Christiano
FINC 520, Spring 2009
Homework 5, due Thursday, May 7.

1. Generate 10,000 observations from the first order autoregressive time series representation:
$$y_t = 0.9y_{t-1} + \varepsilon_t.$$
Then, HP filter the data using the attached software.[1] Recall that the HP filter solves the problem:
$$\min_{\left\{y_t^T\right\}} \sum_{t=1}^{T-1} \left\{ \left(y_t - y_t^T\right)^2 + \lambda \left[\left(y_{t+1}^T - y_t^T\right) - \left(y_t^T - y_{t-1}^T\right) \right]^2 \right\},$$
and the 'HP-filtered' data are
$$y_t^c \equiv y_t - y_t^T.$$
   (a) Graph y_t and y_t^T to verify that the HP filter is operating properly. The y_t^T series should be a smooth version of y_t.
   (b) Calculate $C(0)$, $C(2)$, $C(4)$ using the artificial data, where $C(\tau)$ is the lag τ covariance of y_t^c.
   (c) Repeat the same exercise using the spectral analysis tools we have developed. For this, you will first need the Fourier transform of the HP-filter. To find this, compute the first order necessary condition for optimality satisfied by y_t^T. In lag operator form, this has the representation,
$$y_t = B(L)\, y_t^T,$$
where $B(L)$ is symmetric in positive and negative powers of L, i.e., $B(L) = B\left(L^{-1}\right)$. Note that since $y_t^T = y_t - y_t^c$, this implies:
$$y_t = B(L)\, y_t - B(L)\, y_t^c,$$
or,
$$y_t^c = g(L)\, y_t,$$

[1] The code is HPFAST.m, programmed by Ed Prescott. The call is [d,t]=hpfast(y,s), where y is the column vector of data, s (=1,600) is the value of the multiplier on the HP filter, d denotes the detrended (i.e., 'cyclical') data and t denotes the HP trend.

where the HP filter, $g(L)$ is

$$g(L) = \frac{B(L) - 1}{B(L)}.$$

i. To see what the HP filter does to a time series, graph $g(e^{-i\omega})$ for $\omega \in (0, \pi)$. Note that it looks like a high pass filter: it lets higher frequencies of oscillation through and zeros out the lower frequencies. What is the cutoff between frequencies allowed through and frequencies set to zero? Here, imagine you are working with quarterly data and $\lambda = 1600$.

ii. Use the spectral formulas together with the Riemann approximation to calculate $C(0), C(2), C(4)$. For N sufficiently large, your answer should be the same as in (a).

2. Consider the process

$$x_t = \varepsilon_t - \varepsilon_{t-1}. \tag{1}$$

One might be tempted to 'invert' this process as follows:

$$x_t + x_{t-1} + x_{t-2} + ... = \varepsilon_t.$$

(a) show that the process on the left side of the equality is not well defined in that it is not the limit of a sequence of finite sums of y_t that converges in the mean square sense.

(b) show that there is nevertheless a sense in which ε_t is the one-step-ahead forecast error in y_t. In particular, show that

$$\begin{aligned} P[x_t | x_{t-1}, x_{t-2}, ..., x_{t-n}] &= \sum_{j=1}^{n} d_j x_{t-j}, \\ d_j &= -\left[1 - \frac{j}{n+1}\right]. \end{aligned}$$

(Hint: prove the result for the cases, $n = 2$ and $n = 3$ and the general Prove that

$$E(x_t - P[x_t | x_{t-1}, x_{t-2}, ..., x_{t-n}] - \varepsilon_t)^2 \to 0 \text{ as } n \to \infty.$$

(c) It has been argued that processes with a unit root in the moving average representation, like (1), are extremely unlikely in practice, since most data are measured with at least a little bit of measurement error. That is, even if the true data were measured as in (1), measured data look like x_t^*, where

$$x_t^* = x_t + \upsilon_t,$$

where υ_t white noise and orthogonal to x_t at all leads and lags. (Hint: show that the spectrum of x_t^* is the sum of the spectrum of x_t and of υ_t, and that the unit moving average root in (1) corresponds to a spectrum which is zero at frequency zero.)

3. Prove that the Wold error derived in the handout on spectral analysis and projection problems is not-autocorrelated over time.

4. We motivated the spectral decomposition theorem using the band pass filter and the discrete decomposition of data into a set of sinusoidal functions. This question explores the latter.

   Consider the discrete analog of the spectral representation theorem:

$$\begin{aligned} y_t &= \sum_{j=0}^{(T-1)/2} y_{j,t}, \\ y_{j,t} &= \alpha_j \cos(\omega_j t) + \delta_j \sin(\omega_j t), \end{aligned} \tag{2}$$

   where T is the number of observations on y_t, T is assumed to be odd,

$$\omega_j = \frac{2\pi j}{T},\ j = 0, ..., \frac{T-1}{2},$$

   and α_j and δ_j, $j = 0, ..., (T-1)/2$ are a set of parameters to be determined. At first sight it may look like there are $T+1$ parameters here - 2 for each value of j - and so one more parameters than the number of observations. In fact this is not true, since δ_0 multiplies the zero vector and so can be ignored. Recall that $\cos(x) = \cos(x + 2\pi k)$ and $\sin(x) = \sin(x + 2\pi k)$ for any integer k, so that the j^{th} component of y_t in (2), $y_{j,t}$, has period $2\pi/\omega_j$. Note that higher values of ω_j (i.e., 'higher frequency components of y_t') are associated with cycles of shorter duration, or period.

The equations in (2) can be written in matrix form like this:

$$y = X\beta,$$

where

$$\beta = \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \delta_1 \\ \alpha_2 \\ \delta_2 \\ \vdots \\ \alpha_{\frac{T-1}{2}} \\ \delta_{\frac{T-1}{2}} \end{bmatrix},$$

and X corresponds to the cosine and sine variables in (2) (note how we've ignored δ_0).

(i) Show that

$$\begin{aligned} \frac{1}{T}\sum_{t=1}^{T} \sin(\omega_j t)\sin(\omega_l t) &= \frac{1}{T}\sum_{t=1}^{T} \cos(\omega_j t)\cos(\omega_l t) = \begin{cases} \frac{1}{2} & j = l > 0 \\ 0 & j \neq l \end{cases} \\ \frac{1}{T}\sum_{t=1}^{T} \cos(\omega_j t)\sin(\omega_l t) &= 0, \ \text{all } j, l, \end{aligned}$$

so that X is a square matrix with orthogonal columns. Thus, (2) represents an exact decomposition of a time series (*any* time series, not just the covariance stationary and indeterministic series addressed by the spectral decomposition theorem) into orthogonal sinusoidal components, as in the spectral decomposition theorem.

(ii) Show that

$$\widehat{var}(y_t) = \frac{1}{T}\sum_{t=0}^{T}(y_t - \bar{y})^2 = \sum_{j=0}^{\frac{T-1}{2}} \widehat{var}(y_{j,t}),$$

where

$$\widehat{var}(y_{j,t}) = \frac{1}{2}\left[\alpha_j^2 + \delta_j^2\right], \ j = 1, ..., \frac{T-1}{2},$$

and the α_j's and δ_j's are the unique solution to

$$\beta = \left(X'X\right)^{-1} X'y.$$

(Note that this reduces to $\beta = X^{-1}y$ since X is square and invertible, but it is somewhat easier to understand the vector β by using the more elaborate formula.)

(iii) Denote the k^{th} order sample covariance of $y_1, ..., y_T$ by

$$\begin{aligned}
\hat{\gamma}_k &= \frac{1}{T} \sum_{t=k+1}^{T} \left(y_t - \bar{y}\right)\left(y_{t-k} - \bar{y}\right) \\
\bar{y} &= \frac{1}{T} \sum_{t=1}^{T} y_t,
\end{aligned}$$

and $\hat{\gamma}_{-k} = \hat{\gamma}_k$, for $k = 0, ..., T-1$. Denote the *sample periodogram* by

$$\hat{S}_y\left(\omega\right) = \frac{1}{2\pi} \sum_{k=-T+1}^{T-1} \hat{\gamma}_k e^{-i\omega k}.$$

This is the sample analog of the population spectral density. Show that

$$\hat{S}_y\left(\omega_j\right) = \frac{T}{4\pi} \widehat{var}\left(y_{j,t}\right),$$

so that the periodogram at frequency ω_j corresponds to the variance of the j^{th} frequency component of the data. This is the analog of a similar result that we derived using an argument based on the band pass filter.

5. Consider the industrial production data made available with this homework. There are two versions of the data. One is seasonally unadjusted and the other is seasonally adjusted. The data cover the period, January 1919 to December 2006. Compute the first difference of the log of the data, to obtain the 1,055 monthly growth rates from February 1919 to December 2006.

   (a) Consider the seasonally unadjusted data first. Compute the sample periodogram for frequencies, $\omega_j = 2\pi j/T$, $j = 0, ..., (T-1)/2$

and graph the results. Note how jagged the curve is (throughout, you should only graph the log of the spectrum). This reflects the result (see Hamilton, p. 164) that $\hat{S}_y(\omega)$ and $\hat{S}_y(\omega')$ are approximately independent (for large T) for $\omega \neq \omega'$. Moreover, although $\hat{S}_y(\omega)$ is an unbiased estimator of the true spectrum, $S_y(\omega)$, its variance does not shrink to zero as $T \to \infty$. The lack of precision in the sample spectrum as an estimator of the spectral density is perhaps not surprising. The function being estimated (i.e., the spectrum as a function of frequency) is a high-dimensional object (there is a continuum of frequencies between 0 and π), and no assumptions are made about the structure of the underlying time series representation. This is an example of 'little input' implies 'little output'.

(b) Now consider a more parametric way to estimate the spectrum of the seasonally unadjusted data (call this the 'ar estimator of the spectrum'). Use ordinary least squares to fit a 20 lag scalar ar representation to the data, with a constant term. Compute the spectral density of the resulting ar representation over the same range of frequencies used in (a). Graph the two spectral density estimators in the same picture. Note how one appears to be a smooth version of the other.

(c) Note the local peaks in the spectrum. What period of oscillation do these correspond to?

(d) Apply the ar estimator of the spectrum to the seasonally adjusted data. Graph the spectrum of the adjusted and unadjusted data in the same figure. Are there 'dips' in the spectrum of the seasonally adjusted data, as we were led to expect based on the results for 'optimal seasonal adjustment' in the previous homework? Why or why not?

(e) An alternative strategy that is sometimes used to seasonally adjust monthly data is to regress the data on 12 seasonal dummies (don't include a constant term here, or you'll have perfectly collinear data!)[2] and treat the residual in this regression as the seasonally

[2] That is, let the right hand variables be in the T by 12 matrix, X. The i^{th} column of X has zeros everywhere and a unity in the i^{th} entry. A MATLAB routine that will set up

adjusted data. In the same graph, plot the estimated spectrum of the unadjusted data, the data adjusted by the US government and the data adjusted using the dummy method (in all cases, compute the spectrum using the ar method). Which is the more effective seasonal adjustment procedure, the dummy method or the US government's method?

X with the right structure is:

```
B=eye(12);
X=[];
for ii = 1:88
X=[X
B];
end
X=X([2:end],:);
```