

Dipartimento di Economia e Finanza
Corso di Laurea Magistrale in Finanza dei Mercati e Quantitativa
Cattedra di Scelte di Portafoglio e Gestione del Risparmio

STRUMENTI ECONOMETRICI APPLICATI
A SCELTE DI INVESTIMENTO:
COINTEGRAZIONE PER UNA STRATEGIA
SPECULATIVA

RELATORE:

Prof. Nicola Borri

CANDIDATO:

Cesare Orsini

Matr. 634061

CORRELATORE:

Prof. Giuseppe Ragusa

ANNO ACCADEMICO 2011/2012

Indice

Introduzione	5
1 Cointegrazione e <i>Pair Trading</i>	9
1.1 Simulazione e descrizione dei dati	9
1.1.1 Generazione di processi integrati	9
1.1.2 Serie storiche cointegrate	9
1.2 Radici unitarie e stazionarietà	12
1.3 Cointegrazione e definizione della strategia di <i>trading</i>	18
1.3.1 Individuazione del periodo di operatività	18
1.3.2 Verifica dell'ipotesi di integrazione e identificazione dell'ordine del ritardo	19
1.3.3 Analisi delle relazioni di cointegrazione e stazionarizzazione del processo	28
1.4 Definizione della strategia di <i>trading</i>	36
1.4.1 Stazionarizzazione del <i>trading period</i>	36
1.4.2 Costruzione di due <i>threshold</i>	38
1.4.3 Generazione di un segnale operativo	38
1.4.4 Definizione di un <i>holding period</i>	39
1.5 Applicazione della strategia	46
1.5.1 Premesse al modello	46
1.5.2 Stazionarietà e implementazione della strategia	47
1.5.3 <i>Back Testing</i> della strategia	56
1.6 Implementazione in R	65
A Appendice Teorica	85
A.1 Processo stocastici stazionari e integrati	85
A.1.1 Stazionarietà	85
A.1.2 White Noise	85
A.1.3 Processi a radice unitaria e non stazionari	86
A.1.4 Test di radice unitaria e di stazionarietà	88

A.2	Modelli Vettoriali	91
A.2.1	Modello Vettoriale Autoregressivo	91
A.2.2	Modelli a correzione dell'errore	92
A.3	Il metodo di Johansen	93
A.3.1	Stima del rango di cointegrazione	94
A.3.2	Stima del vettore di cointegrazione	95
	Conclusioni	85
	Bibliografia	96
	*	

Elenco delle figure

1.1	Serie storiche contenute nel campione 1: titolo A e titolo B	14
1.2	Serie storiche contenute nel campione 2: titolo C e titolo D	15
1.3	Serie storiche contenute nel campione 3: <i>future</i> e ETF . . .	16
1.4	<i>Formation period</i> del campione 1	20
1.5	<i>Formation period</i> del campione 2	21
1.6	<i>Formation period</i> del campione 3	22
1.7	Analisi della ACF e della PACF dei residui del VAR relativo ai titoli A -B	25
1.8	Analisi della ACF e della PACF dei residui del VAR relativo ai titoli C - D	26
1.9	Analisi della ACF e della PACF dei residui del VAR relativo al <i>future</i> e all'ETF	27
1.10	Grafico dei <i>formation period</i> stazionarizzati	33
1.11	Esempio di generazione di un segnale operativo	46
1.12	Grafico dei <i>trading period</i> stazionarizzati	49
1.13	Analisi <i>rolling</i> delle componenti dei vettori di cointegrazione	55
1.14	Distribuzione delle perdite	65

Elenco delle tabelle

1.1	Test per radice unitaria e test di stazionarietà	13
1.2	Test per radice unitaria e test di stazionarietà nelle differenze	18
1.3	Intervalli di osservazioni considerati in ogni campione	19
1.4	Test per radice unitaria e test di stazionarietà per il <i>formation period</i> dei tre campioni	19
1.5	Test per radice unitaria e test di stazionarietà per il <i>formation period</i> dei tre campioni nelle differenze	23
1.6	<i>Information criterion</i> applicati sui <i>formation period</i> dei tre campioni	23
1.7	Risultati dei test effettuati sui residui dei VAR stimati . . .	25
1.8	Risultati del test di Johansen per la coppia di titoli A - B .	31
1.9	Test per radice unitaria e test di stazionarietà per i processi S^{1F} , S^{2F} e S^{3F}	32
1.10	Risultati del test di Johansen per la coppia di titoli C - D .	34
1.11	Risultati del test di Johansen per la coppia di titoli f - E .	35
1.12	Dinamica del processo S^{1T} e generazione dei segnali operativi	50
1.13	Dinamica del processo S^{2T} e generazione dei segnali operativi	51
1.14	Dinamica del processo S^{3T} e generazione dei segnali operativi	51
1.15	Test per radice unitaria e test di stazionarietà per i processi S^{1T} , S^{2T} e S^{3T}	53
1.16	Lotti di negoziazione per l'assunzione di posizioni sul portafoglio	57
1.17	Operazioni eseguite sul portafoglio composto da una posizione lunga sul titolo A e una posizione corta sul titolo B	58
1.18	Operazioni eseguite sul portafoglio composto da una posizione lunga sul titolo C e una posizione corta sul titolo D	59
1.19	Operazioni eseguite sul portafoglio composto da una posizione lunga sul <i>future</i> e una posizione corta sull'ETF	59

1.20	Controvalore della posizione e interessi maturati	61
1.21	Valutazioni del rischio assunto su ogni posizione aperta . .	64

Introduzione

Questo lavoro ha l'obiettivo di sviluppare una strategia di investimento mediante strumenti econometrici utili alla sua costruzione e al *Back Testing* della stessa. In particolare si vuole illustrare l'utilizzo di alcuni dei principali strumenti di econometria finanziaria, applicandoli alla costruzione e alla valutazione di uno specifico modello operativo. Tale modello è basato su una strategia di *pair trading* e cerca di trarre vantaggio dalla cointegrazione che caratterizza le serie storiche di alcuni asset finanziari. Questa tipologia di trading, basata sullo studio econometrico delle relazioni di cointegrazione, rappresenta un nuovo approccio che trova sempre più diffusione tra gli operatori presenti sul mercato¹. Questo ha animato la volontà di verificarne la reale applicabilità.

Tra le scelte di investimento presenti sul mercato, considerando quelle che guardano al titolo da un punto di vista valutativo, l'idea generale è quella di vendere *asset* sopravvalutati e acquistare quelli sottovalutati. A tal proposito è possibile determinare l'eventuale *mispricing* di una attività solo se noto il vero valore del titolo in termini assoluti. Spesso quest'ultimo aspetto costituisce un vera e propria incognita su cui è difficile ottenere certezze utili a fare profitto e a tutelare la nostra iniziativa. L'idea alla base del *Pair Trading* pone questa strategia, basata su una analisi relativa del prezzo, come soluzione operativa a questa difficoltà. Tale tecnica sarà illustrata in maniera più efficace nella parte pratica di questo lavoro, per adesso accenniamo all'idea secondo la quale se due *asset* hanno caratteristiche simili allora il prezzo dovrebbe essere più o meno lo stesso, o almeno legato da una relazione di equilibrio. In questa tipologia di analisi il prezzo specifico del titolo assume una importanza relativa, l'unica cosa importante è che i prezzi delle due attività siano stabilmente vicini. Uno scostamento rilevante potrebbe essere dovuto all'*overpricing* di un titolo, all'*underpricing* dell'altro o ad un errore combinato sul *pricing* di entrambi. Il *Pair Trading* fonda la sua operatività sull'assunzione di posizioni di segno opposto. In particolare viene assunta una posizione lunga sul

¹Alcuni *broker* iniziano ad offrire alla propria clientela delle applicazioni che permettono di andare alla ricerca delle strategie basate su relazioni di questo tipo, con tanto di *back test* e significatività statistica.

titolo che ha un prezzo troppo basso ed una posizione corta su quello che ha un prezzo troppo elevato, nell'attesa che il *mispricing* si corregga. La differenza tra i prezzi delle due attività è catturata dalla nozione di *spread*. L'evoluzione di questa misura fornisce i segnali per l'apertura e la chiusura di posizioni nell'ambito di un vero e proprio modello operativo².

La relazione tra il prezzo dei titoli oggetto della strategia potrebbe essere caratterizzata da un equilibrio di lungo periodo o meglio da una relazione di cointegrazione. A tal proposito anticipiamo l'idea dei due econometrici Engle e Granger, che nel 2003 valse loro il Nobel. I dati che arrivano dal mercato evidenziano come, spesso, le serie storiche dei prezzi di strumenti finanziari presentino delle caratteristiche di non stazionarietà. Tra le varie forme di non stazionarietà consideriamo quella dovuta alla presenza di radici unitarie, che tipicamente caratterizza le serie storiche ($I(1)$) del prezzo di *asset* finanziari. Secondo il lavoro dei due econometrici, due o più serie $I(1)$ sono cointegrate se esiste una combinazione lineare di queste che dia luogo ad un processo stazionario. Quindi nel caso di due serie rappresentate dalle variabili x_1 e x_2 , potremo dire che queste sono cointegrate se, essendo entrambe $I(1)$, esiste un vettore $\beta = [1, -\alpha]$ tale che $z = x_1 - \alpha x_2$ è stazionario. La combinazione z è data dal prodotto vettoriale tra il vettore riga β' , detto vettore di cointegrazione e il vettore colonna formato dalle variabili x_1 e x_2 . A questo punto sostituendo la condizione di vicinanza tra i prezzi di due attività con una relazione di equilibrio, potremmo costruire un modello in grado di sfruttarne gli eventuali scostamenti. Infatti, se consideriamo le variabili integrate date dal logaritmo del prezzo di due generici titoli A e B, individuato il vettore di cointegrazione, la loro combinazione lineare è un processo stazionario. Costruito tale processo, incentriamo la nostra strategia sull'apertura e la chiusura di posizioni utili a far profitto sfruttando il fenomeno di *mean-reversion*, tipico dei processi stazionari. In particolare immaginando tale combinazione come un portafoglio composto da una posizione lunga sul titolo A ed una corta sul titolo B, operiamo monitorando l'andamento del processo. Quindi, acquistiamo il portafoglio quando il processo è al di sotto del suo valore di equilibrio e oltre una certa soglia. Successivamente la

²Inoltre il portafoglio costituito nell'ambito di questa strategia *long-short* presenta, per costruzione, una esposizione minima nei confronti del mercato e rendimenti incorrelati rispetto allo stesso, collocandosi così tra le cosiddette *market neutral strategies*. Queste strategie sono costruite assumendo posizioni di segno opposto su diversi strumenti finanziari, in modo tale che il portafoglio abbia una esposizione pressoché nulla rispetto ad ampi movimenti del mercato. L'obiettivo principale è conseguire un profitto dal *mispricing* relativo tra strumenti più o meno collegati, vendendo quelli percepiti come sovrapprezzati e comprando quelli sottoprezzati, evitando così il rischio sistematico. Nei mercati azionari queste strategie si evolvono nello *statistical arbitrage*. Questa tecnica sviluppa una metodologia basata su una analisi delle serie storiche finalizzata all'individuazione di eventuali errori nel *pricing* di azioni. Il *Pair Trading* è una forma specifica di *Statistical Arbitrage*.

posizione è chiusa nel momento in cui il processo è tornato al valore di equilibrio. Speculare è la strategia adottata nel caso in cui il valore della combinazione si trovi al di sopra del suo valore di equilibrio e oltre una certa soglia.

Il lavoro è composto da un unico capitolo centrale e da una Appendice Teorica. Il capitolo centrale descrive da un punto di vista pratico le fasi necessarie alla costruzione del modello operativo. La prima sezione del capitolo è dedicata alle tecniche di simulazione con cui sono state generate le serie storiche utilizzate. Segue la seconda sezione in cui è illustrata l'applicazione degli strumenti econometrici adatti all'analisi di serie storiche integrate. Nella terza sezione è introdotta la nozione di cointegrazione e la sua potenziale applicabilità nell'ambito di una strategia di investimento. In questa parte sono descritte le regole e i parametri su cui è basata l'operatività. A tale scopo vengono definiti anche i meccanismi di generazione dei segnali utilizzati per l'apertura e la chiusura di posizioni. L'ultima sezione descrive la concreta applicazione del modello operativo. In quest'ultima parte la trattazione è sviluppata mediante un duplice approccio. Da un lato si considera la gestione del portafoglio. Tramite il Back Testing della strategia infatti, vengono rilevati i risultati in termini di rischi e rendimento. Questa analisi è effettuata mediante le tecniche di valutazione delle *performance* e del rischio tipiche dell'*asset allocation*. A questa segue lo studio del modello e la valutazione dei suoi risultati da un punto di vista econometrico. Da ultimo l'Appendice teorica sintetizza la teoria econometrica utilizzata nelle applicazioni pratiche sviluppate nel primo capitolo.

La costruzione ed il *Back Testing* della strategia, con particolare riferimento ai test econometrici, vengono implementati mediante l'utilizzo del software statistico R (www.r-project.org). Tale linguaggio di programmazione è stato scelto per la duttilità e per la sua caratteristica di essere liberamente disponibile ed *open source*. Questo fa sì che i risultati di un lavoro di ricerca sviluppato in R siano verificabili da chiunque intenda effettuare un controllo per dare una valutazione del lavoro. A tale scopo sono riportati in Appendice gli *script* della programmazione utilizzata nella costruzione del modello.

Capitolo 1

Cointegrazione e *Pair Trading*

1.1 Simulazione e descrizione dei dati

1.1.1 Generazione di processi integrati

Per analizzare le difficoltà connesse alla costruzione di un modello in grado di ottenere risultati positivi, si è deciso di generare serie storiche con caratteristiche utili a mettere in luce i punti critici della sua costruzione. Siamo partiti da un campione idoneo a illustrare l'applicazione della strategia, per giungere a delle serie molto vicine a quelle di titoli realmente quotati sul mercato. Anticipiamo fin da subito come il numero e la complessità degli strumenti econometrici richiesti siano positivamente correlati con la vicinanza del campione ai dati reali. I *dataset* utilizzati rappresentano la quotazione giornaliera di titoli presenti sul mercato. Più precisamente si tratta di coppie di generiche attività che indicheremo utilizzando le lettere dell'alfabeto. La numerosità campionaria è pari a 1306 osservazioni, circa cinque anni di quotazione. Campioni di queste dimensioni sono utili a mostrare le tecniche che ci hanno consentito di volgere strumenti econometrici, generalmente sfruttati nell'analisi di relazioni di lungo periodo, a una operatività a più breve termine.

1.1.2 Serie storiche cointegrate

Per ottenere delle serie caratterizzate da una consistente relazione di cointegrazione, abbiamo generato due processi autoregressivi stazionari di primo ordine. Gli errori di tali processi sono estrazioni casuali da una distribuzione normale standard multivariata.

In formule:

$$\begin{aligned}x_{At} &= \alpha x_{At-1} + \epsilon_{At} \\x_{Bt} &= \alpha x_{Bt-1} + \epsilon_{Bt}\end{aligned}\tag{1.1}$$

dove α è minore di 1 e ϵ è un *white noise*. A questi due processi sono stati sommati il processo w_t (*random walk*) e due costanti positive. Così facendo abbiamo ottenuto due serie storiche caratterizzate da un *trend* comune e da un *drift*. In modo più formale:

$$\begin{aligned}y_{At} &= x_{At} + \delta w_t + drift_A \\y_{Bt} &= x_{Bt} + \omega w_t + drift_B\end{aligned}\tag{1.2}$$

con w_t uguale a:

$$w_t = \sum_{i=1}^t \epsilon_i\tag{1.3}$$

Come mostrato nella Eq. 1.2 i due processi possono essere scritti come la somma tra un processo $I(1)$ e una parte stazionaria $I(0)$. Con questa tecnica le due serie, la prima per una costante moltiplicativa δ e la seconda per una costante moltiplicativa ω , oscillano attorno allo stesso trend stocastico w_t . La fluttuazione avviene senza mai allontanarsi da esso perché, per costruzione, ambedue tendono a stare nei paraggi del trend comune. Le costanti positive sono state aggiunte come un *drift* e hanno il semplice ruolo di traslare in alto, a nostro piacimento, il livello del prezzo. Dal sistema rappresentato nella Eq. 1.2 abbiamo ottenuto le serie storiche del prezzo del titolo A e del titolo B. È chiaro che la cointegrazione tra le serie nasce proprio dal fatto che, prendendo per esempio la combinazione $y_{At} - \alpha y_{Bt}$, il trend stocastico si riduce. Di conseguenza rimangono solo le oscillazioni attorno a esso e la differenza fra i due autoregressivi stazionari. In altre parole, il risultato della sottrazione è una combinazione tra processi stazionari che genera un'altro processo stazionario. La metodologia fin qui descritta ci fornisce serie storiche caratterizzate da un'elevata relazione di cointegrazione. Tali serie, seppur un po' distanti dalla realtà osservabile sul mercato, costituiscono la condizione ideale per illustrare lo sviluppo della strategia. Inoltre per disporre di dati prossimi a quelli del mercato reale, idonei a mettere in luce i punti critici della costruzione del modello, abbiamo generato serie storiche più complesse e pressoché identiche a quelle di titoli quotati. Per ottenere questo risultato, sono state utilizzate formule di generazione differenti rispetto alla Eq. 1.2. In particolare, mutuando dalla teoria finanziaria classica l'accostamento dei *pattern* azionari alla dinamica tipica di un *random walk*, la serie dei prezzi del titolo C è stata creata mediante questa passeggiata aleatoria:

$$u_t = \sum_{i=1}^t \epsilon_i \quad (1.4)$$

a cui è stato aggiunto un *drift* utile ad alzarne il livello, ottenendo così:

$$y_{Ct} = u_t + drift_C \quad (1.5)$$

Il passo successivo è stato generare un processo autoregressivo simile a quello presente nella Eq. 1.2:

$$x_{Dt} = \alpha x_{Dt-1} + \epsilon_{Dt} \quad (1.6)$$

che tuttavia incide sulla dinamica con un coefficiente α diverso. In questo caso la non stazionarietà è stata ricercata per ciascuna delle componenti del processo generatore. Per questo motivo il coefficiente α della AR(1) è stato posto pari a 0.9, valore prossimo a quelli che renderebbero l'autoregressivo non stazionario. Inoltre con l'obiettivo di rendere più complesso e realistico l'andamento della serie, è stato aggiunto alla stessa il trend deterministico TD . In formule:

$$TD = 1 + \eta t \quad (1.7)$$

dove t è il tempo e η un numero decimale scelto a seconda della velocità desiderata per gli incrementi della componente di trend. In fine per garantire un collegamento, dall'intensità variabile ma persistente, i prezzi del titolo D sono stati generati dalla combinazione del processo y_{Ct} , del *random walk* x_{Dt} e del trend deterministico TD . In modo più formale:

$$y_{Dt} = \phi y_{Ct} + TD + x_{Dt} \quad (1.8)$$

Dalle equazioni 1.5 e 1.8 abbiamo ottenuto le serie storiche del prezzo del titolo C e del titolo D. Queste risultano cointegrate con diverso grado di intensità a seconda dell'intervallo di osservazioni. La relazione assume pertanto valori diversi all'interno del campione. Tale variabilità ci permette di evidenziare i rischi derivanti dall'implementazione della strategia in condizioni sempre più vicine alla realtà. Questi diversi meccanismi di generazione dei dati sono stati utili a un graduale avvicinamento alle condizioni del mercato reale. In fine, la ricerca di un riscontro reale ci ha condotto all'utilizzo di un campione costituito dalle quotazioni giornaliere di un ETF iShares su titoli governativi americani¹ e di un *future* su titoli

¹L'ETF *iShares Barclays 10-20 Year Treasury Bond Fund*, quotato sul NYSE Arca, include le performance dell'indice obbligazionario *Barclays Capital U.S. 10-20 Year Treasury Bond*. L'indice

trentennali del Tesoro ², per il periodo che va dall'8 Giugno 2007 all'8 Giugno 2012. Con i dati descritti fin'ora sono stati costruiti tre campioni, ciascuno costituito da due serie storiche: il campione 1, contenente 1306 osservazioni relative al logaritmo naturale dei prezzi del titolo A e del titolo B; il campione 2 formato da quelli del titolo C e del titolo D; il campione 3 che include il logaritmo naturale delle quotazioni giornaliere del contratto *future* e dell'ETF appena descritti.

1.2 Radici unitarie e stazionarietà

I processi integrati, così come li abbiamo costruiti, hanno delle caratteristiche che li rendono molto interessanti da un punto di vista pratico. Le loro realizzazioni infatti, simulano bene le serie storiche che si incontrano sui mercati finanziari. Esaminando le conseguenze della non stazionarietà di processi di questo tipo sulla nostra strategia, si può affermare che se y_t è $I(1)$, allora Δy_t è $I(0)$ per definizione. In questo caso, tutti gli strumenti per la stima dei parametri che caratterizzano i processi stazionari possono essere utilizzati senza problemi, modellando per esempio: un tasso di rendimento anziché il (logaritmo del) prezzo; il tasso di inflazione anziché (il logaritmo) dell'indice dei prezzi. Tuttavia un approccio di questo tipo presuppone che si sappia preventivamente se una serie è integrata o stazionaria ³ e quindi se rappresentabile con un processo $I(0)$ o $I(1)$. Tale caratteristica nella realtà è molto difficile da stabilire a priori. Nel nostro caso, da un punto di vista operativo, la via più veloce è fornita dai dati stessi. Infatti osservando i grafici delle Figure 1.1, 1.2 e 1.3, è evidente la presenza di andamenti regolari crescenti che si alternano a fasi in cui il trend è decrescente. In nessuno dei grafici si riscontra una oscillazione continua attorno a un valore costante, segno distintivo, da un punto di vista grafico, di un processo stazionario. Pertanto qualora non conoscessimo i processi con cui sono state generate le nostre serie, visualizzando i grafici potremmo dedurre che sarebbero ben approssimate da processi $I(1)$. Tale tecnica, che affiancata da una conoscenza del mercato e dello strumento considerato potrebbe

misura le performance delle obbligazioni del Tesoro statunitense con vita residua compresa tra i dieci e i venti anni. Il fondo investe in un campione rappresentativo di titoli inclusi nell'indice, il paniere così costituito è caratterizzato da un profilo di investimento simile all'indice stesso. Talvolta il fondo è ribilanciato in modo tale da ottenere un replica fisica dell'indice.

²Questo strumento è un *future* scambiato sul CME che ha come sottostante titoli del Tesoro americano, *Treasury bond* trentennali. Nello specifico si tratta di obbligazioni governative *callable* non richiamabili per almeno quindici anni dal primo giorno del mese di consegna e obbligazioni non rimborsabili anticipatamente con una vita residua maggiore o uguale a quindici anni.

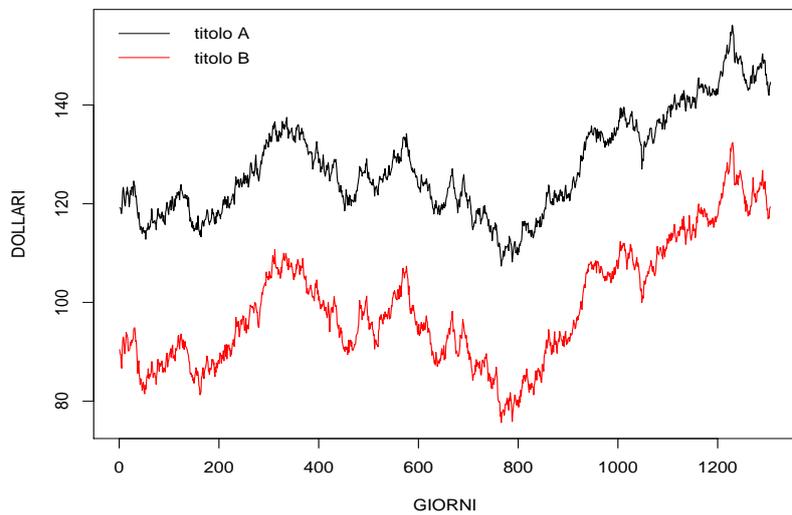
³Questa è indubbiamente una semplificazione, primo perché con un'analisi più approfondita si possono dare esempi di processi che non sono né $I(0)$ né $I(1)$, secondo poiché l'integrazione non è una caratteristica della serie storica, ma del processo stocastico che adottiamo per darne una rappresentazione statistica.

(a) Prezzi				(b) Logaritmo naturale dei prezzi			
Titolo	ADF	PP	KPSS	Titolo	ADF	PP	KPSS
<i>p-value</i>				<i>p-value</i>			
A	0.67	0.68	≤ 0.01	A	0.66	0.68	≤ 0.01
B	0.67	0.67	≤ 0.01	B	0.65	0.65	≤ 0.01
C	0.56	0.36	≤ 0.01	C	0.54	0.28	≤ 0.01
D	0.77	0.72	≤ 0.01	D	0.74	0.68	≤ 0.01
<i>future</i>	0.42	0.44	≤ 0.01	<i>future</i>	0.36	0.39	≤ 0.01
<i>ETF</i>	0.44	0.42	≤ 0.01	<i>ETF</i>	0.34	0.33	≤ 0.01

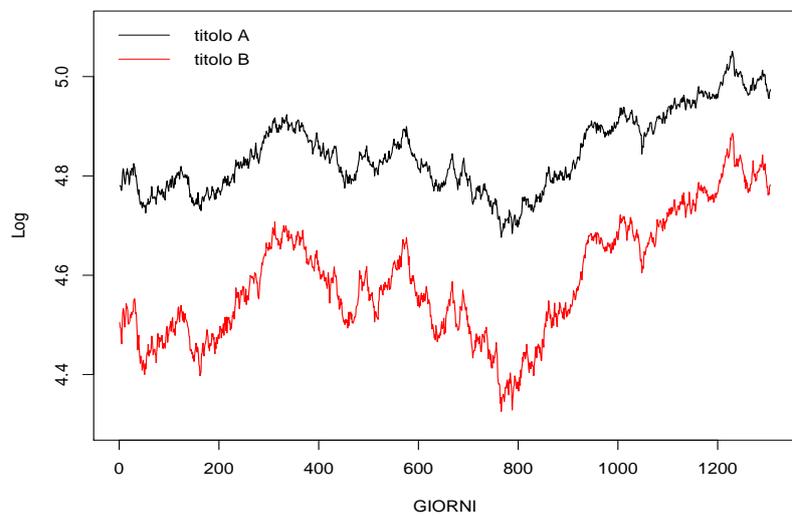
Tabella 1.1: **Test per radice unitaria e test di stazionarietà**

mostrarsi efficace, risulta però troppo semplicistica per almeno tre motivi: in primo luogo, perché un giudizio del genere è piuttosto soggettivo e scarsamente formalizzabile; in secondo luogo, è possibile che un processo sia stazionario attorno a un trend deterministico; infine perché esiste la possibilità che un processo effettivamente $I(1)$ dia luogo a realizzazioni prive di una tendenza particolarmente marcata a salire o a scendere. Per tutte queste ragioni è necessaria una regola di decisione meno arbitraria e più affidabile. Regole di questo tipo sono note come test di radice unitaria, tra questi l'Augmented Dickey-Fuller Test (ADF) e il Phillips-Perron Test (PP) e test di stazionarietà, nel nostro caso il Kwiatkowski-Phillips-Schmidt-Shin Test (KPSS). A proposito di entrambe le tipologie, è ampia la letteratura che tende a metterne in evidenza la debolezza⁴. Nel nostro caso i vari test sono stati utilizzati in modo congiunto con l'obiettivo di trovare conferme coerenti tra i risultati delle varie tipologie. La metodologia utilizzata per interpretare i loro risultati è spiegata nella Sez. 1.5.2. Nella Tabella 1.1 sono contenuti i risultati dei test, ADF, PP e KPSS effettuati sui tre campioni presi in esame. Da questi risulta evidente la presenza di radici unitarie nelle sei serie storiche considerate. In particolare il test ADF, per ciascuna delle serie, restituisce un *p-value* molto elevato.

⁴La questione è trattata in modo più approfondito nella Sez. 1.5.2 tuttavia anticipiamo un commento sui test di radice unitaria: accade spesso che effettuando un test di radice unitaria su una serie storica la quale ragionevolmente dovrebbe fluttuare all'interno di una banda più o meno ampia, non sia possibile rifiutare l'ipotesi di radice unitaria. Questo a volte, si verifica anche con delle serie storiche che notoriamente godono di una fisiologica stazionarietà, come ad esempio i tassi di rendimento dei BOT. In questi casi bisognerebbe mettere da parte l'adesione dogmatica al *p-value*. Piuttosto si dovrebbe ponderare la possibilità che nel campione a nostra disposizione, la serie storica abbia un grado di persistenza tale per cui è meglio, da un punto di vista di aderenza ai dati, pensarla come una realizzazione di un processo $I(1)$ piuttosto che $I(0)$. Questo non vuol dire affermare che la serie sia $I(1)$. In realtà ammesso che abbia senso immaginare la nostra serie come realizzazione di un qualche processo stocastico, noi stiamo solo scegliendo all'interno di una classe limitata di processi (gli ARIMA), la parametrizzazione più appropriata per descrivere i dati. Se poi avessimo osservazioni su migliaia di anni, allora probabilmente il processo più idoneo a rappresentare l'andamento della nostra serie nel tempo sarebbe un $I(0)$. Quindi tutto si riduce a un problema di rappresentazione. Con un test di radice unitaria non stiamo affermando se il processo è $I(1)$ oppure $I(0)$. Stiamo soltanto decidendo se è più conveniente rappresentare il nostro campione con un processo stazionario o integrato.

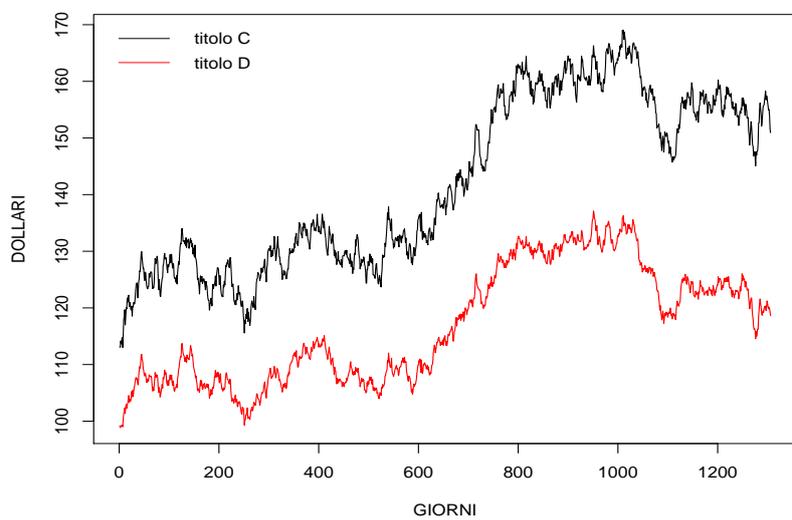


(a) Prezzo

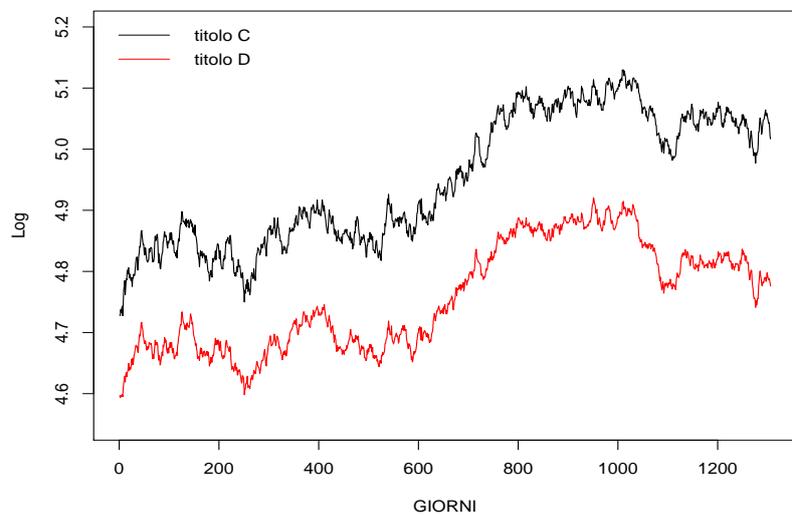


(b) Logaritmo naturale del prezzo

Figura 1.1: Serie storiche contenute nel campione 1: titolo A e titolo B

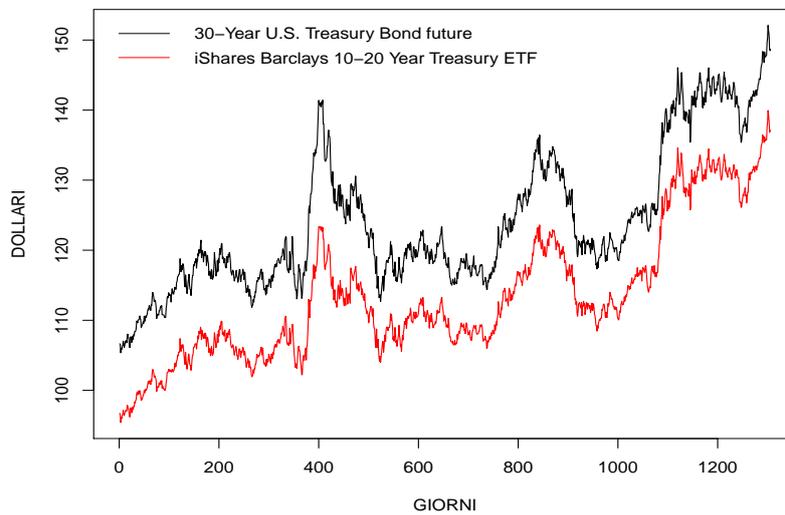


(a) Prezzo

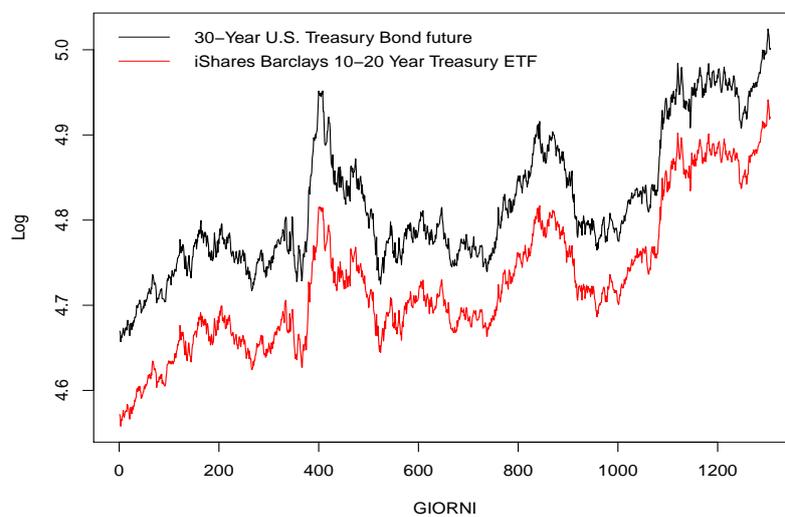


(b) Logaritmo naturale del prezzo

Figura 1.2: Serie storiche contenute nel campione 2: titolo C e titolo D



(a) Prezzo



(b) Logaritmo naturale del prezzo

Figura 1.3: Serie storiche contenute nel campione 3: *future* e *ETF*

Questo risultato ci porta a non poter rifiutare l'ipotesi nulla del test⁵ e quindi a considerare le serie in esame come processi integrati a radice unitaria. Nonostante il test ADF fornisca valori p elevati per tutte le serie storiche considerate, rileviamo una differenza tra i risultati ottenuti per la coppia di titoli A - B e quelli che il test ha prodotto sugli altri campioni. Tali differenze trovano giustificazione nelle tecniche utilizzate per la generazione dei dataset. Come anticipato, l'Augmented DickeyFuller test (come gli altri) ha una potenza molto bassa nel discriminare tra le ipotesi alternative e ha una scarsa validità nel caso in cui i dati presentino *jump* o *structural break* nel processo generativo. Inoltre gli errori della regressione alla base del test ADF, sono assunti essere i.i.d.. Tale ipotesi molto spesso differisce dalla realtà. Tuttavia per rendere più deboli e meno restrittive le nostre assunzioni, ci siamo serviti di un'altro test di radice unitaria, il test Phillips Perron. Un vantaggio di quest'ultimo rispetto al test ADF è la sua maggior robustezza nell'ipotesi di eteroschedasticità dei residui e nell'insensibilità alla specificazione del numero di variabili lagged⁶. Questa proprietà, molto utile alle nostre necessità, rende il test PP più idoneo per un'analisi avente a oggetto dati finanziari. Anche in questo caso i risultati ottenuti ci spingono ad accettare l'ipotesi della presenza di radice unitaria in tutte le serie storiche del nostro dataset. Infatti come mostrato nella Tabella 1.1, i p -value delle statistiche sono molto elevati rispetto al valore soglia 0.05 anche nel caso del test di Phillips-Perron. Infine è stato utilizzato il test KPSS per testare la stazionarietà dei processi esaminati. I valori riportati nella Tabella 1.1 rigettano l'ipotesi nulla di stazionarietà del processo, fornendoci dunque un risposta concorde con i risultati ottenuti dai test di radice unitaria. Come anticipato, siamo giunti alla scelta di processi $I(1)$ mediante una lettura congiunta delle tre statistiche. Infine abbiamo trovato una ulteriore conferma (forse più pratica che teorica) applicando i test alle serie storiche differenziate.

⁵Rinviando all'Appendice Teorica per una trattazione più analitica, ricordiamo sommariamente che l'ipotesi nulla dei test ADF e PP consiste nella presenza di radici unitarie mentre quella del test KPSS è la stazionarietà del processo.

⁶La specificazione del numero di variabili ritardate è un aspetto fondamentale del test ADF poiché necessaria a eliminare la correlazione nei residui della regressione utilizzata nel test. L'inclusione di un ritardo troppo basso comporta stime distorte, mentre un numero di *lag* troppo elevato ne riduce la robustezza.

(a) Prezzi				(b) Logaritmo naturale dei prezzi			
Titolo	ADF	PP	KPSS	Titolo	ADF	PP	KPSS
<i>p-value</i>				<i>p-value</i>			
A	≤ 0.01	≤ 0.01	0.10	A	≤ 0.01	≤ 0.01	0.10
B	≤ 0.01	≤ 0.01	0.10	B	≤ 0.01	≤ 0.01	0.10
C	≤ 0.01	≤ 0.01	0.10	C	≤ 0.01	≤ 0.01	0.10
D	≤ 0.01	≤ 0.01	0.10	D	≤ 0.01	≤ 0.01	0.10
<i>future</i>	≤ 0.01	≤ 0.01	0.10	<i>future</i>	≤ 0.01	≤ 0.01	0.10
<i>ETF</i>	≤ 0.01	≤ 0.01	0.10	<i>ETF</i>	≤ 0.01	≤ 0.01	0.10

Tabella 1.2: **Test per radice unitaria e test di stazionarietà nelle differenze**

Come dimostrano i dati riportati nella Tabella 1.2, le serie nelle differenze sono stazionarie. In particolare gli esiti dei test ADF e PP, con un *p-value* sempre ≤ 0.01 , inducono a rifiutare l'ipotesi nulla di presenza di radici unitarie in tutti e tre i campioni. Inoltre il test KPSS non rigetta l'ipotesi nulla. Pertanto i risultati ottenuti nelle differenze forniscono un ulteriore supporto alle evidenze riscontrate nei livelli.

1.3 Cointegrazione e definizione della strategia di *trading*

1.3.1 Individuazione del periodo di operatività

Giunti a questo punto la definizione della strategia alternerà fasi di costruzione a fasi di *Back Testing* delle stesse. Questa è la formula più idonea per illustrare gli strumenti econometrici utilizzati, le modalità del loro utilizzo e i loro risultati. Per tale motivo questa prima fase è dedicata alla definizione di due periodi: il *formation period* (periodo di formazione) e il *trading period* (periodo di *trading*). Il primo è costituito da 252 giorni di quotazione⁷ e rappresenta l'arco temporale da cui sono stati estrapolati i dati utilizzati nella successiva implementazione del modello. Questa è la fase di analisi delle serie storiche in cui vengono raccolte le informazioni alla base del *trading* attuato nella fase successiva. Il *trading period*, invece, è formato dai 126 giorni di quotazione successivi al *formation period*. In questo arco temporale la strategia viene applicata sul mercato. Quindi, una volta analizzate le informazioni contenute nel periodo di formazione e definiti i parametri del modello, il periodo di *trading* costituisce la fase operativa dello studio. In altre parole un vero e proprio 'banco di prova' per la strategia. Nel periodo di *trading*,

⁷Un anno borsistico.

seguendo i segnali operativi forniti dalla strategia, vengono eseguite operazioni sul mercato, valutandone in seguito la fattibilità nonché i rendimenti e i rischi connessi. La Tabella 1.3 riassume il *trading period* e il *formation period* selezionati in ogni campione.

Campione	<i>formation period</i>	<i>trading period</i>
1	400 – 651	652 – 777
2	272 – 523	524 – 649
3	896 – 1147	1148 – 1273

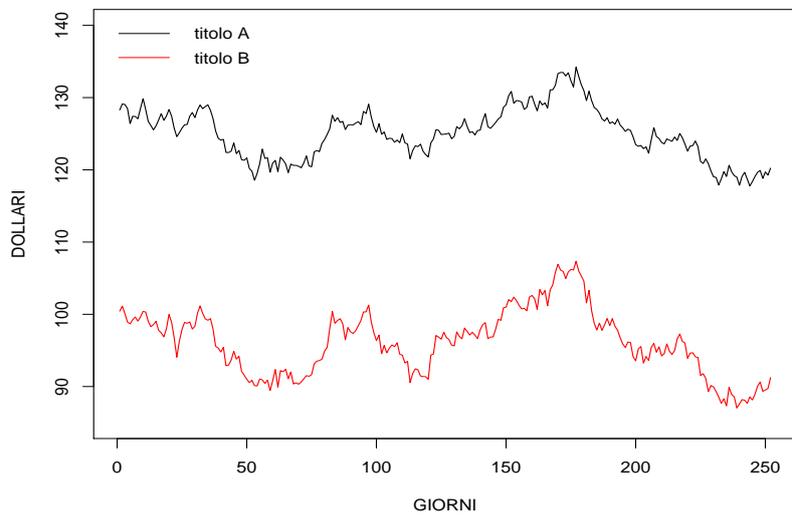
Tabella 1.3: Intervalli di osservazioni considerati in ogni campione

1.3.2 Verifica dell'ipotesi di integrazione e identificazione dell'ordine del ritardo

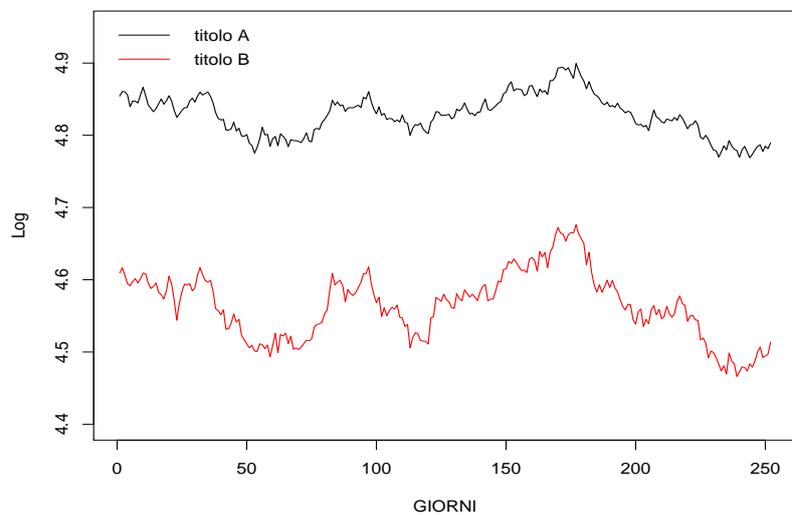
Individuati questi intervalli, l'ultimo giorno di quotazione del *formation period*, ossia il giorno precedente all'inizio del periodo di *trading*, può esser considerato la data di definizione della strategia. Le Figure da 1.4 a 1.6 illustrano l'andamento delle serie dei prezzi e del logaritmo dei prezzi nel *formation period* di ciascun campione. La non stazionarietà di ciascuna delle serie contenute nella Figura 1.4, nella Figura 1.5 e nella Figura 1.6 risulta evidente fin da una prima analisi visiva. I test ADF, PP e KPSS sono stati effettuati sulle serie storiche dei tre *formation period* nei livelli e nelle differenze. Come mostrato nelle Tabelle 1.4 e 1.5, l'ipotesi della presenza di radici unitarie trova ampia conferma negli esiti dei test. Pertanto i tre *subset* in questione mantengono le proprietà di integrazione presenti nel campione dal quale sono stati estratti.

(a) Prezzi				(b) Logaritmo naturale dei prezzi			
Titolo	ADF	PP	KPSS	Titolo	ADF	PP	KPSS
<i>p-value</i>				<i>p-value</i>			
A	0.73	0.69	0.02	A	0.73	0.70	0.02
B	0.57	0.67	0.02	B	0.56	0.67	0.02
C	0.42	0.44	≤ 0.01	C	0.54	0.17	≤ 0.01
D	0.44	0.42	≤ 0.01	D	0.67	0.65	≤ 0.01
<i>future</i>	0.30	0.67	≤ 0.01	<i>future</i>	0.26	0.66	≤ 0.01
<i>ETF</i>	0.34	0.69	≤ 0.01	<i>ETF</i>	0.30	0.67	≤ 0.01

Tabella 1.4: Test per radice unitaria e test di stazionarietà per il *formation period* dei tre campioni

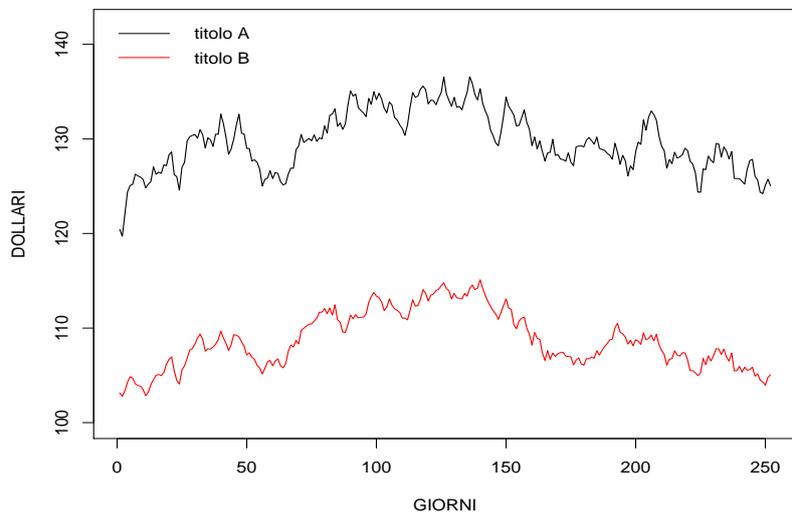


(a) Prezzo

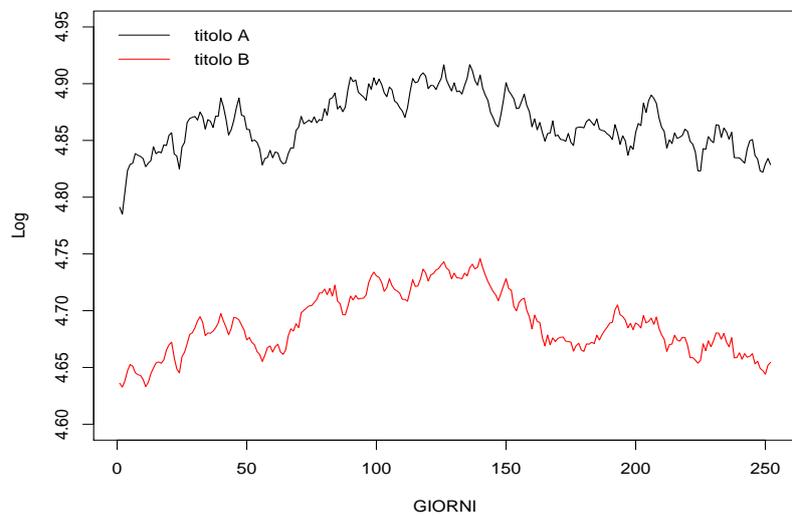


(b) Logaritmo naturale del prezzo

Figura 1.4: ***Formation period*** del campione 1

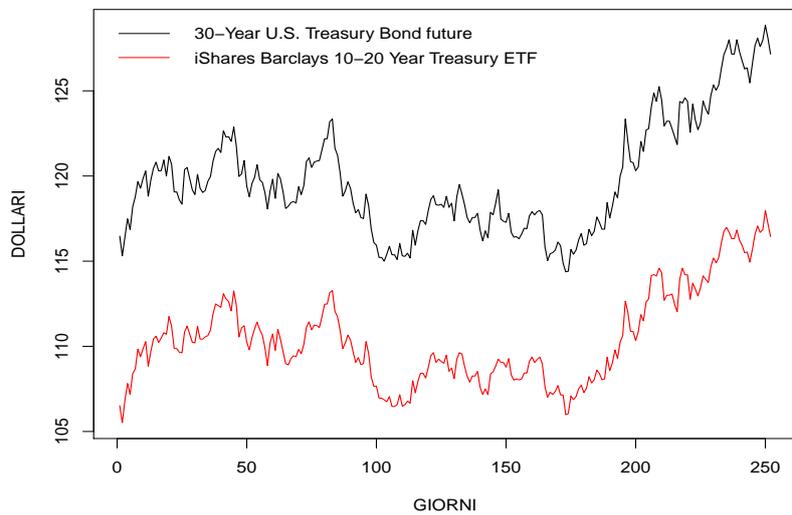


(a) Prezzo

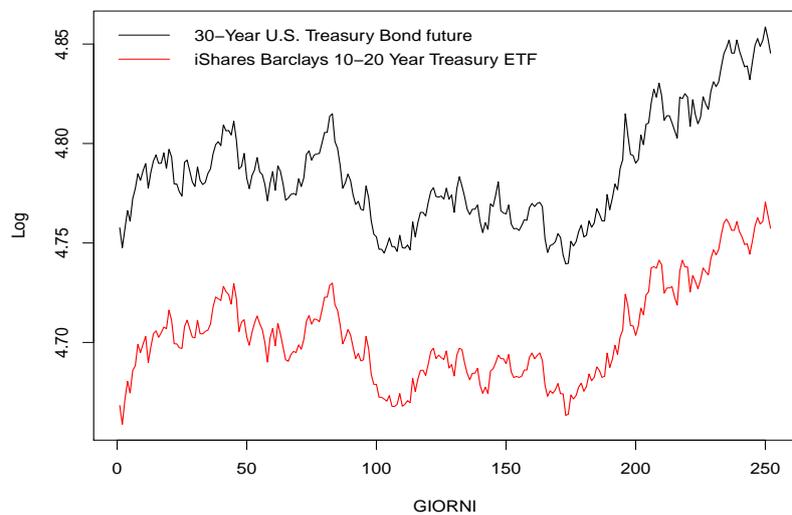


(b) Logaritmo naturale del prezzo

Figura 1.5: *Formation period* del campione 2



(a) Prezzo



(b) Logaritmo naturale del prezzo

Figura 1.6: **Formation period** del campione 3

(a) Prezzi				(b) Logaritmo naturale dei prezzi			
Titolo	ADF	PP	KPSS	Titolo	ADF	PP	KPSS
<i>p-value</i>				<i>p-value</i>			
A	≤ 0.01	≤ 0.01	0.10	A	≤ 0.01	≤ 0.01	0.10
B	≤ 0.01	≤ 0.01	0.10	B	≤ 0.01	≤ 0.01	0.10
C	≤ 0.01	≤ 0.01	0.10	C	≤ 0.01	≤ 0.01	0.10
D	≤ 0.01	≤ 0.01	0.10	D	≤ 0.01	≤ 0.01	0.10
<i>future</i>	≤ 0.01	≤ 0.01	0.10	<i>future</i>	≤ 0.01	≤ 0.01	0.10
<i>ETF</i>	≤ 0.01	≤ 0.01	0.10	<i>ETF</i>	≤ 0.01	≤ 0.01	0.10

Tabella 1.5: **Test per radice unitaria e test di stazionarietà per il *formation period* dei tre campioni nelle differenze**

Passiamo adesso all'identificazione dell'ordine di ritardo utilizzato nello studio delle relazioni di cointegrazione. A tale scopo ci siamo serviti dei cosiddetti *information criterion*: Akaike (AIC), Hannan-Queen (HQ), Schwarz (SC) e il *final prediction error* (FPE)⁸. Mediante questi criteri abbiamo indagato il numero di variabili *lagged* necessarie alla migliore specificazione di un modello di tipo VAR⁹. Nella Tabella 1.6 sono riassunti i risultati dei test AIC, HQ e SC effettuati sui periodi di formazione dei tre campioni.

Campione	AIC	HQ	SC
1	1	1	1
2	1	1	1
3	2	2	2

Tabella 1.6: ***Information criterion* applicati sui *formation period* dei tre campioni**

I valori contenuti nella Tabella 1.6 sono necessari per l'identificazione del ritardo della parte autoregressiva del modello VECM¹⁰ utilizzato per individuare le relazioni di cointegrazione tra le serie storiche considerate. E' facile dimostrare come nel passaggio da una rappresentazione di tipo VAR ad una di tipo VECM, l'ordine del ritardo della componente autoregressiva diminuisca di una unità¹¹. Per tale motivo è utile fare queste considerazioni: riprendendo i dati contenuti nella Tabella 1.6, qualora decidessimo di rappresentare le serie storiche dei titoli A - B e dei titoli C - D mediante un ritardo di ordine superiore, per esempio 2, non incapperemmo

⁸L'informazione fornita da quest'ultimo criterion verrà tralasciata poiché si può dimostrare che al crescere della numerosità campionaria i risultati provenienti da quest'ultimo tenderanno a coincidere con quelli dell'Akaike.

⁹Vectorial autoregressive model.

¹⁰Modello vettoriale a correzione dell'errore, in merito si veda l'Appendice Teorica

¹¹Per la dimostrazione si veda l'Appendice Teorica A.2.2.

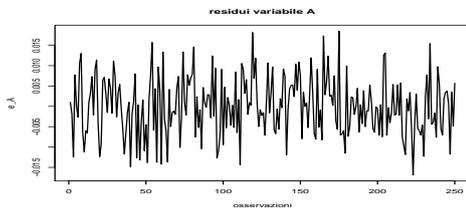
in errori di specificazione. Infatti aumentando di una unità l'ordine del modello autoregressivo rispetto ai valori suggeriti dai criteri di informazione, non avremmo distorsioni nella specificazione, ma solo un modello troppo parsimonioso (e in parte meno consistente). Sulla base di questa assunzione, per le serie contenute nei *formation period* del campione 1 e del campione 2, l'ordine del ritardo scelto è stato incrementato di una unità, portandolo a 2. Per quanto riguarda il campione 3, si è optato per un ritardo di ordine 2. Per poter dimostrare l'affidabilità delle stime dei modelli VAR ottenute con questi ritardi, è stato necessario condurre un'analisi sui residui dei tre modelli. I test effettuati sui residui hanno verificato che questi siano incorrelati e omoschedastici. A tale scopo sono stati effettuati il Portmanteau test (PT) e il Breusch-Godfrey LM test (BG)¹², mentre per verificare la presenza di eteroschedasticità è stato effettuato l'Arch LM test. I risultati ottenuti, riportati nella Tabella 1.7, ci permettono di accettare l'ipotesi nulla di assenza di autocorrelazione dei test PT e BG e di non rigettare l'ipotesi di omoschedasticità dell'Arch test. Unica eccezione è costituita dall'esito dell'Arch test effettuato sui residui del VAR relativo al terzo campione. In questo caso rigettiamo l'ipotesi nulla di assenza di eteroschedasticità. Tuttavia questa rappresenta una imperfezione tollerabile e dovuta, probabilmente, alla elevata quantità di 'noise' caratteristica dei dati empirici utilizzati nel terzo campione. Inoltre gli esiti di queste statistiche trovano supporto nei grafici contenuti nelle Figure 1.7, 1.8 e 1.9. Infatti osservando i *plot* dei residui dei tre modelli stimati, questi sono molto simili al tipico grafico di un *white noise*, mostrando relazioni di dipendenza di entità minima. Da questi grafici, l'analisi visiva dell'ACF e della PACF dei residui ci rassicura sulla corretta specificazione del modello. Le funzioni di autocorrelazione e di autocorrelazione parziale infatti sono prive di evidenze significative. La presenza di rotture delle bande di confidenza, è sporadica e ridotta. Inoltre a conferma della corretta specificazione, l'ACF e la PACF dei residui al quadrato evidenziano una presenza di effetti Arch caratterizzata da una scarsa significatività.

L'analisi sulla corretta specificazione dei modelli condotta fin'ora è lo strumento che abbiamo utilizzato per giungere all'individuazione dell'ordine di ritardo ottimale per il *formation period* di ogni campione. I risultati fin qui descritti supportano la scelta dei valori contenuti nella Tabella 1.7, utilizzati nel proseguo del lavoro.

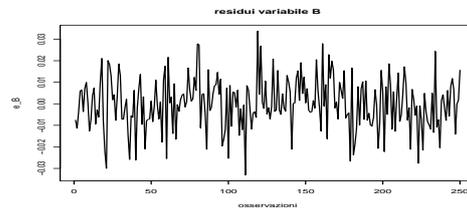
¹²Il Breusch-Godfrey test è un test di autocorrelazione degli errori di un modello di regressione. Questo fa uso dei residui del modello analizzandoli mediante una regressione. L'ipotesi nulla è l'assenza di correlazione seriale per ogni ordine fino a quello prescelto. La visualizzazione delle funzioni di autocorrelazione, di autocorrelazione parziale e i risultati del test Breusch-Godfrey, ci consentono di indagare la presenza di strutture lineari di dipendenza tra i residui. Il Portmanteau test ci permette di individuare la presenza strutture di dipendenza non lineare nella varianza. L'ipotesi nulla è l'assenza di correlazione seriale dovuta a forme di dipendenza non lineare nella varianza

<i>Titoli</i>	<i>PT</i>	<i>BG</i>	<i>ARCH test</i>
	<i>p-value</i>		
<i>A - B</i>	0.78	0.73	0.55
<i>C - D</i>	0.17	0.20	0.97
<i>future - ETF</i>	0.69	0.51	0.00025

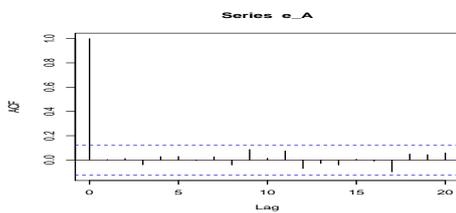
Tabella 1.7: Risultati dei test effettuati sui residui dei VAR stimati



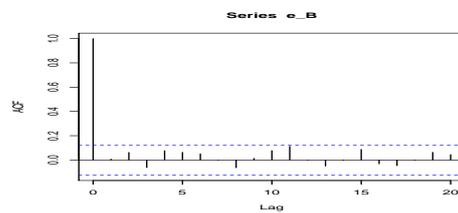
(a) Residui della variabile A



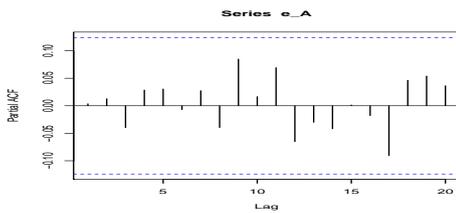
(b) Residui della variabile B



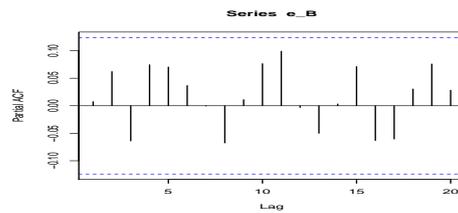
(c) ACF dei residui della variabile A



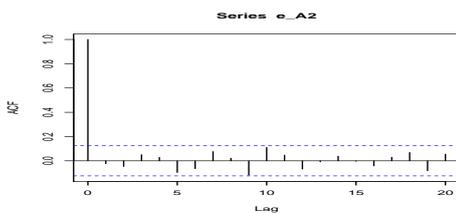
(d) ACF dei residui della variabile B



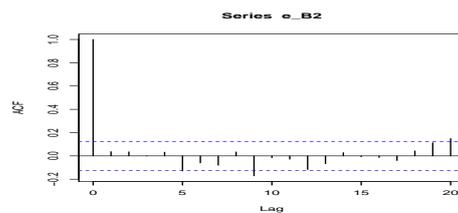
(e) PACF dei residui della variabile A



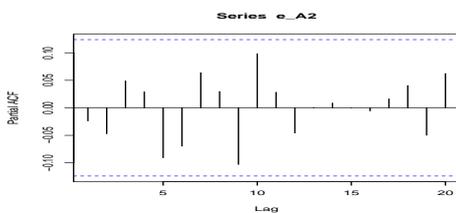
(f) PACF dei residui della variabile B



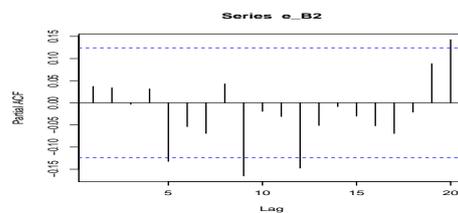
(g) ACF dei residui al quadrato della variabile A



(h) ACF dei residui al quadrato della variabile B

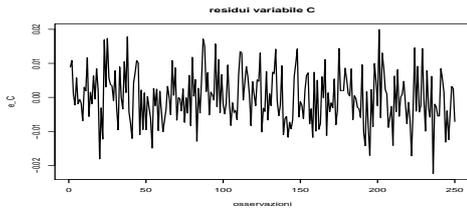


(i) PACF dei residui al quadrato della variabile A

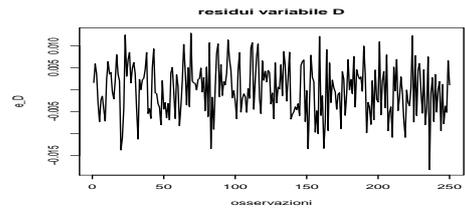


(j) PACF dei residui al quadrato della variabile B

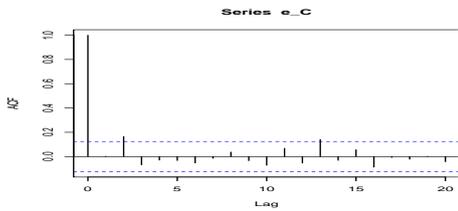
Figura 1.7: Analisi della ACF e della PACF dei residui del VAR relativo ai titoli A -B



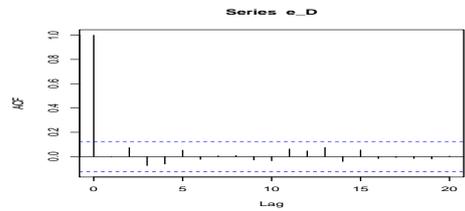
(a) Residui della variabile C



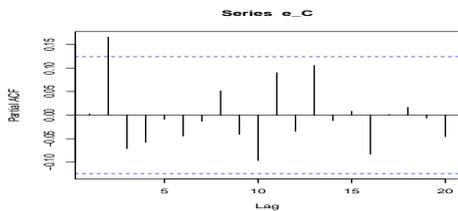
(b) Residui della variabile D



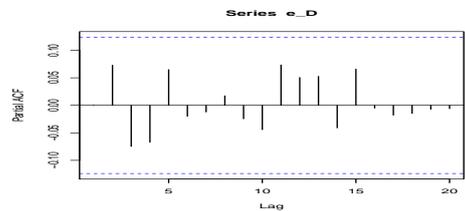
(c) ACF dei residui della variabile C



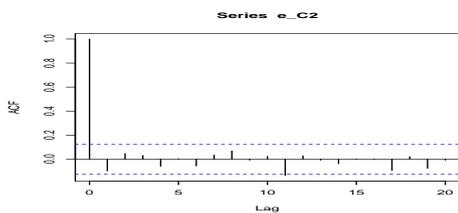
(d) ACF dei residui della variabile D



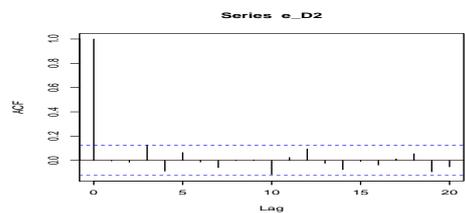
(e) PACF dei residui della variabile C



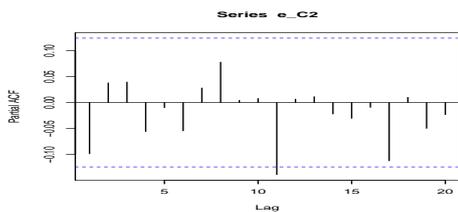
(f) PACF dei residui della variabile D



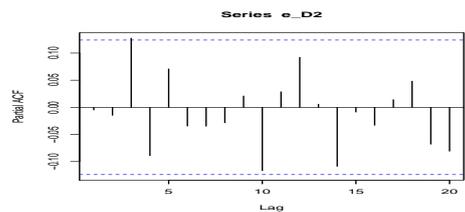
(g) ACF dei residui al quadrato della variabile C



(h) ACF dei residui al quadrato della variabile D

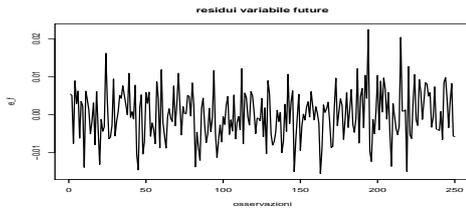


(i) PACF dei residui al quadrato della variabile C

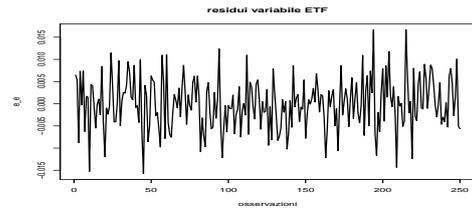


(j) PACF dei residui al quadrato della variabile D

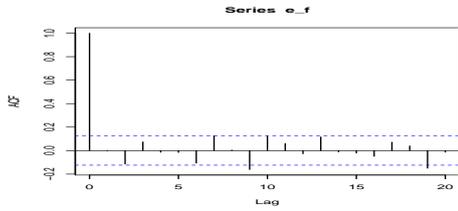
Figura 1.8: Analisi della ACF e della PACF dei residui del VAR relativo ai titoli C - D



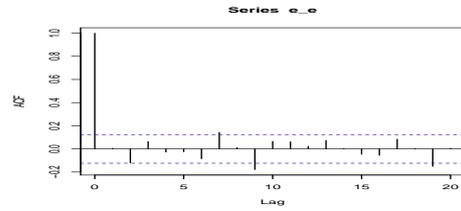
(a) Residui della variabile F



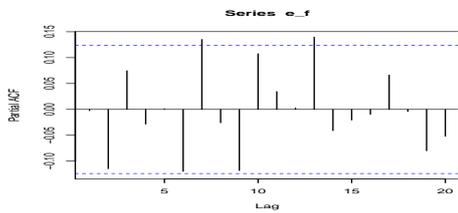
(b) Residui della variabile E



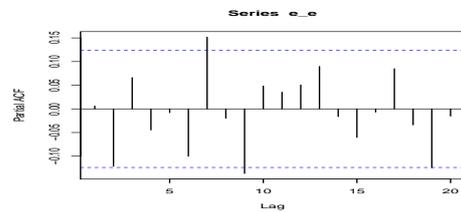
(c) ACF dei residui della variabile F



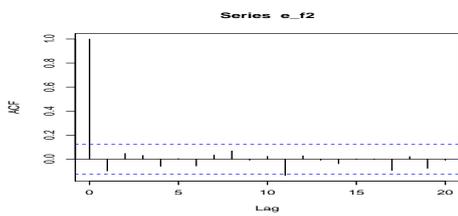
(d) ACF dei residui della variabile E



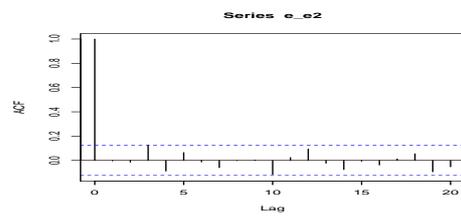
(e) PACF dei residui della variabile F



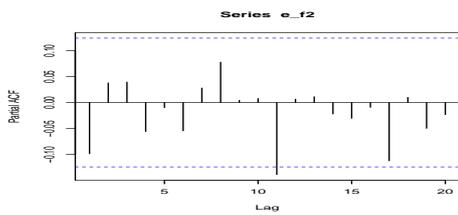
(f) PACF dei residui della variabile E



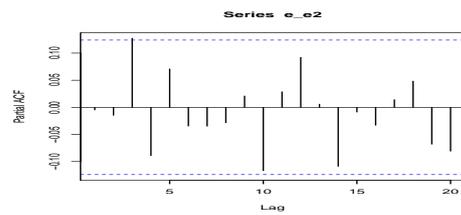
(g) ACF dei residui al quadrato della variabile F



(h) ACF dei residui al quadrato della variabile E



(i) PACF dei residui al quadrato della variabile F



(j) PACF dei residui al quadrato della variabile E

Figura 1.9: Analisi della ACF e della PACF dei residui del VAR relativo al *future* e all'ETF

1.3.3 Analisi delle relazioni di cointegrazione e stazionarizzazione del processo

Dopo aver individuato l'ordine del ritardo più appropriato possiamo procedere all'individuazione delle relazioni di cointegrazione. A tal proposito consideriamo un processo bivariato integrato di ordine 1 del tipo $\mathbf{y}_t = (y_{1t}, y_{2t})$. Supponiamo che sia cointegrato, con vettore di cointegrazione $\beta = [1, -\alpha]$, tale che $\beta' \mathbf{y}_t = y_{1t} - \alpha y_{2t}$ è $I(0)$ ¹³. Si dimostra tramite il famoso Teorema di Rappresentazione di Granger¹⁴ che: *se le serie sono cointegrate allora esiste una rappresentazione in forma ECM; e se esiste una forma ECM e le serie sono integrate allora sono cointegrate*. Sfruttando il legame definito da questo teorema tra le serie cointegrate e la loro rappresentazione in ECM, ci siamo serviti di quest'ultima per individuare le relazioni di cointegrazione tra le serie prese in esame.

Test di Johansen: caratteristiche della statistica

Prima di procedere nello studio delle relazioni di cointegrazione, è necessario effettuare un piccolo *excursus* teorico utile a ricongiungere la teoria alla pratica. Allora iniziamo considerando l'equazione del VAR(p) che abbiamo utilizzato.

Questa presenta una componente di trend:

$$\mathbf{y}_t = K(t) + \sum_{i=1}^p \phi_i \mathbf{y}_{t-i} + \mathbf{a}_t \quad (1.9)$$

¹³La teoria quantitativa della moneta ci fornisce un'ottimo esempio di applicazione in campo economico: *Secondo tale teoria (semplificata), esiste un rapporto stabile fra la quantità di moneta presente in un sistema e il valore delle transazioni effettuate; tale proporzione è detta velocità di circolazione. In formule:*

$$MV = PY$$

Oppure si può dire che la velocità di circolazione è data dal rapporto fra PIL e quantità reale di moneta (M/P), si ha quindi:

$$V = Y/M^*$$

dove M^ è la quantità di moneta reale. Considerando i logaritmi di tali grandezze si ottiene:*

$$v = y - m$$

dove y è il logaritmo del PIL e m è il logaritmo della moneta reale. Una interpretazione integralista di tale teoria condurrebbe ad affermare che v è una quantità fissa. Una meno estrema considererebbe v una quantità che osservata nel tempo, presenta fluttuazioni più o meno persistenti attorno ad un valore centrale di equilibrio. La serie storica di v_t somiglia pertanto alla realizzazione di un processo stazionario che, nonostante nel breve periodo possa presentare scostamenti, tenderà comunque a mantenere un equilibrio. Allora se y_t e m_t sono rappresentabili come processi $I(1)$, sostenere la teoria quantitativa della moneta equivale, più o meno, ad affermare che le serie y_t e m_t cointegrato, e che il vettore di cointegrazione che le lega $[1, -1]$.

¹⁴Si tratta di un teorema abbastanza complesso, la cui dimostrazione esula dagli obbiettivi di questo elaborato. Per un approfondimento più analitico si veda Engle e Granger (1987).

dove $K(t) = k_0 + k_1 t$, k_i sono vettori costanti e \mathbf{a}_t è una sequenza di vettori casuali Gaussiani *i.i.d.* con media zero e matrice di covarianza $\text{Var}(\mathbf{a}_t) = \Sigma$ definita positiva. La sua rappresentazione ECM è uguale a:

$$\Delta \mathbf{y}_t = K(t) + \Pi \mathbf{y}_{t-1} + \sum_{i=1}^{p-1} \phi_i^* \Delta \mathbf{y}_{t-i} + \mathbf{a}_t \quad (1.10)$$

dove $\Pi = -\phi(1)$, $\phi_i^* = -\sum_{j=i+1}^p \phi_j$ per $i = (1, \dots, p-1)$, e $\Delta \mathbf{y}_t = \mathbf{y}_t - \mathbf{y}_{t-1}$. Per identificare il rango della matrice Π abbiamo utilizzato il test di Johansen. La sua statistica test è una χ^2 . Tuttavia come sottolineato dallo stesso Johansen, è necessario specificare correttamente alcuni elementi caratterizzanti del test prima di applicarlo. In caso contrario, un'errata specificazione degli input può portare a risultati distorti. In base agli input scelti, la distribuzione χ^2 di riferimento cambierà poiché cambieranno i gradi di libertà della statistica. Diventa fondamentale in tal senso la specificazione della componente di trend $K(t)$. In particolare è possibile individuare cinque casi:

- **Caso 1:** $K(t) = 0$

In questo caso, nel modello ECM dell'Eq. 1.10, non è presente un termine costante; quindi le componenti di \mathbf{y}_t sono processi $I(1)$ senza *drift* e $\mathbf{w}_t = \beta' \mathbf{y}_t$ ha media uguale a zero.

- **Caso 2:** $k_0 = \alpha \mathbf{c}_0$ e $k_1 = 0$

Questo è il caso di una costante soggetta a restrizioni. La forma ECM diventa:

$$\Delta \mathbf{y}_t = \alpha(\beta' \mathbf{y}_{t-1} + \mathbf{c}_0) + \sum_{i=1}^{p-1} \phi_i^* \Delta \mathbf{y}_{t-i} + \mathbf{a}_t \quad (1.11)$$

In questo caso le componenti di \mathbf{y}_t sono serie $I(1)$ senza *drift* e le serie cointegranti \mathbf{w}_t hanno media pari a \mathbf{c}_0 , diversa da zero.

- **Caso 3:** $k_1 = 0$ e k_0 è senza restrizioni.

Questo è il caso della costante senza restrizioni. La forma ECM diventa:

$$\Delta \mathbf{y}_t = \mathbf{k}_0 + \alpha \beta' \mathbf{y}_{t-1} + \sum_{i=1}^{p-1} \phi_i^* \Delta \mathbf{y}_{t-i} + \mathbf{a}_t \quad (1.12)$$

dove le componenti di \mathbf{y}_t sono serie $I(1)$ con *drift* \mathbf{k}_0 e \mathbf{w}_t può avere una media diversa da zero.

- **Caso 4:** $K(t) = k_0 + \alpha c_1 t$

Questo è il caso di un trend deterministico. In questo caso la forma ECM

diventa:

$$\Delta \mathbf{y}_t = k_0 + \alpha(\beta' \mathbf{y}_{t-1} + c_1 t) + \sum_{i=1}^{p-1} \phi_i^* \Delta \mathbf{y}_{t-i} + \mathbf{a}_t \quad (1.13)$$

Le componenti di \mathbf{y}_t sono serie $I(1)$ con *drift* k_0 e la serie cointegrante \mathbf{w}_t ha un termine di trend lineare $c_1 t$.

- **Caso 5:** $K(t) = k_0 + k_1 t$

In questo caso le serie che compongono y_t sono $I(1)$ con un trend quadratico e w_t ha un trend lineare.

Questi cinque casi sono legati dalla seguente relazione di inclusione:

$$\text{caso 1} \subset \text{caso 2} \subset \text{caso 3} \subset \text{caso 4} \subset \text{caso 5} .$$

Quest'ultima assume una importanza pratica fondamentale per il nostro lavoro. Tale relazione infatti, definisce il tipo di distribuzione da utilizzare in base alla scelta effettuata sulla costante e sulla componente di trend. Queste scelte devono esser coerenti con le caratteristiche delle serie presenti nel *formation period* di ogni campione. Soltanto dopo aver scelto tra i cinque casi presentati, quello che più si avvicina alle caratteristiche della serie che si sta analizzando, possiamo costruire la statistica test. Nel caso 3 e nel caso 5 la statistica si distribuisce come una χ_{n-r}^2 , dove n è il numero delle componenti del sistema e r è il rango della matrice $\mathbf{\Pi}$. Negli altri casi la distribuzione è quella di una χ_r^2 , quindi con r gradi di libertà.

Campione 1

Le serie del *formation period* dei titoli A e B, contenute nella matrice \mathbf{Y}^{1F} di dimensioni (252×2) , sono state sottoposte al test di Johansen, nella sua versione del massimo autovalore¹⁵ e sotto l'ipotesi della presenza di una costante. In questo caso la scelta di inserire una costante nella stima è dovuta alla conoscenza del processo che ha generato le serie (Eq. 1.2) ed ad un'analisi delle caratteristiche del grafico (Figura 1.4). Quest'ultimo mostra una dinamica simile ad una passeggiata aleatoria con deriva positiva. Pertanto l'equazione utilizzata per la stima è riconducibile a quella del caso 3 (Eq. 1.12). I risultati sono riportati nella Tabella 1.8. Il valore assunto dalla statistica in corrispondenza dell'ipotesi di assenza di cointegrazione è pari a 66.57, maggiore del valore critico 20.20 all' 1% di significatività. Ciò implica che si può ragionevolmente rifiutare l'ipotesi di assenza di relazioni di

¹⁵Tutta la parte in merito alle diverse tipologie del test e alle ipotesi sottostanti trova una esaustiva trattazione nell'Appendice Teorica del lavoro.

cointegrazione. Nel caso dell'ipotesi relativa alla presenza di almeno una relazione di cointegrazione, il valore assunto dalla statistica è pari a 3.49, inferiore al valore critico 7.52. Quest'ultimo risultato ci consente di accettare l'ipotesi della presenza di una relazione di cointegrazione ad un livello di significatività del 90%.

H_0	$test$	10%	5%	1%
$r \leq 1$	3.49	7.52	9.24	12.97
$r = 0$	66.57	13.75	15.67	20.20

Tabella 1.8: **Risultati del test di Johansen per la coppia di titoli A - B**

Il test di Johansen individua il vettore di cointegrazione $\beta^1 = [1, -0.613, -2.028]$. Questo vettore risulta normalizzato rispetto la prima componente¹⁶. Un metodo pratico per valutarne la bontà della stima è studiarne la capacità di stazionizzare le serie del *formation period* da cui è stato stimato. Infatti se β^1 è il vettore di cointegrazione si avrà che $\beta^{1'}y_t \sim I(0)$. Questa tecnica consiste nel premoltiplicare il vettore $\beta^{1'}$ per la matrice Ω_1 di dimensione (3×252) , composta per la prima e la seconda riga dalla prima e dalla seconda colonna della matrice \mathbf{Y}^{1F} , e per la terza riga da un vettore di 1.

$$\begin{bmatrix} 1 & -0.613 & -2.028 \end{bmatrix} \times \begin{bmatrix} 4.854 & 4.860 & \dots & 4.789 \\ 4.609 & 4.616 & \dots & 4.513 \\ 1 & 1 & \dots & 1 \end{bmatrix} \quad (1.14)$$

Da questa operazione abbiamo ottenuto un vettore (252×1) contenente i valori del processo denominato S^{1F} . Il grafico contenuto nella Figura 1.10(a) ne evidenzia da subito la stazionarietà. Il processo S^{1F} infatti, presenta oscillazioni continue intorno al suo valore medio mostrando una *mean reversion*, tipica dei processi stazionari. Su questo inoltre, sono stati effettuati i test di radice unitaria e di stazionarietà. L'ordine del ritardo per l'applicazione di queste statistiche è stato identificato mediante il criterio fornito da Schwert (1989) secondo cui l'ordine massimo può essere individuato mediante la formula

$$p_{max} = \left\lceil 12 \left(\frac{T}{100} \right)^{\frac{1}{4}} \right\rceil \quad (1.15)$$

¹⁶A tal proposito dobbiamo sottolineare che il vettore di cointegrazione non è necessariamente unico. Infatti nel caso di un'unica relazione, il vettore associato moltiplicato per una qualsiasi costante λ è ancora un vettore di cointegrazione. Di solito, come in questo caso, la prima componente di β viene utilizzata per normalizzare l'intero vettore. In altre parole si pone $\lambda = \frac{1}{\beta_1}$, allora $\lambda\beta = [1, \frac{\beta_2}{\beta_1}, \dots, \frac{\beta_n}{\beta_1}]$ ottenendo un altro vettore cointegrante.

dove T rappresenta la numerosità del campione. I risultati contenuti nella prima riga della Tabella 1.9¹⁷ (a), (b) e (c) forniscono ulteriore supporto all'ipotesi di stazionarietà. La lettura congiunta di queste evidenze ci garantisce la capacità del vettore β di rendere stazionaria la combinazione lineare ottenuta tramite l'Eq. 1.14.

(a) ADF					
Processo	test	1%	5%	10%	
S^{1F}	-8.54	-3.44	-2.87	-2.57	
S^{2F}	-5.76	-3.44	-2.87	-2.57	
S^{3F}	-4.06	-3.44	-2.87	-2.57	

(b) PP		
Processo	test	Z-stat.
S^{1F}	-103.27	0.10
S^{2F}	-50.52	-0.09
S^{3F}	-43.99	0.57

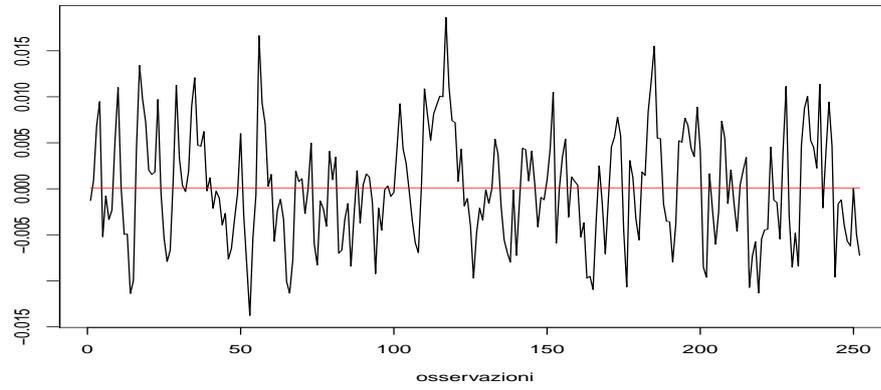
(c) KPSS					
Processo	test	10%	5%	2.5%	1%
S^{1F}	0.254	0.347	0.463	0.574	0.739
S^{2F}	0.079	0.347	0.463	0.574	0.739
S^{3F}	0.416	0.347	0.463	0.574	0.739

Tabella 1.9: **Test per radice unitaria e test di stazionarietà per i processi S^{1F} , S^{2F} e S^{3F}**

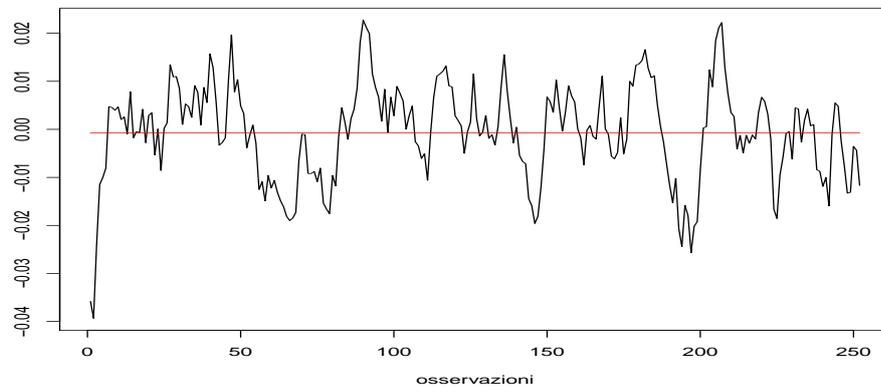
Campione 2

Ora consideriamo il periodo di formazione per la coppia di titoli C e D. I valori di questa finestra di osservazioni sono contenuti nella matrice \mathbf{Y}^{2F} di dimensioni (252×2) . Nonostante l'Eq. 1.8 contenga una componente di trend, il grafico della Figura 1.5 ne evidenzia una ridotta incidenza (per quella sezione). Il *random walk* u_t (Eq. 1.4), comune alle serie del titolo C e del titolo D, sembra avere un effetto dominante nella dinamica che genera i prezzi dei due asset. Per tale motivo, anche in questo caso, la forma del VECM stimato è riconducibile all'equazione del caso 3 (l'Eq. 1.12). Il test di Johansen effettuato sul vettore \mathbf{Y}^{2F} restituisce i

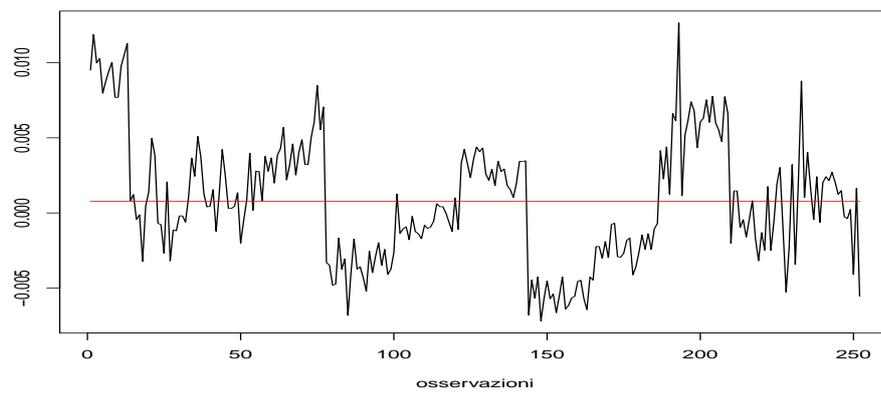
¹⁷Questa volta, per effettuare i test ADF, PP e KPSS è stata utilizzata una diversa funzione del software R. Come si nota, gli output presenti nella Tabella 1.9 sono interpretabili mediante un confronto tra i valori della statistica test e i valori critici, piuttosto che tramite *p-value*. Tale scelta ci permette di distinguere diversi gradi di accettazione.



(a) Processo S^{1F}



(b) Processo S^{2F}



(c) Processo S^{3F}

Figura 1.10: Grafico dei *formation period* stazionarizzati

risultati riportati nella Tabella 1.10. Il vettore di cointegrazione individuato è $\beta^2 = [1, -0.731, -1.437]$. Anche in questo caso, il vettore $\beta^{2'}$ è stato pre-moltiplicato come fatto nella Eq. 1.14, ottenendo $\beta^{2'}\Omega_2 = S^{2F}$. Questa volta la matrice Ω_2 , di dimensione (3×252) , è formata per le prime due righe dalle colonne del vettore \mathbf{Y}^{2F} e per la terza da un vettore di 1. Dalla Figura 1.10 (b) è evidente come il grafico del processo S^{2F} presenti un andamento più o meno simile a quello di una serie stazionaria. Questa ipotesi trova supporto nei risultati dei test di radice unitaria e di stazionarietà effettuati sul processo S^{2F} , contenuti nella seconda riga della Tabella 1.9 (a), (b) e (c). Quindi anche in questo caso il vettore di cointegrazione β^2 , riesce a rendere stazionaria la combinazione ottenuta mediante l'Eq. 1.14. In altre parole è verificata la condizione per cui $\beta^{2'}y_t \sim I(0)$.

H_0	<i>test</i>	10%	5%	1%
$r \leq 1$	3.65	7.52	9.24	12.97
$r = 0$	31.94	13.75	15.67	20.20

Tabella 1.10: **Risultati del test di Johansen per la coppia di titoli C - D**

Campione 3

Nel caso delle serie del future e dell'ETF, la Figura 1.6 motra la presenza, più o meno incidente, di una componente di trend. Tuttavia tale presenza sembra caratterizzare solo una parte delle osservazioni. Inoltre le oscillazioni attorno all'ipotetico trend fanno pensare ad una forma di non stazionarietà tipica di un processo a radice unitaria con drift. Pertanto ad una dinamica differente rispetto a quella di una serie *trend stationary*, che giustificherebbe l'inclusione di una componente di trend nella stima del VECM. Per questo motivo, ancora una volta, la formula scelta per la stima del VECM coincide con l'Eq. 1.12. Consideriamo la matrice \mathbf{Y}^{3F} , di dimensioni (252×2) , contenente le serie del *formation period* dei due asset. Come mostrato nella Tabella 1.11, il test di Johansen sul vettore \mathbf{Y}^{3F} rigetta la nulla di assenza di cointegrazione ad un livello di significatività del 5%. Secondo una lettura poco ortodossa di questo output, rifiutiamo l'ipotesi nulla essendo questa vera, con una probabilità inferiore al 5% . Da un punto di vista inferenziale, tale risultato ci rende meno sicuri rispetto a quelli delle Tabelle 1.8 e 1.10. Inoltre il valore assunto dalla statistica, in corrispondenza dell'ipotesi relativa alla presenza di almeno una relazione di cointegrazione, è pari a 1.07. Essendo tale valore minore del valore critico pari a 7.52, accettiamo l'ipotesi nulla

ad un livello di significatività del 10%. Il vettore di cointegrazione associato è $\beta^3 = [1, -1.026, 0.049]$.

H_0	<i>test</i>	10%	5%	1%
$r \leq 1$	1.07	7.52	9.24	12.97
$r = 0$	18.08	13.75	15.67	20.20

Tabella 1.11: **Risultati del test di Johansen per la coppia di titoli f - E**

Anche in questo caso ripetiamo l'operazione dell'Eq. 1.14. Otteniamo così $\beta^3 \Omega_3 = S^{3F}$, dove la matrice Ω_3 , di dimensioni (3×252) , è costituita per le prime due righe dalle colonne del vettore \mathbf{Y}^{3F} , e per la terza da un vettore di 1. Il processo S^{3F} , derivante da tale operazione, è rappresentato nel grafico della Figura 1.10(c). I risultati dei test di radice unitaria e di stazionarietà effettuati sulla serie del processo S^{3F} sono illustrati nella terza riga della Tabella 1.9 (a), (b) e (c). Rispetto ai processi S^{1F} e S^{2F} , i risultati dei test e la visualizzazione grafica del processo ci rendono meno sicuri in merito alla sua stazionarietà. Il test ADF presenta un valore della statistica pari a -4.06 . Questo valore ci consente di rifiutare l'ipotesi della presenza di radici unitarie, tuttavia non si può fare a meno di notare la vicinanza al valore critico -3.44 . Questo indica che per il processo S^{3F} , il rifiuto dell'ipotesi di radice unitaria non è così netto come accaduto per il processo S^{1F} . Inoltre il risultato del test PP rifiuta l'ipotesi di non stazionarietà. In questo caso il confronto tra il valore della statistica e il valore critico ci permettono un rifiuto più netto. Per quanto riguarda i risultati del test KPSS, ricordiamo che quest'ultimo test è 'a coda destra', quindi si rifiuta l'ipotesi nulla quando la statistica test è maggiore della soglia critica stabilita. Nel caso in esame si ha che la statistica test (0.416) è maggiore della soglia critica al 10% (0.347) e pertanto l'ipotesi di stazionarietà è da rigettare. Anche se l'esito del test KPSS non smentisce in modo secco quelli dei test ADF e PP, questi risultati, se messi a confronto, non possono essere ritenuti robusti. Questa rappresenta una fattispecie che tende a riproporsi durante l'implementazione di queste statistiche, soprattutto se applicate a dati (empirici) di natura finanziaria. In questa situazione, valori delle statistiche altamente significativi lasciano spazio a dei valori che vanno interpretati ricercando coerenza nel confronto con gli esiti di altri test. Pertanto la lettura congiunta dei risultati appena descritti conduce ad una valutazione soggettiva sulla corretta identificazione del vettore β^3 . Ulteriori spiegazioni in merito a tale valutazioni sono fornite nella Sez. 1.5.2.

1.4 Definizione della strategia di *trading*

Una volta identificata la relazione di cointegrazione che lega l'andamento del prezzo di due titoli, procediamo alla definizione della strategia di investimento. L'idea alla base di tale strategia può essere sintetizzata da quanto segue. Consideriamo le serie storiche del logaritmo naturale del prezzo di due titoli x_1 e x_2 , entrambe $I(1)$ e caratterizzate da una significativa relazione di cointegrazione. Identificato il vettore cointegrante $\beta = [1, -\alpha]$ la combinazione lineare¹⁸

$$\beta' \mathbf{x} = \ln(P_{x_1}) - \alpha \ln(P_{x_2}) \quad (1.16)$$

è un processo stazionario che tenderà ad oscillare attorno al suo valore medio. A questo punto, possiamo intendere la combinazione $\ln(P_{x_1}) - \alpha \ln(P_{x_2})$ come un portafoglio costituito da una posizione lunga sul titolo x_1 e una posizione corta sulla proporzione α del titolo x_2 . Ipotizziamo inoltre di fissare due bande che rappresentino un limite superiore e un limite inferiore per le fluttuazioni della combinazione lineare $\beta' \mathbf{x}$. La rottura di una delle bande costituisce un segnale operativo del modello. La strategia infatti, consiste nell'acquistare tale portafoglio in corrispondenza di una rottura del limite inferiore, mantenendo le due posizioni che lo compongono fin quando il processo non è tornato al suo valore di equilibrio. Nel caso di rottura del limite superiore la strategia è speculare. Questo modello si basa sulla proprietà di *mean reversion*, tipica di un processo stazionario. Infatti la chiusura delle posizioni assunte al momento della rottura di uno dei due limiti, avviene in corrispondenza del ritorno al valore medio. L'operatività basata su questo modello richiede una stima preventiva di alcuni parametri. Questi forniscono regole decisionali importanti per il *trading* che si andrà ad effettuare. Di seguito sono descritte le tecniche per la stima di tali parametri necessari a completare l'articolazione del modello.

1.4.1 Stazionarizzazione del *trading period*

Il processo stazionario che fornisce i segnali operativi è ottenuto mediante la stessa operazione illustrata nell'Eq. 1.14. Tuttavia questa volta il vettore di cointegrazione (con o senza costante) è pre-moltiplicato per una matrice composta, per le prime due righe, dalle colonne della matrice contenente i valori del *trading period*. In altre parole, ipotizzando il vettore di cointegrazione β , il valore del processo S^T al tempo t è dato da $\beta' \mathbf{x}_t = \ln(P_{x_1 t}) - \alpha \ln(P_{x_2 t})$ con $t \in \textit{trading period}$. Que-

¹⁸Nel caso in cui il vettore β includa un valore c costante allora la combinazione lineare diventa $\beta' \mathbf{x} = \ln(P_{x_1}) - \alpha \ln(P_{x_2}) + c$.

sta combinazione lineare corrisponde alla stazionarizzazione dei valori contenuti nel *trading period*, che al tempo t deve ancora concludersi. Come mostrato nella Eq. 1.17 il vettore di cointegrazione, per ogni nuova osservazione del periodo di *trading*, deve restituire una combinazione lineare il cui valore S_t^T , unito a quelli precedenti, costituisce una serie storica stazionaria.

$$\begin{array}{rcl}
 t = 1 & \beta \mathbf{x}_1 = & \ln(P_{x_11}) - \alpha \ln(P_{x_21}) \\
 t = 2 & \beta \mathbf{x}_2 = & \ln(P_{x_12}) - \alpha \ln(P_{x_22}) \\
 \dots & \dots & \dots \\
 \dots & \dots & \dots \\
 t = 126 & \beta \mathbf{x}_{126} = & \ln(P_{x_1126}) - \alpha \ln(P_{x_2126})
 \end{array} \left. \vphantom{\begin{array}{rcl} t = 1 \\ t = 2 \\ \dots \\ \dots \\ t = 126 \end{array}} \right\} \begin{array}{l} S_1^T \\ S_2^T \\ \dots \\ \dots \\ S_{126}^T \end{array} \quad (1.17)$$

I fattori che possono incidere sull'efficacia della stazionarizzazione del *trading period* (ossia sulla stazionarietà del processo S_t^T), ottenuta mediante l'Eq. 1.17, hanno una duplice natura. Da un lato è necessaria l'individuazione del vero vettore di cointegrazione, infatti una stima distorta conduce ad uno scarso risultato in termini di stazionarietà. Dall'altro la condizione necessaria è legata alla stabilità della relazione di cointegrazione. A questo secondo fattore è dedicata una sezione apposita della trattazione che segue (Sez. 1.5.2). Tuttavia anticipiamo che durante il *trading period*, il test di Johansen viene ripetuto mediante una *rolling window*¹⁹ di 252 osservazioni. Tale finestra temporale avanza di giorno in giorno, coerentemente con il trascorrere del tempo. Al manifestarsi di un segnale operativo generato dalla rottura di una delle due bande, l'apertura della posizione è subordinata al confronto tra il vettore di cointegrazione utilizzato e quello individuato dalla *rolling window*. Nel caso in cui la differenza tra le rispettive componenti superi una certa soglia, l'operazione non viene eseguita. Questo criterio, pur se poco ortodosso, ci tutela da cambiamenti incisivi nella relazione di cointegrazione.

¹⁹La *rolling analysis* delle serie storiche è spesso utilizzata per valutare la stabilità di un modello statistico attraverso il tempo. La maggiore applicazione è nell'analisi dei dati derivanti da serie storiche finanziarie, nel caso in cui una delle assunzioni chiave è che i parametri del modello statistico siano costanti nel tempo. Una tecnica comune per valutare la costanza dei parametri di un modello è quella di calcolare le stime dei parametri su una *rolling window* che abbia dimensione fissa, muovendosi sull'intero campione. Se i parametri fossero realmente costanti allora le stime ottenute nelle diverse finestre temporali non dovrebbero presentare grandi differenze. Al contrario, evidenze in merito all'instabilità dovuta ad eventuali cambiamenti in un punto interno al campione, dovrebbero essere catturate da queste stime *rolling*.

1.4.2 Costruzione di due *threshold*

Come accennato nella prima parte del paragrafo, questo modello genera un segnale operativo nell'istante in cui il processo S_t^T oltrepassa una delle due soglie relative al limite superiore o al limite inferiore. Si consideri Δ il valore assoluto che individua la distanza tra una delle soglie e il valore medio attorno al quale il processo oscilla. La Figura 1.11 rappresenta un caso in cui Δ è pari a 2. Come si vede, la soglia relativa al limite inferiore (colore blu) dista dal valore medio (colore nero) una misura che in valore assoluto è pari a 2. Illustriamo due approcci che possono essere utilizzati per individuare la misura Δ :

1. Approccio dei percentili

Consideriamo il processo ottenuto stazionarizzando il *formation period* dei due titoli (esempio il processo S^{1F} ricavato mediante l'Eq. 1.14). Dopo aver costruito la funzione di ripartizione delle realizzazioni che lo compongono, selezioniamo il 5° e il 95° percentile. Il valore corrispondente al 95° percentile identifica il limite superiore mentre quello relativo 5° il limite inferiore.

2. Approccio della deviazione standard

Anche in questo caso consideriamo il processo S^{1F} . Dopo aver calcolato la deviazione standard delle realizzazioni che lo compongono, moltiplichiamo per 2 il valore ottenuto. Il risultato ottenuto coincide con il Δ fissato in questo caso. In altre parole la soglia superiore e quella inferiore distano dal valore medio una misura pari a due deviazioni standard.

I due approcci sono abbastanza simili, infatti la frequenza che la funzione di ripartizione associa a valori estremi e indirettamente collegata alla deviazione standard del processo che li genera. Nella fase di *Back Testing* del modello è stata scelta una versione avanzata dell'approccio basato sulla deviazione standard. Nel nostro caso infatti, la deviazione standard misurata nel *formation period* si aggiorna in seguito ad ogni osservazione del periodo di *trading*. Questa versione rolling non genera grosse variazioni, trattandosi infatti di un processo stazionario la deviazione standard dovrebbe presentare un andamento costante. Tuttavia tale evoluzione permette di considerare eventuali variazioni che interessano la variabilità del processo.

1.4.3 Generazione di un segnale operativo

Dopo aver individuato la soglia superiore e la soglia inferiore, illustriamo la dinamica che a seguito della rottura di uno di questi limiti genera un segnale operativo. Consideriamo il caso in cui al tempo t il processo S_t^T si trovi al di sopra del limite

superiore, come avviene nella quinta osservazione della figura Figura 1.11. In questo caso la prima rottura della soglia, quella che avviene nel punto 1 del grafico, non costituisce ancora un vero e proprio segnale. L'apertura delle posizioni infatti, avviene in corrispondenza della rottura successiva, quella che si verifica quando il processo oltrepassa il limite superiore dall'alto verso il basso (punto 2). Lo stesso criterio di identificazione è utilizzato nel caso di rottura della soglia inferiore (colore blu). Questo schema, insieme ad una buona quantità di fiducia in merito alla stazionarietà di S_t^T , ci permette di limitare il rischio (in modo parziale) che al tempo t il processo sia in una fase di *mean aversion* piuttosto che di *mean reversion*. Riuscire a comprendere e a prevedere la dinamica del processo tra il punto 1 ed il punto 2 (o tra il punto 4 e il punto 5) della Figura 1.11 richiede tecniche più sofisticate²⁰ di quelle oggetto della nostra trattazione, pertanto ci atterremo all'utilizzo di questo schema.

1.4.4 Definizione di un *holding period*

Dopo aver definito il criterio per individuare i segnali operativi generati dal modello, illustriamo la tecnica utilizzata per calcolare l'*holding period* della posizione assunta. Come anticipato nella prima parte del paragrafo, considerando la Figura 1.11, il portafoglio venduto in corrispondenza del punto 2 (o acquistato nel punto 5) è acquistato (venduto) al momento in cui il processo S_t^T è tornato al suo valore medio (punto 3 o punto 6). Il tempo che il processo S_t^T impiega in media per ritornare al suo valore di equilibrio costituisce una informazione molto importante, utile a definire una vera e propria regola operativa. Tale misura infatti ci indica il tempo oltre il quale si deve chiudere la posizione, nel caso in cui, dopo un pò di osservazioni, il processo non sia ancora tornato al suo valore medio. Ricordiamo che durante l'implementazione di questa strategia potrebbero verificarsi *structural break* o più in generale variazioni delle condizioni al contorno, in grado di minare la stazionarietà (e quindi la *mean reversion*) di S_t^T . Per tale motivo, posto τ la media del tempo che il processo impiega per completare la sua *mean reversion*, differenze consistenti da questo parametro (fissata una certa soglia di tolleranza) possono costituire un vero e proprio 'campanello d'allarme'. Il valore di τ può essere calcolato mediante una stima campionaria. Il campione utilizzato è costruito in questo modo:

²⁰Ad esempio la creazione di un modello *state-space* per il processo *mean reverting*, che includa anche un termine di *momentum*.

$$\begin{array}{rcl}
t = 1 & \beta \mathbf{x}_1 = & \ln(P_{x_11}) - \alpha \ln(P_{x_21}) \\
\dots & \dots & \dots \\
\dots & \dots & \dots \\
t = 252 & \beta \mathbf{x}_{252} = & \ln(P_{x_1252}) - \alpha \ln(P_{x_2252}) \\
t = 253 & \beta \mathbf{x}_{253} = & \ln(P_{x_1253}) - \alpha \ln(P_{x_2253}) \\
\dots & \dots & \dots \\
t = t & \beta \mathbf{x}_t = & \ln(P_{x_1t}) - \alpha \ln(P_{x_2t}) \\
\dots & \dots & \dots \\
t = 378 & \beta \mathbf{x}_{378} = & \ln(P_{x_1378}) - \alpha \ln(P_{x_2378})
\end{array}
\left. \vphantom{\begin{array}{rcl}
t = 1 & \beta \mathbf{x}_1 = & \ln(P_{x_11}) - \alpha \ln(P_{x_21}) \\
\dots & \dots & \dots \\
\dots & \dots & \dots \\
t = 252 & \beta \mathbf{x}_{252} = & \ln(P_{x_1252}) - \alpha \ln(P_{x_2252}) \\
t = 253 & \beta \mathbf{x}_{253} = & \ln(P_{x_1253}) - \alpha \ln(P_{x_2253}) \\
\dots & \dots & \dots \\
t = t & \beta \mathbf{x}_t = & \ln(P_{x_1t}) - \alpha \ln(P_{x_2t}) \\
\dots & \dots & \dots \\
t = 378 & \beta \mathbf{x}_{378} = & \ln(P_{x_1378}) - \alpha \ln(P_{x_2378})
\end{array}} \right\} \begin{array}{l}
S_1^F \\
\dots \\
\dots \quad \textit{formation period} \\
S_{252}^F \\
S_{253}^T \\
\dots \\
S_t^T \quad \textit{trading period} \\
\dots \\
S_{378}^T
\end{array} \tag{1.18}$$

dove $\{S_1^F, \dots, S_{252}^F\}$ sono i valori del processo ottenuto stazionarizzando il *formation period* e $\{S_{253}^T, \dots, S_t^T\}$ sono i valori del processo ottenuto stazionarizzando la parte del *trading period* trascorsa fino al tempo t . Quindi al tempo t il campione utilizzato per la stima di τ è $\{S_1^F, \dots, S_{252}^F, S_{253}^T, \dots, S_t^T\}$.

Per illustrare le tecniche utilizzate per la stima di questo parametro, introduciamo il processo di Ornstein-Uhlenbeck²¹ (O-U). Tale dinamica trova un diffuso utilizzo nella modellizzazione di processi stocastici *mean reverting*. Consideriamo il processo Aritmetico Ornstein-Uhlenbeck per la variabile stocastica $X(t)$, l'equazione differenziale stocastica associata è

$$dX(t) = \theta(\mu - X(t))dt + \sigma dW(t) \tag{1.19}$$

dove $\theta > 0$ è la velocità di *mean reversion*, μ è il valore medio di lungo periodo verso cui il processo tende a tornare, $W(t)$ ²² è un moto Browniano e σ è la volatilità del processo. Questa equazione ha soluzione esplicita (Kloeden e Platen, 1992)

$$X(t) = X(0)e^{-\theta t} + (1 - e^{-\theta t})\mu + \sigma e^{-\theta t} \int_0^t e^{\theta t} dW(t) \tag{1.20}$$

²¹Noto anche come modello di Vasicek. Quest'altra denominazione è dovuta all'utilizzo del processo di Ornstein-Uhlenbeck per rappresentare la dinamica del tasso di interesse istantaneo nel modello di Vasicek per i tassi di interesse.

²² $W(t)$ è un moto Browniano allora $dW(t) \sim N(0, \sqrt{dt})$.

dove $X(t)$ si distribuisce come una Normale (Dixit e Pindyck, 1994) con media

$$E[X(t)] = X(0)e^{-\theta t} + \mu(1 - e^{-\theta t}) \quad (1.21)$$

e varianza

$$Var[X(t)] = (1 - e^{-2\theta t}) \frac{\sigma^2}{2\theta} \quad (1.22)$$

Per stimare il processo mediante un software generalmente si ricorre ad una sua discretizzazione, se consideriamo l'Eq. 1.20 nel discreto si ottiene

$$X_t = X_{t-1}e^{-\theta \Delta t} + (1 - e^{-\theta \Delta t})\mu + \sigma e^{-\theta \Delta t} + \sigma \sqrt{\frac{1 - e^{-2\theta \Delta t}}{2\theta}} N(0, 1) \quad (1.23)$$

I parametri del processo O-U possono essere stimati dal campione mediante due tecniche:

1. Calibratura con il metodo dei minimi quadrati

Utilizzando la Eq. 1.23, posto

$$a = \mu(1 - e^{-\theta \Delta t}) \quad (1.24)$$

$$b = e^{-\theta \Delta t}$$

regrediamo x_t su x_{t-1} utilizzando la retta di regressione $x_t = a + bx_{t-1} + \varepsilon_t$
Stimati a e b risolviamo ottenendo

$$a = \hat{\mu}(1 - e^{-\theta \Delta t}) \Rightarrow a = \hat{\mu}(1 - b) \Rightarrow \hat{\mu} = \frac{a}{1-b}$$

$$b = e^{-\hat{\theta} \Delta t} \Rightarrow \ln(b) = -\hat{\theta} \Delta t \Rightarrow \hat{\theta} = \frac{-\ln(b)}{\Delta t} \quad (1.25)$$

$$\hat{\sigma} = \frac{sd(\varepsilon_t)}{\sqrt{\frac{1 - e^{-\theta \Delta t}}{2\theta}}} \Rightarrow \hat{\sigma} = \frac{sd(\varepsilon_t)}{\sqrt{\frac{2\theta}{1 - e^{-\theta \Delta t}}}}$$

2. Calibratura con il metodo delle stime di massima verosimiglianza

Possiamo ricavare semplicemente la funzione di probabilità condizionata combinando l'Eq. 1.23 e la funzione di densità di una normale standard

$$P(N_{0,1} = x) = \frac{1}{\sqrt{2\pi\hat{\sigma}^2}} e^{-\frac{x^2}{2}} \quad (1.26)$$

ottenendo la funzione di densità di probabilità condizionata

$$f(x_t|x_{t-1}, \mu, \lambda, \sigma) = \frac{1}{\sqrt{2\pi\hat{\sigma}^2}} e^{\left[-\frac{(x_t - x_{t-1}e^{-\lambda\Delta t} - \mu(1 - e^{-\lambda\Delta t}))^2}{2\hat{\sigma}^2}\right]} \quad (1.27)$$

con

$$\hat{\sigma}^2 = \sigma^2 \frac{1 - e^{-2\lambda\Delta t}}{2\lambda} \quad (1.28)$$

ricavando la funzione di log-verosimiglianza per il set di osservazioni $(x_1 \dots x_n)$

$$\begin{aligned} \ell(\mu, \lambda, \hat{\sigma}) &= \sum_n^{t=1} \ln f(x_t|x_{t-1}; \mu, \lambda, \sigma) \\ &= -\frac{n}{2} \ln(2\pi) - n \ln(\hat{\sigma}) - \frac{1}{2\hat{\sigma}^2} \sum_n^{t=1} [x_t - x_{t-1}e^{-\lambda\Delta t} - \mu(1 - e^{-\lambda\Delta t})]^2 \end{aligned} \quad (1.29)$$

A questo punto posso massimizzare la *log-likelihood* uguagliando a 0 le derivate parziali

$$\begin{aligned} \frac{\partial \ell(\mu, \lambda, \hat{\sigma})}{\partial \mu} &= 0 \\ &= \frac{1}{\hat{\sigma}^2} \sum_{t=1}^n [x_t - x_{t-1}e^{-\lambda\Delta t} - \mu(1 - e^{-\lambda\Delta t})] \\ \mu &= \frac{\sum_{t=1}^n [x_t - x_{t-1}e^{-\lambda\Delta t}]}{n(1 - e^{-\lambda\Delta t})} \end{aligned} \quad (1.30)$$

$$\begin{aligned} \frac{\partial \ell(\mu, \lambda, \hat{\sigma})}{\partial \lambda} &= 0 \\ &= -\frac{-\Delta t e^{-\lambda\Delta t}}{\hat{\sigma}^2} \sum_n^{t=1} [(x_t - \mu)(x_{t-1} - \mu) - e^{-\lambda\Delta t}(x_{t-1} - \mu)^2] \\ \lambda &= \frac{1}{-\Delta t} \ln \left\{ \frac{\sum_n^{t=1} (x_t - \mu)(x_{t-1} - \mu)}{\sum_n^{t=1} (x_{t-1} - \mu)^2} \right\} \end{aligned} \quad (1.31)$$

$$\begin{aligned}
\frac{\partial \ell(\mu, \lambda, \hat{\sigma})}{\partial \hat{\sigma}} &= 0 \\
&= -\frac{n}{\hat{\sigma}} - \frac{1}{\hat{\sigma}^3} [x_t - \mu - e^{-\lambda \Delta t} (x_{t-1} - \mu)]^2 \\
\hat{\sigma}^2 &= -\frac{1}{n} [x_t - \mu - e^{-\lambda \Delta t} (x_{t-1} - \mu)]^2
\end{aligned} \tag{1.32}$$

Le soluzioni del sistema formato da queste condizioni dipendono l'una dall'altra. Tuttavia sia λ che μ sono indipendenti da σ . Il valore di σ può essere trovato una volta che λ e μ sono stati determinati. Quindi per risolvere queste equazioni è sufficiente trovare sia λ che μ . Per semplificare i passaggi cambiamo notazione

$$\begin{aligned}
S_x &= \sum_n^{t=1} x_{t-1} \\
S_{xx} &= \sum_n^{t=1} x_{t-1}^2 \\
S_y &= \sum_n x_t \\
S_{yy} &= \sum_n x_t^2 \\
S_{xy} &= \sum_n^{t=1} x_{t-1} x_t
\end{aligned} \tag{1.33}$$

allora possiamo riscrivere μ e λ come

$$\begin{aligned}
\mu &= \frac{S_y - e^{-\lambda \Delta t} S_x}{n(1 - e^{-\lambda \Delta t})} \\
\lambda &= -\frac{1}{\Delta t} \ln \left\{ \frac{S_{xy} - \mu S_x - \mu S_y + m\mu^2}{S_{xx} - 2\mu S_x + n\mu^2} \right\}
\end{aligned} \tag{1.34}$$

Adesso sostituendo λ in μ otteniamo

$$n\mu = \frac{S_y - \left(\frac{S_{xy} - \mu S_x - \mu S_y + n\mu^2}{S_{xx} - 2\mu S_x + n\mu^2} \right) S_x}{1 - \left(\frac{S_{xy} - \mu S_x - \mu S_y + n\mu^2}{S_{xx} - 2\mu S_x + n\mu^2} \right)} \tag{1.35}$$

da cui eliminando i denominatori si ottiene

$$n\mu = \frac{S_y(S_{xy} - 2\mu S_x + n\mu^2) - (S_{xy} - \mu S_x - \mu S_y + n\mu^2)S_x}{(S_{xy} - 2\mu S_x + n\mu^2) - (S_{xy} - \mu S_x - \mu S_y + n\mu^2)} \quad (1.36)$$

A questo punto raccogliendo i termini

$$n\mu = \frac{(S_y S_{xx} - S_x S_{xy}) + \mu(S_x^2 - S_x S_y) + \mu^2 n(S_y - S_x)}{(S_{xx} - S_{xy}) + \mu(S_y - S_x)} \quad (1.37)$$

e spostando quelli che moltiplicano per μ a sinistra

$$n\mu(S_{xx} - S_{xy}) - \mu(S_x^2 - S_x S_y) = S_y S_{xx} - S_x S_y \quad (1.38)$$

è possibile ottenere le equazioni di massima verosimiglianza

$$\mu = \frac{S_y S_{xx} - S_x S_{xy}}{n(S_{xx} - S_{xy}) - (S_x^2 - S_x S_y)} \quad (1.39)$$

$$\lambda = -\frac{1}{\Delta t} \ln \left\{ \frac{S_{xy} - \mu S_x - \mu S_y + n\mu^2}{S_{xy} - 2\mu S_x + n\mu^2} \right\} \quad (1.40)$$

$$\hat{\sigma}^2 = \frac{1}{n} [S_{yy} - 2a + a^2 S_{xx} - 2\mu(1-a)(S_y - aS_x) + n\mu^2(1-a^2)]$$

$$\sigma^2 = \hat{\sigma}^2 \frac{2\lambda}{1-a^2} \quad (1.41)$$

con $a = e^{-\lambda \Delta t}$.

Le due tecniche illustrate ci restituiscono le stime di $\hat{\mu}$, $\hat{\theta}$ e $\hat{\sigma}$. Per individuare l'**holding period** introduciamo il concetto di *half life* (H)²³. Tale misura rappresenta il tempo che il processo $X(t)$ impiega per raggiungere il punto medio del

²³La versione originale del concetto di *half life* deriva dalla fisica: nella misurazione del tasso di decadimento di una sostanza, l'*half life* coincide con il tempo richiesto da una certa quantità di sostanza per deteriorarsi fino alla metà della sua massa.

segmento che unisce il valore corrente $X(0)$ al valore μ di equilibrio. Nella Figura 1.11 posto $X(0)$ il punto 2 e μ il punto 3, l'*half life* corrisponde al tempo che il processo impiega per raggiungere h . Definiamo ora la relazione diretta che esiste tra la velocità di mean reversion θ e l'*half life* del processo. A tal proposito consideriamo l'Eq. 1.19, il suo valore atteso è

$$E[dX(t)] = \theta(\mu - X(t))dt + 0 \quad (1.42)$$

da cui deriva l'equazione deterministica (eliminato il termine stocastico $\sigma dW(t)$)

$$\frac{dX(t)}{\mu - X(t)} = \theta dt$$

dove $X(t)$ adesso è un valore atteso. A questo punto integrando da $X_0(t_0)$ (noto) al valore atteso del processo in t_1 , $X_1(t_1)$, posto $\Delta t = t_1 - t_0$ si ottiene:

$$-\ln(\mu - X(t))\Big|_{X_0}^{X_1} = \theta \Delta t$$

da cui

$$\ln\left(\frac{X_1 - \mu}{X_0 - \mu}\right) = -\theta \Delta t$$

Posto l'intervallo $\Delta t = t_1 - t_0$ uguale a H , per definizione si ha

$$(X_1 - \mu) = \frac{1}{2}(X_0 - \mu)$$

da cui

$$\left(\frac{X_1 - \mu}{X_0 - \mu}\right) = \frac{1}{2}$$

allora sostituendo si ottiene

$$\ln\left(\frac{1}{2}\right) = -\theta H$$

$$-\ln(2) = -\theta H$$

$$H = \frac{\ln(2)}{\theta} \quad (1.43)$$

Dopo aver stimato H con le due tecniche illustrate, possiamo ottenere $\tau = 2H$.

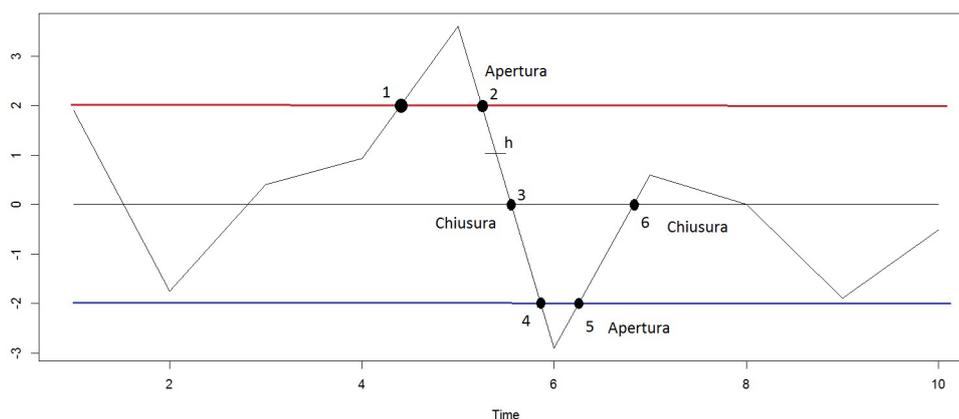


Figura 1.11: **Esempio di generazione di un segnale operativo**

Una volta stimato τ possiamo definire l'*holding period* della nostra posizione come un multiplo di questa misura, oppure soltanto aggiungendo una soglia di tolleranza per le deviazioni dal valore stimato. Come anticipato, il procedimento illustrato ha condotto alla definizione di una vera e propria regola operativa. Da un punto di vista pratico se dopo un lasso di tempo, che supera considerevolmente la nostra stima, il processo S^T non è ancora tornato al suo valore di equilibrio, dovremo iniziare a ponderare un cambiamento nel regime che potrebbe condurre a grossi rischi. Per tale motivo, in questa situazione, superato l'*holding period* si procede alla chiusura delle posizioni che formano la combinazione acquistata o venduta.

1.5 Applicazione della strategia

1.5.1 Premesse al modello

In questo paragrafo assumeremo che, nella fase di applicazione della strategia, valgono le seguenti condizioni:

1. non esistono costi di transazione, questo elimina l'onere di considerare fattori come il *bid-ask spread*, commissioni applicate dal broker e lotti minimi di negoziazione;
2. tutti i profitti sono soggetti alla stessa aliquota fiscale;
3. la valuta in tutti e tre i campioni è il Dollaro;
4. non ci sono limiti sul *leverage*;

5. il denaro può essere dato o preso in prestito allo stesso tasso risk free pari al 5% composto continuamente;
6. qualora il processo S^T raggiunga il suo valore medio o generi un segnale operativo nel periodo compreso tra due osservazioni, l'operazione è eseguita utilizzando il valore della quotazione disponibile più prossima;
7. l'operazione non è eseguita se al momento in cui si genera il segnale operativo le componenti del vettore di cointegrazione hanno subito una variazione superiore al 15%²⁴ (in valore assoluto).

In merito alla prima assunzione va precisato che, per quanto riguarda le osservazioni contenute nel campione 3, queste derivano da strumenti finanziari che presentano condizioni di negoziazione peculiari, sia in merito allo scambio che alla concreta monetizzazione dei rendimenti. Tuttavia la scelta di tali asset è legata alla particolare dinamica seguita dalle loro serie storiche e non alla particolarità degli aspetti tecnici che caratterizzano, ad esempio, la loro negoziazione. Pertanto l'implementazione della strategia su questo campione avverrà assumendo un *environment* semplificato, che non considera l'insieme di frizioni generate dalle caratteristiche tecniche del future e dell'ETF.

1.5.2 Stazionarietà e implementazione della strategia

Creazione dei processi S^{1T} , S^{2T} e S^{3T}

Consideriamo le matrici \mathbf{Y}^{1T} , \mathbf{Y}^{2T} e \mathbf{Y}^{3T} , di dimensioni (252×2) , le cui colonne sono composte dalle osservazioni del *trading period* relativo ad ogni campione. Per il campione 1, la stazionarizzazione del periodo di *trading* avviene mediante l'operazione illustrata nella Eq. 1.14, tuttavia in questo caso la matrice $\mathbf{\Omega}_1$ è sostituita dalla matrice $\mathbf{\Gamma}_1$ (di pari dimensioni), costituita per le prime due righe dalle colonne della matrice \mathbf{Y}^{1T} e per la terza riga da un vettore di 1. Tale sostituzione è stata effettuata anche per il campione 2 e per il campione 3. In altre parole le matrici $\mathbf{\Omega}_2$ e $\mathbf{\Omega}_3$ sono state sostituite dalle matrici $\mathbf{\Gamma}_2$ e $\mathbf{\Gamma}_3$. Anche in questo caso la differenza è dovuta alle prime due righe delle matrici $\mathbf{\Gamma}_2$ e $\mathbf{\Gamma}_3$, contenenti rispettivamente, la matrice \mathbf{Y}^{2T} e la matrice \mathbf{Y}^{3T} . Applicando l'Eq. 1.14 al *trading period* di ogni campione otteniamo:

$$\beta^1 \mathbf{\Gamma}_1 = S^{1T} \quad (1.44)$$

$$\beta^2 \mathbf{\Gamma}_2 = S^{2T} \quad (1.45)$$

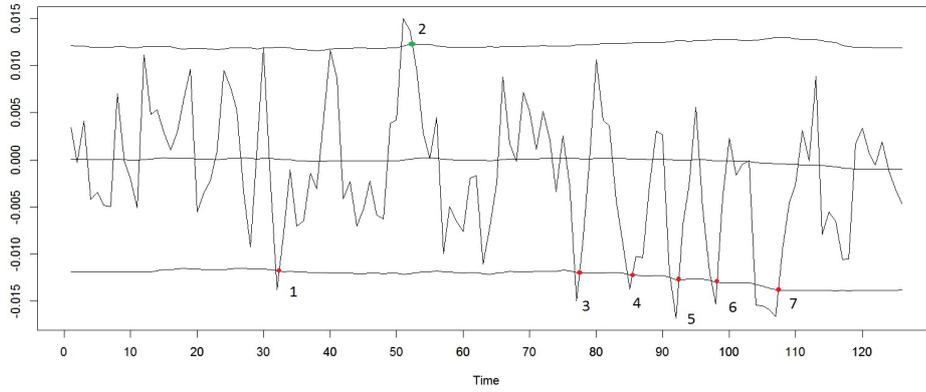
²⁴Si veda Herlemont D. (2004).

$$\beta^3 \Gamma_3 = S^{3T} \quad (1.46)$$

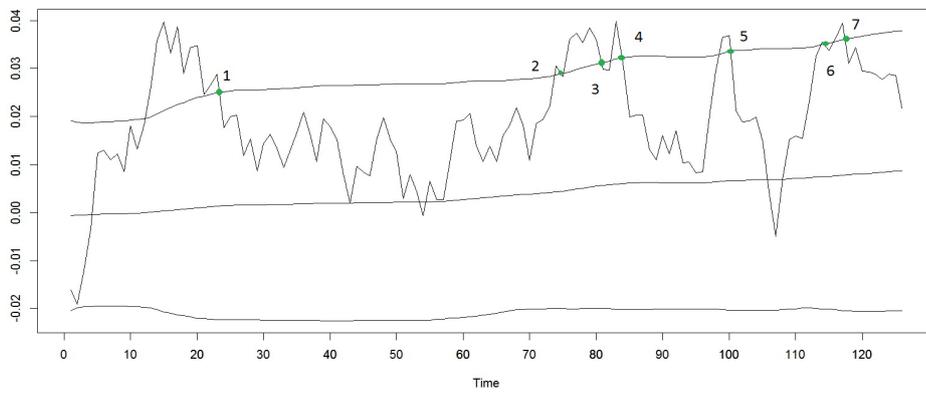
dove β^1 , β^2 e β^3 sono i vettori di cointegrazione stimati. La Figura 1.12 riporta i grafici dei tre processi creati.

Nelle Tabelle 1.12, 1.13 e 1.14 è presentata una sintesi della dinamica seguita dai processi S^{1T} , S^{2T} e S^{3T} . In ognuna di queste tabelle la seconda colonna a sinistra indica le posizioni in cui si sono verificate la prima e la seconda rottura del limite inferiore o superiore. A tal proposito notiamo subito una notevole differenza tra i processi S^{1T} e S^{2T} . Infatti il periodo di tempo che intercorre tra il primo superamento e la generazione del segnale operativo è pari ad un giorno nel campione 1, mentre può raggiungere le 9 osservazioni nel campione 2. Questa differenza è dovuta all'elevata velocità di *mean reversion* che caratterizza il processo S^{1T} . Le colonne denominate *sup.* e *inf.* contengono i valori del limite superiore e inferiore, calcolati mediante il metodo della deviazione standard nella sua evoluzione 'rolling', quindi con valori aggiornati in seguito a ogni osservazione. Con la stessa tecnica è stato calcolato μ ossia la media cui tende il processo, i cui valori sono contenuti nell'ultima colonna. Inoltre in ciascuna delle tre tabelle, le rilevazioni evidenziate in giallo indicano i valori del processo e del relativo limite in corrispondenza di ciascuna delle rotture verificatesi durante il *trading period*.

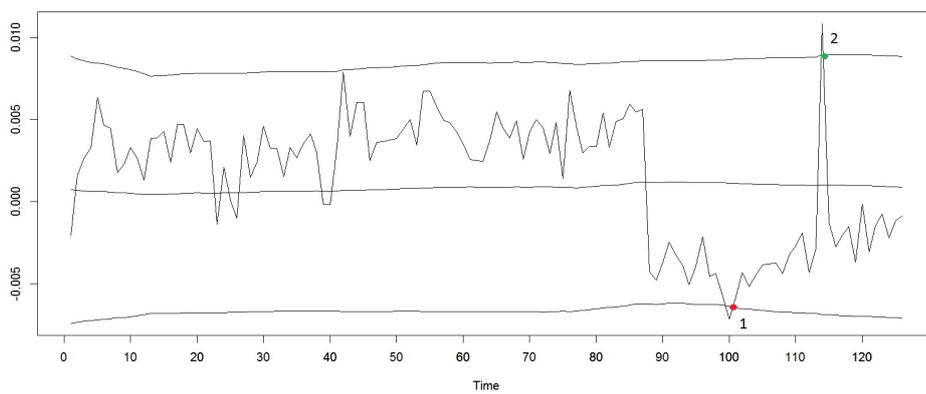
Dalle tabelle Tabelle 1.12, 1.13 e 1.14 emerge una relazione stretta tra possibilità operative e *mean reversion* del processo. Per comprendere questa relazione abbiamo effettuato un'analisi sulla stazionarietà dei processi S^{1T} , S^{2T} e S^{3T} . In altre parole abbiamo testato la capacità del vettore di cointegrazione di rendere stazionaria la combinazione lineare ottenuta mediante i valori del *trading period*. Anche in questo caso siamo ricorsi ai test ADF, PP e KPSS. Prima di illustrarne i risultati riprendiamo una considerazione accennata nella prima parte del capitolo. L'applicazione di questa tipologia di test, come di altri strumenti econometrici, non può prescindere da una preventiva lettura del fenomeno reale. In altre parole la loro funzionalità, dipende dalla capacità dell'utilizzatore di trarne informazioni coerenti con le caratteristiche della serie storica che si sta analizzando e di ponderare che in alcune occasioni la loro significatività è scarsa. A tal proposito è doveroso fare delle precisazioni. Nel tempo questi test si sono evoluti fino a comprendere componenti deterministiche nella rappresentazione delle serie. Tuttavia la storia più recente di questi test ha condotto a risultati sconcertanti in merito alla scarsa potenza dell'applicazione su piccoli campioni. Si è compreso, potremmo dire definitivamente, che questo problema non è altro che il prodotto della sostanziale indistinguibilità, per piccoli campioni, del comportamento di serie stazionarie attorno ad un trend (*trend stationary*) da quello di serie $I(1)$. Questa indistinguibilità, definita 'equi-



(a) Processo S^{1T}



(b) Processo S^{2T}



(c) Processo S^{3T}

Figura 1.12: Grafico dei *trading period* stazionarizzati

<i>op.</i>	<i>time</i>	<i>rottura</i>	S_t^{1T}	<i>sup.</i>	<i>inf.</i>	μ
1	683	<i>prima</i>	-0.013757	0.011965	-0.011787	0.000089
1	684	<i>segnale</i>	-0.007601	0.011965	-0.011862	0.000051
2	702	<i>prima</i>	0.014975	0.012075	-0.012109	-0.000017
2	704	<i>segnale</i>	0.009392	0.012262	-0.011937	0.000162
3	728	<i>prima</i>	-0.015004	0.012192	-0.011933	0.000129
3	729	<i>segnale</i>	-0.008844	0.012214	-0.011994	0.000110
4	736	<i>prima</i>	-0.013693	0.012412	-0.012177	0.000117
4	737	<i>segnale</i>	-0.010269	0.012427	-0.012207	0.000110
5	743	<i>prima</i>	-0.016799	0.012625	-0.012680	-0.000028
5	744	<i>segnale</i>	-0.006843	0.012631	-0.012729	-0.000049
6	749	<i>prima</i>	-0.015297	0.012780	-0.012986	-0.000103
6	750	<i>segnale</i>	-0.004163	0.012776	-0.013010	-0.000117
7	755	<i>prima</i>	-0.015427	0.012698	-0.013214	-0.000258
7	759	<i>segnale</i>	-0.009405	0.012979	-0.013819	-0.000420

Tabella 1.12: **Dinamica del processo S^{1T} e generazione dei segnali operativi**

<i>op.</i>	<i>time</i>	<i>rottura</i>	S_t^{2T}	<i>sup.</i>	<i>inf.</i>	μ
1	1160	<i>prima</i>	0.026059	0.019933	-0.019842	0.000046
1	1171	<i>segnale</i>	0.017667	0.025103	-0.022282	0.001411
2	1221	<i>prima</i>	0.030532	0.028624	-0.019973	0.004326
2	1222	<i>segnale</i>	0.028288	0.028890	-0.019928	0.004481
3	1223	<i>prima</i>	0.036116	0.029336	-0.020023	0.004656
3	1228	<i>segnale</i>	0.029826	0.031178	-0.019898	0.005640
4	1222	<i>prima</i>	0.035937	0.030923	-0.019973	0.005475
4	1222	<i>segnale</i>	0.029826	0.031178	-0.019898	0.005640
5	1230	<i>prima</i>	0.039811	0.031960	-0.020153	0.005904
5	1231	<i>segnale</i>	0.031128	0.032265	-0.020223	0.006021
6	1246	<i>prima</i>	0.036857	0.033597	-0.020367	0.006615
6	1248	<i>segnale</i>	0.021162	0.033706	-0.020380	0.006663
7	1261	<i>prima</i>	0.035499	0.034990	-0.020026	0.007482
7	1262	<i>segnale</i>	0.033799	0.035273	-0.020133	0.007570
8	1263	<i>prima</i>	0.039417	0.035980	-0.020439	0.007770
8	1265	<i>segnale</i>	0.031024	0.036218	-0.020503	0.007858

Tabella 1.13: **Dinamica del processo S^{2T} e generazione dei segnali operativi**

<i>op.</i>	<i>time</i>	<i>rottura</i>	S_t^{3T}	<i>sup.</i>	<i>inf.</i>	μ
1	1247	<i>prima</i>	-0.00717	0.00867	-0.00639	0.00114
1	1248	<i>segnale</i>	-0.00575	0.00869	-0.00646	0.00111
2	1261	<i>prima</i>	0.030532	0.028624	-0.019973	0.004326
2	1262	$\leq \mu$	-0.00126	0.00894	-0.00689	0.00102

Tabella 1.14: **Dinamica del processo S^{3T} e generazione dei segnali operativi**

valenza prossima da un punto di vista osservativo' è sintetizzata in Campbell e Perron (1991) nel modo seguente

“Per campioni finiti, ogni processo stazionario intorno ad un trend può essere approssimato arbitrariamente bene da un processo a radice unitaria (nel senso che le strutture di autocovarianza risulteranno arbitrariamente vicine)”.

Tutto questo comporta che qualora l'analisi di cointegrazione dovesse basarsi su analisi preliminari di integrazione, il livello di confidenza effettivo, successivo all'analisi del rango di cointegrazione, si dovrebbe ridurre drammaticamente al di sotto del livello nominale normalmente assunto²⁵. Fatta questa precisazione (più o meno arrendevole allo stato dell'arte), l'interpretazione dei risultati può essere accompagnata dalle informazioni ottenute mediante la visualizzazione grafica e da una teoria sottostante in merito al comportamento della serie. Nella costruzione del nostro modello i test ADF, PP e KPSS sono stati utilizzati mediante un duplice approccio. Siamo partiti dall'assunto secondo cui nessuna serie $I(1)$ è stazionaria ma non tutte le serie non-stazionarie sono $I(1)$ ²⁶. Quindi, nel testare la presenza di radici unitarie la statistica KPSS è stata utilizzata a conferma dei risultati ottenuti dall'ADF e dal PP (Leybourne e McCabe, 1989). Quando invece l'analisi è stata finalizzata alla verifica della stazionarietà, questa si è poggiata in modo più consistente sulla visualizzazione grafica e sulla teoria sottostante in merito alle caratteristiche del processo. In particolare i risultati dei test ADF, PP e KPSS, riportati nella Tabella 1.17, sono stati interpretati tramite quest'ultimo approccio. I valori riportati nella tabella evidenziano la stazionarietà del processo S^{1T} , questa caratteristica trova supporto nel grafico (a) della Figura 1.12. In questo caso il vettore β^1 stimato sul *formation period* riesce a rendere stazionario il processo S^{1T} . La stazionarietà ne garantisce l'elevata *mean reversion*, che a sua volta costituisce la condizione ideale per l'implementazione di questa strategia. Tale caratteristica ha favorito il continuo generarsi di segnali operativi e la chiusura delle posizioni secondo lo schema base illustrato nella Figura 1.11. Per quanto riguarda il processo S^{2T} i test di radice unitaria rigettano l'ipotesi nulla, rifiutando quindi l'ipotesi che si tratti di un processo $I(1)$. Il test KPSS invece rigetta l'ipotesi nulla di stazionarietà, infatti la statistica test (0.397) è maggiore della soglia critica al 10% (0.347). Il grafico (b) della Figura 1.11 mostra delle oscillazioni che non possono essere definite casuali. A occhio nudo è percepibile come, all'interno della

²⁵Per un tentativo di soluzione del problema si veda Giannini (1992).

²⁶Quest'ultima fattispecie rappresenta un caso tipico che può indurre all'errore di ritenere contraddittori i risultati delle due tipologie di test.

(a) ADF				
Processo	test	1%	5%	10%
S^{1T}	-6.39	-3.46	-2.88	-2.57
S^{2T}	-4.26	-3.46	-2.88	-2.57
S^{3T}	-2.19	-3.46	-2.88	-2.57

(b) PP		
Processo	test	Z-stat.
S^{1T}	-53.11	-1.43
S^{2T}	-21.57	3.56
S^{3T}	-26.54	1.58

(c) KPSS					
Processo	test	10%	5%	2.5%	1%
S^{1T}	0.069	0.119	0.146	0.176	0.216
S^{2T}	0.397	0.347	0.463	0.574	0.739
S^{3T}	1.223	0.347	0.463	0.574	0.739

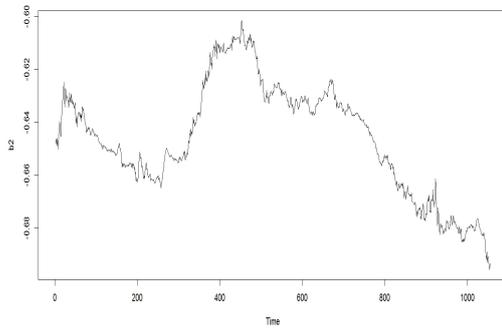
Tabella 1.15: **Test per radice unitaria e test di stazionarietà per i processi S^{1T} , S^{2T} e S^{3T}**

finestra temporale del periodo di *trading*, il vettore di cointegrazione β^2 non riesce a rendere stazionario il processo. Inoltre osservando la soglia superiore, calcolata aggiornando la deviazione standard, è evidente come il processo presenti un trend nella varianza. Questa caratteristica viola una delle condizioni fondamentali della definizione di processo stazionario. Nonostante la Tabella 1.18 mostri rendimenti positivi in corrispondenza delle operazioni effettuate, in questo caso l'implementazione della strategia potrebbe diventare rischiosa. Soprattutto nel caso in cui il processo si trovi oltre una certa soglia, verrebbe meno il vantaggio derivante dal prevedibile ritorno verso la media. Da un punto di vista operativo, l'*holding period* stimato calibrando il modello O-U sul *formation period* stazionarizzato potrebbe diventare una regola decisionale inutile. Infine i risultati relativi al processo S^{3T} ripropongono una situazione più o meno analoga a quella del processo S^{2T} . In questo caso il test ADF non rigetta l'ipotesi relativa alla presenza di radice unitaria e il test KPSS rifiuta in modo netto l'ipotesi nulla di stazionarietà. Il grafico della Figura 1.11(c) mostra un processo che, pur oscillando all'interno delle due soglie senza una forte componente di stagionalità né di deriva, non presenta la caratteristica di *mean reversion* necessaria all'applicazione della nostra strategia. Pertanto l'interpretazione dei risultati e la rappresentazione grafica evidenziano l'impossibilità di usufruire della relazione di cointegrazione individuata per ottenere una combinazione stazionaria.

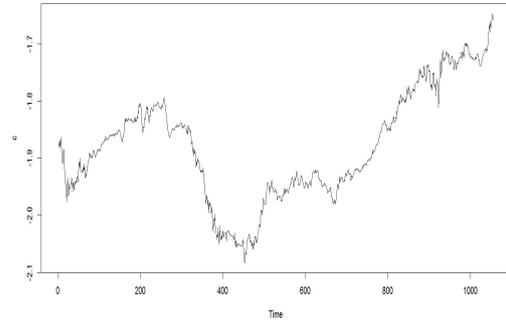
Stabilità del vettore di cointegrazione e applicabilità del modello

Le differenze appena descritte in merito alla stazionarietà dei processi S^{1T} , S^{2T} e S^{3T} sembrano esser collegate all'intensità della relazione di cointegrazione. La misura dei valori assunti dalla statistica di Johansen, riportati nelle Tabelle 1.8, 1.10 e 1.11, sembrano rispecchiare i risultati ottenuti in termini di stazionarietà dei processi S^{1T} , S^{2T} e S^{3T} . Secondo questa idea, considerando la Tabella 1.8, la grandezza del valore critico assunto dalla statistica in corrispondenza dell'ipotesi relativa all'assenza di relazioni di cointegrazione (66.57), giustificherebbe la stazionarietà ottenuta mediante l'Eq. 1.44. Pertanto la non stazionarietà dei processi S^{2T} e S^{3T} sarebbe collegata ai valori più contenuti (rispettivamente 31.94 e 18.08) assunti dalla statistica nelle altre due stime. Nonostante valori della statistica decrescenti sembrino accompagnare una stazionarizzazione sempre peggiore, tale lettura del fenomeno risulta però troppo semplicistica. Questo ci ha spinto allo sviluppo di un'applicazione 'rolling' del test di Johansen, in grado di considerare la mutevolezza delle condizioni al contorno in cui questo viene implementato. In particolare la statistica è stata applicata su ciascun campione mediante delle 'rolling window' di 252 osservazioni²⁷. Le stime mobili delle componenti di ciascun vettore cointegrante sono rappresentate nei grafici della Figura 1.13. Questi evidenziano una consistente variabilità delle componenti di ciascun vettore. In particolare, i grafici relativi alle stime effettuate sul campione 2 e sul campione 3, oltre a considerevoli variazioni, presentano sporadici *shock*. Tale caratteristica potrebbe trovare parziale giustificazione negli shock e nei cambiamenti di regime intrinseci alle serie. L'ipotesi è che questi fattori possano minare la stabilità del vettore stimato su finestre di dimensione ridotta. Tuttavia considerando il grafico relativo alle componenti stimate sul primo campione (Figura 1.13 (a) e (b)), sembrerebbe che la procedura di Johansen non sia robusta, infatti come nei due casi precedenti, spostamenti della finestra temporale possono generare cambiamenti drammatici nella stima delle componenti del vettore. Eppure le serie storiche che costituiscono il campione 1 sono state generate in modo tale che fossero caratterizzate da una forte relazione di cointegrazione. Sembra a questo punto che per stime effettuate su finestre di queste dimensioni, la stabilità del vettore cointegrante possa dipen-

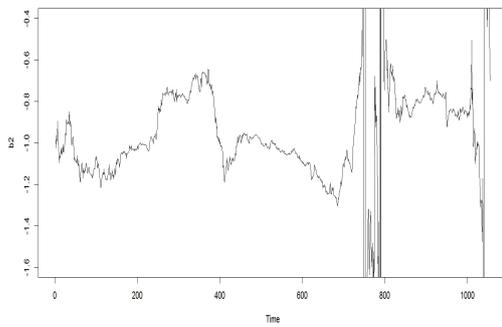
²⁷La scelta dell'ordine di ritardo utilizzato per effettuare il test su ogni finestra temporale è avvenuta mediante una stima *rolling* dei criteri di informazione AC, BIC e SC. Effettuata la stima, in base all' *information criterion* SC, è stato scelto un ritardo pari a 2 per ciascuna delle finestre temporali di tutti e tre i campioni. Il test di Johansen è stato implementato tramite l'Eq. 1.12. Trattandosi di un'applicazione 'rolling' l'ipotesi di costanza di tali parametri potrebbe risultare una limitazione per la bontà della stima. Tuttavia il nostro scopo è il confronto in termini relativi dei valori assunti dalla statistica in ciascuna delle finestre su cui è stato implementato il test. Pertanto l'ipotesi di non variazione di questi parametri non dovrebbe inficiare molto l'obiettivo della nostra analisi.



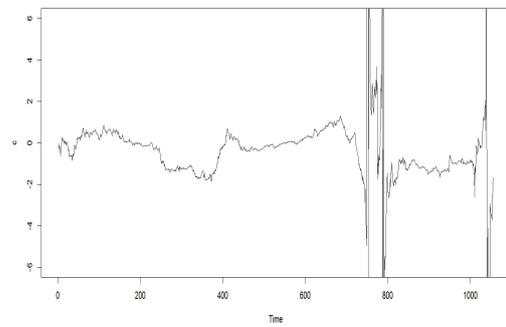
(a) Seconda componente del vettore stimato sul campione 1



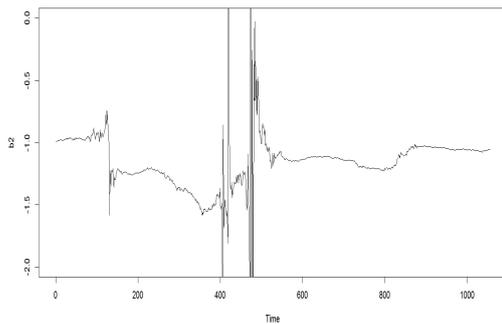
(b) Terza componente del vettore stimato sul campione 1



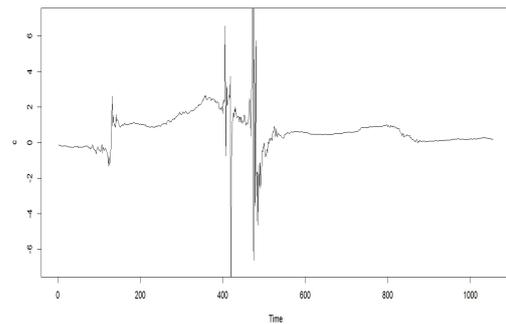
(c) Seconda componente del vettore stimato sul campione 2



(d) Terza componente del vettore stimato sul campione 2



(e) Seconda componente del vettore stimato sul campione 3



(f) Terza componente del vettore stimato sul campione 3

Figura 1.13: *Analisi rolling* delle componenti dei vettori di cointegrazione

dere in modo solo parziale dall'intensità della relazione di cointegrazione. A tal proposito accenniamo ad una problematica introdotta da Phillips (Phillips Peter C. B., 1994) in merito allo strano comportamento degli stimatori di massima verosimiglianza nel caso di campioni finiti. Questa lavoro ha dimostrato infatti che in piccoli campioni stimatori di massima verosimiglianza non hanno momenti finiti. Guardando i grafici, la situazione potrebbe essere anche peggiore poiché le stime *rolling* appaiono caratterizzate anche da una certa componente di correlazione e da sporadici shock. Inoltre come affermato da Zivot (2010), il fatto che la *finite sample distribution* non abbia momenti finiti rende possibile risultati assurdi dell'MLE di Johansen. Pertanto sarebbe interessante avere qualche contributo in merito alla distribuzione delle stime mobili del vettore di cointegrazione. In tal senso sono presenti in letteratura lavori in cui è stata derivata la distribuzione delle stime *rolling* (su un campione finito) di test a radice unitaria (Banerjee, Lumsdain e Stock, 1992).

Tornando al modello operativo costruito, l'identificazione della distribuzione delle stime mobili del vettore cointegrante sul campione finito ci fornirebbe uno strumento molto utile. Ad esempio tramite queste informazioni, si potrebbe subordinare l'apertura di posizioni a delle verifiche sulla stabilità del vettore di cointegrazione. Questo potrebbe tradursi non tanto in un miglioramento della profittabilità della strategia, quanto piuttosto in riduzione dei rischi legati ai cambiamenti di regime o altri fattori che possono minare la stabilità del vettore.

1.5.3 *Back Testing* della strategia

Il paragrafo che segue sintetizza i risultati ottenuti applicando il modello operativo sui tre campioni. A tale scopo è stata sviluppata un'analisi di tipo *Back Testing*. Per fare questo si è collocato il punto di osservazione alla fine del *trading period*, guardando indietro sono state rilevate le operazioni eseguite dal nostro modello fino a quel momento, valutandone i risultati in termini di rischio e rendimento.

Operazioni concluse e rendimento

Le Tabelle 1.17, 1.18 e 1.19 riassumono le operazioni eseguite applicando il modello operativo costruito. Tutte le operazioni sono state effettuate su una combinazione dei due titoli che compongono ogni campione. Per esempio, la Tabella 1.17 riassume i risultati relativi alle posizioni assunte sulla combinazione $A - 0.613B$, dove A e B sono i titoli che costituiscono il campione 1. A tal proposito la colonna denominata *Buy\Sell* identifica la natura della posizione assunta sul portafoglio composto dai due titoli. La Tabella 1.16 contiene la dimensione dei lotti di nego-

posizione	A	B	C	D	f	E
<i>long</i>	10000	-6130	10000	-7310	10000	-10300
<i>short</i>	-10000	6130	-10000	7310	-10000	10300

Tabella 1.16: **Lotti di negoziazione per l'assunzione di posizioni sul portafoglio**

ziazione scelti per ogni coppia di titoli. Le quantità acquistate e vendute in ogni portafoglio rispettano le proporzioni delle componenti di ognuno dei vettori β^1 , β^2 e β^3 . Inoltre le colonne R e R^* riportano il rendimento ottenuto da ogni operazione, rispettivamente al netto e al lordo degli interessi passivi e attivi maturati durante l'*holding period*. A tal proposito, per illustrarne il calcolo, prendiamo come esempio la seconda operazione presente nella Tabella 1.17. In questo caso, in seguito al segnale generato dal processo S^{1T} , in corrispondenza dell'osservazione 704 è stato venduto il portafoglio composto da una posizione lunga su 10000 unità del titolo A e una corta su 6130 unità del titolo B. Il portafoglio è stato riacquistato dopo due giorni di negoziazione (osservazione 706). Il rendimento netto derivante dall'operazione è pari a

$$\begin{aligned} & [\ln(P_{A704}) - \beta_2^1 \ln(P_{B704}) + c] - [\ln(P_{A706}) - \beta_2^1 \ln(P_{B706}) + c] \\ & [\ln(P_{A704}) - \ln(P_{A706})] + \beta_2^1 [\ln(P_{B706}) - \ln(P_{B704})] \\ & [\ln(120.15) - \ln(119.47)] + 0.613 [\ln(88.71) - \ln(87.18)] = 0.92\%. \end{aligned}$$

Come anticipato, il tasso di interesse passivo e attivo è pari al 5% (tasso *risk free*). Nel calcolo degli interessi è stato utilizzato il regime di capitalizzazione continua. Pertanto il rendimento lordo è pari a

$$\left[\ln(120.15e^{0.05 \frac{2}{252}}) - \ln(119.47) \right] + 0.613 \left[\ln(87.18) - \ln(88.71e^{0.05 \frac{2}{252}}) \right] = 0.94\%$$

La colonna τ indica il tempo che i processi S^{1T} , S^{2T} e S^{3T} impiegano per tornare al valore μ . Queste stime sono state ottenute mediante la tecnica descritta nel Par. 1.4.4, in particolare è stato scelto l'approccio basato sul metodo dei minimi quadrati (Eq. 1.25). In corrispondenza di ogni segnale operativo abbiamo calcolato τ , in seguito applicando una soglia di tolleranza discrezionale abbiamo fissato l'*holding period* (*h.p.* nella tabella) per quella posizione. Come riportato nella Tabella 1.17, le posizioni aperte sul portafoglio composto dai titoli A e B sono state chiuse sempre in anticipo rispetto alla scadenza dell'*holding period* fissato alla data dell'apertura. Questo è dovuto alla forte *mean reversion* che caratterizza il processo S^{1T} . Infatti senza tradire le aspettative, ogni posizione è stata chiusa poiché

<i>time</i>	<i>Buy/Sell</i>	P_A	P_B	τ	<i>h.p.</i>	R	R^*	$\Delta\beta_2^1$	Δc
684	<i>B</i>	121.92 \$	93.39 \$						
688	<i>S</i>	123.66 \$	94.61 \$	2.13	4	0.62%	0.59%	1%	1%
704	<i>S</i>	120.15 \$	88.71 \$						
706	<i>B</i>	117.79 \$	87.18 \$	2.22	4	0.92%	0.94%	2%	3%
729	<i>B</i>	117.67 \$	88.32 \$						
731	<i>S</i>	119.46 \$	87.70 \$	2.19	4	1.94%	1.93%	0%	0%
737	<i>B</i>	117.52 \$	88.34 \$						
740	<i>S</i>	115.53 \$	84.07 \$	2.23	4	1.33%	1.27%	0%	0%
744	<i>B</i>	114.24 \$	83.89 \$						
746	<i>S</i>	114.62 \$	82.66 \$	2.26	4	1.24%	1.22%	1%	2%
750	<i>B</i>	115.13 \$	84.59 \$						
751	<i>S</i>	114.84 \$	83.35 \$	2.27	4	0.65%	0.64%	2%	3%
759	<i>B</i>	111.59 \$	81.08 \$						
764	<i>S</i>	110.00 \$	76.88 \$	2.42	5	1.83%	1.79%	3%	4%

Tabella 1.17: **Operazioni eseguite sul portafoglio composto da una posizione lunga sul titolo A e una posizione corta sul titolo B**

il processo è tornato al valore μ e non per decorrenza dell'*holding period*. I dati riportati nelle Tabelle 1.18 e 1.19 evidenziano invece, tempi di *mean reversion* molto più dilatati. Questo ha comportato la chiusura delle posizioni per decorrenza dell'*holding period* prestabilito²⁸. Rappresenta una eccezione l'operazione numero 5 effettuata sul portafoglio composto dai titoli C e D. In questo caso infatti, la posizione è stata chiusa per il completamento della dinamica di *mean reversion*.

Inoltre come anticipato alla fine del Par. 1.4.1, successivamente al manifestarsi di un segnale operativo, l'apertura di ogni posizione è stata subordinata al controllo della variazione del vettore di cointegrazione. Per tale motivo, le ultime due colonne di ogni tabella riportano la variazione assoluta che ha interessato la componente β_2 e la componente c di ogni vettore. Da un punto di vista pratico, prima di aprire una posizione, è stata effettuato il test di Johansen sulla finestra temporale contenente le 252 osservazioni (logaritmo naturale del prezzo) precedenti alla data

²⁸Nel caso dei campioni 2 e 3 non è stato applicato nessun fattore moltiplicativo al parametro τ . Così facendo l'*holding period* è stato definito pari a τ senza alcuna soglia di tolleranza. Tale scelta è legata alla volontà di ottenere un *holding period* fortemente basato sulla stima del parametro, quindi in grado di far risaltare l'efficienza o meno della sua stima.

<i>time</i>	<i>Buy/Sell</i>	P_C	P_D	τ	<i>h.p.</i>	R	R^*	$\Delta\beta_2^2$	Δc
547	<i>S</i>	132.29 \$	109.00 \$						
557	<i>B</i>	131.82 \$	109.20 \$	10	10	0.48%	0.54%	5%	13%
598	<i>S</i>	133.8 \$	109.11 \$						
609	<i>B</i>	132.43 \$	108.79 \$	11	11	0.81%	0.87%	3%	8%
604	<i>B</i>	134.28 \$	109.42 \$						
617	<i>S</i>	132.43 \$	109.98 \$	13	13	1.76%	2%	5%	13%
607	—	134.19 \$	109.12 \$	-	-	-	-	6%	15%
624	<i>B</i>	132.44 \$	108.66 \$						
629	<i>S</i>	132.01 \$	110.68 \$	13	13	1.67%	1.70%	5%	12%
638	—	138.79 \$	113.86 \$	-	-	-	-	14%	34%
641	—	139.27 \$	114.83 \$	-	-	-	-	34%	79%

Tabella 1.18: **Operazioni eseguite sul portafoglio composto da una posizione lunga sul titolo C e una posizione corta sul titolo D**

<i>time</i>	<i>Buy/Sell</i>	P_f	P_E	τ	<i>h.p.</i>	R	R^*	$\Delta\beta_2^3$	Δc
1248	<i>B</i>	132.29 \$	109.00 \$						
1260	<i>S</i>	131.82 \$	109.20 \$	6.12	12	0.28%	0.29%	5%	3%

Tabella 1.19: **Operazioni eseguite sul portafoglio composto da una posizione lunga sul *future* e una posizione corta sull'ETF**

del segnale operativo. Effettuata la stima, la posizione è stata aperta solo nel caso la variazione assoluta non sia stata superiore al 15%. Nella Tabella 1.18, le operazioni 4, 6 e 7 non sono state eseguite proprio a causa di un cambiamento rilevante delle componenti β_2^2 e c interne al vettore β^2 .

Vantaggi derivanti dalla strategia *long-short*

Il portafoglio costruito mediante questo modello rientra nella categoria delle strategie definite *long-short*. Le motivazioni per cui si giunge alla decisione di andare lunghi su uno o più asset e contestualmente corti su qualcos'altro (strategia che ha l'obiettivo di incassare solamente il differenziale di rendimento tra le due posizioni) possono essere le più disparate, in funzione della metodologia utilizzata. Nell'implementazione di questo modello, la combinazione *long-short* è ottenuta mediante un'analisi puramente econometrica ed è finalizzata ad ottenere un processo *mean reverting*. Quindi la combinazione che costituisce il nostro portafoglio è lontana da considerazioni sul *forecast* degli andamenti o dal *sentiment* presente sul mercato. Inoltre va sottolineato che, seppur implementato in un contesto semplificato, questo modello permette di sfruttare i vantaggi tipici di questa strategia. Infatti l'apertura di una posizione lunga o corta sul portafoglio, genera la combinazione *long-short* degli asset che lo compongono. Il controvalore della posizione corta su uno dei due titoli fornisce un ammontare di risorse utile a coprire, in modo totale o parziale, il fabbisogno di risorse generato dalla posizione lunga sull'altro. Questo permette di pagare interessi solo sulla differenza tra i due ammontari. Come definito dalle premesse al modello, l'applicazione di questa strategia è avvenuta in contesto semplificato. Un esempio di tale semplificazione consiste nella possibilità di finanziarsi e investire allo stesso tasso pari al 5% composto continuamente. Tuttavia questo non ci impedisce di apprezzare i benefici descritti che, spesso costituiscono una delle motivazioni per l'assunzione di posizioni di questo tipo da parte di grandi istituzioni finanziarie. A tal proposito, nella Tabella 1.17 sono riportati gli interessi passivi e attivi maturati sulla differenza del controvalore delle posizioni assunte. Naturalmente gli interessi passivi si generano nel caso in cui l'ammontare richiesto dalla posizione lunga è maggiore di quello che si ottiene assumendo la posizione corta. Le operazioni contenute nella Tabella 1.17 (a) infatti, forniscono un esempio di interessi passivi dovuti al maggior ammontare della posizione lunga sul titolo A. La Tabella 1.17 (b) invece, presenta gli interessi maturati effettuando operazioni in cui la posizione corta ha un controvalore maggiore. Inoltre la contemporaneità che caratterizza l'apertura e la chiusura delle posizioni sui due titoli che compongono il portafoglio, consente di conoscere in modo più o meno certo la

(a) Portafoglio composto dai titoli A e B

<i>op.</i>	<i>\$suA</i>	<i>\$suB</i>	$\Delta \$$	<i>Interessi</i>	<i>profitto</i>
1	-1 219 191 \$	572 504 \$	-646 687 \$	-513 \$	9 372 \$
2	1 201 544 \$	-543 775 \$	657 769 \$	261 \$	14 569 \$
3	-1 176 684 \$	541 412 \$	-635 272 \$	-252 \$	21 506 \$
4	-1 175 165 \$	541 530 \$	-633 635 \$	-377 \$	6 614 \$
5	-1 142 417 \$	514 259 \$	-628 158 \$	-249 \$	11 096 \$
6	-1 151 357 \$	518 566 \$	-632 791 \$	-126 \$	4 516 \$
7	-1 115 941 \$	497 039 \$	-618 902 \$	-614 \$	9 246 \$

(b) Portafoglio composto dai titoli C e D

<i>op.</i>	<i>\$suC</i>	<i>\$suD</i>	$\Delta \$$	<i>Interessi</i>	<i>profitto</i>
1	1 322 858 \$	-796 795 \$	526 063 \$	1 045 \$	13 552 \$
2	1 337 965 \$	-797 597 \$	540 368 \$	1 181 \$	12 450 \$
3	1 342 828 \$	-799 879 \$	542 949 \$	1 402 \$	26 929 \$
4	1 324 443 \$	-794 297 \$	530 146 \$	526 \$	19 545 \$

(c) Portafoglio composto dal *future* e dall'ETF

<i>op.</i>	<i>\$suC</i>	<i>\$suD</i>	$\Delta \$$	<i>Interessi</i>	<i>profitto</i>
1	-1 356 250 \$	1 299 345 \$	-56 905 \$	-136 \$	5 422 \$

Tabella 1.20: **Controvalore della posizione e interessi maturati**

dinamica degli interessi maturati alla conclusione dell'operazione.

Misurazione del rischio

Per completare la rilevazione dei risultati ottenuti mediante questo modello è stato misurato il rischio operativo relativo a ciascuna delle posizioni assunte. A tal proposito sono state utilizzate due misure di rischio, il *Value at Risk* e l'*Expected Shortfall*. Il *Value at Risk* è stato calcolato mediante due approcci:

1. Metodo delle Simulazioni Storiche

Questo approccio è stato sviluppato considerando tutte le osservazioni del campione precedenti alla data di apertura della posizione. Con questi dati è stato costruito un *database* contenente n tassi di variazione giornaliera per i prezzi dei titoli che costituiscono il portafoglio. Le variazioni che hanno interessato i due titoli sono state considerate come possibili scenari del giorno successivo a quello in cui si sta effettuando la valutazione. Si dispone così di n scenari alternativi per quel che potrebbe succedere tra 'oggi' e 'domani'. Per esempio lo scenario 1 è quello in cui i tassi di variazione del prezzo dei due titoli sono uguali a quelli registrati nel secondo giorno. Pertanto se p_t è il prezzo di uno dei due titoli alla fine del giorno t con $t = 0 \dots n$ allora il

prezzo che il titolo assumerà domani, ossia alla fine del giorno $n + 1$, è pari nello scenario t a

$$p_n \frac{p_t}{p_{t-1}}.$$

Per ognuno degli n scenari è stata calcolata la variazione di valore del portafoglio. A questo punto si è optato per uno schema di ponderazione delle peso di ogni osservazione. Questa scelta è dovuta alla volontà di dare più peso alle osservazioni più recenti poiché esse riflettono maggiormente le condizioni del mercato in termini di volatilità e di altri fattori di rischio. Lo schema di ponderazione utilizzato è quello in cui i pesi declinano in modo esponenziale. In termini analitici, il peso assegnato al primo scenario (costruito in base alle osservazioni più lontane) è pari a δ ($0 < \delta < 1$) volte il peso assegnato allo scenario 2. A sua volta, il peso assegnato allo scenario 2 è pari a δ volte il peso assegnato allo scenario 3, e così via. Affinché i pesi sommino a 1, il peso attribuito allo scenario t deve essere pari a

$$\frac{\delta^{n-t}(1 - \delta)}{1 - \delta^n}.$$

dove n è il numero degli scenari. Il VaR è stato stimato mettendo in ordine le variazioni di valore del portafoglio (rispetto al valore attuale) relative ai diversi scenari, dalla peggiore alla migliore. Iniziando dalla perdita più elevata, i pesi sono stati sommati fin quando non è stato raggiunto il 1° percentile. La variazione di valore corrispondente all'osservazione selezionata mediante la somma dei pesi rappresenta il VaR all'99%.

2. Metodo Montecarlo

Con questa tecnica è stata generata la distribuzione di probabilità delle variazioni di valore del portafoglio. Per fare questo abbiamo approssimato le variazioni del prezzo dei due titoli mediante un Moto Geometrico Browniano a tempo discreto

$$\begin{aligned} \Delta p_1 &= \mu_1 p_1 \Delta t + \sigma_1 \epsilon_1 \sqrt{t} \\ \Delta p_2 &= \mu_2 p_2 \Delta t + \sigma_2 \epsilon_2 \sqrt{t} \end{aligned} \tag{1.47}$$

dove Δp_1 e Δp_2 rappresentano la variazione del prezzo del titolo 1 e del titolo 2 in un piccolo intervallo di tempo Δt , μ_1 e μ_2 sono i tassi di rendimento attesi

dei titoli, σ_1 e σ_2 sono le volatilità dei due titoli e ϵ_1 ed ϵ_2 rappresentano delle estrazioni casuali da una normale bivariata standardizzata. Nello specifico i valori ϵ_1 ed ϵ_2 sono stati ottenuti mediante la Cholesky *decomposition*, in questo modo

$$\begin{aligned}\epsilon_1 &= x_1 \\ \epsilon_2 &= \rho x_1 + x_2 \sqrt{1 - \rho^2}\end{aligned}\tag{1.48}$$

dove x_1 e x_2 sono dei campioni indipendenti estratti da una normale univariata standardizzata, e ρ è la correlazione tra i rendimenti dei due titoli. I parametri μ , σ e ρ sono stati stimati utilizzando i rendimenti dei due titoli dall'inizio del *formation period* fino alla data dell'operazione. Per valutare il rischio relativo ad ogni posizione, sono state generate 1000 coppie di possibili variazioni $(\Delta p_1, \Delta p_2)$. In corrispondenza di ogni coppia è stata calcolata la variazione di valore del portafoglio rispetto al portafoglio valutato ai prezzi correnti. Successivamente è stata costruita la distribuzione di probabilità di tutte le variazioni rilevate. Il VaR giornaliero ad un livello di confidenza del 99% è pari al 1 percentile di questa distribuzione.

Individuato il VaR giornaliero, per calcolare tale misura su un orizzonte temporale pari all'*holding period* è stata utilizzata la formula

$$VaR \sqrt{h.p.}$$

dove $h.p.$ è l'*holding period* della posizione. Inoltre per calcolare il valore medio delle perdite eccedenti il VaR è stata utilizzata una terza misura di rischio, l'*Expected Shortfall*. Tale misura è stata scelta per tenere in considerazione eventuali accumuli di probabilità sotto la coda sinistra, quindi in corrispondenza di perdite caratterizzate da ammontari estremi.

La Tabella 1.21 riporta le stime per ciascuna delle posizioni aperte, sia quelle ottenute mediante l'approccio Montecarlo (MC_g) che quelle realizzate mediante il metodo delle simulazioni storiche (SS_g). Tali misure sono state calcolate anche per l'*holding period* (h.p.) della posizione ($MC_{h.p.}$, $SS_{h.p.}$), ottenendo una valutazione del rischio per l'intervallo di tempo in cui si attende la *mean reversion* del processo. La Tabella 1.21 riporta inoltre l'*Expected Shortfall* calcolata sulle distribuzioni costruite mediante i due approcci descritti (ES_{MC} , ES_{SS}). Applicando il metodo delle simulazioni storiche, per il VaR delle posizioni aperte nel *trading*

(a) Portafoglio composto dai titoli A e B

<i>op.</i>	MC_g	ES_{MC}	SS_g	ES_{SS}	<i>h.p.</i>	MC_{hp}	SS_{hp}
1	16 342 \$	18 258 \$	17 461 \$	19 167 \$	4	32 684 \$	34 922 \$
2	14 187 \$	16 248 \$	17 017 \$	19 039 \$	4	28 374 \$	34 034 \$
3	15 749 \$	17 944 \$	17 612 \$	18 390 \$	4	31 499 \$	35 223 \$
4	15 779 \$	18 009 \$	17 587 \$	18 368 \$	4	31 559 \$	35 174 \$
5	15 287 \$	17 614 \$	16 426 \$	17 835 \$	4	30 574 \$	32 851 \$
6	15 499 \$	17 853 \$	17 214 \$	17 975 \$	4	30 998 \$	34 429 \$
7	15 014 \$	17 322 \$	16 183 \$	17 413 \$	5	33 572 \$	36 186 \$

(b) Portafoglio composto dai titoli C e D

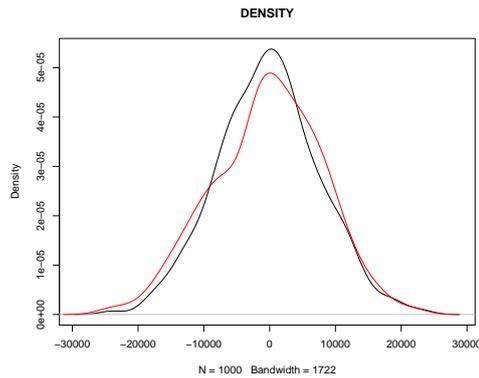
<i>op.</i>	MC_g	ES_{MC}	SS_g	ES_{SS}	<i>h.p.</i>	MC_{hp}	SS_{hp}
1	17 528 \$	20 258 \$	22 129 \$	23 740 \$	10	55 431 \$	69 977 \$
2	17 486 \$	20 187 \$	18 020 \$	20 593 \$	11	57 996 \$	59 766 \$
3	17 513 \$	20 218 \$	17 216 \$	20 461 \$	13	63 145 \$	62 073 \$
4	17 411 \$	20 157 \$	17 120 \$	20 153 \$	14	65 148 \$	61 725 \$

(c) Portafoglio composto dal *future* e dall'ETF

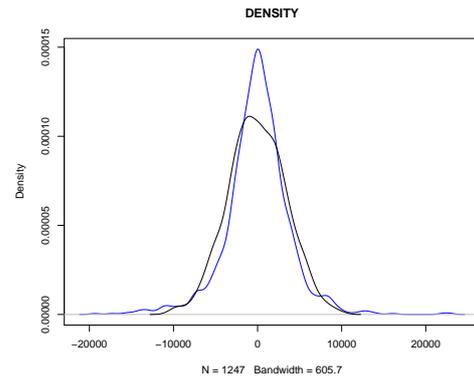
<i>op.</i>	MC_g	ES_{MC}	SS_g	ES_{SS}	<i>h.p.</i>	MC_{hp}	SS_{hp}
1	8 311 \$	9 461 \$	14 007 \$	15 948 \$	12	28 791 \$	48 522 \$

Tabella 1.21: **Valutazioni del rischio assunto su ogni posizione aperta**

period del campione 1, è stato scelto un parametro δ pari a 0.995. In questo modo si è attribuito un peso maggiore alle osservazioni più recenti, privilegiando la loro incidenza rispetto a quelle passate. Per le operazioni effettuate nel *trading period* del campione 2 è stato fatto il contrario, in questo caso infatti, è stato scelto un parametro δ pari a 0.98. Tale scelta è stata dettata dal minor numero di scenari passati a disposizione. In pratica, si è deciso di attribuire un peso maggiore (in misura minima) alle variazioni meno prossime. Dato il maggior numero di scenari passati, per il campione 3 è stato scelto un parametro δ di valore 0.995. Dalla Tabella 1.21 risultano due evidenze particolari. Per quanto riguarda il VaR di ogni operazione, rileviamo una tendenza a sovrastimare il rischio da parte dell'approccio basato sul metodo delle simulazioni storiche. Da questo punto di vista sono particolarmente rappresentative le stime relative all'operazione 1 eseguita nel *trading period* del campione 2 e a quella (l'unica) eseguita nel *trading period* del campione 3. Come mostrato dalla Figura 1.14, la coda sinistra della distribuzione empirica (colore rosso e colore blu), in tutti e due i casi, tende a esser più spessa di quella ottenuta mediante l'approccio Montecarlo. Anche se in misura ridotta tale evidenza tende a riproporsi per ciascuna delle misurazioni effettuate, questo probabilmente è dovuto alla differente numerosità di realizzazioni con cui sono sta-



(a) Prima posizione assunta nel *trading period* del campione 1



(b) Posizione assunta nel *trading period* del campione 3

Figura 1.14: **Distribuzione delle perdite**

te costruite l'una e l'altra distribuzione (ricordiamo che con il metodo Montecarlo le variazioni di valore simulate sono 10000). La seconda evidenza più significativa riguarda l'*Expected Shortfall*. In tal senso, la misurazione effettuata in corrispondenza della prima operazione avvenuta nel *trading period* del campione 2 rileva un notevole differenza tra le due distribuzioni. L'*Expected Shortfall* misurata mediante la distribuzione empirica mostra infatti un accumulo di densità in corrispondenza di perdite caratterizzate da valori più estremi. Pertanto sembra esser stata una scelta avveduta calcolare il valore medio delle perdite eccedenti il VaR utilizzando la distribuzione empirica. Gli scenari passati infatti, sottolineano il rischio che in una situazione negativa, esista una probabilità maggiore rispetto a quanto stimato con il metodo Montecarlo che incorrendo in perdite superiori al VaR, in media queste siano maggiori.

1.6 Implementazione in R

```
library(fBasics)
library(timeDate)
library(timeSeries)
library(zoo)
library(urca)
library(quadprog)
library(lmtest)
library(sandwich)
library(strucchange)
```

```

library(vars)
library(tseries)
library(Sim.DiffProc)
library(sde)
library(depmixS4)
library(PairTrading)
library(distrEx)
library(moments)
library(VaR)
library(fExtremes)

##### GENERAZIONE DEL CAMPIONE 1 E SELEZIONE #####
##### DEL FORMATION PERIOD #####

set.seed(150)
e1 <- rnorm(1306,0,0.5)
e2 <- rnorm(1306,0,0.5)
e3 <- rnorm(1306,0,1)
u1.ar1<-arima.sim(model=list(ar=0.5),innov=e1,n=1306)
u2.ar1<-arima.sim(model=list(ar=0.5),innov=e2,n=1306)
y3<-cumsum(e3)
y1<-0.8*y3+u1.ar1+120
y2<-0.95*y3+u2.ar1+90
#####
y.mat<-data.frame(y1,y2,y3)
y<-log(y.mat[,1:2])
Y<-y.mat[,1:2]
y.f<-y[400:651,]
Y.f<-y.mat[400:651,1:2]
y.t<-y[652:777,]
Y.t<-y.mat[652:777,1:2]
y.ft<-y[400:777,]
Y.ft<-y.mat[400:777,1:2]
Y.ft1<-Y.ft[,1]
Y.ft2<-Y.ft[,2]

#####          GRAFICO DEL CAMPIONE 1          #####
#prezzo

```

```

ts.plot(y1, y2, col = c(1, 2),ylab="DOLLARI",xlab="GIORNI")
legend("topleft", legend = c("titolo A","titolo B"), col = c(1, 2), lty =
c(1),bty="n")
#logaritmo naturale del prezzo
ts.plot(y, col = c(1, 2),ylab="Log",xlab="GIORNI",ylim=c(4.3,5.1))
legend("topleft", legend = c("titolo A","titolo B"), col = c(1, 2), lty =
c(1),bty="n")

##### GENERAZIONE DEL CAMPIONE 2 SELEZIONE #####
##### FORMATION PERIOD E TRADING PERIOD #####
set.seed(171)
sigma <- diag(0.5, 2)
e <- mvrnorm(1306, rep(0, 2), sigma)
u.ar <- arima.sim(1306, model = list(ar = c(0.9)), innov = e[,1])
y2 <- cumsum(e[, 2])+100
trend.det2<-1+0.01*(1:1306)
y1 <- 1.15* y2 + u.ar+trend.det2
y.mat<-cbind(y1,y2)
y.mat<-data.frame(y.mat)
y<-log(y.mat)
Y<-y.mat[,1:2]
y.f<-y[272:523,]
y.t<-y[524:649,]
Y.f<-y.mat[272:523,]
Y.t<- y.mat[524:649,]
y.ft<-y[272:649,]
Y.ft<-y.mat[272:649,1:2]
Y.ft1<-Y.ft[,1]
Y.ft2<-Y.ft[,2]
#####          GRAFICO DEL CAMPIONE 2          #####
# prezzo
ts.plot(y.mat, col = c(1, 2),ylab="DOLLARI",xlab="GIORNI")
legend("topleft", legend = c("titolo C","titolo D"), col = c(1, 2), lty =
c(1),bty="n")
# logaritmo naturale del prezzo
ts.plot(y, col = c(1, 2),ylab="Log",xlab="GIORNI",ylim=c(4.55,5.2))
legend("topleft", legend = c("titolo C","titolo D"), col = c(1, 2), lty =
c(1),bty="n")

```

```

##### TEST DI RADICE UNITARIA E DI #####
##### STAZIONARIETA' EFFETTUATI SULL'INTERO #####
##### CAMPIONE NEI LIVELLI E NELLE DIFFERENZE #####

AD Ftseries<-apply(y.mat[,1:2],2,adf.test)
PPtseries<- apply(y.mat[,1:2],, 2, pp.test)
KPSStseries<- apply(y.mat[,1:2],, 2, kpss.test)
AD Ftseries
PPtseries
KPSStseries
AD Ftseries<-apply(y,2,adf.test)
PPtseries<- apply(y, 2, pp.test)
KPSStseries<- apply(y, 2, kpss.test)
AD Ftseries
PPtseries
KPSStseries
y.mat.diff<-cbind(diff(y.mat[,1],1,1),diff(y.mat[,2],1,1))
y.diff<-cbind(diff(y[,1],1,1),diff(y[,2],1,1))
AD Ftseries<-apply(y.mat.diff,2,adf.test)
PPtseries<- apply(y.mat.diff, 2, pp.test)
KPSStseries<- apply(y.mat.diff, 2, kpss.test)
AD Ftseries
PPtseries
KPSStseries
AD Ftseries<-apply(y.diff,2,adf.test)
PPtseries<- apply(y.diff, 2, pp.test)
KPSStseries<- apply(y.diff, 2, kpss.test)
AD Ftseries
PPtseries
KPSStseries

##### GRAFICO DEL FORMATION PERIOD campione 1 #####
#logaritmo naturale del prezzo
ts.plot(Y.f, col = c(1, 2),ylab="DOLLARI",xlab="GIORNI", ylim=c(85,140))
legend("topleft", legend = c("titolo A","titolo B"), col = c(1, 2), lty =
c(1),bty="n")
# prezzo

```

```
ts.plot(y.f,col = c(1,2),ylab="Log",xlab="GIORNI",ylim=c(4.4,4.95))
legend("topleft", legend = c("titolo A","titolo B"), col = c(1, 2), lty =
c(1),bty="n")
```

```
##### GRAFICO DEL FORMATION PERIOD campione 2 #####
#logaritmo naturale del prezzo
ts.plot(y.f,col = c(1,2),ylab="DOLLARI",xlab="GIORNI",ylim=c(4.6,4.95))
legend("topleft", legend = c("titolo C","titolo D"), col = c(1, 2), lty =
c(1),bty="n")
```

```
#prezzo
ts.plot(Y.f, col = c(1, 2),ylab="Log",xlab="GIORNI", ylim=c(100,141))
legend("topleft", legend = c("titolo C","titolo D"), col = c(1, 2), lty =
c(1),bty="n")
```

```
##### TEST DI RADICE UNITARIA E DI STAZIONARIETA' #####
##### EFFETTUATI SUL FORMATION PERIOD #####
ADFTseries<-apply(y.f,2,adf.test)
PPtseries<- apply(y.f, 2, pp.test)
KPSStseries<- apply(y.f, 2, kpss.test)
ADFTseries
PPtseries
KPSStseries
```

```
##### SELEZIONE DI UN RITARDO PER LA #####
##### SPECIFICAZIONE DI UN MODELLO VAR #####
```

```
VARselect(y.f, lag.max = 12, type="const")
```

```
##### SPECIFICAZIONE DEL MODELLO #####
VAR.c<-VAR(y.f,p=2,type="const")
```

```
##### ANALISI DELLA SPECIFICAZIONE #####
```

```
s.test.PT<-serial.test(VAR.c,lags.bg=4)
s.test.ES<-serial.test(VAR.c,lags.bg=4,type="ES")
s.test.BG<-serial.test(VAR.c,lags.bg=4,type="BG")
a.test<-arch.test(VAR.c,lags.multi=5)
s.test.PT
```

```
s.test.ES
s.test.BG
a.test
```

```
##### ANALISI DELLA ACF E PACF DEI RESIDUI #####
```

```
residuals<-resid(VAR.c)
e_1<-residuals[,1]
e_2<-residuals[,2]
e_1_2<-(e_A)^2
e_2_2<-(e_B)^2
acf(e_1,lag.max=20)
acf(e_2,lag.max=20)
acf(e_1_2,lag.max=20)
acf(e_2_2,lag.max=20)
pacf(e_1,lag.max=20)
pacf(e_2,lag.max=20)
pacf(e_1_2,lag.max=20)
pacf(e_2_2,lag.max=20)
```

```
##### TEST DI JOHANSEN #####
```

```
vecm.eigen.c <- ca.jo(y.f, ecdet = "const", type=c("eigen"), K=2,
spec="transitory")
summary(vecm.eigen.c)
##### ANALISI DEI RESIDUI
plotres(vecm.eigen.c)
```

```
##### STAZIONARIZZAZIONE DEL PROCESSO (formation period) #####
```

```
bc<-vecm.eigen.c@V[,1]
const<-rep(1,252)
Yc<-cbind(y.f,const)
bc.Yc<-t(bc%*%t(Yc))
mean(bc.Yc)
sd(bc.Yc)
m_bc.Yc<-ts(rep(mean(bc.Yc),252))
```

```
##### TEST DI RADICE UNITARIA E DI STAZIONARIETA' #####
```

```
##### EFFETTUATI SUL FORMATION PERIOD #####
##### STAZIONARIZZATO #####
```

```
summary(ur.df(bc.Yc, type = c("drift"), selectlags = c("AIC")))
summary(ur.kpss(bc.Yc))
summary(ur.pp(bc.Yc, use.lag=4))
```

```
##### COSTRUZIONE DI 2 THRESHOLDS #####
```

```
# APPROCCIO PERCENTILI
m<-rep(mean(bc.Yc),length(bc.Yc_2))
ord<-sort(bc.Yc)
oss95percento<-0.05*length(bc.Yc)
oss5percento<-0.95*length(bc.Yc)
oss95percento
oss5percento
percentili<-quantile(bc.Yc, c(.05, .95))
perc5<-percentili[1]
perc95<-percentili[2]
u.bp<-rep(perc95,length(bc.Yc_2))
l.bp<-rep(perc5,length(bc.Yc_2))
ts.plot(cbind(bc.Yc_2,m,u.bp,l.bp))
```

```
# APPROCCIO STANDARD DEVIATION
vol<- apply(bc.Yc, 2, sd)
up<-mean(bc.Yc)+2*vol
lw<-mean(bc.Yc)-2*vol
u.bs<-rep(up,length(bc.Yc_2))
l.bs<-rep(lw,length(bc.Yc_2))
ts.plot(cbind(bc.Yc_2,m,u.bs,l.bs))
```

```
##### STAZIONARIZZAZIONE DEL PROCESSO (trading period) #####
```

```
bc<-vecm.eigen.c@V[,1]
const<-rep(1,126)
Yc<-cbind(y.t,const)
bc.Yc_2<-t(bc%*%t(Yc))
plot(bc.Yc_2,type="l")
```

```

##### STAZIONARIZZAZIONE DEL PROCESSO #####
##### (trading period + formation period) #####
bc<-vecm.eigen.c@V[,1]
const<-rep(1,378)
Yc<-cbind(y.ft[1:378,],const)
bc.Yc_12<-t(bc%*%t(Yc))

#####
##### ANALISI ROLLING #####
#####

##### TRADING PERIOD
g<-length(bc.Yc_2)
sss<-rep(0,126)
for (i in 2:g){
j= bc.Yc_2[1:i]
sss[i]=sd(j)
}
mmm<-rep(0,126)
for (i in 1:g){
mmm[i]=mean(bc.Yc_2[1:i])
}
mm<-as.numeric(mmm)
ss<-as.numeric(sss)
sm<-mm+2*ss
ms<-mm-2*ss
f<- ts(bc.Yc_2)
ts.plot(f,ms,sm,mm)

##### FOORMATION PERIOD + TRADING PERIOD #####
##### funzione per media rolling #####
mr=function(X){
b<-mean(X)
theta=list(output=b)
return(theta)
}
##### funzione per std rolling

```

```

sdr=function(X){
a<-sd(X)
theta=list(output=a)
return(theta)
}
u<-length(bc.Yc_12)
rsd<-rep(0,(u-252))
rm<-rep(0,(u-252))
##### loop per la media
for (i in 252:u){
j= bc.Yc_12[(i-251):i]
mr(j)
rm[i-251]=mr(j)
}
RM<-as.numeric(rm[2:127])
##### loop per la std
for (i in 252:u){
j= bc.Yc_12[(i-251):i]
sdr(j)
rsd[i-251]=sdr(j)
}
RSD<-as.numeric(rsd[2:127])
#### GRAFICO TRADING PERIOD
rup<-(RM+2*RSD)
rlw<-(RM-2*RSD)
ts.plot(cbind(bc.Yc_2,rup,rlw,RM))
axis(1, c(10,20,30,40,50,60,70,80,90,100,110,120))
#### GRAFICO FORMATION PERIOD + TRADING PERIOD
U.BS<-rep(0,378)
L.BS<-rep(0,378)
M<-rep(0,378)
U.BS[1:252]=u.bs
L.BS[1:252]=l.bs
M[1:252]=m
U.BS[253:378]=rup
L.BS[253:378]=rlw
M[253:378]=RM
ts.plot(cbind(bc.Yc_12,U.BS,L.BS,M))

```

```

axis(1,c(20,40,60,80,100,120,140,160,180,200,220,240,260,280,300,320,340,
360))
#### MATRICE SEGNALI OPERATIVI
t.v1<-cbind(Y.t,y.t,bc.Yc_2,rup,rlw,RM)
t.v1

##### FITTING ORNSTEIN-ULEHNBECK PROCESS #####

### MINIMI QUADRATI
ouFit.LS=function(spread,dt) {
n=length(spread)
x=spread[1:(n-1)]
y=spread[2:n]
spread.fit=lm(y~x)
coefs=as.numeric(coefficients(spread.fit))
a=coefs[1]
b=coefs[2]
variance=var(residuals(spread.fit))
lambda=(-log(b))/dt
mu=a/(1-b)
sigma=sqrt(variance*(2*b/(1-exp(-2*lambda*dt))))
theta=list(lambda=lambda, mu=mu,sigma=sigma)
return(theta)
}
ouFit.LS(bc.Yc,1)

### MAXIMUM LIKELYHOOD
ouFit.ML = function(spread, delta) {
n = length(spread)
#delta=n/n # this always evaluates to 1 ???
s1 = spread[1:(n-1)]
s2 = spread[2:n]
Sx = sum(s1)
Sy = sum(s2)
Sxx = sum(s1^2)
Syy = sum(s2^2)
Sxy = sum(s1*s2)
mu =(Sy*Sxx - Sx*Sxy) / ( n*(Sxx - Sxy) - (Sx^2 - Sx*Sy) )

```

```

lambda = -log( (Sxy - mu*Sx - mu*Sy + n*mu^2) / (Sxx -2*mu*Sx +
n*mu^2))/delta
a = exp(-lambda*delta)
sigmah2 = (Syy - 2*a*Sxy + a^2*Sxx - 2*mu*(1-a)*(Sy - a*Sx) +
n*mu^2*(1-a)^2)/n;
sigma = sqrt((sigmah2)*2*lambda/(1-a^2))
list(lambda=lambda, mu=mu,sigma=sigma,sigmah2=sigmah2)
}
ouFit.ML(bc.Yc_12,1)

```

```

#####                HOLDING PERIOD                #####
# minimi quadrati
H.LS<- log(2)/ouFit.LS(bc.Yc,1)$lambda
L.LS<-2*H.LS
L.LS
# mle
H.ML<- log(2)/ouFit.ML(bc.Yc,1)$lambda
L.ML<-2*H.ML
L.ML

```

```

##### ANALISI ROLLING DELLA RELAZIONE                #####
##### DI COINTEGRAZIONE                #####

```

```

vec=function(X){
vecm<-ca.jo(X, ecdet = "const", type=c("eigen"), K=2, spec="transitory")
theta=list(vc=as.vector(vecm@V[,1]),st=vecm@cval[2,1:3],tstat=vecm@teststat[2])
return(theta)
}
u<-length(y[,1])
b1<-rep(0,(u-252))
b2<-rep(0,(u-252))
c<-rep(0,(u-252))
stat1<-rep(0,(u-252))
st1<-rep(0,(u-252))
st2<-rep(0,(u-252))
st3<-rep(0,(u-252))
for (i in 252:u){

```

```

j=y[((i-251):i),(1:2)]
vec(j)
b1[(i-251)]=vec(j)$vc[1]
b2[i-251]=vec(j)$vc[2]
c[i-251]=vec(j)$vc[3]
stat1[i-251]=vec(j)$tstat
st1[i-251]=vec(j)$st[1]
st2[i-251]=vec(j)$st[2]
st3[i-251]=vec(j)$st[3]
}
panel.coint<-cbind(b1,b2,c,stat1,st1,st2,st3)
panel.coint
s1<-ts(st1)
s2<-ts(st2)
s3<-ts(st3)
s<-ts(stat1)
graph<-cbind(s1,s2,s3,s)
ts.plot(graph)

```

```

##### ANALISI ROLLING DELL'ORDINE DI RITARDO #####
##### SECONDO AIC, HQ, SC, FPE #####
VARsel=function(X){
a<-VARselect(X, lag.max = 12, type="const")
theta=list(output=a$selection)
return(theta)
}
u<-length(y[,1])
AIC<-rep(0,(u-252))
HQ<-rep(0,(u-252))
SC<-rep(0,(u-252))
FPE<-rep(0,(u-252))
for (i in 252:u){
j=y[((i-251):i),(1:2)]
VARsel(j)
AIC[(i-251)]=VARsel(j)$output[1]
HQ[(i-251)]=VARsel(j)$output[2]
SC[(i-251)]=VARsel(j)$output[3]
FPE[(i-251)]=VARsel(j)$output[4]
}

```



```

for(i in 1:p-1){
scenari_2[i]=day_2*v2[i]
}

#####          valore portafoglio negli scenari          #####

n_1<-10000
n_2<--7310
v_p<-n_1*day_1+n_2*day_2
v_p
v_ps<-rep(0,p-1)
for(i in 1:p-1){
v_ps[i]=-n_1*scenari_1[i]+(-n_2)*scenari_2[i]
}

losses<-v_p+v_ps
peso=function(i){
l=0.98
nn=p-1
w=((1^(nn-i))*(1-l))/(1-l^n)
theta=list(l1=w)
return(theta)
}

weights<-rep(0,p-1)
for(i in 1:(p-1)){
weights[i]=peso(i)
}
pesi<-as.numeric(weights)

sp<-cbind(losses,pesi)
sort_sp<- sp[order(losses, pesi), ]
sum_pesi<-rep(0,p-1)

ssp<-sort_sp[,2]
for(i in 1:(p-1)){
sum_pesi[i]=sum(ssp[1:i])
}

```

```

}

#####                               VaR HS                               #####
SSSP<-cbind(sort_sp,sum_pesi)
SSSP

#####                               RENDIMENTI E sd DI A e B                               #####

Y.fto1<-Y.ft[1:ca,1]
Y.fto2<-Y.ft[1:ca,2]
do<-length(Y.fto1)

r_1<-rep(0,do)
for(n in 2:do){
r_1[n]<-log(Y.fto1[n]/Y.fto1[n-1])
}

mr_1<-mean(r_1)
sd_1<-sd(r_1)
var_1<-sd_1^2

r_2<-rep(0,do)
for(n in 2:do){
r_2[n]<-log(Y.fto2[n]/Y.fto2[n-1])
}
mr_2<-mean(r_2)
sd_2<-sd(r_2)
var_2<-sd_2^2

##### rendimento medio, std, var, cov, cor #####

r12<-cbind(r_1,r_2)
sd_12<-cbind(sd_1,sd_2)
rho<-cor(r_1, r_2, method = "pearson")
rho
cov12<-rho*sd_1*sd_2

```

```
VCV<-matrix(c(var_1, cov12, cov12 ,var_2),2,2)
VCV<-matrix(c(var_1, cov12, cov12 ,var_2),1,3)
VCV
```

```
##### METODO MONTECARLO #####
```

```
##### SIMULAZIONE MOTO GEOMETRICO BROWNIANO #####
```

```
bmSim=function(X,K,d,v,r1,r2,sigma1,sigma2,q,dt) {
  n=length(X)
  x0=X[n]
  y0=K[n]
  sx2<-(sigma1^2)
  sy2<-(sigma2^2)
  x=matrix(x0,d,v)
  y=matrix(y0,d,v)
  dx=matrix(0,d,v)
  dy=matrix(0,d,v)
  e1=matrix(0,d,v)
  e2=matrix(0,d,v)
  e3=matrix(0,d,v)
  set.seed(1233)
  for(t in 1:v){
    for(h in 1:d)
    e1[h,t]=rnorm(1,mean=0,sd=1)}
  for(t in 1:v){
    for(h in 1:d)
    e3[h,t]=rnorm(1,mean=0,sd=1)}
  for(t in 1:v){
    for(h in 1:d)
    e2[h,t]= e1[h,t]*q+e3[h,t]*(1-q^2)^0.5}
  for(t in 1:v){
    for(h in 2:d)
    dx[h,t]=x[h-1,t]*r1*dt+sigma1*x[h-1,t]*sqrt(dt)*e1[h,t]}
  for(t in 1:v){
    for(h in 2:d)
    x[h,t]=(x[h-1,t]+ dx[h-1,t])}
```

```

for(t in 1:v){
  for(h in 2:d)
dy[h,t]=y[h-1,t]*r2*dt+sigma2*y[h-1,t]*sqrt(dt)*e2[h,t]
}
for(t in 1:v){
  for(h in 2:d)
y[h,t]=(y[h-1,t]+ dy[h-1,t])}

theta=list(S0=x0,G0=y0,dx=dx,dy=dy,e1=e1,e2=e2,e3=e3,x=x,y=y)
  return(theta)
}
#####

n_s<-1000 #<<----- NUMERO DELLE SIMULAZIONI

simulazioni<-bmSim(Y.fto1,Y.fto2,3,n_s,mr_1,mr_2,sd_1,sd_2,rho,1)

simulazioni_1<-simulazioni$x[3,1:n_s]
simulazioni_2<-simulazioni$y[3,1:n_s]

p<-(length(Yo1))
day_1<-Y[(p),1]
day_2<-Y[(p),2]
v_p<-n_1*day_1+n_2*day_2
v_p

p_vs<-rep(0,n_s)
for(i in 1:n_s){
p_vs[i]=(-n_1)*simulazioni_1[i]+(-n_2)*simulazioni_2[i]
}

losses_mc<-v_p+p_vs #<<-- vettore delle perdite con MC

VAR_mc<-quantile(losses_mc, c(0.01))
VAR_mc
VAR_mc*sqrt(hp)

```

```

qp<-quantile(losses_mc, c(0.01))
d<-ifelse(losses_mc<(qp),losses_mc,0)
n<-ifelse(losses_mc<(qp),1,0)
ES.mc<-(sum(d)/sum(n))
ES.mc #<-expected shortfall con montecarlo

###qp<-quantile(losses, c(0.01))
qp<- <---- inserisci VaR calcolato con il metodo delle simulazioni storiche
d<-ifelse(losses<(qp),losses,0)
n<-ifelse(losses<(qp),1,0)
ES.p<-(sum(d)/sum(n))
ES.hs #<-expected shortfall con simulazioni storiche

plot(hist(losses))
plot(density(losses_mc),main="DENSITY")
lines(density(losses),col=c(2))

```

Conclusioni

Alla luce del diffondersi di strategie di mercato basate sull'econometria applicata e su particolari relazioni tra i prezzi di asset finanziari, questo lavoro ha voluto fornire un contributo all'analisi della profittabilità e dell'applicabilità di uno di questi modelli. Lo studio si è focalizzato sul ruolo della relazione di cointegrazione nello sviluppo di una strategia di *Pair Trading*. Si è fatto questo perseguendo un duplice obiettivo. Da un lato è stata curata la costruzione di un vero e proprio modello operativo basato su strumenti di econometria finanziaria. Dall'altro questo modello è stato messo alla prova valutandone aspetti che vanno oltre la capacità di ottenere risultati positivi. Pertanto nel corso della trattazione è stato illustrato in modo sequenziale l'utilizzo degli strumenti econometrici necessari alla costruzione della strategia. Mediante tecniche di simulazione apposite sono stati generati due campioni idonei ai nostri scopi, un terzo è stato ottenuto da quotazioni reali. Su questi sono stati implementati test econometrici che ci hanno consentito di individuare la presenza relazioni di cointegrazione tra le serie contenute in ogni campione. Stimato il vettore di cointegrazione ne è stata illustrata l'applicabilità nell'ambito di una strategia di *Pair Trading*. I dati raccolti in questo modo ci hanno consentito di stimare dei parametri utili a definire delle vere e proprie regole operative su cui basare l'apertura e la chiusura di posizioni. Definite tali regole, il modello è stato applicato sui dati simulati e sul campione reale, procedendo ad un *Back Testing* della strategia. La valutazione ha condotto ad evidenze interessanti. Dal punto di vista della gestione del portafoglio il modello ha ottenuto risultati positivi. Nonostante la sua applicazione sia avvenuta in un contesto privo di frizioni tipiche del mercato reale, la strategia è stata in grado di ottenere un rendimento positivo, con rischi accettabili. Tuttavia il software econometrico utilizzato nella fase di sviluppo ha permesso e indotto un'ulteriore considerazione. Infatti al netto di quelli economici, i risultati dell'applicazione da un punto di vista econometrico sollevano problematiche di non poco conto. In tal senso, i profitti ottenuti nel secondo e nel terzo campione non sono stati realizzati tramite il risultato econometrico più importante, ossia la *mean reversion* del processo ottenuto sfruttando la relazione di cointegrazione. In questi due casi l'identificazione del vettore di cointegrazione

non è stata sufficiente a ottenere un modello operativo automatico che funzionasse in modo fedele allo schema teorico fondante.

Se da un lato quest'ultima evidenza fornisce lo stimolo per un miglioramento dello schema operativo dall'altro ci ripone davanti all'interrogativo in merito alla capacità di comprendere e "battere" le dinamiche (potremmo dire sempre meno governabili) del mercato reale. Nonostante l'utilizzo di strumenti econometrici possa garantire una credibilità più consistente rispetto a strategie basate sul *sentiment* o su altri fattori scarsamente teorizzabili (o se teorizzabili più semplicistici), più volte nello sviluppo del modello si è dovuto prendere atto della scarsa potenza di alcuni test o affidarsi alla stabilità di alcune relazioni. Pertanto il risultato fondamentale è dato dalla completa teorizzazione del modello e dall'identificazione di alcuni punti critici che, alla luce dello stato dell'arte di alcune tecniche di stima, possono costituire il punto di partenza per una sua evoluzione.

Appendice A

Appendice Teorica

A.1 Processo stocastici stazionari e integrati

A.1.1 Stazionarietà

Un processo stocastico y_t è detto stazionario se il suo momento primo e il suo momento secondo sono costanti nel tempo. In altre parole, y_t è stazionario se

- $E(y_t) = \mu_y$ per ogni $t \in T$
- $[(y_t - \mu_y)(y_{t-h} - \mu_y)] = \gamma_h$ per ogni $t \in T$ e tutti gli interi h tali che $t-h \in T$.

La prima condizione indica che tutti i membri di un processo stocastico stazionario hanno la stessa media costante. Quindi una serie storica generata da un processo stocastico stazionario deve fluttuare intorno a un valore medio costante e non deve avere un trend, per esempio. La seconda condizione assicura che la varianza sia costante perché, per $h = 0$, la varianza $\sigma_y^2 = E[(y_t - \mu_y)^2] = \gamma_0$ non dipende da t . Inoltre le covarianze $[(y_t - \mu_y)(y_{t-h} - \mu_y)] = \gamma_h$ non dipendono da t , ma soltanto dalla distanza tra h e t ossia dall'ordine del ritardo. In altre parole i primi due momenti e i momenti incrociati esistono finiti. Se y_t è stazionario possiamo scrivere $y_t \sim I(0)$.

A.1.2 White Noise

Un *white noise* è un processo stocastico che possiede momenti finiti (almeno) fino al secondo ordine; essi sono costanti nel tempo (quindi il processo è stazionario), ma non danno al processo alcuna memoria di sé, quindi, si tratta di un processo stocastico che non esibisce persistenza. In modo più formale:

una serie storica ε_t è definita *white noise* se

- $E(\varepsilon_t) = 0$
- $E(\varepsilon_t^2) = \sum_{i=k}^T (\varepsilon_t)^2 = \sigma^2$
- $\gamma_k = T^{-1} \sum_{i=k}^T (\varepsilon_t)(\varepsilon_{t-k}) = 0$ per $|k| > 0$

Un *white noise* è quindi, un processo composto di un numero infinito di variabili casuali a media zero e varianza costante. Tali variabili sono tutte incorrelate l'una all'altra. Tuttavia questo non significa che esse siano indipendenti. Se però la distribuzione congiunta di tutte le coppie $(\varepsilon_t, \varepsilon_{t-k})$ è una normale bivariata allora tutte le variabili saranno indipendenti. In quest'ultimo caso si parla di *white noise* gaussiano. Del pari, la definizione di un *white noise* multivariato è questa: chiamiamo in questo modo un processo ε_t tale per cui

- $E(\varepsilon_t) = 0$
- $\Gamma_k(\varepsilon_t)(\varepsilon_{t-k}) = 0$ per $|k| > 0$

La definizione di *white noise* multivariato rappresenta quindi il caso generale della definizione di *white noise* univariato. Va notato, peraltro, che Γ è una matrice di varianze e covarianze generica, e pertanto simmetrica e semidefinita positiva, ma non necessariamente diagonale. Di conseguenza, il fatto che un processo multivariato sia un *white noise* esclude la correlazione fra ogni elemento del processo e la storia passata di tutto il processo, ma non esclude che possa esserci correlazione fra elementi contemporanei.

A.1.3 Processi a radice unitaria e non stazionari

Un processo $I(1)$ (integrato di ordine 1) è un processo non stazionario di cui è stazionaria la differenza prima. Più in generale, si definisce come processo $I(d)$ un processo la cui differenza d -esima è stazionaria.

Serie Trend Stationary

Consideriamo il processo

$$y_t = \alpha + \beta t + u_t \tag{A.1}$$

dove u_t è un qualsiasi processo stocastico stazionario a media o . Il processo y_t descritto non è un processo stazionario, poiché $E(y_t) = \alpha + \beta t$, e quindi la media di y_t non è costante (per $\beta \neq 0$). Tuttavia, la non stazionarietà del processo è limitata

a questo aspetto. Se si considerano le variazioni dal *trend*, quel che rimane è un processo stazionario, che può essere analizzato con le tecniche viste qualche pagina fa. È per questo motivo che i processi di questo tipo vengono denominati processi stazionari intorno ad un *trend*, o processi TS (dall'inglese *Trend Stationary*).

Random Walk

w_t è un *random walk* se Δy_t è un *white noise*. Per questo processo vale la relazione $w_t = w_{t-1} + \varepsilon_t$, di conseguenza sostituendo ripetutamente i valori passati di w_{t-1} si ha

$$w_t = w_{t-n} + \sum_{i=0}^{n-1} \varepsilon_{t-i} \quad (\text{A.2})$$

supponendo che in $t = 0$ il processo valga 0 può essere riscritto come

$$w_t = \sum_{i=1}^t \varepsilon_i \quad (\text{A.3})$$

Ad ogni istante t , la media del processo è 0 ma la varianza dipende dal tempo. Tale processo infatti, è costituito dalla somma di t variabili aleatorie indipendenti ed identiche con varianza pari a σ^2 ; ne consegue che la varianza di w_t è $t\sigma^2$, e quindi cresce nel tempo. Inoltre si può dimostrare che la sua covarianza $\gamma_k = \sigma^2 \min(t, s)$. Consegue che w_t non è stazionario. Talvolta questo processo si caratterizza per la presenza di un *trend* deterministico, si parla in questo caso di *random walk* con *drift*. La sua rappresentazione diventa

$$w_t = \sum_{i=0}^t \varepsilon_{t-i} + \mu t \quad (\text{A.4})$$

dove μ è una costante e il secondo termine della somma non è che un *trend* lineare con pendenza μ . Se il *drift*, cioè la costante μ , è positivo, si avrà un processo che tende a salire, ma con fluttuazioni intorno a questo *trend* via via più marcate al passare del tempo. In questo caso si ha che

$$\Delta w_t = \mu + \varepsilon_t \quad (\text{A.5})$$

mentre in casi più in generali si ha

$$\Delta w_t = d_t + \varepsilon_t \quad (\text{A.6})$$

dove d_t è una qualche funzione deterministica del tempo, per esempio $d_t = \delta_0 + \delta_1 t$.

A.1.4 Test di radice unitaria e di stazionarietà

I test per le radici unitarie fanno riferimento ad un sistema d'ipotesi differente rispetto a quello usato dai test per la stazionarietà. In particolare per i test di radice unitaria:

$$H_0 : \phi = 1$$

$$H_1 : |\phi| < 1$$

Quindi sotto l'ipotesi nulla il polinomio caratteristico di un processo AR(1): $\phi(z) = (1 - \phi z) = 0$, ha una radice unitaria¹. Mentre i test di stazionarietà utilizzano il seguente sistema di ipotesi:

$$H_0 : |\phi| < 1$$

$$H_1 : \phi = 1$$

Le statistiche utilizzate per testare questi due sistemi di ipotesi, hanno una distribuzione nulla asintotica non standard e non normale, per tale motivo i loro valori critici vengono solitamente tabulati tramite simulazione.

Augmented Dickey Fuller test

Consideriamo il processo y_t caratterizzato da una struttura ARMA. Per verificare l'esistenza di radici unitarie nella parte AR(p) testiamo le ipotesi $H_0 = \beta = 1$ contro $H_1 = \beta < 1$ usando la regressione

$$y_t = c_t + \beta y_{t-1} + \sum_{i=1}^{p-1} \phi_i \Delta y_{t-i} + \varepsilon_t \quad (\text{A.7})$$

¹Per la stazionarietà di un AR(1) si veda Tsay (2010: capitolo 2)

dove c_t è una funzione deterministica del tempo e $\Delta y_t = y_t - y_{t-1}$ è la serie di y_t nelle differenze. In pratica, c_t può essere 0, una costante o $c_t = b_0 + b_1 t$. Il t ratio di $\widehat{\beta} - 1$,

$$ADF - test = \frac{\widehat{\beta} - 1}{std(\widehat{\beta})}$$

dove $\widehat{\beta}$, che indica la stima di β ottenuta con il metodo dei minimi quadrati, è l'*augmented* DickeyFuller (ADF). Inoltre la differenza prima dell'Eq. A.7 è equivalente a un modello AR(p) con una funzione deterministica c_t . Allora l'Eq. A.7 può essere riscritta come

$$\Delta y_t = c_t + \beta_c y_{t-1} + \sum_{i=1}^{p-1} \phi_i \Delta y_{t-i} + \varepsilon_t$$

dove $\beta_c = \beta - 1$. In questo caso si possono testare le ipotesi $H_0 : \beta_c = 0$ contro $H_1 : \beta_c < 0$, equivalenti a quelle della forma precedente. In tutte e due le forme β è un parametro da stimare; le p variabili differenziate ritardate, Δy_{t-i} , descrivono la struttura degli errori del processo, il valore di p è scelto in modo da rendere serialmente incorrelati gli errori ε_t , che sono anche assunti omoschedastici. Un p troppo piccolo genererà una distorsione del test dovuta alla rimanente correlazione seriale degli errori, mentre un valore elevato di p ridurrebbe la potenza del test. Un metodo pratico per stimare il valore di p , suggerito da Schwert (1989)

$$p_{max} = \left[12 \left(\frac{T}{100} \right)^{\frac{1}{4}} \right] \quad (A.8)$$

dove T rappresenta la numerosità del campione.

Phillips Perron test

Il test di Phillips e Perron (1988) per saggiare l'ipotesi nulla di presenza di radici unitarie, si basa sulla seguente relazione:

$$y_t = \beta c_t + \phi y_{t-1} + u_t \quad (A.9)$$

dove u_t è un processo $I(0)$, senza alcuna assunzione restrittiva sulla sua omoschedasticità. In pratica, la principale differenza tra il test ADF ed il test PP

riguarda l'assunzione di omoschedasticità degli errori. Il test ADF considera una struttura ARMA nell' Eq. A.7 per spiegare la dinamica degli errori assumendoli omoschedastici, mentre il test PP nell'equazione ignora qualsiasi tipo di correlazione seriale negli errori con possibilità di eteroschedasticità. Un vantaggio del test di Phillips Perron rispetto al test ADF è dato dalla maggior robustezza del test PP in merito all'ipotesi di eteroschedasticità degli u_t , inoltre non è condizionato dalla scelta del numero di variabili *lagged* da inserire. Le statistiche test utilizzate nel test PP sono due, che vengono utilizzate in base alla presenza o meno della componente deterministica inserita nella Eq. A.9.

Phillips Perron test

Il test di Kwiatkowsky-Phillips-Schmidt-Shin Test (KPSS) (1992) è un test di stazionarietà, l'ipotesi nulla è la stazionarietà del processo contro l'alternativa che la serie sia non stazionaria o che abbia delle radici unitarie. Questa statistica si basa sul seguente modello

$$\begin{aligned} y_t &= \beta c_t + w_t + u_t \\ w_t &= w_{t-1} + \varepsilon_t \end{aligned} \tag{A.10}$$

dove $\varepsilon_t \sim WN(0, \sigma^2)$, c_t indica la parte deterministica (costante o costante più *trend*), u_t è un processo $I(0)$ che può essere eteroschedastico, mentre w_t è un processo *random walk* e β un parametro da stimare.

Nello specifico, il test saggia il sistema di ipotesi

$$\begin{aligned} H_0 &: \sigma_\varepsilon^2 = 0 \\ H_1 &: \sigma_\varepsilon^2 > 0 \end{aligned} \tag{A.11}$$

Sotto H_0 , si ha che y_t è un processo stazionario in quanto w_t è una costante. La statistica test del KPSS è il moltiplicatore di Lagrange (LM) utilizzato per saggiare il sistema dell'Eq. A.11. Nel loro studio gli autori dimostrarono che il test KPSS converge ad una funzione di un moto Browniano standard che dipende dalla forma dei termini deterministici c_t ma non dai valori dei loro β . Il test KPSS è un test sulla coda destra della sua distribuzione, cosicché l'ipotesi nulla di stazionarietà è rigettata al livello $100 - \alpha$ % se la statistica test KPSS è maggiore del $100(1 - \alpha)$ % quantile della sua distribuzione.

A.2 Modelli Vettoriali

A.2.1 Modello Vettoriale Autoregressivo

Un *vector autoregressive model* (VAR) è un semplice modello vettoriale utile nella modellizzazione delle serie storiche. Una serie storica multivariata \mathbf{y}_t è un processo VAR di ordine 1 (VAR(1)) se è rappresentabile mediante questo modello

$$\mathbf{y}_t = \boldsymbol{\phi}_0 + \mathbf{y}_{t-1}\boldsymbol{\Phi}_{t-1} + \mathbf{a}_t \quad (\text{A.12})$$

dove $\boldsymbol{\phi}_0$ è un vettore $k \times 1$, $\boldsymbol{\Phi}$ è una matrice $k \times k$ e $\{\mathbf{a}_t\}$ è una sequenza di vettori casuali serialmente incorrelati con media 0 e matrice di covarianza $\boldsymbol{\Sigma}$. Nelle applicazioni, la matrice di covarianza $\boldsymbol{\Sigma}$ è richiesta definita positiva altrimenti, la dimensione di \mathbf{y}_t può essere ridotta. In letteratura ricorre l'assunzione secondo cui \mathbf{a}_t è una normale multivariata. Consideriamo il caso bivariato, vale a dire con $k = 2$, $\mathbf{y}_t = (y_{1t}, y_{2t})'$ e $\mathbf{a}_t = (a_{1t}, a_{2t})'$. Allora il modello VAR(1) può essere rappresentato mediante queste due equazioni:

$$\begin{aligned} y_{1t} &= \phi_{10} + \Phi_{11}y_{1,t-1} + \Phi_{12}y_{2,t-1} + a_{1t} \\ y_{2t} &= \phi_{20} + \Phi_{21}y_{1,t-1} + \Phi_{22}y_{2,t-1} + a_{2t} \end{aligned} \quad (\text{A.13})$$

dove Φ_{ij} è la i, j elemento di $\boldsymbol{\Phi}$ e ϕ_{i0} i -esimo elemento di $\boldsymbol{\phi}_0$. Nella prima equazione il termine Φ_{12} illustra la dipendenza lineare di $y_{1,t}$ da $y_{2,t-1}$ in presenza di $y_{1,t-1}$. Pertanto, Φ_{12} è l'effetto condizionale di $y_{2,t-1}$ su $y_{1,t}$ dato $y_{1,t-1}$. Se $\Phi_{12} = 0$, allora y_{1t} non dipende da $y_{2,t-1}$, e il modello mostra che $y_{1,t}$ dipende solo dal suo passato. Allo stesso modo se $\Phi_{21} = 0$, allora la seconda equazione mostra che $y_{2,t}$ non dipende da $y_{1,t-1}$, ma solo da $y_{2,t-1}$. Considerando le due equazioni congiuntamente, se $\Phi_{12} = 0$ e $\Phi_{21} \neq 0$ allora esiste una relazione indiretta tra $y_{1,t}$ e $y_{2,t}$. Se $\Phi_{12} = \Phi_{21} = 0$ allora $y_{1,t}$ e $y_{2,t}$ non hanno un legame. Se $\Phi_{12} \neq 0$ e $\Phi_{21} \neq 0$, allora esiste una relazione retroattiva tra le due serie. Generalizzando al caso $k = p$, la serie storica \mathbf{y}_t può essere rappresentata mediante un modello VAR(p) se essa soddisfa

$$\mathbf{y}_t = \boldsymbol{\phi}_0 + \mathbf{y}_{t-1}\boldsymbol{\Phi}_{t-1} + \dots + \mathbf{y}_{t-p}\boldsymbol{\Phi}_{t-p} + \mathbf{a}_t \quad p > 0, \quad (\text{A.14})$$

quest'ultima può essere riscritta nella sua forma compatta

$$\mathbf{y}_t = \boldsymbol{\phi}_0 + \sum_{i=1}^p \mathbf{y}_{t-i}\boldsymbol{\Phi}_{t-i} + \mathbf{a}_t \quad p > 0, \quad (\text{A.15})$$

dove ϕ_0 e \mathbf{a}_t sono definiti come nel caso del VAR(1), e Φ_j è una matrice di dimensioni $k \times k$.

A.2.2 Modelli a correzione dell'errore

Consideriamo il processo multivariato \mathbf{y}_t le cui componenti y_{it} sono processi $I(1)$. Se la combinazione lineare $\beta' \mathbf{y}_t \sim I(0)$ allora il processo multivariato \mathbf{y}_t è un sistema cointegrato. Il vettore β che ha questa proprietà prende il nome di vettore di cointegrazione. Dato il processo VAR(p)

$$\mathbf{y}_t = \phi_0 + \sum_{i=1}^p \mathbf{y}_{t-i} \Phi_{t-i} + \mathbf{a}_t \quad (\text{A.16})$$

dove \mathbf{a}_t è una sequenza di vettori casuali Gaussiani *i.i.d.* con media zero e matrice di covarianza $\text{Var}(\mathbf{a}_t) = \Sigma$ definita positiva. Assumiamo che $|\phi(z)| = |\mathbf{I} - \sum_{i=1}^p \phi_i z^i| \neq 0$ per $|z| < 1$ ma $|\phi(1)| = 0$. Perciò \mathbf{y}_t ha qualche radice unitaria. Sottraendo \mathbf{y}_{t-1} da entrambi i lati e con un po di algebra otteniamo la sua rappresentazione a correzione dell'errore (ECM)

$$\Delta \mathbf{y}_t = \Pi \mathbf{y}_{t-1} + \sum_{i=1}^{p-1} \phi_i^* \Delta \mathbf{y}_{t-i} + \mathbf{a}_t \quad (\text{A.17})$$

dove $\Pi = -\phi(1)$ è una matrice ($k \times k$), $\phi_i^* = -\sum_{j=i+1}^p \phi_j$ per $i = (1, \dots, p-1)$, e $\Delta \mathbf{y}_t = \mathbf{y}_t - \mathbf{y}_{t-1}$. In questo le matrici coefficienti dell'Eq. A.15 possono essere ottenuti dall'Eq. A.17 in questo modo

$$\begin{aligned} \Phi_i &= \phi_i^* - \phi_{i-1}^*, \quad i = 2, \dots, p \\ \Phi_1 &= \Pi + \mathbf{I} + \phi_1^*. \end{aligned}$$

Consideriamo m il rango di Π . Analizziamo 3 casi interessanti.

1. Rango(Π)=0: questo implica che $\Pi = 0$. Perciò non esistono vettori di cointegrazione. In questo caso \mathbf{y}_t ha k radici unitarie, la soluzione è lavorare direttamente sulle serie nelle differenze, $\Delta \mathbf{y}_t$, ossia con un processo VAR($p-1$).
2. Rango(Π)= k , la matrice Π ha rango pieno quindi non è invertibile, vuol dire che non ci sono radici unitarie, \mathbf{y}_t non è $I(0)$.
3. Rango(Π)= $m > 0$: In questo caso \mathbf{y}_t ha m vettori di cointegrazione e $k - m$ radici unitarie. Esistono $k \times m$ matrici di rango pieno α e β tali che

$$\mathbf{\Pi} = \boldsymbol{\alpha}\boldsymbol{\beta}'.$$

Se il processo generato dalla combinazione $\mathbf{g}_t = \boldsymbol{\beta}'\mathbf{y}_t$ è un processo $I(0)$, posto $\boldsymbol{\beta}_\perp$ una matrice $k \times (k - m)$ una matrice di rango pieno tale che $\boldsymbol{\beta}'_\perp\boldsymbol{\beta} = 0$, allora $\mathbf{y}_t = \boldsymbol{\beta}'_\perp\mathbf{g}_t$ ha $k - m$ radici unitarie e possono essere considerate come $k - m$ trend comuni di \mathbf{y}_t .

Per comprendere in modo più analitico l'informazione contenuta nel modello a correzione dell'errore, posti $k=2$ e $p=1$, riscriviamo la Eq. A.17 in una forma un po più estesa

$$\begin{bmatrix} \Delta y_{1,t} \\ \Delta y_{2,t} \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} \begin{bmatrix} \beta_1 & \beta_2 \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} + \begin{bmatrix} a_{1,t} \\ a_{2,t} \end{bmatrix} \quad (\text{A.18})$$

Consideriamo a questo punto la combinazione ottenuta da $\boldsymbol{\beta}'\mathbf{y}_{t-1}$

$$\beta_1 y_{1,t-1} + \beta_2 y_{2,t-1}$$

questa rappresenta la serie storica delle deviazioni dalla relazione di cointegrazione. In questa ottica, dalla Eq. A.18 si comprende che il movimento ($\Delta\mathbf{y}_t$) di un sistema cointegrato è determinato da due fattori. Uno \mathbf{a}_t è casuale, l'altro $\boldsymbol{\alpha}(\beta_1 y_{1,t-1} + \beta_2 y_{2,t-1})$ è determinato dall'ampiezza, al periodo precedente, della deviazione dalla relazione di cointegrazione, ossia da quella che dal punto di vista interpretativo potremmo chiamare l'entità del disequilibrio al tempo $t - 1$. La matrice $\boldsymbol{\alpha}$ si chiama matrice dei pesi (viene talvolta chiamata matrice dei *loadings*), perché il suo elemento ij ci dice qual è l'effetto sulla i -esima variabile del j -esimo elemento di $\beta_1 y_{1,t-1} + \beta_2 y_{2,t-1}$.

A.3 Il metodo di Johansen

La procedura di Johansen consente di individuare il rango e la matrice di cointegrazione². Si tratta di una procedura in cui si suppone che il sistema cointegrato possa essere rappresentato come un VAR di ordine finito con errori gaussiani. Il primo passo consiste nel rappresentare il VAR in forma ECM

²Nella trattazione che segue verranno riassunte le tappe fondamentali di questa tecnica rinviando a Johansen (1995) per una spiegazione più analitica.

$$\Delta \mathbf{y}_t = \mathbf{\Pi} \mathbf{y}_{t-1} + \sum_{i=1}^{p-1} \phi_i^* \Delta \mathbf{y}_{t-i} + \mathbf{a}_t. \quad (\text{A.19})$$

La stima avviene mediante la tecnica di massima verosimiglianza. In genere lo stimatore di massima verosimiglianza dei parametri di un modello di regressione lineare con errori gaussiani è l'OLS. Tuttavia in questo caso bisogna considerare i vincoli legati alla presenza di cointegrazione: infatti, il rango della matrice $\mathbf{\Pi}$ $k \times k$ è uguale al rango di cointegrazione m , per cui vogliamo una stima di $\mathbf{\Pi}$ con rango ridotto $m < k$ (cosa che l'OLS non fa).

A.3.1 Stima del rango di cointegrazione

Il primo problema è individuare il rango di cointegrazione m . Per fare questo, la procedura di Johansen prevede due test sul rango della matrice $\mathbf{\Pi}$, tutti e due legati al fatto che in una matrice semidefinita positiva il numero di autovalori positivi è uguale al suo rango, e gli altri sono uguali a 0. I test seguono questo iter: per prima cosa, viene definita una matrice \mathbf{M} (rinviando alla letteratura per i dettagli) che per costruzione è semidefinita positiva ed ha lo stesso rango di $\mathbf{\Pi}$. Il vantaggio di lavorare con \mathbf{M} piuttosto che con $\mathbf{\Pi}$ è dato dal fatto che \mathbf{M} è semidefinita positiva. Questo infatti ci assicura che tutti i suoi autovalori siano reali non negativi. Di questa matrice è disponibile una stima consistente $\hat{\mathbf{M}}$, con la conseguenza che gli autovalori di $\hat{\mathbf{M}}$ ($\hat{\lambda}$) sono a loro volta stimatori consistenti degli n autovalori di \mathbf{M} . A questo punto, li si ordina dal più grande $\hat{\lambda}_1$ al più piccolo $\hat{\lambda}_k$ e si imposta un test di azzeramento del più piccolo. Se la nulla di questo test viene rifiutata, quindi se $\lambda_1 > 0$, tutti gli altri saranno positivi e questo indica che la matrice $\mathbf{\Pi}$ ha rango pieno. In questo caso non si va oltre poiché il sistema è stazionario. Se invece l'ipotesi nulla del test di azzeramento è accettata si passa a λ_1 , in questo caso si profilano due alternative:

1. si può effettuare un test con ipotesi nulla $\lambda_{k-1} = 0$, assumendo che $m < k$ e testando l'ipotesi che $m < k - 1$ (variante del massimo autovalore)
2. in alternativa, si può usare un test la cui nulla è $\lambda_k = \lambda_{k-1} = 0$, ossia un test congiunto di azzeramento degli ultimi due autovalori che non dà per scontato che $m < k$ (variante della traccia).

Se l'ipotesi nulla è accettata si prosegue con λ_{k-2} , al primo rifiuto si ottiene una stima di m . Se non si rifiuta mai l'ipotesi nulla allora $m = 0$.

A.3.2 Stima del vettore di cointegrazione

Dopo aver stimato il rango di cointegrazione, si può stimare β . Un problema preliminare che si pone a questo punto è però che la matrice β non è identificata. Infatti, se β è una matrice di cointegrazione, lo è anche $b = \beta L$, dove L è una qualunque matrice $m \times m$ non singolare. Di conseguenza, esiste un numero infinito di matrici ($k \times m$) che sono equivalenti dal punto di vista osservazionale. Una soluzione proposta da Phillips è assumere le prime m righe di β uguali alla matrice identità (rappresentazione triangolare). Quindi possiamo scrivere

$$\hat{\beta} = \begin{bmatrix} I \\ \hat{\beta}_2 \end{bmatrix} \quad (\text{A.20})$$

dove $\hat{\beta}_2$ è libero da vincoli. A questo punto è sufficiente calcolare gli autovettori della matrice \hat{M} .

Dal punto di vista delle proprietà asintotiche, lo stimatore di β ha varie caratteristiche inusuali, la più notevole è quella chiamata superconsistenza. Solitamente quando uno stimatore è consistente, la sua distribuzione collassa a una degenera per $T \rightarrow \infty$ ossia $\hat{\theta} \xrightarrow{p} \theta_0$. Quindi, $\hat{\theta} - \theta_0 \xrightarrow{p} 0$. Tuttavia se moltiplichiamo questa quantità per \sqrt{T} , otteniamo una cosa che non collassa e non diverge, ma si stabilizza su una distribuzione limite. Nei casi che incontriamo di solito $\sqrt{T}(\hat{\theta} \xrightarrow{p} \theta_0)$ diventa, al crescere di T , una variabile aleatoria sempre più prossima a una normale con media 0. In questo caso invece per ottenere una distribuzione limite, la differenza $\hat{\beta} - \beta_0$ deve essere moltiplicata per T anziché per \sqrt{T} . In altre parole, la velocità con cui $\hat{\beta}$ va verso β è molto più grande. Detto in un altro modo, la dispersione dello stimatore è proporzionale a T^{-1} anziché a $T^{-1/2}$. **Questo non vuol dire che in campioni finiti questo stimatore sia particolarmente preciso: è stato notato che la sua distorsione può essere piuttosto seria, e con le ampiezze campionarie che si hanno di solito, il fatto che converga più velocemente non comporta grandi benefici.** Inserita la stima di β nell'ECM si procede alla stima con gli OLS. Gli stimatori che ne escono si distribuiscono come una χ^2 .

Bibliografia

- [1] ALEXANDER, C. (2008), *Market Risk Analysis: Practical Financial Econometrics*, John Wiley & Sons;
- [2] ALEXANDER, C. e DIMITRIU A. (2005), *Indexing, cointegration and equity market regimes*, *International Journal of Finance and Economics*, 10, pp. 213-231;
- [3] BROCKWELL, P. J. e DAVIS, R. A. (1991), *Introduction to Time Series and Forecasting*, Springer-Verlag, 2a ed.;
- [4] CAMPBELL, J. Y. e PERRON, P. (1991), *Pitfalls and opportunities: What macroeconomists should know about unit roots*, in *NBER Macroeconomics Annual 1991*, pp. 141-201, MIT Press;
- [5] DIXIT, A.K. e PINDYCK R.S. (1994), *Investment under Uncertainty*, Princeton University Press, 1994, pp. 468;
- [6] DICKEY, D.A. E FULLER W.A. (1979), *Distribution of the Estimators for Autoregressive Time Series with a Unit Root*, *Journal of the American Statistical Association*, 74, pp. 427-431.
- [7] ENGLE, R. F. e GRANGER, C. W. J. (1987), *Co-integration and Error Correction: Representation, Estimation and Testing*, *Econometrica*, 55 , pp. 251-276;
- [8] FUSS, R. e DIETER K. (2007), *The tactical and strategic value of hedge fund strategies: a cointegration approach*, *Journal of Financial Markets and Portfolio Management*, 21, 2007, pp. 425-444;
- [9] GIANNINI, C. (1992), *Cointegrazione*, *Giornale degli Economisti e Annali di Economia*, Sett.-Dic. 1991, v. 50, pp. 479-498;
- [10] HERLEMONT, D. (2004), *Pairs Trading, Convergence Trading, Cointegration*, YATS Finances & Technologies;

- [11] HULL, J. C. (2008), *Risk Management e Istituzioni Finanziarie*, Pearson;
- [12] HULL, J. C. (2011), *Options Futures and Other Derivatives*, Pearson, 8a ed.;
- [13] JOHANSEN, S. (1995), *Maximum Likelihood Inference in Co-Integrated Vector Autoregressive Processes*, Oxford University Press;
- [14] KLOEDEN, P.E. e PLATEN E. (1992), *Numerical Solution of Stochastic Differential Equations*, Springer-Verlag, 1992, pp. 636;
- [15] LEYBOURNE, S. J. e MCCABE, B. P. M., (1989) *Testing for Coefficient Constancy in Random Walk Models with Particular Reference to the Initial Value Problem*, Empirical Economics, Springer, vol. 14(2), pp. 105-112.
- [16] LUTKEPOHL, H. (2005), *New Introduction to Multiple Time Series Analysis*, Springer-Verlag;
- [17] LUTKEPOHL, H. e KRATZIG, M. (2004), *Applied Time Series Econometrics*, Cambridge University Press;
- [18] MADDALA, G. e KIM, I. (1998), *Unit Roots, Cointegration and Structural Change*, Cambridge University Press;
- [19] PFAFF, B., (2010), *Analysis of Integrated and Cointegrated Time Series with R*, Springer New York, 2a ed.;
- [20] PHILLIPS, P. C. B. (1994), *Some Exact Distribution Theory for Maximum Likelihood Estimators of Cointegrating Coefficients in Error Correction Models*, Econometrica, 62 , pp. 73-93;
- [21] RAMPERTSHAMMER S. (2007), *An Ornstein-Uhlenbeck Framework for Pairs Trading*, Department of Mathematics and Statistics the University of Melbourne;
- [22] SCHWERT G. W. (1989), *Tests for Unit Roots: A Monte Carlo Investigation*, Journal of Business and Economic Statistics, 2, pp. 147-159 ;
- [23] TSAY, R. S. (2010), *Analysis of Financial Time Series*, John Wiley & Sons, 3a ed.;
- [24] VIDYAMURTHYV G. (2004), *Pairs Trading, Quantitative Methods and Analysis*, John Wiley Sons.