Financial Econometrics
HT Week 1 Assignment Answers

February 4, 2010

3.7 **Suppose that y_t follows a random walk then $\Delta y_t = y_t - y_{t-1}$ is stationary.**

(a) **Is $y_t - y_{t-j}$ for and $j \geq 2$ stationary?**

Yes. It is straight forward to show that for a random walk, $y_t = \sum_{i=1}^{t} \epsilon_t + y_0$. From this definition, it is also clear that

$$
\begin{aligned}
y_t - y_{t-j} &= \sum_{i=1}^{t} \epsilon_i + y_0 - \left(\sum_{k=1}^{t-j} \epsilon_k + y_0 \right) \\
&= \sum_{i=0}^{j-1} \epsilon_{t-i}
\end{aligned}
$$

which, by assumption is the sum of j ϵs, all of which are white noise. The only thing left to do is to show that the variance of the sum is finite, which is easy to show using the variance of a sum formula.[1] The final step to show that the autocovariances do not depend on t, which is simple since $y_t - y_{t-j}$ is a MA($j-1$).

(b) **If it is and $\{\epsilon_t\}$ is an i.i.d. sequence of standard normals, what is the distribution of $y_t - y_{t-j}$?**

Since $y_t - y_{t-j} = \sum_{i=0}^{j-1} \epsilon_{t-i}$ is the sum of j IID normal variables with variance σ^2, $y_t - y_{t-j} \sim N(0, j\sigma^2)$.

(c) **What is the joint distribution of $y_t - y_{t-j}$ and $y_{t-h} - y_{t-j-h}$ (Note: The derivation for an arbitrary h is challenging)?**

The key to this problem is to note that since the errors are IID normal, they are also jointly normally distributed,

$$
\begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_{t-1} \\ \epsilon_t \end{bmatrix} \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_t)
$$

[1] Using the Cauchy-Schwartz inequality, it is straightforward to show that $\mathrm{V}[\sum_{i=1}^{k} x_k] \leq (K + K(K-1))\sigma^2 = K^2\sigma^2$ whenever the variance of x_k is finite for all k.

where $\mathbf{0}$ is a t by 1 vector of 0s and $\sigma^2\mathbf{I}_t$ is a diagonal matrix composed of the variance of the ϵs. An alternative method to determine the distribution is to use vectors of "weights" to compute the joint distribution. The weight vectors are just vectors of 1s and 0s that are used to sum the correct ϵs together to form $y_t - y_{t-j}$. Let $\boldsymbol{\epsilon} = [\epsilon_1\ \epsilon_2\ \ldots\ \epsilon_t]'$ be a t by 1 vector of errors. Using this notation,

$$y_t - y_{t-j} = \mathbf{w}_t\boldsymbol{\epsilon} = \sum_{i=1} w_{it}\epsilon_i = 0 + 0 + \ldots + \epsilon_{t-j+1} + \epsilon_{t-j+2} + \ldots + \epsilon_t.$$

Simply by writing out the terms in $\mathbf{w}_{t-i}$ in $y_{t-i} - y_{t-i-j} = \mathbf{w}_{t-i}\boldsymbol{\epsilon}$ that $\mathbf{w}_{t-i} = [\mathbf{0}'_{t-i-j}\ \mathbf{1}'_j\ \mathbf{0}'_i]$ where $\mathbf{0}_k$ is a k by 1 vector of 0s and $\mathbf{1}_k$ is a k by 1 vector of 1s. By stacking the $t - j + 1$ elements in to a matrix, $\mathbf{W}$, the standard properties of a multivariate normal can be used,

$$\mathbf{W} = \begin{bmatrix} \mathbf{w}_j \\ \mathbf{w}_{j+1} \\ \vdots \\ \mathbf{w}_{t-1} \\ \mathbf{w}_t \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & \ldots & 1 & 0 & 0 & \ldots & 0 \\ 0 & 1 & 1 & \ldots & 1 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 1 & 1 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & 1 & \ldots & 1 & 1 & 0 \\ 0 & 0 & \ldots & 0 & 1 & \ldots & 1 & 1 & 1 \end{bmatrix}$$

Using this notation, the vector of differences can be written as

$$\Delta\mathbf{y} = \mathbf{W}\boldsymbol{\epsilon}$$

This is nothing more than a rotation of a multivariate normal random variable and since MV normals are closed under rotations, the answer is complete,and the joint distribution of the two is

$$\begin{bmatrix} y_t - y_{t-j} \\ y_{t-h} - y_{t-j-h} \end{bmatrix} \sim N\left(\mathbf{0}, \begin{bmatrix} j\sigma^2 & \max\left(0, j-h\right)\sigma^2 \\ \max\left(0, j-h\right)\sigma^2 & j\sigma^2 \end{bmatrix}\right)$$

or in the complete form,

$$\mathbf{W}\boldsymbol{\epsilon} \sim N\left(0, \sigma^2\mathbf{W}\mathbf{W}'\right)$$

Note: If it helps in this problem, consider the case where $j = 2$ and $h = 1$.

In this case the joint distribution is

$$\begin{bmatrix} y_t - y_{t-2} \\ y_{t-1} - y_{t-3} \end{bmatrix} \sim N\left(\mathbf{0}, \begin{bmatrix} 2\sigma^2 & \sigma^2 \\ \sigma^2 & 2\sigma^2 \end{bmatrix}\right)$$

3.10 **Answer the following questions about forecast errors.**

(a) **Let $y_t = \phi_0 + \phi_1 y_{t-1} + \epsilon_t$ with the usual assumptions on $\{\epsilon_t\}$. Derive an explicit expression for the 1-step and 2-step ahead forecast errors, $e_{t+h|t} = y_{t+h} - \hat{y}_{t+h|t}$ where $\hat{y}_{t+h|t}$ is the MSE optimal forecast where $h = 1$ or $h = 2$ (what is the MSE optimal forecast?).**

This question require repeated forward recursion until the pattern can be discerned. First recall that the MSE optimal forecast is the conditional mean, which is $\mathrm{E}_t[y_{t+h}] = \sum_{i=0}^{h-1} \phi_1^i \phi_0 + \phi_1^h y_t$ in this problem. Begin with the 1-step ahead forecast,

$$\begin{aligned} e_{t+1|t} &= y_{t+1} - \hat{y}_{t+1|t} \\ &= \phi_0 + \phi_1 y_t + \epsilon_{t+1} - \phi_0 - \phi_1 y_t \\ &= \epsilon_t. \end{aligned}$$

The 2-step can next be derived,

$$\begin{aligned} e_{t+2|t} &= y_{t+2} - \hat{y}_{t+2|t} \\ &= \phi_0 + \phi_1 y_{t+1} + \epsilon_{t+2} - \phi_0 - \phi_1\phi_0 - \phi_1^2 y_t \\ &= \phi_1 y_{t+1} + \epsilon_{t+2} - \phi_1\phi_0 - \phi_1^2 y_t \\ &= \phi_1(\phi_0 + \phi_1 y_t + \epsilon_{t+1}) + \epsilon_{t+2} - \phi_1\phi_0 - \phi_1^2 y_t \\ &= \phi_1\phi_0 + \phi_1^2 y_t + \phi_1\epsilon_{t+1} + \epsilon_{t+2} - \phi_1\phi_0 - \phi_1^2 y_t \\ &= \phi_1\epsilon_{t+1} + \epsilon_{t+2}. \end{aligned}$$

Repeating this for the 3-step ahead case, a pattern emerges, and

$$e_{t+h|t} = \sum_{i=0}^{h-1} \phi_1^i \epsilon_{t+h-i}$$

which shows that the h-step ahead forecast errors, $e_{t+h|t}$, follow a MA(h-1) process.

(b) **What is the autocorrelation function of a time-series of forecast errors $\{e_{t+h|t}\}$, $h = 1$ or $h = 2$. (Hint: Use the formula you derived above)**

The 1-step ahead case is uninteresting since they are a white noise process and their autocorrelations are all 0. The 2-step ahead case is fairly easy to derive using three consecutive forecast errors:

$$\begin{aligned} e_{t+2|t} &= \epsilon_{t+2} + \phi_1\epsilon_{t+1} \\ e_{t+1|t-1} &= \epsilon_{t+1} + \phi_1\epsilon_t \\ e_{t|t-2} &= \epsilon_t + \phi_1\epsilon_{t-1} \end{aligned}$$

Since forecast errors are mean zero (they are the sum of elements of a WN process), the autocovariance is simply the product,

$$\begin{aligned} \gamma_1 &= \mathrm{E}[e_{t+2|t}\, e_{t+1|t-1}] \\ &= \mathrm{E}[(\epsilon_{t+2} + \phi_1\epsilon_{t+1})(\epsilon_{t+1} + \phi_1\epsilon_t)] \\ &= \mathrm{E}[\epsilon_{t+2}\epsilon_{t+1} + \phi_1\epsilon_{t+1}^2 + \phi_1\epsilon_{t+2}\epsilon_t + \phi_1^2\epsilon_{t+1}\epsilon_t] \end{aligned}$$

$$
\begin{aligned}
&= 0 + \phi_1 \mathrm{E}[\epsilon_{t+1}^2] + 0 + 0 \\
&= \phi_1 \sigma^2
\end{aligned}
$$

$$
\begin{aligned}
\gamma_2 &= \mathrm{E}[e_{t+2|t} e_{t|t-2}] \\
&= \mathrm{E}[(\epsilon_{t+2} + \phi_1 \epsilon_{t+1})(\epsilon_t + \phi_1 \epsilon_{t-1})] \\
&= \mathrm{E}[\epsilon_{t+2}\epsilon_t + \phi_1 \epsilon_{t+1}\epsilon_t + \phi_1 \epsilon_{t+2}\epsilon_{t-1} + \phi_1^2 \epsilon_{t+1}\epsilon_{t-1}] \\
&= 0 + 0 + 0 + 0 \\
&= 0
\end{aligned}
$$

(c) **Can you generalize the above to a generic h? (In other words, leave the solution as a function of h).**

By inspection, $\gamma_s = 0$ for any $s \geq 2$. The 3-step ahead case can be shown to produce $\gamma_1 = (\phi_1^3 + \phi_1)\sigma^2 = \phi_1(1 + \phi_1^2)\sigma^2$, $\gamma_2 = \phi_1^2\sigma^2$ and $\gamma_s = 0$ for any $s \geq 3$. The general form can be determined by inspection,

$$
\gamma_s = \begin{cases} \phi_1^s \sigma^2 \sum_{i=0}^{h-1-s} \phi_1^{2i} & \text{for } s \leq h-1 \\ 0 & \text{otherwise} \end{cases}
$$

(d) **How could you test whether the forecast has excess dependence using an ARMA model?**

Since the errors are at most MA($h-1$), a natural test is to regress the series of forecast errors on lagged forecast errors that are *at least* h periods on the past,

$$
\hat{e}_{t+h|t} = \phi_0 + \theta_1 \hat{e}_{t|t-h} + \theta_2 \hat{e}_{t-1|t-h-1} + \ldots + \theta_Q \hat{e}_{t-Q|t-h-Q} + \eta_t \tag{1}
$$

The null hypothesis is $H_0 : \phi_0 = \theta_1 = \theta_2 = \ldots \theta_Q = 0$. Note that the error, η_t, cannot be assumed to be a white noise process and may have an MA(h-1) structure. The easy way to fix this is to use a HAC covariance estimator in the regression where $h-1$ lags are used. This regression is often preferred since it is a special case of an augmented Mincer-Zarnowitz regression:

$$
y_{t+h} = \alpha + \beta \hat{y}_{t+h|t} + \boldsymbol{\gamma} \mathbf{z}_t + \eta_t
$$

becomes

$$
\begin{aligned}
y_{t+h} &= \alpha + 1 \cdot \hat{y}_{t+h|t} + \boldsymbol{\gamma} \mathbf{z}_t + \eta_t \\
y_{t+h} - \hat{y}_{t+h|t} &= \alpha + \boldsymbol{\gamma} \mathbf{z}_t + \eta_t \\
\hat{e}_{t+h|t} &= \alpha + \boldsymbol{\gamma} \mathbf{z}_t + \eta_t
\end{aligned}
$$

Replacing $\boldsymbol{\gamma}$ with $\boldsymbol{\theta}$, $\mathbf{z}_t$ with the lagged errors and α with ϕ_0, the regression in equation 1 can be seen to be a generalized MZ regression where the usual null on the forecast ($\hat{y}_{t+h|t}$) is enforced ($\beta = 1$).

An alternative method to fix the inference in the regression is to include the terms $\hat{e}_{t+h-1|t-1}$, $\hat{e}_{t+h-2|t-2}, \ldots, \hat{e}_{t+1|t-h+1}$ in the regression, although a different null and alternative are needed since the coefficients on these regressors *must not be tested.*