Time Series II – Frequency Domain Methods

John Fricks

Astrostatistics Summer School
Penn State University
University Park, PA 16802

June 8, 2007

Introduction

- In the previous tutorial, we discussed rules for evolution of a time series in the time domain.
- In this tutorial, we will discuss time series from a different perspective; we will look at the frequency components of the data.

Stationarity

- The assumption of stationarity imposes regularity on a time series model.
- We will need repeated observations with the same or similar relationship to one another in order to estimate the underlying relationships between observations.
- There are other ways to do this, but stationarity is the most common and perhaps most basic.

Strictly Stationary

A time series $\ldots, x_{-1}, x_0, x_1, x_2, \ldots$ is *strictly stationary* if for a sequence of times $t_1, t_2, \ldots, t_k$

$$\{x_{t_1}, \ldots, x_{t_k}\}$$

has the same distributions as

$$\{x_{t_1+h}, \ldots, x_{t_k+h}\}$$

for every integer h. In other words,

$$P\{x_{t_1} \leq c_1, \ldots, x_{t_k} \leq c_k\} = P\{x_{t_1+h} \leq c_1, \ldots, x_{t_k+h} \leq c_k\}.$$

Weak Stationarity

An important measure of dependency in time series is autocovariance. This is defined as

$$\gamma(t, s) = E(x_t - \mu_t)(x_s - \mu_s)$$

where $\mu_t = Ex_t$.

The time series x_t is *weakly stationary* if μ_t is constant and $\gamma(s, t)$ depends only on the distance $|s - t|$.

In the case of Gaussian time series, these two concepts of stationarity overlap.

Autocovariance Notation

For a weakly stationary time series, the notation used for autocovariance uses only lag:

$$\gamma(h) = E(x_t - \mu)(x_{t-h} - \mu)$$

where μ is the constant variance.
We also have a concept of the autocorrelation function which we saw in the first section in the ACF plot. The autocorrelation function is defined as

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)}$$

Discrete Fourier Transform of the Time Series

What if we are less interested in how our underlying process evolves in time and are more interested in the variance of the time series at certain frequencies?

We may attempt to apply a Fourier transform to the data. For our time series, $x_1, ..., x_n$, the discrete Fourier transform would be

$$d(\omega_j) = n^{-1/2} \sum_{t=1}^{n} x_t exp(-2\pi i t \omega_j)$$

where $\omega_j = 0, 1/n, ..., (n-1)/n$.

An Alternate Representation

Note that we can break up $d(\omega_j)$ into two parts

$$d(\omega_j) = n^{-1/2} \sum_{t=1}^{n} x_t cos(2\pi i \omega_j t) - in^{-1/2} \sum_{t=1}^{n} x_t sin(2\pi i \omega_j t)$$

which we could write as as a cosine component and a sine component

$$d(\omega_j) = d_c(\omega_j) - i d_s(\omega_j)$$

We may use an inverse Fourier transform to rewrite the data as

$$
\begin{aligned}
x_t &= n^{-1/2} \sum_{j=1}^{n} d(\omega_j) e^{2\pi i \omega_j t} \\
&= n^{-1/2} \sum_{j=1}^{n} d(\omega_j) e^{2\pi i \omega_j t} \\
&= a_0 + n^{-1/2} \sum_{j=1}^{m} d(\omega_j) e^{2\pi i \omega_j t} + n^{-1/2} \sum_{j=m+1}^{n} d(\omega_j) e^{2\pi i \omega_j t} \\
&= a_0 + \sum_{j=1}^{m} \frac{2 d_c(\omega_j)}{n^{-1/2}} cos(2\pi i \omega_j t) + \sum_{j=1}^{m} \frac{2 d_s(\omega_j)}{n^{-1/2}} sin(2\pi i \omega_j t)
\end{aligned}
$$

where $m = \lfloor \frac{n}{2} \rfloor$

- We can think of the Fourier transform as a regression of x_t on sines and cosines.
- The coefficients of this regression are equal to $2/\sqrt{n}$ times the sine part and the cosine part of the Fourier transforms respectively.
- Also note that if our time series is Gaussian, $d_c(\omega_j)$ and $d_s(\omega_j)$ are Gaussian random variables.

The Periodogram

- The periodogram is defined as

$$I(\omega_j) = |d(\omega_j)|^2 = d_c^2(\omega_j) + d_c^2(\omega_j)$$

- If there is no periodic trend in the data, then $Ed(\omega_j) = 0$, and the periodogram expresses the variance of x_t at frequency ω_j.
- If a periodic trend exists in the data, then $Ed(\omega_j)$ will be the contribution to the periodic trend at the frequency ω_j.

The Periodogram

- What are we trying to estimate with the periodogram?
- We can use the periodogram to find periodic trends in the data.
- Is there information left in the periodogram after the trend is removed?
- Assuming that we have a stationary time series, what does the periodogram estimate?

The Spectral Density

The spectral density is the Fourier transform of the autocovariance function

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

for $\omega \in (-0.5, 0.5)$. Note that this is a population quantity. (i.e. This is a constant quantity defined by the model.)

Why is the periodogram an estimate for the spectral density? Let m be the sample mean of our data.

$$\begin{aligned}
I(\omega_j) &= |d(\omega_j)|^2 = n^{-1} \sum_{t=1}^{n} x_t e^{-2\pi i \omega_j t} \overline{\sum_{t=1}^{n} x_t e^{-2\pi i \omega_j t}} \\
&= |d(\omega_j)|^2 = n^{-1} \sum_{t=1}^{n} \sum_{s=1}^{n} (x_t - m)(\overline{x_s} - m) e^{-2\pi i \omega_j (t-s)} \\
&= n^{-1} \sum_{h=-(n-1)}^{(n-1)} \sum_{t=1}^{n-|h|} (x_{t+|h|} - m)(x_t - m) e^{-2\pi i \omega_j (h)} \\
&= \sum_{h=-(n-1)}^{(n-1)} \hat{\gamma}(h) e^{-2\pi i \omega_j (h)} \approx f(\omega_j)
\end{aligned}$$

- Is the periodogram a **good** estimator for the spectral density? Not really!
- The periodogram, $I(\omega_1), \ldots, I(\omega_m)$, attempt to estimate paramters $f(\omega_1), \ldots, f(\omega_m)$. We have nearly the same number of parameters as we have data.
- Moreover, the number of parameters grow as a constant proportion of the data. Therefore, the periodogram is NOT a consistent estimator of the spectral density.

Moving Average

- A simple way to improve our estimates is to use a moving average smoothing technique

$$\hat{f}(\omega_j) = \frac{1}{2m+1} \sum_{k=-m}^{m} I(\omega_{j-k})$$

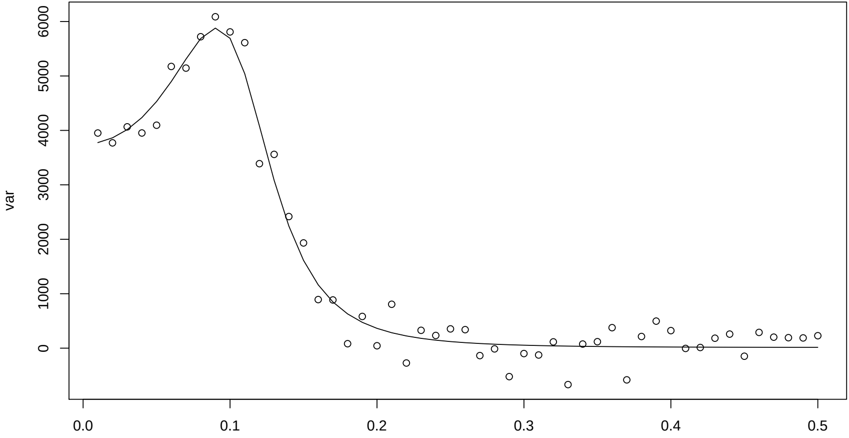
- We can also iterate this procedure of uniform weighting to be more weight on closer observations.

$$\hat{u}_t = \frac{1}{3} u_{t-1} + \frac{1}{3} u_t + \frac{1}{3} u_{t+1}$$

Then, we iterate.

$$\hat{\hat{u}}_t = \frac{1}{3} \hat{u}_{t-1} + \frac{1}{3} \hat{u}_t + \frac{1}{3} \hat{u}_{t+1}$$

Then, substitute to obtain better weights.


var
0
1000
2000
3000
4000
5000
6000
0.0
0.1
0.2
0.3
0.4
0.5

Smoothing Summary

- Smoothing decreases variance by averaging over the periodogram of neighboring frequencies.
- Smoothing introduces bias because the expectation of neighboring periodogram values are similar but not identical to the frequency of interest.
- Beware of oversmoothing!

Tapering

- Tapering corrects bias introduced from the finiteness of the data.
- The expected value of the periodogram at a certain frequency is not quite equal to the spectral density.
- It is affected by the spectral density at neighboring frequency points.
- For a spectral density which is more dynamic, more tapering is required.

Why do we need to taper?

Our theoretical model $..., x_{-1}, x_0, x_1, ...$ consists of a doubly infinite time series. We could think of our data, y_t as the following transformation of the model

$$y_t = h_t x_t$$

where $h_t = 1$ for $t = 1, ..., n$ and zero otherwise. This has repercussions on the expectation of the periodogram of our data.

$$E[I_y(\omega_j)] = \int_{-0.5}^{0.5} W_n(\omega_j - \omega) f_x(\omega) d\omega$$

where $W_n(\omega) = |H_n(\omega)|^2$ and $H_n(\omega)$ is the Fourier transform of the sequence h_t.

The Taper

Specifically,

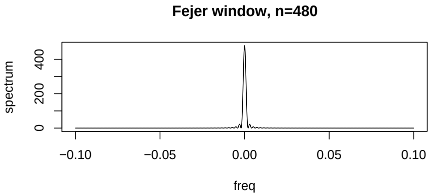
$$H_n(\omega) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} h_t e^{-2\pi i \omega t}$$

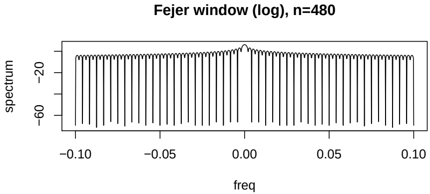
When we put in the h_t above, we obtain a spectral window of

$$W_n(\omega) = \frac{\sin^2(n2\pi\omega)}{\sin^2(\pi\omega)}.$$

We set $W_n(0) = n$.

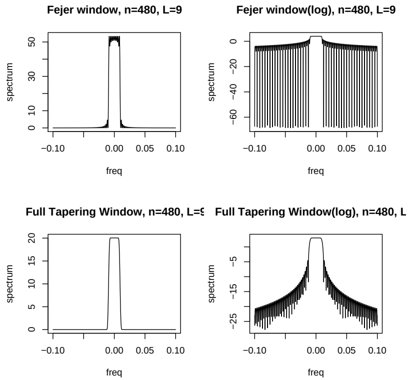
There are problems with this spectral window, namely there is too much weight on neighboring frequencies (sidelobes).

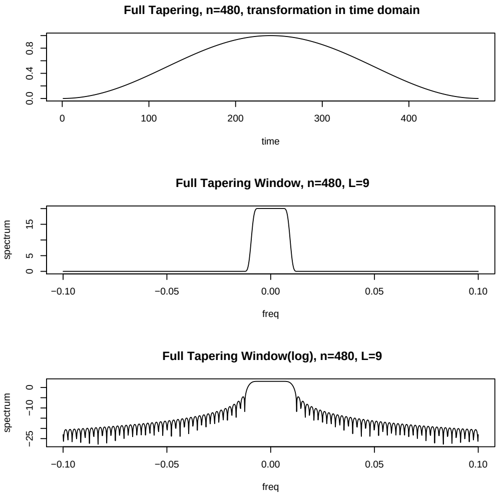


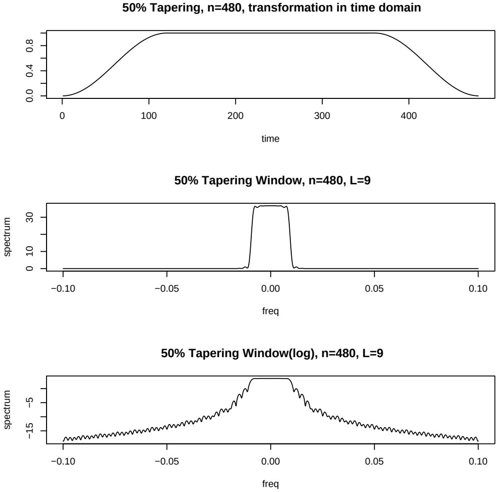


One way to fix this is to use a Cosine taper. We select a transform h_t to be

$$h_t = 0.5\left[1 + cos\left(\frac{2\pi(t-\bar{t})}{n}\right)\right]$$



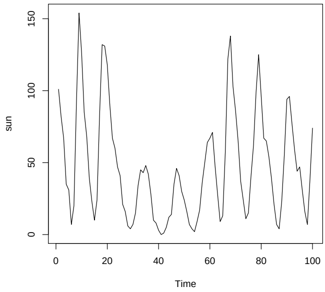

Full Tapering, n=480, transformation in time domain
0.0
0.4
0.8
0
100
200
300
400
time
Full Tapering Window, n=480, L=9
spectrum
0
5
15
−0.10
−0.05
0.00
0.05
0.10
freq
Full Tapering Window(log), n=480, L=9
spectrum
0
−10
−25
−0.10
−0.05
0.00
0.05
0.10
freq


50% Tapering, n=480, transformation in time domain
0.0
0.4
0.8
0
100
200
300
400
time
50% Tapering Window, n=480, L=9
spectrum
0
10
30
−0.10
−0.05
0.00
0.05
0.10
freq
50% Tapering Window(log), n=480, L=9
spectrum
−15
−5
−0.10
−0.05
0.00
0.05
0.10
freq

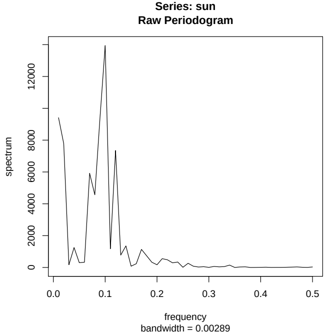
Smoothing and Tapering

- Smoothing introduces bias, but reduces variance.
- Smoothing tries to solve the problem of too many "parameters".
- Tapering decreases bias and introducese variance.
- Tapering attempts to diminish the influence of sidelobes that are introduced via the spectral window.

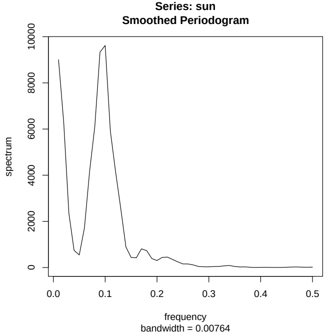
Wolfer sunspots 1770-1869



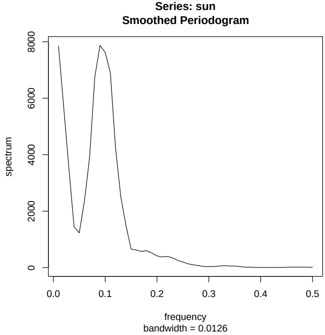
Raw Periodogram



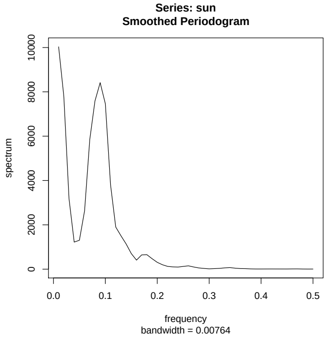
Periodogram with Smoothing Window of 3



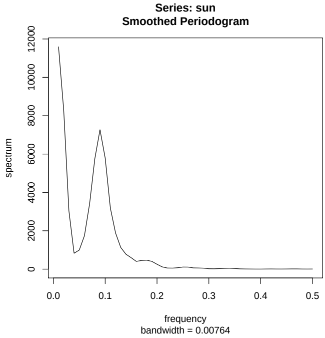
Periodogram with Smoothing Window of 5



Periodogram with Smoothing Window of 3 with Some Tapering

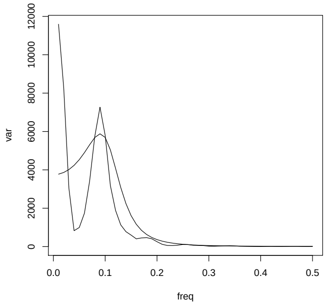


Periodogram with Smoothing Window of 3 with More Tapering



Smoothed Periodogram with ARMA Spectral Density

The smoothed periodogram of the sun spot data with the spectral density of the AR(3) model overlayed.



Dynamic Fourier Analysis

- What can be done for non-stationary data?
- One approach is to decompose our time series as a sum of a non-constant (deterministic) trend plus a stationary "noise" term:

$$x_t = \mu_t + y_t$$

- What if our data instead appears as a stationary model locally, but globally the model appears to shift? One approach is to divide the data into shorter sections (perhaps overlapping) and
- This approach is developed in Shumway and Stoffer. One essentially looks at how the spectral density changes over time.

Wavelets

- We have been using Fourier components as a basis to represent stationary processes and seasonal trends.
- Since we are dealing with finite data, we must use a finite number of terms, and perhaps one could use an alternative basis.
- Wavlets are one option to accomplish this goal. They are particularly well suited to the same situation as dynamic Fourier analysis.

References

- Robert Shumway and David Stoffer. *Time Series Analysis and Its Applications.* Springer NY, 2006.
- Peter Brockwell and Richard Davis. *Time Series: Theory and Methods, Second Ed.* Springer NY, 1991.
- David Brillinger. *Time Series: Data Analysis and Theory.* SIAM, 2001.
- Donald Percival and Andrew Walden. *Spectral Analysis for Physical Applications.* Cambridge University Press, 1993.
- Donald Percival and Andrew Walden. *Wavelet Methods for Time Series Analysis.* Cambridge University Press, 2000.
- Stéphane Mallat. *A Wavelet Tour of Signal Processing.* Academic Press, 1998.