Macroeconometria 2

Mestrado em Economia Monetária e Financeira
Mestrado em Economia

ISCTE-IUL, Dep. de Economia

MODELOS NÃO-LINEARES

Luís Filipe Martins

luis.martins@iscte.pt
http://iscte.pt/~lfsm

Departamento de Métodos Quantitativos,
ISCTE-IUL, Escola de Gestão

Lisboa, Fevereiro de 2010

1 Motivação

- Evidência Empirica (*stylized facts*): Várias séries economicas e financeiras apresentem propriedades estatisticas (momentos e parâmetros) que se distinguem ao longo da amostra. A presença de quebras na estrutura das séries é mais a regra do que a excepção! Algumas séries temporais são mais bem descritas atráves de não-linearidades.

Example 1 *$s_t = i_t^l - i_t^s$, onde s_t é o spread relativamente às taxas de juro nominais de curto (i_t^s) e longo prazo $\left(i_t^l\right)$. A persistência (nível de "correlação temporal") é diferente quando $s_{t-1} \geq 0$ de $s_{t-1} < 0$.*

Example 2 *Para a economia mexicana e no periodo $t = 1978, ..., 198$ considere-se o rácio D_t/P_t onde D_t e P_t são, respectivamente, o volume de dinheiro em dolares e em pesos em contas bancárias. A série $\log D_t - \log P_t$ flutua uniformemente em torno de -2 até 1982 e devido a um choque estrutural subito passa a variar regularmente em torno de -3.25 após 1982. Como modelar esta série?*

Example 3 *US real GNP: Two state*

$$y_t - \mu_{s_t} = \phi_1 \left(y_{t-1} - \mu_{s_{t-1}}\right) + ... + \phi_4 \left(y_{t-4} - \mu_{s_{t-4}}\right) + \varepsilon_t.$$

Finding: $\widehat{\mu}_1 = 1.16\%$; $\widehat{\mu}_2 = -0.36\%$; $\widehat{p}_{11} = 0.90$; $\widehat{p}_{22} = 0.75$; expansões duram em média 10 trimestres; depressões 4.

2 Modelos TAR (Threshold AR)

- A alteração de regime é determinada por uma variavel observada.
- Simples caso (SETAR - self-exciting TAR -, de 1^a ordem, com threshold igual a zero e identicos choques):

$$y_t = \begin{cases} \rho_1 y_{t-1} + \varepsilon_t \text{ se } y_{t-1} > 0 \\ \rho_2 y_{t-1} + \varepsilon_t \text{ se } y_{t-1} \leq 0 \end{cases}, \text{ onde } \rho_1 \neq \rho_2. \tag{1}$$

  Os parametros ρ_1 e ρ_2 descrevem o diferente nivel de persistencia de y_t de $y_{t-1} > 0$ em relação a $y_{t-1} \leq 0$. O valor de zero é interpretado como o sendo o nivel de equilibrio (longo-prazo).

- Estimação de ρ_1, ρ_2 e σ_ε: Neste caso, Minimos Quadrados porque este modelo é equivalente a

$$y_t = \rho_1 I_t y_{t-1} + \rho_2 (1 - I_t) y_{t-1} + \varepsilon_t, \text{ onde } I_t = 1(y_{t-1} > 0), \tag{2}$$

  função indicador.

- Teste a AR: $H0 : \rho_1 = \rho_2$ Estatistica usual F (se o threshold é desconhecido então a estatistica deve ser outra - não standard).
- Extensões:
  - TAR (não é "self-exciting")

$$y_t = \begin{cases} \rho_1 y_{t-1} + \varepsilon_t \text{ se } S_{t-1} > 0 \\ \rho_2 y_{t-1} + \varepsilon_t \text{ se } S_{t-1} \leq 0 \end{cases}, \tag{3}$$

em que S_{t-1} é a variável de estado.

- Threshold diferente de zero (conhecido *versus* desconhecido) e com constante

$$y_t = \begin{cases} \mu_1 + \rho_1 y_{t-1} + \varepsilon_t \text{ se } y_{t-1} > \tau \\ \mu_2 + \rho_2 y_{t-1} + \varepsilon_t \text{ se } y_{t-1} \leq \tau \end{cases}. \tag{4}$$

- Demeaned

$$y_t = \begin{cases} \mu + \rho_1 \left(y_{t-1} - \mu\right) + \varepsilon_t \text{ se } y_{t-1} > \mu \\ \mu + \rho_2 \left(y_{t-1} - \mu\right) + \varepsilon_t \text{ se } y_{t-1} \leq \mu \end{cases}. \tag{5}$$

- Ordem p de desfazamento na parte AR

$$y_t = \begin{cases} \rho_{11} y_{t-1} + \rho_{12} y_{t-2} + ... + \rho_{1p} y_{t-p} + \varepsilon_t \text{ se } y_{t-1} > 0 \\ \rho_{21} y_{t-1} + \rho_{22} y_{t-2} + ... + \rho_{2p} y_{t-p} + \varepsilon_t \text{ se } y_{t-1} \leq 0 \end{cases}. \tag{6}$$

- Ordem k de desfazamento na variável de estado

$$y_t = \begin{cases} \rho_1 y_{t-1} + \varepsilon_t \text{ se } y_{t-k} > 0 \\ \rho_2 y_{t-1} + \varepsilon_t \text{ se } y_{t-k} \leq 0 \end{cases}, \tag{7}$$

onde k é o numero de periodos de desfazamento com que a variavel de estado influencia o regime na variavel principal.

- Diferentes choques

$$y_t = \begin{cases} \rho_1 y_{t-1} + \varepsilon_{1t} \text{ se } y_{t-1} > 0 \\ \rho_2 y_{t-1} + \varepsilon_{2t} \text{ se } y_{t-1} \leq 0 \end{cases}, \text{ onde } V\left(\varepsilon_{1t}\right) \neq V\left(\varepsilon_{2t}\right). \tag{8}$$

- Mais do que 2 regimes; Diferentes ordens de desfazamente entre regimes; Apenas parte dos parametros mudam com o regime; Especificação diferente do tipo AR (eg: STAR); A variavel de estado é diferente em cada regime; ...

- Se o parametro de threshold é desconhecido, então a estimação de minimos quadrados é feita a dois passos (o τ é escolhido com base numa grelha) = OLS não linear (ou MLE).
- Pode-se usar AIC para a selecção de modelos $(p, k, ...)$.

3 Modelos MS (Markov-Switching)

3.1 Introdução e Markov Chains

- Retomemos o exemplo da série $\log D_t - \log P_t \equiv y_t$. Como modela-la? Uma hipótese é usando $AR(1)$ (demeaned) para cada regime, $\begin{cases} y_t - \mu_1 = \phi\left(y_{t-1} - \mu_1\right) + \varepsilon_t, & \text{se } t \leq t_0 \\ y_t - \mu_2 = \phi\left(y_{t-1} - \mu_2\right) + \varepsilon_t, & \text{se } t > t_0 \end{cases}$, $\mu_2 < \mu_1$. Questões:
  (a) Como se estima t_0? Corresponderá de facto a 1982? O que causou esta quebra abrupta?
  (b) Como se estimam os parametros do modelo μ_1, μ_2?
  (c) Como se prevê esta série com base neste modelo?
  (d) Qual é a lei probabilistica (DGP) para $\{y_t\}$?
- Modelo equivalente:
$$y_t - \mu_{s_t^*} = \phi\left(y_{t-1} - \mu_{s_{t-1}^*}\right) + \varepsilon_t, t = 1, ..., T, \tag{9}$$
onde o estado ou regime $s_t^* = 1$ se $t = 1, 2, ..., t_0$ e $s_t^* = 2$ se $t = t_0 + 1, t_0 + 2, ..., T$. No entanto, na prática, s_t não é observado e é aleatória a sua realização. Qual é a lei probabilistica para $\{s_t\}$? Por simplicidade, assumir que a time series, sendo uma sequência de v.a. discretas, é uma cadeia de Markov.
- $\{s_t\}$ $\sim$Cadeia de Markov com N regimes ($\mathcal{S} = \{1, ..., N\}$) e

probabilidades de transição $\{p_{ij}\}_{i,j=1}^{N}: \sum_{j=1}^{N} p_{ij} = 1,$

$$P\left(s_t = j | s_{t-1} = i, s_{t-2} = k, ...\right) = P\left(s_t = j | s_{t-1} = i\right) = p_{ij}, \tag{10}$$

$$\mathbf{P}_{N\times N} = [p_{ij}] = \begin{pmatrix} p_{11} & ... & p_{N1} \\ ... & & ... \\ p_{1N} & ... & p_{NN} \end{pmatrix}.$$

- Irreducible M.C.: Não há regime(s) absorvente(s) (c.f. $\mathbf{P}$).
- Ergodic Irreducible M.C. para $N = 2 : p_{11} < 1, p_{22} < 1, p_{11} + p_{22} > 0$, onde $\pi_1 = P\left(s_t = 1\right) = \frac{1-p_{22}}{2-p_{11}-p_{22}}, \pi_2 = 1 - \pi_1.$

3.2 Modelo de Hamilton e Estimação, Previsão e Inferência

- Geral, linear,

$$y_t = \beta_{s_t}^{'} x_t + \varepsilon_t, t = 1, ..., T, \tag{11}$$

onde $\varepsilon_t \sim N\left(0, \sigma_{s_t}^2\right)$ independente $\forall t$ de s_t que é modelado como uma não observada $N-$state Markov chain.

Example 4

$$\begin{aligned} y_t &= \mu_{s_t} + \phi_{s_t} y_{t-1} + \varepsilon_t, \varepsilon_t \sim i.i.d.N\left(0, \sigma^2\right), N = 2 \\ \alpha &= \left(\mu_1, \mu_2, \phi_1, \phi_2, \sigma^2\right)^{'}; \theta = \left(\alpha^{'}, p_{11}, p_{22}\right)^{'}; j = 1, 2, \end{aligned}$$

- Notas:
  (a) Data $\mathcal{Y}_t = \left(y_t, y_{t-1}, ..., y_{-m}, x_t^{'}, x_{t-1}^{'}, ..., x_{-m}^{'}\right)^{'}$ continuous + Regimes s_t discrete *implica* uma MIXTURE Joint PDF.

(b) Parametros do modelo + Prob. associadas ao regime *implica* MLE com recurso a um processo iterativo na lógica do EM algoritm.

(c) Assume-se

$$f\left(y_t|x_t, \mathcal{Y}_{t-1}, s_t = j; \alpha\right) = f\left(y_t|x_t, \mathcal{Y}_{t-1}, s_t = j, s_{t-1} = i, ..., ; \alpha\right) \tag{12}$$

(d) Assume-se que MC s_t é independente de $\mathcal{Y}_{t-1}$ e x_t (além do erro ε_t).

- Informação probabilistica sobre qual o regime num dado periodo t dada a informação disponivel até esse momento $\mathcal{Y}_t : \widehat{\xi}_{t|t} = E\left(\xi_t|\mathcal{Y}_t\right) = \left(P\left(s_t = 1|\mathcal{Y}_t\right), ..., P\left(s_t = N|\mathcal{Y}_t\right)\right)'$ onde $\xi_t = \mathbf{e}_i$ (i^{th} col. de I_N) quando $s_t = i$. A sequência $\left\{\widehat{\xi}_{t|t}\right\}$ é denominada por *Filtered prob*. (one-step state forecast given $\mathcal{Y}_t$).

- Previsão optima a m passos: $\widehat{\xi}_{t+m|t} = \mathbf{P}^m \widehat{\xi}_{t|t}$.

- *Smoothed prob.* com base em toda a amostra: $\widehat{\xi}_{t|T} = E\left(\xi_t|\mathcal{Y}_T\right)$.

- Previsão para y_t? Demonstra-se que $E\left(y_{t+1}|\mathcal{Y}_t; \theta\right) = h_t'\widehat{\xi}_{t+1|t}$, onde o $(N \times 1)$ vector h_t tem como j^{th} elemento $E\left(y_{t+1}|s_{t+1} = j, \mathcal{Y}_t; \theta\right) = \mu_j + \phi_j y_t$, por exemplo.