

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

村上 弘^{1,a)}

概要：複素エルミート定値一般固有値問題をフィルタを用いて解くことにする。いま一般固有値問題に対応するレゾルベントを考えて、フィルタとして少数（2~4程度）の複素共役対をシフトとするレゾルベントの線形結合の多項式を用いることにする。この枠組みとその他の追加の仮定の下で、レゾルベントのシフトや線形結合の係数および多項式をうまく調整して、フィルタの特性をなるべく都合の良いものにする。ただし本報告では、設計の簡便さの観点から、多項式としてチェビシェフ多項式を用いる。

レゾルベントの作用はシフト行列の連立1次方程式の解法に帰着するが、その連立1次方程式を直接法を用いて解くことに決めた場合は、行列分解を一度計算してその結果を保持して用いることで計算量を節約する。シフトが互いに複素共役であるシフト行列の分解は互いに共有して利用できるため、フィルタの作用を実現するために必要な行列分解の総数は、線形結合に含まれるレゾルベントのシフトである複素共役対の数と同じになり、少数（2~4程度）になる。

キーワード：複素エルミート、固有値問題、フィルタ対角化、レゾルベント、多項式

Solving complex Hermitian-definite generalized eigenproblems by using a polynomial of a linear combination of resolvents as a filter

HIROSHI MURAKAMI^{1,a)}

Abstract: We are to solve complex Hermitian-definite generalized eigenproblems by using a filter. We consider the resolvent which corresponds to the eigenproblem, and construct the filter by a polynomial of a linear combination of those pairs of resolvents whose shifts are a few (2 to 4) complex-conjugate pairs. By this construction with some additional assumptions, to make the filter to have as good properties as possible, we tune the shifts of resolvents and the coefficients of the linear combination and also the polynomial. But in this report, to make the design of the filter simple and convenient, a Chebyshev polynomial is used for the polynomial.

The action of a resolvent is given by the solution of a system of simultaneous equations. If we decide to use some direct methods which use matrix factorizations, to reduce the amount of arithmetic computation, the system of linear equations which corresponds to the action of a resolvent is solved by using the set of matrix factors calculated once and stored. Since the matrix factors can be shared for those matrices whose shifts are a pair of complex-conjugates, the total number of matrix factorizations required to realize the operation of the filter is the same to the number of pairs of resolvents in the linear combination, which is a few (2 to 4).

Keywords: complex Hermitian, eigenproblem, filter diagonalization, resolvent, polynomial

1. はじめに（前回報告からの相違点について）

前回の HPC-165 の報告 [3] では、係数行列 A と B が実対称で B が正定値である実対称定値一般固有値問題 $A\mathbf{v} = \lambda B\mathbf{v}$ に対して、少数のレゾルベントの線形結合の多

¹ 首都大学東京・数理科学専攻
Department of Mathematical Sciences, Tokyo Metropolitan
University

^{a)} mrkmhrsh@tmu.ac.jp

項式をフィルタに用いて、固有値が指定された区間にある固有対の近似を求める方法とその実験例を記述した。今回の報告ではそのフィルタによる解法を枠組みをほとんど保ちながら適用の対象を拡げて、係数行列 A と B が複素エルミートで B が正定値である複素エルミート定値一般固有値問題に対して記述する。

フィルタが複素シフトのレゾルベントの線形結合である場合については、以前の論文 [1] に於いて、フィルタ対角化法の適用範囲を実対称定値の一般固有値問題から複素エルミート定値の一般固有値問題へ拡大したが、今回のフィルタがレゾルベントの線形結合の多項式である場合についても、問題の適用範囲の拡大には同様の方法を用いることができる。フィルタは実ベクトル上の作用素から複素ベクトル上の作用素になり、フィルタを適用して得られた十分多くのベクトルの組から近似固有対の組を構成する方法にも変更が少し加わる。

さらに一般固有値問題の係数行列 A と B が実対称から複素エルミートになることにより、複素シフト ρ のレゾルベント $\mathcal{R}(\rho)$ の作用を与える連立 1 次方程式の係数行列 $C(\lambda) = A - \lambda B$ の種類や性質も変わる。前回の一般固有値問題が実対称定値の場合の係数行列 $C(\rho)$ は、シフト ρ が虚数のときには複素対称であり実数のときには実対称であったが、今回の一般固有値問題が複素エルミート定値の場合には係数行列 $C(\rho)$ は、シフト ρ が虚数のときには複素非対称であり実数のときには複素エルミートである。

また複素シフトのレゾルベントの線形結合 $c_\infty I + \sum_j \gamma_j \mathcal{R}(\rho_j)$ を任意のベクトルに対して作用させる計算の中に現れるシフトが複素共役対であるレゾルベント 2 つの線形結合 $\gamma \mathcal{R}(\rho) + \bar{\gamma} \mathcal{R}(\bar{\rho})$ をうまく扱ってその作用の計算に必要な計算量や作業記憶量などを減らす方法は、一般固有値問題が複素エルミート定値の場合と実対称定値の場合とでは以下に見るように違いがある。

係数 A と B が実対称の場合 与えられた実ベクトル \mathbf{u} に対して、シフトが複素共役対であるレゾルベント 2 つの線形結合を作用させた結果 $\{\gamma \mathcal{R}(\rho) + \bar{\gamma} \mathcal{R}(\bar{\rho})\} \mathbf{u}$ は、 $\mathcal{R}(\bar{\rho}) = \overline{\mathcal{R}(\rho)}$ であることを用いれば、片方の項を作用させた結果の 2 倍の実部 $\text{Re}\{2\gamma \mathcal{R}(\rho) \mathbf{u}\}$ として計算できる。このことから、レゾルベントの作用を与える連立 1 次方程式を解くための行列分解の計算量、行列分解を保持するための記憶量、得られた行列分解を用いてレゾルベントの作用の結果を得るために連立 1 次方程式を解く演算量はそれぞれ、シフトが虚数のものについては簡単に半分にできる。

またシフト ρ に対応する連立 1 次方程式の係数行列 $C(\rho) = A - \rho B$ は複素対称 $\{C(\rho)\}^T = C(\rho)$ であり、特に ρ が実数ならば実対称である。よって計算効率を重視するのであれば、連立 1 次方程式を解くには複素対称行列用の改訂コレスキ法 (modified Cholesky method) を試みることになる。ただし改訂コレスキ分

解の計算は係数行列が正則であってもその数値安定性は不明である。ただし ρ が実数でしかも最小固有値の下界であれば、 $C(\rho)$ は実対称正定値なので通常の実数版の改訂コレスキ法が使えてしかも数値安定である。(虚数のシフトで複素対称版の改訂コレスキ法が数値的に安定でなかった場合には、演算量と分解を保持するための記憶が増加してしまうが、行選択 LU 分解法などの数値安定度のより高い算法の利用に切り替えることも可能である。)

係数 A と B が複素エルミートの場合 与えられた複素ベクトル \mathbf{u} に対して、シフトが複素共役対であるレゾルベント 2 つの線形結合を作用させた結果は、片方の項を作用させた結果の 2 倍の実部にはならないので計算が少し複雑になる。またシフト ρ に対応する連立 1 次方程式の係数行列 $C(\rho) = A - \rho B$ は ρ が実数の場合は複素エルミートであるが、 ρ が虚数の場合には複素非対称である (複素対称でも複素エルミートでもない)。そのためシフト ρ が虚数の場合には、レゾルベントの作用を与える連立 1 次方程式は係数行列 $C(\rho)$ に (行列の転置や複素エルミート転置に対する) 対称性がないので、行選択 LU 分解を用いて解くことにする。シフトが複素共役対である 2 つの連立 1 次方程式にはその係数行列の間に関係 $C(\bar{\rho}) = \{C(\rho)\}^H$ があるので、既に $C(\rho)$ の行選択 LU 分解を得ていれば、 $C(\bar{\rho})$ の行選択 LU 分解は全く演算をせずに直ちに得られる。そのことを用いると、シフトが複素共役対の虚数であるレゾルベント 2 つの線形結合を作用させて得られる結果 $\{\gamma \mathcal{R}(\rho) + \bar{\gamma} \mathcal{R}(\bar{\rho})\} \mathbf{u}$ を計算する際には行列分解は片方のシフト行列についてだけ行なえばよい。ただし、その後で共通の右辺 $B\mathbf{u}$ を持ち係数行列がそれぞれ $C(\rho)$ と $C(\bar{\rho})$ である 2 つの連立 1 次方程式の解それぞれは (片方の行列分解の構成結果を用いて) どちらも計算して求める必要がある。

なお、行と列の両方の選択を伴う LU 分解法 (完全選択) を採用する場合も、互いに複素共役な虚数のシフトを持つ係数行列 $C(\rho)$ と $C(\bar{\rho})$ の間で行列分解の結果を共用できる。

計算量や分解保持のための記憶量を低減するために、行や列の選択をしない LU 分解法を採用する場合には、係数行列が正則であっても分解計算が途中で破綻したり数値不安定を生じる可能性があるため、用心のために連立 1 次方程式の近似解の残差を計算して確認をすることが必要となるであろう。

以下では「複素シフトのレゾルベントの線形結合の多項式」をフィルタとする場合のその構成と実装について、前回の HPC-165 の報告 [3] で記述した実対称定値の一般固有値問題の場合に対する方法をほぼそのままながら複素エルミート定値の一般固有値問題の場合に対して拡張し

た方法を記述する。

2. フィルタ対角化法の概略

行列 A と B が複素エルミートで B が正定値である一般固有値問題 $Av = \lambda Bv$ が与えられたとする。この一般固有値問題の固有値はすべて実数になる。このとき $\lambda \in [a, b]$ を満たす固有対 (λ, \mathbf{v}) をフィルタ対角化法 [10] を用いて解くことにする。

フィルタは固有値が区間 $[a, b]$ に入る固有ベクトルはよく通過するが固有値が $[a, b]$ から離れた固有ベクトルはほとんど阻止するようにうまく構成された線形作用素とする。そのような性質を持つフィルタは、 $[a, b]$ 近傍にある固有値全部に対応する不変部分空間 $S_{[a,b]^*}$ への (近似) 射影作用素である。複素ベクトルをそのようなフィルタを用いて (近似) 射影して得られた像の複素ベクトルを十分多く集めてそれらを分析してうまく線形結合を作ることによって不変部分空間の良い (近似) 基底を構成できる。そのようにして得られた不変部分空間の (近似) 基底に対して Rayleigh-Ritz 法を適用すると必要な固有対の近似が一斉に得られる。

十分多く集めたランダムなベクトルのフィルタによる像から近似対を構成する処理は論文 [1] と同様になる。

3. レゾルベントとその実現

複素数 ρ をシフトとするレゾルベントを以下の式 (1) で定義する。

$$\mathcal{R}(\rho) \equiv (A - \rho B)^{-1} B. \quad (1)$$

ベクトル \mathbf{v} へのレゾルベントの作用は以下の式 (2) になる。

$$\mathbf{x} \leftarrow \mathcal{R}(\rho) \mathbf{v}. \quad (2)$$

式 (2) の作用は以下の連立 1 次方程式 (3) を解いて実現する。

$$C(\rho) \mathbf{x} = B \mathbf{v}. \quad (3)$$

ただし、複素数 ρ をシフトとする上記の連立 1 次方程式 (3) の係数行列 (シフト行列) $C(\rho)$ は以下の式 (4) で定義される。

$$C(\rho) \equiv A - \rho B. \quad (4)$$

シフト ρ が実数の場合は行列 $C(\rho)$ は複素エルミートであるが、シフト ρ が虚数の場合は行列 $C(\rho)$ は複素非対称である。ただしシフトが互いに複素共役である係数行列の間には以下の良い関係 (5) が成立する。

$$C(\bar{\rho}) = \{C(\rho)\}^H. \quad (5)$$

さらにもしも A と B が共に疎行列あるいは帯行列の場合には、係数 $C(\rho)$ もまた同様の疎行列あるいは帯行列になる。

4. レゾルベントを用いたフィルタ

係数 $C(\rho)$ の連立 1 次方程式を解く場合に、 Π を置換として行選択 LU 分解が以下の式 (6) で表わされたとする。

$$\Pi C(\rho) = LU. \quad (6)$$

この式 (6) の複素エルミート転置を作り、式 (5) の関係を用いると、以下の式 (7) が得られる。

$$C(\bar{\rho}) \Pi^T = U^H L^H. \quad (7)$$

この式 (7) はシフト $\bar{\rho}$ の係数行列 $C(\bar{\rho})$ の列選択 LU 分解に他ならないので、複素共役であるシフト $\bar{\rho}$ に対する行列分解の計算を省略できることがわかる。

いま複素エルミート定値一般固有値問題に対して、複素シフトのレゾルベントの線形結合である「楕円フィルタ」[1] で非常に特性の良いフィルタの実現にレゾルベントを 16~32 個用いるとすれば、その数の半分の 8~16 通りの虚数シフトに対する行列分解の構成が必要になる。行列分解の数を減らすために、「シフトが実数のレゾルベント 1 つの多項式」または「シフトが複素共役対の虚数であるレゾルベント 2 つの線形結合の多項式」をフィルタとして採用すれば [2]、必要な行列分解の数は (最少の) 1 通りにできるが、そのようにして得られるフィルタの特性はあまり良くできない。そこで本研究では「レゾルベントの線形結合の多項式」をフィルタに用いて、2~4 対程度の複素共役対をレゾルベントのシフトとして用いる。つまり必要な行列分解の数を 2~4 程度にしてフィルタ特性の改善を試みる。

5. 単一のレゾルベントの多項式によるフィルタ

まず、単一のレゾルベントの多項式によるフィルタ \mathcal{F} について述べる。(次節 6 でこの \mathcal{F} の伝達関数 $g(t)$ に有理関数 $h(t)$ を合成した関数 $g'(t) = g(h(t))$ を作り、 $g'(t)$ を伝達関数とするフィルタ \mathcal{F}' を「複数のレゾルベントの線形結合の多項式」として導いて具体的に構成する。)

いま固有値が固有値分布の下端の区間 $[a, b]$ にある固有対の近似を求めるものとする (ここで a は最小固有値のある下界とする)。レゾルベントの定義は以前の式 (1) とする。いまフィルタ \mathcal{F} が多項式 Q を用いた単一のレゾルベントの多項式として以下の式 (8) で表わされると仮定する。

$$\mathcal{F} = Q(\mathcal{R}(\rho)). \quad (8)$$

そのとき固有対 (λ, \mathbf{v}) に対して以下の式 (9) が成り立つ。

$$\mathcal{F} \mathbf{v} = f(\lambda) \mathbf{v}. \quad (9)$$

ここで $f(\lambda)$ はフィルタ \mathcal{F} の伝達関数であり、 λ の有理関数として以下の式 (10) で与えられる。

$$f(\lambda) = Q\left(\frac{1}{\lambda - \rho}\right). \quad (10)$$

逆に伝達関数が式 (10) の形であれば、それに対応するフィルタは式 (8) で与えられる。

いま区間 $\lambda \in [a, b]$ と $t \in [0, 1]$ の間の線形変換 $\lambda \equiv a + (b - a)t$ を用いて、固有値の座標 λ の正規化座標 t を

定義する. そうして引数 t を用いて表した伝達関数を $g(t)$ ($\equiv f(\lambda)$) とする.

いま, 以下の式 (11) で表わされる 2 つの実数のパラメタ $\mu > 1$ と $\sigma > 0$ を持ち, 定義域が $t \in [0, \infty)$ の 1 次分数関数 $x(t)$ を導入する.

$$x(t) \equiv \frac{\mu + \sigma}{t + \sigma}. \quad (11)$$

もしも伝達関数 $g(t)$ が多項式 P を用いて以下の式 (12)

$$g(t) = P(x(t)) \quad (12)$$

の形に書けるならば, それに対応する伝達関数 $f(\lambda)$ とフィルタ \mathcal{F} は $\gamma \equiv (\mu + \sigma)(b - a)$, $\rho \equiv a - (b - a)\sigma$ とおくと, それぞれ以下の式 (13) で表わされる.

$$\begin{cases} f(\lambda) &= P\left(\frac{\gamma}{\lambda - \rho}\right), \\ \mathcal{F} &= P(\gamma \mathcal{R}(\rho)). \end{cases} \quad (13)$$

そこで, $g(t) = P(x(t))$ を決定する多項式 P および 2 つのパラメタ μ と σ をうまく調節して, フィルタが固有値が区間 $[a, b]$ にある固有ベクトルは良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殆ど阻止する伝達特性を持つようにする.

6. 関数合成による伝達関数

6.1 関数合成前の伝達関数

いま $g(t)$ は関数合成前の伝達関数で, $t \in [0, \infty)$ で有界な実有理関数であるとする. そのとき定義域 $[0, \infty)$ を通過域 (passband) $[0, 1]$, 遷移域 (transitionband) $(1, \mu)$, 阻止域 (stopband) $[\mu, \infty)$ の 3 つの領域に分割する (図 1). 伝達率の閾値 g_p と g_s は, 通過域 $[0, 1]$ に於ける $g(t)$ の最大値は 1 で, 最小値を $g_p (> 0)$ と定義する. さらに阻止域 $[\mu, \infty)$ に於ける $|g(t)|$ の上限を g_s と定義する (図 2).

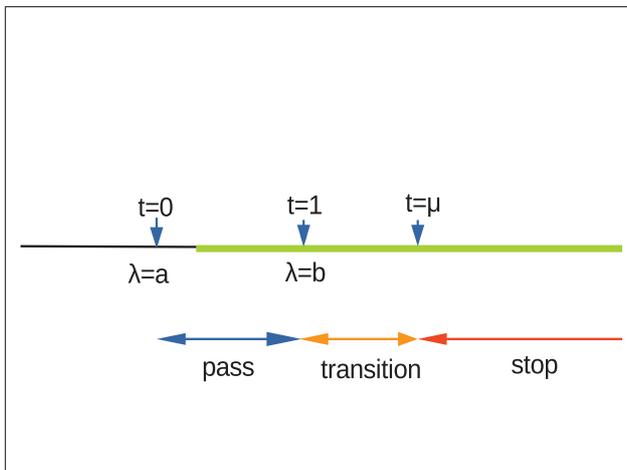


図 1 正規化座標 t と固有値の座標 λ の関係.
通過域 $t \in [0, 1]$; 遷移域 $t \in (1, \mu)$; 阻止域 $t \in [\mu, \infty)$,
緑の部分は固有値の存在が可能な場所

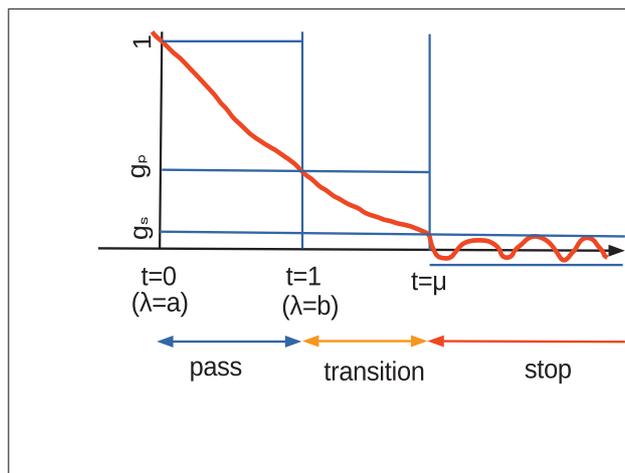


図 2 元の伝達関数 $g(t)$ の概形

6.2 合成に用いる有理関数 $h(t)$ について

いま $h(t)$ は以下の各条件を満たす有理関数とする (有理関数 $h(t)$ の値として $\pm\infty$ も許容することにする).

- 1) 区間 $[0, 1]$ を区間 $[0, 1]$ 全体に写す.
- 2) 区間 $(1, \mu')$ では単調増加で, $h(1) = 1$, $h(\mu') = \mu$.
- 3) $[\mu', \infty)$ では $|h(t) + \sigma| \geq \mu + \sigma$.

そのとき合成関数 $g'(t) \equiv g(h(t))$ もまた有理関数であり, $g'(t)$ の「通過域」, 「遷移域」, 「阻止域」をそれぞれ $[0, 1]$, $(1, \mu')$, $[\mu', \infty)$ に選ぶと, 伝達率の閾値 g_p と g_s はどちらも $g(t)$ と $g'(t)$ に対して共通の値になる.

他方で, $g(t)$ の遷移域の幅は $\mu - 1$ であるが, $g'(t)$ の遷移域の幅は $\mu' - 1$ になる. 関係 $\mu = h(\mu')$ により $\mu > \mu'$ であれば, 合成によって遷移域の幅は縮小されて遮断特性が急峻となり, フィルタの特性は向上する.

合成用の有理関数 $h(t)$ が偶関数のときには $g'(t)$ も偶関数で, 「通過域」, 「遷移域」, 「阻止域」をそれぞれ原点对称に拡げて $|t| \leq 1$, $1 < |t| < \mu'$, $\mu' \leq |t|$ にできる.

合成された有理関数 $g'(t) = g(h(t))$ を伝達関数を持つフィルタ \mathcal{F}' はレゾルベントを用いて具体的に構成できる.

6.3 典型フィルタ 4 種類を模倣した合成

電気回路理論に於ける 4 種類の典型フィルタとしてバターワース型 (Butterworth), チェビシェフ型 (Chebyshev), 逆チェビシェフ型 (inverse-Chebyshev), 楕円型 (elliptic) がある [5], [8].

それらの各構成法の模倣により合成用の有理関数 $h(t)$ を作り, 以下の各式 (14) で表わされる k 次の有理関数とする. (B-合成, C-合成の場合には $h(t)$ は実際には多項式である).

$$h(t) = \begin{cases} t^k & \text{(B-合成)} \\ \frac{1}{2}\{1 + T_k(2t - 1)\} & \text{(C-合成, } k \text{ が奇数)} \\ \frac{1}{2}\{1 + T_k(t)\} & \text{(C-合成, } k \text{ が偶数)} \\ \frac{1 + T_k(\mu')}{1 + T_k(\mu'/t)} & \text{(I-合成)} \\ \frac{1 + R_k(\mu')}{1 + R_k(\mu'/t)} & \text{(E-合成)} \end{cases} \quad (14)$$

上の式 (14) の中で T_k は k 次の第 1 種チェビシェフ多項式を, $R_k(x)$ は k 次の楕円有理関数 (チェビシェフ有理関数) を表わす。

上記の各合成 (B-合成, C-合成, I-合成) に対して関係 $h(\mu') = \mu$ を解くと, μ' を用いて μ を, μ を用いて μ' を, 表わす数式がそれぞれ得られる (それらの具体的な式は前回の HPC-159 の報告 [3] の中に記述がある)。

6.4 合成された伝達関数に対するフィルタの構成法

P を多項式として, 伝達関数が $g'(t) = P(x'(t))$ であるフィルタ \mathcal{F}' を具体的に構成する方法は以下ようになる。

- 1) 有理関数 $x'(t) = x(h(t)) = \frac{\mu + \sigma}{h(t) + \sigma}$ を部分分数分解したものを用いて以下の式 (15) とする (注: 今扱っている種類の有理関数 $h(t)$ の場合には部分分数分解には重複した極は現れない)。

$$x'(t) = c_\infty + \sum_{j=1}^k \frac{c_j}{t - t_j}. \quad (15)$$

- 2) そのとき $x'(t)$ に対応する作用素 \mathcal{X}' は, レゾルベントの線形結合として以下の式 (16) で与えられる。

$$\mathcal{X}' = c_\infty I + \sum_{j=1}^k \gamma_j \mathcal{R}(\rho_j). \quad (16)$$

- 3) すると $g'(t) = P(x'(t))$ に対応するフィルタ \mathcal{F}' は作用素 \mathcal{X}' の多項式 P として以下の式 (17) になる。

$$\mathcal{F}' = P(\mathcal{X}'). \quad (17)$$

いまここで扱っている典型 4 種類の合成のうち E-合成を除く 3 種類 (B-合成, C-合成, I-合成) について, 作用素 \mathcal{X}' をレゾルベントの線形結合で表わした式の中のシフト ρ_j や線形結合係数 γ_j の値は係数行列 A と B が実対称の場合と全く同じものになる。それら各合成の場合の具体的な導出や式は前回の HPC-165 の報告 [3] の中に既に詳しく記述したが, 注意すべき点は, 行列 A と B が実対称である場合に特化した式変形がなされている箇所があることである。たとえばレゾルベントの線形結合である \mathcal{X}' の作用あるいはそれをベクトルへ作用させた結果の式に於いて, シフトが複素共役対であるレゾルベント 2 つの線形結合の作用を片方の 2 倍の実部として計算する形の式である。つ

まり $\gamma \mathcal{R}(\rho) + \bar{\gamma} \mathcal{R}(\bar{\rho})$ を $\text{Re } 2\gamma \mathcal{R}(\rho)$ の形に書いて表しているが, それは行列 A と B が実対称である場合には成立するが, 複素エルミートである場合には成立しないので, 実部をとる演算 Re を用いて書かれている箇所については本報告の内容に沿って記述を改める必要がある。

6.5 作用素 \mathcal{X}' の構成に必要な行列分解の数の削減

実関数 $x'(t)$ の極とレゾルベントのシフトは正規化座標 t と固有値の座標 λ の関係でもって対応する。つまり $x'(t)$ の実数の極には実数のシフトが対応し, $x'(t)$ の複素共役な虚数の極の対には複素共役な虚数のシフトの対が対応する。

作用素 \mathcal{X}' の式の中で, 複素共役対のシフト ρ と $\bar{\rho}$ に対する 2 つの項 $\gamma \mathcal{R}(\rho)$ と $\bar{\gamma} \mathcal{R}(\bar{\rho})$ があるとすると, そのときレゾルベント $\mathcal{R}(\rho)$ と $\mathcal{R}(\bar{\rho})$ に対応する連立 1 次方程式の係数はそれぞれ $C(\rho)$ と $C(\bar{\rho})$ であり, 方程式を「直接法」で解く場合に 2 つの行列分解の間に式 (6) と式 (7) の関係があるので, 行列分解は片方 $C(\rho)$ についてだけ行えばよい。よって必要な行列分解の数は (B-合成, C-合成, I-合成では) シフトが虚数のものは $\lfloor k/2 \rfloor$ 個で, シフトが実数のものは k が奇数の場合に限り 1 個である (表 1)。

表 1 必要な行列分解の数

k	シフトが虚数	シフトが実数
1	0	1
2	1	0
3	1	1
4	2	0
5	2	1
6	3	0
7	3	1
8	4	0

k が奇数の場合にはシフトに実数のものが現れるので実数のシフトと固有値が一致や極端な近接をしない保証のために, フィルタは「下端固有対」用にだけ使う。それに対して k が偶数の場合には実数のシフトは無いのでシフトと固有値の近接が無いことが保証されるから, フィルタは「中間固有対」用に使える (もちろん下端用としても使える)。

6.6 簡易構成型の伝達関数に対する関数合成

有理関数 $h(t)$ の合成により, 伝達関数の閾値 g_p と g_s を変えずに伝達関数の遷移域の幅だけを関係 $\mu = h(\mu')$ に従って $\mu - 1$ から $\mu' - 1$ へ縮小できる。

いま n 次のチェビシェフ多項式 T_n を用いた「簡易構成型」の伝達関数 $g(t)$ は以下の式 (18) であるとする。

$$\begin{cases} g(t) = g_s T_n(2x(t) - 1), \\ x(t) = \frac{\mu + \sigma}{t + \sigma}. \end{cases} \quad (18)$$

その $g(t)$ に $h(t)$ を合成して得られる伝達関数 $g'(t) = g(h(t))$ は以下の式 (19) になる。

$$\begin{cases} g'(t) = g_s T_n(2x'(t) - 1), \\ x'(t) = \frac{\mu + \sigma}{h(t) + \sigma}. \end{cases} \quad (19)$$

伝達関数 $g(t)$ と $g'(t)$ で共通となる閾値 g_s と g_p は標準のパラメタの3つ組 (n, μ, σ) から以下の式 (20) を用いて計算できる。

$$\begin{cases} g_s \leftarrow 1 / \cosh \left(2n \sinh^{-1} \sqrt{\frac{\mu}{\sigma}} \right), \\ g_p \leftarrow g_s \cosh \left(2n \sinh^{-1} \sqrt{\frac{\mu - 1}{\sigma + 1}} \right). \end{cases} \quad (20)$$

合成された伝達関数 $g'(t)$ の例

合成前の簡易型の伝達関数 $g(t)$ の含むパラメタの3つ組を $(n, \mu, \sigma) = (20, 4.0, 4.0)$ に設定した場合には伝達率の閾値は $g_p = 1.2 \times 10^{-3}$, $g_s = 0.98 \times 10^{-15}$ となり、遷移域の幅は $\mu - 1 = 3.0$ となる。それに対して合成用の有理関数 $h(t)$ の種類をそれぞれ B-合成, C-合成, I-合成とし、合成用の次数 $k = 1, 2, 3, 4$ について、合成された伝達関数 $g'(t) = g(h(t))$ を $0 \leq t \leq 5$ の範囲についてグラフにプロットしたものを図 3, 図 4, 図 5 に示す。

B-合成, C-合成については, $k = 1$ のときの $g'(t)$ は合成前の元の $g(t)$ と同じものになるが, I-合成については $k = 1$ のときの $g'(t)$ は合成前の $g(t)$ とは異なる。また $k = 2$ の場合には B-合成と C-合成の $g'(t)$ は一致する。伝達率の閾値は k を変えても変わらないが、遷移域の幅 $\mu - 1$ は k を増すと縮小する傾向がグラフからも確認できる。

6.7 「合成された簡易型フィルタ」の設計と実現

合成前の簡易型のフィルタ \mathcal{F} の伝達関数 $g(t)$ は以前に掲げた式 (18) により与えられる。この $g(t)$ を直接的に決める「標準パラメタの3つ組」は (n, μ, σ) であり、またその特性を特徴付ける「形状パラメタの3つ組」は (μ, g_p, g_s) である。合成されたフィルタの伝達特性を特徴付ける「形状パラメタの3つ組」は (μ', g_p, g_s) である。いま合成に用いる有理関数 $h(t)$ を先に決めてから、「標準パラメタの3つ組」 (n, μ, σ) をうまく決めてそれから得られる「形状パラメタの3つ組」 (μ', g_p, g_s) が希望する値になるべく近くなるようにする。次数 n は離散値なので連続自由度が1つ不足しているために、連続値3つをすべて希望する値に一致させることは一般的にはできない。

「標準パラメタの3つ組」 (n, μ, σ) が決まれば伝達関数 $g'(t)$ が決まり、フィルタ \mathcal{F}' を具体的に構成できる。

6.7.1 「形状パラメタの3つ組」で指定する方法

まず先に、合成用の有理関数 $h(t)$ をその種類 (B-合成, C-合成, I-合成) と次数 k により指定する。そうして $g'(t)$ の「形状パラメタの3つ組」 (μ', g_p, g_s) になるべく希望する値に近くなるように「標準パラメタの3つ組」 (n, μ, σ) を決める。

たとえば等式条件の1つを不等式に緩和した以下の3通りの指定方式を用いる。

図 3 合成された伝達関数 $g'(t)$ の大きさ (B-合成の例)
 $n = 20, \mu = \sigma = 4.0$ ($g_p = 1.2 \times 10^{-3}, g_s = 0.98 \times 10^{-15}$)

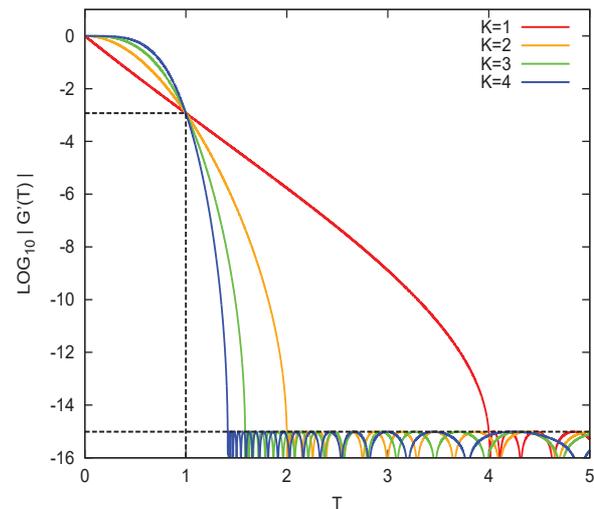


図 4 合成された伝達関数 $g'(t)$ の大きさ (C-合成の例)
 $n = 20, \mu = \sigma = 4.0$ ($g_p = 1.2 \times 10^{-3}, g_s = 0.98 \times 10^{-15}$)

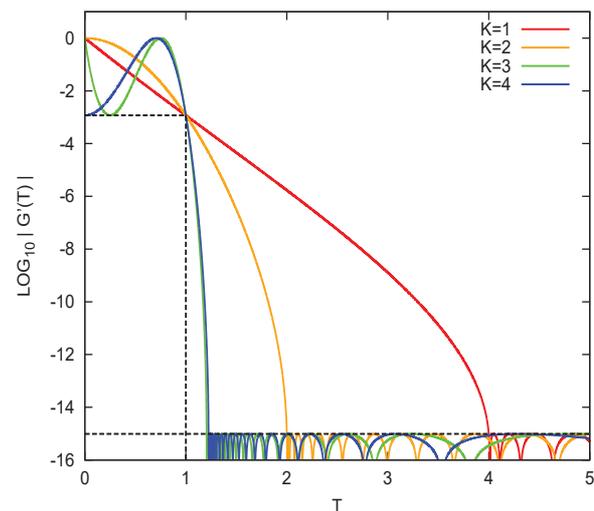
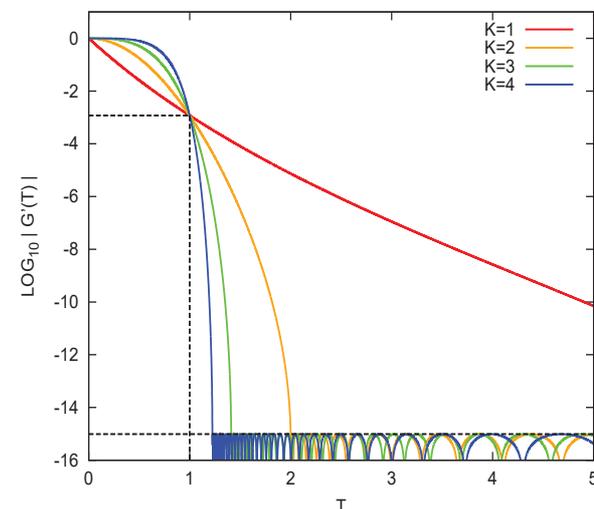


図 5 合成された伝達関数 $g'(t)$ の大きさ (I-合成の例)
 $n = 20, \mu = \sigma = 4.0$ ($g_p = 1.2 \times 10^{-3}, g_s = 0.98 \times 10^{-15}$)



- 1) μ' と g_p は値を指定し, g_s にはその上限値 G_s を指定.
- 2) μ' と g_s は値を指定し, g_p にはその下限値 G_p を指定.
- 3) g_p と g_s は値を指定し, μ' にはその上限値 M' を指定.

上記の各場合について (条件を充足する解があれば) 「標準パラメタの3つ組」 (n, μ, σ) を決定できる (その決定法の詳細は前回の HPC-165 の報告 [3] の中に記述がある).

6.8 「合成で得られた簡易型フィルタ」の作用の実現

合成で得られた簡易型のフィルタ \mathcal{F}' の伝達関数は以前に掲げた式 (19) であるとする. $x'(t)$ に対応する作用素 \mathcal{X}' はレゾルベントの線形結合として以前に掲げた式 (16) で与えられる. すると関数合成による「簡易型のフィルタ」が以下の式 (21) で与えられる. いま $\mathcal{Y}' \equiv 2\mathcal{X}' - I$ とおいて,

$$\mathcal{F}' = g_s T_n(\mathcal{Y}'). \quad (21)$$

いま任意に与えられた複素ベクトルの組 V に対して以下の式 (22) で複素ベクトルの組 $V^{(j)}$ を定義する.

$$V^{(j)} \equiv T_j(\mathcal{Y}')V. \quad (22)$$

そうしてチェビシェフ多項式の漸化式から導かれる以下の漸化式 (23) を用いて, V から始めて $V^{(n)}$ を計算する.

$$\begin{cases} V^{(0)} & \leftarrow V, \\ V^{(1)} & \leftarrow \mathcal{Y}'V, \\ V^{(j)} & \leftarrow 2\mathcal{Y}'V^{(j-1)} - V^{(j-2)} \quad (j \geq 2). \end{cases} \quad (23)$$

すると複素ベクトルの組 V へのフィルタ \mathcal{F}' の作用は以下の式 (24) で与えられる.

$$\mathcal{F}'V = g_s V^{(n)} \quad (24)$$

7. 複素共役なシフトのレゾルベントに対する行列分解の共用について

本節の内容は, 論文 [1] の中の方法を今回のフィルタの構成法 (レゾルベントの線形結合の多項式) に合わせて記述し直したものである.

一般固有値問題が複素エルミート定値の場合には実対称定値の場合と同様に, 関数 $x'(t) = x(h(t))$ が実有理関数であれば, その虚数の極は必ず複素共役対で現れる. よってそれに対応する作用素 \mathcal{X}' の式に含まれるレゾルベントのシフトで虚数のものは必ず複素共役対で現れる. そうしてシフトが互いに複素共役であるレゾルベントの線形結合係数もまた互いに複素共役である.

しかし, 一般固有値問題が実対称定値の場合に比べて複素エルミート定値の場合には, 互いに複素共役なシフトのレゾルベント間の性質の利用法は少し複雑になる. いま行列 A と B が複素エルミートであるとき, シフト行列 $C(\rho) \equiv A - \rho B$ に対しては恒等関係 $C(\bar{\rho}) = \{C(\rho)\}^H$ が成り立つ. このことを利用して計算量や使用記憶量を減らす方法を以下に示す.

7.1 シフト ρ が実数の場合

シフト ρ が実数のとき行列 $C(\rho)$ は複素エルミートである. さらに ρ が一般固有値問題の最小固有値の下界であれば $C(\rho)$ は正定値になる. 複素ベクトル \mathbf{u} へのシフト ρ のレゾルベントの作用は以下の式 (25) で表わされる.

$$\mathbf{v} \leftarrow \mathcal{R}(\rho)\mathbf{u}. \quad (25)$$

いま $\mathcal{R}(\rho) \equiv \{C(\rho)\}^{-1}B$ であるから, まずシフト ρ に依らないベクトル \mathbf{u}' を以下の式 (26) を計算して作る.

$$\mathbf{u}' \leftarrow B\mathbf{u} \quad (26)$$

そうして, それを右辺とする以下の連立1次方程式 (27) を解いてベクトル \mathbf{v} を求める.

$$C(\rho)\mathbf{v} = \mathbf{u}'. \quad (27)$$

この連立1次方程式 (27) の係数行列 $C(\rho)$ は複素エルミートなので, 複素エルミート行列用の改訂コレスキ法 (modified Cholesky method) を用いれば演算の回数や分解を保持するための記憶容量が少なく効率的である. さらに行列が帯であれば帯用に特化した算法が利用できる. (ただし改訂コレスキ分解は, 係数行列が定値であれば数値安定性が保証されるが, 係数行列が不定値の場合には数値安定性は保証されない.)

改訂コレスキ分解が $C(\rho) = LDL^H$ であれば, 以下の式 (28) により連立1次方程式の解は前進消去と後退代入を用いて計算できる.

$$\mathbf{v} \leftarrow L^{-H}D^{-1}L^{-1}\mathbf{u}'. \quad (28)$$

7.2 シフト ρ が虚数の場合

複素エルミート定値一般固有値問題の固有値はすべて実数であることから, シフト ρ が虚数の場合には行列 $C(\rho)$ と $C(\bar{\rho})$ は必ず正則になる.

シフト ρ のレゾルベントの複素ベクトル \mathbf{u} への作用の計算は以下の式 (29) になる.

$$\mathbf{v} \leftarrow \mathcal{R}(\rho)\mathbf{u}. \quad (29)$$

ここで $\mathcal{R}(\rho) \equiv \{C(\rho)\}^{-1}B$ であるから, まずシフト ρ に依らないベクトル \mathbf{u}' を以下の式 (30) を計算して作る.

$$\mathbf{u}' \leftarrow B\mathbf{u}. \quad (30)$$

そうしてそれを右辺とする以下の連立1次方程式 (31) を解いて \mathbf{v} を求める.

$$C(\rho)\mathbf{v} = \mathbf{u}'. \quad (31)$$

この連立1次方程式 (31) の係数 $C(\rho)$ には (転置やエルミート転置に対する) 対称性がないので, 行選択 LU 分解法を用いて解くことにする. 係数行列は正則なので行選択 LU 分解は必ず存在する. 係数行列が帯であれば帯用に特

化させた行選択 LU 分解の算法が利用できる。

いま Π を置換として係数行列 $C(\rho)$ の行選択 LU 分解が以下の式 (32) で与えられるとする。

$$\Pi C(\rho) = LU. \quad (32)$$

すると以下の式 (33) により置換と前進消去と後退代入により連立 1 次方程式 (31) の解が求まる。

$$\mathbf{v} \leftarrow U^{-1}L^{-1}\Pi\mathbf{u}'. \quad (33)$$

同様にシフト $\bar{\rho}$ のレゾルベントのベクトル \mathbf{u} への作用は以下の式 (34) である。

$$\mathbf{w} \leftarrow \mathcal{R}(\bar{\rho})\mathbf{u}. \quad (34)$$

それは以下の連立 1 次方程式 (35) を解いて与える。

$$C(\bar{\rho})\mathbf{w} = \mathbf{u}'. \quad (35)$$

いま以前の行選択 LU 分解の式 (32) のエルミート転置をとれば、以下の新たな関係式 (36) が得られて、既に LU 分解の形をしている (実際 U^H は下三角行列, L^H は上三角行列である)。

$$C(\bar{\rho})\Pi^T = U^HL^H \quad (36)$$

そのことから以下の式 (37) を用いて、前進消去と後退代入と逆置換により連立 1 次方程式 (35) の解が求まる。

$$\mathbf{w} \leftarrow \Pi^TL^{-H}U^{-H}\mathbf{u}'. \quad (37)$$

あるいは行列要素の方の複素共役をとるのを避けて計算するために、以下の式 (38) を用いることもできる。

$$\bar{\mathbf{w}} \leftarrow \Pi^T(L^T)^{-1}(U^T)^{-1}\bar{\mathbf{u}}'. \quad (38)$$

よってレゾルベントの作用を与える連立 1 次方程式を直接法を用いて解く場合には、たとえばシフト ρ の虚部が負の係数行列の分解は省くことにすれば、行列分解の数が減り、全体の行列分解の演算量や記憶参照量を減らせる。

そうしてシフトが複素共役対 ρ と $\bar{\rho}$ である 2 つの連立 1 次方程式の解 \mathbf{v} と \mathbf{w} の両方を (片方の) 行列分解の結果を利用して計算して求める。すると複素ベクトル \mathbf{u} に対してレゾルベントの線形結合 \mathcal{X}' を作用させる計算式の中で、シフトが複素共役対のレゾルベント 2 つの線形結合の寄与は以下の式 (39) で与えられる。

$$\{\gamma\mathcal{R}(\rho) + \bar{\gamma}\mathcal{R}(\bar{\rho})\}\mathbf{u} = \gamma\mathbf{v} + \bar{\gamma}\mathbf{w}. \quad (39)$$

注意： \mathbf{u} が実ベクトルであっても、それに \mathcal{X}' を作用させた結果の $\mathcal{X}'\mathbf{u}$ は一般には実ベクトルにはならない。

上記の説明ではレゾルベントの対を単一のベクトル \mathbf{u} に対して作用させる記述となっているが、実際の計算では複数のベクトルの組をまとめて作用させる。それにより連立 1 次方程式の求解の過程で複数の右辺ベクトルの組をまとめて作り、それに対応する複数の解をまとめて計算することで行列 B やシフト行列の分解結果に対する記憶の走査回数を減らすことができる。

8. おわりに

フィルタ対角化法により複素エルミート定値一般固有値問題 $A\mathbf{v} = \lambda B\mathbf{v}$ の固有対 (λ, \mathbf{v}) に対して指定された区間に固有値 λ があるものを近似して解く、そのためのフィルタとして「レゾルベントの線形結合の多項式」を採用する場合の解法について検討した。解法の定式化は実対称定値一般固有値問題の場合の枠組みをほぼそのまま使える。フィルタの伝達関数には実対称定値一般固有値問題に対するものと同じく実有理関数を採用して、それにより必要なレゾルベントのシフトで虚数のものは必ず複素共役対で現れることを利用すれば、レゾルベントの作用を実現するための連立 1 次方程式のシフト行列の行列分解はシフトの虚部が非負のものについてだけ行なえばよく、必要な行列分解の数を (実対称定値の場合と同様に) 少数 (2~4) に減らせる。ただしシフトが複素共役対である 2 つの連立 1 次方程式の解法では、両方の係数行列の行列分解は共用ができるが、その片方の行列分解を用いて 2 つの連立 1 次方程式のそれぞれの解を両方とも求める処理が必要である点は、片方の解だけを求めればよかった実対称定値一般固有値問題の場合とは異なる。

参考文献

- [1] 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 (2014), pp.57-72.
- [2] 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(2017年10月), pp.168-181.
- [3] 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 (2018年), pp.1-21.
- [4] Anthony P. Austin and Lloyd N. Trefethen: "Computing Eigenvalues of Real Symmetric Matrices with Rational Filters in Real Arithmetic", *SIAM J. Sci. Comput.*, Vol.37, No.3 (2015), pp.A1365-A1387.
- [5] Richard W. Daniels: *Approximation Methods for Electronic Filter Design*, McGraw-Hill, 1974.
- [6] Martin Galgon, Lukas Krämer and Brunno Lang: "The FEAST Algorithm for Large Eigenvalue Problems", *PAMM- Proc. Appl. Math. Mech.*, Vol.11 (2011), pp.747-748.
- [7] Stefan Güttel, Eric Polizzi, Ping Tak Peter Tang and Gautier Viaud: "Zolotarev Quadrature Rules and Load Balancing for the FEAST Eigensolver", *SIAM J. Sci. Comput.*, Vol.37, No.4 (2015), pp.A2100-A2122.
- [8] Miroslav D. Lutovac, Dejan V. Tošić, Brian L. Evans: *Filter Design for Signal Processing*, §12.8, Prentice Hall, 2001.
- [9] Eric Polizzi: "A Density Matrix-based Algorithm for Solving Eigenvalue Problems", *Phys. Rev. B*, Vol.79, No.1 (2009), pp.115112(6pages).
- [10] Sivan Toledo and Eran Rabani: "Very Large Electronic Structure Calculations Using an Out-of-Core Filter-Diagonalization Method", *J. Comput. Phys.*, Vol.180, No.1 (2002), pp.256-269.

付 録

A.1 実験について

A.1.1 計算機システムとプログラミング環境

CPU は Intel Core i7-5960X (8 コア, クロック 3.0GHz, L3 キャッシュ 20MiB, AVX2 拡張命令セット) で, Hyper-Thread 機能や Turbo-Boost 機能は BIOS のメニューでオフにしてある. 主記憶は DDR4-2133MHz (PC4-17000) でクアドチャネル, 16GiB のモジュールを 8 個使用して合計の主記憶容量は 128GiB である. OS は CentOS 7 for x86_64 (64bit 版) である.

プログラムは Fortran90 と OpenMP ディレクティブを入れてコーディングをし, コンパイラは Intel Fortran v15.0.0 for x86_64 で, コンパイラのオプションに "-fast-openmp" を指定して OpenMP による 8 スレッド並列で実行した.

A.1.2 計算の主要部

計算の主要部は, いま X と Y を実あるいは複素の N 次の縦ベクトル m 個を列方向に並べた $N \times m$ 行列とすると, 与えられた X に対してレゾルベント $\mathcal{R}(\rho_j)$ を適用した結果である $Y \leftarrow \mathcal{R}(\rho_j)X$ を求めるところである. 今回のフィルタは $\mathcal{F}' = g_s T_n(\mathcal{Y}')$ であり, $\mathcal{Y}' \equiv 2X' - I$ は以下の式 (A.1) により表される線形作用素 (複素ベクトル上の線形作用素) である.

$$\begin{cases} (2c_\infty - 1)I + \sum_{j=1}^{k/2} \{2\gamma_j \mathcal{R}(\rho_j) + 2\bar{\gamma}_j \mathcal{R}(\bar{\rho}_j)\} & (k \text{ が偶数}), \\ (2c_\infty - 1)I + 2\gamma_R \mathcal{R}(\rho_R) + \sum_{j=1}^{(k-1)/2} \{2\gamma_j \mathcal{R}(\rho_j) + 2\bar{\gamma}_j \mathcal{R}(\bar{\rho}_j)\} & (k \text{ が奇数}). \end{cases} \quad (\text{A.1})$$

A と B を N 次の複素エルミート行列として複素シフト ρ_j のレゾルベントは $\mathcal{R}(\rho_j) \equiv (A - \rho_j B)^{-1} B$ なので, $C^{(j)} \equiv A - \rho_j B$ とすると, $Y^{(j)} \leftarrow \mathcal{R}(\rho_j)X$ の計算は, 各 j について共通な右辺 BX を持つ連立 1 次方程式 $C^{(j)}Y^{(j)} = BX$ を $Y^{(j)}$ について解くことに帰着する. シフト ρ_j が実数であるかまたは虚数であるかに依って $C^{(j)}$ は, 複素エルミート行列あるいは複素非対称行列になるが, さらに以下の例題のように A と B が帯行列ならば $C^{(j)}$ も帯行列になる.

今回の実験に用いたプログラムでは, 帯内部が密の複素帯行列の連立 1 次方程式を, 簡単さためにブロック化などを施さずに古典的な BLAS-2 レベルの帯用の解法 (行列が複素エルミートなら改訂コレスキ法, 複素非対称なら行選択付きの LU 分解法) を用いて解いている. それを Fortran90 で記述したものに対して OpenMP の指示行を単

純に挿入したものである. 現段階の研究の目的は, 手法の数値的性質の調査でありシステムの性能を高度に追求したものにはなっていない. 各種の計算機システム上で (複素エルミートあるいは複素非対称の) 帯行列の連立 1 次方程式を高効率に解く方法はそれぞれで研究課題になりうる.

連立 1 次方程式 $C^{(j)}Y^{(j)} = BX$ を解くには, 各 j について共通する右辺ベクトルの組 BX を作るため, まず X に複素エルミート帯行列 B を乗じる処理 $Z \leftarrow BX$ を行う. そうしてシフトが実数の場合には複素エルミート帯行列用の改訂コレスキ法による行列分解 $C^{(j)} \Rightarrow L^{(j)}D^{(j)}(L^{(j)})^H$ の処理と, その分解結果を用いて前進消去と後退代入を行なって $Y^{(j)}$ を求める処理 $Y^{(j)} \leftarrow (L^{(j)})^{-H} \{(D^{(j)})^{-1}(L^{(j)})^{-1}Z\}$ が要る. シフトが虚数で虚部が正の場合には複素一般 (非対称) 帯行列用の行選択 LU 分解を行う. その分解結果からシフトがその複素共役対の場合の係数行列の列選択 LU 分解が分解計算を新たにせずに得られる. 各レゾルベントに対応する行列分解の結果は (複素共役な虚数のシフトに対するものは虚部が正のものに対するものだけを) 保持して, フィルタの作用の実現のなかでチェビシェフ多項式の次数 n と同じ回数だけ繰り返して用いることになる.

注 1

シフト ρ_j が実数で最小固有値よりも小さい場合には, $C^{(j)}$ は複素エルミートでしかも正定値となるので, 改訂コレスキ法の計算の数値安定性はピボット選択なしでも保証される. しかしそうでない場合には保証されない. シフト ρ_j が虚数のときには $C^{(j)}$ は上下の帯幅の等しい複素非対称の帯行列なので, それに合わせた帯行列用の LU 分解法を用いる. 計算の数値安定性のために行選択を入れて分解計算を行う (行選択をしなければ分解で上下の帯幅は変わらないが, 行選択を行うと下帯幅は変わらないが上帯幅が 2 倍になるため行列格納に必要な記憶量が増え, さらに分解計算の演算量や分解後に連立 1 次方程式を解く演算量も増える).

注 2

今回の実験ではまず j について順次に行列 $C^{(j)}$ の分解を作っておく. そうして各 j についての連立 1 次方程式 $C^{(j)}Y^{(j)} = Z$ を既に作っておいた行列 $C^{(j)}$ の分解結果を利用して解いて $Y^{(j)}$ を求める. 各前進消去と後退代入の処理では, m 個のベクトルの組をまとめて計算することにより $C^{(j)}$ の分解結果への記憶参照の量を節約している.

A.1.3 例題の一般固有値問題について

例題に用いた複素エルミート定値一般固有値問題 $A\mathbf{v} = \lambda B\mathbf{v}$ の係数行列 A と B は, 行列要素が以下の数式で与えられるものとした.

行列次数を N , 対角線を含まない下帯幅は双方とも w_L とする. 行列要素の添字 p と q の範囲は $1, 2, \dots, N$ とする. まず帯外部の行列要素は零とする. 対角要素はそれぞれ $a_{p,p} \equiv p - 1$, $b_{p,p} \equiv p + 1$ とし, 帯内部で対

角線より下側の行列要素は $a_{p,q} \equiv \frac{1}{p+q+1} + \frac{1}{p-q+1}\sqrt{-1}$, $b_{p,q} \equiv \frac{1}{(p-q+1)^2} + \frac{1}{(p+q+1)^2}\sqrt{-1}$ とする. 対角線より上側の行列要素については行列の複素エルミート性から対角線について対称な位置関係にある下側の行列要素の複素共役とする. これにより, 次数 N と下帯幅 w_L を与えたとき A と B が決まる. (更に B が正定値であることの確認が必要であるが, それにはたとえば B の複素エルミート行列用の改訂コレスキ分解が利用できる.)

この一般固有値問題については Sylvester の慣性律により, 実数シフト行列 $C(\rho) \equiv A - \rho B$ の複素エルミート行列用の改訂コレスキ分解の計算結果から (分解の計算が途中で破綻しなければ) 与えた実数のシフト ρ よりも大きい固有値のカウント $I(\rho)$ が得られる. それにより任意の実区間内 $(a, b]$ の固有値の個数が $I(a) - I(b)$ として得られる.

A.1.4 近似対の品質の評価方法

近似固有対 (λ, \mathbf{v}) に対する残差は以下の式 (A.2) である.

$$\mathbf{r} \equiv (A - \lambda B)\mathbf{v} \quad (\text{A.2})$$

この残差に対して, あるノルムによる相対残差 Θ の値を以下の式 (A.3) で定義する.

$$\Theta \equiv \frac{\|\mathbf{r}\|}{\|\lambda B\mathbf{v}\|} \quad (\text{A.3})$$

今回はノルム $\|\cdot\|$ として複素ベクトルの 2-ノルムを採用する. Θ の値は, 近似固有ベクトル \mathbf{v} の正規化には依らない, またこの定義は行列 A と B に共通の定数を乗じても変化しないものになっている. 幾何学的には N 次元空間のなかで 2つの複素ベクトル $A\mathbf{v}$ と $B\mathbf{v}$ の挟む角が θ であれば $\Theta = 2 \sin \frac{\theta}{2}$ で, $\theta \ll 1$ のとき θ にほぼ等しい. ただし $|\lambda|$ が微小な場合には丸め誤差が拡大されて評価としてはよくない.

複数の固有対に対してそれぞれの Θ の値をまとめて一度に計算すれば, その作業の中で行列 A と B の全体の記憶に対する走査回数はそれぞれ 1 回ずつになり, Θ の計算に対する負荷は軽減される.

A.1.5 フィルタ対角化法の概要

フィルタ対角化法では, まず最初に m 個のランダムな複素ベクトルの組 ($N \times m$ 行列) を乱数で生成し, それらを (複素エルミート) B -内積で正規直交化して m 個の複素ベクトルの組 X を作る. そうして X に区間 $[a, b]$ を通過域とするフィルタを作用させて別の m 個のベクトルの組 $Y \leftarrow \mathcal{F}X$ を作る.

これら 2つのベクトルの組 X と Y をもとにして, 用いたフィルタの伝達特性も考慮に入れて, 元の一般固有値問題のある「不変部分空間」を近似する空間の基底を Y の列の適切な線形結合として構成する [1]. その「不変部分空間」は, 実区間 $[a, b]$ のある適切な近傍に入るすべての固有値を持つ固有ベクトルの全体により張られるものである.

その「不変部分空間」の近似空間の基底に 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して採用する. 得られた近似対については (近似対に改良を加える場合はその後), 近似固有値を検査して, 区間 $[a, b]$ にないものについては棄却することができる.

A.1.6 各計算実験の結果のグラフについて

各計算実験の結果について, 3 枚組のグラフを縦に並べて示している.

- 最上段にあるグラフは, 上述の X と Y から作った m 次の (複素エルミート) 行列 $\beta = X^H B Y$ の固有値の分布を減少順に並べて, その順位を横軸にとり, β の固有値の大きさ (絶対値) の対数を縦軸にとってプロットしたものである.

阻止域に於けるフィルタの伝達率の上限値を g_s とし, 計算に用いた数値 (倍精度浮動小数点数) のマシンイプシロンを $\epsilon_{\text{MAC}} = 2.22 \times 10^{-16}$ とするとき, g_s の 10 倍の値と ϵ_{MAC} の 100 倍の値の大きい方の値を β の固有値に対する閾値として設定し, グラフの中のその位置に黒い水平線を描いている.

閾値を下回る β の固有値が多く存在していない場合には, フィルタで処理したベクトルの数 m はまだ十分ではないことを意味する.

- 上から 2 段目のグラフは, m 次の複素エルミート行列 $\alpha = Y^H B Y$ により, m 次の一般固有値問題 $\alpha \mathbf{u} = \phi \beta \mathbf{u}$ を考えて, その m 次の一般固有値問題に対して, β の固有空間を固有値の絶対値の小さい部分を切断して解いて得られた固有値 ϕ , その分布を減少順に, 横軸に順位をと, 縦軸に大きさ $|\phi|$ の対数値をとってプロットしたものである. この ϕ の値は Y で張られる空間に含まれる元の固有値問題の固有対 (λ, \mathbf{v}) のベクトルのフィルタによる伝達率 $f(\lambda)$ の近似値である. グラフに描いた黒い水平線は, ϕ の値分布の中で通過域でのフィルタの伝達率の最小値 g_p から開始して下側に向かって探索して最初に見つかった比較的広い空隙の場所として ϕ に対する閾値に採択された値を示す.
- 最下段のグラフは, フィルタ対角化で得られた近似対について, 横軸にはその固有値をと, 縦軸には式 (A.3) で表される 2-ノルムによる残差の相対値 Θ の対数をとってプロットしたものである.

A.2 例題 1：中間固有対を求めた例 1

複素エルミート定値一般固有値問題 $A\mathbf{v} = \lambda B\mathbf{v}$ に対して、「多項式として Chebyshev 多項式を用いる簡易構成型であるフィルタ \mathcal{F} 」を元にして (B-合成, C-合成, I-合成のそれぞれに対応する) 有理関数 $h(t)$ を合成してフィルタ \mathcal{F}' を構成し、その合成されたフィルタを用いてフィルタ対角化法により近似固有対を求める実験を行なった。チェビシェフ多項式の次数を n 、合成に用いる有理関数の次数を k とする。

既に述べたように実験に用いる一般固有値問題の係数行列 A と B は行列要素が数式を用いて与えられるものであり、行列の次数は $N = 10^6$ 、下帯幅は $w_L = 50$ とした。

合成されたフィルタ \mathcal{F}' の 3 つの形状パラメタに対する条件指定は 3 種類の合成について共通とし、 $\mu' = 1.5$ 、 $g_p = 10^{-1}$ 、 $g_s \leq G_s = 10^{-15}$ に設定した。

求めたい固有対の固有値の範囲を $[a, b] = [1.0000000300, 1.0000000400]$ と設定して実験を行なった。この一般固有値問題の固有対で固有値が区間 $[a, b]$ にあるものは、Sylvester の慣性律を用いた調査により 80 個であることがあらかじめ判っている。

フィルタにより濾過するランダムなベクトルの数は $m = 150$ と設定した。

注記：正規化座標 t によるフィルタの通過域と遷移域を併せた領域 $t \in [-\mu', \mu'] = [-1.5, 1.5]$ と対応する固有値の区間は今の場合は $\lambda \in [1.0000000275, 1.0000000425]$ であり、固有値がその区間に含まれる固有対の数は Sylvester の慣性律による調査により 121 である。そのためフィルタを適用するベクトルの数 m は 121 よりも多くすることが望ましい。

A.2.1 (例題 1) B-合成の場合

B-合成で $h(t)$ の次数が $k = 6$ のときに、指定された 3 つの条件 $\mu' = 1.5$ 、 $g_p = 10^{-1}$ 、 $g_s \leq G_s = 10^{-15}$ を満足する $g'(t)$ を構成し、それを伝達関数とするフィルタ \mathcal{F}' を実験に用いた。使用したフィルタのパラメタを倍精度計算用の有効精度 16 桁で示すと以下ようになる (A.4)。

$$\begin{cases} n &= 22, \\ \mu &= 1.13906\ 25000\ 00000\ \text{E} + 01, \\ \sigma &= 1.41324\ 52177\ 62856\ \text{E} + 01, \\ \mu' &= 1.50000\ 00000\ 00000\ \text{E} + 00, \\ g_p &= 1.00000\ 00000\ 00002\ \text{E} - 01, \\ g_s &= 7.51830\ 69588\ 44429\ \text{E} - 16. \end{cases} \quad (\text{A.4})$$

正規化座標で表した $x'(t)$ の部分分数分解 $x'(t) = c_\infty + \sum_{j=1}^k \frac{c_j}{t-t_j}$ に於ける定数項の係数 c_∞ と虚部が正である極 t_j と極の係数 c_j ($j=1, 2, 3$) を以下の表 A.1 に倍精度計算用の有効精度 16 桁で示す。虚部が負である複素共役な極 \bar{t}_j に対する極の係数はまた複素共役であって \bar{c}_j になる。

注記：表中で $(3.0, 4.0)$ は $3.0 + 4.0\sqrt{-1}$ を表す複素数の表記法である。

表 A.1 B-合成で $k = 6$ の $g'(t)$ の虚部が正の極と極の係数

c_∞	0.0
t_1	(1.34658 43179 40996 E+00, 7.77450 81844 97625 E-01)
c_1	(-4.05319 79212 35245 E-01, -2.34011 49109 04000 E-01)
t_2	(0.00000 00000 00000 E+00, 1.55490 16368 99525 E+00)
c_2	(0.00000 00000 00000 E+00, -4.68022 98218 08001 E-01)
t_3	(-1.34658 43179 40996 E+00, 7.77450 81844 97625 E-01)
c_3	(4.05319 79212 35245 E-01, -2.34011 49109 04000 E-01)

フィルタ対角化法では、 m 個のベクトルの組にフィルタを適用する処理を行なう部分が計算量と計算時間の主要部分である。 $k = 6$ の場合には 3 通りの虚数シフトについて N 次の複素非対称帯行列の行選択 LU 分解を行なう必要がある。虚部が正である 3 つのシフトに対する係数行列の行選択 LU 分解を行うのに掛かった経過時間は 4.39 秒であった。その後 6 個 (3 対の複素共役対) のシフトのレゾルベントの線形結合である作用素 $\mathcal{Y}' = 2\mathcal{X}' - I$ 、その次数 $n = 22$ の Chebyshev 多項式であるフィルタ $\mathcal{F}' = g_s T_n(\mathcal{Y}')$ を $m = 150$ 個のベクトルの組 X に対して作用させてその結果であるベクトルの組 Y を作るのに掛かった時間は 803.6 秒であった。

(帯行列の行列分解の演算量は、行列次数が N で下帯幅が w_L であるときには $O(N w_L^2)$ である。ベクトル 1 個あたりの連立 1 次方程式の演算量は、帯行列の行列分解を既に得た後では $O(N w_L)$ である。よって帯幅が L 倍に増えるとき行列分解に掛かる演算量は L^2 倍になるが、行列分解が得られた後にフィルタの作用を施すのに掛かる演算量は L 倍になるので、帯幅の増大に伴って、行列分解の演算量が全体の演算量の中に占める割合が大きくなっていくであろう。)

実験結果を 3 枚組の図 A.1~図 A.3 に示す。最終的に必要な固有対 80 個が求まっている。

- 1 枚目の β の固有値分布のグラフ (図 A.1) から、閾値以下の切断される固有値が多数あるので、ベクトルの数 $m = 150$ は既に十分であることがわかる。
- 2 枚目の $|\phi|$ の値を示すグラフ (図 A.2) は、伝達率の閾値による切断の様子を示す。
- 3 枚目の 2-ノルムによる相対残差 Θ のグラフ (図 A.3) からは、 Θ の値は固有値の区間の中央付近では 5×10^{-15} 程度で、区間の両端付近で 6×10^{-14} 程度であり、満足な結果であると言える。

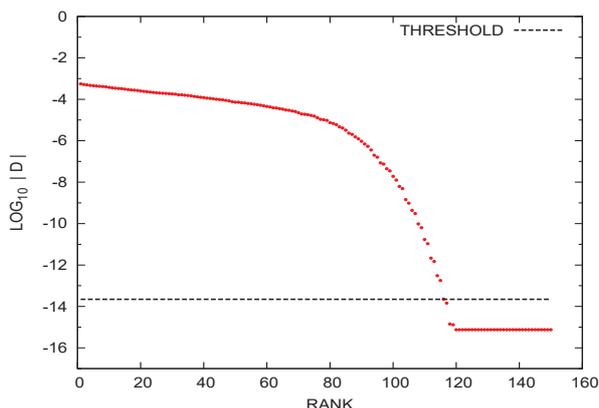


図 A.1 例題 1 (B-合成, $k=6$): β の固有値の分布

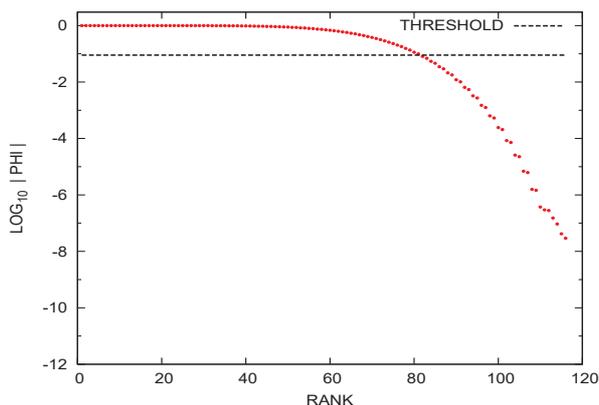


図 A.2 例題 1 (B-合成, $k=6$): $|\phi|$ の値の分布

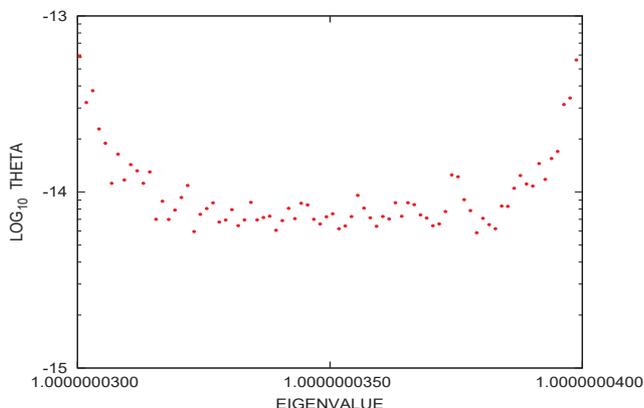


図 A.3 例題 1 (B-合成, $k=6$): 固有値に対する相対残差 Θ

A.2.2 (例題 1) C-合成の場合

C-合成で $h(t)$ の次数が $k = 6$ のときに、指定された 3 つの条件 $\mu' = 1.5$, $g_p = 10^{-1}$, $g_s \leq G_s = 10^{-15}$ を満足する $g'(t)$ を構成し、それを伝達関数とするフィルタ \mathcal{F}' を実験に用いた。使用したフィルタのパラメタを以下に示す (A.5)。

$$\begin{cases} n = 8, \\ \mu = 8.10000\ 00000\ 00000\ E + 01, \\ \sigma = 3.07357\ 82857\ 08572\ E + 00, \\ \mu' = 1.50000\ 00000\ 00000\ E + 00, \\ g_p = 1.00000\ 00000\ 00003\ E - 01, \\ g_s = 1.12934\ 05060\ 07154\ E - 16. \end{cases} \quad (\text{A.5})$$

B-合成に比べて C-合成 (および I-合成) はフィルタを実現する多項式次数 n を小さくできる利点がある。今の場合、B-合成では $n = 22$ であったものが C-合成 (および I-合成) では $n = 8$ で実現ができています。

以下の表 A.2 に、正規化座標で表した $x'(t)$ の部分分数分解 $x'(t) = c_\infty + \sum_{j=1}^k \frac{c_j}{t-t_j}$ に於ける定数項の係数 c_∞ と虚部が正である極 t_j と極の係数 c_j ($j=1, 2, 3$) を示す。

表 A.2 C-合成で $k = 6$ の $g'(t)$ の虚部が正の極と極の係数

c_∞	0.0
t_1	(9.52200 00792 86805 E-01, 2.28535 08490 62925 E-01)
c_1	(-1.56751 46866 84618 E+00, -2.17703 62563 82594 E+00)
t_2	(0.00000 00000 00000 E+00, 4.57070 16981 25850 E-01)
c_2	(0.00000 00000 00000 E+00, -4.35407 25127 65188 E+00)
t_3	(-9.52200 00792 86805 E-01, 2.28535 08490 62925 E-01)
c_3	(1.56751 46866 84618 E+00, -2.17703 62563 82594 E+00)

$k = 6$ に対応する 6 個のレゾルベントの作用を与えるために必要な 3 つの行列の行選択 LU 分解を行うのに掛かった経過時間は 4.39 秒であった。その後 6 個 (3 対の複素共役対) のシフトのレゾルベントの線形結合である作用素 $\mathcal{Y}' = 2\mathcal{X}' - I$, それの次数 $n = 8$ の Chebyshev 多項式であるフィルタ $\mathcal{F}' = g_s T_n(\mathcal{Y}')$ を $m = 150$ 個のベクトルの組 X に対して作用させて新たなベクトルの組 Y を作るのに掛かった時間は 296.3 秒であった。

実験結果を 3 枚組の図 A.4~図 A.6 に示す。最終的に必要な固有対 80 個が求まっている。

- 1 枚目の β の固有値分布のグラフ (図 A.4) から、閾値以下の切断される値が多数あるので、ベクトルの数 $m = 150$ は十分であることがわかる。
- 2 枚目の $|\phi|$ の値を示すグラフ (図 A.5) は、伝達率の閾値による切断の様子を示す。
- 3 枚目の 2-ノルムの相対残差 Θ のグラフ (図 A.6) からは、 Θ の値は固有値の区間に於いて振動しているが、 5×10^{-15} から 5×10^{-14} の範囲であり、満足な結果であると言える。

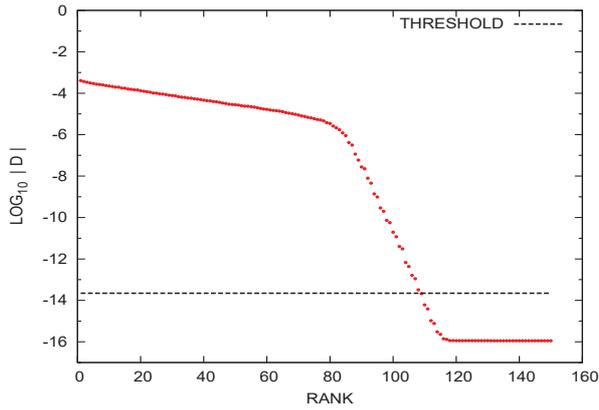


図 A.4 例題 1 (C-合成, $k=6$): β の固有値の分布

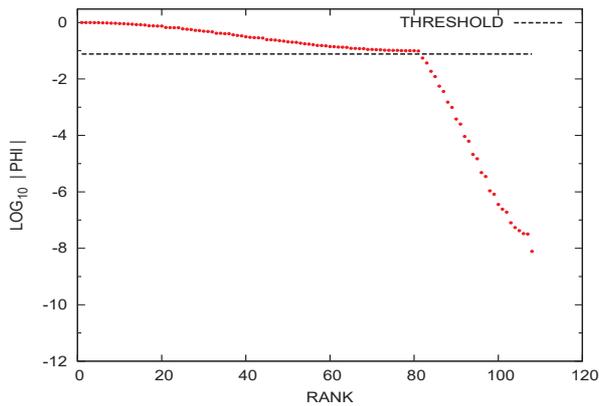


図 A.5 例題 1 (C-合成, $k=6$): $|\phi|$ の値の分布

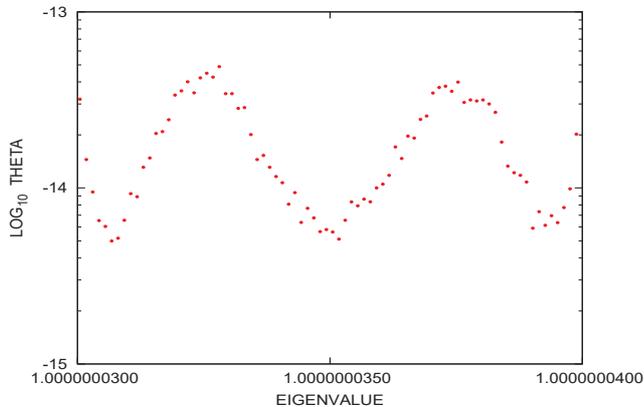


図 A.6 例題 1 (C-合成, $k=6$): 固有値に対する相対残差 Θ

A.2.3 (例題 1) I-合成の場合

I-合成で $h(t)$ の次数が $k = 6$ のときに, 指定された 3 つの条件 $\mu' = 1.5$, $g_p = 10^{-1}$, $g_s \leq G_s = 10^{-15}$ を満足するよう $g'(t)$ を構成し, それを伝達関数とするフィルタ \mathcal{F}' を実験に用いた. 使用したフィルタのパラメタを以下に示す (A.6).

$$\begin{cases} n = 8, \\ \mu = 8.10000\ 00000\ 00000\ E + 01, \\ \sigma = 3.07357\ 82857\ 08572\ E + 00, \\ \mu' = 1.50000\ 00000\ 00000\ E + 00, \\ g_p = 1.00000\ 00000\ 00003\ E - 01, \\ g_s = 1.12934\ 05060\ 07154\ E - 16. \end{cases} \quad (\text{A.6})$$

k が偶数の場合に C-合成と I-合成でこれらの値は一致する.

以下の表 A.3 に, 正規化座標で表した $x'(t)$ の部分分数分解 $x'(t) = c_\infty + \sum_{j=1}^k \frac{c_j}{t-t_j}$ に於ける定数項の係数 c_∞ と虚部が正である極 t_j と極の係数 c_j ($j=1, 2, 3$) を示す.

表 A.3 I-合成で $k = 6$ の $g'(t)$ の虚部が正の極と極の係数

c_∞	0.0
t_1	(1.14961 16304 85944 E+00, 4.33004 73342 27330 E-01)
c_1	(-4.48153 64657 25256 E+00, 1.80184 49046 43251 E-02)
t_2	(0.00000 00000 00000 E+00, 1.74259 06504 37233 E+00)
c_2	(0.00000 00000 00000 E+00, -1.19528 09637 09333 E+01)
t_3	(-1.14961 16304 85944 E+00, 4.33004 73342 27330 E-01)
c_3	(4.48153 64657 25256 E+00, 1.80184 49046 43251 E-02)

フィルタ対角化法では, m 個のベクトルの組にフィルタを適用する処理を行なう部分が計算量と計算時間の主要部分である. $k = 6$ の場合には 3 通りのシフトについて N 次の複素非対称帯行列の行選択 LU 分解を行なう必要がある. 3 つの行列の行選択 LU 分解を行うのに掛かった経過時間は 4.39 秒, その後に 6 個 (3 対の複素共役対) のシフトのレベルメントの線形結合である作用素 $\mathcal{Y}' = 2\mathcal{X}' - I$, その次数 $n = 8$ の Chebyshev 多項式であるフィルタ $\mathcal{F}' = g_s T_n(\mathcal{Y}')$ を $m = 150$ 個のベクトルの組 X に対して作用させて Y を作るのに掛かった時間は 292.4 秒であった.

実験結果を 3 枚組の図 A.7~図 A.9 に示す. 最終的に必要な固有対 80 個が求まっている.

- 1 枚目の β の固有値分布のグラフ (図 A.7) から, 閾値以下で切断される値が多数あることから, ベクトルの数 $m = 150$ は十分であることがわかる.
- 2 枚目の $|\phi|$ の値を示すグラフ (図 A.8) は, 伝達率の閾値による切断の様子を示す.
- 3 枚目の 2-ノルムの相対残差 Θ のグラフ (図 A.9) から, Θ の値は固有値の区間の中央付近では 5×10^{-15} 程度で, 区間の両端付近では 5×10^{-14} 程度であり, 満足な結果であると言える.

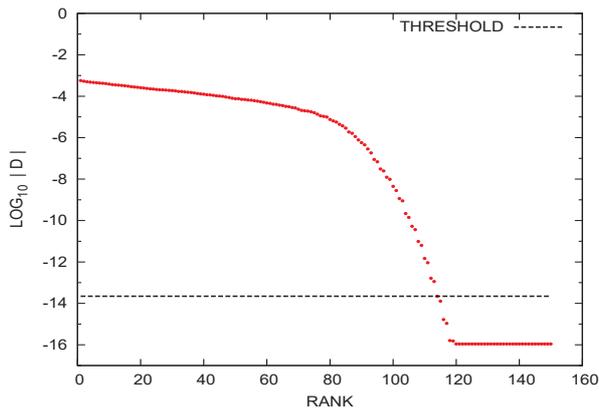


図 A.7 例題 1 (I-合成, $k=6$): β の固有値の分布

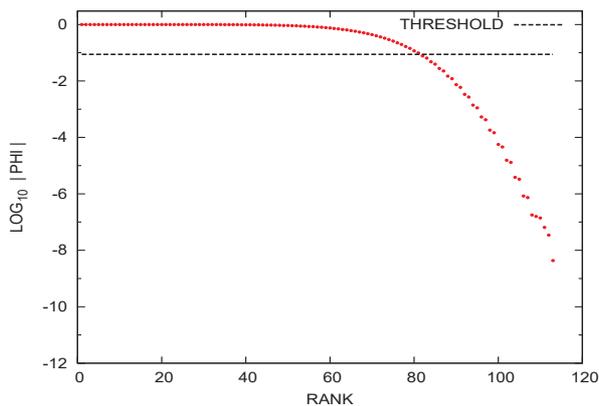


図 A.8 例題 1 (I-合成, $k=6$): $|\phi|$ の値の分布

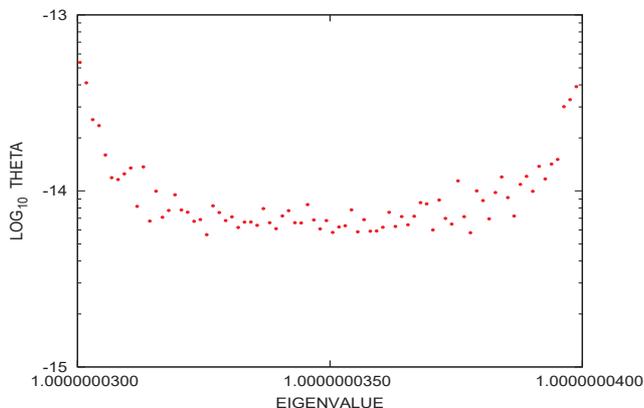


図 A.9 例題 1 (I-合成, $k=6$): 固有値に対する相対残差 Θ

A.3 例題 2 : 中間固有対を求めた例

この実験では用いた一般固有値問題の係数行列 A と B の帯内の行列要素は例題 1 と同じ数式で与えられるものとする. ただし行列次数は $N = 10^6$ で例題 1 と同じにするが, 下帯幅は $w_L = 100$ に拡大した. つまり行列としては例題 1 とは異なるものである.

合成されたフィルタ F' の 3 つの形状パラメタに対する条件指定は 3 種類の合成について共通とし, 例題 1 と同じく $\mu' = 1.5$, $g_p = 10^{-1}$, $g_s \leq G_s = 10^{-15}$ に設定した.

求めたい固有対の固有値の範囲も例題 1 と同じ $[a, b] = [1.0000000300, 1.0000000400]$ に設定した. この一般固有値問題の固有対で固有値が区間 $[a, b]$ にあるものは, Sylvester の慣性律を用いた調査から 32 個であることが判っている.

フィルタにより濾過するランダムなベクトルの数は $m = 60$ と設定した.

注記: 正規化座標 t によるフィルタの通過域と遷移域を併せた領域 $t \in [-\mu', \mu'] = [-1.5, 1.5]$ と対応する固有値の区間は今の場合は $\lambda \in [1.0000000275, 1.0000000425]$ であり, 固有値がその区間に含まれる固有対の数は Sylvester の慣性律による調査により 49 である. そのためフィルタを適用するベクトルの数 m は 49 よりも多くすることが望ましい.

A.3.1 (例題 2) B-合成の場合

B-合成で $h(t)$ の次数が $k = 6$ のときに、指定された3つの条件 $\mu' = 1.5$, $g_p = 10^{-1}$, $g_s \leq G_s = 10^{-15}$ を満足する $g'(t)$ を構成し、それを伝達関数とするフィルタ \mathcal{F}' を実験に用いた。使用したフィルタのパラメタは例題 1 の B-合成の場合のもの (A.4) と同じになる。そうして正規化座標で表した $x'(t)$ の式は例題 1 の B-合成の場合のものと同じになるので、部分分数分解に於ける定数項の係数 c_∞ と虚部が正である極 t_j と極の係数 c_j ($j=1, 2, 3$) についても、例題 1 の B-合成の場合のもの表 A.1 と同じになる。

フィルタを実現するための3通りのシフトに対する行選択 LU 分解に掛かった経過時間は 19.8 秒であった。そうして $m = 60$ 個のベクトルの組にフィルタを適用するのに掛かった経過時間は 547.0 秒であった。

実験結果を3枚組の図 A.10~図 A.12 に示す。最終的に必要な固有対 32 個が求まっている。

- 1枚目の β の固有値分布のグラフ (図 A.10) から、閾値以下の切断される固有値が多数あるので、ベクトルの数 $m = 60$ は既に十分であることがわかる。
- 2枚目の $|\phi|$ の値を示すグラフ (図 A.11) は、伝達率の閾値による切断の様子を示す。
- 3枚目の 2-ノルムによる相対残差 Θ のグラフ (図 A.12) からは、 Θ の値は固有値の区間の中央付近では 3×10^{-15} 程度で、区間の両端付近で 2×10^{-14} 程度であり、満足な結果であると言える。

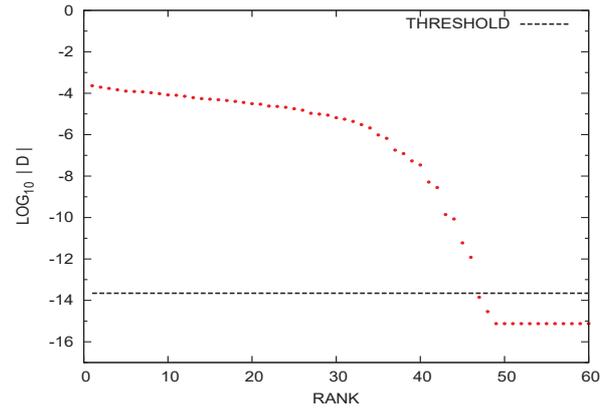


図 A.10 例題 2 (B-合成, $k=6$): β の固有値の分布

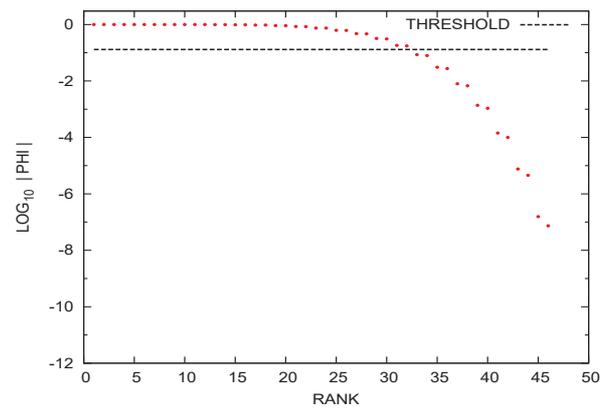


図 A.11 例題 2 (B-合成, $k=6$): $|\phi|$ の値の分布

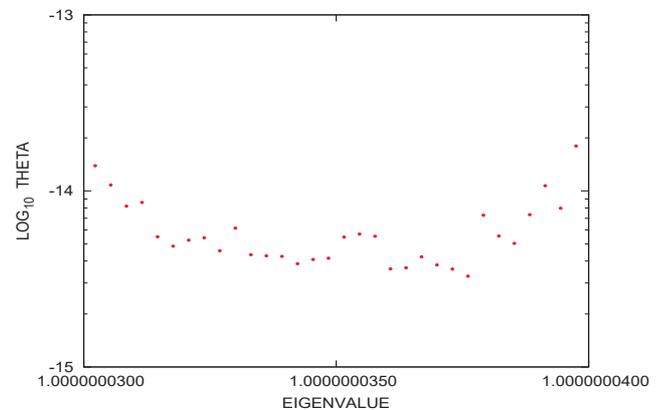


図 A.12 例題 2 (B-合成, $k=6$): 固有値に対する相対残差 Θ

A.3.2 (例題 2) C-合成の場合

C-合成で $h(t)$ の次数が $k = 6$ のときに、指定された 3 つの条件 $\mu' = 1.5$, $g_p = 10^{-1}$, $g_s \leq G_s = 10^{-15}$ を満足する $g'(t)$ を構成し、それを伝達関数とするフィルタ \mathcal{F}' を実験に用いた。使用したフィルタのパラメタは例題 1 の C-合成の場合のもの (A.5) と同じになる。

正規化座標で表した $x'(t)$ は例題 1 の C-合成の場合と同じになるので、その部分分数分解に於ける定数項の係数 c_∞ と虚部が正である極 t_j と極の係数 c_j ($j=1, 2, 3$) も以前の C-合成の場合のもの表 A.2 と同じになる。

$k = 6$ に対応する 6 個のレゾルベントの作用を与えるために必要な 3 つの行列の行選択 LU 分解を行うのに掛かった経過時間は 20.02 秒であった。その後 6 個 (3 対の複素共役対) のシフトのレゾルベントの線形結合である作用素 $\mathcal{Y}' = 2\mathcal{X}' - I$, それの次数 $n = 8$ の Chebyshev 多項式であるフィルタ $\mathcal{F}' = g_s T_n(\mathcal{Y}')$ を $m = 60$ 個のベクトルの組 X に対して作用させて新たなベクトルの組 Y を作るのに掛かった時間は 198.7 秒であった。

実験結果を 3 枚組の図 A.13~図 A.15 に示す。最終的に必要な固有対 32 個が求まっている。

- 1 枚目の β の固有値分布のグラフ (図 A.13) から、閾値以下の切断される値が多数あるので、ベクトルの数 $m = 60$ は十分であることがわかる。
- 2 枚目の $|\phi|$ の値を示すグラフ (図 A.14) は、伝達率の閾値による切断の様子を示す。
- 3 枚目の 2-ノルムの相対残差 Θ のグラフ (図 A.15) からは、 Θ の値は固有値の区間に於いて振動しているが、 5×10^{-15} から 5×10^{-14} の範囲にあり、満足な結果であると言える。

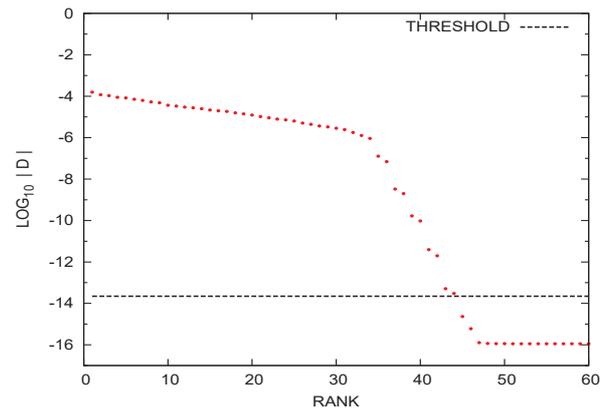


図 A.13 例題 2 (C-合成, $k=6$): β の固有値の分布

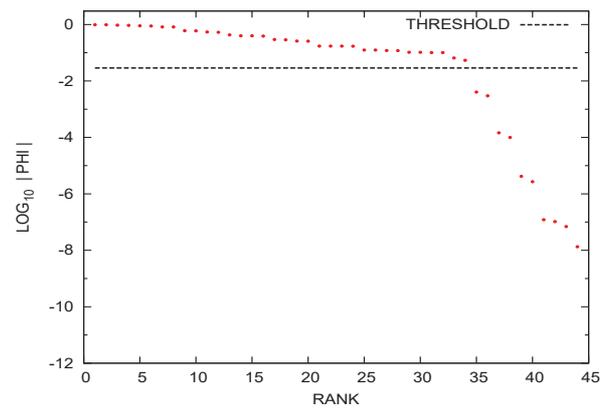


図 A.14 例題 2 (C-合成, $k=6$): $|\phi|$ の値の分布

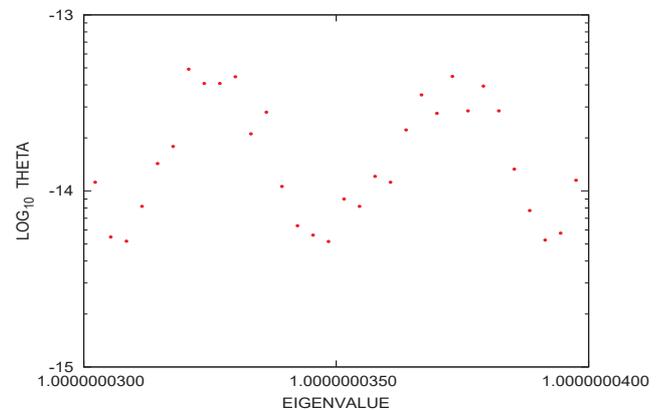


図 A.15 例題 2 (C-合成, $k=6$): 固有値に対する相対残差 Θ

A.3.3 I-合成の場合

I-合成で $h(t)$ の次数が $k = 6$ のときに、指定された3つの条件 $\mu' = 1.5$, $g_p = 10^{-1}$, $g_s \leq G_s = 10^{-15}$ を満足するよう $g'(t)$ を構成し、それを伝達関数とするフィルタ \mathcal{F}' を実験に用いた。使用したフィルタのパラメタは例題1のI-合成の場合のもの (A.6) と同じになる。

正規化座標で表した $x'(t)$ は例題1のI-合成の場合のものと同じになるので、その部分分数分解に於ける定数項の係数 c_∞ と虚部が正である極 t_j と極の係数 c_j ($j=1, 2, 3$) も例題1のI-合成の場合のもの表 A.3 と同じになる。

3つの行列の行選択 LU 分解を行うのに掛かった経過時間は 19.82 秒、その後6個 (3対の複素共役対) のシフトのレゾルベントの線形結合である作用素 $\mathcal{Y}' = 2\mathcal{X}' - I$, その次数 $n = 8$ の Chebyshev 多項式であるフィルタ $\mathcal{F}' = g_s T_n(\mathcal{Y}')$ を $m = 60$ 個のベクトルの組 X に対して作用させて Y を作るのに掛かった時間は 198.8 秒であった。

実験結果を3枚組の図 A.16~図 A.18 に示す。最終的に必要な固有対 32 個が求まっている。

- 1枚目の β の固有値分布のグラフ (図 A.16) から、閾値以下で切断される値が多数あることから、ベクトルの数 $m = 60$ は既に十分であることがわかる。
- 2枚目の $|\phi|$ の値を示すグラフ (図 A.17) は、伝達率の閾値による切断の様子を示す。
- 3枚目の2-ノルムの相対残差 Θ のグラフ (図 A.18) から、 Θ の値は固有値の区間の中央付近では 3×10^{-15} 程度で、区間の両端付近では 2×10^{-14} 程度であり、満足な結果であると言える。

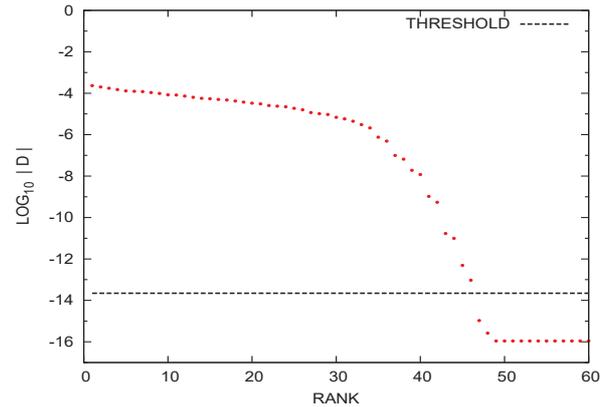


図 A.16 例題2 (I-合成, $k=6$): β の固有値の分布

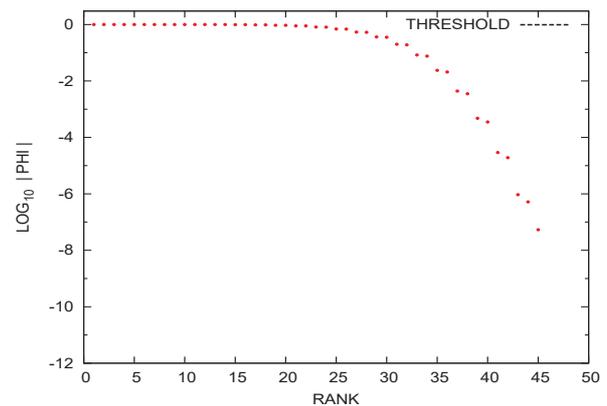


図 A.17 例題2 (I-合成, $k=6$): $|\phi|$ の値の分布

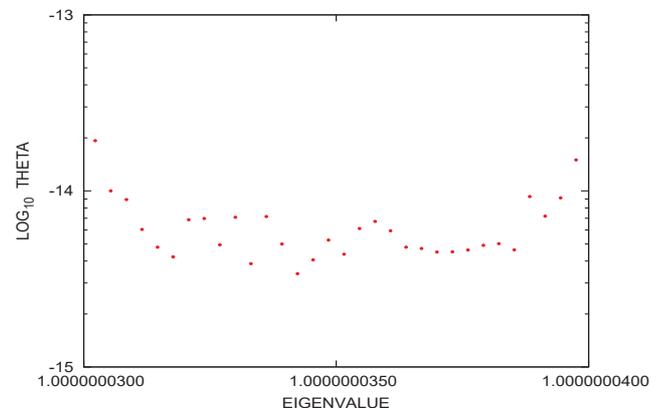


図 A.18 例題2 (I-合成, $k=6$): 固有値に対する相対残差 Θ