

演 題	ケモトリックス総合システムChemishの開発	
発 表 者 ( 所 属 )	橋大樹、荒川正幹、竹内英憲、船津公人 (豊橋技術科学大学) 西村竜一 (三井化学株式会社) 神村基和 (アイセロ化学)	
連 絡 先	〒441-8580 愛知県豊橋市天伯町雲雀ヶ丘 1 - 1 電話番号 : 0532-44-6879 Fax 番号 : 0532-47-9315 E-mail : funatsu@tutkie.tut.ac.jp	
キ ー ワ ー ド	Chemometrics, PCA, PLS, Clustering, Neural Network, Quality Control	
開 発 意 図 適 用 分 野 期 待 効 果 特 徴 など	PCA, PLS, 変数選択、ニューラルネットワーク、クラスタリングなどのケモトリックス計算手法に加え、品質管理 (Quality Control) を行うためのツールが利用できる。操作は GUI を通して簡単に行えるとともに、付属のチュートリアルにより統計解析の初心者であっても短期間に本ソフトウェアを使用できるように配慮されている。	
環 境	適 応 機 種 名	IBM-PC/AT 互換機
	OS 名	Windows95 以上
	ソース言語	C++
	周 辺 機 器	特になし
流 通 形 態 ( 右 の い ず れ か に を つ け て く だ さ い )	<ul style="list-style-type: none"> <li>・日本コンピュータ化学会の無償利用ソフトとする</li> <li>・独自に頒布する</li> <li>・ソフトハウス、出版社等から市販</li> <li>・ソフトの頒布は行なわない</li> <li>・その他                      ・未定</li> </ul>	具 体 的 方 法
		<a href="http://www.quebec.tutkie.tut.ac.jp/cac/">http://www.quebec.tutkie.tut.ac.jp/cac/</a> から試用版のダウンロードが可能。完全版は CAC フォーラムにて配布。

## 1. はじめに

コンピュータが化学の分野で使われるようになってきた 1980 年代後半より、大量の化学データの獲得とデータの処理が可能になってきた。それにともない、大量のデータを数学的、統計学的に処理するための計量化学 (Chemometrics : ケモトリックス) の分野が発展し、さまざまな手法が開発されてきた。これらの手法は化学のあらゆる分野の研究者にとって非常に有用で興味を引くものである。Chemish はケモトリックスに興味を持つ研究者が気軽に利用することができるように、GUI による操作の簡便性、データ処理のしやすさ、結果の表示の見やすさなど配慮して開発されたケモトリックス総合システムである。Chemish を使用することにより、主成分分析、線形重回帰分析、PLS、変数選択、ニューラルネットワーク、クラスタリングなどによるデータ解析をよりの確に短時間で利用することが可能となる。なお、Chemish にはチュートリアルが付いており、操作方のみならず解析を進める上でのポイントなども併せて解説している。図 1 に Chemish の実行画面例を示す。

## 2. Chemish で利用できる計算手法

Chemish では、以下のようなケモメトリックス手法を試用することができる。

- ・ 重回帰分析 (MLR)
- ・ PLS, QPLS
- ・ 主成分分析 (PCA)
- ・ 各種ニューラルネットワーク

Back Propagation

Kohonen

Counter Propagation

- ・ 階層型クラスタリング
- ・ 逆解析

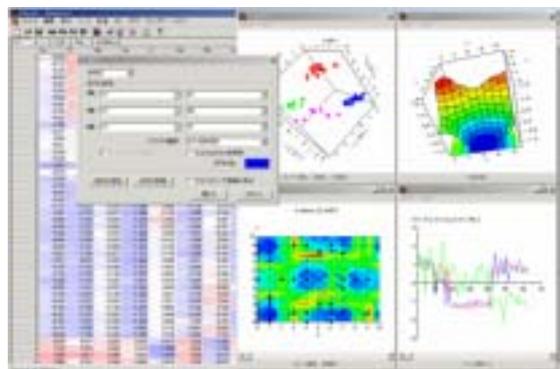


図 1 : Chemish の実行画面

## 3. モデルの評価、予測および逆解析機能

Chemish では、様々な手法でモデリングを

行ってできたモデルに対して、良好なモデルができたかどうか検定を行うことができる。説明変数  $R^2$ 、予測的説明分散  $Q^2$ 、目的変数の実測値と計算値の相関関係などを見ることでモデルの評価を行う。

また、得られたモデルを使用して、様々な値により目的変数値の予測を行うことができる。図 2 は Back Propagation 学習によって得られたモデルを使用して予測を行った例である。説明変数に様々な値を入力することによって、それぞれに対する予測値が得られる。また、このような手動入力だけではなく、ファイル入力による一括予測も可能である。



実測値	説明変数	H0>	H2>	
42000	勉強	-1.549	0.761	
57000	睡眠	0.369	0.518	
32000	テレビ	0.807	0.310	
9000	授業態度	0.315	0.428	
100000	前回	-1.177	1.293	
	バイアス	0.017	0.245	
予測値	目的変数	H0>	H2>	バイアス
85.137	点数	-1.705	1.270	0.131

図 2 : 予測画面

さらに、上の操作で得られたモデルをもとに、逆解析が可能となっている。逆解析とはモデル式の説明変数に値を入れて目的変数値 (予測値) を得る通常の予測とは逆に、目的変数に目的とする値 (たとえば物性値) を入れ、それを満足する説明変数の組 (可能な候補材料のパラメータセット) を創出する操作である。目的物性値や制約条件を入力し、説明変数 X の多次元空間内の点を指定した細かさで探索することで、目的物性値を満足する説明変数 X を求められる。

この他、PLS、QPLS については遺伝的アルゴリズム、MLR では STEPWISE 法による説明変数の最適化を行うことが可能である。Back Propagation ニューラルネットワークでは、クロスバリデーションによる学習回数および中間層のニューロン数それぞれの最適化が可能となっている。

## 4. QC (品質管理) ツール

QC (品質管理) は、製品の大量生産における品質の維持、向上のために重要な課題である。QC を行う場合、製品全体の状態を観測するために、ヒストグラム、パレート図、各種管理図などを利用する。Chemish では、これらの図を容易に描画できるようにするため、QC 計算ツールが利用できる。

## 5. 参考文献

- [1] 宮下芳勝;佐々木慎一,ケモメトリックス,共立出版(1995)
- [2] 中村速男,管理図の作り方と活用,日本規格協会(1984)