ソフトウェア評-UNSCRAMBLER 7.01

相 島 鐵 郎*

Software review : UNSCRAMBLER 7.01

Tetsuo AISHIMA*

* Kikkoman Corporation, Research and Development Division,
399 Noda, Noda-shi, Chiba 278-0037

1. UNSCRAMBLER とは

現在、ケモメトリックス手法¹⁾²⁾を利用するために、欧米の様々な会社や研究機関から各種ソフトウェアが公開されている。それらの中、80年代半ばにノルウェーで開発された UNSCRAMBLER は最も古い製品の一つである。開発者の Dr. Harald Martens 自身が Norwegian Food Research Institute に所属した食品研究者のせい、近赤外 (NIR) データや官能評価データの解析に世界中で広く利用されている。ちなみに Magni 夫人も、UNSCRAMBLER を官能評価データの解析に利用し、多数の優れた論文を著している。本ソフトウェアの功績は、現在最も強力な多変量回帰分析手法と言われる PLS (Partial Least Squares) 回帰分析³⁾を食品や分析化学分野に広めた点にある。しかし近年公開されるバージョンは、PLS 回帰分析にとどまらず、実験計画法、パターン認識手法までも含んでいる。Windows 対応となつてからは、他のソフトウェアと同様に、プルダウン・メニューで操作を進められるので、Word や Excel などの汎用ソフトウェアと同じ感覚で使用できる。

2. 作動環境とデータ入力

日本語、英語を問わず Windows 95 または Windows NT 上で作動するが、Windows 3.1 では使えない。ソフトウェアは CD-ROM で供給され、必要なハードディスク容量は 15 MB である。また MPU としては Pentium 100 MHz 以上が望ましく、RAM は最低 16 MB が必要で、32 MB 以上が推奨されている。多変量回帰分析やパターン認識に供する入力データのフォーマットとして

は、表計算に広く利用されている Excel, Lotus はもとより、最も単純な ASCII、もしスペクトルデータが JCAMP や GRAMS で出力されているならば、そのまま読み込める。

3. 解 析 手 法

利用可能な手法は実験計画法、多変量回帰分析手法およびパターン認識手法に大別できる。

3.1 実験計画法

多数因子中から有用因子を選択するには、要因計画法、部分要因計画法、さらに因子間の相互作用を無視して (因子数 + 1) の実験数で主効果を定量化できる Plackett-Burman 計画が利用できる。有用な 3, 4 個の因子を選択できたら、次のステップとして Box-Behnken 計画や中心複合計画を利用した、応答曲面法による最適条件の探索になる。いずれの計画法も各因子に対応する因子水準の入力は、実際の濃度や温度などでも、記号化した (+1, -1) でもよい。計画法、因子名と上下の水準、繰り返し実験数、中心点での実験数などを入力すると、ランダム化した実験点とそれに対応する実験条件が提示される。指示された条件下で実験を遂行し得られた結果、すなわち応答を入力後、主効果、相互作用、2 次作用などを組み合わせたモデルを選ぶと、分散分析や各因子の効果を定量化した重回帰分析 (Multiple Linear Regression Analysis: MLR) 結果を出力する。出力形式は、三次元応答曲面やそれを投影した等高線プロットなどの図、または数字を選択できる。

3.2 多変量回帰分析

目的変数が単一のデータに対しては MLR の他、主成

* キッコーマン(株)研究本部 (〒278-0037 野田市野田 399)

分回帰分析 (Principal Component Regression Analysis: PCR), PLS1, さらに複数の目的変数を有するデータには PLS2 が準備されている。PLS 回帰分析は MLR と異なり、目的変数と説明変数を共に利用して互いに無相関な PLS 成分を抽出して説明変数とするので、MLR 固有の諸問題を回避できる。これが利用例の多い理由である。では PCR はどうか? PCR で説明変数とする主成分も、説明変数どうしの相関関係を利用して無相関な主成分を抽出することに変わりはない。しかし、主成分分析では説明変数の影響を考慮しないため通常、PLS 回帰分析ほど高い予測精度は期待できない。PLS 回帰分析では、PLS 成分に対する元の変数の重みをプロットすることにより、各説明変数の重要度を定量化して比較できる。MLR からは、この様な知見は得られない。

3.3 パターン認識

目的変数がなく探索的な手法である主成分分析の他、ケモメトリックスの創始者である Dr. Wold の開発した SIMCA (Soft Independent Modeling of Class Analogy) が手法として用意されている。SIMCA では、まず各群ごとの主成分分析からクロスバリデーション法を利用して最適な主成分数を求める。次に、そのようにして得た主成分得点に基づく各群の試料分布を利用して判別する。群どうしの判別状態は、試料分布を二次元に投影した Cooman プロットにより直接見られるし、異常試料の存在もこの図から分かる。また判別分析のように目的変数が質的で群を示すような場合、それを (1, 0) の組合せで表せば、PLS2 を利用した PLS 判別分析が適用できる。もちろん SPSS や SAS などの汎用ソフトウェアのように、多様なクラスタ分析や因子分析などは用意されていないが、化学、食品の分析データに過不足無く対応できるだけの機能は十分完備している。しかし研究者としては、クラスタ分析は UNSCRAMBLER にも是非備えて欲しい解析手法である。

4. 解析結果の出力及びマニュアル

結果をエクスポートするフォーマットは ASCII, JCAMP, MatLab などであり、結果をそのまま Word, Excel, WordPerfect などの汎用ソフトウェアへ移したいユーザーには不満であろう。プリントアウトは画面で見た通りに印刷されるが、カラープリンターが手元に無い場合、白黒で印刷した応答曲面にはかなりの欲求不満を覚える。また、図の大きさの選択機能も豊かとは言えないし、表に用いるフォントやサイズにもかなりの不満を覚える。

一方、英文ではあるが約 500 ページという大部のマニュアルは充実している。各機能についての説明は分かり易いし、手法や用語についてより深い理解を求めるユーザーに対しては、巻末にそれらの数学的な定義も備わっている。マニュアルでは画面をそのまま出力して説明しており、初心者でも食品や化学から例を取ったモデルデータを利用して練習すれば、それほどの困難を感じることなく使用法を修得できるに違いない。

新規購入の価格は約 50 万円とかなり高価であるが、ネットワーク上での使用契約を結べば 1 台当たりの使用料はかなり低減できる。また大学の研究室はアカデミック・ディスカウントにより格段に安く購入することが可能である。より詳しくは、下記に直接コンタクトされたい。

CAMO ASA

Olav Tryggvassonsgt. 24, N-7011 Trondheim, Norway
Tel: +47 7351 4966, Fax: +47 7351 4257
E-mail: camo@camo.no, Web: <http://www.camo.no>

5. 結 論

UNSCRAMBLER 7.01 は実験計画法、応答曲面法を利用した最適化、PLS 回帰分析、SIMCA などの利用には大いに推薦できるソフトウェアである。ただアメリカ製のソフトウェアとは異なり、ヨーロッパのソフトウェアの例に漏れず、定食スタイルと言うよりもむしろアラ・カルト的に構成されている。つまり我々にはややもすると煩雑に感じられる、欧米のレストランにおいて多数のドレッシングやパンから好みの種類を選択しなければならないと同様に、使用に際して決定を迫られる選択肢は各段階で非常に多い。黙って座れば次々に料理が登場するコース・メニューのような状況を期待するユーザーは当初、かなりの戸惑いを感じるに違いない。お金さえ出せば分析機器と同様に、いくらでも多機能のソフトウェアは購入できる。しかし、それを有効に活用するには、ケモメトリックス手法に関する基礎的な知識の修得は不可欠である。これはケモメトリックス手法を正しく利用するためには、避けて通れない関門なのである。

効率的な研究の遂行に不可欠なソフトウェアの客観的な評価レポートは欧米の学会誌では珍しくないが、我が国では宣伝臭が強いと見なされるためか、商業誌はともかく学会誌では余り見かけない。しかし職場はもとより家庭にも高性能のパーソナルコンピュータが広く普及した今日、適切なソフトウェアの選択と使用法の修得は、

私たちを研究者にとどめるか粗大ゴミに貶めるかを左右する一大事であり、このようなレポートが今までなかったことがむしろ不思議でさえある。本評が食品の研究・開発に従事する読者の一助となれば幸いである。

文 献

- 1) 相島鐵郎：ケモメトリックス-新しい分析化学（丸善），（1992）。
- 2) 宮下芳勝・佐々木慎一：ケモメトリックス-化学パターン認識と多変量解析（共立出版），（1995）。
- 3) MARTENS, H. and NÆS, T.: Multivariate Calibration. (John Wiley & Sons, Chichester), (1989).

（平成 11 年 2 月 25 日受理）