新薬開発のための多目的コンジョイント解析

申請中　大阪大学　*河崎 誠　KAWASAKI Makoto
01005194　大阪大学　石井 博昭　ISHII Hiroaki
02202964　大阪大学　杉原一臣　SUGIHARA Kazutomi

1　はじめに

近年，様々な分野において多変量解析による分析がなされている．特に製薬会社における新薬開発では，開発の効率化や成分の特定化を目指し多変量解析を活用している．一方，主としてマーケティングの分野で活用されてきた手法としてコンジョイント解析があり，データによる全体評価からその全体評価をもたらした各属性の部分効用値を求める手法として開発されてきた．Luce と Tukey が 1964 年にコンジョイント解析の公理論的体系化を発表し，コンジョイントという用語を最初に公にした．産業界での最初の適用は，1971 年に Green と Rao が行った．彼らは 5 種類の雑誌に 8 種類の広告のコピーを掲載して，消費者の選好順序評価から，どの組み合わせが最も広告効果があるのかという課題について研究を進めた．

コンジョイント解析にはクラスター解析，TRADE OFF 解析や LINMAP 等様々な手法があるが，最も代表的な手法としては MONANOVA(Monotone Analysis Of Variance) 法がある．MONANOVA 法は，序数尺度で測定された選考順序評価を単調変換した値と，部分効用値の線形モデルにより計算された値との乖離を示すために，多次元尺度における単調回帰と同様の意味を持つストレスという関数を定義し [1]，このストレスを最小にする部分効用値を反復計算によって求める手法である．しかしストレスは凸性をもたないため，最適な部分効用値の導出が不可能となり，解析者が意図を持たずに真の解を探索する場合には，ただ混乱と迷いを与えるだけになる．

分数二次計画法 [3] は目的関数が分数であり，かつ分母分子が二次形式である際の最適化手法である中，分数二次計画法を用いると最適な部分効用値の導出が可能となる．[5] では創薬研究への応用を目指し，上記の手法を拡張した．まず母体となる化合物の特定の位置に様々な置換基を結合させ，新薬として期待される複数のサンプルを挙げる．次にそのサンプルから，あらかじめ行われた実験からどの程度効き目があるのかを，数値的データとして得た活性値を全体評価とし，どの置換基の組み合わせが最も効果的であるかを調べるために，個々の置換基の部分効用値を解析，及び評価した．特に従来では全体評価が単なる順位データでしか与えられなかったが，[5] では全体評価として数値データそのものを部分効用値に反映させた．

しかし実際の創薬研究では，1 つの化合物からは活性値だけでなく，脳内移行性や代謝安定性など，様々な数値データを考慮し最適な化合物を発見しなければならない．そこで本研究では [5] を拡張した，上記の多目的データを全体評価とし，各置換基の部分効用値の導出を行う．

2　分数二次計画法を用いた MONANOVA の定式化

部分効用値ベクトル $\mathbb{b}$ を求めるために，各要因の有無を示す 0-1 デザイン行列 $\mathbb{D}$, 実測値（活性値）ベクトル $\mathbb{Z}$, 加法的コンジョイントモデルのベクトル $\hat{\mathbb{Z}}(=\mathbb{D}\mathbb{b})$ として，野口等は適合度基準として $S^2(\mathbb{b})$ を定義した．

定義（分数二次計画の適合度基準)(ストレスの 2 乗)

$$S^2(\mathbb{b})=\frac{(\mathbb{Z}-\hat{\mathbb{Z}})^T(\mathbb{Z}-\hat{\mathbb{Z}})}{(\hat{\mathbb{Z}}-\bar{\hat{\mathbb{Z}}})^T(\hat{\mathbb{Z}}-\bar{\hat{\mathbb{Z}}})}=\frac{(\mathbb{Z}-\hat{\mathbb{Z}})^T(\mathbb{Z}-\hat{\mathbb{Z}})}{(\hat{\mathbb{Z}}-\mathbb{C})^T(\hat{\mathbb{Z}}-\mathbb{C})}$$

上式の分母は平均値からのずれを，分子は実測値からのずれを表し，$S^2(\mathbb{b})$ の値が小さいほど適合度が高いといえる．分数二次計画法を用いた MONANOVA の解法では，$\mathbb{S}^2(\mathbb{b})$ が最小となるような部分効用値ベクトル $\mathbb{b}$ を求めることになる．またここで $\mathbb{C}$ は

$$
\mathbf{C} = \overline{\mathbf{Db}} = \begin{bmatrix} \dfrac{\hat{z_1} + \hat{z_2} + \cdots + \hat{z_n}}{n} \\ \vdots \\ \dfrac{\hat{z_1} + \hat{z_2} + \cdots + \hat{z_n}}{n} \end{bmatrix}
$$

なる平均値ベクトルである．

このの $S^2(\mathbf{b})$ は順序差を間隔尺度の距離差と見なして，$\overline{\mathbf{Db}}(=\bar{\hat{\mathbf{Z}}}) = \mathbf{C}$ の条件下でこの距離差を最小にする $\mathbf{b}$ を求めるための適合基準であり，この $S^2(\mathbf{b})$ を分数二次計画の適合度基準と呼ぶ．

3 集合被覆問題

コンジョイント解析だけでは単目的しか扱えないので，多目的のデータに応用するために問題を集合被覆(セット・カバーリング) 問題におきかえることにする．$M = \{1, \cdots, m\}$ を有限集合とし，M の部分集合の族を $\{M_1, M_2, \cdots, M_n\}$ とする．このとき，添え字の集合 $N = \{1, \cdots, \}$ の部分集合 F が条件

$$
\bigcup_{j \in F} M_j = M
$$

を満たせば，F は M の被覆であるという．

M_j を被覆に採用する費用を c_j とし，最小費用の被覆を求める問題を集合被覆問題という．この問題は容易に 0-1 整数計画問題として以下のように定式化される．$i \in M, j \in N$ に対して

$$
a_{ij} = \begin{cases} 1 & \text{if } i \in M_j \\ 0 & \text{if } i \notin M_j \end{cases}
$$

なる係数 a_{ij} と，$j \in N$ に対して

$$
x_j = \begin{cases} 1 & \text{if } j \in F \\ 0 & \text{if } j \notin F \end{cases}
$$

を導入すれば，集合被覆問題は

$$
\begin{aligned}
&\textit{minimize} \quad z = \sum_{j=1}^{n} c_j x_j \\
&\textit{subject to} \quad \sum_{j=1}^{n} a_{ij} x_j \geq 1, (i = 1, \cdots, m) \\
&\qquad\qquad\quad x_j \in 0, 1, (j = 1, \ldots, n)
\end{aligned}
$$

と，定式化される．

本研究では上記の問題に類似した整数計画問題を，コンジョイント解析と融合させることにより多目的データの解析を行った．

4 おわりに

コンジョイント解析，及び集合被覆問題に置き換えた場合の詳細な説明については当日の発表で述べる予定である．

謝辞

本研究は，萬有製薬株式会社との共同研究によって遂行したものであります．研究で用いた様々なデータを提供して頂いた萬有製薬株式会社の方々に対し，厚く御礼申し上げます．

参考文献

[1] Kruskal, J. B. [1964] "Multidimensional Scaling by Optimizing Goodness of Fit to a Nonmetric Hypothesis," *Psychometrica*, Vol. 29, No. 1, pp. 1-27.

[2] Ibaraki, I., H. Ishii, J. Iwase, T. Hasegawa, and H. Mine [1976] "Algorithms for Quadratic fractional Programming Problems," *Journal of the Operations Reserch Society of Japan*, Vol. 19, No. 2, pp. 228-244.

[3] Ishii, H. [1978] "Studies on Discrete Programming" *Doctoral Thesis*, pp.167-201.

[4] Beasley, J. E. [1987] "An Algorithm for Set Covering Problem," *European Journal of Operational Research*, Vol. 31, pp. 85-93.

[5] 河崎誠，「新薬開発のためのコンジョイント解析」2000 年度大阪大学工学部応用物理学科卒業論文．