

フィルタの反復適用による 実対称定値一般固有値問題の近似対の改良

村上 弘^{1,a)}

受付日 2018年12月4日, 採録日 2019年4月24日

概要: フィルタ対角化法を用いて実対称定値一般固有値問題の固有対で指定された区間に固有値があるものを近似して求める。フィルタはレゾルベントを用いて構成される。レゾルベントの作用は連立1次方程式を解いて実現されるが、直接法を用いて解くことにするならば、大規模問題を解く場合は行列分解の計算量と特に分解結果を保持する記憶量が制約になりがちである。フィルタとして単一のレゾルベントの多項式を採用すれば、用いるレゾルベントは1つで済むが、そのような構成のフィルタでは伝達関数の特性を十分良くすることができないため、得られる近似対の残差を十分小さくすることができない欠点がある。そこで今回の報告では、ベクトルの組に対して直交化とそれに続くフィルタの適用の組合せを数回反復する方法を示す。この一種の直交反復法により、伝達特性のあまり良くないフィルタを用いてランダムなベクトルの組から始めても、反復により改良された残差の十分小さい近似対が求まることを、実験で例を示す。

キーワード: フィルタ対角化, 固有値問題, レゾルベント, 反復改良

Improvements of Approximate Pairs of a Real Symmetric-definite Generalized Eigenproblem by Iterative Applications of a Filter

HIROSHI MURAKAMI^{1,a)}

Received: December 4, 20118, Accepted: April 24, 2019

Abstract: By using the filter diagonalization method, we solve approximate pairs of a generalized eigenproblem whose eigenvalues are in a specified interval. The filter is constructed by using resolvents. An action of a resolvent is given by solving a simultaneous linear equations. If we are to use some direct method for the solution of the simultaneous linear equations, then both the amount of arithmetic operations for matrix decompositions and especially the amount of storage requirements to hold the matrix factors tend to be the bottle-necks in solving eigenproblem of a large size. If we use a polynomial of a single resolvent for the filter, then the number of resolvents required is only one, which is an advantage. However, with such a construction, it is unable to make the property of the transfer function so good. Therefore magnitudes of residuals of approximate pairs obtained are not made so small, which is a disadvantage. In this report of our experiments, we iterate the following two steps only a few times: 1) an orthogonalization of a set of vectors, 2) an application of a filter to a set of vectors. By this iteration procedure, which is a kind of orthogonal iteration method, we can successfully obtain improved approximate eigenpairs whose residuals are sufficiently small, even by using a filter whose property of transfer is not so good and also started from a set of random vectors. We show several examples of experiments.

Keywords: filter diagonalization, eigenproblem, resolvent, iterative improvement

1. はじめに

実対称定値一般固有値問題 $A\mathbf{v} = \lambda B\mathbf{v}$ の固有対 (λ, \mathbf{v}) で固有値が指定された区間にあるものの近似をフィルタ対

¹ 首都大学東京数理科学専攻
Department of Mathematical Sciences, Tokyo Metropolitan
University, Hachioji, Tokyo 192-0397, Japan

a) mrkmmhrsh@tmu.ac.jp

角化法 (文献 [4], [5], [6], [9]) を用いて求めるとする.

フィルタの構成にはレゾルベント $\mathcal{R}(\rho) \equiv (A - \rho B)^{-1} B$ を用いる. それには以下の種類がある.

- (1) シフトが実数あるいは虚数のレゾルベントの線形結合の実部であるもの (1) (文献 [8], [9], [12], [14], [15], [17], [19]).

$$\mathcal{F} \equiv c_\infty I + \operatorname{Re} \sum_{j=1}^n c_j \mathcal{R}(\rho_j). \quad (1)$$

- (2) シフトが実数 ρ である単一のレゾルベントの実多項式 P であるもの (2) (文献 [21], [22], [23], [25], [33]).

$$\mathcal{F} \equiv P(\mathcal{R}(\rho)). \quad (2)$$

- (3) シフトが虚数 ρ' である単一のレゾルベントの虚部の実多項式 P であるもの (3) (文献 [18], [20], [21], [22], [23], [25], [29], [33]).

$$\mathcal{F} \equiv P(\operatorname{Im}\mathcal{R}(\rho')). \quad (3)$$

- (4) 混合型で「レゾルベントの線形結合の実部」の実多項式 P であるもの (4) (文献 [24], [26], [27], [28], [30], [31], [32]).

$$\mathcal{F} \equiv P\left(\operatorname{Re} \sum_{j=1}^k c_j \mathcal{R}(\rho_j)\right). \quad (4)$$

レゾルベント $\mathcal{R}(\rho)$ のベクトル \mathbf{x} に対する作用 $\mathbf{y} \leftarrow \mathcal{R}(\rho)\mathbf{x}$ は, シフト行列 $C(\rho) \equiv A - \rho B$ を係数とする連立 1 次方程式 $C(\rho)\mathbf{y} = B\mathbf{x}$ を \mathbf{y} について解いて実現する. するとフィルタ対角化法の計算は連立 1 次方程式を解く作業が主要部になる. いまその連立 1 次方程式を直接法で解くことにすると, 固有値問題の行列が大規模であればシフト行列 $C(\rho)$ の行列分解を行う演算量と分解結果を保持する記憶量が計算を実施するうえでの制約要因になりがちなので, 必要なレゾルベントの数をできるだけ少なくしたい. レゾルベントの数が最も少ない 1 つになるフィルタの構成としては, 「シフトが実数の単一のレゾルベントの実多項式」あるいは「シフトが虚数の単一のレゾルベントの虚部の実多項式」があるが, それらはレゾルベントのシフトや多項式のとり方を調整しても特性を理想のものに近づけることが困難である. たとえば遮断特性をきわめて急峻にすることは (レゾルベントを多数用いるのでなければ) 困難である. そのため, 単一のレゾルベントで構成されたフィルタを利用して得られる近似対の残差は十分小さくできず, 近似対を高い精度で得る目的には適さないことがあった.

我々はレゾルベントの数を 1 つではなくて少数 (2 つから 4 つ程度) 用いて構成した混合型のフィルタでは特性をかなり改善できることをすでに見出した (文献 [26], [27], [28], [30], [31], [32]). しかし連立 1 次方程式を直接法で解くのであるかぎりは, 行列分解の演算量とそ

の結果を保持するための記憶量はレゾルベントの数に比例する. そのため可能であれば用いるレゾルベントの数は 1 つのまま, 近似対の精度を改良したいと考える.

そこで本論文では, 特性があまり良くないフィルタを 1 回ではなくて合計 2 回から 4 回程度適用して近似対を改良することを試みる. また本論文では簡単のために毎回同じフィルタを反復して用いる. するとレゾルベントを実現するために解く連立 1 次方程式の係数行列は毎回同じになるので, 連立 1 次方程式を直接法で解くのに必要な行列分解の計算は全体で 1 回だけ行えばよい.

しかしフィルタの適用を単純に繰り返すと, 各「固有ベクトル」に対する伝達率はフィルタの適用回数だけ掛け合わされて違いが拡大する. すると各ベクトルに含まれている「小さい伝達率の固有ベクトル」は, 同じベクトルに含まれている「大きい伝達率の固有ベクトル」の大きい数値に覆われて下位の桁の精度を失ったり消えてしまうことになる. そこで, フィルタを適用する前のベクトルの組に対して (正規) 直交化の処理を毎回施すことにする. それによりベクトルの組の線形独立性を維持することで, 相対的に拡大した「大きい伝達率の固有ベクトル」により「小さい伝達率の固有ベクトル」が覆われて, 丸めの影響により情報の精度を失う傾向を抑制することを狙う. これは古くから構造解析の分野で用いられてきた下端側の固有値の固有対を同時に複数解くための, 逆反復法と直交化を組み合わせた反復手法である「同時逆反復法」(Simultaneous Inverse Iteration Method) (文献 [2], [3], [10], [11], [16]), あるいはその原理である「直交反復法」(Orthogonal Iteration) [7] の類似であるといえる.

フィルタを 1 回だけ適用するこれまでの方法では, 求めたい固有ベクトルについて事前の情報を何も使わずに, 初期ベクトルの組を乱数から作り, それを B -正規直交化してベクトルの組 X を作り, その X に対してフィルタを適用して不要な固有ベクトルを含む割合を減らしたベクトルの組 Y を作る. そうしてその Y を分析して近似対を得ていたわけである. それならば, 初期ベクトルの組を乱数から作るものではなくて Y にして, それを B -正規直交化したベクトルの組 X' を作り, その X' にフィルタを適用して不要な固有ベクトルを含む割合をさらに減らしたベクトルの組 Y' を作れば, その Y' から分析により得られる近似対はより改良されたものになるであろうと考えられる.

本論文では以上の考えに基づいて, 乱数を用いて生成したベクトルの組から始めて, 「 B -正規直交化した後にフィルタを適用する」という処理を 2 回~4 回反復することで近似対を改良する方法を実験してみた例について報告する.

1.1 B -正規直交化について

m 個のベクトルの組 Y に対する B -正規直交化には, 計量 B の特異値分解 (B -SVD) を用いた. 今回の実験で用

いたその具体的な計算法は以下のものである。

まずベクトルの組 Y の計量 B による内積 (Gram 行列) である m 次の正定値対称行列 $G := Y^T B Y$ を作り, 次にその固有値分解 $G =: U D U^T$ を Rutishauser の Jacobi 法 [1] を用いて行い, m 次の対角行列 D と直交行列 U を求める (D の対角要素は減少順にするのが便利である). その U を用いて $Y := Y U$ とすることで計量 B に関して直交化された Y を作る.

しかし元の Y が悪条件であったり, あるいは大規模な Y から G を作る時の丸め誤差の蓄積などにより数値計算で得られた Y の B -直交性が十分ではないことがある. そのような場合には上記の過程をたかだか 2, 3 回ほど繰り返すことにより Y の B -直交性を向上させることを狙う. 繰り返しの途中で G が十分対角に近いと判定されたら, Y の B -直交性は十分であるとす.

Y の B -直交性が十分良好になったとき, 最後に求めた G は対角行列であるが, その各対角要素の平方根 (あるいは Y の各列の計量 B によるノルム) の値が閾値未満である列は除去し, 除去されずに残った各列には対応するノルムの逆数を乗じることで B -正規直交ベクトルの組 Y とす. B -正規直交化後の Y の列の数は, 元の m から閾値による判定で除去された列の数を減じたものになる.

注意点

複数のベクトルの組に対してフィルタを適用する処理では, その組を任意に分割してそれぞれに対してまったく独立にフィルタを適用することにより容易に処理の分散並列化ができるが, 正規直交化の処理はベクトルの組が揃ってからでないと始められないので, データの同期が必要となり分散並列化に対しては不利になる. また正規直交化はベクトルに対する線形的作用素ではないので, 処理の途中に正規直交化が入ると, フィルタは線形でも全体として線形作用素ではなくなることも注意しておく.

2. フィルタ対角化法の概略

いま行列 A と B が実対称で, B は正定値である一般固有値問題 (5) の固有対 (λ, \mathbf{v}) でその固有値 λ が区間 $[a, b]$ にあるものを求めるとする.

$$A \mathbf{v} = \lambda B \mathbf{v} \tag{5}$$

そのために, 固有値が $[a, b]$ にある固有ベクトルはよく通過させるが, 固有値が $[a, b]$ から離れた固有ベクトルはなるべく阻止する線形作用素をうまく構成してフィルタとして用いる. そのようなフィルタは固有値が $[a, b]$ 近傍にある固有ベクトルが張る不変部分空間への射影作用素の近似になる. そうしてフィルタを十分多くの線形独立なベクトルに作用させて得られたベクトルの線形結合をうまく作ることで不変部分空間の近似の基底を構成する (文献 [13] および文献 [17] の 5 章). 得られた基底に対して Rayleigh-Ritz

法を適用することで必要な固有対の近似が得られる.

3. 今回の実験に用いたフィルタ

今回用いたフィルタは単一のレゾルベントの多項式型である. 簡単化のために多項式には Chebyshev 多項式を用いている. するとレゾルベントの数は 1 つになるが, フィルタの伝達特性をあまり良くすることができない (通過域における伝達率の最大と最小の比 $1/g_p$ をあまり小さくすることができない).

固有値の区間 $[a, b]$ が固有値分布の下端である場合には, シフト ρ を実数として, 式 (6) の形のフィルタを採用することができる.

$$\mathcal{F} \equiv g_s T_n(2\gamma \mathcal{R}(\rho) - I). \tag{6}$$

固有値の区間が固有値分布に対して一般的な位置にある場合には, 虚数のシフト ρ' を用いて, 式 (7) の形のフィルタを採用する.

$$\mathcal{F} \equiv g_s T_n(2\gamma' \text{Im} \mathcal{R}(\rho') - I). \tag{7}$$

ここで $T_n(x)$ は x の n 次の第 1 種 Chebyshev 多項式を表し, I は恒等作用素, Im は複素ベクトルに対してその虚部をとり出す作用素である. γ や γ' は実数の定数であり, g_s はフィルタの阻止域における伝達関数の大きさの上限値である.

3.1 今回のフィルタの設計

本論文ではフィルタの指定をパラメタの 3 つ組 (n, μ, g_s) を用いて行う (別の指定の方法も可能である). ここで n は Chebyshev 多項式の次数であり, μ は規格化座標 t による遷移域と阻止域の境界を表す座標, g_s は阻止域におけるフィルタの伝達関数の大きさの上限値である. なお, フィルタの伝達関数の通過域における最大値は 1 であり, 最小値は g_p とする. 以下にフィルタの構成法を示す.

3.2 レゾルベントのシフトに実数を用いる場合

区間 $[a, b]$ の位置は固有値分布の下端にあつて, a は最小固有値 λ_{\min} よりも小さいとする.

与えられたパラメタの 3 つ組 (n, μ, g_s) に対して, 式 (8) により, 実数のシフト ρ とレゾルベントの係数である実数の γ (および g_p) を計算する.

$$\left\{ \begin{array}{l} \sigma \leftarrow \mu / \sinh \left(\frac{1}{2n} \cosh^{-1} \frac{1}{g_s} \right), \\ \rho \leftarrow a - (b - a)\sigma, \\ \gamma \leftarrow (b - a)(\sigma + \mu), \\ g_p \leftarrow g_s \cosh \left\{ 2n \sinh^{-1} \sqrt{(\mu - 1)/(1 + \sigma)} \right\}. \end{array} \right. \tag{8}$$

するとフィルタ \mathcal{F} は式 (9) で与えられる.

$$\mathcal{F} \equiv g_s T_n(2\gamma \mathcal{R}(\rho) - I). \quad (9)$$

いまの場合、 λ に対する正規化座標 t は $\lambda \in [a, b]$ から $t \in [0, 1]$ への線形変換 $t \equiv (\lambda - a)/(b - a)$ により定義される。その t を引数とするフィルタの伝達関数 $g(t)$ は式 (10) で表される。あるいは λ を引数とする伝達関数 $f(\lambda)$ は式 (11) になる。

$$g(t) = g_s T_n\left(2 \frac{\mu + \sigma}{t + \sigma} - 1\right). \quad (10)$$

$$f(\lambda) = g_s T_n\left(2\gamma \frac{1}{\lambda - \rho} - 1\right). \quad (11)$$

3.3 レゾルベントのシフトに虚数を用いる場合

シフトに虚数を用いる場合は、区間 $[a, b]$ の位置を自由に設定できる。

フィルタのパラメタの3つ組 (n, μ, g_s) が与えられたら、式 (12) により、虚数のシフト ρ' とレゾルベントの係数である実数 γ' (および g_p) を計算する。

$$\left\{ \begin{array}{l} \sigma \leftarrow \mu / \sinh\left(\frac{1}{2n} \cosh^{-1} \frac{1}{g_s}\right), \\ \rho' \leftarrow \frac{a+b}{2} + \left(\frac{b-a}{2}\right) \sigma \sqrt{-1}, \\ \gamma' \leftarrow \left(\frac{b-a}{2}\right) \frac{\mu^2 + \sigma^2}{\sigma}, \\ g_p \leftarrow g_s \cosh\left\{2n \sinh^{-1} \sqrt{(\mu^2 - 1)/(1 + \sigma^2)}\right\}. \end{array} \right. \quad (12)$$

するとフィルタ \mathcal{F} は式 (13) で与えられる。

$$\mathcal{F} \equiv g_s T_n(2\gamma' \text{Im } \mathcal{R}(\rho') - I). \quad (13)$$

この場合の λ の正規化座標 t は $\lambda \in [a, b]$ から $t \in [-1, 1]$ への線形変換 $t \equiv (2\lambda - a - b)/(b - a)$ により定義される。その t を引数とする伝達関数 $g(t)$ は式 (14) で表される。あるいは引数 λ の伝達関数 $f(\lambda)$ は以下の式 (15) で表される。

$$g(t) = g_s T_n\left(2 \frac{\mu^2 + \sigma^2}{t^2 + \sigma^2} - 1\right). \quad (14)$$

$$f(\lambda) = g_s T_n\left(2\gamma' \text{Im} \frac{1}{\lambda - \rho'} - 1\right). \quad (15)$$

4. フィルタを正規直交化付きで数回適用する方法の概略

簡単のため、フィルタは同じものを IT 回 (1~4 回) 反復して用いることにする。なおベクトルの組に対して毎回のフィルタ適用の直前に必ず B -正規直交化を行うことにする。

B -正規直交化法には閾値による切断付きの B -特異値分解 (B -SVD) を用いた。その閾値はマシンイプシロンの 100 倍に設定したが、これは特異値の絶対的な値であって相対比ではない。そうして特異値が閾値以下の特異ベクトルは切断 (除去) した。

フィルタを B -正規直交化付きで反復適用する対角化法の概要を (図 1) に示す

(1) フィルタ \mathcal{F} を設定し、用いるレゾルベントの作用を与える連立 1 次方程式の係数行列をここであらかじめ分解してその結果を保持しておく。

(2) m 個のランダムなベクトルの組を生成して Y とする。

(3) 次の処理を $i = 1, 2, \dots, \text{IT}$ と繰り返す。

「 Y を B -正規直交化したものを X に設定し、

その X にフィルタを適用して $Y \leftarrow \mathcal{F}X$ を作る。」

繰り返しの途中で B -正規直交化の際に閾値に基づく切断で有効階数が低下したら、その低下した階数を新たな m (組 X や Y の列の数=ベクトルの数) に設定しなおして処理を続ける。

後は、(X と Y を最後のフィルタ適用時のものとして)、

(4) X と Y および伝達関数の特性を考慮して、 Y の列の線形結合で不変部分空間 ($[a, b]$ 近傍の固有値全てに対する) の近似空間の基底をうまく構成する (方法は文献 [13] および [17] の 5 章)。その基底に Rayleigh-Ritz 法を適用して得られた Ritz 対を元の一般固有値問題の近似対とする。

図 1 フィルタを B -正規直交化付きで反復適用する方法の概略

Fig. 1 Outline of present method which repeats the application of the filter with B -orthonormalization.

5. フィルタの適用による固有ベクトルの含有率の改良

いま対象にしている一般固有値問題のすべての固有対を $(\lambda_i, \mathbf{v}_i)$, $i=1, 2, \dots, N$ とする。任意のベクトル \mathbf{x} にフィルタ \mathcal{F} を適用するとき、その固有ベクトル展開の第 i 番目の固有ベクトル \mathbf{v}_i の係数にはフィルタの伝達率 $f(\lambda_i)$ が乗じられる。今回のフィルタの伝達関数の設定では、固有値 λ が通過域にあれば $g_p \leq f(\lambda) \leq 1$ であり、固有値 λ が阻止域にあれば $|f(\lambda)| \leq g_s (\ll g_p)$ である。よってベクトル \mathbf{x} の固有ベクトル展開はフィルタを適用することで、通過域にある固有値 λ_p に対応する固有ベクトル \mathbf{v}_p に対する展開係数はその $f(\lambda_p)$ 倍に変わり、固有値が阻止域にある固有ベクトルに対する展開係数の大きさはそれぞれ g_s 以下の数が乗じられたものになる。

すると「(多くの) 除去したい固有ベクトルの強度」の「通過させたい p 番目の固有ベクトルの強度」に対する割合は、フィルタの適用により (どれも) 適用前の $g_s/f(\lambda_p)$ 倍以下になる。この比の値は 1 よりも小さく、それは (ある意味で) 「特定の信号 p の強度に対するノイズの強度の比率」がフィルタの適用により低減を受ける倍率である。この倍率が小さいほど、フィルタの適用により固有値が阻止域にある除去したい固有ベクトルの「相対的」な含有率が減る。通過させたい固有ベクトルの伝達率が大きいほどこの倍率は小さくなり、フィルタの適用による品質改良の相対比で考えたときの程度が高い。

通過域にある伝達率が g_p 以上の固有ベクトルのすべてを求める場合には、上記の比の値の上限は g_s/g_p になる。

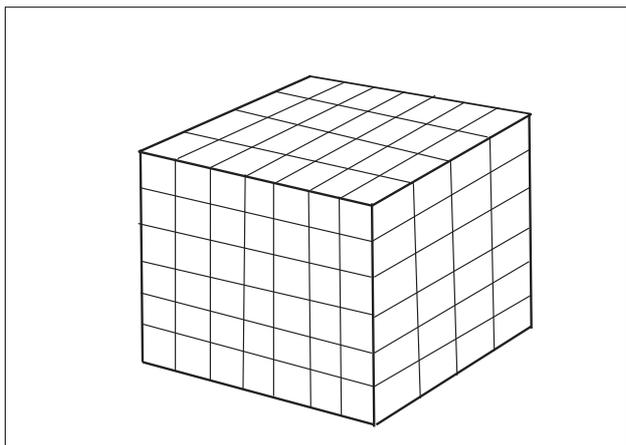


図 2 立方体の FEM 要素分割の概念. 分割例 $(N_1, N_2, N_3) = (3, 5, 6)$

Fig. 2 Concept of finite element subdivision of a cube. For the case of $(N_1, N_2, N_3) = (3, 5, 6)$.

この上限は通過域にある固有値に対する伝達率の「最小値」が閾値 g_p に等しい場合にだけ達成されるものであるから、「最小値」がそれよりも大きければそれだけ、実際に得られる改良は上限から予想されるものよりも良くなる。

6. 実対称定値一般固有値問題の例題

例題として用いた実対称定値一般固有値問題 (16) は、1 辺の長さ π の立方体領域において零 Dirichlet 境界条件を課した 3 次元のラプラシアン固有値問題 (17), その有限要素法 (FEM) による離散化近似により得られるものである。

$$A\mathbf{v} = \lambda B\mathbf{v} \tag{16}$$

$$-\Delta \Psi(x, y, z) = \lambda \Psi(x, y, z) \tag{17}$$

FEM の要素分割は、立方体領域の辺に沿う 3 つの方向をそれぞれ $N_1 + 1, N_2 + 1, N_3 + 1$ の小区間に等分割して得られた各方向の小区間の直積の直方体である (図 2)。FEM 要素内の関数の展開基底には各辺方向の 3 重線形関数を用いた。

行列 A と B の次数は $N = N_1 N_2 N_3$ である。そうして ($N_1 \leq N_2 \leq N_3$ であるとして) 帯幅がなるべく小さくなるように基底関数に番号を適切に付けると、(対角線を含まない) 半帯幅 (下帯幅) は $w_L = 1 + N_1 + N_1 N_2$ になる。

このような FEM 離散化で得られた一般固有値問題に対して、固有値が指定された区間 $[a, b]$ に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。

この例題の固有値の厳密な値は 6.1 節の方法により、簡単な数式から計算で求めることができる。任意の区間内の固有値の数も、厳密値を作ってその区間内のものを数えることで求められる。

6.1 例題の固有値の厳密値を与える式

立方体の各辺の 3 方向を $N_1 + 1, N_2 + 1, N_3 + 1$ に等

分割したとする。そのとき例題の 3 次元問題の固有値は添字の組 (k_1, k_2, k_3) で識別されて、各方向ごとの 1 次元問題の固有値の和の形として式 (18) により表される。

$$E_{(k_1, k_2, k_3)}^{[N_1, N_2, N_3]} = \mathcal{E}_{k_1}^{[N_1]} + \mathcal{E}_{k_2}^{[N_2]} + \mathcal{E}_{k_3}^{[N_3]}, \quad k_i = 1, 2, \dots, N_i. \tag{18}$$

そうして各方向の 1 次元問題の固有値は式 (19) で与えられる。ただし $\theta_k \equiv k\pi/(N + 1)$ である。

$$\mathcal{E}_k^{[N]} = \frac{6k^2 (\sin \theta_k / \theta_k)^2}{(1 + \cos \theta_k)(2 + \cos \theta_k)}, \quad k = 1, 2, \dots, N. \tag{19}$$

6.2 近似対の評価に用いる相対残差

近似対 (λ, \mathbf{v}) の品質を評価する基準として、式 (20) で定義される相対残差 Θ を用いる。この値が小さいほど近似対の品質は良いと評価することにする。

$$\Theta \equiv \frac{\|A\mathbf{v} - \lambda B\mathbf{v}\|}{\|\lambda B\mathbf{v}\|}. \tag{20}$$

この式の値はベクトル \mathbf{v} の規格化にはよらないし、行列 A と B に共通の定数を乗じて変わらない性質を持つ。なお今回はベクトルのノルム $\|\cdot\|$ として 2-ノルムを用いた。

Θ の計算の主要部は、ベクトル \mathbf{v} に対して行列 A と B を乗じる行列ベクトル積であるが、複数のベクトルをまとめて行列 V として、行列積 AV と BV をそれぞれ計算をすれば、行列 A と B の記憶に対する参照を各 1 回にできる。また行列 A と B が疎であればそれだけ計算が容易に行える。

ベクトルのノルムに 2-ノルムを用いた場合の、相対残差 Θ の幾何学的な意味を以下に示す。いま N 次元ユークリッド空間において、図 3 で示される原点 O からの 2 つの位置ベクトル $A\mathbf{v}$ と $\lambda B\mathbf{v}$ の挟む角を θ とする。それら 2 つのベクトルの差が残差ベクトル \mathbf{r} であり、そのノルムが図中の (N 次元) 球の半径である。そうして中心 $\lambda B\mathbf{v}$ で半径 $\|\mathbf{r}\|$ の球面上にあるベクトル $A\mathbf{v}$ が球に接しているときに角 θ は最大になるので、以下の不等式 (21) が成り立つ。

$$\sin \theta \leq \frac{\|\mathbf{r}\|}{\|\lambda B\mathbf{v}\|} = \Theta. \tag{21}$$

注：我々は以前の文献 [33] の中では $2 \sin(\theta/2) = \Theta$ と述べていたが、それは 2 つのベクトル $A\mathbf{v}$ と $\lambda B\mathbf{v}$ のそれぞれの 2-ノルムを等しくするように λ を決めた場合であって、 λ が Rayleigh 商 $(\mathbf{v}^T A\mathbf{v})/(\mathbf{v}^T B\mathbf{v})$ に等しいものとなる我々の場合には正しくなかった。ここで訂正するとともにお詫び申し上げる。

7. フィルタを直交化付きで反復する実験の例

7.1 今回の実験に用いたフィルタについて

フィルタはその伝達特性をパラメタの 3 つ組 (n, μ, g_s)

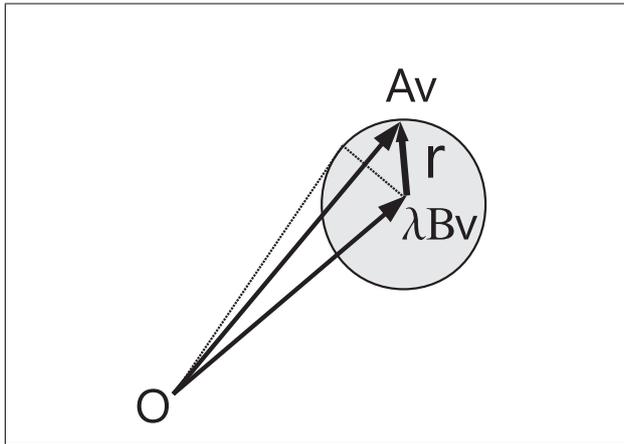


図 3 相対残差 Θ の幾何学的な意味

Fig. 3 Geometrical meaning of the relative residual Θ .

表 1 (下端固有対用フィルタ) 次数 n に対する g_p と g_s/g_p の値
Table 1 Values of g_p and g_s/g_p for degree n (filter for lower exterior eigenpairs).

n	g_p	g_s/g_p
8	8.79884E-9	1.13651E-4
10	4.20592E-8	2.37760E-5
15	4.17183E-7	2.39703E-6
20	1.21554E-6	8.22680E-7

表 2 (中間固有対用フィルタ) 次数 n に対する g_p と g_s/g_p の値
Table 2 Values of g_p and g_s/g_p for degree n (filter for interior eigenvalues).

n	g_p	g_s/g_p
8	5.90737E-7	1.69280E-6
10	4.20226E-6	2.37967E-7
15	5.55703E-5	1.79952E-8
20	1.63167E-4	6.12869E-9

で指定した. 今回は実験の記述の簡単化のために次数 n だけを変えており, 残りの 2 つのパラメタについては $\mu = 1.5$ と $g_s = 10^{-12}$ に固定している. そうして 1 つの反復の中ではつねに同じフィルタを繰り返し用いている.

7.1.1 下端固有対用フィルタ

下端固有対用のフィルタは, 式 (9) で表される簡単な構成のものであり, シフトが実数である単一のレゾルベントの n 次 Chebyshev 多項式である. 伝達関数の形状を代表する 3 つのパラメタである μ , g_s , g_p のうち, 最初の 2 つについては $\mu = 1.5$ と $g_s = 10^{-12}$ に固定している. 残りの 1 つであるフィルタの通過域での伝達率の最小値 g_p はフィルタの次数 n が決まれば式 (8) から計算できる. いま次数 n を 8, 10, 15, 20 としたそれぞれの場合に対する g_p と g_s/g_p の数値を表 1 に掲げる. 通過域内での伝達関数の最大値は 1 であり最小値は g_p であるから, この種類のフィルタは構成は簡単であるが, 伝達率の通過域における一様性が良くないことが分かる. そうして g_s/g_p は「フィルタ

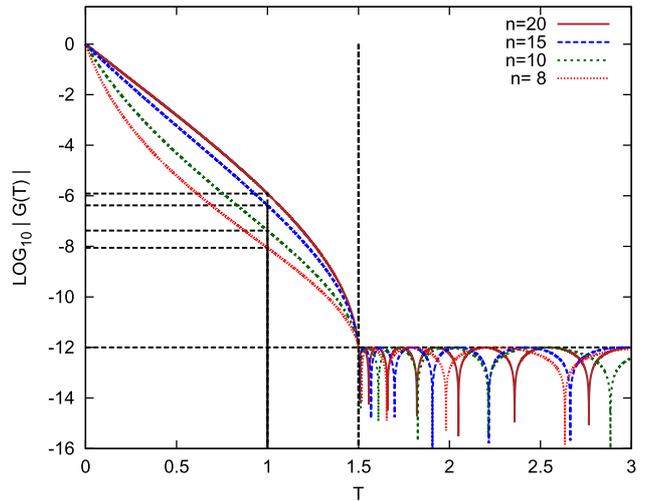


図 4 下端固有対用フィルタの伝達関数の大きさ $|g(t)|$

Fig. 4 Transfer function magnitude $|g(t)|$ of the filter to solve lower exterior eigenpairs.

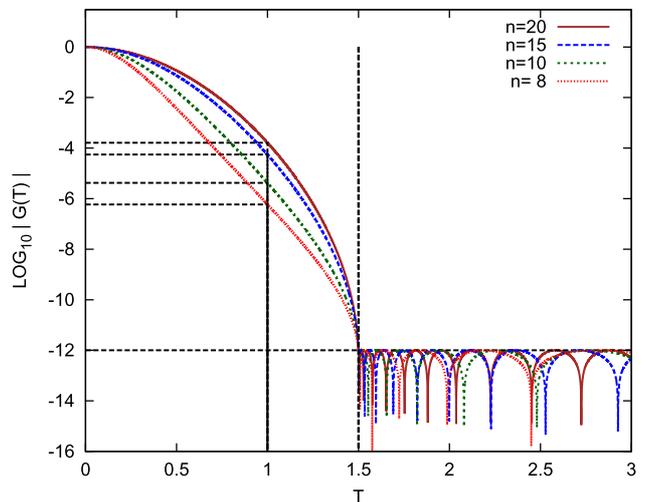


図 5 中間固有対用フィルタの伝達関数の大きさ $|g(t)|$ (右半分)

Fig. 5 Transfer function magnitude $|g(t)|$ of the filter to solve interior eigenpairs.

の阻止域における伝達率の上限值 g_s 」と「通過域伝達率の最小値 g_p 」の比の値であるが, 反復ごとに近似対に含まれる不要な固有ベクトルの割合が減少する比率の上限をこの値により見積もることができよう. つまりこの値が小さいほど, 反復による相対残差の最大値の減少が速いと期待できる.

フィルタの次数 n を 8, 10, 15, 20 とした各場合について, 式 (10) で表された正規化座標 t を引数とする下端固有対用フィルタの伝達関数の大きさ $|g(t)|$ を片対数でグラフにプロットしたものを図 4 に示す.

7.1.2 中間固有対用フィルタ

中間固有対用のフィルタは, 式 (13) で表される簡単な構成のものであり, シフトが虚数である単一のレゾルベントの虚部の n 次 Chebyshev 多項式である. 伝達関数の形状を代表する 3 つのパラメタ μ , g_s , g_p のうち, 最初の 2 つは

$\mu = 1.5$ と $g_s = 10^{-12}$ に固定した. 残りの1つであるフィルタの通過域における伝達率の最小値 g_p はフィルタの次数 n が決まれば式 (12) から計算できる. 次数 n を 8, 10, 15, 20 とした各場合に対する g_p と g_s/g_p の数値を表 2 に掲げる. 通過域において伝達関数の最大値は 1 で最小値は g_p であり, 同じ次数 n では下端固有対用のフィルタに比べると g_p の値が 2 桁程度大きいので, 中間固有対用のフィルタの性質の方が下端固有対用のものよりは良いが, それでもやはり通過域における伝達率の一様性があまり良くないことが分かる. そうして g_s/g_p の値が小さいほど, 反復による相対残差の最大値の減少が速いと期待できる.

フィルタの次数 n を 8, 10, 15, 20 とする各場合に対して, 式 (14) で表される正規化座標 t を引数とする中間固有対用フィルタの伝達関数, その大きさ $|g(t)|$ の右半分 ($t \geq 0$) を片対数でグラフにプロットしたものを図 5 に示す.

単一のレゾルベントを用いた簡易型のフィルタは, 下端固有対用のものよりも中間固有対用の方が伝達特性は良いが, そのかわりにレゾルベントの作用を実現する連立 1 次方程式の係数行列が正定値実対称ではなくて複素対称になること, また方程式の前進後退代入も複素数の演算になること, 複素数のデータを保持するための記憶量は実数の場合の 2 倍になるなどの不利な点もある.

7.2 例題 R1 : 下端固有対を解いた例その 1

この例題 R1 は FEM の要素分割 $(N_1, N_2, N_3) = (20, 30, 40)$ から導かれる一般固有値問題である. その係数行列 A と B の次数 N は 24,000 で下帯幅 w_L は 621 であり, これは小規模な問題である. そうして下端固有対を求める例題として, 固有値が区間 $[a, b] = [0, 30]$ にある固有対を求めてみる. そのような固有対は全部で 54 個ある.

フィルタには下端固有対用のものを用いる. そうして反復では同じフィルタを繰り返し用いる. 固有値が通過域と $\mu = 1.5$ に対応する遷移域の両方をあわせた区間 $[a, b] = [0, 45]$ にある固有対は全部で 106 個ある. そのため, フィルタで濾過するベクトルの数 m は 106 よりも多くすることが望ましい.

例題 R1 : 近似対の相対残差の最大値

フィルタの次数 n を 8, 10, 15, 20 とした各場合に対してベクトルの数 m を 80 から 140 まで 20 刻みでとり, 反復回数 IT を 1 から 4 まで変えて得られた, 固有値が指定区間にある近似対の相対残差の最大値をそれぞれ表に示す (表 3, 表 4, 表 5, 表 6).

フィルタの次数 n を 8, 10, 15, 20 とした各場合について, 横軸に反復回数 IT を 1 から 4 までとり, 縦軸に指定した区間に固有値を持つ近似対の相対残差の最大値をとって, ベクトルの数 m を 80, 100, 120, 140 と 4 通りに変

表 3 例題 R1 : 反復回数に対する相対残差の最大値 ($n = 8$)

Table 3 Example R1: The maximum of relative residuals for iterations (case $n = 8$).

IT	$m = 80$	$m = 100$	$m = 120$	$m = 140$
1	3.5E-02	2.9E-02	2.1E-02	1.7E-02
2	2.8E-04	5.3E-06	2.3E-06	2.2E-06
3	2.6E-06	9.0E-10	2.1E-10	1.6E-10
4	5.8E-08	3.0E-13	1.2E-13	1.2E-13

表 4 例題 R1 : 反復回数に対する相対残差の最大値 ($n = 10$)

Table 4 Example R1: The maximum of relative residuals for iterations (case $n = 10$).

IT	$m = 80$	$m = 100$	$m = 120$	$m = 140$
1	1.1E-02	7.3E-03	4.8E-03	3.3E-03
2	1.0E-04	2.1E-07	9.6E-08	7.6E-08
3	4.5E-07	1.6E-11	2.3E-12	1.3E-12
4	1.3E-08	1.3E-13	1.3E-13	1.3E-13

表 5 例題 R1 : 反復回数に対する相対残差の最大値 ($n = 15$)

Table 5 Example R1: The maximum of relative residuals for iterations (case $n = 15$).

IT	$m = 80$	$m = 100$	$m = 120$	$m = 140$
1	4.8E-03	1.1E-03	8.5E-04	4.3E-04
2	3.4E-05	3.3E-09	8.8E-10	6.8E-10
3	6.2E-08	1.6E-13	1.6E-13	1.6E-13
4	3.5E-10	1.6E-13	1.6E-13	1.6E-13

表 6 例題 R1 : 反復回数に対する相対残差の最大値 ($n = 20$)

Table 6 Example R1: The maximum of relative residuals for iterations (case $n = 20$).

IT	$m = 80$	$m = 100$	$m = 120$	$m = 140$
1	2.6E-03	5.5E-04	2.4E-04	1.4E-04
2	9.3E-06	5.0E-10	1.3E-10	1.1E-10
3	5.9E-08	2.0E-13	1.9E-13	2.0E-13
4	8.0E-11	1.9E-13	2.0E-13	1.9E-13

えて片対数でプロットしたグラフをそれぞれ示す (図 6, 図 7, 図 8, 図 9). 各グラフ中の “expected slope” とラベルされた黒線は, 相対残差の最大値が反復ごとに g_s/g_p 倍に減少すると仮定した場合のグラフの傾きを示したものである. これらの 4 つのグラフから, ベクトルの数 m が 100 未満の場合と 100 以上の場合とでは反復回数 IT の増加にともなう相対残差の最大値の減少の様子がかなり異なっていること, そうしてベクトルの数 m が 100 と 120 の場合の違いは小さく, m が 120 と 140 の場合はほぼ一致していることが分かる.

またこれらのグラフから, フィルタの次数 n が 15 のときは m が 100 以上であれば, 近似対の改良は反復 3 回目で完了していて, 相対残差の最大値は約 10^{-13} であることが分かる.

この例題 R1 で行ったすべての実験の組合せ (フィルタ

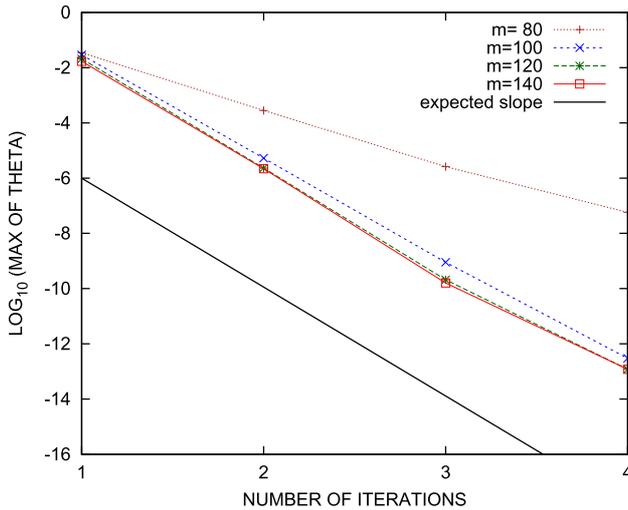


図 6 例題 R1：反復回数に対する相対残差の最大値 ($n = 8$) ($m > 106$ が望ましい)

Fig. 6 Example R1: The maximum of relative residuals for iterations (case $n = 8$) ($m > 106$ is desirable).

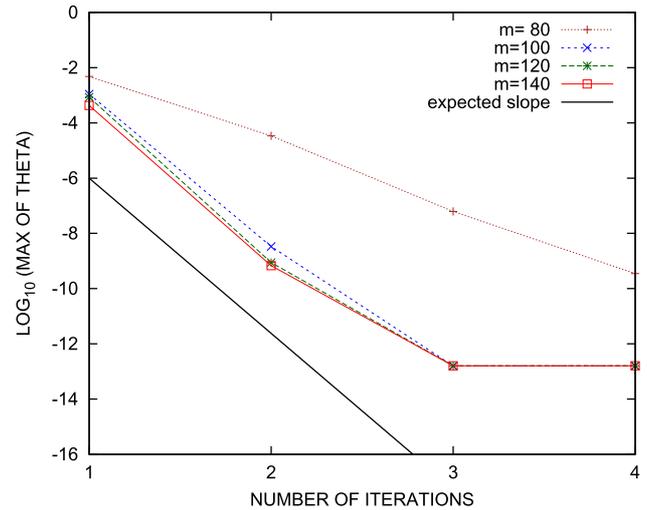


図 8 例題 R1：反復回数に対する相対残差の最大値 ($n = 15$) ($m > 106$ が望ましい)

Fig. 8 Example R1: The maximum of relative residuals for iterations (case $n = 15$) ($m > 106$ is desirable).

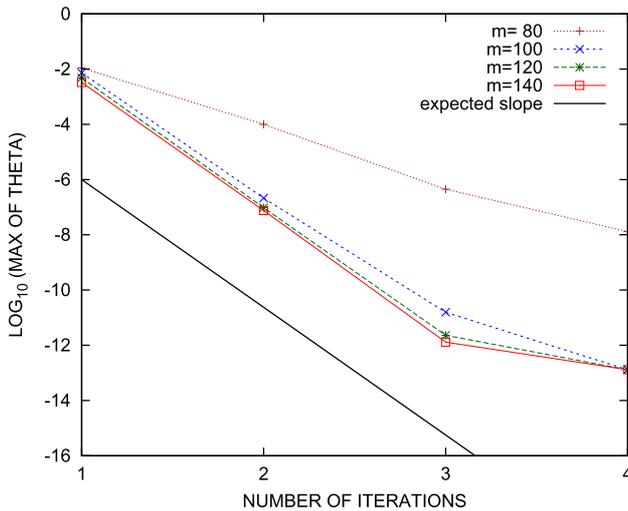


図 7 例題 R1：反復回数に対する相対残差の最大値 ($n = 10$) ($m > 106$ が望ましい)

Fig. 7 Example R1: The maximum of relative residuals for iterations (case $n = 10$) ($m > 106$ is desirable).

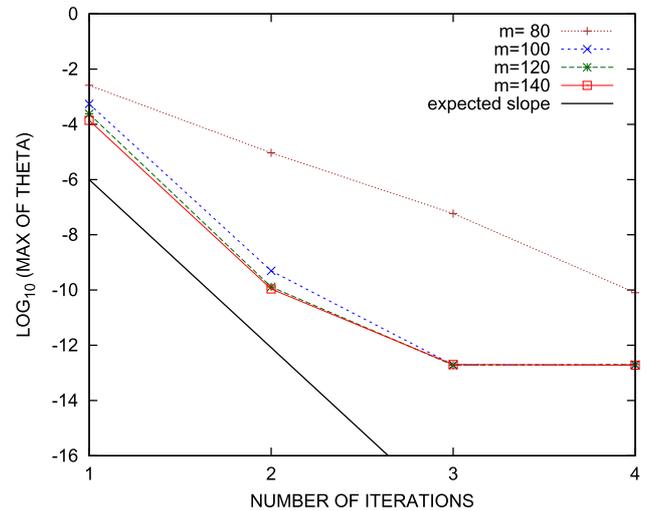


図 9 例題 R1：反復回数に対する相対残差の最大値 ($n = 20$) ($m > 106$ が望ましい)

Fig. 9 Example R1: The maximum of relative residuals for iterations (case $n = 20$) ($m > 106$ is desirable).

の反復回数 IT, フィルタの次数 n , ベクトルの数 m) について, 得られた近似対のうちで固有値が指定された区間にあるものの数は真の固有対の数 54 と一致していた.

例題 R1-Q：近似対の相対残差の最大値 (四倍精度計算)

ここでは特別に計算に用いる浮動小数点の数値と演算をすべて通常精度の IEEE-754 規格の 2 進 64 bit 倍精度から同規格の 2 進 128 bit 四倍精度に単純に置き換えて, 上記の下端固有対を解く例題 R1 と同じ問題を解いてみた例を示す (これを例題 R1-Q と呼ぶことにする).

フィルタの次数 n が 8, 10, 15, 20 の各場合について, 横軸に反復回数 IT を 1 から 8 までとり, 縦軸に指定した区間に固有値を持つ近似対の相対残差の最大値をとって, ベクトルの数 m を 80 から 140 まで 20 刻みでとり, 片対

数でプロットしたグラフをそれぞれ示す (図 10, 図 11, 図 12, 図 13). 各グラフ中の “expected slope” とラベルされた黒線は, 相対残差の最大値が反復ごとに g_s/g_p 倍に減少すると仮定した場合のグラフの傾きを示したものである. いずれの場合にもベクトルの数 m が 120 以上の場合には, 相対残差はほとんど同じ値をとり, 相対残差の最大値は減少が停滞する手前までは反復ごとにほぼ g_s/g_p 倍になって減少していることが分かる.

7.3 例題 R2：下端固有対を解いた例その 2

この例題 R2 で扱う一般固有値問題は, FEM の要素分割 (N_1, N_2, N_3) = (50, 60, 70) から導かれたもので, その係数行列 A と B の次数 N は 210,000 で下帯幅 w_L は 3,051 で

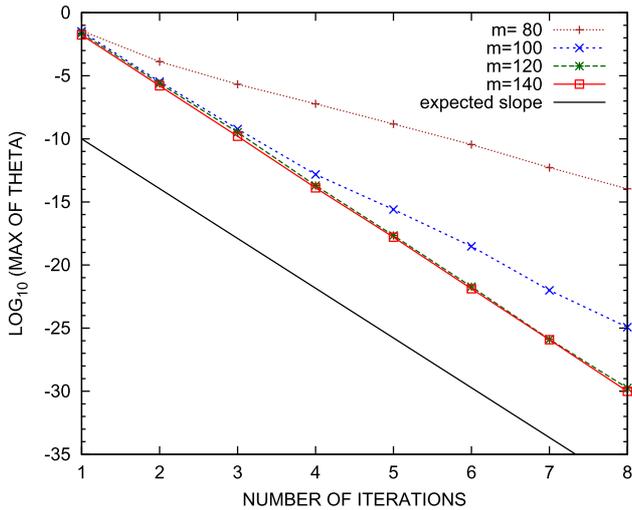


図 10 例題 R1-Q: 反復回数に対する相対残差の最大値 ($n = 8$) ($m > 106$ が望ましい)

Fig. 10 Example R1-Q: The maximum of relative residuals for iterations (case $n = 8$) ($m > 106$ is desirable).

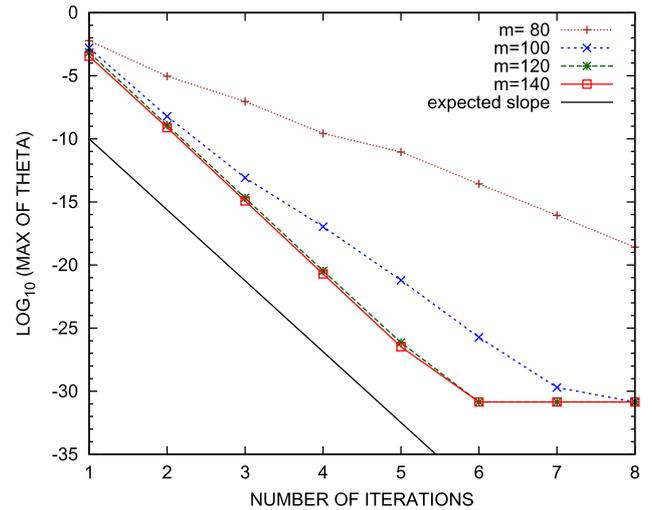


図 12 例題 R1-Q: 反復回数に対する相対残差の最大値 ($n = 15$) ($m > 106$ が望ましい)

Fig. 12 Example R1-Q: The maximum of relative residuals for iterations (case $n = 15$) ($m > 106$ is desirable).

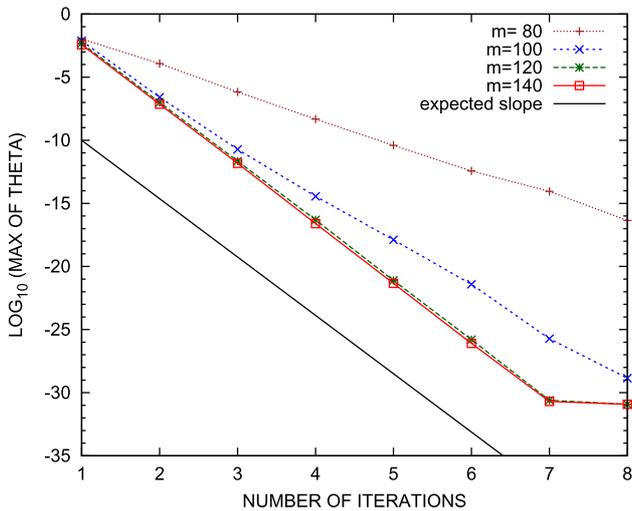


図 11 例題 R1-Q: 反復回数に対する相対残差の最大値 ($n = 10$) ($m > 106$ が望ましい)

Fig. 11 Example R1-Q: The maximum of relative residuals for iterations (case $n = 10$) ($m > 106$ is desirable).

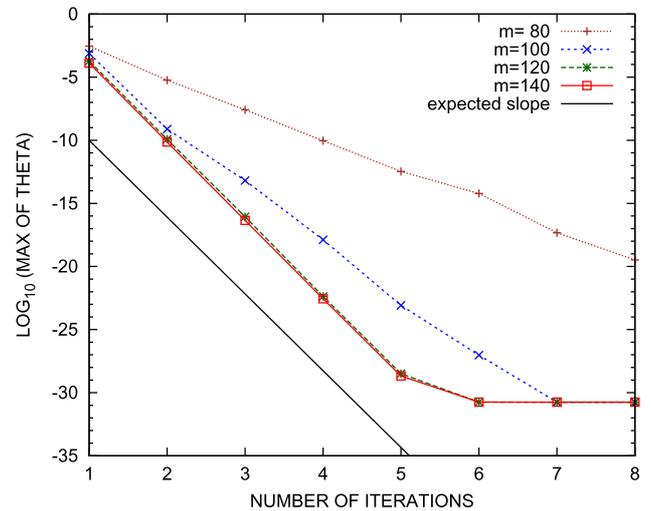


図 13 例題 R1-Q: 反復回数に対する相対残差の最大値 ($n = 20$) ($m > 106$ が望ましい)

Fig. 13 Example R1-Q: The maximum of relative residuals for iterations (case $n = 20$) ($m > 106$ is desirable).

あり、これは中規模な問題である。そうして固有値が下端の区間 $[a, b] = [0, 100]$ にある固有対を求めてみる。そのような固有対は全部で 402 個ある。

フィルタには下端固有対用のものを用いる。固有値が通過域と $\mu = 1.5$ に対する遷移域の両方をあわせた区間 $[a, b'] = [0, 150]$ にある固有対は全部で 764 個ある。そこでフィルタを適用する初期のランダムなベクトルの数 m は 764 よりも多くすることが望ましい。

例題 R2: 近似対の相対残差の最大値

フィルタの次数 n を 8, 10, 15, 20 とした各場合について、ベクトルの数 m と反復回数 IT を変えて得られた近似対のうち区間 $[a, b]$ に固有値があるものについての相対残差の最大値をそれぞれ表に示す (表 7, 表 8, 表 9, 表 10)。

フィルタの次数 n を 8, 10, 15, 20 とした各場合について、横軸に反復回数 IT を 1 から 4 までとり、縦軸に指定区間 $[a, b]$ に固有値を持つ近似対の相対残差の最大値をとって、フィルタを適用するベクトルの数 m を 500, 600, 700, 800 と 4 通りに変えて片対数でプロットしたグラフをそれぞれ示す (図 14, 図 15, 図 16, 図 17)。各グラフ中の“expected slope” とラベルされた黒線は、相対残差の最大値が反復ごとに g_s/g_p 倍に減少すると仮定した場合のグラフの傾きを示したものである。これら 4 つの図から、ベクトルの数 m が 700 までの場合と 800 の場合とでは反復回数 IT の増加にともなう相対残差の最大値の減少の様子に違いがあることが分かる。

またこれらのグラフから、ベクトルの数 m が 800 の場

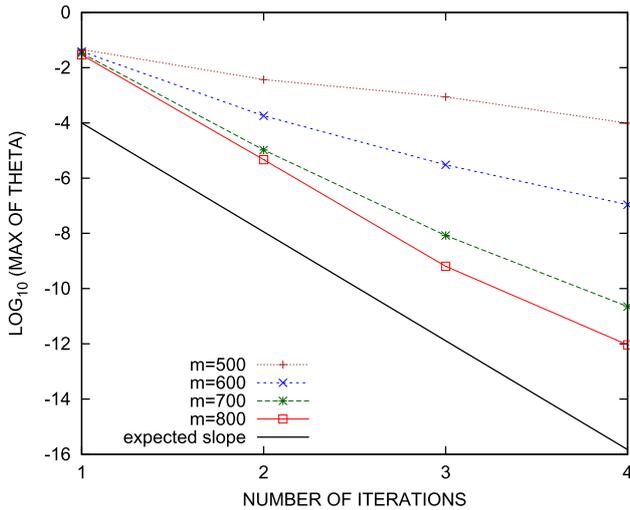


図 14 例題 R2：反復回数に対する相対残差の最大値 ($n = 8$) ($m > 764$ が望ましい)

Fig. 14 Example R2: The maximum of relative residuals for iterations (case $n = 8$) ($m > 764$ is desirable).

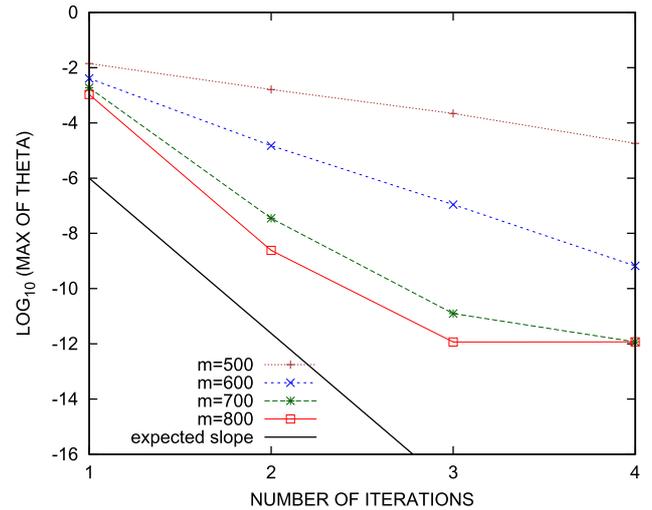


図 16 例題 R2：反復回数に対する相対残差の最大値 ($n = 15$) ($m > 764$ が望ましい)

Fig. 16 Example R2: The maximum of relative residuals for iterations (case $n = 15$) ($m > 764$ is desirable).

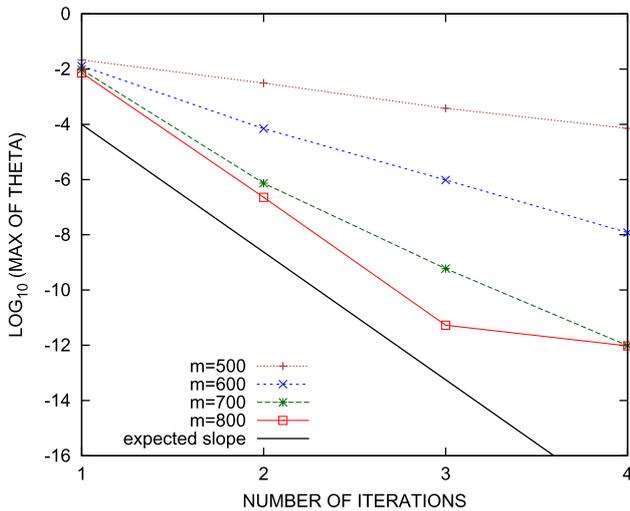


図 15 例題 R2：反復回数に対する相対残差の最大値 ($n = 10$) ($m > 764$ が望ましい)

Fig. 15 Example R2: The maximum of relative residuals for iterations (case $n = 10$) ($m > 764$ is desirable).

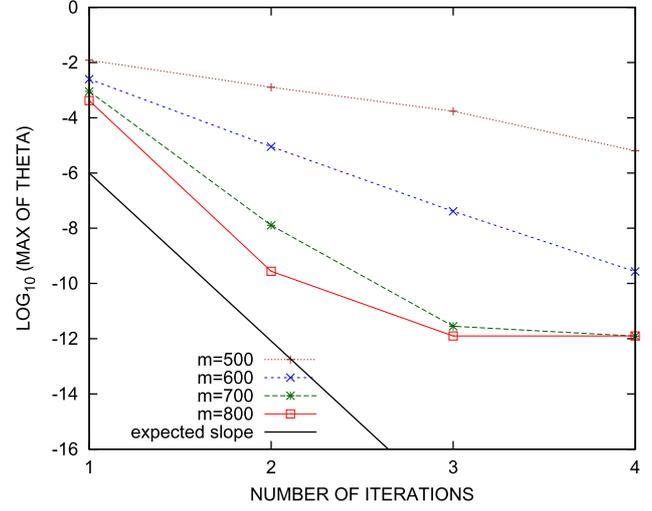


図 17 例題 R2：反復回数に対する相対残差の最大値 ($n = 20$) ($m > 764$ が望ましい)

Fig. 17 Example R2: The maximum of relative residuals for iterations (case $n = 20$) ($m > 764$ is desirable).

合には、近似対の改良はフィルタの次数が $n = 8$ では反復 4 回目まで続いているが、 $n = 10$ では反復 3 回目ですでにほぼ完了していて反復 4 回目で完了し、 $n = 15$ と $n = 20$ では反復 3 回目で完了していること、そうして完了したときの相対残差の最大値が約 10^{-12} であることが分かる。

得られた近似対のうちで固有値が指定された区間 $[0, 100]$ にあるものの数は、反復回数 IT が 2 以上の場合には実験を行ったすべての n と m の組合せについて真の固有対の数 402 と一致している。しかし反復回数 IT が 1 回の場合には一致していない場合がある (表 11, 一致していないものには下線をつけている)。これは固有値の近似が悪ければ区間に含まれる近似固有値の数が変わるためである (固有値分布の下端の区間に固有値がある固有対を求めている

表 7 例題 R2：反復回数に対する相対残差の最大値 ($n = 8$)

Table 7 Example R2: The maximum of relative residuals for iterations (case $n = 8$).

IT	$m = 500$	$m = 600$	$m = 700$	$m = 800$
1	4.6E-02	3.8E-02	3.4E-02	2.9E-02
2	3.7E-03	1.8E-04	1.1E-05	4.7E-06
3	8.7E-04	3.0E-06	8.3E-09	6.4E-10
4	9.8E-05	1.1E-07	2.3E-11	9.2E-13

場合には、フィルタ対角化法による近似対は Rayleigh-Ritz 法で求めることになるので、近似固有値はつねに真の固有値以上となるから、下端区間内にある近似固有値の数は必ず真の固有値の数以下となるので、近似が悪ければ減る傾向にある)。

表 8 例題 R2：反復回数に対する相対残差の最大値 ($n = 10$)

Table 8 Example R2: The maximum of relative residuals for iterations (case $n = 10$).

IT	$m = 500$	$m = 600$	$m = 700$	$m = 800$
1	2.2E-02	1.3E-02	9.5E-03	7.3E-03
2	3.1E-03	7.0E-05	7.4E-07	2.3E-07
3	3.8E-04	9.6E-07	6.0E-10	5.3E-12
4	7.1E-05	1.2E-08	9.7E-13	9.5E-13

表 9 例題 R2：反復回数に対する相対残差の最大値 ($n = 15$)

Table 9 Example R2: The maximum of relative residuals for iterations (case $n = 15$).

IT	$m = 500$	$m = 600$	$m = 700$	$m = 800$
1	1.4E-02	4.1E-03	1.9E-03	1.1E-03
2	1.6E-03	1.5E-05	3.5E-08	2.4E-09
3	2.2E-04	1.1E-07	1.3E-11	1.2E-12
4	1.8E-05	6.6E-10	1.2E-12	1.2E-12

表 10 例題 R2：反復回数に対する相対残差の最大値 ($n = 20$)

Table 10 Example R2: The maximum of relative residuals for iterations (case $n = 20$).

IT	$m = 500$	$m = 600$	$m = 700$	$m = 800$
1	1.2E-02	2.5E-03	9.0E-04	4.2E-04
2	1.3E-03	9.0E-06	1.3E-08	2.8E-10
3	1.7E-04	4.1E-08	2.8E-12	1.3E-12
4	6.4E-06	2.7E-10	1.3E-12	1.3E-12

表 11 例題 R2：区間内に固有値がある近似対の数 (反復回数 IT が 1 のとき)

Table 11 Example R2: Number of approximate pairs whose eigenvalues are in the interval for the case IT=1.

n	$m = 500$	$m = 600$	$m = 700$	$m = 800$
8	<u>392</u>	<u>398</u>	<u>400</u>	<u>401</u>
10	<u>394</u>	<u>401</u>	402	402
15	<u>400</u>	402	402	402
20	<u>401</u>	402	402	402

7.4 例題 R3：下端固有対を解いた例その 3

この例題 R3 では FEM の要素分割を $(N_1, N_2, N_3) = (90, 100, 110)$ とした。それから導かれる一般固有値問題の係数行列 A と B の次数 N は 990,000 で下帯幅 w_L は 9,091 であり、これは大規模な問題である。下端固有対を求める例題として、固有値が区間 $[a, b] = [0, 30]$ にあるものを求めてみる。そのような固有対は全部で 54 個ある。

フィルタには下端固有対用のものを用いる。固有値が通過域と遷移域を合併した区間 $[a, b] = [0, 45]$ に固有対は全部で 108 個あるので、フィルタを適用するベクトルの数 m は 108 よりも多くすることが望ましい。

例題 R3：近似対の相対残差の最大値

フィルタの次数 n を 8, 10, 15, 20 とした各場合について、ベクトルの数 m と反復回数 IT に対して、固有値が指

表 12 例題 R3：反復回数に対する相対残差の最大値 ($n = 8$)

Table 12 Example R3: The maximum of relative residuals for iterations (case $n = 8$).

IT	$m = 80$	$m = 90$	$m = 100$	$m = 110$	$m = 120$
1	1.1E-01	1.5E-01	1.8E-01	1.6E-01	1.7E-01
2	5.1E-04	1.9E-04	6.6E-05	7.1E-05	4.7E-05
3	3.7E-06	2.9E-08	2.3E-08	4.6E-09	3.2E-09
4	1.1E-07	1.1E-10	5.0E-12	3.3E-12	3.3E-12

表 13 例題 R3：反復回数に対する相対残差の最大値 ($n = 10$)

Table 13 Example R3: The maximum of relative residuals for iterations (case $n = 10$).

IT	$m = 80$	$m = 90$	$m = 100$	$m = 110$	$m = 120$
1	6.1E-02	9.3E-02	8.7E-02	8.3E-02	9.2E-02
2	6.8E-05	6.3E-06	2.6E-06	1.8E-06	1.3E-06
3	1.3E-06	4.3E-09	2.1E-10	7.2E-11	1.8E-11
4	2.9E-08	5.7E-12	3.4E-12	3.4E-12	3.4E-12

表 14 例題 R3：反復回数に対する相対残差の最大値 ($n = 15$)

Table 14 Example R3: The maximum of relative residuals for iterations (case $n = 15$).

IT	$m = 80$	$m = 90$	$m = 100$	$m = 110$	$m = 120$
1	3.1E-02	2.5E-02	2.0E-02	1.9E-02	1.6E-02
2	2.4E-05	6.0E-07	8.8E-08	4.9E-08	4.3E-08
3	3.9E-07	2.5E-10	5.8E-12	4.4E-12	4.4E-12
4	1.2E-09	4.4E-12	4.4E-12	4.4E-12	4.4E-12

表 15 例題 R3：反復回数に対する相対残差の最大値 ($n = 20$)

Table 15 Example R3: The maximum of relative residuals for iterations (case $n = 20$).

IT	$m = 80$	$m = 90$	$m = 100$	$m = 110$	$m = 120$
1	2.0E-02	1.7E-02	7.8E-03	8.0E-03	6.7E-03
2	1.2E-05	7.2E-08	2.9E-08	7.8E-09	4.6E-09
3	5.1E-08	3.0E-11	5.7E-12	5.7E-12	5.7E-12
4	2.7E-10	5.7E-12	5.7E-12	5.7E-12	5.7E-12

表 16 例題 R3：区間内に固有値がある近似対の数 (反復回数 IT が 1 回のとき)

Table 16 Example R3: Number of approximate pairs whose eigenvalues are in the interval for the case IT=1.

n	$m = 80$	$m = 90$	$m = 100$	$m = 110$	$m = 120$
8	<u>38</u>	<u>40</u>	<u>41</u>	<u>43</u>	<u>44</u>
10	<u>48</u>	<u>51</u>	<u>52</u>	<u>53</u>	<u>53</u>
15	54	54	54	54	54
20	54	54	54	54	54

定区間にある近似対についての相対残差の最大値を表にしたものをそれぞれ示す (表 12, 表 13, 表 14, 表 15)。

フィルタの次数 n を 8, 10, 15, 20 とした各場合について、横軸に反復回数 IT を 1 から 4 までとり、縦軸に指定した区間に固有値を持つ近似対の相対残差の最大値をとって、ベクトルの数 m を 80 から 120 まで 10 刻みで変えて片

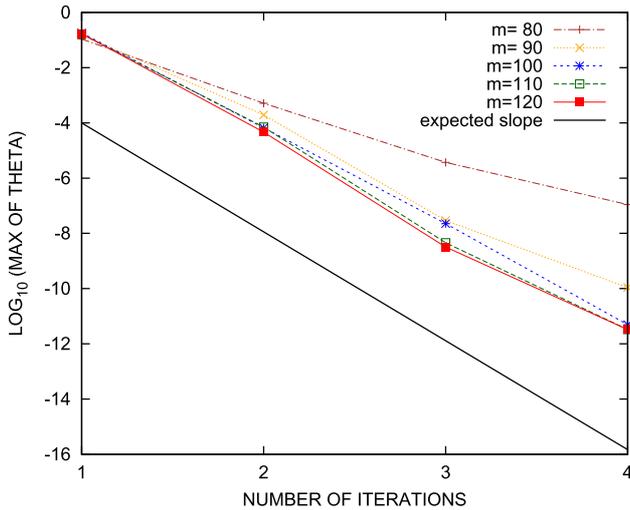


図 18 例題 R3：反復回数に対する相対残差の最大値 ($n = 8$) ($m > 108$ が望ましい)

Fig. 18 Example R3: The maximum of relative residuals for iterations (case $n = 8$) ($m > 108$ is desirable).

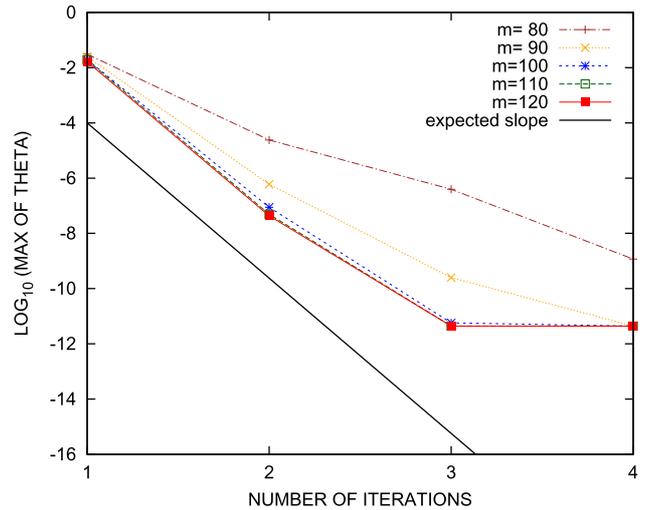


図 20 例題 R3：反復回数に対する相対残差の最大値 ($n = 15$) ($m > 108$ が望ましい)

Fig. 20 Example R3: The maximum of relative residuals for iterations (case $n = 15$) ($m > 108$ is desirable).

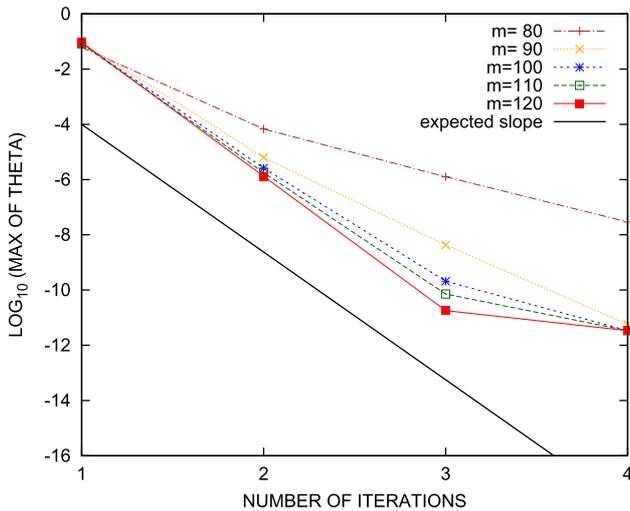


図 19 例題 R3：反復回数に対する相対残差の最大値 ($n = 10$) ($m > 108$ が望ましい)

Fig. 19 Example R3: The maximum of relative residuals for iterations (case $n = 10$) ($m > 108$ is desirable).

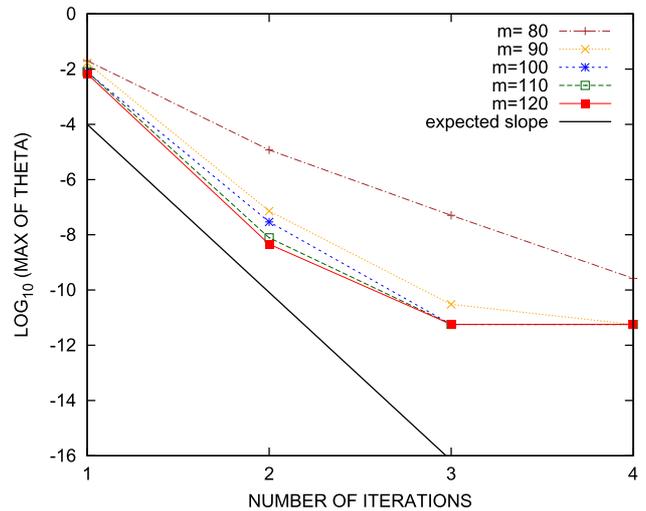


図 21 例題 R3：反復回数に対する相対残差の最大値 ($n = 20$) ($m > 108$ が望ましい)

Fig. 21 Example R3: The maximum of relative residuals for iterations (case $n = 20$) ($m > 108$ is desirable).

対数でプロットしたグラフをそれぞれ示す (図 18, 図 19, 図 20, 図 21). 各グラフ中の “expected slope” とラベルされた黒線は、相対残差の最大値が反復ごとに g_s/g_p 倍に減少すると仮定した場合のグラフの傾きを示したものである。これら 4 つの図から、ベクトルの数 m が 100 以下の場合と 110 以上の場合とでは反復回数 IT の増加にともなう相対残差の最大値の減少の様子が異なっていること、そして m が 110 と 120 の場合の相対残差の最大値の振舞いがほぼ同じであることが分かる。

得られた近似対のうちで固有値が指定された区間にあるものの数は、反復回数 IT が 2 回以上の場合には (実験を行ったすべての n と m 組合せについて) 真の固有対の数 54 と一致したが、反復回数 IT が 1 回の場合には n が 15

と 20 のときは一致しているが、 n が 8 と 10 のときには一致していない (表 16, 一致していないものには下線をつけてある)。これは近似が悪いほど固有値がずれて、固有値が区間内に入る近似対の数が増える傾向があるからである。

7.5 例題 C1：中間固有対を解いた例その 1

この例題 C1 は例題 R1 と同じ FEM の要素分割 $(N_1, N_2, N_3) = (20, 30, 40)$ から導かれた一般固有値問題である。その係数行列 A と B の次数 N は 24,000 で下帯幅 w_L は 621 であり、これは小規模な問題である。そして中間固有対を解く例題として、固有値が区間 $[a, b] = [300, 310]$ にある固有対を求めてみる。そのような固有対は全部で 90

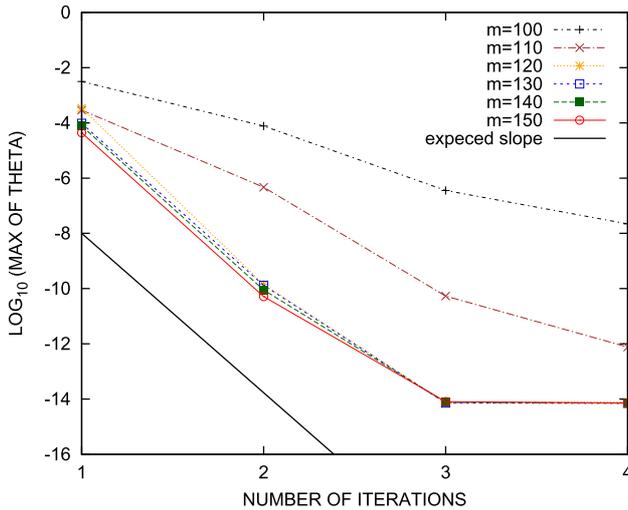


図 22 例題 C1：反復回数に対する相対残差の最大値 ($n = 8$) ($m > 125$ が望ましい)

Fig. 22 Example C1: The maximum of relative residuals for iterations (case $n = 8$) ($m > 125$ is desirable).

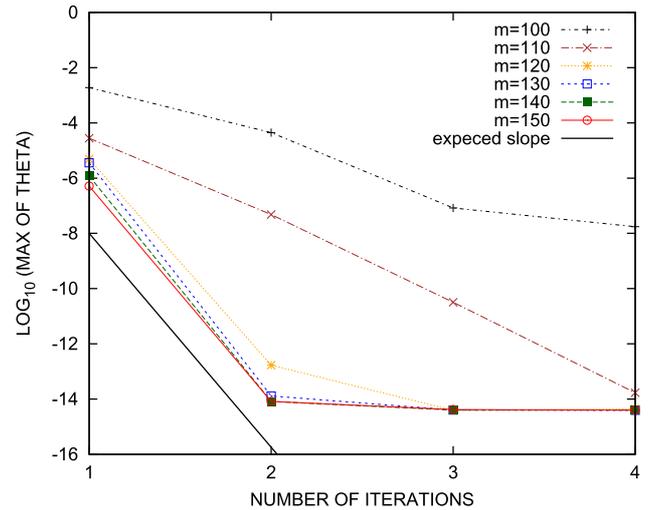


図 24 例題 C1：反復回数に対する相対残差の最大値 ($n = 15$) ($m > 125$ が望ましい)

Fig. 24 Example C1: The maximum of relative residuals for iterations (case $n = 15$) ($m > 125$ is desirable).

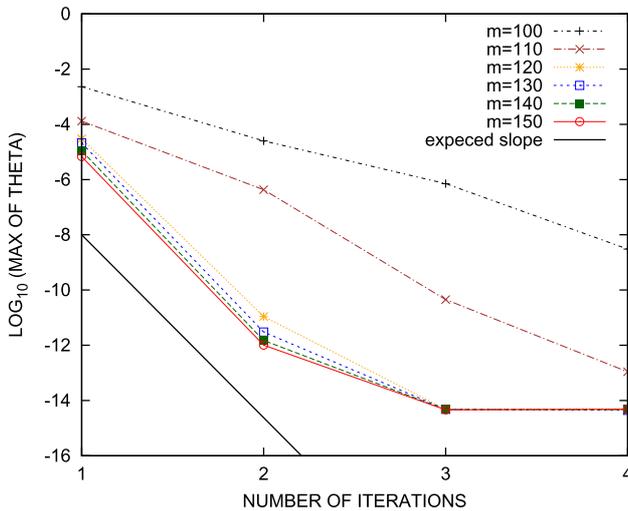


図 23 例題 C1：反復回数に対する相対残差の最大値 ($n = 10$) ($m > 125$ が望ましい)

Fig. 23 Example C1: The maximum of relative residuals for iterations (case $n = 10$) ($m > 125$ is desirable).

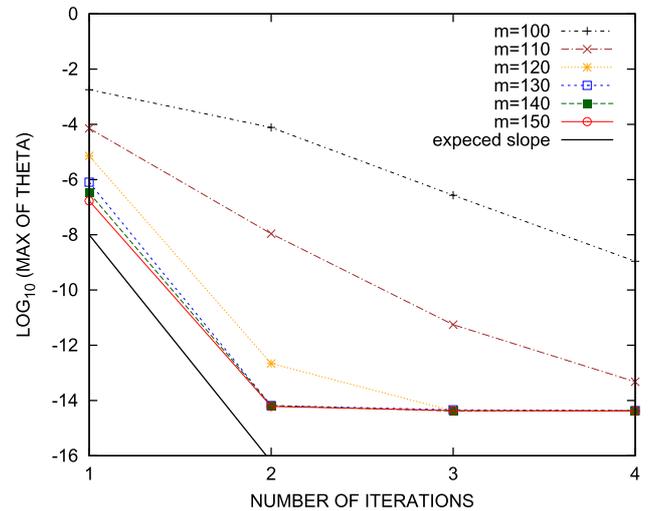


図 25 例題 C1：反復回数に対する相対残差の最大値 ($n = 20$) ($m > 125$ が望ましい)

Fig. 25 Example C1: The maximum of relative residuals for iterations (case $n = 20$) ($m > 125$ is desirable).

個ある。

フィルタは中間固有対用のものを用いる。通過域と $\mu = 1.5$ に対応する遷移域の両方をあわせた区間 $[a', b'] = [297.5, 312.5]$ に固有値がある固有対は全部で 125 個ある。そのため、フィルタを適用する最初のランダムなベクトルの数 m は 125 よりも多くすることが望ましい。

例題 C1：近似対の相対残差の最大値

フィルタの次数 n を 8, 10, 15, 20 とした各場合について、ベクトルの数 m と反復回数 IT に対して、得られた近似対のうち固有値が指定区間 $[a, b]$ にあるものについての相対残差の最大値をそれぞれ表に示す (表 17, 表 18, 表 19, 表 20)。

そうして、フィルタの次数 n が 8, 10, 15, 20 の各場合

について、横軸に反復回数 IT を 1 から 4 までとり、縦軸には指定区間に固有値がある近似対の相対残差の最大値をとって、ベクトルの数 m を 100 から 150 まで 10 刻みで変えて片対数でプロットしたグラフをそれぞれ示す (図 22, 図 23, 図 24, 図 25)。各グラフ中の“expected slope”とラベルされた黒線は、相対残差の最大値が反復ごとに g_s/g_p 倍に減少すると仮定した場合のグラフの傾きを示したものである。これら 4 つのグラフから、ベクトルの数 m が 120 未満の場合と以上の場合とでは反復回数 IT の増加にともなう相対残差の最大値の減少の様子がかなり異なること、そうして相対残差の最大値の振舞いは、 m が 120 以上の各場合にはあまり違いがない様子が分かる。

またこれらのグラフから、ベクトルの数 m が 120 以上

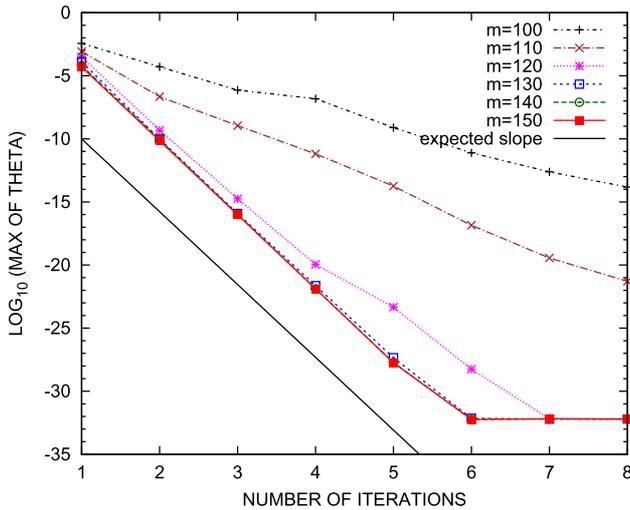


図 26 例題 C1-Q: 反復回数に対する相対残差の最大値 ($n = 8$) ($m > 125$ が望ましい)

Fig. 26 Example C1-Q: The maximum of relative residuals for iterations (case $n = 8$) ($m > 125$ is desirable).

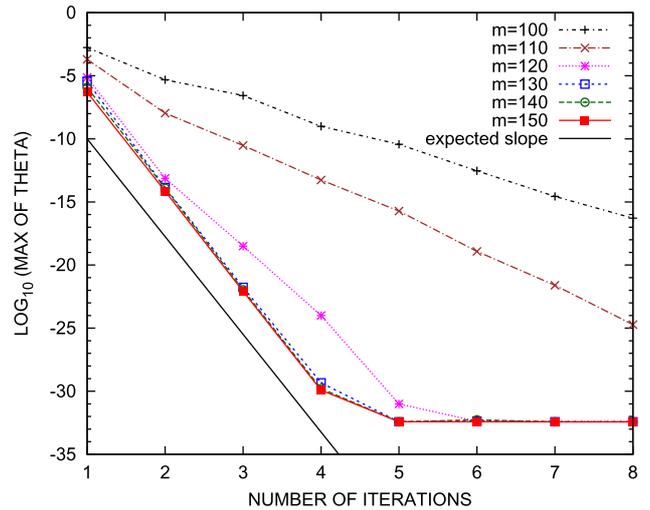


図 28 例題 C1-Q: 反復回数に対する相対残差の最大値 ($n = 15$) ($m > 125$ が望ましい)

Fig. 28 Example C1-Q: The maximum of relative residuals for iterations (case $n = 15$) ($m > 125$ is desirable).

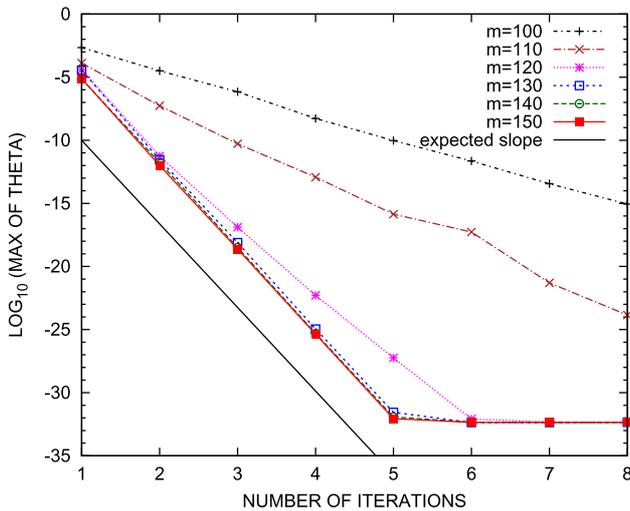


図 27 例題 C1-Q: 反復回数に対する相対残差の最大値 ($n = 10$) ($m > 125$ が望ましい)

Fig. 27 Example C1-Q: The maximum of relative residuals for iterations (case $n = 10$) ($m > 125$ is desirable).

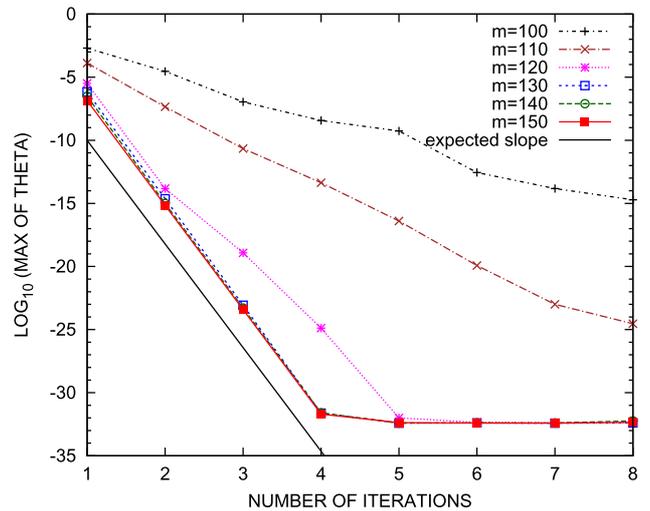


図 29 例題 C1-Q: 反復回数に対する相対残差の最大値 ($n = 20$) ($m > 125$ が望ましい)

Fig. 29 Example C1-Q: The maximum of relative residuals for iterations (case $n = 20$) ($m > 125$ is desirable).

の場合に、近似対の改良はフィルタの次数 n が 8 と 10 のときは反復 3 回目で、次数 n が 15 と 20 のときは m が 120 の場合には反復 3 回目で、 m が 130 以上の場合には反復 2 回目でそれぞれ完了していて、そのとき相対残差の最大値が 10^{-14} 程度であることが分かる。

例題 C1 について行ったすべての実験の組合せ (反復回数 IT, フィルタの次数 n , ベクトルの数 m) において、固有値が指定された区間にある近似対の数は、真の固有対の数 90 と一致した。

例題 C1-Q: 近似対の相対残差の最大値 (四倍精度計算)

ここでは特別に計算に用いる浮動小数点の数値と演算を通常の IEEE-754 規格の 2 進 64 bit 倍精度からすべて同規格の 2 進 128 bit 四倍精度に単純に置き換えて上記の中

間固有対を解く例題 C1 と同じ問題を解いてみた例を示す (これを例題 C1-Q と呼ぶことにする)。

フィルタの次数 n を 8, 10, 15, 20 とした各場合について、横軸に反復回数 IT を 1 から 8 までとり、縦軸には指定区間に固有値がある近似対の相対残差の最大値をとって、ベクトルの数 m を 100 から 150 まで 10 刻みで変えて片対数でプロットしたグラフをそれぞれ示す (図 26, 図 27, 図 28, 図 29)。各グラフ中の “expected slope” とラベルされた黒線は、相対残差の最大値が反復ごとに g_s/g_p 倍に減少すると仮定した場合のグラフの傾きを示したものである。

これら 4 つのグラフのそれぞれについて、ベクトルの数 m が 130 以上の場合には相対残差の最大値はほとんど同じ

表 17 例題 C1：反復回数に対する相対残差の最大値 ($n = 8$)

Table 17 Example C1: The maximum of relative residuals for iterations (case $n = 8$).

IT	$m=100$	$m=110$	$m=120$	$m=130$	$m=140$	$m=150$
1	3.2E-03	2.9E-04	3.4E-04	9.6E-05	7.8E-05	4.4E-05
2	7.7E-05	4.7E-07	1.4E-10	1.3E-10	8.9E-11	5.2E-11
3	3.6E-07	5.4E-11	7.4E-15	7.2E-15	7.6E-15	8.1E-15
4	2.2E-08	7.6E-13	8.0E-15	7.0E-15	7.2E-15	7.1E-15

表 18 例題 C1：反復回数に対する相対残差の最大値 ($n = 10$)

Table 18 Example C1: The maximum of relative residuals for iterations (case $n = 10$).

IT	$m=100$	$m=110$	$m=120$	$m=130$	$m=140$	$m=150$
1	2.3E-03	1.3E-04	3.0E-05	2.1E-05	1.1E-05	6.8E-06
2	2.5E-05	4.3E-07	1.1E-11	3.0E-12	1.5E-12	1.0E-12
3	7.2E-07	4.5E-11	4.8E-15	4.8E-15	4.7E-15	4.6E-15
4	3.0E-09	1.1E-13	5.1E-15	4.4E-15	4.7E-15	4.8E-15

表 19 例題 C1：反復回数に対する相対残差の最大値 ($n = 15$)

Table 19 Example C1: The maximum of relative residuals for iterations (case $n = 15$).

IT	$m=100$	$m=110$	$m=120$	$m=130$	$m=140$	$m=150$
1	1.9E-03	2.8E-05	4.9E-06	3.6E-06	1.2E-06	5.2E-07
2	4.5E-05	4.8E-08	1.7E-13	1.3E-14	8.2E-15	8.3E-15
3	8.3E-08	3.2E-11	3.9E-15	4.0E-15	4.0E-15	4.2E-15
4	1.7E-08	1.7E-14	4.4E-15	3.9E-15	4.0E-15	4.0E-15

表 20 例題 C1：反復回数に対する相対残差の最大値 ($n = 20$)

Table 20 Example C1: The maximum of relative residuals for iterations (case $n = 20$).

IT	$m=100$	$m=110$	$m=120$	$m=130$	$m=140$	$m=150$
1	1.8E-03	7.1E-05	7.2E-06	8.0E-07	3.3E-07	1.7E-07
2	7.7E-05	1.1E-08	2.2E-13	6.6E-15	6.5E-15	6.1E-15
3	2.7E-07	5.6E-12	4.2E-15	4.6E-15	4.2E-15	4.3E-15
4	1.1E-09	4.8E-14	4.3E-15	4.3E-15	4.3E-15	4.2E-15

になっていて、グラフの線が重なっている。そうして反復ごとに相対残差の最大値がほぼ g_s/g_p 倍になって減少していることが分かる。

7.6 例題 C2：中間固有対を解いた例その 2

この例題 C2 もまた例題 R1, C1 と同じ FEM の要素分割 $(N_1, N_2, N_3) = (20, 30, 40)$ から導かれた一般固有値問題である。その係数行列 A と B の次数 N は 24,000 で下帯幅 w_L は 621 であり、これは小規模な問題である。そうして中間固有対を求める例題として、固有値が区間 $[a, b] = [1000, 1010]$ にある固有対を求めてみる。そのような固有対は全部で 92 個ある。

フィルタは中間固有対用のものを用いる。固有値が通過域と $\mu = 1.5$ に対応する遷移域の両方をあわせた区間 $[a', b'] = [997.5, 1012.5]$ にある固有対は全部で 145 個ある

表 21 例題 C2：反復回数に対する相対残差の最大値 ($n = 8$)

Table 21 Example C2: The maximum of relative residuals for iterations (case $n = 8$).

IT	$m=100$	$m=110$	$m=120$	$m=130$	$m=140$	$m=150$
1	1.9E-03	4.5E-04	1.6E-04	9.4E-05	5.5E-05	4.3E-05
2	9.4E-05	2.6E-06	2.2E-07	3.2E-09	5.4E-11	3.7E-11
3	3.0E-05	6.4E-08	1.1E-09	6.2E-12	1.5E-14	1.5E-14
4	2.2E-06	1.2E-09	1.0E-11	1.5E-14	1.6E-14	1.5E-14

表 22 例題 C2：反復回数に対する相対残差の最大値 ($n = 10$)

Table 22 Example C2: The maximum of relative residuals for iterations (case $n = 10$).

IT	$m=100$	$m=110$	$m=120$	$m=130$	$m=140$	$m=150$
1	1.7E-03	2.0E-04	8.0E-05	2.0E-05	1.6E-05	1.3E-05
2	1.7E-04	9.6E-07	7.2E-08	1.0E-09	1.6E-12	7.9E-13
3	1.8E-05	8.3E-09	1.3E-10	2.9E-13	8.9E-15	9.1E-15
4	1.0E-06	8.6E-10	1.7E-12	9.3E-15	9.4E-15	8.7E-15

表 23 例題 C2：反復回数に対する相対残差の最大値 ($n = 15$)

Table 23 Example C2: The maximum of relative residuals for iterations (case $n = 15$).

IT	$m=100$	$m=110$	$m=120$	$m=130$	$m=140$	$m=150$
1	1.1E-03	1.6E-04	4.8E-05	6.7E-06	1.8E-06	1.0E-06
2	1.1E-04	6.8E-07	3.8E-08	9.6E-11	1.9E-14	8.0E-15
3	5.5E-06	4.4E-09	1.9E-10	9.2E-15	5.8E-15	5.6E-15
4	1.9E-06	1.4E-10	1.3E-13	5.7E-15	5.5E-15	5.8E-15

表 24 例題 C2：反復回数に対する相対残差の最大値 ($n = 20$)

Table 24 Example C2: The maximum of relative residuals for iterations (case $n = 20$).

IT	$m=100$	$m=110$	$m=120$	$m=130$	$m=140$	$m=150$
1	1.2E-03	2.8E-04	3.7E-05	5.4E-06	6.5E-07	2.3E-07
2	1.0E-04	7.8E-07	3.8E-08	3.2E-10	6.8E-15	6.9E-15
3	2.9E-05	9.4E-09	1.2E-10	5.2E-14	5.4E-15	5.1E-15
4	1.0E-06	4.6E-11	7.8E-14	5.1E-15	5.2E-15	5.4E-15

ので、フィルタを適用する初期ベクトルの数 m は 145 よりも多くすることが望ましい。

例題 C2：近似対の相対残差の最大値

フィルタの次数 n を 8, 10, 15, 20 とした各場合について、ベクトルの数 m と反復回数 IT に対する相対残差の最大値をそれぞれ表に示す (表 21, 表 22, 表 23, 表 24)。

フィルタの次数 n を 8, 10, 15, 20 とした各場合について、横軸に反復回数 IT を 1 から 4 までとり、縦軸には固有値が指定区間にある近似対の相対残差の最大値をとって、ベクトルの数 m を 100 から 150 まで 10 刻みで変えて片対数でプロットしたグラフをそれぞれ示す (図 30, 図 31, 図 32, 図 33)。各グラフ中の“expected slope”とラベルされた黒線は、相対残差の最大値が反復ごとに g_s/g_p 倍に減少すると仮定した場合のグラフの傾きを示したものである。これら 4 つのグラフから、ベクトルの数 m が 140 未

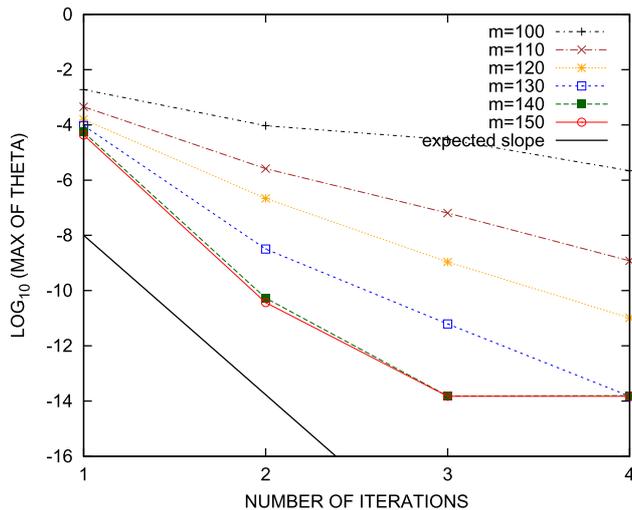


図 30 例題 C2：反復回数に対する相対残差の最大値 ($n = 8$) ($m > 145$ が望ましい)

Fig. 30 Example C2: The maximum of relative residuals for iterations (case $n = 8$) ($m > 145$ is desirable).

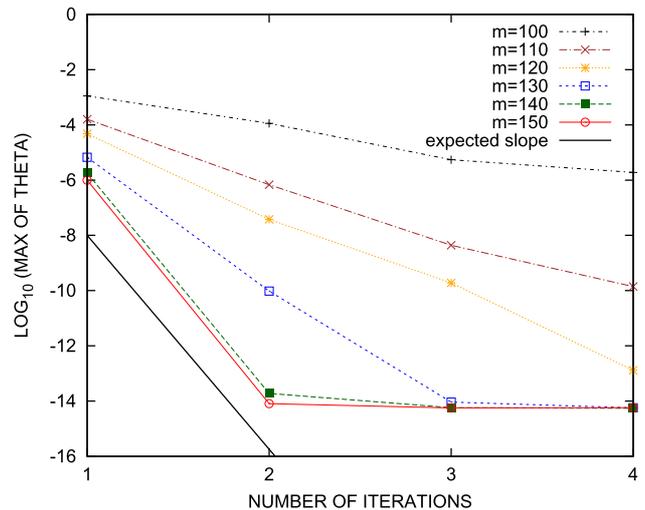


図 32 例題 C2：反復回数に対する相対残差の最大値 ($n = 15$) ($m > 145$ が望ましい)

Fig. 32 Example C2: The maximum of relative residuals for iterations (case $n = 15$) ($m > 145$ is desirable).

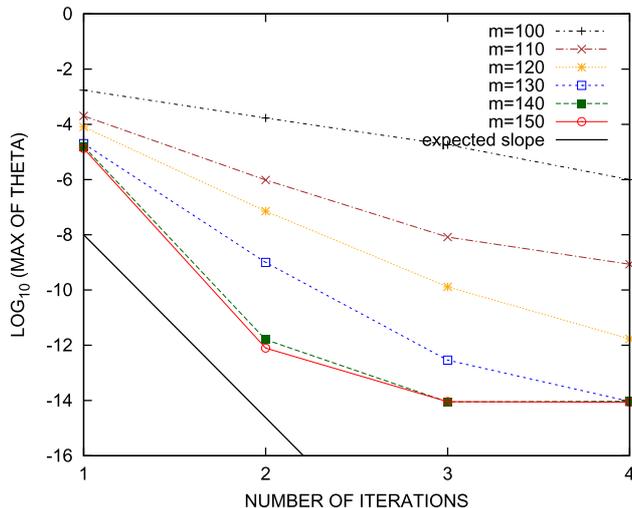


図 31 例題 C2：反復回数に対する相対残差の最大値 ($n = 10$) ($m > 145$ が望ましい)

Fig. 31 Example C2: The maximum of relative residuals for iterations (case $n = 10$) ($m > 145$ is desirable).

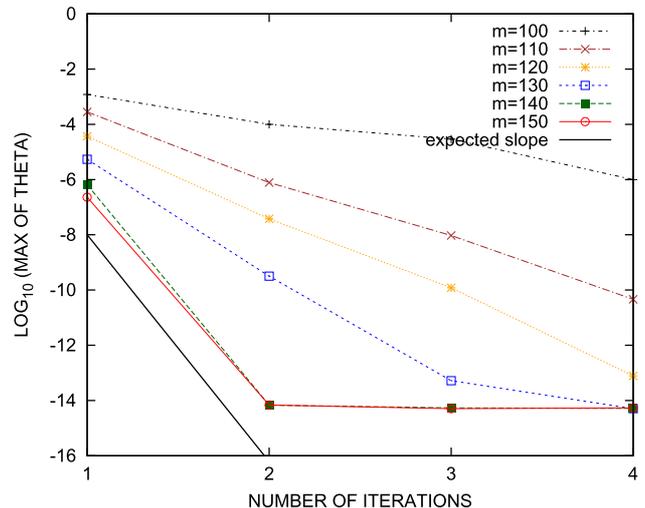


図 33 例題 C2：反復回数に対する相対残差の最大値 ($n = 20$) ($m > 145$ が望ましい)

Fig. 33 Example C2: The maximum of relative residuals for iterations (case $n = 20$) ($m > 145$ is desirable).

満の場合と 140 以上の場合とでは反復回数 IT の増加ともなう相対残差の最大値の減少の様子がかなり違うこと、そうして m が 140 と 150 の場合には違いがほとんどないことが分かる。

またさらにこれらのグラフから、ベクトルの数 m が 140 以上のときに、近似対の改良はフィルタの次数 n が 8 と 10 のときには反復 3 回目で完了しているが、 n が 15 のときにはすでに反復 2 回目でほぼ完了、 n が 20 のときにも反復 2 回目で完了していることが分かる。

この例題 C2 について行ったすべての実験の組合せ (反復回数 IT, フィルタの次数 n , ベクトルの数 m) において、固有値が指定された区間にある近似対の数は、真の固有対の数 92 と一致した。

7.7 例題 C3：中間固有対を解いた例その 3

この例題 C3 は例題 R2 と同じ FEM の要素分割 $(N_1, N_2, N_3) = (50, 60, 70)$ から導かれた一般固有値問題である。その係数行列 A と B の次数 N は 210,000 で下帯幅 w_L は 3,051 であり、中規模な問題である。そうして中間固有対を解く例題として、固有値が区間 $[a, b] = [100, 200]$ にある固有対を求めてみる。そのような固有対は全部で 801 個ある。

フィルタは中間固有対用のものを用いる。固有値が通過域と $\mu = 1.5$ に対応する遷移域の両方をあわせた区間 $[75, 225]$ に固有対は全部で 1,192 個あるので、フィルタを適用する初期ベクトルの数 m は 1,192 よりも多くすることが望ましい。

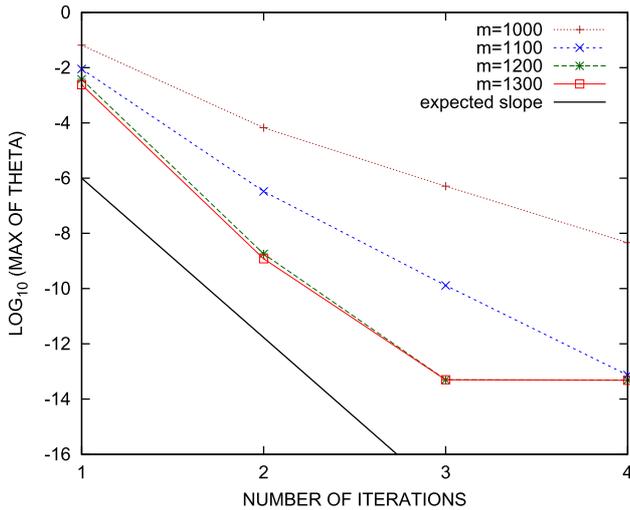


図 34 例題 C3：反復回数に対する相対残差の最大値 ($n = 8$) ($m > 1,192$ が望ましい)

Fig. 34 Example C3: The maximum of relative residuals for iterations (case $n = 8$) ($m > 1,192$ is desirable).

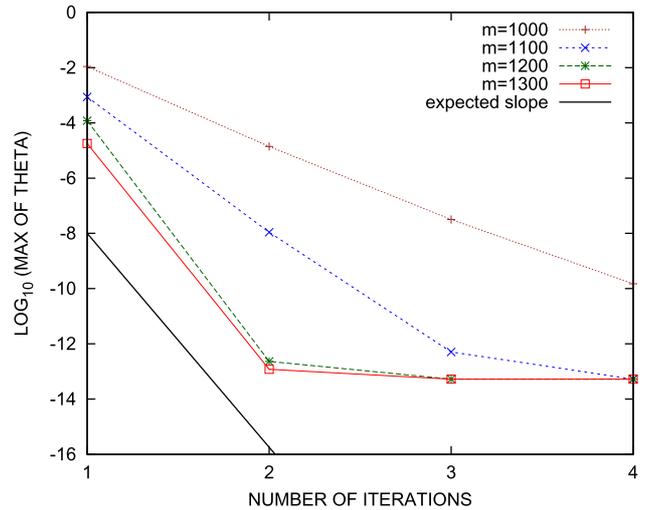


図 36 例題 C3：反復回数に対する相対残差の最大値 ($n = 15$) ($m > 1,192$ が望ましい)

Fig. 36 Example C3: The maximum of relative residuals for iterations (case $n = 15$) ($m > 1,192$ is desirable).

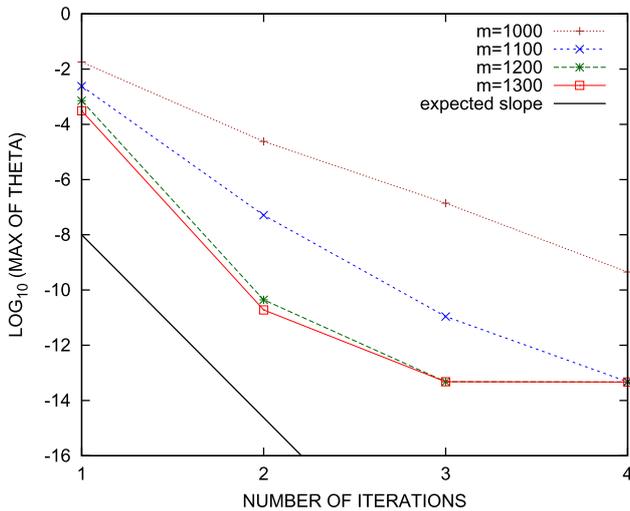


図 35 例題 C3：反復回数に対する相対残差の最大値 ($n = 10$) ($m > 1,192$ が望ましい)

Fig. 35 Example C3: The maximum of relative residuals for iterations (case $n = 10$) ($m > 1,192$ is desirable).

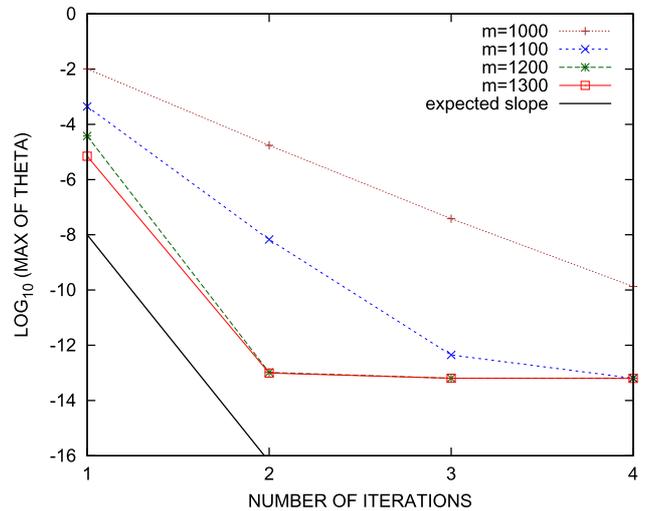


図 37 例題 C3：反復回数に対する相対残差の最大値 ($n = 20$) ($m > 1,192$ が望ましい)

Fig. 37 Example C3: The maximum of relative residuals for iterations (case $n = 20$) ($m > 1,192$ is desirable).

例題 C3：近似対の相対残差の最大値

フィルタの次数 n を 8, 10, 15, 20 とした各場合について、ベクトルの数 m を 1,000 から 1,300 まで 100 刻みで変えて、反復回数 IT を 1 から 4 までとしたものに対する相対残差の最大値を、それぞれ表に示す (表 25, 表 26, 表 27, 表 28)。

またフィルタの次数 n を 8, 10, 15, 20 とした各場合について、横軸に反復回数 IT を 1 から 4 までとり、縦軸に指定した区間に固有値を持つ近似対の相対残差の最大値をとって、ベクトルの数 m を 1,000 から 1,300 まで 100 刻みで変えて片対数でプロットしたグラフをそれぞれ示す (図 34, 図 35, 図 36, 図 37)。各グラフ中の“expected slope”とラベルされた黒線は、相対残差の最大値が反復ご

とに g_s/g_p 倍に減少すると仮定した場合のグラフの傾きを示したものである。これら 4 つのグラフから、ベクトルの数 m が 1,200 未満の場合と 1,200 以上の場合とでは反復回数 IT の増加にともなう相対残差の最大値の減少の様子がかなり違っていること、そして m が 1,200 と 1,300 の場合の結果にはほとんど違いがない様子が分かる。

これらのグラフから、 m が 1,200 以上の場合には、近似対の改良はフィルタの次数 n が 8 と 10 のときには反復 3 回目まで完了して相対残差の最大値は 10^{-13} より小さいこと、 n が 15 のときには反復 2 回目ではほぼ完了して相対残差の最大値は 10^{-13} 程度であることが分かる。

得られた近似対のうち固有値が指定された区間にあるものの数は、ここで示した例題 C3 で行ったすべての実験の

表 25 例題 C3：反復回数に対する相対残差の最大値 ($n = 8$)
Table 25 Example C3: The maximum of relative residuals for iterations (case $n = 8$).

IT	$m=1000$	$m=1100$	$m=1200$	$m=1300$
1	6.5E-02	9.0E-03	3.8E-03	2.4E-03
2	6.7E-05	3.3E-07	1.8E-09	1.2E-09
3	5.1E-07	1.3E-10	5.0E-14	5.0E-14
4	4.6E-09	7.4E-14	4.9E-14	4.8E-14

表 26 例題 C3：反復回数に対する相対残差の最大値 ($n = 10$)
Table 26 Example C3: The maximum of relative residuals for iterations (case $n = 10$).

IT	$m=1000$	$m=1100$	$m=1200$	$m=1300$
1	1.8E-02	2.4E-03	7.2E-04	3.1E-04
2	2.4E-05	5.1E-08	4.4E-11	1.9E-11
3	1.4E-07	1.1E-11	4.8E-14	4.8E-14
4	4.5E-10	4.8E-14	4.7E-14	4.7E-14

表 27 例題 C3：反復回数に対する相対残差の最大値 ($n = 15$)
Table 27 Example C3: The maximum of relative residuals for iterations (case $n = 15$).

IT	$m=1000$	$m=1100$	$m=1200$	$m=1300$
1	1.1E-02	8.7E-04	1.2E-04	1.8E-05
2	1.4E-05	1.1E-08	2.3E-13	1.2E-13
3	3.2E-08	5.1E-13	5.3E-14	5.3E-14
4	1.5E-10	5.3E-14	5.3E-14	5.3E-14

表 28 例題 C3：反復回数に対する相対残差の最大値 ($n = 20$)
Table 28 Example C3: The maximum of relative residuals for iterations (case $n = 20$).

IT	$m=1000$	$m=1100$	$m=1200$	$m=1300$
1	1.0E-02	4.4E-04	3.8E-05	7.0E-06
2	1.7E-05	6.7E-09	1.0E-13	9.9E-14
3	3.8E-08	4.5E-13	6.4E-14	6.4E-14
4	1.3E-10	6.4E-14	6.4E-14	6.4E-14

組合せの範囲では、反復回数 IT が 1 回でフィルタの次数 n が 8 でベクトルの数 m が 1,000 のときにだけ真の固有対の数 801 よりも 1 つだけ多い 802 個になったが、それ以外のすべての場合の組合せについては真の固有対の数 801 に一致した。

8. まとめ

フィルタ対角化法を用いて、実対称定値一般固有値問題の固有対で固有値が指定された区間にあるものを選んで解く方法について扱った。フィルタはレゾルベントを利用して構成するものとする。本論文ではレゾルベントの作用を実現するための連立 1 次方程式は直接法を用いて解くことを前提としている（反復法を用いて解く場合については扱っていない）。その場合、大規模な問題では行列分解に掛かる計算量と特に分解結果を保持するための記憶容量が

計算を実施するうえでの制約要因になりがちである。必要な記憶容量を主記憶上に確保することができない場合には外部記憶を使うことも考えられるが、記憶の転送速度が遅いなどの理由により処理を諦めることになりがちである。そこでそのような場合に向けて、用いるレゾルベントの数（行列分解を行う行列の数）を少なくすることを考察した。

良い特性を持つようにするため、シフトの異なる複数のレゾルベントの線形結合として構成されたフィルタの場合には、連立 1 次方程式を直接法で解くとすると行列分解 1 つ分を格納できる記憶場所が確保できれば、一時には行列分解を 1 つだけ作り、それをただちに用いて連立 1 次方程式を解くことを順次行うことにより、レゾルベントの作用の線形結合を蓄積して得ることでフィルタの作用を実現できる（たとえば優れた特性を持つ楕円フィルタのように、複数のレゾルベントの線形結合であるフィルタであればそれが可能である）。しかしその場合には、用いるレゾルベントと同じ数だけの行列分解を行う必要がある。連立 1 次方程式の直接法による解法では行列分解の演算量が演算量のかなりのあるいはほとんどの割合を占めてしまう可能性が高い。もちろんここでは連立 1 次方程式を直接法を用いて解くことを前提としているので、少なくとも 1 つの行列分解を行うわけであるが、レゾルベントが複数でたとえば 8~12 個の場合は、それだけの数の行列分解を行う必要がある。複数のレゾルベントの線形結合によるフィルタは、もしもフィルタを反復して適用するのであれば、行列分解を格納するための記憶容量が 1 つ分しかない場合には、構成した行列分解を用いて連立 1 次方程式を解き終えたらただちにその構成した行列分解の記憶場所を次の行列の分解を保持するために明け渡す必要がある。すると行列分解の結果は再利用できず、行列分解の回数がフィルタを反復する回数に比例して増えるので、それにより分解に掛かる演算量の合計もまた増えることになる。

少数個（2~4 程度）のレゾルベントの線形結合の多項式型フィルタについても、もしも記憶容量の限界によりレゾルベントに対応するシフト行列の分解を一時には 1 つしか収容できない場合には、やはり行列分解は再利用ができず、各レゾルベントに対応するシフト行列の行列分解を多項式の次数と同じ回数ずつ行う必要がある。たとえば、フィルタがシフトの異なる少数 k 個のレゾルベントの線形結合（の実部）の p 次多項式である場合には、 k 個のレゾルベントに対応するシフト行列の分解をすべて主記憶上に格納できるのであれば、その k 個の行列をそれぞれ 1 回ずつだけ分解すればよいが、たとえば主記憶には行列分解が同時には 1 つ分しか入らない場合には、 k 通りのシフト行列の行列分解を p 回ずつ行う必要があり、行列分解の処理を合計 kp 回だけ行う必要がある。

そこで必要な行列分解の数が最も少ないフィルタとして単一のレゾルベントの多項式型のものを利用することが考

えられる。そのようなフィルタではレゾルベントに対応するシフト行列の行列分解はただ1つだけ行えばよく、その分解結果を主記憶上に格納することができれば、それを利用して単一のレゾルベントの多項式を作用させることができる。しかしその反面、そのような構成のフィルタではあまり良い特性が実現できないので、これまでは残差の十分に小さい近似対は得られなかった（もちろん近似対にそれほど高い精度が要求されない問題や用途もあると思われる）。

そこで今回は、通過域における伝達率の最大最小比が大きくて特性の良くないフィルタと B -正規直交化の処理の組合せを反復することを試みた。フィルタを反復する際に、すでに行列分解が得られていればそれを再利用すればよいので、必要な行列分解は全体で1回だけになる。フィルタの多項式としての次数を p とし、反復回数を IT とすれば、行列分解を1回だけ行った後に、係数行列が共通で右辺ベクトルが m 個ある連立1次方程式を（前進後退代入で）解く回数は、全部で $IT \times p$ 回になる。そのほか、フィルタの適用前には必ず m 個のベクトルの組に対する B -正規直交化が必要になる。

フィルタを適用するベクトルの数 m を、フィルタの通過域と遷移域をあわせた区間に固有値がある固有対の数よりも多くすれば、反復により相対残差はよく減少して、近似対の相対残差の最大値が反復ごとにフィルタの形状パラメタの比 g_s/g_p を乗じた値以下になる振舞い（線形収束）を確認した。つまり用いるフィルタの通過域における特性がそれほど良くななくても、 B -正規直交化とフィルタの適用を組み合わせた処理を数回程度反復することで残差の小さい改良された近似対が得られる。

フィルタを適用する回数を何回にするのかは、近似対の相対残差をどの程度まで小さくすれば目的に対して十分とするのかによるが、そのほかにも用いる初期ベクトルの組の質にも依存する。現在は初期ベクトルの組は一樣乱数を用いて作っているため、フィルタを1回適用した段階での近似対の相対残差の大きさを事前に見積もることは一般には難しい。しかしフィルタを1回適用した結果から得られた近似対の相対残差がある程度小さければ、必要な近似対の相対残差の最大値がフィルタの反復ごとに減少する比率は、 g_s/g_p で上から抑えられることが分かった。用いたフィルタ特性の良さの程度にもよるが、本論文で扱った例題とフィルタの場合には倍精度計算では少数2~4回程度のフィルタの適用で必要な近似対の相対残差の最大値を十分小さくできることを実験で確認した。

参考文献

- [1] Rutishauser, H.: The Jacobi Method for Real Symmetric Matrices, *Numer. Math.*, Vol.9, pp.1-10 (1966). Also in book: Bauer, F.L. (Eds.) *Handbook for Automatic Computation*, Vol.II: *Linear Algebra*, Springer-Verlag (1971).
- [2] Rutishauser, H.: Computational Aspects of F.L. Bauer's Simultaneous Iteration Method, *Numer. Math.*, Vol.13, No.1, pp.4-13 (1969).
- [3] Rutishauser, H.: Simultaneous Iteration Method for Symmetric Matrices, in *Handbook for Automatic Computation*, Vol.II: *Linear Algebra*, pp.284-302, Springer-Verlag (1971). Reprinted from *Numer. Math.*, Vol.16, pp.205-223 (1970).
- [4] 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).
- [5] Polizzi, E.: A Density Matrix-Based Algorithm for Solving Eigenvalue Problems, *Phys. Rev. B*, Vol.79, No.1, pp.115112-115117 (2009).
- [6] Galgon, M., Krämer, L. and Lang, B.: The FEAST Algorithm for Large Eigenvalue Problems, *PAMM*, Vol.11, pp.747-748 (2011).
- [7] Golub, G.H. and Van Loan, C.F.: §8.2.4: Orthogonal Iteration, *Matrix Computations*, 4th ed., The John Hopkins University Press (2013).
- [8] Austin, A.P. and Trefethen, L.N.: Computing Eigenvalues of Real Symmetric Matrices with Rational Filters in Real Arithmetic, *SIAM J. Sci. Comput.*, Vol.37, No.3, pp.A1365-A1387 (2015).
- [9] Güttel, S., Polizzi, E., Tang, P.T.P. and Viaud, G.: Zolotarev Quadrature Rules and Load Balancing for the FEAST Eigensolver, *SIAM J. Sci. Comput.*, Vol.37, No.4, pp.A2100-A2122 (2015).
- [10] 村田健郎, 小国 力, 唐木幸比古: §8.1: ベキ乗法一族, §8.3: レーリー・リッツ法つきの同時逆反復法, §8.5: 一般固有値問題, スーパーコンピュータ: 科学技術計算への適用, 丸善 (1985).
- [11] 村田健郎, 三好俊郎, ドンガラ, J.J., 長谷川秀彦: §11.2: ベキ乗法一族, §11.4: レーリー・リッツ法つきの同時逆反復法, §11.8: 対称行列用の一般固有値問題, 行列計算ソフトウェア: WS, スーパーコン, 並列計算機, 丸善 (1991).
- [12] 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).
- [13] 村上 弘: 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, 情報処理学会論文誌: コンピューティングシステム (ACS35), Vol.4, No.4, pp.1-14 (2011).
- [14] 村上 弘: レゾルベントを用いたフィルタによる固有値問題の解法について, 情報処理学会研究報告, Vol.2012-HPC-133, No.22, pp.1-8 (2012).
- [15] 村上 弘: 実対称定値一般固有値問題の最小側固有値を持つ固有対に対する実数シフトのレゾルベントを組み合わせたフィルタによる解法, 先進的計算基盤システムシンポジウム論文集 2012, pp.81-82 (2012).
- [16] 石上裕之, 木村欣司, 中村佳正: 再直交化付ブロック逆反復法による固有ベクトルの並列計算, HPCS2014 論文集, pp.65-75 (2013).
- [17] 村上 弘: レゾルベントの線形結合をフィルタに用いたエルミート定値一般固有値問題のフィルタ対角化法, 情報処理学会論文誌: コンピューティングシステム (ACS45), Vol.7, No.1, pp.57-72 (2014).
- [18] 村上 弘: レゾルベントの多項式をフィルタとして用いる対角化法について, 情報処理学会研究報告, Vol.2014-HPC-146, No.13, pp.1-4 (2014).
- [19] 村上 弘: 実数シフトのレゾルベントを組み合わせたフィルタによる実対称定値一般固有値問題の下端付近の固有

- 値を持つ固有対の解法, HPCS2015 シンポジウム論文集, Vol.2015, pp.38-51 (2015).
- [20] 村上 弘: 一つのレゾルベントから構成されたフィルタを用いた実対称定値一般固有値問題に対するフィルタ対角化法の実験, 情報処理学会研究報告, Vol.2015-HPC-149, No.7, pp.1-16 (2015).
- [21] 村上 弘: 固有値問題の解法に用いるレゾルベントの多項式型のフィルタの設計, 情報処理学会研究報告, Vol.2016-HPC-153, No.38, pp.1-13 (2016).
- [22] 村上 弘: 実対称定値一般固有値問題の最小側固有対を解くための実数シフトのレゾルベントの多項式によるフィルタの簡易な設計法, 情報処理学会研究報告, Vol.2016-HPC-155, No.44, pp.1-27 (2016).
- [23] 村上 弘: レゾルベントの多項式によるフィルタを用いた実対称定値一般固有値問題の解法, 情報処理学会研究報告, Vol.2016-HPC-157, No.4, pp.1-15 (2016).
- [24] 村上 弘: 実対称定値一般固有値問題を解くためのレゾルベントの多項式型フィルタの設計について, 情報処理学会研究報告, Vol.2017-HPC-158, No.7, pp.1-10 (2017).
- [25] 村上 弘: チェビシェフ展開形で表わされたレゾルベントの多項式によるフィルタの伝達特性の調整, 数理解析研究所講究録, No.2019, pp.96-112 (2017).
- [26] 村上 弘: 実対称定値一般固有値問題を解くための少数のレゾルベントの多項式を用いたフィルタの設計法, 情報処理学会研究報告, Vol.2017-HPC-159, No.4, pp.1-13 (2017).
- [27] 村上 弘: 少数のレゾルベントから構成されたフィルタを用いた実対称定値一般固有値問題の解法, 情報処理学会研究報告, Vol.2017-HPC-160, No.32, pp.1-32 (2017).
- [28] 村上 弘: 少数のレゾルベントで構成された多項式型フィルタによる対称定値一般固有値問題の解法, 情報処理学会研究報告, Vol.2017-HPC-161, No.7, pp.1-13 (2017).
- [29] 村上 弘: レゾルベントの多項式によるフィルタの伝達特性の調整, 数理解析研究所講究録, No.2054, pp.168-181 (2017).
- [30] 村上 弘: 少数のレゾルベントから構成されたフィルタを用いた対称定値一般固有値問題の解法, 情報処理学会研究報告, Vol.2017-HPC-162, No.21, pp.1-34 (2017).
- [31] 村上 弘: 少数のレゾルベントの多項式型フィルタを用いた一般固有値問題の解法, 情報処理学会研究報告, Vol.2018-HPC-165, No.15, pp.1-21 (2018).
- [32] 村上 弘: フィルタにレゾルベントの線形結合の多項式を用いた複素エルミート定値一般固有値問題の解法, 情報処理学会研究報告, Vol.2018-HPC-166, No.10, pp.1-17 (2018).
- [33] 村上 弘: 単一のレゾルベントのチェビシェフ多項式による実対称定値一般固有値問題の解法用の簡易型フィルタ, 情報処理学会論文誌: コンピューティングシステム (ACS64), Vol.12, No.2, pp.1-26 (2019).



村上 弘 (正会員)

1960年生. 1992年北海道大学大学院理学博士号(化学第二学専攻)取得. 現在, 首都大学東京数理科学専攻の准教授. 数理的な問題の数値的あるいは記号的方法による効率あるいは精度の良い解法やその並列化手法の研究に従事.

日本応用数学会, 日本数式処理学会, 日本コンピュータ学会, SIAM, AMS, ACM, IEEE 各会員.

訂正（一部の数式の修正）

この論文の出版後に、不注意により数式中に誤りを入れていたことが判明いたしました。お詫びしてここにその訂正について記述させていただきます。

- 本論文第3ページ目の副節 “3.2 レゾルベントのシフトに実数を用いる場合” 中にある数式 (8) において、

$$\sigma \leftarrow \mu / \sinh \left(\frac{1}{2n} \cosh^{-1} \frac{1}{g_s} \right)$$

とあるのは誤りで、sinh の肩の指数 2 が抜けており、正しくは以下のとおりです。

$$\sigma \leftarrow \mu / \sinh^2 \left(\frac{1}{2n} \cosh^{-1} \frac{1}{g_s} \right)$$

計算に用いたプログラムは正しい式のものになっており、この訂正は実験で行った計算結果には影響はありません。以上です。