Erasmus
School of
Economics

MOOC Econometrics

Lecture 6.3 on Time Series: Specification and Estimation

Dick van Dijk, Philip Hans Franses, Christiaan Heij

Erasmus University Rotterdam

Forecasting

- Past values of time series → Model → Forecast future values
- Notation:

  y_t: time series of interest $(t = 1, \ldots, n)$

  x_t: time series possible explanatory factor (restrict to one)

  $PY_{t-1} = \{y_{t-1}, y_{t-2}, \ldots, y_1\}$: past information on y at time t

  $PX_{t-1} = \{x_{t-1}, x_{t-2}, \ldots, x_1\}$

  Univariate time series forecast model: $\widehat{y}_t = F(PY_{t-1})$

  Forecast model with explanatory factor: $\widehat{y}_t = F(PY_{t-1}, PX_{t-1})$
- Aim: Optimal use of past information to get best forecasts.
- Wish: Forecast error $\varepsilon_t = y_t - \widehat{y}_t$ uncorrelated with past information.

Univariate time series model

- Forecast: $\widehat{y}_t = F(PY_{t-1})$ where $PY_{t-1} = \{y_{t-1}, y_{t-2}, \ldots, y_1\}$.
- Find forecast model F so that $\varepsilon_t = y_t - \widehat{y}_t$ uncorrelated with PY_{t-1}.
- Popular choice: F linear function of p past values:

  $\widehat{y}_t = \alpha + \beta_1 y_{t-1} + \ldots + \beta_p y_{t-p}$.
- $y_t = \widehat{y}_t + \varepsilon_t = \alpha + \beta_1 y_{t-1} + \ldots + \beta_p y_{t-p} + \varepsilon_t$.
- AR(p) model, because ε_t is white noise.

Test question

- Forecast: $\widehat{y}_t = \alpha + \beta_1 y_{t-1} + \ldots + \beta_p y_{t-p}$.
- Forecast error $\varepsilon_t = y_t - \widehat{y}_t$ uncorrelated with y_s for all $s < t$.

Test

Show that ε_t is white noise, i.e., ε_t is uncorrelated with ε_s for all $t \neq s$.

Answer:

- Without loss of generality, consider case $s < t$.
- $\varepsilon_s = y_s - \alpha - \sum_{j=1}^{p} \beta_j y_{s-j}$ linear function of y_r, $r \leq s < t$.
- ε_t is uncorrelated with y_r for all $r < t$, so also uncorrelated with ε_s.

Estimation

- Forecast error: $\varepsilon_t = y_t - \alpha - \sum_{j=1}^{p} \beta_j y_{t-j}$.
- Minimize sum of squared forecast errors: $\sum_{t=p+1}^{n} \varepsilon_t^2$.
- OLS!
- Estimation of ARMA models: Maximum Likelihood.

Time series model with explanatory factor

- Forecast: $\widehat{y}_t = F(PY_{t-1}, PX_{t-1})$.
- Find F such that $\varepsilon_t = y_t - \widehat{y}_t$ uncorrelated with PY_{t-1} and PX_{t-1}.
- Popular choice: linear F:

  $\widehat{y}_t = \alpha + \beta_1 y_{t-1} + \ldots + \beta_p y_{t-p} + \gamma_1 x_{t-1} + \ldots + \gamma_r x_{t-r}$.
- $y_t = \widehat{y}_t + \varepsilon_t = \alpha + \sum_{j=1}^{p} \beta_j y_{t-j} + \sum_{j=1}^{r} \gamma_j x_{t-j} + \varepsilon_t$.

  Autoregressive Distributed Lag model: ADL(p, r).
- Estimation: minimize $\sum_{t=m+1}^{n} \varepsilon_t^2$, where $m = \max(p, r) \rightarrow$ OLS!

Granger causality

- Two variables of interest: y_t and x_t.
- Make ADL model for each variable:

  $y_t = \alpha + \sum_{j=1}^{p} \beta_j y_{t-j} + \sum_{j=1}^{r} \gamma_j x_{t-j} + \varepsilon_t$.

  $x_t = \alpha^* + \sum_{j=1}^{p^*} \beta_j^* x_{t-j} + \sum_{j=1}^{r^*} \gamma_j^* y_{t-j} + \varepsilon_t^*$.
- x_t helps to predict y_t if $\gamma_j \neq 0$ for some j

  y_t helps to predict x_t if $\gamma_j^* \neq 0$ for some j
- x_t is Granger causal for y_t if it helps to predict y_t, whereas y_t does not help to predict x_t.
- Test $H_0 : \gamma_j^* = 0$ for all $j = 1, \ldots, r^*$ by means of F-test.
- Note: Two ADL equations are estimated by OLS per equation.

Consequences of non-stationarity

- Regression assumption A2 not satisfied: regressors y_{t-j} are random.
- Standard OLS t- and F-tests hold true in large enough samples provided all variables in equation are stationary.
- So: First test for non-stationarity before any estimation.
- AR(1): $y_t = \alpha + \beta y_{t-1} + \varepsilon_t$, test $H_0 : \beta = 1$ against $H_1 : -1 < \beta < 1$.
- Rewrite: $\Delta y_t = y_t - y_{t-1} = \alpha + (\beta - 1) y_{t-1} + \varepsilon_t = \alpha + \rho y_{t-1} + \varepsilon_t$

  where $\rho = \beta - 1$
- So: $\Delta y_t = \alpha + \rho y_{t-1} + \varepsilon_t$, test $H_0 : \rho = 0$ against $H_1 : \rho < 0$.
- Reject H_0 of non-stationarity if $t_{\widehat{\rho}} < -2.9$ (not conventional -1.65!).

Test question

Test

Rewrite the AR(2) model $y_t = \alpha + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \varepsilon_t$ as $\Delta y_t = \delta + \rho y_{t-1} + \gamma \Delta y_{t-1} + \varepsilon_t$, and express the parameters (δ, ρ, γ) in terms of $(\alpha, \beta_1, \beta_2)$.

Answer:

- $$\begin{aligned}\Delta y_t &= y_t - y_{t-1}\\ &= \alpha + (\beta_1 - 1) y_{t-1} + \beta_2 y_{t-2} + \varepsilon_t\\ &= \alpha + (\beta_1 + \beta_2 - 1) y_{t-1} - \beta_2 y_{t-1} + \beta_2 y_{t-2} + \varepsilon_t\\ &= \alpha + (\beta_1 + \beta_2 - 1) y_{t-1} - \beta_2 (y_{t-1} - y_{t-2}) + \varepsilon_t\\ &= \alpha + (\beta_1 + \beta_2 - 1) y_{t-1} - \beta_2 \Delta y_{t-1} + \varepsilon_t\end{aligned}$$
- So: $\delta = \alpha$, $\rho = \beta_1 + \beta_2 - 1$, and $\gamma = -\beta_2$.

Augmented Dicky-Fuller test

- Two types of test equations: with or without deterministic trend.
- Test without deterministic trend if data no clear trend direction:
  $$\Delta y_t = \alpha + \rho y_{t-1} + \gamma_1 \Delta y_{t-1} + \ldots + \gamma_L \Delta y_{t-L} + \varepsilon_t$$
- Reject H_0 of non-stationarity if $\boxed{t_{\widehat{\rho}} < -2.9}$
- Test with deterministic trend if data clear trend direction:
  $$\Delta y_t = \alpha + \beta t + \rho y_{t-1} + \gamma_1 \Delta y_{t-1} + \ldots + \gamma_L \Delta y_{t-L} + \varepsilon_t$$
- Reject H_0 of non-stationarity if $\boxed{t_{\widehat{\rho}} < -3.5}$
- Choice lag L: serial correlation check, or AIC/BIC (see Lecture 3).

Summary of Specification and Estimation

- AR model for y_t:

  Step 1: Perform ADF test on y_t.
  → Non-stationarity rejected → model y_t
  → Non-stationarity not rejected → take Δy_t and perform ADF test on Δy_t

  Step 2: Estimate AR model for stationary series by OLS.
- ADL model for y_t with explanatory factor x_t:

  Step 1: Perform ADF tests on y_t and x_t.

  → Take difference until non-stationarity is rejected.

  Step 2: Estimate ADL model for stationary series by OLS.
- One exception: if x_t and y_t are cointegrated.

Cointegration and error correction model

- x_t and y_t are cointegrated if both series are non-stationary, but a linear combination (say $y_t - cx_t$) is stationary.
- $y_t = cx_t$: long-run equilibrium.
- Engle-Granger test for cointegration:
  → Step 1: OLS in $y_t = \alpha + \beta x_t + \varepsilon_t \quad \rightarrow \quad b$ and residuals e_t
  → Step 2: Cointegrated if ADF test on e_t rejects non-stationarity
  $$\Delta e_t = \alpha + \rho e_{t-1} + \gamma_1 \Delta e_{t-1} + \ldots + \gamma_L \Delta e_{t-L} + \omega_t$$
  Critical value $t_{\widehat{\rho}}$: $\underline{-3.4}$ (if extra term βt: $\underline{-3.8}$)
- Error Correction Model (ECM): if x_t and y_t cointegrated, estimate
  $$\Delta y_t = \alpha + \beta_1 (y_{t-1} - bx_{t-1}) + \beta_2 \Delta y_{t-1} + \beta_3 \Delta x_{t-1} + \varepsilon_t$$
  (or more lags for Δy_t and Δx_t)

TRAINING EXERCISE 6.3

- Train yourself by making the training exercise (see the website).
- After making this exercise, check your answers by studying the webcast solution (also available on the website).