25th International Congress on Modelling and Simulation, Darwin, NT, Australia, 9 to 14 July 2023
mssanz.org.au/modsim2023

Predicting from functional brain images using partial least squares for Riemannian manifolds

M. Ryan [a,b], **G. Glonek** [a], **J. Tuke** [a] **and M. Humphries** [a]

[a] *School of Computer and Mathematical Sciences, University of Adelaide, Australia*
[b] *CSIRO, Adelaide, Australia*
Email: matt.ryan@csiro.au

Abstract: The human brain is an anatomically and functionally complex organ highly studied across neuroscience and psychology. One approach to investigating the functional brain is *functional magnetic resonance imaging* (fMRI) – a non-invasive medical imaging technique that creates a high-dimensional spatio-temporal picture from blood flow in the brain. To reduce the complexity of fMRI data, measures such as the *functional connectivity matrix* – the Pearson correlation matrix between R user-specified regions of interest in the brain – may be considered.

Many machine learning models have been applied to fMRI studies to investigate relationships between functional connectivity and phenotype data, with *partial least squares* (PLS) becoming popular among neuroscientists since 1996 (McIntosh et. al., 1996). Developed for chemometric data by Wold (1975), PLS is an extension of multiple linear regression that models high-dimensional data by simultaneously decomposing response and feature variables into a set of related latent variables. In fMRI studies, the upper triangle of the functional connectivity matrix is vectorised to obtain $\mathrm{R}(\mathrm{R}-1)/2$ features per subject which form the inputs into PLS with the intent of predicting clinical outcomes. However, vectorising the functional connectivity matrices fails to account for a key structure of these matrices; functional connectivity matrices are positive definite.

To address the positive definite criteria in our model, we view functional connectivity matrices through the lens of *Riemannian geometry*. A Riemannian manifold is a globally non-linear generalisation of Euclidean space upon which we can perform calculus, discuss tangent vectors, and measure distances between objects. In particular, the set of symmetric positive definite matrices of dimension $\mathrm{R} \times \mathrm{R}$ (of which functional connectivity matrices belong) form a Riemannian manifold with the affine-invariant metric (Pennec et.al., 2006).

We propose a model that generalises PLS to Riemannian manifold data. Specifically, we develop a PLS model such that both the response and feature data can take values in any complete Riemannian manifold. To fit the Riemannian PLS model, we provide an efficient tangent space algorithm that approximates Riemannian PLS by linearising the manifold data to the tangent space at the Fréchet mean, reducing the complexity of the model fitting process.

We present results of Riemannian PLS applied to the Centre for Biomedical Research Excellence (COBRE) dataset (Aine et. al., 2017), which contains functional connectivity and phenotype information on schizophrenic subjects (n = 72) and healthy controls (n = 74). We demonstrate the applicability of Riemannian PLS by using it to predict subject group and age from functional connectivity matrices. With the increasingly complex data available today – especially through multi-modal medical imaging studies – we believe our method may help uncover new, exciting, and intricate relationships in the ever-expanding world of complex data.

REFERENCES

Aine, C. J., Bockholt, H. J., Bustillo, J. R., Cañive, J. M., Caprihan, A., Gasparovic, C., … Calhoun, V. D. , 2017. Multimodal Neuroimaging in Schizophrenia: Description and Dissemination. Neuroinformatics, 15 (4), 343-364.

McIntosh, A. R., Bookstein, F. L., Haxby, J. V., & Grady, C. L., 1996. Spatial Pattern Analysis of Functional Brain Images Using Partial Least Squares. NeuroImage, 3 (3 Pt 1), 143-157.

Pennec, X., Fillard, P., & Ayache, N., 2006. A Riemannian Framework for Tensor Computing. International Journal of Computer Vision, 66, 41-66.

Wold, H., 1975. Soft Modelling by Latent Variables: The Non-Linear Iterative Partial Least Squares (NIPALS) Approach. Journal of Applied Probability, 12 (S1), 117-142.

Keywords: *Machine learning, complex data, fMRI, Riemannian geometry, medical imaging*