

フィルタ対角化法で用いる初期ベクトルの改良について

村上 弘^{1,a)}

概要：固有値問題の固有対で指定した区間に固有値があるものを解く計算をレゾルベントで構成されたフィルタを用いて行う。フィルタに与える最初のベクトルの組は「必要な固有ベクトル」をすべて含んでいる必要がある。そこでこれまでは要素が一樣乱数であるベクトルを十分多く生成して、それを正規直交化して得られるベクトルの組を用いてきた。しかし固有値問題の次数が大きくなるにつれて、そのようなベクトルの組が含む「必要な固有ベクトル」の割合は減少し、それに伴って得られる固有対の精度も低下する。そこでフィルタに最初に与えるベクトルの組を作成する方法として、まず要素が一樣乱数であるベクトルを十分多く生成し、それに手間のあまり掛からない操作を施すことで「不要な固有ベクトル」を含む割合を減らし、それを正規直交化する。この方法について考察する。

キーワード：固有値, 固有ベクトル, フィルタ対角化, レゾルベント, ヤコビ

1. はじめに

与えられた実対称定値一般固有値問題 $A\mathbf{v} = \lambda B\mathbf{v}$ に対して、固有対で固有値が指定された下端あるいは中間の区間にあるものを近似して求める。

最初に与えたベクトルの組に対して（単一のレゾルベントで構成されたあまり特性の良くない）フィルタと正規直交化の操作を組み合わせた処理を2回まで適用することにより、ベクトルの組が含む「必要な固有ベクトル」の割合を高めて近似固有対を求める場合についてであるが、既に前回の報告 [2] の実験で示したように、レゾルベント型のフィルタに最初に与える十分多くのベクトルの組として、これまでのように要素が一樣乱数であるものをそのまま用いるのではなくて、それに一種の前処理として（フィルタに用いるのと同じシフトを持つ）レゾルベントと正規直交化を組み合わせた操作を k 回（1回から数回程度）施してあらかじめ「必要な固有ベクトル」を含む割合を増やしたものをを用いれば、得られる近似固有対の精度はかなり良くなる。この方法はうまく働き、用いるレゾルベントのシフトをフィルタの中で用いるレゾルベントと同じにすれば、連立1次方程式を直接法で解いてレゾルベントの作用を実現している場合には行列分解を共用できるが、レゾルベントの作用を前進後退代入で行う手間、およびレゾルベントの適用ごとに B -正規直交化を行うならばその分の手間も

増える。

そこで、元の固有値問題の行列が（ステンスルで与えられるような）極めて疎なものである場合には、レゾルベントの適用よりも少ない手間ですでに何らかの方法を適用することで要素が一樣乱数であるベクトルの組から「不要な固有ベクトル」を含む割合を減らしたものを作って、それをフィルタに与える最初のものとする 것을考える。このような方法を前回の報告 [2] に於いて特殊な例題について少し試してみたが、今回はその方法をもう少し追求してみた。これはいわゆる（標準固有値問題に対する）多項式フィルタ（Polynomial Filter）であり、文献 [1] などに類似したあるいはより良いと思われる方法の記述が見られる。

2. ベクトルの組の Jacobi 走査の作用の多項式による前処理

1 辺の長さ π の立方体領域を各辺方向を $N_1 + 1, N_2 + 1, N_3 + 1$ の区間に等分割して格子点をとり、境界条件が零固定の3次元ラプラス作用素 $(-\Delta)$ を差分法 (FDM) の中心差分 (7点差分) で離散化近似した行列を A とする。各方向の格子幅をそれぞれ $h_1 = \pi/(N_1 + 1), h_2 = \pi/(N_2 + 1), h_3 = \pi/(N_3 + 1)$ とするとき、7点差分法のステンスルの係数は中心点では $d = 2h_1^{-2} + 2h_2^{-2} + 2h_3^{-2}$ であり、各方向の隣接格子点ではそれぞれ $-h_1^{-2}, -h_2^{-2}, -h_3^{-2}$ となる。係数行列 A は実対称で対角要素はすべて d で、各行について非対角要素の絶対値の和が d 以下なので、Gershgorin の定理からすべての固有値は中心が d で半径が d の複素円盤に含まれる実数なので式 (1) が成り立つ。

¹ 東京都立大学 数理科学専攻
Department of Mathematical Sciences, Tokyo Metropolitan University, Hachioji, Tokyo 192-0397, Japan

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

$$0 \leq \lambda \leq 2d. \quad (1)$$

シフト行列 $A - sI$ の対角分離は式 (2) になる.

$$(d - s)I - (dI - A). \quad (2)$$

このとき式 (3) の Jacobi 反復法の走査 1 回によりベクトル $\mathbf{x}$ から $\mathbf{y}$ を作る.

$$\mathbf{y} \leftarrow (d - s)^{-1}(dI - A)\mathbf{x}. \quad (3)$$

いま行列 A は対角化可能と仮定して, そのすべての固有ベクトルを固定して, それを基底とするベクトル $\mathbf{x}$ の展開を式 (4) とすれば, 同じ基底によるベクトル $\mathbf{y}$ の展開に於いて $\mathbf{v}_\lambda$ の係数は式 (5) になる. つまり $\mathbf{y}$ の展開係数は $\mathbf{x}$ の係数に倍率として式 (6) を乗じたものになる.

$$\mathbf{x} = \sum_{\lambda} c_{\lambda} \mathbf{v}_{\lambda}. \quad (4)$$

$$(d - s)^{-1}(d - \lambda)c_{\lambda}. \quad (5)$$

$$\frac{d - \lambda}{d - s} = 1 - \frac{\lambda - s}{d - s}. \quad (6)$$

いま r を定数として, $\mathbf{y}$ と $\mathbf{x}$ の線形結合 $\mathbf{x}'$ を式 (7) で作る. すると, $\mathbf{x}'$ の各展開係数は $\mathbf{x}$ の各展開係数に式 (8) を乗じたものになる.

$$\mathbf{x}' = r\mathbf{y} + (1 - r)\mathbf{x}. \quad (7)$$

$$1 - rz_{\lambda}, \text{ ただしここで } z_{\lambda} \equiv \frac{\lambda - s}{d - s}. \quad (8)$$

いまの例題では $0 \leq \lambda \leq 2d$ であるから, z_{λ} のとりうる値の範囲はシフト s を零にとるときは式 (9) で, シフト s を正にとるときは (ただし $s < d$ とする) 式 (10) である.

$$0 \leq z_{\lambda} \leq 2. \quad (9)$$

$$-\frac{s}{d - s} \leq z_{\lambda} \leq 2 + \frac{s}{d - s}. \quad (10)$$

たとえば $N_1 = 20$, $N_2 = 30$, $N_3 = 40$ のときには $d = 624.7$ である. 下端固有値の場合にシフト s を零にとると, 行列 A の固有値全体に対応する z_{λ} の範囲は $[0, 2]$ であり, 固有値の区間 $\lambda \in [0, 100]$ には $z_{\lambda} \in [0, 0.1601]$ が対応する.

中間固有値の場合で, 求める固有対の固有値の区間が $[100, 200]$ のときには, その中央の値 $s = 150$ をシフトにとれば, 行列 A の固有値全体に対応する z_{λ} の範囲は $[-0.3160, 2.3160]$ になり, そのうちで固有値の区間 $\lambda \in [100, 200]$ には $|z_{\lambda}| \leq 0.1053$ が対応する.

あるいは中間固有値の場合で, 求める固有対の固有値の区間が $[100, 150]$ のときには, その中央の値 $s = 125$ をシフトにとれば, 行列 A の固有値全体に対応する z_{λ} の範囲は $[-0.2501, 2.2501]$ になり, そのうちで固有値の区間 $\lambda \in [100, 150]$ には $|z_{\lambda}| \leq 0.05003$ が対応する.

2.1 Jacobi 走査 1 回に対応する 1 次式

ここでは定数 r の値を 0.75 と選んだ場合について, $\mathbf{x}$ から $\mathbf{x}'$ を作る場合の z_{λ} に対する固有ベクトル展開の係数に対する倍率 $f(z_{\lambda})$ のグラフを図 1 に示す. さらに $\mathbf{x}$ から $\mathbf{x}'$ を作る処理を 10 回繰り返して適用する場合の固有ベクトル展開係数に対する倍率のグラフを図 2 に, その片対数によるプロットを図 3 に示す.

2.2 Jacobi 走査 2 回に対応する 2 次式

いま $r = r_1$ として $\mathbf{x}$ から $\mathbf{x}'$ を作り, さらに $r = r_2$ として $\mathbf{x}'$ から $\mathbf{x}''$ を作る. これら 2 段階の操作をまとめた場合の固有ベクトル展開の各係数への倍率は 2 次式 (11) になる.

$$f_2(z_{\lambda}) = (1 - r_1 z_{\lambda})(1 - r_2 z_{\lambda}). \quad (11)$$

そこで 2 次式 $f_2(z)$ の振る舞いになるべく都合の良いものとなるように r_1, r_2 をうまく選ぶ. ここでは式 (12) のようにとしてみた.

$$\begin{cases} r_1 = 1/0.7, \\ r_2 = 1/1.75. \end{cases} \quad (12)$$

多項式 $f_2(z)$ のグラフを図 4 に示す. また処理を 10 回繰り返す場合に対応する $|f_2(z)|^{10}$ の片対数グラフを図 5 に示す. さらに公平な比較のために Jacobi 走査を合計 2 回行う処理に対応する $|f_1(z)|^2$ と $|f_2(z)|$ のグラフを図 6 に示す.

2.3 Jacobi 走査 3 回に対応する 3 次式

固有ベクトル展開係数への倍率が, 3 次式 (13) で表される場合を考える.

$$f_3(z_{\lambda}) = (1 - r_1 z_{\lambda})(1 - r_2 z_{\lambda})(1 - r_3 z_{\lambda}). \quad (13)$$

ここでは条件 (14) を満たす 3 次式 (15) を採用する (これが最良というわけではない). その根の逆数である r_1, r_2, r_3 の数値は式 (16) で与えられる.

$$\begin{cases} f_3(0) &= 1, \\ f_3(0.5) &= 0, \\ f_3(1) &= 0.2, \\ f_3(1.5) &= 0.4, \\ f_3(2) &= -0.6. \end{cases} \quad (14)$$

$$f_3(z) = 1 - 4z + (24/5)z^2 - (8/5)z^3. \quad (15)$$

$$\begin{cases} r_1 = 2, \\ r_2 = 1.4472\,13595\,49995\,8, \\ r_3 = 0.5578\,64045\,00042. \end{cases} \quad (16)$$

3 次の多項式 $f_3(z)$ のグラフを図 7 に示す. さらに 10 回反復した場合に対応する $|f_3(z)|^{10}$ の対数グラフを図 8 に

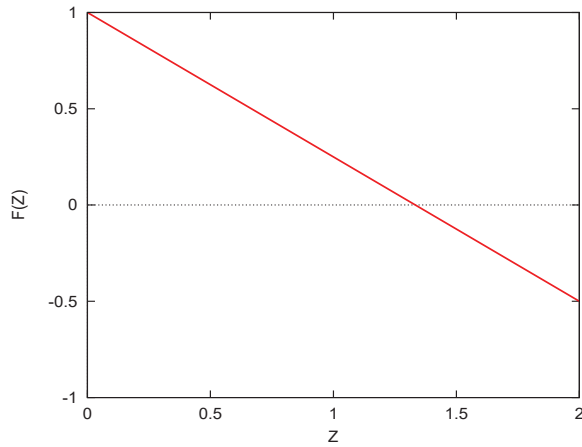


図 1 $f_1(z)$ (1 次式)

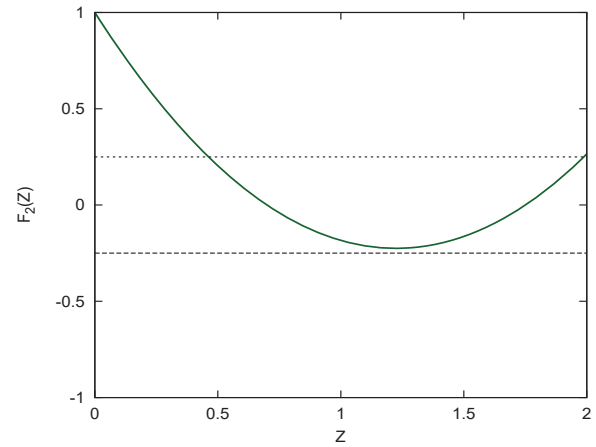


図 4 $f_2(z)$ (2 次式)

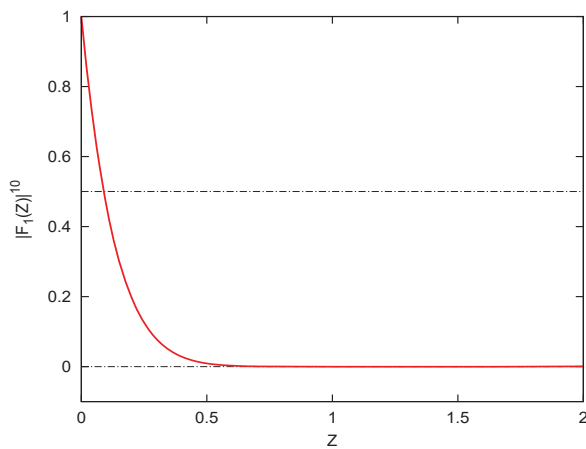


図 2 $|f_1(z)|^{10}$ (1 次式の操作の 10 回適用)

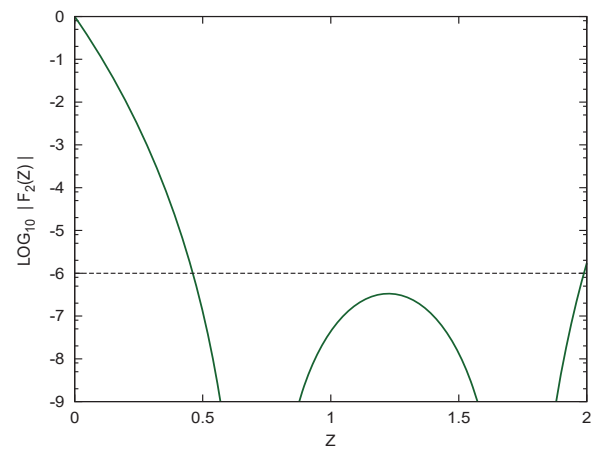


図 5 $|f_2(z)|^{10}$ (2 次式の操作の 10 回適用, 対数プロット)

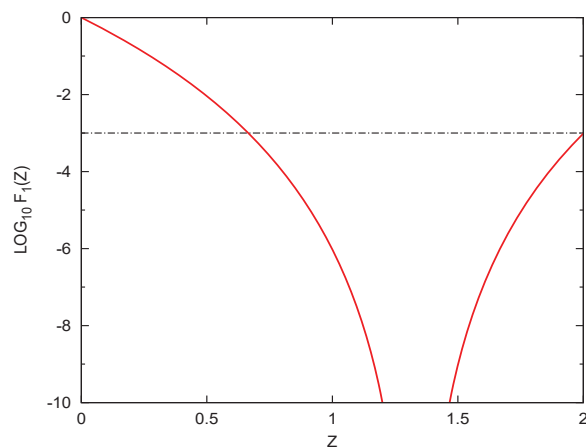


図 3 $|f_1(z)|^{10}$ (1 次式の操作の 10 回適用, 対数プロット)

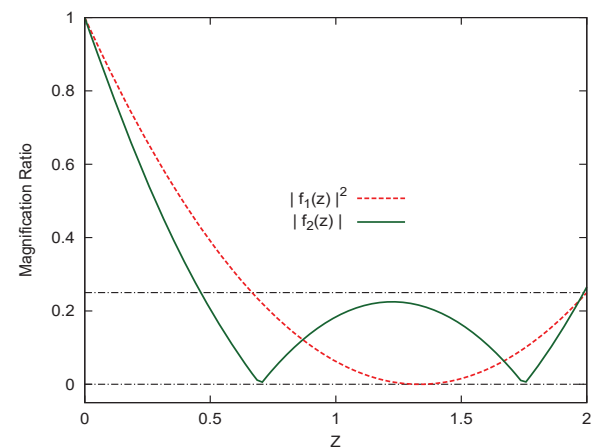


図 6 $|f_2(z)|$ と $|f_1(z)|^2$ (1 次式と 2 次式の比較)

示す。また公平な比較のために Jacobi 走査を合計 6 回行う場合に対応した $|f_1(z)|^6$, $|f_2(z)|^3$, $|f_3(z)|^2$ の各グラフを図 9 に示す。

3. 実験例 (下端側固有値を持つ固有対の近似)

領域が 1 辺の長さ π の立方体内部で境界条件が零固定で

ある 3 次元のラプラス固有値問題を有限要素法 (FEM) により離散化して得られた一般固有値問題 $\mathbf{A}\mathbf{v} = \lambda\mathbf{B}\mathbf{v}$ をフィルタを用いて固有対を近似して解く問題である。有限要素法の各要素は立方体を 3 辺の各方向をそれぞれ $N_1 + 1$, $N_2 + 1$, $N_3 + 1$ に等分割して得られる直方体であり、要素内の基底関数には最も低次の近似である 3 重線形関数を用

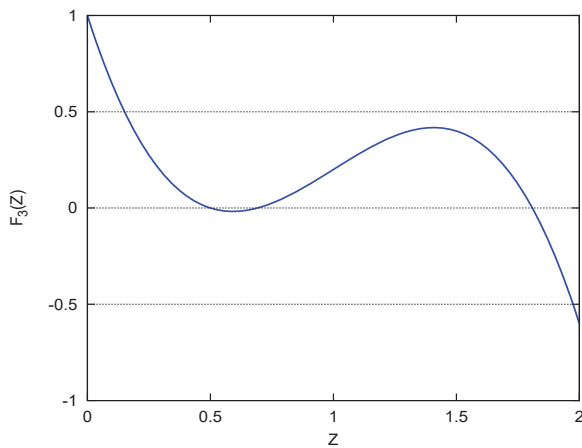


図 7 $f_3(z)$ (3 次式)

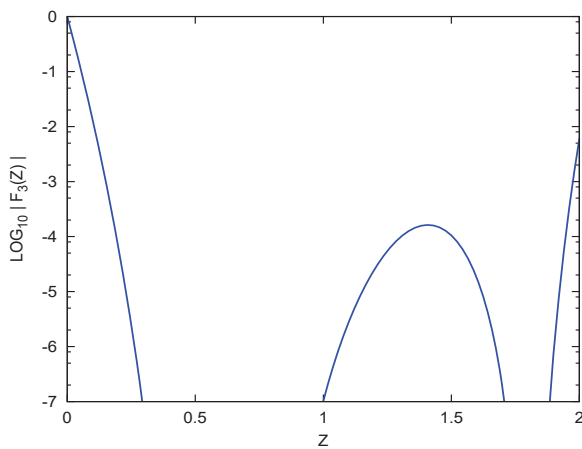


図 8 $|f_3(z)|^{10}$ (3 次式の操作の 10 回適用, 対数プロット)

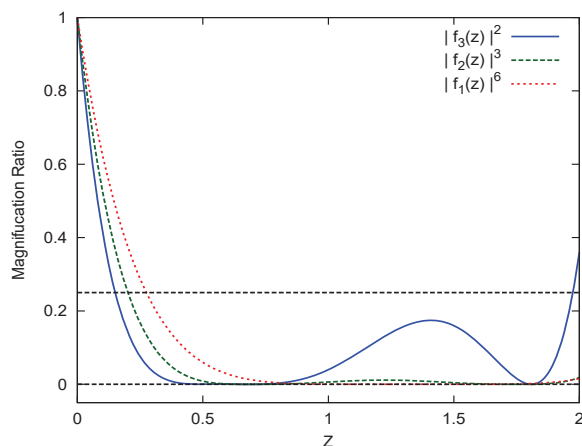


図 9 $|f_1(z)|^6$, $|f_2(z)|^3$, $|f_3(z)|^2$ (1 次, 2 次, 3 次の式の比較)

いた。

最初に与えた m 個のベクトルの組に対してフィルタを 1 回ないし 2 回適用して不変部分空間の近似を作り, それから Rayleigh-Ritz 法により必要な固有対を求めるものとする。下端固有値用のフィルタは前回の報告 [2] で用いたものと同じもので, 単一のレゾルベントで構成された次数

$n = 4$ の Chebyshev 多項式型でかなり低次のものである (他のパラメタは $\mu = 1.5$, $g_{\text{stop}} = 1\text{E-}5$ である)。

フィルタを適用する最初の m 個のベクトルの組は, まず要素が一様乱数のものを作り, それに対して差分法 (FDM) による 3 次元ラプラシアン ($-\Delta$) の 7 点中心差分による離散化である行列 A に対する Jacobi 走査の作用の低次の多項式を ℓ 回反復適用することにより, 行列 A の大きな固有値を持つ固有ベクトルを含む割合を減少させて作成した。ただし Jacobi 走査の多項式の操作を 10 回適用することによりベクトルの組に B -正規直交化を適用している。

これは FEM の離散化で得られた一般固有値問題における下端側固有値の「必要な固有ベクトル」を含む割合を増す操作を FDM の離散化で得られる標準固有値問題の行列 A の下端側固有値の固有ベクトルを含む割合を増す操作で置き換える近似を導入したことになる。いま用いている FEM と FDM の離散化法はどちらも解ベクトルの各要素は連続問題の近似解の格子点上での値をそのまま表しているの, 相互に読み替えのための変換をせずに離散化されたベクトルの値を移行できる。ただし, これは連続問題を FEM で近似して得た一般固有値問題についてベクトルの組が固有値が下端側にある「必要な固有ベクトル」を含む割合を高める作業を, 同じ連続問題を FDM で近似して得た標準固有値問題についてベクトルの組が固有値が下端側にある固有ベクトルを含む割合を高める作業により代用できると仮定していることになる。FEM による一般固有値問題と FDM による標準固有値問題では固有値分布も固有ベクトルも同じではない。そこで FEM 側で固有値が下端側の区間 $[a, b]$ にあるすべての固有ベクトルを求める場合に, そのためのフィルタへの初期ベクトルの組を作るには FDM 側としてどこまでの範囲にある固有値を持つ固有ベクトルを含む割合を高めれば良いのかは自明ではない。たとえば格子分割が $(N_1, N_2, N_3) = (20, 30, 40)$ の場合には, 区間 $[a, b] = [0, 100]$ に固有値が入る固有対の数は FEM では 378 であるが FDM では 460 であり, 固有値が区間 $[a, b] = [100, 200]$ に入る固有対の数は FEM では 684 であるが FDM では 1,225 である。また格子分割が $(N_1, N_2, N_3) = (50, 60, 70)$ の場合には, 区間 $[a, b] = [0, 100]$ に固有値が入る固有対の数は FEM では 402 であるが FDM では 425 であり, 区間 $[a, b] = [100, 200]$ に固有値が入る固有対の数は FEM では 801 であるが FDM では 887 である。このように FEM と FDM では固有値が同じ区間にある固有対の数にはかなりの違いがあり得る。

以下の実験では問題は比較的小規模の $(N_1, N_2, N_3) = (20, 30, 40)$ の場合であり, 求める固有対の固有値の範囲は $[0, 100]$ とした。固有値がそのような範囲にある固有対の正しい数は 378 である。フィルタに最初に与えるベクトルの組の作成では, 一様乱数で作るベクトルの数を $m = 800$ とした。

対称定値一般固有値問題 $A\mathbf{v} = \lambda B\mathbf{v}$ の近似固有対 $(\lambda, \mathbf{v})$ に対する相対残差としては式 (17) を用いた.

$$\frac{\|A\mathbf{v} - \lambda B\mathbf{v}\|_2}{\|\lambda B\mathbf{v}\|_2}. \quad (17)$$

3.1 実験：1 次式 $f_1(z)$ の場合

フィルタを適用する最初のベクトルの組を作成するのに、要素が一様乱数であるベクトル m 個の組に対して、(既に副節 2.1 で示した) 1 次の多項式 $f_1(z) = 1 - 0.75z$ に対応する処理操作を合計 ℓ 回施した場合の例を示す. 回数 $\ell = 0$ は適用しないことを意味し、それ以外は ℓ を 10 から始めて 20, 40, 80, 160 などと 2 倍ずつにした. ただし操作を 10 回適用するごとに B -正規直交化を行っている. 1 次式に対応する処理を ℓ 回適用する場合には、その中の Jacobi 走査の回数は ℓ である.

フィルタ適用 1 回 (IT=1) と 2 回 (IT=2) それぞれの場合について倍精度計算で求めた近似固有対のうちで固有値が指定した区間 $[0, 100]$ にあるものの数を表 1 に (数が正しくないものには下線を付した), それらについて相対残差の最大値を表 2 と図 10 に、固有値に対する相対残差の対数グラフを図 11 と図 12 に示す. 同様に単精度計算で求めた近似固有対のうちで固有値が指定した区間にあるものの数を表 3 に、それらについての相対残差の最大値を表 4 と図 13 に、相対残差の対数グラフを図 14 と図 15 に示す.

表 1 近似固有対の数 (1 次式, $m = 800$, 倍精度計算)

ℓ	IT=1	IT=2
0	275	378
10	<u>376</u>	378
20	378	378
40	378	378
80	378	378
160	378	378
320	378	378
640	378	378

表 2 相対残差の最大値 (1 次式, $m = 800$, 倍精度計算)

ℓ	IT=1	IT=2
0	1.4E-01	8.4E-03
10	2.0E-02	2.4E-04
20	7.6E-03	8.3E-05
40	1.5E-03	1.7E-05
80	1.4E-04	1.1E-06
160	8.6E-07	8.0E-09
320	7.2E-11	5.9E-13
640	1.6E-11	1.2E-13

表 3 近似固有対の数 (1 次式, $m = 800$, 単精度計算)

ℓ	IT=1	IT=2
0	<u>276</u>	378
10	<u>377</u>	378
20	<u>377</u>	378
40	<u>377</u>	378
80	<u>377</u>	378

表 4 相対残差の最大値 (1 次式, $m = 800$, 単精度計算)

ℓ	IT=1	IT=2
0	1.5E-01	8.8E-03
10	2.1E-02	2.9E-04
20	1.0E-02	1.0E-04
40	5.3E-03	5.3E-05
80	8.2E-03	5.4E-05

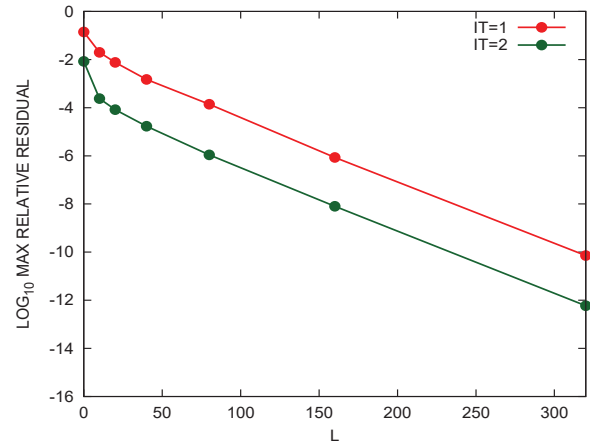


図 10 相対残差の最大値 (1 次式, $m = 800$, 倍精度計算)

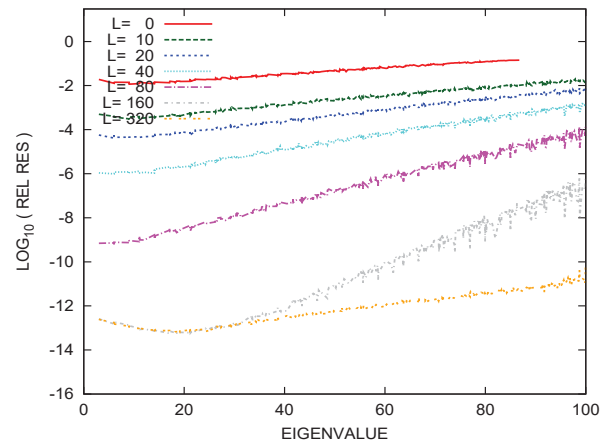


図 11 相対残差 (1 次式, $m = 800$, IT=1, 倍精度計算)

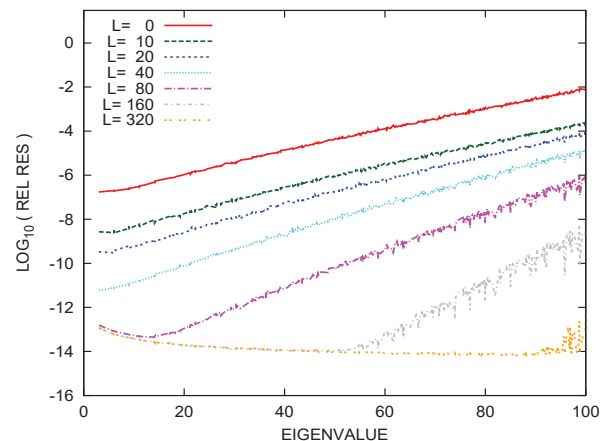


図 12 相対残差 (1 次式, $m = 800$, IT=2, 倍精度計算)

表の数値あるいはグラフから、倍精度計算と単精度計算のどちらについても、一様乱数で作ったベクトルの組に対して Jacobi 走査を元にした 1 次式 $f_1(z)$ に対応する操作を繰り返し適用して作ったベクトルの組をフィルタに与える最初のものであるとすることで、得られる近似解の精度が良くなることわかる。操作回数が 10 回程度の少ない範囲での残差の減少は急であるが、すぐに減少率が一定値に近づき、回数に対する 1 次収束的な振る舞いを見せている。この実験例では倍精度計算では操作を施さない場合に比べて IT=1 の場合でも相対残差の最大値の改良は $\ell = 10$ で 1 桁程度、 $\ell = 20$ で 1 桁半程度、 $\ell = 40$ で 2 桁程度、 $\ell = 80$ で 3 桁程度などとなっている。IT=2 の場合には改良は $\ell = 10$ で 1 桁半程度、 $\ell = 20$ で 2 桁程度、 $\ell = 40$ で 2 桁半程度、 $\ell = 80$ で 3 桁半程度などとなっている。単精度計算の場合も $\ell = 40$ 程度で既に十分良い結果が得られている。

3.2 実験：2 次式 $f_2(z)$ の場合

フィルタを適用する最初のベクトルの組を作成するのに、要素が一様乱数であるベクトル m 個の組に対して、(既に副節 2.2 で示した) 2 次式 $f_2(z) = (1-r_1z)(1-r_2z)$ に対応する処理を ℓ 回施した場合の例を示す。ここで $r_1 = 1/0.7$, $r_2 = 1/1.75$ である。2 次式に対応する処理を ℓ 回適用する場合には、その中での Jacobi 走査の回数は 2ℓ である。

フィルタ適用 1 回 (IT=1) と 2 回 (IT=2) それぞれの場合について倍精度計算で求めた近似固有対のうちで固有値が指定した区間 $[0, 100]$ にあるものの数を表 5 に、それらについての相対残差の最大値を表 6 と図 16 に固有値に対する相対残差の対数グラフを図 17 と図 18 に示す。同様に単精度計算で求めた近似固有対のうちで固有値が指定した区間にあるものの数を表 7 に、それらについての相対残差の最大値を表 8 と図 19 に、相対残差の対数グラフを図 20 と図 21 に示す。

2 次式 $f_2(z)$ に対応する操作を 1 次式 $f_1(z)$ の場合と比較してみると、倍精度計算の場合には 2 次式の操作で $\ell = 80$ のときには相対残差の最大値は IT=1 で $8.0\text{E-}9$, IT=2 で $7.3\text{E-}11$ となっているが、それに対して 1 次式の操作で $\ell = 160$ のときには相対残差の最大値は IT=1 で $8.6\text{E-}7$, IT=2 で $8.0\text{E-}9$ となっていることなどから、Jacobi 走査の回数を同じにした場合には、2 次式の操作を用いる方が相対残差の最大値の減少が早くなっていることがわかる。

3.3 実験：3 次式 $f_3(z)$ の場合

フィルタを適用する最初のベクトルの組を作成するのに、要素が一様乱数であるベクトル m 個の組に対して (既に副節 2.3 で示した) 3 次の多項式 $f_3(z)$ に対応する処理を ℓ 回

施した場合の例を示す。3 次式に対応する処理を ℓ 回適用する場合には、その中での Jacobi 走査の回数は 3ℓ である。

フィルタ適用 1 回 (IT=1) と 2 回 (IT=2) それぞれの場合について倍精度計算で求めた近似固有対のうちで固有値が指定した区間 $[0, 100]$ にあるものの数を表 9 に、それらについての相対残差の最大値を表 10 と図 22 に、固有値に対する相対残差の対数グラフを図 23 と図 24 に示す。

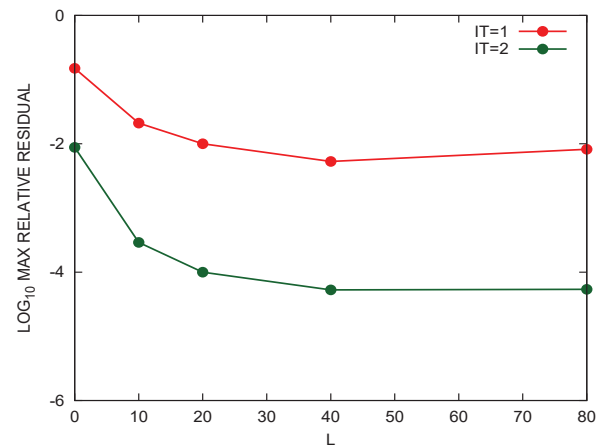


図 13 相対残差の最大値 (1 次式, $m = 800$, 単精度計算)

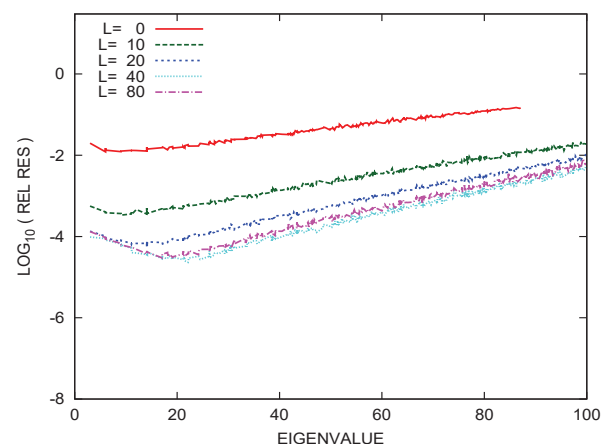


図 14 相対残差 (1 次式, $m = 800$, IT=1, 単精度計算)

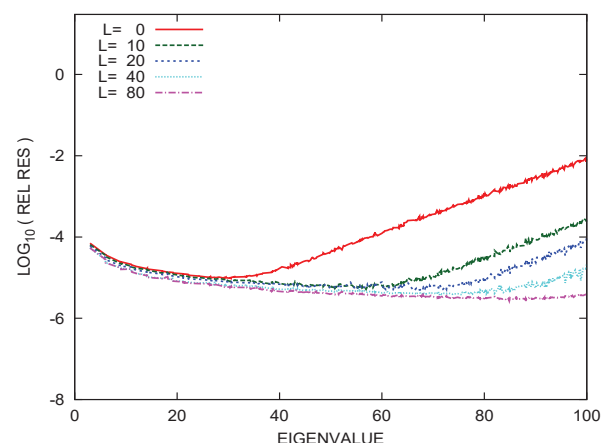


図 15 相対残差 (1 次式, $m = 800$, IT=1, 単精度計算)

表 5 近似固有対の数 (2 次式, $m = 800$, 倍精度計算)

ℓ	IT=1	IT=2
0	275	378
10	378	378
20	378	378
40	378	378
80	378	378
160	378	378
320	378	378

表 6 相対残差の最大値 (2 次式, $m = 800$, 倍精度計算)

ℓ	IT=1	IT=2
0	1.4E-01	9.0E-03
10	3.8E-03	3.3E-05
20	4.3E-04	4.2E-06
40	9.9E-06	1.1E-07
80	8.0E-09	7.3E-11
160	1.3E-11	1.2E-13
320	2.6E-11	1.2E-13

表 7 近似固有対の数 (2 次式, $m = 800$, 単精度計算)

ℓ	IT=1	IT=2
0	275	378
10	378	378
20	378	378
40	378	378
80	378	378

表 8 相対残差の最大値 (2 次式, $m = 800$, 単精度計算)

ℓ	IT=1	IT=2
0	1.4E-01	8.6E-03
10	8.2E-03	5.5E-05
20	7.1E-03	5.3E-05
40	6.3E-03	5.4E-05
80	7.7E-03	5.4E-05

同様に単精度計算で求めた近似固有対のうちで固有値が指定した区間にあるものの数を表 11 に, それらについての相対残差の最大値を表 12 と図 25 に, 相対残差の対数グラフを図 26 と図 27 に示す。

表 9 近似固有対の数 (3 次式, $m = 800$, 倍精度計算)

ℓ	IT=1	IT=2
0	275	378
10	378	378
20	378	378
40	378	378
80	378	378
160	378	378
320	378	378

表 10 相対残差の最大値 (3 次式, $m = 800$, 倍精度計算)

ℓ	IT=1	IT=2
0	1.4E-01	8.5E-03
10	2.8E-02	4.7E-04
20	3.9E-03	7.1E-05
40	9.4E-05	1.5E-06
80	5.7E-08	9.1E-10
160	1.5E-11	1.2E-13
320	1.5E-11	1.2E-13

4. まとめ

連続体の問題を有限要素法で近似して導かれる一般固有値問題で指定した下端側区間に固有値を持つ固有対をフィ

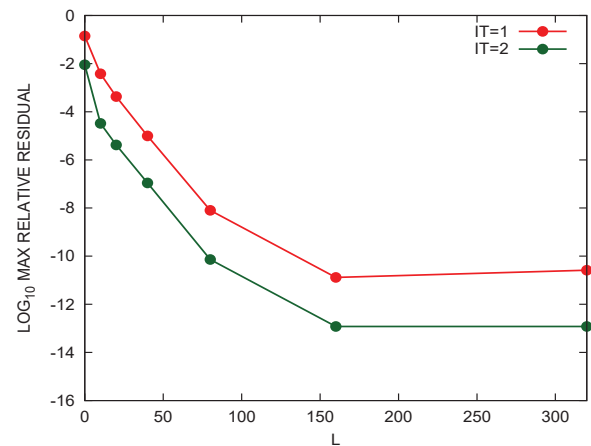


図 16 相対残差の最大値 (2 次式, $m = 800$, 倍精度計算)

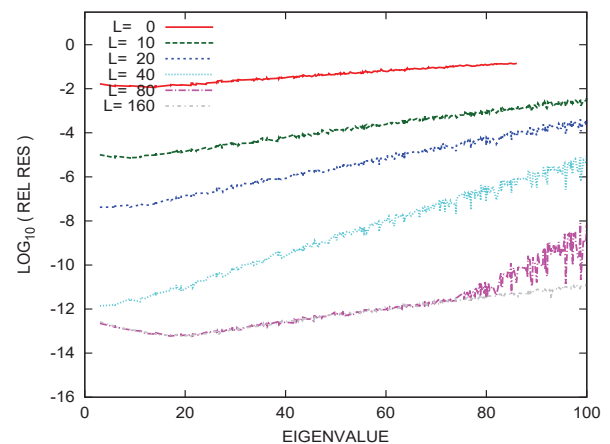


図 17 相対残差 (2 次式, $m = 800$, IT=1, 倍精度計算)

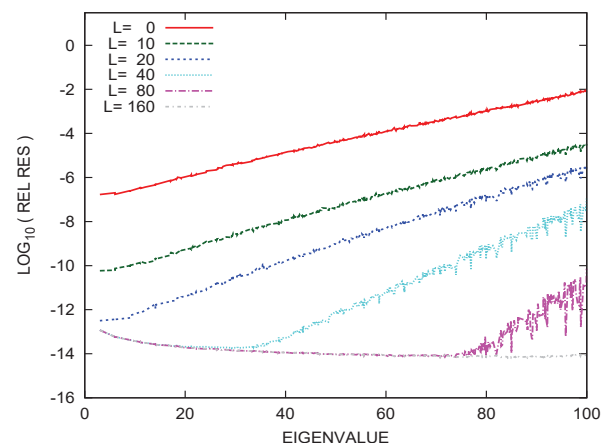


図 18 相対残差 (2 次式, $m = 800$, IT=2, 倍精度計算)

ルタ対角化法で近似して解く。フィルタはレゾルベントで構成されたものを用いることにする。必要な計算資源を抑えるために, 単一のレゾルベントから構成されたフィルタを採用するが, あまり良い特性にはならないのでフィルタを少数回 (今回 2 回まで) 反復して用いることにする。このように構成されたフィルタに最初に与えるベクトルの組

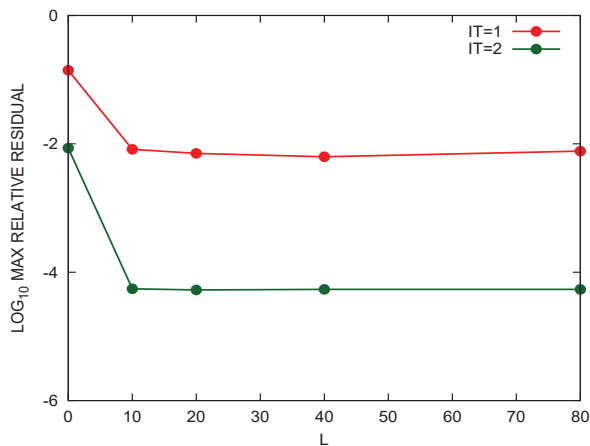


図 19 相対残差の最大値 (2 次式, $m = 800$, 単精度計算)

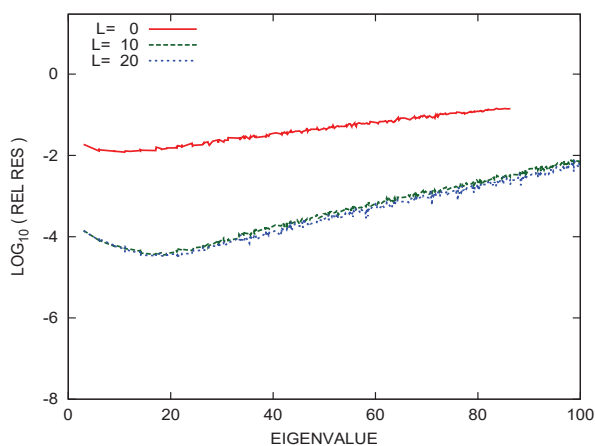


図 20 相対残差 (2 次式, $m = 800$, IT=1, 単精度計算)

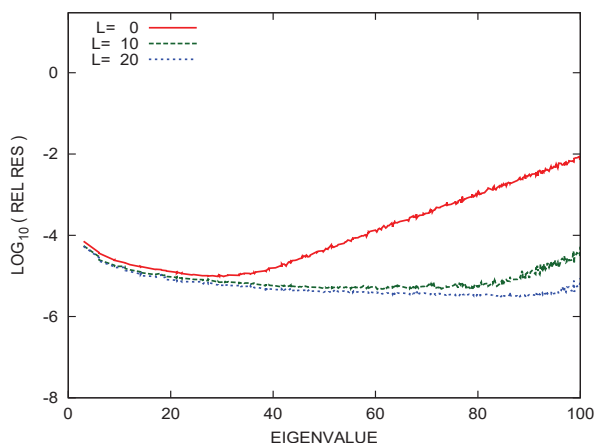


図 21 相対残差 (2 次式, $m = 800$, IT=1, 単精度計算)

として、従来は要素が一様乱数のベクトルの組を正規直交化したものを用いてきた。しかしそれを変更して、一様乱数を要素とするベクトルの組から「不要な固有ベクトル」の割合をなるべく減らして「必要な固有ベクトル」を含む割合を高めて、それを正規直交化してからフィルタに与えればそれにより得られる近似固有対の精度を良くすること

表 11 近似固有対の数 (3 次式, $m = 800$, 単精度計算)

ℓ	IT=1	IT=2
0	275	378
10	377	378
20	378	378
40	378	378
80	378	378

表 12 相対残差の最大値 (3 次式, $m = 800$, 単精度計算)

ℓ	IT=1	IT=2
0	1.5E-01	8.5E-03
10	2.3E-02	4.6E-04
20	7.7E-03	6.2E-05
40	8.4E-03	5.3E-05
80	7.1E-03	5.3E-05

ができる。今回は有限要素法から導かれた一般固有値問題の固有対で固有値が下端付近にあるものを求める場合に、フィルタに与える最初のベクトルの組から「不要な固有ベクトル」を含む割合を減らす手段として、同じ連続問題の差分近似で得られる標準固有値問題を考へて、その係数行列から導かれる Jacobi 反復法の走査の作用の多項式であるフィルタを用いた。そうしてその低次の多項式フィルタを反復適用することで「不要な固有ベクトル」を含む割合を減らして作ったベクトルの組を (レゾルベントの多項式型である) フィルタに対して与えることにより、最終的に得られる一般固有値問題の近似固有対の精度を高めることができた。今回扱った例題は一応簡単な 3 次元ラプラス固有値問題であり、また小規模なものではあるが、実験結果は満足のものであった。

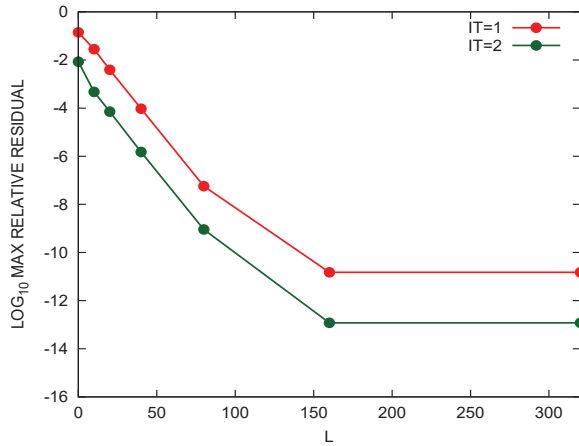


図 22 相対残差の最大値 (3 次式, $m = 800$, 倍精度計算)

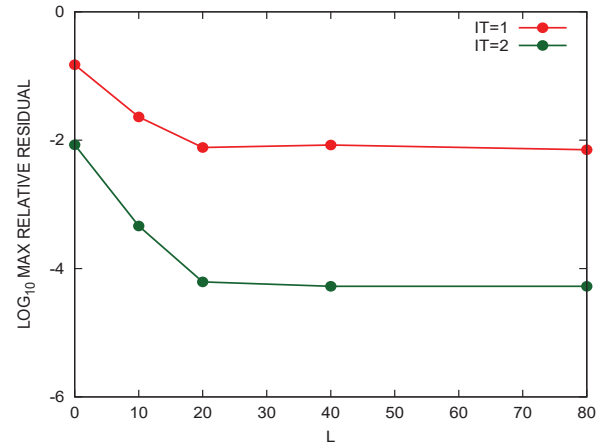


図 25 相対残差の最大値 (3 次式, $m = 800$, 単精度計算)

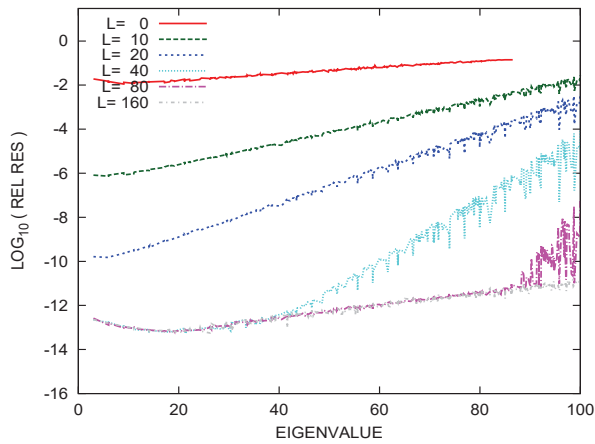


図 23 相対残差 (3 次式, $m = 800$, IT=1, 倍精度計算)

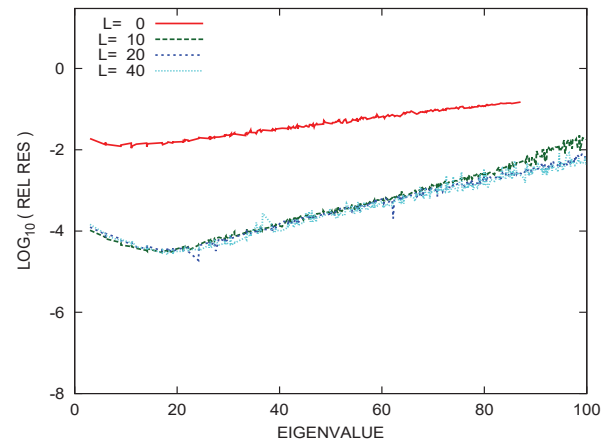


図 26 相対残差 (3 次式, $m = 800$, IT=1, 単精度計算)

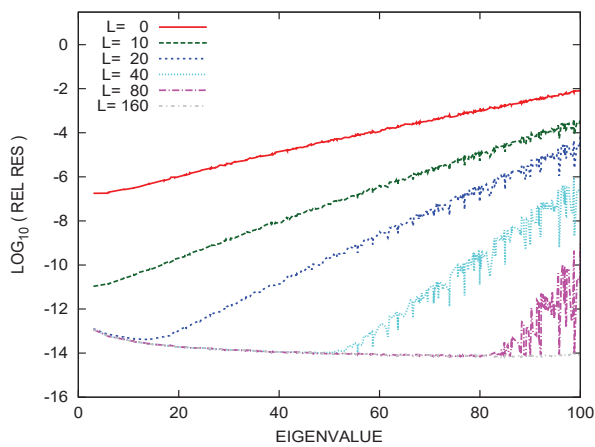


図 24 相対残差 (3 次式, $m = 800$, IT=2, 倍精度計算)

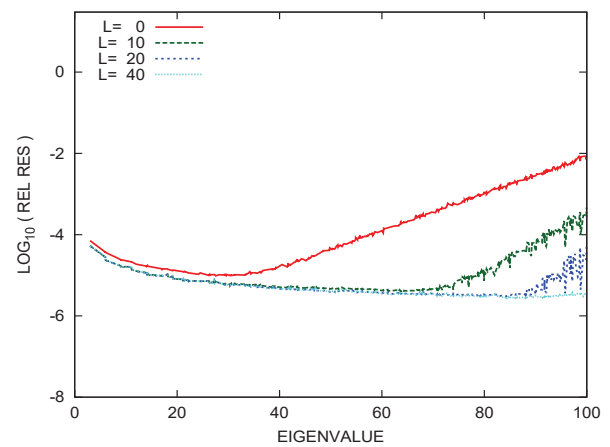


図 27 相対残差 (3 次式, $m = 800$, IT=2, 単精度計算)

付 録

A.1 標準固有値問題の中間固有値を求めた例

1 辺の長さが π の立方体の零固定境界条件を課したラプラス固有値問題を差分法により各座標方向の格子点数を $(N_1, N_2, N_3) = (20, 30, 40)$ に設定して得られる行列 A の標準固有値問題の中間固有値を持つ近似固有対をフィルタを用いて解く例を示す。この場合のフィルタは虚数シフトの単一のレゾルベントの Chebyshev 多項式であり、前回の [2] で用いたものと同じで、そのパラメタは次数 $n = 4$, $\mu = 1.5$, $g_{\text{stop}} = 1\text{E-}5$ である。このフィルタを 2 回まで用いて標準固有値問題の区間 $[a, b] = [100, 150]$ に固有値を持つ固有対を近似して求めた（その正しい数は 494 個である）。

レゾルベントからなるフィルタに最初に与えるベクトルの組を、要素が一様乱数からなる m 個のベクトルの組に対して以下の 2 次式および 4 次式と対応する Jacobi 走査の多項式フィルタを ℓ 回繰り返して適用し、その結果を正規直交化することで作成した。

A.1.1 2 次式を用いた場合の例

2 次式 $g_2(z)$ として式 (A.1) のものを用いた。

$$g_2(z) = 1 - z^2/3 = (1 - z/\sqrt{3})(1 + z/\sqrt{3}). \quad (\text{A.1})$$

これは実根のみを持ち偶関数である。この多項式 $g_2(z)$ の値をプロットしたグラフ（対称性から右半分だけ）を図 A.1 に示す。さらにシフト s を区間 $[a, b]$ の中点 125 に選んだ場合に、横軸の範囲を行列 A のすべての固有値に対応するものにとり、縦軸に $g_2(z)$ に対応する操作を反復した場合の各固有ベクトルに対する倍率をとってグラフにプロットしたものを図 A.2 に示す。また縦軸の倍率を対数にしてプロットしたものを図 A.3 に示す。固有対の固有値の区間 $[a, b] = [100, 150]$ に対応する z の区間は $|z| \leq 0.05003$ である。

ベクトルの数 m を 800, 1000, 1500, 2000 とした各場合について、フィルタに与える m 個のベクトルの組を作るための多項式フィルタの適用回数 ℓ が 0, 10, 20, 40 などとしたときに、最終的に得られた近似固有対のうちで指定された区間に固有値があるものについての相対残差の最大の値を表に示す（倍精度計算の場合が表 A.1, 表 A.2, 表 A.3, 表 A.4 であり、単精度計算の場合が表 A.5, 表 A.6, 表 A.7, 表 A.8 である）。また、ベクトルの数 m を 800, 1000, 1500, 2000 としたそれぞれの場合について、回数 ℓ に対する相対残差の最大値をグラフにプロットした（倍精度計算の場合が図 A.4 であり、単精度計算の場合が図 A.7 である）。

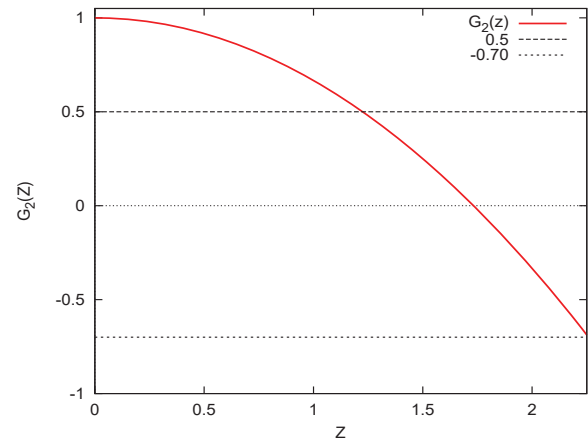


図 A.1 $g_2(z)$ (2 次式, 右半分)

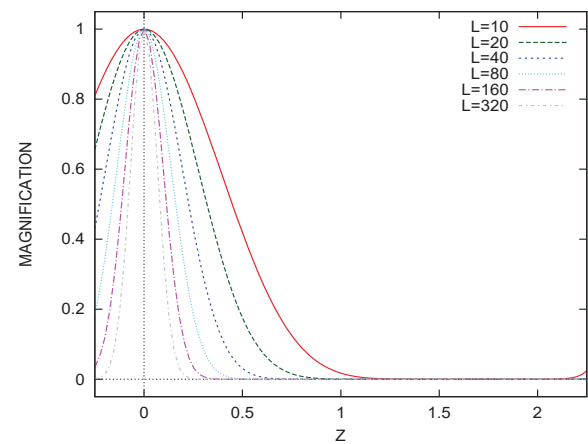


図 A.2 2 次式 $g_2(z)$ と対応する反復による固有ベクトルへの倍率

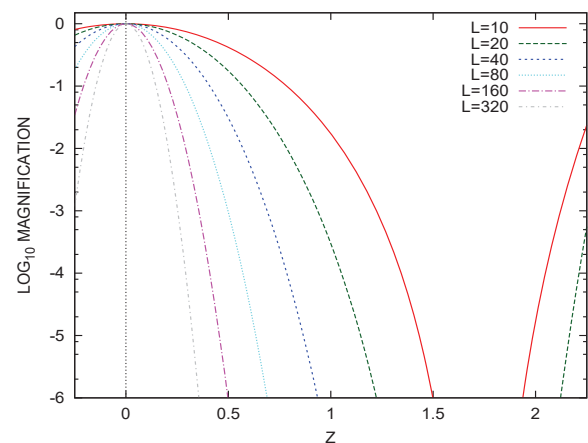


図 A.3 2 次式 $g_2(z)$ と対応する反復による固有ベクトルへの倍率（片対数プロット）

回数 ℓ を変えたときの各近似固有対の固有値に対する相対残差を $m = 1000$ の場合だけについて IT=1 と IT=2 の場合にそれぞれグラフにプロットした（倍精度計算の場合が図 A.5 と図 A.6 であり、単精度計算の場合が図 A.8 と図 A.9 である）。

表 A-1 相対残差の最大値 (2 次式, $m = 800$, 倍精度計算)

ℓ	IT=1	IT=2
0	1.9E-01	4.7E-04
10	9.1E-02	3.7E-05
20	5.6E-02	2.8E-05
40	5.4E-02	2.7E-05
80	3.2E-02	1.6E-05
160	2.5E-02	8.6E-06
320	1.1E-02	4.5E-06
640	4.2E-03	2.2E-06
1280	1.1E-03	5.4E-07

表 A-5 相対残差の最大値 (2 次式, $m = 800$, 単精度計算)

ℓ	IT=1	IT=2
0	1.8E-01	5.3E-04
10	9.8E-02	5.1E-05
20	8.0E-02	4.6E-05
40	6.5E-02	2.5E-05
80	4.0E-02	1.9E-05
160	2.0E-02	1.1E-05
320	1.1E-02	7.6E-06
640	5.8E-03	6.0E-06
1280	2.2E-03	5.7E-06

表 A-2 相対残差の最大値 (2 次式, $m = 1000$, 倍精度計算)

ℓ	IT=1	IT=2
0	1.8E-01	3.2E-04
10	3.5E-02	1.6E-05
20	2.5E-02	1.3E-05
40	1.6E-02	1.0E-05
80	9.9E-03	5.9E-06
160	4.9E-03	3.1E-06
320	2.1E-03	1.3E-06
640	5.6E-04	5.1E-07
1280	6.1E-05	8.3E-08

表 A-6 相対残差の最大値 (2 次式, $m = 1000$, 単精度計算)

ℓ	IT=1	IT=2
0	2.0E-01	3.0E-04
10	3.6E-02	1.8E-05
20	2.9E-02	1.4E-05
40	2.1E-02	1.1E-05
80	9.3E-03	8.3E-06
160	6.0E-03	6.5E-06
320	2.7E-03	5.9E-06
640	2.3E-03	5.6E-06
1280	2.2E-03	5.7E-06

表 A-3 相対残差の最大値 (2 次式, $m = 1500$, 倍精度計算)

ℓ	IT=1	IT=2
0	1.7E-01	1.7E-04
10	8.9E-03	8.0E-06
20	6.4E-03	5.7E-06
40	3.5E-03	4.5E-06
80	2.1E-03	2.7E-06
160	8.1E-04	1.1E-06
320	2.1E-04	2.5E-07
640	1.7E-05	1.7E-08
1280	2.1E-07	1.3E-10

表 A-7 相対残差の最大値 (2 次式, $m = 1500$, 単精度計算)

ℓ	IT=1	IT=2
0	8.7E-02	1.8E-04
10	9.0E-03	1.0E-05
20	6.2E-03	8.4E-06
40	3.9E-03	7.6E-06
80	2.9E-03	6.1E-06
160	2.2E-03	5.6E-06
320	1.6E-03	5.6E-06
640	1.7E-03	5.2E-06
1280	1.7E-03	5.2E-06

表 A-4 相対残差の最大値 (2 次式, $m = 2000$, 倍精度計算)

ℓ	IT=1	IT=2
0	6.6E-02	1.4E-04
10	5.1E-03	5.6E-06
20	3.0E-03	3.6E-06
40	1.8E-03	2.6E-06
80	9.6E-04	1.5E-06
160	3.7E-04	6.0E-07
320	6.5E-05	9.2E-08
640	1.6E-06	3.1E-09
1280	9.0E-10	2.4E-12

表 A-8 相対残差の最大値 (2 次式, $m = 2000$, 単精度計算)

ℓ	IT=1	IT=2
0	6.7E-02	1.4E-04
10	5.2E-03	8.5E-06
20	3.8E-03	6.9E-06
40	2.7E-03	6.3E-06
80	2.1E-03	5.8E-06
160	1.7E-03	5.7E-06
320	1.6E-03	5.2E-06
640	1.4E-03	5.2E-06
1280	1.8E-03	5.1E-06

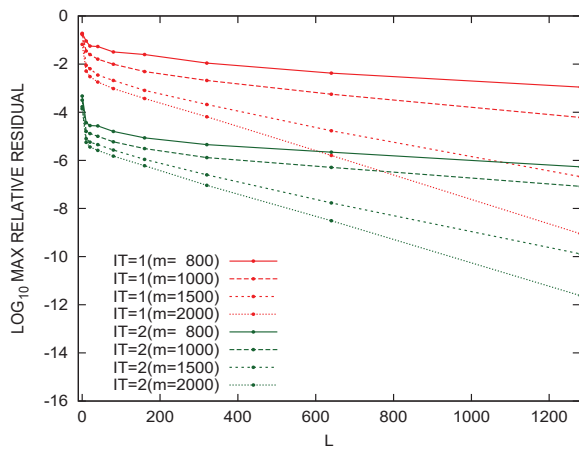


図 A-4 相対残差の最大値 (2 次式, 倍精度計算)

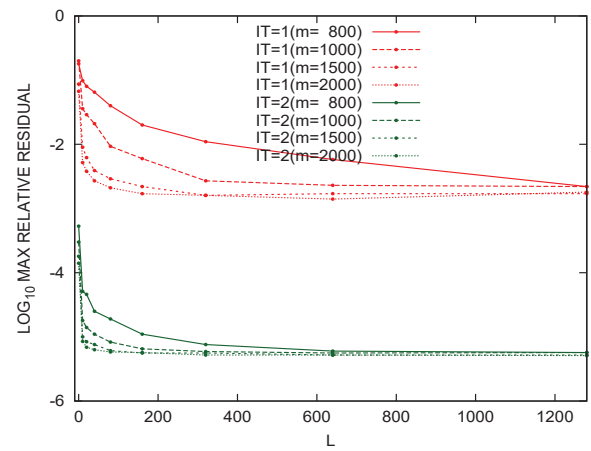


図 A-7 相対残差の最大値 (2 次式, 単精度計算)

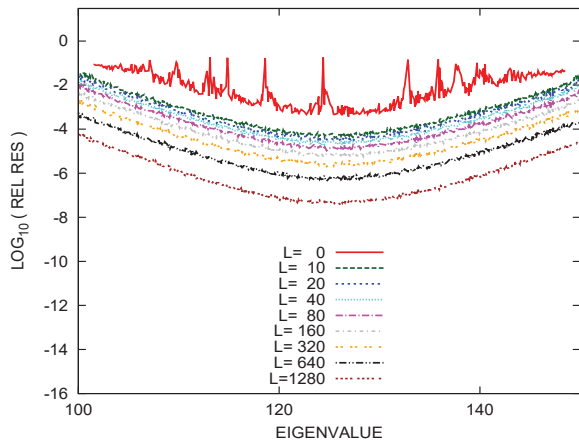


図 A-5 相対残差 (2 次式, $m = 1000$, IT=1, 倍精度計算)

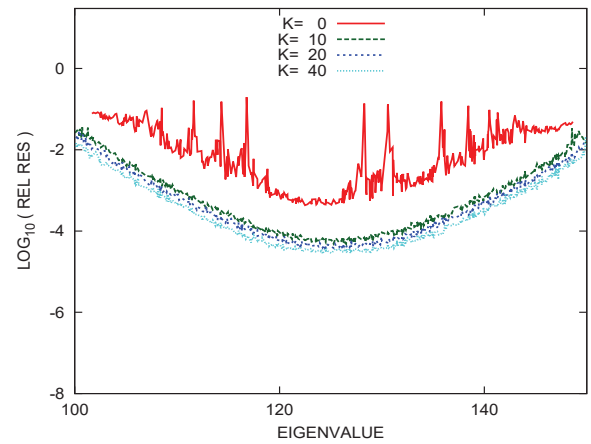


図 A-8 相対残差 (2 次式, $m = 1000$, IT=1, 単精度計算)

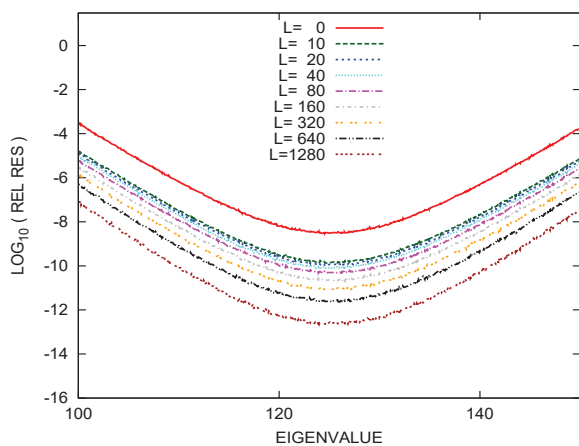


図 A-6 相対残差 (2 次式, $m = 1000$, IT=2, 倍精度計算)

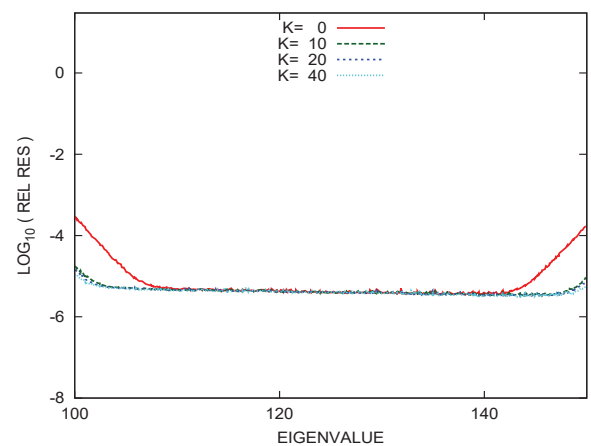


図 A-9 相対残差 (2 次式, $m = 1000$, IT=2, 単精度計算)

A.1.2 4次式を用いた場合の例

4次多項式 $g_4(z)$ として式 (A.2) のものを用いた.

$$\begin{cases} g_4(z) = (1 - r_1 z)(1 - r_2 z)(1 - r_3 z)(1 - r_4 z), \\ r_1 = 0.6805206466816319, \\ r_2 = -r_1, \\ r_3 = 0.4205848896953068, \\ r_4 = -r_3. \end{cases} \quad (\text{A.2})$$

これは実根のみを持ち偶関数である. この多項式 $g_4(z)$ の値を (対称性から右半分だけを) プロットしたグラフを 図 A.10 に示す. さらにシフト s を区間 $[a, b]$ の中点 125 に選んだ場合に, 横軸の範囲を行列 A のすべての固有値に対応するものにとり, 縦軸に $g_4(z)$ に対応する操作を反復した場合の固有ベクトルに対する倍率をとってグラフにプロットしたものを 図 A.11 に示す. また縦軸の倍率を対数にしてプロットしたものを 図 A.12 に示す. 固有対の固有値の区間 $[a, b] = [100, 150]$ に対応する z の区間は $|z| \leq 0.05003$ である.

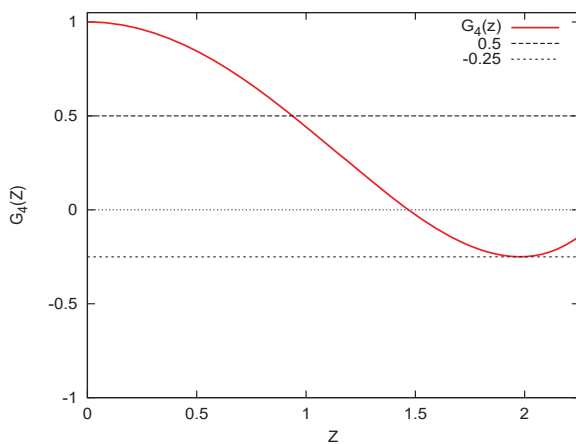


図 A.10 $g_4(z)$ (4次式, 右半分)

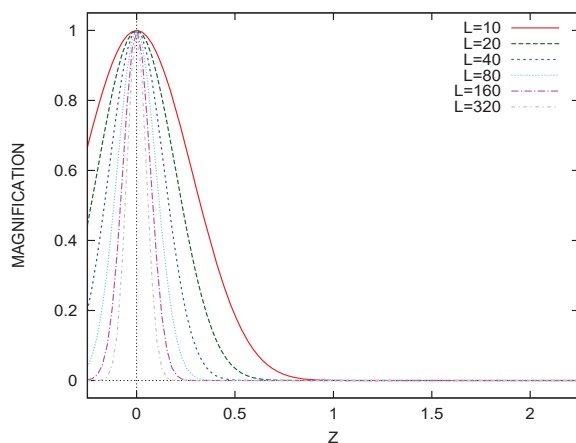


図 A.11 4次式 $g_4(z)$ と対応する反復による固有ベクトルへの倍率

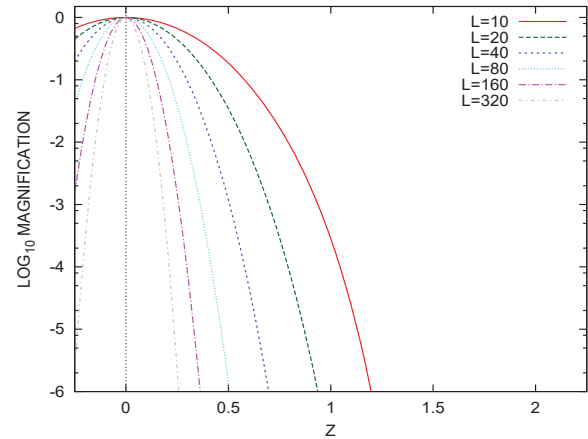


図 A.12 4次式 $g_4(z)$ と対応する反復による固有ベクトルへの倍率 (片対数プロット)

ベクトルの数 m を 800, 1000, 1500, 2000 としたそれぞれの場合について, フィルタに与える m 個のベクトルの組を作るための多項式フィルタの適用回数 ℓ を 0, 10, 20, 40 などとしたときに, 最終的に得られた近似固有対のうちで指定された区間に固有値があるものについての相対残差の最大の値を表に示す (倍精度計算の場合が表 A.9, 表 A.10, 表 A.11, 表 A.12 であり, 単精度計算の場合が表 A.13, 表 A.14, 表 A.15, 表 A.16 である). また, ベクトルの数 m を 800, 1000, 1500, 2000 とした各場合について, 回数 ℓ に対する相対残差の最大値をグラフにプロットした (倍精度計算の場合が図 A.13 であり, 単精度計算の場合が図 A.16 である).

回数 ℓ を変えたときの各近似固有対の固有値に対する相対残差を $m = 1000$ の場合だけについて IT=1 と IT=2 の場合にそれぞれグラフにプロットした (倍精度計算の場合が図 A.14 と図 A.15 であり, 単精度計算の場合が図 A.17 と図 A.18 である).

参考文献

- [1] Haw-Ren Fang and Yousef Saad : "A filtered Lanczos procedure for extreme and interior eigenvalue problems", *SIAM J. Sci. Comput.*, vol.34, No.4 (2012), pp.A2220–A2246.
- [2] 村上弘 : 「フィルタ対角化法に与える初期ベクトルの改良の模索」, 情報処理学会研究報告, vol.2021-HPC-178, no.1 (2021), pp.1–17.

表 A-9 相対残差の最大値 (4 次式, $m = 800$, 倍精度計算)

ℓ	IT=1	IT=2
0	1.7E-01	5.2E-04
10	5.1E-02	2.7E-05
20	4.0E-02	2.4E-05
40	3.8E-02	1.7E-05
80	2.5E-02	1.1E-05
160	1.1E-02	4.9E-06
320	6.4E-03	1.9E-06
640	1.1E-03	6.3E-07
1280	1.1E-04	9.6E-08

表 A-13 相対残差の最大値 (4 次式, $m = 800$, 単精度計算)

ℓ	IT=1	IT=2
0	1.8E-01	5.0E-04
10	6.4E-02	3.8E-05
20	8.3E-02	2.9E-05
40	3.5E-02	1.8E-05
80	2.2E-02	1.4E-05
160	1.3E-02	7.7E-06
320	5.3E-03	5.9E-06
640	2.9E-03	5.6E-06
1280	2.6E-03	5.7E-06

表 A-10 相対残差の最大値 (4 次式, $m = 1000$, 倍精度計算)

ℓ	IT=1	IT=2
0	2.1E-01	2.8E-04
10	2.6E-02	1.2E-05
20	1.8E-02	9.5E-06
40	1.1E-02	6.3E-06
80	5.3E-03	2.8E-06
160	2.1E-03	1.4E-06
320	4.8E-04	5.6E-07
640	8.0E-05	9.8E-08
1280	3.0E-06	3.8E-09

表 A-14 相対残差の最大値 (4 次式, $m = 1000$, 単精度計算)

ℓ	IT=1	IT=2
0	2.0E-01	3.1E-04
10	2.4E-02	1.5E-05
20	2.0E-02	1.1E-05
40	1.1E-02	8.5E-06
80	5.1E-03	6.3E-06
160	2.9E-03	5.8E-06
320	2.0E-03	5.5E-06
640	1.8E-03	5.5E-06
1280	2.3E-03	5.2E-06

表 A-11 相対残差の最大値 (4 次式, $m = 1500$, 倍精度計算)

ℓ	IT=1	IT=2
0	1.9E-01	1.7E-04
10	6.5E-03	6.1E-06
20	4.3E-03	4.3E-06
40	2.1E-03	3.0E-06
80	7.7E-04	1.3E-06
160	2.3E-04	3.0E-07
320	2.2E-05	2.3E-08
640	2.6E-07	1.7E-10
1280	7.0E-11	3.2E-14

表 A-15 相対残差の最大値 (4 次式, $m = 1500$, 単精度計算)

ℓ	IT=1	IT=2
0	1.1E-01	1.8E-04
10	6.4E-03	8.6E-06
20	4.3E-03	7.5E-06
40	3.0E-03	6.4E-06
80	2.2E-03	5.6E-06
160	1.9E-03	5.7E-06
320	1.8E-03	5.2E-06
640	1.5E-03	5.2E-06
1280	1.8E-03	5.1E-06

表 A-12 相対残差の最大値 (4 次式, $m = 2000$, 倍精度計算)

ℓ	IT=1	IT=2
0	6.4E-02	1.3E-04
10	2.8E-03	3.8E-06
20	1.9E-03	2.7E-06
40	1.0E-03	1.6E-06
80	4.2E-04	6.3E-07
160	7.1E-05	1.1E-07
320	2.5E-06	4.3E-09
640	1.6E-09	4.1E-12
1280	3.1E-12	9.7E-15

表 A-16 相対残差の最大値 (4 次式, $m = 2000$, 単精度計算)

ℓ	IT=1	IT=2
0	6.4E-02	1.4E-04
10	3.8E-03	6.9E-06
20	3.5E-03	6.3E-06
40	2.2E-03	5.8E-06
80	1.9E-03	5.7E-06
160	1.4E-03	5.2E-06
320	1.5E-03	5.1E-06
640	1.7E-03	5.1E-06
1280	1.7E-03	5.1E-06

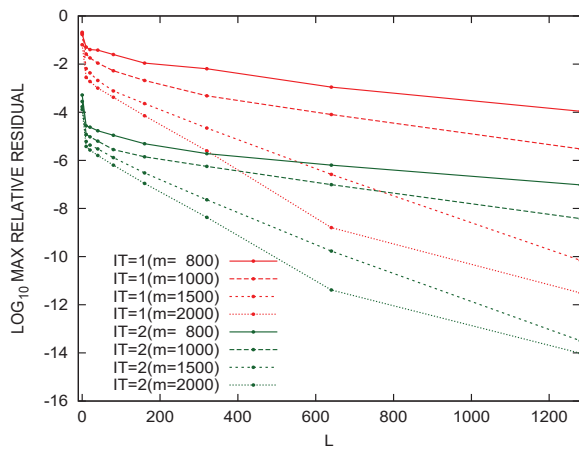


図 A-13 相対残差の最大値 (4 次式, 倍精度計算)

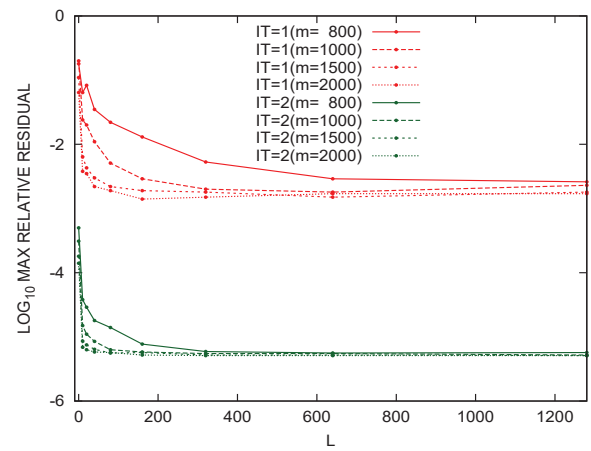


図 A-16 相対残差の最大値 (4 次式, 単精度計算)

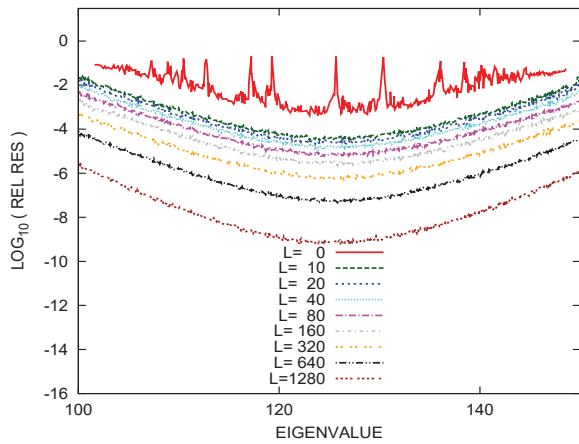


図 A-14 相対残差 (4 次式, $m = 1000$, IT=1, 倍精度計算)

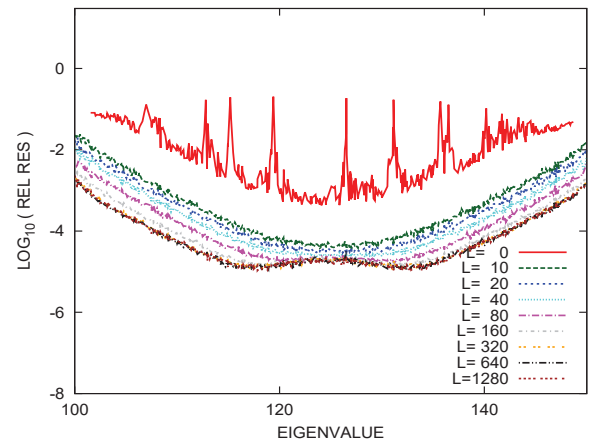


図 A-17 相対残差 (4 次式, $m = 1000$, IT=1, 単精度計算)

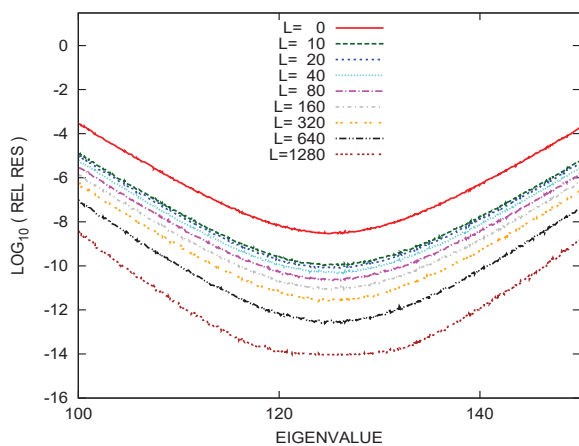


図 A-15 相対残差 (4 次式, $m = 1000$, IT=2, 倍精度計算)

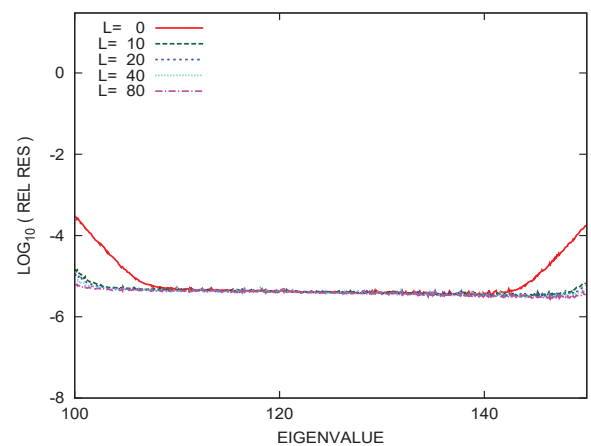


図 A-18 相対残差 (4 次式, $m = 1000$, IT=2, 単精度計算)