Spectral Approach to Modeling Dependence in Multivariate Time Series

Hernando Ombao
Statistics Program, King Abdullah University of Science and Technology (KAUST), Saudi Arabia
E-mail: hernando.ombao@kaust.edu.sa

Abstract. Consider a multivariate time series such as prices of stocks from various sectors, amount of rainfall in many geographical locations, and brain signals from many different locations on the scalp. The goal of this paper is to present the spectral approach to modeling dependence between components of the multivariate time series. There are many measures of dependence - the most popular being cross-correlation or partial cross-correlation. This measure is easy to compute and easy to understand but it coarse in a sense that it is not able to identify the underlying frequencies that are responsible for driving the dependence. In the stock price example, two stocks may be highly correlated but it would be helpful to see if this correlation is driven by the daily fluctuations or by millisecond-level fluctuations. In the neuroscience example, when two brain regions exhibit a high level of correlation, it will be important if this synchronicity is due to low-frequency oscillations or high-frequency oscillations. Here we present an overview of the underlying principles through specific spectral models which decompose the signals into oscillations of various frequencies and then model lead-lag dependence via these oscillations.

1. Introduction
Consider a multivariate time series \( \mathbf{X}_t = [X_1(t), \ldots, X_D(t)]' \) which are time series recordings from \( D \) locations in space (i.e., \( D \) electroencephalogram (EEG) channels on the scalp or \( D \) sectors as in the stock price example). The primary question that we will address in this paper is to assess the nature in which activity in channel \( d_1 \) is associated with activity in channel \( d_2 \).

In more general settings, one asks the additional question of whether or not dependence between channels \( d_1 \) and \( d_2 \) are driven by the other channels. The goal is to present an approach to modeling dependence between components of this \( D \)-dimensional network through the oscillatory components of each univariate time series \( X_d(t) \).

There are many approaches to modeling dependence between components of a multivariate time series. The most common is pairwise cross-correlation which measures...
the strength of linear dependence between two components. Another common measure is partial cross-correlation which gives a measure of strength of the direct linear link between two components (after having removed the linear effect of all the other components in the network). Here, we illustrate how one can dig deeper than the classical correlation (or covariance) in order to understand more complex dependence structure between components. For ease in exposition, we shall develop a model only for $D = 2$ keeping in mind that these ideas are generalizable for higher dimensions.

We now motivate our approach of decomposing observed time series in terms of various oscillations. Under weak stationarity, the Cramér representation of a bivariate time series $X(t) = [X_1(t), X_2(t)]'$ is given by

$$
\begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix} = \int_{-0.5}^{+0.5} \exp(i2\pi\omega t) \begin{pmatrix} dZ_1(\omega) \\ dZ_2(\omega) \end{pmatrix}
$$

which expresses the $d$-th time series $X_d(t)$ (for $d = 1, 2$) as a sum of infinitely many sinusoidal oscillations $\exp(i2\pi\omega t)dZ_\omega(\omega)$ where the frequency $\omega$ lies in $(-0.5, 0.5)$. In this representation, the basis consists of the Fourier waveforms $\exp(i2\pi\omega t)$ with random zero-mean coefficients $[dZ_1(\omega), dZ_2(\omega)]$. Under weak stationarity, the random coefficients are uncorrelated across different frequencies, i.e., $\text{Corr}(dZ_{d1}(\omega), dZ_{d2}(\lambda)) = 0$ for $\omega \neq \lambda$.

One most common measure of dependence, called coherency, gives an indication of synchrony between $X_1(t)$ and $X_2(t)$ for each frequency oscillation $\omega$, i.e.,

$$
\rho_{12}(\omega) = \text{Corr}(\exp(i2\pi\omega t)dZ_1(\omega), \exp(i2\pi\omega t)dZ_2(\omega)) = \text{Corr}(dZ_1(\omega), dZ_2(\omega)).
$$

---

**Figure 1.** Modeling dependence in brain activity through the EEG channels.
Here, coherency is complex-valued and thus its phase captures the lead-lag relationship between $X_1(t)$ and $X_2(t)$. But its squared-modulus, called coherence, lies in the unit interval, i.e.,

$$\eta_{12}(\omega) := |\rho_{12}(\omega)|^2 \in [0, 1].$$

The above motivates an intuitive method in [1] for estimating the true value of coherence via linear filtering which we briefly describe below. Define \(\{\psi_j^\Omega\}\) be absolutely summable filter coefficients with frequency response function with support mainly in the interval \(\Omega \in (0, 0.5)\). In EEG analyses, we measure dependence over various disjoint intervals that correspond the standard bands which are delta \((0 - 4)\) Hertz, theta \((4 - 8)\) Hertz, alpha \((8 - 12)\) Hertz, beta \((12 - 30)\) Hertz and gamma \((30 - 50)\) Hertz.

Next, for observed time series \([X_1(t), X_2(t)]', t = 1, \ldots, T\), denote the \(\Omega\)-filtered time series to be, respectively,

$$X_{1,\Omega}(t) = \sum_j \psi_j^\Omega X_1(t - j) \quad \text{and} \quad X_{2,\Omega}(t) = \sum_j \psi_j^\Omega X_2(t - j)$$

so that the filtered series \(X_{1,\Omega}(t)\) and \(X_{2,\Omega}(t)\) will each have spectra that is concentrated on the frequency band \(\Omega\). Without loss of generality, assume that the mean of the filtered time series are both zero, i.e.,

$$\frac{1}{T} \sum_{t=1}^T X_{d,\Omega}(t) = 0 \quad \text{for} \quad d = 1, 2.$$

In [1], the estimate of coherence at frequency band \(\Omega\) is obtained by shifting one of the filtered time series to obtain the maximum absolute cross-correlation. In the following sections, we develop the proposed spectral approach and then compare the various spectral-based approaches to modeling dependence in multivariate time series.

2. The spectral decomposition of multivariate time series

In this section, we now define the discretized Cramér representation of a weakly stationary time series. First, segment the frequency axis \((0, 0.5)\) into non-overlapping bands \(\Omega_k, k = 1, \ldots, K\) that completely cover \((0, 0.5)\). That is,

(a.) \(\Omega_k \cap \Omega_{k'} = \emptyset\) for \(k \neq k'\)

(b.) \(\cup_k \Omega_k = (0, 0.5)\).

Recall that the notation for \(\Omega_k\)-filtered version of time series \(X_d(t)\) to be \(X_{d,\Omega_k}(t)\). Then, the discretized Cramér representation of \([X_1(t), X_2(t)]'\) is given by

$$\left( \begin{array}{c} X_1(t) \\ X_2(t) \end{array} \right) \approx \left( \frac{\sum_k X_{1,\Omega_k}(t)}{\sum_k X_{2,\Omega_k}(t)} \right)$$

(1)

Here, each \(X_d(t), d = 1, 2\) is decomposed into different oscillations (at specific frequency bands). As noted, under weak stationarity, the oscillations at different frequency bands are uncorrelated, i.e., for any \(d_1\) and \(d_2\), any \(t\) and \(s\), and \(k \neq k'\),

$$\text{Cov}(X_{d_1,\Omega_k}(t), X_{d_2,\Omega_{k'}}(s)) = 0.$$
Figure 2. Each time series $X_1(t)$ and $X_2(t)$ are decomposed into oscillations at various frequency bands: delta (0-4) Hertz, theta (4-8) Hertz, alpha (8-12) Hertz, beta (12-30) Hertz, gamma (30-50) Hertz.

See Figure 2 where each time series $X_1(t)$ and $X_2(t)$ are decomposed into oscillations at various frequency bands. We now illustrate that covariance, as a measure of dependence, can be decomposed into the covariance between the different frequency components. From Equations 1 and 2,

$$\text{Cov}(X_1(t), X_2(s)) \approx \sum_k \text{Cov}(X_{1,\Omega_k}(t), X_{2,\Omega_k}(s)).$$

**Remarks.** From the decomposition of covariance above, one can identify the frequency band $\Omega_k$ that drives the linear association between $X_1(t)$ and $X_2(t)$.

### 3. Different measures of spectral dependence

As already noted, the most common measures of frequency band-specific synchrony is coherence. In practice, we estimate coherence between $X_1(t)$ and $X_2(t)$ at each of the frequency bands by computing the squared sample cross-correlations of each of the filtered components (see [1]). Other approaches to estimating coherence involves computing the Fourier transform and the Fourier periodogram matrix and then smoothing the auto- and cross-periodograms (see [2], [3], [4], [5]).

For non-stationary time series, Loève [5] developed the harmonizable process representation that allowed oscillations at different frequencies to be correlated. Thus,
Figure 3. Each time series $X_1(t)$ and $X_2(t)$ are decomposed into oscillations at various frequency bands. Dual-frequency between the alpha and gamma bands that is evolving across time [10].

under this process, the dual-frequency (band) coherence is defined as

$$\rho_{12}(\Omega_1, \Omega_2) = |\text{Corr}(X_{1,\Omega_1}(t), X_{2,\Omega_2}(t))|^2.$$  

For further details on these processes, see [6], [7], [8], [9].

Under the harmonizable process framework, dual-frequency coherence captures the overall dependence across the entire observation period. This is a limitation since this dependence could vary across time (see Figure 3). Thus, in [10], the evolutionary dual-frequency coherence concept was developed within the context of the discretized harmonizable process.

All the aforementioned coherence measures do not directly measure lead-lag dependence. Here, we propose a spectral cross-oscillatory model below to examine how the oscillatory activity at the $\Omega_1$ band for the first time series may depend on the oscillatory activity at all bands $\Omega_1, \ldots, \Omega_K$ for the second time series:

$$X_{1,\Omega_1}(t) = \sum_{\ell=1}^{L} \sum_{k=1}^{K} X_{2,\Omega_k}(t - \ell) + \epsilon_{1,\Omega_1}(t)$$  \hspace{1cm} (3)

where $\epsilon_{1,\Omega_1}(t)$ is white noise with mean 0 and variance $\sigma^2$. The optimal lag $L$ can be selected using information criteria such as the Akaike information criterion or the Bayesian Schwartz information criterion.
4. Modeling dependence via latent sources

One way to account for dependence between time series $X_1(t)$ and $X_2(t)$ is to model these as functions of latent processes. Define the latent processes as follows: $Z_\delta(t)$ be a latent process with power on the delta band; $Z_\theta(t)$ as the theta process; $Z_\alpha(t)$ as the alpha process; $Z_{\beta}(t)$ as the beta process and $Z_{\gamma}(t)$ as the gamma process. Each of these processes can be approximated by second-order autoregressive processes. In [11], the observed time series $X_1(t)$ and $X_2(t)$ are considered to be linear mixtures of these latent processes:

$$
\begin{pmatrix}
X_1(t) \\
X_2(t)
\end{pmatrix} = 
\begin{pmatrix}
q_{1\delta} & q_{1\theta} & q_{1\alpha} & q_{1\beta} & q_{1\gamma} \\
q_{2\delta} & q_{2\theta} & q_{2\alpha} & q_{2\beta} & q_{2\gamma}
\end{pmatrix}
\begin{pmatrix}
Z_\delta(t) \\
Z_\theta(t) \\
Z_\alpha(t) \\
Z_{\beta}(t) \\
Z_{\gamma}(t)
\end{pmatrix}.
$$

Thus, $X_1(t)$ and $X_2(t)$ are coherent at, say the alpha band, if both $q_{1\alpha} \neq 0$ and $q_{2\alpha} \neq 0$, i.e., both time series contain the alpha latent process.

In the representation above, the latent processes are assumed to be independent. However, for some time series, it is possible for some latent processes to be dependent. Moreover, the relationship could be non-linear. In [12], general models are being studied where the theta and gamma latent processes are coupled in some non-linear way

$$
Z_{\gamma}(t) = A(Z_\theta(t - \ell))Z_{\gamma}(t - \ell).
$$

Here, the amplitude of the gamma latent process at time $t$ may depend on the phase of the theta process at time $t - \ell$ via the function $A(.)$ and in the amplitude of the past gamma process value.

5. Summary

In this paper, we demonstrate different approaches to modeling dependence through oscillatory processes via the Cramér and Loève representations and also via the mixtures (possibly non-linear) of oscillatory latent processes. Through these representations, we formally define various measures of spectral dependence. One advantage of these measures is that they are easily interpretable and that the estimation can be readily implemented via the fast Fourier transform and linear filtering. These measures are demonstrated to be highly informative in analyzing brain signals.

References


