UNIVERSITÀ DEGLI STUDI ROMA TRE
FACOLTÀ DI SCIENZE M.F.N.

Tesi di Laurea in Matematica
di
Maurizio Piedimonte

Sulla stima delle strutture dei tassi di interesse con funzioni Spline

Relatore
Prof. Lucia Caramellino

Correlatore
Prof. Alessandro Ramponi

Il Relatore

Il Candidato

ANNO ACCADEMICO 1999 - 2000
NOVEMBRE 2000

Classificazione AMS : 62H12,91B28
Parole Chiave : Tassi di interesse, Funzioni Spline, Regressione multivariata, Problemi di stima

Maurizio Piedimonte è nato a Roma il 29 Maggio 1974.
Ha conseguito il Diploma di Maturità scientifica presso il Liceo Scientifico statale "Keplero" di Roma nel luglio 1993.
Si è immatricolato al Corso di Laurea in Matematica presso l'Università degli Studi di Roma Tre nell'anno accademico 1993 - 1994.
Ha presentato per la prova di qualificazione all'esame di laurea le seguenti tesine orali :
"La probabilitá condizionata"
"L'eliminazione di Gauss".

Indice

Introduzione

La conoscenza della struttura a termine dei tassi di interesse, ovvero la relazione che caratterizza gli interessi in funzione della maturità di un investimento, risulta essere di grande importanza per la gestione di strumenti finanziari basati sui tassi di interesse, genericamente indicati come *strumenti a reddito fisso*, come ad esempio obbligazioni e swaps.

In generale tali strumenti sono dei contratti in base ai quali un investitore presta del denaro in cambio della promessa di un guadagno futuro. Tale investitore conosce con esattezza i tempi e le quantità dei suoi guadagni, in confronto ad altre forme di investimento caratterizzate da una incertezza del ritorno. Per questa ragione le obbligazioni emesse da governi o grandi istituzioni, per le quali il rischio di perdere il proprio investimento è praticamente nullo, forniscono un indispensabile mezzo per misurare i tassi di interesse offerti dal mercato.

La più importante curva che descrive la struttura a termine dei tassi di interesse, è la cosidetta *funzione di sconto*, dalla quale altre curve, come la curva dei tassi spot o quella dei tassi forward, ben note ai praticanti dei mercati finanziari, possono essere ottenute. Tale funzione, come specificheremo nei capitoli successivi, descrive il valore al momento presente, o valore attualizzato, di una quantità unitaria di denaro che sarà riscosso in un istante di tempo futuro. Sulla base di questa funzione e delle corrispondenti curve spot e forward, varie misure del ritorno di un investimento possono essere definite, cosi' come strategie di copertura del rischio ([2]).

Sfortunatamente la funzione di sconto, i tassi spot ed i tassi forward, non sono direttamente osservabili nel mercato, che fornisce invece solamente i prezzi correnti delle obbligazioni scambiate dagli investitori. Risulta tuttavia che tali prezzi sono legati linearmente alla funzione di sconto che può quindi essere solo stimata usando le informazioni che il mercato fornisce. Da un punto di vista matematico, risulta che tale problema può essere formulato come un problema di regressione lineare. Osserviamo a questo punto che questo tipo di approccio è comunemente chiamato *cross-sectional* in quanto, fissati i prezzi giornalieri delle obbligazioni considerate, produce una stima giornaliera delle curve di interesse, contrariamente all'approccio basato su modelli che descrivono l'evoluzione nel tempo dei tassi considerati (ad esempio [7] o più generalmente [17]).

Poichè inoltre, come è stato osservato da McCulloch ([15]) in poi, la relazione che lega i prezzi delle obbligazioni alla funzione di sconto è perturbato in pratica da alcuni fattori, come ad esempio quelli legati alla tassazione dei guadagni o le differenze praticate tra il prezzo d'acquisto o di vendita o temporanee opportunità di arbitraggio, le tecniche statistiche sono diventate uno strumento essenziale per stimare la funzione di sconto.

In ([15]) McCulloch propone un metodo di stima basato sulle funzioni spline. L'uso di tali funzioni in problemi di regressione, costituisce una vera e propria area di ricerca nell'ambito delle scienze statistiche, i modelli spline di regressione ([10]). Molteplici sono infatti i problemi che possono essere affrontati utilizzando tecniche basate sulle funzioni spline (si veda ad esempio il recente [21]). È stato tuttavia sperimentalmente osservato ([19], [2]) come l'uso di tali funzioni per la stima della struttura a termine dei tassi di interesse, possa produrre dei risultati irrealistici da un punto di vista finanziario, fenomeno sostanzialmente dovuto ad una "cattiva" scelta delle funzioni spline.

Il principale obiettivo di questa tesi è quello di studiare sperimentalmente alcune di queste tecniche basate sulle funzioni spline e di utilizzare dei metodi classici di scelta del modello di regressione per cercare di ovviare ad alcuni dei problemi tipicamente osservati. Più precisamente, poichè il modello spline di regressione è determinato una volta che si siano scelti i nodi della funzione spline approssimante, il problema di scelta del modello di regressione diventa un problema di scelta di una opportuna configurazione di nodi. Abbiamo quindi implementato un algoritmo che produce un'insieme di configurazioni di nodi tra le quali viene scelta quella "ottima" utilizzando diversi criteri classici (R^2 modificato, GCV, AIC e BIC ([12])). Le tecniche proposte sono state quindi sperimentate su tre differenti insiemi di dati "sintetici", ovvero generati da un modello noto.

La struttra di questa tesi è la seguente. Nel primo capitolo descriveremo in dettaglio le obbligazioni con reddito fisso e la loro relazione con la *funzione di sconto*, i tassi di interesse *spot* e *forward*. Il mercato nel quale focalizzeremo la nostra attenzione è quello americano (*US Treasury*), poiché è un mercato molto sicuro, essendo quasi nullo il rischio di non ricevere i soldi investiti, molti titoli sono correntemente quotati ed è comunemente considerato un mercato efficiente . Preciseremo inoltre come la funzione di sconto, il tasso spot e il tasso forward siano legati tra loro e deriveremo la relazione che sussiste con i prezzi delle obbligazioni quotate nel mercato, le loro maturità e le cedole di pagamento.

Nel secondo capitolo deriveremo in dettaglio il problema statistico di regressione e illustreremo diversi metodi proposti in letteratura, principalmente basati sulle funzioni spline, per stimare tali curve. Nel terzo capitolo definiremo formalmente le funzioni spline e le loro principali proprietà e descriveremo i metodi computazionali utilizzati, il metodo dei minimi quadrati e dei minimi quadrati pesati, l'algoritmo di generazione delle configurazioni dei nodi ed i vari criteri di scelta del modello di regressione. Infine, nel quarto capitolo, verranno descritte tutte le prove numeriche effettuate su un insieme di dati test.

Capitolo 1

Il Mercato Fixed Income

1.1 Il mercato delle obbligazioni

Un'obbligazione (*bond*) è uno strumento di credito emesso da istituzioni pubbliche o da compagnie private, e consente a chi ha acquistato l'obbligazione, l'investitore, di ottenere in un certo periodo di tempo un flusso di denaro a seguito del suo investimento.

La principale fonte di insicurezza per un investitore che compra queste obbligazioni, è quella di non ricevere il pagamento dovuto nel periodo specificato (*risk of default*). Nel seguito tratteremo solo il caso di obbligazioni emesse dal Tesoro americano (*US Treasury*), in cui il rischio di default si assume essere nullo.

Quando di un'obbligazione si conosce in anticipo la quantita' di denaro che deve essere retribuita all'investitore e il periodo di tempo che intercorre tra ogni pagamento, l'obbligazione entra nella categoria delle obbligazioni con reddito fisso (*fixed-income bonds*). Tutte le obbligazioni emesse hanno una scadenza o maturita' (*maturity o redemption*) prefissata, che corrisponde al "tempo di vita" dell'obbligazione stessa. Le obbligazioni si dividono in due categorie:le obbligazioni che pagano cedole e quelle che non prevedono cedole, comunemente chiamate *Zero-coupons*.

Le obbligazioni con cedole sono obbligazioni che consentono all'investitore di riscuotere una parte prestabilita del suo capitale ad istanti di tempo prefissati. Questi pagamenti vengono fatti tramite delle cedole, dette *coupons*, la cui frequenza di pagamento varia da mercato a mercato, anche se di solito è semi-annuale o annuale. Quando l'obbligazione giunge alla sua maturita', l'investitore oltre a ricevere l'ultima cedola, riceve anche il valore nominale (*face value*) dell'obbligazione stessa. Ad esempio comprando un'obbligazione del Tesoro americano di valore nominale \$10000 dollari con cedole che restituiscono il 4.375% del valore nominale (indicata nella terminologia del mercato obbligazionario americano come 4 3/8 corrispondente a $4 + 3/8 = 4.375$) e maturità 15 Febbraio 2002, il flusso di cassa garantito all'investitore è di \$437.5 ogni anno fino alla data di maturità in cui viene restituito anche il valore nominale \$10000. Per convenzione del mercato delle obbligazioni americano la cedola di \$437.5 è pagata in due rate semestrali di \$218.75. Viene assunto inoltre come valore nominale di riferimento la quantità \$100.

Le obbligazioni *Zero-coupons* invece retribuiscono tutto il valore degli interessi in un unico pagamento coincidente con la propria maturita' che è al massimo un anno.

Un investitore puo' acquistare nuove obbligazioni direttamente dalle istituzioni che le emettono nel cosidetto primo mercato (*primary market*), oppure acquistare da altri investitori obbligazioni gia' emesse nel secondo mercato (*secondary market*).

Le obbligazioni sono caratterizzate dall'avere un prezzo di vendita e un prezzo di acquisto (*bid-ask price*), che è determinato dalle transazioni che avvengono tra i partecipanti al mercato. Infatti il prezzo di vendita di un'obbligazione aumenta notevolmente nel momento in cui ci sono molte richieste da parte degli investitori, mentre viceversa diminuisce quando pochi la comprano.

Alla luce di quanto abbiamo visto fino ad ora, un'obbligazione è completamente caratterizzata da una scadenza, dalle cedole di pagamento degli interessi, e dal suo prezzo di mercato. Questa caratteristica rende le obbligazioni un utile strumento per misurare i tassi di interesse del mercato ([2]). In questo contesto risultano di fondamentale importanza la funzione di sconto (*discount function*), il tasso di interesse *spot* e il tasso di interesse *forward*, che introdurremo nei paragrafi successivi.

1.2 La funzione di sconto

Il primo passo nello studio dei titoli obbligazionari è quella di quantificare il valore attualizzato di una quantita' di denaro che si ricevera' in futuro.

È preferibile avere \$100 oggi, oppure tra un anno? A prima vista non fa alcuna differenza ma per un investitore comprare delle obbligazioni, in un mercato dove la probabilita' di perdere il proprio investimento è molto bassa, e riscattarli dopo un anno, puo' comportare un determinato guadagno.

Il valore dei soldi non è certamente costante nel tempo, infatti le mille lire degli anni '60, hanno un valore ben diverso nei nostri giorni. Il cambiamento del valore del denaro può essere rappresentato matematicamente mediante una opportuna funzione d, detta *funzione di sconto*.

Tale funzione per un intervallo di tempo $[0, T]$, ci dice quale sara' il valore ad oggi (*attualizzato*) di \$1 alla fine del periodo stesso. È chiaro che $d(0) = 1$, e che è ragionevole supporre che d sia monotona non crescente. In generale, il valore attualizzato di una certa quantita' di denaro che varrà X al tempo T è

$$X \cdot d(T) \tag{1.1}$$

In altre parole $Xd(T)$ rappresenta la quantità di denaro che occorre depositarte in banca al tempo 0 per aver X al tempo T. I valori della funzione d in un insieme finito di istanti di tempo T_1,T_2, ... ,T_N, $d(T_1), d(T_2), \ldots, d(T_N)$, saranno chiamati fattori di sconto (*discount factors*)

Supponiamo per esempio che il fattore di sconto per sei mesi sia $d(.5) = .9825$, allora il valore attualizzato di \$1, tra sei mesi sara' \$.9825. Cosi' il valore attualizzato di \$105 tra sei mesi corrispondera' a $\$105 \cdot \$.9825 = \$103.16$.

Supponiamo di comprare un'obbligazione del valore nominale di \$100 del 15 Febbraio 2000, con cedole 5 1/2, e scadenza 15 Febbraio 2001. Cosi' abbiamo due scadenze T_1, corrispondente ai primi sei mesi, e T_2, che corrisponde alla maturita' dell'obbligazione ovvero un anno. Dopo sei mesi l'investitore ricevera' la prima cedola pari a $\$2.75 = (5 + 1/2)/2$, mentre dato che i sei mesi successivi corrispondono alla scadenza dell'obbligazione, oltre alla cedola, ricevera' anche il valore nominale per un totale di \$102.75 Quindi il valore attualizzato dei pagamenti che ricevera' l'investitore è rispettivamente $\$2.75 \cdot d(T_1)$ e $\$102.75 \cdot d(T_2)$; in totale il valore attualizzato di un'obbligazione con cedole 5 1/2 che restituisce \$100 alla scadenza è:

$$\$2.75 \cdot d(T_1) + \$102.75 \cdot d(T_2)$$

Intuitivamente tale quantità deve corrispondere al prezzo dell'obbligazione. Piu' in generale, sia data un'obbligazione avente valore nominale R, con cedole di pagamento C, maturita' T_M, e tempi di pagamento delle cedole $T_1, T_2, \dots, T_M$. Allora il valore attualizzato di tale obbligazione è dato da:

$$Cd(T_1) + Cd(T_2) + \ldots + (C + R)d(T_M) \tag{1.2}$$

Osserviamo infine, come sia comune in letteratura ([15]) assumere che i pagamenti delle cedole avvengano in tempo continuo piuttosto che a istanti di tempo prefissati. Si suppone pertanto l'esistenza di una funzione di sconto $d(T)$ definita sull'intervallo temporale $[0, T^*]$, $T^* > T_M$ in modo che il valore attualizzato dell'obbligazione sia dato da:

$$C \int_0^{T_M} d(s)ds + Rd(T_M) \tag{1.3}$$

È da sottolineare che l'esistenza della funzione di sconto definita su tutto l'intervallo temporale $[0, T^*]$ è ipotizzata per convenienza, in quanto rende piu' semplici alcuni calcoli.

In generale gli investitori sono interessati a valutare i tassi di interesse piuttosto che la funzione di sconto. Definiremo nei paragrafi successivi due tipi di tassi di interesse, i tassi *spot* e i tassi *forward* e vedremo in che relazione sono con la funzione di sconto.

1.3 Il tasso di interesse Spot

Supponiamo di investire \$100 per un anno con un tasso di interesse dell'8%, guadagnando cosi' \$8 alla fine dell'anno. Come si preferisce che vengano pagati questi \$8?

Si possono avere alla fine dell'anno riscattando così, \$108, oppure prendere \$4 dopo i primi sei mesi, e \$4 alla fine dell'anno. Nell'ultima ipotesi, dopo i primi sei mesi, si possono reinvestire i \$4, guadagnando così più di \$108 alla fine dell'anno. Investendo \$100 per sei mesi, si ha:

$$\$100 \cdot \left(1 + \frac{.08}{2}\right) = \$104$$

dove il termine $(1 + \frac{.08}{2})$ rappresenta il pagamento semi-annuale degli interessi. Per un anno invece, reinvestendo il capitale per ulteriori sei mesi si ha:

$$\$100 \cdot \left(1 + \frac{.08}{2}\right)^2 = \$108.16$$

dove il termine $(1 + \frac{.08}{2})^2$ è al quadrato, poiche' rappresenta il reinvestimento degli interessi riscossi dopo i primi sei mesi.

In generale, indicando con $z(T)$ il tasso di interesse nel periodo di tempo $[0, T]$, investendo un certo valore Y, al tempo 0, otteniamo alla fine di questo periodo la quantità:

$$X(T) = Y \cdot (1 + z(T)) \tag{1.4}$$

Se l'intervallo di tempo viene suddiviso in n sottointervalli di lunghezza ΔT tali che $n\Delta T = T$ e viene effettuato un pagamento parziale corrispondente a un tasso di interesse $z(T)\Delta T$ alla fine di ogni sottoperiodo, si ha:

$$\begin{array}{lll} X(\Delta T) & = Y \cdot (1 + z(T)\Delta T) & \\ X(2\Delta T) & = X(\Delta T)(1 + z(T)\Delta T) & = Y(1 + z(T)\Delta T)^2 \\ \vdots & = \vdots & = \vdots \\ X(n\Delta T) & = Y(1 + z(T)\Delta T)^n & = Y(1 + z(T)\Delta T)^{\frac{T}{\Delta T}} \end{array} \tag{1.5}$$

Definendo $\nu = 1/\Delta T$ la frequenza annuale dei pagamenti, e ponendo $X = X(T)$, la (1.5) diventa:

$$X = Y(1 + z(T)\Delta T)^{\frac{T}{\Delta T}} = Y \cdot \left(1 + \frac{z(T)}{\nu}\right)^{\nu T} \tag{1.6}$$

La formula precedente puo' essere usata per calcolare il tasso di interesse $z(T)$, nel momento in cui si conoscono X, T, Y, infatti:

$$z(T) = \nu \left[\left(\frac{X}{Y} \right)^{\frac{1}{\nu T}} - 1 \right]$$

Il tasso d'interesse $z(T)$ viene chiamato tasso *spot*. Poiche' X è la quantita' di denaro che riceveremo al tempo T, e Y l'investimento iniziale, riscrivendo la (1.6) come:

$$Y = X \cdot \frac{1}{(1 + \frac{z(T)}{\nu})^{\nu T}} \tag{1.7}$$

e confrontandola con la (1.1) possiamo scrivere la relazione che esiste tra la funzione di sconto $d(T)$ e il tasso di interesse spot $z(T)$:

$$d(T) = \frac{1}{(1 + \frac{z(T)}{\nu})^{\nu T}} \tag{1.8}$$

da cui segue:

$$z(T) = \nu \left[\left(\frac{1}{d(T)} \right)^{\frac{1}{\nu T}} - 1 \right] \tag{1.9}$$

È in comune in letteratura considerare il caso di interessi composti con continuita' (*continuos coumponding*), che equivale nella nostra notazione a suddividere l'intervallo $[0, T]$ in sottointervalli sempre più piccoli fino a far tendere $\Delta T \to 0$. In tal caso dalla (1.6) si ottiene:

$$\lim_{\Delta T \to 0} \left(1 + z(T)\Delta T \right)^{\frac{T}{\Delta T}} = \exp(Tz(T))$$

da cui segue:

$$d(T) = e^{-Tz(T)}.$$

Considerando ora il caso in cui si abbiano N scadenze di pagamento $T_1, \ldots, T_N$ ed i corrispondenti fattori di sconto $d(T_1), \ldots, d(T_N)$, assumeremo la seguente definizione:

Definizione 1.3.1 *Dati N fattori di sconto $d(T_1), d(T_2), \dots, d(T_N)$ dove $T_1, T_2, \dots,$ T_N, rappresentano gli istanti di pagamento, definiamo tasso spot la quantita':*

$$z(T_j) = -\frac{1}{T_j} \cdot \log(d(T_j)) \quad j = 1, \dots, N \tag{1.10}$$

1.4 Il tasso di interesse Forward

Un prestito *forward*, è un accordo fatto tra chi presta dei soldi per un certo periodo, e chi li prende in prestito. Il tasso di interesse su un prestito forward specificato al tempo dell'accordo è chiamato *tasso in avanti*. Fissando N istanti di tempo $T_0 = 0, T_1, T_2, \dots, T_N$, il tasso in avanti $\phi(T_{j-1}, T_j)$, è il tasso di interesse stabilito al tempo 0, che si applica per un prestito che avverrà nell'intervallo di tempo (T_{j-1}, T_j) per $j = 1, \dots, N$.

Vediamo in dettaglio la relazione che c'é con il tasso spot e con la funzione di sconto. Supponiamo di fissare dei tempi $T_0, T_1 = 0.5, T_2 = 1, T_3 = 1.5$, e siano $z(T_1) = z(0.5), z(T_2) = z(1), z(T_3) = z(1.5)$ i rispettivi tassi spot. È chiro che $\phi(T_0, T_1) = z(T_1)$ per definizione. Un investimento di \$1 con composizione annuale di interessi, riporta un guadagno al tempo T_2 di $(1 + z(T_2))^{T_2} = 1 + z(1)$. Questo investimento può essere interpretato come la combinazione di un prestito forward nel periodo $(T_0, T_1]$ con un tasso $\phi(T_0, T_1)$ seguito da un altro prestito forward nel periodo $(T_1, T_2]$ con un tasso $\phi(T_1, T_2)$. Poichè il guadagno ottenuto deve essere lo stesso deve essere:

$$(1+\phi(T_0,T_1))^{T_1-T_0} \cdot (1+\phi(T_1,T_2))^{T_2-T_1} = (1+\phi(T_0,T_1)) \cdot (1+\phi(T_1,T_2)) = 1+z(T_2)$$

da cui si puo' ricavare $\phi(T_1, T_2)$. In generale per degli istanti di tempo $T_0 = 0, T_1, T_2, \dots, T_j$ si ha:

$$\begin{array}{lll} \frac{1}{d(T_1)} & = (1+z(T_1))^{T_1} & = (1+\phi(T_0,T_1))^{T_1} \\ \frac{1}{d(T_2)} & = (1+z(T_2))^{T_2} & = (1+\phi(T_0,T_1))^{T_1} \cdot (1+\phi(T_1,T_2))^{T_2-T_1} \\ \vdots & = \vdots & = \vdots \\ \frac{1}{d(T_j)} & = (1+z(T_j))^{T_j} & = (1+\phi(T_0,T_1)^{T_1} \cdot (1+\phi(T_1,T_2))^{T_2-T_1} \cdot \dots \\ & & \cdot (1+\phi(T_{j-2},T_{j-1}))^{T_{j-1}-T_{j-2}} \cdot (1+\phi(T_{j-1},T_j))^{T_j-T_{j-1}} \end{array} \tag{1.11}$$

dove $d(T_j)$ sono i fattori di sconto e $z(T_j)$ e $\phi(T_{j-1}, T_j)$ sono il corrispondente tasso spot e tasso in avanti. Supponiamo di conoscere i fattori di sconto $d(T_1), d(T_2), \dots, d(T_j)$, dalla (1.11) otteniamo:

$$\frac{\frac{1}{d(T_j)}}{\frac{1}{d(T_{j-1})}} = \frac{(1+\phi(T_0,T_1))^{T_1} \cdot (1+\phi(T_1,T_2))^{T_2-T_1} \cdot \ldots \cdot (1+\phi(T_{j-1},T_j))^{T_j-T_{j-1}}}{(1+\phi(T_0,T_1))^{T_1} \cdot (1+\phi(T_1,T_2))^{T_2-T_1} \cdot \ldots \cdot (1+\phi(T_{j-2},T_{j-1}))^{T_{j-1}-T_{j-2}}}$$

che implica:

$$\frac{d(T_{j-1})}{d(T_j)} = (1+\phi(T_{j-1},T_j))^{T_j-T_{j-1}}$$

Ponendo:

$$(1+\phi(T_{j-1},T_j))^{T_j-T_{j-1}} := e^{f(T_j)(T_j-T_{j-1})} \tag{1.12}$$

possiamo dare la seguente definizione:

Definizione 1.4.1 Dati N fattori di sconto $d(t_1), \dots, d(t_N)$, si definisce il tasso forward ad un passo la quantita':

$$f(T_j) = -\frac{1}{T_j - T_{j-1}} \log \frac{d(T_j)}{d(T_{j-1})} \tag{1.13}$$

In generale considerando la funzione di sconto $d(T)$, per un intervallo di tempo arbitrario $[T, T + \triangle T]$, il tasso forward diventa:

$$f(T) = -\frac{1}{\triangle T} \log \frac{d(T+\triangle T)}{d(T)} \tag{1.14}$$

Assumendo la differenziabilita' di d, e facendo il limite per $\triangle T \to 0$ abbiamo che:

$$f(T) = -\frac{d}{dT} \log d(T) \tag{1.15}$$

che usualmente definito come tasso forward istantaneo. La relazione che lega la funzione di sconto con il tasso di interesse forward istantaneo segue dalla (1.15):

$$d(T) = \exp\left(- \int_0^T f(u)du \right) \tag{1.16}$$

1.5 L'equazione delle obbligazioni: arbitraggio e legge del prezzo unico

Come abbiamo osservato nel paragrafo (1.2), risulta intuitivo che il prezzo di mercato di una obbligazione uguagli il suo valore attualizzato ([2]) Poiche' le quotazioni di mercato riportano il prezzo di vendita p^b (*bid*) e di acquisto p^a (*ask*), è usuale considerare come prezzo p dell'obbligazione la media $p = \frac{p^a+p^b}{2}$. Indicheremo inoltre con $\triangle p = p^a - p^b$ lo scarto *bid-ask*. Data quindi un'obbligazione di maturità T_m, risultano definiti gli istanti di tempo $T_1, T_2, \ldots, T_m$, corrispondenti ai pagamenti delle cedole C. Dall'equazione (1.2) segue dunque che il prezzo p di tale obbligazione è:

$$\begin{aligned} p &= \delta(T_1)C + \delta(T_2)C + \ldots + \delta(T_m)(C+R) \\ &= C \textstyle\sum_{i=1}^n \delta(T_i) + \delta(T_m)R \end{aligned} \tag{1.17}$$

dove $d(T_1), d(T_2), \ldots, d(T_m)$, sono i fattori di sconto relativi ai tempi di pagamento e R il valore nominale dell'obbligazione.

Nella pratica le obbligazioni possono essere vendute in qualsiasi momento, cambiando cosi' proprietario da un giorno all'altro; per valutare allora l'effettivo valore dell'obbligazione, bisogna calcolare l'interesse accumulato (accrued interest), sul pagamento della cedola successiva che è calcolato come ([2]):

$$ai = \text{interesse accumulato} = T_0 C \tag{1.18}$$

dove T_0 è la proporzione di tempo passato tra due tempi successivi di pagamento delle cedole. La (1.17) diventa allora:

$$p + ai = C \sum_{i=1}^{n} \delta(T_i) + \delta(T_m) R \tag{1.19}$$

dove p è detto prezzo netto (*clean price*) e $p + ai$ prezzo lordo (*gross price*).

Abbiamo finora considerato la relazione tra il prezzo ed il valore attualizzato espresso tramite la funzione di sconto nel caso di una singola obbligazione. Il mercato tuttavia fornisce ad ogni istante di tempo i prezzi di vendita e di acquisto dell'insieme delle obbligazioni correntemente in vita (*outstanding*) che indicheremo con $\mathcal{B}$. Formalmente rappresenteremo l'i-esima obbligazione $B_i \in \mathcal{B}$ con una tripla $(C_i, T_{M_i}, p_i) \in [0, +\infty) \cdot \Re_+ \cdot \Re_+$, dove $C_i \geq 0$ è la cedola, T_{M_i} la maturita' e p_i è il prezzo corrente. Usualmente si considera un valore nominale R di riferimento fissato: nel caso delle obbligazioni US $R = \$100$.

I prezzi delle obbligazioni *outstanding* non possono essere indipendenti l'uno dall'altro altrimenti si verrebbero a creare come vedremo delle cosidette opportunita' di arbitraggio. In generale ciò è escluso assumendo l'esistenza di una funzione di sconto decrescente $d : [0, +\infty) \to [0, 1]$ con $d(0) = 1$ e tale che:

$$p_i = C_i \sum_{j=1}^{n_i} d(T_j^{(i)}) + 100 d(T_{M_i}^{(i)}) \quad \forall\ B_i \in \mathcal{B} \tag{1.20}$$

dove i tempi $T_j^{(i)}$ sono i j-esimi istanti di pagamento delle cedole C_i relative all'obbligazione B_i. L'equazione (1.20) costituisce il punto di partenza per affrontare il problema di determinare le quantita' di interesse, cioè la funzione di sconto, i tassi spot, i tassi forward, avendo a disposizione i prezzi p_i delle obbligazioni *outstanding*.

Vedremo ora brevemente in cosa consiste un'opportunità di arbitraggio.

Arbitraggio e legge del prezzo unico Le opportunità di arbitraggio sono sostanzialmente legate alla cosidetta *legge del prezzo unico*, che in generale afferma che due strumenti finanziari che garantiscono lo stesso flusso di cassa devono avere il medesimo prezzo.

Consideriamo infatti di comprare il 15 Febbraio 2000 un'obbligazione con maturità 15 Agosto 2000 e cedola 6 7/8 ad un prezzo di \$101.625. Poichè il

cedola	maturita'	prezzo
6 7/8	8/15/00	101.62
5 1/2	2/15/01	101.56
7 3/4	2/15/01	103.75

Tabella 1.1: Cedole, maturità e prezzi di tre obbligazioni del 15 Febbraio 2000

valore attualizzato dell'obbligazione corrisponde al suo prezzo corrente si ha che

$$101.625 = (\frac{1}{2}6.875 + 100)d(.5)$$

che implica $d(.5) = .9824$. Sempre il 15 Febbraio 2000 consideriamo un'altra obbligazione con maturità 15 Febbraio 2001, cedola 5 1/2 e prezzo \$101.5625. Uguagliando il prezzo al valore attualizzato dell'obbligazione otteniamo

$$101.5625 = \frac{1}{2}5.5 \cdot d(.5) + (\frac{1}{2}5.5 + 100) \cdot d(1)$$

da cui, noto $d(.5)$, si ricava $d(1) = .9621$.

Consideriamo ora un ulteriore obbligazione nel giorno 15 Febbraio 2000, con maturità 15 Febbraio 2001, cedola 7 3/4 e prezzo \$103.75.

Utilizzando i fattori di sconto $d(.5)$ e $d(1)$, relativi rispettivamente ai primi sei mesi e alla scadenza dell'obbligazione, già calcolati, possiamo scrivere di nuovo:

$$103.75 = \frac{1}{2}7.75 \cdot d(.5) + (\frac{1}{2}7.75 + 100) \cdot d(1)$$

che corrisponde all'effettivo prezzo di mercato. Abbiamo quindi utilizzato gli stessi fattori di sconto per obbligazioni diverse. Cosa succederebbe infatti se invece di \$103.75, violando la legge del prezzo unico, avessimo \$103.80? La risposta è che si verrebbe a creare una situazione di opportunita' di arbitraggio.

L'opportunita' di arbitraggio è un termine usato dai partecipanti del mercato azionario, per indicare una situazione vantaggiosa dal punto di vista finanziario, permettendo un guadagno certo senza correre alcun rischio.

Nel momento in cui l'obbligazione considerata ha prezzo di \$103.80, invece di \$103.75, si puo' cercare di creare un portafoglio replicante. Tale portafoglio replicante è costituito da un insieme di obbligazioni che garantiscono all'investitore lo stesso flusso di cassa dell'obbligazione data. Supponiamo quindi nel nostro esempio di poter costruire un portafoglio replicante:

- compro \$10.59 dell'obbligazione 6 7/8, 15 Agosto 2000 del valore attualizzato di 101.6250%
- compro $\$1,010.95$ dell'obbligazione 5 1/2, 15 Febbraio 2001 del valore attualizzato 101.5625%
- vendo $\$1,000$ dell'obbligazione 7 3/4, 15 Febbraio 2001 del valore attualizzato 103.8000%

quando acquisto questi titoli, si dice che si assume una posizione short, nel momento in cui vendo l'obbligazione considerata, si assume una posizione long. La posizione short consiste come in un prestito bancario, infatti nel momento in cui compro l'obbligazione, ricevo subito il valore nominale dell'obbligazione con le rispettive cedole, e pago ogni sei mesi i tassi di interesse fino alla scadenza, in cui dovro' pagare anche l'effettivo valore dell'obbligazione stessa. Vediamo i flussi di cassa del portafoglio in esame: il flusso di cassa relativo al 15 Febbraio 2000 sara' positivo, infatti acquisto $10.59 \cdot 101.6250\%$ e $1,010.95 \cdot 101.5625\%$ rispettivamente della prima e della seconda obbligazone, mentre vendo $1,000 \cdot 103.800\%$ dell'obbligazone da replicare, guadagnando cosi' \$50. Il flusso di cassa del 15 Agosto 2000 è zero, poiche' dalle prime due obbligazioni riscuoto rispettivamente i seguenti interessi $\$10.59 \cdot 103.4375\%$ e $1,010.95 \cdot 2.7500\%$, mentre devo pagare $\$1,000 \cdot 3.8750\%$ dell'obbligazione replicata; anche il flusso di cassa del 15 Febbraio 2001 è zero, infatti riscuoto $\$1,010.9489 \cdot 102.750$ e pago $\$1,000.000 \cdot 103.875\%$.

Nel seguente schema vediamo quali sono le condizioni per ottenere un potafoglio replicante:

15/2/00	15/8/00	15/2/01
7 3/4	3.875	103.875
5 1/2	$q_1 \cdot 2.75$	$q_1 \cdot 102.75$
6 7/8	$q_2 \cdot 103.4375$	

Tabella 1.2: Portafoglio replicante

Nella prima colonna sono indicate le cedole, nella seconda colonna e nella terza colonna sono indicati i flussi di cassa dell'obbligazione originale e del portafoglio replicante alla prima data di scadenza e alla maturita' dell'obbligazione. Risolvendo le seguenti equazioni:

$$q_1 \cdot 102.75 = 103.875$$
$$q_1 \cdot 2.75 + q_2 \cdot 103.4375 = 3.875$$

otteniamo q_1 e q_2, e possiamo concludere dicendo che c'é arbitraggio nel momento in cui:

$$101.56 \cdot q_1 + 101.625 \cdot q_2 < 103.875$$

È noto ai praticanti del mercato che quando si verifica una situazione di questo tipo, con il passare del tempo l'arbitraggio tende a svanire . Infatti saranno molti gli investitori che sapranno di tale arbitraggio, e di conseguenza il prezzo delle obbligazioni cambiera' fino ad annullare l'arbitraggio stesso.

Capitolo 2

La stima della struttura a termine con funzioni spline

Nel precedente capitolo abbiamo visto in dettaglio la stretta relazione che c'è tra la funzione di sconto, il tasso di interesse spot e il tasso forward. Il problema pero' è che queste curve non si possono osservare direttamente, ma possono essere solo stimate a partire dalla conoscenza dei prezzi delle obbligazioni, delle loro maturità e delle rispettive cedole di pagamento.

Indicando con N il numero delle obbligazioni *outstanding*, da un punto di vista matematico, si tratta quindi di risolvere un problema di regressione, basato sulle equazioni:

$$p_i = C_i \sum_{j=1}^{n_i} d(T_j^{(i)}) + 100 d(T_{M_i}^{(i)}) \quad \forall\ i = 1, \dots, N \tag{2.1}$$

Nella pratica, come suggerito da McCulloch ([15]) prima e Carleton e Cooper ([4]) poi, non è detto che le equazioni (2.1) siano verificate esattamente: il modello comunemente usato è dato da:

$$p_i = C_i \sum_{l=1}^{n_i} \delta(t_l) + \delta(t_n) R_i + \varepsilon_i \tag{2.2}$$

dove gli ε_i sono un termine di errore di media nulla. Questi errori possono essere causati da differenti fattori: la presenza di uno scarto tra il prezzo di vendita e quello d'acquisto, i costi di transazione e commissivi, problemi di tassazione, e temporanee opportunita' di arbitraggio. In generale si assume che le variabili aleatorie ε_i, con $i = 1, \ldots, N$, sono indipendenti, con varianza σ_i^2 strettamente positiva.

Seguendo McCulloch ([15]) è ragionevole porre tale varianza proporzionale allo scarto bid-ask, quindi $\sigma_i^2 = \sigma^2(\triangle p_i)^2$, dove σ^2 è un fattore di proporzionalita' comune.

In questo capitolo descriveremo i metodi di McCulloch, le spline esponenziali, e i modelli "parsimoniosi"; tutti quanti nello stimare le curve scelte adottano il metodo dei minimi quadrati che verra' descritto nel capitolo 3.

2.1 Metodo di McCulloch

Il metodo di McCulloch si basa sulla stima della funzione di sconto, da cui poi si possono ricavare il tasso spot e il tasso forward in base alle relazioni descritte nel capitolo precedente. Scriviamo prima di tutto la funzione di sconto come combinazione lineare di una base di funzioni che verra' scelta a priori:

$$d(t) = \sum_{j=0}^{k} a_j f_j(t) = 1 + \sum_{j=1}^{k} a_j f_j(t) \tag{2.3}$$

dove le $f_j(t)$ sono k funzioni continue e differenziabili, tali che $f_j(0) = 0$, gli a_j sono i corrispondenti coefficienti per $j = 1, \ldots k$, $a_0 = 1$,in modo che $d(0) = 1$. La scelta di una base di funzioni, è molto importante, in quanto caratterizza la stima della funzione di sconto, e questa scelta verra' discussa nel capitolo successivo. Prendiamo in esame la formula relativa all'i-esima obbligazione

$$p_i = C_i \sum_{l=1}^{n_i} \delta(t_l) + d(t_n) R_i + \varepsilon_i \qquad i = 1, \ldots N, \tag{2.4}$$

dove si ha:

$$
\begin{array}{rcl}
C_i & = & \text{valore della cedola,} \\
R_i & = & \text{valore nominale,} \\
t_l & = & \text{il tempo del pagamento dell'l-esima cedola,} \\
n_i & = & \text{\# di pagamenti fino alla maturita'}
\end{array}
$$

Sostituendo la (2.3) nella (2.4) otteniamo:

$$
\begin{aligned}
p_i &= C_i \sum_{l=1}^{n_i} [1 + \sum_{j=1}^{k} a_j f_j(t_l)] + R_i [1 + \sum_{j=1}^{k} a_j f_j(t_n)] + \varepsilon_i \\
&= nC_i + C_i \sum_{l=1}^{n_i} \sum_{j=1}^{k} a_j f_j(t_l) + R_i + R_i \sum_{j=1}^{k} a_j f_j(t_n) + \varepsilon_i
\end{aligned}
$$

da cui

$$
p_i - nC_i - R_i = \sum_{j=1}^{k} a_j [C_i \sum_{l=1}^{n_i} f_j(t_l) + R_i + R_i f_j(t_n)] + \varepsilon_i
$$

Cosi' ponendo:

$$
\begin{gathered}
y_i = p_i - nC_i - R_i, \\
x_{ij} = C_i \sum_{l=1}^{n_i} f_j(t_l) + R_i + R_i f_j(t_n)
\end{gathered}
$$

dalla (2.4) otteniamo le equazioni:

$$
y_i = \sum_{j=1}^{k} a_j x_{ij} + \varepsilon_i \qquad i = 1, \ldots, N. \tag{2.5}
$$

ovvero

$$
\underline{y} = X \underline{a} + \underline{\varepsilon} \tag{2.6}
$$

dove $\underline{y} = (y_i)_{i=1,\dots,N}$, $X = (x_{i,j})_{i=1,\dots,N,j=1,\dots,k}$ e $\underline{\varepsilon} = (\varepsilon_i)_{i=1,\dots,N}$. Si tratta quindi di un problema di regressione lineare, che puó essere risolto con il metodo dei minimi quadrati, per ottenere la stima dei coefficienti $\hat{a}_j$, da cui segue la corrispondente stima della funzione di sconto:

$$\hat{\delta}(t) = 1 + \sum_{j=1}^{k} \hat{a}_j f_j(t). \tag{2.7}$$

Come osservato in McCulloch ([15]) il modo piu' semplice di scegliere le funzioni f_j e' quella di porre $f_j(t) = t^j$ per $j = 1, \dots, k$, in modo tale che le funzioni approssimanti risultino dei polinomi di grado k. Tale scelta pero' non si dimostra del tutto soddisfacente provoca un'instabilita' nella stima dei parametri nel momento in cui si aumenta il grado k, e soprattutto perchè i polinomi hanno una capacita' di risoluzione uniforme nell'intervallo di interesse. Per risolvere questi problemi, McCulloch suggeri' l'uso delle "spline", che risultano le funzioni approssimanti piu' utilizzate finora ([2]). Funzione spline è diventato sinonimo di funzione polinomiale a tratti (piece polynomials) e trae origine dal fatto che disegna curve "regolari" (cioè con un certo grado di continuita') attraverso punti prefissati. In un primo momento McCulloch aveva utilizzato le spline quadratiche ([15]), che pero' possono creare delle "nocche" nella curva forward in corrispondenza dei punti di raccordo dei polinomi. Questo è causato dal fatto che la derivata seconda della funzione di sconto, e quindi la derivata prima del tasso forward, risulta in questo caso discontinua. Infatti:

$$f^{'}(t) = \left(\frac{\delta^{'}(t)}{\delta(t)}\right) - \left(\frac{\delta^{''}(t)}{\delta(t)}\right). \tag{2.8}$$

Successivamente McCulloch aveva utilizzato le spline cubiche che si erano dimostrate certamente piu' efficaci.

2.2 Spline esponenziali

Una delle principali critiche che sono state mosse all'uso di funzioni polinomiali a tratti, fu quella che la curva stimata del tasso forward, puo' presentare proprieta' indesiderate, come ad esempio una rapida crescita o decrescita per lunghe maturita'.

Vasicek e Fong ([24]) hanno descritto un metodo che puó essere usato per produrre un andamento asintotico piu' realistico della curva forward. Il loro metodo si basa nell'assumere che la funzione di sconto abbia una pendenza esponenziale, in accordo con la moderna teoria dell'equilibrio ([2]). Vasicek e Fong sostengono che le spline normali non producono un buon adattamento locale della funzione di sconto, e inoltre le spline polinomiali usate da McCulloch non possono essere forzate per ottenere una forma esponenziale della funzione di sconto quando le maturita' crescono.

Vasicek e Fong suggeriscono di modificare l'argomento t della funzione di sconto $\delta(t)$ nella seguente forma:

$$t = -\left(\frac{1}{\alpha}\right)\log(1-x) \qquad \text{per } 0 \leq x < 1 \tag{2.9}$$

dove il parametro α rappresenta il valore limite del tasso forward che deve essere stimato.

Questa scelta comporta la trasformazione della funzione di sconto da una funzione approssimativamente esponenziale di t, ad una funzione lineare in x, (dove x rappresenta il tempo trasformato). Le spline polinomiali possono allora essere usate per stimare questa funzione di sconto modificata. Il tutto risulta piu' vantaggioso, in quanto si possono imporre dei vincoli sulla funzione di sconto, come per esempio la non-negativita'.

Vasicek e Fong per stimare la funzione di sconto propongono di utilizzare le spline cubiche, (che verranno descritte nel capitolo successivo): cio' implica che in termini della variabile originale t, la funzione di sconto tra due nodi successivi assume la forma:

$$\delta(t) = b_0 + b_1 e^{-\alpha t} + b_2 e^{-2\alpha t} + b_3 e^{-3\alpha t} \tag{2.10}$$

dove i coefficienti b_i sono i parametri da stimare.

Sebbene Vasicek e Fong sostengono che le spline esponenziali portano a migliori risultati rispetto agli altri metodi usati, non presentarono alcuna prova numerica che testimoniasse la loro tesi. Successivamente Shea ([19]), presentando alcuni risultati empirici, aveva provato che le spline esponenziali non producevano migliori stime della struttura a termine dei tassi di interesse rispetto alle spline polinomiali. Inoltre aveva notato che uno dei fattori che causava l'instabilita' numerica del modello di Vasicek e Fong, era causata dalla trasformazione esponenziale $x = 1 - e^{-\alpha t}$. Piccoli valori di α causano un addensarsi dei valori delle variabili x in alcune porzioni dell' intervallo, provocando instabilita' nella stima del tasso forward per lunghe maturita'.

Langetieg e Smoot ([13]) hanno proposto di stimare la curva dei tassi spot usando le spline cubiche, che è equivalente a stimare la funzione di sconto con una funzione esponenziale:

$$\delta(t) = \exp\left[-t \sum_{j=1}^{k} a_j f_j(t)\right] \tag{2.11}$$

Come Vasicek e Fong, Langetieg e Smoot, sostengono che se la funzione di sconto è veramente una funzione esponenziale, allora la scelta di un modello esponenziale produce migliori risultati piuttosto che stimare la funzione di sconto direttamente con una spline cubica. Al contrario non assumere la funzione di sconto come una esponenziale ma cercando di approssimarla come tale, causera' un adattamento inferiore.

Langetieg e Smoot hanno trovato il loro metodo migliore di quello di Vasicek e Fong, poichè consideravano la trasformazione esponenziale un metodo per approssimare il loro modello esponenziale.

Coleman, Fisher e Ibbotson ([6]), hanno stimato la curva relativa al tasso forward piuttosto che la funzione di sconto. Essi approssimano il tasso forward istantaneo $f(t)$ come una funzione costante a tratti:

$$f(t) = f_j \qquad \text{per } t_{j-1} < t \leq t_j \tag{2.12}$$

Questo è equivalente ad usare una spline esponenziale per approssimare la funzione di sconto: infatti dalla (1.16) per $t_{j-1} < t < t_j$ si ha:

$$\delta(t) = \exp\{-[f_1 t_1 + f_2(t_2 - t_1) + \ldots + f_{j-1}(t_{j-1} - t_1) + f_j(t - t_{j-1})]\} \tag{2.13}$$

Coleman, Fisher e Ibbotson hanno sostenuto che considerare su un certo periodo il tasso forward costante, porta ad una stima migliore invece di usare una spline cubica o quadratica. Nonostante la funzione di sconto, prodotta da questo metodo, risultasse continua, la sua derivata prima è discontinua. Questo pero' non era un problema secondo gli autori, poiche' l'unico vincolo richiesto sulla funzione di sconto è che questa risultasse monotona decrescente, e che ulteriori vincoli non sono giustificati da un punto di vista strettamente finanziario.

2.3 Il problema della stima usando le Spline

Shea ([19]) ha constatato alcuni problemi nell'adottare le funzioni spline per approssimare la struttura a termine dei tassi di interesse. Infatti, egli ha dimostrato che le spline cubiche usate da McCulloch, non vincolano la monotonia della funzione di sconto, producendo cosi' una stima della stessa funzione di sconto che per lunghe maturita' può essere crescente, avendo come conseguenza dei valori negativi della curva del tasso forward.

Un'alternativa proposta da Shea, è di utilizzare dei vincoli "ad hoc" sulla locazione dei nodi e sulla derivata prima di δ. L'introduzione di questi aggiustamenti puo' causare diversi problemi nell'interpretare la struttura a termine dei tassi di interesse; infatti alcuni cambiamenti nelle curve stimate, possono essere attribuite erroneamente ad eventi verificatisi nei mercati, mentre invece dipendono dai vincoli imposti a priori.

I metodi per approssimare queste curve, differiscono l'uno dall'altro, anche nella scelta del numero e della locazione dei nodi. Un numero di nodi molto basso non riesce ad approssimare bene una funzione con una forte pendenza, mentre al contrario un alto numero di nodi, puo' comportare un eccessivo adattamento delle stime ai dati osservati (*overfitting*).

Il metodo adottato da McCulloch ([16]), consiste nello scegliere il numero di nodi k uguale alla radice quadrata del numero delle obbligazioni ($k = \sqrt{n}$), e di posizionarli nei punti $t_1, \ldots, t_k$, in modo tale che $t_1 = 0$ e t_k sia uguale alla piu' lunga scadenza delle obbligazioni considerate, mentre gli altri sottointervalli $[t_i, t_{i+1}]$ contengono approssimativamente lo stesso numero di maturita' osservate. Un vantaggio di questo metodo, è che il posizionamento dei nodi dipende dai dati osservati. Al contrario, un metodo alternativo fu proposto da Litzenberger e Rolfo ([2]), che posizionano i nodi per $t = 1, 5, 10$

anni, che corrisponde rispettivamente al comportamento a breve, medio e lungo termine. Secondo Langetieg e Smoot, quest'ultimo modello ha un migliore comportamento di quello di McCulloch, e puó essere reso ancora piu' efficiente aggiungendo un ulteriore nodo per $t = 0.5$.

2.4 Modelli parsimoniosi

Un metodo alternativo alle spline per la stima della struttura a termine dei tassi di interesse, è quello costituito dall'uso di modelli cosi' detti *parsimoniosi*. L'idea base di questa classe di modelli, è quella di imporre a priori alcune proprieta' della funzione di sconto, come la monotonia, e contemporaneamente di cercare di diminuire il numero di parametri da stimare.

Il modello proposto da Chambers-Carleton-Waldman ([5]), per la funzione di sconto è costituito da un polinomio esponenziale:

$$\delta(t, \alpha) = \exp\left(- \sum_{j=1}^{k} \alpha_j t^j \right). \tag{2.14}$$

I parametri $\alpha_1, \ldots, \alpha_N$ possono essere stimati con il metodo dei minimi quadrati non lineari o con il metodo di massima verosimiglianza, una volta esplicitata la densita' dell'errore nell'equazione (2.4).

Un altro modello è quello di Nelson-Siegel, che proposero esplicitamente una forma funzionale per la curva dei tassi forward:

$$f(t, \alpha) = \beta_0 + \beta_1 e^{-\frac{t}{\tau}} + \beta_2 \frac{t}{\tau} e^{-\frac{t}{\tau}} \tag{2.15}$$

dove $\underline{\alpha} = (\beta_0, \beta_1, \beta_2, \tau)$, sono i parametri da stimare. Di conseguenza, ricordando che

$$\delta(t) = \exp\left(- \int_0^t f(t) dt \right)$$

la funzione di sconto è del tipo:

$$\delta(t,\alpha) = \exp(-t(\beta_0 + (\beta_1+\beta_2)\frac{\tau}{t}(1-e^{-\frac{t}{\tau}}) - \beta_2 e^{-\frac{t}{\tau}})). \tag{2.16}$$

La caratteristica di questo modello consiste nel fatto che la curva forward, è combinazione lineare di tre componenti, $f_s(t) := e^{-\frac{t}{\tau}}$, $f_m(t) := \tau \cdot e^{\frac{t}{\tau}}$ e $f_l(t) := 1$, che rappresentano rispettivamente il comportamento a breve, medio e lungo termine, del tasso forward. I parametri $\beta_0, \beta_1, \beta_2$ misurano rispettivamente il peso di ciascuna componente, mentre τ è il fattore scala del tempo.

In Svensson ([23]) un ulteriore fattore è stato introdotto nella (2.15) per incrementare ulteriormente la flessibilita' del modello di Nelson e Siegel. La curva forward diventa:

$$f(t) = \beta_0 + \beta_1 \exp\left(-\frac{t}{\tau_1}\right) + \beta_2\left[\frac{t}{\tau_1}\exp\left(-\frac{t}{\tau_1}\right)\right] + \beta_3\left[\frac{t}{\tau_2}\exp\left(-\frac{t}{\tau_2}\right)\right] \tag{2.17}$$

dove sono stati aggiunti i due parametri β_3 e τ_2.

In generale i modelli parsimoniosi sono caratterizzati da una minore capacita' di adattamento ai dati, rispetto ai modelli basati sulle spline. Inoltre da un punto di vista computazionale richiedono l'uso di tecniche di minimizzazione non lineare e le stime ottenute sono tipicamente dei minimi locali. Tuttavia hanno molti meno parametri, e non richiedono la scelta del numero e del posizionamento dei nodi.

Nelle figure (2.1) (2.2) (2.3), sono riportate delle famiglie di curve generate dal modello di Nelson e Siegel dando dei valori arbitrari ai parametri $\alpha = (\beta_0,\beta_1,\beta_2,\tau)$. In particolare abbiamo scelto $\beta_0 = 0.08$, $\beta_1 = 0.2 - \beta_0$, $\tau = 1$ e β_2 che varia da -0.15 a 0.6 con un passo di 0.05. Si può notare che la somma $\beta_0 + \beta_1$ indica l'origine del tasso spot e del tasso forward, mentre β_0 rappresenta l'asintoto a cui tendono il tasso spot e forward.

Figura 2.1: Funzione di sconto di Nelson-Siegel

Figura 2.2: Spot generato dalla funzione di sconto di Nelson-Siegel

Figura 2.3: Forward generato dalla funzione di sconto di Nelson-Siegel

Capitolo 3

Funzioni Spline e Metodi Adattativi

3.1 Funzioni Spline

Il metodo di McCulloch è caratterizzato dalla scelta di un insieme di funzioni f_j, con $j = 1, \dots, n$, che hanno un ruolo fondamentale nell'approssimazione della funzione di sconto. Nel voler approssimare una determinata funzione, i polinomi hanno un ruolo importyante come funzioni approssimanti in un determinato intervallo. Se l'intervallo $[a, b]$, in cui si vuole approssimare una data funzione è molto grande, e si hanno molti dati per non avere un grado del polinomio molto alto, si preferisce suddividere l'intervallo $[a, b]$ in piccoli sottointervalli $[\xi_i, \xi_{i+1}]$, con $[a = \xi_0 < \xi_1 < \dots < \xi_n = b]$, in modo che in ogni sottointervallo si abbia un polinomio p_i, con un grado accettabile, che approssima localmente la funzione in esame. Funzioni di questo tipo prendono il nome di spline. In generale possiamo dire che data una funzione $f \in C[a, b]$, e $\triangle := [a = \xi_0 < \xi_1 < \dots < \xi_n = b]$ una suddivisione dell'intervallo $[a, b]$, una funzione spline di grado r e ordine $k = r + 1$ che indicheremo con Sp, con punti di raccordo $\xi_i, i = 1, \dots, n$, è una funzione $\xi \mapsto Sp(\xi)$, con $\xi \in [a, b]$, con le seguenti proprieta':

1. Su ogni intervallo $[\xi_i, \xi_{i+1}]$ con $i = 0, \dots, n-1$, è un polinomio di grado r,

2. la funzione $Sp(\xi)$ e le sue prime $r - 1$ derivate sono continue sul-

l'intervallo $[a, b]$, cioé $Sp(\xi) \in C^{r-1}([a, b])$, tranne al più nei punti ξ,

3. $(D^{m-1}Sp)(\xi_{j+1}^{+}) = (D^{m-1}Sp)(\xi_{j+1}^{-})$ per $m = 1, \ldots, v_j$, dove $v_j \in \{1, \ldots k\}$ è fissato e $(D^{m-1}Sp)(\xi) = \frac{d^{m-1}Sp}{dx^{m-1}}$ calcolata in ξ se $m > 1$, altrimenti $(D^0 Sp)(\xi) = Sp(\xi)$

Questo vuol dire che in ogni punto di raccordo interno i polinomi si uniscono con un certo grado di continuita'. Cosi' una spline di ordine $k = 3$, è localmente un polinomio quadratico, con derivata prima e seconda continua. Se tutti i valori v_j sono k, allora la funzione spline è un singolo polinomio di grado $r - 1$ e di ordine k.

Supponendo di fissare una sequenza di nodi t, è conveniente rappresentare una funzione spline, come combinazione lineare di un insieme fissato (una base) di funzioni spline $M_i(\cdot/r, t)$ per $i = 1, \ldots, n$, di grado r. Data una tale base si può scrivere la funzione spline f di grado k come combinazione lineare di $M_i(\cdot/r, t)$:

$$Sp = \sum_{j=1}^{n} a_i M_i \tag{3.1}$$

per un certo insieme di parametri a_i. Diverse basi spline si possono utilizzare nello studio dell'approssimazione della struttura a termine dei tassi di interesse, come ad esempio le B-spline e quelle usate da McCulloch.

3.2 Le Spline di McCulloch

Supponendo di avere k nodi $\xi_1, \ldots, \xi_k$, McCulloch scelse le seguenti spline cubiche:

per $x < \xi_{j-1}$

$$f_j(x) = 0$$

per $\xi_{j-1} \leq x < \xi_j$

$$f_j(x) = \frac{(x - \xi_{j-1})^3}{6(\xi_j - \xi_{j-1})}$$

per $\xi_j \le x < \xi_{j+1}$

$$f_j(x) = \frac{c^2}{6} + \frac{ce}{2} + \frac{e^2}{2} - \frac{e^3}{6(\xi_{j+1} - \xi_j)}$$

dove

$$c = \xi_j - \xi_{j-1}$$
$$e = x - \xi_j$$

per $\xi_{j+1} \le x$

$$f_j(x) = (\xi_{j+1} - \xi_{j-1}) \left[\frac{2\xi_{j+1} - \xi_j - \xi_{j-1}}{6} + \frac{x - \xi_{j+1}}{2} \right]$$

per $j = k$

$$f_k(x) = x \qquad \forall x$$

Piu' in generale, in Marangio, Ramponi, Bernaschi ([9]) è stata utilizzata la seguente base spline:

$$f_i(x) = x^i \qquad (i = 0, \ldots, r-1)$$

$$f_i(x) = \begin{cases} 0 & \text{se } \ x \le \xi_{i-r+1} \\ (x - \xi_{i-r+1})^r & \text{se } \ \xi_{i-r+1} < x \le \xi_{i-r+2} \\ (x - \xi_{i-r+1})^r - (x - \xi_{i-r+2})^r & \text{se } \ \xi_{i-r+2} < x \end{cases} \tag{3.2}$$

per $i = r, \ldots, r + k - 2$.

L'insieme di tutte le funzioni spline di grado k, formano una spazio lineare di dimensione $k + r - 2$; possiamo infatti scegliere arbitrariamente $k + 1$ coefficienti dei polinomi nel primo intervallo $[\xi_1, \xi_2]$, e un coefficiente per ciascun polinomio nei $k - 2$ intervalli rimanenti $[\xi_i, \xi_{i+1}]$ $(i = 2, \ldots, k-1)$. Gli altri parametri sono fissati dalle condizioni $f^{(n)}(\xi_i^+) = f^{(n)}(\xi_i^-)$ $\forall\, i \in (2, \ldots, k-1)$, e $n \in (0, \ldots, r-1)$.

Se $f_0, \ldots, f_{r+k-2}$ è la base spline con le seguenti condizioni $f_0(x) = 1, f_1(0) = \ldots = f_{r+k-2}(0) = 0$, allora la funzione di sconto diventa:

$$\delta(x, \alpha) = 1 + \sum_{j=1}^{r+k-2} \alpha_j f_j(x)$$

dove $\alpha = (\alpha_1, \ldots, \alpha_{r+k-2})$ sono i coefficienti da stimare.

3.3 Le B-Spline

Definizione 3.3.1 La k-esima differenza divisa di una funzione g nei punti $\tau_i, \ldots, \tau_{i+k}$, è il coefficiente principale del polinomio di ordine $k+1$, che è uguale a g nei punti $\tau_i, \ldots, \tau_{i+k}$. Si denota:

$$[\tau_i, \ldots, \tau_{i+k}]g \tag{3.3}$$

Definizione 3.3.2 Sia $t := (t_i)$ una sequenza non decrescente, l'i-esima B-spline di ordine k si indica $B_{i,k,t}$ ed è definita da:

$$B_{i,k,t}(x) = (t_{i+k} - t_i)[t_i, \ldots, t_{i+k}](\cdot - x)_+^{k-1} \quad \forall\, x \in \Re. \tag{3.4}$$

o equivalentemente dalla seguente ricorrenza:

$$B_{i,1,t}(x) = \begin{cases} 1 & \text{se} \quad t_i \leq x < t_{i+1} \\ 0 & \text{altrimenti} \end{cases}$$

$$B_{i,k,t}(x) = \frac{x - t_i}{t_{i+k-1} - t_i} B_{i,k-1,t}(x) + \frac{t_{i+k} - x}{t_{i+k} - t_{i+1}} B_{i+1,k-1,t}(x) \tag{3.5}$$

La k-esima differenza divisa della funzione $(t-x)_+^{k-1}$ delle due variabili t e x, è scelta fissando x e considerando $(t-x)_+^{k-1}$ nella sola variabile t.

Indicando con $B_i = B_{i,k,t}$ le B-spline hanno le seguenti proprieta':

1. $B_i(x) = 0 \qquad \forall\, x \notin [t_i, t_{i+k}]$

2. se $x \notin [t_i, t_{i+k}]$, l'i-esima B-spline è un polinomio di grado $< k$ nell'intervallo $[t_i, t_{i+k}]$

3. $\sum_i B_i(x) = \sum_{i=r+1-k}^{s-1} B_i(x) = 1 \quad \forall\ t_r < t < t_s$
(è la normalizzazione delle B-spline)

4. $B_i(x)$ è positiva nel suo supporto:
$B_i(x) > 0 \qquad$ per $t_i < t < t_{i+k}$

Definiamo ora l'insieme di tali basi spline:

Definizione 3.3.3 Una funzione spline di ordine k con una sequenza di nodi t, è una combinazione lineare di B-spline di ordine k con la medesima sequenza di nodi t. La collezione di tutte queste funzioni si denota con $\$_{k,t}$:

$$\$_{k,t} := \{\sum_i \alpha_i B_{i,k,t} \;\; t.c. \; \alpha_i \;\in \Re, \forall\ i\} \tag{3.6}$$

Il valore della funzione $\sum_i \alpha_i B_i$ nel punto x è semplicemente il valore della somma $\sum_i \alpha_i B_i(x)$. In base alla definizione (3.3.3) abbiamo il seguente teorema (De Boer):

Teorema 3.3.1 *Sia data una sequenza di punti strettamente crescente* $\xi = (\xi_i)_1^{l+1}$ *(detti breakpoint) e una sequenza di interi non negativi* $v = (v_i)_2^l$*, con* $v_i \le k$*,* $\forall\ i$*, allora:*

$$n := k + \sum_{i=2}^{l}(k - v_i) = kl - \sum_{i=2}^{l} v_i = dim\Gamma_{k,\xi,\nu}$$

dove $\Gamma_{k,\xi,v}$ *indica lo spazio lineare dei polinomi continui a tratti di ordine* k*, con una sequenza di punti* ξ*, che soddisfano le condizioni di continuità specificate da* v*. Sia* $t := (t_i)_1^{n+k}$ *una sequenza non decrescente tale che:*

1. $t_1 \le t_2 \le \ldots \le t_k \le \xi_1$ *e* $\xi_{l+1} \le t_{n+1} \le \ldots \le t_{n+k}$

2. per $i = 2, \ldots, l$ *il punto* ξ_i *occorre esattamente* $k - v_i$ *volte in* t

allora la sequenza $B_1, \ldots, B_n$ di B-spline di ordine k per una sequenza di nodi t, è una base di $\Gamma_{k,\xi,\nu}$, considerata come funzione nell'intervallo $[t_k, t_{n+1}]$:

$$\$_{k,t} = \Gamma_{k,\xi,\nu} \quad in\ [t_k, t_{n+1}]$$

Dimostrazione

Dimostriamo prima che B_i appartiene a $\Gamma_{k,\xi,\nu}$ (come una funzione sull'intervallo $[t_k, t_{n+1}]$). Dalla definizione (3.3.2), la B-spline $B_i = B_{i,k,t}$, è sul fattore scalare $(t_{i+k} - t_i)$, la k-esima differenza divisa nei punti $t_i, \ldots, t_{i+k}$ della funzione $(t-x)_+^{k-1}$. Dalle proprietà delle differenza divise, ci sono coefficienti $d_i, \ldots, d_{i+k}$ che dipendono solo dai numeri $t_i, \ldots, t_{i+k}$ cosi' che, per ogni funzione regolare g, si ha:

$$[t_i, \ldots, t_{i+k}]g = \sum_{r=i}^{i+k} d_r g^{j_r}(t_r) \tag{3.7}$$

con

$$j_r = max(s : r - s \geq i \text{ e } t_{r-s} = t_r), \quad r = i, \ldots, i+k.$$

Di conseguenza,

$$B_i = (t_{i+k} - t_i) \sum_{r=i}^{i+k} d_r (t_r - x)_+^{k-1-j}(k-1)!/(k-1-j_r)! \tag{3.8}$$

Questo mostra che B_i è un polinomio continuo a tratti di ordine k con punti di raccordo nei punti $t_i, \ldots, t_{i+k}$. Rimane da contare il numero di derivate continue di B_i su ciascuno dei punti di raccordo. Per questo notiamo che B_i non può avere un salto nella sua j-esima derivata nei punti di raccordo ξ_j, a meno che, per qualche $r \in [i, i+k]$, si ha $\xi_j = t_r$ e $k-1-j_r = s$. Poichè j_r conta il numero di t_m uguale a t_r e con $i \leq m < r$, ne segue che j_r deve essere minore di $k-\nu_j$ che è il numero totale di t_m uguale a ξ_j e quindi uguale a t_r, per la costruzione di t. Questo vuol dire che $s \geq \nu_j$ e cosi' si ha:

$$\text{jump}_{\xi_j} D^m B_i = 0 \quad \text{per} \quad m = 0, \ldots, \nu_j - 1$$

.

Abbiamo dimostrato che $B_i \in \Gamma_{k,\xi,\nu}$, per ogni i. Poichè ci sono n B_i e la dimensione di $\Gamma_{k,\xi,\nu}$ è anche n, dobbiamo dimostrare che la sequenza $(B_i)_1^n$ è linearmente indipendente per finire la dimostrazione del teorema. L'indipendenaza lineare di $(B_i)_1^n$, segue dal lemma successivo.

□

Lemma 3.3.1 *Sia λ_i un funzionale lineare definito come:*

$$\lambda_i f = \sum_{r=0}^{k-1} (-1)^{k-1-r} \phi^{(k-1-r)}(\tau_i) D^r f(\tau_i), \quad \forall\ f$$

con $\phi(t) = (t_{i+1} - t) \ldots (t_{i+k-1} - t)/(k-1)!$, con τ_i un arbitrario punto nell'intervallo aperto $]t_i, t_{i+k}[$. Allora

$$\lambda_i B_j = \delta_{ij}, \qquad \forall\ j.$$

Dimostrazione:

È chiaro che $\lambda_i B_j$ è un polinomio continuo a tratti come funzione di τ_i, con possibili salti in t_r. È perciò sufficiente dimostrare che $\tau_i \neq t_r$, per ogni r. Dalla (3.7) e dalla (3.8), si ha:

$$\lambda_i B_j = (t_{j+k} - t_j) \sum_{r=j}^{j+k} d_r \lambda_i \left(D_s^{j_r} (s - \cdot)_+^{k-1} \right)$$

con $D_s^{j_r}$ indica la j_r-esima derivata rispetto a s. Calcoliamo $\lambda_i (s - \cdot)_+^{k-1}$. Per $s < \tau_i$, si ha $\lambda_i f = 0$. Per $s > \tau_i$, f è un polinomio uguale a $(s - \cdot)^{k-1}$ nelle vicinanze di τ_i, mentre

$$\lambda_i (s - \cdot)^{k-1} = \sum_{r=0}^{k-1} (-)^{k-1-r} \phi^{(k-1-r)}(\tau_i)(k-1) \cdot \ldots \cdot (k-r)(-)^r (s - \tau_i)^{k-1-r}$$

$$= (-)^{k-1}(k-1)! \sum_{r=0}^{k-1} \left[\phi^{(k-1-r)}(\tau_i)/(k-1-r)! \right] (s - \tau_i)^{k-1-r}$$

Ma l'ultima somma è una serie di Taylor troncata in s di ordine k e poichè ϕ è un polinomio di ordine k, questa somma deve coincidere con ϕ in s. Quindi $\lambda_i (s - \cdot)^{k-1} = (-)^{k-1}(k-1)! \phi_s$.

In definitiva si ha:

$$\lambda_i(s-\cdot)_+^{k-1} = (-)^{k-1}(k-1)!\phi(s)(s-\tau_i)_+^0,$$

e poichè $D_s^r\lambda_i(s-\cdot)_+^{k-1} = \lambda_i(D_s^r(s-\cdot)_+^{k-1})$, concludiamo che:

$$\lambda_i B_j = (t_{j+k}-t_j)(-)^{k-1}(k-1)![t_j,\dots,t_{j+k}]\Phi_i$$

con

$$\Phi_i(s) = \phi(s)(s-\tau_i)_+^0.$$

Calcoliamo ora la k-esima differenza $[t_j,\dots,t_{j+k}]\Phi_i$ di Φ_i dalla definizione (3.3.1) con il coefficiente principale del polinomio di ordine $k+1$ che coincide con Φ_i nei punti $t_j,\dots,t_{j+k}$. Assumendo che $t_i<\tau_i<t_{i+k}$, si ha:

1. Φ_i coincide con ϕ in $t_{i+1}, t_{i+2},\dots$, come un polinomio di ordine $k+1$, ϕ ha coefficiente principale 0, poichè

$$[t_j,\dots,t_{j+k}]\phi_i = 0 \quad \text{per } j = i+1, i+2, \dots;$$

2. ϕ_i è 0 nei punti $t_{i+k-1}, t_{i+k-2},\dots$, poichè

$$[t_j,\dots,t_{j+k}]\phi_i = 0 \quad \text{per } j = i-1, i-2, \dots;$$

3. ϕ_i coincide con il polinomio di ordine $(k+1)$:

$$p(x) = \phi(x)(x-t_i)/(t_{i+k}-t_i),$$

nei punti $t_i,\dots,t_{i+k}$, poichè

$$[t_i,\dots,t_{i+k}]\phi_i = (-)^{k-1}/((k-1)!(t_{i+k}-t_i)).$$

Ma questo mostra che $\lambda_i B_j = \delta_{ij}, \qquad \forall\ j.$
□

Le condizioni di non continuita' di una B-spline, si esprimono dunque attraverso la ripetitivita' di un nodo. Mettendo all'interno dell'intervallo dei singoli nodi, si richiede la continuita' della spline e di tutte le sue derivate. Il teorema (3.3.1) ci permette di costruire una base B-spline per ogni particolare funzione spline nello spazio $\Gamma_{k,\xi,v}$. C'é una stretta relazione tra l'ordine k della spline e le condizioni di continuita' relativi ai punti ξ, in particolare:

$$\text{il numero di condizioni di continuita' in } \xi + \text{il numero di nodi in } \xi = k$$

Il teorema consente di scegliere liberamente i primi e gli ultimi k nodi, ma una scelta conveniente è la seguente:

$$t_1 = t_2 = \ldots = t_k = \xi_1 \qquad t_{n+1} = \cdots = t_{n+k} = \xi_{l+1}$$

che ci permette di includere anche la seguente condizione:

$$v_1 = v_{l+1} = 0$$

In questo modo nei punti estremi dell'intervallo ξ_1, ξ_{l+1}, non abbiamo condizioni di continuita', cosa che invece è verificata all'interno dell'intervallo stesso.

Diamo ora la definizione di una B-rappresentazione:

Definizione 3.3.4 La B-rappresentazione per una funzione $f \ \in \Gamma_{k,\xi,v}$ consiste in:

1. gli interi k e n, che rappresentano rispettivamente l'ordine di f e il numero di parametri lineari ($n = kl - \sum_i v_i$)
2. il vettore $t = (t_i)_1^{n+k}$ contenente la sequenza di nodi in ordine crescente (costruiti con ξ e v come nel teorema)
3. il vettore $\alpha = (\alpha_i)_1^n$ dei coefficienti di f rispetto alla base B-spline $(B_i)_1^n$ su una sequenza di nodi t.

In particolarte il valore di f su x nell'intervallo $[t_k, t_{n+1}]$ è dato da:

$$f(x) = \sum_{i=1}^{n} \alpha_i B_i(x)$$

Un'importante proprieta' delle funzione spline espresse come combinazione lineare delle B-spline, è la possibilita' di esprimere la sua derivata sempre come opportuna combinazione lineare di B-spline. Si puó infatti dimostrare che ([8]):

$$D(\sum_i \beta_i B_{i,k+1}) = \sum_i k \frac{\beta_i - \beta_{i-1}}{t_{i+k} - t_i} B_{i,k}$$

dove D indica l'operatore di derivazione.

3.4 Metodi adattativi

Abbiamo visto nei paragrafi precedenti come l'uso di funzioni spline per stimare la funzione di sconto $d(t)$, e corrispondentemente la struttura a termine dei tassi di interesse, richieda la soluzione del sistema lineare di equazioni (2.5). Nei paragrafi che seguono, dopo aver richiamato brevemente uno dei principali metodi utilizzati per risolvere tale problema, il metodo dei minimi quadrati, descriveremo una classe di algoritmi per la scelta automatica del numero e della posizione dei nodi dell'approssimazione spline. Questi algoritmi sono basati sull'introduzione di opportuni criteri di scelta per il modello di regressione.

3.4.1 Metodo dei minimi quadrati

Supponiamo di avere N osservazioni y_i, con $i = 1, \ldots, N$, di una variabile Y, e p variabili indipendeti $X_1, \ldots, X_p$. Sia x_{in} il valore della variabile relativa all' i-esima variabile indipendente. La relazione che lega le variabili dipendenti Y alle variabili indipendenti X_i, nel modello lineare è espresso mediante il seguente sistema:

$$\begin{aligned}
y_1 =& \quad \beta_0 + \beta_1 x_{11} + \ldots + \beta_p x_{1p} + \varepsilon_1 \\
y_2 =& \quad \beta_0 + \beta_1 x_{21} + \ldots + \beta_p x_{2p} + \varepsilon_2 \\
\vdots =& \quad \vdots \\
y_n =& \quad \beta_0 + \beta_1 x_{n1} + \ldots + \beta_p x_{np} + \varepsilon_n
\end{aligned}$$

che puo' essere scritto in forma di matrice:

$$\underline{y} = X\underline{\beta} + \underline{\varepsilon} \tag{3.9}$$

dove:

$$\underline{\mathbf{y}} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \qquad \mathbf{X} = \begin{pmatrix} 1 & x_{11} & x_{12} & \dots & x_{1p} \\ 1 & x_{21} & x_{22} & \dots & x_{2p} \\ \vdots & \vdots & \vdots & \vdots & \\ 1 & x_{n1} & x_{n2} & \dots & x_{np} \end{pmatrix} \qquad \underline{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{pmatrix} \quad \underline{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

La matrice X è detta matrice di regressione, $\underline{\varepsilon}$ è il termine che si assume essere di media nulla e di componenti indipendenti ed ugualmente distribuite (*omoschedastiche*), e $\beta_0, \beta_1, \dots, \beta_p$, sono i coefficienti da stimare. Il metodo dei minimi quadrati puó essere utilizzato per ottenere tali stime.

Se definiamo con e_n, l'errore n-esimo:

$$e_n = y_n - \beta_1 x_{1n} - \dots - \beta_p x_{pn}, \tag{3.10}$$

la somma quadrata degli errori è:

$$\underline{\xi}^2 = e_1^2 + \dots + e_n^2 = \sum_{i=1}^{n} \left(y_i - \sum_{j=1}^{p} \beta_j x_{ij} \right)^2 = (\underline{y} - X\underline{\beta})^t (\underline{y} - X\underline{\beta}) \tag{3.11}$$

Il metodo dei minimi quadrati, consiste nello scegliere i coefficienti $\beta_0, \dots, \beta_k$ che minimizzano $\underline{\xi}^2$. Differenziando $\underline{\xi}^2$ rispetto a β_k si ha:

$$\frac{\partial \underline{\xi}^2}{\partial \beta_k} = -2 \sum_{i=1}^{n} x_{ik} \left(y_i - \sum_{j=1}^{p} \beta_j x_{ij} \right) \quad k = 1, \dots, p \tag{3.12}$$

Il valore $\hat{\beta}_k$ che rappresenta lo stimatore di β_k, che minimizza $\underline{\xi}^2$ è ottenuto da:

$$\sum_{i=1}^{n} x_{ik}\left(y_i - \sum_{j=1}^{p} \hat{\beta}_j x_{ij}\right) = 0 \quad k = 1, \ldots, p \tag{3.13}$$

riscrivendo la (3.13) sotto forma di matrice si ha:

$$(X^t X)\underline{\hat{\beta}} = X^t \underline{y} \tag{3.14}$$

L'equazione in $\underline{\beta}$

$$(X^t X)\underline{\beta} = X^t \underline{y} \tag{3.15}$$

è chiamata l'equazione normale per il modello lineare generale, ed è un sistema non omogeneo di equazioni lineari. Lo stimatore dei minimi quadrati $\underline{\hat{\beta}}$ di $\underline{\beta}$ dato da:

$$\underline{\hat{\beta}} = (X^t X)^{-1} X^t \underline{y} \tag{3.16}$$

esiste, se esiste $(X^t X)^{-1}$. Quindi la matrice $X^t X$ deve essere invertibile. Affinchè ciò avvenga, la matrice $(X^t X)$ deve avere rango pieno $p+1$. Una condizione necessaria e sufficiente, è che le colonne di X siano linearmente indipendenti, o equivalentemente che nessuna colonna di X sia combinazione lineare delle rimanenti colonne. Se questo avviene, la matrice X si dice multicollineare. Mostriamo che l'equazione (3.15) ha sempre una soluzione unica. Sia $\underline{\hat{\beta}}$ una soluzione di (3.15), allora abbiamo:

$$\begin{aligned}
&(\underline{y} - X\underline{\beta})^t(\underline{y} - X\underline{\beta}) \\
&= [\underline{y} - X\underline{\hat{\beta}} + X(\underline{\hat{\beta}} - \underline{\beta})]^t[\underline{y} - X\underline{\hat{\beta}} + X(\underline{\hat{\beta}} - \underline{\beta})] \\
&= (\underline{y} - X\underline{\hat{\beta}})^t(\underline{y} - X\underline{\hat{\beta}}) + (\underline{\hat{\beta}} - \underline{\beta})^t X^t X(\underline{\hat{\beta}} - \underline{\beta}) + 2(\underline{y} - X\underline{\hat{\beta}})^t X(\underline{\hat{\beta}} - \underline{\beta}) \\
&= (\underline{y} - X\underline{\hat{\beta}})^t(\underline{y} - X\underline{\hat{\beta}}) + (\underline{\hat{\beta}} - \underline{\beta})^t X^t X(\underline{\hat{\beta}} - \underline{\beta}) \\
&= (\underline{y} - X\underline{\hat{\beta}})^t(\underline{y} - X\underline{\hat{\beta}}) + [X(\underline{\hat{\beta}} - \underline{\beta})]^t[X(\underline{\hat{\beta}} - \underline{\beta})]
\end{aligned}$$

poichè

$$\begin{aligned}(\underline{y}-X\underline{\hat{\beta}})^tX(\underline{\hat{\beta}}-\underline{\beta}) &= (\underline{y}^tX-\underline{\hat{\beta}}^tX^tX)(\underline{\hat{\beta}}-\underline{\beta})\\ &= (\underline{y}^tX-\underline{y}^tX)(\underline{\hat{\beta}}-\underline{\beta})\\ &= 0\end{aligned}$$

Inoltre:

$$[X(\underline{\hat{\beta}}-\underline{\beta})]^t[X(\underline{\hat{\beta}}-\underline{\beta})] \geq 0$$

da cui segue che

$$(\underline{y}-X\underline{\hat{\beta}})^t(\underline{y}-X\underline{\hat{\beta}}) \leq (\underline{y}-X\underline{\beta})^t(\underline{y}-X\underline{\beta}) \ \ \forall\ \underline{\beta}$$

Diamo ora la seguente definizione:

Definizione 3.4.1 Uno stimatore della funzione parametrica lineare $L^t\beta$, dove $L = (l_1, \dots, l_p)^t$ è un vettore reale, è definito come $L^t\hat{\beta}$, dove $\hat{\beta}$ è la soluzione dell'equazione (3.15). Se X è di rango pieno, allora $L^t\hat{\beta}$ è unico.

Mostriamo nel teorema che segue, che nel caso del modello lineare di rango pieno, lo stimatore dei minimi quadrati di qualunque funzione lineare parametrica $L^t\hat{\beta}$, è il migliore stimatore lineare non distorto della classe di tutti i stimatori lineari non distorti nel senso della minima varianza. Come caso speciale di questo teorema dimostreremo il teorema di Gauss-Markov, che afferma che lo stimatore dei minimi quadrati $\hat{\beta}_i$ di β_i, è il miglior stimatore lineare non distorto tra tutti gli stimatori lineari non distorti. Se X non ha rango pieno, il teorema vale ancora, purche' $L^t\hat{\beta}$ sia unico per tutte le soluzioni di $X^tX\beta = X^tY$.

Teorema 3.4.1 *Per il modello lineare generale di rango pieno, lo stimatore dei minimi quadrati $L^t\hat{\beta}$ della funzione lineare parametrica $L^t\beta$, dove $\hat{\beta}$ è l'unica soluzione di $X^tX\beta = X^tY$, è il migliore stimatore lineare non distorto.*

Dimostrazione:

$$L^t\underline{\hat{\beta}} = L^t(X^tX)^{-1}X^tY$$

$$E(L^t\underline{\hat{\beta}}) = L^t(X^tX)^{-1}X^tX\underline{\beta} = L^t\underline{\beta}$$

. dove $E(X)$ indica la media della variabile aleatoria X. Ovviamente $L^t\underline{\hat{\beta}}$ è uno stimatore lineare non distorto di $L^t\underline{\beta}$. Sia K^tY, con $K=(k_1,\ldots,k_n)^t$ un altro stimatore lineare non distorto di $L^t\underline{\beta}$. Poiche' K^tY è non distorto si ha:

$$E(K^tY)=K^tX\underline{\beta}=L^t\underline{\beta}.$$

Quindi $K^tX=L^t$. Ora vediamo la varianza di K^tY:

$$\begin{aligned} Var(K^tY) &= Var(K^tY-L^t\underline{\hat{\beta}}+L^t\underline{\hat{\beta}}) \\ &= Var(K^tY-L^t\underline{\hat{\beta}})+Var(L^t\underline{\hat{\beta}})+2Cov(K^tY-L^t\underline{\hat{\beta}},L^t\underline{\hat{\beta}}) \\ &= Var(K^tY-L^t\underline{\hat{\beta}})+Var(L^t\underline{\hat{\beta}}) \\ &\geq Var(L^t\underline{\hat{\beta}}) \end{aligned}$$

poiche'

$$Cov(K^tY-L^t\underline{\hat{\beta}},L^t\underline{\hat{\beta}})=$$

$$\begin{aligned} &= Cov(K^t(Y-X\underline{\hat{\beta}}),K^tX\underline{\hat{\beta}}) \\ &= Cov(K^t[I-X(X^tX)^{-1}X^t]Y,K^tX(X^tX)^{-1}X^tY) \\ &= K^t[I-X(X^tX)^{-1}X^t]E((Y-X\underline{\beta})(Y-X\underline{\beta})^t)X(X^tX)^{-1}X^tK \\ &= \sigma^2K^t[I-X(X^tX)^{-1}X^t]X(X^tX)^{-1}X^tK \\ &= \sigma^2K^t(X-X)(X^tX)^{-1}X^tK \\ &= 0. \end{aligned}$$

Quindi $L^t\underline{\hat{\beta}}$ ha varianza minima nella classe dei stimatori lineari non distorti di L^t.

□

Vediamo ora il teorema di Gauss-Markov:

Teorema 3.4.2 *Lo stimatore dei minimi quadrati* $\underline{\hat{\beta}}_i$ *di* $\underline{\beta}_i$*, per* $i=1,\ldots,k$*, è il migliore stimatore lineare non distorto.*

Dimostrazione

Prendiamo L un vettore con l'i-esima componente uguale a 1 e tutte le altre componenti uguali a zero. Allora $L^t\underline{\beta}=\underline{\beta}_i$. La dimostrazione ora segue dal teorema precedente.

□

Nel caso che stiamo considerando, le componenti ε_i dell'errore non sono supposte essere identicamente distribuite. Infatti la varianza è data da $\sigma^2_{\varepsilon_i}$=$\sigma^2(\triangle p_i)^2 > 0$, dove $\triangle p_i$ è lo scarto bid-ask (*eteroschedasticita'*). Tale problema può essere facilmente risolto (cite Jobson) in questo caso. Infatti il sistema di equazioni (2.5) puó essere scritto in forma compatta:

$$y = X\underline{\beta} + C\underline{\tilde{\varepsilon}} \tag{3.17}$$

dove C è una matrice diagonale tale che:

$$c_{i,j} = \begin{cases} 0 & i \neq j \\ \triangle p_i & i = j, \end{cases}$$

e $\underline{\tilde{\varepsilon}}$ ha componenti indipendenti e identicamente distribuite di stessa varianza σ^2. Moltiplicando per C^{-1} si ha:

$$C^{-1}y = C^{-1}X\underline{\beta} + \underline{\tilde{\varepsilon}} \tag{3.18}$$

che corrisponde ad un modello lineare omoschedastico. L'applicazione del metodo dei minimi quadrati, comporta in questo caso alla minimizzazione di:

$$\underline{\xi}^2 = (C^{-1}\underline{y} - C^{-1}X\underline{\beta})^t(C^{-1}\underline{y} - C^{-1}X\underline{\beta}) = (\underline{y} - X\underline{\beta})^t(C^{-1})^tC^{-1}(\underline{y} - X\underline{\beta}) \tag{3.19}$$

Definendo con $W = (C^{-1})^tC^{-1}$ otteniamo che la funzione da minimizzare è

$$(\underline{y} - X\underline{\beta})^tW(\underline{y} - X\underline{\beta})$$

che corrisponde ad un metodo dei minimi quadrati pesati.([12])

3.4.2 L'Algoritmo Adottativo

Applicando il metodo dei minimi quadrati al sistema (2.5), possiamo ottenere una stima dei coefficienti della funzione spline approssimante. Tuttavia, la matrice di regressione X, è strettamente dipendente dal numero e dalla locazione dei nodi: infatti diverse configurazioni di nodi, modificano le funzioni spline $f_j(t)$, cambiando cosi' la matrice di regressione. La scelta di una configurazione di nodi corrisponde dunque ad un problema di scelta del modello lineare. Esistono una varietà di modi differenti per affrontare il problema di selezione di un modello ([12]) che in generale prevede anche la scelta delle proprietà di continuità della base spline. Nel nostro caso tuttavia, ci limiteremo a considerare solamente il problema relativo alla scelta dei nodi, poichè viene generalmente assunto in letteratura ([2]) che la funzione di sconto sia derivabile fino ad un ordine opportuno. Dunque le spline che abbiamo scelto per le prove numeriche sono continue insieme a tutte le derivate nei punti di raccordo. Se r è l'ordine della spline e k è il numero di punti di raccordo, il numero di parametri da stimare è $r + k - 2$.

Descriviamo nel seguito un algoritmo che produce un insieme di configurazioni di nodi. Supponendo di avere a disposizione un certo numero di obbligazioni con le rispettive cedole e maturita', vediamo come funziona il nostro algoritmo:

1. Partiamo da una configurazione iniziale di k-nodi, predisposti in modo arbitrario, nell'intervallo $[0, 30]$, ottenendo cosi' $k-1$ sottointervalli, in cui cadono le maturita' delle obbligazioni osservate.

2. Aggiungiamo nodi fino ad un massimo di h nodi prefissati, iterando il seguente procedimento:

   (a) Data una configurazione di nodi, per $j = 1, \ldots, \#\text{intervalli}$, aggiungiamo un nodo all'interno del j-esimo sottointervallo, precisamente nella mediana delle maturita', in modo tale che sia a destra che a sinistra, ci siano lo stesso numero di maturita',

   (b) calcoliamo la matrice di regressione, lo stimatore dei minimi quadrati e il residuo $\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$,

   (c) selezioniamo la configurazione di nodi che minimizza il residuo. Avendo aumentato di uno il # degli intervalli, torniamo al passo (2a) fino a quando il numero di nodi aggiunti è uguale ad h.

3. A questo punto abbiamo $h + 1$ configurazioni di nodi, considerando anche quella iniziale. Cominciamo a togliere l nodi a partire dall'ultima configurazione con $k + h$ nodi e $k + h - 1$ intervalli, nel seguente modo:
   (a) Data una configurazione di nodi, per $j = 1, \dots, \#\text{nodi}$:
   (b) togliamo il j-esimo nodo,calcoliamo la matrice di regressione, lo stimatore di dei minimi quadrati e il residuo $\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$,
   (c) selezioniamo la configurazione di nodi che minimizza il residuo, diminuiamo di uno il # dei nodi, e torniamo al passo (3a) fino a quando il numero di nodi tolti è uguale ad l.

Al termine della procedura avro' $h + l + 1$ configurazioni di nodi, i corrispondenti stimatori e loro statistiche. I criteri utilizzati per la scelta delle migliori stime sono ottenuti prendendo in considerazione delle quantità che tipicamente dipendono dalle stime ottenute e dal numero di parametri utilizzati e forniscono come soluzione quel modello che minimizza tale quantità. I criteri che descriveremo nel seguito, sono dei criteri classici nella letteratura statistica della scelta dei modelli. Come descritto nel paragrafo precedente, con il metodo dei minimi quadrati, otteniamo:

$$\hat{y} = X\hat{\beta} = X(X^tX)^{-1}X^ty \tag{3.20}$$

dove $\hat{y}$ sono le quantita' stimate e y sono le quantita' note. Definiamo ora:

1. SSE somma quadratica dei residui (sum of squared residuals)

$$\begin{aligned} SSE = \; & \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = (y - \hat{y})^t(y - \hat{y}) = \\ & (y - X\beta)^t(y - X\beta) = (n - p - 1)s^2 \end{aligned} \tag{3.21}$$

   dove:

   n = numero delle osservazioni
   p = numero parametri, nelle spline r+k-2
   r = grado delle basi spline
   k = numero di nodi

2. SST somma totale quadratica (*total sum of squared*)

$$SST = \sum_{i=1}^{n} (y_i - \overline{y})^2 = y^t y - n\overline{y}^2 \tag{3.22}$$

dove $\overline{y} = \sum_{i=1}^{n} \frac{y_i}{n}$.

3. SSR somma quadratica della regressione (*regression of squares*)

$$SSR = \sum_{i=1}^{n} (\hat{y}_i - \overline{y})^2 = \hat{y}^t \hat{y} - n\overline{y}^2 \tag{3.23}$$

Si definisce allora statistica R^2, la quantità:

$$R^2 = \frac{\sum_{i=1}^{n} (\hat{y}_i - \overline{y})^2}{\sum_{i=1}^{n} (y_i - \overline{y})^2} = \frac{SSR}{SST} = \frac{SST + SSE}{SST} \tag{3.24}$$

che rappresenta la proporzione della variazione totale presente nelle variabili osservate spiegata dal modello di regressione. È chiaro che aumentando il numero p di variabili, il modello di regressione tende ad adattarsi ai dati osservati sempre meglio, fino ad ottenere un adattamento perfetto nel caso in cui $p = n^{-1}$. Per ottenere una migliore misura dell'adeguatezza del modello utilizzato, si può introdurre una funzione che penalizza il numero di variabili esplicative presenti. Ciò permette di introdurre la seguente statistica:

R^2 modificato.

$$R^{'2}(p) = 1 - \frac{n-1}{n-p-1}(1 - R^2) = 1 - \frac{SSE/(n-p-1)}{SST/(n-1)} \tag{3.25}$$

Una statistica utilizzata con successo nell'ambito dell'applicazione delle spline alla regressione, è la seguente:

GCV. La statistica *GCV* (*Generalized Cross Validation*), è stata introdotta da Friedman ([10]) ed è dato da:

$$GCV = \frac{SSE}{n} / \left[1 - \frac{a(J-1)}{n}\right]^2 \tag{3.26}$$

dove J è il numero di componenti delle basi spline. Sostituendo la (3.21) e la (**??**) nella (3.26), si ottiene:

$$\begin{aligned} GCV &= \frac{(n-p-1)s^2}{n} \frac{n^2}{(n-a(J-1))^2} \\ &= \frac{n(\underline{y} - X\underline{b})^t(\underline{y} - X\underline{b})}{[n-a(J-1)]^2} \end{aligned}$$

Il valore di a suggerito in ([21]) è $a = 2.5$.

Concludiamo questa sezione introducendo due classici metodi per la scelta di un modello, il Criterio di Informazione di Akaike -AIC ([1]) ed il Criterio di Informazione Bayesiano -BIC ([18]). Tali criteri sono molto utilizzati in differenti contesti della statistica. Tuttavia, rispetto ai criteri precedenti, R^2 modificato e GCV, richiedono di specificare la distribuzione del rumore. Nel nostro contesto, considerando il modello trasformato..., assumeremo per $\tilde{\xi}$ una distribuzione gaussiana di media nulla e matrice di varianza $\sum = \sigma I$, dove I è la matrice identità. La log-verosimiglianza della osservazione $\underline{\tilde{y}}$ è dunque data da:

$$l(a) = \log f(\underline{\tilde{y}}, \underline{a}) = -\frac{\|\tilde{y} - \tilde{X}\underline{a}\|^2}{2\sigma^2} - \frac{n}{2}\log(\sigma^2) - \frac{n}{2}\log(2\pi) \tag{3.27}$$

dove $\underline{\tilde{y}} = C^{-1}\underline{y}$ e $\tilde{X} = C^{-1}X$ come specificato nella sezione ....

AIC e BIC. I criteri *AIC* e *BIC* differiscono dalla scelta del valore c, nella seguente formula:

$$CRITERIO(p) = -2l(\hat{a_p}) + cp$$

dove $\hat{a_p}$ è la stima di massima verosimiglianza relativa al modello con p parametri.

$$c = \begin{cases} lg(n) & \text{BIC} \\ 2 & \text{AIC} \end{cases} \tag{3.28}$$

Poichè il termine $\frac{n}{2}\log(2\pi)$ non dipende da p, la funzione $CRITERIO$ diventa:

$$CRITERIO(p) = \frac{\|y - X\hat{a}\|^2}{2\sigma^2} + \frac{n}{2}\log(\sigma^2) + cp \tag{3.29}$$

Quando σ^2 non è noto, come nel nostro caso, si può utilizzare nella formula precedente una sua stima $\hat{\sigma^2}$, che come abbiamo visto è

$$\hat{\sigma^2} = \frac{\|y - X\hat{a}\|^2}{n - p - 1} \tag{3.30}$$

Sostituendo la (3.30) nella (3.29) si ottiene:

$$\begin{aligned} CRITERIO(p) &= \frac{\|\tilde{y} - \tilde{X}\hat{a}_p\|^2}{2\hat{\sigma}^2} + \frac{n}{2}\log(\hat{\sigma}^2) + cp \\ &= \frac{n-p-1}{2} + \frac{n}{2}\log\frac{\|\tilde{y} - \tilde{X}\hat{a}_p\|^2}{n-p-1} + cp \end{aligned}$$

Capitolo 4

Prove numeriche

In questo capitolo illustreremo un insieme di risultati numerici ottenuti applicando i metodi descritti nei capitoli precedenti. Al fine di verificare le caratteristiche degli algoritmi utilizzati, abbiamo utilizzato un insieme di prezzi "sintetici", generati dalla (2.4) a partire da una funzione di sconto assegnata a priori. In particolare partendo da un insieme di obbligazioni $B_i, i = 1, \ldots, N$ emesse dal tesoro americano (*US Treasury issues*) con rispettive cedole di pagamento C_i e maturitá t_{M_i}, (riportate nell'appendice), abbiamo calcolato in giorni il periodo di tempo che intecorre tra il giorno della quotazione (14/09/2000) e il giorno di scadenza di ogni singola obbligazione, per ottenere cosi' gli istanti di pagamento delle rispettive cedole $T_k^{(i)}$, per $i = 1, \ldots, N$ e $k = 1, \ldots, N_i$. Abbiamo quindi generato i prezzi mediante la (2.4) utilizzando come funzione di sconto il modello (3.12) proposto da Nelson e Siegel. Dando dei differenti valori ai parametri $\alpha = (\beta_0, \beta_1, \beta_2, \tau)$, abbiamo ottenuto diverse tipologie della funzione di sconto, del tasso spot e del tasso forward. Nelle figure (4.1), (4.2) e (4.3), sono riportate le curve utilizzate corrispondenti ai parametri $\alpha_1 = (\beta_0 = 0.06, \beta_1 = 0, \beta_2 = 0, \tau = 1))$, $\alpha_2 = (\beta_0 = 0.08, \beta_1 = 0.12, \beta_2 = -0.15, \tau = 1))$ e $\alpha_3 = (\beta_0 = 0.08, \beta_1 = 0.12, \beta_2 = 0.6, \tau = 1))$.

In un primo insieme di esperimenti abbiamo stimato tali curve con le basi spline (3.2) e con le basi B-spline, di grado $r = 3, 4, 5$, usando due diverse configurazioni di nodi: quella suggerita da McCulloch, che consiste in 14 nodi posizionati in $0, 0.2, 0.4, 0.6, 0.9, 1.4, 1.8, 2.4, 3.4, 4.3, 6.3, 16.6, 22.8, 29.7$, e quella classica ([2]) con sette nodi in $0, 1, 3, 5, 7, 11, 30$.

Due insiemi di prezzi sono stati utilizzati per le prove numeriche: nel primo insieme non è stato aggiunto alcun rumore $\sigma_{\varepsilon_i} = 0$ mentre nel secondo è stato aggiunto un rumore gaussiano eteroschedastico di media nulla e varianza $\sigma_{\varepsilon_i} = \sigma \Delta p_i$, dove Δp_i è lo scarto bid-ask vero relativo all'obbligazione B_i e $\sigma = 1$. Poichè la stima delle funzioni di sconto risulta essere sempre soddisfacente, riporteremo solo le stime dei tassi spot e dei tassi forward. Lo stimatore dei coefficienti della base spline, è stato ottenuto con il metodo dei minimi quadrati pesati, come descritto nel terzo capitolo, utilizzando la funzione Matlab Regress. Nelle figure la curva originale è riportata con tratto continuo, mentre le stime ottenute utilizzando la base (3.2) sono punteggiate e quelle con le basi B-spline sono tratteggiate.

Il posizionamento dei nodi è indicato dal simbolo $*$ sull'asse delle ascisse.

Figura 4.1: Funzione di sconto di Nelson-Siegel con rispettivi spot forward

Figura 4.2: Funzione di sconto di Nelson-Siegel con rispettivi spot forward

Figura 4.3: Funzione di sconto di Nelson-Siegel con rispettivi spot e forward

Figura 4.4: Spot e Forward con $r = 3$ senza rumore

Figura 4.5: Spot e Forward con $r = 3$ senza rumore

Figura 4.6: Spot e Forward con $r = 4$ senza rumore

Figura 4.7: Spot e Forward con $r = 4$ senza rumore

Figura 4.8: Spot e Forward con $r = 5$ senza rumore

Figura 4.9: Spot e Forward con $r = 5$ senza rumore

Figura 4.10: Spot e Forward con $r = 3$ con rumore

Figura 4.11: Spot e Forward con $r = 3$ con rumore

Figura 4.12: Spot e Forward con $r = 4$ con rumore

Figura 4.13: Spot e Forward con $r = 4$ con rumore

Figura 4.14: Spot e Forward con $r = 5$ con rumore

Figura 4.15: Spot e Forward con $r = 5$ con rumore

configurazione nodi	spline usate	σ^2
McCulloch	B-spline	10.6565
McCulloch	spline 3.2	10.6428
McCulloch	B-spline	10.6233
McCulloch	spline 3.2	10.6246
McCulloch	B-spline	10.6754
McCulloch	spline 3.2	10.6560
classica	B-spline	10.5008
classica	spline 3.2	10.4714
classica	B-spline	10.5162
classica	spline 3.2	10.4708
classica	B-spline	10.5023
classica	spline 3.2	10.4903

Tabella 4.1: Risultati della prima stima

configurazione nodi	spline usate	σ^2
McCulloch	B-spline	10.6565
McCulloch	spline 3.2	10.6428
McCulloch	B-spline	10.6233
McCulloch	spline 3.2	10.6246
McCulloch	B-spline	10.6754
McCulloch	spline 3.2	10.6560
classica	B-spline	10.5008
classica	spline 3.2	10.4714
classica	B-spline	10.5162
classica	spline 3.2	10.4708
classica	B-spline	10.5023
classica	spline 3.2	10.4903

Tabella 4.2: Risultati della seconda stima

configurazione nodi	spline usate	σ^2
McCulloch	B-spline	10.6585
McCulloch	spline 3.2	10.6372
McCulloch	B-spline	10.6323
McCulloch	spline 3.2	10.6267
McCulloch	B-spline	10.6697
McCulloch	spline 3.2	10.6562
classica	B-spline	10.5685
classica	spline 3.2	10.4849
classica	B-spline	10.6687
classica	spline 3.2	10.5838
classica	B-spline	10.5158
classica	spline 3.2	10.5051

Tabella 4.3: Risultati della terza stima

Bibliografia

[1] 0 Akaike H. - Information Theory and the Extension of the Maximum Likelihood Principle, Proc. 2nd. Int. Symp. Information Theory (eds. B.N. Petrioc and F. Csaki), Budapest, Akademiai Kiado, 1973.

[2] N. Anderson, F. Breedon, M. Deacon, A. Derry and G. Murphy, *Estimating and Interpreting the Yield Curve*, John Wiley and Sons, Series in Financial Economics and Quantitative Analysis, Chichester, New York, Brisbane, Toronto and Singapore (1997).

[3] J. Y. Campbell, A. W. Lo and A. C. MacKinlay, *The econometrics of financial markets*, Princeton University Press, Princeton, New Jersey (1997).

[4] W. T. Carleton and I. A. Cooper, *Estimation and uses of the term structure of interest rates*, Journal of Finance **31** (1976) no.4 1067–1083.

[5] D. R. Chambers, W. T. Carleton and D. W. Waldman, *A new approach to estimation of the term structure of interest rates*, Journal of Financial and Quantitative Analysis **19** (1984) no.3 233–252.

[6] T. S. Coleman, L. Fisher and R. G. Ibbotson, *Estimating the term structu re of interest rates from data that include the prices of coupon bonds*, Journal of Fixed Income (1992) 85–116.

[7] J. C. Cox, J. E. Ingersoll and S. A. Ross, *A theory of the term structure of interest rates*, Econometrica **53** (1985) 385–408.

[8] C. De Boor, A practical guide to Splines, Springer-Verlag, New York, (1978).

[9] L. Marangio, A. Ramponi, M. Bernaschi, A critical of techniques for term structure analysis, I.A.C., Rome, (1999).

[10] J.H. Friedman, *Multivariate adaptive regression splines*, The Annals of Statistics **19**, no.1 (1991) 1–141.

[11] A. Kirsch, *An introduction to the mathematical theory of inverse problems*, Applied Mathematical Sciences, vol.120, Springer-Verlag, New York (1996).

[12] J.D. Jobson, Applied Multivariate Data Analysis. Volume I : Regression and Experimental Design, Springer-Verlag , New York, (1991).

[13] T. C. Langetieg and J. S. Smoot, *Estimation of the term structure of interest rates*, Research in Financial Services **1** (1989) 181–222.

[14] B. Malkiel, *Efficient Market Hypothesis*, in *New Palgrave Dictionary of Money and Finance*, eds. P. Newman, M. Milgate and J. Eatwell, Macmillan, London (1992).

[15] J. H. McCulloch, *Measuring the term structure of interest rates*, Journal of Business **44** (1971) 19–31.

[16] J. H. McCulloch, *The tax–adjusted yield curve*, Journal of Finance **30** (1975) no.3 811–830.

[17] 111 M. Musiela, M. Rutkowski, Martingale Methods in Financial Modelling, Springer Verlag, Berlin, (1998).

[18] 111 Schwarz G. - *Estimating the Dimension of a Model*, Ann. Stat., 6, 1978.

[19] G. S. Shea, *Pitfalls in smoothing interest rate term structure data: Equilibrium models and spline approximations*, Journal of Financial and Quantitative Analysis **19** (1984) no.3 253–269.

[20] G. S. Shea, *Interest rate term structure estimation with exponential splines: A note*, Journal of Finance **40** (1985) no.1 319–325.

[21] C.J. Stone, M.H. Hansen, C. Kooperberg, Y.K. Truong, *Polynomial splines and their tensor products in extended linear modeling*, The Annals of Statistics **25**, no. 4 (1997) 1371–1470.

[22] C. R. Nelson and A. F. Siegel, *Parsimonious modeling of yield curves*, Journal of Business **60** (1987) no.4 473–489.

[23] L. Svensson, *Estimating and interpreting forward interest rates: Sweden 1992-94*, International Monetary Fund Working Paper (1994) no.114.

[24] O. A. Vasicek and H. G. Fong, *Term structure modeling using exponential splines*, Journal of Finance **37** (1982) no.2 339–348.

[25] O. A. Vasicek, *An equilibrium characterization of the term structure*, Journal of Financial Economics **5** (1977) 177–188.