Directed Spectral Methods

Definition

Directed spectral measures quantify, in the frequency domain, directed statistical interactions between time-series variables. Most commonly, the various measures are computed based on a (linear) multi-variate autoregressive (MVAR) model of the data. The various methods quantify in different ways the strength of interaction terms in the Fourier transform of the MVAR model. While these methods are sometimes referred to as capturing 'causal' or 'effective' connectivity, they are most properly described as reflecting 'directed functional connectivity' (Friston et al. 2013).

Detailed description

For a multi-variate n-channel process $\boldsymbol{X}(t) = [X_1(t), X_2(t), \ldots, X_n(t)]^{\mathrm{T}}$ (with zero mean), the MVAR model is given by

$$\boldsymbol{X}(t) = \sum_{k=1}^{p} A_k \cdot \boldsymbol{X}(t-k) + \boldsymbol{\epsilon}(t)\,, \tag{1}$$

where the A_k are matrices of regression coefficients, p is the model order, and $\boldsymbol{\epsilon}(t)$ are the residuals (Lütkepohl 2007). The coefficients of this model are uniquely specified by imposing zero correlation between the residuals and the regressors $\boldsymbol{X}(t-k)$, $k = 1, \ldots, p$, and can be derived from the Yule-Walker procedure (Ding et al. 2006). Ideally the model order should be sufficiently high to obtain a good representation of the data, but not so high that overfitting and poor parameter estimation occur. Commonly the model order is selected based on minimizing either the Akaike Information Criterion, or the Bayes Information Criterion, both of which balance the residual variance against the number of coefficients to be estimated (Ding et al. 2006).

Directed spectral measures are computed from the frequency domain representation of the MVAR model:

$$A(\omega) \cdot \tilde{\boldsymbol{X}}(\omega) = \tilde{\boldsymbol{\epsilon}}(\omega)\,, \tag{2}$$

where explicitly in terms of the regression coefficients

$$A(\omega) = I - \sum_{k=1}^{p} z^k A_k\,, \tag{3}$$

$z = e^{-2i\pi\omega\Delta t}$ and Δt is the time between observations. (Let tildes denote Fourier transform and dagger the Hermitian conjugate.) The measures are written in terms of $A(\omega)$, the spectral density (or power) $S(\omega) =: \tilde{\boldsymbol{X}}(\omega) \cdot \tilde{\boldsymbol{X}}^{\dagger}(\omega)$, the covariance matrix of the residuals Σ, and also the transfer matrix $H(\omega) =: A^{-1}(\omega)$, which yields $\tilde{\boldsymbol{X}}(\omega)$ when it acts on the Fourier transformed residuals:

$$\tilde{\boldsymbol{X}}(\omega) = H(\omega) \cdot \tilde{\boldsymbol{\epsilon}}(\omega)\,. \tag{4}$$

Non-parametric estimation. An alternative method for computing directed spectral measures involves obtaining the transfer matrix directly from Fourier and wavelet transforms of the data, obviating the need to explicitly fit an MVAR model (Dhamala et al. 2008; Percival and Walden

2000). An advantage of this approach is that it bypasses the need for model order estimation; however this is replaced with the task of choosing the wavelet prototype and number of tapers appropriately.

Granger-Geweke causality (GGC)

For a pair of variables $\boldsymbol{X}(t) = [X_1(t), X_2(t)]^{\mathrm{T}}$, the Granger-Geweke causality $\mathrm{GGC}_{2\to1}(\omega)$ from X_2 to X_1 at frequency ω quantifies the directed contribution of X_2 to the power of X_1 at frequency ω (Ding et al. 2006; Geweke 1982). It does this by expressing the total power of X_1 as the sum of an 'intrinsic' term and a 'causal' term, and comparing the logarithms of the total power and the intrinsic power. From an MVAR model of the two variables, the power can be expressed via (4) as

$$S(\omega) = H(\omega) \cdot \Sigma \cdot H^{\dagger}(\omega)\,. \tag{5}$$

The decomposition of this into an intrinsic term and a causal term is obtained via the transformation of variables $X_2 \to X_2 - (\sigma_{21}/\sigma_{11})X_1$, which diagonalizes Σ and hence removes the cross-terms from the RHS. (Since GGC is interested in the independent causal contribution from X_2 to X_1, it should by definition be invariant under the addition of a multiple of X_1 to X_2.) After this transformation, we have

$$S_{11}(\omega) = H_{11}(\omega)\sigma_{11}H_{11}^{*}(\omega) + H_{12}(\omega)\sigma_{22}H_{12}^{*}(\omega)\,. \tag{6}$$

GGC is given (after the transformation) by

$$\mathrm{GGC}_{2\to1}(\omega) =: \ln\left\{\frac{S_{11}(\omega)}{H_{11}(\omega)\sigma_{11}H_{11}^{*}(\omega)}\right\}. \tag{7}$$

To analyse a system of n variables, one can compute the above pairwise GGC between each pair of variables by fitting MVAR models successively to each pair of variables. There also exists: (i) the more complicated $\mathrm{GGC}_{j\to i|k_1,\ldots,k_\ell}$, i.e. the conditional GGC from X_j to X_i conditional on $X_{k_1},\ldots,X_{k_\ell}$ (Ding et al. 2006; Geweke 1984), which computes in the same fashion the causal contribution of X_j to the power of X_i that is independent of the causal contributions of $X_{k_1},\ldots,X_{k_\ell}$ to the power of X_i and the intrinsic power of X_i; (ii) 'multivariate GGC' from one group of variables to another group of variables (Barrett et al. 2010; Geweke 1982).

The spectral GGC is related to its time-domain formulation (Ding et al. 2006; Geweke 1982; Granger 1969). The time-domain GGC from X_2 to X_1 quantifies the extent to which the past of X_2 helps predict the future of X_1 over and above the extent to which the past of X_1 predicts its own future. Specifically, it is given by the logarithm of the ratio of the residual variance in a restricted regression of X_1 on its past to the residual variance σ_{11} in the (unrestricted) MVAR model of $[X_1, X_2]^{\mathrm{T}}$. The mean spectral GGC over all frequencies up to the Nyquist frequency is equal to the corresponding time-domain GGC.

Partial Directed Coherence (PDC)

The partial directed coherence $\mathrm{PDC}_{j\to i}(\omega)$ is defined as (Baccalá and Sameshima 2001)

$$\mathrm{PDC}_{j\to i}(\omega) =: \frac{A_{ij}(\omega)}{\sqrt{\sum_{k=1}^{n}|A_{kj}(\omega)|^2}}\,. \tag{8}$$

This measure represents the relative coupling strength of the interaction from X_j to X_i, as compared to all of X_j's interactions as a source with other structures. The PDC ranks the relative interaction strengths with respect to a given signal source, and the normalization is such that

$$0 \leq |\mathrm{PDC}_{j\to i}(\omega)|^2 \leq 1\,, \tag{9}$$
$$\textstyle\sum_{i=1}^{n} |\mathrm{PDC}_{j\to i}(\omega)|^2 = 1\,. \tag{10}$$

Directed Transfer Function (DTF)

The directed transfer function $\mathrm{DTF}_{j\to i}(\omega)$ is defined as (Kaminski and Blinowska 1991)

$$\mathrm{DTF}_{j\to i}(\omega) =: \frac{H_{ij}(\omega)}{\sqrt{\sum_{k=1}^{n} |H_{ik}(\omega)|^2}}\,. \tag{11}$$

This measure represents the coupling strength of the interaction from X_j to X_i, as compared to the sum of those from all variables to X_i. Similar to PDC the normalization is such that

$$0 \leq |\mathrm{DTF}_{j\to i}(\omega)|^2 \leq 1\,, \tag{12}$$
$$\textstyle\sum_{j=1}^{n} |\mathrm{DTF}_{j\to i}(\omega)|^2 = 1\,. \tag{13}$$

Since it makes use of the inverse of the regression matrix, the DTF measure is a linear combination of direct and indirect couplings. The 'direct DTF' (dDTF) variant is an alternative that emphasizes direct connections (Korzeniewska et al. 2003). Note that, unlike for GGC, there is no corresponding time-domain connectivity map for PDC or DTF.

Limitations and extensions

Stationarity. These measures assume that the time-series are covariance stationary. Non-stationary data should be divided into windows that by themselves are approximately stationary (Bressler and Seth, 2011).
Dependence on variables. The measures reflect directed statistical dependencies amongst the particular set of variables chosen, and should not be interpreted as directly reflecting physical causal chains.
Linearity. The linear formulation of these measures can obviously only give information about linear features of signals. Extension to non-linear cases can however present further problems. Marinazzo et al. (2011) review some non-linear approaches to GGC. Note that for (stationary) Gaussian variables, interactions are necessarily linear, and time-domain GGC is equivalent to transfer entropy, enabling an interpretation of GGC in terms of Shannon information flow (Barnett et al. 2009; Barrett et al. 2010).

Neurobiological application

The above limitations should always be kept in mind when applying these measures to neurobiological data. A general principle is that shorter data segments are more stationary than longer data segments, but yield poorer parameter estimates. Thus a trade-off must be made in choosing the segment length; the most appropriate choice will depend on whether the data are 'event-related'

or 'steady-state', and on the estimated model order (Barrett et al. 2012; Cohen 2014). Typically segment lengths range from 500-2000ms (Cohen 2014). These measures are applicable to continuous variables, and therefore to M/EEG or local field potential data, as opposed to spiking data. The sampling rate should be sufficiently high for the Nyquist frequency to be substantially greater than the highest frequency of interest, but not so high that the model order becomes too large; 200-250Hz is typical (Cohen 2014). fMRI typically has too slow a sampling rate to make use of these measures, although novel ultra-rapid acquisition sequences potentially enable access to low frequency interactions (Seth et al. 2013). For M/EEG, pre-processing should include some form of spatial filtering or source localization in order to minimize volume conduction artefacts (Bressler and Seth, 2011; Cohen 2014). Bandpass temporal filtering should be avoided as this smears the data in time and can increase model orders, thus leading to poorer estimates of the measures (although notch filtering to remove line-noise will not contaminate frequencies that are far away from the stop-band; Barnett and Seth 2011).

Summary and comparison of measures

$\mathrm{GGC}_{j\to i}$ is a measure of the directed functional effect that X_j has on X_i, and compares 'intrinsic' and 'causal' contributions to the power in a bivariate autoregressive model (conditional versions of GGC are used if one wants to separate out direct and indirect causal effects). $\mathrm{PDC}_{j\to i}$ captures the relative strength of the direct pairwise coupling from X_j to X_i, compared to the total coupling from X_j to other variables in the frequency domain MVAR model for the set of variables. $\mathrm{DTF}_{j\to i}$ compares the coupling from X_j to X_i with the total coupling from all variables to X_i in the transfer matrix for the set of variables; as such it is a measure of the combined strength of direct and indirect couplings from X_j to X_i. These measures, of which GGC is the most common and perhaps best founded, hold great promise in moving beyond functional localization towards the important problem of dissecting the functional circuits underlying cognition, perception, and behaviour.

References

Baccalá LA, Sameshima K (2001) Partial directed coherence: A new concept in neural structure determination. Biol Cybern 84:463-474

Barnett L, Barrett AB, Seth, AK (2009) Granger causality and transfer entropy are equivalent for Gaussian variables. Phys Rev Lett 103:238701

Barnett LC, Seth, AK (2011) Behaviour of Granger causality under filtering: Theoretical invariance and practical application. J Neurosci Meth 201:404-419

Barrett AB, Barnett L, Seth AK (2010) Multivariate Granger causality and generalized variance. Phys Rev E 81:041907

Barrett AB, Murphy M, Bruno M-A, Noirhomme Q, Boly M, et al. (2012) Granger causality analysis of steady-state electroencephalographic signals during propofol-induced anaesthesia. PLoS ONE 7(1): e29072.

Bressler SL, Seth AK (2011) Wiener-Granger causality: A well established methodology. Neuroimage 58(2):323-329

Cohen MX (2014) Analyzing neural time series data. MIT Press, Cambridge MA

Dhamala M, Rangarajan G, Ding M (2008) Analyzing information flow in brain networks with non-

parametric Granger causality. Neuroimage 41(2):354-362
Ding M, Chen Y, Bressler S (2006) Granger causality: Basic theory and application to neuroscience. In: Schelter S, Winterhalder M, Timmer J (eds) Handbook of time series analysis. Wiley, Wienheim, pp 438-460
Friston K, Moran R, Seth AK (2013) Analyzing connectivity with Granger causality and dynamic causal modelling. Curr Opin Neurobiol 23:1-7
Geweke J (1982) Measurement of linear dependence and feedback between multiple time series. J Am Statist Assoc 77:304-313
Geweke J (1984) Measures of conditional linear dependence and feedback between time series. J Am Statist Assoc 79:907-915
Granger C (1969) Investigating causal relations by econometric models and cross-spectral methods. Econometrica 37:424-438
Kaminski M, Blinowska, K (1991) A new method of the description of the information flow in the brain structures. Biol Cybern 65(3):203-210
Lütkepohl H (2007) New introduction to multiple time series analysis. Springer, Berlin Heidelberg
Korzeniewska A, Manczak M, Kaminski M, Blinowska K, Kasicki S (2003) Determination of information flow direction among brain structures by a modified directed transfer function (dDTF) method. J Neurosci Meth 125(1-2):195-207
Marinazzo D, Liao W, Chen H, Stramaglia S (2011) Nonlinear connectivity by Granger causality 58(2):330-338
Percival D, Walden A (2000) Wavelet methods for time series analysis. Cambridge University Press, Cambridge
Seth AK, Chorley P, Barnett LC (2013) Granger causality analysis of fMRI BOLD signals is invariant to hemodynamic convolution but not downsampling. Neuroimage 65:540-555.