Chapter 15

Granger Causality and the Spectral Density

Consider a linear p-variate stationary time series Z, mean 0. Let $S(\lambda)$ denote the *spectral density matrix*. If $Z^{(i)}$ denotes the ith component of Z, then the ACVF is defined as:

$$\Gamma_{ij}(h) = \mathrm{Cov}(Z_t^{(i)}, Z_{t+h}^{(j)}).$$

The spectral density matrix is the matrix with entries:

$$S_{ij}(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} e^{i\lambda h} \Gamma_{ij}(h).$$

15.1 Representations for Stationary Processes

The stationary process has a *moving average representation* if it can be written as:

$$Z_t = \sum_{s=0}^{\infty} \Theta_s \epsilon_{t-s} \qquad \Theta_0 = I \qquad \{\epsilon_t\} \sim WN(0, \Sigma)$$

Invertibility is equivalent to: $|\Theta(z)| \neq 0$ for all $z \in \mathbb{C} : |z| \leq 1$ where, for a matrix C, $\|C\|$ denotes the square root of the largest eigenvalue of $C^t C$ and $|C| = \sqrt{\det(C^t C)}$. Furthermore, to ensure the process has a well defined covariance structure, it is necessary that $\sum_{s=1}^{\infty} \|\Theta(s)\|^2 < +\infty$.

Doob (Stochastic Processes, John Wiley, New York 1953 pp499 - 500) proves that the existence of such a moving average representation for a stationary time series is equivalent to the existence of the spectral matrix $S_Z(\lambda)$ of Z for almost all frequencies $\lambda \in [-\pi, \pi]$.

Under these assumptions, the mean-squared-error of the one-step-ahead forecast (forecast of Z_t based on $\{Z_s; s \leq t-1\}$ is:

$$|\Sigma| = \exp\left\{\frac{1}{2\pi} \int_{-\pi}^{\pi} \ln |S(\lambda)| d\lambda\right\} > 0.$$

This spectral representation for the mean-squared prediction error was stated (without proof) for univariate time series earlier. It is due to Rozanov.

We restrict attention to series with an $MA(+\infty)$ representation which can be inverted:

$$Z_t = \sum_{j=1}^{\infty} \Phi_j Z_{t-j} + \epsilon_t \tag{15.1}$$

where $\{\epsilon_t\} \sim WN(0, \Sigma)$. As usual, let

$$\Phi(z) = I - \sum_{j=1}^{\infty} z^j \Phi_j.$$

A sufficient condition for existence of an AR($+\infty$) representation is the existence of a constant $c < +\infty$ such that

$$c^{-1} I \preccurlyeq S(\lambda) \preceq cI$$

where, for two matrices A and B, the symbol $A \preceq B$ means that $B - A$ is non-negative definite. This is a result of Rozanov.

Note: not all stationary time series have an AR($+\infty$) representation; recall the MA(1) example of: $Z_t = \epsilon_t + \epsilon_{t-1}$. This does not have such a representation; we showed that $\inf_\lambda S(\lambda) = 0$ in this example.

Now suppose that Z is partitioned into $\binom{X}{Y}$ where X is k-variate and Y is m variate, $k + m = p$. Let S_Z denote S (the subscript indicates the multivariate time series for which this is the spectral density matrix). Use the following partition of $S_Z(\lambda)$:

$$S_Z(\lambda) = \begin{pmatrix} S_X(\lambda) & S_{XY}(\lambda) \\ S_{YX}(\lambda) & S_Y(\lambda) \end{pmatrix}.$$

Both X and Y possess autoregressive representations:

$$\begin{cases} X_t = \sum_{s=1}^{\infty} E_{1s} X_{t-s} + \eta_{1t} & \{\eta_{1t}\} \sim WN(0, C_X) \\ Y_t = \sum_{s=1}^{\infty} G_{1s} Y_{t-s} + \xi_{1t} & \{\xi_{1t}\} \sim WN(0, C_Y). \end{cases}$$

These arise from predicting X only using its own past, respectively Y, only using its own past. The disturbance η_{1t} is the one-step-ahead error when X_t is forecast from its own past alone, similarly ξ_{1t} is the one-step-ahead error when Y_t is forecast from its own past alone. These disturbances are each serially uncorrelated, but may be correlated with each other contemporaneously and at various leads and lags.

These equations denote the linear projections of X_t respectively Y_t on their own pasts.

The equation for Z may be partitioned:

$$\begin{cases} X_t = \sum_{s=1}^{\infty} \Phi_{XX;s} X_{t-s} + \sum_{s=1}^{\infty} \Phi_{XY;s} Y_s + \epsilon_{X;t} \\ Y_t = \sum_{s=1}^{\infty} \Phi_{YY;s} Y_{t-s} + \sum_{s=1}^{\infty} \Phi_{YX;s} X_{t-s} + \epsilon_{Y;t}. \end{cases} \quad (15.2)$$

Since $\epsilon_t = \binom{\epsilon_{Xt}}{\epsilon_{Yt}}$ and $\{\epsilon_t\} \sim WN(0, \Sigma)$, it is clear that the disturbance vectors for this model can only be correlated with each other contemporaneously.

Now consider $\Sigma = \begin{pmatrix} \Sigma_{XX} & \Sigma_{XY} \\ \Sigma_{YX} & \Sigma_{YY} \end{pmatrix}$.

15.2 Useful Representations

Let us pre-multiply the system for $\binom{X}{Y}$ by the matrix:

$$\begin{pmatrix} I_k & -\Sigma_{XY}\Sigma_{YY}^{-1} \\ -\Sigma_{YX}\Sigma_{XX}^{-1} & I_m \end{pmatrix}.$$

This gives a system of equations

$$\begin{cases} X_t = \sum_{s \geq 1} E_{3s} X_{t-s} + \sum_{s=0}^{\infty} F_{3s} Y_{t-s} + e_{Xt} \\ Y_t = \sum_{s=1}^{\infty} G_{3s} Y_{t-s} + \sum_{s=0}^{\infty} H_{3s} X_{t-s} + e_{Yt} \end{cases} \quad (15.3)$$

Note that the transformation, for X_t introduces contemporaneous Y_t and vice versa. Here

$$\binom{e_{Xt}}{e_{Yt}} = \begin{pmatrix} I_k & -\Sigma_{XY}\Sigma_{YY}^{-1} \\ -\Sigma_{YX}\Sigma_{XX}^{-1} & I_m \end{pmatrix} \binom{\epsilon_{Xt}}{\epsilon_{Yt}}.$$

While e_{Xt} and e_{Yt} are correlated with each other, the important point is that (a) e_{Xt} is uncorrelated with ϵ_{Yt} and (b) e_{Yt} is uncorrelated with ϵ_{Xt}. This is a straightforward computation. It follows that e_{Yt} is uncorrelated with Y_t as well as with $\{X_s : s \leq t-1\}$ and $\{Y_s : s \leq t-1\}$.

Now let

$$\widehat{D}(\lambda) = S_{XY}(\lambda) S_Y(\lambda)^{-1}$$

The terms are well defined, by the invertibility condition for the time series. Let D denote the inverse Fourier transform

$$D(s) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \widehat{D}(\lambda) e^{-i\lambda s} d\lambda \qquad s \in \mathbb{Z}.$$

Let

$$W_t := X_t - \sum_{s=-\infty}^{\infty} D(s) Y_{t-s}$$

Theorem 15.1. *The process W, thus defined, is uncorrelated with all $\{Y_s : s \in \mathbb{Z}\}$. From this, it follows that*

$$X_t = \sum_{s=-\infty}^{\infty} D_s Y_{t-s} + W_t$$

is the linear projection of X_t onto $\{Y_s; s \leq t\}$.

Proof Consider the spectral representation of Z: $Z_t = \int_{-\pi}^{\pi} e^{-it\nu} dL(\nu)$ where L is the p-variate orthogonal increment process. Let $L = \binom{L_X}{L_Y}$, the k-variate and m-variate processes corresponding to X respectively Y. Then

$$W_t = \int e^{it\nu} dL_X(\nu) - \sum_s D(s) \int e^{i(t-s)\nu} dL_Y(\nu)$$

so that:

$$\mathbb{E}[W_t \overline{Y}_r]^t] = \int S_{XY}(\nu) e^{-(t-r)\nu} d\nu - \sum_{s=-\infty}^{\infty} D(s) \int S_Y(\nu) e^{-i(t-r+s)\nu} d\nu$$

Using the fact that a convolution of Fourier transforms is the Fourier transform of the product gives that $\mathbb{E}[W_t \overline{Y}_r^t] = 0$ for all r.

The final formula for the projection is a direct consequence of this. □

Similarly, it follows that the spectral density matrix for W is given by:

$$S_W(\lambda) = S_X(\lambda) - S_{XY}(\lambda) S_Y(\lambda)^{-1} S_{YX}(\lambda)$$

The process W has an autoregressive representation

$$W_t = \sum_{s=1}^{\infty} \Phi_{Ws} W_{t-s} + \epsilon_{Wt}$$

and consequently

$$X_t = \sum_{s=1}^{\infty} \Phi_{Ws} X_{t-s} - \sum_{s=0}^{\infty} \Phi_{Ws} \sum_{r=-\infty}^{\infty} D_s Y_{t-s-r} + \epsilon_{Wt}$$

where $\Phi_{W0} = -I$. Grouping the terms gives:

$$X_t = \sum_{s=1}^{\infty} \Phi_{Ws} X_{t-s} + \sum_{s=-\infty}^{\infty} \left(\sum_{r=0}^{\infty} \Phi_{Wr} D(s-r) \right) Y_{t-s} + \epsilon_{Wt}.$$

Since ϵ_{Wt} is a linear function of Y and $\{W_s : s \leq t-1\}$, it is uncorrelated with Y. Since $\{X_s : s \leq t-1\}$ is a linear function of Y and $\{W_s : s \leq t-1\}$, ϵ_{Wt} is uncorrelated with $\{X_s : s \leq t-1\}$. Hence this equation provides the linear projection of X_t on $\{X_s : s \leq t-1\}$ and all Y.

Similarly, we can obtain the projection of Y_t on $\{Y_s : s \leq t\}$ and all X.

15.3 Linear Dependence and Feedback

The measure of *linear feedback from Y to X* is defined as:

$$F_{Y\to X} = \ln \frac{|C_X|}{|\Sigma_{XX}|}.$$

Similarly, the linear feedback from X to Y is:

$$F_{X\to Y} = \ln \frac{|C_Y|}{|\Sigma_{XX}|}.$$

The measure of *instantaeous linear feedback* is:

$$F_{X.Y} = \log \frac{|\Sigma_{XX}||\Sigma_{YY}|}{|\Sigma|}.$$

It is non-zero if and only if the partial correlation between X_t and Y_t conditioned on the entire past history of both processes is zero. Finally, the *measure of linear dependence* is:

$$\widetilde{F}_{X,Y} = \ln \frac{|C_X|`|C_Y|}{|\Sigma|}.$$

Note:

$$\widetilde{F}_{X,Y} = F_{Y\to X} + F_{X\to Y} + F_{X.Y}.$$

We now want to describe the linear feedback in Fourier space and we seek non-negative functions $f_{X\to Y}(\lambda)$ and $f_{Y\to X}(\lambda)$ which represent the transfer in Fourier space.

We use (15.2) and (15.3) as the basis for the transfer function. This may be expressed as:

$$\begin{pmatrix} \Phi_{XX}(B) & \Phi_{XY}(B) \\ G_3(B) & H_3(B) \end{pmatrix} \begin{pmatrix} X_t \\ Y_t \end{pmatrix} = \begin{pmatrix} \epsilon_{Xt} \\ e_{Yt} \end{pmatrix}.$$

The existence of joint autoregressive representation ensures that this can be inverted to express:

$$\begin{pmatrix} X_t \\ Y_t \end{pmatrix} = \begin{pmatrix} A_{11}(B) & A_{12}(B) \\ A_{21}(B) & A_{22}(B) \end{pmatrix} \begin{pmatrix} \epsilon_{Xt} \\ e_{Yt} \end{pmatrix}.$$

We use:

$$X_t = A_{11}(B)\epsilon_{Xt} + A_{12}(B)e_{Yt}$$

Let T_Y denote the correlation matrix of e_{Yt}, then the spectral density of X may be written:

$$S_X(\lambda) = \widehat{A}_{11}(\lambda)\Sigma_X \widehat{A}_{11}^t + \widehat{A}_{12}(\lambda)T_Y \widehat{A}_{12}^t(\lambda)$$

where the hat denotes a Fourier transform.

The measure of linear feedback from Y to X in Fourier space is therefore defined as:

$$f_{Y\to X}(\lambda) = \ln \frac{|S_X(\lambda)|}{|\widehat{A}_{12}(\lambda)\Sigma_X \widehat{A}_{12}^t(\lambda)|}.$$

This is the fraction of the spectral density of X which is due to the disturbance $\{e_{Yt} : t \in \mathbb{Z}\}$.