Complementos de Econometria

Licenciatura em Economia

ISCTE-IUL, Dep. de Economia

MODELOS EM TIME SERIES (PARTE 2)

Luís Filipe Martins

luis.martins@iscte.pt

http://iscte.pt/~lfsm

Departamento de Métodos Quantitativos,

ISCTE-IUL, Escola de Gestão

Lisboa, Outubro de 2009

1 Recapitulação (ver apontamentos)

- MRLM (incluindo o termo independente):

$$\begin{aligned} y_t &= \beta_1 + \beta_2 x_{t2} + ... + \beta_k x_{tk} + u_t = x_t^{'}\beta + u_t, t = 1, ..., T \quad (1) \\ y &= X\beta + u \\ x_{t_{k\times 1}} &= (1, x_{t2}, ..., x_{tk})^{'}, t = 1, ..., T; \beta_{k\times 1} = (\beta_1, \beta_2, ..., \beta_k)^{'}; \\ y_{T\times 1} &= \begin{pmatrix} y_1 \\ ... \\ y_T \end{pmatrix}; X_{T\times k} = \begin{pmatrix} 1 & x_{12} & ... & x_{1k} \\ ... & ... & & ... \\ 1 & x_{T2} & ... & x_{Tk} \end{pmatrix}; u_{T\times 1} = \begin{pmatrix} u_1 \\ ... \\ u_T \end{pmatrix} \end{aligned}$$

- Sob hipoteses Gauss-Markov,

$$\widehat{\beta}_{OLS} = \left(X^{'}X\right)^{-1} X^{'}y = \left(\sum_{t=1}^{T} x_t x_t^{'}\right)^{-1} \sum_{t=1}^{T} x_t y_t \qquad (2)$$

é BLUE (inclui consistencia). Usar formulas de Econometria 1, nomeadamente,

$$\widehat{V\left(\widehat{\beta}|X\right)} = \widehat{\sigma}_u^2 \left(X^{'}X\right)^{-1}. \qquad (3)$$

- MAS, admita-se que em modelos com time series os erros estão autocorrelacionados

$$Cov(u_t, u_s|X) = Cov(u_t, u_s|x_1, ..., x_T) \neq 0, t, s = 1, ..., T, \qquad (4)$$

para algum $t \neq s$. Então, OLS continua a ser consistente mas não é BLUE (há outro estimador - que não o OLS - que é mais eficiente: FGLS) e a formula usual de $V\widehat{\left(\widehat{\beta}|X\right)}$ para o OLS passa a ser

$$\widehat{\Sigma}_{k\times k}=V\widehat{\left(\widehat{\beta}|X\right)}=\left(X^{'}X\right)^{-1}X^{'}\widehat{\Omega}X\left(X^{'}X\right)^{-1}, \quad (5)$$

onde $\Omega_{T\times T}=E(uu'|X)\neq\sigma_u^2I_T$.

- Nestas circunstâncias, sugere-se MANTER o OLS (não usar o FGLS) mas usar uma matriz $\widehat{\Sigma}_{k\times k}$ robusta a autocorrelação. Essa formula deve-se a Newey and West (HAC).

- Pela simplicidade e boas propriedades, assumir que o modelo para o u_t é um AR(1):

$$u_t=\rho u_{t-1}+\varepsilon_t,\varepsilon_t\sim wn\left(0,\sigma_\varepsilon^2\right),t=1,...,T. \quad (6)$$

  Então, u_t é estacionário e de memória que se dissipa para $|\rho|<1$ e, neste caso,

$$E\left(u_t\right)=0;V\left(u_t\right)=\frac{\sigma_\varepsilon^2}{\left(1-\rho^2\right)};\Gamma\left(l\right)=\rho^lV\left(u_t\right);Cor\left(u_t,u_{t+l}\right)=\rho^l. \quad (7)$$

- Testes à (ausência) de autocorrelação dos erros:
  - Durbin-Watson (testa AR(1) em modelos estáticos)
  - $h-$Durbin e $h-alt$ (testa AR(1) em modelos dinâmicos)
  - Breusch-Godfrey (testa AR(p), $u_t=\rho_1u_{t-1}+...+\rho_pu_{t-p}+\varepsilon_t$)

2 Variaveis Dummy, Tendência Deterministica e Sazonalidade

- Em time series também faz sentido usarem-se variaveis Dummy como regressores (ver Econometria 1). Exemplo:

$$C_t = \beta_1 + \beta_2 R_t + \beta_3 D_t + \beta_4 D_t R_t + u_t, \tag{8}$$

em que C_t é o nivel de consumo em t; R_t é o rendimento e D_t é uma Dummy em que $D_t = 1$ para t em tempo de paz e $D_t = 0$ para t em tempo de guerra. Qual é a interpretação de β_3 e β_4? Pode-se (e deve-se) aplicar os testes de Chow de Econometria 1. Outro exemplo é em modelos com sazonalidade (ver em baixo).

- Modelos com tendencia deterministica:

$$\begin{aligned} \text{(linear) } y_t &= \beta_1 + \beta_2 t + u_t, t = 1, ..., T \\ \text{(quadrática) } y_t &= \beta_1 + \beta_2 t + \beta_3 t^2 + u_t, \\ \text{(exponencial) } \log y_t &= \beta_1 + \beta_2 t + u_t. \end{aligned} \tag{9}$$

Podem-se incluir regressores x_t; Estimação pode ser por OLS (ver hipoteses) MAS sem testes t porque a distribuição não é a standard; Usar o R^2;...

- Utilidade: Retirar tendência (deterministica) a y_t ("detrending"). (1) Estimar por OLS o β do modelo com tendencia - $\widehat{\beta}$; (2) Subtrair $\widehat{y}_t = \widehat{\beta}_1 + \widehat{\beta}_2 t$ a y_t para obter a série sem tendencia (detrended) $\widehat{y}_t^d = y_t - \widehat{y}_t$ (que é igual a $\widehat{u}_t$ quando não há x_t). Assim, y_t cresce/decresce com o tempo mas $\widehat{y}_t^d$ fica mais ou menos constante com o tempo.

- Modelos com sazonalidade (por exemplo, y_t é o volume de vendas de brinquedos e os dados são mensais). Sazonalidade reflecte impactos motivados por factores de periodicidade inferior a um ano. Neste caso,

$$y_t = \beta_1 + \beta_2 S_{2t} + ... + \beta_{12} S_{12,t} + u_t, \qquad (10)$$

em que $S_{j,t}$ é uma dummy onde $S_{j,t} = 1$ se t é uma observação do mês j e $S_{j,t} = 0$ se t não é uma observação do mês j.

- Utilidade: Retirar sazonalidade a y_t ("dessasonalização"). (1) Estimar por OLS o β do modelo com $S_t's$ - $\widehat{\beta}$; (2) Subtrair $\widehat{y}_t$ a y_t para obter $\widehat{y}_t^s$.

3 Modelos com desfasamentos e modelos dinâmicos

- Modelo DL(q) (distributed lag of order q = modelo com um numero finito q de desfasamentos):

$$y_t = \alpha_0 + \delta_0 x_t + \delta_1 x_{t-1} + ... + \delta_q x_{t-q} + u_t, t = 1, ..., T \quad (11)$$

Perda de q observações; Estimador OLS com propriedades standard (ver hipoteses do MRLM); Pode haver mais do que um regressor (não incluir a propria variavel dependente y) com desfasamentos;...

- Como

$$E\left(y_t | x_t, ..., x_{t-q}\right) = \alpha_0 + \delta_0 x_t + \delta_1 x_{t-1} + ... + \delta_q x_{t-q}, \quad (12)$$

definem-se os multiplicadores de (i) impacto imediato

$$\delta_0 = \frac{\partial E\left(y_t | x_t, ..., x_{t-q}\right)}{\partial x_t}; \quad (13)$$

(ii) impacto a $s-$passos, $1 \leq s \leq q$ (multiplicador de curto prazo a s periodos)

$$\delta_s = \frac{\partial E\left(y_t | x_t, ..., x_{t-q}\right)}{\partial x_{t-s}}, s = 1, ..., q; \quad (14)$$

e (iii) longo-prazo

$$\delta = \sum_{i=0}^{q} \delta_i = \sum_{i=0}^{\infty} \delta_i. \quad (15)$$

- O multiplicador de LP tambem pode ser encontrado atraves do modelo em equilibrio por $\delta = \frac{\partial y^*}{\partial x^*}$ onde

$$y^* = \alpha_0 + \delta_0 x^* + \delta_1 x^* + ... + \delta_q x^* = \alpha_0 + \sum_{i=0}^{q} \delta_i x^*. \quad (16)$$

- Modelo AR(p) - dinâmico - (já visto antes para os erros do modelo principal):

$$y_t = \rho_0 + \rho_1 y_{t-1} + ... + \rho_p y_{t-p} + u_t, t = 1, ..., T \quad (17)$$

  Estimador OLS com propriedades standard (ver hipoteses do MRLM, nomeadamente erros u_t não autocorrelacionados).

- Modelo ADL(p,q):

$$y_t = \rho_0 + \rho_1 y_{t-1} + ... + \rho_p y_{t-p} + \delta_0 x_t + \delta_1 x_{t-1} + ... + \delta_q x_{t-q} + u_t. \quad (18)$$

  Estimador OLS com propriedades standard (ver hipoteses do MRLM, nomeadamente erros u_t não autocorrelacionados); Pode haver mais do que um regressor (não incluir a propria variavel dependente y) com desfasamentos; Para calcular os multiplicadores de CP devemos ter em atenção que um "choque"de x em t faz variar y em $t, t + 1$ (via y_t)$, t + 2, ...$; Para os multiplicadores de LP recomenda-se o uso do modelo em equilibrio

$$\begin{aligned} y^* &= \rho_0 + \rho_1 y^* + ... + \rho_p y^* + \delta_0 x^* + \delta_1 x^* + ... + \delta_q x^* && (19) \\ &= \rho_0 + \sum_{i=1}^{p} \rho_i y^* + \sum_{i=0}^{q} \delta_i x^* \Leftrightarrow && (20) \end{aligned}$$

$$\left(1 - \sum_{i=1}^{p} \rho_i\right) y^* = \rho_0 + \sum_{i=0}^{q} \delta_i x^* \quad (21)$$

- Modelo ADL(1,1):

$$y_t = \rho_0 + \rho_1 y_{t-1} + \delta_0 x_t + \delta_1 x_{t-1} + u_t, t = 1, ..., T \quad (22)$$

Para $|\rho_1| < 1$, tem-se que

$$y_t = \frac{1}{(1 - \rho_1 L)} \rho_0 + \frac{(\delta_0 + \delta_1 L)}{(1 - \rho_1 L)} x_t + u_t, t = 1, ..., T, \quad (23)$$

onde L é o operador desfasamento ($Ly_t = y_{t-1};\ L^2 y_t = y_{t-2}$). Para $\delta_1 = 0$, e como

$$\frac{1}{(1 - \rho_1 L)} = 1 + \rho_1 L + \rho_1^2 L^2 + \rho_1^3 L^3 + ..., \quad (24)$$

os multiplicadores de CP podem ser obtidos através do DL(∞)

$$y_t = \frac{\rho_0}{(1 - \rho_1)} + \delta_0 \sum_{i=0}^{\infty} \rho_1^i x_{t-i} + u_t. \quad (25)$$

Este DL(∞) tem desfasamentos geometricos (Koyck) porque os coeficientes $\delta_0\rho_1^i$ decaem geometricamente para zero quando i tende para infinito.

- Muitas vezes, a autocorrelação dos erros resulta de uma "má especificação" (escolha) do modelo principal. Senão vejamos... Considere-se o modelo $y_t = \delta x_t + u_t$ em que $u_t = \rho u_{t-1} + \varepsilon_t$. Este modelo é "equivalente" a um ADL(1,1) com erros G.M.

$$y_t = \rho y_{t-1} + \delta x_t - \delta\rho x_{t-1} + \varepsilon_t \tag{26}$$

porque $u_t = \frac{1}{(1-\rho L)}\varepsilon_t$.

4 Testes de raíz unitária

- As propriedades do OLS são as standard desde que as variaveis do MRLM sejam "estacionarias em covariância". Essa hipotese tem sido assumido implicitamente no estudo. Uma serie é "estacionaria" se, dito de uma forma simplista, flutua uniformemente ao longo do tempo em torno de um valor fixo. Estas series não "explodem", não "divergem" no tempo. São variáveis em "equilibrio" (ver multiplicador de LP). Matemáticamente, (1) $E(y_t) = \mu$, constante, para todo t; (2) $V(y_t) = \sigma^2$, constante, para todo t; (3) $Cov(y_t, y_{t-l}) = \gamma_l$, uma função do desfasamento l mas constante para todo t.
- Exemplos de modelos em que y_t é estacionária: $y_t = \beta + u_t$ ou $y_t = \rho y_{t-1} + \varepsilon_t$, para $|\rho| < 1$.
- MAS os seguintes modelos definem variaveis que não são estacionárias ("explodem";"divergem")

$$\begin{aligned} y_t &= \beta_1 + \beta_2 t + u_t && (27) \\ y_t &= y_{t-1} + \varepsilon_t \text{ (passeio aleatório sem deriva)} \\ y_t &= \alpha + y_{t-1} + \varepsilon_t \text{ (passeio aleatório com deriva)} && (28) \end{aligned}$$

- Sabe-se que o estimador OLS no MRLM com variáveis não estacionárias não tem as propriedades standard. Nestas condições, o OLS é superconsistente, a distribuição não é a usual (testes t tendem a ser muito elevados) e o R^2 tende a ser próximo de 1.

- Portanto, convém distinguir/testar a estacionaridade das series antes de proceder à estimação OLS do MRLM.
- Considere-se o modelo AR(1) $y_t = \rho y_{t-1} + \varepsilon_t$. Se $\rho = 1$ o processo não é estacionário (passeio aleatório; I(1); tem raíz unitária); Se $|\rho| < 1$ o processo é estacionário (I(0); não tem raíz unitária).
- Teste DF (Dickey-Fuller):
  - O modelo $y_t = \rho y_{t-1} + \varepsilon_t$ é equivalente a

$$\triangle y_t = \theta y_{t-1} + \varepsilon_t, \tag{29}$$

    onde $\triangle y_t = y_t - y_{t-1}$ e $\theta = (\rho - 1)$.
  - Fazer teste t ao parametro θ para testar $H_0 : \rho = 1$ (não estacionaridade) versus $\rho < 1$ (estacionaridade).
  - A distribuição do teste não é a usual (ver tabela dos valores criticos). One sided (left).
  - Pode ser feito para

$$\triangle y_t = \alpha + \theta y_{t-1} + \varepsilon_t \text{ ou } \triangle y_t = \alpha + \beta t + \theta y_{t-1} + \varepsilon_t \tag{30}$$

    mas as distribuições (valores criticos) são diferentes.
  - Teste ADF assume que ε_t é autocorrelacionado:

$$\triangle y_t = \theta y_{t-1} + \lambda_1 \triangle y_{t-1} + ... + \lambda_p \triangle y_{t-p} + \varepsilon_t, \tag{31}$$

    e a distribuição é a mesma da do teste DF.
- Teste PP (Phillips-Perron) tem as mesmas hipoteses da do DF

mas a estatistica e a distribuição são diferentes

- Teste KPSS inverte as hipoteses (a hipotese nula é de estacionaridade) e a estatistica e distribuição são diferentes das do DF e do PP.

5 Cointegração e modelo com mecanismo de correcção de erros

- Regressão espuria (desprovida de qualquer sentido economico): $y_t = \beta x_t + u_t$ em que y_t e x_t são passeios aleatórios (I(1)) INDEPENDENTES. Como explicado anteriormente, sabe-se que o OLS é superconsistente, a distribuição não é a usual (testes t tendem a ser muito elevados), o R^2 tende a ser próximo de 1 e a DW tende a ser próxima de 0 (isto porque u_t é igualmente tipo I(1)). Este resultado é espurio porque apesar de não haver qualquer relação entre y_t e x_t os testes t são (enganadoramente) significativos!

- Mas há casos em que y_t e x_t são passeios aleatórios (I(1)) mas em que existe $\beta \neq 0$ tal que $y_t - \beta x_t = u_t$ é estacionário (I(0)). Exemplo: Consumo e Produto; Taxas de juro com diferentes maturidades;... Nestas condições, diz-se que y_t e x_t estão cointegradas, u_t é o erro de equilibrio e β é o parametro/vector de cointegração.

- Plots: No caso de relação espuria, as duas series divergem individualmente e não aparentam mover-se par a par. No caso de cointegração, as series divergem individualmente mas há uma "força" que as faz caminhar lado a lado.

- Teste de cointegração de Engle e Granger (EG): Como o que distingue a regressão espuria da relação de cointegração é o processo ε_t (se é I(1) ou I(0)) o teste de cointegração é equivalente a um teste de raiz unitária (DF) aos residuos OLS $y_t - \widehat{\beta} x_t = \widehat{u}_t$. Sob H_0 não há cointegração; Sob H_1 há cointegração. A distribuição do teste não é igual à do DF.

- Se há cointegração, então um modelo a considerar é o ECM (modelo com mecanismo de correção de erros):

$$\triangle y_t = \alpha + \gamma \left(y_{t-1} - \beta x_{t-1} \right) + \phi \triangle x_t + \varepsilon_t, \tag{32}$$

onde todas as variaveis são estacionarias ($\triangle y_t, \triangle x_t$ e $y_{t-1} - \beta x_{t-1}$). Como β não é conhecido, estimação a dois passos OLS: (1) Estimar β em $y_t = \beta x_t + u_t$; (2) Estimar

$$\triangle y_t = \alpha + \gamma \left(y_{t-1} - \widehat{\beta} x_{t-1} \right) + \phi \triangle x_t + \varepsilon_t. \tag{33}$$

Neste modelo, a dinâmica de curto prazo está em $\triangle x_t$ e a de longo prazo (equilibrio) está em $y_t - \widehat{\beta} x_t$. Porque é que "convêm" que γ seja negativo e não nulo ($\gamma < 0$)? Qual é a interpretação para γ?

- Resumindo:
  - Se y_t e x_t são I(0), usar modelo em niveis: $y_t = \beta x_t + u_t$.
  - Se y_t e x_t são I(1) não cointegradas, usar modelo em primeiras diferenças (variações): $\triangle y_t = \beta \triangle x_t + u_t$.
  - Se y_t e x_t são I(1) cointegradas, usar modelo ECM: $\triangle y_t = \alpha + \gamma \left(y_{t-1} - \beta x_{t-1} \right) + \phi \triangle x_t + \varepsilon_t$.

6 Previsão

- Prever uma (ou mais) serie é construir um valor (ou intervalo de confiança) plausivel para o futuro (ainda não observado) com base num modelo e no histórico da mesma.

- Previsão out–of-sample: Do total de observações T, separar em R e $T - R = m$ periodos. As R observações (in-sample) servem para estimar o modelo que será usado e avaliado para efeitos de previsão; As m observações são usadas para calcular os erros de previsão.

- Com base num modelo prever $\widehat{y}_{t+h}$ (h-passos). Normalmente, 1-passo: $\widehat{y}_{t+1}$. Por exemplo, para $y_t = \beta_1 + \beta_2 t + u_t$ temos que $\widehat{y}_{t+1} = \widehat{\beta}_1 + \widehat{\beta}_2 (t+1)$. Por exemplo, para $y_t = y_{t-1} + \varepsilon_t$ temos que $\widehat{y}_{t+1} = y_t$.

- Erro de previsão: $e_{t+1} = y_{t+1} - \widehat{y}_{t+1}$.

- Medida de erro: Como há m previsões e erros de previsão, $e_{R+k}, k = 1, ..., m$, construir RMSE (raiz quadrada do erro quadratico médio): $RMSE = \sqrt{\frac{1}{m} \sum_{k=1}^{m} e_{R+k}^2}$. Entre modelos alternativos para previsão, escolher o que tem um menor RMSE. Outras medidas: Mean absolute error; Mean relative error.

- Previsão por extrapolação: Com base num dado modelo, prever $\widehat{y}_{T+h}$. Como $T + h$ ainda não foi observado não há erro de previsão até esse momento ocorrer.

7 O Modelo ARCH

- O modelo ARCH(1) de Engle, 1982, (autoregressive conditional heteroskedasticity of order one) é útil para modelar a VARIÂNCIA CONDICIONAL (não o valor esperado condicional) quando esta não é constante ao longo do tempo como é, por exemplo, o caso dos retornos de activos (bolsa em geral).

- Modelo em níveis

$$y_t = \beta_1 + \beta_2 x_{t2} + ... + \beta_k x_{tk} + u_t = x_t^{'} \beta + u_t, t = 1, ..., T \quad (34)$$

  que, sob $E(u_t|x_t) = 0$,

$$E(y_t|x_t) = x_t^{'} \beta \quad (35)$$

  e com erros homocedasticos

$$V(u_t|x_t) = \sigma^2 \quad (36)$$

  mas que tem erros heterocedasticos CONDICIONADO no seu passado:

$$V(u_t|u_{t-1}) = E\left(u_t^2|u_{t-1}\right) = \alpha_0 + \alpha_1 u_{t-1}^2, \quad (37)$$

  onde $\alpha_0 > 0, \alpha_1 \geq 0$ para que $V(u_t|u_{t-1}) > 0$. Para "estabilidade" $\alpha_1 < 1$.

- Distinguir Heterocedasticidade, $V(u_t|x_t) = h(x_t)$ (não é o interesse deste modelo), de Heterocedasticidade Condicional - $V(u_t|u_{t-1}, x_t) = V(u_t|u_{t-1}) = f(u_{t-1})$.

- Para que $V(u_t|u_{t-1}) = E\left(u_t^2|u_{t-1}\right)$ temos de assumir que

$E(u_t|u_{t-1}) = 0$, ou seja, os erros não estão autocorrelacionados (não podem ser AR(1)).

- Podendo escrever o modelo como

$$u_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + v_t, \tag{38}$$

onde $E(v_t|u_{t-1}) = 0$, notamos que $y_t = x_t^{'}\beta + u_t$ não é linear nos parametros e por isso não se deve usar OLS. Alternativas: OLS não linear ou Máxima verosimilhança.

- Teste a efeitos ARCH: Teste t a α_1 no modelo

$$\widehat{u}_t^2 = \alpha_0 + \alpha_1 \widehat{u}_{t-1}^2 + v_t, \tag{39}$$

com residuos OLS.

- Plot de variâncias condicionais (estimadas) no tempo:

$$V\widehat{(u_t|u_{t-1})} = \widehat{\alpha}_0 + \widehat{\alpha}_1 \widehat{u}_{t-1}^2. \tag{40}$$