

固有値程式プログラムの信頼性テスト方式

阪大．大型センター 磯本征雄

後藤米子

1. 序文

固有値程式プログラムがライブラリとして登録されるようになり、数値計算に未熟な者にも容易に固有値・固有ベクトルを得ることができるようになった。その反面、アルゴリズムの詳細を熟知しない者にもプログラムを管理し、利用できるようなプログラムの信頼性を確認するテスト方式が必要になる。

あるライブラリ・プログラムの利用者にとって、プログラムが信頼できるとは、望みの精度で解が得られることであり、また精度が望みのものより悪い場合でも必ずその事実を知り得ることであろう。我々の研究目的は、アルゴリズムの詳細に立入ることなくプログラムの信頼性をテストする方式を明らかにすることである。テストは、プログラムが正しく働くこと及び数値解誤差の振舞いに異常がないことを主

な着目点としておこなう。但し，行列は実対称に限定する。

2. 記号の定義

$n \times n$ 次元行列 A に対して，真の固有ベクトルと固有値 $(\bar{x}_i, \lambda_i)$, $i=1, 2, \dots, n$ は既知であるとする。

$$A\bar{x}_i = \lambda_i \bar{x}_i. \quad (2.1)$$

但し，ここでは $|\bar{x}_i| = 1$ と規格化するものと仮定する。

$(\bar{x}_i, \lambda_i)$ に対応する，被テストプログラムにより出力される数値解を $(\hat{x}_i, \hat{\lambda}_i)$ とする。 $(\hat{x}_i, \hat{\lambda}_i)$ を n 組の $(\bar{x}_i, \lambda_i)$ を使って次のように展開する。

$$\hat{\lambda}_i = \lambda_i + \delta\lambda_i. \quad (2.2)$$

$$\begin{aligned} \hat{x}_i &= (1 - \delta_{ii}^2) \bar{x}_i + \delta_{ii}^2 \sum_{j \neq i} \alpha_{ij} \bar{x}_j, \\ &= \bar{x}_i + \delta\bar{x}_i, \quad i=1, 2, \dots, n. \end{aligned} \quad (2.3)$$

$\delta\lambda_i$ は λ_i の誤差である。 $\delta\bar{x}_i$ は $\bar{x}_i$ の誤差ベクトルである。 δ_{ii}^2 は $\delta\bar{x}_i$ の内のベクトル $\bar{x}_i$ に平行な成分の占める割合である。 $\delta_{\perp i}^2$ は $\delta\bar{x}_i$ の内のベクトル $\bar{x}_i$ に垂直な成分の占める割合である。また， α_{ij} は $\delta_{\perp i}^2$ の内のベクトル $\bar{x}_j$ ($j \neq i$) の成分の占める割合である。

テスト行列 A と固有値・固有ベクトルの全体との関係を見やすくするために，行列 Λ と X を導入する。 Λ は，固有値

$\lambda_i (i=1, 2, \dots, n)$ を対角要素とする対角行列である：

$$\Lambda = \begin{bmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ 0 & & \dots \\ & & & \lambda_n \end{bmatrix}. \quad (2.4)$$

X は、固有ベクトル (列ベクトル) を要素とする行列である：

$$X = [\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n]. \quad (2.5)$$

X 及び Λ は次式を満たす： $AX = X\Lambda$ (2.6)

テスト行列 A は、式 (2.6) より $A = X\Lambda X^T$ により生成される。

式 (2.2) と (2.3) におけるパラメター $\delta\lambda_i, |\delta\vec{x}_i|, \delta_{\parallel}^i, \delta_{\perp}^i$ 及び α_{ij} は、 $(\vec{x}_i, \lambda_i)$ 及び $(\vec{x}_i, \lambda'_i)$ により次のようにして与えられる。まず式 (2.2) と (2.3) より、次の関係を得る：

$$\delta\lambda_i = \lambda'_i - \lambda_i, \quad (2.7)$$

$$|\delta\vec{x}_i| = \vec{x}'_i - \vec{x}_i. \quad (2.8)$$

式 (2.3) のベクトルの内積より、次式を得る：

$$\langle \vec{x}'_i \cdot \vec{x}_i \rangle = (1 - \delta_{\parallel}^i)^2 + \delta_{\perp}^{i2} \sum_{j \neq i} \alpha_{ij}^2. \quad (2.9)$$

ここで、ベクトル $\sum_{j \neq i} \alpha_{ij} \vec{x}_j$ を 1 に規格化することにより、次式を得る：

$$\sum_{j \neq i} \alpha_{ij}^2 = 1 \quad (2.10)$$

一方、式 (2.8) より、次の関係を得る：

$$|\delta\vec{x}_i|^2 = \delta_{\parallel}^{i2} + \delta_{\perp}^{i2}. \quad (2.11)$$

式 (2.9), (2.10) 及び (2.11) より、次の関係を得る：

$$\delta_{\parallel}^i = \frac{1}{2} [1 + |\delta\vec{x}_i|^2 - \langle \vec{x}'_i \cdot \vec{x}_i \rangle], \quad (2.12)$$

$$\delta_{\perp}^i = \sqrt{|\delta \vec{x}_i|^2 - \delta_{\parallel}^i{}^2}, \quad (2.13)$$

$$\alpha_{ij} = \langle \vec{x}_j \cdot \vec{x}_i \rangle / \delta_{\perp}^i. \quad (2.14)$$

3. 固有方程式数値解の誤差

固有方程式プログラムの信頼性テストの準備として，数値解誤差の定性的特徴を調べておく．

残差ベクトルの大きさ

$(\vec{x}_i, \lambda_i)$ の行列 A に対する残差ベクトル $\vec{r}_i$ は次式で与えられる．

$$\vec{r}_i = A \vec{x}_i - \lambda_i \vec{x}_i. \quad (3.1)$$

数値解法の過程において，プログラム中では A , $\vec{x}_i$ 及び λ_i は既知である．ここで，計算機により実現可能な最小の相対誤差をととする．この時，相対誤差 ε の程度で式 (3.1) 右辺をゼロに近づけることは可能であろう．行列 A の絶対値最大の固有値を $\lambda_{\max}$ とする．ここで， $\|A\|_2 \cong O(|\lambda_{\max}|)$ ， $|\vec{x}_i| \cong 1 + o(\varepsilon)$ である．したがって $|A \vec{x}_i|$ の大きさは， $O(|\lambda_{\max}|)$ である．以上のことより，式 (3.1) の左辺については，次の近似式を仮定する．

$$|\vec{r}_i| \cong o(\varepsilon \cdot |\lambda_{\max}|). \quad (3.2)$$

このように，残差ベクトルは各々の $(\vec{x}_i, \lambda_i)$ に無関係に，ほぼ一定であろうことが予想される¹⁾．

固有値の分布と誤差

固有値の分布状況は固有値・固有ベクトルの誤差に大きく影響する。²⁾ 以下、このことを調べる。式(3.1)の左より $\vec{x}_j$ を掛けて次式を得る。

$$\langle \vec{x}_j, \vec{\eta}_i \rangle = (\lambda_j - \lambda_i) \langle \vec{x}_j, \vec{x}_i \rangle. \quad (3.3)$$

式(3.3)において $j=i$ とすれば、次式を得る。

$$|\delta \lambda_i| = |\langle \vec{x}_i, \vec{\eta}_i \rangle / \langle \vec{x}_i, \vec{x}_i \rangle|.$$

$\langle \vec{x}_i, \vec{x}_i \rangle = 1 + o(\varepsilon)$ であり、 $\langle \vec{x}_i, \vec{\eta}_i \rangle = o(\varepsilon \cdot |\lambda_{\max}|)$ であることから、次の結果を得る。

$$|\delta \lambda_i| = o(\varepsilon \cdot |\lambda_{\max}|). \quad (3.4)$$

式(3.3)に対して式(2.3)を代入して、次式を得る。

$$\delta \lambda_i^2 = \sum_{j \neq i} \langle \vec{x}_j, \vec{\eta}_i \rangle^2 / (\lambda_j - \lambda_i)^2,$$

$$\delta \lambda_i^2 = [1 - \langle \vec{x}_i, \vec{\eta}_i \rangle / (\lambda_i - \lambda_i)]^2.$$

この2式と式(2.11)より、 $|\delta \vec{x}_i|$ に対して次の関係を得る。

$$|\delta \vec{x}_i|^2 = \left[1 - \frac{\langle \vec{x}_i, \vec{\eta}_i \rangle}{(\lambda_i - \lambda_i)} \right]^2 + \sum_{j \neq i} \frac{\langle \vec{x}_j, \vec{\eta}_i \rangle^2}{(\lambda_j - \lambda_i)^2}. \quad (3.5)$$

式(3.2)及び式(3.4)より $|\langle \vec{x}_i, \vec{\eta}_i \rangle / (\lambda_i - \lambda_i)| = 1 + o(\varepsilon)$ 。故に式(3.5)のオ1項は一般に $o(\varepsilon)$ の大きさである。但し、固有値の誤差 $\delta \lambda_i$ が異常に大きくなれば $|\delta \vec{x}_i|$ も大きくなるであろう。オ2式は、 λ_i に近接する固有値が存在する時、 $|\delta \vec{x}_i|^2$ が大きくなることを示している。²⁾

残差ベクトルと丸め誤差の累積

以上の議論では、行列 A は初期状態のままであり、数値計算の過程で入る丸め誤差の累積効果は無視されている。ところが現実には、数値解を計算する過程で行列 A は何回も変換される。そしてその度に丸め誤差が累積される。この結果、数値解 $(\bar{x}_i, \lambda_i)$ が計算される時点での行列は、すでに誤差 δA_i だけ加わった行列 A'_i になっている：

$$A'_i = A + \delta A_i. \quad (3.6)$$

δA_i を丸め誤差累積効果の大きさを調べるための手掛りとして使う。

A'_i は、次の関係を満たす。

$$A'_i \bar{x}_i = \lambda_i \bar{x}_i. \quad (3.7)$$

式 (3.7) より、式 (3.6) を用いて次式を得る。

$$\delta A_i \bar{x}_i = (\lambda_i - A) \bar{x}_i. \quad (3.8)$$

ここで、 $-\delta A_i \bar{x}_i$ は残差ベクトルである。式 (3.8) をベクトル $\bar{x}_i$ と $\sum_{j=1}^n \alpha_{ij} \bar{x}_j$ の成分に分解して、次の関係式を得る。

$$E_i = \langle \bar{x}_i, \delta A_i \bar{x}_i \rangle = (\lambda_i - \lambda_i)(1 - \delta_i^2). \quad (3.9)$$

$$E_i = \langle \sum_{j=1}^n \alpha_{ij} \bar{x}_j, \delta A_i \bar{x}_i \rangle = \delta_i^2 \sum_{j=1}^n \alpha_{ij}^2 (\lambda_i - \lambda_j). \quad (3.10)$$

式 (3.9) 及び (3.10) は、行列 A の誤差の大きさを示す。但し、これらは演算の繰り返しのための丸め誤差累積効果のほか、入力時点での A の誤差も含んでいる。

4. アルゴリズムに関するテスト方式

アルゴリズムのテストでは、プログラムの総ての部分を駆使する必要がある。すなわち、テスト行列 A としては、様々な異なる性質を備えた行列を使わなければならない。ところが行列 A の性質は、その固有ベクトル X 及び固有値 λ の面から見ることができる。

さて、固有値 λ に注目する時、その性質を次の場合に分類できる。

- 1). λ は最小固有値である。
- 2). λ は他の固有値と縮退している。
- 3). λ は最大固有値でも最小固有値でもない。
- 4). λ は最大固有値である。

(4.1)

1)~4) のいずれも、3次元行列 A により実現でき、したがって固有値の様々な性質に関するテストは3次元空間内にとどめである。固有ベクトルは、正規直交性以外に制約条件は考えられない。したがって固有ベクトルの様々な性質に関しても、3次元空間以上の次元でテストする積極的理由はない。故に我々は3次元行列 A をテスト行列とする。

実対称行列 A の固有値・固有ベクトルは実数及び実ベクトルである。すなわち λ は次式で与えられる。

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}. \quad (4.2)$$

固有ベクトルは回転行列³⁾を使って次式で与えられる。

$$X = \begin{bmatrix} \cos\theta \cos\varphi \cos\psi - \sin\varphi \sin\psi & \cos\theta \sin\varphi \cos\psi + \cos\varphi \sin\psi & -\sin\theta \cos\psi \\ -\cos\theta \cos\varphi \sin\psi - \sin\varphi \cos\psi & -\cos\theta \sin\varphi \sin\psi + \cos\varphi \cos\psi & \sin\theta \sin\psi \\ \sin\theta \cos\varphi & \sin\theta \sin\varphi & \cos\theta \end{bmatrix} \quad (4.3)$$

オイラー角 (θ, φ, ψ) を適当な制約条件の下で変えてゆくことにより，任意の X を機能的に生成できる。テスト行列 A は， Λ 及び X より次式で生成する： $A = X \Lambda X^T$. (4.4)

テストは，固有ベクトル・固有値の内の1つ $(\bar{x}_i, \lambda_i)$ に注目し， $(\bar{x}_i, \lambda_i)$ の変化とこれに対する誤差の大きさ $\delta_{\bar{x}}^i, \delta_{\lambda}^i$ ， $|\delta \bar{x}_i|, |\delta \lambda_i|$ 及び以下の量を測定することにより行なわれる。

$(\bar{x}_i, \lambda_i)$ は，次の近似式を満たす。但し，誤差は $o(\varepsilon | \lambda_{\max})$ 。

$$A \bar{x}_i \cong \lambda_i \bar{x}_i. \quad (4.5)$$

さて，ベクトル $A \bar{x}_i$ と $\lambda_i \bar{x}_i$ の方向余弦は次式で与えられる。

$$\begin{aligned} \cos \Omega &= \langle \bar{x}_i A \bar{x}_i \rangle / \sqrt{\langle \bar{x}_i \cdot \bar{x}_i \rangle \cdot \langle \bar{x}_i A^2 \bar{x}_i \rangle} \\ &= \frac{(1 - \delta_{\bar{x}}^i)^2 \lambda_i + \delta_{\bar{x}}^{i2} \sum_{j \neq i} \alpha_{ij}^2 \lambda_j}{\sqrt{[(1 - \delta_{\bar{x}}^i)^2 + \delta_{\bar{x}}^{i2}] [(1 - \delta_{\bar{x}}^i)^2 \lambda_i^2 + \delta_{\bar{x}}^{i2} \sum_{j \neq i} \alpha_{ij}^2 \lambda_j^2]}}. \end{aligned} \quad (4.6)$$

また，ベクトル $A \bar{x}_i$ と $\lambda_i \bar{x}_i$ の長さの差は次式で与えられる。

$$\begin{aligned} f_i &= |\langle \bar{x}_i A^2 \bar{x}_i \rangle^{1/2} - \langle \bar{x}_i \cdot \lambda_i^2 \bar{x}_i \rangle^{1/2}|, \\ &= |[(1 - \delta_{\bar{x}}^i)^2 \lambda_i^2 + \delta_{\bar{x}}^{i2} \sum_{j \neq i} \alpha_{ij}^2 \lambda_j^2]^{1/2} - |\lambda_i| [(1 - \delta_{\bar{x}}^i)^2 + \delta_{\bar{x}}^{i2}]^{1/2}|. \end{aligned} \quad (4.7)$$

もし，プログラムにおける処理が理想的に速められているならば，式(3.2)の仮定から我々は次の関係を期待できる。

$$\Omega \leq \varepsilon, \quad f_2 \leq \varepsilon \cdot |\lambda_{\max}|. \quad (4.8)$$

テスト結果の評価に当って，テスト行列 A の特性に依存する事項とプログラムの特性に依存する事項を区別しなければならない． $\delta_1, \delta_2, |\delta x_1|, |\delta \lambda_1|$ 等は，行列 A の性質に依存する．上記議論に見るように Ω 及び f_2 は行列 A の性質に直接関係しない，プログラムの性質に依存する量である．以下，具体例により説明する．

図1は，岡やコピ法によるプログラムを，浮動小数点実数有効桁数約10桁の計算機で調べたテスト結果である． λ_2 及び $\lambda_3 = 1.1$ を固定し， λ_1 を $0.1 \sim 1.0$ の範囲で変化させた時の $|\delta x_1|$, $|\delta \lambda_1|/|\lambda_1|$, f_1 , $\sqrt{1 - \cos \Omega}$ の変化の様子が示されている．横軸は λ_1 の値である． $\lambda_{\max} = 1.1$ であることより， $|\delta \lambda_1|/|\lambda_1|$ はほぼ 10^{-10} 程度の大きさになっている． $\lambda_1 \approx 0.9$ 及び ≈ 1.1 において $|\delta x_1|$ が大きくなっているのは，式(3.5)の2項による．一方 $|f_1|$ 及び $\sqrt{1 - \cos \Omega} \approx \Omega/2$ が， 10^{-10} 程度の大きさにおさえられている事より，プログラムが理想的に働いたことを示しているといえよう．

5. 丸め誤差累積効果に関するテスト方式

丸め誤差の累積効果は、計算の手間に由来する。したがって、このためのテストには或る程度大きな次元数の行列を使う必要がある。そして、アルゴリズムのテストのように多数の行列を系統的に生成しこれを解く方法は適さない。

誤差累積効果を見る一方法として、計算の手間の異なる、しかも同じ固有値・固有ベクトルをもつ2行列を解いて、その誤差 $\| \delta x_1 \|, \| \delta x_2 \|, E_1$ 及び E_2 の差異を比較する。この場合にも X 及び Λ の2面から見る。

Λ には、その中の固有値の分布の状態、大小の順序と Λ 内の位置との関係等により、いく通りかのものが考えられる。また、定数倍等の、本来数学的に重要でない操作が、はたしてプログラムにおいても敏感な影響を受けないか否かも調べる必要がある。

X は、行列 A を対角行列 Λ に変換するための変換行列としての意味をもつ。固有方程式解法の手間は、この変換行列 X を得るための繰り返し操作と考えて良いであろう。一方、 $A = X \Lambda X^T$ であり、もし X が単位行列なら A は対角行列となり、行列 A の数値解法上の手間はほとんどかからないであろう。 X が単位行列にくるべて「複雑」な構造をもつならば、その程度に応じて A も対角行列から遠ざかり計算の手間もかかるであろう。但し、「複雑」な構造及び計算の手

向の意味は、解法アルゴリズムにも依存するであろうから、必ずしも自明ではない。この点に関しては未解決である。

以下、具体的テスト例で説明する。テストは、図1と同じ図ヤコビ法のプログラムに対して行った。固有値は大きから小へ定数倍づつ減少する。 X は、単位行列から出発して、任意の2ベクトルのなす平面内の回転を次のベクトルの組合せ順序で次々と繰り返して得られた。

オ1の場合

$$\begin{aligned} & (1,2), (1,3), (1,4), \dots, (1,n-1), (1,n) \\ & (2,3), (2,4), \dots, (2,n-1), (2,n) \\ & \dots \\ & (n/2, n-1), (n/2, n) \end{aligned} \quad (5.1)$$

オ2の場合

$$\begin{aligned} & (n,n-1), (n,n-2), (n,n-3), \dots, (n,2), (n,1) \\ & (n-1,n-2), (n-1,n-3), \dots, (n-1,2), (n-1,1) \\ & \dots \\ & (\frac{n-2}{2}, 2), (\frac{n-2}{2}, 1) \end{aligned} \quad (5.2)$$

オ1とオ2の場合の違いは、行列 X への変換の手順が左上と、右下に集中している点である。

図2及び3には、式(5.1)と(5.2)の X に対するテスト結果を示す。横軸は $\log \lambda_i$ である。図2には各 λ_i ($i=1,2,\dots,50$)

に対する $18\lambda_i/\lambda_i$ と $18\bar{\lambda}_i$ を示した。図3には、同じく $\bar{E}$ を示した。図2と3より、式(3.4), (3.5)の傾向、テスト行列の差異による誤差の微妙な差異、アルゴリズムと δA_i の大さ相等、様々のものを読み取ることが出来る。

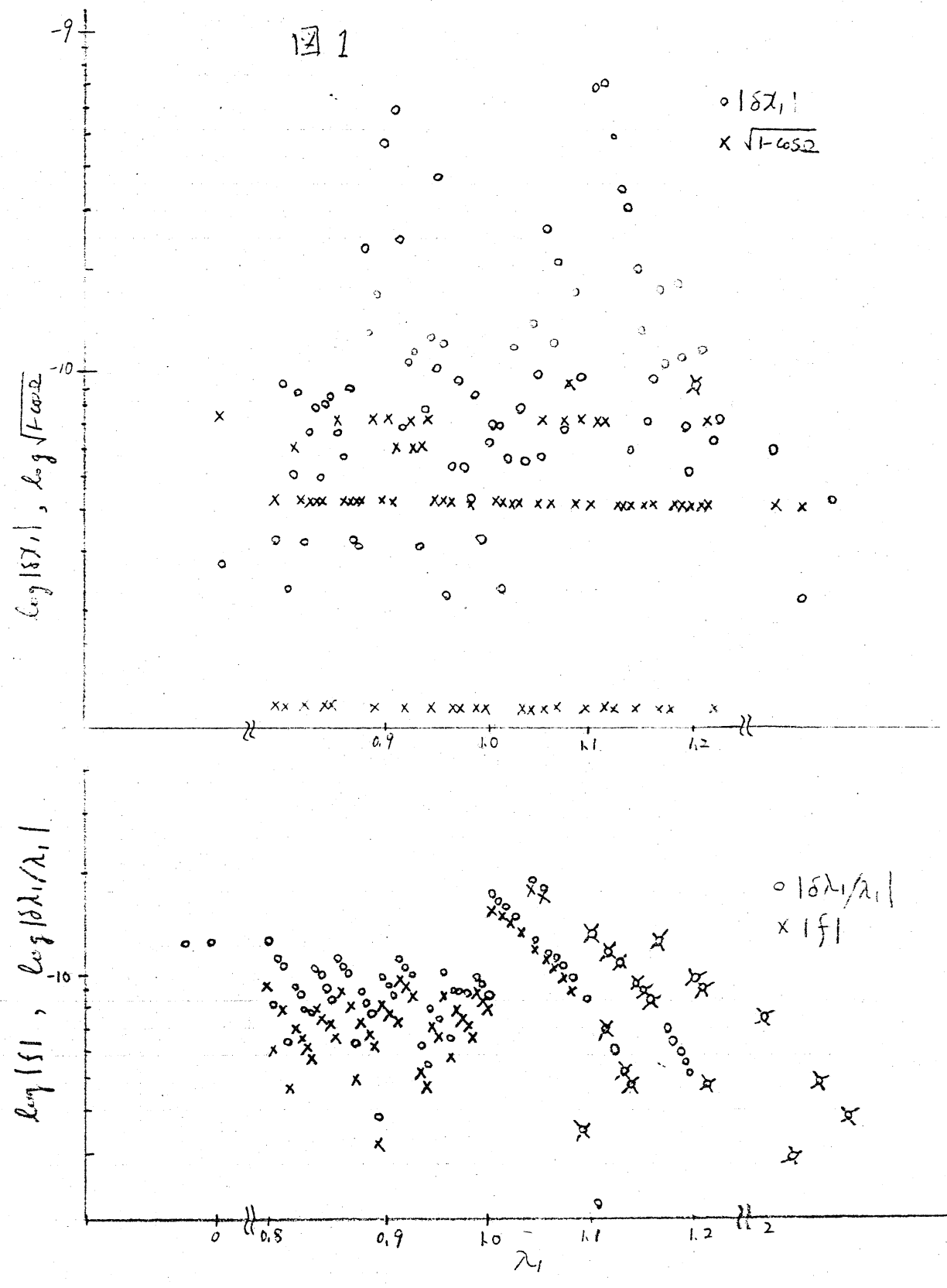
6. まとめ

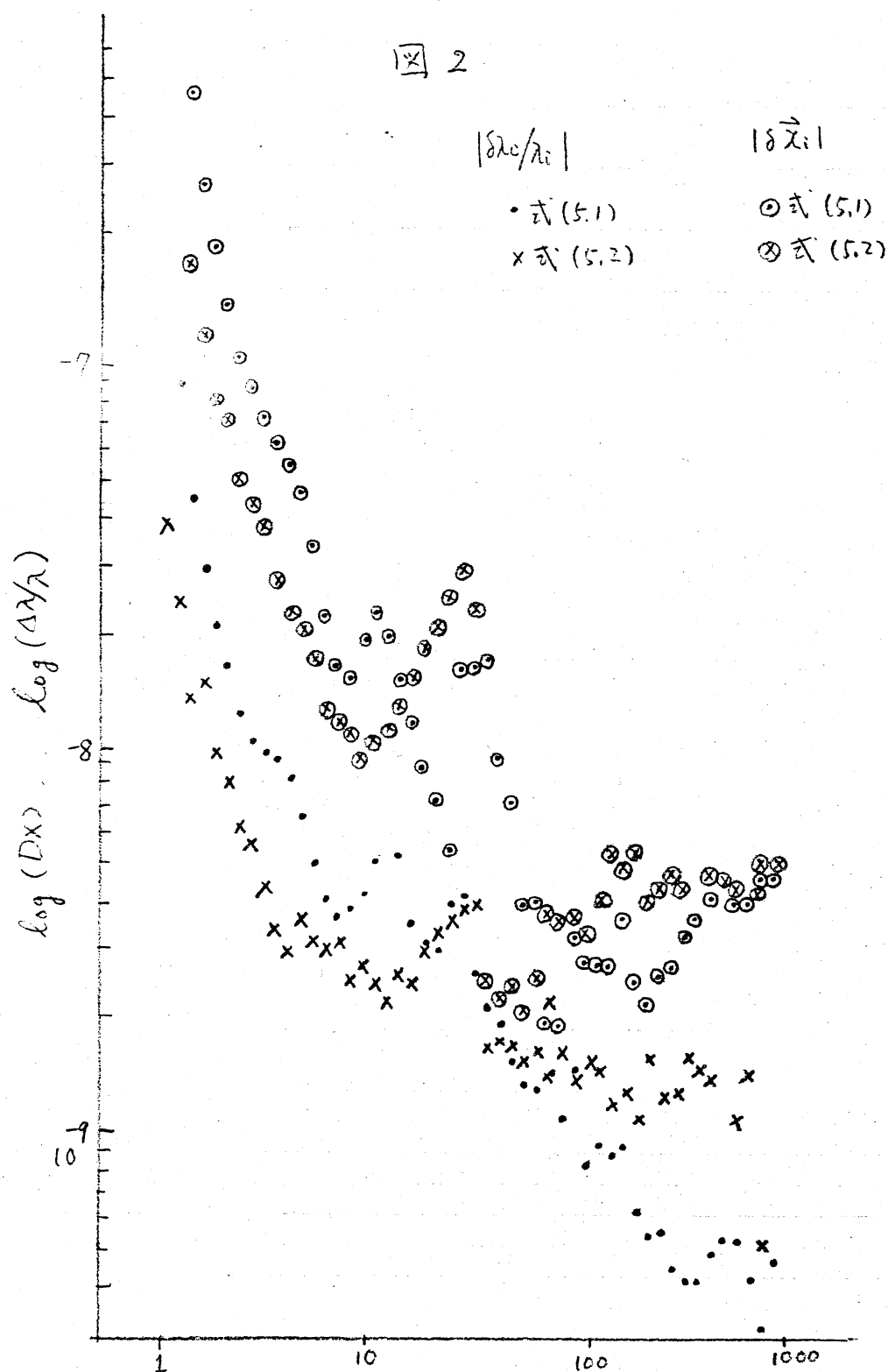
アルゴリズムに関するテスト方式は、一応完了した。丸め誤差累積効果に関するテスト方式については、未だ研究を進めている段階である。特に、 X 及び Λ の与え方と、この結果で与える行列 A の「複雑さ」や計算の手間との関係については、様々のアルゴリズムの可能性を踏えた上で、定性的・定量的に明らかにしなければならないと考えている。

参考文献

- 1). J.H.ウルフソン著/一松 信, 四条忠雄 訳; 基本的演算における丸め誤差解析, 培風館 (1974).
- 2). 遠山 啓; 行列論, 共立全書 (1960)
- 3). 山内恭彦; 回転群とその表現, 岩波書店 (1973)

图 1





[2] 3

$$E_{\perp}^i = \langle \sum_{j \neq i} \alpha_{ij} \vec{\lambda}_j \cdot \delta A_i \vec{\lambda}_i \rangle = \delta_{\perp}^i \sum_{j \neq i} \alpha_{ij}^2 (\lambda_i - \lambda_j)$$

• 式 (5.1)

x 式 (5.2)

