Econ 584 Lab 2
Spring 2006

Eric Zivot

Due: Wednesday, May 3

1 Reading

1. Hamilton, J. (1993), *Time Series Analysis*, chapters 3, 4 and 13 (sections 1-4)
2. Hayashi, F. (2000), *Econometrics*, chapter 6.
3. Cochrane, J. (2005), *Time Series for Macroeconomics and Finance*, chapters 5 and 10.
4. Zivot, E. and J. Wang (2002), chapter 3 and chapter 10 in *Modeling Financial Time Series with S-PLUS*. Springer-Verlag.
5. Zivot, E. (2005). Lecture notes on forecasting, ARMA estimation, state space models, and trend/cycle decompositions.
6. Diebold, F. and R. Mariano (1995). "Comparing predictive accuracy," *Journal of Business and Economic Statistics*, 13, 253-265. Re-printed in *Journal of Business and Economic Statistics*, 20(1), 134-145, January 2002.
7. Clark, P. (1987). "The cyclical component of economic activity," *Quarterly Journal of Economics*. Available in JSTOR.
8. Eviews help topics: Forecasting from an equation; State space models and the Kalman filter.

2 Analytic Questions

1. Let c_t^p denote the (log of) permanent consumption and assume that c_t^p follows a random walk with drift

$$c_t^p = \mu + c_{t-1}^p + \varepsilon_t,\ \varepsilon_t \sim iid\ N(0, \sigma_\varepsilon^2)$$

Assume that observed consumption, c_t, is equal to permanent consumption, c_t^p, plus a Gaussian white noise error term

$$c_t = c_t^p + \eta_t,\ \eta_t \sim iid\ N(0, \sigma_\eta^2)$$

(a) What is a state space representation for c_t? How would you determine the initial state parameters?

(b) Derive the reduced form representation for Δc_t in terms μ, ε_t and η_t.

(c) Determine the autocorrelation function (ACF) for Δc_t. Given that the mean, variance and ACF uniquely determines a covariance stationary and ergodic process, what type of ARMA(p,q) process describes Δc_t? (Hint: see Hamilton pages 102 - 108)

(d) What is the functional relationship between the reduced form model parameters $\mu, \sigma_\varepsilon^2$ and σ_η^2 and the parameters of the ARMA(p,q) process that describes Δc_t? In other words, if you knew the ARMA(p,q) parameters for Δc_t how would you solve for the reduced form parameters $\mu, \sigma_\varepsilon^2$ and σ_η^2?

2. Let y_t be a covariance stationary time series, and consider the prediction of y_{t+h} given information available at time t. Recall, the optimal linear predictor and its associated prediction error are

$$\begin{aligned} y_{t+h|t} &= \mu + \psi_h \varepsilon_t + \psi_{h+1}\varepsilon_{t-1} + \cdots \\ \varepsilon_{t+h|t} &= y_{t+h} - y_{t+h|t} = \varepsilon_{t+h} + \psi_1 \varepsilon_{t+h-1} + \cdots + \psi_{h-1}\varepsilon_{t+1} \end{aligned}$$

Derive the correlation between $\varepsilon_{t+h|t}$ and $\varepsilon_{t+j|t}$ for the pairs $h = 2, j = 1$ and $h = 3, j = 1$, $h = 3, j = 2$.

3. Consider samples of size T from the AR(1) and MA(1) processes

$$\begin{aligned} y_t &= c + \phi y_{t-1} + \varepsilon_t,\ \varepsilon_t \sim WN(0, \sigma_\varepsilon^2) \\ x_t &= \mu + \eta_t + \theta \eta_{t-1},\ \eta_t \sim WN(0, \sigma_\eta^2) \end{aligned}$$

(a) Assuming each process is stationary and ergodic, give the asymptotic distribution for the sample means

$$\begin{aligned} \bar{y} &= T^{-1}\sum_{t=1}^{T} y_t \\ \bar{x} &= T^{-1}\sum_{t=1}^{T} x_t \end{aligned}$$

That is, give the means and variances of the asymptotic distributions for $\bar{y}$ and $\bar{x}$. (Hint: Recall that the long-run variance has the form $LRV = \sum_{j=-\infty}^{\infty} \gamma_j = \sigma^2 \psi(1)^2$ where $\psi(L) = \phi(L)^{-1}\theta(L)$.)

(b) Describe how you could consistently estimate the asymptotic variances of $\bar{y}$ and $\bar{x}$ parametrically and non-parametrically.

(c) What happens to the asymptotic variance of $\bar{y}$ as $\phi \to 1$? Does this result make sense?

(d) What happens to the asymptotic variance of $\bar{x}$ as $\theta \to -1$? Does this result make sense? (Hint: What type of ARMA process results from taking the first difference of the process $x_t = \mu + \eta_t$?)

4. Consider the following unobserved components model for the log of quarterly real GDP y_t:

$$
\begin{aligned}
y_t &= \mu_t + c_t \\
\mu_t &= \delta_{t-1} + \mu_{t-1} + \eta_t,\ \eta_t \sim iid\ N(0, \sigma_\eta^2) \\
\delta_t &= \delta_{t-1} + u_t,\ u_t \sim iid\ N(0, \sigma_u^2) \\
c_t &= \phi_1 c_{t-1} + \phi_2 c_{t-2} + v_t,\ v_t \sim iid\ N(0, \sigma_v^2)
\end{aligned}
$$

where η_t, u_t and v_t are mutually independent. This model was originally proposed by Peter Clark and is known as the "Clark model". Notice that the trend is specified as a random walk with drift, where the drift also follows a random walk.

(a) Give a state space representation for the above process.

(b) How would you specify the distribution of the initial state vector?

(c) Under what restrictions on the model's parameters does the trend become deterministic?

3 Computer Exercises

In this lab you will use Eviews to compare forecasting models using the Diebold-Mariano statistic, and estimate and analyze some ARMA models put into state space form.

3.1 Comparing Forecasting Models

1. Using Eviews, create a new work file with 250 undated observations. Then create a series named Y2 with simulated values from the AR(2) process

$$
\begin{aligned}
y_t &= 1.2 y_{t-1} - 0.4 y_{t-2} + \varepsilon_t,\ \varepsilon_t \sim iid\ N(0, (0.5)^2) \\
y_1 &= y_2 = 0
\end{aligned}
$$

See the Eviews online help for the function nrnd for an example of simulating AR processes.

2. For the simulated series do the following: (1) plot series; (2) plot SACF and PACF; compute descriptive statistics.

3. Using the first 200 observations on Y2, estimate the AR(2) model using the LS command; e.g., LS Y2 C AR(1) AR(2). Report the fit of the equation. From the [View] menu on the Equation window, investigate the ARMA structure, plot the actual, fitted and residuals, and plot the SACF and PACF of the estimated residuals. Comment briefly on what you find. Name the equation AR2FIT, and close the equation window.

4. Using the first 200 observations on Y2, estimate a mis-specified MA(1) model using the LS command; e.g., LS Y2 C MA(1). Report the fit of the equation. From the [View] menu on the Equation window, investigate the ARMA structure, plot the actual, fitted and residuals, and plot the SACF and PACF of the estimated residuals. Comment briefly on what you find and compare with the AR(2) fit. Name the equation MA1FIT.

5. Using observations 201 through 250, compute rolling 1-step ahead (static) forecasts of Y2 from the AR(2) and MA(1) models. Name the AR(2) forecasts Y2FAR2, and name the MA(1) forecasts Y2FMA1. Compare the RMSE and MAE from the two sets of forecasts and comment briefly.

6. From Y2 and the AR(2) and MA(1) forecasts, compute the 1-step ahead forecast errors for observations 201 through 250; e.g., for the AR(2) model the forecast error may be defined using GENR EAR2 = Y2 - Y2FAR2 (don't forget to set the sample correctly). Using the forecast errors from the two models, compute the squared and absolute loss differentials relative to the AR(2) model; e.g., for the squared loss differential use GENR DSQ = EMA1^2 - EAR2^2; for the absolute loss differential use GENR DABS = ABS(EMA1) - ABS(EAR2). Plot these loss differentials. Which model appears to forecast better?

7. Compute the Diebold-Mariano statistic for the loss differentials DSQ and DABS. This can be easily done by regressing each loss differential on a constant and computing the standard error using a HAC (Newey-West type) estimator. The Diebold-Mariano statistic is then the reported t-statistic on the constant. Using these statistics, test the hypothesis that the AR(2) model and the MA(1) model have equal forecasting accuracy using a 5% significance level.

3.2 Working with State Space Models

In this exercise, you will use the real GDP data from economagic that you analyzed in Lab 1. Before doing any estimation, create a new variable that is equal the log of real GDP times 100 and use this variable for all of the estimation. This seemingly innocuous transformation scales the data so that the estimation of the state space model is numerically stable. If you don't scale the data then the state space estimation may not converge.

1. Create a new series by multiplying the log real GDP data by 100. Call this series lrgdp_100. This simple transformation will increase the scale of the error variances and make the estimation of the state space model more stable.

2. Linearly detrend the real GDP data (multiplied by 100) as in lab 1 by regressing on a constant and time trend and take the residuals as the detrended data. Call this data dtlrgdp.

3. Using the detrended data over the period 1947.1 through 1999.1, fit an AR(2) model without a constant using the LS command: LS dtlrgdp AR(1) AR(2). Name the equation LSFIT. Note: you will lose the first two observations due to the lagged values in the AR(2) model. Create dynamic forecasts over the period 2000.1 through 2003.4.

4. Create a state space object named SSAR2 representing an AR(2) model without a constant. See the Eviews help on State Space Models and the Kalman Filter for details on how to do this. Estimate the state space model using the detrended data over the period 1947.1 through 1999.1. Note: you will not lose the first two values because the state space model utilizes the exact likelihood function. Create dynamic forecasts over the period 2000.1 through 2003.4. Compare the results to the LS fit. Should they be the same?

5. Using the state space model, compute the filtered estimates of the state variables. What do these variables represent?

3.3 Estimate Simple Unobserved Components Model

In this exercise, you will estimate a simple UC model for the log-level of quarterly real GDP.

1. Use the series lrgdp_100 for the estimation in this section. This simple transformation will increase the scale of the error variances and make the estimation of the state space model more stable.

2. Create a state space object named SSCLARK for the following unobserved components model

$$
\begin{aligned}
y_t &= \mu_t + c_t \\
\mu_t &= \delta + \mu_{t-1} + \eta_t, \ \eta_t \sim iid\ N(0, \sigma_\eta^2) \\
c_t &= \phi_1 c_{t-1} + \phi_2 c_{t-2} + v_t, \ v_t \sim iid\ N(0, \sigma_v^2)
\end{aligned}
$$

where η_t, u_t and v_t are mutually independent. Estimate this model by maximum likelihood using the log-level of real GDP that you analyzed in Lab 1. Be sure to properly specify the distribution of the initial state vector.

3. Compute the roots of $\phi(z) = 0$ using the estimated values of ϕ_1 and ϕ_2. Briefly discuss the dynamic properties of the estimated cycle.

4. The importance of the random walk trend relative to the AR(2) cycle can be determined by looking at the ratio of σ_η^2 to σ_v^2. Compute this ratio using the maximum likelihood estimates and comment briefly on what you find.

5. Compute and plot the filtered estimates of the signal (observation) and the states. Comment briefly on what you find. Do the filtered estimates of the cycle look like "business cycles"?