Spectral Analysis/The Spectral Domain

March 9, 2019

The material in this set of notes is based on S&S Example 2.8 and S&S Chapter 4.

Theory

The general idea is that there's another nice way to write out models for time series data. So far, we have always talked about modeling each element of a time series as a linear function of past values and/or past errors. Instead, we might think of each element of a time series as a noisy realization of several cyclical processes:

$$x_t = \sum_{k=1}^{r} v_k \cos\left(2\pi\omega_k t\right) + u_k \sin\left(2\pi\omega_k t\right), \; v_k, u_k \overset{i.i.d.}{\sim} \mathcal{N}\left(0, \sigma_k^2\right). \tag{1}$$

There is a very useful theorem called the **spectral representation theorem** that states that if x_t is a stationary process, we can find a **spectral representation** for x_t, i.e. we can find r underlying cyclical processes that vary at rates denoted by $\omega_1, \ldots, \omega_r$ and account for different amounts of the variability $\sigma_1^2, \ldots, \sigma_r^2$ such that (6) holds approximately. Note that it is possible r could be very big! This means that all of our stationary **ARMA**(p, q) models can approximately be written as in (6) certain choices of r, $\omega_1, \ldots, \omega_r$, and $\sigma_1^2, \ldots, \sigma_r^2$, although there may not be nice expressions for r, $\omega_1, \ldots, \omega_r$, and $\sigma_1^2, \ldots, \sigma_r^2$ in terms of our **ARMA**(p, q) parameters $\mu_x, \phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q$, and σ_w^2.

A nice thing about the **spectral representation** of the time series is that it allows

straightforward computation of $\gamma_x(h)$ using some trigonometric identities. When $h = 0$, we have:

$$\begin{aligned}
\gamma_x(0) =& \mathbb{E}[x_t]^2 \\
=& \mathbb{E}\left[\left(\sum_{k=1}^{r} v_k \cos(2\pi\omega_k t) + u_k \sin(2\pi\omega_k t)\right)^2\right] \\
=& \mathbb{E}\left[\sum_{k=1}^{r} \left(v_k \cos(2\pi\omega_k t) + u_k \sin(2\pi\omega_k t)\right)^2\right] + \\
& \mathbb{E}\left[\sum_{k=1}^{r}\sum_{l=1,l\neq k}^{r} \left(v_k \cos(2\pi\omega_k t) + u_k \sin(2\pi\omega_k t)\right)\left(v_l \cos(2\pi\omega_l t) + u_l \sin(2\pi\omega_l t)\right)\right] \\
=& \sum_{k=1}^{r} \mathbb{E}\left[v_k^2\right]\cos^2(2\pi\omega_k t) + \mathbb{E}\left[u_k^2\right]\sin^2(2\pi\omega_k t) \qquad \mathbb{E}[v_k u_k] = 0,\ \mathbb{E}[v_k v_l] = \mathbb{E}[u_k u_l] = 0,\ k \neq l \\
=& \sum_{k=1}^{r} \sigma_k^2. \qquad \sin^2(a) + \cos^2(a) = 1
\end{aligned}$$

More generally, for any $h > 0$ we can write

$$\gamma_x(h) = \sum_{k=1}^{r} \sigma_k^2 \cos(2\pi\omega_k h). \tag{2}$$

Independence of $v_1, \ldots, v_k, u_1, \ldots, u_k$ is the key to straightforward computation of $\gamma_x(h)$, and it is also useful interpretationally as it allows us to decompose a time series with possibly correlated elements into a set of independent random variables.

The actual **spectral representation theorem** is quite technical, and I won't expect you to have a deep understanding of it or ask you to prove it but I will show it and try to explain how it leads us to conclude that any stationary process can approximately be represented by (6) for certain choices of $\omega_1, \ldots, \omega_r$ and $\sigma_1^2, \ldots, \sigma_r^2$. First, we need to define the spectral distribution function $F(w)$.

The spectral distribution function $F(w)$ is closely related to the autocovariance function, $\gamma_x(h)$. It is a monotone non-decreasing function defined on the domain $[-0.5, 0.5]$ that

satisfies:

$$\gamma_x(h) = \int_{-0.5}^{0.5} \exp\{2\pi i\omega h\}\, dF(\omega). \tag{3}$$

By virtue of being monotone non-decreasing and satisfying (3), we can conclude that $F(\omega)$ must have various specific properties - we won't go through them all here, but you can find them in Appendix C of S&S. The important thing to understand is that the spectral distribution function $F(\omega)$ contains the same information as the autocovariance function $\gamma_x(h)$, in the way that the characteristic function of a random variable contains the same information as a random variable's probability density function.

The **spectral representation theorem** states that if x_t is a mean-zero stationary process with spectral distribution function $F(\omega)$ that satisfies (3), then there exists a complex-valued stochastic process $z(\omega)$ on the interval $\omega \in [-0.5, 0.5]$ having stationary uncorrelated increments $dz(\omega)$ such that

$$x_t = \int_{-0.5}^{0.5} \exp\{2\pi i t\omega\}\, dz(\omega), \tag{4}$$

where for $-0.5 \leq \omega_1 \leq \omega_2 \leq 0.5$, $\mathbb{V}[z(\omega_2) - z(\omega_1)] = F(\omega_2) - F(\omega_1)$.

Letting $\omega_1, \ldots, \omega_m$ be an increasing sequence of values from -0.5 to 0.5 and noting that if $z(\omega)$ is stationary, then (4) is equivalent to:

$$x_t = \lim_{r\to\infty} \sum_{k=1}^{r} \exp\{2\pi i t\omega_k\}\, y_k, \tag{5}$$

where the increments $y_k = z(\omega_k) - z(\omega_{k-1})$ are independent, stationary random variables.

We can go from (5) to (6) by applying a bunch of trigonometric identities, and by assuming that the infinite sum in (5) can be reasonably well represented by a finite sum. The technical details are very involved and well beyond the scope of this class, and I have not found a resource that gives a very clear big picture explanation.

Going back to

$$x_t = \sum_{k=1}^{r} v_k \cos\left(2\pi\omega_k t\right) + u_k \sin\left(2\pi\omega_k t\right), \; v_k, u_k \overset{i.i.d.}{\sim} \mathcal{N}\left(0, \sigma_k^2\right). \tag{6}$$

The variances $\sigma_1^2, \ldots, \sigma_r^2$ are specific values of the **spectral density function** of x, which is defined as follows. When x_t is stationary and has a summable autocovariance function $\gamma_x(h)$, i.e.

$$\sum_{h=-\infty}^{\infty} |\gamma_x(h)| < \infty,$$

then it can be written as

$$\gamma_x(h) = \int_{-0.5}^{0.5} \exp\{2\pi i \omega h\} f(\omega) \, d\omega,$$

where $h = \ldots, -2, -1, 0, 1, 2, \ldots$. This is the inverse transform of the **spectral density**, which in turn is given by

$$f(\omega) = \sum_{h=-\infty}^{\infty} \gamma(h) \exp\{-2\pi i \omega h\},$$

where $-0.5 \leq \omega \leq 0.5$. If you are familiar with the Fourier transform, you might recognize that the **spectral density** $f(x)$ as the Fourier transform of the autocovariance function $\gamma_x(h)$. If we refer back to our simple spectral representation for x_t given by (6), each σ_k^2 corresponds to the spectral density function evaluated at frequency ω_k, $\sigma_k^2 = f(\omega_k)$.

The spectral density of a stationary process x_t with summable autocovariances behaves a lot like a probability density function. It corresponds to the derivative of the spectral distribution function $F(\omega)$ and is always nonnegative. However, just as the spectral distribution function is unlike a cumulative distribution function in that $F(0.5) = \gamma_x(0)$ instead of 1, the spectral density function does not always integrate to 1 but rather integrates to $\gamma_x(0)$. The spectral density also has two other nice properties:

- It is symmetric about 0, i.e. $f(\omega) = f(-\omega)$.

- It has period one 1, i.e. $f(\omega + 1) = f(\omega)$.

As a result, we can always just examine $f(\omega)$ on the interval $[0, 0.5]$ without any loss of information - once we know the values of $f(\omega)$ for $0 \leq \omega \leq 0.5$, we know the values of $f(\omega)$ for $-0.5 \leq \omega \leq 0$ and $f(\omega_k + \omega_j)$, where $-0.5 \leq \omega_k \leq 0.5$ and ω_j is any integer.

We can think of the values of the function $f(\omega)$ as the unknown parameters that describe the behavior of a time series x_t when we are working in the **spectral domain**, whereas the values of the function $\gamma_x(h)$ are the unknown parameters that describe the behavior of a time series x_t when we are working in the **time domain**.

Estimation

In practice, we will be interested in estimating the spectral density function $f(\omega)$. It is easiest to think about estimating the spectral density as estimating $\sigma_1^2, \ldots, \sigma_r^2$ based on the model (6), but where the number of cyclical components r is unknown. In practice, what we'll do is use as many cyclical components as possible. Note that (6) looks like a regression problem where $\boldsymbol{x}$ is the response, $\cos(2\pi\omega_k \boldsymbol{t})$ and $\sin(2\pi\omega_k \boldsymbol{t})$ are the covariates with $\boldsymbol{t} = (1, \ldots, n)$, v_k and u_k are the regression coefficients with $\sigma_k^2 = \mathbb{E}\left[\left(v_k^2 + u_k^2\right)/2\right]$.