

モデル化技術の修得をねらった回帰分析教育

大成 幹彦

東京理科大学 工学部 経営工学科

モデル化技術の教育をシステム工学教育の一貫としてとらえ、要求分析から評価にいたる流れの中で、モデル化の課題発見と対応策を自己発見的に修得させることを意図している。モデル化の手段として回帰分析を取り上げる。計算手順を効率的に学ばせて、重点をモデル化技術に置くような教育手順と教材を準備した。表計算ソフトの利用が、計算手順の修得、構築、保存に有効である。教育上重要な点は、表計算上での掃き出し法の説明と実演、模範例での手順と結果の紹介、散布図上で直感的な勾配の求め方、自由課題での変数の選択、意図と因果関係のわかるタイトルの付け方などである。段階的な課題提供によって、受講者の興味を継続し、満足度も向上した。

Education of modeling technology helping heuristic learning

Mikihiko Ohnari

Department of Industrial Management & Engineering, Faculty of Engineering, Science University of Tokyo

In teaching modeling technology in the course of systems approach education, we intend to support students heuristically learning how to define a problem and collect alternative methods to solve it considering its requirements. As a modeling tool regression analysis is selected and teaching schedule and materials are prepared weighing on how to build a model. Spread sheet has revealed effective in learning, constructing and storing the calculation procedure. Many important points in education are experienced. Given students topics and theme to learn step by step, they can keep up their interest and finally obtain their satisfaction.

1. まえがき

コンピュータの利用に当たって要求される能力の一つはシステム思考である。問題点の発見能力とそれらを系統的に表現する能力である。そのためには、対象を把握し、モデル化する能力を訓練する必要がある。コンピュータの教育では、与えられた問題を解く (Solving) という対応よりも、課題や要求を自ら分析して課題の構造を明確にする (Structuring) というアプローチがより重要である。このような視点でモデリングやシミュレーションを教えている。道具として回帰分析を選び、受講者全員に短期間で学ばせ、続いてそれを使ってモデリングの技術を体得させることを意図した。教育上の課題としては、教育手順、計算手段、教材の吟味などがあげられる。これらの課題を、教育現場での体験に基づき評価する。

2. モデリングの教育目的

モデリングの教育課題は、表1に示すようになる。

表1 モデリング教育で教えるべきこと

・モデリング
・モデル化の定義
・目的指向
・写像、抽象化
・変数の選択
・因果関係の認識
・タイトルの付け方
・モデルの構造
・線形 vs. 非線形
・静的 vs. 動的
・シミュレーションモデル vs. 最適化モデル
・利用可能なツール
・回帰分析 (静的線形)
・ARMAモデル (動的線形)
・ニューラルネットワーク (非線形) など
・モデルの妥当性の評価
・Verification
・統計量の評価
・多重共線性の問題
・Validation
・ラーニングデータ
・チェックングデータ

モデリングとは、現実世界の対象を目的に応じて抽象化して思考の世界に写像することである [1]。モデル化したい対象をどのようにとらえて表現するか、何を従属変数に選び、それをどのような独立変数で説明しようとするのか。それらを仮説にまとめて整理しておく。収集したデータは図にプロットして変数間の因果関係を確認し、どのようなモデルにするかと適合度の高いモデルになるかを想像し、モデルの構成とそのためのツールを仮定する。変数間の因果関係の認識は、モデル化の意思の表示であるため、モデル化に際しての最重要事項である。この重要性を認識させるために、モデルに対するタイトルの付け方を指導する。

モデル式が算出できてから、モデルの妥当性・有効性を確かめさせる。分析結果の確認は、統計量があくまで参考値であることを認識して、最終的には自分の目でもってモデルの適用結果を注視すべきである。モデリング教育では、対象に関する課題の明確化から始めて、モデル式の構築・確認までを規範的に実行できる能力を育成する必要がある。

3. モデリング教育の手段としての回帰分析

3.1 回帰分析の工学的教育のねらい

回帰分析の教育的ねらいは、モデル化の手段として回帰分析が使えるようになることである。原理と手順を一通り理解させて、その後は計算手順を隠蔽して、道具として使えるようにする。「隠蔽」とは、作り上げた手順を保存し、その後は必要なとき以外は見なくても済むようにするという意味で使う。道具として使えるためには、その道具が他の道具に比べて有用性が高くなければならない。工学教育では、道具を選択する眼力も付与しなくてはならない。

3.2 回帰分析学習のためのコンピュータの利用法

従来の回帰分析の学習法は、原理を学んだ後に、電卓あるいはポケコンの利用、回帰分析のプログラム作成、汎用統計パッケージの利用などの方策を取っていた。最近ではパソコンの表計算ソフトが使える。あるいは表計算ソフトに含まれる回帰分析ソフトが利用できる [2]。

3.3 回帰分析の教育の内容

(1) 回帰分析の原理と手順の教育: 回帰分析の原理は講義で説明する。最小二乗法で回帰係数が算出されることを理解させ、連立方程式 (正規方程式) の導出を教える。掃き出し法などの計算手順は講義で説明し、その手順を自分で表計算上に構築させて実践で覚えさせる。掃き出し過程の中から回帰係数や多くの統計量が抽出できる。掃き出し過程で直接的に求められない偏相関行列などの統計値は、掃き出し

表2 表計算ソフトでの回帰分析結果の例

A行列 (左まの平方和・積和行列)	正規方程式	記号
0: 1 2 3 29	1 0 0 0	0
1: 26 64 3 67	0 1 0 0	0
2: 3 3 13 17	0 0 1 0	0
3: 29.1 67.2 16.7 85.5	0 0 0 1	0

A...行列 (平方和・積和行列)	正規方程式の係数行列の逆行列	記号
0: 1 2 0.2 2.24	0.08 0 0 0	
1: 0 12.1 -1.97 9	-2 1 0 0	S1.
2: -0 -1.97 12.2 10.9	-0.2 0 1 0	S2.
3: 0 9 10.9 20.4	-2.24 0 0 1	S3.

A...行列	正規方程式の係数行列の逆行列	記号
0: 1 0 0 0.21 0.43 -0.17 -0.04	0 0 0 0	b ⁰ c ₀
1: 0 1 0 0.91 -0.17 0.08 0.01	0 0 0 0	b ¹ c ₁
2: 0 0 1 1.03 -0.04 0.01 0.08	0 0 0 0	b ² c ₂
3: 0 0 0 0.94 -0.21 -0.91 -1.03	1 0 0 0	Sr

A...行列	平方和・積和行列の逆行列	記号
0: 1 0 0 0 0.48 0.03 0.19 -0.23		p ₁
1: 0 1 0 0 0.03 0.96 1.01 -0.56		p ₂
2: 0 0 1 0 0.19 1.01 1.22 -1.1		p ₃
3: 0 0 0 1 -0.23 -0.96 -1.1 1.06		

単相関行列	偏相関行列	記号
0: 1	-1 -0.93 0.95	r ₁ ρ ₁
1: -0.16 1 0.69	-0.93 -1 0.96	r ₂ ρ ₂
2: 0.57 0.69 1	0.95 0.96 -1	r ₃ ρ ₃

結果を利用して追加計算で求める。

計算手順の理解のために、2つの独立変数を含む計算問題とデータを与えて、表計算上に作成した回帰分析の計算例 (表2) をフォローさせる。予測すべき式としては、 $y = x_1 + x_2$ を想定する。

(2) 2つの独立変数の問題: 上記計算例のデータセットから1組か、2組削除して、2つの独立変数の規定問題を解かせる。データを散布図上にプロットさせ、そのデータを吟味して直感的に勾配を描かせる。単回帰分析の場合に散布図上で相関の強いデータから回帰直線を直感的に描くことは易しい。だが、重回帰分析の場合には、散布図上で直感的に勾配を求めることは、一般には難しい。しかし、2つの独立変数の問題で人為的にデータを作る場合には、独立変数の値をいくつかのグループに分け、各グループでその値をほぼ等しくしておくことができる。このときには、散布図上で隠れている変数をグループ化することによって、グループごとに勾配が描ける。

回帰分析結果として得られた回帰係数を散布図上に描かせ、直感的に得た勾配と比較させる。独立変数の数が少なくなった場合の回帰結果を学ばせるため、2独立変数の重回帰結果から、t値の小さいほうの独立変数を強制的に削除して、単回帰分析を実行させる。上記規定問題の解答レポートは、計算、散布図、考察の3つの観点からチェックする。散布図では、縦軸、横軸の名称とキャプションの記入の重要性を指摘する。考察では、得られた結果の数値を引用しながら結果を評価することを促す。

表3 モデル化技術教育の課題と対応

モデル化技術の教育課題	講義と解答講評					解答例	規定問題	変数減少	自由課題
	教育の週								
	1	2	3	4	5				
モデル化の概念									
写像、抽象化の概念	○								◎
変数の抽出		○						○	◎
仮説			○/●	○/●	○/●		○		◎
タイトルの付け方			○		○/●				◎
モデルの構築									
データの収集		○	○		○/●				◎
従属変数、独立変数の選択		○			○/●	○	○	○	◎
モデル式の構成		○	○		○/●	○	○	○	◎
モデル式係数の決定									
係数の決定手法			○			◎	隠蔽	隠蔽	隠蔽
変数選択法			○		●			◎	○
回帰分析の教育課題									
回帰分析の原理（最小二乗法）		○				○	習熟	習熟	習熟
正規方程式の作成		○/●				◎	習熟	習熟	習熟
回帰分析の計算手順（表計算ソフト使用）									
表計算ソフトの使用法	○/●					◎	習熟	習熟	習熟
掃き出し法、逆行列	○/●					◎	隠蔽	隠蔽	隠蔽
統計量の抽出		○/●	●			◎	習熟	習熟	習熟
統計量の追加計算		○/●	●			◎	隠蔽	隠蔽	隠蔽
変数減少法		○	●					◎	○
モデル化技術									
散布図		○	●	●			◎	○	○
直感的勾配の算出		○	●				◎	◎	
多重共線性の問題					○/●				◎
ビジュアルな表現		○	●	●	●		◎		◎
モデル式の検証									
Verification					○/●		○	○	◎
Validation					○/●		○	○	◎
出題と解答の時期									
3元連立方程式の解法	△	●							
規定問題	△	○	●	●					
自由課題	△		○		●				
多重共線性の問題				△	●				

○：講義
●：宿題の解答と講評
△：プリント提供

重要度：◎>○

(3) 自由課題：規定問題に続いて、自分で課題を見つけ、データを集めて分析することを出題する（表4）。自由課題で、モデリングのポイントを教える。モデリングでは、失敗も許容し、それを反省の糧にさせる。失敗を許容すべきところは、自由課題に多くあり、失敗しやすいところが、主要なチェックポイントとなる。題材によっては、多重共線性の問題がでてくる。

4. 教育結果の評価

モデル化技術を教えるには、問題解決法（how to）

だけでなく、課題発見（what to）についても教える必要がある。問題解決法と課題発見は一気に学べない。講義、回帰分析の模範計算例、規定問題、自由課題などを通して順次理解できる。教育課題と教材との対応を表3に示す。この表では教育の流れとどの教材をいつ用いて教えればよいかを、経験に基づいて分析した。まず、解決手段を教える。その原理と手順を教えておく。学習が進むにつれ、手順はできるだけコンピュータにまかせて受講者の意識から隠蔽していき、受講者の関心を課題発見とその解決へと導く。このような視点から、一度学んだことは

速やかに卒業して、それを隠蔽できるようにすること、学習を楽しく継続できること、自己発見的な学習、システムズアプローチで思考することなどの面から教育効果を評価する。

4.1 計算の容易さ

計算の手順は、ミスなく実行できるようになれば、隠蔽することが、受講者の負担を軽減することになる。いつでも隠蔽内容を見ることができるようしておくことは当然である。そのためには、手順を自分でプログラム化させておくのが教育上有効だ。既存の統計パッケージや表計算ソフトに備わっている回帰分析ソフトを利用してもよいが、初心者には既存ソフトの利用は容易ではない。その理由は、それらの組み込みソフトが機能を多く含み過ぎていること、回帰分析の原理を学んだばかりの者には、それらの中から必要な機能を抽出利用することが困難なことによる。

4.2 学習意欲の維持面から計算手段の比較

利用する計算手段によって、学習意欲の継続性と強さが変わってくる。

(1) パソコンが手元にない者は、電卓やポケコンを利用しなければならない。電卓計算では、回帰分析の計算のたびに手順を思い出さないといけない。手続きの隠蔽ができない。このため分析結果に意識を向ける時間と意欲が少なくなる。

(2) プログラミングができる者は、回帰分析のプログラムを組むが、汎用的にするのは手間である。そのため、与えられた変数の数の範囲でプログラムを組むが、独立変数の数を拡張するときに苦勞する。

(3) 表計算ソフト利用: 表計算ソフトを使ったことのある者にとっては、掃き出し法は苦勞なく実行できる。しかし、表計算ソフトに初めて接する者にとっては、表計算と掃き出し法との組み合わせはバリエーションが高いようである。このときに、表計算のセルに計算式を埋め込む方法、他のセルの引用の仕方、さらに他のセルを引用するときに通常表示される相対位置を絶対位置に変換する方法を、掃き出し法の手順の実習の中で教えると、この段階での障害は解決する。とくに表計算ソフトの実演をすると、約30分で、受講生全員が理解し、その後は回帰分析の計算を独りで構築実行できる。

回帰分析計算をプログラムや表計算ソフトで手順化するときの次の課題は、掃き出し法の計算過程からは直接抽出できない単相関係数や偏相関係数、重相関係数、 t 値などの統計量の追加計算である。追加計算の中で、特に戸惑いが多いのは単相関係数、偏相関係数を求めるときに用いる逆行列値を掃き出し結果のどの部分から抽出してくるかである。手順隠

表4 回帰分析自由課題とレポートのまとめ方

身近にあるデータから、回帰モデルを作成せよ。 解答には、モデル構造の仮説とその理由、その代替案、分析結果、結果の評価、考察を入れること。	
1. タイトル	従属変数に及ぼす独立変数の寄与・影響
2. 目的、ねらい	分析目的 仮説: 経験論, 常識 従属変数と主要な独立変数
3. 使用データ	データは記載する。追試ができなくてはならない 出典、測定日時・場所・条件を明記
4. 散布図	横軸と縦軸の名称記入。Captionの記入: 意図の反映
5. 回帰分析	非線形性は独立変数のべき乗で吸収 変数の選択: 変数減少法
6. 分析結果の考察	統計値: t 値 (F値), R , R^2 は目安 散布図上に回帰式または回帰係数を描く 誤差の評価。特に適用範囲での誤差評価をする
7. 今後の課題	仮説, データ, 変数, 手法の見直し 論理的な帰結だけを書く

蔽の視点からは計算ができるようになればよいので、これらの計算は模範計算例の上に表示することにした。その結果、相関係数の算出手順を間違える者が減った。

4.3 自己発見的学習のための配慮

規定問題が計算できるようになったところで、散布図を描いて直感的に得た勾配と回帰係数とを比較させる。独立変数が1つの場合には、散布図上のデータを見ながら最小二乗法的に直線を引くことができる。それは回帰直線に近いものである。独立変数が2つの場合には、散布図上で隠れた変数の値がほとんど等しいものを選んでグループ化することによって、グループごとの直感的勾配を引くことができる。この直感的勾配が回帰係数になることを体験させる。隠れた変数のグループ化を理解させるために x_2 vs. x_1 の散布図を描かせ、各変数の値の近いものでグループ化することを教える。これは直感的勾配を描かせるのに有効である。散布図上で直感的な勾配を描かせる過程で回帰係数の意味は理解できる。変数減少の影響も自習させる。2つの独立変数の回帰式と、1つの独立変数を強制的に削除したときの単回帰結果との比較から、有用な変数の削除が回帰式に及ぼす影響を体得できる。この段階で、帰無仮説や t 検定の理解を深めさせる。

4.4 妥当性の評価

構築したモデルの善し悪しは、評価項目を列挙してから吟味すべきである。評価項目は自由課題の場

合に多く抽出できる。それらは、モデリングの際にどれだけ系統的に考えたか、規範的な発想をしたかを示すことになる。評価項目には、レポートのまとめ方として表4に示すような項目が列挙できる。扱ったデータによっては、多重共線性の問題が発生することがある。この問題は、体験してはじめて回帰分析の前提条件に気付く。多重共線性の問題は、相互に相関の強いデータを独立変数に選んだことによる[3]。失敗は貴重な教育チャンスである。失敗するチャンスを残しておくことも、ヒューリスティックな教育としては重要なことである。

自由課題では変数の選択が重要となる。従属変数とそれを説明する独立変数を規範的な思考から選べるかどうかのポイントとなる。変数選択の善し悪しはタイトルにも反映する。タイトルは、単に変数相互の関係を数学的に記述するという意識では、目的指向の系統的教育の趣旨に沿わない。タイトルの付け方も、自分で付けたタイトルが妥当かどうかを、他人から批判されることによって、初めて自分の問題として認識できるようになる。タイトルでは、課題と解決策を主張しなければならない。レポートでは適切なタイトルを付けることを指導している。タイトルの付け方を出題時に説明しても、十分な理解は得られない。各人が提出してきたタイトルを添削して提示することにより、因果関係を明らかにしたタイトル作りを体験するようである。指導したタイトルの改善例を表5にあげる。

5. アンケート結果

モデリング教育の効果を受講生に出したアンケートによって評価する。2つのグループから集計した。設問は両グループで一部異なっている。表計算ソフトを使っての回帰分析の教育は、ともに67%の受講生が満足している。最初のグループでは、何も実演しなかった(図1)。その後教育したグループでは掃き出し法に必要な表計算の使い方をパソコン上で実演した。(図2)。掃き出し法と表計算の教育の徹底(図2)は、それまで(図1)よりも肯定的な答えが増えている。この結果、モデリング教育の理解度が全般的に向上したことがわかる。ただ、手順は

わかったが、どのようにモデル化したらよいかは、ともにまだ不十分であることが、考察の記述に関する項目などに反映している。手順教育が終わってれば、あとはモデル化の教育に徹することができる。

6. あとがき

教育の世界では、教科主義と経験主義との両教育方針がある[4]。筆者は企業での経験から、未知なることへの挑戦意欲を育てることが重要と信じている。人間が体験に触発されながら自己発見的に課題を解決していくことに期待を掛けて、モデリングの教育をしてきた。問題発見型の人材を育てることが、使命と感じている。そうすることが、長い目で見れば、受講者に喜ばれる教育[5]になると考える。

回帰分析をモデル化技術のわかりやすい手段と位置付けて、段階的に理解させ、できるだけ短い期間で使えるようにさせることを意図した。受講生の満足度と向学心を刺激するようにも配慮してきた。多くの新人がモデリングに興味を持ちシステムズアプローチの1つの過程として体験してくれることを期待している。多くの反復と試行錯誤の結果、問題発見型の人材開発に必要なモデリング技術の教育上の課題として、下記を抽出した。これらは、経験的に得た重点項目である。

「モデリングの要点の短期間での学習、掃き出し法の表計算上での実演と説明、表計算上での回帰分析の手順の構築、理解した後の手順の隠蔽、模範例での解析手順と結果の紹介、散布図上での直感的な勾配の求め方、仮説の大切さ、合目的な変数の選択、意図と因果関係のわかるタイトルの付け方、要求分析、verification と validation の違い」

目的に合ったモデリング技術の修得には、目的に応じた教科項目の選択やウエイト付けが教育上重要になる。

参考文献

- [1] 大成幹彦：シミュレーション工学，オーム社，1993
- [2] 真鍋龍太郎・逆瀬川浩孝・若山邦紘：文科系のコ

表5 回帰分析自由課題のタイトルの改善例

No.	提出タイトル	改良後タイトル
1	中古車の相場	中古車価格の年数依存性
2	鉄道の運賃について距離と価格の関係を分析する	鉄道運賃の遠距離低減則/鉄道運賃の距離依存性
3	バスの走行距離の支払料金への影響	バス運賃の距離依存性
4	バーボンの熟成年数と価格の関係	バーボン価格は熟成年数に依存するか/価格に及ぼす年数の影響
5	マイクロプロセッサの性能を決める要素	MPUの整数演算性能を決めるトランジスタ数と内部動作周波数
6	ノートパソコンの動作速度と価格の相関について	ノートパソコンの価格は動作速度に左右されるか

ンピュータ／応用編 表計算ソフトの活用 岩波書店, 1988.4.13
 [3] 久米均, 飯塚悦功: 回帰分析, 岩波書店, 1993

[4] 水越敏行: 教育方法と評価の立場・観点から, 教育学関連学協会連合第4回大会 1994.10.8
 [5] 教育システム情報学会, Vol. 12, No. 1, p.11

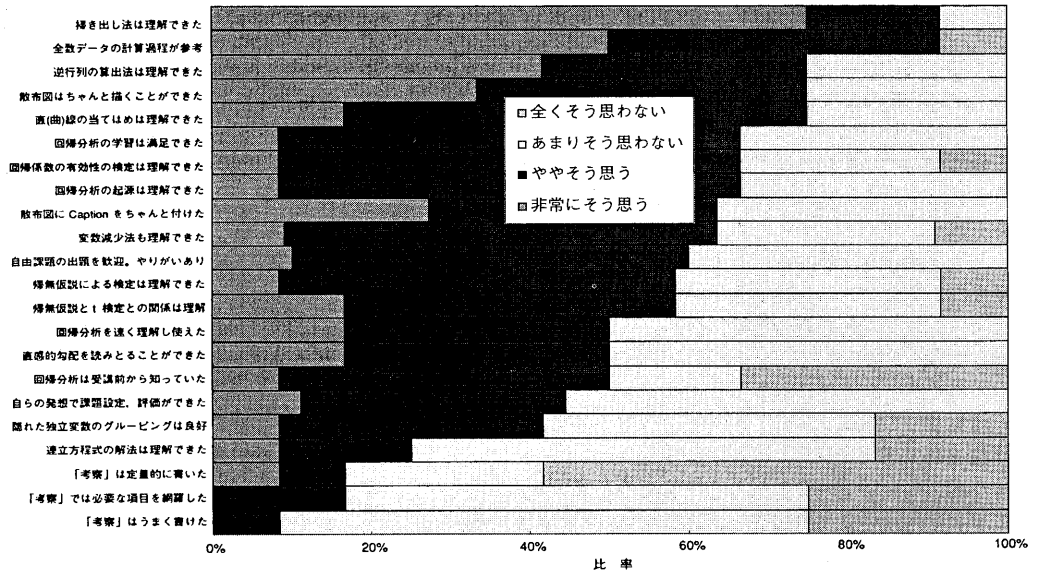


図1 回帰分析教育に関するアンケート結果 (掃き出し法実習なし) 12名

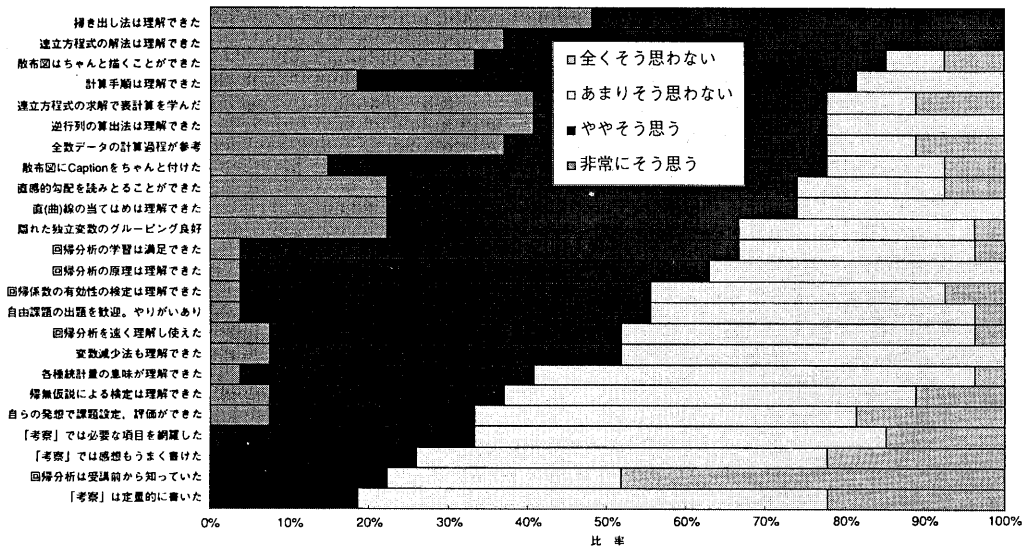


図2 回帰分析教育に関するアンケート結果 (掃き出し法実習あり) 27名