D.C. Economists Minicourse –
What's New in Econometrics: Time Series

Lecture 1

September 26, 2008

Frequency Domain Descriptive Statistics

Outline

0. Introduction to Course
1. Time Series Basics
2. Spectral representation of stationary process
3. Spectral Properties of Filters
   (a) Band Pass Filters
   (b) One-Sided Filters
4. Multivariate spectra
5. Spectral Estimation (a few words – more in Lecture 5)

Introduction to Course

Some themes that have occupied time-series econometricians and empirical macroeconomists in the last decade or so:

1. Low-frequency variability: (i) unit roots, near unit roots, cointegration, fractional models, and so forth; (ii) time varying parameters; (iii) stochastic volatility; (iv) HAC covariance matrices; (v) long-run identification in SVAR

2. Identification: methods for dealing with "weak" identification in linear models (linear IV, SVAR) and nonlinear models (GMM).

3. Forecasting: (i) inference procedures for relative forecast performance of existing models; (ii) potential improvements in forecasts from using many predictors

Some tools used:

1. VARs, spectra, filters, GMM, asymptotic approximations from CLT and LLN.

2. Functional CLT

3. Simulation Methods (MCMC, Bootstrap)

We will talk about these themes and tools. This will not be a comprehensive literature review. Our goal is present some key ideas as simply as possible.

Topics

1. Spectral preliminaries and applications, linear filtering theory (MW)
2. Functional central limit theory and structural breaks (MW)
3. Many instruments/weak identification in GMM I (JS)
4. Many instruments/weak identification in GMM II (JS)
5. Heteroskedasticity- and autocorrelation consistent (HAC) standard errors (MW)
6. The Kalman filter, nonlinear filtering, and Markov Chain Monte Carlo (MW)
7. Recent developments in structural VAR modeling (JS)
8. Econometrics of DSGE models (JS)
9. Stochastic time variation and low-frequency models (MW)
10. Forecast assessment (MW)
11. Dynamic factor models and forecasting with many predictors (JS)
12. Macro modeling with many predictors (JS)

Time Series Basics (and notation)

(References: Hayashi (2000), Hamilton (1994), … , lots of other books)

1. $\{Y_t\}$: a sequence of random variables

2. Stochastic Process: The probability law governing $\{Y_t\}$

3. Realization: One draw from the process, $\{y_t\}$

4. Strict Stationarity: The process is strictly stationary if the probability distribution of $(Y_t, Y_{t+1}, ..., Y_{t+k})$ is identical to the probability distribution of $(Y_\tau, Y_{\tau+1}, ..., Y_{\tau+k})$ for all t, τ, and k. (Thus, all joint distributions are time invariant.)

5. Autocovariances: $\gamma_{t,k} = cov(Y_t, Y_{t+k}^{'})$

6. Autocorrelations: $\rho_{t,k} = cor(Y_t, Y_{t+k}^{'})$

7. Covariance Stationarity: The process is covariance stationary if $\mu_t = E(Y_t) = \mu$ and $\gamma_{t,k} = \gamma_k$ for all t and k.

8. White noise: A process is called white noise if it is covariance stationary and $\mu = 0$ and $\gamma_k = 0$ for $k \neq 0$.

9. Martingale: Y_t follows a martingale process if $E(Y_{t+1} \mid \boldsymbol{F}_t) = Y_t$, where $\boldsymbol{F}_t \subseteq \boldsymbol{F}_{t+1}$ is the time t information set.

10. Martingale Difference Process: Y_t follows a martingale difference process if $E(Y_{t+1} \mid \boldsymbol{F}_t) = 0$. $\{Y_t\}$ is called a martingale difference sequence or "mds."

11. The Lag Operator: L lags the elements of a sequence by one period. $Ly_t = y_{t-1}$, $L^2 y_t = y_{t-2}$,. If b denotes a constant, then $bLY_t = L(bY_t) = bY_{t-1}$.

12. Linear filter: Let $\{c_j\}$ denote a sequence of constants and

$c(L) = c_{-r}L^{-r} + c_{-r+1}L^{-r+1} + \ldots + c_0 + c_1 L + \ldots + c_s L^s$

denote a polynomial in L. Note that $X_t = c(L)Y_t = \sum_{j=-r}^{s} c_j Y_{t-j}$ is a moving average of Y_t. $c(L)$ is sometimes called a linear filter (for reasons discussed below) and X is called a filtered version of Y.

13. AR(p) process: $\phi(\mathrm{L})Y_t = \varepsilon_t$ where $\phi(\mathrm{L}) = (1 - \phi_1 L - \ldots - \phi_p L^p)$ and ε_t is white noise.

14. MA(q) process: $Y_t = \theta(\mathrm{L})\varepsilon_t$ where $\theta(\mathrm{L}) = (1 - \theta_1 L - \ldots - \theta_q L^q)$ and ε_t is white noise.

15. ARMA(p,q): $\phi(L)Y_t = \theta(L)\varepsilon_t$.

16. Wold decomposition theorem (e.g., Brockwell and Davis (1991))
Suppose Y_t is generated by a linearly indeterministic covariance stationary process. Then Y_t can be represented as

$$Y_t = \varepsilon_t + c_1\varepsilon_{t-1} + c_2\varepsilon_{t-2} + \ldots ,$$

where ε_t is white noise with variance σ_ε^2, $\sum_{i=1}^{\infty} c_i^2 < \infty$, and

$\varepsilon_t = Y_t - Proj(Y_t \mid \text{lags of } Y_t)$ (so that ε_t is "fundamental").

17. Spectral Representation Theorem(e.g, Brockwell and Davis (1991)): Suppose Y_t is a discrete time covariance stationary zero mean process, then there exists an orthogonal-increment process $Z(\omega)$ such that

(i) $\mathrm{Var}(Z(\omega)) = F(\omega)$

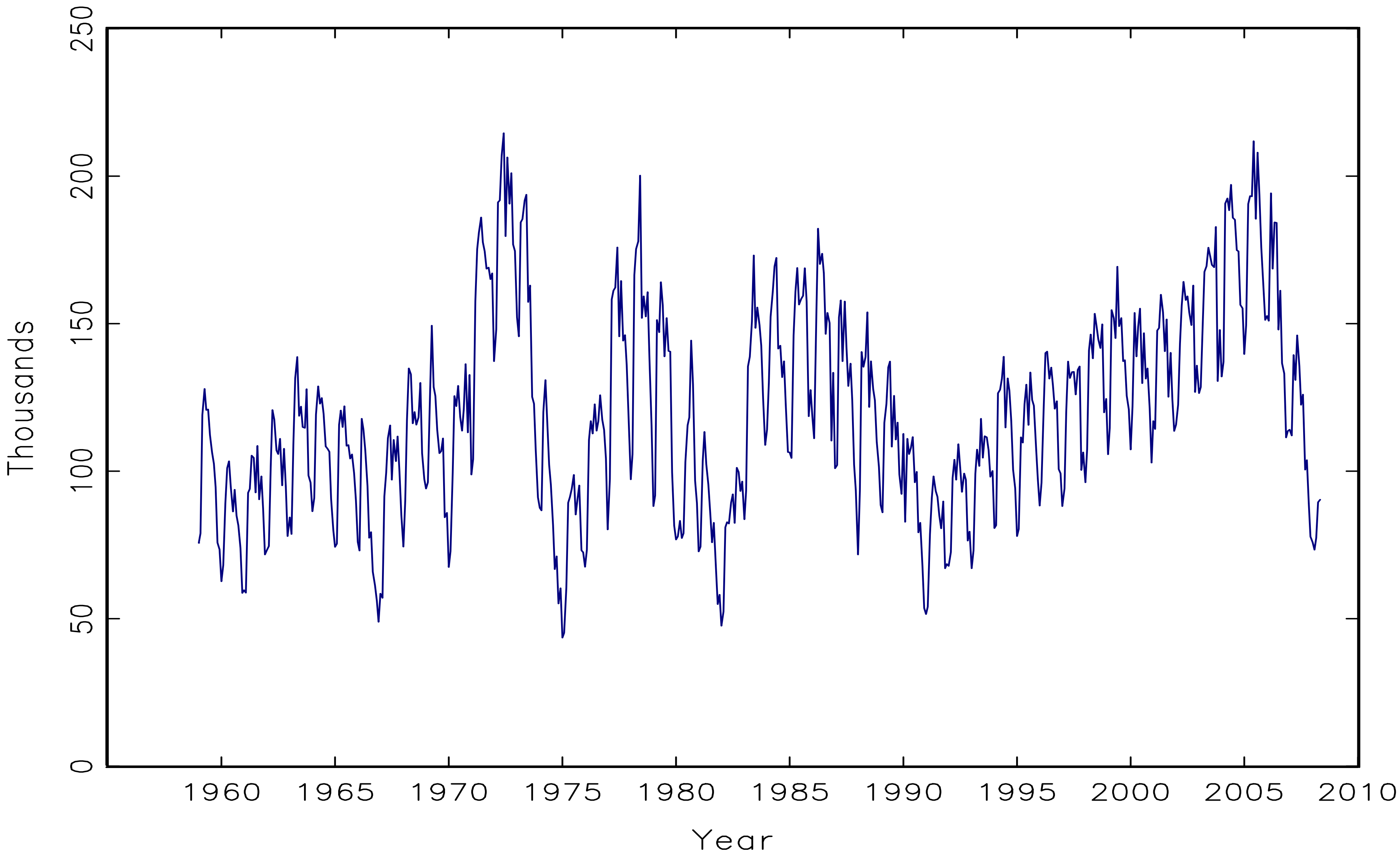
and

(ii) $Y_t = \int_{-\pi}^{\pi} e^{it\omega} dZ(\omega)$

where F is the spectral distribution function of the process. (The spectral density, $S(\omega)$, is the density associated with F.)

This is a useful and important decomposition, and we'll spend some time discussing it.

U.S. Residential Building Permits: 1959:1-2008:5



Some questions

1. How important are the "seasonal" or "business cycle" components in Y_t?

2. Can we measure the variability at a particular frequency? Frequency 0 (long-run) will be particularly important as that is what HAC Covariance matrices are all about.

3. Can we isolate/eliminate the "seasonal" ("business-cycle") component?

4. Can we estimate the business cycle or "gap" component in real time? If so, how accurate is our estimate?

1. Spectral representation of a covariance stationary stochastic process

Deterministic process:

(a) $Y_t = \cos(\omega t)$, strictly periodic with $period = \frac{2\pi}{\omega}$, $Y_0 = 1$, amplitude = 1.

(b) $Y_t = a \times \cos(\omega t) + b \times \sin(\omega t)$, strictly period with $period = \frac{2\pi}{\omega}$, $Y_0 = a$, amplitude = $\sqrt{a^2 + b^2}$

Stochastic Process:

$Y_t = a \times \cos(\omega t) + b \times \sin(\omega t)$, a and b are random variables, 0-mean, mutually uncorrelated, with common variance σ^2.

2^{nd} - moments :

$\mathrm{E}(Y_t) = 0$

$\mathrm{Var}(Y_t) = \sigma^2 \times \{\cos^2(\omega t) + \sin^2(\omega t)\} = \sigma^2$

$$Cov(Y_t, Y_{t-k}) = \sigma^2 \{\cos(\omega t)\cos(\omega(t-k)) + \sin(\omega t)\sin(\omega(t-k))\}$$
$$= \sigma^2 \cos(\omega k)$$

Stochastic Process with more components:

$$Y_t = \sum_{j=1}^{n} \{a_j \cos(\omega_j t) + b_j \sin(\omega_j t)\},$$ $\{a_j, b_j\}$ are uncorrelated 0-mean random variables, with $\mathrm{Var}(a_j) = \mathrm{Var}(b_j) = \sigma_j^2$

2^{nd} - moments :

$\mathrm{E}(Y_t) = 0$

$\mathrm{Var}(Y_t) = \sum_{j=1}^{n} \sigma_j^2$ (Decomposition of variance)

$\mathrm{Cov}(Y_t Y_{t-k}) = \sum_{j=1}^{n} \sigma_j^2 \cos(\omega_j k)$ (Decomposition of auto-covariances)

Stochastic Process with even more components:

$$Y_t = \int_0^{\pi} \cos(\omega t) da(\omega) + \int_0^{\pi} \sin(\omega t) db(\omega)$$

$da(\omega)$ and $db(\omega)$: random variables, 0-mean, mutually uncorrelated, uncorrelated across frequency, with common variance that depends on frequency. This variance function is called the spectrum.

A convenient change of notation:

$$Y_t = a\times\cos(\omega t) + b\times\sin(\omega t)$$

$$= \frac{1}{2}e^{i\omega}(a - ib) + \frac{1}{2}e^{-i\omega}(a + ib)$$

$$= e^{i\omega}Z + e^{-i\omega}\overline{Z}$$

where $i = \sqrt{-1}$ and $e^{i\omega} = cos(\omega) + i\times sin(\omega)$, $Z = \frac{1}{2}(a - ib)$ and $\overline{Z}$ is the complex conjugate of Z.

Similarly

$$Y_t = \int_0^{\pi} \cos(\omega t)da(\omega) + \int_0^{\pi} \sin(\omega t)db(\omega)$$

$$= \frac{1}{2}\int_0^{\pi} e^{i\omega t}(da(\omega) - idb(\omega)) + \frac{1}{2}\int_0^{\pi} e^{-i\omega t}(da(\omega) + idb(\omega))$$

$$= \int_{-\pi}^{\pi} e^{i\omega t} dZ(\omega)$$

where $dZ(\omega) = \frac{1}{2}(da(\omega) - idb(\omega))$ for $\omega \geq 0$ and $dZ(-\omega) = \overline{dZ(\omega)}$ for $\omega > 0$.

Because *da* and *db* have mean zero, so does *dZ*. Denote the variance of $dZ(\omega)$ as $\text{Var}(dZ(\omega)) = \text{E}(dZ(\omega)\overline{dZ(\omega)}) = S(\omega)d\omega$, and using the assumption that *da* and *db* are uncorrelated across frequency $\text{E}(dZ(\omega)\overline{dZ(\omega)'}) = 0$ for $\omega \neq \omega'$.

Second moments of Y:

$$E(Y_t) = E\left\{\int_{-\pi}^{\pi} e^{i\omega t} dZ(\omega)\right\} = \int_{-\pi}^{\pi} e^{i\omega t} E(dZ(\omega)) = 0$$

$$\gamma_k = E(Y_t Y_{t-k}) = E(Y_t \overline{Y}_{t-k}) = E\left\{\int_{-\pi}^{\pi} e^{i\omega t} dZ(\omega) \int_{-\pi}^{\pi} e^{-i\omega(t-k)} \overline{dZ(\omega)}\right\}$$

$$= \int_{-\pi}^{\pi} e^{i\omega t} e^{-i\omega(t-k)} E(dZ(\omega)\overline{dZ(\omega)})$$

$$= \int_{-\pi}^{\pi} e^{i\omega k} S(\omega) d\omega$$

Setting $k = 0$, $\gamma_0 = \text{Var}(Y_t) = \int_{-\pi}^{\pi} S(\omega) d\omega$

Summarizing

1. $S(\omega)d\omega$ can be interpreted as the variance of the cyclical component of Y corresponding to the frequency ω. The period of this component is $period = \dfrac{2\pi}{\omega}$.

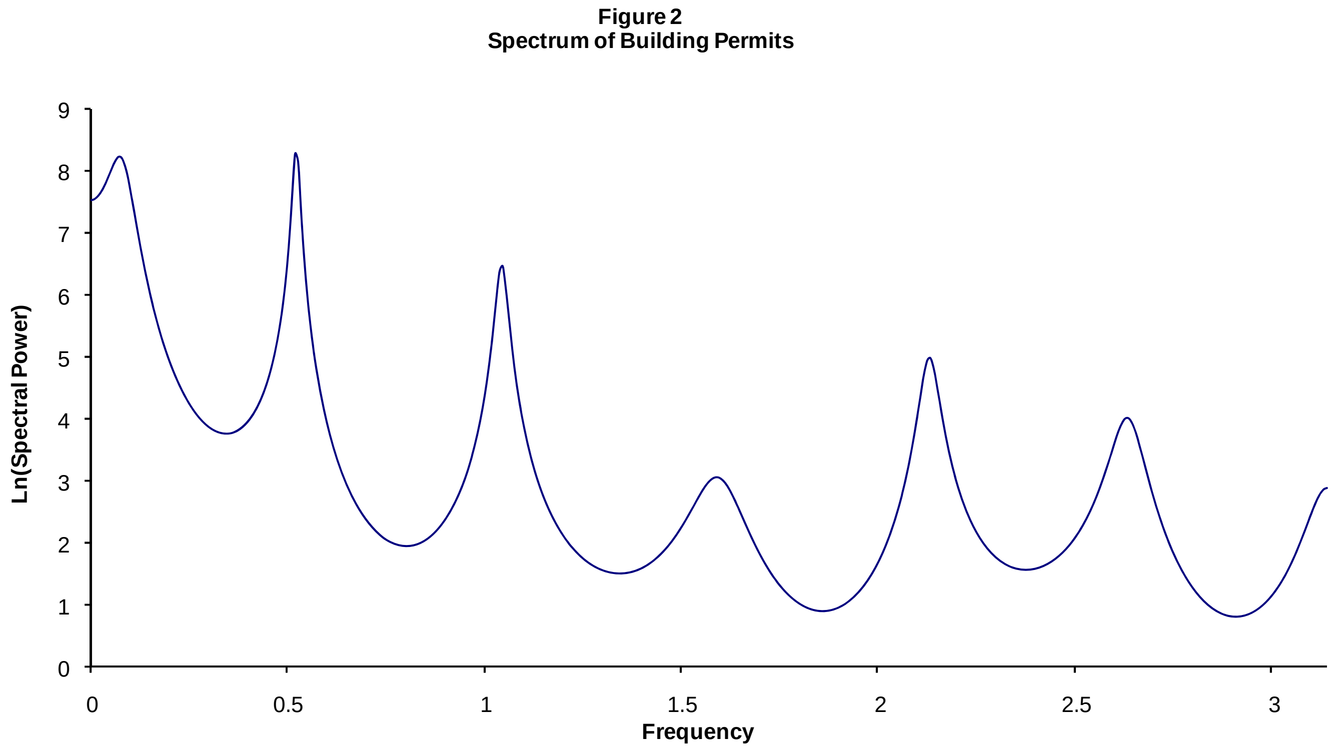
2. $S(\omega) \geq 0$ (it is a variance)

3. $S(\omega) = S(-\omega)$. Because of this symmetry, plots of the spectrum are presented for frequencies $0 \leq \omega \leq \pi$.

4. $\gamma_k = \int_{-\pi}^{\pi} e^{i\omega k} S(\omega) d\omega$ can be inverted to yield

$$S(\omega) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} e^{-i\omega k} \gamma_k = \frac{1}{2\pi} \left\{ \gamma_0 + 2 \sum_{k=1}^{\infty} \gamma_k \cos(\omega k) \right\}$$

The Spectrum of Building Permits



Most of the mass in the spectrum is concentrated around the seven peaks evident in the plot. (These peaks are sufficiently large that spectrum is plotted on a log scale.) The first peak occurs at frequency $\omega = 0.07$ corresponding to a period of 90 months. The other peaks occur at frequencies $2\pi/12$, $4\pi/12$, $6\pi/12$, $8\pi/12$, $10\pi/12$, and π. These are peaks for the seasonal frequencies: the first corresponds to a period of 12 months, and the others are the seasonal "harmonics" 6, 4, 3, 2.4, 2 months. (These harmonics are necessary to reproduce an arbitrary − not necessary sinesoidal − seasonal pattern.)

"Long-Run Variance"

The long-run variance is $S(0)$, the variance of the 0-frequency (or ∞ period component). Since $S(\omega)=\frac{1}{2\pi}\sum_{k=-\infty}^{\infty}e^{-i\omega k}\gamma_k$, then $S(0) = \frac{1}{2\pi}\sum_{k=-\infty}^{\infty}\gamma_k$. As we will see in lecture 5, this plays an important role in statistical inference because (except for the factor 2π) it is the large-sample variance of the sample mean.

2. Spectral Properties of Filters (Moving Averages)

Let $x_t = c(L)y_t$, where $c(L) = c_{-r}L^{-r} + \ldots + c_sL^s$, so that x is a moving average of y with weights given by the c's.

How does $c(L)$ change the cyclical properties of y? To study this, suppose that y is strictly periodic

$$y_t = 2\cos(\omega t) = e^{i\omega t} + e^{-i\omega t}$$

with period $p = \dfrac{2\pi}{\omega}$.

A simple representation for x follows:

$$x_t = \sum_{j=-r}^{s} c_j y_{t-j} = \sum_{j=-r}^{s} c_j [e^{i\omega(t-j)} + e^{-i\omega(t-j)}]$$

$$= e^{i\omega t} \sum_{j=-r}^{s} c_j e^{-i\omega j} + e^{-i\omega t} \sum_{j=-r}^{s} c_j e^{i\omega j} = e^{i\omega t} c(e^{-i\omega}) + e^{-i\omega t} c(e^{i\omega})$$

$c(e^{i\omega})$ is a complex number, say $c(e^{i\omega}) = a + ib$, where $a = Re[c(e^{i\omega})]$ and $b = Im[c(e^{i\omega})]$. Write this number in polar form as

$$c\left(e^{i\omega}\right) = (a^2 + b^2)^{\frac{1}{2}}[\cos(\theta) + i\sin(\theta)] = ge^{i\theta}$$

where $g = (a^2 + b^2)^{\frac{1}{2}} = [c(e^{i\omega})c(e^{-i\omega})]^{\frac{1}{2}}$ and $\theta = \tan^{-1}\left(\frac{b}{a}\right) = \tan^{-1}\left(\frac{Im[c(e^{i\omega})]}{Re[c(e^{i\omega})]}\right)$

Thus

$$x_t = e^{i\omega t} g e^{-i\theta} + e^{-i\omega t} g e^{i\theta}$$
$$= g[e^{i\omega[t-\frac{\theta}{\omega}]} + e^{-i\omega[t-\frac{\theta}{\omega}]}]$$
$$= 2g\cos\left(\omega\left(t-\frac{\theta}{\omega}\right)\right)$$

So that the filter $c(L)$ "amplifies" y by the factor g and shifts y back in time by $\frac{\theta}{\omega}$ time units.

- Note that g and θ depend on ω, and so it makes since to write them as $g(\omega)$ and $\theta(\omega)$.
- $g(\omega)$ is called the filter *gain* (or sometimes the *amplitude gain*).
- $\theta(\omega)$ is called the filter *phase*.
- $g(\omega)^2 = c(e^{i\omega})c(e^{-i\omega})$ is called the *power transfer function* of the filter.

Examples

1. $c(L) = L^2$

$c(e^{i\omega}) = e^{2i\omega} = \cos(2\omega) + i\sin(2\omega)$

so that

$$\theta(\omega) = \tan^{-1}[\frac{\sin 2\omega}{\cos 2\omega}] = 2\omega$$

and $\dfrac{\theta}{\omega} = 2$ time periods. Also $g(\omega) = |c(e^{i\omega})| = 1$.

2. (Sargent (1979)) Kuznets Filter for annual data: Let

$$a(L) = (1/5)(L^{-2} + L^{-1} + L^{0} + L^{1} + L^{2})$$

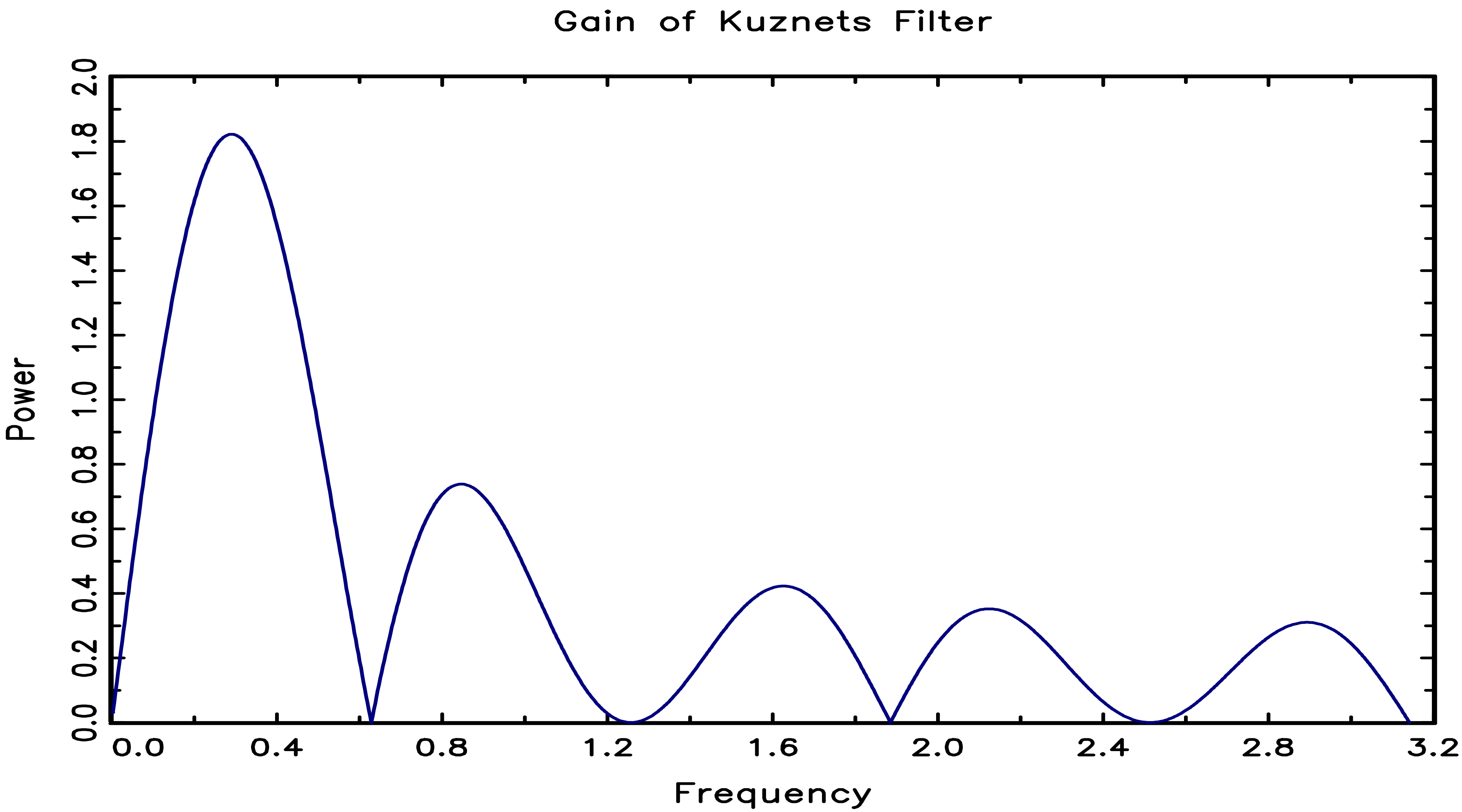
(which ”smooths” the series) and

$$b(L) = (L^{-5} - L^{5})$$

(which forms centered ten-year differences), then the Kuznets filter is

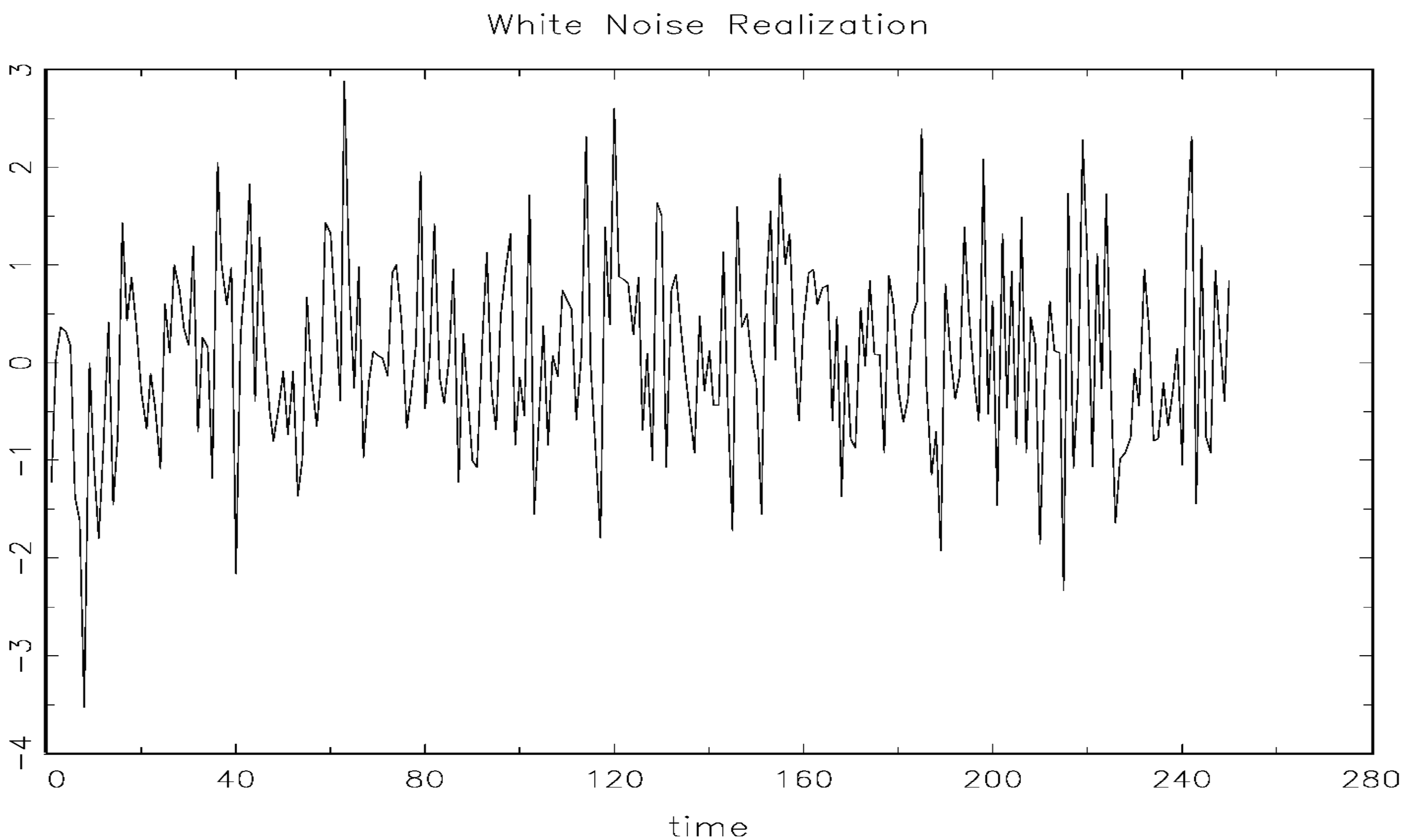
$$c(L) = b(L)a(L)$$

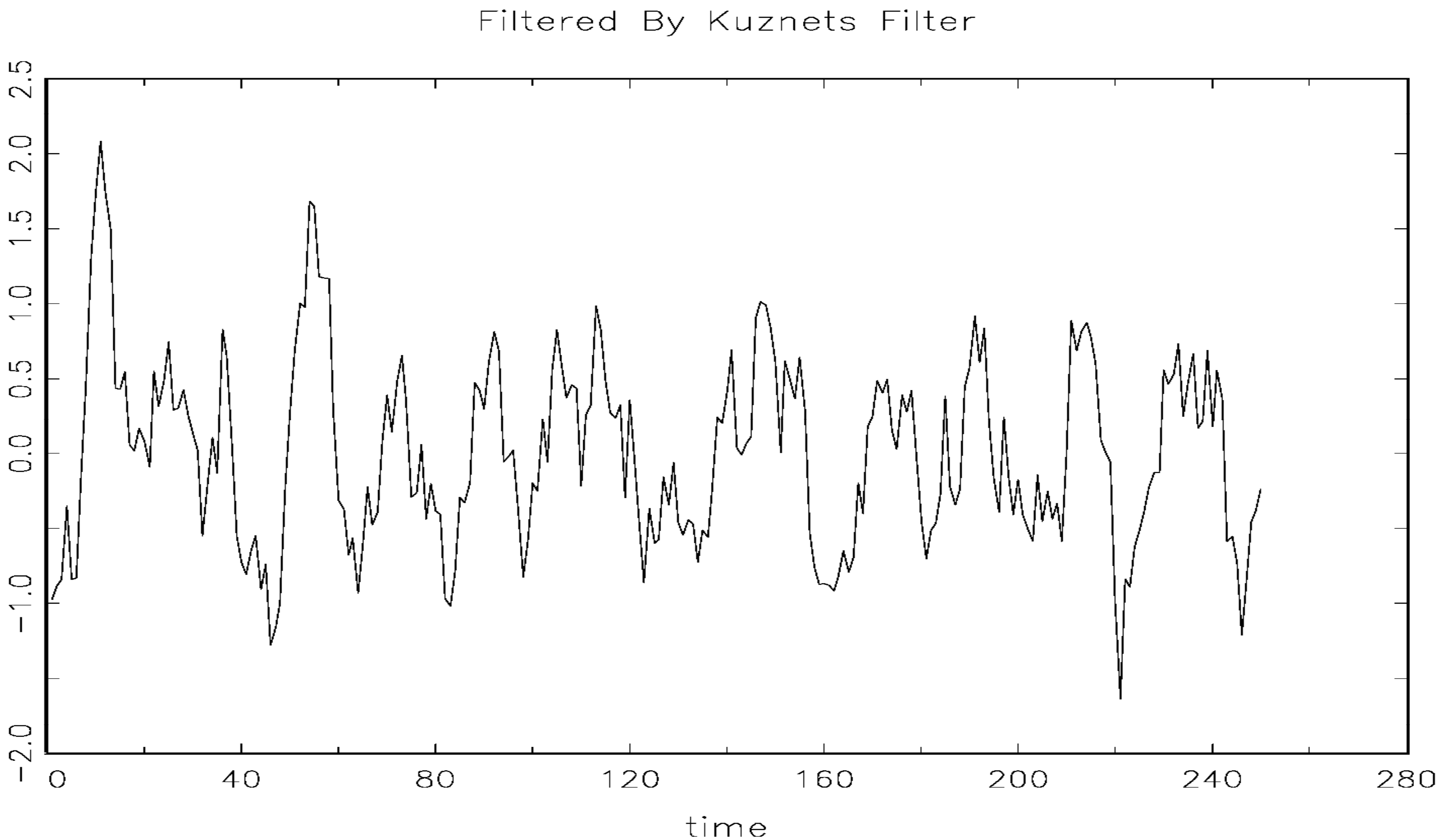
$g(\omega) = |c(e^{i\omega})| = |b(e^{i\omega})|\,|a(e^{i\omega})|$, which are easily computed for a grid of values using Gauss, Matlab, Excel, etc.



Peak at $\omega \approx 0.30$

Period = $2\pi/0.30 \approx 21$ years.


White Noise Realization
3
2
1
0
−1
−2
−3
−4
0
40
80
120
160
200
240
280
time


Filtered By Kuznets Filter
2.5
2.0
1.5
1.0
0.5
0.0
-1.0
-2.0
0
40
80
120
160
200
240
280
time

Example 3: Census X-11 Seasonal Adjustment

The linear operations in X-11 can be summarized by the filter $x_t^{sa} = X11(L)x_t$, $X11$(L) is a 2-sided filter constructed in 8-steps (Young (1968) and Wallis (1974).

8-steps:

X11-1. Form an initial estimate of TC as $\widehat{TC}_t^1 = A_1(L)x_t$, where $A_1(L)$ is the centered 12-month moving average filter $A_1(L) = \sum_{j=-6}^{6} b_j L^j$, with $b_{|6|} = 1/24$, and $b_j = 1/12$, for $-5 \le j \le 5$.

X11-2. Form an initial estimate of $S+I$ as $\widehat{SI}_t^1 = x_t - \widehat{TC}_t^1$

X11-3. Form an initial estimate of S_t as $\hat{S}_t^1 = S_1(L^{12})\widehat{SI}_t^1$, where $S_1(L^{12}) = \sum_{j=-2}^{2} c_j L^{12j}$, and where c_j are weights from a 3×3 moving average (*i.e.,* $1/9, 2/9, 3/9, 2/9, 1/9$).

X11-4. Adjust the estimates of S so that they add to zero (approximately) over any 12 month period as $\hat{S}_t^2 = S_2(L)\hat{S}_t^1$, where $S_2(L) = 1 - A_1(L)$, where $A_1(L)$ is defined in step 1.

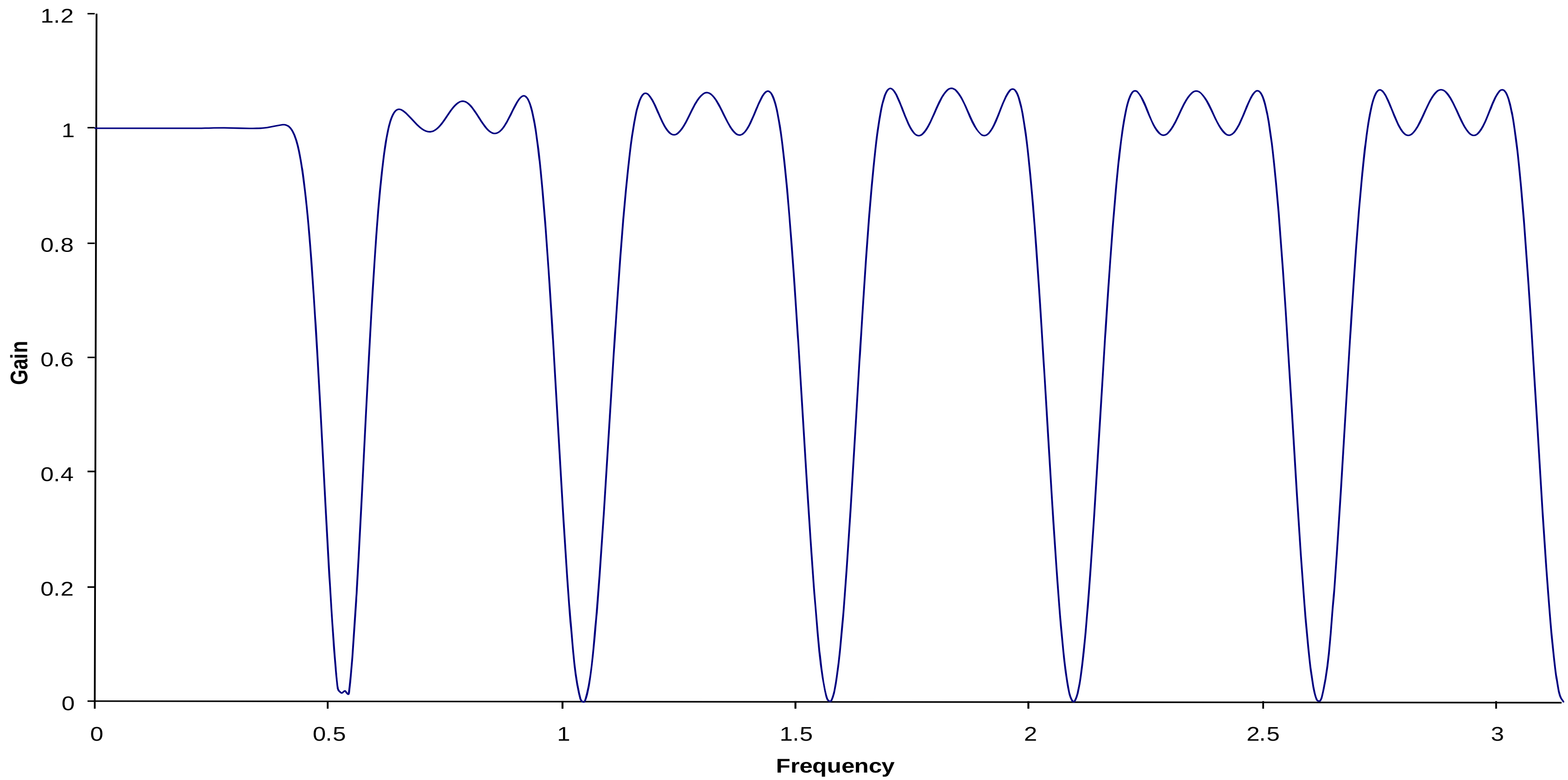
X11-5. Form a second estimate of TC as $\widehat{TC}_t^2 = A_2(L)(x_t^1 - \hat{S}_t^2)$, where $A_2(L)$ denotes a "Henderson" moving average filter. (The 13-term Henderson moving average filter is given by $A_2(L) = \sum_{i=-6}^{6} A_{2,i} L^i$, with $A_{2,0} = .2402$, $A_{2,|1|} = .2143$, $A_{2,|2|} = .1474$, $A_{2,|3|} = .0655$, $A_{2,|4|} = 0$, $A_{2,|5|} = -.0279$, $A_{2,|6|} = -.0194$.)

X11-6. Form a third estimate of S as $\hat{S}_t^3 = S_3(L^{12})(x_t - \widehat{TC}_t^2)$, where $S_3(L^{12}) = \sum_{j=-3}^{3} d_j L^{12j}$, and where d_j are weights from a 3×5 moving average (*i.e.,* $1/15$, $2/15$, $3/15$, $3/15$, $3/15$, $2/15$, $1/15$).

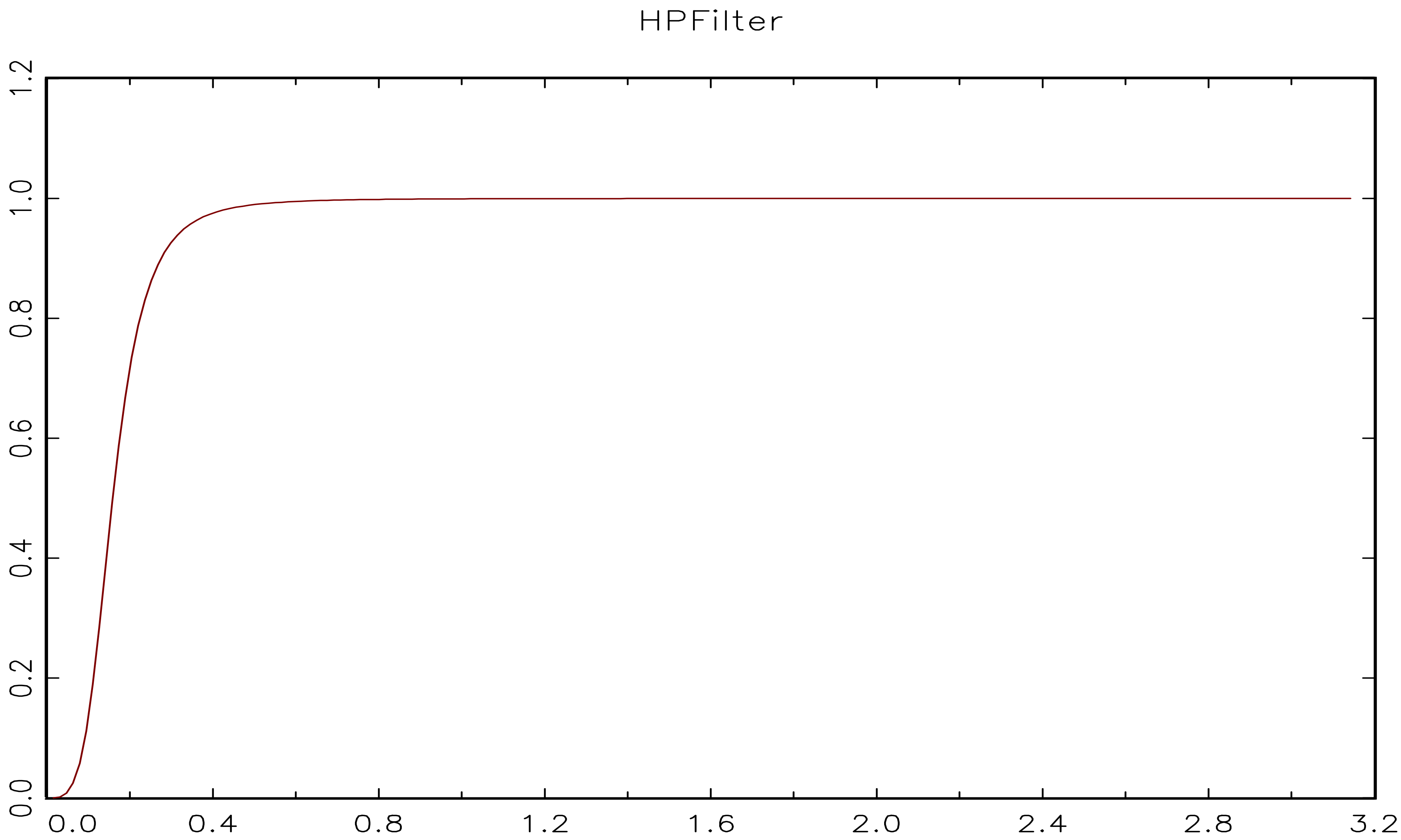
X11-7. Adjust the estimates of S so that they add to zero (approximately) over any 12 month period as $\hat{S}_t^4 = S_2(L)\hat{S}_t^3$, where $S_2(L)$ is defined in step 4.

X11-8. Form a final seasonally adjusted value as $x_t^{sa} = x_t - \hat{S}_t^4$.

Gain of Linear Approximation to X11 (Monthly Data)



Gain of HP Filter (λ = 1600)



Spectra of Commonly Used Stochastic Processes

Suppose that y has spectrum $S_y(\omega)$, and $x_t = c(L)y_t$.

What is the spectrum of x?

Because the frequency components of x are the frequency components of y scaled by the factor $g(\omega)e^{i\theta(\omega)}$, the spectra of x and y are related by

$$S_x(\omega) = g(\omega)^2 S_y(\omega) = c(e^{i\omega})c(e^{-i\omega})S_y(\omega).$$

Let ε_t denote a serially uncorrelated ("white noise") process. Its spectrum is $S_\varepsilon(\omega) = \dfrac{1}{2\pi}\left\{\gamma_0 + 2\sum_{k=1}^{\infty}\gamma_k \cos(\omega k)\right\} = \sigma_\varepsilon^2 / 2\pi$.

If y_t follows an ARMA process, then it can be represented as

$\phi(L)y_t = \theta(L)\varepsilon_t$, or $y_t = c(L)\varepsilon_t$ with $c(L) = \theta(L)/\phi(L)$.

The spectrum of y is then

$$S_y(\omega) = c(e^{i\omega})c(e^{-i\omega})S_\varepsilon(\omega) = \sigma_\varepsilon^2 \frac{\theta(e^{i\omega})\theta(e^{-i\omega})}{\phi(e^{i\omega})\phi(e^{-i\omega})}\frac{1}{2\pi}$$

Writing this out …

$$S_y(\omega) == \sigma_\varepsilon^2 \frac{(1-\theta_1 e^{i\omega} - \ldots - \theta_q e^{iq\omega})(1-\theta_1 e^{-i\omega} - \ldots - \theta_q e^{-iq\omega})}{(1-\phi_1 e^{i\omega} - \ldots - \phi_p e^{ip\omega})(1-\phi_1 e^{-i\omega} - \ldots - \phi_p e^{-ip\omega})}\frac{1}{2\pi}$$

A Classical Problem in Signal Processing: Constructing a Band-Pass Filter (see Baxter and King (1999) for discussion)

Let $c(L)=\sum_{j=-\infty}^{\infty} c_j L^j$. Suppose that we set the phase of $c(L)$ to be equal to 0 (and thus $c(L)$ is symmetric symmetric: $c_j = c_{-j}$) and we want

$$gain(c(L)) = |c(e^{i\omega})| = c(e^{i\omega}) = \begin{Bmatrix} 1 \text{ for } -\underline{\omega} \le \omega \le \underline{\omega} \\ 0 \text{ elsewhere} \end{Bmatrix}$$

(where the second equality follows because $c(L)$ is symmetric).

The calculation is straightforward

Because $c(e^{-i\omega}) = \sum_{j=-\infty}^{\infty} c_j e^{-i\omega j}$, then $c_j = (2\pi)^{-1}\int_{-\pi}^{\pi} e^{i\omega j} c(e^{-i\omega}) d\omega$ (an identity)

Setting the gain equal to unity over the desired frequencies and carrying out the integration yields

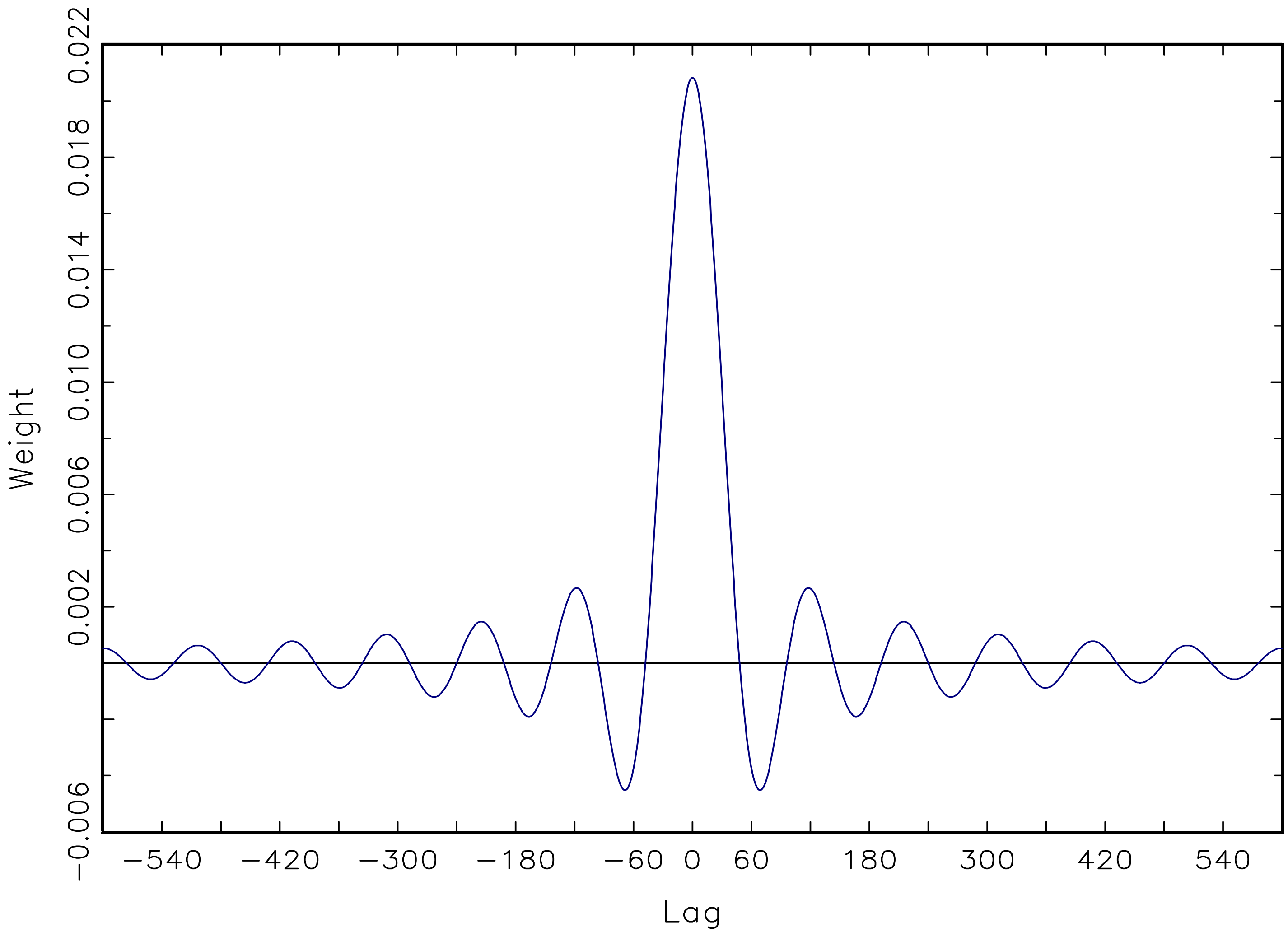
$$c_j = (2\pi)^{-1}\frac{1}{ij}e^{i\omega j}\Big|_{-\underline{\omega}}^{\underline{\omega}} = \begin{cases} \dfrac{1}{j\pi}\sin(\underline{\omega}j) \text{ for } j \neq 0 \\ \dfrac{\underline{\omega}}{\pi} \text{ for } j = 0 \end{cases}$$

Comments:

- The values of c_j die out at the rate j^{-1}
- $1-c(L)$ passes everything except $-\underline{\omega} \leq \omega \leq \underline{\omega}$. Differences of low-pass filters can be used to pass any set of frequencies.
- Baxter and King (1999) show that $c_k(L) = \sum_{j=-k}^{k} c_j L^j$ is an optimal finite order approximation to $c(L)$ in the sense that the gain of $c_k(L)$ is as close (L^2-norm) as possible to the gain of $c(L)$ for a k-order filter.

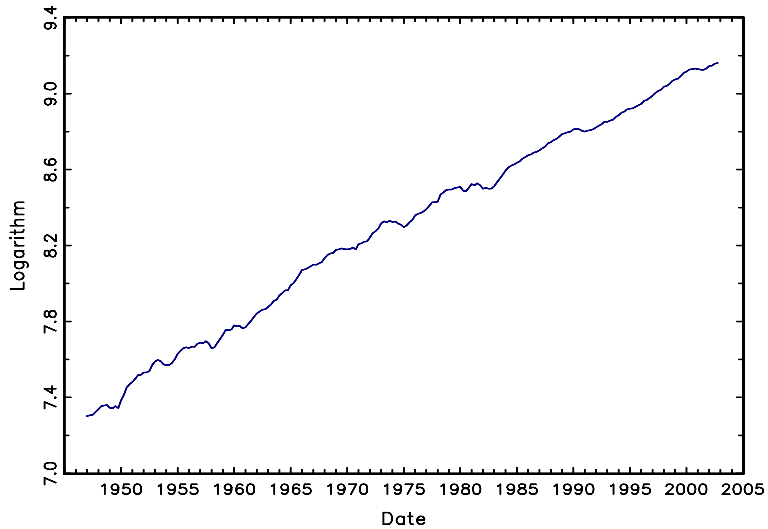
Band-Pass Filter Weights (Monthly Data) for constructing a "Gap"

Periods < 96 Months ($\omega > 2\pi/96$)

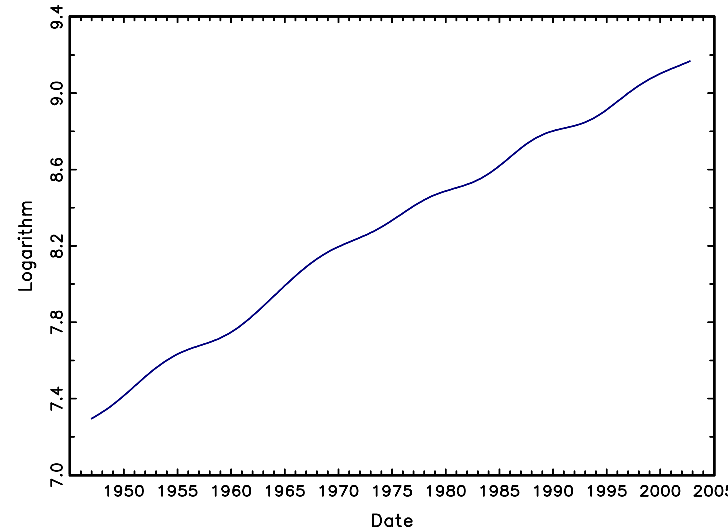


Log of Real GDP

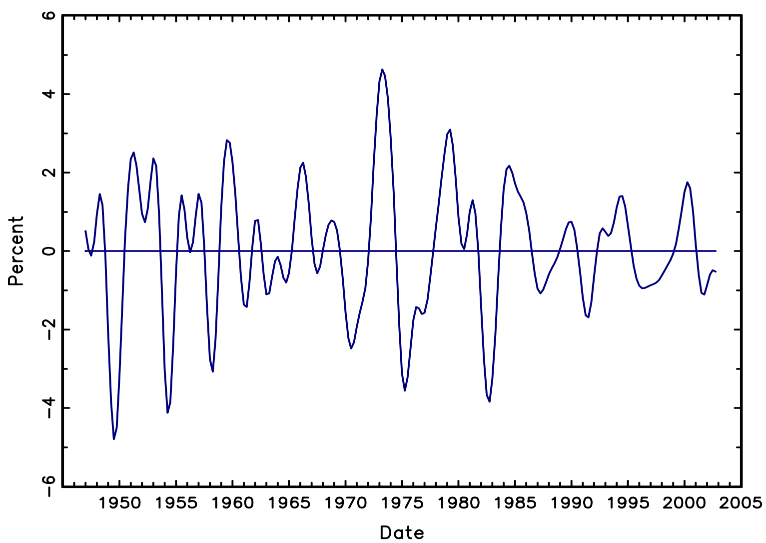
Series



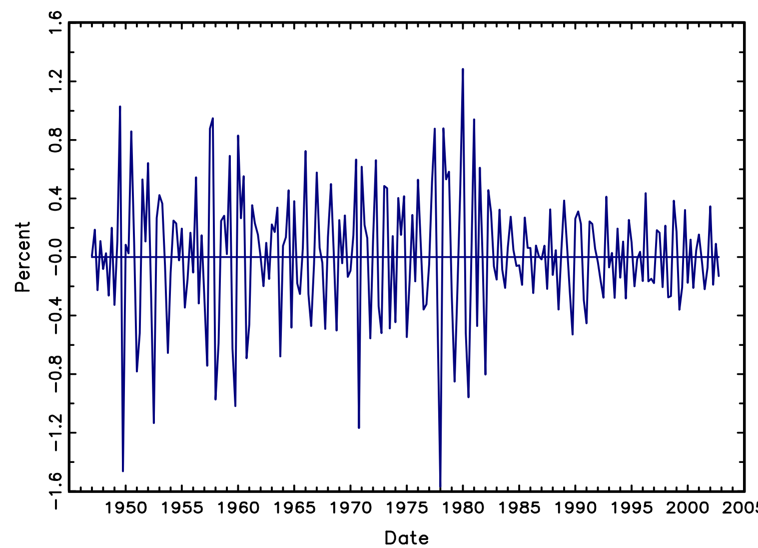
Periods > 32 Quarters



Periods Between 6 and 32 Quarters



Periods < 6 Quarters



Minimum MSE One-Sided Filters

Problem: these filters are two-sided with weights that die out slowly. This introduces "endpoint" problems. Geweke (1978) provides a simple way to implement a minimum MSE estimator using data available in real time. (See Christiano and Fitzgerald (2003) for an alternative approach.)

Geweke's calculation: if $x_t = c(L)y_t = \sum_{i=-\infty}^{\infty} c_i y_{t-i}$,

Optimal estimate of x_t given $\{y_j\}_{j=1}^{T}$:

$$E(x_t \mid \{y_j\}_{j=1}^{T}) = \sum_{i=-\infty}^{\infty} c_i E(y_{t-i} \mid \{y_j\}_{j=1}^{T})$$

$$= \sum_{i=-\infty}^{0} c_{t-i}\hat{y}_i + \sum_{i=1}^{T} c_{t-i} y_i + \sum_{i=T+1}^{\infty} c_{t-i}\hat{y}_i$$

where the $\hat{y}_i$'s denote forecasts and backcasts of y_i constructed from the data $\{y_j\}_{j=1}^{T}$.

These forecasts and backcasts can be constructed using AR, ARMA, VAR or other models. See Findley et. al (1991) for a description of how this is implemented in X-12.

The variance of the error associated with using $\{y_j\}_{j=1}^{T}$ is

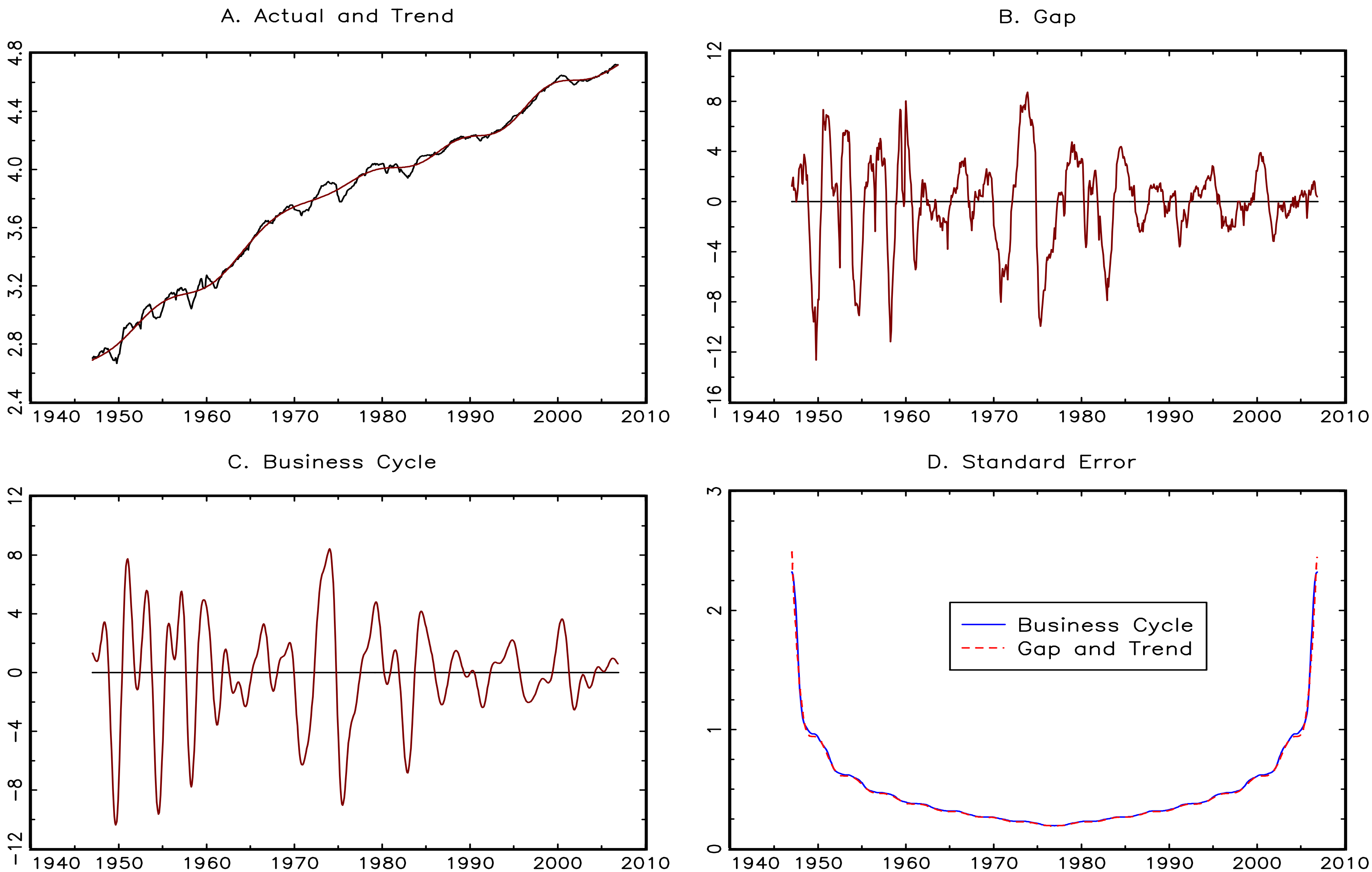
$$var[x_t - E(x_t \mid \{y_j\}_{j=1}^{T}] = var[\sum_{i=-\infty}^{\infty} c_i \{E(y_{t-i} \mid \{y_j\}_{j=1}^{T}) - y_{t-i}\}].$$

This looks messy, but it isn't.

Note: Geweke was concerned with the X11 filter, but his result applies to any linear filter (BandPass, HP, X11, …).

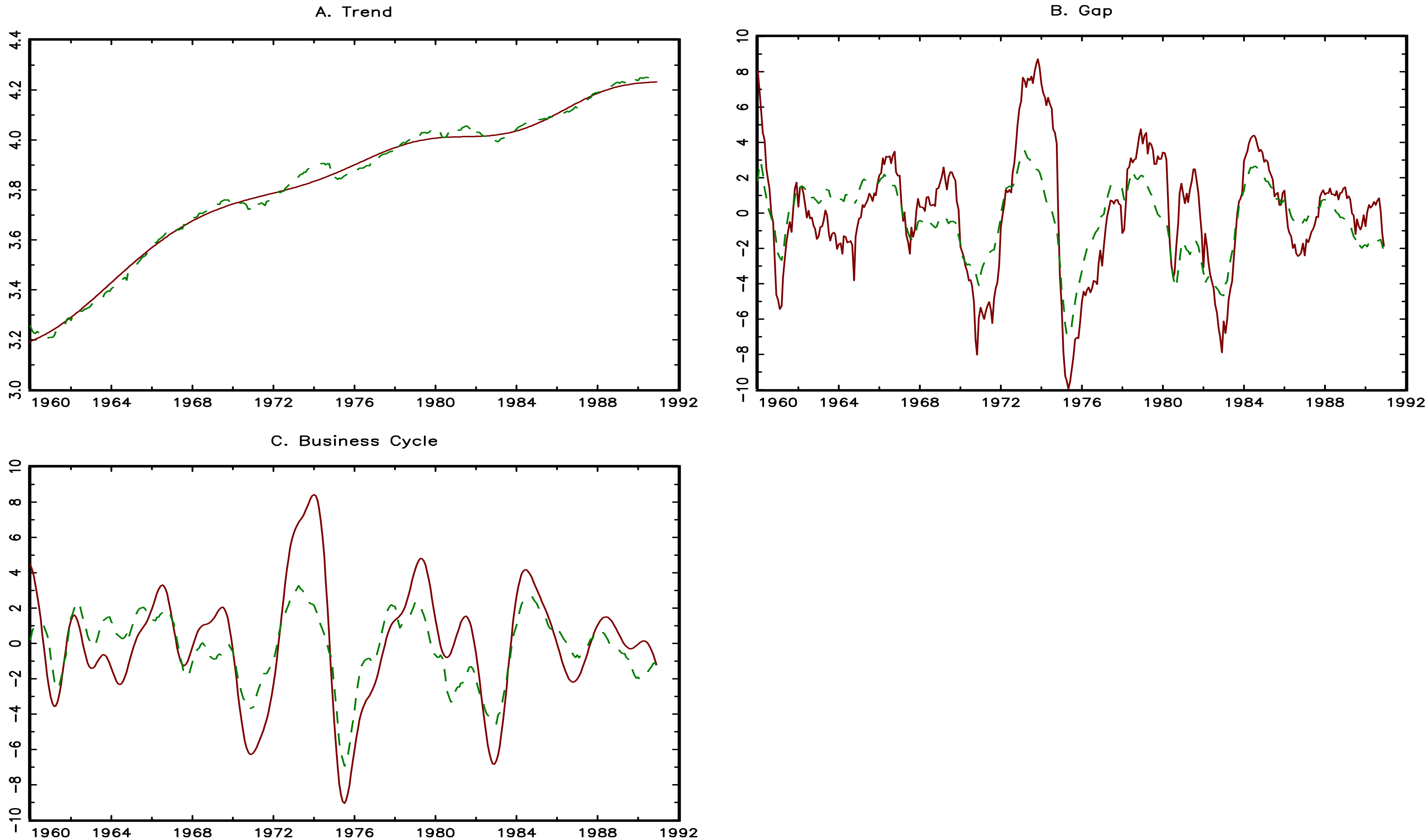
Two-Sided Band-Pass Estimates: Logarithm of the Index of Industrial Production

(From Watson (2007))



Notes: These panels show that estimated values of the band-pass estimates of the trend (periods > 96 months), the gap (periods < 96 months), and the business cycle (periods between 18 and 96 months). Panel D shows the standard errors of the estimates relative to values constructed using a symmetric 100-year moving average. Values shown in panel A correspond to logarithms, while values shown in panels B-D are percentage points.

Two-Sided (Solid) and One-Sided (Dashed) Band-Pass Estimates: Index of Industrial Production



Notes: The solid (red) lines are the two-sided shown in figure 2. The dashed (green) lines are one-sided estimates that do not use data after the date shown on the horizontal axis.

Standard Errors of One-Sided Band-Pass Estimates: AR (Univariate) and VAR (Multivariate) Forecasts

	$Y_{1\text{-}sided}^{Gap}$		$Y_{1\text{-}sided}^{BusinessCycle}$	
Series	AR	VAR	AR	VAR
Ind. Prod.	2.01	1.88	1.88	1.80
Unemp. Rate	0.46	0.41	0.43	0.40
Employment	0.78	0.77	0.75	0.75
Real GDP	1.03	0.86	0.95	0.83

Notes: This table summarizes results for the four series shown in the first column. The entries under "AR" are the standard errors of one-sided band-pass estimates constructed using forecasts constructed by univariate AR models with six lags. The entries under "VAR" are the standard errors of one-sided band-pass estimates constructed using forecasts constructed by VAR models with six lags for monthly models and four lags for quarterly models. The VAR models included the series of interest and first difference of inflation, the term spread, and building permits. Monthly models were estimated over 1960:9-2006:11, and quarterly models were estimated over 1961:III-2006:III.

Regressions Using Filtered Data

Suppose $y_t = x_t'\beta + u_t$ where $E(u_t x_t) = 0$

Consider using the filtered data $y_t^{filtered} = c(L)y_t$ and $x_t^{filtered} = c(L)x_t$.

Write $y_t^{filtered} = x_t^{filtered}{}'\beta + u_t^{filtered}$

Does $E(x_t^{filtered} u_t^{filtered}) = 0$?

This requires that x be "strictly exogenous". (Same reasoning for NOT doing GLS in time series regression.)

3. Multivariate spectra

If y_t is a scalar, $S(\omega) = (2\pi)^{-1} \sum_{j=-\infty}^{\infty} \lambda_j e^{-i\omega j}$ is the spectrum and represents the variance of the complex valued Cramér increment, $dZ(\omega)$, in $y_t = \int_{-\pi}^{\pi} e^{i\omega t} dZ(\omega)$. This can be generalized.

Let Y_t : $n\times 1$ vector, with j'th autocovariance matrix $\Gamma_j = V(Y_t Y'_{t-j})$. Let

$$S(\omega) = (2\pi)^{-1} \sum_{j=-\infty}^{\infty} \Gamma_j e^{-i\omega j}$$

so that $S(\omega)$ is an $n\times n$ matrix. The spectral representation is as above, but now $dZ(\omega)$ is an $n\times 1$ complex-valued random vector with spectral density matrix $S(\omega)$. $S(\omega)$ can be interpreted as a covariance matrix for the increments $dZ(\omega)$. The diagonal elements of $S(\omega)$ are the (univariate) spectra of the series. The off-diagonal elements are the "Cross-spectra".

The cross-spectra are complex valued, with $S_{ij}(\omega) = \overline{S_{ij}(\omega)}$. Cross spectra are often summarized using the following: The real part of $S_{ij}(\omega)$ sometimes called the *co-spectrum* and the imaginary part is called the *quadrature* spectrum. Then, consider the following definitions:

$$Coherence(\omega) = \frac{S_{ij}(\omega)}{\sqrt{S_{ii}(\omega) S_{jj}(\omega)}}$$

$$Gain_{ij}(\omega) = \frac{|S_{ij}(\omega)|}{S_{jj}(\omega)}$$

$$Phase_{ij}(\omega) = \tan^{-1}\left(\frac{-\operatorname{Im}(S_{ij}(\omega))}{\operatorname{Re}(S_{ij}(\omega))} \right)$$

To interpret these quantities, consider two scalars, Y and X with spectra S_Y, S_X, and cross-spectra S_{YX}. Consider the regression of Y_t onto leads and lags of X_t:

$$Y_t = \sum_{j=-\infty}^{\infty} c_j X_{t-j} + u_t = c(L)X_t + u_t$$

Because u and X are uncorrelated at all leads and lags: $S_Y(\omega) = |c(e^{i\omega})|^2 S_Y(\omega) + S_u(\omega)$.

Moreover: $E(Y_t X_{t-k}) = \sum_{j=-\infty}^{\infty} c_j E(X_{t-j} X_{t-k}) = \sum_{j=-\infty}^{\infty} c_j E(X_t X_{t-k+j}) = \sum_{j=-\infty}^{\infty} c_j \gamma_{k-j}$

where γ is denotes the autocovariances of X. Thus

$$S_{YX}(\omega) = (2\pi)^{-1} \sum_{k=-\infty}^{\infty} e^{-i\omega k} \sum_{j=-\infty}^{\infty} c_j \gamma_{k-j} = (2\pi)^{-1} \sum_{j=-\infty}^{\infty} c_j e^{-i\omega j} \sum_{l=-\infty}^{\infty} e^{-i\omega l} \gamma_l = c(e^{-i\omega}) S_X(\omega)$$

Thus, the gain and phase of the cross-spectrum is recognized as the gain and phase of $c(L)$.

4. Spectral Estimation

AR/VAR/ARMA Parametric Estimators: Suppose $\Phi(L)Y_t = \Theta(L)\varepsilon_t$, where Y may be a vector and ε_t is white noise with covariance matrix Σ_ε. The spectral density matrix of Y is

$$S_Y(\omega) = \Phi(e^{i\omega})^{-1}\Theta(e^{i\omega})\Sigma_\varepsilon\Theta(e^{-i\omega})'\Phi(e^{-i\omega})'^{-1}$$

A parametric estimator uses estimated values of the AR, MA parameters and Σ_ε.

Example (VAR(1)): $(I - \Phi L)Y_t = \varepsilon_t$

$$\hat{S}_Y(\omega) = (I - \hat{\Phi}e^{i\omega})^{-1}\hat{\Sigma}_\varepsilon(I - \hat{\Phi}e^{-i\omega})'^{-1}$$

Nonparametric Estimators: To be discussed in Lecture 5 in the context of HAC covariance matrix estimators.