

UNIVERSITÀ DEGLI STUDI DI PISA

FACOLTÀ DI SCIENZE MATEMATICHE, FISICHE E NATURALI

CORSO DI LAUREA IN MATEMATICA

ANNO ACCADEMICO 2005/2006

TESI DI LAUREA

**MODELLIZZAZIONE MATEMATICA E
ANALISI STATISTICA DEI TASSI DI
INTERESSE**

CANDIDATA

Stefania Caramaschi

RELATORE

Prof. Maurizio Pratelli

CONTRORELATORE

Prof. Giorgio Letta

Indice

Introduzione	v
1 Tassi di interesse	1
1.1 Tassi governativi	1
1.2 Tassi interbancari, LIBOR ED EURIBOR	1
1.3 Come misurare i tassi di interesse	2
1.4 L'attivo privo di rischio e il tasso a breve	4
1.5 Obbligazioni senza cedole e i tassi istantanei	5
1.6 La curva dei rendimenti	7
1.7 I tassi futuri	8
1.8 Obbligazioni con cedole e Swaps	10
1.8.1 Interest rate swaps	12
1.9 Rendimenti delle obbligazioni	13
1.10 Caps e Floors	16
1.11 Swaptions	17
2 La modellizzazione matematica dei mercati finanziari	19
2.1 Portafogli autofinanziati	20
2.2 Attivi casuali, arbitraggio e misure martingala	22
2.3 Completezza e assenza di arbitraggio	25
2.4 Valutazione (pricing)	26
2.5 Valutazione dei bonds e misura martingala	27
2.6 Relazioni fra df , dP e dr	28
3 Modelli matematici per i tassi di interesse	32
3.1 Generalità	33
3.2 Equazione fondamentale della struttura a termine	34
3.3 Modellizzazione martingala del tasso a breve	37
3.3.1 Esempi	38
3.4 Inversione della curva dei rendimenti	40
3.5 Perché i modelli multifattoriali per il tasso a breve?	42
3.6 Heath-Jarrow-Morton	44
3.6.1 Introduzione	44

3.6.2	Esistenza di una misura martingala	45
3.6.3	Modellizzazione martingala del tasso futuro	48
4	Analisi statistica della curva dei rendimenti	50
4.1	Analisi in componenti principali di un vettore aleatorio	50
4.2	PCA della curva dei rendimenti	53
4.3	Esempi	55
4.4	Interpretazione e discussione dei risultati	59
4.5	Metodi numerici	65
A	Speranza condizionale	69
A.1	Definizione ed esistenza	69
A.2	Proprietà elementari	70
B	L'integrale stocastico e la formula di Itô	72
B.1	Nozioni fondamentali sui processi stocastici	72
B.2	Martingale	74
B.3	Il Processo di Wiener	74
B.4	Il processo di Wiener multidimensionale	77
B.5	L'integrale stocastico di Itô	77
B.6	Calcolo di Itô	80
B.7	Teorema di Girsanov	81
	Bibliografia	83

Elenco delle figure

1.1	Curva dei rendimenti estratta dal mercato EURO il 13/02/2001	8
4.1	Grafici dei pesi relativi alle prime quattro componenti principali. Dati statunitensi, periodo 93–01	56
4.2	Grafici dei pesi relativi alle prime quattro componenti principali. Dati statunitensi, periodo 01–05	57
4.3	Grafici dei pesi relativi alle prime tre componenti principali. Dati del Regno Unito, periodo 89–92	58
4.4	Grafici dei pesi relativi alle prime quattro componenti principali. Dati del Regno Unito, periodo 88–95	60
4.5	Correlazione fra il tasso a breve e il tasso futuro con diverse maturità	61
4.6	Grafici dei pesi a_1, a_2, a_3 relativi alle prime tre componenti principali ottenuti con $\rho = 0.85$	63
4.7	Grafici della correlazione fra il primo tasso futuro e i successivi per $d=2, 3, 5$ e 10	64
4.8	Superficie della correlazione per $d=3, 5$ e 10	65

Elenco delle tabelle

1.1	Tassi EURIBOR del 10 Agosto 2006	2
1.2	Effetto su 100 euro della frequenza con la quale viene calcolato un tasso di interesse del 10% annuo	3
1.3	Zero-rates	14
1.4	Dati per il metodo bootstrap	15
4.1	Varianza espressa dalle dieci componenti della PCA dei tassi governativi statunitensi	55
4.2	Varianza espressa dalle otto componenti della PCA dei tassi del Regno Unito, periodo 89–92	57
4.3	Pesi a_1, a_2, a_3 relativi alle prime tre componenti principali. Dati del Regno Unito, periodo 89–92	58
4.4	Pesi a_1, a_2, a_3 e corrispondenti autovalori relativi alle prime tre componenti principali. Dati del Regno Unito, periodo 88–95	59
4.5	Varianza espressa dalle nove componenti della PCA dei tassi del Regno Unito, periodo 88–95	60
4.6	Grado di correlazione fra gli incrementi dei tassi futuri statu- nitensi con differenti maturità. Periodo 87–94	61
4.7	Pesi a_1, a_2, a_3 e corrispondenti autovalori relativi alle prime tre componenti principali ottenuti con $\rho = 0.85$	63
4.8	Varianza espressa dalle dieci componenti della PCA dei tassi statunitensi ottenuta con $\rho = 0.85$	64

Introduzione

Sicuramente il tasso di interesse è un “oggetto” con il quale tutti abbiamo avuto a che fare, ad esempio per la richiesta di un prestito, per aver fatto un investimento o semplicemente per avere un conto corrente in una qualsiasi banca. Una definizione semplicistica e intuitiva per il tasso di interesse è quella dell’ammontare di denaro che un debitore promette di pagare al suo creditore. Non appena però ci si addentra nel mondo dei tassi di interesse si ha a che fare con una giungla di definizioni e di tipi diversi di tassi: interbancari o governativi, istantanei o futuri, continuamente o semplicemente composti solo per citarne alcuni. Inoltre i prodotti commerciati sul mercato finanziario derivati dai tassi di interesse sono innumerevoli: opzioni, caps, floors, swaps e molti altri. Lo scopo del primo capitolo sarà proprio quello di familiarizzare con i tassi di interesse e con alcuni dei più semplici prodotti da lui derivati, portando esempi pratici per rendere più concrete le argomentazioni sviluppate.

Il secondo capitolo sarà interamente dedicato alla modellizzazione di un mercato finanziario: da un punto di vista matematico può essere rappresentato tramite uno spazio probabilizzato filtrato $(\Omega, \mathcal{F}, P, \underline{\mathcal{F}})$ sul quale viene definito un processo stocastico a valori in $(\mathbb{R}^{k+1}, \mathcal{B}(\mathbb{R}^{k+1}))$, inteso a modellizzare l’evolversi nel tempo dei prezzi di $k + 1$ attivi finanziari presi come riferimento.

Verrà inoltre affrontata la teoria dell’arbitraggio, introducendo le definizioni di strategia autofinanziata e di portafoglio. In breve l’arbitraggio è equivalente ad investire zero oggi e ricevere domani una quantità di denaro non negativa e positiva con probabilità positiva. Il mercato, talvolta, offre possibilità di arbitraggio. Visto però che tutti gli investitori cercano di sfruttarle essendo molto vantaggiose, si intuisce che il guadagno tende a diminuire sempre più, fino a sparire. Il mercato, quindi, tende, in modo naturale, ad eliminare tali possibilità. Per questo è naturale richiedere che nei nostri modelli non ci siano possibilità di arbitraggio.

Una delle questioni fondamentali che si propone la Finanza matematica è quella di valutare tali prodotti derivati, ovvero di calcolarne il prezzo. Per fare questo si assume la conoscenza del valore del fattore di sconto, ovvero la conoscenza di un continuo di prezzi $P(t, T)$ per ogni T con $T > t$. Questi prezzi, comunque, non sono direttamente osservabili e disponibili sul merca-

to, ma devono essere estrapolati dai prezzi degli strumenti finanziari che in quell'istante vengono commerciati. Il fattore di sconto $P(t, T)$ è la funzione che associa ad ogni istante t il prezzo di un bond che matura all'istante T , in altre parole $P(t, T)$ è il valore all'istante t di 1 euro prelevato all'istante T . Si chiama curva dei rendimenti la funzione $T \mapsto P(t, T)$. L'evoluzione della curva dei rendimenti può essere descritta dalla dinamica di diverse, ma in ogni caso equivalenti, quantità finanziarie, come per esempio il tasso istantaneo $R(t, T)$, i tassi futuri istantanei $f(t, T)$ o i prezzi dei bonds $P(t, T)$. Finché le leggi di trasformazione da una quantità all'altra sono note, la scelta delle variabili indipendenti è una questione puramente di convenienza. Le equazioni che legano le quantità appena menzionate sono:

$$P(t, T) = e^{-\int_t^T f(t, u) du} \quad f(t, T) = -\frac{\partial \ln P(t, T)}{\partial T} \quad P(t, T) = e^{R(t, T)(T-t)},$$

dove $f(t, T)$ è il tasso di interesse stabilito all'istante t per prestare o chiedere in prestito denaro fra l'istante T e $T + \varepsilon$ (con $T > t$), mentre $R(t, T)$ è il tasso previsto dall'istante t alla maturità T .

I modelli più classici adottano come variabile da modellizzare il tasso a breve $r(t)$ (definito come il limite del tasso futuro $f(t, T)$ quando la maturità T tende a t). Si suppone che la sua dinamica segua un'equazione della forma

$$dr(t) = \mu(t, r(t)) dt + \sigma(t, r(t)) dW(t)$$

dove W è un processo di Wiener. La diversa scelta per il termine di *drift* $\mu(t, r(t))$ e della volatilità $\sigma(t, r(t))$ porta a differenti modelli. Verranno brevemente presentati e discussi alcuni dei modelli classici del tasso a breve che presentano, però sia vantaggi che svantaggi:

- una volta conosciuta la legge di r si possono calcolare i prezzi dei derivati tramite la risoluzione di equazioni alle derivate parziali;
- dai prezzi dei bonds (di cui nella maggior parte dei casi si possono dare formule esplicite) si può ricostruire l'intera curva dei rendimenti;
- da un punto di vista economico, però, è piuttosto irragionevole che l'intero mercato sia governato da un'unica variabile;
- è difficile ottenere una struttura realistica per la volatilità dei tassi futuri senza introdurre modelli complicati;
- non appena il modello diventa più realistico, l'adattamento ai dati osservati diventa molto più difficile.

I difetti di questi modelli vengono in parte superati dai modelli in cui la variabile modellizzata è il tasso futuro $f(t, T)$, detti modelli HJM (Heath-Jarrow-Morton).

Nei modelli basati sul tasso a breve, l'evoluzione di una singola quantità, $r(t)$ appunto, caratterizza l'intera curva dei rendimenti. Dunque, la scelta di un modello "povero" per la dinamica di r , risulta un modello povero anche per la curva dei rendimenti. Per esempio, consideriamo un prodotto derivato il cui prezzo dipenda dai tassi $R(t, T_1)$ e $R(t, T_2)$ con maturità rispettivamente 1 anno e 10 anni. In questi modelli i tassi di interesse istantanei $R(t, T)$ risultano perfettamente correlati fra loro per ogni scelta della maturità T . Come conseguenza, un salto nella curva dei rendimenti all'istante t , viene trasmesso invariato a tutte le maturità e la curva, quando il suo punto iniziale (il tasso a breve r_t) compie tale salto, si muove rigidamente nella stessa direzione. Nella realtà, però, i tassi di interesse mostrano una certa decorrelazione (ovvero una non perfetta correlazione). I dati osservati suggeriscono che il grado di correlazione decresce con l'aumentare della distanza fra le maturità.

Si devono cercare allora modelli più soddisfacenti dal punto di vista dell'evoluzione della curva dei rendimenti. I modelli ad un solo fattore sono ancora molto usati quando il prodotto di cui si vuole calcolare il prezzo, non dipende dalla correlazione fra tassi con diverse maturità, ma da un'unico tasso dell'intera curva dei rendimenti (per esempio il tasso a sei mesi). Tuttavia, se i tassi congiunti da cui dipende il nostro prezzo hanno maturità molto vicine fra loro (per esempio sei mesi e un anno), ci si aspetta che la loro correlazione sia abbastanza alta. Così la perfetta correlazione indotta dai modelli ad un solo fattore non è, in linea di principio, inaccettabile. Quando però la correlazione gioca un ruolo rilevante, o comunque quando viene richiesta una maggior precisione, è necessario trovare nuovi approcci che consentano ad un modello di avere una struttura per la correlazione più realistica. Ecco che compaiono i modelli multifattoriali.

Una domanda che viene naturale porsi è: quanti fattori dobbiamo coinvolgere in un modello per soddisfare i nostri propositi pratici? La scelta di tale numero deve essere un buon compromesso fra un'efficiente implementazione numerica e la capacità del modello, sia di rappresentare una struttura realistica per la correlazione (e più in generale per la covarianza), sia di adattarsi, nella maggior parte dei casi e in modo soddisfacente, ai dati osservati nel mercato. Una risposta a questa domanda viene data dall'analisi in componenti principali (in inglese *principal components analysis* da cui la sigla PCA) dell'intera curva dei rendimenti. Essa ci suggerisce che due o tre componenti principali riescono già a mantenere dall'85% al 90% della varianza totale. Possiamo dunque supporre che un modello a due o tre fattori può essere utilizzato per avere risultati soddisfacenti. Nel capitolo 4 verranno fatti alcuni esempi, correlati da tabelle e numerosi grafici che aiuteranno a comprendere meglio l'utilizzo di questa tecnica.

Lo scopo della PCA è quello di ridurre la dimensione di un insieme di dati correlati fra loro, cercando però di mantenerne più alta possibile la varianza. Questa riduzione è ottenuta tramite una trasformazione (lineare) delle

variabili originali in un insieme di nuove variabili, le componenti principali appunto, che risultano fra loro non correlate. Inoltre vengono ordinate in modo tale che già le prime mantengano la maggior parte della varianza delle variabili originali.

Il calcolo delle componenti principali si riduce al calcolo degli autovettori (e corrispondenti autovalori) di una matrice reale, simmetrica e semidefinita positiva, ovvero della matrice delle covarianze. Nell'ultimo paragrafo viene brevemente presentato un metodo numerico per il calcolo degli autovettori e autovalori di una matrice.

L'appendice, lontana dall'essere una trattazione esauriente e completa, richiama sinteticamente tutte le nozioni di probabilità superiore (processi stocastici, processi di Wiener, formula di Itô e integrale stocastico) utilizzati in questa tesi.

Capitolo 1

Tassi di interesse

Il tasso di interesse può essere visto come l'ammontare del denaro che un debitore promette di pagare al suo creditore. Il tasso di interesse che viene applicato dipende dalle situazioni e dal rischio che il debitore ha di non riuscire a pagare (default), per mancanza di fondi, sia gli interessi sia il credito che gli è stato concesso. Più alto è il rischio, più alto è il tasso di interesse.

Ci sono diversi tipi di tassi di interesse e la prima distinzione da fare è tra i *tassi interbancari* (interbank rates) e i *tassi governativi* (treasury rates).

1.1 Tassi governativi

I tassi governativi sono i tassi che un investitore guadagna sui titoli di stato. Questi sono gli strumenti usati da un governo per investire nella propria valuta. Solitamente si assume che non ci sia la possibilità che un governo non riesca a pagare un'obbligazione sulla sua stessa valuta. Il motivo è che ogni governo può sopperire a questa mancanza stampando più moneta. I tassi governativi vengono dunque considerati totalmente privi di rischio, nel senso che, un investitore che acquista dei titoli di stato, è sicuro che sia gli interessi, sia il pagamento principale, gli saranno pagati come promesso.

Tuttavia, i commercianti di derivati (derivatives traders), solitamente non usano i tassi governativi come tassi privi di rischio, piuttosto utilizzano i tassi LIBOR od EURIBOR.

1.2 Tassi interbancari, LIBOR ED EURIBOR

I tassi interbancari sono i tassi di interesse ai quali avvengono gli scambi di depositi e le transazioni fra le banche. Un deposito presso una banca può essere considerato come un prestito a quella banca. Quest'ultima, perciò, dovrà soddisfare alcuni criteri per poter essere in grado di accettare, e quindi sostenere a tale tasso, dei depositi da un'altra banca. Tipicamente, per

soddisfare tali criteri, si intende essere classificati con una doppia A dalle *agenzie di rating*, agenzie che hanno il compito di valutare l'affidabilità di un emittente. Il livello massimo di affidabilità si ha quando si è classificati con una tripla A, ovvero quando la probabilità di default è dello 0.00%, mentre con una doppia A la probabilità è dello 0.02%. Questi tassi, quindi, non sono totalmente privi di rischio, tuttavia sono vicinissimi dall'esserlo.

Il più importante tasso interbancario, solitamente preso come riferimento per i contratti, è il tasso LIBOR (London Interbank Offered Rate) che viene fissato giornalmente a Londra. Un'altro molto usato è il tasso EURIBOR che invece viene fissato a Bruxelles e diffuso giornalmente dalla Federazione Bancaria Europea come media ponderata dei tassi di interesse ai quali le banche operanti nell'Unione Europea cedono i depositi in prestito.

<i>EURIBOR</i>	
Scadenza	Tasso in %
1 mese	3.1040
3 mesi	3.2150
4 mesi	3.2730
6 mesi	3.3940
9 mesi	3.5020
1 anno	3.5880

Tabella 1.1: Tassi EURIBOR del 10 Agosto 2006

1.3 Come misurare i tassi di interesse

Ci sono vari modi di misurare i tassi di interesse: ci si basa, infatti, sulla frequenza con la quale vengono composti che ne rappresenta l'unità di misura.

Supponiamo che un capitale C venga investito per n anni al tasso di interesse R per anno. Se il tasso è composto una volta per anno, il valore dell'investimento al termine degli n anni sarà

$$C(1 + R)^n. \quad (1.3.1)$$

Se invece il tasso R viene composto m volte per anno, il valore finale dell'investimento sarà

$$C\left(1 + \frac{R}{m}\right)^{mn}. \quad (1.3.2)$$

La tabella 1.2 mostra l'effetto prodotto, su 100 euro, dal crescere della frequenza di composizione di un tasso di interesse del 10% annuo.

<i>Frequenza di composizione</i>	<i>Valore di 100 euro al termine dell'anno(euro)</i>
Annuale ($m = 1$)	110.00
Semestrale ($m = 2$)	110.25
Quadrimestrale ($m = 3$)	110.34
Trimestrale ($m = 4$)	110.38
Mensile ($m = 12$)	110.47
Settimanale ($m = 52$)	110.51
Giornaliera ($m = 365$)	110.52

Tabella 1.2: Effetto su 100 euro della frequenza con la quale viene calcolato un tasso di interesse del 10% annuo

Se, per esempio, un tasso del 10% annuo viene misurato componendolo trimestralmente ($m = 4$), significa che ogni tre mesi guadagniamo il 2.5% e gli interessi vengono investiti nuovamente. Quindi, un capitale di 100 euro cresce al valore di $100 \times 1.025^4 = 110.38$ euro al termine dell'anno.

Osserviamo che l'espressione 1.3.2, nel limite per $m \rightarrow \infty$, tende al valore Ce^{Rn} , definito come il valore ottenuto componendo il tasso R continuamente. Investendo dunque, per un anno, un capitale C pari a 100 euro al tasso R del 10% per anno continuamente composto, si ottengono

$$100e^{0.1} = 110.52$$

euro. Si noti come questo valore sia, almeno fino alle prime due cifre decimali, lo stesso di quello riportato nella tabella 1.2 nel caso in cui la frequenza di composizione sia giornaliera, ovvero nel caso $m = 365$. Nelle questioni pratiche i due tassi vengono ritenuti equivalenti. Inoltre per R abbastanza piccolo vale l'approssimazione $e^R = 1 + R + o(R)$.

Indichiamo con R_c il tasso continuamente composto e con R_m l'equivalente tasso di interesse, composto però m volte per anno. Dalle equazioni 1.3.1 e 1.3.2 si ha che

$$Ce^{R_c n} = C(1 + R_m)^{mn}$$

da cui si ricava

$$R_c = m \ln \left(1 + \frac{R_m}{m} \right) \quad \text{e} \quad R_m = m(e^{R_c/m} - 1). \quad (1.3.3)$$

Supponiamo, per esempio, che un creditore quoti un tasso di interesse sui prestiti da lui emessi pari all'8% annuo continuamente composto e che venga pagato trimestralmente. Utilizzando l'espressione per R_m appena trovata ponendo $m = 4$ e $R_c = 0.08$ si ottiene che, l'equivalente tasso di interesse composto trimestralmente, è

$$R_4 = 4(e^{0.08/4} - 1) = 0.0808,$$

cioè l'8.08% per anno. Questo vuol dire che per un prestito di 1.000 euro, viene richiesto un pagamento di interessi pari a 20.20 euro ogni tre mesi.

1.4 L'attivo privo di rischio e il tasso a breve

Il concetto di tasso di interesse appartiene alla nostra vita quotidiana ed è diventato ormai qualcosa di familiare. Quando depositiamo una certa somma di denaro in un conto bancario, ognuno si aspetta di vederla crescere (a qualche tasso di interesse) mano mano che il tempo passa. Inoltre è cosa nota che ricevere una certa somma di denaro oggi non è equivalente a ricevere la stessa somma domani.

Procediamo ora ad esprimere questi concetti in termini rigorosi, iniziando col dare la definizione di attivo privo di rischio. Esso rappresenta un investimento senza rischio, dove il profitto cresce continuamente al tasso privo di rischio prelevato dal mercato ad ogni istante.

Definizione 1.4.1 *Definiamo $B(t)$ come il valore dell'attivo privo di rischio (bank account) all'istante $t \geq 0$. Assumiamo che $B(0) = 1$ e che la sua dinamica sia in accordo con la seguente equazione differenziale:*

$$\begin{cases} dB(t) = r(t)B(t) dt \\ B(0) = 1 \end{cases} \quad (1.4.1)$$

dove $r(t)$ è una funzione positiva del tempo.

Come conseguenza si ha che

$$B(t) = \exp \left(\int_0^t r(s) ds \right). \quad (1.4.2)$$

Questa definizione ci dice che, investire un'unità al tempo zero, produce all'istante t il valore dato dalla 1.4.2 e che $r(t)$ è il tasso istantaneo al quale l'attivo privo di rischio cresce. Solitamente, quando ci si riferisce a questo tasso istantaneo, si intende il tasso a breve, noto anche come *short rate* che verrà definito nel paragrafo 1.7. B è importante per mettere in relazione quantità di denaro disponibili a diversi istanti: per avere un euro nel bank account all'istante T , occorre avere

$$D(t, T) := \frac{B(t)}{B(T)} = \exp \left(- \int_t^T r(s) ds \right)$$

euro all'istante $t \leq T$. $D(t, T)$ viene chiamato *fattore di sconto* (discount factor) ed è stocastico, infatti non è conosciuto con certezza all'istante t . Il fattore di sconto $D(t, T)$ fra due istanti t e T (con $t \leq T$) è l'equivalente al tempo t di una unità prelevata all'istante T .

1.5 Obbligazioni senza cedole e i tassi istantanei

Definizione 1.5.1 *Un'obbligazione senza cedole (zero-coupon bond) con data di maturità T , detta brevemente un T -bond, è un contratto che garantisce al possessore di ricevere la somma di 1 unità (per esempio 1 euro) all'istante T , senza pagamenti intermedi. Il prezzo di un T -bond all'istante $t \leq T$ viene indicato con $P(t, T)$.*

Faremo delle ipotesi per garantire che il mercato dei bonds sia sufficientemente ricco:

Ipotesi 1: esistono T -bonds per ogni $T > 0$;

Ipotesi 2: per ogni fissato T , $\{P(t, T) : 0 \leq t \leq T\}$ è un processo stocastico con $P(t, t) = 1$ per ogni t ;

Ipotesi 3: per ogni fissato t , $P(t, T)$ è differenziabile con continuità nella variabile T .

Se siamo all'istante t , un'obbligazione senza cedole con maturità T , è un contratto che stabilisce il valore presente, per esempio di 1 euro, pagato al tempo T , ovvero alla scadenza del contratto. Che relazione c'è fra il fattore di sconto $D(t, T)$ e il prezzo $P(t, T)$ di un T -bond? La differenza principale fra questi due oggetti sta nel fatto che, il primo è un ammontare equivalente di denaro, mentre il secondo è il valore di un contratto.

Se il tasso di interesse a breve r fosse deterministico, tale sarebbe anche il fattore di sconto; inoltre, necessariamente, $D(t, T) = P(t, T)$ per ogni coppia (t, T) . Se il tasso r è stocastico, il fattore di sconto diventa, all'istante t , una quantità aleatoria che dipende dall'evoluzione di r fra gli istanti t e T . Il prezzo $P(t, T)$ di un T -bond, invece, essendo il valore al tempo t di un contratto con maturità T , deve essere conosciuto all'istante t (deterministico).

Covenzioni del conteggio dei giorni (day-count conventions)

L'ammontare del tempo che intercorre fra l'istante presente t e la data di maturità $T > t$ di un qualsiasi contratto, viene misurata in anni e indicata con $T - t$. Se t e T sono due date espresse come giorno/mese/anno, diciamo $t = (g_1, m_1, a_1)$ e $T = (g_2, m_2, a_2)$, non è molto chiaro cosa si intenda per $T - t$. Il mercato misura in modi diversi il tempo che intercorre fra le due date t e T che, in ogni caso, viene rappresentato come una frazione dell'anno. La particolare scelta della misura del tempo che viene fatta è nota come la convenzione del conteggio dei giorni. Daremo qui solo tre esempi di tali convenzioni:

Actual/365: con questa convenzione un anno dura 365 giorni e $T - t$ è dato da

$$\frac{\text{numero attuale di giorni fra } t \text{ e } T}{365}$$

Actual/360: in questo caso un anno ha 360 giorni e $T - t$ è dato da

$$\frac{\text{numero attuale di giorni fra } t \text{ e } T}{360}$$

30/360: secondo questa convenzione i mesi sono composti da 30 giorni e gli anni da 360, così $T - t$ è dato dalla formula

$$\frac{\min(g_2, 30) + (30 - g_1)^+}{360} + \frac{(m_2 - m_1 - 1)^+}{12} + (a_2 - a_1).$$

Quando si estraggono informazioni sui tassi di interesse è importante sapere con quale convenzione si sta lavorando.

È chiaro che, ogni volta che abbiamo bisogno di sapere il valore attuale di un pagamento che avverrà in un istante futuro, il prezzo dell'obbligazione senza cedola con maturità quel preciso istante, è la quantità fondamentale alla quale si fa riferimento. I prezzi delle obbligazioni senza cedola sono quantità basilari nella teoria dei tassi di interesse, infatti, tutti i tassi possono essere definiti in termini di tali prezzi, come mostreremo qui di seguito e nel paragrafo successivo.

Definizione 1.5.2 *Il tasso istantaneo semplicemente composto prelevato all'istante t con maturità T è il tasso costante, al quale, un investimento di $P(t, T)$ unità deve essere fatto, all'istante t , per produrre una somma pari a 1 unità alla maturità T , quando la crescita avviene proporzionalmente al tempo dell'investimento. È noto anche come spot rate semplicemente composto e viene denotato con $F(t, T)$. In formule:*

$$F(t, T) := \frac{1 - P(t, T)}{(T - t)P(t, T)} \Leftrightarrow P(t, T) (1 + F(t, T)(T - t)) = 1.$$

I tassi LIBOR sono semplicemente composti e sono tipicamente legati ai prezzi delle obbligazioni senza cedola dalla convenzione Actual/360 per il conteggio dei giorni. Ovviamente si può esprimere il prezzo $P(t, T)$ in termini di F come

$$P(t, T) = \frac{1}{1 + F(t, T)(T - t)}.$$

Definizione 1.5.3 *Il tasso istantaneo continuamente composto prelevato all'istante t con maturità T è il tasso costante, al quale, un investimento di $P(t, T)$ unità, fatto all'istante t , cresce continuamente fino alla somma di 1 unità alla maturità T . È noto anche come spot rate continuamente composto e viene denotato con $R(t, T)$. In formule:*

$$R(t, T) := -\frac{\ln P(t, T)}{(T - t)} \Leftrightarrow P(t, T)e^{R(t, T)(T - t)} = 1.$$

Anche in questo caso possiamo esprimere i prezzi delle obbligazioni senza cedola in termini di R :

$$P(t, T) = e^{-R(t, T)(T-t)}.$$

Osserviamo che i tassi di interesse finora introdotti, come del resto i prezzi delle obbligazioni, dipendono in modo evidente dalla maturità: è per questo motivo che si parla di *struttura a termine* sia per i tassi d'interesse che per i prezzi dei bonds.

1.6 La curva dei rendimenti

Dai dati dei tassi di interesse prelevati dal mercato ad un certo istante t , si possono ricavare varie curve, dette, appunto, *curve dei rendimenti*. Esse possono presentare una grande varietà di forme. Tipicamente, però, crescono col crescere delle maturità, ciò vuol dire che i tassi di interesse per investimenti a lungo termine sono più alti rispetto a quelli per investimenti a breve.

Nel mercato delle obbligazioni la curva dei rendimenti viene ottenuta dai dati dei bonds, cioè dalle obbligazioni. Nel mercato monetario, invece, viene estratta dai prezzi di vari prodotti, come ad esempio dagli *swaps* (vedi paragrafo 1.8).

Una curva fondamentale è quella dei prezzi delle obbligazioni senza cedola.

Definizione 1.6.1 *La curva dei prezzi delle obbligazioni senza cedola all'istante t è il grafico della funzione*

$$T \mapsto P(t, T) \quad T > t$$

che, vista la positività dei tassi di interesse, è una funzione decrescente rispetto alla variabile T .

Un'altro esempio di curva dei rendimenti è quello mostrato in figura 1.1. Esso rappresenta il grafico della funzione $T \mapsto R(t, T)$ con $T > t$. Come si può notare la curva non è monotona: inizialmente, infatti, decresce. Questo comportamento invertito si presenta periodicamente nel mercato. La curva estratta dal mercato dell'Euro, spesso, ha un andamento monotono, sebbene nel nostro esempio venga presentata la situazione opposta.

In un certo senso, la curva dei prezzi delle obbligazioni senza cedola apporta meno informazioni rispetto alla curva $T \mapsto R(t, T)$: la prima è monotona, mentre la seconda può assumere svariate forme visto che, coinvolgendo una funzione logaritmica, riflette maggiormente le variazioni dei prezzi delle obbligazioni senza cedola.

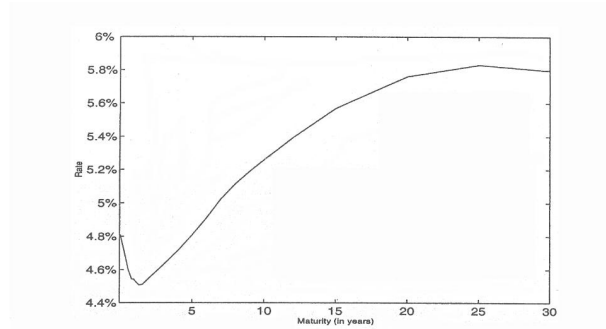


Figura 1.1: Curva dei rendimenti estratta dal mercato EURO il 13/02/2001

1.7 I tassi futuri

Definiremo i tassi futuri (forward rates), che sono caratterizzati da tre istanti, sulla base della seguente costruzione. Supponiamo di essere all'istante t e fissiamo altri due istanti S e T , con $t < S < T$. Vorremmo sottoscrivere un contratto, al tempo t , che ci permetta di avere, facendo un investimento di 1 euro all'istante S , un tasso di ritorno deterministico sull'intervallo $[S, T]$, stabilito all'istante t dal contratto stesso. Questo può essere facilmente realizzato come segue:

1. All'istante t vendiamo un S -bond. Questo ci procura $P(t, S)$ euro con i quali possiamo acquistare esattamente $\frac{P(t, S)}{P(t, T)}$ T -bond. Il nostro investimento netto, all'istante t , è pari a zero.
2. All'istante S siamo obbligati a pagare 1 euro a causa della scadenza dell' S -bond venduto all'istante t .
3. All'istante T , i T -bonds acquistati a t maturano. Riceviamo, così, $\frac{P(t, S)}{P(t, T)} \cdot 1$ euro.

L'effetto netto di questa procedura è l'investimento di 1 euro all'istante S e il guadagno certo di $\frac{P(t, S)}{P(t, T)}$ euro all'istante T . Determiniamo, allora, all'istante t , il tasso di interesse costante $F(t, S, T)$ nel periodo $[S, T]$ come soluzione dell'equazione

$$1 + (T - S)F(t, S, T) = \frac{P(t, S)}{P(t, T)},$$

se pensato come semplicemente composto. Nel caso in cui lo si pensi continuamente composto verrà determinato come soluzione dell'equazione

$$e^{R(t,S,T)(T-S)} = \frac{P(t,S)}{P(t,T)} .$$

Procediamo ora alle definizioni formali.

Definizione 1.7.1 *Il tasso di interesse futuro semplicemente composto prelevato all'istante t per il periodo $[S, T]$ viene denotato con $F(t, S, T)$ ed è definito come*

$$F(t, S, T) := \frac{1}{(T - S)} \left(\frac{P(t, S)}{P(t, T)} - 1 \right) .$$

È noto anche come forward rate semplicemente composto.

Definizione 1.7.2 *Il tasso di interesse futuro continuamente composto prelevato all'istante t per il periodo $[S, T]$ viene denotato con $R(t, S, T)$ ed è definito come*

$$R(t, S, T) := -\frac{\ln P(t, T) - \ln P(t, S)}{(T - S)} .$$

È noto anche come forward rate continuamente composto.

I tassi istantanei $F(t, T)$ e $R(t, T)$, introdotti nel paragrafo 1.5, possono essere ottenuti anche dai due tassi forward appena definiti. Infatti si ha che

$$F(t, T) = F(t, t, T) \text{ e } R(t, T) = R(t, t, T) .$$

Definizione 1.7.3 *Il tasso di interesse futuro istantaneo prelevato all'istante t per la maturità T viene denotato con $f(t, T)$ ed è definito come*

$$f(t, T) := \lim_{T \downarrow S} R(t, S, T) = -\frac{\partial \ln P(t, T)}{\partial T} .$$

È noto anche come forward rate istantaneo.

Osserviamo che, anche in questo caso, si può esprimere il prezzo $P(t, T)$ dell'obbligazione senza cedola in termini del tasso forward istantaneo f . Infatti

$$P(t, T) = P(t, S) \exp \left(- \int_S^T f(t, u) du \right)$$

che, insieme alla richiesta che $P(T, T) = 1$ per ogni T , può essere scritto come

$$P(t, T) = \exp \left(- \int_t^T f(t, u) du \right) .$$

Definizione 1.7.4 *Il tasso di interesse a breve prelevato all'istante t viene denotato con $r(t)$ ed è definito come*

$$r(t) := f(t, t) = \lim_{T \downarrow t} R(t, T) .$$

È noto anche come short rate.

1.8 Obbligazioni con cedole e Swaps

In molti mercati c'è solo un numero ristretto di obbligazioni senza cedola che vengono trattate. Molte obbligazioni includono il pagamento periodico di cedole, i *coupons*. Quelle emesse nei mercati Americani ed Europei, solitamente, prevedono cedole annuali o semestrali.

Definizione 1.8.1 *Un'obbligazione con cedole fisse è un contratto che assicura al possessore il pagamento, a stabilite date future $T_1 < \dots < T_n$, di cedole dal valore fissato c_1, \dots, c_n .*

Un'obbligazione con cedole fisse, dunque, è caratterizzata da:

- n date future $T_1 < \dots < T_n$ (T_n è la data di maturità dell'obbligazione);
- una serie di cedole (deterministiche) c_1, \dots, c_n (i coupons);
- il valore nominale N dell'obbligazione.

Il possessore dell'obbligazione, all'istante T_i , riceve il pagamento c_i per ogni $i = 1, \dots, n$ e l'ultima cedola comprende anche il rimborso del valore nominale N dell'obbligazione.

Poniamo $T_{i+1} - T_i = \delta$ per comodità, anche se questa ipotesi può essere facilmente eliminata. I coupons sono dati da una percentuale del valore nominale, cioè $c_i = K\delta N$, per $i = 1, \dots, n-1$ e per qualche fissato tasso di interesse K , e $c_n = K\delta N + N$.

Il valore $p(t)$ all'istante $t \leq T_1$ dell'obbligazione con cedole fisse è dato dalla somma dei valori attualizzati delle cedole:

$$p(t) = \sum_{i=1}^n P(t, T_i) c_i = \left(K\delta \sum_{i=1}^n P(t, T_i) + P(t, T_n) \right) N$$

Definizione 1.8.2 *Un floating rate note è un contratto che assicura al possessore il pagamento, a stabilite date future $T_0 < \dots < T_n$, di cedole dal valore c_i stabilito all'istante T_{i-1} , per $i = 1 \dots n$.*

Un floating rate note, dunque, è caratterizzato da:

- $n+1$ date future $T_0 < \dots < T_n$ (T_n è la data di maturità del floating rate note);
- una serie di cedole c_1, \dots, c_n (i coupons);
- il valore nominale N del floating rate note.

I pagamenti delle cedole fisse sono ora sostituiti da

$$c_i = (T_i - T_{i-1})F(T_{i-1}, T_i)N,$$

dove $F(T_{i-1}, T_i)$ è un certo tasso di interesse prelevato dal mercato, solitamente il tasso EURIBOR.

Osserviamo che, anche se il pagamento c_i avviene all'istante T_i , il tasso $F(T_{i-1}, T_i)$ è già determinato all'istante T_{i-1} (per questo motivo è stato aggiunto l'istante T_0 alle date T_1, \dots, T_n).

Calcoliamo ora il valore $p(t)$ di questo contratto all'istante $t \leq T_0$. Usando la definizione 1.5.2 di $F(T_{i-1}, T_i)$ otteniamo

$$c_i = \left(\frac{1}{P(T_{i-1}, T_i)} - 1 \right) N.$$

Il valore di c_i all'istante t , ovvero il suo valore attualizzato, è ottenuto moltiplicando c_i per il fattore di sconto $P(t, T)$:

$$P(t, T_i)c_i = NP(t, T_i) \left(\frac{1}{P(T_{i-1}, T_i)} - 1 \right) = N(P(t, T_{i-1}) - P(t, T_i)).$$

Nell'ultima eguaglianza si è usato che

$$\frac{P(t, T_i)}{P(T_{i-1}, T_i)} = P(t, T_{i-1}), \quad (1.8.1)$$

ovvero che $P(t, T_{i-1})$ è il valore di $1/P(T_{i-1}, T_i)$ all'istante t . Infatti:

1. All'istante t compriamo un T_{i-1} -bond al costo di $P(t, T_{i-1})$.
2. All'istante T_{i-1} riceviamo una unità con la quale acquistiamo esattamente $1/P(T_{i-1}, T_i)$ T_i -bonds. A questo istante il nostro investimento netto è zero.
3. All'istante T_i riceviamo $1/P(T_{i-1}, T_i)$ unità a causa della maturità dei T_i -bonds acquistati all'istante precedente.

Abbiamo, quindi, investito all'istante t la somma $P(t, T_{i-1})$ per ottenere all'istante T_i la somma $1/P(T_{i-1}, T_i)$.

Per calcolare il valore $p(t)$ del contratto dobbiamo sommare ai valori all'istante t delle cedole, il valore nominale N attualizzato, ottenendo la semplicissima formula

$$p(t) = \sum_{i=1}^n N(P(t, T_{i-1}) - P(t, T_i)) + P(t, T_n)N = P(t, T_0)N$$

Ciò significa che questo contratto equivale a N unità all'istante T_0 . In particolare se $t = T_0$ si ha che $p(T_0) = N$.

1.8.1 Interest rate swaps

Un interest rate swap è un contratto che permette di scambiare un pagamento ad un certo tasso fisso con uno a tasso variabile (tipicamente al tasso EURIBOR). Ci sono diversi tipi di interest rate swaps.

Un *payer interest rate swap* è caratterizzato da:

- $n + 1$ date future $T_0 < \dots < T_n$ con $T_i - T_{i-1} \equiv \delta$ (T_n è la data di maturità dello swap);
- il tasso di interesse fisso K ;
- il valore nominale N dello swap.

Come in precedenza l'ipotesi di equidistanza delle date è solo di convenienza e può essere facilmente eliminata.

I pagamenti avvengono solo alle date T_1, \dots, T_n . All'istante T_i , per $i = 1, \dots, n$, il possessore del contratto paga $K\delta N$ e riceve $F(T_{i-1}, T_i)\delta N$, cioè al netto riceve $c_i = (F(T_{i-1}, T_i) - K)\delta N$. Per calcolare il valore di tale quantità all'istante t , utilizziamo la definizione 1.5.2 di $F(T_{i-1}, T_i)$ e l'eguaglianza 1.8.1, ottenendo

$$\begin{aligned} P(t, T_i)c_i &= P(t, T_i) \left(\frac{1 - P(T_{i-1}, T_i)}{\delta P(T_{i-1}, T_i)} - K \right) \delta N \\ &= P(t, T_i) \left(\frac{1}{P(T_{i-1}, T_i)} - 1 - \delta K \right) N \\ &= (P(t, T_{i-1}) - P(t, T_i) - P(t, T_i)\delta K)N. \end{aligned} \tag{1.8.2}$$

Il valore $\Pi_p(t)$ dello swap all'istante $t \leq T_0$ è dunque ottenuto sommando i valori attualizzati dei singoli pagamenti appena trovati:

$$\Pi_p(t) = N \left(P(t, T_0) - P(t, T_n) - K\delta \sum_{i=1}^n P(t, T_i) \right).$$

Un *receiver interest rate swap* è, invece, un contratto che permette al possessore di scambiare un pagamento a tasso variabile con uno a tasso fisso. Il suo valore $\Pi_r(t)$ all'istante $t \leq T_0$ è, ovviamente, dato da

$$\Pi_r(t) = -\Pi_p(t).$$

La questione, ora, è come determinare il tasso K . Il tasso swap futuro $R_{swap}(t)$ all'istante $t \leq T_0$ è il tasso fisso K tale che $\Pi_p(t) = \Pi_r(t) = 0$, ovvero tale che

$$N \left(P(t, T_0) - P(t, T_n) - K\delta \sum_{i=1}^n P(t, T_i) \right) = 0$$

da cui

$$R_{swap}(t) = \frac{P(t, T_0) - P(t, T_n)}{\delta \sum_{i=1}^n P(t, T_i)} .$$

Possiamo dare una rappresentazione alternativa per $R_{swap}(t)$, talvolta più comoda. Sfruttando la definizione 1.7.1 da cui si ricava $P(t, T_{i-1}) - P(t, T_i) = F(t, T_{i-1}, T_i) \delta P(t, T_i)$, possiamo riscrivere la 1.8.2 come

$$N \delta P(t, T_i) (F(t, T_{i-1}, T_i) - K).$$

Sommando otteniamo

$$\Pi_p(t) = N \delta \sum_{i=1}^n P(t, T_i) (F(t, T_{i-1}, T_i) - K)$$

da cui si ricava $R_{swap}(t)$ come media pesata dei tassi futuri semplicemente composti, risolvendo rispetto a K l'equazione $\Pi_p(t) = 0$. Dunque

$$R_{swap}(t) = \sum_{i=1}^n w_i(t) F(t, T_{i-1}, T_i)$$

dove i pesi sono dati da

$$w_i(t) = \frac{P(t, T_i)}{\sum_{j=1}^n P(t, T_j)} .$$

1.9 Rendimenti delle obbligazioni

Per un'obbligazione $P(t, T)$ senza cedole il rendimento (o zero-rate) è semplicemente il tasso istantaneo $R(t, T)$ continuamente composto, cioè

$$P(t, T) = e^{-R(t, T)(T-t)} .$$

Come abbiamo già detto, la maggior parte delle obbligazioni prevede il pagamento di cedole intermedie.

Sia $p(t)$ il prezzo all'istante t di una tale obbligazione con: date di pagamento $T_1 < \dots < T_n$, cedole fisse c_1, \dots, c_n e valore nominale N come nella definizione 1.8.1. Ricordiamo che il suo prezzo è

$$p(t) = \sum_{i=1}^n P(t, T_i) c_i , \quad t \leq T_1 .$$

Tale prezzo, da solo, non può determinare lo zero-rate, in quanto parte del rendimento viene realizzato sotto forma di cedole precedenti la maturità dell'obbligazione. Si parla, allora, di *rendimento interno* dell'obbligazione. Si

tratta cioè del tasso costante $y(t)$ (sul periodo $[t, T_n]$) continuamente composto che genera il prezzo di mercato dell'obbligazione, ovvero l'unica soluzione dell'equazione

$$p(t) = \sum_{i=1}^n c_i e^{y(t)(T-t)}.$$

Ad esempio, consideriamo il caso in cui gli zero-rates siano quelli riportati nella tabella 1.3.

<i>Maturità (anni)</i>	<i>Zero-rate (%)</i>
0.5	5.0
1.0	5.8
1.5	6.4
2.0	6.8

Tabella 1.3: Zero-rates

Supponiamo che, un'obbligazione che matura fra 2 anni con un valore nominale di 100 euro, preveda il pagamento di cedole semestrali al tasso del 6% per anno. Il prezzo dell'obbligazione all'istante $t = T_1$ è dato da

$$p(t) = 3e^{-0.05 \times 0.5} + 3e^{-0.058 \times 1.0} + 3e^{-0.064 \times 1.5} + 103e^{-0.068 \times 2.0} = 98.39 \text{ euro.}$$

Il suo rendimento interno è dato dalla soluzione dell'equazione

$$3e^{-y \times 0.5} + 3e^{-y \times 1.0} + 3e^{-y \times 1.5} + 103e^{-y \times 2.0} = 98.39$$

Questa equazione non lineare del tipo $f(y) = 0$, può essere risolta con vari metodi, fra i quali ricordiamo brevemente quello di Newton-Raphson: si parte da una stima y_0 della soluzione e si producono stime successive y_i sempre migliori, utilizzando la formula $y_{i+1} = y_i - f(y_i)/f'(y_i)$.

Risolvendo la nostra equazione si trova $y = 6.76\%$.

Passiamo ora a illustrare come calcolare gli zero-rates utilizzando le obbligazioni con cedole fisse. L'approccio più conosciuto è il così detto *metodo bootstrap*.

Consideriamo i prezzi di 5 diverse obbligazioni riportati nella tabella 1.4.

Poiché le prime tre obbligazioni non prevedono il pagamento di cedole, gli zero-rates corrispondenti alle maturità di tali obbligazioni possono essere facilmente calcolati. L'obbligazione che matura in 3 mesi prevede un rendimento di 2.5 euro su un investimento iniziale di 97.5 euro. Quindi, lo zero-rate (trimestralmente composto) a 3 mesi, è pari al $(4 \times 2.5)/97.5 = 10.256\%$

<i>Valore nominale (euro)</i>	<i>Maturità (anni)</i>	<i>Valore della cedola*</i>	<i>Prezzo dell'obbl.</i>
100	0.25	0	97.5
100	0.50	0	94.9
100	1.00	0	90.0
100	1.50	8	96.0
100	2.00	12	101.6

Tabella 1.4: Dati per il metodo bootstrap

* Metà della cedola viene pagata ogni 6 mesi.

per anno. Utilizzando la formula 1.3.3 per calcolare il corrispondente tasso continuamente composto, otteniamo

$$4 \ln \left(1 + \frac{0.10256}{4} \right) = 0.10127,$$

ovvero il 10.127% per anno. Ragionando alla stessa maniera, calcoliamo gli zero-rate (continuamente composti) a 6 e 12 mesi, che sono rispettivamente pari al 10.469% e al 10.536% per anno. La quarta obbligazione matura in 1.5 anni: dopo 6 mesi viene pagata una cedola pari a 4 euro, così come al termine del primo anno, mentre alla maturità viene pagata una cedola di 104 euro.

Denotiamo con $R_{1.5}$ lo zero-rate continuamente composto a 1.5 anni. Per calcolarlo basta risolvere la facile equazione seguente:

$$4e^{-0.10469 \times 0.5} + 4e^{-0.10536 \times 1.0} + 104e^{-R_{1.5} \times 1.5} = 96.$$

Svolgendo i calcoli otteniamo $e^{-1.5R_{1.5}} = 0.85196$, da cui

$$R_{1.5} = -\frac{\ln(0.85196)}{1.5} = 0.10681.$$

Analogamente si può calcolare lo zero-rate R_2 continuamente composto a 2 anni, usando, oltre ai tassi a 0.5, 1, 1.5 anni appena calcolati, anche le informazioni sulla quinta obbligazione riportata nella tabella. Dall'equazione

$$6e^{-0.10469 \times 0.5} + 6e^{-0.10536 \times 1.0} + 6e^{-0.10681 \times 1.5} + 106e^{-R_2 \times 2.0} = 101.6$$

si ricava $R_2 = 0.10808$, ovvero 10.808%.

Osserviamo, però, che nella pratica non si hanno obbligazioni per qualsiasi maturità. L'approccio più usato è quello di interpolare i prezzi delle obbligazioni prelevati dal mercato. Per esempio, supponiamo di sapere che un'obbligazione che matura in 2.3 anni con cedole del 6% viene venduta a 98 euro e, invece, una che matura in 2.7 anni con cedole del 6.5%, viene venduta a 99 euro. Si può allora assumere che, un'obbligazione che matura in 2.5 anni con cedole del 6.5%, venga venduta a 98.5 euro.

1.10 Caps e Floors

Un *caplet* con data di inizio T e data di saldo $T + \delta$ paga al possessore la differenza fra un semplice tasso $F(T, T + \delta)$ prelevato dal mercato (per esempio il tasso EURIBOR) e un fissato tasso K . All'istante $T + \delta$ verrà pagata la quantità

$$\delta (F(T, T + \delta) - K)^+. \quad (1.10.1)$$

Un *cap* è un contratto formato da una serie di caplets. Può essere visto come un payer swap in cui lo scambio di pagamento avviene solo se ha un valore positivo.

Un cap è caratterizzato da:

- $n + 1$ date future $T_0 < \dots < T_n$ con $T_i - T_{i-1} \equiv \delta$ (T_n è la data di maturità del cap);
- il tasso di interesse fisso K .

I pagamenti avvengono alle date T_1, \dots, T_n . All'istante T_i il possessore del cap riceve

$$\delta (F(T_{i-1}, T_i) - K)^+.$$

Sia $t \leq T_0$. Con

$$Cpl(t, i) \quad i = 1, \dots, n$$

indicheremo il prezzo all'istante t dell' i -esimo caplet con data di inizio T_{i-1} e data di saldo T_i e con

$$Cp(t) = \sum_{i=1}^n Cpl(t, i)$$

indicheremo il prezzo all'istante t del cap.

Un cap protegge il possessore da tassi di interesse troppo crescenti. Garantisce, infatti, che il tasso di interesse pagato per un prestito a tasso variabile non superi il predeterminato tasso K . Per esempio, supponiamo che una compagnia abbia da pagare un debito di una unità al tasso EURIBOR e che abbia il timore che nel futuro il tasso EURIBOR diventi troppo alto. Tale compagnia vorrebbe tutelarsi, bloccando il pagamento ad un certo tasso fissato K . Entra quindi in un cap: paga il tasso EURIBOR L e riceve $(L - K)^+$ dal contratto del cap. Così facendo, la compagnia paga un netto di

$$L - (L - K)^+ = \min(L, K),$$

ovvero al massimo paga K .

Un *floor* è l'inverso di un cap. Protegge, dunque, il possessore da tassi di interesse troppo bassi. Un floor è un contratto formato da una serie di

floorlets, il cui pagamento è, con le stesse notazioni usate sopra, all'istante T_i

$$\delta(K - F(T_{i-1}, T_i))^+.$$

Scriveremo

$$Fl(t, i) \quad i = 1, \dots, n$$

per indicare il prezzo all'istante t dell' i -esimo floorlet e

$$Fl(t) = \sum_{i=1}^n Fl(t, i)$$

per indicare il prezzo all'istante t del floor.

Caps e floors sono strettamente legati agli swaps. Infatti, grazie all'eguaglianza $(L - K)^+ - (K - L)^+ = L - K$, si può facilmente verificare che

$$Cp(t) - Fl(t) = \Pi_p(t),$$

dove $\Pi_p(t)$ è il valore all'istante t di un payer swap con tasso fissato K , valore nominale pari a 1 e stesse date di pagamento del cap e del floor.

Poniamo $t = 0$. Il cap (floor) è detto *at-the-money* (ATM) se

$$K = R_{swap}(0) = \frac{P(0, T_0) - P(0, T_n)}{\delta \sum_{i=1}^n P(0, T_i)}.$$

È detto, invece, *in-the-money* (ITM) se $K < R_{swap}(0)$ ($K > R_{swap}(0)$) ed è detto *out-of-the-money* (OTM) se $K > R_{swap}(0)$ ($K < R_{swap}(0)$).

Osserviamo che, se abbiamo una sola data T_1 di pagamento, il cap collassa ad un caplet con data di inizio T_0 e data di saldo T_1 . In questo caso particolare

$$R_{swap}(0) = F(0, T_0, T_1) = \frac{1}{\delta} \left(\frac{P(0, T_0) - P(0, T_1)}{P(0, T_1)} \right).$$

Quindi il caplet risulta ATM se $K = F(0, T_0, T_1)$ e ITM se $K < F(0, T_0, T_1)$.

1.11 Swaptions

Introduciamo ora l'ultimo derivato dei tassi di interesse che tratteremo, chiamato *swaptions*. Ci sono due tipi principali di swaptions, la versione *payer* e quella *receiver*. Un payer (receiver) swaption europeo con tasso fisso K è un'opzione che dà il diritto (e non l'obbligo) al possessore di entrare in un payer (receiver) swap con tasso fisso K ad una certa data futura, detta maturità dello swaption. Solitamente, tale data coincide con la data di inizio dello swap sottostante. La lunghezza $T_n - T_0$ dello swap è chiamata *tenore* dello swaption.

Richiamiamo l'espressione del valore di un payer swap con tasso fisso K all'istante di inizio T_0

$$\Pi_p(T_0, K) = N \sum_{i=1}^n P(T_0, T_i) \delta(F(T_0, T_{i-1}, T_i) - K).$$

Quindi il valore dello swaption con tasso fisso K alla maturità T_0 è

$$N \left(\sum_{i=1}^n P(T_0, T_i) \delta(F(T_0, T_{i-1}, T_i) - K) \right)^+.$$

Contrariamente al caso del cap, questa quantità non può essere decomposta in prodotti elementari, come i caplets ed è questa la principale differenza fra i caps/floors e gli swaptions. Infatti, in generale vale la disuguaglianza (e non l'uguaglianza) seguente

$$\left(\sum_{i=1}^n P(T_0, T_i) \delta(F(T_0, T_{i-1}, T_i) - K) \right)^+ \leq \sum_{i=1}^n P(T_0, T_i) \delta(F(T_0, T_{i-1}, T_i) - K)^+.$$

In particolare, questo, ci dice che il valore di un payer swaption è sempre minore del valore del corrispondente contratto cap.

Capitolo 2

La modellizzazione matematica dei mercati finanziari

Un modello per un mercato finanziario con $k + 1$ *attivi finanziari* (azioni, bonds, opzioni, ...) può essere rappresentato tramite uno spazio probabilizzato (Ω, \mathcal{F}, P) sul quale vengono definiti una filtrazione $\underline{\mathcal{F}} = (\mathcal{F}_t)_{t \in \mathbb{I}}$, che supponiamo verificare le condizioni abituali, e un processo stocastico $S = (S_0, \dots, S_k)$, che supponiamo essere una semimartingala continua. Tale ipotesi viene fatta per rendere S un buon processo integratore, ovvero per dare senso a scritture del tipo $\int_0^t H(s) dS(s)$. $S_i(t)$ è interpretato come il prezzo, all'istante t , di una unità dell' i -esimo attivo finanziario. L'attivo S_0 ha un ruolo particolare, quello di *numerario*: dividendo ogni prezzo per la quantità $S_0(t)$, si definiscono le unità dei restanti attivi attualizzati rispetto a quello scelto come numerario. Dunque lo assumiamo positivo, ovvero $S_0(t) > 0$ P -q.c. Come convenzione, inoltre, assumiamo che $S_0(0) = 1$. In generale, come numerario, può essere scelto un qualunque attivo, come per esempio un bond a 3 anni, ma, per convenienza, in molti casi, viene scelto l'attivo privo di rischio B definito come nel paragrafo 1.4.

Esempio. Il più semplice e non banale esempio di mercato è rappresentato dal modello Black-Scholes. In questo modello c'è un tasso di interesse deterministico r e un attivo stocastico S_1 . Denotando S_0 con B e S_1 con S , il modello è così definito:

$$\begin{cases} dB(t) = rB(t) dt \\ dS(t) = \alpha S(t) dt + \sigma d\tilde{W}(t) \end{cases}$$

dove \tilde{W} è un processo di Wiener sotto P e la filtrazione è quella generata da \tilde{W} . Il tasso di interesse r , il tasso di rendimento locale α e la volatilità σ sono, in questo modello, costanti deterministiche. Osserviamo che il tasso di rendimento di B , dato da

$$\frac{dB(t)}{B(t)} = r dt,$$

è deterministico. Al contrario, quello di S dato da

$$\frac{dS(t)}{S(t)} = \alpha dt + \sigma d\tilde{W},$$

è stocastico.

2.1 Portafogli autofinanziati

Definizione 2.1.1 *Una strategia è un qualsiasi processo vettoriale prevedibile e limitato*

$$h(t) = [h_0(t), H_1(t), \dots, H_k(t)] \in \mathbb{R}^{k+1}.$$

Definizione 2.1.2 *Il valore all'istante t del portafoglio corrispondente alla strategia h è definito come*

$$V(t; h) = h_0(t)S_0(t) + \sum_{i=1}^k H_i(t)S_i(t).$$

Definizione 2.1.3 *Una strategia h (per cui abbiano senso le espressioni sotto il segno di integrale) si dice autofinanziata se*

$$dV(t; h) = h_0(t) dS_0(t) + \sum_{i=1}^k H_i(t) dS_i(t),$$

oppure

$$V(t; h) = V(0; h) + \int_0^t h_0(s) dS_0(s) + \sum_{i=1}^k \int_0^t H_i(s) dS_i(s).$$

Osserviamo che la seconda espressione è semplicemente la versione integrale della prima.

L'interpretazione delle prime due definizioni è semplice: $H_i(t)$ sono le unità dell' i -esimo attivo che abbiamo nel portafoglio all'istante t , mentre $h_0(t)$ sono le unità di numerario (è stata usata la notazione $h = (h_0, H_1, \dots, H_k)$ proprio per rendere evidente la scelta del numerario). $V(t; h)$ è il valore complessivo del portafoglio. La terza definizione è intesa a formalizzare l'idea intuitiva di una strategia in cui l'acquisto di un nuovo attivo viene finanziato solo dalla vendita di attivi già presenti nel portafoglio. Infatti, la variazione istantanea del valore del portafoglio è data dalla somma delle variazioni dei prezzi dei soli attivi S_0, \dots, S_k moltiplicati per il numero di attivi posseduti in quell'istante.

Ci sarà utile in seguito esprimere il valore del portafoglio corrispondente ad una data strategia tramite i *relativi pesi* ($u_0(t), \dots, u_k(t)$), dove $u_i(t)$ è definito come la proporzione del valore del portafoglio che, all'istante t , viene investita nell' i -esimo attivo.

Definizione 2.1.4 Consideriamo una strategia $h = [h_0(t), H_1(t), \dots, H_k(t)]$. I relativi pesi sono definiti come

$$u_0(t) = \frac{h_0(t)S_0(t)}{V(t; h)}, \quad u_i(t) = \frac{H_i(t)S_i(t)}{V(t; h)} \quad i = 1, \dots, k$$

Scelto $S_0(t)$ come numerario, diamo la seguente definizione:

Definizione 2.1.5 Il processo vettoriale dei prezzi attualizzati $\tilde{S} = (\tilde{S}_0(t), \dots, \tilde{S}_k(t))$ è definito da

$$\tilde{S}(t) = \frac{S(t)}{S_0(t)}.$$

Definizione 2.1.6 Il valore attualizzato del portafoglio corrispondente alla strategia h è definito come

$$\tilde{V}(t; h) = \frac{V(t; h)}{S_0(t)}.$$

Può essere utile la seguente caratterizzazione delle strategie autofinanziate:

Proposizione 2.1.7 Una strategia h è autofinanziata se e solo se

$$d\tilde{V}(t; h) = \sum_{i=1}^k H_i(t) d\tilde{S}_i(t),$$

oppure

$$\tilde{V}(t; h) = \tilde{V}(0; h) + \sum_{i=1}^k \int_0^t H_i(s) d\tilde{S}_i(s).$$

Dimostrazione.

Per rendere più leggera la notazione, omettiamo la dipendenza sia dal tempo che dalla strategia h .

Osserviamo che, posto $\tilde{X} = \frac{X}{S_0}$ vale

$$d\tilde{X} = \frac{dX}{S_0} - \tilde{X} \frac{dS_0}{S_0}. \quad (2.1.1)$$

Supponiamo che la strategia h sia autofinanziata. Applicando la formula

2.1.1 con $\tilde{X} = \tilde{V}$ e svolgendo i calcoli si ottiene:

$$\begin{aligned}
d\tilde{V} &= \frac{dV}{S_0} - \tilde{V} \frac{dS_0}{S_0} \\
&= h_0 \frac{dS_0}{S_0} + \sum_{i=1}^k H_i \frac{dS_i}{S_0} - \tilde{V} \frac{dS_0}{S_0} \\
&= \left(h_0 - \frac{V}{S_0} \right) \frac{dS_0}{S_0} + \sum_{i=1}^k H_i \left(d\tilde{S}_i + \tilde{S}_i \frac{dS_0}{S_0} \right) \\
&= \left[h_0 - \left(h_0 + \sum_{i=1}^k H_i \tilde{S}_i \right) + \sum_{i=1}^k H_i \tilde{S}_i \right] \frac{dS_0}{S_0} + \sum_{i=1}^k H_i d\tilde{S}_i \\
&= \sum_{i=1}^k H_i d\tilde{S}_i
\end{aligned}$$

Viceversa, supponiamo che valga $d\tilde{V} = \sum_{i=1}^k H_i d\tilde{S}_i$ e dimostriamo che la strategia è autofinanziata:

$$\begin{aligned}
dV &= S_0 d\tilde{V} + \tilde{V} dS_0 \\
&= S_0 \sum_{i=1}^k H_i d\tilde{S}_i + \tilde{V} dS_0 \\
&= S_0 \sum_{i=1}^k H_i \left(\frac{dS_i}{S_0} - \tilde{S}_i \frac{dS_0}{S_0} \right) + \frac{V}{S_0} dS_0 \\
&= \sum_{i=1}^k H_i dS_i - \sum_{i=1}^k H_i \tilde{S}_i dS_0 + \left(h_0 + \sum_{i=1}^k H_i \tilde{S}_i \right) dS_0 \\
&= h_0 dS_0 + \sum_{i=1}^k H_i dS_i \quad \square
\end{aligned}$$

2.2 Attivi casuali, arbitraggio e misure martingala

Elementi fondamentali della teoria dell'arbitraggio sono i derivati finanziari. Essi sono attivi che, in qualche modo, vengono definiti usando gli attivi sottostanti S_0, \dots, S_k . Tipici esempi di derivati sono le opzioni, i caps, i floors e gli interest rate swaps. La formalizzazione matematica di un derivato è molto semplice.

Definizione 2.2.1 *Un attivo casuale con maturità T (o contingent claim) è una qualsiasi variabile aleatoria $X \in L^0(\mathcal{F}_T, P)$, cioè un'arbitraria variabile aleatoria \mathcal{F}_T -misurabile.*

Un attivo casuale è un contratto che specifica l'ammontare stocastico X di denaro che dovrà essere pagato al possessore all'istante T .

Esempio. Nel modello Black-Scholes il più famoso esempio di attivo casuale è quello di un'opzione call Europea su S con prezzo strike K e data di maturità T . Questo contratto è definito da

$$X = \max[S(T) - K, 0]$$

e dà il diritto, e non l'obbligo, di comprare all'istante T una unità dell'attivo S al prezzo prestabilito K .

Definizione 2.2.2 *Una strategia autofinanziata h si dice un arbitraggio se il valore del portafoglio corrispondente verifica le seguenti condizioni:*

1. $V(0; h) = 0$;
2. $V(T; h) \geq 0$ P -q.c.;
3. $P\{V(T; h) > 0\} > 0$.

Se non esistono strategie arbitraggio per ogni $T \in \mathbb{I}$ ($T < +\infty$), allora diremo che il mercato è privo di arbitraggio.

In pratica l'arbitraggio è equivalente ad investire zero oggi e ricevere domani una quantità di denaro non negativa e positiva con probabilità positiva. Il mercato, talvolta, offre possibilità di arbitraggio. Visto però che tutti gli investitori cercano di sfruttarle essendo molto vantaggiose, si intuisce che il guadagno tende a diminuire sempre più, fino a sparire. Il mercato, quindi, tende, in modo naturale, ad eliminare tali possibilità. Per questo è naturale richiedere che nei nostri modelli non ci siano possibilità di arbitraggio.

Un tipico esempio di un potenziale arbitraggio si ha quando, oltre all'usuale attivo privo di rischio B , ne è presente un altro:

Proposizione 2.2.3 *Supponiamo che esista un processo U che abbia la seguente dinamica:*

$$dU(t) = k(t)U(t) dt,$$

dove k è un qualche processo misurabile e adattato. Se il mercato è privo di arbitraggio, allora necessariamente

$$k = r \quad dt \otimes dP - q.c.$$

dove con r si denota il tasso a breve.

Dimostrazione.

Consideriamo il processo attualizzato

$$\tilde{U}(t) = \frac{U(t)}{B(t)} = U(0) \exp \int_0^t (k(s) - r(s)) ds,$$

da cui si ricava

$$d\tilde{U} = \frac{dU}{B} - \tilde{U} \frac{dB}{B} = k \frac{U}{B} - \tilde{U} r = (k - r)\tilde{U}$$

Allora, ponendo $h(t) = I_{\{k(t) > r(t)\}}$, è facile mostrare che si ottiene una strategia autofinanziata il cui corrispondente valore attualizzato del portafoglio \tilde{V} è dato da

$$\tilde{V}(t; h) = \int_0^t h(s) d\tilde{U}(s) = \int_0^t \left(I_{\{k(s) > r(s)\}} (k(s) - r(s)) \tilde{U}(s) \right) ds \geq 0.$$

L'assenza di arbitraggio richiede che, posto

$$\mathcal{N} := \{(t, \omega) \mid k(t, \omega) > r(t, \omega)\} \subseteq [0, T] \times \Omega$$

(che è misurabile), si abbia

$$0 = E[\tilde{V}(T)] = \int_{\mathcal{N}} \underbrace{\left(I_{\{k(t, \omega) > r(t, \omega)\}} (k(t, \omega) - r(t, \omega)) \right)}_{>0 \text{ su } \mathcal{N}} dt \otimes dP.$$

Ciò accade solo se \mathcal{N} ha misura nulla.

Si conclude la dimostrazione usando la stessa argomentazione, ma cambiando i segni. \square

Investighiamo ora quando un dato modello è privo di arbitraggio e lo strumento fondamentale che si usa è la misura martingala.

Definizione 2.2.4 Diremo che una misura di probabilità Q sullo spazio (Ω, \mathcal{F}) è una misura martingala se:

1. $Q \sim P$, cioè $Q(A) = 0 \Leftrightarrow P(A) = 0$ per ogni $A \in \mathcal{F}$;
2. Il processo dei prezzi attualizzati \tilde{S} è una martingala sotto Q .

L'insieme delle misure martingale viene denotato con \mathcal{P} .

Che relazione intercorre fra un modello privo di arbitraggio e l'esistenza di una misura martingala? Nel caso in cui l'insieme dei tempi \mathbb{I} sia discreto e finito, le due condizioni si equivalgono grazie al *primo teorema fondamentale dell'asset pricing*:

Teorema 2.2.5 (Dalang, Morton e Willinger) Sia $\mathbb{I} \subset \mathbb{N}$ finito. Un modello è privo di arbitraggio se e solo se esiste una misura martingala.

J. M. Harrison e S. Pliska nel 1981 dimostrarono la stessa cosa nell'ipotesi più generale che \mathbb{I} fosse discreto (cioè \mathbb{I} sottoinsieme di \mathbb{N}) ma, in compenso, richiedevano che Ω fosse uno spazio finito. Nel caso a tempi continui, le cose si complicano: è sempre vero che l'esistenza di una misura martingala esclude arbitraggi ma, per dimostrare il viceversa, F. Delbaen e W. Schachermayer hanno sostituito la condizione di non arbitraggio con quella di *no free lunch with vanishing risk*. Da ora in poi, quando parleremo di assenza di arbitraggio, intenderemo l'esistenza di una misura martingala.

2.3 Completezza e assenza di arbitraggio

Definizione 2.3.1 1. Un attivo casuale X con maturità T si dice replicabile se esiste una strategia autofinanziata h tale che il corrispondente valore del portafoglio verifica

$$V(T; h) = X \quad P\text{-}q.c.$$

V verrà chiamato portafoglio replicante.

2. Il mercato si dice completo se ogni attivo casuale è replicabile.

Osserviamo che, da un punto di vista puramente finanziario, il portafoglio replicante è indistinguibile dall'attivo X .

La nozione di completezza e quella di arbitraggio sono strettamente legate dal seguente teorema, noto come *secondo teorema fondamentale dell'asset pricing*:

Teorema 2.3.2 (Dalang, Morton e Willinger) Sia $\mathbb{I} \subset \mathbb{N}$ finito. Un modello privo di arbitraggio è completo se e solo se la misura martingala è unica.

Per una dimostrazione si può consultare [10].

Cerchiamo di dare alcuni criteri per stabilire velocemente se un dato modello è completo e/o privo di arbitraggio. L'argomentazione sarà puramente euristica.

Supponiamo che il processo dei prezzi S dei $k + 1$ attivi sottostanti sia governato da r fonti aleatorie. Non daremo qui una definizione precisa di fonte aleatoria ma, un tipico esempio, è un processo di Wiener. In particolare, nei modelli basati sul processo di Wiener, il numero delle fonti aleatorie è la dimensione di tale processo.

Ogni nuovo attivo che aggiungiamo al modello (senza incrementare r) fa aumentare la potenziale possibilità di creare una strategia di arbitraggio. Per avere allora un modello privo di arbitraggio, il numero k degli attivi sottostanti (escluso quello preso come numerario) deve essere minore del numero r delle fonti aleatorie. D'altra parte, però, ogni nuovo attivo aggiunto ci procura nuove possibilità di replicare un dato attivo casuale. La completezza, quindi, richiede che il numero degli attivi k sia maggiore di r . Riassumiamo tutto nel seguente teorema, che può essere considerato come un *Meta-Teorema*, di cui però non daremo una formulazione rigorosa. Comunque, in casi concreti, può esserne data una dimostrazione precisa.

Teorema 2.3.3 Sia k il numero degli attivi presenti nel modello (escluso l'attivo preso come numerario) e sia r il numero delle fonti aleatorie. Allora:

1. Il modello è privo di arbitraggio se e solo se $k \leq r$;

2. Il modello è completo se e solo se $k \geq r$;
3. Il modello è completo e privo di arbitraggio se e solo se $k = r$.

Esempio. Riprendiamo l'esempio del modello Black-Scholes descritto a pagina 19. Vi erano solo due attivi sottostanti, S e l'attivo privo di rischio B . La fonte aleatoria, in questo caso, è una sola, rappresentata dal processo di Wiener \tilde{W} . Dunque $k = r = 1$. Usando il teorema 2.3.3, ci si aspetta che il modello risulti completo e privo di arbitraggio, come, in realtà, si può dimostrare essere.

2.4 Valutazione (pricing)

Poniamoci ora il problema di calcolare un prezzo *equo* $\Pi(t; X)$ per un dato attivo casuale X con maturità T . Assumiamo che $\mathcal{P} \neq \emptyset$, ovvero che esista una misura martingala Q . Ci sono due approcci:

- $\Pi(t; X)$ è un prezzo equo per l'attivo X se il mercato *allargato*, ovvero ottenuto aggiungendo agli attivi S_0, \dots, S_k un attivo il cui prezzo sia proprio $\Pi(X)$, non consente arbitraggi.

Sia Q la misura martingala per il mercato esteso $[\Pi(X), S_0, \dots, S_k]$. Per definizione, il processo dei prezzi attualizzati, è una martingala sotto Q , ovvero

$$\frac{\Pi(t; X)}{S_0(t)} = E^Q \left[\frac{\Pi(T; X)}{S_0(T)} \middle| \mathcal{F}_t \right] = E^Q \left[\frac{X}{S_0(T)} \middle| \mathcal{F}_t \right].$$

Quindi un prezzo equo per X è dato dalla formula:

$$\Pi(t; X) = S_0(t) E^Q \left[\frac{X}{S_0(T)} \middle| \mathcal{F}_t \right], \quad (2.4.1)$$

dove Q è una misura martingala per il mercato S_0, \dots, S_k dato a priori. Osserviamo che differenti scelte della misura martingala portano a differenti processi dei prezzi per gli attivi casuali.

- Se l'attivo X è replicabile da una strategia h , allora l'unico prezzo equo è dato da $\Pi(t; X) = V(t; h)$.

Supponiamo dunque che l'attivo casuale X sia replicabile tramite una strategia autofinanziata h . Visto che, da un punto di vista finanziario, un attivo casuale X e il suo portafoglio replicante $V(t; h)$ sono equivalenti, il prezzo equo $\Pi(t; X)$ per X è dato da

$$\Pi(t; X) = V(t; h) \quad (2.4.2)$$

Che relazione intercorre fra la 2.4.1 e la 2.4.2? E se X fosse replicabile da due diverse strategie?

Osserviamo che, definendo il prezzo equo come nella 2.4.2, vale l'egualianza

$$\frac{\Pi(t; X)}{S_0(t)} = \tilde{V}(t; h)$$

e, dato che \tilde{V} è una martingala sotto Q , tale sarà anche $\frac{\Pi(t; X)}{S_0(t)}$. Si ottiene, quindi

$$\Pi(t; X) = S_0(t) E^Q \left[\frac{X}{S_0(T)} \middle| \mathcal{F}_t \right] = V(t; h). \quad (2.4.3)$$

La 2.4.3 vale per ogni scelta della strategia replicante e per ogni misura martingala Q .

In un modello completo, il prezzo di ogni attivo casuale, è **univocamente** determinato dalla condizione di assenza di arbitraggio. Il prezzo è unico perché, tale attivo, in un certo senso, è superfluo e può essere sostituito dal portafoglio replicante. In un modello incompleto, invece, la richiesta di non arbitraggio, non è più sufficiente a determinare l'unicità del prezzo di un attivo casuale: abbiamo più misure martingala, ognuna delle quali produce un prezzo diverso, in accordo con l'assenza di arbitraggio. La questione, allora, è quale misura martingala bisogna usare per calcolare i prezzi. La risposta è semplice: la misura martingala viene scelta dal mercato stesso!

2.5 Valutazione dei bonds e misura martingala

Vogliamo in questo paragrafo introdurre una misura martingala, questa volta nel mercato delle obbligazioni senza cedole. A tale scopo dobbiamo specificare la scelta del numerario. La più ovvia, ma anche la più comune è quella dell'attivo privo di rischio B . Per mantenere traccia di B come numerario, aggiungiamo una precisazione alla definizione di misura martingala:

Definizione 2.5.1 *Sia $(\Omega, \mathcal{F}, P, \mathcal{F})$ uno spazio probabilizzato filtrato sul quale viene definito il processo dei prezzi delle obbligazioni senza cedole. Una misura martingala Q è detta neutrale al rischio (risk neutral) se si usa come numerario l'attivo privo di rischio, cioè se, per ogni fissata maturità $T \in \mathbb{I}$, il processo*

$$\tilde{P}(t, T) = \frac{P(t, T)}{B(t)}, \quad 0 \leq t \leq T$$

è una martingala sotto Q .

Consideriamo un attivo casuale X con maturità T . Sia poi Q una misura martingala neutrale al rischio. Possiamo subito calcolare il processo dei

prezzi $\Pi(t; X)$ per X , utilizzando la formula 2.4.1 in cui sostituiamo a $S_0(t)$ il numerario $B(t) = e^{-\int_0^t r(s) ds}$:

$$\Pi(t; X) = E^Q \left[X e^{-\int_t^T r(s) ds} \middle| \mathcal{F}_t \right]. \quad (2.5.1)$$

In particolare, il processo dei prezzi P per un T -bond, è

$$P(t, T) = E^Q \left[e^{-\int_t^T r(s) ds} \middle| \mathcal{F}_t \right],$$

dato che $X = P(T, T) = 1$.

Come ulteriore esempio di applicazione della formula 2.4.1, calcoliamo il prezzo $\Pi_p(t)$ di un payer interest rate swap, usando le stesse notazioni del paragrafo 1.8.1. Ricordiamo che, ad ogni istante T_i , il possessore del contratto swap riceve $(F(T_{i-1}, T_i) - K)\delta N$.

$$\begin{aligned} \Pi_p(t) &= \sum_{i=1}^n E^Q \left[e^{-\int_t^{T_i} r(s) ds} (F(T_{i-1}, T_i) - K)\delta N \middle| \mathcal{F}_t \right] \\ &= \sum_{i=1}^n E^Q \left[e^{-\int_t^{T_i} r(s) ds} \left(\frac{1 - P(T_{i-1}, T_i)}{\delta P(T_{i-1}, T_i)} - K \right) \delta N \middle| \mathcal{F}_t \right] \\ &= N \sum_{i=1}^n E^Q \left[e^{-\int_t^{T_i} r(s) ds} \frac{1}{P(T_{i-1}, T_i)} - e^{-\int_t^{T_i} r(s) ds} (1 + \delta K) \middle| \mathcal{F}_t \right] \\ &= N \sum_{i=1}^n (P(t, T_{i-1}) - (1 + \delta K)P(t, T_i)) \\ &= N \left(P(t, T_0) - P(t, T_n) - \delta K \sum_{i=1}^n P(t, T_i) \right). \end{aligned}$$

L'espressione appena trovata per il prezzo $\Pi_p(t)$ è, ovviamente, la stessa trovata a pagina 12.

2.6 Relazioni fra df , dP e dr

In questo paragrafo mostreremo quali relazioni formali devono intercorrere fra df , dP e dr .

Supponiamo che le dinamiche del tasso a breve r , del prezzo di un T -bond e del tasso futuro siano le seguenti:

Tasso a breve

$$dr(t) = a(t) dt + b(t) dW(t) \quad (2.6.1)$$

Prezzo di un T -bond

$$dP(t, T) = P(t, T)m(t, T) dt + P(t, T)v(t, T) dW(t) \quad (2.6.2)$$

Tasso futuro

$$df(t, T) = \alpha(t, T) dt + \sigma(t, T) dW(t) \quad (2.6.3)$$

Con W intendiamo ovviamente un processo di Wiener, che può essere anche vettoriale; in questo caso le volatilità $\sigma(t, T)$ e $v(t, T)$ sono dei vettori riga.

Avremo bisogno, per svolgere i calcoli in modo agevole, di alcune ipotesi.

Ipotesi 1. Per ogni fissato ω e t $m(t, T)$, $v(t, T)$, $\alpha(t, T)$ e $\sigma(t, T)$ sono derivabili con continuità nella variabile T . La derivata parziale rispetto a T verrà indicata con $m_T(t, T)$, $v_T(t, T)$, \dots

Ipotesi 2. Tutti i processi sono assunti essere abbastanza regolari da permettere sia di derivare sotto il segno di integrale, sia di scambiare l'ordine di integrazione.

Queste ipotesi sono state formulate appositamente e, ovviamente, si vorrebbero dare delle condizioni che implicino le proprietà desiderate. Questo si può fare, ad un altro prezzo però, quello di complicare notevolmente la complessità tecnica. Il risultato principale sulle relazioni cercate è contenuto nella proposizione seguente. Osserviamo che tali relazioni sussistono indipendentemente dalla misura di probabilità considerata e, in particolare, non assumiamo che il mercato sia privo di arbitraggio.

Proposizione 2.6.1 1. Se $P(t, T)$ verifica la 2.6.2, allora il tasso futuro soddisfa

$$df(t, T) = \alpha(t, T) dt + \sigma(t, T) dW(t),$$

dove α e σ sono date da

$$\begin{cases} \alpha(t, T) = v_T(t, T) v(t, T) - m_T(t, T) \\ \sigma(t, T) = -v_T(t, T). \end{cases}$$

2. Se $f(t, T)$ verifica la 2.6.3, allora il tasso a breve soddisfa

$$dr(t) = a(t) dt + b(t) dW(t),$$

dove a e b sono date da

$$\begin{cases} a(t) = f_T(t, t) + \alpha(t, t) \\ b(t) = \sigma(t, t). \end{cases}$$

3. Se $f(t, T)$ verifica la 2.6.3, allora $P(t, T)$ soddisfa

$$\begin{aligned} dP(t, T) &= P(t, T) \left(r(t) + A(t, T) + \frac{1}{2} \|B(t, T)\|^2 \right) dt \\ &+ P(t, T) B(t, T) dW(t), \end{aligned}$$

dove

$$\begin{cases} A(t, T) = - \int_t^T \alpha(t, s) ds \\ B(t, T) = - \int_t^T \sigma(t, s) ds. \end{cases} \quad (2.6.4)$$

Dimostrazione.

1. Applichiamo la formula di Itô al processo $\ln P(t, T)$:

$$\begin{aligned} d \ln P(t, T) &= \frac{dP(t, T)}{P(t, T)} - \frac{1}{2} \frac{P^2(t, T) \|v(t, T)\|^2}{P^2(t, T)} dt \\ &= \left(m(t, T) - \frac{1}{2} \|v(t, T)\|^2 \right) dt + v(t, T) dW(t). \end{aligned}$$

Poiché per definizione $df(t, T) = -\frac{\partial P(t, T)}{\partial T}$, si ha

$$df(t, T) = (-m_T(t, T) + v_T(t, T)v(t, T)) dt - v_T(t, T) dW(t).$$

2. Integrando la dinamica 2.6.3 e ricordando che $r(t) = f(t, t)$, otteniamo

$$r(t) = f(0, t) + \int_0^t \alpha(s, t) ds + \int_0^t \sigma(s, t) dW(s). \quad (2.6.5)$$

Possiamo scrivere

$$\alpha(s, t) = \alpha(s, s) + \int_s^t \alpha_T(s, u) du$$

$$\sigma(s, t) = \sigma(s, s) + \int_s^t \sigma_T(s, u) du,$$

e, inserendole nella 2.6.5, abbiamo

$$\begin{aligned} r(t) &= f(0, t) + \int_0^t \alpha(s, s) ds + \int_0^t \int_s^t \alpha_T(s, u) dud s \\ &+ \int_0^t \sigma(s, s) dW(s) + \int_0^t \int_s^t \sigma_T(s, u) dud W(s). \end{aligned}$$

Scambiando l'ordine di integrazione, si ha la tesi.

3. Scriviamo $P(t, T)$ nella forma $P(t, T) = e^{Y(t, T)}$ dove

$$Y(t, T) = - \int_t^T f(t, s) ds. \quad (2.6.6)$$

Utilizziamo la forma integrale della 2.6.3

$$f(t, s) = f(0, s) + \int_0^t \alpha(u, s) du + \int_0^t \sigma(u, s) dW(u)$$

e inseriamola nella 2.6.6:

$$\begin{aligned} Y(t, T) &= - \int_t^T f(0, s) ds - \int_0^t \int_t^T \alpha(u, s) ds du - \int_0^t \int_t^T \sigma(u, s) ds dW(u) \\ &= - \int_0^T f(0, s) ds - \int_0^t \int_u^T \alpha(u, s) ds du - \int_0^t \int_u^T \sigma(u, s) ds dW(u) \\ &\quad + \int_0^t f(0, s) ds + \int_0^t \int_u^t \alpha(u, s) ds du + \int_0^t \int_u^t \sigma(u, s) ds dW(u) \\ &= Y(0, T) - \int_0^t \int_u^T \alpha(u, s) ds du - \int_0^t \int_u^T \sigma(u, s) ds dW(u) \\ &\quad + \int_0^t f(0, s) ds + \int_0^t \int_0^s \alpha(u, s) du ds + \int_0^t \int_0^s \sigma(u, s) dW(u) ds. \end{aligned}$$

Usando il fatto che $r(s) = f(s, s)$ e integrando l'equazione 2.6.3 sull'intervallo $[0, s]$, otteniamo che

$$r(s) = f(0, s) + \int_0^s \alpha(u, s) du + \int_0^s \sigma(u, s) dW(u).$$

Possiamo riscrivere $Y(t, T)$ nella forma

$$Y(t, T) = Y(0, T) + \int_0^t r(s) ds - \int_0^t \int_u^T \alpha(u, s) ds du - \int_0^t \int_u^T \sigma(u, s) ds dW(u).$$

Ponendo A e B come nella 2.6.4, si ha che

$$dY(t, T) = (r(t) + A(t, T)) dt + B(t, T) dW(t).$$

Applichiamo ora la formula di Itô al processo $P(t, T) = e^{Y(t, T)}$

$$\begin{aligned} dP(t, T) &= P(t, T) dY(t, T) + \frac{1}{2} P(t, T) \|B(t, T)\|^2 dt \\ &= P(t, T) \left(r(t) + A(t, T) + \frac{1}{2} \|B(t, T)\|^2 \right) dt \\ &\quad + P(t, T) B(t, T) dW(t). \quad \square \end{aligned}$$

Capitolo 3

Modelli matematici per i tassi di interesse

Originariamente, la modellizzazione dei tassi di interesse, si basava sull'assunzione di una specifica dinamica per il solo processo del tasso a breve r . I modelli, in cui l'unica variabile modellizzata è il tasso a breve, costituiscono una classe veramente limitata, tuttavia sono storicamente l'approccio più vecchio ed hanno alcune interessanti caratteristiche computazionali. In particolare, gran parte della teoria (dell'hedging e del pricing), viene sviluppata con equazioni alle derivate parziali (PDE) e spesso è possibile trovare formule analitiche per i prezzi dei bonds e dei derivati. Inoltre, modellizzare direttamente il tasso a breve r , è molto conveniente: le quantità fondamentali, infatti, come i prezzi dei bonds e tutti gli altri tipi di derivati, sono definiti, dalla condizione di assenza di arbitraggio, come la speranza condizionale di una funzione del tasso r . L'esistenza di una misura martingala implica che, in assenza di arbitraggio, il prezzo $\Pi(t; X)$ all'istante t di un contratto che paga X all'istante T , sia dato dalla 2.5.1, ovvero

$$\Pi(t; X) = E^Q \left[X e^{-\int_t^T r(s) ds} \middle| \mathcal{F}_t \right].$$

In particolare, come già visto in precedenza, il prezzo $P(t, T)$ all'istante t di un'obbligazione senza cedole con maturità T , essendo l'equivalente all'istante t di una unità prelevata a T , cioè $X = 1$, è dato da

$$P(t, T) = E^Q \left[e^{-\int_t^T r(s) ds} \middle| \mathcal{F}_t \right].$$

Grazie a quest'ultima espressione, è chiaro che, ogni volta che riusciamo a caratterizzare la distribuzione $e^{-\int_t^T r(s) ds}$, con la dinamica per r scelta, condizionando con le informazioni disponibili all'istante t , siamo in grado di calcolare i prezzi dei bonds. E, come abbiamo già osservato nel primo capitolo, dai prezzi dei bonds possiamo ricavare tutti gli altri tipi di tassi di interesse, così, l'intera curva dei rendimenti, è caratterizzata dalle proprietà

della distribuzione di r . Nonostante tutti questi vantaggi, è piuttosto irragionevole che l'intero mercato venga governato da un'unica variabile. Nascono così approcci diversi, come ad esempio i modelli a due o più fattori o i modelli basati sul tasso futuro. Di questi tratteremo brevemente in seguito, mentre ora ci addentreremo nella formalizzazione dei modelli basati sul tasso a breve.

3.1 Generalità

Assumiamo che, la dinamica del tasso a breve r , sotto la misura di probabilità P del mondo reale, ovvero quella originale, sia

$$dr(t) = \mu(t, r(t)) dt + \sigma(t, r(t)) d\tilde{W}(t), \quad (3.1.1)$$

dove

- $\mu(t, r)$ e $\sigma(t, r)$ sono date funzioni reali, abbastanza regolari da garantire all'equazione differenziale stocastica 3.1.1 di avere un'unica soluzione;
- \tilde{W} è un processo di Wiener sotto P ;
- la filtrazione è quella generata da \tilde{W} .

Supponiamo che l'unico processo dei prezzi conosciuto a priori sia l'attivo privo di rischio B con

$$\begin{cases} dB(t) = r(t)B(t) dt \\ B(0) = 1. \end{cases} \quad (3.1.2)$$

La prima naturale domanda da porsi è se, specificare la dinamica di r sotto P e la richiesta che il mercato dei bonds sia privo di arbitraggio, siano sufficienti per determinare univocamente i prezzi dei bonds. La risposta è no, e la ragione è che il mercato specificato a priori è incompleto. Infatti abbiamo a disposizione solo l'attivo privo di rischio e non abbiamo possibilità di formare dei portafogli interessanti; l'unica cosa che possiamo fare è semplicemente depositare tutto il nostro capitale in banca e aspettare che cresca in accordo con la 3.1.2. In particolare non c'è alcuna possibilità di replicare nessun attivo casuale, nemmeno quelli semplici associati alle obbligazioni senza cedole. Questa mancanza di completezza può essere vista anche come conseguenza della non unicità della misura martingala del mercato dato a priori. Infatti, essendo il processo dei prezzi attualizzati quello banale, ovvero

$$\tilde{S}(t) = \frac{S_0(t)}{S_0(t)} \equiv 1,$$

si ha che, **ogni** misura di probabilità Q con $Q \sim P$, è una misura martingala.

Usando il teorema 2.3.3 con $k = 0$ e $r = 1$, ci si aspetta che il mercato sia privo di arbitraggio, ma incompleto.

Anche nel modello Black-Scholes, descritto a pagina 19 ha una sola fonte aleatoria e un solo attivo S . La differenza fondamentale fra i due modelli è che, al contrario di S , il tasso a breve r non è il prezzo di un attivo.

Come abbiamo appena visto, i prezzi dei bonds con maturità diverse non sono unici. Tuttavia essi non potranno assumere una forma qualunque, visto che, per soddisfare la richiesta di assenza di arbitraggio, dovranno verificare certe relazioni. Lo scopo del prossimo paragrafo sarà proprio quello di trovare tali relazioni.

3.2 Equazione fondamentale della struttura a termine

Assumiamo che:

Ipotesi 1. Il mercato dei T -bonds sia privo di arbitraggio per ogni scelta della maturità T ;

Ipotesi 2. La dinamica del tasso a breve r sia, sotto P , data da:

$$dr(t) = \mu(t, r(t)) dt + \sigma(t, r(t)) d\tilde{W}(t);$$

Ipotesi 3. Il processo dei prezzi per un T -bond sia della forma

$$P(t, T) = F(t, r(t), T),$$

dove F è una funzione sufficientemente regolare da permettere tutti i calcoli successivi.

Il ragionamento che faremo per trovare le relazioni fra i prezzi dei bonds con diverse maturità che ne caratterizzano la struttura a termine (si veda pagina 7), può essere riassunto nel seguente schema:

1. Fissiamo due maturità S e T e costruiamo una strategia basata su S -bonds e T -bonds.

Per sottolineare il fatto che la maturità è fissata, scriveremo $F^T(t, r)$ in luogo di $F(t, r, T)$. Inoltre con F_t^T intenderemo $\frac{\partial F^T}{\partial t}$.

2. Scegliremo i relativi pesi di una strategia in modo da avere un portafoglio, la cui dinamica sia

$$dV(t) = k(t)V(t) dt,$$

ovvero priva del termine di disturbo $d\tilde{W}$.

3. Questo ci porta ad aver creato una “ipotetica banca” con $k(t)$ come tasso a breve. Per escludere arbitraggi, grazie alla proposizione 2.2.3, dobbiamo avere che la relazione $k(t) = r(t)$ valga P -q.c.

Questa condizione ci procurerà un’equazione alle derivate parziali.

Iniziamo ora il programma appena schematizzato. Usando la formula di Itô otteniamo che, la dinamica del prezzo di un T -bond, è

$$\begin{aligned} dF^T &= \frac{\partial F^T}{\partial t} dt + \frac{\partial F^T}{\partial r} dr + \frac{1}{2} \sigma^2 \frac{\partial^2 F^T}{\partial^2 r} dt \\ &= \left(\frac{\partial F^T}{\partial t} + \frac{\partial F^T}{\partial r} \mu + \frac{1}{2} \sigma^2 \frac{\partial^2 F^T}{\partial^2 r} \right) dt + \frac{\partial F^T}{\partial r} \sigma d\tilde{W}. \end{aligned}$$

Poniamo

$$\begin{cases} \alpha_T = \frac{F_t^T + \mu F_r^T + \frac{1}{2} \sigma^2 F_{rr}^T}{F^T} \\ \sigma_T = \frac{\sigma F_r^T}{F^T} \end{cases} \quad (3.2.1)$$

in modo da poter scrivere

$$dF^T = F^T \alpha_T + F^T \sigma_T d\tilde{W}. \quad (3.2.2)$$

Consideriamo il valore del portafoglio V corrispondente alla strategia (h^S, h^T)

$$V = h^S F^S + h^T F^T.$$

I relativi pesi sono definiti come:

$$u^S = \frac{h^S F^S}{V} \quad \text{e} \quad u^T = \frac{h^T F^T}{V} \quad \left(\Rightarrow u^S + u^T = \frac{h^S F^S + h^T F^T}{V} = 1 \right).$$

Quindi

$$dV = h^S dF^S + h^T dF^T = V \left(u^S \frac{dF^S}{F^S} + u^T \frac{dF^T}{F^T} \right).$$

Sostituiamo a dF^T l’espressione 3.2.2, ottenendo

$$dV = V(u^S \alpha_S + u^T \alpha_T) dt + V(u^S \sigma_S + u^T \sigma_T) d\tilde{W}. \quad (3.2.3)$$

Cerchiamo i pesi in modo tale che il termine in $d\tilde{W}$ non ci sia, ovvero

$$\begin{cases} u^S + u^T = 1 \\ u^S \sigma_S + u^T \sigma_T = 0 \end{cases}$$

La soluzione di tale sistema è data da

$$\begin{cases} u^S = -\frac{\sigma_T}{\sigma_S - \sigma_T} \\ u^T = \frac{\sigma_S}{\sigma_S - \sigma_T} \end{cases}$$

L'equazione 3.2.3 assume ora la forma

$$dV = V \left(\frac{\alpha_T \sigma_S - \alpha_S \sigma_T}{\sigma_S - \sigma_T} \right) dt.$$

Per la proposizione 2.2.3, visto che si escludono arbitraggi, il processo fra le parentesi tonde deve eguagliare il tasso a breve r , ovvero

$$\frac{\alpha_T \sigma_S - \alpha_S \sigma_T}{\sigma_S - \sigma_T} = r \quad \text{cioè} \quad \frac{\alpha_S(t) - r(t)}{\sigma_S(t)} = \frac{\alpha_T(t) - r(t)}{\sigma_T(t)}.$$

Il membro sinistro dell'ultima equazione non dipende da T , come del resto quello destro non dipende da S . Abbiamo, quindi, la seguente proposizione:

Proposizione 3.2.1 *Assumiamo che il mercato dei bonds sia privo di arbitraggio. Allora esiste un processo λ tale che la relazione*

$$\frac{\alpha_T(t) - r(t)}{\sigma_T(t)} = \lambda(t) \quad (3.2.4)$$

valga per ogni t e per ogni scelta della maturità T .

C'è una naturale interpretazione di questo risultato. Ricordiamo che r è il tasso di rendimento dell'attivo privo di rischio, mentre α_T è, per l'equazione 3.2.2, il rendimento locale di un T -bond. Il numeratore $\alpha_T - r$ rappresenta una differenza nel guadagno, ovvero indica quanto meglio stiamo facendo rispetto al caso privo di rischio, cioè rispetto ad investire nel titolo privo di rischio. Quando dividiamo tale quantità per σ_T , stiamo dividendo per l'ammontare del rischio al quale siamo soggetti, misurato dalla volatilità σ_T del T -bond. Quindi λ può essere interpretato come il premio (per unità di volatilità) per il rischio corso nell'investire nel T -bond. Un termine usato per il processo λ è *il prezzo del rischio* (the market price of risk).

Possiamo ottenere ulteriori informazioni dall'equazione 3.2.4, inserendovi le definizioni di α_T e σ_T date dalla 3.2.1. Svolgendo le dovute semplificazioni e tenendo conto che $P(T, T) = F^T(T, r) = 1$, otteniamo l'equazione fondamentale della struttura a termine:

Proposizione 3.2.2 *Assumiamo che valgano le ipotesi 1, 2 e 3 di pagina 34. Allora F^T soddisfa l'equazione*

$$F_t^T + (\mu - \lambda \sigma) F_r^T + \frac{1}{2} \sigma^2 F_{rr}^T - r F^T = 0, \quad (3.2.5)$$

$$F^T(T, r) = 1.$$

Segue dall'equazione 3.2.4 che λ è della forma $\lambda = \lambda(t, r)$, quindi la 3.2.5 è un'equazione alle derivate parziali. Il problema però è che λ non è determinato dal modello. Quindi per risolverla occorre specificare λ .

Possiamo ottenere anche un'ulteriore rappresentazione (Feynman-Kač) della soluzione dell'equazione 3.2.5.

Proposizione 3.2.3 *La funzione $F(t, r, T)$ può essere espressa nella forma*

$$F(t, r, T) = E_{t,r}^Q \left[\exp \left(- \int_t^T r(s) ds \right) \right],$$

dove la dinamica di r sotto Q è data da

$$\begin{aligned} dr(s) &= (\mu - \lambda \sigma) ds + \sigma dW(s) \\ r(t) &= r. \end{aligned} \quad (3.2.6)$$

Osserviamo che la dinamica di r sotto Q è stata ottenuta dalla dinamica sotto P attraverso una trasformazione di Girsanov e che ovviamente Q è una misura martingala. Come abbiamo già osservato λ non è determinato entro il modello, quindi dobbiamo specificarlo a priori. In effetti λ , comparendo nella formula della densità $\frac{dQ}{dP}$, collega direttamente la misura di probabilità originale P con quella martingala Q . Specificare λ , dunque, è equivalente a specificare la misura martingala, che non è determinata dal modello. Sono gli agenti, nel mercato concreto, che (implicitamente) determinano Q e λ .

Se siamo interessati principalmente a calcolare i prezzi dei derivati dei tassi di interesse, possiamo modellizzare direttamente la dinamica di r sotto Q (e λ è così implicito) visto che, per questo scopo, la misura di probabilità originale P non è necessaria. Questo approccio viene chiamato *modellizzazione martingala*.

3.3 Modellizzazione martingala del tasso a breve

Modellizziamo ora il tasso a breve r direttamente sotto una fissata misura martingala Q .

Assumiamo che, sotto Q , la dinamica di r sia

$$dr(t) = \mu(t, r(t)) dt + \sigma(t, r(t)) dW(t),$$

dove W è un processo di Wiener sotto Q .

Enunciamo ora la proposizione che contiene l'equazione fondamentale generale della struttura a termine.

Proposizione 3.3.1 *Sia X un attivo casuale con maturità T della forma $X = \Phi(r(T))$. Allora il processo dei prezzi privo di arbitraggio è dato da $\Pi(t, X) = F(t, r(t))$ dove F è una soluzione dell'equazione*

$$F_t(t, r) + \mu(t, r)F_r(t, r) + \frac{1}{2}\sigma^2(t, r)F_{rr}(t, r) - rF(t, r) = 0,$$

$$F(T, r) = \Phi(r).$$

In particolare, i prezzi dei T -bonds, sono dati da $P(t, T) = F^T(t, r(t))$, dove

$$F_t^T + \mu F_r^T + \frac{1}{2}\sigma^2 F_{rr}^T - rF^T = 0,$$

$$F^T(T, r) = 1.$$

3.3.1 Esempi

Lontano dall'essere una lista esauriente, diamo qui alcuni esempi dei modelli classici per il tasso a breve, con la convenzione che, se i parametri dipendono dal tempo, verrà indicato esplicitamente, in caso contrario saranno considerati costanti.

1. Vasiček (1977)

$$dr = (b - ar) dt + \sigma dW$$

2. Cox-Ingersoll-Ross (CIR, 1985)

$$dr = (b - ar) dt + \sigma\sqrt{r} dW$$

3. Dothan (1978)

$$dr = ar dt + \sigma r dW$$

4. Black-Derman-Toy (1990)

$$dr = a(t)r dt + \sigma(t)r dW$$

5. Ho-Lee (1986)

$$dr = \Phi(t) dt + \sigma dW$$

6. Hull-White (estensione Vasiček) (1990)

$$dr = (\Phi(t) - a(t)r) dt + \sigma(t) dW$$

7. Hull-White (estensione CIR) (1985)

$$dr = (\Phi(t) - a(t)r) dt + \sigma(t)\sqrt{r} dW$$

Non volendoci addentrare troppo nel dettaglio, parleremo in maniera del tutto informale delle caratteristiche principali dei modelli appena elencati. Iniziamo, per esempio, con il modello Vasiček : la dinamica del tasso a breve segue un'equazione differenziale stocastica lineare che può essere risolta esplicitamente e la legge di r risulta normale. Inoltre, sia le espressioni che le distribuzioni di molte quantità legate ai tassi di interesse, sono facilmente ottenibili. Il prezzo per un T -bond, dato dalla usuale formula

$$P(t, T) = E^Q \left[e^{-\int_t^T r(s) ds} \middle| \mathcal{F}_t \right],$$

può essere calcolato come una semplice espressione dipendente da a , b , σ e $r(t)$; una volta nota la funzione $T \mapsto P(t, T; a, b, \sigma, r(t))$, siamo in grado di conoscere l'intera curva dei rendimenti all'istante t . Questo significa che, la

curva iniziale dei rendimenti all'istante $t = 0$, è un output del modello, che dipende dai parametri a, b, σ e dalla condizione iniziale $r(0) = r_0$.

Uno dei difetti principali di questo modello è che i tassi possono assumere valori negativi con probabilità positiva, anche se in realtà molto piccola. Un modello invece che esclude tale possibilità è il modello CIR, grazie all'aggiunta del termine \sqrt{r} nella parte stocastica della dinamica di $r(t)$. Se i parametri a, b e σ variano entro limiti ragionevoli, questo modello implica tassi di interesse positivi. Inoltre mantiene una certa trattabilità analitica, nonostante, rispetto al modello Vasiček sia più realistico.

Anche il modello Dothan assicura tassi di interesse positivi q.c. per ogni $t > 0$ e questo perché $r(t)$ risulta avere una legge lognormale (cioè la legge di $\ln r$ è normale). Uno degli svantaggi di questo modello, che troviamo anche nel modello Black-Derman-Toy, è l'esplosione del titolo privo di rischio. Si può dimostrare, infatti, che

$$E \left[e^{\int_0^{\Delta t} r(s) ds} \middle| \mathcal{F}_0 \right] = +\infty.$$

Ciò significa che, in un intervallo di tempo arbitrariamente piccolo, possiamo, in media, fare infiniti soldi partendo da una unità (per esempio 1 euro).

Un problema classico che coinvolge i modelli più semplici, come il Vasiček, il CIR e il Dothan, è che non sono in grado di riprodurre, in modo soddisfacente, la curva osservata dei rendimenti, essendo più un output che un input del modello. Indichiamo con $P^*(0, T)$ i prezzi dei bond osservati sul mercato all'istante $t = 0$. Una volta osservati, siamo in grado di costruire la curva $T \mapsto P^*(0, T)$, che vorremmo il nostro modello fosse in grado di incorporare. Per questo motivo, dobbiamo forzare i parametri presenti nel modello a creare una curva che sia più vicina possibile a quella osservata nel mercato, ovvero alla $T \mapsto P^*(0, T)$. Nonostante i prezzi $P^*(0, T)$ vengano osservati solo per un numero finito di maturità T_i , tre parametri (a, b e σ , per esempio) non sono sufficienti a riprodurre in modo soddisfacente la curva dei rendimenti. Inoltre, alcune forme della curva $T \mapsto R^*(0, T)$ (come ad esempio quella descritta a pagina 7), non possono mai essere ottenute, per nessun valore dei parametri.

Un modo per risolvere questo problema è introdurre dei parametri che variano nel tempo. Questa è la caratteristica principale dei modelli Ho-Lee, Hull-White (sia estensione Vasiček che CIR). Del problema della stima dei parametri parleremo più dettigliatamente nel prossimo paragrafo.

Una proprietà comune a quasi tutti i modelli elencati (tranne il Dothan e il Black-Derman-Toy) è quella di essere dei modelli affini, nel senso che i prezzi dei bonds possono essere scritti nella forma

$$P(t, T) = e^{A(t, T) - B(t, T)r(t)},$$

con A e B opportune funzioni deterministiche. Si può dimostrare che un modello

$$dr(t) = \mu(t, r(t)) dt + \sigma(t, r(t)) dW(t)$$

è affine se μ e σ sono della forma

$$\begin{aligned}\mu(t, r) &= \alpha(t)r + \beta(t) \\ \sigma(t, r) &= \sqrt{\gamma(t)r + \delta(t)},\end{aligned}\tag{3.3.1}$$

ovvero che la struttura affine per μ e σ^2 sono condizioni sufficienti a garantire che il modello sia affine. Inoltre, se il modello è affine e μ e σ^2 sono indipendenti dal tempo, allora tali coefficienti sono necessariamente affini, cioè della forma $\mu(r) = \alpha r + \beta$ e $\sigma(r) = \sqrt{\gamma r + \delta}$.

Concludiamo questo paragrafo con alcune considerazioni. Nella scelta di un modello per il tasso a breve r , ovviamente si deve tener conto sia dei vantaggi che degli svantaggi. In particolare, bisognerebbe porsi alcune domande:

- La dinamica di r permette di avere tassi di interesse positivi, ovvero $r(t) > 0$ q.c. per ogni t ?
- Quale legge per r implica la dinamica scelta?
- I prezzi dei bond $P(t, T) = E^Q \left[e^{-\int_t^T r(s) ds} \middle| \mathcal{F}_t \right]$ (e quindi anche i tassi futuri, istantanei e swap) sono esplicitamente calcolabili?
- I prezzi dei derivati (come caps, floors, swaptions) sono esplicitamente calcolabili?
- Il modello è *mean-reverting*, nel senso che il valore atteso di $r(t)$ tende ad un valore costante nel limite per $t \rightarrow +\infty$, mentre la sua varianza non esplode?

3.4 Inversione della curva dei rendimenti

Supponiamo di aver scelto un modello specifico per il tasso a breve, per esempio il modello Vasiček. Per implementarlo dobbiamo stimare i parametri a , b e σ e a tale scopo sembra naturale utilizzare i dati delle serie storiche. Ma, avendo scelto di modellizzare la dinamica di r sotto la misura martingala Q , i parametri sono dei Q -parametri, mentre i dati vengono ovviamente osservati sotto la probabilità originale P . Così la loro stima non è più un problema statistico standard. Per evitarlo potremmo modellizzare r direttamente sotto la probabilità originale P . Così facendo, però, come già osservato, non siamo più in grado di determinare i prezzi dei bonds, dato che vengono calcolati utilizzando Q . Andiamo incontro, quindi, al problema della stima del kernel di Girsanov che ci permette di passare da P a Q , ma questo problema è più o meno equivalente alla stima dei Q -parametri. Osserviamo comunque che, dal confronto fra l'equazione 3.1.1 e l'equazione 3.2.6, la volatilità è la stessa sia sotto P che sotto Q .

Come abbiamo già osservato, ci sono più misure martingala e dobbiamo determinare con quale stiamo lavorando. Poiché la misura martingala non è specificata entro il modello, ma piuttosto scelta dal mercato, le informazioni necessarie a determinarla saranno estratte dai prezzi osservati direttamente nel mercato. Questo viene fatto *calibrando il modello ai dati* (fitting the yield curve) che può essere schematizzato così:

- Fissare un modello concreto che coinvolga un vettore parametrico α :

$$dr(t) = \mu(t, r(t); \alpha) dt + \sigma(t, r(t); \alpha) dW(t).$$

- Risolvere, per ogni maturità T , l'equazione fondamentale

$$\begin{cases} F_t^T + \mu F_r^T + \frac{1}{2} \sigma^2 F_{rr}^T - r F^T = 0 \\ F^T(T, r) = 1. \end{cases}$$

Questo ci fornisce il teorico prezzo $P(t, T; \alpha) = F^T(t, T; \alpha)$;

- Osserviamo oggi ($t = 0$), dal mercato, i prezzi $P(0, T)$ per ogni T . Li indicheremo con $\{P^*(0, T), T \geq 0\}$.
- Scegliamo ora α in modo tale che, la teorica curva $\{P(0, T; \alpha), T \geq 0\}$, si avvicini il più possibile a $\{P^*(0, T), T \geq 0\}$. Così abbiamo stimato il nostro vettore parametrico α^* .
- Usando questa procedura abbiamo fissato Q , richiedendo che, la dinamica di r sotto la probabilità martingala Q , sia

$$dr(t) = \mu(t, r(t); \alpha^*) dt + \sigma(t, r(t); \alpha^*) dW(t).$$

Idealmente vorremmo trovare α^* tale che

$$P(0, T; \alpha^*) = P^*(0, T) \text{ per ogni } T \geq 0,$$

ma questo è un sistema infinito dimensionale (un'equazione per ogni T). Quindi, se stiamo lavorando con un modello contenente un vettore parametrico finito dimensionale, non c'è speranza di ottenere un perfetto adattamento ai dati osservati. Questo è un grande problema anche perché i prezzi dei derivati sono molto sensibili ai cambiamenti del processo dei prezzi sottostante. È naturale dunque costruire modelli con un vettore parametrico infinito dimensionale e, un modo, è quello di lasciare uno o più parametri, in un modello finito dimensionale, come funzioni deterministiche del tempo. Questo è esattamente l'approccio usato nei modelli Ho-Lee e Hull-White (sia estensione CIR che Vasiček). Va comunque osservato che, l'introduzione di

un vettore parametrico infinito dimensionale, porta una notevole instabilità numerica nella stima dei parametri.

C'è un approccio, completamente differente, per risolvere il problema dell'adattamento del modello ai dati osservati, quello di D. Heath, R. Jarrow e A. Morton (HJM). Nel loro modello viene presa, come condizione iniziale per la curva dei tassi futuri, proprio la curva osservata sul mercato avendo così, praticamente per definizione, un perfetto adattamento ai dati osservati. A questo modello sarà dedicato il paragrafo 3.6.

3.5 Perché i modelli multifattoriali per il tasso a breve?

Più che descrivere nel dettaglio i modelli multifattoriali, vorremmo motivarne il loro utilizzo, mettendo in evidenza i punti deboli dei modelli ad un solo fattore. Abbiamo presentato fino ad ora solo modelli in cui il tasso a breve r_t costituisce la coordinata fondamentale con la quale viene caratterizzata l'intera curva dei rendimenti. Infatti, la conoscenza del tasso a breve e delle proprietà della sua distribuzione, portano alla conoscenza dei prezzi dei bonds, attraverso l'usuale relazione

$$P(t, T) = E^Q \left[e^{-\int_t^T r(s) ds} \middle| \mathcal{F}_t \right].$$

Una volta calcolati tali prezzi ad un dato istante t , si può ricostruire l'intera curva dei rendimenti, sempre allo stesso istante t . Così, l'evoluzione di una singola quantità, è sufficiente a caratterizzare l'intera curva dei rendimenti. Dunque, la scelta di un modello povero per la dinamica di r risulta un modello povero anche per tale curva. Per spiegare meglio questa affermazione, consideriamo per un momento il modello Vasiček, descritto brevemente nel paragrafo 3.3.1. Ricordiamo che questo modello prevede che la dinamica di r segua l'equazione differenziale stocastica lineare

$$dr = (b - ar) dt + \sigma dW$$

e che r sia così una variabile gaussiana. Inoltre, essendo un modello affine, i prezzi dei bonds assumono la forma

$$P(t, T) = e^{A(t, T) - B(t, T)r(t)}.$$

Da qui possiamo calcolare facilmente tutti i tipi di tassi di interesse in termini di r , come ad esempio il tasso istantaneo continuamente composto $R(t, T)$ introdotto nel paragrafo 1.5:

$$R(t, T) := -\frac{\ln P(t, T)}{(T - t)} = -\frac{A(t, T)}{(T - t)} + \frac{B(t, T)}{(T - t)} r_t := a(t, T) + b(t, T) r_t. \quad (3.5.1)$$

Consideriamo, all'istante t , un attivo casuale X che dipenda dalla legge congiunta di due tassi del tipo 3.5.1. Per esempio, supponiamo che la data di maturità di X sia $T_2 = t + 10$ (anni) e quella di inizio sia $T_1 = t + 1$ (anni). Il pagamento X all'istante T_2 dipenderà quindi dalla legge congiunta dei tassi $R_1 = R(t, T_1)$ a un anno e $R_2 = R(t, T_2)$ a dieci anni. In particolare, visto che viene coinvolta la legge congiunta, la correlazione fra questi due tassi ha un ruolo cruciale. Nel modello Vasiček questa correlazione può essere facilmente calcolata:

$$\text{Corr}[R_1, R_2] = \frac{\text{Cov}[R_1, R_2]}{\sqrt{\text{Var}[R_1]\text{Var}[R_2]}} \equiv 1.$$

Ciò significa che, ad ogni istante e per ogni maturità, i tassi istantanei continuamente composti sono perfettamente correlati: per esempio, il tasso a trent'anni prelevato all'istante t risulta perfettamente correlato con quello a tre mesi, prelevato allo stesso istante. Come conseguenza, un salto nella curva dei rendimenti all'istante t , viene trasmesso invariato a tutte le maturità e la curva, quando il suo punto iniziale (il tasso a breve r_t) compie tale salto, si muove rigidamente nella stessa direzione. Nella realtà, però, i tassi di interesse mostrano una certa decorrelazione (ovvero una non perfetta correlazione). I dati osservati suggeriscono che il grado di correlazione decresce con l'aumentare della distanza fra le maturità. Si devono cercare allora modelli più soddisfacenti dal punto di vista dell'evoluzione della curva dei rendimenti. I modelli ad un solo fattore (come ad esempio Hull-White), sono ancora molto usati quando il derivato di cui si vuole calcolare il prezzo, non dipende dalla correlazione fra tassi con diverse maturità, ma da un'unico tasso dell'intera curva dei rendimenti (per esempio il tasso a sei mesi). Tuttavia, se i tassi congiunti da cui dipende il nostro pagamento X hanno maturità molto vicine fra loro (per esempio sei mesi e un anno), ci si aspetta che la loro correlazione sia abbastanza alta. Così la perfetta correlazione indotta dai modelli ad un solo fattore non è, in linea di principio, inaccettabile. Quando però la correlazione gioca un ruolo rilevante, o comunque quando viene richiesta una maggior precisione, è necessario trovare nuovi approcci che consentano ad un modello di avere una struttura per la correlazione più realistica. Ecco che compaiono i modelli multifattoriali e in particolare quelli a due fattori. Supponiamo di sostituire il modello Vasiček con il suo ipotetico modello a due fattori:

$$\begin{aligned} r_t &= x_t + y_t \\ dx_t &= (b_x - a_x x_t) dt + \sigma_x dW_1(t) \\ dy_t &= (b_y - a_y y_t) dt + \sigma_y dW_2(t), \end{aligned}$$

dove la correlazione istantanea fra le fonti aleatorie è data da $\rho dt = dW_1 dW_2$. Si può dimostrare che, per questo tipo di modelli, vale ancora che i prezzi dei bonds sono funzioni affini, questa volta dei due fattori x e y

$$P(t, T) = e^{A(t, T) - B^x(t, T) x_t - B^y(t, T) y_t}$$

e che la correlazione è data da

$$\text{Corr}[R_1, R_2] = \text{Corr}[b^x(t, T_1)x_t + b^y(t, T_1)y_t, b^x(t, T_2)x_t + b^y(t, T_2)y_t] \neq 1$$

(dove $b^x(t, T) := B^x(t, T)/(T - t)$ e $b^y(t, T) := B^y(t, T)/(T - t)$).

Questa volta la correlazione fra i tassi R_1 e R_2 dipende fortemente dalla correlazione fra i due fattori x e y che, a sua volta, dipende anche da quella istantanea ρ . Bisognerebbe ora discutere sulla flessibilità della struttura della covarianza che si è guadagnata e chiedersi se questa sia sufficiente per i propositi pratici, ma questo ci allontana dai nostri scopi. Per una trattazione più approfondita si può consultare [4, capitolo 4].

Una domanda che viene naturale porsi è: quanti fattori dobbiamo coinvolgere in un modello per soddisfare i nostri propositi pratici? Ovviamente, quello che abbiamo presentato a due fattori, può essere esteso a tre o più fattori. La scelta di tale numero deve essere un buon compromesso fra un'efficiente implementazione numerica e la capacità del modello, sia di rappresentare una struttura realistica per la correlazione (e più in generale per la covarianza), sia di adattarsi, nella maggior parte dei casi e in modo soddisfacente, ai dati osservati nel mercato. Una risposta a questa domanda viene data dall'analisi in componenti principali dell'intera curva dei rendimenti. Essa ci suggerisce che, sotto la probabilità originale P , due o tre componenti principali riescono già a mantenere dall'85% al 90% della varianza totale. Dato che la struttura della covarianza istantanea dello stesso processo, quando si passa dalla probabilità P alla misura martingala Q , non cambia, possiamo pensare che, anche nel mondo neutrale al rischio, un modello a due o tre fattori può essere utilizzato per avere risultati soddisfacenti.

All'analisi in componenti principali verrà dedicato il capitolo successivo.

3.6 Heath-Jarrow-Morton

3.6.1 Introduzione

Fino a questo punto abbiamo analizzato modelli per i tassi di interesse in cui il tasso a breve è l'unica variabile modellizzata. Come già osservato più volte, questo porta sia vantaggi che svantaggi. Riassumiamoli brevemente:

- una volta conosciuta la legge di r si possono calcolare i prezzi dei derivati tramite la risoluzione di equazioni alle derivate parziali;
- dai prezzi dei bonds (di cui nella maggior parte dei casi si possono dare formule esplicite) si può ricostruire l'intera curva dei rendimenti;
- da un punto di vista economico, però, è piuttosto irragionevole che l'intero mercato sia governato da un'unica variabile;
- è difficile ottenere una struttura realistica per la volatilità dei tassi futuri senza introdurre modelli complicati;

- non appena il modello diventa più realistico, l'adattamento ai dati osservati descritto nel paragrafo 3.4 diventa molto più difficile.

Un'alternativa ai modelli basati sul tasso a breve fu proposta da D. Heath, R. Jarrow e A. Morton nel 1992: modellizzare direttamente l'intera curva dei tassi futuri, prendendo come variabile (infinito dimensionale) il tasso futuro istantaneo $f(t, T)$.

Un'altra differenza fra i modelli basati su r e quelli basati su f è che, nei primi, per specificare la dinamica di r servivano sia $\mu(t, r(t))$ che $\sigma(t, r(t))$, mentre negli ultimi è sufficiente specificare la volatilità σ . Infatti, Heath, Jarrow e Morton, dimostrarono, imponendo la condizione di non arbitraggio, che la dinamica del tasso futuro

$$df(t, T) = \alpha(t, T) dt + \sigma(t, T) dW(t)$$

(e quindi anche $\alpha(t, T)$) era completamente specificata dalla sola volatilità $\sigma(t, T)$.

Si parla, più che di modello, di “metodologia” HJM: svilupparono, infatti, in un contesto più generale, una teoria entro la quale trovarono spazio numerosi modelli basati sul tasso futuro. In questi modelli, il problema fondamentale, è definire un'opportuna funzione per la volatilità $\sigma(t, T)$. Solo una ristretta classe di funzioni implica un processo Markoviano per il tasso a breve che rende i calcoli per le valutazioni più agevoli. Per esempio, questo accade quando σ è della forma

$$\sigma(t, T) = a(t)b(T),$$

con a e b funzioni strettamente positive e deterministiche del tempo.

Le forme standard, come ad esempio

$$\sigma(t, T) = \sigma, \sigma e^{-\lambda(T-t)}, \sigma_0 + (T-t)\sigma_1, \dots$$

rendono il modello trattabile ma non realistico.

Ci sono poi volatilità che implicano, sia per il tasso a breve, sia per quello futuro, distribuzioni normali, come nel caso in cui σ sia del tipo

$$\sigma(t, T) = \sigma_0[\beta(T-t) + 1]e^{-\lambda(T-t)},$$

con σ_0, β e λ costanti non negative.

3.6.2 Esistenza di una misura martingala

Procediamo con la formalizzazione matematica di un modello basato sul tasso futuro f e diamo delle condizioni per l'esistenza di una misura martingala.

Consideriamo, sullo spazio probabilizzato filtrato $(\Omega, \mathcal{F}, P, \underline{\mathcal{F}})$, un processo di Wiener \tilde{W} d -dimensionale. La filtrazione è quella generata da \tilde{W} .

Assumiamo che, per ogni fissato T , la dinamica del tasso futuro $f(t, T)$, sotto la misura di probabilità originale P , sia data da

$$df(t, T) = \alpha(t, T) dt + \sigma(t, T) d\tilde{W}(t), \quad (3.6.1)$$

dove, per ogni T , $\alpha(t, T)$ e $\sigma(t, T)$ sono processi adattati. Assumiamo inoltre l'esistenza dell'usuale attivo privo di rischio B .

Osserviamo subito che 3.6.1 è un sistema stocastico infinito dimensionale (un'equazione per ogni fissato T).

Come condizione iniziale ($t = 0$), usiamo la curva dei tassi futuri osservati dal mercato all'istante $t = 0$, cioè

$$f(0, T) = f^*(0, T), \quad \forall T \geq 0$$

così, **per costruzione**, otteniamo automaticamente un perfetto adattamento del modello ai dati osservati. Il problema descritto nel paragrafo 3.4 viene quindi completamente evitato.

Vista la corrispondenza

$$P(t, T) = e^{-\int_t^T f(t, s) ds},$$

la determinazione dei tassi futuri $\{f(t, T), T > 0\}$ è equivalente a quella dei prezzi dei bonds $\{P(t, T), T > 0\}$.

Dato che abbiamo d fonti aleatorie e un numero infinito di attivi (un bond per ogni maturità T) andiamo incontro al rischio di aver introdotto delle opportunità di arbitraggio. Cerchiamo allora di capire quando il mercato dei bonds, indotto dal modello 3.6.1, sia privo di arbitraggio, nel senso che esiste una misura martingala. Il risultato principale è il seguente:

Proposizione 3.6.1 *Assumiamo che la dinamica della famiglia dei tassi futuri sia data dalla 3.6.1. Allora esiste una misura martingala Q se e solo se esiste un processo adattato $g(t)$ con le seguenti proprietà:*

1. *Il processo*

$$\mathcal{E}(g \star \tilde{W})_t := \exp \left(\int_0^t g(s) \cdot d\tilde{W}(s) - \frac{1}{2} \int_0^t \|g(s)\|^2 ds \right) \quad (3.6.2)$$

è una martingala sotto P .

2. *Per ogni $T \geq 0$ e per ogni $t \leq T$*

$$\alpha(t, T) = \sigma(t, T) \cdot \int_t^T \sigma(t, s) ds - \sigma(t, T) \cdot g(t). \quad (3.6.3)$$

Dimostrazione.

Per la proposizione 2.6.1 sappiamo che dP deve essere della forma

$$\begin{aligned} dP(t, T) &= P(t, T) \left(r(t) + A(t, T) + \frac{1}{2} \|B(t, T)\|^2 \right) dt \\ &+ P(t, T) B(t, T) d\tilde{W}(t), \end{aligned}$$

dove

$$\begin{cases} A(t, T) = - \int_t^T \alpha(t, s) ds \\ B(t, T) = - \int_t^T \sigma(t, s) ds. \end{cases} \quad (3.6.4)$$

Lavorando con la filtrazione generata da \tilde{W} sappiamo che ogni misura $Q \sim P$ è ottenuta da una trasformazione di Girsanov e, denotando il kernel di Girsanov con g , otteniamo che, la dinamica dei prezzi dei bonds sotto Q , è data da

$$\begin{aligned} dP(t, T) &= P(t, T) \left(r(t) + A(t, T) + \frac{1}{2} \|B(t, T)\|^2 + B(t, T)g(t) \right) dt \\ &+ P(t, T) B(t, T) dW(t), \end{aligned}$$

dove W è un processo di Wiener sotto Q . Inoltre Q è una misura martingala se e solo se il tasso di rendimento locale di un T -bond eguaglia il tasso a breve r , cioè se e solo se

$$A(t, T) + \frac{1}{2} \|B(t, T)\|^2 + B(t, T)g(t) = 0 \quad \forall T > 0 \text{ e } \forall t \leq T.$$

Questa relazione vale identicamente per ogni T , quindi, derivando rispetto a T , otteniamo la condizione 3.6.3.

La condizione 3.6.2 garantisce che Q sia una misura di probabilità. \square

Riprendiamo l'equazione 3.6.3

$$\alpha(t, T) = \sum_{i=1}^d \sigma_i(t, T) \int_t^T \sigma_i(t, s) ds - \sum_{i=1}^d \sigma_i(t, T) g_i(t).$$

Questo è un sistema lineare di equazioni infinito dimensionale per determinare il vettore d -dimensionale $g(t)$, una volta date α e σ . Il sistema è sovradeterminato. Questo significa che non possiamo specificare a piacere tutti i termini $\alpha(t, T)$ e $\sigma(t, T)$, se vogliamo un mercato dei bonds privo di arbitraggio. Come fare? Possiamo schematizzare il procedimento come segue:

- Specifichiamo arbitrariamente la volatilità $\sigma(t, T)$ per ogni maturità T .

- Data $\sigma(t, T)$ come sopra, fissiamo d maturità di riferimento T_1, \dots, T_d . Solo per tali maturità specifichiamo anche i termini $\alpha(t, T_1), \dots, \alpha(t, T_d)$.
- Per ogni fissato t la matrice $(\sigma_i(t, T_j))_{i,j}$ sarà invertibile (nel caso generico) così il kernel di Girsanov $(g_1(t), \dots, g_d(t))$ è univocamente determinato (per ogni fissato t) come soluzione del sistema lineare di equazioni

$$\alpha(t, T_j) = \sum_{i=1}^d \sigma_i(t, T_j) \int_t^T \sigma_i(t, s) ds - \sum_{i=1}^d \sigma_i(t, T_j) g_i(t) \quad j = 1, \dots, d$$

- Ora i termini $\alpha(t, T)$ per le maturità differenti da T_1, \dots, T_d sono univocamente definiti dalla relazione 3.6.3.

In termini dei prezzi dei bonds, questo significa che possiamo esprimere la volatilità dei prezzi esattamente per d bonds presi come riferimento. I prezzi di tutti gli altri saranno determinati dalla struttura dei prezzi di quelli presi come riferimento e dalla richiesta di non arbitraggio. Questo è in accordo con il teorema 2.3.3: visto che, in questo caso, abbiamo d fonti aleatorie, ci si aspetta che il mercato sia completo e privo di arbitraggio se e solo se specifichiamo i prezzi per d bonds.

Una questione naturale è chiedersi, sotto quali condizioni, la matrice $(\sigma_i(t, T_j))_{i,j}$ risulti invertibile per ogni t . Ci sono dei risultati in questa direzione ma, oltre ad essere molto tecnici, esulano dallo scopo di questa tesi. Per una trattazione più approfondita, rimandiamo a [3, pagine 95–98].

3.6.3 Modellizzazione martingala del tasso futuro

Torniamo alla questione della modellizzazione martingala e supponiamo, quindi, di specificare la dinamica del tasso futuro direttamente sotto la misura martingala Q . Vista la relazione $r(t) = f(t, t)$, ogni modello basato sul tasso a breve può essere equivalentemente tradotto in termini del tasso futuro. E, come abbiamo già visto, se specifichiamo la dinamica di r sotto Q i prezzi dei T -bonds sono dati dalla formula

$$P(t, T) = E^Q \left[e^{-\int_t^T r(s) ds} \middle| \mathcal{F}_t \right].$$

Invece, nel caso della modellizzazione di $f(t, T)$ sotto Q , specificare il tasso futuro è equivalente a specificare i prezzi dei bonds, visto che vale la relazione

$$P(t, T) = e^{-\int_t^T f(t, s) ds}.$$

Affinché queste due formule valgano simultaneamente, dobbiamo imporre delle condizioni per $\alpha(t, T)$ e per $\sigma(t, T)$. Il risultato principale è il seguente:

Proposizione 3.6.2 *Assumiamo che Q sia una misura martingala per il mercato dei bonds e che la dinamica di $f(t, T)$ sotto Q sia data da*

$$df(t, T) = \alpha(t, T) dt + \sigma(t, T) dW(t),$$

dove W è un processo di Wiener d -dimensionale sotto Q . Allora:

1. Per ogni $T > 0$ e per ogni $t \leq T$

$$\alpha(t, T) = \sigma(t, T) \cdot \int_t^T \sigma(t, s) ds \quad Q - q.c.$$

2. La dinamica dei prezzi dei bonds sotto Q è data da

$$dP(t, T) = P(t, T)r(t) dt + P(t, T)B(t, T) dW(t),$$

dove $B(t, T)$ è dato dalla 3.6.4.

Dimostrazione.

Il fatto di modellizzare direttamente sotto una misura martingala implica che il kernel di Girsanov $g(t)$ è zero e dunque la tesi segue immediatamente dalla proposizione 3.6.1. \square

La morale della proposizione 3.6.2 è che, quando specifichiamo la dinamica di $f(t, T)$ (sotto Q), possiamo liberamente specificare la struttura della volatilità $\sigma(t, T)$ mentre il parametro $\alpha(t, T)$ viene già univocamente determinato dalla scelta di σ .

Osserviamo che, visto che $Q \sim P$, il processo della volatilità è il medesimo sia sotto Q che sotto P .

Una specie di algoritmo per un modello HJM può essere schematizzato così:

1. specificare la volatilità $\sigma(t, T)$ (questa è la parte della modellizzazione);
2. il parametro $\alpha(t, T)$ è univocamente determinato dalla relazione

$$\alpha(t, T) = \sigma(t, T) \cdot \int_t^T \sigma(t, s) ds;$$

3. osservare oggi ($t = 0$) dal mercato la curva dei tassi futuri

$$\{f^*(0, T), T \geq 0\};$$

4. integrare per ottenere

$$f(t, T) = f^*(0, T) + \int_0^t \alpha(s, T) ds + \int_0^t \sigma(s, T) \cdot dW(s);$$

5. calcolare i prezzi dei bonds tramite la formula

$$P(t, T) = e^{-\int_t^T f(t, s) ds};$$

6. passare ora alla valutazione dei derivati.

Capitolo 4

Analisi statistica della curva dei rendimenti

Perchè modellizzare l'intera struttura a termine dei tassi di interesse? In fondo non è necessario se dobbiamo calcolare il prezzo di una singola opzione call Europea su un bond (vedi pagina 23). Ma, se consideriamo il prezzo di un attivo casuale anche dei più semplici come un cap, un floor o uno swap, esso dipende dai pagamenti intermedi agli istanti T_1, \dots, T_n , dove n può essere anche 20 (per esempio uno swap a 10 anni con pagamenti semiannuali). La valutazione quindi di questi prodotti richiede la modellizzazione dell'intera struttura delle covarianze e la stima storica di una matrice così grande non è più trattabile da un punto di vista statistico. Occorre dunque imporre una struttura alla correlazione presente fra le varie quantità finanziarie della struttura a termine. Questa struttura può essere stabilita specificando la dinamica di poche variabili, ad esempio le componenti principali di cui tratteremo in questo capitolo.

4.1 Analisi in componenti principali di un vettore aleatorio

Lo scopo dell'utilizzo dell'analisi in componenti principali (in inglese *principal components analysis* da cui la sigla PCA) è quello di ridurre la dimensione di un insieme di dati correlati fra loro, cercando però di mantenerne più alta possibile la varianza. Questa riduzione è ottenuta tramite una trasformazione (lineare) delle variabili originali in un insieme di nuove variabili, le componenti principali appunto, che risultano fra loro non correlate. Inoltre vengono ordinate in modo tale che già le prime mantengano la maggior parte della varianza delle variabili originali.

Il calcolo delle componenti principali si riduce al calcolo degli autovettori (e corrispondenti autovalori) di una matrice reale, simmetrica e semidefinita positiva. Entriamo nel dettaglio.

Supponiamo che X sia un vettore di k variabili aleatorie reali e che, per i nostri scopi, sia di particolare importanza tener conto della loro struttura delle covarianze. A meno che il numero k non sia piccolo, non è di aiuto guardare semplicemente alle k varianze o a tutte le $\frac{1}{2}k(k-1)$ covarianze. Si vorrebbe tener conto solo di d variabili con $d \ll k$, ottenute come combinazioni lineari delle variabili X_1, \dots, X_k , che riescano in ogni caso a mantenere la maggior parte delle informazioni contenute nelle varianze e covarianze delle variabili originali.

Il primo passo è quello di cercare un vettore v appartenente a \mathbb{R}^k di norma unitaria che renda massima la varianza

$$\text{Var}[\langle v, X \rangle],$$

dove $\langle v, X \rangle = \sum_{i=1}^k v_i X_i$.

Introduciamo la matrice C delle covarianze (che supponiamo invertibile), definita come

$$C_{ij} = \text{Cov}[X_i, X_j] \quad i, j = 1, \dots, k.$$

Osserviamo che C è una matrice $k \times k$ reale, simmetrica e semidefinita positiva, quindi possiamo trovare una matrice ortogonale A (cioè $A^T = A^{-1}$) tale che $C = ADA^T$, dove con D denotiamo la matrice diagonale di elementi $\lambda_1, \dots, \lambda_k$. Ovviamente le colonne a_1, \dots, a_k della matrice A sono gli autovettori normalizzati di C , mentre $\lambda_1 \geq \dots \geq \lambda_k \geq 0$ sono i corrispondenti autovalori che supponiamo essere ordinati in modo decrescente (in caso contrario è sufficiente permutare le righe di C).

La covarianza $\text{Cov}[\langle v, X \rangle, \langle w, X \rangle]$, sfruttando la linearità, può essere scritta nella forma

$$\begin{aligned} \text{Cov}\left[\sum_{i=1}^k v_i X_i, \sum_{j=1}^k w_j X_j\right] &= \sum_{i,j=1}^k v_i w_j \text{Cov}[X_i, X_j] \\ &= \sum_{i,j=1}^k v_i w_j C_{ij} \\ &= \langle Cv, w \rangle, \end{aligned}$$

da cui si ricava la forma quadratica

$$\text{Var}[\langle v, X \rangle] = \text{Cov}[\langle v, X \rangle, \langle v, X \rangle] = \langle Cv, v \rangle.$$

Quindi

$$\max_{\|v\|=1} \text{Var}[\langle v, X \rangle] = \max_{\|v\|=1} \langle Cv, v \rangle = \langle Ca_1, a_1 \rangle = \lambda_1.$$

Chiameremo $p_1 = \langle a_1, X \rangle$ la *prima componente principale*, $p_2 = \langle a_2, X \rangle$ la *seconda componente principale* e così via, mentre le componenti dell'autovettore a_i verranno chiamati i *pesi relativi alla i -esima componente principale*.

Verifichiamo che effettivamente le nuove variabili p_i sono non correlate: ponendo $p = (p_1, \dots, p_k)^T$ possiamo scrivere $p = A^T X$ da cui

$$\text{Cov}[p_i, p_j] = \sum_{p,q=1}^k A_{ip}^T \text{Cov}[X_p, X_q] A_{jq}^T = (A^T C A)_{ij} = \lambda_i \delta_{ij},$$

dove $\delta_{i,j}$ è il simbolo di Kronecker. In particolare ricaviamo che $\text{Var}[p_i] = \lambda_i$.

Osserviamo che la varianza totale viene preservata essendo la matrice A della trasformazione ortogonale:

$$\sum_{i=1}^k \text{Var}[X_i] = \sum_{i=1}^k C_{ii} = \sum_{i=1}^k \sum_{j=1}^k \lambda_j A_{ij}^2 = \sum_{i=1}^k \lambda_i = \sum_{i=1}^k \text{Var}[p_i].$$

La varianza della variabile p_i , come già osservato, è l'autovalore λ_i , la cui grandezza determina l'importanza della variabile p_i stessa. La chiave statistica infatti è il rapporto

$$\nu_i = \frac{\lambda_i}{\sum_{j=1}^k \lambda_j}. \quad (4.1.1)$$

Esso rappresenta la porzione della varianza totale espressa dall' i -esima componente principale.

Poniamo ora $\tilde{w} = (D^{1/2})^{-1} p$ e $w = \tilde{w} - \mu[\tilde{w}]$, dove con $D^{1/2}$ denotiamo la matrice diagonale di elementi $\sqrt{\lambda_1}, \dots, \sqrt{\lambda_k}$ e con $\mu[\tilde{w}_i]$ la media di \tilde{w}_i . Allora risulta

$$\mu[w] = 0 \quad \text{e} \quad \text{Cov}[w_i, w_j] = \text{Cov}[\tilde{w}_i, \tilde{w}_j] = \delta_{ij}.$$

Utilizzando le quantità appena introdotte possiamo scrivere il vettore X nella forma

$$\begin{aligned} X &= Ap = AD^{1/2} \tilde{w} = AD^{1/2} (w + \mu[\tilde{w}]) = \mu[X] + AD^{1/2} w \\ &= \mu[X] + \sum_{j=1}^k a_j \sqrt{\lambda_j} w_j, \end{aligned}$$

in componenti

$$X_i = \mu[X_i] + \sum_{j=1}^k A_{ij} \sqrt{\lambda_j} w_j.$$

Se k è molto grande risulterà difficile avere un'idea del comportamento delle variabili X_1, \dots, X_k . Comunque, se gli autovalori λ_j della matrice delle covarianze C sono molto vicini a zero per $j = d+1, \dots, k$ per qualche numero $d < k$, possiamo affermare che, in un certo senso, la "vera dimensione" del vettore aleatorio X è d . In questo modo il vettore X può essere accuratamente approssimato da

$$X \approx \mu[X] + \sum_{j=1}^d a_j \sqrt{\lambda_j} w_j. \quad (4.1.2)$$

Sebbene X sia a valori in uno spazio di dimensione k , in effetti vive nel sottospazio affine generato solo dai primi d autovettori a_1, \dots, a_d . Questa sarà esattamente la situazione che ci si presenterà quando applicheremo l'analisi in componenti principali alla curva dei tassi futuri. In questo caso, k corrisponderà al numero delle date di maturità dei tassi futuri presi in considerazione per le quali saranno disponibili delle osservazioni. Vedremo che gli autovalori della matrice delle covarianze decrescono rapidamente così l'effettiva dimensione sarà circa tre o quattro.

La seguente scrittura sarà utile in seguito: poniamo

$$\sigma_i := \sqrt{\text{Var}[X_i]} = (C_{ii})^{1/2} = \left(\sum_{j=1}^k A_{ij}^2 \lambda_j \right)^{1/2}$$

$$v_i := \frac{X_i - \mu[X_i]}{\sigma_i} = \frac{\sum_{j=1}^k A_{ij} \sqrt{\lambda_j} w_j}{\sigma_i} \quad i = 1, \dots, k$$

in modo da ottenere

$$\mu[v_i] = 0 \quad \text{Var}[v_i] = 1 \quad X_i = \mu[X_i] + \sigma_i v_i$$

e

$$\text{Corr}[X_i, X_j] = \text{Cov}[v_i, v_j] = \frac{C_{ij}}{\sigma_i \sigma_j} = \frac{\sum_{m=1}^k A_{im} A_{jm} \lambda_m}{\sigma_i \sigma_j}. \quad (4.1.3)$$

4.2 PCA della curva dei rendimenti

Supponiamo che X sia un vettore aleatorio di dimensione k . Ogni componente X_i potrebbe rappresentare un tasso a termine con una certa maturità T_i , in modo tale che l'intero vettore X sia in grado di descrivere, in modo abbastanza esauriente, l'intera struttura a termine della curva dei rendimenti.

Le osservazioni, per esempio giornaliere, delle variabili X_1, \dots, X_k vengono raccolte in una matrice $\mathbf{x} = (x_{ij})$ di grandi dimensioni, in cui la riga $x^{(i)} = (x_{i1}, \dots, x_{ik})$ è una realizzazione del vettore aleatorio $X^{(i)}$. La manipolazione dei dati prevista dalla PCA è dovuta ad alcuni risultati teorici. Se $(X^{(n)})_{n \geq 1}$ è una successione di variabili aleatorie reali isonome (nel nostro caso a valori in $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$) la legge forte dei grandi numeri assicura che le stime

$$\mu^N = \frac{1}{N} \sum_{i=1}^N X^{(i)}$$

e

$$C_{lm}^N = \frac{1}{N-1} \sum_{i=1}^N (X_l^{(i)} - \mu_l^N)(X_m^{(i)} - \mu_m^N)$$

convergono quasi certamente per $N \rightarrow +\infty$ al vettore delle medie μ e agli elementi (l, m) della matrice delle covarianze C di una qualsiasi variabile aleatoria isonoma alle $X^{(i)}$. Per garantire questa convergenza però, viene richiesta, oltre all'integrabilità delle $X^{(i)}$, la loro indipendenza.

In pratica sono i dati contenuti nella matrice \mathbf{x} ad essere usati per calcolare stime empiriche $\hat{\mu}^N$ e \hat{C}^N di μ^N e C^N . Tipicamente sono calcolate come

$$\hat{\mu}^N = \frac{1}{N} \sum_{i=1}^N x^{(i)}$$

e

$$\hat{C}_{lm}^N = \frac{1}{N-1} \sum_{i=1}^N (x_l^{(i)} - \hat{\mu}_l^N)(x_m^{(i)} - \hat{\mu}_m^N)$$

e la legge dei grandi numeri assicura che, se la taglia N del campione è sufficientemente grande, le stime empiriche $\hat{\mu}^N$ e \hat{C}^N sono abbastanza vicine ai valori μ del comune vettore delle medie delle $X^{(i)}$ e C della comune matrice delle covarianze.

Nelle nostre applicazioni l'ipotesi di indipendenza delle $X^{(i)}$ corrisponde all'indipendenza dei tassi a termine da un giorno all'altro e questo è piuttosto inverosimile. Per evitare questo problema si può procedere in diversi modi. Uno di questi consiste nell'assumere che la successione $(X^{(n)})_{n \geq 1}$ sia stazionaria, ovvero che la legge del blocco $[X^{(n)}]_{n \geq 1}$ sia identica a quella del blocco $[X^{(n+1)}]_{n \geq 1}$. Sotto tale ipotesi rimane la convergenza quasi certa delle stime μ^N e C^N verso i valori μ e C del vettore delle medie e della matrice delle covarianze, sebbene la velocità di convergenza non sia la stessa.

In seguito quindi supporremo implicitamente che le righe $(x^{(n)})$ della matrice \mathbf{x} siano realizzazioni della successione stazionaria $(X^{(n)})$. Per interpretare i risultati della PCA dobbiamo ricordarci di verificare che il modello utilizzato abbia questa proprietà. In ogni caso, quando la stazionarietà è in dubbio, il metodo standard consiste nel “differenziare” le osservazioni e calcolare la PCA degli incrementi giornalieri dei tassi di interesse. Vediamo come questo può essere fatto. Supponiamo di avere delle osservazioni dei tassi a termine che coprono un periodo di almeno alcuni anni e più precisamente, agli istanti t_i per $i = 1, \dots, N$, abbiamo le osservazioni delle k variabili aleatorie $r_{t_i}(T_j)$ per $j = 1, \dots, k$. Poniamo

$$x_{ij} = r_{t_{i+1}}(T_j) - r_{t_i}(T_j) \quad i = 1, \dots, N \quad j = 1, \dots, k.$$

Esso rappresenta la i -esima osservazione della variabile aleatoria X_j . X_{ij} rappresenta l'incremento fra gli istanti t_i e t_{i+1} del tasso a termine $r(T_j)$ con maturità T_j . Applicando l'analisi in componenti principali al vettore X si ricavano i pesi a_1, \dots, a_k relativi alle k componenti principali. Prima di essere utilizzati, i dati sono stati manipolati, o meglio estrapolati (si veda pagina 14) ed interpolati (si veda pagina 15), per garantire ogni giorno osservazioni dei tassi con le stesse maturità.

Riportiamo qui di seguito alcuni esempi.

4.3 Esempi

PCA dei tassi governativi statunitensi

In [5] sono presi in considerazione i dati sulle osservazioni giornaliere dei tassi governativi statunitensi con maturità a 1, 3 e 6 mesi e 1, 2, 3, 5, 7, 10, 20 e 30 anni. I dati coprono un periodo che va dal 01/10/1993 al 12/08/2005. Le maturità estreme creano qualche problema: le osservazioni dei tassi a breve termine infatti non sono sempre disponibili, mentre il ritiro nel 2001 dei bonds a lunga durata rende difficoltoso estrapolare i dati per le maturità a 30 anni. Per questo motivo si è applicata l'analisi in componenti principali a due sottoperiodi, dividendo i dati in due sottoinsiemi disgiunti.

<i>Periodo 93-01</i>			<i>Periodo 01-05</i>	
Comp. princ.	Var. espressa (%)	Var. cumulativa (%)	Var. espressa (%)	Var. cumulativa (%)
1	79.7	79.7	82.5	82.5
2	19.0	98.7	9.5	92.0
3	1.0	99.7	3.8	95.8
4	0.1	99.8	1.8	97.6
5	0.1	99.9	0.9	98.5
6	0.1	100	0.5	99.0
7	0.0	100	0.3	99.3
8	0.0	100	0.3	99.6
9	0.0	100	0.2	99.8
10	0.0	100	0.2	100

Tabella 4.1: Varianza espressa dalle dieci componenti della PCA dei tassi governativi statunitensi

Il primo periodo va dal 01/10/1993 al 31/07/2001 e vengono considerate, in questo caso, le 10 maturità $T = 0.25, 0.50, 1, 2, 3, 5, 7, 10, 20, 30$ anni. Nella tabella 4.1 vengono riportate, in percentuale, le proporzioni ν_j date dalla 4.1.1, relative alle 10 componenti principali calcolate. Nella figura 4.1 sono disegnati i grafici dei pesi relativi alle prime quattro componenti principali.

Consideriamo ora i dati dal 01/08/2001 al 12/08/2005. Questo periodo copre 1008 giorni lavorativi e viene utilizzato lo stesso numero di maturità del periodo precedente: è stata eliminata quella a 30 anni e introdotta quella

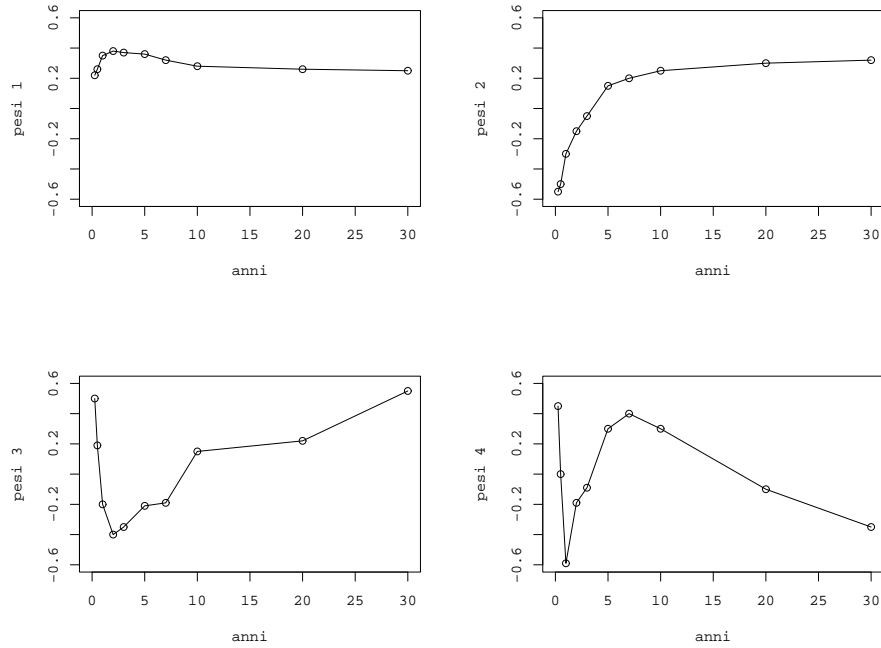


Figura 4.1: Grafici dei pesi relativi alle prime quattro componenti principali. Dati statunitensi, periodo 93–01

a 1 mese. I risultati sono pressoché gli stessi e sono riportati sempre nella tabella 4.1 e nella figura 4.2.

PCA dei tassi a termine del Regno Unito

In [12, capitolo 2] è stata applicata l'analisi in componenti principali ai dati provenienti dal mercato del Regno Unito. Il periodo preso in considerazione copre circa 3 anni, dal 1989 al 1992, e l'intero spettro delle maturità è stato diviso in 8 parti, ottenendo le osservazioni del vettore aleatorio X le cui componenti X_i sono gli incrementi del tasso futuro $R(t, S, T)$:

$$X_i = R(t + \Delta t, t + \Delta t + \tau_{i-1}, t + \Delta t + \tau_i) - R(t, t + \tau_{i-1}, t + \tau_i),$$

dove $0 = \tau_0 < \tau_1 < \dots < \tau_8$. I risultati così ottenuti sono riportati nelle tabelle 4.2 e 4.3 e nella figura 4.3.

PCA dei tassi a termine del Regno Unito (Nunes e Webber)

Nel 1997 Nunes e Webber investigarono sulla struttura dei tassi a termine del Regno Unito, utilizzando i dati osservati dal Settembre 1988 al febbraio

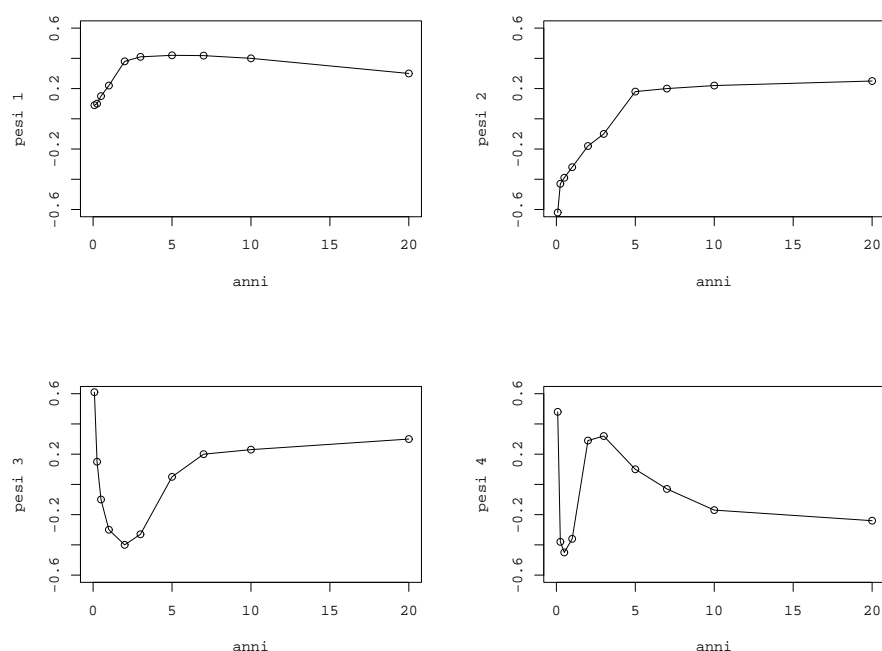


Figura 4.2: Grafici dei pesi relativi alle prime quattro componenti principali. Dati statunitensi, periodo 01–05

Comp. princ.	Var. espressa (%)	Var. cumulativa (%)
1	92.170	92.170
2	6.930	99.100
3	0.614	99.714
4	0.240	99.954
5	0.031	99.985
6	0.004	99.989
7	0.002	99.991
8	0.000	99.995

Tabella 4.2: Varianza espressa dalle otto componenti della PCA dei tassi del Regno Unito, periodo 89–92

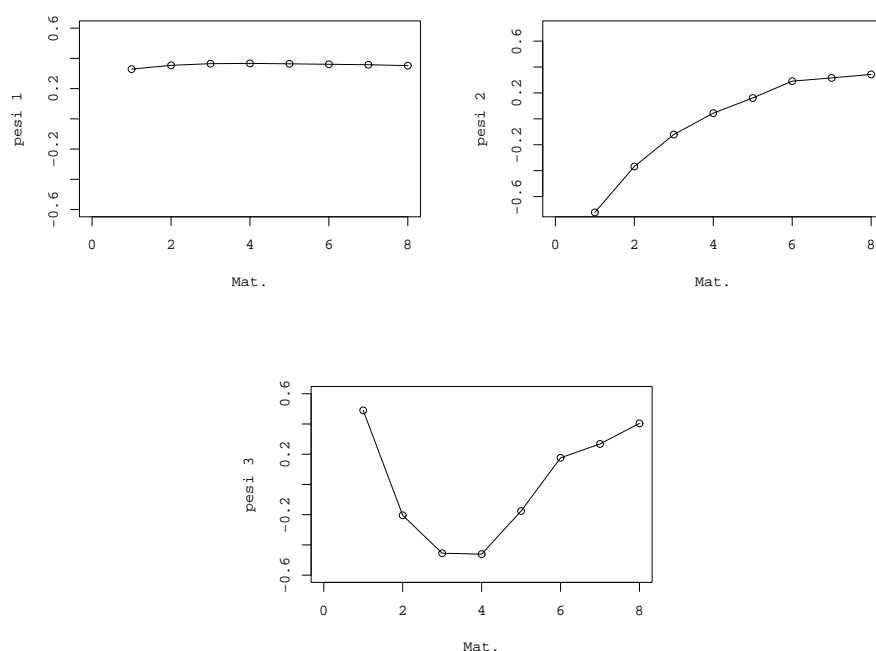


Figura 4.3: Grafici dei pesi relativi alle prime tre componenti principali. Dati del Regno Unito, periodo 89–92

Maturità	a_1	a_2	a_3
1	0.329	-0.722	0.49
2	0.354	-0.368	-0.204
3	0.365	-0.121	-0.455
4	0.367	0.044	-0.461
5	0.364	0.161	-0.176
6	0.361	0.291	0.176
7	0.358	0.316	0.268
8	0.352	0.343	0.404

Tabella 4.3: Pesi a_1, a_2, a_3 relativi alle prime tre componenti principali. Dati del Regno Unito, periodo 89–92

1995. I dati riguardavano il tasso LIBOR a 3, 6 e 12 mesi e i tassi swap a 2, 3, 4, 5, 7 e 10 anni, per un totale di 1609 osservazioni di un vettore con 9 componenti. Gli autovettori e le relative varianze sono riportati nella tabella 4.4 e 4.5, mentre i grafici nella figura 4.4. Questo esempio è stato tratto da [8, capitolo 16].

Maturità	a_1	a_2	a_3
3 mesi	-0.3158	0.4693	-0.7280
6 mesi	-0.3439	0.4747	0.0636
1 anno	-0.3748	0.4251	0.6772
2 anni	-0.3531	-0.1542	-0.0512
3 anni	-0.3423	-0.1962	-0.0596
4 anni	-0.3354	-0.2463	-0.218
5 anni	-0.3309	-0.2634	-0.0157
7 anni	-0.3057	-0.2957	-0.0021
10 anni	-0.2903	-0.3065	0.0201
Autovalore	0.0400	0.0128	0.0037

Tabella 4.4: Pesi a_1, a_2, a_3 e corrispondenti autovalori relativi alle prime tre componenti principali. Dati del Regno Unito, periodo 88–95

4.4 Interpretazione e discussione dei risultati

In generale, maggiore è la correlazione fra le variabili originali, maggiore sarà la varianza espressa dalla prima componente principale e più efficiente risulta la trasformazione in componenti principali. Molti studi empirici sono stati fatti sul comportamento delle serie storiche dei tassi futuri con differenti maturità e sulla loro correlazione. Per esempio, Brown e Schaefer analizzarono i dati dei tassi statunitensi dal 1987 al 1994 e osservarono che il grado di correlazione fra questi tassi decresceva con l'aumentare della distanza fra le maturità ma non con velocità costante (si veda tabella 4.6 e figura 4.5).

Come si può notare anche da tutti i grafici dei pesi riportati in queste pagine, le componenti principali ottenute da una PCA sono state osservate avere le seguenti caratteristiche:

1. I pesi a_{i1} relativi alla prima componente principale hanno stesso segno e modulo pressoché eguale, conferendo al loro grafico un andamento piatto. Possono quindi essere interpretati come il “livello medio” della curva dei rendimenti.
2. I pesi a_{i2} relativi alla seconda componente hanno approssimativamente

Comp. princ.	Var. espressa (%)	Var. cumulativa (%)
1	64.6	64.6
2	20.7	85.3
3	5.9	91.3
4	3.6	94.9
5	2.9	97.8
6	0.9	98.7
7	0.5	99.3
8	0.4	99.7
9	0.3	100.0

Tabella 4.5: Varianza espressa dalle nove componenti della PCA dei tassi del Regno Unito, periodo 88–95

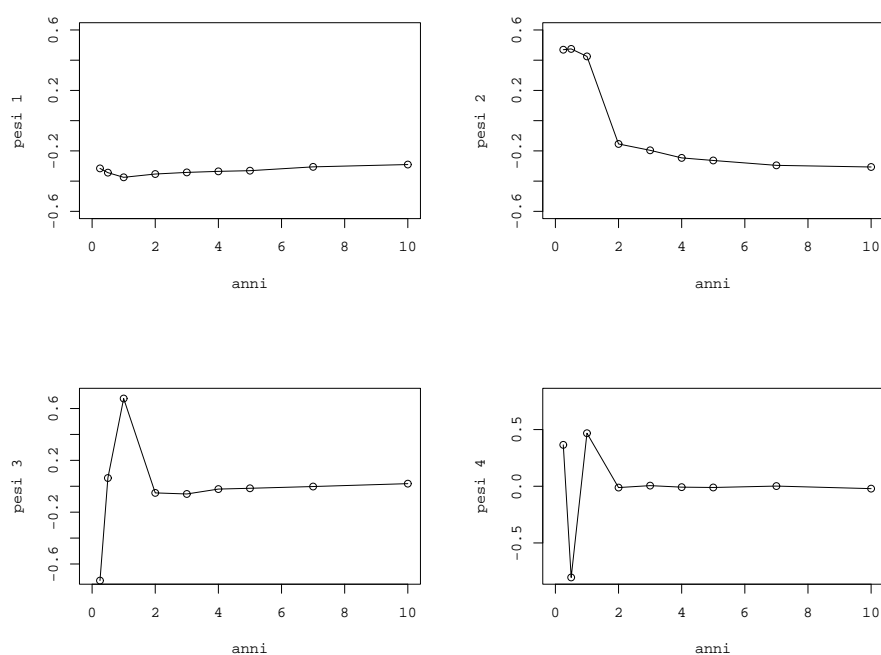


Figura 4.4: Grafici dei pesi relativi alle prime quattro componenti principali. Dati del Regno Unito, periodo 88–95

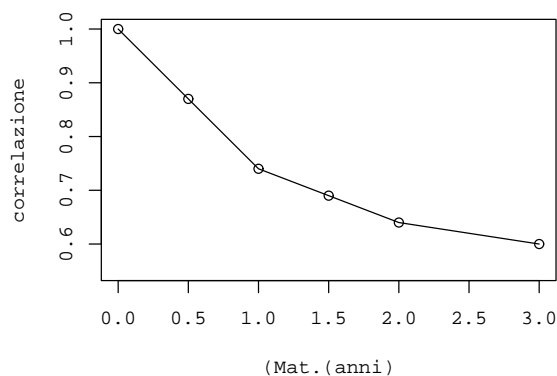


Figura 4.5: Correlazione fra il tasso a breve e il tasso futuro con diverse maturità

lo stesso modulo ma segno opposto agli estremi dell'intero spettro delle maturità. Vengono interpretati come la “pendenza” della curva.

3. La terza componente è costituita dai pesi a_{i3} di egual segno e quasi stesso modulo agli estremi dello spettro delle maturità, mentre nel centro da pesi di segno opposto. Il loro grafico suggerisce che, la terza componente principale, catturi la “curvatura” della curva dei rendimenti.

I grafici dei pesi relativi alle successive componenti risultano molto disturbati e sembrano non avere alcuna interpretazione. Osserviamo del resto che solo

Maturità (anni)	0.00	0.50	1.00	1.50	2.00	3.00
0.00	1.00	0.87	0.74	0.69	0.64	0.60
0.50	0.87	1.00	0.96	0.93	0.90	0.85
1.00	0.74	0.96	1.00	0.99	0.95	0.92
1.50	0.69	0.93	0.99	1.00	0.97	0.93
2.00	0.64	0.90	0.95	0.97	1.00	0.95
3.00	0.60	0.85	0.92	0.93	0.95	1.00

Tabella 4.6: Grado di correlazione fra gli incrementi dei tassi futuri statunitensi con differenti maturità. Periodo 87–94

la prima componente esprime già una varianza pari all'80–90% di quella totale e che le prime tre insieme arrivano ad esprimere fino al 95–99%.

Questi dati suggeriscono che l'effettiva “dimensione” della curva dei rendimenti è circa tre e che, se la approssimiamo con opportune combinazioni lineari costituite dalle sole prime tre componenti, l'errore commesso è veramente piccolo e la perdita di informazioni è minima. Attenzione però, la perdita di informazioni relative alla varianza è minima. Per quanto riguarda la struttura della correlazione? La perdita di informazioni risulta maggiore, non a caso la PCA si concentra più sugli elementi diagonali della matrice delle covarianze che non sugli altri. Per focalizzare meglio e in concreto la discussione riportiamo un esempio ([12, capitolo 2]).

Faremo due ipotesi:

1. la curva dei rendimenti viene interamente descritta da k tassi futuri a n mesi;
2. la correlazione ρ fra i tassi futuri con maturità T_1 e T_2 è data da

$$\rho(T_1, T_2) = \exp(-\beta(T_2 - T_1)). \quad (4.4.1)$$

Quest'ultima ipotesi è in accordo con i dati riportati nella tabella 4.6. I dati osservati provengono dal mercato statunitense e riguardano 10 tassi futuri X_i a 6 mesi (coprendo così uno spettro delle maturità pari a 5 anni), con il primo tasso che matura in 2 anni. Questi sono i tassi che entrano in gioco, per esempio, nel calcolo del prezzo di un qualsiasi swaption (si veda pagina 17) che matura in 2 anni con un tenore di 5.

Dopo aver imposto una struttura alla correlazione del tipo 4.4.1 si può applicare una PCA ai dati utilizzando la matrice delle covarianze

$$\text{Cov}[X_i, X_j] = \sigma_i \sigma_j \exp(-\beta(T_j - T_i)) \quad j > i,$$

ottenuta utilizzando le volatilità σ_i e σ_j estratte dal mercato. Gli autovettori e i corrispondenti autovalori della matrice risultante sono riportati nella tabella 4.7 e 4.8 e sono stati ottenuti ponendo $\rho = 0.85$.

I grafici dei pesi relativi alle prime tre componenti principali (figura 4.7), anche se ottenuti da una matrice nota a priori e non ottenuta da una serie di dati disturbati estratti dal mercato, confermano le interpretazioni di livello medio, pendenza e curvatura.

Possiamo ora approssimare il vettore X , le cui componenti sono i dieci tassi futuri X_i , con combinazioni lineari delle sole prime d componenti principali, con $d \ll 10$:

$$X \approx \mu[X] + \sum_{j=1}^d a_j \sqrt{\lambda_j} w_j.$$

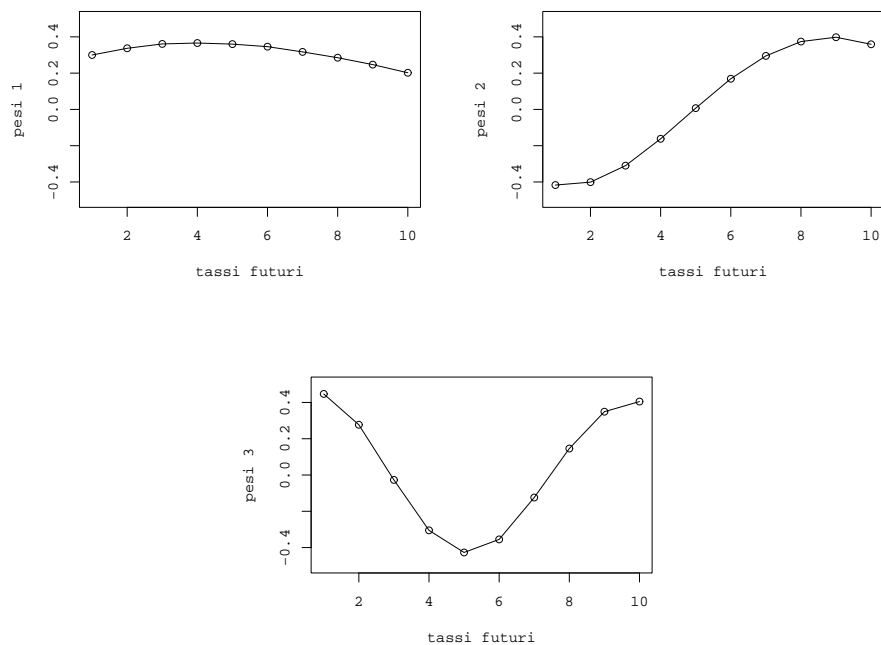


Figura 4.6: Grafici dei pesi a_1, a_2, a_3 relativi alle prime tre componenti principali ottenuti con $\rho = 0.85$

Maturità (anni)	a_1	a_2	a_3
2.0	0.300	-0.417	0.447
2.5	0.337	-0.401	0.277
3.0	0.361	-0.310	-0.027
3.5	0.366	-0.162	-0.305
4.0	0.360	0.007	-0.427
4.5	0.346	0.169	-0.355
5.0	0.317	0.295	-0.124
5.5	0.285	0.374	0.146
6.0	0.247	0.398	0.349
6.5	0.202	0.359	0.405
Autovalore	0.316	0.118	0.051

Tabella 4.7: Pesì a_1, a_2, a_3 e corrispondenti autovalori relativi alle prime tre componenti principali ottenuti con $\rho = 0.85$

Comp. princ.	Var. espressa (%)	Var. cumulativa (%)
1	74.4	74.4
2	12.7	87.1
3	4.9	92.0
4	2.6	94.6
5	1.6	96.2
6	1.1	97.3
7	0.8	98.1
8	0.7	98.8
9	0.7	99.5
10	0.5	100.0

Tabella 4.8: Varianza espressa dalle dieci componenti della PCA dei tassi statunitensi ottenuta con $\rho = 0.85$

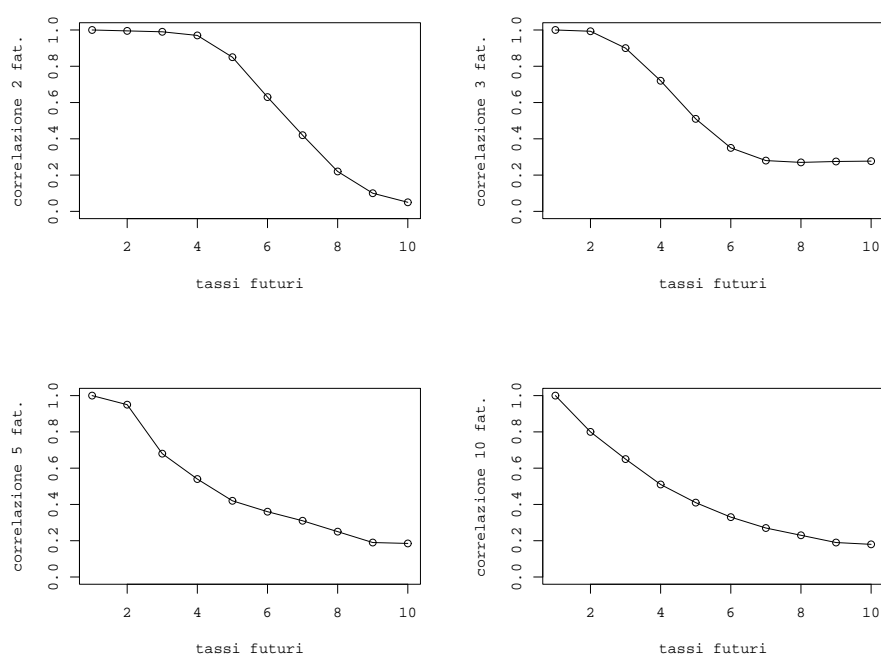


Figura 4.7: Grafici della correlazione fra il primo tasso futuro e i successivi per $d=2, 3, 5$ e 10

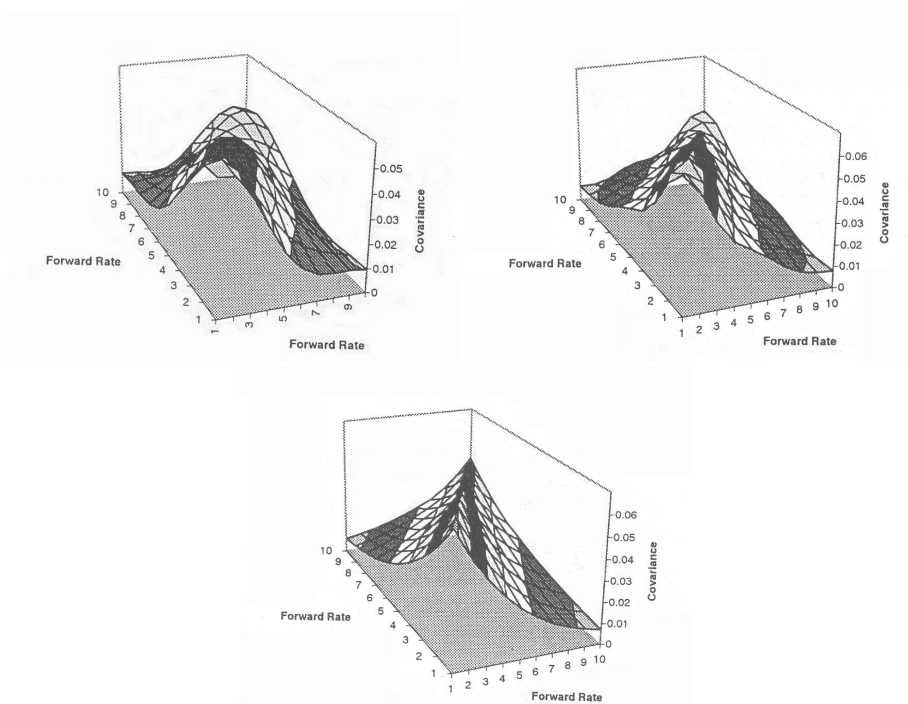


Figura 4.8: Superficie della correlazione per $d=3, 5$ e 10

La correlazione calcolata usando solo d componenti (o fattori) è data dalla 4.1.3 arrestando la somma a d . Nella figura 4.7 viene riportata la correlazione fra il primo tasso e i successivi per diversi valori di d , mentre nella figura 4.8 viene riportata l'intera superficie della correlazione. Come si può notare, la forma della funzione correlazione cambia in base al numero di fattori utilizzati per l'approssimazione. Utilizzando pochi fattori, la correlazione risulta concava in corrispondenza delle maturità che si trovano all'inizio dello spettro, mentre è convessa per le lunghe maturità. Già con 5 fattori la forma mostra una convessità positiva per quasi tutte le maturità. Per ottenere dunque una correlazione di tipo esponenziale come abbiamo imposto all'inizio occorrono circa 5 o 6 fattori contro i 2 o 3 suggeriti dalla PCA.

4.5 Metodi numerici

Concludiamo questo capitolo presentando brevemente un metodo numerico per il calcolo degli autovettori e autovalori di una matrice.

Ipotizziamo che gli autovalori siano tutti distinti, un'ipotesi ragionevole data la natura empirica della matrice delle covarianze. Poiché tale matrice è simmetrica e definita positiva, essa possiede tutte le caratteristiche di cui

abbiamo bisogno per poter applicare il *metodo delle potenze*. L'autovettore di modulo maggiore viene calcolato tramite una procedura iterativa. Data la matrice delle covarianze C di dimensione k e denotati con $\lambda_1, \dots, \lambda_k$ i suoi autovalori ordinati in modo decrescente, fissiamo un vettore $\mathbf{t}_0 \in \mathbb{R}^k$ e generiamo la successione $\{\mathbf{y}_i\}$, $i = 1, 2, \dots$, così definita

$$\begin{aligned} \mathbf{y}_0 &= \mathbf{t}_0 \\ \mathbf{y}_i &= C\mathbf{y}_{i-1}, \quad i = 1, 2, \dots \end{aligned} \quad (4.5.1)$$

Poiché tutti gli autovettori $\mathbf{a}_1, \dots, \mathbf{a}_k$ sono linearmente indipendenti, il vettore \mathbf{t}_0 può essere espresso per mezzo della combinazione lineare

$$\mathbf{t}_0 = \sum_{j=1}^k \alpha_j \mathbf{a}_j$$

e si supponga scelto in modo tale che α_1 sia diverso da zero; risulta quindi

$$\begin{aligned} \mathbf{y}_i &= C^i \mathbf{t}_0 = \sum_{j=1}^k \alpha_j C^i \mathbf{a}_j = \sum_{j=1}^k \alpha_j \lambda_j^i \mathbf{a}_j \\ &= \lambda_1^i \left[\alpha_1 \mathbf{a}_1 + \sum_{j=2}^k \alpha_j \left(\frac{\lambda_j}{\lambda_1} \right)^i \mathbf{a}_j \right]. \end{aligned} \quad (4.5.2)$$

Indicate con $y_r^{(i)}$ e $a_r^{(j)}$ le r -esime componenti dei vettori \mathbf{y}_i e \mathbf{a}_j , si ha, per gli indici h per cui le componenti $y_h^{(i)}$ e $a_h^{(1)}$ siano diverse da zero

$$\frac{y_h^{(i+1)}}{y_h^{(i)}} = \lambda_1 \frac{\alpha_1 a_h^{(1)} + \sum_{j=2}^k \alpha_j \left(\frac{\lambda_j}{\lambda_1} \right)^{i+1} a_h^{(j)}}{\alpha_1 a_h^{(1)} + \sum_{j=2}^k \alpha_j \left(\frac{\lambda_j}{\lambda_1} \right)^i a_h^{(j)}}, \quad (4.5.3)$$

e poiché $|\lambda_j/\lambda_1| < 1$ per $j \geq 2$ si ha

$$\lim_{i \rightarrow +\infty} \frac{y_h^{(i+1)}}{y_h^{(i)}} = \lambda_1.$$

Quindi da un certo indice i in poi l'autovalore λ_1 può essere approssimato mediante uno dei rapporti $y_h^{(i+1)}/y_h^{(i)}$.

Con questo metodo si può approssimare anche l'autovettore \mathbf{a}_1 . Dalla 4.5.2 risulta infatti

$$\lim_{i \rightarrow +\infty} \frac{\mathbf{y}_i}{\lambda_1^i} = \alpha_1 \mathbf{a}_1$$

e quindi, per $h = 1, \dots, k$ per cui la componente $a_h^{(1)}$ sia diversa da zero, si ha

$$\lim_{i \rightarrow +\infty} \frac{y_h^{(i+1)}}{\lambda_1^i} = \alpha_1 a_h^{(1)} \quad \text{e} \quad \lim_{i \rightarrow +\infty} \frac{\mathbf{y}_i}{y_h^{(i)}} = \frac{\mathbf{a}_1}{a_h^{(1)}}.$$

Poiché per i sufficientemente elevato l'indice m di una componente di massimo modulo rimane costante, la successione $\mathbf{y}_i/y_m^{(i)}$ converge all'autovettore normalizzato in norma ∞ . Ricordiamo che $\|\mathbf{a}_1\|_\infty := \max_{h=1,\dots,k} a_h^{(1)}$

Questo metodo richiede ad ogni passo il calcolo del prodotto della matrice C per un vettore: se C non è sparsa come nel nostro caso, ogni passo richiede k^2 operazioni moltiplicative. Operando però in aritmetica finita, con la 4.5.1 dopo pochi passi si possono presentare condizioni di underflow o di overflow (i numeri trovati sono rispettivamente troppo piccoli o troppo grandi per essere rappresentati come numeri in virgola mobile). Per evitare che ciò accada è necessario eseguire ad ogni passo la normalizzazione del vettore ottenuto, costruendo una successione $\{\mathbf{t}_i\}$ per $i = 1, 2, \dots$ così definita

$$\begin{aligned} \mathbf{u}_i &= C\mathbf{t}_{i-1} \\ \mathbf{t}_i &= \frac{\mathbf{u}_i}{\|\mathbf{u}_i\|_2}, \quad i = 1, 2, \dots \end{aligned} \quad (4.5.4)$$

dove \mathbf{t}_0 è scelto in modo tale che $\|\mathbf{t}_0\|_2 = 1$.

Si ha allora

$$\mathbf{t}_i = \frac{1}{\gamma_i} \mathbf{y}_i = \frac{1}{\gamma_i} C^i \mathbf{t}_0 \quad \text{dove} \quad \gamma_i = \prod_{j=1}^i \|\mathbf{u}_j\|_2,$$

e poiché

$$\mathbf{u}_{i+1} = \frac{1}{\gamma_i} C^{i+1} \mathbf{t}_0,$$

operando come nella 4.5.3 si ha che il rapporto fra le h -esime componenti di \mathbf{u}_{i+1} e \mathbf{t}_i , per gli indici h per cui $t_h^{(i)}$ e $a_h^{(i)}$ siano diverse da zero, è dato da

$$\frac{u_h^{(i+1)}}{t_h^{(i)}} = \lambda_1 \frac{\alpha_1 a_h^{(1)} + \sum_{j=2}^k \alpha_j \left(\frac{\lambda_j}{\lambda_1}\right)^{i+1} a_h^{(j)}}{\alpha_1 a_h^{(1)} + \sum_{j=2}^k \alpha_j \left(\frac{\lambda_j}{\lambda_1}\right)^i a_h^{(j)}}.$$

Quindi

$$\lim_{i \rightarrow +\infty} \frac{u_h^{(i+1)}}{t_h^{(i)}} = \lambda_1.$$

Per la normalizzazione del vettore \mathbf{u}_i è possibile utilizzare anche altre norme vettoriali, come per esempio la norma ∞ . Se la matrice C può essere scritta nella forma $C = ADA^T$ con A ortogonale e D diagonale come nel nostro caso, è più conveniente utilizzare la norma 2 in quanto si può dimostrare che la convergenza all'autovalore λ_1 è molto più veloce (rispetto alla norma ∞) e in particolare che

$$\sigma_i = \mathbf{t}_i^T \mathbf{u}_{i+1} = \lambda_1 \left[1 + O\left(\frac{\lambda_2}{\lambda_1}\right)^{2i} \right].$$

Come criterio di arresto si può utilizzare la condizione

$$\|\mathbf{u}_{i+1} - \sigma_i \mathbf{t}_i\|_2 < \varepsilon$$

che, oltre ad essere facilmente applicabile, fornisce una maggiorazione per l'errore assoluto: infatti si può dimostrare che esiste un autovalore λ della matrice C tale che

$$|\lambda - \sigma_i| \leq \frac{\|(C - \sigma_i I)\mathbf{t}_i\|_2}{\|\mathbf{t}_i\|_2} = \|\mathbf{u}_{i+1} - \sigma_i \mathbf{t}_i\|_2 < \varepsilon.$$

Per trovare l'autovalore successivo e il corrispondente autovettore basta applicare il metodo delle potenze alla matrice

$$C_2 = C - \lambda_1 \mathbf{a}_1 \mathbf{a}_1^T.$$

Appendice A

Speranza condizionale

Daremo per scontate le nozioni tradizionalmente oggetto dei corsi di probabilità elementare (si veda [11]).

Consideriamo fissati uno spazio probabilizzato (Ω, \mathcal{F}, P) e una sottotribù \mathcal{A} di \mathcal{F} . Si denoti con Q la restrizione di P ad \mathcal{A} .

A.1 Definizione ed esistenza

Definizione A.1.1 *Data una variabile aleatoria reale integrabile X , si chiama versione della speranza condizionale di X rispetto alla sottotribù \mathcal{A} una variabile aleatoria Y , misurabile rispetto ad \mathcal{A} e integrabile secondo Q , che verifica la relazione*

$$\int_A X dP = \int_A Y dP$$

per ogni elemento A di \mathcal{A} .

Teorema A.1.2 *Data una variabile aleatoria reale integrabile X esiste sempre un'unica (a meno di equivalenza modulo Q) versione della speranza condizionale di X rispetto alla sottotribù \mathcal{A} . Inoltre una tale versione si può prendere positiva se X è positiva.*

Dimostrazione.

Ci si può limitare a dimostrare la seconda affermazione. Infatti la prima si riconduce alla seconda considerando la decomposizione $X = X^+ - X^-$. Si supponga dunque che X sia positiva. Consideriamo allora la misura ν , definita come

$$\nu(A) = \int_A X dP \quad \forall A \in \mathcal{F}.$$

Si riconosce essere (in modo banale) assolutamente continua rispetto a P . Perciò la sua restrizione μ alla sottotribù \mathcal{A} è assolutamente continua rispetto

all'analogia restrizione di P , ovvero rispetto a Q . Grazie al teorema di Radon-Nikodým μ ammette una densità Y rispetto a Q , cioè

$$\mu(A) = \int_A X dP = \int_A Y dQ \quad \forall A \in \mathcal{A}$$

dove Y è una variabile aleatoria positiva appartenente a $\mathcal{L}^1(Q)$. Rimane da osservare che, se Y è una variabile aleatoria misurabile rispetto ad \mathcal{A} , allora è integrabile secondo Q se e soltanto se lo è secondo P . Inoltre in questo caso si ha che

$$\int Y dP = \int Y dQ.$$

Questo basta per concludere che Y è una versione della speranza condizionale di X rispetto ad \mathcal{A} .

Per quanto riguarda l'unicità, basta osservare che, se Y_1 e Y_2 sono due versioni della speranza condizionale di X rispetto ad \mathcal{A} , si ha che

$$\int_A X dP = \int_A Y_1 dP = \int_A Y_2 dP \quad \forall A \in \mathcal{A},$$

ovvero che le due misure

$$\nu_1(A) = \int_A Y_1 dP \quad \text{e} \quad \nu_2(A) = \int_A Y_2 dP$$

coincidono sulla tribù \mathcal{A} . Le due densità Y_1 e Y_2 devono essere necessariamente equivalenti modulo Q .

Definizione A.1.3 *Data una variabile aleatoria reale integrabile X , si chiama speranza condizionale di X rispetto alla sottotribù \mathcal{A} la classe di equivalenza (in $L^1(Q)$) di una versione della speranza condizionale di X rispetto alla sottotribù \mathcal{A} . La si denota con il simbolo*

$$E^P[X|\mathcal{A}],$$

oppure semplicemente con

$$E[X|\mathcal{A}]$$

se è chiaro il riferimento alla misura di probabilità.

A.2 Proprietà elementari

Proposizione A.2.1 *Siano X e Y due variabili aleatorie reali integrabili.*

1. $E[aX + bY|\mathcal{A}] = aE[X|\mathcal{A}] + bE[Y|\mathcal{A}]$ per ogni a e b numeri reali.
2. $E[E[X|\mathcal{A}]] = E[X]$.
3. Se X è misurabile rispetto alla tribù \mathcal{A} allora $E[X|\mathcal{A}] = X$.

4. Se X è indipendente dalla tribù \mathcal{A} allora $E[X|\mathcal{A}] = E[X]$.
5. Se \mathcal{B} è una sottotribù di \mathcal{A} allora $E[E[X|\mathcal{A}]|\mathcal{B}] = E[X|\mathcal{B}]$.
6. Se Y è misurabile rispetto alla tribù \mathcal{A} e limitata allora $E[XY|\mathcal{A}] = YE[X|\mathcal{A}]$.

Appendice B

L'integrale stocastico e la formula di Itô

B.1 Nozioni fondamentali sui processi stocastici

Fissiamo uno spazio probabilizzato (Ω, \mathcal{F}, P) , uno spazio misurabile (E, \mathcal{E}) e un insieme ordinato (\mathbb{I}, \leq) . Un processo $(X_t)_{t \in \mathbb{I}}$ avente \mathbb{I} come insieme degli indici è una famiglia di variabili aleatorie su (Ω, \mathcal{F}, P) a valori in (E, \mathcal{E}) . Un altro modo di pensare un processo stocastico è come un'applicazione X dall'insieme $\mathbb{I} \times \Omega$ in E tale che l'applicazione $X_t = X(t, \cdot)$ sia una variabile aleatoria. Fissato invece $\omega \in \Omega$, $t \mapsto X_t(\omega)$ è un'applicazione dall'insieme \mathbb{I} in E detta *traiettoria del processo* X_t associata ad ω .

Generalmente si pensa ad un processo come all'evolversi nel tempo di un fenomeno: è quindi naturale riferirsi a (E, \mathcal{E}) come allo *spazio degli stati* e a \mathbb{I} come all'*insieme dei tempi* (che nei casi più comuni è \mathbb{N} o \mathbb{R}^+). X potrebbe rappresentare, nell'ambito della finanza, il prezzo di un qualsiasi attivo (derivato, azione, ...). Ad un certo istante (futuro) t , $X_t(\omega)$ è una quantità aleatoria. Ogni ω rappresenta uno stato dell'economia in corrispondenza del quale abbiamo il prezzo $X_t(\omega)$. Alternativamente, fissata una realizzazione dell'economia, abbiamo l'evolversi del prezzo $X(\cdot, \omega)$ nel tempo.

Nel caso in cui un processo sia a valori nella retta reale \mathbb{R} munita della tribù boreliana $\mathcal{B}(\mathbb{R})$, lo chiameremo *processo reale*.

Dato un processo X su (Ω, \mathcal{F}, P) , consideriamo per ogni t la tribù \mathcal{G}_t generata dal blocco delle variabili aleatorie X_s con $s \leq t$. La famiglia $\underline{\mathcal{G}} = (\mathcal{G}_t)_{t \in \mathbb{I}}$ è detta *filtrazione naturale di X* . Risulta che $\underline{\mathcal{G}}$ sia una famiglia crescente di sottotribù, ovvero fissati s, t con $s \leq t$ si ha che $\mathcal{G}_s \subseteq \mathcal{G}_t \subseteq \mathcal{F}$. Più in generale si può definire una filtrazione dati solo uno spazio misurabile (Ω, \mathcal{F}) qualsiasi e un insieme ordinato (\mathbb{I}, \leq) :

Definizione B.1.1 Una *filtrazione sullo spazio misurabile* (Ω, \mathcal{F}) avente \mathbb{I} come insieme dei tempi è una famiglia $\underline{\mathcal{F}} = (\mathcal{F}_t)_{t \in \mathbb{I}}$ di sottotribù di \mathcal{F} avente \mathbb{I} come insieme degli indici e tale che si abbia $\mathcal{F}_s \subseteq \mathcal{F}_t$ per ciascuna coppia

s, t di elementi di \mathbb{I} con $s \leq t$. Uno spazio probabilizzato (Ω, \mathcal{F}, P) sul quale è stata definita una filtrazione $\underline{\mathcal{F}}$ si chiama spazio probabilizzato filtrato e lo si denota con $(\Omega, \mathcal{F}, P, \underline{\mathcal{F}})$.

Definizione B.1.2 Fissata una filtrazione $\underline{\mathcal{F}} = (\mathcal{F})_{t \in \mathbb{I}}$ su (Ω, \mathcal{F}, P) , un processo X avente come insieme dei tempi una parte \mathbb{J} di \mathbb{I} si dice adattato a $\underline{\mathcal{F}}$ se la variabile aleatoria X_t è \mathcal{F}_t -misurabile per ogni elemento t di \mathbb{J} .

Fissata una filtrazione $\underline{\mathcal{F}}$ è ovvio che, affinché un processo X sia adattato a $\underline{\mathcal{F}}$, è necessario e sufficiente che per ogni indice t si abbia $\mathcal{G}_t \subseteq \mathcal{F}_t$, dove con $(\mathcal{G}_t)_{t \in \mathbb{I}}$ viene denotata la filtrazione naturale di X .

Definizione B.1.3 Si dice che la filtrazione $\underline{\mathcal{F}} = (\mathcal{F})_{t \in \mathbb{I}}$ su (Ω, \mathcal{F}, P) verifica le condizioni abituali se possiede le seguenti proprietà:

1. $\underline{\mathcal{F}}$ è continua a destra, ovvero per ogni $t \in \mathbb{I}$ vale $\mathcal{F}_t = \bigcap_{u > t} \mathcal{F}_u$.
2. $\underline{\mathcal{F}}$ è completa, ovvero ogni elemento di \mathcal{F} trascurabile secondo P appartiene a \mathcal{F}_t per ogni $t \in \mathbb{I}$.

Supporremo sempre che le filtrazioni prese in considerazione verifichino le condizioni abituali.

Definizione B.1.4 Dati due processi X e Y definiti sullo stesso spazio diremo che Y è una modificazione di X se per ogni fissato t le due variabili aleatorie X_t e Y_t coincidono quasi certamente.

È utile considerare accanto a questa nozione, anche un concetto di equivalenza molto più restrittivo.

Definizione B.1.5 Due processi X e Y definiti sullo stesso spazio sono indistinguibili se per quasi ogni ω le due traiettorie $X(\cdot, \omega)$ e $Y(\cdot, \omega)$ coincidono.

La differenza fra le definizioni B.1.4 e B.1.5 è che, nella prima, l'insieme trascurabile sul quale le due variabili X_t e Y_t non coincidono dipende dall'istante t , mentre nella definizione B.1.5 vi è un unico insieme trascurabile al di fuori del quale le due traiettorie coincidono.

Una famiglia di processi particolarmente interessanti è quella costituita dai processi reali continui a destra e con limite sinistro finito: essi si indicano con la sigla *rcll* (dall'inglese right continuous left limited) oppure con la sigla *càdlàg* (dal francese continu à droit limitée à gauche). In modo analogo si possono considerare i processi continui a sinistra indicati con la sigla *lc* oppure *càg*.

Definizione B.1.6 La tribù \mathcal{O} generata da tutti i processi adattati *rcll* si chiama tribù opzionale. Un processo si dice opzionale se è \mathcal{O} -misurabile.

Definizione B.1.7 La tribù \mathcal{P} generata da tutti i processi adattati *lc* si chiama tribù prevedibile. Un processo si dice prevedibile se è \mathcal{P} -misurabile.

B.2 Martingale

Introduciamo ora dei processi con particolari proprietà, le martingale appunto.

Definizione B.2.1 *Dato uno spazio probabilizzato (Ω, \mathcal{F}, P) munito di una filtrazione $\underline{\mathcal{F}} = (\mathcal{F})_{t \in \mathbb{I}}$, un processo reale X avente come insieme dei tempi una parte \mathbb{J} di \mathbb{I} è una martingala se valgono le condizioni seguenti:*

1. X è adattato a $\underline{\mathcal{F}}$.
2. La variabile X_t è integrabile per ogni elemento t di \mathbb{J} .
3. Per ogni coppia s, t di elementi di \mathbb{J} con $s \leq t$ la variabile aleatoria X_s è una versione della speranza condizionale di X_t rispetto a \mathcal{F}_s , ovvero

$$X_s \in E[X_t | \mathcal{F}_s]. \quad (\text{B.2.1})$$

Se invece si denota impropriamente con X_s la classe di equivalenza in $L^1(P)$ avente X_s come rappresentante la condizione B.2.1 può essere riscritta come

$$X_s = E[X_t | \mathcal{F}_s].$$

Dalle proprietà della speranza condizionale elencate nella proposizione A.2, si ha che la condizione B.2.1 è equivalente a

$$E[X_t - X_s | \mathcal{F}_s] = 0.$$

Una martingala può rappresentare l'evoluzione, al variare del tempo, del capitale di un giocatore durante il corso di un gioco equo. Il valore atteso ad un certo istante futuro t del processo, conosciuta la storia del processo stesso anteriore all'istante t , è il suo valore attuale.

Date due martingale X e Y (rispetto alla stessa filtrazione $\underline{\mathcal{F}}$), si riconosce facilmente che anche $aX + bY$ è una martingala per ogni coppia a, b di numeri reali. Dunque l'insieme costituito da tutte le martingale su (Ω, \mathcal{F}, P) rispetto alla filtrazione $\underline{\mathcal{F}}$ è uno spazio vettoriale.

B.3 Il Processo di Wiener

Supponiamo fissato uno spazio probabilizzato (Ω, \mathcal{F}, P) e su di esso una filtrazione $\underline{\mathcal{F}}$ avente \mathbb{R}_+ come insieme dei tempi.

Definizione B.3.1 *Un processo X su (Ω, \mathcal{F}, P) avente \mathbb{R}_+ come insieme dei tempi è detto additivo (rispetto alla filtrazione $\underline{\mathcal{F}}$) se è adattato a $\underline{\mathcal{F}}$ e se per ciascun numero reale positivo s il blocco $[X_{t+s} - X_s]_{t \in \mathbb{R}_+}$ è indipendente dalla tribù \mathcal{F}_s .*

Definizione B.3.2 Un processo reale W su (Ω, \mathcal{F}, P) avente \mathbb{R}_+ come insieme dei tempi è detto un processo di Wiener (o moto browniano) relativo alla filtrazione $\underline{\mathcal{F}}$ se

1. $W_0 = 0$ $P - q.c.$
2. W è additivo (rispetto a $\underline{\mathcal{F}}$).
3. W ha traiettorie continue.
4. Per ogni coppia s, t di numeri reali positivi $(W_{t+s} - W_s)$ ha legge $\mathcal{N}(0, t)$.

Osserviamo innanzitutto che l'ipotesi che W sia nullo in 0 è inessenziale: in effetti le proprietà alle quali deve sottostare un processo di Wiener riguardano soltanto gli incrementi e dunque non dipende dalla variabile aleatoria iniziale. Inoltre se W è un processo di Wiener allora per ogni coppia s, t di numeri reali positivi la variabile aleatoria $(W_{t+s} - W_s)$ è indipendente da W_s . Più in generale, per ogni partizione finita $0 \leq t_1 < t_2 < \dots < t_n < +\infty$,

$$(W_{t_1}, W_{t_2} - W_{t_1}, \dots, W_{t_n} - W_{t_{n-1}})$$

è una famiglia di variabili aleatorie indipendenti e la legge della generica variabile $(W_{t_j} - W_{t_{j-1}})$, ovvero la legge $\mathcal{N}(0, t_j - t_{j-1})$, dipende da t_j e t_{j-1} solo tramite la loro differenza.

Il processo di Wiener ha particolare importanza nell'ambito della Finanza matematica in quanto si presta bene a modellizzare l'andamento di un prezzo le cui variazioni non dipendano dal passato: infatti, se $\underline{\mathcal{F}}$ è la tribù generata da un processo di Wiener W , allora per ogni coppia s, t di numeri reali positivi $W_{t+s} - W_s$ è indipendente da \mathcal{F}_s , cioè dal passato.

Proposizione B.3.3 Sia W un processo di Wiener su (Ω, \mathcal{F}, P) rispetto alla filtrazione $\underline{\mathcal{F}}$. Allora:

1. $(W_t)_{t \geq 0}$ è una martingala.
2. $(W_t^2 - t)_{t \geq 0}$ è una martingala.
3. $(\exp(\sigma W_t - \sigma^2/2t))_{t \geq 0}$ è una martingala.

Dimostrazione.

La prima affermazione è conseguenza delle proprietà elementari della speranza condizionale; si ha infatti, per ogni coppia s, t di numeri reali positivi

$$E[W_{t+s} - W_s | \mathcal{F}_s] = E[W_{t+s} - W_s] = 0.$$

Per quanto riguarda la seconda affermazione, innanzitutto osserviamo che per ogni coppia s, t di numeri reali positivi si ha:

$$E[(W_{t+s} - W_s)^2 | \mathcal{F}_s] = E[W_{t+s}^2 + W_s^2 | \mathcal{F}_s] - 2W_s E[W_{t+s} | \mathcal{F}_s] = E[W_{t+s}^2 - W_s^2 | \mathcal{F}_s].$$

Ne segue, dunque, sfruttando il fatto che $(W_{t+s} - W_s)$ è indipendente dalla tribù \mathcal{F}_s e che ha legge $\mathcal{N}(0, t)$,

$$\begin{aligned} E[(W_{t+s}^2 - (t+s)) - (W_s^2 - s) | \mathcal{F}_s] &= E[(W_{t+s} - W_s)^2 | \mathcal{F}_s] - t \\ &= E[(W_{t+s} - W_s)^2] - t = 0. \end{aligned}$$

Anche la seconda affermazione è così provata. Per quanto riguarda l'ultima, innanzitutto ricordiamo che, se Y ha legge normale ridotta, allora, per ogni numero reale λ

$$E[e^{\lambda Y}] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{\lambda x} e^{-x^2/2} dx = e^{\lambda^2/2}.$$

Per ogni coppia s, t di numeri reali positivi si ha

$$\begin{aligned} E[e^{\sigma W_{t+s} - \sigma^2(t+s)/2} | \mathcal{F}_s] &= e^{\sigma W_s - \sigma^2(t+s)/2} E[e^{\sigma(W_{t+s} - W_s)} | \mathcal{F}_s] \\ &= e^{\sigma W_s - \sigma^2(t+s)/2} E[e^{\sigma(W_{t+s} - W_s)}] \\ &= e^{\sigma W_s - \sigma^2(t+s)/2} e^{\sigma^2 t/2} \\ &= e^{\sigma W_s - \sigma^2 s/2}. \end{aligned}$$

Questo conclude la dimostrazione. \square

Di facile dimostrazione è la seguente proposizione:

Proposizione B.3.4 *Supponiamo fissato uno spazio probabilizzato (Ω, \mathcal{F}, P) e su di esso un processo reale W nullo in 0 e con traiettorie continue. Le condizioni seguenti sono allora equivalenti:*

1. W è un processo di Wiener rispetto alla filtrazione naturale.
2. W è una famiglia gaussiana e centrata e, per ogni coppia s, t di numeri reali positivi si ha:

$$E[W_t W_s] = t \wedge s.$$

Riassumiamo le “proprietà patologiche” del processo di Wiener nella proposizione seguente. Per una trattazione più approfondita e per una dimostrazione si consulti [9, paragrafo 2.9].

Proposizione B.3.5 *Sia W un processo di Wiener su (Ω, \mathcal{F}, P) . Allora:*

1. *L'insieme costituito dagli elementi $\omega \in \Omega$ per i quali la traiettoria $W(\cdot, \omega)$ ammette in qualche punto di \mathbb{R}_+ derivata destra inferiore e derivata destra superiore entrambe finite è contenuto in un evento trascurabile. Un enunciato analogo vale per le derivate sinistre.*
2. *Per quasi ogni ω , la traiettoria $W(\cdot, \omega)$ non è monotona su alcun intervallo proprio (e quindi possiede un insieme denso di punti di massimo locale).*
3. *Per quasi ogni ω , la traiettoria $W(\cdot, \omega)$ assume massimi diversi su ogni coppia d'intervalli propri, compatti e disgiunti, con estremi razionali (e quindi non possiede alcun punto di massimo locale che sia stretto).*

B.4 Il processo di Wiener multidimensionale

Supponiamo fissato uno spazio probabilizzato (Ω, \mathcal{F}, P) e, su di esso, una filtrazione $\underline{\mathcal{F}} = (\mathcal{F}_t)_{t \geq 0}$.

Definizione B.4.1 *Siano W^1, \dots, W^d processi di Wiener relativi alla filtrazione $\underline{\mathcal{F}}$. Si supponga che i blocchi $[W_t^1]_{t \geq 0}, \dots, [W_t^d]_{t \geq 0}$ siano variabili aleatorie indipendenti. Si dice allora che il processo W (sullo spazio (Ω, \mathcal{F}, P) , a valori in $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$), definito da*

$$W_t = [W_t^1, \dots, W_t^d],$$

è un processo di Wiener d -dimensionale (relativo alla filtrazione $\underline{\mathcal{F}}$).

Grazie all'indipendenza delle componenti del processo W , è immediato riconoscere che il processo W verifica le proprietà seguenti:

1. Il processo W è adattato a $\underline{\mathcal{F}}$ e per ogni coppia s, t di numeri reali positivi la variabile aleatoria $(W_{t+s} - W_s)$ è indipendente dalla tribù \mathcal{F}_s .
2. Per ogni coppia s, t di numeri reali positivi la legge su $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ della variabile aleatoria $(W_{t+s} - W_s)$ coincide con la misura prodotto $\mathcal{N}(0, t)^{\otimes d}$. In particolare questa legge è dotata di densità (rispetto alla misura di Borel-Lebesgue) n_t , definita da

$$n_t(x) = (2\pi t)^{-n/2} \exp(-|x|^2/(2t)).$$

3. Il processo W ha traiettorie continue.

È facile riconoscere che queste tre proprietà caratterizzano completamente un processo di Wiener di dimensione d , nel senso che, se W è un processo a valori nello spazio misurabile $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ e verifica le tre proprietà appena elencate, allora esso è un processo di Wiener d -dimensionale.

B.5 L'integrale stocastico di Itô

L'obiettivo di questo paragrafo è quello di definire, dato un processo H e un processo di Wiener W , un processo del tipo

$$I(H)_t = \int_0^t H_s dW_s.$$

Non è possibile definirlo come

$$I(H)_t = \int_0^t H_s \frac{dW_s}{ds} ds,$$

in quanto il processo di Wiener ha traiettorie non derivabili per quasi ogni $\omega \in \Omega$. Un altro approccio, senz'altro più efficace, è quello di definire l'integrale tramite la misura di Stieltjes corrispondente al processo integratore; è necessario, però, che le sue traiettorie siano a variazione limitata, proprietà non posseduta dal processo di Wiener.

Iniziamo col costruire l'integrale stocastico per un particolare insieme di processi H , i *processi semplici*, per estenderlo poi, in vari passi, ad una classe più ampia possibile di processi. Riportiamo qui solo i risultati principali. Per una trattazione più approfondita rimandiamo a [9, capitolo 3].

Supponiamo fissato uno spazio probabilizzato (Ω, \mathcal{F}, P) e su di esso una filtrazione $\underline{\mathcal{F}} = (\mathcal{F}_t)_{t \geq 0}$ e un processo di Wiener $W = (W_t)_{t \geq 0}$ relativo alla filtrazione $\underline{\mathcal{F}}$. T è un numero reale finito strettamente positivo e indicheremo con \mathbb{I} l'intervallo $[0, T]$.

Definizione B.5.1 *Un processo $H = (H_t)_{t \in \mathbb{I}}$ è un processo semplice se può essere scritto nella forma*

$$H_t(\omega) = \sum_{i=1}^n \phi_i(\omega) I_{[t_{i-1}, t_i]}(t),$$

dove con $I_{[t_{i-1}, t_i]}$ indichiamo l'indicatrice dell'intervallo $]t_{i-1}, t_i]$, ϕ_i è un processo limitato e $\mathcal{F}_{t_{i-1}}$ -misurabile e $0 = t_0 < t_1 < \dots < t_n = T$.

Definizione B.5.2 *Sia H un processo semplice. L'integrale stocastico di H_t in dW_t è il processo continuo $(I(H)_t)_{t \in \mathbb{I}}$ definito per ogni $t \in]t_k, t_{k+1}]$ come*

$$I(H)_t = \sum_{i=1}^k \phi_i(W_{t_i} - W_{t_{i-1}}) + \phi_{k+1}(W_t - W_{t_k}).$$

Osserviamo che $I(H)_t$ può essere scritto nella forma

$$I(H)_t = \sum_{i=1}^n \phi_i(W_{t_i \wedge t} - W_{t_{i-1} \wedge t})$$

che prova la continuità di $t \mapsto I(H)_t$. Scriveremo $\int_0^t H_s dW_s$ invece di $I(H)_t$.

Proposizione B.5.3 *Se H è un processo semplice, allora:*

1. $(\int_0^t H_s dW_s)$ è una martingala (rispetto a $\underline{\mathcal{F}}$) continua.
2. $E[(\int_0^t H_s dW_s)^2] = E[\int_0^t H_s^2 ds]$.
3. $E \left[\sup_{t \in \mathbb{I}} \left| \int_0^t H_s dW_s \right|^2 \right] \leq 4E[\int_0^T H_s^2 ds]$.

Per una dimostrazione consultare [10, pagina 36].

Ora che abbiamo definito l'integrale stocastico per processi semplici ed enunciato alcune sue proprietà, cerchiamo di estendere la definizione ad una classe \mathcal{H} di processi adattati più ampia, definita come

$$\mathcal{H} = \left\{ (H_t)_{t \in \mathbb{I}} \text{ adattati con } E\left[\int_0^T H_s^2 ds\right] < +\infty \right\}.$$

Proposizione B.5.4 *Esiste un'unica mappa lineare J da \mathcal{H} nello spazio delle martingale rispetto a $\underline{\mathcal{F}}$ continue definite su \mathbb{I} tale che:*

1. *Se $(H_t)_{t \in \mathbb{I}}$ è un processo semplice $J(H)_t = I(H)_t$ $P - q.c.$ per ogni $t \in \mathbb{I}$.*
2. *Se $t \in \mathbb{I}$, $E[J(H)_t^2] = E[\int_0^t H_s^2 ds]$.*

La mappa J è unica nel senso che, se J e J' soddisfano le proprietà precedenti, allora per ogni $t \in \mathbb{I}$

$$J(H)_t = J'(H)_t \quad P - q.c.$$

Per $H \in \mathcal{H}$ denotiamo $J(H)_t$ con $\int_0^t H_s dW_s$.

La dimostrazione di questa proposizione, che può essere trovata in [10, pagina 38], si basa sul fatto che se $H \in \mathcal{H}$ allora esiste una successione $(H^n)_{n \geq 0}$ di processi semplici tale che

$$\lim_{n \rightarrow +\infty} E\left[\int_0^T |H_s - H_s^n|^2 ds\right] = 0.$$

Una dimostrazione di questo risultato si può trovare in [9, pagina 134].

Facciamo un ulteriore passo nell'estensione della definizione di integrale stocastico a processi che soddisfano una condizione di integrabilità più debole rispetto ai processi appartenenti ad \mathcal{H} . Introduciamo la classe $\tilde{\mathcal{H}}$ così definita

$$\tilde{\mathcal{H}} = \left\{ (H_t)_{t \in \mathbb{I}} \text{ adattati con } \int_0^T H_s^2 ds < +\infty \quad P - q.c. \right\}.$$

Proposizione B.5.5 *Esiste un'unica mappa lineare \tilde{J} da $\tilde{\mathcal{H}}$ nello spazio vettoriale dei processi continui definiti su \mathbb{I} tale che:*

1. *Se $(H_t)_{t \in \mathbb{I}}$ è un processo semplice $\tilde{J}(H)_t = I(H)_t$ $P - q.c.$ per ogni $t \in \mathbb{I}$.*
2. *Se $(H^n)_{n \geq 0}$ è una successione di elementi di $\tilde{\mathcal{H}}$ tali che $\int_0^T (H_s^n)^2 ds$ converge a 0 in probabilità, allora $\sup_{t \in \mathbb{I}} |\tilde{J}(H^n)_t|$ converge a 0 in probabilità.*

Scriviamo $\int_0^t H_s dW_s$ per $\tilde{J}(H)_t$.

Sottolineiamo il fatto che, in questo caso, $(\int_0^t H_s dW_s)_{t \in \mathbb{I}}$ non è più necessariamente una martingala.

Fino ad ora si è definito l'integrale stocastico rispetto ad un processo di Wiener. È possibile estendere il concetto di integrale stocastico fino al caso più generale in cui il processo integratore X sia una semimartingala continua, ovvero sia un processo adattato che ammette una decomposizione del tipo $X = M + A$ dove M è una martingala locale continua e A è un processo continuo e adattato con $A_0 = 0$ q.c. e localmente a variazione finita. Tale decomposizione è unica a meno di modificazioni. Per una trattazione completa consultare [9, capitolo 3].

B.6 Calcolo di Itô

Introduciamo ora alcuni risultati di calcolo differenziale basato sull'integrale stocastico rispetto ad un processo di Wiener, con lo scopo di riuscire a differenziare una funzione del tipo $t \mapsto f(W_t)$.

L'esempio seguente vuole solo dimostrare che una semplice estensione del calcolo differenziale classico fallisce. Proviamo a differenziare la funzione $t \mapsto W_t^2$ in termini di " dW_t ". Per le funzioni $f(t)$ differenziabili e nulle in zero vale

$$f(t)^2 = 2 \int_0^t f(s) f'(s) ds = 2 \int_0^t f(s) df(s).$$

Nel caso "stocastico" è impossibile avere una formula analoga, cioè

$$W_t^2 = 2 \int_0^t W_s dW_s.$$

Infatti $(\int_0^t W_s dW_s)_t$ è una martingala nulla in zero, dato che $E[\int_0^T W_s^2 ds] < +\infty$. Inoltre è non negativa in quanto $\int_0^t W_s dW_s = W_t^2/2$ e, una martingala non negativa nulla in zero, non può essere altro che identicamente nulla.

Iniziamo con l'introdurre la classe dei processi per i quali è possibile applicare la formula di Itô.

Definizione B.6.1 Dato uno spazio probabilizzato (Ω, \mathcal{F}, P) e su di esso una filtrazione $\mathcal{F} = (\mathcal{F}_t)_{t \geq 0}$ e un processo di Wiener $(W_t)_{t \geq 0}$ relativo ad \mathcal{F} , si dice che un processo reale $(X_t)_{t \in \mathbb{I}}$ è un processo di Itô se per ogni $t \in \mathbb{I}$ può essere scritto nella forma

$$X_t = X_0 + \int_0^t K_s ds + \int_0^t H_s dW_s \quad P - q.c.,$$

dove

1. X_0 è una variabile aleatoria \mathcal{F}_0 -misurabile.
2. $(K_t)_{t \in \mathbb{I}}$ e $(H_t)_{t \in \mathbb{I}}$ sono processi adattati alla filtrazione $\underline{\mathcal{F}}$.
3. $\int_0^T |K_s| ds < +\infty$ $P - q.c.$
4. $\int_0^T |H_s|^2 ds < +\infty$ $P - q.c.$

Proposizione B.6.2 *La decomposizione dei processi di Itô è unica a meno di modificazioni.*

Teorema B.6.3 (formula di Itô) *Sia $(X_t)_{t \in \mathbb{I}}$ un processo di Itô della forma*

$$X_t = X_0 + \int_0^t K_s ds + \int_0^t H_s dW_s.$$

Se $f \in C^2(\mathbb{R})$, allora

$$f(X_t) = f(X_0) + \int_0^t f'(X_s) dX_s + \frac{1}{2} \int_0^t f''(X_s) d[X, X]_s \quad P - q.c.,$$

dove per definizione

$$[X, X]_s = \int_0^s H_s^2 ds$$

e

$$\int_0^t f'(X_s) dX_s = \int_0^t f'(X_s) K_s ds + \int_0^t f'(X_s) H_s dW_s.$$

Analogamente se $(t, x) \mapsto f(t, x) \in C^{1,2}([0, T] \times \mathbb{R})$, allora

$$\begin{aligned} f(t, X_t) &= f(0, X_0) + \int_0^t f'_s(s, X_s) ds \\ &+ \int_0^t f'_x(s, X_s) dX_s + \frac{1}{2} \int_0^t f''_{xx}(s, X_s) d[X, X]_s \quad P - q.c. \end{aligned}$$

B.7 Teorema di Girsanov

Ci limitiamo al solo enunciato del teorema.

Teorema B.7.1 (di Girsanov) *Su uno spazio probabilizzato (Ω, \mathcal{F}, P) sia $(\tilde{W}_t)_{t \in \mathbb{I}}$ un processo di Wiener d -dimensionale e $\underline{\mathcal{F}} = (\mathcal{F}_t)_{t \in \mathbb{I}}$ la sua filtrazione naturale. Se Q è una misura di probabilità equivalente a P , allora esiste un processo adattato $(g_t)_{t \in \mathbb{I}}$ che verifica $\int_0^T \|g_s\|^2 ds < +\infty$ P -q.c. e tale che la densità $\frac{dQ}{dP} |_{\mathcal{F}_t}$ sia data da*

$$\frac{dQ}{dP} |_{\mathcal{F}_t} = \mathcal{E}(g \star \tilde{W})_t := \exp \left(\int_0^t g_s \cdot d\tilde{W}_s - \frac{1}{2} \int_0^t \|g_s\|^2 ds \right). \quad (\text{B.7.1})$$

inoltre, il processo

$$W_t := \tilde{W}_t - \int_0^t g_s ds$$

è un processo di Wiener sotto la misura di probabilità Q .

Viceversa, se $(g_t)_{t \in \mathbb{I}}$ è un processo adattato che verifica la condizione $\int_0^T \|g_s\|^2 ds < +\infty$ P -q.c. e tale che $(\mathcal{E}(g \star \tilde{W})_t)$ sia una martingala, allora la B.7.1 definisce una misura di probabilità Q equivalente a P .

Una condizione sufficiente affinché $(\mathcal{E}(g \star \tilde{W})_t)$ sia una martingala è che valga

$$E \left[\exp \left(\frac{1}{2} \int_0^T \|g_s\|^2 ds \right) \right] < +\infty \quad (\text{condizione di Novikov}).$$

Per una dimostrazione sia del teorema di Girsanov che della condizione di Novikov, rimandiamo a [9, paragrafo 3.5].

Bibliografia

- [1] D. BINI, M. CAPOVANI, O. MENCHI. *Metodi Numerici per l'Algebra Lineare*. Zanichelli.
- [2] T. BJÖRK. *Arbitrage Theory in Continuous Time*. Oxford (1998).
- [3] T. BJÖRK. Interest Rate Theory. *Financial mathematics*. Lecture Notes in Mathematics, Vol. 1656, 53–122 (1996) .
- [4] D. BRIGO, F. MERCURIO. *Interest Rate Models-Theory and Practice*. Springer (2001).
- [5] R. CARMONA, M. TEHRANCHI. *Interest Rate Models: an Infinite Dimensional Stochastic Analysis Perspective*. Springer (2006).
- [6] D. FILIPOVIĆ. *Fixed Income Models*. Appunti del corso, Princeton University (Autunno 2002).
- [7] J. C. HULL. *Options, Futures and Other Derivatives*. Sesta edizione, Prentice Hall (2006).
- [8] J. JAMES, N. WEBBER. *Interest Rate Modelling*. John Wiley & Sons (2000).
- [9] I. KARATZAS, S. E. SHREVE. *Brownian Motion and Stochastic Calculus*. Seconda edizione, Springer (1991).
- [10] D. LAMBERTON, B. LAPEYRE. *Introduction to Stochastic Calculus Applied to Finance*. Chapman & Hall (1996).
- [11] G. LETTA. *Probabilità Elementare*. Zanichelli (1993).
- [12] R. REBONATO. *Interest-Rate Option Model*. John Wiley & Sons (1996).