Lecture 12

Spectral Analysis of Time Series II

Readings: CC08 Chapter 14; BD16 Chapter 4, Chapter 10.1; SS17 Chapter 1.5-1.6, Chapter 4.4-Chapter 4.6, Chapter 4.8, Chpater 5.5

MATH 8090 Time Series Analysis
Week 12

Whitney Huang
Clemson University

Agenda

1. **Review**
2. **Southern Oscillation Index Example**
3. **Parametric Spectral Estimation**
4. **Lagged Regression Models**

Review: Nonparametric Spectral Estimation

- Periodogram: $I(\omega_j) = |d(\omega_j)|^2$, where

$$d(\omega_j) = n^{-\frac{1}{2}} \sum_{t=1}^{n} y_t e^{-2\pi i \omega_j t}, \ \omega_j = \frac{j}{n}, \ j = 0, 1, \cdots, n-1$$

  - $\frac{I(\omega_j)}{\frac{1}{2} f(\omega_j)} \overset{\approx i.i.d}{\sim} \chi^2_2, \ j = 1, \cdots, m = \frac{n-1}{2} \Rightarrow \mathbb{E}[I(\omega_j)] \approx f(\omega_j)$ (unbiased)
  - But $\mathrm{Var}[I(\omega_j)] \approx f^2(\omega_j)$ (inconsistent because it does not go to 0 as $n \to \infty$)

- Smooth the periodogram
  - Averaged periodogram: $\bar{f}(\omega_j) = \frac{1}{L} \sum_{k=-l}^{l} I(\omega_{j+k}), \quad L = 2l + 1$
  - Smoothed periodogram: $\bar{f}(\omega_j) = \sum_{k=-l}^{l} W_l(k) I(\omega_{j+k})$

- Pointwise CI for $f(\omega_j)$:

$$\frac{\nu \bar{f}(\omega_j)}{\chi^2_\nu(1-\alpha/2)} \le f(\omega_j) \le \frac{\nu \bar{f}(\omega_j)}{\chi^2_\nu(\alpha/2)}, \quad \nu = \frac{2}{\sum_{k=-l}^{l} W_l^2(k)}$$

Spectral ANOVA

- For odd $n = 2m + 1$, the inverse transform can be written

$$y_t - \bar{y} = \frac{2}{\sqrt{n}} \sum_{j=1}^{m} \left[d_{\cos}(\omega_j) \cos(2\pi\omega_j t) + d_{\sin}(\omega_j) \sin(2\pi\omega_j t) \right].$$

- Square and sum over t; orthogonality of sines and cosines implies that

$$\begin{aligned} \sum_{t=1}^{n} (y_t - \bar{y})^2 &= 2 \sum_{j=1}^{m} \left[d_{\cos}(\omega_j)^2 + d_{\sin}(\omega_j)^2 \right] \\ &= 2 \sum_{j=1}^{m} I(\omega_j) \end{aligned}$$

We have partitioned $\sum_{t=1}^{n} (y_t - \bar{y})^2$ into $2 \times \sum_{j=1}^{m} I(\omega_j)$. This leads to Spectral ANOVA $\Rightarrow$ total variance = sum of frequency components

Spectral ANOVA (Cont'd)

Source	df	SS	MS
ω_1	2	$2I(\omega_1)$	$I(\omega_1)$
ω_2	2	$2I(\omega_2)$	$I(\omega_2)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
ω_m	2	$2I(\omega_m)$	$I(\omega_m)$
Total	$2m = n - 1$	$\sum(y_t - \bar{y})^2$	

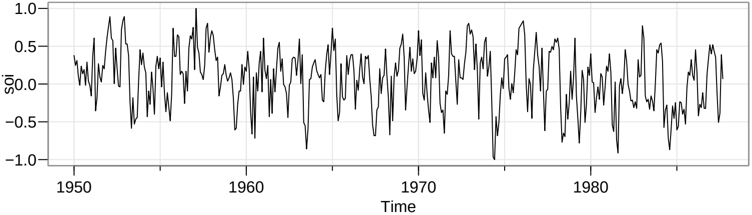
Toy example:

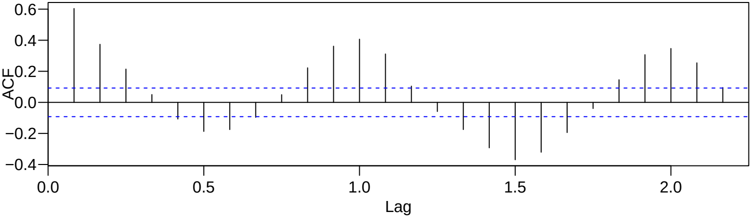
```
> x <- c(1, 2, 3, 2, 1) - mean(x)
> c1 <- cos(2 * pi * (1:5) * (1 / 5)); s1 <- sin(2 * pi * (1:5) * (1 / 5))
> c2 <- cos(2 * pi * (1:5) * (2 / 5)); s2 <- sin(2 * pi * (1:5) * (2 / 5))
> omega1 <- cbind(c1, s1); omega2 <- cbind(c2, s2)
> anova(lm(x ~ omega1 + omega2))
Warning in anova.lm(lm(x ~ omega1 + omega2)) :
  ANOVA F-tests on an essentially perfect fit are unreliable
Analysis of Variance Table

Response: x
          Df  Sum Sq Mean Sq F value Pr(>F)
omega1     2 2.74164 1.37082
omega2     2 0.05836 0.02918
Residuals  0 0.00000
```

Example: Southern Oscillation Index (SOI)

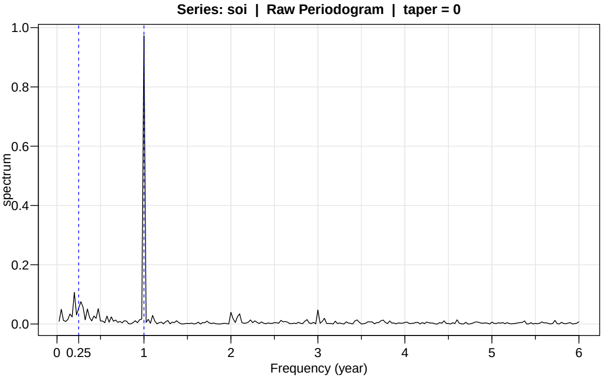
Southern Oscillation Index (SOI) for a period of 453 months ranging over the years 1950-1987





What are the hidden periods of SOI?

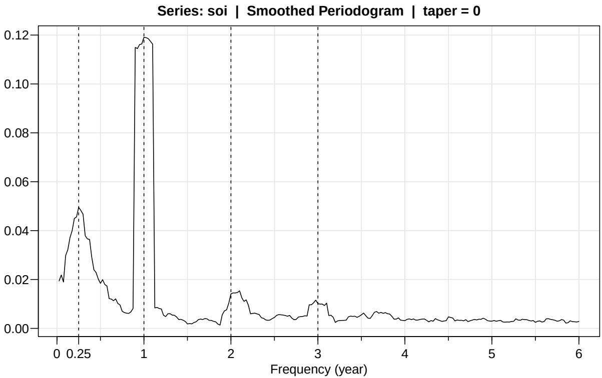
SOI Example: Raw Periodogram



An approximate 95% confidence interval for $f(\omega)$:

ω	Period	Power	Lower	Upper
$\frac{1}{48}$	4 years	0.0537	0.0146	2.1222
$\frac{1}{12}$	1 year	0.9722	0.2636	38.4011

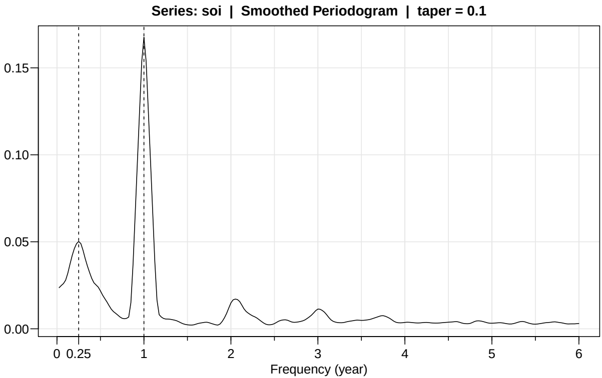
SOI Example: Averaged Periodogram (Daniell with $l = 4$)



An approximate 95% confidence interval for $f(\omega)$:

ω	Period	Power	Lower	Upper
$\frac{1}{48}$	4 years	0.0495	0.0279	0.1113
$\frac{1}{12}$	1 year	0.1191	0.0670	0.2677

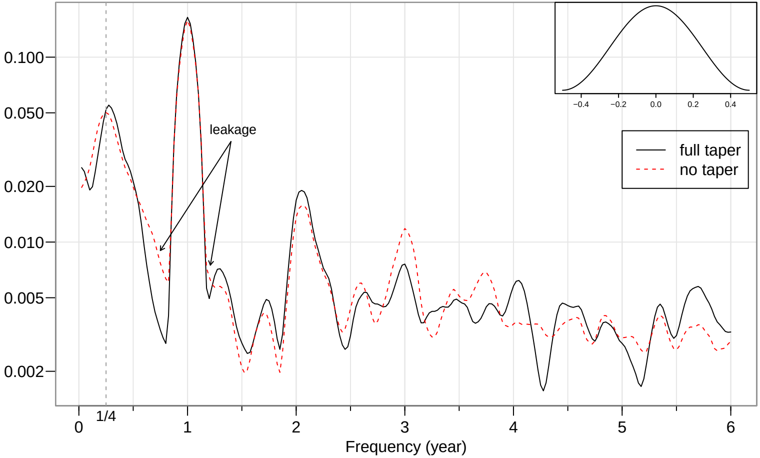
SOI Example: Smoothed Periodogram (modified Daniell $c(3,3)$)



An approximate 95% confidence interval for $f(\omega)$:

ω	Period	Power	Lower	Upper
$\frac{1}{48}$	4 years	0.0502	0.0283	0.1129
$\frac{1}{12}$	1 year	0.1675	0.0943	0.3767

SOI Example: Apply Tapering to Alleviate Spectral Leakage



The tapered spectrum does a better job in separating the yearly cycle $\omega = 1$ and the El Niño cycle $\omega = \frac{1}{4}$

Seasonally Adjusted SOI [Source: Peter Bloomfield's ST 730 Lecture Notes]

- The Southern Oscillation Index data provided by Shumway and Stoffer is not seasonally adjusted, which explains the substantial peaks in the periodogram at the annual frequency

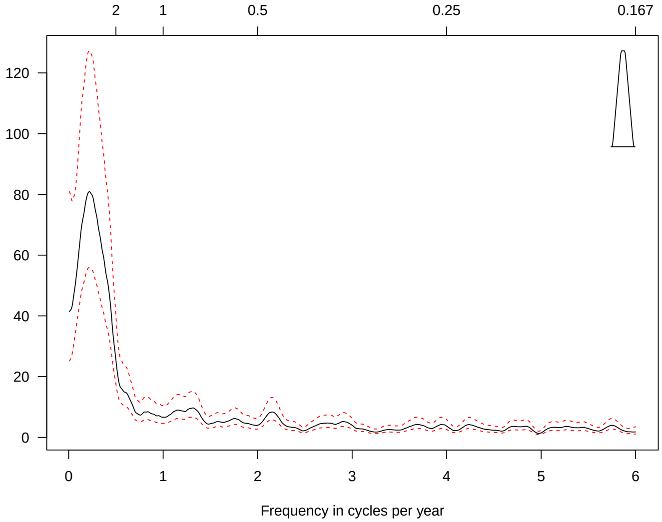
Seasonally Adjusted SOI [Source: Peter Bloomfield's ST 730 Lecture Notes]

- The Southern Oscillation Index data provided by Shumway and Stoffer is not seasonally adjusted, which explains the substantial peaks in the periodogram at the annual frequency
- So the series is non-stationary, and has neither an autocovariance function nor a spectral density function

Seasonally Adjusted SOI [Source: Peter Bloomfield's ST 730 Lecture Notes]

- The Southern Oscillation Index data provided by Shumway and Stoffer is not seasonally adjusted, which explains the substantial peaks in the periodogram at the annual frequency
- So the series is non-stationary, and has neither an autocovariance function nor a spectral density function
- A more sensible analysis uses the seasonally adjusted series. (Bloomfield did this by fitting a seasonal means model using data from 1876-2010.)

SOI Example from Bloomfield: Smoothed Periodogram



Note that the peak at the annual frequency disappear due to the removal of the annual cycle

Parametric versus Nonparametric Estimation

- Parametric estimation: estimate a model that is specified by a fixed number of parameters
- Nonparametric estimation: estimate a model that is specified by a number of parameters that can grow as the sample grows

The smoothed periodogram estimates we have considered are **nonparametric**: the estimates of the spectral density can be parameterized by estimated values at ω_j's. As $n \uparrow$, the number of distinct frequency values increases

The time domain models we considered are **parametric**. For example, an ARMA(p,q) process can be completely specified with $p + q + 1$ parameters

Parametric Spectral Estimation

The typical approach is to use the maximum likelihood parameter estimates $(\hat{\phi}_1, \cdots, \hat{\phi}_p, \hat{\sigma}^2)$ for the parameters of an AR(p), and then compute $f(\omega)$ for this estimated AR model:

$$\hat{f}(\omega) = \frac{\hat{\sigma}^2}{|\hat{\phi}(e^{-2\pi\omega})|^2}$$

For large n,

$$\mathrm{Var}(\hat{f}(\omega)) \approx \frac{2p}{n} f^2(\omega)$$

- The bias decreases as $p \uparrow$, the number of parameters increase, as one can model more complex spectra

Parametric Spectral Estimation

The typical approach is to use the maximum likelihood parameter estimates $(\hat{\phi}_1, \cdots, \hat{\phi}_p, \hat{\sigma}^2)$ for the parameters of an AR(p), and then compute $f(\omega)$ for this estimated AR model:

$$\hat{f}(\omega) = \frac{\hat{\sigma}^2}{|\hat{\phi}(e^{-2\pi\omega})|^2}$$

For large n,

$$\mathrm{Var}(\hat{f}(\omega)) \approx \frac{2p}{n} f^2(\omega)$$

- The bias decreases as $p\uparrow$, the number of parameters increase, as one can model more complex spectra
- The variance increase linealy with p

ARMA Spectral Estimation

- Sometimes ARMA models are used instead
- Estimate the parameters of an ARMA(p,q) model and compute its spectral density:

$$\hat{f}(\omega) = \hat{\sigma}^2 \left| \frac{\hat{\theta}(e^{-2\pi i \omega})}{\hat{\phi}(e^{-2\pi i \omega})} \right|^2 .$$

- However, it is more common to use large AR models, rather than ARMA models

Parametric versus Nonparametric Spectral Estimation

- The main advantage of parametric spectral estimation over nonparametric is that it often gives better frequency resolution of a small number of peaks

Parametric versus Nonparametric Spectral Estimation

- The main advantage of parametric spectral estimation over nonparametric is that it often gives better frequency resolution of a small number of peaks
- This is especially important if there is more than one peak at nearby frequencies

Parametric versus Nonparametric Spectral Estimation

- The main advantage of parametric spectral estimation over nonparametric is that it often gives better frequency resolution of a small number of peaks
- This is especially important if there is more than one peak at nearby frequencies
- The disadvantage of parametric spectral estimation is the inflexibility due to the use of the restricted class of ARMA models.

Parametric Spectral Estimation: Summary

Given data $y_1, y_2, \cdots, y_n$,

1. Estimate the AR parameters $(\phi_1, \phi_2, \cdots, \phi_p, \sigma^2)$ using maximum likelihood or Yule-Walker/least squares, choose a suitable model order p using AIC or BIC

Parametric Spectral Estimation: Summary

Given data $y_1, y_2, \cdots, y_n$,

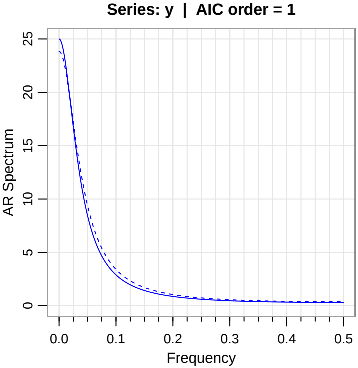
1. Estimate the AR parameters $(\phi_1, \phi_2, \cdots, \phi_p, \sigma^2)$ using maximum likelihood or Yule-Walker/least squares, choose a suitable model order p using AIC or BIC

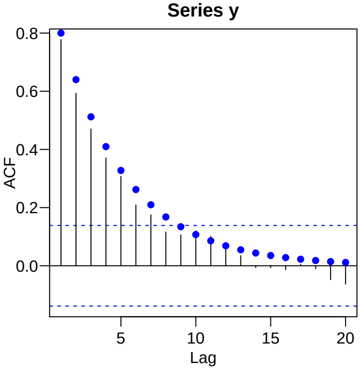
2. Use the estimates $(\hat{\phi}_1, \hat{\phi}_2, \cdots, \hat{\phi}_p, \hat{\sigma}^2)$ to compute the estimated spectral density:

$$\hat{f}(\omega) = \frac{\hat{\sigma}^2}{\left|\hat{\phi}(e^{-2\pi i\omega})\right|^2}$$

Example: AR(1) with $\phi = 0.8$

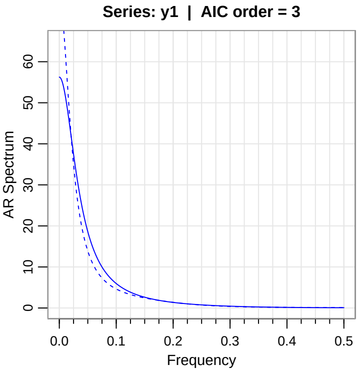
1. Use AIC to select p, the order of the AR model
2. Use the estimates $(\hat{\phi}_1, \hat{\phi}_2, \cdots, \hat{\phi}_p, \hat{\sigma}^2)$ to compute the estimated spectral

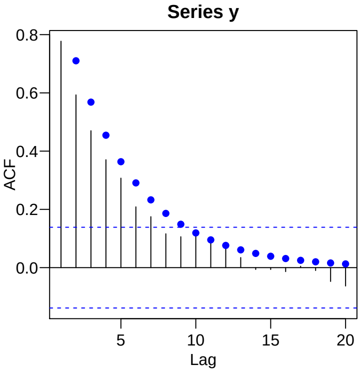




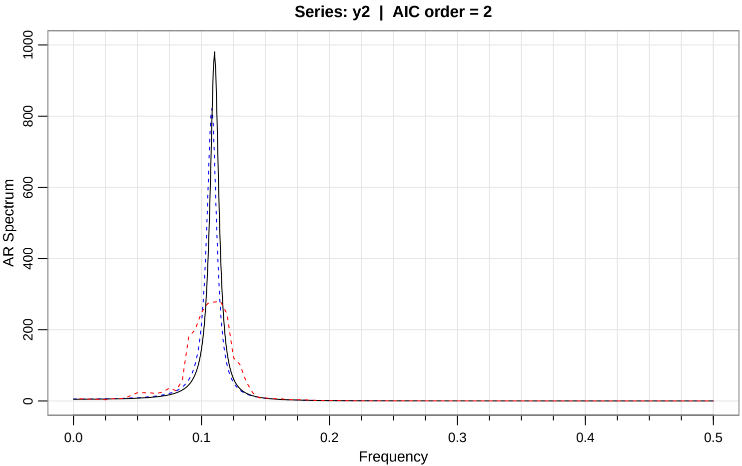
Example: ARMA$(1, 1)$ with $\phi = 0.8$ and $\theta = 0.5$

1. Use AIC to select p, the order of the AR model
2. Use the estimates $(\hat{\phi}_1, \hat{\phi}_2, \cdots, \hat{\phi}_p, \hat{\sigma}^2)$ to compute the estimated spectral



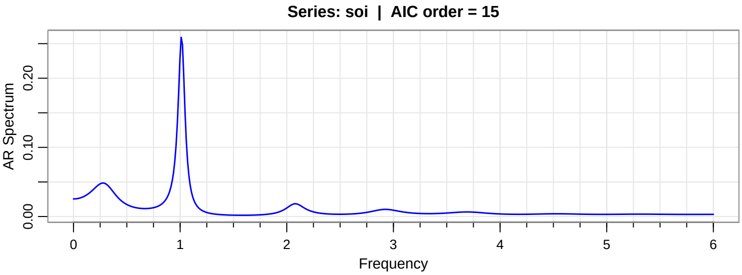


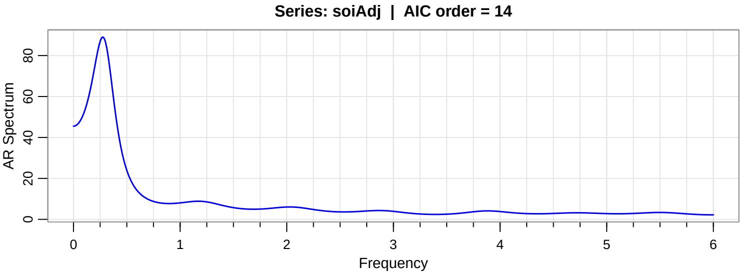
Example: AR(2) with $\phi_1 = 1.5$ and $\phi_2 = -0.95$



- Black: True
- Blue: Parametric Estimate
- Red: Non-parametric Estimate (Daniell $l = 3$)

SOI Example


Series: soi | AIC order = 15
AR Spectrum
0.00
0.10
0.20
0
1
2
3
4
5
6
Frequency


Series: soiAdj | AIC order = 14
AR Spectrum
0
20
40
60
80
0
1
2
3
4
5
6
Frequency

Lagged Regression Models

A lagged regression model expresses the output time series as a linear combination of current and past (and possibly future) values of an input series:

$$Y_t = \sum_{h=-\infty}^{\infty} \beta_h X_{t-h} + V_t,$$

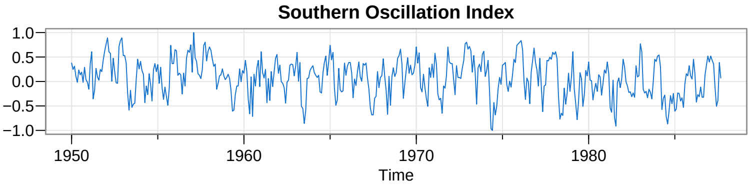
where

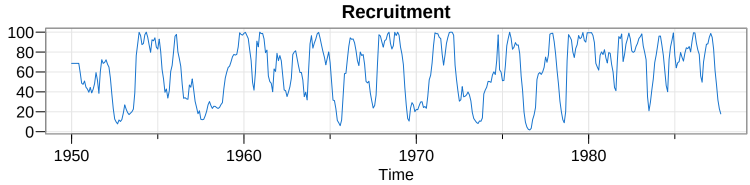
- X_t: observed input time series
- Y_t: observed output time series
- V_t: stationary noise process

Applications:

- Identify the (best linear) dynamic relationship between two time series, X_t and Y_t
- Forecast the future of Y_t based on X_t (often setting $\beta_h = 0$ for $h < 0$ to use only past inputs)

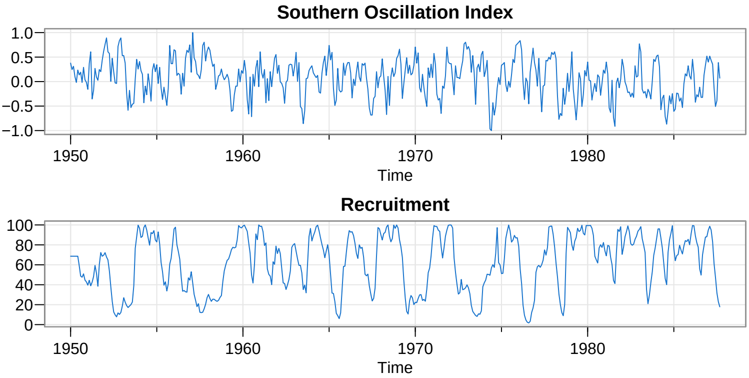
An Example of Lagged Regression Model





- We may wish to identify how the values of the recruitment series is related to the SOI

An Example of Lagged Regression Model



- We may wish to identify how the values of the recruitment series is related to the SOI
- We may wish to predict future values of recruitment from the SOI

Lagged Regression Models: Approaches

- Time domain: model the input series, extract the white time series driving it ("prewhitening"), regress with transformed output series
  - Cross-covariance function
  - Cross-correlation function
- Frequency domain: Calculate the input's spectral density, and the cross-spectral density between input and output, and find the transfer function relating them, in the frequency domain
  - Cross spectrum
  - Coherence

Cross-Covariance

Recall that the autocovariance function of a stationary process $\{Y_t\}$ is

$$\gamma_X(h) = \mathbb{E}\left[(X_{t+h} - \mu_X)(X_t - \mu_X)\right].$$

The cross-covariance function of two jointly stationary processes $\{Y_t\}$ and $\{X_t\}$ is

$$\gamma_{XY}(h) = \mathbb{E}\left[(X_{t+h} - \mu_X)(Y_t - \mu_Y)\right].$$

Note: Jointly stationary = constant means, autocovariances depending only on the lag h, and cross-covariance depends only on h

Cross-Correlation

The cross-correlation function of jointly stationary $\{X_t\}$ and $\{Y_t\}$ is

$$\rho_{XY}(h) = \frac{\gamma_{XY}(h)}{\sqrt{\gamma_X(0)\gamma_Y(0)}}.$$

Notice that $\rho_{XY}(h) = \rho_{YX}(-h)$ but $\rho_{XY}(h)$ is not necessarily equal to $\rho_{XY}(-h)$

Example: Suppose that $Y_t = \beta X_{t-\ell} + W_t$ for $\{X_t\}$ stationay and uncorrelated with $\{W_t\}$, and $\{W_t\}$ a zero mean white noise. Then $\{X_t\}$ and $\{Y_t\}$ are jointly stationary, with $\mu_Y = \beta\mu_X$,

$$\gamma_{XY}(h) = \beta\gamma_X(h+\ell).$$

- If $\ell > 0$, we say X_t leads Y_t
- If $\ell < 0$, we say X_t lags Y_t

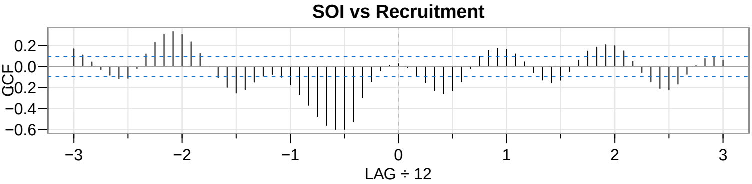
Sample Cross-Covariance and Sample Cross-Correlation

The sample cross-covariance is

$$\hat{\gamma}_{XY}(h) = \frac{1}{n}\sum_{i=1}^{n-h}(x_{t+h} - \bar{x})(y_t - \bar{y})$$

for $h \geq 0$. Then sample CCF is

$$\hat{\rho}_{XY}(h) = \frac{\hat{\gamma}_{XY}(h)}{\sqrt{\hat{\gamma}_X(0)\hat{\gamma}_Y(0)}}$$



Example: CCF of SOI and recruitment has a peak at $h = -6$. Thus, SOI leads recruitment by 6 months

Lagged Regression in the Time Domain

Suppose we wish to fit a lagged regression model of the form

$$Y_t = \beta(B)X_t + V_t = \sum_{j=0}^{\infty} \beta_j X_{t-j} + V_t,$$

where X_t is an observed input series, Y_t is the observed output series, and V_t is a stationary noise process, uncorrelated with X_t

Lagged Regression in the Time Domain

Suppose we wish to fit a lagged regression model of the form

$$Y_t = \beta(B)X_t + V_t = \sum_{j=0}^{\infty} \beta_j X_{t-j} + V_t,$$

where X_t is an observed input series, Y_t is the observed output series, and V_t is a stationary noise process, uncorrelated with X_t

One approach (pioneered by Box and Jenkins) is to fit ARMA models for X_t and V_t, and then find a simple rational representation for $\beta(B)$. This is the transfer function models

Lagged Regression in the Time Domain

$$Y_t = \beta(B)X_t + V_t = \sum_{j=0}^{\infty} \beta_j X_{t-j} + V_t,$$

For example:

$$X_t = \frac{\theta_X(B)}{\phi_X(B)} W_t,$$

$$V_t = \frac{\theta_V(B)}{\phi_V(B)} Z_t,$$

$$\beta(B) = \frac{\delta(B)}{\omega(B)} B^d$$

Notice the delay B^d, indicating that Y_t lags X_t by d steps

Lagged Regression in the Time Domain

How do we choose all of these parameters?

- Fit $\theta_X(B)$, $\phi_X(B)$ to model the input series $\{X_t\}$

Lagged Regression in the Time Domain

How do we choose all of these parameters?

- Fit $\theta_X(B)$, $\phi_X(B)$ to model the input series $\{X_t\}$
- Prewhiten the input series by applying the inverse operator $\phi_X(B)/\theta_X(B)$:

$$\tilde{Y}_t = \frac{\phi_X(B)}{\theta_X(B)} Y_t = \beta(B) W_t + \frac{\phi_X(B)}{\theta_X(B)} V_t$$

Lagged Regression in the Time Domain

How do we choose all of these parameters?

- Fit $\theta_X(B)$, $\phi_X(B)$ to model the input series $\{X_t\}$
- Prewhiten the input series by applying the inverse operator $\phi_X(B)/\theta_X(B)$:

$$\tilde{Y}_t = \frac{\phi_X(B)}{\theta_X(B)} Y_t = \beta(B) W_t + \frac{\phi_X(B)}{\theta_X(B)} V_t$$

- Calculate the cross-correlation of $\tilde{Y}_t$ with W_t,

$$\gamma_{\tilde{Y},W}(h) = \mathbb{E}\left[\tilde{Y}_{t+h} W_t\right] = \mathbb{E}\left[\sum_{j=0}^{\infty} \beta_j W_{t+h-j} W_t\right] = \sigma_W^2 \beta_h$$

to give an indication of the behavior of $\beta(B)$

Lagged Regression in the Time Domain

How do we choose all of these parameters?

- Fit $\theta_X(B)$, $\phi_X(B)$ to model the input series $\{X_t\}$
- Prewhiten the input series by applying the inverse operator $\phi_X(B)/\theta_X(B)$:

$$\tilde{Y}_t = \frac{\phi_X(B)}{\theta_X(B)} Y_t = \beta(B) W_t + \frac{\phi_X(B)}{\theta_X(B)} V_t$$

- Calculate the cross-correlation of $\tilde{Y}_t$ with W_t,

$$\gamma_{\tilde{Y},W}(h) = \mathbb{E}\left[\tilde{Y}_{t+h} W_t\right] = \mathbb{E}\left[\sum_{j=0}^{\infty} \beta_j W_{t+h-j} W_t\right] = \sigma_W^2 \beta_h$$

  to give an indication of the behavior of $\beta(B)$

- Estimate the coefficients of $\beta(B)$ and hence fit an ARMA model for the noise series V_t

Lagged Regression in the Time Domain

The prewhitening step inverts the linear filter $X_t = \frac{\theta_X(B)}{\phi_X(B)} W_t$. Then the lagged regression is between the transformed Y_t and a white series W_t. This makes it easy to determine a suitable lag

Example: In the SOI/recruitment series, we treat SOI as an input, estimate an AR(1) model, prewhiten it, and consider the cross-correlation between the transformed recruitment series and the prewhitened SOI. This shows a large peak at lag -5 (corresponding to the SOI series leading the recruitment series)

This sequential estimation procedure ϕ_X, θ_X, then β, then ϕ_V, θ_V is rather ad hoc. State space methods (ARMAX model) offer an alternative, and they are also convenient for vector-valued input and output series

Lagged Regression in the Frequency Domain: Coherence

To analyze lagged regression in the frequency domain, we'll need the notion of coherence, the analog of cross-correlation in the frequency domain

Define the cross-spectrum as the Fourier transform of the cross-correlation,

$$f_{XY}(\omega) = \sum_{h=-\infty}^{\infty} \gamma_{XY}(h) e^{-2\pi i \omega h},$$

$$\gamma_{XY}(h) = \int_{-\frac{1}{2}}^{\frac{1}{2}} f_{XY}(\omega) e^{2\pi i \omega h}\, d\omega,$$

provided that $\sum_{h=-\infty}^{\infty} |\gamma_{XY}(h)| < \infty$

Notice that $f_{XY}(\omega)$ is complex: $f_{XY}(\omega) = c_{XY}(\omega) - i q_{XY}(\omega)$. Also, $\gamma_{YX}(h) = \gamma_{XY}(-h)$ implies $f_{YX}(\omega) = \overline{f_{XY}(\omega)}$

$$\Rightarrow c_{YX}(\omega) = c_{XY}(\omega) \quad \text{and} \quad q_{YX}(\omega) = -q_{XY}(\omega)$$

Coherence

- The squared coherence function is

$$\rho^2_{Y,X}(\omega) = \frac{|f_{YX}(\omega)|^2}{f_X(\omega) f_Y(\omega)}.$$

  measures the strength of the relationship between X_t and Y_t at frequency ω

- $\rho^2_{Y,X}(\omega)$ is an analog of R^2, it measures the fraction of variance in Y_t at frequency ω, $f_Y(\omega)$, explained by X_t

- $\rho^2_{Y,X}(\omega) = |\rho_{Y,X}(\omega)|^2$, where

$$\rho_{Y,X}(\omega) = \frac{f_{YX}(\omega)}{\sqrt{f_X(\omega) f_Y(\omega)}}$$

Estimating Squared Coherence

Recall that we estimated the spectral density using the smoothed squared modulus of the DFT of the series,

$$\begin{aligned}\bar{f}_X(\omega_j) &= \frac{1}{L}\sum_{k=-(L-1)/2}^{(L-1)/2} |d_X(\omega_j)|^2 \\ &= \frac{1}{L}\sum_{k=-(L-1)/2}^{(L-1)/2} d_X(\omega_{j+k})\overline{d_X(\omega_{j+k})}.\end{aligned}$$

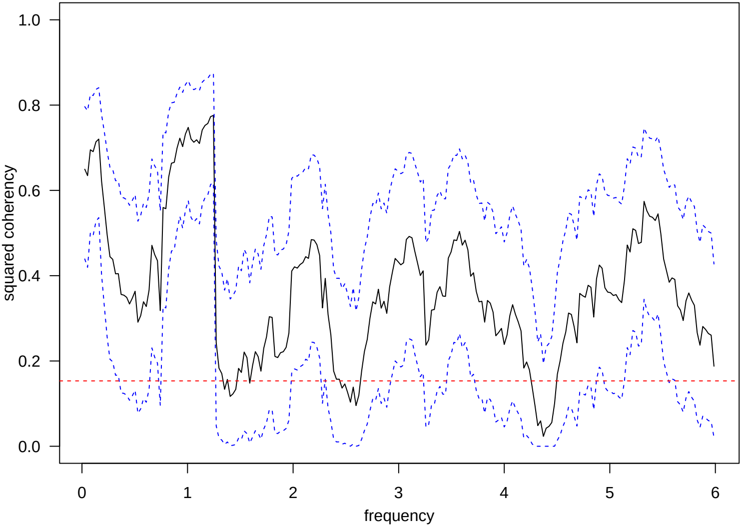
We can estimate the cross spectral density using the same sample estimate,

$$\bar{f}_{XY}(\omega_j) = \frac{1}{L}\sum_{k=-(L-1)/2}^{(L-1)/2} d_X(\omega_{j+k})\overline{d_Y(\omega_{j+k})}$$

Also, we can estimate the squared coherence using these estimates,

$$\bar{\rho}^2_{Y,X}(\omega) = \frac{|\bar{f}_{YX}(\omega)|^2}{\bar{f}_X(\omega)\bar{f}_Y(\omega)}.$$

Estimating Squared Coherence: SOI/Recruitment Example



Recall Lagged Regression Models

$$Y_t = \sum_{j=-\infty}^{\infty} \beta_j X_{t-j} + V_t$$

The projection theorem tells us that the coefficients that minimize the mean squared error,

$$\mathbb{E}\left[\left(Y_t - \sum_{j=-\infty}^{\infty} \beta_j X_{t-j}\right)^2\right]$$

satisfy the orthogonality conditions

$$\mathbb{E}\left[\left(Y_t - \sum_{j=-\infty}^{\infty} \beta_j X_{t-j}\right) X_{t-k}\right] = 0, \quad k = 0, \pm 1, \pm 2, \cdots$$

Taking the expectations inside leads to the normal equations

$$\sum_{j=-\infty}^{\infty} \beta_j \gamma_X(k-j) = \gamma_{YX}(k), \quad k = 0, \pm 1, \pm 2, \cdots$$

Lagged Regression Models in the Frequency Domain

We could solve these equations for the β_j using the sample autocovariance and sample cross-covariance. But it is more convenient to use estimates of the spectra and cross-spectrum because convolution with $\{\beta_j\}$ in the time domain is equivalent to multiplication by the Fourier transform of $\{\beta_j\}$ in the frequency domain

We replace the autocovariance and cross-covariance with the inverse Fourier transforms of the spectral density and cross-spectral density in the orthogonality conditions, i.e., replace

$$\sum_{j=-\infty}^{\infty} \beta_j \gamma_X(k-j) \quad k = 0, \pm 1, \pm 2, \cdots$$

by

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} \sum_{j=-\infty}^{\infty} \beta_j e^{2\pi i \omega (k-j)} f_X(\omega)\, d\omega$$

Lagged Regression Models in the Frequency Domain

This gives, for $k = 0, \pm 1, \pm 2, \cdots$,

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} \sum_{j=-\infty}^{\infty} \beta_j e^{2\pi i\omega(k-j)} f_X(\omega)\, d\omega = \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{2\pi i\omega k} f_{YX}(\omega)\, d\omega,$$

$$\Rightarrow \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{2\pi i\omega k} B(\omega) f_X(\omega)\, d\omega = \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{2\pi i\omega k} f_{YX}(\omega)\, d\omega,$$

where $B(\omega) = \sum_{j=-\infty}^{\infty} e^{-2\pi i\omega j} \beta_j$ is the Fourier transform of the coefficient sequence β_j. Since the Fourier transform is unique, the orthogonality conditions are equivalent to

$$\boxed{B(\omega) f_X(\omega) = f_{YX}(\omega).}$$

Then we may take

$$\hat{B}(\omega_k) = \frac{\hat{f}_{YX}(\omega_k)}{\hat{f}_X(\omega_k)}$$

Lagged Regression Models in the Frequency Domain

We can write the mean squared error at the solution as follows

$$
\begin{aligned}
\mathbb{E}\left[\left(Y_t - \sum_{j=-\infty}^{\infty} \beta_j X_{t-j}\right) Y_t\right] &= \gamma_Y(0) - \sum_{j=-\infty}^{\infty} \beta_j \gamma_{XY}(-j) \\
&= \int_{-\frac{1}{2}}^{\frac{1}{2}} \left(f_Y(\omega) - B(\omega) f_{XY}(\omega)\right) d\omega \\
&= \int_{-\frac{1}{2}}^{\frac{1}{2}} f_Y(\omega) \left(1 - \frac{f_{YX}(\omega) f_{XY}(\omega)}{f_X(\omega) f_Y(\omega)}\right) d\omega \\
&= \int_{-\frac{1}{2}}^{\frac{1}{2}} f_Y(\omega) \left(1 - \frac{|f_{YX}(\omega)|^2}{f_X(\omega) f_Y(\omega)}\right) d\omega \\
&= \int_{-\frac{1}{2}}^{\frac{1}{2}} f_Y(\omega)(1 - \rho_{Y,X}^2(\omega))\, d\omega.
\end{aligned}
$$

$$\Rightarrow \text{MSE} = \int_{-\frac{1}{2}}^{\frac{1}{2}} f_Y(\omega)(1 - \rho_{YX}^2(\omega))\, d\omega$$

$$\Rightarrow f_V(\omega) = \left(1 - \rho_{Y,X}^2(\omega)\right) f_Y(\omega)$$

Lagged Regression Models in the Frequency Domain

$$\text{Recall MSE} = \int_{-\frac{1}{2}}^{\frac{1}{2}} f_Y(\omega)(1 - \rho_{Y,X}^2(\omega))\, d\omega.$$

Thus, $\rho_{Y,X}^2(\omega)$ indicates how the variance of $\{Y_t\}$ at a frequency ω is accounted for by $\{X_t\}$. Compare with the corresponding decomposition for random variables:

$$\mathbb{E}(Y - \beta X) = \sigma_Y^2(1 - \rho_{Y,X}^2)$$

We can estimate the β_j in the frequency domain:

$$\hat{B}(\omega_k) = \frac{\hat{f}_{YX}(\omega_k)}{\hat{f}_X(\omega_k)}.$$

We can approximate the inverse Fourier transform of $\hat{B}(\omega)$,

$$\hat{\beta}_j = \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{2\pi i \omega j} \hat{B}(\omega)\, d\omega$$

via the sum,

$$\hat{\beta}_j = \frac{1}{M} \sum_{k=0}^{M-1} \hat{B}(\omega_j) e^{-2\pi i \omega_k j}.$$

Lagged Regression Models in the Frequency Domain

Here is the procedure:

1. Estimate the spectral density $f_X(\omega)$ and cross-spectral density $f_{YX}(\omega)$

Lagged Regression Models in the Frequency Domain

Here is the procedure:

1. Estimate the spectral density $f_X(\omega)$ and cross-spectral density $f_{YX}(\omega)$
2. Compute the transfer function $\hat{B}(\omega)$:

$$\hat{B}(\omega_k) = \frac{\hat{f}_{YX}(\omega_k)}{\hat{f}_X(\omega_k)}.$$

Lagged Regression Models in the Frequency Domain

Here is the procedure:

1. Estimate the spectral density $f_X(\omega)$ and cross-spectral density $f_{YX}(\omega)$
2. Compute the transfer function $\hat{B}(\omega)$:
$$\hat{B}(\omega_k) = \frac{\hat{f}_{YX}(\omega_k)}{\hat{f}_X(\omega_k)}.$$
3. Take the inverse Fourier transform to obtain the impulse response function β_j:
$$\hat{\beta}_j = \frac{1}{M} \sum_{k=0}^{M-1} \hat{B}(\omega_j) e^{-2\pi i \omega_k j}.$$