

マルチブロック成分法による脳画像解析

佐賀大学医学部 川口 淳

1. 概要

近年の医学ビッグデータ解析は、これまで個々に解析されていた複数のデータセットを統合的に解析することによって、疾患の特徴づけなどを多様な角度から同時に行っている。脳画像解析ではマルチモダルと呼ばれ、脳の形態と機能などの側面からの脳病態が評価される。本研究では多種モダリティ脳画像間の関連性を明らかにする方法として、膨大なデータにおいて有効な情報を保持しつつ次元縮小を行った上で統計学的モデリングを行う中での次元縮小法を提案する。この方法ではモダリティの情報のみならず、アウトカムの情報を組み込んで解釈のしやすい結果を導くように工夫している。またスパースモデリングを用いアウトカムに関連する変数の特定を行う。さらに元の次元に戻して結果を表示する事が可能である。

2. 方法

n 人の被験者を考える。 X_m を $n \times p_m$ ($m = 1, 2, \dots, M$) 行列。 M はブロック (モダリティ) 数である。 Z を n 次元アウトカムベクトルとする。 X を要約するマルチブロックスコア (成分) は

$$\mathbf{t} = \sum_{m=1}^M b_m X_m \mathbf{w}_m$$

$\mathbf{w} = (\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_M)^\top$ をサブ重みベクトルと呼び、 $\mathbf{b} = (b_1, b_2, \dots, b_M)^\top$ をスーパー重みベクトルと呼び、その求め方として $\max L_0(\mathbf{w})$ という最適化問題を提案する。ただし、

$$L(\mathbf{b}, \mathbf{w}) = (1 - \mu) \mathbf{t}^\top \mathbf{t} + \mu \mathbf{t}^\top \mathbf{Z} - \sum_{m=1}^M P_{\lambda_m}(\mathbf{w}_m)$$

また $\|\mathbf{w}_m\|^2 = 1$, $\|\mathbf{b}\|^2 = 1$ という制約を設け、 $0 \leq \mu \leq 1$, $P_\lambda(x) = 2\lambda|x|$ である。このとき、 \mathbf{w}_m に関する coordinate update は $\tilde{\mathbf{w}}_m = h_{\lambda_m}(b_m X_m^\top \{(1 - \mu)\mathbf{u} + \mu\mathbf{Z}\})$ として与えられる。ただし $h_\lambda(y) = \text{sign}(y)(|y| - \lambda)_+$ である。方法はアルツハイマー病研究における解析に応用される。

参考文献

Kawaguchi A, Yamashita F (2017). Supervised Multiblock Sparse Multivariable Analysis with Application to Multimodal Brain Imaging Genetics. *Biostatistics*, 18(4) 651-665.

Kawaguchi A (2017). Supervised Dimension Reduction Methods for Brain Tumor Image Data Analysis. In *Frontiers of Biostatistical Methods and Applications in Clinical Oncology*, Matsui S and Crowley J. (Ed.) 401-412