



UNIVERSITA' DEGLI STUDI DI PADOVA
DIPARTIMENTO DI SCIENZE ECONOMICHE ED AZIENDALI
"M.FANNO"

DIPARTIMENTO DI SCIENZE STATISTICHE

CORSO DI LAUREA IN ECONOMIA

PROVA FINALE

**"ALCUNI RECENTI SVILUPPI NELL'ANALISI DELLE SERIE
STORICHE"**

RELATORE:

PROF.SSA BISAGLIA LUISA

LAUREANDO: IAVARONE LUCA

MATRICOLA N. 1115794

ANNO ACCADEMICO 2017 – 2018

Alla mia famiglia

Indice

INTRODUZIONE

p. 3

PARTE PRIMA

1. Concetti base delle serie storiche

1.1 <i>Definizione di serie storica</i>	5
1.2 <i>Approccio classico alle serie storiche</i>	6
1.3 <i>Serie storiche e processi stocastici</i>	11
1.4 <i>La stazionarietà</i>	12
1.5 <i>Il rasoio di Occam</i>	14

2. Analisi univariata delle serie storiche

2.1 <i>L'operatore ritardo</i>	16
2.2 <i>Processi white noise</i>	17
2.3 <i>Processi a media mobile (MA)</i>	17
2.4 <i>Processi autoregressivi (AR)</i>	18
2.5 <i>Processi a radice unitaria (random walk)</i>	21
2.6 <i>Processi ARMA</i>	22
2.7 <i>Integrazione e processi ARIMA</i>	23

3. Analisi multivariata delle serie storiche

3.1 <i>Processi VAR</i>	28
1.2 <i>Stabilità e stazionarietà dei processi VAR</i>	29
1.3 <i>Processi VMA</i>	30
1.4 <i>Cointegrazione e processi VEC</i>	31
1.5 <i>Cointegrazione: approccio di Engle-Granger e caso bivariato</i>	34

PARTE SECONDA

4. Introduzione all'analisi strutturale

4.1 *Introduzione storica ai modelli SVAR* 43

4.2 *Causalità di Granger e causalità istanea* 45

5. Modelli SVAR

5.1 *Definizione dei processi SVAR* 48

5.2 *Il problema di identificazione* 49

5.3 *Funzione di risposta all'impulso* 52

5.4 *Un'applicazione SVAR* 53

5.5 *Conclusioni sui modelli SVAR* 63

PARTE TERZA

Conclusioni 65

Appendici 66

Bibliografia 80

Introduzione

Il presente lavoro si occupa di serie storiche, ed in particolare dei modelli statistici *Vector Autoregressive*, utilizzati nell'econometria applicata. Esso è strutturato come segue:

- Nella “Parte Prima” sono esposti i concetti fondamentali e i modelli impiegati per la comprensione e l'analisi delle serie storiche in senso economico. Specificatamente, essa si articola in tre capitoli: nel corso del primo capitolo, si descrivono le serie storiche in quanto processi stocastici, scandagliandone le componenti fisiologiche e definendo le strutture analitiche che costituiscono l'approccio metodologico adottato; nel secondo capitolo, si inquadrano tali processi stocastici secondo le forme e le relazioni statistico-matematiche nella dimensione univariata; nel terzo capitolo, le procedure individuate si generalizzano al caso multivariato.
- Nella “Parte Seconda”, si offre una disamina dell'analisi strutturale possibile con gli strumenti illustrati: il quarto capitolo è incentrato sul dibattito storico che ha portato all'adozione dei nuovi strumenti statistici illustrati nella Prima Parte per il *policy-making process*, specificando le basi teoriche cui essi poggiano; nel quinto capitolo si evidenziano approfonditamente le potenzialità del loro impiego nell'indagine macroeconomica.
- Nella “Parte Terza”, si conclude riguardo i risultati derivati. In appendice si riportano, in ordine: approfondimenti riguardo le nozioni di stazionarietà in sistemi dinamici ed interpretazione della funzione di risposta all'impulso con i numeri complessi; le stime con il codice R utilizzato.

Parte prima

1. Concetti base delle serie storiche¹

1.1 Definizione di serie storica

La statistica inferenziale si preoccupa di analizzare ed elaborare i dati, mediante la strumentazione tecnica di cui essa dispone. I dati raccolti possono configurarsi come dati *cross-section*, con cui si intendono osservazioni condotte su una pluralità di unità statistiche, oppure come *serie storiche*, ovvero osservazioni su una o più grandezze, protratte nel tempo. Utilizzando le parole di Piccolo (2000), una **serie** è “una distribuzione di frequenza organizzata rispetto ad un criterio qualitativo”: configurandosi come fenomeni che seguono un’evoluzione non deterministica, il modello probabilistico di riferimento risulta essere il processo stocastico.

Definizione 1.1. Un **processo stocastico** si definisce come una successione di variabili casuali

$$\{Y_t: -\infty < t < +\infty\}$$

ordinate nel tempo, con relazioni di dipendenza interne arbitrarie. Un esperimento conduce ad una particolare successione numerica $\{y_t: -\infty < t < +\infty\}$, detta *traiettoria* del processo.

▪

Nell’analisi delle serie storiche, il fenomeno da prevedere viene trattato come una sorta di *black box*, adottando un approccio che cerca di ricostruire il “sentiero” passato del fenomeno per ottenerne la previsione, modellandolo rispetto al tempo. La previsione viene realizzata con l’assunzione che le regolarità di comportamento osservate si mantengano in futuro.

Una serie storica che contiene i valori passati di una sola variabile si definisce *univariata*; se tali valori si riferiscono a più di una variabile, si definisce *multivariata*.

Una serie storica, inoltre, può essere *continua* o *discreta*. In una serie storica con tempo continuo, le osservazioni sono prese in ogni istante; con tempo discreto, in precisi istanti. La misura della temperatura durante il giorno, ad esempio, sarà continua; la produzione di un’impresa o la crescita della popolazione sono misurate a intervalli regolari. Una serie continua può essere facilmente trasformata in discreta raggruppando i dati in un arco temporale più ampio.

¹Per il realizzo di questa parte, si è scelta l’impostazione di Adhikari e Agrawal (2013); si ringraziano Davidson et al. (1978), Grazzini (2012), Hipel, McLeod (1994) e Piccolo (2000) per importanti contributi ripresi nell’esposizione. Le rappresentazioni grafiche realizzate con R seguono la traccia di Shumway e Stoffer (2018). La figura 1.1 è tratta da Walter (2008).

1.2 Approccio classico alle serie storiche

Una serie storica può considerarsi la risultante di quattro componenti:

- *Componente stagionale*: se presente, la serie è influenzata da componenti stagionali. Ad esempio, una serie che raffigura la vendita di gelati durante l'anno è chiaramente influenzata dalla stagionalità. Tali serie sono anche dette periodiche, poiché il ciclo stagionale si ripete in un periodo fisso.
- *Componente ciclico*: si osserva in una serie che presenta andamenti in aumento ed in diminuzione non legati ad un periodo fisso, e pertanto diverge dalla stagionalità. In una serie economica, il pattern ciclico è determinato dalle espansioni e contrazioni dell'economia.
- *Trend*: andamento crescente o decrescente di lungo periodo.
- *Componente occasionale*: variazioni accidentali causate da eventi imprevedibili.

Tali componenti non vanno considerate come se fossero comportamenti stagni; al contrario, è molto probabile riscontrare combinazioni delle stesse.

I metodi di scomposizione tendono ad individuare due pattern: la *stagionalità* ed il *trend-ciclo*, che a sua volta può essere scomposto in una componente di fondo, il trend, e una componente dipendente da fenomeni congiunturali, il ciclo.² I metodi di scomposizione non sono utilizzati direttamente nella previsione, ma sono utili nella descrizione dei dati storici a disposizione.

Una serie storica può essere descritta mediante un modello additivo od uno moltiplicativo. Per il modello *additivo* vale:

$$y_t = TC_t + S_t + U_t$$

Mentre per il *moltiplicativo*:

$$y_t = TC_t \cdot S_t \cdot U_t$$

Che si può riscrivere come:

$$\log y_t = \log TC_t + \log S_t + \log U_t$$

² In realtà, l'identificazione e la stima della tendenza da sempre rappresentano un grande problema per gli statistici, perché il trend, come tutte le altre componenti di una serie storica, è una variabile latente (non osservabile) e la definizione di movimento di lungo periodo non è del tutto chiara: esso può rivelarsi parte di un ciclo se l'intervallo delle osservazioni della serie è abbastanza ampio. È per tale motivo che si aggrega il movimento generato come *trend-ciclo*.

Il metodo delle *medie mobili* consente di “lisciare” la serie, eliminando oscillazioni stagionali ed erratiche, calcolando una nuova serie storica in cui il termine relativo ad un determinato tempo è il risultato della media di k termini della serie originaria a loro a loro adiacenti (sottounità stagionali).³

Se k è dispari, ciascuna MA si riferisce al tempo centrale sull’intervallo in cui è stata calcolata.

Ad esempio, con $k = 3$, $MA3(y_t) = \frac{(y_{t-1} + y_t + y_{t+1})}{3}$.

Se k è pari, è necessario calcolare la media di due medie mobili contigue per centrare il valore.

Ad esempio, con $k = 4$, $MA4(y_t)_1 = \frac{(y_{t-2} + y_{t-1} + y_t + y_{t+1})}{4}$; $MA4(y_t)_2 = \frac{(y_{t-1} + y_t + y_{t+1} + y_{t+2})}{4}$.

Quindi, $MA4(y_t) = \frac{MA4(y_t)_1 + MA4(y_t)_2}{2}$.

Applichiamo i concetti esposti con R. Seguiamo l’impostazione di Davidson et al. (1978), con dati diversi. Si mostra la decomposizione di serie storiche nelle diverse componenti. Usiamo come esempi i dataset relativi alla produzione di birra australiana, ed il numero dei passeggeri aerei, già presenti nei pacchetti R.

È importante la scelta tra modello additivo e moltiplicativo. Osservando l’andamento di una serie, ci si può chiedere: l’ampiezza delle fluttuazioni stagionali aumenta all’aumentare dei valori della serie? Se sì, usiamo il modello moltiplicativo; altrimenti, quello additivo.

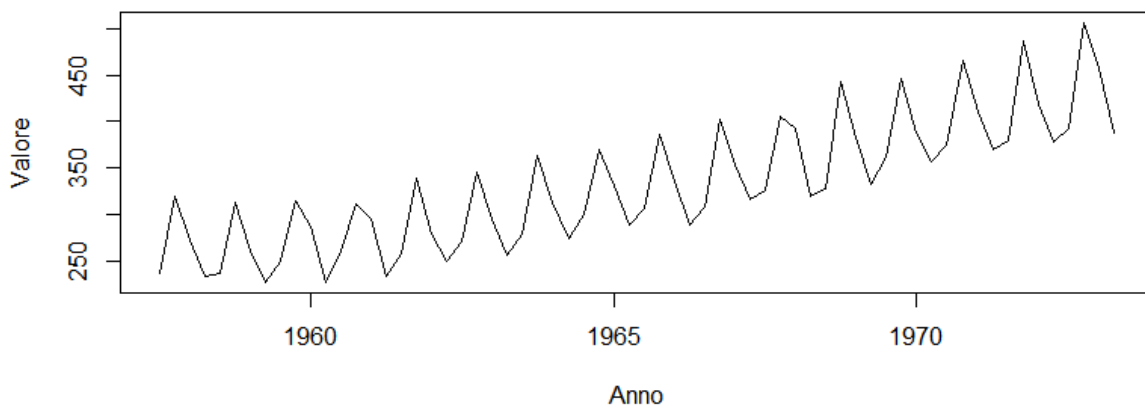


Fig.1.1. Produzione di birra australiana.

³ È bene accertarsi in maniera oggettiva della significativa presenza di stagionalità; in caso contrario, si incorre nel rischio di effetti imprevisti, come il fenomeno Slutsky-Yule: una serie del tutto erratica, se filtrata con medie mobili, può sembrare ben strutturata, scambiando un andamento del tutto casuale per un andamento ciclico.

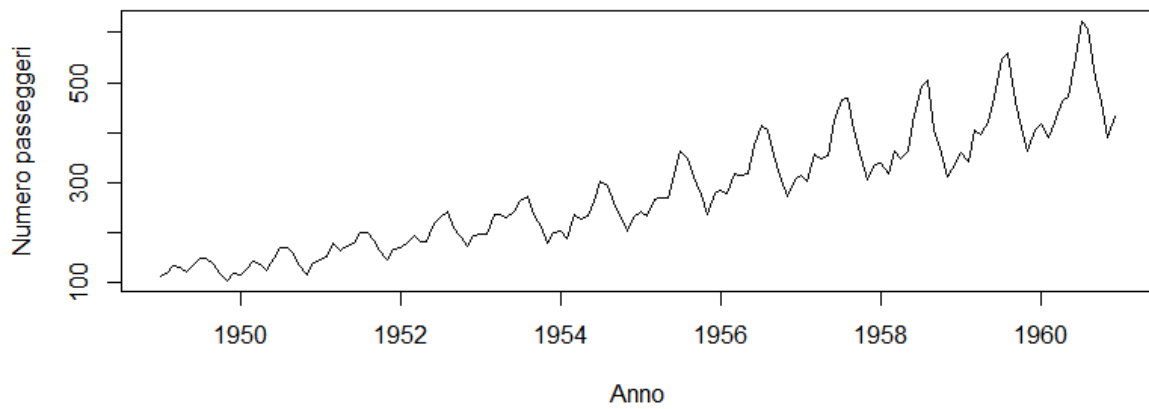


Fig.1.2. Numero di passeggeri aerei.

Per la prima serie, si nota che la variazione stagionale sembra costante e non cambia quando cresce il valore della serie. Quindi, scegliamo il modello additivo.

Nella seconda, invece, le fluttuazioni stagionali sono legate all'incrementare della serie, quindi è opportuno il modello moltiplicativo (interazione tra trend e stagionalità).

Troviamo il trend. Per decomporre una serie, filtriamo con la media mobile: osservando che la produzione di birra australiana e il numero dei passeggeri aerei hanno entrambe una componente stagionale annuale e che le serie hanno, rispettivamente, osservazioni trimestrali e mensili, applichiamo alla serie della birra MA(4), mentre a quella dei passeggeri MA(12).

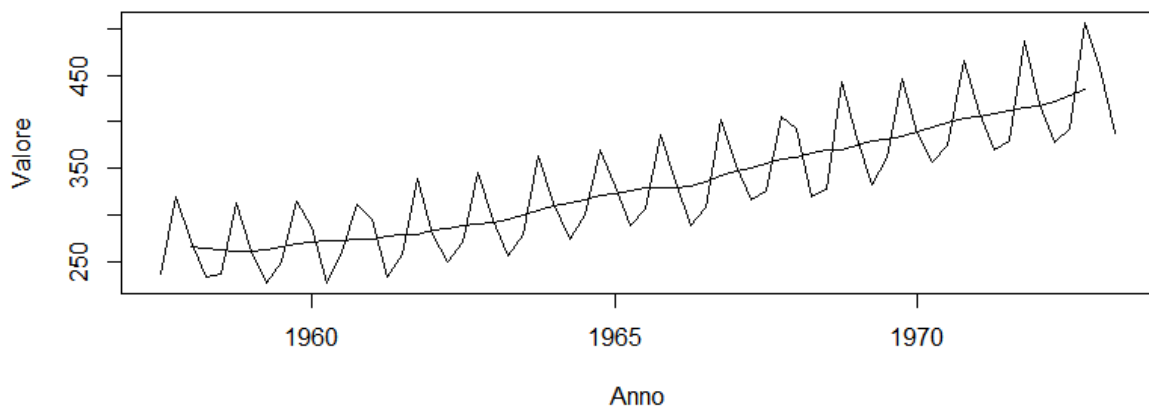


Fig.1.3. Applicazione della media mobile alla produzione di birra australiana. Rappresenta il trend stimato.

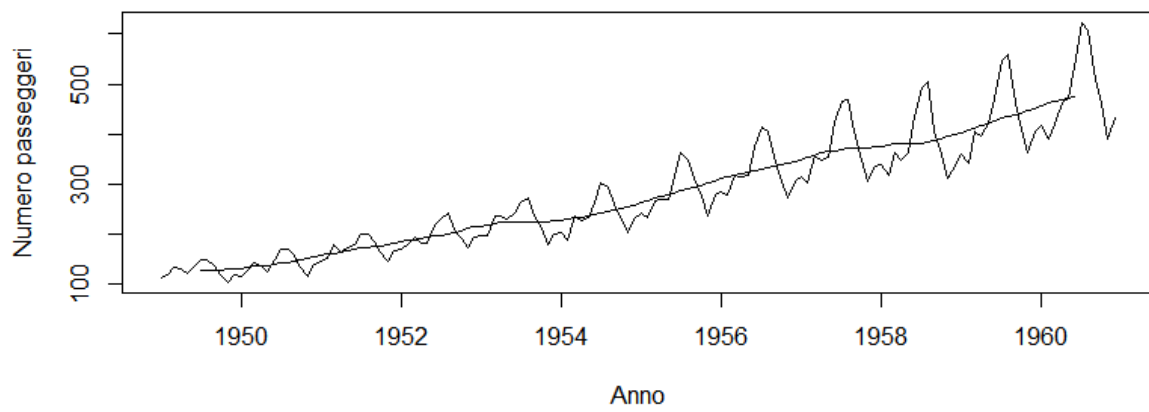


Fig.1.4. Applicazione della media mobile al numero dei passeggeri aerei. Rappresenta il trend stimato.

Detrendizziamo le serie. Rimuovendo il trend, osserveremo solo la stagionalità.

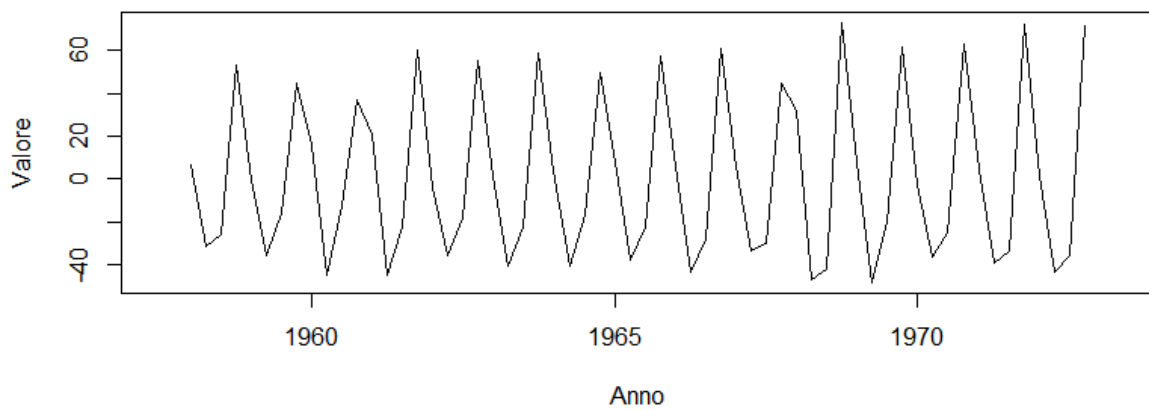


Fig.1.5. Serie detrendizzata, produzione birra australiana.

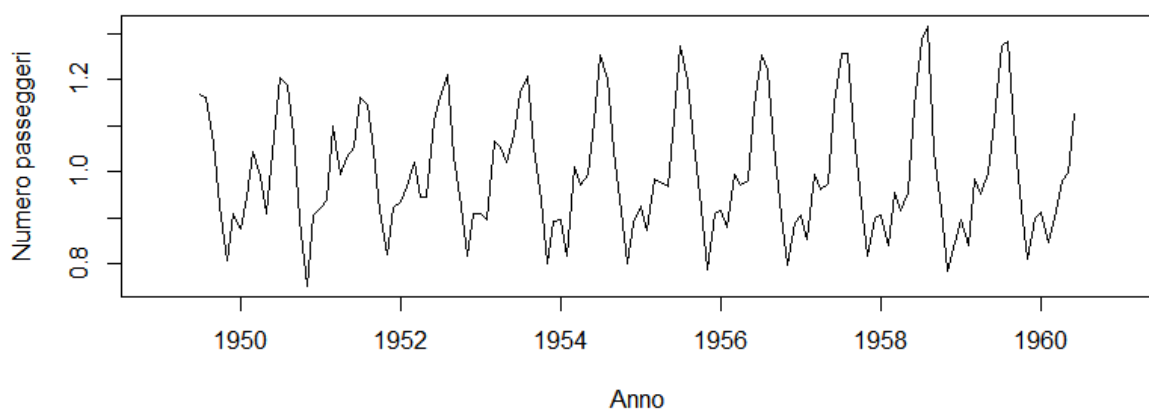


Fig.1.6. Serie detrendizzata, numero passeggeri aerei.

Identifichiamo la stagionalità media. Una volta sottratto il trend dalla serie, si raggruppano i risultati ordinati per mese, o trimestre, e fatta la media.

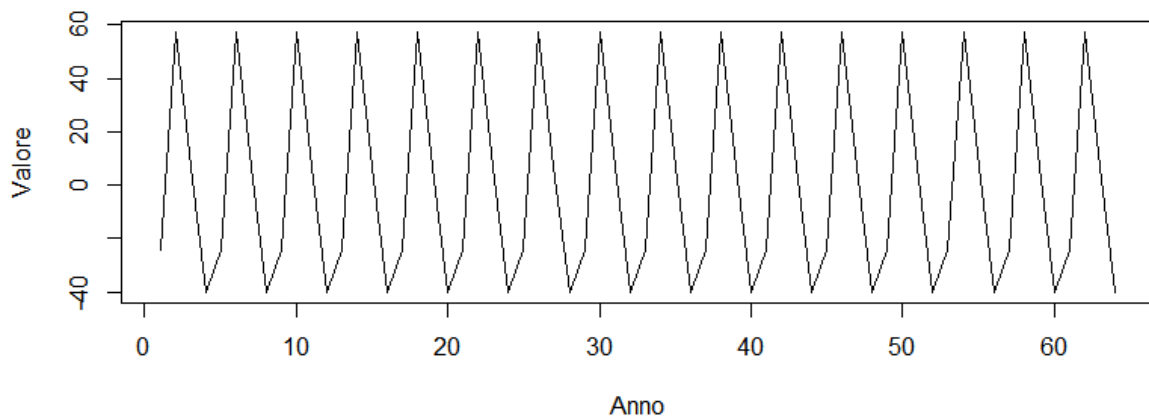


Fig.1.7. Stagionalità media, produzione birra australiana.

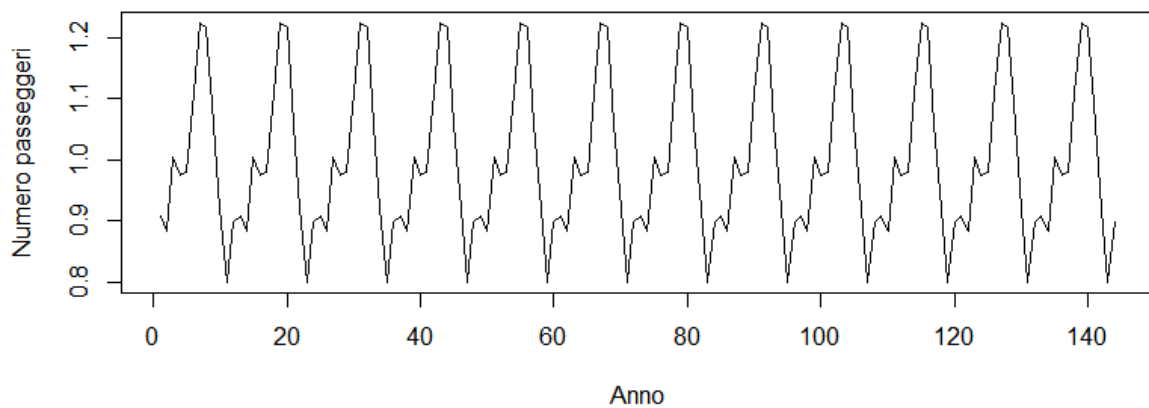


Fig.1.8. Stagionalità media, numero passeggeri aerei.

La componente irregolare? È la parte residua.

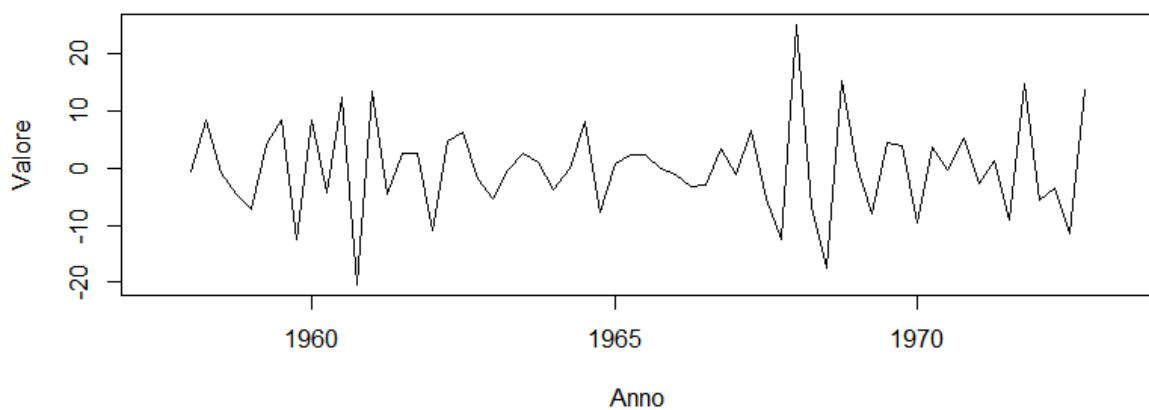


Fig.1.9. Random noise, produzione birra australiana.

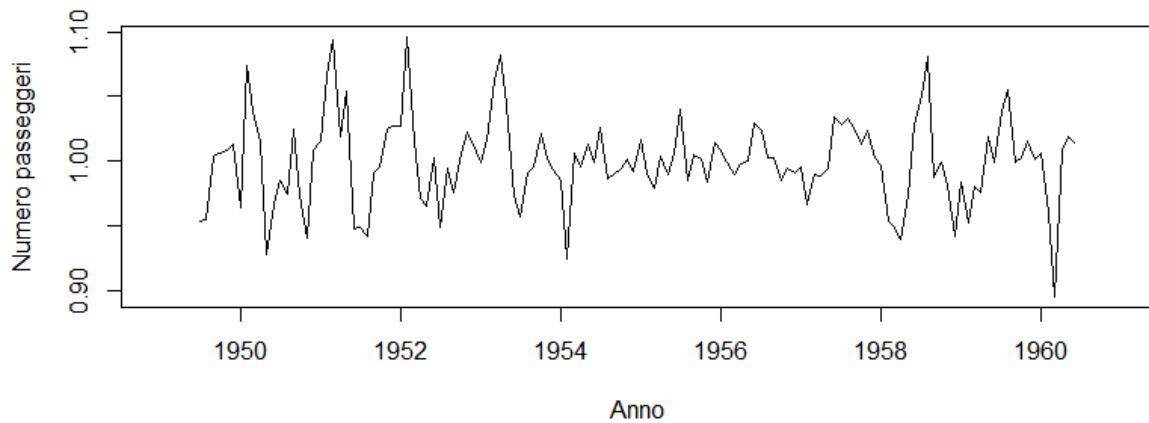


Fig.1.10. Random noise, numero passeggeri aerei.

In conclusione, nell'approccio classico le componenti vengono rappresentate mediante funzioni matematiche note: il trend è una costante oppure una funzione polinomiale; la stagionalità è una ripetizione regolare di effetti ogni s periodi, con s ad indicare la frequenza della serie: $S_t = \sum_{i=1}^s \delta_i D_{it}$, con $\sum_{i=1}^s \delta_i = 0$, $D_{it} = \begin{cases} 1, & \text{stagione } i \\ 0, & \text{altrimenti} \end{cases}$ δ_i i coefficienti che misurano l'effetto. La componente irregolare è l'unica di natura stocastica, solitamente rappresentata come un rumore bianco (*white-noise*), incorrelato con gli errori passati.

Vi sono notevoli inconvenienti nella decomposizione classica delle serie:

- occorre ipotizzare una precisa forma funzionale per le componenti, implicando notevoli dosi di soggettività;
- le componenti sono molto “rigide”: nell'approccio classico si usa un metodo “globale”, con cui ogni osservazione contribuisce in modo analogo alla previsione. Oggi si preferisce stimare i parametri con metodi “locali”: il trend è una somma di una componente deterministica e un disturbo casuale, la cui varianza determinerà di volta in volta un maggiore o minore allontanamento dal trend deterministico; la somma degli effetti stagionali su s periodi non è esattamente pari a 0, ma ad una v.c. con media nulla.

1.3 Serie storiche e processi stocastici

Una serie storica è considerata come una realizzazione finita di un processo stocastico: non possiamo prevedere con certezza cosa accadrà in futuro. Richiamandone la definizione, una sequenza di dati osservati è una particolare traiettoria del processo: tali informazioni vengono

poi sfruttate per comprendere il processo generatore di dati, e quindi costruire un modello che risulti essere efficace nella previsione degli andamenti futuri⁴.

Un'assunzione ricorrente è che le variabili di una serie storica siano *indipendenti e identicamente distribuite* (i.i.d.), seguendo una distribuzione normale. Invero, è interessante osservare che le serie storiche non sono i.i.d.: esse seguono un pattern più o meno regolare nel tempo. Ad esempio, se la temperatura odierna è molto alta, è probabile lo sia anche domani. Ecco perché i risultati delle previsioni di una serie sono spesso vicini ai valori attuali.

1.4 La stazionarietà

Il concetto di stazionarietà di un processo stocastico può essere visto come una forma di “equilibrio statistico”: *i momenti*, come la media e la varianza, *non dipendono dal tempo*. È una condizione necessaria per la costruzione di modelli di previsione, riducendo notevolmente la complessità matematica⁵. Esistono due definizioni di stazionarietà: in senso stretto e in senso debole.

Definizione 1.2. Una serie storica si dice **stazionaria in senso forte** se il comportamento di un insieme di valori:

$$\{x_{t_1}, x_{t_2}, \dots, x_{t_k}\}$$

risulta essere identico allo stesso insieme spostato nel tempo di un intervallo h :

$$\{x_{t_1+h}, x_{t_2+h}, \dots, x_{t_k+h}\}$$

per ogni $k = 1, 2, \dots$, per ogni istante $t_1, t_2, \dots, t_k$, e ogni spostamento $h = 0, \pm 1, \pm 2, \dots$.

■

Per la stazionarietà in senso forte, la funzione di densità congiunta, $F(t_1, t_2, \dots, t_n)$ delle n variabili aleatorie, scelte arbitrariamente, $X(t_1), X(t_2), \dots, X(t_n)$, deve essere invariante per *qualsiasi* traslazione temporale, ossia deve dipendere unicamente dagli n istanti solo per il tramite delle $n - 1$ differenze, per ogni possibile insieme di variabili casuali: $\tau_k \equiv t_{k+1} - t_k$, $k = 1, 2, \dots, n - 1$.

⁴ Con *data generating process* (DGP) si intende propriamente il fenomeno sottostante alla generazione dei dati, mentre con *modelli* si definiscono quelle strutture algoritmiche che cercano di descrivere ed emulare il fenomeno, seppur imperfettamente.

⁵ Si vogliono evitare, essenzialmente, *break strutturali* nei sottocampioni di una serie, che hanno come conseguenza valori dei parametri diversi.

Definizione 1.3. Una serie storica x_t si dice **stazionaria in senso debole** se:

- la funzione del valor medio, μ_t , è costante e non dipende dal tempo t ;
- la funzione di autocovarianza, $\gamma(s, t)$, dipende da s e t solo attraverso la distanza $|s - t|$, definita l'autocovarianza come:

$$\gamma(s, t) = \text{cov}(x_s, x_t) = E[(x_s - \mu_s)(x_t - \mu_t)]$$

▪

Il *valore medio* costituisce il baricentro di una v.c. e mutua tutte le proprietà e caratteristiche della media aritmetica⁶. La *varianza* misura di quanto una variabile si discosti dal suo valore atteso. La *covarianza* di quanto una variabile vari al variare dell'altra, esprimendo la reciproca dipendenza. L'*autocovarianza* è la covarianza tra le v.c. del medesimo processo stocastico spaziate fra di loro di un determinato sfasamento temporale⁷. L'*autocorrelazione* (ACF, *AutoCorrelation Function*) si ottiene dall'autocovarianza, rapportandola agli scarti quadratici medi. In formule: $\rho(t) = \frac{E[(x_s - \mu_s)(x_t - \mu_t)]}{\sigma_{x_s} \sigma_{x_t}}$.

La stazionarietà in senso debole richiede che i momenti secondi non dipendano dal tempo: in tal caso, il processo x_t si dice *stazionario in senso debole*. Poiché essa pone condizioni solamente sui momenti secondi della serie, anziché su tutte le possibili distribuzioni, nella ricerca scientifica si lavora principalmente con quest'ultima. Quando si parlerà di stazionarietà, pertanto, si intenderà la stazionarietà in senso debole.

Spesso le serie non sono stazionarie. Citando Hipel e McLeod (1994), maggiore è l'arco di tempo ove si registrano le osservazioni, maggiore è la possibilità che le serie siano non stazionarie, ad esempio a causa di trend o stagionalità. Varie tecniche sono adottate per trasformare le serie in stazionarie.

⁶ Sia X una v.c.: $E[X]$ è sempre compreso tra il valore minimo e massimo dei valori che assume la v.c.; $E[X - \mu] = 0$, essendo $E(\cdot)$ un operatore lineare; il valore medio è l'unico valore che minimizza $E[X - \mu]^2$; e così via. Resta fondamentale il distinguo concettuale tra la *sintesi di una variabile casuale* e la *sintesi di una variabile statistica*. Da Piccolo (2000).

⁷ Un'assunzione implicita contenuto nel testo è che, per $j \rightarrow \infty$, $\text{cov}(x_t, x_{t-j}) \rightarrow 0$. Tale proprietà assume il nome di **ergodicità**. In breve, all'aumentare dell'ampiezza del campione, aumenta anche l'informazione in nostro possesso. Osservare il processo per un tempo "abbastanza lungo" equivale a poter osservare "quasi tutte" le sottosequenze che il processo è in grado di generare: è pertanto possibile dimostrare che la memoria del processo, per lunghi orizzonti, è debole e una sua realizzazione abbastanza lunga è equivalente all'osservazione di un gran numero di realizzazioni *indipendenti* fra loro. Per approfondimenti, Grazzini (2012).

1.5 Il rasoio di Occam

Per costruire un modello efficace, ci ispiriamo al principio del *rasoio di Occam*: tra i vari modelli in grado di rappresentare i dati di una serie, preferiremo sempre il più semplice di tutti, ossia quello con il minor numero di parametri. Come spiegato da Hipel e McLeod (1994), Occam, teologo, filosofo e francescano inglese vissuto a cavallo tra il 1200 e il 1300, sostenne come ci si dovesse liberare dai concetti superflui per intendere la realtà secondo l'imperscrutabile volontà di Dio, assente da regole e leggi che limiterebbero la libertà d'azione: come una *lama che taglia di netto gli orpelli*. Tale concetto è stato poi ripreso dalle scienze moderne nella forma di *parsimonia informativa*: ridurre la complessità dei fenomeni a una regola semplice, sciogliendone i nodi.

Più il numero di parametri cresce, più aumenta il rischio di *overfitting*: un modello assurdo e sbagliato potrebbe descrivere i dati perfettamente, ma risulterebbe inutile nell'effettuare previsioni.

Rappresentiamo quanto detto su assi cartesiani: si disegnano delle osservazioni per $t = 6$. Qual è la nostra previsione per il successivo istante temporale? Siano date quattro curve polinomiali che descrivono il comportamento delle osservazioni. Maggiore è l'ordine del polinomio, maggiore è l'accuratezza per le osservazioni effettuate; ma ancor più distorta sarà la previsione nel periodo successivo.

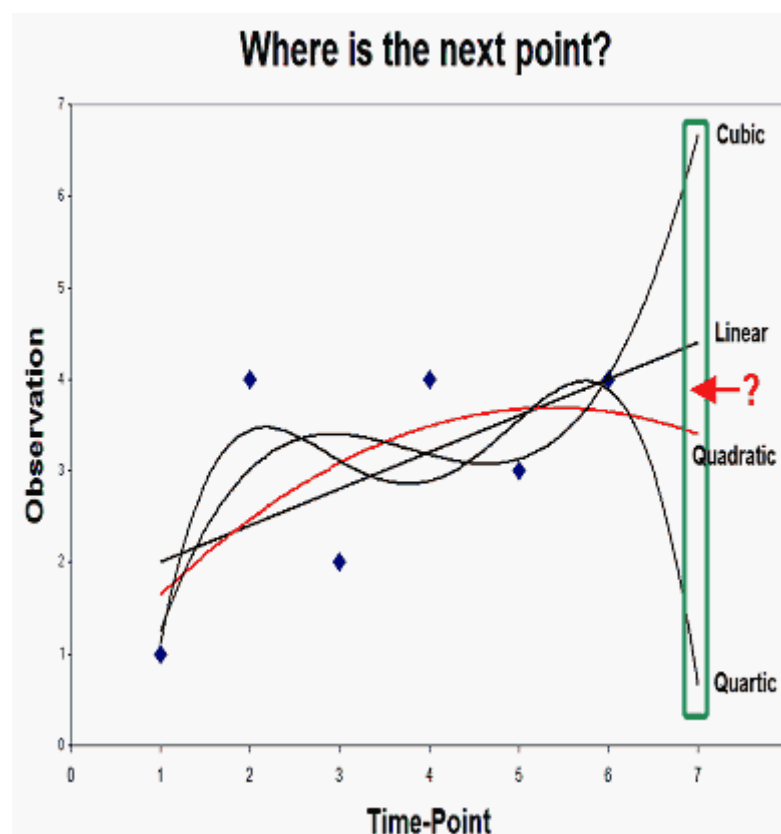


Fig.1.1. Rappresentazione del problema di *overfitting*. Tratta da “Conceptualizing Overfit or Over-parametization” di Steve Walter.

In conclusione, si deve prestare particolare attenzione nella costruzione di un modello, evitando di inserire variabili ridondanti.

2. Analisi univariata delle serie storiche⁸

La regressione classica è spesso insufficiente nello spiegare le relazioni dinamiche di un processo stocastico. Tuttavia, se abbracciamo l'idea che il passato influenza il presente, possiamo costruire dei modelli per stimare i valori attuali basandosi sulla correlazione con i valori precedenti (*ritardati*).

Utilizzando delle osservazioni ordinate secondo il criterio temporale, se esse contengono delle informazioni circa il futuro andamento della variabile *target*, è possibile effettuare una *previsione* di quest'ultima secondo una funzione $f(\cdot)$:

$$\hat{x}_{T+h} = f(x_t, x_{t-1}, \dots)$$

In questo capitolo si effettua una rassegna della famiglia dei processi stocastici *ARIMA*, largamente utilizzati in econometria.

2.1 L'operatore ritardo

Le serie storiche sono, sostanzialmente, serie numeriche. La manipolazione di tali sequenze può essere effettuata tramite l'operatore ritardo, applicabile sia a serie di variabili casuali (ossia i processi stocastici) che di numeri (ossia le rispettive traiettorie), sfalsando di un intervallo temporale, di lunghezza k , i valori della serie.

Definizione 2.1. Si definisce *operatore ritardo*:

$$Lx_t = x_{t-1}$$

un operatore lineare tale per cui l'applicazione ad una sequenza x_t produce valori differenziati di un periodo. Se l'applicazione viene effettuata k volte, si otterrà una serie x_{t-k} , ovvero valori differenziati di k periodi.

■

Si deve prestare attenzione sull'*applicazione* (e non *moltiplicazione*) di L . Ad esempio, date due sequenze x_t e y_t , definiamo $z_t = x_t y_t$. È evidente che $z_{t-1} = x_{t-1} y_{t-1}$, ma si potrebbe addurre illogicamente che: $z_{t-1} = x_{t-1} y_{t-1} = Lx_t Ly_t = L^2 x_t y_t = L^2 z_t = z_{t-2}$.

⁸ In merito a questo capitolo, sono stati fondamentali i contributi di Lütkepohl (2005), Tsay (2002). Per i processi ARMA, Mikusheva (2007). Per l'integrazione ed il caso illustrato, Granger, Newbold (1974) e Shumway, Stoffer (2018). La figura 2.1 è tratta da Banerjee et al. (1993). Si veda in bibliografia per maggiori dettagli.

2.2 Processi white-noise

Definizione 2.2. Un processo *white-noise* è tale per cui:

$$u_t \sim i.i.d. N(0, \sigma^2)$$

con le seguenti proprietà:

- $E[u_t] = E[u_t | u_{t-1}, u_{t-2}, \dots] = E[u_t | \text{informazioni disponibili a } t-1] = 0$
- $E[u_t u_{t-j}] = \text{cov}[u_t u_{t-j}] = 0$
- $\text{var}[u_t] = \text{var}[u_t | u_{t-1}, u_{t-2}, \dots] = \text{var}[u_t | \text{informazioni disponibili a } t-1] = \sigma^2$

■

Le proprietà prima e seconda rispecchiano l'assenza di *correlazione seriale* o *prevedibilità* tra gli errori. Per la terza proprietà vale l'*omoschedasticità condizionale*, quindi la varianza degli errori non dipende da quelli precedenti e rimane costante.

2.3 Processi a media mobile (MA)

Un processo a media mobile $MA(q)$ utilizza gli errori passati della serie come variabili esplicative.

Definizione 2.3. Un processo a media mobile di ordine q , abbreviato $MA(q)$, è definito come:

$$x_t = u_t + B_1 u_{t-1} + \dots + B_q u_{t-q}$$

dove q rappresentano i ritardi e $B_1, \dots, B_q$ sono parametri. Gli errori sono trattati come *white-noise*, con media zero e varianza σ^2 .

■

Applicando l'operatore ritardo, si può riscrivere:

$$x_t = (1 + B_1 L + \dots + B_q L^q) u_t$$

o, più concisamente:

$$x_t = B(L) u_t$$

I processi MA sono stazionari per ogni valore di $B_1, \dots, B_q$. Analizzandone i momenti, risulta che:

- $E[x_t] = E[u_t + B_1 u_{t-1} + \dots + B_q u_{t-q}] = (B_1 + \dots + B_q) E[u_t + u_{t-1} + \dots + u_{t-q}] = 0$
- $\text{var}(x_t) = E[x_t^2] = \sigma_x^2 \sum_{i=0}^q B_i^2$

$$\bullet \quad cov(x_t x_{t-j}) = \begin{cases} f(j) \neq 0, & j = 1, 2, \dots, q \\ 0, & j = q + 1, q + 2, \dots \end{cases}$$

Il processo $MA(q)$ è caratterizzato dal fatto che l'autocovarianza diventa zero dopo il q -esimo ritardo.

Un'altra peculiarità dei processi MA , che tornerà utile successivamente, è la funzione di risposta all'impulso: essa descrive l'effetto di un *innovation process*⁹ sui valori di una serie in un determinato istante.

Definizione 2.4. In un processo $MA(q)$ del tipo:

$$x_t = u_t + B_1 u_{t-1} + \dots + B_q u_{t-q}$$

la funzione di risposta all'impulso sarà:

$$\frac{dx_t}{du_{t-j}} = \begin{cases} B_j, & j = 1, 2, \dots, q \\ 0, & j = q + 1, q + 2, \dots \end{cases}$$

▪

La funzione di risposta all'impulso è uno degli strumenti macroeconomici di maggior popolarità, poiché consente di comprendere i rapporti causa-effetto tra i fenomeni economici e quantificarne l'impatto.

2.4 Processi autoregressivi (AR)

I processi AR sono uno dei modelli più utilizzati nell'analisi univariata delle serie.

Definizione 2.5. Si definisce *modello autoregressivo* di ordine p , abbreviato $AR(p)$, un modello della forma:

$$x_t = A_1 x_{t-1} + \dots + A_p x_{t-p} + u_t,$$

$$t = 0, \pm 1, \pm 2, \dots,$$

con u_t errore *white-noise*: $E[u_t] = 0, E[u_t u_s] = \sum u, E[u_t u'_s] = 0, s \neq t$.

▪

⁹ Nella letteratura riguardante le serie storiche, il termine *innovation* è utilizzato intercambiabilmente con *errore*. La sottile differenza rispetto all'analisi di dati cross-section è che in quest'ultimi manca un criterio temporale. Essendo invece le osservazioni nelle serie storiche ordinate rispetto al tempo, un cambiamento nel termine d'errore al tempo $t + 1$ porterà nuove informazioni rispetto a quelle disponibili al tempo t . È per tale ragione che l'errore è anche chiamato *innovation*.

Il termine costante è omissso per semplicità.

In un modello $AR(p)$ il valore futuro di una variabile è una combinazione lineare di p osservazioni ed un errore casuale.

Definizione 2.6. Usando l'operatore ritardo si può riscrivere:

$$A(L)x_t = u_t,$$

con $A(L) = 1 - A_1L - \dots - A_pL^p$, ossia un polinomio di grado p nell'operatore ritardo.

■

Le caratteristiche del processo possono essere studiate ricorrendo alla sua forma più semplice, ovverosia con un solo ritardo. Mostriamo come da un processo $AR(1)$ ci si possa ricondurre ad un processo $MA(\infty)$. Scriviamo un processo $AR(1)$:

$$x_t = Ax_{t-1} + u_t$$

$$(1 - AL)x_t = u_t$$

$$x_t = (1 - AL)^{-1}u_t$$

Che significato assume $(1 - AL)^{-1}$? Essendo una serie geometrica, si può riscrivere:

$$\frac{1}{(1 - AL)} = 1 + AL + A^2L^2 + A^3L^3 + \dots$$

Derivando che:

$$x_t = (1 + AL + A^2L^2 + A^3L^3 + \dots)u_t = C(L)u_t$$

ossia un processo $MA(\infty)$ con $C_i = A^i$ e media zero.

Definizione 2.7. Un processo $AR(1)$, del tipo:

$$x_t = Ax_{t-1} + u_t$$

può essere trasformato in un processo a media mobile $MA(\infty)$.

■

Per le proprietà della serie geometrica, deve valere $|A| < 1$ affinché il processo sia stazionario. Inoltre, le radici del polinomio caratteristico devono trovarsi fuori dal cerchio unitario (ugualmente al caso con p ritardi, con polinomio caratteristico $1 - A_1z - A_2z^2 - \dots - A_pz^p$).

Definizione 2.8. Un processo $AR(1)$ è stazionario se $|A| < 1$. Infatti, se la condizione è rispettata, la somma della serie esiste ed è convergente¹⁰. In formule:

$$\sum_{k=0}^{\infty} A^k L^k = \lim_{n \rightarrow \infty} \sum_{k=0}^n A^k L^k = \lim_{n \rightarrow \infty} \frac{1 - (AL)^{n+1}}{1 - AL} = \frac{1}{1 - AL}$$

■

Definizione 2.9. L'equazione caratteristica associata all'equazione differenziale $(1 - AL)x_t = u_t$ è $1 - AL = 0$. Ponendo $L = z$, troviamo che $z = \frac{1}{A}$ e quindi $A = \frac{1}{z}$. Se $\left|\frac{1}{A}\right| > 1 \Rightarrow |A| < 1$, il processo è stabile.

■

Analizziamo i momenti di $AR(1)$. Il *valore medio* sarà:

$$\mu = E[x_t] = E[Ax_{t-1}] + E[u_t] = A\mu + 0 = A\mu$$

L'espressione precedente è vera in due casi: se $\mu = 0$ e quindi vera per qualsiasi valore di A , oppure se $A = 1$ e quindi vera per ogni μ , con la media del processo indeterminata. In quest'ultimo caso, si parla di processo *a radice unitaria*, poiché il valore per cui $A(z) = 0$ è 1. Nel dettaglio il paragrafo 2.5.

Per la *varianza* vale:

$$\begin{aligned} var(x_t) &= var(Ax_{t-1} + u_t) = var(Ax_{t-1}) + var(u_t) + 2cov(Ax_{t-1}, u_t) \\ &= A^2 var(x_{t-1}) + var(u_t) \end{aligned}$$

Se la serie è stazionaria, la sua varianza è costante; essendo $var(x_t) = \gamma_x(0)$, si può riscrivere:

$$\gamma_x(0) = A^2 \gamma_x(0) + \sigma_u^2$$

Pertanto:

$$\gamma_x(0) = \sigma_u^2 \frac{1}{(1 - A^2)}$$

Si noti che per avere una varianza finita e positiva abbiamo bisogno che il denominatore sia maggiore di zero, dunque $|A| < 1$: se valesse 1, la funzione perderebbe di significato e per valori prossimi a 1, $\gamma_x(0) \rightarrow +\infty$.

Calcoliamo l'*autocovarianza* per un intervallo temporale pari a 1:

¹⁰ Una serie *convergente* è una serie tale che il limite delle sue somme parziali è finito. Per una serie divergente, il limite della successione risulterebbe infinito, ed il processo si definirebbe *esplosivo* ($|A| > 1$). Per approfondimenti, si veda Bertsch, Dal Passo, Giacomelli (2011).

$$\begin{aligned}\gamma_x(1) &= \text{cov}(x_t, x_{t-1}) = \text{cov}(Ax_{t-1} + u_t, x_{t-1}) = \text{cov}(Ax_{t-1}, x_{t-1}) + \text{cov}(u_t, x_{t-1}) \\ &= \text{cov}(Ax_{t-1}, x_{t-1}) + 0 = A\text{cov}(x_{t-1}, x_{t-1}) = A\gamma_x(0)\end{aligned}$$

Il coefficiente di correlazione di Pearson equivale a:

$$\frac{\gamma_x(1)}{\gamma_x(0)} = A$$

provando che A quantifica il legame tra le variabili ritardate.¹¹

2.5 Processi a radice unitaria (random walk)

Si è detto che per assicurare la stazionarietà deve valere $|A| < 1$. Ma cosa succederebbe se $|A| = 1$? Si consideri, per semplicità, un processo $AR(1)$ del tipo:

$$x_t = Ax_{t-1} + u_t, \quad u_t \sim i.i.d. (0, \sigma^2)$$

Questo può essere riscritto come:

$$x_t = x_0 + \sum_{i=0}^{t-1} u_{t-i}$$

Un processo di questo tipo è definito *random walk*.

Definizione 2.10. Si definisce *random walk (passeggiata aleatoria)* un processo del tipo:

$$x_t = x_{t-1} + u_t, \quad u_t \sim WN$$

La presenza di un elemento precedente nella definizione di x_t cambia le caratteristiche del processo stocastico. Se poniamo $x_0 = 0$, si ottiene:

$$x_1 = 0 + u_1 = u_1$$

$$x_2 = x_1 + u_2 = u_1 + u_2$$

$$\vdots$$

$$x_t = \sum_{t=1}^T u_t$$

con momenti di primo e secondo ordine pari a:

$$E[x_t] = 0 \quad \text{Var}[x_t] = \text{Var}\left[\sum_{t=1}^T u_t\right] = T\sigma^2$$

¹¹ Si noti come il coefficiente di correlazione assume notoriamente valori compresi tra -1 e 1. Ancora, $|A| < 1$.

Essendo la varianza non costante, bensì funzione di t , un processo *random walk* non è stazionario.

▪

Un *random walk* non è stazionario, ma può divenirlo se applicata una differenza di ordine pari a 1.

Definizione 2.11. Un *random walk* è stazionario in differenza¹². In formule:

$$\Delta x_t = x_t - x_{t-1} = x_{t-1} + u_t - x_{t-1} = u_t$$

▪

Se si inserisse una costante deterministica diversa da 0, si otterrebbe un processo *random walk con drift*:

$$x_1 = v + 0 + u_1 = v + u_1$$

$$x_2 = v + x_1 + u_2 = 2v + u_1 + u_2$$

⋮

$$x_t = v t + \sum_{t=1}^T u_t$$

Supponendo sempre $x_0 = 0$. Allora, la media sarà pari a:

$$E(y_t) = v t$$

Cioè la costante deterministica comporta la presenza di un *trend lineare* nella media. Il processo si definisce non stazionario *in media* (oltreché in varianza).

2.6 Processi ARMA

Combinando i processi *AR* e *MA*, si ottengono i processi *ARMA*.

Definizione 2.12. Un processo *ARMA*(p, q) è definito come:

$$x_t = A_1 x_{t-1} + \dots + A_p x_{t-p} + u_t + B_1 u_{t-1} + \dots + B_q u_{t-q}$$

¹² Un processo stazionario per la sua differenza è anche detto **integrato** di ordine 1, $x_t \sim I(1)$. Per dettagli, si faccia riferimento al paragrafo 2.7.

$$x_t = \mu_t + \sum_{i=1}^p A_i x_{t-i} + \sum_{j=1}^q B_j \mu_{t-j}$$

con $p \neq 0$ che esprime l'ordine del processo autoregressivo e $q \neq 0$ i termini di errore del processo a media mobile.

▪

Il modello *ARMA* può essere visto come un processo autoregressivo con errori correlati. Ricordando che:

$$u_t = A(L)x_t$$

$$x_t = B(L)u_t$$

Una forma più compiuta è:

$$A(L)x_t = B(L)u_t$$

Ci sono due aspetti da sottolineare: il processo è stazionario se le radici di $A(L)$ sono al di fuori del cerchio unitario: quindi, solo la parte *AR* è coinvolta; inoltre, dopo il q -esimo ritardo, l'autocorrelazione segue il pattern di *AR*.

2.7 Integrazione e processi ARIMA

I modelli *AR*, *MA*, *ARMA* possono essere utilizzati solo con dati stazionari. Nella realtà, molte serie storiche, come quelle socio-economiche e di business, mostrano un comportamento non stazionario. Pertanto, si generalizza ai modelli *ARIMA* per includere anche il caso di non-stazionarietà.

Nei modelli *ARIMA*, le serie non stazionarie diventano tali *differenziando* i dati della serie. In formule, i modelli $ARIMA(\rho, d, q)$ sono definiti come:

$$A(L)(1 - L)^d x_t = B(q)u_t$$

con d che definisce il livello di differenziazione. Un esempio famoso di modello *ARIMA* è il *random walk* ($ARIMA(0,1,0)$).

Vi sono sostanziali differenze fra serie $I(0)$ e $I(1)$.

- Se $x_t \sim I(0)$, allora i valori della serie fluttuano attorno ad una media costante; la varianza di x_t è finita; uno *shock* aleatorio ha solo un effetto temporaneo sul valore di x_t ; le autocorrelazioni decrescono velocemente all'aumentare dei ritardi, e la loro somma è finita.

- Se $x_t \sim I(1)$ con $x_0 = 0$, è assente la *mean-reversion*¹³, implicando periodi, anche protratti a lungo nel tempo, con andamenti crescenti o decrescenti molto decisi; la varianza di x_t tende a infinito quando $t \rightarrow \infty$; uno *shock* aleatorio ha un effetto permanente sulla serie; le autocorrelazioni tendono a 1 all'aumentare dei lag, per $t \rightarrow \infty$.

$$y_t = \varphi y_{t-1} + \epsilon_t$$

	$ \varphi < 1$	$\varphi = 1$
Varianza	Finita	Illimitata
Autocorrelazioni	$\rho_i = \varphi^i$	$\rho_i = \sqrt{1 - \frac{i}{t}}$
<i>mean-reverting</i>	Sì	No
Memoria	Temporanea	Permanente

Fig.2.1. Processi $I(0)$ e $I(1)$ a confronto ($AR(1)$ vs *random walk*). Tratta da Banerjee et al. (1993).

Un risultato interessante è che la somma di una serie $I(0)$ e di una $I(1)$ è una serie $I(1)$. Esso è conosciuto come la *scomposizione di Beveridge e Nelson*, o BN ¹⁴. Definiamo il caso generale.

Definizione 2.13. Se:

$$x_t \sim I(d_x)$$

$$y_t \sim I(d_y)$$

Allora vale:

$$z_t = ax_t + by_t \sim I[\max(d_x, d_y)]$$

con a e b costanti.

■

Lavorare con serie non stazionarie, trattandole come se lo fossero, porta a risultati fuorvianti. Uno dei più noti è quello della *regressione spuria*. Illustriamo il caso procedendo secondo

¹³ Termine largamente adottato dal linguaggio finanziario, che indica la fluttuazione dei valori in rialzo ed in ribasso intorno alla media della serie.

¹⁴ Nella pratica, una scomposizione BN è utilizzata in macroeconometria quando si tratta di separare trend e ciclo di una serie storica: si stima un modello $ARMA$ sulle differenze prime, quindi si scompongono i parametri stimati. È uno dei vari strumenti per raggiungere questo scopo e non è esente da critiche: su tutte è quella di supporre aprioristicamente che la serie $I(1)$ sia un *random walk*.

Granger e Newbold (1974). Prendiamo due processi *random walk* con errori *white-noise* indipendenti fra loro.

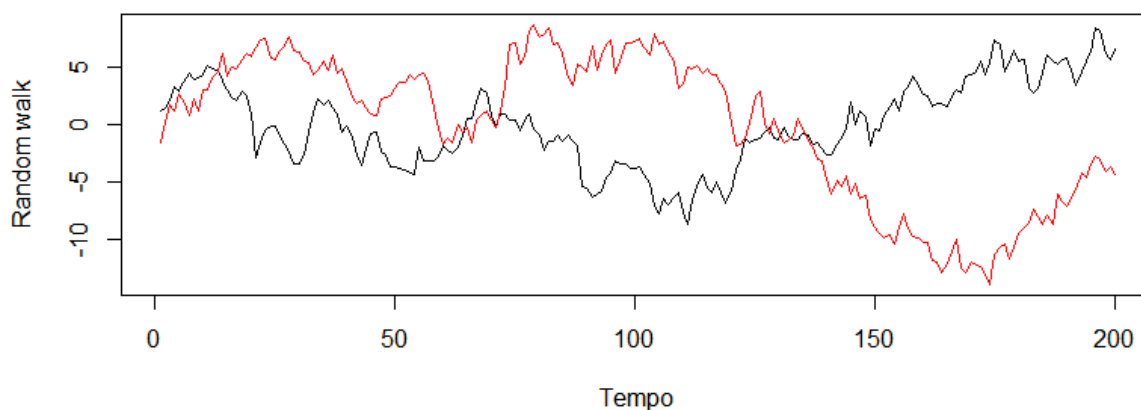


Fig.2.2. Due *random walk* a confronto.

Essendo processi altamente autoregressivi, non dovremmo trovare traccia di relazioni significative tra loro. Ma se impostiamo un modello di regressione lineare OLS, del tipo $y_t = v + x_t + u_t$, i risultati sono:

Residuals				
Min.	1Q	Median	3Q	Max.
-6.7150	-2.2883	-0.2849	2.2245	7.4873

Coefficients				
	Estimate	Std. Error	t-value	Pr(> t)
Intercept	-0.12793	0.21082	-0.607	0.545
rwalk2.ts	-0.34997	0.03393	-10.314	<2e-16 ***

Signif. codes:
 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.981 on 198 degrees of freedom
 Multiple R-squared: 0.3495, Adjusted R-squared: 0.3462

Tab.2.1. Stime prodotte dalla regressione di *rwalk1* su *rwalk2* con OLS.

Il p-value è significativo! Ma la funzione di autocorrelazione ACF mostra che l'assunzione di incorrelazione seriale tra i residui è violata.

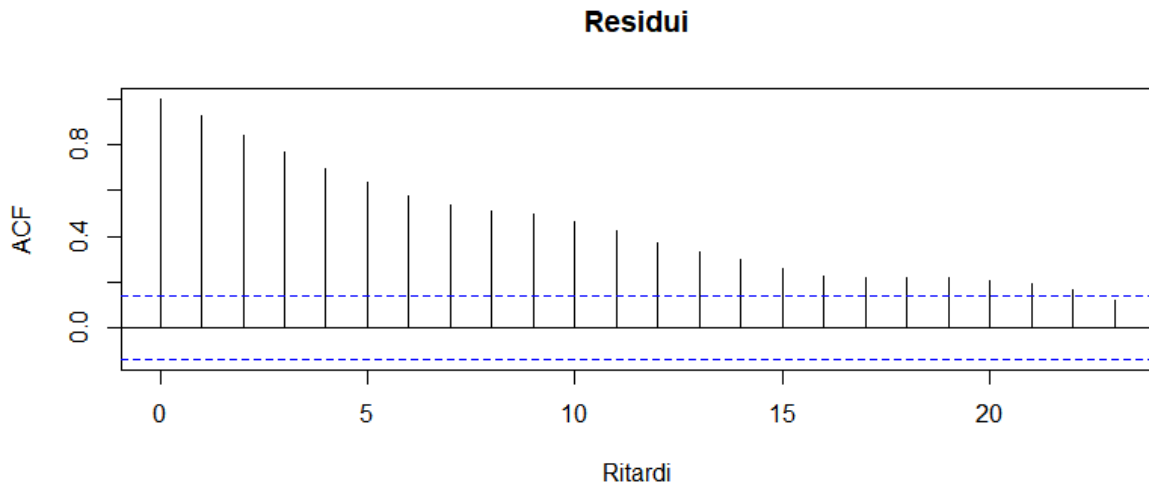


Fig.2.3. ACF del modello OLS.

Le conseguenze dell'avere errori autocorrelati sono: stime inefficienti dei coefficienti; previsioni sub-ottimali; test di significatività sui coefficienti non validi (i test t portano spesso al rifiuto dell'ipotesi nulla, tanto più tanto quanto è grande il campione).¹⁵

¹⁵ Una regressione fra variabili integrate può avere un senso: questo accade quando le variabili a destra e sinistra sono cointegrate. Si veda il paragrafo 3.4.

3. Analisi multivariata delle serie storiche¹⁶

È raro che dei fenomeni complessi, come quelli che si studiano in economia, possano essere descritti mediante una singola variabile. È naturale ricorrere a più variabili casuali o vettori aleatori in un contesto dinamico, avvalendosi dell'analisi *multivariata*.

I *momenti primi e secondi* (valore medio e matrice di covarianza) di un processo multivariato sono definiti come:

$$E[\mathbf{x}_t] = \mu$$

$$\Gamma_h(0) = E[\mathbf{x}_t - \mu][\mathbf{x}_t - \mu]'$$

dove, se il processo $\mathbf{x}_t$ ha n elementi, μ è un vettore $n \times 1$ e Γ_h è una matrice $n \times n$. L' i -esimo elemento sulla diagonale della matrice rappresenta la varianza di x_{ij} , dove l' (i, j) -esimo elemento di $\Gamma_h(0)$ è la covarianza fra x_{it} e x_{jt} .

Un errore *white-noise multivariato* sarà un processo $\mathbf{u}_t$ definito come:

$$E[\mathbf{u}_t] = 0$$

$$\Gamma_h = E[\mathbf{u}_t \mathbf{u}_{t-h}] = \begin{cases} \Sigma & h = 0 \\ 0 & h \neq 0 \end{cases}$$

risultando molto simile al caso univariato. Si noti che Σ indica una matrice di varianze e covarianze generica, e pertanto simmetrica e semidefinita positiva, ma non necessariamente diagonale. Di conseguenza, il fatto che un processo multivariato sia un *white-noise* esclude la correlazione fra gli elementi correnti del processo e tutta la sua storia passata, ma non esclude la possibilità di correlazione fra i vari elementi correnti.

Anche l'*operatore ritardo* può essere applicato nel caso in cui $\mathbf{x}_t$ sia un vettore, similmente al caso univariato. Generalizzando al caso di polinomi di ordine p , si può scrivere:

$$A(L)\mathbf{x}_t = A_0\mathbf{x}_t + A_1\mathbf{x}_{t-1} + \dots + A_p\mathbf{x}_{t-p}$$

Il fatto di poter interpretare $A(L)$ come una matrice di polinomi comporta che l'inversione di tali operatori segue le normali regole di inversioni di matrici.

¹⁶ In relazione alla redazione di questo capitolo, si è ricorso ai lavori di Lütkepohl (2005). Per i processi *VMA*, Pfaff (2008). Il caso di cointegrazione si avvale dei contributi di Brooks (2008), Johansen (2014), Murray (1994), Stigler (2010). Relativamente alla notazione matematica, di notevole aiuto Bertsch, Dal Passo, Giacomelli (2011).

3.1 Processi VAR

Un processo stocastico VAR (*Vector Autoregression*) è utilizzato nel catturare interdipendenze lineari nelle serie storiche multiple.

Definizione 3.1. Semplificando la notazione scrivendo:

- $\mathbf{x}_t = (x_{1t}, \dots, x_{Kt})'$, un insieme di variabili endogene per $k = 1, \dots, K$,
- $\mathbf{v} = (v_1, \dots, v_K)'$, un vettore ($K \times 1$) che rappresenta le variabili deterministiche e consente di lavorare nell'ipotesi di una media non nulla $E[\mathbf{x}_t]$,
- $A_i = \begin{bmatrix} \alpha_{11,i} & \cdots & \alpha_{1K,i} \\ \vdots & \ddots & \vdots \\ \alpha_{K1,i} & \cdots & \alpha_{KK,i} \end{bmatrix}$, i coefficienti della matrice,

la forma del modello VAR di ordine p potrà scriversi come:

$$\mathbf{x}_t = \mathbf{v} + A_1 \mathbf{x}_{t-1} + \dots + A_p \mathbf{x}_{t-p} + \mathbf{u}_t,$$

dove $\mathbf{u}_t = (u_{1t}, \dots, u_{Kt})'$ forma una sequenza i.i.d. (indipendenti ed identicamente distribuiti) di K vettori casuali con valore atteso uguale a zero: $E[u_{Kt}] = 0$, per $K = 1, 2, \dots, n$.

■

Osservando un VAR(1) del tipo:

$$x_t = v + A_1 x_{t-1} + u_t$$

e stabilendo che la ricezione dei dati osservati inizi al tempo $t = 1$, si ottiene:

$$\begin{aligned} x_1 &= v + A_1 x_0 + u_1 \\ x_2 &= v + A_1 x_1 + u_2 = v + A_1(v + A_1 x_0 + u_1) + u_2 \\ &= (I_k + A_1)v + A_1^2 x_0 + A_1 u_1 + u_2,^{17} \\ &\vdots \end{aligned}$$

$$x_t = (I_k + A_1 + \dots + A_1^{t-1})v + A_1^t x_0 + \sum_{i=0}^{t-1} A_1^i u_{t-i}$$

Pertanto, le distribuzioni congiunte (e non) delle $\mathbf{x}_t$ sono unicamente determinate dalle distribuzioni, congiunte e non, di $x_0, u_1, \dots, u_t$.

¹⁷ I_k è una matrice identità, ossia una matrice quadrata ($n \times n$) in cui tutti gli elementi nella diagonale sono costituiti dal numero 1, mentre tutti i restanti sono 0. La proprietà fondamentale di I_k è che $A \cdot I_k = A$, $I_k \cdot B = B$.

3.2 Stabilità e stazionarietà dei processi VAR

Per dimostrare la stazionarietà, si procede come segue. Consideriamo un VAR(p) del tipo:

$$\mathbf{x}_t = A_1 \mathbf{x}_{t-1} + \dots + A_p \mathbf{x}_{t-p} + \mathbf{u}_t$$

Usando l'operatore ritardo, $A(L)$, si può riscrivere:

$$A(L)\mathbf{x}_t = \mathbf{u}_t,$$

con $A(L) = I_k - A_1 L - \dots - A_p L^p$. L'operatore ritardo riduce il grado di $\mathbf{x}_t$.

$$\mathbf{x}_t = A(L)^{-1} \mathbf{u}_t$$

Sostituiamo l'operatore L con $z \in \mathbb{C}$ per studiare le proprietà algebriche del polinomio. Il determinante di $A(z)$ viene detto *polinomio caratteristico*, poiché dalle proprietà della matrice A dipenderà la dinamicità del sistema studiato.

Definizione 3.2. Il processo si dice stabile se:

$$\det A(z) = \det(I_k - A_1 z - \dots - A_p z^p) \neq 0$$

per $z \in \mathbb{C}$, $|z| \leq 1$. Allora, le radici caratteristiche del polinomio risultano stazionarie e si trovano fuori dal cerchio unitario.

■

Ad esempio, sia:

$$x_t = Ax_{t-1} + u_t, \text{ con } A = \begin{bmatrix} 3 & 6 \\ 1 & 4 \end{bmatrix}$$

allora:

$$A(z) = I - Az = \begin{pmatrix} 1 - 3z & -6z \\ -z & 1 - 4z \end{pmatrix}$$

$$\det(A(z)) = (1 - 3z)(1 - 4z) - (-6z)(-z) = 1 - 7z + 12z^2 - 6z^2 = 6z^2 - 7z + 1$$

Poiché è un'equazione di secondo grado a coefficienti complessi, si può risolvere in $\mathbb{C}$ con la stessa formula valida in $\mathbb{R}$:

$$z = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

Troviamo che $z_1 = 1$, $z_2 = \frac{1}{6}$. Ne risulta che il processo è esplosivo, poiché ha una matrice dentro il cerchio unitario.

Altrimenti, avremmo potuto utilizzare il metodo “*companion matrix*”, ossia della matrice compagna, calcolando gli autovalori della matrice, che non sono altro che l'inverso delle radici

caratteristiche, ottenute ponendo $L = z$. Le soluzioni sarebbero invece cadute dentro il cerchio unitario, contrariamente alla ricerca delle radici z .

Ad esempio, poiché un processo $\text{VAR}(p)$ può essere riscritto come un processo $\text{VAR}(1)$, del tipo:

$$\xi_t = A\xi_{t-1} + \mathbf{u}_t^{18}$$

dove:

$$\xi_t = \begin{pmatrix} \mathbf{y}_t \\ \vdots \\ \mathbf{y}_{t-p+1} \end{pmatrix}, \quad A = \begin{bmatrix} A_1 & A_2 & \dots & A_{p-1} & A_p \\ I & 0 & \dots & 0 & 0 \\ 0 & I & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & I & 0 \end{bmatrix}, \quad \mathbf{v}_t = \begin{pmatrix} \mathbf{u}_t \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

$$(I - AL)\xi_t = \mathbf{v}_t$$

Cerchiamo la soluzione calcolando i valori per cui $\det(\lambda I - A) = 0$. Riprendendo l'esempio precedente, gli autovalori corrispondenti della matrice A sono:

$$\det(A - \lambda Id) = \det \begin{pmatrix} 3 - \lambda & 6 \\ 1 & 4 - \lambda \end{pmatrix} = (3 - \lambda)(4 - \lambda) - 6 = \lambda^2 - 7\lambda + 6$$

Risolvendo:

$$\lambda_{1,2} = \frac{7 \pm \sqrt{49 - 24}}{2} = \frac{7 \pm 5}{2}$$

Giungiamo alle soluzioni caratteristiche $\lambda_1 = 1$, $\lambda_2 = 6$. Notiamo che sono proprio l'inverso delle radici z ; allora, per gli autovalori associati alla matrice A , si verificherà che esse siano dentro il cerchio unitario.

3.3 Processi VMA

Si prenda in esame un processo autoregressivo di ordine p con notazione precedentemente descritta, del tipo:

$$\mathbf{x}_t = \mathbf{v} + A_1\mathbf{x}_{t-1} + \dots + A_p\mathbf{x}_{t-p} + \mathbf{u}_t,$$

$$t = 0, \pm 1, \pm 2, \dots,$$

con $\mathbf{u}_t$ errore *white-noise*: $E[\mathbf{u}_t] = 0, E[\mathbf{u}_t\mathbf{u}_s] = \Sigma \mathbf{u}, E[\mathbf{u}_t\mathbf{u}_s'] = 0, s \neq t$.

I processi VAR possono essere trasformati in $\text{VMA}(\infty)$ secondo la decomposizione a media mobile di Wold.

¹⁸ Le componenti deterministiche sono soppresse nella notazione senza perdita di informazioni.

Definizione 3.3. Se il processo è stabile, allora esiste una rappresentazione:

$$\mathbf{x}_t = \Phi_0 \mathbf{u}_t + \Phi_1 \mathbf{u}_{t-1} + \Phi_2 \mathbf{u}_{t-2} + \dots$$

con $\Phi_0 = I_k$ e $\Phi_s = \sum_{j=1}^s \Phi_{s-j} A_j$, $s = 1, 2, \dots$

■

Un $VAR(p)$ può scriversi nella forma di $VAR(1)$ tale che

$$A(L)\mathbf{x}_t = \mathbf{u}_t$$

Sia

$$\Phi(L) = \sum_{i=0}^{\infty} \Phi_i L^i$$

un operatore tale che $\Phi(L)A(L) = I_k$. Se moltiplichiamo $A(L)\mathbf{x}_t = \mathbf{u}_t$ per $\Phi(L)$, si ottiene:

$$\mathbf{x}_t = \Phi(L)\mathbf{u}_t = \sum_{i=0}^{\infty} \Phi_i \mathbf{u}_{t-i}$$

L'operatore $\Phi(L)$ è quindi l'inverso di $A(L)$ e talvolta si denota come $A(L)^{-1}$. Il processo è definito se $A(L)$ è invertibile, quindi se $|A(z)| \neq 0$ per $|z| < 1$.

Richiamando quanto visto nel paragrafo 2.3, sotto tale prospettiva $\mathbf{x}_t = D(\mathbf{u}_t)$: $\mathbf{x}_t$ è un vettore delle n variabili osservabili, $\mathbf{u}_t$ rappresenta l'insieme degli *shock* $\mathbf{w}$ accaduti sino al periodo t , $D(\cdot)$ è la funzione di risposta all'impulso.

3.4 Cointegrazione e processi VEC

Si definisce cointegrazione “il caso in cui due o più serie temporali con trend stocastici si muovono congiuntamente in modo simile nel lungo periodo, tanto che sembrano possedere lo stesso trend”¹⁹. Consideriamo combinazioni lineari di processi univariati.

Siano due serie $\mathbf{x}_{1,t}$ e $\mathbf{x}_{2,t}$ entrambe $I(d)$: allora, la combinazione $\mathbf{z}_t = \mathbf{x}_{2,t} - \boldsymbol{\beta} \mathbf{x}_{1,t}$ è ancora $I(d)$.

Supponiamo adesso che $\mathbf{z}_t \sim I(d - b)$, con $b > 0$:

- se $d = b = 1$, allora $\mathbf{x}_{1,t}$ e $\mathbf{x}_{2,t}$ sono entrambe $I(1)$ e presentano componenti di lungo periodo dominanti;

¹⁹ Peracchi, F., 2012. *Cointegrazione*, Dizionario di Economia e Finanza, Enciclopedia Treccani.

- se, invece, $\mathbf{z}_t \sim I(0)$ (stazionario), le componenti di lungo periodo delle due serie devono *controbilanciarsi* per dar luogo a un processo stazionario. In altre parole, se due serie si muovono insieme, seppur individualmente non stazionarie, può esistere un valore del parametro $\boldsymbol{\beta}$ tale che la combinazione lineare risulti stazionaria, eliminando il trend comune. In formule, $\mathbf{x}_{2,t} = \boldsymbol{\beta}\mathbf{x}_{1,t}$. Lo scalare $\mathbf{z}_t$ diventa la misura delle deviazioni del sistema $(\mathbf{x}_{1,t}, \mathbf{x}_{2,t})$ dal suo equilibrio e viene chiamato *errore di equilibrio*.²⁰

Se due serie presentano ordini di integrazioni differenti, allora esse non possono cointegrare, date le diverse grandezze delle varianze.

Definizione 3.4. Dato un insieme di variabili $\{x_{1t}, \dots, x_{kt}\} \sim I(1)$, se esiste una combinazione lineare tale che:

$$\beta_1 x_{1t} + \dots + \beta_k x_{kt} = \boldsymbol{\beta}' \mathbf{x}_t, \text{ trend-stazionario}^{21}$$

$\beta_j \neq 0, j = 1, \dots, k$, allora tutte le variabili x sono cointegrate di ordine 1. In generale, se $\mathbf{x} \sim I(d)$, $\boldsymbol{\beta}' \mathbf{x} \sim I(d - b)$. Il vettore $\boldsymbol{\beta}$ è chiamato vettore di cointegrazione.

■

Ci sono tre modi per decidere se due variabili possono essere cointegrabili:

- Assunzioni tecniche derivanti dalla teoria economica;
- Trend stocastici comuni;
- Test statistici per la cointegrazione.

Definizione 3.5. Dato un processo $VAR(p)$:

$$\mathbf{x}_t = \mathbf{A}_1 \mathbf{x}_{t-1} + \dots + \mathbf{A}_p \mathbf{x}_{t-p} + \mathbf{u}_t$$

²⁰ Le variabili vanno dove vogliono, eppure esiste tra loro una relazione che vale sempre. Romanticamente, si può pensare a due variabili $I(1)$ come a due innamorati eternamente in viaggio, ma sempre legati. Murray (1994) usa la similitudine di un ubriaco con il suo cane: usciti insieme dal bar, fanno due passeggiate *random walk* per rientrare a casa. Il cane è senza guinzaglio, cosicché ad ogni passo l'ubriaco lo richiama. Il cane smette di vagabondare senza una precisa direzione, abbaiando al padrone. Ognuno sente l'altro. Il cane allora pensa: "Non posso allontanarmi troppo, altrimenti mi punirà rinchiudendomi dentro casa!", ed il padrone, a sua volta: "Non posso allontanarmi troppo, altrimenti mi sveglierà a notte fonda abbaiando per rientrare dentro casa!". Dunque, entrambi si riavvicinano, colmando il gap che li separava: ecco il *meccanismo di correzione d'errore*.

²¹ Trend-stazionario significa che, sottratto il trend deterministico, il processo diventa $I(0)$.

Esiste un *modello di correzione d'errore* tale che (ignorando termini costanti e tenendo conto che $\mathbf{x}_t = \mathbf{x}_{t-1} + \Delta\mathbf{x}_t$):

$$\Delta\mathbf{x}_t = \mathbf{\Pi}\mathbf{x}_{t-1} + \sum_{i=1}^{p-1} \mathbf{A}_i^* \Delta\mathbf{x}_{t-i} + \mathbf{u}_t$$

con $\mathbf{A}_i^* = -\sum_{j=i+1}^p \mathbf{A}_j$, $j = 1, \dots, p-1$; $\mathbf{\Pi} = -(\mathbf{I} - \mathbf{A}_1 - \dots - \mathbf{A}_p) = -\mathbf{A}(z=1)$.

■

Interpretiamo tale rappresentazione. Se $\mathbf{\Pi} = \mathbf{0}$, $\mathbf{x}_t$ è un *random walk* multivariato e non esiste cointegrazione; se $\mathbf{\Pi}$ ha rango pieno²², allora le variabili sono stazionarie perché $\mathbf{\Pi}$ è invertibile ($\mathbf{\Pi}^{-1}\Delta\mathbf{x}_t = \mathbf{x}_{t-1} + \dots + \mathbf{x}_{t-p} + \mathbf{\Pi}^{-1}\mathbf{u}_t$); se $\mathbf{\Pi}$ ha rango m , $0 < m < k$, allora c'è cointegrazione. Scriviamo: $\mathbf{\Pi} = \boldsymbol{\alpha}\boldsymbol{\beta}'$, $(k \times k) = (k \times m)[(k \times m)']$, dove le colonne di $\boldsymbol{\beta}$ contengono m vettori di cointegrazione e le colonne di $\boldsymbol{\alpha}$ m vettori di aggiustamento.

Ponendo $\Delta\mathbf{x} = \mathbf{0}$, otteniamo la relazione di lungo periodo

$$\mathbf{\Pi}\mathbf{x}^* = \mathbf{0}$$

$$\mathbf{\Pi}\mathbf{x}^* = \boldsymbol{\alpha}(\boldsymbol{\beta}'\mathbf{x}^*) = \mathbf{0}$$

Nel caso in cui $0 < \text{Rango}(\mathbf{\Pi}) = \text{Rango}(\boldsymbol{\alpha}) = m < k$, il numero di soluzioni del sistema di equazioni lineari diverse da zero è m : $\boldsymbol{\beta}'\mathbf{x}^* = \mathbf{0}_{m \times 1}$.

Se in $t-1$ assistiamo ad una deviazione,

$$\mathbf{\Pi}\mathbf{x}^* = \boldsymbol{\xi}_{t-1} \neq \mathbf{0}$$

I coefficienti in $\boldsymbol{\alpha}$ moltiplicati per gli errori $\boldsymbol{\beta}'\mathbf{x}_{t-1}$ producono un aggiustamento ($\Delta\mathbf{x}_t$). Ad esempio, se $\mathbf{x}_t$ devia dall'equilibrio di lungo periodo diventando “troppo grande”, $\boldsymbol{\xi}_{t-1} > 0$, $\mathbf{x}_t$

²² Si definisce *rango* o caratteristica di una matrice l'ordine del determinante più alto estraibile da una matrice che sia diverso da zero. Altrimenti detto, è il numero di colonne, o righe, linearmente indipendenti. Sia una matrice $m \times n$. Con $m < n$, essa avrà rango pieno quando le sue m righe sono linearmente indipendenti; con $m > n$, la matrice avrà rango pieno quando le sue n colonne sono linearmente indipendenti. Se $m = n$, la matrice avrà rango pieno se le sue righe o colonne sono linearmente indipendenti. Se una matrice è quadrata, si può anche guardare squisitamente al suo determinante: se diverso da zero, allora avrà rango pieno.

deve decrescere per garantire l'equilibrio di lungo periodo: pertanto, $\Delta \mathbf{x}_t < 0$; se diventa “troppo piccola”, sarà l'inverso.²³

Consideriamo ora un $VAR(p)$ con $\mathbf{x}_t \sim I(1)$. Allora, $\Delta \mathbf{x}_t \sim I(0)$ e $\Pi = -(\mathbf{I} - \mathbf{A}_1)$. Nel modello a correzione d'errore, $\Pi = \alpha\beta'$, $\Pi\mathbf{x}_{t-1}$ *error-correction term* (ECM). Il caso di nostro interesse è se Π ha rango m , $0 < m < k$: allora ci sono m colonne linearmente indipendenti in β ; m combinazioni lineari stazionarie $\beta'\mathbf{x}_t$; $\mathbf{x}_t$ possiede $(k - m)$ radici unitarie (trend stocastici comuni). Pertanto, $k = m + (k - m)$, k variabili $I(1)$, m relazioni di cointegrazione, $(k - m)$ trend.

3.5 Cointegrazione: approccio di Engle-Granger e caso bivariato

Quale potrebbe essere una strategia di modellizzazione dei dati se si sospetta essi siano non stazionari e cointegrati? Vi sono almeno tre metodi: Engle-Granger, Engle-Yoo, Johansen. In particolare, secondo Engle-Granger un'equazione di lungo periodo (ossia, una cointegrazione) esiste solo se gli errori sono stazionari.

Definizione 3.6. Si definisce *test di Engle-Granger*:

$$\mathbf{x}_t, y_t \sim I(1)$$

$$y_t = \alpha + \mathbf{x}_t'\beta + \mathbf{u}_t$$

$$\begin{cases} H_0: \mathbf{u}_t \sim I(1), & \text{no cointegrazione} \\ H_1: \mathbf{u}_t \sim I(0), & \text{cointegrazione} \end{cases}$$

Se $\mathbf{u}_t$ è stazionario, $\mathbf{x}$ e y sono cointegrati.

■

Il metodo di Engle-Granger è una “*single-equation technique*”²⁴ e condotto in due step:

- Step 1

Accertarsi che tutte le variabili siano $I(1)$. Dopodiché, stimare con OLS la regressione cointegrata. Non è possibile effettuare inferenza sui coefficienti stimati: si può solo calcolare il valore dei parametri. Testare sui residui della regressione l'ipotesi di stazionarietà. Se sono $I(0)$, procedere al secondo step; altrimenti, stimare un modello che contenga squisitamente le differenze prime.

²³ È bene specificare che, per definizione, il meccanismo di correzione d'errore è asintotico. La convergenza necessita sempre di un tempo infinito: pertanto, i coefficienti di aggiustamento non devono essere legati ad un'interpretazione temporale.

²⁴ Brooks, C., 2008. *Introductory Econometrics for Finance*, p. 341.

- Step 2

Usare i residui come variabile esplicativa. La combinazione lineare stazionaria di variabili non stazionarie è il vettore di cointegrazione. È possibile inferire sulla regressione al secondo stadio.

Effettuiamo una simulazione in R sulla cointegrazione. Vogliamo analizzare il legame dei tassi di interesse tra i titoli di stato americani scadenziati a 3 mesi e a 12 mesi. Ci serviamo dei dati della Federal Reserve Bank. Costruiamo una struttura dei rendimenti basata sulle ipotesi di *market efficiency* e aspettative razionali:

$$(1 + R_{1y,t}) = (1 + R_{3m,t})(1 + f_{3m,1y,t})$$

con $R_{1y,t}$ che indica il tasso di interesse sul prestito dei *bond* americani a 3 mesi, $R_{3m,t}$ il tasso di interesse sul prestito a 12 mesi e $f_{3m,1y,t}$ il tasso *forward* tra i 3 mesi e 1 anno. L'implicazione della teoria delle aspettative è che i due tassi siano cointegrati con il vettore di cointegrazione uguale a 1. Le osservazioni, con frequenza mensile, partono dal 1962, e sono troncate al 2006.

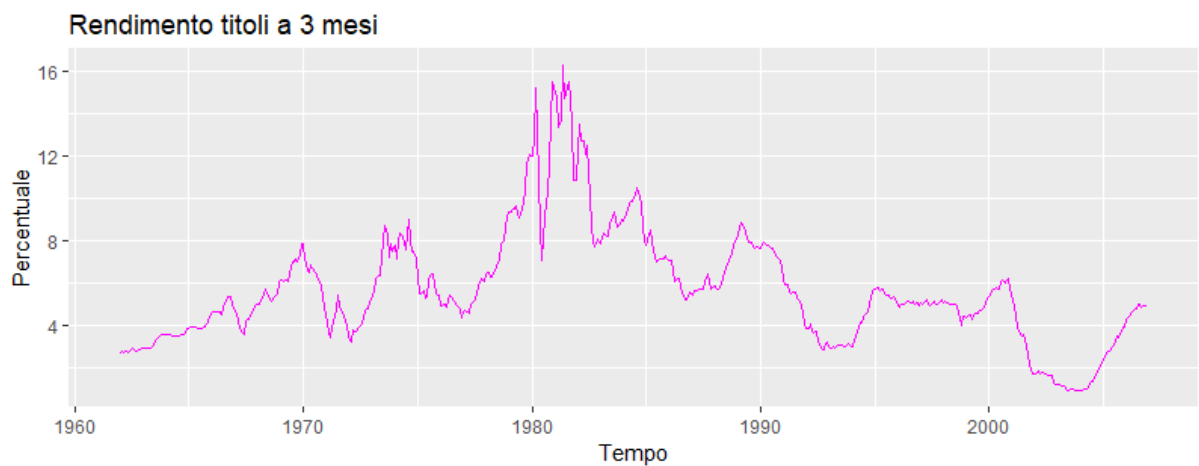


Fig.3.1. Rendimenti dei titoli di stato americani a 3 mesi.

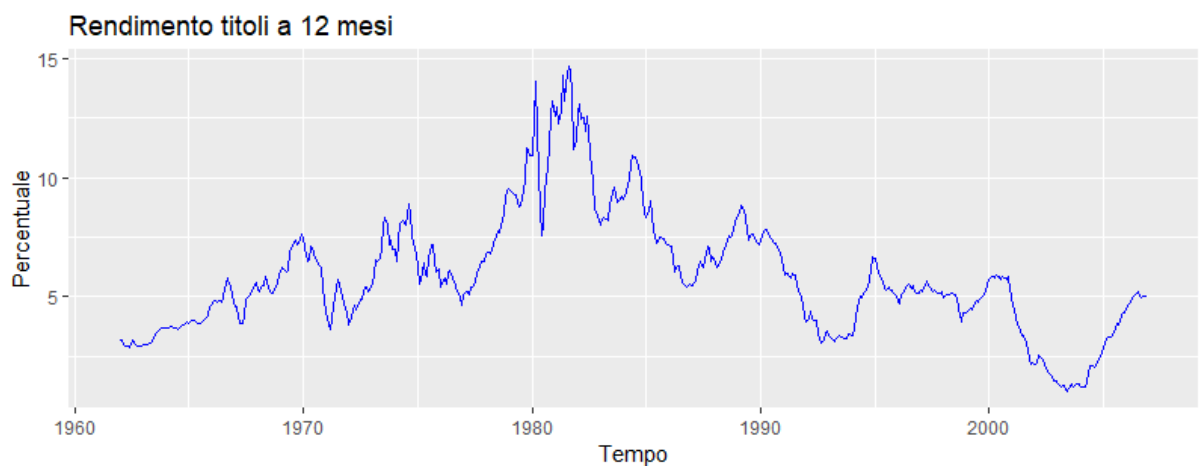


Fig.3.2. Rendimenti dei titoli di stato americani a 12 mesi.

Calcoliamo un modello $AR(1)$ con stime OLS, prima sulla serie dei titoli a 3 mesi, poi quella annuale.

```
Coefficients:
      1
0.9836

order selected 1
sigma^2 estimated as 0.244
```

Fig.3.3. Stime $AR(1)$ titoli a 3 mesi.

```
Coefficients:
      1
0.9862

order selected 1
sigma^2 estimated as 0.1797
```

Fig.3.4. Stime $AR(1)$ titoli a 12 mesi.

Entrambi i coefficienti sono vicini a 1. Usiamo il test di Dickey-Fuller aumentato per verificare la presenza di una radice unitaria nelle serie. Se accettiamo l'ipotesi nulla, il processo è non stazionario.²⁵

Per il caso a 3 mesi,

```
Dickey-Fuller = -2.3212,
Lag order = 0,
p-value = 0.4424
alternative hypothesis: stationary
```

```
Dickey-Fuller = -3.1218,
Lag order = 1,
p-value = 0.1035
alternative hypothesis: stationary
```

```
Dickey-Fuller = -2.4892,
Lag order = 8,
p-value = 0.3712
alternative hypothesis: stationary
```

Per il caso a 12 mesi,

```
Dickey-Fuller = -2.1571,
Lag order = 0,
p-value = 0.5118
alternative hypothesis: stationary
```

²⁵ Un aspetto condiviso dai test *unit root* è di non essere efficaci nel distinguere tra radici unitarie e valori vicini ad esse. Qui è usato il DF per la sua comodità, giacché esplicita direttamente i p -value. Vi sono diversi test alternativi più potenti del Dickey-Fuller, ma nessuno di essi è infallibile.

```

Dickey-Fuller = -2.9619,
Lag order = 1,
p-value = 0.1712
alternative hypothesis: stationary

```

```

Dickey-Fuller = -2.4382,
Lag order = 8,
p-value = 0.3928
alternative hypothesis: stationary

```

In nessun caso rigettiamo l'ipotesi nulla. Pertanto, le due serie sono integrate, $I(1)$. Rappresentiamo le due serie insieme con il termine *spread*, definito come:

$$diff = R_{1y,t} - R_{3m,t}$$

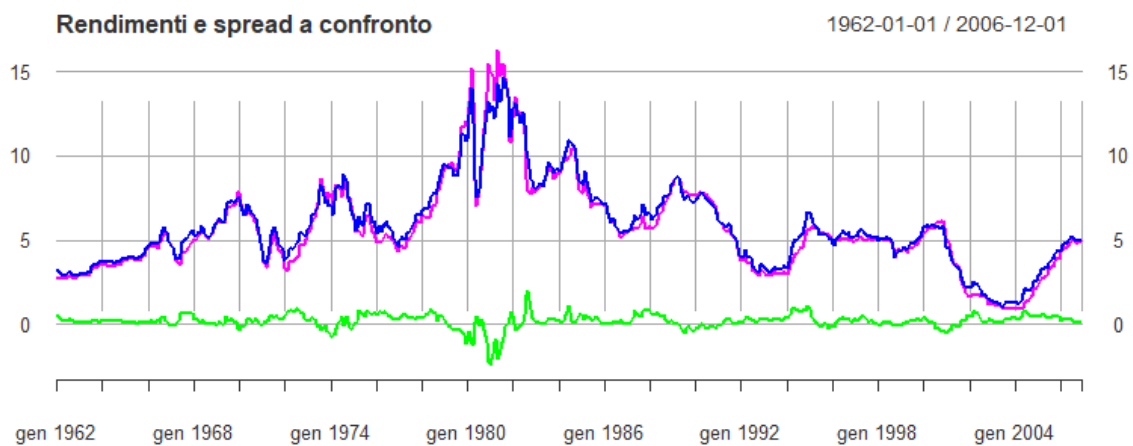


Fig.3.5. Rendimenti e spread a confronto.

I due tassi di interesse si muovono assieme, e sono ovvi candidati per la cointegrazione. Testiamo per $I(0)$ sulla differenza.

```

Dickey-Fuller = -6.0127,
Lag order = 0,
p-value = 0.01
alternative hypothesis: stationary
p-value smaller than printed p-value

```

```

Dickey-Fuller = -7.3329,
Lag order = 1,
p-value = 0.01
alternative hypothesis: stationary
p-value smaller than printed p-value

```

```

Dickey-Fuller = -4.7383,
Lag order = 8,
p-value = 0.01
alternative hypothesis: stationary
p-value smaller than printed p-value

```

Rigettiamo l'ipotesi nulla, definendo la differenza stazionaria, $I(0)$.

Usiamo un modello lineare che non tiene conto dei valori ritardati. Regrediamo:

$$R_{1y,t} = \alpha + \beta R_{3m,t} + \varepsilon_t$$

Residuals				
Min.	1Q	Median	3Q	Max.
-1.90051	-0.20566	-0.03407	0.21571	1.96860

Coefficients				
	Estimate	Std. Error	t-value	Pr(> t)
Intercept	0.623605	0.036016	17.32	<2e-16 ***
as.matrix (TB3MS)	0.925227	0.005718	161.80	<2e-16 ***

Signif. codes:
0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.3633 on 538 degrees of freedom
Multiple R-squared: 0.9799 Adjusted R-squared: 0.9798

Tab.3.1. Stime OLS, modello lineare.

Mentre la teoria economica suggerisce:

$$R_{1y,t} = R_{3m,t}$$

noi stimiamo:

$$R_{1y,t} \cong 0.62 + 0.93R_{3m,t}$$

Lo stimatore OLS fornisce risultati consistenti ($\hat{\beta} \rightarrow \beta$ per $T \rightarrow \infty$) ma non possono essere utilizzati per inferire.

Verifichiamo la stazionarietà dei residui con test DF.

Dickey-Fuller = -6.3586,
Lag order = 0,
p-value = 0.01
alternative hypothesis: stationary
p-value smaller than printed p-value

```
Dickey-Fuller = -7.7544,
Lag order = 1,
p-value = 0.01
alternative hypothesis: stationary
p-value smaller than printed p-value
```

```
Dickey-Fuller = -5.0887,
Lag order = 8,
p-value = 0.01
alternative hypothesis: stationary
p-value smaller than printed p-value
```

Rifiutiamo l'ipotesi nulla e accettiamo che i residui sono stazionari. Pertanto, le due serie sono cointegrate.

Costruiamo adesso un modello dinamico che includa i ritardi delle variabili: poiché il *long rate* eccede in valore lo *short rate*, ci aspettiamo un termine di correzione d'errore positivo affinché lo *short rate* cresca. I modelli sono i seguenti:

- Modello 1: $\Delta R_{3m,t} = \mu_1 + \phi_{11}\Delta R_{3m,t-1} + \phi_{12}\Delta R_{3m,t-2} + \gamma_{11}\Delta R_{1y,t-1} + \gamma_{12}\Delta R_{1y,t-2}$
- Modello 2: $\Delta R_{1y,t} = \mu_2 + \phi_{21}\Delta R_{3m,t-1} + \phi_{22}\Delta R_{3m,t-2} + \gamma_{21}\Delta R_{1y,t-1} + \gamma_{22}\Delta R_{1y,t-2}$
- Modello 1 con ECM (error correction term): $\Delta R_{3m,t} = \mu_1 + \phi_{11}\Delta R_{3m,t-1} + \phi_{12}\Delta R_{3m,t-2} + \gamma_{11}\Delta R_{1y,t-1} + \gamma_{12}\Delta R_{1y,t-2} + \beta_1(R_{1y,t-1} - \alpha R_{3m,t-1})$
- Modello 2 con ECM (error correction term): $\Delta R_{1y,t} = \mu_2 + \phi_{21}\Delta R_{3m,t-1} + \phi_{22}\Delta R_{3m,t-2} + \gamma_{21}\Delta R_{1y,t-1} + \gamma_{22}\Delta R_{1y,t-2} + \beta_2(R_{1y,t-1} - \alpha R_{3m,t-1})$

Le stime OLS sono:

- Modello 1:

```
Residuals:
    Min       1Q   Median       3Q      Max
-3.03568 -0.12947  0.01026  0.15149  2.09470

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)      0.003088   0.018630   0.166   0.8684
L(d(zoo.mtbill3m), 1) 0.006669   0.091499   0.073   0.9419
L(d(zoo.mtbill3m), 2) -0.065869   0.091191  -0.722   0.4704
L(d(zoo.mtreas), 1)   0.536818   0.107053   5.014 7.26e-07 ***
L(d(zoo.mtreas), 2)  -0.200056   0.109298  -1.830   0.0678 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.4317 on 532 degrees of freedom
Multiple R-squared:  0.1937, Adjusted R-squared:  0.1877
F-statistic: 31.96 on 4 and 532 DF, p-value: < 2.2e-16
```

- Modello 2:

```

Residuals:
    Min       1Q   Median       3Q      Max
-2.21012 -0.13161 -0.00381  0.17465  1.32022

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)      0.002935   0.015875   0.185   0.8534
L(d(zoo.mtbill3m), 1) -0.167990   0.077971  -2.155   0.0316 *
L(d(zoo.mtbill3m), 2)  0.143564   0.077708   1.847   0.0652 .
L(d(zoo.mtreas), 1)    0.632402   0.091225   6.932 1.20e-11 ***
L(d(zoo.mtreas), 2)   -0.384039   0.093138  -4.123 4.33e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.3679 on 532 degrees of freedom
Multiple R-squared:  0.1933, Adjusted R-squared:  0.1872
F-statistic: 31.86 on 4 and 532 DF, p-value: < 2.2e-16

```

- Modello 1 (con ECM):

```

Residuals:
    Min       1Q   Median       3Q      Max
-2.86704 -0.13249  0.00645  0.13327  2.40077

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)     -0.03472   0.02069  -1.678   0.09388 .
L(d(zoo.mtbill3m), 1)  0.08513   0.09238   0.922   0.35720
L(d(zoo.mtbill3m), 2)  0.04472   0.09415   0.475   0.63501
L(d(zoo.mtreas), 1)    0.45335   0.10766   4.211 2.99e-05 ***
L(d(zoo.mtreas), 2)   -0.30906   0.11124  -2.778   0.00566 **
L(zoo.mtreas - zoo.mtbill3m) 0.18983   0.04773   3.977 7.94e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.4258 on 531 degrees of freedom
Multiple R-squared:  0.217, Adjusted R-squared:  0.2097
F-statistic: 29.44 on 5 and 531 DF, p-value: < 2.2e-16

```

- Modello 2 (con ECM):

```

Residuals:
    Min       1Q   Median       3Q      Max
-2.12170 -0.13203 -0.00607  0.17098  1.37612

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)    -0.009187   0.017854  -0.515   0.6071
L(d(zoo.mtbill3m), 1) -0.142834   0.079723  -1.792   0.0738 .
L(d(zoo.mtbill3m), 2)  0.179019   0.081246   2.203   0.0280 *
L(d(zoo.mtreas), 1)    0.605640   0.092907   6.519 1.65e-10 ***
L(d(zoo.mtreas), 2)   -0.418988   0.095995  -4.365 1.53e-05 ***
L(zoo.mtreas - zoo.mtbill3m) 0.060861   0.041189   1.478   0.1401
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.3674 on 531 degrees of freedom
Multiple R-squared:  0.1966, Adjusted R-squared:  0.189
F-statistic: 25.98 on 5 and 531 DF, p-value: < 2.2e-16

```

Poiché la differenza prima di $R_{3m,t}$ ha un coefficiente significativo nel termine a correzione d'errore (la differenza $R_{1y,t} - R_{3m,t}$ con lag uno), mentre quella di $R_{1y,t}$ ne ha uno non significativo, si può concludere che il tasso $R_{1y,t}$ sia esogeno nel lungo periodo, opinione condivisa da diversi economisti. La velocità di aggiustamento nel breve periodo con cui è corretto il disequilibrio del periodo precedente è del 18,983% per mese.

Parte seconda

4. Introduzione all'analisi strutturale²⁶

In questo capitolo, si presenta dapprima il *percursor* storico che ha portato all'adozione dei modelli *VAR* nell'analisi macroeconometrica, ed in particolare dei modelli *SVAR*, al giorno d'oggi i più utilizzati nella determinazione del processo di *policy-making*. Di seguito, si espone il concetto di *G-causality* di Granger, il quale formalizza i rapporti di causa-effetto in senso statistico, e prelude ai modelli *SVAR* presentati nel quinto capitolo.

4.1 Introduzione storica ai modelli *SVAR*

I modelli *VAR* furono introdotti negli anni '80, in risposta alla critica di Lucas del 1976 che aveva evidenziato le fragilità metodologiche dei *modelli strutturali*, basati sui sistemi di equazioni simultanee (SEM, acronimo di *Structural Equation Modeling*).

Da un punto di vista concettuale, i *VAR* sono dei modelli statistici di serie storiche multiple dinamiche in cui ogni variabile viene posta in relazione con tutte le altre variabili ritardate di un certo numero di periodi, senza alcun vincolo da imporre a priori riguardo la specificazione dei parametri: l'obiettivo dei modelli strutturali era la traduzione delle relazioni economiche, basate sulla teoria, e quindi deterministiche per definizione, in equazioni statistiche, perciò stocastiche. Si procedeva dapprima nella stima empirica dei coefficienti delle variabili di interesse appartenenti al sistema economico analizzato; successivamente, ci si chiedeva se un'azione sulle variabili di *policy*, trattate come esogene, e quindi sotto il controllo dei *policy-maker*, avrebbe prodotto un effetto sulle variabili di interesse, e nel caso con che grado di intensità.

In formule, un modello strutturale può essere così rappresentato:

$$B\mathbf{Y}_t = d_t + \sum_{i=1}^p C_i \mathbf{Y}_{t-i} + \sum_{j=0}^q D_j \mathbf{X}_{t-i} + \mathbf{u}_t$$

dove con $\mathbf{Y}$ si indicano le variabili endogene, mentre con $\mathbf{X}$ le variabili esogene.

B indica le relazioni tra le variabili del vettore $\mathbf{Y}$; C_i e D_i sono matrici di parametri; d_t è un vettore che include variabili fisse; $\mathbf{u}_t$ è un errore *white noise* con media nulla ed errori non autocorrelati, $\mathbf{u}_t \sim (0, \Sigma_u)$.

²⁶ L'introduzione del presente capitolo è fedele al dibattito econometrico intrapreso da Lucas (1976) e proseguito da Sims (1980). Per l'analisi di causalità, si riprende Granger (1988).

Non si può procedere per una stima diretta, essendovi possibili elementi non nulli nella matrice B al di fuori della diagonale, che indica correlazione fra il termine di errore e le variabili esplicative, $E[\mathbf{u}_t \mathbf{X}_{t-i}] \neq 0$ (endogeneità): è necessario ricorrere alle informazioni contenute in altre variabili non incluse nella specificazione (strumenti), per ottenere lo stimatore a variabili strumentali, oppure servendosi di un approccio a due stadi (TSLS).²⁷

Bisognerà quindi passare ad un modello *in forma ridotta*, in cui tutte le variabili endogene sono espresse unicamente in funzione delle variabili esogene e dei parametri relativi al modello.

Passando in forma ridotta, il nostro modello strutturale diventa:

$$\mathbf{Y}_t = d_t + \sum_{i=1}^p A_i \mathbf{Y}_{t-i} + \sum_{j=0}^q D_j^* \mathbf{X}_{t-i} + \varepsilon_t$$

con $A_i = B^{-1}C_i$, $D_j^* = B^{-1}D_j$, $\varepsilon_t = B^{-1}u_t$. La forma ridotta può essere stimata come un comune modello dinamico.

Nel 1976, Lucas sottolinea come gli agenti economici abbiano un comportamento “*forward-looking*”: vale a dire, i valori delle variabili economiche di interesse al tempo t sono influenzate dalle aspettative sul futuro andamento dell'economia al tempo $t + 1$, formulate in base alle informazioni in loro possesso. L'adozione di nuove *policy* modifica perciò il contenuto informativo su cui si basa il processo decisionale, comportando una conseguente rivalutazione delle aspettative: il risultato è che uno studio condotto sugli effetti di determinate politiche economiche mediante un modello di equazioni simultanee risulterebbe essere fuorviante, poiché i parametri variano al variare della *policy*. Occorrerebbe conoscere i cosiddetti “*deep-parameters*”, relativi alle preferenze individuali, la tecnologia disponibile, i vincoli di spesa, per descrivere adeguatamente come gli individui formano le proprie aspettative.

Negli anni '80, Sims introduce i modelli *VAR* come risposta al “fallimento” dell'approccio tradizionale, criticando le restrizioni imposte dalla teoria economica senza alcun test empirico sui dati²⁸, seguendo un approccio diverso: partendo da un modello basato sui dati empirici e

²⁷ Intuitivamente, le variabili esplicative x sono caratterizzate da una variabilità “buona”, che non dipende dall'errore u_t , e da una variabilità “cattiva”, che invece dipende dall'errore u_t . La variabilità cattiva non consente di utilizzare OLS, rendendolo distorto; eliminandola, possiamo ottenere uno stimatore non distorto, dove lo strumento z_i è legato a x_i (rilevanza, $E[z_i x_i] \neq 0$), ma non a u_t (ipotesi di esogeneità). Teniamo solo la variabilità in x_i dovuta alla variabilità in z_i , che non dipende dal termine d'errore.

²⁸ Sull'argomento è famoso un esempio basato su equazioni simultanee circa la domanda e l'offerta di un prodotto agricolo. La temperatura è utilizzata come variabile strumentale esogena per identificare la domanda: l'ipotesi è che la temperatura non influenzi la domanda di prodotti agricoli direttamente. Se però considerassimo le

sulla teoria statistica, si identificano le relazioni tra le variabili di interesse. Tutte le variabili del sistema economico sono trattate come endogene, non essendovi nessuna supposizione a priori basantesi sulla teoria economica.

I modelli *VAR* non hanno l'obiettivo di descrivere l'intera economia in larga scala: l'interesse del ricercatore si concentra su un numero ristretto di variabili economiche X . Essi sono dei modelli in forma ridotta, ovverosia tutte le variabili correnti al tempo t sono poste in funzione delle variabili ritardate: non esiste perciò alcuna distinzione tra variabili endogene ed esogene.

Un *VAR* di ordine p , $Y_t = \sum_{i=1}^p A_i Y_{t-i} + \varepsilon_t$, con Y vettore ($n \times 1$) di variabili rilevanti, A_i matrici ($n \times n$) di coefficienti, ε_t errore *white-noise*, sfrutta le correlazioni nel sistema per fare *previsione*, *analisi di causalità in senso Grangeriano*, ed *analisi strutturale* in particolare, con cui è possibile dare una struttura economica al sistema, sfruttando una particolare tipologia dei processi *VAR*, ossia i modelli *SVAR* (*Structural VAR*) con cui si possono condurre analisi circa gli effetti di politiche economiche (*impulse response function*).

4.2 Causalità di Granger e causalità istantanea

Nel 1969, Granger introdusse il concetto di causalità, o, alternativamente, “*G-causality*”. In breve, l'idea generale è che *la causa di un effetto non può avvenire dopo l'effetto stesso*.

Se una variabile x influenza la variabile z , allora l'informazione contenuta nella variabile x dovrebbe migliorare la previsione della variabile z . Chiamando con Ω_t l'insieme delle informazioni disponibili sino all'istante t , possiamo definire $z_t(h|\Omega_t)$ come il miglior predittore della variabile z dopo h intervalli di tempo, dato il set Ω_t , minimizzando il *mean-squared error*, $MSE = \sum_t (h|\Omega_t)^{29}$.

aspettative, “pur essendo certi che le preferenze dei consumatori in America non siano influenzate dal clima in Brasile, è pur vero che i consumatori americani, dopo aver letto sul giornale di un'ondata di freddo che avrebbe investito il Brasile con gravi conseguenze sui raccolti, potrebbero decidere di accumulare scorte di caffè in previsione di un suo aumento del prezzo”. Perciò, le variabili che condizionano la curva di offerta si inseriscono in quelle di domanda, e viceversa, per tramite delle aspettative future. Da Sims (1980), p.6.

²⁹ Si definisce errore quadratico medio, o *mean-squared error* (*MSE*) di uno stimatore $\hat{\theta}$ rispetto al parametro θ , la differenza $MSE(\hat{\theta}) = E[(\hat{\theta} - \theta)^2]$, che si dimostra essere uguale alla somma della varianza e della distorsione dello stimatore rispetto al valore del parametro (*bias*). In formule, $MSE(\hat{\theta}) = Var(\hat{\theta}) + (Bias(\hat{\theta}, \theta))^2$. L'errore quadratico medio è perciò una misura per giudicare la qualità di uno stimatore, e precisamente di quanto i valori predetti si discostano rispetto ai valori osservati.

Definizione 4.1. In termini probabilistici, si dice che il processo $\mathbf{x}_t$ causa $\mathbf{z}_t$ in senso Grangeriano se:

$$\sum_t (h|\Omega_t) < \sum_t (h|\Omega_t\{\mathbf{x}_s: s \leq t\})$$

dove $\Omega_t\{\mathbf{x}_s: s \leq t\}$ contiene tutto l'insieme delle informazioni dell'universo Ω_t , eccetto il passato e presente del processo $\mathbf{x}_t$.

▪

Se $\mathbf{z}_t$ può essere predetto in modo più efficace con le informazioni aggiuntive contenute nel processo $\mathbf{x}_t$, allora si dice che $\mathbf{x}_t$ causa $\mathbf{z}_t$ in senso Grangeriano. Un risultato correlato e di immediata intuizione è quello per cui un processo $\mathbf{z}_t$ si definisce “non G-causato” da $\mathbf{x}_t$ se $\mathbf{z}_t(h|\Omega_t)$ non utilizza informazioni contenute nel processo $\mathbf{x}_t$.

Se $\mathbf{x}_t: (n \times 1)$ causa $\mathbf{z}_t: (m \times 1)$, e, a sua volta, $\mathbf{z}_t$ causa $\mathbf{x}_t$, allora il processo $(\mathbf{z}'_t, \mathbf{x}'_t)$ si definisce *sistema feedback*³⁰.

Definizione 4.2. Si dice che vi è *causalità istantanea* tra $\mathbf{z}_t$ e $\mathbf{x}_t$ se:

$$\sum_z (1|\Omega_t \cup \mathbf{x}_{t+1}) \neq \sum_z (1|\Omega_t)$$

▪

In altre parole, se aggiungiamo al periodo t l'informazione $\mathbf{x}_{t+1}$ al set Ω_t , essa migliora immediatamente la previsione di $\mathbf{z}$ al tempo $t + 1$, quindi $\mathbf{z}_{t+1}$. La causalità istantanea è anche simmetrica: se riscontrata tra $\mathbf{x}_t$ e $\mathbf{z}_t$, allora vale anche tra $\mathbf{z}_t$ e $\mathbf{x}_t$.

Si riscontrano alcune problematiche riguardo la definizione di causalità precedentemente data:

- è criticabile la scelta dell'errore quadratico medio come misura di precisione del forecasting. Ovviamente, parametri diversi di riferimento portano ad una diversa definizione di causalità;
- la scelta del set di informazioni Ω_t : in un mondo in cui non si ha accesso a tutte le informazioni disponibili, e quindi in presenza di asimmetrie informative, l'ottimo predittore di Ω_t non può essere determinato. È per tale motivo che usualmente il ricercatore utilizza una definizione più debole di causalità, sostituendo l'universo Ω_t

³⁰ Il termine *feedback* è utilizzato nel riferirsi a una situazione in cui due o più sistemi dinamici sono connessi in modo tale che ogni sistema influenza l'altro.

con le informazioni contenute nel passato e presente del processo studiato:
 $\{z_s, x_s | s \leq t\}$.

5. Modelli SVAR³¹

Usando le parole di Stock e Watson (2001), i macroeconometrici vogliono quattro cose:

- Descrivere serie storiche macroeconomiche;
- Fare previsioni;
- Comprendere grazie ai dati la vera struttura della macroeconomia;
- Dare suggerimenti ai *policy maker* sulle politiche da adottare.

I modelli *VAR* strutturali sono una famiglia di strumenti statistici in grado di soddisfare tali richieste.

5.1 Definizione dei processi SVAR

Definizione 5.1. Partendo da un processo K -dimensionale stazionario $VAR(p)$ del tipo:

$$\mathbf{x}_t = A_1 \mathbf{x}_{t-1} + \dots + A_p \mathbf{x}_{t-p} + \mathbf{u}_t$$

Si omette la componente deterministica ν del processo per studiare unicamente la componente stocastica del processo generatore di dati. Nella sua forma strutturale è definito come:

$$A\mathbf{x}_t = A_1^* \mathbf{x}_{t-1} + \dots + A_p^* \mathbf{x}_{t-p} + B\boldsymbol{\varepsilon}_t$$

assumendo che gli errori strutturali $\boldsymbol{\varepsilon}_t$ siano *white-noise* e che i coefficienti delle matrici A_i^* per $i = 1, \dots, p$, differiscano dalla forma ridotta. Si noti che i residui in forma ridotta possono essere calcolati da un modello *SVAR* come $\mathbf{u}_t = A^{-1}B\boldsymbol{\varepsilon}_t$ e la sua matrice di varianze e covarianze come $E[\mathbf{u}_t \mathbf{u}_t'] = \Sigma_{\mathbf{u}} = A^{-1}BB^T A^{-1^T}$.

■

Le matrici dei coefficienti sono diverse rispetto alla forma ridotta, e per questo le indichiamo con l'asterisco. Se moltiplicassimo per la matrice inversa di A , risulterebbe:

$$\mathbf{x}_t = A^{-1} \cdot A_1^* \mathbf{x}_{t-1} + \dots + A^{-1} \cdot A_p^* \mathbf{x}_{t-p} + A^{-1} \cdot B\boldsymbol{\varepsilon}_t$$

$$\mathbf{x}_t = A_1 \mathbf{x}_{t-1} + \dots + A_p \mathbf{x}_{t-p} + \mathbf{u}_t$$

³¹ Si ringraziano Fernández-Villaverde, Rubio-Ramírez (2010); Lütkepohl (2005); Pfaff (2008); Rubio-Ramírez, J.F., Waggoner, D.F., Zha, T. (2010); Stock, Watson (2001); Uhlig (2005).

5.2 Il problema di identificazione

Nei modelli *SVAR* sorge un *problema di identificazione*: nel passaggio dalla forma ridotta, non possiamo stimare in modo unico i parametri strutturali, poiché ci sono più parametri nella forma strutturale che in quella ridotta. Tuttavia, se supponessimo di avere a disposizione alcune informazioni circa il sistema strutturale, potremmo utilizzarle per identificare lo *SVAR* dalla forma ridotta.

Definizione 5.2. Possono distinguersi tre diversi tipi di modelli *SVAR*³²:

- *Modello A*: la matrice B è identica, I_k (il numero minimo di restrizioni per l'identificazione è $\frac{K(K-1)}{2}$);
- *Modello B*: la matrice A è identica, I_k (il numero minimo di restrizioni per l'identificazione è lo stesso del modello A);
- *Modello AB*: restrizioni su entrambe le matrici (il minimo numero di restrizioni per l'identificazione è $K^2 + \frac{K(K-1)}{2}$).

■

L'idea di base di Sims era quella di trattare tutte le variabili come endogene, stimare un modello senza restrizioni in forma ridotta, e con stime OLS (che si dimostra essere consistenti) identificare gli *shock* strutturali. Imporre un certo ordine è fondamentale, giacché in una rappresentazione a media mobile (*MA*) gli *shock* strutturali non sono in relazione con gli altri simultaneamente. Un esempio rende chiari i concetti esposti.

Supponiamo di avere tre variabili: Δy , la crescita reale del PIL; π , l'inflazione; i , il tasso di

policy. Se x_t è un vettore (3×1) tale che: $x_t = \begin{bmatrix} \Delta y_1 & \Delta y_2 & \dots & \Delta y_T \\ \pi_1 & \pi_2 & \dots & \pi_T \\ r_1 & r_2 & \dots & r_T \end{bmatrix}$.

Un *SVAR*(1) stazionario è:

$$Ax_t = A_1^* x_{t-1} + \varepsilon_t$$

con $B = I_k$, A e A_1^* matrici (3×3), ε_t vettore (3×1) di processi *white noise* non osservabili.

Si può rappresentare in forma matriciale o come sistema di equazioni lineari:

³² Si intende definire il fenomeno nei termini utili ai nostri obiettivi espositivi, esulando diversi aspetti tecnici. Esaustivi, in questo senso, Lütkepohl (2005) e Rubio-Ramírez, J.F., Waggoner, D.F., Zha, T. (2010).

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} \Delta y_t \\ \pi_t \\ r_t \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix} \begin{bmatrix} \Delta y_{t-1} \\ \pi_{t-1} \\ r_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{\Delta y_t} \\ \varepsilon_{\pi_t} \\ \varepsilon_{r_t} \end{bmatrix}$$

$$\begin{cases} a_{11}\Delta y_t + a_{12}\pi_t + a_{13}r_t = b_{11}\Delta y_{t-1} + b_{12}\pi_{t-1} + b_{13}r_{t-1} + \varepsilon_{\Delta y_t} \\ a_{21}\Delta y_t + a_{22}\pi_t + a_{23}r_t = b_{21}\Delta y_{t-1} + b_{22}\pi_{t-1} + b_{23}r_{t-1} + \varepsilon_{\pi_t} \\ a_{31}\Delta y_t + a_{32}\pi_t + a_{33}r_t = b_{31}\Delta y_{t-1} + b_{32}\pi_{t-1} + b_{33}r_{t-1} + \varepsilon_{r_t} \end{cases}$$

Essendo $\boldsymbol{\varepsilon}_t = (\varepsilon'_{\Delta y_t}, \varepsilon'_{\pi_t}, \varepsilon'_{r_t}) \sim N(0, I)$, si può leggere $\varepsilon_{\Delta y_t}$ come *shock* aggregato; ε_{π_t} come *shock* “cost-push”³³; ε_{r_t} come *shock* della politica monetaria.

La matrice A dà problemi poiché include tutte le relazioni contemporanee tra le altre variabili endogene. La soluzione è moltiplicare per A^{-1} la forma ridotta del modello, trasferendo la *dipendenza contemporanea* delle variabili endogene (data da A) negli errori “modificati” $\mathbf{u}_t$. Ora la correlazione di $\mathbf{u}_t$ è diversa da I : $\mathbf{u}_t \sim N(0, \Sigma_u)$.

Non possiamo interpretare degli errori in forma ridotta $\mathbf{u}$ come fossero *shock* strutturali. Essendo combinazioni lineari di diversi *shock* $\boldsymbol{\varepsilon}$, è difficile dire quale sia la sua natura.

Si prenda una forma ridotta del tipo:

$$\mathbf{x}_t = A_1 \mathbf{x}_{t-1} + \mathbf{u}_t, \quad A_1 = A_1^* A^{-1}, \quad \mathbf{u}_t = A^{-1} B \boldsymbol{\varepsilon}_t$$

E si voglia giungere alla sua forma strutturale:

$$A \mathbf{x}_t = A_1^* \mathbf{x}_{t-1} + B \boldsymbol{\varepsilon}_t$$

Sapendo che:

$$\sum_{\mathbf{u}} = E[\mathbf{u}_t \mathbf{u}_t'] = E[A^{-1} \boldsymbol{\varepsilon} (A^{-1} \boldsymbol{\varepsilon})'] = A^{-1} \sum_{\boldsymbol{\varepsilon}} (A^{-1})' = A^{-1} A^{-1'}$$

poiché $\sum_{\boldsymbol{\varepsilon}} = I$. Conoscendo A^{-1} abbiamo finito, giacché $\boldsymbol{\varepsilon}_t = A \mathbf{u}_t$, $A_1 = A_1^* A^{-1}$.

Scrivendo $\sum_{\mathbf{u}} = A^{-1} A^{-1'}$ in forma matriciale,

³³ L’inflazione *cost-push* è un tipo di inflazione causata dalla crescita del costo di beni e servizi di grande importanza (ad esempio, la crisi petrolifera di cui ha avuto esperienza l’Occidente negli anni ’70).

$$\Sigma_u = E[\mathbf{u}_t \mathbf{u}_t'] \equiv \begin{bmatrix} \sigma_{u1}^2 & \sigma_{u1,u2}^2 & \sigma_{u1,u3}^2 \\ - & \sigma_{u2}^2 & \sigma_{u2,u3}^2 \\ - & - & \sigma_{u3}^2 \end{bmatrix} = A^{-1} A^{-1'} \equiv \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}^{-1} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}^{-1'} \quad 34$$

Possiamo derivare un sistema di equazioni. Ma ci sono 9 incognite (gli elementi di $A^{-1}A^{-1'}$) e solamente 6 equazioni (poiché la matrice di varianza-covarianza è simmetrica): pertanto, il sistema non è identificato.

Esistono diversi schemi di identificazione che ovviano al problema. Qui studiamo il più conosciuto: la *decomposizione di Cholesky*.

Definizione 5.3. Una matrice X simmetrica e semidefinita positiva può essere scomposta come:

$$X = \mathbf{P}'\mathbf{P}$$

con $\mathbf{P}$ matrice triangolare superiore e $\mathbf{P}'$ matrice triangolare inferiore. In formule, $X = \begin{bmatrix} a & b \\ b & c \end{bmatrix} =$

$$\begin{bmatrix} \sqrt{a} & \frac{b}{\sqrt{a}} \\ 0 & \sqrt{c - \frac{b^2}{a}} \end{bmatrix}.$$

■

Assumiamo che A sia triangolare inferiore:

$$\begin{bmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} \Delta y_t \\ \pi_t \\ r_t \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix} \begin{bmatrix} \Delta y_{t-1} \\ \pi_{t-1} \\ r_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{\Delta y_t} \\ \varepsilon_{\pi_t} \\ \varepsilon_{r_t} \end{bmatrix}$$

In altre parole, $\varepsilon_{\Delta y_t}$ impatta simultaneamente tutte le variabili, $\Delta y_t, \pi_t, r_t$; ε_{π_t} impatta simultaneamente su π_t, r_t , non su Δy_t ; ε_{r_t} impatta solo su r_t . Adesso si hanno 6 incognite e 6 equazioni.

Le assunzioni si applicano anche alla sua inversa, $A^{-1} = \begin{bmatrix} \hat{a}_{11} & 0 & 0 \\ \hat{a}_{21} & \hat{a}_{22} & 0 \\ \hat{a}_{31} & \hat{a}_{32} & \hat{a}_{33} \end{bmatrix}$. Se moltiplichiamo per A^{-1} , diventa:

³⁴ La varianza di $\mathbf{x} = [x_0, \dots, x_t]$ è $\sum_{t=0}^T \frac{(x_t - \bar{x})^2}{N}$. Poiché i residui hanno media 0, $\Gamma_0 = \sum_{t=0}^T \frac{x_t^2}{N}$. In un VAR bivariato

$$\text{avremmo: } \mathbf{u}_t \mathbf{u}_t' = \begin{bmatrix} u_1^1 & u_1^2 & \dots & u_1^T \\ u_2^1 & u_2^2 & \dots & u_2^T \\ \vdots & \vdots & \ddots & \vdots \\ u_T^1 & u_T^2 & \dots & u_T^T \end{bmatrix} \begin{bmatrix} u_1^1 & u_1^2 \\ u_2^1 & u_2^2 \\ \vdots & \vdots \\ u_T^1 & u_T^2 \end{bmatrix} = \begin{bmatrix} \sum_{t=0}^T (u_t^1)^2 & \sum_{t=0}^T (u_t^1 u_t^2) \\ - & \sum_{t=0}^T (u_t^2)^2 \end{bmatrix}. \quad \text{Perciò, } E[\mathbf{u}_t \mathbf{u}_t'] =$$

$$\begin{bmatrix} \text{VAR}[u^1] & \text{COV}[u^1 u^2] \\ - & \text{VAR}[u^2] \end{bmatrix}.$$

$$\begin{cases} \Delta y_t = \dots + \hat{a}_{11}\varepsilon_{\Delta y_t} \\ \pi_t = \dots + \hat{a}_{21}\varepsilon_{\Delta y_t} + \hat{a}_{22}\varepsilon_{\pi_t} \\ r_t = \dots + \hat{a}_{31}\varepsilon_{\Delta y_t} + \hat{a}_{32}\varepsilon_{\pi_t} + \hat{a}_{33}\varepsilon_{r_t} \end{cases}$$

Si implementa questo schema di identificazione via decomposizione di Cholesky di Σ_u :

$$\Sigma_u = \mathbf{P}'\mathbf{P}$$

con $\mathbf{P}'$ triangolare inferiore. Si noti che $\Sigma_u = A^{-1}A^{-1'}$ e che A è triangolare inferiore. Segue che $\mathbf{P}' = A^{-1}$, e c'è identificazione.

5.3 Funzione di risposta all'impulso

La *G-causality* potrebbe non rivelarci tutte le interazioni tra le variabili del sistema. È spesso di interesse conoscere la *risposta* di una variabile all'*impulso* di un'altra variabile in un sistema multivariato, come accennato nel paragrafo 2.3.

Definizione 5.4. Una funzione di risposta all'impulso (*IRF*, *Impulse-Response Function*) descrive l'evoluzione di una variabile di interesse lungo un precipuo orizzonte temporale continuo, $0 < t < +\infty$, successivamente ad uno *shock* avvenuto ad un determinato istante.

▪

Una *IRF* risponde alla domanda di quale sia la risposta nei valori correnti e futuri in ognuna delle variabili in seguito ad un incremento unitario di uno degli *shock* strutturali, assumendo ritorni zero nei successivi periodi e che gli altri errori rimangano zero (analisi dell'effetto *moltiplicativo*). Ad esempio, in un *VAR* bivariato $\mathbf{x}_t = (x'_{1t}, x_{2t}^1)$, si definisce un vettore di impulsi esogeni $\mathbf{s}_t$ che vogliamo imporre agli *shock* strutturali del sistema:

Tempo t	1	2	h
Impulso a $\varepsilon_1 (s_{1,t})$	$s_{1,1} = 1$	$s_{1,2} = 0$	$s_{1,h} = 0$
Impulso a $\varepsilon_2 (s_{2,t})$	$s_{2,1} = 0$	$s_{2,2} = 0$	$s_{1,h} = 0$

Come precedentemente illustrato, nel caso univariato *IRF* è $\frac{dx_t}{du_{t-j}}$. Se, anziché uno scalare, x_t fosse un vettore colonna con dimensioni $n \times 1$, allora avremmo una rappresentazione ibrida di un processo *VAR* (1), del tipo:

$$\mathbf{x}_t = A\mathbf{x}_{t-1} + A^{-1}\mathbf{s}_t$$

La IRF è data da $\begin{cases} IRF_1 = A^{-1}\mathbf{s}_1 \\ IRF_t = A \cdot IRF_{t-1} \end{cases}, t = 2, \dots, h$. La porzione della deviazione di $x_{i,t}$ dalla sua media dovuta agli *shock* strutturali ε_i si ottiene riscrivendo le osservazioni dei dati come somma cumulativa di *shock* strutturali.

5.4 Un'applicazione SVAR

I modelli *SVAR* consentono ai *policy-maker* di prevedere i cambiamenti macroeconomici. Diamone una dimostrazione, utilizzando dati della Federal Reserve.

Prendiamo in considerazione tre serie storiche: UNRATE (la disoccupazione); FEDFUNDS (i tassi di interesse con cui le banche in eccesso di liquidità prestano denaro alle altre con depositi overnight³⁵); CPIAUCSL (indice dei prezzi al consumo dei cittadini urbani, ossia la media ponderata di un paniere di beni e servizi che, confrontata nel tempo, permette di riconoscere periodi di inflazione e deflazione).

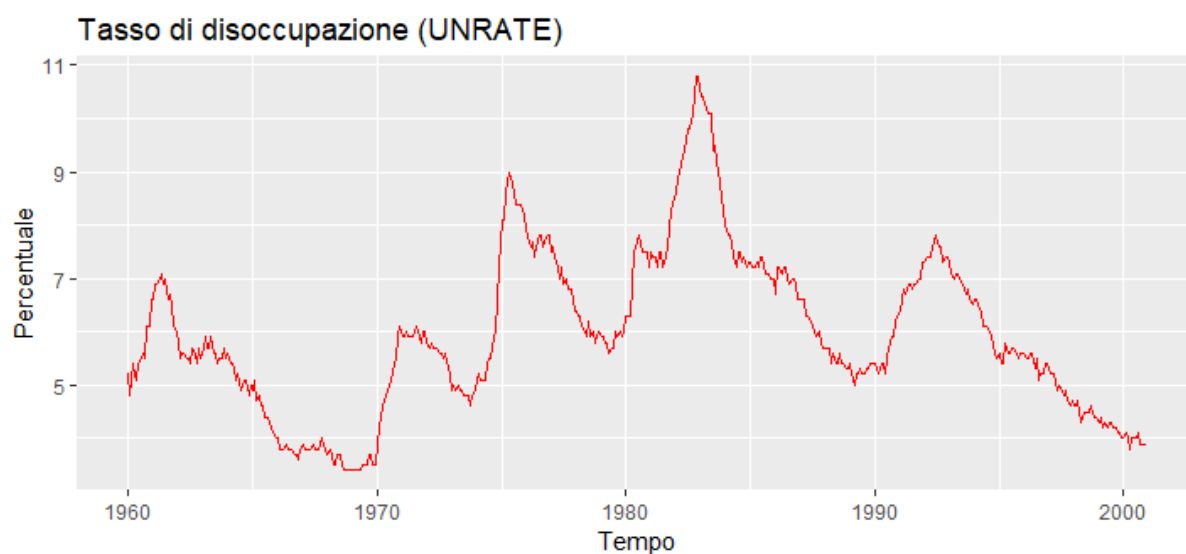


Fig.5.1. Tasso di disoccupazione (UNRATE) dal 1960 a fine 2000.

³⁵ Questo tipo di tassi di interesse è determinato tra le banche che danno a prestito e quelle che lo ricevono. Tuttavia, è largamente influenzato dalla FED tramite operazioni a mercato aperto (acquisizione o vendita di titoli di stato), accrescendo o diminuendo la liquidità comprando o vendendo *bond*. Se l'economia cresce troppo velocemente, per contenere l'inflazione sono auspicabili tassi più alti, e viceversa. Si guarda con grande interesse a questo tipo di tassi giacché influenzano altri tassi a lungo termine come mutui, prestiti e risparmi dei consumatori.

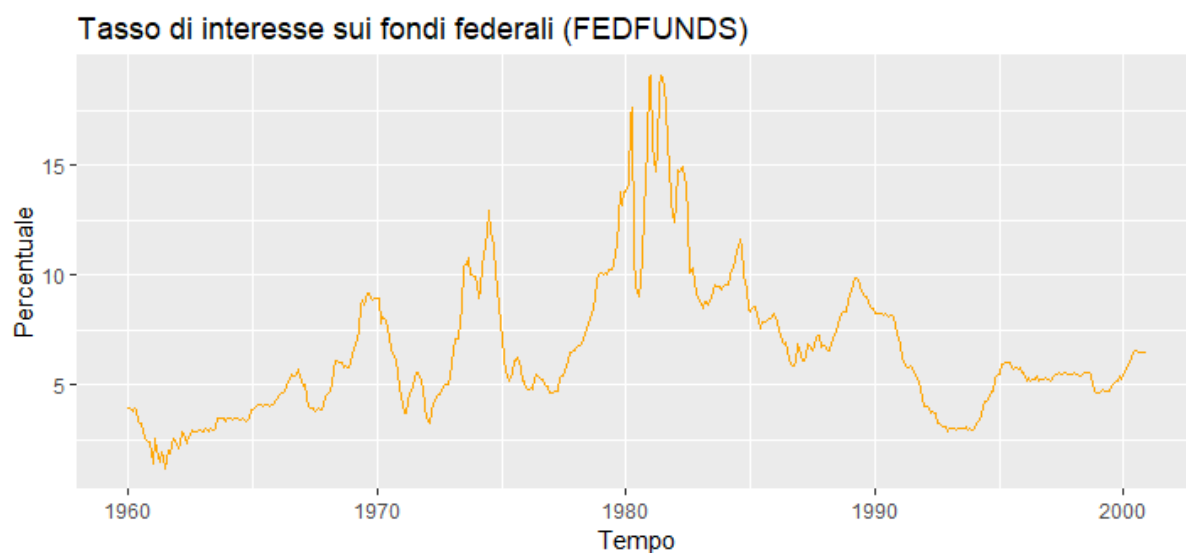


Fig.5.2. Tasso di interesse effettivo sul prestito di fondi federali dal 1960 a fine 2000.

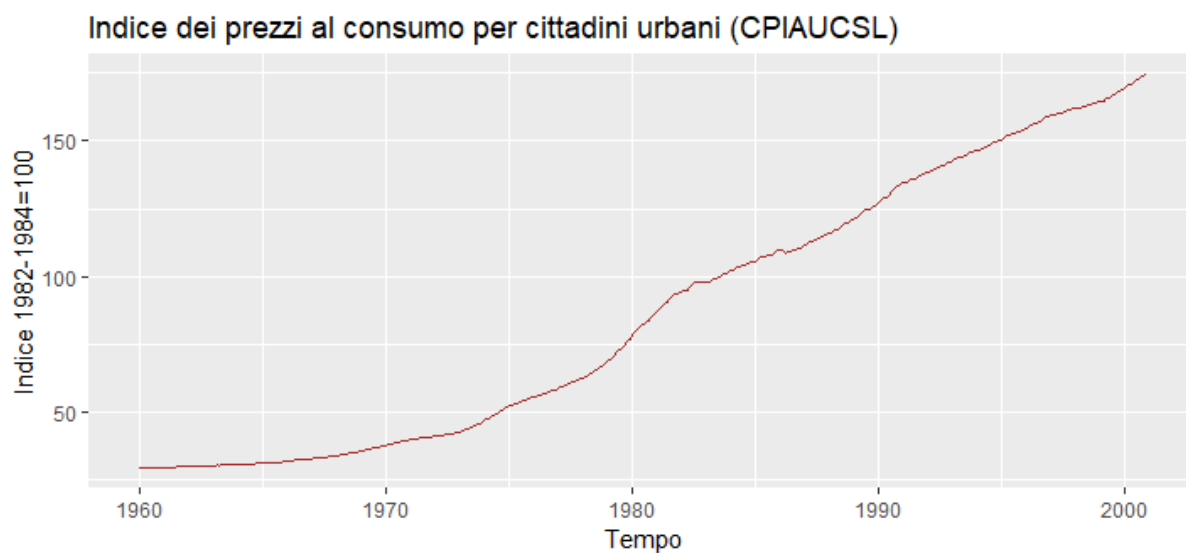


Fig.5.3. Indice dei prezzi al consumo per cittadini urbani dal 1960 a fine 2000.

Le funzioni di autocorrelazione e autocorrelazione correlazione (ACF e PACF)³⁶ sono:

³⁶ ACF, al lag k , esprime la correlazione tra osservazioni separate da k intervalli; PACF, al lag k , esprime la correlazione tra osservazioni separate da k intervalli, tenendo conto dei valori nel mezzo dell'intervallo.

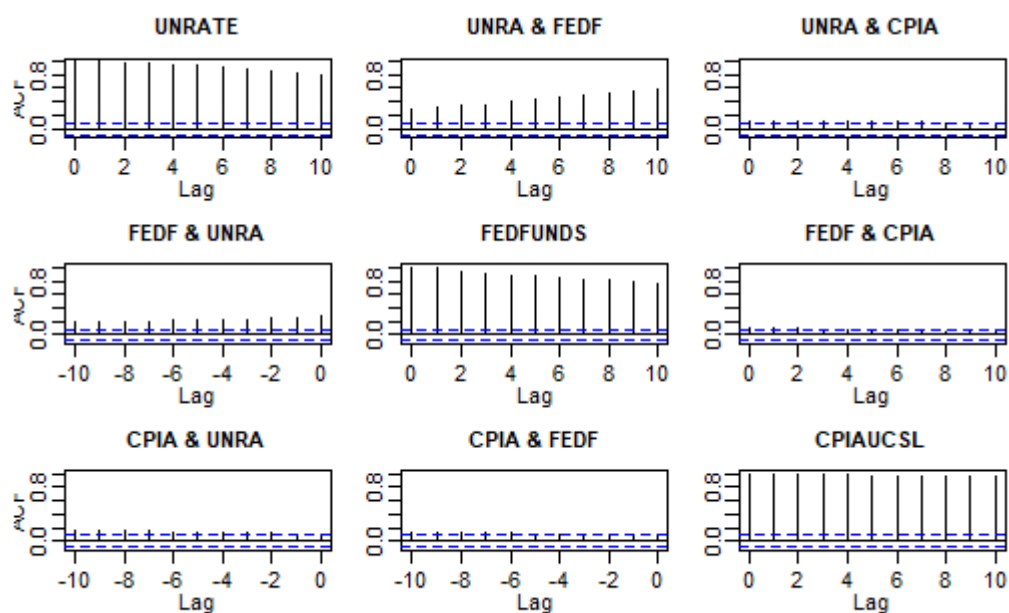


Fig.5.4. Funzione ACF.

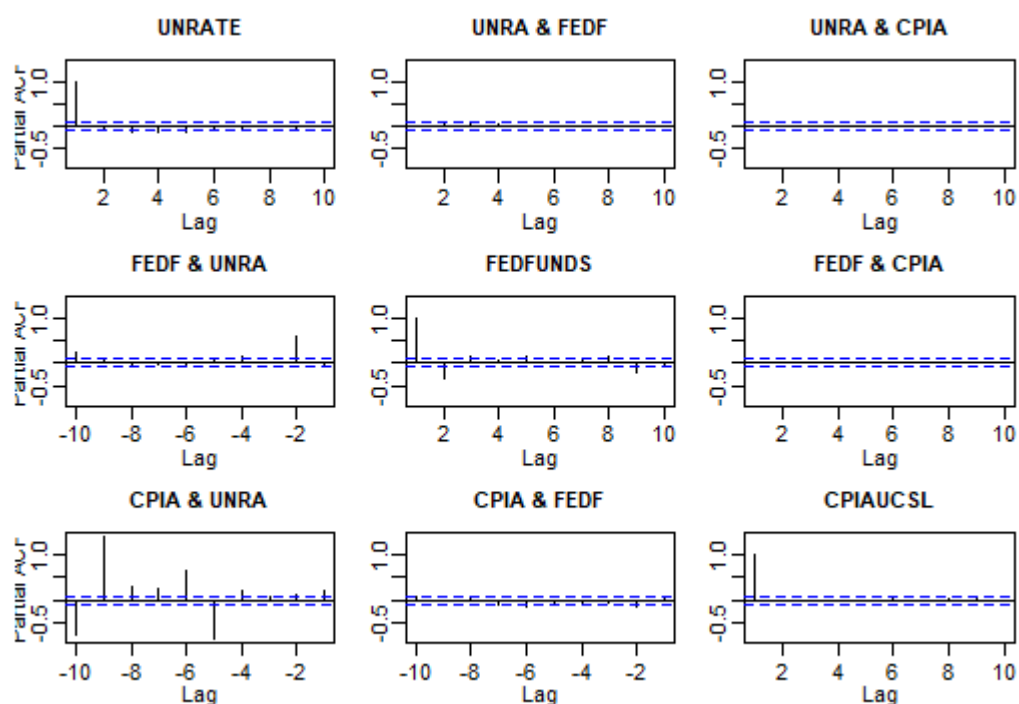


Fig.5.5. Funzione PACF.

Costruiamo un modello VAR. Per scegliere appropriatamente la lunghezza del ρ -esimo *lag*, confrontiamo diversi criteri che minimizzano il numero di ritardi: ciò risulta utile nel caso in cui ve ne siano diversi con valori molto simili, e si voglia garantire parsimonia informativa.

AIC(n)	HQ(n)	SC(n)	FPE(n)
12	5	2	12

Scegliamo lo *Schwartz Criterion* (ossia, il BIC), che penalizza in maggior misura rispetto all'*Akaike Information Criterion* (AIC) il numero di parametri che possono condurre ad *overfitting*. Si simulano *IRF* con metodo *bootstrap*³⁷, e sono:

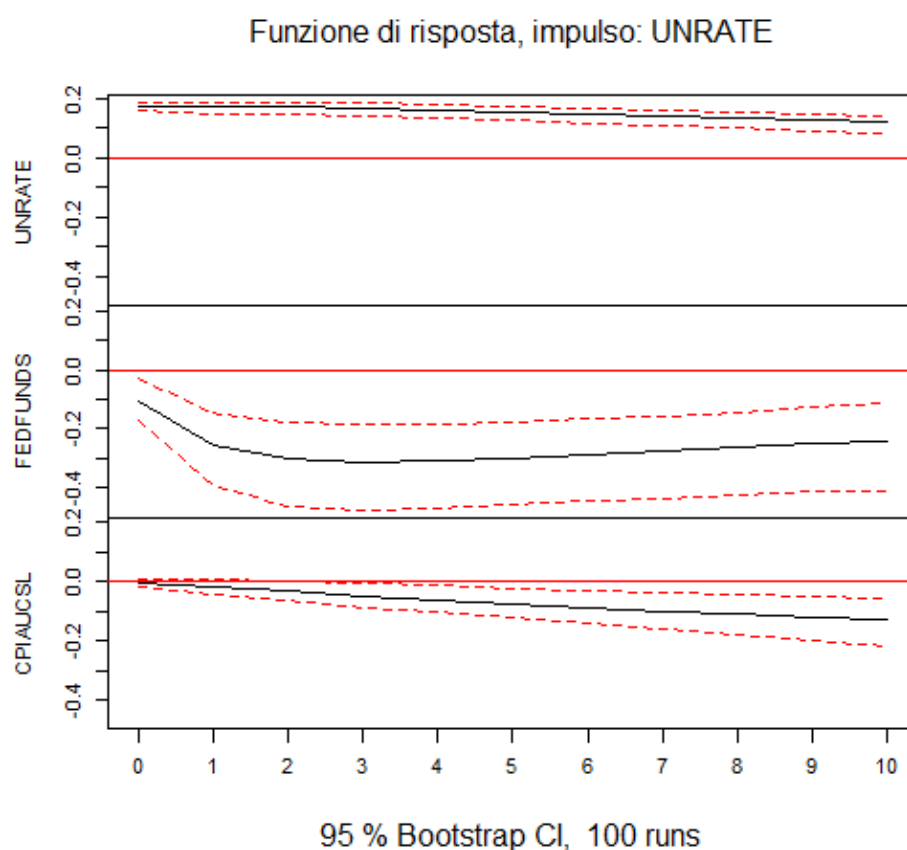


Fig.5.6. Funzione di risposta all'impulso del tasso di disoccupazione.

³⁷ Il *metodo bootstrap* è una tecnica di simulazione condizionata che prevede l'uso di osservazioni reali come base delle simulazioni. Viene introdotto da Efron nel 1979 come metodo informatico al fine di determinare l'accuratezza della stima di una statistica: egli propone la stima dell'*errore standard* (e, in pratica, intervalli di confidenza) introducendo il campione bootstrap $\mathbf{x}^* = (x_1^*, \dots, x_n^*)$, ottenuto estraendo n volte dal campione originario (con ripetizione, quindi trascurando l'ordine di estrazione) $\mathbf{x} = (x_1, \dots, x_n)$.

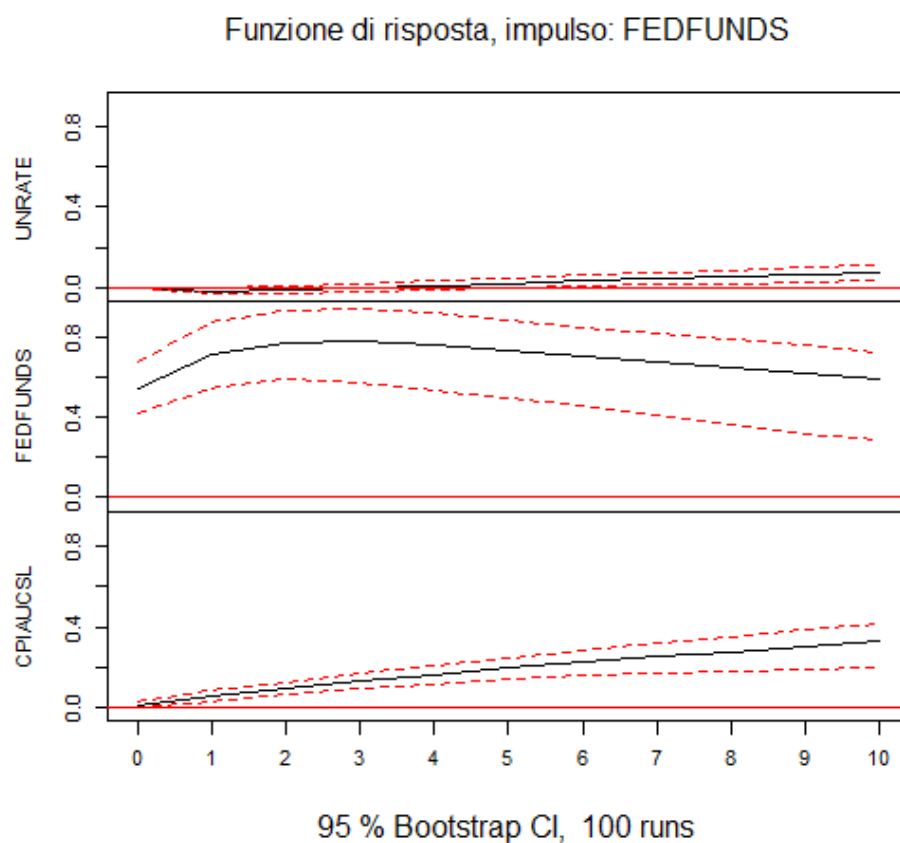


Fig.5.7. Funzione di risposta all'impulso del tasso di interesse sui fondi federali.

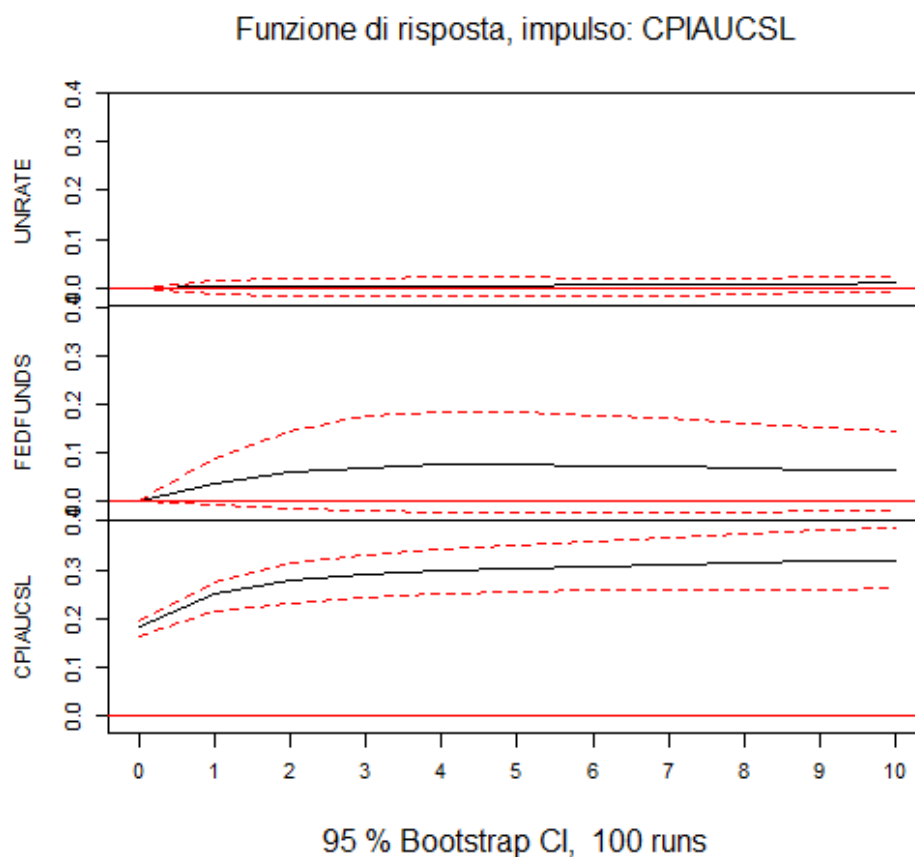


Fig.5.8. Funzione di risposta all'impulso dell'indice dei prezzi al consumo.

Stimando il *VAR* così costruito, si riscontra presenza di radici unitarie: pertanto, il processo non è stabile. Non abbiamo interesse nei risultati così ottenuti giacché le *IRF* non convergono a zero, implicando che gli *shock* hanno effetti permanenti, e ciò è irragionevole in molti casi.

Costruiamo ora un modello *VAR* con serie differenziate, scegliendo anzitutto l'ordine di p .

AIC(n)	HQ(n)	SC(n)	FPE(n)
12	4	2	12

Il BIC è 2. Le stime corrispondenti sono riportate in appendice. ACF e PACF sono:

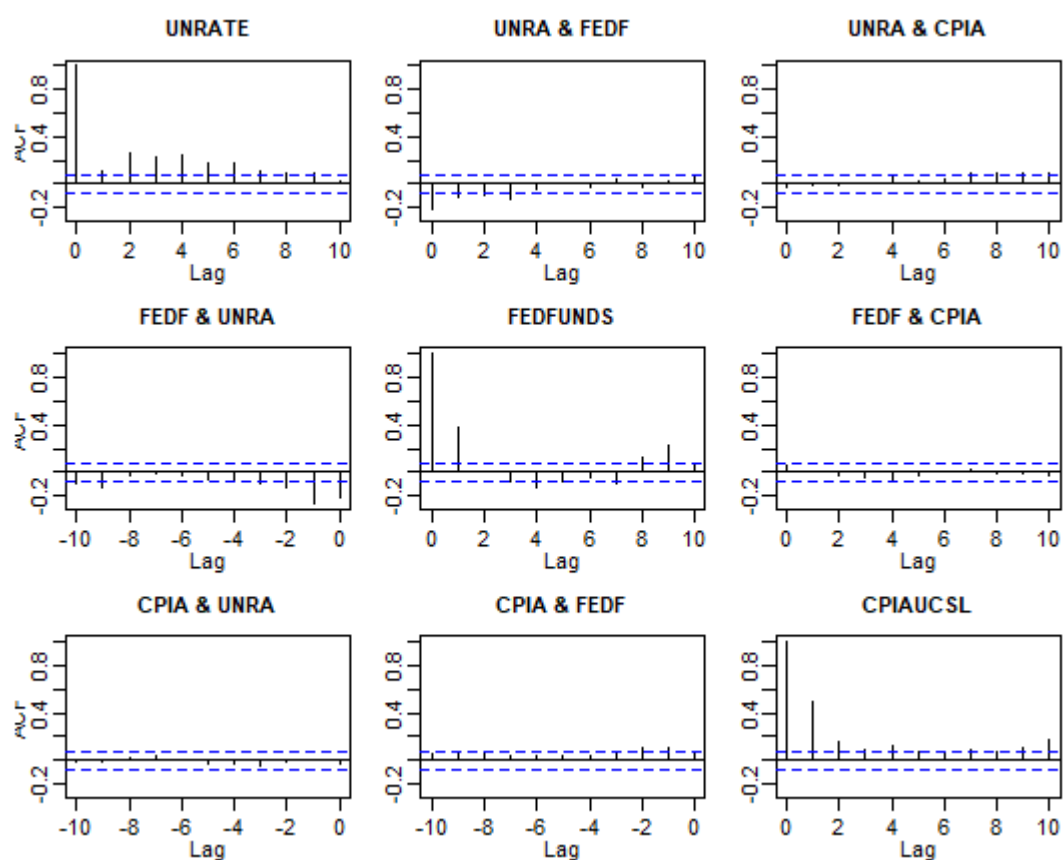


Fig.5.9. Funzione ACF.

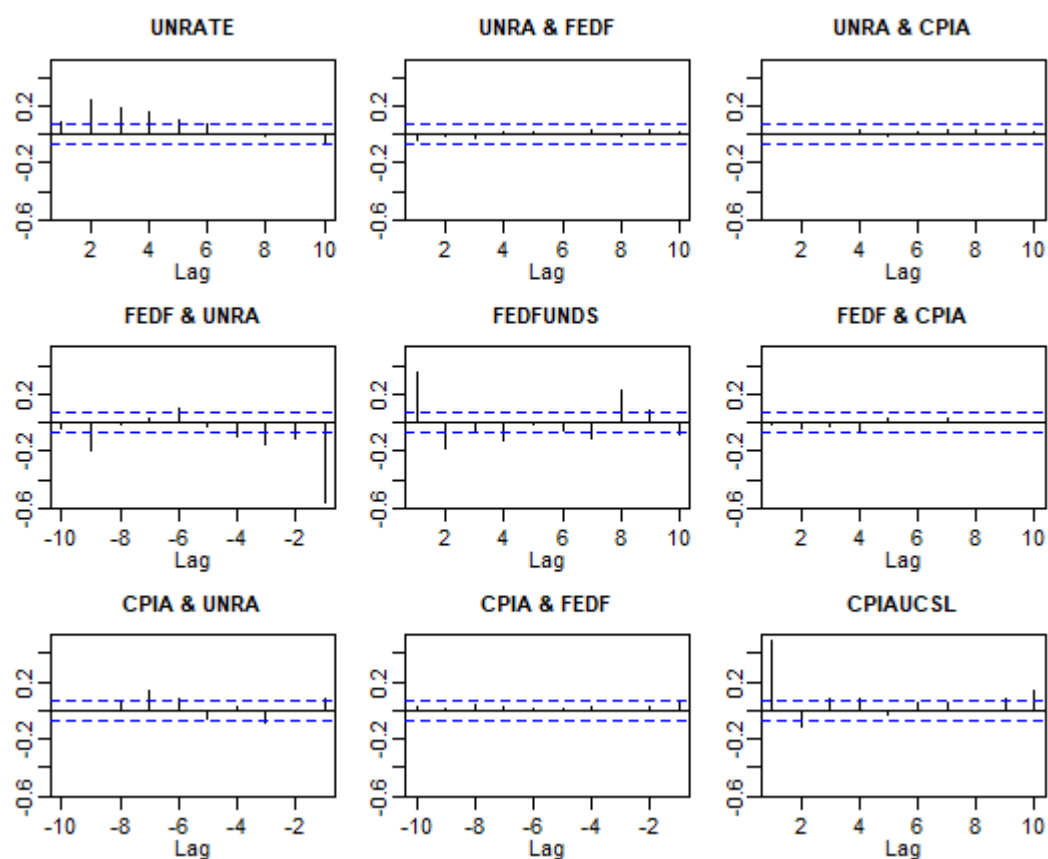


Fig.5.10. Funzione PACF.

Le funzioni di risposta all'impulso sono:

Funzione di risposta, impulso: UNRATE

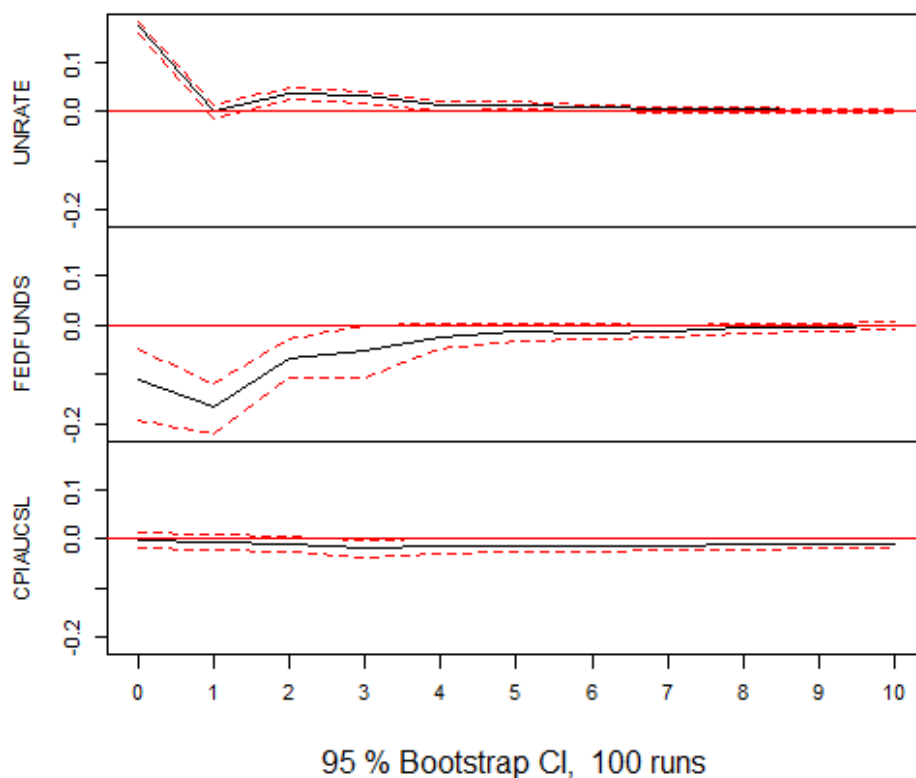


Fig.5.11. Funzione di risposta all'impulso della disoccupazione.

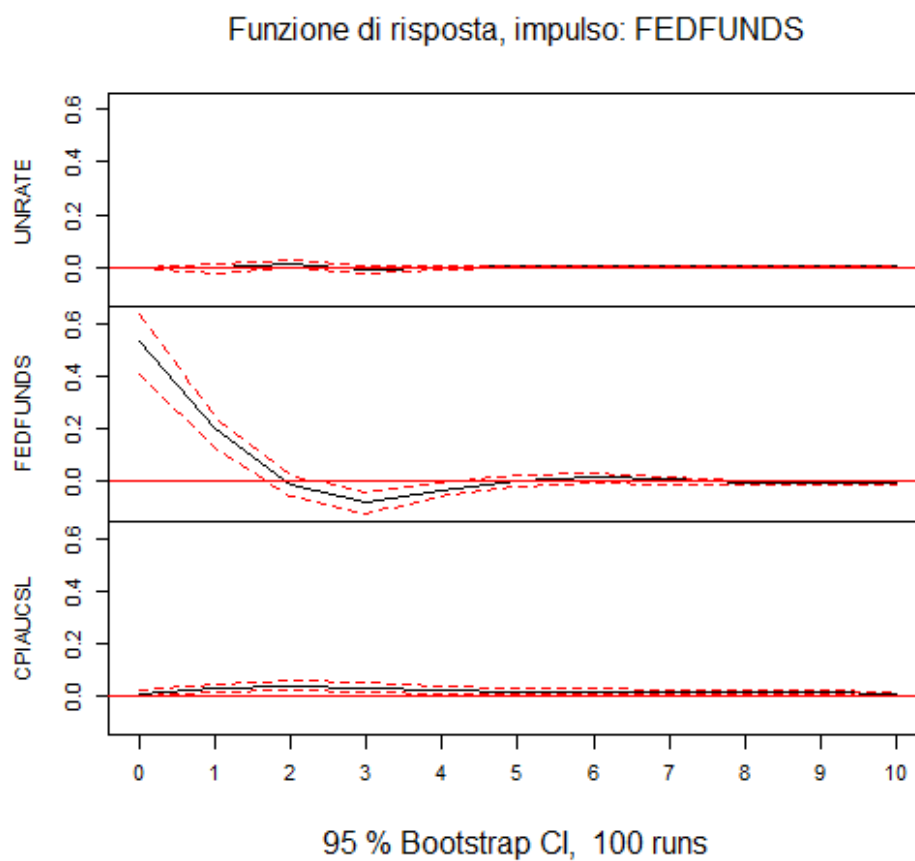


Fig.5.12. Funzione di risposta all'impulso del tasso di interesse sui fondi federali.

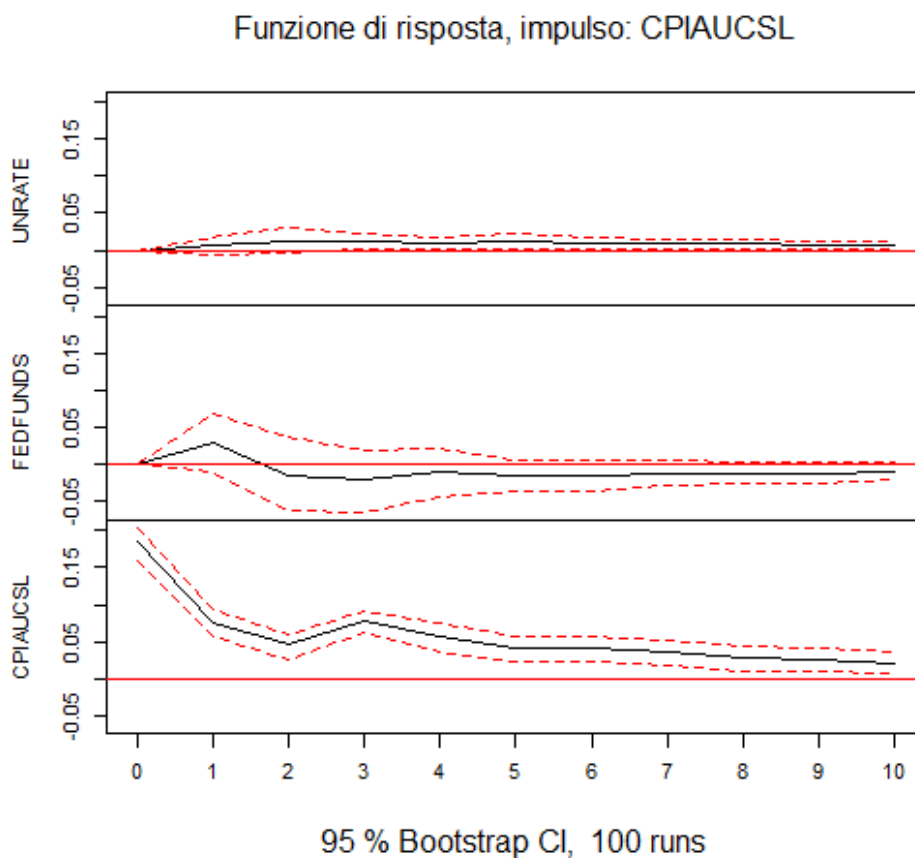


Fig.5.13. Funzione di risposta all'impulso dell'indice dei prezzi al consumo.

Interpretando le diverse funzioni di risposta all'impulso per il modello VAR con serie differenziate, si nota che:

- Quando la disoccupazione sale, il tasso di interesse tende a decrescere nel tempo (interessi più bassi per stimolare l'economia e quindi l'occupazione); l'impatto sull'inflazione non è significativo.
- Quando il tasso di interesse effettivo sul prestito di fondi federali sale, c'è una modesta crescita dell'inflazione (in linea con la policy della FED di aumentare i tassi per contenerla); l'impatto sulla disoccupazione non è significativo.
- Quando l'inflazione sale, la disoccupazione tende a crescere leggermente, così come i tassi sui fondi FED, in modo più deciso.

Con uno SVAR (*A model*), ed A matrice triangolare inferiore, si stima:

$$\hat{A} = \begin{pmatrix} & U & P & I \\ U & 1.00000 & 0.00000 & 0 \\ P & -0.02167 & 1 & 0 \\ I & 0.49220 & -0.09676 & 1 \end{pmatrix}$$

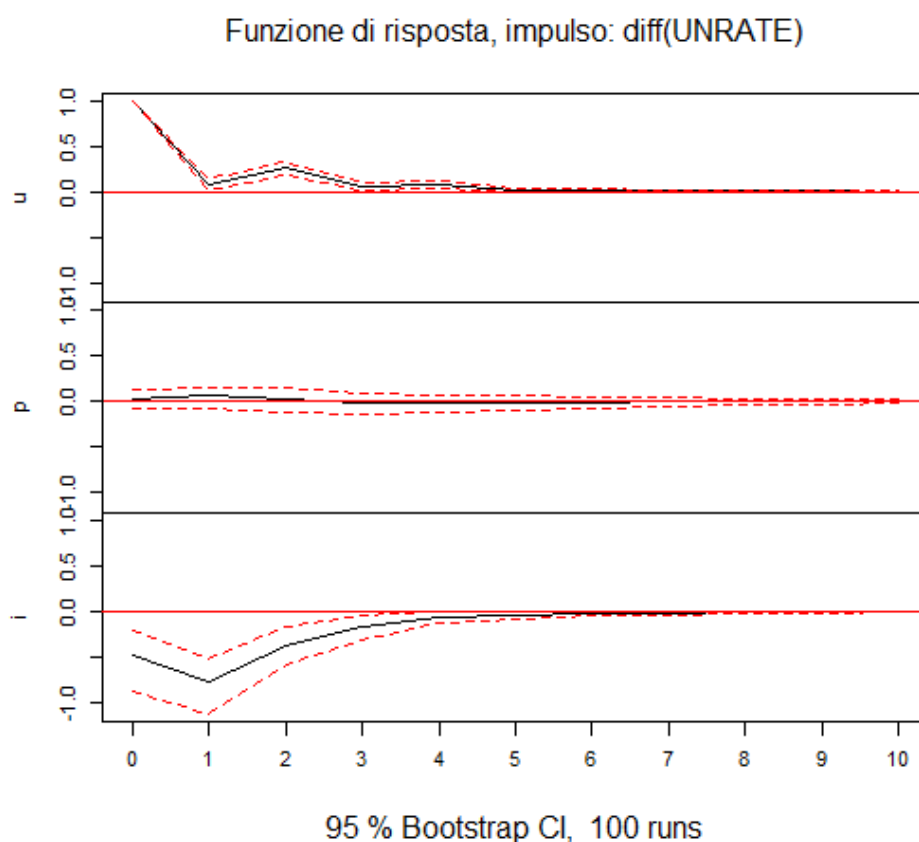


Fig.5.14. Funzione di risposta all'impulso diff(UNRATE), SVAR.

Funzione di risposta, impulso: diff(FEDFUNDS)

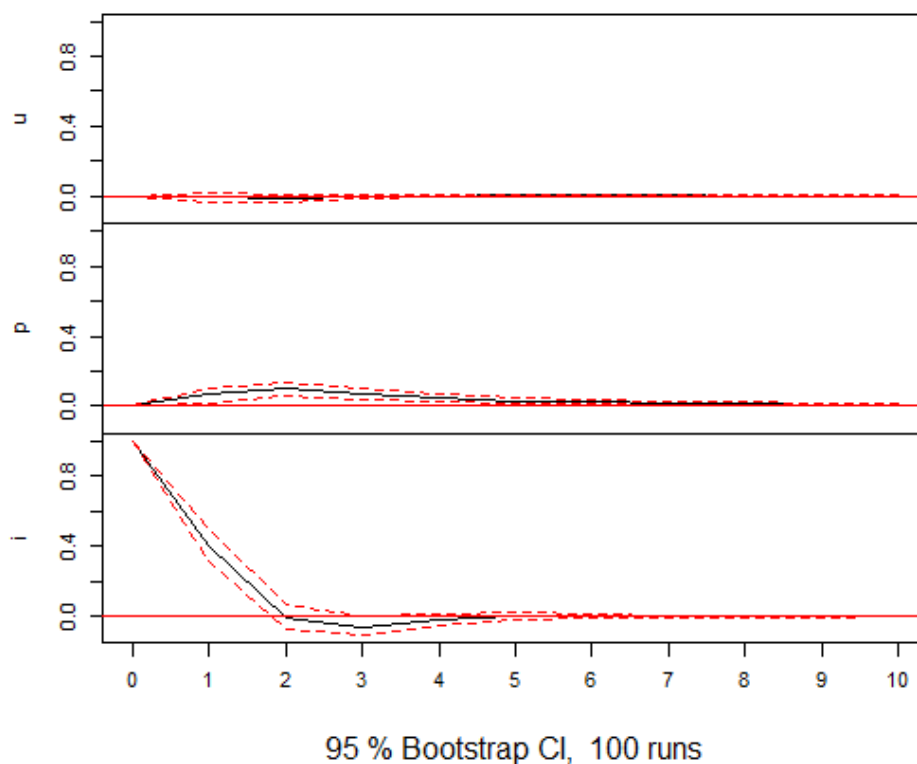


Fig.5.15. Funzione di risposta all'impulso diff(FEDFUNDS), SVAR.

Funzione di risposta, impulso: diff(CPIAUCSL)

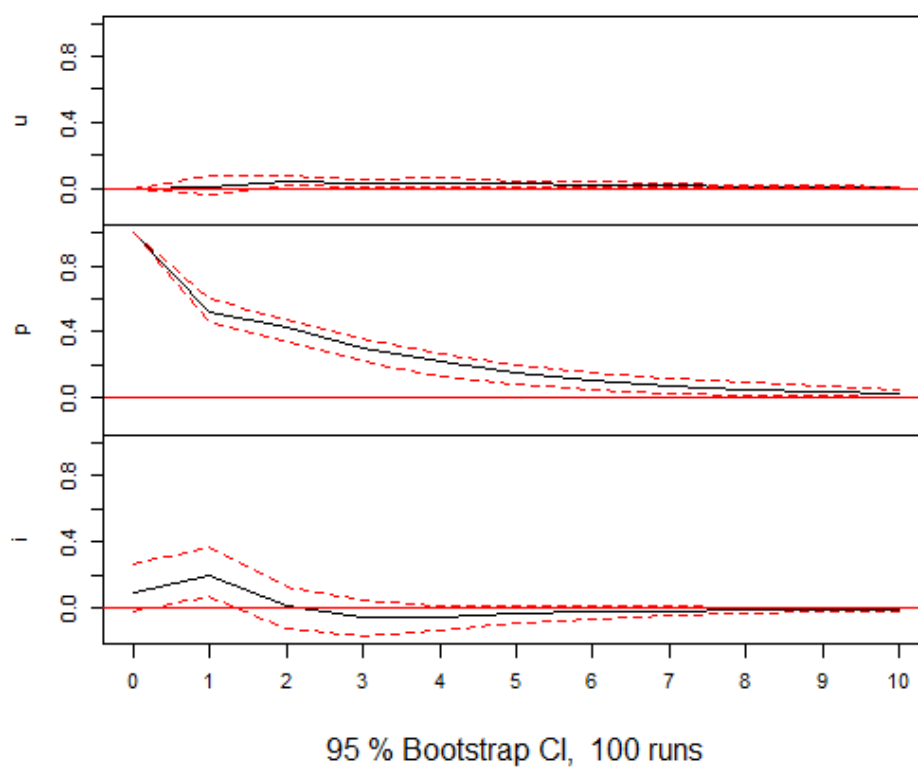


Fig.5.16. Funzione di risposta all'impulso diff(CPIAUCSL), SVAR.

5.5 Conclusioni sui modelli SVAR

I modelli *SVAR* offrono un approccio attraente per il processo di stima: richiedono un numero minimo di assunzioni per l'identificazione parametrica e sono relativamente semplici da utilizzare grazie all'utilizzo di software commerciali e dataset liberamente scaricabili dal web. Nelle mani di ricercatori particolarmente abili, i modelli *SVAR* hanno contribuito a comprendere le fluttuazioni macroeconomiche, l'importanza dei diversi *shock* economici, generando fortunati e fruttuosi confronti e dibattiti tra i macroeconomisti.

Tuttavia, i modelli *SVAR* sono stati anche criticati, essenzialmente per tre aspetti, riportati di seguito:

- Si è argomentato che gli *shock* economici analizzati mediante modelli *SVAR* non assomigliano agli *shock* misurati da altri meccanismi, come quello delle aspettative future sull'andamento dei prezzi.
- Gli *shock* simulati da un modello *SVAR* possono riflettere variabili omesse dal modello, e se esse sono correlate con le variabili incluse, la stima degli *shock* sarà distorta.
- I risultati *SVAR* sono particolarmente sensibili alle restrizioni di identificazione. In merito a tale critica c'è l'opinione secondo cui gli schemi di identificazione sono un risultato forzato, costruiti dai ricercatori di modo da ottenere delle risposte "ragionevoli". Citando Uhlig (2005), "se uno schema di identificazione incontra la saggezza convenzionale, allora esso sarà un successo; altrimenti, esso assomiglierà ad un puzzle, ad un fallimento": il rischio è che gli economisti diano una personale visione dei dati con assunzioni aprioristiche, dietro "il mantello della statistica inferenziale".

Parte terza

Conclusioni

Una convinzione ferrea di Granger era che l'econometria, teorica e applicata, dovesse avere un valore diretto verso i *policy-makers*.

In questo lavoro, si è ricostruito il *background* concettuale attuale dell'analisi economica, riprendendo diversi risultati dal cuore della letteratura odierna riguardo le serie storiche: tratteggiandone il volto, si è successivamente posta l'attenzione sia a situazioni singole che dinamiche, applicando le nozioni esposte con un esempio bivariato sulla struttura a termine dei tassi di interesse; da qui ci siamo calati nel campo dell'economia pura, studiando come gli strumenti descritti potessero risultare utili a tale scopo. Si è simulato un modello a tre variabili per osservare l'effetto nel tempo di *shock* esogeni nelle singole variabili, e come essi si ripercuotono nel sistema, sotto certe condizioni.

A

Stabilità e stazionarietà dei sistemi dinamici³⁸

A.1 Sistemi di equazioni lineari del primo ordine

Si osservi il seguente *sistema di equazioni differenziali lineari omogenee di primo ordine*:

$$\begin{cases} y'_1 = a_{11}y_1 + a_{12}y_2 + \dots + a_{1n}y_n \\ y'_2 = a_{21}y_1 + a_{22}y_2 + \dots + a_{2n}y_n \\ \vdots \\ y'_n = a_{n1}y_1 + a_{n2}y_2 + \dots + a_{nn}y_n \end{cases}$$

per $x \in I$, dove I è un intervallo aperto (ossia con estremi non compresi) e le funzioni $a_{ij} \in C(I)$ sono assegnate. Posto:

$$A(x) = \begin{pmatrix} a_{11}(x) & \dots & a_{1n}(x) \\ \vdots & \ddots & \vdots \\ a_{n1}(x) & \dots & a_{nn}(x) \end{pmatrix}, \quad x \in I$$

il sistema si può riscrivere in forma compatta come

$$\mathbf{y}' = A(x)\mathbf{y}, \quad x \in I.$$

La combinazione lineare delle soluzioni $\mathbf{y}_i$, $\mathbf{z} := \alpha_1\mathbf{y}_1 + \dots + \alpha_m\mathbf{y}_m$, è l'unica soluzione del problema di Cauchy:

$$\begin{cases} \mathbf{z}' = A(x)\mathbf{z} \\ \mathbf{z}(x_0) = \mathbf{0} \end{cases}$$

Più specificamente, m soluzioni $\mathbf{y}_1, \dots, \mathbf{y}_m$ si dicono linearmente indipendenti se:

$$\alpha_1\mathbf{y}_1(x) + \dots + \alpha_m\mathbf{y}_m(x) = \mathbf{0} \quad \text{per ogni } x \in I \Rightarrow \alpha_1 = \dots = \alpha_m = 0$$

Si può ora caratterizzare la *soluzione generale* (o *integrale generale*) del sistema $\mathbf{y}' = A(x)\mathbf{y}$:

$$\tilde{\mathbf{y}} \text{ è soluzione di } \mathbf{y}' = A(x)\mathbf{y} \text{ in } I \Leftrightarrow \tilde{\mathbf{y}}(x) = \sum_{i=1}^n a_i \mathbf{y}_i(x) \text{ se } x \in I$$

Se la matrice $A(x)$ non dipende da x , $A(x) = A$, è possibile determinare esplicitamente n soluzioni linearmente indipendenti di $\mathbf{y}' = A\mathbf{y}$. L'idea generale è di cercare n soluzioni del tipo $\mathbf{y}_i(x) = e^{\lambda_j x} \mathbf{v}_j$, con λ_j e $\mathbf{v}_j \in \mathbb{R}^n - \{0\}$ scelti opportunamente.

Sostituendo nel sistema $\mathbf{y}' = A\mathbf{y}$ e dividendo per $e^{\lambda_j x}$ si ottiene:

³⁸ Si veda Bertsch (2011) e Sasso (2012). La figura A.1 è tratta da Smith (2002).

$$\lambda_j \mathbf{v}_j = A \mathbf{v}_j \Leftrightarrow (A - \lambda_j \text{Id}) = 0$$

dove Id è la matrice identità, con 1 sulla diagonale e 0 altrimenti:

$$A - \lambda \text{Id} = \begin{pmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{pmatrix}$$

Quindi λ_i è un autovalore della matrice A e $\mathbf{v}_j$ un autovettore associato. Ma un autovalore (ovverosia una soluzione dell'equazione caratteristica $\det(A - \lambda \text{Id}) = 0$, un'equazione polinomiale di grado n) non è necessariamente un numero reale. Nel caso di due equazioni:

$$\begin{cases} u' = au + bv \\ v' = cu + dv \end{cases} \quad \text{per } x \in R, \text{ ovvero } A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Sarà:

$$\begin{aligned} \det(A - \lambda \text{Id}) &= \det \begin{pmatrix} a - \lambda & b \\ c & d - \lambda \end{pmatrix} = (a - \lambda)(d - \lambda) - bc = ad - \lambda a - \lambda d + \lambda^2 \\ &= \lambda^2 - (a + d)\lambda + ad - bc = 0 \end{aligned}$$

Si potranno avere:

- due soluzioni reali e distinte, λ_1 e λ_2 ;
- due soluzioni complesse coniugate, $\lambda = \mu \pm i\omega \in C$, ($\omega \neq 0$);
- una sola soluzione reale, $\lambda \in R$.

A.2 Il concetto di stabilità

Consideriamo l'equazione differenziale del primo ordine:

$$y' = f(y) \quad \text{in } R$$

In questo caso, f non è in funzione di x , pertanto l'equazione si definisce *autonoma*.

Se a è uno zero di f , tale che $f(a) = 0$, la funzione costante $y(t) = a$, con $t \in R$ è soluzione della nostra EDO; essa è anche detta *soluzione stazionaria*. Il termine “stazionaria” fa riferimento a una situazione dinamica, ovvero a un problema che dipende dal tempo t . Pertanto, prendiamo come variabile indipendente t .

Intuitivamente, si può dire che:

- Un punto y_0 si dice punto di equilibrio *stabile* se il sistema ritorna allo stato di equilibrio dopo una piccola perturbazione.

- Un punto y_0 si dice punto di equilibrio *instabile* se il sistema si allontana dallo stato di equilibrio dopo una piccola perturbazione.

Si dimostra che:

- Se esiste $\varepsilon_0 > 0$ tale che $f \geq 0$ in $[a - \varepsilon_0, a]$ e $f \leq 0$ in $(a, a + \varepsilon_0]$ allora $y = a$ è stabile;
- Se esiste $\varepsilon_0 > 0$ tale che $f > 0$ in $[a - \varepsilon_0, a]$ e $f < 0$ in $(a, a + \varepsilon_0]$ allora $y = a$ è asintoticamente stabile;
- Se esiste $\varepsilon_0 > 0$ tale che $f < 0$ in $[a - \varepsilon_0, a]$ e $f > 0$ in $(a, a + \varepsilon_0]$ allora $y = a$ è instabile.

Nel caso si abbiano delle radici caratteristiche minori o maggiori di 1 in modulo, l'immagine seguente mostra che il processo risulta stazionario o esplosivo.

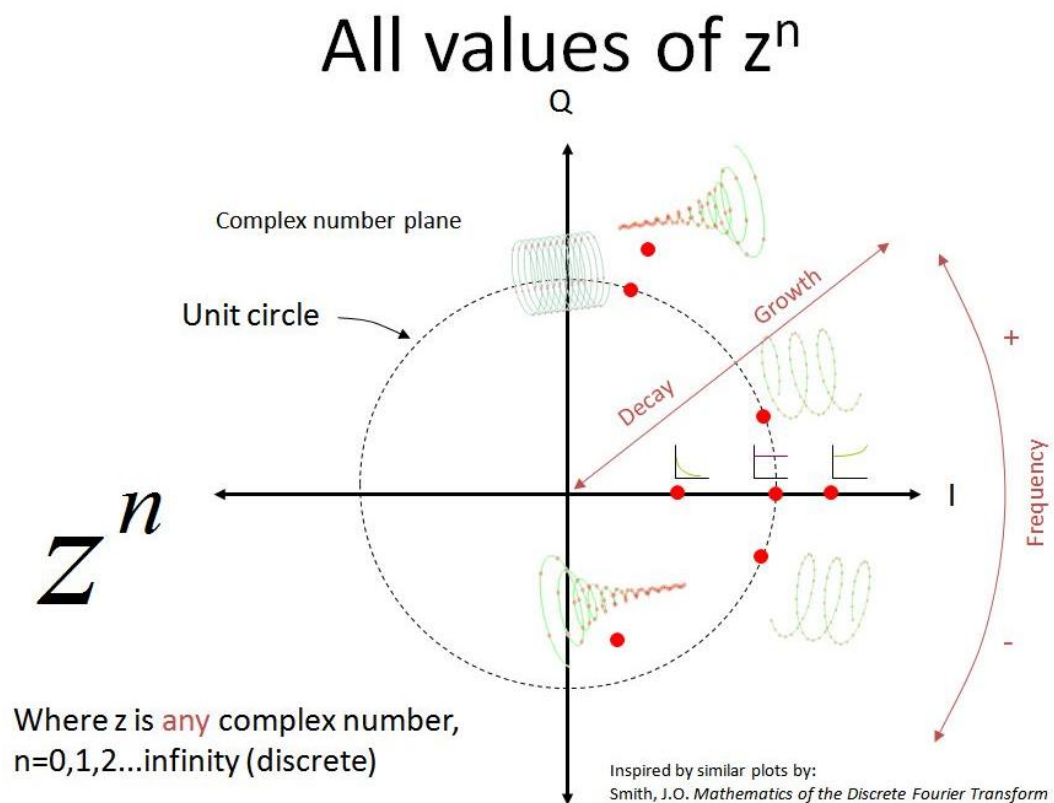


Fig.A.1. Comportamento del processo dentro e fuori il cerchio unitario.

B

Cenni sui numeri complessi³⁹

Il presupposto su cui nasce tale appendice riguardante i numeri complessi si rintraccia nella volontà di rilevare alcuni aspetti della funzione di risposta all'impulso.

Vi sono semplici equazioni che in R non hanno soluzioni: $x^2 + 1 = 0$, ad esempio. È necessario estendere ulteriormente il campo dei numeri reali introducendo un nuovo insieme numerico, l'insieme dei *numeri complessi*, tale che le operazioni ivi definite continuino a godere delle proprietà valide in R .

B.1 Costruzione e terminologia dei numeri complessi

Si introducano le seguenti operazioni per l'insieme delle coppie (a, b) di numeri reali:

$(a, b) + (c, d) = (a + c, b + d)$, operazione di addizione;

$(a, b) \cdot (c, d) = (ac - bd, ad + bc)$, operazione di moltiplicazione.

Chiameremo l'insieme R^2 , strutturato con queste operazioni, insieme dei numeri complessi, e lo denoteremo con la lettera C . Un numero complesso è quindi una *coppia ordinata* di numeri reali. A differenza di R , qui non è definibile una relazione d'ordine tale che:

- Per ogni $x, y, z \in C$, se $x < y \Rightarrow x + z < y + z$;
- Per ogni $x, y, z \in C$, con $z > 0$, se $x < y \Rightarrow xz < yz$.

Definendo l'unità immaginaria i come il numero complesso $(0, 1)$, risulta possibile scrivere ogni numero complesso nella forma:

$$(a, b) = (a, 0) + (0, b) = (a, 0) + (b, 0) \cdot (0, 1)^{40} = a + bi,$$

detta forma algebrica di un numero complesso.

Un'importante proprietà del numero complesso i è che il quadrato coincide con l'opposto dell'unità del campo C :

$$(0, 1) \cdot (0, 1) = (0 - 1, 0 + 0) = (-1, 0)$$

Semplificando la notazione, scrivendo semplicemente a anziché $(a, 0)$, otteniamo:

$$i^2 = -1.$$

³⁹ Per approfondimenti, Bertsch (2011) e Sasso (2012).

⁴⁰ $(b, 0) \cdot (0, 1) = (b \cdot 0 - 0 \cdot 1, b \cdot 1 + 0 \cdot 0) = (0 - 0, b + 0) = (0, b)$.

La rappresentazione *in forma algebrica* di un generico numero complesso (a, b) , indicato con z , è quindi:

$$z = a + bi,$$

di cui il numero a è la parte *reale* del numero complesso, mentre b ne è la parte *immaginaria*.

Definiamo inoltre *coniugato* un numero complesso con la parte immaginaria invertita di segno. In forma algebrica:

$$\bar{z} = a - bi$$

Si osserva che $z \cdot \bar{z}$ è per forza un numero reale positivo:

$$z \cdot \bar{z} = (a + bi)(a - bi) = a^2 - (b^2 i^2) = a^2 + b^2.$$

B.2 Rappresentazione grafica dei numeri complessi

Una naturale rappresentazione geometrica dei numeri complessi è nel piano cartesiano, detto anche piano di Gauss, con l'asse delle ascisse detto *asse reale*, mentre l'asse delle ordinate definito come *asse immaginario*.

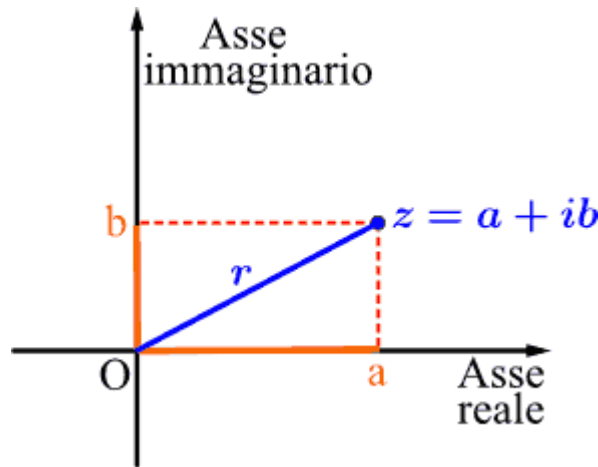


Fig.A.1. Rappresentazione di un piano di Gauss.

Un diverso tipo di rappresentazione è con un *sistema di coordinate polari*: esso prevede di fissare un punto, detto *polo*, ed una semiretta avente come origine il polo, detta *asse polare*. Allora, un punto P può ancora essere identificato come una coppia ordinata di numeri reali, ma al posto dell'ascissa e dell'ordinata, si scrive:

$$P = (r, \theta)$$

con r che identifica la distanza di P dal polo (modulo), e θ che rappresenta l'angolo formato dall'asse polare con la semiretta $\overline{OP}$ (argomento).

Applicando il primo teorema sui triangoli rettangoli⁴¹, si ottiene:

$$x = r \cos \theta, \quad y = r \sin \theta$$

Sia $z = a + bi$ e $P(a, b)$ un punto nel piano di Gauss. Usando le coordinate polari, possiamo riscrivere z come:

$$z = r(\cos \theta + i \sin \theta),$$

con $r \geq 0$ e $0 \leq \theta \leq 2\pi$. In tal caso il numero complesso si dice espresso in *forma trigonometrica*. Si dimostra che, dati z_1 e z_2 , per la moltiplicazione:

$$z_1 z_2 = r_1 r_2 [\cos(\theta_1 + \theta_2) + i \sin(\theta_1 + \theta_2)]$$

mentre per la divisione:

$$\frac{z_1}{z_2} = \frac{r_1}{r_2} \cdot [\cos(\theta_1 - \theta_2) + i \sin(\theta_1 - \theta_2)]$$

B.3 Potenze e radici in C

Per le *potenze*, vale il Teorema di De Moivre:

$$z^n = r^n [\cos n\theta + i \sin n\theta]$$

Applicando De Moivre, possiamo trovare le radici n -esime di un numero complesso. Per radice n -esima, dato un numero complesso w ed un intero positivo n , si intende un numero z tale che:

$$z^n = w$$

Se conosciamo la forma trigonometrica di w , che risulta:

$$w = \rho(\cos \varphi + i \sin \varphi)$$

Le radici n -esime di w sono i numeri complessi $z = r(\cos \theta + i \sin \theta)$ tali che $z^n = w$. Con De Moivre tale condizione diventa:

$$r^n(\cos n\theta + i \sin n\theta) = \rho(\cos \varphi + i \sin \varphi)$$

B.4 Forma esponenziale dei numeri complessi

Per agevolare i calcoli, un'altra forma dei numeri complessi molto utilizzata è quella esponenziale. Si dimostra che un numero immaginario puro può essere scritto come:

⁴¹ Un cateto è uguale al prodotto dell'ipotenusa per il seno dell'angolo opposto, $\sin \theta$, o per il coseno dell'angolo adiacente, $\cos \theta$.

$$e^{i\theta} = \cos \theta + i \sin \theta$$

Riprendendo la forma trigonometrica dei numeri complessi,

$$z = r(\cos \theta + i \sin \theta)$$

è chiaro che:

$$z = re^{i\theta}$$

detta *forma esponenziale* di un numero complesso. Essa ci consente di dare una scrittura più semplice alle operazioni di moltiplicazione, divisione, ed alla formula di De Moivre. In ordine:

$$r_1 e^{i\theta_1} \cdot r_2 e^{i\theta_2} = r_1 r_2 e^{i(\theta_1 + \theta_2)}$$

$$\frac{r_1 e^{i\theta_1}}{r_2 e^{i\theta_2}} = \frac{r_1}{r_2} \cdot e^{i(\theta_1 - \theta_2)}$$

$$(re^{i\theta})^n = r^n e^{in\theta}$$

La prima uguaglianza, $e^{i\theta} = \cos \theta + i \sin \theta$, è detta anche *prima formula di Eulero*. Ponendo $-\theta$, si ricava:

$$e^{-i\theta} = \cos \theta - i \sin \theta$$

che è invece la *seconda formula di Eulero*.

Se le sommiamo, e dividiamo per due, otteniamo la *terza*:

$$\cos \theta = \frac{e^{i\theta} + e^{-i\theta}}{2}$$

Se le sottraiamo, sempre dividendo per due, otteniamo la *quarta*:

$$\sin \theta = \frac{e^{i\theta} - e^{-i\theta}}{2i}$$

B.5 Interpretazione statistico-economica

Se l'ordine di un processo autoregressivo $AR(p)$ è maggiore di 1, allora tali processi possono avere un andamento *ciclico* se fra le radici del polinomio $A(z)$ c'è una coppia di numeri complessi coniugati. In tal caso, i processi assumono un andamento ciclico, e le fasi oscillano intorno ad un valore medio. Essi risultano ottimi per descrivere cicli economici. Supponiamo di avere un polinomio in L di ordine 2 con radici complesse coniugate:

$$A(L) = 1 - \varphi_1 L - \varphi_2 L^2 = (1 - zL)(1 - \bar{z}L)$$

Posto $|z| < 1$ per assicurare l'invertibilità, scriviamo:

$$A(L)^{-1} = (1 - zL)^{-1}(1 - \bar{z}L)^{-1} = (1 + zL + z^2L^2 + \dots)(1 + \bar{z}L + \bar{z}^2L^2 + \dots)^{42}$$

che equivale a:

$$A(L)^{-1} = 1 + (z + \bar{z})L + (z^2 + z\bar{z} + \bar{z}^2) + (z^3 + z^2\bar{z} + \bar{z}^2z + \bar{z}^3)L^3 + \dots$$

e, più in generale:

$$A(L)^{-1} = 1 + \theta_1 L + \theta_2 L^2 + \dots$$

con $\theta_k = \sum_{j=0}^k z^j z^{k-j}$.

È interessante notare che:

$$\theta_{k+1} = z^{k+1} + \bar{z}\theta_k = \bar{z}^{k+1} + z\theta_k$$

da cui si ricava:

$$z^{k+1} - \bar{z}^{k+1} = (z - \bar{z})\theta_k$$

Ora effettuiamo alcune semplici mutazioni algebriche. Ricordando che:

$$\sin \theta = \frac{e^{i\theta} - e^{-i\theta}}{2i}$$

diventa:

$$r^k \sin(\theta k) = \frac{z^k - \bar{z}^k}{2i} \quad 43$$

svolgendo θ_k diventa:

$$\theta_k = \frac{z^{k+1} - \bar{z}^{k+1}}{(z - \bar{z})} = \frac{2i \cdot r^{k+1} \sin(\theta(k+1))}{2i \cdot r \cdot \sin \theta} = r^k \cdot \frac{\sin(\theta(k+1))}{\sin \theta}$$

Sostanzialmente, la funzione di risposta all'impulso è il risultato di due funzioni: $\frac{\sin(\theta(k+1))}{\sin \theta}$, funzione periodica in k con periodo $T = \frac{2\pi}{\theta}$ (quindi più θ è piccolo, più è lungo il ciclo); r^k , invece, restringe l'ampiezza delle oscillazioni, poiché si è supposto $|\rho| < 1$, ed ovviamente $\rho^k \rightarrow 0$ per $k \rightarrow \infty$.

⁴² Sono, cioè, delle serie geometriche convergenti del tipo:

$$\frac{1}{1 - zL} = \sum_{k=0}^{\infty} z^k L^k$$

$$\frac{1}{1 - \bar{z}L} = \sum_{k=0}^{\infty} \bar{z}^k L^k$$

⁴³ Usando $z = re^{i\theta}$.

C

Stime R

Paragrafo 5.4

VAR con serie non differenziate:

Endogenous variables: UNRATE, FEDFUNDS, CPIAUCSL
 Deterministic variables: const
 Sample size: 490
 Log Likelihood: -90.684
 Roots of the characteristic polynomial:
 1.002 0.9863 0.9524 0.4675 0.3314 0.08405

Estimation results for equation UNRATE:

	Estimate	Std. Error	t value	Pr(> t)
UNRATE.l1	0.97239	0.04593	21.171	< 2e-16
FEDFUNDS.l1	-0.02928	0.01363	-2.148	0.03222
CPIAUCSL.l1	0.01744	0.04114	0.424	0.67176
UNRATE.l2	0.01157	0.04557	0.254	0.79974
FEDFUNDS.l2	0.04348	0.01373	3.168	0.00163
CPIAUCSL.l2	-0.01777	0.04121	-0.431	0.66642
const	0.02390	0.03558	0.672	0.50212

Residual standard error: 0.177 on 483 degrees of freedom
 Multiple R-Squared: 0.9865, Adjusted R-squared: 0.9864
 F-statistic: 5891 on 6 and 483 DF, p-value: < 2.2e-16

Estimation results for equation FEDFUNDS:

	Estimate	Std. Error	t value	Pr(> t)
UNRATE.l1	-0.65364	0.14326	-4.563	6.42e-06
FEDFUNDS.l1	1.31042	0.04252	30.816	< 2e-16
CPIAUCSL.l1	0.20253	0.12832	1.578	0.1152
UNRATE.l2	0.64608	0.14213	4.546	6.93e-06
FEDFUNDS.l2	-0.33631	0.04281	-7.856	2.60e-14
CPIAUCSL.l2	-0.20311	0.12854	-1.580	0.1147
const	0.20704	0.11098	1.866	0.0627

Residual standard error: 0.5521 on 483 degrees of freedom
 Multiple R-Squared: 0.9709, Adjusted R-squared: 0.9706
 F-statistic: 2689 on 6 and 483 DF, p-value: < 2.2e-16

Estimation results for equation CPIAUCSL:

	Estimate	Std. Error	t value	Pr(> t)
UNRATE.l1	0.0003679	0.0473065	0.008	0.9938
FEDFUNDS.l1	0.0647816	0.0140419	4.613	5.08e-06
CPIAUCSL.l1	1.3723826	0.0423743	32.387	< 2e-16
UNRATE.l2	-0.0010824	0.0469318	-0.023	0.9816
FEDFUNDS.l2	-0.0395123	0.0141368	-2.795	0.0054
CPIAUCSL.l2	-0.3713534	0.0424442	-8.749	< 2e-16
const	-0.0654237	0.0366463	-1.785	0.0748

Residual standard error: 0.1823 on 483 degrees of freedom
 Multiple R-Squared: 1, Adjusted R-squared: 1
 F-statistic: 5.755e+06 on 6 and 483 DF, p-value: < 2.2e-16

Covariance matrix of residuals:

	UNRATE	FEDFUNDS	CPIAUCSL
UNRATE	0.031327	-0.018662	-0.001303
FEDFUNDS	-0.018662	0.304767	0.008291
CPIAUCSL	-0.001303	0.008291	0.033232

Correlation matrix of residuals:

	UNRATE	FEDFUNDS	CPIAUCSL
UNRATE	1.00000	-0.19099	-0.04038
FEDFUNDS	-0.19099	1.00000	0.08239
CPIAUCSL	-0.04038	0.08239	1.00000

VAR con serie differenziate:

Endogenous variables: UNRATE, FEDFUNDS, CPIAUCSL
 Deterministic variables: const
 Sample size: 488
 Log Likelihood: -69.438
 Roots of the characteristic polynomial:
 0.8369 0.7659 0.584 0.584 0.5755 0.5755 0.4907 0.4907 0.3088

Estimation results for equation UNRATE:

	Estimate	Std. Error	t value	Pr(> t)
UNRATE.l1	-0.007647	0.045642	-0.168	0.8670
FEDFUNDS.l1	-0.010946	0.014641	-0.748	0.4551
CPIAUCSL.l1	0.033734	0.040703	0.829	0.4076
UNRATE.l2	0.220669	0.044850	4.920	1.19e-06
FEDFUNDS.l2	0.016837	0.015397	1.094	0.2747
CPIAUCSL.l2	0.060812	0.044099	1.379	0.1685
UNRATE.l3	0.182936	0.045599	4.012	6.99e-05
FEDFUNDS.l3	-0.027506	0.014294	-1.924	0.0549
CPIAUCSL.l3	0.015690	0.040408	0.388	0.6980
const	-0.034330	0.013372	-2.567	0.0106

Residual standard error: 0.1714 on 478 degrees of freedom
 Multiple R-Squared: 0.1238, Adjusted R-squared: 0.1073
 F-statistic: 7.507 on 9 and 478 DF, p-value: 2.636e-10

Estimation results for equation FEDFUNDS:

	Estimate	Std. Error	t value	Pr(> t)
UNRATE.l1	-0.712680	0.145160	-4.910	1.25e-06
FEDFUNDS.l1	0.371252	0.046564	7.973	1.15e-14
CPIAUCSL.l1	0.160947	0.129450	1.243	0.214364
UNRATE.l2	-0.147333	0.142641	-1.033	0.302175
FEDFUNDS.l2	-0.179049	0.048968	-3.656	0.000284
CPIAUCSL.l2	-0.189085	0.140252	-1.348	0.178238
UNRATE.l3	-0.204277	0.145021	-1.409	0.159604
FEDFUNDS.l3	-0.069421	0.045461	-1.527	0.127407
CPIAUCSL.l3	0.046866	0.128511	0.365	0.715507
const	-0.004283	0.042528	-0.101	0.919823

Residual standard error: 0.5451 on 478 degrees of freedom
 Multiple R-Squared: 0.2256, Adjusted R-squared: 0.211
 F-statistic: 15.47 on 9 and 478 DF, p-value: < 2.2e-16

Estimation results for equation CPIAUCSL:

	Estimate	Std. Error	t value	Pr(> t)
UNRATE.11	0.007148	0.049135	0.145	0.88439
FEDFUNDS.11	0.046389	0.015762	2.943	0.00341
CPIAUCSL.11	0.415128	0.043818	9.474	< 2e-16
UNRATE.12	0.010148	0.048283	0.210	0.83361
FEDFUNDS.12	0.032147	0.016575	1.939	0.05303
CPIAUCSL.12	0.067344	0.047474	1.419	0.15668
UNRATE.13	-0.026752	0.049088	-0.545	0.58603
FEDFUNDS.13	0.005058	0.015388	0.329	0.74252
CPIAUCSL.13	0.291014	0.043500	6.690	6.26e-11
const	0.067658	0.014395	4.700	3.41e-06

Residual standard error: 0.1845 on 478 degrees of freedom
 Multiple R-Squared: 0.4855, Adjusted R-squared: 0.4758
 F-statistic: 50.11 on 9 and 478 DF, p-value: < 2.2e-16

Covariance matrix of residuals:

	UNRATE	FEDFUNDS	CPIAUCSL
UNRATE	0.0293761	-0.019046	-0.0005205
FEDFUNDS	-0.0190462	0.297133	0.0057060
CPIAUCSL	-0.0005205	0.005706	0.0340444

Correlation matrix of residuals:

	UNRATE	FEDFUNDS	CPIAUCSL
UNRATE	1.00000	-0.20386	-0.01646
FEDFUNDS	-0.20386	1.00000	0.05673
CPIAUCSL	-0.01646	0.05673	1.00000

G-causality:

- i: diff(FEDFUNDS)
- u: diff(UNRATE)
- p: diff(CPIAUCSL)

Granger causality H_0 : i do not Granger-cause u and p

F-Test = 5.8191, df1 = 4, df2 = 2085, p-value = 0.0001177

H_0 : No instantaneous causality between: i and u p

Chi-squared = 23.486, df = 2, p-value = 7.945e-06

▪

Granger causality H_0 : p do not Granger-cause u i

F-Test = 2.755, df1 = 4, df2 = 2085, p-value = 0.02661

H_0 : No instantaneous causality between: p and u i

Chi-squared = 1.9288, df = 2, p-value = 0.3812

▪

Granger causality H_0 : u do not Granger-cause p i

F-Test = 9.8104, df1 = 4, df2 = 2085, p-value = 7.322e-08

H_0 : No instantaneous causality between: u and p i

Chi-squared = 22.118, df = 2, p-value = 1.574e-05

■

Stime SVAR:

Type: A-model
 Sample size: 700
 Log Likelihood: -2979.771
 Method: direct
 Number of iterations: 104
 Convergence code: 0

LR overidentification test:

LR overidentification

$\chi^2 = 5700$, df = 3, p-value <2e-16

Estimated A matrix:

	u	p	i
u	1.00000	0.00000	0
p	0.01331	1.00000	0
i	0.48595	-0.1045	1

Estimated standard errors for A matrix:

	u	p	i
u	0.0000	0.0000	0
p	0.2317	0.0000	0
i	0.2317	0.1611	0

Estimated B matrix:

	u	p	i
u	1	0	0
p	0	1	0
i	0	0	1

Covariance matrix of reduced form residuals (*100):

	u	p	i
u	100.000	-1.331	-48.73
p	-1.331	100.018	11.10
i	-48.734	11.101	124.84

Coefficienti funzione di risposta all'impulso, SVAR:

• U

	u	p	i
[1,]	1.0000000000	0.021668507	-0.490104182
[2,]	0.083209284	0.061157636	-0.772203178
[3,]	0.259152963	0.019737096	-0.383131844
[4,]	0.056072336	-0.011082010	-0.171576333
[5,]	0.073231166	-0.012293922	-0.064977148

```
[6,] 0.021285677 -0.010153050 -0.043332208
[7,] 0.019707766 -0.007088389 -0.026563944
[8,] 0.006904569 -0.005848617 -0.016323083
[9,] 0.005360183 -0.004594834 -0.007988910
[10,] 0.002051155 -0.003567887 -0.004054813
[11,] 0.001346028 -0.002629364 -0.001950655
```

Lower Band, CI= 0.95

```
      u      p      i
[1,] 1.0000000000 -0.12689359 -0.779228265
[2,] 0.0251992745 -0.08260518 -0.966016616
[3,] 0.1837668881 -0.11121622 -0.587508020
[4,] 0.0195646024 -0.11604327 -0.264560067
[5,] 0.0402544671 -0.10157207 -0.134540436
[6,] 0.0040124400 -0.08056714 -0.085877547
[7,] 0.0063866159 -0.06279909 -0.054985667
[8,] -0.0016241296 -0.04885442 -0.036419296
[9,] 0.0002995426 -0.03709253 -0.021272550
[10,] -0.0032569093 -0.02854892 -0.013922326
[11,] -0.0021031639 -0.02120500 -0.008801315
```

Upper Band, CI= 0.95

```
      u      p      i
[1,] 1.0000000000 0.12883366 -0.250665418
[2,] 0.154946460 0.18586173 -0.565627005
[3,] 0.314776753 0.18251035 -0.182406584
[4,] 0.099164421 0.10240301 -0.039228707
[5,] 0.112494676 0.10749596 0.008865000
[6,] 0.043908562 0.08220681 -0.008065914
[7,] 0.038529365 0.06816732 -0.008878152
[8,] 0.018562747 0.04825683 -0.003252080
[9,] 0.014805985 0.03500381 0.002223766
[10,] 0.008631998 0.02617573 0.003391914
[11,] 0.006625901 0.02016041 0.004798860
```

• I

```
      u      p      i
[1,] 0.0000000000 0.000000000 1.000000000
[2,] -0.008195513 0.063965929 0.401686484
[3,] -0.016044532 0.097183574 -0.014525669
[4,] -0.004543129 0.072166884 -0.064801535
[5,] 0.001348546 0.044831555 -0.019410095
[6,] 0.003494948 0.029205522 0.002380545
[7,] 0.003005030 0.020993088 0.001456661
[8,] 0.002468441 0.015457894 -0.002392775
[9,] 0.001951498 0.011120400 -0.003235158
[10,] 0.001562109 0.007846059 -0.002463893
[11,] 0.001197014 0.005517910 -0.001676535
```

Lower Band, CI= 0.95

```
      u      p      i
[1,] 0.000000e+00 0.000000000 1.000000000
[2,] -3.481126e-02 0.020346218 0.338541615
[3,] -3.637846e-02 0.063182660 -0.085173373
[4,] -1.572377e-02 0.045161045 -0.121105904
[5,] -5.616306e-03 0.025794785 -0.042170927
[6,] -6.418696e-04 0.015290611 -0.017838684
[7,] 1.797742e-05 0.010091340 -0.012127935
[8,] 3.713990e-04 0.007166462 -0.010482995
[9,] 3.826789e-04 0.004802968 -0.008849989
[10,] 3.362767e-04 0.002958005 -0.006333429
```

[11,] 2.787768e-04 0.001765385 -0.004480757

Upper Band, CI= 0.95

	u	p	i
[1,]	0.000000000	0.000000000	1.000000000
[2,]	0.015169866	0.10223670	0.4618049371
[3,]	0.007414558	0.13045142	0.0447408095
[4,]	0.006607467	0.09887589	-0.0221722430
[5,]	0.009078313	0.06117079	0.0108304329
[6,]	0.008524827	0.04056963	0.0208725806
[7,]	0.006931736	0.03033114	0.0127934485
[8,]	0.005457856	0.02319648	0.0037319918
[9,]	0.004461260	0.01809122	0.0013414258
[10,]	0.003525252	0.01411998	0.0009353561
[11,]	0.002731088	0.01099786	0.0007035983

• P

	u	p	i
[1,]	0.000000000	1.000000000	0.096761583
[2,]	0.009686413	0.52532704	0.197001095
[3,]	0.045197234	0.42567260	0.012593821
[4,]	0.029666562	0.30492532	-0.058528827
[5,]	0.034417685	0.21737494	-0.053965559
[6,]	0.025998823	0.15270850	-0.038056889
[7,]	0.021942890	0.10798194	-0.027493497
[8,]	0.016155011	0.07655311	-0.021837028
[9,]	0.012351595	0.05427667	-0.017076624
[10,]	0.009013404	0.03838922	-0.012894095
[11,]	0.006650642	0.02712015	-0.009429288

Lower Band, CI= 0.95

	u	p	i
[1,]	0.000000000	1.000000000	-0.07292671
[2,]	-0.0418366878	0.456618228	0.07099816
[3,]	-0.0048352069	0.359849423	-0.10102168
[4,]	-0.0015413531	0.231784358	-0.17827796
[5,]	-0.0018999334	0.143145677	-0.14812841
[6,]	0.0016347314	0.088392128	-0.10051368
[7,]	0.0006725920	0.053728153	-0.06909704
[8,]	0.0010710261	0.033632576	-0.05169834
[9,]	0.0005950135	0.020979400	-0.03768213
[10,]	0.0006335438	0.013000201	-0.02798737
[11,]	0.0003702770	0.007624173	-0.02074909

Upper Band, CI= 0.95

	u	p	i
[1,]	0.000000000	1.000000000	0.233669352
[2,]	0.06736077	0.59316825	0.340182044
[3,]	0.08199750	0.48038690	0.112632087
[4,]	0.06879937	0.35981115	0.027883700
[5,]	0.06285974	0.27567661	0.009847740
[6,]	0.04966948	0.20570792	0.008207871
[7,]	0.04351374	0.15590249	0.009118752
[8,]	0.03340326	0.11825675	0.005260751
[9,]	0.02500184	0.09016262	0.002709396
[10,]	0.01823179	0.06876242	0.001515143
[11,]	0.01369839	0.05209639	0.001009614

Bibliografia

Adhikari, R., Agrawal, R.K., 2013. *An Introductory Study on Time Series Modeling and Forecasting*, LAP Lambert Academic Publishing, Nuova Delhi, pp. 12-16, pp. 18-23.

Banerjee, A., et al. (1993). *Co-integration, Error Correction, and the Econometric Analysis of Non-Stationary Data*, OUP Catalogue, Oxford University Press.

Bertsch M., Dal Passo R., Giacomelli L., 2011, *Analisi matematica*, 2^a edizione. McGraw-Hill, pp. 482-519.

Brooks, C., 2008. *Introductory Econometrics for Finance*, 2^a edizione. Cambridge University Press, Cambridge, pp. 318-365.

Davidson, J.E.H., et al., 1978. *Econometric Modelling of the Aggregate Time-Series Relationship Between Consumers' Expenditure and Income in the United Kingdom*, The Economic Journal, Royal Economic Society, vol. 88(352), pp. 661-692.

Fernández-Villaverde, J., Rubio-Ramírez, J.F., 2010. *Structural Vector Autoregressions*. In: Durlauf S.N., Blume L.E. (eds) *Macroeconometrics and Time Series Analysis*, The New Palgrave Economics Collection, Palgrave Macmillan, Londra.

Granger, C.W.J., 1988. *Some recent development in a concept of causality*, Journal of Econometrics, Elsevier, vol. 39 (1-2), pp. 199-211.

Granger, C.W.J., Newbold, P., 1974. *Spurious regressions in econometrics*, Journal of Econometrics, Elsevier, vol. 2(2), pp. 111-120.

Grazzini, J., 2012. *Analysis of the Emergent Properties: Stationarity and Ergodicity*, Journal of Artificial Societies and Social Simulation, vol. 15(2), pp. 1-7.

Hipel, K.W., McLeod, A.I., 1994. "Time Series Modelling of Water Resources and Environmental Systems", Elsevier, Amsterdam, pp. 16-19, pp. 63-69.

Johansen, S., 2014. *Times Series: Cointegration*, Center for Research in Econometric Analysis of Time Series, Danish National Research Foundation, Research Paper 2014-38.

Lucas, R., 1976. *Econometric policy evaluation: A critique*, Carnegie-Rochester Conference Series on Public Policy, Elsevier, vol. 1(1), pp. 19-46.

Lütkepohl, H., 2005. *New introduction to Multiple Time Series analysis*. Springer Verlag, Heidelberg, pp. 1-66, pp. 357-412, pp. 645-676.

- Mikusheva, A., 2007. *Course materials for 14.384 Time Series Analysis*, MIT OpenCourseWare (<http://ocw.mit.edu>), Massachusetts Institute of Technology.
- Murray, M., 1994. *A drunk and her dog: an illustration of cointegration and error correction*. The American Statistician, American Statistical Association, vol. 48 (1), pp. 37-39.
- Peracchi, F., 2012. *Cointegrazione*, Dizionario di Economia e Finanza, Enciclopedia Treccani.
- Pfaff, B., 2008. *VAR, SVAR and SVEC Models: Implementation Within R Package vars*. Journal of Statistical Software, 27, pp. 1-32.
- Piccolo, D., 2000. *Statistica*, 2^a edizione. Il Mulino, Bologna, p. 39, pp. 310-320, pp. 833-871.
- Rubio-Ramírez, J.F., Waggoner, D.F., Zha, T., 2010. *Structural Vector Autoregressions: Theory of Identification and Algorithms for Inference*, Review of Economic Studies, Oxford University Press, vol. 77(2), pp. 665-696.
- Shumway, R.H., Stoffer, D.S., 2018. *Time Series Analysis Using the R Statistical Package*, Free Dog Publishing, pp. 33-51, pp. 58-67, pp. 85-88.
- Sims, C.A., 1980. *Macroeconomics and reality*, Econometrica, The Econometric Society, vol. 48(1), pp. 1-48.
- Sasso, L., 2012. *Matematica a colori 5. Edizione Blu*. Petrini, Novara, pp. 655-668.
- Smith, J.O., 2002. *Mathematics of the Discrete Fourier Transform*, CCRMA, Stanford University, p. 93.
- Stigler, M., 2010. *Threshold cointegration: overview and implementation in R*, R package version 0.7-2.
- Stock, J.H., Watson, M.W., 2001. *Vector Autoregressions*, Journal of Economic Perspectives, American Economic Association, vol. 15(4), pp. 101-115.
- Tsay, R.S., 2002. *Analysis of Financial Time Series*, Wiley Series in Probability and Statistics, John Wiley and Sons Inc., U.S.A., pp. 22-79.
- Uhlig, H., 2005. *What are the effects of monetary policy on output? Results from an agnostic identification procedure*. Journal of Monetary Economics, Elsevier, vol. 52, pp. 381-419.
- Walter, S., 2008. *Conceptualizing Overfit or Over-parametization*, Rasch Measurement Transactions, vol. 22(2), p. 1135.