

階層型成分法によるマルチモダル脳画像解析

佐賀大学医学部 川口 淳

1. 概要

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として SMSMA (Supervised Multi-Block Sparse Multivariable Analysis) 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. SMSMA 法

n 人の被験者を考える。 X_m を $n \times p_m$ ($m = 1, 2, \dots, M_X$) 行列, Y_m を $n \times q_m$ ($m = 1, 2, \dots, M_Y$) 行列とする。 M_X, M_Y 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 と Y の関連性を表す階層型スコア (成分) を X と Y それぞれに対して求める。

$$\mathbf{t} = \sum_{m=1}^{M_X} b_{X,m} X_m \mathbf{w}_{X,m}, \quad \mathbf{u} = \sum_{m=1}^{M_Y} b_{Y,m} Y_m \mathbf{w}_{Y,m}$$

$\mathbf{w}_X = (\mathbf{w}_{X,1}, \mathbf{w}_{X,2}, \dots, \mathbf{w}_{X,M_X})^\top$, $\mathbf{w}_Y = (\mathbf{w}_{Y,1}, \mathbf{w}_{Y,2}, \dots, \mathbf{w}_{Y,M_Y})^\top$ をサブ重みベクトルと呼び, $\mathbf{b}_X = (b_{X,1}, b_{X,2}, \dots, b_{X,M_X})^\top$ と $\mathbf{b}_Y = (b_{Y,1}, b_{Y,2}, \dots, b_{Y,M_Y})^\top$ をスーパー重みベクトルと呼び, スコア同士及びスコアとアウトカムの共分散の罰則付き最大化によって求める方法を提案する。すなわち, $\max L_0(\mathbf{w}_X, \mathbf{w}_Y)$ という最適化問題を考える事になる。ただし,

$$L(\mathbf{b}_X, \mathbf{w}_X, \mathbf{b}_Y, \mathbf{w}_Y) = \mu_{XY} \mathbf{t}^\top \mathbf{u} + \mu_{XZ} \mathbf{t}^\top \mathbf{Z} + \mu_{YZ} \mathbf{u}^\top \mathbf{Z} - \sum_{m=1}^{M_X} P_{\lambda_{X,m}}(\mathbf{w}_{X,m}) - \sum_{m=1}^{M_Y} P_{\lambda_{Y,m}}(\mathbf{w}_{Y,m})$$

また $\|\mathbf{w}_X\|^2 = 1$, $\|\mathbf{w}_Y\|^2 = 1$, $\|\mathbf{b}_X\|^2 = 1$, $\|\mathbf{b}_Y\|^2 = 1$ という制約を設け, $0 \leq \mu_{XZ}, \mu_{YZ}, \mu_{XY} \leq 1$, $\mu_{XY} = 1 - \mu_{XZ} - \mu_{YZ}$, $P_\lambda(x) = 2\lambda|x|$ である。このとき, $\mathbf{w}_{X,m}$ に関する coordinate update は $\tilde{\mathbf{w}}_{X,m} = h_{\lambda_{X,m}}(b_{X,m} \mathbf{X}_m^\top \{\mu_{XY} \mathbf{u} + \mu_{XZ} \mathbf{Z}\})$ として与えられる。ここで $h_\lambda(y) = \text{sign}(y)(|y| - \lambda)_+$ である。他の重みに関しても類似のものが与えられ, 重みを算出するアルゴリズムが構築される。

3. 応用

アルツハイマー病研究における 2 モダリティ脳画像 (MRI, PET) そして遺伝子情報 (SNP) を用いた解析に応用した。

参考文献

Kawaguchi A, Yamashita F (2017). Supervised Multiblock Sparse Multivariable Analysis with Application to Multimodal Brain Imaging Genetics. *Biostatistics*, in press.