

意思決定支援のための3次元 データ・モデル可視化方法の検討

谷口 洋司 古賀 明彦 広瀬 正

(株)日立製作所システム開発研究所

データを用いたビジネス分野の意思決定プロセス（データの検証、データを用いた対象のモデル化）を、データ可視化技術を利用して支援する方式について述べる。ビジネス分野の意思決定においては、ノイズを多く含んだデータしか利用できない場合が多い。このような状況で意思決定を行うためには、データ要因間の関係を見出し、ユーザの持つ経験や知識を最大限に引き出し、データの情報不足を補う必要がある。本稿では、データ検証支援機能（3次元プロット表示・補助線表示・詳細情報表示・4面図表示）とモデル構築支援機能（モデルの3次元表示・モデルのアニメーション表示）を提案する。

3-D DATA AND MODEL VISUALIZATION METHOD FOR DECISION MAKING

Yoji Taniguchi Akihiko Koga Tadashi Hirose

Systems Development Laboratory,
Hitachi, Ltd.

Midousuji-S.G Bldg. 10F 3-6-1 Bakuroh-machi Chuo-ku, Osaka-shi, 541 Japan

This paper describes 3-D data and model visualization for decision making. In business fields, the data which we can use is very noisy in many cases, so it's difficult to find relations of the data and to make effective models for reasonable decision making. We propose following facilities to make good the lack of information by extracting user's experience and knowledge. (1)3-D plot display, (2)drawing assistance line, (3)query detail information, (4)4-plane figure, (5)model 3-D display, (6)model animation 3-D display.

1 はじめに

計算機利用形態は人の定型的な作業の支援や作業代替から、人の（意思決定や思考、発想などの）非定型的な作業の支援へと発展していく。そこでは、従来以上にインタラクティブな使い方に適した計算機システムが求められる。そのシステム構築の考え方は「人間中心」でなければならない。計算機システムに関するユーザインタフェース研究の課題の一つはこの「インタラクティブな環境の実現」である。

計算機パワーの飛躍的な増大で情報処理そのもののパワーについての制約は緩くなった。しかし、入出力、特に人との間の入出力処理の改善は未開発な状態である。入力側（計算機から見れば出力側）は人間の感覚にマッチした、人にとって親切な情報表示が求められる。出力側、すなわち情報表示技術には、データ可視化技術と呼ばれ、多量のデータベース内データの視覚的な表示、気流の流れや磁場の状態など不可視データの視覚的な表示が行われている¹。

一方、意思決定プロセスの一つに対象のモデル化がある。操作対象プロセスの性質を把握するプロセスで、より正確な（正しい意思決定に必要な）モデル化が求められる。経験的なデータ、過去の実績データなどからの

ような操作（意思決定）を行うとどのような結果が得られるのかを予測するためのモデルを生成する方法として重回帰分析などが知られている。また、新しい試みとしてニューラルネットワークの学習機能を用いる方法が提案されている²。

我々は3次元表示を利用することにより、ユーザが潜在的に持つ知識を引き出し、データに乗っているノイズの除去、データを説明する適切なモデルの構築、モデルを用いた意思決定を一貫して行うビジュアル意思決定支援システムを提案した³。本システムは、3次元表示を用いたデータ検証支援機能、モデル構築支援機能、計画機能によって、ユーザの意図や知識を引き出す点に特徴がある（図1）。本稿では、いくつかの実データへの適用を通じて検討した、データ検証支援・モデル構築支援機能におけるデータ・モデル可視化方法を提案する。

2 データ検証支援機能

2.1 データ可視化機能

データ可視化技術の目的は、科学技術計算結果を対話的に処理することにより、「自然現象を解明すること」であり、津波や嵐の現

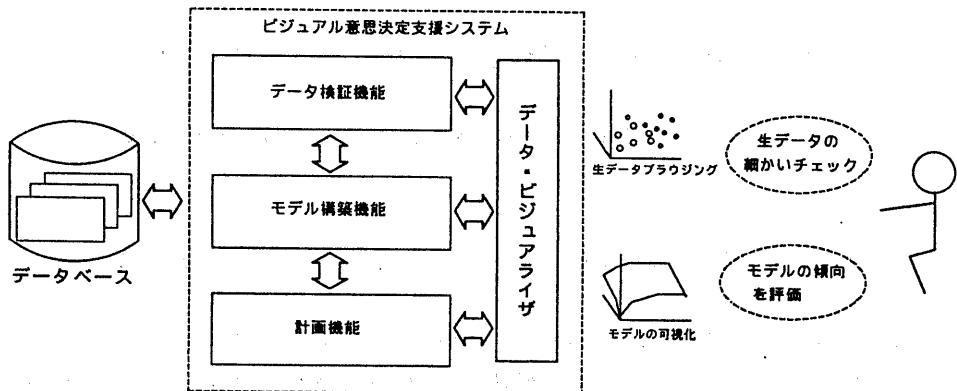


図1 ビジュアル意思決定支援システム

象、物質の破壊現象、あるいは熱の伝搬の可視化に適用されてきた。

ここで云うデータ可視化機能は、単にグラフィカルにデータを表示するだけでなく、ユーザの

- (1) 解析計算結果を対話的にさまざまな方法で見える
 - (2) いつでも任意にパラメータを変更でき、その都度再計算・再表示ができる
- といった操作を可能とするものである。これによって、ユーザは、
- (1) どのパラメータが重要なのかという問題解決の糸口をつかむ
 - (2) 正確に現象を把握する
 - (3) 多数のパラメータを総合的に調べて設計を最適化することができ、効率的にデータを評価できる。

2.2 データ検証支援機能

データを用いたビジネス分野の意思決定の場面においては、データ全体の傾向を掴むことだけでなく、個々のデータの占める位置、詳細を知りたいことがある。担当者にとっては、個々のデータそのものに関する知見や発想や理解があるのである。また、個々のデータを見つけないという欲求も生じる。

データ可視化（3次元表示）を利用して、個々のデータの占める位置を明らかにし、さらにそのデータの詳細情報を簡単に得られるようにすることにより、ユーザが潜在的に持つ知識を引き出し、発想を支援するデータ検証機能を検討した。ある製品の開発状況と納期の関係を表すデータの例（図2）を用いて説明する。

製品名	規模	重量	不具合案件	納期
A	123	56.3	0.66	201
B	99	39.1	0.92	148
C	87	45.7	0.85	193
.
.
.

図2 データの例

(1) データの3次元プロット表示

データを、任意の3項目（規模、重量など）を座標軸とした3次元空間上の点として、遠近感を出して3次元プロットし、リアルタイムの視点変更、座標軸変更などを可能とした（図3）。

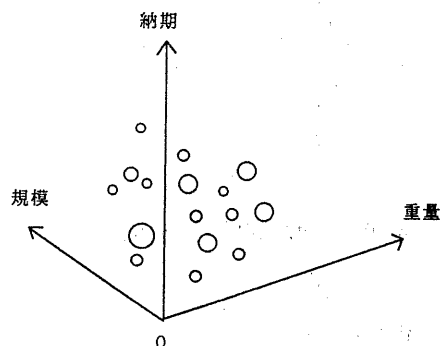


図3 データ3次元プロット表示

一般的に、数値データをグラフ化する場合、散布図にするか折れ線グラフにするか、またどのデータ項目をグラフにするかなど決定すべき要因は多い。最初から、目的がはっきりしている場合にはそれにあつた方法でグラフ化すればよいが、データをどのように利用するかを検討するような場合には、取り合えずそのままのデータを眺めて、データ項目間の関係を知る必要がある。

(2) 補助線表示

2次元グラフにおいては、各データの座標値をわかりやすくするために、目盛やグリッドなどが使用されるが、3次元表示では、目盛やグリッドが多くなると、ユーザにとって見づらくなる。

補助線表示機能では、目盛や補助線を指定された点ごとに表示する方式を採用した。即ち、3次元表示されたデータを表す点をマウスでピックアップすると、点からXY平面に降ろした垂線と、垂線とXY平面との交点からX軸Y軸にそれぞれ降ろした垂線が表示され、更に点の各軸の座標値が座標軸の近傍に表示される（図4）。

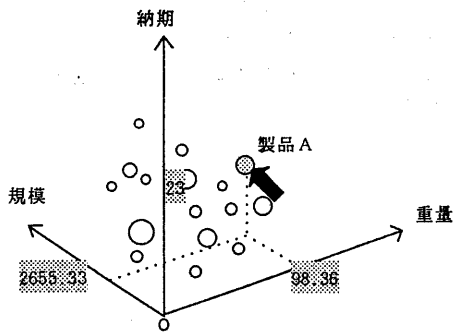


図4 補助線表示機能

(3) 詳細情報表示

データを3次元プロット表示することにより、いろいろな軸でデータを分析することができるが、XYZ軸に設定されている3つのデータ項目以外について比較することはできない。類似したデータと思ったデータであっても、他のデータ項目も類似しているとは限らない。

詳細情報表示機能では、指定された点ごとに詳細情報を表示する方式を採用した。即ち、3次元表示されたデータを表す点をマウスでピックアップすると、ウィンドウが開き、データ名（製品名）とすべてのデータ項目値が表示される（図5）。

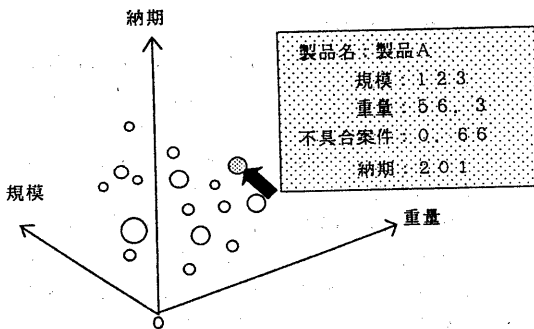


図5 詳細情報表示機能

(4) 4面図表示

データを3次元空間上に点として、3次元プロット表示している時に、各点の正確な位置関係を把握するためには、視点を変更（図形を回転）して、各座標軸の方向からデータを眺めればよいが、遠近感が出ているとかえって2つの座標（データ項目）間の関係を掴みにくいという問題がある。

4面図表示機能では、3次元プロット表示とともに、3つの座標軸方向からの2次元プロット表示（遠近感なし）を同時に行う方式を採用した。即ち、CADなどで用いられる3面図に加えて3次元プロットを同時に表示するのである。

4面図表示機能と上記補助線表示・詳細情報表示機能を併用すれば、データ間の関係やデータ項目間の関係をより多くの観点から同時に掴むことができる（図6）。

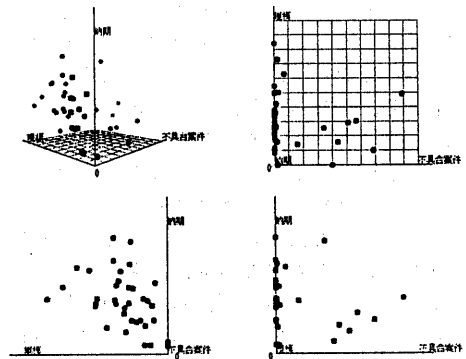


図6 4面図表示機能

2.3 データ検証機能の効果

データをプロットすることにより、突出したデータの検出ができ、さらに詳細情報をみることにより、そのデータがノイズデータなのか、特異データなのかの判断ができる。

また、例えば、「不具合案件の多い製品は納期が短い」（図7）といった、相関分析手法ではなかなか発見できないデータ項目間の関係が、データをプロットすることによって発見できる。さらにそのような関係を持ったデータを直接検証することができる。

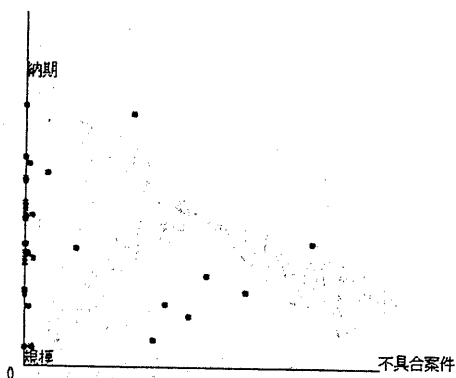


図7 不具合案件と納期の関係

3 モデル構築支援機能

3.1 モデル可視化へのアプローチ

経験的なデータ、過去の実績データなどからどのような操作（意思決定）を行うとどのような結果が得られるのかを予測するためのモデルを生成する方法として、重回帰分析やニューラルネットワークの学習機能などがある。

これらの手法で構築したモデルの評価には、モデルの予測値と実績データとの誤差の二乗平均値が利用される場合が多い。しかし、この方法だけでモデルの入出力特性を正確に把握することはできない。重回帰モデルは、パラメータの値によって、モデルの入出力特性をある程度知ることができるが、ニューラルネットワークは、学習パラメータ（ネットワークのウェイト）の値では、入出力特性が把握できず、ブラックボックスとみられることが多い。

我々は、モデルの入出力特性を把握するために、モデルを3次元表示することを検討し、ユーザの持つイメージとのマッチングを図った。

3.2 モデルの3次元表示

モデルの3次元表示機能では、モデルの2つの入力変数と1つの出力変数の関係を滑らかな面として表示する方式を採用した。即ち、

2つの入力変数をそれぞれの定義域内で一定の割合で変動させ、その時の出力変数の値を滑らかな面で補間して表示するのである。但し、入力変数が3つ以上の場合には、表示しない入力変数は任意の値に固定し、表示する2つの入力変数のみを変動させる。同じデータから構築した重回帰モデルとニューロモデルを3次元表示した例を図8に示す。入出力特性が大きく異なることがわかる。

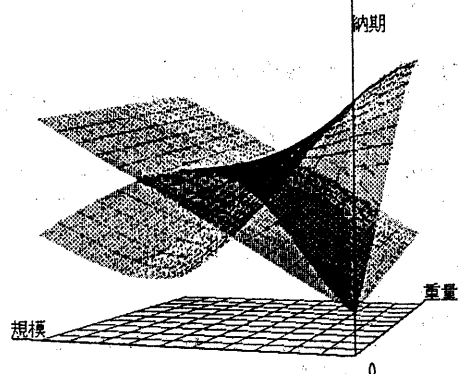


図8 重回帰・ニューロモデル

3.3 データ・モデルの同時表示

(1) モデルの次元とデータの次元

構築されたモデルが、データの間をいかにかにうまく説明（学習）しているかを確認するために、モデルとデータを同時に3次元表示することが考えられる。しかし、データのデータ項目が4つ以上の場合、データが表示されている3次元空間とモデルが表示されている3次元空間は別のものである。

このことを例1を用いて説明する。

(例1)

[データ]

データ項目:

「規模」「重量」「不具合案件」「納期」

3次元表示:

X軸=「規模」

Y軸=「重量」

Z軸=「納期」

[モデル]

入力変数:

「規模」「重量」「不具合案件」

出力変数:

「納期」

3次元表示:

X軸=「規模」

Y軸=「重量」
Z軸=「納期」
「不具合案件」=0.5

データは、X軸=「規模」、Y軸=「重量」、Z軸=「納期」、「不具合案件」=任意として3次元表示されるのに対し、モデルは「不具合案件」を0.5に固定して3次元表示されるので、たまたま「不具合案件」=0.5であるデータ以外は同じ3次元空間には表示できない。

(2) 同じ3次元空間での表示

データ・モデルの同時表示機能では、構築されたモデルが、データの関係をいかにうまく説明（学習）しているかを確認するために、モデルを表示している3次元空間上にデータを補正して表示する方式を採用した。例1のモデルを表示した3次元空間には、

X軸=「規模」
Y軸=「重量」
Z軸=「不具合案件0.5のときのモデルの予測値
+予測誤差」

但し、予測誤差=データの納期の値-モデルによる予測値とした、データを3次元プロットするのである。データを補正して表示した例（図9）と、そのまま表示した例（図10）を示す。

これにより、構築されたモデルが、データの関係をいかにうまく説明しているか、また、どのデータが突出したデータかを視覚的にユーザが確認することができる。特に、ニューロモデルに適用した場合には、ニューロモデルの入出力特性を把握でき、さらに教師データをどのように学習しているかを把握することができる。

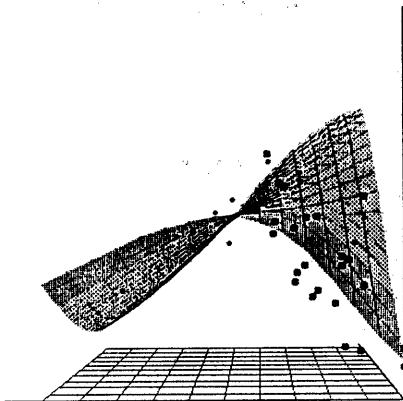


図9 データ・モデル同時表示（補正あり）

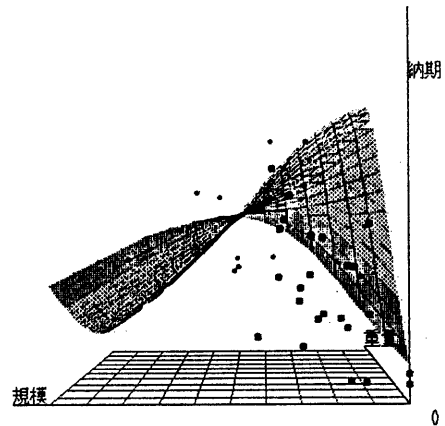


図10 データ・モデル同時表示（補正なし）

3.4 モデルのアニメーション表示

(1) ニューロモデルと学習誤差

良好なニューラルネットワークを構築する（学習させる）ためには、その学習回数が必要な要因となる。一般的には、教師データとニューラルネットワークの予測との誤差の二乗和を評価基準として、学習の収束状態を判断し、学習を終了させることが多い。しかし、この判断基準だけではニューラルネットワークが過度な学習によって教師データにオーバーフィッティングしてしまう可能性がある。

(2) ニューロモデルのアニメーション表示

モデルのアニメーション表示機能では、学習時（構築中）のニューロモデルの入出力特性を学習回数毎に表示する方式を採用した。即ち、学習によって刻々と変化していくニューラルネットワークの入出力特性を学習回数と共に表示するのである（図11）。これにより、ユーザはニューラルネットワークの学習の度合を判断し、必要に応じて学習を終了させることにより、過度な学習を防ぐことができる。さらに、誤差の二乗和による判断との併用により、良好なニューラルネットワークが構築できる。

4 おわりに

データを用いたビジネス分野の意思決定プ

ロセス（データの検証、データを用いた対象のモデル化）を、データ可視化技術を利用して支援する方式を検討した。

現場の担当者には、個々のデータそのものに関する知見や発想や理解があると考え、個々のデータの占める位置を明らかにし、データの詳細情報を簡単に得られるデータ検証機能を提案した。

また、データの間関係を説明するモデルの入出力関係を滑らかな面として3次元表示する方式を提案した。これにより、従来、ブラックボックスとされていたモデル（例えば、ニューラルネットワーク）の入出力特性を掴むことができる。

5 参考文献

- [1] 木崎：解析データ統合能力を強化するビジュアルイゼーション・ソフト、日経CG 1992.5 pp 97-106
- [2] David E. Rumelhart, James L. McClelland and PDP Research Group(甘利 俊一 監訳) :Parallel Distributed Processing(PDP モデルー認知科学とニューロン回路網の探索)産業図書(1989.2)
- [3] 谷口、他：三次元グラフィックスを用いたビジュアル意思決定支援システムの提案、第45回情報処理学会全国大会 7M-07
- [4] 原田：数値データから情報を引き出すデータ・ビジュアルイゼーション、日経メカニカル 1992.3.16 pp 20-34

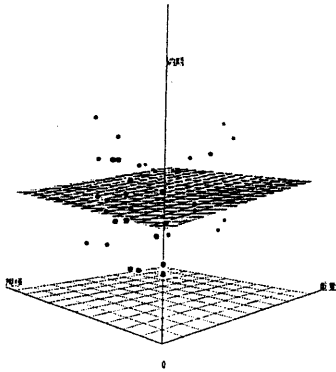


图11.1 20回学习

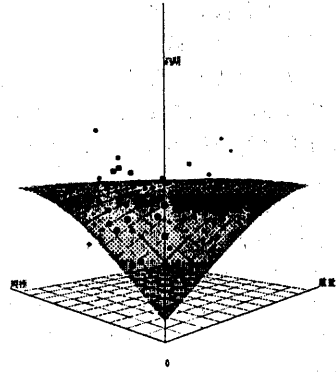


图11.4 200回学习

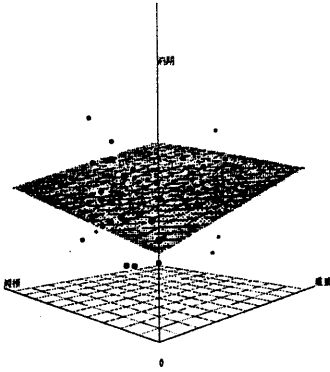


图11.2 50回学习

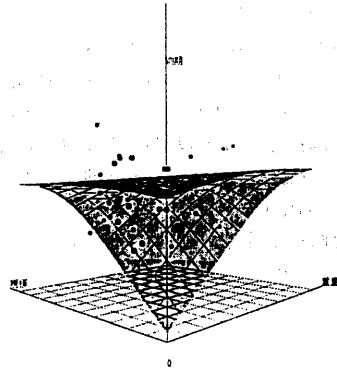


图11.5 300回学习

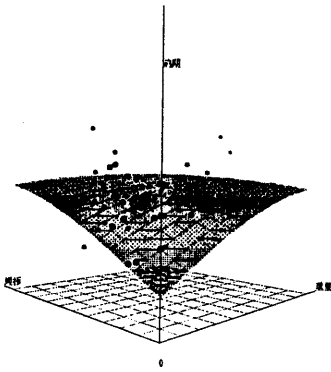


图11.3 100回学习

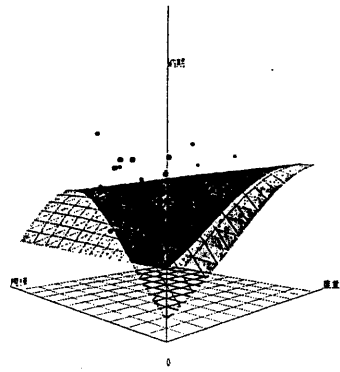


图11.6 500回学习