

多様性を前提とした社会経済現象の構造化手法に関する研究

東京工業大学 紀伊雅敦 Masanobu KII

東京工業大学 土井健司 Kenji DOI

1. はじめに

社会経済現象をあらわすモデルの構造は一般に現象に関する公理,あるいは仮定から演繹的に導かれる。したがってそのモデルの妥当性は,現象の再現性ととも、この前提となる公理,仮定により保証される。一方,現象が多様性を有し,事前知識のみから現象の構造を表すことが困難な場合,あるいは演繹的に導かれるモデルにおいて必要とされるデータが不足している場合には,得られているデータのみから現象を構造化する手法が必要とされる。データに基づく構造化手法として,回帰分析に代表される数多くの手法が提案されている。本論ではこれらの方法について概説し,そのひとつである人工ニューラルネットワーク(Artificial Neural Network; ANN)¹⁾を用いた,効率的な構造化手法について提案する。

2. 構造化手法の概要

従来,様々な現象の構造化手法が提案されているが,それらは一般に,説明したい現象をデータから推定するものであるといえる。ここで,AHPなどの意思決定手法なども人間の知識,感覚を通じた現象の構造化手法であると考えられるが,アンケート,インタビューなどを用いており,構造化の観点からは特殊な手法である。

ここでは社会経済データからその現象構造を導く一般的な手法について概説する。その手法についてまとめたものが表—1である。これらの手法はその仮定する構造の特性により,線形,非線形に分けられる。

線形の構造化手法としては,線形回帰モデル,LISREL等があげられる。これらは,パラメータ値の統計的安定性,構造解釈の容易性などの利点がある反面,その構造上の制約から,現象が非線形性,非単調性

表—1 代表的な構造化手法

線形	線形回帰モデル LISREL
非線形	人工ニューラルネットワーク(ANN) GMDH

などの多様性を有する場合,十分な再現性が得られない,すなわち現象の構造を適切に把握できない可能性がある。また LISREL は説明変数の合成変量による線形モデルを推定するものである。ただし,合成変量の構造はアприオリに与える必要があり,構造確認的な手法であるといえる。

次に,非線形の構造化手法として,ANNとGMDH(Group Method of Data Handling)²⁾があげられる。これらは共に任意の関数を高精度に近似することが可能な手法として位置付けられているものである。ANNは複数の単調な非線形関数を層状に組み合わせることにより任意の関数の近似を与える入出力システムを実現するものである。しかしながら,バックプロパゲーションといわれるパラメータ推定方法については,計算量が膨大である,大域的最適性が保証されないといった問題がある。GMDHはある説明変数の組み合わせによる,複数の線形回帰モデルの重ね合わせで現象を表すモデルを構築している。そのためパラメータ推定は線形モデルにおける手法を用いればよいことから,ANNと比較して高速に行うことが可能である。ただし,GMDHは,回帰モデルの重ね合わせにおいてシングモイド変換を行う場合,ANNで表しうる関数族の部分族となることが示されている³⁾。従って,GMDHは最適性に関してANNと同様の問題を有する。本論ではモデル分析における現象の構造化手法としてANNに着目し,その汎化を考慮した効率的なパラメータ推定手法について提案する。

3. モデル評価の効率化

多くのモデル選択において情報量基準による方法が用いられてきた。これは、モデルに用いる変数、パラメータ数を変化させて得られる複数のモデルの比較を行うものであるが、ANNのようにパラメータ数を自由に設定可能な場合、比較すべきモデルの数は非常に多くなる。Watanabe⁴⁾は、パラメータ数をパラメータに関して微分可能な関数とし、かつ従来の情報量基準を一般化した形式の修正情報量基準を提案している。これによりデータに基づく期待される誤差を最小とするパラメータとパラメータ数を同時に推定可能となる。ここではこの情報量基準について概説する。

まず ANN が表すモデルを $\varphi(x; \theta)$ とする。またこのモデルの出力の確率分布を $N(\varphi(x; \theta), \sigma)$ とする。すると、情報量基準 I_α は、パラメータ数 m 、サンプル数 n 、出力の確率密度関数 $p(\theta; y|x)$ により θ の関数として、

$$I_\alpha(\theta) = 2n L_{emp}(\theta) + A F_\alpha(\theta) \quad (1)$$

$$\text{但し, } L_{emp}(\theta) = -\frac{1}{n} \sum_{i=1}^n \log p(\theta; y_i | x_i) \quad (2)$$

$$F_\alpha(\theta) = \sum_{i=1}^m f_\alpha(\theta_i) \quad (3)$$

と表される。 $f_\alpha(\theta)$ は次の2つの条件を満たす。

$$(1) f_\alpha(\theta) \rightarrow f_0(\theta) \text{ when } \alpha \rightarrow 0$$

$$F_0(\theta) = \sum_{i=1}^m f_0(\theta_i)$$

$$f_0(\theta) = \begin{cases} 0 & (\text{if } \theta = 0) \\ 1 & (\text{if } \theta \neq 0) \end{cases}$$

$$(2) |\theta_1| \leq |\theta_2| \rightarrow 0 \leq f_\alpha(\theta_1) \leq f_\alpha(\theta_2) \leq 1$$

この $f_\alpha(\theta)$ の例として $f_\alpha(\theta) = 1 - \exp(-\theta^2/\alpha^2)$ 等があげられている。この $I_\alpha(\theta)$ は θ に関して2階微分可能であり、この情報量基準をパラメータ推定の目的関数として用いることが可能である。但し、 $I_\alpha(\theta)$ は $A=2$ 、 $\alpha \rightarrow 0$ の時、AICであり、予測誤差の理論上の期待値を最小化するものとなる。文献4)では A を感度により決定していたが、ここでは Cross Validation を用いた A の推定方法について提案する。

4. Cross Validation による情報量基準の選択

まず、与えられた入出力データセット (S) を g 個のサブセットに分割し、その j 番目のサブセット (S_j) を除いたデータセット ($S - S_j$) を用い、 $I_\alpha(\theta)$ を最小とするパラメータの推定量を $\hat{\theta}(-j)$ とする。このとき Cross Validation による平均対数尤度の期待値は

$$L_{CV} = -\frac{1}{g} \sum_{j=1}^g \frac{1}{n_j} \sum_{(x_i, y_i) \in S_j} \log p(\hat{\theta}(-j); y_i | x_i) \quad (4)$$

で与えられる。但し、 n_j は S_j のデータ数である。そこで $I_{CV} = 2N L_{CV}$ とおき、 $I_{CV} \approx I_\alpha(\hat{\theta})$ を仮定すると、式(2)より

$$A = \frac{2n(L_{CV} - L_{emp}(\hat{\theta}))}{F_\alpha(\hat{\theta})} \quad (5)$$

とおける。これは、 L_{emp} と比較して L_{CV} が大きい、すなわち、データセット ($S - S_j$) に対するモデルの誤差と比較して (S_j) に対するモデルの誤差が大きい場合、パラメータ数で表される項の影響を大きくすることにより、モデルの汎化性能を向上させるための I_α を設定するものである。

5. おわりに

社会経済現象のモデル分析において、現象の構造化は必要不可欠であるが、現象に関するデータや知識が限られている場合、利用可能なデータから帰納的に構造化を行う必要がある。本研究では非線形に対応した構造化手法である ANN において、その期待対数尤度の近似値を極大化するモデルの推定手法を提案した。従来、モデル選択では、想定する全モデルについてパラメータ推定を行わなくてはならなかったが、この方法により、効率的な構造化が可能になるものと考えられる。

- 1) 豊田秀樹; 非線形多変量解析—ニューラルネットによるアプローチ—, 朝倉書店, 1996
- 2) 池田三郎; GMDH(変数組み合わせ計算法)の基礎と応用-I-, システムと制御, vol.23, No.12, pp.710-717, 1979
- 3) 清水英範, 児島利治; 人工衛星画像のファジィ分類法への GMDH の適用可能性—ニューラルネットワークとの比較を中心として—, 写真測量とリモートセンシング, vol.33, No.2, 1994
- 4) S. Watanabe; A Modified Information Criterion for Automatic Model and Parameter Selection in Neural Network Learning, IEICE Transaction on Information and systems, Vol. E78-D, No.4, 1995