

対話型モデリング支援システム

中森 義輝

1. はじめに

数学的モデリングの一般のプロセスは、モデルタイプの選択、パラメータ推定、およびモデル検証の3段階からなる。すなわち、与えられた入力データを用いて、指定されたクラスのモデルの中から、数学的モデルを求める。つまり、モデルのパラメータ推定以前に、モデルの構造がわかっていなければならないという問題がある。したがって、問題領域専門家のどんな知識を、どう利用するかが重要なポイントとなる。

近年、コンピュータ支援によるモデリングの概念や設計指針が、多くの研究者によって提案されるようになった。そこでは人間とコンピュータ、あるいは人間と人間との対話をいかに円滑にするかが重視される。ここで紹介する対話型モデリング支援システム (IMSS: Interactive Modeling Support System)¹⁾は、構造モデリング+統計モデリング+インタフェースからなるコンピュータソフトである。

IMSSは現在約180の対話場面からなり、PC98上(一部はSUN3上)で稼働している。紙面の都合で全容を説明できないので、現在話題のファジィモデル²⁾の構築におけるIMSSの使用法について簡単に紹介する。

2. モデルの記述法

数値データからモデルを構築するさいには次のような静的記述を用いる。

- (i) 説明変数集合: $I = \{x_1, x_2, \dots\}$
- (ii) 被説明変数集合: $O = \{y_1, y_2, \dots\}$

ここに、集合 I は制御変数と過去の状態を表わす変数を含み、また、集合 O は現在の状態変数と出力変数を含む。ただし、時間ステップ1年というような社会問題を扱う場合は、集合 I の中に現在の状態変数の一部を加え

ることもある。

求めるファジィモデルは次のようなルールの集合からなるモデル群である。

$$\begin{aligned} \text{ルール } R^K: & \text{if } x_1 \text{ is } A_1^K, \text{ and } x_2 \text{ is } A_2^K, \dots, \\ & \text{then } y_i = c_{i0}^K + \sum_j c_{ij}^K \cdot x_j, \quad i=1, 2, \dots \end{aligned}$$

ここで、条件部に登場する変数を前件部変数、結論部に登場する説明変数を後件部変数という。また、 A_1^K, A_2^K, \dots はファジィ部分集合で、それらのメンバーシップ関数 $A_1^K(x_1), A_2^K(x_2), \dots$ を規定する定数を前件部パラメータといい、結論部の線形式の係数を後件部パラメータという。ファジィモデルによる予測値は次式で与えられる²⁾。

$$y_i^* = \frac{\sum_K w^K \cdot y_i^* K / \sum_K w^K}{\text{均}} \quad (\text{ルールについての加重平均})$$

$$w^K = \prod_j A_j^K(x_j^*) \quad (\text{前件部変数についての積})$$

ここに、 x_j^* は入力値、 $y_i^* K$ はルール R^K からの出力値である。

ファジィモデル同定の課題は、

- (a) 前件部変数の選択 (データ空間のファジィ分割)
- (b) ファジィ部分集合のメンバーシップ関数の同定
- (c) 結論部の線形式の変数選択と係数の同定

という3つのサブ問題からなる。図1にIMSSによるファジィモデル同定プロセスの概略を示す。また以下では、その主要部分の簡単な説明を行なう。

3. データの観察とクラスタリング

まずデータの前処理とクラスタリングについての注意事項をまとめる。

変数集合の定義: 変数集合をあらかじめ適切に定義することは困難なケースが多いので、モデリング過程で変数の追加・修正が必要となる。そのさいの主要な観点は次の2つである。

- (i) 時間遅れあるいは非線形変換された変数を用いて、ダイナミックあるいは非線形関係の同定を可能とす

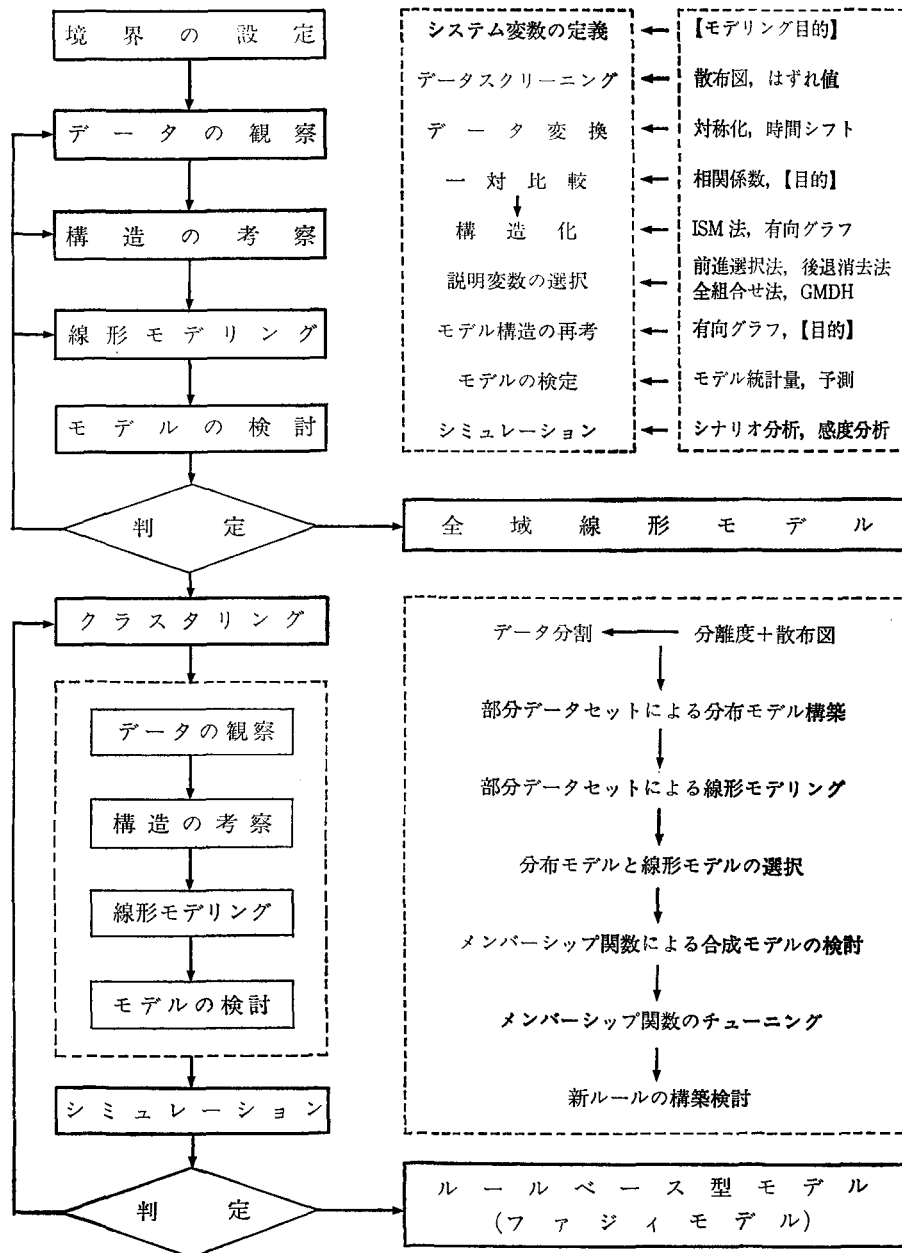


図 1 対話型モデリング支援システムによるファジィモデルの同定プロセス

る。

(ii) 分布の非対称性と非線形関係はしばしば共存するから、分布の対称化に心がける。

観測データの観察：はずれ値の影響や、多重共線性問題を回避するためにデータスクリーニングを行なう。データセットへのアクセシビリティを重視し、データの図示（単独あるいは複数の散布図）に心がける。

クラスタリング：データの散布状況を参照して、以下の手順によりファジィ部分集合を構築する。

- (1) クラスタリング手法（Ward法を用いる）により、各説明変数の標準化データを重複を許して2分割する。
- (2) これに対応して他の変数（被説明変数を含む）の標準化データを2分割する。
- (3) データ分離度（クラスターの重心間距離にデータ

数に関する重みを考慮)を計算する。

(4) データ分離度の大きいものから順に、データの散布状況を観察し、どの変数でデータ分割するかを決定する。ここでの留意点は部分データセットから変数間の線形関係をうまく抽出することであり、データ分離度は参考とするだけである。

4. 問題構造化

ここでは、2項関係の導入によりモデルシステムの構造を考察する。

2項関係: 入出力変数の明確な区別や状態変数間の関係づけが困難な場合、一対比較により因果関係を導入し、全体の構造をグラフ表示して変数間関係の考察を行なう。また関係の強度を導入し、説明変数選択時に選択のされ方に序列をいれる。ここで、因果関係の最初の入力にネガティブとする。すなわち、関係ないと断定できるところのみを指定する。推移的埋込法⁹⁾を利用し、部分的一対比較により直接的に可到達関係を入力することもできる。ただしここでは、間接的影響も注意深く考慮する必要がある。

関係のグラフ表示⁹⁾: 変数を頂点とし、関係を矢線であらわす有向グラフを描いて、モデル構造について考察する。描く有向グラフの種類は、隣接グラフ、可到達グラフ、および骨格グラフである。修正ルーチンにおいてはサイクルセットの分割、矢線の追加・削除、隣接行列の直接修正などが対話的に行なえる。

5. 線形モデリング

ここでの主要なステップは、説明変数の選択、モデルパラメータの同定、モデル構造のグラフ表示、およびモデル構造の修正である。

説明変数選択法: オプションとして、前進選択法、後退消去法、全組合せ選択法、および GMDH⁹⁾を用意している。入力した変数間関係のある基準にしたがって参照して説明変数を選択する。

モデル構造のグラフ表示: モデルの構造を有向グラフ表示し、目的にかなう構造であるかどうかを検討する。有向グラフの種類は、隣接グラフ、可到達グラフ、および骨格グラフである。ここでは、狭いスクリーン上でわかりやすく見せるための工夫をしている⁹⁾。

モデル構造の修正: 修正法は、フォーマットの修正、サイクル内関係の修正、階層構造の修正、および隣接関係の直接修正である。修正を行なえば線形モデリングを

繰り返す。

6. モデル検証

主要なフェーズはモデル単純化、統計的解析、およびシミュレーション(シナリオ分析、感度分析)である。

モデル単純化: 1つの線形式の中に同一サイクル内の変数が数多く存在したり、多くの異なるサイクルからの変数がいくつか存在するような場合、モデル構造を不変に保ったままでモデルを単純化する。

統計的解析: さまざまなモデル統計量により、モデルの説明力および予測力を検定する。それらは、推定された回帰係数の標準誤差、自由度修正決定係数、残差プロット、残差の標準偏差、予測の信頼区間などである。また、個々の回帰係数が0かどうかを判定するt検定、すべての回帰係数が0かどうかを判定するF検定も可能である。

シナリオ分析: 入力変数の値を想定し、その他の変数の挙動を眺める。入力変数同士も相関を保って動いている場合が多いので、入力値の組合せに注意しなければならない。

感度分析: 出力変数の1次式の最適化(線形計画法)に対してその他の変数の挙動を眺める。あるいは、変数の変域変更の影響(感度分析)を視察する。このモードはシナリオ分析とは反対の側からモデルを眺めるシミュレータである。

7. ルールベース型モデル

部分データセットによっては線形モデルが不適当なケースがあるから、その場合の後件部は線形モデルのかわりに分布モデル(ファジィ集合にもとづく)とする。実際には、分布と線形の両モデルのうちどちらを採用するかを、ある種のシミュレーションによって決定する。入力変数同士といえども実際には相関を持って動いているので、ある入力変数の値が決められたとき、他の入力変数の許容入力範囲がスクリーン上で変更されるなどの工夫をしている⁷⁾。

このようにしていくつかのルールが求められたとき、メンバーシップ関数のチューニングによりルールの協調化を計る⁷⁾。ここでの2つの観点は次のとおりである。

(i) ルールが合成される非線形部分を、実データと予測値とを比較しながら修正する(内挿問題)。

(ii) メンバーシップ関数やサポートセットを修正することにより、モデルの有効範囲を決定する(外挿問題)。

Computer Today

9月号 / 8月18日発売 / 定価930円

最新Xウィンドウシステムのすべて

Xウィンドウのすべて 篠田陽一・今泉貴史

はじめに
ウィンドウシステムの基礎知識
ウィンドウシステムの利用
X libプログラミング
ツールキットプログラミング
ウィンドウの将来

アルゴリズムアニメーション 榎原博之
ユーザインタフェースワークステーション
とネットワーク 笠原孝雄

＜新連載＞
プログラミングとロジシャン 野崎昭弘
MS-DOSシェルプログラムの技法 木下 恂
Cの高速コーディング 太田昌孝
アセンブラ入門 玉井 浩

月刊誌

数理科学

8月号 / 発売中 / 定価960円

パラドックスの効用

パラドックスの効用	今井 功
パラドックスから見た科学史	金子 務
数学パラドックスの世界	中村義作
バナッハ・タルスキの逆理	前原昭二
言葉と想い	清水達雄
ソリトンのパラドックス	戸田盛和
パラドックスの常識	広田良吾
美の進化 性淘汰のパラドックス	巖佐 庸
寄生のパラドックス	佐々木 顕
子殺しのパラドックス	山村則男
レオンチェフの逆理	武隈慎一
民主主義と自由主義のパラドックス	川本隆史

■最新刊 好評発売中

情報処理システム入門

浦 昭二・市川照久共編 / A5 / 定価1751円

▶価格表示は、税込み価格となっています。

サイエンス社

東京都千代田区神田須田町2-4 安部徳ビル

☎03(256)1091 振替 東京7-2387

場合によっては、新しいルールを構築する必要がある。

8. おわりに

モデルに要求される詳細はモデルの使用目的に依存するが、モデルは必要かつ最小で操作性の高いものでなければならない。そのためには対象のある局面のみを大胆にカットしなければならない。これは問題の性格のみならず、モデラーの価値観に負うところが大きい。対話型モデリング支援システムの最も大きな利点は、モデリングを通じて問題が見えてくることである。

参考文献

- [1] Nakamori Y.: Development and Application of an Interactive Modeling Support System. *Automatica*, Vol.25, No.2 (1989), 185/206.
- [2] Takagi T. and Sugeno M.: Fuzzy Identification of Systems and Its Applications to Modeling and Control. *IEEE Trans. on Syst., Man and Cybern.*, Vol.SMC-15, No.1 (1985) 116/132.
- [3] Warfield, J.N.: *Societal Systems: Planning, Policy and Complexity*. John Wiley, New York, 1976.
- [4] Warfield, J.N.: Toward Interpretation of Complex Structural Models. *IEEE Trans. on Syst. Man and Cybern.*, Vol.SMC-4, No. 5 (1974), 405/417.
- [5] Ivakhnenko, A.G.: The Group Method of Data Handling, A Rival of the Method of Stochastic Approximation. *Soviet Automatic Control*, Vol.13, No.3 (1968), 43/55.
- [6] 中森義輝, 村崎好則: 階層グラフの最適頂点配置のための発見的アルゴリズム. 計測自動制御学会論文集, Vol.25, No.3 (1989), 369/376.
- [7] 中森義輝, 馬屋原一孝, 甲斐沼美紀子, 森田恒幸: 混合型フェジィモデル構築について, 第5回フェジィ・システム・シンポジウム講演論文集, 1989, 337-342.