Global PCA of Local Moments

With Applications to MRI Segmentation

Jacob M. Maronge[1, 2], John Muschelli[3], Ciprian M. Crainiceanu[3]

[1]University of Wisconsin - Madison, Department of Statistics

[2]University of Wisconsin - Madison, Department of Biostatistics and Medical Informatics

[3]Johns Hopkins Bloomberg School of Public Health, Department of Biostatistics

Contact Information:
Jacob M. Maronge
Department of Statistics
University of Wisconsin - Madison
1300 University Ave, Madison, WI, 53706

Email: jmmaronge@gmail.com
Website: https://jmmaronge.github.io

Introduction

Every image can be vectorized. However, its meaning, interpretation, and complexity is encapsulated in the collection of all neighborhoods of all locations. We present a framework for studying the information stored within these neighborhoods. Such matrices are very large and store information inefficiently, but they provide a useful theoretical framework for representation of imaging information. Here we propose to exploit this theoretical framework to introduce simple methods to quantify the variation in multimodal images based on the shared information across local spatial neighborhoods and subjects.

Challenges

- Imaging data is very large.
- When we consider local neighborhoods the size of these data are immediately increased by a factor of the size of the neighborhoods we consider.
- It is difficult to concisely describe this complicated data structure in a way that is useful for studies of association with health outcomes.

Here we consider the case when multimodal imaging is available for multiple subjects.

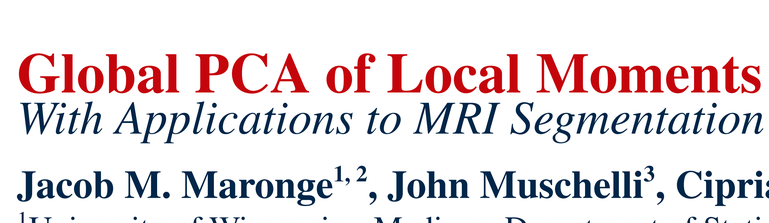

Figure 1: From left to right: DWI, FLAIR, T1w, and T2w images for subject 05, 2015 ISLES Challenge

We would like to use our approach to study these multimodal images and characterize how well this framework predicts ischemic stroke lesions in patients.

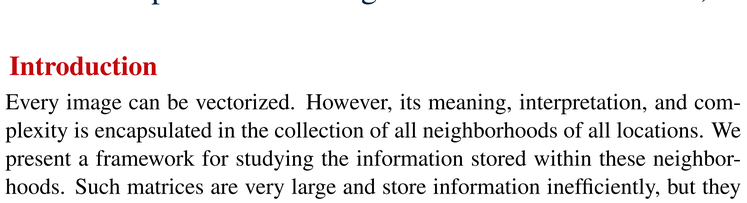

Figure 2: From left to right: FLAIR image, gold standard manual segmentation, and prediction example

Objectives

1. Decompose observed variability images at the population level.
2. Describe and quantify the main directions of variation.
3. Use these directions of variation to improve segmentation and studies of association with health outcomes.

Materials and Methods

To achieve these objectives we propose to decompose the local variability of various moments of the image intensities: the image intensity, image intensity square, and so on. Consider the following mock example in 2-D:

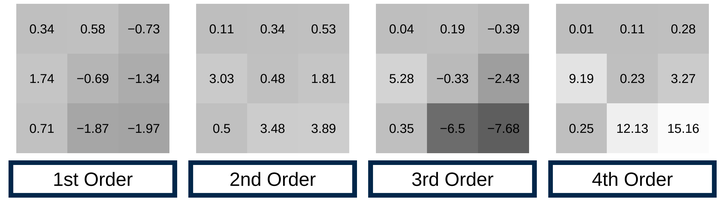

Figure 3: 2-D mock example of moments of local neighbors

$$X_{ij} = (\underbrace{0.34, 0.58, \ldots, -1.97}_{\text{1st Order}}, \underbrace{0.11, \ldots, 3.89}_{\text{2nd Order}}, \underbrace{0.04, \ldots, -7.68}_{\text{3rd Order}}, \underbrace{0.01, \ldots, 15.16}_{\text{4th Order}}).$$

This gives the jth row for subject i. If we are interested in considering additional imaging modalities, we simply add those as additional columns in the matrix. We perform this operation for each voxel for each subject and stack these vectors into a (potentially large) matrix $\mathbf{X}$. Next,

- Center and scale the columns of $\mathbf{X}$.
- Perform PCA on centered and scaled matrix.
- Calculate principal component scores.
- Use first Q principal component scores as predictors and manual segmentation as outcome to train a model to perform segmentation of stroke lesions.

Computational Issues

We have 2 challenges computationally: 1.) we need to perform PCA on a very large matrix $\mathbf{X}$ with dimensions $N \times P$, where $N >> P$, 2.) we would like to center and scale the columns of $\mathbf{X}$ before decomposing. Let $\bar{x}$ denote the vector of column means, s denote the vector of column standard deviations, and $\mathbf{Y}$ denote the centered and scaled version of $\mathbf{X}$. For the PCA, all we need is the left singular values, therefore we can decompose,

$$\mathbf{C} = \mathbf{Y}^T\mathbf{Y} = \mathbf{V}\mathbf{D}^T\mathbf{U}^T\mathbf{U}\mathbf{D}\mathbf{V}^T = \mathbf{V}\mathbf{D}^2\mathbf{V}^T,$$

which is $P \times P$ in dimension instead of $N \times P$. We still need to calculate $\mathbf{C}$, and we would like to do so efficiently. Notice,

$$\mathbf{C} = \frac{\sum_i \mathbf{X}_i^T\mathbf{X}_i - N\bar{x}\bar{x}^T}{ss^T}. \quad (1)$$

Where $\mathbf{X}_i$ is the block of $\mathbf{X}$ for the ith subject. Each piece of equation 1 can be calculated by reading in one subject at a time. This is implemented in the R package MEDALS (Memory Efficient Decomposition for Analysis of Local neighborhood moments for Segmentation) and is available at https://github.com/JMMaronge/medals.

Results

All results shown are from the 2015 ISLES SISS Challenge. Below we show partial ROC (pROC) curves for prediction of stroke lesions for each subject, as well as overall training and overall testing results. We also show an example of how our method performed on a particular subject.

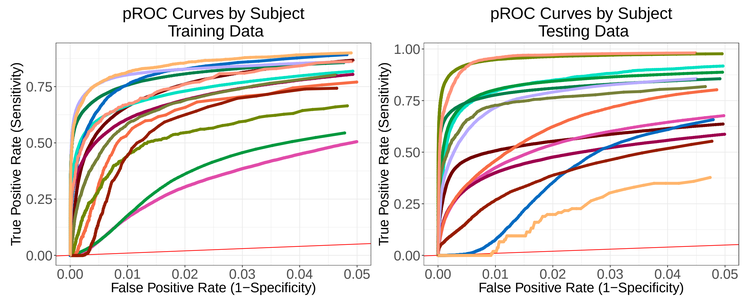

Figure 4: Subject level pROC curves for training set and testing set

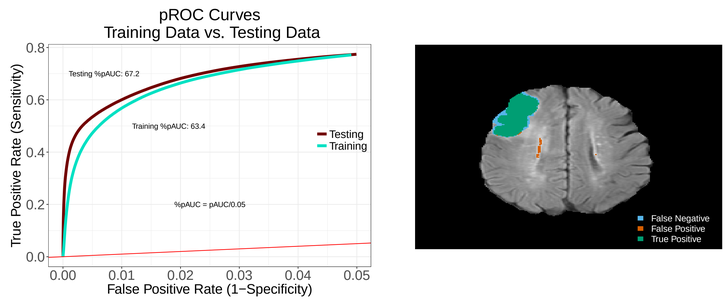

Figure 5: (left) Overall training results and overall testing results (right) Example of performance on one subject

Conclusions

- Carefully using information contained within neighborhoods allows for a richer insight of imaging data structures.
- Using local neighborhoods creates difficulties with storing the data, but these can be overcome by leveraging structure of the data.
- Characterizing variation in neighborhoods allows for interesting applications in studies of association with health outcomes.

References

[1] John Muschelli, Elizabeth M Sweeney, Natalie L Ullman, Paul Vespa, Daniel F Hanley, and Ciprian M Crainiceanu. Pitchperfect: Primary intracranial hemorrhage probability estimation using random forests on ct. *NeuroImage : Clinical*, 14:379–390, 2017.

[2] Vadim Zipunnikov, Brian Caffo, David M. Yousem, Christos Davatzikos, Brian S. Schwartz, and Ciprian Crainiceanu. Multilevel functional principal component analysis for high-dimensional data. *J Comput Graph Stat.*, 20(4):852–873, 2011.