STAT 804

Lecture 19 Notes

Processes with Periodic Components

Some of the series we have looked at have had clear annual cycles, returning to high levels in the same month every year. In our analysis of such processes we have tried to model the mean μ_t as a periodic function $\mu_{t+12} = \mu_t$. Sometimes we have fitted specific periodic functions to μ_t – writing $\mu_t = \alpha \cos(2\pi t/12) + \beta \sin(2\pi t/12)$.

Another process we studied, that of sunspot numbers, also seems to show a clearly periodic component, though now the frequency or period of the oscillation is not so obvious. In this section of the course we investigate the notion of decomposing a general stationary time series into simple periodic components. We will take these components to be cosines and sines. We will be focussing on problems in which the period is not prespecified, that is problems more like the sunspot data than the annual cycle examples.

For a statistician, the simplest description of what we will do is to say that we will examine the correlation between our series and sines and cosines of various periods. We will use these correlations in several ways:

- We will look for periods for which the correlation is very high to seek evidence of an underlying entirely periodic component with that particular period.

- We will study the behaviour of these correlations as a function of period (or frequency) for ARMA processes. The resulting plots can be used to identify specific ARMA models and to test the quality of an ARMA fit.

- We will study the effect of filtering on these correlations. Different filters can remove periodic components (and diminish the correlation between the filtered process and the sine or cosine of corresponding frequency).

- We can use this information about the effect of filtering to discuss the fitting of models to pairs of series in which we assume that a series Y is a filtered version of a series X, both of which we have data from, and try to estimate the filter.

Periodic Functions

A periodic function f on the real line has the property that $f(t+d) = f(t)$ for some d and all t. The smallest possible choice of d is the period of f. The frequency of f in cycles per time unit, is $1/d$. The most famous periodic functions are the trigonometric functions $\sin(\omega t + \phi)$ and its relatives. This function has period $2\pi/\omega$ and frequency $\omega/(2\pi)$ cycles per time unit. Often, for trigonometric functions it is convenient to refer to ω as the frequency; the units now are radians per time point.

The achievement of Fourier was to recognize that essentially any function f with period 1 can be represented as a sum of functions $\sin(2\pi kt)$ or $\cos(2\pi kt)$. The tactic is to suppose that

$$f(t) = a_0 + \sum_{k=1}^{\infty} a_k \cos(2\pi kt) + \sum_{k=1}^{\infty} b_k \sin(2\pi kt)$$

To discover the values of the coefficients we make use of the orthogonality properties:

$$\int_0^1 \cos(2\pi kt)\cos(2\pi jt)\, dt = 1(j = k)/2$$

$$\int_0^1 \sin(2\pi kt)\sin(2\pi jt)\, dt = 1(j = k)/2$$

and

$$\int_0^1 \cos(2\pi kt)\sin(2\pi jt)\, dt = 0\,.$$

Now multiply $f(t)$ by say $\cos(2\pi kt)$ and integrate from 0 to 1. Expanding the integral using the supposed expression of f as a sum gives us

$$\int_0^1 f(t)\cos(2\pi kt)\, dt = a_k \int_0^1 \cos^2(2\pi kt)\, dt = 1/2\,.$$

Similarly $b_k = 2\int_0^1 f(t)\sin(2\pi kt)\, dt$.

Mathematically the fact that we can derive a formula for the coefficients is far from proving that the resulting sum actually represents f; the key missing piece of the proof is that any function whose Fourier coefficients are all 0 is essentially the 0 function.

Correlation between functions

The integrals in the previous section can be thought of as analogous to covariances and variances. For instance a Riemann sum for

$$\int_0^1 \cos(2\pi kt)\cos(2\pi jt)\,dt$$

is

$$\sum_{\ell=0}^{n-1} \cos(2\pi k\ell/n)\cos(2\pi j\ell/n)/n$$

which is an average product. In fact it is possible to show that

$$\sum_{\ell=0}^{n-1} \cos(2\pi k\ell/n)/n = 0$$

So that the average product is just a "sample" covariance. It is also possible to evaluate the average product exactly to see that

$$\sum_{\ell=0}^{n-1} \cos(2\pi k\ell/n)\cos(2\pi j\ell/n)/n = 1(j=k)/2$$

exactly. When $j = k$ this becomes a variance, equal to $1/2$ so that the correlation is just the covariance times 2, which is 0 in any case when $j \neq k$.

Interpreting all the integrals above, then, as covariances we see that all the sines are uncorrelated with each other and with all the cosines and all the cosines are uncorrelated with each other.

Notice particularly that the sine with frequency j and the cosine with frequency j are uncorrelated. This has an important implication for looking for components at frequency j cycles per time unit in a time series: if we want a certain frequency we have to consider both the cosine and the sine at that frequency. An alternative summary of what we need is to consider the trigonometric identity

$$\sin(\omega t + \phi) = \cos(\phi)\sin(\omega t) + \sin(\phi)\cos(\omega t)\,.$$

When we look for a component with frequency ω we will allow ourselves to adjust the number ϕ, called the *phase*, in order to maximize the correlation with our data. This is equivalent to adjusting the coefficients $\cos(\phi)$ and $\sin(\phi)$ to maximize a correlation with the right hand side of the trigonometric identity.

Complex Exponentials

Many of the identities in this subject are more easily derived using complex variables. In particular, the identity

$$e^{i\theta} = \cos(\theta) + i\sin(\theta)$$

where $i^2 = -1$ permits any series in sines and cosines to be rewritten in terms of exponentials. We can then often use tricks involving geometric sums to simplify our algebra.

For instance we can write

$$\cos(2\pi kt) = \frac{\exp(2\pi kti) + \exp(-2\pi kti)}{2}$$

and

$$\sin(2\pi kt) = \frac{\exp(2\pi kti) - \exp(-2\pi kti)}{2i}$$

These permit us to rewrite the expansion (1) in the form

$$f(t) = \sum_{-\infty}^{\infty} c_k \exp(2\pi kti)$$

where $c_k = (a_k - ib_)/2k$ for $k > 0$, $c_0 = a_0$ and $c_k = (a_k + ib_k)/2$ for $k < 0$. In fact

$$c_k = \int_0^1 f(t) \exp(-2\pi kti)\, dt\,.$$

Fourier transforms

For functions which are not periodic we can proceed by a further approximation Suppose f is defined on the real line and fix a large value of T. Define

$$g(t) = f(-T/2 + tT)\,.$$

Then g is defined on $[0, 1]$ and

$$g(t) = \sum \exp(2\pi kti) \int_0^1 g(s) \exp(2\pi ksi)\, ds$$

according to (1) above. Re-express the conclusion in terms of f to get

$$f(u) = \frac{1}{T} \sum \exp(2\pi ki(u + T/2)/T \int_{-T/2}^{T/2} f(v) \exp(-2\pi ki(v + T/2)/T) \, dv$$

which simplifies to

$$f(u) = \frac{1}{T} \sum \exp(2\pi \frac{k}{T} ui) \int_{-T/2}^{T/2} f(v) \exp(-2\pi \frac{k}{T} vi) \, dv$$

You should recognize this sum as a Riemann sum for the integral

$$\int_{-\infty}^{\infty} \exp(2\pi xui) \int_{-T/2}^{T/2} f(v) \exp(-2\pi xvi) \, dv \, du$$

which then converges as $T \to \infty$ to the expression

$$f(u) = \int_{-\infty}^{\infty} \exp(2\pi xui) \int_{-\infty}^{\infty} f(v) \exp(-2\pi xvi) \, dv \, du$$

The function

$$\hat{f}(x) = \int_{-\infty}^{\infty} f(v) \exp(-2\pi xvi) \, dv$$

is called the Fourier transform of f and we have derived a Fourier inversion formula. [WARNING: no proofs here! This integral will exist for, for example, f which are integrable over all the real line.] This inversion formula expresses the function f as a linear combination of sines and cosines, though there are infinitely many frequencies involved.

Transforms of Stochastic Processes

We now seek to apply these ideas with the function f being our stochastic process X. We have several difficulties:

- X is not periodic.

- X is often only a discrete time function and in any case our data is inevitably discrete time.

- Even for continuous time X, the integral in the Fourier transform will typically not converge since X does not even go to 0 at plus or minus infinity.

The discrete nature of X leads us to the study of a discrete approximation to the integral:

$$\sum_{t=0}^{T-1} X_t \exp(i2\pi\omega t)$$

This object has real part

$$\sum_{t=0}^{T-1} X_t \cos(2\pi\omega t)$$

and imaginary part

$$\sum_{t=0}^{T-1} X_t \sin(2\pi\omega t)$$

so that apart from the means not being 0 we are studying the sample co-variance with sines and cosines at frequency ω. We now study the statistical properties of these objects and then try to interpret them.

Suppose that X is a mean 0 stationary time series with autocovariance function C. We define the discrete Fourier transform of X as

$$\hat{X}(\omega) = \frac{1}{\sqrt{T}} \sum_{t=0}^{T-1} X_t \exp(i2\pi\omega t)\,.$$

Our choice to divide by the square root of T is motivated by the recognition that the sum of T terms typically has a standard deviation on the order of $\sqrt{T}$ leading us to expect that $\hat{X}$ will have a standard deviation which has a reasonable limit as $T \to \infty$.

We begin by computing moments of $\hat{X}$. Since $\hat{X}$ is complex valued we have to think about what these moments should be. One way to think about this is to view $\hat{X}$ as a vector with two components, the real and imaginary parts. This would give $\hat{X}$ a mean and a 2 by 2 variance covariance matrix. Also of interest however will be the expected modulus squared of $\hat{X}$, namely

$$\mathrm{E}[|\hat{X}(\omega)|^2] = \mathrm{E}[\hat{X}(\omega)\overline{\hat{X}(\omega)}]$$

where $\bar{z}$ is the complex conjugate of z. (If $z = x + iy$ with x and y real then $\bar{z} = x - iy$.)

Since the Xs have mean 0 we see that

$$\mathrm{E}\hat{X}(\omega) = 0$$

(you should note that the expected value of a complex valued random variable is computed by finding the expected value of the real and imaginary parts). Then

$$\mathrm{E}[|\hat{X}(\omega)|^2] = \frac{1}{T} \sum_{s=0}^{T-1} \sum_{t=0}^{T-1} \exp(i2\pi\omega(s-t))\mathrm{E}(X_s X_t)$$

The expected values are just $C(s-t)$. We can gather together all the terms involving $C(0)$, all those involving $C(1)$ and so on to find

$$\mathrm{E}[|\hat{X}(\omega)|^2] = \frac{1}{T} \left(TC(0) + (T-1)(e^{i2\pi\omega} + e^{-2\pi\omega}) + \cdots \right)$$

which simplifies to

$$C(0) + (1 - 1/T)C(1)(e^{i2\pi\omega} + e^{-2\pi\omega}) + (1 - 2/T)C(2)(e^{i4\pi\omega} + e^{-4\pi\omega}) \cdots$$

As $T \to \infty$ the coefficients of $C(k)$ converges to 1 and we see (using $C(k) = C(-k)$)

$$\mathrm{E}[|\hat{X}(\omega)|^2] = \sum_{-\infty}^{\infty} C(k) \exp(i2\pi\omega k).$$

The right hand side of this expression is defined to be the **spectral density**, or **power spectrum**, of X:

$$f_X(\omega) = \sum_{-\infty}^{\infty} C(k) \exp(i2\pi\omega k).$$

There are a number of ways to look at spectral densities and the discrete Fourier transform:

- The discrete Fourier transform is actually a re-representation of the data because we can recover the data from the transform by an inverse transform:

$$\frac{1}{\sqrt{T}} \sum_{k=0}^{T-1} \exp(-i2\pi kt/T)\hat{X}(k) = \frac{1}{T} \sum_{k=0}^{T-1} \exp(-i2\pi kt/T) \sum_{s=0}^{T-1} \exp(i2\pi ks/T)X_s$$

$$= \frac{1}{T} \sum_{s=0}^{T-1} X_s \sum_{k=0}^{T-1} \exp(i2\pi k(s-t)/T)$$

  For $s = t$ the sum over k is simply T while for $s \neq t$ the sum can be done as a geometric series and seen to be 0. Thus the inside sum just picks out the term $s = t$ giving X_t as the inverse transform.

- Thus the DFT decomposes X into trigonometric functions of various frequencies with $\hat{X}(k/T)$ being the weight on the component at frequency k/T. The spectral density is the limit of the variance of that weight or an approximation to the variance of the component of X at frequency k/T.

- The spectral density is a transform of the autocovariance function of X. In particular

$$\int_0^1 f_X(\omega) \exp(i2\pi\ell\omega)\, d\omega = C_X(\ell).$$

- Since

$$\sum_{k=0}^{T-1} \exp(i2\pi kt/T) = 0$$

for any integer t we see that $\hat{X}(k/T)$ is, apart from a factor of $sqrtT$, a complex number whose real part is the sample covariance between X and $\cos(2\pi kt/T)$ and whose imaginary part is the sample covariance between X and $\sin(2\pi kt/T)$. Consider computing the covariance between X and $a\cos(2\pi kt/T) + b\sin(2\pi kt/T)$ with a and b chosen to maximize the covariance subject to $a^2 + b^2 = 1$. The resulting coefficients are found by multiple regression of X_t on the cosine and sine. Using the fact that $\sum_{k=0}^{T-1} \cos(2\pi kt/T)\sin(2\pi kt/T) = 0$ we can check that the covariance is maximized by taking a and b proportional to the real and imaginary parts of $\hat{X}(k/T)$ respectively and that the covariance with this linear combination is $|\hat{X}(k/T)|^2$. This calculation requires t to be a non-zero integer. In practice we usually apply these techniques to $X - \bar{X}$.

- In view of the last calculation it is useful to examine the behaviour of $\hat{X}$ when X has a purely periodic component, say $X_t = \cos(\theta t + \phi) + W_t$ where W is a stationary mean 0 series. We use $\cos(\theta t + \phi) = (\exp(i(\theta t + \phi)) + \exp(-i(\theta t + \phi)))/2$ to see that

$$\hat{X}(\omega) = \frac{1}{2\sqrt{T}}\left[\frac{e^{i\phi}[1 - e^{i(\omega+\theta)T}]}{1 - e^{i(\omega+\theta)}} + \frac{e^{i\phi}[1 - e^{i(\omega-\theta)T}]}{1 - e^{i(\omega-\theta)}}\right] + \hat{W}(\omega)$$

The leading term converges to 0 unless $\omega = \theta$ or $\omega = -\theta$. Restricting our attention to positive frequencies, when $\omega = \theta$ we get

$$\hat{X}(\omega) = e^{i\phi}\sqrt{T}/2 + \hat{W}(\omega)$$

8

If we move ω around and look at the size of $\hat{X}$ when we hit a frequency at which there is a purely periodic component the modulus of $\hat{X}$ will suddenly contain a component proportional to $\sqrt{T}$. This permits us to look for truly periodic components by looking for spikes or peaks in plots of the modulus squared of the DFT.

- We will see later that if two series Y and X are related by Y being a filtered version of X then the spectral densities have a very simple relation to one another in terms of some property of the filter. We can use this fact to actually estimate the filter itself when this is unknown.