一般化固有値も重複度を考えてちょうど n 個だけ存在することが分かる。

[1] つまり、P と a から $b = Pa$ を計算するとき、P^{-1} が疎であるならば、連立 1 次方程式 $P^{-1}b = a$ を解くことによって b を求める、という手順が採用できる。

[2] 微少な変形をすることによっても定義できる？もっとも B が正則でないような場合まで含めて素朴に考えるのはきっと危険だ。

注意 2.2 (計算向きの方法ではない) A, B が対称であっても $\widetilde{A}$ は対称であるとは限らず、A, B が疎であっても $\widetilde{A}$ は疎であるとは限らず、$\widetilde{A}$ の固有値問題に帰着することによって固有値を計算するのはあまり勧められる方法ではない (よほど小規模な問題でない限り問題外であろう [3])。

2.3 B が正定値である場合にその Cholesky 分解を用いる方法

(この項の要点は式番号をふった式にある。)

B が正定値 (実対称で固有値がすべて正) である場合、Cholesky 分解

$$B = U^T U \tag{2}$$

を代入して (U は正則な上三角行列であり、特に対角成分が正であるものを選ぶことができ、その場合は一意的に決定される)

$$Ax = U^T U x.$$

両辺に $\left(U^T\right)^{-1}$ をかけると

$$\left(U^T\right)^{-1} Ax = Ux.$$

補題 2.3 P が正則であるとき、$\left(P^T\right)^{-1} = (P^{-1})^T$. 特に正則な対称行列の逆行列は対称行列である。

証明 $Q := P^{-1}$ とすると、$PQ = I$. これから $Q^T P^T = (PQ)^T = I^T = I$. ゆえに $\left(P^T\right)^{-1} = Q^T = (P^{-1})^T$. ■

なお、$\left(P^T\right)^{-1}$ のことをしばしば P^{-T} と書く人が多い (個人的には悪趣味だと感じているが)。

そこで

$$A' := (U^{-1})^T A U^{-1}, \quad y := Ux \tag{3}$$

とおくと明らかに

$$Ax = \lambda Bx \iff A'y = \lambda y, \tag{4}$$

$$x = 0 \iff y = 0. \tag{5}$$

A が実対称であれば、A' も実対称である。これは次の補題から分かる。

補題 2.4 P が n 次対称行列であるとき、任意の n 次正方行列 Q について $P' := Q^T P Q$ は対称である。

証明

$$(P')^T = (Q^T P Q)^T = Q^T P^T (Q^T)^T = Q^T P Q = P'. ■$$

[3] しかし、まったく役に立たないというわけではない。ずっと以前、某先生の卒研の学生が加藤敏夫先生の変分法の解説を勉強していて、微分方程式の固有値問題の数値計算に取り組んでいたとき、結局は Mathematica で $\det(\lambda I - B^{-1}A) = 0$ を解くのが一番だった、ということがあった。「I 君の固有値問題」 (http://nalab.mind.meiji.ac.jp/~mk/labo/text/eigen-I.pdf) を見よ。

補題 2.5 A, B は実対称行列で、B は正定値、$B = UU^T$ を Cholesky 分解 (すなわち U は上三角行列) とするとき、$A' := (U^{-1})^T AU^{-1}$ とおくと、

$$\det(\lambda I - A') = \det(\lambda I - B^{-1}A) \quad (\lambda \in \mathbf{C}).$$

証明

$$\begin{aligned}\det(\lambda I - A') &= \det(\lambda UU^{-1} - (U^{-1})^T AU^{-1}) = \det(\lambda U^{-1}U - U^{-1}(U^T)^{-1}A) \\ &= \det(\lambda I - (U^T U)^{-1}A) = \det(\lambda I - B^{-1}A). \blacksquare\end{aligned}$$

ここまでをまとめておこう。

定理 2.6 (対称な一般化固有値問題のコレスキー分解に基づく変換) A, B が n 次実対称行列で、B は正定値であるとき、B の Cholesky 分解 $B = U^T U$ (U は対角成分が正である上三角行列) を用いて、

$$A' := (U^{-1})^T A U$$

とおくと、次の (1)–(3) が成り立つ。

(1) A' は実対称行列である (ゆえにその固有値はすべて実数であり、対応する固有ベクトルで $\mathbf{R}^n$ の正規直交基底をなすものが存在する)。

(2) A' の固有値は (A, B) の一般化固有値に等しい。すなわち、

$$\sigma(A') := \{\lambda \in \mathbf{C}; \exists x \in \mathbf{C}^n \setminus \{0\} \quad \text{s.t.} \quad A'x = \lambda x\},$$
$$\sigma(A, B) := \{\mu \in \mathbf{C}; \exists y \in \mathbf{C}^n \setminus \{0\} \quad \text{s.t.} \quad Ay = \mu By\}$$

とおくとき、$\sigma(A) = \sigma(A, B)$. さらに

$$\det(\lambda I - A') = \det(\lambda - B^{-1} A)$$

であるので重複度まで込めて等しい。

(3) A' の固有値を重複度の分だけ並べて書いたものを

$$\lambda_1, \lambda_2, \cdots, \lambda_n,$$

これらに対応する A' の固有ベクトルからなる $\mathbf{R}^n$ の正規直交基底を

$$v_1, v_2, \cdots, v_n$$

とするとき、

$$u_i := U^{-1} v_i \quad (i = 1, 2, \cdots, n)$$

とおくと、

$$A u_i = \lambda_i B u_i \quad (i = 1, 2, \cdots, n),$$

$$\langle u_i, u_j \rangle = \delta_{ij} \quad (i, j = 1, 2, \cdots, n).$$

ただし $\langle \cdot, \cdot \rangle$ は次式で定義される $\mathbf{R}^n$ の内積である。

$$\langle x, y \rangle := (Bx, y) = (x, By) = (Ux, Uy) \quad (x, y \in \mathbf{R}^n).$$

証明 (1), (2) と (3) の前半は済んでいる。(3) の後半については

$$\langle u_i, u_j \rangle = (Uu_i, Uu_j) = (v_i, v_j) = \delta_{ij}. \blacksquare$$

さて、実対称行列 A' の標準固有値問題については、Rayleigh 商

$$\frac{(A'y,y)}{(y,y)} = \frac{y^T A' y}{y^T y}$$

が重要な役目を果たすが、これについては次の補題が成り立つ。

補題 2.7 A, B が n 次実対称行列で、B は正定値であるとき、B の Cholesky 分解 $B = U^T U$ (U は対角成分が正である上三角行列) を用いて、

$$A' := (U^{-1})^T A U$$

とおくと、任意の $x \in \mathbf{R}^n$ に対して

$$\frac{(Ax,x)}{(Bx,x)} = \frac{(A'y,y)}{(y,y)}, \quad y := Ux. \tag{6}$$

証明　最初に $x \neq 0 \Leftrightarrow y \neq 0$ であることを注意しておく。

$$\frac{(Ax,x)}{(Bx,x)} = \frac{(AU^{-1}y, U^{-1}y)}{(Ux,Ux)} = \frac{((U^{-1})^T AU^{-1}y, y)}{(y,y)} = \frac{(A'y,y)}{(y,y)}.\blacksquare$$

系 2.8 A, B が n 次実対称行列で、B は正定値であるとき、

$$\max \sigma(A,B) = \max_{x\in\mathbf{R}^n\setminus\{0\}} \frac{(Ax,x)}{(Bx,x)}, \quad \min \sigma(A,B) = \min_{x\in\mathbf{R}^n\setminus\{0\}} \frac{(Ax,x)}{(Bx,x)},$$

$$\max\{|\lambda|; \lambda \in \sigma(A,B)\} = \max_{x\in\mathbf{R}^n\setminus\{0\}} \left|\frac{(Ax,x)}{(Bx,x)}\right|.$$

より詳しく min-max 原理も成り立つ。すなわち固有値全部を大きい方から順に (重複度の分だけ繰り返して) 並べたものを

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$$

とするとき、$1 \leq k \leq n$ なる k に対して、

$$\max_{\dim V=k} \min_{x\in V\setminus\{0\}} \frac{(Ax,x)}{(Bx,x)} = \text{大きい方から } k \text{ 番目の固有値} = \lambda_k,$$

$$\min_{\dim V=k} \max_{x\in V\setminus\{0\}} \frac{(Ax,x)}{(Bx,x)} = \text{小さい方から } k \text{ 番目の固有値} = \lambda_{n-k+1}.$$

ここで $\dim V = k$ は V が $\mathbf{R}^n$ のすべての k 次元部分空間を走ることを意味する。

注意 2.9 (数値計算する上での注意) ただし、A, B が疎行列であっても A' は疎行列であるとは限らない (ほとんどの場合は密行列になる)。

なお、冪乗法の系統のアルゴリズム (逆反復法, シフト法) を採用する場合には、事前にシフトをすませておく工夫が有用である。後述の Lanczos 法の節を見よ。

2.4 split Cholesky 分解を用いる方法

LAPACK で採用されている方法である。残念ながら適当な参考文献が入手できていない。LAPACK のドライバー・ルーチン `spbstf` のソース・プログラム `dpbstf.f` 中の注釈くらいしか参考になるものがない。

http://www.netlib.org/cgi-bin/netlibfiles.pl?filename=/lapack/double/dpbstf.f

なお、LAPACK のソースをサーチするには

http://www.cs.colorado.edu/~jessup/lapack/applets/search.html

が便利である。

B を

$$B = S^T S$$

の形に分解するのだが、通常の Cholesky 分解とは異なり、S は上三角行列ではなく、

$$S = \begin{pmatrix} U & 0 \\ M & L \end{pmatrix}, \quad U = (n + k_d)/2 \text{ 次上三角行列}, \quad L = (n - k_d)/2 \text{ 次下三角行列},$$

のような行列である。ここで k_d は半バンド幅である。

$$A' = P^{-T} A P^{-1}$$

とおくと、前小節の補題 2.4 により A' は対称であることが分かる。

A, B が帯行列である場合、A' も (やや帯幅は増えるものの) 帯行列になる。

3 二分法

(準備中 — 村田先生の本から)

4 Lanczos 法

Lanczos 法の利点

一般固有値問題は、行列の疎性を保存したまま標準固有値問題に帰着するのが困難であるわけだが、Lanczos 法 (実対称行列の 3 重対角化のアルゴリズムである) ではその問題を回避できる。その理由は、Lanczos 法は Lanczos 原理 (Arnoldi 原理と呼ぶべき？) に基づいているため、問題の行列 A の成分そのものは必要でなく、任意のベクトル x との積 Ax さえ計算できればよいということにある。

標準固有値問題への帰着

$$Ax = \lambda Bx$$

必要ならばシフトを行う。すなわち任意に選んだ $\widetilde{\lambda} \in \mathbf{R}$ に対して、

$$Ax - \widetilde{\lambda}Bx = (\lambda - \widetilde{\lambda})Bx.$$

そこで

$$\widetilde{A} \stackrel{\text{def.}}{=} A - \widetilde{\lambda}B, \quad \mu \stackrel{\text{def.}}{=} \lambda - \widetilde{\lambda}$$

とおくと

$$\widetilde{A}x = \mu Bx.$$

B の Cholesky 分解を

$$B = U^T U$$

として、上の式に代入すると

$$\widetilde{A}x = \mu U^T Ux.$$

左から $\mu^{-1}U\widetilde{A}^{-1}$ をかけると (ただし $\widetilde{A}^{-1} = (\widetilde{A})^{-1}$)

$$\mu^{-1}Ux = U\widetilde{A}^{-1}U^T Ux.$$

そこで

$$A' \stackrel{\text{def.}}{=} U\widetilde{A}^{-1}U^T, \quad \nu \stackrel{\text{def.}}{=} \mu^{-1}, \quad y \stackrel{\text{def.}}{=} Ux$$

とおくと、

$$A'y = \nu y$$

という標準固有値問題に帰着された。

これに Lanczos 法を適用して、対称三重対角行列 T の固有値問題

$$Tz = \nu z$$

に変換する。

T を計算で求めるのが楽であること

A, B が帯行列であっても、A' が疎性を持つことは期待できないが、任意に与えられた $x \in \mathbf{R}^n$ に対して、

$$A'x = U\widetilde{A}^{-1}U^T x$$

を計算するのは簡単である。帯三角行列 U^T, U の掛け算は容易であるし、$\widetilde{A}^{-1}$ の掛け算 $\widetilde{A}^{-1}z$ は対称帯行列係数の連立 1 次方程式

$$\widetilde{A}w = z$$

を解くことであるから、$\widetilde{A}$ を Cholesky 分解しておけば非常に効率的に計算可能である。したがって Lanczos 法に現れる計算は楽である。

T についての固有値問題の解法

まず T の固有値は A' の固有値に等しく、一般化固有値問題 $\widetilde{A}x = \mu Bx$ の固有値の逆数であることに注意しよう。

T は三重対角であるから、例えば二分法 (bisection method) を使うこともできる。

あるいは冪乗法を行うと、T の絶対値最大の固有値、すなわち $\widetilde{A}x = \mu Bx$ の絶対値最小の固有値の逆数が求まる。ゆえに元の問題 $Ax = \lambda Bx$ の、$\widetilde{\lambda}$ に最も近い固有値が求められることになる。シフト $\widetilde{\lambda} = 0$ とすると、絶対値最小の固有値が求まることになる。

(この辺り矢川・青山 [2] を参考にしたが、絶対値最小の固有値にこだわっているのは、扱っている問題の性質によるのだと思う。この点は整理する必要がある。)

5 精度保証付き数値計算

偏微分方程式の精度保証付き数値計算で、一般化固有値問題の固有値の絶対値の最大値 (これってスペクトル半径と言って良い？) の精度の良い上界が必要になる。そのため一般化固有値問題についても研究が行われている。当の中尾充宏先生のグループが色々試していて、その成果は渡部・山本・中尾 [3] に詳しく説明されている。

また、http://www.cc.kyushu-u.ac.jp/RD/watanabe/RESERCH/paper_list.html には院生の修士論文が色々ある。

桂田研の学生向けに読書ノートが

http://nalab.mind.meiji.ac.jp/~mk/labo/members/reading-wyn1999.pdf

においてある。

A 雑題

A.1 ノルム

以下紹介する命題の証明は、桂田 [6] で読める。

$\mathbf{R}^n$ は自然に $\mathbf{R}$ 上のベクトル空間であるが、任意の $p \in [1, \infty)$ に対して

$$\|x\|_p := \left(\sum_{j=1}^{n} |x_j|^p\right)^{1/p} \qquad (x = (x_1, \cdots, x_n)^T \in \mathbf{R}^n),$$

また $p = \infty$ に対して

$$\|x\|_p = \|x\|_\infty := \max_{1 \le j \le n} |x_j| \quad (x = (x_1, \cdots, x_n)^T \in \mathbf{R}^n)$$

とおくと (本当は n にも依存するわけだが、混乱は起らないと考えられるので、単に $\|\cdot\|_p$ と書くことにする)、任意の $1 \le p \le \infty$ に対して $\|\cdot\|_p$ は $\mathbf{R}^n$ のノルムとなる。以下 $\|\cdot\|_p$ を $\mathbf{R}^n$ の p ノルムと呼ぶことにする。

実数を成分とする m 行 n 列の行列全体の集合を $\mathbf{R}^{m\times n}$ と書く。$m=n$ ならば $\mathbf{R}^{m\times n}$ は多元環、$m\neq n$ であっても $\mathbf{R}^{m\times n}$ はベクトル空間の構造を持つ。

$A\in\mathbf{R}^{m\times n}$ は線型写像 $\mathbf{R}^n\ni x\mapsto Ax\in\mathbf{R}^m$ を引き起こすが、$\mathbf{R}^n$ の任意のノルム $\|\cdot\|$, $\mathbf{R}^m$ の任意のノルム $\|\cdot\|'$ に対して、

$$\|A\|'':=\sum_{x\neq 0}\frac{\|Ax\|'}{\|x\|}=\sum_{\|x\|=1}\|Ax\|'=\sum_{\|x\|\le 1}\|Ax\|'\quad (A\in\mathbf{R}^{m\times n})$$

とおくと、

$$\|A\|''<\infty\quad (A\in\mathbf{R}^{m\times n})$$

となり、$\|\cdot\|''$ は $\mathbf{R}^{m\times n}$ のノルムとなる。これを $\|\cdot\|$, $\|\cdot\|'$ から誘導される作用素ノルムと呼ぶ。

任意の $A\in\mathbf{R}^{n\times n}$ に対して、A のスペクトル半径 (A の固有値の絶対値の最大値) を $r(A)$ と書く。

補題 A.1 (作用素ノルム≧スペクトル半径) $\|\cdot\|$ を $\mathbf{R}^{n\times n}$ の任意の作用素ノルムとするとき、

$$\|A\|\ge r(A)\quad (A\in\mathbf{R}^{n\times n}).$$

$\mathbf{R}^n$ に p ノルムを与えたとき、$\mathbf{R}^{n\times n}$ の作用素ノルム

$$\|A\|:=\sup_{x\in\mathbf{R}^n\setminus\{0\}}\frac{\|Ax\|_p}{\|x\|_p}\quad (A\in\mathbf{R}^{n\times n})$$

を $\mathbf{R}^{n\times n}$ の p ノルムと呼ぶことにして、やはり $\|\cdot\|_p$ という記号で表わす。

命題 A.2 (行列の p ノルム) 任意の $n\in\mathbf{N}$, $A=(a_{ij})\in\mathbf{R}^{n\times n}$ に対して、

$$\|A\|_1=\max_{1\le j\le n}\sum_{i=1}^n|a_{ij}|,\quad \|A\|_\infty=\max_{1\le i\le n}\sum_{j=1}^n|a_{ij}|,$$
$$\|A\|_2=\sqrt{r(AA^T)}=\sqrt{r(A^TA)}\le\sqrt{\|A\|_1\|A\|_\infty}.$$

特に A が対称ならば

$$\|A\|_1=\|A\|_\infty,\quad \|A\|_2=r(A).$$

二つ目の等式は、より一般に A が正規行列 ($AA^T=A^TA$) ならばなりたつ。

A.2 Rayleigh 商と min-max 原理

桂田 [7] を見よ。

前項に引き続き行列 A のスペクトル半径を $r(A)$ と書くことにする。

$A\in\mathbf{R}^{n\times n}$, $x\in\mathbf{R}^n$, $x\neq 0$ とするとき、

$$R_A(x):=\frac{(Ax,x)}{(x,x)}=\frac{x^TAx}{x^Tx}$$

を A の x に対する Rayleigh 商と呼ぶ。($R_A(x)$ という記号はここだけのものである。)

- x が A の固有値 λ に属する固有ベクトルならば、$R_A(x) = \lambda$ である。
- A と x が与えられたとき、$\|Ax - \lambda x\|$ を最小にする λ は $R_A(x)$ である。
- x が A の固有値 λ に属する固有値ベクトルに近いとき、$R_A(x)$ は λ の良い近似値となる (ここには書かないが具体的な評価式が得られる)。

命題 A.3 (min-max 原理) $A \in \mathbf{R}^{n\times n}$ が対称であるとき、A の最大固有値、最小固有値を $\lambda_{\max}, \lambda_{\min}$ とすると、

$$\left\{\frac{(Ax,x)}{(x,x)}; x \in \mathbf{R}^n \setminus \{0\}\right\} = [\lambda_{\min}, \lambda_{\max}],$$

$$\sup_{x\in\mathbf{R}^n\setminus\{0\}} \frac{(Ax,x)}{(x,x)} = \lambda_{\max}, \quad \inf_{x\in\mathbf{R}^n\setminus\{0\}} \frac{(Ax,x)}{(x,x)} = \lambda_{\min}, \quad \sup_{x\in\mathbf{R}^n\setminus\{0\}} \left|\frac{(Ax,x)}{(x,x)}\right| = r(A).$$

より詳しく、A の固有値を (重複度の分だけ繰り返して並べて) $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_n$ とすると、

$$\begin{aligned}
\min_{\dim V=k} \max_{x\in V\setminus\{0\}} R_A(x) &= \text{小さい方から } k \text{ 番目の固有値} = \lambda_{n-k+1} \quad (1 \le k \le n),\\
\max_{\dim V=k} \min_{x\in V\setminus\{0\}} R_A(x) &= \text{大きい方から } k \text{ 番目の固有値} = \lambda_k \quad (1 \le k \le n).
\end{aligned}$$

ただし “$\dim V = k$” は「$\mathbf{R}^n$ のすべての k 次元線型部分空間 V について」という意味である。

同じことを次のように表現してある本も多い。

$$\begin{aligned}
\min_{\dim S=j-1} \max_{x\in S^\perp\setminus\{0\}} R_A(x) &= \text{大きい方から } j \text{ 番目の固有値} = \lambda_j \quad (1 \le j \le n),\\
\max_{\dim S=j-1} \min_{x\in S^\perp\setminus\{0\}} R_A(x) &= \text{小さい方から } j \text{ 番目の固有値} = \lambda_{n-j+1} \quad (1 \le j \le n).
\end{aligned}$$

(厳密には命題の後半を min-max 原理という。証明については例えば ??? を見よ。)

A.3 正値対称行列, Cholesky 分解

Cholesky 分解については桂田 [5] を見よ。

$A = (a_{ij}) \in M^{n\times n}$ が上三角行列であるとは、対角線より下にあるすべての成分が 0, つまり

$$i > j \quad \Longrightarrow \quad a_{ij} = 0$$

を満たすことをいう。

命題 A.4 (正値対称行列の Cholesky 分解の存在) $A \in \mathbf{R}^{n\times n}$ が正値対称行列ならば

$$A = U^T U \tag{7}$$

を満たす上三角行列 U が存在する。特に U の対角成分がすべて 0 であるものに限ると U は一意的である。

与えられた A に対して、(7) を満たす U を求めることを「A を Cholesky 分解する」という。

命題 A.5 (正値対称行列の平方根) $A \in \mathbf{R}^{n\times n}$ が正値対称行列ならば

$$A = B^2$$

を満たす正値対称行列 B が存在する。

証明　A が実対称であるから、ある実直交行列 P を取ると

$$P^T AP = \operatorname{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n) \quad (\lambda_i \text{ は } A \text{ の固有値})$$

となることはよく知られている。

$$A = P \operatorname{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n) P^T$$

であるが、

$$B := P \operatorname{diag}\left(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \cdots, \sqrt{\lambda_n}\right) P^T$$

とおくと、B は明らかに実対称で、$P^T P = I$ に注意すると

$$\begin{aligned} B^2 &= P \operatorname{diag}\left(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \cdots, \sqrt{\lambda_n}\right) P^T P \operatorname{diag}\left(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \cdots, \sqrt{\lambda_n}\right) P^T \\ &= P \operatorname{diag}\left(\sqrt{\lambda_1}^2, \sqrt{\lambda_2}^2, \cdots, \sqrt{\lambda_n}^2\right) P^T = P \operatorname{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n) P^T = A. \blacksquare \end{aligned}$$

A.4　misc

命題 A.6 (帯行列の積は帯行列) A, B は n 次正方行列で、それぞれ半バンド幅が m_A, m_B であるとする。つまり

$$|i-j| > m_A \implies a_{ij} = 0, \quad |i-j| > m_B \implies b_{ij} = 0.$$

このとき AB の半バンド幅は $m_A m_B$ 以下である。すなわち

$$|i-j| > m_A m_B \implies (AB)_{ij} = 0.$$

あれ $m_A + m_B$ かな？証明を書け。

B 文献について

一般化固有値問題のアルゴリズムについて、参考になったのは、村田 [1], 矢川・青山 [2], LAPACK USERS' GUIDE だったりする。古い本が冴えないのが多い中、村田先生の本が役に立つのはさすがだなあ、と思う。

一般化固有値の絶対値の上界評価については、とにかく渡部・山本・中尾充宏 [3] が便利である。

ベクトル、行列のノルムについては、桂田 [6] が便利である。行列の (標準) 固有値問題のアルゴリズムについては、桂田 [7] とその続編 (将来マージ予定) の桂田 [8] を見よ。正値対称行列の Cholesky 分解については、桂田 [5] を見よ。

参考文献

[1] 村田 健郎, 線形代数と線形計算法序説, サイエンス社 (1986).

[2] 矢川 元基・青山裕司, 有限要素固有値解析, 森北出版 (2001).

[3] 渡部善隆・山本野人・中尾充宏, 一般化固有値問題の精度保証付き計算とその応用, 応用数理学会論文誌, VOL.9 NO.3 SEPTEMBER (1999), pp.137–150.
なお、http://www.cc.kyushu-u.ac.jp/scp/system/library/PROGRAM_LIBRARIES/Nakao/vpgep/vpgep.pdf も参照せよ。

[4] G. H. Golub and C. F. Van Loan, Matrix Computations, third edition, The Johns Hopkins University Press (1996).

[5] 桂田 祐史, Cholesky 分解ノート, http://nalab.mind.meiji.ac.jp/~mk/labo/text/cholesky.pdf においてある。

[6] 桂田 祐史, 連立 1 次方程式 III 「自分が使うための線形代数」, http://nalab.mind.meiji.ac.jp/~mk/labo/text/linear-eq-3.pdf においてある。

[7] 桂田 祐史, 行列の固有値問題, http://nalab.mind.meiji.ac.jp/~mk/labo/text/eigenvalues.pdf においてある。

[8] 桂田 祐史, 固有値問題ノートの補足, http://nalab.mind.meiji.ac.jp/~mk/labo/text/eigenvalues-add.pdf においてある。

[9] 桂田 祐史, 一般化固有値問題, http://nalab.mind.meiji.ac.jp/~mk/labo/text/generalized-eigenvalue-problem.pdf においてある。
この文書である。

[10] 桂田 祐史, 『渡部・山本・中尾の一般化固有値問題の論文 (1999) を読む』, http://nalab.mind.meiji.ac.jp/~mk/labo/members/reading-wyn1999.pdf
渡部善隆・山本野人・中尾充宏 [3] の読書ノート。

[11] N. Yamamoto and M. T. Nakao, Numerical verifications of solutions for elliptic equations in non-convex polygonal domains, Numerische Mathematik, Vol.65 (1993), pp. 503–521.

[12] T. Yamamoto, Error bounds for computed eigenvalues and eigenvectors, Numerische Mathematik, Vol.34 (1980), pp. 189–199.

[13] T. Yamamoto, Error bounds for computed eigenvalues and eigenvectors II, Numerische Mathematik, Vol.40 (1982), pp. 201–206.

[14] S. M. Rump, Verification Methods for Dense and Sparse Systems of equations, in Jürgen Herzberger (Editor), Topics in Validated Computations (Proceedings of the IMACS-GAMM International Workshop on Validated Computation, Oldenburg, Germany, August 30th–September 3rd, 1993), Elsevier Science, North Holland, Amsterdam, 1994, pp. 63–135.