

論文 / 著書情報
Article / Book Information

題目(和文)	σ - 逆行列とその応用に関する研究
Title(English)	
著者(和文)	佐藤誠
Author(English)	Makoto Sato
出典(和文)	学位:工学博士, 学位授与機関:東京工業大学, 報告番号:甲第989号, 授与年月日:1978年3月26日, 学位の種別:課程博士, 審査員:
Citation(English)	Degree:Doctor of Engineering, Conferring organization: Tokyo Institute of Technology, Report number:甲第989号, Conferred date:1978/3/26, Degree Type:Course doctor, Examiner:
学位種別(和文)	博士論文
Type(English)	Doctoral Thesis

0-逆行列 とその 応用 に関する研究

指導教官

飯島 泰蔵 教授

提出者

大学院博士課程電気工学専攻

佐 藤 誠

目 次

第 1 章 序 論

- 1. 1 研究の背景とその目的 1
- 1. 2 研究の概要 4

第 2 章 σ -逆行列の定義とその性質

- 2. 1 まえがき 5
- 2. 2 定義と基本性質 5
- 2. 3 最小分散推定と σ -逆行列 18

第 3 章 σ -逆行列のスペクトル特性

- 3. 1 まえがき 26
- 3. 2 直方行列のスペクトル理論 26
- 3. 3 σ -逆行列のスペクトル特性 34
- 3. 4 σ -逆行列の摂動 40

第 4 章 分割行列の σ -逆行列

- 4. 1 まえがき 44
- 4. 2 分割行列の σ -逆行列 44
- 4. 3 最小分散推定の逐次推定 49

第 5 章 σ -逆行列の計算法

- 5. 1 まえがき 60
- 5. 2 σ -逆行列の計算法 I —直接法— 60

5.3	σ -逆行列の計算法 II - 逐次計算法 -	67
5.4	Moore-Penrose 逆行列の正則化と σ -逆行列	69
第 6 章 最適分散性特徴パラメータ		
6.1	まえがき	72
6.2	特徴パラメータ	73
6.3	Doubly Stochastic 行列	75
6.4	最適分散性特徴パラメータ	78
6.5	最適分散性特徴パラメータの応用	90
6.6	あとがき	93
第 7 章 線形観測システムの最適化問題		
7.1	線形観測システムの最適化	95
7.2	最適化問題 I	97
7.3	最適化問題 II	102
	付録	110
第 8 章 曲線のあてはめ問題と σ -逆行列		
8.1	曲線のあてはめ問題	113
8.2	σ -逆行列による曲線のあてはめ	115
8.3	線形観測システムにおける σ の推定方法	118
第 9 章 σ -逆行列の画像処理への応用		
9.1	まえがき	128
9.2	投影像の観測システム	129
9.3	断層像の再構成システム	136
9.4	投影角の最適化	143
9.5	あとがき	154

第10章 結論

156

第1章 序論

謝辞

157

参考文献

158

1.1

発表論文一覧

161

正交行列 A に対し

$$AA^T = A^T A = I \quad (1.1)$$

となる行列 A^T が存在すると、この行列 A^T を行列 A の逆行列という。但し、ここで、 I は単位行列である。逆行列の概念は、行列に関する基本的な概念の一つであるが、正交行列も正則な行列にのみ対応する行列である。

この逆行列の概念の拡張として、長方形列を含む任意の行列に対して存在し、逆行列のもつある種の性質を保持（という行列のことを一般逆行列という）逆行列のどのような性質に着目するかにより、種々の一般逆行列が存在する。その中でも特に重要な一般逆行列として、Moore-Penrose 逆行列がある。

1955年、R. Penrose は、任意の行列（長方形列も含む） A に対して、次の四つの行列方程式を満足する行列 X が一意に存在することも示した⁽¹⁷⁾。

$$AXA = A \quad (1.2)$$

$$XAX = X \quad (1.3)$$

$$(AX)^* = AX \quad (1.4)$$

$$(XA)^* = XA \quad (1.5)$$

第1章 序論

1. 1 研究の背景とその目的

正方行列 A に対し

$$AA^{-1} = A^{-1}A = I \quad (1.1)$$

となる行列 A^{-1} が存在するとき、この行列 A^{-1} を行列 A の逆行列という。但し、ここで、 I を単位行列である。逆行列の概念は、行列に関する基本的な概念の一つであるが、正方でしかも正則な行列にのみ存在する行列である。

この逆行列の概念の拡張として、長方形行列を含む任意の行列に対して存在し、逆行列のもつある種の性質を保持している行列のことを一般逆行列という。逆行列のどの性質に着目するかにより、種々の一般逆行列が存在する。その中でも特に重要な一般逆行列として、Moore-Penrose 逆行列がある。

1955年、R. Penrose は、任意の行列（長方形行列も含む） A に対して、次の四つの行列方程式を満足する行列 X が一意に存在することを示した⁽¹⁷⁾。

$$AXA = A \quad (1.2)$$

$$XAX = X \quad (1.3)$$

$$(AX)^* = AX \quad (1.4)$$

$$(XA)^* = XA \quad (1.5)$$

ここで、 A^* は行列 A の随伴行列である。この行列 X は、既に1920年代に、E. H. Moore により、全く別の方法で定義付けられ、研究がなされた。このことから、行列 X のことを、Moore-Penrose 逆行列と呼び、 A^+ と表わす。

この Moore-Penrose 逆行列は、連立方程式の最小二乗解を与える行列であり、また統計的推定問題における最小二乗推定あるいは最良線形不偏推定のための行列でもある。このように、Moore-Penrose 逆行列は、その数学的構造が整っているとともに、応用分野も広いため、数多くの研究がなされている。

推定問題は、科学、工学ならびに社会科学の分野における最も重要な問題の一つである。推定問題の多くは、いわゆる統計的線形モデルとして取扱うことができる。いま、次のような線形観測システムを考える。

$$Z = AX + V \quad (1.6)$$

ここで、 X は未知ベクトル、 A は既知の行列、 Z は観測ベクトルである。観測時に伴う観測誤差が伴う。ベクトル V が観測誤差を表わす項である。この線形観測システムにおける推定問題とは、得られた観測ベクトル Z から、未知ベクトル X の推定量 $\hat{X}$ を求めることである。

このような推定問題では、推定のための最適基準の与え方により、様々な推定量が得られる。最適基準として、不偏な（あるいは偏り最小の）線形推定量の中で、真値 X と推定量 $\hat{X}$ との間の誤差分散 $E(\|\hat{X} - X\|^2)$ が最小となるもの、という規準を選んだときの推定量が、最良線形不偏推定量である。この推定量は、幾つかの仮定の下に、いわゆる最小二乗推定量と呼ばれる推定量と同じものとなる。最小二乗推定量は行列 A の Moore-Penrose 逆行列 A^+ を用いて、次のように与えられる。

$$\hat{X} = A^+ Z \quad (1.7)$$

最小二乗推定の注目すべき点は、推定の際に、ベクトル x やベクトル y の統計的性質に関する事前の情報を必要としなければならないことである。このことは、(1.7)において用いられている Moore-Penrose 逆行列 A^+ が行列 A だけから一意に定まることから明らかである。

対象に関する統計的性質についての事前の情報を必要としなければならないことは、最小二乗推定の一つの大きな利点であり、統計処理の分野において、最小二乗推定が古くから良く使われていることの理由の一つである。しかし、この特徴は、最小二乗推定の一つの限界をも示している。いま、対象とする線形観測システムの統計的性質が事前により明らかな場合には、その情報を十分に活かした推定を行なわなければならない。最小二乗推定では、このような事前の情報の活用ということが十分に行なえないのである。

ベクトル x とベクトル y の平均ならびに共分散といった統計的性質が事前に分かっている場合には、真値 x と推定量 $\hat{x}$ との間の誤差分散 $E(\| \hat{x} - x \|^2)$ が小さいという意味で、最小二乗推定よりも良い推定が存在する。この推定のことを最小分散推定と呼ぶ。この推定は、最適規準として、単に誤差分散 $E(\| \hat{x} - x \|^2)$ を最小にするという規準を用いたときの最適な推定のことである。

画像処理あるいはパターン認識などの分野における推定問題では、推定すべき未知ベクトルの次元に比べて、用いることのできる観測データの次元は決して多くなは。このため、事前に対象ならびに観測誤差の統計的性質を十分に調べて、その情報を活用しなければならない。このような立場から見ると、最小二乗推定に比べて、最小分散推定がより重要な推定方法として位置付けられる。

以上述べてきたことを背景として、本研究では、最小二乗推定のための数学的道具である Moore-Penrose 逆行列に対応した、最小分散推定のための数学的道具を求め、その性質を明らかにすることを目的とする。

1. 2 本研究の概要

本研究では、はじめに第2章において、最小分散推定のための数学的道具として、 G -逆行列という概念を与え、その代数的な定義付けを行なう。さらに、 G -逆行列のもつ基本性質、最小分散推定における役割を明らかにする。第3章においては、 G -逆行列の数学的構造を解明するために、 G -逆行列のスペクトル特性を求める。 G -逆行列と Moore-Penrose 逆行列との間の類似性及び、相違性は、すべてこのスペクトル特性の上に集約されている。

第4章、ならびに第5章では、実用的見地から見た G -逆行列の構造ならびに性質を明らかにする。第4章では、分割行列の G -逆行列を求め、最小分散推定の逐次推定の方法を明らかにする。第5章では、 G -逆行列の計算法を、直接法と逐次計算法の二つに分けて考察する。

第5章までの前半が、 G -逆行列の理論を扱ったものである。第6章からの後半は、 G -逆行列の応用に関する研究である。

第6章では、線形観測システムの解析において重要な概念である最適分散性特徴パラメータについて明らかにして、その応用例を与える。この章は、次の第7章の準備にもなっている。第7章では、線形観測システムの最適化問題を扱う。この章では、二種類の最適化問題を考察する。なお、最適化問題Ⅱにおいては、第6章の最適分散性特徴パラメータの性質を利用する。

第8章では、 G -逆行列の最も基本的な応用として、曲線のお2はめ問題を扱う。従来の最小二乗法による処理方法との比較を行なう。第9章では、 G -逆行列の応用として、画像処理への応用を考える。ここで扱う画像処理は、与えられた断層像をその投影像から再構成する問題である。この問題の処理において、各章で得られた G -逆行列に関する諸性質が有効に活かされている。

G -逆行列の数学的理論を扱う第2章と第3章だけは、行列の係数体として複素数を与えた。残りの章は、すべて実数を係数体としている。

第2章 σ -逆行列の定義とその性質

2.1 まえがき

σ -逆行列の定義の方法として、(i) 代数的な方法、(ii) 構成的な方法、(iii) 最小分散相定を与える行列としての定義付け。の三つが考えられるが、本章では、(i) の代数的な方法による定義付けを行なう。Moore-Penrose 逆行列は R. Penrose⁽¹⁷⁾ の与えた四つの行列方程式を満たす行列として代数的に定義付けられる。2.2 にあいて、この四つの行列方程式に対応する二つの行列方程式を用いて、 σ -逆行列の定義付けを行なう。さらに、 σ -逆行列のもつ基本性質を明らかにする。

2.3 では、統計的線形モデルシステムでの最小分散相定を与える行列としての σ -逆行列の役割を明らかにする。

2.2 定義と基本性質

σ -逆行列は、二つの行列方程式により、次のように代数的に定義される。

[定義 2.1] A を任意の $m \times n$ 行列とする。 σ を正の実数とする。このとき、次の二つの行列方程式

$$AXA + \sigma^2 X^* = A \quad (2.1)$$

$$(XA)^* = XA \quad (2.2)$$

を満たす $n \times m$ 行列 X を行列 A の σ -逆行列と呼ぶ。行列 A の σ -逆行列を A^σ と表わす。

なお、ここで A^* は、行列 A の随伴行列である。

定義 2.1 で与えた σ -逆行列は、任意の $m \times n$ 行列 A に対して、必ず存在して、しかも常に一意に定まる。 σ -逆行列の存在性の一意性を明らかにしたのが次の定理である。

[定理 2.1] 任意の $m \times n$ 行列 A に対して、その σ -逆行列 A^σ は一意に存在する。そして、行列 A の σ -逆行列 A^σ は次式で与えられる。

$$A^\sigma = A^* (AA^* + \sigma^2 I)^{-1} \quad (2.3)$$

ここで、 I は、単位行列である。

(証明) (2.2) を (2.1) に代入する。

$$A(XA)^* + \sigma^2 X^* = A$$

$$\therefore (AA^* + \sigma^2 I) X^* = A \quad (2.4)$$

ここで、 σ は正の実数であるから、行列 $AA^* + \sigma^2 I$ は正値エルミート行列である。ゆえに、行列 $AA^* + \sigma^2 I$ の逆行列 $(AA^* + \sigma^2 I)^{-1}$ が存在する。よって、

$$X^* = (AA^* + \sigma^2 I)^{-1} A$$

となる。これより、

$$X = A^*(AA^* + \sigma^2 I)^{-1} \quad (2.5)$$

が求まる。

逆に、(2.5) で与えた行列 X が (2.1) 及び (2.2) を満足することは、明らかである。

(証明終)

定理 2.1 において、正の実数 σ が、 σ -逆行列 A^σ の存在性と一意性のために重要な役割を演じていることが分かる。

σ -逆行列のもつ幾つかの基本性質を、系 2.1 に示す。

[系 2.1] 行列 A の σ -逆行列 A^σ は、次の性質をもつ。

$$(a) \quad A^\sigma = (A^*A + \sigma^2 I)^{-1} A^* \quad (2.6)$$

$$(b) \quad (AA^\sigma)^* = A A^\sigma \quad (2.7)$$

$$(c) \quad (A^*)^\sigma = (A^\sigma)^* \quad (2.8)$$

(d) 行列 A がユニタリ行列であるとするとする。このとき、

$$A^\sigma = \frac{1}{1+\sigma^2} A^* \quad (2.9)$$

(e) $\lambda \neq 0$ なる実数 λ に対して、次式が成り立つ。

$$(\lambda A)^\sigma = \frac{1}{\lambda} A^{\sigma'} \quad (2.10)$$

そこで、 $\sigma' = \frac{\sigma}{|\lambda|}$ である。

(f) 行列 U と V が、ともにユニタリ行列であるとする。このとき、

$$(UAV)^\sigma = V^* A^\sigma U^* \quad (2.11)$$

となる。

(証明) (a) A^σ の左側から、行列 $A^*A + \sigma^2 I$ を掛ける。

$$\begin{aligned} & (A^*A + \sigma^2 I) A^\sigma \\ &= (A^*A + \sigma^2 I) A^* (AA^* + \sigma^2 I)^{-1} \\ &= A^* (AA^* + \sigma^2 I) (AA^* + \sigma^2 I)^{-1} \\ &= A^* \end{aligned} \quad (2.12)$$

そこで、行列 $A^*A + \sigma^2 I$ は、正値対称行列であるから、逆行列 $(A^*A + \sigma^2 I)^{-1}$ が存在する。ゆえに、

$$A^\sigma = (A^*A + \sigma^2 I)^{-1} A^*$$

となる。

(b) (2.6) を用いることにより.

$$\begin{aligned} & (AA^\sigma)^* \\ &= \{ A(A^*A + \sigma^2 I)^{-1} A^* \}^* \\ &= A(A^*A + \sigma^2 I)^{-1} A^* \\ &= AA^\sigma \end{aligned} \tag{2.13}$$

となる。

(c)

$$\begin{aligned} & (A^\sigma)^* \\ &= \{ A^*(AA^* + \sigma^2 I)^{-1} \}^* \\ &= (AA^* + \sigma^2 I)^{-1} A \\ &= (A^*)^\sigma \end{aligned} \tag{2.14}$$

(d) 行列 A が $U = \sigma$ 行列ならば, $AA^* = I$ となる。これより,

$$\begin{aligned} & A^\sigma \\ &= A^*(AA^* + \sigma^2 I)^{-1} \end{aligned}$$

$$= \frac{1}{1+\sigma^2} A^* \quad (2.15)$$

(e)

$$\begin{aligned} & (\lambda A)^\sigma \\ &= \bar{\lambda} A^* (\lambda |\lambda|^2 A A^* + \sigma^2 I)^{-1} \\ &= \frac{\bar{\lambda}}{|\lambda|^2} A^* (A A^* + \frac{\sigma^2}{|\lambda|^2} I)^{-1} \\ &= \frac{1}{\lambda} A^{\sigma'} \end{aligned} \quad (2.16)$$

ここで、 $\sigma' = \frac{\sigma}{|\lambda|}$ とおく。

(f)

$$\begin{aligned} & (UAV)^\sigma \\ &= (UAV)^* \{ UAV (UAV)^* + \sigma^2 I \}^{-1} \\ &= V^* A^* U^* \{ U (A A^* + \sigma^2 I) U^* \}^{-1} \\ &= V^* A^* (A A^* + \sigma^2 I)^{-1} U^* \\ &= V^* A^\sigma U^* \end{aligned} \quad (2.17)$$

系 2.1, (b) は、の逆行列 A^σ が、(2.1) と (2.2) の行列方程式の他に次の行列方程式を満足していることを示している。

$$(AX)^* = AX \quad (2.18)$$

このことから、(2.1), (2.2) 及び (2.18) の代数的構造が互いに独立であることが分かる。 σ -逆行列の定義として、(2.1) と (2.18) の二つの行列方程式を用いることにより、全く同様の議論を展開することが可能である。

系 2.1, (b), (c), (e) 及び (f) は、Moore-Penrose 逆行列 A^+ に対しても成り立つ性質である。Moore-Penrose 逆行列 A^+ の一つ重要な性質として、反射性という性質がある。即ち、

$$(A^+)^+ = A \quad (2.19)$$

となる。 σ -逆行列 A^σ には、この反射性は成り立たない。

σ -逆行列 A^σ の簡単な例を示す。

[例 2.1]

(a) a をスカラーとする。このとき、

$$a^\sigma = \frac{\bar{a}}{|a|^2 + \sigma^2} \quad (2.20)$$

となる。

(b)

$$I^\sigma = \frac{1}{1 + \sigma^2} I \quad (2.21)$$

(c)

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} \quad (2.22)$$

このとき、行列 A の σ -逆行列 A^σ は、次のようになる。

$$\begin{aligned}
 A^\sigma &= (A^*A + \sigma^2 I)^{-1} A^* \\
 &= \begin{bmatrix} 2+\sigma^2 & 1 \\ 1 & 2+\sigma^2 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \\
 &= \frac{1}{(3+\sigma^2)(1+\sigma^2)} \begin{bmatrix} 2+\sigma^2 & -1 \\ -1 & 2+\sigma^2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \\
 &= \frac{1}{(3+\sigma^2)(1+\sigma^2)} \begin{bmatrix} 2+\sigma^2 & 1+\sigma^2 & -1 \\ -1 & 1+\sigma^2 & 2+\sigma^2 \end{bmatrix} \quad (2.23)
 \end{aligned}$$

σ -逆行列 A^σ と Moore-Penrose 逆行列 A^+ との間の関係を明らかにするために、次に二つの補助定理を与える。

[補助定理 2. A]^(注) 行列 A に対し、行列 X が行列 A の Moore-Penrose 逆行列 A^+ であるための必要十分条件は、次の二つの条件である。

$$(a) \quad XA = P_{R(A^*)} \quad (2.24)$$

$$(b) \quad AX = P_{R(A)} \quad (2.25)$$

ここで、 $P_{R(A)}$ は、行列 A の値域 $R(A)$ への正射影である。

[補助定理 2. B] A を任意のエルミート行列とする。このとき、次の式

(注) 既に証明されている定理、補助定理、系については、本文のアルファベットで番号付けを行なった。

が成り立つ。

$$P_R(A) = \lim_{\delta \rightarrow 0} (A + \delta I)^{-1} A \quad (2.26)$$

$$= \lim_{\delta \rightarrow 0} A (A + \delta I)^{-1} \quad (2.27)$$

この二つの補助定理と (2.3) 及び (2.6) より、次の定理を得る。

[定理 2.2]

$$\lim_{\sigma \rightarrow +0} A^\sigma = A^+ \quad (2.28)$$

(証明) $X = \lim_{\sigma \rightarrow +0} A^\sigma$ とおく。このとき、次式が成り立つ。

$$\begin{aligned} XA &= \lim_{\sigma \rightarrow +0} (A^*A + \sigma^2 I)^{-1} A^*A \\ &= P_R(A^*A) \end{aligned} \quad (2.29)$$

$$\begin{aligned} AX &= \lim_{\sigma \rightarrow +0} AA^* (AA^* + \sigma^2 I)^{-1} \\ &= P_R(AA^*) \end{aligned} \quad (2.30)$$

こゝで $R(A^*A) = R(A^*)$, $R(AA^*) = R(A)$ であるから、補助定理 2. A より、 X は Moore-Penrose 逆行列 A^+ である。

(証明終)

Moore-Penrose 逆行列の場合と同様に、 σ -逆行列 A^σ も重み付き σ -逆行列に拡張することができる。

〔定義 2.2〕 A を任意の $m \times n$ 行列, σ を正の実数とある。行列 R と行列 S をそれぞれ m 次, n 次の正值エルミート行列とある。このとき、次の二つの行列方程式

$$AXAS + \sigma^2 R X^* = AS \quad (2.31)$$

$$(XAS)^* = XAS \quad (2.32)$$

を満たす $n \times m$ 行列 X を行列 A の重み付き σ -逆行列と呼ぶ。行列 A の重み付き σ -逆行列を A_{SR}^σ と表わす。

重み付き σ -逆行列 A_{SR}^σ に対して、次の定理が成り立つ。

〔定理 2.3〕 行列 A の重み付き σ -逆行列 A_{SR}^σ は、一意に存在して、次式を満足する。

$$A_{SR}^\sigma = SA^*(ASA^* + \sigma^2 R)^{-1} \quad (2.33)$$

(証明) (2.32) を (2.31) に代入する。

$$A(XAS)^* + \sigma^2 R X^* = AS$$

$$(ASA^* + \sigma^2 R) X^* = AS \quad (2.34)$$

ここで、 σ は正の実数、 R と S とは、ともに正値エルミート行列であるから、行列 $ASA^* + \sigma^2 R$ も正値エルミート行列である。よって、

$$X^* = (ASA^* + \sigma^2 R)^{-1} AS \quad (2.35)$$

となる。ゆえに、(2.33) が求まる。

(証明終)

重み付きの逆行列 A_{SR}^σ は、 σ 逆行列を用いて次のように表わすこともできる。

[系 2.2]

$$A_{SR}^\sigma = S^{\frac{1}{2}} B^\sigma R^{-\frac{1}{2}} \quad (2.36)$$

ここで、 $B = R^{-\frac{1}{2}} A S^{\frac{1}{2}}$ である。

(証明)

$$\begin{aligned} A_{SR}^\sigma &= SA^*(ASA^* + \sigma^2 R)^{-1} \\ &= S^{\frac{1}{2}} B^* (BB^* + \sigma^2 I)^{-1} R^{-\frac{1}{2}} \\ &= S^{\frac{1}{2}} B^\sigma R^{-\frac{1}{2}} \end{aligned}$$

(証明終)

重み付きの逆行列 A_{SR}^σ のもつ幾つかの性質を次の系に示す。これは、 σ 逆

行列 A^σ に関する性質. 系 2.1 に対応するものである。

[系 2.3]

$$(a) \quad A_{SR}^\sigma = (A^* R^{-1} A + \sigma^2 S^{-1})^{-1} A^* R^{-1} \quad (2.37)$$

$$(b) \quad (R^{-1} A A_{SR}^\sigma)^* = R^{-1} A A_{SR}^\sigma \quad (2.38)$$

$$(c) \quad (A^*)_{R^{-1}S^{-1}}^\sigma = (A_{SR}^\sigma)^* \quad (2.39)$$

$$(d) \quad (\lambda A)_{SR}^\sigma = \frac{1}{\lambda} A_{SR}^{\sigma'} \quad (2.40)$$

ここで $\lambda \neq 0$ であり, $\sigma' = \frac{\sigma}{|\lambda|}$ である。

(証明) (a) $B = R^{-\frac{1}{2}} A S^{\frac{1}{2}}$ とおくと, 系 2.2 及び系 2.1, (a) より, 次のようになる。

$$\begin{aligned} A_{SR}^\sigma &= S^{\frac{1}{2}} B^\sigma R^{-\frac{1}{2}} \\ &= S^{\frac{1}{2}} (B^* B + \sigma^2 I)^{-1} B^* R^{-\frac{1}{2}} \\ &= S^{\frac{1}{2}} \{ S^{\frac{1}{2}} (A^* R^{-1} A + \sigma^2 S^{-1}) S^{\frac{1}{2}} \}^{-1} S^{\frac{1}{2}} A^* R^{-1} \\ &= (A^* R^{-1} A + \sigma^2 S^{-1})^{-1} A^* R^{-1} \end{aligned}$$

(b) (2.37) より, $R^{-1} A A_{SR}^\sigma$ がエルミート行列となることは明らかである。

(c) 同様. (2.37) より得らる.

$$\begin{aligned}
 & (A_{SR}^\sigma)^* \\
 &= \{ S A^* (A S A^* + \sigma^2 R)^{-1} \}^* \\
 &= (A S A^* + \sigma^2 R)^{-1} A S \\
 &= (A^*)_{R^{-1} S^{-1}}^\sigma
 \end{aligned}$$

(d) $B = R^{-\frac{1}{2}} A S^{\frac{1}{2}}$ とおく.

$$(\lambda A)_{SR}^\sigma = S^{\frac{1}{2}} (\lambda B)^\sigma R^{-\frac{1}{2}}$$

==で. 系 2.1, (e) の性質を用いることにより.

$$\begin{aligned}
 & (\lambda A)_{SR}^\sigma \\
 &= S^{\frac{1}{2}} \cdot \frac{1}{\lambda} B^{\sigma'} R^{-\frac{1}{2}} \\
 &= \frac{1}{\lambda} S^{\frac{1}{2}} B^{\sigma'} R^{-\frac{1}{2}} \\
 &= \frac{1}{\lambda} A_{SR}^{\sigma'}
 \end{aligned}$$

とある. 但し, $\sigma' = \frac{\sigma}{|\lambda|}$ である.

(証明終)

2.3 最小分散推定と σ -逆行列

σ -逆行列の推定問題における役割を明らかにする。

ここで対象とする観測システムは、次のような統計的線形観測システムである。

$$x = Ax + w \quad (2.41)$$

x は、 n 次元の未知ベクトル、 A は既知の $n \times m$ 行列、 w は、 m 次元の観測ベクトルである。 w は、観測時に加わる観測誤差を表わす m 次元ベクトルである。

ベクトル x とベクトル w は、互いに独立な確率ベクトルであり、その平均と共分散が次のように与えられている。

$$E(x) = 0, \quad \text{Cov}(x) \stackrel{(\text{注})}{=} S \quad (2.42)$$

$$E(w) = 0, \quad \text{Cov}(w) = \sigma^2 R \quad (\sigma > 0) \quad (2.43)$$

ここで、 S と R は、ともに正値エルミート行列であり、 σ は、正の実数である。 σ の値は、行列 R を何らかの方法で正規化することにより、具体的に定まる。なお、行列 R の正規化の方法としては、(i) $\text{tr} R = m$, (ii) $\text{tr} R = 1$, (iii) $|R| = 1$ などが考えられる。ここで $\text{tr} A$ は、行列 A のトレースを表わす。

(2.42) 及び (2.43) のような統計的性質をもつ線形観測システムをここでは、 $(A, S, \sigma^2 R)$ と表わすことにする。特に、行列 S と行列 R がともに単位行列の場合は、単に (A, σ^2) と表わす。

線形観測システム $(A, S, \sigma^2 R)$ における推定問題とは、得られた観測ベ

(注) $\text{Cov}(x)$ は、ベクトル x の共分散を表わす。即ち、

$$\text{Cov}(x) = E(xx^*)$$

となる。

クトル $\mathbf{x}$ から、未知ベクトル $\mathbf{x}$ の推定量 $\hat{\mathbf{x}}$ を求めることである。推定量 $\hat{\mathbf{x}}$ としては、推定における最適基準の選び方により、様々なものが得られる。その中で、単に推定量 $\hat{\mathbf{x}}$ と真値 $\mathbf{x}$ との間の誤差分散 $E(\|\hat{\mathbf{x}} - \mathbf{x}\|^2)$ の最小化を最適基準としたものが、最小分散推定量である。

$\mathbf{G}$ -逆行列は、この最小分散推定量を与える行列である。

[定理 2.4] 観測システム $(A, S, \sigma^2 R)$ における未知ベクトル $\mathbf{x}$ の最小分散推定量 $\hat{\mathbf{x}}$ は、次式で与えられる。

$$\hat{\mathbf{x}} = A_S^{\sigma} \mathbf{z} \quad (2.44)$$

このとき、誤差共分散 $\text{Cov}(\hat{\mathbf{x}} - \mathbf{x})$ は

$$\text{Cov}(\hat{\mathbf{x}} - \mathbf{x}) = (I - A_S^{\sigma} A) S \quad (2.45)$$

となる。

(証明) 最小分散推定量 $\hat{\mathbf{x}}$ が、ある $n \times m$ 行列 K により、次式で与えられるものとする。

$$\hat{\mathbf{x}} = K \mathbf{z} \quad (2.46)$$

このとき、

$$\hat{\mathbf{x}} - \mathbf{x} = (KA - I) \mathbf{x} + K \mathbf{v} \quad (2.47)$$

となる。これより、誤差共分散 $\text{Cov}(\hat{\mathbf{x}} - \mathbf{x})$ は次のようになる。

$$\begin{aligned}
& \text{Cov}(\hat{X} - X) \\
&= E((\hat{X} - X)(\hat{X} - X)^*) \\
&= (KA - I)S(KA - I)^* + \sigma^2 K R K^* \\
&= K(ASA^* + \sigma^2 R)K^* - KAS - SA^*K^* + S
\end{aligned} \tag{2.48}$$

ここで、

$$K = A_{SR}^{\sigma} + K_1 \tag{2.49}$$

とおく。 K_1 は $n \times m$ 行列である。このとき、 $\text{Cov}(\hat{X} - X)$ は

$$\begin{aligned}
& \text{Cov}(\hat{X} - X) \\
&= (A_{SR}^{\sigma} + K_1)(ASA^* + \sigma^2 R)(A_{SR}^{\sigma} + K_1)^* \\
&\quad - (A_{SR}^{\sigma} + K_1)AS - SA^*(A_{SR}^{\sigma} + K_1)^* + S \\
&= K_1(ASA^* + \sigma^2 R)K_1^* + (I - A_{SR}^{\sigma}A)S
\end{aligned} \tag{2.50}$$

となる。

推定量 $\hat{X}$ と真値 X との間の誤差分散 $E(\|\hat{X} - X\|^2)$ は、 $\text{Cov}(\hat{X} - X)$ のトレースにより与えられる。即ち、

$$E(\|\hat{X} - X\|^2) = \text{tr} \text{Cov}(\hat{X} - X) \tag{2.51}$$

となる。いま、任意の半正値エルミート行列 A に対して、 $A \geq 0$ となることに注意する。(2.50) の右辺、第一項 $K_1(A S A^* + \sigma^2 R) K_1^*$ は、半正値エルミート行列であるから、このトレースは、常に非負である。このことから、(2.51) の誤差分散が最小となるのは、 $K_1 = 0$ のときであることが分かる。即ち、ベクトル x の最小分散推定量 $\hat{x}$ は、(2.44) で与えられる。

このとき、誤差共分散 $\text{cov}(\hat{x} - x)$ が (2.45) となることも、(2.50) より明らかである。

(証明終)

本節では、未知ベクトル x の平均 $E(x)$ が零ベクトルとなる場合を対象としている。しかし、未知ベクトル x の平均 $E(x)$ が、

$$E(x) = x_0 \quad (2.52)$$

となる場合でも、 $x' = x - x_0$ は平均が零ベクトルの確率ベクトルとなることに注意すると、最小分散推定量 $\hat{x}$ が重み付きの逆行列 $A S R$ を用いて与えられる。即ち、

$$\hat{x} = x_0 + A S R^{-1} (y - A x_0) \quad (2.53)$$

となる。

行列 S と行列 R とが、ともに単位行列の場合が、次の系である。

[系 2.4] 線形観測システム (A, σ^2) における未知ベクトル x の最小分散推定量 $\hat{x}$ は、次式で与えられる。

$$\hat{x} = A^T x \quad (2.54)$$

このとき、誤差共分散 $\text{Cov}(\hat{X} - X)$ は次のようになる。

$$\text{Cov}(\hat{X} - X) = I - A^{\sigma} A \quad (2.55)$$

最小分散推定に関連して、さらに次の二つの系が成り立つ。

[系 2.5] T を任意の $p \times n$ 行列とする。このとき、ベクトル $Y = TX$ の最小分散推定量 $\hat{Y}$ は次式で与えられる。

$$\begin{aligned} \hat{Y} &= T \cdot \hat{X} \\ &= T \cdot A_{SR}^{\sigma} \cdot X \end{aligned} \quad (2.56)$$

(証明) 定理 2.4 の証明と同じように $\hat{Y}$ が次式で与えられるとする。

$$\hat{Y} = K \cdot X \quad (2.57)$$

そこで、 K は $p \times m$ 行列である。このとき、

$$\hat{Y} - Y = (KA - T)X + KV \quad (2.58)$$

となる。このより、誤差共分散 $\text{Cov}(\hat{Y} - Y)$ は、

$$\begin{aligned} &\text{Cov}(\hat{Y} - Y) \\ &= K(A^{\sigma} A^* + \sigma^2 R)K^* - K A^{\sigma} T^* - T A^{\sigma} K^* + T S T^* \end{aligned} \quad (2.59)$$

となる。ここで、

$$K = T A_{SR}^{\sigma} + K_1 \quad (2.60)$$

と置く。但し、 K_1 は $p \times m$ 行列である。このとき、

$$\begin{aligned} \text{Cov}(\hat{Y} - Y) \\ = K_1 (A S A^* + \sigma^2 R) K_1^* + T (I - A_{SR}^{\sigma} A) S T^* \end{aligned} \quad (2.61)$$

となる。これより、定理 2.4 の証明と同じように、誤差分散 $E(\|\hat{Y} - Y\|^2)$ は、 $K_1 = 0$ のときに最小となる。

(証明終)

[系 2.6] ベクトル X の最小分散推定量 $\hat{X} = A_{SR}^{\sigma} X$ は、任意の半正値エルミート行列 Q に対して、

$$E(\|\hat{X} - X\|_Q^2) = E((\hat{X} - X)^* Q (\hat{X} - X)) \quad (2.62)$$

を最小にする推定量である。但し、ここで $\|X\|_Q = (X^* Q X)^{\frac{1}{2}}$ である。

(証明) 系 2.5 より、 $Q^{\frac{1}{2}} A_{SR}^{\sigma} X$ は、 $Y = Q^{\frac{1}{2}} X$ の最小分散推定量である。これより $\hat{X} = A_{SR}^{\sigma} X$ は、

$$\begin{aligned} E(\|Q^{\frac{1}{2}} \hat{X} - Q^{\frac{1}{2}} X\|^2) \\ = E((\hat{X} - X)^* Q (\hat{X} - X)) \end{aligned}$$

を最小化する。

(証明終)

以上、線形制御システムにおける σ -逆行列の役割を明らかにした。Moore-Penrose 逆行列が、最小二乗推定あるいは、最良線形不偏推定を与える行列である σ に対して、 σ -逆行列は、最小分散推定を与える行列である。

なお、推定問題の双対問題である制御問題における σ -逆行列の役割も考えることができる。いま、次のような線形制御システムを考える。

$$\dot{x} = A x \quad (2.63)$$

ここで、 A は既知の $m \times n$ 行列、 x は制御ベクトル、 $\dot{x}$ は被制御ベクトルである。いま、被制御ベクトルの目標値を b とする。このとき、制御ベクトル x に対する被制御ベクトル $\dot{x}$ と目標値 b との間の誤差は、

$$\|\dot{x} - b\|^2 = \|A x - b\|^2 \quad (2.64)$$

で評価することができる。

また、制御ベクトル x による制御に要するコストが、ベクトル x のノルム $\|x\|^2$ に比例するものとする。このとき、目標値からの誤差と制御に要するコストを考慮して、

$$\|A x - b\|^2 + \sigma^2 \|x\|^2 \quad (2.65)$$

の最小化を線形制御システムの最適化基準とする。

このとき、(2.65)を最小化する最適制御ベクトル $\bar{x}_b$ は、行列 A の σ -逆行列 A^σ と目標値 b とを用いて、次のように与えられる。

$$\bar{X}_b = A^T b$$

(2.66)

目標値からの誤差の評価 (2.64) 及び 制御のためのコストの評価に、それぞれ、正値エルミート行列を用いたノルムの二乗を採用した場合は、(2.66) の逆行列を、重み付きの逆行列に置き換えた形に一般化される。

第3章 σ -逆行列のスペクトル特性

3.1 まえがき

σ -逆行列は、任意の行列に対して二つの行列方程式により代数的に定義される。本章では、 σ -逆行列の数学的構造をより詳しく解明するために、 σ -逆行列のスペクトル特性を明らかにする。

σ -逆行列は一般の直方行列に対して定義されている。このため、従来の正規行列に対するスペクトル理論をそのまま適用することはできない。そこで、直方行列のためのスペクトル理論をはじめに3.2で導入する。この理論は、R. Penrose, C. Lanzas, M. R. Hestenes らにより展開されたものであり、適用範囲の広い理論である。この理論を用いて、3.3で σ -逆行列のスペクトル特性を明らかにする。更にその応用として、 σ -逆行列の振動の問題を3.4で扱う。

3.2 直方行列のスペクトル理論

A を n 次正規行列とする。 A の相異なる固有値を $\{\lambda_i : i=1, \dots, k\}$ とおく。このとき、行列 A は次のようにスペクトル表現することができる。

$$A = \sum_{i=1}^k \lambda_i P_i \quad (3.1)$$

但し、 P_i は次のような正射影である。

$$P_i P_j = 0, \quad i \neq j \quad (3.2)$$

$$\sum_{i=1}^k P_i = I \quad (3.3)$$

本節では、 Π の正規行列に対するスペクトル理論を一般の直交行列に拡張する。はじめに、特異値分解の拡張である UDV^* 分解について明らかにする。

[定理 3. A] A を階数 r の $m \times n$ 行列とする。 $\alpha(A) = \{\alpha_1, \alpha_2, \dots, \alpha_r\}$ を行列 A の特異値とし、次式を満たすものとする。

$$\alpha_1 \geq \alpha_2 \geq \dots \geq \alpha_r > 0 \quad (3.4)$$

さらに、 $d(A) = \{d_1, d_2, \dots, d_r\}$ を次式を満たす任意の複素数とする。

$$|d_i| = \alpha_i, \quad i = 1, \dots, r \quad (3.5)$$

$\{u_1, u_2, \dots, u_r\}$ を行列 AA^* の固有ベクトルとする。即ち、

$$AA^* u_i = \alpha_i^2 u_i, \quad i = 1, \dots, r \quad (3.6)$$

$$(u_i, u_j) = \delta_{ij}^{(\text{註})}, \quad i, j = 1, \dots, r \quad (3.7)$$

さらに、 $\{v_1, v_2, \dots, v_r\}$ を次のように定める。

$$v_i = \frac{1}{d_i} A^* u_i, \quad i = 1, \dots, r \quad (3.8)$$

このとき、 $\{v_1, v_2, \dots, v_r\}$ は、行列 A^*A の固有ベクトルとなり、

(註) δ_{ij} は Kronecker の δ である。即ち、 $\delta_{ii} = 1$ で、 $i \neq j$ ならば、 $\delta_{ij} = 0$ である。

正規直交系をなす。即ち、

$$A^* A v_i = \alpha_i^2 v_i, \quad i = 1, \dots, r \quad (3.9)$$

$$(v_i, v_j) = \delta_{ij}, \quad i, j = 1, \dots, r \quad (3.10)$$

更に、

$$u_i = \frac{1}{d_i} A v_i, \quad i = 1, \dots, r \quad (3.11)$$

が成り立つ。

逆に $\{v_1, v_2, \dots, v_r\}$ を (3.9) 及び (3.10) を満たすべしとし、 $\{u_1, u_2, \dots, u_r\}$ を (3.11) で定義すると、 $\{u_1, u_2, \dots, u_r\}$ は、(3.6), (3.7) 及び (3.8) を満足する。

(3.9) は、(3.5), (3.6) 及び (3.8) より証明できる。即ち、

$$\begin{aligned} A^* A v_i &= \frac{1}{d_i} A^* A A^* u_i \\ &= d_i A^* u_i \\ &= |d_i|^2 v_i \\ &= \alpha_i^2 v_i \end{aligned}$$

また、(3.10) は、次のようにして証明できる。

$$\begin{aligned}
 (v_i, v_j) &= \frac{1}{d_i d_j} (A^* u_i, A^* u_j) \\
 &= \frac{1}{d_i d_j} (A A^* u_i, u_j) \\
 &= \frac{d_i}{d_j} (u_i, u_j) \\
 &= \delta_{ij}
 \end{aligned}$$

(3.11) は, (3.8) と (3.6) とから容易に求まる。

この定理 3. A から, つぎの定理が導かれる。

[定理 3. B] A を階数 r の $m \times n$ 行列とする。 $d(A) = \{d_1, d_2, \dots, d_r\}$ は, 定理 3. A と同じものとする。このとき, 次式を満たす m 次のユニタリ行列 U と n 次のユニタリ行列 V が存在する。

$$D = U^* A V = \left[\begin{array}{cc|c} d_1 & 0 & \vdots \\ & \ddots & 0 \\ 0 & d_r & \vdots \\ \hline & 0 & 0 \end{array} \right] \quad (3.12)$$

(証明) $\{u_1, u_2, \dots, u_r\}$ と $m-r$ 個のベクトル $\{u_{r+1}, \dots, u_m\}$ を加えて, m 次ベクトル空間の正規直交基底を作る。但し, $\{u_{r+1}, \dots, u_m\}$ は次式を満たす。

$$A^* u_i = 0, \quad i = r+1, \dots, m \quad (3.13)$$

行列 U を次のように定義する。

$$U = [u_1 \cdots u_r u_{r+1} \cdots u_m] \quad (3.14)$$

U は m 次のユニタリ行列である。同様に、 $\{v_1, v_2, \dots, v_r\}$ に $n-r$ 個のベクトル $\{v_{r+1}, \dots, v_n\}$ を加えて、 n 次元ベクトル空間の正規直交基底を作る。但し、 $\{v_{r+1}, \dots, v_n\}$ は次式を満たす。

$$A v_j = 0, \quad j = r+1, \dots, n \quad (3.15)$$

行列 V を次のように定義する。

$$V = [v_1 \cdots v_r v_{r+1} \cdots v_n] \quad (3.16)$$

V は n 次のユニタリ行列である。

この U と V に対して、行列

$$D = U^* A V = [d_{ij}] \quad (3.17)$$

は、次式を満たす。

$$d_{ij} = u_i^* A v_j = 0, \quad i > r \text{ または } j > r \quad (3.18)$$

また、 $i, j = 1, \dots, r$ に対しては、

$$d_{ij} = u_i^* A v_j$$

$$= \frac{1}{d_j} u_i^* A A^* u_j$$

$$= d_j u_i^* u_j$$

$$= d_j \delta_{ij}. \quad (3.19)$$

(3.18) と (3.19) より、行列 D が (3.12) を満足することが示された。

(証明終)

この定理 3. B より、行列 A は次のようにユニタリ行列 U , V と対角行列 D を用いて分解されることが分かる。

$$A = U D V^* \quad (3.20)$$

この分解を $U D V^*$ 分解と呼ぶ。この分解は、特異値分解の拡張となっている。

次に、部分等長作用素について、その性質を明らかにしておく。

n 次元ベクトル空間から m 次元ベクトル空間への線形作用素 U が、その零空間の直交補空間において、ノルムを保存するとき、即ち、

$$\|Ux\| = \|x\| \quad \text{for all } x \in N(U)^\perp \quad (3.21)$$

が成り立つとき、 U を部分等長作用素という。

部分等長作用素の幾つかの性質を次の定理に示す。

[定理 3. C] U を $m \times n$ 行列とする。次の 8 つの命題は同値である。

- (a) U は部分等長作用素である。
- (a*) U^* は部分等長作用素である。
- (b) U^*U は正射影である。
- (b*) UU^* は正射影である。
- (c) $UU^*U = U$
- (c*) $U^*UU^* = U^*$
- (d) $U^* = U^+$
- (e) U^+ は部分等長作用素である。

以上の準備の下に、正規行列のスペクトル理論の直方行列への拡張を、次の定理に示す。

[定理 3. D] A を階数 r の $m \times n$ 行列とする。 $d(A) = \{d_1, d_2, \dots, d_r\}$ は定理 3. A と同じものとする。このとき、次式を満たす r 個の部分等長作用素 $\{E_i : i = 1, \dots, r\}$ が存在する。

$$E_i E_j^* = 0, \quad E_i^* E_j = 0,$$

$$1 \leq i \neq j \leq r \quad (3.22)$$

$$E_i E_i^* A = A E_i^* E_i, \quad i=1, \dots, r \quad (3.23)$$

但し、 E は次式で与えられる。

$$E = \sum_{i=1}^r E_i \quad (3.24)$$

E 自身も部分等長作用素である。そして、行列 A は、次のようにして表わすことができる。

$$A = \sum_{i=1}^r d_i E_i \quad (3.25)$$

この定理の部分等長作用素 $\{E_i : i=1, \dots, r\}$ は、定理 3. B の UDV^* 分解から求めることができる。(3.20) の行列 A の UDV^* 分解を $\{u_1, u_2, \dots, u_r\}$ 及び $\{v_1, v_2, \dots, v_r\}$ を用いて表現すると、

$$A = \sum_{i=1}^r d_i u_i v_i^* \quad (3.26)$$

となる。ここで、

$$E_i = u_i v_i^*, \quad i=1, \dots, r \quad (3.27)$$

とおくと、(3.25) の表現が得られる。(3.27) の E_i が部分等長作用素であることは、定理 3. C, (c) を調べればよい。即ち、

$$E_i E_i^* E_i$$

$$\begin{aligned}
&= u_i v_i^* (u_i v_i^*)^* u_i v_i^* \\
&= u_i (v_i^* v_i) (u_i^* u_i) v_i^* \\
&= u_i v_i^* \\
&= E_i
\end{aligned} \tag{3.28}$$

となるから、 E_i は部分等長作用素である。 E_i が (3.22), (3.23) を満足することも容易に示すことができる。また、(3.24) の行列 E が部分等長作用素であることも、定理 3.C, (c) を確かめることにより、示すことができる。

このように、直方行列のスペクトル表現は UDV^* 分解を書き換えたものである。正規行列のスペクトル表現 (3.1) における固有値 λ_i に対応したものが d_i であり、正射影 P_i に対応したものが、部分等長作用素 E_i であることが分かる。

3.3. σ -逆行列のスペクトル特性

[定理 3.1] A を階数 r の $m \times n$ 行列とする。行列 $D, U, \text{ および } V$ を定理 3.B と同じものとする。このとき、行列 A の σ -逆行列は、次のように与えられる。

$$A^\sigma = V D^\sigma U^* \tag{3.29}$$

ここで D^σ は決式のような $n \times m$ の対角行列である。

$$D^\sigma = \left[\begin{array}{ccc|cc} \frac{d_1}{d_1 \bar{d}_1 + \sigma^2} & & 0 & & \\ & \ddots & & & \\ 0 & & \frac{dr}{dr \bar{r} + \sigma^2} & & \\ \hline & & & 0 & \\ & 0 & & & 0 \end{array} \right] \quad (3.30)$$

(証明) (3.29) は (3.20) と系 2.1, (f) より明らかである。\$D^\sigma\$ は、次のようにして求めることができる。

$$D^\sigma = D^* (D D^* + \sigma^2 I)^{-1}$$

$$= D^* \left[\begin{array}{ccc|ccc} d_1 \bar{d}_1 + \sigma^2 & & 0 & & & \\ & \ddots & & & & \\ 0 & & dr \bar{r} + \sigma^2 & & & \\ \hline & & & \sigma^2 & & \\ & 0 & & & 0 & \\ & & & & 0 & \sigma^2 \end{array} \right]$$

$$= \left[\begin{array}{ccc|cc} \frac{d_1}{d_1 \bar{d}_1 + \sigma^2} & & 0 & & \\ & \ddots & & & \\ 0 & & \frac{dr}{dr \bar{r} + \sigma^2} & & \\ \hline & & & 0 & \\ & 0 & & & 0 \end{array} \right]$$

(証明終)

比較として、定理 3.1 に対応する Moore-Penrose 逆行列 A^+ の表現を次に示す。

$$A^+ = V D^+ U^* \quad (3.31)$$

ここで、 D^+ は次式で与えられる。

$$D^+ = \left[\begin{array}{cc|cc} \frac{1}{d_1} & & 0 & \\ & \ddots & & 0 \\ & 0 & \frac{1}{d_r} & \\ \hline & & & 0 \\ 0 & & & 0 \end{array} \right] \quad (3.32)$$

(3.30) と (3.32) を比較すると分かるように、 σ -逆行列 A^σ において、 $\sigma \rightarrow 0$ とすると、Moore-Penrose 逆行列 A^+ となることが分かる。

ここで、 σ -逆行列 A^σ のスペクトル表現のために次のスカラー関数を定義する。

$$g(z; \sigma) \equiv \frac{\bar{z}}{z \bar{z} + \sigma^2}, \quad \sigma > 0 \quad (3.33)$$

この関数 $g(z; \sigma)$ は、 σ をパラメータにもつ、スカラー関数である。このスカラー関数 $g(z; \sigma)$ を用いて、 σ -逆行列 A^σ のスペクトル表現が次の定理で与えられる。

[定理 3.2] A を階数 r の $m \times n$ 行列とする。 $d(A) = \{d_1, d_2, \dots, d_r\}$ と $u = \{E_i : i=1, \dots, r\}$ はそれぞれ定理 3.1, 定理 3.2 と同じも

のとする。このとき、行列 A の σ -逆行列 A^σ は次のように表わされる。

$$A^\sigma = \sum_{i=1}^r g(d_i; \sigma) E_i^* \quad (3.34)$$

但し、 $g(z; \sigma)$ は、(3.33) のスカラー関数である。

(証明) 定理 3. D および定理 3. I より明らかである。

(証明終)

比較のために Moore-Penrose 逆行列 A^+ のスペクトル表現を次に示す。

$$A^+ = \sum_{i=1}^r \frac{1}{d_i} E_i^* \quad (3.35)$$

(3.34) と (3.35) から明らかなように、Moore-Penrose 逆行列のスペクトル特性は、スカラー関数 $1/z$ で与えられるのに対して、 σ -逆行列のスペクトル特性は、スカラー関数 $g(z; \sigma)$ により与えられる。

z が実数の場合の $g(z; \sigma)$ を図 3. 1 に示す。図 3. 1 では、 σ が 1. 0, 0.5, 0.25 の場合について示した。この図から明らかなように、 σ が零に近づくと、 $g(z; \sigma)$ は $1/z$ に近づき、Moore-Penrose 逆行列のスペクトル特性に近づく。但し、 $g(z; \sigma)$ の $1/z$ への収束は、一様なものではない。 $z=0$ の近傍においては、 $g(z; \sigma)$ と $1/z$ との間には、大きな差が存在している。このように、 σ -逆行列 A^σ と Moore-Penrose 逆行列 A^+ との間の類似性及び相違性は、このスペクトル特性を与えるスカラー関数 $g(z; \sigma)$ と $1/z$ に集約されてくる。

スカラー関数 $g(z; \sigma)$ は $1/z$ と異なり、任意の $\sigma > 0$ に対して、連続で有界な関数である。この性質が、 σ -逆行列 A^σ 自身にも反映している。Moore-Penrose 逆行列 A^+ は、行列 A の関数として見たとき、必ずしも

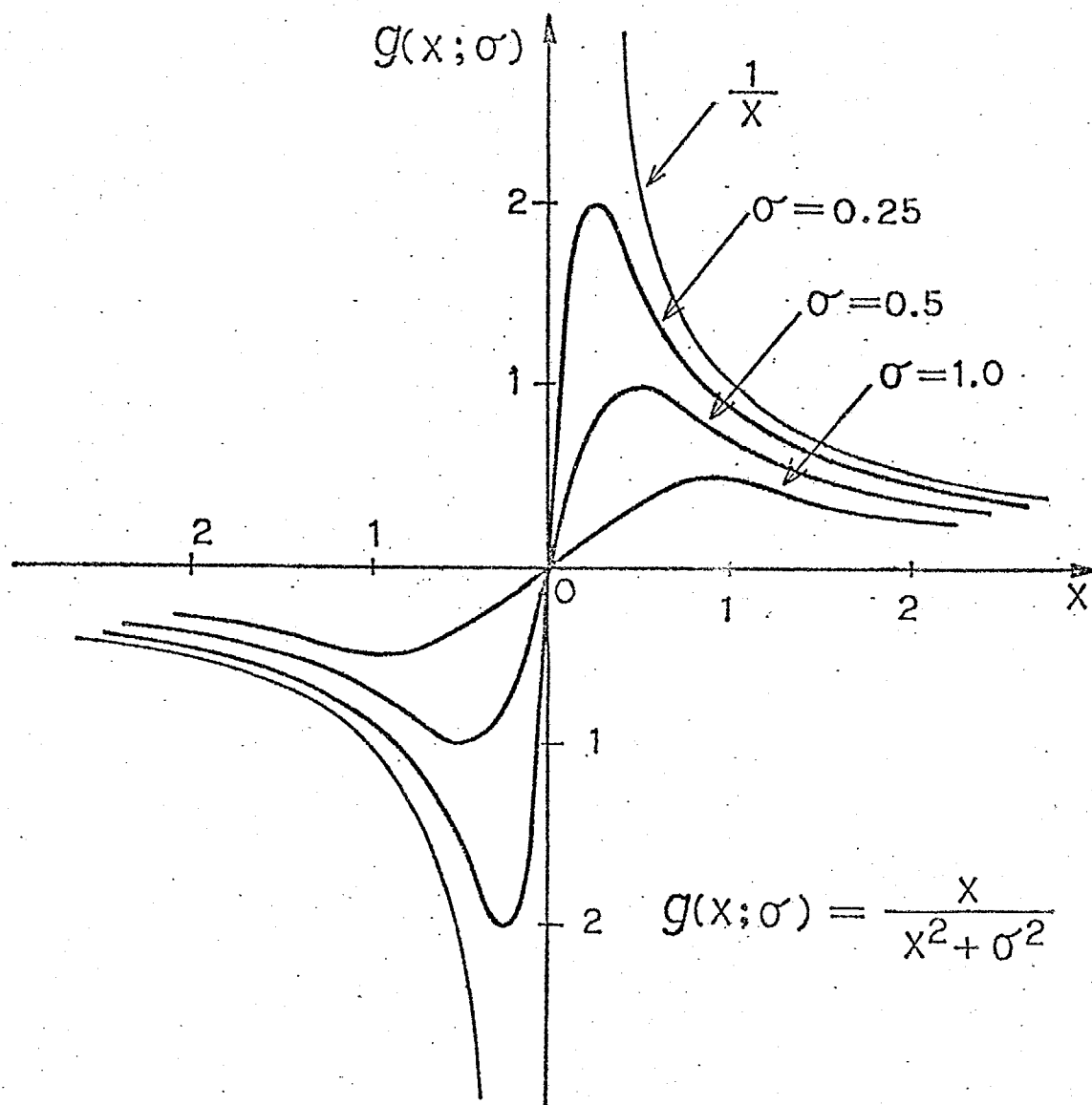


図 3.1 σ -逆行列のスペクトル特性

連続関数ではない。これに対して、 σ -逆行列 A^σ は、行列 A の連続な関数となっている。例として、次の行列 A を考えてみよう。

$$A = \begin{bmatrix} 1 & 0 \\ 0 & \varepsilon \end{bmatrix} \quad (3.36)$$

この行列 A の σ -逆行列 A^σ は、

$$A^\sigma = \begin{bmatrix} \frac{1}{1+\sigma^2} & 0 \\ 0 & \frac{\varepsilon}{\varepsilon^2+\sigma^2} \end{bmatrix} \quad (3.37)$$

となる。一方、行列 A の Moore-Penrose 逆行列 A^+ は、次のようになる。

$$A^+ = \begin{cases} \begin{bmatrix} 1 & 0 \\ 0 & 1/\varepsilon \end{bmatrix}, & \varepsilon \neq 0 \text{ の場合} \\ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, & \varepsilon = 0 \text{ の場合} \end{cases} \quad (3.38)$$

$\varepsilon \rightarrow 0$ のとき、 σ -逆行列 A^σ の $(2,2)$ 成分 $\varepsilon/(\varepsilon^2+\sigma^2)$ は、零に近づき $\varepsilon=0$ において連続性が保たれてゐる。これに反して、Moore-Penrose 逆行列の場合には、 $\varepsilon=0$ において、 $(2,2)$ 成分の連続性が保たれてゐない。

連続性の他にスカラー関数 $g(\varepsilon; \sigma)$ のもつ、もう一つの特徴、有界性に関連して、次の系が成り立つ。

[系 3.1] 行列 A の σ -逆行列 A^σ のノルム^(注) に関して、次の不等式が成り立つ。

$$\|A^\sigma\| \leq \frac{1}{2\sigma} \quad (3.39)$$

(証明) 定理 3.2 より、次式が得られる。

(注) 行列のノルムとして、スペクトルノルムを用いる。即ち、

$$\|A\| = \max \{ \sqrt{\lambda} : \lambda \text{ は } A^*A \text{ の固有値} \}$$

$$\|A^\sigma\| = \max_i |g(d_i; \sigma)| \quad (3.40)$$

また、スカラー関数 $g(z; \sigma)$ に関しては、次の不等式が成り立つ。

$$|g(z; \sigma)| = \frac{|z|}{|z|^2 + \sigma^2} \leq \frac{1}{2\sigma} \quad (\text{等号は } |z| = \sigma \text{ のとき}) \quad (3.41)$$

(3.40) と (3.41) より、(3.39) は明らかである。

(証明終)

＝のように、 σ -逆行列 A^σ は、行列 A の連続な関数であり、しかも有界性が保証されている。このことは、 σ -逆行列 A^σ が安定した、実現しやすい数学的構造をもっていることを示している。

3.4 σ -逆行列の摂動

本節では、 σ -逆行列の摂動、特に σ に関する摂動の問題を扱う。

[定理 3.3] $\sigma = \sigma_0 + \delta$ とする。このとき、次式が成り立つ。

$$A^\sigma = A^{\sigma_0} + (\sigma - \sigma_0) B(\sigma_0) + O(\delta^2) \quad (\text{注}) \quad (3.42)$$

但し、行列 $B(\sigma)$ は、 σ を $n \times n$ とする行列であり、次式で与えられる。

$$B(\sigma) \equiv \sum_{i=1}^n \frac{-2\bar{d}_i \sigma}{(d_i \bar{d}_i + \sigma^2)^2} E_i^* \quad (3.43)$$

(注) $O(\delta^2)$ は、各成分の位数が δ^2 の行列を意味する。

こゝで、 $d(A) = \{d_1, d_2, \dots, d_r\}$ 及 $u^i \in E_i : i=1, \dots, r$ はそれぞれ、定理 3.A と定理 3.D と同じものである。

(証明) σ -逆行列のスペクトル特性を与えるスカラー関数 $g(z; \sigma)$ を σ についてテーラー展開する。

$$\begin{aligned} g(z; \sigma) &= g(z; \sigma_0 + \delta) \\ &= g(z; \sigma_0) + (\sigma - \sigma_0) \left\{ \frac{d}{d\sigma} g(z; \sigma) \right\}_{\sigma=\sigma_0} \\ &\quad + (\sigma - \sigma_0)^2 \left\{ \frac{d^2}{d\sigma^2} g(z; \sigma) \right\}_{\sigma=\sigma'} \end{aligned} \tag{3.44}$$

但し、 σ' は $\sigma' = \sigma_0 + \theta \delta$ ($0 < \theta < 1$) を満たす実数である。

こゝで、

$$\frac{d}{d\sigma} g(z; \sigma) = \frac{-2\bar{z}\sigma}{(z\bar{z} + \sigma^2)^2} \tag{3.45}$$

であるから、定理 3.2 より、(3.42) が求まる。

(証明終)

定理 3.3 より、 δ が十分に小さい場合には、 σ -逆行列 A^σ に関する次の近似式が成り立つことが分かる。

$$A^\sigma \doteq A^{\sigma_0} + (\sigma - \sigma_0) \cdot B(\sigma_0) \tag{3.46}$$

行列 A の σ -逆行列 A^σ は、線形観測システム (A, σ^2) の最小分散推定を与える行列である。いま、線形観測システムが (A, σ_0^2) から (A, σ^2) に変動したとする。即ち、観測時に加わる雑音レベルが σ_0^2 から σ^2 に変化しただけである。この場合、推定システムも観測システムの変動に適応して変化させなければならぬ。具体的には、用いる σ -逆行列を A^{σ_0} から A^σ に変えなければならぬ。このように σ -逆行列 A^σ を、 σ の変動に従ってその都度、再構成するのは容易でない。

いま、 σ の変動が、 σ_0 を中心として微小なものであるならば (3.46) の近似式を用いることができる。 σ -逆行列 A^{σ_0} と $B(\sigma_0)$ だけを予め用意することにより、近似的に A^σ を実現することができる。この方法により、観測時の雑音レベルの変動に対して適応性のある推定システムを比較的容易に構成することが可能となる。

定理 3.3 において、 $\delta \rightarrow 0$ とおくことにより、 σ -逆行列 A^σ の $n \times n$ マトリクスに関する導関数を求めることができる。即ち、

$$\begin{aligned} \frac{d}{d\sigma} A^\sigma &= \lim_{\sigma' \rightarrow \sigma} \frac{A^{\sigma'} - A^\sigma}{\sigma' - \sigma} \\ &= \lim_{\delta \rightarrow 0} \frac{\{A^\sigma + \delta B(\sigma) + O(\delta^2)\} - A^\sigma}{\delta} \\ &= \lim_{\delta \rightarrow 0} \left\{ B(\sigma) + \frac{1}{\delta} O(\delta^2) \right\} \\ &= B(\sigma) \end{aligned} \quad (3.47)$$

となる。このように行列 $B(\sigma)$ は、 σ -逆行列 A^σ の $n \times n$ マトリクスに関する導関数とみなすことができる。この導関数 $B(\sigma)$ は、行列 A の σ -逆行列 A^σ と Moore-Penrose 逆行列 A^+ とを用いて次のように表わすことができる。

[系 3.2]

$$\begin{aligned}\frac{d}{d\sigma} A^\sigma &= B(\sigma) \\ &= -2\sigma A^\sigma (A^+)^* A^\sigma\end{aligned}\tag{3.48}$$

(証明) (3.45) は、次のように表わすことが出来る。

$$\frac{-2\bar{z}\sigma}{(z\bar{z}+\sigma^2)^2} = -2\sigma \frac{g(z;\sigma)^2}{\bar{z}}\tag{3.49}$$

ここで、部分等長作用素 E_i が (3.22) 及び (3.28) を満足することに注意すると (3.48) が得られる。即ち、

$$\begin{aligned}B(\sigma) &= \sum_{i=1}^r -2\sigma \frac{g(d_i;\sigma)^2}{d_i} E_i^* \\ &= -2\sigma \left(\sum_{i=1}^r g(d_i;\sigma) E_i^* \right) \left(\sum_{i=1}^r \frac{1}{d_i} E_i \right) \left(\sum_{i=1}^r g(d_i;\sigma) E_i^* \right) \\ &= -2\sigma A^\sigma \left(\sum_{i=1}^r \frac{1}{d_i} E_i^* \right)^* A^\sigma \\ &= -2\sigma A^\sigma (A^+)^* A^\sigma\end{aligned}\tag{3.50}$$

となる。

(証明終)

第4章 分割行列の σ -逆行列

4.1 まえがき

線形観測システムが規模が大きくなると、その推定システムを求めるには、高次元の行列の σ -逆行列が必要となる。 σ -逆行列を求めるには、行列の逆行列演算が必要であり、行列の次元が高くなると急速に演算ステップ数が増大する。そこで、できるだけ高次元の逆行列演算を避けるために、 σ -逆行列の逐次計算を行なう必要が生じる。そのための準備として、本章では、分割行列の σ -逆行列の陽表現を求める。

また、分割行列の σ -逆行列の陽表現は、最小分散推定の逐次推定システムを構成するために必要となるものである。4.3では、最小分散推定の逐次推定の問題を、最小二乗推定の場合と比較しながら検討する。

4.2 分割行列の σ -逆行列

はじめに次のような行列の分割を掲かう。

$$A = [U : V] \quad (4.1)$$

ここで、 A は、 $m \times n$ 行列である。また、 U は、 $m \times r$ 行列、 V は $m \times (n-r)$ 行列とする。この分割行列に対して、次の定理が成り立つ。

[定理 4.1]

$$[U:V]^{\sigma} = \begin{bmatrix} U^{\sigma} - U^{\sigma} V K \\ \hline K \end{bmatrix} \quad (4.2)$$

==で

$$K = (V^T S V + \sigma^2 I)^{-1} V^T S \quad (4.3)$$

$$S' = I - U U^{\sigma} \quad (4.4)$$

である。

(証明) 定理 2.1 より, $A^{\sigma} = A^T (A A^T + \sigma^2 I)^{-1}$ (注) であるから, 行列 X が, A^{σ} であることを示すには,

$$X (A A^T + \sigma^2 I) = A^T \quad (4.5)$$

となることを示せばよい。(4.1) より,

$$A A^T + \sigma^2 I = U U^T + V V^T + \sigma^2 I$$

である。この右辺を(4.2)の右辺に掛ける。

$$\begin{aligned} & X (U U^T + V V^T + \sigma^2 I) \\ &= \begin{bmatrix} U^{\sigma} (I - V K) (U U^T + V V^T + \sigma^2 I) \\ \hline K (U U^T + V V^T + \sigma^2 I) \end{bmatrix} \end{aligned} \quad (4.6)$$

(注) 本章以上では、係数体として実数を用いるので、行列 A の随伴行列 A^* は、転置行列 A^T に置き換えられる。

こゝで、

$$\begin{aligned} & S(UU^T + \sigma^2 I) \\ &= (I - UU^T)(UU^T + \sigma^2 I) \\ &= UU^T + \sigma^2 I - UU^T \\ &= \sigma^2 I \end{aligned} \tag{4.7}$$

に注意すると、 $K(UU^T + VV^T + \sigma^2 I)$ は次のようになる。

$$\begin{aligned} & K(UU^T + VV^T + \sigma^2 I) \\ &= KVV^T + (V^T S V + \sigma^2 I)^{-1} V^T S (UU^T + \sigma^2 I) \\ &= (V^T S V + \sigma^2 I)^{-1} V^T S V V^T + \sigma^2 (V^T S V + \sigma^2 I)^{-1} V^T \\ &= (V^T S V + \sigma^2 I)^{-1} (V^T S V + \sigma^2 I) V^T \\ &= V^T \end{aligned} \tag{4.8}$$

また、 $U^T(I - VK)(UU^T + VV^T + \sigma^2 I)$ は次のようになる。

$$\begin{aligned} & U^T(I - VK)(UU^T + VV^T + \sigma^2 I) \\ &= U^T(UU^T + VV^T + \sigma^2 I) - U^T V V^T \\ &= U^T(UU^T + \sigma^2 I) \\ &= U^T \end{aligned} \tag{4.9}$$

(4.8) と (4.9) の結果を (4.6) に代入することにより, (4.5) が成り立つことは明らかである。

(証明終)

なお, この定理において, 行列 $V^T S V + \sigma^2 I$ は正値対称行列であるから, その逆行列が常に存在することに注意する。いま, 行列 U の σ -逆行列 U^σ が既に求められている場合には, (4.3) の行列 K を求めることにより, 行列 A の σ -逆行列 A^σ が得られる。行列 K は, $n-r$ 次元の逆行列演算により求めることができるので, 直接に行列 A から求めるよりは, 容易に σ -逆行列 A^σ を求めることができる。

この定理 4.1 において, 特に $r = n-1$ とした場合が次の系 4.1 である。

[系 4.1]

$$[U : a]^\sigma = \begin{bmatrix} U^\sigma - b h^T \\ \hline h^T \end{bmatrix} \quad (4.10)$$

ここで,

$$b = U^\sigma a \quad (4.11)$$

$$h = \frac{S a}{a^T S a + \sigma^2} \quad (4.12)$$

$$S = I - U U^\sigma \quad (4.13)$$

である。

(証明) 定理 4.1 において, 行列 V をベクトル a に置き換えること

により、容易に求められる。

(証明終)

次に、下のような分割行列について考える。

$$A = \begin{bmatrix} U \\ -\frac{U}{V} \end{bmatrix} \quad (4.14)$$

ここで、 A は $m \times n$ 行列である。また、 U は $r \times n$ 行列、 V は $(m-r) \times n$ 行列である。

この分割行列に対して、次の定理が成り立つ。証明は、定理 4.1 と系 2.1, (c) より明らかである。

[定理 4.2]

$$\begin{bmatrix} U \\ -\frac{U}{V} \end{bmatrix}^{\sigma} = \left[(I - KV)U^{\sigma} \mid K \right] \quad (4.15)$$

ここで、

$$K = SV^T (VSV^T + \sigma^2 I)^{-1} \quad (4.16)$$

$$S = I - U^{\sigma}U \quad (4.17)$$

である。

さらに、定理 4.2 において、 $r=m-1$ とおくことにより、次の系が求まる。

[系 4.2]

$$\begin{bmatrix} U \\ -a^T \end{bmatrix}^\sigma = [(I - ha^T) U^\sigma \mid h] \quad (4.18)$$

==で、

$$h = \frac{S a}{a^T S a + \sigma^2} \quad (4.19)$$

$$S = I - U^\sigma U \quad (4.20)$$

である。

4.3 最小分散推定の逐次推定

(4.14) の行列の分割は、線形観測システム (A, σ^2) を次のように二つのシステムに分割して考えることに対応している。

$$\begin{cases} \mathbb{Z}_1 = U \mathbb{X} + W_1 \\ \mathbb{Z}_2 = V \mathbb{X} + W_2 \end{cases} \quad (4.21)$$

==で、

$$\mathbb{Z} = \begin{bmatrix} \mathbb{Z}_1 \\ \mathbb{Z}_2 \end{bmatrix}, \quad W = \begin{bmatrix} W_1 \\ W_2 \end{bmatrix} \quad (4.22)$$

である。

はじめに観測 U が行なわれて、観測ベクトル z_1 が得られたとする。このときのベクトル x の最小分散推定量 $\hat{x}_1$ は、次のように与えられる。

$$\hat{x}_1 = U^T z_1 \quad (4.23)$$

次に、新たに観測 V が行なわれて、観測ベクトル z_2 が追加されたとする。このとき、先に求めた推定量 $\hat{x}_1$ を修正して、 z_1 と z_2 を用いたベクトル x の最小分散推定量 $\hat{x}$ を求めなければならぬ。 $\hat{x}$ は、次式で与えられる。

$$\hat{x} = \begin{bmatrix} U \\ V \end{bmatrix}^T \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \quad (4.24)$$

== で、定理 4.2 を用いると、

$$\hat{x} = (I - KV)U^T z_1 + K z_2 \quad (4.25)$$

となる。但し、ここで K は定理 4.2 と同じものである。更に、この式を整理すると次のようになる。

$$\hat{x} = \hat{x}_1 + K(z_2 - V\hat{x}_1) \quad (4.26)$$

この式が、新たな観測ベクトル z_2 の追加による $\hat{x}_1$ の修正の方法を与えている。右辺の第二項が、観測ベクトル z_2 による修正項になっている。

次に観測データが一つずつスカラーとして得られるような観測システムを考えよう。第 k 次の観測システムが次のように与えられる。

$$r_k = a_k^T x + v_k, \quad k = 1, \dots, n \quad (4.27)$$

ここで、 a_k は、行列 A の第 k 行ベクトルである。

いま、第 k 次までの観測が行われたとしよう。このときのベクトル x の最小分散推定量 $\hat{x}_k$ は次式で与えられる。

$$\hat{x}_k = A_k^T z_k \quad (4.28)$$

ここで、

$$A_k = \begin{bmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_k^T \end{bmatrix}, \quad z_k = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_k \end{bmatrix} \quad (4.29)$$

である。

いま、次のような観測が新たに行われたとする。

$$z_{k+1} = a_{k+1}^T x + v_{k+1} \quad (4.30)$$

このとき、ベクトル x の最小分散推定量 $\hat{x}_{k+1}$ は

$$\hat{x}_{k+1} = A_{k+1}^T z_{k+1} \quad (4.31)$$

となる。ここで、

$$A_{k+1} = \begin{bmatrix} A_k \\ a_{k+1}^T \end{bmatrix}, \quad (4.32)$$

$$\mathbb{Z}_{k+1} = \begin{bmatrix} \mathbb{Z}_k \\ \zeta_{k+1} \end{bmatrix} \quad (4.33)$$

である。

(4.32) の分割行列に対して、系 4.2 を適用すると、

$$A_{k+1}^\sigma = [(I - h_{k+1} a_{k+1}^T) A_k^\sigma \ ; \ h_{k+1}] \quad (4.34)$$

となる。但し、

$$h_{k+1} = \frac{(I - A_k^\sigma A_k) a_{k+1}}{a_{k+1}^T (I - A_k^\sigma A_k) a_{k+1} + \sigma^2} \quad (4.35)$$

である。

以上のことから、ベクトル $\hat{x}$ の最小分散推定量 $\hat{x}$ の逐次推定システムが次のような形で与えられる。

$$\begin{aligned} \hat{x}_{k+1} &\equiv A_{k+1}^\sigma \mathbb{Z}_{k+1} \\ &= (I - h_{k+1} a_{k+1}^T) A_k^\sigma \mathbb{Z}_k + h_{k+1} \zeta_{k+1} \\ &= \hat{x}_k + h_{k+1} (\zeta_{k+1} - a_{k+1}^T \hat{x}_k) \end{aligned} \quad (4.36)$$

ここで、

$$\hat{x}_0 = 0$$

である。

(4.36)の右辺第一項は、第 k 次までの観測によるベクトル x の最小分散推定量である。第二項が、新しい観測データ z_{k+1} による修正項である。 $a_{k+1}^T \hat{x}_k$ は、ベクトル x を $\hat{x}_k$ と推定した場合の第 $k+1$ 次の観測の推定量に当る。その推定誤差が、 $z_{k+1} - a_{k+1}^T \hat{x}_k$ である。新しい推定量 $\hat{x}_{k+1}$ は、 $\hat{x}_k$ に、この推定誤差 $z_{k+1} - a_{k+1}^T \hat{x}_k$ に比例した修正項を加えることにより得られる。修正ベクトル h_{k+1} は、(4.35)で与えられるが、このベクトルは観測データに依存せずに定まるものである。

なお、修正ベクトル h_{k+1} を算出するには、 $I - A_k^T A_k$ という行列を求めなければならぬが、これは、次のような形で逐次に求めることができる。

$$I - A_k^T A_k = (I - h_k a_k^T)(I - A_{k-1}^T A_{k-1}) \quad (4.37)$$

(4.27)のような逐次観測システムの最小二乗推定も、(4.36)と全く同じ形式で逐次推定することができる。但し、修正ベクトル h_{k+1} の構成方法に若干の違いがある。比較のために、最小二乗推定の場合の修正ベクトル h_{k+1} を次に示す。

$$h_{k+1} = \begin{cases} \frac{(I - A_k^+ A_k) a_{k+1}}{a_{k+1}^T (I - A_k^+ A_k) a_{k+1}}, & \text{の場合, } (I - A_k^+ A_k) a_{k+1} \neq 0 \\ \frac{A_k^+ A_k^{T+} a_{k+1}}{1 + a_{k+1}^T A_k^+ A_k^{T+} a_{k+1}}, & \text{その他の場合.} \end{cases} \quad (4.38)$$

観測データがスカラー量として逐次に得られる場合の逐次推定システムを明らかにした。つぎに、より一般的な逐次観測過程の場合の逐次推定問題について考察する。

いま、観測過程が次のように N 個の観測システムに分割されておけるとする。

$$x_i = A_i x + w_i, \quad i=1, \dots, N \quad (4.39)$$

ここで、未知ベクトル x は、確率ベクトルであり、その平均及び共分散が次のように与えられているとする。

$$E(x) = x_0 \quad (4.40)$$

$$\text{Cov}(x - x_0) = S \quad (4.41)$$

また、観測誤差ベクトル w_i は、ベクトル x に対して独立であり、しかも各 i に対して互いに独立な確率ベクトルである。その平均及び共分散は次のように与えられているとする。

$$E(w_i) = 0, \quad i=1, \dots, N \quad (4.42)$$

$$\text{Cov}(w_i) = \sigma^2 R_i, \quad i=1, \dots, N \quad (4.43)$$

このとき、ベクトル x の最小分散推定量 $\hat{x}$ は、 N 個の観測ベクトル $\{x_i : i=1, \dots, N\}$ を同時に用いて、直接に求めることも可能である。しかし、観測ベクトルの次元が高くなると、高次元の逆行列演算が必要となるため、必ずしも容易に推定量を求めることはできない。そこで、観測ベクトル $x_i, i=1, \dots, N$ を順次に用いて、最小分散推定量 $\hat{x}$ を逐次推定する必要がある。

はいめに、次のような記法を定める。

$$A_{(i)} = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_i \end{bmatrix} \quad (4.44)$$

$$Z(i) = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_i \end{bmatrix} \quad (4.45)$$

$$R(i) = \begin{bmatrix} R_1 & & 0 \\ & R_2 & \\ 0 & & \ddots \\ & & & R_i \end{bmatrix} \quad (4.46)$$

$\hat{X}(i)$: 第 i 段階までの観測ベクトル $\{z_1, z_2, \dots, z_i\}$ を用いたときのベクトル X の最小分散推定量.

$S(i)$: 推定量 $\hat{X}(i)$ と真値 X との間の誤差共分散.

ベクトル X の最小分散推定量 $\hat{X}$ は、次の定理に示すように、 $\hat{X}(i)$ と $S(i)$ とを同時に逐次推定することにより、求めることができる。

[定理 4.3]

$$\hat{X}(i) = \hat{X}(i-1) + A_i^T S_{(i-1)}^{-1} R_i (z_i - A_i \hat{X}(i-1)), \quad (4.47)$$

$$S(i) = (I - A_i^T S_{(i-1)}^{-1} A_i) S(i-1), \quad (4.48)$$

$i=1, \dots, N$

但し、 $\hat{X}_0 = X_0$, $S(0) = S$ である。

(証明) $\hat{X}_{(i-1)}$ は、行列 $A_{(i-1)}$ の σ -逆行列 $A_{(i-1)}^{\sigma} S_{R(i-1)}$ を用いて次のように与えられる。

$$\hat{X}_{(i-1)} = X_0 + A_{(i-1)}^{\sigma} S_{R(i-1)} (Z_{(i-1)} - A_{(i-1)} X_0) \quad (4.49)$$

この $\hat{X}_{(i-1)}$ を (4.47) に代入して整理すると

$$\hat{X}_{(i)} = X_0 + G (Z_{(i)} - A_{(i)} X_0) \quad (4.50)$$

となる。但し、ここで行列 G は、次式である。

$$G = [(I - A_i^{\sigma} S_{R(i-1)} R_i) A_{(i-1)}^{\sigma} S_{R(i-1)} : A_{(i-1)}^{\sigma} S_{R(i-1)}] \quad (4.51)$$

これより、証明すべきことは、

$$G = A_{(i)}^{\sigma} S_{R(i)} \quad (4.52)$$

である。

ここでは、 $A_{(i)}$ の分割行列の σ -逆行列を求めることにより、 $A_{(i)}^{\sigma} S_{R(i)}$ が (4.51) の行列 G となることを示す。

系 2.2 より、

$$A_{(i)}^{\sigma} S_{R(i)} = S^{\frac{1}{2}} (R_{(i)}^{-\frac{1}{2}} A_{(i)} S^{\frac{1}{2}})^{\sigma} R_{(i)}^{-\frac{1}{2}} \quad (4.53)$$

である。ここで

$$U = R_{(i-1)}^{-\frac{1}{2}} A_{(i-1)} S^{\frac{1}{2}} \quad (4.54)$$

$$V = R_i^{-\frac{1}{2}} A_i S^{\frac{1}{2}} \quad (4.55)$$

とおく。

$$A_{(i)}^{\sigma} S_{R(i)} = S^{\frac{1}{2}} \begin{bmatrix} U \\ -V \end{bmatrix}^{\sigma} R_{(i)}^{-\frac{1}{2}} \quad (4.56)$$

となる。分割行列の σ -逆行列に関する定理 4.2 を用いて、(4.56)を変形する。

$$\begin{aligned} & A_{(i)}^{\sigma} S_{R(i)} \\ &= S^{\frac{1}{2}} [(I - KV) U^{\sigma} \mid K] R_{(i)}^{-\frac{1}{2}} \\ &= [S^{\frac{1}{2}} (I - KV) U^{\sigma} R_{(i)}^{-\frac{1}{2}} \mid S^{\frac{1}{2}} K R_{(i)}^{-\frac{1}{2}}] \\ &= [(I - S^{\frac{1}{2}} K R_{(i)}^{-\frac{1}{2}} A_i) A_{(i-1)}^{\sigma} S_{R(i-1)} \mid S^{\frac{1}{2}} K R_{(i)}^{-\frac{1}{2}}] \end{aligned} \quad (4.57)$$

但し、

$$K = Q V^T (V Q V^T + \sigma^2 I)^{-1} \quad (4.58)$$

$$Q = I - U^{\sigma} U \quad (4.59)$$

である。ここで、

$$\begin{aligned} & S^{\frac{1}{2}} Q S^{\frac{1}{2}} \\ &= S^{\frac{1}{2}} (I - U^{\sigma} U) S^{\frac{1}{2}} \end{aligned}$$

$$\begin{aligned}
&= (I - S^{\frac{1}{2}} U^{\sigma} R_{(i-1)}^{-\frac{1}{2}} R_{(i-1)}^{\frac{1}{2}} U S^{-\frac{1}{2}}) S \\
&= (I - A_{(i-1)} S_{(i-1)}^{\sigma} A_{(i-1)}) S \\
&= S_{(i-1)}
\end{aligned} \tag{4.60}$$

となることに注意すると、行列 $S^{\frac{1}{2}} K R_i^{-\frac{1}{2}}$ は次のようになる。

$$\begin{aligned}
&S^{\frac{1}{2}} K R_i^{-\frac{1}{2}} \\
&= S^{\frac{1}{2}} Q V^T (V Q V^T + \sigma^2 I)^{-1} R_i^{-\frac{1}{2}} \\
&= S_{(i-1)} A_i^T (A_i S_{(i-1)} A_i^T + \sigma^2 R_i)^{-1} \\
&= A_i S_{(i-1)}^{\sigma} R_i
\end{aligned} \tag{4.61}$$

(4.61) を (4.57) に代入することにより、(4.52) が成り立つことが分かる。

(証明終)

定理 4.3 に示すように最小分散相定量の逐次推定システムは、 σ -逆行列を用いて非常に簡明な形で与えられることが分かる。第 i 段階の観測による逐次推定の方法は、(4.47) より分かるように、未知ベクトル X の平均値が分散が、

$$E(X) = \hat{X}_{(i-1)} \tag{4.62}$$

$$\text{Cov}(X - \hat{X}_{(i-1)}) = S_{(i-1)} \tag{4.63}$$

である場合の通常最小分散推定を行なえばよいのである。

この第 i 段階での逐次推定には、前段階でのベクトル X の推定量 $\hat{X}_{(i-1)}$ の他に、その誤差共分散 $S_{(i-1)}$ が必要である。そのために、 $S_{(i)}$ も $\hat{X}_{(i)}$ と同時に求める必要がある。その逐次計算の方法を与えているのが (4.48) である。

第5章 σ -逆行列の計算法

5.1. まえがき

σ -逆行列は、線形観測システムにおける最小分散推定のための数学的道具として与えられた概念である。実用的側面からみた場合に、ある数学的道具が有用であるためには、実際にその道具を得るための計算法が簡単なものでなくてはならない。Moore-Penrose 逆行列は、その代数的構造は非常に整ったものであるが、実際にこの行列を求めることは必ずしも容易ではない。これは第3章で明らかにしたように、行列 A の Moore-Penrose 逆行列 A^+ が行列 A の関数として見たとき、連続な関数でないことから容易に推測することができる。これに対して、 σ -逆行列 A^σ は行列 A の連続関数であり、しかも有界性が保証されている。このことから、 σ -逆行列 A^σ の計算法は、Moore-Penrose 逆行列 A^+ よりも容易であることが予測できる。

本章では σ -逆行列の計算法を、直接法と逐次計算法に分けて、それぞれ 5.2 と 5.3 で明らかにする。また、5.4 では、Moore-Penrose 逆行列の計算のための一つの技法である正則化と σ -逆行列との関連を明らかにする。

5.2 σ -逆行列の計算法 I —直接法—

はじめに、Moore-Penrose 逆行列 A^+ の直接計算法について簡単にふれておく。

行列 A を階数 r の $m \times n$ 行列とする。いま、階数 r が次の条件を満足する場合を考えよう。

$$r = \min(m, n) \quad (5.1)$$

このときには、行列 A の Moore-Penrose 逆行列 A^+ はつきりの二つの式のいずれかにより、容易に計算することができる。

$$(a) \quad A^+ = A^T (A A^T)^{-1}, \quad r = m \text{ の場合} \quad (5.2)$$

$$(b) \quad A^+ = (A^T A)^{-1} A^T, \quad r = n \text{ の場合} \quad (5.3)$$

しかし、一般には、行列 A は (5.1) が成り立つとは必要ない。そこで、行列 A を何らかの方法で最大階数行列に分解しなければならない。最大階数行列への分解とは次のようなものである。

$$A = F G \quad (5.4)$$

ここで、 F と G は、それぞれ階数 r の $m \times r$ 行列と $r \times n$ 行列である。

この分解ができれば、 A^+ は次のようにして求めることができる。

$$A^+ = G^T (G G^T)^{-1} (F^T F)^{-1} F^T \quad (5.5)$$

このように Moore-Penrose 逆行列の計算法は、(i) 最大階数行列への分解 と (ii) 最大階数行列の Moore-Penrose 逆行列の計算 の二つのステップに分けられる。(i) の最大階数行列への分解には、

(a) ガウス消去法による LU 分解

(b) ハウスホルダー変換

(c) グラム-シュミットの直交化

などの方法が考えられる。各方法にはそれぞれ一長一短があり、決定版といったアルゴリズムはないようである。

いずれの方法を用いても、(i) の最大階数行列への分解を数値計算で実行する場合には、一つの原理的な困難性が存在する。それは、与えられた行列の階数を知ることが安定に実行できない点である。これは、行列の階数という概念そのものに起因するものである。行列の階数は、行列の行ベクトルあるいは列ベクトルの中で互いに独立なベクトルの数の最大数として定義される。数値計算の場合、必然的に数値の丸め誤差が伴うために、ベクトルの独立性の判定を安定に実行することができないのである。

このように、与えられた行列の階数が予め明らかでない場合には、Moore-Penrose 逆行列の数値計算に常に不安定性が存在する。それと比べると、 σ -逆行列の数値計算は、非常に扱い易いものである。

σ -逆行列の計算には、定理 2.1 あるいは系 2.1, (a) の表現式をそのまま用いるがよい。即ち、

$$(a) \quad A^\sigma = A^T (A A^T + \sigma^2 I)^{-1} \quad (5.6)$$

$$(b) \quad A^\sigma = (A^T A + \sigma^2 I)^{-1} A^T \quad (5.7)$$

ここで、逆行列演算の次元を考慮すると、 $m < n$ のときは、(5.6) を用い、 $m \geq n$ のときは (5.7) を用いるのがよい。ここでは、 $m \geq n$ の場合について考える。

(5.7) において、行列 $A^T A + \sigma^2 I$ は、 $\sigma > 0$ であるから、正値対称行列である。正値対称行列の逆行列演算には、コレスキー分解という有用な行列分解方法がある。コレスキー分解とは、次のような分解である。いま、行列

A が正値対称行列ならば

$$A = L L^T \quad (5.8)$$

と分解することができる。ここで、行列 L は、下三角行列^(注)である。

このコレスキー分解の数値アルゴリズムは、普通の消去法で必要となるピボットの選択といった判断操作が必要でなく、非常にコンパクトなものである。下三角行列 L の逆行列は容易に求めることができるから、コレスキー分解を行なうことにより、正値対称行列の逆行列演算は簡易することが可能となる。

コレスキー分解を用いた σ -逆行列の計算法は、次のようなステップに分けることができる。

[σ -逆行列の計算法Ⅰ—直接法—]

1. $U = A^T A + \sigma^2 I$ を求める。
2. コレスキー分解により、 $L L^T = U$ となる下三角行列 L を求める。
3. 下三角行列 L の逆行列 L^{-1} を求める。
4. $A^\sigma = (L^{-1})^T L^{-1} A^T$ により、 A^σ を求める。

この σ -逆行列の計算法に従った計算例を次に示す。

[例 5.1] 次の行列 A の σ -逆行列を求める。

(注) 行列 $A = (a_{ij})$ において、 $i < j$ ならば $a_{ij} = 0$ であるような行列を、下三角行列という。

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}$$

$$\begin{aligned} 1. \quad U &= \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} + \sigma^2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 2+\sigma^2 & 1 \\ 1 & 2+\sigma^2 \end{bmatrix} \end{aligned}$$

2. コレスキー分解により L を求める。

$$L = \begin{bmatrix} \sqrt{2+\sigma^2} & 0 \\ \frac{1}{\sqrt{2+\sigma^2}} & \sqrt{\frac{(3+\sigma^2)(1+\sigma^2)}{(2+\sigma^2)}} \end{bmatrix}$$

3.

$$L^{-1} = \begin{bmatrix} \frac{1}{\sqrt{2+\sigma^2}} & 0 \\ \frac{-1}{\sqrt{(3+\sigma^2)(2+\sigma^2)(1+\sigma^2)}} & \sqrt{\frac{2+\sigma^2}{(3+\sigma^2)(1+\sigma^2)}} \end{bmatrix}$$

4.

$$A^\sigma = (L^{-1})^T L^{-1} A^T$$

$$= \frac{1}{(3+\sigma^2)(1+\sigma^2)} \begin{bmatrix} 2+\sigma^2 & 1+\sigma^2 & -1 \\ -1 & 1+\sigma^2 & 2+\sigma^2 \end{bmatrix}$$

σ -逆行列を求めるサブルーチンのフローチャートリストを図 5.1 に示す。

通常の逆行列を求めるプログラムにおいては、与えられた行列が正則であるか否かを判定するために、判定値を与えなければならぬ。これは、Moore-Penrose 逆行列のプログラムでも同様であり、行列の階数を決定するための判定値を与えなければならぬ。しかし、この判定値の与え方には、明確な基準がない。対象とする行列や、計算過程に加わる丸め誤差の影響などに依存して定められなければならず、実際には、経験的に与えるしかない。

σ -逆行列の計算のためのサブルーチンでは、図 5.1 に示すように、このような判定値の設定は必要としない。その代わりに正の実数 σ を与えれば良い。この σ の値の設定は、対象とする線形観測システムの統計的性質が明らかであれば、一意に定まるものである。仮に、統計的性質が十分に明らかでない場合でも、 σ の統計的意味合いがは、ざりしるるので、具体的な予測の下に、 σ の値を設定することが出来る。

σ -逆行列の計算の安定性の問題について考えてみる。 σ -逆行列を求める上で問題となる演算は、行列 $U = ATA + \sigma^2 I$ の逆行列演算だけである。この行列 U の逆行列演算の安定性を調べるために、行列 U の条件数を求める。いま、行列 ATA の最大固有値 M 、最小固有値 (零も含む) m とすると、行列 U の条件数 $\text{cond}(A)$ は次のようになる。

$$\text{cond}(A) = \frac{M + \sigma^2}{m + \sigma^2} \quad (5.9)$$

条件数は、小さい値である方が、安定した数値計算が行なえる。(5.9) より $\sigma (> 0)$ の存在が、 σ -逆行列の計算の安定性に重要な役割を果たしていることが分かる。行列 ATA は非負行列であるから、最小固有値 m は最悪の場合で

ISN	SOURCE STATEMENT
00001	SUBROUTINE SINV(A,U,SIGMA,MR,NR,NC)
	C THIS ROUTINE CALCULATES THE SIGMA-INVERSE OF A
	C AND STORES THE TRANSPOSE OF IT IN A.
	C MR = FIRST DIMENSION NO. OF A,
	C NR = NO. OF ROWS IN A.
	C NC = NO. OF COLUMNS IN A.
	C U IS THE BOOKKEEPING MATRIX.
	C SIGMA MUST BE POSITIVE.
00002	C
	DIMENSION A(MR,NC),U(NC,NC)
00003	C
00004	DO 10 J=1,NC
00005	DO 10 I=1,J
00006	U(I,J)=0.0
00007	DO 10 K=1,NR
00008	10 U(I,J)=U(I,J)+A(K,I)*A(K,J)
00009	DO 20 I=1,NC
	20 U(I,I)=U(I,I)+SIGMA
00010	C
00011	U(1,1)=SQRT(U(1,1))
00012	DO 30 I=2,NC
00013	30 U(1,I)=U(1,I)/U(1,1)
00014	DO 40 J=2,NC
00015	JA=J-1
00016	DO 41 K=1,JA
00017	41 U(J,J)=U(J,J)-U(J,K)*U(J,K)
00018	U(J,J)=SQRT(U(J,J))
00019	IF(J.EQ,NC) GO TO 40
00020	JB=J+1
00021	DO 42 I=JB,NC
00022	U(I,J)=U(J,I)
00023	DO 43 K=1,JA
00024	43 U(I,J)=U(I,J)-U(I,K)*U(J,K)
00025	42 U(I,J)=U(I,J)/U(J,J)
	40 CONTINUE
00026	C
00027	DO 50 I=1,NC
00028	50 U(I,I)=1.0/U(I,I)
00029	DO 51 K=2,NC
00030	JA=NC+1-K
00031	DO 51 J=1,JA
00032	IA=K+J-2
00033	U(J,IA+1)=0.0
00034	DO 52 I=J,IA
00035	52 U(J,IA+1)=U(J,IA+1)-U(IA+1,I)*U(J,I)
	51 U(J,IA+1)=U(J,IA+1)*U(IA+1,IA+1)
00036	C
00037	DO 60 I=1,NR
00038	DO 60 J=1,NC
00039	X=0.0
00040	JA=NC+1-J
00041	DO 61 K=1,JA
00042	61 X=X+A(I,K)*U(K,JA)
00043	60 A(I,JA)=X
00044	DO 70 I=1,NR
00045	DO 70 J=1,NC
00046	X=0.0
00047	DO 71 K=J,NC
00048	71 X=X+A(I,K)*U(J,K)
	70 A(I,J)=X
00049	C
00050	RETURN
	END

図 5.1 G-逆行列のサブルーチン

零である。このときでも、(5.9) の分母には、 σ^2 が存在するために、条件数 $\text{cond}(A)$ が発散してしまうことがない。 σ の値は、大きいほど大きい程、条件数 $\text{cond}(A)$ は 1 に近づき、数値計算はより安定することが分かる。

5.3. σ -逆行列の計算法Ⅱ —逐次計算法—

σ -逆行列の逐次計算法を求めるには、第4章で明らかにした分割行列の σ -逆行列を利用すればよい。

行列 A を $m \times n$ 行列とする。行列 A の第 k 列ベクトルを a_k とし、行列 A_k を

$$A_k = [a_1 \ a_2 \ \dots \ a_k] \quad (5.10)$$

とする。行列 A_k の分割行列 $[A_{k-1} | a_k]$ に対し、系 4.1 を適用することにより、 σ -逆行列の逐次計算法が得られる。

[σ -逆行列の計算法Ⅱ — 逐次計算法 —]

$$A_k^\sigma = \begin{bmatrix} A_{k-1}^\sigma & -b_k h_k^T \\ \hline & h_k^T \end{bmatrix}, \quad k=2, \dots, n \quad (5.11)$$

ここで

$$b_k = A_{k-1}^\sigma a_k \quad (5.12)$$

$$h_k = \frac{(I - A_{k-1} A_{k-1}^\sigma) a_k}{a_k^T (I - A_{k-1} A_{k-1}^\sigma) a_k + \sigma^2} \quad (5.13)$$

である。

A_1^+ から始めて $k=2$ から $k=n$ まで (5.11) の計算を繰り返すことにより、行列 A の σ -逆行列 A^+ が求められる。この逐次計算法では、行列の逆行列演算は使われていない。

この σ -逆行列の逐次計算法に対応した Moore-Penrose 逆行列の逐次計算法を次に示す。

$$A_k^+ = \begin{bmatrix} A_{k-1}^+ - \frac{b_k h_k^T}{h_k^T b_k} \\ -\frac{b_k h_k^T}{h_k^T b_k} \end{bmatrix} \quad (5.14)$$

ここで、

$$b_k = A_{k-1}^+ a_k \quad (5.15)$$

$$c_k = a_k - A_{k-1} b_k \quad (5.16)$$

$$h_k^T = \begin{cases} c_k^+ & , \quad c_k \neq 0 \text{ の場合} \\ \frac{b_k^T A_{k-1}^+}{1 + b_k^T b_k} & , \quad c_k = 0 \text{ の場合} \end{cases} \quad (5.17)$$

である。

(5.11) と (5.14) とからも明らかのように σ -逆行列と Moore-Penrose 逆行列の逐次計算法は、形式的にはよく似ている。但し、ベクトル h_k の構成の仕方に若干の違いがある。Moore-Penrose 逆行列の逐次計算法では、直接法と同様に、数値計算上の原理的な困難性が存在している。それは、(5.17) におけるベクトル c_k が零ベクトルであるか否かの判断操作である。この判

断は、丸め誤差の存在する数値計算では、安定に行なうことができなり操作である。これに対して、 Φ 逆行列の逐次計算では、(5.13)より明らかのように Moore-Penrose 逆行列の逐次計算法で見られるような判断操作を必要としなため、安定した計算法であることが分かる。

5.4. Moore-Penrose 逆行列の正則化と Φ 逆行列

5.2, 5.3 から明らかのように、 Φ 逆行列に比べて Moore-Penrose 逆行列の数値計算には、不安定性が存在している。これは、第2章で示したように Moore-Penrose 逆行列のスペクトル特性が $1/\varepsilon$ というスカラー関数であり、 $\varepsilon=0$ の近傍において不連続、非有界であることが原因となってくる。いわゆる不適切問題と呼ばれる種類の問題から導かれる行列では、行列の特異値に非常に小さな値が生じるために、この行列の Moore-Penrose 逆行列の数値計算での不安定性を避けることができる。

この問題に対する数値計算上の対策として、種々の方法が考えられている。行列の階数が予め明らかの場合には、それを考慮することにより、数値計算上の不安定性を取除くことができる。しかし、一般には、行列の階数は明らかでない。このような場合の数値計算上の技法として、正規化 (regularization) という方法がある。この方法は、A. N. Tikhonov⁽²⁴⁾ により導入された方法である。

次の連立一次方程式を考える。

$$A x = b \quad (5.18)$$

ここで、 A は $m \times n$ 行列、 b は与えられた m 次元ベクトル、 x は未知の n 次元ベクトルである。このとき、行列 A の Moore-Penrose 逆行列 A^+ に

よる解 $\hat{x} = A^+ b$ は、 x のノルム最小の最小二乗解を与える。即ち、

$$\|A x - b\| = \inf_x \|A x - b\| \quad (5.19)$$

を満足する x の中でノルム $\|x\|$ が最小の解となる。

いま、行列 A の特異値の中に零に近い値があると、Moore-Penrose 逆行列 A^+ は、数値計算上、安定に求めることができないため、解 $\hat{x} = A^+ b$ 自身も不安定なものとなる。そこで、最適化の規準を次のように変更する。即ち最小二乗解では、 $\|A x - b\|$ の最小化を計ったが、それを次の

$$\|A x - b\|^2 + \lambda \|x\|^2 \quad (\lambda > 0) \quad (5.20)$$

という値の最小化に置き換える。この (5.20) の第二項は、 x のノルムの平方 $\|x\|^2$ に λ を掛けた一定の罰金 (penalty) である。この項を最適化規準に予め加えておくことにより、解の安定化を計ることができるようである。

実際には、(5.20) の最小解は次式で与えられる。

$$\hat{x} = R_\lambda \cdot b \quad (5.21)$$

但し、

$$R_\lambda \equiv (A^T A + \lambda I)^{-1} A^T \quad (5.22)$$

である。この行列 R_λ は、Moore-Penrose 逆行列 A^+ と異なり、 $\lambda > 0$ に対して、安定に数値計算できる。このことから、行列 A^+ の代わりに行列 R_λ を用いることにより、安定した解を求めることができる。この方法が正規化と呼ばれるものである。行列 R_λ は正規化行列と呼ばれてゐる。

行列 R_λ と Moore-Penrose 逆行列 A^+ との間には、次のような漸近的な

関係がある。

$$A^+ = \lim_{\lambda \rightarrow +0} R_\lambda \quad (5.23)$$

このことから、 λ を十分に小さく与えれば、正規化により得られる解が、ノルム最小の最小二乗解に近いものとなることが分かる。

この正規化という方法は、あくまでも最小二乗解を求めるための計算上の技法として導入されたものである。この方法において、正規化パラメータ λ をどのように選ぶか、という問題が残されている。 λ を大きく選べば、それだけ行列 R_λ の計算は安定化するが、得られた解が、最小二乗解とは異なったものとなる。 λ をより小さく選べば、最小二乗解により近い解が得られるが、その代りに、行列 R_λ の計算が不安定となる。このように解の精度化と安定化という要求は、相反するものであるから、各場面に依じて適当に λ を選ぶ必要がある。

正規化という手法は、連立方程式 (5.18) の最小二乗解を安定に求めるための計算上の技法として導入されたものであるが、この範囲で考える限り、パラメータ λ の選定に対する明確な基準を与えることはできない。しかし、この問題を統計的な線形観測システムでの推定問題として見直してみると、パラメータ λ に関して、より積極的な意味付けが可能となる。(5.22) の正規化行列 R_λ をみてみると、これはちょうど $\sigma^2 = \lambda$ としたときの行列 A の σ -逆行列 A^σ に他ならない。このことは、正規化による解 $\hat{x} = R_\lambda b$ が、連立方程式 (5.18) を統計的な線形観測システム (A, λ) とみなしたときの x の最小分散推定量であることを意味している。このような観点に立てば、正規化パラメータ λ も線形観測システムでの観測誤差の統計的性質である σ^2 に対応していることが明らかになり、選定の基準が明確なものとなるのである。

第 6 章 最適分散性特徴パラメータ

6.1 まえがき

文字図形や画像のように位相構造をもつパターンは、有限次元ベクトル空間の元として近似的にとりえることができる。いま、この空間の次元を N としよう。このとき 1 組の正規直交系を定めることにより、パターンは N 個の特徴パラメータにより表現することができる。

特徴パラメータの中でも、特に基本特徴パラメータは重要なものである。このパラメータはパターンの近似という立場から与えられたものである。パターンを N 個以下の L 個の特徴パラメータを用いて近似する場合に、基本特徴パラメータは最良近似を与えるパラメータとなる。一方、等分散性特徴パラメータは、全く別の立場から、飯島により提案された特徴パラメータである。⁽¹⁰⁾ この特徴パラメータは、各パラメータの分散がみな等しい値をもつものであり、均一な精度をもつ処理機構を用いて情報処理を行なう場合に、処理精度の向上という立場から最適な特徴パラメータとなることが知られている。⁽¹⁰⁾

しかし、実際には、処理機構が必ずしも均一な構造をもたない場合がある。このような場合には、等分散性特徴パラメータは最適なものとはならない。それぞれの処理機構に合わせて特徴パラメータを選ぶ必要が生じてくる。

最適分散性特徴パラメータは、以上のような要請のもとに提案されたものである。この特徴パラメータは分散がある最適性の下にあらかじめ定められた値をもつような特徴パラメータとして定義される。

本章では、このような最適分散性特徴パラメータの実現可能性について考察する。さらに、応用として、パターン間の内積演算処理の場面にこの特徴パラメータの適用を行なう。

6.2 特徴パラメータ

本章で対象とするパターンは、 N 次元実ベクトル空間の元として与えられるとする。パターンは、 N 次元確率ベクトル $\mathbf{x}$ であり、その情報源の統計的性質、すなわち、その平均、及び共分散が次のように与えられているとする。

$$E(\mathbf{x}) = \mathbf{0} \quad (6.1)$$

$$E(\mathbf{x}\mathbf{x}^T) = \mathbf{S} \quad (6.2)$$

ここで、行列 $\mathbf{S}$ は正値対称行列である。

$\{\phi_i : i=1, \dots, N\}$ を任意の N 次元正規直交系とする。このとき、ベクトル $\mathbf{x}$ は次のように表わされる。

$$\mathbf{x} = \sum_{i=1}^N a_i \phi_i \quad (6.3)$$

ここで、

$$a_i = (\mathbf{x}, \phi_i)^{(注)}, \quad i=1, \dots, N \quad (6.4)$$

である。

$\mathbf{a} = (a_1 \ a_2 \ \dots \ a_N)^T$ とおくと、ベクトル $\mathbf{x}$ とベクトル $\mathbf{a}$ とは一対一に対応する。一般に、このようにして得られた $\{a_i : i=1, \dots, N\}$ をパターン $\mathbf{x}$ の特徴パラメータという。

正規直交系は任意に選ぶことができるが、その中で特に特徴パラメータ $\{a_i\}$ が次の統計的性質を満たすような正規直交系をモード関数系（あるいは、 $K-L$ 直交系）と呼ぶ。

(注). $(\mathbf{x}, \mathbf{y})$ は ベクトル $\mathbf{x}$ とベクトル $\mathbf{y}$ との間の内積を表わす。

$$E(a_i a_j) = \lambda_i \delta_{ij}, \quad i, j = 1, \dots, N \quad (6.5)$$

但し、 $\{\lambda_i : i = 1, \dots, N\}$ は行列 S の固有値であり、次式を満足する。

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N \geq 0 \quad (6.7)$$

このモード関数系による特徴パラメータを基本特徴パラメータという。基本特徴パラメータは、パターンとの近似の立場から求められた特徴パラメータである。いま、 N 個以下の L 個の特徴パラメータだけを用いてパターン x を近似することを考える。このとき、近似による誤差の二乗平均

$$E \| x - \sum_{i=1}^L a_i \phi_i \|^2 \quad (6.8)$$

が最小となる、という意味で基本特徴パラメータは最適なパラメータである。

一方、等分散性特徴パラメータとは、次の条件を満足する特徴パラメータのことである。

$$E(a_i^2) = \mu, \quad i = 1, \dots, N \quad (6.9)$$

但し、

$$\mu = \frac{1}{N} \sum_{i=1}^N \lambda_i \quad (6.10)$$

である。この特徴パラメータは、情報処理の精度化の立場から与えられたものである。例えば、同程度の信頼性をもつ積演算回路を用いた内積演算処理において、この等分散性特徴パラメータが最適な特徴パラメータであることが明らかにされている。(10)

6.3 Doubly Stochastic 行列

n 次実行列 $A = (a_{ij})$ が次の三つの条件を満足するとき、この行列 A を doubly stochastic 行列 (ds 行列と略す) という。

$$(a) \quad a_{ij} \geq 0, \quad i, j = 1, \dots, n \quad (6.11)$$

$$(b) \quad \sum_{i=1}^n a_{ij} = 1, \quad j = 1, \dots, n \quad (6.12)$$

$$(c) \quad \sum_{j=1}^n a_{ij} = 1, \quad i = 1, \dots, n \quad (6.13)$$

最も簡単な ds 行列は順列行列である。順列行列とは、各行各列に 1 がちょうど一つあり、他はすべて 0 であるような行列である。例えば、次のような行列である。

$$\begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

n 次の ds 行列全体のなす集合を $K(n)$ と書くと、 $K(n)$ は n 次実行列全体のなす n^2 次元ユークリッド空間の凸部分集合である。この凸部分集合 $K(n)$ の構造を明らかにしたのが、次のバーコフ、フォンノイマンの定理である。

[定理 6.A] ds 行列全体のなす集合 $K(n)$ は凸部分集合であり、その端点集合は n 次の順列行列全体である。

この定理より、任意の ds 行列 A は、必ず幾つかの順列行列 $P_1, P_2, \dots, P_m$ の凸一次結合の形で書くことができる。即ち、

$$A = \sum_{i=1}^m \alpha_i P_i \quad (6.14)$$

但し、

$$\alpha_i \geq 0, \quad i=1, \dots, m, \quad (6.15)$$

$$\sum_{i=1}^m \alpha_i = 1 \quad (6.16)$$

である。

この ds 行列 A, B の積 $C = AB$ もやはり ds 行列である。このことから、 n 次元実ベクトルの間に次のような関係を定めることができる。

[定義 6.A] この n 次元実ベクトル X, Y に対して、 $D \in K(n)$ が存在して、 $Y = DX$ となるとき、次の関係

$$X \succ Y \quad (6.17)$$

が成り立つとする。

この関係が半順序関係となることは明らかである。定理 6.A より、ベクトル Y が、すべての順列行列 $\{P_i : i=1, \dots, n!\}$ を用いてつくられる $\text{convex hull} \{P_i X : i=1, \dots, n!\}$ に属することと、関係 $X \succ Y$ が成り立つことが同値となる。

任意に与えられた n 次元ベクトル X, Y に対して、上の関係が成り立つか否かを判定するには、どのようにしたらよいであろうか。この判定法が、

Schur (1923)により与えられている:

$X \succ Y$ となるための必要十分条件は、各自然数 k ($1 \leq k \leq n-1$) に対して $\Omega = \{1, 2, \dots, n\}$ の k 個の元よりなる部分集合 J 全体の族を $\mathcal{M}_k$ として、

$$\max_{J \in \mathcal{M}_k} \left(\sum_{j \in J} x_j \right) \geq \max_{J \in \mathcal{M}_k} \left(\sum_{j \in J} y_j \right), \quad (6.18)$$

$$k = 1, 2, \dots, n-1$$

となり、しかも

$$\sum_{i=1}^n x_i = \sum_{i=1}^n y_i \quad (6.19)$$

が成り立つことである。

なお、条件 (6.18) は、 X と Y の成分を並べかえて、

$$\begin{cases} x_1 \geq x_2 \geq \dots \geq x_n \\ y_1 \geq y_2 \geq \dots \geq y_n \end{cases} \quad (6.20)$$

としておけば、

$$\begin{cases} x_1 \geq y_1 \\ x_1 + x_2 \geq y_1 + y_2 \\ \dots \dots \dots \\ x_1 + x_2 + \dots + x_{n-1} \geq y_1 + y_2 + \dots + y_{n-1} \end{cases} \quad (6.21)$$

と同値である。

この判定法にちなんで、 ds 行列による一次変換を Schur 変換と呼ぶ。

6. 4 最適分散性特徴パラメータ

最適分散性特徴パラメータは等分散性特徴パラメータの自然な拡張として定義される。

[定義 6. 1] $\{\mu_i : i=1, \dots, N\}$ は次の条件を満足するものとする。

$$\mu_i \geq 0, \quad i=1, \dots, N \quad (6.22)$$

$$\sum_{i=1}^N \mu_i = \sum_{i=1}^N \lambda_i \quad (6.23)$$

このとき、次の条件を満足する特徴パラメータ $\{a_i : i=1, \dots, N\}$ を $\{\mu_i\}$ に応じた最適分散性特徴パラメータと呼ぶ。

$$E(a_i^2) = \mu_i, \quad i=1, \dots, N \quad (6.24)$$

このように、最適分散性特徴パラメータは、 $\{\mu_i\}$ に応じて定まるものである。とく $\{\mu_i\}$ として、

$$\mu_i = \frac{1}{N} \sum_{i=1}^N \lambda_i, \quad i=1, \dots, N \quad (6.25)$$

とおくと、等分散性特徴パラメータに一致する。

最適分散性特徴パラメータは、 $\{\mu_i\}$ に応じて定まるものであるが、ここで問題となるのは、(6.22)及び(6.23)を満たす任意の $\{\mu_i\}$ に対して最適分散性特徴パラメータを実際に求めることができるか否か、ということである。この問題を明らかにしたのが次の定理である。

[定理 6.1] 最適分散性特徴パラメータの存在するための必要十分条件は次式である。

$$\mu < \lambda \quad (6.26)$$

但し、

$$\mu = (\mu_1 \ \mu_2 \ \dots \ \mu_N)^T$$

$$\lambda = (\lambda_1 \ \lambda_2 \ \dots \ \lambda_N)^T$$

である。

(証明) あらかじめモード関数系 $\{\phi_i : i=1, \dots, N\}$ 及び基本特徴パラメータ $\{a_i : i=1, \dots, N\}$ を求めておく。任意の特徴パラメータ $\{b_i : i=1, \dots, N\}$ は N 次直交行列 U を用いて

$$b = U a \quad (6.27)$$

と表わすことができる。但し、 $b = (b_1 \ b_2 \ \dots \ b_N)^T$ である。また、この特徴パラメータを有する正規直交系 $\{\psi_i : i=1, \dots, N\}$ は次のようにして求めることができる。

$$(\psi_1 \ \psi_2 \ \dots \ \psi_N)^T = U (\phi_1 \ \phi_2 \ \dots \ \phi_N)^T \quad (6.28)$$

ここで、ベクトル b の共分散を求めると、

$$\begin{aligned} E(b b^T) \\ &= E(U a a^T U^T) \\ &= U \Lambda_N U^T \end{aligned} \quad (6.29)$$

となる。但し、

$$\Lambda_N = \begin{bmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ 0 & & \ddots \\ & & & \lambda_N \end{bmatrix} \quad (6.30)$$

である。これより、証明すべきことは、

$$U \Lambda_N U^T = \begin{bmatrix} \mu_1 & & * \\ & \mu_2 & \\ * & & \ddots \\ & & & \mu_N \end{bmatrix} \quad (6.31)$$

を満たす直交行列 U が存在するための必要十分条件を求めることである。

<必要性> $U = (u_{ij})$ を用いて λ と μ の間の関係を明らかにする。

$$\mu_i = \sum_{j=1}^N u_{ij} \lambda_j u_{ij} = \sum_{j=1}^N u_{ij}^2 \lambda_j, \quad j=1, \dots, N$$

これより、次式が求められる。

$$\mu = \begin{bmatrix} u_{11}^2 & u_{12}^2 & \cdots & u_{1N}^2 \\ u_{21}^2 & u_{22}^2 & \cdots & u_{2N}^2 \\ \vdots & \vdots & \ddots & \vdots \\ u_{N1}^2 & u_{N2}^2 & \cdots & u_{NN}^2 \end{bmatrix} \preceq \lambda \quad (6.32)$$

==で、 U は直交行列であるから行列 (u_{ij}^2) はdS行列である。ゆえに、

$$\mu \prec \lambda \quad (6.33)$$

となる。

・〈十分性〉 式(6.31)を満足する直交行列 U を具体的に構成することにより十分性を示す。

一般性を失なうことなく、

$$\mu_1 \geq \mu_2 \geq \cdots \geq \mu_N \quad (6.34)$$

が成り立つものとする。Schurの判定法により、次の不等式が成り立つ。

$$\sum_{i=1}^k \lambda_i \geq \sum_{i=1}^k \mu_i, \quad k=1, \dots, N-1 \quad (6.35)$$

いま、次のような行列 U_1 を構成する。

$$U_1 \Lambda_N U_1^T = \begin{bmatrix} \mu_1 & & * \\ \hline & \lambda'_2 & & 0 \\ * & & \lambda'_3 & \\ & 0 & & \ddots & \\ & & & & \lambda'_N \end{bmatrix} \quad (6.36)$$

但し, $\{\lambda'_i : i = 2, \dots, N\}$ は次式を満足する。

$$\sum_{i=2}^k \lambda'_i \geq \sum_{i=2}^k \mu_i, \quad i = 2, \dots, N-1 \quad (6.37)$$

このような行列 U_1 が求めれば, 次に右下の $N-1$ 次の小行列に対して同様な方法で U_2 を構成して,

$$U_2 U_1 \Lambda_N U_1^T U_2^T = \begin{bmatrix} \mu_1 & & & * \\ & \mu_2 & & \\ & & \lambda''_3 & 0 \\ * & & & \ddots \\ & & 0 & \lambda''_N \end{bmatrix} \quad (6.38)$$

とすることが出来る。以下, 同様にして $U_3, U_4, \dots, U_{N-1}$ が求められる。そして,

$$U = U_{N-1} U_{N-2} \cdots U_1 \quad (6.39)$$

とすると, 行列 U は (6.31) を満足する。

以上のことから, 直交行列 U_1 の構成法が明らかになれば, 十分性は証明されたことになる。

いま, 行列 U_1 を次のようにおく。

$$U_1 = R(\theta_{1N}) R(\theta_{1N-1}) \cdots R(\theta_{13}) R(\theta_{12}) \quad (6.40)$$

ここで, $R(\theta_{ij})$ は, i, j 成分のなす二次元平面上での回転である。すなわち,

$$R(\theta_{ij}) = I_{N-2} \oplus U_0(\theta_{ij}), \quad (6.41)$$

$$U_0(\theta_{ij}) = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \quad (6.42)$$

ここで、 I_{N-2} は $N-2$ 次の単位行列、 $\oplus$ は直和を表す。

この行列 U_1 により、(6.36.) の形の変換がなされることは明らかである。
(6.37) を満足するように $\{\theta_{1k} : k=2, \dots, N\}$ を適当に選び、 $U_1 \Lambda_N U_1^T$ の $(1,1)$ 成分を μ_1 にすればよい。

まず、二次元の場合について考えてみる。

$$U_0(\theta) \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} U_0(\theta)^T = \begin{bmatrix} x_1 & * \\ * & x_2 \end{bmatrix} \quad (6.43)$$

とおく。

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \cos^2 \theta & \sin^2 \theta \\ \sin^2 \theta & \cos^2 \theta \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} \quad (6.44)$$

となる。

いま、 θ を 0 から $\pi/4$ まで変化させると、 x_1 は λ_1 から $(\lambda_1 + \lambda_2)/2$ に減少し、 x_2 は λ_2 から $(\lambda_1 + \lambda_2)/2$ に増加する。行列 $U_0(\theta)$ により、 $(1,1)$ 成分から $(2,2)$ 成分に移すことのできる値は、 θ が 0 から $\pi/4$ に増加するに従って、0 から $(\lambda_1 - \lambda_2)/2$ にまで変化することが分かる。

以上のことから、(6.40) の行列 U_1 において、 θ_{1k} を順次に 0 から $\pi/4$ に増加させることにより、 $(1,1)$ 成分の値 λ_1 は次第に減少する。そして、その減少分が (k,k) 成分に移動する。 λ_1 が μ_1 になるまで、 θ_{1k} を調整

すればよい。なお、このとき、 $\lambda_k < \mu_k$ である k についてのみ、 θ_{1k} の調整を行なう。具体的に θ_{1k} を定める手続きを以下に示す。

<ステップ1> $\Delta_1 = \lambda_1 - \mu_1$ とする。 $\Delta_1 = 0$ のときは、 $\theta_{1k} = 0$, $k = 2, \dots, N$ として手続きを終える。 $\Delta_1 > 0$ のときには、 $2 \rightarrow k$ として、ステップ2に進む。

<ステップ2> $\Delta_k = \mu_k - \lambda_k$ を求める。 $\Delta_k \leq 0$ のときは、 $\theta_{1k} = 0$ とする。 $k+1 \rightarrow k$ として、ステップ2の始めにもどる。 $\Delta_k > 0$ のときは、次の二つの場合に分けられる。 $\Delta_k < \Delta_1$ のときは、ステップ3に進む。 $\Delta_k \geq \Delta_1$ のときは、ステップ4に進む。

<ステップ3> λ_k が μ_k になるように θ_{1k} を決める。即ち、

$$\begin{bmatrix} \lambda_1 - \Delta_k \\ \mu_k \end{bmatrix} = \begin{bmatrix} \cos^2 \theta_{1k} & \sin^2 \theta_{1k} \\ \sin^2 \theta_{1k} & \cos^2 \theta_{1k} \end{bmatrix} \cdot \begin{bmatrix} \lambda_1 \\ \lambda_k \end{bmatrix} \quad (6.45)$$

を満たす θ_{1k} を求める。

$$\sin^2 \theta_{1k} = \frac{\mu_k - \lambda_k}{\lambda_1 - \lambda_k}$$

より、次のように θ_{1k} が定まる。

$$\theta_{1k} = \sin^{-1} \sqrt{\frac{\mu_k - \lambda_k}{\lambda_1 - \lambda_k}} \quad (6.46)$$

この段階で λ_1 は $\lambda_1 - \Delta_k$ となるから、 $\Delta_1 - \Delta_k \rightarrow \Delta_1$, $k+1 \rightarrow k$ として、ステップ2にもどる。

<ステップ4> このとき、 λ_1 を μ_1 にすることができ。即ち、

$$\begin{bmatrix} \mu_1 \\ \lambda_k + \Delta_1 \end{bmatrix} = \begin{bmatrix} \cos^2 \theta_{1k} & \sin^2 \theta_{1k} \\ \sin^2 \theta_{1k} & \cos^2 \theta_{1k} \end{bmatrix} \cdot \begin{bmatrix} \lambda_1 \\ \lambda_k \end{bmatrix} \quad (6.47)$$

を満たす θ_{1k} を求める。

$$\sin^2 \theta_{1k} = \frac{\lambda_1 - \mu_1}{\lambda_1 - \lambda_k}$$

より、次のように θ_{1k} が定まる。

$$\theta_{1k} = \sin^{-1} \sqrt{\frac{\lambda_1 - \mu_1}{\lambda_1 - \lambda_k}} \quad (6.48)$$

この段階で (1, 1) 成分は μ_1 になるから、以下の θ_{1k} をすべて 0 とし、手続きを終る。

k が 2 から N へ増すに従って Δ_1 の値は次第に零に近づくから、必ずある時点でステップ4に進み、手続きは終る。以上のようにして、(6.36) を満たす直交行列 U_1 が求められる。なお、 λ_1 と μ_1 との差は、 $\lambda_k < \mu_k$ である (k, k) 成分に移動するが、 k の若い順番に移動が行なわれることから、(6.37) の条件も満たれていくことが分かる。

以上で定理は証明された。

(証明終)

この定理の十分性の証明は、そのまま最適分散性特徴パラメータを与える正規直交系の構成法を与えておく。なお、この構成法は対称行列の固有値及び、

固有ベクトルを求めるためのヤコビ法と対照的な方法である。

〔例〕 次の場合の最適分散性特徴パラメータを求めてみる。

$$S = \begin{bmatrix} 8 & 0 & 0 \\ 0 & 3 & 2 \\ 0 & 2 & 3 \end{bmatrix}, \quad \mu = \begin{bmatrix} 6 \\ 6 \\ 2 \end{bmatrix}$$

まず、はじめに行列 S を対角化する。

$$U_0 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ 0 & \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix}$$

とすると、

$$S_0 = U_0 S U_0^T = \begin{bmatrix} 8 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

となる。行列 U_0 の行ベクトルがモード関数系であり、 $\lambda = (8 \ 5 \ 1)^T$ である。

μ は、 $(6, 2, 2)$ 及び $(6, 2, 3)$ を満足し、さらに Schwarz の判定法により、 $(6, 2, 6)$ も満足していることがわかる。ゆえに、 μ に応ずる最適分散性特徴パラメータが存在する。

実際は、

$$U_1 = \begin{bmatrix} \sqrt{\frac{2}{3}} & -\sqrt{\frac{1}{3}} & 0 \\ \sqrt{\frac{1}{3}} & \sqrt{\frac{2}{3}} & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad U_2 = \begin{bmatrix} \sqrt{\frac{5}{6}} & 0 & -\sqrt{\frac{1}{6}} \\ 0 & 1 & 0 \\ \sqrt{\frac{1}{6}} & 0 & \sqrt{\frac{5}{6}} \end{bmatrix}$$

とおく。

$$U_2 U_1 S_0 U_1^T U_2^T = \begin{bmatrix} 6 & \sqrt{\frac{5}{3}} & \sqrt{5} \\ \sqrt{\frac{5}{3}} & 6 & \sqrt{\frac{1}{3}} \\ \sqrt{5} & \sqrt{\frac{1}{3}} & 2 \end{bmatrix}$$

となる。上の行列の対角成分が $(6, 6, 2)$ となっている。これより、次の行列

$$U = U_2 U_1 U_0 = \begin{bmatrix} \frac{\sqrt{5}}{3} & \frac{\sqrt{3}-\sqrt{5}}{6} & \frac{-\sqrt{3}-\sqrt{5}}{6} \\ \frac{\sqrt{3}}{3} & \frac{\sqrt{3}}{3} & \frac{\sqrt{3}}{3} \\ \frac{1}{3} & \frac{-1-\sqrt{15}}{6} & \frac{-1+\sqrt{15}}{6} \end{bmatrix}$$

の行ベクトルが求める最適分散性特徴パラメータを与える正規直交系である。

N 次元ベクトル空間のパターン $\times$ は、 N 個の特徴パラメータにより表現されるが、更に冗長性を加えて M ($> N$) 個の特徴パラメータにより表現することもできる。このための数学的道具として、擬似直交系という概念が提案されている。^{(11), (15)} いま、 M 個の擬似直交系を $\{\phi_i : i=1, \dots, M\}$ とする。このとき、パターン $\times$ は次のように表わされる。

$$\times = \sum_{i=1}^M b_i \phi_i \quad (6.49)$$

==で、

$$b_i = (x, \phi_i) \quad , \quad i = 1, \dots, M \quad (6.50)$$

である。この形式は、(6.3)と(6.4)とまったく同じであるが、特徴パラメータの数が N から M に増えている。

このように冗長性を加味した特徴パラメータに対して、最適分散性特徴パラメータを次のように定義することができる。

いま、 $\{\mu_i : i = 1, \dots, M\}$ が次の条件を満足するものとする。

$$\mu_i \geq 0 \quad , \quad i = 1, \dots, M \quad (6.51)$$

$$\sum_{i=1}^M \mu_i = \sum_{i=1}^N \lambda_i \quad (6.52)$$

このとき、擬似直交系による特徴パラメータ $\{b_i : i = 1, \dots, M\}$ が次の条件を満足するとき、これを $\{\mu_i\}$ に依じた最適分散性特徴パラメータと呼ぶ。

$$E(b_i^2) = \mu_i \quad , \quad i = 1, \dots, M \quad (6.53)$$

この条件を行列表現すれば、次のようになる。

$$P S P^T = \begin{bmatrix} \mu_1 & & & \\ & \mu_2 & & * \\ & & \ddots & \\ * & & & \mu_M \end{bmatrix} \quad (6.54)$$

==で、行列 P は、 $\{\phi_i\}$ を行ベクトルにもつ擬似直交行列である。

任意の擬似直交行列 P は、次の行列 P_0

$$P_0 = \underbrace{\begin{bmatrix} 1 & & & 0 \\ & 1 & & \\ & & \ddots & \\ 0 & & & 1 \\ \hline & & & & 0 \end{bmatrix}}_N \left. \vphantom{\begin{bmatrix} 1 & & & 0 \\ & 1 & & \\ & & \ddots & \\ 0 & & & 1 \\ \hline & & & & 0 \end{bmatrix}} \right\} M \quad (6.55)$$

と、 M 次の直交行列 V を用いて、

$$P = V P_0 \quad (6.56)$$

と表わすことができる。これより、(6.54)の左辺は次のようになる。

$$P S P^T = V S' V^T \quad (6.57)$$

但し、 S' は、次の M 次対称行列である。

$$S' = \begin{bmatrix} S & 0 \\ 0 & 0 \end{bmatrix} \quad (6.58)$$

これより、(6.54)を満足する行列 P を求める問題は、 M 次元ベクトル空間において共分散行列 S' の場合の直交行列 V を求める問題に帰着される。定理 6.1 で示した方法により、まず直交行列 V を求め、(6.56) に従って行列 P を求めれば良い。求めた擬似直交行列 P の行ベクトルが求める最適分散性特徴ベクトル x を与える擬似直交系となる。

6.5 最適分散性特徴パラメータの応用

最適分散性特徴パラメータの応用として、二つのパターン間の内積演算の精度化の問題を考える。

いま、内積演算を実行するために M ($\geq N$) 個の積演算回路が用意されているとする。 M 個の積演算回路には、次のような比例性ノイズが各回路に独立に加わるものとする。

$$a_i \cdot b_i \rightarrow a_i \cdot b_i (1 + \Delta_i), \quad i = 1, \dots, M \quad (6.59)$$

ここで、 Δ_i は情報源とは独立な確率変数であり、次のような性質をもつものとする。

$$E(\Delta_i) = 0, \quad i = 1, \dots, M \quad (6.60)$$

$$E(\Delta_i^2) = \gamma_i, \quad i = 1, \dots, M \quad (6.61)$$

この M 個の積演算回路をすべて活用するために、擬似直交系を用いて、パターン X を M 個の特徴パラメータで表現する。このとき、パターン X とパターン Y との間の内積 (X, Y) は次のようにして求まる。

$$(X, Y) = \sum_{i=1}^M a_i \cdot b_i \quad (6.62)$$

但し、 $\{a_i : i = 1, \dots, M\}$ および $\{b_i : i = 1, \dots, M\}$ は、擬似直交系 $\{c_i : i = 1, \dots, M\}$ を用いて、次式のように与えられるパターン X およびパターン Y の特徴パラメータである。

$$\begin{cases} a_i = (\mathbf{x}, \phi_i) \\ b_i = (\mathbf{y}, \phi_i) \end{cases}, \quad i = 1, \dots, M \quad (6.63)$$

a_i と b_i との間の積演算に i 番目の積演算回路を用いるものとする。また、 $\{a_i \cdot b_i : i = 1, \dots, M\}$ の総和演算には演算誤差は入らないとする。このとき、内積演算に生ずる演算誤差の二乗平均は次のようになる。

$$\begin{aligned} \bar{\varepsilon}^2 &= E \left\{ \left(\sum_{i=1}^M a_i b_i (1 + \Delta_i) - \sum_{i=1}^M a_i b_i \right)^2 \right\} \\ &= E \left\{ \left(\sum_{i=1}^M a_i b_i \Delta_i \right)^2 \right\} \\ &= E \left(\sum_{i=1}^M r_i \cdot a_i^2 b_i^2 \right) \end{aligned}$$

ここで、パターン $\mathbf{x}$ とパターン $\mathbf{y}$ はともに同じ情報源から互いに独立に取り出されるものとする。このとき、平均二乗誤差 $\bar{\varepsilon}^2$ は次のようになる。

$$\bar{\varepsilon}^2 = \sum_{i=1}^M r_i \cdot p_i^2 \quad (6.64)$$

但し、 p_i は特散パラメータ a_i の分散である。即ち、

$$E(a_i^2) = p_i, \quad i = 1, \dots, M \quad (6.65)$$

である。

内積演算処理の精度化を計るには (6.64) の $\bar{\varepsilon}^2$ が最小となるように、 $\{p_i\}$ を定めれば良い。 $\{p_i\}$ には、次の制約条件がある。

$$\sum_{i=1}^M p_i = \sum_{i=1}^N \lambda_i \quad (6.66)$$

これより、未定係数 k を用いて、

$$J = \sum_{i=1}^M r_i \cdot p_i^2 + k \left(\sum_{i=1}^M p_i - \sum_{i=1}^N \lambda_i \right) \quad (6.67)$$

とおき、 J の各 p_i についての変分を零とおくことにより、 $\bar{\varepsilon}^2$ を最小にする $\{p_i\}$ が求まる。これを $\{p_i^*\}$ とおくと、

$$p_i^* = \frac{k}{r_i}, \quad i=1, \dots, M \quad (6.68)$$

となる。ここで定数 k は次式で与えられる。

$$k = \left(\sum_{i=1}^N \lambda_i \right) / \left(\sum_{i=1}^M \frac{1}{r_i} \right) \quad (6.69)$$

以上のことから、

$$\mu = (p_1^* \ p_2^* \ \dots \ p_M^*)^T \quad (6.70)$$

に対応した最適分散性特徴パラメータが求まる特徴パラメータとなる。この特徴パラメータが実現できるか否かは、この μ が定理 6.1 の条件 (6.26) を満足するか否かを調べなければならぬ。一般に情報源の固有値 $\{\lambda_i\}$ のばらつきに比べて、 $\{r_i\}$ のばらつきが十分に小さい場合には、この条件 (6.26) は満足される。

最適分散性特徴パラメータを用いた場合の内積演算の平均二乗誤差 $\bar{\varepsilon}_0^2$ は、次のようになる。

$$\bar{\varepsilon}_0^2 = \left(\sum_{i=1}^N \lambda_i \right)^2 / \left(\sum_{i=1}^M \frac{1}{r_i} \right) \quad (6.71)$$

比較のために、等分散性特徴パラメータを用いた場合の平均二乗誤差 $\bar{\varepsilon}_1^2$ を求めると、

$$\bar{\varepsilon}_1^2 = \frac{1}{M^2} \left(\sum_{i=1}^M r_i \right) \left(\sum_{i=1}^N \lambda_i \right)^2 \quad (6.72)$$

となる。 $\bar{\varepsilon}_0^2$ と $\bar{\varepsilon}_1^2$ との間の比を求めてみる。

$$\frac{\bar{\varepsilon}_0^2}{\bar{\varepsilon}_1^2} = \frac{H}{A} \quad (6.73)$$

ここで、 H と A は次式で与えられる。

$$A \equiv \frac{1}{M} \sum_{i=1}^M r_i \quad (6.74)$$

$$H \equiv \left(\frac{1}{M} \sum_{i=1}^M \frac{1}{r_i} \right)^{-1} \quad (6.75)$$

このように、 $\bar{\varepsilon}_0^2$ と $\bar{\varepsilon}_1^2$ との間の比は、 $\{r_i\}$ の調和平均と相加平均との間の比となることが分かる。 $\{r_i\}$ のばらつきが大きくなると、調和平均は相加平均に比べて小さな値をとる。このことは、 $\{r_i\}$ のばらつきが大きくなればなる程、最適分散性特徴パラメータを用いることの効果が著しいことを示している。

6.6 あとがき

一般に位相構造をもつパターンはベクトル空間の元としてとらえることができる。情報源の統計的性質、例えば、平均とか共分散といったものは、対象とするパターンに固有の性質である。一方、このパターンの処理には、あるハードウェアによる処理機構が用いられるが、処理機構自身も信頼性に関して固有性をもっている。このときに、パターンの性質と処理機構の性質とをうまく整合させるようにパターンを取り扱う必要がある。本章で与えた最適分散性特徴パラメータは、このような要請の下に提案されたものである。

本章では、最適分散性特徴パラメータの応用として、パターンの間の内積演算処理について考えたが、この特徴パラメータは、この他にパターンの伝送、記憶、更に一般の情報処理においても、重要な役割を演ずるものと期待される。

第7章 線形観測システムの最適化問題

7.1 線形観測システムの最適化

統計的線形観測システム $(A, S, \sigma^2 R)$ における未知ベクトル x の最小分散推定量 $\hat{x}$ は、行列 A の重み付きの逆行列 $A S R$ を用いて次式で与えられる。

$$\hat{x} = A S R^{-1} y \quad (7.1)$$

ここで、 y は得られた観測ベクトルである。このとき、真値 x と最小分散推定量 $\hat{x}$ との間の平均二乗誤差は次のようになる。

$$E(\|x - \hat{x}\|^2) = \text{tr}((I - A S R^{-1} A) S) \quad (7.2)$$

いま、未知ベクトル x の再現性を問題とするならば、この $E(\|x - \hat{x}\|^2)$ を線形観測システムの最適化基準とするのが適當である。(7.2) を行列 A の汎関数として、 $J(A)$ と表わすことにする。本章で取り扱う線形観測システムの最適化とは、この $J(A)$ を最小とする行列 A を求めることである。但し、実際には、与えられた状況により、実現可能な行列の集合が予め限られている。実現可能な行列 A の集合 M とおくと、線形観測システムの最適化問題は、

$$J(A_0) = \min_{A \in M} J(A) \quad (7.3)$$

を満たす行列 $A_0 \in M$ を求めることである。

線形観測システム及びその推定システムをまとめて図式化すると、図7.1の

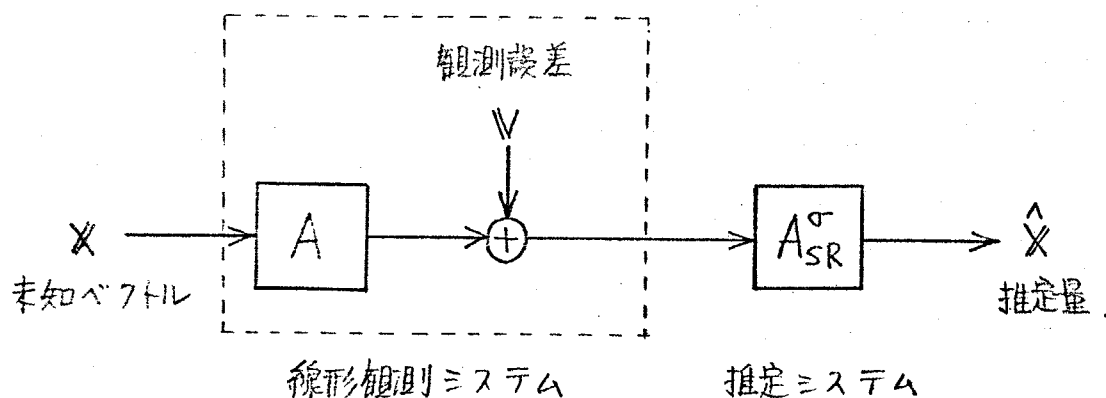


図 7.1 線形観測システムとその推定システム

ようになる。二つのシステムをまとめると、一つのシステムとしてみると、ちょうど未知ベクトル $\hat{x}$ の並列伝送系とみることが出来る。ここで、行列 A は変調器であり、行列 A_{SR}^{σ} が復調器である。このことから、線形観測システムの最適化問題は、並列伝送系の最適化問題と等価であることが分かる。

行列 A の実現可能な集合 $\mathcal{M}$ は、それぞれ状況に従って定まるものである。本章では、次の二つの場合について、その最適化問題を検討する。

$$\text{I. } \mathcal{M} = \{A \mid A S A^T = P_0\} \quad (7.4)$$

$$\text{II. } \mathcal{M} = \{A \mid a_i^T S a_i = p_i, i=1, \dots, m\} \quad (7.5)$$

但し、IIにおいて、 a_i は行列 A の第 i 行ベクトルである。

線形観測システム $(A, S, \sigma^2 R)$ において、観測ベクトル z の各成分 z_i が電気的信号として与えられているものとする。観測誤差が加わらない場合の観測ベクトル z の総平均電力 $E(\|z\|^2)$ および各成分 z_i の平均電力 $E(z_i^2)$ は、次のようになる。

$$\begin{aligned}
& E(\|z\|^2) \\
&= E(\tau z z^T) = E(\tau A x x^T A^T) \\
&= \tau A S A^T \tag{7.6}
\end{aligned}$$

$$\begin{aligned}
& E(z_i^2) \\
&= E(a_i^T x x^T a_i) \\
&= a_i^T S a_i \tag{7.7}
\end{aligned}$$

このことから、最適化問題Ⅰは、観測ベクトル z の総平均電力が p_0 に抑えられている場合の行列 A の最適化である。また、最適化問題Ⅱは、観測ベクトル z の各成分 z_i の平均電力が それぞれ p_i に抑えられている場合の行列 A の最適化である。

並列伝送系の立場からみた最適化の問題として、最適化問題Ⅰ及び最適化問題Ⅱは、すでにそれぞれ参考文献(16)及び(12)において取り扱われているが、本章では、線形観測システムの最適化の問題として、それぞれ、7.2と7.3において検討を行なう。なお、本章では、観測ベクトル z の次元 m が未知ベクトル x の次元 n よりも大きい場合について考察する。また、行列 R は、 $\tau R = m$ の形に正規化されているものとする。

7.2. 最適化問題Ⅰ

本節では、行列 A の実現可能な集合 $\mathcal{A}$ が (7.4) で与えられている場合の線形観測システム $(A, S, \sigma^2 R)$ の最適化問題を考える。

行列 S と行列 R はともに正値対称行列であるから、次のように対角化するこ

とができる。

$$S = G S_0 G^T, \quad S_0 = \text{diag}(s_1, s_2, \dots, s_n) \quad (7.8)$$

$$R = H R_0 H^T, \quad R_0 = \text{diag}(r_1, r_2, \dots, r_m) \quad (7.9)$$

ここで、 G 及び H は、それぞれ n 次と m 次の直交行列である。また、行列 R は、 $\text{tr } R = m$ の形に正規化されているから、

$$\sum_{i=1}^m r_m = m \quad (7.10)$$

が成り立つ。ここで、行列 A を次のように表わす。

$$A = H B G^T \quad (7.11)$$

このとき、 $J(A)$ (以後、単に J と書く) は次のようになる。

$$J = \text{tr} (I - C B) S_0 \quad (7.12)$$

但し、行列 C は、

$$C = S_0 B^T (B S_0 B^T + \sigma^2 R_0)^{-1} \quad (7.13)$$

である。また、(7.4) の条件 $\text{tr } A S A^T = p_0$ は次のようになる。

$$\text{tr } B S_0 B^T = p_0 \quad (7.14)$$

これより、行列 A の最適化問題は、制約条件 (7.14) の下に (7.12) の J を最

小化する行列 B を求める問題となる。ここで、 L を

$$L = \frac{1}{2} (I - CB) S_0 + \lambda (\frac{1}{2} B S_0 B^T - P_0) \quad (7.15)$$

とおく。但し、 λ はラグランジエの未定係数である。この L の行列 B に関する変分を求める。

$$\delta L = -\frac{1}{2} (\delta C \cdot B + C \cdot \delta B) S_0 + \lambda B S_0 \cdot \delta B \quad (7.16)$$

ここで、 $S B$ に対する行列 C の変分 δC には、次のような関係式が成り立つ。
(7.13) より、

$$C (B S_0 B^T + \sigma^2 R_0) = S_0 B^T \quad (7.17)$$

である。いま、 $B \rightarrow B + \delta B$, $C \rightarrow C + \delta C$ とおくと、

$$C (\delta B S_0 B^T + B S_0 \delta B^T) + \delta C (B S_0 B^T + \sigma^2 R_0) = S_0 \delta B^T$$

$$\delta C (B S_0 B^T + \sigma^2 R_0) = (I - CB) S_0 \delta B^T - C \cdot \delta B \cdot S_0 B^T$$

両辺の右側から $C^T = (B S_0 B^T + \sigma^2 R_0)^{-1} B S_0$ を掛けると、

$$\delta C \cdot B S_0 = (I - CB) S_0 \cdot \delta B^T C^T - C \cdot \delta B S_0 B^T C^T \quad (7.18)$$

となる。

(7.18) の右辺を (7.16) の $\delta C \cdot B \cdot S_0$ に代入すると、

$$\delta L = -\text{tr}(\mathbf{I} - \mathbf{C}\mathbf{B})\mathbf{S}_0\delta\mathbf{B}^T\mathbf{C}^T - \text{tr}\mathbf{C}\mathbf{S}\mathbf{B}\mathbf{S}_0(\mathbf{I} - \mathbf{B}^T\mathbf{C}^T) + 2\lambda\text{tr}\mathbf{B}\mathbf{S}_0\delta\mathbf{B}$$

となる。これより、次式が得られる。

$$\delta L = 2\text{tr}\{\lambda\mathbf{B} - \mathbf{C}^T(\mathbf{I} - \mathbf{C}\mathbf{B})\}\mathbf{S}_0\delta\mathbf{B}^T \quad (7.19)$$

行列 $\mathbf{B}$ が J の極値を与えるためには、任意の $\delta\mathbf{B}$ に対して $\delta L = 0$ とならなければならない。ゆえに、

$$\lambda\mathbf{B} = \mathbf{C}^T(\mathbf{I} - \mathbf{C}\mathbf{B}) \quad (7.20)$$

が行列 $\mathbf{B}$ の必要条件となる。この条件式から、行列 $\mathbf{B}$ 及び行列 $\mathbf{C}$ はともに次のような対角行列でなければならない (付録)。

$$\mathbf{B} = \begin{bmatrix} \mathbf{B}_0 \\ \vdots \\ \mathbf{B}_0 \\ \vdots \\ 0 \end{bmatrix}, \quad \mathbf{B}_0 = \text{diag}(b_1, b_2, \dots, b_n) \quad (7.21)$$

$$\mathbf{C} = \begin{bmatrix} \mathbf{C}_0 & \vdots & 0 \end{bmatrix}, \quad \mathbf{C}_0 = \text{diag}(c_1, c_2, \dots, c_n) \quad (7.22)$$

次に、行列 $\mathbf{B}$, $\mathbf{C}$ の対角成分の最適化を行なう。はじめに $\{c_i\}$ は、(7.13) より、次のようになる。

$$c_i = \frac{s_i b_i}{s_i b_i^2 + \sigma^2 r_i}, \quad i = 1, \dots, n \quad (7.23)$$

ここで、(7.20) の対角成分を比較すると

$$\lambda b_i = \frac{s_i b_i}{s_i b_i^2 + \sigma^2 r_i} \cdot \frac{\sigma^2 r_i}{s_i b_i^2 + \sigma^2 r_i} \quad (7.24)$$

となり、これより、 b_i が定められる。即ち、

$$b_i = \frac{1}{\sqrt{s_i}} \left(\sqrt{\frac{\sigma^2 r_i s_i}{\lambda}} - \sigma^2 r_i \right)^{\frac{1}{2}} \quad (注) \quad i=1, \dots, n \quad (7.25)$$

但し、定数 λ は、制約条件(7.14)より定まる。即ち、

$$\lambda = \frac{\sigma^2}{(p_0 + n\sigma^2)^2} \left(\sum_{i=1}^n \sqrt{s_i r_i} \right)^2 \quad (7.26)$$

である。

(7.25) 及び (7.26) より、 J の最小値 J^* は、次のようになる。

$$J^* = \text{tr} (I - CB) S_0$$

$$= \sum_{i=1}^n \frac{\sigma^2 r_i s_i}{s_i b_i^2 + \sigma^2 r_i}$$

$$= \sum_{i=1}^n \sqrt{\sigma^2 \lambda r_i s_i}$$

$$\therefore J^* = \frac{\sigma^2}{p_0 + n\sigma^2} \left(\sum_{i=1}^n \sqrt{r_i s_i} \right)^2 \quad (7.27)$$

以上で、 J の最小化はなされた。但し、ここでは、はじめに行列 S と R を、(7.8)、(7.9) のように対角化して議論を進めたが、各対角行列の対角成分の配置は自由である。この配置は、(7.27) が最小化するように与えなければなら

(注) 括弧の中が負になるときは、若干の修正が必要である。

らなけ。このためには、 $\{s_i\}$, $\{r_i\}$ を次のように与えれば良い。このとき、 J が最小となるのは、明らかである。

$$\begin{cases} s_1 \geq s_2 \geq \dots \geq s_n \\ r_1 \leq r_2 \leq \dots \leq r_n \leq r_{n+1} \leq \dots \leq r_m \end{cases} \quad (7.28)$$

以上の結果をまとめる。

〔最適化問題Ⅰ〕 線形観測システム $(A, S, \sigma^2 R)$ における行列 A の実現可能な集合を (7.4) とする。このとき、 J を最小とする行列 A は、次式で与えられる。

$$A = H \tilde{B} G^T \quad (7.29)$$

ここで、 H と G とは、(7.28) を満足するように行列 S 及び R を対角化する直交行列であり、(7.8), (7.9) をそれぞれ満足している。行列 $\tilde{B}$ は、(7.21) の形の対角行列であり、対角成分 b_i は (7.25) で与えられるものである。

なお、このときの J の最小値 J^* は (7.27) に与えられている。

7.3 最適化問題Ⅱ

本節では、行列 A の実現可能な集合が (7.5) で与えられている場合の線形観測システム $(A, S, \sigma^2 R)$ の最適化問題を考える。

(7.5) の集合における最適化問題は、最適化問題Ⅰのように厳密な最適化、

が容易となる。本節では、観測ベクトル区に加わる観測誤差が各成分に対して互いに独立に加わる場合、即ち、行列 R が対角行列である場合について、最適化問題を考える。

(7.5) の条件 $\{a_i^T S a_i = p_i, i=1, \dots, m\}$ を行列表現すると、

$$A S A^T = P \quad (7.30)$$

となる。但し、行列 P は、次式で与えられる。

$$P \equiv \begin{bmatrix} p_1 & & & \\ & p_2 & & * \\ & & \ddots & \\ * & & & p_m \end{bmatrix} \quad (7.31)$$

ここで、行列 P の対角成分以外は任意の値でよい。

行列 A を次のように表わす。

$$A = R^{\frac{1}{2}} B S^{-\frac{1}{2}} \quad (7.32)$$

ここで、行列 B は、 $m \times n$ 行列である。このとき、 J は、

$$J = \frac{1}{2} \{ I - B^T (B B^T + \sigma^2 I)^{-1} B \} S \quad (7.33)$$

となる。また、条件 (7.30) は次のようになる。

$$R^{\frac{1}{2}} B B^T R^{\frac{1}{2}} = P \quad (7.34)$$

ここで行列 R が対角行列であれば、(7.34) は次のように書き換えることができる。

ける。

$$B B^T = Q \quad (7.35)$$

$$Q \equiv \begin{bmatrix} q_1 & & & \\ & q_2 & & * \\ & & \ddots & \\ * & & & q_m \end{bmatrix} \quad (7.36)$$

ここで、 Q の対角成分 $\{q_i\}$ は、 $\{p_i\}$ と行列 R の対角成分 $\{r_i\}$ を用いて次のように与えられる。

$$q_i \equiv \frac{p_i}{r_i}, \quad i = 1, \dots, m \quad (7.37)$$

行列 A の最適化問題は、制約条件(7.35)の下に、(7.33)を最小にする行列 B を求める問題となる。ここで、更に行列 B を次のように特異値分解表現する。

$$B = U D V^T \quad (7.38)$$

行列 U と V は、それぞれ m 次と n 次の直交行列であり、行列 D は次のような対角行列である。

$$D \equiv \begin{bmatrix} D_0 \\ \hline 0 \end{bmatrix} \quad (7.39)$$

$$D_0 \equiv \text{diag}(d_1 \ d_2 \ \dots \ d_m) \quad (7.40)$$

任意の直交行列 B は仮らず (7.38) の形に表現することが出来る。このとき J は、

$$\begin{aligned} J &= \text{tr} \{ I - V D^T (D D^T + \sigma^2 I)^{-1} D V^T \} S \\ &= \text{tr} V D_1 V^T S \end{aligned} \quad (7.41)$$

となる。但し、行列 D_1 は、次式で与えられる。

$$D_1 \equiv \text{diag} (d'_1 \ d'_2 \ \dots \ d'_n) , \quad (7.42)$$

$$d'_i \equiv \frac{\sigma^2}{d_i^2 + \sigma^2} \quad , \quad i = 1, \dots, n \quad (7.43)$$

制約条件 (7.25) は、

$$U D D^T U^T = Q \quad (7.44)$$

となる。

そこで、行列 U 、行列 D 及び行列 V についてそれぞれ最適化を行なう。(7.44) の制約条件に行列 V が関係してゐないことに注意して、はじめに行列 V の最適化を行なう。いま、 $V = (V_{ij})$ とすると、(7.41) より、

$$J = \sum_{i=1}^n \sum_{j=1}^n d'_i s_j V_{ij}^2 \quad (7.45)$$

となる。この J の行列 V に対する変分を求める。いま、行列 V の次のような変分をとる。但し、ここで $1 \leq k \neq l \leq n$ である。

$$V \rightarrow R_{kl}(\delta\theta) V \quad (7.45)$$

但し、 $R_{kl}(\theta)$ は k, l 成分のつくる二次元平面上での角度 θ の回転を表わす行列である。このとき、行列 V の成分 $\{V_{ki}, V_{li}, i=1, \dots, n\}$ が次のように変化する。

$$\begin{cases} V_{ki} \rightarrow V_{ki} \cos \delta\theta - V_{li} \sin \delta\theta \\ V_{li} \rightarrow V_{ki} \sin \delta\theta + V_{li} \cos \delta\theta \end{cases} \quad (7.46)$$

, $i = 1, \dots, n$

このときの J の変分 δJ_{kl} は

$$\begin{aligned} \delta J_{kl} &= \sum_{i=1}^n d_i \left[s_k \{ (V_{ki} \cos \delta\theta - V_{li} \sin \delta\theta)^2 - V_{ki}^2 \} \right. \\ &\quad \left. + s_l \{ (V_{ki} \sin \delta\theta + V_{li} \cos \delta\theta)^2 - V_{li}^2 \} \right] \\ &= 2 \delta\theta (s_k - s_l) \sum_{i=1}^n d_i' V_{ki} V_{li} + O(\delta\theta^2) \quad (7.47) \end{aligned}$$

となる。行列 V が J の極値を与えるためには、 $\delta J_{kl} = 0$ とななければならない。一般に、 $s_k \neq s_l$ であるから、

$$\sum_{i=1}^n d_i' V_{ki} V_{li} = 0 \quad (7.48)$$

となる。

(7.48) を行列表現すると次のようになる。

$$V D_1 V^T = D_2 \quad (7.49)$$

但し、行列 D_2 は、次式で与えられる対角行列である。

$$D_2 \equiv \text{diag} (d_1'' d_2'' \cdots d_n'') \quad (7.50)$$

$$d_k'' \equiv \sum_{i=1}^n d_i' v_{ki}^2, \quad k=1, \dots, n \quad (7.51)$$

(7.49) により、行列 D_1 と行列 D_2 は共に対角行列であるから、行列 V は単位行列でなければならぬ。^(注)

$V=I$ とし、 J を求めると

$$J = \sum_{i=1}^n \frac{\sigma^2 s_i}{d_i^2 + \sigma^2} \quad (7.52)$$

となる。この J には、行列 U は関係してはいない。ゆえに、残りは、対角行列 D の最適化だけである。但し、行列 D の最適化を行なう際には、制約条件 (7.44) を満足する直交行列 U の存在する範囲で最適を行なう必要がある。この問題を直接扱うのは容易でない。このため、本節では、はじめに次の条件

$$U D D^T U^T = U Q \quad (7.53)$$

の下で対角行列 D の最適化問題を解く。そして、その最適解に対して、(7.44) を満たす直交行列 U が存在するための条件を逐に求めることにする。

(7.53) は、制約条件 (7.44) が成り立つための一つの必要条件である。この (7.53) は次のように表わすこともできる。

(注) 数学的には、置換行列も (7.49) の形式を満足するが、物理的に解釈すれば置換行列は観測値の番号の付け替えに過ぎないので、ここでは考慮しない。

$$\sum_{i=1}^n d_i^2 = \sum_{i=1}^m g_i \quad (7.54)$$

そこで

$$L = \sum_{i=1}^n \frac{\sigma^2 s_i}{d_i^2 + \sigma^2} + k \left(\sum_{i=1}^n d_i^2 - \sum_{i=1}^m g_i \right) \quad (7.55)$$

とおく。kはラグランジュ乗数係数である。各 d_i^2 について L の変分 δL_i を求める。

$$\delta L_i = - \frac{\sigma^2 s_i}{(d_i^2 + \sigma^2)^2} + k \quad (7.56)$$

対角行列 D が J の極値を与えるためには、 $\delta L_i = 0$ でなければならぬ。
 従って、 $\{d_i\}$ を求めると

$$d_i = (\lambda \sqrt{s_i} - \sigma^2)^{\frac{1}{2}} \quad (\text{注}) \quad (7.57)$$

となる。但し、定数 λ は、制約条件 (7.54) より定まる。即ち、

$$\lambda = \frac{\sum_{i=1}^m g_i + n \sigma^2}{\sum_{i=1}^n \sqrt{s_i}} \quad (7.58)$$

となる。このとき、 J の最小値 J^* は次のようになる。

$$J^* = \frac{\sigma^2}{\sum_{i=1}^m g_i + n \sigma^2} \left(\sum_{i=1}^n \sqrt{s_i} \right)^2 \quad (7.59)$$

(注). 括弧の中が負になる場合は若干の修正が必要である。

$\{d_i\}$ の最適値は、(7.57) で与えられるが、これは、条件 (7.53) の下で得られたものである。(7.57) の $\{d_i\}$ に対して、(7.44) を満足する直交行列 U が常に存在する保証はない。(7.44) を満たす直交行列 U が存在するための条件を求める必要がある。(7.44) の式が第 6 章、6.4, (6.31) と同じ形式であることに注意すると、この条件は、ちょうど、共分散として、 DD^T をもつ情報源に対して、 $\{g_i\}$ に応じた最適分散性特徴 $11^{\circ} \times -\rho$ の存在するための条件と一致することが分かる。定理 6.1 より、(7.44) の直交行列 U が存在するための条件は、次のようにして与えられる。

$$g_i < d_i \quad (\text{注}) \quad (7.60)$$

但し、ベクトル d とベクトル g は、それぞれ行列 DD^T と行列 Q の対角成分より成る次の m 次元ベクトルである。

$$d \equiv (d_1^2 \ d_2^2 \ \cdots \ d_n^2 \ 0 \ \cdots \ 0)^T \quad (7.61)$$

$$g \equiv (g_1 \ g_2 \ \cdots \ g_n \ g_{n+1} \ \cdots \ g_m)^T \quad (7.62)$$

以上の結果をまとめると、

「最適化問題 II」 線形観測システム $(A, S, \sigma^2 R)$ における行列 A の実現可能な集合を (7.5) とする。このとき、 J を最小とする行列 A は、次式で与えられる。

$$A = R^{\frac{1}{2}} \tilde{U} \tilde{D} S^{-\frac{1}{2}} \quad (7.63)$$

ここで、 $\tilde{D}$ は (7.39) の形の対角行列であり、 $\{d_i\}$ の対角成分 $\{d_i\}$ は、(7.57)

(注) $\succ$ は、定義 6.A で与えた半順序関係である。

で与えられる。また、行列 $\tilde{U}$ は、(7.44) を満足する行列である。このとき、 J の最小値 J^* は (7.59) である。

なお、この最適化において、直交行列 $\tilde{U}$ が存在するためには、(7.60) の条件が満足されるなければならない。

付録 (行列 B と行列 C がともに対角行列となることの証明)^(注)

$$\begin{aligned} C &= S_0 B^T (B S_0 B^T + \sigma^2 R_0)^{-1} \\ &= (B^T R_0^{-1} B + \sigma^2 S_0^{-1})^{-1} B^T R_0^{-1} \end{aligned} \quad (7.64)$$

であることに注意して、(7.20) の両辺の右側から $S_0 (B^T R_0^{-1} B + \sigma^2 S_0^{-1})$ を掛ける。このとき、(7.20) の右辺は、

$$\begin{aligned} &\text{右辺} \\ &= C^T (I - C B) S_0 (B^T R_0^{-1} B + \sigma^2 S_0^{-1}) \\ &= C^T S_0 (B^T R_0^{-1} B + \sigma^2 S_0^{-1}) - C^T C (B S_0 B^T + \sigma^2 R_0) R_0^{-1} B \\ &= C^T S_0 (B^T R_0^{-1} B + \sigma^2 S_0^{-1}) - C^T S_0 B^T R_0^{-1} B \\ &= \sigma^2 C^T \end{aligned}$$

となる。これより、次式が得られる。

$$\lambda B S_0 (B^T R_0^{-1} B + \sigma^2 S_0^{-1}) = \sigma^2 C^T \quad (7.65)$$

両辺の左側より、 $S_0 B^T$ を掛ける。

(注) この証明は、文献(16)における証明法に準拠している。

$$\lambda S_0 B^T B S_0 (B^T R_0^{-1} B + \sigma^2 S_0^{-1}) = \sigma^2 S_0 B^T (B S_0 B^T + \sigma^2 R_0)^{-1} B S_0 \quad (7.66)$$

ここで、 $S_0 B^T B S_0$ と $B^T R_0^{-1} B + \sigma^2 S_0^{-1}$ はともに対称行列であり、しかもその積も対称行列であることが分かる。ゆえに、この二つの行列は同時に対角化することができる。即ち、

$$S_0 B^T B S_0 = T S_1 T^T \quad (7.67)$$

$$B^T R_0^{-1} B + \sigma^2 S_0^{-1} = T S_2 T^T \quad (7.68)$$

となる。ここで、 T は直交行列、 S_1 と S_2 はともに対角行列である。(7.67), (7.68) を (7.66) に代入して整理すると、

$$\lambda T S_1 S_2 T^T = \sigma^2 (S_0 - \sigma^2 T S_2^{-1} T^T)$$

となる。これより、1次式が得られる。

$$T (\lambda S_1 S_2 + \sigma^4 S_2^{-1}) T^T = \sigma^2 S_0 \quad (7.69)$$

ここで、右辺は対角行列、また左辺の $(\lambda S_1 S_2 + \sigma^4 S_2^{-1})$ も対角行列であるから、行列 T は単位行列でなければならぬ。即ち、

$$S_0 B^T B S_0 = S_1 \quad (7.70)$$

$$B^T R_0^{-1} B + \sigma^2 S_0^{-1} = S_2 \quad (7.71)$$

となる。行列 B を次のように特異値分解する。

$$B = U D V^T \quad (7.72)$$

但し、 U と V とはそれぞれ m 次と n 次の直交行列であり、 D は次のような対角行列である。

$$D \equiv \begin{bmatrix} D_0 \\ \vdots \\ 0 \end{bmatrix}, \quad D_0 \equiv \text{diag}(d_1, d_2, \dots, d_n) \quad (7.73)$$

(7.72) を (7.70) に代入して整理すると。

$$V D^T D V^T = S_0^{-2} S_1 \quad (7.74)$$

ここで、右辺は対角行列であり、また左辺の $D^T D$ も対角行列であるから、 V は単位行列でなければならぬ。次に (7.72) を (7.71) に代入して整理すると。

$$U^T R_0^{-1} U = D^{-2} (S_2 - \sigma^2 S_0^{-1}) \quad (7.75)$$

となる。これより、やはり U も単位行列でなければならぬ。

以上のことから、 $B = D$ となり、行列 B は対角行列であることが分かる。行列 C も (7.13) より、行列 B が対角行列ならば、やはり対角行列となることが分かる。

第8章 曲線のあてはめ問題と σ -逆行列

8.1 曲線のあてはめ問題

統計的データの処理問題として最も基本的な問題の一つとして、曲線のあてはめ問題がある。本章では、 σ -逆行列による最小分散推定の立場から、曲線のあてはめ問題を扱って、従来の最小二乗推定による処理方法との対比を行なう。

独立変数 x の離散的な値 $x_1, x_2, \dots, x_n$ における観測値 $z_1, z_2, \dots, z_n$ が与えられたとき、点 (x_ν, z_ν) , $\nu=1, \dots, n$ にならばくうまくあてはまる簡単な曲線 $z = f(x)$ を求めることを、一般に曲線のあてはめ問題という。本章では、あてはめに用いる関数として、次数 m の回帰多項式を考える。回帰多項式は一般に次式で与えられる。

$$f(x) = \sum_{i=0}^m \beta_i \phi_i(x) \quad (8.1)$$

ここで、 $\phi_i(x)$ は、 i 次の多項式である。

(8.1) において、係数 $\{\beta_i; i=0, \dots, m\}$ を定めることによりあてはめる曲線が定まる。いま、本来の関数が (8.1) の形が多項式であり、各観測点 x_i における観測誤差がなれとすれば、係数 $\{\beta_i\}$ と観測値 $\{z_i\}$ との間には、次の関係が成り立つ。

$$z = \Phi \cdot \beta \quad (8.2)$$

但し、

$$z \equiv (z_1 \ z_2 \ \dots \ z_n)^T \quad (8.3)$$

$$\Phi \equiv \begin{bmatrix} \phi_0(x_1) & \phi_1(x_1) & \cdots & \phi_m(x_1) \\ \phi_0(x_2) & \phi_1(x_2) & \cdots & \phi_m(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_0(x_n) & \phi_1(x_n) & \cdots & \phi_m(x_n) \end{bmatrix} \quad (8.4)$$

$$\mathbf{b} \equiv (\beta_0 \ \beta_1 \ \cdots \ \beta_m)^T \quad (8.5)$$

である。

連立方程式 (8.2) を解くことにより、観測値 $\mathbf{z}$ から、係数ベクトル $\mathbf{b}$ を求めることができる。実際には、観測値 $\{z_i\}$ には、観測誤差が含まれているから、(8.2) の最小二乗解を求めることになる。即ち、与えられた $\mathbf{z}$ に対して、

$$\|\mathbf{z} - \Phi \mathbf{b}\|^2 \quad (8.6)$$

が最小となる $\hat{\mathbf{b}}$ を求める。このベクトル $\hat{\mathbf{b}}$ は、行列 Φ の Moore-Penrose 逆行列 Φ^+ を用いて、

$$\hat{\mathbf{b}} = \Phi^+ \mathbf{z} \quad (8.7)$$

として求めることができる。最小二乗解 $\hat{\mathbf{b}} = (\hat{\beta}_0 \ \hat{\beta}_1 \ \cdots \ \hat{\beta}_m)^T$ により、あてはめ曲線は次のようになる。

$$\hat{f}(x) = \sum_{i=0}^m \hat{\beta}_i \phi_i(x) \quad (8.8)$$

以上が、最小二乗法による曲線のあてはめ問題の処理方法である。観測値 $\{z_i\}$ の数 n が回帰多項式の次数 m に比べて十分大きい場合には、この方

法により、問題なく曲線のおてはめを行なうことができる。しかし、 n が m に比べて十分に大きくない場合には、いわゆる、おてはめ過ぎ (over fitting) の問題が生いてくる。回帰多項式の自由度、即ち次数 m に比べて、観測値の数が十分でなると、各観測値に含まれている観測誤差に過敏に反応してしまう結果、信頼性のある曲線のおてはめができなくなるのである。

この問題を最小二乗法の範囲で解決するには、回帰多項式の次数 m を、観測値の数 n に比べて十分に小さく選ぶしか方法がない。しかし、あまり次数を小さく選んでしまうと、回帰多項式の表現力が落ち、真の曲線からの偏りが大きくなるので注意しなければならぬ。このように、最小二乗法による曲線のおてはめにおいては、回帰多項式の次数 m を適切に選ぶことが非常にむずかしい問題となるのである。

8.2 の逆行列による曲線のおてはめ

最小二乗法による曲線のおてはめにおいては、観測値の数 n の大きさに対して、回帰多項式の次数 m を十分に小さく設定しておく必要がある。本来、対象としている曲線がある程度の多様性をもっており、回帰多項式の次数 m が予めある大きさに定められているとする。しかも、その次数の大きさに比べて、十分な数の観測値が得られない場合を考えよう。このとき、回帰多項式の次数を下げずに、しかも安定した曲線のおてはめを行なうにはどうしたらよいであろうか。

このためには、曲線のおてはめの推定モデルの中に観測値に含まれている観測誤差の統計的性質を組み込むことにより、観測誤差の影響を過度に受けずに曲線のおてはめを行なう必要がある。いま、観測点 x_i における観測値 y_i には、観測誤差 ε_i が含まれているものとする。このとき、(8.2) は、次のような統計的枠組観測システムとして捉え直すことができる。

$$\mathbf{z} = \mathbf{X} \mathbf{b} + \mathbf{e} \quad (8.9)$$

但し、 $\mathbf{e} \equiv (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)^T$ である。

ベクトル $\mathbf{b}$ とベクトル $\mathbf{e}$ がともに確率ベクトルであり、その平均が零ベクトル、共分散がそれぞれ $\mathbf{S}$ と $\sigma^2 \mathbf{R}$ で与えられるならば、(8.9) は、線形観測システム $(\mathbf{X}, \mathbf{S}, \sigma^2 \mathbf{R})$ となる。このとき、行列 $\mathbf{X}$ の重み付きの逆行列 $\mathbf{X} \mathbf{S} \mathbf{R}$ を用いることにより、ベクトル $\mathbf{b}$ の最小分散推定量 $\hat{\mathbf{b}}$ が求められる。即ち、

$$\hat{\mathbf{b}} = \mathbf{X} \mathbf{S} \mathbf{R} \mathbf{z} \quad (8.10)$$

となる。これは、ちょうど (8.7) の行列 $\mathbf{X}$ の Moore-Penrose 逆行列 $\mathbf{X}^+$ を $\mathbf{X} \mathbf{S} \mathbf{R}$ に置き換えた形式となっている。

このように、対象となる曲線の統計的性質及び観測誤差の統計的性質が事前に分かっている場合には、行列 $\mathbf{X}$ の $\mathbf{X} \mathbf{S} \mathbf{R}$ を用いることにより、曲線のあてはめ問題を解くことができる。この方法では、観測誤差の大きさを統計的に評価した推定を行なうために、観測値の数が、回帰多項式の次数 m に比べて十分に大きくなる場合でも、観測誤差の影響に過敏となり、信頼性のある曲線のあてはめが可能となる。

$\mathbf{X} \mathbf{S} \mathbf{R}$ の逆行列を用いた曲線のあてはめの数値実験例を次に示す。

ここでは、次のような回帰多項式を考える。

$$f(x) = \sum_{i=0}^m \beta_i x^i \quad (8.11)$$

このとき、行列 $\mathbf{X}$ は

$$\Phi = \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^m \\ 1 & x_2 & x_2^2 & \cdots & x_2^m \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^m \end{bmatrix} \quad (8.12)$$

となる。数値実験では $m=11$ とし、区間 $[-1, 1]$ を 24 等分して、24 個の観測点 $\{x_i: i=1, \dots, 24\}$ を定めた。 $\{\beta_i: i=0, \dots, 11\}$ の推定は、次式で行なった。

$$\hat{b} = \Phi^T \Phi^{-1} \Phi^T \cdot x \quad (8.13)$$

実験では、真の曲線を図 8.1, (a) に示すように、 $f(x) = 4x^3 - 3x$ とした。各観測点に加わる観測誤差は、図 8.1, (b) のように与えた。この値は乱数表を用いて与えたものであり $\sigma = 0.01$ とした。図 8.1, (c) の $\times$ 印が観測値 $\{z_i: i=1, \dots, 24\}$ である。

の逆行列 $\Phi^T \Phi$ による曲線のあてはめの結果を図 8.2, (a) ~ (f) に示す。この実験では、曲線の推定における σ の値として、10.0 ~ 0.0 の間で 8 個の値を定め、各 σ に対して、推定した回帰多項式 $\hat{f}(x)$ を求め、同時に $\hat{f}(x)$ と真の曲線との間の二乗誤差を算出した。図 8.2, (d) が $\sigma = 0.01$ の場合の曲線のあてはめの様子である。これが観測誤差の σ の値に等しい σ を用いたときの結果である。 σ が 0.01 より大きな値の場合 (図 8.2, (a) ~ (c)) は、観測誤差の存在を過度に評価している。このために、推定した回帰多項式 $\hat{f}(x)$ は、観測値への適合性よりもより滑らかな曲線であることを重視したのとなっている。特に $\sigma = 10.0$ の場合は、顕著な例であり、殆ど観測値を無視した直線に近い曲線があてはめられている。逆に σ が 0.01 より小さい場合 (図 8.2, (e) ~ (f)) には、観測値への適合性をより重視した曲線のあてはめが行なわれる。 σ が小さくなればなる程、推定した関数が観測誤差によ

と引き摺られる傾向が強くなる。図 8.2, (c) は、 $\sigma=0.0$ の場合である。これは、ちょうど最小二乗法による推定の結果である。いわゆる、あてはめ過ぎの現象が明らかに現われている。

各 σ に対する推定多項式 $\hat{f}(x)$ と真の曲線との間の平均二乗誤差 (MSQDEV) をプロットしたグラフが 図 8.3 である。このグラフからも、 $\sigma=0.01$ の場合に最も良い曲線のあてはめが行なわれているのが明らかである。

最小二乗法による曲線のあてはめの場合と、 σ -逆行列を用いた曲線のあてはめの場合の観測誤差の影響を比較するため、各観測点における観測誤差と曲線の推定誤差をプロットしたのが、図 8.4, (a), (b) である。図 8.4, (a) は、 $\sigma=0.01$ の σ -逆行列を用いた場合であり、図 8.4, (b) は、最小二乗法 ($\sigma=0.0$) の場合である。図 8.4, (b) の最小二乗法による曲線のあてはめの場合の方が、図 8.4, (a) の σ -逆行列による場合に比べて、より大きく観測誤差の影響を受けていることが分かる。

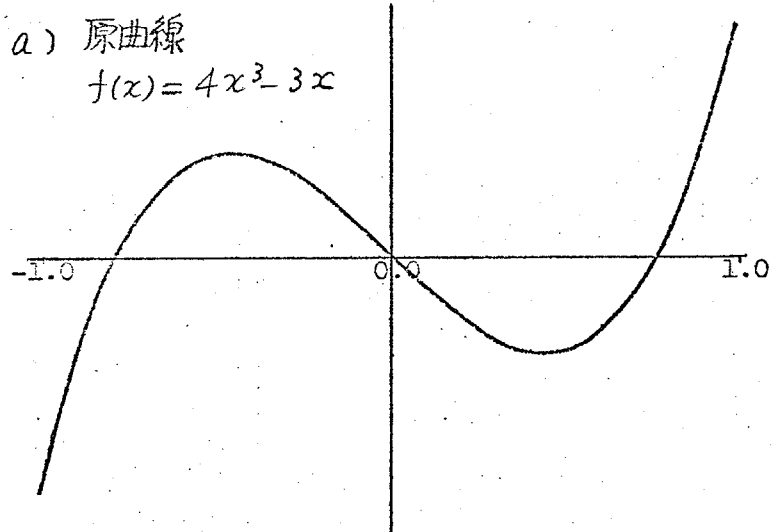
8.3 線形観測システムにおける σ の推定方法

対象とする曲線 $f(x)$ が観測誤差の統計的性質が既知の場合には、 σ -逆行列を用いた最小分散推定を行なうことにより、最小二乗法よりも優れた曲線のあてはめが可能である。しかし、実際には、観測点での観測値だけが与えられており、統計的性質に関する事前の情報がない場合、あるいは非常に乏しい場合が少なくない。このような場合にも、 σ -逆行列による曲線のあてはめを行なうには、何人らかの方法により、予め統計的性質を推定しておく必要がある。本節では、線形観測システム (A, σ^2) における σ の推定方法について考察を行なう。

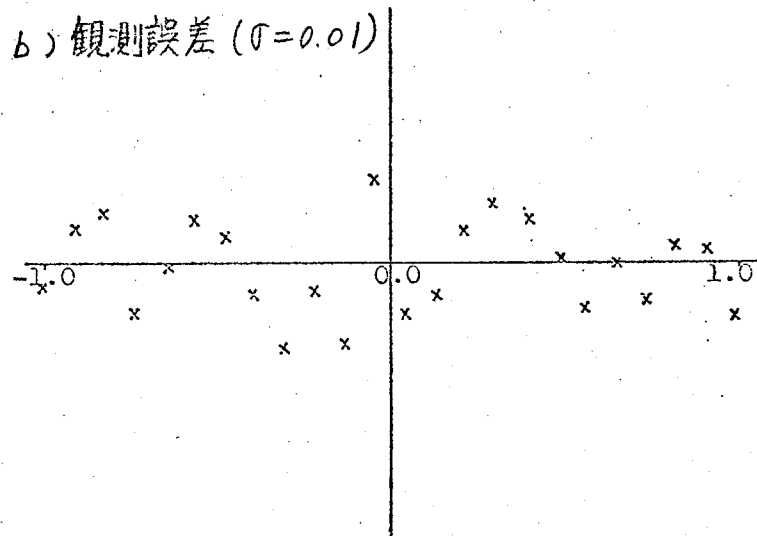
8.2 の数値実験例からも明らかをように、いま、 σ の推定値 $\hat{\sigma}$ として、現実の線形観測システム (A, σ^2) の σ の値を用いているならば、統計的な平均の意味で最も良い曲線のあてはめが行なえるであろう。そこで、予想される

(a) 原曲線

$$f(x) = 4x^3 - 3x$$



(b) 観測誤差 ($\sigma = 0.01$)



(c) 観測値

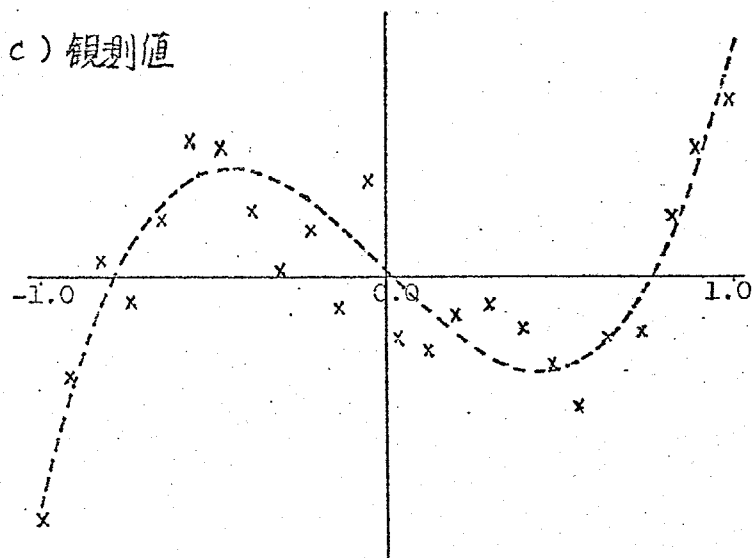
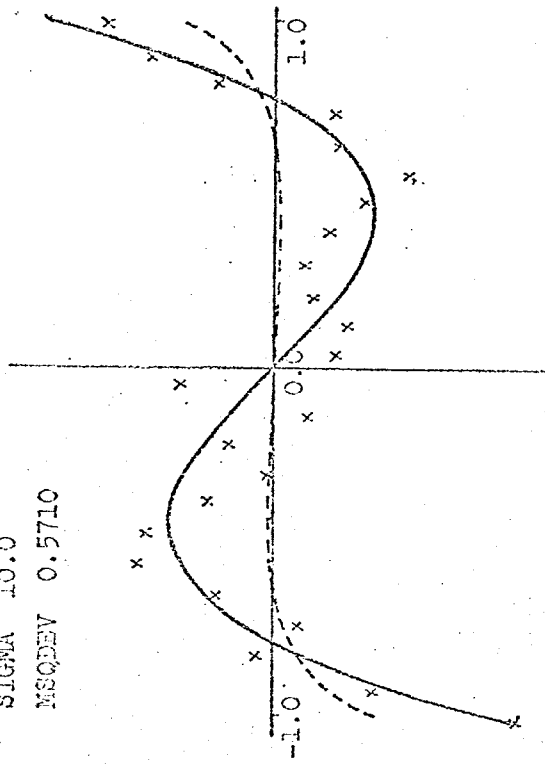
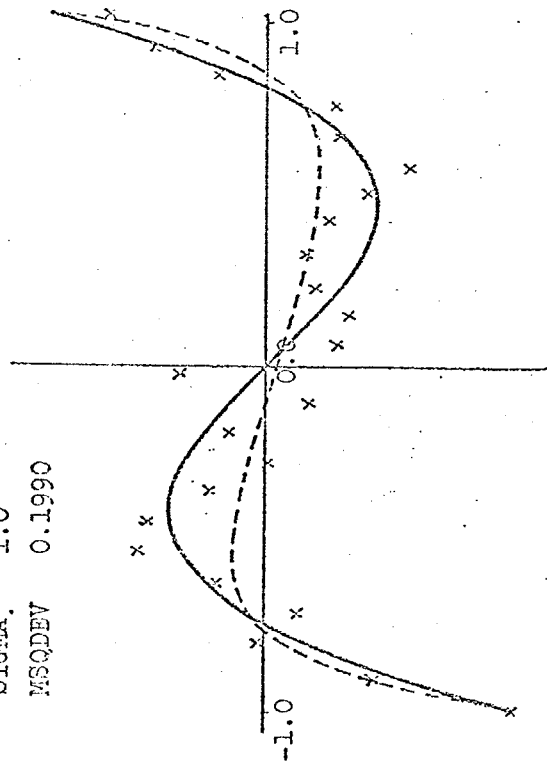


図 8.1 曲線のあてはめ問題

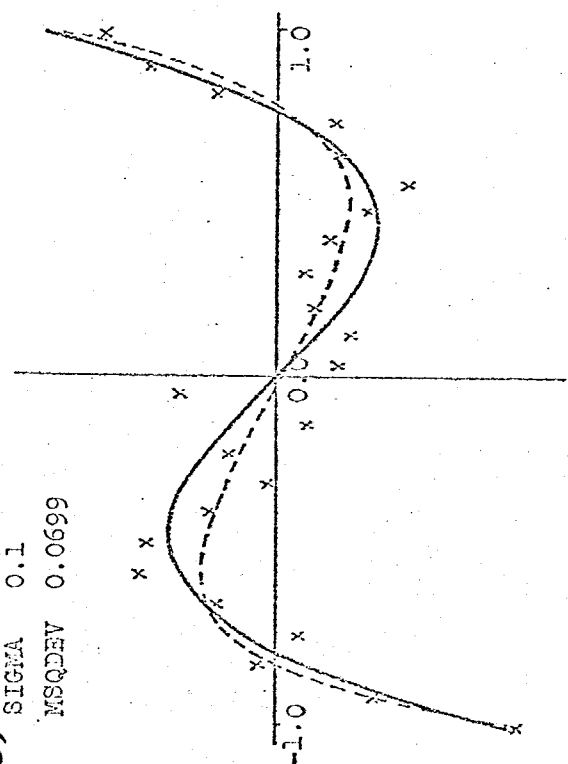
(a) SIGMA 10.0
MSQDEV 0.5710



(b) SIGMA 1.0
MSQDEV 0.1990



(c) SIGMA 0.1
MSQDEV 0.0699



(d) SIGMA 0.01
MSQDEV 0.0394

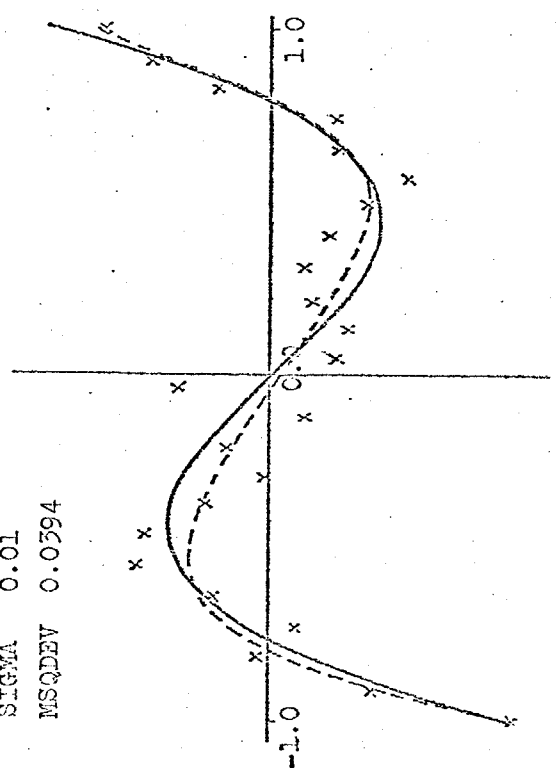
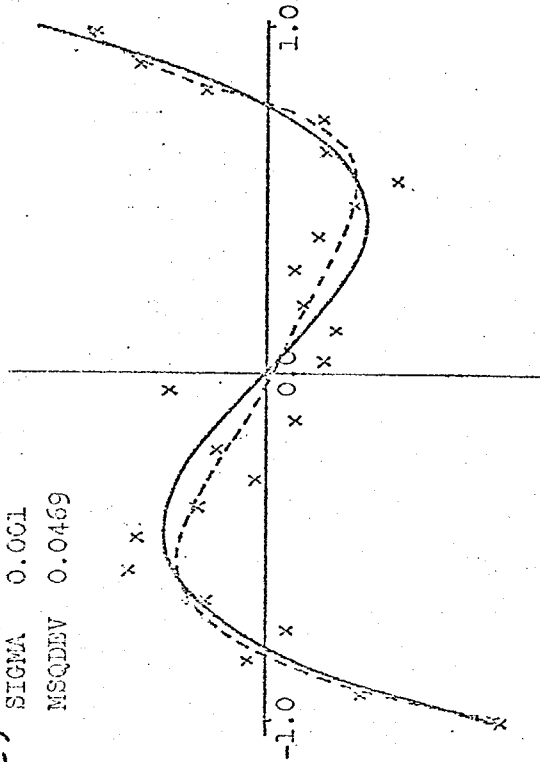
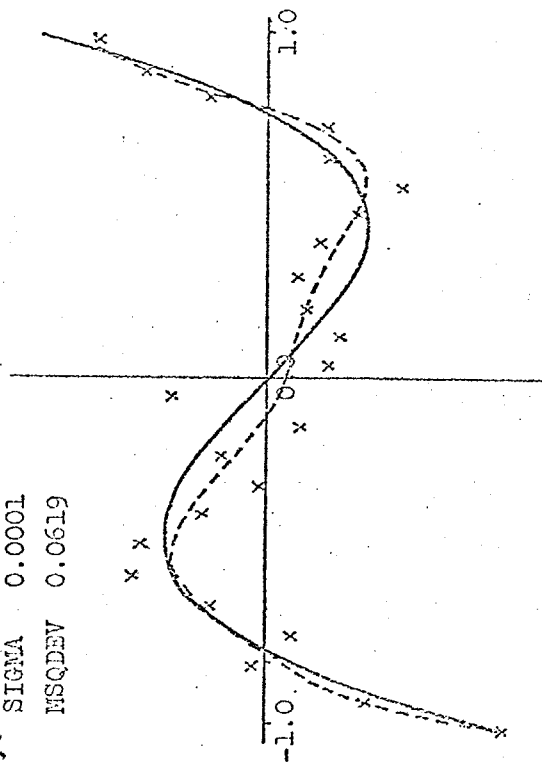


図 8.2 σ -逆行列による曲線のあてはめ (a) ~ (d)

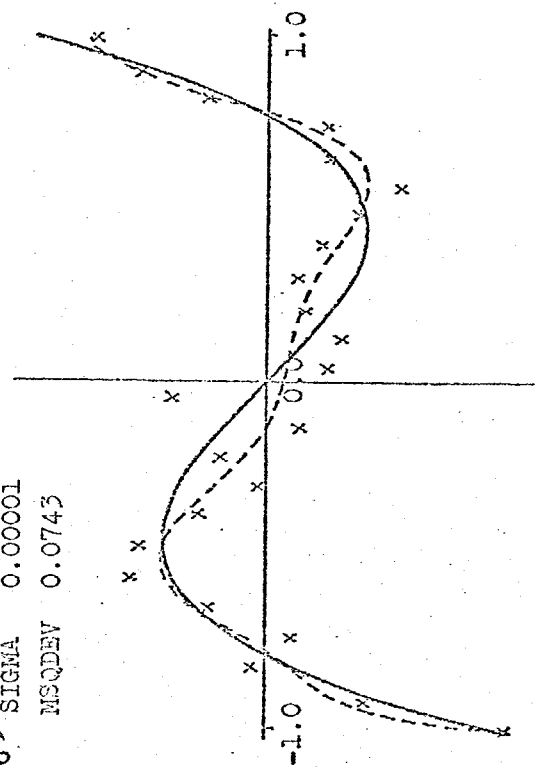
(e) SIGMA 0.061
MSQDEV 0.0469



(f) SIGMA 0.0001
MSQDEV 0.0619



(g) SIGMA 0.00001
MSQDEV 0.0743



(h) SIGMA 0.0
MSQDEV 0.0359

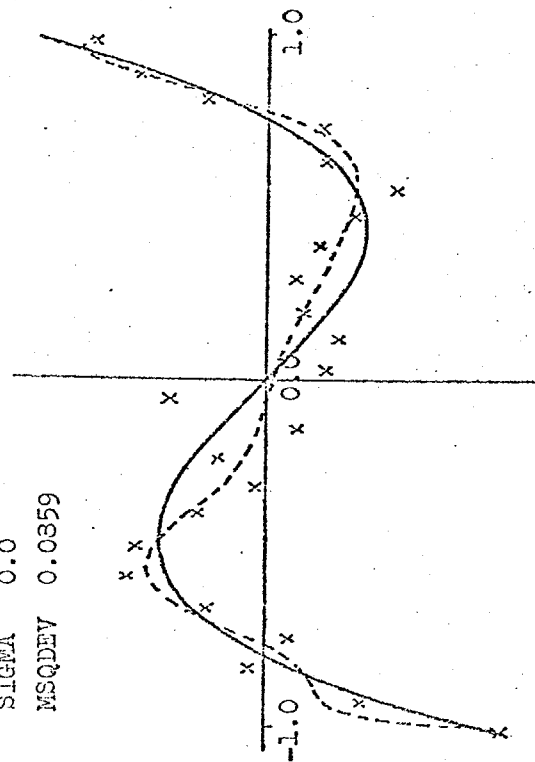


図 8.2. σ -逆行列による曲線のあてはめ (e)~(h)

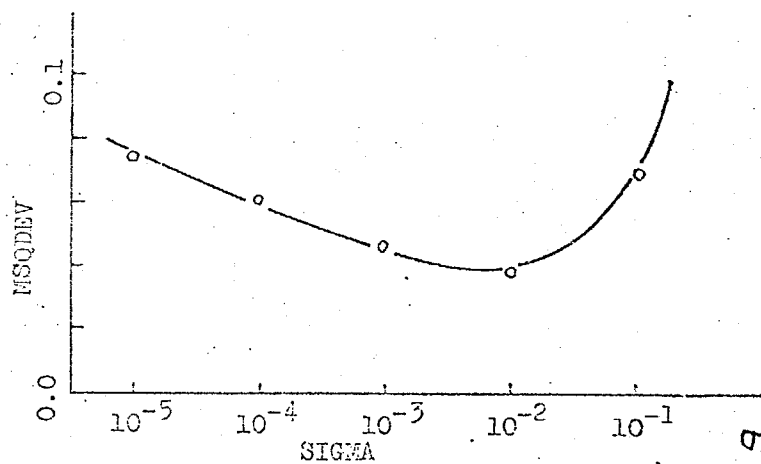
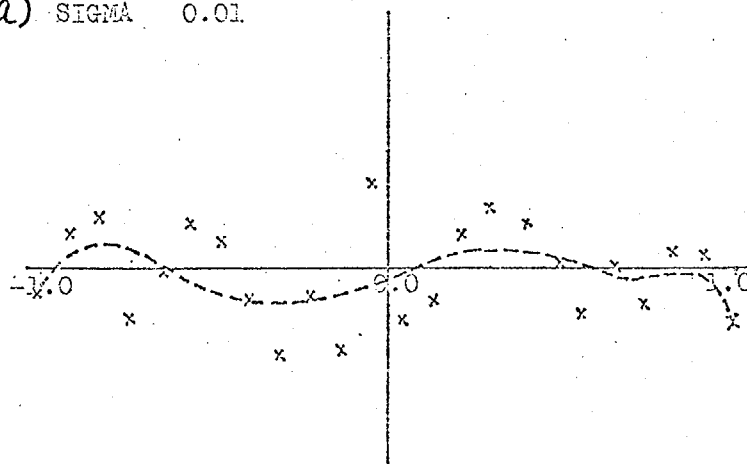


図 8.3 σ に対する平均二乗誤差

(a) SIGMA 0.01



(b) SIGMA 0.0 (最小二乗法)

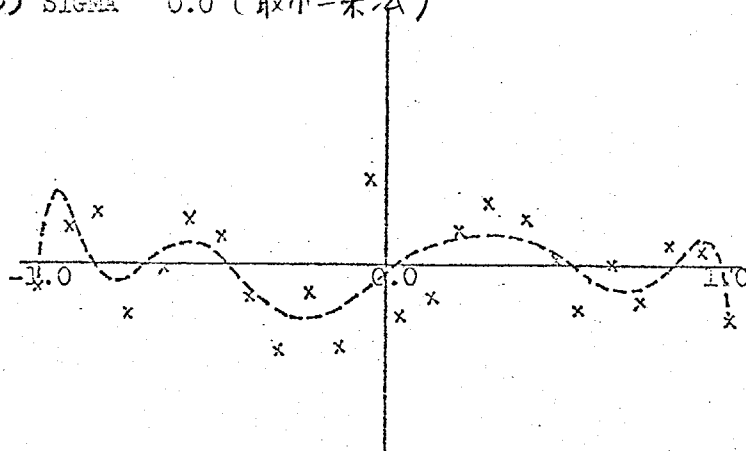


図 8.4 観測誤差と曲線の推定誤差

幾つかの σ の値について、 σ -逆行列 A^σ による曲線のおくはめを行ない、推定した回帰多項式と真の曲線との間の平均二乗誤差が最小となる σ の値を求め、それを、 σ の推定値 $\hat{\sigma}$ とすればよい。しかし、実際には、対象とする曲線の真値を推定するために σ -逆行列を使うのであるから、真の曲線は勿論分っている。ゆえに、上のような方法で σ の適合性を評価することは不可能である。そこで、ここでは、与えられた観測値だけを用いて σ の適合性を検定するモデルを次のように構成する。

(1) 線形観測システム (A, σ^2) を次のように二つのシステムに分割する。

$$\begin{cases} z_1 = A_1 x + w_1 \\ z_2 = A_2 x + w_2 \end{cases} \quad (8.14)$$

但し、

$$z = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}, \quad A = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}, \quad w = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \quad (8.15)$$

である。

σ の推定値 $\hat{\sigma}$ を与えて、行列 A_1 の σ -逆行列 $A_1^{\hat{\sigma}}$ を求める。観測ベクトル z_1 から、 σ -逆行列 $A_1^{\hat{\sigma}}$ を用いて、ベクトル x の最小分散推定量 $\hat{x}_1$ を得る。即ち、

$$\hat{x}_1 = A_1^{\hat{\sigma}} z_1 \quad (8.16)$$

(2) (1)で求めたベクトル x の推定量 $\hat{x}_1$ を用いて、観測ベクトル z_2 の推定を行なう。即ち、

$$\hat{x}_2 = A_2 \hat{x}_1 \quad (8.17)$$

である。

(3) 実際には得られていない観測ベクトル x_2 とその推定量 $\hat{x}_2$ との間の平均二乗誤差 $\|x_2 - \hat{x}_2\|^2$ を求める。

この平均二乗誤差は、 σ の推定値 $\hat{\sigma}$ の関数である。予想される幾つかの σ の推定値 $\hat{\sigma}$ に対して、上の手続により、平均二乗誤差を求め、その値が最小になる $\hat{\sigma}$ を求める σ の推定値とする。

本来は、ベクトル x とその推定量 $\hat{x}$ との間の誤差から、 σ の適合性を評価したるのであるが、それは不可能である。その代りに、観測ベクトル x_2 とその推定量 $\hat{x}_2$ との間の誤差を調べることにより、 σ の適合性を調べようというのが、(1) ~ (3) に示した方法である。観測ベクトル x_1, x_2 に加わっている観測誤差 w_1 と w_2 は互いに独立であるから、 $\|x - \hat{x}\|^2$ の代りに $\|x_2 - \hat{x}_2\|^2$ が σ の適合性の検定に利用できることが分かる。実際に、 $\|x_2 - \hat{x}_2\|^2$ の統計的平均 $E\|x_2 - \hat{x}_2\|^2$ は、 $\hat{\sigma} = \sigma$ の近傍において最小値を取ることが次に示される。いま、

$$f(\hat{\sigma}) = E\|x_2 - \hat{x}_2\|^2 \quad (8.18)$$

とおく。このとき、

$$\begin{aligned} f(\hat{\sigma}) &= E\|(A_2 x + w_2) - A_2 A_1^{\hat{\sigma}} (A_1 x + w_1)\|^2 \\ &= E\|A_2 (I - A_1^{\hat{\sigma}} A_1) x\|^2 + E\|w_2\|^2 + E\|A_2 A_1^{\hat{\sigma}} w_1\|^2 \end{aligned}$$

$$\begin{aligned}
&= \text{tr} A_2 (I - A_1 \hat{\sigma} A_1) (I - A_1 \hat{\sigma} A_1)^T A_2^T \\
&\quad + n_2 \sigma^2 + \sigma^2 \text{tr} A_2 A_1 \hat{\sigma} A_1 \hat{\sigma}^T A_2^T \\
&= \text{tr} A_2 A_1 \hat{\sigma} (A_1 A_1^T + \sigma^2 I) A_1 \hat{\sigma}^T A_2^T \\
&\quad - 2 \text{tr} A_2 A_1 \hat{\sigma} A_1 A_2^T + \text{tr} A_2 A_2^T + n_2 \sigma^2
\end{aligned} \tag{8.19}$$

但し、 n_2 は、ベクトル $\mathbb{R}_2$ の次元である。 $n=2$ 。

$$A_1 A_1^T + \sigma^2 I = K \tag{8.20}$$

$$\hat{\sigma}^2 - \sigma^2 = \alpha \tag{8.21}$$

とあって、 $f(\hat{\sigma})$ を整理すると、次のようになる。

$$\begin{aligned}
&f(\hat{\sigma}) \\
&= \text{tr} A_2 A_1^T (K + \alpha I)^{-1} K (K + \alpha I)^{-1} A_1 A_2^T \\
&\quad - 2 \text{tr} A_2 A_1^T (K + \alpha I)^{-1} A_1 A_2^T + \text{tr} A_2 A_2^T + n_2 \sigma^2 \\
&= \text{tr} A_2 A_1^T \{ (K + \alpha I)^{-1} K (K + \alpha I)^{-1} - 2 (K + \alpha I)^{-1} \\
&\quad + K^{-1} \} A_1 A_2^T + f(\sigma)
\end{aligned} \tag{8.22}$$

$n=2$ (8.22) の括弧 $\{ \}$ の中を整理すると、

$$\begin{aligned}
&(K + \alpha I)^{-1} K (K + \alpha I)^{-1} - 2 (K + \alpha I)^{-1} + K^{-1} \\
&= (K + \alpha I)^{-1} \{ K - 2 (K + \alpha I) + (K + \alpha I) K^{-1} (K + \alpha I) \} (K + \alpha I)^{-1} \\
&= \alpha^2 (K + \alpha I)^{-1} K^{-1} (K + \alpha I)^{-1}
\end{aligned} \tag{8.23}$$

となる。=4より、 $f(\hat{\sigma})$ は次のようになる。

$$f(\hat{\sigma}) = x^2 \text{tr} A_2 A_1^T (K+xI)^{-1} K^{-1} (K+xI)^{-1} A_1 A_2^T + f(\sigma) \quad (8.24)$$

いま、 $\hat{\sigma} = \sigma$ の近傍においては、 $x \neq 0$ であるから、(8.24) は更に次のようになる。

$$f(\hat{\sigma}) = x^2 \text{tr} A_2 A_1^T (A_1 A_1^T + \sigma^2 I) A_1^T A_2^T + f(\sigma) + O(x^3) \quad (8.25)$$

=4より、 $f(\hat{\sigma}) = E \|z_2 - \hat{z}_2\|^2$ は、 $\hat{\sigma} = \sigma$ の近傍において最小値を取ることが分かる。

(1), (2), (3) による σ の推定を、8.3の数値実験例について行なった結果を図8.5に示す。この例では、 z_1 として、観測点 $x_1, x_3, \dots, x_{23}$ の12個の観測値 $z_1, z_3, \dots, z_{23}$ を割り当て、 z_2 として、観測点 $x_2, x_4, \dots, x_{24}$ の12個の観測値 $z_2, z_4, \dots, z_{24}$ を割り当てた。 10^{-5} から 1.0 までの幾つかの $\hat{\sigma}$ に対して、 $\|z_2 - \hat{z}_2\|^2$ の値をプロットしたものが、図8.5の実線を示すグラフである。

この実験例では、 $\hat{\sigma} = 0.03$ の近傍において $\|z_2 - \hat{z}_2\|^2$ が最小値となるから、 $\hat{\sigma} = 0.03$ が σ の推定値となる。実際の σ の値は、 $\sigma = 0.01$ であるから、比較的よい σ の推定値が求められることが分かる。

$\|z_2 - \hat{z}_2\|^2$ が $\hat{\sigma} = \sigma$ の近傍において最小となるといふことは、(8.25)に示すようにあくまでも統計的な性質として成り立つものである。図8.5の実験例では、グラフが単峰性となり、比較的真の σ の値の近傍において最小値が得られているが、観測誤差の加わり方の如何により、図8.5の点線を示すよ

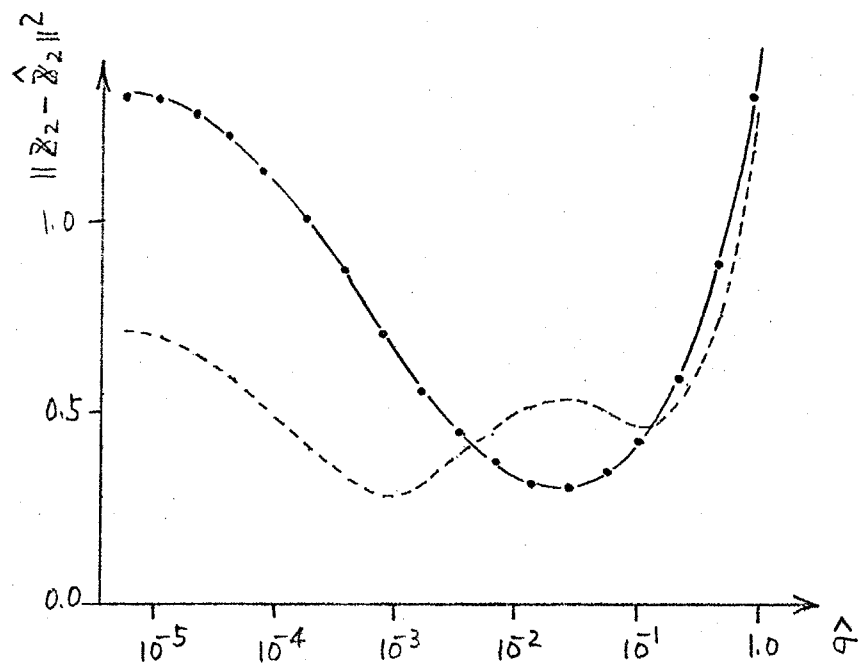


図 8.5 $\hat{\sigma}$ に対する $\|z_2 - \hat{z}_2\|^2$ のグラフ

うな双峰性のグラフとなる場合もある。この場合には、必ずしも真値に近い σ の推定を行なうことはできない。

第 9 章 σ -逆行列の画像処理への応用

9.1 まえがき

本章では σ -逆行列の画像処理への応用として、投影像からの断層像再構成の問題を扱う。

三次元物体の断層像をその投影像から再構成する問題は、天体物理学、顕微鏡検査、工業的な非破壊検査、放射線学などの多くの分野において応用される問題である。近年は、特に X 線による医用診断への実用化が可能となり、CT (Computerized Tomography) と呼ばれる新しい画像処理技術の分野を確立しつつある。

断層像の再構成の方法にも種々のものが提案されている。代数的方法、同時逐次近似法 (SIRT)、フーリエ法、コンホリユーショニ法などが主なものである。これらの方法は、いずれも豊富な投影像の観測と高速度、大容量の情報処理技術を前提としたものである。しかし、投影像の観測が十分に行なえない場合、あるいは、再構成のための情報処理をできるだけ簡易化する必要がある場合には、対象とする断層像の統計的性質及び投影像の観測に加わる測定誤差の統計的性質を十分に考慮に入れた断層像の再構成をしなければならぬ。本章では、投影像の観測過程を一つの統計的線形観測システムとして捉え、 σ -逆行列を用いた最小分散推定により、断層像の再構成を行なうことにより、上の問題に対する解決方法を与えている。

9.2 では、投影像の観測過程の記述を行ない、線形観測システムを明らかにする。本章では、断層像をガウス型重み関数と二変数多項式の積として表現する。この表現形式は、直交座標系の回転に対して不変な形式であるために、線形観測システムの記述が正確にかつ簡明に行なうことのできる利点を有して

いる。9.3 において、断層像の再構成システムを明らかにする。第4章、4.3の結果を用いて、断層像の逐次推定システムを構成し、さらに幾つかの仮定の下に、システムの簡易化を行なう。

一般に断層像の再構成に用いられる投影像のデータ量は膨大なものであり、かなりの処理時間を必要とする。このために、より少ない投影像の観測で、しかも十分な精度の断層像が再構成できるように、投影角の最適化の問題が起きてくる。また、X線による医用診断においては、医学的立場からX線の低線量化が望まれており、この点からも、投影角の最適化問題は重要な問題である。本章では、9.4 においてこの問題を扱う。

9.2 投影像の観測システム

直交座標系 (x, y) における断層像 $f(x, y)$ をガウス型重み関数と二変数多項式との積で表現する。即ち、

$$f(x, y) = \frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}} \sum_{m=0}^M \sum_{k=0}^m f_{mk} \varphi_{mk}(x, y) \quad (9.1)$$

ここで、

$$\varphi_{mk}(x, y) \equiv \frac{x^k y^{m-k}}{\sqrt{k!(m-k)!}} \quad (9.2)$$

である。(9.1) において、二変数多項式の次数 M を十分に大きくとることに
より、断層像を精度良く表現することが出来る。更に、座標系 (x, y) を適当
に選ぶことにより、低い次数 M でも断層像の大域的構造を良く表現することが
出来る。

直交座標系 (x, y) と角度 θ をなす新しい直交座標系 (x', y') により、断層像 $f(x, y)$ を表現し直したものを $f^\theta(x', y')$ とおく。 $f^\theta(x', y')$ は、やはり (9.1) と同じ表現形式で表わされる。即ち、

$$f^\theta(x', y') = \frac{1}{2\pi} e^{-\frac{x'^2 + y'^2}{2}} \sum_{m=0}^M \sum_{k=0}^m f_{mk}^\theta \varphi_{mk}(x', y') \quad (9.3)$$

である。これは、(9.1) で与えたガウス型重み関数と二変数多項式が共に座標系の回転に対して不変な形式であるためである。

断層像 $f(x, y)$ の角度 θ 方向への投影像を $p(u; \theta)$ とおく。 $p(u; \theta)$ は、 $f^\theta(x, y)$ を y について積分することにより得られる。即ち、

$$p(u; \theta) = \int_{-\infty}^{\infty} f^\theta(u, y) dy \quad (9.4)$$

である。ここで、 $p(u; \theta)$ は、次のように一次元のガウス型重み関数と M 次一変数多項式との積として表現される。

$$p(u; \theta) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} \sum_{k=0}^M p_k^\theta u^k \quad (9.5)$$

投影像の観測により、 $p(u; \theta)$ の係数 $\{p_k^\theta : k=0, \dots, M\}$ の値が、各投影角 $\{\theta_i : i=1, \dots, N\}$ に対して測定される。断層像の再構成とは、この係数 $\{p_k^{\theta_i} : k=0, \dots, M, i=1, \dots, N\}$ の観測値を用いて、

$f(x, y)$ の係数 $\{f_{mk} : m=0, \dots, M, k=0, \dots, m\}$ を推定することである。 $\{f_{mk}\}$ と $\{p_k^\theta\}$ との間の線形関係を明らかにすることにより、投影像の線形観測システムが求められる。

ここで、各係数を用いて、次のベクトルを定義する。

$$f_i \equiv \begin{bmatrix} f_{i0} \\ f_{i1} \\ \vdots \\ f_{iM} \end{bmatrix}, \quad f \equiv \begin{bmatrix} f_0 \\ f_1 \\ \vdots \\ f_M \end{bmatrix},$$

$$f_i^\theta \equiv \begin{bmatrix} f_{i0}^\theta \\ f_{i1}^\theta \\ \vdots \\ f_{iM}^\theta \end{bmatrix}, \quad f^\theta \equiv \begin{bmatrix} f_0^\theta \\ f_1^\theta \\ \vdots \\ f_M^\theta \end{bmatrix},$$

$$P^\theta \equiv \begin{bmatrix} p_0^\theta \\ p_1^\theta \\ \vdots \\ p_M^\theta \end{bmatrix}$$

(9.6)

はじめに、ベクトル f とベクトル f^θ との間の関係を明らかにする。
この関係は、 $\varphi_{mk}(x, y)$ と $\varphi_{mk}(x', y')$ との間の線形関係に対応している。
直交座標系 (x, y) と (x', y') との間には

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} a & b \\ -b & a \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \quad (9.7)$$

という関係がある。ここで

$$a = \cos \theta, \quad b = \sin \theta \quad (9.8)$$

である。これより、 $\varphi_{mk}(x', y')$ は次のようになる。

$$\begin{aligned} \varphi_{mk}(x', y') &= \frac{(ax+by)^k (-bx+ay)^{m-k}}{\sqrt{k!(m-k)!}} \\ &= \sum_{h=0}^k \sum_{g=0}^{m-k} (-1)^{m-k-g} \frac{\sqrt{k!(m-k)!}}{h!g!(k-h)!(m-k-g)!} \\ &\quad \times a^{k-h+g} b^{m-k+h-g} x^{m-h-g} y^{h+g} \quad (9.9) \end{aligned}$$

いま、 $k' = m - h - g$ とおき、二重総和の順序を次のように変更する。

$$\sum_{h=0}^k \sum_{g=0}^{m-k} \Rightarrow \sum_h \sum_{k'=0}^m \quad (9.10)$$

但し、 h は、(9.10) の分母の階乗の中が負にならない範囲で動かすものとする。このとき、

$$\begin{aligned} \varphi_{mk}(x', y') &= \sum_h \sum_{k'=0}^m (-1)^{k'-k+h} \frac{\sqrt{k!(m-k)!}}{h!(m-k'-h)!(k-h)!(k'-k+h)!} \\ &\quad \times a^{m+k-k'-2h} b^{k'-k+2h} x^{k'} y^{m-k'} \quad (9.11) \end{aligned}$$

これより、次式が求まる。

$$\varphi_{mk}(x', y') = \sum_{k'=0}^m a_{mkk'}^{\theta} \varphi_{mk'}(x, y) \quad (9.12)$$

但し、

$$a_{mkk'}^{\theta} \equiv (-1)^{k'-k} \sum_h (-1)^h \frac{\sqrt{k!(m-k)!k'!(m-k')!}}{h!(m-k'-h)!(k-h)!(k'-k+h)!} \\ \times (\cos \theta)^{m+k-k'-2h} (\sin \theta)^{k'-k+2h} \quad (9.13)$$

である。(9.12)より、ベクトル φ と φ^{θ} との間の線形関係が次のようになる。

$$\varphi^{\theta} = A(\theta) \cdot \varphi \quad (9.14)$$

ここで、行列 $A(\theta)$ は、回転角 θ をパラメータにもつ行列であり、次のように各次数ごとの直和として与えられる。

$$A(\theta) \equiv A_0(\theta) + A_1(\theta) + \dots + A_M(\theta)^{(\text{注})}$$

$$= \begin{bmatrix} A_0(\theta) & & & 0 \\ & A_1(\theta) & & \\ & & \ddots & \\ 0 & & & A_M(\theta) \end{bmatrix} \quad (9.15)$$

但し、行列 $A_m(\theta)$ は次式で与えられる。

$$A_m(\theta) = [a_{mkk'}^{\theta}] \quad (9.16)$$

次にベクトル φ^{θ} とベクトル p^{θ} との間の関係を明らかにする。

(注) 本章では、混同の恐れのない限り、各次数ごとの直和を単に行列の和の形で表現する。

この関係は、(9.4) の積分を行なうことにより求められる。

$$\begin{aligned}
 p(u; \theta) &= \int_{-\infty}^{\infty} f^{\theta}(u, y) dy \\
 &= \int_{-\infty}^{\infty} \frac{1}{2\pi} e^{-\frac{u^2+y^2}{2}} \sum_{m=0}^M \sum_{k=0}^m f_{mk}^{\theta} \varphi_{mk}(u, y) dy \\
 &= \frac{1}{2\pi} e^{-\frac{u^2}{2}} \sum_{m=0}^M \sum_{k=0}^m f_{mk}^{\theta} u^k \left(\int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} \frac{y^{m-k}}{\sqrt{k!(m-k)!}} dy \right)
 \end{aligned}
 \tag{9.17}$$

一般に、

$$\int_{-\infty}^{\infty} x^m e^{-\frac{x^2}{2}} dx = \begin{cases} (m-1)!! \sqrt{2\pi} & m: \text{偶数} \\ 0 & m: \text{奇数} \end{cases}
 \tag{9.18}$$

が成り立つことに注意すると、(9.17) は更に次のようになる。

$$\begin{aligned}
 p(u; \theta) &= \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} \sum_{m=0}^M \sum_{k=0}^m b_{mk} f_{mk}^{\theta} u^k \\
 &= \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} \sum_{k=0}^M \left(\sum_{m=k}^M b_{mk} f_{mk}^{\theta} \right) u^k
 \end{aligned}
 \tag{9.19}$$

そこで、 b_{mk} は次式で与えられる。

$$b_{mk} = \begin{cases} \frac{(2\ell-1)!!}{\sqrt{k!(m-k)!}} & : m-k=2\ell \text{ の場合} \\ 0 & : m-k=2\ell-1 \text{ の場合} \end{cases} \quad (9.20)$$

(9.5) と (9.19) とから、次式が求まる。

$$p_k^\theta = \sum_{m=k}^M b_{mk} f_{m,k}^\theta \quad (9.21)$$

これより、 f^θ と p^θ との間の関係は次のようになる。

$$p^\theta = B f^\theta \quad (9.22)$$

但し、行列 B は次式で与えられる。

$$B \equiv \begin{bmatrix} \boxed{\begin{matrix} b_{00} & b_{10} & 0 \\ & 0 & b_{11} \end{matrix}} & \cdots & \boxed{\begin{matrix} b_{M0} & & 0 \\ & b_{M1} & \\ 0 & & \ddots \\ & & & b_{MM} \end{matrix}} \\ 0 & & \end{bmatrix} \quad (9.23)$$

(9.14) の行列 $A(\theta)$ と (9.22) の行列 B を用いて、投影像の観測システムが構成される。

N 個の投影角 $\theta_i : i=1, \dots, N$ に対して、投影像 $p(u; \theta_i)$ が観測される。実際には、係数 $p_k^{\theta_i} : k=0, \dots, M$ が観測されるのであるが、各観測値には、観測誤差が加わる。このことを考慮して、投影像の観測システムを構成すると次のようになる。

[投影像の観測システム]

$$\begin{aligned} z_i &= p^{\theta_i} + w_i \\ &= BA(\theta_i)\varphi + w_i, \quad i=1, \dots, N \end{aligned} \quad (9.24)$$

z_i の成分 z_{ik} が、係数 $p_k^{\theta_i}$ の観測値である。 z_{ik} は、 $p_k^{\theta_i}$ と観測誤差 v_{ik} との和として表わされる。

この観測システムにおいて、断層像の係数ベクトル φ 及び観測誤差ベクトル w_i が共に互いに独立な確率ベクトルであり、その平均及び共分散が次のように与えられているとする。

$$\begin{aligned} E(\varphi) &= \varphi_0, \quad \text{Cov}(\varphi - \varphi_0) = S, \\ E(w_i) &= 0, \quad \text{Cov}(w_i) = \sigma^2 R_i, \\ i &= 1, \dots, N \end{aligned} \quad (9.25)$$

このとき、(9.24) の観測システムは、統計的線形観測システムとなり、第4章、4.3 で扱った線形観測システムの逐次観測過程 (4.39) とみることが出来る。

9.3 断層像の再構成システム

第4章、4.3、定理 4.3 の最小分散推定の逐次推定システムを適用することにより、断層像の再構成システムが求められる。投影角 θ_i の投影像の観測段階における断層像の係数ベクトル φ の推定量を $\hat{\varphi}_{(i)}$ とおく。定理 4.3 の行

列 A_i に (9.24) の行列 $BA(\theta_i)$ を代入すると、次のように断層像の再構成システムが求まる。

$$\left\{ \begin{aligned} \hat{f}_i &= \hat{f}_{i-1} + (BA(\theta_i))_{S_{i-1}R_i}^\sigma (z_i - BA(\theta_i) \hat{f}_{i-1}) \end{aligned} \right. \quad (9.26)$$

$$S_i = \{I - (BA(\theta_i))_{S_{i-1}R_i}^\sigma BA(\theta_i)\} S_{i-1} \quad (9.27)$$

$$, \quad i=1, \dots, N$$

但し、 $\hat{f}_{(0)} = f_0$, $S_{(0)} = S$ である。

実際にこの逐次推定システムを忠実に実行するには、係数ベクトル f の逐次計算の他に、(9.27) に示した共分散行列 S_i の逐次計算を同時に行なわなければならない。一般に断層像を表現するには、多項式の次数 M を十分大きく選ばねばならない。共分散行列 S は、 $(M+1)(M+2)/2$ 次の正方行列であるから、 M が大きな値であると、行列 S は、非常に高い次元の行列となる。この行列を逐次に計算するのは容易ではない。そこで、本節では、若干の仮定と近似を行なうことにより、共分散行列 S の逐次計算を必要としないうちに、再構成システムを求める。

はじめに、断層像及び観測誤差の統計的性質に関して、次の二つの仮定を設ける。

$$(a) \quad S = s_0 I + s_1 I_1 + \dots + s_M I_M \quad (9.28)$$

$$(b) \quad R_i = R, \quad i=1, \dots, N \quad (9.29)$$

共分散行列 S が、(9.28) の形であれば、次式が成り立つ。

$$\begin{aligned}
& A(\theta) S A(\theta)^T \\
&= \sum_{k=0}^M s_k A_k(\theta) I_k A_k(\theta)^T \\
&= \sum_{k=0}^M s_k I_k \\
&= S
\end{aligned} \tag{9.30}$$

このことは、対象とする断層像の統計的性質が、どの角度の直交座標系からみても同じであることを示している。統計的性質として、扱おう対象が原点を中心に等方性をもつ場合には、仮定(a)は満足される。

仮定(b)は、どの方向の投影に対しても、観測に加わる観測誤差が同じ統計的性質をもつことを意味している。

さらに、次の近似式を導入する。

$$(c) \quad B^T R^{-1} B \doteq \sum_{k=0}^M b_k I_k \tag{9.31}$$

ここで $b_k: k=0, \dots, M$ の値は、行列 R に依存して決められる。このような形の近似がどの程度、有効なものであるか、その評価は容易でない。但し、特に行列 R が対角行列のような場合には、(9.31)の近似がかなり有効であることが確かめられている。

仮定(a), (b) 及び近似(c) が成り立つと、共分散行列 $S(i)$ の逐次推定は、次のように簡易化される。

$$S(i) = \sum_{k=0}^M s(i, k) I_k, \quad i=0, 1, \dots, N \tag{9.32}$$

但し、 $s(i, k)$ は次式で与えられる。

$$S(i, k) = \frac{S_k}{1 + \frac{S_k b_k}{\sigma^2} i} \quad (9.33)$$

証明は、 i に関する帰納法により行なうことができる。 $i=0$ のときは、明らかに成り立つ。 $i=l+1$ においては、(9.33) が成り立つと仮定する。このとき、 $S(l+1)$ は、次のようになる。

$$\begin{aligned} S(l+1) &= (I - (BA(\theta_{l+1}))_{S(l), R}^{\sigma} BA(\theta_{l+1})) S(l) \\ &= (S(l)^{-1} + \frac{1}{\sigma^2} A(\theta_{l+1})^T B^T R^{-1} B A(\theta_{l+1}))^{-1} \\ &= \left\{ S(l)^{-1} + \frac{1}{\sigma^2} A(\theta_{l+1})^T \left(\sum_{k=0}^M b_k I_k \right) A(\theta_{l+1}) \right\}^{-1} \\ &= \left\{ \sum_{k=0}^M \left(\frac{1}{S(l, k)} + \frac{b_k}{\sigma^2} \right) I_k \right\}^{-1} \quad (9.34) \end{aligned}$$

==で、

$$\begin{aligned} &\frac{1}{S(l, k)} + \frac{b_k}{\sigma^2} \\ &= \frac{1 + \frac{S_k b_k}{\sigma^2} l}{S_k} + \frac{b_k}{\sigma^2} \\ &= \frac{1 + \frac{S_k b_k}{\sigma^2} (l+1)}{S_k} \\ &= \frac{1}{S(l+1, k)} \quad (9.35) \end{aligned}$$

となるから $i = l+1$ においても (9.32) が成り立つ。

以上のことから、帰納法により、(9.32) が証明された。

さらに、この結果を用いることにより、行列 $BA(\theta_i)$ の逆行列 $(BA(\theta_i))_{S(i-1)}^{-1}$ は次のようになる。

$$\begin{aligned} & (BA(\theta_i))_{S(i-1)}^{-1} \\ &= (A(\theta_i)^T B^T R^{-1} BA(\theta_i) + \sigma^2 S_{S(i-1)}^{-1})^{-1} A(\theta_i)^T B^T R^{-1} \\ &= \left(\sum_{k=0}^M b_k I_k + \sigma^2 \sum_{k=0}^M \frac{1}{S(i-1, k)} I_k \right)^{-1} A(\theta_i)^T B^T R^{-1} \\ &= \left(\sum_{k=0}^M \frac{S(i-1, k)}{\sigma^2 + b_k S(i-1, k)} I_k \right) A(\theta_i)^T B^T R^{-1} \quad (9.36) \end{aligned}$$

以上のことから、断層像 $f(x, y)$ の再構成システムは、次のようになる。

[断層像の再構成システム]

$$\begin{aligned} \hat{f}_{(i)} &= \hat{f}_{(i-1)} + Q_{(i)} A(\theta_i)^T B^T R^{-1} (z_i - BA(\theta_i) \hat{f}_{(i-1)}) \\ & \quad , i = 1, \dots, N \quad (9.37) \end{aligned}$$

但し、 $\hat{f}_{(0)} = f_0$ であり、行列 $Q_{(i)}$ は、(9.33) の $S(i, k)$ を用いて、次式で与えられる。

$$Q_{(i)} = \sum_{k=0}^M \frac{S(i-1, k)}{\sigma^2 + b_k S(i-1, k)} I_k \quad (9.38)$$

仮定 (a), (b) ならびに近似 (c) を与えることにより, 断層像の再構成システムは, (9.37) の形に単純化される。この再構成システムでは, 係数ベクトル $\mathbf{g}$ の逐次推定だけが行われ, その共分散行列 $\mathbf{S}$ の逐次計算は不要になっている。ここまでのシステムを簡易化すると, 形式的には, いわゆる SIRT と呼ばれる同時逐次近似法と類似したものとなる。

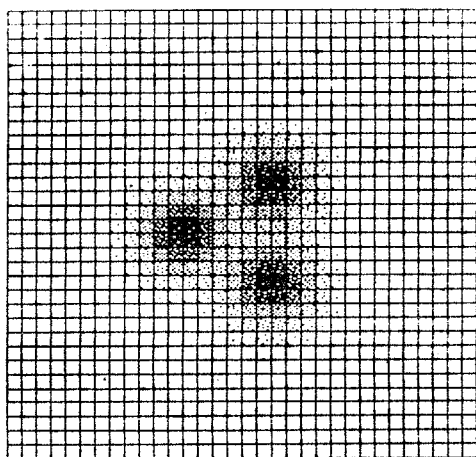
(9.37) の右辺, 第二項が, 第 i 段階の観測ベクトル $\mathbf{g}_i$ による修正項である。この修正項における $\mathbf{A}(\mathbf{D}_i)^T \mathbf{B}^T \mathbf{R}^{-1}$ という変換は, ちょうど SIRT における逆投影 (back projection) と呼ばれる操作に対応したものである。行列 $\mathbf{Q}(i)$ は, SIRT におけるダンピング係数に対応した効果をもつものである。

ダンピング係数は, 逐次近似法において, 近似値の収束性の安定化のために修正項の効果を弱める働きをもつものである。従来の SIRT においては, このダンピング係数の与え方が問題となっていた。本節では, 最小分散推定の逐次推定の立場から, 再構成システムの簡易化を行うことにより, 最適なダンピング係数を理論的に導出できたと考えられる。

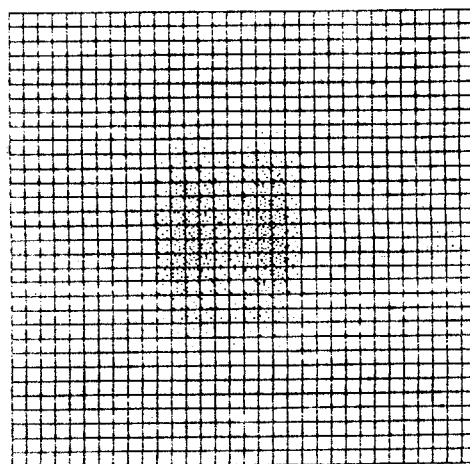
(9.33) より明らかのように, $\sigma(i, k)$ は, 観測の数 i の増加とともに, 次第に小さな値となる。これより, (9.38) の行列 $\mathbf{Q}(i)$ の対角成分も i の増加とともに小さくなる。このことは, ダンピング効果が観測の数 i が増えるに従って大きくなることを示している。その変化の程度は, 観測誤差の大きさに, 即ち, σ の値に依存している。(9.33), (9.38) より明らかのように, σ が小さいほど, 行列 $\mathbf{Q}(i)$ の対角成分は, より速く零に近づく。このことは, σ の値が小さい程度, 再構成される断層像の収束が速いことを意味している。

(9.37) で与えた, 投影像からの断層像の再構成システムの簡単な数値実験の結果を, 図 9.1, (a) ~ (f) に示す。

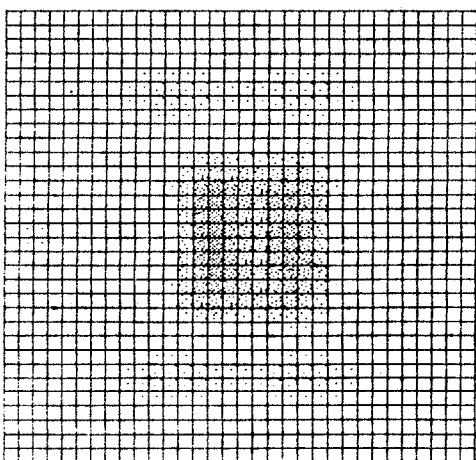
図 9.1, (a) が, 原断層像である。ガウス型の関数を三つ集めて作った像である。断層像の逐次推定の結果が 図 9.1, (b) ~ (f) である。逐次推定の



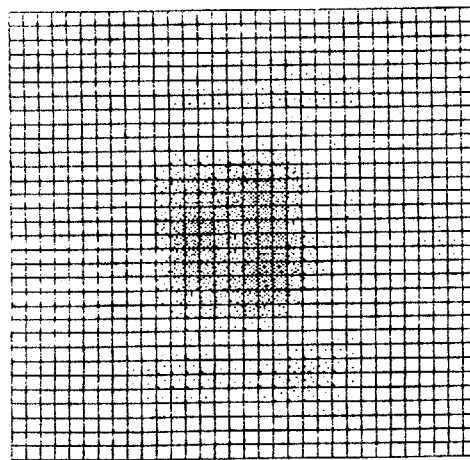
(a) 原断層像



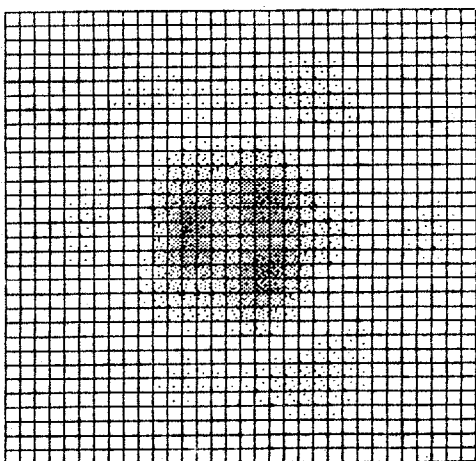
(b) $i=1$



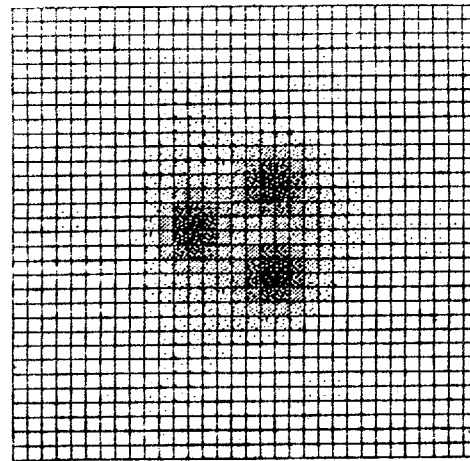
(c) $i=2$



(d) $i=3$



(e) $i=4$



(f) $i=20$

図 9.1 投影像からの断層像再構成 (i : 投影像の数)

初期段階 ($i=1\sim 4$) において、急速に原断層像に近い断層像に逐次相定されていく様子がよく表われている。図 9.1, (十) は、観測の数 $i=20$ における相定断層像である。ほぼ原断層像に等しいものが得られている。この段階になると、ダンピング効果がかなり大きなものとなり、新しい観測による断層像の修正はほとんど行なわれなくなる。

なお、この数値実験では、多項式の次数 $M=8$ とし、其分散行列 S と R はともに単位行列とした。また、係数ベクトル $\mathbf{f}$ の初期値 $\mathbf{f}_{(0)}$ として、零ベクトルを用いた。

9.4. 投影角の最適化

本節では、投影像の観測システムにおける投影角の最適化の問題を考察する。この問題は、第7章で論じた線形観測システムの最適化問題の具体例となる。

9.2 で求めた投影像の観測システム (9.24) は、次のように N 個の投影像の観測システムをまとめることによる。一つの線形観測システムとして表わすことができる。

$$\mathbf{z} = A \mathbf{f} + \mathbf{v} \quad (9.39)$$

ここで、ベクトル $\mathbf{z}$, $\mathbf{v}$ および行列 A は次式で与えられる。

$$\mathbf{z} \equiv \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_N \end{bmatrix}, \quad \mathbf{v} \equiv \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_N \end{bmatrix} \quad (9.40)$$

$$A \equiv \begin{bmatrix} BA(\theta_1) \\ BA(\theta_2) \\ \vdots \\ BA(\theta_N) \end{bmatrix} \quad (9.41)$$

行列 A は、 $\{\theta_1, \theta_2, \dots, \theta_N\}$ をパラメータにもつ行列である。いま、このパラメータ $\{\theta_i\}$ を動かしてできる行列 A の集合 $\mathcal{M}$ とおく。最適化の基礎として、断層像の真値 f とその最小分散推定量 $\hat{f}$ との間の推定誤差の二乗平均を採用すると、投影角の最適化問題は、次式

$$J(A_0) = \min_{A \in \mathcal{M}} J(A) \quad (9.42)$$

を満たす行列 A_0 を $\mathcal{M}$ の中から求めることである。但し、 $J(A)$ は、

$$\begin{aligned} J(A) &\equiv E \| f - \hat{f} \|^2 \\ &= \text{tr} (I - A \Sigma_R A^T) S \end{aligned} \quad (9.43)$$

である。

この問題を一般の場合に於いて解くことは容易ではない。本節では、適宜に幾つかの仮定を設ける。この最適化問題を解くことにする。

断層像 $f(x, y)$ の大域的な特徴を最も良く表わすのは、二変数多項式の一次までの項であるから、 $M=1$ とする。このとき、行列 $BA(\theta)$ は次のような 2×3 行列となる。

$$BA(\theta) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\sin \theta & \cos \theta \end{bmatrix} \quad (9.44)$$

これより、行列 A は次のようになる。

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\rho \sin \theta_1 & \cos \theta_1 \\ 1 & 0 & 0 \\ 0 & -\rho \sin \theta_2 & \cos \theta_2 \\ \vdots & \vdots & \vdots \\ 1 & 0 & 0 \\ 0 & -\rho \sin \theta_N & \cos \theta_N \end{bmatrix} \quad (9.45)$$

また、係数ベクトル μ と観測誤差ベクトル V の共分散 S 及び R は次のように与えられるものとする。

$$S = \begin{bmatrix} S_0 & 0 & 0 \\ 0 & S_1 & 0 \\ 0 & 0 & S_1 \end{bmatrix} \quad (9.46)$$

$$R = \begin{bmatrix} R_1 & & 0 \\ & R_2 & \\ 0 & & R_N \end{bmatrix} \quad (9.47)$$

但し、 R_i は、次式で与えられる。

$$R_i = \begin{bmatrix} r_0(\theta_i) & 0 \\ 0 & r_1(\theta_i) \end{bmatrix} \quad (9.48)$$

(9.47) と (9.48) は、投影像に加わる観測誤差は、各投影に対して互に

独立であり、 γ の分散が投影角 θ の関数として与えられていることを示している。

以上の仮定の下に、最適化の評価規準である断層像の相定誤差の二乗平均、 $J(A)$ を求める

$$\begin{aligned}
 J(A) &= \text{tr} (I - A S_R^{-1} A) S \\
 &= \text{tr} \{ S - S A^T (A S A^T + \sigma^2 R)^{-1} A S \} \\
 &= \text{tr} \left(\frac{1}{\sigma^2} A^T R^{-1} A + S^{-1} \right)^{-1} \quad (9.49)
 \end{aligned}$$

ここで、はじめに行列 $A^T R^{-1} A$ を求める。

$$\begin{aligned}
 A^T R^{-1} A &= \sum_{i=1}^N A(\theta_i)^T B^T R_i^{-1} B A(\theta_i) \\
 &= \sum_{i=1}^N \begin{bmatrix} \frac{1}{\gamma_0(\theta_i)} & 0 & 0 \\ 0 & \frac{\sin^2 \theta_i}{\gamma_1(\theta_i)} & \frac{-\sin \theta_i \cos \theta_i}{\gamma_1(\theta_i)} \\ 0 & \frac{-\sin \theta_i \cos \theta_i}{\gamma_1(\theta_i)} & \frac{\cos^2 \theta_i}{\gamma_1(\theta_i)} \end{bmatrix} \quad (9.50)
 \end{aligned}$$

すなわち、 $\frac{1}{\sigma^2} A^T R^{-1} A + S^{-1}$ は次のようになる。

$$\frac{1}{\sigma^2} A^T R^{-1} A + S^{-1} = \begin{bmatrix} a & 0 & 0 \\ 0 & b & d \\ 0 & d & c \end{bmatrix} \quad (9.51)$$

ここで、 a, b, c および d は、次式で与えられる。

$$a \equiv \frac{1}{S_0} + \frac{1}{\sigma^2} \sum_{i=1}^N \frac{1}{r_0(\theta_i)} \quad (9.52)$$

$$b \equiv \frac{1}{S_1} + \frac{1}{\sigma^2} \sum_{i=1}^N \frac{\sin^2 \theta_i}{r_1(\theta_i)} \quad (9.53)$$

$$c \equiv \frac{1}{S_1} + \frac{1}{\sigma^2} \sum_{i=1}^N \frac{\cos^2 \theta_i}{r_1(\theta_i)} \quad (9.54)$$

$$d \equiv -\frac{1}{\sigma^2} \sum_{i=1}^N \frac{\cos \theta_i \sin \theta_i}{r_1(\theta_i)} \quad (9.55)$$

$J(A)$ は、(9.51) により次式で与えられる。

$$\begin{aligned} J(A) &= \det \begin{bmatrix} a & 0 & 0 \\ 0 & b & d \\ 0 & d & c \end{bmatrix}^{-1} \\ &= \frac{1}{a} + \frac{b+c}{bc-d^2} \end{aligned} \quad (9.56)$$

すなわち、 $J(A)$ は次のようになる。

$$J(A) = \frac{1}{\frac{1}{s_0} + \frac{1}{\sigma^2} m_0} + \frac{2 \left(\frac{1}{s_1} + \frac{1}{2\sigma^2} m_1 \right)}{\left(\frac{1}{s_1} + \frac{1}{2\sigma^2} m_1 \right)^2 - \frac{1}{4\sigma^2} m_2} \quad (9.57)$$

但し、 m_0, m_1, m_2 は次式で与えられる。

$$m_0 \equiv \sum_{i=1}^N \frac{1}{r_0(\theta_i)} \quad (9.58)$$

$$m_1 \equiv \sum_{i=1}^N \frac{1}{r_1(\theta_i)} \quad (9.59)$$

$$m_2 \equiv \left(\sum_{i=1}^N \frac{\cos 2\theta_i}{r_1(\theta_i)} \right)^2 + \left(\sum_{i=1}^N \frac{\sin 2\theta_i}{r_1(\theta_i)} \right)^2 \quad (9.60)$$

この $J(A)$ を最小にする $\theta_i: i=1, \dots, N$ が求めるものである。(9.48) の $r_j(\theta)$, $j=1, 2$ が一般の θ の関数の場合に、 $J(A)$ を最小にする θ_i を求めるのは容易でない。はじめに次のような特別な場合を考えてみる。

(a) $r_j(\theta) = r_j$ ($j=0, 1$) の場合

この場合、観測誤差の統計的性質が投影角に依存せずに一定であり、 m_0, m_1 及び m_2 は次のようになる。

$$m_0 = \frac{N}{r_0}, \quad m_1 = \frac{N}{r_1} \quad (9.61)$$

$$m_2 = \frac{1}{Y_1^2} \left\{ \left(\sum_{i=1}^N \cos 2\theta_i \right)^2 + \left(\sum_{i=1}^N \sin 2\theta_i \right)^2 \right\} \quad (9.62)$$

m_0 と m_1 はともに θ に関係のない定数となる。これより、 $J(A)$ を最小化するには、(9.57)の第二項の分母を最大化すればよい。 $m_2 \geq 0$ に注意すると、 $m_2 = 0$ のとき、 $J(A)$ は最小値をとることが分かる。このとき、 $\{\theta_i\}$ は、次式で与えられる。

$$\theta_i = \theta_0 + (i-1) \frac{\pi}{N}, \quad i=1, \dots, N \quad (9.63)$$

但し、 θ_0 は、 $0 \leq \theta_0 < \frac{\pi}{N}$ を満たす任意の定数である。

このように、投影像の観測システムが等方性の場合、即ちどの方向の投影に対しても、加わる観測誤差の影響が同程度であるような場合には、投影角を π/N ごとに等間隔に選ぶことにより、最適な観測システムが得られる。

医用診断に用いられるCTの観測装置の多くは、その投影角が等間隔に選ばれている。これは、投影像の観測システムが等方性の場合、最適な観測システムを実現しているとみることができる。

(b) 一般の $Y_j(\theta)$, ($j=0, 1$) の場合

この場合は、 $J(A)$ の最小化を厳密に解くことは容易でない。しかし、 θ の関数 $Y_j(\theta)$, $j=0, 1$ の変動がそれほど大きくなり場合には、近似解を求めることができる。

(9.57) における m_0, m_1 および m_2 はすべて投影角 $\{\theta_i\}$ に依存している。ここで、 m_0 と m_1 は、それぞれ $Y_0(\theta_i)$, $Y_1(\theta_i)$ の調和平均に対応しているので、 $Y_0(\theta)$, $Y_1(\theta)$ がともに θ の関数として変動の大きくなるものであるならば、 $\{\theta_i\}$ の選び方に対して m_0, m_1 はあまり大きな変動を示す

ことはなり。これに対して、 m_2 は、 $\{\theta_i\}$ の並び方により、その値は大きく変化する。このことは、 $J(A)$ の最小化において、 m_2 の項が最も大きな影響を与えることを示している。そこで、ここでは、 $J(A)$ の最小化問題を m_2 の最大化問題に置き換える。

(9.60) より、 $m_2 \geq 0$ であるから、 $m_2 = 0$ とする $\{\theta_i\}$ が得られれば、それが求める最適解である。 $m_2 = 0$ とする $\{\theta_i\}$ は、投影角の数 N が十分に大きいときには、次のようにして近似的に求めることができる。

いま、一般性を失うことなく、 $\{\theta_i: i=1, \dots, N\}$ は

$$0 \leq \theta_1 < \theta_2 < \dots < \theta_N < \pi \quad (9.64)$$

を満たすものとする。ここで、次のような単調関数 $\theta(x)$ を導入する。

$$\begin{cases} \theta(0) = 0, & \theta(1) = \pi \\ \theta(\alpha + \frac{i-1}{N}) = \theta_i \end{cases} \quad (9.65)$$

但し、 α は、 $0 \leq \alpha < \frac{1}{N}$ を満たす定数である。

一般に、 N が十分に大きい場合には、任意の関数 $f(\theta)$ に対して、次の近似式が成り立つ。

$$\begin{aligned} \sum_{i=1}^N f(\theta_i) &\doteq N \int_0^1 f(\theta(x)) dx \\ &= N \int_0^\pi f(\theta) \cdot \frac{dx}{d\theta} \cdot d\theta \end{aligned} \quad (9.66)$$

これより、 m_2 は次のように近似することができる。

$$m_2$$

$$\doteq N^2 \left\{ \left(\int_0^\pi \frac{\cos 2\theta}{r_1(\theta)} \cdot \frac{dx}{d\theta} \cdot d\theta \right)^2 + \left(\int_0^\pi \frac{\sin 2\theta}{r_1(\theta)} \cdot \frac{dx}{d\theta} \cdot d\theta \right)^2 \right\} \quad (9.67)$$

いま、

$$\int_0^\pi \cos 2\theta \, d\theta = \int_0^\pi \sin 2\theta \, d\theta = 0 \quad (9.68)$$

に注意すると、 $m_2 \doteq 0$ とするには、次式が成り立たなければならない。

$$r_1(\theta) \frac{d\theta}{dx} = k \quad \text{--- 一定} \quad (9.69)$$

両辺を積分すると、

$$\int_0^\theta r_1(\theta) \, d\theta = \int_0^x r_1(\theta) \frac{d\theta}{dx} \, dx = kx \quad (9.70)$$

となる。そこで、 $\theta = \pi$ とおくことにより、 k が定まる。即ち、

$$k = \int_0^\pi r_1(\theta) \cdot d\theta \quad (9.71)$$

(9.70) と (9.71) より、次式が得られる。

$$x = \int_0^\theta r_1(\theta) \, d\theta / \int_0^\pi r_1(\theta) \, d\theta \quad (9.72)$$

(9.72) の右辺は、 θ の関数である。これを $R(\theta)$ とおくと、 $r_1(\theta) > 0$ より、この関数は、 θ に関する単調増加関数である。ゆえに、その逆関数 $R^{-1}(x)$ が

存在する。即ち、

$$\theta(x) = R^{-1}(x) \quad (9.73)$$

となる。(9.65)と(9.73)より、 $\{\theta_i\}$ は次のようにして得られる。

$$\theta_i = R^{-1}\left(\alpha + \frac{i-1}{N}\right), \quad i = 1, \dots, N \quad (9.74)$$

但し、

$$R(\theta) \equiv \int_0^\theta r_1(\theta) d\theta / \int_0^\pi r_1(\theta) d\theta \quad (9.75)$$

である。

以上のことから、一般の $r_0(\theta)$, $r_1(\theta)$ の場合の最適な投影角 $\{\theta_i\}$ は、次のようにして求まる。はじめに、(9.75)に従って、関数 $R(\theta)$ を求める。この関数の逆関数 $R^{-1}(x)$ を求め、(9.74)により、 $\{\theta_i\}$ を定める。なお、 $\{\theta_i\}$ を求める際に必要な関数は、 $r_1(\theta)$ だけで、 $r_0(\theta)$ を必要としないうちに注意する。なお、このときの $J(A)$ は次のようになる。

$$J(A) = \frac{1}{\frac{1}{s_0} + \frac{1}{\sigma^2} m_0} + \frac{2}{\frac{1}{s_1} + \frac{1}{2\sigma^2} m_1} \quad (9.76)$$

簡単な例として、 $r_1(\theta) = r_1(1 + \rho \sin \theta)$ の場合を考えてみる。このとき、

$$\int_0^\theta r_1(\theta) d\theta = r_1(\theta - \omega \theta + 1) \quad (9.77)$$

であるから、 $R(\theta)$ は次のようになる。

$$R(\theta) = \frac{1}{\pi+2} (\theta - \cos \theta + 1) \quad (9.78)$$

これより、最適な投影角 θ_i が、(9.74) に従って求められる。なお、実際に求める場合には、 $R(\theta)$ のグラフを利用することができる。 $N=10$ の場合を図 9.2 に示す。 $R(\theta)$ のグラフの縦軸の区間 $[0, 1]$ を 10 等分して、その値を与える θ が求める投影角となる。

図 9.2 では、投影角 θ_i は、 $\theta = \frac{\pi}{2}$ の近く、即ち、 $Y_i(\theta)$ が高い値をもつ方向に密に選ばれている。これは、統計的性質として、観測誤差の大きな方向ほど、より細かく投影を行なう必要のあることを意味している。

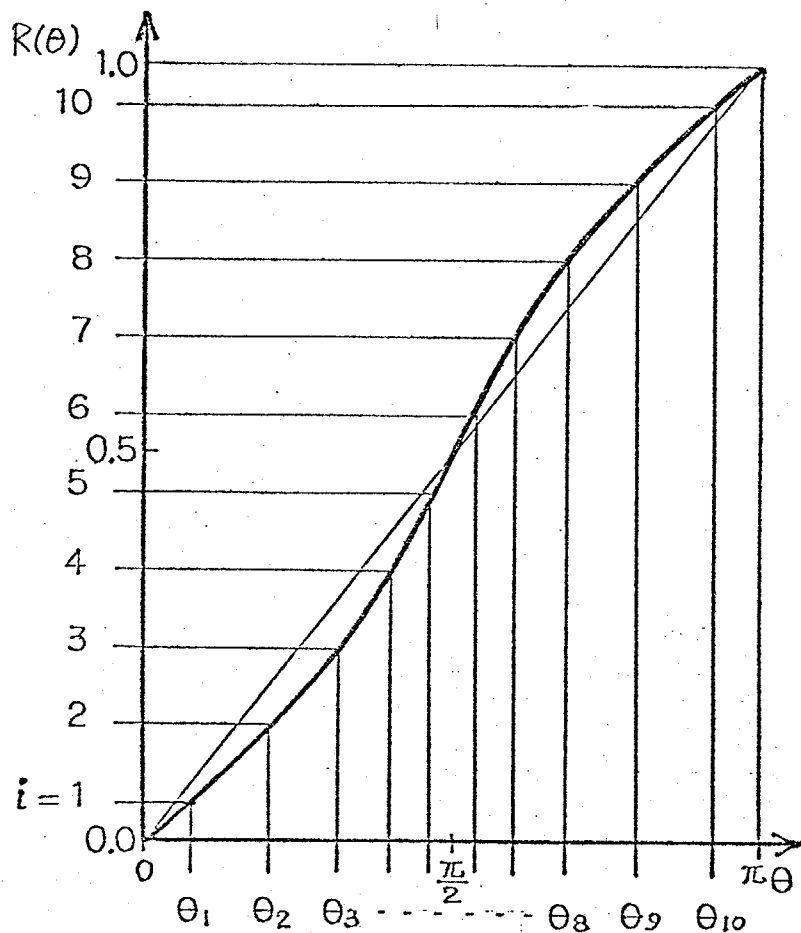


図 9.2. $R(\theta)$ と最適投影角 ($N=10$)

9.5 あとがき

Φ -逆行列の画像処理への応用として、三次元物体の断層像をその投影像から再構成する問題を考えた。本章の断層像の再構成システムの特徴として、次の二点があげられる。

(a) ガウス型重み関数と二変数多項式の積による断層像の表現。

(b) Φ -逆行列を用いた最小分散推定による断層像の逐次推定

ガウス型重み関数と二変数多項式の積により断層像を表現することの第一の利点は、投影像の観測過程を正確に、且つ簡明に記述できることである。また、この表現形式は、比較的低次元で断層像の大域的な構造を捉えるのに適している。逆に断層像の微細な構造を再現する必要がある場合には、この表現形式は必ずしも適したものではない。それは、多項式表現のもつ“くせ”が次数が高くなるに従い顕著になり、画像の表現力が落ちるためである。また、次数が高くなると、多項式の係数ベクトルから、断層像自身を作り出すことに比較的手間のかかることも問題となる。

以上のことから、本章で示した再構成システムは、少ない投影像の観値から断層像の大域的構造を効率良く求める必要がある場面において有効なシステムとなると思われる。

断層像の推定方法として、 Φ -逆行列を用いた最小分散推定を採用したことにより、対象とする断層像ならびに観測誤差の統計的性質を十分に活かすことのできる再構成システムが得られている。逆に、それらの統計的性質に関する事前の情報が手許にない場合には、本章で与えた再構成システムを用いることができない。このことから、扱う断層像がある特定の対象に限定されている、その統計的性質が明らかであり、しかも観測誤差の統計的性質も十分に調べられている

場合に、本章の断層像の再構成システムが有効である。

投影像の観測過程を統計的な線形観測システムとして捉えることの必要性は、9.3の投影角の最適化の問題においても明らかである。対象とする断層像の統計的性質と観測誤差の統計的性質を考慮に入れることにより、はじめにこの最適化問題は定式化できるのである。従来の最小二乗法の立場からは、この投影角の最適化の問題を取り扱うことはできない。

第 10 章 結論

最小分散推定は、最小二乗推定とともに線形観測システムの推定問題において重要な推定方法である。本研究では、最小分散推定のための数学的道具として、 σ -逆行列という概念を提案した。 σ -逆行列は最小分散推定を与える行列であり、このことは、ちょうど Moore-Penrose 逆行列が最小二乗推定を与える行列であることに対応している。

σ -逆行列は、任意の行列（直方行列も含む）に対して必ず一意に存在する行列である。Moore-Penrose 逆行列が、四つの行列方程式の解として定義されるのに対して、 σ -逆行列は、二つの行列方程式の解として定義される。 σ -逆行列と Moore-Penrose 逆行列との間の類似性及び相違性は、それぞれの行列のスペクトル特性の上に集約されている。Moore-Penrose 逆行列のスペクトル特性は、スカラー関数 $1/\lambda$ で与えられるのに対して、 σ -逆行列のスペクトル特性は、スカラー関数 $g(\lambda; \sigma)$ で与えられる。この関数 $g(\lambda; \sigma)$ は、関数 $1/\lambda$ と異なり、連続でしかも有界な関数である。このスペクトル特性の持つ性質が、 σ -逆行列の性質にも大きく反映している。

Moore-Penrose 逆行列は、代数的構造としては、非常に整ったものであるが、実用性の立場からみると必ずしも扱いやすい行列ではなから。それは、Moore-Penrose 逆行列をもとに行列の関数とみたときに、必ずしも連続な関数とはならないことによる。これに対して、 σ -逆行列は、もとの行列の関数としてみたときに連続な関数であり、しかも有界性が保証されている。このように実用性の観点からみると、 σ -逆行列は、Moore-Penrose 逆行列よりもはるかに扱いやすい安定した行列である。この性質は、直接に σ -逆行列を数値計算する場合、あるいは最小分散推定を逐次推定する場合などに、顕著にあらわれてくる。

Φ -逆行列に関連した問題として、線形観測システムの最適化の問題を明らかにした。なお、この問題のための準備として、最適分散性分散パラメータの概念を明らかにした。

Φ -逆行列の具体的応用として、第8章では、曲線のみを扱う問題を扱い、第9章では、画像処理の問題である投影像からの断層像再構成の問題を検討した。これらの応用において、 Φ -逆行列が使用性の高い数学的道具であることが確かめられた。

Φ -逆行列の理論の残された問題として、この概念のヒルベルト空間上の作用素への拡張の問題が残されている。また、非線形作用素への拡張も今後の問題として残されている。

謝 辞

本研究を行うにあたり、終始ご指導、ご鞭撻いただいた東京工業大学飯島泰蔵教授に心から感謝の意を表す。また、貴重な御意見、御討議をいただいた東京工業大学小川英光氏、電子統合研究所磯道義典氏に厚くお礼申し上げます。更に、本研究にご協力いただいた東京工業大学堀江睦彦氏および東京工業大学飯島研究室の諸氏に感謝する。

参考文献

- (1) Akaike, H. : " A new look at the statistical model identification ", IEEE Trans. on Automatic Control, vol. AC-19, No. 6, p. 716 (1974-12)
- (2) Albert, A. : " Regression and the Moore-Penrose pseudoinverse ", Academic Press, Inc. (1972)
- (3) Bellman, R. : " Introduction to Matrix Analysis ", McGraw-Hill (1960)
- (4) Ben-Israel, A., Greville, T. N. E. : " Generalized inverses: theory and applications ", John Wiley & Sons, Inc. (1974)
- (5) Bouillion, T. L., Odell, P. L. : " Generalized inverse matrices ", John Wiley & Sons, Inc. (1971)
- (6) Budinger, T. F., Gullberg, G. T. : " Three dimensional reconstruction in nuclear medicine emission imaging ", IEEE Trans. on Nuclear Science, vol. NS-21, p. 2 (1974)
- (7) Cline, R. E. : " Representations for the generalized inverse of a partitioned matrix ", SIAM J. Appl. Math. vol. 12, p. 588 (1964)
- (8) Erdelyi, I., Ben-Israel, A. : " Extremal solutions of linear equations and generalized inversion between Hilbert spaces ", J. of Math. Analysis and Applications, vol. 39, p. 298 (1972)

- (9) 右田勝久："線形システムの観測と固定", コロナ社 (1976)
- (10) 飯島泰蔵："等分散性特徴パラメータの抽出理論", 電子通信学会オートマトンと自動制御研究会資料 (1966-10)
- (11) 飯島泰蔵："擬似直交性理論とその空間回路網への応用", 昭和48年電子通信学会全国大会, S8-1 (1973)
- (12) Lee, K-H, Peterson: "Optimal linear coding for vector channels", IEEE Trans. on Communications, vol. COM-24, No. 12, p. 1283 (1976-12)
- (13) Luenberger, D. G.: "Optimization by Vector Space Methods", John Wiley & Sons (1969)
- (14) Mirsky, L.: "Results and problems in the theory of doubly-stochastic matrices", Z. Wahrsh. vol. 1, p. 319 (1963)
- (15) 小川英光, 飯島泰蔵："擬似直交性理論", 電子通信学会論文誌(D), vol. 58-D, No. 5, p. 271 (1975-5)
- (16) 尾佐竹倫, 田中英彦："最適通信方式マトリクス変換による伝送路と情報の整合問題—", 電子通信学会論文誌(A), vol. 52-A, No. 12, p. 513 (1969-12)
- (17) Penrose, R.: "A generalized inverse for matrices", Proc. Cambridge Philos. Soc. vol. 51, p. 406 (1955)
- (18) Penrose, R.: "On best approximate solutions of linear matrix equations", Proc. Cambridge Philos. Soc. vol. 52, p. 17 (1956)
- (19) Porter, W. A., Williams, J. P.: "A note on the minimum effort control problem", J. of Math. Analysis and Applications, vol. 13, p. 251 (1966)
- (20) Rao, C. R., Mitra, S. K.: "Generalized inverse of

matrices and its applications", John Wiley & Sons, Inc. (1971) (邦訳. 渋谷, 田辺, 一般逆行列とその応用, 東京図書)

- (21) Rust, B., Burrus, W.R., Schneeburger, C.S. : "A simple algorithm for computing the generalized inverse of a matrix", *Communications of the ACM*, vol. 9, No. 5, p. 381 (1966-5)
- (22) Stewart, G.W. : "On the continuity of the generalized inverse.", *SIAM J. Appl. Math.*, vol. 17, No. 1, p. 33 (1969-1)
- (23) Tikhonov, A.N. : "The stability of algorithms for the solution of degenerate systems of linear algebraic equations", *USSR Computational Math. and Math. Physics*, vol. 6, p. 181 (1965)
- (24) Тихонов, А.Н. : "О решение некорректно поставленных задач и методе регуляризации", *Доклады Академии Наук СССР*, том. 151, No. 3, p. 501 (1963)
- (25) Тихонов, А.Н. : "О регуляризации некорректно поставленных задач", *Доклады Академии Наук СССР*, том. 153, No. 1, p. 49 (1963)
- (26) Wedin, R.A. : "Perturbation theory for pseudo-inverses", *BIT*, vol. 13, p. 217 (1973)

発表論文 - 覧

1. 論文

- (1) 佐藤, 飯島: "空間情報伝送系における比例性ノイズ抑制法", 電子通信学会論文誌(D), vol. 58-D, No. 5 (1975-5)
- (2) M. Sato, T. Iijima: " σ -Inverse of a Matrix", 電子通信学会論文誌(E), (採録済)
- (3) 佐藤, 飯島: "最適分散性特徴パラメータ", 電子通信学会論文誌(D), (採録済)

2. 口頭発表(1): 研究会

- (1) 佐藤, 飯島: "空間情報伝送系におけるノイズ抑制法", 電子通信学会パターン認識と学習研究会, PRL 73-45 (1973-7)
- (2) 佐藤, 飯島: "擬似アダマール変換とその応用について", 電子通信学会パターン認識と学習研究会, PRL 75-39 (1975-10)
- (3) 佐藤, 飯島: " σ -逆行列一定義とその性質", 電子通信学会パターン認識と学習研究会, PRL 76-64 (1977-1)
- (4) 佐藤, 飯島: " σ -逆行列とそのスペクトル特性", 電子通信学会パターン認識と学習研究会, PRL 76-81 (1977-2)
- (5) 佐藤, 飯島: " σ -逆行列-CTにおける投影角の最適化-", 電子通信学会パターン認識と学習研究会, PRL 76-84 (1977-3)
- (6) 佐藤, 飯島: "最適分散性特徴パラメータ", 電子通信学会パターン認識と学習研究会, PRL 77-11 (1977-5)

- (7) 佐藤, 飯島: " σ -逆行列-とその計算法-", 電子通信学会パターン認識と学習研究会, PRL 77-18 (1977-6)
- (8) 佐藤, 飯島: " σ -逆行列の応用-線形システムの最適化-", 電子通信学会パターン認識と学習研究会, PRL 77-26 (1977-7)
- (9) 佐藤, 堀江, 飯島: " 最小分散推定による投影像からの断層像再構成", 電子通信学会パターン認識と学習研究会, PRL 77-42 (1977-10)

3. 口頭発表(2): 大会

- (1) 佐藤, 飯島: " 疑似アダマール変換", 昭和50年度電子通信学会全国大会, No. 960 (1975-3)
- (2) 佐藤, 飯島: " 射影図形からの二次元図形再構成", 昭和51年度電子通信学会統合全国大会, No. 1219 (1976-3)
- (3) 佐藤, 飯島: " 逐次推定による射影図形からの二次元図形再構成", 昭和52年度電子通信学会統合全国大会, No. 1167 (1977-3)
- (4) 佐藤, 飯島: " σ -逆行列", 昭和52年度電子通信学会情報部門全国大会, No. 170 (1977-7)