

MATH 545: INTRODUCTION TO TIME SERIES ANALYSIS

WINTER 2025

COINTEGRATION IN A SIMPLE MODEL

Suppose

$$X_t = X_{t-1} + Z_{1,t}$$
$$Y_t = 0.6X_t + Z_{2,t}.$$

Clearly $\{X_t\}$ is non-stationary; also,

$$Y_t - Y_{t-1} = 0.6(X_t - X_{t-1}) + Z_{2,t} - Z_{2,t-1} = 0.6Z_{1,t} + Z_{2,t} - Z_{2,t-1}$$

so $\{Y_t\}$ is non-stationary, being stationary only after differencing. The model encapsulates a long-run relationship between the two non-stationary series, in that $Y_t - 0.6X_t$ is stationary. Regressing Y_t on X_t allows the cointegrating relationship to be discovered.

Our task is to predict *changes* in Y_t on the basis of *past* values of the series, that is, data up to and including time point $t - 1$. We could try to model changes in Y_t via a linear model based on the stationary differenced series

$$(Y_t - Y_{t-1}) = \psi_0 + \psi_1(X_{t-1} - X_{t-2}) + \psi_2(Y_{t-1} - Y_{t-2}) + Z_t$$

or

$$Y_t^* = \psi_0 + \psi_1 X_{t-1}^* + \psi_2 Y_{t-1}^* + Z_t$$

say. However, this model does not exploit the long-run relationship between Y_t and X_t , so we add in a third term to make the *error correction model*

$$Y_t^* = \psi_0 + \psi_1 X_{t-1}^* + \psi_2 Y_{t-1}^* + \gamma(Y_{t-1} - 0.6X_{t-1}) + Z_t$$

that reflects the fact that if indeed there is a long-run relationship between $\{Y_t\}$ and $\{X_t\}$, then a positive discrepancy – occurring by chance – at time $t - 1$, so that $Y_{t-1} > 0.6X_{t-1}$, will have to be ‘corrected’ at time t . Consequently, the difference $Y_t - Y_{t-1}$ should be smaller (or more negative) than it would be if the equilibrium was observed at time $t - 1$ (ie $Y_{t-1} \approx 0.6X_{t-1}$).

Note that in fact any historical error correction term, that is

$$(Y_{t-j} - \alpha_0 - \alpha_1 X_{t-j}) \quad j \geq 1$$

could be used in this model to act as an error corrector; the calculation is based on the assumption that the sequence of error correction terms is stationary, and at time t all the historical terms are available.

Implementation

```
set.seed(1231456)
n<-1000
z1<-rnorm(n)
z2<-rnorm(n)
x<-cumsum(z1);
y<-0.6*x+z2
```

We first fit the model

$$Y_t^* = \psi_0 + \psi_1 X_{t-1}^* + \psi_2 Y_{t-1}^* + Z_t \tag{1}$$

```
ystar<-diff(y)[-1]
ystar.1<-diff(y)[-n-1]
xstar.1<-diff(x)[-n-1]
fit0<-lm(ystar~xstar.1+ystar.1)
summary(fit0$coef)
##      Min.    1st Qu.    Median      Mean    3rd Qu.      Max.
## -0.488897 -0.248162 -0.007427 -0.050976  0.167985  0.343397
#R-squared
summary(fit0)$r.squared
## [1] 0.2090031
```

This shows that the differenced series are useful in forecasting the change in Y_t . Now we try the error correction model assuming the known relationship between $\{X_t\}$ and $\{Y_t\}$:

$$Y_t^* = \psi_0 + \psi_1 X_{t-1}^* + \psi_2 Y_{t-1}^* + \gamma(Y_{t-1} - 0.6X_{t-1}) + Z_t \quad (2)$$

where $(Y_{t-1} - 0.6X_{t-1}) = Z_{2,t}$.

```
error<-z2; error.lagged<-error[-c(1,n)]
xstar<-diff(x); ystar<-diff(y)
diff.dat<-data.frame(embed(cbind(xstar,ystar),2))
colnames(diff.dat)<-c("xstar","ystar","xstar.1","ystar.1")
ecm.reg0<-lm(ystar~xstar.1+ystar.1+error.lagged,data=diff.dat)
round(summary(ecm.reg0)$coef,6)

##              Estimate Std. Error    t value Pr(>|t|)
## (Intercept) -0.088895   0.037175  -2.391272 0.016975
## xstar.1      0.012735   0.042330   0.300859 0.763585
## ystar.1      0.018745   0.036017   0.520442 0.602872
## error.lagged -1.019659   0.050902 -20.031906 0.000000

#R-squared
summary(ecm.reg0)$r.squared

## [1] 0.4364913
```

This analysis clearly shows that the error correction term is influential, with

$$\hat{\gamma} = -1.01966$$

with standard error 0.05090.

We now consider estimating the model

$$Y_t = \alpha_0 + \alpha_1 X_t + Z_t$$

to estimate the cointegrating vector and repeating the analysis

```
lr.reg<-lm(y~x)
summary(lr.reg)$coef

##              Estimate Std. Error    t value Pr(>|t|)
## (Intercept) -0.1815281 0.058808321  -3.086776 0.00207923
## x            0.5913306 0.004196762 140.901641 0.00000000
```

We obtain the residuals $\{\hat{z}_t^*\}$ from the fit, and proceed with the model

$$Y_t^* = \psi_0 + \psi_1 X_{t-1}^* + \psi_2 Y_{t-1}^* + \gamma \hat{z}_{t-1}^* + Z_t \quad (3)$$

```
error<-residuals(lr.reg)
error.lagged<-error[-c(1,n)]
xstar<-diff(x); ystar<-diff(y)
diff.dat<-data.frame(embed(cbind(xstar,ystar),2))
colnames(diff.dat)<-c("xstar","ystar","xstar.1","ystar.1")
ecm.reg1<-lm(ystar~xstar.1+ystar.1+error.lagged,data=diff.dat)
round(summary(ecm.reg1)$coef,6)

##              Estimate Std. Error    t value Pr(>|t|)
## (Intercept) -0.007139   0.036830  -0.193850 0.846333
## xstar.1      0.012904   0.042136   0.306246 0.759482
## ystar.1      0.025036   0.035974   0.695947 0.486625
## error.lagged -1.031839   0.050928 -20.260887 0.000000

#R-squared
summary(ecm.reg1)$r.squared

## [1] 0.440193
```

Note that the R-squared quantities for the models in equations (2) and (3) are higher than that for (1), indicating the predictive advantage in adding the error correction term. This analysis clearly shows that the error correction term is influential, with

$$\hat{\gamma} = -1.03184$$

with standard error 0.05093.

Finally, for illustration, we demonstrate that any historical error correction term may be included, for example, z_{t-2}^* rather than z_{t-1}^* , using the model

$$Y_t^* = \psi_0 + \psi_1 X_{t-1}^* + \psi_2 Y_{t-1}^* + \gamma z_{t-2}^* + Z_t \quad (4)$$

```
##           Estimate Std. Error  t value Pr(>|t|)
## (Intercept) -0.007139   0.036830  -0.19385 0.846333
## xstar.1      0.623062   0.041230  15.11177 0.000000
## ystar.1     -1.006803   0.036112 -27.87975 0.000000
## error.lagged -1.031839   0.050928 -20.26089 0.000000
## [1] 0.440193
```

This produces an identical fit for γ , although the estimates for the other regression coefficients change.

Theoretical analysis of the Error Correction Model: Considering the model in (2), we have by standard theory that the least-squares estimators converge to

$$\mathbb{E}[\mathbf{X}^\top \mathbf{X}]^{-1} \mathbb{E}[\mathbf{X}^\top Y]$$

where $Y \equiv Y_t^*$, and

$$\mathbf{X} = [1, X_{t-1}^*, Y_{t-1}^*].$$

is formed from the stationary series. Now

$$\mathbf{X}^\top \mathbf{X} = \begin{bmatrix} 1 & X_{t-1}^* & Y_{t-1}^* \\ X_{t-1}^* & (X_{t-1}^*)^2 & X_{t-1}^* Y_{t-1}^* \\ Y_{t-1}^* & X_{t-1}^* Y_{t-1}^* & (Y_{t-1}^*)^2 \end{bmatrix}$$

and we have immediately that $\mathbb{E}[X_{t-1}^*] = \mathbb{E}[Y_{t-1}^*] = 0$. Also, if $\phi = 0.6$, we have

$$\mathbb{E}[(X_{t-1}^*)^2] = \mathbb{E}[(X_{t-1} - X_{t-2})^2] = \mathbb{E}[Z_{1,t-1}^2] = \sigma_1^2$$

$$\mathbb{E}[X_{t-1}^* Y_{t-1}^*] = \mathbb{E}[(X_{t-1} - X_{t-2})(\phi(X_{t-1} - X_{t-2}) + Z_{2,t-1} - Z_{2,t-1})] = \phi \mathbb{E}[Z_{1,t-1}^2] = \phi \sigma_1^2$$

$$\mathbb{E}[(Y_{t-1}^*)^2] = \mathbb{E}[(\phi(X_{t-1} - X_{t-2}) + Z_{2,t-1} - Z_{2,t-1})^2] = \mathbb{E}[(\phi Z_{1,t-1} + Z_{2,t-1} - Z_{2,t-1})^2] = \phi^2 \sigma_1^2 + 2\sigma_2^2.$$

Thus

$$\mathbb{E}[\mathbf{X}^\top \mathbf{X}] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \sigma_1^2 & \phi \sigma_1^2 \\ 0 & \phi \sigma_1^2 & \phi^2 \sigma_1^2 + 2\sigma_2^2 \end{bmatrix} \quad \therefore \quad \mathbb{E}[\mathbf{X}^\top \mathbf{X}]^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{\phi^2}{2\sigma_2^2} + \frac{1}{2\sigma_1^2} & -\frac{\phi}{2\sigma_2^2} \\ 0 & -\frac{\phi}{2\sigma_2^2} & \frac{1}{2\sigma_2^2} \end{bmatrix}$$

Also

$$\mathbb{E}[\mathbf{X}^\top Y] = \begin{bmatrix} Y_t^* \\ Y_t^* X_{t-1}^* \\ Y_t^* Y_{t-1}^* \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -\sigma_2^2 \end{bmatrix}$$

as

$$\mathbb{E}[Y_t^* Y_{t-1}^*] = \mathbb{E}[(\phi Z_{1,t} + Z_{2,t} - Z_{2,t-1})(\phi Z_{1,t-1} + Z_{2,t-1} - Z_{2,t-2})] = -\mathbb{E}[Z_{2,t-1}^2]$$

so

$$\mathbb{E}[\mathbf{X}^\top \mathbf{X}]^{-1} \mathbb{E}[\mathbf{X}^\top Y] = \begin{bmatrix} 0 \\ \frac{\phi}{2} \\ -\frac{1}{2} \end{bmatrix}$$

For the model in (3), we have

$$\mathbf{X} = [1, X_{t-1}^*, Y_{t-1}^*, Z_{2,t-1}].$$

Now we have

$$\mathbf{X}^\top \mathbf{X} = \begin{bmatrix} 1 & X_{t-1}^* & Y_{t-1}^* & Z_{2,t-1} \\ X_{t-1}^* & (X_{t-1}^*)^2 & X_{t-1}^* Y_{t-1}^* & X_{t-1}^* Z_{2,t-1} \\ Y_{t-1}^* & X_{t-1}^* Y_{t-1}^* & (Y_{t-1}^*)^2 & Y_{t-1}^* Z_{2,t-1} \\ Z_{2,t-1} & X_{t-1}^* Z_{2,t-1} & Y_{t-1}^* Z_{2,t-1} & (Z_{2,t-1})^2 \end{bmatrix}$$

where

$$\mathbb{E}[X_{t-1}^* Z_{2,t-1}] = 0$$

$$\mathbb{E}[Y_{t-1}^* Z_{2,t-1}] = \mathbb{E}[(\phi(X_{t-1} - X_{t-2}) + Z_{2,t-1} - Z_{2,t-2})Z_{2,t-1}] = \mathbb{E}[(\phi Z_{1,t-1} + Z_{2,t-1} - Z_{2,t-2})Z_{2,t-1}] = \sigma_2^2$$

Thus

$$\mathbb{E}[\mathbf{X}^\top \mathbf{X}] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \sigma_1^2 & \phi\sigma_1^2 & 0 \\ 0 & \phi\sigma_1^2 & \phi^2\sigma_1^2 + 2\sigma_2^2 & \sigma_2^2 \\ 0 & 0 & \sigma_2^2 & \sigma_2^2 \end{bmatrix} \quad \therefore \quad \mathbb{E}[\mathbf{X}^\top \mathbf{X}]^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{\phi^2}{\sigma_2^2} & -\frac{\phi}{\sigma_2^2} & \frac{\phi}{\sigma_2^2} \\ 0 & -\frac{\phi}{\sigma_2^2} & \frac{1}{\sigma_2^2} & -\frac{1}{\sigma_2^2} \\ 0 & \frac{\phi}{\sigma_2^2} & -\frac{1}{\sigma_2^2} & \frac{2}{\sigma_2^2} \end{bmatrix}$$

Also

$$\mathbb{E}[\mathbf{X}^\top Y] = \begin{bmatrix} Y_t^* \\ Y_t^* X_{t-1}^* \\ Y_t^* Y_{t-1}^* \\ Y_t^* Z_{2,t-1} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -\sigma_2^2 \\ -\sigma_2^2 \end{bmatrix}$$

as

$$\mathbb{E}[Y_t^* Z_{2,t-1}] = \mathbb{E}[(\phi Z_{1,t} + Z_{2,t} - Z_{2,t-1})Z_{2,t-1}] = -\mathbb{E}[Z_{2,t-1}^2].$$

Therefore

$$\mathbb{E}[\mathbf{X}^\top \mathbf{X}]^{-1} \mathbb{E}[\mathbf{X}^\top Y] = \begin{bmatrix} 0 \\ 0 \\ 0 \\ -1 \end{bmatrix}$$