

各種周回積分型固有値解法の関係性について

今倉 暁^{*}, 杜 磊[†], 櫻井 鉄也^{*,‡}

^{*}筑波大学, [†]大連理工大学, [‡]JST/CREST

A map of contour integral-based eigensolvers for solving generalized eigenvalue problems

Akira Imakura^{*}, Lei Du[†], Tetsuya Sakurai^{*,‡}

^{*}University of Tsukuba, [†]Dalian University of Technology, [‡]JST/CREST

E-mail: imakura@cs.tsukuba.ac.jp

1 はじめに

本論文では, 複素平面上の指定された領域 $\Omega \subset \mathbb{C}$ 内部の固有値および対応する固有ベクトルを求める一般化固有値問題

$$A\mathbf{x}_i = \lambda_i B\mathbf{x}_i, \quad A, B \in \mathbb{C}^{n \times n}, \quad \mathbf{x}_i \in \mathbb{C}^n \setminus \{\mathbf{0}\}, \quad \lambda_i \in \Omega \subset \mathbb{C} \quad (1)$$

を解くことを考える. ここで, 領域 Ω の境界上の点 z において $zB - A$ は正則であるとする.

このような一般化固有値問題 (1) に対する有力な解法として, 2003 年に櫻井・杉浦によって周回積分型固有値解法 (SS-Hankel 法) が提案された [11]. その基本となるアイディアは, 一般化固有値問題 (1) の固有値を極に持つ有理関数 $f(z) := \mathbf{u}^H(zB - A)^{-1}B\mathbf{v}$, $\mathbf{u}, \mathbf{v} \in \mathbb{C}^n$ を考え, 領域 Ω 内部の極を Cauchy の積分公式に基づく Kravanja らの手法 [9] により求めるというものである. 解析関数 $f(z)$ の領域 Ω 内部の極を求める Kravanja らの手法 [9] により, 対象の一般化固有値問題 (1) は, Hankel 行列を係数に持つ小規模な一般化固有値問題に帰着される.

その後, 上記の SS-Hankel 法を基にした高精度・高安定な改良法として, Rayleigh-Ritz 過程に基づく改良法 (SS-RR 法) [12], ブロック化に基づく改良法 (block SS-Hankel 法, block SS-RR 法) [4, 5] および Arnoldi 法に基づく拡張法 (block SS-Arnoldi 法) [6] などが提案されている. また, 異なるタイプの周回積分型固有値解法として, Rayleigh-Ritz 過程を用いた加速部分空間反復法に基づく FEAST 法が Polizzi によってエルミート一般化固有値問題向けに提案され, その後拡張・改良が行われている [10, 13].

一方, 周回積分型固有値解法は非線形固有値問題に対しても研究が行われており, 非線形固有値問題向け block SS-Hankel 法 [1, 2], および block SS-RR 法 [14] が提案された. また, 2012 年に Beyn によって異なるタイプの周回積分型固有値解法 (以下 Beyn 法と呼ぶ) が提案された [3].

本論文では, 代表的な周回積分型固有値解法である, block SS-Hankel 法 [4], block SS-RR 法 [5], block SS-Arnoldi 法 [6], FEAST 法 [10] および非線形固有値解法である Beyn 法 [3] を一般化固有値問題 (1) に適用した際のアプローチに着目し, 各解法の数理的な

関係性について解析を行う。本論文では簡単のため行列束 A, B が対角化可能であると仮定する。なお、行列束 A, B が対角化できない場合についても同様に議論可能であり、その場合の議論の詳細については別に記載する予定である。

本論文では以下の表記を用いる。行列 $V = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m] \in \mathbb{C}^{n \times m}$ に対し、 $\mathcal{R}(V) := \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m\}$ を行列 V の Range 空間とする。また、 $A \in \mathbb{C}^{n \times n}$ に対し、 $\mathcal{K}_k^\square(A, V) := \mathcal{R}([V, AV, \dots, A^{k-1}V])$ を block Krylov 部分空間とする。また、本論文では MATLAB 表記を用いる。

2 各種周回積分型固有値解法の概要

第1節で記したように、本論文では簡単のため行列束 A, B が対角化可能であると仮定する。すなわち、次式を満たすような正則行列 $Y^{-1} = [\tilde{\mathbf{y}}_1, \tilde{\mathbf{y}}_2, \dots, \tilde{\mathbf{y}}_n]^H$ および $X = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n]$ が存在する。

$$Y^{-1}(zB - A)X = z \begin{bmatrix} I_r & \\ & O_{n-r} \end{bmatrix} - \begin{bmatrix} \Lambda_r & \\ & I_{n-r} \end{bmatrix},$$

ここで、 $\Lambda_r = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_r)$ である。一般化固有値問題 $A\mathbf{x}_i = \lambda_i B\mathbf{x}_i$ は r 個の有限の固有値 $\lambda_1, \lambda_2, \dots, \lambda_r$ および $n - r$ 個の無限固有値を持ち、対応する左/右固有ベクトルは $\tilde{\mathbf{y}}_i, \mathbf{x}_i$ である。また、 $\tilde{X}^H = [\tilde{\mathbf{x}}_1, \tilde{\mathbf{x}}_2, \dots, \tilde{\mathbf{x}}_n]^H = X^{-1}$ とし、 $X_r := [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_r]$, $\tilde{X}_r := [\tilde{\mathbf{x}}_1, \tilde{\mathbf{x}}_2, \dots, \tilde{\mathbf{x}}_r]$, $X_\Omega := [\mathbf{x}_i | \lambda_i \in \Omega]$, $\tilde{X}_\Omega := [\tilde{\mathbf{x}}_i | \lambda_i \in \Omega]$, $\Lambda_\Omega := \text{diag}(\lambda_i | \lambda_i \in \Omega)$ と置く。

$L, M \in \mathbb{N}$ を入力パラメータ、 $V \in \mathbb{C}^{n \times L}$ を入力行列とする。また、行列 $S \in \mathbb{C}^{n \times LM}$, $S_k \in \mathbb{C}^{n \times L}$ を以下のように周回積分により定義する。

$$S := [S_0, S_1, \dots, S_{M-1}], \quad S_k := \frac{1}{2\pi i} \oint_{\Gamma} z^k (zB - A)^{-1} B V dz. \quad (2)$$

この時、行列 S および S_k に対し、以下の定理が成り立つ。詳細は [7] などを参照されたい。

定理 1 行列 S_k は以下のように展開される。

$$S_k = X_\Omega \Lambda_\Omega^k \tilde{X}_\Omega^H V = C_\Omega^k S_0, \quad C_\Omega := X_\Omega \Lambda_\Omega \tilde{X}_\Omega^H.$$

定理 2 一般化固有値問題 (1) の固有値数 m が $m \leq LM$ を満たすとし、入力行列 V が $V^H X_\Omega$ がフルランクとなる行列であるとする。この時、以下が成り立つ。

$$\mathcal{R}(S) = \mathcal{R}(X_\Omega).$$

定理 2 は部分空間 $\mathcal{R}(S)$ が対象の固有ベクトル $\mathbf{x}_i, \lambda_i \in \Omega$ のみを含むことを意味する。各種周回積分型固有値解法は上記の定理 1, 2 に基づき解法が設計され、周回積分 (2) を N 点台形則等の数値積分により

$$\hat{S} := [\hat{S}_0, \hat{S}_1, \dots, \hat{S}_{M-1}], \quad \hat{S}_k := \sum_{j=1}^N \omega_j z_j^k (z_j B - A)^{-1} B V \quad (3)$$

のように近似することでアルゴリズムが導出される. ここで, z_j は積分点, ω_j は対応する重みであり,

$$\sum_{j=1}^N \omega_j z_j^k \begin{cases} = 0, & k = 0, 1, \dots, N-2 \\ \neq 0, & k = -1 \end{cases}$$

を満たす.

以下では, 各種周回積分型固有値解法のアルゴリズムの概略を示す.

2.1 block SS-Hankel 法

block SS-Hankel 法 [4] は SS-Hankel 法の拡張法として提案された手法である. 複素モーメント $\mu_k^\square := V^H S_k \in \mathbb{C}^{L \times L}$ を要素に持つ block Hankel 行列 $H_M, H_M^< \in \mathbb{C}^{LM \times LM}$ を

$$H_M^\square := \begin{pmatrix} \mu_0^\square & \mu_1^\square & \cdots & \mu_{M-1}^\square \\ \mu_1^\square & \mu_2^\square & \cdots & \mu_M^\square \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{M-1}^\square & \mu_M^\square & \cdots & \mu_{2M-2}^\square \end{pmatrix}, \quad H_M^{\square<} := \begin{pmatrix} \mu_1^\square & \mu_2^\square & \cdots & \mu_M^\square \\ \mu_2^\square & \mu_3^\square & \cdots & \mu_{M+1}^\square \\ \vdots & \vdots & \ddots & \vdots \\ \mu_M^\square & \mu_{M+1}^\square & \cdots & \mu_{2M-1}^\square \end{pmatrix},$$

と定義する. この時, 定理 1 に基づき, 一般化固有値問題 (1) の固有対 $(\lambda_i, \mathbf{x}_i), \lambda_i \in \Omega$ を block Hankel 行列 $H_M, H_M^<$ を係数とする一般化固有値問題

$$H_M^{\square<} \mathbf{t}_i = \theta_i H_M^\square \mathbf{t}_i \quad (4)$$

の固有対を用いて $(\lambda_i, \mathbf{x}_i) = (\theta_i, S \mathbf{t}_i)$ のように計算できる [4].

実用上は, 数値積分 (3) により計算される (近似) 複素モーメント $\hat{\mu}_k^\square := V^H \hat{S}_k \in \mathbb{C}^{L \times L}$ により定義される (近似) block Hankel 行列 $\hat{H}_M^\square, \hat{H}_M^{\square<} \in \mathbb{C}^{LM \times LM}$ が用いられる. また, $\hat{H}_M^\square$ を, 特異値分解により数値ランク $\hat{m}$ を用いて低ランク近似

$$\hat{H}_M^\square = [U_1, U_2] \begin{bmatrix} \Sigma_1 & O \\ O & \Sigma_2 \end{bmatrix} \begin{bmatrix} W_1^H \\ W_2^H \end{bmatrix} \approx U_1 \Sigma_1 W_1^H$$

を行うことで, 計算コストの削減および数値安定性の改善を図る. この時, 対象の固有値問題 (1) は, $\hat{m}$ 次の標準固有値問題

$$U_1^H \hat{H}_M^{\square<} W_1 \Sigma_1^{-1} \mathbf{t}_i = \theta_i \mathbf{t}_i$$

に帰着される. block SS-Hankel 法のアルゴリズムを Algorithm 1 に示す.

2.2 block SS-RR 法

定理 2 は, 対象の固有対 $(\lambda_i, \mathbf{x}_i), \lambda_i \in \Omega$ が部分空間 $\mathcal{R}(S)$ による Rayleigh-Ritz 過程に基づき

$$Q^H A Q \mathbf{t}_i = \theta_i Q^H B Q \mathbf{t}_i$$

Algorithm 1 The block SS-Hankel method**Input:** $L, M, N \in \mathbb{N}, V \in \mathbb{C}^{n \times L}, (z_j, \omega_j)$ for $j = 1, 2, \dots, N$ **Output:** Approximate eigenpairs $(\tilde{\lambda}_i, \tilde{\mathbf{x}}_i)$ for $i = 1, 2, \dots, \hat{m}$

- 1: Compute $\hat{S}_k = \sum_{j=1}^N \omega_j z_j^k (z_j B - A)^{-1} B V$, and set $\hat{S} = [\hat{S}_0, \hat{S}_1, \dots, \hat{S}_{M-1}]$
- 2: Set block Hankel matrices $\hat{H}_M^\square, \hat{H}_M^{\square <}$
- 3: Compute SVD of $\hat{H}_M^\square$: $\hat{H}_M^\square = [U_1, U_2][\Sigma_1, O; O, \Sigma_2][W_1, W_2]^H$
- 4: Compute eigenpairs $(\theta_i, \mathbf{t}_i)$ of $U_1^H \hat{H}_M^{\square <} W_1 \Sigma_1^{-1} \mathbf{t}_i = \theta_i \mathbf{t}_i$,
and compute $(\tilde{\lambda}_i, \tilde{\mathbf{x}}_i) = (\theta_i, \hat{S} U_1 \mathbf{t}_i)$ for $i = 1, 2, \dots, \hat{m}$

Algorithm 2 The block SS-RR method**Input:** $L, M, N \in \mathbb{N}, V \in \mathbb{C}^{n \times L}, (z_j, \omega_j)$ for $j = 1, 2, \dots, N$ **Output:** Approximate eigenpairs $(\tilde{\lambda}_i, \tilde{\mathbf{x}}_i)$ for $i = 1, 2, \dots, \hat{m}$

- 1: Compute $\hat{S}_k = \sum_{j=1}^N \omega_j z_j^k (z_j B - A)^{-1} B V$, and set $\hat{S} = [\hat{S}_0, \hat{S}_1, \dots, \hat{S}_{M-1}]$
- 2: Compute SVD of $\hat{S}$: $\hat{S} = [U_1, U_2][\Sigma_1, O; O, \Sigma_2][W_1, W_2]^H$
- 3: Compute eigenpairs $(\theta_i, \mathbf{t}_i)$ of $U_1^H \hat{A} U_1 \mathbf{t}_i = \theta_i U_1^H B U_1 \mathbf{t}_i$,
and compute $(\tilde{\lambda}_i, \tilde{\mathbf{x}}_i) = (\theta_i, U_1 \mathbf{t}_i)$ for $i = 1, 2, \dots, \hat{m}$

を解くことにより計算できる事を意味している。ここで、 $Q = \text{orth}(S)$ である。

block SS-RR 法 [5] は上記のアイディアに基づき提案された解法であり、部分空間として数値積分により近似された $\mathcal{R}(\hat{S}) \approx \mathcal{R}(S)$ を用いる。また、実用上の観点から $\hat{S}$ の特異値分解に基づく低ランク近似を

$$\hat{S} = [U_1, U_2] \begin{bmatrix} \Sigma_1 & O \\ O & \Sigma_2 \end{bmatrix} \begin{bmatrix} W_1^H \\ W_2^H \end{bmatrix} \approx U_1 \Sigma_1 W_1^H$$

と置くと、帰着される小規模固有値問題は

$$U_1^H \hat{A} U_1 \mathbf{t}_i = \theta_i U_1^H B U_1 \mathbf{t}_i$$

と書くことができる。近似固有対は $(\tilde{\lambda}_i, \tilde{\mathbf{x}}_i) = (\theta_i, U_1 \mathbf{t}_i)$ として得られる。block SS-RR 法のアルゴリズムを Algorithm 2 に示す。

2.3 FEAST 法

FEAST 法 [10] はエルミート一般化固有値問題に対して提案された周回積分型固有値解法であり、そのアルゴリズムは Rayleigh-Ritz 過程を用いた加速部分空間反復法に基づく。エルミート一般化固有値問題 $A \mathbf{x}_i = \lambda_i B \mathbf{x}_i$ に対する Rayleigh-Ritz 過程を用いた加速部分空間反復法では、部分空間を加速オペレータ $\rho(B^{-1}A)$ に基づき生成する。加速オペレータを $\rho(B^{-1}A) = B^{-1}A$ と設定すると、加速部分空間反復法は Rayleigh-Ritz 過程を用いた通常の部分空間反復法となり、入力ベクトル本数 L に応じて絶対値の大きい方から L 個の固有値および対応する固有ベクトルが得られる。

Algorithm 3 The FEAST method**Input:** $L, N \in \mathbb{N}, V_0 \in \mathbb{C}^{n \times L}, (z_j, \omega_j)$ for $j = 1, 2, \dots, N$ **Output:** Approximate eigenpairs $(\tilde{\lambda}_i, \tilde{\mathbf{x}}_i)$ for $i = 1, 2, \dots, L$

- 1: **for** $k = 1, 2, \dots$, until convergence **do**:
- 2: Compute $\hat{S}_0^{(k)} = \sum_{j=1}^N \omega_j (z_j B - A)^{-1} B V_{k-1}$
- 3: Compute eigenpairs $(\theta_i^{(k)}, \mathbf{t}_i^{(k)})$ of $\hat{S}_0^{(k)H} A \hat{S}_0^{(k)} \mathbf{t}_i = \theta_i \hat{S}_0^{(k)H} B \hat{S}_0^{(k)} \mathbf{t}_i$,
and compute $(\tilde{\lambda}_i^{(k)}, \tilde{\mathbf{x}}_i^{(k)}) = (\theta_i^{(k)}, \hat{S}_0^{(k)} \mathbf{t}_i^{(k)})$ for $i = 1, 2, \dots, L$
- 4: Set $V_k = \hat{S}_0^{(k)} [\tilde{\mathbf{x}}_1^{(k)}, \tilde{\mathbf{x}}_2^{(k)}, \dots, \tilde{\mathbf{x}}_L^{(k)}]$
- 5: **end for**

FEAST 法は定理 2 の性質に基づき、加速オペレータを

$$\rho(B^{-1}A) = \sum_{j=1}^N \omega_j (z_j B - A)^{-1} B \approx \frac{1}{2\pi i} \oint_{\Gamma} (zB - A)^{-1} B dz$$

と設定することで、領域 Ω 内部の固有値および対応する固有ベクトルを計算する。また、FEAST 法では周回積分の近似として、Gauss-Legendre 積分が用いられる。FEAST 法のアルゴリズムは Algorithm 3 のように書かれる。

2.4 block SS–Arnoldi 法

定理 1, 2 より、 $C_\Omega := X_\Omega \Lambda_\Omega \tilde{X}_\Omega$ に着目すると、以下の 2 つの定理が導出される [6]。

定理 3 一般化固有値問題 (1) の固有値数 m が $m \leq LM$ を満たすとし、入力行列 V が $V^H X_\Omega$ がフルランクとなる行列であるとする。この時、以下が成り立つ。

$$\mathcal{R}(S) = \mathcal{K}_M^\square(C_\Omega, S_0).$$

定理 4 一般化固有値問題 (1) の固有値数 m が $m \leq LM$ を満たすとし、入力行列 V が $V^H X_\Omega$ がフルランクとなる行列であるとする。この時、一般化固有値問題 (1) は行列 C_Ω の固有値問題

$$C_\Omega \mathbf{x}_i = \lambda_i \mathbf{x}_i, \quad \mathbf{x}_i \in \mathcal{R}(S) = \mathcal{K}_M^\square(C_\Omega, S_0) \quad (5)$$

と等しい右固有対を持つ。

定理 3, 4 から、一般化固有値問題 (1) の固有対 $(\lambda_i, \mathbf{x}_i)$, $\lambda_i \in \Omega$ は行列 C_Ω の固有値問題 (5) に対し、block Krylov 部分空間 $\mathcal{K}_M^\square(C_\Omega, S_0)$ を用いた block Arnoldi 法に基づき計算できる事がわかる。この時、行列 C_Ω は陽的には生成されず、block Arnoldi 過程の各反復で必要となる行列 C_Ω に対する行列積計算は、周回積分に基づき計算される。周回積分を数値積分で近似することで block SS–Arnoldi 法のアルゴリズムが導出される (Algorithm 4)。

なお、文献 [6] で提案された block SS–Arnoldi 法は、他の解法で用いられている行列の低ランク近似に基づく計算コストの削減および数値安定性の改善は行われていない。block SS–Arnoldi 法の低ランク近似を含む改良法については、現在研究が進められている [8]。

Algorithm 4 The block SS-Arnoldi method

Input: $L, M, N \in \mathbb{N}, V \in \mathbb{C}^{n \times L}, (z_j, \omega_j)$ for $j = 1, 2, \dots, N$

Output: Approximate eigenpairs $(\tilde{\lambda}_i, \tilde{\mathbf{x}}_i)$ for $i = 1, 2, \dots, LM$

- 1: Solve $Y_j = (z_j B - A)^{-1} B V$ for $j = 1, 2, \dots, N$
- 2: $W_0 = \sum_{j=1}^N \omega_j Y_j$
- 3: Compute QR decomposition of W_0 : $W_0 = W_1 R$
- 4: Set $\Psi_{1,j} = R^{-1}$ for $j = 1, 2, \dots, N$
- 5: **for** $k = 1, 2, \dots, M$ **do**:
- 6: $\tilde{\Psi}_{k,j} = z_j \Psi_{k,j}$ for $j = 1, 2, \dots, N$
- 7: $\tilde{W}_k = \sum_{j=1}^N \omega_j Y_j \tilde{\Psi}_{k,j}$
- 8: **for** $i = 1, 2, \dots, k$ **do**:
- 9: $H_{i,k} = W_i^H \tilde{W}_k$
- 10: $\tilde{\Psi}_{k,j} = \tilde{\Psi}_{k,j} - \Psi_{i,j} H_{i,k}$ for $j = 1, 2, \dots, N$
- 11: $\tilde{W}_k = \tilde{W}_k - W_i H_{i,k}$
- 12: **end for**
- 13: Compute QR decomposition of $\tilde{W}_k$: $\tilde{W}_k = W_{k+1} H_{k+1,k}$
- 14: $\Psi_{k+1,j} = \tilde{\Psi}_{k,j} H_{k+1,k}^{-1}$ for $j = 1, 2, \dots, N$
- 15: **end for**
- 16: Set $W = [W_1, W_2, \dots, W_M]$ and $H_M = \{H_{i,j}\}_{1 \leq i,j \leq M}$
- 17: Compute eigenpairs $(\theta_i, \mathbf{z}_i)$ of $H_M \mathbf{t}_i = \theta_i \mathbf{t}_i$,
and $(\tilde{\lambda}_i, \tilde{\mathbf{x}}_i) = (\theta_i, W \mathbf{t}_i)$ for $i = 1, 2, \dots, LM$

2.5 Beyn 法

Byen 法 [3] は周回積分型の非線形固有解法である。本小節では、Byen 法を一般化固有値問題 (1) に適用した場合のアルゴリズムの概略を示す。

行列 S_0 の特異値分解を $S_0 = U \Sigma W^H$ とする。この時、定理 1 に基づき、一般化固有値問題 (1) の固有対 $(\lambda_i, \mathbf{x}_i)$, $\lambda_i \in \Omega$ を標準固有値問題

$$U^H S_1 W \Sigma^{-1} \mathbf{t}_i = \theta_i \mathbf{t}_i$$

の固有対を用いて $(\lambda_i, \mathbf{x}_i) = (\theta_i, U \mathbf{t}_i)$ のように計算できる [3]。実用上は、数値積分により得られる行列 $\hat{S}_0$ を特異値分解し、数値ランク $\hat{m}$ を用い低ランク近似を行う。

$$\hat{S}_0 = [U_1, U_2] \begin{bmatrix} \Sigma_1 & O \\ O & \Sigma_2 \end{bmatrix} \begin{bmatrix} W_1^H \\ W_2^H \end{bmatrix} \approx U_1 \Sigma_1 W_1^H. \quad (6)$$

この時、対象の固有値問題 (1) は、標準固有値問題

$$U_1^H \hat{S}_1 W_1 \Sigma_1^{-1} \mathbf{t}_i = \theta_i \mathbf{t}_i \quad (7)$$

に帰着される。Beyn 法を一般化固有値問題に適用した際のアルゴリズムを Algorithm 5 に示す。

Algorithm 5 The Beyn method**Input:** $L, N \in \mathbb{N}, V \in \mathbb{C}^{n \times L}, (z_j, \omega_j)$ for $j = 1, 2, \dots, N$ **Output:** Approximate eigenpairs $(\tilde{\lambda}_i, \tilde{\mathbf{x}}_i)$ for $i = 1, 2, \dots, \hat{m}$

- 1: Compute $\hat{S}_0, \hat{S}_1$, where $\hat{S}_k = \sum_{j=1}^N \omega_j z_j^k (z_j B - A)^{-1} B V$
- 2: Compute SVD of $\hat{S}_0$: $\hat{S}_0 = [U_1, U_2][\Sigma_1, O; O, \Sigma_2][W_1, W_2]^H$
- 3: Compute eigenpairs $(\theta_i, \mathbf{t}_i)$ of $U_1^H \hat{S}_1 W_1 \Sigma_1^{-1} \mathbf{t}_i = \theta_i \mathbf{t}_i$,
and compute $(\tilde{\lambda}_i, \tilde{\mathbf{x}}_i) = (\theta_i, U_1 \mathbf{t}_i)$ for $i = 1, 2, \dots, \hat{m}$

3 各種周回積分型固有値解法の関係性

第2節で示されたように、各種周回積分型固有値解法は、行列 S および S_k の性質に基づき解法が設計され、周回積分を数値積分として近似することで具体的なアルゴリズムが導出されている。本節では、近似行列 $\hat{S}$ および $\hat{S}_k$ のもつ性質に着目し、数値積分に基づくアルゴリズム同士の関係性について解析する。

3.1 block SS-RR 法, FEAST 法

数値積分の近似度を示すフィルター関数 $f(\lambda_i)$ を

$$f(\lambda_i) := \sum_{j=1}^N \frac{\omega_j}{z_j - \lambda_i}$$

と定義する。この時、以下の定理が成り立つ。例えば、文献 [7] などを参照されたい。

定理 5 行列 $\hat{S}_k$ は以下のように展開される。

$$\hat{S}_k = X_r \Lambda_r^k f(\Lambda_r) \tilde{X}_r^H V = C_r^k \hat{S}_0, \quad C_r = X_r \Lambda_r \tilde{X}_r^H.$$

また、行列 $\hat{S}$ は

$$\hat{S} = X_r f(\Lambda_r) \tilde{X}_r^H [V, C_r V, C_r^2 V, \dots, C_r^{M-1} V]$$

と書くことができる。ここで、 $f(\Lambda_r) = \text{diag}(f(\lambda_1), f(\lambda_2), \dots, f(\lambda_r))$ である。

第2節で示されたように部分空間 $\mathcal{R}(S)$ や $\mathcal{R}(S_k)$ は対象の固有ベクトル $\mathbf{x}_i, \lambda_i \in \Omega$ のみを含む。一方で、定理5から $\mathcal{R}(\hat{S})$ や $\mathcal{R}(\hat{S}_k)$ はフィルター関数 $f(\lambda_i)$ に応じて対象の固有ベクトル成分 $\mathbf{x}_i, \lambda_i \in \Omega$ を多く含む部分空間であることがわかる。

block SS-RR 法は、一般化固有値問題 $A\mathbf{x}_i = \lambda_i B\mathbf{x}_i$ に対し、部分空間 $\mathcal{R}(\hat{S})$ を用いて Rayleigh-Ritz 過程に基づき固有対を計算する手法である。block SS-RR 法のアルゴリズム (Algorithm 2) は、2.2節で示されたように行列 $\hat{S}$ の特異値分解に基づく低ランク近似を用いて近似され導出される。定理5より、フィルター関数 $f(\lambda_i)$ に応じて部分空間 $\mathcal{R}(\hat{S})$ が対象の固有ベクトル成分を多く含むため、Rayleigh-Ritz 過程により、対象の固有対を精度よく得ることができる。詳細な誤差解析は文献 [7] を参照されたい。

一方, FEAST 法は 2.3 節で示されたように, Rayleigh-Ritz 過程を用いた加速部分空間反復法に基づき, その各反復では block SS-RR 法と同様に一般化固有値問題 $A\mathbf{x}_i = \lambda_i B\mathbf{x}_i$ に対し, 部分空間として $\mathcal{R}(\hat{S}_0)$ を用いた直交化無しの Rayleigh-Ritz 過程が行われている. block SS-RR 法と同様に, フィルター関数 $f(\lambda_i)$ に応じて部分空間 $\mathcal{R}(\hat{S}_0)$ が対象の固有ベクトル成分を多く含むため, Rayleigh-Ritz 過程により対象の固有対を精度よく得ることができる.

3.2 block SS-Hankel 法, Beyn 法, block SS-Arnoldi 法

定理 5 および行列 C_r の定義より, 以下の 2 つの定理が成り立つ.

定理 6 一般化固有値問題 (1) の固有値数 m が $m \leq LM$ を満たすとし, 入力行列 V が $V^H X_\Omega$ がフルランクとなる行列であるとする. この時, 以下が成り立つ.

$$\mathcal{R}(\hat{S}) = \mathcal{K}_M^\square(C_r, \hat{S}_0).$$

定理 7 一般化固有値問題 (1) の固有値数 m が $m \leq LM$ を満たすとし, 入力行列 V が $V^H X_\Omega$ がフルランクとなる行列であるとする. この時, 一般化固有値問題

$$A\mathbf{x}_i = \lambda_i B\mathbf{x}_i, \quad \lambda_i \neq \infty$$

は行列 C_r の固有値問題

$$C_r \mathbf{x}_i = \lambda_i \mathbf{x}_i, \quad \mathbf{x}_i \in \mathcal{R}(\hat{S}) = \mathcal{K}_M^\square(C_r, \hat{S}_0) \quad (8)$$

と等しい右固有対を持つ.

本小節では, 定理 6, 7 に基づき block SS-Hankel 法, Beyn 法および block SS-Arnoldi 法の関係性を示す.

block SS-Hankel 法の複素モーメントは, 定理 5 より,

$$\hat{\mu}_k^\square = V^H \hat{S}_k = V^H C_r \hat{S}_{k-1} = \dots = V^H C_r^k \hat{S}_0$$

と表現することができる. ここで,

$$\tilde{S} := [V, C_r^H V, (C_r^H)^2 V, \dots, (C_r^H)^{M-1} V]$$

と定義すると, block Hankel 行列 $\hat{H}_M^\square, \hat{H}_M^{\square<}$ はそれぞれ

$$\hat{H}_M^\square = \tilde{S}^H \hat{S}, \quad \hat{H}_M^{\square<} = \tilde{S}^H C \hat{S}$$

と書け, block SS-Hankel 法で得られる一般化固有値問題 (4) は

$$\tilde{S}^H C_r \hat{S} \mathbf{t}_i = \theta_i \tilde{S}^H \hat{S} \mathbf{t}_i,$$

と表すことができる。これは、block SS-Hankel 法が一般化固有値問題 $A\mathbf{x}_i = \lambda_i B\mathbf{x}_i$ ではなく、標準固有値問題 (8) に対して、近似固有ベクトル $\tilde{\mathbf{x}}_i$ を $\mathcal{R}(\hat{S})$ から探索し、残差 $\mathbf{r}_i := C_r \tilde{\mathbf{x}}_i - \tilde{\lambda}_i \tilde{\mathbf{x}}_i$ に対して $\mathbf{r}_i \perp \mathcal{R}(\tilde{S})$ の Petrov-Galerkin 型の直交条件を課す射影法として解釈することができることを意味する。block SS-Hankel 法のアルゴリズム (Algorithm 1) は、標準固有値問題 (8) に対する Petrov-Galerkin 型の射影法で得られる一般化固有値問題に対し低ランク近似をすることにより得られる。

同様に Beyn 法で得られる固有値問題 (7) の係数行列は $\hat{S}_1 = C_r \hat{S}_0$ の関係式および行列 $\hat{S}_0$ の特異値分解 (6) に基づき、

$$U_1^H \hat{S}_1 W_1 \Sigma_1^{-1} = U_1^H C_r \hat{S}_0 W_1 \Sigma_1^{-1} = U_1^H C_r U_1$$

と変形できる。従って、Beyn 法は標準固有値問題 (8) に対し、行列 $\hat{S}_0$ の低ランク近似により得られる部分空間 $\mathcal{R}(U_1)$ を用いた Rayleigh-Ritz 過程に基づく射影法であると解釈できる。

一方、定理 7 に基づき、block SS-Arnoldi 法は標準固有値問題 (8) に対し、block Krylov 部分空間 $\mathcal{K}_M^\square(C_r, \hat{S}_0)$ を用いた block Arnoldi 法を適用したアルゴリズムとして解釈できる。

3.3 各解法の関係性のまとめ

3.1 節で示されたように、block SS-RR 法および FEAST 法はどちらも一般化固有値問題 $A\mathbf{x}_i = \lambda_i B\mathbf{x}_i$ に対する Rayleigh-Ritz 過程に基づくアルゴリズムであり、それぞれ部分空間として $\mathcal{R}(\hat{S})$ および $\mathcal{R}(\hat{S}_0)$ が用いられている。FEAST 法は block SS-RR 法に対し $M = 1$ とし基底の直交化・低ランク近似を行わないことで簡略化する一方で、加速部分空間反復法に基づき反復計算を行うアルゴリズムである。

一方で 3.2 節で示されたように、block SS-Hankel 法、Beyn 法および block SS-Arnoldi 法は、block SS-RR 法や FEAST 法と異なり、一般化固有値問題 $A\mathbf{x}_i = \lambda_i B\mathbf{x}_i$ ではなく標準固有値問題 (8) に対する射影法として解釈できる。block SS-Hankel 法は部分空間を $\mathcal{R}(\hat{S})$ とした Petrov-Galerkin 過程に基づく射影法、Beyn 法は部分空間を $\mathcal{R}(\hat{S}_0)$ とした Rayleigh-Ritz 過程に基づく射影法、そして block SS-Arnoldi 法は部分空間を $\mathcal{R}(\hat{S}) = \mathcal{K}_M^\square(C_r, \hat{S}_0)$ とした block Arnoldi 法に基づく解法である。

また、block SS-Hankel 法、block SS-RR 法および block SS-Arnoldi 法は部分空間として $\mathcal{R}(\hat{S})$ を用いるため、各積分点で解く線形方程式の右辺ベクトル数 L に対し、使用可能な最大の部分空間のサイズは LM となる。一方で、FEAST 法と Beyn 法は部分空間として $\mathcal{R}(\hat{S}_0)$ を用いるため、各積分点で解く線形方程式の右辺ベクトル数 L に対し、使用可能な最大の部分空間のサイズは L となる。このため同じ部分空間サイズを使用する場合、 $\mathcal{R}(\hat{S}_0)$ を用いる FEAST 法と Beyn 法は、 $\mathcal{R}(\hat{S})$ を用いる block SS-Hankel 法、block SS-RR 法および block SS-Arnoldi 法に比べ部分空間生成に対し多くの計算コストを必要とすることが予想される。

3.1 節で示されたように、FEAST 法は block SS-RR 法を $M = 1$ として簡略化したアルゴリズムに対応する。一方で、 $M \geq 2$ とした Beyn 法に対応するアルゴリズムはこれま

Algorithm 6 An extended Beyn method

Input: $L, M, N \in \mathbb{N}, V \in \mathbb{C}^{n \times L}, (z_j, \omega_j)$ for $j = 1, 2, \dots, N$

Output: Approximate eigenpairs $(\tilde{\lambda}_i, \tilde{x}_i)$ for $i = 1, 2, \dots, \hat{m}$

- 1: Compute $\hat{S}_k = \sum_{j=1}^N \omega_j z_j^k (z_j B - A)^{-1} B V$,
and set $\hat{S} = [\hat{S}_0, \hat{S}_1, \dots, \hat{S}_{M-1}], \hat{S}_+ = [\hat{S}_1, \hat{S}_2, \dots, \hat{S}_M]$
 - 2: Compute SVD of $\hat{S}$: $\hat{S} = [U_1, U_2][\Sigma_1, O; O, \Sigma_2][W_1, W_2]^H$
 - 3: Compute eigenpairs (θ_i, t_i) of $U_1^H \hat{S}_+ W_1 \Sigma_1^{-1} t_i = \theta_i t_i$,
and compute $(\tilde{\lambda}_i, \tilde{x}_i) = (\theta_i, U_1 t_i)$ for $i = 1, 2, \dots, \hat{m}$
-

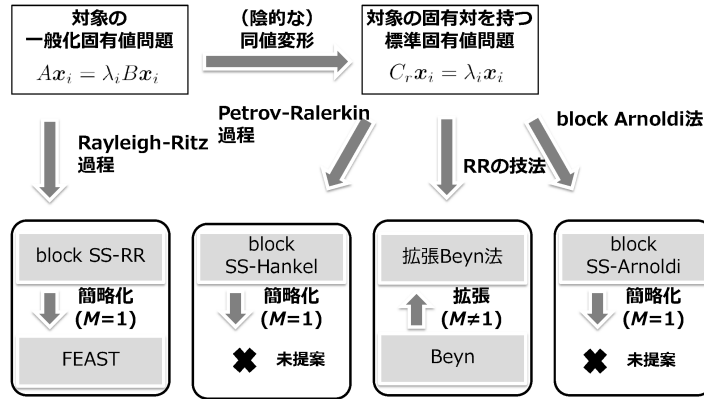


Fig. 1: A map of contour integral-based eigensolvers.

で提案されていない。本論文では、Beyn 法のアイディアに基づき部分空間として $\mathcal{R}(\hat{S})$ を用いるアルゴリズム（拡張 Beyn 法）を提案する。

2.5 節および 3.2 節で示されたように、Beyn 法では行列 $\hat{S}_0, \hat{S}_1$ が $\hat{S}_1 = C_r \hat{S}_0$ の関係性を満たす事に着目し、 $\hat{S}_0$ の特異値分解および行列 $\hat{S}_1$ に基づき固有値問題 (7) を導出している。使用する部分空間を $\mathcal{R}(\hat{S}_0)$ から $\mathcal{R}(\hat{S})$ と変更するため、行列 $\hat{S}$ の特異値分解および

$$\hat{S}_+ := [\hat{S}_1, \hat{S}_2, \dots, \hat{S}_M] = C_r \hat{S}$$

を用いる。拡張 Beyn 法のアルゴリズムを Algorithm 6 に示す。

各種周回積分型固有値解法、block SS-Hankel 法、block SS-RR 法、FEAST 法、block SS-Arnoldi 法、Beyn 法および拡張 Beyn 法の関係図を Fig. 1 に示す。

4 数値実験

本論文の主たる目的は各種の周回積分型固有値解法の数理的な関係性を解析することであるが、本節では、 $M = 1, 2, 4, 8$ とした block SS-Hankel 法、block SS-RR 法、block SS-Arnoldi 法および拡張 Beyn 法の有効性について、簡単な数値実験を通し比較を行う。

モデル問題として、5000 次元の乱数・密・実対称一般化固有値問題を考え、区間 $[-1, 1]$ に存在する 200 個の固有値および対応する固有ベクトルを求めることを考える。積分路

Table 1: Computational results of block SS-Hankel, block SS-RR, block SS-Arnoldi and extended Beyn method with $M = 1, 2, 4$ and 8 .

Method	M	$\hat{m}$	Time [sec.]				residual norm	
			t_{LU}	t_{Solve}	t_{Other}	t_{Total}	$\max_{\lambda_i \in \Omega} \ \mathbf{r}_i\ _2$	$\min_{\lambda_i \in \Omega} \ \mathbf{r}_i\ _2$
SS-Hankel	1	391	79.45	130.43	22.76	232.65	1.90E-13	1.32E-14
	2	421	80.08	69.32	21.52	170.93	1.17E-12	1.38E-14
	4	509	80.12	40.00	25.32	145.44	1.51E-11	4.77E-14
	8	733	80.81	27.49	35.47	143.76	6.74E-09	1.25E-11
SS-RR	1	398	79.71	127.25	33.67	240.63	3.52E-14	1.89E-14
	2	413	81.43	67.17	34.30	182.90	1.77E-11	1.89E-14
	4	447	80.25	39.57	36.80	156.61	8.29E-14	2.24E-14
	8	541	81.10	27.18	39.03	147.30	7.66E-13	5.98E-14
SS-Arnoldi	1	800	79.65	126.25	61.39	267.28	5.96E-10	6.32E-14
	2	800	80.04	67.84	45.17	193.06	5.12E-10	1.77E-13
	4	800	79.39	39.41	45.07	163.88	1.55E-08	8.06E-14
	8	800	80.70	27.09	49.05	156.83	2.38E-09	6.69E-14
Ex Beyn	1	398	80.73	127.53	36.35	244.61	3.53E-14	1.91E-14
	2	413	80.09	67.53	36.23	183.85	1.68E-10	1.95E-14
	4	447	80.96	39.75	36.66	157.36	8.33E-14	2.28E-14
	8	541	80.46	27.12	40.94	148.53	7.67E-13	6.10E-14

Γ を中心 0, 長径 2, 短径 0.2 の楕円とし, パラメータ (L, M) は $LM = 800$ となるように $(800, 1), (400, 2), (200, 4), (100, 8)$ と設定した. また, 積分点数は $N = 32$ とした. この時, 問題の対称性を利用することで, 解くべき線形方程式の本数は $N/2 = 16$ となる. 各解法の低ランク近似は行列の最大特異値を σ_1 とすると, 特異値が $\sigma_i/\sigma_1 \geq 10^{-14}$ を満たす特異値 σ_i および対応する特異ベクトルを用いた.

プログラムは Fortran 90 で実装され, 全ての計算は倍精度演算で行った. 数値実験は Intel Xeon CPU E5-2450 (2.10GHz) 8 コアを使用し, 各積分点での計算のみ MPI により並列化を行った.

各解法の実験結果を Table 1 に示す. まず, 計算時間について考察する. 表中の t_{LU} は各線形方程式の LU 分解にかかる計算時間, t_{Solve} は前進後退代入時間, t_{Other} は行列の特異値分解や小規模固有値問題の求解時間, t_{Total} は合計計算時間を示す. Table 1 に示されるように, 計算時間の多くの部分は各積分点での線形方程式の求解 ($t_{LU} + t_{Solve}$) が占める. 特に, $M = 1$ とした場合では $M = 8$ の場合と比べ各方程式の右辺ベクトル数が 8 倍必要であり, t_{Solve} に多くの計算時間がかかる. 結果として M を小さくするに連れて合計計算時間 (t_{Total}) は大きく増加している.

一方 t_{Other} に着目すると, 現状の block SS-Arnoldi 法では行列の低ランク近似に基づく計算コストの削減法が適用されておらず, 他の解法と比較して計算時間が多くかかっている. また, block SS-Hankel 法, block SS-RR 法および拡張 Beyn 法では, $M = 1$ と $M = 8$ を比較すると, $M = 1$ の場合の方が数値ランク (小規模固有値問題のサイズ) $\hat{m}$ が小さくなり, t_{Other} が小さくなる結果が得られた. また, block SS-Hankel 法は他の解法と異な

り行列の直交化を行わないため, t_{Other} が小さくなっていることがわかる.

次に, 計算精度について考察する. 直交化を行わない block SS-Hankel 法および低ランク近似を行わない block SS-Arnoldi 法が, 特に $M = 8$ の場合に, 他の解法と比較して精度が悪い結果が得られた. 一方で, block SS-RR 法および拡張 Beyn 法については $M = 8$ とした場合についてもよい精度の固有対が得られている.

5 まとめ

本論文では, 代表的な周回積分型固有値解法である, block SS-Hankel 法, block SS-RR 法, block SS-Arnoldi 法, FEAST 法および Beyn 法を一般化固有値問題 (1) に適用した際のアロリズムの数理的な関係性について解析を行った. 本論文の解析により, block SS-RR 法および FEAST 法が一般化固有値問題 $A\mathbf{x}_i = \lambda_i B\mathbf{x}_i$ に対する射影法であるのに対し, block SS-Hankel 法, block SS-Arnoldi 法および Beyn 法は標準固有値問題 $C\mathbf{x}_i = \lambda_i \mathbf{x}_i$ に対する射影法として解釈できる事が示された. また, 各解法で使用する部分空間の違いに着目し, Beyn 法の $M \geq 2$ とした拡張法である拡張 Beyn 法を提案した. 我々の数値実験から, 各解法は $M \geq 2$ と設定することにより, $M = 1$ の場合と比較し計算時間を削減できる事が示された.

今後の課題は行列束が対角化できない場合の解析や非線形固有値問題に対する各解法の数理的な関係性の解析である. また, 各解法を実問題に適用した際の有効性をより詳細に比較することも今後の課題としてあげられる.

謝辞

本研究はH P C I 戦略プログラム 分野5「物質と宇宙の起源と構造」, JST/CREST, KAKENHI(Nos. 25104701, 25286097, 25870099) の援助を受けた.

参考文献

- [1] Asakura, J., Sakurai, T., Tadano, H., Ikegami, T., Kimura, K.: A numerical method for nonlinear eigenvalue problems using contour integrals. JSIAM Letters, **1**(2009), 52–55.
- [2] Asakura, J., Sakurai, T., Tadano, H., Ikegami, T., Kimura, K.: A numerical method for polynomial eigenvalue problems using contour integral. Japan J. Indust. Appl. Math, **27**(2010), 73–90.
- [3] Beyn, W.-J.: An integral method for solving nonlinear eigenvalue problems. Lin. Alg. Appl., **436**(2012), 3839–3863.

- [4] Ikegami, T., Sakurai, T., Nagashima, U.: A filter diagonalization for generalized eigenvalue problems based on the Sakurai-Sugiura projection method. *J. Comput. Appl. Math.*, **233**(2010), 1927–1936.
- [5] Ikegami, T., Sakurai, T.: Contour integral eigensolver for non-Hermitian systems: a Rayleigh-Ritz-type approach. *Taiwan. J. Math.*, **14** (2010), 825–837.
- [6] Imakura, A., Du, L. Sakurai, T.: A block Arnoldi-type contour integral spectral projection method for solving generalized eigenvalue problems, *Applied Mathematics Letters*, **32** (2014), 22–27.
- [7] Imakura, A., Du, L. Sakurai, T.: Accuracy analysis on the Rayleigh-Ritz type of the contour integral based eigensolver for solving generalized eigenvalue problems, Technical Report of Department of Computer Science, University of Tsukuba (CS-TR), CS-TR-14-23.
- [8] 今倉 暁, 杜 磊, 櫻井 鉄也, Communication-Avoiding Arnoldi 版周回積分型固有値解法, 日本応用数理学会 2014 年度年会.
- [9] Kravanja, P., Sakurai, T., van Barel, M.: On locating clusters of zeros of analytic functions. *BIT*, **39** (1999), 646–682.
- [10] Polizzi, E.: A density matrix-based algorithm for solving eigenvalue problems. *Phys. Rev. B*, **79** (2009), 115112.
- [11] Sakurai, T., Sugiura, H.: A projection method for generalized eigenvalue problems using numerical integration. *J. Comput. Appl. Math.*, **159** (2003), 119–128.
- [12] Sakurai, T. Tadano, H.: CIRRR: a Rayleigh-Ritz type method with counter integral for generalized eigenvalue problems. *Hokkaido Math. J.*, **36** (2007), 745–757.
- [13] Tang, P. T. P., Polizzi, E.: FEAST as a subspace iteration eigensolver accelerated by approximate spectral projection. *SIAM J. Matrix Anal. Appl.*, **35**(2014), 354–390.
- [14] Yokota, S., Sakurai, T.: A projection method for nonlinear eigenvalue problems using contour integrals. *JSIAM Letters*, **5**(2013), 41–44.