DATA ANALYTICS

Unit 3: Spectral Analysis of Time Series

Jyothi R.
Department of Computer Science and Engineering

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, $x_1, \ldots, x_n$, the discrete Fourier transform would be

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

Interpreting DFT and Another Representation

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

- which we can write as a cosine component and a sine component

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

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 $E[d(\omega_j)] = 0$, and the Periodogram expresses the variance of x_t at frequency ω_j.

- If a periodic trend exists in the data, then $E[d(\omega_j)]$ 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 auto covariance 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.)

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.

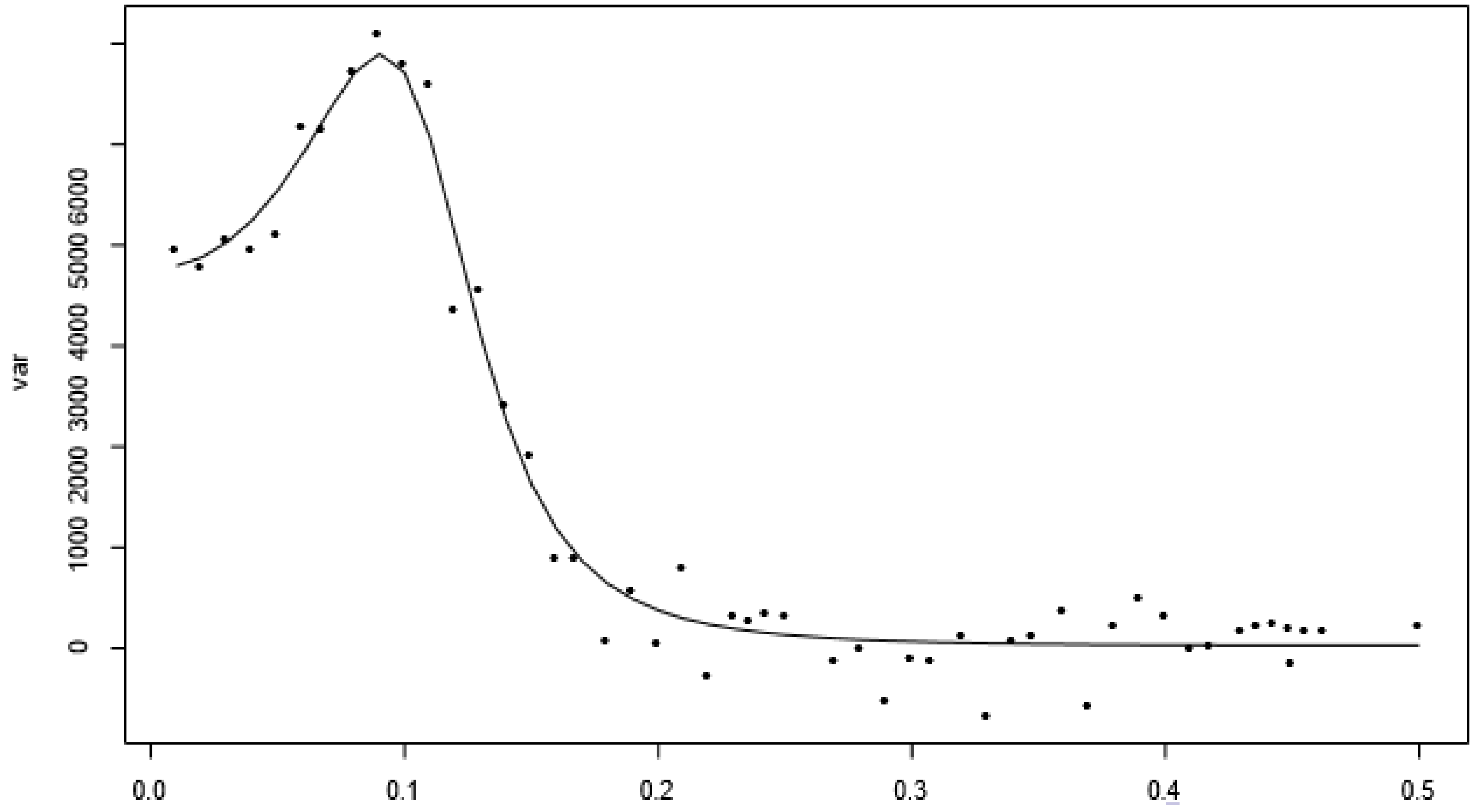
$$\hat{u}_t = \frac{1}{3}u_{t-1} + \frac{1}{3}u_t + \frac{1}{3}u_{t+1}$$

- Then, we iterate.
- Then, substitute to obtain better weights.

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

DATA ANALYTICS

Moving Average



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 over smoothing!

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 $t = 1, \ldots, n$ 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,

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

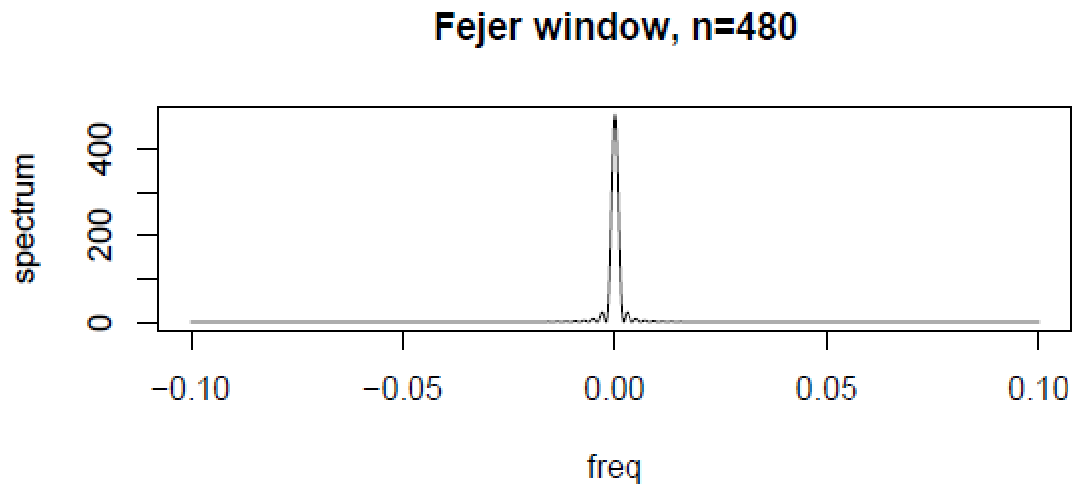
When we put in the h_t above, we obtain a spectral window of

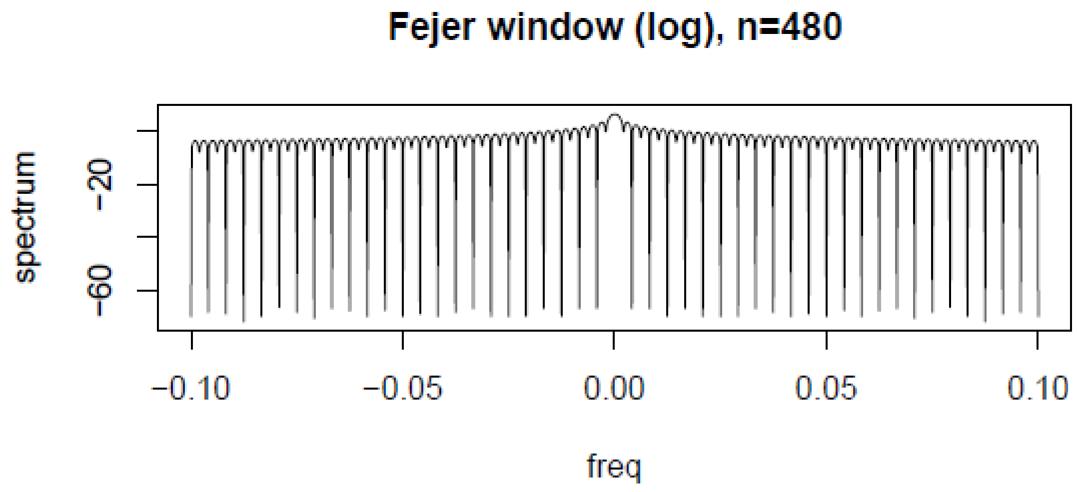
$$W_n(\omega) = \frac{\sin^2(n2\pi\omega)}{\sin^2(\pi\omega)}.$$

We set $W_n(0) = n$.

Smoothing and Tapering

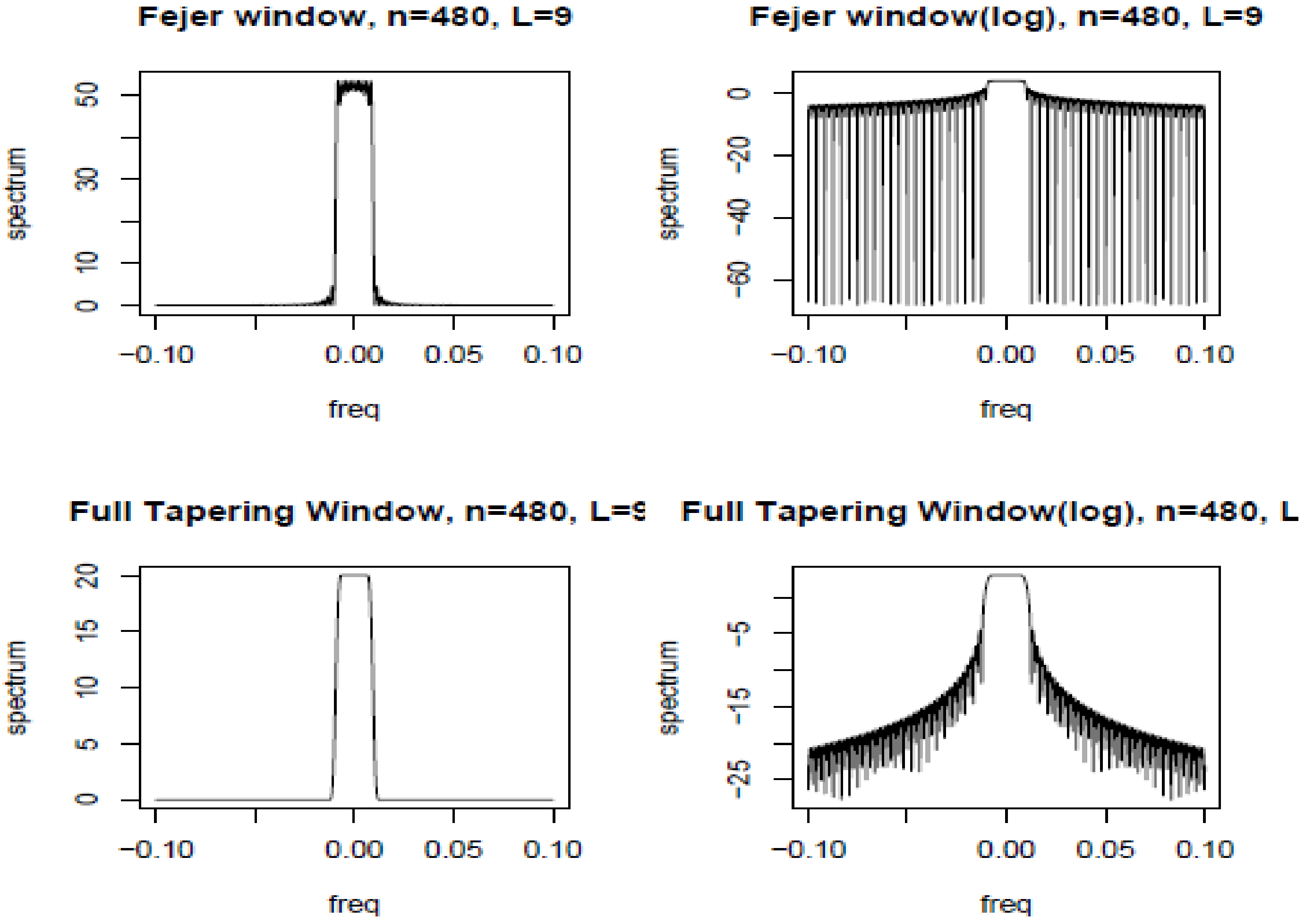
- There are problems with this spectral window, namely there is too much weight on neighboring frequencies (sidelobes).





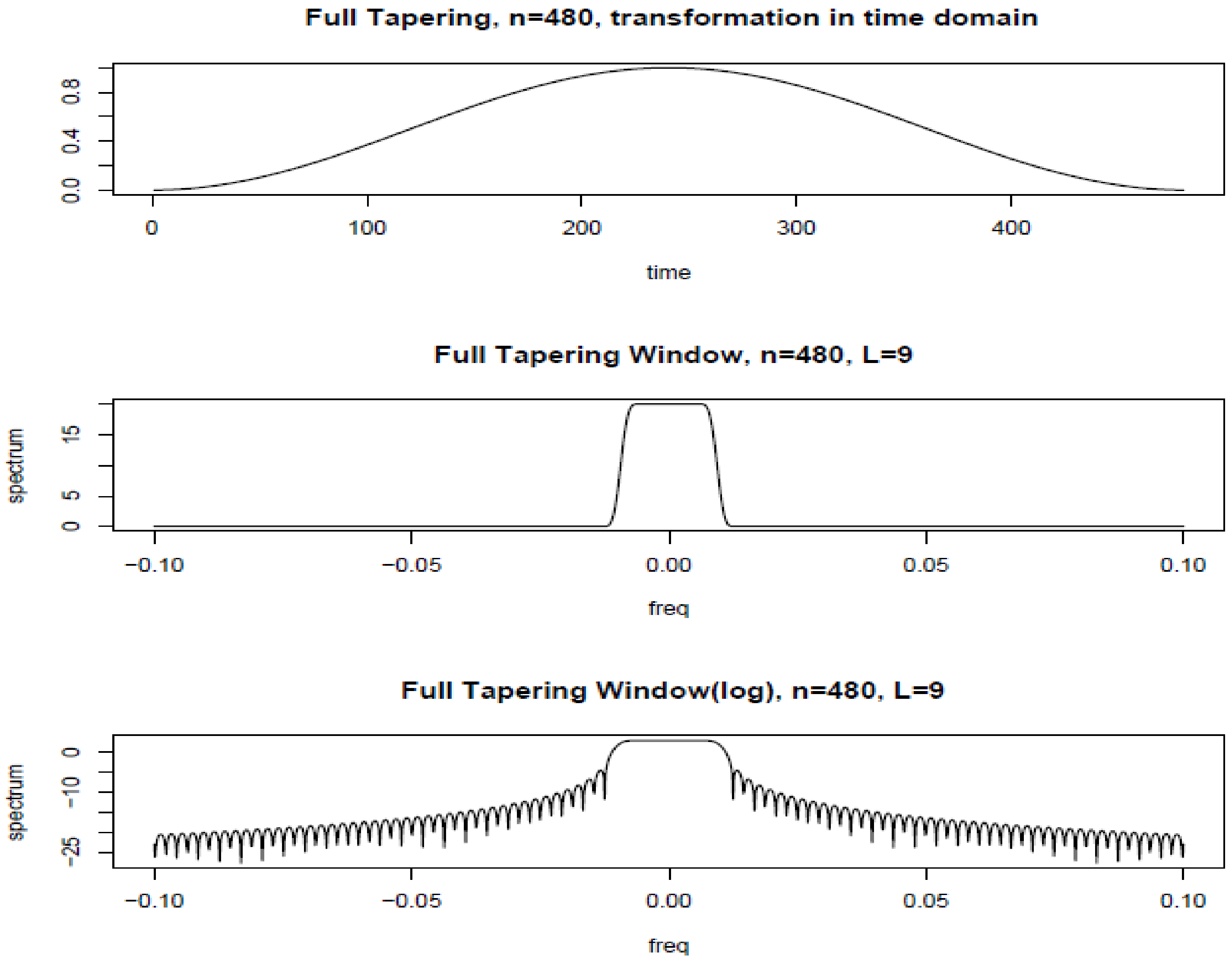
Smoothing and Tapering

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



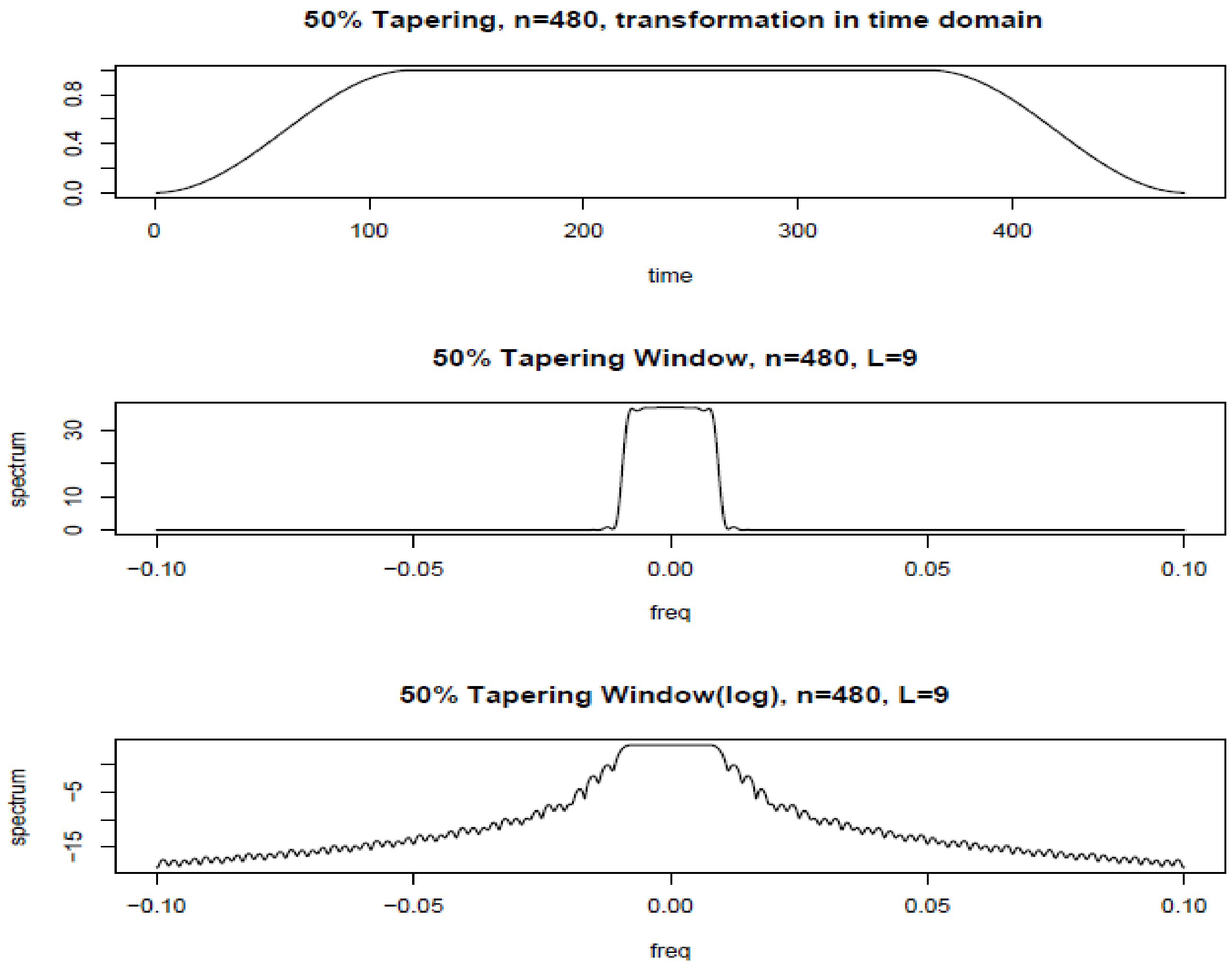
$$h_t = 0.5\left[1 + \cos\left(\frac{2\pi(t - \bar{t})}{n}\right)\right]$$

Smoothing and Tapering



DATA ANALYTICS

Smoothing and Tapering

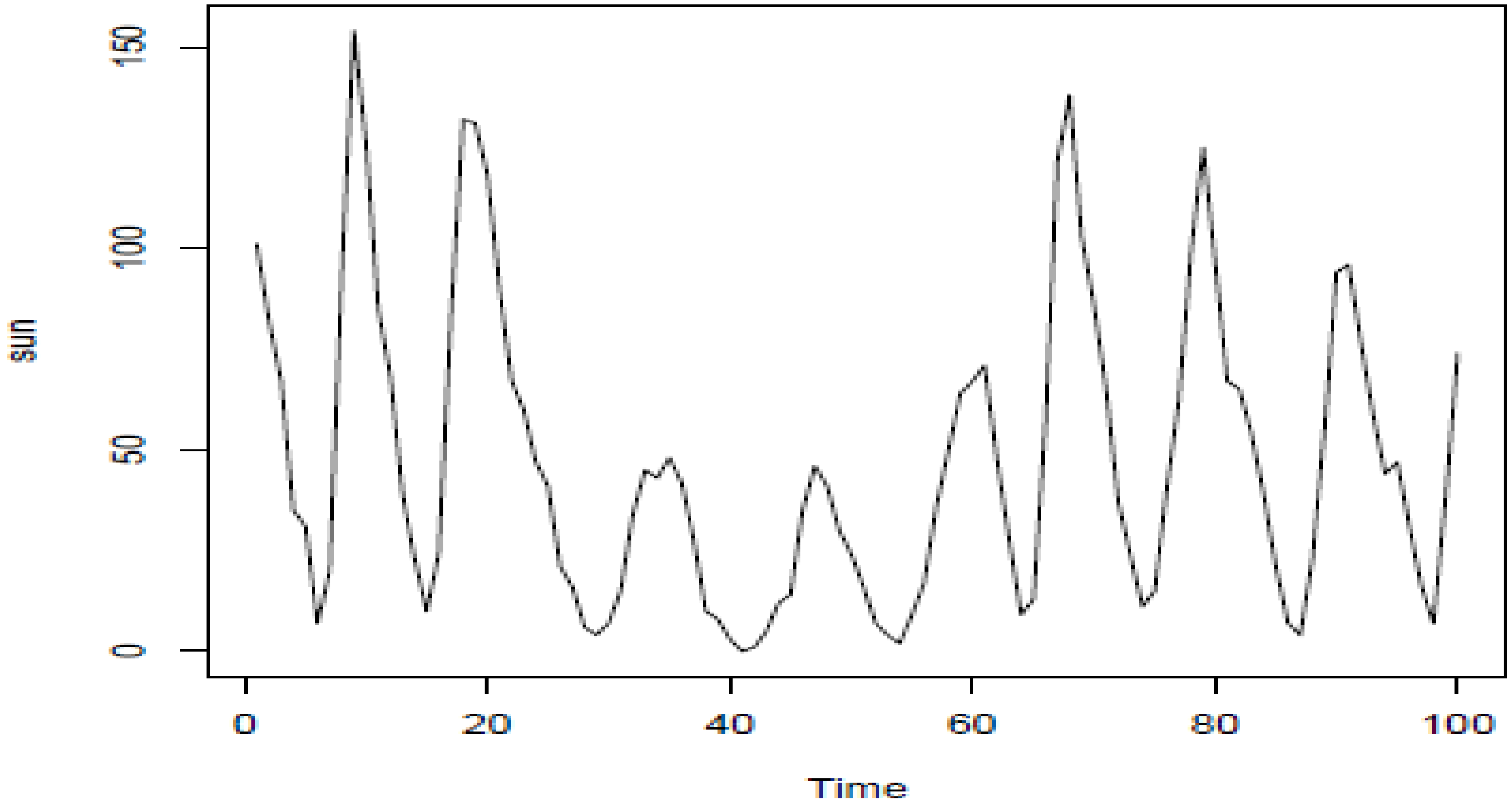


Smoothing and Tapering

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

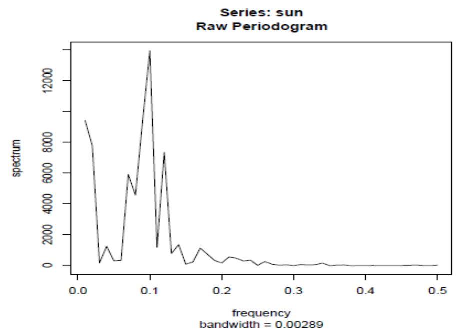
Examples

Wolfer sunspots 1770-1869

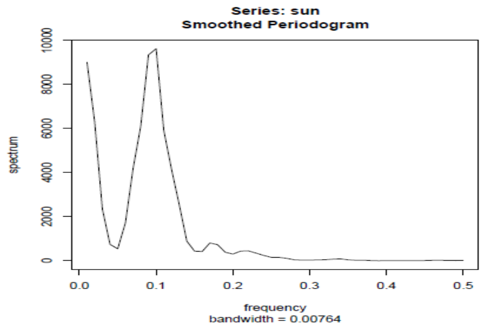


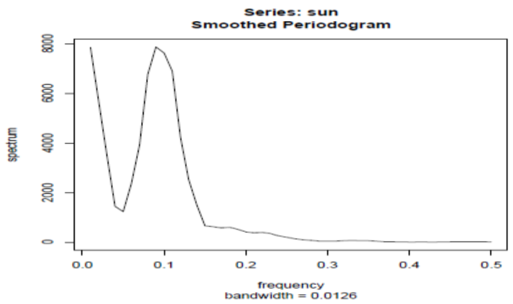
Examples: Smoothed Periodogram with ARMA Spectral Density

Raw Periodogram

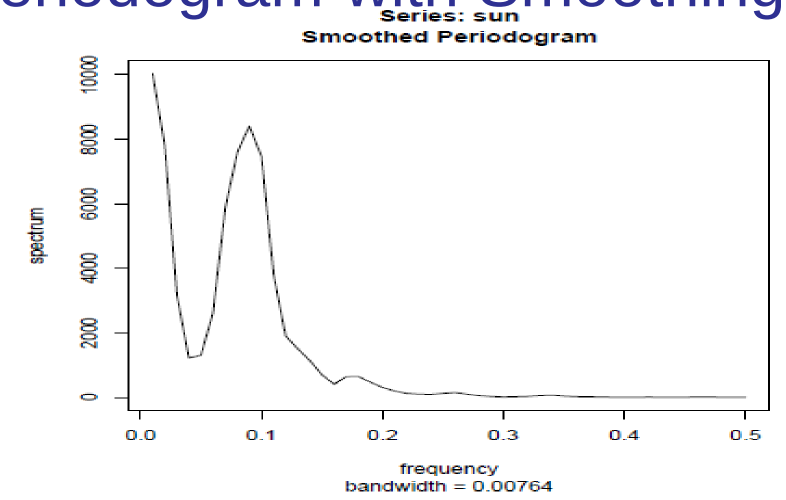


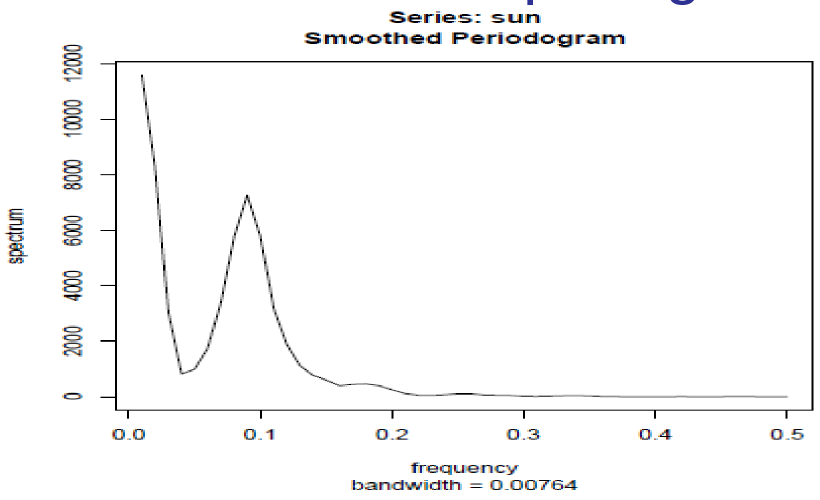
Periodogram with Smoothing Window of 3, 5





Periodogram with Smoothing Window of 3 with Tapering and more tapering





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.

- Wavelets are one option to accomplish this goal. They are particularly well suited to the same situation as Dynamic Fourier analysis.

References

Text Book:

“Business Analytics, The Science of Data-Driven Making”, U. Dinesh Kumar, Wiley 2017

Image Courtesy

https://www.abs.gov.au/websitedbs/D3310114.nsf/home/Time+Series+Analysis:+The+Basics

https://otexts.com/fpp2/stationarity.html
https://www.abs.gov.au/websitedbs/D3310114.nsf/home/Time+Series+Analysis:+The+Basics
https://bookdown.org/rdpeng/timeseriesbook/spectral-analysis.html
https://www.stat.berkeley.edu/~bartlett/courses/153-fall2010/lectures/15.pdf
https://astrostatistics.psu.edu/su07/fricks_2timeseries07.pdf)
https://blog.octo.com/en/time-series-features-extraction-using-fourier-and-wavelet-transforms-on-ecg-data/
https://jmread.github.io/talks/Time_Series_AI.pdf

THANK YOU

Jyothi R
Assistant Professor, Department of Computer Science
jyothir@pes.edu

Practice Quiz

5. In a pure auto-regressive process, AR(*p*), the value of *p* can be identified using

(a) Auto-correlation function

(b) Partial auto-correlation function

(c) Auto-correlation and partial auto-correlation function

(d) Ljung–Box test

Practice Quiz

6. Power of a forecasting model is calculated using

(a) Root mean square error (RMSE)

(b) Theil's coefficient

(c) Mean absolute percentage error (MAPE)

(d) Bayesian information criteria (BIC)

7. A necessary condition for accepting a time-series forecasting model is

(a) The residuals should follow a normal distribution

(b) The residuals should be white noise

(c) The residuals should be black noise

(d) The residuals should follow a normal distribution and the *R*-square should be high

Practice Quiz

8. In an ARIMA model, differencing is carried out

(a) To convert a stationary process to a non-stationary process

(b) To convert a non-stationary process to a stationary process

(c) To remove seasonal fluctuations from the data

(d) To remove cyclical fluctuations from the data

Practice Quiz

9. Overall fitness of a forecasting model is checked using

(a) Durbin–Watson Test

(a) Theil coefficient

(c) Ljung–Box test

(d) Dickey–Fuller test

DATA ANALYTICS

Practice Quiz

10. Presence of non-stationarity is checked using

(a) Durbin−Watson Test

(b) Theil coefficient

(c) Ljung−Box test

(d) Dickey−Fuller test