Macroeconometria 1

Mestrado em Economia Monetária e Financeira

Mestrado em Economia

ISCTE-IUL, Dep. de Economia

MODELOS LINEARES (ARIMA)

Luís Filipe Martins

luis.martins@iscte.pt

http://iscte.pt/~lfsm

Departamento de Métodos Quantitativos,

ISCTE-IUL, Escola de Gestão

Lisboa, Setembro de 2009

1 Motivação

- (Parte 1: Análise Univariada) Evidência Empirica (*stylized fact*): As variáveis de natureza economica apresentam dependencia temporal (memória) a três níveis:
  (a) A memória dissipa-se rapidamente (função geométrica) com o passado.
  (b) A memória é longa (função hiperbolica), i.e., há dependencia relativamente significativa até um passado longicuo.
  (c) A memória é permanente e significativa.
- (Parte 2: Análise Multivariada - Ver VAR e Cointegração)

2 Introdução e Algumas Definições

- Objectivo**:** Modelar Time Series (séries cronológicas). *Exemplo*: Modelos ARIMA
  - Simplicidade.
  - Ajustamento e Previsão.
  - Lineares; Univariados e (Não)Estacionaridade.
- Time Series (ou processo estocástico), $\{Y_t\}_{t=1}^{T} = \{Y_1, Y_2, ..., Y_T\}$ é uma sequência de variáveis aleatórias.
- Realização da Time Series, $\{y_t\}_{t=1}^{T} = \{y_1, y_2, ..., y_T\}$ é o gráfico com as observações de uma variável time series.

Example 1 *taxa de desemprego, inflação, índice PSI-20, ...*

- Y_t é estritamente estacionário se a função distribuição conjunta de $\{Y_{t_1}, ..., Y_{t_m}\}$ não muda com o tempo.
- Y_t é estacionário em covariância se para todo t e l, $E(Y_t) = \mu$ e $Cov(Y_t, Y_{t-l}) = \Gamma(l)$.
- estritamente estacionário$\Rightarrow$estacionário em covariância.
- O processo estocástico estacionário Y_t possui memória que se dissipa se $\Gamma(l) \to 0$ quando $l \to \infty$.
- Notação: Para Y_t univariado, $Cov(Y_t, Y_{t-l}) = \gamma_l$.
- Variância e Autocorrelação: $\gamma_0 = V(Y_t)\,; Corr(Y_t, Y_{t-l}) = \rho_l = \frac{\gamma_l}{\gamma_0}$.
- ρ_l é a função autocorrelação (FAC).
- Função de autocorrelação parcial (FACP): efeito que o passado do processo tem sobre o mesmo, *ceteris paribus,* i.e. coeficientes de correlação parcial através de uma autoregressão.
- Momentos Amostrais (equivalente aos momentos da população para uma amostra de dimensão T):

$$\overline{y} = \frac{1}{T}\sum_{t=1}^{T} y_t; \widehat{\gamma}_l = \frac{1}{T}\sum_{t=l+1}^{T}(y_t - \overline{y})(y_{t-l} - \overline{y}).$$

- Operador Desfasamento (Lag), L :

$$\begin{aligned} LY_t &= Y_{t-1}, LY_{t-1} = Y_{t-2}, ... \\ L^2Y_t &= L(Lx_t) = LY_{t-1} = Y_{t-2}; L^jY_t = Y_{t-j}. \end{aligned}$$

- Primeiras Diferenças:

$$\begin{aligned} (1-L)Y_t &= Y_t - Y_{t-1} = \triangle Y_t; \\ L\triangle Y_t &= \triangle LY_t = \triangle Y_{t-1} = Y_{t-1} - Y_{t-2}, ... \end{aligned}$$

3 Modelos ARMA

- Y_t é estacionário e sem componente sazonal.

3.1 White Noise (Ruido Branco)

- $Y_t \equiv \varepsilon_t \sim wn\left(0, \sigma_\varepsilon^2\right)$, o seu estudo justifica-se pela sua simplicidade e base de construção de modelos mais sofisticados.

- ε_t white noise gaussiano se $\varepsilon_t \sim i.i.d.N\left(0, \sigma_\varepsilon^2\right)$.

- Propriedades:

$E(\varepsilon_t) = 0, V(\varepsilon_t) = \gamma_0 = \sigma_\varepsilon^2 < \infty, Cov(\varepsilon_t, \varepsilon_{t-l}) = \gamma_l = 0, l \neq 0.$

- FAC: $\rho_0 = 1; \rho_l = 0, l \neq 0.$

- Teorema da Decomposição de Wold (para os modelos ARMA): Y_t estacionário em covariância$\Rightarrow$

$$Y_t = \mu + \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} = \mu + \varepsilon_t + \psi_1 \varepsilon_{t-1} + \psi_2 \varepsilon_{t-2} + ...,$$

em que $\sum \psi_j^2 < \infty, \psi_0 = 1.$

3.2 Modelo Autoregressivo (AR)

- Y_t é uma combinação linear do seu passado:

$AR(1) : Y_t = c + \phi Y_{t-1} + \varepsilon_t, \varepsilon_t \sim wn\left(0, \sigma_\varepsilon^2\right), t = 1, ..., T.$

- Parcimónio e capta a dependência/memória das séries. Para

$c = 0$,

$$\begin{aligned} Y_t &= \phi(\phi Y_{t-2} + \varepsilon_{t-1}) + \varepsilon_t = ... \\ &= \phi^j Y_{t-j} + \varepsilon_t + \phi\varepsilon_{t-1} + ... + \phi^{j-1}\varepsilon_{t-(j-1)} \\ &\equiv \phi^t Y_0 + \sum_{j=0}^{t-1} \phi^j \varepsilon_{t-j} = \phi^t Y_0 + \sum_{i=1}^{t} \phi^{t-i}\varepsilon_i. \end{aligned}$$

- Se $|\phi| < 1$ então " $\lim_{j\to\infty} \phi^j Y_{t-j} \to 0$" e, por isso,

$$Y_t = \sum_{j=0}^{\infty} \phi^j \varepsilon_{t-j} = \sum_{i=1}^{\infty} \phi^{t-i}\varepsilon_i.$$

- Portanto, para que o $AR(1)$ seja estacionário e de memória que se dissipa, $|\phi| < 1$. Neste caso,

$$E(Y_t) = 0; V(Y_t) = \gamma_0 = \frac{\sigma_\varepsilon^2}{1-\phi^2}; \gamma_l = \phi^l \gamma_0; \rho_l = \phi^l.$$

- Nota: Se $|\phi| < 1$, $\frac{1}{1-\phi L} = 1 + \phi L + \phi^2 L^2 + ...$
- FAC: Decaimento geometrico/exponencial $\rightarrow$ Memória curta, $|\rho_l| \leq C r^{|l|}$, com $0 < r < 1, C > 0$.
- Mais geral,

$$\begin{aligned} AR(p) \quad &: \quad Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + ... + \phi_p Y_{t-p} + \varepsilon_t \\ &\Leftrightarrow \phi(L) Y_t = c + \varepsilon_t. \end{aligned}$$

- Polinómio autoregressivo: $\phi(L) = 1 - \phi_1 L - \phi_2 L^2 - ... - \phi_p L^p$
- Estacionaridade se as raizes de $\phi(L)$ estão fora do circulo

unitário, i.e., a(s) solução(ões) de $1-\phi_1 z-\phi_2 z^2-...-\phi_p z^p=0$. Condição necessária para estacionaridade: $\phi_1+\phi_2+...+\phi_p<1$.

Example 2 $AR(2): \phi_1+\phi_2<1, \phi_2-\phi_1<1, |\phi_2|<1, \phi_1, \phi_2 \in \Re$.

- FAC: Premultiplicar as equações de Yule-Walker, $Y_t-\mu=\phi_1(Y_{t-1}-\mu)+...+\phi_p(Y_{t-p}-\mu)+\varepsilon_t$, por $Y_{t-j}-\mu$ e aplicar o valor esperado:

$$\begin{aligned}
\gamma_l &= \phi_1\gamma_{l-1}+\phi_2\gamma_{l-2}+...+\phi_p\gamma_{l-p}, \qquad && l>0\\
\gamma_l &= \phi_1\gamma_1+\phi_2\gamma_2+...+\phi_p\gamma_p+\sigma^2, \qquad && l=0 \text{ e}\\
\rho_l &= \phi_1\rho_{l-1}+\phi_2\rho_{l-2}+...+\phi_p\rho_{l-p},
\end{aligned}$$

- FACP: 1 para $l=0$; ϕ_j para $l=j=1,...,p$ e 0 para $l>p$.

3.3 Modelo Médias Móveis (MA)

- Y_t é uma função linear dos choques do presente e passado.
- Processo médias móveis de primeira ordem:

$$MA(1): Y_t=\mu+\varepsilon_t-\theta\varepsilon_{t-1}, \varepsilon_t \sim wn\left(0,\sigma_\varepsilon^2\right), t=1,...,T.$$

- Propriedades:

$$\begin{aligned}
E(Y_t) &= \mu; V(Y_t)=\gamma_0=\sigma_\varepsilon^2(1+\theta^2);\\
\gamma_1 &= -\theta\sigma_\varepsilon^2; \gamma_l=0, l>1\\
\rho_1 &= -\frac{\theta}{1+\theta^2}; \rho_l=0, l>1.
\end{aligned}$$

- Conclusão: $MA(1)$ é estacionário para qualquer valor real de μ, θ e tem memória curta (até ao lag 1). A FAC é zero para $l > 1$.
- Se $|\theta| < 1$, o $MA(1)$ é invertível e corresponderá a um $AR(\infty), Y_t = \varepsilon_t - \theta Y_{t-1} - \theta^2 Y_{t-2} + ...$ Portanto, a FACP tem um decaimento exponencial para zero.
- Invertibilidade: Raiz do polinómio médias móveis $\theta(L) = 1 - \theta L$ está fora do circulo unitário.
- Mais geral,

$$MA(q) \; : \; Y_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - ... - \theta_q \varepsilon_{t-q} \Leftrightarrow Y_t = \mu + \theta(L)\varepsilon_t.$$
$$\theta(L) \; = \; 1 - \theta_1 L - ... - \theta_q L^q.$$

- FAC zero para $l > q$; FACP tem um decaimento exponencial para zero se $\theta(L)$ é invertível.

3.4 Modelo ARMA

- Combina AR e MA.

$$ARMA(p, q) : \phi(L) Y_t = c + \theta(L)\varepsilon_t.$$

- $ARMA(1,1) \; : \; Y_t \; = \; \phi Y_{t-1} + \varepsilon_t - \theta \varepsilon_{t-1}$, tem FAC $\rho_1 = \frac{(1-\phi\theta)(\phi-\theta)}{1+\theta^2-2\phi\theta}, \rho_l = \phi \rho_{l-1}$.

Modelo	FAC	FACP
AR	decaimento exponencial ou de amortecimento sinusoidal	zero para $l > p$
MA	zero para $l > q$	decaimento exponencial ou de amortecimento sinusoidal
$ARMA$	decaimento exponencial ou de amortecimento sinusoidal	decaimento exponencial ou de amortecimento sinusoidal

4 Modelos ARIMA e SARIMA

4.1 ARIMA

- Y_t é não estacionário e sem componente sazonal.
- Definição: $Y_t \sim I(0)$ se o processo Y_t é estacionário. O processo Y_t é $I(d)$ se são necessárias d primeiras diferenças para o tornar estacionário: $Y_t \sim I(d) : \triangle^d Y_t \sim I(0)$. O processo $I(d)$ diz-se que tem d raízes unitárias (no seu polinómio autoregressivo).
- Muitas variáveis económicas e financeiras são não estacionárias. Apresentam uma tendência. Estocástica ou Deterministica? Aqui, é estocástica. Normalmente, $d = 1$; alguns exemplos de $d = 2$.

Example 3 $AR(1), Y_t = \alpha + \rho Y_{t-1} + \varepsilon_t$ *é estacionário se* $|\rho| < 1 : \triangle Y_t = (\rho - 1)\left(Y_{t-1} - \frac{\alpha}{1-\rho}\right) + \varepsilon_t$. *Se* $\rho = 1, \triangle Y_t = \varepsilon_t$.

Example 4 *Passeio Aleatório sem deriva:* $Y_t = Y_{t-1} + \varepsilon_t \Leftrightarrow Y_t = Y_0 + \sum_{j=1}^{t} \varepsilon_j$.

Example 5 *Passeio Aleatório com deriva:* $Y_t = \alpha + Y_{t-1} + \varepsilon_t \Leftrightarrow \triangle Y_t = \alpha + \varepsilon_t \Leftrightarrow Y_t = Y_0 + \alpha t + \sum_{j=1}^{t} \varepsilon_j$.

- FAC de Y_t não decai para zero pois o efeito de um choque é permanente. Na verdade, a FAC teórica não existe. MAS a FAC de $\triangle^d Y_t$ decai para zero. Memória infinita.
- Conclusão: o processo Y_t é um $ARIMA\,(p, d, q)$ se, não sendo estacionário em niveis, torna-se estacionário após d diferenças. Por exemplo, o processo $ARIMA\,(p, 1, q)$ é estacionário nas primeiras diferenças pois remove a tendência. estocástica.
- Concretamente,
$$ARIMA(p, d, q) : \phi(L)\triangle^d Y_t = c + \theta(L)\varepsilon_t.$$
- Identificação de p, q : FAC e FACP do processo $\triangle^d Y_t$.
- Identificação de d : Testes de raizes unitárias (ver apontamentos sobre raizes unitárias) ...
- Condições de Estacionaridade e Invertabilidade de $\triangle^d Y_t$: Raizes dos polinómios $\phi(L)$ e $\theta(L)$.

4.2 SARIMA

- Y_t tem componente sazonal e é/não é estacionário **(**$d = 0$ versus $d > 0)$: $SARIMA(p, d, q) \times (P, D, Q)_S$:
$$\phi(L)\Phi(L^S)\left(1 - L^S\right)^D (1 - L)^d Y_t = c + \theta(L)\Theta(L^S)\varepsilon_t,$$
$$\begin{aligned} \Phi(L^S) \;&=\; 1 - \Phi_S L^S - \Phi_{2S} L^{2S} - ... - \Phi_{PS} L^{PS} \\ \Theta(L^S) \;&=\; 1 - \Theta_S L^S - \Theta_{2S} L^{2S} - ... - \Theta_{PS} L^{PS}. \end{aligned}$$

Example 6 $SARIMA(0,1,1) \times (0,1,1)_3$

- S define a frequência dos dados: Dados mensais para $S = 12, \ldots$
- Identificação de d, D : Testes de raizes unitárias simples e sazonais ...
- Ao processo $\left(1 - L^S\right)^D (1 - L)^d Y_t$:
  - Identificação de P, Q : FAC e FACP nos lags sazonais, $S, 2S, 3S, \ldots$
  - Identificação de p, q : FAC e FACP nos lags não sazonais.

5 Metodologia de Box-Jenkins

- Na prática, temos um método: Transformar os dados (log estabiliza a variância; diferenças estacionariza a série;...), identificação do processo (d,p,q,...), estimação dos parâmetros, inferência/testes, previsão in-sample e out-of-sample.

- PASSO 1: Transformar os dados
  - Estabilizar a variância (transformação de Box-Cox) e estacionar a série (d, D) .

- PASSO 2: Identificação do processo
  - Identificação de p, q, P, Q pela comparação entre as FAC e FACP Empiricas e as FAC e FACP Teóricas.

  - FAC Empirica (correlograma):

$$\begin{aligned}
\widehat{\mu} &= \frac{1}{T}\sum_{t=1}^{T} y_t; \widehat{\gamma}_0 = \frac{1}{T}\sum_{t=1}^{T}(y_t - \widehat{\mu})^2 \\
\widehat{\gamma}_l &= \frac{1}{T}\sum_{t=l+1}^{T}(y_t - \widehat{\mu})(y_{t-l} - \widehat{\mu}); \widehat{\rho}_l = \frac{\widehat{\gamma}_l}{\widehat{\gamma}_0}.
\end{aligned}$$

  - Bandas da FAC Empirica: $\pm\frac{2}{\sqrt{T}}$ porque $\sqrt{T}\widehat{\rho} \xrightarrow[H_0]{d} N(0,1)$.

  - FACP Empirica: Estimação AR!

  - Estratégia: Especifico para Geral (Down-Top).

- PASSO 3: Estimação
  - OLS, MLE, ...

- PASSO 4: Inferência/Testes de Diagonóstico
  - Testes sobre os parâmetros do modelo, ...
  - Residuos Ruido Branco? Testes de Autocorrelação aos residuos do modelo (DW, BP, LB, ...), ...
  - Critérios de Selecção de modelos (AIC, BIC, ...), ...
- PASSO 5: Previsão
  - In-sample e Out-of-sample ...
  - Representação $AR \rightarrow$ Valor esperado condicional: $\widehat{y}_{t+k} = E(y_{t+k}|y_t, y_{t-1}, ...)$. Eg: $AR(p)$:

$$\begin{aligned} \widehat{y}_{t+k} &= c + \phi_1 E(y_{t+k-1}|y_t, y_{t-1}, ...) \\ &\quad +\phi_2 E(y_{t+k-2}|y_t, y_{t-1}, ...) + ... + \phi_p E(y_{t+k-p}|y_t, y_{t-1}, ...). \end{aligned}$$

  Note-se que, por exemplo, $E(y_t|y_t, y_{t-1}, ...) = y_t$ mas que $E(y_{t+1}|y_t, y_{t-1}, ...) \equiv \widehat{y}_{t+1} \neq y_{t+1}$.

  - Erro de previsão In-sample: $e_t = y_t - \widehat{y}_t$.
  - Performance do modelo: Minimizar medidas

$$RMSE, MAE, MAPE, U, ...$$