STATS 207: Time Series Analysis
Autumn 2020

Lecture 13: Spectral Regression and Principal Components

Dr. Alon Kipnis

October 26th 2020

Genera Info

- HW3 is Due Monday 11/2/2020.

Outline

Frequency-Domain Regression (Multivariate Case)

Spectral Principal Components

Optimum Filtering and Signal Extraction

So Far..

- **Periodogram** – **component of data variance explainable by sinusoids**.
- **Spectral density** – discrete Fourier transform of the covariance function. Gives **typical size** of **periodogram**.
- **Coherence** gives "correlation at frequency j" between two series.
- We **smooth** the **periodogram** to **estimate** the **spectral density** and **coherence**.

Applications:

- Discovering sinusoids in noise.
- Frequency Domain Regression (SOI & Recruitment).

Today:

- Frequency Domain Regression – The Multivariate Case.
- Spectral Principal Components.

Frequency-Domain Regression (Multivariate Case)

Spectral Representation of a Vector Stationary Process (review)

- Example 4.20 (and Property 4.18):
  Consider a **jointly stationary** bivariate process (x_t, y_t). The **autocovariance matrix** is

$$\Gamma(h) = \begin{pmatrix} \gamma_x(h) & \gamma_{xy}(h) \\ \gamma_{yx}(h) & \gamma_y(h) \end{pmatrix},$$

  and the **spectral matrix** is

$$\boldsymbol{f}(\omega) = \begin{pmatrix} f_x(\omega) & f_{xy}(\omega) \\ f_{yx}(\omega) & f_y(\omega) \end{pmatrix}.$$

- **Note**: Hermitian symmetry: $\Gamma^*(\omega) = \Gamma(\omega)$.
- Obvious extensions to **higher dimensions**.

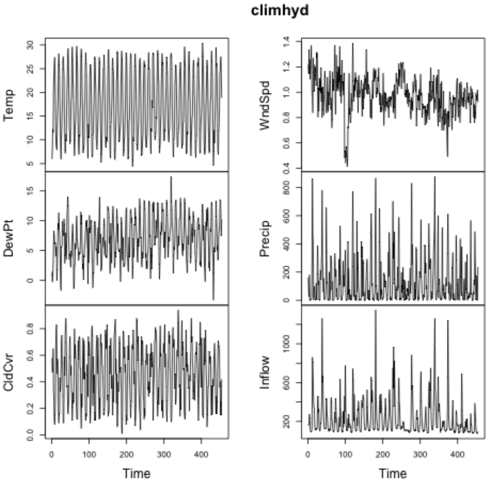
Regression for Jointly Stationary Series

- Suppose $(\boldsymbol{x}_t)$ and (y_t) are **jointly stationary**.
- (y_t) is the **response** series (to be predicted).
- $(\boldsymbol{x}_t)$ is the **predictor** series (covariates, features).
- $(\boldsymbol{x}_t)$ is **vectorial** in general, with q **components**.
- Example 7.1: (y_t) is **monthly** Lake Shasta Inflow; $(\boldsymbol{x}_t)$ has $q = 5$ climate covariates...

```
str(climhyd)
```

```
'data.frame':   454 obs. of  6 variables:
 $ Temp  : num  5.94 8.61 12.28 11.61 20.28 ...
 $ DewPt : num  1.436 -0.285 0.857 2.696 5.7 ...
 $ CldCvr: num  0.58 0.47 0.49 0.65 0.33 0.28 0.11 0.08 0.15 0.27 ...
 $ WndSpd: num  1.22 1.15 1.34 1.15 1.26 ...
 $ Precip: num  160.53 65.79 24.13 178.82 2.29 ...
 $ Inflow: num  156 168 173 273 233 ...
```

Example 7.1: `climhyd` Dataset



(Apply log transform to `Inflow`, `sqrt` transform to `Precip`)

Regression Model

- **Scalar** case (Section 4.8, Lecture 12):

$$y_t = \sum_{r=\infty}^{\infty} \beta_{r1} x_{t-r,1} + v_t$$

  - (y_t) is the response (e.g. inflow).
  - (x_{t1}) single observed input.
  - (β_{r1}) sequence of filter coefficients.
  - (v_t) is a stationary noise uncorrelated with $(\boldsymbol{x}_t)$
- **Vector** case (Section 7.3)

$$\begin{aligned} y_t &= \sum_{j=1}^{q} \sum_{r=-\infty}^{\infty} \beta_{jr} x_{t-r,j} + v_t \\ &= \sum_{r=-\infty}^{\infty} \boldsymbol{\beta}_r' \boldsymbol{x}_{t-r} + v_t \end{aligned}$$

  - $\boldsymbol{x}_t = (x_{t1}, \ldots, x_{tq})$ **vector** of observed inputs.
  - $(\boldsymbol{\beta}_r)$ is a **vector** sequence of coefficients.

Stationarity Assumptions

- $(\boldsymbol{x}_t', y_t)$ stationary $(q+1)$-component vector time series (coordinates are jointly stationary pairwise).
- **Spectral Matrix**:

$$\boldsymbol{f}(\omega) = \begin{pmatrix} f_{x_1}(\omega) & f_{x_1x_2}(\omega) & \cdots & f_{x_1y}(\omega) \\ \vdots & & \ddots & \vdots \\ f_{yx_1}(\omega) & f_{yx_2}(\omega) & \cdots & f_y(\omega). \end{pmatrix}$$

- Absolute summability of cross-covariances

$$\sum_{h=-\infty}^{\infty} |h||\Gamma_{jk}(h)| < \infty, \quad 1 \le j,k \le q+1.$$

- y_t, $\{x_{tk}\}$, $k = 1, \ldots, q$ are linear processes.

Derivation of Frequency-Domain Regression, I

- MSE Criterion

$$\text{MSE}_t = \mathbb{E}\left[\left(y_t - \sum_{r=-\infty}^{\infty} \boldsymbol{\beta}_r' \boldsymbol{x}_{t-r}\right)^2\right], \qquad t \text{ arbitrary, e.g. } 0.$$

- "Orthogonality principle":

$$\mathbb{E}\left[\left(y_t - \sum_{r=-\infty}^{\infty} \boldsymbol{\beta}_r' \boldsymbol{x}_{t-r}\right) \boldsymbol{x}_{t-s}'\right] = 0'$$

$$\Leftrightarrow \qquad \mathbb{E}\left[y_t \boldsymbol{x}_{t-s}'\right] = \mathbb{E}\left[\sum_{r=-\infty}^{\infty} \boldsymbol{\beta}_r' \boldsymbol{x}_{t-r} \boldsymbol{x}_{t-s}'\right]$$

$$\Leftrightarrow \qquad \gamma_{yx}'(s) = \sum_{r=-\infty}^{\infty} \boldsymbol{\beta}_r' \Gamma_x(s-r), \qquad \forall s \in \mathbb{Z}.$$

- These are the **Normal Equations**. Analog of Yule Walker.

Derivation of Frequency-Domain Regression, II

$$\sum_{r=-\infty}^{\infty} \boldsymbol{\beta}_r' \Gamma_x(s-r) = \boldsymbol{\gamma}_{yx}'(s), \qquad \forall s \in \mathbb{Z}.$$

- Applying **Spectral representation** to both sides:

$$\boldsymbol{\gamma}_{yx}'(s) = \int_{-\frac{1}{2}}^{\frac{1}{2}} \boldsymbol{f}_{yx}'(\omega) e^{2\pi i \omega s} d\omega,$$

$$\Gamma_x(u) = \int_{-\frac{1}{2}}^{\frac{1}{2}} \boldsymbol{f}_x(\omega) e^{2\pi i \omega u} d\omega.$$

$$\sum_{r=-\infty}^{\infty} \boldsymbol{\beta}_r' \Gamma_x(s-r) = \int_{-\frac{1}{2}}^{\frac{1}{2}} \overbrace{\sum_{r=-\infty}^{\infty} \boldsymbol{\beta}_r' e^{-2\pi i \omega r}}^{\boldsymbol{B}'(\omega)} \boldsymbol{f}_x(\omega) e^{2\pi i \omega s} d\omega.$$

- From **uniqueness of Fourier coefficients**:

$$\boldsymbol{B}'(\omega) \boldsymbol{f}_x(\omega) = \boldsymbol{f}_{yx}'(\omega) = \boldsymbol{f}_{xy}^*(\omega), \qquad \forall \omega \in (-1/2, 1/2).$$

Derivation of Frequency-Domain Regression, III

- **Frequency response** of **regression filter** (mono-frequency):

$$\boldsymbol{B}'(\omega) = \boldsymbol{f}_x^{-1}(\omega)\boldsymbol{f}_{xy}^*(\omega), \qquad \forall \omega \in (-1/2, 1/2).$$

- **Impulse response** of **regression filter**:

$$\boldsymbol{\beta}_t = \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{2\pi i \omega t} \boldsymbol{B}(\omega) d\omega, \qquad \forall t \in \mathbb{Z}.$$

- **Finite-cutoff regression filter**:

$$\boldsymbol{\beta}_t^M = \frac{1}{M} \sum_{k=0}^{M-1} e^{2\pi i \omega_{k:n} t} \boldsymbol{B}(\omega_{k:n}), \qquad \omega_{k:n} = k/n.$$

- **MSE** relation:

$$\text{MSE} = \text{MSE}_t = \text{MSE}_0 = \mathbb{E}\left[\left(y_0 - \sum_{r=-\infty}^{\infty} \boldsymbol{\beta}_r' \boldsymbol{x}_{-r}\right)^2\right]$$

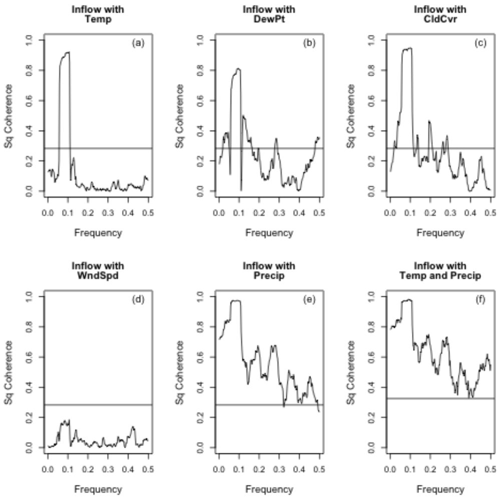
$$= \int_{-\frac{1}{2}}^{\frac{1}{2}} f_y(\omega)\left[1 - \rho_{yx}^2(\omega)\right] d\omega,$$

where

$$\rho_{yx}^2(\omega) = \frac{\boldsymbol{f}_{xy}^*(\omega)\boldsymbol{f}_x^{-1}(\omega)\boldsymbol{f}_{xy}(\omega)}{f_y(\omega)}$$

is the squared **multiple coherence**, which is a scalar function in the range $[0, 1]$.

Example 7.1 `climhyd` Coherencies



F-statistic significant threshold is superposed.

Estimation of Frequency-Domain Regression

- Sample **Spectral Matrix**:

$$\bar{\boldsymbol{f}}_x(\omega_k) = \frac{1}{L} \sum_{\ell=-m}^{m} \boldsymbol{d}_x(\omega_k + \ell/n)\boldsymbol{d}_x^*(\omega_k + \ell/n).$$

- Sample **Cross-Spectrum Vector**:

$$\bar{\boldsymbol{f}}_{xy}(\omega_k) = \frac{1}{L} \sum_{\ell=-m}^{m} \boldsymbol{d}_x(\omega_k + \ell/n)d_y^*(\omega_k + \ell/n).$$

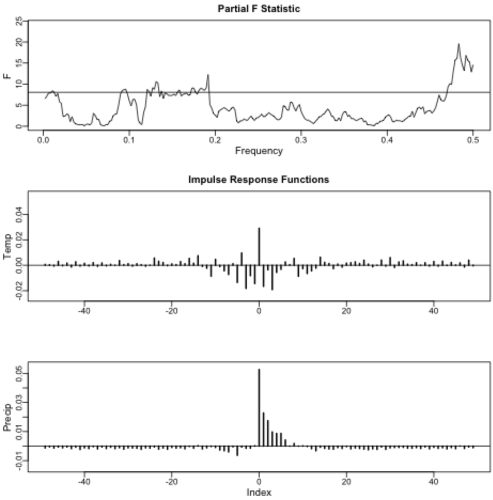
- Sample **Regression Matrix** (mono-frequency):

$$\hat{\boldsymbol{B}}(\omega) = \bar{\boldsymbol{f}}_x^{-1}(\omega)\bar{\boldsymbol{f}}_{xy}^*(\omega).$$

- Sample **Regression Filter** ($L = 2M + 1$):

$$\hat{\boldsymbol{\beta}}_t^M = \frac{1}{M} \sum_{k=0}^{M-1} e^{2\pi i t \omega_{k:n}} \hat{\boldsymbol{B}}(\omega_{k:n}).$$

Example 7.1 `climhyd` Regression on Temperature, Precipitation



F-statistic of temperate 'on top of' precipitation; quantile 0.001.

Comments of F Statistic

Suppose $q = q_1 + q_2$ predictors. Does a model with q terms $\boldsymbol{x}_t = (x_{t1}, \ldots, x_{tq})$ improve over q_1 terms $\boldsymbol{u}_t = (x_{t1}, \ldots, x_{tq_1})$ alone?

- **Full** Model:

$$\bar{\boldsymbol{f}}_{y|x}(\omega) = \bar{f}_y(\omega) - \bar{\boldsymbol{f}}^*_{xy}(\omega)\bar{\boldsymbol{f}}_x^{-1}(\omega)\bar{\boldsymbol{f}}_{xy}(\omega).$$

- **Reduced** Model:

$$\bar{\boldsymbol{f}}_{y|u}(\omega) = \bar{f}_y(\omega) - \bar{\boldsymbol{f}}^*_{uy}(\omega)\bar{\boldsymbol{f}}_u^{-1}(\omega)\bar{\boldsymbol{f}}_{uy}(\omega).$$

- **Power** due to **left out** q_2 terms:

$$\text{SSR}(\omega) = L \cdot \left[\bar{\boldsymbol{f}}_{y|u}(\omega) - \bar{\boldsymbol{f}}_{y|x}(\omega)\right].$$

- **Power** due to **error**:

$$\text{SSE}(\omega) = L \cdot \bar{\boldsymbol{f}}_{y|x}(\omega).$$

- F-statistic:

$$F = \frac{L-q}{q_2}\frac{\text{SSR}(\omega)}{\text{SSE}(\omega)} \quad \text{is asymptotically } F^{2q_2}_{2(L-q)}.$$

Comments of F Statistic (cont'd)

- In Example 4.21 we used

$$\frac{\bar{\rho}^2_{yx}}{1-\bar{\rho}^2_{yx}}(L-1) \quad \text{(asymptotically } F^2_{2(L-1)}\text{)}.$$

to test against $\rho^2_{xy}(\omega) = 0$ (equivalently, $B(\omega) = 0$).

- In Example 7.1 we used

$$\frac{L-q}{q_2}\frac{\text{SSR}(\omega)}{\text{SSE}(\omega)} \quad \text{(asymptotically } F^{2q_2}_{2(L-q)}\text{)}$$

to test whether **a model** with $q = q_1 + q_2$ terms $(x_{t1}, \dots, x_{tq})$ **improves** over a model with q_1 terms $(x_{t1}, \dots, x_{tq_1})$.

- The first case corresponds to $q_1 = 0$ and $q_2 = 1$.

Spectral Principal Components

Classical Principal Components

- $\boldsymbol{x} \sim \mathcal{N}(0, \Sigma), \qquad \Sigma \in \mathbb{R}^{n \times n}$.
- **Maximum variance** of **Linear combination**:

$$\max_{\boldsymbol{c} \neq 0} \frac{\mathrm{Var}(\boldsymbol{c}'\boldsymbol{x})}{\boldsymbol{c}'\boldsymbol{c}}.$$

- **Algebraic** form:

$$\max_{\boldsymbol{c} \neq 0} \frac{\boldsymbol{c}'\Sigma\boldsymbol{c}}{\boldsymbol{c}'\boldsymbol{c}} = \lambda_1(\Sigma) \qquad \text{(top \textbf{eigenvalue} of } \Sigma)$$

$$\boldsymbol{e}_1 \equiv \arg\max_{\boldsymbol{c} \neq 0} \frac{\boldsymbol{c}'\Sigma\boldsymbol{c}}{\boldsymbol{c}'\boldsymbol{c}} \qquad \text{(corresponding \textbf{eigenvector})}.$$

- **Best approximation** to full vector $\boldsymbol{x}$ by **scalar RV**:

$$\min_{Y, \boldsymbol{c} \in \mathbb{R}^n} \mathbb{E}\left[\|\boldsymbol{x} - Y\boldsymbol{c}\|_2^2\right],$$

is attained by $Y = \boldsymbol{e}_1'\boldsymbol{x}$.

Spectral Principal Components

- **Algebraic Form**:

$$\lambda_1(\boldsymbol{f}_x(\omega)) \equiv \max_{\boldsymbol{c}=\boldsymbol{c}(\omega)\neq 0} \frac{\boldsymbol{c}'\boldsymbol{f}_x(\omega)\boldsymbol{c}}{\boldsymbol{c}'\boldsymbol{c}}$$

$$\boldsymbol{e}_1(\omega) \equiv \arg\max_{\boldsymbol{c}=\boldsymbol{c}(\omega)\neq 0} \frac{\boldsymbol{c}'\boldsymbol{f}_x(\omega)\boldsymbol{c}}{\boldsymbol{c}'\boldsymbol{c}}$$

(top **eigenvalue** of $\boldsymbol{f}_x(\omega)$ and corresponding **eigenvector**).

- **Best approximation** to vector time series $(\boldsymbol{x}_t)$ by a **scalar time series** with **deterministic vector filter**

$$\hat{\boldsymbol{x}}_t = \sum_{r=-\infty}^{\infty} y_{t-r}\boldsymbol{c}_r \qquad \text{solving} \qquad \min_{(y_t),(\boldsymbol{c}_t)} \mathbb{E}\left[\|\boldsymbol{x}_t - \hat{\boldsymbol{x}}_t\|_2^2\right],$$

is achieved by

$$\boldsymbol{c}_r^* = \int_{-\frac{1}{2}}^{\frac{1}{2}} \boldsymbol{e}_1(\omega)e^{2\pi i\omega r}\,d\omega.$$

Empirical Spectral Principal Components

- **Empirical Spectral Density**

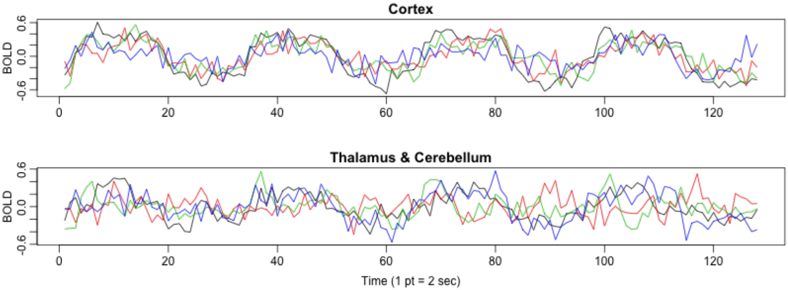
$$\bar{\boldsymbol{f}}_x(\omega_k) = \frac{1}{L}\sum_{\ell=-m}^{m} \boldsymbol{d}_x(\omega_k + \ell/n)\boldsymbol{d}_x^*(\omega_k + \ell/n)$$

- **Algebraic Form**:

$$\lambda_1\left(\bar{\boldsymbol{f}}_x(\omega)\right) \equiv \max_{\boldsymbol{c}=\boldsymbol{c}(\omega)\neq 0} \frac{\boldsymbol{c}'\bar{\boldsymbol{f}}_x(\omega)\boldsymbol{c}}{\boldsymbol{c}'\boldsymbol{c}} \qquad \text{(top } \textbf{eigenvalue}\text{)},$$

$$\boldsymbol{e}_1(\omega) \equiv \arg\max_{\boldsymbol{c}=\boldsymbol{c}(\omega)\neq 0} \frac{\boldsymbol{c}'\boldsymbol{f}_x(\omega)\boldsymbol{c}}{\boldsymbol{c}'\boldsymbol{c}} \qquad \text{(corresponding } \textbf{eigenvector}\text{)}.$$

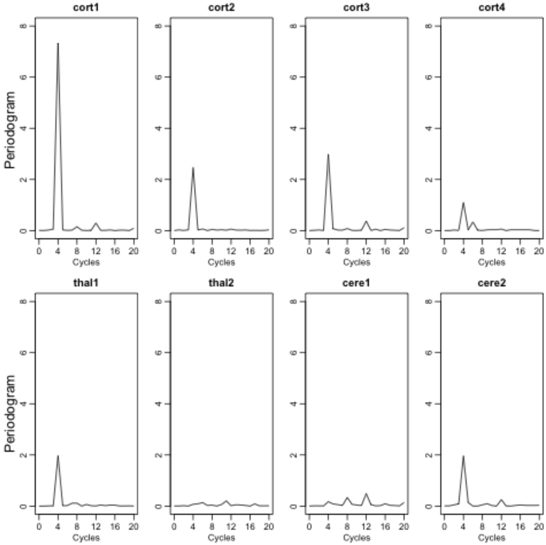
Example 7.14: fMRI Data



- Brain activity from 8 sensors: $\boldsymbol{x}_t = (x_{t1}, \ldots, x_{t8})$, $t = 1, \ldots, 128$.
  - Sensors: $x_{t1}, x_{t2}, x_{t3}, x_{t4}$ in the **cortex**.
  - Sensors: x_{t5}, x_{t6} in the **thalamus**.
  - Sensors : x_{t7}, x_{t8} in the **cerebellum**.
- One sample every two sec for 256 sec.
- **Stimuli**: **brush** on the hand for 32 sec, **stop** for 32 sec.
- **Stimulus Frequency**: $\omega^* \equiv 4/128$ (4 cycles in 128 samples).

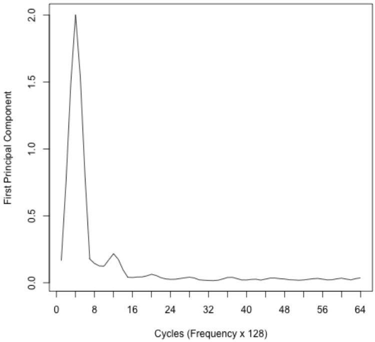
Example 7.14: Individual Periodograms

The individual periodograms of $x_{t1}, \ldots, x_{t8}$:



Example 7.14: Individual Periodograms

The estimated spectral density $\hat{\lambda}_1(j/128)$, $j = 1, \ldots, 128$, for the first **principal component series**:



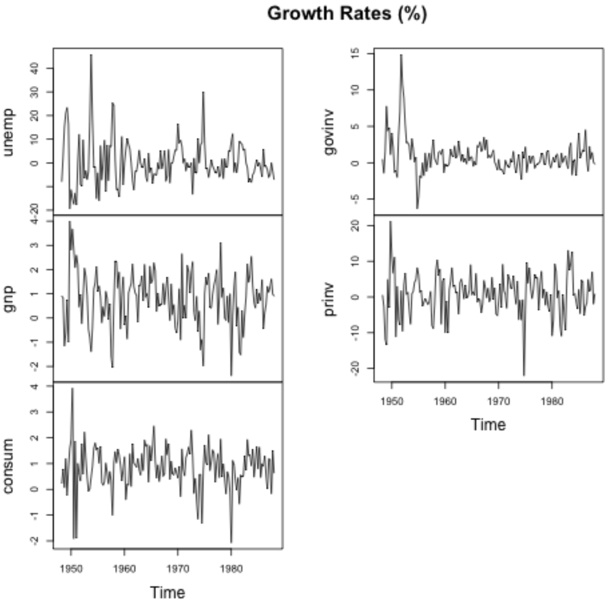
Magnitudes of the PC Vector at the **Stimulus Frequency** (4/128):

Sensor	1	2	3	4	5	6	7	8
$\hat{\boldsymbol{e}}_1(4/128)$	.64	.36	.36	.22	.32	.05	.13	.39

Example 7.14: Macroeconomic Data

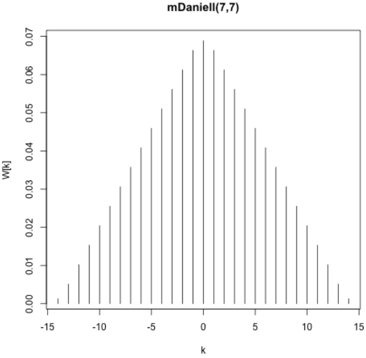
- $x_t = (x_{t1}, \ldots, x_{t5})$. **Growth Rates** of:
  - Unemployment (`unemp`)
  - GNP (`gnp`)
  - Consumption (`consum`)
  - Government Investment (`govinv`)
  - Private Investment (`prinv`)
- **Quarterly** measurements.
- **Interesting Frequencies**:
  - $\omega_{4y} = 0.25$ (once every four years).
  - $\omega_{4y} = 0.125$ (once every eight years)
- Detrended. Standardized (to variance 1).

Example 7.14: Data



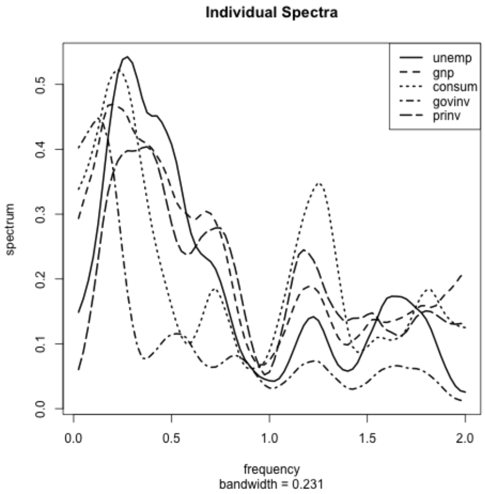
Example 7.14: Periodogram Smoothing Kernel

```
plot(kernel("modified.daniell", L))
```



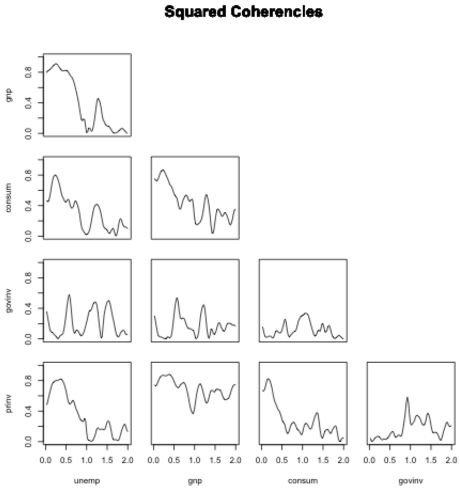
Example 7.14: Individual Spectra

```
plot(gr.spec, log="no",col=1, main="Individual Spectra", lty=1:5, lwd=2)
legend("topright", colnames(econ5), lty=1:5, lwd=2)
```



Example 7.14: Coherencies

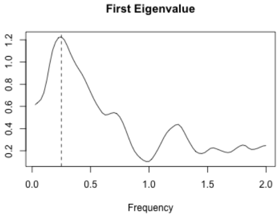
```
plot.spec.coherency(gr.spec, ci=NA,  main="Squared Coherencies")
```

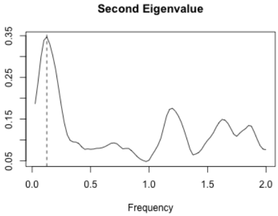


Example 7.14: Principal Components

```
# PCs
n.freq = length(gr.spec$freq)
lam = matrix(0,n.freq,5)
for (k in 1:n.freq) lam[k,] = eigen(gr.spec$fxx[,,k], symmetric=TRUE, only.

par(mfrow=c(2,1), mar=c(4,2,2,1), mgp=c(1.6,.6,0))
plot(gr.spec$freq, lam[,1], type="l", ylab="", xlab="Frequency", main="Firs
abline(v=.25, lty=2)
plot(gr.spec$freq, lam[,2], type="l", ylab="", xlab="Frequency", main="Seco
abline(v=.125, lty=2)
e.vec1 = eigen(gr.spec$fxx[,,10], symmetric=TRUE)$vectors[,1]
e.vec2 =  eigen(gr.spec$fxx[,,5], symmetric=TRUE)$vectors[,2]
round(Mod(e.vec1), 2);  round(Mod(e.vec2), 3)
```





Optimum Filtering and Signal Extraction

Optimum Filtering (Wiener-Kolmogorov Problem)

- Model:

$$y_t = \sum_{r=-\infty}^{\infty} \beta_r x_{t-r} + v_t,$$

  where

  - (v_t) is a **noise process** uncorrelated of (x_t).
  - (β_r) are **known**.
  - (y_t) is **observed**. (x_t) is **not observed**.

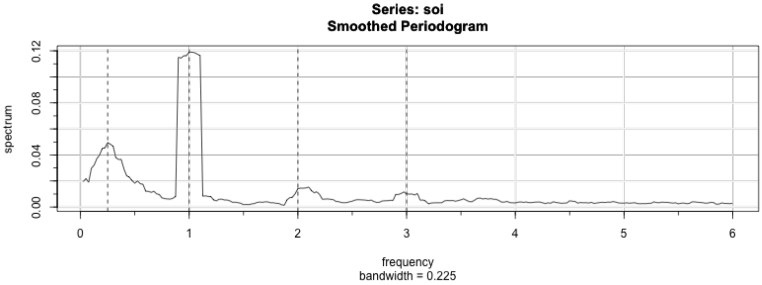
- **Goal**: Find an estimator for (x_t) of the form

$$\hat{x}_t = \sum_{r=-\infty}^{\infty} a_r y_{t-r}.$$

- Solved by Norbert Wiener and Andrey Kolmogorov in the 1940's.

Example 4.25: Estimating El-Ninõ Signal

Spectrum of SOI:



Frequency-Domain Approach

- **Orthogonality principle**: optimum (a_r) satisfies

$$\mathbb{E}\left[\left(x_t - \sum_{r=-\infty}^{\infty} a_r y_{t-r}\right) y_{t-s}\right] = 0 \Rightarrow \sum_{r=-\infty}^{\infty} a_r \gamma_y(s-r) = \gamma_{xy}(s)$$

  for all $s = 0, \pm 1, \pm 2, \ldots$.

- Use **spectral representation** and properties of **linear filters**:

$$A(\omega) f_y(\omega) = f_{xy}(\omega)$$

$$A(\omega)\left(|B(\omega)|^2 f_x(\omega) + f_v(\omega)\right) = B^*(\omega) f_x(\omega)$$

  where: $A(\omega)$, $B(\omega)$ are the frequency responses of (a_r), (β_r).

- **Optimal** filter's frequency response:

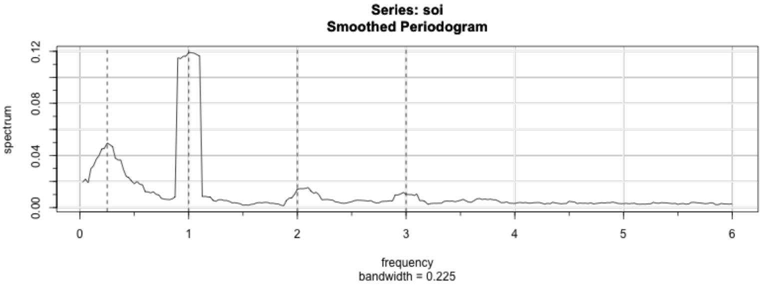
$$A(\omega) = \frac{B^*(\omega)}{|B(\omega)|^2 f_x(\omega) + f_v(\omega)}$$

- **Minimized** MSE

$$MSE^* = \int_{-\frac{1}{2}}^{\frac{1}{2}} \left[1 - \frac{|B(\omega)|^2}{|B(\omega)|^2 f_x(\omega) + f_v(\omega)}\right] f_x(\omega) d\omega$$

Example 4.25: Estimating El-Ninõ Signal

- Spectrum of SOI:



- Assume:

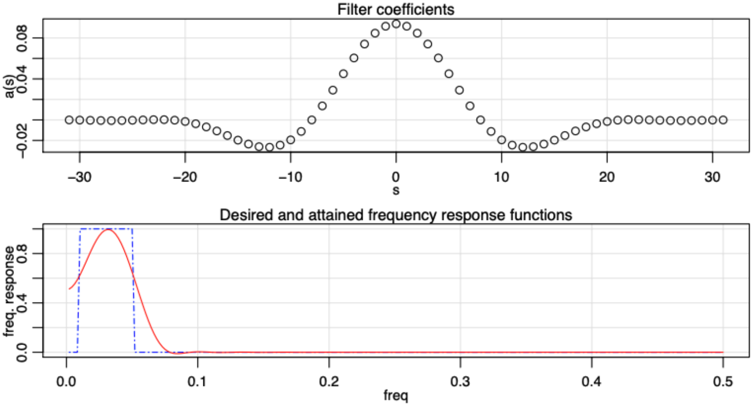
$$SOI_t = x_t + v_t,$$

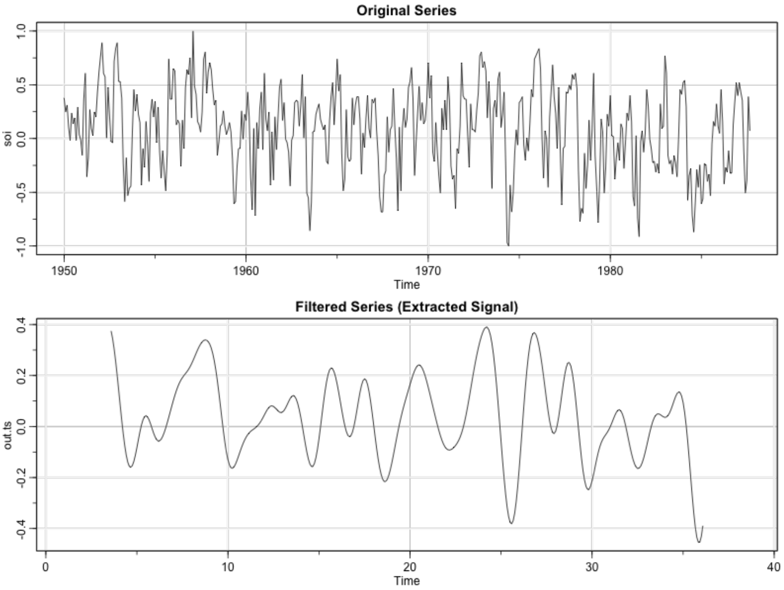
where

$$f_x(\omega) = \begin{cases} f_{SOI}(\omega) & \omega < 1/7 \\ 0 & \omega > 1/7 \end{cases}$$

Note: $B(\omega) = 1$ or $\beta_t = \delta_t$.

Example 4.25: Extracting El-Ninõ Signal (cont'd)




Original Series
soi
1.0
0.5
0.0
-0.5
-1.0
1950
1960
1970
1980
Time
Filtered Series (Extracted Signal)
out.ts
0.4
0.2
0.0
-0.2
-0.4
0
10
20
30
40
Time