Chapter 14

Spectral Analysis (II)

14.1 Linear Filters, Interpolation and detection

Linear Filters Let $\{X_t\}$ be a stationary time series and let $\{Y_t\}$ be defined as:

$$Y_t = h(B)X_t$$

The relation between the spectral properties of $\{X_t\}$, the input and $\{Y_t\}$, the output are summarised in Theorem 14.1.

Consider a filter $h(B)$ and the test functions $e^{i\lambda t}$ as input. Then

$$(h(B)e^{i\lambda \cdot})_t = \sum_{k=-\infty}^{\infty} h_{t-k}e^{i\lambda k} = \sum_{j=-\infty}^{\infty} h_j e^{i\lambda(t-j)} = e^{i\lambda t} \sum_{j=-\infty}^{\infty} h_j e^{-i\lambda j} = e^{i\lambda t} h(e^{-i\lambda}).$$

Theorem 14.1. *Let $\{X_t\}$ be a stationary input in a stable TLF $h(B)$ and let $\{Y_t\}$ be the output, i.e. $Y = h(B)X$. Then*

1. $\mathbb{E}[Y_t] = h(1)\mathbb{E}[X_t]$*;*
2. Y_t *is stationary;*
3. $F_Y(\lambda) = \int_{(-\pi,\lambda]} |h(e^{-i\nu})|^2 \, dF_X(\nu)$*.*

Proof

1. This follows from:

$$\mathbb{E}[Y_t] = \sum_{k=-\infty}^{\infty} h_{t-k}\mathbb{E}[X_k] = \mathbb{E}[X_k] \sum_{k=-\infty}^{\infty} h_{t-k} = \mathbb{E}[X_k]h(1).$$

2. This follows from the stationarity of X.

3. This follows from:

$$\begin{aligned}
\mathbf{C}(Y_{t+h}, Y_t) &= \mathbf{C}\Big(\sum_{j=-\infty}^{\infty} h_j X_{t+h-j}, \sum_{k=-\infty}^{\infty} h_k X_{t-k}\Big) \\
&= \sum_{j,k=-\infty}^{\infty} h_j \overline{h_k} \mathbf{C}(X_{t+h-j}, X_{t-k}) \\
&= \sum_{j,k=-\infty}^{\infty} h_j \overline{h_k} \gamma_X(h-j+k) \\
&= \sum_{j,k=-\infty}^{\infty} h_j \overline{h_k} \int_{(-\pi,\pi]} e^{i\lambda(h-j+k)}\, dF_X(\lambda) \\
&= \int_{(-\pi,\pi]} e^{i\lambda h} \Big(\sum_{j=-\infty}^{\infty} h_j e^{-i\lambda j}\Big)\Big(\overline{\sum_{k=-\infty}^{\infty} h_k e^{-i\lambda k}}\Big)\, dF_X(\lambda) \\
&= \int_{(-\pi,\pi]} e^{i\lambda h} h(e^{-i\lambda})\overline{h(e^{-i\lambda})}\, dF_X(\lambda) = \int_{(-\pi,\pi]} e^{i\lambda h} |h(e^{-i\lambda})|^2\, dF_X(\lambda).
\end{aligned}$$

□

Interpolation Let $\{X_t,\, t \in \mathbb{Z}\}$ be a real stationary time series with mean 0 and spectral density f, where $f(\lambda) \geq A > 0$ for all $\lambda \in [-\pi, \pi]$. Assume that the entire time series has been observed except at the time point $t = 0$. The *best linear interpolator* $\widehat{X}_0$ of X_0 is defined by

$$\widehat{X}_0 = P_{\overline{\text{spa}}\{X_t,\, t\neq 0\}} X_0.$$

Let X_t have spectral representation $X_t = \int_{(-\pi,\pi]} e^{it\lambda}\, dZ(\lambda)$. Set

$$\mathcal{H}_0 = \overline{\text{spa}}\{e^{it\cdot}, t \neq 0\} \subset L^2(F).$$

Then

$$\widehat{X}_0 = \int_{(-\pi,\pi]} g(\lambda)\, dZ(\lambda),$$

where $g(\cdot) = P_{\mathcal{H}_0} 1$. By the projection theorem, it follows that $g \in \mathcal{H}_0$ is the unique solution of

$$\mathbb{E}[(X_0 - \widehat{X}_0)\overline{X_t}] = \int_{-\pi}^{\pi} (1 - g(\lambda)) e^{-it\lambda} f(\lambda)\, d\lambda = 0 \quad \text{for } t \neq 0.$$

Any solution of the projection equations must satisfy:

$$(1 - g(\lambda)) f(\lambda) = k \quad \text{or} \quad g(\lambda) = 1 - \frac{k}{f(\lambda)}.$$

It is enough to see that g above *is* a solution. The problem is to determine k so that $g \in \mathcal{H}_0$. This requires:

$$0 = \int_{-\pi}^{\pi} g(\nu)\, d\nu = \int_{-\pi}^{\pi} 1 - \frac{k}{f(\nu)}\, d\nu = 2\pi - \int_{-\pi}^{\pi} \frac{k}{f(\nu)}\, d\nu,$$

from which:

$$k = \frac{2\pi}{\int_{-\pi}^{\pi} \frac{d\nu}{f(\nu)}}$$

Thus

$$\widehat{X}_0 = \int_{(-\pi,\pi]} \left(1 - \frac{2\pi}{f(\lambda) \int_{-\pi}^{\pi} \frac{d\nu}{f(\nu)}}\right) dZ(\lambda).$$

Now consider the *mean square interpolation error* $\mathbb{E}\left[(\widehat{X}_0 - X_0)^2\right]$.

$$\begin{aligned} \mathbb{E}\left[(\widehat{X}_0 - X_0)^2\right] &= \int_{-\pi}^{\pi} |1 - g(\lambda)|^2 f(\lambda)\, d\lambda = \int_{-\pi}^{\pi} \frac{|k|^2}{f(\lambda)^2} f(\lambda)\, d\lambda \\ &= k^2 \cdot \frac{2\pi}{k} = 2\pi k = \frac{4\pi^2}{\int_{-\pi}^{\pi} \frac{d\lambda}{f(\lambda)}}. \end{aligned}$$

Example 14.1 (AR(1) process)**.**

For the AR(1) process, it is straightforward to compute that:

$$f(\lambda) = \frac{\sigma^2}{2\pi} \left| \frac{1}{1 - \phi_1 e^{-i\lambda}} \right|^2 = \frac{\sigma^2}{2\pi} \frac{1}{1 - 2\phi_1 \cos\lambda + \phi_1^2}.$$

Since

$$\int_{-\pi}^{\pi} \frac{d\lambda}{f(\lambda)} = \frac{2\pi}{\sigma^2} \int_{-\pi}^{\pi} (1 - 2\phi_1 \cos(\lambda) + \phi_1^2)\, d\lambda = \frac{4\pi^2}{\sigma^2}(1 + \phi_1^2),$$

it follows that:

$$\begin{aligned} \widehat{X}_0 &= \int_{(-\pi,\pi]} \left(1 - \frac{1}{2\pi(1+\phi_1^2) f(\lambda)}\right) dZ(\lambda) = \int_{(-\pi,\pi]} \left(1 - \frac{1 - \phi_1(e^{-i\lambda} + e^{i\lambda})}{1 + \phi_1^2}\right) dZ(\lambda) \\ &= \int_{(-\pi,\pi]} (\phi_1 e^{-i\lambda} + \phi_1 e^{i\lambda}) dZ(\lambda) = \phi_1 X_{-1} + \phi_1 X_1 \end{aligned}$$

and

$$\mathbb{E}[(\widehat{X}_0 - X_0)^2] = \frac{\sigma^2}{1 + \phi_1^2}.$$

□

Detection Suppose that the stationary time series $\{X_t,\, t \in \mathbb{Z}\}$ is a disturbed signal. That is, it is the sum of a signal $\{S_t,\, t \in \mathbb{Z}\}$ and a noise $\{N_t,\, t \in \mathbb{Z}\}$, where the signal and the noise are independent stationary time series with means 0 and spectral densities f_S and f_N respectively. (Note that the noise is not assumed to be white noise.) Assume that the entire time series $X_t = S_t + N_t$ has been observed. The *best linear detector* $\widehat{S}_0$ of S_0 is defined by

$$\widehat{S}_0 = P_{\overline{\text{spa}}\{X_t,\, t\in\mathbb{Z}\}} S_0,$$

where $\overline{\text{spa}}\{X_t,\, t \in \mathbb{Z}\}$ is a Hilbert sub-space of the Hilbert space $\overline{\text{spa}}\{S_t,\; N_t,\, t \in \mathbb{Z}\}$.

It follows from the projection theorem that $\widehat{S}_0$ is the unique solution of

$$\mathbb{E}[(S_0 - \widehat{S}_0)\overline{X_t}] = 0 \quad \text{for all } t.$$

Let S_t and N_t have spectral representations

$$S_t = \int_{(-\pi,\pi]} e^{it\lambda}\, dZ_S(\lambda) \quad \text{and} \quad N_t = \int_{(-\pi,\pi]} e^{it\lambda}\, dZ_N(\lambda)$$

respectively. Then X_t has spectral representation

$$X_t = \int_{(-\pi,\pi]} e^{it\lambda}\,(dZ_S(\lambda) + dZ_N(\lambda)),$$

where Z_S and Z_N are independent. Thus

$$\widehat{S}_0 = \int_{(-\pi,\pi]} g(\lambda)\,(dZ_S(\lambda) + dZ_N(\lambda)),$$

for some function $g \in L^2(F_S + F_N)$. Now,

$$\begin{aligned}
0 &= \mathbb{E}[(S_0 - \widehat{S}_0)\overline{X_t}] \\
&= \mathbb{E}\left[\left(\int_{(-\pi,\pi]} dZ_S(\lambda) - \int_{(-\pi,\pi]} g(\lambda)\,(dZ_S(\lambda) + dZ_N(\lambda))\right)\int_{(-\pi,\pi]} e^{-it\lambda}\,\overline{(dZ_S(\lambda) + dZ_N(\lambda))}\right] \\
&= \int_{(-\pi,\pi]} e^{-it\lambda} f_S(\lambda)\,d\lambda - \int_{(-\pi,\pi]} e^{-it\lambda} g(\lambda)(f_S(\lambda) + f_N(\lambda))\,d\lambda \\
&= \int_{(-\pi,\pi]} e^{-it\lambda}\,(f_S(\lambda) - g(\lambda)(f_S(\lambda) + f_N(\lambda)))\,d\lambda.
\end{aligned}$$

It follows that

$$f_S(\lambda) - g(\lambda)(f_S(\lambda) + f_N(\lambda)) = 0 \quad \text{or} \quad g(\lambda) = \frac{f_S(\lambda)}{f_S(\lambda) + f_N(\lambda)}$$

From this we get the best linear detector

$$\widehat{S}_0 = \int_{(-\pi,\pi]} \frac{f_S(\lambda)}{f_S(\lambda) + f_N(\lambda)}\,(dZ_S(\lambda) + dZ_N(\lambda)),$$

and

$$
\begin{aligned}
\mathbb{E}[(S_0 - \widehat{S}_0)^2] &= \mathbb{E}\left[S_0^2\right] - \mathbb{E}\left[\widehat{S}_0^2\right] \\
&= \int_{(-\pi,\pi]} f_S(\lambda)\,d\lambda - \int_{(-\pi,\pi]} |g(\lambda)|^2 \left(f_S(\lambda) + f_N(\lambda)\right)\,d\lambda \\
&= \int_{(-\pi,\pi]} f_S(\lambda)\,d\lambda - \int_{(-\pi,\pi]} \left|\frac{f_S(\lambda)}{f_S(\lambda)+f_N(\lambda)}\right|^2 (f_S(\lambda)+f_N(\lambda))\,d\lambda \\
&= \int_{(-\pi,\pi]} \left(f_S(\lambda) - \frac{f_S^2(\lambda)}{f_S(\lambda)+f_N(\lambda)}\right)\,d\lambda \\
&= \int_{-\pi}^{\pi} \frac{f_S(\lambda)f_N(\lambda)}{f_S(\lambda)+f_N(\lambda)}\,d\lambda.
\end{aligned}
$$

14.2 Estimating the spectral density

Recall that the ACVF $\gamma(\cdot)$ of a stationary time series has the spectral representation

$$\gamma(h) = \int_{(-\pi,\pi]} e^{ih\nu}\,dF(\nu).$$

If, furthermore, $\sum_{-\infty}^{\infty} |\gamma(h)| < \infty$, then the spectral density function $f(\cdot)$ defined by

$$f(\lambda) = \frac{1}{2\pi}\sum_{n=-\infty}^{\infty} e^{-in\lambda}\gamma(n).$$

is well defined, $F(\lambda) = \int_{-\pi}^{\lambda} f(\nu)\,d\nu$ and

$$\gamma(h) = \int_{-\pi}^{\pi} e^{ih\lambda} f(\lambda)\,d\lambda.$$

Attention is now restricted to the situation where $\{X_t\}$ is a stationary time series with mean μ and with absolutely summable covariance; $\sum_{-\infty}^{\infty} |\gamma(h)| < \infty$, so that the spectral density $f(\cdot)$ is well defined. The n variables $X_1, \ldots, X_n$ are observed.

The periodogram The first empirical approximation to $f(\cdot)$ is the so-called *periodogram.* Set

$$\omega_j = \frac{2\pi j}{n}, \qquad -\pi < \omega_j \le \pi$$

and

$$F_n := \{j \in \mathbb{Z},\, -\pi < \omega_j \le \pi\} = \left\{-\left[\frac{n-1}{2}\right], \ldots, \left[\frac{n}{2}\right]\right\},$$

where $[x]$ denotes the integer part of x.

Definition 14.2. *The* periodogram $I_n(\cdot)$ *of* $\{X_1, \ldots, X_n\}$ *is defined by*

$$I_n(\omega_j) = \frac{1}{n}\left|\sum_{t=1}^{n} X_t e^{-it\omega_j}\right|^2, \quad j \in F_n.$$

The values of the periodogram may be expressed in terms of the sample covariance, as indicated in the following proposition.

Proposition 14.3. *The periodogram satisfies:*

$$I_n(\omega_j) = \begin{cases} n|\overline{X}|^2 & \text{if } \omega_j = 0, \\ \sum_{|k|<n} \widehat{\gamma}(k) e^{-ik\omega_j} & \text{if } \omega_j \neq 0, \end{cases}$$

where $\widehat{\gamma}(k) = \frac{1}{n}\sum_{t=1}^{n-|k|}(X_t - \overline{X})(X_{t+|k|} - \overline{X})$ *and* $\overline{X} = \frac{1}{n}\sum_{j=1}^{n} X_j$*.*

Proof The result is clear when $\omega_j = 0$. Consider $\omega_j \neq 0$. Then

$$I_n(\omega_j) = \frac{1}{n}\left|\sum_{t=1}^{n} X_t e^{-it\omega_j}\right|^2 = \frac{1}{n}\sum_{s=1}^{n} X_s e^{-is\omega_j} \sum_{t=1}^{n} X_t e^{it\omega_j}.$$

Since $\sum_{t=1}^{n} e^{-is\omega_j} = 0$, this may be written as:

$$\begin{aligned} I_n(\omega_j) &= \frac{1}{n}\sum_{s=1}^{n}\sum_{t=1}^{n}(X_s - \overline{X})e^{-is\omega_j}(X_t - \overline{X})e^{it\omega_j} \\ &= \sum_{k=-n+1}^{n-1} e^{ik\omega_j}\frac{1}{n}\sum_{s=1\vee(1-k)}^{n\wedge(n-k)}(X_s - \overline{X})(X_{k+s} - \overline{X}) \\ &= \sum_{|k|\leq n-1} e^{ik\omega_j}\widehat{\gamma}(k). \end{aligned}$$

The last equality follows from the definition of $\widehat{\gamma}(k)$. □

The periodogram has now been defined for the *Fourier frequencies* $\frac{2\pi j}{n}$; it is extended to all $\omega \in (-\pi, \pi]$ (or equivalently $(0, 2\pi]$) in the following way.

Definition 14.4 (Extension of the periodogram)**.** *For any* $\omega \in [-\pi, \pi]$ *define*

$$I_n(\omega) = \begin{cases} I_n(\omega_k) & \text{if } \omega_k - \pi/n < \omega \leq \omega_k + \pi/n \text{ and } 0 \leq \omega \leq \pi, \\ I_n(-\omega) & \text{if } \omega \in [-\pi, 0). \end{cases}$$

For $\omega \in [0, \pi]$, let

$$g(n,\omega) = \frac{2\pi j}{n} \qquad j = \min \operatorname{argmin}_k \left\{\left|\omega - \frac{2\pi k}{n}\right|\right\}$$

(the multiple of $\frac{2\pi}{n}$ closest to ω, taking the the smaller one if there are two). For $\omega \in [-\pi, 0)$, $g(n,\omega) = g(n,-\omega)$. With this notation,

$$I_n(\omega) = I_n(g(n,\omega)).$$

Theorem 14.5. *The following convergence results hold:*

$$\mathbb{E}\left[I_n(0)\right] - n\mu^2 \to 2\pi f(0) \quad \text{as } n \to \infty$$

and

$$\mathbb{E}\left[I_n(\omega)\right] \to 2\pi f(\omega) \quad \text{as } n \to \infty \text{ if } \omega \neq 0.$$

If $\mu = 0$ then $I_n(\omega)$ converges uniformly to $2\pi f(\omega)$ on $[-\pi, \pi)$

Proof Omitted. It is straightforward and may be found in Brockwell and Davis. □

For a strictly linear time series $\{X_t\}$ with mean 0; that is

$$X_t = \sum_{j=-\infty}^{\infty} \psi_j Z_{t-j}, \qquad \{Z_t\} \sim \mathrm{IID}(0, \sigma^2)$$

the statistical behaviour of the periodogram has been well analysed in the literature. The following theorem gives some key properties.

Theorem 14.6. *Let $\{X_t\}$ be a strictly linear time series with*

$$\mu = 0, \qquad \sum_{j=-\infty}^{\infty} |\psi_j||j|^{1/2} < \infty \quad \text{and} \quad \mathbb{E}[Z^4] < \infty.$$

Then

$$\mathbf{C}(I_n(\omega_j), I_n(\omega_k)) = \begin{cases} 2(2\pi)^2 f^2(\omega_j) + O(n^{-1/2}) & \text{if } \omega_j = \omega_k = 0 \text{ or } \pi, \\ (2\pi)^2 f^2(\omega_j) + O(n^{-1/2}) & \text{if } 0 < \omega_j = \omega_k < \pi, \\ O(n^{-1}) & \text{if } \omega_j \neq \omega_k. \end{cases}$$

Proof Omitted. This may also be found in Brockwell and Davies. It requires heavy computation, but the steps are all reasonably obvious. □

14.2.1 Smoothing the periodogram

When estimating the periodogram, the same number of parameters are estimated as observations. That is, if there are n observations, then the periodogram is constructed from the estimates of $\gamma(0), \ldots, \gamma(n-1)$. The deficiencies of the periodogram may be seen from a few examples; the periodogram for 'white noise' may not have a discernible pattern, but it will be far from constant, when $f(\lambda) = \frac{\sigma^2}{2\pi}$ for all $\lambda \in (-\pi, \pi]$ for $\mathrm{WN}(0, \sigma^2)$. A first attempt may be to consider

$$\frac{1}{2\pi} \sum_{|k| \leq m} \frac{1}{2m+1} I_n(\omega_{j+k}).$$

More generally, the following class of estimators is considered.

Definition 14.7. *The estimator* $\widehat{f}(\omega) = \widehat{f}(g(n,\omega))$ *with*

$$\widehat{f}(\omega_j) = \frac{1}{2\pi} \sum_{|k| \le m_n} W_n(k) I_n(\omega_{j+k}),$$

where

$$\begin{cases} m_n \to \infty \quad \text{and} \quad m_n/n \to 0 \quad \text{as} \quad n \to \infty, \\ W_n(k) = W_n(-k), \quad W_n(k) \ge 0, \quad \text{for all } k, \\ \sum_{|k| \le m_n} W_n(k) = 1, \\ \sum_{|k| \le m_n} W_n^2(k) \to 0 \quad \text{as} \quad n \to \infty \end{cases}$$

is called a discrete spectral average estimator *of* $f(\omega)$.

If $\omega_{j+k} \notin [-\pi, \pi]$ *the term* $I_n(\omega_{j+k})$ *is evaluated by defining* I_n *to have period* 2π.

Theorem 14.8. *Let* $\{X_t\}$ *be a strictly linear time series with*

$$\mu = 0, \qquad \sum_{j=-\infty}^{\infty} |\psi_j| |j|^{1/2} < \infty \quad \text{and} \quad \mathbb{E}[Z^4] < \infty.$$

Then

$$\lim_{n \to \infty} \mathbb{E}\left[\widehat{f}(\omega)\right] = f(\omega)$$

and

$$\lim_{n \to \infty} \frac{1}{\sum\limits_{|k| \le m_n} W_n^2(k)} \mathbf{C}(\widehat{f}(\omega), \widehat{f}(\lambda)) = \begin{cases} 2f^2(\omega) & \text{if } \omega = \lambda = 0 \text{ or } \pi, \\ f^2(\omega) & \text{if } 0 < \omega = \lambda < \pi, \\ 0 & \text{if } \omega \ne \lambda. \end{cases}$$

Proof Omitted. □

Remark If $\mu \neq 0$, then the term $I_n(0)$ is omitted from the construction and instead:

$$\widehat{f}(0) := \frac{1}{2\pi} \left(W_n(0) I_n(\omega_1) + 2 \sum_{k=1}^{m_n} W_n(k) I_n(\omega_{k+1}) \right).$$

For each appearance of $I_n(0)$ in the formula for $\widehat{f}(\omega_j)$, it is replaced with $\widehat{f}(0)$. □

Example 14.2.

$$W_n(k) = \begin{cases} 1/(2m_n + 1) & \text{if } |k| \le m_n, \\ 0 & \text{if } |k| > m_n, \end{cases}$$

and

$$\mathbf{V}(\widehat{f}(\omega)) \sim \begin{cases} \frac{1}{m_n} f^2(\omega) & \text{if } \omega = 0 \text{ or } \pi, \\ \frac{1}{m_n} \frac{f^2(\omega)}{2} & \text{if } 0 < \omega < \pi. \end{cases}$$

□

The *lag window spectral estimator* is defined as follows:

Definition 14.9. *An estimator* $\widehat{f}_L(\omega)$ *of the form*

$$\widehat{f}_L(\omega) = \frac{1}{2\pi} \sum_{|h| \le r_n} w(h/r_n) \widehat{\gamma}(h) e^{-ih\omega}$$

where

$$\begin{cases} r_n \to \infty \quad \text{and} \quad r_n/n \to 0 \quad \text{as} \quad n \to \infty, \\ w(x) = w(-x), \quad w(0) = 1, \\ |w(x)| \le 1, \quad \text{for all } x, \\ w(x) = 0, \quad \text{for } |x| > 1 \end{cases}$$

is called a lag window spectral estimator *of* $f(\omega)$.

Discrete spectral average estimators and lag window spectral estimator are essentially the same.

The *spectral window* is defined as:

$$W(\omega) = \frac{1}{2\pi} \sum_{|h| \le r_n} w(h/r_n) e^{-ih\omega}.$$

The lag window spectral estimator may be expressed in terms of the following slightly different extension of the periodogram:

$$\tilde{I}_n(\omega) = \sum_{|h| \le n} \widehat{\gamma}(h) e^{-ih\omega}.$$

Note that

$$\widehat{\gamma}(h) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ih\lambda} \tilde{I}_n(\lambda)\, d\lambda.$$

It follows that

$$
\begin{aligned}
\widehat{f}_L(\omega) &= \frac{1}{(2\pi)^2}\sum_{|h|\le r_n} w(h/r_n)\int_{-\pi}^{\pi} e^{-ih(\omega-\lambda)}\tilde{I}_n(\lambda)\,d\lambda \\
&= \frac{1}{(2\pi)^2}\int_{-\pi}^{\pi}\left(\sum_{|h|\le r_n} w(h/r_n)e^{-ih(\omega-\lambda)}\right)\tilde{I}_n(\lambda)\,d\lambda \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} W(\omega-\lambda)\tilde{I}_n(\lambda)\,d\lambda \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} W(\lambda)\tilde{I}_n(\omega+\lambda)\,d\lambda \\
&\simeq \frac{1}{2\pi}\sum_{|j|\le[n/2]} W(\omega_j)\tilde{I}_n(\omega+\omega_j)\,\frac{2\pi}{n} \\
&\simeq \frac{1}{2\pi}\sum_{|j|\le[n/2]} W(\omega_j)I_n(g(n,\omega)+\omega_j)\,\frac{2\pi}{n}.
\end{aligned}
$$

Here $\widehat{f}_L(\omega)$ is approximated by a discrete spectral average estimator with weights

$$W_n(j) = 2\pi W(\omega_j)/n, \quad |j| \le [n/2].$$

It is straightforward to show that

$$\sum_{|j|\le[n/2]} W_n^2(j) \simeq \frac{r_n}{n}\int_{-1}^{1} w^2(x)\,dx.$$

The following theorem holds:

Theorem 14.10. *Let $\{X_t\}$ be a strictly linear time series with $\mu = 0$, $\sum_{j=-\infty}^{\infty} |\psi_j||j|^{1/2} < \infty$ and $\mathbb{E}\left[Z^4\right] < \infty$. Then*

$$\lim_{n\to\infty} \mathbb{E}\left[\widehat{f}_L(\omega)\right] = f(\omega)$$

and

$$
\mathbf{V}(\widehat{f}_L(\omega)) \sim \begin{cases} \frac{r_n}{n} 2f^2(\omega)\int_{-1}^{1} w^2(x)\,dx & \text{if } \omega = 0 \text{ or } \pi \\ \frac{r_n}{n} f^2(\omega)\int_{-1}^{1} w^2(x)\,dx & \text{if } 0 < \omega < \pi. \end{cases}
$$

Proof Omitted. □

Example 14.3 (The Rectangular or Truncated Window)**.**

For this window,

$$
w(x) = \begin{cases} 1 & \text{if } |x| \le 1, \\ 0 & \text{if } |x| > 1, \end{cases}
$$

and

$$\mathbf{V}(\widehat{f}_L(\omega)) \sim \frac{2r_n}{n} f^2(\omega) \quad \text{for } 0 < \omega < \pi.$$

□

Example 14.4 (The Blackman-Tukey Window)**.**

For this window,

$$w(x) = \begin{cases} 1 - 2a + 2a\cos x & \text{if } |x| \leq 1, \\ 0 & \text{if } |x| > 1, \end{cases}$$

and

$$\mathbf{V}(\widehat{f}_L(\omega)) \sim \frac{2r_n}{n}(1 - 4a + 6a^2) f^2(\omega) \quad \text{for } 0 < \omega < \pi.$$

Note that $\widehat{f}_L(\omega) = a\widehat{f}_T(\omega - \pi/r_n) + (1-2a)\widehat{f}_T(\omega) + a\widehat{f}_T(\omega + \pi/r_n)$ where $\widehat{f}_T$ is the truncated estimate. This estimate is easy to compute. Usual choices of a are 0.23 (The Tukey-Hamming estimate) or 0.25 (The Tukey-Hanning estimate). □

14.3 Spectral Properties of Multivariate Time Series

Firstly, assume that

$$\sum_{h=-\infty}^{\infty} |\gamma_{ij}(h)| < \infty, \quad i, j = 1, \ldots, m. \tag{14.1}$$

Definition 14.11 (The cross spectrum)**.** *Let* $\{\underline{X}_t,\ t \in \mathbb{Z}\}$ *be an* m*-variate stationary time series whose ACVF satisfies* (14.1)*. The function*

$$f_{jk}(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} e^{-ih\lambda} \gamma_{jk}(h), \quad -\pi \leq \lambda \leq \pi, \ \ j \neq k,$$

is called the cross spectrum *or* cross spectral density *of* $\{X_{tj}\}$ *and* $\{X_{tk}\}$*. The matrix*

$$f(\lambda) = \begin{pmatrix} f_{11}(\lambda) & \ldots & f_{1m}(\lambda) \\ \vdots & & \\ f_{m1}(\lambda) & \ldots & f_{mm}(\lambda) \end{pmatrix}$$

is called the spectrum *or* spectral density matrix *of* $\{\underline{X}_t\}$.

It follows from direct calculations that

$$\Gamma(h) = \int_{-\pi}^{\pi} e^{ih\lambda} f(\lambda)\, d\lambda.$$

The function $f_{jj}(\lambda)$ is the spectral density of $\{X_{tj}\}$ and therefore non-negative and symmetric about zero. However, since $\gamma_{ij}(\cdot)$, $i \neq j$, is not in general symmetric about zero, the cross spectral density is typically complex-valued. The spectral density matrix $f(\lambda)$ is non-negative definite for all $\lambda \in [-\pi, \pi]$.

Theorem 14.12 (Multivariate Spectral Distribution)**.** $\Gamma(\cdot)$ *is the ACVF of an m-variate stationary time series $\{\underline{X}_t,\, t \in \mathbb{Z}\}$ if and only if*

$$\Gamma(h) = \int_{(-\pi,\pi]} e^{ih\lambda}\, dF(\lambda), \quad h \in \mathbb{Z},$$

where $F(\cdot)$ is an $m\times m$ matrix distribution, i.e. $F_{jk}(-\pi) = 0$, $F_{jk}(\cdot)$ is right-continuous and $F(\mu)-F(\lambda)$ is non-negative definite for all $\lambda \leq \mu$.

Proof Left as an exercise: similar to the univariate case. □

Similarly, a multivariate time series has a spectral representation.

Theorem 14.13 (Multivariate Spectral Representation)**.** *Let $\{\underline{X}_t\}$ be a multivariate stationary time series, with mean $\underline{\mu}$. Then $\{\underline{X}_t - \underline{\mu}\}$ has the representation*

$$\underline{X}_t - \underline{\mu} = \int_{(-\pi,\pi]} e^{it\lambda}\, d\underline{Z}(\lambda)$$

where $\{\underline{Z}(\lambda),\, \lambda \in [-\pi, \pi]\}$ is an m-variate process whose components are complex-valued and satisfy, when integrating against test functions $f, g : [-\pi, \pi] \to \mathbb{R}$, $f, g \in L^2(F)$

$$\mathbb{E}\left[\int_{(-\pi,\pi]} \int_{(-\pi,\pi]} f(\lambda)g(\nu)dZ_j(\lambda)dZ_k(\nu)\right] = \int_{(-\pi,\pi]} f(\lambda)g(\lambda)dF_{jk}(\lambda).$$

Proof Left as an exercise: similar to the univariate case. □

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:

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

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:

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

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

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

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:

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

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 $\{e_{Yt} : t \in \mathbb{Z}\}$.