Modelos de Regressão Dinâmica + ECM

Pedro Costa Ferreira

FGV|IDE

Big Data e Data Science

November 14, 2016

Introdução

- Modelos dinâmicos estudam a relação entre variáveis observadas em instantes de tempo diferentes;
- Em diferentes contextos estamos interessados em estudar se o comportamento de uma determinada variável (dependente) é influenciado por uma ou mais variáveis (explicativas);
- Podemos, por exemplo, investigar se o Índice de Preços ao Consumidor Amplo (IPCA) em um determinado mês influencia a maneira como os consumidores brasileiros formam suas expectativas de inflação em meses subsequentes.

Modelo clássico de regressão linear - uma pequena revisão

- Os modelos que assumem estrutura linear entre variável dependente e variáveis explicativas são chamados modelos de regressão linear ou modelos lineares;

- Considere a variável dependente Y_t, observada ao longo do tempo, e k variáveis explicativas $\{X_{1,t}, X_{2,t}, \dots X_{k,t}\}$. O modelo linear usualmente encontrado na literatura pode ser escrito como

$$\boxed{\mathrm{Y}_t = \beta_0 + \beta_1 \mathrm{X}_{1,t} + \beta_2 \mathrm{X}_{2,t} + \cdots + \beta_k \mathrm{X}_{k,t} + \epsilon_t}$$

Onde:
β_0 é um nível global;
β_j's, $j \in \{1, 2, \dots, k\}$ são os parâmetros correspondentes aos respectivos efeitos isolados de cada $X_{j,t}$ sobre Y_t;
ϵ_t é o erro do modelo no tempo t.

- A construção dos modelos de regressão linear é fundamentada na aceitação dos seguintes pressupostos sobre o erro ϵ_t:

  i. Exogeneidade estrita: as variáveis explicativas X_k são estritamente exógenas com respeito ao termo de erro ϵ_t de maneira que

  $$E(\epsilon_t \mid \boldsymbol{X}) = 0, \quad t = 1, 2, \dots, T$$

  onde $\boldsymbol{X}$ inclui todas as k variáveis explicativas e todos os instantes de tempo t:

  $$\boldsymbol{X} = \begin{bmatrix} X_{1,1} & X_{2,1} & \cdots & X_{k-1,1} & X_{k,1} \\ X_{1,2} & X_{2,2} & \cdots & X_{k-1,2} & X_{k,2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ X_{1,T} & X_{2,T} & \cdots & X_{k-1,T} & X_{k,T} \end{bmatrix}.$$

Modelo clássico de regressão linear - uma pequena revisão

ii. Ausência de colinearidade perfeita: nenhuma variável explicativa $X_{k,t}$ é constante ou pode ser expressa como uma função linear de outras covariáveis.

iii. Homocedasticidade: a variância do erro é a mesma para todas as observações, ou seja, $Var(\epsilon_t \mid \boldsymbol{X}) = \sigma^2$, $t = 1, 2, \ldots, T$.

iv. Ausência de correlação serial: os termos de erro são independentes (condicionalmente a $\boldsymbol{X}$), ou seja, $Cov(\epsilon_t, \epsilon_{t-s} \mid \boldsymbol{X}) = 0$, $s = 1, 2, \ldots, T-1$.

v. Normalidade: os ϵ_t's são identicamente distribuídos com $\epsilon_t \sim N(0, \sigma^2)$.

- Os pressupostos expostos anteriormente são razoáveis em contextos onde as observações são independentes. Na análise de séries temporais, entretanto, algumas suposições frequentemente não são satisfeitas.

- Em particular, a suposição de que as variáveis explicativas são independentes de toda a história de Y (exogeneidade estrita) e que choques em Y no período t não persistem em $t+1$ (correlação serial) são usualmente violadas.

Modelo clássico de regressão linear - uma pequena revisão

- O impacto sobre Y de um choque em determinada variável explicativa X pode não ocorrer imediatamente, sendo Y afetado somente após alguns instantes de tempo, ou mesmo Y pode afetar seu próprio valor em tempos posteriores.

- A omissão dessa dinâmica pode induzir correlação serial nos erros, sendo interessante nessas circunstâncias optar pelo uso de modelos dinâmicos.

Modelos autorregressivos com defasagem distribuída (ADL)

Modelos autorregressivos com defasagem distribuída (ADL)

A forma geral do modelo ADL com p defasagens para Y e q defasagens para X, denotado por ADL(p,q), é dada por

$$\boxed{\phi(L)Y_t = \alpha + \theta(L)X_t + \nu_t} \tag{1}$$

Onde:
$\phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p$, cujas raízes não pertencem ao círculo unitário, o que significa dizer que Y é estacionário;
$\theta(L) = \theta_0 + \theta_1 L + \theta_2 L^2 + \cdots + \theta_q L^q$;
L é um operador de defasagem tal que $L^k Y_t = Y_{t-k}$.

Podemos reescrever (1) como:

$$\boxed{\mathrm{Y}_t = \alpha + \phi_1 Y_{t-1} + \cdots + \phi_p Y_{t-p} + \theta_0 X_t + \theta_1 X_{t-1} + \cdots + \theta_q X_{t-q} + \nu_t} \tag{2}$$

Supondo que os erros $\boldsymbol{\nu} = \{\nu_1, \ldots, \nu_T\}$ são ruídos brancos, o modelo (2) pode ser estimado via mínimos quadrados ordinários.

Modelos autorregressivos com defasagem distribuída (ADL)

⇝ Escolhendo as defasagens p e q

Diferentes métodos estão disponíveis para avaliar o número apropriado de defasagens no modelo, não existindo um "método correto". A escolha é, portanto, usualmente feita pela combinação de métodos.

a) Testes de significância dos parâmetros

   Podemos começar com um número elevado de defasagens e avaliar a significância do coeficiente associado à maior defasagem.

b) Critérios de informação

   Escolhemos o número de defasagens com base no modelo que retorna o menor valor desses critérios.

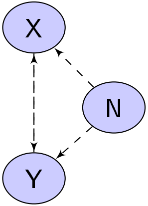
O problema da não estacionariedade
Regressões Espúrias

O problema da não estacionariedade - Regressões Espúrias

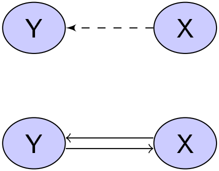
Mera Coincidência

Consumo de Frango per capita $\Longleftrightarrow$ Importação de petróleo United States

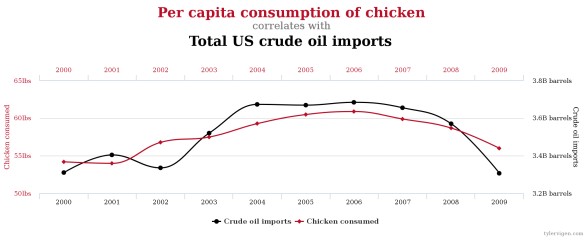
Fator comum

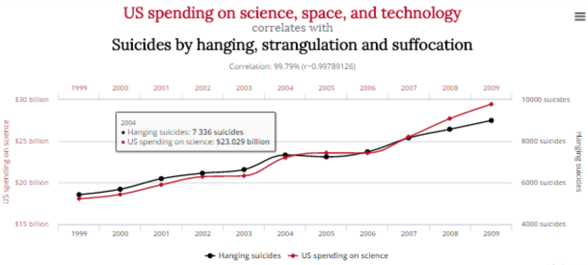


Causalidade

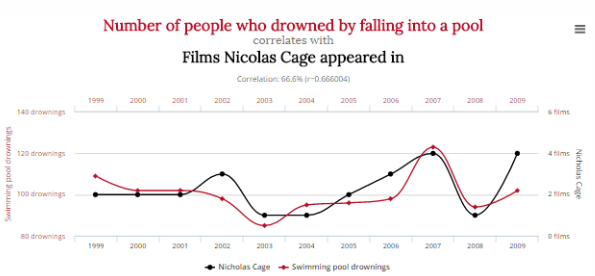


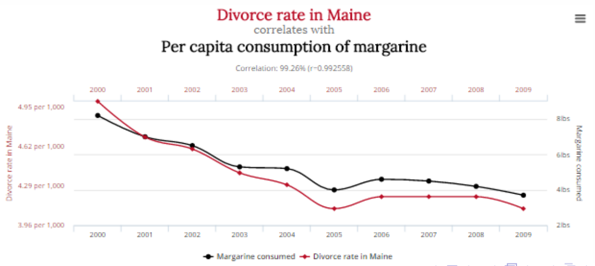
O problema da não estacionariedade - Regressões Espúrias





O problema da não estacionariedade - Regressões Espúrias





O problema da não estacionariedade - Regressões Espúrias

Quero estimar: $Y_t = a + bX_t + \varepsilon_t$

Assumindo que: $\begin{bmatrix} Y_t = \phi Y_{t-1} + u_t \\ X_t = \alpha X_{t-1} + w_t \end{bmatrix} \Longrightarrow$ Independentes

Experimento de Granger e Newbold (1974)

Se $\phi = \alpha = 1$
X_t e Y_t **NÃO** estacionárias

$\Longrightarrow$

- R^2 altos e DW baixos
- Alta chance de rejeitar $H_0 : b = 0$
- Razão t não segue t de Student
- Estatística F não segue distribuição F

Como proceder?

- Remover a tendência (Detrending)?
  - Pode não resolver! Possibilidade de tendência estocástica
- Diferenciar até estacionariedade?
  - Perda de informação de longo prazo (t. econômica)
- Possível solução
  - Teoria da Cointegração

Teoria de Cointegração

Verificar Estacionariedade (Testes de Raízes Unitárias)

- Séries Estacionárias, usar econometria clássica;
- Séries Não Estacionárias, verificar Cointegração
  - Séries Cointegradas Modelo de Correção de Erros
  - Séries Não Cointegradas Modelo sem correção de Erros

Um pouco de História ...

- Box & Jenkins (1970) Processos estocásticos não estacionários/integrados (Modelos ARIMA)
- Granger e Newbold (1974) Econometria clássica não vale se variáveis do modelo são séries temporais não estacionárias (Regressões Espúrias)
- Em particular, se as séries Não Estacionárias forem independentes obtém-se **regressões espúrias**;
- Diferenciar séries até estacionariedade não resolve, perde-se informações de **longo-prazo**;
- O que fazer?...

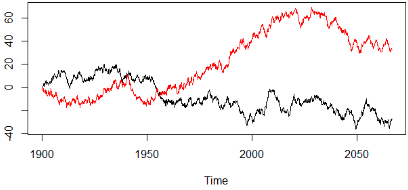
Um pouco de História ...

Teoria da Cointegração

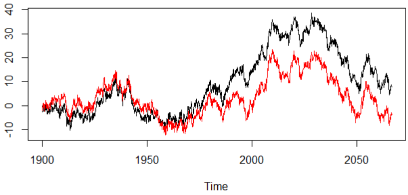
- **Granger (1983)** introduz o conceito de cointegração na literatura;
- **Granger e Engle (1987)** estabelecem relação entre cointegração e os modelos de correção de erros;
- **Década de 1990** proliferam trabalhos teóricos e empíricos;
- **2003** Granger e Engle ganham o Prêmio Nobel de Economia

A Drunk and her dog: an illustration of cointegration and error correction

Two RW process



Two RW process with cointegration



Código R disponível em: github.com/pedrocostaferreira

Conceitos iniciais

- Sejam duas séries não estacionárias: $\begin{cases} Y_t \sim I(1) \\ X_t \sim I(1) \end{cases}$

- Seja a regressão: $Y_t = a + bX_t + \varepsilon_t$

- X_t e Y_t serão cointegradas se $\varepsilon_t \sim I(0)$

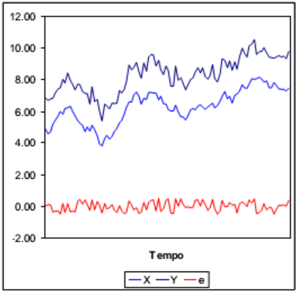
- X_t e Y_t NÃO serão cointegradas se $\varepsilon_t \sim I(1)$

Implicações

- Se Y e X são cointegradas, então:
  - Tendência estocástica comum;
  - Tendências estocásticas se cancelam mutuamente;
  - Relação de equilíbrio de longo prazo;
  - Relação de curto prazo;
  - Desvios no equilíbrio de longo prazo são **transientes**;
  - A regressão de Y contra X **não é espúria**;
- Se Y e X NÃO são cointegradas, então:
  - Tendências estocásticas são independentes;
  - Só há relação de curto prazo;
  - Desvios não tendem a se corrigir, são persistentes;
  - A regressão de Y contra X é espúria;

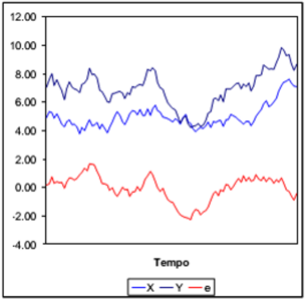
Ilustração

Cointegração



$\varepsilon_t \sim I(0)$ (ruído branco)

Não Cointegração



$\varepsilon_t \sim I(1)$ (passeio aleatório)

Propriedades

- Se $X_t \sim I(0)$ então $a + bX_t \sim I(0)$
- Se $Y_t \sim I(0)$ e $X_t \sim I(0)$ então $aY_t + bX_t \sim I(0)$
- Se $Y_t \sim I(1)$ e $X_t \sim I(0)$ então $aY_t + bX_t \sim I(1)$
- Se $Y_t \sim I(1)$ e $X_t \sim I(1)$ então (em geral) $aY_t + bX_t \sim I(1)$

Definição de Cointegração

- Se $Y_t \sim I(1)$ e $X_t \sim I(1)$, e existir uma combinação linear $aY_t + bX_t \sim I(0)$, então y_t e X_t são cointegradas.

Teste de Cointegração (Engle e Granger, 1987)

i. Computar a regressão cointegrante

$$\boxed{Y_t = \hat{a} + \hat{b}X_t + \hat{\varepsilon}_t}$$

ii. Aplicar o teste ADF sobre os resíduos

$$\boxed{\Delta\hat{\varepsilon}_t = \delta + \mu_t + \lambda \hat{\varepsilon_{t-1}} + w_t}$$

$H_0 : \lambda = 0$ (Y e X são cointegradas)
$H_a : \lambda < 0$ (Y e X não são cointegradas)

Modelo de Correção de Erro

Em caso de cointegração:

$$\boxed{\Delta Y_t = \alpha_0 + \alpha_1 \hat{\varepsilon_{t-1}} + \sum_{i=1}^{p} \beta_i \Delta Y_{t-i} + \sum_{j=1}^{k} \gamma_j \Delta X_{t-j} + u_t}$$

(Modelo de Correção de Erro)

onde: $\hat{\varepsilon_t} = Y_t - \hat{a} - \hat{b} X_t$ (Resíduos da Equação Cointegrante)

Em caso de NÃO cointegração:

$$\boxed{\Delta Y_t = \alpha_0 + \sum_{i=1}^{p} \beta_i \Delta Y_{t-i} + \sum_{j=1}^{k} \gamma_j \Delta X_{t-j} + u_t}$$

(Modelo de S/ Correção de Erro)

Observações

- Se X e Y forem $I(1)$ e cointegradas, mas só uma possuir também tendência determinística, deve-se incluir a variável t como explicativa na equação cointegrante;

- Se X e Y forem $I(1)$ e cointegradas e ambas possuírem tendência determinística, deve-se verificar no teste de cointegração se a série de erros tem tendência determinística também. Se tiver, inclua a variável t como explicativa na equação cointegrante;

Exemplo Gujarati (cap 21, pág 730)

- Seja a equação:

$$\hat{DPC}_t = \underset{(-7,48)}{-171,44} + \underset{(119,87)}{0,96}RPD_t$$

$$R^2 = 0,99 \qquad \text{d=0,53}$$

O que essas estatísticas nos dizem?

- Pelo lado positivo?
- Pelo lado negativo?

- Vamos testar e ver se as séries são não estacionárias

$$\Delta \hat{DPC}_t = 91,7 + 0,77t - \underset{\text{(-1,32)}}{0,04}DPC_{t-1} \qquad \Delta \hat{RPD}_t = 326,2 + 2,99t - \underset{\text{(-2,57)}}{0,15}RPD_{t-1}$$

Exemplo Gujarati (cap 21, pág 730)

Estatística τ: $-1,32$; $-2,57$, maiores que o valor crítico;

Não rejeito H_0: Raíz unitária

Neste caso, os testes t e F usuais são não válidos Regressão Espúria

Como resolver o problema??

Distribuição Empírica Cumulativa de Tau

	Nível de Significância			
N	1%	2.50%	5%	10%
Estatística Tau: Sem intercepto e sem tendência				
25	-2.66	-2.26	-1.95	-1.60
50	-2.62	-2.25	-1.95	-1.61
100	-2.60	-2.24	-1.95	-1.61
250	-2.58	-2.23	-1.95	-1.62
300	-2.58	-2.23	-1.95	-1.62
inf	-2.58	-2.23	-1.95	-1.62
Estatística Tau_U: Com intercepto e sem tendência				
25	-3.75	-3.33	-3.00	-2.62
50	-3.58	-3.22	-2.93	-2.60
100	-3.51	-3.17	-2.89	-2.58
250	-3.46	-3.14	-2.88	-2.57
300	-3.44	-3.13	-2.87	-2.57
inf	-3.43	-3.12	-2.86	-2.57
Estatística Tau_Tau: Com intercepto e com tendência				
25	-4.380	-3.950	-3.600	-3.240
50	-4.150	-3.800	-3.500	-3.180
100	-4.040	-3.730	-3.450	-3.150
250	-3.990	-3.690	-3.430	-3.130
300	-3.980	-3.680	-3.420	-3.130
inf	-3.960	-3.660	-3.410	-3.120

Exemplo Gujarati (cap 21, pág 730)

- Vamos testar a cointegração

Teste de Engle Granger (AEG)

Como vimos, quero testar se a combinação linear entre as duas variáveis é estacionária

$$\boxed{u_t = DPC_t - \beta_1 - \beta_2 RPD_t}$$

Submetemos o resíduo da regressão ao teste ADF

$$\Delta\hat{u} = \underset{(-5,67)}{-0,27\hat{u_{t-1}}}$$

Dado os valores críticos de Engle-Granger a estatística t rejeita a hipótese nula de raiz unitária, portanto as séries se cointegram.

Exemplo Gujarati (cap 21, pág 730)

- Solução: Mecanismo de Correçao de Erro

Seguindo a metodologia proposta por Engle e Granger, temos a seguinte equação estimada:

$$\Delta D\hat{P}C = 11,6 + \underset{(4,17)}{0,29}\Delta RPD_t - \underset{(-3,60)}{0,08}\hat{u_{t-1}}$$

$\Uparrow$

Se significante, nos diz qual é a proporção do desequilíbrio em DPC em um período é corrigida no período seguinte

- Conclusão

Alterações a curto prazo na RPD tem efeitos positivos sobre o DPC e que cerca de 0,08 entre o valor efetivo e o valor de equilíbrio de DPC é corrigida a cada trimestre.

Aplicação à expectativa de inflação dos consumidores

A Sondagem do Consumidor da FGV coleta mensalmente informações de mais de 2100 brasileiros em sete das principais capitais do país (Porto Alegre, São Paulo, Rio de Janeiro, Belo Horizonte, Salvador, Brasília e Recife). Cerca de 75% destes entrevistados vêm respondendo aos quesitos relacionados às expectativas de inflação.

A pergunta quantitativa possui a seguinte formulação: na sua opinião, de quanto será a inflação brasileira nos próximos 12 meses?

Os resultados da pergunta quantitativa divulgados nesse relatório são obtidos em duas etapas, da seguinte forma:

- Agregação, sem ponderação, das respostas individuais em cada um dos 28 níveis de agregação (4 níveis de renda e 7 capitais);
- Agregação dos diferentes níveis geográficos e de renda por pesos determinados pelo consumo.

Literatura

- FERREIRA, P.G.C. et. al. **Análise de Séries Temporias em R: um curso introdutório**. Disponível em: https://pedroferreira.shinyapps.io/timeseries/
- Spurious correlations. Available at: http://tylervigen.com/spurious-correlations
- Shiny App - Time Series - https://pedroferreira.shinyapps.io/timeseries/
- COWPERTWAIT, P. S. P., METCALFE, A. V. **Introductory Time Series with R.** Springer. 2009.
- GUJARATI, D. **Econometria Básica.** São Paulo: Makron Books. Quarta Edição, 2006.