Notes on the article : A Graph Signal Processing perspective on functional brain imaging

Ines Vati

2023–10–22

1 Expected project outcomes

- reading the paper, implement it in Python and launch experiments on real time series

- report (5 pages) + source code (jupyter notebook)

- oral presentation 10 min

1.1 TODO

section to be modified
- get familiar with pygsp : https://pygsp.readthedocs.io/en/stable/tutorials/intro.html

- clean notebook

- push a notebook with the graph built on the BOLD5000 dataset

- read paper about graph randomization : A Spectral Method for Generating Surrogate Graph Signals

1.2 What have been done so far ?

- tuto intro filtering pygsp https://pygsp.readthedocs.io/en/stable/tutorials/intro.html. This tuto illustrate the notion of graph fourier basis and graph filtering.

- apply alignement and liberality filtering on BOLD5000

2 Contributions

Functional graph (nodes associated to brain regions and edges to structural or functional connections) are built based on functional connectivity, which is a pairwise measure of statistical interdependency between pairs of regional activity traces. This article is a review on GSP for brain imaging which aims at :

- discuss their potential to integrate brain structure, contained in the graph itself, with brain function, residing in the graph signals

- show how brain activity can be meaningfully filtered based on concepts of spectral modes derived from brain structure

- how GSP tools can be applied to provide insights into the dynamics of functional brain activity.

The connectivity structure can be measured directly by counting fiber tracts in diffusion-weighted MRI or can be inferred from fMRI BOLD measurements.

They introduce some graph signal operations :

- filtering

- randomization preserving anatomical smoothness

- localized decompositions

3 Section II: Brain graphs and brain signals

contruction of a graph from structural connectivity and considering brain activity as graph signal

A weighted graph $\mathcal{G} = (\mathcal{V}, A)$ where $\mathcal{V} = 1, \ldots, N$ is a set of N nodes associated with specific brain regions and $A \in \mathbb{R}_+^{N \times N}$ is a weighted adjacency matrix with entries $A_{i,j}$ representing the strength of physical connection between brain regions i and j. In other words, $A_{i,j}$ corresponds to the strength of axonal connection measured by a simple count of the number of streamlines (estimated individual fibers) that connect the regions, estimated with diffusion spectrum imaging (DSI). $x \in \mathbb{R}^N$ is acquired brain activity signals such that the value of ith component x_i quantify neural activity in brain region i.

4 Section III: GSP for Neuroimaging

derive the graph spectral domain by the eigendecomposition of a graph shift operator. Such eigenmodes provide robust representation of the connectome in health and disease.

- Graph laplacian : $L = D - A$

- Symmetric Normalized Graph Laplacian : $L_{sym} = D^{-1/2} L D^{-1/2}$

- Random-walk graph laplacian : $L_{rw} = D^{-1} L$

- Graph shift operator (assuming henceforth that S is diagonalizable) $S = V \Lambda V^{-1}$ where Λ is a diagonal matrix containing the eigenvalues λ_k
  get a transformation that characterizes exchanges between neighboring nodes.
  using adjacency matrix A as shift vector vs using graph laplacian

Several well known signal processing problem (V for graph Laplacian as a shift operator is a common solution)

- Laplacian embedding : find a mapping of graph nodes on a line so that connected nodes stay as close as possible

- Graph cut problem : partitio a graph with subcommities of nodes with as few cross-connection as possible

4.1 Graph Fourier transform

Definition 1. *Graph Fourier Transform Consider a signal $x\mathbb{R}^N$ and a graph shift operator $S = V\Lambda V^{-1} \in \mathbb{R}^{N \times N}$. Then, $\tilde{x} = V^T x$ and $x = V\tilde{x}$ form a GFT pair.*

A GFT is equivalent to DFT for cyclic graphs.

4.2 Graph signal filtering

We define diagonal filtering matrix G, where $[G]_i = g(\lambda_i)$ is the frequency response for the graph frequency associated with λ_i, and retrieve the filtered signal as $Y_G = VGV^{-1}X$.

e. g. ideal low-pass frequency : $g(\lambda_i) = 1$ for λ_i corresponding to low frequencies and $g(\lambda_i) = 0$ otherwise.

Filters are usually defined in the spectral domain.

4.3 Generation of graph surrogate signal

We want to assess significance of obtained results throught statistical testing (invalidate the null hypothesis).

e. g. we want to test if the observed signal is significantly different from a random signal[1]. We want to preserve the structure of the graph while randomizing the signal.

For correlated data, the Fourier phase randomization procedure $Y = XF^H\phi_{time}F$ where ϕ_{time} contains random phase factors. In the graph setting, this procedure can be generalized $Y = V\phi_{graph}V^T X$ where ϕ_{graph} is a diagonal matrix containing random phase factors (random sign flips of the graph spectral coefficients).

The surrogate signal Y is a random signal with the same power spectrum as X but with randomized phase and the nonstationary spatial effects are destroyed.

---

[1]Parametric test like t-test. Nonparametric tests like permutation test (=re-randomization ou shuffle test) : The null hypothesis is that all samples come from the same distribution $H_0 : F = G$. Under the null hypothesis, the distribution of the test statistic is obtained by calculating all possible values of the test statistic under possible rearrangements of the observed data.

They applied the method developped in [1] to compute excursions.

Dans le contexte du traitement du signal et de l'analyse de séries temporelles, la méthode de *randomization* de phase implique la création d'un jeu de données de "substitution" ou *surrogate data* à partir d'une série temporelle donnée. Les *surrogate data* servent de base nulle ou de référence par rapport à laquelle les propriétés statistiques de la série d'origine peuvent être comparées. Le but de générer des données préservant certaines propriétés de la série temporelle comme son spectre/power spectrum density ou son autocovariance tout en randomisant d'autres aspects. Plus précisément, dans le processus de randomisation de phase, les ensembles de données de substitution sont créés en randomisant les phases de la série temporelle originale tout en maintenant inchangées les amplitudes des composantes de Fourier. Ce processus perturbe les relations temporelles et les motifs dans les données tout en préservant le spectre de puissance global.

Cette méthode est souvent utilisée dans le cadre de tests d'hypothèses sur la structure non aléatoire ou les composantes oscillatoires présentes dans une série temporelle. En comparant les mesures statistiques ou les caractéristiques des données d'origine à celles des données de substitution, on peut évaluer si les motifs observés sont statistiquement significatifs ou simplement dues au hasard. Par exemple, cette technique est utilisée en neurosciences pour évaluer la signification des oscillations neuronales observées ou des motifs de synchronisation entre les neurones.

De même que pour la méthode classique où les amplitudes des coefficients de la transformée de Fourier sont conservées dans les données de substitution, aprés avoir transformé le signal du graphe dans le domaine spectrale de fourier, ils effectuent une randomisation en permutant ou en générant aléatoirement le signe des coefficients de Fourier $\tilde{x} = V^T x$.

4.4 Wavelets and Slepians on graph

- Wavelets : a function $\psi \in L^2(\mathbb{R})$ is an orthonormal wavelet if it can be used to define a Hilbert basis, that is complete and orthonormal system for the Hilbert space $L^2(\mathbb{R})$

- Slepians functions: a set of orthogonal functions that are maximally concentrated in a given frequency band. They are commonly used in spectral analysis and as window functions for reducing spectral leakage

[2] and improving frequency resolution

a localized decomposition for graph signals that is based on a generalization of Slepians functions:
$\rightarrow$ problem of retrieving a signal $x \in \mathbb{R}^N$ that is maximally concentrated within a subset of nodes. As the global concentration of a signal is given by $x^T x$, we end up maximizing : μ

5 Section IV: application to the context of attention task

Tp study the association between brain signal and attention switching, they decomposed the functional brain response into two components :

- one that represents alignment with structural connectivity (regions that activate together are also physically wired)

- one that describes liberality (areas that exhibits high signal variability)

1. graph signal filtering with two different filtering matrices
   ψ_{Al} : low pass filtering for alignment; so that $Y_{\phi_{Al}} = V\psi_{Al}V^T X$ is the transformed functional data in which only the then lowest frequency modes are expressed
   ψ_{Li} : high-pass filtering; for liberality, only the ten highest frequency modes
   to derive a subject-specific alignment and liberality concentration, they compute the norms of the filtered signal, averaged across all temporal samples of a given subject.

2. derive subject-specific alignment or liberality (a scalar) : average norm of filtered signal across all temporal samples

3. relate signal alignment and liberality to cognitive performance of each subject :
   they compute the partial Pearson's correlation between their concentration measures and "switch cost" (response time during switching

---

[2]Spectral leakage means that the energy from one frequency component can "leak" into adjacent frequency bins, making it challenging to accurately identify frequencies.

task blocks compared to non-switching task blocks)
age, gender, and motion were included as covariate to remove their impact from the results

6   Section V: Perspectives

excursions : popular moments in time when entering a regime of strong alignment, or liberality, with the underlying brain structure.

1. resolving excursions in alignment and liberality regimes

    (a) Graph surrogate method : generate 1000 null signal matrices using graph domain randomization $\rightarrow$ operator ϕ_{graph}

    (b) combined with filtering operations to generate null data for aligned and liberal signal components $Y = V\phi_{graph}\phi_{Al}V^{T}X$ and $Y =$ respectively

    (c) threshold the filtered signal, at an α-level of 5% to locate significant signal excursions

    (d) other approach with classical Fourier phase randomization procedure

2. combining graph excursion with Fourier analysis

3. probing excursion with a subgraph using Slepians

7   Ideas / Remarks

data ? https://bold5000-dataset.github.io/website/index.html
download only data from Openneuro for at least 2 participants.

Experiment :

- 4 subjects named CSI1, CSI2, CSI3, CSI4

- Functional task data acquisition 15 sessions :

- 9 or 10 functional scans of slow event-related 5000 scene data (labeled 5000scenes) runs were acquired during the same scanning sessions

- 0 or 1 functional localizer scans used to define scene selective regions (labeled localizer)

- each event.json file lists each stimulus, the onset time, and the participant's response (participants performed a simple valence task)

- Anatomical data acquisition session: 16 Anatomical Data: T1 weighted MPRAGE scan, a T2 weighted SPACE, diffusion spectrum imaging

Pre-processing : fMRIprep: Preprocessed data for all functional data of CSI1 through CSI4 (In the derivatives folder)

- preprocessed both in T1w image space and on surface space. Functional data was motion corrected, susceptibility distortion corrected, and aligned to the anatomical data using bbregister.

- refer to the paper for the details on preprocessing.

- Reports resulting from fMRI prep, which include the success of anatomical alignment and distortion correction, among other measures of pre-processing success are all listed in the sub-CSIX.html files.

All MRI and fMRI data provided is with Siemens pre-scan normalization filter.

use nilearn to create functional connectome :

8   Understand the data

8.1   What is fMRI ?

Functional magnetic resonance imaging (fMRI) is based on the fact that when local neural activity increases, increases in metabolism and blood flow lead to fluctuations of the relative concentrations of oxyhaemoglobin (the red cells in the blood that carry oxygen) and deoxyhaemoglobin (the same red cells after they have delivered the oxygen). Oxyhaemoglobin and deoxyhaemoglobin

have different magnetic properties (diamagnetic and paramagnetic, respectively), and they affect the local magnetic field in different ways. To record cerebral activity during functional sessions, the scanner is tuned to detect this "Blood Oxygen Level Dependent" (BOLD) signal.

Brain activity is measured in sessions that span several minutes, during which the participant performs some cognitive task and the scanner acquires brain images, typically every 2 or 3 seconds (the time between two successive image acquisition is called the Repetition time, or TR).

Before fMRI images can be used to do meaningful comparisons, they must be processed to ensure that the voxels that are being compared represent the same brain regions, irrespective of the variability in size and shape of the brain and its microarchitecture across different subjects in the experiment. The process is called spatial registration or spatial normalization. During this procedure, the voxels of all the brain images are 'registered' to correspond to the same region of the brain. Usually, the images are registered to a standard 'template' brain image. Once this is done, the coordinate of a voxel are in the same space as the template. nilearn package does not perform spatial preprocessing; it only does statistical analyses on the voxel time series. For preprocessing functions, users are referred to Nipype or fMRIPrep.

8.2 BOLD5000 dataset

Spatial normalization to the ICBM 152 Nonlinear Asymmetrical template version 2009c was performed.

References

[1] E. Pirondini, A. Vybornova, M. Coscia, and D. Van De Ville, "A spectral method for generating surrogate graph signals", IEEE Signal Processing Letters **23**, 1275–1278 (2016).