最小2乗推定の一手法とその手動制御系への応用[†]

小 畑 秀 文*

An Algorithm for Least-Squares Estimation and Its Application to a Manual Control System

Hidefumi KOBATAKE*

This paper presents an algorithm for least-squares estimation of the states of a dynamic system. This method based on a modification of the Gauss-Newton method possesses characteristics of both the steepest descent method and the Gauss-Newton method. Consequently, it has the advantage of quick convergence over the steepest descent method and the advantage of good stability over the Gauss-Newton method. This method is useful when the observer of the states of the system is very noisy and/or when the initial estimation errors are large. A computer simulation substantiates the above remarks. The application to the parameter estimation of a response model of the human operator in a control system is also shown.

1. 序 論

未知システムの入出力データからその動特性を推定する問題に関してはすでに多くの研究がなされている。それらの多くはシステムの構造は既知であるとし、それに含まれる未知パラメータを推定しようとするものであり、古くは最大傾斜法^{1), 2)}か Gauss-Newton 法³⁾に基づいたものが多く、近年ではカルマン・フィルタをその基本としたものが主流となっている。

本論文では古いタイプに属する Gauss-Newton 法を変形した一つのシステムの同定法を提案する。これは D. W. Marquardt の方法⁴⁾と同様に最大傾斜法のもつ安定性と Gauss-Newton 法のもつ収束の早さとをそれぞれ適度に取り入れることができ、初期推定値が大きな誤差を含んでいたり、観測雑音が大い場合など適度の安定性と収束速度とを要求されるときに有

効と思われる。本法はすでに得られている入出力データに対してのパラメータの最適化には持論のこと、観測値が得られるごとに推定を行なう逐次推定にも用いることができる。

本手法の有効性を示すために、観測雑音がきわめて大きい系の一例として予測という比較的困難な作業を含む手動制御系を取上げ、人間の動特性同定に本手法を用いた結果を、計算機シミュレーションと実際の系への適用例とで示す。

2. 最小2乗法に基づくパラメータ推定法

2.1 問題の記述

取扱う系は構造既知のシステムとその観測系から成るとし、システムの状態変数と観測出力はつぎのように離散値系として与えられるものとする。

$$x_k = g(x_{k-1}, u_{k-1}) \quad (1)$$

$$y_k = h(x_k) + n_k \quad (2)$$

x, u, y, n はそれぞれ $n \times 1$ 状態ベクトル, $r \times 1$ 外力ベクトル, $m \times 1$ 出力ベクトル, $m \times 1$ 観測雑音ベクトル, g, h は $n \times 1$ および $m \times 1$ の既知非線形関数ベクトルであり, k は時刻を表すものとする。 n は加算的に重畳する観測雑音であり、その平均値は0と仮定する。

関数 g に未知パラメータ a が含まれている場合を考えると、 a の値に時間的な変動はなく一定と考えれば、

$$a_k = a_{k-1} \quad (3)$$

と書けるから、 a を遷移方程式(3)を満たす状態変数とみなすことができる。そこで以下ではパラメータ推定問題を一般的な状態変数の推定問題として取扱う。

時刻 i での状態変数の推定値を $\hat{x}_i$ とする。問題は各時刻 $i(i=1, 2, \dots, l)$ での観測値 y_i と外力 u_i とを知ったとき、観測値と推定値の誤差に関する2次形式

$$E(l) = \sum_{i=1}^l \{y_i - h(\hat{x}_i)\}^T R_i \{y_i - h(\hat{x}_i)\} \quad (4)$$

[†] 第10回計測自動制御学会学術講演会にて一部発表(昭和46・8)

* 東京大学宇宙航空研究所 東京都目黒区駒場 4-6-1

* Institute of Space and Aeronautical Science, University of Tokyo, Meguro-ku, Tokyo

(Received February, 4, 1974)

を最小にするように $\bar{x}_i$ を決定することである。ここに T は転置を表わし、 R_i は $m \times m$ 正値対称行列であり、

$$\bar{x}_i = g(\bar{x}_{i-1}, u_{i-1}) \quad (5)$$

である。推定はデータがすでに得られているときには初期値 $\bar{x}_0$ を、逐次推定の場合には現在値 $\bar{x}_i$ を求める問題になる。

2.2 記号の説明

本論文では非線形関数を Taylor 展開してつぎのように近似する。

$$g(x) \doteq g(x_0) + G(x_0)(x - x_0) \quad (6)$$

$$h(x) \doteq h(x_0) + H(x_0)(x - x_0) \quad (7)$$

$$E(l) \doteq E_0(l) + e_0^T(x - x_0) + \frac{1}{2}(x - x_0)^T S_0(x - x_0) \quad (8)$$

ここに G, H は $n \times n, m \times n$ の Jacobian matrix であり、

$$G(x) = \frac{\partial g}{\partial x}, \quad H(x) = \frac{\partial h}{\partial x} \quad (9)$$

e_0 は $n \times 1$ ベクトル、 S_0 は $n \times n$ 行列であり、

$$e_0 = \frac{\partial E(l)}{\partial x} \bigg|_{x_0}^T, \quad S_0 = \frac{\partial^2 E(l)}{\partial x^2} \bigg|_{x_0} \quad (10)$$

である。 G は遷移行列と呼ばれるが、その逆行行列を M とする。すなわち、

$$M_k = \left(\frac{\partial x_i}{\partial x_{k-1}} \right)^{-1} \quad (11)$$

2.3 従来の方法

パラメータ最適化の方法の代表的なものに最大傾斜法と Gauss-Newton 法とがある。前者は $E(l)$ の最大傾斜ベクトル方向への修正を繰返してゆくもので、この方法の利点は「修正係数を適当に小さくすれば試行回数を増してゆく」と確実に最適解に収束することである。ただし修正係数を定める一般的方法はなく、経験的に定めなければならないこと、最適解に近づく収束が遅くなるという二つの欠点がある。一方後者は $E(l)$ を状態変数に関する 2 次形式で (8) 式のように近似し、それが最小となるように推定値を求めるものである。一般に極小値付近では $E(l)$ は状態変数の最適解からの偏差の 2 次形式で近似できると考えてよいから、推定値が最適値に近い場合には収束はきわめて早い。ただし最大傾斜法による修正ベクトル Δx_{SD} と Gauss-Newton 法による修正ベクトル Δx_{GN} とのなす角を θ とするとき、

$$\theta < 90^\circ \quad (12)$$

が成立する⁴⁾が、 θ が 80° を越えることはめずらしいことではなく、修正方向が最大傾斜方向と大きく異なる

ことがあるため、推定値が最適値に近くない場合には必ずしも収束するとは限らない。

2.4 Gauss-Newton 法の変形

Gauss-Newton 法を変形することにより最大傾斜法と Gauss-Newton 法のもつ特性の中間の特性をもたせることができる。その変形法とは 2 次形式で近似した誤差の 2 乗和 $E(l)$ に対し最大傾斜法を一定回数だけ適用し、その結果得られた推定値に基づいて改めて $E(l)$ の 2 次形式近似を求めて同様の手順を繰返すものである。(8) 式より α を修正係数として最初の修正ベクトルを

$$\Delta x_1 = -\alpha e_0^0 \quad (13)$$

と定める。ただし $e_0^0 = e_0$ である。この修正の結果、最大傾斜ベクトルは

$$e_0^1 = e_0^0 + S_0 \Delta x_1 = (I - \alpha S_0) e_0^0 \quad (14)$$

となる。 I は単位行列である。この e_0^1 から第 2 回目の修正量 Δx_2 が得られる。以下同様の手順であらじめ定められた回数 N だけ修正を繰返せば、新しい推定値 $\hat{x}_0$ は

$$\hat{x}_0 = \bar{x}_0 + \sum_{i=1}^N \Delta x_i^0 \quad (15)$$

となる。この $\hat{x}_0$ を $\bar{x}_0$ として改めて $E(l)$ の 2 次形式近似を求め同じ手順を繰返す。

2.5 変形法の解析

(15) 式右辺第 2 項の修正ベクトルを $\Delta \bar{x}_0$ とすれば、

$$\Delta \bar{x}_0 = V_{GN} + V_{SD} \quad (16)$$

と書ける。ただし、

$$V_{GN} = -S_0^{-1} e_0^0 \quad (17)$$

$$V_{SD} = (I - \alpha S_0)^N S_0^{-1} e_0^0 \quad (18)$$

である。 V_{GN} は (8) 式の最小値を与える修正量であり、Gauss-Newton 法より得られる修正ベクトルそのものである。この修正ベクトルの方向を V_{SD} だけ補正したところが本手法の特徴であり、その結果つぎのような性質が生れる (証明は付録参照)。

(1) $N=1$ のとき修正方向は最大傾斜方向となる。

(2) $0 < \alpha < 2/\lambda_{\max}$ ($\lambda_{\max}$ は S_0 の最大固有値) のとき、 N の増大と共に $\Delta \bar{x}_0$ は V_{GN} に漸近してゆき、 $N \rightarrow \infty$ で両者は一致する。すなわ

ち、 $N \rightarrow \infty$ ち、 $V_{SD} \rightarrow 0$ となる。

(3) $0 < \alpha < 1/\lambda_{\max}$ のとき

(i) N の増大と共に $\Delta \bar{x}_0$ は最大傾斜ベクトル方向から単調に V_{GN} に漸近してゆく。そのときノルム $\|\Delta \bar{x}_0\|$ は単調増加である。

(ii) α が大きいほど V_{GN} への収束は早い。

(iii) 修正量は、固有値が大きき $E(l)$ に与える影響力が大きい固有ベクトル方向ほど大きく、 $E(l)$ に与える影響力の小さい固有ベクトル方向ほど修正量は押えられる。

すなわち、本法は安定性と収束速度の重要度に応じて α と N を選ぶことにより修正方向を e_0 と V_{GN} の間の任意の方向に選べるという特徴をもつ。これは D. W. Marquardt の提案した方法⁴⁾の特性と同じであるが、本法にはつぎのような特徴がある。

(i) 荷重行列 S_0 の逆行列を求める必要がない。

(ii) 修正係数 α はきわめて簡単に決定できる。

(i) は繰返し数 N が小さければ計算量の点で大きな利点となり、(ii) の α は収束を保証するものではないが、あとで示すように安定性は著しく増大することから、本法は最大傾斜法では収束が遅く、Gauss-Newton 法では不安定となるような“中間領域”において有効な方法といえよう。

2.6 α と N の範囲

本節では前述の“中間領域”での α と N の実用的な値を与えることにする。

α : 上限は S_0 の最大固有値で決まるが繰返しの都度固有値を求めるのは実用的とはいえず、簡単に求められる近似値で代用する必要がある。ここで、

$$\text{tr}(S_0) = \sum_{j=1}^n \lambda_j \quad (19)$$

なる関係を用いる。 $\text{tr}(\cdot)$ は trace を意味し、 S_0 の対角要素の総和である。 S_0 は正値行列で、その固有値 λ_j は非負であるから、 α が $2/\lambda_{\max}$ 以下になるためには

$$\alpha = \frac{a}{\text{tr}(S_0)} \quad (0 < a < 2.0) \quad (20)$$

であればよい。一定の修正量を得るには α は大きいほど N は小さくてすむ。しかも本法が前節(3)の性質をできるだけ保つようにするため、 $\alpha \approx 1/\lambda_{\max}$ とする必要があるが、特殊な場合を除き実用上は $a=2.0$ としてよい。

N : 修正量は N が一定でも扱うシステムの構造やその状態変数の値に直接依存するので一般的に述べることは困難であるが、つぎに示す例から N は大体 $1 \leq N \leq 5$ の間で選定するのがよいと思われる。Fig. 1(a) に修正量 $\Delta \bar{x}$ と修正方向の N による変化を示す。荷重行列 S としては対角要素 $s_{ii} = 2^{i-1}$ ($i=1, 2, \dots, 5$) である対角行列をとり、最大傾斜ベクトルとして

$$e_1 = (2^0, 2^{-1}, 2^{-2}, 2^{-3}, 2^{-4})^T, \quad e_2 = (1, 1, 1, 1, 1)^T, \\ e_3 = (2^0, 2^1, 2^2, 2^3, 2^4)^T \quad (21)$$

の3種類を考えた。 e_1 は V_{GN} とのなす角度が小さ

Fig. 1 Relationship between the correction vector and the iteration time (a) $\text{NORM } \|\Delta x_i\|$ is normalized by $\|V_{GN}\|$. ANGLE means the angle between the vector $\Delta \bar{x}_i$ and V_{GN} normalized by that between e_i and V_{GN} (b) Norm ratio of $\Delta \bar{x}$ and V_{GN} in case of $N=2$

い例であり、 $N=1$ ではほぼ1に近い修正量を得る。 e_3 は V_{GN} と方向が大きく異なる場合で、得られる修正量は小さい。安定性が増すのはこのためである。本図より繰返し数が数回程度のところが前述の“中間領域”に相当するものと考えられる。

次元数による修正量の変化の一例を Fig. 1 (b) に示す。 e は各要素が総て1である n 次元ベクトル、 S は対角要素が等比例である対角行列 ($s_{ii} = 0.5s_{i-1, i-1}$, $s_{11} = 1, i=1, 2, \dots, n$) を用いた。この組み合わせは V_{GN} と e の方向の違いの平均的なものと考えられる。この例でも知れるように $N=2$ でも大体 V_{GN} の50%前後の修正量を得られることが期待でき、 N が小さくてすむことから本法の利点(i)が生かされているといえよう。

2.7 逐次推定

本法は逐次推定にそのまま適用することができる。時刻 l で新しい観測値が得られたとき、状態変数の推定値 $\hat{x}_l$ をつぎのように定める。

$$\hat{x}_l = \hat{x}_l + \Delta \hat{x}_l \quad (22)$$

ここに

$$\Delta \hat{x}_l = -\alpha \sum_{i=1}^N e_i^{i-1} \quad (23)$$

$$e_i^i = (I - \alpha S_i) e_i^{i-1} \quad (24)$$

$$\mathbf{e}_i^0 = \mathbf{M}_i^T \mathbf{e}_{i-1}^N - 2\mathbf{H}_i^T \mathbf{R}_i \{\mathbf{y}_i - \mathbf{h}(\hat{\mathbf{x}}_i)\} \quad (25)$$

$$\mathbf{S}_i = \mathbf{M}_i^T \mathbf{S}_{i-1} \mathbf{M}_i + 2\mathbf{H}_i^T \mathbf{R}_i \mathbf{H}_i \quad (26)$$

$$\alpha = a / \text{tr}(\mathbf{S}_i) \quad (27)$$

$$\hat{\mathbf{x}}_i = \mathbf{g}(\hat{\mathbf{x}}_{i-1}, \mathbf{u}_{i-1}) \quad (28)$$

この場合にも $\mathbf{S}_i$ の逆行列は求める必要がなく、荷重を事前に決めかねるとき $\mathbf{S}_0 = (0)$ とできるから取扱いは簡単になる。

3. 変形法によるパラメータ推定実験

本節では観測雑音が大きい系のパラメータ推定に本法を適用し、その有効性を示す。取扱う系は制御者が入力信号の未来値を予測する手動制御系であり、その動作モデルの未知パラメータを推定の対象とした。まず計算機シミュレーションにより、推定が精度よく行なわれることを示し、つぎに実際の系の入出力関係を用いて推定を行ない、その結果が妥当なものであること、およびそれから興味ある関係式を見出したことを述べる。

3.1 予測動作と非線形性

予測動作とは制御者に対して入力信号を表示して与え、その未来値を連続的に予測させ、その予測値を制御者の出力（予測系の出力でもある）とするものである。表示にはブラウン管オシロスコープを用い、入出力をいずれも上下に振れる 2 輝点で表わし、入力輝点に対し出力輝点を予測時間に相当する距離だけ先行させて水平に掃引する。入力輝点の描く波形は 2 次元波形であり、それに先行する出力輝点はこの 2 次元波形の外挿点とみることでもできる。制御者は制御棒を手で前後に操作することにより出力輝点の位置を決める（実験装置の詳細は文献 5）を参照）。

制御者の予測動作の特性に関してはすでにその基本的性質を明らかにしたが⁹⁾、レムナントと呼ばれる雑音成分が出力の 50% 前後にもなること、予測時間 γ が大きくなるにつれて予測出力に非線形性が顕著に現われることなどが特徴である。入力に正弦波を用いて実験した結果つぎのことが知れた。この傾向は不規則入力に対しても同様である。

(1) 出力波形は主に奇数次高調波を含み、基本波に対するその割合 ρ は γ に対し単調増加と考えられる (Fig. 2 参照)。

(2) γ が大きくなると出力は矩形波に近い波形になる。

すなわちこの非線形特性は原点对称な一種の飽和特性と考えられる。制御者の予測は基本的には微分動作によることが明らかにされているが、上の事実から非線

Fig. 2 Power ratio of the third-order harmonics and the fundamental wave in the prediction output of the subject. f_0 and γ mean the frequency in Hz of the sinusoidal input and the prediction time span respectively

Fig. 3 Model of the prediction task of the operator

形要素はこの微分が行なわれたあとにあると思われる。Fig. 3 は以上を単純化した数式モデルである。飽和要素は簡単のため $\tanh(\cdot)$ で近似し、予測は 1 次微分によるものとした。飽和要素のあとの 1 次遅れは制御者の最も簡単で一般性のある伝達特性を表わすものである。このモデルに従えば制御者出力 $x(t)$ は次式で記述される。

$$\frac{dx}{dt} + T_L x = K \tanh \left\{ C \left(u + T_A \frac{du}{dt} \right) \right\} \quad (29)$$

実際の予測出力 $y(t)$ は $x(t)$ に (29) 式で記述できないレムナント $n(t)$ が加わったものでつぎのようになる。

$$y(t) = x(t) + n(t) \quad (30)$$

3.2 計算機シミュレーション

(29), (30) 式は離散値系に書きかえられるから前者をシステム方程式、後者をその観測系とみれば本法が適用できる。計算機により逐次推定実験をつぎのように行なった。システムは 0.1 秒ごとの離散値系とし、系の観測および推定も 0.1 秒ごとに行なった。外力 u は中心周波数 0.3 Hz の帯域雑音を計算機内で発生させて用いた。その標準偏差は $\sigma_u = 0.4$, $\sigma_{\dot{u}} = 0.86$ である。誤差の 2 乗和は観測時刻を始点として過去に向かって時定数 0.2 秒で指数的に減衰する重みをつけた。パラメータの値は $T_L = K = 3.0$, $C = 1.0$, $T_A = 1.5$ とした。 T_A だけは後述するように本法の有効性を確かめるために大きくし、ほかは実際の系に近い値とした。 $a = 2.0$ である。 T_L , K , C , T_A を時間変化の無

Table 1 Number of times of incorrect estimation

k	N	
	2	5
1	1	8
2	5	7
3	1	5
4	0	4

い状態変数とし、(29)式の微分を時間間隔 0.1 秒の差分で近似すれば、系は5次元遷移方程式となる。未知パラメータの初期推定値は順に、2.0, 1.0, 0.2, 0.1 とした。推定は10種類の白色の観測雑音列に対し、その標準偏差 σ_n を4段階 ($\sigma_n = (k/5)\sigma_x$, $k=1, 2, 3, 4$) に変えて N を2と5の場合について比較検討を行なった。 $k=1$ でも雑音と初期推定誤差とが共に大きいため Gauss-Newton 法ではいずれも不安定であった。

外力 u に関する成分 Cu とその微分値 $\dot{u}$ に関する成分 $CT_A\dot{u}$ とが出力 x に及ぼす影響力を比較すると、

$$\frac{CT_A\sigma_{\dot{u}}}{C\sigma_u} > 3 \quad (31)$$

の関係があり、微分値 $\dot{u}$ に関する成分が3倍以上の力をもっている。このことと、初期推定値 $\hat{C}$, $\hat{T}_A$ が共に0に近いことから、雑音が大きければ、 $\dot{u}$ に関してみれば正しく推定され、 u に関しては符号が逆になる値、すなわち、

$$\hat{C} < 0 \text{ かつ } \hat{T}_A < 0 \quad (32)$$

の領域に落込み、そこで安定しやすい。この領域が一つの安定領域であることは直感的にも明らかである。

Table 1 は推定値が(32)式の領域に収束してしまった回数を比較したもので、Gauss-Newton 法に比べて安定性において数段優ることがわかる。 $N=2$ のパラメータ推定値の平均値の変化を **Fig. 4** に示した。図は2秒おきの値の変化のみを示した。 $k=1$ および3のとき10回のうち1回だけ推定値が(32)式の領域に収束しており、図はそれらも含めた平均値である。図の推定値が真の値よりも小さめになっているのはそのためであり、その影響を除けば推定値は20秒位で真の値に近いところまで達していることが知れた。これから本法が初期推定誤差が大で雑音の大きい系に対して有効であることが知れよう。

3.3 予測動作への応用

Table 2 Estimated parameter values

Input	Subject	r	T_L	K	C	T_A	σ_u	$\sigma_{\dot{u}}$
sinusoidal 0.2 Hz	TF	0	1.89	3.42	1.20	0.36	0.66	1.36
		0.2	2.44	3.33	1.10	0.57	0.65	1.39
		0.4	2.46	3.31	1.00	0.77	0.66	1.37
sinusoidal 0.3 Hz	HK	0.1	2.45	2.73	1.27	0.47	0.60	1.22
		0.2	2.67	2.56	1.08	0.63	0.60	1.27
		0.3	2.50	2.43	1.02	0.91	0.61	1.22
sinusoidal 0.3 Hz	KH	0.1	2.73	2.79	1.35	0.41	0.69	1.42
		0.2	2.77	2.80	1.18	0.59	0.68	1.42
		0.3	3.01	2.86	1.16	0.81	0.69	1.42
random 1 random 2	HK	0.2	2.70	2.24	1.40	0.48	0.39	—
		0.2	2.68	2.20	1.28	0.70	0.40	—
random 1 random 2	TF	0.2	2.68	2.29	1.33	0.11	0.52	—
		0.2	2.37	2.43	1.35	0.40	0.54	—

推定に用いたデータの長さは51.2秒間で始めの20秒間は初期推定誤差の影響を避けるため除外し、後半の31.2秒間の推定値の平均を制御者のパラメータ推定値とした。入力信号は不規則入力1, 2の2種類(注1)および正弦波であり、それらに対する場合の推定値を **Table 2** に示す。正弦波のデータは **Fig. 2** のそれと同一のものであり、不規則入力の場合も含め高調波の強さや出力の入力に対する位相進み量からみて、これらパラメータ値は妥当なものであることが知れた。

Fig. 4 Time history of the estimated parameter values. These were estimated at every 0.1 second and are shown at every 2 seconds in this figure

(注 1) それぞれ m 系列信号を2次系フィルタ $\frac{2}{(S+1)(0.3S+1)}$ $\frac{1}{S^2+0.3S+2}$ に通したものの。

表から各入力信号に対し、各被験者とも、 T_L , K , C の値は γ によらずほぼ一定とみられる。これに対し T_A だけは γ の増加に従って著しく値が大きくなっている。このとき C が減少し、 K がそれを補う分だけ増加するとすれば非線形性は増すことはないが、推定結果は非線形性の強さがほとんど T_A の大きさのみによって決ることを示している。これは制御者の入力の観測量のうち u に関する成分 Cu の平均的な大きさ（すなわち、その標準偏差）が常に一定に保たれているといいかえることができる。これを式の形で表現すれば

$$u^* \doteq \text{constant} \cdot \frac{u + T_A \ddot{u}}{\sigma_u} \quad (33)$$

と表わされ、入力の観測量がその振幅の平均的な大きさを示す標準偏差 σ_u からの相対的变化量に近いものであることを示している。人間が外部の刺激から受ける感覚量は絶対的な刺激量によるのではなく、相対的な変化量に比例するとされ、それは Weber-Fechner の法則と呼ばれているが、(33)式も一種の Weber-Fechner の法則とみることも可能と思われる。

4. 結 論

最大傾斜法と Gauss-Newton 法を組合わせた方法を提案し、それが雑音が大であったり初期推定誤差が大である場合など、適度な安定性と収束速度とが要求される中間領域において有効であることを示した。その一つの応用例として制御者の予測動作にみられる非線形特性を取上げ、数式モデルのパラメータの推定に適用し、制御者の入力観測量が入力の平均的な大きさに対する相対的变化量であると思われることを示した。このことはさらに他の実験によっても確認することが望ましいと思われる。

終わりに、本研究を行なうにあたり終始変らずご指導いただいた東京大学工学部（現在、千葉大学工学部）磯部孝教授に深く感謝の意を表す。また日頃ご指導いただく東京大学宇宙航空研究所五十嵐寿一教授、石井泰助教授に謝意を表す。

参 考 文 献

- 1) H. J. Kelly: Methods of Gradients, in Optimization Techniques, Leitman, G. Ed., Academic Press (1962)
- 2) Nasser E. Nahi: Estimation Theory and Applications, John Wiley & Sons (1969)

- 3) G. A. Bekey and R. B. McGhee: Gradient Methods for the Optimization of Dynamic System Parameters, in Computing Methods in Optimization Problems, Academic Press (1964)
- 4) Donald W. Marquardt: An Algorithm for Least-Squares Estimation of Nonlinear Parameters, J. Soc. Indust. Appl. Math., 11-2 (1963)
- 5) 小畑：制御者の予測動作の特性とその適応性および最適性に関する研究，計測自動制御学会論文集，9-5，577/587 (1973)

《 付 録 》

2.5 節 証明

(8)式に関して行列 S_0 が対角化するように $\Delta x = P \Delta z$ なる座標軸の変換を行なうと

$$\begin{aligned} E(l) &= E_0(l) + \mathbf{e}_0^T P \Delta z + \frac{1}{2} \Delta z^T [P^T S_0 P] \Delta z \\ &\triangleq E_0(l) + \mathbf{r}_0^T \Delta z + \frac{1}{2} \Delta z^T \mathbf{T}_0 \Delta z \end{aligned} \quad (\text{A-1})$$

と書ける。ここに $\mathbf{T}_0$ は対角行列であり、その要素を $\lambda_{11}, \lambda_{22}, \dots, \lambda_{nn}$ とする。座標系 $\mathbf{z}$ において変形法を適用すれば、

$$\begin{aligned} \Delta \mathbf{z}^N &= \mathbf{v}_{GN} + \mathbf{v}_{SD} \\ &= \begin{bmatrix} 1 - (1 - \alpha \lambda_{11})^N & & 0 \\ & 1 - (1 - \alpha \lambda_{22})^N & \\ 0 & & \ddots \\ & & & 1 - (1 - \alpha \lambda_{nn})^N \end{bmatrix} \mathbf{v}_{GN} \end{aligned} \quad (\text{A-2})$$

となる。ただし

$$\mathbf{v}_{GN} = -\mathbf{T}_0^{-1} \mathbf{r}_0 \quad (\text{A-3})$$

である。これから $|1 - \alpha \lambda_{ii}| < 1$ のときには本法の性質(2), (3)の(ii), (iii)がただちに導かれる。また $\mathbf{v}_{GN}$ と $\mathbf{v}_{SD}$ とのなす角を ϕ とすると

$$\begin{aligned} \cos \phi &= \frac{\mathbf{v}_{GN}^T \mathbf{v}_{SD}}{\|\mathbf{v}_{GN}\| \cdot \|\mathbf{v}_{SD}\|} \\ &= \frac{-\sum_{i=1}^n (1 - \alpha \lambda_{ii})^N v_i^2}{\left(\sum_{j=1}^n v_j^2\right)^{1/2} \cdot \left(\sum_{k=1}^n (1 - \alpha \lambda_{kk})^{2N} v_k^2\right)^{1/2}} \end{aligned} \quad (\text{A-4})$$

となる。ただし $v_i (i=1, 2, \dots, n)$ は $\mathbf{v}_{GN}$ の各要素である。これから簡単な計算により

$$\frac{\partial}{\partial N} \cos \phi > 0 \quad (\text{A-5})$$

が導かれ、 $(\partial \phi / \partial N) > 0$ となる。(A-2)式から $\|\Delta \mathbf{z}_0^N\|$ が単調増加であることは明らかであり、性質(3)の(i)が求められる。