動的逆最適化による評価関数構造の決定

正員 章 宏 (九州工大)

正員 石川 眞 澄 (九州工大)

Structure Determination of a Criterion Function by Dynamic Inverse Optimization

Hong Zhang, Member, Masumi Ishikawa, Member (Kyushu Institute of Technology)

Dynamic inverse optimization generates a quadratic criterion function under which given input and output sequences become optimal for a known state equation model. In this paper we propose to use statistical test to decide which state variables are included in a criterion function. To evaluate the effectiveness of the proposed method, simulation experiments are carried out. Statistical test successfully determines which state variables are included in a criterion function.

キーワード：動的逆最適化，ネットワーク学習，統計的検定，評価関数構造

1. ま え が き

近年，逆問題に関する様々な研究が各分野で盛んに行われている⁽¹⁾。その主な理由の一つは，逆問題を解くことによりある事象間の因果法則を明らかにできることである。我々は，動的最適化問題⁽²⁾の逆問題を動的逆最適化問題と呼ぶ。動的逆最適化問題も逆問題の一種であり，既知の制御対象に対してある入出力時系列が与えられた場合，この入出力時系列を最適とする評価関数を求めることを目的とする。この問題解決に当たって，筆者らはネットワーク学習により動的逆最適化問題を解く方法を提案した⁽³⁾。この方法は，動的逆最適化問題を解くための最適性条件をニューラルネットワークにより表現し，この学習により，与えられた観測時系列に対応する評価関数を求めるものである。

動的逆最適化問題を解くことにより，観測時系列を動的最適化という観点から解釈できる。これにより状態変数の中のどの部分が評価に含まれているか，および各状態変数の相対的の重みを知ることが期待される。前者の例として例えば，腕のリーチング運動に関して，提案されている様々な仮説^{(4)・(6)}のうち，いずれが妥当であるかをデータに基づいて結論付けることが考えられる。

しかし，これまでの動的逆最適化問題の解法⁽³⁾においては，場合によっては最適性条件を満たさない解が得られること，ノイズのある場合には評価パラメータ値が零であるかどうかを決定できないという二つの問題点があった。前者の問題点に対しては，動的逆最適化問題を解く二段階学習を繰り返し緩和的に解く方法を提案する。後者は，ノイズを含む観測時系列に対してネットワーク学習によって得られた評価関数はノイズの影響を受けるため，真の評価パ

ラメータが零であっても，その推定値が零になることはなく，しかも評価パラメータ行列の対角成分が非負という制約のため，評価パラメータ値が零であるかどうかを決定することは一層困難である。

本論文では，ネットワーク学習で求めた評価関数にどの状態変数が含まれているかを，統計的検定によって決定する方法を提案する。ここでは帰無仮説として，ある状態変数が評価関数に含まれていないという仮説を考える。このような統計的検定を実行するためには，この帰無仮説の下では期待値が零になり，それ以外では非零になる統計量を採用することが必要である。

すぐに考え付くのは，当該の状態変数に対応する評価パラメータを統計量とすることである。真の評価パラメータが零の場合，この推定値がノイズのため正になったり負になったりすれば良いのであるが，実際には評価パラメータ行列が半正定であるという制約下で動的逆最適化問題を解いているので，帰無仮説の下でその期待値が零になるという条件を満足しない。従って評価パラメータを統計量とすることはできない。

本論文では，評価関数の当該評価パラメータに関する偏微分値を統計量とすることを提案する。ただし偏微分値は当該評価パラメータが零の場合で評価するものとする。真の評価パラメータが零の場合，この偏微分値は当然零となり，ノイズがある場合には，偏微分値は零を中心に正負にばらつく。また真の評価パラメータが正值の場合には，この偏微分値は負となる。従って，上記の統計量の持つべき必要条件を満足する。

この統計量は，ある変数の時間積分であるので，正規性を仮定して差し支えない。従って，幾つかの入出力系列に

基づいて t 値を求め、 t 検定により偏微分値の期待値が零であるという帰無仮説の統計的検定が可能になる。これは、伝達関数モデルやインパルス応答モデルの次数の妥当性を統計的に検定するという著者の一人の過去の研究で用いたアイデアと同一である⁽⁷⁾⁽⁸⁾。本論文では、この統計量を具体的に計算する方法も併せて提案する。

本提案の有効性を検証するために、簡単な制御系を対象としてノイズを含む幾つかの観測時系列に対して、それぞれの動的逆最適化問題を解き、その結果に基づいて統計的検定を行ない、評価関数の構造を決める実験を行なう。

本論文の構成は、まずは第2章で動的逆最適化問題を解く方法を必要最小限に述べる。次に、第3章で本稿に用いた統計的検定の手法を述べる。そして、第4章でノイズを含む観測時系列に対応する評価関数を推定し、統計的検定により評価関数構造の決定を行なった結果を示す。

2. 動的逆最適化問題を解く方法

ここで、ネットワーク学習による動的逆最適化問題を解く方法を必要最小限に述べる。

〈2・1〉 動的順・逆最適化に関する定式化 次の離散線形時不変系を考える。

$$\mathbf{x}_{k+1} = A_d \mathbf{x}_k + B_d \mathbf{u}_k \quad \dots\dots\dots (1)$$

$$\mathbf{y}_k = C_d \mathbf{x}_k \quad \dots\dots\dots (2)$$

ただし、 $A_d \in \mathbb{R}^{n \times n}$ 、 $B_d \in \mathbb{R}^{n \times m}$ と $C_d \in \mathbb{R}^{l \times n}$ はそれぞれ制御対象系の定数行列である。

上記の制御対象に対して、評価行列 $Q(\geq 0) \in \mathbb{R}^{n \times n}$ 、 $P_N(\geq 0) \in \mathbb{R}^{n \times n}$ と $R(> 0) \in \mathbb{R}^{m \times m}$ が与えられた場合、評価関数

$$J = \frac{1}{2} \mathbf{x}_N^T P_N \mathbf{x}_N + \frac{1}{2} \sum_{k=1}^{N-1} (\mathbf{x}_k^T Q \mathbf{x}_k + \mathbf{u}_k^T R \mathbf{u}_k) \quad \dots\dots (3)$$

を最小にする制御入力時系列 $\{\mathbf{u}_k\} (k=1, 2, \dots, N-1)$ を求めるのは動的最適化問題である。そこで(3)式に示すLQレギュレータ問題を解くには、Hamiltonian 関数

$$H_k = \frac{1}{2} (\mathbf{x}_k^T Q \mathbf{x}_k + \mathbf{u}_k^T R \mathbf{u}_k) + \mathbf{l}_{k+1}^T (A_d \mathbf{x}_k + B_d \mathbf{u}_k)$$

を用いる。上式における $\mathbf{l}_{k+1}$ は共状態ベクトルである。

そして、Hamiltonian 関数 H_k を $\mathbf{l}_{k+1}$ 、 $\mathbf{x}_k$ と $\mathbf{u}_k$ に関して偏微分すると、次の制御対象の状態方程式と共状態方程式および制御入力の停留条件が得られる。

$$\mathbf{x}_{k+1} = \frac{\partial H_k}{\partial \mathbf{l}_{k+1}} = A_d \mathbf{x}_k + B_d \mathbf{u}_k \quad \dots\dots\dots (4)$$

$$\mathbf{l}_k = \frac{\partial H_k}{\partial \mathbf{x}_k} = Q \mathbf{x}_k + A_d^T \mathbf{l}_{k+1} \quad \dots\dots\dots (5)$$

$$0 = \frac{\partial H_k}{\partial \mathbf{u}_k} = R \mathbf{u}_k + B_d^T \mathbf{l}_{k+1} \quad \dots\dots\dots (6)$$

初期状態ベクトル $\mathbf{x}_0$ が与えられた場合、状態ベクトル $\mathbf{x}_k$ と共状態ベクトル $\mathbf{l}_k$ の関係は二点境界値問題として関係付けられ、即ち、 $\mathbf{l}_k = P_k \mathbf{x}_k (k \leq N)$ が成り立つ。ただし、 $P_k \in \mathbb{R}^{n \times n}$ は時刻 k における対称正定行列であり、リカッチ方程式で定まる。そこで、時刻 $k \rightarrow \infty$ の定常状態の場合を考えると、リカッチ方程式は代数リカッチ方程式 (algebraic Riccati equation) に帰着され、次式で表わされる。

$$P = Q + A_d^T P A_d - A_d^T P B_d (R + B_d^T P B_d)^{-1} B_d^T P A_d \quad (7)$$

このとき、状態ベクトル $\mathbf{x}_k$ と共状態ベクトル $\mathbf{l}_k$ の対応関係は次式のようにになる。

$$\mathbf{l}_k = P \mathbf{x}_k; \quad k \leq N \quad \dots\dots\dots (8)$$

従って、制御対象系の定常状態フィードバック制御則が

$$\mathbf{u}_k = -(R + B_d^T P B_d)^{-1} B_d^T P A_d \mathbf{x}_k = -F_d \mathbf{x}_k \quad \dots\dots (9)$$

で与えられる。ただし、 $F_d \in \mathbb{R}^{m \times n}$ は定常状態フィードバックゲイン行列である。安定化制御のため、代数リカッチ方程式の解 $P > 0$ が満たされることが必要である⁽⁹⁾。

以上の定式化により、 $P > 0$ と(4)(5)(6)式が、動的逆最適化問題を解くための必要十分条件である。他方、動的逆最適化問題は制御対象の入出力時系列 $\{\mathbf{u}_k\}$ 、 $\{\mathbf{x}_k\}$ が与えられた場合、 $Q \geq 0$ 、 $R > 0$ の範囲内で評価関数 J を最小化する Q 、 R を推定する。ただ議論の簡単化のため、以下では、評価行列 R 、 Q を対角行列と仮定し、また R の評価重みの最大値を常に1となるよう Q の値を比例的に変化させるという正規化を行なう。

〈2・2〉 ネットワークの構造と学習 動的逆最適化問題を解くための最適性条件をネットワークの構築により表現するには、(8)式を(5)(6)式にそれぞれ代入し、共状態ベクトル $\mathbf{l}_k$ を含まない状態差分式に変換する必要がある。即ち、

$$(Q - P) \mathbf{x}_k = -A_d^T P \mathbf{x}_{k+1} \quad \dots\dots\dots (10)$$

$$R \mathbf{u}_k = -B_d^T P \mathbf{x}_{k+1} \quad \dots\dots\dots (11)$$

このとき、(4)式の状態方程式と(9)式の定常状態フィードバック制御則を同時に考えると、動的逆最適化問題を解く線形ネットワークの構造は図1のようになる⁽³⁾。

図1の各ブロックはネットワークの各状態層を、矢印の実線と破線はそれぞれの層間結合重みを表す。層間結合重み A_d 、 A_d^T 、 B_d 、 B_d^T と C_d は不変であり、層間結合重み R 、 Q 、 P と $-F_d$ は可変である。ネットワークの出力層は $\mathbf{e}_{k+1}^1$ 層と $\mathbf{e}_{k+1}^2$ 層からなる。(10)と(11)式の両辺が等しいことを、 $\mathbf{e}_{k+1}^1$ 層と $\mathbf{e}_{k+1}^2$ 層にそれぞれの両辺の差を出力させ、出力層 $\mathbf{e}_{k+1}^1$ と $\mathbf{e}_{k+1}^2$ の出力目標に零ベクトルを与えることにより表現している。

ネットワーク学習の手順は次の二段階からなる。第一段階において、初期状態ベクトル $\mathbf{x}_0$ と観測時系列 $\{\mathbf{y}_k\} \in \mathbb{R}^l$

($l \leq n$) を図1のネットワークの y_{k+1} 層に与えて、制御対象の閉ループモデルの状態フィードバックゲイン行列 F_d を学習により求める。この場合、観測時系列 $\{y_k\}$ はダイナミックなシステムの出力から得られるので、制御対象の閉ループモデル同定の学習には通時的誤差逆伝搬法⁽¹⁰⁾を用いる。第二段階において、閉ループモデルを用いて、零ベクトルを出力層の e_{k+1}^1 層と e_{k+1}^2 層に与えて、状態フィードバックゲイン行列 F_d に対応する評価行列 R , Q および行列 P を学習により求める。この場合の学習には誤差逆伝搬法⁽¹¹⁾のバッチモードを用いる。また学習時、評価行列 $Q \geq 0$, $R > 0$ と代数リカッチ方程式の解 $P > 0$ を保証するにはネットワーク学習と相性のよい固有値修正法を用いる⁽³⁾。

図1 動的逆最適化問題を解くネットワーク構造
Fig.1. The structure of a neural network for solving dynamic inverse optimization problems.

ただこの二段階の学習において、与えられたノイズを含む観測時系列に対して第一段階で求めた状態フィードバックゲイン行列 F_d を固定し、第二段階で評価行列 R , Q および行列 P を求めても、ネットワークの出力平均自乗誤差が十分に小さくならない場合がある。これは以下のような場合である。第一段階で F_d を求める場合には、行列 Q の大きさには無関係に、与えられた時系列だけから求まる。しかし第二段階で Q , R , P を求める場合、真の Q と推定時の $\hat{Q}$ の行列の次元が同じ場合には出力平均自乗誤差が十分小さくなるが、後者の次元の方が小さい場合には出力平均自乗誤差が十分小さくならない。これは本来 F_d は Q , R , P と密接な関連があるのに、これらと独立に F_d を求めたためである。言い換えれば、第一段階で求める F_d と第二段階で求める Q , R , P が整合的でないためである。

この問題を解決するには、次の二つの方法が考えられる。第一の方法は、学習プロセスを一段階にする方法である。この場合には全体としてリカレント学習を行なうことになるが、(7) 式の代数リカッチ方程式の存在のため誤差信号を伝播させることが困難である。第二の方法は、二段階の学習を繰り返し緩和的行なう方法である。この場合には、状態フィードバックゲイン行列 F_d と評価行列 R , Q との整合性を取るため、求めた評価行列 R , Q で状態フィードバックゲイン行列 F_d' を計算し、この F_d' に対応する評価行列 R' , Q' を学習により求める。このような手順を収束

するまで繰り返し行なう。本論文ではこの緩和的な学習法を用いる。

3. 統計的検定

ある状態変数が評価関数 J に含まれているか否かを統計的検定により決定することを提案する。この際、次の t 検定⁽¹²⁾を用いる。ある母集団からの観測値 $\{X_1, X_2, \dots, X_n\}$ が互いに独立で平均 μ , 分散 σ^2 の正規分布に従うとする。ここで、母集団の観測値の期待値が θ_0 であるという帰無仮説

$$H_0: \mu = \theta_0, \quad \sigma \text{ は未知}$$

を検定するとき

$$|t| > t_\alpha \quad \dots \dots \dots (12)$$

を棄却域とする。ここで、

$$t = \frac{\bar{\mu} - \theta_0}{\sqrt{\hat{\sigma}^2/n}} \quad \dots \dots \dots (13)$$

ただし $\bar{\mu} = \sum_i X_i/n$, $\hat{\sigma}^2 = \sum_i (X_i - \bar{\mu})^2/(n-1)$ であり、 t_α は自由度 $n-1$, 有意水準 α の棄却点である。(12) 式が成立すれば、帰無仮説を棄却できる。

評価関数 J に状態変数 x_i が含まれていない場合、評価関数の評価パラメータ q_{ii} に関する偏微分値の期待値が零である。ただし $q_{ii} = 0$ における偏微分を求めるものとする。従って帰無仮説は、次式で表わされる。

$$H_0: E \left[\frac{\partial J}{\partial q_{ii}} \bigg|_{q_{ii}=0} \right] = 0 \quad \dots \dots \dots (14)$$

(14) 式における統計量 $\frac{\partial J}{\partial q_{ii}} \big|_{q_{ii}=0}$ に関して、ネットワークにおいて学習時の誤差逆伝搬法の結合重み q_{ii} の修正量は次式で表わされる。

$$\Delta q_{ii} = -\eta \frac{\partial J}{\partial q_{ii}} \quad \dots \dots \dots (15)$$

ただし、 η は学習率である。

従ってここでの統計量 d_i は $-\Delta q_{ii}/\eta$ で与えられる。(11) 式の η は定数のため、 Δq_{ii} は偏微分値と比例的であり、 $\{\Delta q_{ii}\}$ により統計的検定を行なって差し支えない。もちろん学習時の出力平均自乗誤差が最小となる場所で Δq_{ii} の値を統計量として用いる。

なおネットワーク学習の際、評価行列 R , Q に対して正規化を行なった。正規化に伴い評価行列 R , Q を比例的に拡大・縮小すると、(3) 式より評価関数 J も同様に比例的に拡大・縮小する。また (15) 式の偏微分値も同じく比例的に拡大・縮小し、従って (13) 式で与えられる t 値は正規化による影響を全く受けない。このことより、評価行列 R , Q に対する正規化は評価関数の構造決定結果には何らの影響を及ぼさないことがわかる。

4. シミュレーション実験

シミュレーション実験においては、下記に示す制御対象と初期状態の値を用いる。

$$\begin{aligned} \mathbf{x}_{k+1} &= \begin{pmatrix} 0.9996 & 0.0204 \\ -0.0408 & 1.0404 \end{pmatrix} \mathbf{x}_k + \begin{pmatrix} 0.0020 \\ 0.2040 \end{pmatrix} \mathbf{u}_k, \\ \mathbf{y}_k &= (1 \ 0) \mathbf{x}_k, \quad \mathbf{x}_0 = (1 \ 0)^T. \end{aligned}$$

制御対象の開ループの行列 A_d の固有値の絶対値 ($\lambda_1 = 1.02 + 1.0204i$, $\lambda_2 = 1.02 - 1.0204i$) は 1 より大きいので、この制御対象は不安定系であり、また対 (A_d, B_d) と対 (C_d, A_d) のランクが共に 2 であるから、当該の制御対象は安定化可能である。

なお、ネットワーク学習の時、学習率 0.001~0.06、モーメンタム 0.5 であり、ネットワーク学習の目標パターンとする時系列の長さは 100 である。

シミュレーション実験に用いた時系列 $\{y_k^i\} (i = 1, \dots, 4)$ に対応した評価行列は以下の通りである。

$$\begin{aligned} \{y_k^1\} &\Leftrightarrow R_1 = 1, \quad Q_1 = \begin{pmatrix} 0.7 & 0 \\ 0 & 0 \end{pmatrix}, \\ \{y_k^2\} &\Leftrightarrow R_2 = 1, \quad Q_2 = \begin{pmatrix} 0.7 & 0 \\ 0 & 0.05 \end{pmatrix}, \\ \{y_k^3\} &\Leftrightarrow R_3 = 1, \quad Q_3 = \begin{pmatrix} 0.7 & 0 \\ 0 & 0.1 \end{pmatrix}, \\ \{y_k^4\} &\Leftrightarrow R_4 = 1, \quad Q_4 = \begin{pmatrix} 0.7 & 0 \\ 0 & 0.2 \end{pmatrix}. \end{aligned}$$

ここで、一例として時系列 $\{y_k^1\}$ の様子を図 2 に示す。

図 2 実験用の時系列 $\{y_k^1\}$
Fig.2. Output sequence $\{y_k^1\}$ for experiment.

以下のシミュレーション実験において、正規性のノイズ $N(0, 0.1^2)$ を時系列 $\{y_k^i\} (i = 1, \dots, 4)$ に加えたものが与えられる目標パターンとする。一例として時系列 $\{y_k^1\}$ にノイズを加えた観測時系列の様子を図 3 に示す。このような観測時系列を 10 通り用意する。一般に正規性ノイズは、

図 3 時系列 $\{y_k^1\}$ にノイズ $N(0, 0.1^2)$ を加えた一例
Fig.3. An example of output sequence $\{y_k^1\}$ with observation noise, $N(0, 0.1^2)$.

$[-\infty, \infty]$ の値を取り得るが、平均値から標準偏差の 2~3 倍離れると、確率密度が急激に減少するため、ノイズが異常に大きくる確率は極めて小さい。図 3 もこのことを示している。

〈4・1〉 学習の結果 ここで、時系列 $\{y_k^1\}$ にノイズを加えた 10 個の観測時系列のそれぞれに対してネットワーク学習を行ない、求まった評価行列 $\hat{Q}_1$ の結果を表 1 に示す。

表 1 時系列 $\{y_k^1\}$ にノイズ $N(0, 0.1^2)$ のある場合の推定された評価行列 $\hat{Q}_1$ の結果. ($q_{22}=0$)

Table 1. A result of the estimated criterion matrix $\hat{Q}_1$ when observation noise, $N(0, 0.1^2)$, is added to output sequence $\{y_k^1\}$. ($q_{22}=0$).

Parameter	Max.	Min.	Mean	S.D.
$\hat{q}_{11}$	1.0035	0.6346	0.7675	0.1176
$\hat{q}_{22}$	0.0622	0.0000	0.0165	0.0172

表 1 より、評価関数の評価パラメータはノイズの分散の影響を受けて、大きく変動することがわかる。この場合、推定された評価パラメータ q_{22} の平均値、標準偏差を見ると、状態変数 x_2 が評価に含まれているかどうかははっきりと断定できない。時系列 $\{y_k^2\}$, $\{y_k^3\}$ と $\{y_k^4\}$ にノイズを加えたすべての観測時系列に対してネットワーク学習を行なった場合も同様である。

〈4・2〉 t 検定 ここで、前節で述べた手順に基づいて、時系列 $\{y_k^i\} (i = 1, \dots, 4)$ にノイズを加えた観測時系列に対してネットワーク学習を行ない、「状態変数 x_2 が評価に含まれていない」という帰無仮説 H_0 に対して統計的検定を行なう。

例として時系列 $\{y_k^1\}$ にノイズ $N(0, 0.1^2)$ を加えたすべての観測時系列に対して学習の結果を示す。この場合、得られた $\{\Delta q_{22}\}$ の値を表 2 に示す。

ここで、 $\{\Delta q_{22}\}$ の値を求めた場合、緩和的学習は一回で収束した。時系列 $\{y_k^1\}$ にノイズ $N(0, 0.1^2)$ を加えた観測時系列に対して t 検定を行なった結果を表 3 に示す。

表 3 より、真の評価パラメータ $q_{22}=0$ の場合に限って、

表2 時系列 $\{y_k^1\}$ にノイズ $N(0, 0.1^2)$ のある場合の $\{\Delta q_{22}\}$ の値

Table 2. The value of $\{\Delta q_{22}\}$ to output sequence $\{y_k^1\}$ with observation noise, $N(0, 0.1^2)$. ($q_{22}=0$).

No.	Δq_{22}	No.	Δq_{22}
1	0.0107768	6	-0.0355057
2	0.0198349	7	-0.0954470
3	0.0381521	8	0.1907210
4	0.0979111	9	0.0654002
5	-0.0852062	10	-0.0007726

表3 時系列 $\{y_k^i\} (i=1, \dots, 4)$ にノイズ $N(0, 0.1^2)$ のある場合の t 検定の結果。
 $t_{0.05}=2.262$.

Table 3. A result of the t test for output sequences $\{y_k^i\} (i=1, \dots, 4)$ with observation noise, $N(0, 0.1^2)$. $t_{0.05}=2.262$.

q_{22}	t	$ t > t_\alpha?$	Decision
0	0.763	No	accept H_o
0.05	3.203	Yes	reject H_o
0.1	15.45	Yes	reject H_o
0.2	9.939	Yes	reject H_o

仮説 $H_o: \partial J / \partial q_{22}|_{q_{22}=0} = 0$ は棄却されない。これより、結論として状態変数 x_2 が評価関数 J に含まれていないという仮説を棄却できない。他の真の評価パラメータ $q_{22} \neq 0$ の場合については、 $H_o: \partial J / \partial q_{22}|_{q_{22}=0} = 0$ は棄却される。即ち、状態変数 x_2 が評価関数 J に含まれていないという仮説を棄却できる。これは妥当な結論である。

〈4・3〉 符号検定 t 検定は、統計量が正規分布に従う場合に適用可能である。ここでの統計量は時間積分により得られるので正規性を仮定して差し支えないと考えられる。しかし、たとえ正規性を仮定しなくてもノンパラメトリック検定の一つである符号検定 (sign test) ⁽¹³⁾ を適用することが可能である。

ある母集団における $\{X_1, X_2, \dots, X_n\}$ があり、しかも X_i の分布は任意でよい。 X_i が正であるサンプルの個数を n とする。この統計量 n は次の 2 項分布に従う。

$$P(n) = {}_N C_n \left(\frac{1}{2}\right)^N \dots\dots\dots (16)$$

そこで、次式が成立すれば、帰無仮説を棄却できる。

$$\left| n - \frac{N}{2} \right| > n_\alpha \dots\dots\dots (17)$$

ただし、 n_α は次式によって定められる。

$$\sum_{n=0}^{[\frac{N}{2}-n_\alpha]} {}_N C_n \left(\frac{1}{2}\right)^N \leq \frac{\alpha}{2} < \sum_{n=0}^{[\frac{N}{2}-n_\alpha]+1} {}_N C_n \left(\frac{1}{2}\right)^N$$

上式における $[\]$ はガウス符号であり、即ち、 $[x]$ が x を超えない最大の整数を表す。

有意水準 $\alpha=0.05$ 、観測値の個数 $N=10$ の時、符号検定の仮説の棄却域は $n_\alpha=\{0, 1, 9, 10\}$ である。表3と同じケースについて、符号検定を行なった結果を表4に示す。

表4により、真の評価パラメータ $q_{22}=0.1$ と $q_{22}=0.2$ の場合に対して、仮説 $H_o: \partial J / \partial q_{22}|_{q_{22}=0} = 0$ は棄却される。これより、結論として状態変数 x_2 が評価関数 J に含まれていないという仮説を棄却できる。一方、真の評価パラメータ $q_{22}=0$ と $q_{22}=0.05$ の場合に対して、仮説 $H_o: \partial J / \partial q_{22}|_{q_{22}=0} = 0$ は棄却されない。これより、結論として状態変数 x_2 が評価関数 J に含まれていないという仮説を棄却できない。真の評価パラメータ $q_{22}=0.05$ の場合において、 t 検定と符号検定の結果は異なっている。これは、正規性を仮定していないため符号検定の方が t 検定よりも若干検出力が小さいためである。

表4 時系列 $\{y_k^i\} (i=1, \dots, 4)$ にノイズ $N(0, 0.1^2)$ のある場合の符号検定の結果

Table 4. A result of the sign test for output sequences $\{y_k^i\} (i=1, \dots, 4)$ with observation noise, $N(0, 0.1^2)$.

q_{22}	n	$n - \frac{N}{2}$	$ n - \frac{N}{2} > n_\alpha?$	Decision
0	6	No		accept H_o
0.05	8	No		accept H_o
0.1	10	Yes		reject H_o
0.2	10	Yes		reject H_o

以上よりノイズを含む観測時系列に対応する評価関数構造の決定を行なうことに成功した。

5. ま と め

本論文では、ネットワーク学習による動的逆最適化問題の解法を用いて、ノイズを含む観測時系列に対応する評価関数を求め、評価関数にどの状態変数が含まれているかを統計的検定の手法により決定する方法を提案した。

本提案の有効性を検証するため、簡単な制御対象を用いてネットワーク学習により、ノイズを含む様々な観測時系列に対応する評価関数の構造を決定した。評価関数のある評価パラメータに関する偏微分値を統計量として求め、 t 検定を用いて評価関数に対応する状態変数が含まれているかどうかを決定することに成功した。また同様な統計量に対して符号検定も行なった。両者の検定結果には大差がない。これらのことにより、与えられたノイズを含む観測時系列に対応する評価関数構造を決定でき、本提案の有効性が確認された。

統計的検定においては、帰無仮説が真であるにも関わらずこれを棄却する第一種の過りを一定として、帰無仮説が誤りであるにも関わらずこれを採択する第二種の過りを最小にする方法を採用している。第二種の過りを1から引いた確率、すなわち帰無仮説が誤りである際にこれを棄却する確率を検出力と呼んでいる。ノイズの分散が大きくなれば、(13) 式の分子があまり変化しないのに分母だけが大きくなるため、 t 値が小さくなる。従って帰無仮説が誤りである場合に、これを棄却し難くなる。即ち、ノイズの分散が大きくなると、評価関数の構造決定の際の検出力が小さくなる。これに関する詳細は別の機会に発表したい。

本提案手法は、システムの次元数が大きくなった場合にも原理的には同様に適用可能である。ただ、大規模な動的逆最適化問題（例えば次元数 $n=100$ ）を解く場合には、ネットワーク学習のローカルミニマム問題が深刻になり、また評価行列 Q の半正定性と行列 P の正定性の確保も困難となる。従って現実には大規模な動的逆最適化問題を解くのは困難である。ある程度の規模の動的逆最適化問題（たとえ次元数 $n=10$ ）であれば、その困難さの程度はデータに依存するものの、解くことが可能であると考えている。大規模化の問題についてはより詳細な検討が必要であり、今後の課題として残されている。

本提案手法の腕のリーチング運動への適用について簡単に述べる。この問題は脳の運動中枢が水平面内の2点間到達運動において、可能な無数の軌道の中から一つの最適な軌道を決めていることを検証するものである。腕のキネマティックな特徴を解釈するための様々な評価関数の一つであるトルク変化最小モデル C_τ ⁽⁵⁾ は下記の通りである。

$$C_\tau = \frac{1}{2} \int_0^{t_f} \sum_{i=1}^m \left(\frac{d\tau_i}{dt} \right)^2 dt \dots\dots\dots (18)$$

ただし、 t_f は運動時間、 τ_i は m 個の関節のうち、 i 番目の関節で発生するトルクである。トルクが力 $\times$ 距離で表われ、また力は加速度 $\times$ 質量で表わされる。従って評価関数 C_τ は座標 x の3次微分の自乗和で表わされる。評価行列 Q のうち、 x の3次微分に対応する箇所の重みが零であるかどうかを本提案手法で検定することが考えられる。

同一の軌道データを用いて他の評価関数候補に対しても同様の検定が行なえるので、脳の運動中枢が水平面内の2点間到達運動において、どの評価関数が妥当であるかの議論の結着に期待できる。ただ現実には例えば(18)式に示すように評価関数には座標 x の3次微分以外の項が含まれていないので、このまま本提案手法を適用することは無理である。実際の適用に際しては何らかの工夫が必要とされよう。これは今後の課題として残されている。

（平成12年4月3日受付，同12年7月3日再受付）

文 献

- (1) チャールズ W. グロエッチ著，金子 晃，山本昌宏，滝口孝志：数理科学における逆問題，サイエンス社，1996。
- (2) F. L. Lewis: *Optimal Control*, John Wiley & Sons, Inc., 1986.
- (3) 章 宏，石川眞澄：“ニューラルネットワーク学習による動的逆最適化問題の解法”，電気学会論文誌，**120-C**, 4, 481~488 (2000)
- (4) W. Abend, E. Bizzi and P. Morasso: “Human arm trajectory formation,” *Brain*, **105**, 331~348 (1982)
- (5) Y. Uno, M. Kawato and R. Suzuki: “Formation and control of optimal trajectory in human multijoint arm movement: Minimum torque-change model,” *Biological Cybernetics*, **61**, 89~101 (1989)
- (6) 川人光男：脳の計算理論，産業図書，1996。
- (7) 茅 陽一：“伝達関数モデルの妥当性の検定”，計測自動制御学会論文集，**5**, 4, 54~63 (1969)

- (8) 茅 陽一，石川眞澄：“インパルス応答モデルの妥当性の検定”，計測自動制御学会論文集，**6**, 4, 49~57 (1970)
- (9) 兒玉慎三，須田信英：システム制御のためのマトリクス理論，計測自動制御学会，1978。
- (10) R. J. Williams and D. Zipser: “Gradient-Based Learning Algorithms for Recurrent Networks and Their Computational Complexity,” Y. Chauvin and D. E. Rumelhart Eds., *Backpropagation: Theory, Architectures and Applications*, LEA, Inc, 1995.
- (11) D. E. Rumelhart, et al: “Parallel Distributed Processing,” The MIT Press, 1986.
- (12) 矢野健太郎監修，宮本敏雄訳編：数学ハンドブック，森北出版，1985。
- (13) 竹内 啓編集委員代表：統計学辞典，東洋経済新報社，1997。

章 宏（正員）1957年12月10日生。1991年九州大学大学院工学研究科修士課程修了。同年九州工業大学情報工学部助手，現在に至る。主としてニューロコンピューティング，逆最適化問題等に関する研究に従事。

石川 眞澄（正員）1947年3月8日生。69年東京大学工学部電気工学科卒業。74年東京大学大学院工学系研究科博士課程修了。工学博士。同年通産省電子技術総合研究所入所。同所情報科学部主任研究官。86-87年カリフォルニア大学サンディエゴ校客員研究員。同所情報科学部主任研究官。86-87年カリフォルニア大学サンディエゴ校客員研究員。90年より九州工業大学情報工学部制御システム工学科教授。2000年より同大学大学院生命体工学研究科長。ニューロコンピューティングの研究に従事。81年電気学会論文賞。日本神経回路学会，電子情報通信学会，日本認知科学会，IEEEなどの会員。