MNE-Python: Improved AR Connectivity

About me

- Jordan Drew (github username: jadrew43)
- University of Washington / Electrical and Computer Engineering / 5th year / expected graduation August 2023
- Time zone: Los Angeles, USA (PDT)

Previous Experience

- I have processed multiple neuroimaging datasets and produced preliminary data regarding the dynamic functional connectivity between regions of interest during an attention task. The connectivity measurement implements a state-space model which utilizes sensor data and a forward map, developed from subjects' MRI scans informing structural anatomy, to predict the activity across regions of interest. A vector autoregressive model is then implemented in order to determine the cross-regional influence of regions of interest over time. Within these iterative steps of the state-space model implementation, the pipeline additionally utilizes an Expectation Maximization (EM) algorithm to ensure the best fit of the time-varying autoregressive (AR) coefficients using regularization parameters to ensure that the coefficients are not too large as well as promoting temporal smoothness.
- I have used the model's outputs to build graphical representations of functional connectivity networks, where each node represents some region of interest, and directed edges indicate the inferred direction of information flow between regions. This representation can contain weighted edges, where the weight is determined by the average value of AR coefficients within some time of interest. I've also studied how to compute the statistical difference between these graphical representations and determine which edges contribute the most to those significant differences.

Project Information

1) Sub-org: MNE-Python
2) Abstract: Currently, MNE-Python implements vector autoregressive modeling as a method of measuring functional connectivity across sensors used in neuroimaging. The purpose of this project will be to extend the measures of functional connectivity using a state-space modeling approach utilizing the observed variables (sensor measurements) to estimate functional connectivity across latent variables (neurological regions of interest). The output of this method will provide time-varying autoregressive (AR) coefficients, indicative of the connection strength between any two latent variables. The model's output can be used to draw a directed graph as a visualization tool to understand information flow across latent variables. Statistical methods will be implemented in order to determine statistical differences between different network configurations.

3) Detailed description:
    a) Package the VAR/state-space model for measuring functional connectivity described in Yang et al. 2016 as a core method for MNE-Connectivity. Design new function to complement current VAR model with tutorial **outlining the steps of the VAR/state-space model** applied to neuroimaging data, using sensor data as the observed variables and the activity across regions of interest as the latent variables.
    b) Build tutorial including design of functional connectivity across a small number (2-3) of latent variables. Users can manipulate the strength of connections across the latent variables to design ground truth. Latent and observed variable data will be synthesized from the ground truth. Synthesized data will be used as an input to the state-space model. The output will **return an estimate of connectivity across time resembling the ground truth**.
    c) Using an autoregressive model of order 1, where every latent variable is measured in both lead (time = t) and lag (time = t+1) positions, we can infer the direction of information flow between any two latent variables. Estimated connectivity values will be used to **draw directed graphs to provide a visualization tool for functional connectivity across regions of interest**.
    d) The Hotelling T^2 test will be implemented as in Ginestet et al. 2017 in order to **calculate the statistical differences between directed networks of latent variables**, where nodes indicate regions of interest and edges indicate functional connectivity between those regions. The statistics will inform the user of which edges contribute to the significant differences between functionally connected networks, and the relative contribution of edges to those differences.

4) Weekly timeline:
    a) **Community bonding period (May 20 - June 12):** *prepare*
        i) Learn protocols for making pull requests and commits to MNE github
        ii) Set up hardware to contribute to MNE
        iii) Make and get merged two PRs for small feature/bug/enhancement changes to learn the GitHub workflow
    b) **Week 1 (June 13):** *implement state-space model regression with basic kalman filter algorithm*
        i) Observed variables - data from sensor space
        ii) Latent variables - estimated activity from regions of interest
    c) **Week 2 (June 20):** *implement improved EM algorithm (Yang et al. 2016)*
        i) E: estimate latent variable activity using observed variables
        ii) M: fit AR coefficients using regularization parameters
    d) **Week 3 (June 27):** *build toy example using state-space model*
    e) **Week 4 (July 4):** *develop tutorial for measuring connectivity*
        i) Use toy example alone
        ii) Use toy example + EM algorithm (will use simulation to build ground truth and compare with and without EM algorithm)
    f) **Week 5 (July 11):** *traveling*

g) **Week 6 (July 18):** *traveling*
h) **Week 7 (July 25):** *build code for designing ground truth*
    i) Construct 3x3 matrix, each component contains vector to represent time-varying AR coefficients, using SourceSimulator
    ii) Use Hann window to define AR values over fixed number of samples
    iii) Allow for temporal jitter - varying start time for Hann window (can start at start of time period, be centered in time, or end at end of time duration)
    iv) Allow for amplitude variations of AR values (range of -0.2 to 0.2)
i) **Week 8 (August 1):** *build code for simulation data*
    i) Simulate latent and observed variable activity using ground truth connectivity values from latent variable pairs
    ii) Structure data to simulate multiple trials and subjects (show that less data means worse estimates)
    iii) Use subject data as input to state-space model
    iv) Plot ground truth and model output together
j) **Week 10 (August 15):** *build network/graphical representation*
    i) Use Networkx to transform model output into graph/network
        (1) Implement weighted and unweighted versions
    ii) Use Networkx to draw directed graph with customizable features
k) **Week 11 (August 22):** *prep for Hotelling T2*
    i) Use network to compute adjacency and degree matrices
        (1) For weighted and unweighted versions
    ii) Compute Laplacian matrix
    iii) Bootstrap data for multiple estimates of connectivity
    iv) Compute covariance matrix
    v) Compute T1: statistical differences between network and null network
    vi) Compute T2: statistical differences between networks A and B
    vii) Linear decomposition of stats for significance per edge
l) **Week 12 (August 29):** *code freeze*
    i) Finish tests and evals
m) **Final Week (September 5):** *submission*
    i) Submit product and final mentor eval

5) <u>Other Commitments</u>:
    a) **May 20- June 3:** on East Coast with limited availability (can work asynchronously)
    b) **July 9 - 15:** attending conference; presenting preliminary results from previously mentioned state-space model to compare functional connectivity between different subject groups
    c) **July 18 - 25:** vacation
    d) **September 9 - 12:** vacation (can have final submitted by September 8)