

実関数の帰納推論 (II) : 近似推論

Kalvis Apsītis¹

Setsuo Arikawa²
(有川 節夫)

Rūsiņš Freivalds³

Eiju Hirowatari²
(廣渡 栄寿)

Carl H. Smith¹

¹Department of Computer Science, University of Maryland, USA. : {kalvis,smith}@cs.umd.edu.

²Department of Informatics, Kyushu University. : {arikawa,eiju}@i.kyushu-u.ac.jp.

³Institute of Mathematics and Computer Science, University of Latvia, Latvia. : rusins@cc.lu.lv.

1 始めに

本稿では、例からの実関数の帰納推論について議論する。我々は、計算可能実関数 [4, 5, 6, 8, 9] や区間関数 [1, 7] の両方の特徴を合わせ持つ帰納的実関数を提案し、与えられた例から帰納的実関数を推論する学習モデルを提案した [2]。この学習モデルは、Gold の極限同定 [3] の自然な拡張となっている。

例からの実関数の帰納推論を実行する際に、目標関数そのものを学習できることが、実関数の学習の中で最も良い場合である。有理区間上で定義される帰納的実関数の帰納的に枚挙可能な集合は、帰納推論可能であることを示した [2]。しかしながら、学習目標である関数の集合が帰納的に枚挙可能であるとは限らない。それどころか、帰納的実関数で表されない場合もあり得る。そのような場合、与えられた数値データから、関数を計算するアルゴリズムの無限列を出力し、そのアルゴリズムが計算する帰納的実関数列が学習目標に収束していくような近似推論による極限学習が適している。本稿では、まず、このような学習モデルの4つの成功基準を提案する。また、有限回を除いて推測の表す関数と目標関数との誤差を正しく見積もることができる有限学習モデルについても議論をする。有限近似推論機械は、入力データとして、[2] で定義された例示だけでなく、目標関数の定義域の端点の例示も受け取ることによって、目標概念に収束していく関数列を計算するアルゴリズムの列を出力するだけでなく、有限回を除いて目標関数との誤差を正しく見積もりを行うことができる。そして、ある共通の定義域上で定義されている連続関数の集合に注目して、それらの極限学習可能性、有限学習可能性について議論する。

2 準備

帰納的実関数や関数の例示や推論機械等の定義は、特に断らない限り、[2] と同様である。本節では、以後の議論に必要な実関数間の距離を定める測度について準備を行なう。

定義 1 \mathcal{X} を関数の集合とする。次の条件を満たす写像 $d: \mathcal{X} \times \mathcal{X} \rightarrow R \cup \{\infty\}$ を測度と呼ぶ。

1. $(\forall h_1, h_2 \in \mathcal{X})[d(h_1, h_2) \geq 0]$,
2. $(\forall h_1, h_2 \in \mathcal{X})[d(h_1, h_2) = 0 \Leftrightarrow h_1 = h_2]$,
3. $(\forall h_1, h_2 \in \mathcal{X})[d(h_1, h_2) = d(h_2, h_1)]$,
4. $(\forall h_1, h_2, h_3 \in \mathcal{X})[d(h_1, h_2) + d(h_2, h_3) \geq d(h_1, h_3)]$.

本稿では、ある区間 I で定義される連続関数の集合上の測度として、以下のように定義される最大測度を用いる。

$$d(h_1, h_2) = \sup_{x \in I} |h_1(x) - h_2(x)|.$$

補題 1 σ を関数 h の例示とし, $\mathcal{A}_{h'}$ をある帰納的実数関数 h' を計算するアルゴリズムとする. h と h' は, ともにある区間 I 上で定義されているとする. このとき, $\sigma[n]$ と $\mathcal{A}_{h'}$ を入力として受け取り, h と h' の距離 d_n を計算し, 出力する手続き $\text{DISTANCE}_I(\sigma[n], \mathcal{A}_{h'})$ が存在する. さらに, $\text{DISTANCE}_I(\sigma[n], \mathcal{A}_{h'})$ の出力の列 $\{d_n\}$ は, $\lim_{n \rightarrow \infty} d_n = d(h, h')$ となる単調数列である.

$\text{DISTANCE}_I(\sigma[n], \mathcal{A}_{h'})$ は, 実際に, 以下のように構成できる.

```

Algorithm:  $\text{DISTANCE}_I(\sigma[n], \mathcal{A}_{h'})$ 
begin
   $d_n = 0$ ;
  let  $\sigma[n] = \langle w_1, \dots, w_n \rangle$ ;
  for  $k := 1$  to  $n$  do begin
    let  $w_k = \langle \langle p_k, \alpha_k \rangle, \langle q_k, \beta_k \rangle \rangle$ ;
    if  $[p_k - \alpha_k, p_k + \alpha_k] \subseteq I$  then begin
       $\langle q', \beta' \rangle := \mathcal{A}_{h'}(\langle p_k, \alpha_k \rangle)$ ;
      if  $[q' - \beta', q' + \beta'] \cap [q_k - \beta_k, q_k + \beta_k] = \emptyset$  then begin
         $d_{n,k} := |q' - q_k| - |\beta'| - |\beta_k|$ ;
        if  $d_{n,k} > d_n$  then  $d_n := d_{n,k}$ 
      end
    end
  end
  end;
  output  $d_n$ 
end.

```

h, h_n を区間 I 上で定義される関数とする. $\lim_{n \rightarrow \infty} d(h, h_n) = 0$ となるとき, 関数列 $\{h_n\}$ は, h に収束するという.

3 近似推論

3.1 学習の成功基準

推論機械 \mathcal{M} が関数 h を近似推論することに成功するとは, \mathcal{M} によって出力される推測の列がアルゴリズム $\mathcal{A}_{h'}$ に収束し, $\mathcal{A}_{h'}$ によって計算される関数 h' と関数 h との距離が十分に小さくなる場合のことをいう. 本節では, “推論機械 \mathcal{M} が目標関数 h を近似推論する” という成功基準として, 4 つの成功基準を提案する.

推論機械の推測が計算する関数列が目標関数に収束することは, 最も基本的な極限学習の成功基準である.

定義 2 h を連続関数とする. ある推論機械 \mathcal{M} が存在し, 任意の h の例示 σ に対して, \mathcal{M} によって出力される推測の列 $\{A_{h_n}\}$ が存在して, $\lim_{n \rightarrow \infty} d(h, h_n) = 0$ となるとき, \mathcal{M} は h を列極限学習するといい, $h \in \text{SEQLIM}(\mathcal{M})$ と表す.

帰納的実数関数の集合 \mathcal{T} に対して, 任意の $h \in \mathcal{T}$ を列極限学習する推論機械 \mathcal{M} が存在するとき, \mathcal{T} は SEQLIM 学習可能であるという. また, SEQLIM によって, SEQLIM 学習可能な全ての集合のクラスを表す.

推論機械の推測が計算する関数列の部分列が目標関数に収束するという, SEQLIM 学習可能性より弱い成功基準が考えられる.

定義 3 h を連続関数とする. ある推論機械 \mathcal{M} が存在し, 任意の h の例示 σ に対して, \mathcal{M} によって出力される推測の列 $\{A_{h_n}\}$ が存在して, $\lim_{k \rightarrow \infty} d(h, h_{k_n}) = 0$ となるとき, \mathcal{M} は h を部分列極限学習するといい, $h \in \text{SUBLIM}(\mathcal{M})$ と表す. ただし, $\{h_{k_n}\}$ は, $\{h_n\}$ の部分列である.

帰納的実数関数の集合 T に対して、任意の $h \in T$ を部分列極限学習する推論機械 M が存在するとき、 T は *SUBLIM* 学習可能であるという。また、*SUBLIM* によって、*SUBLIM* 学習可能な全ての集合のクラスを表す。

次の定義は、Gold による学習モデル [3] に最も近い成功基準である。ここで取り扱う推論機械 M は、データとともに誤差限界 $\varepsilon > 0$ を受け取り、 M によって出力される推測の列 \mathcal{A}_{h_n} は、目標関数との距離が ε 以内である帰納的実数関数 h^* を計算するアルゴリズムに収束しなければいけない。

有限個を除いて、 $h_n = h^*$ となる場合、推測の列 $\{\mathcal{A}_{h_n}\}$ が \mathcal{A}_{h^*} に収束するという。

定義 4 h を連続関数とする。ある推論機械 M が存在し、任意の h の例示 σ と任意の $\varepsilon > 0$ に対して、 M によって出力される推測 $\mathcal{A}_{h_n} = M(\sigma[n], \varepsilon)$ が存在して、推測の列 $\{\mathcal{A}_{h_n}\}$ が $\lim_{k \rightarrow \infty} d(h, h^*) \leq \varepsilon$ となる \mathcal{A}_{h^*} に収束するとき、 M は h を近似的に極限学習するといひ、 $h \in \text{PRELIM}(M)$ と表す。

帰納的実数関数の集合 T に対して、任意の $h \in T$ を近似的に極限学習する推論機械 M が存在するとき、 T は *PRELIM* 学習可能であるという。また、*PRELIM* によって、*PRELIM* 学習可能な全ての集合のクラスを表す。

学習目標である関数を h とし、 σ を h の例示とする。各データ $\sigma[n]$ を受け取るたびに、関数を計算するアルゴリズム \mathcal{A}_{h_n} と、 h_n と h の誤差 ε_n の組を推測として出力する推論機械を考える。 $d(h_n, h) \leq \varepsilon_n$ となる場合のみ、正しく誤差を見積もっているとする。推測を出力した後で、その推測 $\langle \mathcal{A}_{h_n}, \varepsilon_n \rangle$ は誤った見積もりをしていることを示すデータを受け取る場合がある。こういう場合には、誤った誤差の見積もりをしていた証拠を得た時点で、推論機械は、その推測の削除を行なうようにする。以上のように構成される推論機械を用いて、正しく誤差を見積もり、極限において学習目標である関数 h を計算するアルゴリズムに収束するような削除されない推測の無限列を得ることを、近似学習の成功基準とすることができる。

定義 5 h を連続関数とする。ある推論機械 M が存在し、任意の h の例示 σ と任意の $n \in N$ に対して、 M は組 $\langle \mathcal{A}_{h_n}, \varepsilon_n \rangle$ を出力し、 $k < n$ となる推測 $\langle \mathcal{A}_{h_k}, \varepsilon_k \rangle$ を次の条件を満たすように削除することができるとき、 M は h を強極限学習するといひ、 $h \in \text{STRLIM}(M)$ と表す。

1. $d(h_k, h) > \varepsilon_k$ となる全ての推測 $\langle \mathcal{A}_{h_k}, \varepsilon_k \rangle$ は、いつか必ず削除される。
2. 各推測 $\langle \mathcal{A}_{h_{n_i}}, \varepsilon_{n_i} \rangle$ が決して削除されないような、無限数列 $n_1 \leq n_2 \leq \dots$ が存在する。
3. 削除されない推測の列 $\{\langle \mathcal{A}_{h_{n_i}}, \varepsilon_{n_i} \rangle\}$ に対して、 $\lim_{i \rightarrow \infty} \varepsilon_{n_i} = 0$ が成り立つ。

帰納的実数関数の集合 T に対して、任意の $h \in T$ を強極限学習する推論機械 M が存在するとき、 T は *STRLIM* 学習可能であるという。また、*STRLIM* によって、*STRLIM* 学習可能な全ての集合のクラスを表す。

3.2 実関数の近似学習

本節では、提案した4つの成功基準の学習能力の比較を行なう。近似推論を行なう際には、推論機械の推測が計算する関数と、目標関数がどのくらいの距離にあるかを、入力として受け取るデータだけから、実際に求めなければいけない。補題 1は、その距離を計算する手続き $\text{DISTANCE}_T(\sigma[n], \mathcal{A}_{h'})$ が構成可能であることを示している。

$\text{SEQLIM} \subseteq \text{SUBLIM}$, $\text{PRELIM} \subseteq \text{SUBLIM}$, $\text{STRLIM} \subseteq \text{SUBLIM}$ であることが、その定義から明らかであり、*SUBLIM* が最も一般的な学習方法であることがわかる。

SUBLIM 学習可能性は、極限学習の中で最も制約の弱い成功基準である。そこで、まず、どのような関数の集合が *SUBLIM* 学習可能であり、どのような関数の集合がそうでないかを調べる。

任意の $h_\mu \in T$ に対して、 $h_\lambda \neq h_\mu$, つまり $d(h_\lambda, h_\mu) > \varepsilon_\lambda$ となる実数 $\varepsilon_\lambda > 0$ が存在するとき、関数 $h_\lambda \in T$ は孤立しているという。

定理 1 \mathcal{X} を共通の定義域をもつ連続関数の集合とし, $T \subseteq \mathcal{X}$ とする. T が非加算個の孤立関数を含むとき, T は *SUBLIM* 学習可能でない.

例 1 T を半開区間 $(0, 1]$ 上で定義されている一様有界連続関数の集合とする. このとき, T は *SUBLIM* 学習可能でない.

\mathcal{X} と \mathcal{Y} を関数の集合とする. 任意の関数 $h \in \mathcal{X}$ に対して, \mathcal{Y} の中のある関数によって十分に近似可能であるとき, \mathcal{Y} が \mathcal{X} に稠密であるという.

定理 2 \mathcal{X} を共通の定義域をもつ連続関数の集合とし, $\mathcal{Y} \subseteq \mathcal{X}$ を帰納的実数関数の帰納的に枚挙可能な集合とする. \mathcal{Y} が \mathcal{X} に稠密であるとき, \mathcal{X} は *SUBLIM* 学習可能である.

例 2 T を閉区間 I 上で定義されている全ての連続関数の集合とする. このとき, T は *SUBLIM* 学習可能である.

連続関数を近似推論により極限学習するためには, 目標関数の定義域が有理閉区間であることが重要である. *SUBLIM* 学習可能性以外の学習可能性について, 以下の定理が得られる.

定理 3 T を有理閉区間 I 上で定義される全ての連続関数の集合とする. このとき, T は *SEQLIM* 学習可能であり, *STRLIM* 学習可能である.

補題 2 $STRLIM \subseteq PRELIM$.

例 2 と同様に, 有理閉区間上で定義される全ての連続関数の集合もまた, *SEQLIM*, *STRLIM* の意味でも学習可能である. また, 補題 2 から *PRELIM* 学習可能であることもわかる.

4 有限学習

有限学習は, 極限学習に有限回を除いて推測の表す関数と目標関数との誤差を正しく見積もらなければならないという制限を加えた学習である. 前節で導入した 4 つの学習の成功基準に対して, それぞれ対応する有限学習の 4 つの成功基準が考えられる. 有理閉区間上で定義される全ての連続関数の集合を対象としたときに, それらの 4 つの学習可能性は同等であった. そこで, 本節では, *STRLIM* 学習可能性に相当する成功基準に着目し, どのような関数の集合がその基準のもとで有限学習可能であるかを議論する.

定義 6 h を連続関数とする. ある学習機械 M が存在し, 任意の h の例示 σ に対して, $\sigma[n]$ を入力として受け取った M の出力が組 $\langle A_{h_n}, \varepsilon_n \rangle$ であり, $d(h, h_n) \leq \varepsilon_n$ かつ $\lim_{k \rightarrow \infty} \varepsilon_n = 0$ となるとき, M は h を有限近似するといひ, $h \in STRFIN(M)$ と表す. ただし, 有限回だけ $\varepsilon_n = \infty$ となることを許す.

帰納的実数関数の集合 T に対して, 任意の $h \in T$ を有限近似する推論機械 M が存在するとき, T は *STRFIN* 学習可能であるという. また, *STRFIN* によって, *STRFIN* 学習可能な全ての集合のクラスを表す.

M は推測を削除することができないので, 全ての誤差 ε_n を正しく見積もらなければならない. 明らかに, *STRFIN* 学習可能性は, *STRLIM* 学習可能性の制限になっている. *STRFIN* 学習可能性の条件は, 正しく誤差の見積もりを行いつつ, 関数を補完したり, 積分を計算したり, 微分方程式を解いたりする数値解析の分野で, 通常要求されていることであり, ごく自然なものである.

全ての連続関数の集合は *STRFIN* 学習可能でない. 実際, ある関数に関してどんなに多くのデータを受け取ったとしても, 推測機械に出力される推測が表す関数が, 目標関数と全く同じものであるという保証を与えることはできない. そこで, 目標関数を単調関数に制限する. しかし, 単調関数ですら, 次の例で示すように *STRFIN* 学習可能でない場合がある.

例 3 $\varepsilon, \delta \in \mathbb{R}$ とし, $\delta > 0$ とする. T を次のような区間線形関数 $h_{\varepsilon, \delta} : [0, 1] \rightarrow \mathbb{R}$ 全体の集合とする.

$$h_{\varepsilon, \delta}(x) = \begin{cases} \varepsilon - \frac{x}{\delta}\varepsilon & \text{if } x \in [0, \delta], \\ 0 & \text{if } x \in [\delta, 1]. \end{cases}$$

明らかに, 全ての $h_{\varepsilon, \delta}$ は単調関数である. しかしながら, T は *STRFIN* 学習可能でない.

この例は, 十分に小さな区間 $[0, \delta]$ で関数の違いが存在するとき, その関数の違いを表す証拠をデータ $\sigma[n]$ から見つけることができないことを示している. [2] で与えた関数の例示は, *STRFIN* 学習モデルに適していない. なぜならば, どんなに大量のデータを受け取ったとしても, 任意の精度の端点 $h(0), h(1)$ の値を知ることはできないからである. そこで, 有理閉区間 $[a, b]$ が与えられたとき, その境界である a と b の例示の仕方を次のように定義する.

定義 7 $[a, b]$ を有理閉区間とし, $h : [a, b] \rightarrow \mathbb{R}$ を連続関数とする. 次のようなデータ列 w'_1, w'_2, \dots を考える.

$$w'_k = \langle \langle a, 0 \rangle, \langle q'_k, \beta'_k \rangle \rangle, \quad k > 0.$$

各 k に対して, $h(a) \in [q'_k - \beta'_k, q'_k + \beta'_k]$ かつ $\lim_{k \rightarrow \infty} \beta'_k = 0$ となるとき, データ列 w'_1, w'_2, \dots を境界条件 $h|_{x=a}$ の例示と呼ぶ. 境界条件 $h|_{x=b}$ の例示も同様に定義される.

定理 4 T を閉区間 $I = [a, b]$ 上で定義されている全ての連続非減少関数の集合とする. このとき, 任意の $h \in T$ を *STRFIN* 学習する推論機械 M が存在する. ただし, M は, h の任意の例示とともに境界条件 $h|_{x=a}$ と $h|_{x=b}$ の任意の例示 w'_k と w''_k を入力として受け取るものとする.

5 終わりに

本稿では, 誤差を含む数値データから, 関数を計算するアルゴリズムの無限列を出力し, そのアルゴリズムが計算する帰納的実数関数が学習目標である実関数に収束していくような極限学習モデルを提案した. また, *SEQLIM*, *SUBLIM*, *PRELIM*, *STRLIM* の4つの成功基準を提案し, ある共通の区間上で定義されている連続関数の集合に対して, それらの学習能力の比較を行なった. 最も制約の弱い成功基準である *SUBLIM* 学習可能性ですら, 半開区間 $(0, 1]$ 上で定義されている一様有界連続関数の集合を学習することができない. しかし, 有理閉区間上で定義される全ての連続関数の集合においては, 4つの成功基準のいずれでも学習可能であるという結果を得た. さらに, 推論機械が, 推測の列を出力するだけでなく, 有限回を除いてその推測が表す関数と目標関数との誤差を正しく見積もることのできる有限学習可能性を提案した. この推論機械では, 学習すべき関数 h が有理閉区間上で定義されているとき, その境界点 $(x, h(x))$ のデータを $\langle \langle x, 0 \rangle, \langle q, \beta \rangle \rangle$ のように, x の誤差を無視した形で与えることによって, 閉区間 $[a, b]$ 上で定義されている全ての連続非減少関数の集合を *STRFIN* 学習可能であることを示した.

参考文献

- [1] G. Alefeld and J. Herzberger. *Introduction to Interval Computations*. Academic Press, London, England, 1983.
- [2] K. Apsitis, S. Arikawa, R. Freivalds, E. Hirowatari, and C.H. Smith. 実関数の帰納推論 (I): 厳密推論. In 98冬のLAシンポジウム予稿集, pp. 29-1 - 29-6, 1998.
- [3] E.M. Gold. Language identification in the limit. *Information and Control*, Vol. 10, pp. 447-474, 1967.
- [4] A. Grzegorzcyk. Computable functionals. *Fundamenta Mathematicae*, Vol. 42, pp. 168-202, 1955.
- [5] A. Grzegorzcyk. On the definitions of computable real continuous functions. *Fundamenta Mathematicae*, Vol. 44, pp. 61-71, 1957.

- [6] K. Ko. *Complexity Theory of Real Functions*. Birkhäuser, 1991.
- [7] R.E. Moore. *Interval Analysis*. Prentice-Hall, 1966.
- [8] A. Mostowski. On computable sequences. *Fundamenta Mathematicae*, Vol. 44, pp. 37–51, 1957.
- [9] M. B. Pour-El and J.I. Richards. *Computability in Analysis and Physics*. Springer-Verlag, 1988.