Processes with Periodic Components

Some series looked at have had clear annual cycles, returning to high levels in the same month every year.

Our analysis: tried to model mean μ_t as a periodic function $\mu_{t+12} = \mu_t$.

Sometimes fitted specific periodic functions to μ_t – writing $\mu_t = \alpha \cos(2\pi t/12) + \beta \sin(2\pi t/12)$.

Series of sunspot numbers, also seems to show a clearly periodic component, though now the frequency or period of the oscillation is not so obvious.

Will decompose general stationary time series into simple periodic components.

Take components to be cosines and sines.

Focus on problems in which the period is not prespecified (problems more like the sunspot data than the annual cycle examples).

Use correlation between series and sines and cosines of various periods:

- Look for periods for which correlation is very high: seek evidence of purely periodic component with that period.

- Study correlations as function of period (or frequency) for ARMA processes. Use plots to identify specific ARMA models and to test quality of ARMA fit.

- Study effect of filtering on correlations: remove periodic components (diminish correlation between filtered process and sine or cosine of corresponding frequency).

- Use information about filtering to fit models to pairs of series where series Y is filtered version of series X, given data from X, Y and estimate filter.

Periodic Functions

Defn: A function f on real line is *periodic* if $f(t+d) = f(t)$ for some d and all t.

Smallest possible choice of d is *period* of f.

The *frequency* of f in cycles per time unit, is $1/d$.

Best known periodic functions: trigonometric functions $\sin(\omega t + \phi)$ and relatives.

This function has period $2\pi/\omega$ and frequency $\omega/(2\pi)$ cycles per time unit.

Often convenient to refer to ω as frequency; units now are radians per time point.

Fourier: essentially any function f with period 1 can be represented as a sum of functions $\sin(2\pi k t)$ or $\cos(2\pi k t)$.

Tactic: suppose that

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

Find coefficients using orthogonality:

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

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

and

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

Multiply f by $\cos(2\pi kt)$; integrate from 0 to 1.

Expand integral using assumed expression of f as a sum:

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

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

Mathematically: derivation of formula for coefficients is not proof proving that resulting sum actually represents f.

Missing piece: proof that any function whose Fourier coefficients are all 0 is essentially the 0 function.

Correlation between functions

Integrals in previous section are analogous to covariances and variances. E.g. 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:

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

So: average product is a "sample" covariance.

Can evaluate average product exactly to see

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

exactly.

When $j = k$ this is a variance, equal to 1/2.

So correlation is 2×covariance: 0 when $j \neq k$.

Summary: sines are uncorrelated with each other and with all cosines and all cosines are uncorrelated with each other.

Notice particularly: sine with frequency j and cosine with frequency j are uncorrelated.

Important implication for looking for components at frequency j cycles per time unit in a time series: must consider both cosine and sine at that frequency.

Alternative summary: consider trigonometric identity

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

To look for component with frequency ω adjust ϕ, called *phase*, to maximize correlation with data.

Equivalent to adjusting coefficients $\cos(\phi)$ and $\sin(\phi)$ to maximize correlation with right hand side of trigonometric identity.

Complex Exponentials

Many identities easier with complex variables.

Basic identity: for $i^2 = -1$

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

Use to rewrite series in sines and cosines in terms of exponentials.

Can often use tricks involving geometric sums to simplify algebra.

For instance:

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

and

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

So can rewrite expansion (1) in form

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

where $c_k = (a_k - ib_k)/2$ 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

Nonperiodic functions: make further approximation.

Suppose f defined on real line. Fix 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 conclusion in terms of f:

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

which simplifies to

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

Recognize this as Riemann sum for integral

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

which converges as $T \to \infty$ to

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

Defn: the function

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

is *Fourier transform* of f.

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.

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

Apply these ideas with f being stochastic process X. Several difficulties:

- X is not periodic.
- X often only discrete time function; data always discrete time.
- Even for continuous time X, Fourier transform integral typically doesn't converge:

$$X \not\to 0 \text{ as } t \to \pm\infty.$$

Discrete X leads to study of discrete approximation to 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: apart from means not being 0 studying sample covariance with sines and cosines at frequency ω.

Statistical properties and interpretation?

Suppose X mean 0 stationary time series, autocovariance function C.

Defn: *discrete Fourier transform* of X is

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

Division by $\sqrt{T}$ motivated by recognition that sum of T terms typically has standard deviation on order of $\sqrt{T}$.

So expect SD of $\hat{X}$ will have reasonable limit as $T \to \infty$.

First compute moments of $\hat{X}$.

Moments of complex valued $\hat{X}$?

One way: view $\hat{X}$ as vector with two components, the real and imaginary parts.

Gives $\hat{X}$ a mean and a 2 by 2 variance covariance matrix.

Also of interest: expected modulus squared of $\hat{X}$,:

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

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

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

Note expected value of complex valued random variable is computed by finding expected value of real and imaginary parts.

Then

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

Expected values are $C(s-t)$.

Gather together all terms involving $C(0)$, all those involving $C(1)$, etc.:

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

which simplifies to

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

As $T \to \infty$ coefficient of $C(k)$ converges to 1.

Use $C(k) = C(-k)$ to see

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

Defn: *Spectral density*, or *power spectrum*, of X:

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

Interpretations of spectral density and discrete Fourier transform:

- Discrete Fourier transform is rerepresentation of the data: can recover data from transform by inverse transform:

$$
\begin{aligned}
\frac{1}{\sqrt{T}} & \sum_{k=0}^{T-1} \exp\left(\frac{-i2\pi kt}{T}\right) \hat{X}(k/T) \\
&= \frac{1}{T} \sum_{k=0}^{T-1} \exp\left(\frac{-i2\pi kt}{T}\right) \sum_{s=0}^{T-1} \exp\left(\frac{i2\pi ks}{T}\right) X_s \\
&= \frac{1}{T} \sum_{s=0}^{T-1} X_s \sum_{k=0}^{T-1} \exp\left(\frac{i2\pi k(s-t)}{T}\right)
\end{aligned}
$$

For $s = t$ sum over k is T

For $s \neq t$ sum can be done as geometric series — get 0.

So inside sum just picks out term $s = t$ giving X_t as the inverse transform.

- So DFT decomposes X into trigonometric functions of various frequencies: $\hat{X}(k/T)$ is weight on component at frequency k/T.

- Spectral density is limit of variance of that weight or an approximation to variance of component of X at frequency k/T.

- Spectral density is transform of ACF of X:

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

- Since for any integer $t \neq 0, \pm T, \pm 2T, \ldots$

$$\sum_{k=0}^{T-1} \exp(i 2\pi k t/T) = 0$$

we see $\hat{X}(k/T)$ is, apart from a factor of $\sqrt{T}$, a complex number:

  - real part is sample covariance between X and $\cos(2\pi kt/T)$

  - imaginary part is sample covariance between X and $\sin(2\pi kt/T)$.

- Compute covariance between X and

$$a\cos(2\pi kt/T) + b\sin(2\pi kt/T).$$

Choose a, b to maximize covariance subject to $a^2 + b^2 = 1$.

Resulting coefficients found by multiple regression of X_t on the cosine and sine.

Since

$$\sum_{k=0}^{T-1} \cos(2\pi kt/T)\sin(2\pi kt/T) = 0$$

can check covariance maximized by taking a and b proportional to real and imaginary parts of $\hat{X}(k/T)$ respectively.

Also: covariance with this linear combination is $|\hat{X}(k/T)|^2$.

Calculation requires t to be a non-zero integer.

In practice apply techniques to $X - \bar{X}$.

- Later: if series Y is a filtered version of series X then spectral densities have simple relation to one another in terms of some property of filter.

  Can use this fact to estimate the filter itself when this is unknown.

Properties of Fourier Series

Fourier series for function f truncated to order K, namely

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

where coefficients given by Fourier integrals gives best possible approximation to f as a linear combination of these sines and cosines in the following sense.

Try to choose c_k and d_k to minimize

$$\int_0^1 [f(t) - f_K(t)]^2 \, dt$$

where

$$f_K(t) = c_0 + \sum_{k=1}^{K} c_k \cos(2\pi k t) + \sum_{k=1}^{K} d_k \sin(2\pi k t)$$

Square out integrand, integrate term by term.

Remember sines and cosines are *orthogonal*:

$$\int_0^1 [f(t) - f_K(t)]^2 \, dt$$

$$= \int_0^1 f^2(t)\, dt - 2c_0 \int_0^1 f(t)\, dt$$

$$- 2\sum_{k=1}^{K} \int_0^1 f(t) \left[c_k \cos(2\pi kt) + d_k \sin(2\pi kt)\right] \, dt$$

$$+ \sum_{k=1}^{K} (c_k^2 + d_k^2)$$

Take derivative with respect to c_j to get

$$c_j - 2\int_0^1 f(t) \cos(2\pi jt)\, dt$$

which is 0 when c_j is Fourier coefficient.

Meaning: Fourier series is best possible approximation to a function f by a trigonometric polynomial of this type.

BUT, conclusion depends quite heavily on how we measure quality of approximation.

Example: Fourier approximations to each of 3 functions on [0,1]: the line $y = x$, the quadratic $y = x(1 - x)$ and the square well $y = 1(x < 0.25) + 1(y > 0.75)$.

For each plot: pictures get better as K improves.

But well shaped plot shows effects of Gibb's phenomenon: near the discontinuity in f there is an overshoot which is very narrow and spiky.

Overshoot size does not depend on order of approximation.

Similar discontinuity implicit for $y = x$:

Fourier approximations have period 1.

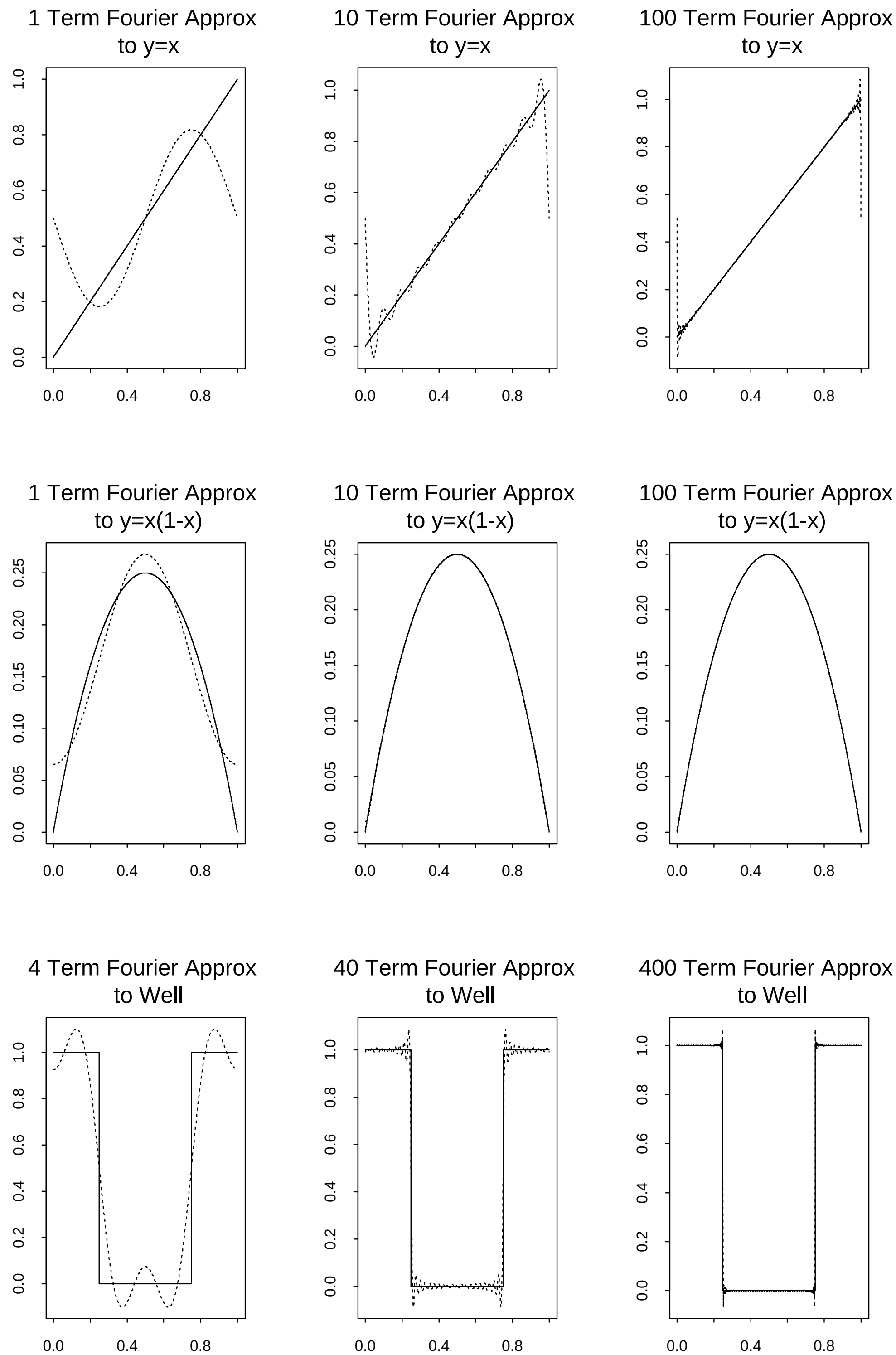
So approximations equal at 0,1 but $y = x$ is not.

Quadratic function has $f(0) = f(1)$; Fourier approximation is much better.

My S-plus plotting code:

```
lin <-  function(k)
{
   x <- seq(0, 1, length = 5000)
   kv <- 1:k
   sv <- sin(2 * pi * outer(x, kv))
   y <-  - sv %*% (1/(pi * kv)) + 0.5
   plot(x, x, xlab = "", ylab = "",
     main = paste(as.character(k),
     "Term Fourier Approximation to y=x"),
     type = "l")
        lines(x, y, lty = 2)
}
```

shows the use of the outer function and the paste function as well as how to avoid loops using matrix arithmetic.


1 Term Fourier Approx
to y=x
10 Term Fourier Approx
to y=x
100 Term Fourier Approx
to y=x
1 Term Fourier Approx
to y=x(1-x)
10 Term Fourier Approx
to y=x(1-x)
100 Term Fourier Approx
to y=x(1-x)
4 Term Fourier Approx
to Well
40 Term Fourier Approx
to Well
400 Term Fourier Approx
to Well
0.0
0.2
0.4
0.6
0.8
1.0
0.05
0.10
0.15
0.20
0.25