Probabilistic Inference : Changepoints and Cointegration

Chris Bracegirdle

c.bracegirdle@cs.ucl.ac.uk

26th April 2013

Outline

1 Changepoints
   - Modelling Approaches
   - Reset Models : Probabilistic Inferece

2 Cointegration
   - Introduction to Cointegration
   - Cointegration the Bayesian way

3 Intermittent Cointegration

Changepoints

- Boundaries of a partition of time-series into disjoint, contiguous segments

- The process generating each segment is independent of other segments.

- Conditioned on a partition $\mathcal{P}$, the data in a segment $\{a : b\} \subset \{1 : T\}$ are independent of other data,

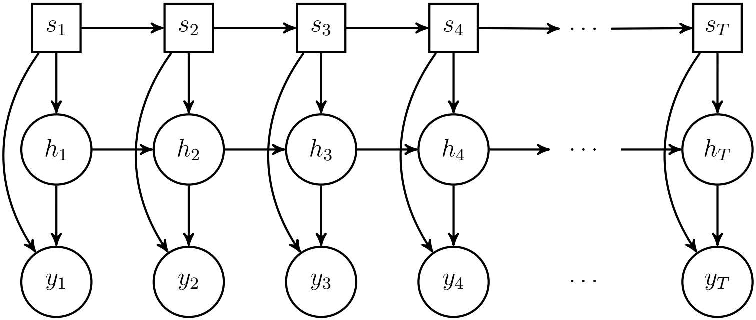
$$p(y_{a:b}|\mathcal{P}) \equiv p(y_{a:b}|y_{1:a-1}, y_{b+1:T}, \mathcal{P})$$

Changepoint Detection

- Binary Segmentation
  - divide-and-conquer, greedy strategy for identifying change-points
  - cost function calculated for all locations, compared with the cost of no changepoint
  - segments are then sub-divided
  - greedy algorithm — not necessarily optimal
- Segment Neighbourhoods
  - exact algorithm based on dynamic programming
- various optimisations

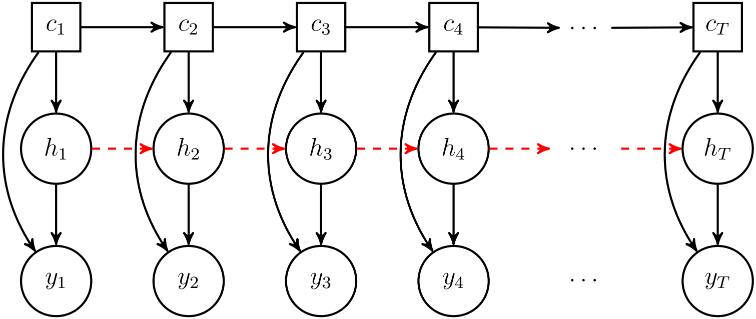
Reset Models : Probabilistic Inference

- Inspired by Switching Latent Markov Models
  - For S possible latent possible states inference scales $O\left(S^T\right)$
  - Intractable for all but very short samples



Reset Models : Probabilistic Inference

- Reset model (Bracegirdle & Barber, 2011)
  - Change in state index $\Rightarrow$ break in latent dynamics
  - A form of changepoint model
  - Filtering scales $O\left(T^2\right)$, smoothing $O\left(T^3\right)$.



Outline

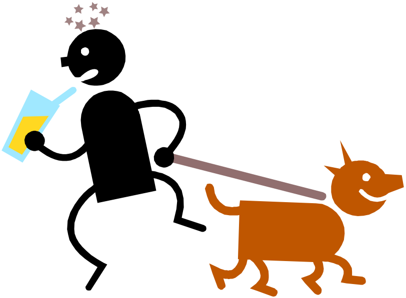
1 Changepoints
Modelling Approaches
Reset Models : Probabilistic Inferece

2 Cointegration
Introduction to Cointegration
Cointegration the Bayesian way

3 Intermittent Cointegration

What is "Cointegration"?

- Two series $x_{1:T}$ and $y_{1:T}$, both unpredictable random walks
- Relationship may be predictable: **drunk walking a dog**



- Significant interest: statistical arbitrage, pairs trading

A Concrete Example

- Suppose

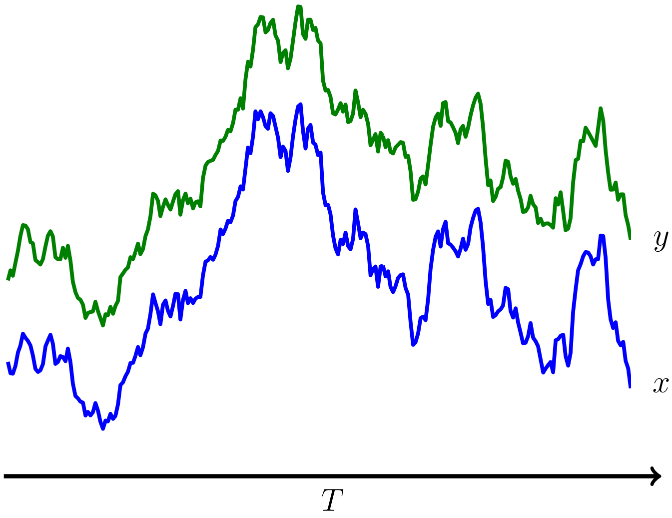
$$x_{t+1} = x_t + \epsilon_{t+1}, \quad y_{t+1} = y_t + \epsilon_{t+1}, \quad \epsilon_t \sim \mathcal{N}(0, 1)$$

- x and y are unpredictable random walks
- **BUT** the difference

$$x_{t+1} - y_{t+1} = x_t - y_t$$

is **perfectly** predictable

What does it look like?



Greek and German Bond Yields

Motivation for Cointegration

- We seek to write y_t in linear terms in x_t.

- This is simple regression $\Longrightarrow$ Use least-squares estimation.

$$\boldsymbol{\beta} = \left(\mathbf{X}^\top \mathbf{X}\right)^{-1} \mathbf{X}^\top \mathbf{y}.$$

- OLS is B.L.U.E.?

Spurious Regression

- If x_t and y_t are correlated, may be due to a third variable z_t.
- Control for the effect of z_t on x_t and y_t, the partial effect between x_t and y_t is removed; the relationship is then explained.
- Initial relation between x_t and y_t was a spurious relation.

Spurious Regression

- Additional complication with non-stationary series x_t, y_t.

- Granger & Newbold (1974): x_t and y_t assumed to be independent random walks defined by

$$x_t = x_{t-1} + \eta_t^x, \quad \eta_t^x \sim \mathcal{N}\left(0, \sigma_x^2\right)$$
$$y_t = y_{t-1} + \eta_t^y, \quad \eta_t^y \sim \mathcal{N}\left(0, \sigma_y^2\right)$$

  estimate the regression

$$y_t = \beta_0 + \beta_1 x_t + \epsilon_t$$

- $\Longrightarrow$ large proportion of simulations, a significant relationship detected

Spurious Regression

- x_t, y_t independent $\Longrightarrow$ true value of β_1 = 0

- However, residuals $\epsilon_t = y_t - \beta_0$ form a random walk with mean $\langle \epsilon_t \rangle = \langle y_1 \rangle - \beta_0$
  Consistency. For stationary data $x_{i,t}$ need *contemporaneous exogeneity*,

$$\left\langle \epsilon_t \right\rangle_{p(\epsilon_t | x_{1:p,t})} = 0.$$

  Bias. Need *strict exogeneity*,

$$\left\langle \epsilon_t \right\rangle_{p\left(\epsilon_t \middle| x_{1:p,1:T}\right)} = 0.$$

  $\Longrightarrow$ OLS is biased

Classical Method: Engle-Granger

1. Use ordinary least squares to estimate $y_t = \alpha + \beta x_t$
2. Use Dickey-Fuller (DF) test to check for unit root in the residual $\epsilon_t = y_t - \alpha - \beta x_t$

Dickey-Fuller (DF) Test

- For a simple autoregressive model

  $$\epsilon_t = \phi\epsilon_{t-1} + \eta_t, \quad \eta_t \sim \mathcal{N}\left(0, \sigma^2\right)$$

  DF-test is for the presence of a unit root ($\phi = 1$)

- Normally written as

  $$\Delta\epsilon_t = (\phi - 1)\epsilon_{t-1} + \eta_t = \delta\epsilon_{t-1} + \eta_t$$

  and test $\mathcal{H}_0 : \delta = 0$

- Test statistic:
  if t^* >critical value $\Rightarrow$ do not reject $\mathcal{H}_0$, i.e. unit root exists

The Problem with Engle-Granger

- OLS is **consistent**, but for small samples has **bias** (Watson & Teelucksingh, 2002)

- OLS effectively assumes $\phi = 0$ in the autoregression

$$\epsilon_t = \phi \epsilon_{t-1} + \eta_t, \quad \eta_t \sim \mathcal{N}\left(0, \sigma^2\right), \quad |\phi| < 1$$

- **Spurious regression**: under $\mathcal{H}_0$ of the DF test, $\epsilon_{1:T}$ is a random walk ($\phi = 1$) so the OLS estimation was spurious (Granger & Newbold, 1974)

$\Longrightarrow$ Difficult to reconcile OLS and DF assumptions

Plan of Attack

- Build a more consistent model, incorporate uncertainty in ϕ
- Can we find a better estimation of α, β?
- Can we improve the test for cointegration?

Cointegration for Our Purposes

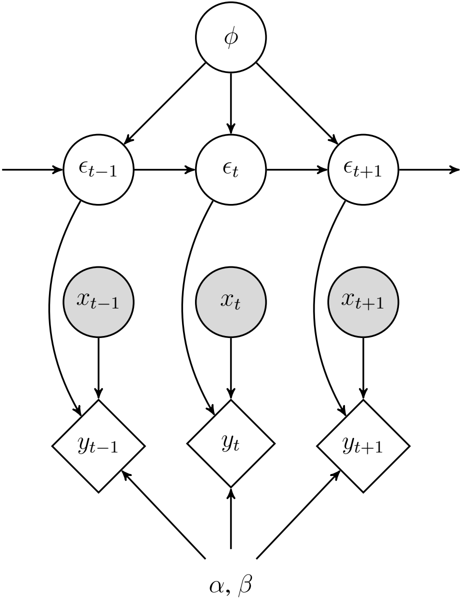
- x and y are cointegrated if $\exists$ stationary linear combination
- We write

$$y_t = \alpha + \beta x_t + \epsilon_t$$

where $\epsilon_{1:T}$ forms a stationary process,

$$\epsilon_t = \phi \epsilon_{t-1} + \eta_t, \quad \eta_t \sim \mathcal{N}\left(0, \sigma^2\right), \quad |\phi| < 1$$

A Generative Model



- Pick α, β by maximum likelihood
- OLS assumes $\phi = 0$
- We **integrate** ϕ

Estimating α, β

- Set parameters α,β with maximum likelihood

$$p(y_{1:T}|x_{1:T}, \alpha, \beta) = \int_\phi p(\phi)\, p(\epsilon_{1:T}|\phi)$$

- Latent variable; use Expectation Maximisation (EM)
- Require moments $\langle\phi\rangle$, $\langle\phi^2\rangle$ from posterior

Inference of ϕ

- Define

$$p(y_t|x_t, \epsilon_t) = \delta(y_t - \alpha - \beta x_t - \epsilon_t)$$
$$p(\epsilon_t|\epsilon_{t-1}, \phi) = \mathcal{N}\left(\epsilon_t \middle| \phi\epsilon_{t-1}, \sigma^2\right)$$

- Cointegration requires $|\phi| < 1$, so prior

$$p(\phi) = \mathcal{U}(\phi|\,(-1,1)) = \tfrac{1}{2}\,[\phi \in (-1,1)]$$

- $\epsilon_{1:T}$ stationary when each $\langle \epsilon_t \rangle = 0$ and $\langle \epsilon_t^2 \rangle$ is constant.
  - Variance satisfies $\left\langle \epsilon_{t+1}^2 \right\rangle = \phi^2 \left\langle \epsilon_t^2 \right\rangle + \sigma^2$
  - so choose $p(\epsilon_1|\phi) = \mathcal{N}\left(\epsilon_1 \middle| 0, \frac{\sigma^2}{(1-\phi^2)}\right)$

Inference of ϕ

- Posterior is given by a truncated Gaussian with prefactor,

$$p(\phi|\epsilon_{1:T}) \propto p(\phi)\,\sqrt{1-\phi^2}\mathcal{N}\left(\phi\middle|\frac{\hat{e}_{12}}{\hat{e}_1},\frac{\sigma^2}{\hat{e}_1}\right)$$

  where

$$\hat{e}_{12} \equiv \textstyle\sum_{t=2}^{T}\epsilon_t\epsilon_{t-1}, \qquad \hat{e}_1 \equiv \textstyle\sum_{t=3}^{T}\epsilon_{t-1}^2$$

Estimation Algorithm

$\{\alpha, \beta, \sigma^2\} \leftarrow \textsc{LinearRegression}\left(x_{1:T}, y_{1:T}\right)$

repeat

  $\epsilon_{1:T} \leftarrow y_{1:T} - \alpha - \beta x_{1:T}$

  $\{l_{\mathrm{C}}, \langle\phi\rangle, \langle\phi^2\rangle\} \leftarrow \textsc{CointInference}\left(\epsilon_{1:T}, \sigma^2\right)$

  $\{\alpha, \beta, \sigma^2\} \leftarrow \mathrm{EM}\left(x_{1:T}, y_{1:T}, \langle\phi\rangle, \langle\phi^2\rangle\right)$

until convergence

A Bayesian Cointegration Test

- Seek confirmation that $\epsilon_{1:T}$ is stationary
- Compare likelihoods of the cointegration model with new random walk (RW) model
- RW model is easy—just take $\phi = 1$!

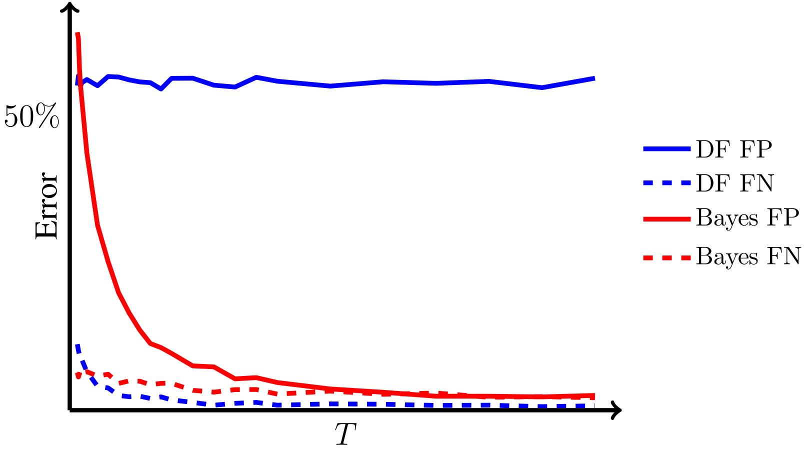
A Bayesian Cointegration Test

- Compare likelihoods of the cointegration model with random walk (RW) model

$$\frac{\text{Random Walk}}{\text{Cointegration}} : \frac{p(y_{1:T}|x_{1:T};\phi = 1)}{p(y_{1:T}|x_{1:T};|\phi| < 1)}$$
$$= \frac{p(\epsilon_{1:T}|\phi = 1)}{p(\epsilon_{1:T}|\,|\phi| < 1)} \equiv \frac{l_{\text{RW}}}{l_{\text{C}}}$$

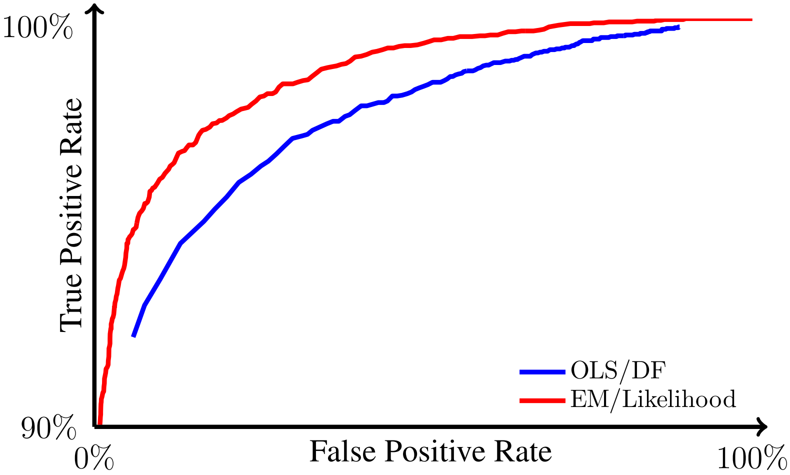
- If $\frac{l_{\text{RW}}}{l_{\text{C}}} <$ threshold, series $x_{1:T}$, $y_{1:T}$ are cointegrated

Detecting Cointegration – Does it Work?



5,000 generated series pairs for each T, 50% cointegrated.
FP = false positive rate. FN = false negative rate.
DF = OLS/DF test (5% significance). Bayes = EM/Bayes factor (threshold $\exp 2$)

Receiver Operating Characteristic Curve



10,000 generated series pairs, 50% cointegrated, $T = 100$.
EM = estimation in Bayesian cointegration model by ML
Likelihood = Bayes factor comparison

Outline

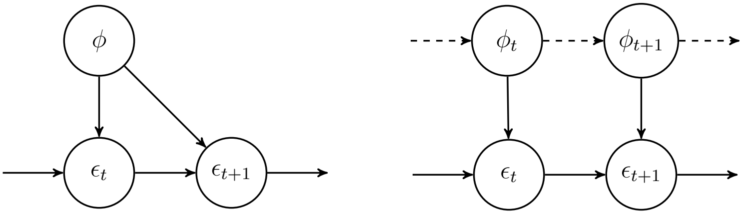
1 Changepoints
 Modelling Approaches
 Reset Models : Probabilistic Inferece

2 Cointegration
 Introduction to Cointegration
 Cointegration the Bayesian way

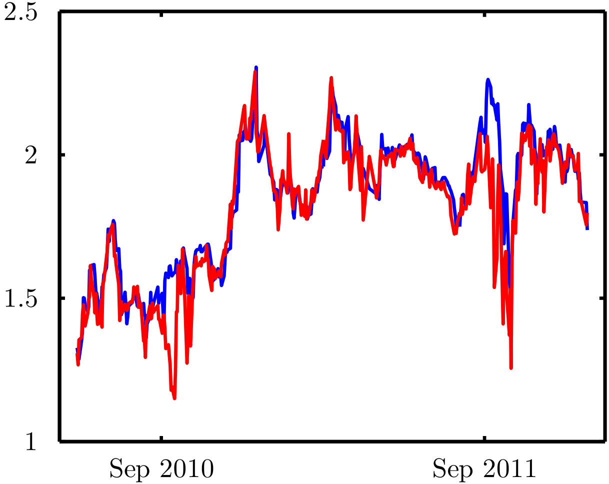
3 Intermittent Cointegration

Intermittent Cointegration

- What if cointegration can be ‘switched off’?
- Example: Interconnector gas pipeline
- Place the model in reset framework, allow changepoints (Bracegirdle & Barber, 2012)

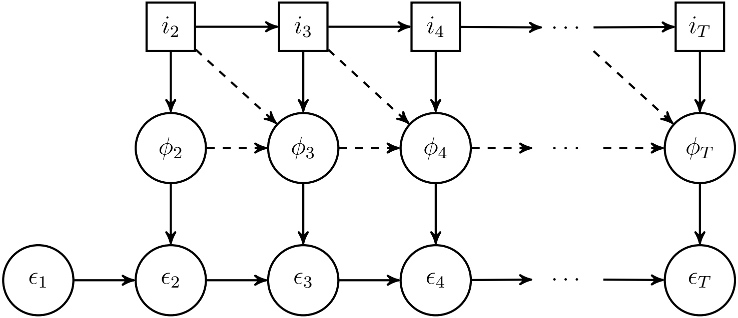


Interconnector – Prices



UK prices from `www.nationalgrid.com`, Zeebrugge `www.apxendex.com`.
Blue=x_t=Zeebrugge, Red=y_t=UK

Intermittent Cointegration Model



- Time-varying ϕ_t set up as piecewise-constant
- Use binary switch i_t, 1 represents random walk, 0 represents cointegration

Intermittent Cointegration Model

- Initial distribution for ϕ_2 is

$$p(\phi_2|i_2) = \begin{cases} p^1(\phi_{\phi_2}) = \delta(\phi_2 - 1) & i_2 = 1 \\ p^0(\phi_{\phi_2}) = \mathcal{U}(\phi_2|\,(-1,1)) & i_2 = 0 \end{cases}$$

- Piecewise-constant transition

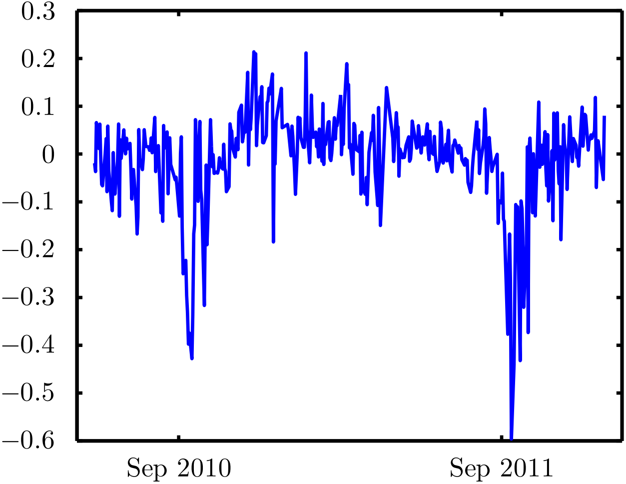
$$p(\phi_t|\phi_{t-1}, i_t, i_{t-1}) = \begin{cases} p^1(\phi_{\phi_t}) & i_t = 1 \\ \delta(\phi_t - \phi_{t-1}) & i_t = 0, i_{t-1} = 0 \\ p^0(\phi_{\phi_t}) & i_t = 0, i_{t-1} = 1. \end{cases}$$

- State switch user specified $p(i_1)$, $p(i_t|i_{t-1})$

Intermittent Cointegration Model

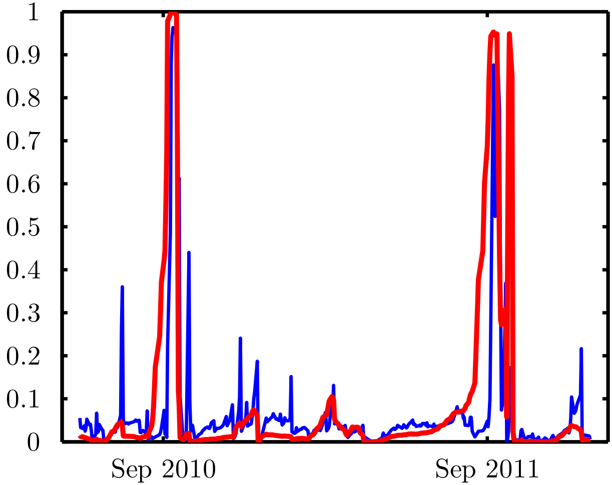
- Inference for $p(\phi_t, i_t|\epsilon_{1:T})$ obtained by filtering and correction smoothing
- Taking $\phi_{2:T}$ and $i_{2:T}$ as latent variables, apply EM to estimate parameters α, β, σ^2

Interconnector – Residuals



Maintenance occurs each September.

Interconnector – RW Posterior



Maintenance occurs each September.

Summary

- Exact and approximate inference algorithms for changepoints

- Two novel techniques for cointegration
  - Estimate α, β in a Bayesian framework for ϕ
  - Check for stationarity by comparing likelihood with RW

- Extensible cointegration model allows for intermittent cointegration

References

Bracegirdle, C. and Barber, D. Switch-Reset Models : Exact and Approximate Inference. In *AISTATS*, volume 15. JMLR, 2011.

Bracegirdle, C. and Barber, D. Bayesian Conditional Cointegration. In *29th International Conference on Machine Learning (ICML)*, 2012.

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

Watson, P. K. and Teelucksingh, S. S. *A practical introduction to econometric methods: Classical and modern*. University of West Indies Press, 2002.