推定 Hessian を利用したノンパラメトリック PWA モデルの構築法

藤 本 悠 介*・丸 田 一 郎*・杉 江 俊 治*

Construction of Non-parametric PWA Models with Estimated Hessian

Yusuke FUJIMOTO*, Ichiro MARUTA* and Toshiharu SUGIE*

This paper is concerned with construction of non-parametric PWA (Piecewise Affine) models. Though their effectiveness is demonstrated in the literature, the existing identification method suffers from heavy computation load. This load is mainly due to the large data set which is required for representing the model. This paper proposes a method to construct non-parametric PWA models represented with less data points based on estimated Hessian.

Key Words: nonlinear system identification, PWA model, non-parametric model

1. はじめに

近年、制御対象の複雑化や、高度な制御への要求から、非線形モデリングの重要性がますます高くなってきている。特に、入出力データからモデルを構築する非線形システム同定は多くの注目を集めており、これまでにさまざまなモデルを用いた同定法が提案されてきた¹⁾。それらのモデルの中でも、区分的アフィン (Piecewise Affine: 以下 PWA) モデルは、表現できる関数の広さとその単純さから制御のためのモデルとして有用であり、その同定法についても数多くの研究がなされている (たとえば文献 2))。中でも丸田らは、多数のデータ点からなるノンパラメトリックな表現を用いた PWA モデルによる同定法を提案している³⁾。

このモデルでは、多数のデータ点とそれらの線形内挿から対象の入出力関係を表現する (Fig. 1 参照)。このモデルは、ノンパラメトリックな写像の表現に基づいたものであるために自由度が高く、実データに対しても有効であることが確認されている^{3), 4)}。一方、この同定法は計算負荷が大きく、状態もしくは入力の次元が大きいシステムには適用しにくいという問題点がある。この計算負荷は、主としてモデル表現に要するデータ点の多さに起因する。そこで本稿では、観測データから推定した Hessian からある種の重要度を定め、効率的にモデル表現のためのデータ点を配置し、モデルの精度を劣化させることなく計算負荷を大幅に軽減させた新しい同定法を提案する。

なお、著者らの先行研究と本研究との違いについて触れて

Fig. 1 Illustration of non-parametric PWA model

おく。著者らはこれまでに、ノンパラメトリック PWA モデルの表現に要するデータ点を削減することを目的とした研究を行なっている。このような削減を行なえば、実際にモデルを利用する際に、計算時間を短縮し、かつメモリへの負荷を低減することができる。この目的を達成するために、著者らは数千以上に上る多数の変数をもつ ℓ_1 最適化問題を複数回解くという手法を提案した⁵⁾。この手法は一定の有効性を示したが、モデルを構築する際の計算負荷が大きく、計算時間が膨大になりがちであるという問題点を有していた。一方本研究では、点配置を効率的に行なうことで、同定の際の計算負荷を小さくすることを目的とする。結果的に得られるモデルは先行研究と同様に少数のデータ点で表現されるものとなるが、さらにこのモデルを高速に構築しようとするのが本研究である。

本稿の構成は以下のとおりである。まず 2 章で、データ集合に基づいた PWA 写像によるモデル表現について、簡単に説明する。続いて 3 章で本稿で考察する問題を設定し、4 章で提案法について説明する。その後 5 章で数値例を、6 章で実験例を示し、提案法の有効性を検証する。

〔記法〕 $\mathbb{R}$ で実数の集合を表わす。集合 $\{d^1, \dots, d^N\}$ を

* 京都大学大学院情報学研究科 京都市左京区吉田本町

* Graduate School of Informatics, Kyoto University, Yoshida-Honmachi, Sakyo-ku, Kyoto

(Received June 26, 2014)

(Revised September 22, 2014)

$\{d^i\}_{i=1}^N$ で表わす. ベクトル $\mathbf{a}$ に対し, $\mathbf{a}^T$ で転置を表わす. ベクトル $\mathbf{a}, \mathbf{b}$ に対し, $\mathbf{a} \leq \mathbf{b}$ でベクトルの各要素に対して不等号が成立していることを示す. $\|\cdot\|$ でベクトルおよび行列の 2 ノルムを表わす. 2 階連続微分可能な関数 $f(\mathbf{x})$ に対し, $\frac{\partial^2 f}{\partial \mathbf{x}^2}$ でその Hessian を表わす. P をユークリッド空間上の点の集合とすると, $\text{Co}(P)$ を P の凸包, $\mathcal{T}(P)$ を P の Delaunay 単体分割とする.

2. ノンパラメトリック PWA モデルの表現法

この章では, 丸田らが提案したデータ集合に基づいた PWA 写像によるモデル表現法³⁾について, 簡単に説明する.

写像の定義域を $\mathbb{R}^d$, 値域を $\mathbb{R}$ とする. 定義域上の入力 $\mathbf{x}^i \in \mathbb{R}^d$ および値域上の出力 $y^i \in \mathbb{R}$ からなるデータ対 $\begin{pmatrix} \mathbf{x}^i \\ y^i \end{pmatrix}$

を N 個含むデータ集合を, $D = \left\{ \begin{pmatrix} \mathbf{x}^i \\ y^i \end{pmatrix} \right\}_{i=1}^N$ とおく. このとき, データ集合 D に基づく PWA 写像 $\Pi_D: \mathbb{R}^d \mapsto \mathbb{R}$ を, 以下の手続きにより値 $\Pi_D(\mathbf{x})$ が得られる写像とする.

(1) Delaunay 単体分割 $\mathcal{T}(\{\mathbf{x}^i\}_{i=1}^N)$ から, $\mathbf{x}$ を含む

Delaunay 単体 $V = \{\mathbf{x}^{v^i}\}_{i=1}^{d+1}$ を抽出する.

(2) 抽出した単体上のデータ $\{\mathbf{x}^{v^i}\}_{i=1}^{d+1}$ から, 線形補間により $\Pi_D(\mathbf{x})$ の値を定める.

定義域が一次元の場合, この写像は Fig. 1 に示したような折れ線のグラフとなる. また, 二次元の場合は, Fig. 2 に示すように三角形の面の集合により曲面を近似したものとなる. 上の手続きは, これらの概念を一般の次元へ拡張したものとなっている. Delaunay 単体分割については付録 A を参照されたい.

以降では, $\mathbf{x} \in \mathbb{R}^d$, 関数 $f(\mathbf{x})$ を $f: \mathbb{R}^d \mapsto \mathbb{R}$ として, こ

Fig. 2 Illustration of data-based PWA map ($d=2$)

の関数をデータ集合 $\hat{D} = \left\{ \begin{pmatrix} \hat{\mathbf{x}}^i \\ \hat{y}^i \end{pmatrix} \right\}_{i=1}^{\hat{N}}$ に基づく PWA 写像 $\Pi_{\hat{D}}: \mathbb{R}^d \mapsto \mathbb{R}$ で表わすことを考える. ただし, $\hat{\mathbf{x}}^i \in \mathbb{R}^d, \hat{y}^i \in \mathbb{R}$ である.

3. 問題設定

つぎに, 本稿における同定問題の設定について説明する.

今, つぎの式に従って生成される y_0^i と $\mathbf{x}_0^i$ から, 観測データ集合 $D_0 = \left\{ \begin{pmatrix} \mathbf{x}_0^i \\ y_0^i \end{pmatrix} \right\}_{i=1}^{N_0}$ が得られたとする.

$$y_0^i = f(\mathbf{x}_0^i) + \eta^i \quad (1)$$

ここで, f はシステムの入出力特性を記述する非線形関数であり, η^i は観測ノイズである.

注意 1. 動的システムであっても, (1) 式の形式で挙動を記述できる. 非線形一入力一出力システムの入力を u , 出力を z とすると, このシステムの動特性は以下の非線形微分方程式で記述される.

$$\dot{z}^{(n)} = f(z, \dot{z}, \dots, z^{(n-1)}, u, \dot{u}, \dots, u^{(m)}) \quad (2)$$

このとき, $\mathbf{x} = (z, \dot{z}, \dots, z^{(n-1)}, u, \dot{u}, \dots, u^{(m)})^T, y = z^{(n)}$ とすることで, 動的システムを (1) 式の形式に帰着させることができる. $d = n + m$ となっていることに注意されたい. 同様に, (1) の形式で状態空間表現を行なうこともできる. m 入力 n 状態の非線形動的システムを考える. このシステムの状態を $\mathbf{z} \in \mathbb{R}^n$, 入力を $\mathbf{u} \in \mathbb{R}^m$ とすると, $\mathbf{x} = \begin{pmatrix} \mathbf{z} \\ \mathbf{u} \end{pmatrix}$ とすることで, (1) の形式に帰着させられる. この場合, $\mathbf{z} = (z_1, z_2, \dots, z_n)^T$ として, $\dot{z}_i = f_i(\mathbf{x})$ となる関数 $f_i: \mathbb{R}^d \mapsto \mathbb{R}$ を n 種類用意すればよい. $\square$

注意 2. 動的システムの場合は特に, $\mathbf{x}_0^i$ をノイズなく観測することは難しい. しかしそのような場合でも, (1) の形で観測を行なっていると考えることは可能である. すなわち, η^i が $\mathbf{x}_0^i$ へのノイズの影響も含んでいるとみなすことができる. $\square$

ここでの目標は, 観測データ集合 D_0 から, ノンパラメトリック PWA モデル $\Pi_{\hat{D}}$ が f をよく近似するようにデータ集合 $\hat{D} = \left\{ \begin{pmatrix} \hat{\mathbf{x}}^i \\ \hat{y}^i \end{pmatrix} \right\}_{i=1}^{\hat{N}}$ を構成することである. ただしこのとき, $\hat{N}$ はなるべく小さい値となることが望ましい. また, 構築するモデル $\Pi_{\hat{D}}$ の定義域は有界な凸多面体 Ω であるとし, その頂点集合が $\{\mathbf{x}_\Omega^i\}_{i=1}^{N_\Omega}$ で与えられているものとする. したがって具体的に考察する問題は, Ω における $\hat{\mathbf{x}}^i$ の効率的な配置と, 対応する $\hat{y}^i$ の推定である.

$\hat{\mathbf{x}}^i$ の配置が最終的なモデルの精度に与える影響について, 図を用いて簡単に補足する. Fig. 3(a) および Fig. 3(b) は,

Fig. 3 Influence of distribution of $\hat{\mathbf{x}}^i$ ($d = 1, \hat{N} = 5$)

$d = 1, \hat{N} = 5$ の場合についてデータ点の配置の与える影響を表わした模式図である。特に Fig. 3(a) では、データ点を x 方向で等間隔となるよう配置している。この場合、各データ点の y 座標をどのように定めても、Fig. 3(a) の配置では Fig. 3(b) のモデルほどの精度は得られない。このように、 $f(\mathbf{x})$ をデータ集合に基づいた PWA 写像 $\Pi_{\hat{D}}$ により表わす場合、データ点の配置が最終的なモデルの精度に大きな影響を与える。

4. 推定 Hessian に基づく点配置および同定法

4 章では、3 章で設定した問題に対し、効率的な $\hat{\mathbf{x}}^i$ の配置を利用した同定法を提案する。提案法では、以下の二つのステップでモデルを構築する。

Step 1 関数 f の Hessian を推定し、それを用いて $\{\hat{\mathbf{x}}^i\}_{i=1}^{\hat{N}}$ を構築する。

Step 2 各 $\hat{\mathbf{x}}^i$ に対応する $\hat{y}^i$ を、以下の最小二乗問題を解くことにより決定する。

$$\text{minimize} \quad \sum_{i=1}^{N_0} (y_0^i - \Pi_D(\mathbf{x}_0^i))^2 \quad (3)$$

Step 1 については、4.1 節で詳細に述べる。提案法の特徴は、この推定した Hessian に基づいた点配置と、Step 2 における最小二乗法の利用である。

本章の構成は以下のとおりである。まず 4.1 節で、Hessian に基づいた点配置アルゴリズムを提案する。その後、4.2 節で関連した定理とともにアルゴリズムの妥当性について説明し、4.3 節で最小二乗問題 (3) について補足する。最後に 4.4 節で計算量について述べる。

4.1 推定 Hessian に基づく点配置アルゴリズム

この節では、 $\hat{\mathbf{x}}^i$ を効率的に配置するアルゴリズムを提案する。このアルゴリズムではまず、一様分布に基づいて多数の点を凸多面体 Ω 上に散布し、各点における重要度に対応する値を計算する。そしてこの重要度が大きい順に点を選択していき、 $\{\hat{\mathbf{x}}^i\}_{i=1}^{\hat{N}}$ とする。ただしこの重要度は、点を選択するごとに更新される。

本アルゴリズムでは、与えられた点 $\mathbf{x}_q$ での $\left\| \frac{\partial^2 f}{\partial \mathbf{x}^2} \right\|$ を利用して、重要度を計算する。観測データからこの値を推定する方法について、本稿では以下に述べるような単純な二次形式フィッティングを用いる。なおこの方法は、Just-in-Time モデリングなどで用いられているものと同一である⁶⁾。

今、観測データ集合 $D_0 = \left\{ \begin{pmatrix} \mathbf{x}_0^i \\ y_0^i \end{pmatrix} \right\}_{i=1}^{N_0}$ と要求点 $\mathbf{x}_q \in \mathbb{R}^d$ が与えられたとする。このとき $\mathbf{x}_q$ から最も近い k 点を $\{\mathbf{x}_0^i\}_{i=1}^{N_0}$ から選択し、そのインデックスを $i_1, \dots, i_k$ とする。これらの k -最近傍点に対し、つぎの評価関数 (4) が最小となるように A_q, B_q, c_q を設定する。ただし $A_q \in \mathbb{R}^{d \times d}, B_q \in \mathbb{R}^{1 \times d}, c_q \in \mathbb{R}$ であり、 A_q は対称行列であるとする。

$$\sum_{m=1}^k \left\{ y_0^{i_m} - \left(\frac{1}{2} (\mathbf{x}_0^{i_m} - \mathbf{x}_q)^T A_q (\mathbf{x}_0^{i_m} - \mathbf{x}_q) + B_q (\mathbf{x}_0^{i_m} - \mathbf{x}_q) + c_q \right) \right\}^2 \quad (4)$$

これは $\mathbf{x}_q$ の近傍で、二次の Taylor 展開により f を近似していることに対応する。また、 k は平滑化の度合いを決めるパラメータであり、 $k \geq \frac{1}{2}(d+1)(d+2)$ のとき (4) の最小化が解けるようになる。このとき、 $\|A_q\|$ が $\mathbf{x}_q$ における $\left\| \frac{\partial^2 f}{\partial \mathbf{x}^2} \right\|$ を近似していると考えられることができる。表記の簡単のため、以降では与えられた $D_0, \mathbf{x}_q, k$ に対応する $\|A_q\|$ を返す関数を、 $\text{EstHesse}(D_0, \mathbf{x}_q, k)$ と定義する。

注意 3. ここでは、二次近似によって Hessian を推定する方法を示した。二次近似を用いる利点は、より高次の近似に比べ高速であることと、明快な計算で Hessian を推定できることである。ただし Algorithm 1 において、二次近似以外の手法で Hessian を推定しても、アルゴリズムに本質的な違いはないことに注意されたい。□

以上を踏まえたうえで、 $\{\hat{\mathbf{x}}^i\}_{i=1}^{\hat{N}}$ を構築する点配置アルゴリズムを提案する。アルゴリズムにおけるパラメータは、初期散布点の数 $\hat{N}_0$ 、最終的に得られるデータ集合中の点の数 $\hat{N}$ 、そして Hessian 推定に用いる平滑化パラメータ k の三つである。 D_0, Ω は与えられていると仮定していることに注意されたい。また、アルゴリズム中において、 r_i は $\hat{\mathbf{x}}_0^i$ における Hessian の推定値を示し、 l_{NN}^i は、各初期散布点 $\hat{\mathbf{x}}_0^i$ から最も近い選択済みの点までの距離を示す。

このアルゴリズムでは、初期散布点集合 $\{\hat{\mathbf{x}}_0^i\}_{i=1}^{\hat{N}_0}$ の中から、各グループにおいて評価関数 $r_i \cdot (l_{\text{NN}}^i)^2$ を最大とする $\hat{\mathbf{x}}_0^i$ を選択していく。また、 $\{\hat{\mathbf{x}}^i\}_{i=1}^{\hat{N}_0}$ を Ω の頂点集合としたの

Algorithm 1 Construct $\{\hat{\mathbf{x}}^i\}_{i=1}^{\hat{N}}$

Input : $D_0, \Omega, \hat{N}_0, \hat{N}, k$
Output : $\{\hat{\mathbf{x}}^i\}_{i=1}^{\hat{N}}$

Sample $\{\hat{\mathbf{x}}_0^i\}_{i=1}^{\hat{N}_0}$ from uniformly distribution over Ω
 $r_i \leftarrow \text{EstHesse}(D_0, \hat{\mathbf{x}}_0^i, k) \quad (i = 1, 2, \dots, \hat{N}_0)$
 $\{\hat{\mathbf{x}}^i\}_{i=1}^{\hat{N}_\Omega} \leftarrow \{\hat{\mathbf{x}}_0^i\}_{i=1}^{\hat{N}_\Omega}$
for $m = 1 + N_\Omega$ **to** $\hat{N}$ **do**
 $l_{\text{NN}}^i \leftarrow \min_{j \in \{1, \dots, m-1\}} \|\hat{\mathbf{x}}_0^i - \hat{\mathbf{x}}^j\| \quad (i = 1, 2, \dots, \hat{N}_0)$
 $i_{\text{next}} \leftarrow \text{argmax}_i r_i \cdot (l_{\text{NN}}^i)^2$
 $\hat{\mathbf{x}}^m \leftarrow \hat{\mathbf{x}}_0^{i_{\text{next}}}$
end for

は、最終的な出力である $\{\hat{\mathbf{x}}^i\}_{i=1}^{\hat{N}}$ が確かに Ω を覆うようにするためである。

このアルゴリズムにおいて、 $r_i \cdot (l_{\text{NN}}^i)^2$ はループごとに更新される重要度としての役割を果たす。これにより選択される点は、その段階で最も誤差を生じうる点を近似したものである。このような点を追加することで、 $|f(\mathbf{x}) - \Pi_D(\mathbf{x})|$ の最大値を効率よく小さくしていくことができる。また、停止条件について、ここでは選択する点の数 $\hat{N}$ で規定したが、 $r_i \cdot (l_{\text{NN}}^i)^2$ の最大値を基準とする方法も考えられる。これらの詳細については 4.2 節で述べる。

ここで、ノイズ η^i がこのアルゴリズムに与える影響について補足しておく。ノイズの影響は、Hessian 推定値 r_i のみに現れる点に注意されたい。したがって、Hessian を適切に推定することでノイズの影響を小さくできる。ここで示したような k -最近傍での二次形式フィッティングにより Hessian を推定する場合は、 k の適切な選択が、ノイズ η^i の影響軽減に重要な役割を果たす。一般に、 k を大きくするとノイズの影響は軽減されるが、過度に大きくすると二次近似の妥当性が失われてしまう。 k の選択指標の一つとして、たとえば文献 6) では Final Prediction Error (FPE) ⁷⁾ を用いている。 $D_0, \mathbf{x}_q, k$ を固定したとき、 V_k を評価関数 (4) の最小値とする。また、 $M = \frac{1}{2}(d+1)(d+2)$ とするとき、以下の評価関数を考える。

$$FPE(k) = V_k \frac{1 + M/k}{1 - M/k} \quad (5)$$

このような評価関数を最小化する k を用いることで、ノイズの影響を軽減しつつ二次近似の妥当性を保つような Hessian の推定値を得ることができる。

4.2 アルゴリズム中の評価関数について

4.2 節では、Algorithm 1 における評価関数 $r_i \cdot (l_{\text{NN}}^i)^2$ の意味について述べる。上で述べたように、この評価関数により選択される点は最も誤差を生じうる点を近似したものである。以降では、この近似の根拠となる定理を示す。

まず、線形内挿による非線形関数の近似において誤差の上界を与える結果を、ノンパラメトリック PWA モデルについて示す^{8),9)}。関数 f について、つぎを仮定する。

(A1) $f(\mathbf{x})$ は 2 階連続微分可能である。

また、データ集合 $\hat{D} = \left\{ \begin{pmatrix} \hat{\mathbf{x}}^i \\ \hat{y}^i \end{pmatrix} \right\}_{i=1}^{\hat{N}}$ に関して、以下を仮定する。

(A2) 任意の i について、 $\hat{y}^i = f(\hat{\mathbf{x}}^i)$ が成立する。

このとき、Delaunay 単体分割 $\mathcal{T}(\{\hat{\mathbf{x}}^i\}_{i=1}^{\hat{N}})$ 中の任意の単体 V_j 上で、元の関数とモデルの誤差の上界が与えられる。 V_j の外接円の中心および半径をそれぞれ $\mathbf{C}_{\text{cs},j}, R_{\text{cs},j}$ とし、 $\bar{r}_j = \sup_{\mathbf{x} \in V_j} \left\| \frac{\partial^2 f}{\partial \mathbf{x}^2} \right\|$ とする。 r_i が各点での Hessian の 2 ノルムの推定値であったのに対し、 $\bar{r}_j$ は各単体上での Hessian の 2 ノルムの上限であることに注意されたい。このとき、以下の不等式が成り立つ。

$$|f(\mathbf{x}) - \Pi_D(\mathbf{x})| \leq \frac{\bar{r}_j}{2} (R_{\text{cs},j}^2 - \|\mathbf{x} - \mathbf{C}_{\text{cs},j}\|^2) \quad \forall \mathbf{x} \in V_j \quad (6)$$

V_j の最小内包円 (付録 A 参照) の中心と半径を $\mathbf{C}_{\text{ms},j}, R_{\text{ms},j}$ とする。このとき、(6) の右辺は $\mathbf{x} = \mathbf{C}_{\text{ms},j}$ で最大値 $\frac{\bar{r}_j}{2} R_{\text{ms},j}^2$ をとる⁸⁾。なお、(A1) および (A2) の下でこの上界はタイトなものであり、より高階の微分を用いずとも十分に誤差を評価できる。以降では、アルゴリズム中の評価関数 $r_i \cdot (l_{\text{NN}}^i)^2$ を最大化する点が、 $\mathbf{C}_{\text{ms},j^*}$ を近似していることを示す。ここで、 j^* は (6) 右辺の最大値を最大化する j 、すなわち $\text{argmax}_j \frac{\bar{r}_j}{2} R_{\text{ms},j}^2$ とする。

まず、いくつかの表記を導入する。点群 $\hat{\mathcal{X}} = \{\hat{\mathbf{x}}^i\}_{i=1}^{\hat{N}}$ が与えられたとき、関数 $l_{\text{NN}}(\mathbf{x})$ を以下のように定義する。

$$l_{\text{NN}}(\mathbf{x}) = \min_{i \in \{1, \dots, \hat{N}\}} \|\mathbf{x} - \hat{\mathbf{x}}^i\| \quad (7)$$

これは、各点 $\mathbf{x}$ に対し、最も近い $\hat{\mathcal{X}}$ 中の点までの距離を対応づける関数である。この関数は $\hat{\mathcal{X}}$ の関数でもあるが、表記の簡単のため省略する。つぎに、 $\bar{r}_j$ に相当する値を返す関数を定義する。関数 f と点群 $\hat{\mathcal{X}} = \{\hat{\mathbf{x}}^i\}_{i=1}^{\hat{N}}$ が与えられたとする。このとき、区分的定数関数 $\bar{r}(\mathbf{x})$ を、以下の手順で値を返す関数とする。

(1) $\mathcal{T}(\hat{\mathcal{X}})$ 中の $\mathbf{x}$ を含む単体 V_j を抽出する。

(2) $\bar{r}_j = \sup_{\mathbf{x} \in V_j} \left\| \frac{\partial^2 f}{\partial \mathbf{x}^2} \right\|$ を返す。 $\mathbf{x}$ が複数の単体の境界上にある場合、各単体に対応する $\bar{r}_j$ のうち、最大となるものを返す。

l_{NN} と同様、これは $\hat{\mathcal{X}}$ や f の関数でもあるが、表記の簡単のため省略する。以上を用いると、 $r_i \cdot (l_{\text{NN}}^i)^2$ に関連するつぎの定理が証明できる。

定理 1. V_{j^*} の外接円中心が V_{j^*} に含まれる、すなわち外接円と最小内包円が一致すると仮定すると、以下の等式が成り立つ。

$$\mathbf{C}_{\text{cs},j^*} = \text{argmax}_{\mathbf{x} \in \text{Co}(\hat{\mathcal{X}})} \bar{r}(\mathbf{x}) \cdot (l_{\text{NN}}(\mathbf{x}))^2 \quad (8)$$

$$\frac{\bar{r}_{j^*}}{2} R_{\text{cs},j^*}^2 = \frac{1}{2} \max_{\mathbf{x} \in \text{Co}(\hat{\mathcal{X}})} \bar{r}(\mathbf{x}) \cdot (l_{\text{NN}}(\mathbf{x}))^2 \quad (9)$$

□

証明は付録とする。

この定理の解釈を述べる。仮定 (A1), (A2) が満たされる
とき、関数 f とそのモデル $\Pi_{\hat{D}}$ との誤差は、不等式 (6) の
形で抑えられる。 j^* の定義から、 $\frac{\bar{r}_{j^*}}{2} R_{\text{ms},j^*}^2$ はすべての j に
ついての (6) 右辺の最大値であり、外接円に関する仮定から
(6) 右辺を最大化する $\mathbf{x}$ が $\mathbf{C}_{\text{CS},j^*}$ である。すなわちこれらの
仮定の下で、関数 f とモデル $\Pi_{\hat{D}}$ の誤差上界を最大化する
点と、 $\bar{r}(\mathbf{x}) \cdot \left(l_{\text{NN}}(\mathbf{x})\right)^2$ を最大化する点は一致する。そして
この $\bar{r}(\mathbf{x}) \cdot \left(l_{\text{NN}}(\mathbf{x})\right)^2$ を緩和したものが、アルゴリズム中
の評価関数 $r_i \cdot \left(l_{\text{NN}}^i\right)^2$ である。 Hessian の 2 ノルムの変化
が緩やかなとき、 $\bar{r}(\mathbf{x}^i)$ は r^i で近似できるため、この定理は
Algorithm 1 の妥当性を示す根拠といえる。また、定理の後
半部分より、 $\frac{1}{2} r^i \cdot \left(l_{\text{NN}}^i\right)^2$ が誤差上界の最大値を近似したも
のともみなすことができる。したがってこれを監視することで、
選択する点の数 $\hat{N}$ を決定する方法も考えられる。

4.3 $\hat{\mathbf{y}}^i$ 推定の最小二乗問題による定式化

$\hat{\mathbf{y}}^i$ の推定に関する最小化問題 (3) について補足する。この
最小二乗問題は、モデル出力と観測データとの誤差二乗和を
最小化していることに対応する。ここで、 $\Pi_{\hat{D}}(\mathbf{x})$ はベクトル
 $[\hat{\mathbf{y}}^1, \dots, \hat{\mathbf{y}}^{\hat{N}}]^T$ に対し線形となるため、(3) は以下のように書
き換えられる。

$$\text{minimize } \|\mathbf{y}_0 - K\hat{\mathbf{y}}\|^2 \quad (10)$$

ただしここで $\mathbf{y}_0, \hat{\mathbf{y}}$ はそれぞれ $[\mathbf{y}_0^1, \dots, \mathbf{y}_0^{N_0}]^T \in \mathbb{R}^{N_0}$,
 $[\hat{\mathbf{y}}^1, \dots, \hat{\mathbf{y}}^{\hat{N}}]^T \in \mathbb{R}^{\hat{N}}$ である。また、 $K \in \mathbb{R}^{N_0 \times \hat{N}}$ は観測
データとモデル出力の誤差を計算するための行列であり、 $K\hat{\mathbf{y}}$
の第 i 行が $\Pi_{\hat{D}}(\mathbf{x}_0^i)$ となるような行列である。(10) は最小二
乗問題であり、高速かつ安定に解くことができる。これを解
くことで $\hat{\mathbf{y}}^i$ を求め、データ点集合 $\hat{D}$ を構成する。

丸田らが提案した方法³⁾では $\hat{\mathbf{x}}^i$ の配置について考えられ
ていなかったため、非常に多くの $\hat{\mathbf{x}}^i$ を Ω 上に一様分布する
乱数に従って生成し、それらすべてをモデル表現に用いていた。
このとき $\hat{\mathbf{y}}^i$ の推定については、過適合を防ぐペナルティ
項を (10) に加え、 ℓ_1 最適化に帰着させていた。一方提案法
では、 $\hat{\mathbf{y}}^i$ の推定に先立ってモデルのデータ点 $\begin{pmatrix} \hat{\mathbf{x}}^i \\ \hat{\mathbf{y}}^i \end{pmatrix}$ の数を
大幅に削減するため、過適合は起こりにくい。このため最小
二乗法による推定が可能となった。

4.4 アルゴリズムの計算量

Table 1 は、Algorithm 1 の計算量を表にしたものである。
この表において、Hessian 推定部分と点を選択する部分を陽
に分けた。これは Hessian の推定方法は、提案法における本
質的な部分ではないからである。

Table 1 Table of computational complexity

	Computational complexity
Hessian estimation	$O(k\hat{N}_0(\log N_0 + d^2))$
Data selection	$O(\hat{N}_0\hat{N}\log \hat{N})$

以降では、この章で述べたアルゴリズムの有効性を、数値
例および実験例を通して検証する。

5. 数値例による提案法の有効性検証

この章では、4 章で述べた提案法による $\hat{\mathbf{x}}^i$ の配置が確かに
適切であることを、数値例により検証する。

$d = 1$ として、関数 $f: \mathbb{R} \mapsto \mathbb{R}$ を、つぎのように定義する。

$$f(x) = e^{-0.6x} (3 \cos 3x + 4 \sin 4x + 5 \cos 5x) \quad (11)$$

このような関数 f に対し、(1) に示したような観測を行なっ
た。ただし η^i は、平均 0、分散 1 の正規分布に従う独立なノ
イズとする。このような観測を Ω 上で一様に 1000 点分生成
し、プロットしたものが Fig. 4 である。ただし Ω は以下の
ように定めた。

$$\Omega = \{x \mid -2 \leq x \leq 2\} \quad (12)$$

この観測データ集合に対し、提案法を適用した。各パラメー
タは、 $\hat{N}_0 = 2000, \hat{N} = 8, k = 22$ とした。このとき得られた
ノンパラメトリック PWA モデルを Fig. 5 に示す。実線が提
案法によるモデル、破線が比較のため示した八つの点を Ω 上
に等間隔に配置した場合の結果である。比較例での y 座標は、
提案法と同様最小二乗法により定めている。特に $x = -0.6$
近辺で、等間隔にデータ点を配置した場合のモデルの精度が
大きく劣化していることがわかる。

この例では、関数 f が既知であるため、得られたモデル $\hat{f}$

Fig. 4 Distribution of observed data generated by (11)

Fig. 5 Illustration of non-parametric PWA model constructed by proposed method

と関数 f との誤差を計算することができる．ここでは，誤差を以下に示す関数の 2 ノルムで測ることとする．

$$\sqrt{\int_{-2}^2 |f(x) - \hat{f}(x)|^2 dx} \quad (13)$$

このとき，提案法により得られたモデルの誤差は 3.916, Ω 上で等間隔にデータ点を配置したモデルの誤差は 7.039 であった．提案法は等間隔にデータ点を配置する場合に比べ，約 44% 誤差を削減している．これより提案法によるデータ点配置は，ただデータ点を等間隔に配置する場合よりも適切であるといえる．

6. モータに対する提案法の適用

つぎに，提案法を実システムであるモータの同定に対して適用し，その有効性について検討する．実験に用いたモータはトルク指令値 τ [N·m] を入力としており，今回は角速度 ω [rad/s] を出力とみなす．このとき状態方程式

$$\dot{\omega} = g(\tau, \omega) \quad (14)$$

を記述する関数 g の良い近似となる PWA 写像 $\Pi_{\hat{D}}(\mathbf{x})$ を同定する．ただしこのモータは 1 ms ごとにエンコーダにより角度を計測しており，これを差分することにより角速度および角加速度を求めている．入出力データを得る際には，一様分布する乱数に基づいて 0.1 s ごとにトルク指令値を変更し，18000 s の間，0.1 s ごとに $(\tau, \omega, \dot{\omega})$ をサンプリングした．

同定領域 Ω は，以下のように設定した．

$$\Omega = \left\{ \begin{pmatrix} \tau \\ \omega \end{pmatrix} \mid \begin{pmatrix} -13.5 \\ -2.5 \end{pmatrix} \leq \begin{pmatrix} \tau \\ \omega \end{pmatrix} \leq \begin{pmatrix} 13.5 \\ 2.5 \end{pmatrix} \right\} \quad (15)$$

このとき観測したデータから Ω に収まったデータ点 179977 点を，観測データ集合 $D_0 = \left\{ \begin{pmatrix} x_0^i \\ y_0^i \end{pmatrix} \right\}_{i=1}^{179977}$ とした．また，各パラメータは $\hat{N}_0 = 100000, \hat{N} = 35, k = 3000$ とした．これより得られたノンパラメトリック PWA モデルを図示したものを **Fig. 6** に示す．ただし，境界付近では端点を結んだ直線が現れ構造がわかりにくいので，やや内側の領域を図示している．また， $\mathbf{x}$ の各要素を規格化し，要素間のオーダーの違いから生じる問題を回避している． $\hat{N} = 35$ とした結果，PWA モデルのモードに相当する Delaunay 単体の数は 64 となった．このモデルと同じ観測データから同定した Hammerstein-Wiener モデル，および実機に対して，**Fig. 7** に示す検証用入力を印加した．ただしこの入力値はトルク指令値 τ [N·m] である．各モデルと実機の応答を **Fig. 8** に示す．灰色の細い実線が実機の応答を 300 周期分プロットしたもの，細い実線が比較のために示した Hammerstein-Wiener モデルの出力，太い実線が提案法によるモデルの出力である．なお，モータの特性は温度などによっても変化するため，300 周期分のプロットは帯状に見えるものとなっている．ただし Hammerstein-Wiener モデルは MATLAB 2013a System Identification Toolbox

Fig. 6 Model of motor constructed by proposed method

Fig. 7 Illustration of validation input

Fig. 8 Responses to the validation input

によって推定されたものであり，推定時の設定では線形部分を二次の離散時間伝達関数とし，入力遅れを 1 ステップとした．また，サンプリングタイムは 0.01 s とした．非線形項は区分的アフィンなものとし，そのモードの数は 10 とした．提案法では連続時間モデルを構築したため，提案法に関してはサンプリングタイム 0.001 s でシミュレーションを行なっている．

モータの非線形性は複雑なものであり，Hammerstein-Wiener モデルでは完全には表現できていないが，提案法ではその挙動をよく再現できていることがわかる．したがって，提案法は実システムに対しても有効な，汎用性の高い同定法である．

つぎに，Algorithm 1 による点配置の有効性を示すため，

Fig. 9 Responses to validation input

四つの方法でノンパラメトリック PWA モデルを構築し、その応答を見る。四つの方法の詳細は以下のとおりである。

提案法 $\hat{x}^i$ の配置に Algorithm 1 を利用し、少数の点からモデルを構成。 $\hat{y}^i$ は最小二乗法を用いて推定。

既存法 (1) $\hat{x}^i$ の配置はランダムで、多数の点からモデルを構成。 ℓ_1 最適化を用いて $\hat{y}^i$ を推定。

既存法 (2) まず既存法 (1) の方法で同定した後に、先行研究⁵⁾の方法で点を削減。

ランダム配置 $\hat{x}^i$ の配置はランダムで、少数の点からモデルを構成。 $\hat{y}^i$ は最小二乗法を用いて推定。

これらを比較する目的は、Algorithm 1 による点配置が確かに適切なものであると確かめることである。

Fig. 9(a) と Fig. 9(b) は、これら四種類のモデルに対して Fig. 7 に示した検証用入力を印加したときの応答を示している。ただし、図の視認性のために、Fig. 9(a) に提案法と既存法 (1)・(2) によるモデルの応答を、Fig. 9(b) にランダム配置によるモデルの応答を示している。どちらの図も、Fig. 8 と同様に灰色の細い実線で実機の応答を 300 周期分プロットしている。

まず、提案法と既存法 (1)・(2) を比較した、Fig. 9(a) の結果について述べる。ここでは太い実線、太い破線、細い実線が、それぞれ提案法、既存法 (1)、既存法 (2) によって構成されたモデルの出力となっている。この図から、提案法と既存法 (1) のモデルの応答はほとんど一致することがわかる。したがって提案法は、十分多くのデータ点でモデルを構築したときの精度と同程度の精度を少数のデータ点で達成しているといえる。一方、応答が実機のものとあまり一致しないことから、既存法 (2) のモデルの精度は高くはないことがわかる。

これは、35 点というデータ点数が、既存法 (2) でモデルを構築する際には不十分であることによる。

つぎに、提案法とランダム配置を比較した Fig. 9(b) について述べる。実線が提案法、破線がランダム配置によるモデルの応答を示している。ランダムな配置によるモデルは、実機の応答をある程度再現しているが、摩擦により生じる不連続性をあまり再現できていない。たとえば 1s から 1.5s までの間は、実機では摩擦により角速度が 0 となる。提案法によるモデルは実機と同様に一度停止する挙動を示しているが、ランダム配置によるモデルでは、この間も角速度が 0 にならない。これは、摩擦が $\omega = 0$ 近辺で生じるため、この領域に多くデータ点を配置しなければ摩擦を表現できるモデルとならないことによる。このように、システムが局所的に強い非線形性をもつような場合、ランダムな配置ではその非線形性をうまく再現できない。

以上の結果より、Algorithm 1 による点配置は、少数のデータ点からノンパラメトリック PWA モデルを構築する上で確かに有用であるといえる。また、ランダムに点を配置する場合と比べ、システムの特性をより効率的に記述していることがわかる。

最後に、計算時間について述べる。本稿では、実装メモリが 12.0 GB であり、プロセッサが Intel Core i7 CPU 920 (2.67 GHz) であるようなコンピュータ上で、MATLAB を用いて計算を行なった。このとき、既存法 (2) でモデルを構築するのに要した時間は、約 1900s であった。ただしここで、既存法 (2) における ℓ_1 最適化は、等価な二次計画問題に帰着して MATLAB 上の関数 quadprog を用いて計算した。一方提案法による同定に要する時間は、約 180s であった。なお、180s のうち Hessian 推定に要する時間は約 175s であり、計算時間の大半を占めている。この結果から提案法は、少数の点からなるノンパラメトリック PWA モデルを既存法に比べて高速に構築できることがわかる。

7. おわりに

本稿では、各点における推定 Hessian を用いてアフィン化誤差を見積もり、それを利用して点を配置するノンパラメトリック PWA モデルの新たな同定法を提案した。また、数値例、実験例を通して提案法の有効性を確認した。提案法ではモデル表現に用いる点の配置を選択的に行なっているため、既存法に比べてモデル表現に要するデータ点の数が大きく削減されている。また、提案法は最小二乗法により推定を行なうため、 ℓ_1 最適化を用いる既存法に比べて高速である。このため状態または入力の数が多いシステムへの適用が、既存法に比べて容易であると考えられる。今後の課題として、ノンパラメトリック PWA モデルの精度を理論的に保証することが挙げられる。

謝辞 本研究の一部は、JSPS 科学研究費補助金 #25249058 (基盤 A)、#25820179 (若手 B) の助成を受けた。記して、謝意を表したい。

参 考 文 献

- 1) O. Nelles: *Nonlinear System Identification: From classical approaches to neural networks and fuzzy models*, Springer (2001)
- 2) A. Garulli, S. Paoletti and A. Vicino: A survey on switched and piecewise affine system identification, *Preprints of the 16th IFAC Symposium on System Identification*, 344/355, IFAC (2012)
- 3) 丸田, 杉江: L_1 最適化によるデータ圧縮に基づく PWA モデルの同定, 第 11 回制御部門大会, 162-2-4 (2011)
- 4) 丸田, 足立: 区分的線形モデルを用いたシステム同定法の画像処理への応用, 第 41 回制御理論シンポジウム, 169/174 (2012)
- 5) 藤本, 丸田, 杉江: ノンパラメトリック区分的アフィンモデルの ℓ_1 最適化に基づく単純化, システム制御情報学会論文誌, **27-4**, 141/149 (2014)
- 6) A. Stenman, F. Gustafsson and L. Ljung: Just in time models for dynamical systems, *Proceedings of the 35th IEEE Decision and Control*, 1996, **1**, 1115/1120 (1996)
- 7) H. Akaike: Fitting autoregressive models for prediction, *Annals of the institute of Statistical Mathematics*, **21-1**, 243/247 (1969)
- 8) S. Waldron: The error in linear interpolation at the vertices of a simplex, *SIAM Journal on Numerical Analysis*, **35-3**, 1191/1200 (1998)
- 9) J. Shewchuk: What is a good linear finite element? Interpolation, conditioning, anisotropy, and quality measures, preprint, 11th International Meshing Roundtable (2002)
- 10) D. Du and F. Hwang, editors: *Computing in Euclidean Geometry*, 2nd edition, World Scientific Publishing (1995)
- 11) V.T. Rajan: Optimality of the Delaunay triangulation in $\mathbb{R}^d$, *Discrete & Computational Geometry*, **12-1**, 189/202 (1994)

《 付 録 》

A. Delaunay 単体分割について

$\mathbf{x}^i \in \mathbb{R}^d$ とする. 点集合 $P = \{\mathbf{x}^i\}_{i=1}^N$ が与えられたとき, $\text{Co}(P)$ を複数の単体 ($d+1$ 個の頂点からなる d 次元多面体) で重複なく覆うことを考える. ただしこれらの単体は, その頂点が P 内の点であり, かつまた P のすべての点は少なくとも一つの単体の頂点であるとする. このような単体の集合を構成することを単体分割と呼ぶ. $d=2, N=9$ の場合にこれを図示したのが **Fig. A.1** である.

Delaunay 単体分割はそのような分割の一つである. 一般次元の場合の Delaunay 単体分割については, 文献 10) が詳しい. この分割の一つの特徴は, 各単体の外接円が空円 (empty circle) となることである. すなわち, 各外接円はその内部に P 内の点をもたない. また, Delaunay 単体分割はいくつか

Fig. A.1 Illustration of Delaunay tessellation ($d=2, N=9$)

の最適性をもつことが知られている¹¹⁾. 特に, Delaunay 単体分割における最大の最小内包円半径は, ほかの任意の単体分割での最大の最小内包円半径以下である. なお, 最小内包円とは, 単体の境界および内部をすべて含むような円のうち, 半径が最小のもののことをいう. 単体の外接円中心が単体に含まれる場合は外接円と最小内包円は一致する. 外接円中心が単体に含まれない場合, 単体の各ファセットの最大外接円が最小内包円となる.

B. 定理 1 の証明

まず二つの補題を証明する. 一つ目は最小内包円の性質に関するものである.

補題 1. $\hat{\mathbf{x}}^i \in \mathbb{R}^d$ とする. $\{\hat{\mathbf{x}}^i\}_{i=1}^{d+1}$ を頂点集合とする単体を V とし, その最小内包円の半径と中心をそれぞれ $R_{\text{ms}}, C_{\text{ms}}$ とする. このとき, 以下が成り立つ.

$$l_{\text{NN}}(\mathbf{x}) \leq R_{\text{ms}} \quad \forall \mathbf{x} \in V \quad (\text{B.1})$$

Proof. 一般性を失うことなく C_{ms} を $\mathbb{R}^d$ の原点にとることができる. このとき, 任意の $\mathbf{x} \in V$ は, 以下のような凸結合の形で表わすことができる.

$$\mathbf{x} = \sum_{i=1}^{d+1} a_i \hat{\mathbf{x}}^i \quad \left(a_i \geq 0, \sum_{i=1}^{d+1} a_i = 1 \right) \quad (\text{B.2})$$

インデックス I をつぎのようにおく.

$$I = \underset{i}{\operatorname{argmax}} \mathbf{x}^T \hat{\mathbf{x}}^i \quad (\text{B.3})$$

このとき $\|\mathbf{x}\|^2$ はつぎの不等式を満たす.

$$\|\mathbf{x}\|^2 = \mathbf{x}^T \left(\sum_{i=1}^{d+1} a_i \hat{\mathbf{x}}^i \right) \leq \mathbf{x}^T \left(\sum_{i=1}^{d+1} a_i \hat{\mathbf{x}}^I \right) = \mathbf{x}^T \hat{\mathbf{x}}^I \quad (\text{B.4})$$

これは, 左辺が正であることに注意すると $\|\mathbf{x}\|^2 - 2\mathbf{x}^T \hat{\mathbf{x}}^I \leq 0$ と書き直せる. これより以下の不等式が成り立つ.

$$\begin{aligned} \|\mathbf{x} - \hat{\mathbf{x}}^I\|^2 &= \|\mathbf{x}\|^2 - 2\mathbf{x}^T \hat{\mathbf{x}}^I + \|\hat{\mathbf{x}}^I\|^2 \\ &\leq \|\hat{\mathbf{x}}^I\|^2 = R_{\text{ms}}^2 \end{aligned} \quad (\text{B.5})$$

$l_{\text{NN}}(\mathbf{x})$ は最寄りの $\hat{\mathbf{x}}^i$ までの距離であり, $l_{\text{NN}}(\mathbf{x}) \leq \|\mathbf{x} - \hat{\mathbf{x}}^I\|$ なので, 補題 1 は成り立つ. $\square$

補題 1 を用いると, つぎの補題が証明される.

補題 2. 点群 $\hat{\mathcal{X}} = \{\hat{\mathbf{x}}^i\}_{i=1}^N$ が与えられたとし, V_j を $\hat{\mathcal{X}}$ の Delaunay 単体分割中の単体とする. V_j の外接円中心が V_j に含まれると仮定する. このとき, 以下の等式が成り立つ.

$$\max_{\mathbf{x} \in V_j} l_{\text{NN}}(\mathbf{x}) = R_{\text{cs},j} \quad (\text{B.6})$$

$$\operatorname{argmax}_{\mathbf{x} \in V_j} l_{\text{NN}}(\mathbf{x}) = C_{\text{cs},j} \quad (\text{B.7})$$

$\square$

Proof. Delaunay 単体分割の特徴（外接円が空円となる．付録 A 参照）から， $l_{\text{NN}}(\mathbf{C}_{\text{cs},j}) = R_{\text{cs},j}$ である．仮定より $\mathbf{C}_{\text{cs},j} \in V_j$ なので， $\max_{\mathbf{x} \in V_j} l_{\text{NN}}(\mathbf{x}) \geq R_{\text{cs},j}$ が成り立つ．一方補題 1 より，任意の $\mathbf{x} \in V_j$ で $l_{\text{NN}}(\mathbf{x}) \leq R_{\text{cs},j}$ である．よって (B.6)，(B.7) が成り立つ． $\square$

以上の補題から，定理 1 を証明する．

Proof. まず (9) を証明する．補題 1 より，任意の V_j に対して以下が成り立つ．

$$\frac{\bar{r}_j}{2} \max_{\mathbf{x} \in V_j} \left(l_{\text{NN}}(\mathbf{x}) \right)^2 \leq \frac{\bar{r}_j}{2} R_{\text{ms},j}^2 \quad (\text{B.8})$$

また，補題 2 と仮定より，単体 V_{j^*} で以下が成り立つ．

$$\frac{\bar{r}_{j^*}}{2} \max_{\mathbf{x} \in V_{j^*}} \left(l_{\text{NN}}(\mathbf{x}) \right)^2 = \frac{\bar{r}_{j^*}}{2} R_{\text{cs},j^*}^2 \quad (\text{B.9})$$

以降では，つぎの式が成立することを示す．

$$\frac{1}{2} \max_{\mathbf{x} \in \text{Co}(\hat{\mathcal{X}})} \bar{r}(\mathbf{x}) \cdot \left(l_{\text{NN}}(\mathbf{x}) \right)^2 = \frac{\bar{r}_{j^*}}{2} \max_{\mathbf{x} \in V_{j^*}} \left(l_{\text{NN}}(\mathbf{x}) \right)^2 \quad (\text{B.10})$$

任意の $\mathbf{x} \in \text{Co}(\hat{\mathcal{X}})$ に対し， $\mathbf{x} \in V_k$ となるインデックス k を対応させる． $\mathbf{x}$ が隣接する単体 $(V_{k_1}, V_{k_2}, \dots)$ の境界上に位置する場合は， $k = \arg\max_j \bar{r}_{k_j}$ とする．このとき， j^* の定義と (B.8)，(B.9) から，任意の $\mathbf{x} \in \text{Co}(\hat{\mathcal{X}})$ に対して以下が成り立つ．

$$\begin{aligned} \frac{1}{2} \bar{r}(\mathbf{x}) \cdot \left(l_{\text{NN}}(\mathbf{x}) \right)^2 &\leq \frac{\bar{r}_k}{2} \max_{\mathbf{x} \in V_k} \left(l_{\text{NN}}(\mathbf{x}) \right)^2 \\ &\leq \frac{\bar{r}_k}{2} R_{\text{ms},k}^2 \\ &\leq \frac{\bar{r}_{j^*}}{2} R_{\text{cs},j^*}^2 \\ &= \frac{\bar{r}_{j^*}}{2} \max_{\mathbf{x} \in V_{j^*}} \left(l_{\text{NN}}(\mathbf{x}) \right)^2 \quad (\text{B.11}) \end{aligned}$$

仮定より $\mathbf{C}_{\text{cs},j^*} \in \text{Co}(\hat{\mathcal{X}})$ なので，(B.10) が成り立つ．したがって以下の関係を得る．

$$\frac{1}{2} \max_{\mathbf{x} \in \text{Co}(\hat{\mathcal{X}})} \bar{r}(\mathbf{x}) \cdot \left(l_{\text{NN}}(\mathbf{x}) \right)^2 = \frac{\bar{r}_{j^*}}{2} R_{\text{cs},j^*}^2 \quad (\text{B.12})$$

以上で定理 1 の (9) が証明された．(8) は，上の議論と補題 2 から得られる． $\square$

丸 田 一 郎 (正会員)

2011 年京都大学大学院情報学研究科システム科学専攻博士後期課程修了．同年日本学術振興会特別研究員 PD (慶應義塾大学訪問研究員)．2012 年から京都大学大学院情報学研究科特定助教となり現在に至る．システム同定の研究に従事．博士 (情報学)．システム制御情報学会の会員．

杉 江 俊 治 (正会員・フェロー)

1978 年京都大学大学院工学研究科精密工学専攻修士課程修了．84 年同博士課程研究指導認定退学．78～80 年日本電信電話公社勤務．84 年大阪府立大学工学部助手，88 年京都大学工学部助教授，97 年同教授を経て，現在京都大学大学院情報学研究科教授．ロバスト制御理論とその応用に関する研究などに従事．工学博士．日本ロボット学会，システム制御情報学会の会員．IEEE のフェロー．

[著 者 紹 介]

藤 本 悠 介 (学生会員)

1990 年生．2013 年京都大学工学部物理工学科卒業．同年より同大学大学院情報学研究科修士課程に入学し，現在に至る．システム制御情報学会の会員．