Copyright

by

Daniel Alejandro Noble Hernandez

2021

The Master's Thesis Committee for Daniel Alejandro Noble Hernandez
Certifies that this is the approved version of the following Master's Thesis:

Extending PrAGMATiC: Modeling covariances between responses across the human cortex

APPROVED BY
SUPERVISING COMMITTEE:

Alexander G. Huth, Supervisor

Etienne Vouga

Extending PrAGMATiC: Modeling covariances between responses across the human cortex

by

Daniel Alejandro Noble Hernandez

Thesis

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Master of Science in Computer Science

The University of Texas at Austin

May 2021

Dedication

To my Mum and Dad

Acknowledgements

I’d like to express my gratitude to my thesis advisor, Dr. Alexander G. Huth, for taking me on as a student and guiding me through this research and the ensuing thesis. His passion for his field and good nature was infectious throughout this process, and his patience in teaching me and insight at the frequent roadblocks I’ve hit have been crucial in my completion of this thesis.

I’m grateful to Dr. Etienne Vouga for his ready willingness to serve as a reader on my Master’s Thesis Committee.

Thanks to my friends for their steady support, listening to me panic about my occasional incompetence as a programmer when I’ve felt most out of my depth.

I’m grateful to Brick, for keeping me company every night, squashed into an armchair with me.

Abstract

Extending PrAGMATiC: Modeling covariances between responses across the human cortex

Daniel Alejandro Noble Hernandez, MSCompSci

The University of Texas at Austin, 2021

Supervisor: Alexander G. Huth

The human brain can be understood and organised as a topographic map of cortical areas. Dividing the brain into these distinct subsections has long been an ongoing effort, with a number of often disagreeing attempts to create this map – representative across individuals – being made over time. Both surgical and computational techniques have been broadly utilized in this pursuit, focusing on characterizing these chosen cortical areas based on their structural and functional similarities across individuals. Because the general anatomy is fairly well-understood now, computational methods are favoured; a popular approach taken involves measuring the responses of areas of the cortex according to functional magnetic resonance imaging (fMRI) data. We take a related – but modified – approach in this paper, delineating a model that uses the covariances oof the responses across the cortex rather than the cortical responses themselves. An extension of the existing hierarchical, Bayesian, probabilistic and generative model PrAGMATiC, our mathematical formulation of the model assumes and samples from an underlying Wishart, rather than Gaussian, distribution. This will allow the model to learn parameters to describe the functional covariances of responses at

vertices across the cortex. Since direct comparisons of functional values, rather than their covariances, are not readily achieved in resting state fMRI, this formulation will be able to identify cortical parcellations using resting state fMRI data by providing a framework under which comparisons are possible.

Table of Contents

Chapter 1: Introduction

The human cortex—a millimeters-thick layer on the surface of the brain dictating who we are as individuals—is one of the most complex pieces of matter that we know of. The cortex not only orients us to the world around us and directs our interactions with it, but is responsible for our cognition, languages, emotions, and personalities, amongst other higher mental functions. How it functions and how it is organized—both at a low and high level—are fascinating, and, as yet, largely unsolved questions. As is currently understood in modern neuroscience, the cortex of the human brain is functionally organized in a hierarchical manner [1]. To provide some context regarding current ideas of cortical organization, we briefly present a top-down understanding of the brain's organization.

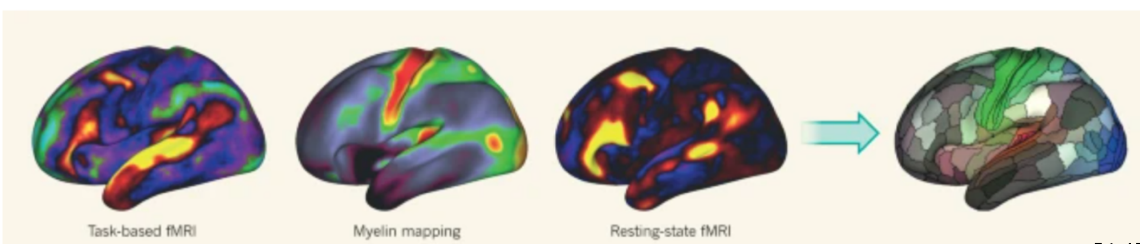
1.1 Early mapping efforts

At the topmost level, the cortex is divided into four lobes—frontal, parietal, temporal, and occipital—each responsible for a broad array of functions; taking the parietal lobe for instance, it is widely responsible for sensory processing and higher visual processing [13]. A more specific schema of organization put forth by Korbinian Brodmann lays out 52 aptly named Brodmann areas—many of which have been further subdivided over time—based on the neuron-level organization he observed when staining neurons across the cortex [15]. As it happened, these purely anatomically based parcellations proved to closely correlate with the actual cognitive functions of the areas they represented. One such area, Brodmann area 17, maps the primary visual cortex, receiving raw visual information directly from the retina, after being relayed through the lateral geniculate nucleus; area 17, commonly termed V1, is responsible for the first low-level processing of visual data before feeding it to higher areas of the cortex [10].

The anatomical maps revealed by Brodmann were later refined by functional neurophysiologists. Early efforts by Wilder Penfield and his colleagues led to the uncovering of the motor cortex—encompassing Brodmann areas 1, 2 and 3—on the precentral gyrus of the frontal lobe and the somatosensory cortex—Brodmann area 4—on the postcentral gyrus of the parietal lobe. The somatosensory cortex is the part of the brain that receives and processes the sensation of touch, whilst the motor cortex directs voluntary movement. Penfield, a neurosurgeon, mapped these areas on the brains of epileptic patients during surgery; he stimulated parts of the cortex using electrodes and took note of the motions and perceived sensations that each area's stimulation evoked [9]. In a more general sense, the other areas and sub-areas of the cortex serve both specific and hugely varied cognitive functions.

1.2 Computational methods

Since Penfield's time, significant advances in medical imaging—particularly with magnetic resonance imaging (MRI)—have been made, allowing the focus to shift to non-invasive studies to capture enormous amounts of brain data [7]. The organization of the brain discussed above is based largely on the structural anatomy of the underlying neurons found in each region, and these modern methods, relying in particular on fMRI data, have allowed for much more fine mappings of cortical areas. Although significant progress has been made in this regard, much of the cortex and its functions remains to be mapped definitively; furthermore, the connectivities of brain areas differ between individuals [5], presenting further considerations in coming up with a general map. Until fairly recently, such parcellations have been generated using a single property—such as functional connectivity, neuronal architecture, or function. This has biased the resulting cortical maps in a sense, as though significant overlap of areas should be expected across each of these generated maps, they are likely to vary along each of these properties. In their 2016 paper, Glasser et al. address precisely this problem, setting out a multi-modal approach of parcellating the cortex. Rather than use a single property as the basis for their model, utilizing data offered by the Human Connectome Project, they took measures of cortical myelin content and cortical thickness to capture structure, task fMRI from seven tasks to capture cortical function, and used resting state fMRI to capture functional connectivity in their model [4]. The general approach used by Glasser et al. is illustrated below in Figure 1. The authors point out that in this manner, not only were they able to confirm information from a single modality with that from the others, but they were also able to capture more fine-grained details in their maps that would have been missed by a single property.



[14]

Figure 1: Task-based fMRI captured cortical function, myelin mapping captured structure, and resting state fMRI captured functional connectivity in the multi-modal approach employed by Glasser et al.

Though the above multi-modal approach is robust, it is not without its drawbacks—there were costs associated with switching between what are sometimes conflicting modalities, as discussed by Schaefer et al. in their 2018 paper. There remains significant interest in generating parcellations of the cortex based on one of the properties—functional connectivity. Resting state fMRI has been used broadly to cluster areas with similar functional connectivity patterns based on these rfMRI signals [11].

Rather than directly compare functional responses across the cortex, we seek to provide a framework that compares the covariances of the functional responses. The PrAGMATiC model upon which our work is based assumes that functional values are directly comparable across subjects, but for resting state fMRI, no such first-order relationships exist [5]. However, covariances of these values appear to be conserved across individuals, providing an opportunity to make these direct comparisons. Our model is an extension of PrAGMATiC, under which we assume that the covariance of the functional values across brain areas is drawn from a Wishart distribution. The rest of the paper provides necessary background on the model that inspired this work, discusses the mathematical basis for the model's learning, presents preliminary results, and discusses the implications of the work and future directions that the research will take.

Chapter 2: Background

PrAGMATiC—a probabilistic and generative model of areas tiling the cortex—is a hierarchical Bayesian model that learns parameters to generate functional maps that are common across the cortices of a set of subjects [5] and it is the basis for the work described below. We begin with a brief overview of the difficulty posed by high dimensional models and how PrAGMATiC addresses the issue to predict cortical maps, followed by a discussion of PrAGMATiC's learning formulation, which is the subject of extension.

2.1 Cortical map prediction

Though more well-established maps of parts of the cortex, such as the visual and motor cortices, have low dimensional structure—the retina shows a one-to-one mapping onto the visual cortex within 3 dimensions for instance—most of the cortex does not have a known functional map, and defining these functional dimensions is a complex task [5]. An established method of obtaining these functional dimensions is to apply independent component analysis (ICA) to resting state fMRI data [3]. A number of studies using ICA on brain data have revealed components similar to the more well-established visual cortical networks and have come up with similar functional areas across subjects, indicating the effectiveness of this approach [2]. Nonetheless, as Huth et al. discuss in their 2017 paper, quantifying these functional maps is complicated by the variability across subjects given that, though similar, all subjects' cortical anatomies are unique, and is further complicated by the intrinsic difficulty of visualizing and understanding higher dimensional data.

To address the problems presented by higher dimensional cortical maps, Huth et al. designed PrAGMATiC, a Bayesian model that generates functional maps across the cortex shared by a set of subjects. PrAGMATiC treats each subject's functional data as drawn from a single probability distribution underlying the full subject population, and this probability distribution is split into two parts—an arrangement model distribution and an emission model distribution [5]. The arrangement model features a set of hidden variables H representing the physical arrangement of control points across the cortex upon which the functional maps are ordered; $P(H; L, K)$ is the distribution over arrangements constrained by the lengths, L, and spring constants, K, of springs that connect each of the control points [5]. The emission model is described by a set V of functional data, with $P(V|H, M)$ as the distribution over the functional data given the arrangement, allowing functional data to be drawn from a Gaussian distribution with a unique mean, M, for each area [5]. Combining the distributions gives the joint distribution for a single subject, s:

$$P(V_s, H_s|\ M, L, K) = P(H_s; L, K)P(V_S, |H_s M) \quad (1)$$

which serves as the basis for the model's learning [5]. The two main components of the model, arrangement and emission, are illustrated in Figure 2.

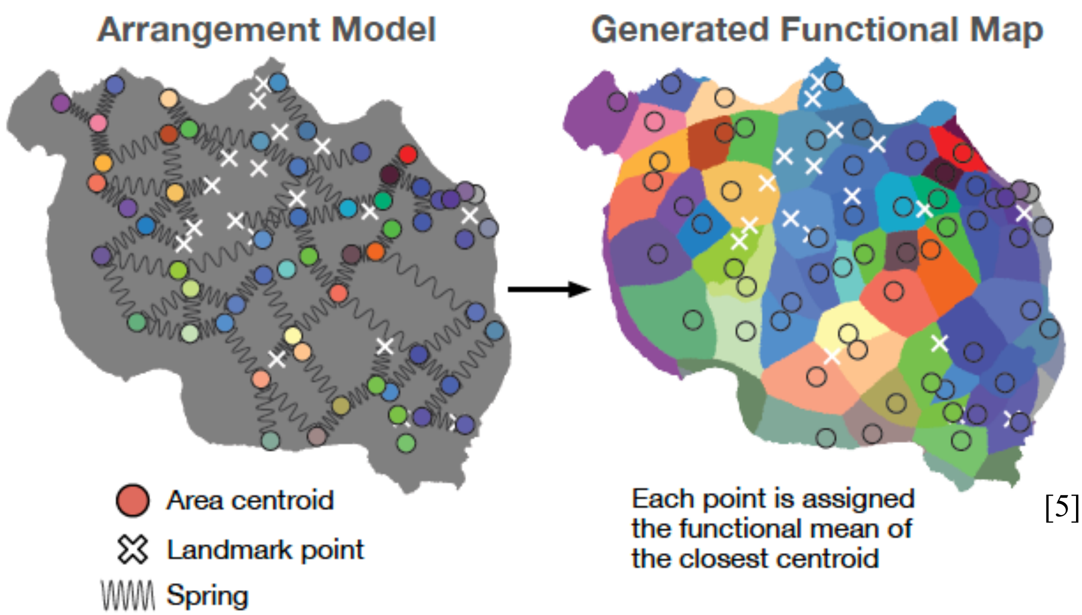


Figure 2: The arrangement model shown on the left illustrates the arrangement of the centroids across the cortical surface, joined by springs that allow variation in exact location across subjects, but the same organisation relative to one another. The emission model predicts the functional data based on the functional means of areas, and generates a map as seen on the right.

2.2 PrAGMATiC learning

Fitting PrAGMATiC is centered on learning each of the parameters – the area functional means, M, as well as the spring lengths and constants, L and K. In order to do that, Huth et al. carried out maximum likelihood estimation (MLE); they maintained j Markov chains for each of the s subjects, obtaining a gradient for each chain by drawing a sample for each learning step t:

$$H_s^{j,t} \sim P(H_s | H_s^{j,t-1}; L^{t-1}, K^{t-1}) \quad (2)$$

Following each Gibbs sweep, Huth et al. describe parameter learning by computing the gradients for each parameter, M, L, and K, and taking a step down each gradient, updating the current estimates:

$$M^t = M^{t-1} - \epsilon_M \frac{\partial \mathcal{L}(M^{t-1}; V^{obs})}{\partial M^{t-1}} \tag{3}$$

$$L^t = L^{t-1} - \epsilon_L \frac{\partial \mathcal{L}(L^{t-1}; V^{obs})}{\partial L^{t-1}} \tag{4}$$

$$K^t = K^{t-1} - \epsilon_K \frac{\partial \mathcal{L}(K^{t-1}; V^{obs})}{\partial K^{t-1}} \tag{5}$$

The area functional means, M, of the original PrAGMATiC represent the functional values that are assumed to be directly comparable, when no such first-order relationships exist between them across subjects in resting state fMRI experiments. This is the part of the model that we extend; we seek to replace the functional means with the covariance of the functional values.

Chapter 3: Model Extension

Our extension to PrAGMATiC requires the replacement of the area functional means with a new parameter, the covariances of the functional values, thus allowing direct comparisons across subjects. Below, we discuss the general reasoning and formulation of this concept within the model, and provide the mathematical basis delineating the form that these changes take and how they were derived.

3.1 Learning covariance

In our extension, we replace the area functional means M with Ω, the functional covariance between areas. In this way, we can define each element of the matrix $\Omega_{i,j}$ as the average covariance between centroids i and j. In order to model the covariance between responses across cortical locations, we took a resting state time course of length T and assumed that the covariance of T functional values across the visible units was drawn from a Wishart distribution; when a covariance matrix is not known, it can often be reliably sampled from the Wishart [12]. As such, with N_V visible units – termed vertices through the rest of our description – and Σ as a symmetrical N_V x N_V matrix, we illustrate the sampling as $V^T V \sim W_{N_V}(\Sigma, T)$.

Given that Ω represents the functional covariance between each pair of centroids, and bearing in mind that each centroid contained a number of vertices, we note that $\hat{\Sigma}$, our estimate of the covariance between pairs of vertices, is a blown-up version of Ω, and we relate the two to one another by defining a binary matrix $B(H)$ such that:

$$\hat{\Sigma}(\Omega, \mathrm{H}) = B(H)\Omega B^T(H) \tag{6}$$

Thus, with these definitions in mind, we considered the formulation of the total probability of the observed data given the parameters $P(V_s^{obs}; L, K, M)$, which was defined by the expectation of the full probability distribution over H. We replaced this equation in the original model with the following:

$$P(V^T V | L, \Omega) = \sum_h P(V^T V | h; \Omega) P(h; L) \tag{7}$$

We sought to use this formulation to learn the covariance parameters, spring lengths, and spring constants to maximize the log likelihood of the covariances across the subjects, changing equation (3) above to:

$$\Omega^t = \Omega^{t-1} - \epsilon_\Omega \frac{\partial \mathcal{L}(\Omega^{t-1}; V^{obs})}{\partial \Omega^{t-1}} \tag{8}$$

3.2 Mathematical basis

Although the gradients for the spring likelihoods and spring constants were not modified, the computation for the gradient with respect to Ω had to be determined. We defined the gradient using the probability density function of the Wishart distribution:

$$\frac{\partial \mathcal{L}(\Omega, V_{obs})}{\partial \Omega} = \frac{\partial}{\partial \Omega} \ln \left(P_W\left(V^T V; \hat{\Sigma}, T\right)\right) \tag{9}$$

$$= \frac{\partial}{\partial \Omega} \ln \left(\frac{|V^T V|^{(T-p-1)/2} e^{-tr(\hat{\Sigma}^{-1} V^T V)/2}}{2^{Tp/2} \left|\hat{\Sigma}\right|^{T/2} \Gamma_p(\mathrm{T}/2)} \right) \tag{10}$$

We found that though three of the resulting five terms went away upon derivation, and though the fourth simplified nicely too, we had significant difficulty finding the derivative of the second term, and were left with:

$$\frac{\partial \mathcal{L}(\Omega, V_{obs})}{\partial \Omega} = -\frac{1}{2} \frac{\partial}{\partial \Omega} tr\left(\hat{\Sigma}^{-1} V^T V\right) + \frac{T(\Omega^{-1})^T}{2} \tag{11}$$

Without a clear-cut path to determining the derivative of the second term $-\frac{1}{2} \frac{\partial}{\partial \Omega} tr\left(\hat{\Sigma}^{-1} V^T V\right)$ theoretically, we turned to an analytical approach. Part of the difficulty of computing this term of (11) came from the substitution of $\hat{\Sigma}^{-1}$with $(\mathrm{B}\Omega B^T)^{-1}$. Since B was a non-square matrix, it was not possible to take its inverse as required by this derivation. Thus, we considered the pseudo-inverse of B, B^+ – allowing us to formulate such a scenario – and used this to define the generalized inverse of $\hat{\Sigma}$ as:

$$\hat{\Sigma}^{-1} = {B^+}^T \Omega^{-1} B^+ \tag{12}$$

We used this formulation of the generalized inverse of $\hat{\Sigma}$ in our derivations, and in this manner, we were no long constrained by the non-invertibility of B. Even with this correction, we still faced a difficulty with calculating the partial derivative of the second term, so we employed the following analytical approach. Given the dimensions of each of the matrices involved, we suspected that the derivative of the second term would work out as:

$$B^T\hat{\Sigma}^{-1}V^TV\hat{\Sigma}^{-1}B \tag{13}$$

We defined U as:

$$U = V(B\Omega B^T)^{-1}B \tag{14}$$

Using the pseudoinverse of B as discussed above allowed us to simply U as:

$$U = V(B^T)^{-1}\Omega^{-1} \tag{15}$$

Using equation (14) to rewrite (13) then told us that under our assumption, the second term could be computed with equation (15) simply as U^TU. With this presumed reduction mind, we sought to test our suspicion by running a simulation for it.

Chapter 4: Experiments

Though our suspected reduction for our formula simplified nicely and would be relatively computationally efficient, we still needed to prove that it was correct, given that it was largely supposition. In order to show that it was sound, we implemented a simulation for it in a python testbed, as discussed below.

4.1 Programming simulation

The idea behind setting up the programming testbed was to construct an environment in which we defined some true Ω, from which we could compute a true Σ. Then, beginning with a random Ω, we used the suspected equation for the gradient, $\frac{\partial \mathcal{L}(\Omega, V_{obs})}{\partial \Omega}$, to update the random Ω and approximate the true Ω. We used a Jupyter notebook to set up a simulation in which we defined a small number of centroids as 10, the number of vertices per centroid as 50, and the number of timepoints T in V as 200. We used these values to construct B as a 500 x 10 binary matrix of repeating identity matrices, visualized for clarity, along with its pseudoinverse in Figure 1 below. The figure provides an intuitive look at how a pseudoinverse functions for a non-square matrix, where though stretched along the x-axis in the pseudoinverse by a factor of 5, the values themselves are a smaller by a factor of 5.

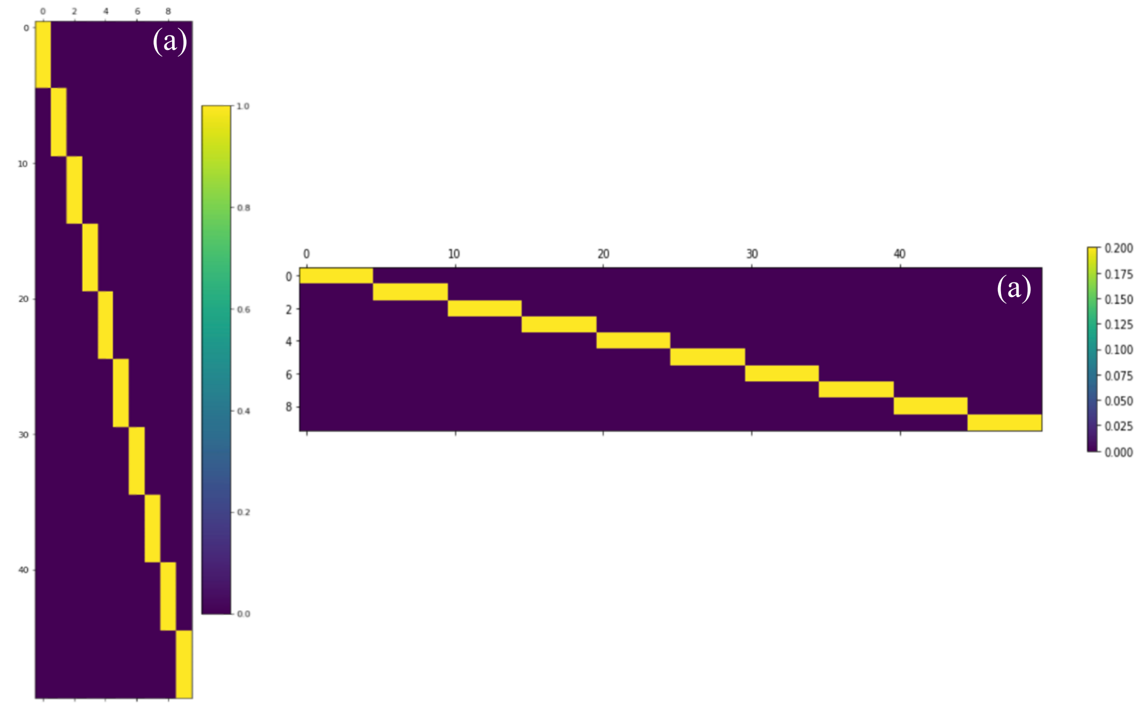


Figure 1: The left plot (a) shows matrix B as defined above, whilst the right plot (b) illustrates its pseudoinverse, B^{+}.

We then randomly generated a true Ω and used this, along with the above B and B^{T} to define a true Σ. The two are displayed below in Figure 2; as expected, they appear exactly the same, only differing in size given our construction of Σ as a blown-up version of Ω.

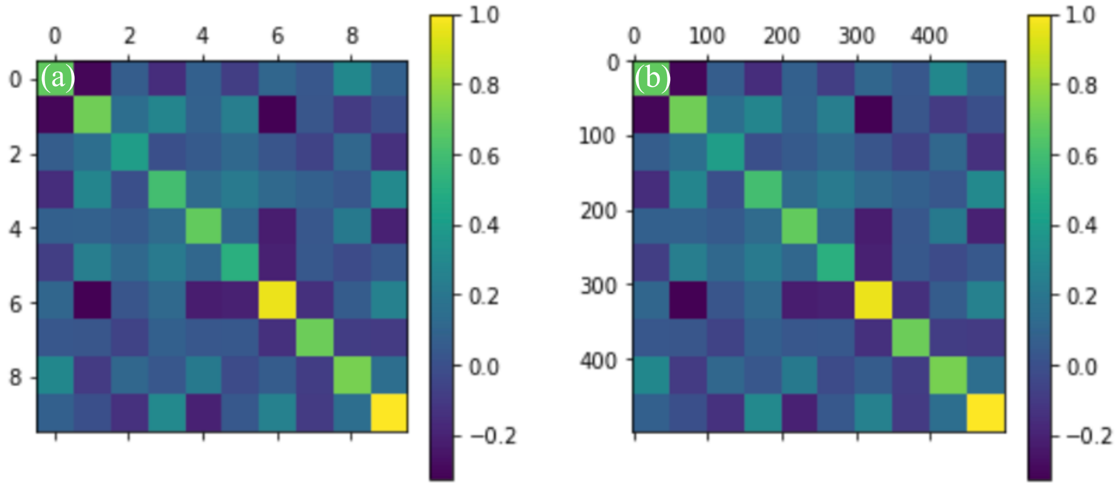


Figure 2: The left plot (a) shows the defined true Ω, whilst the right plot (b) shows the corresponding true Σ.

Following this, since we could not sample $V^T V$ directly from a Wishart distribution in the testbed, we sampled V from $P(V|\Sigma)$ by drawing a V from a multivariate normal distribution with true Σ as the covariance of the distribution. Using this sampled value of V, we were able to compute $V^T V$, and we defined a random omega which we'd update over iterations.

Having set up the experiment and all of the necessary matrices, we ran our learning algorithm based upon the full equation:

$$\frac{\partial \mathcal{L}(\Omega, V_{obs})}{\partial \Omega} = U^T U - \Omega^{-1} T \tag{16}$$

For each epoch, we computed the gradient using equation (16) and with this gradient, updated our estimate of Ω with equation (8). We ran the learning algorithm over 5,000 iterations and found that it converged with an error of 0.0494 by the end of the iterations. We illustrate the convergence in Figure 3 below, which displays the errors over the learning.

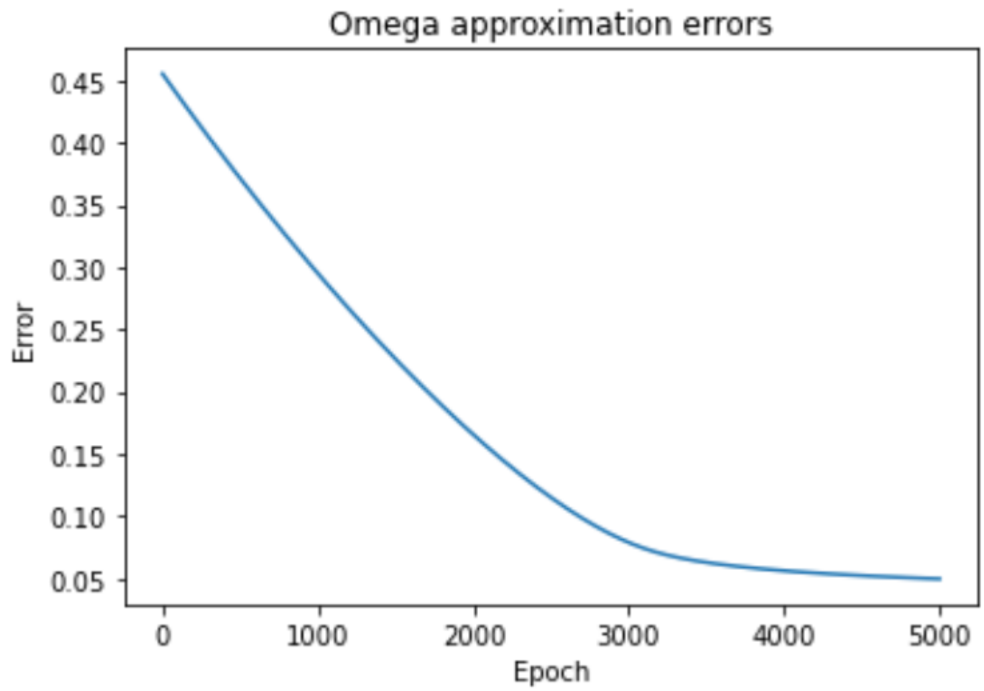


Figure 3: Diminishing error of random Ω approximation over iterations, converging at 4.94% error after 5000 iterations.

To confirm the convergence suggested by Figure 3, we illustrate the original random Ω, which was simply initialized as the identity matrix, alongside the approximated random Ω and the true Ω in Figure 4. This comparison clearly shows the convergence of random Ω to true Ω based on the similarities of Figure 4(b) and Figure 4(c).

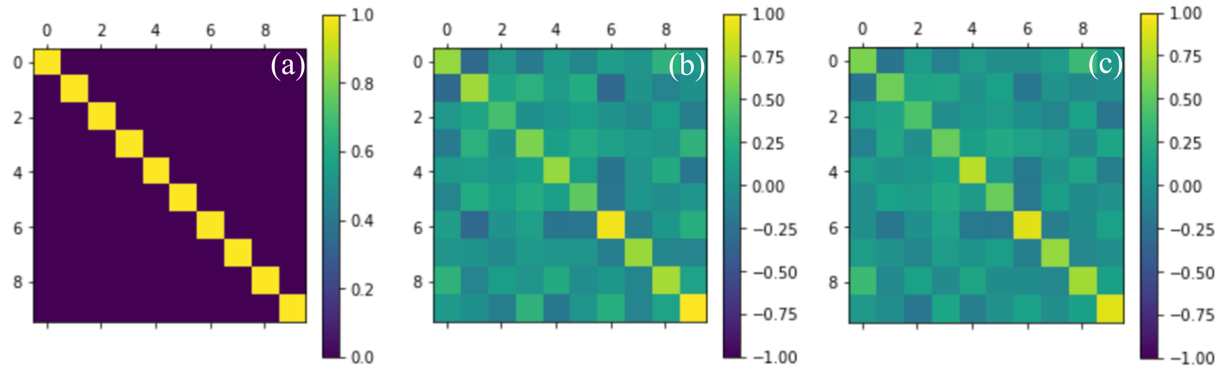


Figure 4: Plot (a) displays the original random Ω generated as the identity matrix, (b) illustrates the approximation of this matrix by the end of 5000 iterations, and (c) shows the true Ω.

The results illustrated in Figure 3 and Figure 4 confirmed our supposition and allowed us to use the proposed equation (16) to determine the Ω gradient, giving us a fairly clear-cut path forward.

Chapter 5: Discussion and Conclusion

With the correctness of the underlying mathematical basis for the model established by our preliminary results, we discuss the implications of our research and the next steps we intend to take for our model extension.

5.1 Implications and future direction

The results derived from the programming testbed confirmed the correctness of equation (16) in determining the Ω gradient; at this stage, we have derived the necessary changes needed to be made in the mathematical formulation of the original PrAGMATiC to replace the area functional means with covariances of the functional values.

Though we have made the changes in the code based on the preliminary results, due to time constraints, we have not finalized the changes to the dataset to run the model with visualizations effectively. This leaves a clear avenue for future direction on the ongoing research – to implement the necessary changes to PrAGMATiC so as to visualize the results of learning the covariances.

Toward this end, we note first that it is nontrivial to visualize the learned Ω given that the matrix has dimensions of number of centroids by number of centroids, when we ultimately would like to visualize this as a 1 by number of vertices vector. With this line of reasoning, it is possible to find the first eigenvector $\boldsymbol{v}$ of Ω to project that onto the brain. This can be achieved by computing $\boldsymbol{v} * B$. The first eigenvector of the covariance matrix represents the ordering of the values whose covariances are being computed along one dimension. In more detail, this first eigenvector of the covariance matrix, sometimes referred to as the principal gradient of connectivity, has been shown to account for the positions of networks over the primary sensory and motor regions [8]. Values that covary significantly will end up in similar places, whilst values that do not covary as much will appear in different places along that dimension. This is how we will approach the problem of visualizing the covariance learning of our extension of PrAGMATiC.

5.2 Conclusion

Extending PrAGMATiC as a covariance-based model presents a novel approach of generating functional maps across the cortex. Building off of the original PrAGMATiC algorithm's formulation of the functional mapping of the brain as a probability distribution that can be used to sample subjects, our modification to the model will estimate the covariance of the functional values rather than directly estimating the means of the areas. This will allow us to make direct comparisons between subjects – an assumption that PrAGMATiC has to make – thus enabling us to generate cortical parcellations using exclusively resting state fMRI.

References

[1] Ballard, D. H. Brain computation as hierarchical abstraction. Cambridge, The MIT Press, 2015.

[2] Damoiseaux, J. S., Rombouts, S. A. R. B., Barkhof, F., Scheltens, P., Stam, C. J., Smith, S. M., & Beckmann, C. F. Consistent resting-state networks across healthy subjects. *PNAS,* 103(37):13848-13853 (2006). doi: 10.1073/pnas.0601417103

[3] De Luca, M., Beckmann, C. F., De Stefano, N., Matthews, P. M., & Smith, S. M. fMRI resting state networks define distinct modes of long-distance interactions in the human brain. *NeuroImage,* 29(4):1359-1367 (2006). doi: 10.1016/j.neuroimage.2005.08.035

[4] Glasser, M., Coalson, T., Robinson, E., Hacker, C. D., Harwell, J., Yacoub, E., Ugurbil, K., Andersson, J., Beckmann, C. F., Jenkinson, M., Smith, S. M., Van Essen, D. C. A multi-modal parcellation of human cerebral cortex. *Nature,* 536:11-18 (2016). https://doi.org/10.1038/nature18933 doi: 10.1038/nature18933

[5] Huth, A. G., Griffiths, T. L., Theunissen, F. E., & Gallant, J. L. PrAGMATiC: a Probabilistic and Generative Model of Areas Tiling the Cortex. *arXiv: Quantitative Methods* (2015). url: https://arxiv.org/abs/1504.03622

[6] Khare, K. & Rajaratnam, B. Wishart distributions for decomposable covariance graph models. *The Annals of Statistics*, 39(1):514-555 (2011). doi: 10.1213/10-AOS841

[7] Kwong, K. K., Bellieveau, J. W., Cesler, D. A., Goldberg, I. E., Weisskoff, R. M., Poncelet, B. P., Kennedy, D. N., Hoppel, B. E., Cohen, M. S., Turner, R., *et al.* Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. *Proc Natl Acad Sci U S A,* 89(12):5675-9 (1992). doi: 10.1073/pnas.89.12.5675

[8] Margulies, D. S., Ghosh, S. S., Goulas, A., Falkiewicz, M., Huntenburg, J. M., Langs, G., Bezgin, G., Eickhoff, S. B., Castellanos, F. X., Petrides, M., Jefferies, E., & Smallwood, J. Situating the default-mode network along a principle gradient of macroscale cortical organization. *PNAS*. 113(44):12574-12579 (2016). doi: 10.1073/pnas.1608282113

[9] Penfield, W. & Boldrey, E. Somatic motor and sensory representation in the cerebral cortex of man as studied by electrical stimulation. *Brain: A Journal of Neurology*, 60:389-443. doi: 10.1093/brain/60.4.389

[10] Rao, R. P. N. & Ballard, D. H. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, 2:79-87 (1999). doi: 10.1038/4580

[11] Schaefer, A., Kong, R., Gordon, E. M., Laumann, T. O., Zuo, X. N., Holmes, A. J., Eickhoff, S. B., & Yeo, B. T. T. Local-global parcellation of the human cerebral

cortex from intrinsic functional connectivity MRI. *Cereb Cortex,* 28(9):3095-3114 (2018). doi: 10.1093/cercor/bhx179

[12] Sellentin, E. & Heavens, A. F. Parameter inference with estimated covariance matrices. *Monthly Notices of the Royal Astronomical Society,* 456:132-136 (2016). doi: 10.1093/mnrasl/slv190

[13] Wandell, B. A., Dumoulin, S. O., & Brewer, A. A. Visual field maps in human cortex. *Neuron,* 56(2):366-383 (2007). doi: 10.1016/j.neuron.2007.10.012

[14] Yeo, B. T. T. & Eickhoff, S. B. A modern map of the human cerebral cortex. *Nature* 536:152-154 (2016). doi: 10.1038/nature/18914

[15] Zilles, K. Brodmann: a pioneer of human brain mapping—his impact on concepts of cortical organization. *Brain,* 141(11):3262-3278 (2018). doi: 10.1093/brain/awy273