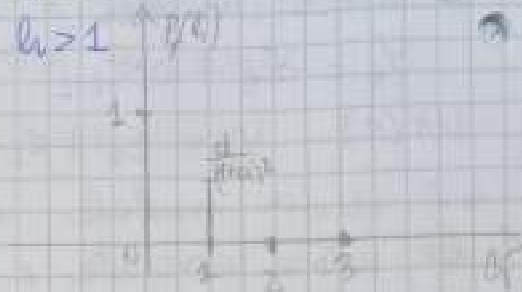
12/11/15

$\Rightarrow COV(Y_t, Y_{t-h}) \stackrel{\Delta}{=} \gamma(h) = \begin{cases} (1+a^2)\sigma^2_\varepsilon & \text{se } h=0 \\ a\,\sigma^2_\varepsilon & \text{se } h=1 \\ 0 & \text{se } h>1 \end{cases}$

$\Rightarrow \rho_Y(h) = \begin{cases} 1 & \text{se } h=0 \\ \frac{a}{(1+a^2)} & \text{se } h=1 \\ 0 & \text{se } h>1 \end{cases}$



OSS La correlazione è diversa da 0 fino al ritardo 1 in MA(1)

$\Rightarrow$ in MA(2) sarà $\neq 0$ fino al ritardo 2.

$\Rightarrow$ ... IN MA(q) CORR $\neq 0$ fino al ritardo q.

$\Rightarrow$ (CRITICA DI YULE, MA(·) VA A LISCIARE IL PROCESSO. ?)

• $Y_t \sim MA(2)$: $Y_t = \varepsilon_t + a_1 \varepsilon_{t-1} + a_2 \varepsilon_{t-2}$ $\varepsilon_t \sim WN(0, \sigma^2_\varepsilon)$

$\Rightarrow$ si può verificare che

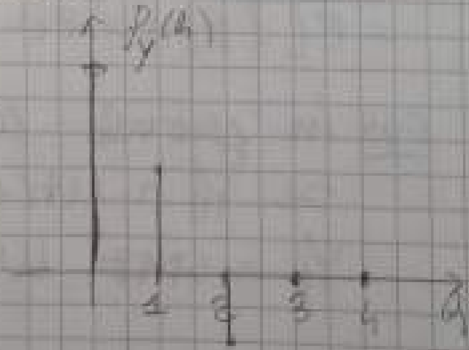
$\forall a_1, a_2$ FINITI $\Rightarrow Y_t$ STAB. IN COV.

DIM $E[Y_t] = 0$

$VAR[Y_t] = (1 + a_1^2 + a_2^2)\sigma^2_\varepsilon$

$\Rightarrow \gamma(h) = \begin{cases} (1 + a_1^2 + a_2^2)\sigma^2_\varepsilon & \text{se } h=0 \\ (a_1 + a_1 \cdot a_2)\sigma^2_\varepsilon & \text{se } h=1 \\ a_2 \sigma^2_\varepsilon & \text{se } h=2 \\ 0 & \text{se } h>2 \end{cases}$

OSS Si dice che il PROC. MA(2) ha memoria più lunga di MA(1)

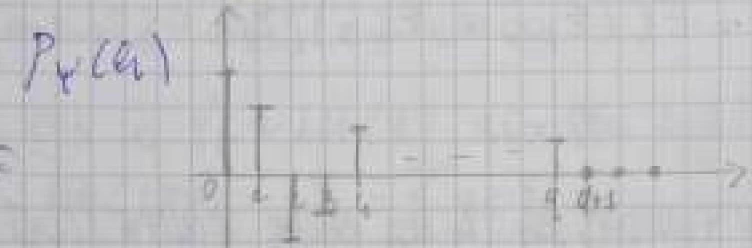


$Y_t \sim MA(q)$: $Y_t = \varepsilon_t + \sum_{s=1}^{q} a_s \varepsilon_{t-s}$ $\quad \varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$

$\forall a_1, \dots, a_q$ $\quad Y_t$ È STAZIONARIO IN COVARIANZA.

$E[Y_t] = 0$

$$\gamma(h) = \begin{cases} (1 + a_1^2 + a_2^2 + \dots + a_q^2)\,\sigma_\varepsilon^2 & \text{se } h = 0 \\ \neq 0 & \text{se } 1 \le h \le q \\ 0 & \text{se } h > q \end{cases}$$



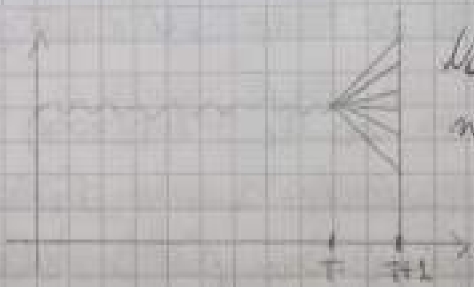
OSS q è il cut-off del processo: punto da cui si fa il salto a 0.

$\Rightarrow$ MEMORIA DI q.

OSS $Y_t \sim$ I.I.D.

$E[Y_t] = \mu$

$VAR[Y_t] = \sigma_\varepsilon^2$ $\quad \forall t$



Non c'è correlazione nel tempo.

Mi interessa Y_{T+1}, LA MIGLIOR PREVISIONE È $E(Y_{t+1}) \overset{IID}{=} \mu =$
$= E[E[Y_{t+1} | \{Y_t, \dots\}]]$

Ma se Y_t fossero correlati nel tempo:
$E[Y_{t+1}] = \mu \neq E[Y_{t+1} | \{Y_t, \dots, Y_1\}]$ HO DUE PREVISIONI DIVERSE.

- Per prevedere μ uso $\Rightarrow \hat{\mu} = \frac{1}{T} \sum_{t=1}^{T} y_t$ = MEDIA CAMPIONARIA che avrà $VAR(\hat{\mu})$.

OSS In generale $VAR[\widehat{E[Y_{t+1} | \{Y_t, \dots, Y_1\}]}] \le VAR(\hat{\mu})$
per cui cercheremo una stima di quello condizionato

$\Rightarrow$

Quindi consideriamo il modello MA:

- Per MA(1) sarà meglio $E[Y_{t+1} | \underline{Y}_t]$ di $E[Y_{t+1}]$ (se prendo $E[Y_{t+2} | \underline{Y}_t] = E[Y_{t+2}]$)
- Per MA(2) sarà correlato fino a ritardo 2, e così via
- Supponiamo di avere MA($+\infty$) (in teoria si può fare, in pratica no) $Y_t = \varepsilon_t + \sum_{j=1}^{+\infty} a_j \varepsilon_{t-j} \; (= \sum_{j=0}^{+\infty} a_j \varepsilon_{t-j} \text{ con } a_0 = 1)$, $\varepsilon_t \sim WN(0, \sigma^2_\varepsilon)$

$E(Y_t) = 0$

$\gamma(h) = E[Y_t \cdot Y_{t-h}] = E[(\sum_{j=0}^{+\infty} a_j \varepsilon_{t-j})(\sum_{j=0}^{+\infty} a_j \varepsilon_{t-j-h})] =$

$= \sigma^2_\varepsilon \sum_{j=0}^{+\infty} a_j a_{j+h}$

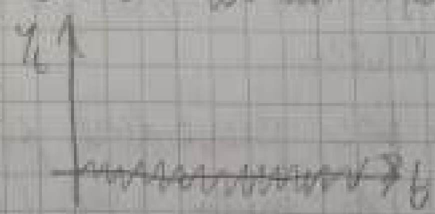
$\Rightarrow \gamma(0) = \sigma^2_\varepsilon \sum_{j=0}^{+\infty} a_j^2$

(OSS NON È DETTO CHE SIA FINITA, MA A NOI SERVE FINITA COSÌ CON TUTTE FINITE $\Rightarrow$ STAZ. IN COVARIANZA.)

- Se $\sum_{j=0}^{+\infty} a_j^2 < +\infty \Rightarrow Y_t \sim MA(+\infty)$ è STAZ. IN COV.

Condizione di quadrato-sommabilità $\Rightarrow$ ci aspettiamo che valori più vicini sono più correlati tra loro che quelli distanti. Se è così dovrebbe risultare soddisfatta la condizione, perché più cresce l'indice più diminuisce la quantità fino a 0 (ORIGINALE, MA COMUNQUE)
(es il PIL di oggi non è correlato con quello di 100 anni fa)

- Fin ora abbiamo sempre visto $E[Y_t] = 0 \Rightarrow$



- DOBBIAMO GENERALIZZARE IL MODELLO

▲ $MA(+\infty)$: $Y_t = \mu + \sum_{s=0}^{+\infty} a_s \varepsilon_{t-s}$ (μ è componente deterministica)
$\Rightarrow E[Y_t] = \mu \quad \forall t$.

■ Th di WOLD (dà una rappresentazione per i proc. stoc. in cui) TUTTI I PROCESSI STAZIONARI IN COVARIANZA SENZA COMPONENTI DETERMINISTICHE e $\forall t\ E(Y_t) = 0$ SI POSSONO RAPPRESENTARE CON UN $MA(+\infty)$ t.c.

$$Y_t = \sum_{s=0}^{+\infty} a_s \varepsilon_{t-s} \quad \text{con } a_0 = 1, \quad \varepsilon_t \sim WN(0, \sigma^2_\varepsilon)$$
$$\text{e } \sum_{s=0}^{+\infty} a_s^2 < +\infty$$

OSS Se $Y_t \sim MA(1)$ risulta uguale a $MA(+\infty)$ $Y_t = \sum_{s=0}^{+\infty} a_s \varepsilon_{t-s}$ con $a_0 = 1$ con $a_j = 0\ \forall j > 1$

● PROCESSI AR (AUTOREGRESSIVI) (E ARMA)

OSS $Y_t = \alpha + \beta X_t + u_t$ = MODELLO DI REG. LIN. STATICO (X_t è dello stesso periodo)

$\Rightarrow Y_t = \alpha + \beta Y_{t-1} + u_t$ = AUTOREGRESSIONE DI ORDINE 1 $\Rightarrow$ regredisco sui valori passati della variabile dipendente = MODELLO DI REG. LIN. DINAMICO (poiché c'è legame nel tempo).

▲ $Y_t \sim AR(p)$ = MODELLO AUTOREGRESSIVO DI ORDINE p.
$Y_t = \alpha + \theta_1 Y_{t-1} + \theta_2 Y_{t-2} + \dots + \theta_p Y_{t-p} + \varepsilon_t$ (eq. a diff. finite) con $\varepsilon_t \sim WN(0, \sigma^2_\varepsilon)$ (massimo ritardo p: regressione con p valori del passato).

OSS Se $\theta_1 = 0 \Rightarrow$ cmq è di ordine p. Conta il max ritardo.

- $\underset{\text{INPUT}}{\varepsilon_t} \longrightarrow AR(p) \longrightarrow \underset{\text{OUTPUT}}{(Y_t)_t}$

▲ $Y_t \sim AR(1)$ MODELLO AUTOREGRESSIVO DI ORDINE 1
$$Y_t = \underset{\text{costante}}{\theta} Y_{t-1} + \varepsilon_t, \quad \text{con } \varepsilon_t \sim WN(0, \sigma^2_\varepsilon)$$
(CASO GENERALE è θ_t = DIPENDENTE DAL TEMPO)

è equazione ALLE DIFF. FINITE STOCASTICA LIN. e A COEF. COSTANTI.

Poiché $\forall t$, $\Rightarrow Y_{t-1} = \theta Y_{t-2} + \varepsilon_{t-2} \Rightarrow$

$\Rightarrow Y_t = \theta(\theta Y_{t-2} + \varepsilon_{t-1}) + \varepsilon_t = \theta^2 Y_{t-2} + \theta \varepsilon_{t-1} + \varepsilon_t =$

(ma $Y_{t-2} = \theta Y_{t-3} + \varepsilon_{t-2}$) $= \theta^3 Y_{t-3} + \theta^2 \varepsilon_{t-2} + \theta \varepsilon_{t-1} + \varepsilon_t =$

$= \dots$ itero fino a $t-n$) $\Rightarrow$

$Y_t = \theta^n Y_{t-n} + \sum_{j=0}^{n-1} \theta^j \varepsilon_{t-j}$ SI PUÒ SCRIVERE Y_t COME C.L. DEL VALORE DEL PERIODO PRECEDENTE E SOMMATORIA DEL PROCESSO WN.

▲ Supponiamo $|\theta|<1$ $(-1<\theta<1)$ CONDIZIONE DI STAZ. IN COV. ASINTOTICA $\Rightarrow$ per $n \to +\infty$ $\theta^n \to 0$

cioè se $|\theta|<1 \Rightarrow \lim_{n\to+\infty} \theta^n = 0$

$Y_t \xrightarrow[n\to+\infty]{} \sum_{j=0}^{+\infty} \theta^j \varepsilon_{t-j} \sim MA(\infty)$

$Y_t = \sum_{j=0}^{+\infty} \theta^j \varepsilon_{t-j}$ è staz. in cov $\Leftrightarrow \sum_{j=0}^{+\infty} (\theta^j)^2 < +\infty$

ma $\sum_{j=0} (\theta^2)^j = 1 + \theta^2 + (\theta^2)^2 + (\theta^2)^3 + \dots$ SERIE GEOMETRICA DI RAGIONE θ^2

è CONVERGENTE SE $|\theta^2|<1 \Rightarrow |\theta|<1$ che vale per (HP) mem.

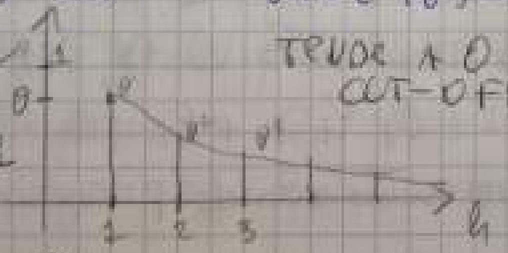
Inoltre la serie converge a $\frac{1}{1-\theta^2}$

- ASINTOTICAMENTE $\Rightarrow$ ($\Leftrightarrow |\theta|<1$)

$E(Y_t) \to 0$ $\quad VAR(Y_t) \to \sigma_\varepsilon^2 \frac{1}{1-\theta^2}$ $\quad \gamma(h) \to \sigma_\varepsilon^2 \frac{\theta^h}{1-\theta^2}$

$\rho(h)$

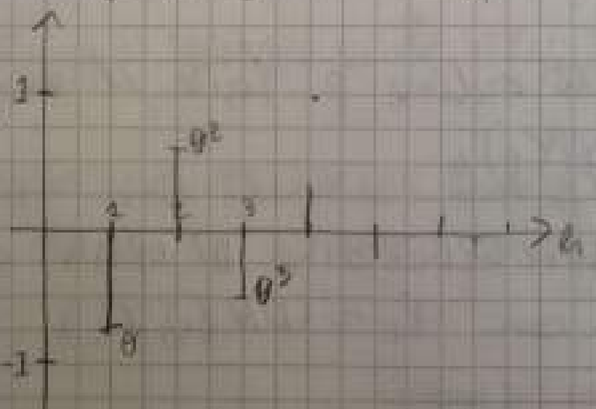
$0<\theta<1$



TENDE A 0 E NON HA UN CUT-OFF.

$\rho(h)$

$-1<\theta<0$



▲ AR(1): $Y_t = \theta Y_{t-1} + \varepsilon_t$, $\varepsilon_t \sim WN(\ )$

$Y_0 \doteq$ CONDIZIONE INIZIALE è il primo valore che consideriamo che può essere noto o V.A.

Noi Y_0 = V.A. t.c. (H) $Y_0 \perp^{\text{incorrelato}} \{\varepsilon_1, \varepsilon_2, \dots\}$

$\Rightarrow$ Partendo da Y_0 e andando in avanti risolvo Y_t

$\Rightarrow Y_t = \theta^t Y_0 + \sum_{s=0}^{t-1} \theta^s \varepsilon_{t-s}$ (analogo ad altra lezione)

OSS AR(1) NON È STAZ. IN COVARIANZA perché

$E[Y_t] = \theta^t E[Y_0] \doteq \mu_t \ (\neq \mu, \forall t)$

$VAR[Y_t] = \theta^{2t} VAR[Y_0] + \sigma^2_\varepsilon \sum_{s=0}^{t-1} \theta^{2s} \doteq \sigma^2_t$

} dipendono in generale da t, $\forall \theta \neq 0$

C.V.D. ■

OSS se (H) $|\theta| < 1 \Rightarrow$ NON C'È IN GENERALE STAZ. IN COV., MA C'È STAZIONARIETÀ ASINTOTICA PER AR(1)

- se $|\theta| < 1$ $\lim_{t \to +\infty} E[Y_t] = 0$

$\lim_{t \to +\infty} VAR[Y_t] = \frac{\sigma^2_\varepsilon}{1-\theta^2}$

I MOMENTI PRIMI E SECONDI CONVERGONO A COSTANTI (non dip. da t)

C.V.D. ■

▲ $|\theta| < 1$ È CONDIZIONE DI STAZIONARIETÀ ASINTOTICA PER AR(1)

OSS se $|\theta| < 1$

- $Y_t = \theta^t Y_0 + \underbrace{\sum_{s=0}^{t-1} \theta^s \varepsilon_{t-s}}_{X_t \text{ dip. } Y_0} =$

$\Rightarrow (Y_t - X_t) = \theta^t Y_0 \xrightarrow[t \to +\infty]{\text{se } |\theta|<1} 0 \quad \forall Y_0$

dove $X_t \xrightarrow[t \to +\infty]{} \sum_{s=0}^{+\infty} \theta^s \varepsilon_{t-s} \sim MA(\infty)$

$\Rightarrow Y_t \xrightarrow[t \to +\infty]{\text{se } |\theta|<1} \sum_{s=0}^{+\infty} \theta^s \varepsilon_{t-s} \sim MA(+\infty)$

(NB) LA CONVERGENZA È IN MEDIA QUADRATICA cioè

$\lim_{t \to +\infty} E[(Y_t - X_t)^2] = 0$

$\Rightarrow \left(E[Y_t] \to 0 \;,\; VAR(Y_t) \to \frac{\sigma_\varepsilon^2}{1-\theta^2} \;,\; \rho(k) \to \theta^k \right) \quad \forall Y_0 \text{ se } |\theta|<1$

Oss Nei modelli AR(1) c'è fase iniziale in cui conta condizione iniziale e poi se c'è condizione di staz. asintotica via via conta sempre meno e ottengo momenti asintotici sopra.

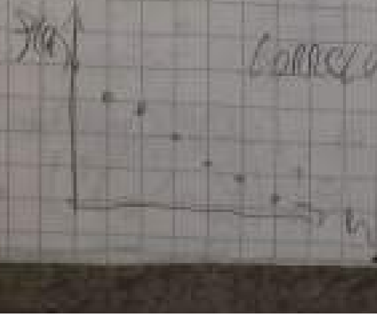
Oss È idea simile ad EQUILIBRIO STABILE IN ECONOMIA (se non sono in eq. ⇒ ne sono attratto, se sono in eq. ⇒ ci starò per sempre)
Qui analogo si arriva a MA(∞) da AR(1) con qualsiasi condizione iniziale (sotto HP) e se a Y_0 do distr. asintotica ⇒ Y_t ha immediatamente raggiunto EQ.

$X_t \sim MA(\infty) \Rightarrow$
$\Rightarrow E[X_t]=0 \qquad VAR[X_t] = \frac{\sigma_\varepsilon^2}{1-\theta^2}$

Th. $Y_t \sim AR(1) \quad |\theta|<1$
se $E[Y_0]=0 \qquad VAR[Y_0] = \frac{\sigma_\varepsilon^2}{1-\theta^2} \doteq \sigma_Y^2$
$\Rightarrow Y_1, Y_2, \ldots \Rightarrow E[Y_t]=0 \;\; \forall t \geq 0$
$VAR[Y_t] = \sigma_Y^2 \quad \forall t \geq 0$ C.V.D. ■

Oss In questo sui libri di testo si scrive AR(1) con $|\theta|<1$ è STAZIONARIO (non solo asintotico) perché (HP) $E[Y_0]=0$, $VAR[Y_0] = \frac{\sigma_\varepsilon^2}{1-\theta^2}$, se no in generale non c'è subito equilibrio, ma c'è fase di transizione

Th Se $Y_t \sim AR(1)$ con $|\theta|<1$
$\Rightarrow \exists!$ RAPPRESENTAZIONE MA(∞) $\sim Y_t$
t.c. $Y_t = \sum_{j=0}^{+\infty} \theta^j \varepsilon_{t-j}$



CORRELOGRAMMA

OSS Se $|\theta|=1 \Rightarrow$ non avremmo staz. nemmeno asintotica.

▲ OPERATORE RITARDO: L è un applicazione che si applica a successioni numeriche e di VA (PROC. STOC. A PARAMETRO DISCRETO) t.c. $LY_t = Y_{t-1} \quad \forall t$ trasla indietro di un punto

$\Rightarrow L(LY_t) = LY_{t-1} = Y_{t-2} \quad \forall t$

$\Rightarrow L^m Y_t = Y_{t-m}$

$L^0 Y_t \triangleq Y_t \qquad \Rightarrow L^0 \triangleq 1$

OSS Se $Y_t = c \quad \forall t$ cioè $\{Y_t\}$ è succ. cost.

$\Rightarrow L^m Y_t = L^m c = c \quad \forall m, t$

$\Rightarrow AR(1)$: $Y_t = \theta Y_{t-1} + \varepsilon_t \Rightarrow L^0 Y_t = \theta L Y_t + \varepsilon_t \Rightarrow$

$\Rightarrow 1 \cdot Y_t - \theta L Y_t = \varepsilon_t \Leftrightarrow (1-\theta L) Y_t = \varepsilon_t$

(abbiamo costruito $(1-\theta L) Y_t$ = POLINOMIO DI GRADO 1 NELL'OPERATORE RITARDO $= L$

$\Rightarrow$ ▲ $AR(p)$: $Y_t - \sum_{j=1}^{p} \theta_j Y_{t-j} = \varepsilon_t \Rightarrow$

$\Rightarrow (1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_p L^p) Y_t = \varepsilon_t$

POLINOMIO DI GRADO p NELL'OPERATORE RITARDO.

$\theta_p(L) \overset{NOT}{=}$ POLINOMIO DI GRADO p IN L.

$\Rightarrow AR(p)$: $\theta_p(L) Y_t = \varepsilon_t$ (scrittura compatta) $\varepsilon_t \sim WN(\,)$

$\Rightarrow \theta_p(L)$ è FILTRO LINEARE CHE APPLICATO A PROCESSO Y_t LO TRASFORMA IN UN PROC. $WN(.)$

$\Rightarrow$ faccio al contrario per risolvere e trovare la soluzione per il proc. $Y_t \Rightarrow$ idea:

$$Y_t = [\theta_p(L)]^{-1} \varepsilon_t \qquad \varepsilon_t \sim WN$$

NB DEVE $\exists [\theta_p(L)]^{-1} \Rightarrow$ SERVONO CONDIZIONI DI INVERTIBILITÀ DEL POLINOMIO $\theta_p(L)$

▲ CONDIZIONE DI INVERTIBILITÀ DI $\theta_p(L)$
⇒ ∃! $\theta_p(L)^{-1} \doteq \Psi_\infty(L)$ (polinomio in L, che sarà di ordine ∞)
t.c. $\theta_p(L) \cdot \Psi_\infty(L) = 1 = L^0$

OSS Non tutti i polinomi sono invertibili

▲ CONDIZIONE DI INVERTIBILITÀ IN AR(1)
Se $|\theta| < 1$ ⇒ ∃ RAPPRESENTAZIONE AR(1) ~ MA(∞)
⇒ $\underline{Y_t} = \sum_{s=0}^{+\infty} \theta^s \varepsilon_{t-s} = \sum_{s=0}^{+\infty} [\theta^s L^s] \varepsilon_t = \underline{\Psi_\infty(L) \cdot \varepsilon_t}$
($\Psi_\infty(L)$ è polinomio di ordine ∞ in L)

AR(1) : $\theta_1(L) Y_t = \varepsilon_t \overset{|\theta|<1}{\Leftrightarrow} Y_t = \Psi_\infty(L) \cdot \varepsilon_t$: RAPPR. MA(∞) DI Y_t
polinomio invertibile
⇒ PER AR(1) $\theta_1(L) = (1 - \theta L)$ È INVERTIBILE ⇔
$|\theta| < 1$
⇒ $(\theta(L))^{-1} = \Psi_\infty(L) = \sum_{s=0}^{+\infty} \theta^s L^s = 1 + \theta L + \theta^2 L^2 + \dots$

OSS CONDIZIONE DI INVERTIBILITÀ DI $\theta(L)$ ⇔ (coincide)
⇔ CONDIZIONE DI STAZIONARIETÀ ASINTOTICA DI AR(1)
⇒ Generalizzato per AR(P)

▲ EQUAZIONE CARATTERISTICA DEL PROC. AR(P)
Es AR(2) : $(1 - \theta_1 L - \theta_2 L^2) Y_t = \varepsilon_t$
⇒ EQUAZIONE CARATTERISTICA: è $\theta(z)$ sostituendo a L, z (che non è più un'applicazione, ma una variabile ∈ ℂ = complessi)
& poniamo = 0.
⇒ EQUAZIONE CARATTERISTICA $\theta_p(z) = 0$
Es $1 - \theta_1 z - \theta_2 z^2 = 0$
⇒ calcolo le soluzioni (radici) dell'equazione caratteristica (∈ ℂ ⇒ ∃ sempre p soluzioni) ⇒ $z_1, \dots, z_p$
⇒ CONDIZIONE DI INVERTIBILITÀ DI AR(2)
Se $|z_1| > 1 \wedge |z_2| > 2$ ⇒ AR(2) È STAZ. IN COV. ASINTOTICA
e $\theta_2(L)$ È INVERTIBILE

⇒ AR(2) sub (MA): $Y_t = (1-\theta_1 L - \theta_2 L^2)^{-1} \cdot \varepsilon_t$

⇒ OSS Spesso nei software non vengono calcolate le radici z_j $j=1,\dots,p$ ma le RADICI INVERSE $\lambda_j = z_j^{-1}$ $\forall j=1,\dots,p$

⇒ CONDIZIONE DI STAZ. EQUIVALENTE

$|\lambda_j|<1$ $\forall j=1,\dots,p$

OSS $(1-\theta_1 L-\theta_2 L^2) \overset{\text{fattorizzo}}{=} \underbrace{(1-\lambda_1 L)\,(1-\lambda_2 L)}_{\text{RADICI INVERSE}}$

⇒ $\theta(L)\, Y_t = \varepsilon_t \sim WN$

$(1-\lambda_2 L)\,[(1-\lambda_1 L) Y_t] = \varepsilon_t \sim WN$

⇒ Cioè per ottenere $\theta(L)$ posso prendere Y_t e applicare il filtro $(1-\lambda_1 L)$ e al processo risultante applicare il filtro $(1-\lambda_2 L)$.

OSS IL POLINOMIO INVERSO $= (1-\theta_1 L-\theta_2 L^2)^{-1} = (1-\lambda_1 L)^{-1}(1-\lambda_2 L)^{-1}$ ($\approx$ J)

• AR(2): $\theta_2(L)\, Y_t = \varepsilon_t \rightarrow Y_t = \overbrace{\theta(L)^{-1}}^{MA(\infty)} \varepsilon_t$ ≥ (HP) $|\lambda_1|<1$ $|\lambda_2|<1$

⇒ $Y_t = \underbrace{(1-\lambda_1 L)^{-1}}_{\psi_{\infty 1}(L)}\ \underbrace{(1-\lambda_2 L)^{-1}}_{\psi_{\infty 2}(L)} \cdot \varepsilon_t$

⇒ $Y_t = (1+\lambda_1 L+\lambda_1^2 L^2+\lambda_1^3 L^3+\dots)(1+\lambda_2 L+\lambda_2^2 L^2+\lambda_2^3 L^3+\dots)\varepsilon_t$

⇒ cioè: Rappresentazione MA(∞) di AR(2)

⇒ $\Psi_\infty(L) = 1+\psi_1 L+\psi_2 L^2+\psi_3 L^3+\dots$

▲ AR(p) $\theta(L)\, Y_t = \varepsilon_t$

eq. caratteristica $\theta(z)=0$

⇒ CONDIZIONE DI INVERTIBILITÀ, STAZIONARIETÀ ARMONIA $|z_j|>1$ $\forall j$

⇒ RAPPR. MA(∞) $Y_t = \underbrace{\theta(L)^{-1}}_{\psi_\infty(L)} \cdot \varepsilon_t$

$\theta(L) = \prod_{i=1}^{p} (1-\lambda_i L)$

$\theta(L)^{-1} = \prod_{i=1}^{p} (1-\lambda_i L)^{-1} = 1+\lambda_i L+\lambda_i^2 L^2+\dots$

17/11/15

AR(p): $\Theta_p(L)\, Y_t = \varepsilon_t$, $\varepsilon_t \sim WN(0, \sigma^2_\varepsilon)$

Non è in generale staz in cov, ma lo è in senso asintotico

(4.2) CONDIZIONI DI STAZIONARIETÀ IN COV. ASINTOTICA

$z \in \mathbb{C} \mid \Theta_p(z) = 0$, $z_1, \dots, z_p$ radici , $|z_i| > 1 \; \forall i$

$\Leftrightarrow \lambda_i = z_i^{-1}$ $|\lambda_i| < 1$ $\forall i$.

RICHIAMO: NUMERI COMPLESSI

- NUMERI IMMAGINARI

▲ $i \doteq$ UNITÀ IMMAGINARIA t.c. $i^2 \doteq -1$

$\Rightarrow \sqrt{-1} = \sqrt{i^2} = \pm i$

▲ NUMERO IMMAGINARIO $\doteq b\, i$ con $b \in \mathbb{R}$

▲ NUMERO COMPLESSO $\doteq a + b\, i$ con $a, b \in \mathbb{R}$

OSS Se equazione di 2° grado non ha soluzioni reali, ma $\Delta < 0$

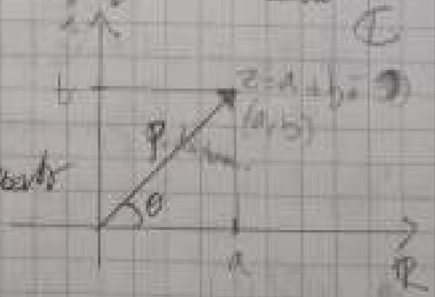
$\Rightarrow \exists$ 2 SOLUZIONI COMPLESSE $z_{1,2}$

$z_{1,2} = a \pm b \cdot i$ $z_1 = a + b \cdot i$ $z_2 = a - bi$

$\Rightarrow$ le soluzioni sono due numeri complessi coniugati

▲ MODULO DI UN NUMERO COMPLESSO z: $\doteq \rho = |z|$ [se z è reale coincide con valore assoluto, se no $\neq$]

$\rho = |z| = \sqrt{a^2 + b^2}$



OSS Un numero complesso può essere individuato tramite le coordinate polari $\doteq (\rho, \theta)$

$z \doteq a + ib = \rho\,(\cos(\theta) + i \sin(\theta))$

$\Rightarrow z^t = \rho^t\,(\cos(t\theta) + i \sin(\theta t))$

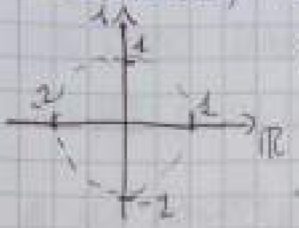
$\Rightarrow \lambda = z^{-1} \doteq \underset{|\lambda|}{\underset{\shortparallel}{\rho^{-1}}}\,[\cos(-\theta) + i \sin(-\theta)]$

(OSS $\rho \doteq |z|$ $\lambda \doteq z^{-1}$ $\Rightarrow \rho_\lambda = |\lambda| = |z^{-1}| = \frac{1}{|z|}$?)

OSS $\frac{z}{x} = z \cdot x^{-1}$

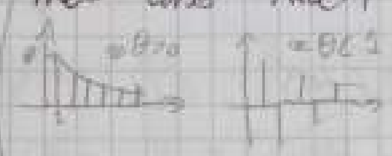
OSS Quando abbiamo numeri complessi entrano in gioco funzioni periodiche. ⇒ le vedremo nel correlogramma.

▲ AR(2) STAZ. IN COV. $|\lambda_1|<1 \wedge |\lambda_2|<1 \Leftrightarrow |z_1|>1 \wedge |z_2|>1$



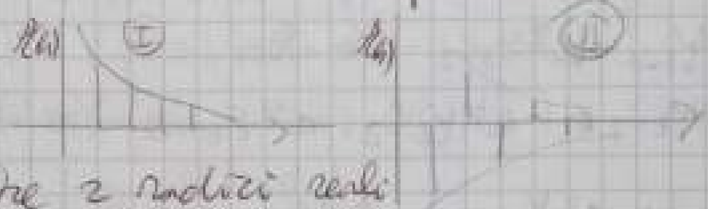
cioè tutte le radici λ_i all'interno del cerchio di raggio unitario nel piano complesso ($\Leftrightarrow z_i$ all'esterno)

• $\rho(h)$ (nel caso AR(1) avevamo due possibilità:



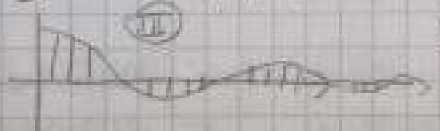
correlogramma cui il valore assoluto converge a zero con andamento esponenziale)

• se AR(2)



$\rho(h)$ I, II se 2 radici reali

III se λ_1, λ_2 COMPLESSE CONIUGATE ($|\lambda_1| = |\lambda_2|$)



andamento periodico ma smorzato ($\Rightarrow 0$)

OSS Correlogramma $\rho(h)$ è collegato ad andamento della previsione ⇒ III rappresenta andamento CICLICO (boom e poi recessione)

OSS Le radici complesse vanno sempre in coppia (CONIUGATE)
es AR(3) ⇒ radici o (3 reali) ∨ (1 reale ∧ 2 complesse coniugate)

OSS $|\rho(h)| \xrightarrow[h \to +\infty]{} 0$ in modo esponenziale per AR(1)

OSS se CORRELOGRAMMA mostra che $\rho(h) = 0 \ \forall h > q$
⇒ è MA(q) (o proc. ARMA(p,q))

Ma per proc. AR(p) non basta il CORRELOGRAMMA per individuare p ⇒ useremo CORRELOGRAMMA PARZIALE

• MA(1): $Y_t = \varepsilon_t + a\varepsilon_{t-1}$ è sempre STAZ IN COV.

$Y_t = (1 + aL)\varepsilon_t = a_1(L)\varepsilon_t$

Oss Potrebbe essere utile (per la stima) TRASFORMARE

il MA(1) $\longrightarrow$ AR(+∞)

• AR(+∞): $Y_t = \sum_{j=1}^{+\infty} \theta_j Y_{t-j} + \varepsilon_t$, $\varepsilon_t \sim WN(\ )$

$\Rightarrow$ se $\exists [a_p(L)]^{-1} \Rightarrow a(L)^{-1} Y_t = \varepsilon_t$ (AR(+∞))

• RICAVO FORMA AR(∞) DELL' MA(1)

MA(1) = $Y_t = a_1(L) \cdot \varepsilon_t \Rightarrow$ AR(+∞). $a(L)^{-1} Y_t = \varepsilon_t$

DIM MA(1) $\stackrel{\Delta}{=} Y_t = \varepsilon_t + a\varepsilon_{t-1}$ (ma $Y_{t-1} = \varepsilon_{t-1} + a\varepsilon_{t-2} \Rightarrow \varepsilon_{t-1} = Y_{t-1} - a\varepsilon_{t-2}$)

$\Rightarrow Y_t = \varepsilon_t + aY_{t-1} - a^2\varepsilon_{t-2}$ (e itero...)

$\Rightarrow Y_t = \varepsilon_t + aY_{t-1} - a^2 Y_{t-2} + a^3\varepsilon_{t-3}$ (e itero...)

$\Rightarrow Y_t = \varepsilon_t + aY_{t-1} - a^2Y_{t-2} + a^3Y_{t-3} - a^4Y_{t-4} + \dots + a^{n-1}Y_{t-n+1}$

$- a^n\varepsilon_{t-n}$ (n pari)

OSS se $a \geq 1 \Rightarrow \varepsilon_{t-n}$ non sparisce mai neanche per $n \to +\infty$

$\Rightarrow$ MA(1) NON È SEMPRE INVERTIBILE cioè RAPPRESENTABILE

COME AR(+∞) DEVE SPARIRE $a^n\varepsilon_{t-n}$ per $n \to +\infty$

• $|a| < 1 \Rightarrow a^n \xrightarrow[n\to+\infty]{} 0$ INVERTIBILE

$\Rightarrow$ MA(1) $\longrightarrow$ AR(∞)

$Y_t = aY_{t-1} - a^2Y_{t-2} + a^3Y_{t-3} - \dots + \varepsilon_t \Leftrightarrow \theta_{\infty}(L)Y_t = \varepsilon_t$

• CONDIZIONE $|a| < 1 \Rightarrow \exists! \, [a_1(L)]^{-1} \stackrel{\Delta}{=} \theta_{\infty}(L)$
INVERTIBILITÀ DI $a_1(L)$

e ricavo $\theta_{\infty}(L)$:

$Y_t - aY_{t-1} + a^2Y_{t-2} - a^3Y_{t-3} + \dots = \varepsilon_t$

$\Rightarrow [a_1(L)]^{-1} = \theta_{+\infty}(L) = 1 - aL + a^2L^2 - a^3L^3 + \dots$

$\Rightarrow \theta_{+\infty}(L)Y_t = \varepsilon_t$

$\Rightarrow$ Se $Y_t = \varepsilon_t + a\,\varepsilon_{t-1}$ con $|a|<1$ $\Rightarrow$ il proc. MA(1) è STAZ. IN COV e INVERTIBILE.

OSS Quando il polinomio $a_q(L)$ è invertibile $\Rightarrow$ il proc. MA(q) è DETTO INVERTIBILE ($\neq$ AR(p) che quando $\theta_p(L)$ invertibile $\Rightarrow$ AR(p) è DETTO STAZIONARIO)

▲ MA(q) è INVERTIBILE ($\Rightarrow$ = AR($+\infty$)) $Y_t = a_q(L)\,\varepsilon_t$

$\Leftrightarrow$ $a_q(z)=0$, $z\in\mathbb{C}$ $\Rightarrow$ soluzioni z_i, $i=1,\dots,q$

è invertibile $\Leftrightarrow$ $|z_i|>1$ $\forall i$ $\vee$ $|\lambda_i|<1$ $\forall i$ $(\lambda_i \doteq z_i^{-1})$

OSS MA(1): $Y_t = (1+aL)\varepsilon_t$ $\quad \lambda_1 = -a$

$\Rightarrow \theta_{+\infty}(L) = (1+aL)^{-1} = (1-(-a)L)^{-1} \doteq$ ($\Rightarrow$ sappiamo quando è invertibile (VEDO AR(1))) $= 1+\theta L+\theta^2L^2+\dots$ se $|\theta|<1$ (con $\theta \doteq -a$)

$\Rightarrow (1+aL)^{-1} = 1-aL+a^2L^2-a^3L^3+a^4L^4+\dots$ se $|a|<1$

OSS PER STIMARE e PREVEDERE PROC MA(q) SERVE CHE SIANO INVERTIBILI.

▲ ARMA(p,q) è MODELLO CON PARTE AR(p) e PARTE MA(q) (MISTO).

$$Y_t - \sum_{j=1}^{p}\theta_j Y_{t-j} = \underbrace{\varepsilon_t - \sum_{j=1}^{q} a_j\,\varepsilon_{t-j}}_{\text{MA(q)}} \qquad \varepsilon_t \sim WN$$

$$\underset{WN}{\varepsilon_t} \longrightarrow \boxed{MA(q)} \longrightarrow \underset{MA(q)}{X_t} \longrightarrow \boxed{AR(p)} \longrightarrow \underset{ARMA(p,q)}{Y_t}$$

OSS MA e AR sono equivalenti in caso di condizioni STAZ. INV. cambia solo il n° parametri (FINITO VS INFINITO)

$\Rightarrow$ ARMA(p,q) SONO MODELLI MISTI DEI PROC. PURI AR(p), MA(q) (di ordine finito) $\Rightarrow$ SOTTO CONDIZIONI possiamo vedere ARMA(p,q) IN MA($+\infty$) o AR($+\infty$)

• ARMA(p,q): $\theta_p(L)\, Y_t = a_q(L)\,\varepsilon_t$

⇒ se $\exists!\,(\theta(L))^{-1}$ ⇒ $Y_t = \underbrace{\theta(L)^{-1} a(L)}\cdot\varepsilon_t \triangleq \psi_\infty(L)\,\varepsilon_t \equiv$

(quando $\triangleq$ MA(+∞)

• sotto condizioni invertibilità $|z_i|>1\ \forall i$ (z_i t.c. $\theta(z_i)=0$)

⇒ ARMA(p,q) è asintoticamente staz. in cov. ⇔ $\exists!\ \theta(L)^{-1}$

es ARMA(1,q) è asint. staz. in cov ⇔ $|\theta_1|<1$

▲ ARMA(p,q) è invertibile ⇒ AR(+∞)

$\theta(L)\,Y_t = a(L)\cdot\varepsilon_t$ se $\exists!\ a(L)^{-1}$

⇒ $a(L)^{-1}\cdot\theta(L)\,Y_t = \varepsilon \triangleq \theta_\infty(L)\,Y_t = \varepsilon_t$ AR(+∞)

• il proc. ARMA(p,q) è invertibile se $\exists!\ a(L)^{-1}$ ⇔ $z_i: a(z)=0$ ⇒ $\forall i\ |z_i|>1$

oss Y_t ARMA(p,q) se staz. in cov. ⇒ Y_t MA(+∞) (e non c'è cost / intercetta = 0)

⇒ $E[Y_t]=0$

dim $Y_t = \sum_{s=0}^{+\infty}\psi_s\,\varepsilon_{t-s} \overset{MA(\infty)}{\Rightarrow} E[Y_t] = 0+\dots+0=0$ c.v.d.

es ARMA(1,1) $Y_t-\theta Y_{t-1} = \varepsilon_t + a\,\varepsilon_{t-1}$

⇒ $(1+\theta L+\theta^2L^2+\dots)(1+aL)$ ricavo tutti i ψ_j

(ma in generale lo fa il PC.)

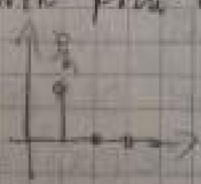
• CORRELOGRAMMA $\rho(k)$

IN UN MODELLO MISTO LA COMPONENTE AR È DOMINANTE

(MA(q) ha memoria finita fino a q, ma AR rimane!)

es AR(1)

MA(1)



⇒ CORRELOGRAMMA GUIDATO DA PARTE AR

⇒ Qualitativamente ARMA(p,q) HA CORRELOGRAMMA COME AR(p)

OSS (ma) fino al $h = q$ entra in gioco anche parte $MA(q)$ alterando il correlogramma di $AR(p)$, ma $\forall h > q$ è quello di $AR(p)$

$\Rightarrow$ PROBLEMA SE VOGLIO STIMARE MODELLO MISTO E GUARDO CORRELOGRAMMA $\Rightarrow$ SE È $MA(q)$ o AR, ma se non è $MA(q)$ $\Rightarrow$ non so se è quale AR ARMA sia $\Rightarrow$ GUARDO CORRELOGRAMMA PARZIALE

- APPROCCIO DI BOX - JENKINGS

13/11/15

OSS Vantaggio $ARMA(p,q)$ rispetto a modelli $AR(p)$, $MA(q)$ è che sono più parsimoniosi (stimano meno parametri)

Andrebbero bene anche AR o MA al posto di ARMA, ma con n° parametri molto alto. $ARMA(p,q) \xrightarrow{\text{spesso}} MA(+\infty) = AR(+\infty)$

($MA(\infty)$ ma coeff sempre + piccoli $\Rightarrow$ approx con $MA(q)$ $q >> 0$
Analogo per $AR(p)$)

ARMA(p,q) SONO MODELLI PARSIMONIOSI (basta stimare pochi parametri)

OSS PROBLEMA: SI POSSONO AVERE MODELLI ARMA (p, q) CON RADICI COMUNI

(es) ARMA(1,1) $\quad \theta(L) Y_t = a(L) \varepsilon_t \Rightarrow (1-\theta L) Y_t = (1 + aL) \varepsilon_t$

$\lambda_1 = \theta \quad \Rightarrow |\theta| < 1 \quad \quad \lambda_1 = -a \Rightarrow |-a| < 1$

$\Rightarrow$ SE RADICI COMUNI $\Rightarrow \theta = -a$

$\Rightarrow (1-\theta L) Y_t = (1+aL)\varepsilon$ ma posso semplificare

$\Rightarrow Y_t = \varepsilon_t \quad \Rightarrow$ ARMA(1,1) CON RADICI COMUNI $\Rightarrow$ È WN

OSS IN GENERALE SE ARMA (p,q) CON RADICI COMUNI UNA PARTE AR e MA SI ELIDE.

$\Rightarrow$ PROBLEMA DI IDENTIFICAZIONE E DI STIMA

$\Rightarrow$ OSS Data una S.S. SIAMO IN GRADO DI MODELLIZZARE E STIMARE SOLO PROC ARMA(p,q) CHE NON HA RADICI COMUNI

$\Rightarrow$ È PROBLEMA PRATICO (non solo teorico)

OSS Se trovo coef. ARMA(1,1) e ottengo $\theta = -\alpha$ so ottengo CORRISPONDENZA a WN,

⇒ ATTENZIONE: SE RADICI COMUNI POTREI COMPLICARE IL MODELLO INUTILMENTE
⇒ torno a modello più semplice (pure test statistici hanno problemi in caso di radici comuni)

OSS PROBLEMA RADICI COMUNI:
⇒ • Uso meno parametri possibili in ARMA(p, q)
• Se DATASET è grande ⇒ USO MODELLO AR(p) CON p GRANDE invece di ARMA(p, q)

(1) NEI CONCRETI MODELLI CARICATI SU DATI [illegible] ⇒ AR(15)

OSS SE VOGLIO PROC. A MEDIA ≠ 0 $E[Y_t] \neq 0$
(per MA(q) basta aggiungere una costante) PER ARMA(p, q)

(2) ARMA(2, q) ma vale $\forall p \neq 2$
⇒ AGGIUNGO COSTANTE α ⇒ VERO VALORE ATTESO DEL PROC.

ARMA(2, q): $\theta(L) Y_t = \alpha + a(L)\varepsilon_t$

OSS $\theta(L) = 1 - \theta_1 L - \theta_2 L^2$

$\theta(z) = 0$ eq. CARATTERISTICA $z \in \mathbb{C}$

$\theta(z) = 1 - \theta_1 z - \theta_2 z^2 = f : \mathbb{C} \to$

⇒ soluzioni z_1, z_2 ⇒ condizione $|z_1| > 1$, $|z_2| > 1$

⇒ $\theta(L) = (1 - \lambda_1 L)(1 - \lambda_2 L)$

$$\left(\begin{aligned} &\Rightarrow \theta(1) = 1 - \theta_1 - \theta_2 = (1 - \lambda_1)(1 - \lambda_2) \\ &\Rightarrow [\theta(1)]^{-1} = \frac{1}{1 - \theta_1 - \theta_2} = \left(\frac{1}{1-\lambda_1}\right) \cdot \left(\frac{1}{1-\lambda_2}\right) = \theta(L)^{-1} \end{aligned} \right)$$

OSS Per trovare $E[Y_t]$ sfrutto rappresentazione $Y_t \sim MA(\infty)$

⇒ $Y_t = [\theta(L)]^{-1} \cdot \alpha + \underbrace{[\theta(L)]^{-1} \cdot a(L)\varepsilon_t}_{MA(\infty)}$

$$\left(\begin{aligned} &\text{con } \theta^{-1}(L) = (1 - \lambda_1 L)^{-1}(1 - \lambda_2 L)^{-1} = \\ &= (1 + \lambda_1 L + \lambda_1^2 L^2 + \dots)(1 + \lambda_2 L + \lambda_2^2 L^2 + \dots) \end{aligned} \right)$$

⇒ Modello con 2 intercette ≠ è simile per ACCO e analisi [illegible] molto ≠.

23/11/15

• DAT MACRO (importanti)

• CORRELOGRAMMA di PIL USA

⇒ ACF che va giù lentamente a dire che prima è LENTO (...) la PACF ⇒ SEMBRA AR(1), ma per ACF $\hat{\rho}(1) = \theta$ ⇒ $\hat{\theta} = 1$ ⇒

⇒ PROBLEMA: Radice unitaria.

⇒ MODELLO $Y_t = Y_{t-1} + \varepsilon_t$ ⇒ AR(1) NON STAZ.

– è Y_t ~ RANDOM WALK ⇒ MODELLO NON È STAZ.

IN QUESTO CASO DIFF. PRIME sono staz.

OSS Queste ACF e PACF sono tipiche di un proc. I(1) non stazionario (ACF va a 0 in modo lento, (lin. non exp) ⇒ CORRELAZIONE ALTA anche dopo molti periodi, per presenza di un trend)

OSS ∃ TEST di RADICE UNITARIA.

OSS Nelle ACF, PACF ci sono 2 bande $\pm 1{,}96/\sqrt{T}$ che è soglia del TEST $H_0: \rho(1) = 0$ vs $H_1: \rho(1) \neq 0$

⇒ INTERVALLO DI CONFIDENZA $\rho(1)$ AL 95%

$$\Pr\left[\hat{\rho}(1) - \frac{1{,}96}{\sqrt{T}} \le \rho(1) \le \hat{\rho}(1) + \frac{1{,}96}{\sqrt{T}}\right] = IC.$$

⇒ TEST vedo se 0 sta in IC o no.

Analogamente posso vedere se la stima sta fuori o dentro LA BANDA.

Se $\hat{\rho}(1) \notin \left[-\frac{1{,}96}{\sqrt{T}}, \frac{1{,}96}{\sqrt{T}}\right]$ ⇒ RIF H_0 al 5%

TEST DI BARTLETT.

OSS se modello non è staz. ⇒ va trasformato in qualcosa di stazionario.

▲ APPROCCIO di BOX-JENKINS (ARIMA)

① PRE-PROCESSING: Guarda s.s. Y_t e se non è STAZ. TRASFORMALA IN STAZIONARIA (modello ARIMA ⇒ ARMA delle DIFFERENZE (prime)) ⇒ $\Delta Y_t \doteq Y_t - Y_{t-1} \doteq (1-L)Y_t$

⇒ elimino trend lineare se non basta ⇒ diff 2ª etc. FINO A STAZ.

⇒ GRETL ⇒ d_PIL

⇒ [illegible] d_PIL PACF ≈ AR(2), ma ACF va giù lentamente ⇒ potrebbe essere ancora non staz.

OSS In letteratura interpreto diff prime indice dei prezzi ≈ INFLAZIONE e si pensa sia $I(2)$ come si vede dal grafico ⇒

GRETL ⇒ d_d_PIL

OSS C'è un problema di ETEROGENEITÀ per staz. per la MEDIA, ma non staz. perché VARIANZA NON È COSTANTE NEL TEMPO. ⇒ USO IL LOG.

GRETL LOG. STABILIZZA LA VARIANZA.

la log-PIL non è staz. IN MEDIA ⇒

$\Rightarrow Y_t \to \ln(Y_t)$ $\quad \Delta \ln(Y_t) = \ln(Y_t) - \ln(Y_{t-1})$

$(1-L)\ln(Y_t) \simeq \dfrac{\Delta Y_t}{Y_{t-1}}$ ⇒ tasso di variazione trimestrale

$\Rightarrow \Delta(\ln(Y_t)) \times 100 \simeq$ tasso di variazione trimestrale in termini %.

OSS Di solito si lavora su tasso annualizzato (%)

$\Delta \ln(Y_t) \times 100 \times 4$ ⇒ È MISURA DEL TASSO DI INFLAZIONE.

⇒ GRETL ⇒ Definisco nuova variabile $d\_log(Y_t)*4*100 =$
= INFLAZIONE ⇒
≤ SEMBRA ANCORA NON STAT.
⇒ d_INFLAZIONE sembra stat.

OSS Nel dataset lavorando con indice dei prezzi al consumo CPI, con $\ln(CPI) = X_t$ e poi diff. seconda $\Delta\Delta X_t = \Delta^2 X_t$

▲ $\Delta^d X_t = \underbrace{\Delta \cdot \ldots \cdot}_{d \text{ volte}} \Delta X_t$

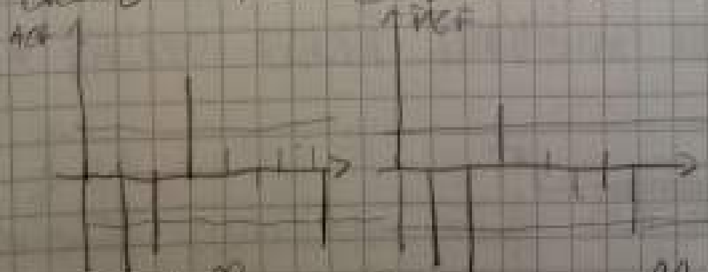
▲ $d \triangleq$ GRADO/ORDINE DI INTEGRAZIONE DELLA VARIABILE X_t = N° MINIMO DI DIFFERENZE DA APPLICARE AFFINCHÉ IL PROCESSO SIA STAZIONARIO.

⇒ ▲ MODELLO ARIMA (p, d, q)
parto da Y_t ⇒ poi diff ⇒ $\Delta^d Y_t \triangleq X_t$ STAZIONARIA IN COV. ⇒ ($Y_t \triangleq$ è INTEGRATA DI ORDINE $d = I(d)$)

OSS $X_t \sim$ STAZ IN COV. SI DICE PROC. $X_t \sim I(0)$ INTEGRATO DI ORDINE 0.
⇒ $X_t \sim ARMA(p, q)$ STAZ. IN COV.

BOX e JENKINS:
- 2ª FASE IDENTIFICO ORDINI p, q.

GRETL ACF e PACF di d_INFLAZIONE



⇒ Potrebbe essere ARMA(p,q) AR(p) MA(q) con q>>0

è difficile leggere correlogramma ⇒ identifico modello e ordine con criteri informativi

OSS NEI MODELLI ARIMA LA PARTE AR È DOMINANTE (MA viene dopo) ⇒ di solito si preferisce lavorare con PROC. AR(p) e se dopo aver fatto le stime mi accorgo che errori non sono WN ⇒ aggiungo parte MA.

OSS Modelli AR(p) si possono stimare con OLS.

⇒ OLS: ⇒ Dependent variable e lags (2)

⇒ GRAPH ⇒ RESIDUI e vedo se WN. (a occhio è difficile) ⇒ TEST per la correlazione dei residui

⇒ GRAPH CORRELOGRAMMA DEI RESIDUI (lunghezza lag è ritardo 8.

OSS Si fa arrivare fino a ritardo 8 (con dati trimestrali) ritardo di 2 anni su dati economici.

Scanned by CamScanner

24/11/12

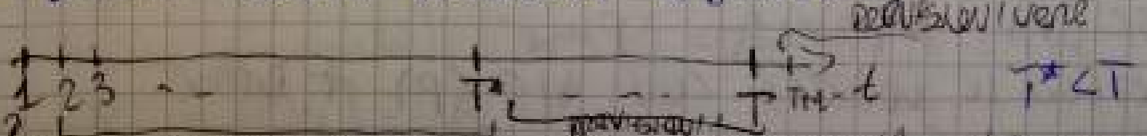
OSS S.S. trimestrale se c'è stagionalità
⇒ vedo componente stagionale del correlogramma
ma nelle freq. stagionali $\hat{\rho}(4 \cdot n)$ la vedo nei multipli di 4 (se trimestrali) (in genere sempre meno intensi)

OSS Nella nostra S.S. NON C'È STAGIONALITÀ È STATA PROPOST. S.S. DESTAGIONALIZZATA ⇒ se rimuovo male succede che non si vede ritardo dopo un anno, non si vede $\hat{\rho}(6)$ sig. (POSSIBILE SPIEGAZIONE)

OSS Dato che obiettivo modello è prevedere ⇒ non ci interessa troppo teoria economica o fase di stima ⇒ VOGLIO VEDERE
= BONTÀ PREVISIONI CON PSEUDO PREVISIONI
⇒ Prima di fare le stime, metto via dei dati per previsioni.
⇒ $Y_t = \beta_0 + \beta_1 Y_{t-1} + \varepsilon_t \quad \forall t$ AR(1) CAMPIONE $Y_1, ..., Y_t$
$y_2 = \beta_0 + \beta_1 y_1 + \varepsilon_1$ (PRIMA eq. DI CUI HO DATI)

⇒ SE USASSI TUTTI I DATI AVREI T-1 EQUAZIONI
$Y_t = \beta_0 + \beta_1 Y_{t-1} + \varepsilon_t \quad t = 2, ..., T$
forma matriciale $\underline{Y} = X \underline{\beta} + \underline{\varepsilon}$ con $\underline{Y} = \begin{bmatrix} Y_2 \\ \vdots \\ Y_T \end{bmatrix}$ $X = \begin{bmatrix} 1 & Y_1 \\ \vdots & \vdots \\ 1 & Y_{T-1} \end{bmatrix}$
⇒ la DIMENSIONE DI CAMPIONE DI STIMA
è $(T-1)$ (per AR(1)), per AR(2) è $(T-2)$...

• VALUTO BONTÀ PREVISIVA MODELLO.



PRE-CAMPIONE CAMPIONE (DI STIMA) OUT-OF-SAMPLE (fuori dal campione di stima)
⇒ PREVISIONI IN-SAMPLE (all'interno del campione di stima)
⇒ PREVISIONI OUT-OF-SAMPLE = PSEUDO PREVISIONI di cui abbiamo in realtà il valore vero. ⇒ OTTENGO ERRORI DI STIMA ⇒
Valuto bontà previsiva del modello e se va bene
⇒ STIMO IL MODELLO SU TUTTI I DATI DISPONIBILI e FACCIO

PREVISIONI VERE

⇒ OGGI lavoro fuori un anno OUT-OF-SAMPLE
⇒ Sottocampione campione 1960:1 - 2004:1

● CRITERI INFORMATIVI (Utili per confrontare qualsiasi modello, non solo ARIMA(p, i, q)) (LIBRO S.W. PAG 614)

▲ $AIC(p) = \ln\left(\frac{SSR(p)}{T}\right) + (p+1)\frac{2}{T}$

PER MODELLO AR(p)

$SSR(p) = \sum_t \hat{\varepsilon}_t^2$ (OSS $\frac{SSR(p)}{T}$ è stima (distorta) DELLA VARIANZA DEL PROC. ε_t → NON È MISURA DI FIT DI ADATTAMENTO DEL MODELLO AI DATI)

OSS • $\ln\left(\frac{SSR(p)}{T}\right)$ è misura di cattivo adattamento ai dati

• $(p+1)\frac{2}{T}$ è PENALIZZAZIONE PER MODELLI NON PARSIMONIOSI

OSS $p+1 \stackrel{\circ}{=} K$ = n° parametri da stimare (p coef. + costante)

▲ $BIC(p) = \ln\left(\frac{SSR(p)}{T}\right) + (p+1)\frac{\ln(T)}{T}$

Analogo ad AIC(p), ma penalizzazione maggiore per piccoli campioni.

⇒ SELEZIONE MODELLO

Calcolo AR(p) $\forall p = 0, \dots, p_{max}$

⇒ AIC(p) $\forall p = 0, \dots, p_{max}$

⇒ SCELGO IL MODELLO CON AIC(p) + BASSO

OSS Come R^2 fit ↑ con p↑ ⇒ per questo c'è penalizzazione

OSS AIC tende a farci scegliere un modello SOVRAPARAMETRIZZATO. BIC è CRITERIO INFORMATIVO CONSISTENTE ⇒ TENDE (A +∞) A SCEGLIERE IL MODELLO "CORRETTO". ⇒ Tende ad essere + PARSIMONIOSO

OSS Confronto va fatto a parità di DIMENSIONE CAMPIONARIA e INFO CAMPIONARIA. I modelli vanno tutti stimati sullo stesso campione di stima.

⇒ PROCEDURA CORRETTA DI SELEZIONE MODELLO:

① SEPARO DATI

② scelgo che modelli voglio indagare (es. ARMA(p,q)) e ordine MAX ⇒ CONFRONTO MODELLI SU STESSO CAMPIONE.

⇒ GRETL MODELLO AR(3) ⇒ SIC(3) = 692,9926 AIC(3) = 664,..
MODELLO AR(4) ⇒ SIC(4) = 679,25 AIC(4) = 663,37

OSS METODO DI STIMA OLS VA BENE PER MODELLI AR(P), MA NON VA BENE PER MA(q) O ARMA(p,q)

● METODI DI STIMA PER AR(P). DATI $y_1, \dots, y_T$

• METODO OLS: ⇒ Uso come CAMPIONE DI STIMA (NON TUTTO IL CAMPIONE, MA PERDO IL PRECAMPIONE) = $y_{p+1}, \dots, y_T$ e SI CONDIZIONA ALLE PRIME P OSSERVAZIONI (PRE-CAMPIONE).

OSS Non usiamo info su $y_1, \dots, y_p$ come distr. prob. ma $y_1, \dots, y_p$ sono NUMERI A CUI CI CONDIZIONIAMO.

OSS Potremmo voler utilizzare info distributiva (o sui primi momenti) $Y_t = \beta_0 + \beta_1 Y_{t-1} + \varepsilon_t$ $|\beta_1| < 1$

$$Y_1 \sim E[Y_1] = \frac{\beta_0}{1-\beta_1}$$

$$VAR[Y_1] = \frac{\sigma^2_\varepsilon}{1-\beta_1^2}$$

▲ METODO DI STIMA DI MAX VEROSIMIGLIANZA (qui sfruttiamo anche $E(Y_1)$ $VAR[Y_1]$ legate con β_0, β_1, ma serve DISTRIBUZIONE COMPLETA)

(H0) $Y_t \sim N(\mu_Y, \sigma^2_Y)$

$\Rightarrow$ (H0) $(Y_1, \dots, Y_T) \sim N^{(T)}$

$\Rightarrow f(y_1, \dots, y_T) \overset{\text{ass. cond.}}{=} f(y_1) \cdot \underbrace{f(y_2, \dots, y_T | y_1)}_{\Downarrow}$

OSS Metodo OLS sfrutta info su $\Downarrow$ e manda $f(y_1)$

METODO DI MAX VEROSIMIGLIANZA COMPLETA SFRUTTA TUTTA INFO.

OSS Però gli OLS non richiedono (H0) DISTRIBUTIVA.

OSS Se uso INFO $f(y_2, \dots, y_T | y_1)^{\sim N(\cdot)}$ e non $f(y_1)^{\sim N(\cdot)}$

$\Rightarrow$ USO METODO DI MAX VEROSIMIGLIANZA CONDIZIONATA (al pre-campione)

OSS IN (H0) DI NORMALITÀ SI DIMOSTRA CHE STIME AR(P) SONO UGUALI CON METODO OLS O DI MAX VEROSIMIGLIANZA CONDIZIONATA, SE USO INFO COMPLETA $\Rightarrow$ METODO DI MAX VEROSIMIGLIANZA COMPLETA $\Rightarrow$ EQUAZIONI CHIAVE NON LINEARI $\Rightarrow$ NO OLS $\Rightarrow$ STIME NUMERICHE.

OSS Se campione è grande possiamo trascurare info pre campione (poco importante $\Rightarrow$ uso OLS)

- STIMA MA ARMA.

OSS MA o ARMA NON SONO PIÙ MODELLI DI REGRESSIONE
→ NON POSSO USARE METODO DI STIMA OLS.

⇒ ▲ METODO DI STIMA DI MAX VEROSIMIGLIANZA.
(sub HP $\varepsilon_t \sim N(0, \sigma^2_\varepsilon)$)

es MA(1) $Y_t = \varepsilon_t + a\,\varepsilon_{t-1}$ (⇒ non posso fare reg. perché ε_t e ε_{t-1} non sono osservabili (LATENTI))

▲ NLS (NON LINEAR LEAST SQUARES)

OSS ENTRAMBI I METODI FUNZIONANO SOLO PER PROC. INVERTIBILI

$\Rightarrow Y_t = \sum_{j=1}^{+\infty} \varphi_j Y_{t-j} + \varepsilon_t$ con $\varphi_j = a^j$

⇒ IDEA: TRONCO E STIMO CON MLE o NLS (VEDO LIBRO INVERTIBILITÀ)

NLS ⇒ Metodi numerici.

GRETL stima modello ARMA
⇒ IN GRETL può da radici e non λ AR(2) vede 2 radici complesse e 2 reali MA |z|<1 ⇒ MODELLO STAZ INVER.

OSS Se t'interessa lag 8 e non tutti i precedenti è cmq AR(8) con vincoli sui coef. ⇒ stime e previsioni con AR(8)

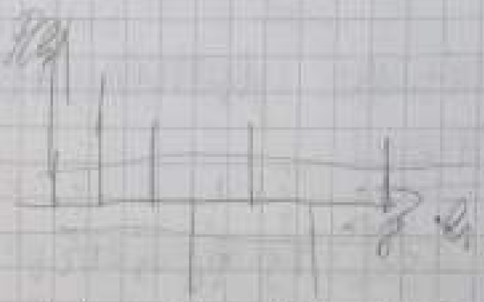
Scanned by CamScanner

• APPROCCIO BOX - JENKINS

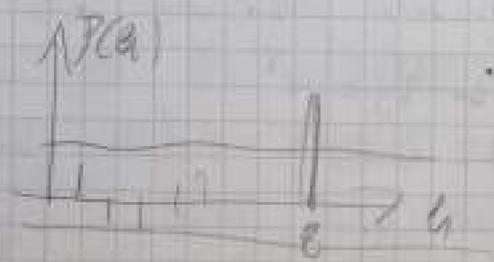
1. IDENTIFICAZIONE MODELLO e ORDINE p, q.
2. STIMA
3. DIAGNOSTICHE per CORRETTA SPEC MODELLO (Test Corr WN.)
   - ⇒ Se non va bene torno a 1
   - Se va bene ⇒ 4
4. PREVISIONI OUT-OF-SAMPLE.

...

OSS ⇒ $Y_t \sim MA(8)$ ⇒



ma



⇒ $Y_t \sim MA(8)$ VINCOLATO

1/4/16

• DATI: PIL ITALIANO TRIMESTRALE DESTAGIONALIZZATO IN TERMINI REALI 1960:1 - 2012:2 OECD

• Se no uso dati Moodle, ma problema dati in periodo di crisi ⇒ non c'è crescita negli ultimi anni.
Ma cmq PIL ha trend ⇒ USO DIFF. PRIME = tasso di crescita

X CASA ESERCIZIO: ① STIMO MODELLO PER DATI '60:1-2004:4 e PREVEDO 2005:1-2007:4 previsioni
② STIMO su tutto '60:1-2007:4 stime 2008:1-2012:2 Previsioni (dovrebbero andare male a causa crisi) (le previsioni statiche dopo un po' il modello si adatta)

② ARCH MODELLI

OSS Nei modelli ARMA standard visto la varianza condizionale è costante nel tempo (essendo proc. stazionario) ma voglio AR $\varepsilon_t \sim D.M.$ $\Rightarrow E[\varepsilon_t | F_{t-1}] = 0$

⇒ Non sappiamo nulla sulla varianza

$Var(\varepsilon_t) = ?$ $\quad Var[\varepsilon_t | F_{t-1}] = ?$

⇒ Assunzione $\varepsilon_t \sim D.M.$ Vincola ① $E[\varepsilon_t] = 0$ ∧
① $cov(\varepsilon_t, \varepsilon_{t-h}) = 0 \; \forall h$, ma potrebbe essere anche non staz.

⇒ Aggiungo ⓑ su varianza di $\varepsilon_t \sim D.M.$
e in più ⓒ ETEROSCHEDASTICITÀ CONDIZIONATA ⇒ ε_t è detto PROC. DI INNOVAZIONE.

(1) $\varepsilon_t \sim DM$

(2) $VAR[\varepsilon_t | F_{t-1}] = \sigma^2_\varepsilon$ OMOSCHEDASTICITÀ [condizionale]

$F_{t-1} = I_{t-1}$

$\Rightarrow VAR(\varepsilon_t) = \sigma^2_\varepsilon$

$\Rightarrow$ se $\varepsilon_t \sim DM$. e $VAR(\varepsilon_t) = \sigma^2_\varepsilon \Rightarrow$

$\varepsilon_t \sim$ PROCESSO DI INNOVAZIONE (componente non prevista della variabile dipendente)

OSS $\varepsilon_t \sim$ proc. di innovazione $\Rightarrow \varepsilon_t \sim WN(0, \sigma^2_\varepsilon)$

OSS (2) $\varepsilon_t \sim$ proc. innovazione è restrittiva in molti problemi reali (es inflazione) $\Rightarrow$

Meglio (1) più debole $\varepsilon_t \sim D.M.$, (come modelli ARCH) in cui la varianza condizionata può variare nel tempo. [$= f$ (assetto informativo)]

$VAR(\varepsilon_t | F_{t-1}] = h(F_{t-1})$

(ma si può comunque avere $VAR(\varepsilon_t) = \sigma^2_\varepsilon \Rightarrow$ STAZ.)

$\Rightarrow$ (1) • $VAR[\varepsilon_t | F_{t-1}) = h(F_{t-1})$ (presenza di eteroschedasticità da es. con test di White)

• $VAR(\varepsilon_t) = \sigma^2_\varepsilon$ (perché proc. ARMA è staz.)

$\Rightarrow$ Nel caso dei modelli AR = modello di regressione dinamica con OLS + $se(\hat\beta_j)$ ROBUSTI DI WHITE ALL'ETEROSCHEDASTICITÀ

$\Rightarrow$ ENGLE test LM (eteroschedasticità)

$\Rightarrow$ con H_0: omoschedasticità

(Infatti si sa che le Δ inflazione sono eteroschedastiche)

⇒ la fine inferenza statistica sui parametri non va bene stime standard ⇒ stime in GRETL per ss. con errori consistenti all'eteroschedasticità (HAC) ⇒
⇒ stesse stime coef, ma ≠ SE e test.

• Altro test che vale per tutti modelli dinamici di regressione (test per AR, AR-DL ma non ARMA perché non è modello di regressione)

• TEST DI BREUSCH-GODFREY (in GRETL)
IDEA TEST: quando c'è autocorrelazione nei residui si potrebbe sfruttare l'info dei residui.
cioè: MODELLO DI REG DINAMICA

$$Y_t = X_t^T \beta + u_t \qquad u_t \overset{H_0}{\sim} WN(0, \sigma_u^2)$$ cioè incorrelati nel tempo

$H_0: u_t \sim WN$ vs $H_1: u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + \rho_3 u_{t-3} + \rho_4 u_{t-4} + v_t$ con $v_t \sim WN$ (es test per autocorr residui fino a ritardo 4)

cioè $H_1: u_t \sim AR(4)$

⇒ sotto H_1: $Y_t = X_t^T \beta + \rho_1 u_{t-1} + \dots + \rho_4 u_{t-4} + v_t$

⇒ $H_0: \rho_1 = \rho_2 = \rho_3 = \rho_4 = 0$

⇒ IDEA: fai reg. OLS e fai F-test su H_0, ma IL PROBLEMA CHE u_t SONO NON OSSERVABILI ⇒ LI RIMPIAZZO CON RESIDUI (stima consistente degli errori ⇒ va bene per campioni grandi)

OSS Inoltre è più corretto rimpiazzare Y_t con residuo $\hat{u}_t$

⇒ REGRESSIONE AUSILIARIA

$$\hat{u}_t = X_t^T B + \rho_1 \hat{u}_{t-1} + \dots + \rho_q \hat{u}_{t-q} + v_t$$

⇒ Analogo a test di WHITE R^2 della REG AUSILIARIA $= R_u^2$

⇒ $(T-q) R_u^2 \overset{H_0}{\underset{T \to \infty}{\sim}} \chi^2(q)$

#OSSERVAZIONI DELLA REG. AUSILIARIA

OSS Ovviamente potenza del test si riduce al crescere dei ritardi testati ⇒ MEGLIO PROCEDERE IN VIA ITERATIVA:

Test H_1: $u_t \sim AR(1)$ ⇒ se trovo autocorr ⇒ STOP.

se non rif. H_0: u_t incorr ⇒ testo $AR(p)$, $p \geq 2$

OSS al crescere di p? ⇒ TRADE-OFF:

- POTENZA TEST ↓ vs • CONSIDERO MAGGIORI ALTERNATIVE POSSIBILI (es. almeno testo ritardi fino a un anno)

- GARCH (previsione OUT-OF-SAMPLE)

PREVISIONE (SINGOLA, E PIÙ ACCURATE) CON REG.

OSS INTERVALLI DI CONFIDENZA PREVISIONE ...

- modelli anche con AR(4) e confronto MSE

OSS Prevedere l'inflazione dal passato è difficile ⇒ CATTIVE PREVISIONI ⇒ è modello BENCHMARK con cui confrontare previsioni modelli teoria economica.

⇒ svolgo conti:

$$(1+\lambda_1 L+\lambda_1^2 L^2+\dots)(1+\lambda_2 L+\lambda_2^2 L^2+\dots)\cdot\alpha =$$

(oss $L^n\alpha=\alpha$ cost) $=$

$$=(1+\lambda_1 L+\lambda_1^2 L^2+\dots)(1+\lambda_2\alpha+\lambda_2^2\alpha+\dots)=$$

$$=(1+\lambda_1 L+\lambda_1^2 L^2+\dots)\underbrace{(1+\lambda_2+\lambda_2^2+\dots)}_{\text{serie geom di ragione } \lambda_2 \text{ e } |\lambda_2|<1}\alpha =$$

$$=(1+\lambda_1 L+\lambda_1^2 L^2+\dots)\left[\left(\frac{1}{1-\lambda_2}\right)\alpha\right]_{\text{cost}} =$$ ⇒ analogo a prima

$$=\underbrace{(1+\lambda_1+\lambda_1^2+\dots)}_{\text{serie geom}}\left[\left(\frac{1}{1-\lambda_2}\right)\alpha\right]=\left(\frac{1}{1-\lambda_1}\right)\left(\frac{1}{1-\lambda_2}\right)\alpha$$

$$\Rightarrow [\theta(1)]^{-1}\cdot\alpha=\left(\frac{1}{1-\lambda_1}\right)\cdot\left(\frac{1}{1-\lambda_2}\right)\cdot\alpha$$

pure $[\theta(1)]^{-1}=\frac{1}{1-\theta_1-\theta_2}$

$$Y_t=\frac{\alpha}{1-\theta_1-\theta_2}+\underbrace{[\theta(L)]^{-1}\alpha(L)\cdot\varepsilon_t}_{MA(\infty)}$$

$$\Rightarrow E[Y_t]=\frac{\alpha}{1-\theta_1-\theta_2}$$ in generale $$E[Y_t]=\theta(L)^{-1}\cdot\alpha=$$

$$=\theta(1)^{-1}\cdot\alpha$$

QUESTA FORMULA VALE SEMPRE [illegible] $$=\frac{\alpha}{1-\left(\sum_{j=1}^{p}\theta_j\right)}=\frac{\alpha}{\theta(1)}$$

Oss se $\theta(1)=0$ ⇒ il proc. non è staz in cov. (e non si può scrivere quello scritto sopra)

$$\theta(L)=1-\theta_1 L-\dots-\theta_p L^p$$

$$\theta(1)=1-\theta_1-\dots-\theta_p$$

Non è stazionario perché se $\theta(1)=0$ ⇒ per $\theta(z)=0$ ⇒ $z=1$ è uno zero della funzione ⇒ proc. ha una radice unitaria ⇒ non staz. ($\cancel{|z|>1}$)

Cioè se $\theta_1+\theta_2+\dots+\theta_p=1$ ⇒ ∃ almeno una radice unitaria (per la prima riga) ⇒ non c'è staz in cov.

▲ $\hat{ACF} \doteq \hat{\rho}(h)$ = AUTO CORRELATION FUNCTION

▲ $\hat{PACF} \doteq$ FUNZIONE DI AUTOCORRELAZIONE PARZIALE.

OSS $\rho(h) = corr(Y_t, Y_{t-h})$

$$PACF(h) \stackrel{\Delta}{=} corr(Y_t, Y_{t-h} \mid \overbrace{Y_{t-1}, Y_{t-2}, \dots, Y_{t-h+1}}^{\text{valori che stanno in mezzo}})$$

$\Rightarrow$ (come in reg. lin $Y_t = \alpha + \beta_1 x_t + \beta_2 z_t + \varepsilon_t$ a parità del resto)

OSS $PACF(1) = \rho(1)$ (In mezzo non c'è nulla)

$AR(2): Y_t = \alpha + \underbrace{\beta_1 Y_{t-1}}_{} + \beta_2 Y_{t-2} + \varepsilon_t$ (so controllo Y_{t-1})

$\Rightarrow PACF(2) = \beta_2$ = coef di $AR(2)$ relativo a Y_{t-2}

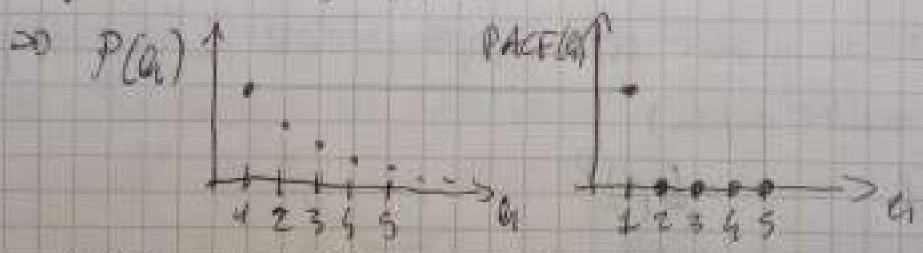
$\Rightarrow$ in $AR(p) \Rightarrow PACF(p) = \theta_p$

OSS Può succedere (succede spesso) che $\rho(h) \neq 0$, ma $PACF(h) = 0$ cioè avendo già inserito altri ritardi più vicini $\Rightarrow$ inserire nel modello anche Y_{t-h} è inutile.

(ES) $Y_t \sim AR(1) \qquad Y_t = \alpha + \theta_1 Y_{t-1} + \varepsilon_t$

e $|\theta_1| < 1 \quad , \theta_1 > 0$

$\Rightarrow$



OSS $AR(1) = AR(p)$ con $\theta_h = 0 \quad \forall h > 1$

$\Rightarrow PACF(h) = 0 \quad \forall h > 1$

OSS PACF DI $AR(p)$ è $\simeq$ ACF DI $MA(q)$

OSS PACF MI CONSENTE DI IDENTIFICARE ORDINE p DI $AR(p)$

$p = \max\{h \mid PACF(h) \neq 0\}$

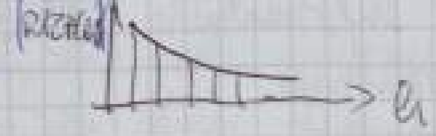
OSS PACF DI $AR(p)$ SI ANNULLA DOPO p RITARDI

Scanned by CamScanner

OSS MA(1) $\xrightarrow{\text{invertibile}}$ AR(+∞) ⇒ coef. AR(+∞) $\varphi_j = a^j$

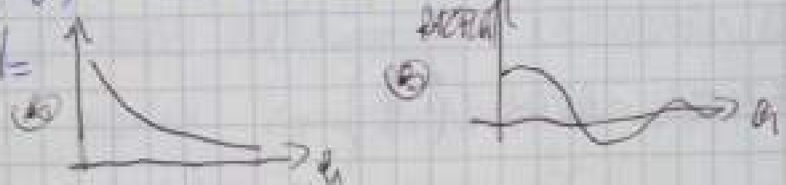
diversi tutti da 0, ma sempre più piccoli

⇒ PACF(1) = ρ(1) e poi decresce a zero in modo esponenziale



OSS ARMA(p,q) $\xrightarrow{(H)}$ AR(+∞)

⇒ |PACF(h)| =



⇒ PACF(h) è dello stesso tipo degli MA

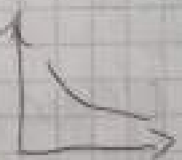
OSS • per proc. AR(p) ⇒ PACF ci dà ordine p

• per proc. MA(q) ⇒ ACF ci basta per ordine q

[ma anche PACF ci conferma)

• ARMA(p,q) ⇒ |ρ(h)|

|PACF|



ACF e PACF possono avere natura simile

⇒ insieme ci permettono di capire che è proc. misto

OSS Approccio BOX-JENKINS

Fase di identificazione modello ARMA(p,q) di p e q è fatta tramite ACF, PACF

OSS ACF e PACF non sono noti ⇒ stimati da S.S. e poi per capire se stat. sig. ≠ 0 ⇒ faccio test statistici e interpreto i grafici (OSS richiede esperienza)

(S.S. cap. 14 libro)

OSS ∃ ulteriore strumento: CRITERI INFORMATIVI:

- AKAIKE - AIC
- BAYESIANO DI SCHWARZ BIC
- SIC

OSS Idea: confronto dei modelli possibili tramite il criterio informativo

(es) $AR(p)$ $p = 1, \dots, p_{max}$ ← max ritardo (es. 15)

⇒ stimo 15 modelli $AR(p)$ ⇒ li confronto e scelgo il migliore in base al criterio informativo (es. AIC)

t.c. MINIMIZZA CRITERIO INFORMATIVO

27/11/15

▲ BOX-JENKINS

1. PRE-PROCESSING
2. IDENTIFICAZIONE ARMA $(\hat{p}, \hat{q})$
3. STIMA
4. ANALISI DI CORRETTA SPECIFICAZIONE DEL MODELLO
   - NO → torna a 2.
   - SI → 5.
5. PREVISIONE

OSS (4) In particolare analisi per vedere se residui WN.

(2) grafici e criteri informativi

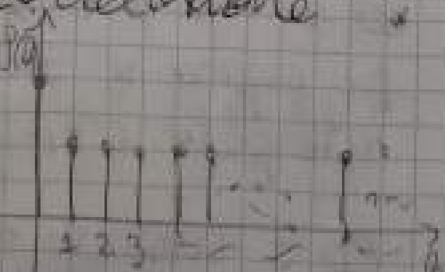
▲ $(Y_t)_t$ PROCESSO SCAMBIABILE (IN SENSO DEBOLE (RISPETTO AI PRIMI DUE MOMENTI)) è particolare proc. stazionario t.c.

$E[Y_t] = \mu$

$VAR[Y_t] = \sigma_Y^2$

$\gamma(h) = COV(Y_t, Y_{t-h}) = \begin{cases} \sigma_Y^2 & h = 0 \\ \gamma \neq 0, & h \neq 0 \end{cases}$ (OSS SARÀ >0)

è proc. staz. in cov, ma in cui la correlazione è costante per ogni ritardo h.



OSS In generale proc. SCAMBIABILE ha memoria infinita ⇒ c'è staz., ma non valgono alcune proprietà importanti sulla somma dei momenti del processo.

⑥ Data s.s. $y_1, \dots, y_T$ ⇒ la media campionaria:
$\bar{Y} = \frac{1}{T} \sum_{t=1}^{T} Y_t$ non è BUONO STIMATORE NON È CONSISTENTE per la media (è solo corretto). $\bar{Y} \not\to E(Y_t) = \mu$

oss la correlazione che rimane nel tempo distrugge alcune proprietà dello stimatore (la varianza dello stimatore non va a zero) ⇒ NON BASTA (HP) STAZIONARIETÀ PER STIMARE IN MODO CONSISTENTE I MOMENTI DEL PROCESSO.

⑤ Analoga a media campionaria non consistente anche la varianza campionaria non lo è:
$\hat{\sigma}_Y^2 = \frac{1}{T} \sum_{t=1}^{T} (Y_t - \bar{Y})^2 \not\to Var[Y_t] = \sigma_Y^2$
se proc. è scambiabile

⇒ Come stimiamo $\hat{\gamma}(h)$? data s.s. $y_1, \dots, y_T$
$\hat{\gamma}(h) = \frac{1}{T} \sum_{t=h+1}^{T} (y_t - \bar{y})(y_{t-h} - \bar{y})$
(poiché $\gamma(h) = E[(Y_t - E(Y_t))(Y_{t-h} - E(Y_{t-h}))]$)

oss Cmq noi stimiamo $\hat{\gamma}(h)$ per valori non troppo grandi, in generale a meno di info economiche dopo un po' è difficile ottenere le correlazioni sig. (⇒ per $T >> 0$ $T-h \simeq T$)
$\hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)}$

Il correlogramma dei residui $\hat{\rho}_{\hat{u}}(h)$ analogo a $\hat{\rho}(h)$ ma al posto di $\{y_t\}$ uso $\{\hat{u}_t\}$

oss SE PROCESSO SCAMBIABILE ⇒ neanche $\hat{\rho}(h) \not\to \rho(h)$

SE PROC. È SCAMBIABILE NON SI VERIFICA LA CONVERGENZA IN PROB DEI MOMENTI CAMPIONARI AI MOMENTI SPECIFICI DEL PROCESSO.

cioè i momenti calcolati nel tempo non convergono ai momenti in un periodo se dati correlati

⇒ SERVONO ALTRI VINCOLI (HP) SUL PROC. AFFINCHÉ CI SIA CONVERGENZA (in prob) DEI MOMENTI CAMPIONARI AI MOMENTI GRANDI TEORICI. (⇒ BUONE STIME).

(HP) ▲ PROCESSO DEBOLMENTE DIPENDENTE (ASINTOTICAMENTE INDIPENDENTE)

(IDEA: è sempre che (HP) $h \uparrow$ ⇒ v.c. sempre meno correlate ⇒ $\rho(h) \to 0$ velocemente.)

LA DISTR. CONGIUNTA DI (Y_{t_1}, Y_{t_1+h}) come si comporta al crescere di h. ($\perp\!\!\!\perp$ ⇒ CONGIUNTA = PRODOTTO MARGINALI)

$$\Rightarrow \left| F(Y_{t_1}, Y_{t_1+h}) - F(Y_{t_1}) \cdot F(Y_{t_1+h}) \right|$$

OSS SE SONO DEBOLMENTE DIPENDENTI,

$$\Rightarrow \lim_{h \to +\infty} \left| F(Y_{t_1}, Y_{t_1+h}) - F(Y_{t_1}) \cdot F(Y_{t_1+h}) \right| = 0 \quad \forall t_1$$

ma questo non deve valere solo per la congiunta bivariata, ma per ogni congiunta:

$$\lim_{h \to +\infty} \left| F(Y_{t_1}, Y_{t_2}, \dots, Y_{t_n}, Y_{t_1+h}, Y_{t_2+h}, \dots, Y_{t_n+h}) - F(Y_{t_1}, \dots, Y_{t_n}) \cdot F(Y_{t_1+h}, \dots, Y_{t_n+h}) \right|$$

$\forall$ ..

OSS ASINTOTICA $\perp\!\!\!\perp$ ⇒ ASINTOTICA INCORRELAZIONE

▲ ASINTOTICA INCORRELAZIONE PER PROC. STAZ. $\rho(h) \to 0$

Th se (Y_t) è proc. asintoticamente indipendente e staz. in covarianza $\Rightarrow \{Y_t\}_t$ è ergodico per il 1° e 2° momento.

$\bar{Y} \xrightarrow{P} \mu = E[Y_t]$
$\hat{\sigma}^2_Y \xrightarrow{P} \sigma^2_Y$
$\hat{\gamma}(h) \xrightarrow{P} \gamma(h)$
$\hat{\rho}(h) \xrightarrow{P} \rho(h)$

convergenza in prob. dei momenti campionari ai processi staz. (per i primi momenti)

▲ ERGODICITÀ: un proc. $\{Y_t\}_t$ si dice ergodico per il primo momento se $\bar{Y} \xrightarrow{P} E[Y_t] = \mu$

▲ ergodico per il primo e 2° momento $\Leftrightarrow$
$\bar{Y} \xrightarrow{P} \mu$, $\hat{\sigma}^2_Y \xrightarrow{P} \sigma^2_Y$, $\hat{\gamma}(h) \xrightarrow{P} \gamma(h)$

OSS Per buone stime non basta stazionarietà del proc., ma serve anche ERGODICITÀ
$\Rightarrow$ condizione suff. è ASINTOTICA INDIPENDENZA.

OSS Un proc. scambiabile non è ergodico.

OSS se $\{Y_t\}_t \sim WN(0, \sigma^2_Y)$ $\Rightarrow$ $\hat{\rho}_Y(h) \underset{T>>0}{\approx} N(0; \frac{1}{T})$
$\Rightarrow \hat{\sigma}_{\hat{\rho}(h)} = \frac{1}{\sqrt{T}}$

(ecco perchè bande in gretl hanno questa forma)

OSS $\hat{\rho}(h) \perp \hat{\rho}(j)$ $h \neq j$ asintoticamente
($\Rightarrow \simeq \perp$ in grandi campioni)

$\Rightarrow$ ha senso fare il test $H_0: \begin{cases} \rho(1) = 0 \\ \rho(2) = 0 \\ \vdots \\ \rho(H) = 0 \end{cases}$ a partire da $\{\hat{\rho}(h)\}_{h=1,\dots,H}$

$\Rightarrow$ a partire da $\hat{\rho}(h) \underset{T>>0}{\overset{H_0}{\approx}} N(0, \frac{1}{T})$ voglio trovare stat test

$\Rightarrow \sqrt{T} \cdot (\hat{\rho}(h)) \underset{T>>0}{\overset{H_0}{\approx}} N(0, 1)$ $\perp$ per $h = 1, \dots, H$

▲ ⇒ TEST DI BOX-PIERCE (Q-TEST)

$$\sum_{a=1}^{H} [\sqrt{T}\, \hat{\rho}(a)]^2 \underset{T\to\infty}{\overset{H_0}{\approx}} \chi^2(H)$$

OSS In realtà per campioni piccoli il test non funziona molto ⇒ aggiustato pesando + correlazioni per h piccole ⇒

▲ TEST LJUNG-BOX (MA IDEA ANALOGA) $\underset{T\to\infty}{\overset{H_0}{\approx}} \chi^2(H)$

⇒ Uso test Ljung-Box (Q-TEST GRETL)

• TEST DI BARTLET $H_0: \rho(a)=0 \quad \forall h$ FISSATO,
BASATO su $\hat{\rho}(a) \underset{T\to\infty}{\overset{H_0}{\approx}} N(0, \frac{1}{T})$ è il test che sfrutto se confronto correlogramma con banda. (GRETL)

• Come leggo Q-TEST IN GRETL
Ho tanti Q-TEST ognuno dice che tutti i $\rho(a)=0$ sig $\forall h \leq$ LAG-MAX

③ STIMA: per AR(p) PROC. STAZ. IN COV. ed ERGODICO

Modello: $Y_t = \beta_0 + \sum_{j=1}^{p} \theta_j Y_{t-j} + \varepsilon_t \qquad \varepsilon_t \sim WN(0, \sigma^2_\varepsilon)$

OSS Nei proc. AR(p) abbiamo sempre come input WN. ε_t che sappiamo essere incorrelato con Y_t, ma non sappiamo se i regressori non sono correlati (cioè $\varepsilon_t \perp Y_{t-j} \quad j = 1, \dots, p$) $\Rightarrow \hat{OLS}$ CONSISTENTE.
$\Rightarrow$ Servono altre HP.

• Dati S.S. $\underbrace{y_1, \dots, y_p}_{\text{PRE-CAMPIONE}}, \underbrace{y_{p+1}, \dots, y_T}_{\text{CAMPIONE}}$

$\Rightarrow$ tratto precampione non come v.c., ma come numeri noti.

$\Rightarrow$ (HP) $(y_1, \dots, y_p) \perp \varepsilon_t \quad t = p+1, \dots,$
($\Rightarrow$ basta per la consistenza, ma non è molto chiaro)

$\Rightarrow$ (HP) (STD. libro).

HP ① $E[\varepsilon_t | Y_{t-1}, Y_{t-2}, \dots, Y_{t-p}, \dots] = 0 \quad \forall t$

(OSS BASTEREBBE CONDIZIONARE A $Y_{t-1}, \dots, Y_{t-p}$, ma se usiamo l'infinito passato $\Rightarrow$ interpretazione è più forte.)

① $\Rightarrow$ PROPRIETÀ:

OSS ① $E[\varepsilon_t | Y_{t-1}, Y_{t-2}, \dots, Y_{t-p}, \dots] \; \forall t = 0$
$\Rightarrow E[\varepsilon_t | Y_{t-1}, \dots, Y_{t-p}] \quad \forall t$

OSS ⓘⓘ $E[Y_t | Y_{t-1}, \dots] = \beta_0 + \sum_{j=1}^{p} \theta_j Y_{t-j}$
$\shortparallel$
$E[Y_t | Y_{t-1}, \dots, Y_{t-p}]$ ②//

(Vedo dim. su modello di REG. LIN.)

(iii) $E[\varepsilon_t]=0 \quad \forall t$

(iv) (È SBAGLIATA SU LIBRO S.W.) $COV(\varepsilon_t, \varepsilon_{t-h})=0 \ , h\neq 0$

(v) $COV(\varepsilon_t, Y_{t-h})=0 \quad \forall h>0$

DIM X CASA (LIBRO S.W.) OSS solo (iv) è PROPRIETÀ NUOVA

OSS (IP) (1) CONDIZIONANDO A INFINITO PASSATO È MEGLIO DI (IP) $E[\varepsilon_t | Y_{t-1}, \dots, Y_{t-p}]$ PERCHÉ?

(ES) AR(1) $Y_t = \beta_0 + \theta Y_{t-1} + \varepsilon_t$

(IP) (1) $E[\varepsilon_t | Y_{t-1}]=0$

$\Rightarrow E[Y_t | Y_{t-1}] = \beta_0 + \theta Y_{t-1} = g(Y_{t-1})$ F.REG.

ma in generale $E[Y_t | Y_{t-1}, Y_{t-2}] = \delta_0 + \delta_1 Y_{t-1} + \delta_2 Y_{t-2} = g(Y_{t-1}, Y_{t-2})$ F.REG.

OSS In generale F.REG. è funzione di tutte le variabili condizionanti (o a seconda della f.reg. (es AR(1) e AR(2))

$\Rightarrow$ (IP) $E[\varepsilon_t | Y_{t-1}, \dots]=0$ $\wedge$ AR(1) (sempre AR(1))

$\Rightarrow E[Y_t | \underbrace{Y_{t-1} \dots}] = g(I_{t-1}) \overset{AR(1)}{=} \beta_0 + \theta Y_{t-1}$

I_{t-1} = information set dell'infinito passato fino a Y_{t-1}

cioè solo (IP) (1) POSSO AFFERMARE CHE MODELLO AR(1) È IL MODELLO CORRETTO non serve usare AR(p), $p>1$, perché ritardi >1 non sono influenti considerato ritardo di ordine 1.

$\Rightarrow$ AR(p): $Y_t = \beta_0 + \sum_{j=1}^{p} \theta_j Y_{t-j} + \varepsilon_t$

(IP) (1) $E[\varepsilon_t | I_{t-1}]=0$ (ASSUNZIONE FONDAMENTALE CHE ASSUMEREMO SEMPRE)

OSS HP① $\Rightarrow \varepsilon_t \sim$ PROC. DIFFERENZA DI MARTINGALA RISPETTO AD INSIEME INFORMATIVO I_{t-1}.

PROC. DIFF. DI MARTINGALA = PROC. PER IL QUALE IL VALORE CORRENTE DI Y_t BASATO SU TUTTA L'INFO PASSATA È PARI A ZERO. cioè $(X_t)_t \sim$ PROC. DIFF. MARTINGALA

$\Leftrightarrow E[X_t | I_{t-1}] = 0$ con $I_{t-1} = \{X_{t-1}, X_{t-2}, \dots\}$ = insieme passato di X_t

② $E[|X_t|] < +\infty \Rightarrow$ valore assoluto finito $\Rightarrow$
$\Rightarrow \exists E[X_t]$ FINITO (CONDIZIONE TECNICA)

(OSS X_t condizionatamente NEI GARCH)

- NOTAZIONE DM ≡ DIFFERENZA DI MARTINGALA.

Th CONDIZIONE ① $E[X_t | I_{t-1}] = 0 \Rightarrow$
$\Rightarrow$ I $E[X_t] = 0$
$\Rightarrow$ II $cov(X_t, X_{t-h}) = 0 \quad \forall h > 0$ (PROC DM non è AUTOCORR.)
$\Rightarrow$ MIGLIOR PREVISIONE DM DEL VALORE FUTURO BASATA SUI VALORI CORRENTI È ZERO (non è prevedibile dal passato)

OSS PER STIMARE AR(p) + (OLS) CON OLS
HP ① $E[\varepsilon_t | I_{t-1}] = 0$ ($\Rightarrow$ è AR(1) p=1)
④ ∄ PERFETTA COLLINEARITÀ ($\Leftrightarrow$ X DI RANGO PIENO = p+1)
③ $\exists E[Y_t^4] \in]0, +\infty[\quad \forall t$
② (su dati cross section HP campioni IID. su dati temporali) non ha senso $\Rightarrow$ HP②
PROC. STAZ E ASINTOTICAMENTE INDIPENDENTE.

$$\Rightarrow \underset{(T-p)\times 1}{\underline{Y}} = X\underline{\beta} + \underline{\varepsilon} \quad AR(p) \qquad t = p+1, \dots, T$$

• PROPRIETÀ STIMATORE OLS (cont. AR)

▲ $\hat{\beta}_{OLS} = (X^T X)^{-1} X^T \underline{Y}$ ⇒ PROPRIETÀ:

① CONSISTENZA: $\hat{\beta}_{OLS} \xrightarrow[T \to \infty]{P} \beta$

② ASINTOTICA NORMALITÀ CONGIUNTA: $\hat{\beta}_{OLS} \underset{T \to \infty}{\approx} N(\underline{\beta}, \Sigma_{\underline{\beta}})$

Oss NON È CORRETTO (NOVITÀ)
$E[\hat{\beta}_{OLS}] \neq \underline{\beta}$ (È DISTORTO)

Per (3 bis) correttezza $\hat{\beta}_{OLS}$ per s.s. serve (H2) $E[\underline{\varepsilon}|X] = \underline{0}$
ma in s.s. non è possibile.
Perché vorrebbe dire $E[\varepsilon_t|X] = 0 \ \forall t = p+1, \dots, T$
ma es. $t = p+1 \Rightarrow E[\varepsilon_{p+1}|X] \neq 0$ perché $cov(\varepsilon_{p+1}, X)$
ma X contiene tutti i valori della Y es. primo ritardo
$\underline{Y}_{-1} = [Y_p, Y_{p+1}, \dots, Y_{T-1}]^T \Rightarrow cov(\varepsilon_{p+1}, Y_{p+1})$ è
ovviamente $\neq 0$. perché $Y_{p+1} = \beta_0 + \sum_{s=1}^{p} \phi_s Y_{t-s} + \varepsilon_{p+1}$
analogo $\forall t$.
$\Rightarrow E[\underline{\varepsilon}_t|X] \neq \underline{0} \Rightarrow$ OLS NON È CORRETTO

Oss su modelli dinamici (che tra regressori hanno ritardi della variabile dipendente) ⇒ stimatore OLS è distorto, ma consistente (per campioni grandi va bene)

Oss SE MODELLO DINAMICO CON ERRORI AUTOCORRELATI
(es) $Y_t = \beta_0 + \phi Y_{t-1} + u_t \Rightarrow Y_{t-1}$ correlato con u_t
⇒ OLS È INCONSISTENTE DISTORTO E ASINTOTICAMENTE NORMALE ⇒ STIME E TEST DA EVITARE
⇒ SE CON ERRORI AUTOCORRELATI
(⇒ va evitato)

30/11/15

▲ $Y_t \sim ARIMA(p,1,q)$

cioè $\Delta Y_t \sim I(0)$ e $\Delta Y_t \sim ARMA(p,q)$ STAT IN COV.

Posto $X_t = \Delta Y_t$ t.c. $\theta_p(L) X_t = \alpha_q(L) \varepsilon_t$

$\Rightarrow$ le radici λ_i della parte AR t.c. $|\lambda_i| < 1$ $\forall i = 1, \dots, p$

OSS $X_t = (1 - 1 \cdot L) Y_t$ (differenze prime significa applicare polinomio AR(1) con radice unitaria)

$\Rightarrow \theta(L)(1-L) Y_t = \alpha(L) \varepsilon_t$

$$\underbrace{\left[\prod_{i=1}^{p} (1-\lambda_i L)\right]}_{\text{PARTE AR STAZ.}} \underbrace{(1-L)^d}_{\text{PARTE AR NON STAZ.}} Y_t = \alpha_q(L) \varepsilon_t \Rightarrow ARIMA(p,d,q) \quad (\sim I(d))$$

OSS $ARIMA(p,d,q)$ contiene esattamente d radici unitarie nella parte AR.

OSS $ARIMA(p,d,q)$ è PROC OMOGENEO NON STAZ., omogeneo perché dopo d diff. diventa staz.

$\Rightarrow$ POLINOMIO AUTOREGRESSIVO È DI ORDINE $p+d$: $\left[\prod_{i=1}^{p}(1-\lambda_i L)\right](1-L)^d$

$\Rightarrow$ OSS $ARIMA(p,d,q) \equiv ARMA(p+d,q)$ NON STAZ. IN COV.

Ma questa forma di non staz è molto particolare. NON STAZIONARIETÀ OMOGENEA $\Rightarrow$ passare alle differenze d-me porta ad un processo stazionario

• PREVISIONE ARIMA (p,1,1)

$X_t = \Delta Y_t \sim AR(1) \qquad X_t = Y_t - Y_{t-1} \Rightarrow Y_t = X_t + Y_{t-1}$

$\Rightarrow$ previsioni $\hat{X}_{T+h|T}$ · previsioni $\hat{Y}_{T+h|T} = Y_{T} + \hat{X}_{T+h|T}$

OSS ▲ PROC. DIFF. DI MARTINGALA (D.M.) : ε_t

① $E[|\varepsilon_t|] < +\infty$

② $E[\varepsilon_t | I_{t-1}] = 0$

$\Rightarrow E[\varepsilon_t] = 0 \wedge cov(\varepsilon_t, \varepsilon_{t-h}) = 0 \quad \forall h \neq 0$

$\Rightarrow E[\varepsilon_{t+h} | I_t] = 0 \quad \forall h \geq 1$

(dove I_{t-1} = ASSETTO INFORMATIVO = σ-ALGEBRA GENERATA DAL PASSATO DI ε_t =

$\doteq I_{t-1} \doteq \{\varepsilon_{t-1}, \dots\}$)

~~OSS In generale $E[\varepsilon_{t-k} | I_t] \neq 0 = \varepsilon_{t-k} \quad \forall t \geq 1$~~

OSS $E[\varepsilon_t | F_{t-1}] = 0$ e non uso gli assetti informativi crescenti nel tempo $F_{t-1} \subset F_t$

$\Rightarrow$ si dice che ε_t è proc. DM rispetto a F_{t-1} (assetto info. più generale generato anche da v.A $\neq \varepsilon_t$)

$\Rightarrow$ vale ancora $E[\varepsilon_t] = 0 \wedge E[\varepsilon_{t+h} | F_t] = 0 \quad \forall h \geq 1$

ma non più $\cancel{cov(\varepsilon_t, \varepsilon_{t-h}) = 0} \quad \forall h \neq 0$

perché F_t contiene passato di altro proc. stoc. $\doteq X_t$

$\Rightarrow cov(\varepsilon_t, X_{t-h}) = 0 \quad \forall h \neq 0. \quad (\forall X_{t-h} \in F_{t-1})$

② PREVISIONE ARMA(1,1) STAZ. IN COV e INVERTIBILE

$Y_t = \theta Y_{t-1} + a\,\varepsilon_{t-1} + \varepsilon_t \qquad |\theta| < 1 \quad |a| < 1$

$\Rightarrow \varepsilon_t \sim$ D.M. RISPETTO A $F_{t-1} = \{Y_{t-1}, Y_{t-2}, \dots\}$

($\Rightarrow$ so (SIM) che F_{t-1} è equivalente a $I_{t-1} = \{\varepsilon_{t-1}, \varepsilon_{t-2}, \dots\}$

Poiché in rappresentazione AR(+∞) $Y_t = \sum_{j=1}^{+\infty} \varphi_j Y_{t-j} + \varepsilon_t$

$\Rightarrow \varepsilon_t = Y_t - \sum_{j=1}^{+\infty} \varphi_j Y_{t-j} \Rightarrow$ da $\{\varepsilon_t\}$ ricavo $\{Y_t\}$ e

VICEVERSA GRAZIE AD INVERTIBILITÀ E STAZIONARIETÀ) oss Grazie ad equivalenza aretti informativi:

$E[\varepsilon_t | F_{t-1}] = E[\varepsilon_t | I_{t-1}] = 0$, $cov(\varepsilon_t, \varepsilon_{t-q}) = 0 \quad \forall q$

• PREVISIONE (scelta del previsore dipende da criterio di ottimalità)

$\underbrace{Y_1, \ldots, Y_T}_{\text{OSSERVATI o S.S}}, \underbrace{Y_{T+1}, Y_{T+2}, \ldots}_{\text{ciò che voglio prevedere}}$

$\Rightarrow$ PREVISORE: $\hat{Y}_{T+1|T} = g_1(F_T)$, ← ALEATORIA = V.A. FUNZ. DELL'ISTANTE T

$\hat{Y}_{T+2|T} = g_2(F_T)$, ..., $\hat{Y}_{T+k|T} = g(F_T)$

• Le fine previsioni ci condizioniamo sempre all'ultimo dato osservato e andiamo avanti k passi a prevedere. (PREVISIONI DINAMICHE).

• Se invece prevedo solo un passo avanti (cioè non uso la dinamica del modello per andare avanti nel tempo oltre ad 1 passo) $\Rightarrow$ ASSORBO NUOVA INFO $\Rightarrow \hat{Y}_{T+1|T} =$ = PREVISIONI STATICHE)

• oss ERRORE DI PREVISIONE $= (Y_{T+k} - \hat{Y}_{T+k|T}) = V.A.$

$\Rightarrow$ Voglio minimizzare $E[(Y_{T+k} - \hat{Y}_{T+k|T})^2 | F_T]$ errore quadratico medio (è F. perdita simmetrica influenza non sempre adeguata)

$\Rightarrow \min_{\{g_k\}} E[(Y_{T+k} - g_k(F_T))^2 | F_T] \Rightarrow$ SOLUZIONE è

$g_k(F_T) \triangleq E[Y_{T+k} | F_T] \triangleq \hat{Y}_{T+k|T}$ OTTIMALE

(tra le g_k t.c. $\exists$ FINITO $E(g_k^2(\cdot))$)

EQUAZIONE MODELLO AL TEMPO T+1: $Y_{T+1} = \theta Y_T + a\varepsilon_T + \varepsilon_{T+1}$

⇒ Voglio calcolare $\hat{Y}_{T+1|T} = E[Y_{T+1}|F_T] =$ (sostituisco)

$= E[\theta Y_T + a\varepsilon_T + \varepsilon_{T+1}|F_T] \overset{lin}{=} \theta E[Y_T|F_T] + a E[\varepsilon_T|F_T] +$
$+ E[\varepsilon_{T+1}|F_T] = \hat{\theta} Y_T + \hat{a}\varepsilon_t + \underset{(WN)}{0}$

⇒ PREVISIONE ARMA(1,1) UN PASSO AVANTI: $\hat{Y}_{T+1|T} = \hat{\theta} Y_T + \hat{a}\varepsilon_t$

ma ε_t è latente ⇒ come faccio a prevedere?
⇒ devo stimare ε_t (⊛ potrei pensare E media di $E(\varepsilon_t)=0$, ma posso fare di meglio) nella fase di stima abbiamo stimato tutti gli errori: STIMA $\hat{Y}_t$ FILTRO ⇒ stima residui $\hat{\varepsilon}_t$ ⇒ uso queste info (se no ⊛ nella previsione modello ARMA(p,q) = AR(p))

• PREVISIONE 2 PASSI AVANTI (in modo ricorsivo)

$\hat{Y}_{T+2|T} = E[Y_{T+2}|F_T] = E[\theta Y_{T+1} + a\varepsilon_{T+1} + \varepsilon_{T+2}|F_T] =$
$\overset{lin}{=} \theta E[Y_{T+1}|F_T] + a E[\varepsilon_{T+1}|F_T] + E[\varepsilon_{T+2}|F_T] =$

(parte MA(q) sparisce dalla previsione dopo passo q ⇒ 0)
⇒ parte predominante è parte AR(p))

$= \theta \hat{Y}_{T+1|T} = \theta^2 Y_T + a\theta\varepsilon_t$

⇒ $\hat{Y}_{T+2|T} = \hat{\theta}\hat{Y}_{T+1|T} = \hat{\theta}^2 Y_T + \hat{a}\hat{\theta}\hat{\varepsilon}_t$

OSS PREVISIONE ARMA(1,1) contiene già come casi particolari formule previsioni AR(1), MA(1)

• AR(1): $\hat{Y}_{T+1|T} = \hat{\theta} Y_T$ $\quad \hat{Y}_{T+2|T} = \hat{\theta}\hat{Y}_{T+1|T} = \hat{\theta}^2 Y_T$

OSS PER ARMA(1,1) la parte MA(1) NON ENTRA DIRETTAMENTE DOPO h>1, MA CONTINUA AD ENTRARE INDIRETTAMENTE, PREVISIONI ≠ DA AR(1)

OSS $\hat{Y}_{T+K|T} \xrightarrow[K \to +\infty]{} E[Y_t] \triangleq \mu$ SE PROC. $Y_t \sim ARMA(p,q)$ È STAZ. IN COV.

(All'inizio conta direttamente parte MA, poi solo AR, ma asintoticamente, x la stazionarietà sparisce tutto)

⇒ SE FAI PREVISIONI TANTO AVANTI NEL TEMPO (previsioni dinamiche) ⇒ tanto vale abbandonare il modello ARMA(p,q) e stimare μ che x la stazionarietà.

$\hat{\mu} = \frac{1}{T} \sum Y_t$ = MEDIA ARITMETICA DEI DATI.

(perché sotto stazionarietà $\theta^k \xrightarrow[k \to \infty]{} 0$)

OSS Oltre a dare previsione, siamo interessati a dare precisione della previsione: $VAR(\hat{Y}_{T+K|T})$ = Varianza del previsore

OSS $VAR[\hat{Y}_{T+K|T}] \leq VAR[Y_t]$ la varianza del previsore non è + grande della varianza del processo (che è quella della media campionaria)

⇒ Nel lungo periodo ≃ uguali, ma nel breve periodo:

$VAR(\hat{Y}_{T+K|T}) \leq VAR(\hat{\mu})$

⇒ $\hat{Y}_{T+K|T}$ È PIÙ PRECISO DI $\hat{\mu}$, ma se proc. staz. in cov ⇒ asintoticamente =.

$VAR(\hat{Y}_{T+K|T}) \xrightarrow[K \to \infty]{} VAR(Y_t)$

N.B. Per prevedere nel breve periodo conviene usare modelli ARMA(p,q), asintoticamente però è equivalente usare la media campionaria sia come previsione in termini precisione della stima

VEDO LIBRO Trova espressione generale ARMA(p,q) che generalizza ARMA(1,1) (ma stessa idea)

ES AR(2) $Y_{t+k} = \theta_1 Y_{t+k-1} + \theta_2 Y_{t+k-2} + \varepsilon_{t+k}$

⇒ $\hat{Y}_{T+K|T} = \theta_1 \hat{Y}_{T+K-1|T} + \theta_2 \hat{Y}_{T+K-2} + \hat{\varepsilon}_{T+K|T}$

Scanned by CamScanner

cioè valutare il modello in termini di previsioni.

- se $T+k-1 \leq T \Rightarrow$ sostituisco Y_{T+k-1} valore vero e non il suo previsore etc... vedi libro.

⊙ $AR(2)$ ad un passo è ottenuto

$Y_{T+1|T} = \theta_1 Y_T + \theta_2 Y_{T-1} + \varepsilon_{T+1|T}$ $= 0$

⇒ Nel Modello uso anche altre variabili da teoria economica
(es CURVA DI PHILLIPS) uso TASSO DI DISOCCUPAZIONE

(GRAFICO LHUR ⇒ Modello AR(p) su LHUR (distribuzione ...)
ma obiettivo non è prevedere TASSO DI DISOCCUPAZIONE, ma
INFLAZIONE.

• CURVA DI PHILLIPS DI BREVE PERIODO
c'è correlazione negativa tra INFLAZIONE
e TASSO DI DISOCCUPAZIONE.

⇒ se $U_{t-1} \uparrow \Rightarrow \Delta INF_t \downarrow$

(Disocc ↓ ⇒ imprese ad assumere ⇒ economia a pieno ...)
INF ↑

⇒ GRAFICO ⇒ Modello OLS $\Delta INF_t = f(\text{const}, LHUR_{t-1})$

GRAFICO MODELLO AR-DL $\Delta INF_t = f_{t,s}(\text{const}, LHUR_{t-1}; \Delta INF_{t-1})$

31/12/15 GRETL:

▲ MODELLO AR-DL (o ADL):

AR-DL (4,4) $\Rightarrow Y_t = f(Y_{t-1}, \dots, Y_{t-4}, X_{t-1}, \dots, X_{t-4})$ $\forall \varepsilon_t$

OSS Potremmo mettere nell'info anche X_t, ma dal punto di vista previsionale non la X_t solo!

OSS MODELLO AR-DL È MODELLO DI REGRESSIONE (DINAMICO) $\Rightarrow$ posso fare test per modelli di reg. OLS...

GRETL AR DL (4,4) vedo che è meglio (AIC) di AR(4), ma non tutti i regressori sono significativi con T-TEST ma perché regressori (OLS) sono correlati tra loro ma li posso buttare via subito (SE + GRANDI) $\Rightarrow$ FACCIO F-TEST CHE LI CONSIDERA CONGIUNTAMENTE.

$\Rightarrow$ F-TEST WALD MI DICE CHE È MEGLIO AR-DL (2,2)

$\Rightarrow$ se residui del nuovo modello non sono autocorrelati $\Rightarrow$ accetto nuovo modello (se no no!)

($\Rightarrow$ ACF residui, test B-G di autocorr. residui)

$\Rightarrow$ TEST a GRAFICI ok $\Rightarrow$ MODELLO AR-DL (2,2)

NELLA STIMA, POTREI VOLER VERIFICARE SE ANCHE IN PREVISIONE

OSS TEST WHITE NON SEGNALA MANCO ETEROSCHEDASTICITÀ, mentre quello semplice (solo i quadrati) SEGNALA ETEROSCHEDASTICITÀ (perché se troppi regressori non segnala per potenza test è)

OSS NOI ABBIAMO USATO ERRORI ROBUSTI $\Rightarrow$ STIME PER CAMPIONI GRANDI VANNO BENE CMQ

(OLS DISTORTI, MA CONSISTENTI e ASINTOTICAMENTE NORMALI)

⇒ (HP) $\{(Y_t, X_t)\}$ è un processo bivariato staz. in cov.
(HP modello AR DL)

• oss stime migliorate, ma R^2 basso
× OBT. previsioni out-of-sample (1 passo avanti) (stanche)

• MODELLO AR-DL (2,2)
$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \delta_1 X_{t-1} + \delta_2 X_{t-2} + \varepsilon_t$$
(HP) $E(\varepsilon_t \mid F_{t-1}) = 0$, dove $F_{t-1} = \left\{ \begin{matrix} X_{t-1}, X_{t-2}, \dots \\ Y_{t-1}, Y_{t-2}, \dots \end{matrix} \right\}$
$\varepsilon_t \sim$ D.M. rispetto a $F_{t-1} = I_{t-1} \supseteq \{\varepsilon_{t-1}, \varepsilon_{t-2}, \dots\}$
⇒ (ε_t WN)

•
⇒ previsore ad un passo: $E[Y_t \mid F_{t-1}] = \beta_0 + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \delta_1 X_{t-1} + \delta_2 X_{t-2}$
$\hat{Y}_{T+1|T} = E[Y_{T+1} \mid F_T]$ $\forall T$
⇒ varianza dell'errore di previsione un passo avanti (modello corretto)
$VAR(\hat{Y}_{T+h|T+h-1} - Y_{T+h}) = VAR(\varepsilon_{T+h}) = \sigma^2_\varepsilon$ se c'è omoschedasticità
Con $S.E.R.^2 = \hat{\sigma}^2_\varepsilon$ ⇒ creo nell'intervallo di confidenza previsioni sub (HP) $\varepsilon_{T+h} \sim N(0, \sigma^2_\varepsilon)$ (HP) normalità

• oss Tutto il ragionamento è basato sul fatto che i parametri siano noti ⇒ unica fonte di incertezza è ε_t (ma non è così!) ⇒ devo stimare i parametri! ⇒ 2 fonti di incertezza/errori
① sui parametri che vanno stimati
② bontà modello su σ^2_ε
⇒ intervalli di previsione si allargano di più, ma in campioni grandi se $(\hat{\beta}) \xrightarrow{p} 0$ ⇒ variabilità dell'errore è molto molto più grande della variabilità dei parametri (che di solito viene trascurata)

OSS ERRORI DI PREV AD UN PROCESSO I(1), PER IL CRESCERE DEI PASSI h ⇒ ERRORE MA(h) ⇒
⇒ LA PREVISIONE SI SCOSTA DI PIÙ FINO A CHE RAGGIUNGE LA VARIANZA MARGINALE DEL PROCESSO
vedi libro Verbeek

- MODELLO ARDL:
⇒ LA RELAZIONE TRA LE DUE VARIABILI SI DICE STABILE SE LA PARTE AR È STAZIONARIA.
(cioè in cui (⇔) IL POLINOMIO AUTOREGRESSIVO È INVERTIBILE)

▲ MODELLO AR-DL: (in generale) $\underbrace{\theta(L)}_{AR(p)} Y_t = \beta_0 + \underbrace{\delta(L)}_{DL(p)} X_t + \varepsilon_t$

⇒ es AR-DL(2,2)

$$(1-\theta_1 L-\theta_2 L^2)Y_t = (\delta_1+\delta_2 L^2)X_t + \beta_0 + \varepsilon_t$$

⇒ ▲ ARDL È STABILE ⇔
⇔ $\theta(L)$ È INVERTIBILE (quella che nel modello AR(p) è condizione di staz. in cov. nel modello ARDL è condizione di stabilità)

XCASA: ESERCIZIO SU PIL

OSS STABILITÀ? LA RELAZIONE TRA 2 VARIABILI (I PARAMETRI) È COSTANTE NEL TEMPO O È CAMBIATA?
(es INTRODUZIONE NUOVI FATTORI es EURO)
⇒ es $Y_t = \beta_{0,t} + \beta_{1,t} X_t + \varepsilon_t$ ⇒ RELAZIONE CAMBIA NEL TEMPO.
Se c'è un cambiamento brusco ⇒ c'è un break strutturale.

⇒ IN τ^* C'È BREAK STRUTTURALE.

⇒ TEST DI CHOW: individuato dove c'è BREAK strutturale ⇒ test di CHOW quando stimo prima e dopo T^* per vedere se sono simili.

- Il Test di CHOW può essere implementato come un F-TEST inserendo due regressori DUMMIES = 1 nel secondo periodo che modificano intercetta e inclinazione.
- $Y_t = \beta_0 + \beta_1 X_t + \beta_3 I_{t>T^*} + \beta_4 I_{t>T^*} \cdot X_t + \varepsilon_t$

⇒ Test CHOW $H_0: \beta_3 = \beta_4 = 0$ ⇒ F-TEST $F_{(2, +\infty)}$

- ⇒ Nel caso c'è davvero Break ⇒ a scopi previsivi uso solo info dopo test.

OOER ⇒ CHOW TEST ⇒ Brk (se non ho info) ⇒ METÀ DEL CAMPIONE. p-value = 0.42 ⇒ NON SEMBRA CHE C'È BREAK

OSS Un MIGLIORAMENTO DEL CHOW-TEST è QLR come CHOW TEST, MA LA DATA DEL TEST VIENE SCELTA IN MODO STATISTICO;

IDEA: Viene fatto CHOW-TEST ∀ possibile data di Break
- e viene riportata la statistica test F del test di CHOW ∀ data.

OSS Escludo prime e ultime osservazioni (perché se no ho problemi nella distr. test.)

⇒ TEST QLR: Prima parte da DATA BREAK PIÙ PROBABILE (con valore + alto della statistica Wald)

⇒ POTREI FARE CHOW TEST in quella data come Break. MA PROCEDIMENTO È SBAGLIATO! Perché data del test non è stata selezionata in modo esogeno, ma dal campione (ERRORE)

(è un problema di test multipli)

(inoltre il sup delle statistiche CHOW TEST non è + una F o χ^2, ma non è nota ⇒ c'è distribuzione campionaria statistica tabulata (con Monte Carlo)

⇒ vedo se sta fuori da soglia stat test simulata

⇒ vedo se c'è stabilità parametri

⇒ QLR test c'è break ⇒ stimo modello su campione post break.

⇒ i test CUSUM (CUSUM quadrati (+ [illegible])) per stabilità dei parametri.

⇒ modo carino per vedere stabilità è grafici stime parametri al crescere osservazioni.

(All'inizio + rumore, ma se poi si stabilizzano ⇒ ⇒ stabilità

(PREVEDIBILITÀ)

(CAP 14) 7/12/15

- **TEST DI CAUSALITÀ DI GRANGER** (nome un po' fuorviante)

OSS CAUSALITÀ in realtà non spiega relazione causa effetto, ma nel senso di PREVEDIBILITÀ.

In un modello dinamico (in cui $Y_t = \Delta INFL_t$ $X_t = U_t$)

$\Rightarrow$ vogliamo prevedere $\hat{Y}_{t+1} = g(Y_t, Y_{t-1}, Y_{t-2}, \dots ; X_t, X_{t-1}, \dots)$

IDEA: se dopo aver usato tutta l'info su $Y_t, Y_{t-1}, \dots$ passata e corrente se l'info passata e corrente su $X_t, X_{t-1}, \dots$ è utile cmq per prevedere Y_{t+1} $\Rightarrow$

$\{X_t\} \xrightarrow{G.C.} \{Y_t\}$

- $\Rightarrow$ Usiamo modello AR-DL per testare GRANGER CAUSALITY

TEST di G-C.

H_0: Non c'è G.C.

$\Rightarrow$ se (HP) RELAZIONE TRA VARIABILI È LINEARE

$\Rightarrow$ (HP) Modello ARDL (p_1, p_2)

$\Rightarrow H_0: \delta_1 = \dots = \delta_{p_2} = 0$ (coef. DL stat. sign. $= 0$)

$\Rightarrow$ FACCIO TEST F per MODELLO.

(GRETL ARDL(?,2) TEST F omette variabili,
$\Rightarrow$ VEDIAMO CHE DISOCCUPAZIONE G.C. INFLAZIONE)

(CAP 15) • STIMA DEGLI EFFETTI CAUSALI DINAMICI

• CROSS-SECTION

(A1) relazione lineare fra 2 variabili $Y_i = \beta_0 + \beta_1 X_i + u_i$, $i = 1, \dots, n$

⇒ (A1) MODELLO

① $E[u_i | X_i] = 0$
② $(Y_i, X_i) \sim$ i.i.d. } ⇒ $E[u_i | X_1, \dots, X_n] = 0$

⇒ $COV[u_i, X_j] = 0 \quad \forall i, \forall j$ anche diversi

OSS (A1) $E[\underline{u} | X] = \underline{0}$ è NECESSARIA PER LA CORRETTEZZA DELLO STIMATORE.

⇒ Analogamente su S.S.

• TIME-SERIES

Modello $Y_t = \beta_0 + \beta_1 X_t + u_t \quad t = 1, \dots, T$

⇒ $\underline{Y} = X\underline{\beta} + \underline{u}$

⇒ (A1) $E[\underline{u} | X] = \underline{0}$ cioè $E[u_t | X_1, \dots, X_T] = 0 \quad \forall t$

⇒ OLS È CORRETTO (sotto (A1))

Ma questa (A1) non è valida nei modelli dinamici

(es $X_t = Y_{t-1}$) ⇒ OLS DISTORTI, MA CONSISTENTI

OSS (A1) $E[\underline{u} | X] = 0 \Rightarrow cov(u_t, X_\tau) = 0 \quad \forall t, \tau = 1, \dots, T$

cioè errore corrente (u_t) è INCORRELATO CON I VALORI PASSATI, CORRENTI e FUTURI DELLE X.

OSS $\boxed{COV(u_t, X_\tau) = 0 \quad \forall t, \tau = 1, \dots, T}$ è chiamata

CONDIZIONE DI ESOGENEITÀ FORTE (o stretta)

(è (A1) molto forte che su S.S. economiche spesso salta)

⇒ X_t è detto STRETTAMENTE ESOGENO (A1) (rispetto)

▲ CONDIZIONE DI ESOGENEITÀ (DEBOLE)
$cov(u_t, X_{t-s}) = 0 \quad \forall s \geq 0$
cioè l'errore corrente non è correlato con il valore correnti e passati di X

↳ X_t è debolmente ESOGENO. (non sappiamo se valori futuri sono correlati con u_t)

▲ X_t è REGRESSORE ENDOGENO se $cov(u_t, X_t) \neq 0$

⇒ Nei modelli deve capire da valutazione del fenomeno il tipo di regressori e ...

OSS (1) ESOGENEITÀ FORTE ⇒ OLS CORRETTO E CONSISTENTE
(2) ESOGENEITÀ DEBOLE ⇒ OLS DISTORTO E CONSISTENTE
(3) ENDOGENEITÀ ⇒ OLS DISTORTO E INCONSISTENTE
(come cross section)

OSS NEI MODELLI DINAMICI (es. AR(1)) NON VALE ESOGENEITÀ

$AR(1) \quad Y_t = \beta_0 + \beta_1 \overset{=Y_{t-1}}{X_t} + u_t$
⇒ $cov(u_t, Y_{t-s}) = 0 \quad s > 0$, ma $cov(u_t, Y_t) \neq 0$
e $cov(u_t, Y_{t+1}) \neq 0$ ..

⇒ NEL MODELLO AR(1) IL REGRESSORE $X_t = Y_{t-1}$ è SOLO DEBOLMENTE ESOGENO (vale per AR(p))
⇒ OLS DISTORTI, MA CONSISTENTI (regressori non sono endogeni)

(b) Per capire se $\exists$ esogeneità forte o debole [illegible] verità.

(a) $Y_t=\beta_0+\beta_1 X_t+u_t$, u_t WN.

In econometria quando X_t ha effetti su Y_t e Y_t ha effetti su X_t (simultanea causalità / interdipendenza)

cioè $\begin{cases} Y_t=\beta_0+\beta_1 X_t+u_t \\ X_t=\gamma_0+\gamma_1 Y_t+v_t \end{cases}$ FEEDBACK

$\Rightarrow cov(u_t,X_t)=\overset{\neq 0}{\gamma_1}\,\overset{\neq 0}{cov(u_t,Y_t)}=cov(u_t,v_t)\neq 0$

$\Rightarrow X_t$ è un regressore endogeno

(L1 cap 3 (simultanea causalità), 12)

(c) Se c'è feedback, ma feedback dinamico.

$Y_t=\beta_0+\beta_1 X_t+u_t$, u_t WN

$X_t=\gamma_0+\gamma_1 Y_{t-1}+v_t$

$\Rightarrow cov(X_t,u_t)=0$ ma $cov(X_{t+1},u_t)\neq 0$

(dim) $cov(X_{t+1},u_t)=\overset{\neq 0}{\gamma_1}\,\overset{\neq 0}{cov(Y_t,u_t)}+\overset{=0\text{ perché WN}}{cov(v_{t+1},u_t)}$

$\Rightarrow X_t$ è debolmente esogeno. (il futuro della variabile è correlato con u_t, ma non valori correnti e passati)

- STIMARE (un modo molto calssico) GLI EFFETTI CAUSALI DINAMICI.
Vogliamo dati $\{Y_t\}, \{X_t\}$ capire effetto statico
- $= \frac{\partial Y_t}{\partial X_t}$ ed effetto dinamico $\frac{\partial Y_{t+h}}{\partial X_t}$
cioè effetti di variazione di X_t su Y corrente o futura. Modelli Dinamici ci permettono di guardare effetti dinamici ⇒ UTILE in macroeconomia (es effetto politica economica)
oss Noi vogliamo stimare causa-effetto ⇒ X_t ⇒ per interpretare in questo modo X_t DEVE ESSERE ESOGENA.
- Se fosse una variabile di controllo ⇒ X_t STRET. ESOGENA (non dipende da nulla, mente ovvero ⇒ incorrellata con gli errori) ⇒ quasi come esperimento controllato nel tempo
⇒ (AR) $\{X_t\}$ ESOGENA (Almeno debolmente)
⇒ Vogliamo stimare effetti dinamici
$\frac{\partial Y_{t+h}}{\partial X_t} = \frac{\partial Y_t}{\partial X_{t-h}}$ (MOLTIPLICATORI)
- ⇒ IL MODELLO (completo) CHE CI DICE CIÓ CHE CI INTERESSA È DL
DL(∞) $Y_t = \beta_0 + \delta_0 X_t + \delta_1 X_{t-1} + \delta_2 X_{t-2} + \dots + u_t$
Y_t è spiegata da valori correnti e passati X_t.
⇒ DL(∞) : $Y_t = \beta_0 + \sum_{j=0}^{+\infty} \delta_j X_{t-j} + u_t$
- $Y_t = \beta_0 + \delta(L) \cdot X_t + u_t$
⇒ $\delta(L) = \delta_0 + \delta_1 L + \delta_2 L^2 + \dots$
Questo è modello teorico di riferimento, ma non è stimabile (∞ parametri) e poi pensiamo che dopo
- un po' il passato non è così influente allora approx DL(∞) con DL(q)

$DL(q)$: $Y_t = \beta_0 + \delta_0 X_t + \dots + \delta_q X_{t-q} + u_t$

$\Rightarrow Y_t = \beta_0 + \delta_q(L) X_t + u_t$ dove $\delta(L) = \delta_0 + \delta_1 L + \delta_2 L^2 + \dots + \delta_q L^q$

OSS I MOLTIPLICATORI SONO I COEF. $\delta_1, \dots, \delta_q$

ma per leggere $DL(q)$ come modello di regressione dobbiamo mettere HP su errore

HP ① $E[u_t \mid \overbrace{X_t, X_{t-1}, \dots}^{F_t}] = 0$

(ma modello $\delta_{q+i} \cdot X_{t-(q+i)}$ con $\delta_{q+i} = 0 \quad \forall i \geq 1$)

X_t e tutti i suoi ritardi sono VARIABILI DEBOLMENTE ESOGENE

$\Rightarrow$ OLS DISTORTO, MA CONSISTENTE

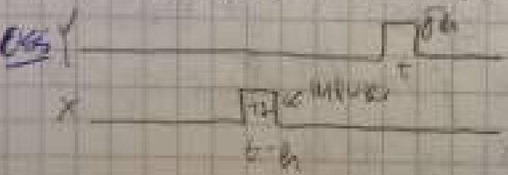
$\Rightarrow E[Y_t \mid F_t] = \beta_0 + \delta_q(L) X_t$

$\Rightarrow \dfrac{\partial E[Y_t \mid F_t]}{\partial X_{t-h}} = \delta_h \qquad h = 0, 1, \dots, q$

effetto parziale della variazione $\Delta X_{t-h} = +1$ CETERIS PARIBUS di quanto in media è variata la Y_t

OSS (▲ $\Delta X_{t-h} = +1$ è detto IMPULSO UNITARIO in $t-h$
$\Delta X_{t-s} = 0 \quad \forall s \neq h$)

OSS



EFFETTO IMPULSO SOLO DOPO h PERIODI
IL RESTO = 0

▲ δ_0 = MOLTIPLICATORE DI IMPATTO

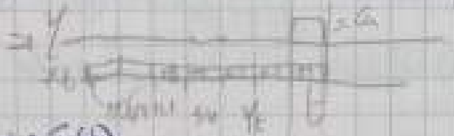
▲ $\delta_1, \dots, \delta_h$ = MOLTIPLICATORI RITARDATI (o ad interim)

OSS l'effetto su Y se c'è variazione su X_t $\forall t$ è la somma dei vari moltiplicatori (in termini attesi) cioè $\Delta E[Y_t | F_t] = \delta_0 \cdot \Delta X_t + \delta_1 \cdot \Delta X_{t-1} + \delta_2 \cdot \Delta X_{t-2} + \dots + \delta_q \cdot \Delta X_{t-q}$ = effetto complessivo è somma dei singoli effetti (perché Modello è lineare).

⇒ Qui esperimento non è più effetto di un impulso, ma mantiene la variazione per + periodi

es DL(4)

⇒ aggiungessimo variazioni precedenti ⇒ effetto è comunque c_4; $c_4 = c_5 = c_6 = \dots$ perché Modello DL(4) ⇒ dopo un po' effetto va a zero



⇒ $c_q = (\delta_0 + \dots + \delta_q)$ = MOLTIPLICATORE CUMULATO è EFFETTO DELLA VARIAZIONE NEL LUNGO PERIODO. = MOLTIPLICATORE CUMULATO DI LUNGO PERIODO = effetto asintotico di una variazione unitaria della X_t $\forall t$ su Y (sul punto futuro della Y).

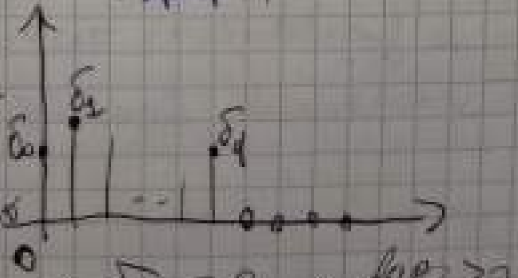
⇒ MOLTIPLICATORI IMPATTO e RITARDATI ci dicono dinamica nel periodo.

⇒ MOLTIPLICATORE DI LUNGO PERIODO: ⇒ Statica COMPARATA vedo se ΔX oggi quando Y non si muove più ⇒ nuovo equilibrio.

▲ FUNZIONE DI RISPOSTA ALL'IMPULSO (I.R.F.)

è grafico dei moltiplicatori = profilo degli effetti dinamici

è interessante per capire fenomeno



es se reazione non è immediata ⇒ $\delta_0 = 0$ e dopo >0

⇒ fa capire effetto nel tempo del ΔX su Y

OSS IRF - per un IMPULSO ISOLATO, ma si può fare anche per impulsi costanti nel tempo

OSS Altrimenti se ignoti ⇒ useremo stime OLS.

(LIBRO vede es. [illegible] X_t [illegible]) [illegible]

$Sh(P_t) = f(\Delta X_t, X_{t-1}, \ldots)$ → Modello DL

⇒ il [illegible] è [illegible] [illegible] [illegible] [illegible]

4/11/15

• EFFETTO DI LUNGO PERIODO?

DL(q): $Y_t = \beta_0 + \delta(L)X_t + u_t$

$c_h = \delta_0 + \delta_1 + \dots + \delta_h \qquad 0 \le h \le q$

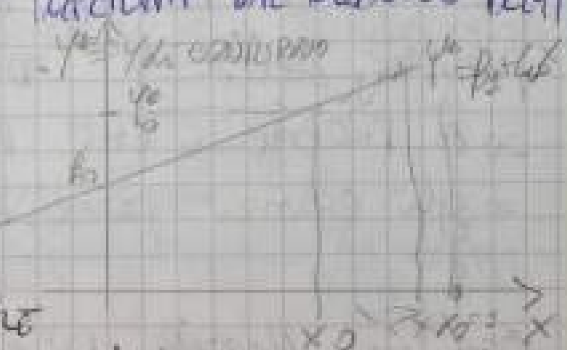
$c_q = \delta_0 + \delta_1 + \dots + \delta_q = \delta(1)$

$\Rightarrow c_\infty = c_q$ = EFFETTO ASINTOTICO/DI LUNGO PERIODO

$\Rightarrow$ se shock u_t fossero nulli $\Rightarrow$ esisterebbe relazione statica deterministica di lungo periodo. se la X non varia $\Rightarrow$

se $u_t = 0 \ \forall t$, $X_t = X \ \forall x$ $\Rightarrow$ RELAZIONE STATICA DI LUNGO PERIODO $Y^* = \beta_0 + (\delta_0 + \delta_1 + \dots + \delta_q)X$ = RELAZIONE STATICA

LINEARE DI L.P. (DI EQUILIBRIO) DETERMINISTICA IMPLICATA DAL MODELLO DL(q)



la retta rappresenta tutti i possibili equilibri del sistema dinamico (in un mondo deterministico no shock)

LA RETTA È DETTA ATTRATTORE perché se mi sposto il sistema dinamico tende a riportarsi in equilibrio sulla retta = ATTRATTORE

oss se si verifica variazione permanente $X_0 \to X_0 + 1$ (es cambia il clima) $\Rightarrow$ il sistema dinamico porta a nuovo equilibrio $Y^* = Y_0^* + \delta_0$

$\Rightarrow c_\infty = \frac{\Delta Y^*}{\Delta X}$ = MOLTIPLICATORE ASINTOTICO $\Rightarrow$ la variazione dei valori di equilibrio (asintotica) a fronte del ΔX unitario. (in generale variazione è $\Delta Y^* = c_\infty \cdot \Delta x$)

• INFERENZA SUI PARAMETRI DEL MODELLO DL [illegible]

$Y_t = \beta_0 + \delta_0 X_t + \dots + \delta_q X_{t-q} + u_t$ (HP)

① $E[u_t | X_t, \dots] = 0$

② (Y_t, X_t) congiuntamente stat e asintoticamente indipendenti

③ $0 < E[Y^{8+\varepsilon}] < +\infty$ con $\varepsilon > 0$

sia X sia Y hanno più di 8 momenti finiti

④ ASSENZA DI PERFETTA COLLINEARITÀ

OSS (HP) ③ non serve per stima OLS, ma è (HP) tecnica ..

OSS Modello DL(q) non AR-DL ⇒ non entra in gioco passato della Y ⇒

⇒ È PROBABILE CHE u_t e u_{t-h} SONO CORRELATI TRA LORO

⇒ gli errori possono essere serialmente correlati

(N.B. stesso SW su errori pag OSS DL non ADL)

OSS Se Modello DL(q) ma in verità Y_t dipende anche da Y_{t-1} ⇒ Y_{t-1} è variabile rilevante

⇒ Y_{t-1} sta dentro u_t ⇒ x Y_{t-1} è correlato con u_t ⇒ analogamente Y_{t-2} sta dentro u_{t-1}

⇒ se $corr(Y_{t-1}, Y_{t-2}) \neq 0$ ⇒

⇒ $cov(u_t, u_{t-1}) \neq 0$

⇒ C'È CORRELAZIONE SERIALE DEGLI ERRORI

OSS sappiamo che MODELLO ADL CON u_t AUTOCORRELATI ⇒ OLS DISTORTI e INCONSISTENTI

Ma x modello DL con u_t AUTOCORRELATI

⇒ OLS DISTORTI, MA CONSISTENTI e ASINTOTICAMENTE NORMALI

(perché X debolmente esogeno)

⇒ IMPORTANTE CHE MODELLI DL SIANO CORRETTAMENTE

SPECIFICATO, metto tutti i q ritardi rilevanti della x
⇒ anche se gli errori sono correlati, non è un
problema in grandi campioni

OSS PROBLEMA: in tal caso la stima della matrice di
VAR/COV degli stimatori che useremo per inferenza sui
parametri (es T-test...)
$\widehat{VAR}(\hat{\beta}) \doteq \hat{\Sigma}_{\hat{\beta}}$ ⇒ STIMATORE CLASSICO E DI WHITE
sono inconsistenti perché c'è autocorrelazione negli
errori (PROBLEMA NUOVO)

⇒ DEVO USARE STIMATORE CONSISTENTE SIA AD ETEROSCHE-
DASTICITÀ SIA AD AUTOCORRELAZIONE.

- STIMATORE DI NEWEY-WEST $\hat{\Sigma}_{\hat{\beta}, NW} \xrightarrow[T\to\infty]{P} \Sigma_{\hat{\beta}}$
è consistente sia in caso di eteroschedasticità sia
di autocorrelazione degli errori u_t.
(HAC = Heteroschedasticity-Autocorrelation consistent)
SOTTO HP③ $0 < E[Y^{8+h}] < +\infty$
OSS STIMATORE RICHIEDE DI SPECIFICARE PARAMETRO
m = Parametro di troncamento
- IDEA: dovrei stimare $\hat{\rho}_{\hat{u}}(1), ..., \hat{\rho}_{\hat{u}}(m-1)$ cioè considerare
i valori della ACF residui, dovrei considerarli
tutti, ma in pratica fix m e considero i primi
m ritardi e m dovrebbe essere scelto ~~xxx~~ in
base alla grandezza del campione:
$(m = 0{,}75 \cdot T^{1/3}$ No studio formula)
T = grandezza campione)

- ⇒ NW È HAC

(LIBRO 1 S-W)
NO FORMULE MA IDEA E SCELTA
NW

OSS Se errori fanno WN ⇒ stimatore NW ≤ BETTER
⇒ lo generalizza, ma robusto ad ANCORE.

OSS Si può usare Modello DL(q) → per stimare effetti causali dinamici
• Ma posso usare ARDL con errori WN ⇒ per STIMARE GLI EFFETTI CAUSALI DINAMICI (come prima) ?
SI SE LA VARIABILE X_t È FORTEMENTE ESOGENA.
(≠ DL in cui basta DEBOLMENTE ESOGENA)

• Modello ARDL(p,q) qui non per previsione ⇒ X_t entra
$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \dots \beta_p Y_{t-p} + \delta_0 X_t + \dots + \delta_q X_{t-q} + U_t$$
⇒ Modello si può stimare con OLS se X_t è debolmente ESOGENA (Y_{t-1}, lo è per forza)
⇒ SOTTO DEBOLE ESOGENEITÀ ⇒ OLS DISTORTI, MA CONSISTENTI
Ma non troviamo subito MOLTIPLICATORI perché X_{t-1} influisce non su Y_t ma su Y_{t-1} ⇒
$$\Rightarrow \frac{\partial E[Y_t|\dots]}{\partial X_t} = \delta_0 \quad \text{ma} \quad \frac{\partial E[Y_t|\dots]}{\partial X_{t-1}} \neq \delta_1$$
(ma $= \delta_0\beta_1 + \delta_1$)
ma moltiplicatori ignoti ⇒ STIMA OLS BASTA DEBOLE ESOGENEITÀ, MA SE VOGLIAMO INTERPRETARE MOLTIPLICATORI COME EFFETTI CAUSALI DINAMICI serve (H) + FORTE
X FORTEMENTE ESOGENA (vedi dopo)

(2) DL(0), $\beta_0 = 0$ $\quad Y_t = \delta_0 X_t + u_t$

(1) $E[u_t | X_t, X_{t-1}, \dots] = 0$

$u_t \sim AR(1)$ stat.

$\Rightarrow u_t = 0 + \rho u_{t-1} + \varepsilon_t$, $\varepsilon_t \sim WN(0, \sigma^2_\varepsilon)$

(1.b) modello DL con errori autocorrelati non modello ARMA $\Rightarrow$ GLS

$u_t (1-\rho L) = \varepsilon_t$

$u_t = (1-\rho L)^{-1} \varepsilon_t$

$Y_t = \delta_0 X_t + u_t = \delta_0 X_t + (1-\rho L)^{-1} \varepsilon_t$

$\Rightarrow (1-\rho L) Y_t = (1-\rho L) \delta_0 X_t + \varepsilon_t$, $\varepsilon_t \sim WN$

$\Rightarrow$ ARDL(1,1) con X_t corrente)

$\Rightarrow Y_t = \rho \underline{Y_{t-1}} + \delta_0 \underline{X_t} - \rho \delta_0 \underline{X_{t-1}} + \varepsilon_t$

$\Rightarrow$ è STABILE (condizione di stabilità ARDL è che parte AR sia invertibile)

$\Rightarrow$ ho trasformato DL(0) in ARDL(1,1) stabile.

$\Rightarrow$ Qui non ha senso DL è + parsimonioso, ma in generale passare da DL a ARDL potrebbe anche portare a modello più parsimonioso.

(4) Se $X_t \sim$ STRETT. ESOGENA

$\Rightarrow$ OLS DISTORTO ma CONSISTENTE e ASINTOTICAMENTE NORMALE

$\Rightarrow$ STIMATORE DI WHITE $\hat{\Sigma}_\beta$.

OSS DL $\to$ ARDL $\Rightarrow$

$\Rightarrow$ torniamo indietro ma ci perdiamo δ_1:

Modello ARDL $\quad Y_t = \rho Y_{t-1} + \delta_0 X_t + \delta_1 X_{t-1} + \varepsilon_t$

$\Rightarrow$ per stimare effetti causali DINAMICI torno a modello DL (come facevamo con polinomio autoregressivo in ARMA) OSS ARDL(p,q) = DL(∞)

$\Rightarrow$ Riscrivo $(1-\rho L)Y_t = (\delta_0+\delta_1 L)X_t + \varepsilon_t$

$\Rightarrow$ Invertendo $Y_t = \underbrace{(1-\rho L)^{-1}(\delta_0+\delta_1 L)}_{\text{polinomio di ordine } \infty} X_t + (1-\rho L)^{-1}\varepsilon_t$

(perché $(1-\rho L)^{-1} = (1+\rho L+\rho^2L^2+\rho^3L^3+\dots)$)

$\Rightarrow (1-\rho L)^{-1}(\delta_0+\delta_1 L) \doteq \varphi_\infty(L)$

$\Rightarrow Y_t = \varphi_\infty(L)X_t + u_t =$ (DL(∞))

$= \sum_{j=0}^{+\infty} \varphi_j X_{t-j} + u_t$

$\Rightarrow$ Stimo effetti causali dinamici:

$\Rightarrow \dfrac{\partial E[Y_t|\dots]}{\partial X_{t-h}} = \varphi_h$ = MOLTIPLICATORI

$\Rightarrow$ ricavo moltiplicatori da

$(1+\rho L+\rho^2L^2+\rho^3L^3+\dots)(\delta_0+\delta_1 L) = \delta_0+\delta_1 L+\rho\delta_0 L+\rho\delta_1 L^2+$
$+\delta_0\rho^2L^2+\dots = \underbrace{\delta_0}_{\varphi_0} + \underbrace{(\delta_1+\rho\delta_0)}_{\varphi_1}L + \underbrace{(\rho\delta_1+\rho^2\delta_0)}_{\varphi_2}L^2+\dots$

$\Rightarrow \varphi_0 = \delta_0,\ \varphi_1 = \delta_1+\rho\delta_0,\ \varphi_2 = \rho\delta_1+\rho^2\delta_0, \dots$

OSS è partiamo da DL(0) in cui δ_0 $\delta_h = 0\ \forall h>0$
$\Rightarrow$ qui ho rappresentazione $\neq$, ma formulazioni
equivalenti $\Rightarrow \varphi_1=\varphi_2=\dots=0$

IDEA:
OSS È SEMPRE POSSIBILE SCRIVERE UN MODELLO DL(q)
CON ERRORI AUTOCORRELATI AD UN PROC. ARDL(p, q_2)
CON ERRORI WN, MA SERVE SAPERE IL PROCESSO
AR(p) CHE GUIDA GLI ERRORI u_t (stazionario)
$\Rightarrow$ DA MODELLO ARDL(p, q) CON ERRORI WN POSSIAMO
PASSARE AD UN MODELLO DL($+\infty$) E DAI SUOI
COEFFICIENTI abbiamo φ_j = MOLTIPLICATORI PER
GLI EFFETTI CAUSALI.

OSS ovviamente nel passare da una rappresentazione all'altra entrano in gioco vincoli non lineari sul modello ARDL
⇒ se consideriamo vincoli (es $\delta_1 = -\beta\delta_0$)
⇒ LE DUE RAPPRESENTAZIONI DL(q) e DL(∞) COINCIDONO

OSS I coef. dei modelli sono ignoti ⇒ stima e i dati dovrebbe fare sì che vincoli escano fuori.

IDEA APPLICAZIONI:
ho dati ⇒ faccio modello ARDL di cui non so ordine ⇒ stimo dai dati ⇒ ricavo rappresentazione DL(∞) e vedo quali coef. sono sig. e vincoli

N.B. I MOLTIPLICATORI DEL MODELLO DL(∞) POSSONO ESSERE INTERPRETATI COME EFFETTI CAUSALI SOLO SE X È FORTEMENTE ESOGENA

OSS invece di questo casino posso fare direttamente stima DL(q), ma ARDL può essere più parsimonioso e portare a stessi risultati.

VEDI L1 S-W

ARDL (p, q) X È STRETT. ESOGENA.

$\Rightarrow \theta(L) Y_t = \beta_0 + \delta(L) X_t + \varepsilon_t \qquad \varepsilon_t \sim WN$

(con X_t strett. esogena)

⇒ STABILE $\Rightarrow \exists!\ \theta^{-1}(L)$

$$Y_t = \frac{\beta_0}{\theta(1)} + \theta(L)^{-1} \cdot \delta(L) X_t + \theta(L)^{-1} \cdot \varepsilon_t \qquad \text{RAPPRESENTAZIONE DL}(\infty)$$

$$\Rightarrow Y_t = \frac{\beta_0}{\theta(L)} + \varphi_{\infty}(L) X_t + \dots = DL(\infty)$$

$$\Rightarrow c_1 = \varphi_0 + \varphi_1, \quad c_2 = \varphi_0 + \varphi_1 + \varphi_2, \dots$$

⇒ POICHÉ C'È STABILITÀ ⇒ ∃ FINITO MOLTIPLICATORE CUMULATO ASINTOTICO.

$$C_\infty = \sum_{i=0}^{\infty} \varphi_{ei} = \varphi_\infty(1) = \frac{\alpha(1)}{\theta(1)} < +\infty \qquad C_{T\infty} = \beta_{T\infty}$$ = PENDENZA ASINTOTICA DELLA RETTA.

si ricava da modello ADL

$$Y^* = \frac{\beta_0}{\theta(1)} + \beta_{T\infty} X \quad \Rightarrow \Delta Y^* = \beta_{T\infty} \Delta X$$

⇒ $\beta_{T\infty}$ ci dice di quanto si sposta valore di equilibrio della Y rispetto a variazione della X (statica comparata)

AR 15 TUTTO E PUNTO CHE 16