

## 多次元データモデリングソフトウェアの開発

荒川 正幹, 船津 公人\*

東京大学大学院 工学系研究科 化学システム工学専攻, 〒113 - 8656 東京都文京区本郷7丁目3番1号

\*e-mail: funatsu@chemsys.t.u-tokyo.ac.jp

(Received: October 27, 2003; Accepted for publication: February 20, 2004; Published on Web: May 10, 2004)

近年のコンピュータの進歩にともない、化学の分野においては日々多くの数値データが蓄積されている。そして、これらのデータから効率的に意味のある知識を抽出するために、多くのケモメトリックス手法が開発、提案されている。通常これらのデータは多次元/多変量であり、従来の統計ソフトウェアでは扱うのが困難であった。そこで、本研究ではこれらのデータを直接扱うことを目的とした、多次元データモデリングソフトウェアの開発を行った。本ソフトウェアを用いることにより、多次元の数値データを用いた解析が可能となる。ユーザは散布図、棒グラフ、折れ線グラフなどの様々なグラフを描くことによって、データの構造を視覚的に捉えることができる。また、Multi-way PLS 法や Kohonen ニューラルネットワークといったモデリング手法を用いて、より詳細な解析を行うことが可能である。

キーワード: 多次元データ, Multi-way PLS 法, Kohonen ニューラルネットワーク

### 1 はじめに

近年化学の分野においては、コンビナトリアルケミストリやハイスループットスクリーニング (HTS) の発達にともない、多くの情報が日々蓄積されている。これらの情報の中からいかにして意味のある知識を抽出できるかが、効率的な創薬の鍵を握っているといえる。また、構造活性相関 (QSAR)、構造物性相関 (QSPR) をはじめとするケモメトリックスに関する研究も盛んに行われており、数値データの有効活用の重要性が高まっている。

通常これらの数値データは多次元/多変量であり、既存の統計ソフトウェアでは取り扱うことが困難である。そこで本研究では、このような場面において多次元データの操作、解析、グラフ描画などを行うことを目指し、多次元データモデリングソフトウェアの開発を行った。本システムでは、通常の表計算ソフトウェアで取り扱われるような行列データに加え、多次元配列データを扱うことが可能である。また、様々な形式でグラフを描画することによって、このようなデータの構造を視覚的に捉えることが可能となっている。さ

らに、これらのデータに対し、Multi-way PLS 法による回帰分析、Kohonen ニューラルネットワークの学習などを行い、より詳細な解析を行うことも可能である。

これらのケモメトリックス手法を化学データへと応用した研究として、文献 [1, 2] を挙げるができる。

文献 [1] は、CoMFA 法による 3D-QSAR 解析において、適切な配座および重ね合わせルールを推定するために、Multi-way PLS 法を用いたものである。はじめに CoMFA フィールド行列を、異なる配座および重ね合わせルールを用いて複数計算する。次にこれらの行列を結合し、説明変数となる配列を作成する。1つの CoMFA フィールドが行列、つまり 2 次元の配列であるため、説明変数配列はこれに配座、重ね合わせルールの次元を加えた 4 次元の配列となる。そして、この 4 次元配列を説明変数、生物学的活性値を目的変数として Multi-way PLS 法による回帰分析を行う。その結果得られる回帰係数を解析することにより、活性を説明するのに最も適した配座および重ね合わせルールを推定することができる。最後に選択された配座と重ね合わせルールを用いて通常の CoMFA を行い、3D-QSAR モデルを作成する。文献 [1] では、この解

析手順を cholecystokinin-B (CCK-B) の拮抗剤である benzodiazepine 誘導体へと適用し、良好な結果を得ている。

文献 [2] は、Kohonen ニューラルネットワークおよび Multi-way PLS 法を応用することにより、分子表面上の静電、疎水ポテンシャルを用いた 3D-QSAR 解析を行った例である。分子表面における静電、疎水ポテンシャルはその化合物の生物学的な活性を説明するために極めて重要であると考えられる。しかし、これらのポテンシャルと生物学的活性値との間で QSAR モデルを構築しようとする場合、各分子構造によって分子表面の形状が異なることが問題となる。この手法では、Kohonen ニューラルネットワークを利用することによって、この問題点を克服している。まずサンプル化合物のなかから 1 つの分子構造をテンプレート分子として選択し、その分子表面の形状を Kohonen ニューラルネットワークにより 2 次元マップへと写像する。つまり、分子表面上に一定間隔で設定されたサンプリング点の 3 次元座標を入力ベクトルとし、Kohonen ニューラルネットワークによる非線形写像を行う。この操作により分子表面はそのトポロジカルな関係を維持しつつ、2 次元のマップへと写像されることとなる。そしてこの発火情報をもとにして Kohonen マップと同じサイズの行列に、分子表面上の静電、疎水ポテンシャルをそれぞれ写像する。こうして作成されたマップを 1 つの配列に結合したものを説明変数として用いることにより、3D-QSAR 解析が可能となる。説明変数配列は 4 次元の配列となるため、モデリングには Multi-way PLS 法を使用している。そして、回帰モデルの回帰係数値を分子表面上に逆射影することによって、活性にとって重要であると思われる領域を特定することが可能となる。この手法を 36 のエストロゲン拮抗剤を用いた 3D-QSAR 解析に適用し、良好な結果を得ている [2]。

上記 2 つの研究においては、いずれも多次元配列を取り扱う必要があり、Kohonen ニューラルネットワーク、Multi-way PLS 法を用いることでこれらを可能にしている。1 つ目の配座、重ね合わせルール推定手法では、Multi-way PLS 法を用いることによって、配座と重ね合わせルールを同時に推定することが可能である。これらは互いに密接に関係しているため、同時に取り扱うことが重要であり、そのためには Multi-way PLS 法などの多次元配列を扱う手法が不可欠である。2 つ目の分子表面上の静電、疎水ポテンシャルを用いた 3D-QSAR 解析においても、Multi-way PLS 法を用

いることで、静電ポテンシャルと疎水ポテンシャルの両方を考慮したモデルを作成することに成功している。また、分子表面の形状の違いを取り除くために、Kohonen ニューラルネットワークによる非線形写像を行っている。これらのケモメトリックス手法は非常に有益なものであるが、まだ一般的な統計ソフトウェアには実装されていないのが現状である。そこで本研究では、このような解析を GUI によって簡便に行うことができるシステムの開発を行った。

## 2 手法

本システムでは、多次元配列データの操作や各種グラフ描画による多次元情報の可視化の他、PLS 法、Multi-way PLS 法、Kohonen ニューラルネットワークなどのケモメトリックス手法を用いたデータ解析が可能である。以下では本システムに実装されている解析手法および本システムで用いる多次元配列のためのファイル形式について説明する。

### 2.1 PLS 法、Multi-way PLS 法

PLS 法はケモメトリックスの分野などで広く用いられている線形重回帰分析手法である。説明変数  $X$ 、目的変数  $Y$  から、それぞれ  $T$ 、 $U$  と呼ばれるスコアを抽出しモデリングを行う。その基本式を以下に示す。

$$\begin{aligned} X &= TP' + E \\ Y &= UQ' + F \end{aligned} \quad (1)$$

ここで  $P$ 、 $Q$  はローディング、 $E$  と  $F$  は残差である。PLS 法は、通常の最小 2 乗法による線形重回帰と比較し、頑健で予測的なモデルが得られるという特徴を持っている。また、サンプル数よりも記述子数の方が多い状況であってもモデルが構築できるという利点もある。線形重回帰分析手法であるため、上記の式を変形することにより、以下の回帰式を得ることが可能である。

$$Y = XB + E \quad (2)$$

ここで  $B$  は回帰係数、 $E$  はモデル残差である。PLS 法においては目的変数の数は複数であってもよい、 $Y$  および  $B$  はベクトルではなく行列となる。目的変数が 1 つ、つまり  $Y$  が 1 次元の列ベクトルである場合は  $B$  も列ベクトルとなる。

通常の PLS 法による回帰分析においては、説明変数  $X$  および目的変数  $Y$  は 2 次元の配列、つまり行列で表現される。しかし、扱うべきデータはときとして 3 次元以上の多次元配列となる。このような場面において安定した回帰モデルを得るための手法として提案されたのが、Multi-way PLS 法 [3] である。これは PLS 法における  $X$ 、 $Y$  を 3 次元以上の多次元配列へと拡張した手法であり、データが本質的に多次元の性質を持つ場合には、通常の PLS 法に比べ頑健なモデルが得られることが知られている。以下に Multi-way PLS 法の基本式を示す。

$$\begin{aligned} \underline{X} &= T \otimes \underline{W}' + \underline{E} \\ \underline{Y} &= U \otimes \underline{C}' + \underline{E} \end{aligned} \quad (3)$$

$\underline{X}$  は説明変数、 $T$  は  $\underline{X}$  から抽出されたスコア、 $\underline{W}$  はウェイト、 $\underline{E}$  が誤差である。 $\underline{Y}$ 、 $U$ 、 $\underline{C}$ 、 $\underline{E}$  についても同様である。 $\underline{X}$ 、 $\underline{Y}$  などのアンダーラインは、これらが行列ではなく、多次元の配列であることを示している。Multi-way PLS 法は通常の PLS 法を一般化したものとみなすことができ、 $\underline{X}$  および  $\underline{Y}$  が行列である場合には、Multi-way PLS 法と通常の PLS 法は全く同じ結果を与える。回帰式についても通常の PLS 法と同様に、以下の式によって表現される。

$$\underline{Y} = \underline{X} \otimes \underline{B} + \underline{E} \quad (4)$$

回帰係数  $\underline{B}$  も同様に多次元配列となる。例えば説明変数が 4 次元の配列、目的変数が 1 次元の列ベクトルである場合、回帰係数は 3 次元配列となる。

## 2.2 Kohonen ニューラルネットワーク

Kohonen ニューラルネットワーク [4] は Kohonen によって 1982 年に提案された、教師なし学習を行うニューラルネットワークの一種である。多次元空間の情報を、トポロジーを保持したまま 2 次元のマップ上へと非線形写像する能力を持つことから、多変量データの次元縮約やクラスタリングを行うためのツールとして応用されている。Figure 1 に Kohonen ニューラルネットワークの概要図を示す。ネットワークは入力層と出力層からなっており、出力層の 2 次元平面上にはニューロンが並べられている。まずサンプルが入力層を通してネットワークへと入力され、勝者ニューロンの探索が行われる。出力層上の各ニューロンの重みベクトルと、入力ベクトルとの間で次式によって距離が計算され、最も近いニューロンが勝者ニューロンとして選択される。

$$d_j = \sqrt{\sum_{i=1}^m (x_i - w_{ij})^2} \quad (5)$$

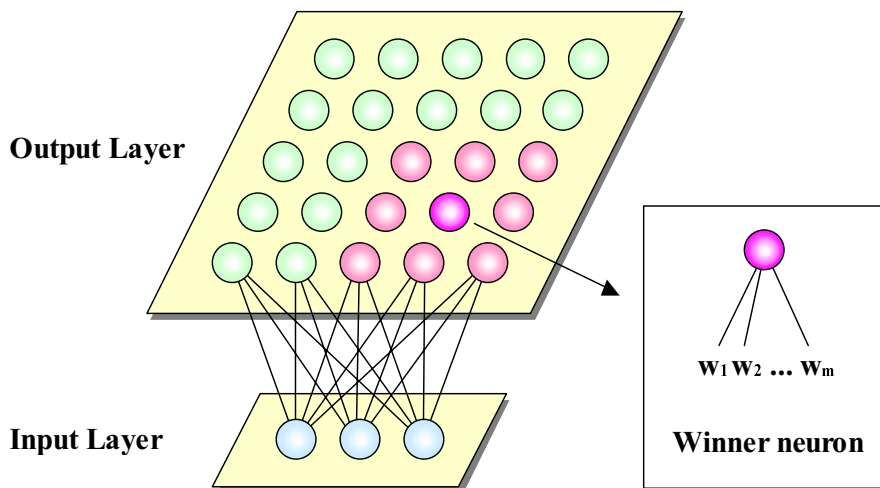


Figure 1. Structure of KNN.

ここで  $x_i$  は入力ベクトルの  $i$  番目の値、 $w_{ij}$  はニューロン  $j$  の重みベクトルの  $i$  番目の値、 $m$  は記述子の数、 $d_j$  はニューロン  $j$  と入力ベクトルとの間の距離を示す。そして選択された勝者ニューロンとその近傍のニューロンの重みベクトルを、入力ベクトルへと近づく方向へと修正する。

$$w_{ij^*}(t+1) = w_{ij^*}(t) + \alpha[x_i - w_{ij^*}(t)] \quad (6)$$

ここで、 $j^*$  は勝者ニューロン、 $t$  は学習回数、 $\alpha$  は学習係数を示す。すべての入力ベクトルについてこの操作を行うことで1世代の学習が完了する。そして、 $\alpha$  の値を減少させながら十分収束するまでこれを繰り返すことにより、自己組織化学習が行われる。その結果、入力空間において近い距離にあるサンプルは出力層上においても近い位置へ、遠くにあるサンプルは遠い位置へと写像される。つまり、トポロジカル情報を保持したまま、入力空間のサンプルが2次元マップへと非線形写像される。

### 2.3 多次元配列のためのファイル形式

統計ソフトウェアなどで一般に用いられている csv などのファイル形式では、多次元配列の情報を取り扱うことが困難である。そこで、本システム独自のファイル形式 (mdd 形式) を定義した。mdd ファイルは複数の多次元配列を格納することが可能なテキストファイルである。ファイルの先頭に、格納されている配列の個数が記述され、それに続いて各配列のデータが記述される。各配列の情報は、配列の名前、次元数、各次元の要素数、各次元の名前、各次元の各要素の名前、数値データの順に記述される。mdd ファイル形式を用いることにより、本システムの入出力となる複数の多次元配列を、1つのファイルへと格納することが可能となる。

## 3 実行例

以下に本システムの実行例を示す。1つ目は Kohonen ニューラルネットワークの学習を行った例、2つ目は Multi-way PLS 法を用いた配座、重ね合わせルール選択手法 [1] の計算例である。

	0	1	2	3	4	5	
マップ_縦	0	1.412200	1.379900	1.341400	1.102500	-0.667900	-0.803600
マップ_横	1	1.421700	1.369700	1.328900	1.217700	-0.659400	-0.860600
Fe	2	1.401900	1.393300	1.401400	1.141400	-0.601200	-0.817900
None	3	1.368000	1.373100	1.279800	0.952400	-0.365000	-0.769100
None	4	1.043600	1.281400	0.265300	0.185900	-0.268700	-0.418600
None	5	-0.336600	-0.315100	-0.394600	-0.345800	-0.300200	-0.225800
None	6	-0.385000	-0.404200	-0.378400	-0.352100	-0.290100	-0.576000
配列_1	7	-0.312200	-0.343400	-0.388100	-0.379100	-0.385500	-0.982500
発火位置	8	0.666700	0.143100	-0.318800	-0.401600	-0.459900	-1.015700
重みベクトル_X	9	1.306300	0.931300	0.479900	-0.200900	-0.691700	-0.813800
	10	1.319700	1.092200	1.129800	0.500700	-0.623100	-0.779200
	11	1.268300	1.234100	1.189200	1.130800	-0.557800	-0.759000

Figure 2. Main window of the software.

Figure 2 に本システムのメインウィンドウを示す。図左下のリストボックスは、配列一覧部であり、現在システムに読み込まれている配列の名前が一覧表示される。ここでは“配列\_1”と、それを用いて Kohonen ニューラルネットワークの学習を行った結果である“発火位置”、“重みベクトル\_X”の3つの配列が読み込まれている。この配列一覧部において選択されている配列の次元情報が左上のコンボボックス群によって示され、そこで任意の2つの次元を指定することによって配列内の値がウィンドウ右側の行列表示部に表示される。

ここでは、配列“重みベクトル\_X”を例として説明を行う。この配列は  $10 \times 10 \times 10$  の3次元配列であり、各次元の名前は、マップ\_縦、マップ\_横、記述子となっている。それぞれの次元の各要素には以下のラベルがつけられている。

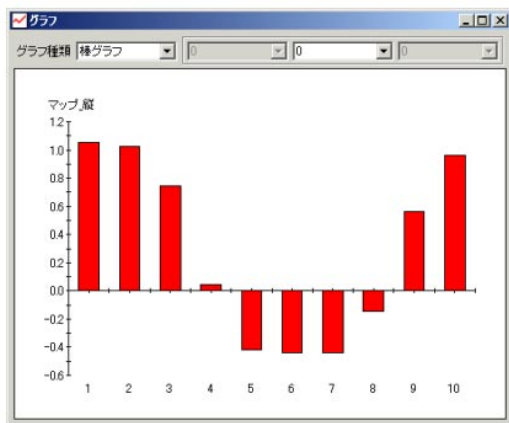
マップ\_縦 : 0, 1, 2, 3, 4, 5, 6, 7, 8, 9

マップ\_横 : 0, 1, 2, 3, 4, 5, 6, 7, 8, 9

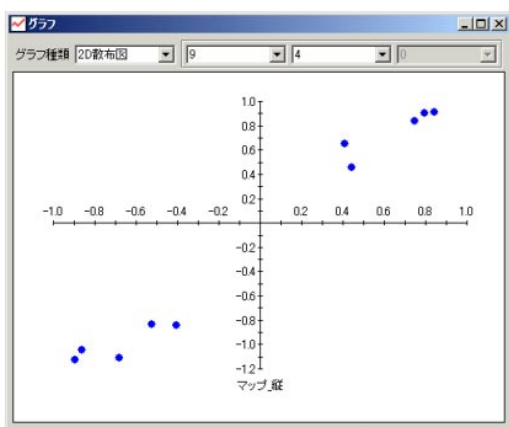
記述子 : Fe, Ti, Ba, Ca, K, Mn, Rb, Sr, Y, Zr

ウィンドウ左上の次元情報を格納するためのコンボボックスは5つ用意されており、5次元までの配列を取り扱うことが可能である。ここでは、配列が3次元であるため、上の3つのみが使われており、残りの2つについては None と表示されている。各コンボボックスには先頭に次元名が格納され、それに続いて各要素のラベルが格納される。例えば1つ目のコンボボックスの内容は以下ようになる。

マップ\_縦, 0, 1, 2, 3, 4, 5, 6, 7, 8, 9



(a) Bar graph



(b) 2D scatter plot

Figure 3. Graph examples.

これらのコンボボックスを用いて多次元配列の任意のスライスを指定することによって、その値をウィンドウ右側の行列表示部に表示する。固定する次元についてはその要素ラベルを、その他の次元については次元の名前をコンボボックスで指定する。この例では、3つ目の次元を Fe に固定しており、右側の行列では、行がマップ\_縦に、列がマップ\_横に対応している。このような仕組みによって、多次元配列の任意のスライスを指定し、行列状のデータとして画面上で確認することが可能となる。また表示されたスライスを csv ファイルとして書き出すことも可能である。

また、これらのデータを用いてグラフを描画することも可能である。Figure 3 に本システムによって描画したグラフの例を示す。本システムでは、棒グラフ、折れ線グラフ、2次元/3次元散布図、曲面グラフの描画が可能である。各グラフはマウスにより回転、拡大、縮小、移動などの操作が可能である。



Figure 4. Configuration dialog of Kohonen neural network.

Figure 4 に Kohonen ニューラルネットワークの設定ダイアログを示す。このダイアログにおいて、計算条件を設定し、計算開始ボタンを押すことによって、Kohonen ニューラルネットワークの学習を行うことができる。設定可能な計算条件は、マップサイズ、学習回数、学習範囲、学習率の値と、トラスマッピング、オートスケーリング、データシャッフルのオン/オフである。

マップサイズは出力層の一辺のニューロン数、学習回数は学習を繰り返す数である。学習範囲は重みベクトルの修正を行う際、勝者ニューロンの周囲のどれだけの範囲のニューロンを修正するかを指定する値である。例えば学習範囲が 5 であれば、各学習において第 5 近傍までのニューロンが入力ベクトルへ近づく方向へと修正される。学習率は式 (6) における  $\alpha$  の初期値であり、学習回数とともに減少していく。設定可能な値の範囲は 0 から 1 である。大きな値を指定すれば少ない学習回数で収束するが、大きすぎると適切な写像が行われない可能性がある。トラスマッピングの設定をオンにすると、出力層としてトラス状の平面が用いられる。トラスマッピングでは、マップの上辺と下辺、右辺と左辺がそれぞれ隣接しているものとして、近傍範囲の計算が行われる。これにより、多次元空間から 2次元平面への写像が無理なく行われることが多いため、Kohonen ニューラルネットワークでは一般的にトラスマッピングが用いられる。オートスケーリングの設定をオンにすると、入力サンプルの各記述子が平均 0、分散 1 となるようにスケールされる。これにより、各記述子の平均値や分散の違いに影響を受けないモデリングが可能となる。記述子間で単位が異なる場合など、その平均値や分散に特に意味がない場合には、通常前処理としてオートスケーリングを行う。データシャッフルの設定をオンにすると、

入力サンプルの学習の順番をランダムにすることができる。Kohonen ニューラルネットワークの学習では、先に入力されたサンプルよりも、後に入力されたサンプルの影響が大きくなる事が知られているが、この設定を用いることにより、学習の順番による影響を少なくすることが可能である。

学習条件の設定を行った後、計算開始ボタンを押すことにより Kohonen ニューラルネットワークの学習が行われ、計算結果が自動的にシステムに読み込まれる。計算結果は、“発火位置”、“重みベクトル”という2つの配列として得られる。“発火位置”は2次元の配列であり、それぞれの入力サンプルがマップ上のどの位置に発火するのかが示したものである。各サンプルについて、それが発火するニューロンの X 座標と Y 座標およびそのときの重みベクトルと入力ベクトルとの誤差の値が格納されている。“重みベクトル”は3次元の配列であり、各記述子に対する重みの値が格納されている。1つのマップが2次元であるため、これに記述子の次元が加わり3次元の配列となる。得られた重みベクトルを用いて Figure 5 に示すような曲面グラフを描くことによって、マップを視覚化することができる。重みの値が面の高さによって表現されている。この図を見ることによって Kohonen マップの様子を視覚的に捉えることが可能である。また、この図はマウスにより、移動、拡大、縮小などの操作が可能である。さらに、色やフォント、座標軸などについての様々な設定を行うことも可能である。

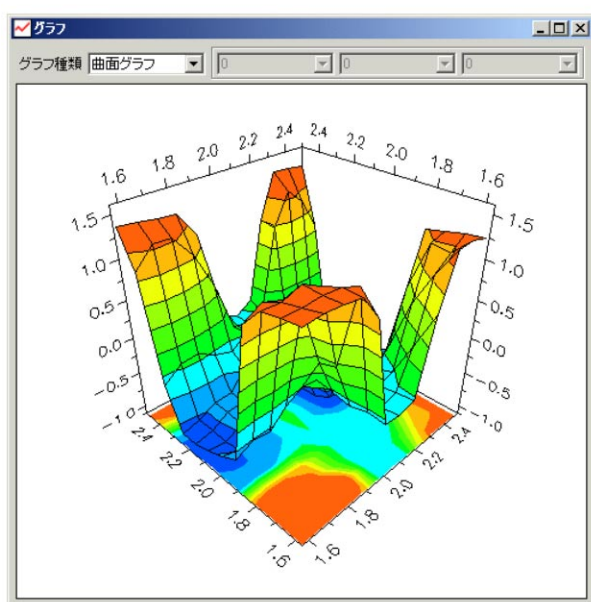


Figure 5. Example of surface graph.

2つ目の例として、Multi-way PLS 法を用いた配座、重ね合わせルール選択手法 [1] の計算例を示す。CCK-B の拮抗剤である benzodiazepine 誘導体 16 化合物について、14 種類の配座および 3 種類の重ね合わせルールを用いて CoMFA フィールド値の計算を行い、説明変数  $X$  を作成した。説明変数は4次元の配列であり、サイズは  $16$  (化合物数)  $\times$   $5096$  (記述子数)  $\times$   $14$  (配座数)  $\times$   $3$  (重ね合わせルール数) である。この計算は、本研究室で別途開発されたソフトウェアによって行い、mdd ファイルを介して、システムへと入力された。次に生物学的活性値を目的変数として、Multi-way PLS 法による回帰分析を行った。Figure 6 に Multi-way PLS 法の設定ダイアログを示す。

説明変数、目的変数のコンボボックスには、現在システムに読み込まれている配列の一覧が表示される。そのなかから説明変数、目的変数として用いる配列を選択し、計算を行う最大成分数、センタリング、スケールリング、クロスバリデーションの有無を指定した後、計算開始ボタンを押すことにより Multi-way PLS 法による回帰分析が行われる。ここでは、センタリング、クロスバリデーションを行い、3成分までの計算を行った。計算結果として、“R<sup>2</sup> 乗”、“回帰係数”、“スコア-T”、“Q<sup>2</sup> 乗”、“予測値”の配列が得られ、メインウィンドウの配列一覧部へ追加される。

次に回帰係数の2乗和の比率を計算し、活性配座および適切な重ね合わせルールの推定を行った。説明変数が4次元の配列であるため、得られる回帰係数は3次元の配列であり、そのサイズは  $5096$  (記述子数)  $\times$   $14$  (配座数)  $\times$   $3$  (重ね合わせルール数) となる。この配列を用いて文献 [1] の定義に従い、下記の式によって WC と WA の2つの指標を求めた。

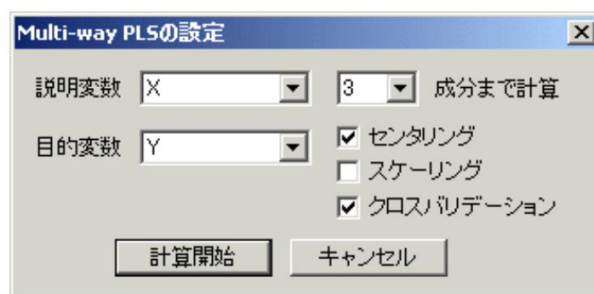


Figure 6. Configuration dialog of Multi-way PLS.

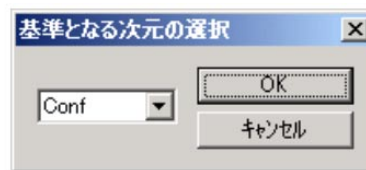
$$WC(l) = \sum_i \sum_k^{field\ align} (B_{ilk}^2) / \sqrt{\sum_j^{conf} \left[ \sum_i \sum_k^{field\ align} (B_{ijk}^2) \right]^2} \quad (7)$$

$$WA(l) = \sum_i \sum_j^{field\ conf} (B_{ijl}^2) / \sqrt{\sum_k^{align} \left[ \sum_i \sum_j^{field\ conf} (B_{ijk}^2) \right]^2} \quad (8)$$

この値が最大となるものが、活性配座および適切な重ね合わせルールであると考えられる。Figure 7にWCの計算例を示す。Multi-way PLS法によって得られた回帰係数の配列を選択した状態で、メニューから[2乗和の比率を計算]を選択することにより(a)のダイアログが表示される。そして、基準とする次元を指定し計算を行うことにより、(b)に示すような結果が得られる。

#### 4 まとめ

多次元/多変量の数値データを効果的に処理することを目標に、多次元データモデリングソフトウェアの開発を行った。数値データをもとにしてグラフを描くことにより、的確にデータの様子を把握することが可能である。また、Multi-way PLS法、Kohonenニューラルネットワークなどのモデリング手法を利用することによって、詳細な解析を行うことができる。本システムは今後Web上で一般に公開することを予定している。



(a) Configuration dialog.

Conf	比率
Conf_1	0.183169
Conf_2	0.266881
Conf_3	0.312619
Conf_4	0.192903
Conf_5	0.327651
Conf_6	0.196827
Conf_7	0.289515
Conf_8	0.178442
Conf_9	0.356288
Conf_10	0.206942
Conf_11	0.306280
Conf_12	0.309801
Conf_13	0.220924
Conf_14	0.300280

(b) WC values.

Figure 7. Calculation of WC.

#### 参考文献

- [1] K. Hasegawa, M. Arakawa, K. Funatsu, *Comput. Biol. Chem.*, **27**, 211-216 (2003).
- [2] K. Hasegawa, M. Arakawa, K. Funatsu, *Comput. Biol. Chem.*, **27**, 381-386 (2003).
- [3] R. Bro, *J. Chemom.*, **10**, 47-61 (1996).
- [4] T. Kohonen, *Self-Organizing Maps*, 3rd Ed., Springer (2000).

# Development of Multi-Dimensional Data-Modeling Software

Masamoto ARAKAWA and Kimito FUNATSU\*

Department of Chemical System Engineering, The University of Tokyo  
Hongo 7-3-1, Bunkyo-ku, Tokyo 113-8656, Japan  
*\*e-mail: funatsu@chemsys.t.u-tokyo.ac.jp*

In the domain of chemistry, many experimental data are being accumulated by combinatorial chemistry and high throughput screening (HTS) and so on. In order to obtain benefit information from these data, a lot of methodologies of chemometrics have been developed. Since usually these numerical data are represented as multi-dimensional arrays, traditional statistical software cannot handle these data directly. In this study, we developed data modeling software, which handles multi-dimensional arrays effectively. By using this software we are able to analyze multi-dimensional data. It is possible to understand the structure of multi-dimensional data by using graphical representation such as bar graph, line graph and so on. And the data modeling method, Multi-way PLS (partial least squares) and Kohonen neural network are available for the analysis of multi-dimensional data.

**Keywords:** Multi-dimensional data, Multi-way PLS, Kohonen neural network