

Indice

Introduzione	iii
1 I modelli markoviani per la valutazione dei titoli derivati	1
1.1 Introduzione	1
1.2 Il tempo continuo	2
1.3 Il modello di Vasicek	5
1.4 Il modello di Cox, Ingersoll e Ross	7
1.5 La volatilità stocastica	9
1.5.1 Le serie storiche finanziarie e la modellistica ARCH	11
1.5.2 I modelli ARCH e le EDS	13
2 Metodi di Stima per i parametri di Equazioni Differenziali Stocastiche	15
2.1 Introduzione	15
2.2 Soluzione di una Equazione Differenziale Stocastica	15
2.3 Soluzioni numeriche delle EDS e schemi di discretizzazione	18
2.4 Lo Stimatore di Massima Verosimiglianza in tempo continuo	20
2.5 La Pseudo Massima Verosimiglianza	21
2.6 Una stima con il Metodo Generalizzato dei Momenti	25
2.7 Gli stimatori di Inferenza Indiretta	26
2.8 Una rassegna di esperimenti Monte Carlo	29
3 Le condizioni dei Momenti di Hansen e Scheinkman per i processi di Markov in tempo continuo	30
3.1 Introduzione	30
3.2 L'operatore infinitesimale di un processo di Markov omogeneo	31
3.3 L'operatore infinitesimale come generatore di condizioni di ortogonalità	33
3.4 Applicazione delle condizioni C1 e C2 alle EDS "mean-reverting"	34
3.5 Lo stimatore GMM	38
3.6 Uno studio Monte Carlo	42
3.6.1 Stima del processo di Ornstein-Uhlenbeck	42
3.6.2 Stima del processo "Radice Quadrata"	49
3.7 Spunti per ulteriori ricerche	55

4	Applicazione dei HS-GMM al mercato statunitense	57
4.1	Introduzione	57
4.2	I dati	57
4.3	Le stime dei parametri	59
4.4	La valutazione dei titoli	61
4.5	Conclusioni	64
A	L'integrale stocastico di Ito	67
B	La Soluzione approssimata della EDS lineare (2.10)	70
C	Il Metodo Generalizzato dei Momenti	72
D	La matrice di covarianza dei GMM con parametri di disturbo	75
E	Le condizioni di regolarità per i GMM con parametri di disturbo	78
	Bibliografia	80

Introduzione

Fino ad una trentina d'anni fa, il complesso dei metodi matematici impiegati dall'analisi finanziaria, noto con il nome di Calcolo Attuariale, era quasi esclusivamente rivolto alla determinazione del valore attuale probabile di un'attività finanziaria.

Tre articoli fondamentali, apparsi a nove anni di distanza l'uno dall'altro cambiarono in maniera radicale questa situazione:

- Shape nel 1964, e in seguito Lintner nel 1965, proposero il Modello di Equilibrio per Attività Finanziarie, il *Capital Asset Pricing Model*, meglio noto con l'acronimo di CAPM, che introdusse la nozione di rischio sistematico di un titolo, misurato da un certo coefficiente di regressione;
- Black e Scholes nel 1973 mostrarono che il prezzo teorico di un'opzione di acquisto su un'azione è la soluzione di una certa equazione alle derivate parziali.

In seguito a questi tre articoli, è nata la necessità di utilizzare tecniche matematiche spesso molto diverse e sofisticate: Ottimizzazione, Programmazione Dinamica, Calcolo Differenziale Stocastico, Equazioni alle derivate Parziali e Controllo Ottimo.

Da un punto di vista economico e finanziario, si possono distinguere tre grandi temi nello sviluppo dei lavori di ricerca degli ultimi 25 anni: l'*ottimizzazione del portafoglio e dei consumi*, la *determinazione del prezzo di un'attività finanziaria* e l'*arbitraggio*. In modo intuitivo, si può definire l'ipotesi di assenza di opportunità di arbitraggio come l'impossibilità di realizzare guadagni con investimenti iniziali nulli o negativi. Sotto tale ipotesi, in particolare, esiste una legge di probabilità secondo la quale il prezzo di ogni titolo è uguale al valore atteso attualizzato dei suoi rendimenti futuri. Attraverso tale risultato il secondo e il terzo tema risultano strettamente legati; infatti i principali modelli per la determinazione dei prezzi dei titoli sono stati sviluppati sotto ipotesi di assenza opportunità di arbitraggio.

L'approccio oggi più largamente impiegato per la definizione dei modelli per la determinazione dei prezzi delle attività finanziarie trae gran parte delle sue fondamenta dal lavoro di Black e Scholes, i quali hanno avuto il merito di determinare una formula esplicita per i prezzi delle opzioni di tipo europeo che non dipende dalle ipotesi sulle preferenze degli operatori e che è funzione di variabili osservabili e della loro dinamica. Il mercato descritto da questi modelli è regolato dall'ipotesi di assenza di opportunità di arbitraggio non rischioso; gli agenti hanno aspettative conformi sullo stato futuro, e la variabile che guida il mercato è, frequentemente, il tasso di interesse istantaneo. In generale, in tali modelli, si ipotizza che la dinamica dei processi, sia di prezzo che di rendimento, possa essere descritta tramite una *Equazione Differenziale Stocastica* del tipo:

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t$$

dove $\{W_t, t > 0\}$ è un moto browniano, μ e σ sono funzioni continue rispetto a t e prendono il nome di deriva e di coefficiente di diffusione. La soluzione di tale Equazione

Differenziale Stocastica è un processo di Markov a traiettorie continue noto come *Processo di Diffusione*.

L'impiego delle Equazioni Differenziali Stocastiche equivale alla teoria della passeggiata aleatoria per i prezzi dei titoli finanziari, teoria che afferma la non prevedibilità dei rendimenti a causa dell'efficienza informativa dei mercati.

Se si ipotizza che il tasso di interesse istantaneo segua una Equazione Differenziale Stocastica (EDS), utilizzando l'ipotesi di assenza di opportunità di arbitraggio e i risultati del calcolo differenziale stocastico (infatti per tali equazioni non sono applicabili le regole del calcolo differenziale tradizionale), si ricava il prezzo di attività finanziaria derivata, cioè di un titolo il cui valore dipende dai valori di variabili sottostanti più elementari. Tale prezzo è determinato dalla così detta *Equazione di Valutazione* del prezzo di equilibrio ed è la soluzione di una certa equazione alle derivate parziali, dipendente dal titolo che ne è il presupposto, dal tasso di interesse e dai parametri che caratterizzano l'Equazione Differenziale Stocastica che descrive la dinamica del tasso di interesse stesso.

A seguito dell'assunzione dell'ipotesi di assenza di opportunità di arbitraggio per la caratterizzazione della struttura dei prezzi di equilibrio resta definito il *premio di rischio*, cioè un compenso richiesto per accettare il rischio indotto dalle fluttuazioni della variabile che guida il mercato.

In tale impostazione rientrano in particolare, oltre al modello di Black e Scholes, i modelli di Vasicek (1977) e di Cox, Ingersoll e Ross (1985).

Per l'impiego dall'Equazione di valutazione è necessario conoscere il valore dei parametri delle EDS presenti nel modello considerato.

La difficoltà della stima da dati discreti dei parametri di una Equazione Differenziale Stocastica attraverso i metodi di Massima Verosimiglianza è fortemente legata alla difficoltà della determinazione della probabilità di transizione del processo di Markov che ne è soluzione. La ricerca si è quindi orientata verso lo studio di metodi alternativi che si possono raggruppare in due approcci: quelli di Pseudo Massima Verosimiglianza e quelli "Simulation Based". L'approccio di Pseudo Massima Verosimiglianza nasce dall'impossibilità di determinare le soluzioni analitiche per molte Equazioni Differenziali Stocastiche ed impiega per lo più approssimazioni di tali soluzioni o trasformazioni delle Equazioni Differenziali in equazioni alle differenze finite. Gli stimatori di Pseudo Massima Verosimiglianza (Gourieroux, Monfort e Trognon 1984) possono essere definiti come stimatori di Massima Verosimiglianza calcolati a partire da un modello sbagliato. Le stime tendono verso uno pseudo valore che non necessariamente coincide con il valore vero del parametro oggetto di studio. Se i momenti condizionali primo e secondo della funzione di Pseudo Verosimiglianza sono esattamente specificati, si ottengono stime consistenti per i parametri coinvolti in questi modelli; qualora il momento secondo non sia correttamente specificato, si otterranno stime consistenti solo per i parametri coinvolti nel momento primo. Quest'ultima situazione è abbastanza frequente per le Equazioni Differenziali Stocastiche, per cui spesso le stime dei coefficienti della deriva risultano consistenti, ma quelle dei parametri coinvolti dal coefficiente di diffusione risultano altrettanto spesso inconsistenti. Tale problema è particolarmente fastidioso, in quanto i parametri di diffusione intervengono in maniera rilevante nell'Equazione di Valutazione dei prezzi delle attività finanziarie.

I metodi “Simulation Based” sono di recente sviluppo e si possono in genere ricondurre al metodo dei Momenti. L’idea è quella di poter determinare le stime dei parametri di un modello di difficile valutazione attraverso una procedura di simulazione e stima che ricorre all’uso di modelli ausiliari. Tra questi metodi si ricordano quello dei Momenti Simulati (Duffie e Singleton 1993) e dell’Inferenza Indiretta (Gourieroux, Monfort e Renault 1993, Gallant e Tauchen 1994). Tale approccio consente di determinare stime consistenti dei parametri di interesse, ma presenta lo svantaggio di richiedere un’elevata complessità di elaborazione legata alla difficoltà di utilizzare con successo nell’ottimizzazione della funzione obiettivo algoritmi ricorsivi che richiedano il calcolo della matrice Jacobiana.

Hansen e Scheinkman (1995), sfruttando le proprietà dell’operatore infinitesimale del processo $\{X_t, t > 0\}$ soluzione di una Equazione Differenziale Stocastica, hanno ricavato due tipi di condizioni di ortogonalità. Con l’ausilio di tali condizioni e tramite la scomposizione spettrale dell’operatore infinitesimale, gli autori pervengono ad uno stimatore non parametrico per il coefficiente di diffusione di una Equazione Differenziale Stocastica.

Queste condizioni appaiono, però, di un certo interesse anche per analisi parametriche delle Equazioni Differenziali Stocastiche: infatti le condizioni sono valide per ogni funzione $\phi(X_t)$ infinitamente derivabile e limitata, e con esse è possibile costruire una procedura di stima secondo il Metodo Generalizzato dei Momenti (GMM). L’uso di tali condizioni presenta il particolare vantaggio di non richiedere la conoscenza della funzione di Verosimiglianza del processo soluzione della Equazione Differenziale Stocastica considerata.

In questo lavoro viene presentato uno stimatore GMM, derivato dalle condizioni di Hansen e Scheinkman per un’opportuna scelta della funzione $\phi(X_t)$, per i parametri di diffusione delle Equazioni Differenziali Stocastiche che caratterizzano i modelli per la valutazione dei titoli finanziari.

Il lavoro è organizzato come segue.

Nel Capitolo 1 verranno illustrati i principali modelli per la determinazione dei prezzi dei titoli delle attività finanziarie.

Nel Capitolo 2, dopo una breve rassegna della teoria delle soluzioni delle Equazioni Differenziali Stocastiche, verranno illustrati alcuni dei principali metodi di stima per i parametri di tali Equazioni.

Nel Capitolo 3 verrà presentato lo stimatore GMM derivato dalle condizioni proposte da Hansen e Scheinkman ed attraverso studi Monte Carlo verrà condotta una verifica delle sue proprietà econometriche e un confronto con gli stimatori di Pseudo Massima Verosimiglianza e di Inferenza Indiretta.

Nel Capitolo 4 infine verrà presentata la valutazione di attività finanziarie derivate per il mercato statunitense mediante l’applicazione dello stimatore GMM proposto.

Capitolo 1

I modelli markoviani per la valutazione dei titoli derivati

1.1 Introduzione

L'attuale modellazione statistica dei processi finanziari (sia che si tratti del processo dei prezzi sia dei tassi dei rendimenti) ipotizza spesso che tali processi siano *markoviani in tempo continuo*, come pure che i loro equivalenti in tempo discreto siano markoviani di ordine 1. Nei paragrafi che seguono verranno descritti le motivazioni che hanno suggerito questa interpretazione dei mercati finanziari e alcuni tra i modelli maggiormente impiegati nella recente letteratura, con particolare riferimento alle attività derivate. Un'attività finanziaria *derivata* o più semplicemente un *derivato* (come i *futures*, *options*, *swaps* ed altre) è un titolo il cui valore dipende da quelli di altre variabili più elementari.

Nella maggior parte dei casi le variabili sottostanti i derivati sono i prezzi delle altre attività finanziarie scambiate. Ad esempio un'opzione di acquisto di tipo europeo (*call option*) è un contratto in base al quale un operatore acquista il diritto di comprare ad una certa data (*data di esercizio*), ad un fissato prezzo (*prezzo d'esercizio*), un certo titolo (*titolo di riferimento*) in possesso di un altro operatore. Tale diritto ha ovviamente un costo, il prezzo dell'opzione appunto, che sarà lecito supporre in qualche relazione con la *maturità* dell'opzione stessa, cioè il periodo di tempo che intercorre dalla data del contratto alla data di esercizio, e dal prezzo del titolo di riferimento. Se alla data di esercizio il prezzo del titolo di riferimento indicato sul mercato è minore del prezzo d'esercizio, l'opzione di acquisto non viene esercitata e l'operatore perderà il prezzo pagato per l'opzione.

Dato quindi il prezzo di esercizio, la data di esercizio e i rendimenti associati al titolo di riferimento un rilevante problema pratico è la valutazione delle attività derivate, cioè la determinazione del loro prezzo.

L'approccio oggi più largamente impiegato ha le sue fondamenta dal lavoro di Black e Scholes (1973) i quali hanno avuto il merito di determinare una formula esplicita per i prezzi delle opzioni di tipo europeo, che non dipende dalle ipotesi sulle preferenze degli operatori e che è funzione di variabili osservabili e della loro dinamica. Tale approccio si fonda sulle implicazioni dell'ipotesi di assenza di opportunità di arbitraggio non rischioso nel mercato considerato. Per assenza di opportunità di arbitraggio si intende

l'impossibilità di lucrare guadagni positivi con investimenti iniziali nulli o negativi. Nel modello di Black e Scholes questo si traduce nell'ipotesi che un titolo derivato (l'opzione in particolare), o meglio il suo rendimento, deve poter essere duplicato da un portafoglio di titoli esistenti, quindi attraverso opportune strategie di compravendita. Il valore di tale portafoglio darà il prezzo cercato.

Uno dei fattori che più influenzano i prezzi dei derivati è la *volatilità*. In modo intuitivo si può definire la volatilità di un titolo come la misura dell'incertezza che si ha sui movimenti futuri del prezzo dell'attività considerata. In altre parole i prezzi dei titoli presentano una variabilità, che sotto ipotesi di efficienza dei mercati finanziari è provocata dall'arrivo casuale di nuove informazioni sul rendimento futuro dei titoli di riferimento.

A partire dal lavoro di Black e Scholes sono stati proposti molti modelli per l'interpretazione della dinamica dei prezzi delle attività derivate.

Nel prossimo paragrafo verranno introdotti i principali concetti utilizzati nei modelli per la valutazione dei titoli derivati che saranno presi in considerazione nel corso di questo lavoro.

1.2 Il tempo continuo

Un'attività finanziaria che non paga dividendi è caratterizzata dal suo prezzo di mercato S_t al tempo t . Alla data t , il prezzo di mercato futuro $S_{t+\delta}$ (dove $\delta > 0$ indica lo scarto, misurato in anni tra il periodo corrente t e la data $t + \delta$) è una variabile aleatoria. Dato l'insieme delle informazioni disponibili al tempo t , si può considerare la distribuzione dei rendimenti $S_{t+\delta}/S_t$ nel periodo $[t, t+\delta]$ condizionale alle informazioni disponibili. La letteratura decompone il rendimento in due parti, una sistematica espressa dal valore atteso dei rendimenti, ed una che modella le variazioni casuali del prezzo in risposta ad eventi che condizionano il mercato, come l'arrivo di notizie non attese. La descrizione della dinamica dei rendimenti necessita quindi di uno strumento che modelli l'interazione tra la componente sistematica del fenomeno e le fluttuazioni casuali che lo perturbano. Pertanto, indicato con \mathbf{E}_t il valore atteso condizionale a tali informazioni si può scrivere:

$$\mathbf{E}_t(S_{t+\delta}/S_t) = S_t^{-1} \mathbf{E}_t S_{t+\delta}$$

Per δ tendente a zero si assume che il tasso dei rendimenti istantaneo composto tenda ad un valore finito $\mu(t)$ e quindi si può scrivere:

$$\left. \frac{\partial}{\partial \tau} \mathbf{E}_t(S_\tau) \right|_{\tau=t} = \mu(t) S_t.$$

Allo stesso modo viene definita la varianza condizionale:

$$Var_t(S_{t+\delta}/S_t) = S_t^{-2} Var_t(S_{t+\delta})$$

e si assume che $\delta^{-1} Var_t(S_{t+\delta}/S_t)$ tenda ad un valore finito $\sigma^2(t)$ per δ che va a zero. Da cui:

$$\left. \frac{\partial}{\partial \tau} Var_t(S_\tau) \right|_{\tau=t} = \sigma^2(t) S_t^2$$

Questi argomenti suggeriscono che la dinamica del prezzo S_t di un titolo può essere rappresentata mediante le *Equazioni Differenziali Stocastiche* (EDS nel seguito) o *processi di Ito* del tipo:

$$dS_t = \mu(t)S_t dt + \sigma(t)S_t dW_t \quad (1.1)$$

dove il processo $\{W_t, t > 0\}$ è un moto Browniano, ed in particolare un processo di Markov. Il termine $\mu(t)S_t$ prende il nome di *deriva* del processo di diffusione e il termine $\sigma(t)S_t$ prende il nome di *diffusione*.

Le EDS possono essere considerate come l'estensione ad un contesto aleatorio delle equazioni differenziali ordinarie che permettono di analizzare l'evoluzione di una struttura (come un mercato finanziario, appunto) il cui sistema di regole deterministiche (es. le clausole contrattuali) è perturbato da eventi stocastici di un certo tipo. Per tali Equazioni non sono applicabili le regole del calcolo differenziale tradizionale. In particolare la formula di derivazione viene sostituita dalla *formula di Ito* che è alla base del Calcolo Differenziale Stocastico. Le EDS, le loro soluzioni analitiche e numeriche, i metodi e le problematiche della stima dei loro parametri e gli aspetti inferenziali verranno dettagliatamente illustrate nei Capitoli seguenti.

Con tale formulazione resta definita la cosiddetta volatilità istantanea data da:

$$\sigma_t = \left[\lim_{\delta \rightarrow 0} \delta^{-1} \text{Var}_t(S_{t+\delta}/S_t) \right]^{1/2}$$

L'equazione (1.1) con $\mu(t) = \mu$ e $\sigma^2(t) = \sigma^2$ costanti per ogni t , è alla base del lavoro di Black e Sholes.

In modo equivalente, utilizzando le formule di attualizzazione, si può anche dire che il prezzo S_t dopo t periodi è dato circa da $S_t = S_0 e^{\mu t}$ dove μ indica il valore atteso del tasso di rendimento e S_0 il prezzo del titolo al tempo $t = 0$. Tale valore può essere letto come soluzione dell'equazione differenziale $dS = \mu S dt$ con valore iniziale S_0 . Ma in pratica il prezzo ha una dinamica influenzata da una certa variabilità; si può allora fare la ragionevole ipotesi che la volatilità sia costante in un breve periodo di tempo δ . Definendo quindi σ^2 come la varianza delle fluttuazioni del prezzo del titolo, $\sigma^2 S_t^2 \delta$ è la varianza del prezzo S_t durante il periodo δ . Questi argomenti suggeriscono che la dinamica di S_t si possa rappresentare con un'equazione del tipo (1.1).

L'uso delle EDS per la modellazione delle dinamiche dei titoli equivale alla teoria della passeggiata aleatoria per i prezzi dei titoli stessi, teoria che afferma la non prevedibilità dei rendimenti a causa dell'efficienza informativa dei mercati finanziari (Fama 1970). Quindi si assume che i rendimenti su periodi consecutivi equispaziati $[t + \delta, t + \delta + 1]$, $\delta = 1, 2, \dots$ siano indipendenti e identicamente distribuiti. In questa ottica è ragionevole assumere che il valore atteso e la varianza del tasso dei rendimenti $\log(S_{t+\delta}/S_t)$ sia proporzionale alla maturità del titolo.

In letteratura l'espressione del tipo (1.1) è generalmente utilizzata per la rappresentazione della dinamica dei prezzi dei titoli, in questo caso il valore $\mu(t)$ assume il significato di tasso istantaneo degli investimenti non rischiosi. Un'altra specificazione di frequente uso per il termine di deriva è $\mu(t)S_t = k(\theta - S_t)$, cioè una funzione lineare in S_t . In tal caso se k è una costante positiva, tutti gli scarti da S_t rispetto a θ diminuiscono in valore assoluto con orizzonte temporale δ a velocità esponenziale. Per

questo il processo del tipo:

$$dS_t = k(\theta - S_t)dt + \sigma S_t dW_t, \quad (1.2)$$

dove si è assunto $\sigma(t) = \sigma$ costante per ogni t , prende il nome di processo con forza di richiamo verso il valore di equilibrio θ o *mean reverting*. Questa specificazione viene preferibilmente impiegata per la rappresentazione della dinamica dei tassi di interesse. Le considerazioni fatte a proposito della volatilità inducono a considerare dei processi di diffusione con elasticità della varianza $\gamma \geq 0$ costante:

$$dS_t = k(\theta - S_t)dt + \sigma S_t^\gamma dW_t \quad (1.3)$$

Dire che questa elasticità è positiva implica assumere che la varianza istantanea di S_t tende a zero quando S_t tende a zero; tale proprietà assicura in particolare che il processo $\{S_t, t > 0\}$ che rappresenta un prezzo o un tasso di interesse assuma sempre valori non negativi. In particolare con $\gamma > 1/2$ o con $\gamma = 1/2$ e $\sigma^2 < 2k\theta$ il processo $\{S_t, t > 0\}$ non assume mai valore zero.

Di particolare interesse è il processo “radice quadrata”:

$$dS_t = k(\theta - S_t)dt + \sigma\sqrt{S_t}dW_t$$

che è stato utilizzato per modellare il tasso di interesse istantaneo da Cox, Ingersoll e Ross (1985).

Per quanto riguarda l'informazione al tempo t , la modellistica finanziaria tende a riassumere l'informazione disponibile alla data t nei valori di un certo numero di variabili dette *variabili di stato* che caratterizzano l'economia descritta da tali modelli. L'ipotesi che viene generalmente assunta è che le dinamiche di tali variabili seguano a loro volta una EDS.

La volatilità svolge un ruolo centrale per la determinazione del prezzo delle attività derivate. Modelli come quello di Black e Scholes sono tra i più impiegati per lo studio dei derivati nonostante le loro assunzioni siano in genere violate, ed in particolare è l'ipotesi che σ^2 sia una costante è spesso non verificata.

Nei paragrafi successivi, nell'ambito della modellistica basata sulle EDS, saranno descritti il modello di Vasicek (1977) e quello di Cox, Ingersoll e Ross (1985), che si fondano sull'ipotesi di volatilità costante, e sarà illustrato come il loro successo sia dovuto alla possibilità di determinare delle formule di prezzo dipendenti da sole variabili osservabili (il tasso di interesse e la sua dinamica in particolare). Gli autori inoltre soffermano la loro attenzione sulla struttura per scadenza dei tassi di interesse¹.

A tali modelli sarà fatto costante riferimento nella restante parte di questo lavoro.

Verranno inoltre brevemente descritti anche i principali modelli che sono stati sviluppati sotto l'ipotesi di volatilità non costante nel tempo e che prendono il nome di *modelli a volatilità stocastica*.

¹Con *struttura per scadenza* o struttura a termine si intende la relazione esistente tra i tassi di rendimento di titoli aventi le medesime caratteristiche quanto a dividendi, cedole pagate (in genere si suppone tale valore nullo), rischiosità, e clausole contrattuali, ma con scadenze diverse. La struttura per scadenza contiene tutti i fattori di attualizzazione dei rendimenti futuri ed è quindi indispensabile per analizzare tutte le operazioni economiche proiettate nel futuro.

1.3 Il modello di Vasicek

Il contesto generale in cui si colloca tale modello è quello della determinazione del prezzo di un titolo derivato sotto ipotesi di assenza di opportunità arbitraggio. Tali prezzi alla data t sono realizzazioni di processi stocastici e dipendono quindi dal verificarsi di un evento aleatorio. Tale evento è il realizzarsi di un processo che descrive la variabile di stato considerata la cui dinamica segue una EDS. Le regole contrattuali del mercato dipendono dalla sola data corrente t e dalla variabile di stato. Formulata in questo contesto, l'ipotesi di assenza di opportunità di arbitraggio implica che tutti i portafogli non rischiosi che si possono costruire in questo mercato abbiano un rendimento istantaneo uguale al tasso di interesse delle attività non rischiose, cioè dei titoli a rendimento fisso.

In questo modello l'unica variabile di stato è il tasso di interesse istantaneo r_t la cui dinamica è descritta dalla EDS:

$$dr_t = k(\theta - r_t)dt + \sigma dW_t \quad (1.4)$$

Il tasso di interesse istantaneo è non rischioso perché è il tasso di rendimento di un titolo che vale una lira tra un istante (cioè alla scadenza del titolo) e quindi per l'assenza di rischio di insolvenza dell'emittente non è soggetto a incertezza nel prezzo nel prossimo istante. Ciò non impedisce che r_t vari nel tempo in modo aleatorio, cioè che sia un processo stocastico.

Si consideri ora il problema della determinazione del prezzo di un titolo a rendimento fisso senza cedole, che paga una lira alla data di scadenza T (ad esempio un'obbligazione), e si indichi con $P(r, t, T)$ tale prezzo all'istante t , quando il tasso assume il valore r . Seguendo Vasicek (1977), in tale contesto si può dimostrare che esiste una funzione λ dipendente da t e da r tale per cui il prezzo $P(r, t, T)$ è soluzione dell'equazione:

$$\frac{1}{2}\sigma^2 P_{rr} + k(\theta - r)P_r + P_t - \lambda r P_r - rP = 0$$

con condizione al contorno $P(r, t, T) = 1$ dove T indica la data di scadenza del titolo, P_r la derivata prima del prezzo rispetto ad r , P_{rr} la derivata seconda, P_t la derivata prima di P rispetto a t .

La funzione λ assume il significato di *premio di rischio*² ed è legata all'incertezza

²Nella letteratura dell'analisi dei mercati finanziari il *premio di rischio* costituisce una delle determinanti del rendimento atteso da un'attività finanziaria. Gli investitori avversi al rischio, a parità di rendimenti, si orientano verso investimenti che presentano un rischio più basso. Di conseguenza l'investitore sarà disposto a sopportare un rischio elevato per un investimento soltanto se tale rischio è accompagnato da un compenso, un premio, in termini di maggior rendimento atteso. Nel premio di rischio si distingue una componente detta *non sistematica* e una detta *sistematica o di mercato*. La prima componente è legata al rischio di insolvenza di chi emette i titoli in cui si investe. La seconda componente dipende dalle fluttuazione dei valori dei titoli ed è legata al ciclo degli affari e quindi alle condizioni del mercato. Mentre la prima componente del rischio può essere eliminata attraverso opportune strategie di diversificazione del portafoglio, la seconda è intrinseca nel mercato. Per ulteriori approfondimenti si rimanda a Blake (1990, Cap.2)

sulla variabile di stato r ; in particolare indica che il valore atteso dei rendimenti che eccedono il tasso delle attività non rischiose è proporzionale alla sua sensibilità al rischio legato ad r . In questo contesto esiste una legge di probabilità detta *corretta per il rischio*, tale che:

$$P(r, t, T) = \mathbf{E}_t \left[\exp \left(- \int_t^T r(s) ds \right) \right]$$

dove \mathbf{E}_t indica il valore atteso condizionato all'insieme di informazioni disponibili al tempo t , ed esprime, sotto ipotesi di assenza di arbitraggio, una formula di attualizzazione estesa al tempo continuo. Tale espressione è nota in letteratura come Equazione di Valutazione. Il prezzo del titolo risulta uguale al valore atteso della somma attualizzata dei pagamenti promessi dal titolo (cedole, dividendi o il rimborso finale). Supponendo inoltre che il premio di rischio λ sia una costante, (assunzione per la verità difficile da giustificare) si può scrivere che:

$$P(r, t, T) = \exp \left[-(T-t)R_\infty + a_t(R_\infty - r) + \frac{\sigma^2}{4k} a_t^2 \right] \quad (1.5)$$

dove

$$a_t = \frac{1 - e^{-k(T-t)}}{k} \quad \text{e} \quad R_\infty = \theta - \frac{\lambda}{k} - \frac{\sigma^2}{2k^2}$$

Questa formula dipende solo da $(T-t)$, cioè dalla maturità del titolo non rischioso e questo è dovuto al fatto che sotto la probabilità corretta per il rischio il processo soluzione $\{r_t\}$ è stazionario. Per un titolo non rischioso come quello considerato fino a questo momento, che paga una lira alla scadenza, si ha che

$$P(r, t, T) = \exp(-(T-t)R(r, t, T))$$

dove $R(r, t, T)$ è il rendimento riferito alla maturità. Leggendo questa formula in funzione della maturità si ottiene la struttura per termine alla data t quando il tasso istantaneo è r .

Quindi:

$$R(r, t, T) = R_\infty - \frac{a_t}{T-t}(R_\infty - r) - \frac{\sigma^2 a_t^2}{4k(T-t)}$$

indica la struttura per scadenza cercata.

Il termine R_∞ indica il tasso di lungo periodo sul mercato obbligazionario, infatti posto $T-t = \tau$ si ha che

$$\lim_{\tau \rightarrow \infty} R(r, t, T) = R_\infty,$$

da cui il simbolismo adottato, e $R_\infty - r$, detto *spread*, indica lo scarto tra il tasso istantaneo e quello di lungo periodo.

Jamshidian (1989) ha derivato sotto le ipotesi del modello di Vasicek la formula per la determinazione del prezzo di un'opzione di vendita su un titolo a rendimento fisso senza cedole. Sia, quindi, $C(r, t, T; s, K)$ il prezzo dell'opzione, K il prezzo di esercizio, T la data di esercizio dell'opzione, s la maturità del titolo a rendimento fisso, r il tasso

di interesse e $P(r, t, s)$ il prezzo del titolo a rendimento fisso derivato dalle formule precedentemente descritte. Il prezzo deve rispettare la condizione finale:

$$C(r, t, T; s, K) = \max[P(r, T, s) - K, 0]$$

sotto vincolo $t \leq T \leq s$ altrimenti l'opzione non verrà mai esercitata.

La formula del prezzo a cui si perviene è:

$$C(r, t, T; s, K) = P(r, t, s)N(h) - KP(r, t, T)N(h - \psi) \quad (1.6)$$

dove

$$\begin{aligned} h &= \ln[P(r, t, s)/P(r, t, T)K]/\psi + \psi/2 \\ \psi &= \sigma(r_s)(1 - e^{-k(s-T)})/k \\ \sigma^2(r_s) &= \sigma^2(1 - e^{-2k(s-t)})/2k \end{aligned}$$

e $N(\cdot)$ indica il valore della funzione di ripartizione di una variabile aleatoria normale standard.

1.4 Il modello di Cox, Ingersoll e Ross

Nel modello di Black e Scholes i prezzi delle attività derivate, ricavati sotto l'ipotesi di non arbitraggio, sono dipendenti da quelli di altre attività; in particolare il prezzo del titolo su cui si basa l'opzione e il tasso di interesse del mercato risultano esogeni rispetto al prezzo dell'opzione. Il modo più rigoroso per rendere endogeni i prezzi di tutte le attività finanziarie è quello di costruire un modello di equilibrio generale, intertemporale e stocastico per un'economia di produzione e scambio, con operatori massimizzanti, avversi al rischio e dotati di aspettative razionali che interagiscono in mercati competitivi dove le opportunità di investimento variano in modo aleatorio. Il modello di Cox, Ingersoll e Ross (1985) è un importante esempio di questo approccio. Tale modello assume che un'equazione del tipo (1.1) deve essere soddisfatta da tutti i prezzi delle attività finanziarie e può essere utilizzata per trovare i loro valori di equilibrio.

Gli autori in particolare analizzano il problema della struttura per scadenza dei tassi di interesse.

Il modello di Cox, Ingersoll e Ross (CIR nel seguito) assume che la sola variabile di stato che influenza l'economia sia le variazioni nelle opportunità di produzione e che la sua dinamica sia modellata da una EDS del tipo "Radice Quadrata". Gli autori assumono anche che la dinamica del tasso di interesse di equilibrio possa essere espressa da:

$$dr_t = k(\theta - r_t)dt + \sigma\sqrt{r_t}dW_t. \quad (1.7)$$

Si consideri il problema della valutazione di un titolo a rendimento fisso senza cedole, che paga una lira alla scadenza. Gli autori determinano che il prezzo di questo titolo dipende completamente dalla struttura per scadenza ed indicano in quale modo vada

determinato il premio di rischio λ . L'equazione fondamentale per il prezzo di un titolo a rendimento fisso $P(r, t, T)$ è:

$$\frac{1}{2}\sigma^2 r P_{rr} + k(\theta - r)P_r + P_t - \lambda r P_r - rP = 0$$

con condizione al contorno $P(r, T, T) = 1$, dove T indica la data di scadenza del titolo, P_r la derivata prima del prezzo rispetto ad r , P_{rr} la derivata seconda, P_t la derivata prima di P rispetto a t .

Da questo viene dedotto che il tasso di rendimento atteso del titolo è $r + (\lambda r P_r / P)$. Si nota in particolare che l'equazione del prezzo dipende solo da una variabile aleatoria che è il tasso di interesse istantaneo. Pertanto il prezzo del titolo considerato è dato da:

$$P(r, t, T) = F(t, T)e^{-G(t, T)r} \quad (1.8)$$

dove

$$F(t, T) = \left[\frac{2\phi e^{[(k+\lambda+\phi)(T-t)]/2}}{(\phi+k+\lambda)(e^{\phi(T-t)}-1)+2\phi} \right]^{2k\theta/\sigma^2}$$

$$G(t, T) = \frac{2(e^{\phi(T-t)}-1)}{(\phi+k+\lambda)(e^{\phi(T-t)}-1)+2\phi}$$

$$\phi = ((k + \lambda)^2 + 2\sigma^2)^{1/2}$$

Il prezzo così determinato è funzione decrescente convessa del tasso di interesse e crescente del tempo (della maturità). I parametri del processo del tasso di interesse hanno i seguenti effetti: il prezzo è funzione decrescente del livello medio del tasso di interesse θ e crescente concava (decrecente convessa) rispetto a k se il tasso di interesse è più grande (piccolo) di θ . Inoltre il prezzo è una funzione crescente concava del parametro di rischio λ ; questo perché alti valori di λ indicano una alta covarianza tra il tasso di interesse e la ricchezza investita. Il prezzo del titolo è funzione crescente e concava di σ^2 . Alti valori di σ^2 indicano maggiore incertezza sulle opportunità future di ottenere rendimenti positivi.

La dinamica del prezzo del titolo non rischioso considerato è data dalla EDS:

$$dP = r(1 - \lambda G(t, T))Pdt - G(t, T)P\sigma\sqrt{r}dW$$

che nel contesto della sola variabile di stato considerata esprime la correlazione negativa tra i rendimenti sul titolo e le variazioni del tasso di interesse.

I titoli sono in genere quotati sui mercati in termini di rendimenti piuttosto che di prezzi. Allora per un titolo non rischioso come quello considerato fino a questo momento, che paga una lira alla scadenza, si ha che $P(r, t, T) = \exp(-(T-t)R(r, t, T))$ dove $R(r, t, T)$ è il rendimento alla maturità, quindi:

$$R(r, t, T) = (rG(t, T) - \log F(t, T))/(T-t)$$

che rappresenta quindi la struttura per scadenza.

All'avvicinarsi della scadenza, $R(r, t, T)$ tende ad r indipendentemente dal valore dei parametri e per T che tende all'infinito si ha il valore limite $R(r, t, \infty) = \frac{2k\theta}{\phi+k+\lambda}$.

Gli autori presentano anche una formula per la determinazione del prezzo di un titolo i cui rendimenti dipendono dal tasso di interesse, come un'opzione sul titolo a rendimento fisso, e quindi per la previsione della dinamica di tale prezzo. Si consideri quindi un'opzione di vendita sul titolo a rendimento fisso, e siano $C(r, t, T; s, K)$ il prezzo dell'opzione, K il prezzo di esercizio, T la data di esercizio, s la maturità del titolo a rendimento fisso, e r il tasso di interesse. Il prezzo deve rispettare la condizione finale:

$$C(r, t, T; s, K) = \max[P(r, T, s) - K, 0]$$

sotto i vincoli: $t \leq T \leq s$ e $K \leq F(T, s)$ altrimenti l'opzione non verrebbe mai esercitata.

La formula per la valutazione del prezzo a cui si perviene è:

$$C(r, t, T; s, K) = P(r, t, s) \chi^2 \left(2r^*[\phi + \psi + G(T, s)]; \frac{4k\theta}{\sigma^2}; \frac{2\phi^2 r e^{\gamma(T-t)}}{\phi + \psi + G(T, s)} \right) - K P(r, t, T) \chi^2 \left(2r^*[\phi + \psi]; \frac{4k\theta}{\sigma^2}; \frac{2\phi^2 r e^{\gamma(T-t)}}{\phi + \psi} \right), \quad (1.9)$$

dove

$$\begin{aligned} \gamma &= ((k + \lambda)^2 + 2\sigma^2)^{1/2} \\ \phi &= \frac{2\gamma}{\sigma^2(e^{\gamma(T-t)} - 1)} \\ \psi &= (k + \lambda + \gamma)/\sigma^2 \\ r^* &= \left[\log \left(\frac{F(T, s)}{K} \right) \right] / G(T, s) \end{aligned}$$

e $\chi^2(., ., ., .)$ è la legge di probabilità χ^2 non centrata e r^* è il valore critico del tasso di interesse che verifica $K = P(r^*, T, s)$.

Il prezzo dell'opzione risulta funzione crescente della maturità quando si considera fissata la data di scadenza del titolo sottostante, ed è funzione crescente del tasso di interesse.

1.5 La volatilità stocastica

Come abbiamo accennato nel paragrafo 1.2, l'ipotesi di volatilità costante è spesso irrealistica; di conseguenza, se la varianza condizionale dei prezzi dei titoli varia nel tempo, le ipotesi alla base della formula di Black e Scholes e degli altri modelli illustrati cadono. Per l'estensione al caso della volatilità non costante, sono state quindi proposte in letteratura specificazioni alternative di modelli che in questo paragrafo saranno presentati brevemente, rimandando il lettore interessato al lavoro di Ghysels, Harvey e Renault (1995).

Un modo spesso impiegato per ottenere qualche stima della volatilità è quello dell'inversione della formula del prezzo. Come brevemente descritto le formule per la valutazione dei prezzi dei derivati sono ricavate sotto ipotesi di assenza di opportunità

di arbitraggio. In particolare se questa condizione è verificata il prezzo delle opzioni non dipende dalle preferenze degli operatori ed esiste la probabilità corretta per il rischio con la quale è conveniente ricavare la formula del prezzo dei titoli derivati. Si può assumere che tale legge di probabilità appartenga ad una famiglia parametrica $\{g_\omega, \omega \in \Omega\}$; allora, indicato con $C_t = S_t f[\sigma_S, x_t, \omega_0]$ il prezzo dei titoli considerati, ad esempio di un'opzione, dove ω_0 è il valore vero dei parametri, e supposto che la funzione f sia biunivoca per ogni coppia (x_t, ω_0) , invertendo la formula del prezzo si ha la *volatilità implicita istantanea*:

$$\sigma_{imp}(\omega) = f^{-1}[S_t, C_t, x_t, \omega].$$

Tale definizione della volatilità implicita risulta utile nella misura in cui si riescono ad ottenere stime sufficientemente accurate di ω_0 . Nel caso delle opzioni si ricorre spesso alla volatilità implicita che si ottiene invertendo la formula per le opzioni di Black e Scholes, assegnando a C_t il valore osservato del prezzo dell'opzione.

Alcune regolarità empiriche sui derivati e sulla volatilità implicita forniscono indicazioni utili per la specificazione di modelli a volatilità non costante. Tra queste sono di particolare interesse:

- a) la forte leptocurtosi e asimmetria della distribuzione dei rendimenti dei titoli;
- b) il *leverage effects* che indica la tendenza per le variazioni dei prezzi dei titoli ad essere negativamente correlate con le variazioni della volatilità nei titoli;
- c) la persistenza della volatilità; specialmente i dati ad alta frequenza presentano evidenze di radici unitarie;
- d) gli effetti di calendario che evidenziano come informazioni che giungono agli operatori durante la chiusura dei mercati si riflettono sui prezzi dei titoli alla riapertura dei mercati. Nei giorni che seguono i fine settimana e le feste la volatilità è più alta che negli altri giorni, ma non tanto quanto ci si aspetterebbe se il tasso di arrivo delle informazioni fosse costante;
- e) l'effetto *smile*. Se il mercato seguisse le condizioni di Black e Scholes, allora le volatilità implicite corrispondenti alle opzioni che si riferiscono allo stesso titolo dovrebbero coincidere con il parametro σ della EDS che descrive la dinamica del titolo di riferimento, cosa che non è in generale verificata. La volatilità implicita di Black e Scholes dipende dal tempo t , dalla maturità $T-t$, dal rapporto $x_t = \log S_t / KP(t, T-t)$, dove $P(t, T-t)$ indica il prezzo del titolo non rischioso di riferimento. Se la volatilità implicita viene impiegata per il calcolo del prezzo dei titoli derivati, questa produce una distorsione nel prezzo detta *effetto smile* dal comportamento a U presentato dalla volatilità implicita al variare del prezzo di esercizio K . Il raggio di curvatura aumenta quando diminuisce la maturità, e inoltre la volatilità implicita non si comporta in maniera simmetrica a seconda che C_t sia minore o maggiore di K . Appare quindi naturale sviluppare modelli nei quali la volatilità possa essere stocastica.

Il modello di Hull e White (1987) affronta questo problema. In particolare viene assunto che la volatilità stessa sia una variabile di stato, e che il prezzo S_t del titolo e la volatilità σ_t siano dei processi di diffusione con innovazioni correlate. Il modello può

essere scritto come:

$$\left\{ \begin{array}{l} dS_t = \mu_t S_t dt + \sigma_t S_t d\mathcal{W}^S_t \\ \sigma_t^2 = f(Y_t) \\ dY_t = \gamma_t dt + \delta_t d\mathcal{W}^\sigma_t \\ Cov(\mathcal{W}^S, \mathcal{W}^\sigma) = \rho_t dt \end{array} \right. \quad (1.10)$$

dove $\mathcal{W} = (\mathcal{W}^S, \mathcal{W}^\sigma)$ è un moto browniano bivariato, $f : \mathbb{R} \rightarrow \mathbb{R}^+$ è una funzione nota, e ρ è il coefficiente di correlazione istantaneo tra il processo di prezzo S_t e il processo di volatilità σ_t . Stime empiriche hanno mostrato che in generale l'ipotesi di correlazione non nulla non può essere rifiutata. In questo modello restano definiti due premi di rischio relativi all'incertezza legata al processo del prezzo e a quello di volatilità.

Il contesto generale è ancora quello di un mercato dove vale l'ipotesi di assenza di opportunità di arbitraggio; esiste quindi la legge di probabilità neutrale al rischio associata al premio di rischio della volatilità sotto la quale viene ricavata la formula del prezzo. Nel caso in cui $\rho_t = 0$ e con premio di rischio per la volatilità non dipendente dal prezzo S , gli autori hanno determinato una formula per la valutazione del prezzo di un'opzione di tipo europeo che generalizza quella di Black e Scholes.

Un aspetto importante di tale modello è che volatilità è una variabile latente, si pone quindi il problema della relazione tra σ_t non osservabile e i parametri ignoti del modello. Inoltre si presentano diversi problemi inferenziali e computazionali originati dalla specificazione in tempo continuo e dalla presenza della variabile latente σ_t : ad esempio non è applicabile il metodo della Massima Verosimiglianza per la stima dei parametri del modello. Una discussione di questi aspetti si trova in Pastorello, Renault e Touzi (1994).

Un approccio assai diffuso in letteratura per affrontare il problema della volatilità non costante è quello di considerare i modelli a varianza condizionale eteroschedastica, ponendo quindi l'accento sullo studio delle serie storiche finanziarie attraverso la modellistica ARCH.

1.5.1 Le serie storiche finanziarie e la modellistica ARCH

Gli ARCH (*Autoregressive Conditional Heteroskedasticity*, Engle 1982), i GARCH (*Generalized Autoregressive Conditional Heteroskedasticity*, Bollerslev 1986), gli ARCH esponenziali (EGARCH, Nelson 1991), gli ARV (*Autoregressive Random Variance*, Taylor 1986), sono spesso impiegati per gli studi sulle serie storiche finanziarie. Un'ampia rassegna di tali studi è presente in Bollerslev, Engle e Nelson (1994).

Il processo ARCH propone una distinzione tra momento secondo condizionato e non condizionato analoga a quella dei modelli ARMA per la media. La formulazione originaria di un modello di regressione con errori ARCH(p) nell'ipotesi di distribuzione gaussiana è del tipo:

$$y_t = x_t' \beta + \epsilon_t$$

$$(\epsilon_t | I_t) \sim N(0, h_t) \quad ,$$

$$h_t = \alpha_0 + \sum_{j=1}^p \alpha_j \epsilon_{t-j}^2$$

dove x_t è un vettore di variabili esogene o endogene ritardate, β è il vettore dei coefficienti di regressione e I_t è l'informazione disponibile. Le condizioni di non negatività sui parametri per la varianza sono $\alpha_i > 0, i = 0, 1 \dots p$. Il processo lineare ARCH è stazionario in covarianza se e solo se tutte le radici dell'equazione caratteristica ad esso associata risultano esterne al cerchio unitario; inoltre presenta disturbi incorrelati ma non indipendenti poiché è presente autocorrelazione tra i quadrati delle innovazioni, e il processo non condizionato è leptocurtico anche se la distribuzione è gaussiana. Tali proprietà hanno reso popolare questa classe di modelli per l'analisi delle serie storiche finanziarie caratterizzate da non indipendenza, leptocurtosi e volatilità non costante nel tempo. Nell'uso di tali modelli è opportuna l'imposizione di restrizioni parametriche sulla funzione di varianza per garantire l'accettabilità economica delle stime ottenute.

Una naturale estensione del processo ARCH lineare è ottenibile specificando una funzione della varianza condizionata che segue un processo ARMA. Il modello GARCH offre una struttura più flessibile dell'equazione di varianza condizionata che evita una formulazione complessa della struttura dei ritardi e conserva le caratteristiche del processo ARCH utili nelle applicazioni. La formulazione di un modello di regressione con errori GARCH (p,q) nell'ipotesi di distribuzione gaussiana è del tipo:

$$y_t = x_t' \gamma + \epsilon_t$$

$$(\epsilon_t | I_t) \sim N(0, h_t) \quad ,$$

$$h_t = \alpha_0 + \sum_{j=1}^p \alpha_j \epsilon_{t-j}^2 + \sum_{j=1}^q \beta_j h_{t-j}$$

dove x_t è un vettore di variabili esogene o endogene ritardate e γ è il vettore dei parametri incogniti. Le condizioni di non negatività sui parametri per la varianza sono $\alpha_i \geq 0, i = 0, 1 \dots p, \beta_i \geq 0, i = 1, 2 \dots q$. Un processo GARCH risulta stazionario in senso debole con $var = \alpha_0 (1 - \sum \alpha_j - \sum \beta_j)^{-1}$ e $cov(\epsilon_t, \epsilon_s) = 0$, per tutti i $t \neq s$ e se e solo se $\sum_{j=1}^p \alpha_j + \sum_{j=1}^q \beta_j < 1$. Il modello GARCH permette una specificazione più parsimoniosa, nel senso che un GARCH di ordine basso, esempio un GARCH(1,1), ha proprietà simili a quelle ad un ARCH di ordine più elevato. Inoltre individua una maggiore persistenza della volatilità.

Day e Lewis (1992) hanno verificato che un modello GARCH(1,1) stimato per la varianza condizionata dell'indice dei rendimenti dei titoli di riferimento di alcune opzioni considerate nella loro analisi fornisce informazioni statisticamente significative in aggiunta alla volatilità implicita ricavata dall'inversione della formula di Black e Scholes. Attraverso questo tipo di approccio si possono ben cogliere alcune delle regolarità empiriche sopra ricordate come la curtosi delle serie osservate, e la persistenza della volatilità. Bisogna tener presente comunque che mentre tali modelli colgono bene la

dipendenza temporale dei momenti secondi condizionati, essi devono essere interpretati con attenzione nelle loro implicazioni economiche.

Un modello a volatilità stocastica ARV può essere scritto come:

$$y_t = \sigma_t \epsilon_t \quad t = 1, 2, \dots, T \quad (1.11)$$

dove $y_t = \log(S_t/S_{t-1}) - \mu$ indica il processo dei rendimenti, $\log \sigma_t^2$ segue un processo AR(1) e gli ϵ_t vengono assunti indipendenti ed identicamente distribuiti di varianza ignota. Questo modello, introdotto da Taylor (1986), presenta alcune difficoltà in fase di stima della varianza dei rendimenti, che però possono essere superate tramite l'utilizzo di una modellistica ARCH, ma coglie molto bene il comportamento non simmetrico che spesso si riscontra nei prezzi dei titoli. La varianza della (1.11) può assumere quindi una formulazione GARCH(1,1) del tipo:

$$\sigma_t^2 = \gamma + \alpha y_t^2 + \beta \sigma_{t-1}^2 \quad t = 1, \dots, T$$

La motivazione dell'uso di tale modellistica discende dalle sue prestazioni in fase di previsione. Le diverse applicazioni presenti in letteratura evidenziano che la curtosi delle serie storiche studiate è colta in maniera più accurata da modelli GARCH con errori non gaussiani, ma non cattura l'asimmetria delle distribuzioni dei rendimenti nella stessa misura del modello ARV. Per risolvere questo problema Nelson (1991) propone di utilizzare i modelli EGARCH dove $\log \sigma_t^2$ è supposta essere una funzione delle osservazioni passate in modulo e al quadrato.

1.5.2 I modelli ARCH e le EDS

Un importante risultato, dovuto a Nelson (1990), lega la modellistica ARCH alle EDS: gli ARCH infatti costituiscono un'approssimazione dei processi di diffusione.

Si consideri dunque un sistema economico, ad esempio i mercati finanziari precedentemente descritti, caratterizzato da EDS del tipo (1.1), dove le variabili di stato e il termine di diffusione non siano osservabili. È possibile formulare un processo generatore dei dati di tipo ARCH che sia simile al vero processo, nel senso che la distribuzione delle traiettorie generate dall'ARCH e quelle del processo di diffusione tendono ad "avvicinarsi" all'aumentare del grado di approssimazione.

In maniera più rigorosa: si consideri il modello (1.10) di Hull e White scritto nella forma:

$$\begin{cases} d \log(S_t) = \left(\mu_t - \frac{\sigma_t^2}{2} \right) dt + \sigma_t d\mathcal{W}^S_t \\ d \log(\sigma_t^2) = \beta(\gamma - \log(\sigma_t^2))dt + \delta d\mathcal{W}^\sigma_t \\ Cov(\mathcal{W}^S, \mathcal{W}^\sigma) = \rho dt \end{cases} \quad (1.12)$$

Approssimando tale modello alle differenze finite attraverso il metodo di Eulero, che verrà descritto nel paragrafo 2.3, e posto $s_t = \log(S_t)$, Melino e Turnbull (1990) ottengono:

$$\begin{cases} s_{t+h} = s_t + \left(\mu - \frac{\sigma_t^2}{2}\right) h + \sqrt{h}\sigma_t\epsilon_{1,t+h} \\ \log(\sigma_{t+h}^2) = \log(\sigma_t^2) + h\beta[\gamma - \log(\sigma_t^2)] + \delta\sqrt{h}\epsilon_{2,t+h} \end{cases}$$

dove $t = h, 2h, 3h\dots$ e $(\epsilon_{1,t}, \epsilon_{2,t})$ è una normale bivariata con vettore delle medie nullo e matrice di covarianze:

$$Var(\epsilon_1, \epsilon_2) = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}.$$

Nelson fornisce un insieme di condizioni sufficienti per cui il sistema appena descritto converga al (1.12) per h che tende a zero. Bisogna notare però che tale approssimazione non è un ARCH; in particolare per $\rho \neq 1$ σ_t^2 non è semplicemente una funzione dei dati osservati in tempo discreto. Per ottenere un'approssimazione di tipo ARCH, Nelson suggerisce di sostituire l'approssimazione del processo per il termine di diffusione con:

$$\log(\sigma_{t+h}^2) = \log(\sigma_t^2) + h\beta[\gamma - \log(\sigma_t^2)] + \sqrt{h}g(\epsilon_{1,t+h}),$$

dove g è una funzione misurabile con $\mathbf{E}[|g(\epsilon_{1,t+h})|^{2+\phi}] < \infty$, per qualche $\phi > 0$ e

$$Var(\epsilon_{1,t}, g(\epsilon_{1,t})) = \begin{bmatrix} 1 & \rho\delta \\ \rho\delta & \delta^2 \end{bmatrix}.$$

Una proposta per la funzione g è:

$$g(\epsilon_{1,t}) = \rho\delta\epsilon_{1,t} + \delta \left(\frac{1-\rho^2}{1-2/\pi}\right)^{1/2} \left[|\epsilon_{1,t}| - \left(\frac{2}{\pi}\right)^{1/2}\right].$$

Attraverso questa scelta di g si perviene in particolare ad un modello EGARCH. L'approssimazione con la modellistica ARCH permette di ottenere una formulazione del modello di più facile trattazione computazionale e la possibilità di impiegare le stime di Massima Verosimiglianza.

Capitolo 2

Metodi di Stima per i parametri di Equazioni Differenziali Stocastiche

2.1 Introduzione

La stima da dati discreti dei modelli di Markov in tempo continuo è oggetto di una vasta letteratura. Fin dagli inizi si è visto che i metodi di stima basati sui metodi di Massima Verosimiglianza di EDS non lineari, come quella del modello CIR, possono essere difficili da utilizzare a causa del costo computazionale implicato dalla necessità di risolvere l'equazione di Fokker e Plank per la determinazione della funzione di Verosimiglianza (Lo 1988).

La ricerca si è quindi orientata verso lo studio di metodi alternativi che si possono raggruppare in due approcci: quelli di Pseudo Massima Verosimiglianza (PML nel seguito) e quelli "Simulation Based". L'approccio di PML nasce dall'impossibilità di determinare la soluzione analitica per molte EDS ed impiega per lo più approssimazioni di tali soluzioni o trasformazioni delle EDS in equazioni alle differenze finite. I metodi "Simulation Based" sono di recente sviluppo e si possono in genere ricondurre al metodo dei Momenti. L'idea è quella di poter determinare le stime dei parametri di un modello di difficile valutazione attraverso la simulazione di modelli approssimati. Tra questi metodi ricordiamo quello dei Momenti Simulati e l'Inferenza Indiretta.

In questo capitolo, dopo aver descritto la teoria delle soluzioni analitiche delle EDS, saranno illustrati alcuni dei principali metodi di stima di PML e il metodo dell'Inferenza Indiretta. Verranno anche presentati i principali metodi di discretizzazione come mezzo per ottenere modelli approssimati di facile valutazione. Sarà fatto particolare riferimento alle EDS utilizzate nei modelli finanziari come quello di Vasicek e il modello CIR illustrati nel capitolo precedente. Mostriamo infine alcuni risultati di simulazioni Monte Carlo per i metodi illustrati, presentate da Bianchi et al. (1995).

2.2 Soluzione di una Equazione Differenziale Stocastica

In questo paragrafo e nel successivo saranno brevemente illustrate le soluzioni analitiche e numeriche di una EDS seguendo i testi di Arnold (1975) e Kloeden e Platen (1982) ai quali rimandiamo il lettore per una trattazione più completa.

Si consideri uno spazio di probabilità $(\Omega, \mathcal{F}, \mathcal{P})$ su cui è definito un processo di Moto Browniano $\{\mathcal{W}_t, t \geq 0\}$. Si indichi con (\mathcal{F}_t) la famiglia delle più piccole σ -algebre che rende \mathcal{W}_t misurabile, tali che $(\mathcal{F}_t) \in \mathcal{F}$.

Definizione. Un processo stocastico $\{X_t, t \in [t_0, T]\}$, definito su $(\Omega, \mathcal{F}, \mathcal{P})$, è detto *soluzione* nell'intervallo $[t_0, T]$ della EDS:

$$dX_t = \mu(X_t, \alpha)dt + \sigma(X_t, \alpha)d\mathcal{W}_t \quad (2.1)$$

con valore iniziale $X_{t_0} = c$, costante o variabile aleatoria di assegnata legge di probabilità, se gode delle seguenti proprietà:

1. X_{t_0} è \mathcal{F}_{t_0} misurabile.
2. Le funzioni $\mu(X_t, \alpha)$ e $\sigma(X_t, \alpha)$, dette *coefficienti* della EDS, sono \mathcal{F}_t misurabili $\forall t \in [t_0, T]$ e la funzione $\mu(X_t, \alpha)$ ammette integrale finito secondo Lebesgue in $[t_0, T]$.

3.

$$X_t = \int_{t_0}^T \mu(X_s, \alpha) ds + \oint_{t_0}^T \sigma(X_s, \alpha) d\mathcal{W}_t$$

dove \oint indica l'*integrale stocastico*¹ secondo Ito.

□

L'esistenza e l'unicità della soluzione $\{X_t\}$ è garantita dal seguente

Teorema 2.2.1 *La EDS (2.1) ammette $\{X_t\}$ come soluzione che soddisfa la condizione iniziale $X_0 = c$ e tale soluzione è unica se*

1.

$$\forall t \in [t_0, T], \quad |\mu(x) - \mu(y)| + |\sigma(x) - \sigma(y)| \leq K |x - y|$$

2.

$$\forall t \in [t_0, T], \quad |\mu(x)|^2 + |\sigma(x)|^2 \leq K^2(1 + |x|^2).$$

□

¹Per una breve introduzione all'integrale stocastico secondo Ito si veda l'Appendice A

Quasi tutte le traiettorie del processo $\{X_t\}$ sono funzioni continue in $[t_0, T]$ ed $\{X_t\}$ ammette momenti finiti di ordine p se e solo se $\mathbf{E}|X_{t_0}|^p < \infty$.

Si consideri la seguente EDS lineare generale

$$dX_t = (A(t)X_t + a(t))dt + (B(t)X_t + b(t))dW_t \quad (2.2)$$

dove i coefficienti A, a, B, b sono misurabili e limitati in $[t_0, T]$.

Tale EDS viene detta *autonoma* se i coefficienti A, a, B, b non dipendono da t . L'equazione (2.2) ha per soluzione:

$$X_t = \Phi_t \left(c + \int_{t_0}^t \Phi_s^{-1} (a(s) - B(s)b(s)) ds + \int_{t_0}^t \Phi_t^{-1} b(s) dW_s \right) \quad (2.3)$$

dove

$$\Phi_t = \exp \left(\int_{t_0}^t \left(\frac{A(s) - 1}{2B(s)^2} \right) ds + \int_{t_0}^t B(s) dW_s \right)$$

è la soluzione dell'equazione omogenea

$$d\Phi_t = A(t)\Phi_t dt + B(t)\Phi_t dW_t \quad \text{con } \Phi_{t_0} = 1$$

Nel caso $B(t) = 0$ la soluzione (2.3) è un processo gaussiano se e solo se il valore iniziale è una costante o una variabile aleatoria gaussiana.

si può enunciare inoltre il seguente teorema:

Teorema 2.2.2 *Se la EDS (2.1) ammette il processo $\{X_t\}$ come soluzione, allora $\{X_t\}$ è un processo di Markov nell'intervallo $[t_0, T]$ il cui valore iniziale in t_0 ha la legge di probabilità di c .* \square

La probabilità di transizione è indicata da $P(X_t \in B \mid X_s = s)$ per ogni B boreliano su \mathbb{R} . Quindi la EDS (2.2) ha come soluzione, condizionatamente a $X_s = x$:

$$\begin{aligned} X_t(s, x) = & \Phi(t, s) \left(x + \int_s^t \Phi(u, r)^{-1} (a(r) - B(r)b(r)) dr \right. \\ & \left. + \int_s^t \Phi(u, r)^{-1} b(r) dW_s \right) \end{aligned} \quad (2.4)$$

dove

$$\Phi(t, s) = \exp \left(\int_s^t \left(\frac{A(r) - 1}{2B(r)^2} \right) dr + \int_s^t B(r) dW_s \right)$$

Per alcune EDS lineari la soluzione (2.4) può essere scritta esplicitamente. Questo è il caso dell'equazione di Ornstein-Uhlenbeck che è alla base del modello di Vasicek e dell'equazione "radice quadrata" del modello CIR.

Infatti nel caso dell'equazione di Ornstein-Uhlenbeck:

$$dX_t = k(\theta - X_t)dt + \sigma dW_t$$

si ha che la legge di probabilità di $\{X_t\}$ condizionatamente al valore X_s , con $s < t$, è una legge normale del tipo:

$$X_t | X_s \sim \mathcal{N}\left(\theta(1 - e^{-k(t-s)}) + X_s e^{-k(t-s)}, \frac{\sigma^2}{2k}(1 - e^{-2k(t-s)})\right) \quad (2.5)$$

e legge di probabilità marginale $\{X_t\}$ tende ad una normale di media θ e varianza $\frac{\sigma^2}{2k}$ per t che tende all'infinito (da cui il nome di “media di lungo periodo” per il parametro θ).

Nel caso della EDS del modello CIR

$$dX_t = k(\theta - X_t)dt + \sigma\sqrt{X_t}dW_t$$

la legge di probabilità di $\{X_t\}$ condizionatamente al valore X_s ($s < t$) è una legge χ^2 non centrata $\chi^2[2cx_s, 2q + 2, 2u]$ di parametri:

$$\begin{aligned} c &= \frac{2k}{\sigma^2(1 - \exp(-k(t-s)))} \\ u &= cx_s e^{(-k(t-s))} \\ q &= \frac{2k\theta}{\sigma^2} - 1 \end{aligned} \quad (2.6)$$

Il valore atteso condizionale è:

$$\mathbf{E}[X_t | X_s] = \theta(1 - e^{(-k(t-s))}) + X_s e^{-k(t-s)}$$

la varianza condizionale è:

$$\mathbf{Var}[X_t | X_s] = \frac{\sigma^2}{k} X_s (e^{-k(t-s)} - e^{-2k(t-s)}) + \theta \frac{\sigma^2}{2k} (1 - e^{-k(t-s)})^2$$

e la legge di probabilità marginale $\{X_t\}$ tende ad una legge Gamma di media θ e varianza $\frac{\sigma^2\theta}{2k}$.

2.3 Soluzioni numeriche delle EDS e schemi di discretizzazione

Una delle principali difficoltà che si incontrano nella risoluzione delle EDS è legata al calcolo dell'integrale di Ito che compare nella (2.3) e alla possibilità che non si riesca ad esprimere la soluzione in forma chiusa. Tali difficoltà hanno condotto allo sviluppo di metodi per il calcolo numerico di approssimazioni alle soluzioni per assegnati valori iniziali.

Uno dei metodi più ampiamente utilizzati è quello dell'*approssimazione in tempo discreto*, nel quale l'equazione differenziale in tempo continuo viene sostituita da un'equazione alle differenze finite. Si generano valori Y_1, \dots, Y_n ... che approssimano $X_{t_1}(t_0, x_0), \dots, X_{t_n}(t_0, x_0)$... per un'assegnata partizione $t_0 \leq t_1 \leq \dots \leq t_n = T$ dell'intervallo $[t_0, T]$. Queste approssimazioni risultano essere abbastanza accurate se gli

incrementi del tempo $\Delta_n = t_{n+1} - t_n$ con $n=0,1,2,\dots$ sono sufficientemente piccoli.

La più semplice approssimazione discreta è lo *schema di Eulero* Per la EDS (2.1) lo schema ha la forma

$$Y_{n+1} = \mu(Y_n, \alpha)Y_n\Delta_n + \sigma(Y_n, \alpha)\Delta\mathcal{W}_n$$

per $n=0,1,\dots, N-1$ con valore iniziale $X_0 = c$ e dove $\Delta_n = \tau_{n+1} - \tau_n$ con $t_0 = \tau_0 < \tau_1 < \dots < \tau_N = t$ è una partizione dell'intervallo $[t_0, T]$ e $\Delta\mathcal{W}_n = \mathcal{W}_{\tau_{n+1}} - \mathcal{W}_{\tau_n}$.

La variabile casuale $\Delta\mathcal{W}_n$ è una normale di media zero e varianza Δ_n . Nel caso si conosca la soluzione di una EDS si può valutare un errore di approssimazione usando il criterio dell'errore assoluto.

Definizione. Sia $\{X_t\}$ la soluzione di una EDS e sia Y_{t_τ} la sua approssimazione discreta di passo τ . Tale approssimazione è convergente in senso forte a $\{X_t\}$ al tempo t se

$$\lim_{\tau \rightarrow 0} \mathbf{E}(|X_t - Y_{t_\tau}|) = 0$$

inoltre l'approssimazione discreta Y_t è detta convergente in senso forte di ordine $\gamma > 0$ al tempo t se esiste una costante positiva C , non dipendente da τ e un valore τ_0 tale che

$$\epsilon(\tau) = \mathbf{E}(|X_t - Y_{t_\tau}|) \leq C\tau^\gamma \quad \forall \tau \in (0, \tau_0)$$

□

In particolare l'approssimazione di Eulero risulta convergente in senso forte di ordine $\gamma = 0.5$.

Accanto all'approssimazione di Eulero si può considerare l'approssimazione di *Milstein*. In tale schema si sostituisce lo sviluppo di Taylor di primo ordine della funzione $\sigma(X_t, \alpha)$ nella (2.1), e si applica l'approssimazione di Eulero precedentemente illustrata. Si perviene ad un'espressione del tipo:

$$Y_{n+1} = Y_n + [\mu(Y_n) - \frac{1}{2}\sigma(Y_n)\sigma'(Y_n)]\tau + \sigma(Y_n)\Delta\mathcal{W}_n + \frac{1}{2}\sigma(Y_n)\sigma'(Y_n)[\Delta\mathcal{W}_n]^2$$

dove $\sigma'(Y_n)$ indica la derivata prima della della funzione $\sigma(Y_n)$ rispetto a Y . Tale schema di approssimazione risulta convergente in senso forte di ordine $\gamma = 1$.

Il processo di discretizzazione deve inoltre garantire la stabilità numerica delle soluzioni.

Definizione. L'approssimazione discreta Y_{t_τ} è numericamente stabile in senso probabilistico per una data EDS se per ogni intervallo finito $[t_0, T]$, esiste una costante positiva Δ_0 tale che per ogni $\epsilon > 0$ e $\forall \delta \in (0, \Delta_0)$ si ha che:

$$\lim_{|Y_0^\delta - \tilde{Y}_0^\delta| \rightarrow 0} \sup_{t_0 \leq t \leq T} P(|Y_{n_\tau}^\delta - \tilde{Y}_{n_\tau}^\delta| > \epsilon) = 0$$

dove \tilde{Y}_0^δ indica l'approssimazione discreta ottenuta a partire dal valore iniziale Y_0^δ . □

Si può verificare che il metodo di Eulero è numericamente stabile.

2.4 Lo Stimatore di Massima Verosimiglianza in tempo continuo

I metodi di Massima Verosimiglianza sono certamente tra i più impiegati dagli statistici per i problemi di inferenza. Nel caso delle EDS tuttavia l'esplicitazione della funzione di Verosimiglianza è spesso tutt'altro che semplice. Per poter risolvere questo problema, gli studi si sono orientati alla ricerca di una formula della Verosimiglianza che non richieda l'uso degli integrali stocastici. Sono quindi stati proposti per i parametri delle EDS gli stimatori di Massima Verosimiglianza in Tempo Continuo (CTML), che saranno illustrati brevemente in questo paragrafo, rimandando il lettore interessato al testo di Lipster e Shirayev (1974).

Si consideri un processo di Markov omogeneo $\{X_t, t \in [0, T]\}$ soluzione dell'equazione differenziale stocastica (EDS)

$$dX_t = \mu(X_t, \alpha)dt + \sigma(X_t, \alpha)dW_t. \quad (2.7)$$

dove W_t è un moto Browniano e α rappresenta un vettore di parametri.

Seguendo Liptser e Shirayev(1974, cap.7) le stime di Massima Verosimiglianza in Tempo Continuo (CTML) dei parametri α sono ottenute eguagliando a zero il vettore

$$\begin{aligned} \mathcal{S}_t(x, \alpha) = \frac{\partial}{\partial \alpha} \ln \mathcal{L}_t(x, \alpha) &= \int_0^T \left(\frac{\partial}{\partial \alpha} \mu(x_s, \alpha) \right) \sigma^2(x_s, \alpha) dx_s \\ &\quad - \int_0^T \left(\frac{\partial}{\partial \alpha} \mu(x_s, \alpha) \right) \sigma^2(x_s, \alpha) \mu(x_s, \alpha) ds \end{aligned}$$

dove $\mathcal{L}_t(x, \alpha)$ indica il rapporto di verosimiglianza. Dal Teoremi del Limite della teoria delle martingale segue che le stime di Massima verosimiglianza $\hat{\alpha}$ sono consistenti e asintoticamente normali. La covarianza quadratica di \mathcal{S}_t :

$$\mathcal{I}_T(x, \alpha) = \int_0^T \left(\frac{\partial}{\partial \alpha} \mu(x_s, \alpha) \right) \sigma^2(x_s, \alpha) \mu(x_s, \alpha) ds,$$

sotto condizioni di regolarità, può essere considerata come la matrice di informazione di Fisher.

Nel caso particolare del modello CIR

$$dX_t = k(\theta - X_t)dt + \sigma\sqrt{X_t}dW_t$$

per il parametri k , θ e σ si ha

$$\begin{aligned} \hat{k} &= \frac{\int_0^T \frac{ds}{x_s} \int_0^T dx_s - T \int_0^T \frac{dx_s}{x_s}}{T^2 - \int_0^T x_s ds \int_0^T \frac{ds}{x_s}} \\ \hat{\theta} &= \frac{T \int_0^T dx_s - \int_0^T x_s ds \int_0^T \frac{dx_s}{x_s}}{\int_0^T \frac{ds}{x_s} \int_0^T dx_s - T \int_0^T \frac{dx_s}{x_s}} \end{aligned}$$

$$\sigma = \frac{\int_0^T [dx_s]^2}{\int_0^T x_s dt}$$

e matrice di Informazione pari a:

$$\mathcal{I}_t(k, \theta) = \begin{bmatrix} \int_0^T \frac{(\theta - r_s)^2}{\sigma^2(x_s)} ds & \int_0^T \frac{k(\theta - r_s)^2}{\sigma^2(x_s)} ds \\ \int_0^T \frac{k(\theta - r_s)^2}{\sigma^2(x_s)} ds & \int_0^T \frac{k^2}{\sigma^2(x_s)} ds \end{bmatrix}$$

Nella pratica, non è tuttavia possibile registrare un insieme di osservazioni continue, ma solo un insieme di osservazioni discrete ottenibili da una partizione $0 = t_0 < t_1 < \dots < t_n = T$ di $[0, T]$, dove $\delta_n = t_{n+1} - t_n$ è indipendente da p e tale che $\lim_{n \rightarrow \infty} \delta_n = 0$. Si possono quindi sostituire gli integrali continui con somme discrete. Per questi stimatori sono state dimostrate proprietà di consistenza e limiti in distribuzione. Per ulteriori approfondimenti si rimanda il lettore ai testi di Le Breton (1976) e Genon-Catalot e Jacob (1993).

2.5 La Pseudo Massima Verosimiglianza

Una delle principali ipotesi che garantiscono le proprietà delle stime di Massima Verosimiglianza è che la legge stocastica che regola il manifestarsi del fenomeno in esame sia riconducibile ad una specifica famiglia parametrica di distribuzioni di probabilità; in altre parole che il modello sia correttamente specificato. In diverse circostanze non siamo in grado di affermare che questa condizione sia verificata, di regola per la mancanza di informazioni sufficienti sul fenomeno in esame. Ma può anche darsi il caso che, pur con la certezza che il modello sia stato correttamente specificato, la trattazione analitica o numerica della funzione di Verosimiglianza risulti impossibile. In tali casi bisogna limitarsi alla massimizzazione di una funzione di Verosimiglianza associata ad una famiglia di leggi di probabilità che non contiene necessariamente la vera distribuzione, ma che è compatibile con le ipotesi sui momenti. Tale metodo di stima è detto di *Pseudo Massima Verosimiglianza*² (PML nel seguito).

Gourieroux e Monfort (1993) e Gourieroux, Monfort e Trognon (1984) presentano le condizioni generali sotto le quali il metodo di PML basato sui momenti condizionali fornisce stime consistenti ed asintoticamente normali e dimostrano l'esistenza di un limite inferiore per la varianza asintotica.

Alla luce delle prospettate difficoltà per la soluzione esplicita della (2.4), il metodo di PML può essere utilmente applicato per la stima dei parametri delle EDS.

²In letteratura sono presenti sia la dizione di “Pseudo” che di “Quasi” Verosimiglianza. La “Pseudo” è presente soprattutto nei lavori sui criteri di informazione e nella scuola francese (Gourieroux, Monfort 1984 cap.8). “Quasi” è presente soprattutto nella scuola anglosassone e nei lavori di White (1994). Il termine “Quasi” è utilizzato con diverso significato nella scuola francese per gli stimatori a due stadi con parametri di disturbo, ai quali faremo riferimento nel corso di questo lavoro. Per evitare confusione ed ambiguità ci atterremo alla terminologia in uso nella scuola francese.

Si consideri dunque il modello generale (*cfr.* Cap.1):

$$dX_t = k(\theta - X_t)dt + \sigma X_t^\gamma dW_t \quad \gamma \geq 0. \quad (2.8)$$

che, come mostrato è uno dei più utilizzati per rappresentare la dinamica di un tasso di interesse. In particolare per $\gamma = 0$ si ha il modello di Vasicek (1975), per $\gamma = 1/2$ si ha il processo “radice quadrata” di Cox-Ingersoll-Ross (1985) e per $\gamma = 1$ si ha il modello di Brennan e Schwartz (1980).

Nella pratica si osserva solo una serie temporale $X_{t_1}, \dots, X_{t_i}, \dots, X_{t_T}$, con $t_i < t_{i+1}$, $i=1, 2, \dots, T$, cioè una realizzazione in tempo discreto del processo $\{X_t\}$. Dalla proprietà di Markov di $\{X_t\}$ si ricava che la distribuzione di $\{X_{t_{i+1}}\}$ condizionata a $X_{t_i}, X_{t_{i-1}}, \dots, X_{t_1}$ è pari quella di $\{X_{t_{i+1}}\}$ condizionata a X_{t_i} . Supponendo che le osservazioni siano equispaziate ed indicata con $\delta t = t_i - t_{i-1}$, l'unità temporale di osservazione, l'equazione (2.4) può essere utilizzata per determinare la distribuzione di $X_{t_i+\delta t} | X_{t_i}$.

In particolare nel caso della (2.8) per $\gamma = 0$ dalla (2.5) si ottiene una rappresentazione AR(1) stazionaria e gaussiana del processo dei valori osservati:

$$X_{t_i+\delta t} = \theta(1 - e^{-k\delta t}) + e^{-k\delta t}X_{t_i} + \epsilon_t \quad (2.9)$$

dove ϵ_t è la legge normale $N(0, \frac{\sigma^2}{2k}(1 - e^{-2k\delta t}))$, i cui parametri σ, θ, k sono tutti strettamente positivi ed identificabili.

L'approssimazione discreta della soluzione del processo rappresentata dalla precedente equazione diviene più accurata al tendere di δt a zero; in questo senso quindi la (2.5) pur essendo l'esatta distribuzione condizionata di $X_{t_i+\delta t} | X_{t_i}$ costituisce una Pseudo-Verosimiglianza.

Sempre per la (2.8) nel caso di $\gamma = 1/2$, dalla (2.4) si ha per $X_{t_i+\delta t} | X_{t_i}$ una distribuzione χ^2 non centrata con i parametri indicati nelle formule (2.6). Come nel caso precedente l'approssimazione discreta della legge χ^2 non centrata conduce ad una funzione di Pseudo Verosimiglianza nel senso sopra descritto, che risulta poco usata in letteratura per la stima dei parametri. Nel caso di $\gamma = 1$, invece, la distribuzione $X_{t_i+\delta t} | X_{t_i}$ non è conosciuta, si è quindi costretti ad utilizzarne una qualche approssimazione.

Si consideri ora la EDS più generale del tipo:

$$dX_t = k(\theta - X_t)dt + \sigma g(X_t, \gamma)dW_t. \quad (2.10)$$

dove γ indica eventuali parametri e g è una funzione nota. Tale espressione racchiude anche la (2.8) per $g(X_t, \gamma) = X_t^\gamma$. La soluzione della (2.10) assume la forma³:

$$X_{t+\delta t} = \theta(1 - e^{-k\delta t}) + e^{-k\delta t}X_t + e^{-k(t+\delta t)} \int_t^{t+\delta t} \sigma g(X_s, \gamma) e^{ks} dW_s \quad (2.11)$$

Nel caso in cui $g(X_t, \gamma)$ sia una funzione lineare in $\{X_t\}$ si può ricondurci alla (2.4).

Negli altri casi per la soluzione dell'integrale di Ito

$$\int \sigma g(X_s, \gamma) \exp(k\delta t) dW_s$$

³Per la determinazione dell'espressione seguente si veda l'Appendice B

si può considerare la seguente approssimazione (supposto δt abbastanza piccolo):

$$e^{-k(t+\delta t)} * \int_t^{t+\delta t} \sigma g(X_s, \gamma) e^{ks} dW_s \cong e^{-k\delta t} \sigma g(X_t, \gamma) [W_{t+\delta t} - W_t].$$

L'integrale ha quindi approssimativamente una legge normale del tipo:

$$N(0, e^{-2k\delta t} \sigma^2 g^2(X_t, \gamma) \delta t).$$

Con tale approssimazione si può ottenere dalla (2.11) una rappresentazione AR(1) del processo dei valori osservati a cui è associata una Pseudo Verosimiglianza di tipo normale. Si ha quindi:

$$X_{t_i+\delta t} = \theta(1 - e^{-k\delta t}) + e^{-k\delta t} X_{t_i} + \epsilon_{t_i+\delta t} \quad (2.12)$$

dove

$$\epsilon_{t_i+\delta t} \sim N(0, e^{-2k\delta t} \sigma^2 g^2(X_{t_i}, \gamma) \delta t).$$

In questa impostazione per la stima dei parametri della (2.12) si possono utilizzare vari metodi.

Uno stimatore di Pseudo Massima Verosimiglianza (PML) può essere definito in modo intuitivo come uno stimatore di Massima Verosimiglianza calcolato a partire da un modello errato; la stima quindi tenderà verso uno "pseudo valore vero" che non necessariamente coincide con il valore vero del parametro oggetto di studio. L'approccio di PML basato sui momenti condizionati è diretto alla stima della media condizionale ($\mathbf{E}[x_{t_i+\delta t} | x_{t_i}]$ nel nostro caso) e possibilmente della varianza condizionale.

Si indichi allora con $m(x_{t_i}, \alpha) = \mathbf{E}[x_{t_i+\delta t} | x_{t_i}]$ la media condizionale dove α è il vettore dei parametri di interesse e m una funzione nota; in pratica supponiamo che la media condizionata sia ben specificata.

Si indichi inoltre con $v(x_{t_i}, \gamma) = \text{Var}([x_{t_i+\delta t} | x_{t_i}])$ la varianza condizionata, assunta correttamente specificata, dove γ è il vettore dei parametri che la caratterizzano. di cui si possa ottenere una stima consistente.

Gourieroux e Monfort (1993) e Gourieroux, Monfort e Trognon (1989) mostrano che le stime di PML dei parametri α presenti nel momento primo convergono al valore vero se la Pseudo Verosimiglianza appartiene alla famiglia *esponenziale lineare*⁴. Ad esempio la variabile aleatoria Normale, con varianza nota, appartiene alla famiglia esponenziale lineare.

Qualora il vettore γ abbia elementi in comune con α gli Autori mostrano che le stime di PML dei parametri α e γ convergono ai valori veri se la Pseudo Verosimiglianza

⁴Una famiglia di densità di probabilità $f(x, \alpha)$ indicizzata dal parametro α , appartiene alla famiglia *esponenziale lineare* se può essere scritta nella forma $f(x, \alpha) = \exp[A(\alpha) + B(x) + C(\alpha)x]$ dove α è la media della distribuzione. Per una descrizione della famiglia esponenziale lineare si veda Gourieroux e Monfort (1993)

appartiene alla famiglia *esponenziale quadratica*⁵.

Da quanto detto appare chiaro che il modello (2.5) soddisfa le condizioni richieste per le stime di PML, mentre per la (2.9) siamo in presenza di un modello con eteroschedasticità di forma ignota. In questa situazione si può solo affermare che condizionatamente alla specificazione della varianza le stime di PML dei parametri della media sono consistenti.

Nel caso in cui nell'espressione della varianza condizionata, ben specificata o meno, non compaiono parametri si possono utilizzare anche i Minimi Quadrati per la stima del momento primo condizionato. Sempre con riferimento al modello (2.12) e posto $a = e^{-k\delta t}$ e $b = \theta(1 - e^{-k\delta t})$, si ha che:

$$(\hat{a}, \hat{b}) = \underset{a, b}{\operatorname{argmin}} \sum_{i=1}^{n-1} \frac{(X_{t_i+\delta t} - a - bX_{t_i})^2}{g^2(X_{t_i})}$$

Gli stimatori \hat{a} e \hat{b} così ottenuti sono consistenti e più efficienti degli stimatori OLS per δt abbastanza piccolo. Pertanto le stime di θ, k, σ possono essere ricavate da:

$$e^{-\hat{k}\delta t} = \hat{a} \quad \hat{\theta}(1 - e^{-\hat{k}\delta t}) = \hat{b}$$

$$\hat{\sigma}^2 = \frac{e^{2\hat{k}\delta t}}{(n-1)\delta t} \sum_{i=0}^{n-1} \frac{\hat{u}_{i+1}^2}{g^2(X_{t_i})}$$

dove \hat{u}_{i+1} sono i residui di regressione. Le stime presentate sono consistenti per δt che tende a zero con velocità $1/n$.

Qualora dei parametri di interesse siano presenti nell'espressione della varianza (es. nella (2.5) il parametro di regressione k oppure il parametro γ nel caso della (2.8)) il metodo della PML sarebbe preferibile per l'efficienza delle stime ottenute.

Alternativamente si può ricorrere alle stime dei Minimi Quadrati Ponderati Iterati (Davidson e Mackinnon 1993), detti anche Minimi Quadrati Quasi Generalizzati dalla scuola francese (Gourieroux e Monfort 1989 Cap. 8), che risultano asintoticamente efficienti. Il metodo di stima consiste in una procedura a due stadi: nel primo stadio si calcolano le stime OLS di a e b e quindi i residui di regressione \hat{u}_i . Tramite i Minimi Quadrati Non Lineari:

$$\hat{\gamma} = \underset{\gamma}{\operatorname{argmin}} \sum_{i=1}^{n-1} (\hat{u}_{i+1}^2 - g^2(x_t, \gamma))^2.$$

i residui \hat{u}_i permettono di ottenere stime consistenti dei parametri contenuti nell'espressione della varianza.

⁵Una famiglia di densità di probabilità $f(x, \alpha, \gamma)$ indicizzata dai parametri α, γ appartiene alla famiglia *esponenziale quadratica* se può essere scritta nella forma $f(x, \alpha, \gamma) = \exp[A(\alpha, \gamma) + B(x) + C(\alpha, \gamma)x + D(\alpha, \gamma)x^2]$ dove α e γ sono rispettivamente la media e la varianza della distribuzione. Per una descrizione della famiglia esponenziale quadratica si veda Gourieroux e Monfort (1993).

Nel secondo stadio si utilizza la stima dell'eteroschedasticità così ottenuta per il calcolo delle stime di a e b di seconda approssimazione attraverso i Minimi Quadrati Ponderati:

$$(\tilde{a}, \tilde{b}) = \underset{a, b}{\operatorname{argmin}} \sum_{i=0}^{n-1} \frac{(X_{t_i+\delta t} - a - bX_{t_i})^2}{g^2(X_{t_i}, \tilde{\gamma})}.$$

Si calcolano nuovamente i residui di regressione \tilde{u}_t e si imposta nuovamente il calcolo dei NLS per la stima dei parametri della varianza di secondo stadio $g^2(X_t, \tilde{\gamma})$. Le stime di k, θ, σ possono quindi essere ottenuti usando le espressioni inverse precedentemente illustrate.

Naturalmente anche attraverso l'uso di una delle approssimazioni discrete (es. quella di Eulero) si perviene ad una PML. Un esempio noto in letteratura è quello dell'equazione di Ornstein-Uhlenbeck. Infatti applicando a tale equazione la discretizzazione di Eulero si ha un modello AR(1) del tipo:

$$x_t - x_{t-\delta t} = k\theta\delta t - k\delta t x_{t-\delta t} + \sigma\epsilon_t$$

e quindi

$$x_t = k\theta\delta t + (1 - k\delta t)x_{t-\delta t} + \sigma\epsilon_t \quad (2.13)$$

con $\epsilon_t \sim N(0, \delta t)$. Appare evidente, confrontando la (2.9) con la (2.13), che la stima di PML di k ottenuta con questo ultimo modello, in genere con l'impiego delle stime OLS, è distorta.

2.6 Una stima con il Metodo Generalizzato dei Momenti

Illustriamo ora brevemente il metodo di stima proposto da Chan et al. (1992) fondato sul Metodo Generalizzato dei Momenti⁶ di Hansen (1982) (GMM nel seguito). Sotto condizioni di ergodicità del processo soluzione della EDS (2.8) è possibile costruire delle condizioni di ortogonalità fondate sulla discretizzazione di Eulero nella forma:

$$x_{t+1} - x_t = k\theta - kx_t + \epsilon_{t+1} \quad t = 1, 2, \dots, T \quad (2.14)$$

dove $\mathbf{E}[\epsilon_t] = 0$ e $\mathbf{E}[\epsilon_t^2] = \sigma^2 x_t^{2\gamma}$. L'approccio proposto dagli Autori consiste nel ricavare dalla (2.14) un insieme di vincoli sui momenti primo e secondo di ϵ_t che costruiscono un sistema di equazioni da usare per i GMM.

Posto $\alpha = (k, \theta, \sigma, \gamma)$ il vettore dei parametri di interesse, le restrizioni proposte sono:

$$\mathbf{E}(f_t(\alpha)) = 0 \quad \text{dove} \quad f_t(\alpha) = \begin{bmatrix} \epsilon_t \\ \epsilon_t x_{t-1} \\ \epsilon_t^2 - \sigma^2 x_{t-1}^{2\gamma} \\ (\epsilon_t^2 - \sigma^2 x_{t-1}^{2\gamma}) x_{t-1} \end{bmatrix}$$

Lo stimatore GMM è ottenuto come:

$$\hat{\alpha} = \underset{\alpha}{\operatorname{argmin}} \left(\frac{1}{T} \sum_{t=1}^T f_t(\alpha) \right)' D_t \left(\frac{1}{T} \sum_{t=1}^T f_t(\alpha) \right) \quad (2.15)$$

⁶Per una descrizione del Metodo Generalizzato dei Momenti si veda l'Appendice C

dove D_t è una matrice definita positiva. Hansen (1982) ha dimostrato che scegliendo come D_t la matrice di covarianza del vettore $f_t(\alpha)$, si ottengono gli stimatori GMM di α con la più piccola matrice di covarianza asintotica. Una stima consistente della matrice di covarianza è stata proposta da Newey e West (1987).

Sotto ipotesi nulla che i vincoli siano validi e che la matrice D_t converga per t che tende all'infinito alla matrice di covarianza di $f_t(\alpha)$

$$T \left(\frac{1}{T} \sum_{t=1}^T f_t(\alpha) \right)' D_t \left(\frac{1}{T} \sum_{t=1}^T f_t(\alpha) \right) \sim \chi^2_{(J-K)}$$

dove J è il numero delle restrizioni e K è il numero dei parametri da stimare. Tale statistica fornisce una misura della bontà di adattamento del modello.

2.7 Gli stimatori di Inferenza Indiretta

I modelli econometrici conducono spesso a formulazioni piuttosto complesse delle leggi delle variabili endogene, condizionali alle esogene, che rendono estremamente complicato il calcolo delle stime dei parametri oggetto di studio. Sono classici esempi in letteratura i modelli a scelta multipla discreta (Gourieroux e Monfort 1993b), i modelli non lineari a parametri aleatori, i modelli ARCH con variabili latenti e i modelli finanziari in tempo continuo basati sulle EDS.

Come illustrato nei paragrafi precedenti, per la stima di questi ultimi modelli, per i quali non è agevole derivare la forma esplicita della funzione di verosimiglianza, è usuale ricorrere ad una delle approssimazioni descritte, in particolare al metodo di Eulero, e quindi applicare il metodo della PML. Questo metodo fornisce stime consistenti per δt che tende a zero quando il numero delle osservazioni cresce. Tale circostanza è chiaramente poco realistica in quanto l'unità temporale di osservazione δt è generalmente considerata fissata al crescere della lunghezza della serie osservata T . Inoltre una importante causa di inconsistenza delle stime dei parametri coinvolti nella varianza è l'errata specificazione della forma di eteroschedasticità.

Nella letteratura recente sono stati proposti due metodi per ottenere stime consistenti. Il primo si basa su condizioni sui momenti soddisfatte dalle osservazioni per poter applicare il Metodo Generalizzato dei Momenti (GMM). Tale approccio è stato seguito da Hansen e Scheinkman (1995) quando le condizioni sono derivate dall'operatore infinitesimale di una EDS e sarà diffusamente illustrato nel capitolo successivo.

Sempre al Metodo dei Momenti si rifanno quelli che calibrano i parametri del modello con qualche momento marginale primo o secondo. Ad esempio Duffie e Singleton (1990, 1993), suggeriscono di utilizzare il Metodo dei Momenti simulati quando i momenti teorici hanno una forma poco trattabile. Altre applicazioni sono state quelle di McFadden (1989), e di Pakes e Pollard (1989).

Il secondo metodo consiste nella possibilità di applicare l'usuale approccio basato sull'approssimazione discreta del modello e successivamente di correggere la distorsione asintotica attraverso simulazione, e questo è il caso dei metodi di Inferenza Indiretta introdotti da Smith(1993), Gourieroux, Monfort e Renault (1993) e da Gallant e Tauchen (1994).

In modo generale, sia (M) il modello strutturale oggetto di studio, definito da:

$$y_t = f(y_{t-1}, \underline{x}_t, u_t, \alpha)$$

dove f è una funzione nota, \underline{x}_t è il vettore delle variabili esogene, u_t è un disturbo con distribuzione di probabilità nota, α è il vettore dei parametri che caratterizzano il modello. Si supponga che (M) sia caratterizzato da una funzione di Verosimiglianza $\mathcal{L}(\underline{y}_t, \alpha)$ di difficile trattazione, e si consideri un modello approssimato (ausiliario) (M_\star) , caratterizzato dai parametri β , con funzione di Verosimiglianza $\lambda_\star(\underline{y}_t, \beta)$ di più facile trattazione.

Il metodo dell'Inferenza Indiretta sfrutta la simulazione fondata sul modello strutturale (M) per correggere la distorsione asintotica di $\hat{\beta}_\star$ fondate su (M_\star) . È importante sottolineare che in questo approccio i modelli (M) e (M_\star) possono essere caratterizzati da parametri diversi, nel senso che i parametri del modello (M_\star) possono non avere lo stesso significato fisico e statistico di quelli di (M) anche se sono in eguale numero. La procedura dell'Inferenza Indiretta può essere descritta nei seguenti passi:

1. Si calcolano le stime $\hat{\beta}$ di PML del modello (M_\star) , usando i valori osservati \underline{y}_t :

$$\hat{\beta} = \operatorname{argmax}_\beta \lambda_\star(\underline{y}_t; \beta);$$

2. Si generano disturbi casuali u_t dalla distribuzione nota e si simula il modello (M) condizionatamente ad un valore assegnato del vettore α ottenendo i valori simulati $\tilde{y}_t(\alpha)$;
3. Si calcolano le stime $\tilde{\beta}(\alpha)$ di PML dei parametri del modello ausiliario (M_\star) usando i valori simulati $\tilde{y}_t(\alpha)$:

$$\tilde{\beta} = \operatorname{argmax}_\beta \lambda_\star(\tilde{y}_t(\alpha); \beta);$$

4. La stima di Inferenza Indiretta del vettore α proposta da Gouriéroux, Monfort e Renault(1993) è ottenuta risolvendo il seguente problema di minimo:

$$\hat{\alpha} = \operatorname{argmin}_\alpha [\hat{\beta} - \tilde{\beta}(\alpha)]' \Omega^{-1} [\hat{\beta} - \tilde{\beta}(\alpha)]. \quad (2.16)$$

dove Ω è una matrice simmetrica definita positiva.

Gallant e Tauchen(1994) hanno proposto una versione modificata dello stimatore (2.16), ottenuto considerando direttamente la funzione di score del modello ausiliario. Si ha quindi che:

$$\hat{\alpha} = \operatorname{argmin}_\alpha \frac{\partial \lambda_\star}{\partial \beta'}(\tilde{y}_t(\alpha); \hat{\beta}) \Sigma^{-1} \frac{\partial \lambda_\star}{\partial \beta}(\tilde{y}_t(\alpha); \hat{\beta}); \quad (2.17)$$

dove Σ è una matrice definita positiva.

Le matrici Ω e Σ ottimali sono quelle che minimizzano la varianza asintotica delle stime di Inferenza Indiretta. Per una loro espressione analitica si rimanda il lettore a

Gourieroux, Monfort e Renault (1993) e Gallant e Tauchen (1994).

Gli stimatori (2.16) e (2.17) associati con la matrice di pesi ottimale hanno la stessa accuratezza asintotica. Inoltre, se i modelli (M) e (M_*) hanno lo stesso numero di parametri e non ci sono problemi di identificazioni, le stime sono indipendenti dalle matrici dei pesi Ω e Σ ed asintoticamente equivalenti. Infatti dalla (2.16) si tratta di risolvere l'equazione in α : $\hat{\beta} = \tilde{\beta}(\alpha)$ e $\frac{\partial \lambda_*}{\partial \beta}(\tilde{y}_t(\alpha); \hat{\beta}) = 0$ dalla (2.17); dalla definizione di $\tilde{\beta}(\alpha)$ discende l'equivalenza.

La scelta del modello ausiliario è strettamente legata alla specificità del problema oggetto di studio. Per la stima dei parametri delle EDS in genere si sceglie il modello discretizzato oppure in forza al risultato di Nelson (1990) ricordato nel paragrafo 1.5.2 un modello ARCH. Bisogna inoltre sottolineare l'elevata complessità di elaborazione e alcuni problemi di calcolo di questo metodo. Infatti ogni algoritmo numerico per la soluzione di problemi di ottimizzazione non lineari, come quelli implicati dalle (2.16) e (2.17), è caratterizzato da una successione di vettori α che convergono alla soluzione del problema. L'uso di algoritmi di tipo Newton implicano il calcolo numerico della matrice Jacobiana, che in questo contesto coinvolge il calcolo di numerosi insiemi di dati simulati $\tilde{y}_t(\alpha)$. Pertanto sarebbe preferibile l'utilizzo di algoritmi che non richiedono il calcolo della matrice Jacobiana.

Un esempio di applicazione dell'Inferenza Indiretta per le EDS è presentato da Gourieroux, Monfort e Renault (1993). Gli autori rivolgono la loro attenzione all'equazione di Ornstein-Uhlenbeck

$$dX_t = k(\theta - X_t)dt + \sigma dW_t.$$

Essi presentano un confronto tramite esercizi Monte Carlo, di 200 replicazioni su serie di 250 osservazioni, il comportamento delle stime di ML basate sulla soluzione (2.5), di PML basati sulla discretizzazione di Eulero (2.13) e le stime di Inferenza Indiretta dove la (2.5) viene usata per la simulazione del modello strutturale e la (2.13) è impiegata come modello ausiliario. Da tali esercizi emerge che le stime di θ e le varianze Monte Carlo sono le stesse per tutti i metodi considerati e risultano non distorte. Le stime di k e σ sono come atteso distorte per il metodo di PML, mentre le stime di Inferenza Indiretta presentano un sostanziale miglioramento nella distorsione rispetto alle stime PML ma rispetto a queste presentano una maggiore varianza Monte Carlo. La metodologia dell'Inferenza Indiretta può essere estesa ai modelli con variabili latenti. Una recente applicazione è quella di Pastorello, Renault e Touzi (1994) per il modello di Hull e White (1987), illustrato nel paragrafo 1.5, scritto nella forma:

$$\begin{cases} dS_t = \mu(S_t, \sigma_t) + \sigma_t dW_t^s \\ d\sigma_t^2 = k(\theta - \sigma_t^2)dt + \psi dW_t^\sigma \end{cases} \quad (2.18)$$

dove i moti browniani W_t^s e W_t^σ sono supposti incorrelati. Gli Autori sono interessati alla stima dei parametri θ , k e ψ del processo di volatilità; per la simulazione del modello strutturale viene considerata l'approssimazione di Eulero del sistema (2.18), mentre viene proposto come modello ausiliario un EGARCH(1,0) secondo il suggerimento di Nelson (1990).

2.8 Una rassegna di esperimenti Monte Carlo

Bianchi et al.(1994) presentano una valutazione del comportamento di 7 distinti stimatori tra i CTML, i GMM, la PML, e l'Inferenza Indiretta, descritti nei paragrafi precedenti, attraverso esperimenti Monte Carlo. La EDS di riferimento è la (1.6) del modello CIR descritto nel paragrafo 1.4. I dati estratti dalla EDS del modello CIR sono ottenuti mediante la simulazione di traiettorie del processo soluzione seguendo gli schemi di Eulero, Milstein e Talay. Gli Autori mettono in rilievo che non sono state riscontrate differenze significative nei risultati di stima dovute al metodo di generazione dei dati impiegato.

Per la procedura di simulazione l'unità di tempo è un anno, l'intervallo di tempo di generazione è pari ad un'ora, l'intervallo di osservazione è pari a un dato osservato ogni trenta giorni, l'intervallo di simulazione va da 1 a 200 anni.

Il comportamento degli stimatori è differente per il tre parametri oggetto di studio.

Nel caso di σ^2 lo stimatore migliore tra quelli descritti risulta quello di CTML, mentre quello dei GMM per piccoli campioni presenta una distorsione fino al 10%.

Il parametro di media di lungo periodo θ è stimato in generale senza distorsione significativa. Il miglior stimatore risulta ancora il CTML. Per il parametro k tutti gli stimatori si presentano ampiamente distorti nei piccoli campioni e convergono al valore vero per ampiezze campionarie elevate. Asintoticamente l'errore standard delle stime decresce così come la distorsione di k è decrescente ed è non significativa per grandi campioni.

Capitolo 3

Le condizioni dei Momenti di Hansen e Scheinkman per i processi di Markov in tempo continuo

3.1 Introduzione

Il problema della stima da dati discreti dei parametri di un'Equazione Differenziale Stocastica attraverso i metodi di Massima Verosimiglianza è fortemente legato alla difficoltà della determinazione della probabilità di transizione del processo di Markov che ne è soluzione.

Come mostrato nel Capitolo precedente, in letteratura sono stati sviluppati approcci di stima alternativi fondati principalmente sulla discretizzazione del processo, come il metodo di Pseudo Massima Verosimiglianza e l'Inferenza Indiretta per simulazione. Hansen e Scheinkman (1995) sfruttando le proprietà dell'operatore infinitesimale del processo $\{X_t, t > 0\}$ soluzione di una EDS, hanno proposto due insiemi di condizioni di ortogonalità, valide per ogni trasformazione $\phi(X_t)$ infinitamente derivabile, con quali si può costruire una procedura di stima per i parametri di una EDS secondo il Metodo Generalizzato dei Momenti.

In questo capitolo, dopo una rassegna sull'operatore infinitesimale, saranno illustrate le condizioni di Hansen e Scheinkman e verrà derivato, attraverso un'opportuna scelta della trasformazione $\phi(X_t)$, uno stimatore GMM per i parametri delle EDS utilizzate nei modelli descritti nel Capitolo 1. Sarà mostrata anche un'espressione della matrice di covarianza asintotica per gli stimatori ottenuti. Verranno infine presentati risultati di simulazioni Monte Carlo per la verifica del comportamento asintotico e nei piccoli campioni degli stimatori proposti e si confronteranno i risultati con quelli ottenuti con i metodi della Pseudo Massima Verosimiglianza e dell'Inferenza Indiretta.

Nella restante parte di questo lavoro viene fatta l'ipotesi che i processi di Markov, soluzioni delle EDS considerate, siano ergodici e reversibili¹. La proprietà di reversibilità in particolare è auspicabile perché consente di limitare l'analisi ai soli ope-

¹Un processo $\{X_t, t \in \mathbb{R}\}$ è detto *reversibile* se $\forall \tau \in \mathbb{R}, \forall t_1, t_2, \dots, t_n \in \mathbb{R}$ le distribuzioni di $(X_{t_1}, X_{t_2}, \dots, X_{t_n})$ e $(X_{\tau-t_1}, X_{\tau-t_2}, \dots, X_{\tau-t_n})$ sono le stesse. In particolare la distribuzione di X_t è indipendente da t .

ratori infinitesimali autoaggiunti².

3.2 L'operatore infinitesimale di un processo di Markov omogeneo

In questo paragrafo sarà introdotto il concetto di operatore infinitesimale di un processo di Markov omogeneo e ne saranno illustrate le principali proprietà. A tale scopo sarà fatto riferimento ai testi di Arnold (1975) e Dynkin(1965) ai quali si rimanda il lettore per ulteriori approfondimenti.

Sia $\{X_t, t \geq 0\}$ un processo di Markov definito sullo spazio di probabilità $(\Omega, \mathcal{F}_t, \mathcal{P}_*)$, dove \mathcal{F}_t è una famiglia di σ -algebre, e a valori in $(\mathbb{R}, \mathcal{B}, \mathcal{P})$, dove \mathcal{B} è la classe di Borel e \mathcal{P} è la misura di probabilità immagine su \mathcal{B} .

Definizione. Un processo di Markov $\{X_t, t \geq 0\}$ è omogeneo rispetto al tempo se la probabilità di transizione $\mathcal{P}(s, y, t, B) = \mathcal{P}(X_t \in B \mid X_s = y)$ è stazionaria, cioè se la condizione

$$\mathcal{P}(s + u, y, t + u, B) = \mathcal{P}(s, y, t, B)$$

è identicamente soddisfatta per $0 \leq s \leq t$ e per $0 \leq s + u \leq t + u$ e $\forall B$ boreliano in \mathbb{R} . Da cui si può scrivere $\mathcal{P}(t, y, x) = \mathcal{P}(X_t \mid X_0 = y)$. \square

Si indichi con $L^2(\mathcal{P})$ lo spazio delle funzioni $\phi : \mathbb{R} \rightarrow \mathbb{R}$, misurabili secondo Borel, per cui esiste finito $\int_{\mathbb{R}} \phi^2 d\mathcal{P}$. Tale spazio dotato di norma $\langle \phi, \psi \rangle = \int_{\mathbb{R}} \phi\psi d\mathcal{P}$ e prodotto scalare $\|\phi\| = \sqrt{\int_{\mathbb{R}} \phi^2 d\mathcal{P}}$ è uno spazio di Hilbert.

Al processo $\{X_t\}$ è associata la famiglia di operatori lineari $\{\mathcal{T}_t : t \geq 0\}$:

$$\mathcal{T}_t\phi(y) = \mathbf{E}[\phi(X_t) \mid X_0 = y] = \int_{\mathbb{R}} \phi(x)\mathcal{P}(t, y, dx) \quad (3.1)$$

dove $\mathcal{P}(t, y, dx) = \mathcal{P}(X_t \mid x_0 = y)$.

Tali operatori sono limitati e godono della proprietà di semigrupp, cioè:

$$\mathcal{T}_{t+s} = \mathcal{T}_t\mathcal{T}_s = \mathcal{T}_s\mathcal{T}_t, \quad t, s, t + s \geq 0$$

Definizione. L'Operatore infinitesimale \mathcal{A} di un processo di Markov omogeneo $\{X_t, t_0 \leq t \leq T\}$, è definito dal:

$$\lim_{t \rightarrow 0} \frac{\mathcal{T}_t(\phi) - \phi}{t} = \mathcal{A}\phi$$

per $\phi : \mathbb{R} \rightarrow \mathbb{R}$ in $L^2(\mathcal{P})$. Questo limite esiste ed è finito se ϕ ha derivata prima rispetto a t continua e derivata seconda rispetto ad x continua e a crescita limitata. Il dominio D dell'operatore infinitesimale \mathcal{A} è un sottospazio proprio di $L^2(\mathcal{P})$ di tutte le funzioni ϕ per le quali il limite esiste. \square

Una proprietà dei processi di Markov con traiettorie continue è che le loro probabilità di transizione $\mathcal{P}(t, x, B)$ sono univocamente definite dall'operatore \mathcal{A} .

È possibile verificare che:

²Gli operatori autoaggiunti saranno illustrati nel paragrafo 3.3

1. per il Teorema fondamentale del Calcolo Integrale:

$$\mathcal{A} \int^t \mathcal{T}_s \phi ds = \mathcal{T}_t \phi - \phi \quad \forall \phi \in L^2(\mathcal{P})$$

2. Gli operatori \mathcal{A} e \mathcal{T} commutano, cioè si verifica che:

$$\mathcal{T}_t \phi - \phi = \int^t \mathcal{A} \mathcal{T}_s \phi ds = \int^t \mathcal{T}_s \mathcal{A} \phi ds \quad \forall \phi \in \mathcal{D}$$

Si vedrà ora come l'operatore \mathcal{A} assuma rilievo in relazione ai processi di diffusione.

Definizione. Un processo di Markov omogeneo $\{X_t, t \geq 0\}$ con traiettorie quasi ovunque continue è detto *processo di diffusione* se la probabilità di transizione $\mathcal{P}(y, t, B)$ soddisfa le seguenti condizioni $\forall s \geq 0$ e $\epsilon > 0$:

1.

$$\lim_{t \rightarrow s} \frac{1}{t-s} \int_{|x-y| \geq \epsilon} \mathcal{P}(y, t, dx) = 0$$

2. Esiste una funzione $\mu(s, y)$ tale che

$$\lim_{t \rightarrow s} \frac{1}{t-s} \int_{|x-y| \leq \epsilon} (x-y) \mathcal{P}(y, t, dx) = \mu(y)$$

3. Esiste una funzione $\sigma(s, y)$ a valori non negativi tale che

$$\lim_{t \rightarrow s} \frac{1}{t-s} \int_{|x-y| \leq \epsilon} (x-y)(x-y)' \mathcal{P}(y, t, dx) = \sigma(y)$$

□

Le funzioni μ e σ sono dette coefficienti del processo di diffusione; in particolare μ è detto *deriva* e σ è detto *diffusione*.

Si può verificare che per i processi di diffusione le probabilità di transizione, sotto assunzioni di regolarità, sono univocamente determinate dai coefficienti del processo. Ad ogni processo di diffusione con coefficienti μ e σ è associato l'*operatore differenziale* o *operatore di Dynkin*, la cui applicazione ad una generica funzione ϕ del dominio è:

$$\mathcal{D}\phi = \mu(s, x) \frac{\partial}{\partial x_t} \phi + \frac{1}{2} \sigma(s, x) \frac{\partial^2}{\partial x_t^2} \phi$$

per ogni funzione ϕ limitata per $t \geq 0$ e ivi due volte differenziabile rispetto ad x .

Dalla definizione di operatore infinitesimale \mathcal{A} e dalle condizioni 2 e 3 della precedente definizione si ricava che (Arnold 1975, pag 156):

$$\mathcal{A}\phi = \mathcal{D}\phi. \quad (3.2)$$

Nel Teorema 2.2.2 viene affermato che $\{X_t\}$ soluzione di una EDS del tipo: $dX_t = \mu(t, X_t)dt + \sigma(t, X_t)d\mathcal{W}_t$ è un processo di Markov; sotto alcune condizioni esso è anche un processo di diffusione:

Teorema 3.2.1 *La EDS: $dX_t = \mu(t, X_t)dt + \sigma(t, X_t)d\mathcal{W}_t$ soddisfa le condizioni del Teorema 2.2.1. Se le funzioni μ e σ sono continue rispetto a t , la soluzione $\{X_t\}$ è un processo di diffusione per $t \geq 0$ con deriva e parametro di diffusione pari a μ e σ . In particolare la soluzione di una EDS autonoma è un processo di diffusione omogeneo.* □

3.3 L'operatore infinitesimale come generatore di condizioni di ortogonalità

Sia $\{X_t\}$ processo di Markov omogeneo, quindi $\mathbf{E}[\phi(x_t)]$ è indipendente dal tempo e la sua derivata, se esiste, rispetto a t è nulla. Ricordando che $\mathcal{T}_t\phi(y) = \mathbf{E}[\phi(x_t) | x_0 = y]$ per la Legge del Valore Atteso Iterato si ha che

$$\int_{\mathbb{R}} \phi d\mathcal{P} = \int_{\mathbb{R}} \mathcal{T}_t\phi d\mathcal{P} \quad \forall \phi \in L^2(\mathcal{P})$$

Quindi, derivando rispetto a t , sfruttando il fatto che \mathcal{A} e \mathcal{T}_t commutano in D , e per la legge del Valore Atteso Iterato si può scrivere:

$$\begin{aligned} \frac{\partial}{\partial t} \mathbf{E}[\phi(x_t)] &= \frac{\partial}{\partial t} \mathbf{E}[\mathbf{E}[\phi(x_{t+1}) | x_t]] = \\ \mathbf{E}\left[\frac{\partial}{\partial t} \mathbf{E}[\phi(x_{t+1}) | x_t]\right] &= \mathbf{E}[\mathcal{A}\phi(x_t)] = 0 \end{aligned}$$

Da quanto sopra Hansen e Scheinkman ricavano la seguente Condizione dei Momenti:

$$\mathbf{C1} : \mathbf{E}[\mathcal{A}\phi(x_t)] = 0 \quad \forall \phi \in D$$

dove D è il dominio dell'operatore infinitesimale \mathcal{A} .

Attraverso l'uso degli operatori aggiunti e sfruttando la proprietà di commutazione tra \mathcal{A} e \mathcal{T}_t si può pervenire ad un secondo insieme di condizioni di ortogonalità. Si ricorda che si definisce *operatore aggiunto* di \mathcal{T} l'operatore \mathcal{T}^* tale che $\langle \phi^*, \mathcal{T}\phi \rangle = \langle \phi, \mathcal{T}^*\phi^* \rangle$; si indichi inoltre con \mathcal{A}^* l'operatore aggiunto di \mathcal{A} e con D^* il suo dominio in $L^2(\mathcal{P})$.

Dal fatto che \mathcal{A} e \mathcal{T}_t commutano si ha $\langle \mathcal{T}\mathcal{A}\phi, \phi^* \rangle = \langle \mathcal{A}\mathcal{T}\phi, \phi^* \rangle$ e per gli operatori aggiunti si ha che $\langle \mathcal{A}\mathcal{T}\phi, \phi^* \rangle = \langle \phi, \mathcal{A}^*\mathcal{T}^*\phi^* \rangle$. Quindi sostituendo una relazione nell'altra e applicando la Legge del Valore Atteso Iterato si ha:

$$\begin{aligned} \langle \mathcal{T}\mathcal{A}\phi, \phi^* \rangle &= \langle \phi, \mathcal{A}^*\mathcal{T}^*\phi^* \rangle \\ \mathbf{E}[\mathbf{E}[\mathcal{A}\phi | x_0]\phi^*] &= \mathbf{E}[\mathbf{E}[\mathcal{A}^*\phi^* | x_0]\phi] \end{aligned}$$

Da quanto sopra gli Autori ricavano la seguente Condizione dei Momenti

$$\mathbf{C2} : \mathbf{E}[(\mathcal{A}\phi(x_{t+1}))\phi^*(x_t)] = \mathbf{E}[(\mathcal{A}^*\phi^*(x_t))\phi(x_{t+1})] \quad \forall \phi \in D \text{ e } \phi^* \in D^*.$$

Appare evidente che **C2** include **C1** per ϕ^* funzione costante.

Hansen e Scheinkman proseguono il loro lavoro introducendo i criteri in base ai quali un operatore possa soddisfare le condizioni **C1** e **C2**. In altre parole gli autori si preoccupano di garantire le condizioni per le quali attraverso l'uso di \mathcal{A} , che soddisfa le **C1** e **C2**, si possano identificare il deriva e la diffusione di una EDS associata ad un processo di Markov reversibile. Nella pratica poi, gli Autori si pongono nelle condizioni di poter impiegare l'operatore di Dynkin che garantisce tale identificazione a meno di un fattore di scala.

Essi sono inoltre particolarmente interessati a pervenire ad uno stimatore non parametrico per la diffusione $\sigma(x_t)$ di una EDS generale del tipo (2.1) che ammetta come soluzione un processo di Markov reversibile. Tale stimatore viene ricavato attraverso la decomposizione spettrale unica, garantita dalle condizioni **C1** e **C2**, dell'operatore infinitesimale \mathcal{A} , supposto autoaggiunto.

Un caso particolare di tale stimatore è stato utilizzato da Ait-Sahalia(1994) per lo studio dei tassi “overnight” per gli Stati Uniti dal 1973 al 1993. Le condizioni **C1** e **C2** hanno quindi nel lavoro di Hansen e Scheinkman un ruolo del tutto strumentale per la determinazione dello stimatore non parametrico.

Tali condizioni tuttavia appaiono di un certo interesse per un'analisi parametrica delle EDS perché possono costituire la base di un approccio GMM per la stima dei parametri. L'uso di tali condizioni presenta il particolare vantaggio di non richiedere la conoscenza della funzione di Verosimiglianza del processo soluzione della EDS, evitando quindi il problema della soluzione della (2.4) o dell'approssimazione della Verosimiglianza.

Restano aperte due questioni: la prima è legata alla scelta della funzione ϕ da utilizzare nelle **C1** e **C2**. La seconda è legata ad un problema noto da tempo nella letteratura degli operatori differenziali: tali operatori sono identificati a meno di un parametro di scala. Tale questione, come sarà illustrato nel corso del prossimo paragrafo, risulta particolarmente fastidiosa per le EDS del tipo (2.8) in quanto tale fattore di scala viene a coincidere con uno dei parametri di interesse.

La restante parte di questo capitolo è dedicata alla derivazione dalle **C1** e **C2**, subordinatamente ad una scelta ragionata della funzione ϕ , di condizioni dei momenti di agevole interpretazione e calcolabilità, e quindi ad una proposta di stima GMM per le EDS del tipo (2.8) che fornisca una soluzione al problema del fattore di scala.

3.4 Applicazione delle condizioni **C1** e **C2** alle EDS “mean-reverting”

L'operatore differenziale o di Dynkin introdotto nel paragrafo 3.2 appare essere naturale candidato per un concreto utilizzo delle condizioni **C1** e **C2**. Tale operatore infatti è esplicitamente dipendente dai coefficienti della EDS, e risulta autoaggiunto per i processi di diffusione ergodici; il suo impiego pertanto comporta il notevole vantaggio di semplificare le espressioni derivanti dallo sviluppo delle condizioni e di introdurre in forma esplicita i parametri di interesse.

Come primo passo bisogna far riferimento ad una misura di probabilità che garantisca la stazionarietà del processo soluzione $\{X_t\}$.

Si consideri la seguente EDS

$$dx_t = \mu(x_t)dt + \sigma(x_t)dW_t \quad (3.3)$$

dove W_t è un moto Browniano. Seguendo Karlin e Taylor(1981) si introduce la *funzione di scala* S la cui derivata è:

$$s(y) = \exp \left[- \int^y 2[\mu(z)/\sigma^2(z)] dz \right].$$

Utilizzando la misura di probabilità \mathcal{P} con densità

$$p(y) = \frac{1}{s(y)\sigma^2(y) \int [s(z)\sigma^2(z)]^{-1} dz}$$

per inizializzare il processo, si ha che \mathcal{P} è l'unica distribuzione stazionaria che può essere associata alla soluzione della EDS. Pertanto, sotto la misura \mathcal{P} e sotto l'ipotesi che la soluzione $\{X_t\}$ sia un processo di Markov ergodico e reversibile, si può verificare che l'operatore infinitesimale \mathcal{A} introdotto nel paragrafo 3.2 è *autoaggiunto*, quindi $\mathcal{A} = \mathcal{A}^*$, ed è l'operatore di *Dynkin*; pertanto si ha che:

$$\mathcal{A}\phi = \mathcal{D}\phi = \mu(x_t)\frac{\partial}{\partial x_t}\phi(x_t) + \frac{1}{2}\sigma^2(x_t)\frac{\partial^2}{\partial x_t^2}\phi(x_t) \quad (3.4)$$

Si può considerare ora l'applicazione della condizione **C1** al modello generale:

$$dx_t = k(\theta - x_t)dt + \sigma x_t^\gamma dW_t \quad \gamma \geq 0 \quad (3.5)$$

La scelta della funzione ϕ è legata alla definizione dell'operatore infinitesimale e al fatto che in fase di stima i valori attesi presenti nelle espressioni delle condizioni **C1** e **C2** vengono approssimati attraverso il calcolo dei corrispondenti valori attesi campionari dalle serie osservate di lunghezza finita. Per giustificare questa approssimazione attraverso la Legge dei Grandi Numeri, e per garantire le proprietà asintotiche delle stime, è necessaria l'ipotesi che sotto la misura di probabilità \mathcal{P} il processo sia ergodico e che siano verificate alcune condizioni sul dominio degli operatori \mathcal{A} e \mathcal{A}^* . Hansen e Scheinkman (1995) dimostrano che è ancora valido il Teorema del Limite Centrale se i domini D e D^* sono sottoinsiemi propri dello spazio $L^2(\mathcal{P})$ costituiti da funzioni infinitamente derivabili.

Dal seguente teorema si possono quindi trarre delle valide indicazioni sulla scelta della funzione ϕ :

Teorema 3.4.1 *Se \mathcal{P} è una misura di probabilità su \mathbb{R} tale che per qualche $\alpha > 0$*

$$\int \exp(\alpha |x|) d\mathcal{P} < \infty$$

allora tutti i momenti di \mathcal{P} esistono e la successione dei polinomi ortogonali è completa in $L^2(\mathcal{P})$. Nel caso in cui \mathcal{P} è una legge gaussiana i polinomi di Hermite:

$$\phi_i(x_t) = (-1)^i \exp\left(\frac{1}{2}x_t^2\right) \left(\frac{\partial^i}{\partial x_t^i} \exp\left(-\frac{1}{2}x_t^2\right)\right),$$

con $i=0,1,2,\dots$ e $\phi_0(x_t) = 1$, costituiscono una base ortogonale completa. □

Tale teorema permette allora di individuare una classe di funzioni per le quali sono verificate le condizioni richieste; si ritiene, quindi, opportuno restringere l'attenzione ai polinomi di Hermite e alle funzioni esponenziali:

$$\phi(x_t) = \exp(\alpha x_t) \quad \text{con } \alpha \in \mathbb{R}.$$

Quest'ultima presenta il vantaggio di poter essere facilmente ricondotta alla funzione generatrice dei momenti della variabile aleatoria $\{X_t\}$.

Applicando quindi la condizione **C1** al modello (3.5) per $\phi(x_t) = \exp(\alpha x_t)$, $\alpha \in \mathbb{R}$ si ha che:

$$\mathbf{E}[\mathcal{D}\phi(x_t)] = \mathbf{E}[k(\theta - x_t)\frac{\partial}{\partial x_t}\exp(\alpha x_t) + \frac{1}{2}\sigma^2 x_t^{2\gamma}\frac{\partial^2}{\partial x_t^2}\exp(\alpha x_t)] = 0$$

Per la linearità del valore atteso e indicata con $\mathcal{M}_{x_t}(\alpha) = \mathbf{E}[e^{\alpha x_t}]$ la funzione generatrice dei momenti, si ottiene la seguente espressione in α :

$$\mathcal{G}(\alpha) = \theta\mathcal{M}_{x_t}(\alpha) - \frac{\partial\mathcal{M}_{x_t}(\alpha)}{\partial\alpha} + \frac{\alpha\sigma^2}{2k}\frac{\partial^{2\gamma}\mathcal{M}_{x_t}(\alpha)}{\partial\alpha^{2\gamma}} = 0$$

Calcolando la derivata i -esima rispetto ad α nel punto $\alpha = 0$ ed indicato con η il vettore dei parametri, si può scrivere:

$$\mathbf{E}[f_i(x_t, \eta)] = \theta\mathbf{E}[x_t^i] + i\frac{\sigma^2}{2k}\mathbf{E}[x_t^{2\gamma+i-1}] - \mathbf{E}[x_t^{i+1}] = 0 \quad i=0,1,2,\dots \quad (3.6)$$

Con procedimento del tutto analogo si può applicare la condizione **C2** alle funzioni $\phi(x_t) = \exp(\alpha x_t)$ e $\phi^*(x_t) = \exp(\beta x_t)$.

$$\mathbf{E}[\phi^*(x_t)\mathcal{D}\phi(x_{t+1}) - \phi(x_{t+1})\mathcal{D}\phi^*(x_t)] = 0$$

$$\begin{aligned} & \mathbf{E}[\exp(\beta x_t)[k(\theta - x_{t+1})\frac{\partial}{\partial x_{t+1}}\exp(\alpha x_{t+1}) + \frac{1}{2}\sigma^2 x_{t+1}^{2\gamma}\frac{\partial^2}{\partial x_{t+1}^2}\exp(\alpha x_{t+1})]] = \\ & = \mathbf{E}[\exp(\alpha x_{t+1})[k(\theta - x_t)\frac{\partial}{\partial x_t}\exp(\beta x_t) + \frac{1}{2}\sigma^2 x_t^{2\gamma}\frac{\partial^2}{\partial x_t^2}\exp(\beta x_t)]] \end{aligned}$$

$$\begin{aligned} & \alpha\theta\mathbf{E}[\exp(\alpha x_{t+1} + \beta x_t)] - \alpha\mathbf{E}[x_{t+1}\exp(\alpha x_{t+1} + \beta x_t)] + \\ & \quad \frac{\alpha^2}{2}\frac{\sigma^2}{k}\mathbf{E}[x_{t+1}^{2\gamma}\exp(\alpha x_{t+1} + \beta x_t)] = \\ & = \beta\theta\mathbf{E}[\exp(\alpha x_{t+1} + \beta x_t)] - \beta\mathbf{E}[x_t\exp(\alpha x_{t+1} + \beta x_t)] \\ & \quad + \frac{\beta^2}{2}\frac{\sigma^2}{k}\mathbf{E}[x_t^{2\gamma}\exp(\alpha x_{t+1} + \beta x_t)] \end{aligned}$$

Calcolando la derivata seconda mista rispetto ad α e β nel punto $(\alpha = 0, \beta = 0)$, si ottiene:

$$2\theta\mathbf{E}[x_t^2 x_{t+1}] + \frac{\sigma^2}{k}\mathbf{E}[x_{t+1}^{2\gamma} x_t^2] - 2\theta\mathbf{E}[x_{t+1}^2 x_t] - \frac{\sigma^2}{k}\mathbf{E}[x_t^{2\gamma} x_t^2] = 0$$

Altre espressioni di questo tipo possono essere ricavate per derivazioni successive miste sempre nel punto $(\alpha = 0, \beta = 0)$. Non è emersa però in modo evidente una formula

ricorsiva come per l'espressione (3.6).

Si applichino ora le condizioni **C1** ai polinomi di Hermite di grado i . Dalla formulazione generale del polinomio di Hermite per $i=1,2,3,\dots$ si ha che:

$$\phi_0(x_t) = 0, \phi_1(x_t) = x_t, \phi_2(x_t) = x_t^2 - 1, \phi_3(x_t) = x_t^3 - 3x_t, \phi_4(x_t) = x_t^4 - 6x_t^2 + 3, \text{ etc, etc ...}$$

Sostituendo tali polinomi nella condizione **C1** abbiamo:

$$\begin{aligned} \mathbf{E}[\mathcal{D}\phi_1(x_t)] &= \mathbf{E}[k(\theta - x_t)], \\ \mathbf{E}[\mathcal{D}\phi_2(x_t)] &= \mathbf{E}[(\theta - x_t)x_t + \frac{\sigma^2}{2k}x_t^{2\gamma}], \\ \mathbf{E}[\mathcal{D}\phi_3(x_t)] &= \mathbf{E}[(\theta - x_t)(x_t^2 - 1) + \frac{\sigma^2}{k}x_t^{2\gamma+1}], \\ \mathbf{E}[\mathcal{D}\phi_4(x_t)] &= \mathbf{E}[(\theta - x_t)(x_t^3 - 3x_t) + \frac{3\sigma^2}{2k}x_t^{2\gamma}(x_t^2 - 1)], \dots \end{aligned}$$

Si può quindi scrivere il seguente sistema di condizioni di ortogonalità:

$$\begin{aligned} \mathbf{E}[\mathcal{D}\phi_{i+1}(x_t)] &= \mathbf{E}[f_i(x_t, \eta)] \\ &= \mathbf{E}[(\theta - x_t)\phi_i(x_t) + i\frac{\sigma^2}{2k}x_t^{2\gamma}\phi_{i-1}(x_t)] = 0 \end{aligned} \quad (3.7)$$

con $i = 0, 1, 2, \dots$ e $\phi_{-1}(x_t) = 0$. Come si può notare, le condizioni (3.6) all'aumentare dell'ordine di derivazione i sviluppano la successione dei momenti del processo $\{X_t\}$; le condizioni (3.7) per $i = 1, 2$ danno il momento primo e secondo della distribuzione marginale di $\{X_t\}$, e quindi coincidono con le (3.6), e per $i > 2$ forniscono come vincoli delle combinazioni lineari di tutti i momenti fino al grado $(i + 1)$ -esimo.

Le condizioni (3.6) e (3.7) forniscono allora validi vincoli sui momenti del processo soluzione $\{X_t\}$ che possono essere utilmente impiegate per una stima di tipo GMM dei parametri delle EDS, in particolare nei casi delle EDS utilizzate nei modelli di Vasicek e CIR illustrati nel Capitolo 1.

Le condizioni di Hansen e Scheinkman ora descritte non sembrano invece particolarmente utili nel caso dei modelli a fattori latenti (Gourieroux e Monfort 1994). Si consideri, ad esempio, il modello di Hull e White, descritto brevemente nel paragrafo 2.7, nella forma studiata da Pastorello, Touzi e Renault (1993):

$$\begin{cases} dX_t = \mu X_t dt + \sigma_t X_t d\mathcal{W}_t^x \\ d \log \sigma_t = k(\theta - \log \sigma_t) dt + \sigma d\mathcal{W}_t^\sigma. \end{cases} \quad (3.8)$$

Si introduca il processo bivariato $y_t = \begin{pmatrix} X_t \\ \log \sigma_t \end{pmatrix}$ a cui è associata:

$$dy_t = \begin{bmatrix} \mu y_{1t} \\ k(\theta - y_{2t}) \end{bmatrix} dt + \begin{bmatrix} y_{1t} \exp(y_{2t}) & 0 \\ 0 & \sigma \end{bmatrix} d\mathcal{W}_t.$$

In modo analogo a quanto precedentemente descritto applicando l'operatore di Dynkin si ha:

$$\mathcal{D}\phi(y_t) = \mu y_{1t} \frac{\partial}{\partial y_1} \phi(y_t) + k(\theta - y_{2t}) \frac{\partial}{\partial y_2} \phi(y_t) + \frac{1}{2} \left(y_{1t}^2 \exp(2y_{2t}) \frac{\partial^2}{\partial y_1^2} \phi(y_t) + \sigma^2 \frac{\partial^2}{\partial y_2^2} \phi(y_t) \right).$$

Poiché la sola variabile osservabile è $\{X_t\}$, che nel caso del modello di Hull e White indica il prezzo di un titolo, la condizione dei momenti si può basare su un operatore di Dynkin che dipenda solo da y_{1t} . Questo significa imporre che:

$$y_{1t} \frac{\partial}{\partial y_1} \phi(y_t), \quad \frac{\partial}{\partial y_2} \phi(y_t), \quad y_{1t}^2 \exp(2y_{2t}) \frac{\partial^2}{\partial y_1^2} \phi(y_t), \quad \frac{\partial^2}{\partial y_2^2} \phi(y_t)$$

siano simultaneamente indipendenti da y_{2t} e ciò si verifica se e solo se la funzione $\phi(y_t)$ non dipende da y_{2t} . In questa situazione dunque l'approccio proposto non fornisce condizioni dei momenti interessanti.

3.5 Lo stimatore GMM

Come illustrato nel paragrafo 2.5, il metodo di Pseudo Massima Verosimiglianza permette di stimare in modo consistente i parametri presenti nella media condizionata dalle variabili esplicative di un modello ma non quelli coinvolti dalla varianza condizionata.

L'approccio di tipo GMM proposto in questo lavoro consente in linea generale di risolvere i problemi legati ai parametri della varianza dovuti alla cattiva specificazione del modello approssimato. La metodologia generale dei GMM si fonda sul fatto che se un modello è correttamente specificato allora certe condizioni sui momenti sono verificate dal valore vero dei parametri.

Nel contesto logico nel quale opera l'approccio GMM proposto, non è presente errata specificazione. Infatti le condizioni (3.6) e (3.7), che costituiscono il modello, sono derivate esclusivamente dalle proprietà dell'operatore di Dynkin, senza dover esplicitare la funzione di Verosimiglianza associata al modello oggetto di studio. Quindi un'eventuale errata specificazione sarebbe derivata solo dall'uso improprio di tale operatore. L'impiego dell'operatore di Dynkin è giustificato se il processo oggetto di studio gode delle proprietà di ergodicità, reversibilità e stazionarietà, che nel nostro contesto si assumono verificate.

Le condizioni di ortogonalità (3.6) e (3.7), però, mettono bene in evidenza un risultato già noto in letteratura sugli operatori infinitesimali di una EDS: tali operatori sono identificati a meno di un fattore di scala. Nel nostro caso tale fattore di scala viene a coincidere con il parametro k della EDS. Il problema risulta poi particolarmente fastidioso in quanto tale parametro diventa indistinguibile dalla diffusione σ che è di particolare interesse nel contesto delle applicazioni delle EDS per i modelli finanziari. Si noti peraltro che il parametro γ della (3.5) risulta ben identificato.

Tale problema può rendere di per sè vano l'utilizzo della metodologia GMM proposta a meno di non riesaminare i termini del problema della stima dei parametri della (3.5)

e considerare come parametri di interesse i soli σ e γ . In tale impostazione i parametri k e θ possono essere trattati come *parametri di disturbo*.

Tra i diversi criteri proposti in letteratura per la trattazione dei parametri di disturbo è stato ritenuto opportuno seguire il metodo della “concentrazione” (Burguete, Galant e Souza, 1982). L’idea è quella di ottenere per i parametri di interesse σ e γ delle stime GMM fondate sulle (3.6), o sulle (3.7), condizionati a stime consistenti dei disturbi k e θ ottenute per altra via. Gourieroux, Monfort e Renault (1991) dimostrano che le stime GMM *a due stadi* e quelle con parametri di disturbo stimabili in modo consistente, sono asintoticamente normali e consistenti sotto alcune estensioni delle condizioni di regolarità proposte da Hansen (1982).

Si indichi dunque con $\eta = (\sigma, \gamma)$ e con $\nu = (k, \theta)$ rispettivamente il vettore dei parametri di interesse e di disturbo; pertanto le restrizioni dei momenti assumono la forma: $\mathbf{E}[f_i(x_t, \eta, \nu)] = 0$.

Come mostrato nel paragrafo 2.5, i parametri $\nu = (k, \theta)$ sono stimati consistentemente attraverso il metodo della PML.

Considerata quindi la Pseudo Verosimiglianza:

$$x_{t+1} | x_t \sim \mathcal{N}(\theta(1 - \exp(-k)) + \exp(-k)x_t, \sigma^2 \exp(-k)x_t^{2\gamma}) \quad (3.9)$$

siano $\hat{\nu} = (\hat{k}, \hat{\theta})$ le stime ottenute dalla condizione dei Minimi Quadrati:

$$\mathbf{E}[x_{t+1} | x_t] = \theta(1 - \exp(-k)) + \exp(-k)x_t.$$

Pertanto la (3.6) assume la forma:

$$\mathbf{E}[f_i(x_t, \eta, \hat{\nu}) | \hat{\nu}] = \hat{\theta} \mathbf{E}[x_t^i] + i \frac{\sigma^2}{2\hat{k}} \mathbf{E}[x_t^{2\gamma+i-1}] - \mathbf{E}[x_t^{i+1}] = 0 \quad \forall i \in \mathbb{N}, \quad (3.10)$$

e la (3.7) assume la forma:

$$\mathbf{E}[f_i(x_t, \eta, \hat{\nu}) | \hat{\nu}] = \mathbf{E} \left[(\hat{\theta} - x_t) \phi_i(x_t) + i \frac{\sigma^2}{2\hat{k}} x_t^{2\gamma} \phi_{i-1}(x_t) | \hat{\nu} \right] = 0$$

con $i = 0, 1, 2, \dots$ e $\phi_{-1}(x_t) = 0$.

Le stime GMM dei parametri $\eta = (\sigma, \gamma)$ sono determinate dalla soluzione del problema di minimo della forma quadratica:

$$\hat{\eta}(\Omega) = \underset{\eta}{\operatorname{argmin}} \left(\frac{1}{T} \sum_{t=1}^T f(x_t, \eta, \hat{\nu}) \right)' \Omega_t^{-1} \left(\frac{1}{T} \sum_{t=1}^T f(x_t, \eta, \hat{\nu}) \right) \quad (3.11)$$

dove, come di consueto, il valore atteso è stimato tramite la media campionaria, Ω_t è una matrice simmetrica definita positiva ed f indica il vettore delle condizioni.

Un’espressione della matrice di varianza e covarianza asintotica delle stime GMM così ottenute può essere derivata adattando alle nostre esigenze quella proposta da Gourieroux, Monfort e Renault (1991) per le stime GMM a due stadi. Si consideri quindi le condizioni dei Momenti: $\mathbf{E}[f(x_t, \eta, \nu)] = 0$ dove η è il vettore dei parametri

di interesse, e ν è il vettore dei parametri di disturbo stimabili in modo consistente. Si suppone che le stime $\hat{\nu}$ siano ottenibili attraverso il metodo GMM applicato su restrizioni del tipo $\mathbf{E}[g_i(x_t, \nu)] = 0$.

Lo stimatore dei GMM $\hat{\eta}(\Omega)$ può essere ottenuto come soluzione della condizione di minimo di primo ordine:

$$\left(\frac{1}{T} \sum_{t=1}^T \frac{\partial f}{\partial \eta'}(x_t, \eta(\Omega_t), \hat{\nu}) \right) \Omega_t^{-1} \left(\frac{1}{T} \sum_{t=1}^T f(x_t, \eta(\Omega_t), \hat{\nu}) \right) = 0$$

Dallo sviluppo del primo ordine intorno al valore vero dei parametri³ η_0 e ν_0 si ottiene:

$$\begin{aligned} \sqrt{T}(\hat{\eta}(\Omega_t) - \eta_0) \approx \\ - \left[\mathbf{E} \frac{\partial f}{\partial \eta'} \Omega_t^{-1} \mathbf{E} \frac{\partial f}{\partial \eta} \right]^{-1} \mathbf{E} \frac{\partial f}{\partial \eta'} \Omega_t^{-1} \left[\frac{1}{\sqrt{T}} \sum_{t=1}^T f_t + \mathbf{E} \frac{\partial f}{\partial \nu'} \sqrt{T}(\hat{\nu} - \nu_0) \right] \end{aligned} \quad (3.12)$$

dove

$$f_t = f(x_t, \eta_0, \nu_0), \quad \mathbf{E} \frac{\partial f}{\partial \eta'} = \mathbf{E} \frac{\partial f(x_t, \eta_0, \nu_0)}{\partial \eta'}, \quad \mathbf{E} \frac{\partial f}{\partial \nu'} = \mathbf{E} \frac{\partial f(x_t, \eta_0, \nu_0)}{\partial \nu'}.$$

Se il vettore $\frac{1}{\sqrt{T}} \sum_{t=1}^T f_t$ e il vettore $\sqrt{T}(\hat{\nu} - \nu_0)$ sono asintoticamente normali⁴, allora $\sqrt{T}(\hat{\eta}(\Omega_t) - \eta_0)$ è asintoticamente normale con matrice di covarianza pari a:

$$Var(\hat{\eta}(\Omega_t) - \eta_0) =$$

$$\left[\mathbf{E} \frac{\partial f}{\partial \eta'} \Omega_t^{-1} \mathbf{E} \frac{\partial f}{\partial \eta} \right]^{-1} \mathbf{E} \frac{\partial f}{\partial \eta'} \Omega_t^{-1} \Omega_\star \Omega_t^{-1} \mathbf{E} \frac{\partial f}{\partial \eta} \left[\mathbf{E} \frac{\partial f}{\partial \eta'} \Omega_t^{-1} \mathbf{E} \frac{\partial f}{\partial \eta} \right]^{-1} \quad (3.13)$$

dove

$$\Omega_\star = Var \left(\frac{1}{\sqrt{T}} \sum_{t=1}^T f_t + \mathbf{E} \frac{\partial f}{\partial \nu'} \sqrt{T}(\hat{\nu} - \nu_0) \right). \quad (3.14)$$

Come di consueto esiste una scelta ottimale per la matrice dei pesi Ω_t tale che minimizza la varianza dello stimatore $\hat{\eta}(\Omega_t)$. Ponendo infatti $\Omega_t = \Omega_\star$ la (3.13) diventa:

$$Var(\hat{\eta}(\Omega_\star) - \eta_0) = \left(\mathbf{E} \frac{\partial f}{\partial \eta'} \Omega_\star^{-1} \mathbf{E} \frac{\partial f}{\partial \eta} \right)^{-1}. \quad (3.15)$$

Un'espressione della matrice Ω_\star può essere ricavata in base alle considerazioni che seguono. Si suppone che le stime $\hat{\nu}$ siano ottenibili attraverso il metodo GMM applicato alle restrizioni $\mathbf{E}[g(x_t, \nu)] = 0$. Applicando lo sviluppo di Taylor alla condizione di

³Per una descrizione completa di tale derivazione si veda l'Appendice D

⁴Tale condizione è verificata in base al Teorema del Limite Centrale se il vettore f e il vettore g , che esprime le condizioni dei GMM per la stima di ν , soddisfano le condizioni di regolarità proposte da Gourieroux, Monfort e Renault (1991) riportate nell'Appendice E

minimo di primo ordine della forma quadratica per la stima GMM del vettore ν , in modo analogo a quanto sopra illustrato si ha che:

$$\sqrt{T}(\hat{\nu} - \nu_0) = - [\Gamma' \Sigma \Gamma]^{-1} \Gamma' \Sigma \frac{1}{\sqrt{T}} \sum_{t=1}^T g_t$$

dove $g_t = g(x_t, \nu)$, $\Gamma = \mathbf{E}[\frac{\partial g_t}{\partial \nu}]$, e Σ è pari alla scelta ottimale $(\text{Var}(\frac{1}{T} \sum_{t=1}^T g_t))^{-1}$. Dalla teoria generale dei GMM si ha che:

$$\text{Var}(\sqrt{T}(\hat{\nu} - \nu_0)) = [\Gamma' \Sigma \Gamma]^{-1}.$$

Tenuto conto di quanto detto la matrice (3.14) può essere sviluppata nel modo seguente:

$$\begin{aligned} \Omega_\star &= \text{Var} \left(\frac{1}{\sqrt{T}} \sum_{t=1}^T f_t + \mathbf{E} \frac{\partial f}{\partial \nu'} \sqrt{T}(\hat{\nu} - \nu_0) \right) \\ &= \text{Var} \left(\frac{1}{\sqrt{T}} \sum_{t=1}^T f_t \right) + \mathbf{E} \frac{\partial f}{\partial \nu'} \text{Var}(\sqrt{T}\hat{\nu} - \nu_0) \mathbf{E} \frac{\partial f}{\partial \nu} \\ &\quad - 2 \mathbf{E} \frac{\partial f}{\partial \nu'} [\Gamma' \Sigma \Gamma]^{-1} \Gamma' \Sigma \text{Cov} \left(\frac{1}{\sqrt{T}} \sum_{t=1}^T g_t, \frac{1}{\sqrt{T}} \sum_{t=1}^T f_t \right) \\ &= \Omega(f_t) + \mathbf{E} \frac{\partial f}{\partial \nu'} [\Gamma' \Sigma \Gamma]^{-1} \mathbf{E} \frac{\partial f}{\partial \nu} - 2 \mathbf{E} \frac{\partial f}{\partial \nu'} [\Gamma' \Sigma \Gamma]^{-1} \Gamma' \Sigma \mathbf{E} \left[\frac{1}{\sqrt{T}} \sum_{t=1}^T g_t \frac{1}{\sqrt{T}} \sum_{t=1}^T f_t \right], \end{aligned} \tag{3.16}$$

dove $\Omega(f_t)$ indica la matrice di covarianza per le condizioni dei GMM relativa a $\frac{1}{\sqrt{T}} \sum_{t=1}^T f_t$.

Come si può notare nella matrice proposta si distinguono in particolare due componenti: una è la matrice di covarianza della stima GMM dei parametri η che discende dalla teoria tradizionale, la seconda è il contributo della varianza della stima dei parametri di disturbo. A proposito della stima di tali componenti si possono fare alcune considerazioni: si ha ragione di ritenere che le condizioni (3.6) impiegate nei GMM proposti presentino un certo grado di autocorrelazione seriale, di cui è necessario tener conto nella stima della matrice $\Omega(f_t)$. In letteratura sono stati proposti diversi stimatori della forma:

$$\Omega(f_t) = \frac{T}{T-p} \sum_{j=-T+1}^{T-1} \mathcal{K} \left(\frac{j}{S_T} \right) \Phi_T(j) \tag{3.17}$$

dove \mathcal{K} è un Nucleo a valori reali, S_T è un parametro di troncamento, p è l'ordine di autocorrelazione, (che in molte applicazioni risulta ignoto), e

$$\Phi_T(j) = \frac{1}{T} \sum_{t=j+1}^T [f_t f'_{t-j}].$$

Nel corso di questo lavoro verrà utilizzata la versione della (3.17) proposta da Newey e West (1987):

$$\Omega(f_t) = \Phi_T(0) + \sum_{j=1}^p \left(1 - \frac{j}{p+1}\right) [\Phi_T(j) + \Phi_T(j)']. \quad (3.18)$$

Nella derivazione della matrice Ω_* è stato supposto che la stima dei parametri di disturbo possa essere ottenuta attraverso i GMM con condizioni del tipo $\mathbf{E}[g(x_t, \nu)] = 0$, mentre nel corso della presentazione degli stimatori HS-GMM si sono indicate le stime dei Minimi Quadrati tratte dalla Pseudo Verosimiglianza (3.9). Tali assunzioni non sono in contraddizione in quanto sia le stime dei Minimi Quadrati sia quelle di PML possono essere ricondotte a stime di tipo GMM. Nel primo caso infatti le condizioni di ortogonalità tra le variabili esogene e gli errori possono essere lette come condizioni GMM, nel secondo caso ciò avviene per le condizioni di primo ordine del problema di massimo che determina le stime (Newey e McFadden, 1994).

3.6 Uno studio Monte Carlo

Come è stato mostrato nel corso di questo capitolo, lo stimatore GMM (3.11) derivato dalle condizioni di Hansen e Scheinkman (HS-GMM nel seguito) si presenta come una delle possibili soluzioni al problema della stima consistente del parametro di diffusione di una EDS. In particolare esso può presentare un'alternativa al metodo dell'Inferenza Indiretta.

In questo paragrafo sarà illustrata una valutazione degli aspetti econometrici dello stimatore proposto attraverso uno studio Monte Carlo e sarà operato un confronto con il metodo dell'Inferenza Indiretta. Gli stimatori HS-GMM saranno applicati al processo di Ornstein-Uhlenbeck usato da Vasicek (1977) e al processo "radice quadrata" che è alla base del modello CIR (1985).

Il modello di riferimento è dato dalla EDS (3.5) e saranno distinti due casi: quello in cui il parametro γ di elasticità della varianza è incognito, che verrà indicato come modello non vincolato, e quello in cui tale parametro è noto, ed assume i valori $\gamma = 0$ e $\gamma = 0.5$, che verrà indicato come modello vincolato.

3.6.1 Stima del processo di Ornstein-Uhlenbeck

Nel Capitolo 2 è stato brevemente presentato il processo di Ornstein-Uhlenbeck

$$dX_t = k(\theta - X_t)dt + \sigma dW_t \quad (3.19)$$

come modello vincolato dalla EDS (3.5) per $\gamma = 0$. Per il processo (3.19) è stata determinata nel paragrafo 2.2 la soluzione condizionata a $X_s = x$, con $s < t$, espressa dalla (2.5). Supponendo quindi di osservare le realizzazioni del processo $\{X_t\}$ ad intervalli equispaziati $\delta t = t_i - t_{i-1}$, dalla (2.5) si può ricavare la legge $X_{t_i+\delta t} | X_{t_i}$ che

può essere letta come discretizzazione “esatta” del processo (3.19). Si ha, quindi, che:

$$X_{t_i+\delta t} = \theta(1 - e^{-k\delta t}) + e^{-k\delta t}X_{t_i} + \sigma \left(\frac{1 - e^{-2k\delta t}}{2k} \right)^{1/2} \epsilon_t \quad (3.20)$$

dove $\epsilon_t \sim \mathcal{N}(0, 1)$ e $\delta t = t_i - t_{i-1}$, $i=1,2,\dots,T$. Le traiettorie del processo di Ornstein-Uhlenbeck possono quindi essere simulate tramite la (3.20).

A questo proposito Gourieroux, Monfort e Renault (1993) suggeriscono la seguente strategia: le realizzazioni del processo (3.19) vengono osservate ad ogni istante t_i . La simulazione di una traiettoria in tempo continuo sottostante il processo discreto osservato può avvenire generando tra $X_{t_{i-1}}$ e X_{t_i} un numero di osservazioni intermedie. Il reciproco di tale numero viene indicato dagli Autori come *passo di simulazione* δt ed essi suggeriscono $\delta t = 0.1$. Di conseguenza il processo osservato X_{t_i} , $i = 0, 1, \dots, T$ consisterà nell'estrazione dalla traiettoria simulata di un dato ogni 10, definendo quindi implicitamente un *passo di osservazione* $\Delta t = 1$. Pertanto la (3.20) viene assunta come modello generatore dei dati per $\delta t = 0.1$, mentre la funzione di Pseudo Verosimiglianza dalla quale saranno ricavate le stime di PML sarà la legge del processo:

$$X_{t+1} = \theta(1 - e^{-k}) + e^{-k}X_t + \sigma \left(\frac{1 - e^{-2k}}{2k} \right)^{1/2} \epsilon_t \quad (3.21)$$

dove $\epsilon_t \sim \mathcal{N}(0, 1)$ e $t=1,2,\dots,T$.

Si consideri ora il caso in cui si voglia procedere alla stima del modello (3.19).

In tale contesto il vettore dei parametri di interesse si riduce al solo σ e le condizioni dei HS-GMM (3.10) assumono una forma del tipo:

$$\mathbf{E}[f_i(x_t, \sigma, \hat{\nu}) | \hat{\nu}] = \hat{\theta}\mathbf{E}[x_t^i] + i\frac{\sigma^2}{2k}\mathbf{E}[x_t^{i-1}] - \mathbf{E}[x_t^{i+1}] = 0 \quad \forall i \in \mathbb{N} \quad (3.22)$$

Da tale equazione per $i = 1$ si ottiene l'espressione della varianza della distribuzione marginale di lungo periodo (Arnold 1975, pag 135) $X_t \sim \mathcal{N}(\theta, \frac{\sigma^2}{2k})$; pertanto per la stima HS-GMM (3.11) di σ basta considerare la sola condizione di ortogonalità:

$$\mathbf{E}[f_1(x_t, \sigma, \hat{\nu}) | \hat{\nu}] = \hat{\theta}\mathbf{E}[x_t] + \frac{\sigma^2}{2k} - \mathbf{E}[x_t^2] = 0 \quad (3.23)$$

e la varianza delle stime sarà indipendente dalla matrice dei pesi scelta, che sarà pertanto posta pari alla matrice identità. Tale condizione può essere ottenuta anche a partire dalle (3.7) per $i = 1$, assegnato il valore $\gamma = 0$. Le stime dei parametri di disturbo sono ottenute in accordo con quanto descritto nel paragrafo precedente. Si ricorda che in questo contesto la legge (3.21) viene indicata come Pseudo Verosimiglianza in quanto rappresenta un'approssimazione discreta alla soluzione del processo che è tanto più accurata al tendere di δt a zero. Il valore proposto $\delta t = 0.1$ come passo di simulazione dei dati è sufficiente perché le stime basate su tale Pseudo Verosimiglianza presentino una distorsione trascurabile.

Le stime di Inferenza Indiretta sono ottenute sulla base dello stimatore (2.16) proposto da Gourieroux, Monfort e Renault (1993) e descritto nel capitolo precedente, dove il

modello generatore dei dati è dato dalla (3.20), e dove il modello ausiliario è dato dalla discretizzazione di Eulero del processo (3.19):

$$X_{t+1} = k\theta + (1 - k)X_t + \sigma\epsilon_t$$

con $\epsilon_t \sim \mathcal{N}(0, 1)$.

poiché il modello strutturale e il modello ausiliario hanno lo stesso numero di parametri, lo stimatore (2.16) non dipende dalla matrice dei pesi Ω che pertanto viene assunta pari alla matrice identità.

Nella Tabella 1 sono riportati i risultati di un esperimento Monte Carlo di 200 repliche effettuato su serie osservate di lunghezza $T=250$ con $\theta = 0.1, k = 0.8, \sigma = 0.06$ in accordo a quanto presentato da Gourieroux, Monfort e Renault (1993).

Come si può notare lo stimatore HS-GMM per il parametro di diffusione σ presenta un comportamento simile a quello degli altri stimatori proposti, ed in particolare il suo errore standard è migliore di quello delle stime di Inferenza Indiretta.

Tab.1: Processo Ornstein-Uhlenbeck Num. Repl.=200, T=250

Metodo di Stima	Parametri Valore vero	θ 0.1	k 0.8	σ 0.06
PML	Media	0.1001 (0.0050)	0.8219 (0.1211)	0.0599 (0.0039)
	Distorsione	0.0001	0.0019	-.0001
Ind.Inf	Media	0.0997 (0.0064)	0.8187 (0.2047)	0.0611 (0.0064)
	Distorsione	-.0003	-.0187	0.0011
HS-GMM	Media	—	—	0.0599 (0.0039)
	Distorsione			-.0001

Nella Tabella 2 sono riportati i risultati di un esperimento Monte Carlo di 200 repliche su serie di lunghezza $T=1000$, allo scopo di presentare una verifica della consistenza dello stimatore proposto.

Tab.2: Processo Ornstein-Uhlenbeck Num. Repl.=200, T=1000

Metodo di Stima	Parametri Valore vero	θ 0.1	k 0.8	σ 0.06
PML	Media	0.0999 (0.0023)	0.8052 (0.0616)	0.0600 (0.0019)
	Distorsione	-.0001	0.0052	—
Ind.Inf	Media	0.0995 (0.0033)	0.8064 (0.0921)	0.0600 (0.0024)
	Distorsione	-.0005	0.064	—
HS-GMM	Media	—	—	0.0600 (0.0019)
	Distorsione	—	—	—

Procediamo ora alla stima del modello (3.5) non vincolato,

$$dX_t = k(\theta - X_t)dt + \sigma X_t^\gamma d\mathcal{W}_t.$$

In tale contesto le condizioni di ortogonalità presentano la forma data dalla (3.10) e lo stimatore HS-GMM è dato dalla (3.11); anche in questo caso sono stati impiegati i HS-GMM esattamente identificati, pertanto le stime saranno basate sulle condizioni (3.10) per $i = 1, 2$:

$$\begin{cases} \mathbf{E}[f_1(x_t, \eta, \hat{\nu}) | \hat{\nu}] = \hat{\theta}\mathbf{E}[x_t] + \frac{\sigma^2}{2k}\mathbf{E}[x_t^{2\gamma}] - \mathbf{E}[x_t^2] = 0 \\ \mathbf{E}[f_2(x_t, \eta, \hat{\nu}) | \hat{\nu}] = \hat{\theta}\mathbf{E}[x_t^2] + \frac{\sigma^2}{k}\mathbf{E}[x_t^{2\gamma+1}] - \mathbf{E}[x_t^3] = 0 \end{cases} \quad (3.24)$$

con matrice dei pesi pari alla matrice identità.

Le stime di PML sono basate sulla (2.12) del capitolo precedente con $g(X_t, \gamma) = X_t^\gamma$ e passo di osservazione $\Delta t = 1$; pertanto la funzione di Pseudo Verosimiglianza avrà una forma del tipo:

$$X_{t+1} = \theta(1 - e^{-k}) + e^{-k} X_t + e^{-k} \sigma X_t^\gamma \epsilon_t$$

dove $\epsilon_t \sim \mathcal{N}(0, 1)$.

Le stime di Inferenza Indiretta sono ottenute in base allo stimatore (2.16) dove il modello generatore dei dati è la discretizzazione di Eulero della (2.10) con $g(X_t, \gamma) = X_t^\gamma$:

$$X_{t+1} = \theta k \delta t + (1 - k \delta t) X_t + \sigma \delta t X_t^\gamma \epsilon_t \quad (3.25)$$

con passo di simulazione $\delta t = 0.1$ e passo di osservazione $\Delta t = 1$. Il modello ausiliario può essere determinato in base alle seguenti considerazioni: (analogamente a Calzolari e Fiorentini (1994)) posto $a = \theta k$ e $b = (1 - k)$, la (3.25) può essere riscritta come:

$$\frac{X_{t+1} - a - bX_t}{X_t^\gamma} = u_t \quad (3.26)$$

con $u_t \sim N(0, \sigma^2)$ i.i.d. Indicato, quindi, con $\beta = (\theta, k, \sigma, \gamma)$ il vettore dei parametri del modello ausiliario, con y_t la variabile endogena e con x_t la variabile endogena ritardata, la (3.26) può essere scritta nella forma:

$$f(y_t, x_t, \beta) = u_t.$$

Quest'ultima espressione può essere letta come caso particolare di un sistema di equazioni non lineari simultanee (sistema che in questo caso è costituito da una sola equazione), al quale è possibile applicare la stima di Massima Verosimiglianza a Informazione Completa non lineare (non linear FIML) proposta da Amemiya (1977). Tale stima è ottenuta iterando fino a convergenza un metodo di variabili strumentali del tipo:

$$\hat{\beta}_{(m+1)} = \hat{\beta}_{(m)} - [\hat{G}'_{(m)} G_{(m)}]^{-1} \hat{G}'_{(m)} \hat{u}_{(m)}$$

dove $G'_{(m)}$ è la matrice jacobiana calcolata nel punto $\hat{\beta}_{(m)}$, sia $g_{t,(m)} = \frac{\partial f}{\partial \beta} |_{\hat{\beta}_{(m)}}$ la sua t -esima colonna, e $\hat{G}'_{(m)}$ è una matrice le cui colonne sono date da:

$$\hat{g}_{t,(m)} = g_{t,(m)} - \left(\frac{1}{T} \sum_{s=1}^T \frac{\partial g_s}{\partial u_s} \right) \hat{u}_{t,(m)}$$

calcolata nel punto $\hat{\beta}_{(m)}$. Attraverso questa procedura di stima vengono calcolati i vettori $\hat{\beta}$ e $\tilde{\beta}$ dello stimatore (2.16). Anche in questo caso lo stimatore (2.16) è indipendente dalla matrice dei pesi Ω che viene pertanto posta pari alla matrice identità.

I risultati dell'esperimento Monte Carlo con 200 replicazioni su serie di lunghezza $T=250$, $\theta = 0.25$, $k = 0.8$, $\sigma = 0.06$, $\gamma = 0$ sono riportati nella Tabella 3. Come si può vedere gli HS-GMM hanno un comportamento pessimo: il parametro di diffusione è ampiamente distorto, così come il parametro γ ed in entrambi i casi l'errore standard è di ampiezza rilevante. Migliore, ma non completamente soddisfacente è il comportamento delle stime di Inferenza Indiretta.

Tab.3: Processo Ornstein-Uhlenbeck Num. Repl.=200, T=250

Metodo di Stima	Parametri Valore vero	θ 0.25	k 0.8	σ 0.06	γ 0.0
PML	Media	0.2496 (0.0047)	0.8234 (0.1317)	0.0137 (0.0088)	0.0941 (0.0137)
	Distorsione	-.0004	0.0234	-.0463	0.0941
IND.INF	Media	0.2512 (0.0067)	0.8236 (0.1737)	0.0831 (0.0318)	0.1790 (0.2359)
	Distorsione	0.0012	0.0236	0.0231	0.1790
HS-GMM	Media	—	—	0.1212 (0.0914)	0.4124 (0.3919)
	Distorsione	—	—	0.0612	0.4124

Nella Tabella 4 sono presentati i risultati dell'esperimento Monte Carlo condotto sullo stesso modello per 100 replicazioni su serie osservate di lunghezza T=1000. Come si può osservare, all'aumentare della lunghezza delle serie, le stime dei HS-GMM presentano dei miglioramenti sostanziali soprattutto per il parametro di diffusione: i valori stimati di σ e γ continuano a presentare una evidente distorsione, ma il valore vero è interno all'intervallo di confidenza per i parametri. Anche le stime di Inferenza Indiretta presentano un miglioramento, mentre le stime di PML risultano ampiamente distorte.

Tab.4 Processo Ornstein-Uhlenbeck Num. Repl.=100, T=1000

Metodo di Stima	Parametri Valore vero	θ 0.25	k 0.8	σ 0.06	γ 0.0
PML	Media	0.2498 (0.0024)	0.8010 (0.0563)	0.0098 (0.0021)	0.0327 (0.0574)
	Distorsione	-.0002	0.0010	-.0502	0.0723
Ind.Inf	Media	0.2503 (0.0038)	0.8147 (0.0761)	0.0645 (0.0144)	0.0324 (0.1510)
	Distorsione	0.0003	0.0147	0.0045	0.0078
HS-GMM	Media	—	—	0.0715 (0.0119)	0.1233 (0.1107)
	Distorsione	—	—	0.0115	0.1233

E' importante sottolineare, inoltre, che durante gli esercizi Monte Carlo descritti nelle Tabelle 3 e 4 si sono verificati casi abbastanza frequenti, per tutti i metodi di stima proposti, nei quali la procedura di ottimizzazione non ha raggiunto la convergenza.

Per la verifica della consistenza degli stimatori proposti sono stati condotti esperimenti Monte Carlo per lunghezze delle serie $T=250, 500, 1000, 3000, 10.000, 150.000$. I risultati di tali esperimenti sono stati raccolti nelle Figure 1 e 2 che mostrano rispettivamente l'andamento della stima per σ e γ .

Come si può notare le stime di PML presentano asintoticamente, come atteso, un'ampia distorsione. Le stime di Inferenza Indiretta e i HS-GMM tendono alla convergenza verso i valori veri dei parametri, ma le stime dei HS-GMM mostrano una convergenza sensibilmente più lenta rispetto a quelle di Inferenza Indiretta.

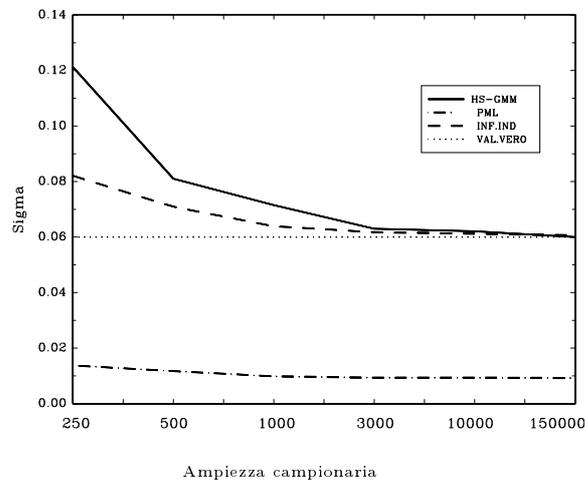


Fig.1 Processo Ornstein-Ulhenbeck: Stime di Sigma

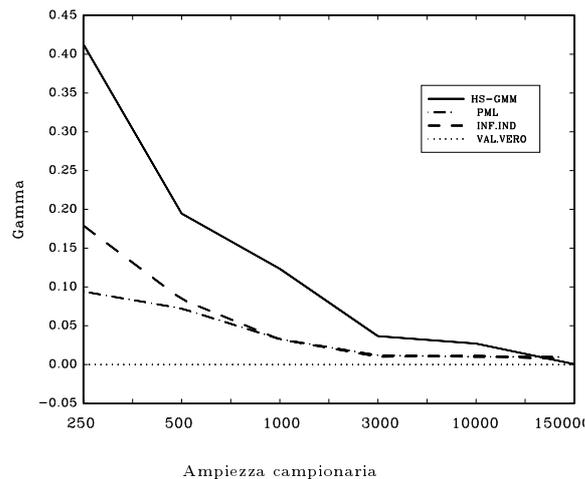


Fig.2 Processo Ornstein-Ulhenbeck: Stime di Gamma

Infatti specie per il parametro γ occorrono più di 150.000 osservazioni perché la con-

vergenza verso il valore vero diventi tangibile. Questo fa sì che l'uso dei HS-GMM non presenti nessun vantaggio significativo rispetto al metodo dell'Inferenza Indiretta, come invece si è verificato nel caso del modello vincolato.

Una spiegazione di tale fenomeno è da ricercarsi nel fatto che le condizioni (3.24) esprimono vincoli sui momenti della distribuzione marginale del processo soluzione $\{X_t\}$, che risulta essere una distribuzione di lungo periodo, pertanto si rende necessario un elevato numero di osservazioni perché la funzione obiettivo tenda a presentare un punto di minimo in un intorno del valore vero dei parametri.

3.6.2 Stima del processo “Radice Quadrata”

Nel Capitolo 2 è stato brevemente presentato il processo “Radice Quadrata”

$$dX_t = k(\theta - X_t)dt + \sigma\sqrt{X_t}d\mathcal{W}_t \quad (3.27)$$

come modello vincolato dalla EDS generale (3.5) per $\gamma = 0.5$. Nel paragrafo 2.2 è stata descritta la soluzione condizionata a $X_s = x$, con $s < t$, espressa dalla (2.4) che è stato visto possedere una distribuzione di tipo χ^2 non centrata con parametri indicati nelle (2.6). DeVroye (1986, pag 469) propone un algoritmo, che sarà assunto come modello simulatore dei dati, per la generazione di numeri pseudo casuali estratti da una variabile aleatoria χ^2 non centrata. Tale algoritmo si basa sull'osservazione che la legge χ^2 non centrata appartiene alla famiglia delle distribuzioni di Bessel. Pertanto la generazione di numeri pseudo casuali tratte da queste distribuzioni si basa sulla generazione di dati da una distribuzione Gamma i cui parametri sono dati da numeri pseudo casuali a loro volta estratti da una distribuzione di Poisson e trasformati in modo da rispettare i vincoli dati dalle (2.6).

La soluzione della (3.27) illustrata nel capitolo precedente è data da un χ^2 non centrato di media condizionale $\mathbf{E}[X_{t+1} | X_t] = \theta(1 - e^{-k}) + e^{-k}X_t$ e varianza condizionale

$$\sigma^2(X_{t+1} | X_t) = X_t \frac{\sigma^2}{k}(e^{-k} - e^{-2k}) + \theta \frac{\sigma^2}{2k}(1 - e^{-k})^2.$$

La funzione di Pseudo Verosimiglianza con momento primo e secondo correttamente specificati sarà quindi data dalla legge del processo:

$$X_{t+1} = \theta(1 - e^{-k}) + e^{-k}X_t + \sigma(X_{t+1} | X_t)\epsilon_t \quad (3.28)$$

dove $\epsilon_t \sim \mathcal{N}(0, 1)$.

Si consideri il problema della stima del modello (3.27). In tale contesto, analogamente a quanto visto per il processo di Ornstein-Uhlenbeck, il vettore dei parametri di interesse si riduce al solo σ e le condizioni dei HS-GMM (3.10) assumono la forma del tipo:

$$\mathbf{E}[f_i(x_t, \sigma, \hat{\nu}) | \hat{\nu}] = \hat{\theta}\mathbf{E}[x_t^i] + i \frac{\sigma^2}{2\hat{k}}\mathbf{E}[x_t^i] - \mathbf{E}[x_t^{i+1}] = 0 \quad \forall i \in \mathbb{N} \quad (3.29)$$

Ponendo $i = 1$ nella (3.29) si ottiene $var(X_t) = \frac{\theta\sigma^2}{2\hat{k}}$, che risulta essere la varianza della distribuzione marginale di $\{X_t\}$ di lungo periodo (Feller 1957) $X_t \sim \Gamma(\frac{2k\theta}{\sigma^2}, \frac{2k}{\sigma^2})$; pertanto per la stima HS-GMM di σ basta considerare la sola condizione di ortogonalità:

$$\mathbf{E}[f_1(x_t, \sigma, \hat{\nu}) | \hat{\nu}] = \hat{\theta}\mathbf{E}[x_t] + \frac{\sigma^2}{2\hat{k}}\mathbf{E}[x_t] - \mathbf{E}[x_t^2] = 0 \quad (3.30)$$

Le stime dei parametri di disturbo sono ottenute in accordo con quanto descritto nel paragrafo 3.5, come pure le stime di PML sono ottenute in base alla (3.28). Le stime di Inferenza Indiretta sono ottenute in base allo stimatore (2.16), dove il modello generatore dei dati è dato dalla discretizzazione di Eulero del processo (3.27):

$$X_{t+1} = k\theta + (1 - k)X_t\delta t + \sigma\sqrt{X_t\delta t}\epsilon_t$$

dove $\epsilon_t \sim \mathcal{N}(0, 1)$ e $\delta t = 0.1$ è il passo di simulazione in accordo con quanto proposto da Gourieroux, Monfort e Renault (1993); la stessa discretizzazione con passo di osservazione pari a uno è impiegata come modello ausiliario.

Nella Tabella 5 si riportano i risultati di un esperimento Monte Carlo di 200 replicazioni effettuato su serie osservate di lunghezza $T=250$, con $\theta = 0.25$, $k = 0.8$, $\sigma = 0.06$.

Tab.5: Processo Radice Quadrata Num. Repl.=200, T=250

Metodo di Stima	Parametri Valore vero	θ 0.1	k 0.8	σ 0.06
PML	Media	0.1001 (0.0014)	0.8286 (0.1471)	0.0647 (0.0061)
	Distorsione	0.0001	0.0086	0.0047
Ind.Inf	Media	0.1000 (0.0021)	0.8183 (0.1956)	0.0607 (0.0060)
	Distorsione	0.0000	0.0183	0.0007
HS-GMM	Media	—	—	0.0603 (0.0050)
	Distorsione			0.0003

Come si può notare gli HS-GMM si comportano in modo simile agli stimatori di Inferenza Indiretta e presentano un errore standard migliore.

Nella Tabella 6 si riportano i risultati di un esperimento Monte Carlo di 200 replicazioni su serie di lunghezza $T=1000$, allo scopo di presentare una verifica del comportamento dello stimatore proposto all'aumentare dell'ampiezza campionaria.

Tab.6: Processo Radice Quadrata Num. Repl.=200, T=1000

Metodo di Stima	Parametri Valore vero	θ 0.1	k 0.8	σ 0.06
PML	Media	0.1000 (0.0007)	0.8100 (0.0668)	0.0601 (0.0019)
	Distorsione	—	0.0100	0.0001
Ind.Inf	Media	0.0998 (0.0010)	0.8068 (0.0886)	0.0600 (0.0025)
	Distorsione	-.0002	0.0068	—
HS-GMM	Media	—	—	0.0602 (0.0030)
	Distorsione	—	—	0.0002

Procediamo ora alla stima del modello (3.5) non vincolato

$$dX_t = k(\theta - X_t)dt + \sigma X_t^\gamma dW_t.$$

In tale contesto le condizioni di ortogonalità presentano la forma data dalla (3.10) e lo stimatore HS-GMM è dato dalla (3.11). Anche in questo caso sono stati considerati i HS-GMM esattamente identificati, pertanto le stime saranno basate sulle condizioni (3.24) in modo analogo al processo di Ornstein-Uhlenbeck. Le stime di PML sono basate sulla (2.12) del capitolo precedente con $g(X_t, \beta) = X_t^\gamma$ e passo di osservazione $\Delta t = 1$; pertanto la funzione di Pseudo Verosimiglianza avrà una forma del tipo:

$$X_{t+1} = \theta(1 - e^{-k}) + e^{-k} X_t + e^{-k} \sigma X_t^\gamma \epsilon_t$$

dove $\epsilon_t \sim \mathcal{N}(0, 1)$.

Le stime di Inferenza Indiretta sono ottenute in base allo stimatore (2.16) dove il modello generatore dei dati è dato dalla discretizzazione di Eulero (3.25) sopra descritta, con passo di simulazione $\delta t = 0.1$ e passo di osservazione pari a 1, mentre il modello ausiliario è ricavato in accordo con quanto descritto nel paragrafo 3.6.2.

I risultati dell'esperimento Monte Carlo con 200 repliche su serie di lunghezza $T=250$, $\theta = 0.25$, $k = 0.8$, $\sigma = 0.06$, $\gamma = 0.5$ sono riportati nella Tabella 7. Anche in questo caso gli HS-GMM hanno un comportamento pessimo ed in particolare gli intervalli di confidenza per i parametri contengono lo zero. Gli stimatori di Inferenza Indiretta presentano una distorsione minore anche se gli errori standard sono tali per cui lo stimatore non risulta molto preciso.

Un esperimento Monte Carlo di 100 repliche condotto su serie di lunghezza $T=1000$, i cui risultati sono riportati nella Tabella 8, non ha evidenziato miglioramenti significativi dovuti all'aumentare dell'ampiezza campionaria sia per le stime di inferenza Indiretta, sia per gli HS-GMM.

Tab.7: Processo Radice Quadrata Num.Repl.=200, T=250

Metodo di Stima	Parametri Valore vero	θ 0.25	k 0.8	σ 0.06	γ 0.5
PML	Media	0.2501 (0.0025)	0.8456 (0.1251)	0.0159 (0.0332)	0.4055 (0.4006)
	Distorsione	0.0001	0.0456	-.0441	-.0945
Ind.Inf	Media	0.2502 (0.0030)	0.8284 (0.1566)	0.0783 (0.0407)	0.6027 (0.3056)
	Distorsione	0.0002	0.0284	0.0183	0.1027
HS-GMM	Media	—	—	0.3085 (0.4030)	0.8010 (0.9800)
	Distorsione			0.2485	0.3010

Tab.8: Processo Radice Quadrata Num. Repl.=100, T=1000

Metodo di Stima	Parametri Valore vero	θ 0.25	k 0.8	σ 0.06	γ 0.5
PML	Media	0.2499 (0.0011)	0.8133 (0.0612)	0.0067 (0.0050)	0.3093 (0.2446)
	Distorsione	-.0001	0.0133	0.0007	-.1907
Ind.Inf	Media	0.2502 (0.0017)	0.8130 (0.0738)	0.0721 (0.0376)	0.5415 (0.3599)
	Distorsione	.0002	0.0133	0.0121	0.0415
HS-GMM	Media	—	—	0.0994 (0.1031)	0.6400 (0.4875)
	Distorsione			0.0394	0.1400

L'origine del cattivo comportamento dello stimatore può essere investigato tramite uno studio del comportamento in fase di stima della funzione obiettivo del problema di minimo (3.11) di cui i parametri stimati sono soluzione. Per agevolare tale studio la funzione obiettivo è stata modificata in modo che dipenda da un solo parametro. Nella fase di stima, nella (3.24) il valore atteso viene sostituito, come di consueto, dalla media campionaria della serie oggetto di studio, e quindi, poiché si opera con un sistema di condizioni GMM esattamente identificato, dalla prima equazione delle (3.24) è possibile ricavare il parametro σ e sostituirlo nella seconda condizione di ortogonalità considerata, ottenendo quindi una funzione del solo parametro $\gamma \geq 0$.

Nelle Figure 3,4,5 e 6 è riportato l'andamento della funzione obiettivo al variare di γ calcolata su una serie osservata di lunghezza $T=250, 1000, 10000, 150000$. Si può notare come all'aumentare della lunghezza della serie osservata la funzione obiettivo tenda a presentare un minimo assoluto per $\gamma = 0.5$ come atteso. Si osserva, però, che affinché si raggiunga una forma accettabile di tale funzione occorrono più di 150000 dati e che nell'intorno del punto di minimo essa si presenta abbastanza piatta da produrre una certa variabilità delle stime. Questa situazione fa sì che gli HS-GMM non presentino nel caso del modello non vincolato nessun vantaggio significativo per la loro applicazione rispetto al metodo dell'Inferenza Indiretta contrariamente a quanto verificatosi nel caso del modello vincolato.

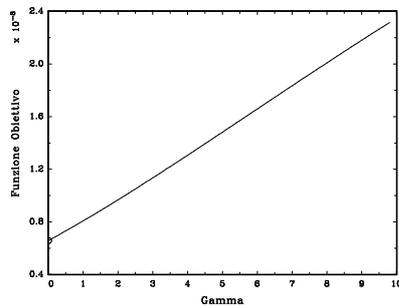


Fig.3 T=250

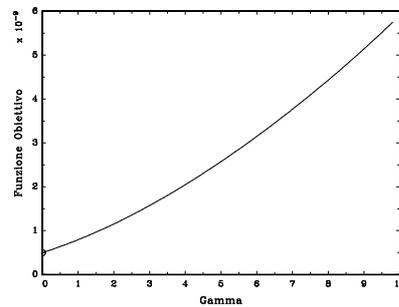


Fig.4 T=1000

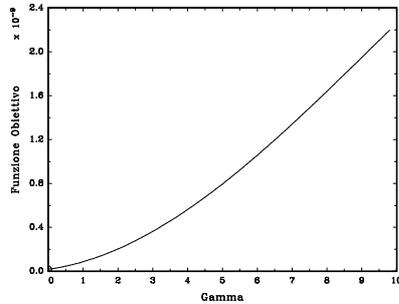


Fig.5 T=10000

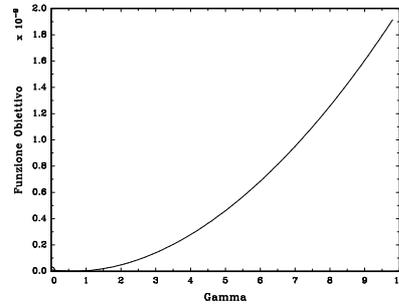


Fig.6 T=150000

Tale situazione è determinata in primo luogo dall'elevato grado di correlazione tra le condizioni (3.24) che tende a rendere piatta la funzione obiettivo, e dal fatto che sia le condizioni (3.6), sia le (3.7), si basano sulla distribuzione marginale del processo $\{X_t\}$ che è una distribuzione di lungo periodo. Pertanto per numerosità poco elevate delle serie la funzione tende ad essere abbastanza piatta da causare una variabilità molto elevata nelle stime. Durante gli esercizi Monte Carlo, sempre per ampiezze campionarie limitate, si sono verificati casi abbastanza frequenti nei quali la procedura di

ottimizzazione non ha raggiunto la convergenza. Quindi, in condizioni di modello non vincolato è necessario un numero di osservazioni elevato per poter ottenere una funzione obiettivo che permetta di discriminare con un certo grado di affidabilità il parametro di elasticità della varianza.

Analogamente al processo di Ornstein-Uhlenbeck sono stati condotti esperimenti Monte Carlo per lunghezze delle serie $T=250, 500, 1000, 3000, 10.000, 150.000$ per la verifica della consistenza degli stimatori proposti. I risultati di tali esperimenti sono stati raccolti nelle Figure 7 e 8 che mostrano rispettivamente l'andamento della stima di σ e γ all'aumentare della lunghezza delle serie.

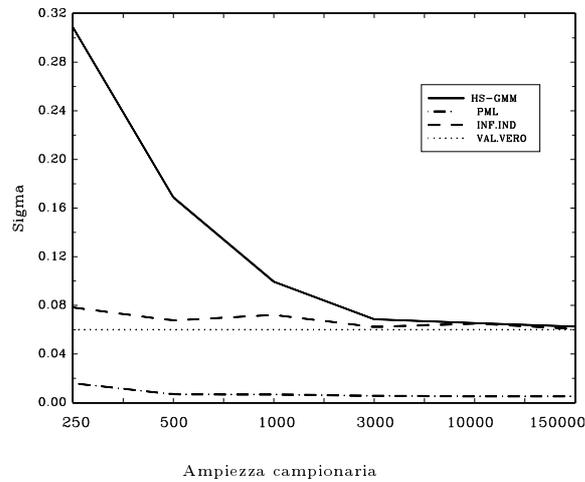


Fig.7 Processo Radice Quadrata: Stime di Sigma

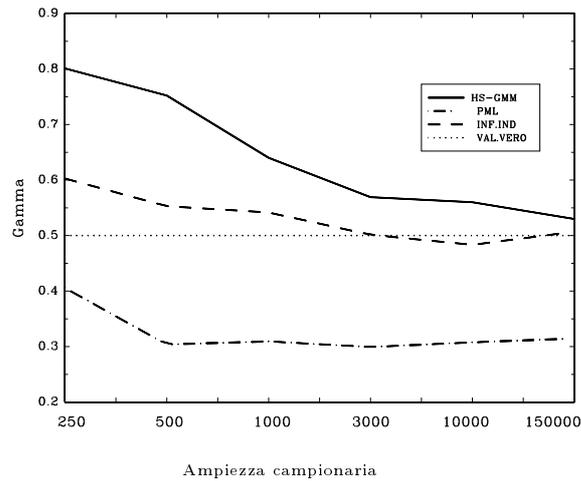


Fig.8 Processo Radice Quadrata: Stime di Gamma

Come si può notare entrambe le stime tendono alla convergenza verso i valori veri dei parametri, mentre le stime di PML presentano, come atteso, un'ampia distorsione.

3.7 Spunti per ulteriori ricerche

Nella teoria dei GMM l'impiego di un numero di condizioni di ortogonalità superiore a quello dei parametri, si parla quindi di GMM sovraidentificati, trova la sua giustificazione nel fatto che ogni condizione pone un vincolo che deve essere rispettato dai dati: più sono le informazioni di cui si dispone, maggiori sono i vincoli che si possono imporre, migliore sarà la qualità delle stime che si possono ottenere. Tale situazione ha un limite naturale nel fatto che non tutte le condizioni di ortogonalità che si possono imporre sono apportatrici di informazioni rilevanti per la stima. Nel caso dei GMM sovraidentificati le stime non sono più indipendenti dalla matrice dei pesi adottata; inoltre per poter utilizzare una stima consistente della matrice dei pesi ottimale si rende necessaria una procedura di ottimizzazione a due stadi.⁵

Sia per il processo di Ornstein-Uhlenbeck, sia per il processo Radice Quadrata sono stati condotti esercizi Monte Carlo basati sui HS-GMM sovraidentificati per verificare se l'introduzione di ulteriori condizioni di ortogonalità potesse fornire un miglioramento della qualità delle stime.

Mediante la matrice di Newey-West, descritta nel paragrafo 3.5, è stato riscontrato un forte grado di correlazione tra le condizioni (3.10) per $i = 1, \dots, 10$. Tale situazione rende estremamente instabile il calcolo dell'inversa della matrice di covarianza tra le condizioni $\Omega(f_t)$, e della matrice Ω_* necessaria per il calcolo dello stimatore (3.11) con matrice dei pesi ottimale. Per lo stesso motivo il calcolo della matrice di covarianza delle stime HS-GMM risulta problematico.

Come si può osservare, le condizioni (3.10) per $i = 1, 2, \dots$ altro non sono che la stessa condizione moltiplicata ogni volta per x_t . Inoltre al crescere di i la variabile x_t^i diventa sempre più piccola, poiché i processi generano valori inferiori all'unità, aggravando i problemi di instabilità numerica. Questo fa sì che ogni condizione aggiuntiva non apporti nessuna informazione significativamente rilevante, pur determinando la successione dei momenti della distribuzione marginale $\{X_t\}$. Inoltre uno studio grafico della funzione obiettivo di secondo stadio del problema (3.11) ha rilevato che essa si presenta molto piatta in una vasta area intorno al punto di minimo, e che tale situazione non migliora in maniera sensibile al crescere della lunghezza delle serie osservate, come evidenziato dalle Figure 9 e 10 che si riferiscono al processo di Ornstein-Uhlenbeck, e dalle Figure 11 e 12 che si riferiscono al processo "Radice Quadrata".

⁵Una breve descrizione dei GMM sovraidentificati è presente nell'Appendice C

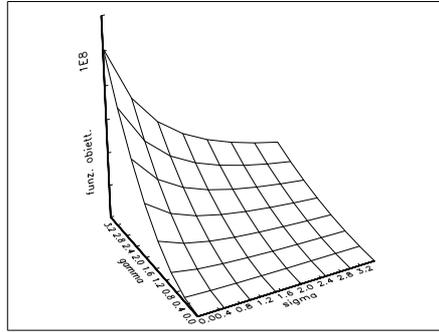


Fig.9 T=250

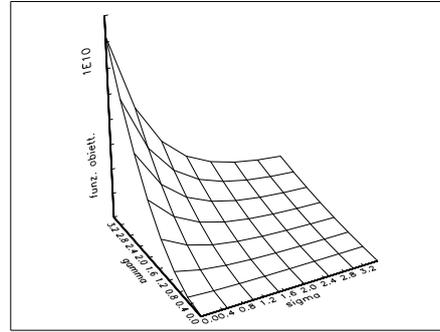


Fig.10 T=10000

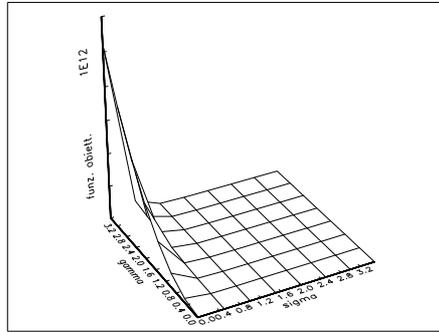


Fig.11 T=250

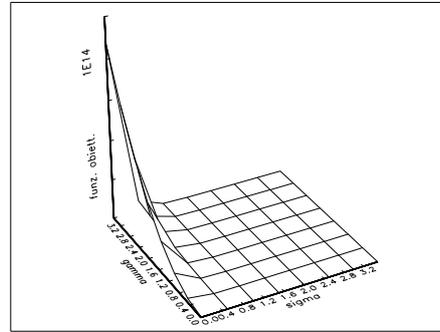


Fig.12 T=10000

In base a tali considerazioni si è quindi ritenuto opportuno non riportare gli esercizi Monte Carlo per gli HS-GMM sovraidentificati. Una giustificazione teorica a questa situazione può essere ricercata nelle relazioni che legano i momenti del processo $\{X_t\}$ soluzione della EDS (3.5). In particolare Arnold (1975, cap 7), mostra che i momenti di ordine pari della soluzione $\{X_t\}$ sono combinazione lineare dei momenti di ordine pari inferiori. Questa situazione porterebbe a riconsiderare la scelta della funzione ϕ da applicare alle condizioni **C1** in modo che essa non conduca ad una combinazione lineare dei momenti della distribuzione di $\{X_t\}$, o che porti ad una stabilizzazione della correlazione tra le condizioni di ortogonalità.

Capitolo 4

Applicazione dei HS-GMM al mercato statunitense

4.1 Introduzione

In questo capitolo sarà presentata un'applicazione dei HS-GMM ai mercati finanziari statunitensi. A seguito dei risultati presentati nel Capitolo precedente, non verrà presa in considerazione la stima del modello non vincolato (3.5) limitando quindi l'attenzione ai modelli di Vasicek (1977) e Cox, Ingersoll e Ross (1985).

La stima dei parametri θ , k e σ delle EDS che caratterizzano i due modelli sarà condotta secondo i metodi della PML, dell'Inferenza Indiretta e dei HS-GMM. Sarà determinato il prezzo di un titolo a rendimento fisso senza cedole (*zero-coupon bond*), e quello di un'Opzione di vendita secondo le relazioni individuate dai modelli considerati e descritte nel Capitolo 1.

4.2 I dati

I dati a disposizione si riferiscono ai tassi di interesse *overnight* per i fondi federali degli Stati Uniti per il periodo 2 Gennaio 1973 - 31 Dicembre 1993 e rappresentano il tasso d'interesse calcolato giornalmente dalla Federal Reserve Bank di New York. Poiché si tratta di osservazioni giornaliere per tassi d'interesse su depositi interbancari di brevissima durata, è auspicabile che essi costituiscano una "proxy" assai efficace dei tassi di interesse istantanei descritti nel Capitolo 1.

I fondi federali sono riserve presso la Federal Reserve Bank trasferibili tra istituzioni di deposito. Le transazioni di fondi federali permettono a tali istituzioni di ottenere guadagni positivi su fondi in eccesso sulle riserve richieste che altrimenti resterebbero infruttuosi. Tale tasso riflette l'entità dei prestiti interbancari allo scoperto di brevissimo periodo, che durano cioè meno di una notte. La serie è costruita mediando i tassi tra tutti gli operatori con pesi proporzionali alla grandezza dei prestiti. Poiché i tassi di interesse pagati sulle altre attività finanziarie di breve periodo "Treasure-Bills" si muovono in stretta relazione con il tasso sui fondi federali, appare naturale fondare la determinazione del prezzo (*pricing*) di titoli a rendimento fisso senza cedole e quello di altre attività finanziarie derivate come le opzioni di vendita o di acquisto di tipo

europeo sulla dinamica di tale tasso.

Nella Figura 13 è riportata la serie storica dei tassi “overnight” annualizzati, cioè trasformati come ordine di grandezza in modo che l’unità temporale di riferimento sia l’anno. Nella Tabella 9 sono riportate le principali statistiche descrittive relative alla serie storica analizzata.

Attraverso il test di Dickey-Fuller aumentato è stata sottoposta a verifica l’ipotesi nulla di non stazionarietà della serie. Il test di Dickey-Fuller aumentato è stato calcolato sul modello:

$$\Delta X_t = \mu + \alpha X_{t-1} + \sum_{j=1}^p \alpha_j \Delta X_{t-1} + u_t$$

dove X_t indica il tasso di interesse “overnight”, $p = 30$ è l’ordine di autocorrelazione, e u_t sono errori che seguono un processo AR(1) stazionario¹. L’ipotesi di non stazionarietà viene rifiutata al livello del 90%; tuttavia è bene ricordare che a causa della scarsa potenza di questo test il rifiuto dell’ipotesi nulla lascia spazio a molti dubbi sulla stazionarietà della serie, che è quindi da ritenersi solo probabile.

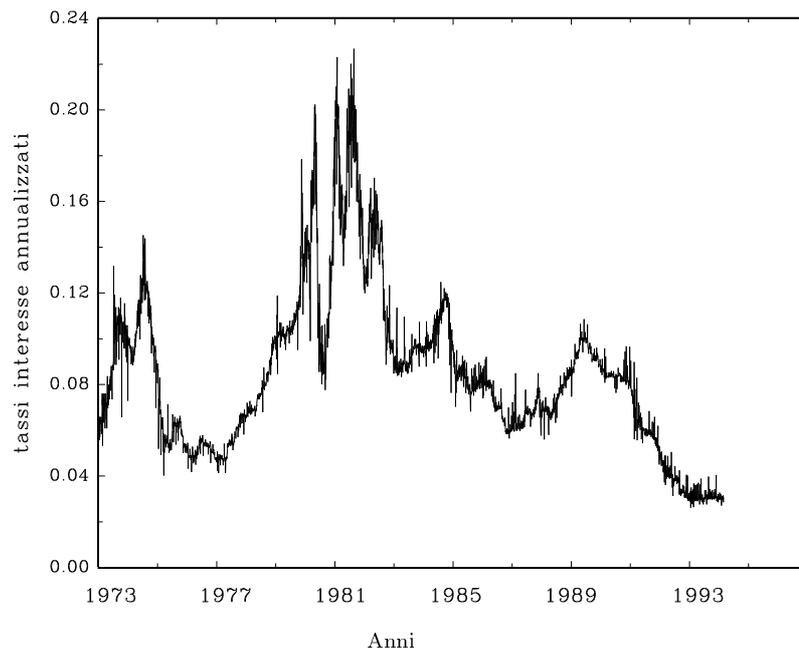


Fig.13: Tassi di Interesse annualizzati

¹Per maggiori dettagli si rimanda il lettore interessato al testo di Davidson e MacKinnon (1993), cap.20 pag 711.

Tab.9: Statistiche descrittive

	Tassi Interesse spot
Num. osser.	5289 8 osservaz. outliers cancellate
Media	0.08445
Dev. Stand	0.03502
Curtosi	4.255
Augmented Dickey-Fuller	-2.58
H_0 : Non stazion.	Rifutata al 90% (val. critico= -2.57)

4.3 Le stime dei parametri

Come illustrato nei capitoli precedenti il modello di Vasicek (1977) descrive la dinamica del tasso di interesse a brevissimo periodo tramite la EDS (3.19). La stima dei parametri θ , k e σ^2 è stata condotta con i metodi della PML, dell'Inferenza Indiretta e dei HS-GMM secondo le procedure illustrate nel paragrafo 3.6.2, e i risultati sono riportati nella Tabella 10.

Tab.10: Stima dei Parametri del Modello di Vasicek

Metodo di Stima	θ	k	σ^2
PML	0.083585 (0.009804)	0.008071 (0.001376)	2.0228×10^{-5} (1.4984×10^{-7})
Ind.Inf	0.083200 (0.002996)	0.007398 (0.004215)	2.0259×10^{-5} (1.0566×10^{-6})
HS-GMM	—	—	2.0747×10^{-5} (1.0172×10^{-6})

Le stime riportate nella Tabella 10 sono riferite alla serie dei tassi giornalieri annualizzati. Come si può osservare i risultati sono simili per tutti i metodi di stima impiegati, in accordo con quanto verificato nel Capitolo precedente attraverso gli esercizi Monte Carlo.

La varianza delle stime di Inferenza Indiretta è stata calcolata utilizzando l'espressione individuata da Gourieroux, Monfort e Renault (1993) nella forma:

$$Var_{asy}(\sqrt{n}(\hat{\alpha} - \alpha_0)) = \left[\frac{\partial \beta(\alpha_0)}{\partial \alpha} J_0 I_0^{-1} J_0 \frac{\partial \beta(\alpha_0)}{\partial \alpha'} \right]^{-1}$$

dove $\alpha = (\theta, k, \sigma^2)$ è il vettore dei parametri di interesse, α_0 indica il valore vero dei parametri, $\hat{\alpha}$ il valore stimato sulla base della (2.16), e

$$J_0 = -plim \frac{\partial^2 \lambda_{\star}}{\partial \beta \partial \beta'}(\underline{y}_t, \beta_0), \quad I_0 = Var_{asy} \left(\sqrt{n} \frac{\partial \lambda_{\star}}{\partial \beta}(\underline{y}_t, \beta_0) \right),$$

dove β_0 indica il vettore dei parametri ausiliari, calcolati sui dati osservati \underline{y}_t , che corrispondono al valore vero α_0 dei parametri di interesse. La varianza delle stime dei HS-GMM è stata calcolata sulla base dell'espressione (3.15) proposta nel Capitolo 3. Come si può osservare dagli errori standard riportati nella Tabella 10, tutte le stime risultano statisticamente significative ed in particolare gli errori standard delle stime del parametro σ^2 secondo l'Inferenza Indiretta e gli HS-GMM risultano dello stesso ordine di grandezza.

Nel modello di Cox Ingersoll e Ross (1985) la dinamica del tasso di interesse a brevissimo periodo è descritta tramite la EDS (3.25). La stima dei parametri θ, k e σ^2 è stata condotta con i metodi della PML, dell'Inferenza Indiretta e dei HS-GMM secondo le procedure illustrate nel paragrafo 3.6.3, e i risultati sono riportati nella Tabella 11.

Tab.11: Stima dei Parametri del Modello CIR

Metodo di Stima	θ	k	σ^2
PML	0.083200 (0.009790)	0.005784 (0.001387)	2.0341×10^{-6} (3.9724×10^{-8})
Ind.Inf	0.085410 (0.001208)	0.005361 (0.001946)	2.0389×10^{-6} (1.0305×10^{-7})
HS-GMM	—	—	1.7932×10^{-6} (1.2364×10^{-7})

Le stime riportate nella Tabella 11, analogamente a quanto fatto per la Tabella 10, sono riferite alla serie dei tassi giornalieri annualizzati. Anche in questo caso i risultati sono simili per tutti i metodi di stima impiegati. Le varianze delle stime sono state

calcolate in accordo con quanto precedentemente descritto, e tutte le stime risultano statisticamente significative. Si noti come anche in questo caso gli errori standard delle stime del parametro σ^2 secondo l’Inferenza Indiretta e gli HS-GMM risultano dello stesso ordine di grandezza.

4.4 La valutazione dei titoli

Una volta determinato il valore dei parametri di può quindi procedere alla valutazione di un titolo a rendimento fisso senza cedole, secondo i modelli di Vasicek e CIR. Il calcolo di tale prezzo è basato sulla formula (1.5) e sulla formula (1.8). I prezzi sono stati calcolati per assegnati valori del tasso istantaneo annualizzato e della maturità espressa in anni.

Il premio di rischio λ per i dati presi in esame è stato calcolato da Ait-Sahalia (1994) utilizzando una regressione spline sulla curva dei rendimenti dei titoli “zero coupon” implicata dai titoli con scadenza alla fine di ogni mese tratta da *The Wall Street Journal*.

Le stime sono $\lambda_{Vas} = -0.0295$ e $\lambda_{CIR} = -0.00141$.

I prezzi corrispondono al valore nominale del titolo pari a 100 lire. Nelle Tabelle 12, 13 e 14 sono riportati i prezzi calcolati con le stime dei parametri θ, k, σ ottenute con la stima di PML, dei HS-GMM e di Inferenza Indiretta, rispettivamente. Dei due elementi riportati in ogni casella quello in alto è il prezzo secondo il modello Vasicek e quello in basso è il prezzo secondo il CIR.

Tab.12: Prezzi “zero-coupon” con stime PML dei parametri

Tasso istantaneo annualizzato → ***** Maturità ↓	0.02	0.04	0.06	0.08	0.10	0.12	0.14
0.5	98.6653 99.0000	97.6855 98.0160	96.7155 97.0418	95.7551 96.0773	94.8042 95.1224	93.8628 94.1769	92.9307 93.2408
1	96.6832 98.0005	94.7764 96.0641	92.9072 94.1660	91.0748 92.3055	89.2786 90.4816	87.5178 88.6939	85.7917 86.9414
5	64.4094 90.0409	58.3961 81.5610	52.9443 73.8796	48.0015 66.9217	43.5201 60.6191	39.4571 54.9100	35.7734 49.7387

Tab.13: Prezzi “zero-coupon” con stime HS-GMM dei parametri

Tasso istantaneo annualizzato → ***** Maturità ↓	0.02	0.04	0.06	0.08	0.10	0.12	0.14
0.5	98.6663 99.0000	97.6865 98.0160	96.7165 97.0418	95.7561 96.0773	94.8052 95.1224	93.8637 94.1769	92.9317 93.2408
1	96.6873 98.0005	94.7804 96.0641	92.9111 94.1660	91.0786 92.3055	89.2823 90.4816	87.5215 88.6939	85.7953 86.9414
5	64.4755 90.0409	58.4561 81.5610	52.9987 73.8796	48.0508 66.9217	43.5648 60.6191	39.4976 54.9100	35.8101 49.7386

Tab.14: Prezzi “zero-coupon” con stime Inferenza Indiretta dei parametri

Tasso istantaneo annualizzato → ***** Maturità ↓	0.02	0.04	0.06	0.08	0.10	0.12	0.14
0.5	98.6688 99.0002	97.6888 98.0161	96.7186 97.0418	95.7580 96.0772	94.8069 95.1221	93.8653 94.1766	92.9331 93.2404
1	96.6969 98.0013	94.7892 96.0645	92.9191 94.1660	91.0859 92.3051	89.2888 90.4809	87.5272 88.6927	85.8004 86.9399
5	64.6155 90.0591	58.5734 81.5689	53.0963 73.8791	48.1314 66.9142	43.6307 60.6060	39.5509 54.8924	35.8526 49.7175

Come si può notare, le valutazioni ottenute non differiscono in maniera apprezzabile al variare dei metodi di stima dei parametri e questo permette di affermare in particolare che la differenza riscontrata nella stima del parametro σ^2 per il modello CIR tra i metodi di Inferenza Indiretta e i HS-GMM non è significativa ai fini dell'esercizio di valutazione dei titoli.

A parità di tasso di interesse e all'aumentare della maturità i prezzi diminuiscono, come atteso. Si nota in particolare che le valutazioni ottenute sono abbastanza elevate per bassi tassi di interesse, soprattutto per il modello CIR, per una maturità di 5 anni, e viceversa sono piuttosto basse per alti tassi di interesse per il modello di Vasicek sempre per un maturità di 5 anni. Parte di questo effetto è dovuto al valore calcolato del premio di rischio.

Una volta determinato il valore del titolo a rendimento fisso si può calcolare quello di un'opzione di acquisto di tipo europeo su tale titolo.

Come descritto nel Capitolo 1 un'opzione di acquisto è un contratto che dá il diritto di comprare un determinato titolo ad un certo istante (scadenza) oppure entro un certo periodo di tempo, ad un prezzo fissato (prezzo di esercizio). Quando può essere esercitata solo alla scadenza l'opzione viene detta di *tipo europeo*, mentre se essa può essere esercitata entro la data di scadenza viene detta di *tipo americano*. Le opzioni, inoltre, sono indicate come *nella moneta*, *fuori la moneta* e *alla moneta* se il prezzo del titolo di riferimento è rispettivamente maggiore, minore, o uguale al prezzo di esercizio K dell'opzione. Chiaramente un'opzione sarà esercitata solo se è "nella moneta". Pertanto si è espresso il prezzo di esercizio K come proporzione del corrispondente prezzo del bond; quindi per prezzi di esercizio inferiori a 1 si ha l'opzione nella moneta, per prezzi di esercizio superiori ad 1 si ha l'opzione fuori dalla moneta per il prezzi di esercizio uguale ad 1 si ha l'opzione alla moneta.

Nella Tabella 15 sono riportati prezzi di un'opzione di vendita di tipo europeo su un titolo a rendimento fisso con maturità a 5 anni, calcolati in base alla formula (1.6), descritta nel paragrafo 1.3, e derivata secondo le ipotesi del modello di Vasicek.

Tab.15: Prezzi Opzione Vasicek su "zero-cupon" a 5 anni con stime HS-GMM dei parametri

Tasso ist. annualiz	Matur. Opzione	Prezzo di esercizio				
		0.96	0.98	1	1.02	1.04
0.04	0.25	3.1111	2.2559	1.2392	0.9346	0.4876
0.04	0.5	3.6889	2.7965	1.7547	1.0461	0.5460
0.04	1	5.2737	4.1872	2.9672	2.2005	1.4045
0.14	0.25	3.1080	2.4000	1.4065	0.9952	0.5116
0.14	0.5	3.8745	3.2207	2.5437	1.6805	1.3563
0.14	1	5.9808	5.7062	5.0886	4.4771	3.8624

La maturità dell'opzione è espressa in anni. Tutti i prezzi si riferiscono ad un titolo

di riferimento che paga 100 lire alla scadenza. A parità di tasso di interesse, i prezzi mostrano il comportamento atteso al variare del prezzo di esercizio e della maturità dell'opzione.

Nella Tabella 16 sono riportati prezzi di un'opzione di vendita di tipo europeo su un titolo a rendimento fisso con maturità a 5 anni, calcolati in base alla formula (1.9), descritta nel paragrafo 1.4, e derivata secondo le ipotesi del modello CIR.

La legge di probabilità χ^2 non centrata presente nelle formula è stata generata in base all'algoritmo suggerito da Devroye (1986) e ricordato nel paragrafo 3.6. Tutti i prezzi si riferiscono ad un titolo di riferimento che paga 100 lire alla scadenza, e analogamente all'esercizio precedente la maturità dell'opzione è espressa in anni.

Tab.16: Prezzi Opzione CIR su "zero-cupon" a 5 anni con stime HS-GMM dei parametri

Tasso ist. annualiz	Matur. Opzione	Prezzo di esercizio				
		0.96	0.98	1	1.02	1.04
0.04	0.25	6.4865	5.0140	3.7991	2.8109	2.0213
0.04	0.5	6.5250	5.0495	3.8316	2.8411	2.0496
0.04	1	6.6340	5.1481	3.9207	2.9218	2.1207
0.14	0.25	7.3281	6.7913	6.2885	5.8178	5.3772
0.14	0.5	7.4176	6.8768	6.3702	5.8958	5.4517
0.14	1	7.4453	6.9036	6.3962	5.9210	5.4761

Anche in questo caso, a parità di tasso di interesse, i prezzi mostrano il comportamento atteso al variare del prezzo di esercizio e della maturità dell'opzione.

4.5 Conclusioni

Il problema della stima da dati discreti dei parametri delle EDS è strettamente legato alla difficoltà di determinare la probabilità di transizione del processo di Markov che ne è soluzione. Tale difficoltà rende spesso impraticabili i metodi di stima fondati sulla funzione di Verosimiglianza e costringe gli statistici ad utilizzare modelli approssimati. L'uso di tali modelli conduce quindi ad una errata specificazione che ha tra le sue più vistose conseguenze la inconsistenza delle stime ottenute.

In questo lavoro è stato proposto uno stimatore di tipo GMM per i parametri di una EDS derivato dalle condizioni di ortogonalità per i processi di Markov in tempo

continuo recentemente introdotte da Hansen e Scheinkman (1995) (HS-GMM). Tale stimatore presenta il vantaggio, di non richiedere la conoscenza del processo soluzione della EDS oggetto di studio. Inoltre, essendo fondati sulle sole proprietà dell'operatore di Dynkin, gli HS-GMM sono derivati in un contesto logico nel quale non è presente errata specificazione del modello, pertanto essi forniscono stime consistenti dei parametri di interesse.

Tramite esercizi Monte Carlo, inoltre, sono state condotte verifiche sulle proprietà econometriche dei HS-GMM sia nei piccoli che nei grandi campioni, proponendo in particolare un confronto con gli stimatori di Inferenza Indiretta. La scelta dell'Inferenza Indiretta come metro di paragone è stata motivata dal successo che tale metodo sta raccogliendo in letteratura per la stima dei parametri delle EDS, grazie alla sua proprietà di eliminare la distorsione delle stime dovuta all'uso di modelli approssimati. Il metodo di Inferenza Indiretta presenta, però, alcuni problemi di applicabilità dovuti alla sua elevata complessità di elaborazione. In particolare, oltre alla difficoltà di utilizzare con successo nell'ottimizzazione della funzione obiettivo algoritmi ricorsivi che richiedano il calcolo della matrice Jacobiana, il metodo dell'Inferenza Indiretta presenta una certa sensibilità al criterio di ottimizzazione scelto per la stima del modello ausiliario. Infatti risultati empirici hanno verificato che un'opportuna scelta di tale criterio incide sensibilmente sulle possibilità di convergenza della procedura di ottimizzazione e sulla qualità delle stime ottenibili.

Gli HS-GMM si propongono quindi come valido metodo di stima alternativo per la loro agevole calcolabilità.

A causa della identificazione a meno di un fattore di scala delle condizioni di ortogonalità sui cui si basa lo stimatore HS-GMM proposto, l'attenzione è stata concentrata sui soli parametri del coefficiente di diffusione delle EDS, ed in particolare sono state prese in considerazione l'equazione di Ornstein-Uhlenbeck e l'equazione "Radice Quadrata" impiegate nei modelli di Vasicek (1977) e di Cox, Ingersoll e Ross (1985). Dagli esercizi Monte Carlo è emerso che nel caso delle EDS di Ornstein-Uhlenbeck e "Radice Quadrata" lo stimatore HS-GMM ha un comportamento paragonabile agli stimatori di Inferenza Indiretta, e rispetto a questi presenta una migliore precisione. Buoni risultati sulla validità applicativa dello stimatore HS-GMM sono stati ottenuti anche dall'applicazione dei modelli di Vasicek e CIR al mercato statunitense per la valutazione di opzioni di vendita di tipo europeo. Le stime ottenute per i parametri del modello di Vasicek tramite il metodo dell'Inferenza Indiretta e i HS-GMM sono simili per tutti i metodi di stima considerati ed in particolare corrispondono a quelle ottenute sugli stessi dati da Ait-Sahalia (1994). Anche per il modello CIR le stime ottenute sono simili per tutti i metodi considerati ma differiscono, specie per il parametro di attrazione verso la media k , da quelle ottenute da Ait-Sahalia (1994), il quale ha impiegato per tale stima lo stimatore GMM suggerito da Chan e alt.(1992).

E' stata presa in considerazione anche la EDS con elasticità della varianza costante, indicata nel corso di questo lavoro come modello non vincolato. L'interesse in questa equazione deriva dal fatto che sia il processo di Ornstein-Uhlenbeck, che il processo "Radice Quadrata" possono essere considerati come suoi casi particolari, fissato il valore dell'elasticità della varianza. Pertanto la restrizione sul valore di tale elasticità può essere sottoposta a verifica di ipotesi con l'obiettivo di discriminare quale sia il

processo che ha condizionato il verificarsi dei dati osservati e quindi fornire un criterio in base al quale poter scegliere tra i modelli proposti per la valutazione delle attività finanziarie quello che meglio si adatta ai dati. In letteratura non sono stati trovati esempi di applicazione del metodo di Inferenza Indiretta per tale modello non vincolato.

Gli esercizi Monte Carlo, nel caso del modello non vincolato hanno evidenziato alcuni importati risultati. I HS-GMM hanno mostrato nei piccoli campioni un comportamento gravemente insoddisfacente, e proprietà asintotiche piuttosto deboli. Le stime di Inferenza Indiretta hanno mostrato un andamento nei piccoli campioni significativamente migliore dei HS-GMM, anche se non completamente soddisfacente, mentre sono stati ottenuti buoni risultati in termini di consistenza. E' importante sottolineare, però, che nel corso degli esercizi Monte Carlo per il modello non vincolato con ampiezze campionarie non elevate si sono verificati casi abbastanza frequenti nei quali le procedure di ottimizzazione di entrambi i metodi di stima non hanno raggiunto la convergenza.

L'origine del cattivo funzionamento dei HS-GMM è da ricercarsi nell'elevato grado di autocorrelazione tra le condizioni di ortogonalità che rende la funzione obiettivo piuttosto piatta, e nel fatto che queste si basano sulla legge di probabilità marginale del processo soluzione che è una legge di lungo periodo. Per le stesse ragioni i HS-GMM sovraidentificati non hanno fornito risultati degni di rilievo.

Quanto illustrato suggerisce lo schiudersi di ulteriori filoni di ricerca. Nuove indagini possono essere volte ad individuare condizioni di ortogonalità, derivate da quelle generali indicate da Hansen e Scheinkman (1995), alternative a quelle sviluppate in questo lavoro che tendano a risolvere in particolare il problema dell'autocorrelazione. E' interessante, inoltre, investigare le possibilità di un ulteriore sviluppo metodologico e computazionale della procedura di ottimizzazione del metodo di Inferenza Indiretta per l'EDS con elasticità della varianza non vincolata al fine di pervenire ad un miglioramento della qualità delle stime nei piccoli campioni.

I risultati ottenuti lasciano aperti, infine, temi di ricerca sulla possibilità di impiego dei HS-GMM per EDS con specificazioni alternative a quelle delle equazioni di Ornstein-Uhlenbek e "Radice Quadrata" del coefficiente di diffusione per l'estensione al caso della volatilità non costante, in particolare alla diffusione dipendente esplicitamente dal tempo. La principale limitazione all'uso di queste EDS in finanza è legata alla difficoltà della risoluzione della equazione alle derivate parziali che fornisce la formula di valutazione del prezzo delle attività derivate.

Appendice A

L'integrale stocastico di Ito

Le equazioni differenziali stocastiche svolgono un ruolo centrale in molte applicazioni della matematica alle scienze naturali e all'ingegneria. Tali equazioni contengono elementi che caratterizzano le strutture fisiche dei fenomeni e dell'ambiente circostante che esse descrivono e sono sperimentalmente determinate. Tuttavia questi fenomeni sono spesso perturbati da eventi casuali e quindi la loro descrizione tramite le equazioni differenziali ordinarie appare poco realistica.

Le Equazioni Differenziali Stocastiche (EDS) sono state sviluppate quindi con l'intento di fornire uno strumento matematico che permetta una più realistica formulazione delle equazioni per le scienze applicate.

In modo intuitivo le EDS possono essere definite come equazioni differenziali che contengono elementi casuali. Tali elementi possono essere coefficienti o qualche processo stocastico. Di conseguenza, la soluzione di una EDS sarà un processo stocastico ed il problema sarà quello di determinare le sue proprietà probabilistiche.

Si consideri quindi l'equazione differenziale ordinaria:

$$dx = \mu(x, t)dt \quad (\text{A.1})$$

di punto iniziale x_0 .

La (A.1) può essere considerata come una EDS degenera perchè priva del termine aleatorio. La (A.1) è equivalente all'equazione integrale:

$$x(t) = x_0 + \int_{t_0}^t \mu(x_s, s) ds$$

sotto condizioni di regolarità per la funzione μ , che vengono utilizzate per garantire l'esistenza e l'unicità della soluzione.

Si consideri ora il problema empirico:

$$dX_t = \mu(X_t, t)dt + \sigma(X_t, t)\epsilon_t dt \quad (\text{A.2})$$

di punto iniziale x_0 , dove il termine $\sigma(X_t, t)\epsilon_t$ descrive delle fluttuazioni casuali interagenti con il sistema considerato.

L'equazione (A.2) costituisce un modello generale di EDS adeguato per numerosi sistemi naturali ed ingegneristici soggetti a perturbazioni casuali variabili nel tempo.

Uno degli esempi più noti della (A.2) è l'equazione di Langevin per il moto Browniano comunemente riconosciuta come l'origine delle EDS.

Si ricorda che un moto Browniano $\{W_t, t > 0\}$ è un processo di Markov definito nello spazio di probabilità $(\Omega, \mathcal{F}, \mathcal{P})$ valori iniziale $W_0 = 0$, ad incrementi indipendenti, con $W_t - W_s \sim \mathcal{N}(0, (t - s))$.

Seguendo questa impostazione il termine ϵ_t è una variabile aleatoria gaussiana standard per ogni t , e $\sigma(X_t, t)$ è un fattore di intensità spesso dipendente dal tempo. Per la (A.2) si può quindi definire l'equazione integrale soluzione:

$$X_t = x_0 + \int_{t_0}^t \mu(X_s, s) ds + \int_{t_0}^t \sigma(X_s, s) \epsilon_s ds. \quad (\text{A.3})$$

Osservando che l'integrale indefinito di ϵ_t è un moto Browniano, cioè $dW_t = \epsilon_t dt$, la (A.3) può essere riscritta nella forma :

$$X_t = c + \int_{t_0}^t \mu(X_s, s) ds + \int_{t_0}^t \sigma(X_s, s) dW_s \quad (\text{A.4})$$

che è impiegata usualmente in letteratura.

Le condizioni richieste per l'esistenza e l'unicità della soluzione indicata della (A.4) sono state illustrate nel paragrafo 2.2.

Il problema della soluzione della (A.4) è legata al moto Browniano: infatti le traiettorie di W_t , rispetto a cui si integra, non sono a variazione limitata, pertanto il secondo integrale della (A.4) non può essere un integrale di Riemann-Stieltjes. Si necessita quindi di una nuova definizione di integrale che permetta di risolvere questo problema, cioè di un *integrale stocastico*.

Gli approcci di maggiore interesse sono stati proposti da Ito (1951)¹ e da Stratonovich (1966)². L'*integrale stocastico di Ito*, ricopre un ruolo di grande rilievo per l'applicazione delle EDS specialmente in forza della sua proprietà di essere, come funzione dell'estremo superiore di integrazione, una martingala.

L'integrale stocastico di Ito viene dapprima definito per funzioni a scalini, e successivamente esteso per ogni funzione arbitraria attraverso approssimazione di tali funzioni tramite quelle a gradini.

Si consideri, quindi, uno spazio di probabilità $(\Omega, \mathcal{F}, \mathcal{P})$ su cui è definito un processo di moto Browniano $\{W_t, t \leq 0\}$ e si indichi con (\mathcal{F}_t) la famiglia delle più piccole σ -algebre che rende W_t misurabili, tali che $(\mathcal{F}_t) \in \mathcal{F}$. Sia $t_0 < t_1 < \dots < t_n = T$ una partizione dell'intervallo di integrazione $[t_0, T]$; una funzione σ è una funzione a gradini se $\sigma(s) = \sigma(t_{i-1})$ per ogni $s \in [t_{i-1}, t_i)$ con $i = 1, 2, \dots, n$. Per ogni funzione a gradini si definisce *integrale stocastico secondo Ito* di σ rispetto al moto Browniano W_t :

$$\int_{t_0}^t \sigma(s) dW_s = \sum_{i=1}^n \sigma(t_{i-1})(W_{t_i} - W_{t_{i-1}}).$$

Il valore atteso di tale integrale è nullo, e si ha che

$$\int_{t_0}^t (a\sigma_1(s) + b\sigma_2(s)) dW_s = a \int_{t_0}^t \sigma_1(s) dW_s + b \int_{t_0}^t \sigma_2(s) dW_s.$$

¹Ito, K. (1951), On Stochastic Differential Equation *American Mathematical Society*, n.4

²Stratonovich, R.L. (1966), A new representation for stochastic integrals and equations, *SIAM J. Control*, vol. 4

Per ogni arbitraria funzione σ \mathcal{F}_t -misurabile esiste una successione di funzioni a gradini tale che $\lim_{n \rightarrow \infty} \int_{t_0}^t |\sigma(s) - \sigma_n(s)|^2 ds = 0$ con probabilità uno, si definisce allora

$$\int_{t_0}^t \sigma_n(s) d\mathcal{W}_s = \sum_{i=1}^n \sigma_n(t_{i-1})(\mathcal{W}_{t_i} - \mathcal{W}_{t_{i-1}}).$$

Si può dimostrare che $\lim_{n \rightarrow \infty} \int_{t_0}^t \sigma_n(s) d\mathcal{W}_s = I(\sigma)$ in probabilità, dove $I(\sigma)$ è una variabile casuale indipendente dalla specifica scelta della successione σ_n .

Definizione. Per ogni funzione σ , \mathcal{F}_t -misurabile, definita in $[t_0, T]$, l'*integrale stocastico secondo Ito* di σ rispetto al moto Browniano \mathcal{W}_t su $[t_0, T]$ è definito come la variabile aleatoria $I(\sigma)$ quasi-ovunque unicamente determinata da:

$$\int_{t_0}^t \sigma(s) d\mathcal{W}_s = \lim_{n \rightarrow \infty} \int_{t_0}^t \sigma_n(s) d\mathcal{W}_s$$

dove σ_n è una successione di funzioni a gradini che approssimano σ nel senso che $\lim_{n \rightarrow \infty} \int_{t_0}^t |\sigma(s) - \sigma_n(s)|^2 ds = 0$ q.o. \square

Tale integrale gode di alcune proprietà tra le quali ricordiamo:

1. Il valore dell'integrale resta immutato se la funzione σ viene cambiata per ogni s in un insieme di misura di Lebesgue nulla.
2. L'integrale stocastico come funzione del suo estremo superiore è una *martingala* rispetto alla famiglia di σ -algebre \mathcal{F}_t
3. Sempre come funzione del suo estremo superiore, le traiettorie dell'integrale stocastico sono quasi ovunque continue.
4. Gli incrementi dell'integrale stocastico sono indipendenti
5. Se la funzione σ è indipendente da $\omega \in \Omega$ spazio degli eventi, l'integrale stocastico è un processo Gaussiano in $[t_0, T]$ di media nulla. Questo è ad esempio il caso di σ funzione costante.

Un esempio classico del calcolo dell'integrale di Ito è il seguente:

$$\int_{t_0}^t \mathcal{W}_s d\mathcal{W}_s = \frac{1}{2}(\mathcal{W}_t^2 - \mathcal{W}_{t_0}^2) - \frac{1}{2}(t - t_0).$$

Il lettore interessato ad una rigorosa introduzione di base alle EDS e all'integrale secondo Ito può consultare il testo di Arnold (1974); il testo di Sobczyk (1991) approfondisce l'approccio applicativo verso la fisica e l'ingegneria. Il riferimento al testo di Ikeda e Watanabe (1989), dove è presente una trattazione approfondita delle EDS, è rivolto al lettore già in possesso di nozioni sul tema affrontato e di ampie conoscenze matematiche.

Appendice B

La Soluzione approssimata della EDS lineare (2.10)

Si consideri l'EDS (2.10) di valore iniziale $X_0 = c$:

$$dX_t = k(\theta - X_t)dt + \sigma g(X_t, \gamma)dW_t.$$

La determinazione della seguente soluzione approssimata segue quella proposta da Renault (1989) ed ha un carattere intuitivo. Si consideri l'equazione differenziale deterministica:

$$\frac{dX}{dt} + kX = b(t)$$

con valore iniziale $X_0 = c$. Com'è noto la soluzione generale di tale equazione ha la forma:

$$X(t) = ce^{-kt} + e^{-kt} \int_0^t e^{ks} b(s) ds \quad (\text{B.1})$$

Ponendo

$$b(s) = k\theta + \sigma^2 g^2(X_t, \gamma) \dot{W}(t)$$

e sostituendo nella (B.1) si ha l'espressione (2.12), ossia:

$$X_{t+\delta t} = \theta(1 - e^{-k\delta t}) + e^{-k\delta t} X_t + e^{-k(t+\delta t)} \oint_t^{t+\delta t} \sigma g(X_s, \gamma) e^{ks} dW_s$$

Nel caso in cui $g(X_t, \gamma)$ sia una funzione lineare in X_t la soluzione viene ricondotta alla (2.4).

Nel caso in cui $g(X_t, \gamma)$ sia funzione non lineare in X_t , l'integrale di Ito

$$\oint \sigma g(X_s, \gamma) \exp(k\delta t) dW_s$$

si può approssimare come segue. Si consideri una partizione dell'intervallo $[0, t]$ del tipo $0 < t_1 < \dots < t_k < \dots < t$ per $k = 1, 2, \dots$ e con passo $t_{k+1} - t_k = \delta t$ sufficientemente piccolo. Nell'intervallo $[t_k, t_k + \delta t]$ la funzione integranda dell'integrale di Ito può essere considerata costante, pertanto si può scrivere la seguente soluzione approssimata:

$$e^{-k(t+\delta t)} * \int_t^{t+\delta t} \sigma g(X_s, \gamma) e^{ks} dW_s \cong e^{-k\delta t} \sigma g(X_t, \gamma) [W_{t+\delta t} - W_t].$$

L'integrale ha quindi, in un intervallo di ampiezza δt , approssimativamente una legge normale del tipo:

$$N(0, e^{-2k\delta t} \sigma^2 g^2(X_t, \gamma) \delta t).$$

Si può allora scrivere:

$$X_{t_i+\delta t} = \theta(1 - e^{-k\delta t}) + e^{-k\delta t} X_{t_i} + \epsilon_{t_i+\delta t}$$

dove

$$\epsilon_{t_i+\delta t} \sim N(0, e^{-2k\delta t} \sigma^2 g^2(X_{t_i}, \gamma) \delta t).$$

Appendice C

Il Metodo Generalizzato dei Momenti

L'idea del Metodo Generalizzato dei Momenti (GMM) nasce dall'osservazione che se il modello è esattamente specificato, allora il valore vero dei parametri di interesse verificherà certe condizioni sui momenti. Quindi tali restrizioni possono essere usate non solo come test per la specificazione del modello ma anche per definire i parametri. Si supponga, per esempio, che la variabile X abbia una distribuzione del tipo $\phi(x, \theta)$ con $\mathbf{E}[X] = \theta$, allora deve essere $\mathbf{E}[X - \theta] = 0$.

Uno dei vantaggi dei GMM come metodo di stima è che esso ammette modelli che sono consistenti con diversi Processi Generatori dei Dati (DGP). Ogni DPG è ammissibile se soddisfa un numero relativamente piccolo di restrizioni sui momenti, cosa che ad esempio non è possibile con il metodo di Massima Verosimiglianza dove il modello deve essere completamente specificato.

La condizione sui momenti che sotto esatta specificazione del modello sarà verificata dal valore vero dei parametri θ_0 sarà del tipo:

$$\mathbf{E}[f_i(x_t, \theta_0)] = 0 \quad i = 1, 2, \dots, J$$

e verrà indicata col nome di *condizione di ortogonalità*, J indica il numero di condizioni considerate, e k il numero dei parametri di interesse.

Si supponga di osservare X_1, X_2, \dots, X_T variabili i.i.d. allora dal Teorema del Limite Centrale possiamo affermare che

$$\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T f_i(x_t, \theta) = \mathbf{E}[f_i(x_t, \theta)] \quad q.o. \quad i = 1, 2, \dots, J.$$

Si può quindi impostare una funzione obiettivo per il problema di stima del tipo Errore Quadratico Medio minimo. Poichè il valore medio delle condizioni di ortogonalità è nullo per ipotesi per il valore vero dei parametri θ_0 , si ha che lo stimatore dei GMM è dato da:

$$\hat{\theta}(D_t) = \operatorname{argmin}_{\theta} \left(\frac{1}{T} \sum_{t=1}^T f(x_t, \theta) \right)' D_t \left(\frac{1}{T} \sum_{t=1}^T f(x_t, \theta) \right)$$

dove D_t è una matrice di pesi simmetrica definita positiva, convergente a D_0 definita positiva e f indica il vettore delle condizioni. Sotto le condizioni di regolarità proposte da Hansen (1982) lo stimatore $\hat{\theta}(D_t)$ è consistente, asintoticamente normale,

con matrice di covarianza dipendente dalla matrice dei pesi D_t scelta:

$$Var(\sqrt{T}(\hat{\theta} - \theta_0)) \approx (\Gamma' D_t \Gamma)^{-1} \Gamma' D_t \Omega D_t \Gamma (\Gamma' D_t \Gamma)^{-1}$$

dove $\Gamma = \left(\frac{1}{T} \sum_{i=1}^T \frac{\partial}{\partial \theta} f(x_t, \theta) \right)$ e Ω indica la matrice di covarianza delle f . Condizione necessaria e sufficiente perchè le stime GMM siano efficienti è che la matrice dei pesi D_t sia uguale a Ω^{-1} inversa della matrice di covarianza per le condizioni f . In base a tale scelta si ha:

$$Var(\sqrt{T}(\hat{\theta} - \theta_0)) \approx (\Gamma' \Omega^{-1} \Gamma)^{-1}.$$

Si possono distinguere due casi: quello in cui il numero delle condizioni prese in esame è pari al numero dei parametri, quindi $k = J$, e si parlerà di GMM esattamente identificati, e quello in cui $k < J$, parleremo di GMM sovraidentificati. Nel primo caso la forma quadratica che definisce il problema di ottimo assume valore zero nel punto minimo, e la condizione di minimo di primo ordine non dipende dalla matrice dei pesi D_t . In tal caso dunque la varianza dello stimatore non dipende dai pesi scelti e risulta minima. Nel caso dei GMM sovraidentificati la corretta determinazione della matrice Ω assume un ruolo fondamentale per l'efficienza delle stime. Nel caso in cui non si abbia ragione di ritenere che le condizioni f siano serialmente autocorrelate, allora:

$$\Omega \approx \left(\frac{1}{T} \sum_{t=1}^T f(x_t, \hat{\theta}) \right)' \left(\frac{1}{T} \sum_{t=1}^T f(x_t, \hat{\theta}) \right).$$

Qualora si ritenga che le condizioni f presentino un certo grado di autocorrelazione seriale, cosa abbastanza frequente nello studio delle serie storiche, la matrice Ω deve essere stimata correttamente. In generale l'ordine di autocorrelazione non è noto e deve essere determinato in base a considerazioni esterne alla procedura di stima.

Sono molte le stime di Ω proposte in letteratura in letteratura e si rifanno per lo più a criteri non parametrici. La più nota è quella proposta da Newey e West (1987).

Quindi nel caso sovraidentificato la procedura di stima avviene in due passi:

1. Si determina la stima $\tilde{\theta}$ di prima approssimazione scegliendo come pesi la matrice identità. Le stime così ottenute sono consistenti ma inefficienti.
2. Si calcola Ω per $\tilde{\theta}$ e si determina

$$\hat{\theta}(\Omega) = \underset{\theta}{\operatorname{argmin}} \left(\frac{1}{T} \sum_{t=1}^T f(x_t, \theta) \right)' \Omega^{-1}(\tilde{\theta}) \left(\frac{1}{T} \sum_{t=1}^T f(x_t, \theta) \right)$$

dove $\Omega(\tilde{\theta}) = Var \left(\frac{1}{T} \sum_{t=1}^T f(x_t, \tilde{\theta}) \right)$. Le stime $\hat{\theta}$ così determinate risultano consistenti ed efficienti.

Hansen (1982) propone un test per sottoporre a verifica la significatività delle condizioni sovraidentificate, ovvero se considerare un numero di condizioni maggiori a quello dei parametri apporta un miglioramento statisticamente significativo alla qualità delle stime. Tale test si basa sul fatto che la quantità:

$$T \left(\frac{1}{T} \sum_{t=1}^T f(x_t, \hat{\theta}) \right)' D_t \left(\frac{1}{T} \sum_{t=1}^T f(x_t, \hat{\theta}) \right) \sim \chi_{(J-k)}^2$$

quando D_t tende a Ω^{-1} .

I metodi di stima dei Minimi Quadrati (LS), delle Variabili Strumentali (IV) e il metodo della Massima Verosimiglianza (ML) possono essere ricondotti ad una procedura di tipo GMM (Newey e McFadden, 1994).

Si indichi con $\mathcal{L}(\theta | x)$ la funzione di Verosimiglianza e sia $\hat{\theta} = \operatorname{argmax}_{\theta} \mathcal{L}(\theta | x)$ la stima di ML. La condizione di minimo di primo ordine è data: $\frac{\partial}{\partial \theta} \mathcal{L}(\theta | x) = 0$, quindi $\mathbf{E} \left[\frac{\partial}{\partial \theta} \mathcal{L}(\theta | x) \right] = 0$ possono essere lette come condizioni di ortogonalità. Si ha inoltre che Γ è il valore atteso della matrice hessiana e Ω è il valore atteso dell'outer-product delle derivate prime. Poichè il sistema delle condizioni di ortogonalità è esattamente identificato la matrice di covarianza per le stime coincide, come atteso, con l'inversa della matrice di informazione di Fisher.

Nel metodo dei LS nel modello lineare classico si suppone che $\mathbf{E}[Xe] = 0$, cioè si richiede che le variabili esplicative siano incorrelate con gli errori. Tale condizione può essere letta come una restrizione di ortogonalità dei GMM, e in modo analogo per il metodo delle IV dove la condizione di incorrelazione tra le variabili strumentali e gli errori possono essere lette come condizioni di ortogonalità dei GMM. Per ulteriori approfondimenti e una trattazione più completa dei GMM si rimanda il lettore interessato al testo di Davidson e MacKinnon (1993, cap.17), al lavoro di Ogaki (1992) e a quello di Newey e McFadden (1994).

Appendice D

La matrice di covarianza dei GMM con parametri di disturbo

La matrice di covarianza per le stime GMM con parametri di disturbo è stata ricavata seguendo la dimostrazione fornita da Gourieroux, Monfort e Renault (1991) per le stime GMM a due stadi.

Siano

$$\mathbf{E}[f(x_t, \eta, \hat{\nu}) | \hat{\nu}] = 0$$

le condizioni di ortogonalità per i parametri di interesse η dove $\hat{\nu}$ sono stime consistenti dei parametri di disturbo ν , e sia

$$\mathcal{G} = \mathbf{E} \left[\frac{\partial f}{\partial \eta} \right] = \mathbf{E} \left[\frac{\partial f(x_t, \eta_0, \nu_0)}{\partial \eta} \right]$$

e $\Omega(f) = \Omega(f(x_t, \eta_0, \nu_0))$ la matrice di covarianza tra le condizioni f , dove η_0 e ν_0 indicano il valore vero dei parametri. Siano

$$\mathbf{E}[g(x_t, \nu)] = 0$$

le condizioni di ortogonalità per la stime GMM dei parametri di disturbo ν e sia

$$\Gamma = \mathbf{E} \left[\frac{\partial g}{\partial \nu} \right] = \mathbf{E} \left[\frac{\partial g(x_t, \nu_0)}{\partial \nu} \right]$$

e con $\Sigma(g) = \Sigma(g(x_t, \nu_0))$ la matrice di covarianza tra le condizioni g .

Lo stimatore GMM con parametri di disturbo è ottenuto come soluzione del problema di minimo:

$$\hat{\eta}(W) = \underset{\eta}{\operatorname{argmin}} \left(\frac{1}{T} \sum_{t=1}^T f(x_t, \eta, \hat{\nu}) \right)' W_T^{-1} \left(\frac{1}{T} \sum_{t=1}^T f(x_t, \eta, \hat{\nu}) \right)$$

La condizione di minimo di primo ordine è data da:

$$\frac{1}{T} \sum_{t=1}^T \frac{\partial f}{\partial \eta'}(x_t, \hat{\eta}, \hat{\nu}) W_T^{-1} \frac{1}{T} \sum_{t=1}^T f(x_t, \hat{\eta}, \hat{\nu}) = 0$$

Si consideri l'espansione del primo ordine di $\frac{1}{T} \sum_{t=1}^T f(x_t, \hat{\eta}, \hat{\nu})$ in un intorno di (η_0, ν_0) e moltiplicando tutto per \sqrt{T} si ha:

$$\begin{aligned} \frac{1}{\sqrt{T}} \sum_{t=1}^T f(x_t, \hat{\eta}, \hat{\nu}) &\approx \frac{1}{\sqrt{T}} \sum_{t=1}^T f(x_t, \eta_0, \nu_0) + \frac{1}{T} \sum_{t=1}^T \frac{\partial f}{\partial \eta}(x_t, \eta_0, \nu_0) \sqrt{T}(\hat{\eta} - \eta_0) \\ &+ \frac{1}{T} \sum_{t=1}^T \frac{\partial f}{\partial \nu}(x_t, \eta_0, \nu_0) \sqrt{T}(\hat{\nu} - \nu_0) \end{aligned} \quad (\text{D.1})$$

Sostituendo nella condizione di primo ordine:

$$\begin{aligned} \frac{1}{T} \sum_{t=1}^T \frac{\partial f}{\partial \eta'}(x_t, \hat{\eta}, \hat{\nu}) W_T^{-1} \left(\frac{1}{\sqrt{T}} \sum_{t=1}^T f(x_t, \eta_0, \nu_0) \right. \\ \left. + \frac{1}{T} \sum_{t=1}^T \frac{\partial f}{\partial \eta}(x_t, \eta_0, \nu_0) \sqrt{T}(\hat{\eta} - \eta_0) + \frac{1}{T} \sum_{t=1}^T \frac{\partial f}{\partial \nu}(x_t, \eta_0, \nu_0) \sqrt{T}(\hat{\nu} - \nu_0) \right) = 0 \end{aligned}$$

Sotto condizioni di regolarità, per il Teorema del Limite Centrale si ha:

$$\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \frac{\partial f}{\partial \eta}(x_t, \hat{\eta}, \hat{\nu}) = \mathbf{E} \left[\frac{\partial f}{\partial \eta} \right] = \mathcal{G}$$

e dove $\frac{\partial f}{\partial \eta} = \frac{\partial f}{\partial \eta}(x_t, \eta_0, \nu_0)$.

Quindi la condizione di primo ordine diventa:

$$\mathbf{E} \left[\frac{\partial f}{\partial \eta'} \right] W^{-1} \left(\frac{1}{\sqrt{T}} \sum_{t=1}^T f + \mathbf{E} \left[\frac{\partial f}{\partial \eta} \right] \sqrt{T}(\hat{\eta} - \eta_0) + \mathbf{E} \left[\frac{\partial f}{\partial \nu} \right] \sqrt{T}(\hat{\nu} - \nu_0) \right) = 0$$

dove $f = f(x_t, \eta_0, \nu_0)$.

Sviluppando il prodotto e ricordando che $\mathcal{G} = \mathbf{E} \left[\frac{\partial f}{\partial \eta} \right]$ si ha che:

$$\mathcal{G}' W^{-1} \frac{1}{\sqrt{T}} \sum_{t=1}^T f + \mathcal{G}' W^{-1} \mathcal{G} \sqrt{T}(\hat{\eta} - \eta_0) + \mathcal{G}' W^{-1} \mathbf{E} \left[\frac{\partial f}{\partial \nu'} \right] \sqrt{T}(\hat{\nu} - \nu_0) \approx 0.$$

Ricavando $\sqrt{T}(\hat{\eta} - \eta_0)$ si ha:

$$\begin{aligned} \sqrt{T}(\hat{\eta} - \eta_0) \\ \approx (\mathcal{G}' W^{-1} \mathcal{G})^{-1} \mathcal{G}' W^{-1} \frac{1}{\sqrt{T}} \sum_{t=1}^T f + (\mathcal{G}' W^{-1} \mathcal{G})^{-1} \mathcal{G}' W^{-1} \mathbf{E} \left[\frac{\partial f}{\partial \nu'} \right] \sqrt{T}(\hat{\nu} - \nu_0) \end{aligned}$$

e quindi

$$\sqrt{T}(\hat{\eta} - \eta_0) \approx (\mathcal{G}' W^{-1} \mathcal{G})^{-1} \mathcal{G}' W^{-1} \left[\frac{1}{\sqrt{T}} \sum_{t=1}^T f + \mathbf{E} \left[\frac{\partial f}{\partial \nu'} \right] \sqrt{T}(\hat{\nu} - \nu_0) \right]$$

Se la funzione

$$\frac{1}{\sqrt{T}} \sum_{t=1}^T f$$

e lo stimatore $\sqrt{T}(\hat{\nu} - \nu_0)$ sono congiuntamente asintoticamente normali allora $\sqrt{T}(\hat{\eta} - \eta_0)$ è asintoticamente normale, con matrice di covarianza asintotica data da:

$$Var_{asy}\sqrt{T}(\hat{\eta} - \eta_0) = (\mathcal{G}'W^{-1}\mathcal{G})^{-1}\mathcal{G}'W^{-1}\Omega_{\star}W^{-1}\mathcal{G}(\mathcal{G}'W^{-1}\mathcal{G})^{-1}$$

dove

$$\Omega_{\star} = Var_{asy} \left(\frac{1}{\sqrt{T}} \sum_{t=1}^T f + \mathbf{E} \left[\frac{\partial f}{\partial \nu'} \right] \sqrt{T}(\hat{\nu} - \nu_0) \right)$$

Quest'ultima matrice è sviluppata nel corso del paragrafo 3.5.

In accordo con la teoria dei GMM anche in questo caso esiste la scelta ottimale della matrice dei pesi W che è $W = \Omega_{\star}$ che minimizza la varianza asintotica della stime. Per questa specifica scelta si ha:

$$Var_{asy}\sqrt{T}(\hat{\eta} - \eta_0) = (\mathcal{G}'\Omega_{\star}^{-1}\mathcal{G})^{-1}$$

Appendice E

Le condizioni di regolarità per i GMM con parametri di disturbo

Condizioni di regolarità per l'esistenza, la consistenza e l'asintotica normalità degli stimatori GMM con parametri di disturbo possono essere dedotte da quelle proposte da Gourieroux, Monfort e Renault (1991) per gli stimatori GMM a due stadi.

Si pone:

1. Il processo $\{X_t, t \in [t_0, T]\}$ è stazionario ed ergodico.
2. I parametri di interesse η appartengono ad un sottoinsieme compatto B di \mathbb{R}^k , dove k è il numero di parametri di interesse, e il valore vero η_0 di η è interno a B .
3. La funzione $f(x_t, \eta, \nu)$ delle condizioni di ortogonalità è misurabile rispetto a x e differenziabile rispetto a η e a ν per ogni η , e ν appartenenti a B .
4. Il valore atteso $\mathbf{E}[f(x_t, \eta, \hat{\nu})]$ esiste per ogni η e ν .
5. Il sistema $\mathbf{E}[f(x_t, \eta, \hat{\nu})] = 0$ ammette η_0 come unica soluzione.
6. Il rango della matrice $\mathbf{E} \left[\frac{\partial f}{\partial \eta}(x_t, \eta, \nu) \right]$ è k .
7. Il processo $Z_t = \begin{pmatrix} f(x_t, \eta_0, \nu_0) \\ g(x_t, \nu_0) \end{pmatrix}$, $-\infty < t < +\infty$, dove $g(x_t, \nu_0)$ indicano le condizioni di ortogonalità per la stima consistente dei parametri di disturbo, è tale per cui $\mathbf{E}(Z_0 Z_0')$ esiste ed è finita e $\mathbf{E}(Z_0 | Z_{-j}, Z_{-j-1}, \dots)$ converge in media quadratica a zero per j che tende all'infinito.
8. Il processo $Z_j = \mathbf{E}(Z_0 | Z_{-j}, Z_{-j-1}, \dots) - \mathbf{E}(Z_0 | Z_{-j-1}, Z_{-j-2}, \dots)$, per $j \geq 0$ è tale che $\sum_{j=0}^{\infty} \mathbf{E}(Z_j' Z_j)^{\frac{1}{2}}$ è finita.

Le condizioni 7 ed 8 implicano che

$$\mathbf{E} \begin{bmatrix} f(x_t, \eta_0, \nu_0) \\ g(x_t, \nu_0) \end{bmatrix} = 0, \quad -\infty < t < +\infty$$

e forniscono condizioni sufficienti per l'applicazione del Teorema del Limite Centrale per processi stazionari ed ergodici (Hansen 1982).

Bibliografia

- Ait-Sahalia, Y. (1994), Nonparametric Pricing of Interest Rate Derivative Securities, *W.P.*, Graduate School of Business, University of Chicago.
- Amemiya, T. (1977) The Maximum Likelihood and the Nonlinear Three-Stage Least Squares Estimators in General Nonlinear Simultaneous Equation Model, *Econometrica*, vol. 45, 955-968.
- Arnold, L. (1975) *Stochastic Differential Equations* New York, John Wiley and Sons.
- Bianchi, C., R. Cesari, L. Panattoni (1995) Alternative Estimation on Cox Ingersoll Ross Model of the Term Structure of Interest Rates: a Monte Carlo Comparison, *Banca d'Italia, Temi di Discussioni* n. 326.
- Bernnan, M.J., E.S. Schwartz (1982) An Equilibrium Model of Bond Pricing and Test of Market Efficiency, *Journal of Financial and Quantitative Analysis*, vol. 17, 301-329.
- Black, F., M. Scholes (1973) The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, vol. 81, 637-659.
- Blake D. (1990) *Financial Market Analysis* McGraw-Hill, London
- Bollerslev T. (1986) Generalized Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics*, vol 31, 307-327
- Bollerslev T., R.F. Engle, D.B. Nelson (1994) ARCH models, in *Handbook of Econometrics*, vol. 4, by F.Engle and D. McFadden, Elsevier North-Holland, Amsterdam, 2959-3038.
- Burguete, J., R. Gallant, G. Souza (1982) On Unification of Asymptotic Theory of Nonlinear Econometric Models, *Econometric Reviews*, 151-190.
- Calzolari, G., G. Fiorentini (1994) Conditional Heteroskedasticity in Nonlinear Simultaneous Equations, *European University Institute, Firenze* WP ECO n.94/44.
- Cox, J.C., J. Ingersoll and S.A. Ross (1985) A Theory of Term Structure of Interest Rates, *Econometrica*, vol. 53, n.2, 385-408.

- Day T.E., C.M. Lewis (1992), Stock Market Volatility and Information Content of Stock Index Options, *Journal of Econometrics*, vol 52, 267-288
- Davidson, R. and J.G. Mackinnon (1993), *Estimation and Inference in Econometrics*, Oxford University Press.
- Devroye, L. (1986) *Nonuniform Random Variate Generation* Springer-Verlag, New York
- Dynkin, E.B. (1965) *Markov Process*, Springer-Velag, Berlin.
- Duffie, D. and K. Singleton (1993), Simulated Moments Estimation of Markov Models of Asset Prices, *Econometrica*, vol. 61,929-952.
- Engle R.F. (1982) Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Inflation, *Econometrica*, vol 50, 987-1007
- Feller, W. (1951) Two Singular Diffusion Problems, *Annals of Mathematics*, vol. 54, 173-182.
- Gallant, R. and G. Tauchen (1992) Which Moment to Match? WP, Duke University.
- Genon-Catalot, V. and J. Jacob (1993) On the Estimation of the Diffusion Coefficient for Multidimensional Diffusion Processes, *Annales de l'Istitut Henri Poincaré*, vol. 29, n. 1, 119-51.
- Ghysels, E., A. Harvey, E. Renault (1995) Stochastic volatility, in *Handbook of Statistics* vol. 14 (forthcoming), North-Holland, Amsterdam.
- Gourieroux, C., A. Monfort (1989) *Statistique et Modèles Econométriques vol.2*, Economica, Paris
- Gourieroux, C., A. Monfort (1993) Pseudo Likelihood Methods, in *Handbook of Statistics*, vol. 11, by G.S. Maddala, C.R. Rao and H.D. Vinod, North Holland, Amsterdam.
- Gourieroux, C., A. Monfort (1993b) Simulation Based Inference: a Survey with Special Reference to Panel Data Models, *Journal of Econometrics*, vol.59, 5-33.
- Gourieroux, C., A. Monfort (1994) *Simulation Based Econometric Methods*, C.O.R.E.(forthcoming), Louvain.
- Gourieroux, C., A. Monfort, E. Renault (1991) Two Stage Generalized Method of Moments with Applications to Regressions with Heteroskedasticity of Unknown Form, *WP n.91-15 GREMAQ*, Université Toulouse 1.
- Gourieroux, C., A. Monfort, E. Renault (1993) Indirect Inference, *Journal of Applied Econometrics*, vol. 8, 85-118.

- Gourieroux, C., A. Monfort, A. Trognon (1984) Pseudo Maximum Likelihood Methods: Theory, *Econometrica*, vol. 52, 681-700.
- Hansen, L.P. (1982) Large Sample Properties of Generalized Method of Moment Estimators, *Econometrica*, vol 50, 1029-1054.
- Hansen, L.P. and J. Scheinkman (1995) Back to the Future: Generating Moment Implications for Continuous Time Markov Process, *Econometrica*, vol 63, n.4, 767-804.
- Hull, J. and A. White (1987) The Pricing Options on Assets with Stochastic Volatility, *Journal of Finance*, vol. 3, 281-300.
- Ikeda, N. and S. Watanabe (1989) *Stochastic Differential Equations and Diffusion Processes* North Holland, Amsterdam.
- Jamshidian, F. (1989), An Exact Bond Option Formula, *The Journal of Finance*, vol. 44, n.1, 205-209.
- Karlin, S. and H.M. Taylor (1981) *A Second Course in Stochastic Processes*, Academic Press, New York.
- Kloeden, P.E. and E. Platen (1992) *Numerical Solution of Stochastic Differential Equations*, Springer-Verlag, Berlin.
- Lipster, R.S. and A.N. Shiriyayev (1977) *Statistics of Random Processes I: General Theory*, Springer-Verlag, Berlin.
- Lo, A. (1988) Maximum Likelihood Estimation on Generalized Ito Processes with Discretely Sampled Data, *Journal of Econometric Theory*, vol. 4, 213-247.
- Melino, A. and S.M. Turnbull (1990) Pricing Foreign Currency Options with Stochastic Volatility, *Journal of Econometrics*, vol. 45, 239-265.
- Nelson, D.B. (1990) ARCH Models as Diffusion Approximations, *Journal of Econometrics*, vol. 45, 7-38.
- Nelson, D.B. (1991) Conditional Heteroskedasticity in Asset Returns: a New Approach, *Econometrica*, vol 59, 347-370
- Newey, W. and D. McFadden (1994) Large Sample Theory and Hypothesis Testing, in *Handbook of Econometrics*, vol. 4, by F.Engle and D. McFadden, Amsterdam Elsevier North-Holland, 2113-2245.
- Ogaki, M. (1992) Generalized Method of Moment. Econometric Applications, in *Handbook of Statistics*, vol. 11, by G.S. Maddala, C.R. Rao and H.D. Vinod, North Holland, Amsterdam.
- Pakes, A., D. Pollard (1989) Simulation and Asymptotics of optimization Estimators, *Econometrica*, vol. 57, 1027-1058.

- Pastorello, S., E. Renault and N. Touzi (1994), Statistical Inference for Random Variance Option Pricing, *Quaderno n. 31*, Dipartimento di Scienze Economiche, Università di Padova.
- Renault, E. (1989) Statistique des Modeles Markoviens d'Evaluation d'Actifs Financiers, *mimeo GREMAQ*, Université Toulouse 1.
- Rudin, W. (1973) *Functional Analysis*, McGraw-Hill Book Co, New York.
- Smith, A. (1990) Three Essays on the Solution and Estimation of Dynamic Macroeconometric Models, *Ph.D Dissertation*, Duke University.
- Smith, A. (1993) Estimating Nonlinear Time-Series Models Using Simulated Vector Autoregressions, *Journal of Applied Econometrics*, vol. 8 63-84
- Sobczyk, S. (1991) *Stochastic Differential Equations with applications to physics and Engineering*, Kluwer Academic Publishers, London
- Taylor S. (1986) *Modelling Financial Time Series*, John Wiley and Sons, New York
- Vasicek, O.A. (1977) An Equilibrium Characterization of the Term Structure, *Journal of Financial Economics*, vol. 5, n. 2, 177-88.
- White, H. (1994) *Estimation, Inference and Specification Analysis*, Cambridge University Press.