

L'applicazione delle reti neurali nel *pricing* dei derivati creditizi

Alessandro Ludovici*

Università «G. d'Annunzio» di Chieti-Pescara

Nel presente lavoro si illustra un nuovo approccio al pricing dei derivati creditizi, strumenti finanziari innovativi in grado di immunizzare un portafoglio titoli dal rischio di default degli emittenti, utilizzando le reti neurali. Dopo una trattazione essenziale delle principali problematiche inerenti l'utilizzo di questi strumenti statistici non lineari, particolare enfasi, vista la loro diffusione, viene posta sulle caratteristiche dei Credit Default Swap e sulle peculiarità degli approcci strutturale e in forma ridotta proposti per la loro analisi. Nella parte finale del lavoro si mostra l'efficacia delle reti neurali nell'approssimare la valutazione dei derivati creditizi, e nell'aumentare l'anticipo nella previsione del default. [JEL Classification: C45, G12, G32, G33]

1. - Introduzione

La storia recente dei mercati finanziari mostra come, allo sviluppo impetuoso del processo di innovazione finanziaria, che ha investito tutte le componenti strutturali degli stessi, si sia associato l'impegno costante da parte degli operatori nell'individuare

* <alessandro.ludovici@prometeia.it>. L'Autore ringrazia la Prof.ssa Eliana Angelini per il costante supporto durante la redazione di questo lavoro. Un ringraziamento anche al Prof. Andrea Roli per i preziosi suggerimenti e alla Dott.ssa Maila Bozzetto per l'aiuto nell'analisi dei dati. Nel ringraziare infine i *referee* del Premio Costa per i commenti ricevuti, un pensiero conclusivo va alla memoria del Prof. Antonio Cirillo, i cui insegnamenti hanno reso possibile questo lavoro. L'Autore rimane responsabile di eventuali errori o imprecisioni.

metodologie computazionali più efficienti, in grado di costituire un efficace supporto dinamico di analisi.

Nel presente lavoro si approfondisce un possibile indirizzo nello studio di questo fenomeno, focalizzando l'attenzione sul mercato degli strumenti derivati, che ormai da decenni costituisce l'origine di quasi tutti i più importanti fenomeni innovativi, e sulle reti neurali, metodo computazionale che ha attirato in misura sempre crescente l'interesse degli operatori. Più in particolare, si tenta di analizzare la capacità di questi strumenti di approssimare in maniera praticamente esatta i risultati di complesse e non lineari metodologie di calcolo di valori finanziari.

Nella prima sezione si procede ad un'analisi essenziale dell'architettura delle reti neurali, focalizzando l'attenzione soprattutto sui concetti che risultano maggiormente utili per l'applicazione agli strumenti finanziari.

Nella seconda sezione si affrontano i derivati creditizi nella prima parte, a motivo del rapido sviluppo che ha caratterizzato la loro storia recente come efficaci vie di gestione del rischio di credito: tra di essi enfasi particolare viene rivolta all'analisi dei *Credit Default Swap*, in quanto tra i diversi strumenti risultano essere di sicuro i più diffusi; nella seconda parte invece si procede alla descrizione del modello *CreditGrades*TM, proposto da importanti istituzioni finanziarie come primo tentativo di introdurre maggiore trasparenza nella valutazione di questi strumenti di gestione del rischio di credito. La trattazione si conclude evidenziando come, per quanto riguarda questo comparto dei mercati finanziari, le reti neurali costituiscono un valido strumento di calcolo: non essendosi infatti ancora pervenuti in letteratura alla elaborazione in forma chiusa di una formula di valutazione per i CDSs, in grado di legare gli *spread* di mercato alle variabili sottostanti specifiche di ciascuna impresa analizzata, la rete neurale può, come si vedrà, sopperire efficacemente a questa mancanza, affrontando il problema della determinazione della forma funzionale da un punto di vista statistico. Una ulteriore elaborazione effettuata nello studio concerne invece uno dei maggiori vettori aerei internazionali, la compagnia statunitense UAL Corporation. L'intervallo temporale riguarda il perio-

do 1997-2004: i dati necessari per l'elaborazione sono stati ottenuti tramite i *database* del Federal Reserve System e Bloomberg™. L'elaborazione effettuata mostra come, sebbene tutti i modelli proposti segnalino il deterioramento del merito creditizio con un certo anticipo, CreditGrades™ risulta particolarmente efficiente sotto questo profilo. L'efficienza viene ulteriormente potenziata dall'utilizzo della rete neurale realizzata, che consente di aumentare in maniera apprezzabile il numero di trimestri di anticipo con cui il deterioramento del merito creditizio viene segnalato.

2. - Reti neurali: architettura e applicazioni

Le reti neurali sono state utilizzate in svariati campi di studio, quali ad esempio l'ingegneria, la medicina, la fisica ed altri. Per quanto riguarda i motivi che hanno portato in questo studio ad utilizzare questo tipo di strumenti tra le varie forme di approssimazione non lineare note in letteratura, è necessario ricordare innanzitutto che le reti neurali non richiedono di specificare preliminarmente la relazione esistente tra *input* e *output* del fenomeno studiato: la rete infatti, consentendo, come si vedrà, di determinare un'approssimazione funzionale di tipo non parametrico del fenomeno studiato, si configura come uno strumento dotato di elevate caratteristiche di flessibilità e adattabilità. Inoltre, le reti neurali sono già abbastanza conosciute¹ in ambito economico-finanziario e costituiscono il rappresentante principale della famiglia dei sistemi di apprendimento non lineare. Occorre aggiungere anche che non esistono, per quanto riguarda la letteratura disponibile, studi che applicano tali strumenti ai derivati creditizi, ragione che ha quindi spinto naturalmente ad esplorare questo innovativo campo di analisi. È infine necessario ricordare che, al di là delle caratteristiche di efficacia che verranno meglio illustrate nel seguito della trattazione,

¹ Tra i primi lavori in ambito economico-finanziario cfr. MALLIARIS M.E. - SALCHENBERGER L. (1993).

le reti neurali si connotano anche per tempi di addestramento, di *tuning* dei parametri e di esecuzione relativamente contenuti, poichè utilizzano tecniche di ottimizzazione note da tempo in letteratura. Sebbene le relative strutture si differenzino notevolmente tra loro, è possibile enucleare alcuni principi fondamentali riguardanti essenzialmente il funzionamento di tali strumenti operativi. La parte restante di questa sezione è pertanto dedicata alla presentazione dei principali risultati in materia disponibili in letteratura. Inoltre, è importante impostare la trattazione sottolineando che, per analizzare le dinamiche finanziarie, risultano efficaci reti relativamente poco complesse, perlomeno rispetto a quelle di altri settori².

A. - Architettura delle reti neurali

Una rete neurale associa ad un set di variabili *input* $\{x_i\}$, $i = 1, 2, \dots k$ un set di una o più variabili *output* $\{y_j\}$, $j = 1, 2, \dots h$. Caratteristica essenziale della rete neurale rispetto agli altri metodi di approssimazione è l'utilizzo di uno o più strati nascosti (*hidden layers*), in cui le variabili subiscono una trasformazione di tipo logistico o logsigmoideale: caratteristica questa che, come si vedrà, attribuisce a questi strumenti particolare efficienza nel modellare processi statistici non lineari³.

Nella rete neurale *feedforward* alla elaborazione sequenziale tipica dei metodi di approssimazione lineare si associa l'elaborazione parallela. Infatti mentre nell'elaborazione sequenziale alle variabili *input* osservate vengono attribuiti particolari pesi tramite i neuroni dello strato di *input*, in quella parallela i neuroni dello strato nascosto operano ulteriori trasformazioni al fine di aumentare l'efficienza predittiva. Le connessioni (tra neuroni *input* e neuroni degli strati nascosti, e tra questi e i neuroni *output*) vengono definite sinapsi. La rete neurale *feed-*

² DOLCINO F. - GIANNINI C. - ROSSI E. (1998, p. 1). Per un utile inquadramento del fenomeno in termini generali, cfr. FLOREANO D. - NOLFI S. (1993, pp. 15-50), nonché GORI M. (2003).

³ McNELIS P.D. (2005, p. 21).

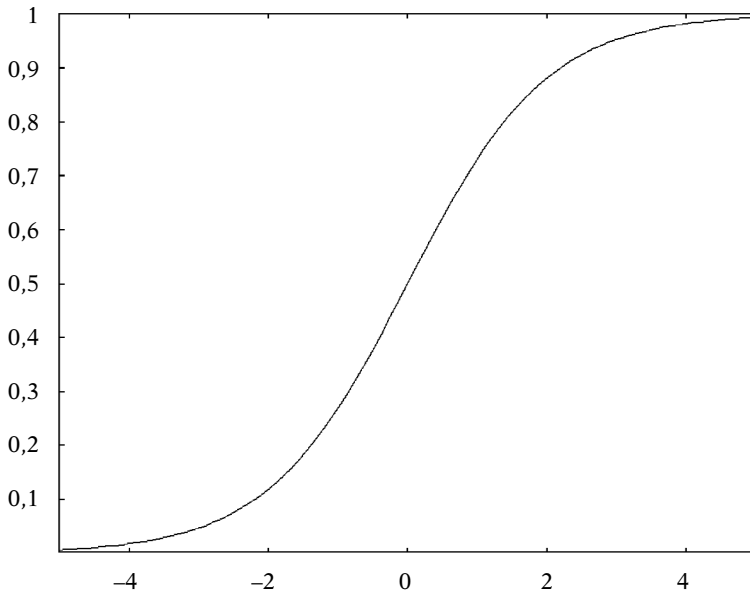
forward con un solo strato nascosto è quella più semplice e allo stesso tempo più utilizzata soprattutto in campo economico e finanziario.

I neuroni processano quindi le variabili *input* in due modi: una prima volta formandone combinazioni lineari e successivamente compiendo delle trasformazioni di queste combinazioni tramite una particolare funzione, tipicamente la funzione logsigmoideale, illustrata nel grafico 1.

Caratteristica fondamentale di questa funzione è l'asintoticità rispetto ai valori 0 e 1, che risulta particolarmente consona ai fenomeni economici, i quali solitamente per valori molto alti o molto bassi delle variabili indipendenti risultano asintotici in quanto piccole variazioni di queste variabili si traducono in variazioni di entità estremamente ridotta delle variabili dipendenti.

GRAF. 1

FUNZIONE LOGSIGMOIDALE



A livello analitico, la rete neurale può essere descritta dalla seguente serie di equazioni⁴:

$$(1) \quad n_{k,t} = \omega_{k,0} + \sum_{i=1}^m \omega_{k,i} x_{i,t}$$

$$(2) \quad N_{k,t} = L(n_{k,t}) = \frac{1}{1 + e^{-n_{k,t}}}$$

$$(3) \quad y_t = \gamma_0 + \sum_{k=1}^q \gamma_k N_{k,t}$$

in cui $L(n_{k,t})$ rappresenta la funzione di attivazione logsigmoideale. Si tratta di un sistema con m variabili *input* x_i e q neuroni. Una combinazione lineare di queste variabili *input*, osservate al tempo t , con il vettore dei coefficienti (o pesi dei neuroni di *input* che dir si voglia) $\omega_{k,i}$ e il termine costante (*bias*) $\omega_{k,0}$ forma la variabile $n_{k,t}$. Questa variabile viene poi trasformata dalla funzione logistica e diviene un neurone $N_{k,t}$ al tempo o all'osservazione t . Il set di q neuroni al tempo o all'osservazione t viene quindi combinato linearmente con il vettore dei coefficienti γ_k e sommato al termine costante γ_0 per ottenere l'*output* y_t relativo al tempo o all'osservazione t , rappresentante il valore stimato dalla rete per la variabile oggetto di analisi in base all'*input* fornito. La rete neurale *feedforward* associata alla funzione di attivazione logsigmoideale viene spesso definita *multi-layer perceptron* o rete MLP.

Un problema di complessità elevata potrebbe essere affrontato ampliando questa struttura, e quindi utilizzando due (rispettivamente N e P) o più strati nascosti⁵:

$$(4) \quad n_{k,t} = \omega_{k,0} + \sum_{i=1}^m \omega_{k,i} x_{i,t}$$

⁴ McNELIS P.D. (2005, p. 25).

⁵ McNELIS P.D. (2005, p. 33).

$$(5) \quad N_{k,t} = L(n_{k,t}) = \frac{1}{1 + e^{-n_{k,t}}}$$

$$(6) \quad p_{l,t} = \rho_{l,0} + \sum_{k=1}^s \rho_{l,k} N_{k,t}$$

$$(7) \quad P_{l,t} = \frac{1}{1 + e^{-p_{l,t}}}$$

$$(8) \quad y_t = \gamma_0 + \sum_{l=1}^q \gamma_l P_{l,t}$$

Aggiungere un secondo strato aumenta il numero di parametri (pesi) da stimare di un fattore $(s + 1)(q - 1) + (q + 1)$, in quanto la rete con un solo strato nascosto, con m input e s neuroni ha $(m + 1)s + (s + 1)$ parametri, mentre la stessa rete con due strati nascosti e q neuroni nel secondo strato nascosto ha $(m + 1)s + (s + 1)q + (q + 1)$ parametri. Tuttavia lo svantaggio di questi modelli per la complessità non consiste tanto nel numero di parametri, che in ogni caso fa perdere gradi di libertà se il campione è limitato ed implica tempi di *training* superiori, quanto nella maggiore probabilità che la rete converga su un ottimo locale piuttosto che globale. È stato comunque dimostrato che una rete neurale con due strati di neuroni è in grado di approssimare qualsiasi funzione non lineare⁶. Un ulteriore punto di forza di questo strumento consiste appunto nel fatto che non si limita ad approssimare un fenomeno in base ad una presunta forma funzionale da adattare, ma procede alla determinazione di quale sia la forma funzionale che meglio riesce a descrivere il fenomeno oggetto di analisi.

Nel grafico 2 è illustrata una rete con un numero multiplo di variabili *output*⁷.

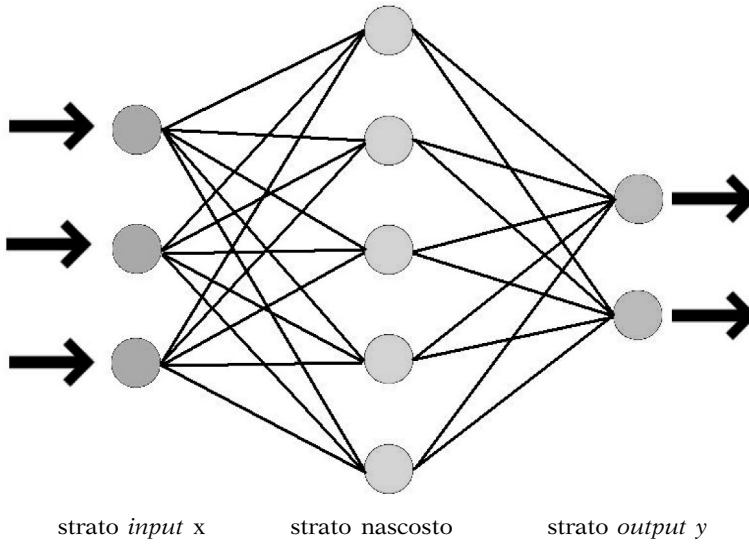
Una rete neurale con uno strato nascosto e due variabili *output* è descritta dal seguente sistema di equazioni⁸:

⁶ DOLCINO F. - GIANNINI C. - ROSSI E. (1998, p. 20); WHITE H. (1990, pp. 540, ss).

⁷ DOLCINO F. - GIANNINI C. - ROSSI E. (1998, p. 2); MARCO G. - VARETTO F. (1994, p. 24).

⁸ McNELIS P.D. (2005, p. 36).

RETE NEURALE CON UNO STRATO NASCOSTO
E DUE NEURONI OUTPUT



$$(9) \quad n_{k,t} = \omega_{k,0} + \sum_{i=1}^m \omega_{k,i} x_{i,t}$$

$$(10) \quad N_{k,t} = L(n_{k,t}) = \frac{1}{1 + e^{-n_{k,t}}}$$

$$(11) \quad y_{1,t} = \gamma_{1,0} + \sum_{k=1}^q \gamma_{1,k} N_{k,t}$$

$$(12) \quad y_{2,t} = \gamma_{2,0} + \sum_{k=1}^q \gamma_{2,k} N_{k,t}$$

È possibile osservare che l'aggiunta di un'ulteriore variabile *output* implica la stima di ulteriori $(q + 1)$ parametri, pari al numero di neuroni dello strato nascosto aumentato di un'unità. Quindi aggiungere variabili *output* implica aumentare il numero

dei parametri da stimare in ragione del numero dei neuroni dello strato nascosto, e non delle variabili *input*. L'utilizzo di una rete neurale con *output* multipli ovviamente ha senso solo laddove questi siano strettamente correlati ovvero dipendano dallo stesso set di variabili *input*: esempi potrebbero essere la struttura temporale dei tassi di inflazione o dei tassi di interesse.

Una delle critiche più spesso rivolte a questi strumenti è quella di essere sostanzialmente delle *black-box*: domande cioè come la natura dei parametri, le ragioni della scelta del loro numero, di quello dei neuroni, di quello degli strati nascosti, i motivi che legano l'architettura della rete alla struttura dei sottostanti fenomeni da spiegare non trovano risposta in simili teorizzazioni⁹.

Il rischio, quando si costruiscono modelli basati su un alto numero di parametri, è che l'estrema ecletticità che ne deriva, consentendo di spiegare tutto e il contrario di tutto finisca per non portare alcun contributo conoscitivo. Tuttavia, va comunque notato che innanzitutto si tratta di una critica che può essere mossa a qualsiasi modello di approssimazione statistica: non solo alle reti neurali quindi, ma anche ai modelli lineari, alla regressione univariata e multivariata e così via. Le reti neurali, in particolare, consentono di spiegare l'andamento di fenomeni per loro natura fortemente irregolari, sulle cui motivazioni dell'andamento è quindi difficile identificare precise relazioni di causa-effetto. In un certo senso quindi la critica *black-box* costituisce paradossalmente anche uno dei maggiori punti di forza delle reti neurali. In ogni caso, la semplicità con cui è possibile aumentare il numero dei parametri della rete non deve mai far dimenticare l'importanza, in ogni modello, della chiarezza e della semplicità delle assunzioni¹⁰.

B. - Normalizzazione dei valori

In generale, una rete neurale non è in grado di analizzare dati o restituire soluzioni in valore assoluto: specialmente nel caso

⁹ McNELIS P.D. (2005, p. 55).

¹⁰ McNELIS P.D. (2005, p. 57).

in cui dovessero presentarsi dati di valore insolitamente elevato o ridotto, potrebbero infatti verificarsi problemi di *overflow* o di *underflow* delle stime. Quando invece sono usate funzioni di tipo sigmoideale, diviene indispensabile preprocessare i dati: questa famiglia di funzioni infatti presenta un codominio di tipo $[0, 1]$ ovvero $[-1, 1]$, per cui i valori devono essere uniformati a questi intervalli altrimenti l'*output* della rete risulterebbe privo di significato, restituendo la soglia superiore o inferiore in corrispondenza di tutti i differenti valori superiori o inferiori ad un determinato limite. In altre parole, per una grande quantità di dati non normalizzati all'intervallo i neuroni semplicemente trasmetterebbero il valore soglia, e con ciò andrebbe persa larga parte dell'informazione. Per quanto riguarda i metodi¹¹, la riduzione lineare trasforma la serie di valori x_k nella serie \hat{x}_k , utilizzando le seguenti formule:

$$(13) \quad \hat{x}_{k,t} = \frac{x_{k,t} - \min(x_k)}{\max(x_k) - \min(x_k)}$$

laddove si voglia ottenere un *output* compreso tra 0 ed 1, ovvero

$$(14) \quad \hat{x}_{k,t} = 2 \frac{x_{k,t} - \min(x_k)}{\max(x_k) - \min(x_k)} - 1$$

laddove si voglia ottenere un *output* compreso tra -1 ed 1.

C. - Processo di apprendimento

Una volta normalizzati i valori, si pone il problema della stima dei parametri (pesi) tramite il processo noto come apprendimento (*training*) della rete neurale¹². In via preliminare è necessario notare che le modalità attraverso cui tale processo può es-

¹¹ McNELIS P.D. (2005, p. 64).

¹² DOLCINO F. - GIANNINI C. - ROSSI E. (1998, pp. 33, ss.); McNELIS P.D. (2005, p. 65); MARCO G. - VARETTO F. (1994, p. 25).

sere portato a termine sono due: una prima, detta “non supervisionata”, che consiste nel sottoporre alla rete i soli dati di *input* del campione, in modo che questa attui una scomposizione in *cluster* degli stessi cogliendo le affinità esistenti tra loro; alla seconda, detta “supervisionata”, che verte sull’analisi sia dei dati di *input* che di *output*, per l’importanza che riveste nella presente trattazione sono principalmente riferibili le considerazioni che seguono.

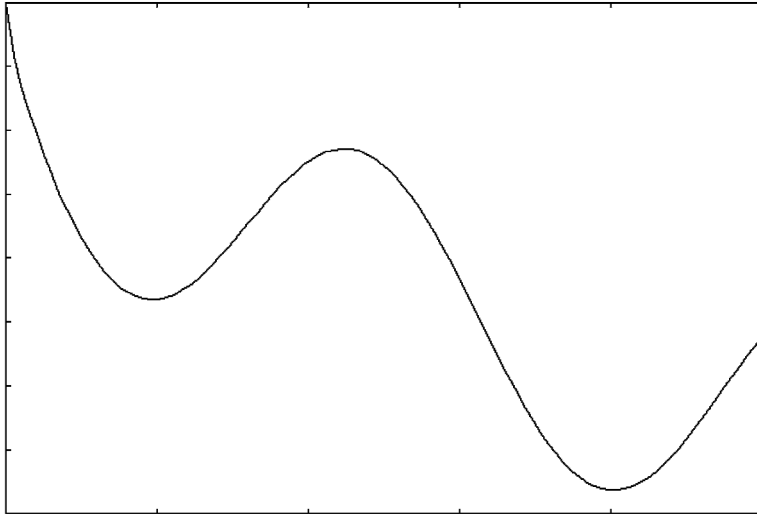
Il *training* della rete è un problema molto più complesso della stima dei parametri di un modello lineare, data la natura di alta complessità non lineare delle reti neurali. Per questi motivi possono esistere numerose soluzioni di ottimo locale, che quindi non minimizzano lo scostamento tra i risultati forniti dalla rete e i valori effettivi che si intende stimare. In breve, in qualsiasi modello non lineare è necessario iniziare la stima dei parametri in base a condizioni che rappresentano ipotesi di larga approssimazione sul valore degli stessi. Tuttavia, come si vedrà, dalla bontà di queste ipotesi iniziali dipende la capacità del processo di stima dei parametri di convergere all’ottimo globale: se infatti le ipotesi iniziali si situano nelle vicinanze, invece che dell’ottimo globale, di un ottimo locale¹³, con tutta probabilità è quest’ultimo che verrà raggiunto.

Questa problematica è illustrata nel grafico 3: il *set* di parametri iniziali (o pesi dei neuroni) se scelto casualmente potrebbe situarsi ovunque sull’asse delle ascisse: laddove si situasse nelle vicinanze di un minimo locale, il processo di addestramento della rete condurrebbe verso questo. Più precisamente, si osserverà in seguito come il processo di addestramento della rete si completa quando si raggiunge un punto in cui la derivata della funzione di perdita si annulla: è appena il caso di ricordare che questa condizione, oltre agli ottimi globali, identifica anche gli ottimi locali nonché i punti di sella. In quest’ottica, si può anticipare sin da ora che il coefficiente di apprendimento, che indica la sensibilità della rete al processo di addestramento, laddove fosse scelto in misura troppo ridotta comporterebbe l’impossibilità della rete di

¹³ DOLCINO F. - GIANNINI C. - ROSSI E. (1998, pp. 29, ss.).

GRAF. 3

ESEMPIO DI SUCCESSIONE DI MINIMO LOCALE E GLOBALE



“sfuggire” agli ottimi locali; laddove invece fosse troppo elevato, potrebbe portare il processo di addestramento ad oscillare continuamente lontano dai punti di ottimo, e la rete quindi divergerebbe.

In termini analitici, è possibile illustrare il processo di apprendimento di una rete con due strati nascosti, per cui occorre quindi determinare il *set* di parametri $\Omega = \{\omega_{k,i}, \rho_{l,k}, \gamma_l\}$.

Il problema consiste¹⁴ nella minimizzazione della funzione di perdita Ψ , definita come la somma dei quadrati delle differenze tra gli *output* campionari osservati y e gli *output* forniti dalla rete \hat{y} :

$$(15) \quad \min_{(\Omega)} \Psi(\Omega) = \sum_{t=1}^T (y_t - \hat{y}_t)^2$$

$$(16) \quad \hat{y}_t = f(x_t; \Omega)$$

¹⁴ McNELIS P.D. (2005).

in cui T è il numero delle osservazioni del vettore *output* y , e $f(x_i; \Omega)$ rappresenta la rete neurale. $\Psi(\Omega)$ è una funzione non lineare di Ω . Tutte le ottimizzazioni non lineari iniziano con una ipotesi iniziale sulla soluzione Ω_0 e cercano ulteriori, migliori soluzioni fino a trovare la migliore possibile entro un ragionevole numero di iterazioni. Molteplici sono le metodologie per condurre questa ricerca¹⁵: alcune fanno riferimento a metaeuristiche quali possono essere gli algoritmi genetici, in alternativa al metodo classico di ottimizzazione chiamato “della discesa del gradiente” o di Newton-Raphson. In ogni caso l'algoritmo di minimizzazione della funzione errore prosegue fino all'ultima iterazione n , o in alternativa si può impostare un criterio di tolleranza¹⁶. Ai fini di evitare gli ottimi locali, una soluzione potrebbe essere quella di determinare una prima convergenza del processo, e poi ripeterlo con un set di parametri iniziali casuali differenti per verificare se la soluzione cambia. Alternativamente, si potrebbe procedere con numerosi processi e poi determinare quello che fornisce la migliore soluzione.

Tuttavia, i maggiori problemi si pongono quando aumenta il numero dei parametri o la forma funzionale della rete diventa particolarmente complessa. Paul John Werbos propose all'inizio degli anni '70 un'alternativa al metodo del gradiente chiamata metodo di retropropagazione¹⁷. Si tratta di un metodo molto flessibile, capace di ovviare ai problemi causati dalla stima della matrice Hessiana nella discesa del gradiente, e di sicuro si tratta del metodo attualmente più usato. Nel passaggio da un'iterazione alla successiva nel processo di stima dei parametri, la matrice Hessiana inversa è infatti rimpiazzata da una matrice identità avente dimensione pari al numero k dei parametri, moltiplicata per il parametro di apprendimento ρ :

$$(17) \quad (\Omega_1 - \Omega_0) = -H_0^{-1}Z_0 = -\rho Z_0$$

Per evitare oscillazioni questo parametro viene scelto nell'in-

¹⁵ McNELIS P.D. (2005, p. 67).

¹⁶ DOLCINO F. - GIANNINI C. - ROSSI E. (1998, p. 44).

¹⁷ McNELIS P.D. (2005, p. 69); MARCO G. - VARETTO F. (1994, p. 27).

tervallo $[0,05, 0,5]$ e inoltre può anche essere endogeno, cioè assumere diversi valori quando il gradiente scende e il processo sembra convergere; o infine si possono adottare parametri differenti per i diversi coefficienti. Rimane comunque il problema della scelta di questo parametro, e dell'esistenza di minimi locali. Inoltre, bassi valori del parametro di apprendimento, sebbene evitino come anticipato l'oscillazione, possono prolungare inutilmente la convergenza del processo di minimizzazione. Questo può tuttavia essere accelerato aggiungendo un *momentum* per cui all'iterazione n avremo¹⁸:

$$(18) \quad (\Omega_n - \Omega_{n-1}) = -\rho Z_{n-1} + \mu(\Omega_{n-1} - \Omega_{n-2})$$

Pertanto, con μ in genere pari a 0,9, si consente al processo di variazione dei parametri di muoversi più velocemente fuori da una "pianura" nella superficie degli errori.

Vediamo ora quali sono i metodi per valutare l'efficacia degli *output* forniti dalla rete.

Relativamente alla valutazione della bontà delle stime prodotte dalla rete, si utilizza in genere l'indice *R*-quadro (*goodness of fit*) per quanto riguarda la capacità della rete di stimare proprio i dati con cui è stata addestrata, ed il *Root mean squared error* per quanto riguarda la capacità della rete di generalizzare le stime al di fuori del campione utilizzato per l'addestramento; in altre parole, suddiviso il campione in due parti, l'una (*in sample*) verrà utilizzata per addestrare la rete, e l'altra (*out of sample*), in genere pari al 25% dei dati complessivi, verrà utilizzata per valutare la capacità della rete di stimare dati provenienti dalla stessa popolazione oggetto di analisi, ma non usati per l'addestramento.

Per quanto riguarda comunque la quantità complessiva di dati necessari¹⁹, indubbiamente una rete neurale richiede la stima di molti più coefficienti rispetto ad esempio ad un modello lineare, implicando questo la necessità di un ampio campione. Sicura-

¹⁸ McNELIS P.D. (2005, p. 70).

¹⁹ DOLCINO F. - GIANNINI C. - ROSSI E. (1998, p. 24), ove si analizzano anche i concetti di "errore di stima" ed "errore di approssimazione".

mente la disponibilità di ampi campioni migliora le capacità predittive della rete, ma implica anche una maggiore durata del relativo processo di addestramento. Inoltre, la disponibilità di un ampio campione non sempre è un aspetto positivo, specie per quanto riguarda il mondo della finanza dove la possibilità di poter indagare dati molto lontani nel tempo porta a distorsioni nei modelli, in quanto i relativi fenomeni sottostanti tendono a variare con estrema rapidità e quindi dati remoti nel tempo non sono più in alcuna connessione con essi.

3. - Derivati creditizi

«I derivati creditizi²⁰ (*credit derivatives*) sono contratti il cui valore finale dipende dal merito creditizio (*creditworthiness*) di una o più entità commerciali o sovrane»²¹. In questa parte si procederà ad un'analisi dei principali contributi disponibili in letteratura riguardanti la tematica della loro valutazione, mirata anche ad introdurre il modello CreditGrades™, che si è affermato come metodologia di primo piano nella valutazione della probabilità di *default*. Iniziamo dunque dai *Credit Default Swap* (CDSs), che negli ultimi anni hanno conquistato la netta superiorità²² in termini di volumi scambiati nel mercato dei derivati creditizi, mercato che tuttavia rimane essenzialmente *over the counter* in quanto non si è ancora pervenuti all'istituzione centralizzata di un sistema di scambi, se non in misura piuttosto contenuta ad esempio in Europa con la creazione dell'indice *iTraxx*.

In generale, i CDSs²³ sono contratti che offrono protezione contro il rischio di insolvenza (evento creditizio o *credit event*)²⁴ di una specifica impresa, denominata soggetto di riferimento (*reference entity*). L'acquirente la protezione ottiene il diritto di ven-

²⁰ J.P. MORGAN (2000, pp. 7, ss.).

²¹ HULL J.C. (2003, p. 708).

²² FRANCIS C. - KAKODKAR A. - MARTIN B. (2003, pp. 4, ss.).

²³ CAPUTO NASSETTI F. - FABBRI A. (2000, p. 89); O'KANE D. - TURNBULL S. (2003, p. 1); J.P. MORGAN (2000, pp. 12, ss.); FRANCIS C. - KAKODKAR A. - MARTIN B. (2003, pp. 10, ss.).

²⁴ CAPUTO NASSETTI F. - FABBRI A. (2000, pp. 31, ss. e p. 59).

dere alla pari una specifica obbligazione (obbligazione di riferimento o *reference obligation*)²⁵, emessa dall'impresa, quando si verifica l'evento creditizio, il cui valore nominale è il capitale nozionale (*notional principal*) del contratto²⁶. Detto diritto è ovviamente ottenuto dietro corresponsione di un prezzo, costituito da una serie di pagamenti periodici, calcolati applicando il prezzo (*spread*) quotato al capitale di riferimento, che si concluderanno con la scadenza del contratto o al verificarsi dell'evento creditizio, che solitamente comporta il pagamento di un rateo finale. Come in ogni contratto derivato, la liquidazione²⁷ può aver luogo con la consegna fisica delle obbligazioni dietro corresponsione del loro valore nominale, ovvero per contanti: in questo caso viene posta in essere una complessa procedura, in cui l'agente incaricato (*calculation agent*) determina il prezzo medio Z dell'obbligazione²⁸ ad una data successiva di un prefissato numero di giorni a quello di verifica dell'evento creditizio, e provvede a liquidare al compratore della protezione una somma pari al $(100 - Z)\%$ del capitale nozionale.

Il mercato dei derivati creditizi, come si è detto, è essenzialmente un mercato OTC, ove quindi svolgono un ruolo determinante²⁹ gli intermediari finanziari in veste di *market maker*: essi rendono noto in ogni istante un prezzo *bid* che sono disposti a pagare per acquistare la protezione ad una data scadenza e su un certo *reference entity*, ed un prezzo *ask* superiore a cui sono disposti a vendere la protezione.

I derivati creditizi sono strumenti molto efficaci per gestire e modificare il profilo di rischio di un portafoglio di esposizioni finanziarie, che risulta ridotto o aumentato nella misura in cui si acquistano o si vendono protezioni. Inoltre, vendendo e acquistando protezioni su *reference entity* operanti in differenti settori, è possibile procedere ad una rapida diversificazione del portafoglio.

²⁵ CAPUTO NASSETTI F. - FABBRI A. (2000, p. 31).

²⁶ HULL J.C. (2003).

²⁷ CAPUTO NASSETTI F. - FABBRI A. (2000, p. 34), ove si trovano approfondimenti sulle diverse modalità di *settlement*.

²⁸ HULL J.C. (2003, p. 709).

²⁹ HULL J.C. (2003); FRANCIS C. - KAKODKAR A. - MARTIN B. (2003, pp. 8, ss.).

La trattazione prosegue illustrando due tra i principali modelli proposti in letteratura per la valutazione di questi strumenti, il primo riconducibile all'approccio *reduced form* ed il secondo riconducibile all'approccio strutturale. Va comunque sottolineato che entrambi i modelli sono applicabili unicamente all'analisi di società emittenti titoli quotati.

A. - Il modello di Hull e White

Gli approcci *reduced form* ipotizzano l'esistenza di una correlazione tra *default* dell'impresa e particolari indicatori macroeconomici della congiuntura, definiti *background factors*. Più specificamente, il modello di Hull e White, alla cui analisi è dedicata questa sezione, si basa sulla fondamentale considerazione che la differenza nel valore di due titoli obbligazionari aventi analoghe caratteristiche, di cui il primo emesso dall'emittente pubblico considerato per definizione privo di rischio, e l'altro emesso da un qualsiasi altro differente emittente, costituisce la stima espressa dal mercato dei costi del *default*. Pertanto, analizzando l'andamento di tale differenziale, è possibile estrapolare in qualsiasi momento la stima espressa dal mercato della probabilità di *default* dell'emittente sulla base delle quotazioni dei titoli obbligazionari.

Tutto ciò premesso, nel seguito si procederà all'illustrazione del modello, assunte le ipotesi semplificatrici di neutralità verso il rischio, di mutua indipendenza delle variabili e di possibile verifica dell'evento creditizio solo alle date di pagamento t_i , $i = 1, \dots, n$, in modo da evitare le complicazioni legate al conteggio del rateo. Si assume inoltre che la probabilità di *default* del venditore la protezione sia nulla, ma su questo si avrà modo di tornare nel seguito. Indicheremo con:

- T : la vita residua in anni del contratto;
- p_i : la probabilità di insolvenza al tempo t_i ;
- \hat{R} : il tasso atteso di recupero, ossia le somme che si prevede di poter comunque incassare dal *reference entity* a seguito del verificarsi dell'evento creditizio;

- $u(t_i)$: il valore attuale, per euro e su base annua, dei pagamenti da 0 a t_i ;
- $e(t_i)$: il valore attuale di un pagamento effettuato al tempo t_i per un importo pari a $t_i - t^*$ euro, ove t^* rappresenta la data di pagamento immediatamente antecedente t_i ;
- $v(t_i)$: il valore attuale di € 1 ricevuto al tempo t_i ;
- w : il pagamento annuo effettuato dal compratore per euro di nozionale;
- s : il valore di w che rende nullo il valore del *credit default swap*;
- π : la probabilità che entro la scadenza non si verifichi l'evento creditizio;
- $A(t_i)$: gli interessi unitari maturati sulla *reference obligation* al tempo t_i .

Come già sottolineato in precedenza, l'approccio assume che le probabilità di *default* sono esogene in quanto desumibili dalle quotazioni dei titoli obbligazionari emessi dal *reference entity* o da contratti derivati³⁰ quali gli *asset swap*. Si avrà pertanto:

$$(19) \quad \pi = 1 - \sum_{i=1}^n p_i$$

Il valore attuale dei pagamenti è:

$$(20) \quad w \sum_{i=1}^n (u(t_i) + e(t_i)) p_i + wu(T) \pi$$

mentre il *payoff* atteso del CDS risulta essere:

$$(21) \quad 1 - \hat{R}(1 + A(t_i)) = 1 - \hat{R} - A(t_i) \hat{R}$$

con valore attuale data l'ipotesi di neutralità verso il rischio pari a³¹:

³⁰ HULL J.C. (2003, pp. 676, ss.).

³¹ HULL J.C. (2003, p. 712).

$$(22) \quad \sum_{i=1}^n (1 - \hat{R} - A(t_i)\hat{R})v(t_i)p_i$$

Di conseguenza, il valore corrente di un CDS per chi compra la protezione è pari alla differenza tra il valore attuale del *payoff* atteso e il valore attuale dei pagamenti:

$$(23) \quad \sum_{i=1}^n (1 - \hat{R} - A(t_i)\hat{R})v(t_i)p_i - w \sum_{i=1}^n (u(t_i) + e(t_i))p_i - wu(T)\pi$$

da cui avremo s (CDS *spread*)³², ossia il valore di w che annulla il valore corrente del contratto:

$$(24) \quad s = \frac{\sum_{i=1}^n (1 - \hat{R} - A(t_i)\hat{R})v(t_i)p_i}{\sum_{i=1}^n (u(t_i) + e(t_i))p_i + u(T)\pi}$$

che dovrà essere corrisposto annualmente dal compratore. Nella tavola 1 è esposto un esempio numerico³³ di utilizzo della formula illustrata; si ipotizza la durata quinquennale del contratto, con pagamenti annuali del premio, e un titolo di riferimento con tasso cedolare del 12% annuo. Si ipotizza inoltre che le probabilità di *default* estrapolate dai dati di mercato come precedentemente esposto siano quelle indicate in tabella, e che il *default* possa avvenire solo alla fine di ciascun anno, mentre il tasso atteso di recupero è posto pari al 40%. Ne segue uno *spread* su base annua pari a 542 *basis point*.

Laddove invece si supponesse che l'insolvenza possa verificarsi in qualsiasi momento, p_i dovrà essere sostituito da $q(t)$, ossia dalla funzione di densità di probabilità delle insolvenze al tempo t , e si avrà pertanto:

³² HULL J.C. (2003).

³³ HULL J.C. (2003); BENKERT C. (2004, p. 81), O'KANE D. - TURNBULL S. (2003, pp. 5, ss.); FRANCIS C. - KAKODKAR A. - MARTIN B. (2003, pp. 16, ss.), per approcci alternativi.

TAV. 1

ESEMPIO DI UN CONTRATTO CDS
(nostre elaborazioni)

	T_i	R	p_i	$A(t_i)$	$v(t_i)$	$(1-R-A(t_i)R)p_i v(t_i)$	$u(t_i)$	$u(t_i)p_i$
	1	0,4	0,05000	0,12000	0,95238	0,02629	0,95238	0,04762
	2	0,4	0,07000	0,12000	0,90703	0,03505	1,85941	0,13016
	3	0,4	0,09000	0,12000	0,86384	0,04292	2,72325	0,24509
	4	0,4	0,10500	0,12000	0,82270	0,04768	3,54595	0,37232
	5	0,4	0,12000	0,12000	0,78353	0,05190	4,32948	0,51954
			0,43500			0,20383		1,31473
π			0,56500					
$\sum u(t_i)p_i + u(T)\pi$			3,76089					
S			0,05420					

$$(25) \quad s = \frac{\int_0^T (1 - \hat{R} - A(t_i)\hat{R})v(t)q(t)dt}{\int_0^T (u(t) + e(t))q(t)dt + u(T)\pi}$$

Per meglio comprendere la natura dello *spread* s è possibile utilizzare una argomentazione di quasi-arbitraggio³⁴. Un acquisto contemporaneo della *reference obligation* e di un CDS di pari scadenza scritto su di essa elimina il rischio di insolvenza. Se indichiamo con y il rendimento alla scadenza dell'obbligazione, risulta evidente che l'annullamento del rischio di credito implica una riduzione di y in misura pari a s , eventualmente non protraentesi sino alla scadenza dell'obbligazione ma fino alla verifica dell'evento creditizio. In sostanza, $y - s$ risulta essere un tasso di rendimento privo di rischio, che pertanto deve essere allineato al tasso *risk free* r di pari scadenza al fine di precludere ogni possibilità di arbitraggio. Tuttavia il compratore della protezione non si vedrebbe corrispondere alcun interesse per il periodo intercorrente tra la data di pagamento dell'ultima cedola e il verificarsi dell'evento creditizio, e inoltre non è detto che il tasso di rendimento che egli riesca a realizzare da quel momento fino alla originaria scadenza del contratto sia pari a r ³⁵.

³⁴ HULL J.C. (2003, p. 713).

³⁵ HULL J.C. (2003, p. 714).

Il tasso atteso di recupero risulta essere l'unica variabile da stimare³⁶ per valutare s , in quanto tutte le altre variabili sono definite nel contratto CDS ($T, w, A(t)$) o desumibili dal mercato ($p_i, u(t), e(t), v(t)$): tuttavia esso ha una limitata incidenza sul CDS *spread*, in quanto le sue influenze, diretta sulle probabilità di *default* ed inversa sul *payoff* in caso di evento creditizio³⁷, tendono ad elidersi. L'incidenza può però essere elevata nel caso dei CDS non *standard*, come ad esempio i *binary* CDS³⁸ che garantiscono al verificarsi dell'evento creditizio un pagamento indipendente dal tasso atteso di recupero, che pertanto va ad influenzare la sola probabilità di *default*.

Altri CDS di tipo non *standard* sono i *basket* CDS³⁹, in cui i soggetti di riferimento sono molteplici ed il pagamento è previsto laddove si verifichi il *default* di uno qualsiasi di loro. L'*add-up basket* CDS è sostanzialmente un portafoglio di CDS ordinari, mentre nel *first-to-default basket* CDS⁴⁰ il contratto si conclude quando si verifica il primo *default*. Questi ultimi strumenti risultano di natura alquanto complessa, e per la loro valutazione occorre procedere tramite simulazioni con il metodo Monte Carlo⁴¹: tuttavia, al crescere della correlazione tra i *default*⁴², il valore del *first-to-default basket* CDS diminuisce, poiché esso risulta meno "diversificato": se invece i *reference entity* sono poco correlati tra loro, sarà più probabile che si verifichi almeno un *default* rispetto al caso in cui sono poco diversificati, e quindi sarà più probabile che venga erogato il pagamento promesso nel contratto in caso di *default*, ragion per cui il valore del contratto stesso aumenta⁴³.

³⁶ HULL J.C. (2003, p. 715); CAPUTO NASSETTI F. - FABBRI A. (2000, pp. 142, ss.) ove si trovano approfondimenti sulla relativa tecnica di stima.

³⁷ HULL J.C. - WHITE A. (2000, p. 19).

³⁸ HULL J.C. (2003, p. 716).

³⁹ HULL J.C. (2003).

⁴⁰ J.P. MORGAN (2000, p. 28), FRANCIS C. - KAKODKAR A. - MARTIN B. (2003, pp. 84, ss.).

⁴¹ HULL J.C. (2003); CAPUTO NASSETTI F. - FABBRI A. (2000, pp. 190, ss.), ove si trovano numerosi approfondimenti sulle tecniche di valutazione basate sulla probabilità condizionata di eventi.

⁴² FRANCIS C. - KAKODKAR A. - MARTIN B. (2003, pp. 89, ss.).

⁴³ HULL J.C. (2003).

È infine necessario considerare anche la probabilità di *default* del venditore della protezione⁴⁴: in tale eventualità infatti, il compratore dovrà stipulare un nuovo CDS con un altro venditore, e ovviamente lo *spread* da corrispondere sarà superiore laddove si sia verificato uno scadimento del merito creditizio del soggetto di riferimento. Quindi, al fine di tener conto degli effetti sul contratto del *default* del venditore, è necessario considerare sia la probabilità di scadimento del merito creditizio del soggetto di riferimento, sia la correlazione di questo evento con il *default* del venditore. Anche in questo caso è possibile quindi procedere unicamente attraverso simulazioni con il metodo Monte Carlo, e per questa via è stato dimostrato⁴⁵ che l'incidenza del rischio di *default* del venditore sul CDS *spread* è direttamente proporzionale alla sopra citata correlazione. Appare comunque evidente che l'acquisto di una protezione debba essere effettuato presso un venditore il cui *default* sia quanto meno possibile correlato a quello del soggetto di riferimento.

B. - Il modello CreditGrades™

Il modello CreditGrades™ per la valutazione del rischio di credito nasce dall'esperienza di operatività sui mercati di quattro tra le principali istituzioni finanziarie: Deutsche Bank, Goldman Sachs, Riskmetrics Group e J.P. Morgan, nel tentativo di creare uno "standard di trasparenza" sui mercati del credito. Dopo il rilascio nel 1997 del modello CreditMetrics™, impostosi come una sorta di *standard de facto* nella valutazione del rischio di credito a livello di portafoglio, nel 2002, a detta delle stesse istituzioni fondatrici del progetto⁴⁶, era sentita l'esigenza di uno strumento in grado di esprimere una valutazione del rischio di credito a livello di singola esposizione: questo sia in seguito alla sempre più pervasiva attività di vigilanza esercitata dalle autorità in questo

⁴⁴ HULL J.C. (2003, pp. 717, ss.).

⁴⁵ HULL J.C. (2003, p. 718); su questi concetti cfr. ANGELINI E. (2002, pp. 276, ss.).

⁴⁶ FINGER C. - FINKELSTEIN V. - LARDY J.P. - PAN G. - TA T. - TIERNEY J. (2002, p. 1).

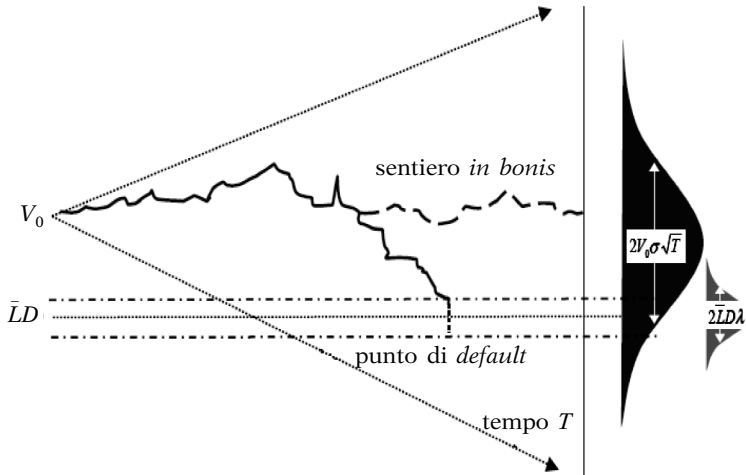
importantissimo ambito dell'intermediazione finanziaria, sia in quanto i mercati negli anni si sono evoluti nel senso di una crescente attenzione verso le tematiche del rischio, potendosi ad esempio citare in merito l'impetuoso sviluppo del mercato dei derivati creditizi visto nella parte precedente. Un ruolo essenziale in questo contesto hanno poi giocato le "crisi" di importanti *corporation*, fattesi sempre più numerose sino ad assumere nel 2001 connotati decisamente preoccupanti. Al di là del celeberrimo "caso Enron", le crisi societarie in quell'anno «furono notevoli sia in termini di incidenza unitaria che di frequenza. I 29 *default* verificatisi totalizzarono più di un miliardo di dollari di debito per ciascuno»⁴⁷.

Il modello CreditGrades™ è riconducibile al cosiddetto "approccio strutturale", che cioè, a partire dagli studi di Black, Scholes e Merton, utilizza la teoria di valutazione delle opzioni per stimare la probabilità di *default* dell'impresa analizzata, al fine di istituire una sorta di collegamento tra il mercato del credito ed il mercato azionario. Infatti gli azionisti della società, secondo questo approccio, qualora il valore delle attività superi quello del debito (e quindi l'opzione risulta *in the money*, cioè esercitabile) procederanno al rimborso, mentre in caso contrario abbandoneranno le attività ai creditori: la loro posizione è quindi assimilabile ai detentori di un'opzione *call* venduta dai creditori. Ne deriva dunque che la scadenza del debito ed il suo ammontare costituiscono rispettivamente la scadenza e lo *strike price* dell'opzione. Pertanto, la probabilità che l'opzione risulti *in the money* è speculare alla probabilità che si verifichi il *default* dell'impresa. Intenzione più volte sottolineata dagli autori di questo modello era quella di amplificare al massimo le caratteristiche di questo approccio, e cioè di utilizzare per le stime solo ed esclusivamente dati aziendali desumibili dalle dinamiche di mercato, senza fare ricorso alcuno a *database* proprietari suscettibili di minare per l'appunto le caratteristiche di trasparenza del modello⁴⁸. Questa parte è dedicata all'analisi degli aspetti essenziali del modello, come delineato nel do-

⁴⁷ FINGER C. - FINKELSTEIN V. - LARDY J.P. - PAN G. - TA T. - TIERNEY J. (2002).

⁴⁸ FINGER C. - FINKELSTEIN V. - LARDY J.P. - PAN G. - TA T. - TIERNEY J. (2002, pp. 2, ss.).

PROCESSO STOCASTICO DELLE ATTIVITÀ DELL'IMPRESA
E PUNTO DI *DEFAULT**



* FINGER C. - FINKELSTEIN V. - LARDY J.P. - PAN G. - TA T. - TIERNEY J. (2002, p. 7).

cumento tecnico di presentazione redatto dai suoi autori. Uno dei tratti peculiari riguarda la definizione del punto di *default*, poiché in questo modello esso viene infatti assunto come stocastico. Si legge al riguardo nel *technical document*: «Non è possibile ipotizzare come noto il livello di *leverage* dell'impresa, se non nel momento in cui effettivamente il *default* si verifica. La natura stocastica del punto di *default* consente di ammettere anche che il valore delle attività dell'impresa sia più vicino ad esso di quanto fosse altrimenti ipotizzato. Questo conduce a *spread* di breve termine superiori»⁴⁹, ovviando in tal modo ad uno dei principali punti di debolezza dell'approccio strutturale. Brevemente, riassumiamo le caratteristiche essenziali di quest'ultimo⁵⁰:

⁴⁹ FINGER C. - FINKELSTEIN V. - LARDY J.P. - PAN G. - TA T. - TIERNEY J. (2002, p. 5).

⁵⁰ FINGER C. - FINKELSTEIN V. - LARDY J.P. - PAN G. - TA T. - TIERNEY J. (2002, p. 6); per una descrizione sintetica del modello *cfr.* STAMICAR R. - FINGER C.C. (2005, pp. 2, ss.).

- le attività dell'impresa seguono un processo stocastico V ;
- il *default* avviene quando detto processo oltrepassa la soglia costituita dal punto di *default*;
- tale punto di *default* deve necessariamente corrispondere ad una qualche frazione del totale delle passività; nel modello dipende dal tasso di recupero medio L , supposto stocastico.

In relazione all'ultimo punto, la natura stocastica di L è confermata anche da analisi empiriche svolte al riguardo citate nel *technical document*, che hanno mostrato l'estrema variabilità di questo valore, influenzato ad esempio dalla natura finanziaria ovvero operativa del dissesto, dalle prospettive di risanamento o di liquidazione dell'impresa e così via. L è quindi distribuito in maniera lognormale con media \bar{L} e deviazione *standard* λ ; indicando con D il debito per azione, si avrà:

$$(26) \quad LD = \bar{L}De^{\left(\lambda z - \frac{\lambda^2}{2}\right)}$$

La probabilità che l'impresa risulti *in bonis* al tempo t è quindi uguale alla probabilità che V non scenda al di sotto del punto di *default* LD prima del tempo t , per cui si avrà⁵¹:

$$(27) \quad P(t) = N\left(-\frac{A_t}{2} + \frac{\log(d)}{A_t}\right) - dN\left(-\frac{A_t}{2} - \frac{\log(d)}{A_t}\right)$$

$$(28) \quad d = \frac{V_0 e^{\lambda^2}}{LD}$$

$$(29) \quad A_t^2 = \sigma^2 t + \lambda^2$$

In queste equazioni $N(x)$ indica la funzione di distribuzione cumulativa normale *standard* della variabile x , mentre σ indica la volatilità degli *asset* del *reference entity*.

⁵¹ FINGER C. - FINKELSTEIN V. - LARDY J.P. - PAN G. - TA T. - TIERNEY J. (2002).

Ottenuta dunque la probabilità di *default* ($1 - P(t)$), si pone il problema di tradurre questi risultati in adeguati *spread* creditizi⁵². A tal fine è necessario introdurre il tasso di recupero R che, a differenza di L definito come tasso di recupero medio di tutte le diverse tipologie di debito, si riferisce invece alla particolare forma di intervento creditizio cui lo *spread* è applicato.

Posta la funzione di densità di probabilità per il momento del *default*⁵³

$$(30) \quad f(t) = -\frac{dP(t)}{dt}$$

la probabilità cumulativa di *default* fino al tempo t sarà

$$(31) \quad 1 - P(0) + \int_0^t dsf(s)$$

Riferendo l'analisi ad un CDS con scadenza t e *spread* c , il valore attuale dei pagamenti attesi dovuti alle perdite sarà

$$(32) \quad (1 - R)(1 - P(0) + \int_0^t dsf(s)e^{-rs})$$

ove r come sempre rappresenta il tasso di interesse *risk free*, supposto costante. Il valore attuale dei flussi di pagamenti attesi relativi allo *spread* sarà

$$(33) \quad c \int_0^t dsP(s)e^{-rs}$$

Dalla differenza di questi due pagamenti consegue quindi il prezzo del CDS⁵⁴; lo *spread* di equilibrio, che, eguagliando i pagamenti relativi alle perdite attese a quelli relativi allo *spread*, annulla il valore del CDS, sarà quindi⁵⁵

⁵² FINGER C. - FINKELSTEIN V. - LARDY J.P. - PAN G. - TA T. - TIERNEY J. (2002), p. 45).

⁵³ FINGER C. - FINKELSTEIN V. - LARDY J.P. - PAN G. - TA T. - TIERNEY J. (2002).

⁵⁴ FINGER C. - FINKELSTEIN V. - LARDY J.P. - PAN G. - TA T. - TIERNEY J. (2002), p. 46).

⁵⁵ FINGER C. - FINKELSTEIN V. - LARDY J.P. - PAN G. - TA T. - TIERNEY J. (2002).

$$(34) \quad \hat{c} = r(1-R) \frac{1-P(0)+H(t)}{P(0)-P(t)e^{-rt}-H(t)}$$

ove

$$(35) \quad H(t) = e^{r\xi}(G(t+\xi)-G(\xi))$$

$$(36) \quad G(t) = d^{z+\frac{1}{2}} N\left(-\frac{\log(d)}{\sigma\sqrt{(t)}} - z\sigma\sqrt{(t)}\right) + d^{-z+\frac{1}{2}} N\left(-\frac{\log(d)}{\sigma\sqrt{(t)}} + z\sigma\sqrt{(t)}\right)$$

$$(37) \quad z = \sqrt{\left(\frac{1}{4} + \frac{2r}{\sigma^2}\right)}$$

$$(38) \quad \xi = \frac{\lambda^2}{\sigma^2}$$

La problematica inerente la determinazione del valore e della volatilità σ delle attività dell'impresa viene quindi affrontata in maniera del tutto peculiare dal modello, sempre volto a coniugare il rigore metodologico con il tentativo di semplificare la stima dei parametri⁵⁶: prima di tutto si focalizza l'attenzione sul lungo termine, ossia $t > \lambda^2 / \sigma^2$: in caso contrario infatti sulla probabilità di *default* influisce principalmente il parametro λ ; inoltre, tenendo sempre presente come punto di riferimento il lemma di Itô, la distanza dal *default* viene definita come il numero di deviazioni *standard* su base annuale che separano il valore presente delle azioni dell'impresa dal punto di *default*⁵⁷:

$$(39) \quad \eta = \frac{1}{\sigma} \log\left(\frac{V}{LD}\right) = \frac{V}{\sigma_s S} \frac{\partial S}{\partial V} \log\left(\frac{V}{LD}\right)$$

ove con S si indica il corso dell'azione dell'impresa e con σ_s se ne indica la volatilità. A questo punto, è possibile esaminare le seguenti condizioni al contorno⁵⁸:

⁵⁶ FINGER C. - FINKELSTEIN V. - LARDY J.P. - PAN G. - TA T. - TIERNEY J. (2002, p. 10).

⁵⁷ STAMICAR R. - FINGER C.C. (2005, p. 10).

⁵⁸ FINGER C. - FINKELSTEIN V. - LARDY J.P. - PAN G. - TA T. - TIERNEY J. (2002).

– al tendere verso il punto di *default*, S/LD risulta molto minore di 1, pertanto si ha $V = S + LD \approx LD$ ed è quindi possibile scrivere l'approssimazione

$$(40) \quad V \approx LD + \frac{\partial V}{\partial S} S$$

e sostituendo nella (39) avremo:

$$(41) \quad \eta \approx \frac{1}{\sigma_s}$$

– per valori elevati di S/LD è invece possibile assumere che $S/V \rightarrow 1$, cioè il tasso di incremento di V eguaglia quello di S , e risulta

$$(42) \quad \eta \approx \frac{1}{\sigma_s} \log\left(\frac{S}{LD}\right)$$

Ne deriva che tutte queste condizioni possono essere contemporaneamente soddisfatte dalle seguenti relazioni:

$$(43) \quad V = S + LD$$

$$(44) \quad \eta = \frac{(S+LD)}{\sigma_s S} \log\left(\frac{S+LD}{LD}\right)$$

Per cui per V_0 avremo

$$(45) \quad V_0 = S_0 + \bar{LD}$$

$$(46) \quad \sigma = \sigma_s \frac{S}{S + \bar{LD}}$$

Quest'ultima equazione mostra che, per una volatilità σ delle attività costante, la volatilità ed il prezzo delle azioni sono legati da una relazione inversa: ragion per cui all'approssimarsi del *default* la volatilità delle azioni raggiunge livelli molto elevati, e questo risulta in linea⁵⁹ con il fenomeno empirico del *volatility smi-*

⁵⁹ FINGER C. - FINKELSTEIN V. - LARDY J.P. - PAN G. - TA T. - TIERNEY J. (2002, p. 11).

le. Dato tuttavia che nei mercati azionari possono anche osservarsi più o meno temporanei andamenti erratici dei prezzi, il modello suggerisce di assumere a riferimento i valori non transitori \hat{S} e $\hat{\sigma}_s$, volatilità azionaria che può essere sia storica che implicita, ai fini di determinare un livello della volatilità delle attività da considerarsi stabile⁶⁰.

Si avrà dunque

$$(47) \quad \sigma = \hat{\sigma}_s \frac{\hat{S}}{\hat{S} + \bar{L}D}$$

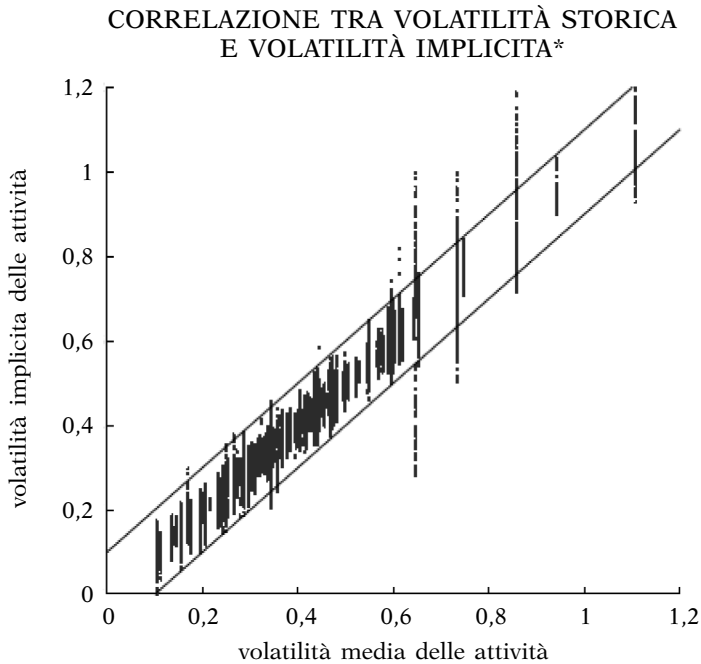
Per la verifica empirica del modello, descritta nel dettaglio nel *technical document*, occorre preliminarmente determinare uno stimatore della volatilità delle attività⁶¹, elemento non osservabile direttamente sul mercato, e assunto come costante nel lungo periodo dal modello. Ponendosi questo nel filone cosiddetto strutturale, occorre comunque che tale stimatore dipenda in qualche modo dalla volatilità azionaria, posti come parametro di verifica i CDSs *spread* osservati sul mercato. Per la verifica gli autori del modello hanno utilizzato un campione costituito da 122 imprese, operanti sia nel settore industriale che in quello finanziario, relativamente all'andamento dei CDSs *spread* a 5 anni nel periodo maggio 2000 - agosto 2001. I dati provengono dal *database* J.P. Morgan per quanto riguarda i valori dei tassi di recupero specifici di ciascuna *reference obligation*, e un totale di 6.194 quotazioni è stato utilizzato. Una volta fissati gli usuali valori $\bar{L} = 0,5$ e $\lambda = 0,3$ e ricordando le citate condizioni al contorno (40)-(42), la volatilità delle attività implicita nel modello è il valore di σ che consente di ottenere un CDS *spread* pari a quello osservato sul mercato, e si nota che storicamente tale valore appare piuttosto stabile⁶²: il grafico 5 evidenzia che solo per 4 imprese lo scostamento tra il valore di lungo periodo e quello osservato supera il 10%. Pertanto si procede con il calcolo della volatilità storica per cia-

⁶⁰ FINGER C. - FINKELSTEIN V. - LARDY J.P. - PAN G. - TA T. - TIERNEY J. (2002).

⁶¹ FINGER C. - FINKELSTEIN V. - LARDY J.P. - PAN G. - TA T. - TIERNEY J. (2002), p. 17).

⁶² FINGER C. - FINKELSTEIN V. - LARDY J.P. - PAN G. - TA T. - TIERNEY J. (2002).

GRAF. 5



* FINGER C. - FINKELSTEIN V. - LARDY J.P. - PAN G. - TA T. - TIERNEY J. (2002, p. 19).

scuna impresa, e, secondo gli autori del modello, i risultati migliori al riguardo si ottengono⁶³ sulla base di un periodo di osservazione compreso tra 750 e 1.000 giorni: si fa notare tuttavia che i risultati più aderenti sono relativi alle imprese più rischiose, e ciò è dovuto al fatto che mentre per queste ultime esiste una stretta relazione tra andamento delle azioni e andamento del debito, per i prenditori caratterizzati da elevato merito creditizio si configura invece una accentuata⁶⁴ *volatility skew*, in quanto il *default* di questi sarebbe in pratica assimilabile ad un caso di *deep out of the money*. Utilizzando quindi una finestra temporale di

⁶³ FINGER C. - FINKELSTEIN V. - LARDY J.P. - PAN G. - TA T. - TIERNEY J. (2002, p. 18).

⁶⁴ FINGER C. - FINKELSTEIN V. - LARDY J.P. - PAN G. - TA T. - TIERNEY J. (2002, p. 19).

1.000 giorni, l'88% dei valori stimati cade entro un *range* del 10% dal valore assunto come reale, cioè la volatilità implicita nel CDS *spread* a 5 anni osservato sul mercato. Prove successive hanno mostrato che detto risultato, essendo invariante rispetto all'epoca storica di osservazione, risulta robusto sia per il settore industriale sia per quello del credito⁶⁵.

Si è già avuto modo per l'appunto di osservare che il deterioramento del merito creditizio si accompagna spesso a incrementi della volatilità e degli *spread* applicati sul debito: date le difficoltà che si incontrano nell'utilizzo della volatilità storica, il modello è stato esteso⁶⁶ per utilizzare la volatilità implicita nelle opzioni ad un anno scritte sulle azioni del *reference entity*: questo consente maggiore accuratezza delle stime e il superamento delle difficoltà insite nella redazione per tempi discreti dei prospetti di bilancio⁶⁷. L'utilizzo della volatilità implicita quindi⁶⁸, anche se ovviamente produce stime più variabili, consente di catturare con maggiore rapidità i segnali lanciati dalle quotazioni di mercato.

Dal grafico 6 risulta evidente come l'approssimazione fornita dal modello utilizzando la volatilità implicita (linea tratteggiata), anche non considerando la correzione per il *volatility skew* (linea sottile), fornisce una stima del CDS *spread* (linea spezzata) nettamente migliore di quella ottenibile con l'utilizzo della volatilità storica (linea a puntini). Infine, sono descritte nel *technical document* le misure di associazione⁶⁹ tra i risultati del modello e quelli effettivamente osservati sul mercato: per quanto riguarda l'indice di correlazione, occorre tener presente che maggior peso relativo assumono le alte probabilità di *default* delle imprese caratterizzate da merito creditizio più basso: anche considerando questi fattori di disturbo, l'indice si situa sempre al di sopra del 60%, con una probabilità minima di corretta classificazione del merito creditizio pari al 75%.

⁶⁵ FINGER C. - FINKELSTEIN V. - LARDY J.P. - PAN G. - TA T. - TIERNEY J. (2002).

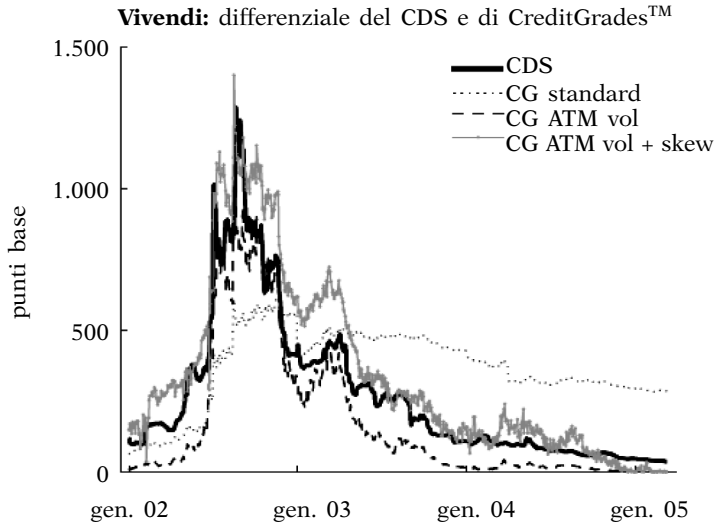
⁶⁶ STAMICAR R. - FINGER C.C. (2005, p. 1).

⁶⁷ HULL J.C. - NELKEN I., WHITE A. (2004, pp. 6, ss.).

⁶⁸ STAMICAR R. - FINGER C.C. (2005, pp. 26).

⁶⁹ FINGER C. - FINKELSTEIN V. - LARDY J.P. - PAN G. - TA T. - TIERNEY J. (2002, p. 28).

ANDAMENTO DEGLI *SPREAD* STIMATI
RELATIVI ALLA SOCIETÀ VIVENDI
AL VARIARE DELLA TIPOLOGIA DI VOLATILITÀ CONSIDERATA*



* STAMICAR R. - FINGER C.C. (2005, p. 18).

C. - Applicazione delle reti neurali al modello CreditGrades™

La parte finale di questo lavoro è dedicata all'illustrazione dei risultati ottenuti dagli esperimenti di applicazione delle reti neurali condotti. Scelta dell'architettura della rete, numero degli strati e dei neuroni, processi di apprendimento e relative misure di efficienza, valutazione della capacità predittiva della rete sono tutte problematiche⁷⁰ di difficile e soprattutto non univoca risposta, a cui solo l'esperienza e la verifica delle diverse metodologie esistenti può trovare soluzione. Gli approcci proposti vanno quindi visti essenzialmente come punti di partenza, passi obbligati verso analisi più approfondite e quindi più specifiche, che difficilmente si prestano ad una descrizione in termini generali.

Nella restante parte di questa sezione si esporranno le po-

⁷⁰ McNELIS P.D. (2005, p. 109).

tenzialità delle reti neurali nella approssimazione del *pricing* dei derivati creditizi utilizzando dati artificiali⁷¹, cioè generati sulla base del modello CreditGrades™. Il motivo primario che spinge verso una tale scelta è che in questo modo non esistono limiti alle dimensioni dei campioni disponibili, né per quelli *in sample* né per quelli *out of sample*, oltre ovviamente alle caratteristiche di aderenza alla realtà del modello appena viste.

Occorre quindi generare un campione casuale concernente le variabili indipendenti del modello (S, D, σ_s, t, r, R), ricordando che siamo in presenza di equazioni omogenee in D ⁷², per cui si può porre $D = 1$ e sostituire S con S/D : per quanto riguarda i *range* di variazione, si è posto $25\% < S/D < 625\%$ e $10\% < \sigma_s < 90\%$, valori⁷³ largamente prevalenti sul mercato, e nella stessa ottica si è posto $3\% < r < 7\%$ e $3 < t < 7$ anni. L'*output* della rete è costituito da due valori, il primo dei quali rappresenta il CDS *spread* (che per le restrizioni poste sopra risulterà pari al massimo a 1.500 *basis point*, valore anche qui più che adeguato), ed il secondo indicante la probabilità di *default*: si tratta di variabili strettamente correlate tra loro, e quindi è possibile utilizzare un'unica rete neurale per la loro stima. Il campione *in sample* è costituito da 500 osservazioni.

TAV. 2

VALUTAZIONE DELLA STIMA DELLA RETE NEURALE
(nostre elaborazioni)

tipologia di errore	valore
<i>R</i> -quadro <i>spread</i>	0,99
<i>R</i> -quadro prob. default	0,99
Root mean squared error <i>spread</i>	0,00147
Root mean squared error prob. <i>default</i>	0,00744

La tavola 2 mostra i valori dell'*R*-quadro e del *root mean squared error*: come è facile osservare, si tratta di risultati di elevata

⁷¹ McNELIS P.D. (2005, p. 115); HANKE M. (1997, p. 21).

⁷² STAMICAR R. - FINGER C.C. (2005, p. 14).

⁷³ In alternativa sarebbe possibile utilizzare i dati di volatilità dai relativi mercati, allo scopo di focalizzare meglio il tipo di problema per la rete neurale.

concordanza. La rete neurale *multi-layer perceptron* è stata addestrata in modalità “supervisionata” per 1.000 epoche di apprendimento utilizzando l’algoritmo *backpropagation*, con un solo *layer* nascosto di 19 neuroni. Entrambi tali valori derivano da un accurato studio preliminare condotto secondo il metodo della *cross-validation*⁷⁴, ossia scomponendo, come descritto nella precedente sezione 2, il campione in una parte *in sample* da utilizzare per l’addestramento di reti differenti per numero di neuroni e di strati nonché per epoche di apprendimento, ed una parte *out of sample* necessaria per determinare quale rete minimizza l’errore di previsione. In osservanza dei risultati analitici già descritti secondo cui con lo strumento rete neurale è possibile approssimare qualsiasi fenomeno non lineare, risulta pertanto evidente che anche la nuova applicazione alla sottofamiglia dei derivati creditizi cattura efficacemente la variabilità insita nelle dinamiche di mercato: una interessante direzione di ulteriore approfondimento dell’analisi, cui sono orientati gli sviluppi futuri dello studio, è sicuramente quella di sostituire al campione di valori casuali, dei valori reali rilevati sul mercato. Come evidenziato a suo tempo, vista anche la novità dell’applicazione dello strumento rete neurale ai derivati creditizi, nel presente lavoro si è invece scelto di operare mediante dati artificiali anche perché, al di là di considerazioni riguardanti la disponibilità delle informazioni, in ogni caso qualsiasi tentativo di stima di fenomeni economici tramite reti neurali presuppone⁷⁵ l’assunzione di un modello di base in grado di individuare le variabili supposte esplicative del fenomeno analizzato. In questo contesto si è ritenuta particolarmente proficua l’esposizione del modello CreditGrades™, per le sue caratteristiche di semplicità e di robustezza oltre che per la sua ampia diffusione tra gli operatori. Tuttavia, date le osservazioni sopra formulate in merito all’aderenza alla realtà di questo modello, l’estensione dei risultati della rete neurale ai dati di mercato non porrebbe difficoltà. Non essendosi pervenuti in letteratura ad una

⁷⁴ McNELIS P.D. (2005, pp. 115, ss.); su questi concetti cfr. BISHOP C.M. (1995) e HECHT-NIELSEN R. (1991).

⁷⁵ Per vari esempi al riguardo cfr. McNELIS P.D. (2005, pp. 145, ss.), in particolare pp. 148, ss.

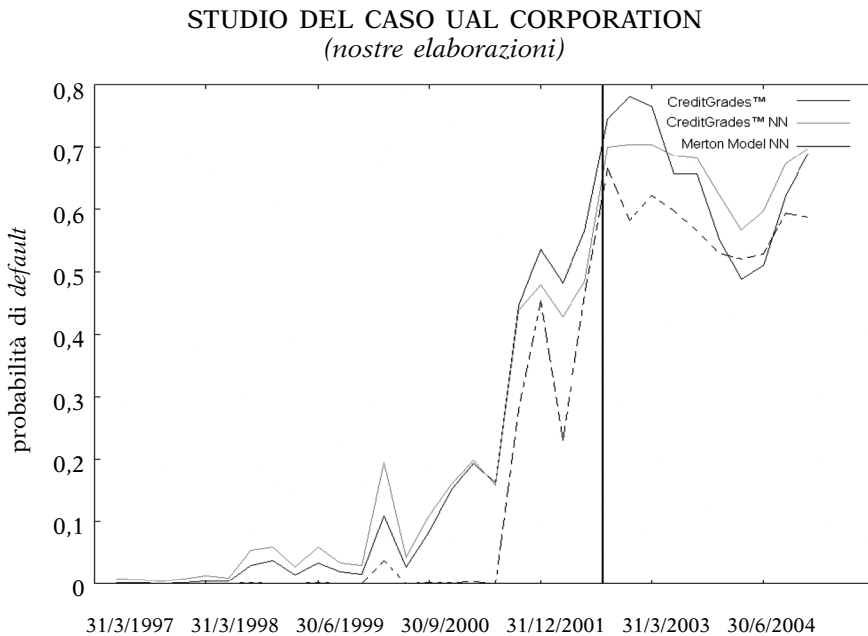
univoca determinazione della forma funzionale di valutazione dei CDSs, la rete neurale può quindi essere vista come un efficace strumento di elaborazione in grado di sopperire a questa mancanza.

D. - Previsione del default: il caso UAL Corporation

In conclusione della trattazione riportiamo nel grafico 7 un'elaborazione riguardante l'impresa statunitense UAL Corporation, operante nel settore dei trasporti aerei come uno dei maggiori vettori internazionali. La società richiese il 12 settembre del 2002 di aderire al capitolo 11 della normativa americana in materia di dissesti, dichiarando l'avvenuto risanamento finanziario il 1° febbraio 2006.

Utilizzando i dati esposti nella tavola 3, sono state calcolate le probabilità di *default* per ciascun trimestre del periodo marzo

GRAF. 7



'97-dicembre '04 di questa impresa, in base a 3 differenti modelli: il modello CreditGrades™ (in nero), la rete neurale addestrata secondo le risultanze del modello CreditGrades™ (in grigio), ed un'altra rete neurale addestrata secondo le risultanze del classico modello di Merton (linea tratteggiata). Nel grafico la linea verti-

TAV. 3

DATI RELATIVI AL CASO UAL CORPORATION
(nostre elaborazioni)

Data	Leverage D/E	Vol. imp. %	Tasso risk-free
31/03/97	0,9008	35,5400	0,0602
30/06/97	0,8120	36,4100	0,0567
30/09/97	0,7095	35,9500	0,0547
31/12/97	0,7878	38,3700	0,0551
31/03/98	0,9293	40,2400	0,0541
30/06/98	1,1247	33,0900	0,0538
30/09/98	1,4788	46,5000	0,0441
31/12/98	1,7558	43,5600	0,0453
31/03/99	1,3761	39,3000	0,0472
30/06/99	1,5614	46,5200	0,0507
30/09/99	1,5238	38,9200	0,0522
31/12/99	1,3534	40,3100	0,0598
31/03/00	1,7447	68,9900	0,0628
30/06/00	1,6977	38,7800	0,0608
30/09/00	2,5322	44,4500	0,0607
31/12/00	3,6114	43,7700	0,0532
31/03/01	4,2346	46,1000	0,0409
30/06/01	3,9570	39,0200	0,0372
30/09/01	9,3055	78,9300	0,0249
31/12/01	13,6767	92,4500	0,0217
31/03/02	11,3719	68,7800	0,0270
30/06/02	16,0735	90,6400	0,0206
30/09/02	72,5431	287,1200	0,0153
31/12/02	5,9430	325,9800	0,0132
31/03/03	9,4754	345,5400	0,0119
30/06/03	9,3012	233,1100	0,0109
30/09/03	6,1978	226,0700	0,0115
31/12/03	4,7573	170,0500	0,0126
31/03/04	4,8408	143,7500	0,0120
30/06/04	4,9365	152,0600	0,0209
30/09/04	9,9011	203,4300	0,0221
31/12/04	7,9689	256,7400	0,0275

cale nera indica la data del *default*, ed è facile osservare come tutti e tre i modelli presentati segnalino con largo anticipo il deterioramento della condizione finanziaria della società attraverso l'incremento della probabilità di *default* misurata sulle ordinate: tuttavia da questo caso emerge la maggiore capacità segnaletica relativa del modello CreditGrades™ rispetto al classico modello di Merton, e inoltre la particolare rete neurale addestrata sulla base delle sue risultanze in una certa misura “amplifica” le buone caratteristiche predittive del modello CreditGrades™, incrementando l'anticipo temporale con cui il deterioramento del merito creditizio viene segnalato attraverso l'aumento della probabilità di *default*. Sono indicati nella tavola 3 i dati utilizzati per l'analisi, ottenuti dal *database* Bloomberg™ per quanto riguarda il *leverage* e la volatilità azionaria implicita, e dal *database* del Federal Reserve System per quanto riguarda il tasso di interesse a durata annua costante dei Treasury Bills, assunto come tasso *risk free*. L'orizzonte temporale dell'analisi è parimenti annuo, mentre il tasso di recupero specifico è posto uguale al 50%.

BIBLIOGRAFIA

- ANGELINI E., «Il profilo di rischio di un portafoglio con derivati creditizi: l'elaborazione di un modello di simulazione Montecarlo», *Banca Impresa Società*, n. 2, a. XXI, 2002, pp. 259-282.
- BISHOP C.M., *Neural Networks for Pattern Recognition*, Oxford, Oxford University Press, 1995.
- CAPUTO NASSETTI F. - FABBRI A., *Trattato sui contratti derivati di credito*, Milano, EGEA, 2000.
- DOLCINO F. - GIANNINI C. - ROSSI E., *Reti neurali artificiali per l'analisi e la previsione di serie finanziarie*, Milano, Collana studi del Credito Italiano, 1998.
- FINGER C. - FINKELSTEIN V. - LARDY J.P. - PAN G. - TA T. - TIERNEY J., *CreditGrades Technical Document*, New York, RiskMetrics Group, 2002.
- FLOREANO D. - NOLFI S., «Reti neurali: algoritmi di apprendimento, ambiente di apprendimento, architettura», *Giornale Italiano di Psicologia*, a. XX, febbraio 1993, pp. 15-50.
- FRANCIS C. - KAKODKAR A. - MARTIN B., *Credit Derivative Handbook 2003*, London, Merrill Lynch Global Securities Research and Economics Group, 2003.
- GORI M., «Introduzione alle reti neurali artificiali», *Mondo Digitale*, AICA, n. 4, settembre 2003.
- HANKE M., «Neural Network Approximation of Option-Pricing Formulas for Analytically Intractable Option-Pricing Models», *Journal of Computational Intelligence in Finance*, settembre-ottobre, 1997.
- HECHT-NIELSEN R., *Neurocomputing*, Reading (MA), Addison-Wesley Publishing Company, 1991.
- HULL J.C., *Opzioni, futures e altri derivati*, Milano, Il Sole 24Ore S.p.A., 2003.
- HULL J.C. - NELKEN I. - WHITE A., «Merton's Model, Credit Risk, and Volatility Skews», *Journal of Credit Risk*, n. 1, 2004, pp. 1-27.
- HULL J.C. - WHITE A., «Valuing Credit Default Swaps I: No Counterparty Default Risk», *Journal of Derivatives*, vol. 8, n. 1, autunno 2000, pp. 29-40.
- J.P. MORGAN, *The J.P. MORGAN guide to credit derivatives*, Londra, Risk Publications, 2000.
- MALLIARIS M.E. - SALCHENBERGER L., «A Neural Network Model for Estimating Option Prices», *Journal of Applied Intelligence*, n. 3, 1993, pp. 193-206.
- MARCO G. - VARETTO F., *Diagnosi delle insolvenze e reti neurali*, Roma, Bancaria Editrice, 1994.
- MCNELIS P.D., «Neural Networks in Finance», San Diego, *Elsevier Academic Press*, 2005.
- O'KANE D. - TURNBULL S., «Valuation of Credit Default Swaps», *Quantitative Credit Research Quarterly*, Lehman Brothers, vol. 2003-Q1/Q2, aprile, 2003.
- STAMICAR R. - FINGER C.C., *Incorporating Equity Derivatives into the CreditGrades Model*, New York, RiskMetrics Group, 2005.
- WHITE H., «Connectionist Nonparametric Regression: Multilayer Feed-Forward Networks Can Learn Arbitrary Mappings», *Neural Networks*, n. 3, 1990, pp. 535-549.