Lecture 7 - Forecasting with VEC Models

Pedro Valls[1]

[1]Sao Paulo School of Economics - FGV
CEQEF-FGV

July 26, 2024

FGV EESP

CENTRO DE ESTUDOS QUANTITATIVOS EM ECONOMIA E FINANÇAS

Table of Contents

Outline

- Introduction
- Spurious regression
- Cointegration and error correction models
- Engle and Granger cointegration test
- Johansen cointegration test
- MA representation of cointegrated processes
- Example using simulated series
- Forecasting in the presence of cointegration
- The effect of stochastic trends on forecast
- Example using simulated series

Table of Contents

Introduction

- Integrated variables can be made stationary by differencing
- However in a multivariate context, there also exists the possibility that linear combinations of integrated variables are stationary, this is the case of cointegration
- Specification of cointegrated processes were implicit in the so-called error correction models proposed by [Davidson et al., 1978].
- Cointegration was introduced in a series of papers by [Granger, 1983], [Granger and Weiss, 1983] and [Engle and Granger, 1987].
- We consider the consequences of cointegration for modelling and forecasting.

Table of Contents

Spurious Regression I

- [Granger and Newbold, 1974] considered two unrelated random walks y_{1t} and y_{2t} such that

$$y_{it} = y_{it-1} + \varepsilon_{it} \quad \text{where} \quad \varepsilon_{it} \sim WN(0,1) \quad i = 1,2 \tag{1}$$

- and they are independent of each other
- And we regress y_{1t} onto y_{2t} namely:

$$y_{2t} = \beta_0 + \beta_1 y_{1t} + u_t \tag{2}$$

- Simulation evidence reporter by [Granger and Newbold, 1974] featured properties that later were formalized by [Phillips, 1986] in particular:
  - $\widehat{\beta}_1$ does not converge in probability to zero but instead converges in distribution to a non-normal random variable not necessarily centered at zero.

Spurious Regression II

- The usual $t-$statistics for testing $\beta_1 = 0$ diverges as $T \to \infty$
- The usual R^2 from the regression converges to unity as $T \to \infty$

- But if we regress

$$d(y_{2t}) = \alpha_0 + \alpha_1 d(y_{1t}) + v_t \tag{3}$$

- simulation evidence
  - $\widehat{\alpha}_1$ converge in probability to zero and the distribution will converges to $N(0,1)$ as $T \to \infty$
  - The usual $t-$statistics for testing $\beta_1 = 0$ converges to $N(0,1)$ as $T \to \infty$
  - The usual R^2 from the regression converges to zero as $T \to \infty$
- Use the following commands in R
- Sprious Regression

Table of Contents

Cointegration and error correction models I

- We can write a $VAR(p)$ in levels of y_t

$$\Phi(L)y_t = \varepsilon_t \quad \text{with} \quad \varepsilon_t \sim NI(0, \Sigma) \tag{4}$$

as

$$\Delta y_t = \underbrace{-\left(I - \Phi_1 - \cdots - \Phi_p\right)}_{\Pi} y_{t-1} + \Gamma_1 \Delta y_{t-1} + \cdots + \Gamma_{p-1} \Delta y_{t-p+1} + \varepsilon_t \tag{5}$$

where

$$\Gamma_i = -(\Phi_{i+1} + \cdots + \Phi_p) \tag{6}$$

- When y_t is weakly stationary we know that $|\Phi(z)| = 0$ has all the roots outside the unit circle.

Cointegration and error correction models II

- In the presence of unit roots i.e. $|\Phi(1)| = 0$ the matrix $\Phi(1) = \Pi$ has reduced rank.
- So the rank of Π will determine if the variables are integrated and cointegrated.
- If the rank is r, with $0 \leq r \leq m$.
- When $r = 0$ equation (5) reduces to

$$\Delta y_t = \Gamma_1 \Delta y_{t-1} + \cdots + \Gamma_{p-1} \Delta y_{t-p+1} + \varepsilon_t \tag{7}$$

  and the valid representation is a $VAR(p-1)$ in first differences.
- When $r = m$ all the components of y_t are stationary and the valid representation is (4) a $VAR(p)$ in level

Cointegration and error correction models III

- When $0 < r < m$ we can write Π as

$$\Pi = \underset{m\times r}{\alpha} \cdot \underset{r\times m}{\beta'} \tag{8}$$

where β contains the coefficients of the r independent stationary combination of the m integrated variables, known as cointegrated vector

$$\beta' y_{t-1} \sim I(0) \tag{9}$$

and α is the load matrix given the importance of each cointegration relationship in each equation.

- In this case we can rewrite (5) in a Vector Error Correction Model (VECM) given by

$$\Delta y_t = -\alpha\beta' y_{t-1} + \Gamma_1 \Delta y_{t-1} + \cdots + \Gamma_{p-1}\Delta y_{t-p+1} + \varepsilon_t \tag{10}$$

Cointegration and error correction models IV

- Some comments about equations (4), (7) and (10)
  - If the components of y_t are non-stationary and we estimate (4) OLS estimates of the coefficients remains consistent but the asymptotic distribution are non-standard.
  - If the components of y_t are non-stationary and co-integrate and we estimate (7) the model is mispecified so OLS will be inconsistent.
  - If the components of y_t are non-stationary and co-integrate and we estimate (4) the estimates are consistent but inefficient

Table of Contents

Engle and Granger cointegration test I

- [Engle and Granger, 1987] suggest the following two-stage procedure to test cointegration.
    - First estimate by OLS the regression:

$$y_{1t} = \beta_2 y_{2t} + \cdots + \beta_m y_{mt} + \epsilon_t \tag{11}$$

    - Second test whether ϵ_t has a unit root.
        - If it does there is no cointegration
        - If it does not then the variables are cointegrated.
- The critical values for the test of unit root are different from the Dickey & Fuller Test (see [MacKinnon, 1991] or [MacKinnon, 1996])

Table of Contents

Johansen cointegration test I

- In order to apply the Johansen procedure we need the additional assumption that the VAR error are normally distributed
- It is a sequentially tests for the following hypotheses:

$$\begin{array}{ccc} & H_0 & H_1 \\ (1) & r=0 & r=1 \\ (2) & r\leq 1 & r=2 \\ (3) & r\leq 2 & r=3 \\ \vdots & \vdots & \vdots \\ (m-1) & r\leq m-1 & r=m \end{array}$$

- Two test statistics are use:
  a. the trace test; and
  b. the maximum eigenvalue test

both with an asymptotic $\aleph^2$ distribution.

Johansen cointegration test II

- The cointegration coefficients and the loading are not uniquely identified.
- We have to impose a priori restrictions on the coefficients of α and β
- For example it can be shown that if

$$\beta' = \left[\underset{r \times r}{I} : \widetilde{\beta} \right] \tag{12}$$

- then α and $\widetilde{\beta}$ are exactly identified

Table of Contents

MA representation of cointegrated processes I

- A stationary process admits a $MA(\infty)$ representation

$$\Delta y_t = C(L)\varepsilon_t \tag{13}$$

and using the same steps as in [Beveridge and Nelson, 1981] decomposition we have

$$y_t = \underbrace{C(1)\sum_{j=1}^{t}\varepsilon_t}_{PC_t} + \underbrace{C^*(L)\varepsilon_t}_{TC_t} \tag{14}$$

where PC_t and TC_t stand for permanent and transitory components respectively and we have the following relationship between (13) and (14)

$$C(L) = C^*(L)(I - L) + C(1) \tag{15}$$

MA representation of cointegrated processes II

- The presence of cointegration imposes constrains on $C(1)$.
- Given that
$$\beta' y_t \sim I(0) \tag{16}$$
it must be that
$$\beta' C(1) = 0 \tag{17}$$
as stationary variables should not be influenced by stochastic trends.
- From (17) $C(1)$ is singular and has rank $m - r$.
- The singularity of $C(1)$ implies that the long-run covariance of Δy_t
$$C(1)\Sigma C'(1) \tag{18}$$
is singular and has rank $m - r$.
- From (10) and defining the $m \times (m - r)$ matrices of rank $m - r$ $\beta_\perp$ and $\alpha_\perp$ such that

MA representation of cointegrated processes III

- $\alpha'_\perp \alpha = 0$, $\beta'_\perp \beta = 0$.
- $rank(\alpha, \alpha_\perp) = n$ and $rank(\beta, \beta_\perp) = n$
- $(\alpha'_\perp \Gamma(1) \beta_\perp)^{-1}$ exists where $\Gamma(1) = I_m - \sum_{i=1}^{p-1} \Gamma_i$
- $\beta_\perp (\alpha'_\perp \Gamma(1) \beta_\perp)^{-1} \alpha'_\perp + \alpha' (\alpha'_\perp \Gamma(1) \beta_\perp)^{-1} \beta' = I_m$

- From (10) and defining the $m \times (m-r)$ matrices of rank $m-r$, $\beta_\perp$ and $\alpha_\perp$ such that
  - A necessary and sufficient condition for $\beta' y_t$ to be $I(0)$ is that

$$\alpha'_\perp \Gamma(1) \beta_\perp$$

    has full rank
  - the BN decomposition is given by (14) where

$$C(1) = \beta_\perp (\alpha'_\perp \Gamma(1) \beta_\perp)^{-1} \alpha'_\perp \tag{19}$$

MA representation of cointegrated processes IV

- Notice that

$$\begin{aligned} \beta' C(1) &= \beta' \beta_\perp (\alpha'_\perp \Gamma(1) \beta_\perp)^{-1} \alpha'_\perp = 0 && (20) \\ C(1)\alpha &= \beta_\perp (\alpha'_\perp \Gamma(1) \beta_\perp)^{-1} \alpha'_\perp \alpha = 0 && (21) \end{aligned}$$

- From (10) and defining the $m \times (m - r)$ matrices of rank $m - r$, $\beta_\perp$ and $\alpha_\perp$ such that
  - The common trends in y_t are extracted using

$$\begin{aligned} PC_t &= C(1) \sum_{j=1}^{t} \varepsilon_t \\ &= \beta_\perp (\alpha'_\perp \Gamma(1) \beta_\perp)^{-1} \alpha'_\perp \sum_{j=1}^{t} \varepsilon_t \\ &= \xi \alpha'_\perp \sum_{j=1}^{t} \varepsilon_t && (22) \end{aligned}$$

MA representation of cointegrated processes V

where $\xi = \beta_\perp(\alpha'_\perp \Gamma(1)\beta_\perp)^{-1}$

- The common trends are the linear combinations

$$\alpha'_\perp \sum_{j=1}^{t} \varepsilon_t \tag{23}$$

- We can estimate both PC_t in (22) and the stochastic trends in (23) from the estimated VECM parameters and residuals
- Other authors suggested permanent components that directly depend on y_t
- For example [Gonzalo and Granger, 1995], denoted by GG proposed

$$PC_{GGt} = \alpha'_\perp y_t \tag{24}$$

while [Johansen, 1995], denoted by J, proposed

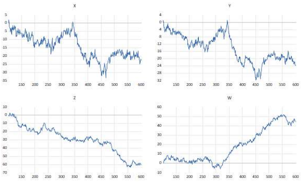
$$PC_{Jt} = \beta'_\perp y_t \tag{25}$$

Table of Contents

Example using simulated series I

- Use the following commands in R
  simulated VEC

Figure: 7.9.1

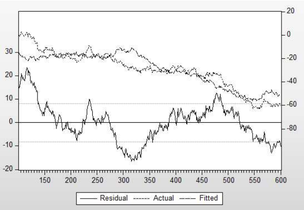


Example using simulated series II

Dependent Variable: Z
Method: Least Squares
Date: 11/05/18 Time: 09:04
Sample: 101 600
Included observations: 500

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-15.19438	0.521170	-29.15435	0.0000
W	-0.781675	0.020377	-38.36033	0.0000

R-squared	0.747146	Mean dependent var	-29.44014
Adjusted R-squared	0.746638	S.D. dependent var	16.24381
S.E. of regression	8.176334	Akaike info criterion	7.044357
Sum squared resid	33292.51	Schwarz criterion	7.061215
Log likelihood	-1759.089	Hannan-Quinn criter.	7.050972
F-statistic	1471.515	Durbin-Watson stat	0.025929
Prob(F-statistic)	0.000000		



Example using simulated series III

Date: 11/05/18 Time: 09:04
Sample: 101 600
Included observations: 500

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.982	0.982	485.03	0.000
		2	0.966	0.035	954.89	0.000
		3	0.950	0.006	1410.2	0.000
		4	0.933	-0.028	1850.5	0.000
		5	0.916	-0.021	2275.6	0.000
		6	0.899	-0.004	2685.9	0.000
		7	0.882	0.012	3082.4	0.000
		8	0.864	-0.069	3463.3	0.000
		9	0.845	-0.027	3828.5	0.000
		10	0.827	0.007	4179.0	0.000
		11	0.808	-0.029	4514.3	0.000
		12	0.792	0.070	4837.0	0.000

Breusch-Godfrey Serial Correlation LM Test:
Null hypothesis: No serial correlation at up to 2 lags

F-statistic	7358.918	Prob. F(2,496)	0.0000
Obs*R-squared	483.6990	Prob. Chi-Square(2)	0.0000

Test Equation:
Dependent Variable: RESID
Method: Least Squares
Date: 11/05/18 Time: 09:04
Sample: 101 600
Included observations: 500
Presample missing value lagged residuals set to zero.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.002426	0.094292	-0.025725	0.9795
W	-0.001015	0.003687	-0.275167	0.7833
RESID(-1)	0.955355	0.044914	21.27092	0.0000
RESID(-2)	0.030316	0.044972	0.674116	0.5006

R-squared	0.967398	Mean dependent var	-1.77E-14
Adjusted R-squared	0.967201	S.D. dependent var	8.168137
S.E. of regression	1.479292	Akaike info criterion	3.628972
Sum squared resid	1085.399	Schwarz criterion	3.662689
Log likelihood	-903.2430	Hannan-Quinn criter.	3.642202
F-statistic	4905.946	Durbin-Watson stat	1.786401
Prob(F-statistic)	0.000000		

Example using simulated series IV

Null Hypothesis: RES_SPURIOUS has a unit root
Exogenous: None
Lag Length: 0 (Automatic - based on SIC, maxlag=17)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-2.067356	0.0373
Test critical values:	1% level	-2.569604	
	5% level	-1.941459	
	10% level	-1.616273	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(RES_SPURIOUS)
Method: Least Squares
Date: 11/05/18 Time: 09:04
Sample (adjusted): 102 600
Included observations: 499 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
RES_SPURIOUS(-1)	-0.014878	0.007196	-2.067356	0.0392

R-squared	0.007008	Mean dependent var	-0.051145
Adjusted R-squared	0.007008	S.D. dependent var	1.315590
S.E. of regression	1.310972	Akaike info criterion	3.381417
Sum squared resid	855.8871	Schwarz criterion	3.389860
Log likelihood	-842.6636	Hannan-Quinn criter.	3.384730
Durbin-Watson stat	2.064196		

Heteroskedasticity Test: ARCH

F-statistic	11288.06	Prob. F(1,497)	0.0000
Obs*R-squared	477.9561	Prob. Chi-Square(1)	0.0000

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Date: 11/05/18 Time: 09:04
Sample (adjusted): 102 600
Included observations: 499 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.340202	1.090778	1.228666	0.2198
RESID^2(-1)	0.976105	0.009187	106.2453	0.0000

R-squared	0.957828	Mean dependent var	66.25579
Adjusted R-squared	0.957743	S.D. dependent var	98.19149
S.E. of regression	20.18472	Akaike info criterion	8.851729
Sum squared resid	202489.2	Schwarz criterion	8.868613
Log likelihood	-2206.506	Hannan-Quinn criter.	8.858355
F-statistic	11288.06	Durbin-Watson stat	1.902846
Prob(F-statistic)	0.000000		

Example using simulated series V

Heteroskedasticity Test: White
Null hypothesis: Homoskedasticity

F-statistic	25.03014	Prob. F(2,497)	0.0000
Obs*R-squared	45.75390	Prob. Chi-Square(2)	0.0000
Scaled explained SS	49.43220	Prob. Chi-Square(2)	0.0000

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Date: 11/05/18 Time: 09:04
Sample: 101 600
Included observations: 500

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	105.9238	6.971600	15.19361	0.0000
W^2	0.080848	0.018589	4.349281	0.0000
W	-5.060419	0.894271	-5.658708	0.0000

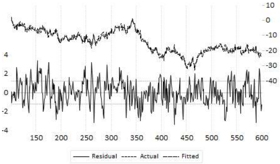
R-squared	0.091508	Mean dependent var	66.58502
Adjusted R-squared	0.087852	S.D. dependent var	98.36891
S.E. of regression	93.94865	Akaike info criterion	11.92936
Sum squared resid	4386695.	Schwarz criterion	11.95464
Log likelihood	-2979.339	Hannan-Quinn criter.	11.93928
F-statistic	25.03014	Durbin-Watson stat	0.048138
Prob(F-statistic)	0.000000		

Dependent Variable: Y
Method: Least Squares
Date: 11/05/18 Time: 09:04
Sample: 101 600
Included observations: 500

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.072796	0.119112	-0.611160	0.5414
X	0.992715	0.007413	133.9165	0.0000

R-squared	0.972981	Mean dependent var	-14.23798
Adjusted R-squared	0.972927	S.D. dependent var	7.442292
S.E. of regression	1.224546	Akaike info criterion	3.247009
Sum squared resid	746.7570	Schwarz criterion	3.263867
Log likelihood	-809.7522	Hannan-Quinn criter.	3.253624
F-statistic	17933.64	Durbin-Watson stat	0.864215
Prob(F-statistic)	0.000000		

Example using simulated series VI



Date: 11/05/18 Time: 09:04
Sample: 101 600
Included observations: 500

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.567	0.567	161.57	0.000
		2	0.301	-0.030	207.22	0.000
		3	0.176	0.026	222.93	0.000
		4	0.124	0.029	230.68	0.000
		5	0.067	-0.025	232.95	0.000
		6	0.048	0.021	234.13	0.000
		7	0.019	-0.023	234.32	0.000
		8	-0.016	-0.032	234.45	0.000
		9	0.005	0.045	234.46	0.000
		10	0.018	0.005	234.63	0.000
		11	0.045	0.042	235.66	0.000
		12	0.000	-0.065	235.66	0.000

Example using simulated series VII

Breusch-Godfrey Serial Correlation LM Test:
Null hypothesis: No serial correlation at up to 2 lags

F-statistic	118.9739	Prob. F(2,496)	0.0000
Obs*R-squared	162.1013	Prob. Chi-Square(2)	0.0000

Test Equation:
Dependent Variable: RESID
Method: Least Squares
Date: 11/05/18 Time: 09:04
Sample: 101 600
Included observations: 500
Presample missing value lagged residuals set to zero.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.100255	0.098728	-1.015469	0.3104
X	-0.006930	0.006152	-1.126318	0.2606
RESID(-1)	0.584431	0.044830	13.03661	0.0000
RESID(-2)	-0.023327	0.045150	-0.516656	0.6056

R-squared	0.324203	Mean dependent var	-2.60E-15
Adjusted R-squared	0.320115	S.D. dependent var	1.223318
S.E. of regression	1.008688	Akaike info criterion	2.863147
Sum squared resid	504.6564	Schwarz criterion	2.896864
Log likelihood	-711.7867	Hannan-Quinn criter.	2.876377
F-statistic	79.31596	Durbin-Watson stat	1.986120
Prob(F-statistic)	0.000000		

Null Hypothesis: RES_REG has a unit root
Exogenous: None
Lag Length: 0 (Automatic - based on SIC, maxlag=17)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-11.69294	0.0000
Test critical values:	1% level	-2.569604	
	5% level	-1.941459	
	10% level	-1.616273	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(RES_REG)
Method: Least Squares
Date: 11/05/18 Time: 09:04
Sample (adjusted): 102 600
Included observations: 499 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
RES_REG(-1)	-0.431956	0.036942	-11.69294	0.0000

R-squared	0.215404	Mean dependent var	-0.002764
Adjusted R-squared	0.215404	S.D. dependent var	1.138373
S.E. of regression	1.008342	Akaike info criterion	2.856494
Sum squared resid	506.3432	Schwarz criterion	2.864936
Log likelihood	-711.6952	Hannan-Quinn criter.	2.859807
Durbin-Watson stat	1.965120		

Example using simulated series VIII

Heteroskedasticity Test: ARCH

F-statistic	40.67933	Prob. F(1,497)	0.0000
Obs*R-squared	37.75296	Prob. Chi-Square(1)	0.0000

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Date: 11/05/18 Time: 09:04
Sample (adjusted): 102 600
Included observations: 499 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.086027	0.108772	9.984474	0.0000
RESID^2(-1)	0.274914	0.043103	6.378035	0.0000

R-squared	0.075657	Mean dependent var	1.496496
Adjusted R-squared	0.073797	S.D. dependent var	2.035377
S.E. of regression	1.958835	Akaike info criterion	4.186577
Sum squared resid	1907.007	Schwarz criterion	4.203461
Log likelihood	-1042.551	Hannan-Quinn criter.	4.193203
F-statistic	40.67933	Durbin-Watson stat	2.000109
Prob(F-statistic)	0.000000		

Heteroskedasticity Test: White
Null hypothesis: Homoskedasticity

F-statistic	3.031216	Prob. F(2,497)	0.0492
Obs*R-squared	6.025527	Prob. Chi-Square(2)	0.0492
Scaled explained SS	5.534536	Prob. Chi-Square(2)	0.0628

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Date: 11/05/18 Time: 09:04
Sample: 101 600
Included observations: 500

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.239362	0.319089	3.884068	0.0001
X^2	-0.003172	0.001580	-2.007289	0.0453
X	-0.075213	0.047784	-1.574012	0.1161

R-squared	0.012051	Mean dependent var	1.493514
Adjusted R-squared	0.008075	S.D. dependent var	2.034430
S.E. of regression	2.026199	Akaike info criterion	4.256182
Sum squared resid	2040.425	Schwarz criterion	4.281470
Log likelihood	-1061.046	Hannan-Quinn criter.	4.266105
F-statistic	3.031216	Durbin-Watson stat	1.466664
Prob(F-statistic)	0.049150		

Example using simulated series IX

Figure: Engle-Granger Cointegration

Date: 11/05/18 Time: 09:04
Series: Y X
Sample: 101 600
Included observations: 500
Null hypothesis: Series are not cointegrated
Cointegrating equation deterministics: C
Automatic lags specification based on Schwarz criterion (maxlag=17)

Dependent	tau-statistic	Prob.*	z-statistic	Prob.*
Y	-11.69294	0.0000	-215.5458	0.0000
X	-11.88740	0.0000	-220.9395	0.0000

*MacKinnon (1996) p-values.

Intermediate Results:

	Y	X
Rho - 1	-0.431956	-0.442765
Rho S.E.	0.036942	0.037247
Residual variance	1.016753	1.021282
Long-run residual variance	1.016753	1.021282
Number of lags	0	0
Number of observations	499	499
Number of stochastic trends**	2	2

**Number of stochastic trends in asymptotic distribution

Example using simulated series X

Figure: Johansen´s Test

Unrestricted Cointegration Rank Test (Trace)

Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	0.05 Critical Value	Prob.**
None *	0.199868	111.0575	12.32090	0.0000
At most 1	2.79E-05	0.013918	4.129906	0.9233

Trace test indicates 1 cointegrating eqn(s) at the 0.05 level
* denotes rejection of the hypothesis at the 0.05 level
**MacKinnon-Haug-Michelis (1999) p-values

Unrestricted Cointegration Rank Test (Maximum Eigenvalue)

Hypothesized No. of CE(s)	Eigenvalue	Max-Eigen Statistic	0.05 Critical Value	Prob.**
None *	0.199868	111.0436	11.22480	0.0000
At most 1	2.79E-05	0.013918	4.129906	0.9233

Max-eigenvalue test indicates 1 cointegrating eqn(s) at the 0.05 level
* denotes rejection of the hypothesis at the 0.05 level
**MacKinnon-Haug-Michelis (1999) p-values

Unrestricted Cointegrating Coefficients (normalized by b'*S11*b=I):

Y	X
-0.928721	0.929446
0.047329	0.015097

Unrestricted Adjustment Coefficients (alpha):

D(Y)	-0.060282	0.005440
D(X)	-0.551664	0.004232

1 Cointegrating Equation(s): Log likelihood -1432.409

Normalized cointegrating coefficients (standard error in parentheses)

Y	X
1.000000	-1.000781
	(0.00604)

Adjustment coefficients (standard error in parentheses)

D(Y)	0.055986
	(0.04325)
D(X)	0.512341
	(0.05692)

Example using simulated series XI

Figure: VEC estimation

Vector Error Correction Estimates
Date: 11/05/18 Time: 10:00
Sample (adjusted): 103 600
Included observations: 498 after adjustments
Standard errors in ( ) & t-statistics in [ ]

Cointegrating Eq:	CointEq1
Y(-1)	1.000000
X(-1)	-1.000781
	(0.00604)
	[-165.677]

Error Correction:	D(Y)	D(X)
CointEq1	0.055986	0.512341
	(0.04325)	(0.05692)
	[1.29458]	[9.00149]
D(Y(-1))	-0.087143	-0.183899
	(0.06393)	(0.08415)
	[-1.36301]	[-2.18547]
D(X(-1))	0.007405	0.037536
	(0.04605)	(0.06061)
	[0.16079]	[0.61928]
R-squared	0.006919	0.166186
Adj. R-squared	0.002906	0.162817
Sum sq. resids	534.5117	925.8795
S.E. equation	1.039145	1.367649
F-statistic	1.724298	49.32876
Log likelihood	-724.2490	-861.0473
Akaike AIC	2.920679	3.470069
Schwarz SC	2.946044	3.495434
Mean dependent	-0.049468	-0.049426
S.D. dependent	1.040658	1.494736

Determinant resid covariance (dof adj.)	1.093056
Determinant resid covariance	1.079927
Log likelihood	-1432.409
Akaike information criterion	5.784776
Schwarz criterion	5.852416
Number of coefficients	8

Example using simulated series XII

Figure: B-N Decomposition

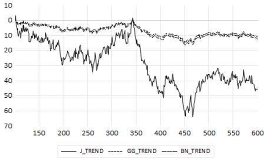


Table of Contents

Forecasting in the presence of cointegration

- We use the VECM representation (10) where r and the unkown parameters are replaced byh their ML estimates, i.e.

$$\Delta\widehat{y}_{T+h} = -\widehat{\alpha}\widehat{\beta}'\widehat{y}_{T+h-1} + \widehat{\Gamma}_1\Delta\widehat{y}_{T+h-1} + \cdots + \widehat{\Gamma}_{p-1}\Delta\widehat{y}_{T+h-p+1} \quad (26)$$

- for $h = 1, \cdots, H$ where as usual forecast values on the right hand side are replaced by actual realizations when available.
- Forecast for the level can be obtained as

$$\widehat{y}_{T+h} = \widehat{y}_T + \Delta\widehat{y}_{T+1} + \cdots + \Delta\widehat{y}_{T+h} \quad (27)$$

- [Clements and Hendry, 1998] present detailed Monte Carlo experiment to rank the forecasts from ECMs and VAR in differences and levels, finding that in general the ECM forecasts should be preferred, as long as there is not substancial model mis-specification.

Table of Contents

The effect of stochastic trends on forecast I

- Let us start from the $MA(\infty)$ representation (13).
- Assuming $\varepsilon_j = 0$, $j \leq 0$ and $y_0 = 0$ we can rewrite (13) as

$$y_t = \sum_{i=1}^{t} \sum_{j=0}^{t-i} C_j \varepsilon_i \tag{28}$$

- and

$$\begin{aligned} y_{T+h} &= \sum_{i=1}^{T+h} \sum_{j=0}^{T+h-i} C_j \varepsilon_i \\ &= \sum_{i=1}^{T} \sum_{j=0}^{T+h-i} C_j \varepsilon_i + \sum_{i=T+1}^{T+h} \sum_{j=0}^{T+h-i} C_j \varepsilon_i \end{aligned} \tag{29}$$

The effect of stochastic trends on forecast II

- Thus

$$\widehat{y}_{T+h} = E(y_{T+h}|y_T) = \sum_{i=1}^{T} \sum_{j=0}^{T+h-i} C_j \varepsilon_i \tag{30}$$

- Given that

$$\lim_{h\to\infty} \sum_{j=0}^{T+h-i} C_j = C(1) \tag{31}$$

- and if the C_j decay repidly, we can write

$$\sum_{j=0}^{T+h-i} C_j \approx C(1) \tag{32}$$

- and

$$\widehat{y}_{T+h} \approx C(1) \sum_{i=1}^{T} \varepsilon_i \tag{33}$$

The effect of stochastic trends on forecast III

- so that, at least for long horizons h the forecasts are driven by the stochastic trends.
- Note that

$$\beta' \widehat{y}_{T+h} \approx \beta' C(1) \sum_{i=1}^{T} \varepsilon_i = 0 \tag{34}$$

- so that for long horizon forecasts are tied together by the presence of cointegration among the variables.
- From (29) and (30) we have

$$e_{T+h} = y_{T+h} - \widehat{y}_{T+h} = \sum_{i=T+1}^{T+h} \sum_{j=0}^{T+h-i} C_j \varepsilon_i = \sum_{i=1}^{h} \sum_{j=0}^{h-i} C_j \varepsilon_{T+i} \tag{35}$$

The effect of stochastic trends on forecast IV

- and

$$Var(e_{T+h}) = \sum_{i=1}^{h} \left[\left(\sum_{j=0}^{h-i} C_j \right) \Sigma \left(\sum_{j=0}^{h-i} C_j' \right) \right] \tag{36}$$

- The variance of the forecasts for the levels grows with the forecast horizon, while $\lim_{h \to \infty} \frac{Var(e_{T+h})}{h}$ converges to a well defined matrix.
- Finally $\lim_{h \to \infty} Var(\beta' e_{T+h})$ converges to a proper matrix.

Table of Contents

Example using simulated series

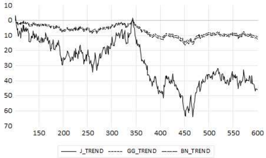


Table of Contents

Example using simulated series

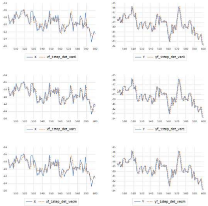


Table of Contents

Example using simulated series

VAR0	y			x		
	1 step ahead, deterministic	1 step ahead, stochastic			1 step ah...	1 step ahead, stochastic
RMSFE	0.926000	0.920000		RMSFE	1.302000	1.302000
MAE	0.758000	0.754000		MAE	1.004000	1.006000
VAR1	y			x		
	1 step ahead, deterministic	1 step ahead, stochastic			1 step ah...	1 step ahead, stochastic
RMSFE	0.920000	0.919000		RMSFE	1.374000	1.376000
MAE	0.750000	0.751000		MAE	1.095000	1.095000
VECM	y			x		
	1 step ahead, deterministic	1 step ahead, stochastic			1 step ah...	1 step ahead, stochastic
RMSFE	0.935000	0.935000		RMSFE	1.305000	1.302000
MAE	0.767000	0.767000		MAE	1.002000	0.998000

Table of Contents

Example using simulated series

- Aula7_2020

Table of Contents

Reference I

Beveridge, S. and Nelson, Charles, R. (1981).
A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of business cycle.
Journal of Monetary Economics, 7:151–174.

Clements, M. P. and Hendry, D. F. (1998).
Forecasting Economic Time Series.
Cambridge University Press.

Davidson, J. E. H., Hendry, D. F., Srba, F., and Yeo, S. (1978).
Econometric modelling of the aggregate time-series relationship between consumers´ expenditure and income in the united kingdom.
Economic Journal, 88:661–692.

Reference II

Engle, R. F. and Granger, C. W. J. (1987).
Co-integration and error correction: Representation, estimation, and testing.
Econometrica, 55(2):251–276.

Gonzalo, J. and Granger, C. W. J. (1995).
Estimation of common long-memory components in cointegrated systems.
Journal of Business & Economic Statistics, 13(1):27–35.

Granger, C. W. J. (1983).
Co-integrated variables and error-correcting models.
Technical Report 93-13, UCSD.

Granger, C. W. J. and Newbold, P. (1974).
Spurious regressions in econometrics.
Journal of Econometrics, 2(2):111–120.

Reference III

Granger, C. W. J. and Weiss, A. A. (1983).
Time series analysis of error-correction models.
In *Studies in Econometrics, Time Series, and Multivariate Statistics*, pages 255–278. Elsevier.

Johansen, S. (1995).
Likelihood-based Inference in Cointegrated Vector Autoregressive Models.
Oxford University Press.

MacKinnon, J. G. (1991).
Critical values for cointegration tests.
In Engle, R. F. and Granger, C. W. J., editors, *Long- Run Economic Relationships*, pages 267–276. Oxford University Press.

Reference IV

MacKinnon, J. G. (1996).
Numerical distribution functions for unit root and cointegration tests.

Journal of Applied Econometrics, 11(6):601–618.

Phillips, P. C. B. (1986).
Understanding spurious regressions in econometrics.
Journal of Econometrics, 33:311–340.