

対称一般固有値問題の フィルタ作用素を用いた不変部分空間の近似構成

村 上 弘^{†1}

大規模な実対称定値一般固有値問題 $A\mathbf{v} = \lambda B\mathbf{v}$ の固有対のうちで、固有値が区間 $[a, b]$ にあるものをフィルタ対角化法で近似して求める。フィルタ対角化法はフィルタの出力ベクトルの組が張る部分空間に対して subspace 法を適用して近似固有対を求める。フィルタにはシフト量が複素数のレゾルベントの線形結合を用いる。フィルタからの出力ベクトルの組は、通常は数値的階数が多く落ちているので、そのままでは subspace 法に与える部分空間の基底としては適切ではない。そのためこれまでの方法では、フィルタの出力ベクトルの組を B -計量で特異値分解し、特異値が閾値以上の特異ベクトルだけを集めた B -正規直交基底を構成し、それを subspace 法に与える基底として用いてきた。今回の新しい方法では、フィルタへの入力ベクトルの組とフィルタからの出力ベクトルの組の両方およびフィルタの伝達特性の情報を用いて、区間 $[a, b]$ に対応する固有値を持つ不変部分空間の近似基底となるように B -正規直交基底を構成する。係数行列が帯の場合の実対称定値一般固有値問題を解いた数値実験例をいくつか示す。

Construction of the Approximate Invariant Subspace of a Symmetric Generalized Eigenproblem by the Filter Operator

HIROSHI MURAKAMI^{†1}

By the filter diagonalization method, for a real symmetric definite generalized eigenproblem $A\mathbf{v} = \lambda B\mathbf{v}$ of large size, only those eigenpairs are solved whose eigenvalues are in the interval $[a, b]$. The filter diagonalization method applies the subspace method to the subspace which is spanned by the set of filter's output vectors to give approximations of eigenpairs. The filter we use is a linear combination of resolvents with complex shifts. Since the set of filter's output vectors usually has many rank deficiencies, the direct application of the subspace method to the set is not appropriate. Therefore, previously the set of B -orthonormal basis for the subspace method has been constructed by the use of SVD of the filter's output vectors in B -metric with the cut-off of singu-

lar values below a threshold. Our new method makes use both information of input and output vectors of the filter and the property of the filter operator to construct a set of B -orthonormal basis so that the set spans an approximation of the invariant subspace whose eigenvalues are in the interval $[a, b]$. Some results of numerical experiments are shown which solved real symmetric definite generalized eigenproblems for the cases both coefficient matrices are banded.

1. はじめに

フィルタ対角化法¹⁾⁻⁷⁾を用いると、実対称標準固有値問題と同様の方法で実対称定値一般固有値問題に対しても、固有値分布のある狭い範囲に固有値がある固有対だけを複数同時に求めることができる。そこで、大規模な実対称定値一般固有値問題 $A\mathbf{v} = \lambda B\mathbf{v}$ の区間 $[a, b]$ に固有値がある固有対の近似をフィルタ対角化法を用いて求めることにする。いまシフト量 τ のレゾルベントを $\mathcal{R}(\tau) = (A - \tau B)^{-1}B$ と表すとき、本論文でフィルタとして用いる作用素はシフト量が複素数のレゾルベントのいくつかの線形結合で表せるものとする。

レゾルベントの作用は連立 1 次方程式を解いて実現するので、係数 $C = A - \tau B$ の連立 1 次方程式が高速に解けるならば、固有値問題の近似解法は連立 1 次方程式の解法に帰着できることになる。固有値問題の係数行列が密ならば利点がないが、疎行列の場合には有利になりうる。たとえば帯幅の小さい行列であれば帯行列専用の直接解法を用いて、ランダム疎行列であれば（必要に応じて前処理を加えた）反復法を用いて、大規模な実対称固有値問題が解ける。今回の実験例では係数 A, B が帯幅の小さい行列の場合だけを扱い、帯行列専用の LU 分解を用いて係数が帯行列 C の連立 1 次方程式を解いている。

フィルタ対角化法では、十分多くのベクトルの組 X をフィルタへの入力として与え、フィルタから出力されたベクトルの組 Y の張る部分空間の中から subspace 法を用いて近似固有対を取り出す（本論文において「subspace 法」とは Rayleigh-Ritz 法、すなわち与えられたベクトルの組が張る部分空間の中で一般固有値問題の Ritz 値と Ritz ベクトルの組（Ritz 対）を近似固有対として求めることを意味するものとする）。フィルタから出力されたベクトルの組 Y は通常は数値的階数が多く落ちているので、subspace 法を適用する部分空間の基底として Y をそのまま用いるのは適切ではない。

^{†1} 首都大学東京 数理情報科学専攻

Department of Mathematics and Information Sciences, Tokyo Metropolitan University

従来の方法では、フィルタの出力ベクトルの組 Y に B -計量の特異値分解を適用して、閾値以上の特異値を持つ特異ベクトルだけを集めて作った B -正規直交基底を subspace 法を適用する部分空間の基底として用いてきた。しかし閾値の設定には明確な原理がなかった。さらに、 Y を計量 B で特異値分解する際には行列 A はまったく参照されないの、この方法で得られる B -正規直交基底は元の固有値問題の不変部分空間の基底の良い近似にはならない。

今回の新しい方法では、出力ベクトルの組 Y と入力ベクトルの組 X の両方の情報、およびフィルタ作用素の伝達関数の特性値を参照することで、「区間 $[a, b]$ に固有値を持つ不変部分空間」の基底の良い近似となるように B -正規直交基底をうまく構成できるので、それに subspace 法を適用して必要な近似固有対が求められる。

2. フィルタとその伝達関数

フィルタ対角化法では、大規模な N 次の実対称定値な一般固有値問題 (GEVP): $Av = \lambda Bv$ の区間 $[a, b]$ 内に固有値がある固有対を選択的に求めるために、通過帯域が $[a, b]$ のフィルタを用いる。フィルタ作用素 \mathcal{F} は複素シフト量 τ のレゾルベント $\mathcal{R}(\tau) \equiv (A - \tau B)^{-1} B$ の線形結合: $\mathcal{F} = c_\infty I + \sum_{p=1}^{2n} \gamma_p \mathcal{R}(\tau_p)$ であるとする。そのとき、固有値が λ の固有ベクトル v に対しては $\mathcal{F}v = f(\lambda)v$ が成立する。ここで、 $f(\lambda) = c_\infty + \sum_{p=1}^{2n} \gamma_p / (\lambda - \tau_p)$ は λ の有理関数であり、固有値 λ を持つ固有ベクトルに対する伝達率を与えるので伝達関数と呼ばれる。フィルタ \mathcal{F} と伝達関数 $f(\lambda)$ に含まれているパラメータの次数 n 、複素シフト量 τ_p 、複素係数 γ_p 、実係数 c_∞ は、伝達関数の特性に対して課された制約条件を満たすように決める。今回の実対称定値一般固有値問題では固有値も固有ベクトルもすべて実数の範囲にとれるので、フィルタを実ベクトルに作用させた結果が再び実ベクトルとなるように、 \mathcal{F} は実の演算子で、 $f(\lambda)$ も実の有理関数とする。

アナログ回路のフィルタ設計法^{8),9)} に登場する典型的な 4 種類のフィルタである Butterworth, Chebyshev, inverse Chebyshev, elliptic と同じ伝達率の特性を持つフィルタがレゾルベントの線形結合により構成できる^{7),10)}。Butterworth フィルタの場合のレゾルベントのシフト量の分点分布は、複素平面内の円周上の等分点であり、文献 2), 3) で用いられているフィルタの分点分布と同じである (ただし、文献 2) の方法は、モーメント展開を経由する定式化であり、フィルタはレゾルベントの線形結合の形で表されたものではない)。

2.1 フィルタの設計

いま $\lambda \in [a, b]$ から $t \in [-1, 1]$ への線形変換により λ の正規化座標 t を定義する

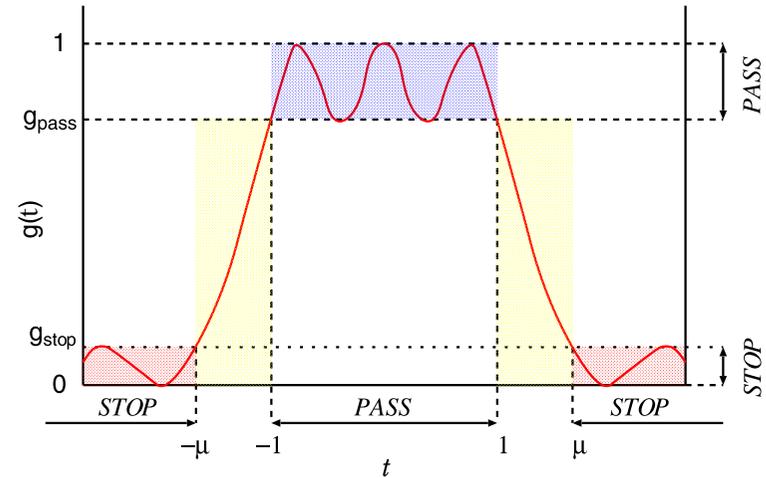


図 1 フィルタの伝達関数 $g(t)$ の概念図

Fig. 1 Conceptual graph of the filter's transfer function $g(t)$.

($\lambda = \frac{1}{2}(a+b) + \frac{1}{2}(b-a)t$ である)。そのとき $|t| \leq 1$ は通過帯域 (passband) に、 $|t| \geq \mu > 1$ は阻止帯域 (stopbands) に、途中の $1 < |t| < \mu$ は遷移帯域 (transitionbands) にそれぞれ対応する。フィルタの種類は伝達関数 $f(\lambda) = g(t)$ の関数形の違いである。伝達関数 $g(t)$ の値には制約条件を課して、通過帯域での $g(t)$ の下限値を g_{pass} 、阻止帯域での $g(t)$ の上限値を g_{stop} とする (図 1 参照)。簡単化のため g_{pass} についての条件は tight であるとする。

典型的な 4 種類のフィルタに対しては、制約条件の 3 組 ($\mu, g_{\text{pass}}, g_{\text{stop}}$) を指定すると、条件を満たせる次数 n の最小値が決まり、 n をその値以上に設定すれば $g(t)$ が完全に決まる (もしも制約条件の 3 組を (μ, g_{pass}, n) で与える場合には、条件を満たせる g_{stop} の最小値が決まり、 g_{stop} をその値以上に設定すれば $g(t)$ が完全に決まる。同様に制約条件の 3 組を ($g_{\text{pass}}, g_{\text{stop}}, n$) で与える場合には、条件を満たせる μ の最小値が決まり、 μ をその値以上に設定すれば $g(t)$ が完全に決まる)。

関数 $g(t)$ が決まると、伝達関数 $f(\lambda) = g(t)$ の複素極の位置と極の係数から、通過帯域が $\lambda \in [a, b]$ であるフィルタ \mathcal{F} のパラメータが完全に決まる。

フィルタの設計の詳細 (特に典型的な 4 種類のフィルタ) については、すでに文献 7), 10), 11) に記述したので今回は記述を省略する。

3. Subspace 法に与える基底の構成 (従来の方法)

ランダムな N 次ベクトルを十分多く m 個 ($m \ll N$) 集め, それを計量 B で正規直交化することで線形独立性を高めた $N \times m$ 行列をベクトルの組 X とする ($X^T B X = I_m$ である. ただし I_m は m 次の単位行列). この組 X をフィルタ \mathcal{F} への入力として, 各ベクトルごとに独立に作用させて, $N \times m$ 行列である出力ベクトルの組 Y を作る. Y の張る空間は固有値が区間 $[a, b]$ の近傍にある固有ベクトルで張られたものになる. そうして subspace 法により Y の張る部分空間内で元の GEVP の近似対を求める.

フィルタの性質から, Y は一般には数値的階数が多く落ちた行列になる. 数値的安定化のため従来は, この $N \times m$ 行列 Y を計量 B で特異値分解する (すなわち, $Y = U \Sigma V^T$, $U^T B U = I_m$, $V^T V = I_m$ となる $N \times m$ の B -正規直交である特異ベクトルの組 U , 非負である特異値が減少順に並んだ m 次の対角行列 Σ , m 次の直交行列 V を求める). そうして, 特異値が閾値よりも大きい特異ベクトル r ($\leq m$) 個だけを集めた組 Z を作り, (Z はすでに B -正規直交系であるから) 小さい次数 r の対称行列 $Z^T A Z$ を係数とする標準固有値問題に帰着させてきた.

特異ベクトルの組は Y の張る空間を計量 B でよく説明する軸ではあるが, 元の GEVP を参照しないで構成されるので, 各特異ベクトルは固有値が $[a, b]$ 近傍の固有ベクトルの線型結合である.

切断の閾値を大きくすると, Z に含まれる基底の個数が減って部分空間が狭まるので一般に subspace 法の近似度は下がる. しかし閾値を過度に小さくとると, 特異値の小さいベクトル (丸め誤差の占める割合が多く有効精度が落ちている) も基底 Z に参加して, subspace 法で得られる近似対に残差の大きい「偽の固有対」が多く現れるようになる.

4. Subspace 法に与える基底の構成 (今回の新しい方法)

フィルタ \mathcal{F} がレゾルベントの線形結合の場合には, 元の GEVP: $A\mathbf{v} = \lambda B\mathbf{v}$ の固有対 (λ, \mathbf{v}) のベクトル \mathbf{v} は, また同時に $\mathcal{F}\mathbf{v} = f(\lambda)\mathbf{v}$ を満たすので \mathcal{F} の固有ベクトルにもなる, という性質を利用する. 逆は必ずしも成り立たない. そこで, \mathcal{F} の不変部分空間は元の GEVP の不変部分空間になることを利用して \mathcal{F} の固有ベクトルから元の GEVP の固有ベクトルを復元する.

まず Y の張る部分空間内に制限された作用素 \mathcal{F} の近似固有対 (ϕ, \mathbf{v}) を求める (方法は後述. またベクトル \mathbf{v} は B -正規直交系になるようにとれる).

得られた \mathcal{F} の近似固有対 (ϕ, \mathbf{v}) のうちで, 区間 $[a, b]$ に対する伝達関数 $f(\lambda)$ の値域の「近傍」に ϕ の値が含まれているようなベクトル \mathbf{v} を集めて組 Z を作る (レゾルベントの線形結合によるフィルタの設計パラメータで表せば, $\lambda \in [a, b]$ に対する $f(\lambda)$ の値域は $[g_{\text{pass}}, 1]$ である ($f(\lambda)$ から c_∞ の項を省略する場合には $[g_{\text{pass}} - c_\infty, 1 - c_\infty]$ となる)).

そうして, この組 Z が張る部分空間内での元の GEVP: $A\mathbf{v} = \lambda B\mathbf{v}$ の近似固有対を subspace 法で求める.

注 意

数値計算では ϕ の値は丸め誤差などの影響で摂動を受けて真値からずれるので, その影響により本当は必要な近似固有対 (ϕ, \mathbf{v}) のベクトルを落としてしまわないようにする必要がある. 誤差の影響を考慮するとき ϕ の真値が区間 $[a, b]$ に対する $f(\lambda)$ の値域に含まれている可能性があれば, そのような ϕ の値を持つ対 (ϕ, \mathbf{v}) のベクトル \mathbf{v} も組 Z に参加させておく必要がある. ϕ の閾値を下げると, Z で張られた近似不変部分空間の固有値の範囲が広がる. ただし, ϕ の絶対値が小さい対 (ϕ, \mathbf{v}) のベクトル \mathbf{v} は, フィルタで強く減衰するため数値相殺による丸め誤差を含む割合が大きいので ϕ に反比例して精度が低い. するとそのようなベクトルを Z に含めると, subspace 法で得られる元の GEVP の近似対の精度が劣化する可能性がある. そこで ϕ の閾値は区間 $(g_{\text{stop}}, g_{\text{pass}})$, あるいは c_∞ の項を省略する場合には ϕ の閾値は区間 $(\max(0, g_{\text{stop}} - c_\infty), g_{\text{pass}} - c_\infty)$, の内にある値で, 0 にはあまり近くない正の値に設定する.

不変部分空間の基底の近似 Z を構成する場合に注意すべき他の点は, ϕ の真値が重複している場合には, 重複している ϕ の値を持つ固有対 (ϕ, \mathbf{v}) のベクトル \mathbf{v} を Z にすべて含めるかあるいはまったく含めないかのどちらかにしなければならず, 部分的に含まれてはならないことである. ϕ の真値が重複していても, 数値計算による近似値は摂動を受けて重複が解けているのが普通であり, 真値が重複しているかどうかは ϕ の近似値をみても判断できない. そのため, ϕ の近似値の分布の中である程度大きな間隙がある場所に閾値を設定する. これにより, 真値は重複しているが摂動により重複が解けている ϕ の値を閾値で分断するリスクを避ける. どの程度大きい間隙ならば十分であるかは, 誤差限界の評価による.

4.1 \mathcal{F} の近似固有対の解法

Ritz-Galerkin 法を用いて以下の手順で, 組 Y の張る部分空間内での \mathcal{F} の近似固有対 (ϕ, \mathbf{v}) を求める.

固有値方程式 $\mathcal{F}\mathbf{v} = \phi \mathbf{v}$ に対して $\mathbf{v} = Y\mathbf{u}$ とおき, 両辺に左から $X^T B$ を乗じると, 小さい次数 m の GEVP である $\alpha \mathbf{u} = \phi \beta \mathbf{u}$ を得る. ここで $\alpha = X^T B \mathcal{F} Y$, $\beta = X^T B Y$

である． \mathcal{F} がレゾルベントの線形結合の形式であるという仮定から容易に導かれる関係 $B\mathcal{F} = \mathcal{F}^T B$ を用いると $\alpha = Y^T B Y$ が分かり、 α は半正定値対称である．同様に $\beta = X^T B F X = Y^T B X = \beta^T$ であるから β も対称である．

典型的 4 種類のフィルタでは $f(\lambda)$ が非負の実関数であることから β は半正定値になる (ただし、フィルタの係数 c_∞ は、次数 n が偶数でフィルタの種類が inverse Chebyshev あるいは elliptic の場合にだけ零でない正の値を持つが、その大きさが微小である場合に零と見なして省略する近似を行う場合には、 β は微小な大きさの負の固有値を持ちうる)．

フィルタの性質から通常、 Y の特異値には微小な値が多いので、係数 α, β は両方とも悪条件になるので、そのことを考慮して固有値方程式 $\alpha u = \phi \beta u$ は以下の特別な方法で解く．

4.2 固有値方程式 $\alpha u = \phi \beta u$ の解法

表式に Y が積の形で α には 2 回、 β には 1 回、入っているので、行列 β は行列 α よりも数値的条件が良い．そこで絶対値の小さい固有値を持つ固有ベクトルに対しても計算精度が高い Rutishauser の Jacobi 法¹²⁾ を用いて β を固有値分解して、固有値を減少順に (負のものは後に) 並べて $\beta = Q^T D Q$ とする

直交変換 $u \equiv Q w$ により $G \equiv Q \alpha Q^T$ とおくと、方程式は $G w = \phi D w$ となる．

きわめて小さい閾値 ϵ (実験で用いた ϵ の値はマシンイプシロン ϵ_M の 100 倍とした．IEEE 754 規格の 64-bit 倍精度ではマシンイプシロンの値は約 2.22×10^{-16} である) を決めて、 ϵ 未満の D の対角要素と対応する行と列を G, D, w から省いて、切断された固有値方程式 $\hat{G} \hat{w} = \phi \hat{D} \hat{w}$ を得る．さらに、変換 $\hat{w} \equiv \hat{D}^{1/2} z$ により $H \equiv \hat{D}^{-1/2} \hat{G} \hat{D}^{-1/2}$ とおいて得られる対称標準固有値問題 $H z = \phi z$ を Jacobi 法で解く．

その固有対 (ϕ, z) のベクトル z から逆変換により対応する \hat{w} を得て、切断された行の自由度に零を補って w を得て、さらに u を得て、 $v = Y u$ を得る．

(注: Jacobi 法で解いた z は正規直交系になり、 u は β -正規直交系になる．そのとき $v^{(i)} = (1/\sqrt{\phi^{(i)}}) \cdot Y u^{(i)}$ と規格化すれば、 $v^{(i)}$ は B -正規直交になる．)

5. 数値実験例

計算機システム

数値実験に用いた計算機システムの仕様は次のものである．CPU は intel Core i7-920 (2.66 GHz, 8 MB L3, 4 コア, Hyperthread 機能オフ) で主記憶 24 GB (triple channel; 6×4 GB DDR3 1333 MHz PC3-10600 DIMM)．コンパイラは intel Fortran/OpenMP v12.0 for intel 64 (コンパイルオプション `-fast -openmp`) で、演算には IEEE 754 規格

の 64-bit 倍精度浮動小数点数を用いた．OS は Fedora 14 for intel 64 である．

OpenMP によるスレッド並列化について

実験例はすべて上記の 1 CPU (intel Core i7-920) のシステムを用いて行った．この CPU はコアを 4 個持ちマルチコア並列処理が可能である．使用したコンパイラ (intel Fortran) は Fortran90 言語で書かれたプログラムに OpenMP の指示行を適切に挿入することで、比較的容易にスレッド並列化による並列処理が記述できる．今回の各実験では、1 スレッド並列化なし (1 コア実行) による実行時間と 4 スレッド並列化 (4 コア実行) による実行時間とを、フィルタ対角化法の部分とその後の 2 回の Rayleigh 商逆反復についてそれぞれあげた．実験の各例では以下で示すように、4 スレッドで並列化したことにより計算の経過時間が 1/3 以下に短縮され、安価な PC のマルチコア CPU の計算機性能がより良く発揮できた．ただし、記載している経過時間は、並列化を行った場合も行わなかった場合も、あくまでも現状のプログラム実装で計算した場合の参考値であって、記憶参照のパターンや、連立 1 次方程式の解法、帯行列ベクトル積、 B -正規直交化、 B -SVD、小規模な対称行列の Jacobi 対角化などの今後の実装上の改良によりまだ時間短縮の余地はあると考えている．

フィルタ対角化法の計算の主要部は、フィルタ \mathcal{F} を入力ベクトルの組 X に作用させて出力ベクトルの組 Y を作る処理であり、ここまでの処理は従来の方法でも新しい方法でもまったく同じ計算をしている．フィルタ作用素が複数の異なる複素シフト量が τ_p であるレゾルベント $\mathcal{R}(\tau_p)$ の複素係数 γ_p による線形結合であることを用いて、異なる p それぞれについて $\mathcal{R}(\tau_p)X$ を作るたびに γ_p を乗じて (その実部を) Y に加えている．並列化を行う場合は、この異なる p に対する作業を OpenMP を用いて 4 スレッドに分配して処理した．

また Rayleigh 商逆反復の処理では、複数の近似対に対する処理を 4 スレッドで分配して同時並列に行った．

レゾルベントの作用の計算法

複素数 τ をシフトとするレゾルベント $\mathcal{R}(\tau) = (A - \tau B)^{-1} B$ をベクトルの組 X に作用させる計算は、 $C = A - \tau B$ とおくと A, B の実対称性から C は複素対称行列で、連立 1 次方程式の組 $CZ = BX$ をベクトルの組 Z について解くことに帰着される．

いま A, B が帯行列で半帯幅が h である場合には、 C も半帯幅 h となり、連立 1 次方程式は帯行列専用の行ピボット交換付きの LU 分解法¹³⁾ を用いて安定に解くことができる．その場合 LU 分解の L は下半帯幅が h で、 U は上半帯幅 $2h$ 以下となる．

複素シフト τ が実数でなければ C は必ず正則なので、行ピボット交換を省ける可能性もあるが、数値的に不安定となる可能性も考えられる．また C の対称性 $C^T = C$ を利用する

修正 Cholesky 法 $C = LDL^T$ やそれに対角ピボット交換を加えた方法では、たとえ C が正則でも対角ピボットに零もしくは絶対値の小さい値が現れて計算が不安定となる可能性がある。数値安定性のためには対称性を保ちながらピボット選択を行いながら帯幅の増加を抑えるような複雑な処理を行う必要が生じる。そのため現状では C の対称性をあえて使わずに、帯行列に対する行ピボット交換付きの LU 分解を用いている。

逆反復法で固有ベクトルを改良する計算では、近似固有値 λ に対して実対称行列 $C = A - \lambda B$ を係数とする連立 1 次方程式を解くが、これも C の対称性を使わずに、帯行列に対する行ピボット交換付きの LU 分解^{13)–15)} を用いて解いている。

近似固有対の精度の評価方法

近似固有対 (λ, \mathbf{v}) のベクトル \mathbf{v} が B -正規化されていれば、残差ベクトル $\mathbf{r} \equiv (A - \lambda B)\mathbf{v}$ の B^{-1} -ノルムである $\Delta \equiv \sqrt{\mathbf{r}^T B^{-1} \mathbf{r}}$ は固有値の誤差上界であり、 λ から距離 Δ 以内に真の固有値が必ずある。これは標準固有値問題の場合の Wilkinson の上界の簡単な拡張となっている。ただし、固有ベクトルの近似が悪いときには固有値の誤差上界としては過大評価で、固有値の精度推定には逆反復による改良で得られる近似固有値の変動の大きさの方が良い。

実験の例題について

以下の例題 1 から 3 において、フィルタ出力の組 Y から subspace 法に与える基底 Z の構成法として、切断付き SVD を用いた従来の方法と提案の新しい方法とを比較実験する。

例題 1 は、従来の方法では subspace 法に与える部分空間の基底は、不変部分空間の近似空間を張るようには構成されていないので、得られる近似対には一般に残差の大きい「偽の固有対」も含まれる。特異値の閾値（基底の規模）の選択によっては「偽の固有対」が固有値が指定区間の外にある近似対としてだけではなくて、指定区間の中にある近似対として出現する可能性があることを示す例である（新しい方法では subspace 法に与える部分空間の基底は、不変部分空間の近似空間を張るように構成しているので「偽の固有対」は現れない）。

例題 2 と例題 3 の GEVP の係数行列は同じであるが、固有値を求める区間および用いたフィルタの特性が異なる。例題 2 のフィルタは elliptic で区間 $[100, 150]$ 、 $n = 16$ 、 $\mu = 1.1$ 、 $g_{\text{pass}} = 0.5$ 、 $g_{\text{stop}} = 10^{-14}$ で特性が急峻で高性能な例として取り上げ、例題 3 のフィルタは Chebyshev で、 $n = 4$ 、 $\mu = 2$ 、 $g_{\text{pass}} = 0.5$ 、 $g_{\text{stop}} = 1.1 \times 10^{-4}$ で特性がゆるやかで性能があまり良くない例として取り上げている。例題 2 ではフィルタの性能が非常に良かったためか、最終的に得られた必要な近似対の精度は従来の方法と新しい方法で同等であった。例題 3 では、従来の方法では特異値の閾値（subspace 法に与える基底の規模）の選択はかなり難しく、得られる結果が良くなるように試行錯誤を行って見つけた閾値を採用したが、新

しい方法ではそのような試行錯誤は必要なかった。

各実験で示した経過時間の表の項目で「乱数ベクトルの生成」の経過時間の値が 1 スレッドと 4 スレッドで（ほぼ）同じになっているのは、実際には乱数ベクトルの組の生成は並列化をしていないからである。これは乱数で生成される初期ベクトルの組の値を同じに保つ方法として最も簡単であり、また経過時間の全体に占める割合が小さいので並列化の必要性を特に感じなかったからでもある。

5.1 例題 1（次数 $n = 12$ の elliptic フィルタ）

この例題の係数行列 A, B はどちらも次数 $N = 10^5$ の帯行列で、半帯幅は $h = 100$ である。帯内 ($|p - q| \leq h$) の行列要素は、 $A_{p,q} = pq / \sqrt{p^2 + q^2}$ 、 $B_{p,q} = 1 / (p + q - 1) + \delta_{p,q}$ で与えた。ただし $p, q = 1, 2, \dots, N$ で、 $\delta_{p,q}$ は Kronecker 記号である。この固有値問題の固有対で固有値が区間 $[50, 100]$ 内のものは 110 個である。

入力ベクトルの組 X は、乱数ベクトル $m = 150$ 個を B -正規直交化して作成した。使用したフィルタは次数 $n = 12$ の elliptic で、 $\mu = 1.4$ 、 $g_{\text{pass}} = 0.5$ (tight)、 $g_{\text{stop}} = 1.0 \times 10^{-15}$ である。フィルタの係数 c_{∞} は 10^{-15} 程度と小さいので省いた。

従来の方法（例題 1）

B -SVD の特異値分布をグラフ（図 2）に示す。特異値の閾値を 10^{-9} とすると、130 個の特異ベクトルが切断の結果として得られた。その 130 個のベクトルからなる基底に対して subspace 法を適用したところ、得られた近似固有対で固有値が区間 $[50, 100]$ 内にあるものは 111 個であった。これらの 111 個の近似対だけについてグラフ（図 3）に、横軸に近似固有値、縦軸に残差のノルム Δ の値をプロットした。グラフ中の IT0 がフィルタ対角化法による残差のノルムである。この IT0 のグラフからフィルタ対角化法により得られた近似固有対の中に残差が非常に大きい偽の固有対で固有値が区間内のものが 1 個あること、その 1 個以外の 110 個の近似固有対の残差のノルムの大きさは 10^{-6} 程度で、近似固有値の相対精度は少なくとも $7 \sim 8$ 桁程度あることが分かる。IT0 の下側にある IT1、IT2 はほとんど重なっているが、フィルタ対角化法で得られた近似固有対を Rayleigh 商逆反復によりそれぞれ 1 回、2 回改良を施したものである。IT0 で大きい残差を示している偽の固有対 1 個は逆反復で改良すると、固有値が変化して区間 $[50, 100]$ の外にはみ出たのでグラフの IT1、IT2 のプロットには含まれていない。IT1、IT2 のプロットに含まれている逆反復で改良した近似固有対 110 個については、残差のノルムが 10^{-10} 程度であることから、固有値の相対精度は少なくとも 12 桁程度はあることが分かる。

経過時間は、1 スレッド実行の場合はフィルタ対角化全体が 240.1 秒、逆反復 2 回が 378.2

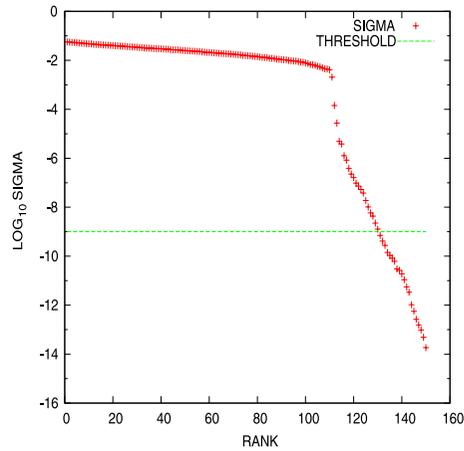


図 2 例題 1, 従来の方法: 特異値分布 ($m = 150$, 閾値 10^{-9})

Fig. 2 Exam-1, by the current method: Distribution of singular values ($m = 150$, threshold = 10^{-9}).

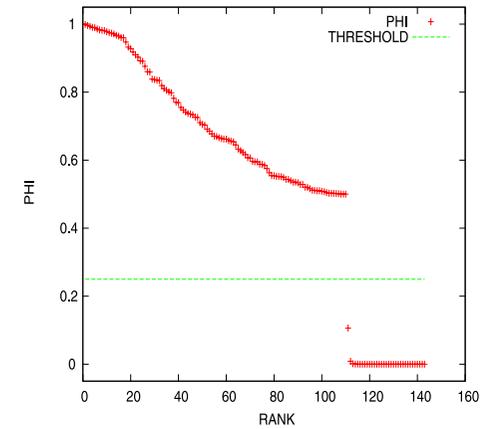


図 4 例題 1, 新しい方法: ϕ の値の分布 ($m = 150$, 閾値 0.25)

Fig. 4 Exam-1, by the present method: Distribution of ϕ ($m = 150$, threshold = 0.25).

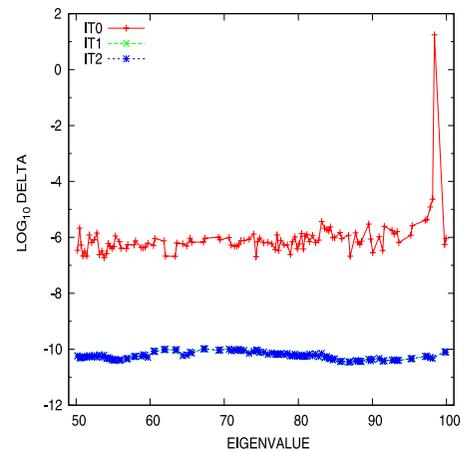


図 3 例題 1, 従来の方法: 近似固有対の残差のノルム (区間 $[50, 100]$, $m = 150$)

Fig. 3 Exam-1, by the current method: Norms of residuals of approximated eigenpairs (interval $[50, 100]$, $m = 150$).

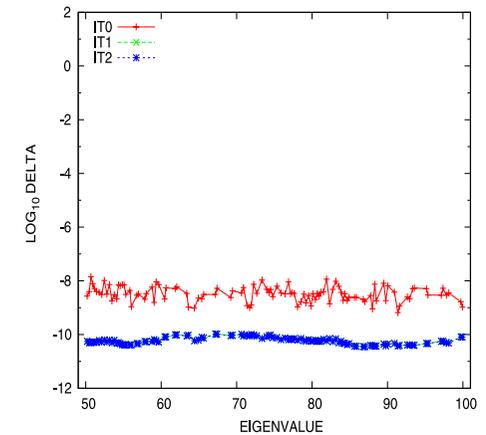


図 5 例題 1, 新しい方法: 近似固有対の残差のノルム (区間 $[50, 100]$, $m = 150$)

Fig. 5 Exam-1, by the present method: Norms of residuals of approximated eigenpairs (interval $[50, 100]$, $m = 150$).

秒で、4 スレッド並列実行の場合はフィルタ対角化全体が 70.4 秒、逆反復 2 回が 108.3 秒であった。経過時間の内訳を表 1 に示す。

新しい方法 (例題 1)

グラフ (図 4) に「縮小されたフィルタ作用素の固有値」 ϕ の分布を示す。 $g_{\text{pass}} = 0.5$ であり、 10^{-15} 程度である c_{∞} は省いたので、通過帯域での ϕ の下限は約 0.5000000 である。 ϕ の閾値を 0.25 とすると、固有対の個数と等しい 110 個のベクトルの組を得た。その 110 個のベクトルからなる基底に対して subspace 法を適用したところ、得られた近似固有対の固有値はすべて区間 $[50, 100]$ 内にあった。得られた近似固有対についてグラフ (図 5) に、横軸に近似固有値を、縦軸には残差のノルム Δ の値をプロットした。グラフ中の IT0 がフィルタ対角化法による残差のノルムである。その下側にある IT1, IT2 はほとんど重なっているが、フィルタ対角化法で得られた近似固有対を Rayleigh 商逆反復によりそれぞれ 1 回、2 回改良を施したものである。この IT0 のグラフからフィルタ対角化法による近似固有対の残差のノルムの大きさは $10^{-9} \sim 10^{-8}$ 程度で、固有値の相対精度は少なくとも 10 桁 ~ 11 桁程度あることが分かる。IT1, IT2 のグラフからは、逆反復で改良された近似固有対の固有値の相対精度は少なくとも 12 桁程度あることが分かる。

経過時間は、1 スレッド実行の場合はフィルタ対角化が 226.3 秒、逆反復 2 回が 320.2 秒であり、4 スレッド並列実行の場合はフィルタ対角化が 66.2 秒、逆反復 2 回が 92.6 秒であった。経過時間の内訳を表 2 に示す。

5.2 例題 2 (次数 $n = 16$ の elliptic フィルタ)

この例題の係数行列 A, B はどちらも次数 $N = 5 \times 10^5$ の帯行列で、半帯幅は $h = 100$ である。帯内 ($|p - q| \leq h$) の行列要素は $A_{p,q} = \max(p, q) - 1$, $B_{p,q} = 1/(p + q - 1) + \delta_{p,q}$ とした。ただし $p, q = 1, 2, \dots, N$ で、 $\delta_{p,q}$ は Kronecker 記号である。この固有値問題の固有対で固有値が区間 $[100, 150]$ 内のものは 99 個である。

入力ベクトルの組 X は、乱数ベクトル $m = 150$ 個を B -正規直交化して作成した。フィルタは次数 $n = 16$ の elliptic で、 $\mu = 1.1$, $g_{\text{pass}} = 0.5$ (tight), $g_{\text{stop}} = 10^{-14}$ である。フィルタの係数 c_{∞} は 10^{-14} 程度と小さいので省いた。

従来の方法 (例題 2)

B -SVD の特異値分布をグラフ (図 6) に示す。特異値の閾値 10^{-9} を用いて切断すると 104 個の特異ベクトルの組が得られた。その 104 個の特異ベクトルの組からなる基底に対して subspace 法を適用したところ、得られた近似固有対のうちで固有値が区間 $[100, 150]$ 内にあるものはちょうど 99 個であった。グラフ (図 7) に、これら 99 個の近似対について、

表 1 例題 1, 従来の方法: 経過時間 (秒) の内訳

Table 1 Exam-1, by the current method: The breakdown of elapsed times in seconds.

	1 スレッド計算	4 スレッド計算
フィルタ対角化全体	240.08	70.35
乱数ベクトル生成	0.33	0.33
B -正規直交化	2.49	1.10
フィルタの適用	210.85	61.27
B -SVD による基底作成	16.83	5.21
Subspace 法	9.58	2.44
Rayleigh 商逆反復 2 回	378.22	108.30

表 2 例題 1, 新しい方法: 経過時間 (秒) の内訳

Table 2 Exam-1, by the present method: The breakdown of elapsed times in seconds.

	1 スレッド計算	4 スレッド計算
フィルタ対角化全体	226.25	66.17
乱数ベクトル生成	0.33	0.33
B -正規直交化	2.49	1.10
フィルタの適用	211.91	61.24
不変部分空間の基底作成	4.58	1.78
Subspace 法	6.95	1.74
Rayleigh 商逆反復 2 回	320.25	92.61

横軸に近似固有値、縦軸に残差のノルム Δ の値をプロットした。グラフ中の IT0 がフィルタ対角化法による残差のノルムである。その下側にある IT1, IT2 はほとんど重なっているが、フィルタ対角化法で得られた近似固有対を Rayleigh 商逆反復によりそれぞれ 1 回、2 回改良を施したものである。これら 99 個の近似固有対の残差のノルムの大きさは IT0 のグラフでは $10^{-8} \sim 10^{-6}$ 程度で、フィルタ対角化法による近似固有値の相対精度は少なくとも 8 桁程度あること、また IT1, IT2 のグラフでは 10^{-9} 程度なので、改良後の相対精度は少なくとも 11 桁程度あることが分かる。

経過時間は、1 スレッド実行の場合はフィルタ対角化が 1,530 秒、逆反復 2 回が 1,515 秒であり、4 スレッド並列実行の場合はフィルタ対角化が 448 秒、逆反復 2 回が 428 秒であった。経過時間の内訳を表 3 に示す。

新しい方法 (例題 2)

β の固有値分解を 2.2×10^{-14} を閾値として切断すると、次元は 106 となった。「縮小されたフィルタ作用素の固有値」 ϕ の分布をグラフ (図 8) に示す。 $g_{\text{pass}} = 0.5$ から、通過帯域

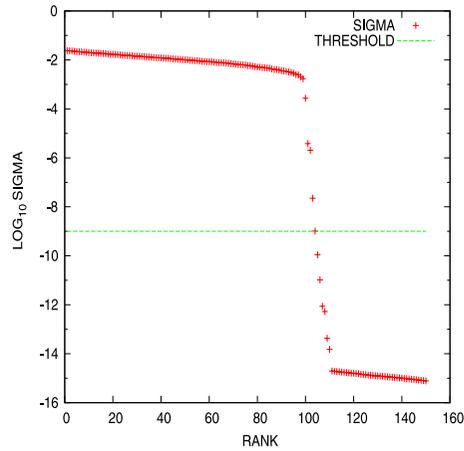


図 6 例題 2, 従来の方法: 特異値分布 ($m = 150$, 閾値 10^{-9})

Fig. 6 Exam-2, by the current method: Distribution of singular values ($m = 150$, threshold = 10^{-9}).

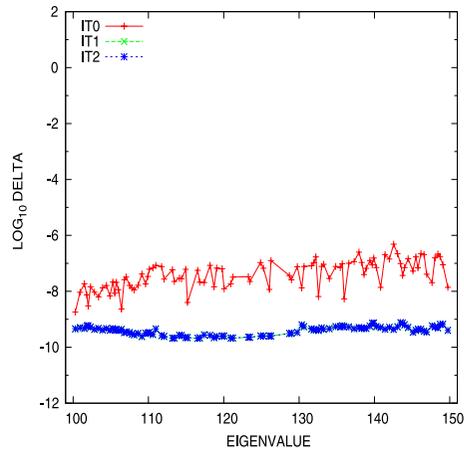


図 7 例題 2, 従来の方法: 近似固有対の残差のノルム (区間 $[100, 150]$, $m = 150$)

Fig. 7 Exam-2, by the current method: Norms of residuals of approximated eigenpairs (interval $[100, 150]$, $m = 150$).

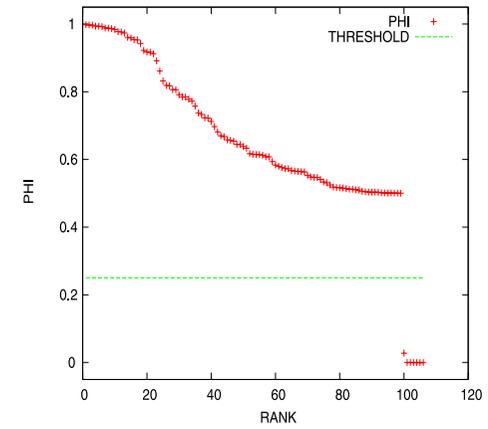


図 8 例題 2, 新しい方法: ϕ の値の分布 ($m = 150$, 閾値 0.25)

Fig. 8 Exam-2, by the present method: Distribution of ϕ ($m = 150$, threshold = 0.25).

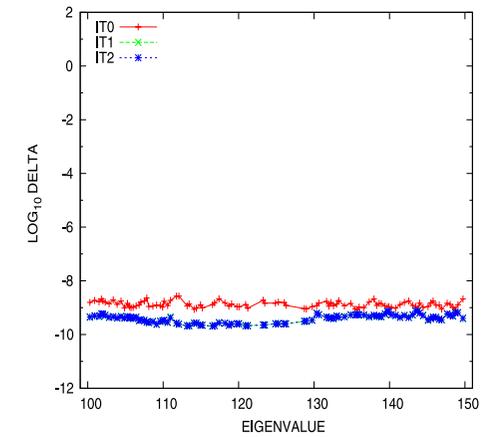


図 9 例題 2, 新しい方法: 近似固有対の残差のノルム (区間 $[100, 150]$, $m = 150$)

Fig. 9 Exam-2, by the present method: Norms of residuals of approximated eigenpairs (interval $[100, 150]$, $m = 150$).

表 3 例題 2, 従来の方法: 経過時間 (秒) の内訳

Table 3 Exam-2, by the current method: The breakdown of elapsed times in seconds.

	1 スレッド計算	4 スレッド計算
フィルタ対角化全体	1,530.43	447.74
乱数ベクトル生成	1.64	1.65
B-正規直交化	12.32	5.44
フィルタの適用	1,404.35	407.95
B-SVD による基底作成	80.90	24.97
Subspace 法	31.23	7.74
Rayleigh 商逆反復 2 回	1,515.17	428.44

表 4 例題 2, 新しい方法: 経過時間 (秒) の内訳

Table 4 Exam-2, by the present method: The breakdown of elapsed times in seconds.

	1 スレッド計算	4 スレッド計算
フィルタ対角化全体	1,477.86	429.42
乱数ベクトル生成	1.64	1.65
B-正規直交化	12.31	5.45
フィルタの適用	1,412.42	407.03
不変部分空間の基底作成	22.13	8.09
Subspace 法	29.36	7.21
Rayleigh 商逆反復 2 回	1,443.39	414.44

での ϕ の下限は約 0.5000000 である。 ϕ の閾値を 0.25 とすると、ちょうど固有対の個数に等しい 99 個の B-正規直交ベクトルからなる基底を得た。その基底に対して subspace 法を適用して得られた近似固有対の固有値は、すべて区間 [100, 150] 内にあった。グラフ (図 9) には得られた近似固有対について、横軸に近似固有値を、縦軸には残差のノルム Δ の値をプロットした。グラフ中の IT0 がフィルタ対角化法による残差のノルムである。その下側にある IT1, IT2 はほとんど重なっているが、フィルタ対角化法で得られた近似固有対を Rayleigh 商逆反復によりそれぞれ 1 回、2 回改良を施したものである。この IT0 のグラフから、フィルタ対角化法による近似固有対の残差のノルムの大きさは 3×10^{-9} 以下で、フィルタ対角化法による近似固有値の相対精度は少なくとも 10 桁程度あることが分かる。

経過時間は、1 スレッド実行の場合はフィルタ対角化が 1,478 秒、逆反復 2 回が 1,443 秒であり、4 スレッド並列実行の場合はフィルタ対角化が 429 秒、逆反復 2 回が 414 秒であった。経過時間の内訳を表 4 に示す。

5.3 例題 3 (低次 $n = 4$ の Chebyshev フィルタ)

この例題の係数行列 A, B はどちらも次数 $N = 5 \times 10^5$ の帯行列で、半帯幅は $h = 100$ である。帯内 ($|p - q| \leq h$) の行列要素は、 $A_{p,q} = \max(p, q) - 1$, $B_{p,q} = 1/(p + q - 1) + \delta_{p,q}$ と与えた。ただし、 $p, q = 1, 2, \dots, N$ で、 $\delta_{p,q}$ は Kronecker 記号である。この固有値問題の固有対で固有値が区間 [150, 200] 内のものは 85 個ある。

入力ベクトルの組 X は、乱数ベクトル $m = 200$ 個を B-正規直交化して作成した。フィルタは次数 $n = 4$ の Chebyshev で、 $\mu = 2$, $g_{\text{pass}} = 0.5$ (tight), $g_{\text{stop}} = 1.1 \times 10^{-4}$ とした。係数 c_∞ の値は Chebyshev フィルタではつねに零である。

従来の方法 (例題 3)

B-SVD の特異値分布をグラフ (図 10) に示す。特異値の閾値を 10^{-3} とすると切断の結果 96 個の B-特異ベクトルの組が得られた (この閾値は結果が比較的良好になるように選んだものである)。その 96 個の B-特異ベクトルの組を基底として subspace 法を適用して得られた近似固有対 96 個のうちで、区間 [150, 200] 内に固有値があるものは 86 個であった。これは固有対の正しい個数よりも 1 個多いが、近似固有対に逆反復を 1 回適用すると、区間の上端に非常に近い固有値 199.98426 を持つ対 1 個が改良されて固有値が少し変化して 200.03413 と区間外に出たので、区間内に固有値を持つ近似固有対は 85 個となった。グラフ (図 11) に、これら 86 (85) 個の近似固有対について、横軸に近似固有値、縦軸に残差のノルム Δ の値をプロットした。グラフ中の IT0 がフィルタ対角化法による近似固有対 86 個の残差のノルムで、IT0 の下側の IT1, IT2 はそれぞれフィルタ対角化法で得られた 86 個の近似固有対を Rayleigh 商逆反復により 1 回、2 回改良を施して得られた近似固有対のうちで固有値が区間 [150, 200] 内の 85 個分のプロットである。IT0 で示されたフィルタ対角化法で得られた 86 個の近似固有対の残差のノルムの大きさが 10^0 程度で、近似固有値の相対精度は少なくとも 2 桁程度あることが分かる。

経過時間は、1 スレッド実行の場合はフィルタ対角化が 618 秒、逆反復 2 回が 1,400 秒であり、4 スレッド並列実行の場合はフィルタ対角化が 195 秒、逆反復 2 回が 396 秒であった。経過時間の内訳を表 5 に示す。

新しい方法 (例題 3)

β の最小固有値は 1.1×10^{-9} で、 β の固有値の切断は閾値 2.2×10^{-14} では生じなかった。「縮小されたフィルタ作用素の固有値」 ϕ の分布を減少順にグラフ (図 12) に示す。 $g_{\text{pass}} = 0.5$, $c_\infty = 0$ から、通過帯域での ϕ の下限の値は 0.5 である。 $\phi_{85} = 0.5000367$, $\phi_{86} = 0.4891279$, $\phi_{87} = 0.4238465$, $\phi_{88} = 0.2716056$, $\phi_{89} = 0.2230796$ なので、 ϕ の

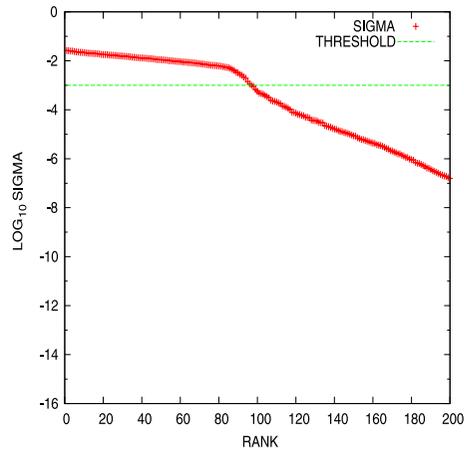


図 10 例題 3, 従来の方法: 特異値分布 ($m = 200$, 閾値 10^{-3})

Fig. 10 Exam-3, by the current method: Distribution of singular values ($m = 200$, threshold = 10^{-3}).

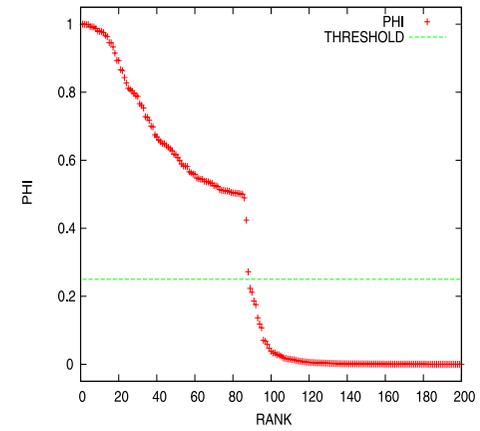


図 12 例題 3, 新しい方法: ϕ の値の分布 ($m = 200$, 閾値 0.25)

Fig. 12 Exam-3, by the present method: Distribution of ϕ ($m = 200$, threshold = 0.25).

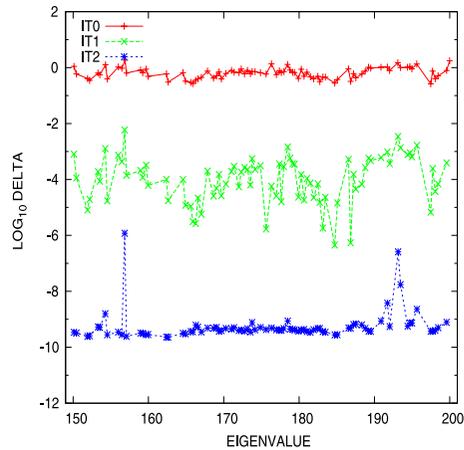


図 11 例題 3, 従来の方法: 近似固有対の残差のノルム (区間 [150, 200], $m = 200$)

Fig. 11 Exam-3, by the current method: Norms of residuals of approximated eigenpairs (interval [150, 200], $m = 200$).

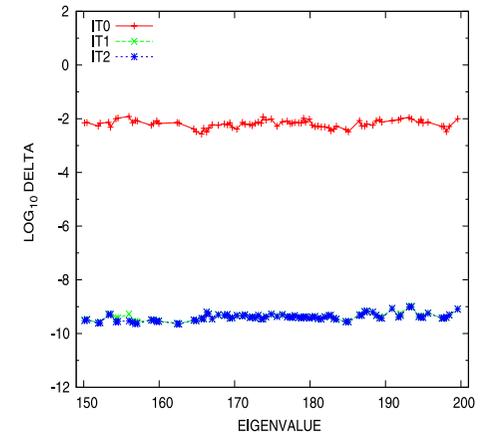


図 13 例題 3, 新しい方法: 近似固有対の残差のノルム (区間 [150, 200], $m = 200$)

Fig. 13 Exam-3, by the present method: Norms of residuals of approximated eigenpairs (interval [150, 200], $m = 200$).

61 対称一般固有値問題のフィルタ作用素を用いた不変部分空間の近似構成

表 5 例題 3, 従来の方法: 経過時間 (秒) の内訳

Table 5 Exam-3, by the current method: The breakdown of elapsed times in seconds.

	1 スレッド計算	4 スレッド計算
フィルタ対角化全体	617.72	195.35
乱数ベクトル生成	2.19	2.20
B-正規直交化	23.62	9.20
フィルタの適用	438.67	131.78
B-SVD による基底作成	125.56	45.30
Subspace 法	27.68	6.87
Rayleigh 商逆反復 2 回	1,399.68	395.92

表 6 例題 3, 新しい方法: 経過時間 (秒) の内訳

Table 6 Exam-3, by the present method: The breakdown of elapsed times in seconds.

	1 スレッド計算	4 スレッド計算
フィルタ対角化全体	524.99	206.38
乱数ベクトル生成	2.20	2.19
B-正規直交化	23.63	9.20
フィルタの適用	442.55	170.49
不変部分空間の基底作成	32.26	18.43
Subspace 法	24.36	6.08
Rayleigh 商逆反復 2 回	1,285.07	360.83

85 番目までが通過帯域に対応している．単純に ϕ の閾値を 0.25 に設定すると，真の固有対の個数よりも 3 個多い 88 個の B-正規直交ベクトルの組が得られた．その 88 個のベクトルからなる基底に subspace 法を適用して得られた元の GEVP の近似固有対のうちで近似固有値が区間 [150, 200] 内にあるものはちょうど 85 個で，残り 3 個は近似固有値が区間の外にあった．得られた近似固有対のうち固有値が区間 [150, 200] 内のものについてグラフ (図 13) に横軸に近似固有値，縦軸に残差のノルム Δ の値をプロットした．グラフ中の IT0 はフィルタ対角化法による残差のノルムである．IT0 の下側にある IT1, IT2 はほとんど重なっているが，フィルタ対角化法で得られた近似固有対を Rayleigh 商逆反復によりそれぞれ 1 回，2 回改良を施したものである．この IT0 のグラフからフィルタ対角化法による近似固有対の残差のノルムの大きさは 10^{-2} 程度で，固有値の相対精度は少なくとも 4 桁程度はあることが分かる．また IT1, IT2 のグラフから逆反復で改良した近似固有対の固有値の相対精度が少なくとも 11 桁程度あることも分かる．

経過時間は，1 スレッド実行の場合はフィルタ対角化が 525 秒，逆反復 2 回が 1,285 秒で

あり，4 スレッド並列実行の場合はフィルタ対角化が 206 秒，逆反復 2 回が 361 秒であった．経過時間の内訳を表 6 に示す．

5.4 実験結果に対する若干の考察

例題 1~3 の実験において，フィルタ対角化法 (Rayleigh 商逆反復は含めない) により得られた区間内に固有値がある固有対の残差が新しい方法では従来法に比べて小さい傾向が見られるが，その理由としてはまず従来法では特異値の切断に用いる閾値を安全側に大きめとしていることがあげられるが，他の理由として考えられることは，新しい方法では問題の対称性を利用して \mathcal{F} をベクトルの組 X に 1 度適用するだけで実質的にフィルタ作用素 \mathcal{F} を 2 度重ねて適用しているのと同様の効果を得ていることがあげられる．

作用素 \mathcal{F} の固有値問題を部分空間 X 上で近似して解く場合：

フィルタ作用素 \mathcal{F} の固有値方程式 $\mathcal{F}\mathbf{v} = \phi\mathbf{v}$ を X の張る空間に制限して $\mathbf{v} = X\mathbf{u}$ において，方程式の残差と X が計量 B で直交する条件を要請すると， $(X, \mathcal{F}X\mathbf{u} - \phi X\mathbf{u})_B = 0$ から， $X^T B \mathcal{F} X \mathbf{u} = \phi X^T B X \mathbf{u}$ を得るが， $X^T B X = I_m$ ， $X^T B \mathcal{F} X = \beta$ であるから結局，実対称行列 β の標準固有値問題 $\beta\mathbf{u} = \phi\mathbf{u}$ を得る．これを解いて固有対 (ϕ, \mathbf{u}) を得たら， \mathcal{F} の近似対は $(\phi, X\mathbf{u})$ となる．

作用素 \mathcal{F} の固有値問題を部分空間 Y 上で近似して解く場合：

フィルタ作用素 \mathcal{F} の固有値方程式 $\mathcal{F}\mathbf{v} = \phi\mathbf{v}$ を $Y = \mathcal{F}X$ の張る空間に制限して $\mathbf{v} = Y\mathbf{u}$ とおき，方程式の残差と X が計量 B で直交する条件を要請すれば， $(X, \mathcal{F}Y\mathbf{u} - \phi Y\mathbf{u})_B = 0$ から， $X^T B \mathcal{F} Y \mathbf{u} = \phi X^T B Y \mathbf{u}$ を得るが， $X^T B \mathcal{F} Y = \alpha$ ， $X^T B Y = \beta$ であるから結局，実対称行列 α, β の一般固有値問題 $\alpha\mathbf{u} = \phi\beta\mathbf{u}$ を得る．これを解いて固有対 (ϕ, \mathbf{u}) を得たら， \mathcal{F} の近似対は $(\phi, Y\mathbf{u})$ となる．

このように，フィルタ作用素 \mathcal{F} の固有値問題を近似して解くための部分空間を，ランダムな B-正規直交ベクトルの組 X が張る空間内ではなくて， X にフィルタを 1 回適用することであらかじめ不要な固有ベクトルを含む割合を減らした Y が張る空間内とすることで，より高い精度が得られると考えられる．

6. ま と め

フィルタ対角化法では，通過帯域 $[a, b]$ のフィルタ \mathcal{F} を用意する．次にフィルタへの入力ベクトルの組として，十分に多い m 個の乱数ベクトルを B-正規直交化した組 X を作る．その X の各ベクトルに \mathcal{F} を作用させて得られる組 Y を作り，フィルタからの出力ベクトルの組とする．組 Y の各ベクトルの線形結合を適切に選んで部分空間の基底 Z を構成し，

subspace 法に Z を与えて元の固有値問題の近似固有対を求める．

従来の方法では，計量 B で Y を特異値分解し，特異値が閾値以上の特異ベクトルを集めることで B -正規直交である部分空間の基底 Z を構成してきた．フィルタの性能が良ければ通常は満足に働くが，特異値の切断の閾値の合理的な決定法は見出しにくい．切断付きの SVD では，subspace 法に与える基底 Z を元の一般化固有値問題の係数 A とは無関係に，出力ベクトルの組 Y と計量 B だけを用いて構成するので，一般には Z の張る空間は不変部分空間の良い近似にはならず，subspace 法を Z に適用して得られる近似固有対には残差の大きい偽の固有対が混ざる．偽の固有対の固有値は通常は区間外にあるが，区間内に現れることもある．

今回の新しい方法では， Y の張る空間内において固有値の区間 $[a, b]$ と対応する不変部分空間を近似して，組 Z がその B -正規直交基底となるように構成する．そのようにして得られた不変部分空間の近似を張る基底 Z に subspace 法を適用することで元の一般化固有値問題の近似固有対を求める．

レゾルベントの線形結合を通過帯域 $[a, b]$ のフィルタ \mathcal{F} として採用した場合には，GEVP の固有対 (λ, \mathbf{v}) に対応して，作用素 \mathcal{F} は固有対 (ϕ, \mathbf{v}) を持つ．ただし， $\phi = f(\lambda)$ はベクトル \mathbf{v} のフィルタによる伝達率で， f は伝達関数である．この性質を利用することで，subspace 法に与えるための（固有値が $[a, b]$ の）不変部分空間を近似する空間の基底 Z が構成できる．

フィルタ作用素 \mathcal{F} の固有値方程式 $\mathcal{F}\mathbf{v} = \phi\mathbf{v}$ を Y の張る空間内に \mathbf{v} を制限して解いた近似固有対 (ϕ, \mathbf{v}) のうちで， ϕ の値が（通過帯域 $\lambda \in [a, b]$ に対する $f(\lambda)$ の値域である）区間 $[g_{\text{pass}}, 1]$ の近傍にあるベクトル \mathbf{v} を集めて（計量 B で正規化して）基底 Z とする．

数値実験の結果からは，新しい方法は従来の方法よりもかなり優れている．特にフィルタの弁別性能が低くて，遷移帯域に固有値が多く分布する場合でも，それに対応して入力ベクトルの個数を多くとれば，透過帯域内の固有値に対応する不変部分空間の基底の近似を適切に構成できる．

参 考 文 献

- 1) Toledo, S. and Rabani, E.: Very Large Electronic Structure Calculations Using an Out-of-Core Filter-Diagonalization Method, *J. Comput. Phys.*, Vol.180, No.1, pp.256–269 (2002).
- 2) Sakurai, T. and Sugiura, H.: A projection method for generalized eigenvalue problems using numerical integration, *J. Comp. Appl. Math.*, Vol.159, pp.119–128

(2003).

- 3) Polizzi, E.: Density-matrix-based algorithm for solving eigenvalue problems, *Phys. Rev. B*, Vol.79, No.11, p.115112 (6pages) (2009).
- 4) Ikegami, T., Tadano, H., Umeda, H. and Sakurai, T.: Hierarchical parallel algorithm to solve large generalized eigenproblems, HPCS2010 論文集, pp.107–114 (2010).
- 5) 村上 弘: 帯対称定値一般固有値問題のフィルタ対角化法の実験, 情報処理学会研究報告, Vol.2007-HPC-110, No.6, pp.31–36 (2007).
- 6) 村上 弘: レゾルベントの線形結合によるフィルタ対角化法, 情報処理学会論文誌: コンピューティングシステム (ACS21), Vol.49, No.SIG2, pp.66–87 (2008).
- 7) 村上 弘: 固有値が指定された区間内にある固有対を解くための対称固有値問題用のフィルタの設計, 情報処理学会論文誌: コンピューティングシステム (ACS31), Vol.3, No.3, pp.1–21 (2010).
- 8) Daniels, R.: *Approximation Methods for Electronic Filter Design*, McGraw-Hill (1974).
- 9) Lutovac, M., Tošić, D. and Evans, B.: *Filter Design for Signal Processing*, Prentice Hall (2001).
- 10) 村上 弘: フィルタ対角化法の帯域通過フィルタの最適化, 情報処理学会研究報告, Vol.2010-HPC-124, No.3, pp.1–8 (2010).
- 11) 村上 弘: 楕円フィルタによる実対称定値一般固有値問題のフィルタ対角化法の実験, 情報処理学会研究報告, Vol.2010-HPC-125, No.1, pp.1–10 (2010).
- 12) Rutishauser, H.: The Jacobi method for real symmetric matrices, *Numer. Math.*, Vol.9, No.1, pp.1–10 (1966).
- 13) Martin, R. and Wilkinson, J.: Solution of Symmetric and Unsymmetric Band Equations and the Calculations of Eigenvectors of Band Matrices, *Numer. Math.*, Vol.9, No.4, pp.279–301 (1967).
- 14) 村田健郎, 小国 力, 唐木幸比古: スーパーコンピュータ科学技術計算への適用, 丸善 (1985).
- 15) 小国 力 (編), 村田健郎, 三好俊郎, ドンガラ, J.J., 長谷川秀彦 (著): 行列計算ソフトウェア, 丸善 (1991).

付 録

A.1 入力ベクトルの個数 m の選択法の考察

算法の前提とした「入力ベクトルの個数 m を十分に大きくとる」ことについて，若干の考察を試みる．

フィルタの通過帯域 $\lambda \in [a, b]$ が正規化座標 $t \in [-1, 1]$ に対応しているとき，関係

$\lambda = \frac{1}{2}(a+b) + \frac{1}{2}(b-a)t$ を用いて, $t = -\mu$ に $\lambda = a'$ を, $t = \mu$ に $\lambda = b'$ をそれぞれ対応させる. 区間 $[a', b']$ はフィルタの通過帯域 $[a, b]$ の両側に遷移帯域を加えて上げたものである. 以下で, 固有値が $[a, b]$ にある固有対の個数を \mathcal{N} とし, 固有値が $[a', b']$ にある固有対の個数を \mathcal{N}' とする.

m 個の B -正規直交ベクトルの組を X とし, フィルタを X に作用させた出力の組を Y とする. いま「フィルタの伝達率は阻止帯域では数値的に微小で零と見なせる」が十分に成立していると仮定すると, Y の計量 B での数値的な階数は \mathcal{N}' を超えない.

フィルタの伝達率は最大値が 1 (恒等演算子の項を省く場合には 1 よりわずかに小さい) としているので, たとえば阻止帯域における伝達率の上限値が浮動小数点数のマシンepsilon (IEEE 754 の倍精度 64 bit 浮動小数点数では 2.22×10^{-16}) 程度のたとえば 10^{-14} や 10^{-15} などであればほぼ零と見なしてよいし, マシンepsilon の程度よりもかなり大きな値, たとえば 10^{-12} や 10^{-8} などであってもそれが十分に小さい値であると見なせる場合には仮想的にそのレベルの演算精度でフィルタを作用させる計算を行ったと思うことができる.

$m < \mathcal{N}'$ の場合には, Y の m 個のベクトルの線形結合により固有値が $[a', b']$ にある固有ベクトル (全部で \mathcal{N}' 個) を表すことは自由度が不足しているので不可能である. このことから, 求めようとしている固有値が $[a, b]$ にある固有対もうまく近似できないかまたは近似の精度が出ない可能性がある.

乱数ベクトルを元にして B -正規直交化で得られた X は, ほとんどの場合に一般的 (generic) な状況にあるので, $\mathcal{N}' \leq m$ であれば計量 B での Y の数値的な階数はちょうど \mathcal{N}' になる. 実際には m の値はちょうど \mathcal{N}' にしないで, 少し大きな値にするのがよい. その理由は乱数を用いて一般的 (generic) 状況を作り出しているが, その数値的な generic の程度 (一種の条件数) があまり良くない状況が生じる確率を下げるためである.

フィルタの遷移帯域に固有値を持つ固有対の個数が多い場合に (これは伝達関数の特性がどれほど急峻であっても固有値分布が遷移帯域に密集していれば遭遇する), $\mathcal{N}' \leq m$ とせずに $\mathcal{N} \leq m < \mathcal{N}'$ とした場合には, 区間 $[a', b']$ にある固有値を重複を込めて考え, 固有値に対応するフィルタの伝達率を減少順に並べたとするとき, 第 $m+1$ 番目の伝達率が (伝達関数の最大値 1 に比べて) 「数値的に微小で零と見なせる」かどうかによって近似対の精度が支配される.

さて, 以上の考察に従うならば, m の値を固有値が $[a', b']$ にある固有対の個数 \mathcal{N}' よりも少し大きくなるようにとればよい. たとえば以下のような方法が考えられる.

- (1) 何らかの根拠が経験により, \mathcal{N}' がその上限が既知の場合には, それに基づいて $\mathcal{N}' < m$ となる少し大きめの m を設定する. そうしてこの m の値を用いても以降の計算で必要となる記憶量や演算量が過大にはならないと判断したら計算を進める (過大になる場合には, 元の区間 $[a, b]$ を分割して扱うことを検討する).
- (2) \mathcal{N}' の値について予備知識がないか, 推定値に誤差がともなう場合には, 必要な記憶量や演算量が過大とならない範囲での m について, たとえば m を少しずつ段階的に増やして様子を見ながら計算を行うことが可能である (入力ベクトルの個数 m を増加させる場合は, 計算をまったく最初からやり直すのではなくて, B -正規直交ベクトルの組 X を増加分と対応して拡張分の列だけを計算し, その X の拡張分の列にフィルタを作用させて組 Y の拡張分の列を作る, などとなるべく処理を追加的に行うことが可能である). 一般的な状況にある X をフィルタに作用させても得られた Y の数値的な階数が m となってしまう数値的な階数低下が十分に起きない状況や, 得られた近似固有対の誤差評価で, たとえば残差が十分に小さくない状況であれば, 使用した m の値は十分でないとして m をさらに増加させる (m を増やしつづけてある限度を超えたら, 区間 $[a, b]$ の再分割を検討する).
- (3) 実対称定値一般固有値問題では固有値が指定した区間にある固有対の個数は Sylvester の慣性律を用いて求めることができ, 2 分法などの区間縮小法で固有値を求めるのに用いられている¹⁵⁾. 実対称行列 A, B に対して, 実数のシフト λ を与えて実対称行列 $C = A - \lambda B$ を作り, たとえば修正 Cholesky 法で $C = LDL^T$ と分解できるなら, 固有値が区間 $[a, b]$ にある固有対の個数 \mathcal{N} はシフト $\lambda = a, b$ における対角行列 D の符号数の差として求められる. 同様に固有値が区間 $[a', b']$ にある固有対の個数 \mathcal{N}' も D のシフト $\lambda = a', b'$ での符号数の差として求められる (ただし両端のシフトで C の修正 Cholesky 法による分解が破綻せずにできたとする). この方法は A, B が幅の狭い帯行列であれば高速に実行しうる.

A.2 固有値方程式 $\alpha u = \phi \beta u$ の固有ベクトル展開を用いた考察

縮小された固有値方程式 $\alpha u = \phi \beta u$ について理解を深めるため, 固有ベクトル展開を用いて状況の分析を試みる.

いま元の固有値問題の固有対にその伝達率 $\phi = f(\lambda)$ が広義単調減少順となるように番号をつけておくものとする. そうして第 j 番目の固有ベクトルを第 j 列に並べた N 次の正方行列を V とする. 固有ベクトルは B -正規直交にとるので $V^T B V = I_N$ である. また元の固有値問題の固有値を固有ベクトルの順番に対角に並べた N 次対角行列を Λ とすれば

$AV = BVA$ が成り立つ．さらに第 j 番目の固有ベクトルの伝達率 $\phi^{(j)}$ を第 j 対角要素に並べた N 次対角行列を Φ とおくと $FV = V\Phi$ である．フィルタへの m 個の入力ベクトルの組 X は $N \times m$ 行列で， B -正規直交化されていれば $X^T B X = I_m$ である．固有ベクトルによる X の展開係数を表す $N \times m$ 行列 C が存在して $X = VC$ と書けて， X の B -正規直交性から $C^T C = I_m$ で， C はフルランクである． $Y = FX = FVC = V\Phi C$ である．

「縮小された \mathcal{F} の固有値問題」の係数である m 次対称行列は $\alpha = Y^T B Y = C^T \Phi^2 C$ ， $\beta = X^T B Y = C^T \Phi C$ と書ける．いま，フィルタの伝達特性から $(m+1)$ 番目以降の伝達率 $\phi^{(j)}$ の大きさが十分に無視できるように m を大きくとってあったと仮定する（たとえば阻止帯域での伝達関数の上限値 g_{stop} が無視できる小さい値であると見なせるときに，固有値が通過帯域または遷移帯域のどちらかにある固有対の個数よりも m を大きくとる）．そのとき，行列 Φ をその m 次の首位行列 $\hat{\Phi}$ までで切断する近似を行う．それに対応して C を最初の m 行までで切断した m 次正方形行列 \hat{C} を用いると， $\alpha \approx \hat{C}^T \hat{\Phi}^2 \hat{C}$ ， $\beta \approx \hat{C}^T \hat{\Phi} \hat{C}$ と近似できる．さらに \hat{C} の条件はあまり悪くないと仮定する（乱数を用いて作った X に由来している \hat{C} の条件は通常はあまり悪くない．そうでなければ乱数ベクトルをとり直すか個数 m を増やすことで改善が期待できる）．すると「縮小された \mathcal{F} の固有値問題」 $\alpha \mathbf{u} = \phi \beta \mathbf{u}$ は，この Φ の切断近似の下で $\hat{C} \mathbf{u} = \mathbf{w}$ とおくと，対角行列を係数とする一般固有値問題 $\hat{\Phi}^2 \mathbf{w} = \phi \hat{\Phi} \mathbf{w}$ となり，その第 j 番目の固有値は $\hat{\Phi}$ の第 j 対角要素の値である $\phi^{(j)}$ となることが分かる．ただし， C や \hat{C} は既知量ではないので，固有値方程式 $\alpha \mathbf{u} = \phi \beta \mathbf{u}$ をそのままなんらかの数値的方法で解くことになる．その際には丸め誤差の影響で固有値 ϕ の値が微小なものに対する近似解は良い精度が得られないであろう．しかし今の場合には固有値の範囲が $\phi \in [g_{\text{pass}}, 1]$ であるような固有対 (ϕ, \mathbf{u}) だけを解けばよく， g_{pass} は実際には（た

たとえば 0.5 程度の）微小ではない値を設定するので，困難の程度は少ない．

付記：いま同様に元の固有値問題 $A \mathbf{v} = \lambda B \mathbf{v}$ を $\mathbf{v} \in \text{span}(Y)$ に制限する近似で得られる方程式 $A Y \mathbf{u} = \lambda B Y \mathbf{u}$ の残差と X との通常の内積から生じる方程式は，小さい次数 m の行列 $\gamma = X^T A Y$ ， $\beta = X^T B Y$ を係数とする対称一般固有値問題 $\gamma \mathbf{u} = \lambda \beta \mathbf{u}$ になる（ $\gamma = C^T \Lambda \Phi C$ が示せるので γ は対称である）． Φ の切断近似の下で， $\hat{C} \mathbf{u} = \mathbf{w}$ とおくと，対角行列を係数とする一般固有値問題 $\hat{\Lambda} \hat{\Phi} \mathbf{w} = \lambda \hat{\Phi} \mathbf{w}$ になり，その第 j 番目の固有値は「数学的には」元の固有値問題の固有値 $\lambda^{(j)}$ となることが分かる．ただし， C や \hat{C} は既知量ではないので，実際には対称一般固有値問題 $\gamma \mathbf{u} = \lambda \beta \mathbf{u}$ をそのままなんらかの数値的方法で解くことになる．その場合には $\phi^{(j)}$ の値が微小であればそれだけ対応する $\lambda^{(j)}$ の近似値は丸め誤差の影響で精度が失われうる（実際には $\gamma \mathbf{u} = \lambda \beta \mathbf{u}$ を単に解いただけでは得られた各近似固有解 (λ, \mathbf{u}) の j との対応も，対応する $\phi^{(j)}$ の値も分からない）．極端な場合には， λ の近似値は元の固有値問題の固有値から大きく外れた無意味な値にもなりうる．

(平成 23 年 1 月 28 日受付)

(平成 23 年 5 月 17 日採録)



村上 弘 (正会員)

現在，首都大学東京・数理情報科学専攻の准教授．現在の研究分野は，数学または科学の問題の，数値的あるいは記号的方法による，精度あるいは効率の良い解法やその並列化である．1992 年に北海道大学の理学博士号（化学第二学専攻）を取得．