6K-03　関数関係とならない教師データを伴う３層ニューラルネットの学習

＊下川 信祐　河野 芳江　寺前 裕之
(株) エイ・ティ・アール環境適応通信研究所

1　はじめに

人とモノの関係が深まり複雑になってきた今日，高機能・高性能な良いものを創ることよりも，将来の市場をイメージし，これに向けて的確なプランを建てて実行することがますます重要になって来ています．私達は，人とモノの複雑な関係に分けいって，その特徴を調べ，新たな人とモノの関係を築くプランを建てるのに役立つ，概念的な枠組と方法を検討しています [1, 2].

デザインコンセプトを実現してゆく方法として，様々な物事の背後に潜む関係の中から連続体のものを，ニューラルネット等を用いて取り出し，予測や評価に利用することを検討しています (図 1).

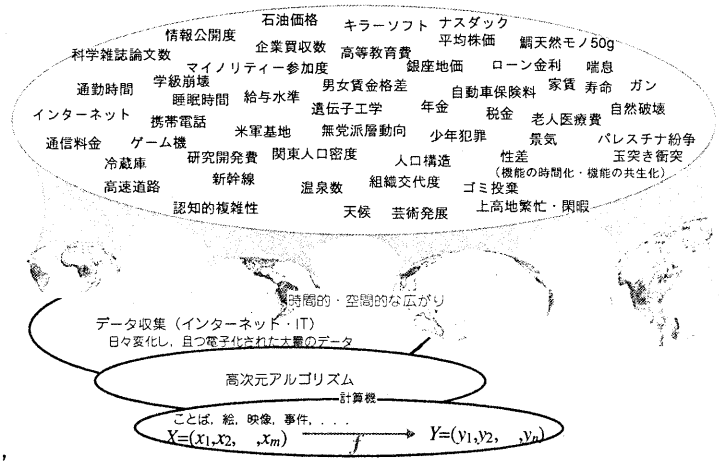

図１．様々な物事の背後に潜む関係を取り出して予測・評価する.

2　高次元性，緩やかな関係

抽出の対象となる関係は，通常，かなり高次元の空間に配位したものとなります．この時，必ずしも関数関係とは限らない，緩やかな関係を抽出する必要がしばしば生じます．例えば，収集できるデータが関係の抽出に有効な密度を伴うには，データの次元をあまり大きくできません．次元が減らされると関係性の‘独立変数’を尽くすことが難しくなり，‘独立変数’の値に対応する‘従属変数’の値は多値となって関係性にあそびが生じます.

一方，関数関係は予測・評価として利用しやすく，計算機による表現や扱いもやさしくなります．そこで，必ずしも関数関係ではない関係を関数関係を通して取り出すことを考えます.

3　商空間

本検討では，従属変数の自由度を落して関数関係が得られる場合を考察します．即ち，従属変数の空間 Y に適当な同値関係 $\sim$ を導入して得られる商空間 $Y/\sim$ 上の変数に対して，データが関数関係となるとします．この時，商空間 $Y/\sim$ は一般に非線形 (多様体) となるので，ニューラルネットなどの出力変数に直接コードすることは良策とはいえません．そこで，従属変数の空間 Y を出力変数にコードし，誤差評価関数 $Er(g)$ を通常のものに変えて，つぎのようにとります.

$$Er(g) = \sum_{k=1}^{N_{\rm data}} \rho^2(\pi(\mathcal{N}(u_k, g)), \pi(v_k)). \tag{1}$$

ここで，$\{(u_k, v_k)\}_{k=1}^{N_{\rm data}}$ は教師データ，$\mathcal{N}(u, g)$ は，結合状態 g における，入力 u に対する多層ニューラルネットの出力，π は商写像 $\pi\colon Y \to Y/\sim$, ρ^2 は商空間上の適当な距離関数です．通常の誤差評価関数では，多値データによって収束が阻まれるところを，商写像による弱い誤差評価関数を導入して回避します.

‘Learning in 3-layered neural network with the teacher data which are not functionally related.’

SHIMOGAWA, Shinsuke (simogawa@acr.atr.co.jp), KOHNO, Yoshie (kohno@acr.atr.co.jp), TERAMAE, Hiroyuki (teramae@acr.atr.co.jp).

ATR Adaptive Communications Research Labratories.

2-2-2 Hikaridai, Seika-cho, Souraku-gun, Kyoto Pref. 619-0288.

4 評価のための問題，多項式の根

多値データを伴う評価モデルとして，多項式の根を射影空間上で求めることをネットに模倣させる問題を扱います．以下ではこれを説明します．出発点は，1変数代数方程式：

$$a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 = 0 \tag{2}$$

です．係数 $(a_n, a_{n-1}, \cdots, a_1, a_0)$ を与えて，方程式の解 $(\alpha_1, \alpha_2, \cdots, \alpha_n)$ をニューラルネットを用いて近似的に求めることを考えます．学習データを，根と係数の関係

$$a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 = a_n (x - \alpha_1)(x - \alpha_2) \cdots (x - \alpha_n)$$

を用いて作ります． $a_n = 1$ とせず，低次の方程式を含む形でのネットの模倣特性に興味があります．しかし，このままでは， $a_n \sim 0$ の近くで，データが退化し，ネットの学習がうまくゆきません．これは代数方程式(2)自身が， $a_n = 0$ で退化することに対応しています．そこで，低次の方程式を含みならが，退化の無い方程式にする必要があります．これは方程式(2)を次のように射影化することで果たせます：

$$a_n X^n + a_{n-1} X^{n-1} Z + \cdots + a_1 X Z^{n-1} + a_0 Z^n = 0. \tag{3}$$

同次方程式(3)の零点集合は平面 $\mathbb{k}^2$, $\mathbb{k} = \mathbb{R}$(実数体) or $\mathbb{C}$(複素数体)，上の原点を通る幾つかの直線となり，平面 $\mathbb{k}^2$ 上の原点を通る直線全体である一次元射影空間 $\mathbb{P}$ 上の点集合をさだめます．原点を除く平面上の点 $(Z, X) \in \mathbb{k}^2 - (0)$ に対し (Z, X) と原点を通る直線を $[Z, X]$ と記します (同次座標,[3])． x が方程式(2)の解であることと， $[1, x]$ が方程式(3)の解であることは同値です．射影化したとき，根と係数の関係は，次のようになります．

$$a_n X^n + a_{n-1} X^{n-1} Z + \cdots + a_1 X Z^{n-1} + a_0 Z^n = (p_1 X - q_1 Z)(p_2 X - q_2 Z) \cdots (p_n X - q_n Z). \tag{4}$$

この恒等式によって定まる関係

$$u = (a_n, \cdots, a_0) \rightsquigarrow v = (p_1, q_1, \cdots, p_n, q_n) \iff \text{恒等式(4)成立},$$

が，ニューラルネット $\mathcal{N}$ に模倣させたい緩やかな関係です．実際，対応関係 $\rightsquigarrow$ は，そのままでは関数関係ではなく，商写像 π: $(\mathbb{k}^2 - (0))^n \ni v = (p_1, q_1, \cdots, p_n, q_n) \mapsto ([p_1, q_1], \cdots, [p_n, q_n]) \in (\mathbb{P})^n$ と結合させて(局所的には) 関数関係となります．

そこで，射影空間 $\mathbb{P}$ 上の距離関数 $\rho_{\mathbb{P}}^2([a, b], [c, d]) = |ad - bc|^2 / (|a|^2 + |b|^2)(|c|^2 + |d|^2)$ を用いて， $(\mathbb{P})^n$ 上の距離関数 $\rho^2 := \sum_{j=1}^n \rho_{\mathbb{P}}^2$ をとり，誤差関数(1)によってニューラルネット

$$\mathcal{N}(g) : \mathbb{k}^{n+1} \ni u = (a_n, \cdots, a_1) \mapsto v = (p_1, q_1, \cdots, p_n, q_n) \in (\mathbb{k}^2 - (0))^n$$

の学習をこころみます．

5 まとめ

関数関係とならない緩やかな関係を商空間の導入によって関数関係として取り出す方法の検討について述べました．関数関係でない教師データを多層ニューラルネットに与えて弱い誤差評価関数による学習が考えられます．多項式の根の問題は，そのための評価問題として利用可能です．数値例については，講演で紹介する予定です．

参考文献

[1] 下川，大田原，“インターネットを含むコミュニケーションツールの利用調査とユーザー意識 – 人とモノの新たな関係を築くために –，” 電子情報通信学会誌， Vol.83, No.10, pp.740-742(2000).

[2] 下川，新上，大田原，河野 “横断するデザイン，一人とモノの関係からの戦略と方法 -,” 日本 OR 学会創立 40 周年記念事業特別研究助成プロジェクト「情報通信ネットワークの新しい性能評価法に関する総合的研究シンポジウム」報文集，けいはんな，京都府精華町，Jan. 2001.

[3] 河田 敬義，『アフィン幾何・射影幾何』 岩波講座 基礎数学 2，岩波， 1976.