SERIE STORICHE:
MODELLI NON LINEARI

Giovanni Fonseca
Dipartimento di Economia
Università degli Studi dell'Insubria - Varese

Politecnico di Torino
20 novembre 2002

MODELLI NON LINEARI

- Siano

$\{\varepsilon_t\}_{t\in\mathbb{Z}}$ successione di v.c. IID,
assolutamente continue rispetto
alla misura di Lebesgue su $\mathbb{R}$,

$$\underline{X}_t = (X_t, \ldots, X_{t-p+1})',$$
$$\underline{\varepsilon}_t = (\varepsilon_t, \ldots, \varepsilon_{t-q+1})',$$
$$Y_t = (\underline{X}'_t, \underline{\varepsilon}'_t)'.$$

$$\downarrow$$

$$X_t = f(Y_{t-1}, \varepsilon_t)$$

- Modellizzazione della media condizionata:

$$X_t = f(Y_{t-1}) + \varepsilon_t$$

  e.g. ARMA non lineari (NARMA), modelli a soglia (SETAR), modelli di tipo bilineare non markoviani;

- Modellizzazione della varianza condizionata:

$$X_t = f(Y_{t-1})\varepsilon_t$$

  e.g. modelli eteroschedastici (ARCH) e loro generalizzazioni, modelli di tipo bilineare markoviano;

- Modellizzazione congiunta di media e varianza condizionata:

$$X_t = f_1(Y_{t-1}) + f_2(Y_{t-1})\varepsilon_t.$$

RASSEGNA DI MODELLI

- Modelli a soglia: TAR, SETAR, TARMA, STAR, TARSO, TARSC

- Modelli bilineari: BL, TBL

- Modelli eteroschedastici: ARCH, GARCH, TARCH, M-ARCH...

- Modelli con coefficienti aleatori: RCA, DSM

- Modelli deterministici caotici

- Combinazioni di modelli

- FAR^1= Fractional AR
  FAR^2= Functional AR

STABILITA' DEL MODELLO
(stazionarietà)

Nel caso di modelli ARMA per dimostrare la stazionarietà si utilizzano tecniche che sfruttano profondamente l'assunzione di linearità $\longrightarrow$ NON appropriate per modelli non lineari.

$\{Y_t = (\underline{X}_t', \underline{\varepsilon}_t')'\}$ è una catena di Markov

$$\downarrow$$

Utilizzo di tecniche sviluppate per l'analisi della stabilità di una catena di Markov.

E' necessario quindi dare una rappresentazione markoviana al processo generato dal modello.
$AR(1)$ è una catena di Markov
$MA(1)$ non è una catena di Markov, non possiede una rappresentazione markoviana univariata, ma bivariata (X_t, ε_t)

CATENE DI MARKOV (MC)

• $\{Z_t\}_{t \in \mathbb{Z}}$ è una catena di Markov definita su uno spazio $(\mathsf{Z}, \mathcal{B})$ se per $z_i \in \mathsf{Z}$, $A \in \mathcal{B}$

$$\Pr[Z_t \in A | Z_{t-1} = z_1, Z_{t-2} = z_2, \ldots] =$$

$$= \Pr[Z_t \in A | Z_{t-1} = z_1].$$

Proprietà rilevanti:

• Omogeneità rispetto al tempo
• Aperiodicità

• φ-irriducibilità
Sia $\varphi(\cdot)$ una misura σ-finita su $(\mathsf{Z}, \mathcal{B})$.

$$\forall A \in \mathcal{B} \ : \ \varphi(A) > 0 \Rightarrow \sum_{n=1}^{\infty} P^n(z, A) > 0 \quad \forall z \in \mathsf{Z}.$$

• T-continuità

STAZIONARIETA' DI UNA MC

• π è una misura invariante finita per una MC $\{Z_t\}$ se

$$\pi(A) = \int_{\mathsf{Z}} P(z, A)\pi(dz).$$

$$\Downarrow$$

$Z_0 \sim \pi \;\Rightarrow\; Z_t \sim \pi \; \forall t$ e quindi $\{Z_t\}$ è un processo stazionario.

• Ergodicità di una MC: la probabilità di transizione a n passi $P^n(z, \cdot\,)$ converge verso la misura invariante (unica) π

$$\lim_{n \to \infty} \|P^n(z, \cdot\,) - \pi(\,\cdot\,)\| = 0 \quad \forall z.$$

• Ergodicità geometrica: convergenza verso π con progressione geometrica

$$\lim_{n \to \infty} \rho^{-n} \|P^n(z, \cdot\,) - \pi(\,\cdot\,)\| = 0 \quad \rho \in (0, 1).$$

• Per una MC: ergodicità $\Rightarrow$ stazionarietà

DRIFT-CONDITIONS
(condizioni di deriva)

Sono criteri che permettono di determinare se una catena di Markov è “stabile”
(Meyn e Tweedie, 1993)

Sia $\{Z_t\}$ una MC omogenea, aperiodica, φ-irriducibile, T-continua e
sia $V : \mathsf{Z} \to I\!R$ una funzione misurabile, non-negativa, *norm-like*.

• $\{Z_t\}$ è *ergodica* se esistono un insieme *compatto* C e una costante $b < \infty$ tali che

$$\mathsf{E}[V(Z_t)|Z_{t-1} = z] \leq V(z) - 1 + b\mathbb{1}_C(z).$$

• $\{Z_t\}$ è *geometricamente ergodica* se esistono un insieme *compatto* C e due costanti $b < \infty$ e $\delta \in (0, 1)$ tali che

$$\mathsf{E}[V(Z_t)|Z_{t-1} = z] \leq \delta V(z) + b\mathbb{1}_C(z).$$

CRITERIO ALTERNATIVO
(Liu e Susko, 1992)

Sia $\{Z_t\}$ una MC omogenea tale che

$$\lim_{A \downarrow \emptyset} \sup_{x \in K} P(x, A) = 0$$

per ogni insieme K compatto (countable additivity condition).

Esiste una misura finita invariante π per $\{Z_t\}$.

$$\Updownarrow$$

Esistono $V : \mathsf{Z} \to I\!R$ funzione misurabile, non-negativa, *norm-like*
e P_0 misura di probabilità iniziale per Z_0 tali che

$$\sup_t \mathsf{E}_{P_0}\{\mathsf{E}[V(Z_t)|Z_0 = z]\} < \infty$$

Applicazioni ai modelli per serie storiche

Applicazioni dei risultati sulla stabilità di MC ai modelli non-lineari per serie storiche si possono trovare in

- Meyn e Tweedie (1993)
- Fonseca e Tweedie (2002)
- Tong, (1990)
- Bhattacharya e Lee (1995)
- Chan e Tong (1985)
- Cline e Pu (1999)
- Tjostheim (1990)

In generale si ha che per $Y_t = f_1(Y_{t-1}) + f_2(Y_{t-1})\varepsilon_t$ se

- $\{\varepsilon_t\}$ sono IID, assolutamente continui con densità positiva su $I\!R$ e valore atteso finito
- f_1 limitata sui compatti
- $0 < f_2(y) < c < \infty \ \forall y \in I\!R$
- $\lim_{y \to \pm\infty} \frac{f_1(y)}{y} < 1$

$$\lim_{y \to -\infty} \frac{f_1(y)}{y} \times \lim_{y \to +\infty} \frac{f_1(y)}{y} < 1$$

$\Rightarrow$ Y_t geometricamente ergodica.
(In Ferrante et al. (2002) ipotesi diverse per f_1, f_2)

MODELLI A SOGLIA

Ideati per studiare fenomeni con ciclicità

- SETARMA(m;p,q)

$$X_t = \sum_{j=1}^{m} \left\{ a_0^{(j)} + \sum_{i=1}^{p} a_i^{(j)} X_{t-i} + \sum_{i=1}^{q} b_i^{(j)} \varepsilon_{t-i} \right\} \mathbb{1}_{R_j}(\underline{X}_{t-1}) + \varepsilon_t$$

con $\bigcup_{j=1}^{m} R_j = \mathbb{R}^p$

- Rappresentazione markoviana SETAR(m;p) ($b_i^{(j)} = 0 \ \forall i, j$)

$$\underline{X}_t = \sum_{j=1}^{m} \left\{ A^{(j)} \underline{X}_{t-1} + B^{(j)} \underline{\varepsilon}_{t-i} \right\} \mathbb{1}_{R_j}(\underline{X}_{t-1}) + (\varepsilon_t, 0, \ldots, 0)'$$

$$A^{(j)} = \begin{bmatrix} a_1^{(j)} & a_2^{(j)} & \cdots & a_p^{(j)} \\ 1 & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ 0 & \cdots & 1 & 0 \end{bmatrix}$$

SETAR(m;p)

• Stazionarietà

$$\max_{j=1,\ldots,m} \sum_{i=1}^{p} |a_i^{(j)}| < 1$$

• Stima e identificazione
Tong (1980): *grid search* + *AIC(k)*

$\underline{X}_{t-1} \in R_j \longrightarrow X_{t-d} \in (r_{j-1}, r_j]$

GRID SEARCH: si stimano (max ver cond) modelli per diversi valori di $d = 1, \ldots, p$ e valori soglia $r_j \in I\!R$

AIC(k): per ogni modello stimato si calcola AIC

$\Rightarrow$ Si sceglie il modello con AIC minore.

CASI PARTICOLARI

SETAR(2;1) $X_t = a_1 X_{t-1}^{+} + a_2 X_{t-1}^{-} + \varepsilon_t$
(Petruccelli e Woolford (1984))

• Ergodicità (cond. necessaria e sufficiente)

$$a_1 < 1 \quad a_2 < 1 \quad a_1 a_2 < 1$$

• Stima: OLS con separazione dei dati osservati $\Rightarrow$ stime consistenti.

SETAR(m;1) (Chan et al. (1985))

$$X_t = \sum_{j=1}^{m} \left\{ a_0^{(j)} + a_1^{(j)} X_{t-i} \right\} 1\!\!1_{(r_{j-1}, r_j]}(X_{t-1}) + e_n$$

• Ergodicità se vale una delle seguenti:

1) $a_1^{(1)} < 1, \quad a_1^{(m)} < 1, \quad a_1^{(1)} a_1^{(m)} < 1$

2) $a_1^{(1)} = 1, \quad a_1^{(m)} < 1, \quad a_0^{(1)} > 0$

3) $a_1^{(1)} < 1, \quad a_1^{(m)} = 1, \quad a_0^{(m)} < 0$

4) $a_1^{(1)} = 1, \quad a_1^{(m)} = 1, \quad a_0^{(m)} < 0 < a_0^{(1)}$

5) $a_1^{(1)} a_1^{(m)} = 1 \quad a_1^{(1)} < 0 \quad a_0^{(m)} + a_1^{(m)} a_0^{(1)} > 0$

N.B: dipendono solo dai parametri dei regimi $1, m$

• Stima: OLS con divisione delle osservazioni fra i vari regimi. Consistenza delle stime se le soglie r_i sono *note*

CASI PARTICOLARI (2)

STAR Smooth Transition AR
Ideati per stimare il parametro di soglia r

$$X_t = a_0^{(1)} + \sum_{i=1}^{p} a_i^{(1)} X_{t-i} +$$
$$+ \left(a_0^{(2)} + \sum_{i=1}^{p} a_i^{(2)} X_{t-i} \right) F\left(\frac{X_{t-d} - r}{z} \right) + \varepsilon_t$$

con $F(\,\cdot\,) \in [0,1]$ (e.g. funzione logistica, funzione di ripartizione gaussiana) e
$z \geq 0$ parametro di smoothing
$z \to 0 \Rightarrow STAR \to SETAR$

• Ergodicità $p = d = 1$

$$a_1^{(1)} < 1,\; (a_1^{(1)} + a_1^{(2)}) < 1,\; a_1^{(1)}(a_1^{(1)} + a_1^{(2)}) < 1$$

• Stima: *grid search* su vari valori di r, z con stima LS e scelta del modello definitivo tramite AIC.

MODELLI BILINEARI BL(p,q,P,Q)

Subba Rao (1981), Subba Rao e Gabr (1984). Fenomeni con esplosioni improvvise (e *casuali*)

$$X_t = \sum_{i=1}^{p} a_i X_{t-i} + \sum_{j=1}^{q} b_j \varepsilon_{t-j} + \\ + \sum_{i=1}^{P} \sum_{j=1}^{Q} c_{i,j} X_{t-i} \varepsilon_{t-j} + \varepsilon_t$$

- Struttura delle autocorrelazioni di ARMA(p,1)

- $X_t = \beta X_{t-k} \varepsilon_{t-k+m} + \varepsilon_t$
  $k \geq 2, \quad 1 \leq m \leq k-1$

$$\Rightarrow \ \gamma_s = \begin{cases} \frac{1}{1-\beta^2} & s = 0 \\ 0 & s \neq 0 \end{cases}$$

$\Rightarrow$ WN

- Stazionarietà debole (mediante calcolo dei momenti del processo) di BL(p,0,p,1)

$$\rho(A \otimes A + B \otimes B\sigma^2) < 1$$

con $\rho(A) = \|A\|_2$ (massimo autovalore di A)

IDENTIFICAZIONE E STIMA

Si *assume* l'invertibilità per stimare il modello.

• *Grid search* e AIC per determinare gli ordini del modello p, q, P, Q.

• Stima con max ver cond e procedura di Newton-Raphson per massimizzazione numerica.

BL(1,0,1,1)

$X_t = aX_{t-1} + bX_{t-1}\varepsilon_{t-1} + \varepsilon_t$

• Stazionarietà debole: $a^2 + b^2\sigma^2 < 1$

• Rappresentazione markoviana
$X_t = (a + b\varepsilon_{t-1})X_{t-1} + \varepsilon_t = Z_{t-1} + \varepsilon_t$
$Z_t = (a + b\varepsilon_t)X_t = (a + b\varepsilon_t)(Z_{t-1} + \varepsilon_t)$
$\Rightarrow Z_t$ è MC

• Ergodicità di Z_t: $|a| + |b|\mathrm{E}[|\varepsilon_1|] < 1$
$\Rightarrow$ stazionarietà forte di X_t

MODELLI BILINEARI A SOGLIA

$$X_t = aX_{t-1} + \left[b_1 \mathbb{1}_{(-\infty,c)}(X_{t-1}) + b_2 \mathbb{1}_{[c,\infty)}(X_{t-1})\right] X_{t-1}\varepsilon_{t-1} + \varepsilon_t$$

- Rappresentazione markoviana

$$Z_t = \left(a + \left[b_1 \mathbb{1}_{(-\infty,c)}(X_t) + b_2 \mathbb{1}_{[c,\infty)}(X_t)\right] \varepsilon_t\right) X_t$$
$$X_t = Z_{t-1} + \varepsilon_t$$
$$Z_t = \left(a + \left[b_1 \mathbb{1}_{(-\infty,c)}(Z_{t-1} + \varepsilon_t) + b_2 \mathbb{1}_{[c,\infty)}(Z_{t-1} + \varepsilon_t)\right] \varepsilon_t\right) \times$$
$$\times \left(Z_{t-1} + \varepsilon_t\right)$$

- Stazionarietà : $|a| + \max\{|b_1|, |b_2|\}\mathrm{E}[|\varepsilon_1|] < 1$

STIMA E IDENTIFICAZIONE

Si assume l'invertibilità e $\{\varepsilon_t\}$ gaussiani

• Si determina un insieme di plausibili valori soglia $\{c_1, \ldots, c_r\}$ (e.g. i quantili della funzione di ripartizione empirica della serie osservata) e per ogni valore $c = c_i$ si stima il modello (*grid search*)

• Data la serie osservata $\{y_1, \ldots, y_T\}$, per il calcolo della verosimiglianza, si fissano i valori iniziali della serie $Y_1 = y_1$ e dei residui $e_1 = 0$. I residui successivi si determinano dall'equazione ricorsiva del modello

$$e_t = y_t - a y_{t-1} - b(y_{t-1}, c) e_{t-1}$$

• Per ottenere lo stimatore di max ver del parametro $\theta_c = (a, b_1, b_2)'$ è sufficiente minimizzare

$$Q(\theta_c) = \sum_{t=2}^{T} e_t^2$$

IDENTIFICAZIONE E STIMA (2)

• Indichiamo lo *score* con

$$g(\theta_c) = \left(\frac{dQ(\theta_c)}{da}, \frac{dQ(\theta_c)}{db_1}, \frac{dQ(\theta_c)}{db_2}\right)$$

e la matrice delle derivate seconde con

$$H(\theta_c) = \left[\frac{d^2Q(\theta_c)}{d\theta_{c,i}d\theta_{c,j}}\right]_{i=1,2,3 \;\; j=1,2,3}$$

dove $\theta_{c,1} = a,\ \theta_{c,2} = b_1,\ \theta_{c,3} = b_2$ e $d\theta_{c,i}d\theta_{c,i} = d\theta^2_{c,i}$

• Per minimizzare $Q(\theta_c)$ si utilizza la procedura ricorsiva di Newton-Raphson. Si espande in serie di Taylor $g(\widehat{\theta}_c)$ in un intorno di $\widehat{\theta}_c = \theta_c$

$$g(\widehat{\theta}_c) = 0 = g(\theta_c) + H(\theta_c)(\widehat{\theta}_c - \theta_c)$$

da cui $\theta_c^{(k+1)} = \theta_c^{(k)} - H^{-1}(\theta_c^{(k)})g(\theta_c^{(k)})$

• La procedura converge verso un valore $\widehat{\theta}_c$, minimo locale (e sperabilmente globale) di $Q(\theta_c)$

• Le derivate di $Q(\theta_c)$ dipendono dai parametri e quindi servono dei valori di partenza per a, b_1, b_2. Sono adatti valori nulli o stima di a da un modello AR(1).

IDENTIFICAZIONE E STIMA (3)

- Si calcola la varianza residua

$$\widehat{\sigma}_c^2 = \frac{1}{T-1} Q(\widehat{\theta}_c)$$

e si determina

$$AIC_c(4) = (T-1)\ln \widehat{\sigma}_c^2 + 2 \times 4$$

- I parametri stimati finali saranno quelli corrispondenti al valore AIC_c più basso al variare di c in $\{c_1, \ldots, c_r\}$