

UNIVERSITA' DEGLI STUDI DI PADOVA

Facoltà di Scienze Statistiche

Corso di Laurea Specialistica in Scienze Statistiche,
Economiche, Finanziarie ed Aziendali

Tesi di Laurea

**Memoria lunga e volatilità: modelli, confronti e
applicazioni.**

Relatore : Ch.mo Prof. SILVANO BORDIGNON

Laureando: MARCO PAVAN

ANNO ACCADEMICO 2004-2005

*Ai miei genitori,
a mio fratello,
ai miei nonni
e a Federica*

Sommario

1. INTRODUZIONE	7
1.1 MODELLI PER LA VOLATILITÀ E MEMORIA LUNGA	8
1.2 COME SI ARTICOLA LA TESI.....	10
2. LA VOLATILITÀ: CARATTERISTICHE, MISURE E MODELLI.....	11
2.1 CARATTERISTICHE DELLA VOLATILITÀ	11
2.2 COSA GENERA VOLATILITÀ	13
2.3 LE MISURE DI VOLATILITÀ	13
2.3.1 <i>Varianza mobile</i>	14
2.3.2 <i>Varianza mobile e aggregazione</i>	15
2.3.3 <i>Varianza RiskmetricsTM</i>	15
2.4 MODELLI PER LA VOLATILITÀ	16
3. CENNI DI ANALISI SPETTRALE E MODELLI ARFIMA	19
3.1 ALCUNI STUDI SULLA DENSITÀ SPETTRALE	19
3.1.1 <i>La funzione generatrice delle autocovarianze</i>	21
3.1.2 <i>Lo spettro del processo</i>	21
3.1.3 <i>Il significato dello spettro</i>	23
3.1.4 <i>Il periodogramma</i>	25
3.2 IL MODELLO ARFIMA: SPECIFICAZIONE E PROPRIETÀ	26
4. FRACTIONALLY INTEGRATED GENERALIZED AUTOREGRESSIVE CONDITIONAL HETEROSKEDASTICITY (FIGARCH).....	29
4.1 DEFINIZIONE E PROPRIETÀ DEL MODELLO FIGARCH.....	30
4.2 STIMA E INFERENZA.....	34
4.2.1 <i>Stima della Massima Verosimiglianza</i>	35
4.2.2 <i>Stima di Quasi Massima Verosimiglianza</i>	37

5. MODELLI A VOLATILITÀ STOCASTICA CON MEMORIA LUNGA (LMSV)	41
5.1 LA PERSISTENZA NELLA VOLATILITÀ	41
5.2 UN MODELLO PER LA VOLATILITÀ STOCASTICA CON MEMORIA LUNGA	42
5.3 STIMA DEL PARAMETRO DI MEMORIA LUNGA	45
5.3.1 <i>Il metodo R/S o del rango riscaldato</i>	46
5.3.2 <i>Il metodo del log-periodogramma</i>	48
5.4 STIMA DI UN MODELLO LMSV	50
5.4.1 <i>Stima di FDQML</i>	52
5.4.2 <i>FDQML “Accresciuto”</i>	54
6. ANALISI EMPIRICHE.....	59
6.1 ANALISI DEI DATI AUTOSTRADE	59
6.1.1 <i>Modelli a volatilità stocastica per Autostrade S.p.A</i>	68
6.2 ANALISI DEI DATI MIBTEL	70
6.2.1 <i>Modelli a volatilità stocastica per l’indice Mibtel</i>	79
6.3 CONCLUSIONI	81
7. APPENDICE.....	83
7.1 ALCUNE FUNZIONI UTILI	83
7.2 STUDIO DI SIMULAZIONI.....	90
7.3 LO SPLITTING METHOD	92
BIBLIOGRAFIA.....	94

Capitolo 1

1. INTRODUZIONE

L'interesse per i processi a integrazione frazionaria è nato negli anni '50 nel campo dell'ingegneria, in particolare nell'idrologia: il primo studio fu proposto da Hurst (1951) a proposito delle proprietà di riempimento dei bacini idrici. Anche se il primo studio applicato a dati finanziari è dovuto a Mandelbrot (1971), questo tipo di modellistica ha iniziato a destare vero interesse negli econometrici solo all'inizio degli anni '80. Fino al 1980 lo strumento maggiormente utilizzato per gli studi di serie storiche economiche era la modellistica ARMA; tale classe di modelli presenta, insieme all'indubbio pregio della semplicità, alcune grosse rigidità. Prima di continuare, è meglio specificare che cosa si intende per "integrazione", ovvero si ricorre al concetto di integrazione di un certo ordine per un processo per indicare il numero di volte che si deve applicare l'operatore differenza per ottenere un processo stazionario. Formalmente, un processo si dice integrato di ordine n , e si indica con la notazione $I(n)$, se, differenziato n volte, genera un processo stazionario. Per estensione si attribuisce l'aggettivo di "integrato" o "stazionario" anche alla serie osservata di realizzazioni del processo; ad esempio, il processo dato da $y_t = y_{t-1} + \varepsilon_t$, ove $\varepsilon_t \sim IID(0, \sigma_\varepsilon^2)$, è del tipo $I(1)$. Qualora gli ε_t abbiano struttura più complessa (ad esempio, siano un processo ARMA), la presenza di integrazione rende il processo più generale definito come ARIMA (*AutoRegressive – Integrated – Moving Average*), che rappresentano una classe di modelli introdotta pressoché simultaneamente agli ARMA ordinari.

Tuttavia, non tutti i fenomeni possono essere descritti da modelli semplici, e, in effetti, alcune serie presentano difficoltà di adattamento all'ipotesi che siano generate da un processo ARMA. In particolare, in alcuni casi a fronte di autocorrelazioni teoriche, che in un modello ARMA decadono esponenzialmente, quelle empiriche calcolate a partire dai valori di una serie seguono un sentiero di discesa molto più "morbido", suggerendo l'inadeguatezza della semplicità del modello adottato.

In sostanza, la conseguenza dell'adattare un modello ARMA a dati che presentano tale andamento nell'autocorrelazione empirica è che il processo "dimentica il proprio passato" in modo più lento rispetto a quello che ci si sarebbe aspettato.

Le prime applicazioni dei modelli a integrazione frazionaria a dati macroeconomici riguardano l'evoluzione nel tempo del PNL reale, in particolare il dibattito se tale variabile sia $I(1)$ oppure $I(0)$ con trend deterministico. Studi in questa direzione sono stati portati avanti da Diebold e Rudebusch (1989) e Sowell (1992b). Ancora più interessanti sono i risultati ottenuti relativamente agli indici dei prezzi al consumo da Baillie, Chung e Tieslau (1996). Si è infatti osservato che i dati di tutti i paesi osservati, meno quelli del Giappone, presentavano evidenza di comportamento $I(d)$.

Molto più interessanti per questo lavoro sono però i risultati che si sono ottenuti nella modellazione a integrazione frazionaria delle serie finanziarie ad alta frequenza. Studi sono stati condotti nel campo dei tassi di interesse (Crato e Rothman, 1994, Tsay, 2000), ma soprattutto in quello dei rendimenti giornalieri assoluti o quadrati dei mercati azionari e dei cambi. Le numerose ricerche in questo campo (si vedano ad esempio Crato e de Lima, 1994, Baillie e Bollerslev, 1994, Granger e Ding, 1996a e 1996b, Baillie, Bollerslev e Mikkelsen, 1996, Bollerslev e Mikkelsen, 1996, Cheung, 1996, Andersen e Bollerslev, 1997a) hanno messo in evidenza la presenza di integrazione frazionaria nei rendimenti assoluti e nella varianza condizionata.

1.1 Modelli per la volatilità e memoria lunga

L'incertezza, oggigiorno, è cruciale nella modellazione delle serie storiche finanziarie. In molte delle teorie riguardanti i titoli azionari, l'incertezza associata al loro prezzo è un fattore importante nella determinazione del premio per il

rischio. Questa incertezza è usualmente messa in relazione con la varianza condizionale, che è variabile nel tempo, nel tentativo di assecondare al meglio il comportamento di diverse serie storiche finanziarie. In particolare molte serie finanziarie, come i rendimenti di un titolo azionario, non hanno una struttura complessa nelle autocorrelazioni ma presentano persistenza nelle autocorrelazioni di alcune trasformazioni come i quadrati o i moduli dei rendimenti.

I lavori di Engle (1982) e Taylor (1986) hanno posto le basi per i due principali filoni di letteratura utilizzati nel modellare l'eteroschedasticità condizionale. Il modello ARCH di Engle (1982) e le successive estensioni considera la varianza condizionale come funzione dei quadrati delle osservazioni passate. La seconda tendenza si riferisce ai modelli a volatilità stocastica (SV), nei quali si pone che la volatilità sia generata da un processo stocastico cosicché sia l'equazione per la media che per la varianza hanno un termine d'errore.

Una grande gamma di ricerche ha suggerito che la volatilità di molte serie storiche finanziarie evidenzia una forte persistenza che non può essere modellata semplicemente utilizzando i modelli ARCH o SV. Per modellare questo comportamento Engle e Bollerslev (1986) hanno proposto il modello Integrated Generalized ARCH (IGARCH). Sebbene i modelli IGARCH abbiano molte somiglianze con i modelli ARIMA per la media condizionale, vi sono diversi punti molto diversi. Per inciso, sebbene non sia debolmente stazionario, l'IGARCH è strettamente stazionario (Nelson, 1990). Inoltre la funzione di autocorrelazione dei quadrati non è costante e decresce esponenzialmente.

L'esistenza di memoria lunga nelle potenze dei valori assoluti dei rendimenti di numerosi titoli finanziari ha aperto una nuova branca di ricerca che include la persistenza nella modellazione della varianza condizionale. In questo ambito, Baillie, Bollerslev e Mikkelsen (1996) hanno proposto il Fractionally IGARCH (FIGARCH). Più recentemente Harvey (1998) e Breidt, Crato e de Lima (1998) hanno proposto il modello Long Memory Stochastic Volatility (LMSV) che modella il termine della log-volatilità come un modello Fractional ARIMA (ARFIMA). Per entrambi questi modelli ci sono vantaggi e svantaggi che verranno presi in esami nei capitoli successivi.

1.2 Come si articola la tesi

Dopo una breve introduzione al problema, nel Capitolo 2 si illustreranno le principali caratteristiche della volatilità, assieme alle sue misure più frequentemente utilizzate; inoltre verranno introdotti formalmente, ma in maniera sintetica, i modelli ARCH, GARCH, EGARCH e SV, che serviranno da base per formalizzare i due modelli che sono d'interesse principale per questa tesi, ovvero il FIGARCH e LMSV.

Anche il Capitolo 3 è stato scritto per introdurre i modelli ARFIMA, che verranno diffusamente richiamati in seguito, e alcuni strumenti dell'analisi spettrale (come il periodogramma e lo spettro) che verranno utilizzati per ricavare la metodologia di stima del modello LMSV.

Nel Capitolo 4 verrà trattato il modello FIGARCH: in particolare verrà ricavato, tramite alcuni accorgimenti, dal modello IGARCH e verranno esposte anche le sue principali proprietà, come il lento decadimento iperbolico della funzione di autocorrelazione. Inoltre verranno presentati i metodi di stima più frequenti, che sono la massima verosimiglianza e la quasi massima verosimiglianza.

Nel Capitolo 5 verrà analizzato l'altro modello di principale interesse di questa tesi, ovvero l'LMSV. Dopo aver mostrato le caratteristiche principali di questa classe di modelli, verranno presentate due procedure di stima del parametro di integrazione frazionaria d . Infine verrà presentata nel dettaglio una delle varie tipologie di stima, che fa capo all'approssimazione di Whittle, e una sua variante introdotta di recente da Deo, Hurvich e Lu (2005).

Nel Capitolo 6 verranno messe in pratica le nozioni introdotte analizzando due serie reali, il titolo Autostrade S.p.A. e l'indice Mibtel. Come si vedrà in seguito, l'introduzione della memoria lunga porterà a dei modelli che assecondano meglio i dati rispetto a quelli con memoria corta.

Infine nell'appendice verranno riportate alcune funzioni scritte in linguaggio R che sono state utilizzate per compiere le analisi empiriche.

Capitolo 2

2. La volatilità: caratteristiche, misure e modelli

In questo capitolo verranno introdotti alcuni argomenti inerenti alla volatilità: in primo luogo verranno descritte le sue principali caratteristiche e alcune possibili cause della sua manifestazione. In seguito verranno descritte alcune misure differenti della volatilità, come la varianza mobile o la varianza RiskmetricsTM, e introdotti i principali modelli, come ARCH, GARCH ed altri.

2.1 Caratteristiche della volatilità

La maggior parte delle ricerche macroeconomiche sono focalizzate sull'equazione della media della variabile che si analizza. Se ad esempio si studia l'andamento del tasso di disoccupazione, è d'interesse vedere da quali altre variabili esso è influenzato, e quale livello di disoccupazione si può ragionevolmente attendere in futuro. Poca attenzione è in genere dedicata ai momenti successivi della distribuzione della variabile, a cominciare dalla varianza.

Le ricerche che hanno tentato di isolare delle regolarità nei rendimenti dei mercati finanziari hanno però dato scarsi risultati: tali rendimenti appaiono difatti (specialmente quelli dei mercati dei cambi) decisamente poco autocorrelati. D'altra parte, per i gestori di portafoglio e gli analisti finanziari, non è rilevante solo riuscire a prevedere i rendimenti futuri delle attività finanziarie, ma anche la loro volatilità, in modo da poter misurare e coprire adeguatamente il rischio.

Definire il concetto di volatilità è complesso e non può portare a risultati univoci; in prima approssimazione si può dire che la volatilità misura la dispersione dei rendimenti che una certa attività finanziaria è suscettibile di generare. Volendo essere più formali, e attenendosi a un contesto tipicamente statistico, si osserva che, se si tratta il rendimento come una variabile aleatoria con una certa distribuzione nota, si può identificare, dato un livello di confidenza, un intervallo all'interno del quale si troverà, al dato livello di confidenza, il rendimento futuro.

Una caratteristica della volatilità dei titoli azionari è che non è direttamente osservabile; ad esempio la volatilità giornaliera non è direttamente osservabile dato che esiste una sola osservazione giornaliera. Se sono disponibili dei dati infragiornalieri, come i rendimenti ogni 5 minuti, allora si può stimare la volatilità giornaliera; inoltre la non osservabilità della volatilità rende difficile la valutazione della performance previsiva dei modelli con eteroschedasticità condizionale.

Sebbene la volatilità dei titoli azionari sia una variabile non direttamente osservabile, possiede delle caratteristiche che sono comunemente osservate nei rendimenti della attività. Per prima cosa esistono degli effetti di clustering (ovvero la volatilità può essere alta in certi periodi e più bassa in altri). Secondo, la volatilità evolve nel tempo in maniera continua, cioè sono rari i salti di volatilità. Terzo, può succedere che la volatilità presenti correlazione, magari di lungo periodo o di tipo stagionale (quest'ultima nei dati intragiornalieri). Quarto, l'informazione che si accumula nelle giornate di chiusura dei mercati ha l'effetto di aumentare la volatilità, per cui quest'ultima tende ad essere più elevata nei giorni di riapertura (il che è noto anche come effetto *week-end*). Quinto, la volatilità sembra reagire in maniera differente a grossi incrementi o decrementi dei prezzi. Alcuni modelli sono stati appositamente proposti per catturare questa caratteristica: per esempio, il modello EGARCH è stato sviluppato per catturare l'asimmetria nella volatilità generata da grandi incrementi o grandi decrementi dei rendimenti delle attività finanziarie.

2.2 Cosa genera volatilità

La risposta ad una tale domanda può essere solo parziale, dato che non vi è consenso sulle cause delle variazioni della volatilità. Tuttavia, una possibile spiegazione dell'insorgere e persistere della volatilità si riferisce al fatto che i prezzi si muovono sulla base delle reazioni degli agenti economici ai flussi d'informazione (le innovazioni) e che le informazioni di dominio pubblico a livello macroeconomico (dati su offerta di moneta, produzione industriale, inflazione disoccupazione), arrivano esse stesse a grappoli e quindi c'è un raggruppamento di innovazioni più importanti di altre, che hanno cioè un maggiore impatto (quindi caratterizzate da una varianza più elevata). Al tempo stesso, si può pensare che le caratteristiche dell'attività di negoziazione che si svolge sui vari mercati del mondo, grazie alla presenza di fuso orario, possa generare un accumulo di notizie che arrivano da altri mercati aperti e che si trasmette ad altri mercati al momento dell'apertura.

Fra le spiegazioni alternative più elaborate, una può essere quella di un effetto psicologico a catena dovuto alla sorpresa di certe notizie che segnalano incertezza circa l'andamento dell'economia (dati sulla disoccupazione o dati preliminari sull'inflazione) che hanno bisogno di ulteriore conferma. Un segnale di possibile recessione che, se confermato, porterebbe ad un ribasso dei prezzi (innovazione negativa) può semplicemente generare dubbio e nervosismo con possibili serie di conferme e smentite successive che vengono eventualmente amplificate. Tali segnali rendono i mercati maggiormente attenti a quello che viene annunciato successivamente.

2.3 Le misure di volatilità

Un indice di variabilità tipicamente utilizzato come misura della volatilità dei rendimenti è la varianza calcolata sul periodo campionario (varianza storica o realized volatility al quadrato):

$$\sigma^2 = \frac{1}{T-1} \sum_{t=1}^T (r_t - \bar{r})^2$$

La media aritmetica dei rendimenti \bar{r} è il parametro intorno al quale si calcolano gli scostamenti al quadrato nell'espressione della deviazione standard. Essa può essere sostituita con altre misure di posizione alternative come la media mobile.

Come già detto in precedenza, il fatto che la volatilità sia variabili nel tempo è una delle regolarità che si riscontrano empiricamente: dall'analisi del comportamento dei rendimenti, è agevole evidenziare che ci sono dei periodi in cui la varianza tende a rimanere bassa e altri in cui tende a rimanere elevata. Se si misura la volatilità su tutto il periodo di analisi non si riesce a cogliere questo aspetto, perché non si riesce a vedere questo alternarsi di periodi di alta e bassa volatilità. Il problema è dunque quello di misurare la volatilità in modo da evidenziarne l'evoluzione temporale. Di seguito vengono illustrati alcuni modi per conseguire questo risultato.

2.3.1 Varianza mobile

Si scinda il periodo di osservazione T in tanti intervalli di ampiezza τ e ogni volta, aggiungendo un'osservazione più recente ed eliminandone una più vecchia, si ricalcoli la varianza mantenendo costante il numero di osservazioni su cui essa è calcolata:

$$\sigma_t^2 = \frac{1}{\tau} \sum_{i=t-\tau+1}^t (r_i - \bar{r}_t)^2$$

Anche il rendimento medio $\bar{r}_t = \frac{1}{\tau} \sum_i r_i$ viene ricalcolato di volta in volta togliendo l'osservazione più vecchia ed aggiungendo la più recente.

Questa misura, molto utilizzata, presenta un andamento molto smussato tipico di ogni media mobile; raramente si trovano dei salti in questa misura e ciò accade solo quando o il primo rendimento del periodo precedente o quello più recente hanno valori abnormi.

2.3.2 Varianza mobile e aggregazione

Una misura alternativa della volatilità può essere ottenuta come una sorta di varianza mobile calcolando la misura all'interno del mese e attribuendo lo stesso valore a tutte le osservazioni del mese. Per esempio, la varianza di gennaio sarà data da:

$$\sigma_{gen}^2 = \frac{1}{N_{gen} - 1} \sum_{i=1}^{N_{gen}} (r_{i,gen} - \bar{r}_{gen})^2,$$

dove N_{gen} è il numero di giorni di apertura di gennaio; analogamente la varianza di febbraio sarà:

$$\sigma_{feb}^2 = \frac{1}{N_{feb} - 1} \sum_{i=1}^{N_{feb}} (r_{i,feb} - \bar{r}_{feb})^2$$

Anche in questo caso si ottiene una misura di volatilità che varia nel tempo.

2.3.3 Varianza RiskmetricsTM

Si tratta di una diversa misura di volatilità costruita dalla J.P. Morgan¹, anch'essa variabile nel tempo. La volatilità al tempo t è calcolata come combinazione convessa della stima della volatilità effettuata al tempo precedente $t-1$, e del quadrato dello scarto del rendimento r_t al tempo t e una media, possibilmente variabile nel tempo, dei rendimenti \bar{r}_t :

$$\hat{\sigma}_t^2 = \lambda \hat{\sigma}_{t-1}^2 (1 - \lambda) (r_t - \bar{r}_t)^2$$

dove $0 \leq \lambda \leq 1$. Se λ è posto uguale ad 1 si torna al caso di variabilità costante: in pratica non si fa mai influenzare dalle innovazioni connesse con l'osservazione più recente. Se λ è uguale a zero, si ottiene come stima della varianza lo scarto più recente della media al quadrato. Per quanto riguarda la scelta di λ , J.P. Morgan suggerisce, basandosi sull'esperienza, di utilizzare un valore di $\lambda \approx 0.9$.

¹ Si veda www.riskmetrics.com.

2.4 Modelli per la volatilità

Una delle regolarità empiriche presentate dalle serie dei rendimenti è la persistenza di autocorrelazione in trasformazioni dei rendimenti, come il quadrato e i moduli. Engle(1982), sfruttando questa caratteristica, suggerisce che l'andamento della varianza di alcune serie storiche sia di tipo condizionatamente autoregressiva: sulla base dell'informazione disponibile al tempo $t-1$, la varianza al tempo t è esprimibile come funzione dei valori più recenti delle innovazioni al quadrato. Tale modello viene chiamato ARCH, ovvero *AutoRegressive Conditional Heteroskedasticity*.

Secondo Engle(1982), un modello ARCH(p) può essere espresso come:

$$y_t = \varepsilon_t \sigma_t$$

$$\varepsilon_t \sim IID(0,1) \text{ e di conseguenza } y_t | I_{t-1} \sim (0, \sigma_t^2)$$

$$\sigma_t^2 = \omega + \alpha_1 y_{t-1}^2 + \dots + \alpha_p y_{t-p}^2 = \omega + \alpha(B) y_t^2$$

in cui B è l'operatore ritardo e $\alpha(B) = \alpha_1 B + \dots + \alpha_p B^p$.

Bollerslev (1986), generalizzò il processo ARCH nel processo GARCH (G sta per *Generalized*) e, in particolare, un processo GARCH(p,q) può essere definito dalle seguenti condizioni:

$$y_t = \varepsilon_t \sigma_t$$

$$\varepsilon_t \sim IID(0,1) \text{ e di conseguenza } y_t | I_{t-1} \sim (0, \sigma_t^2)$$

$$\sigma_t^2 = \omega + \alpha_1 y_{t-1}^2 + \dots + \alpha_p y_{t-p}^2 + \beta_1 \sigma_{t-1}^2 + \dots + \beta_q \sigma_{t-q}^2$$

$$= \omega + \alpha(B) y_t^2 + \beta(B) \sigma_t^2$$

in cui B è l'operatore ritardo e $\alpha(B) = \alpha_1 B + \dots + \alpha_p B^p$ e $\beta(B) = \beta_1 B + \dots + \beta_q B^q$; per la non negatività della varianza condizionata bisogna imporre le condizioni $\omega > 0, \alpha_i \geq 0, \beta_j \geq 0$.

Una specificazione alternativa del modello GARCH fu presentata da Nelson (1991) che ha proposto l'EGARCH (Exponential GARCH) che, nel caso di un EGARCH(1,1), assume la forma

$$\log \sigma_t^2 = \alpha_0 + \beta_1 \log \sigma_{t-1}^2 + \alpha_1 \left(\left| \frac{y_{t-1}}{\sigma_{t-1}} \right| - \sqrt{\frac{2}{\pi}} \right) + \gamma \frac{y_{t-1}}{\sigma_{t-1}}$$

Qualsiasi valore sia attribuito ai parametri, la trasformazione esponenziale assicura la non negatività della varianza; infatti l'espressione può essere riscritta come:

$$\sigma_t^2 = \alpha_0^* (\sigma_{t-1}^2)^{\beta_1} \exp\left(\frac{\gamma \cdot y_{t-1} + \alpha_1 |y_{t-1}|}{\sigma_{t-1}}\right)$$

dove $\alpha_0^* = \exp\left(\alpha_0 - \alpha_1 \sqrt{\frac{2}{\pi}}\right)$

Nel modello EGARCH il termine $\beta_1 \log \sigma_{t-1}^2$ cattura l'effetto di persistenza nella volatilità; dato che l'espressione ha un termine autoregressivo, la stazionarietà è assicurata dalla condizione $0 < \beta < 1$. La sua dimensione determinerà quanto rapido sia l'assorbimento degli shock passati.

Il termine $\left(\frac{y_{t-1}}{\sigma_{t-1}} \middle| - \sqrt{\frac{2}{\pi}}\right)$ è una variabile casuale a media zero nel caso in cui le

innovazioni standardizzate siano distribuite normalmente² che consente di tenere conto della possibilità di una reazione asimmetrica proporzionale alle innovazioni.

L'effetto misurato dal termine espresso in valore assoluto $\left|\frac{y_{t-1}}{\sigma_{t-1}}\right|$ è evidentemente

indipendente dal fatto che gli shock siano positivi o negativi, mentre l'effetto asimmetrico viene evidenziato dall'ultimo elemento dell'espressione, $\gamma \frac{y_{t-1}}{\sigma_{t-1}}$,

dove y_{t-1} può assumere qualunque segno; il segno atteso per il parametro γ sarà negativo (in questo caso si parla di effetto *leverage*), in quanto ci si aspetta un effetto amplificatore sulla volatilità nel caso di innovazioni negative ed un impatto ridotto sulla volatilità nel caso di innovazioni positive.

Una tipologia alternativa di modelli per la volatilità sono quelli che considerano un termine di disturbo nell'equazione della varianza condizionale; questi vengono chiamati modelli a volatilità stocastica e, ad esempio, possono essere del tipo:

$$y_t = \varepsilon_t \sigma_t$$

$$\sigma_t = \sigma \exp\left(\frac{u_t}{2}\right)$$

$$\phi(B)u_t = \eta_t$$

² Nelson specifica il suo modello in maniera più generale riferendosi ad una variabile casuale z_t con distribuzione Generalized Error, ponendo questo termine uguale a $z_{t-1} - E(z_{t-1})$

dove $\sigma > 0$, $\{\eta_t\}$ sono *i.i.d.* $N(0, \sigma_\eta^2)$, $\{\eta_t\}$ e $\{\varepsilon_t\}$ sono indipendenti. Nel Capitolo 5 verranno ripresi e analizzati in maniera più profonda, ipotizzando che il processo u_t possa essere un processo a memoria lunga come un ARFIMA.

Capitolo 3

3. Cenni di analisi spettrale e modelli ARFIMA

L'idea naturale nello studio delle serie storiche è quella di analizzare una serie in funzione delle realizzazioni che assume a diversi istanti di tempo (si parla anche di analisi nel dominio del tempo). Questa impostazione non deve essere però presa per assoluta. Infatti si può rappresentare una serie anche come somma di funzioni periodiche con l'obiettivo di individuare dei cicli rilevanti nel comportamento della serie. Questo filone di ricerca prende il nome di analisi spettrale, o analisi nel dominio della frequenza.

3.1 Alcuni studi sulla densità spettrale

Altre incongruenze tra i risultati ottenuti dai dati reali e quelli che ci si sarebbero attesi sulla base della modellistica esistente provenivano dagli studi sulla densità spettrale, che gli analisti di serie storiche avevano iniziato a utilizzare all'inizio degli anni '60.

In prima approssimazione si può dire che l'analisi spettrale si occupa di scomporre la serie in cicli di diversa frequenza, e di analizzarne la dipendenza sulla base dell'andamento di una funzione che descrive il contributo di ciascuno di questi cicli all'andamento complessivo della serie; tale funzione prende il nome di densità spettrale. L'area racchiusa sotto la funzione di densità spettrale misura la

varianza della serie, mentre le aree “parziali” indicano la parte della varianza complessiva che è imputabile a cicli di frequenza inferiore a quella considerata.

Sulla base di tale risultato possiamo osservare che:

- un processo in cui le osservazioni sono incorrelate (ad esempio un white noise) deve avere spettro piatto a tutte le frequenze;
- la varianza di un processo che ha densità spettrale alta a frequenza basse è dovuta in larga parte a cicli di bassa frequenza, per cui la dipendenza di lungo termine del processo è più forte di quella di breve termine;
- la densità spettrale di un processo ciclico ha un picco in corrispondenza della frequenza del ciclo stesso;
- lo spettro di processi non stazionari è esplosivo per frequenze tendenti a zero;

Nel corso degli anni molti ricercatori hanno stimato gli spettri di diverse serie storiche, comprese anche quelle riguardanti i titoli azionari. Granger (1966), tentando di trovare un punto in comune fra tutte queste stime, ricavò che in tutte le serie analizzate la maggior parte della varianza è spiegata da cicli di frequenza bassa, il che indica che si ha una forte dipendenza delle osservazioni da quelle passate. Non si tratta, a dire il vero, di un risultato sorprendente in sé, visto che si ritiene che buona parte delle serie economiche possiedano dipendenza di questo tipo.

Il risultato insolito è che utilizzando gli strumenti ARMA tradizionali non si riusciva a liberare la serie dalla dipendenza (e a restituire quindi uno spettro dei residui piatto). La densità spettrale dei residui appariva infatti del tipo disegnato in Figura 3.1.

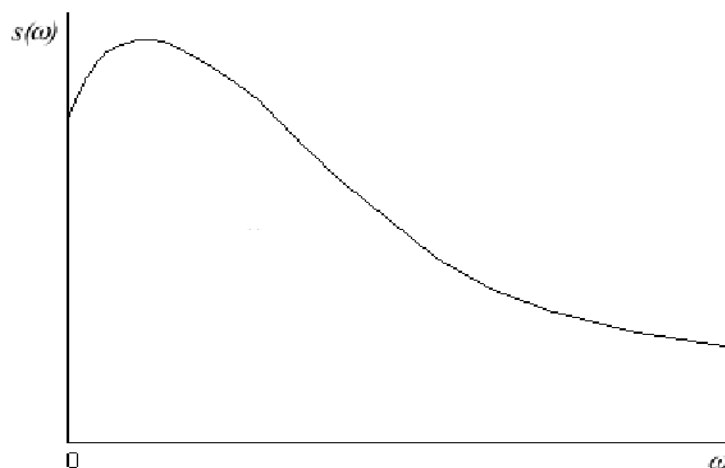


Figura 3. 1

Si può osservare che la densità spettrale resta forte alle basse frequenze, segno che permangono relazioni di lungo periodo che la modellistica ARMA non è stata in grado di cogliere. Si rendono quindi necessari modelli più adatti a cogliere tale dinamica di lungo termine.

3.1.1 La funzione generatrice delle autocovarianze

Sia $\{y_t\}$ un processo stazionario in covarianza. Si definisce la funzione generatrice delle autocovarianze come:

$$g(z) = \sum_{j=-\infty}^{+\infty} \gamma_j z^j$$

dove z indica uno scalare complesso. Quest'ultima espressione rappresenta un modo per raccogliere in un polinomio le autocovarianze di un generico processo: si ha infatti una somma di termini in z in cui il coefficiente indica il valore dell'autocovarianza, mentre l'esponente di z indica il ritardo a cui tale autocovarianza si riferisce. Si supponga ad esempio che $\{y_t\}$ sia un processo MA(1): per tale processo vale $\gamma_0 = (1 + \theta^2)\sigma^2$ e $\gamma_1 = \gamma_{-1} = \theta\sigma^2$. Si avrà allora:

$$\begin{aligned} g(z) &= (\theta\sigma^2)z^{-1} + [(1 + \theta^2)\sigma^2]z^0 + (\theta\sigma^2)z^1 = \\ &= (1 + \theta z)(1 + \theta z^{-1})\sigma^2 \end{aligned}$$

3.1.2 Lo spettro del processo

Limitando i valori di z che giacciono sul cerchio unitario (ovvero $z = e^{-i\omega}$), cosicché la funzione g dipenda esclusivamente da ω , si ottiene che lo spettro del processo è dato da:

$$s(\omega) = \frac{g(e^{-i\omega})}{2\pi} = \frac{1}{2\pi} \sum_{j=-\infty}^{+\infty} \gamma_j e^{-i\omega j} \quad (3.2)$$

Si noti che

$$e^{-i\omega j} = \cos(\omega j) - i \sin(\omega j)$$

e che, poichè $\{y_t\}$ è un processo stazionario in covarianza, si ha $\gamma_{-j} = \gamma_j$. Riscrivendo l'espressione dello spettro utilizzando queste ultime due proprietà si ha:

$$s(\omega) = \frac{\gamma_0}{2\pi} [\cos 0 - i \sin 0] + \frac{1}{2\pi} \left\{ \sum_{j=1}^{+\infty} \gamma_j [\cos(\omega j) + \cos(-\omega j) - i \sin(\omega j) - i \sin(-\omega j)] \right\}$$

che, utilizzando semplici proprietà trigonometriche, diventa:

$$s(\omega) = \frac{1}{2\pi} \left[\gamma_0 + 2 \sum_{j=1}^{+\infty} \gamma_j \cos(\omega j) \right] \quad (3.1)$$

Da quest'ultima espressione si apprezzano le proprietà principali dello spettro: la simmetria attorno allo zero (poiché $\cos(\omega)$ è simmetrico) e la periodicità, anch'essa dovuta alla periodicità del coseno. Pertanto, conoscendo lo spettro nell'intervallo $[0, \pi]$, si può ricavarlo su tutto \mathbf{R} .

Riprendendo l'esempio sopra, lo spettro di un processo MA (1) è dato da:

$$\begin{aligned} s(\omega) &= \frac{\sigma^2}{2\pi} (1 + \theta e^{-i\omega})(1 + \theta e^{i\omega}) = \\ &= \frac{\sigma^2}{2\pi} (1 + \theta(e^{-i\omega} + e^{i\omega}) + \theta^2) \end{aligned}$$

e, ricordando che

$$e^{-i\omega} + e^{i\omega} = \cos(\omega) - i \sin(\omega) + \cos(\omega) + i \sin(\omega) = 2 \cos(\omega)$$

in definitiva si ottiene che

$$s(\omega) = \frac{\sigma^2}{2\pi} (1 + 2\theta \cos \omega + \theta^2)$$

In questo modo lo spettro è stato costruito a partire dalle autocovarianze: si può fare anche l'operazione inversa, cioè ricavare le autocovarianze a partire dallo spettro. Si può dimostrare che:

$$\gamma_j = \int_{-\pi}^{\pi} s(\omega) e^{i\omega j} d\omega = \int_{-\pi}^{\pi} s(\omega) \cos(\omega j) d\omega.$$

Un ultimo risultato che necessita d'essere menzionato è il "teorema di rappresentazione spettrale", in quanto riguarda la possibilità, per un generico processo $\{y_t\}$, di passare da una rappresentazione classica in funzione del tempo a una rappresentazione che esprima il processo come somma di funzioni

periodiche. In particolare, un processo stazionario Y_t di media μ possiede una rappresentazione del tipo

$$Y_t = \mu + \int_0^{\pi} \alpha(\omega) \cos(\omega t) d\omega + \int_0^{\pi} \beta(\omega) \sin(\omega t) d\omega$$

dove $\alpha(\omega), \beta(\omega)$ sono variabili casuali a media nulla e, come al solito, ω indica la frequenza angolare generalmente espressa in radianti.

3.1.3 Il significato dello spettro

Se si prova a ricavare γ_0 , cioè la varianza del processo, dall'ultima espressione, si ottiene

$$\gamma_0 = \int_{-\pi}^{\pi} s(\omega) d\omega$$

Ciò significa, in termini più intuitivi, che l'area delimitata dallo spettro tra $-\pi$ e π corrisponde alla varianza. Questo risultato può essere esteso a un generico ω_1

compreso in $[-\pi, \pi]$: la quantità $\int_{-\omega_1}^{\omega_1} s(\omega) d\omega = 2 \int_0^{\omega_1} s(\omega) d\omega$ rappresenterà la

porzione di varianza del processo che è spiegata da cicli con frequenza inferiore a ω_1 .

La situazione sarà più chiara confrontando alcuni spettri. Nella figura 3.2 si è tracciato lo spettro di un processo MA (1) con $\psi = 1/2$ e varianza, per semplicità, unitaria. Ci si aspetta che un processo AR (1), a parità del parametro, abbia una dipendenza di lungo periodo più forte: ciò risalta subito confrontando la figura 3.2 con la figura 3.3, dove si è tracciato lo spettro di un AR (1) con $\phi = 1/2$ e varianza unitaria. Nella figura 3.4 invece viene presentato lo spettro di un processo random walk, cioè un AR (1) con radice unitaria: in questo caso, con una dipendenza di lungo termine fortissima, lo spettro è addirittura esplosivo per $\omega \rightarrow 0$.

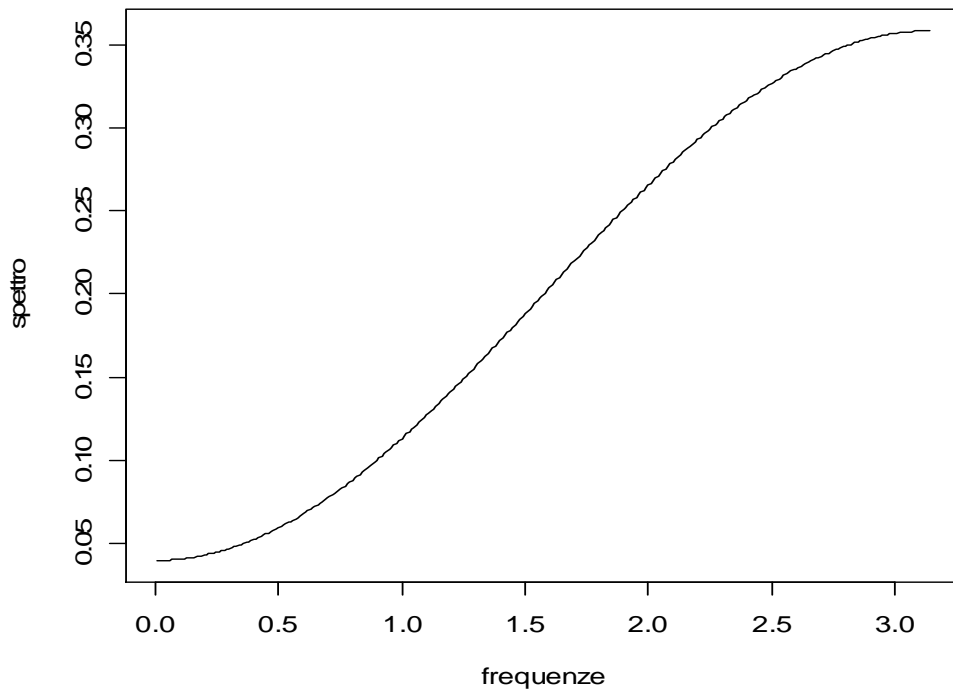


Figura 3. 2

Lo spettro di un processo MA(1) con $\theta = 1/2$

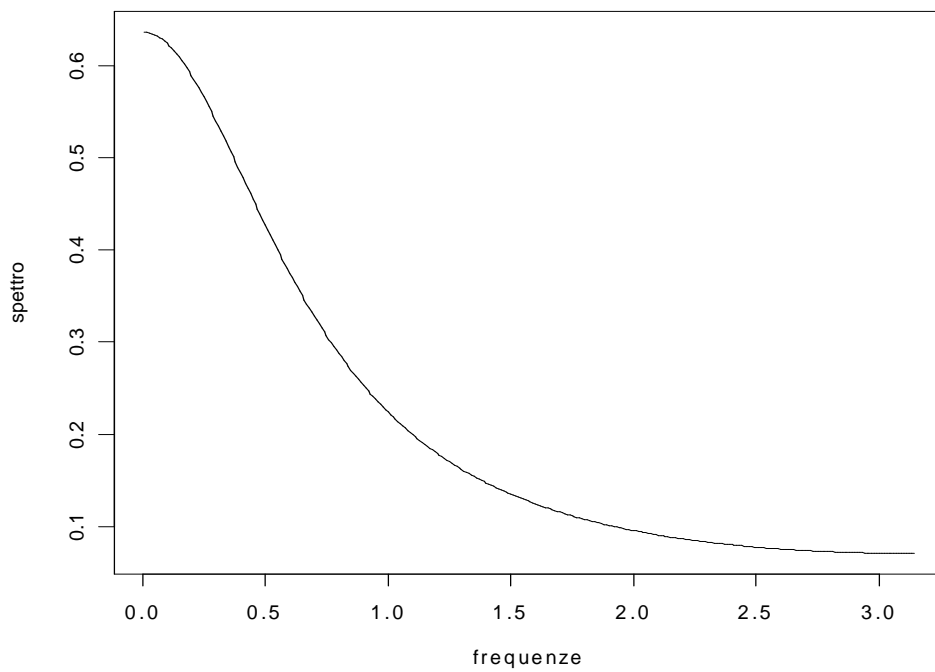


Figura 3. 3

Lo spettro di un processo AR(1) con $\phi = 1/2$

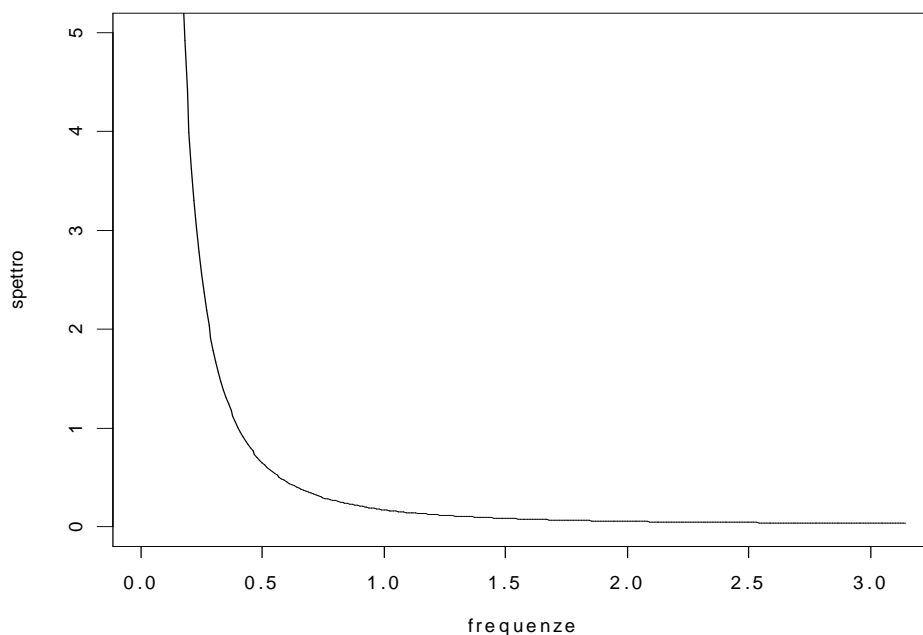


Figura 3. 4

Lo spettro di un random walk

3.1.4 Il periodogramma

Nel paragrafo 3.1.2 è stato presentato il procedimento necessario per costruire lo spettro a partire dalla forma analitica del processo generatore dei dati. All'atto pratico, però, ci si trova il più delle volte nella situazione opposta: si ha a disposizione un insieme di dati ma non si conosce la forma esatta del processo che li ha generati. Verrà presentato adesso il modo in cui stimare lo spettro a partire dai dati campionari, la cui stima prende il nome di periodogramma.

Si supponga di avere una serie di T osservazioni per un processo $\{ y_t \}$, stazionario e con processo generatore dei dati ignoto. Si possono stimare le autocovarianze a partire dai dati:

$$\hat{\gamma}_j = \hat{\gamma}_{-j} = \frac{1}{T} \sum_{t=j+1}^T (y_t - \bar{y})(y_{t-j} - \bar{y}) \quad (3.3)$$

dove \bar{y} indica la media campionaria delle osservazioni.

A questo punto basta inserire la (3.3) nella (3.2):

$$\hat{s}(\omega) = \frac{1}{2\pi} \sum_{j=-T+1}^{T+1} \hat{\gamma}_j e^{-i\omega j}$$

Ovviamente anche per il periodogramma vale quanto derivato nei paragrafi precedenti per lo spettro, in particolare la periodicità, la simmetria e l'espressione alternativa:

$$\hat{s}(\omega) = \frac{1}{2\pi} \left[\hat{\gamma}_0 + 2 \sum_{j=1}^{T-1} \hat{\gamma}_j \cos(\omega j) \right]$$

3.2 Il modello ARFIMA: specificazione e proprietà

Il più semplice modello della classe ARFIMA è il fractional white noise (detto, appunto, ARFIMA(0,d,0)). Per questo modello, la funzione di autocovarianze è data da

$$\gamma_k = \frac{\Gamma(1-2d)\Gamma(k+d)}{\Gamma(d)\Gamma(1-d)\Gamma(k+1-d)}$$

Quest'ultima espressione può essere anche scritta come (Hosking, 1981):

$$\gamma_k = \frac{(-1)^k \Gamma(-2d+1)}{\Gamma(-k+1-d)\Gamma(k+1-d)}$$

Sulla base di queste espressioni si ha $\gamma_0 = \frac{\Gamma(1-2d)}{\Gamma^2(1-d)}$ e quindi, passando alle autocorrelazioni:

$$\begin{aligned} \rho_k = \frac{\gamma_k}{\gamma_0} &= \frac{\Gamma(1-2d)\Gamma(k+d)\Gamma^2(1-d)}{\Gamma(d)\Gamma(1-d)\Gamma(k+1-d)\Gamma(1-2d)} = \\ &= \frac{\Gamma(1-d)\Gamma(k+d)}{\Gamma(d)\Gamma(k-d+1)} \cong \\ &\cong \frac{\Gamma(1-d)}{\Gamma(d)} k^{2d-1} \end{aligned}$$

Da quest'ultima espressione risalta un'importante caratteristica del fractional white noise, cioè l'andamento delle autocorrelazioni. Infatti si nota immediatamente che, per $0 < d < 1/2$, le autocorrelazioni decadono iperbolicamente (soddisfacendo la condizione di memoria lunga).

Per quanto riguarda il caso $-1/2 < d < 0$, non molto interessante da un punto di vista pratico, ci si limita ad osservare che ciò indica un processo a memoria corta e "antipersistente", secondo la definizione di Mandelbrot (1977). Le autocorrelazioni sono tutte negative (a parte, ovviamente, ρ_0 , che vale 1) e

tendono a zero al crescere di k , mentre la densità spettrale è dominata da frequenze alte e tende a zero al decrescere di ω .

Inoltre si ha anche che:

$$\lim_{\omega \rightarrow 0} s(\omega) = \omega^{-2d}$$

Nella figura 3.5 viene riportato a titolo di esempio lo spettro di un processo ARFIMA(0, 1/4, 0).

Dopo aver identificato le proprietà del più semplice processo a integrazione frazionaria, è naturale procedere oltre e definire un generico processo ARFIMA(p,d,q). Questo modello lo si ottiene combinando una rappresentazione a integrazione frazionaria con la consueta modellistica ARMA. Così facendo si ottiene un processo “a due velocità”: la dipendenza di breve periodo sarà catturata dai coefficienti ARMA, quella di lungo periodo dalla presenza del coefficiente di integrazione frazionaria d .

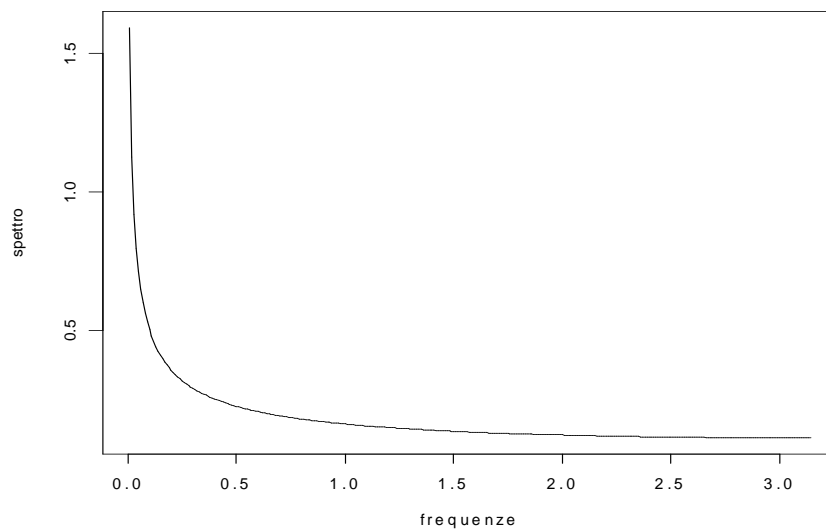


Figura 3.5

Lo spettro di un processo ARFIMA(0, 1/4, 0)

Sulla base della consueta notazione di Box e Jenkins (1976) si definisce il processo $\{y_t\}$ come ARFIMA (p,d,q) se il suo processo generatore dei dati è:

$$(1 - B)^d \phi(B)(y_t - \mu) = \theta(B)\eta_t$$

con η_t un white noise a varianza unitaria (in seguito si assumerà soventemente che $\mu = 0$). Ci si può aspettare che un processo ARFIMA (p,d,q) goda delle stesse

proprietà asintotiche del fractional white noise, visto che la differenza sta solo nella dipendenza di breve periodo. In effetti è dimostrato da Hosking (1981) che:

- $\{y_t\}$ è stazionario se $d < 1/2$ e tutte le radici di $\phi(B)$ giacciono fuori dal cerchio unitario;
- $\{y_t\}$ è invertibile se $d > -1/2$ e tutte le radici di $\theta(B)$ giacciono fuori dal cerchio unitario;
- $\lim_{\omega \rightarrow 0} \omega^{2d} s(\omega)$ esiste ed è finito
- $\lim_{k \rightarrow +\infty} k^{1-2d} \rho_k$ esiste ed è finito

Capitolo 4

4. Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity (FIGARCH)

Recenti studi sulla modellazione della media condizionale di molte serie finanziarie hanno messo in luce il ruolo della persistenza degli shocks e la necessità di testare la presenza di radici unitarie nella rappresentazione autoregressiva dei processi; da questi studi è emerso che la stretta distinzione fra $I(0)$ e $I(1)$ può essere troppo restrittiva. Per questo sono stati proposti dei modelli in cui $I(d)$ con $0 < d < 1$; a differenza dei processi $I(0)$, in cui la propagazione degli shocks sulla media è esponenziale, o dei processi $I(1)$, in cui la persistenza è infinita, nei processi $I(d)$ il tasso di decadimento dei coefficienti della rappresentazione $MA(\infty)$ è iperbolico. Analogamente a queste questioni poste per la media condizionale, simili domande diventano rilevanti nel modellare la varianza condizionale.

Queste osservazioni hanno portato Engle e Bollerslev (1986) a formulare un nuovo tipo di processi, chiamati IGARCH (Integrated GARCH), che possiedono molte delle caratteristiche dei processi a radice unitaria, o $I(1)$, per la media; però, studi condotti da de Lima, Breidt e Crato (1994), Dacorogna et al. (1993), Ding, Granger ed Engle (1993) e Harvey (1993) documentano l'apparente presenza di memoria lunga nelle autocorrelazioni dei quadrati o nei valori assoluti dei rendimenti di svariati titoli azionari; questo fatto ha condotto Baillie, Bollerslev e

Mikkelsen (1996) a formulare una nuova classe di modelli, chiamati FIGARCH (Fractionally Integrated GARCH). Questo filone rappresenta tutt'oggi una delle applicazioni più interessanti e più promettenti dell'integrazione frazionaria; difatti l'utilizzo di un modello a integrazione frazionaria consente di avere una modellazione per la varianza condizionale più aderente alla realtà e, grazie alla memoria lunga, avere dei vantaggi operativi assai rilevanti.

Per rendersene conto basta passare in rassegna tutte quelle situazioni in cui la modellistica ARCH tradizionale viene utilizzata per stimare la volatilità.

Per fare alcuni esempi particolarmente significativi: calcolo del *Value at Risk* di un certo portafoglio, prezzaggio delle opzioni secondo Black e Scholes (1973). L'importanza della modellistica a memoria lunga per la varianza condizionata è evidente: se cerchiamo di prevedere la volatilità su un orizzonte lungo utilizzando un tradizionale ARCH otterremo, a causa della memoria corta, una previsione poco distante o addirittura coincidente con la varianza non condizionata; ciò significa, all'atto pratico, trascurare l'eteroschedasticità. Al contrario, sfruttando la proprietà di memoria lunga, si potrà ottenere una previsione più significativa e più sensibile alle osservazioni passate.

Quella delle previsioni sulla volatilità è una tematica particolarmente interessante sia per gli accademici che per gli operatori professionali. Se con l'introduzione della modellistica ARCH si è potuto gettare uno sguardo sulla dinamica di breve periodo della varianza condizionata, con l'utilizzo dell'integrazione frazionaria si può fare un passo avanti e volgere l'attenzione al medio e lungo periodo.

4.1 Definizione e proprietà del modello FIGARCH

I processi FIGARCH combinano alcune caratteristiche dei processi integrati frazionari per la media (ARFIMA) con i processi GARCH per la varianza condizionale; in particolare i processi FIGARCH implicano un lento decadimento di tipo iperbolico per i coefficienti dei ritardi dei quadrati delle innovazioni nella funzione della varianza condizionale.

Considerando la specificazione data da Bollerslev (1986), un processo GARCH(p,q) può essere definito dalle seguenti condizioni:

$$y_t = z_t \sigma_t = \varepsilon_t$$

$$z_t \sim IID(0,1) \text{ e di conseguenza } \varepsilon_t | I_{t-1} \sim (0, \sigma_t^2)$$

$$\begin{aligned} \sigma_t^2 &= \omega + \alpha_1 y_{t-1}^2 + \dots + \alpha_p y_{t-p}^2 + \beta_1 \sigma_{t-1}^2 + \dots + \beta_q \sigma_{t-q}^2 \\ &= \omega + \alpha(B) y_t^2 + \beta(B) \sigma_t^2 \end{aligned}$$

in cui B è l'operatore ritardo e $\alpha(B) = \alpha_1 B + \dots + \alpha_p B^p$ e $\beta(B) = \beta_1 B + \dots + \beta_q B^q$; per la non negatività della varianza condizionata bisogna imporre le condizioni $\omega > 0, \alpha_i \geq 0, \beta_j \geq 0$.

In alternativa, tenendo conto della relazione $y_t^2 = \sigma_t^2 + v_t$, il processo GARCH(p,q) può essere riscritto come un ARMA(m,p) su y_t^2 , con $m = \max\{p, q\}$, in cui il processo $\{v_t\}$ è interpretato come "innovazione" della varianza condizionata; infatti si ha

$$\begin{aligned} y_t^2 - v_t &= \omega + \alpha(B) y_t^2 + \beta(B) (y_t^2 - v_t) \\ &= \omega + \alpha(B) y_t^2 + \beta(B) y_t^2 - \beta(B) v_t \end{aligned}$$

in definitiva si ottiene:

$$[1 - \alpha(B) - \beta(B)] y_t^2 = \omega + [1 - \beta(B)] v_t$$

La persistenza che si osserva suggerisce che l'equazione $[1 - \alpha(B) - \beta(B)] = 0$ possa avere una radice unitaria. In questo caso si può scrivere

$$\phi(B)(1 - B) y_t^2 = \omega + [1 - \beta(B)] v_t \quad (4.1)$$

dove $\phi(B)$ è un polinomio d'ordine m-1:

$$\phi(B) = [1 - \alpha(B) - \beta(B)](1 - B)^{-1} = (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_{m-1} B^{m-1})$$

con $\phi_i = \alpha_i + \beta_i$.

Il processo illustrato dall'Eq. (4.1) viene chiamato IGARCH(p,q), ovvero Integrated GARCH. Per permettere di avere persistenza e memoria lunga nella varianza condizionata, l'operatore differenza prima viene sostituito da un operatore di differenza frazionaria $(1 - B)^d$, la cui espansione è data da

$$(1 - B)^d = \frac{(j - d - 1)!}{j!(-d - 1)!} B^j \quad \text{con } j = 1, 2, \dots, \infty \quad (4.2)$$

Fatte queste considerazioni, il modello per y_t^2 si ottiene dall'Eq. (4.1), ovvero

$$\phi(B)(1-B)^d y_t^2 = \omega + [1 - \beta(B)]v_t \quad (4.3)$$

dove si ha che $0 < d < 1$ e tutte le radici di $\phi(B)$ e di $[1 - \beta(B)]$ sono esterne al cerchio unitario. Si può notare dall'Eq. (4.3) che i rendimenti (residui) al quadrato differenziati in maniera frazionaria, $(1-B)^d y_t^2$, seguono un processo ARMA(m,q) e, di conseguenza, i residui (rendimenti) al quadrato y_t^2 seguono un processo ARFIMA(m,d,q). Sfruttando la solita relazione $y_t^2 = \sigma_t^2 + v_t$ si può riscrivere l'Eq. (4.2) nelle seguenti

$$\begin{aligned} \phi(B)(1-B)^d y_t^2 &= \omega + [1 - \beta(B)]v_t \\ \phi(B)(1-B)^d y_t^2 &= \omega + [1 - \beta(B)](y_t^2 - \sigma_t^2) \\ [1 - \beta(B)]\sigma_t^2 &= \omega + [1 - \beta(B) - \phi(B)(1-B)^d]y_t^2 \end{aligned}$$

Dunque l'espressione per la varianza condizionata di un modello FIGARCH(m,d,q) è data da:

$$\begin{aligned} \sigma_t^2 &= \omega + \beta(B)\sigma_t^2 + [1 - \beta(B) - \phi(B)(1-B)^d]y_t^2 \\ &= \omega + \sum_{i=1}^q \beta_i \sigma_{t-i}^2 + [1 - \beta(B) - \phi(B)(1-B)^d]y_t^2 \end{aligned} \quad (4.4)$$

Quando $0 < d < 1$, la dinamica di breve periodo viene descritta dai coefficienti di $\phi(B)$ e $\beta(B)$, mentre il parametro d della differenza frazionaria modella le caratteristiche di lungo periodo della volatilità; analogamente a ciò che si verifica per la classe IGARCH, la presenza di un valore $\omega > 0$ indica la presenza di un drift, e implica quindi che il processo non sia stazionario.

Relativamente ai coefficienti, vale la pena rimarcare l'osservazione di Chung (1999) in merito al significato del coefficiente ϕ . Nel FIGARCH di Baillie Bollerslev e Mikkelsen (1996) non viene imposto alcun vincolo su tale coefficiente, cosicché non è necessariamente $\phi(B) = 1 - \alpha(B) - \beta(B)$ come nel GARCH; in questa situazione un FIGARCH(r,0,m) non è equivalente a un GARCH(r,m) e il coefficiente ϕ assume un significato autonomo ancora da chiarire.

Un'espressione alternativa per la varianza condizionata è:

$$\begin{aligned}\sigma_t^2 &= \omega[1 - \beta(q)]^{-1} + \left\{1 - [1 - \beta(B)]^{-1} \phi(B)(1 - B)^d\right\} y_t^2 \\ &= \omega[1 - \beta(q)]^{-1} + \lambda(B)y_t^2\end{aligned}\quad (4.5)$$

in cui si è posto che $\lambda(B) = \lambda_1 B + \lambda_2 B^2 + \dots$ e $\beta(q) = \beta_1 + \beta_2 + \dots + \beta_q$.

Ad esempio, la formulazione per un modello FIGARCH(1,d,1) è

$$\sigma_t^2 = \omega + \beta_1 \sigma_{t-1}^2 + \left[1 - \beta_1 B - (1 - \phi_1 B)(1 - B)^d\right] y_t^2$$

dove $\phi_1 = \alpha_1 + \beta_1$ è definito come nell'EQ. (4.1). Quest'ultima espressione può essere riscritta equivalente mente nella forma ARCH infinito

$$\sigma_t^2 = \frac{\omega}{1 - \beta_1} + \left[1 - \frac{(1 - \phi_1 B)(1 - B)^d}{(1 - \beta_1 B)}\right] y_t^2 = \frac{\omega}{1 - \beta_1} + \lambda_1(B)y_t^2$$

dove $\lambda_1(B) = \sum_{i=1}^{\infty} \lambda_{1,i} B^i$. In maniera analoga, il modello FIGARCH(1,d,0) si ottiene

ponendo $\phi_1 = 0$ nell'espressione precedente; in questo caso si ha

$$\sigma_t^2 = \omega + \beta_1 \sigma_{t-1}^2 + \left[1 - \beta_1 B - (1 - B)^d\right] y_t^2$$

e la forma ARCH infinito diventa

$$\sigma_t^2 = \frac{\omega}{1 - \beta_1} + \left[1 - \frac{(1 - B)^d}{(1 - \beta_1 B)}\right] y_t^2 = \frac{\omega}{1 - \beta_1} + \lambda_2(B)y_t^2$$

dove $\lambda_2(B) = \sum_{i=1}^{\infty} \lambda_{2,i} B^i$. Seguendo Bollerslev e Mikkelsen (1996), i coefficienti

ARCH nell'operatore ritardo $\lambda_1(B)$ hanno la forma ricorsiva

$$\lambda_{1,i} = \phi_1 - \beta_1 + d$$

per $i=1$ e

$$\lambda_{1,i} = \beta_1 \lambda_{1,i-1} + \left[\frac{(i-1-d)}{i} - \phi_1\right] \delta_{i-1}$$

per $i=2,3,\dots$ dove $\delta_i = \frac{\delta_{i-1}(i-1-d)}{i}$ è una funzione ricorsiva per i coefficienti

nell'espressione di $(1 - B)^d$ nell'Eq. (4.2). I coefficienti dell'operatore ritardo $\lambda_2(B)$ si ottengono ponendo $\phi_1 = 0$ nelle due espressioni per $\lambda_{1,i}$.

Quindi, per il processo FIGARCH(m,d,q) scritto nell'Eq. (4.3), affinché si abbia che la varianza condizionata sia non negativa è necessario che tutti i coefficienti della rappresentazione ARCH infinito nell'Eq. (4.5) siano non negativi,

cioè $\lambda_i \geq 0$ per $i=1,2,\dots$. Dall'Eq. (4.5) si può notare che la varianza condizionale di un modello FIGARCH, come per un IGARCH, dipende da infiniti ritardi di y_t^2 ; un'altra similitudine con il modello IGARCH è che anche per un FIGARCH il momento secondo della distribuzione non condizionata di y_t è infinito e dunque il processo chiaramente non è debolmente stazionario. Tuttavia, come dimostrato da Nelson (1990a) per un IGARCH(1,1) ed esteso in generale per un IGARCH(p,q) da Bougerol e Picard (1992), se $0 \leq d \leq 1$ il modello IGARCH è strettamente stazionario ed ergodico: lo stesso fatto vale per un modello FIGARCH(p,d,q) e discende da una diretta estensione della dimostrazione per il caso dell'IGARCH(p,q). Comunque, le principali differenze fra un IGARCH ed un FIGARCH risiedono nel modo in cui la funzione di autocorrelazioni di y_t^2 tende a zero e nel modo in cui eventuali shocks nella varianza condizionata esauriscono il loro effetto. Mentre nei modelli IGARCH uno shock persiste indefinitamente, per i modelli GARCH e FIGARCH prima o poi cessano di far sentire il loro effetto: per un GARCH a velocità esponenziale, per un FIGARCH a velocità iperbolica.

4.2 Stima e Inferenza

La stima di un convenzionale modello GARCH non presenta particolari difficoltà poiché le restrizioni sui parametri del modello per assicurare la positività della varianza condizionale sono sempre verificate. L'approccio più comune per procedere con la stima è assumere la normalità condizionale del processo, ovvero che $y_t | I_{t-1} \sim N(0, \sigma_t^2)$. Sotto questa assunzione, o se il processo generatore dei dati è noto, si può applicare il metodo della massima verosimiglianza.

Nel caso di un modello FIGARCH, la complessità aumenta poiché nell'espressione (4.2) del filtro $(1-B)^d$ vi sono dei coefficienti negativi che non possono essere eliminati; questo implica che la classe di modelli più semplice per la varianza condizionale a memoria lunga, FIGARCH(0,d,0), non può essere definita perché vi sono dei coefficienti negativi nell'espressione (4.4) che non possono essere estromessi. Per il più generale modello FIGARCH(p,d,q), non è disponibile un'espressione generale delle restrizioni sui parametri ma, come hanno fatto notare Bollerslev e Mikkelsen (1996), le restrizioni necessarie per

specifici modelli FIGARCH(p,d,q) possono essere ottenute studiando caso per caso.

Ad esempio, Per un modello FIGARCH(1,d,1) queste sono

$$\beta_1 - d \leq \phi_1 \leq \frac{(2-d)}{3}$$

e

$$d \left[\phi_1 - \frac{(1-d)}{2} \right] \leq \beta_1 (\phi_1 - \beta_1 + d)$$

mentre Chung (1999) suggerisce

$$0 \leq \phi_1 \leq \beta_1 \leq d \leq 1$$

4.2.1 Stima della Massima Verosimiglianza

Assumendo la normalità condizionale, è possibile stimare (4.4) massimizzando il logaritmo naturale della verosimiglianza di un convenzionale modello ARCH:

$$\log L(y_1, y_2, \dots, y_T | \theta) = -\frac{T \log(2\pi)}{2} - \frac{1}{2} \sum_{t=1}^T \left[\log(\sigma_t^2) + \frac{y_t^2}{\sigma_t^2} \right]$$

dove $\theta' \equiv (\omega, d, \beta_1, \dots, \beta_p, \phi_1, \dots, \phi_q)$ e σ_t^2 si ottiene approssimativamente in accordo con l'Eq. (4.4) usando il filtro differenziale $(1-B)^d$ troncato in qualche punto oltre il quale il contributo dei termini successivi alla sommatoria è trascurabile.

A questo punto è bene aprire una breve parentesi: se le osservazioni y sono indipendenti, la funzione di densità di probabilità congiunta può essere scritta come:

$$f(y_1, y_2, \dots, y_T | \theta) = f(y_1 | \theta) \cdot f(y_2 | \theta) \cdots f(y_T | \theta) = \prod_{t=1}^T f(y_t | \theta)$$

Se invece le osservazioni y non sono indipendenti, la densità congiunta dell'intero campione può essere fattorizzata sequenzialmente come:

$$f(y_1, y_2, \dots, y_T | \theta) = f(y_T | y_{T-1}, \dots, y_1; \theta) \cdot f(y_{T-1} | y_{T-2}, \dots, y_1; \theta) \cdot f(y_{p+1} | y_p, y_{p-1}, \dots, y_1; \theta) \cdot f(y_p, y_{p-1}, \dots, y_1; \theta)$$

Si ottiene così un prodotto di funzioni di densità di probabilità di una variabile casuale in un certo istante condizionata ai valori assunti dalle variabili casuali

precedenti e dalla funzione di probabilità congiunta di un certo numero p di variabili casuali riferite ai periodi iniziali (il vantaggio è che talvolta si conosce la distribuzione condizionata).

Detto ciò la verosimiglianza diviene

$$L(\theta) = f(y_1, y_2, \dots, y_T | \theta) = f(y_p, y_{p-1}, \dots, y_1; \theta) \cdot \prod_{t=p+1}^T f(y_t | I_{t-1}; \theta)$$

e la log-verosimiglianza è

$$l(\theta) = \log L(y_1, y_2, \dots, y_T | \theta) = \log f(y_p, y_{p-1}, \dots, y_1; \theta) + \sum_{t=p+1}^T \log f(y_t | I_{t-1}; \theta)$$

Per quanto riguarda il termine delle condizioni iniziali, $f(y_p, y_{p-1}, \dots, y_1; \theta)$, si possono fare alcune considerazioni:

- Se p è piccolo si può fare riferimento alla distribuzione non condizionata (sempre se è nota);
- A volte si sostituiscono i valori $y_{-1}, y_{-2}, \dots, y_{-p}$ con dei valori precampionati (procedura di *backcasting*);
- Ai fini della procedura di stima, se T è sufficientemente grande, il primo addendo può essere trascurato (verosimiglianza condizionata).

Un problema che merita di essere affrontato è quello dei valori pre-campionari. In genere nelle stime ARCH si suole condizionare a valori pari alla varianza non condizionata del processo. Ma come è stato già osservato, per il FIGARCH come per l'IGARCH, la varianza non condizionata non esiste. Pertanto si può procedere in due modi: o si condiziona a valori “eliminati” dalla serie originale, oppure, seguendo la proposta di Baillie, Bollerslev e Mikkelsen (1996), si utilizza, al posto della varianza non condizionata, la media dei residui al quadrato. Comunque le condizioni sui valori precampionati non hanno effetti sulla distribuzione asintotica degli stimatori e dei test statistici.

Detto ciò, se si indica con $\hat{\theta}_T = \arg \max_{\theta} l(\theta)$, ovvero il vettore dei valori dei parametri che rendono massima la log-verosimiglianza, si hanno le seguenti proprietà asintotiche di $\hat{\theta}_T$:

Normalità: $\sqrt{T}(\hat{\theta}_T - \theta) \sim N(0, V_{as})$

$$\hat{\theta}_T \sim N(\theta, T^{-1}V_{as})$$

in cui si indica con θ il vero valore dei parametri e con $V_{as} = J^{-1}$ data da

$$J = E \left[- \frac{\partial^2 l_t(\theta)}{\partial \theta \partial \theta'} \right] = E \left[\frac{\partial l_t(\theta)}{\partial \theta} \cdot \frac{\partial l_t(\theta)}{\partial \theta'} \right]$$

dove le derivate sono calcolate in $\theta = \hat{\theta}$.

Una stima di J è data da:

$$\hat{J} = - \frac{1}{T} \sum_{t=1}^T \frac{\partial^2 l_t(\theta)}{\partial \theta \partial \theta'} \Big|_{\theta=\hat{\theta}} = \frac{1}{T} \sum_{t=1}^T \frac{\partial l_t(\theta)}{\partial \theta} \cdot \frac{\partial l_t(\theta)}{\partial \theta'} \Big|_{\theta=\hat{\theta}}$$

Ci sono alcune considerazioni da fare riguardanti questa procedura di stima. In primo luogo, il troncamento del filtro differenziale $(1-B)^d$ implica che la stima di massima verosimiglianza è solo approssimativa, sebbene il grado di approssimazione possa essere ridotto incrementando il punto di troncamento. In secondo luogo, sempre a causa di questo troncamento, le stime possono essere poco precise in piccoli campioni; sebbene questo non sia un grosso problema per i data sets finanziari, questo comunque implica la perdita di dati e quindi di informazioni. In terzo luogo, il troncamento implica che la serie utilizzata non è la “vera” serie differenziata frazionalmente e quindi viene introdotta della distorsione nella procedura di stima .

4.2.2 Stima di Quasi Massima Verosimiglianza

In molte applicazioni pratiche con dati ad alta frequenza si è notato che le innovazioni standardizzate, $z_t = y_t \sigma_t^{-1}$, sono leptocurtiche e non sono normali i.i.d. In questa situazione (o anche quando il modello non è correttamente specificato) si può utilizzare una procedura di Quasi Massima Verosimiglianza.

Si supponga di avere una serie di dati $\{y_t\}$ generati a partire da una certa distribuzione $G(y)$, della quale non si è in grado di descrivere a priori la forma. Sia invece $\{F(y|\theta) : \theta \in \Theta\}$ una famiglia di distribuzioni di probabilità nota in cui si ritiene (erroneamente o no) sia inclusa la vera distribuzione $G(y)$.

Per motivare l'utilizzo della massima verosimiglianza anche nel caso in cui $G(y)$ non sia in $\{F(y|\theta):\theta\in\Theta\}$ viene introdotto il criterio d'informazione di Kullback-Leibler (KLIC) definito come

$$KLIC(g : f) = E_g \left[\ln \frac{g(y)}{f(y|\theta)} \right] = E_g[\ln g(y)] - E_g[\ln f(y|\theta)] \quad (4.6)$$

Il KLIC è una misura della discrepanza tra la vera distribuzione G e la distribuzione $F(y|\theta)$.

Si può dimostrare che

$$KLIC(g : f) \geq 0$$

$$KLIC(g : f) = 0 \quad \Leftrightarrow \quad f = g \text{ quasi ovunque}$$

Il primo addendo della (4.6) è una costante ignota, dunque per minimizzarla si deve trovare il valore θ^* che massimizza $E_g[\ln f(y|\theta)]$; se il modello è correttamente specificato θ^* è quel valore che rende nulla la discrepanza, altrimenti è il valore che la minimizza senza mai annullarla. Uno stimatore naturale di $E_g[\ln f(y|\theta)]$ è:

$$l_T(\theta | y) = \frac{1}{T} \sum_{t=1}^T \ln f(y_t | \theta)$$

che, a meno di un fattore moltiplicativo T^{-1} non è altro che la funzione di log-verosimiglianza del modello mal specificato: la funzione viene detta funzione di quasi log-verosimiglianza. Lo stimatore di quasi verosimiglianza

$$\tilde{\theta}_T = \arg \max_{\theta} l_T(\theta)$$

è dunque uno stimatore di θ^* e può essere interpretato come stimatore di “minima ignoranza” rispetto al vero modello.

White (1982) ha dimostrato, sotto alcune condizioni, che:

$$p \lim_{T \rightarrow +\infty} \tilde{\theta}_T = \theta^*$$

Questa proprietà consente di utilizzare con una certa sicurezza la stima di massima verosimiglianza anche quando non si è sicuri di aver specificato bene il modello. Infatti, in caso di errata specificazione, si avrà una stima di quasi-massima verosimiglianza che convergerà, per $T \rightarrow +\infty$, al vettore θ^* . Tale vettore, pur non corrispondendo necessariamente alla soluzione “vera”,

minimizzerà comunque, per come è definito, l'ignoranza sulla struttura del modello.

Sfortunatamente la consistenza e la normalità asintotica per le stime QML sono state formalmente dimostrate solo per un IGARCH(1,1) e per un FIGARCH(1,d,0), mentre non sono ancora dimostrate per un generico IGARCH(p,q) e per un generico FIGARCH(p,d,q). Però, nei casi sopraccitati, si può dimostrare sotto alcune ipotesi³, si veda Lee e Hansen (1994), che la stima di quasi massima verosimiglianza $\tilde{\theta}_T$ è consistente e distribuita asintoticamente normalmente, ovvero

$$\sqrt{T}(\tilde{\theta}_T - \theta^*) \sim N(0, V_{as})$$

$$\tilde{\theta}_T \sim N(\theta^*, T^{-1}V_{as})$$

dove $V_{as} = J^{-1} \cdot I \cdot J^{-1}$, ed inoltre

$$I = E_0 \left[\frac{\partial l_t(\theta)}{\partial \theta} \cdot \frac{\partial l_t(\theta)}{\partial \theta'} \right] \quad J = E_0 \left[- \frac{\partial^2 l_t(\theta)}{\partial \theta \theta'} \right]$$

dove J è la matrice Hessiana, I è il prodotto dei gradienti ed E_0 sta ad indicare che i valori attesi vengono fatti rispetto alla vera distribuzione. Le matrici I e J possono essere stimate a partire dai dati sostituendo il valore atteso E_0 con la media empirica e l'ignoto parametro θ con il suo stimatore $\tilde{\theta}_T$

$$\hat{J}_T = - \frac{1}{T} \sum_{t=1}^T \frac{\partial^2 l_t(\tilde{\theta}_T)}{\partial \theta \theta'}$$

$$\hat{I}_T = \frac{1}{T} \sum_{t=1}^T \frac{\partial l_t(\tilde{\theta}_T)}{\partial \theta} \cdot \frac{\partial l_t(\tilde{\theta}_T)}{\partial \theta'}$$

Il comportamento di questo stimatore è stato messo alla prova su dati simulati. Baillie, Bollerslev e Mikkelsen (1996) fissano M a mille e studiano otto differenti

³ Le ipotesi sono:

- i. z_t è stazionario ed ergodico
- ii. z_t^2 è non degenere
- iii. $E_{T-1}(z_t^4) \leq k < \infty$
- iv. $\sup_t E_{T-1}[\log(\beta_1 + \alpha_1 z_t^2)] < \infty$

specificazioni FIGARCH. In tutti i modelli la distorsione per \tilde{d} è molto bassa, e raddoppiando la dimensione del campione si riduce all'incirca di $\sqrt{2}$, chiaro segnale della \sqrt{T} – consistenza. Si osserva inoltre che le stime più precise di d si hanno quando il processo generatore dei dati è un GARCH ($d = 0$) o un IGARCH ($d = 1$); ciò indica che lo stimatore è in grado di distinguere tra i modelli e può quindi essere utilizzato anche senza un'analisi preventiva sulla presenza di integrazione frazionaria.

Stimando invece i dati simulati con un semplice GARCH (1, 1) si ottiene un altro risultato importante: la somma dei coefficienti si avvicina al cerchio unitario; ciò indica che, se si utilizzano esclusivamente dei modelli GARCH, l'integrazione frazionaria può essere scambiata per la presenza di una radice unitaria.

Relativamente all'efficienza degli stimatori QML, è facile rendersi conto che essa non sussiste. Pur senza darne una prova formale, si osserva che, utilizzando uno stimatore QML, si tralascia parte dell'informazione disponibile, cioè la forma dei termini di disturbo; in questo senso la matrice d'informazione ha valori più elevati rispetto a quelli di una semplice stima di massima verosimiglianza, e lo stimatore QML è meno efficiente.

Capitolo 5

5. Modelli a Volatilità Stocastica con Memoria Lunga (LMSV)

Molte ricerche suggeriscono che la volatilità condizionale di molti titoli finanziari presenta memoria lunga: questo tipo di persistenza non viene modellato in maniera appropriata da modelli tipo GARCH o a volatilità stocastica standard (con memoria corta). In questa sezione, seguendo la specificazione data da Breidt, Crato e de Lima (1998), verrà specificata una classe di modelli chiamata LMSV (Long Memory Stochastic Volatility) che permette di catturare la persistenza nella volatilità condizionale.

Questi tipi di modelli sono costruiti incorporando un processo ARFIMA in un modello a volatilità stocastica standard.

5.1 La persistenza nella volatilità

Un processo debolmente stazionario ha “memoria corta” quando, chiamata $\rho(h)$ la sua funzione di autocorrelazione, si ha:

$$|\rho(h)| \leq Cr^{|h|} \quad \text{con } C > 0, 0 < r < 1$$

A differenza dei processi con memoria corta, la cui funzione di autocorrelazione (ACF) ha un decadimento esponenziale, un processo debolmente stazionario ha memoria lunga se $\rho(h)$ ha un decadimento iperbolico, ovvero:

$$\rho(h) \sim Ch^{2d-1} \text{ con } h \rightarrow \infty, C \neq 0, d < 0.5$$

Ovviamente un comportamento di questo tipo ha i suoi effetti anche sullo spettro. Infatti, data la relazione

$$s(\omega) = \frac{1}{2\pi} \left[\gamma_0 + 2 \sum_{j=1}^{+\infty} \gamma_j \cos(\omega j) \right] \quad (5.1)$$

dove $s(\omega)$ indica lo spettro, si osserva che, se il processo ha memoria lunga, si ha

$$\sum_{j=-\infty}^{\infty} \gamma_j = 2 \sum_{j=0}^{\infty} \gamma_j = +\infty$$

Per $\omega \rightarrow 0$, $\cos(\omega j) \rightarrow 1$ per cui il secondo addendo della (5.1) non converge. Ciò implica che:

$$\lim_{\omega \rightarrow 0} s(\omega) = +\infty$$

5.2 Un modello per la volatilità stocastica con memoria lunga

Un approccio alternativo per descrivere l'evoluzione della volatilità delle serie storiche finanziarie è di introdurre un'innovazione nell'equazione della varianza condizionale di y_t . In maniera analoga ad un EGARCH, per assicurare la positività della varianza condizionale, il modello con volatilità stocastica (SV) usa $\log(\sigma_t^2)$; quindi un possibile modello è definito da

$$y_t = \varepsilon_t \sigma_t$$

$$(1 - \alpha_1 B + \dots + \alpha_m B^m) \ln(\sigma_t^2) = \alpha_0 + v_t$$

dove $\{\varepsilon_t\}$ sono *i.i.d.* $N(0,1)$, $\{v_t\}$ sono *i.i.d.* $N(0, \sigma_v^2)$, $\{\varepsilon_t\}$ e $\{v_t\}$ sono indipendenti e tutti gli zeri del polinomio $1 - \sum_{i=1}^m \alpha_i B^i$ sono in modulo maggiori di uno.

L'introduzione delle innovazioni $\{v_t\}$ sostanzialmente incrementa la flessibilità del modello nel descrivere l'evoluzione di σ_t^2 , ma incrementa la difficoltà nella stima dei parametri: questo è dovuto al fatto che per ogni shock y_t il modello utilizza due innovazioni, ε_t e v_t . Più recentemente il modello a volatilità

stocastica è stato ampliato, usando l'idea della differenza frazionaria, per catturare la presenza di memoria lunga nella volatilità.

Un semplice modello con volatilità stocastica a memoria lunga (LMSV) è dato da

$$y_t = \varepsilon_t \sigma_t$$

$$\sigma_t = \sigma \exp\left(\frac{u_t}{2}\right)$$

$$(1-B)^d u_t = \eta_t$$

dove $\sigma > 0$, $\{\eta_t\}$ sono *i.i.d.* $N(0, \sigma_\eta^2)$, $\{\eta_t\}$ e $\{\varepsilon_t\}$ sono indipendenti e $-0.5 < d < 0.5$.

Dato che è stato considerato il caso in cui $\{u_t\}$ ⁴ sia gaussiano, ne deriva che y_t è sia strettamente che debolmente stazionario. Se si indica con $\gamma(\cdot)$ la funzione di autocovarianza di $\{u_t\}$ allora la struttura di covarianza di y_t la si ottiene dalle proprietà della distribuzione lognormale:

$$E[y_t] = 0 \quad \text{Var}(y_t) = \exp\{\gamma(0)/2\} \sigma^2$$

$$\text{Cov}(y_t, y_{t+h}) = 0 \quad \text{con } h \neq 0$$

Quindi il processo $\{y_t\}$ ⁵ è una sequenza white noise; anche il processo $\{y_t^2\}$ è sia strettamente che debolmente stazionario (anche in questo caso i suoi momenti possono essere ottenuti dalle proprietà della distribuzione lognormale).

⁴ Inoltre, per il processo $\{u_t\}$, si ha che (Brockwell e Davis, 1991, Bertelli e Caporin, 2002):

$$f(\lambda) = \frac{\sigma_\eta^2}{2\pi} \left| 1 - e^{-i\lambda} \right|^{-2d}, \quad -\pi \leq \lambda \leq \pi$$

$$\gamma(0) = \sigma_\eta^2 \frac{\Gamma(1-2d)}{\Gamma^2(1-d)}$$

$$\gamma(k) = \sigma_\eta^2 \frac{\Gamma(k+d)\Gamma(1-2d)}{\Gamma(k+1-d)\Gamma(1-d)\Gamma(d)}$$

$$\rho(h) = \frac{\Gamma(h+d)\Gamma(1-d)}{\Gamma(h-d-1)\Gamma(d)}, \quad h = 1, 2, \dots$$

⁵ L'eccesso di curtosi del processo $\{y_t\}$ è dato, quando l'errore $\{\varepsilon_t\}$ è gaussiano, da:

$$\frac{E[y_t^4]}{E[y_t^2]^2} - 3 = 3(\exp\{\gamma(0)\} - 1)$$

La serie può essere analizzata in maniera più semplice dopo essere stata trasformata in un processo stazionario:

$$\begin{aligned} x_t &= \log y_t^2 = \log \sigma^2 + u_t + \log \varepsilon_t^2 \\ &= (\log \sigma^2 + E[\log \varepsilon_t^2]) + u_t + (\log \varepsilon_t^2 - E[\log \varepsilon_t^2]) \\ &= \mu + u_t + e_t \end{aligned}$$

dove $\{e_t\}$ è *i.i.d.* con media zero e varianza σ_e^2 ; se, per esempio, ε_t è distribuito come una normale standard, allora $\log \varepsilon_t^2$ si distribuisce come il logaritmo⁶ di una variabile casuale χ_1^2 , $E[\log \varepsilon_t^2] = -1.27$, $\sigma_e^2 = \frac{\pi^2}{2}$ (Wishart, 1947). Si può notare che il processo $\{x_t\}$ è la somma fra un processo gaussiano a memoria lunga e un termine d'errore non gaussiano fra loro indipendenti; la struttura di covarianza $\{x_t\}$ è

$$\begin{aligned} E[x_t] &= \mu \\ \gamma_x(0) &= \text{Var}(x_t) = \sigma_u^2 + \sigma_e^2 \\ \gamma_x(h) &= \text{Cov}(x_t, x_{t+h}) = \gamma(h) + \sigma_e^2 \cdot I_{\{h=0\}} \end{aligned}$$

dove $I_{\{h=0\}}$ assume valore uno se $h=0$ e zero altrimenti. Quindi, la funzione di autocorrelazione di x_t è data da (per ulteriori chiarimenti si veda Pérez e Ruiz, 2002)

$$\rho_x(k) = \frac{\gamma(k)}{\gamma(0)} = \sigma_\eta^2 \frac{\Gamma(k+d)\Gamma(1-2d)}{\Gamma(k+1-d)\Gamma(1-d)\Gamma(d)} \cdot \frac{1}{\sigma_u^2 + \sigma_e^2}$$

Più in generale, il termine $\{u_t\}$ può essere modellato da un ARFIMA(p,d,q); in questo caso si ha che

$$(1-B)^d \phi(B)u_t = \theta(B)\eta_t \quad \eta_t \text{ i.i.d. } N(0, \sigma_\eta^2)$$

La densità spettrale di $\{x_t\}$, chiamata $f_\beta(\cdot)$, quindi sarà data da:

$$f_\beta(\omega_j) = \frac{\sigma_\eta^2}{2\pi} \frac{|\theta(-i\omega_j)|^2}{|1 - \exp(-i\omega_j)|^{2d} |\phi(-i\omega_j)|^2} + \frac{\sigma_e^2}{2\pi}, \quad -\pi \leq \omega_j \leq \pi$$

⁶ In particolare, $E[\log \varepsilon_t^2] = \Psi\left(\frac{1}{2}\right) + \log 2$ dove $\Psi(\cdot)$ è la funzione digamma

dove $\omega_j = \frac{2\pi \cdot j}{T}$ è la j -esima frequenza di Fourier e il vettore dei parametri è dato da $\beta = (d, \sigma_\eta^2, \sigma_\varepsilon^2, \phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q)'$.

Per poter essere utilizzata nelle applicazioni pratiche, la densità spettrale $f_\beta(\cdot)$ verrà riscritta⁷ (si veda l'appendice 7.1) utilizzando seni e coseni e alcune proprietà dei numeri complessi.

Ad esempio, il più semplice tipologia che possa essere considerato è il modello LMSV(0,d,0) che ha forma del tipo

$$y_t = \sigma \exp\left\{\frac{(1-B)^{-d} \eta_t}{2}\right\} \varepsilon_t$$

mentre un'espressione più generale è data dal modello LMSV(1,d,1)

$$y_t = \sigma \exp\left\{\frac{(1-\phi B)^{-1}(1-B)^{-d}(1+\theta B)\eta_t}{2}\right\} \varepsilon_t$$

Un vantaggio dei modelli LMSV è che non ci sono restrizioni sui parametri, come invece accade nei modelli FIGARCH, e il più semplice modello frazionario per la varianza condizionale con $p, q = 0$ e $d \neq 0$ può essere considerato.

5.3 Stima del parametro di memoria lunga

Per ottenere stime del parametro di memoria lunga esistono diversi metodi, che variano tra approcci parametrici e non parametrici; qui si prendono in considerazione due alternative: il metodo del log-periodogramma (semiparametrico) e il metodo R/S o del rango riscaldato (non parametrico).

⁷ Le proprietà a cui si fa riferimento sono le seguenti:

- $|a + bi| = |a - bi| = \sqrt{(a + bi)(a - bi)} = \sqrt{a^2 + b^2}$
- $e^{i\omega} = \cos \omega + i \sin \omega$

Quest'ultimo risultato viene chiamato teorema di De Moivre ed ha un utilizzo particolarmente interessante. Infatti, si supponga di avere un numero complesso $z = |a + bi|$ giacente sul cerchio unitario e di volerlo esprimere in funzione dell'angolo ω al centro del cerchio; se z ha modulo uno si ha, per i risultati trigonometrici fondamentali $a = \cos \omega$ e $b = \sin \omega$, che $z = \cos \omega + i \sin \omega$. Quindi, in definitiva, $z = e^{i\omega} = e^{-i\omega}$

L'esistenza di memoria lunga nella volatilità delle serie storiche finanziarie verrà analizzata prendendo in considerazione due sue tradizionali misure ovvero i quadrati della serie e il logaritmo dei quadrati della serie.

5.3.1 Il metodo R/S o del rango riscaldato

Questo tipo di procedura, chiamata test del rango riscaldato o del rango sulle deviazioni standard, fu introdotta da Hurst (1951) e sviluppata successivamente da altri autori.

Sia $\{y_1, \dots, y_N\}$ la serie storica in cui si vuole testare la presenza di memoria lunga. La statistica R/S è il rango delle somme parziali delle deviazioni dei valori osservati dalla loro media, riscaldato con le deviazioni standard. Per ogni $n > 0$, il *rango aggiustato* R è dato da:

$$R(n) = \max_{1 \leq h \leq n} \sum_{t=1}^h (y_t - \bar{y}) - \min_{1 \leq h \leq n} \sum_{t=1}^h (y_t - \bar{y})$$

mentre il *fattore di normalizzazione* S è dato da

$$S(n) = \frac{1}{\sigma_n}$$

dove $\bar{y} = \frac{1}{N} \sum_{t=1}^N y_t$ e $\sigma_n = \sqrt{\frac{1}{N} \sum_{t=1}^N (y_t - \bar{y})^2}$ ed n può essere considerato come un

ritardo. Per tenere conto di possibili dipendenze di breve periodo, Lo (1991) ha suggerito di considerare una versione modificata in cui fattore di normalizzazione viene definito come la radice quadrata di uno stimatore consistente della varianza campionaria

$$S(n, q) = \frac{1}{\hat{\sigma}_n(q)} \quad \text{dove} \quad \hat{\sigma}_n(q) = \sigma_n^2 + 2 \cdot \sum_{j=1}^q \omega_j(q) \cdot \hat{\gamma}(q)$$

dove $\hat{\gamma}(q)$ è l'usuale stimatore della funzione di covarianza e

$$\omega_j(q) = 1 - \frac{j}{q+1} \quad q < n$$

e q rappresenta il massimo ritardo di cui si vuole tenere conto. In questo modo la deviazione standard campionaria viene sostituita dalla radice quadrata della stima di varianza di lungo periodo mediante lo stimatore di Newey-West con parametro

di lisciamiento q ; in tal modo si tengono in considerazione anche le autocovarianze diverse da zero in un processo a memoria corta.

Per quanto riguarda la scelta di q , non c'è un valore ottimale unanimemente riconosciuto nella letteratura; comunque, Andrews (1991) suggerisce di utilizzare

$$q = \left(\frac{3n}{2}\right)^{\frac{1}{3}} \cdot \left(\frac{2\hat{\rho}}{1-\hat{\rho}^2}\right)^{\frac{2}{3}}$$

Chiariti i ruoli di R e di S , la statistica R/S viene definita come

$$Q(n, q) = \frac{R(n)}{S(n, q)}$$

Il coefficiente di Hurst H viene stimato come

$$\hat{H}(n, q) = \frac{\log Q(n, q)}{\log n}$$

Se è presente solo memoria corta allora $\hat{H}(n, q)$ converge a $1/2$; se invece vi è la presenza di memoria lunga allora $\hat{H}(n, q)$ converge ad un valore maggiore di $1/2$ (Mandelbrot e Taqqu, 1979). Se il processo soddisfa una serie di condizioni di regolarità, inclusa l'esistenza di momenti di ordine superiore a $4 + \delta$, con $\delta > 0$, Lo (1991) ha mostrato che, sotto l'ipotesi nulla di memoria corta, la statistica

$$Q_n^* = \frac{1}{\sqrt{n}} Q(n, q) \rightarrow V$$

converge al rango di un ponte browniano. La funzione di distribuzione di questo rango, la si chiami F_V , è data da

$$F_V(v) = \sum_{k=-\infty}^{\infty} (1 - 4v^2 k^2) e^{-2v^2 k^2}$$

Una delle caratteristiche più importanti di questa statistica è la sua robustezza rispetto a cambiamenti di scala e a deviazioni anche estreme dalla normalità (come, per esempio, varianza infinita).

Nella pratica il metodo R/S può essere sostanzialmente classificato come una procedura grafica che si articola nel seguente modo. Si suddivide la serie in k sottoserie ognuna di numerosità $m = \frac{N}{k}$ e si calcolano le statistiche $Q_i(n)$ ai punti

iniziali $t_i = i \cdot \frac{N}{k} + 1$ per $i = 0, 1, \dots, k-1$ e $n \leq N - t_i$. Si ottengono così molti valori

campionari di $Q(n)$ per un dato n . In particolare, se $n < \frac{N}{k}$ si ottengono k diverse

stime di $Q(n)$ (una per ogni gruppo), se $n \geq N - \frac{N}{k}$ si ottiene una singola stima di $Q(n)$. Detto ciò si consideri il fatto che, per processi FGN (Fractional Gaussian Noise) e ARFIMA(p,d,q), Mandelbrot (1975) ha mostrato la seguente relazione asintotica

$$E(Q_n) \sim C \cdot n^H \quad \text{per } n \rightarrow \infty$$

dove H è il coefficiente di Hurst (in letteratura è ormai consolidata la relazione $H = d + 0.5$). Dunque

$$\log[E(Q_n)] \sim \log C + H \log n$$

Infine, per ottenere una stima del coefficiente H , e quindi di d , è sufficiente considerare il grafico di $\log[E(Q_n)]$ contro $\log n$ (questo grafico viene chiamato *pox-plot* per la statistica R/S), adattare una retta ai punti del grafico e considerare la stima del coefficiente angolare. Se i dati non presentano alcuna dipendenza si ha $H = 0.5$ ($d = 0$), mentre valori di $H > 0.5$ indicano la presenza di memoria lunga. Come ultima osservazione si deve avere l'accortezza di utilizzare solo ritardi n sufficientemente grandi (per la relazione asintotica), ma non troppo (si ottengono poche stime di $Q(n)$).

5.3.2 Il metodo del log-periodogramma

Una stima asintoticamente non distorta della funzione di densità spettrale è data dal periodogramma

$$I(\omega) = \frac{1}{2\pi n} \left| \sum_{t=1}^n y_t e^{-i\omega t} \right|^2$$

e per essa si ha

$$\lim_{\omega \rightarrow 0} E[I(\omega)] = f(\omega) \cong C|\omega|^{1-2H} \quad (5.4)$$

dove H rappresenta il coefficiente di Hurst. Geweke e Porter-Hudack (1983) hanno introdotto il metodo del log-periodogramma (o regressione ai minimi quadrati nel dominio delle frequenze) in seguito sviluppato e migliorato da Robinson (1995).

Tale metodo consiste nel:

1. calcolare il periodogramma alle frequenze di Fourier

$$\omega_{j,N} = \frac{2\pi j}{N}, \quad j=1,2,\dots,[(N-1)/2]$$

con $[.]$ = funzione parte intera

2. applicare i minimi quadrati ordinari all'equazione

$$\log(I(\omega_{j,N})) = \alpha + \beta \log(\omega_{j,N}) + e_j$$

dove $e_j \sim i.i.d.$ a media nulla e $\beta = -2d$

3. ricavare \hat{d} (o \hat{H}) a partire da $\hat{\beta}$.

E' opportuno evidenziare alcuni aspetti critici rilevanti per questa metodologia di stima. La relazione (5.4) vale asintoticamente, cioè solo per basse frequenze (ovvero nelle vicinanze dello zero); stimare d utilizzando tutte le ordinate del periodogramma in $(-\pi, \pi)$ porta quindi a stimatori fortemente distorti. Geweke e Porter-Hudack suggeriscono di utilizzare $m = g(N)$ frequenze con $g(\cdot)$ tale che

$$\lim_{N \rightarrow \infty} g(N) = \infty \quad \lim_{N \rightarrow \infty} \frac{g(N)}{N} = 0$$

Nella pratica essi suggeriscono di utilizzare $m = \sqrt{N}$. In tal caso \hat{d} è una stima consistente e asintoticamente distribuita come una variabile casuale Normale, ma con un tasso di convergenza più lento di $N^{-1/2}$, essendo

$$\sqrt{m} \cdot (\hat{d} - d) \rightarrow \frac{\pi}{24} \cdot N(0,1)$$

Quindi, rispetto ad uno stimatore parametrico ottimale, l'efficienza di tale stimatore è zero. Inoltre, Agiaklogou et al. (1993), mettono in luce che, quando il processo generatore dei dati è un ARFIMA(p,d,q) con p e/o $q \neq 0$, il metodo del log-periodogramma comporta una notevole distorsione nella stima del parametro di lungo periodo; tale distorsione dipende dai valori dei parametri che descrivono il comportamento a breve del processo ed è più evidente per grandi valori positivi del parametro AR o per grandi valori negativi del parametro MA.

Nonostante tali problemi, tale tecnica rimane ancora oggi il solo metodo non parametrico per stimare d per cui sia anche possibile derivare la distribuzione asintotica dello stimatore (e quindi test e intervalli di fiducia per d).

5.4 Stima di un modello LMSV

Negli ultimi anni sono stati proposti diversi metodi per stimare i modelli LMSV. Hsu e Breidt (1997), in un contesto bayesiano, hanno ottenuto la distribuzione a posteriori dei parametri e le stime delle volatilità utilizzando un algoritmo basato sul Gibbs sampling. Altri due algoritmi basati su un approccio bayesiano sono stati proposti da So (1999) e Chan e Petris(2002); queste tre procedure sono computazionalmente intensive e difficilmente generalizzabili in modelli più complicati come, per esempio, quando l'equazione delle log-volatilità ha una componente ARMA.

Deo e Hurvich (2001) consigliano di utilizzare uno stimatore semiparametrico di d basato su uno stimatore GPH di Geweke e Porter-Hudak (1983) per i modelli ARFIMA e hanno ottenuto la sua distorsione e la sua varianza asintotica. La normalità asintotica dello stimatore GPH si basa sulla gaussianità del logaritmo delle volatilità: Bollerslev e Wrigth (2000) hanno mostrato che le proprietà dello stimatore GPH del parametro d sono in stretta relazione con la frequenza dei dati campionari.

Wrigth (1999) ha proposto uno stimatore GMM ed ha dimostrato la sua consistenza e normalità asintotica quando $d < 1/4$; comunque, numerosi studi hanno messo in luce che usualmente le stime di d per dati ad alta frequenza sui rendimenti sono comprese fra 0.3 e 0.47 (Andersen et al. 2001). Quindi, il caso $d > 1/4$ è più interessante da un punto di vista empirico.

Infine, Harvey(1998) e Breidt et al. (1998) hanno proposto di stimare i parametri di un modello LMSV tramite la quasi massima verosimiglianza nel dominio delle frequenze (FDQML). Questa procedura sarà oggetto dell'interesse del prossimo paragrafo e verrà utilizzata nelle successive analisi empiriche.

Nella tabella della pagina seguente viene fornito un riassunto delle principali caratteristiche degli stimatori presentati in questo paragrafo per un generico modello LMSV.

Metodo	Autori	Proprietà Asintotiche	Prove su campioni finiti	Vantaggi e limiti
QML	Harvey (1998) Breidt et al. (1998)	Consistenza	Proprietà povere quando $d \sim 0.5$ e/o $\sigma^2_\eta \sim 0$	Stima della volatilità
MCMC	Hsu e Breidt (1997) So (1999) Chan e Petris (2000)	Basato su simulazioni	Non disponibile	Computazionalmente intensivo Difficilmente generalizzabile Stima della volatilità
Semi-parametrico	Deo e Hurvich(2001)	Risultati solo per la stima di d Asintoticamente normale sotto certe condizioni	Non esaustivo Basato su disegno Monte Carlo	
GMM	Wright (1999)	Consistente e asintoticamente normale per $d < 1/4$	Simile a QML	Comportamento asintotico sconosciuto per $d > 1/4$

5.4.1 Stima di FDQML

L'esatta verosimiglianza del vettore dei parametri β dato il campione (y_1, \dots, y_T) implica un integrale n -dimensionale e, di conseguenza, è molto difficile da valutare. Una semplice alternativa alla verosimiglianza gaussiana nel dominio temporale è la massimizzazione della sua rappresentazione nel dominio delle frequenze. Detto ciò, la procedura di quasi massima verosimiglianza sfrutta il fatto, accennato in precedenza, che la serie trasformata⁸ $x_t = \log y_t^2$ può essere espressa come la somma fra un processo gaussiano a memoria lunga e un errore a media nulla non gaussiano. Gli stimatori QML si ottengono trattando la serie x_t come una serie gaussiana e massimizzando la verosimiglianza di x_t . Anche se x_t non è un processo gaussiano, Deo (1995) ha mostrato che gli stimatori risultanti sono \sqrt{T} consistenti e asintoticamente normali. Nella pratica, è più semplice minimizzare l'approssimazione nel dominio delle frequenze della log-verosimiglianza negativa gaussiana nel dominio temporale, chiamata approssimazione di Whittle, che è data da

$$L_n(\beta) = 2\pi n^{-1} \sum_{k=1}^{\lfloor n/2 \rfloor} \left\{ \log f(\omega_k) + \frac{I_n(\omega_k)}{f_\beta(\omega_k)} \right\} \quad (5.5)$$

dove $[.]$ denota la funzione parte intera, $f_\beta(\omega_k)$ è la densità spettrale introdotta precedentemente, $\omega_{k,N} = \frac{2\pi k}{N}$ è la k -esima frequenza di Fourier e

$$I_n(\omega_k) = \frac{1}{2\pi n} \left(\sum_{t=1}^n x_t \cos \omega_k t \right)^2 + \frac{1}{2\pi n} \left(\sum_{t=1}^n x_t \sin \omega_k t \right)^2$$

⁸ La scelta di utilizzare $\log y_t^2$ invece che y_t sembra essere giustificata da evidenze empiriche. Deo e Hurvich (2003) hanno proposto una spiegazione in termini di effetti dei valori outlier per il fatto che il modulo e i quadrati dei rendimenti hanno autocorrelazioni campionarie più piccole rispetto al logaritmo dei quadrati dei rendimenti; presumibilmente un fenomeno simile sussiste per il grado di persistenza implicato dai periodogrammi. In verità, Wright (2002) ha mostrato che, utilizzando simulazioni basate su modelli LMSV e di tipo ARCH, stimatori parametrici di d basati sul periodogramma hanno una distorsione negativa minore se vengono utilizzati i logaritmi dei quadrati dei rendimenti anziché i moduli o i quadrati dei rendimenti.

è la k -esima ordinata del periodogramma normalizzato. Breidt et al. (1998) hanno dimostrato che lo stimatore QML è consistente⁹; inoltre Deo (2002) ha provato la normalità asintotica dello stimatore QML ottenuto massimizzando la verosimiglianza nel dominio temporale e ed ha congetturato che gli stessi risultati possono essere estesi allo stimatore di Whittle. Con riferimento ad un campione finito, Breidt et al. hanno concluso che le proprietà dello stimatore sono soddisfacenti; comunque, Pérez e Ruiz (2001) hanno esteso le loro analisi ed hanno concluso che, quando i parametri sono vicini alla non-stazionarietà ($d \approx 0.5$) e/o alla omoschedasticità ($\sigma_\eta^2 \approx 0$), le proprietà dello stimatore QML sono povere, tali che sono necessari campioni di grandissime dimensioni per ottenere buoni risultati inferenziali.

Ci sono alcuni aspetti critici inerenti a questa procedura di stima che devono essere evidenziati:

- Tutti i risultati precedenti valgono per $d \in (0,0.5)$ o per $H \in (0.5,1)$ anche se risultati di simulazione confermano l'affidabilità delle stime anche per valori di $H \in (0,1)$ e $d \in (-0.5,0.5)$. L'estensione al caso in cui H o d non appartengono a questi intervalli, ovvero quando si è nel caso non stazionario, è un problema ancora aperto.
- I precedenti risultati dovrebbero essere estesi al caso in cui il processo osservato non è gaussiano.
- La condizione di identificabilità del teorema nella nota 9 viene verificata se σ_e^2 è conosciuto in base ad una distribuzione assunta per ε_t : ad esempio, se $\varepsilon_t \sim N(0,1)$ allora $\sigma_e^2 = \frac{\pi^2}{2}$. Se σ_e^2 non è conosciuto, il modello è identificato solo se la componente ARFIMA non è un white noise, ovvero se $\phi_p \neq 0$ per qualche p , $\theta_q \neq 0$ per qualche q , oppure $d \neq 0$.

⁹ In merito a questo, Breidt et al. (1998) hanno dimostrato il seguente teorema:

Si assuma che il vettore dei parametri $\beta = (d, \sigma_\eta^2, \sigma_e^2, \phi_1, \dots, \phi_p, \theta_1, \dots, \theta_p)'$ sia un elemento dello spazio parametrico compatto Θ e si assuma che se $f_{\beta_1}(\omega) \equiv f_{\beta_2}(\omega)$ per ogni $\omega \in [0, \pi]$ allora $\beta_1 = \beta_2$, dove $f_\beta(\cdot)$ è definito come in precedenza. Si chiami $\hat{\beta}_n$ il vettore dei parametri che minimizza (5.5) su Θ e sia β_0 il vettore dei veri valori dei parametri; allora $\hat{\beta}_n \rightarrow \beta_0$ quasi certamente.

Nel prossimo paragrafo verrà presentato un metodo, introdotto da Deo et al. (2005), in cui si prova ad ovviare al problema della non gaussianità di x_t .

Infine, bisogna ricordare che tutti i risultati precedenti sono definiti a livello asintotico mentre la maggior parte delle volte si è interessati alle proprietà degli stimatori nel caso di campioni limitati.

5.4.2 FDQML “Accresciuto”

Premesso che per questo stimatore non esiste alcuna teoria asintotica, Deo et al. (2005) suggeriscono di considerare il fatto che, dato che $x_t = \log y_t^2$ non è gaussiano, delle stime migliori possono essere ottenute applicando il metodo di Whittle a qualche trasformazione di $|y_t|$ che sia più vicina alla gaussianità del logaritmo dei quadrati dei rendimenti. Una potenziale trasformazione è $|y_t|^c$ scegliendo c in modo che, per esempio, l'indice di asimmetria teorico di $|y_t|^c$ sia 0, in modo che la sua distribuzione sia più vicina alla gaussianità di x_t . Per poter applicare la procedura FDQML alla serie trasformata $|y_t|^c$, si devono prima affrontare due questioni:

- i. La scelta dell'esponente c .
- ii. Il calcolo della densità spettrale di $|y_t|^c$.

Per quanto riguarda il primo punto, Deo et al. (2005) suggeriscono una procedura a due stadi. Nella prima fase si applica la procedura FDQML alla serie x_t in modo da ottenere una prima stima dei parametri del modello LMSV; usando queste stime, si può calcolare l'indice di asimmetria teorico di $|y_t|^c$ per ogni valore di c nella maniera seguente. Per prima cosa si può notare che l'indice di asimmetria $E\left[|y_t|^c - \mu_c\right]^3$, dove $\mu_c = E\left[|y_t|^c\right]$, può essere riscritto, a meno di una costante moltiplicativa, come

$$E\left[|y_t|^c - \mu_c\right]^3 = E|y_t|^{3c} - \mu_c^3 - 3\mu_c E|y_t|^{2c} + 3\mu_c^2 E|y_t|^c$$

Utilizzando la funzione generatrice dei momenti di una variabile casuale Normale¹⁰ e ponendo $\sigma_u^2 = \text{Var}(u_t)$, si ha, per ogni numero intero positivo k , che

$$E|y_t|^{kc} = E|\varepsilon_t|^{kc} \exp\left\{\sigma_u^2 k^2 \frac{c^2}{8}\right\} \sigma^{kc}$$

Seguendo Harvey (1998), quando ε_t si distribuisce come una Normale standard si ha che

$$E|\varepsilon_t|^c = \frac{2^{c/2} \Gamma((c+1)/2)}{\Gamma(1/2)}$$

Combinando queste tre ultime espressioni si può ottenere una espressione esplicita per l'indice di asimmetria di $|y_t|^c$: a questo punto si può minimizzare l'indice di asimmetria in funzione di c . Deo et al. (2005) inoltre hanno minimizzato una somma pesata, con pesi equivalenti, dei quadrati dell'indice di asimmetria e dell'eccesso di curtosi ottenendo però dei valori di c molto simili al caso in cui si considera solo l'indice di asimmetria.

Dopo aver trovato un opportuno valore di c , la stima di Whittle richiede la densità spettrale della trasformazione $|y_t|^c$; sfortunatamente non esiste una semplice formula per la densità spettrale di $|y_t|^c$ come quella descritta per x_t . Comunque, come detto in precedenza, si sa che per ogni serie storica stazionaria con covarianza data da $c_j(\beta)$ al ritardo j , la densità spettrale $f_\beta(\omega)$ alla frequenza ω è data da

$$f_\beta(\omega) = \frac{1}{2\pi} \sum_{|j|<\infty} c_j(\beta) \exp\{ij\omega\}$$

La funzione di covarianza $c_j(\beta)$ di $|y_t|^c$ può essere ottenuta notando che, per $a, b > 0$, si ha

$$E|y_t|^a = \sigma^a E|\varepsilon_t|^a \exp\left\{a^2 \frac{\sigma_u^2}{8}\right\}$$

e

¹⁰ Se z si distribuisce come una variabile casuale Normale, allora la sua funzione generatrice dei momenti è data da

$$g(t) = \exp\left\{\mu t + \sigma^2 \frac{t^2}{2}\right\} \quad \text{dove } E(z) = \mu \text{ e } \text{Var}(z) = \sigma^2$$

$$E(|y_t|^a |y_{t-j}|^b) = \sigma^{a+b} E|\varepsilon_t|^a E|\varepsilon_t|^b \exp\left\{(a^2 + b^2) \frac{\sigma_u^2}{8} + ab \frac{\gamma_u(j)}{4}\right\}$$

dove $\gamma_u(j) = Cov(u_t, u_{t-j})$. A questo punto, per calcolare la funzione di covarianza $\gamma_u(j)$ di un processo ARFIMA, Deo et al. (2005) propongono di utilizzare un semplice algoritmo proposto da Bertelli e Caporin (2002), chiamato “splitting method” (questo metodo verrà trattato nell’appendice 7.3); combinando questo algoritmo con l’espressione di $E|\varepsilon_t|^c$ data in precedenza si può facilmente calcolare $c_j(\beta)$.

Infine, per applicare il metodo di Whittle si deve troncare in un qualche punto l’espressione di $f_\beta(\omega)$ per poter rendere la procedura computazionalmente fattibile; comunque, una scelta sfortunata del punto di troncamento può portare ad una densità spettrale negativa, cosa che non è certamente auspicabile. Quindi, per ovviare a questo problema, è stato proposto di pesare la funzione di covarianza $c_j(\beta)$ con un’appropriata funzione regolare w_j prima del troncamento, che assicura una densità spettrale positiva data da

$$f_{\beta,M}(\omega) = \frac{1}{2\pi} \sum_{|j|<M} w_j c_j(\beta) \exp\{ij\omega\}$$

dove M è il punto di troncamento. Deo et al. (2005) nel loro lavoro propongono di utilizzare come sequenza dei pesi la nota “Bartlett window”, che assicura di ottenere densità positive e che è data da $w_j = 1 - \frac{|j|}{M}$; si può mostrare che, per ogni valore $\omega > 0$, $f_{\beta,M}(\omega) \rightarrow f_\beta(\omega)$ per $M \rightarrow \infty$.

Per testare la bontà di questo stimatore, Deo et al. (2005) hanno generato un campione di 30000 dati da un modello LMSV con la seguente struttura:

$$y_t = \sigma \exp\left\{\frac{h_t}{2}\right\} \varepsilon_t \quad \text{con } \sigma > 0$$

e h_t segue un modello ARFIMA(1,d,0) dato da

$$(1 - \alpha B)(1 - B)^d h_t = \eta_t$$

indipendente da ε_t ; questa simulazione si basa su 300 repliche del suddetto modello.

Gli autori confrontano le stime ottenute utilizzando questa procedura con quella tradizionale FDQML e, riassumendo, concludono che:

- i. Lo stimatore FDQML “accresciuto” offre stime sensibilmente migliori per quanto riguarda il parametro autoregressivo α della componente ARFIMA e la varianza dell’errore (σ_η^2) nel processo h_t .
- ii. Per il parametro di differenziazione d , il nuovo metodo offre stime molto simili al metodo classico; per il parametro σ_e^2 invece, lo stimatore FDQML risulta migliore.
- iii. Per quanto riguarda la variabilità delle stime, lo stimatore FDQML “accresciuto” apporta miglioramenti per i parametri α e d , mentre la variabilità risulta leggermente superiore per le stime di σ_η^2 e σ_e^2 .

Capitolo 6

6. Analisi empiriche

In questo capitolo verranno presentate le analisi inerenti due serie storiche finanziarie, Autostrade e Mibtel, concentrando l'attenzione sulla varianza: in particolare si adatteranno ai dati i modelli presentati nei capitoli precedenti, ovvero i modelli FIGARCH e LMSV.

Per quanto riguarda gli applicativi, i modelli ARMA e GARCH sono stati stimati tramite il programma Econometric Views (EViews), i modelli FIGARCH tramite il programma S-Plus e i modelli a volatilità stocastica sono stati stimati con delle apposite routine compilate in linguaggio R.

6.1 Analisi dei dati Autostrade

I dati che sono stati utilizzati in questa sezione si riferiscono all'andamento dei prezzi del titolo Autostrade S.p.A. tra il 15 marzo 1996 e il 23 maggio 2005 per un totale di 2184 osservazioni. I dati mancanti per le festività non sono stati riempiti da alcun valore, per questo nei grafici delle varie analisi non verrà mai riportato il tempo ma l'indice dell'osservazione.

Poiché nella parte inerente i modelli a volatilità stocastica si utilizzerà una trasformazione dei dati, ovvero il logaritmo dei quadrati di rendimenti, si è resa necessaria una modifica dei dati originari. Infatti, per non avere il problema di calcolare il logaritmo di 0, i rendimenti nulli sono stati sostituiti con la media fra il rendimento precedente e quello seguente.

Infine, i dati antecedenti al 1 gennaio 1999 sono stati forniti in lire: per essere rapportati in euro sono stati divisi per 1936,27 ovvero la parità fra Lira e Euro. Tutti i dati sono stati reperiti su www.finance.yahoo.com.

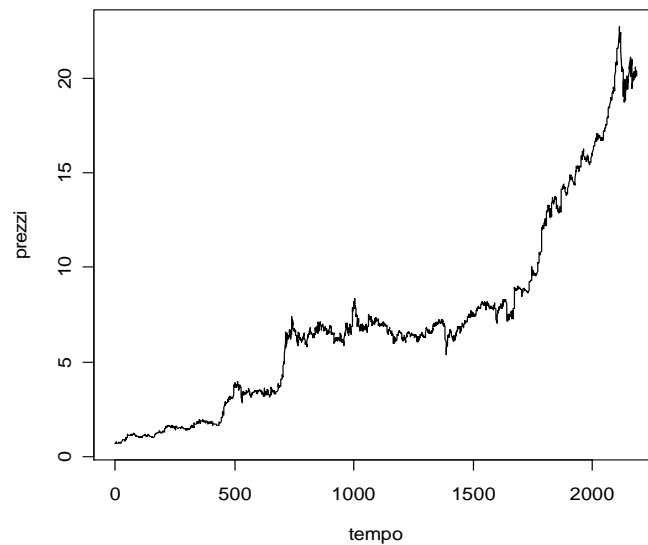


Figura 6. 1

Prezzi del titolo Autostrade S.p.A

Nella figura 6.1 si può notare l'andamento crescente che ha avuto il titolo Autostrade in questi anni, soprattutto dall'estate 2002 (che corrisponde circa all'intervallo fra 1700-1800) fino all'inizio 2005, dove non ha avuto, per il momento, dei risultati brillanti come il recente passato.

Nelle figure 6.2 e 6.3 vengono rappresentati gli andamenti dei rendimenti (calcolati come differenze dei logaritmi dei prezzi) e i rendimenti al quadrato.

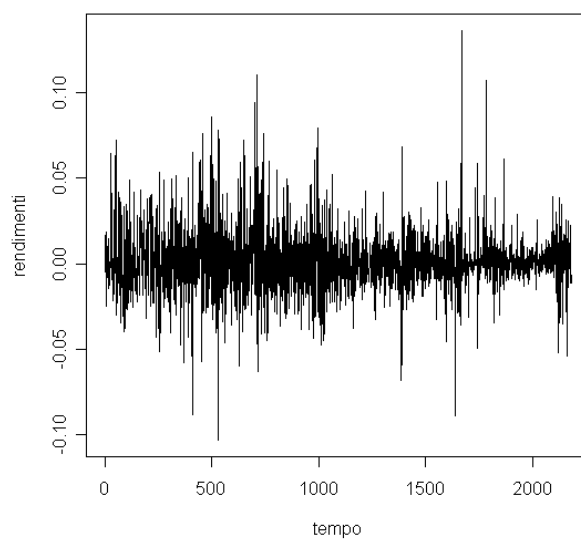


Figura 6. 2

Rendimenti giornalieri del titolo Autostrade S.p.A

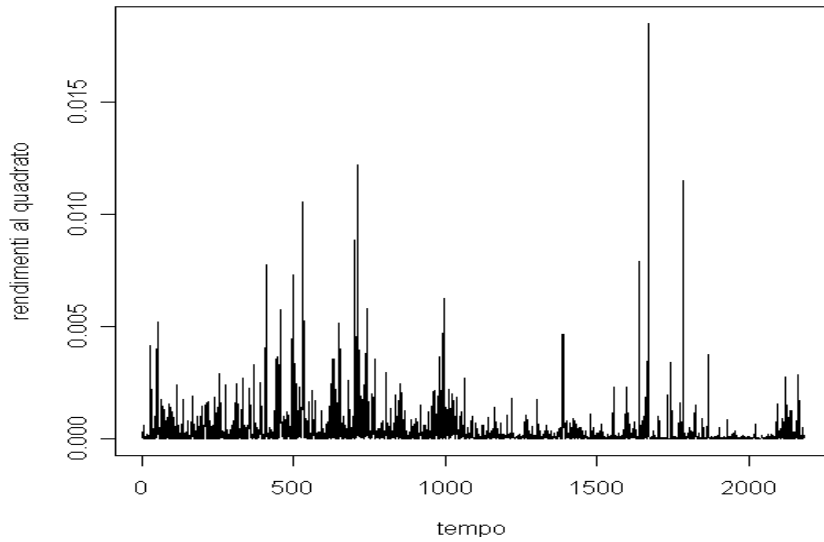


Figura 6. 3

Rendimenti giornalieri al quadrato del titolo Autostrade S.p.A

A questo punto è necessario analizzare l'autocorrelazione e l'autocorrelazione parziale dei rendimenti per cercare l'eventuale presenza di un qualche tipo di dinamica sulla media.

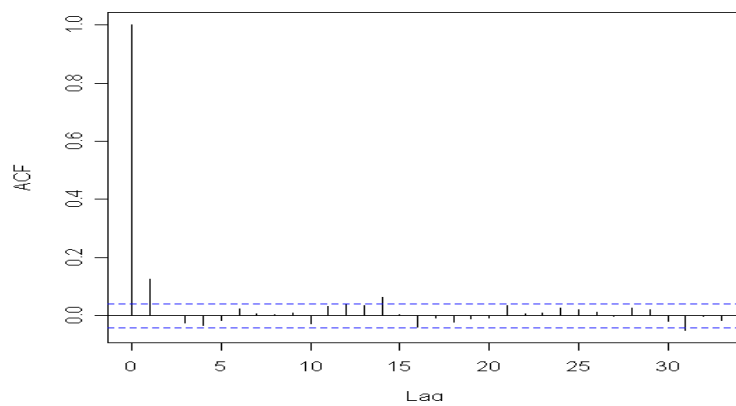


Figura 6. 4

Autocorrelazione dei rendimenti

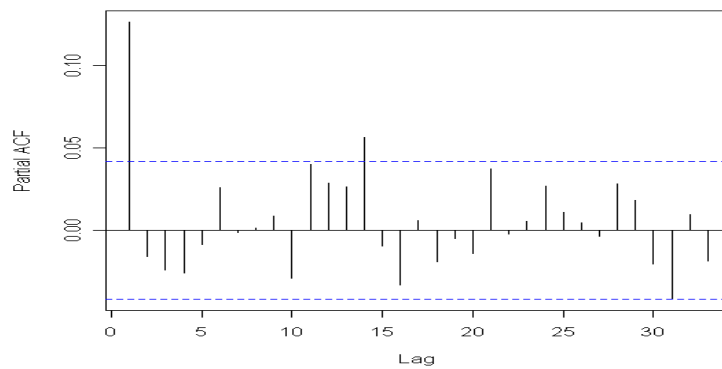


Figura 6. 5

Autocorrelazione parziale dei rendimenti

Questi due grafici sembrano indicare la presenza di una qualche componente ARMA sulla media. Dopo alcuni tentativi, il modello migliore sembra essere un AR(1). Nella tabella 6.1 vengono presentati i risultati della regressione.

Variabile	Coefficiente	Standard error	Statistica t	P-value
μ	0.001613	0.000469	3.441393	0.0006
AR(1)	0.126545	0.021247	5.955844	0.0000
Statistica F			35.47207	
P-value			0.000000	
Log-verosimiglianza			5541.844	
AIC			-5.074857	
SC			-5.069644	

Tabella 6. 1

Stima AR(1) dei dati della figura 6.2

Si nota chiaramente la significatività sia del coefficiente AR(1) che della costante. Forte è anche la presenza di effetti ARCH (visibili già dai “grappoli” di volatilità presenti nelle figure 6.2 e 6.3), come testimonia la significatività della statistica F della tabella 6.2 e il grafico delle autocorrelazioni dei residui al quadrato della regressione.

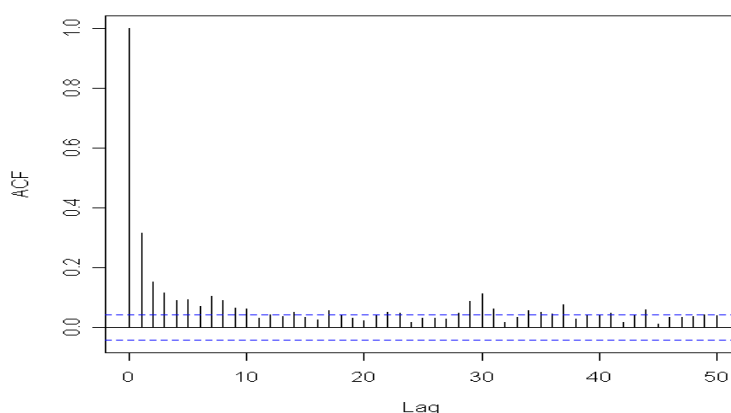


Figura 6. 6

Autocorrelazioni per i residui al quadrato del modello della tabella 6.1

Il test ARCH, difatti, da risultati piuttosto espliciti

Variabile	Coefficiente	Standard error	Statistica t	P-value
c	0.000250	0.0000195	12.84906	0.0000
$Res^2(-1)$	0.315089	0.020332	15.49752	0.0000
Statistica F			240.1794	
P-value			0.000000	
TR^2			216.5295	
P-value			0.000000	

Tabella 6. 2

Test ARCH-LM sui residui della tabella 6.1

Prima di procedere con l'analisi della varianza condizionata si può osservare dalla figura 6.7 che i residui presentano la caratteristica di non normalità che ci si attendeva: in particolare si rileva dell'asimmetria positiva (0.658247) e un marcato eccesso di curtosi (6.87139), tanto che la statistica di Jarque-Bera vale addirittura 1520.201; in questo contesto, quindi, è opportuno procedere nella stima dei modelli FIGARCH tramite l'uso della quasi massima verosimiglianza.

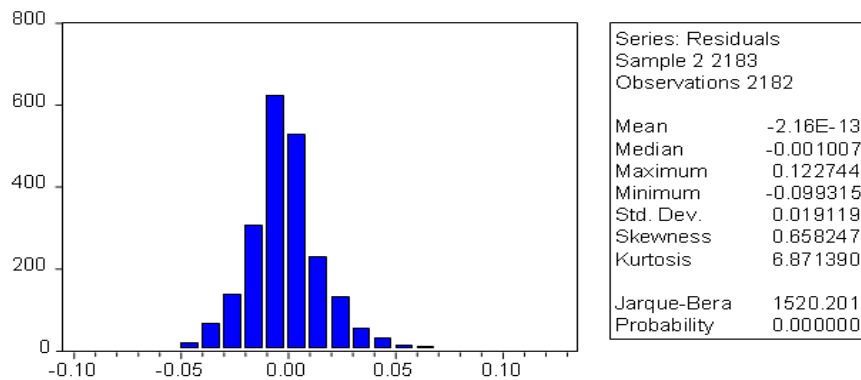


Figura 6.7

Statistiche descrittive e di normalità dei residui della tabella 6.1

A questo punto, dopo aver identificato il modello ottimale per la media, si può passare all'introduzione dell'integrazione frazionaria. L'utilità di una modellazione a memoria lunga è evidente se si esaminano le autocorrelazioni dei rendimenti al quadrato e del loro logaritmo, il modulo dei rendimenti e il periodogramma del logaritmo dei rendimenti al quadrato, che presenta un marcato picco a frequenza molto basse.

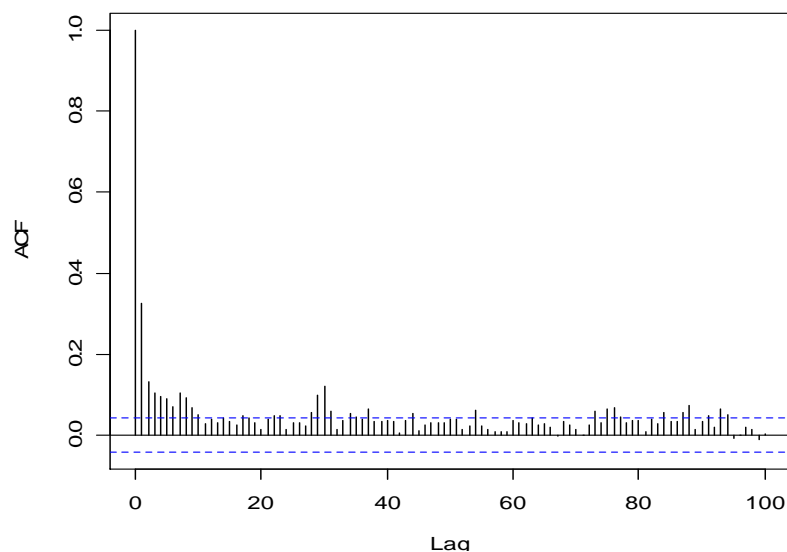


Figura 6.8

Autocorrelazioni dei rendimenti al quadrato

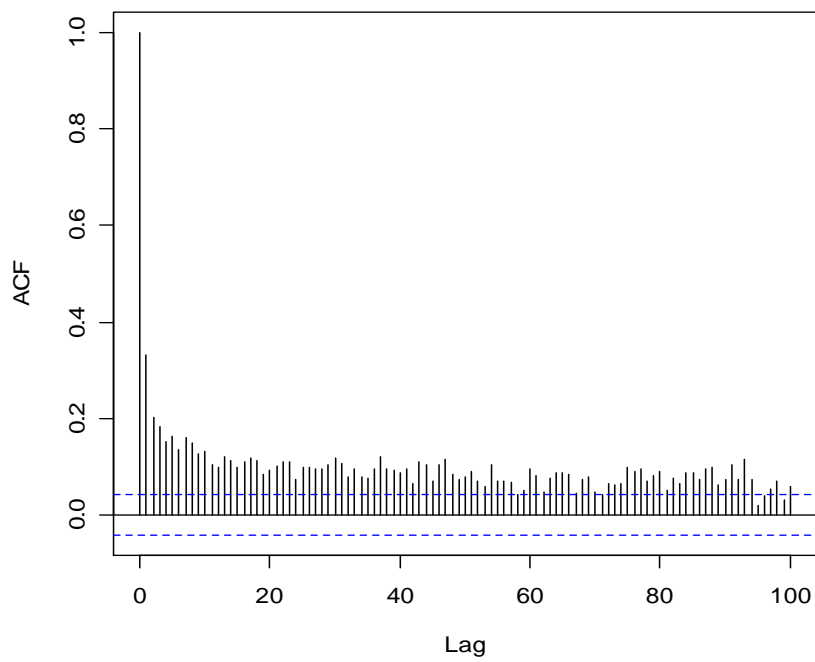


Figura 6. 9
Autocorrelazioni dei moduli dei rendimenti

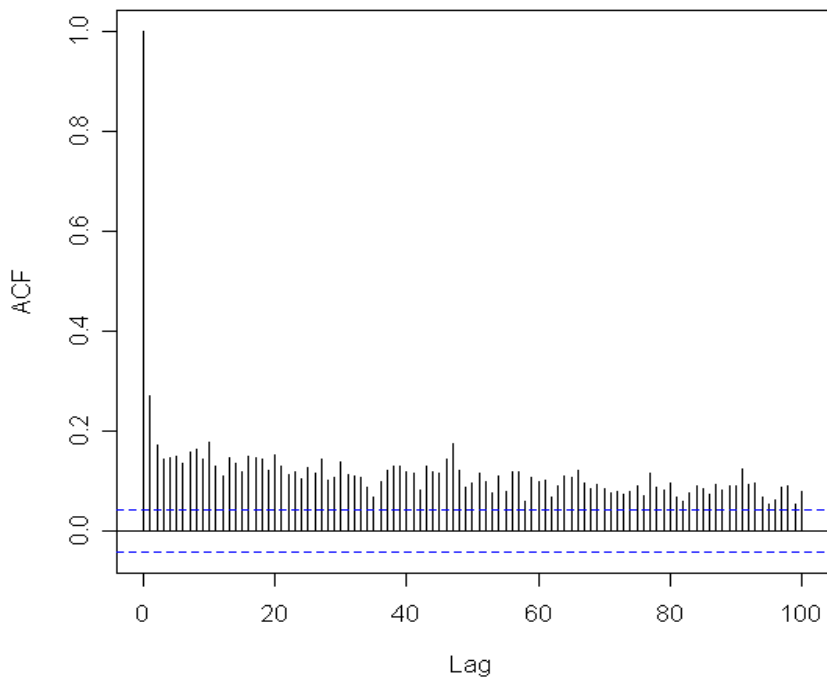


Figura 6. 10
Autocorrelazioni dei logaritmi dei rendimenti al quadrato

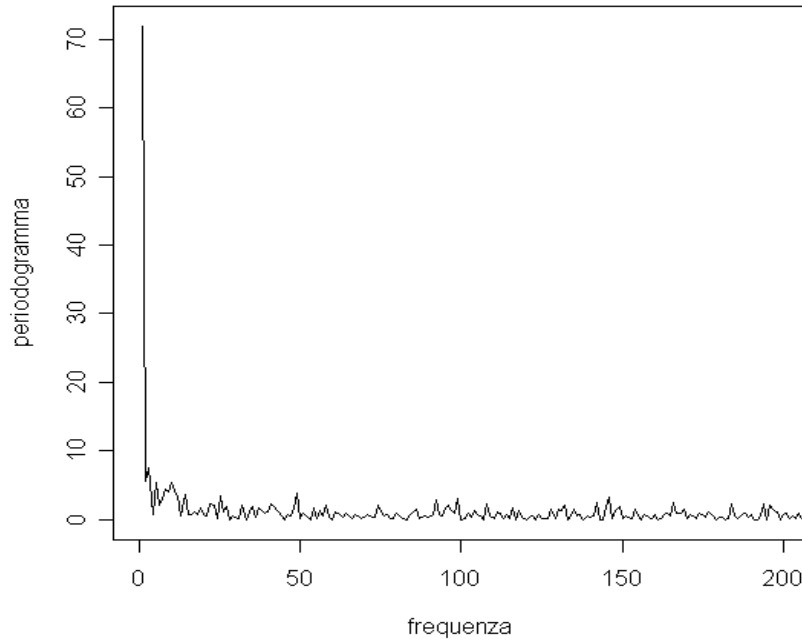


Figura 6.11

Periodogramma del logaritmo dei rendimenti al quadrato

Il primo modello che verrà stimato è un FIGARCH(1,d,0). I risultati sono riportati nella tabella 6.3

Variabile	Coefficiente	Standard error	Statistica t	P-value
μ	0.00117184	0.00030845	3.7992	0.0008
$AR(1)$	0.07423420	0.02747622	2.7018	0.0003
ω	0.00003218	0.0000167	1.9228	0.0027
$GARCH(1)$	-0.05023829	0.05382022	-0.9334	0.1753
d	0.30851129	0.04467157	6.9062	0.0000
<i>Log-verosimiglianza</i>			5801.783	
<i>AIC</i>			-5.310841	
<i>SC</i>			-5.297812	

Tabella 6.3

Stima QML FIGARCH (1,d,0) sui dati della figura 6.2

La log verosimiglianza incrementa rispetto al modello della tabella 6.1 ed entrambi i criteri informativi concordano nell'accettare la nuova tipologia del modello (infatti diminuiscono entrambi). Il coefficiente di integrazione frazionaria è decisamente significativo (e anche, ovviamente, statisticamente diverso da uno). Si osserva inoltre che il coefficiente GARCH(1) ha perso significatività. Alla luce di questo fatto, provo a stimare per i dato un modello FIGARCH(0,d,0).

Variabile	Coefficiente	Standard error	Statistica t	P-value
μ	0.00090184	0.00040845	2.2085	0.0136
$AR(1)$	0.0482546	0.02847622	1.69455	0.0445
ω	0.00008326	0.0000185	4.5005	0.0000
d	0.53511326	0.0546613	9.7896	0.0000
<i>Log-verosimiglianza</i>			5638.669	
<i>AIC</i>			-5.162317	
<i>SC</i>			-5.151894	

Tabella 6. 4

Stima QML FIGARCH (0,d,0) sui dati della figura 6.2

Come si può notare la perdita di verosimiglianza è netta ed entrambi i criteri informativi propendono nel rifiutare questo modello.

Come ultimo tentativo provo a stimare un modello FIGARCH(1,d,1):

Variabile	Coefficiente	Standard error	Statistica t	P-value
μ	0.00129927	0.00039544	3.286	0.0005
$AR(1)$	0.08991265	0.02489656	3.611	0.0002
ω	0.0001305	0.0000555	2.3526	0.0009
$ARCH(1)$	0.532536	0.004903	51.626	0.0000
$GARCH(1)$	0.07536	0.248582	0.3709	0.3554
d	0.263104	0.0000218	144562	0.0000
<i>Log-verosimiglianza</i>			5810.23	
<i>AIC</i>			-5.317664	
<i>SC</i>			-5.302029	

Tabella 6. 5

Stima QML FIGARCH (1,d,1) sui dati della figura 6.2

Questa nuova parametrizzazione apporta dei miglioramenti sia in termini di varianza sia in termini di criteri informativi, pertanto la tengo come quella definitiva. Nella figura 6.12 viene tracciato l'andamento della varianza condizionata secondo il modello della tabella 6.5:

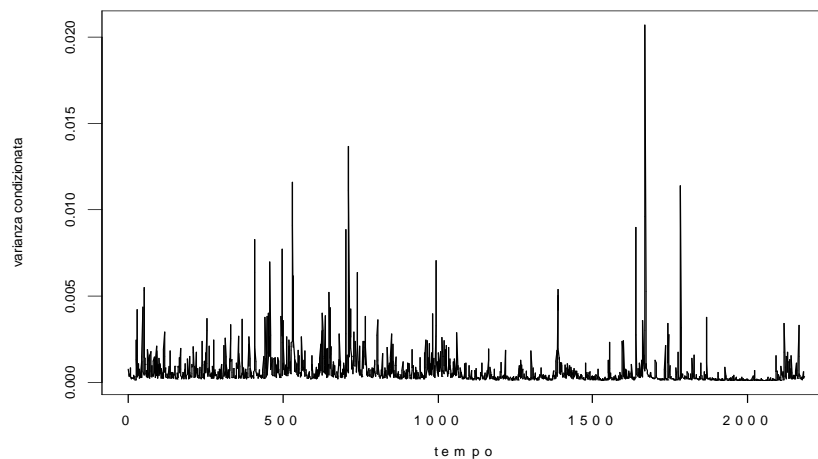


Figura 6. 12

Varianza condizionata secondo il modello della tabella 6.5

Nella figura 6.13 invece viene riportato il grafico dei residui standardizzati del modello che, in linea di massima, ha un andamento soddisfacente (in particolare si noti la scomparsa dei clustering della volatilità) se si esclude qualche valore anomalo.

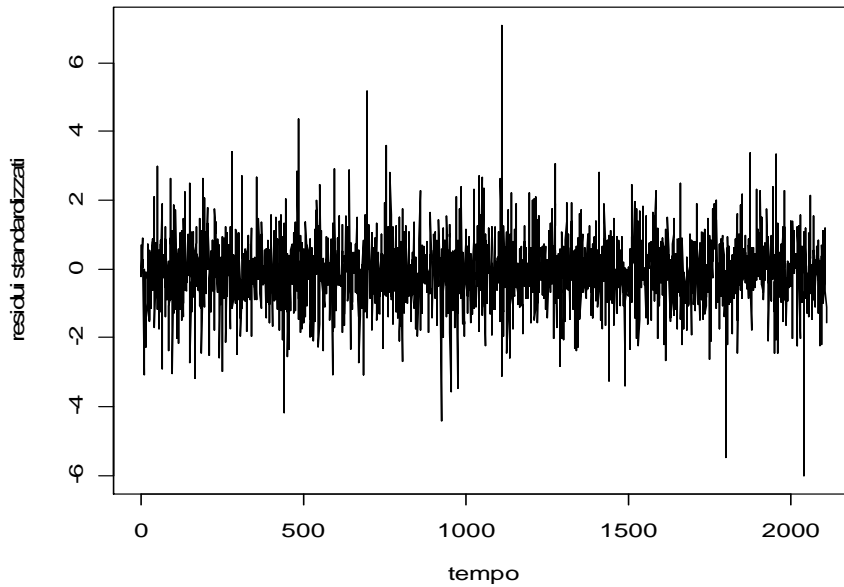


Figura 6.13

Residui standardizzati del modello della tabella 6.5

Nella figura sottostante vengono riportate le statistiche relative ai residui sopraccitati:

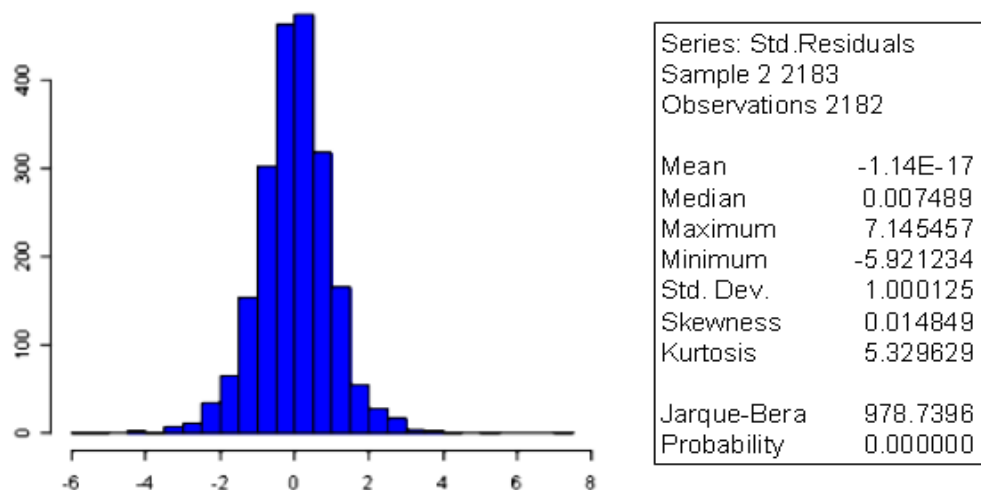


Figura 6.14

Statistiche descrittive e di normalità dei residui della tabella 6.5

Da questi grafici si possono estrarre alcune informazioni. Si osserva che non si riesce ad avvicinarsi alla normalità perché permane un'eccessiva curtosi dovuta alla presenza di troppi valori sulle code. Ciò indica chiaramente, come si era ipotizzato in partenza, che la distribuzione dei termini di disturbo non è normale, ed è quindi stato saggio procedere utilizzando la quasi-massima verosimiglianza.

6.1.1 Modelli a volatilità stocastica per Autostrade S.p.A

In questa sezione si forniranno le stime ottenute modellando la serie dei logaritmi dei rendimenti al quadrato ipotizzando diverse specificazioni del modello LMSV; la scelta del modello migliore verrà effettuata comparando il valore della verosimiglianza nel punto di massimo e i criteri informativi AIC e SC, che tengono conto della verosimiglianza e del numero dei parametri. In sintesi i modelli stimati saranno del tipo

$$x_t = \log y^2 = u_t + e_t$$

dove, come al solito, $\{e_t\}$ è i.i.d. con media 0 e varianza σ_e^2 indipendente da $\{u_t\}$ e $\{u_t\}$ è un processo ARFIMA(p,d,q),

$$(1 - B)^d \phi(B)u_t = \theta(B)\eta_t$$

dove $\{\eta_t\}$ è un processo i.i.d. con media 0 e varianza σ_η^2 .

Le stime ottenute sono riassunte nelle seguenti tabelle:

Modello	$\hat{\sigma}_\eta^2$	$\hat{\sigma}_e^2$	\hat{d}
<i>LMSV(0,d,0)</i>	0.9782168	2.8122984	0.3940289
<i>Log-verosimiglianza</i>			1.714121
<i>AIC</i>			0.001178084
<i>SC</i>			0.008995476

Tabella 6. 6

Stime FDQML LMSV(0,d,0) della serie $\log y^2$

Modello	$\hat{\sigma}_\eta^2$	$\hat{\sigma}_e^2$	\hat{d}	θ
<i>LMSV(0,d,1)</i>	0.2377645	3.2254380	0.3928131	-0.9551657
<i>Log-verosimiglianza</i>				1.707830
<i>AIC</i>				0.002100018
<i>SC</i>				0.01252321

Tabella 6. 7

Stime FDQML LMSV(0,d,1) della serie $\log y^2$

Modello	$\hat{\sigma}_\eta^2$	$\hat{\sigma}_e^2$	\hat{d}	ϕ
<i>LMSV(1,d,0)</i>	0.7293759	2.9994503	0.3889478	0.1349167
<i>Log-verosimiglianza</i>			1.712067	
<i>AIC</i>			0.002096137	
<i>SC</i>			0.01251933	

Tabella 6. 8

Stime FDQML LMSV(1,d,0) della serie $\log y^2$

Modello	$\hat{\sigma}_\eta^2$	$\hat{\sigma}_e^2$	\hat{d}	ϕ	θ
<i>LMSV(1,d,1)</i>	0.3638957	3.1194204	0.3938267	-0.2569355	-0.9936482
<i>Log-verosimiglianza</i>			1.706338		
<i>AIC</i>			0.003017556		
<i>SC</i>			0.01604654		

Tabella 6. 9

Stime FDQML LMSV(1,d,1) della serie $\log y^2$

Osservando i risultati ottenuti, si può notare che tutti gli indici indicano il modello LMSV(0,d,0), come il migliore. Nella figura 6.15 viene rappresentato il periodogramma dei dati osservati e la stima dello spettro implicata dal modello stimato in precedenza. Si può notare che, a parte il primo valore, la stima si adatta, almeno visivamente, ai dati empirici.

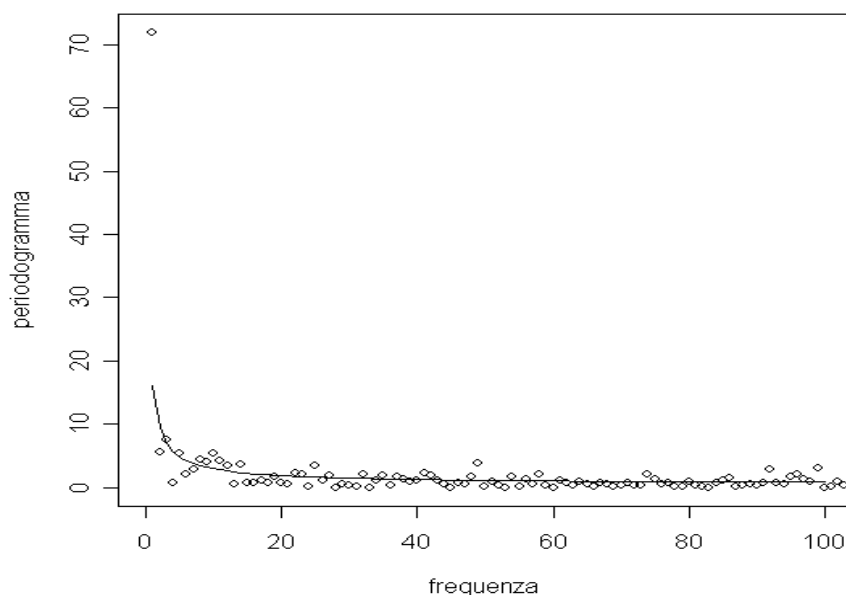


Figura 6. 15

Periodogramma dei dati Autostrade S.p.A con stima dello spettro secondo la tabella 6.6

Per concludere, riporto in figura 6.16 le autocorrelazioni empiriche e teoriche implicate dai dati e dal modello: si può notare che per i primi valori c'è una sottostima ma, proseguendo con i ritardi, la stima si porta su un livello accettabile.

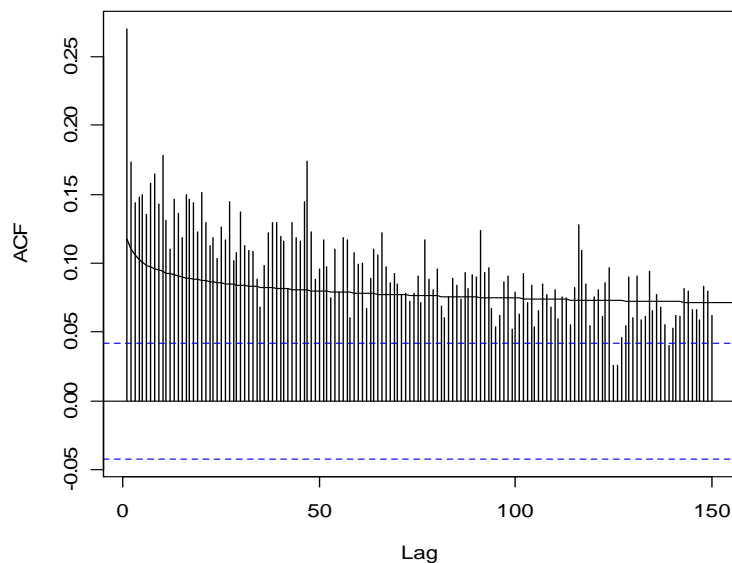


Figura 6.16

Autocorrelazioni empiriche della serie $\log y^2$ e ACF teoriche implicate dai parametri del modello nella tabella 6.6

6.2 Analisi dei dati Mibtel

Analogamente per quanto fatto per la serie Autostrade S.p.A, in questo paragrafo verranno riportate le analisi inerenti l'indice Mibtel. I dati che vengono utilizzati si riferiscono all'andamento del valore dell'indice tra il 2 gennaio 1997 e il 20 maggio 2005 per un totale di 2112 osservazioni. Anche in questo caso i dati mancanti per le festività non sono stati riempiti da alcun valore, per questo nei grafici delle varie analisi non verrà mai riportato il tempo ma l'indice dell'osservazione.

Come prima, poiché nella parte inerente i modelli a volatilità stocastica si utilizzerà una trasformazione dei dati, ovvero il logaritmo dei quadrati di rendimenti, si è resa necessaria una modifica dei dati originari. Infatti, per non avere il problema di calcolare il logaritmo di 0 (in questo caso però gli zeri erano molto pochi), i rendimenti nulli sono stati sostituiti con la media fra il rendimento precedente e quello seguente.

Tutti i dati sono stati reperiti su www.finance.yahoo.com.

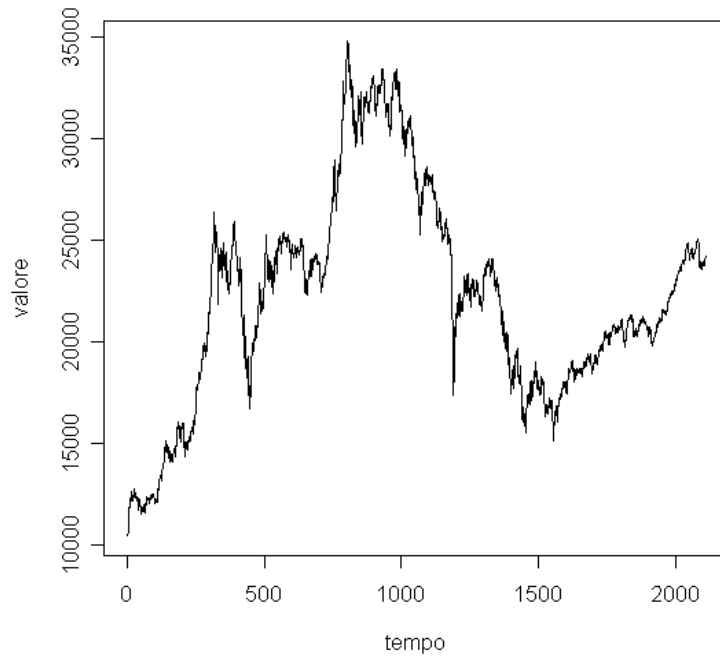


Figura 6. 17

Andamento dell'indice Mibtel

Nella figura 6.17 si può notare la caratteristica di molti titoli azionari di quel periodo, ovvero la crescita che si ha avuto dall'inizio 1999 (circa il cinquecentesimo dato) fino all'inizio del 2002 (circa il millesimo dato).

Nelle figure 6.18 e 6.19 vengono rappresentati gli andamenti dei rendimenti (anche in questo caso calcolati come differenze dei logaritmi dei prezzi) e i rendimenti al quadrato.

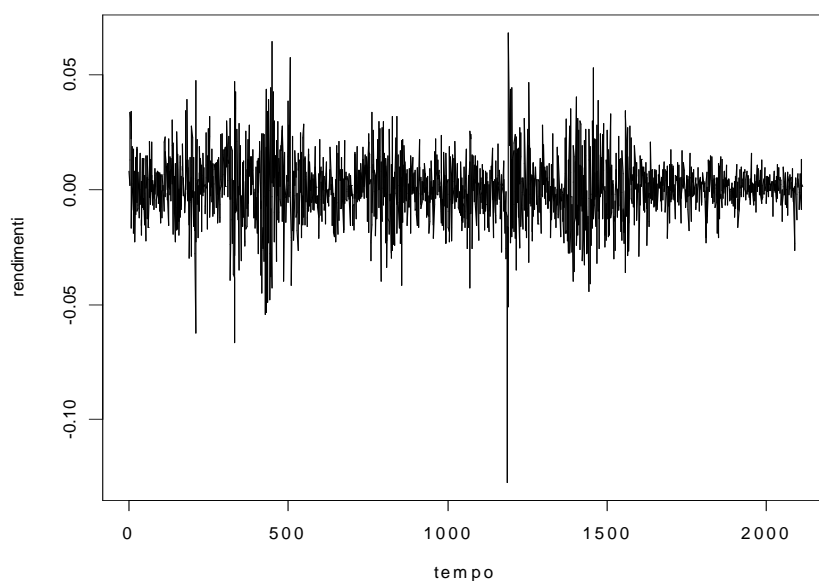


Figura 6. 18

Rendimenti dell'indice Mibtel

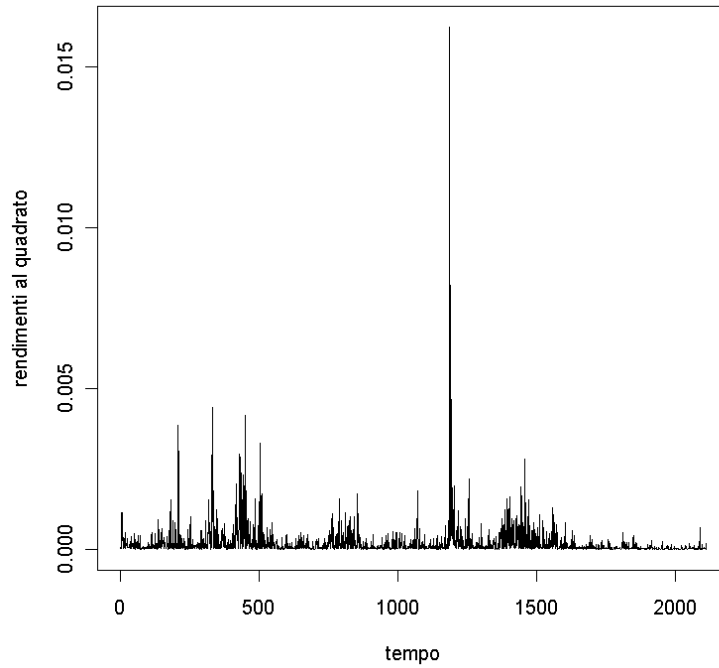


Figura 6. 19

Rendimenti al quadrato dell'indice Mibtel

A questo punto si può passare all'analisi sui livelli: per farla è bene osservare l'ACF e la PACF dei rendimenti rappresentate nelle seguenti figure.

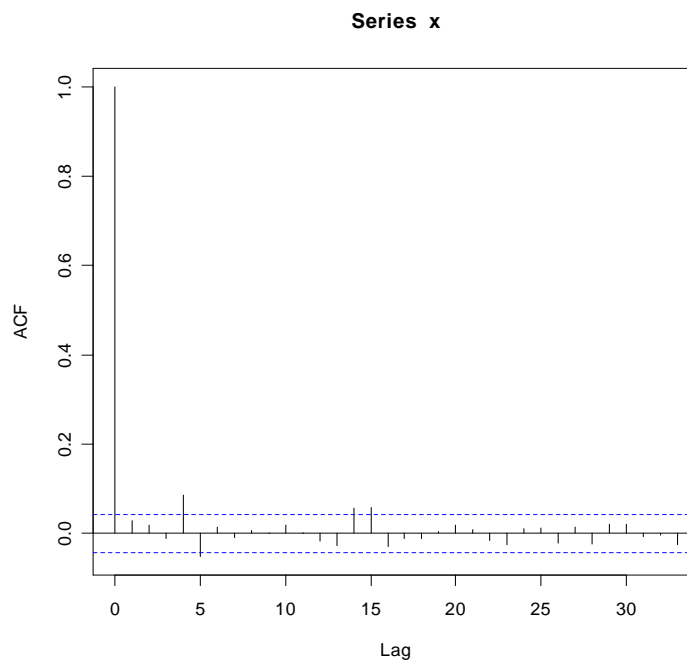


Figura 6. 20

Autocorrelazione dei rendimenti

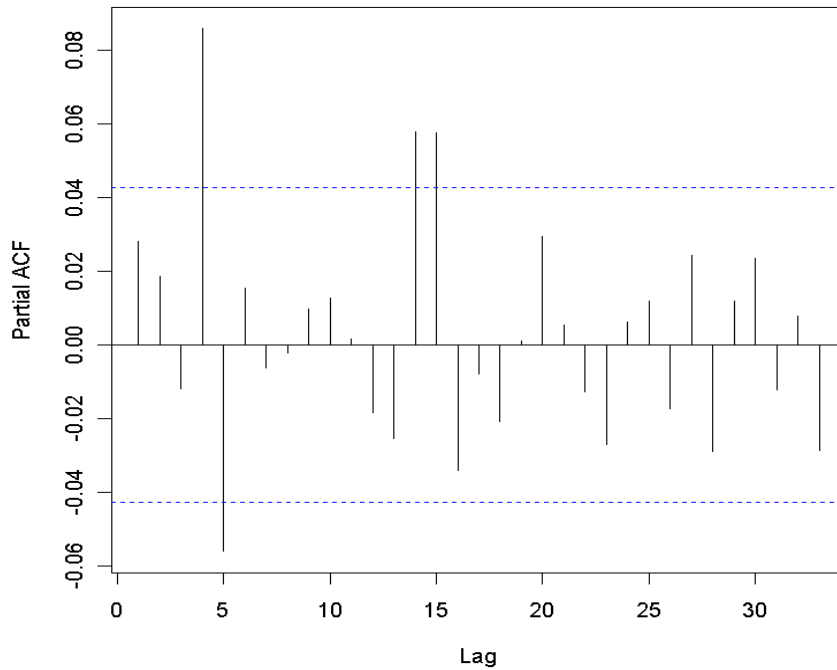


Figura 6. 21

Autocorrelazione parziale dei rendimenti

Anche questa volta i due grafici sembrano indicare la presenza di una qualche componente ARMA sulla media. Dopo alcuni tentativi, il modello migliore sembra essere un AR(4) senza costante. Nella tabella 6.10 vengono presentati i risultati della regressione.

Variabile	Coefficiente	Standard error	Statistica t	P-value
AR(4)	0.086223	0.021678	3.977365	0.0001
<i>Log-verosimiglianza</i>			6058.74	
<i>AIC</i>			-5.750113	
<i>SC</i>			-5.747430	

Tabella 6. 10

Stima AR(4) sui dati della figura 6.18

Stimato il comportamento sulla media, si può passare all'analisi della varianza condizionata. Il correlogramma dei residui al quadrato del modello della tabella 6.10 e i risultati della tabella 6.11 indicano la presenza di qualche dinamica sulla varianza condizionata.

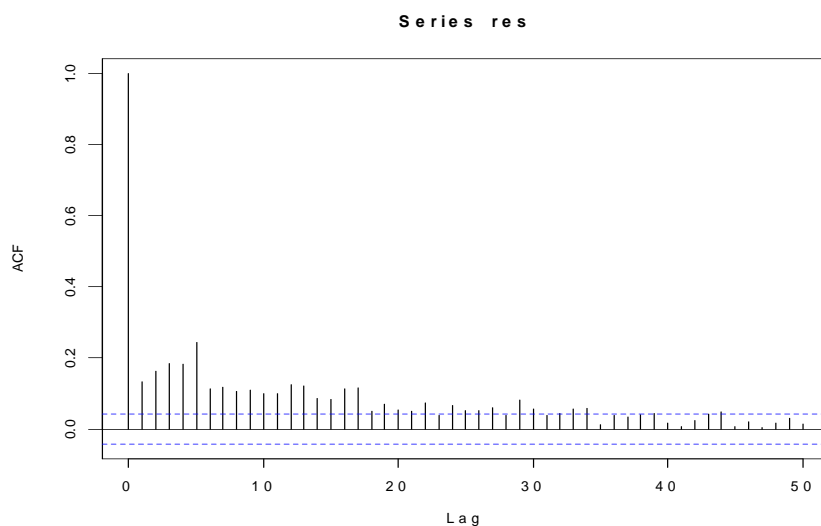


Figura 6. 22

Autocorrelazione dei residui al quadrato della regressione della tabella 6.10

Variabile	Coefficiente	Standard error	Statistica t	P-value
<i>c</i>	0.000161	1.15E-05	14.02266	0.0000
<i>Res²(-1)</i>	0.134182	0.021598	6.212854	0.0000
<i>Statistica F</i>			38.59955	
<i>P-value</i>			0.000000	
<i>TR²</i>			37.94020	
<i>P-value</i>			0.000000	

Tabella 6. 11

Test ARCH-LM sui residui della tabella 6.10

Prima di procedere con la stima della varianza condizionata, si può notare che i residui della tabella 6.10 presentano la caratteristica di non normalità, visto il marcato eccesso di curtosi (8.184751) e il valore del test *JB* di Jarque-Bera (2413.043). Quindi, anche in questo caso, è opportuno procedere con una procedura di quasi massima verosimiglianza.

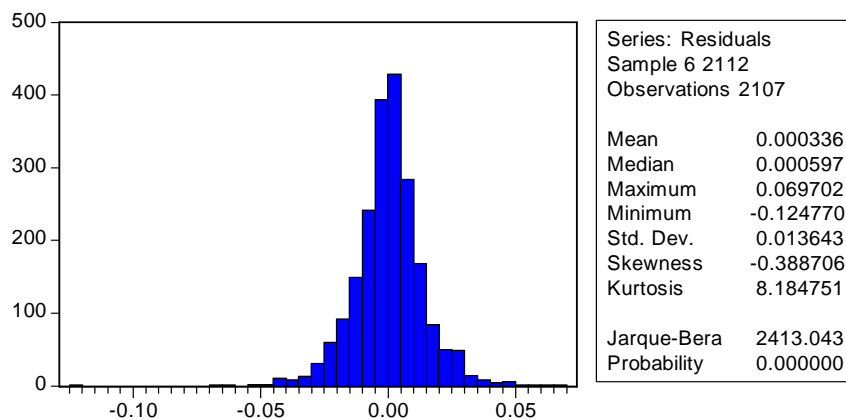


Figura 6. 23

Statistiche descrittive e di normalità dei residui della tabella 6.10.

A questo punto si può passare all'integrazione frazionaria. Anche in questo caso i seguenti grafici indicano l'occorrenza dell'integrazione frazionaria; nelle figure 6.24, 6.25, 6.26 si nota che le autocorrelazioni rimangono significative per molti ritardi mentre nella figura 6.27 si notano chiaramente dei picchi a frequenze molto basse.

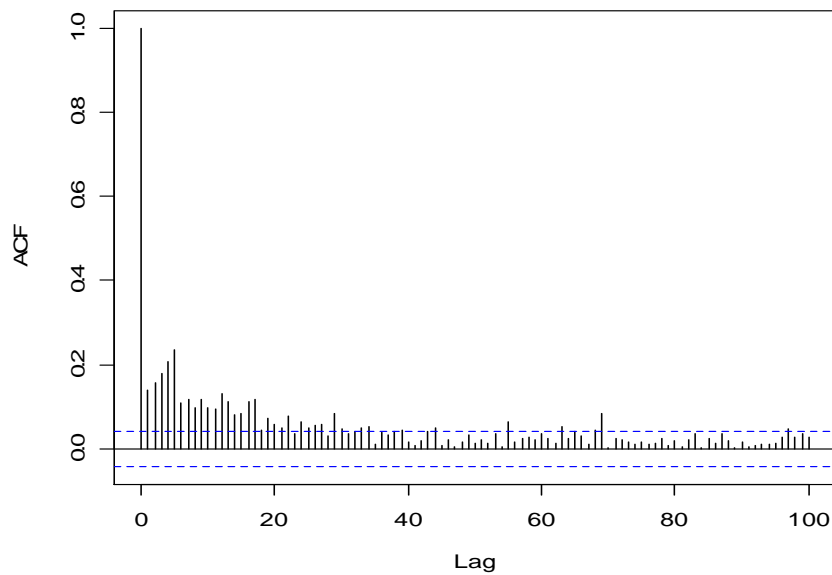


Figura 6. 24

Autocorrelazione dei rendimenti al quadrato

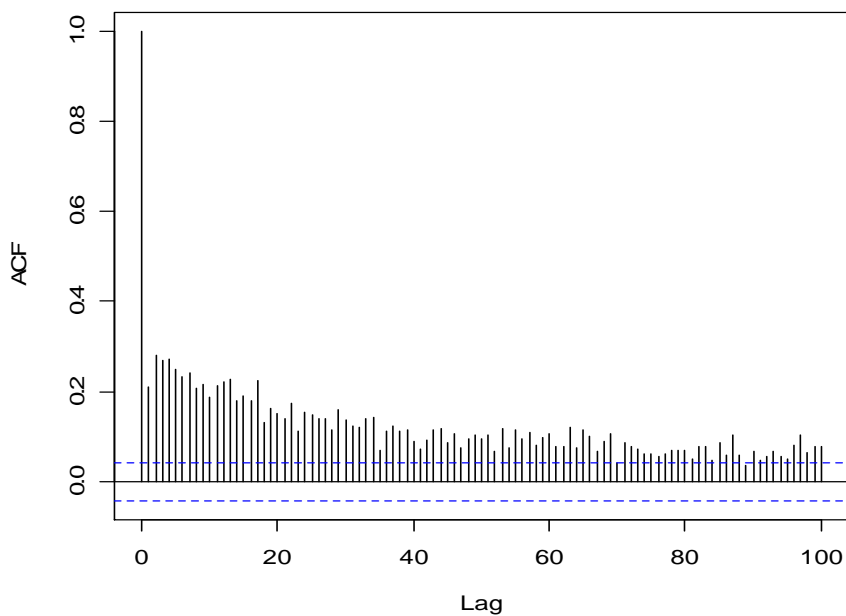


Figura 6. 25

Autocorrelazioni del modulo dei rendimenti

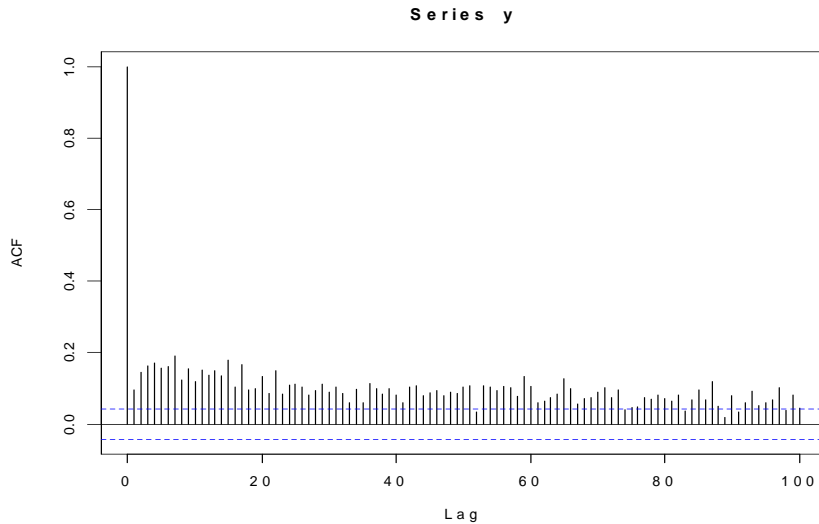


Figura 6. 26

Autocorrelazioni dei logaritmi dei rendimenti al quadrato

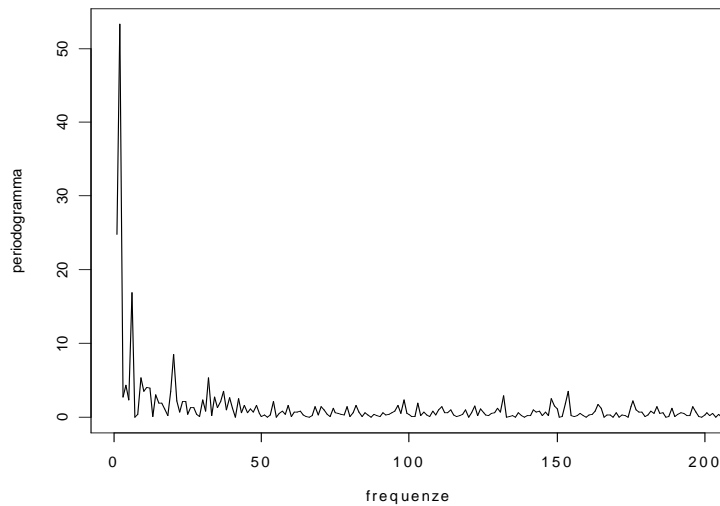


Figura 6. 27

Periodogramma dei logaritmi dei rendimenti al quadrato

Come in precedenza il primo modello che verrà stimato è un FIGARCH(1,d,0).

Nella tabella seguente vengono riportati i risultati relativi a tale modello:

Variabile	Coefficiente	Standard error	Statistica t	P-value
<i>AR(4)</i>	0.0786795	0.0284	2.7709	0.0028
ω	0.00009096	0.0000087	10.4457	0.0000
<i>GARCH(1)</i>	-0.45527445	0.07421	-6.135	0.0000
<i>d</i>	0.39077962	0.06665	5.8632	0.0000
<i>Log-verosimiglianza</i>			6343.56	
<i>AIC</i>			-6.006215	
<i>SC</i>			-5.9955	

Tabella 6. 12

Stima QML FIGARCH(1,d,0) sui dati della figura 6.18

Si può notare che la verosimiglianza aumenta rispetto al modello AR(4) (che aveva un valore di 6058.74) e anche i due criteri informativi AIC e SC concordano nel reputare migliore il modello FIGARCH(1,d,0); comunque, come nel caso del titolo Autostrade, inserisco nell'equazione della varianza condizionata un termine ARCH(1) e di seguito riporto i risultati ottenuti:

Variabile	Coefficiente	Standard error	Statistica t	P-value
AR(4)	0.06549	0.02336442	2.803	0.0026
ω	0.000002228	0.00003813	0.05844	0.4767
GARCH(1)	0.692361	0.36963441	1.87309	0.0305
ARCH(1)	-0.356498	0.00094103	-37.884	0.0000
d	0.356913	0.02935	12.161	0.0000
Log-verosimiglianza			6389.36	
AIC			-6.04866	
SC			-6.035265	

Tabella 6. 13

Stima QML FIGARCH(1,d,1) sui dati della figura 6.18

Con questo nuovo modello la costante ω è diventata non significativa, mentre tutti gli altri parametri sono statisticamente diversi da zero ad un livello del 5%; inoltre la verosimiglianza e i due criteri informativi concordano nel selezionare quest'ultimo modello come il migliore per i dati dell'indice Mibtel.

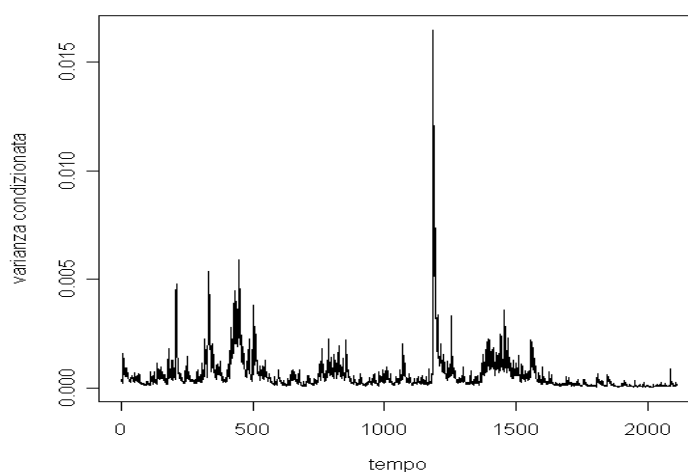


Figura 6. 28

Varianza condizionata implicata dal modello della tabella 6.13

Nella figura sottostante viene rappresentato il grafico dei residui standardizzati che, a differenza dei titoli Autostrade S.p.A, non ha un andamento del tutto soddisfacente in quanto permangono ancora alcuni cluster di volatilità. Forse, per poter rendere questi residui migliori bisognerebbe ipotizzare qualche altra tipologia di modelli per questi dati: poiché l'interesse di questa tesi è rivolta verso

i modelli FIGARCH, non mi addentro ulteriormente nell'analisi e presento il miglior modello delle classe sin qui discussa.

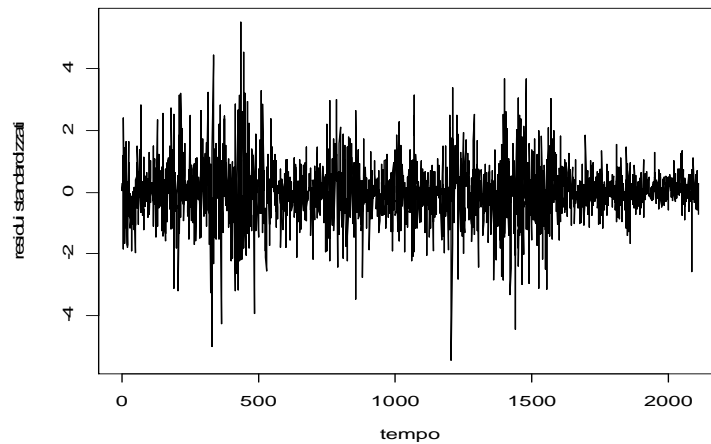


Figura 6. 29

Residui standardizzati del modello della tabella 6.13

Nella figura sottostante vengono riportate le statistiche relative ai residui standardizzati:

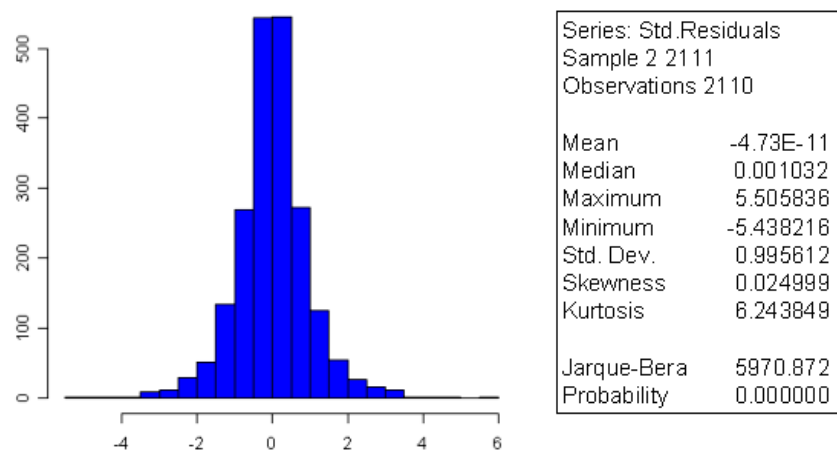


Figura 6. 30

Statistiche descrittive e di normalità dei residui della tabella 6.13

Come in precedenza, si può osservare che non si riesce ad avvicinarsi alla normalità perché permane un'eccessiva curtosi dovuta alla presenza di troppi valori sulle code. Ciò indica chiaramente, come si era ipotizzato anche in questo caso, che la distribuzione dei termini di disturbo non è normale, ed è quindi stato saggio procedere utilizzando la quasi-massima verosimiglianza.

6.2.1 Modelli a volatilità stocastica per l'indice Mibtel

In questa sezione si forniranno le stime ottenute modellando la serie dei logaritmi dei rendimenti al quadrato dell'indice Mibtel ipotizzando diverse specificazioni del modello LMSV; la scelta del modello migliore, in maniera analoga al titolo Autostrade S.p.A, verrà effettuata comparando il valore della verosimiglianza nel punto di massimo e i criteri informativi AIC e SC. I modelli stimati saranno ancora una volta del tipo

$$x_t = \log y^2 = u_t + e_t$$

dove, come al solito, $\{e_t\}$ è i.i.d. con media 0 e varianza σ_e^2 indipendente da $\{u_t\}$ e $\{u_t\}$ è un processo ARFIMA(p,d,q),

$$(1-B)^d \phi(B)u_t = \theta(B)\eta_t$$

dove $\{\eta_t\}$ è un processo i.i.d. con media 0 e varianza σ_η^2 .

Le stime ottenute sono riassunte nelle seguenti tabelle:

Modello	$\hat{\sigma}_\eta^2$	$\hat{\sigma}_e^2$	\hat{d}
<i>LMSV(0,d,0)</i>	0.1003540	4.6780865	0.4553162
<i>Log-verosimiglianza</i>		8.228045	
<i>AIC</i>		-0.004953145	
<i>SC</i>		0.003083212	

Tabella 6. 14

Stime FDQML LMSV(0,d,0) della serie $\log y^2$

Modello	$\hat{\sigma}_\eta^2$	$\hat{\sigma}_e^2$	\hat{d}	θ
<i>LMSV(0,d,1)</i>	2.9503149	1.9657618	0.4501743	-0.6028725
<i>Log-verosimiglianza</i>			8.196615	
<i>AIC</i>			-0.00397595	
<i>SC</i>			0.006739194	

Tabella 6. 15

Stime FDQML LMSV(0,d,1) della serie $\log y^2$

Modello	$\hat{\sigma}_\eta^2$	$\hat{\sigma}_e^2$	\hat{d}	ϕ
<i>LMSV(1,d,0)</i>	0.5786444	4.2107636	0.4651192	0.5706072
<i>Log-verosimiglianza</i>			8.219023	
<i>AIC</i>			-0.003997180	
<i>SC</i>			0.006717964	

Tabella 6. 16

Stime FDQML LMSV(1,d,0) della serie $\log y^2$

Modello	$\hat{\sigma}_\eta^2$	$\hat{\sigma}_e^2$	\hat{d}	ϕ	θ
<i>LMSV(1,d,1)</i>	0.1932187	4.1076622	0.3695957	0.7337923	-0.9967632
<i>Log-verosimiglianza</i>			7.782812		
<i>AIC</i>			-0.002636487		
<i>SC</i>			0.01075744		

Tabella 6. 17

Stime FDQML LMSV(1,d,1) della serie $\log y^2$

Anche in questo caso il miglior modello per i dati è un LMSV(0,d,0), come riportato nella tabella 6.14; in particolare si nota un valore di $\hat{d} = 0.4553162$, ovvero prossimo alla non stazionarietà. Infine, nelle seguenti figure vengono riportati i grafici del periodogramma con la stima dello spettro e il grafico delle autocorrelazioni empiriche di $\log y^2$ con le autocorrelazioni teoriche implicate dal modello dal modello della tabella 6.14.

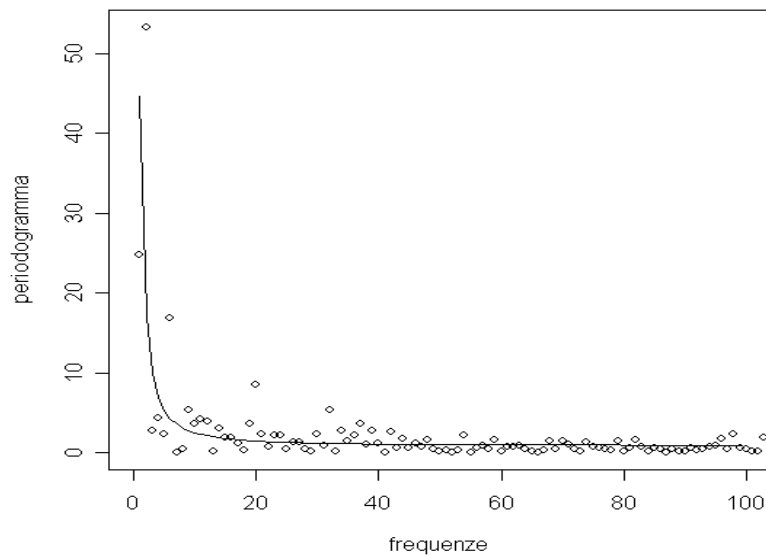


Figura 6. 31

Periodogramma dei dati Mibtel con stima dello spettro secondo la tabella 6.14

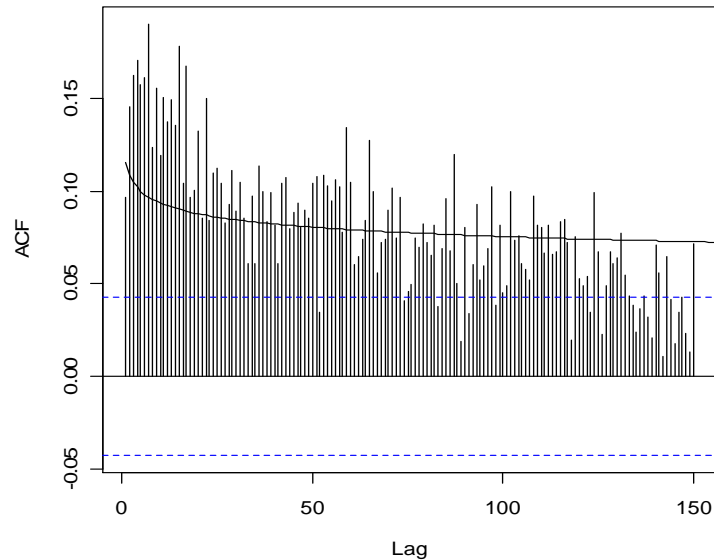


Figura 6.32

Autocorrelazioni empiriche della serie $\log y^2$ e ACF teoriche implicate dai parametri del modello nella tabella 6.14

In questo caso si può notare che le autocorrelazioni teoriche tendono a sottostimare quelle empiriche per i primi valori mentre tendono a sovrastimarle proseguendo con i ritardi.

6.3 Conclusioni

Studi condotti su campioni reali hanno suggerito l'idea che ci sia la presenza di memoria lunga nella varianza condizionale di alcune attività finanziarie, come i tassi di cambio, indici di mercato e anche titoli azionari. Lo scopo di questa tesi era di esporre quali fossero le caratteristiche, la struttura e le procedure di stima di modelli che ammettono memoria lunga nella volatilità condizionale.

Nel Capitolo 2 sono stati introdotti i modelli a volatilità stocastica con memoria lunga (LMSV), che rappresentano una famiglia di modelli analiticamente trattabili che possono essere stimati usando gli strumenti tipici dell'analisi dei processi stazionari: in particolare i modelli LMSV sono stati costruiti a partire dalla famiglia dei modelli a memoria lunga ARFIMA, i quali sono ben conosciuti e da cui i modelli LMSV "ereditano" le proprietà. Inoltre è stato provato che lo stimatore FDQML proposto per questi modelli è consistente e prove di

simulazioni su campioni finiti hanno mostrato che delle buone proprietà e che può essere utilizzato nell'ambito delle serie storiche finanziarie.

Nel Capitolo 3 è stata presentata un'altra famiglia di modelli introdotta da Baillie, Bollerslev e Mikkelsen (1996) chiamata fractionally integrated generalized autoregressive conditional heteroskedasticity (FIGARCH): anche questa classe di modelli presentava caratteristiche compatibili con la dipendenza di lungo raggio mostrata dai moduli e dai quadrati dei rendimenti delle attività finanziarie. In particolare degli shock sulla varianza condizionale vengono riassorbiti ad un tasso iperbolico determinato dal parametro di differenziazione frazionaria mentre la dinamica di breve periodo viene modellata dai parametri del convenzionale GARCH. Inoltre lo stimatore di quasi massima verosimiglianza del modello FIGARCH è stato dimostrato essere \sqrt{T} -consistente e asintoticamente normale.

Infine nel Capitolo 4 sono state prese in considerazione due serie di dati reali, l'azione Autostrade S.p.A. e l'indice Mibtel, delle quali è stata stimata la varianza condizionale tramite i modelli introdotti nei capitoli precedenti. Per entrambe le serie si è notato un miglioramento nelle stime passando da memoria corta a memoria lunga (incremento della verosimiglianza nei modelli FIGARCH, autocorrelazione teorica che assecondava meglio l'empirica per i modelli LMSV).

Capitolo 7

7. Appendice

In questa appendice verranno presentate tutte le routine da me sviluppate in linguaggio R per le applicazioni pratiche dei modelli teorici presentati nei capitoli 4 e 5; in pratica ho scritto delle funzioni che calcolano i valori delle formule espresse in precedenza.

Inoltre riporterò uno studio di simulazioni che è stato effettuato per testare la bontà delle funzioni scritte.

Infine verrà presentata in dettaglio la procedura “Splitting Method” sviluppata da Bertelli e Caporin (2002), utile se si vuole calcolare la funzione di autocorrelazione di un processo ARFIMA(p,d,q).

7.1 Alcune funzioni utili

Le prima funzione che presenterò riguarda il calcolo del periodogramma di una serie stazionaria x .

```
periodogram<-function(x)
{
t <- length(x)
tt <- trunc((t-1)/2)
y <- matrix(0,ncol=1,nrow=tt)
i <- c(1:t)
for (j in 1:tt)
{
cos_j <- cos(2*pi*j*i/t)
sen_j <- sin(2*pi*j*i/t)
y[j] <- ((sum(x*cos_j))^2+(sum(x*sen_j))^2)/(2*pi*t)
}
```

```

    }
  return (y)
}

```

Di seguito presenterò le funzioni che calcolano la verosimiglianza dei vari modelli LMSV(p,d,q) che ho utilizzato nelle analisi. Ogni funzione può essere ottimizzata con l'algoritmo che si predilige; in questa tesi si è scelto di utilizzare la funzione `optim` del software R.

1. Modello LMSV(0,d,0)

```

lvs0d0<- function(theta)
{
  s2ni<-theta[1]
  s2e<-theta[2]
  d<-theta[3]
  t<-length(y)
  tt<-trunc((t-1)/2)
  spec<-matrix(0,ncol=1,nrow=tt)
  for (j in 1:tt)
  {
    spec[j]=(s2ni/(2*pi))*abs(2*sin(pi*j/t))^(2*d)+(s2e/(2*pi))
  }
  lv=((2*pi)/t)*sum(log(spec)+period/spec)
  return(lv)
}

```

2. Modello LMSV(1,d,0)

```

lvs1d0<- function(theta)
{
  s2ni<-theta[1]
  s2e<-theta[2]
  d<-theta[3]
  fi1<-theta[4]
  t<-length(y)
  tt<-trunc((t-1)/2)
  spec<-matrix(0,ncol=1,nrow=tt)
  for (j in 1:tt)
  {
    spec[j]=(s2ni/(2*pi))*(1/(((abs(2*sin(pi*j/t)))^(2*d))*(1+fi1^2-
2*fi1*cos(2*pi*j/t))))+(s2e/(2*pi))
  }
  lv=((2*pi)/t)*sum(log(spec)+period/spec)
  return(lv)
}

```

3. Modello LMSV(0,d,1)

```
lvs0d1<- function(theta)
{
  s2ni<-theta[1]
  s2e<-theta[2]
  d<-theta[3]
  th1<-theta[4]
  t<-length(y)
  tt<-trunc((t-1)/2)
  spec<-matrix(0,ncol=1,nrow=tt)
  for (j in 1:tt)
  {
    spec[j]=(s2ni*(1+th1^2+2*th1*cos(2*pi*j/t)))/(2*pi*(abs(2*sin(pi*
    j/t)))^(2*d))+(s2e/(2*pi))
  }
  lv=((2*pi)/t)*sum(log(spec)+period/spec)
  return(lv)
}
```

4. Modello LMSV(1,d,1)

```
lvs1d1<- function(theta)
{
  s2ni<-theta[1]
  s2e<-theta[2]
  d<-theta[3]
  fi1<-theta[4]
  th1<-theta[5]
  t<-length(y)
  tt<-trunc((t-1)/2)
  spec<-matrix(0,ncol=1,nrow=tt)
  for (j in 1:tt)
  {
    spec[j]=(s2ni*(1+th1^2+2*th1*cos(2*pi*j/t)))/((2*pi)*((abs(2*sin(
    pi*j/t)))^(2*d))*(1+fi1^2-2*fi1*cos(2*pi*j/t)))+(s2e/(2*pi))
  }
  lv=((2*pi)/t)*sum(log(spec)+period/spec)
  return(lv)
}
```

OSSERVAZIONE: nel loro lavoro Breidt et al. (1998) affermano che il termine

$$\left|1 - e^{-i\lambda}\right|^{-2d} = \left|2 \sin\left(\frac{\lambda}{2}\right)\right|^{-2d} = (\sqrt{2 - 2 \cos \lambda})^{-2d}$$

della densità spettrale introduce nella verosimiglianza un termine proporzionale a

$$d \sum_{k=1}^{\lfloor n/2 \rfloor} \log(2 - 2 \cos \lambda) 2\pi T^{-1}$$

il cui integrale è improprio ma converge asintoticamente a 0. Gli autori affermano che se questo termine viene eliminato dalla verosimiglianza si hanno degli effetti trascurabili nello stimatore.

In realtà, durante lo studio di simulazioni effettuato (che sarà presentato nel prossimo paragrafo), è stato notato che, su campioni di 1024 osservazioni e 500 repliche, la funzione con questa modifica sottostimava sensibilmente il parametro d'integrazione frazionaria d e quindi non è stata utilizzata.

Comunque, per completezza informativa, riporto di seguito le versioni modificate della verosimiglianza:

1. Modello LMSV(0,d,0)

```
lvs0d0<- function(theta)
{
  s2ni<-theta[1]
  s2e<-theta[2]
  d<-theta[3]
  t<-length(y)
  tt<-trunc((t-1)/2)
  spec<-matrix(0,ncol=1,nrow=tt)
  mod<-matrix(0,ncol=1,nrow=tt)
  for (j in 1:tt)
  {
    spec[j]=(s2ni/(2*pi))*((2-2*cos(2*pi*j/t))^(d))+s2e/(2*pi))
    mod[j]= s2ni+s2e*(2-2*cos(2*pi*j/t))^d
  }
  lv=((2*pi)/t)*sum(log(mod)+period/spec)
  return(lv)
}
```

2. Modello LMSV(1,d,0)

```
lvs1d0<- function(theta)
{
  s2ni<-theta[1]
  s2e<-theta[2]
  d<-theta[3]
  fi1<-theta[4]
  t<-length(y)
  tt<-trunc((t-1)/2)
  spec<-matrix(0,ncol=1,nrow=tt)
```

```

mod<-matrix(0,ncol=1,nrow=tt)
for (j in 1:tt)
{
spec[j]=(s2ni/(2*pi))*(1/(((abs(2*sin(pi*j/t)))^(2*d))*(1+fi1^2-
2*fi1*cos(2*pi*j/t))))+(s2e/(2*pi))
mod[j]=(s2ni+s2e*(1+fi1^2-2*fi1*cos(2*pi*j/t))*(2-
2*cos(2*pi*j/t))^d)/(1+fi1^2-2*fi1*cos(2*pi*j/t))
}
lv=((2*pi)/t)*sum(log(mod)+period/spec)
return(lv)
}

```

3. Modello LMSV(0,d,1)

```

lvs0d1<- function(theta)
{
s2ni<-theta[1]
s2e<-theta[2]
d<-theta[3]
th1<-theta[4]
t<-length(y)
tt<-trunc((t-1)/2)
spec<-matrix(0,ncol=1,nrow=tt)
mod<-matrix(0,ncol=1,nrow=tt)
for (j in 1:tt)
{
spec[j]=(s2ni*(1+th1^2+2*th1*cos(2*pi*j/t)))/(2*pi*(abs(2*sin(pi*
j/t)))^(2*d))+s2e/(2*pi))
mod[j]=s2ni*(1+th1^2+2*th1*cos(2*pi*j/t))+s2e*(2-
2*cos(2*pi*j/t))^d
}
lv=((2*pi)/t)*sum(log(mod)+period/spec)
return(lv)
}

```

4. Modello LMSV(1,d,1)

```

lvs1d1<- function(theta)
{
s2ni<-theta[1]
s2e<-theta[2]
d<-theta[3]
fi1<-theta[4]
th1<-theta[5]
t<-length(y)
tt<-trunc((t-1)/2)
spec<-matrix(0,ncol=1,nrow=tt)
mod<-matrix(0,ncol=1,nrow=tt)

```

```

for (j in 1:tt)
{
spec[j]=(s2ni/(2*pi))*(1/(((abs(2*sin(pi*j/t)))^(2*d))*(1+fi1^2-
2*fi1*cos(2*pi*j/t))))+(s2e/(2*pi))
mod[j]=(s2ni*(1+th1^2+2*th1*cos(2*pi*j/t))+s2e*(1+fi1^2-
2*fi1*cos(2*pi*j/t))*(2-2*cos(2*pi*j/t))^d)/(1+fi1^2-
2*fi1*cos(2*pi*j/t))
}
lv=((2*pi)/t)*sum(log(mod))+period/spec)
return(lv)
}

```

Funzione per stimare l'autocorrelazione di un modello LMSV(0,d,0)

```

acf0d0<-function(k,s2ni,s2e,d)
{
v<-s2ni*gamma(1-2*d)*(gamma(1-d))^2
y<-array(1,k)
for (j in 1:k)
{
y[j]=(s2ni*gamma(j+d)*gamma(1-2*d))/(gamma(j-d+1)*gamma(1-
d)*gamma(d))/(v+s2e)
}
return(y)
}

```

Funzioni che servono per disegnare il grafico di alcuni spettri

```

#calcolo dello spettro di AR(1) con fi=0.5
spettro<-function(t)
{
x<-ts(c(1:314)/100)
y<-array (1,314)
for (j in 1:t)
{
i<-j/100
y [j]<- (1/(2*pi))*(1/(1-cos(i)+0.25))
}
return(plot(x,y,xlab="frequenze",ylab="spettro",type="l"))
}
spettro(314)

```

```

#calcolo dello spettro si un ARFIMA(0,d,0) con d=0.25
spettro<-function(t)
{
x<-ts(c(1:314)/100)
y<-array (1,314)
for (j in 1:t)

```



```

{
i<-j/100
y [j]<- (1/(2*pi))*abs(2*sin(i/2))^(0.5)
}
return(plot(x,y,xlab="frequenze",ylab="spettro",type="l"))
}
spettro(314)

```

Funzioni che calcolano il valore della verosimiglianza di un modello

FIGARCH: di queste funzioni ne sono state create sei varianti, tre per i dati Mibtel e tre per i dati Autostrade, che riguardavano i casi dei modelli FIGARCH(0,d,0), FIGARCH(1,d,0) e FIGARCH(1,d,1).

Per non appesantire ulteriormente questa sezione, riporto solo il caso di un FIGARCH(1,d,1) per i dati Mibtel:

```

llfi1d1<-function(ar,w,a,b,d)
{
ntr=1000
t<-length(y)
ressq<-matrix(1,ncol=1,nrow=t)
phi<-matrix(0,ncol=1,nrow=ntr)
ee<-matrix(0,ncol=1,nrow=ntr+t)
zeta<-matrix(0,ncol=1,nrow=t)
h<-matrix(0,ncol=1,nrow=t)
s<-matrix(0,ncol=1,nrow=ntr)
for(k in 1:ntr)
{
phi[k]=exp(log(gamma(k-d))-log(gamma(k+1)))/gamma(-d)
}
ressq[1:4]=0
for(j in 5:t)
{
ressq[j]=(y[j]-ar*y[j-1])^2
}
mrq<-mean(ressq)
for(j in 1:2211)
{
if(j<=100) ee[j]=mrq else ee[j]=ressq[j-ntr]
}
for(j in 1:t)
{
for(k in 1:ntr)
{
s[k]=phi[k]*ee[j+ntr-k]
}
}
}

```

```

    zeta[j]=sum(s)
  }
  h[1]=w+ee[ntr]-zeta[1]
  for(j in 2:t)
  {
    h[j]=w+ee[j+ntr]-zeta[j]+b*h[j-1]-b*ee[j+ntr-1]+a*zeta[j-1]+b*zeta[j-1]
  }
  lv<-(-t/2)*log(2*pi)-sum(log(h)+ressq/h)
  return(lv)
}

```

7.2 Studio di simulazioni

Per testare la bontà delle funzioni che calcolano la verosimiglianza, è stato eseguito uno studio su simulazioni di modelli LMSV con diversi parametri per valutare la distorsione delle stime. In questo esperimento sono stati considerati dei campioni di lunghezza $T=1024$ e tre classi di modelli, ovvero LMSV(0,d,0), LMSV(1,d,0) e LMSV(1,d,1): per ogni modello sono state considerate diverse combinazioni di parametri. Tutti i risultati riportati nella tabella 5.1 sono stati ottenuti da 500 simulazioni dello stesso modello. Le serie simulate sono state generate con i seguenti comandi in R:

```

e<-rnorm(1024)
x<-ts(fracdiff.sim(1024, ar=*, ma=*, d = *)$series)
z<-exp(x/2)*e
y<-log(z^2)

```

Con questi comandi R pone le due varianze σ_η^2 e σ_e^2 uguali a uno. Infine le stime dei parametri riportate nella tabella sono state calcolate come media dei parametri stimati delle 500 simulazioni e anche gli standard error sono quelli campionari.

Modello	d	ϕ	θ
LMSV(0,0.4,0)	0.3869182 (0.1491559)	***	***
LMSV(0,0.2,0)	0.1396432 (0.1573790)	***	***
LMSV(0,0,0)	-0.03366038 (0.2704920)	***	***
LMSV(0,-0.2,0)	-0.1979597 (0.5032754)	***	***
LMSV(0,-0.4,0)	-0.3660663 (0.6025499)	***	***
LMSV(0.8,0,0)	-0.00575805 (0.2389066)	0.832866 (0.4960998)	***
LMSV(0.8,0.2,0)	0.1802617 (0.1995823)	0.753652 (0.369801)	***
LMSV(0.4,0.4,0)	0.4089075 (0.06795546)	0.30809031 (0.1459602)	***
LMSV(0.4,0.4,0.4)	0.3720265 (0.1969325)	0.33748192 (0.4659155)	0.3142390 (0.5099634)
LMSV(0.4,0.4,0.3)	0.3332638 (0.1133840)	0.37177756 (0.5311291)	0.23776107 (0.6037161)
LMSV(0.8,0.2,0.3)	0.2876595 (0.1650787)	0.77157954 (0.2904797)	0.3430231 (0.3879637)

Tabella 7. 1

Risultati di simulazioni dello stimatore FDQLM con $T=500$

In merito a questi risultati si possono trarre alcune conclusioni:

- Le deviazioni standard delle stime sono pressoché comparabili con quelle ottenute da Breidt et al. (1998) nel loro lavoro.
- Inoltre gli stessi autori affermano che aumentando l'ampiezza del campione le stime dei parametri migliorano sensibilmente (la distorsione e la dispersione delle stime diminuisce significativamente).
- Lo stimatore FDQML fornisce stime meno distorte e più precise quando il parametro di integrazione frazionaria d è positivo (che, all'atto pratico, risulta essere il caso più interessante).
- E' possibile, durante le simulazioni, ottenere dei valori outlier molto grandi, soprattutto con campioni piccoli.

- Infine, si può notare che le stime, sebbene vicine al vero valore del parametro, sono sempre leggermente inferiori, ovvero c'è una sistematica piccola sottostima.

Tutte queste considerazioni hanno fatto sì che questo stimatore potesse essere utilizzato per stimare delle serie reali, anche per la sua semplicità d'uso.

7.3 Lo Splitting Method

In questa sezione riporterò una procedura introdotta da Bertelli e Caporin (2002) utile se si desidera calcolare la funzione di autocovarianze di un processo ARFIMA(p,d,q).

L'assunzione base di questo metodo si fonda sulla decomposizione dell'autocovarianza di un processo ARFIMA(p,d,q) che porta a identificare una relazione tra l'autocovarianza del processo nel suo complesso e le sue due componenti fondamentali: in processo ARMA e il processo a integrazione frazionaria. Se si considera un processo stazionario a memoria lunga $\{z_t\}$ con

funzione di autocovarianze $\gamma^*(j)$, verificato che $\sum_{j=-\infty}^{+\infty} |\psi_j| < \infty$, il processo

ARMA(p,q) $x_t = \psi(B)z_t$ è stazionario e la sua funzione di autocovarianze può essere scritta come

$$\gamma(k) = \sum_{h=-\infty}^{+\infty} \tilde{\gamma}(h) \gamma^*(k-h) \quad (7.1)$$

dove $\tilde{\gamma}(h)$ rappresenta l'autocovarianza della pura componente ARMA data da

$$\tilde{\gamma}(h) = \sum_{j=0}^{\infty} \psi_j \psi_{j-h}$$

Detto ciò, la sua convergenza verso zero è abbastanza veloce cosicché si può troncare l'espressione (7.1) scegliendo un m abbastanza piccolo. Il valore di m può essere scelto osservando i parametri del processo ARMA: se $p=0$ deve essere $m=p$ senza alcuna approssimazione sulla covarianza; in questo caso tutti i metodi sono equivalenti; se $p \neq 0$, m dipende dal valore del coefficiente AR, aumentando con il suo valore assoluto.

Un'altra caratteristica di questa metodologia è il superamento dell'assunzione della radice unitaria sulla componente autoregressiva, cosicché l'equazione (7.1) è ottenuta dalla funzione di autocovarianze del processo ARMA, che non richiede restrizioni.

Bibliografia

ARTECHE J., 2005, *Gaussian semiparametric estimation in long-memory in stochastic volatility and signal plus noise models*, Journal of Econometrics.

AZZALINI A., 2001, *Inferenza statistica. Una introduzione basata sul concetto di verosimiglianza*, Springer.

BAILLIE R. T., BOLLERSLEV T., MIKKELSEN H. O., 1996, *Fractionally integrated generalized autoregressive conditional heteroskedasticity*, Journal of Econometrics.

BERTELLI S., CAPORIN M., 2002, *A note on calculating autocovariances of long-memory processes*, Journal of time series analysis.

BISAGLIA L., 2004, *Serie storiche economiche (c.p.)*, lucidi del corso.

BOLLERSLEV T., 1986, *Generalized autoregressive conditional heteroskedasticity*, Journal of Econometrics.

BORDIGNON S., 2004, *Serie storiche economiche (c.p.)*, lucidi del corso.

BREIDT F. J., CRATO N., DE LIMA P., 1998, *The detection and estimation of long-memory in stochastic volatility*, Journal of Econometrics.

- BROTO C., RUIZ E., 2004, *Estimation methods for stochastic volatility model*, Journal of Economic Surveys.
- CAPORIN M., 2002, *FIGARCH models: Stationarity, estimation methods and the identification problem*, working paper 0202, GRETA.
- CHUNG C. F., 1999, *Estimating the FIGARCH model*, unpublished working paper, National Taiwan University.
- DEO R. S., HURVICH C. M., 2001, *On the log periodogram regression estimator of the memory parameter in long-memory stochastic volatility model*, Econometric Theory.
- DEO R., HURVICH C., LU Y., 2005, *Forecasting realized volatility using a long-memory stochastic volatility model: estimation, prediction and seasonal adjustment*, Journal of Econometrics.
- DI FONZO T., LISI F., 2001, *Complementi di statistica economica: Analisi delle serie storiche univariate*, Cleup.
- DOORNIK J.A., OOMS M., 1999, *A package for Estimating Forecasting and Simulating arfima models*, Oxford: www.nuff.ox.ac.uk/Users/Doornik.
- GALLO G. M., LOMBARDI M. J., 2005, *Processi a memoria lunga e frazionalmente integrati GARCH*, Quaderni 2005/02, Università degli studi di Firenze.

GALLO G. M., PACINI B., 2002, *Metodi quantitativi per i mercati finanziari*, Carocci.

HOSKING J.R.M., 1981, *Fractional differencing*, Biometrika.

HURVICH C. M., RAY B. K., 2003, *The local Whittle estimator of long-memory stochastic volatility*, Journal of Financial Econometrics.

LISI F., 2004, *Serie storiche finanziarie (c.p.)*, lucidi del corso.

PACE L., SALVAN A., 2001, *Introduzione alla statistica: II inferenza, verosimiglianza, modelli*, Cedam.

PEREZ A., RUIZ E., 2003, *Properties of the sample autocorrelations of non-linear transformations in Long Memory Stochastic Volatility models*, Journal of Financial Econometrics.

TSAY R., 2002, *Analysis of financial time series*, Wiley.

WILKINS N., 2004, *Indirect estimation of long-memory volatility model*, netec.mcc.ac.uk.

Ringraziamenti

Ringrazio il mio relatore, il professore Silvano Bordignon, che mi ha ascoltato e consigliato durante questi mesi nei quali ho scritto la tesi.

Ringrazio la professoressa Luisa Bisaglia e il professore Francesco Lisi per la generosità e l'aiuto che mi hanno offerto per risolvere i miei problemi nelle analisi empiriche.

Ringrazio la mia famiglia che mi ha sempre sostenuto durante questi cinque anni sia moralmente che economicamente.

Ringrazio i miei amici Alessandro, Manuele, Antonio, Stefano, Nadir, Diego e tutti quelli che ho conosciuto in questi cinque anni, con i quali ho condiviso gioie e dolori della vita universitaria.

Infine, ma non meno importante, ringrazio la mia fidanzata Federica che ormai da molti anni mi accompagna nella mia vita e, come dico sempre anche a lei, penso che non riuscirei a trovare nessuna ragazza che mi vuole bene come lei.