Time series

Gabriela Ciołek

The model of the ordinary least squares is considered under the following hypotheses:

- H1 : The model is linear in $x_{i,t}$
- H2 : the values $x_{i,t}$ are observed without error
- H3 : $\mathbb{E}[u] = 0$ the mean of the error is equal to zero
- H4: $\mathbb{E}[u^t u] = \sigma^2 I_n$, the variance of the error is constant
- H5 : $\mathbb{E}[u_t u_{t+1}]$, the errors are not correlated
- H6 : $Cov(x_{i,t}, u_t)$, the error is independent of the explanatory variable

- The violation of hypothesis $H5$ concerns time series where the off-diagonal elements of the covariance matrix of the errors are nonzero.
- In this case, the obtained OLS estimators are unbiased but no longer have a minimal variance.
- We need to identify new estimators and techniques for detecting possible autocorrelation of the errors.

- The residuals of the regression can be correlated in series.
- The nature of dependence between residuals needs to be tested.
- The most popular and simple model to be tested is the AR (1) model (other models can be tried, we start with the simplest try).
- IN TP6 we test the presence of correlations in series of type AR (1).

- Assume that the residuals are given by

$$u_t = \rho u_{t-1} + e_t$$

- Assume that $|\rho| < 1$ (statbility conditon), and that e_t are independent, zero mean random variables with variance σ_e^2.
- In the model AR(1), the null hypothesis H_0 assumes that the errors are not correlated in the series

$$H_0 : \rho = 0.$$

- We do a regression

$$\hat{u}_t = \rho \hat{u}_{t-1} + e_t$$

  and recover $\hat{\rho}$
- We calculate t-statistic $t_{\hat{\rho}}$.

GLS Method

- If we reject H_0, we need to transform the data in order to apply the generalized least squares method.
- If we assume the hypothesis of an autocorrelation of order 1, the model is written:

$$Y = X\beta + u$$

$$u_t = \rho u_{t-1} + e_t,$$

where e_t has zero mean and variance σ_e^2.

Data transformation

$Var(u_t) = \frac{\sigma_e^2}{1-\rho^2} = \sigma_u^2$ and $\mathbb{E}[u_t u_{t-i}] = \rho^i \sigma_u^2$. Thus, the matrix of the variance of residuals u is given by

$$\Omega_u = \mathbb{E}[u^t u] = \frac{\sigma_e^2}{1-\rho^2} \times$$

$$\times \begin{pmatrix} 1 & \rho & \rho^2 & \dots & \rho^{T-1} \\ \rho & 1 & & & \\ \rho^2 & & & & \\ \dots & & & & \\ \dots & & & & \\ \rho^{T-1} & & & & 1 \end{pmatrix}$$

- We denote by P be the matrix such that $\Omega_u^{-1} = P'P$ (decomposition of Cholesky).
- Suppose that Ω has the eigenvalues $\lambda_1, \cdots, \lambda_T$.
- By Cholesky's decomposition we obtain

$$\Omega = S\Lambda S',$$

where Λ is a diagonal matrix with the diagonal elements $(\lambda_1, \cdots, \lambda_T)$ and S is an orthogonal matrix.

- We have

$$\begin{aligned}\Omega^{-1} &= S^{-1}\Lambda^{-1}S'^{-1}\\ &= S^{-1}\Lambda^{-1/2}\Lambda'^{-1/2}S'^{-1}\\ &= PP',\end{aligned}$$

where $P = S^{-1}\Lambda^{-1/2}$ and $\Lambda^{-1/2}$ is a diagonal matrix with the diagonal elements $(\sqrt{\lambda_1}, \cdots, \sqrt{\lambda_T})$.

- It can be proved that $P\Omega P' = I_T$..
- Multiply $y = X\beta + u$ by P and get

$$Py = PX\beta + Pu.$$

- Set $y^0 = Py, \; X^0 = PX$ and $u^0 = Pu$.
- We retrieved the classical linear regression model

$$y^0 = X^0\beta + u^0,$$

where $\mathbb{E}[u^0] = 0$ and $\mathbb{E}[u^0 u^{0\prime}] = \sigma^2 P\Omega P' = \sigma^2 I_T$.

The GLS estimator of β is given by

$$\begin{aligned}\hat{\beta}_{GLS} &= (X^{0\prime}X^{0})^{-1}X^{0\prime}y^{0}\\ &= (X'P'PX)^{-1}X'P'Py\\ &= (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}y.\end{aligned}$$

The model $PY = PX + Pu$ possess independent and homoscedastic errors. P is given by:

$$P = \begin{pmatrix} \sqrt{1-\rho^2} & 0 & 0 & \cdots & & 0 \\ -\rho & 1 & 0 & & & 0 \\ 0 & -\rho & 1 & & & 0 \\ \cdots & & & & & \\ \cdots & & & & & 0 \\ 0 & & & & -\rho & 1 \end{pmatrix}$$

Granger causality test

- Causal effect: ceterus paribus change in one variable has an effect on another variable
- A variable X is causal to variable Y if X is the cause of Y or Y is the cause of X.
- Granger causality is a limited notion of causality, where past values of one series (x_t) are useful for predicting future values of (y_t) after past values of y_t have been controlled for.

Granger causality

Consider two series y_t and z_t such that

$$y_t = \delta_0 + \alpha_1 y_{t-1} + \gamma_1 z_{t-1} + \alpha_2 y_{t-2} + \gamma_2 z_{t-2} + \cdots$$

and

$$z_t = \eta_0 + \beta_1 y_{t-1} + \rho_1 z_{t-1} + \beta_2 y_{t-2} + \rho_2 z_{t-2} + \cdots ,$$

where each equation contains an error that has zero expected value given past information on y and z. Such equations allow us to test whether, after controlling for past y, past z help to forecast y_t.

Generally, we say that z **Granger causes** y if

$$E(y_t|I_{t-1}) \neq E(y_t|J_{t-1}), \tag{1}$$

where I_{t-1} contains past information on y and z, and J_{t-1} contains only information on past y. When (1) holds, past z is useful, in addition to past y, for predicting y_t.

Once we assume a linear model and decide how many lags of y should be included in $E(y_t|y_{t-1}, y_{t-2}, \cdots)$, we can easily test the null hypothesis that z does not Granger cause y. To be more specific, suppose that $E(y_t|y_{t-1}, y_{t-2}, \cdots)$ depends on only three lags:

$$y_t = \delta_0 + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \alpha_3 y_{t-3} + u_t$$

$$E(u_t|y_{t-1}, y_{t-2}, \cdots) = 0.$$

Now, under the null hypothesis that z does not Granger cause y, any lags of z that we add to the equation should have zero population coefficients. If we add z_{t-1}, then we can simply do a t test on z_{t-1}. If we add two lags of z, then we can do F test for joint significance of z_{t-1} and z_{t-2} in the equation

$$y_t = \delta_0 + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \alpha_3 y_{t-3} + \gamma_1 z_{t-1} + \gamma_2 z_{t-2} + u_t.$$

In practice we need to decide on which lags of y and z to include:

- start by estimating an autoregressive model for y and performing t and F tests to determine how many lags of y should appear
- with annual data, the number of lags is typically small, say one or two
- with quarterly or monthly data, there are usually many more lags
- once an autoregressive model for y has been chosen, we can test for lags of z
- the choice of lags of z is less important because, when z does not Granger cause y, no set of lagged z's should be significant

Exercise 1

- Consider the model (see TP5):

$$i3t = \beta_0 + \beta_1 inft + \beta_2 deft + u_t$$

-
- Apply the transformation $Y \to PY$ and calculate the series $y_t - y_{t-1}$ for $t = 2, \cdots, T$ and $\sqrt{1-\rho^2}y_1$ for the first observation.
- Do the same for X.
- Obtain OLS estimators for the transformed data

Exercise 2

- Consider the model:

$$i3_t = \beta_0 + \beta_1 inf_{t-1} + \beta_2 inf_{t-2} + \beta_3 def_{t-1} + \beta_4 def_{t-2} + u_t$$

  for $t = 1, \cdots, n$.
- Test the significance of the inflation and deficit (Granger causality test): $H_0 : \beta_1 = \beta_2 = 0$ then $H_0 : \beta_3 = \beta_4 = 0$.

$$y = intdef(:, 2);$$

$$[n, k] = size(intdef)$$

$$X = [intdef(:, [3, 6])]; \textit{without constants, because we have dynamic model}$$

$$y = intdef(3 : n, 2);$$

$$X\_lag = [X(2 : n - 1, :)]$$

$$X\_lag2 = [X(1 : n - 2, :)]$$

$$X = [X\_lagX\_lag2]$$

$$[n, k] = size(X)$$

$$beta = inv(X' * X) * X' * y$$

$$u = y - X * beta;$$

Unrestricted model:

$$SSR0 = u' * u;$$

Restricted model (the inflation is removed):

$$X = X(:, [2, 4]);$$

$$\beta 1 = inv(X' * X) * X' * y;$$

$$u1 = y - X * \beta 1;$$

$$SSR1 = u1' * u1;$$

$$F = ((SSR1 - SSR0)/SSR0) * ((n - k)/2);$$

$$p = fdis\_prb(F, 2, n - k)$$

Restricted model (the deficit is removed)

$$X = [X\_lagX\_lag2];$$

$$X = X(:, [1, 3]);$$

$$beta1 = inv(X' * X) * X' * y;$$

$$u1 = y - X * beta1;$$

$$SSR1 = u1' * u1;$$

$$F = ((SSR1 - SSR0)/SSR0) * ((n - k)/2);$$

$$p = fdis\_prb(F, 2, n - k)$$

We obtain respectively $p = 1.1491 - 14$ and $p = 1.4639 - 4$ so we reject H_0 each time.