Brain subcortical functional activity analysis with fMRI data

Ha Nguyen

May 14, 2024

1 Problem Statement

In this project, we apply Vector Autoregressive (VAR) models to functional magnetic resonance imaging (fMRI) time series to examine the temporal relationships between functional activity among brain regions. We consider the time series data of 49 healthy adults observed over same time horizon across four complete resting state fMRI runs of $p = 86$ brain regions, each scan with $n = 1200$ time points.

In my current research, we consider n samples of fMRI signals $\{X_t\}_{t=1}^{n} \sim \mathbb{N}_p(0, \Sigma)$. The off-diagonal entries of the precision matrix $\Omega = \Sigma^{-1}$ encode partial correlation information $\omega_{ij} = Cor(X_{tr}, X_{ts} \mid X_{\setminus t\{r,s\}})$. Under i.i.d assumption, we would like to incorporate some prior knowledge, such as structural connectivity (SC) of the brain, in our estimation of functional connectivity Ω.

As an extension, we explore the dependence of fMRI time series on their past signals in this project. Under stationary assumption, we run VAR(l) model and estimate the coefficients matrices A_j:

$$X_t = \sum_{j=1}^{l} A_j X_{t-j} + \epsilon_t, \epsilon_t \sim \mathbb{N}(0, \Sigma)$$

In this project, we focus on $p = 6$ main subcortical regions: left and right thalamus-proper, hippocampus and amygdala. Previous analyses indicated that the block structure of functional connectivity, measured as Pearson correlation of fMRI time series, is consistent across subjects (Figure 1). In this analysis, we present data visualization of one subject for illustration, and apply the same VAR model to all subjects. We then explore the relationship between the AR coefficients and SC, which represents the integrity of inter-regional white matter pathways estimated from diffusion MRI (dMRI).

2 Data

The data set used here is a publicly available, high-resolution, preprocessed magnetic resonance imaging data from the Human Connectome Project

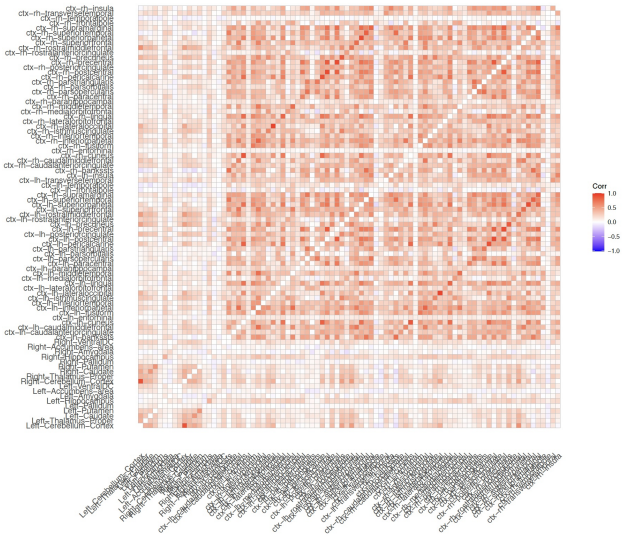

Figure 1: Functional connectivity of a subject

(HCP)—Young Adult S1200 release [Van Essen et al., 2013].

Preprocessing: MRIs were acquired on a Siemens Skyra 3 T scanner at Washington University in St. Louis. Each subject underwent four gradient-echo EPI resting-state functional MRIs (rsfMRI), with TR = 720 ms, TE = 33.1 ms, 2.0 mm isotropic voxels, FoV = 208 × 180 mm^2, flip angle = 52°, matrix = 104 × 90, 72 slices. Two 15 min rsfMRI runs were acquired at each of the two sessions. In the end, the data consisted of 1200 time points for each run, for a total of 4800 time points for each subject over four scans. Each run of each subject's rfMRI was preprocessed by the HCP consortium [Smith et al., 2013]. Gray matter was parcellated into 86 cortical, subcortical and cerebellar regions using FreeSurfer [Fischl et al., 1999] and regional time series were calculated as the mean of the time series in each voxel of a given region.

86×86 structural connectivity matrix: tractograms were used to produce region-of-interest-volume normalised pairwise structural connectivity. More details on how SC was obtained can be found in

Dhamala et al. [2021].

3 Visualization

We analyze time series data from six subcortical regions of a participant (Figure 2 and 3). The data does not exhibit any outliers, shifts, specific trends, or seasonality. The volatility is consistent over time for all times series. Movements appear to be similar between corresponding regions in the left and right hemispheres, with some slightly persistent upward or downward runs within these time series.

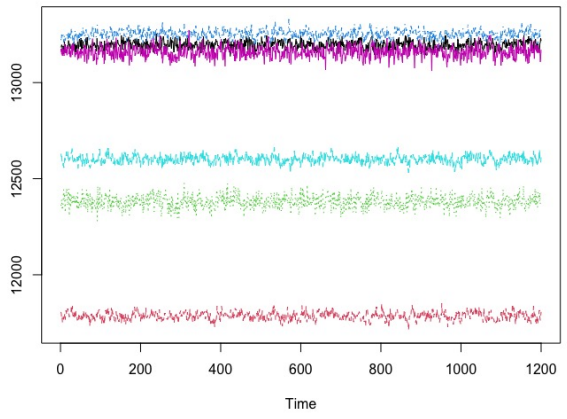

Figure 2: Six subcortical regions

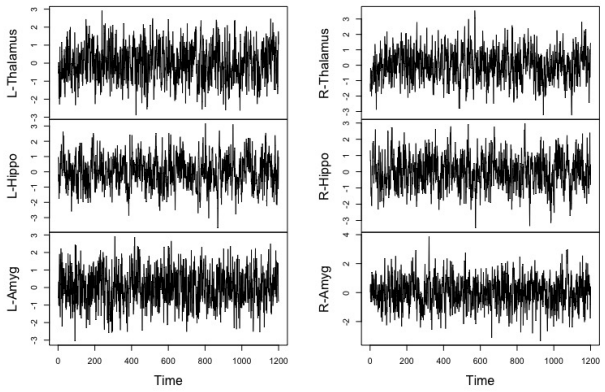

Figure 3: Normalized times series of all regions

The ACF plots for all time series show exponential decay, and the PACF suggest significant lags up to lag 7 for all regions (Figure 4). The only exception is the left hippocampus, where the PACF plot shows significant correlations at lag 1, 2, 27, and 29.

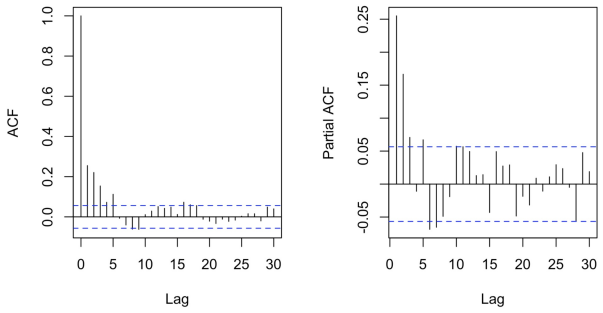

Figure 4: ACF and PACF of the left thalamus

Figure 5 displays the cross-correlations of regions in the left hemisphere, indicating significant cross-correlations of the hippocampus with the thalamus up to lag 6, and with the amygdala up to lag 5. This pattern also appears in the right hemisphere with cross-correlations up to lag 4.

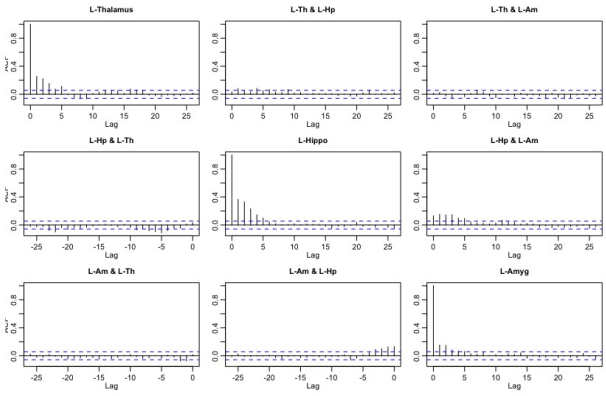

Figure 5: Cross-correlation of left hemisphere

Figure 6 shows high cross-correlations up to lag 5 for the thalamus between the left and right hemispheres. The hippocampus and amygdala exhibit similar patterns up to lags 4 and 3, respectively.

4 Analysis

4.1 Check stationarity

Augmented Dickey-Fuller Test shows that all six times series are stationary with p-values < 0.01.

4.2 VAR(2) model

Using AIC criteria for a parsimonious model, it appears that $l = 2$ would be the best fit.

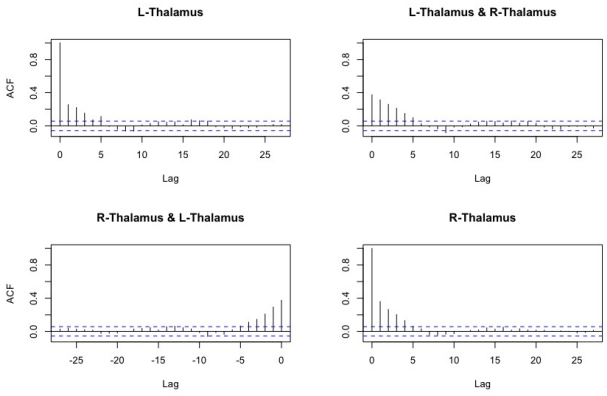

Figure 6: Cross-correlation between left and right thalamus

```
selected order: aic =  2
selected order: bic =  1
selected order: hq =  2
Summary table:
       p     AIC     BIC      HQ     M(p) p-value
 [1,]  0 -0.4515 -0.4515 -0.4515   0.0000  0.0000
 [2,]  1 -0.9971 -0.8444 -0.9395 716.1235  0.0000
 [3,]  2 -1.0876 -0.7822 -0.9725 177.0937  0.0000
 [4,]  3 -1.0753 -0.6172 -0.9028  55.8763  0.0184
 [5,]  4 -1.0592 -0.4484 -0.8291  51.0612  0.0494
 [6,]  5 -1.0382 -0.2747 -0.7506  45.2700  0.1383
 [7,]  6 -1.0177 -0.1015 -0.6726  45.4803  0.1337
 [8,]  7 -0.9859  0.0830 -0.5832  32.3119  0.6447
 [9,]  8 -0.9611  0.2605 -0.5009  40.1695  0.2906
[10,]  9 -0.9236  0.4507 -0.4059  25.5530  0.9023
[11,] 10 -0.9043  0.6227 -0.3291  45.8568  0.1257
```

Figure 7: Select lag order for VAR model

In Figure 8, we run a VAR(2) model without the intercept. All AR coefficients have standard errors between 0.02 and 0.03. Since the standard errors are of similar magnitude, we highlight the coefficients with higher values in yellow. (*) denotes the coefficients with t-statistics > 2. Here, $y_{1,t}, y_{2,t}, y_{3,t}, y_{4,t}, y_{5,t}, y_{6,t}$ represent the left thalamus, hippocampus, amygdala, and their counterparts in the right hemisphere, respectively.

The results show that all time series, except for the left amygdala, exhibit dependence on their past returns. Notably, apart from the left amygdala, all time series show dependence on the past returns of their counterparts in the opposite hemisphere (around 0.1 to 0.2). Additionally, the left hippocampus and amygdala exhibit low dependence on each other's past returns (around 0.06). The residual covariance matrix indicates that corresponding regions on both sides of the brain are concurrently correlated (0.1 to 0.2).

$$\begin{pmatrix} y_{1,t} \\ y_{2,t} \\ y_{3,t} \\ y_{4,t} \\ y_{5,t} \\ y_{6,t} \end{pmatrix} = \begin{pmatrix} 0.1168^{*} & 0.0525 & 0.0501 & 0.19584^{*} & 0.0307 & -0.0150 \\ -0.0178 & 0.2119^{*} & 0.0704 & 0.00892 & 0.1588^{*} & -0.0124 \\ -0.0582 & 0.0651^{*} & 0.0939^{*} & -0.03514 & 0.0259 & 0.0756^{*} \\ 0.1457^{*} & 0.0435 & 0.0257 & 0.23802^{*} & 0.0388 & -0.0541 \\ -0.0316 & 0.1655^{*} & 0.0383 & 0.03441 & 0.2077^{*} & -0.0368 \\ 0.0181 & 0.0560 & 0.1099^{*} & -0.04380 & 0.0850 & 0.1249^{*} \end{pmatrix} \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \\ y_{3,t-1} \\ y_{4,t-1} \\ y_{5,t-1} \\ y_{6,t-1} \end{pmatrix} +$$

$$\begin{pmatrix} 0.0932^{*} & -0.02069 & -0.0381 & 0.11332^{*} & 0.02599 & -0.00169 \\ -0.0476 & 0.17639^{*} & 0.0628^{*} & -0.01739 & 0.02958 & -0.01498 \\ -0.0393 & 0.03743 & 0.0905^{*} & -0.01315 & -0.00625 & 0.06363^{*} \\ 0.0621^{*} & 0.00252 & -0.0206 & 0.10051^{*} & 0.02675 & -0.01700 \\ 0.0535 & 0.13834^{*} & 0.0154 & -0.04180 & 0.10613^{*} & 0.01436 \\ 0.0100 & -0.01156 & 0.0901^{*} & -0.00904 & -0.03243 & 0.09979^{*} \end{pmatrix} \begin{pmatrix} y_{1,t-2} \\ y_{2,t-2} \\ y_{3,t-2} \\ y_{4,t-2} \\ y_{5,t-2} \\ y_{6,t-2} \end{pmatrix} + \begin{pmatrix} a_{1,t} \\ a_{2,t} \\ a_{3,t} \\ a_{4,t} \\ a_{5,t} \\ a_{6,t} \end{pmatrix}$$

$$\text{where } Cov(a_t) = \Sigma = \begin{pmatrix} 0.8476 & 0.0027 & 0.0446 & 0.2093 & 0.0523 & -0.0242 \\ 0.0027 & 0.7865 & 0.0499 & 0.0208 & 0.2036 & 0.0513 \\ 0.0446 & 0.0499 & 0.9295 & -0.0183 & 0.0551 & 0.1239 \\ 0.2093 & 0.0208 & -0.0183 & 0.8150 & 0.0764 & -0.0170 \\ 0.0523 & 0.2036 & 0.0551 & 0.0764 & 0.7849 & 0.0965 \\ -0.0242 & 0.0513 & 0.1239 & -0.0170 & 0.0965 & 0.9070 \end{pmatrix}$$

Figure 8: VAR(2) model

Most of the coefficients, besides the ones commented, are insignificant. This aligns with our expectation since the times series are resting state fMRI.

4.3 Model diagnostics

Model diagnostic plots and the Ljung-Box test shows that the residuals of each times series meet the white noise assumption. We also note that although there are moderate correlations in the residual of corresponding regions on both sides of the brain, there is no Granger causality relationship.

Since the focus of this project is on estimating the AR coefficients, we did not perform any forecasting.

4.4 All participants

We run separate VAR(2) models without intercept for each of the 49 participants with only the significant coefficients. We then explore the relationship between fMRI temporal dependence (AR coefficients) and strength of white matter tracts (SC) of the same subcortical regions across the two hemispheres.

Figure 9 suggests there is no relationship between SC and the temporal dependence of a region on the right hemisphere on lag 1 signals of its left counterpart. The thalamus and hippocampus generally have higher correlations compared to the amygdala. We notice the same pattern with the temporal dependence of the right region on lag 1 signals of its left counterpart.

Figure 10 exhibits a similar pattern for lag 2, although the correlations in the thalamus and hippocampus tend to be lower in magnitude compared to lag 1. Overall, SC does not seem to be a good candidate for prior information when discovering temporal relationship of fMRI signals.

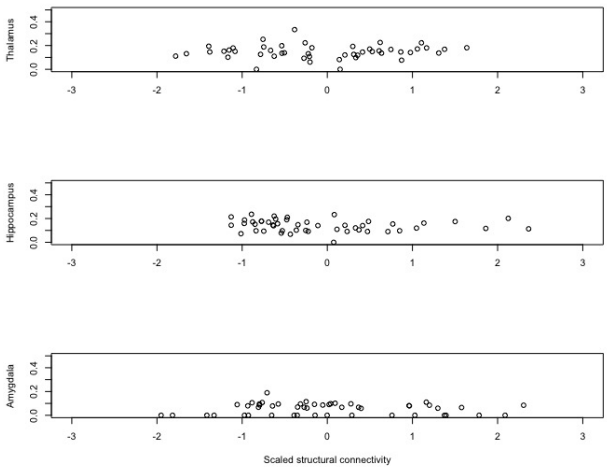

Figure 9: Temporal dependence of a region on the right hemisphere on lag 1 signals of its left counterpart

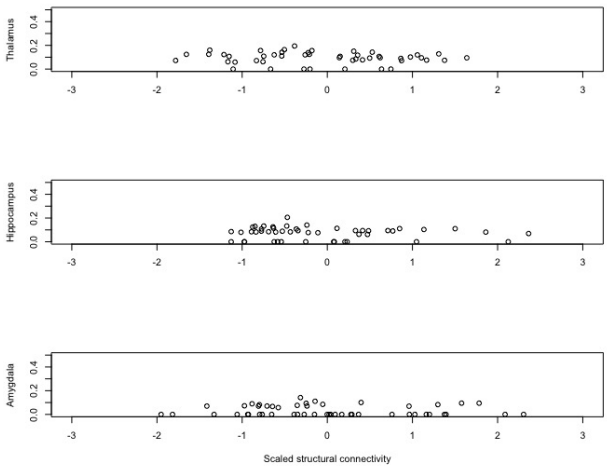

Figure 10: Temporal dependence of a region on the right hemisphere on lag 2 signals of its left counterpart

5 Conclusion

In this study, we apply VAR(2) models to fMRI time series data to investigate the dynamic relationships between functional activities in six main subcortical brain regions. Our analysis reveals that most regions, except for the left amygdala, depend on their past values and those of their counterparts in the opposite hemisphere. The left hippocampus and amygdala also showed interdependence on each other's past signals. The residual covariance matrix shows concurrent correlations between corresponding regions on both sides of the brain.

Model diagnostics indicate that the residuals meet the white noise assumption and there was no evidence of Granger causality between corresponding regions. Additionally, our exploration of the relationship between AR coefficients and SC suggest that SC might not be a suitable candidate for prior information in discovering lag relationships.

A good next step would be to apply VAR(l) model to all $p = 86$ regions of the brain. In this case, it would be beneficial to explore a sparse estimator for the AR coefficients such as lasso.

References

Ekta Dhamala, Keith W. Jamison, Abhishek Jaywant, Simon Dennis, and Amy Kuceyeski. Distinct functional and structural connections predict crystallised and fluid cognition in healthy adults. *Human Brain Mapping*, 42(10):3102–3118, Jul 2021. doi: 10.1002/hbm.25420.

Bruce Fischl, Martin I. Sereno, and Anders M. Dale. Cortical surface-based analysis: Ii. inflation, flattening, and a surface-based coordinate system. *NeuroImage*, 9(2):195–207, 1999. ISSN 10538119. doi: 10.1006/nimg.1998.0396.

S. M. Smith, C. F. Beckmann, J. Andersson, E. J. Auerbach, J. Bijsterbosch, G. Douaud, E. Duff, D. A. Feinberg, L. Griffanti, M. P. Harms, et al. Resting-state fmri in the human connectome project. *Neuroimage*, 80:144–168, 2013.

D C Van Essen, S M Smith, D M Barch, T E Behrens, E Yacoub, K Ugurbil, W-M H Consortium, et al. The wu-minn human connectome project: an overview. *Neuroimage*, 80:62–79, 2013.