

Homework 9: Stabilità strutturale delle matrici di connettività funzionale

In questo homework a partire da dati fMRI acquisiti in condizione di resting state si andrà a caratterizzare la struttura della connettività funzionale a singolo soggetto di una popolazione di soggetti sani. La derivazione ed il pre-processing applicato al dato funzionale saranno parte integrante dello studio che sarà basato su metodologie derivate dalla graph theory. Nella cartella Data sono contenuti i dati funzionali di 22 soggetti acquisiti presso il dipartimento di neuroradiologia dell'ospedale Borgo Trento di Verona. Per ciascun soggetto i dati sono contenuti in una sottocartella A2MBxx, dove xx = 1,...,22 indica il numero del soggetto.

Scopo del lavoro

Lo scopo del lavoro è di caratterizzare, nella popolazione dei soggetti sani, le distribuzioni dei descrittori tipici dei network derivate da functional connectivity (FC) sulla base di metodi di graph theory. Si vuole inoltre valutare l'impatto dell'operazione di thresholding delle matrici FC sugli indicatori.

Tipicamente, infatti, tali descrittori vengono utilizzati in abito di ricerca neurologica per discriminare tra condizione patologica e non, perciò il loro impegno può essere fondamentale anche in ambito clinico.

Dati e informazioni di pre-processing

Il preprocessing delle immagini funzionali include i seguenti step: correzione per slice-timing effect, correzione per il movimento, coregistrazione su spazio T1w e successiva normalizzazione in spazio MNI152.

Inoltre, l'immagine T1w di ciascun soggetto è stata segmentata in materia bianca (WM), grigia (GM) e liquido cefalorachidiano (CSF), e parcellizzata in regioni di interesse secondo uno specifico atlante. L'analisi delle immagini funzionali verrà svolta nello spazio T1w. Per ciascuna delle regioni di interesse presenti nell'atlante è stato estratto il segnale fMRI preprocessato medio, i dati ottenuti sono contenuti nel file atlas_sampled_timeseries.mat.

In aggiunta a questi dati si forniscono anche i risultati di una principal component analysis condotta separatamente sui voxel relativi a WM e CSF. Questi segnali verranno utilizzati in una seconda fase di preprocessing del dato (denoising).

Per ogni soggetto la relativa cartella contiene due file .mat:

estimated_motion.mat contiene le seguenti variabili:

- *motion* = regressori di movimento (225 x 6) (6 = 3 traslazioni [mm] + 3 rotazioni [deg]);
- *comp_white* = componenti principali (225 time x 224 space) dall'analisi PCA su WM;
- *var_white* = varianza (%) spiegata dalle componenti principali *comp_white*;
- *CSFprinc* = componenti principali (225 time x 224 space) dall'analisi PCA su CSF;
- *explained_variance_CSF* = varianza (%) spiegata dalle componenti principali *CSFprinc*.

atlas_sampled_timeseries.mat contiene una struttura (*malROI*) con i seguenti campi:

- *tac* = time course segnale fMRI pre-processato (totale 99 ROI x 225 campioni temporali)

- *name* = descrizione (nome anatomico e sigla) della regione cerebrale di interesse a cui è riferita la serie temporale (nel campo *time_course*) (10 x 99)
- *label* = identificatore ROI (riferite al doc: *Label-Information.xlsx*)

L'atlante utilizzato in questo studio consiste in una matrice 3D di valori interi in cui ogni numero indicizza (campo *label*) una specifica regione anatomica corticale, come descritto nel documento allegato *Label-Information.xlsx*. L'appartenenza delle ROI descritte in tale documento ad aree anatomicamente corticali, subcorticali o altro è desumibile, noto il nome anatomico della ROI da ulteriori numerosi riferimenti on-line.

Analisi richieste

- 1) In seguito ad un errore di valutazione, alcune delle 99 ROI indicate, sono potenzialmente non corrispondenti ad aree cerebrali corticali o sottocorticali. In base ai dati a disposizione fMRI e descrizioni dell'atlante identificare eventuali regioni estranee (non corticali né sottocorticali) e riportarne gli identificatori, è necessario rimuovere tali regioni dal set di confronto? Se sì, perché? (consiglio: rispondete a questo quesito per ultimo).
- 2) I segnali delle regioni di interesse presentano una deriva lineare nel tempo? Se sì, è questa costante in tutte le ROI?
Applicare se ritenuto necessario un filtraggio temporale passa-alto adeguato (es. Butterworth, frequenza taglio 1/200 Hz o superiore se ritenuto necessario) e mantenere separate le dinamiche fMRI filtrate dalle non filtrate.
- 3) Utilizzando, come visto a lezione, un modello lineare di regressione, rimuovere il contributo del movimento e volume parziale relativo a WM e CSF dalle dinamiche fMRI.
Vanno considerate come regressori di volume parziale le sole componenti principali che spiegano almeno il 50% della varianza per WM e 70% per CSF. Sulla base di tale criterio scegliere il numero di componenti adatte per ciascun soggetto ed identificare un numero di componenti mediamente adatto all'intero pool di soggetti. Si riporti il numero di componenti scelte per ciascun soggetto e il numero medio e se ne discutano brevemente le implicazioni. Tipicamente la regressione del movimento viene svolta utilizzando le dinamiche dei parametri stimati in fase di riallineamento e la loro derivata al prim'ordine. Si provi ad ipotizzare la motivazione di tale approccio.
Considerando come confound i regressori di volume parziale WM e CSF (col numero di componenti scelto a livello gruppo), il movimento e la sua derivata, svolgere la regressione lineare di tali contributi per ogni soggetto.
- 4) Per valutare l'efficacia delle precedenti operazioni, visualizzare graficamente l'andamento del segnale regredito da confounds, filtrato e non filtrato (se si è applicato un filtro riportare anche gli spettri) per la ROI 60 (Occipital Pole) per del soggetto A2MB20. Costruire infine per

semplicità una matrice *processed_fmri* (225x99x22) contenente tutte le dinamiche fMRI di ogni ROI regredite per i confound di tutti i soggetti ed eventualmente pre-filtrate.

- 5) Calcolare per ogni soggetto le matrici FC mediante correlazione di Pearson. Memorizzare coerentemente il p-value associato a tali stime (ad esempio vedi comando *corr*, Matlab). Ripetere tale operazione sostituendo lo stimatore di correlazione di Pearson con la correlazione parziale non regolarizzata (ad esempio vedi *partialcorr*, Matlab).
- 6) Applicare sogliatura hard alle matrici ottenute rimuovendo tutte le stime di connettività che non superano la soglia di significatività al 5%, con correzione per confronti multipli (es. Bonferroni oppure una misura ritenuta adeguata).
- 7) Visualizzare in un'immagine con un subplot 2x2 le matrici FC ottenute con le due metodologie appena richieste ed i relativi p-value associati (riportare solamente un soggetto, ad esempio A2MB20).
- 8) Calcolare la matrice FC di gruppo su tutti i soggetti e discuterne la validità rappresentativa nei singoli soggetti. Si scelga almeno un criterio adeguato per la stima della connettività di gruppo discutendone le implicazioni (es. media delle FC singolo soggetto).
- 9) Per ogni soggetto si calcoli a partire dalla FC stimata e sogliata la *global efficiency* (GE) e *characteristic path length* (CPL) utilizzando le funzioni fornite (*charpath.m*, *efficiency_wei.m*). Rappresentare graficamente le distribuzioni sui soggetti di tali descrittori, separandoli per metodo di calcolo della FC utilizzato.
- 10) Le distribuzioni dei due descrittori sui 22 soggetti sono sensibili al metodo utilizzato per calcolare le matrici FC? Quale descrittore di network possiede minore variabilità tra i soggetti? La variabilità osservata è funzione del metodo di calcolo della FC?

Consegnare via email (bertoldo@dei.unipd.it)

- 1) una breve relazione (metodi usati, problemi, unità di misura dei parametri, ...) **non superiore alle 6 pagine (non serve inserire nel report le figure generate dal codice stesso).**
- 2) codici Matlab commentati ed ordinati, separatamente per ogni studente
- 3) file results_NOME_COGNOME.mat con:
 - a. una matrice ($N_{ROI} \times 225 \times 22$) o array di celle ($N_{ROI} \times 225 \times 22$) contenente (per ogni soggetto ed ogni ROI) la serie temporale di segnale BOLD se ritenuto necessario filtrata e regredita per il contributo di movimento, WM e CSF.
 - b. la matrice FC calcolata ($N_{ROI} \times 22 \times 2$) per ogni soggetto con i 2 metodi di correlazione. Siano riportati e giustificati esplicitamente i parametri utilizzati nel calcolo di tali matrici.
 - c. una matrice in cui siano riportati i descrittori di network richiesti (GE, CPL) per ogni soggetto e condizione di calcolo della FC ($2 \times 2 \times 22$).