

### カルマン・フィルタと最小自乗法

NISHIYA, Takanobu / 西谷, 隆亘

---

(出版者 / Publisher)

法政大学工学部

(雑誌名 / Journal or Publication Title)

Bulletin of the Faculty of Engineering, Hosei University / 法政大学工学部  
研究集報

(巻 / Volume)

19

(開始ページ / Start Page)

93

(終了ページ / End Page)

109

(発行年 / Year)

1983-03

(URL)

<https://doi.org/10.15002/00004089>

# カルマン・フィルタと最小自乗法

西谷隆亘\*

## An Approach to the Kalman Estimation Problem by the Least Square Method

Takanobu NISHIYA

### Summary

In the case that the system dynamic equations are

$$\begin{cases} \mathbf{x}(k+1) = \Phi(k+1; k)\mathbf{x}(k) + \Theta(k+1; k)\mathbf{u}(k) + \mathbf{v}(k) \\ \mathbf{y}(k) = \mathbf{D}(k)\mathbf{x}(k) + \mathbf{w}(k) \end{cases} \quad (\text{I})$$

the Kalman's estimation problem is solved plain by the least square method, in terms of the algebraical expression, minimizing the square errors of the state variable  $\mathbf{x}$ .

The solution of the problem is in the form of a cyclic calculation as follows;

$$\begin{cases} \mathbf{A}(k) = \mathbf{P}(k|k-1)\mathbf{D}'(k)[\mathbf{D}(k)\mathbf{P}(k|k-1)\mathbf{D}'(k) + \mathbf{R}(k)]^{-1} \\ \hat{\mathbf{x}}^0(k|k) = \{\mathbf{I} - \mathbf{A}(k)\mathbf{D}(k)\}\hat{\mathbf{x}}^0(k|k-1) + \mathbf{A}(k)\mathbf{y}(k) \\ \hat{\mathbf{x}}^0(k+1|k) = \Phi(k+1; k)\hat{\mathbf{x}}^0(k|k) + \Theta(k+1; k)\bar{\mathbf{u}}(k) \\ \mathbf{P}(k|k) = \{\mathbf{I} - \mathbf{A}(k)\mathbf{D}(k)\}\mathbf{P}(k|k-1) \\ \mathbf{P}(k+1|k) = \Phi(k+1; k)\mathbf{P}(k|k)\Phi'(k+1; k) \\ \quad + \Theta(k+1; k)\mathbf{S}(k)\Theta'(k+1; k) + \mathbf{Q}(k) \end{cases} \quad (\text{II})$$

where  $\mathbf{I}$ : unit matrix

$$\mathbf{Q}(k) = E[\mathbf{v}(k)\mathbf{v}'(k)] \quad \mathbf{S}(k) = E[\bar{\mathbf{u}}(k)\bar{\mathbf{u}}'(k)]$$

$$\mathbf{R}(k) = E[\mathbf{w}(k)\mathbf{w}'(k)]$$

and

$$\begin{cases} \hat{\mathbf{y}}^0(k|k) = \mathbf{D}(k)\hat{\mathbf{x}}^0(k|k) \\ \hat{\mathbf{y}}^0(k+1|k) = \mathbf{D}(k+1)\hat{\mathbf{x}}^0(k+1|k) \end{cases} \quad (\text{III})$$

The method of the identification of the coefficient matrices of the system is also explained by the Kalman filter theory itself.

### Résumé

Lors que les équations dynamiques du système sont comme (I), la solution du problème par R.E. KALMAN (1960) s'exprime dans le cadre de la méthode du moindre carrée. Les résultats sont obtenus en formes cycliques (II) et (III).

D'ailleurs l'identification des matrices coefficients du système s'explique employé la théorie du Kalman filtre.

### § 1. 諸 論

カルマン・フィルタは提唱者の R.E. KALMAN (1969) に因んで名付けられた線形フィルタで

\* 土木工学科

ある<sup>1)</sup>。

難解だといわれるカルマン・フィルタは、ヒルベルト空間での直交射影や条件付期待値の概念を用いて説明されている。しかし、カルマン・フィルタは最小自乗法の原理を時系列データの推定に適用したものという観点からも導けるので、カルマン・フィルタの後述する特徴

- 1) 状態量の導入
- 2) 逐次式表現

が念頭にあれば、通常の最小自乗法と全く同等のレベルあるいは手法で理解が可能である。

本稿では、離散時間系の線形推定におけるカルマン・フィルタを最小自乗法との関連を考慮して、平易に誘導することが試みられている。こうすることにより、カルマン・フィルタの理解が容易になり、更に新しい応用面も開け、正しい認識が得られることと思われる。

## §2. 推定問題とカルマン・フィルタ

ここでいうフィルタとは、時系列データから誤差を除去して、真値を推定する手続きのことであり、計算によりこれが行われる場合は、その計算式のことをフィルタと呼ぶ。

計算式の係数が時間の関数であれば、非定常と呼ばれ、そうでなければ定常と呼ばれる。また、式の型が変数の1次式なら線形、そうでなければ非線形と呼ばれるが、非線形問題の難しさを論ずる際には、通常の最小自乗法では、変数の1次式であることよりは、係数に関して1次式であることの方が重要であることを付言しておく。

データの得られる実際の現象をモデル化して、その機能を式で表現したものまで含めてモデルと呼ばれる。このモデルの設定——定式化及び式の中での係数の具体的な数値を決めること——を同定というが、慣用的には、係数の具体的な数値を決めることのみを「同定」と呼んでいるようである。モデルが同定された後に、そのモデルにより推定が行われる。したがって、推定問題で一番重要なのは、モデルの設定である。その際に、モデルは可観測、可制御などの条件を充足していることが必要で、この点には留意しなければならない<sup>2)</sup>。

カルマン・フィルタでは、線形モデルが同定されているものとして、すなわち、現象が定式化され、その式中の係数が既知のものとして、推定が行われるのに対し、通常の最小自乗法では、モデルの設定がなされていることだけを前提として、「同定」がなされるのである。

推定問題は一般に、ステップ $k$ までの時系列データの観測値が与えられた場合、推定が行われるステップ $j$ に応じて

- i) 予測問題； $j > k$
- ii) フィルタ問題； $j = k$
- iii) 平滑問題； $j < k$

と分類されるが、カルマン・フィルタは形式的には、これら3つの問題の解を包括している。

以下では、今まで述べてきたことを時間離散データの場合について見ることにする。

<記号の説明>

英大文字は、確率変数、英小文字でその実現値 (観測値) を表し、小文字の太字はベクトル (例えば  $\mathbf{x}$ ) 大文字の太字はマトリクス (例えば  $\mathbf{X}$ ) を表すことを原則とする。ただし、確率変数のベクトルは、英大文字の太字で表されることもある。

ベクトルとマトリクスの転置は ' によって表す。

$E[ \ ]$  …… 平均値の作用素

— …… 平均値を表す (例えば  $E[x] = \bar{x}$ )。

^ …… 推定値

^0 …… 最適推定値

~ …… 誤差

§3. 確率変数の推定値と最小自乗法

確率変数の推定は、その生起する確率に基づいてなされ、最も起こりそうな値が推定される。最も起こりそうな値という時、直観的には、2通りが考えられる。すなわち、確率変数の従う確率密度関数が、最大値をとる時の確率変数の値をさす場合と、期待値をさす場合である。特に、正規分布の場合は両者は一致する。

これは、推定値と実現値の差 (誤差と呼ぶ) になるべく小さくなるように推定すること、その時、誤差の絶対値が最小になるように、あるいは、誤差の自乗の平均値 (期待値) が最小になるように考慮されることと、数学的な手続きは同等である。しかし、数学的な取り扱いが容易であることから、一般には後者が推定の条件に採用されることが多い。本論でも、誤差の自乗の期待値が最小となるような制約により議論の展開がなされる。

さて、制約条件を満たす推定値は、最適推定値と呼ばれるが、それは次のようになる。

実現値を  $\mathbf{z}$ 、推定値を  $\hat{\mathbf{z}}$  とすると、制約条件は、

$$E[(\mathbf{z} - \hat{\mathbf{z}})(\mathbf{z} - \hat{\mathbf{z}})'] \rightarrow \text{minimum} \tag{3-1}$$

である。(3-1)式を満足する最適推定値  $\hat{\mathbf{z}}^0$  は

$$\frac{\partial}{\partial \hat{\mathbf{z}}} \{E[(\mathbf{z} - \hat{\mathbf{z}})(\mathbf{z} - \hat{\mathbf{z}})']\} = 0 \text{ により}$$

$$\hat{\mathbf{z}}^0 = E[\mathbf{z}] \tag{3-2}$$

となり、これは直観的なものと符合する。

従って、最小自乗の原理は次のようにいうことができる。

最小自乗の原理；『ある確率変数の誤差の自乗の期待値を最小にする最適推定値は、その確率変数の期待値である』<sup>3)</sup>。

以上に見られるような最小自乗原理による推定は、不偏ではあるが、有効性は持たない。誤差が正規性ならば、最尤推定と一致し、更にシステムパラメータの確率密度も既知ならば、ベイズ推定もできる<sup>4)</sup>。最小自乗推定、最尤推定、ベイズ推定については、LEE (1964) や COX (1964) らの優れた解説があるので、それらを参照されたい<sup>5, 6)</sup>。

### 3.1 最小自乗法

最小自乗法は、一般に、確率変数がパラメータを含む、他の確率変数の関数で表される時、その期待値 (推定式) と実際に得られる値 (実現値) との差の自乗の平均 (分散) が最小になるように、そのパラメータを定め、その時のパラメータの値を使って、確率変数を推定する方法である。

例えば、確率変数  $Y$  が他の確率変数  $X$  の関数で (当然のことながら関数型は既知である)

$$Y=f(X; \alpha, \beta) \quad \text{ただし, } \alpha, \beta \text{ はパラメータ} \quad (3-3)$$

と表される時、の  $Y$  推定値  $\hat{Y}$  は

$$\hat{Y}=E[Y]=E[f(X; \alpha, \beta)]=\bar{f}(X; \alpha, \beta) \quad (3-4)$$

である。  $X$  の実現値が  $x$ 、  $Y$  の実現値を  $y$  とすると、

$$V=E[(y-\hat{Y})^2]=E[(y-\bar{f}(x; \alpha, \beta))^2] \rightarrow \text{minimum} \quad (3-5)$$

となるような  $\alpha, \beta$  を推定する (同定)。然る後、推定された  $\alpha, \beta$  の値を  $\hat{\alpha}=a, \hat{\beta}=b$  とすると、その時、

$$\hat{Y}^0=\bar{f}(x; a, b) \quad (3-6)$$

により推定される値を最適推定値と呼ぶ (推定)。推定誤差  $\varepsilon=Y-\hat{Y}^0$  は  $E[\varepsilon]=0$  であり、相互に独立である。なわち、  $x_i$  に対応する誤差を  $\varepsilon_i$ 、  $x_j$  に対応する誤差を  $\varepsilon_j$  とする時、

$$E[\varepsilon_i \varepsilon_j]=\delta_{ij} \quad \text{ただし, } \delta_{ij} \text{ はクロネッカのデルタ} \quad (3-7)$$

### 3.2 線形な関係式の場合

最も簡単な場合は

$$Y=\alpha X+\beta+\varepsilon \quad (3-8)$$

である。いま(3-8)式にしたがう確率変数の組  $(X, Y)$  の実現値  $(x, y)$  がある時、最小自乗法の意味で係数  $\alpha, \beta$  を推定し、  $Y$  の推定式  $\hat{Y}$  を決定する。(3-8)式はベクトル表示すると

$$Y=\mathbf{a}'\mathbf{X}+\varepsilon \quad \text{ただし, } \mathbf{a}'=(\alpha, \beta), \mathbf{X}'=(X, 1) \quad (3-8)'$$

$$\hat{Y}=E[Y]=\mathbf{a}'E[\mathbf{X}] \quad (3-9)$$

$E[\mathbf{X}]=\mathbf{x}=(x, 1)$  と仮定すると、誤差の自乗平均は

$$V=E[(y-\hat{Y})^2]=E[(y-\mathbf{a}'\mathbf{x})^2]=E[y^2]-2\mathbf{a}'E[\mathbf{x}y]+\mathbf{a}'E[\mathbf{x}\mathbf{x}']\mathbf{a} \quad (3-10)$$

$V$  を最小にする  $\mathbf{a}'$  の推定値  $\hat{\mathbf{a}}'=(a, b)$  は

$$\frac{dV}{d\mathbf{a}}=0; \quad \mathbf{a}'E[\mathbf{x}\mathbf{x}']=E[\mathbf{y}\mathbf{x}'] \quad (3-11)$$

を満足するものとして得られる (付録A参照)。(3-11)式は通常、正規方程式と呼ばれるものである。

$$\therefore \hat{a}' = E[yx'] \{E[xx']\}^{-1} \quad (3-12)$$

したがって、

$$\hat{Y}^0 = \hat{a}'x \quad (3-13)$$

となる。更に一般的な型は

$$Y = AX + \epsilon \quad (3-14)$$

ここに  $Y$ ; 確率変数ベクトル ( $q \times 1$ )

$A$ ; 係数マトリクス ( $q \times p$ )

$X$ ; 確率変数ベクトル ( $p \times 1$ )

$\epsilon$ ; 誤差ベクトル ( $q \times 1$ )

$$E[\epsilon] = 0$$

で表現され、最小自乗解は

$$\hat{A}^0 = E[yx'] \{E[xx']\}^{-1} \quad (3-15)$$

$$\hat{Y}^0 = \hat{A}^0 x \quad (3-16)$$

となる (付録B参照)。

$\hat{X} \neq x$ ,  $\hat{Y} \neq y$  の時の誤差規範は

$$J = (y - \hat{Y})(y - \hat{Y})' + (x - \hat{X})(x - \hat{X})' \quad (3-17)$$

となり、 $E[J]$  を最小にするように係数を決定する問題となるが、本論の主題とは、直接の関係がないので、これについては、別の機会に議論することにする。

#### § 4. 時系列データの線形回帰模型

時系列の場合の線形回帰模型 (動的モデル) を考察する。時系列を扱う場合は、一般には係数は時間の関数である (非定常)。

次のような時間に依存する2つの変数  $Y$ ,  $X$  を結ぶ線形の式を考える。

$$Y(t) = A(t; s)X(s) + w(t) \quad (4-1)$$

ここに  $Y(t)$ ; 時刻  $t$  での変数ベクトル ( $q \times 1$ )

$A(t; s)$ ; 時刻  $s$  での  $X(s)$  と  $Y(t)$  を結ぶ遷移マトリクス ( $q \times p$ )

$X(s)$ ; 時刻  $s$  での変数ベクトル ( $p \times 1$ )

$w(t)$ ; 時刻  $t$  での誤差ベクトル ( $q \times 1$ )

$$E[w(t)] = 0$$

がある時、最小自乗法の意味での  $Y(t)$  の推定値  $\hat{Y}^0(t)$  は

$$\hat{Y}^0(t) = E[Y(t)] = A(t; s)E[X(s)] + E[w(t)] = A(t; s)\bar{X}(s) = A(t; s)\hat{X}(s) \quad (4-2)$$

で与えられる。 $A(t; s)$ が既知ならば、 $\hat{X}(s)$ を知ることにより $\hat{Y}^0(t)$ を計算することができる。しかし、現実の問題では、 $A(t; s)$ は未知で、 $\hat{X}(s)$ も不明な場合が普通である。変数の観測値 $x(s)$ 、 $y(t)$ があれば、 $\hat{X}(s)=x(s)$ と見なし、最小自乗法の意味で、 $A(t; s)$ を推定する。すなわち、

$$J=(y(t)-\hat{Y}(t))(y(t)-\hat{Y}(t))' \quad (4-3)$$

とおき、 $E[J]$ を最小にするような $A(t; s)$ を求めると(付録B参照)

$$\hat{A}^0(t; s)=E[y(t)x'(s)]\{E[x(s)x'(s)]\}^{-1} \quad (4-4)$$

となり、右辺の平均値は時間の関数となるが、一般には、最早や簡単には計算できない。定常かつエルゴード過程であれば、(4-4)式は観測値から計算される。その時、最小自乗法の意味での最適推定値 $\hat{Y}^0(t)$ は

$$\hat{Y}^0(t)=\hat{A}^0(t; s)\hat{X}(s) \quad (4-5)$$

となる。

$X(s)$ が直接には観測されないで、他の観測可能な変数(確率変数でも確定変数でもいずれも可)により

$$X(s)=\theta(s; u)m(u)+v(s)$$

ここに  $\theta(s; u)$ ; 時刻  $u$  での  $m(u)$  と時刻  $s$  での  $X(s)$  を結ぶマトリクス  
( $p \times r$ )

$m(u)$ ; 時刻  $u$  で観測可能なベクトル ( $r \times 1$ )

$v(s)$ ; 時刻  $s$  での誤差ベクトル ( $p \times 1$ )

$$E[v(s)]=0$$

で与えられる場合は

$$\hat{X}^0(s)=\bar{X}(s)=\theta(s; u)E[m(u)]+E[v(s)]=\theta(s; u)\bar{m}(u)=\hat{X}(s) \quad (4-7)$$

となり、 $\theta(s; u)$ 、 $\bar{m}(u)$ が既知でかつ、 $A(t; s)$ も既知である場合、(4-2)式により $\hat{Y}^0(t)$ を得ることができる。 $A(t; s)$ 、 $\theta(s; u)$ が未知の場合でも、観測値( $m(u)$ 、 $y(t)$ )が与えられれば、(4-7)式を(4-2)式へ代入することにより、 $B(t; u)=A(t; s)\theta(s; u)$ とおき、(4-5)式までの議論と全く同様に、 $\hat{Y}^0(t)$ は計算される。しかしながら、興味があるのは、更に一般的な、 $X(s)$ が自己回帰的な場合、すなわち、

$$X(s)=\Phi(s; u)X(u)+\theta(s; u)m(u)+v(s) \quad (4-8)$$

ここに  $\Phi(s; u)$ ; 時刻  $u$  での  $X(u)$  と時刻  $s$  での  $X(s)$  を結ぶ遷移マトリクス ( $p \times p$ )

の時である。この場合は、観測値( $m(u)$ 、 $y(t)$ )があっても、 $X$ に関する情報がない限り、直接代入によっても、係数が未知の時は、解決できない。しかし、係数が既知ならば、逐次推定により $\hat{Y}^0(t)$ は求めることができる。KALMANは次章で述べるように「状態遷移」の概念を用いて、

逐次式による解法を提案したのである。

## § 5. 最小自乗法によるカルマン・フィルタの誘導

### 5.1 カルマンの推定問題

時間離散型の線形動的システムの次のような表現を考える。

$$\begin{cases} \mathbf{x}(j+1) = \Phi(j+1; j)\mathbf{x}(j) + \Theta(j+1; j)\mathbf{u}(j) + \mathbf{v}(j) \end{cases} \quad (5-1)$$

$$\begin{cases} \mathbf{y}(j) = \mathbf{D}(j)\mathbf{x}(j) + \mathbf{w}(j) \end{cases} \quad (5-2)$$

ここに  $\mathbf{x}(j)$ ; ステップ  $j$  の時の状態量を表わすベクトル ( $p \times 1$ )

(観測不能であってもかまわない)。

$\mathbf{y}(j)$ ; ステップ  $j$  の時の観測量を表すベクトル ( $q \times 1$ )

$\mathbf{u}(j)$ ; ステップ  $j$  の時の入力または制御量を表わすベクトル  
( $r \times 1$ )

$\Phi(j+1; j)$ ; ステップ  $j$  の状態量をステップ ( $j+1$ ) のものへ遷移させる  
マトリクス ( $p \times p$ )。(遷移行列)

$\Theta(j+1; j)$ ; ステップ  $j$  の入力をステップ ( $j+1$ ) での状態量へ寄与させ  
るマトリクス ( $p \times r$ )。(駆動行列)

$\mathbf{D}(j)$ ; 状態量と観測量を結ぶマトリクス ( $q \times p$ )。(観測行列)

$\mathbf{v}(j), \mathbf{w}(j)$ ; 各々 ( $p \times 1$ ), ( $q \times 1$ ) の次元を持つ誤差ベクトル。

(5-1)式は状態方程式, (5-2)式は観測方程式と呼ばれることもある。

一般に, 誤差は分散はもつが, 期待値はゼロで, 相互に独立であるから

$$E[\mathbf{v}(j)] = \mathbf{0}; \quad E[\mathbf{w}(j)] = \mathbf{0} \quad (5-3)$$

$$E[\mathbf{v}(j)\mathbf{w}'(j)] = \mathbf{0} \quad (5-4)$$

が成り立つ。また, 分散は次のように表す。

$$E[\mathbf{v}(j)\mathbf{v}'(j)] = \mathbf{Q}(j); \quad E[\mathbf{w}(j)\mathbf{w}'(j)] = \mathbf{R}(j) \quad (5-5)$$

(5-1)式から(5-5)式までの関係があり,  $\Phi, \Theta, \mathbf{D}$  が既知の時, (5-2)式より  $\mathbf{y}(j)$  の最適推定値  $\hat{\mathbf{y}}^0(j)$  を求めるには, (3-2)式の原理より, (5-3)式を考慮して

$$\hat{\mathbf{y}}^0(j) = E[\mathbf{y}(j)] = \mathbf{D}(j)E[\mathbf{x}(j)] + E[\mathbf{w}(j)] = \mathbf{D}(j)\mathbf{x}^0(j) \quad (5-6)$$

であり,  $\mathbf{x}^0(j)$  を求める必要がある。 $\mathbf{x}^0(j)$  は(5-1)式により

$$\begin{aligned} \mathbf{x}^0(j+1) &= E[\mathbf{x}(j+1)] = \Phi(j+1; j)E[\mathbf{x}(j)] + \Theta(j+1; j)E[\mathbf{u}(j)] + E[\mathbf{v}(j)] \\ &= \Phi(j+1; j)\mathbf{x}^0(j) + \Theta(j+1; j)\bar{\mathbf{u}}(j) \quad \text{ただし, } \bar{\mathbf{u}}(j) = E[\mathbf{u}(j)] \end{aligned} \quad (5-7)$$

であるから, 逐次的に求めることができるので, (5-6)式により  $\hat{\mathbf{y}}^0(j)$  は推定できる。しかし, この線形推定は  $\mathbf{x}(j)$  の推定如何にかかっている。そこで, 推定精度を向上させるには, 既知の観測データを有効に利用することが必要である。そうして, この問題は『観測値  $\{\mathbf{y}(j); j=0,$



1, 2, ..., k} が与えられた時、状態ベクトル  $x(j)$  の最小自乗法の意味での最適推定値  $\hat{x}^0(j)$  を求めること』に達着する。これがいわゆるカルマンの推定問題といわれるものである<sup>3)</sup>。

## 5.2 カルマン・フィルタの誘導 (カルマンの推定問題の解)

線形動的システムに関する式(5-1), (5-2)がある時、KALMAN は観測値の蓄積とともに推定の精度は向上するものとし、次式が成立するものとした。

### Kalman の仮定

$$\hat{x}^0(j|k) = \hat{x}^0(j|k-1) + A(j)\tilde{y}(k|k-1) \quad (5-8)$$

$$\text{ここに } \tilde{y}(k|k-1) = y(k) - \hat{y}^0(k|k-1) \quad (5-9)$$

$\hat{x}^0(j|k)$  は  $k$  ステップまでの観測値が既知の場合の  $x(j)$  の最適推定値を表す。

$\hat{x}^0(j|k-1)$ ,  $\hat{y}^0(k|k-1)$  などと同様である。

$A(j)$  は比例係数のマトリクス ( $p \times q$ )。

(5-8)式は、 $k$  ステップまでの観測値  $\{y(j); j=0, 1, 2, \dots, k\}$  がある場合の  $j$  ステップでの状態量の最適推定値  $\hat{x}^0(j|k)$  は、 $(k-1)$  ステップまでの観測値しかない場合の最適推定値  $\hat{x}^0(j|k-1)$  に対して、 $k$  ステップでの観測値  $y(k)$  と  $(k-1)$  ステップでの  $k$  ステップの最適推定値  $\hat{y}^0(k|k-1)$  との差  $\tilde{y}(k|k-1)$  に比例する量だけ修正されることを意味している。 $A(j)$  は未知のマトリクスであり、後に示すようにして決定されるが、KALMAN gain と呼ばれることもある。(5-8), (5-9)式は、この系がフィードバック系であることを示している。また、これは形式的には  $j$  と  $k$  の関係により、予測問題 ( $j > k$ )、フィルタ問題 ( $j = k$ )、平滑問題 ( $j < k$ ) に対応する。

(5-8)式を補助式として、(5-1), (5-2)式により状態量  $x$  の最適推定値を求める訳である。

(5-7)式より

$$\hat{x}^0(j+1|k) = \Phi(j+1; j)\hat{x}^0(j|k) + \Theta(j+1; j)\bar{u}(j) \quad (5-10)$$

となる。(5-6)式より、 $j = k$  とおくと

$$\hat{y}^0(k|k-1) = D(k)\hat{x}^0(k|k-1) \quad (5-11)$$

を得るので、(5-9)式は

$$\tilde{y}(k|k-1) = y(k) - D(k)\hat{x}^0(k|k-1) \quad (5-9)'$$

となる。したがって、(5-8)式は

$$\hat{x}^0(j|k) = \hat{x}^0(j|k-1) + A(j)\{y(k) - D(k)\hat{x}^0(k|k-1)\} \quad (5-12)$$

となる。

$A(j)$  が決まれば、(5-10), (5-12)式により逐次計算を行うことができるが、新たに得られる実際のデータは1ステップごとのものであるから、サイクリックな計算の便宜上からは、 $j = k$  の時、最も有効に働く。

### $A(j)$ の決定

$x(j)$  の推定誤差を

$$\hat{x}(j|k) = x(j) - \hat{x}^0(j|k) \quad (5-13)$$

のように表現すると, 最適推定値  $\hat{x}^0(j|k)$  は

$$P(j|k) = E[\hat{x}(j|k)\hat{x}'(j|k)] \quad (5-14)$$

を最小にするように最小自乗法の意味で推定される。一方, (5-8)式と(5-13)式から

$$\hat{x}(j|k) = \hat{x}(j|k-1) - A(j)\hat{y}(k|k-1) \quad (5-13)'$$

であるから,  $A(j)$  は,  $P(j|k)$  を  $A(j)$  で微分してゼロとおくことにより得られる「正規方程式」から求まる。すなわち,

$$\frac{dP(j|k)}{dA'(j)} = 0; \quad E[\hat{x}(j|k-1)\hat{y}'(k|k-1)] - A(j)E[\hat{y}(k|k-1)\hat{y}'(k|k-1)] = 0 \quad (5-15)$$

を解くことにより,  $A(j)$  は決定される。

ところで, (5-9)'式は

$$\hat{y}(k|k-1) = D(k)\hat{x}(k|k-1) + w(k) \quad (5-9)''$$

と書き替えられるので, (5-15)式の左辺第1項は, 誤差は相互に独立であることを考慮すると

$$E[\hat{x}(j|k-1)\hat{y}'(k|k-1)] = E[\hat{x}(j|k-1)\hat{x}'(k|k-1)]D'(k) \quad (5-16)$$

となる。また, 第2項の  $E[\ ]$  は(5-3)~(5-5)式を考慮して

$$\begin{aligned} E[\hat{y}(k|k-1)\hat{y}'(k|k-1)] &= D(k)E[\hat{x}(k|k-1)\hat{x}'(k|k-1)]D'(k) + R(k) \\ &= D(k)P(k|k-1)D'(k) + R(k) \end{aligned} \quad (5-17)$$

となるから,  $A(j)$  は結局, 次のようになる。

$$A(j) = E[\hat{x}(j|k-1)\hat{x}'(k|k-1)]D'(k)[D(k)P(k|k-1)D'(k) + R(k)]^{-1} \quad (5-18)$$

(5-15)式あるいは(5-18)式が得られる時

$$\begin{aligned} P(j|k) &= E[\hat{x}(j|k-1)\hat{x}'(j|k-1)] - A(j)E[\hat{y}(k|k-1)\hat{x}'(j|k-1)] \\ &= P(j|k-1) - A(j)E[\hat{y}(k|k-1)\hat{x}'(j|k-1)] \end{aligned}$$

であるが, 右辺第2項  $E[\ ]$  は, (5-16)式の転置であるから

$$P(j|k) = P(j|k-1) - A(j)D(k)E[\hat{x}(k|k-1)\hat{x}'(j|k-1)] \quad (5-19)$$

となる。一方, (5-1)式と(5-10)式から

$$\hat{x}(j+1|k) = \Phi(j+1; j)\hat{x}(j|k) + \Theta(j+1; j)\hat{u}(j) + v(j) \quad (5-20)$$

を得るから

$$\begin{aligned} P(j+1|k) &= E[\hat{x}(j+1|k)\hat{x}'(j+1|k)] = \Phi(j+1; j)P(j|k)\Phi'(j+1; j) \\ &\quad + \Theta(j+1; j)S(j)\Theta'(j+1; j) + Q(j) \end{aligned} \quad (5-21)$$

ただし,  $S(j) = E[\hat{u}(j)\hat{u}'(j)]$

となる。

(5-18), (5-19), (5-21)の三式により, 逐次式が構成され, 先の(5-10), (5-12)の両式と組み

合わされて、カルマンの推定問題の解が得られたことになる。

### 5.3 逐次計算手順

以上得られた結果を計算が便利なように、サイクリックな逐次式を構成する  $j=k$  の場合について、まとめると次のようになる。

$$\left\{ \begin{array}{l} \text{状態方程式} \quad \mathbf{x}(k+1) = \Phi(k+1; k)\mathbf{x}(k) + \Theta(k+1; k)\mathbf{u}(k) + \mathbf{v}(k) \end{array} \right. \quad (5-22)$$

$$\left\{ \begin{array}{l} \text{観測方程式} \quad \mathbf{y}(k) = \mathbf{D}(k)\mathbf{x}(k) + \mathbf{w}(k) \end{array} \right. \quad (5-23)$$

に対する逐次計算手順は

$$\left\{ \begin{array}{l} \mathbf{A}(k) = \mathbf{P}(k|k-1)\mathbf{D}'(k)[\mathbf{D}(k)\mathbf{P}(k|k-1)\mathbf{D}'(k) + \mathbf{R}(k)]^{-1} \end{array} \right. \quad (5-24)$$

$$\left\{ \begin{array}{l} \hat{\mathbf{x}}^0(k|k) = \hat{\mathbf{x}}^0(k|k-1) + \mathbf{A}(k)\{\mathbf{y}(k) - \mathbf{D}(k)\hat{\mathbf{x}}^0(k|k-1)\} \\ \quad = \{\mathbf{I} - \mathbf{A}(k)\mathbf{D}(k)\}\hat{\mathbf{x}}^0(k|k-1) + \mathbf{A}(k)\mathbf{y}(k) \end{array} \right. \quad (5-25)$$

$$\left\{ \begin{array}{l} \hat{\mathbf{x}}^0(k+1|k) = \Phi(k+1; k)\hat{\mathbf{x}}^0(k|k) + \Theta(k+1; k)\bar{\mathbf{u}}(k) \end{array} \right. \quad (5-26)$$

$$\left\{ \begin{array}{l} \mathbf{P}(k|k) = \{\mathbf{I} - \mathbf{A}(k)\mathbf{D}(k)\}\mathbf{P}(k|k-1) \end{array} \right. \quad (5-27)$$

$$\left\{ \begin{array}{l} \mathbf{P}(k+1|k) = \Phi(k+1; k)\mathbf{P}(k|k)\Phi'(k+1; k) \\ \quad + \Theta(k+1; k)\mathbf{S}(k)\Theta'(k+1; k) + \mathbf{Q}(k) \end{array} \right. \quad (5-28)$$

ただし、 $\mathbf{I}$ は単位マトリクスである。

$$\mathbf{Q}(k) = E[\mathbf{v}(k)\mathbf{v}'(k)]$$

$$\mathbf{R}(k) = E[\mathbf{w}(k)\mathbf{w}'(k)] \quad \mathbf{S}(k) = E[\bar{\mathbf{u}}(k)\bar{\mathbf{u}}'(k)]$$

である。

出力  $\mathbf{y}$  の最適推定値は、次式で与えられる。

$$\left\{ \begin{array}{l} \hat{\mathbf{y}}^0(k|k) = \mathbf{D}(k)\hat{\mathbf{x}}^0(k|k) \end{array} \right. \quad (5-29)$$

$$\left\{ \begin{array}{l} \hat{\mathbf{y}}^0(k+1|k) = \mathbf{D}(k+1)\hat{\mathbf{x}}^0(k+1|k) \end{array} \right. \quad (5-30)$$

係数マトリクス  $\Phi$ ,  $\Theta$ ,  $\mathbf{D}$  や分散マトリクス  $\mathbf{Q}$ ,  $\mathbf{R}$ ,  $\mathbf{S}$  などは、各ステップでは既知として、1つ前のステップの  $\mathbf{P}(k|k-1)$  と  $\hat{\mathbf{x}}^0(k|k-1)$  を用いると、(5-25)式はフィルタ問題、(5-26)式は予測問題の解となっている。 $k=0$  から計算を始めるとすれば、初期値  $\mathbf{P}(0|-1)$ ,  $\hat{\mathbf{x}}^0(0|-1)$  が必要となるが、最も簡単な仮定は、 $\hat{\mathbf{x}}^0(0|-1)=0$  である。その時、 $\hat{\mathbf{x}}(0|-1)=\mathbf{x}(0)$  であり、 $\mathbf{P}(0|-1) = E[\hat{\mathbf{x}}(0|-1)\hat{\mathbf{x}}'(0|-1)] = E[\mathbf{x}(0)\mathbf{x}'(0)]$  となる。したがって、 $\mathbf{x}(0)$  を適当に仮定すればよい。

(5-27)式は対称マトリクスであるので、対称の形で表現する。

$$\begin{aligned} \mathbf{P}(k|k) &= \{\mathbf{I} - \mathbf{A}(k)\mathbf{D}(k)\}\mathbf{P}(k|k-1)\{\mathbf{I} - \mathbf{A}(k)\mathbf{D}(k)\}' \\ &\quad + \{\mathbf{I} - \mathbf{A}(k)\mathbf{D}(k)\}\mathbf{P}(k|k-1)\{\mathbf{A}(k)\mathbf{D}(k)\}' \end{aligned} \quad (5-27)'$$

となるが、

$$\mathbf{A}(k) = \mathbf{I} - \mathbf{A}(k)\mathbf{D}(k) \quad (5-31)$$

とおくと

$$\mathbf{P}(k|k) = \mathbf{A}(k)\mathbf{P}(k|k-1)\mathbf{A}'(k) + \mathbf{A}(k)\mathbf{P}(k|k-1)\mathbf{D}'(k)\mathbf{A}'(k) \quad (5-27)''$$

一方, (5-24)式より

$$\begin{aligned} \Delta(k)\{D(k)P(k|k-1)D'(k)+R(k)\} &= P(k|k-1)D'(k) \\ \therefore \Delta(k)R(k) &= P(k|k-1)D'(k) - \Delta(k)D(k)P(k|k-1)D'(k) \\ &= [I - \Delta(k)D(k)]P(k|k-1)D'(k) = \Lambda(k)P(k|k-1)D'(k) \end{aligned}$$

上式を(5-27)''式の第2項へ代入すると

$$P(k|k) = \Lambda(k)P(k|k-1)\Lambda'(k) + \Delta(k)R(k)\Delta'(k) \tag{5-32}$$

となり,  $P(k|k-1)$ ,  $R(k)$  は対称マトリクスであるので,  $P(k|k)$  も対称形となる。

#### 5.4 同定問題への応用<sup>1)</sup>

これまで述べてきたカルマン・フィルタは係数マトリクス  $\Phi$ ,  $\Theta$ ,  $D$  は既知として, その場合の精度の向上を考えたけれども, 実際の問題では, 係数自身が未知であることが多い。最小自乗法では, 制約条件を満たすように係数を定める (同定) が, カルマン・フィルタにおいても, 状態量を表すベクトル  $x$  が観測可能な場合は, フィルタの計算過程と同様な方法で, 係数を定めることができる。係数も時間と共に遷移して, 正しい値に収束して行くという考えで, (5-22)式における変数と係数の役割を交替させて計算される。しかし, この計算では, 係数が必ずしも, 元の推定値  $\hat{x}^0$  を求める時の制約条件を満たしてはいない。以下に(5-24)式を用いて,  $\Phi$ ,  $\Theta$  を推定する方法を述べる。 $D$  についても, 全く同様な方法で推定できる。

(5-22)式における係数の要素を

$$\left. \begin{aligned} \Phi(k+1; k) &= \begin{bmatrix} \phi_{11} & \phi_{12} & \dots & \phi_{1p} \\ \phi_{21} & \phi_{22} & \dots & \phi_{2p} \\ \vdots & & & \\ \phi_{p1} & \phi_{p2} & \dots & \phi_{pp} \end{bmatrix} = \begin{bmatrix} \phi_1' \\ \phi_2' \\ \vdots \\ \phi_p' \end{bmatrix} \\ \text{ただし, } \phi_i' &= [\phi_{i1}, \phi_{i2}, \dots, \phi_{ip}] \\ \Theta(k+1; k) &= \begin{bmatrix} \theta_{11} & \theta_{12} & \dots & \theta_{1r} \\ \theta_{21} & \theta_{22} & \dots & \theta_{2r} \\ \vdots & & & \\ \theta_{p1} & \theta_{p2} & \dots & \theta_{pr} \end{bmatrix} = \begin{bmatrix} \vartheta_1' \\ \vartheta_2' \\ \vdots \\ \vartheta_p' \end{bmatrix} \\ \text{ただし, } \vartheta_i' &= [\theta_{i1}, \theta_{i2}, \dots, \theta_{ir}] \end{aligned} \right\} \tag{5-33}$$

とすると, (5-22)式は

$$\begin{aligned} x(k+1) &= \begin{bmatrix} \phi_1' \\ \phi_2' \\ \vdots \\ \phi_p' \end{bmatrix} x(k) + \begin{bmatrix} \vartheta_1' \\ \vartheta_2' \\ \vdots \\ \vartheta_p' \end{bmatrix} u(k) + v(k) \\ &= \underbrace{\begin{bmatrix} x'(k) & & & 0 \\ & x'(k) & & \\ & & \ddots & \\ 0 & & & x'(k) \end{bmatrix}}_{(p \times p)} \begin{bmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_p \end{bmatrix} + \underbrace{\begin{bmatrix} u'(k) & & & 0 \\ & u'(k) & & \\ & & \ddots & \\ 0 & & & u'(k) \end{bmatrix}}_{(p \times r)} \begin{bmatrix} \vartheta_1 \\ \vartheta_2 \\ \vdots \\ \vartheta_p \end{bmatrix} + v(k) \end{aligned}$$

$$= \begin{bmatrix} x'(k) & & 0 & u'(k) & & 0 \\ & x'(k) & & & u'(k) & 0 \\ & & \ddots & & & \\ 0 & & & x'(k) & & 0 \\ & & & & & u'(k) \\ & & & & & & u'(k) \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_p \\ \vartheta_1 \\ \vartheta_2 \\ \vdots \\ \vartheta_p \end{bmatrix} + v(k)$$

となり、

$$M(k) = \begin{bmatrix} x'(k) & & 0 & u'(k) & & 0 \\ & x'(k) & & & u'(k) & 0 \\ & & \ddots & & & \\ 0 & & & x'(k) & & 0 \\ & & & & & u'(k) \\ & & & & & & u'(k) \end{bmatrix} \quad (5-34)$$

$$h'(k) = [\phi_1', \phi_2', \dots, \phi_p' : \vartheta_1', \vartheta_2', \dots, \vartheta_p']$$

$$z(k) = x(k+1)$$

とおくと、結局、(5-22)式は

$$z(k) = M(k)h(k) + v(k) \quad (5-35)$$

となる。 $\{z(j); j=0, 1, 2, \dots, k\}$  および  $\{M(j); j=0, 1, 2, \dots, k\}$  が既知であるとして、 $h(k)$  を推定することになるが、(5-35)式は観測方程式(5-23)式と同型である。したがって、更にもう1つの遷移方程式

$$h(k+1) = \phi(k+1; k)h(k) + \omega(k) \quad (5-36)$$

をつくり、 $\phi(k+1; k)$  を既知とすれば、前と同様な議論ができる。結果のみ示すと、

$$\left. \begin{aligned} k(k) &= \Pi(k|k-1)M'(k)[M(k)\Pi(k|k-1)M'(k) + Q(k)]^{-1} \\ \hat{h}^0(k|k) &= \hat{h}^0(k|k-1) + k(k)\{z(k) - M(k)\hat{h}^0(k|k-1)\} \\ \hat{h}^0(k+1|k) &= \phi(k+1; k)\hat{h}^0(k|k) \\ \Pi(k|k) &= \{I - k(k)M(k)\}\Pi(k|k-1) \\ \Pi(k+1|k) &= \phi(k+1; k)\Pi(k|k)\phi'(k+1; k) + \Omega(k) \end{aligned} \right\} \quad (5-37)$$

ただし、 $I$  は単位マトリクス

$$\Omega(k) = E[\omega(k)\omega'(k)]$$

であり、(5-37)式により係数マトリクスは推定される。しかし、この場合も  $\phi(k+1; k)$  をいかに決めるかという問題が依然として残る。 $h(k)$  が時間によらない場合は、 $\phi(k+1; k) = I$  とおけば、 $h(k)$  は一定の値に収束することが予想される。

$D$  が未知の場合も、状態量を表すベクトル  $x$  が観測可能なら、(5-23)式で、 $x$  と  $D$  の役割を交替すれば、(5-35)式以下の議論が、そのまま成立する。

## §6. 結 論

今まで述べてきたカルマン・フィルタの説明は、平均値の概念、誤差の性質など数学的には厳密な議論をしないで、進めてきた。実在のデータは理想的な数学的条件に適うものはまれなので、応用する側としては、データの統計的性質などは、発散するような特別な事態にならない限りは、さ程、気にする必要はなかろう<sup>8,9)</sup>。

また、データ構造も理論的には

$$\text{(共通成分)} + \text{(その時点での固有成分)} + \text{(誤差)} \quad (\text{イ})$$

と分解されるが、実際の数値計算上での分離は困難で、最小自乗法では

$$\text{(誤差以外の成分)} + \text{(誤差)} \quad (\text{ロ})$$

の形でしか処理できない。

ロケット軌道のように物理的法則などからモデルが完全に同定される例もあるが、実際の問題では係数マトリクス自身は未知であることが多い。カルマン・フィルタでは、単一の過程内では係数の同定をしないので、あらかじめ回帰式などにより係数を求めておくことも考えられよう。しかし、カルマン・フィルタでも、5.4節で見たように、係数が未知の場合、係数推定のための別のカルマン・フィルタの計算を行うことにより、係数だけを別に推定することができる訳である。係数がステップごとに決まって行くという非定常性をもたせれば、非線形が非定常逐次計算に隠れる場合があり、非線形推定もこれで処理できる可能性がある。そうはいつても、係数を求める非定常計算を行うカルマン・フィルタの係数が必要になるという自己矛盾をきたし、やはり困難である。ついでに、もう1つ難点を挙げると、変数の個数を増すと係数の個数が急速に増え、記憶容量が増えたり、安定性がなくなることが報告されている<sup>10,11)</sup>。

また、推定精度の向上のために、過去のデータを用いるため、どうしても外挿的性質は避けられない。

要は、モデルの設定や同定が一番肝腎なことである。

ともあれ、逐次式による計算法はカルマン・フィルタの優れた点であり、

- 1) 計算機で計算する場合、データ記憶のメモリが節約できる。
- 2) 非定常問題の解も可能である。

などの利点があり、オンライン適応制御には適しているであろう。実用に当たっては、対象とする系とサンプリング間隔など吟味しなければならない点が多々あることに留意する必要がある。

## 謝 辞

本稿をおわるにあたり、文献その他で種々の御教示を戴いた本学工学部電気工学科の白井五郎、計測制御専攻の渡辺嘉二郎の両助教授に謝意を表す。

## 参考文献

- 1) Kalman, R.E., "A New Approach to Linear Filtering and Prediction Problems", *Trans. ASME, Ser. D, J. Basic Eng.*, Vol. 82, 1960, pp. 35-45.
- 2) 藤井澄二他3名, 「制御工学Ⅲ (岩波講座基礎工学20)」, 岩波書店, 1968年7月, pp. 310-318.
- 3) Kuo, B.C., "Discrete-Data Control Systems", Prentice-Hall, pp. 302-323.
- 4) 鈴木 胖, 「4. 不規則データによるシステム同定」, 計測と制御, 第14巻第1号, 昭和50年1月, pp. 132-139.
- 5) Lee, R.C.K., "Optimal Estimation, Identification of Control", M.I.T. Press, 1964, pp. 35-78.
- 6) Cox, H., "On the Estimation of State Variable and Parameters for Noisy Dynamic Systems", *IEEE Trans. Automat. Contr.*, Vol. AC-9, Jan. 1964, pp. 5-12.
- 7) 相良節夫, 「同定問題」, 計測と制御, 第8巻第4号, 昭和44年4月, pp. 268-280.
- 8) Fitzgerald, R.J., "Divergence of the Kalman Filter", *IEEE Trans. Automat. Contr.*, Vol. AC-16, No. 6, Dec. 1971, pp. 736-747.
- 9) Thé, G., "Kalman-filter divergence due to process-noise decoupling", *Proc. IEE*, Vol. 121, No. 6, June 1974, pp. 525-528.
- 10) 西村敏充, 「米国の宇宙計画におけるカルマン・フィルタの応用—I」, システムと制御, Vol. 18, No. 10, 1974, pp. 560-566.
- 11) 西村敏充, 「米国の宇宙計画におけるカルマン・フィルタの応用—II」, システムと制御, Vol. 18, No. 11, 1974, pp. 613-620.

以下, カルマン・フィルタおよびそれに関連するものを扱った文献.

- 12) Kalman, R.E. and R.S. Bucy, "New Results in Linear Filtering and Prediction Theory", *Trans. ASME, Ser. D, J. Basic Eng.*, Vol. 83, 1961, pp. 95-108.
- 13) 有本 卓, 「カルマン・フィルタ」, 計測と制御, 第9巻第8号, 昭和45年8月, pp. 590-598.
- 14) Jazwinski, A.H., "Stochastic Processes and Filtering Theory", Academic Press, 1970.
- 15) 日野幹雄, 「Kalman の予測推定理論の平易な誘導について」, 東工大・土木工学科研究報告, No. 15, 1973年12月, pp. 91-99.
- 16) 日野幹雄, 「カルマン・フィルタと多変量解析を組み合わせた確率過程の予測法」, 土木学会論文報告集第228号, 1974年8月, pp. 45-54.
- 17) 有本 卓, 「カルマン・フィルタ」, 産業図書, 昭和52年1月.
- 18) Kailath, T., "A View of Three Decades of Linear Filtering Theory", *IEEE Trans. Inform. Theory*, Vol. IT-20, No. 20, Mar. 1974, pp. 146-181.
- 19) 日野幹雄, 「水文流出系予測へのカルマン・フィルタ理論の適用」, 土木学会論文報告集第221号, 1974年1月, pp. 39-47.
- 20) 日野幹雄, 森, 吉川, 「カルマン・フィルタによる大気汚染の予測」, 土木学会論文報告集第224号, 1974年4月, pp. 79-90.
- 21) 日野幹雄, 石川, 穴戸, 「カルマン・フィルタによる洪水予測の二・三の例」, 第19回水理講演会論文集, 1975年2月, pp. 77-82.
- 22) Beck, B. and P. Young, "Systematic Identification of DO-BOD Model Structure", *J. of the Environmental Engineering Div., ASCE*, Vol. 102, No. EE5, Oct. 1976, pp. 909-927.

## 付録A. ベクトルに関する微分

## 1. スカラー変数によるベクトル変量の微分

ベクトル  $y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_q \end{bmatrix}$  のスカラー変数  $x$  による微分は

$$\frac{dy}{dx} = \begin{bmatrix} \frac{dy_1}{dx} \\ \frac{dy_2}{dx} \\ \vdots \\ \frac{dy_q}{dx} \end{bmatrix} \quad (A-1)$$

と定義する。

2. ベクトル変数によるスカラー変量の微分

ベクトル  $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix}$  によるスカラー変量  $C$  の微分は

$$\frac{dC}{dx} = \left[ \frac{\partial C}{\partial x_1}, \frac{\partial C}{\partial x_2}, \dots, \frac{\partial C}{\partial x_p} \right], \quad \frac{dC}{dx'} = \begin{bmatrix} \frac{\partial C}{\partial x_1} \\ \frac{\partial C}{\partial x_2} \\ \vdots \\ \frac{\partial C}{\partial x_p} \end{bmatrix} \quad (A-2)$$

と定義する。

3. ベクトルの微分 (ヤコビアン)

$$\frac{dy}{dx} = \begin{bmatrix} \frac{\partial y_1}{\partial x_1}, \frac{\partial y_1}{\partial x_2}, \dots, \frac{\partial y_1}{\partial x_p} \\ \frac{\partial y_2}{\partial x_1}, \frac{\partial y_2}{\partial x_2}, \dots, \frac{\partial y_2}{\partial x_p} \\ \vdots \\ \frac{\partial y_q}{\partial x_1}, \frac{\partial y_q}{\partial x_2}, \dots, \frac{\partial y_q}{\partial x_p} \end{bmatrix} \quad (A-3)$$

$$\frac{dy'}{dx'} = \begin{bmatrix} \frac{\partial y_1}{\partial x_1}, \frac{\partial y_2}{\partial x_1}, \dots, \frac{\partial y_q}{\partial x_1} \\ \frac{\partial y_1}{\partial x_2}, \frac{\partial y_2}{\partial x_2}, \dots, \frac{\partial y_q}{\partial x_2} \\ \vdots \\ \frac{\partial y_1}{\partial x_p}, \frac{\partial y_2}{\partial x_p}, \dots, \frac{\partial y_q}{\partial x_p} \end{bmatrix} \quad (A-3)'$$

と定義する。

$$\frac{dx}{dx} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix} = I, \quad \text{同様に } \frac{dx'}{dx'} = I \quad (A-4)$$



4. 2次形式の微分

$A$  を正方行列 ( $p \times p$ ) とする時,  $q=p$  としてスカラー量  $x' Ay$  の微分を考える。ただし,  $x, y$  ともに他のベクトル  $t$  の関数とする。

$$\frac{d}{dt'}(x' Ay) = \frac{dx'}{dt'}(Ay) + \frac{dy'}{dt'}(A'x) \tag{A-5}$$

$\therefore$  (スカラー量だから  $x' Ay = (x' Ay)' = y' A' x$ )

注) (A-5)式を求めるのに

$$\frac{d}{dt'}(x' Ay) = \frac{dx'}{dt'}(Ay) + x' A \cdot \frac{dy}{dt'}$$

とすると, 右辺第二項の演算は未定義なので実行不可能である。

$$\frac{d}{dx'}(x' Ax) = (A + A')x \tag{A-6}$$

$A$  が対称行列すなわち  $A' = A$  なら

$$\left. \begin{aligned} \frac{d}{dx'}(x' Ax) &= 2Ax \\ \frac{d}{dx}(x' Ax) &= 2x' A \end{aligned} \right\} \tag{A-7}$$

付録 B. 自乗誤差の微分

未知マトリクス

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{q1} & a_{q2} & \cdots & a_{qp} \end{bmatrix} = \begin{bmatrix} a_1' \\ a_2' \\ \vdots \\ a_q' \end{bmatrix}$$

とし,  $y$  の推定値が  $(q \times 1)$

$$\hat{y} = Ax \quad (p \times 1)$$

で与えられる時

$$J = (y - \hat{y})(y - \hat{y})'$$

に極値 (最小値) を与える  $A$  は,  $J$  の各要素を  $a_k$  で微分することにより求められる。

$$J = (yy') - A(xy') - (yx')A' + A(xx')A' \tag{B-1}$$

$(xy') = \{z_1, z_2, \dots, z_q\}$  と表わすと  $J$  の要素は

$$\{J_{ij}\} = \{y_i y_j\} - a_i' \cdot z_j - z_i' \cdot a_j + a_i'(xx')a_j$$

各項はスカラー量であることに注意して, 微分すれば,

$$\frac{\partial \{J_{ij}\}}{\partial a_k'} = \frac{da_i'}{da_k'} \{(xx')a_j - z_j\} + \frac{da_j'}{da_k'} \{(xx')a_i - z_i\}$$

$$\left. \begin{aligned} i=k; \quad j=i & \quad \frac{\partial J_{kk}}{\partial a_k'} = 2\{(xx')a_k - z_k\} \\ \text{(対角要素)} & \\ j \neq i & \quad \frac{\partial J_{kj}}{\partial a_k'} = (xx')a_j - z_j \\ \text{(非対角要素)} & \\ i \neq k; \quad j=k & \quad \frac{\partial J_{ik}}{\partial a_k'} = (xx')a_i - z_i \\ j \neq k & \quad \frac{\partial J_{ij}}{\partial a_k'} \equiv 0 \end{aligned} \right\}$$

求める  $A$  は  $\frac{\partial \{J_{ij}\}}{\partial a_k'} = 0$  とおくことにより求められるが、上式の中で意味をもつものは、

$$(xx')A' - (xy') = 0 \tag{B-2}$$

あるいは

$$A(xx') - (yx') = 0 \tag{B-3}$$

である。

形式的に上述のことを

$$\frac{dJ}{dA} = 2\{(xx')A' - (xy')\} = 0 \tag{B-2}'$$

$$\frac{dJ}{dA'} = 2\{A(xx') - (yx')\} = 0 \tag{B-3}'$$

と表すことにする。(B-2)式、(B-3)式より  $A' = (xx')^{-1}(xy')$  または  $A = (yx')(xx')^{-1}$  (B-1)式の平均値

$$E[J] = E[yx'] - A\{E[xy']\} - E[yx']A' + A\{E[xx']\}A' \tag{B-4}$$

の極値を扱う時も同じ議論により

$$\frac{dE[J]}{dA'} = 2\{A\{E[xx']\} - E[yx']\} = 0 \tag{B-5}$$

$$\therefore A = E[yx']\{E[xx']\}^{-1} \tag{B-6}$$

を得る。