

Modelli di interazione spaziale: presupposti teorici e aspetti applicativi

Elena Cefis e Giuseppe Espa

1. Introduzione

I modelli di interazione mirano a modellare la dipendenza o l'autocorrelazione spaziale tra unità. Sorti nell'ambito della statistica spaziale, tali modelli sono stati recentemente utilizzati in economia per spiegare l'interazione tra agenti, ovvero come il comportamento di un agente venga influenzato dal comportamento di altri agenti presenti nella stessa economia (mercato, luogo, etc.). In relazione al fatto che il comportamento di un agente venga influenzato dal comportamento della totalità degli altri agenti oppure che venga influenzato da solo un sottoinsieme di essi si distinguono modelli di interazione globale e modelli di interazione locale.

L'interazione locale presuppone che gli agenti vengano influenzati dal comportamento di agenti ad essi, in qualche senso, vicini¹. L'idea di vicinato introduce un'esplicita dimensione spaziale. Gli agenti vengono pensati come fisicamente posizionati su un ambiente di riferimento, ovvero vengono considerati come nodi di un grafo o come celle di un reticolo.

La prima importante applicazione dei modelli di interazione spaziale in ambito economico si ha con Föllmer nel 1974 che rimuove due ipotesi fondamentali della teoria microeconomica convenzionale, permettendo che le preferenze di un agente possano essere stocastiche e che possano dipendere dal comportamento degli altri agenti. In generale, i modelli di interazione spaziale hanno avuto un'ampia applicazione in economia poiché rappresentano uno strumento utile sia per l'analisi di problemi specifici, sia per la creazione di un nuovo paradigma teorico alternativo a quello dell'equilibrio generale. In particolare i modelli dinamici di interazione spaziale sono stati utilizzati per individuare i fondamenti microeconomici (nel comportamento degli individui) delle dinamiche aggregate. La struttura teorica macroeconomica standard si basa infatti sull'ipotesi dell'agente rappresentativo. Tale ipotesi presuppone che gli individui siano omogenei e che agiscano in isolamento, ovvero che interagiscono tra loro solo indirettamente attraverso il meccanismo dei prezzi di mercato. I modelli dinamici di interazione spaziale permettono di derivare le dinamiche aggregate (collettive) dall'interazione di agenti eterogenei. Una diretta conseguenza di ciò è rappresentata dal fatto che le fluttuazioni del comportamento aggregato possono avere origine non solo da shock esogeni al sistema, ma anche endogenamente dalla semplice interazione tra gli agenti.

I modelli di interazione spaziale sono stati proficuamente applicati in contesti specifici per spiegare fenomeni economici in cui gli agenti possono rivedere periodicamente le loro scelte e in cui l'economia è caratterizzata da una forte frammentazione sociale, economica e/o geografica, da informazione imperfetta ed incompleta e da forti esternalità di rete che spingono gli agenti a comportamenti

¹ Il concetto di vicinato è da intendersi nel senso più ampio possibile, potendosi considerare come "vicini" tutti gli agenti rilevanti in base ad una specifica relazione, come ad esempio i settori interagenti all'interno di una tavola input/output.

imitativi. Sono quindi molteplici i campi in cui sono stati utilizzati tali modelli. Orlean (1990, 1992) e Kirman (1991) studiano i fenomeni di contagio delle opinioni nei mercati finanziari, mostrando come i cambiamenti delle opinioni presenti sui mercati, causati dall'interazione stocastica tra gli individui, riescano a spiegare l'estrema volatilità dei prezzi dei titoli e la formazione di bolle speculative.

David e Foray (1992) applicano un modello di interazione spaziale (e più precisamente di filtrazione) per spiegare la diffusione di una tecnologia innovativa in imprese caratterizzate da forti differenze nel livello di informazione che possiedono sulle caratteristiche della nuova tecnologia e nella capacità di riconoscere i) l'utilità derivante dall'adozione della nuova tecnologia e ii) gli effetti delle esternalità di rete derivanti dalla coordinazione. Sempre nel medesimo contesto Dalle (1994) analizza il processo di diffusione tecnologica in un ambiente caratterizzato da forte incertezza e informazione imperfetta (riguardo le caratteristiche delle tecnologie esistenti) che spesso implicano l'esistenza di esternalità di rete.

Durlauf (1993) studia la crescita di un sistema economico caratterizzato da una molteplicità di industrie eterogenee (cioè caratterizzate da funzioni di produzione distinte) che interagiscono intensamente tra loro. Il modello genera equilibri multipli a cui è associata una semplice struttura probabilistica che dimostra come le interazioni locali possono influire sull'equilibrio aggregato.

Infine Haining (1984, 1990) utilizza i modelli di interazione spaziale per spiegare la differenziazione di prezzo di un bene omogeneo in una singola area urbana.

I modelli di interazione spaziale sono strumenti estremamente utili per studiare gli aspetti spaziali dell'economia (in questo caso al concetto di "vicinato" diamo un'accezione più strettamente geografica). Esistono infatti interazioni geografico-economiche importanti non destinate a scomparire nel breve tempo. Nonostante la crescente dematerializzazione e internazionalizzazione della produzione Quah (1996) ha mostrato come la locazione fisica e gli spillover geografici siano fattori importanti per spiegare la dinamica delle distribuzioni dei redditi regionali. Come Krugmann (1992) ha evidenziato, i modelli ad interazione spaziale si rivelano buoni strumenti in relazione allo sviluppo di due campi dell'economia: da una parte lo studio empirico della dinamica di crescita delle regioni e dall'altra l'analisi teorica di sistemi caratterizzati da rendimenti crescenti e forti esternalità locali.

All'interno dei modelli di interazione locale esistono molte tipologie di modelli in relazione alla struttura delle interazioni degli agenti: strutture più o meno omogenee, statiche e simmetriche (per un'esauriente rassegna sulla loro applicazione in economia si veda Fagiolo, 1996). Il presente lavoro è focalizzato sull'analisi dei modelli basati sui campi aleatori di Gibbs e di Markov su reticoli. Si vogliono presentare in modo semplice e nello stesso tempo esaustivo i presupposti statistici su cui si basa la teoria dei campi aleatori (o "Random Fields"), non senza trascurare i principali aspetti applicativi concernenti la stima e la simulazione di tali modelli. L'obiettivo è quello di fornire uno

strumento di lavoro utilizzabile in un ampio ventaglio di contesti, da quelli più propriamente teorici a quelli più specificatamente applicativi.

Dopo una breve introduzione sull'origine storica della statistica spaziale (paragrafo 2), vengono introdotti alcuni elementi derivati dalla statistica fisica sui quali si basa l'analisi delle interazioni spaziali, quali il concetto di clique, la nozione di stazionarietà dei processi aleatori spaziali e le definizioni di campi aleatori di Gibbs e di Markov (paragrafi 3-6). Nel paragrafo 7 vengono presentati gli automodelli per variabili qualitative (sconnesse e ordinabili) e quantitative, con particolare riguardo al modello auto-logistico (che corrisponde al modello del ferromagnetismo di Ising). Successivamente (paragrafo 8) vengono riportati in modo sintetico altri modelli di interazione spaziale che rappresentano approcci alternativi nel caso particolare di processi gaussiani. Nel paragrafo 9 vengono presentati due metodi per la stima dei parametri dei modelli di interazione: la massima verosimiglianza e la pseudo-verosimiglianza. Infine vengono discussi i metodi di simulazione di un processo auto-normale e illustrati due algoritmi per la simulazione di un campo aleatorio markoviano a stati discreti (paragrafo 10). Il paragrafo 11 conclude.

2. Origini storiche ed evoluzione della Statistica Spaziale

La letteratura statistica riguardante l'analisi di dati spaziali ha avuto origine circa quaranta anni fa come ramo dell'analisi delle serie storiche e dell'inferenza da processi stocastici temporali grazie ai lavori pionieristici di Whittle (1954) e Bartlett (1955). Questa circostanza ha caratterizzato marcatamente lo sviluppo della disciplina nei suoi primi anni di vita. Infatti per molti anni, ed in particolare con un ritmo accelerato a partire dagli anni '70, è fiorita una gran quantità di studi i cui sforzi erano nella direzione di estendere a due dimensioni i procedimenti messi a punto per serie di dati ordinati nel tempo.

La convinzione di fondo in questi anni è stata che i problemi incontrati lungo questo percorso (ad esempio la non unidirezionalità di una serie spaziale) fossero ostacoli da rimuovere e che al termine di tale processo l'analisi delle serie spaziali avrebbe avuto la sua codificazione come era accaduto alla sua più celebre antenata analisi delle serie storiche (Box e Jenkins, 1976). La maggior parte degli studi di questo periodo sono ottimamente riassunti in volumi quali Bennett (1979) e Cliff e Ord (1981).

A partire dal principio degli anni '80, tuttavia, si sono verificate due circostanze destinate a mutare radicalmente la prospettiva degli studiosi di statistica spaziale.

In primo luogo una serie di studi hanno dimostrato come alcuni dei problemi emersi, lungi dall'essere puri incidenti di percorso, minavano dal fondo la possibilità di estendere a due dimensioni i risultati dell'analisi delle serie storiche. Tra questi vanno sicuramente ricordati il problema della scarsa rilevanza pratica dell'ipotesi di stazionarietà per processi aleatori nello spazio (Haining, 1990; Ripley, 1988), la

presenza di innegabili effetti di bordo sfortunatamente troppo spesso ignorati (Ripley, 1988), i problemi di natura asintotica in schemi cartesiani regolari finiti e, infine, il problema dell'unità areale modificabile ossia dell'indeterminatezza dei procedimenti di analisi spaziale al mutare dei confini delle aree studiate (Arbia, 1989).

In secondo luogo lo sviluppo accelerato e la diffusione su larga scala di sistemi automatici di calcolo hanno reso disponibili in questi anni una ricchissima serie di informazioni in formato digitale facilmente accessibile e trattabile. In particolare la diffusione di immagini computerizzate nelle più svariate discipline (quali, ad esempio, le immagini da satellite ma anche immagini da radar, da aereo, da microscopi elettronici ed immagini di fenomeni clinici quali le ecografie, le tomografie assiali e le risonanze magnetiche nucleari) e di mappe automatiche quali quelle generate dai moderni Sistemi Geografici Informativi (Burrough, 1986; Arbia, 1993a) (ad esempio nella pianificazione ambientale o per il controllo ambientale) ha posto una serie di nuove domande agli studiosi di statistica spaziale.

In queste nuove situazioni, nelle quali le competenze dello statistico sempre più spesso si associano a quelle di fisici e di ingegneri, il paradigma di riferimento non è più unicamente quello dell'analisi delle serie storiche. In tal senso alcuni concetti mutuati dalla statistica fisica (come, ad esempio, quello di *stato* introdotto in meccanica da Gibbs) sono destinati a svolgere un ruolo di crescente importanza nella statistica spaziale.

3. L'interazione nello spazio

Sia $S = \{s_1, s_2, \dots, s_n\}$ un insieme finito di n località che possono essere considerate sia come punti di un grafo sia come celle di un reticolo regolare². Sia inoltre Ω l'insieme di tutte le possibili realizzazioni di una variabile aleatoria X sullo spazio S . Ogni elemento di Ω definirà quindi una possibile configurazione spaziale del fenomeno considerato.

È logico attendersi che l'insieme S , che rappresenta le località, mostri una qualche struttura geografica. Dal momento che, come espresso nella *Prima legge della geografia* formulata da Tobler (1970), *everything is related to everything else, but near things are more related than distant things*, risulta necessario formalizzare i meccanismi con cui gli eventi interagiscono nello spazio. In questa direzione la nozione di vicinanza fra oggetti diviene fondamentale. Formalizzando con l'ausilio delle probabilità condizionate, il sito s_j è un *vicino* del sito s_i , con $i \neq j$, se la probabilità

$$P(X_i = x_i \mid \text{tutte le } X_k = x_k, k \neq i) \quad [3.1]$$

² Nel prosieguo, la generica località verrà indicata indifferentemente con la notazione s_i appena introdotta o con l'indice i che la caratterizza.

dipende da x_j , la realizzazione della variabile aleatoria X_j nella località s_j . Dalla definizione [3.1] consegue immediatamente la definizione di un insieme di elementi $C\{i\} \subset S$ che siano contigui ad un generico sito s_i . L'insieme $C\{i\}$ è tale da racchiudere in sé l'informazione su s_i contenuta in tutte le unità del sistema, cioè è l'insieme sufficiente a spiegare la dipendenza di s_i dagli altri elementi dello spazio geografico di riferimento (Benedetti, 1994). Volendo esprimere in maniera formale i concetti poc'anzi illustrati, il sottoinsieme dei vicini della località s_i è costituito dalle località s_j tali che

$$P(X_i = x_i | X_j = x_j, \forall j \neq i) = P(X_i = x_i | X_j = x_j, j \neq i, j \in C\{i\}). \quad [3.2]$$

Il modo classico di specificare l'insieme $C\{i\}$ per ogni località s_i ($i=1, 2, \dots, n$) è quello di definire una matrice di connettività C il cui generico elemento $c_{ij}=c_{ji}$ indica l'appartenenza o meno ($c_{ij}=1$ o 0 rispettivamente) di s_j al sottoinsieme delle località contigue all'i-ma.

Molti criteri riguardanti la specificazione della struttura della matrice di contiguità sono stati suggeriti nella letteratura (Cliff e Ord, 1981); di tali principi i più diffusi nelle applicazioni pratiche sono i seguenti (cfr. Figura 3.1):

criterio della torre: $C\{i\} = \{s_{i-1,j}, s_{i,j+1}, s_{i+1,j}, s_{i,j-1}\}$;

criterio dell'alfiere: $C\{i\} = \{s_{i-1,j-1}, s_{i-1,j+1}, s_{i+1,j+1}, s_{i+1,j-1}\}$;

criterio della regina: $C\{i\} = \{s_{i-1,j}, s_{i,j+1}, s_{i+1,j}, s_{i,j-1}, s_{i-1,j-1}, s_{i-1,j+1}, s_{i+1,j+1}, s_{i+1,j-1}\}$.

Un modo alternativo di specificare un sistema di vicinato che si è rivelato di grande utilità nelle applicazioni concernenti dati disposti su reticoli regolari scaturisce dalla definizione di *clique*. Un *clique* è un sottoinsieme di siti contenuto in S tale che tutte le coppie di località che lo compongono sono mutuamente vicine, dove la relazione di vicinato è definita in senso probabilistico dalla relazione (3.2).

$i-1, j-1$	$i-1, j$	$i-1, j+1$
$i, j-1$	i, j	$i, j+1$
$i+1, j-1$	$i+1, j$	$i+1, j+1$

Fig.3.1 Criteri di vicinato nella specificazione delle matrici di connettività.

Singole località costituiscono per definizione dei *cliques*. Inoltre se R_i denota l'insieme dei *cliques* che contengono l' i -mo sito, allora un sistema di vicinato sarà la classe ordinata $R = \{R_1, R_2, \dots, R_n\}$, dove R denoterà l'insieme di tutti i *cliques* contenuti in una struttura di vicinato. Infine va precisato che genericamente i *cliques* sono a due a due non disgiunti (Dubes e Jain, 1989).

Il sistema di vicinato più utilizzato nell'analisi di dati disposti su sistemi cartesiani regolari considera come vicini del primo ordine di una generica cella le quattro caselle che hanno con quel sito un lato in comune. I vicini del secondo ordine saranno così i quattro *pixels* che condividono con il generico sito solo uno spigolo (cfr. Figura 3.2). In Figura 3.3 sono riportate le tipologie di *clique* che esistono quando in un reticolo regolare il sistema di vicinato è ristretto al primo e al secondo ordine.

2	1	2
1	i	1
2	1	2

Fig.3.2 Ordine di vicinato di un generico sito i .

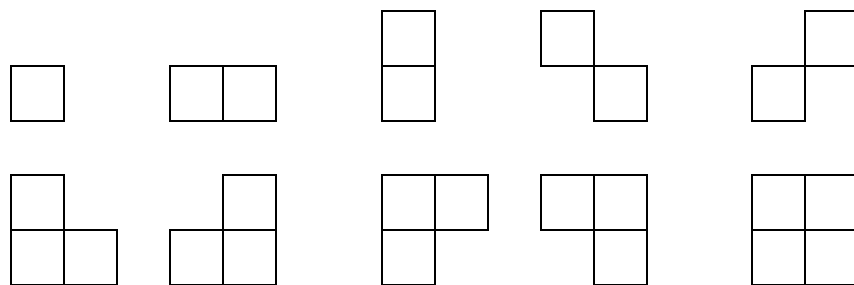


Fig. 3.3 Tipologie di *clique* in un sistema di vicinato di primo e di secondo ordine.

Deve essere sottolineato che il concetto di sistema di vicinato permette di distinguere tra vicini di primo ordine e vicini di secondo ordine, ovvero permette di ordinare per importanza i vicini rilevanti. Questa possibilità è assai importante perchè permette di modellare relazioni di diversa importanza tra gli agenti oppure di introdurre in modo indiretto diverse tipologie di agenti i cui comportamenti pesano in modo differente sull'evoluzione del fenomeno preso in considerazione. Nella letteratura

economica la possibilità di modellare l'interazione tra agenti che possiedono sistemi di vicinato con ordini di vicinato composti (vicini di primo grado e vicini di secondo, etc.) non è quasi mai stata presa in considerazione. Piuttosto è stato sottolineato come limite dei modelli di interazione spaziale basati sui campi di Markov l'impossibilità di introdurre agenti che esercitino differenti gradi di influenza (ci riferiamo in particolare al limite rappresentato dall'impossibilità di introdurre istituzioni e organizzazioni intermedie che solitamente è stato attribuito a tali modelli (Hors e Lordon, 1995)).

4. Stazionarietà dei processi aleatori spaziali

Un processo stocastico spaziale è un insieme di n variabili aleatorie (X) indicizzate tramite un insieme I . Tale insieme di indici può essere costituito da coppie di numeri reali individuando così dei punti all'interno di una superficie continua (Diggle, 1983) oppure da coppie di numeri interi specificando in tal modo, per esempio, le posizioni delle celle all'interno di un grigliato regolare (Haining, 1990). Un processo stocastico è individuato univocamente dalla funzione di densità congiunta $f_{1,2,\dots,n}(X_1, X_2, \dots, X_n)$ delle n variabili casuali (Ripley, 1981, pag. 9). La necessità di fare inferenza sui parametri incogniti di tale funzione di densità impone che detta distribuzione possieda alcune proprietà.

Il concetto più importante da questo punto di vista è quello di stazionarietà del processo, concetto che implica che le proprietà del processo non mutino nello spazio. In senso stretto ciò significa che considerato un sottoinsieme della funzione di densità congiunta, la legge di distribuzione non muta se il sottoinsieme viene traslato (omogeneità) o ruotato (isotropia) rispetto al grigliato. Una conseguenza della stazionarietà in senso stretto è che essa è trasmessa ad ogni ordine di momenti. Condizioni più deboli possono essere quindi derivate semplicemente imponendo la stazionarietà dei momenti solo fino ad un certo ordine. Quando le variabili casuali sono distribuite normalmente la stazionarietà spaziale debole (fino al secondo ordine) implica stazionarietà spaziale in senso stretto dal momento che i processi Gaussiani sono individuati univocamente dai momenti del primo e del secondo ordine.

Formalmente, sia I una generica disposizione di siti all'interno di un reticolato regolare che si suppone essere di dimensioni finite. Siano inoltre $s=(s_1, s_2)$ e $i=(i_1, i_2)$ due punti qualunque in I . Infine siano $\gamma(s, i)$ e $R(s, i)$ rispettivamente la covarianza e la correlazione tra due variabili casuali in posizione s e i :

$$\gamma(s, i) = E\{[X(s) - E(X(s))][X(i) - E(X(i))]\}$$

$$R(s, i) = \gamma(s, i) / [\gamma(s, s)\gamma(i, i)]^{1/2}$$

dove il simbolo E denota il valore atteso.

La stazionarietà dei momenti fino al secondo ordine significa formalmente che medie e varianze non devono cambiare da cella a cella (indipendenza da \mathbf{s}):

$$E[X(\mathbf{s})] = \mu_s = \mu \forall \mathbf{s} \text{ e } \text{Var}[X(\mathbf{s})] = \sigma_s^2 = \sigma^2 \forall \mathbf{s}. \quad [4.1]$$

Ad affiancare le [4.1] ci sono poi le condizioni su $\gamma(\mathbf{s}, \mathbf{i})$ e $R(\mathbf{s}, \mathbf{i})$. In particolare se, oltre ad essere valide le [4.1], $\gamma(\mathbf{s}, \mathbf{i})$ dipende solo dalla posizione relativa di \mathbf{s} ed \mathbf{i} e non da quella assoluta, cioè dipende esclusivamente dalla lunghezza e dal verso del vettore $\mathbf{h}=(h_1, h_2)$ condotto tra \mathbf{s} ed \mathbf{i} , allora $\gamma(\mathbf{s}, \mathbf{i})$ ed $R(\mathbf{s}, \mathbf{i})$ sono indipendenti dalla posizione ed il processo spaziale è detto debolmente omogeneo. In questo caso si avrà il seguente risultato:

$$\gamma(\mathbf{s}, \mathbf{i}) = \gamma(\mathbf{h}) \text{ e } R(\mathbf{s}, \mathbf{i}) = R(\mathbf{h}).$$

Se $\gamma(\mathbf{s}, \mathbf{i})$ e $R(\mathbf{s}, \mathbf{i})$ non dipendono da come il vettore \mathbf{h} è orientato ma solo dalla sua lunghezza allora il processo spaziale sarà isotropico in senso debole e si potrà scrivere:

$$\gamma(\mathbf{s}, \mathbf{i}) = \gamma(r)$$

dove $r = (h_1^2 + h_2^2)^{1/2}$. Per inciso si noti che per le definizioni riportate si assume che la varianza sia positiva e finita (Ripley, 1981, pag. 9-10).

È importante sottolineare, a conclusione di questa sezione, che la stazionarietà è una proprietà tanto utile quanto irrealistica dei processi che si suppone abbiano generato i dati spaziali. Se così fosse infatti, bisognerebbe essere disposti ad accettare che la correlazione tra una variabile osservata ad esempio a Roma ed a Milano sia la stessa di qualsiasi altra coppia di paesi, purchè distanti 500km, il che è in conflitto con il più elementare buon senso. Inoltre, nella pratica, l'analisi spaziale ha a che fare con aree che sono fisicamente limitate o con sottoaree di più vaste zone che sono anch'esse limitate. Per tali motivi risulta spesso più utile e realistico stabilire delle condizioni di *quasi*

o *locale* omogeneità e isotropia, condizioni per le quali le proprietà richieste valgono per un'area specifica e non sussistono per punti situati al di fuori del dominio spaziale considerato. Alcuni metodi per l'individuazione di zone di stazionarietà locale e di stima dei parametri incogniti per dati generati da leggi di questo tipo si trovano in Arbia e Benedetti (1990).

5. Campi aleatori di Gibbs

Un campo aleatorio di Gibbs discreto (o Gibbs Random Field, abbreviato in GRF, seguendo la dizione anglosassone) relativo al grigliato definito dall'insieme S e dalla matrice \mathbf{C} (cfr. paragrafo 3.), è una particolare misura di probabilità π definita sullo spazio degli eventi Ω tale che:

$$\pi(\mathbf{X} = \mathbf{x}) = Z^{-1} \exp[-U(\mathbf{x})/T]. \quad [5.1]$$

La funzione $U(\mathbf{x})$ è detta *funzione di energia* ed assume la seguente forma analitica:

$$U(\mathbf{x}) = \sum_{c \in Q} V_c(\mathbf{x})$$

dove c sono i *cliques* dell'insieme S . In generale il termine V_c rappresenta il *potenziale* dello stato di Gibbs; si tratta di una funzione associata ad ogni *clique* che dipende esclusivamente dai valori assunti da \mathbf{X} nelle località che costituiscono il *clique* c .

Nella relazione [5.1], la quale fornisce un modello globale per un sistema spaziale discreto e regolare, Z e T sono delle costanti. Di queste la prima, che è una costante di normalizzazione la quale prende il nome di funzione di partizione, è tale che:

$$Z = \sum_{\mathbf{x}} \exp[-U(\mathbf{x})/T].$$

La costante Z di cui alla relazione precedente è generalmente non computabile, eccezion fatta per alcune semplici circostanze. L'impossibilità di calcolare Z pone considerevoli problemi in fase di campionamento e stima.

La seconda costante (T) rappresenta invece la temperatura all'interno di sistemi fisici. A basse temperature la configurazione assume una forma regolare; al contrario alte temperature inducono una distribuzione di tipo caotico (al limite per $T \rightarrow \infty$ la distribuzione è interamente dovuta al caso). L'inverso del parametro, $\delta=1/T$, può essere pertanto visto come una misura del grado di casualità di una configurazione spaziale (Geman e Geman, 1984; Ripley, 1988). La relazione [5.1] esprime il fatto che la probabilità che ha ciascuna configurazione di realizzarsi in pratica dipende dalla funzione di energia e da un parametro che descrive la regolarità della configurazione stessa (T) (Arbia, 1993b).

A chiusura di questa breve digressione sulle distribuzioni di Gibbs (anche dette distribuzioni di Boltzman) preme sottolineare in questa sede che un GRF si parametrizza mediante la scelta di un potenziale. Ad esempio Derin e Elliott (1987) hanno utilizzato la seguente funzione di clique:

$$V_c = \begin{cases} \xi_c & \text{se tutti i siti nel } \textit{clique } c \text{ hanno stesso valore} \\ -\xi_c & \text{altrimenti.} \end{cases}$$

Il parametro ξ_c quantifica l'influenza del *clique* c sulle caratteristiche del GRF.

Quando gli unici potenziali non nulli sono, con riferimento alla Figura 3.1, quelli del tipo $\{(i, j)\}$, $\{(i-1, j), (i, j)\}$ e $\{(i, j-1), (i, j)\}$, possibilmente con correzioni ai bordi del grigliato, allora il GRF è detto processo di interazione a due a due.

6. Campi aleatori markoviani e teorema di Hammersley e Clifford

Un campo aleatorio di Gibbs descrive le proprietà globali di un sistema spaziale in termini della distribuzione di probabilità congiunta [5.1] per i valori relativi a tutti i *pixels*. Esiste però una seconda classe di misure di probabilità per processi spaziali che, a differenza degli stati di Gibbs, non ha alcun riferimento apparente con la fisica. Si tratta dell'estensione dei processi markoviani ad uno spazio a due dimensioni. Tali modelli, chiamati campi aleatori markoviani (o Markov Random Fields, abbreviato in MRF, seguendo la dizione anglosassone) vengono definiti mediante proprietà locali.

Un campo aleatorio, relativamente ad un sistema di vicinato, è un MRF discreto o continuo se, rispettivamente, la sua distribuzione di probabilità o la sua funzione di densità soddisfa le tre proprietà seguenti:

i) condizione di positività:

$$P(\mathbf{X} = \mathbf{x}) > 0 \quad \forall \mathbf{x} \in \Omega. \quad [6.1]$$

Questa stretta condizione di positività della probabilità permette che il sistema non venga bloccato in uno stato. Questa è una delle proprietà che ha reso questi modelli interessanti in ambito economico. Il fatto che la probabilità di un agente di fare una delle scelte ammissibili sia strettamente positiva (qualunque sia la scelta e qualunque sia la configurazione dell'intero sistema) permette che il sistema possa uscire da stati di *lock-in*. Quando formulati in modo dinamico questi modelli danno la possibilità di spiegare perchè, ad esempio in un modello di diffusione tecnologica, una tecnologia si espanda, rimanga dominante per un determinato periodo e successivamente venga rimpiazzata da un'altra tecnologia. Questa possibilità di uscire da stati di *lock-in* non esiste in altri modelli utilizzati per modellare processi di diffusione, come, ad esempio, i modelli basati sulle urne di Polya (per una rassegna generale su questo si veda Dosi, Ermoliev, Kaniovski, 1994).

ii) proprietà di Markov:

$$P(X_i = x_i | X_j = x_j, \forall j \neq i) = P(X_i = x_i | X_j = x_j, j \neq i, j \in C\{i\})^3; \quad [6.2]$$

Tale proprietà afferma che l'agente *i-esimo* viene influenzato solo dal comportamento degli agenti vicini e non dal comportamento di tutti gli agenti presenti nell'intero sistema.

.	.	.
.	<i>i</i>	.
.	.	.

Fig.6.1 L'insieme $\tilde{o}i$ quando il *clique* R_i contiene tutti i vicini del secondo ordine.

iii) condizione di omogeneità:

$$P(X_i = x_i | X_j = x_j, \forall j \neq i) = P(X_i = x_i | X_j = x_j, j \neq i, j \in C\{i\})$$

ovvero la forma funzionale è la stessa per tutte le celle che costituiscono il sistema spaziale di riferimento. In altre parole, gli agenti si comportano tutti secondo le stesse regole di comportamento, anche se questo non implica che tutti gli agenti si comportino ugualmente poichè le “modalità” di interazione (ciò che più avanti viene identificato con i parametri β_{ij}) possono essere localmente diverse.

Specificare un potenziale cioè stabilire una funzione di *clique* come, ad esempio, quella utilizzata da Derin e Elliott (1987) e riportata nel paragrafo precedente, consente di formulare un modello globale mentre ipotizzare un modello per le probabilità

³ La relazione può essere riscritta con un utile cambio di notazione nei seguenti termini:

$$P(X_i = x_i | \mathbf{X}_{n \setminus i} = \mathbf{x}_{n \setminus i}) = P(X_i = x_i | \mathbf{X}_{\partial i} = \mathbf{x}_{\partial i}) \quad [6.2bis]$$

dove l'indice $n \setminus i$ fa riferimento a tutte le n celle che costituiscono il sistema spaziale di riferimento esclusa l' i -ma mentre il sottoinsieme ∂i comprende tutti i siti che si trovano nelle vicinanze della cella i stessa. Ad esempio se il *clique* Q_i contiene tutti i vicini del secondo ordine, ∂i sarà l'insieme mostrato in Figura 6.1.

condizionate di cui alla relazione [6.2] poc'anzi illustrata permette di stabilire un modello locale.

Hammersley e Clifford hanno dimostrato, in un teorema molto celebrato circolato per anni in forma privata (Hammersley e Clifford, 1971) e successivamente pubblicato da altri autori (Preston, 1973; Grimmett, 1973 e Besag, 1974), che esiste un unico GRF per ogni MRF e viceversa purchè il campo aleatorio di Gibbs sia definito in termini di *cliques* in un sistema di vicinato (equivalenza Markov-Gibbs). Quindi ogni MRF in generale non è un GRF così come ogni GRF non costituisce un MRF.

Definire un campo aleatorio markoviano specificando le proprietà locali senza riguardo ai *cliques* od al sistema di vicinato può condurre a modelli la cui distribuzione congiunta non esiste. Modelli così costruiti potrebbero essere inconsistenti e di discutibile utilità nelle applicazioni pratiche. Allo stesso modo costruire un campo aleatorio di Gibbs mediante una distribuzione di probabilità congiunta che ignora la struttura di *clique* può portare a modelli la cui distribuzione locale non mostra una ragionevole dipendenza spaziale (Dubes e Jain, 1989). Un siffatto modello potrebbe essere difficile da simulare e le usuali procedure per la stima dei parametri potrebbero non essere applicabili.

Per le ragioni descritte si limiterà l'attenzione ai campi aleatori costruiti con riferimento ad una struttura di *clique* che possano essere espressi sia mediante un GRF che tramite un MRF. Volendo analizzare con maggiore dettaglio le affermazioni di cui sopra, si considerino le condizioni markoviane espresse nelle [6.2] e [6.2bis].

L'approccio condizionato alla modellistica spaziale consiste nell'ipotizzare un modello per tali probabilità condizionate e poi derivare a posteriori la struttura di probabilità congiunta che ne consegue (Arbia, 1993b). Tuttavia, al fine di garantire che la struttura di probabilità congiunta del processo spaziale esista e sia unica, è necessario imporre alcune restrizioni circa la forma funzionale delle probabilità condizionate. Tali restrizioni sono identificate dal teorema di Hammersley e Clifford, teorema che ha solo di recente ricevuto considerevole attenzione ma che costituisce lo strumento essenziale alla costruzione di schemi spaziali validi attraverso l'approccio delle probabilità condizionate.

Si definisca pertanto la grandezza⁴:

$$Q(\mathbf{x}) = \ln\{P(\mathbf{x})/P(\mathbf{0})\}, \quad \forall \mathbf{x} \in \Omega$$

dove, come già detto, $P(\mathbf{x})$ rappresenta la densità congiunta di una singola realizzazione di un processo \mathbf{X} e $P(\mathbf{0})$ la densità congiunta di una realizzazione del processo che

⁴ Intuitivamente $Q(\mathbf{x})$ misura quanto più è verosimile la configurazione riportata al numeratore rispetto alla configurazione di soli zero.

consiste di soli zeri per la quale vale la condizione di positività ($P(\mathbf{0}) > 0$).

Il teorema di Hammersley e Clifford identifica la forma più generale di $Q(\mathbf{x})$ che assicuri una struttura di probabilità congiunta valida e coerente per il sistema considerato (Besag, 1974, pag.197). Tale forma funzionale è la seguente:

$$Q(\mathbf{x}) = \sum_i x_i G_i(x_i) + \sum_i \sum_{j>i} x_i x_j G_{ij}(x_i x_j) + \sum_i \sum_{j>i} \sum_{k>j>i} x_i x_j x_k G_{ijk}(x_i x_j x_k) + \dots + x_1 x_2 \dots x_n G_{12\dots n}(x_1 x_2 \dots x_n) \quad [6.3]$$

dove le sommatorie sono da intendersi da 1 a n e le funzioni G sono non nulle solo se gli indici costituiscono dei *cliques*. Dalla relazione [6.3] possono essere derivate distribuzioni di probabilità condizionate valide. Infatti:

$$\begin{aligned} \exp[Q(\mathbf{x}) - Q(\mathbf{x}_i)] &= \exp\{\ln[P(\mathbf{x})/P(\mathbf{0})] - \ln[P(\mathbf{x}_i)/P(\mathbf{0})]\} = \frac{P(\mathbf{x})}{P(\mathbf{x}_i)} = \\ &= \frac{P(x_i | x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n) P(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)}{P(0 | x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n) P(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)} = \frac{P(x_i | x_j, j \in C\{i\})}{P(0 | x_j, j \in C\{i\})} \end{aligned} \quad [6.4]$$

dove $\mathbf{x}_i = (x_1, \dots, x_{i-1}, 0, x_{i+1}, \dots, x_n)$.

L'importanza del risultato [6.3] non discende solo dal garantire l'esistenza di una funzione di probabilità congiunta date certe assunzioni sulle probabilità condizionate, ma principalmente dal suggerire i vincoli da imporre a tali distribuzioni al fine di ottenere delle rappresentazioni valide. È interessante notare come le densità congiunte possano essere sempre scomposte nelle interazioni tra coppie di variabili aleatorie, tra terne di variabili aleatorie, etc. Tale proprietà riveste una notevole importanza in quanto, essendo possibile eliminare i termini relativi ad effetti dovuti a dipendenze superiori al primo ordine, nella pratica lo studio dei MRF è limitato ai cosiddetti auto-modelli (Besag, 1974), ossia ai processi per i quali la [6.3] si riduce a:

$$Q(\mathbf{x}) = \sum_i x_i G_i(x_i) + \sum_i \sum_{j>i} \beta_{ij} x_i x_j. \quad [6.5]$$

Quindi

$$\exp[Q(\mathbf{x}) - Q(\mathbf{x}_i)] = \frac{P(x_i | \bullet)}{P(0 | \bullet)} = \exp\left[x_i \left\{ G_i(x_i) + \sum_{j=1}^n \beta_{ij} x_j \right\}\right] \quad [6.6]$$

dove i parametri β_{ij} rappresentano l'interazione tra coppie di variabili aleatorie e sono tali che $\beta_{ij}=0$ a meno che $j \in C\{i\}$ e $\beta_{ij}=\beta_{ji}$. Essendo tali parametri l'elemento chiave della forma semplificata di processo markoviano multidimensionale presentata nella [6.6], le restrizioni imposte dal teorema di Hammersley e Clifford riguardano essenzialmente i valori assumibili dai parametri β_{ij} .

Dalle relazioni [6.4] e [6.5] è possibile quindi ricavare valide formalizzazioni probabilistiche per le tipologie di dati spaziali più comuni (Besag, 1974). Tali modelli possono essere classificati ulteriormente rispetto alla forma funzionale che $P(\cdot)$ assume così da imporre l'introduzione di termini quali auto-logistico, auto-binomiale e auto-normale per identificare specifici schemi spaziali. Detti schemi verranno esaminati in dettaglio nel corso del prossimo paragrafo.

7. Alcuni modelli di interazione spaziale

7.1 Auto-modelli per variabili qualitative

Il caso più semplice dal quale si intende partire è la situazione in cui la variabile aleatoria X sia di tipo dicotomico (presenza/assenza di una certa caratteristica). Tale caso è quello che si verifica, per fare un parallelo con l'analisi delle immagini digitali, quando si dispone di immagini in bianco e nero senza toni di grigio intermedi.

Le probabilità per cui è necessario derivare un modello valido sono del tipo:

$$p_i = P(X_i = 1 | X_j = x_j, j \neq i, j \in C\{i\}) \quad [7.1]$$

e

$$q_i = P(X_i = 0 | X_j = x_j, j \neq i, j \in C\{i\}) \quad [7.2]$$

con $p_i + q_i = 1$.

Le relazioni [7.1] e [7.2] rappresentano rispettivamente le probabilità di presenza e assenza di un fenomeno nel pixel i -mo condizionatamente alla presenza/assenza del medesimo fenomeno nelle celle che costituiscono il vicinato di i . Se in particolare i soli parametri non nulli sono quelli associati con i *cliques* composti di singoli pixels e di coppie di pixels, si ottiene il modello auto-logistico per il quale vale la relazione (Besag, 1974):

$$Q(\mathbf{x}) = \sum_{i=1}^n \alpha_i x_i + \sum_i \sum_{j>i} \beta_{ij} x_i x_j. \quad [7.3]$$

Dalle espressioni [6.4] e [6.6] del paragrafo precedente nonchè dalla [7.3] segue che:

$$\begin{aligned}
& \frac{P(X_i = x_i | X_j = x_j, j \neq i, j \in C\{i\})}{P(X_i = 0 | X_j = x_j, j \neq i, j \in C\{i\})} = \exp[\mathcal{Q}(\mathbf{x}) - \mathcal{Q}(\mathbf{x}_i)] = \\
& = \exp \left[\sum_{i=1}^n \alpha_i x_i + \sum_i \sum_{j>i} \beta_{ij} x_i x_j - \left(\sum_{j=1}^{i-1} \alpha_j x_j + \alpha_i 0 + \sum_{j=i+1}^n \alpha_j x_j + \sum_j \sum_{t>j; j, t \neq i} \beta_{jt} x_j x_t + \sum_{j=1}^n \beta_{ij} 0 x_j \right) \right] = \\
& = \exp \left[\alpha_i x_i + \sum_{j=1}^n \beta_{ij} x_i x_j \right] = \exp \left[x_i \left(\alpha_i + \sum_{j=1}^n \beta_{ij} x_j \right) \right] \tag{7.4}
\end{aligned}$$

dove $\beta_{ij}=0$ a meno che $j \in C\{i\}$ e $\beta_{ij}=\beta_{ji}$.

Dato che i valori assumibili da X sono 1 oppure 0 si ottengono le relazioni seguenti:

$$p_i = P(X_i = 1 | X_j = x_j, j \neq i, j \in C\{i\}) = \frac{\exp \left[\alpha_i + \sum_{j=1}^n \beta_{ij} x_j \right]}{1 + \exp \left[\alpha_i + \sum_{j=1}^n \beta_{ij} x_j \right]} \tag{7.5}$$

e

$$q_i = P(X_i = 0 | X_j = x_j, j \neq i, j \in C\{i\}) = \frac{1}{1 + \exp \left[\alpha_i + \sum_{j=1}^n \beta_{ij} x_j \right]} \tag{7.6}$$

che rispettano le condizioni di coerenza in quanto p_i+q_i e che costituiscono l'analogo del modello logistico classico (Cox, 1970, cap.1) eccezion fatta per le variabili esplicative che sono esse stesse realizzazioni del processo. Le [7.5] e [7.6] mostrano come viene modificata la nostra aspettativa di *successo* e di *insuccesso* nella cella i -ma sulla base dei risultati ottenuti nei *pixels* ad essa adiacenti.

I parametri β_{ij} rappresentano l'interazione tra *cliques* di dimensione due, con valore critico per $\beta = \sinh^{-1} \cong 0.88$ oltre il quale (asintoticamente) le realizzazioni del processo saranno dominate da una sola classe, mentre gli α_i sono dei parametri locali

utili per tenere conto nel modello della presenza di trend spaziali⁵ o dell'influenza esercitata da altre variabili esplicative⁶. Nel caso di interazione spaziale nulla si avrà ovviamente $\beta_{ij}=0$ per ogni indice i e j e, così:

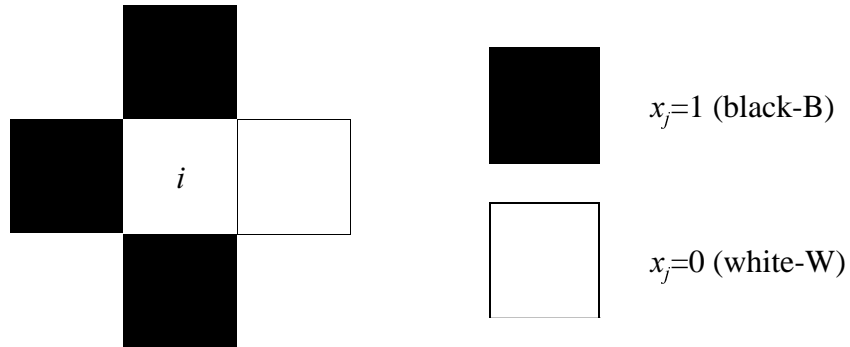
$$P(X_i = x_i | X_j = x_j, j \neq i, j \in C\{i\}) = \frac{\exp(\alpha_i)}{1 + \exp(\alpha_i)}.$$

Tale probabilità non dipende dai vicini della cella i -ma, ma solo dalle condizioni locali nella cella stessa. In caso contrario il parametro β_{ij} regola il livello di interazione spaziale ossia l'intensità con la quale le condizioni del contesto locale modificano le aspettative di successo (Arbia, 1993b). In altre parole, i parametri β_{ij} misurano l'intensità dell'interazione tra gli agenti, per cui al crescere dei parametri aumenta l'influenza che gli agenti vicini esercitano sull'agente i -mo.

È utile, a questo punto, mostrare i concetti poc'anzi illustrati con un semplice esempio (Arbia e Espa, 1996). Si supponga, per semplicità di esposizione, che $\alpha_i = \alpha = 1 \ \forall i$ e che inoltre $\beta_{ij} = \beta w_{ij}$, dove β è una costante e $w_{ij} \in W$ rappresenta l'elemento generico di una matrice di connettività binaria. In termini più formali $w_{ij} \in W$ è tale che:

$$w_{ij} = \begin{cases} 1 & \text{se } j \in C\{i\} \\ 0 & \text{altrimenti} \end{cases}.$$

Si consideri inoltre la seguente configurazione spaziale:



⁵ Al riguardo è sufficiente porre, ad esempio, $\alpha_i = f(\text{lat}_i, \text{long}_i)$ dove $(\text{lat}_i, \text{long}_i)$ sono le coordinate geografiche dell' i -ma cella e specificare per f una forma lineare, quadratica, etc.

⁶ Tale influenza può essere catturata dal modello ponendo $\alpha_i = g(Z_i, U_i, \dots)$ dove (Z_i, U_i, \dots) sono variabili esplicative correlate con la variabile oggetto di indagine.

La probabilità che la cella i -ma presenti un valore pari ad 1 sarà, in base alla [7.5] ed alle assunzioni formulate:

$$p_i = \frac{\exp\left(1 + \sum_{j \in C\{i\}} \beta x_j\right)}{\left[1 + \exp\left(1 + \sum_{j \in C\{i\}} \beta x_j\right)\right]}.$$

Nel caso di assenza di interazione spaziale tra le località, la probabilità di successo è semplicemente

$$p_i = e^1 / (1 + e^1) = 0.73$$

che rappresenta la probabilità che *a priori* (prima cioè di avere cognizione di ciò che accade nel contesto spaziale della località i) si attribuisce all'evento $x_i=1$.

Nel nostro esempio tre vicini della località i -ma presentano valore 1 ed un vicino valore 0. Pertanto si avrà

$$p_i = \frac{\exp(1 + 3\beta)}{[1 + \exp(1 + 3\beta)]}.$$

Nel caso di moderata interazione spaziale, ad esempio $\beta=1/3$, si avrà $p_i=0.88$ che è superiore al valore atteso non condizionato (0.73). Nel caso di una elevata interazione spaziale (ad esempio $\beta=1$) si avrà invece che la probabilità di successo condizionata *esploderà* ad un valore molto prossimo ad 1 ($p_i=0.98$). Infine, nel caso di interazione spaziale negativa (ad esempio $\beta=-1/3$) si avrà $p_i=0.5$, ossia la probabilità a priori (0.73) viene ritoccata verso il basso.

Il comportamento del modello risulta facilmente comprensibile studiando l'effetto del parametro β sulla probabilità condizionata p_i al variare della configurazione al contorno della cella i (si veda al riguardo la Figura 7.1).

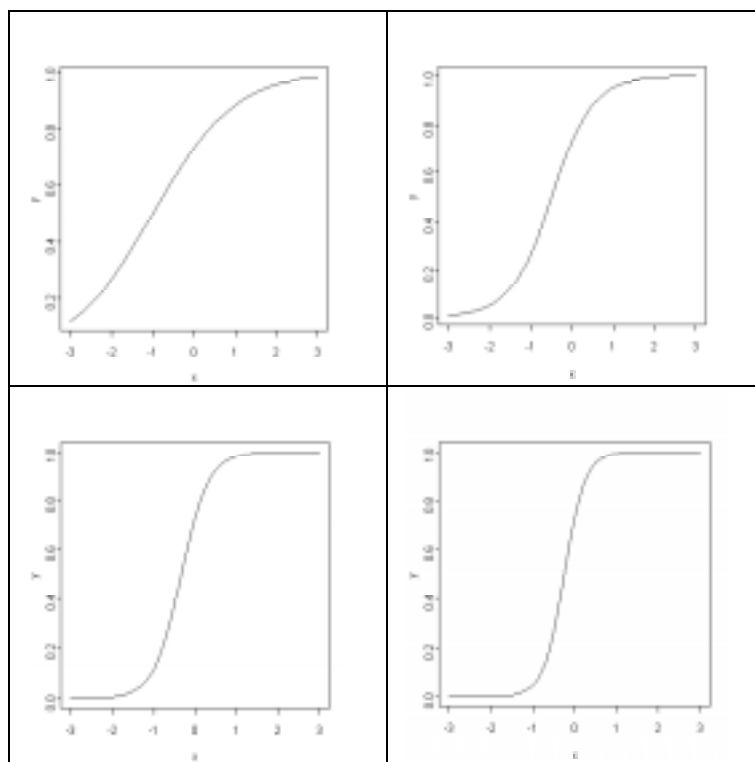


Fig.7.1 Andamento della probabilità condizionata p_i di presenza di una caratteristica nella località i al variare del parametro β (ascissa) ed al variare della configurazione della medesima caratteristica nelle celle circostanti ad i . Da sinistra verso destra e dall'alto al basso il vicinato è così costituito: 1B e 3W; 2B e 2W; 3B e 1W; 4B e 0W.

Il modello auto-logistico, che corrisponde in fisica al modello del ferromagnetismo di Ising, è stato ampiamente utilizzato in ambito economico adattandolo a contesti assai diversi tra loro.

Föllmer(1974) applica il modello auto-logistico all'analisi dell'equilibrio generale per rimuovere due ipotesi fondamentali su cui si basava la teoria microeconomica standard: i) le caratteristiche di un agente economico, e in particolare le sue preferenze, sono dati fissi iniziali; ii) le caratteristiche di un agente sono indipendenti dall'ambiente in cui l'agente opera. Nel modello di Föllmer gli agenti sono caratterizzati dalle loro preferenze e dotazioni iniziali che ne definiscono lo "stato". Il comportamento di ogni agente (cioè quale stato l'agente decide di assumere), e in particolare la scelta delle sue preferenze, è influenzato dall'ambiente in cui si trova, ovvero dal comportamento di un numero finito di altri agenti, l'insieme dei suoi "vicini rilevanti". Lo stato di ogni agente è così una variabile stocastica che dipende dallo stato in cui si trovano i vicini dell'agente. Più precisamente, la struttura di interazione tra gli agenti è definita come un campo aleatorio di Markov e Föllmer costruisce una *Ising economy* che altro non è che la trasposizione del modello di Ising in tale contesto. Il modello di Ising gli permette di analizzare la robustezza dei risultati di equilibrio generale ovvero se l'introduzione della

stocasticità delle preferenze e dell'interdipendenza degli agenti intacca seriamente l'esistenza dei "price equilibria".

Inserendosi nel filone di ricerca riguardante i processi di diffusione tecnologica precedentemente esplorato da Arthur, Ermoliev e Kaniovski (1987), Arthur (1989), David (1988,1992) e da David e Foray (1992), Dalle (1994) utilizza un modello auto-logistico per modellare un processo di diffusione di due tecnologie in un'economia caratterizzata da informazione imperfetta. L'ipotesi di informazione imperfetta implica che l'idea di agente rappresentativo diventi un concetto irrilevante (Kirman, 1992) e che al suo posto debba essere presa in considerazione una molteplicità di agenti eterogenei, tra loro interdipendenti. Sempre dall'ipotesi di informazione imperfetta Dalle fa discendere l'esistenza di due tipi di informazioni che sono disponibili ad ogni agente: i) l'informazione interna che è legata alla storia passata e alla traiettoria tecnologica propria di ogni impresa; e ii) l'informazione esterna che a sua volta può essere locale (se l'impresa la riceve dai suoi "vicini rilevanti") o globale (se l'informazione è distribuita uniformemente nell'economia da una fonte centralizzata). Utilizzando un modello auto-logistico e avvalendosi del teorema di Hammersley-Clifford, Dalle mostra come si ottenga diversificazione nell'adozione delle tecnologie se le imprese ricevono solo informazione interna e informazione esterna locale. Se la componente interna dell'informazione prevale sulla componente esterna locale (ovvero quando l'intensità dell'interazione è al di sopra di una determinata soglia) si ottiene una configurazione in cui entrambe le tecnologie coesistono in proporzioni più o meno identiche. Se la componente esterna locale prevale su quella interna si ottiene una configurazione "strutturata", in cui vi è una chiara predominanza di una tecnologia mentre l'altra sopravvive in nicchie. Nel momento in cui l'informazione esterna globale viene introdotta la standardizzazione viene (quasi) raggiunta: l'industria viene intrappolata da effetti di lock-in e le prospettive di un'evoluzione sono piuttosto rare.

Durlauf (1993, 1994) utilizza il modello auto-logistico per prendere in considerazione il ruolo delle complementarità e dei mercati incompleti in un modello di crescita economica. L'autore analizza l'evoluzione di un'economia composta da un numero finito di industrie caratterizzate da funzioni di produzione aggregate non-convesse e da complementarità tecnologiche che creano legami intertemporali tra le funzioni di produzione delle industrie. Le complementarità sono date dalle interazioni tra le industrie, solitamente da quelle tecnologicamente vicine. Durlauf definisce un equilibrio aggregato come una misura di probabilità congiunta che caratterizza molti agenti. La specificazione microeconomica del modello produce un insieme di misure di probabilità che descrivono le azioni di ogni agente individuale condizionate alla storia dell'economia. Un equilibrio esiste quando queste misure di probabilità condizionate sono compatibili con la misura di probabilità congiunta sopra tutti gli agenti e in tutti i periodi. Le industrie scelgono un livello di produzione in base agli spillovers che ricevono dalle industrie tecnologicamente vicine, poichè non esiste un meccanismo di mercato dove le industrie possono interagire. Il comportamento dinamico di un'economia definita in tale modo è fortemente influenzato dall'intensità delle interazioni. Al crescere del numero di industrie "vicine" e dell'influenza di queste sulle

decisioni di ogni altra, l'economia sperimenta una transizione di fase dall'ergodicità alla non ergodicità: in presenza di forti complementarità tra industrie emergono forti fluttuazioni nell'output aggregato e l'economia nel suo complesso subisce fasi cicliche di boom e crisi.

Haining (1984, 1990) sempre utilizzando questo modello cerca di spiegare la distribuzione geografica del prezzo della benzina negli anni 1982-83 in alcuni punti della città di Sheffield (GB). La scelta di un modello di interazione spaziale ben si presta all'analisi del fenomeno oggetto di studio. Infatti è noto che nel Regno Unito il prezzo della benzina è lasciato libero di variare da un distributore all'altro, come del resto accade in Italia dal 1991. Tuttavia, contrariamente a quanto accade nel nostro paese, nel Regno Unito si assiste ad elevate differenze tra i vari dettaglianti. In particolare il periodo di rilevazione considerato da Haining è stato caratterizzato da una elevata variabilità dei prezzi della benzina a causa del crollo generalizzato registratosi a seguito del ritrovamento del greggio nella Manica e della successiva ripresa inflazionistica. Per di più in tale mercato è ragionevole ipotizzare un'interazione spaziale nella fissazione dei prezzi in quanto il bene è perfettamente omogeneo senza variazioni di qualità da un distributore all'altro ed, inoltre, vi è una piena visibilità dei prezzi i quali, di norma, appaiono scritti su tabelloni luminosi posti al di sopra del distributore stesso.

Lo schema auto-logistico costruito da Haining è un modello di competizione pura che enfatizza il comportamento di quei dettaglianti che, in periodi di caduta dei prezzi, adottano politiche di prezzo aggressive volte all'aumento delle quote di mercato (Schendel e Balestra, 1969). Tale modello descrive la distribuzione di equilibrio dei prezzi dicotomizzati nel modo seguente:

$$X_i = \begin{cases} 1 & \text{se l'i - mo dettagliante fissa un prezzo} \in [s_1, s_2] \\ 0 & \text{altrimenti} \end{cases}$$

in termini di probabilità condizionate.

Una proprietà interessante del modello auto-logistico, sulla quale si tornerà nella sezione 10.2, è che la varianza di $\sum X_i$ tende a decrescere per valori crescenti del parametro di interazione spaziale β (*phase transition*). Tradotto in termini economici ciò significa che durante una guerra dei prezzi, quando la competizione tra siti diventa intensa, la varianza totale dei prezzi può decrescere.

La situazione più generale per ciò che concerne le applicazioni pratiche (Haining, 1990) si presenta quando si osserva una variabile qualitativa sconnessa che può assumere un numero finito C di valori⁷. Si ponga allora:

⁷ Preme sottolineare che la variabile qualitativa presa in considerazione non deve essere necessariamente dicotomica, ma può assumere un numero finito a piacere di valori, contrariamente a

$$Q(\mathbf{x}) = \sum_{k=1}^C \alpha_k n_k - \sum_k \sum_{l>k} \beta_{kl} n_{kl} \quad [7.7]$$

dove n_k è il numero di celle che presentano la classe k mentre n_{kl} è il numero di coppie di località (*cliques*) che presentano rispettivamente le classi k e l . Il significato dei parametri α_k e β_{kl} è analogo a quello illustrato per il modello auto-logistico.

Dalle espressioni [6.4] e [6.6] nonché dalla relazione [7.7] discende con passaggi analoghi a quelli che hanno condotto alla [7.4] che

$$\frac{P(X_i = k | X_j = x_j, j \neq i, j \in C\{i\})}{P(X_i = 0 | X_j = x_j, j \neq i, j \in C\{i\})} = \exp \left[\alpha_k + \sum_{l=1}^C \beta_{kl} u_i(l) \right]$$

dove $u_i(l)$ è il numero di vicini del pixel i che presentano la classe l .

Se $\beta_{kl} = \beta$ e nell'ipotesi che le classi siano tra loro scambiabili così da poter ritenere il parametro α_k ridondante, si ottiene il modello introdotto da Strauss (1977):

$$P(X_i = k | X_j = x_j, j \neq i, j \in C\{i\}) = \frac{\exp[\beta u_i(k)]}{\sum_{l=1}^C \exp[\beta u_i(l)]}. \quad [7.8]$$

Una variante del modello di Strauss si ha nel caso in cui le classi posseggano un ordinamento naturale.

In tale caso si può supporre che la variabile aleatoria X_i abbia una distribuzione di probabilità condizionata di tipo binomiale associata ad una probabilità di successo θ_i dipendente dai valori riscontrati nel vicinato del pixel i . Sotto le assunzioni poc'anzi formulate la variabile aleatoria X_i assume la classe k ($k=1,2,...,C$) con probabilità pari a:

$$P(X_i = k | X_j = x_j, j \neq i, j \in C\{i\}) = \binom{C}{k} \theta_i^k (1 - \theta_i)^{C-k}.$$

In tal caso, a seguito del teorema di Hammersley e Clifford, si avrà

$$\ln \left\{ \frac{\theta_i}{1 - \theta_i} \right\} = \alpha_i + \sum_j \beta_{ij} x_j$$

quanto viene affermato in alcune sedi. Il fatto che solitamente le applicazioni dei modelli basati sui MRF vengano fatte per semplicità considerando solo due possibili valori non può essere individuato come un limite di tali modelli (Hors e Lordon, 1995, p.17).

e quindi:

$$\theta_i = \frac{\exp\left[\alpha_i + \sum_j \beta_{ij} x_j\right]}{1 + \exp\left[\alpha_i + \sum_j \beta_{ij} x_j\right]}.$$

Il processo così definito è detto auto-binomiale (Besag, 1974).

Nel caso particolare in cui $C=2$, caso in cui non esistono strutturalmente problemi di ordinamento, sia il modello di Strauss che quello auto-binomiale si riducono al modello auto-logistico.

7.2 Lo schema gaussiano

In numerose applicazioni pratiche (si vedano, per una rassegna, Besag, 1986; Ripley, 1988 e Mardia, 1990) si può supporre che la variabile casuale \mathbf{X} sia di tipo assolutamente continuo e che abbia distribuzione condizionata gaussiana:

$$P\left(X_i = x_i \mid X_j = x_j, j \neq i, j \in C\{i\}\right) = (2\pi\sigma_i^2)^{-1/2} \exp\left[-\frac{1}{2}\sigma_i^{-2} \left\{X_i - \mu_i - \sum_{j \neq i} \beta_{ij} (X_j - \mu_j)\right\}^2\right] \quad [7.9]$$

dove $P(\cdot)$ indica la densità di probabilità condizionata.

Se l'espressione [7.9] è vera per ogni cella i si può, in modo equivalente, usare la notazione seguente (Morrison, 1967):

$$\begin{aligned} E\left(X_i \mid X_j = x_j, j \in C\{i\}\right) &= \mu_i + \sum_{j \neq i} \beta_{ij} (X_j - \mu_j) \\ \text{Var}\left(X_i \mid X_j = x_j, j \in C\{i\}\right) &= \sigma_i^2 \end{aligned}$$

dove, al solito, $\beta_{ij} \neq 0$ solo se i *pixel* i e j costituiscono un clique.

La densità congiunta delle \mathbf{X} è (Besag, 1974):

$$\mathbf{X} \sim \text{MVN}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \text{ con } \boldsymbol{\Sigma} = \boldsymbol{\sigma}^2 (\mathbf{I} - \mathbf{B})^{-1}$$

dove \mathbf{B} è una matrice di generico elemento β_{ij} .

Questo modello, molto utilizzato nella sua forma semplificata con il solo parametro ρ dell'espressione $\mathbf{B} = \rho\mathbf{C}$, è stato originariamente denominato auto-normale, ma risulta ormai meglio noto come autoregressivo condizionato (CAR) (Ripley, 1981; Benedetti, 1994) nella contrapposizione con il modello autoregressivo simultaneo (SAR) espresso tipicamente dalla relazione

$$X_i = \mu_i + \sum_{j \neq i} \beta_{ij} (X_j - \mu_j) + \varepsilon_i$$

dove $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ sono delle variabili casuali gaussiane indipendenti di media 0 e varianza σ^2 .

8. Altri modelli di interazione spaziale

I modelli analizzati nella sezione precedente sono molto convincenti da un punto di vista logico. Il loro utilizzo è giustificato dalla definizione stessa delle distribuzioni di probabilità condizionate delle singole variabili aleatorie in ciascuna cella nonché dalle proprietà *fisiche* ereditate dalla connessione esistente con i processi di Gibbs. In conseguenza di ciò tali modelli sono quelli maggiormente usati nella letteratura, sia metodologica che applicata, più recente. Tuttavia, nel caso particolare dei processi gaussiani, sono stati proposti negli anni alcuni approcci alternativi che verranno qui di seguito riportati, in maniera molto sintetica, per ragioni di completezza.

Si consideri il caso in cui \mathbf{X} si distribuisce secondo una normale multivariata di vettore di medie μ e di matrice di varianze e covarianze \mathbf{V} tale che $\text{Cov}(X_i X_j) = \sigma_{ij}$. Nel caso di modelli di processi aleatori gaussiani stazionari (Arbia, 1989) l'unico problema è costituito dal trovare il modo di parametrizzare gli elementi della matrice \mathbf{V} in termini di un numero minimo di parametri. Tali modelli possono essere ottenuti specificando direttamente una forma plausibile per \mathbf{V} oppure specificando uno schema di interazione tra variabili aleatorie dal quale \mathbf{V} possa essere derivata (Haining, 1990, pag. 75).

Sono riconducibili al primo modo di procedere modelli come quello utilizzato da Mardia e Watkins (1989) i quali hanno considerato la funzione che modella la variazione locale a piccoli livelli di scala:

$$\sigma_{ij}(h) = \begin{cases} \sigma^2 (1 - \alpha^{-1}|h|)^4 & |h| < \alpha \\ 0 & |h| > \alpha \end{cases}$$

dove h rappresenta la distanza tra i siti i e j . Tale approccio è noto come la rappresentazione diretta del processo (si veda Mardia, 1990) o modello DR (Direct Representation).

Un modo di modellizzare gli elementi della matrice V che fa invece riferimento al secondo degli approcci discussi sopra è quello di fare ricorso ad un modello SAR la cui espressione analitica è stata riportata nella precedente sezione: Tale modello è stato proposto da Whittle nel 1954 e può senza dubbio essere considerato il padre dei modelli di processi aleatori spaziali. Lo schema SAR, come già detto, è definito dall'equazione autoregressiva

$$X_i - \mu_i - \sum_{j \neq i} \beta_{ij} (X_j - \mu_j) = \varepsilon_i$$

con $\{\varepsilon_i\}$ un insieme di variabili aleatorie di tipo rumore bianco spaziale e β_{ij} un termine di interazione spaziale tale che $\beta_{ij} \in \{B\}$.

In tale caso la matrice di varianze e covarianze sarà definita da (Haining, 1990):

$$V = \sigma^2 [(I - B)'(I - B)]^{-1}$$

con V soggetta ad una serie di restrizioni che assicurino l'invertibilità del termine $[(I - B)'(I - B)]$.

Riscrivendo, per evitare confusione nelle notazioni, la matrice di varianze e covarianze del modello CAR mostrato nel paragrafo precedente come $\Sigma = \sigma^2 (I - C)^{-1}$ è facile verificare come, semplicemente ponendo $C = B + B' - B'B$ ogni processo SAR sia anche CAR (Ripley, 1981, pag.90).

Infatti:

$$\frac{\Sigma}{\sigma^2} = (I - C)^{-1} = (I - B - B' + B'B)^{-1} = [(I - B) - B'(I - B)]^{-1} = [(I - B)'(I - B)]^{-1} = \frac{V}{\sigma^2}.$$

Non vale il viceversa o, per lo meno, vale in modo piuttosto intuitivo. Sia $B = (I - L)^{-1}$ dove LL' è la scomposizione di Cholesky della matrice $(I - C)$. Allora:

$$\begin{aligned} \frac{V}{\sigma^2} &= (I - B)^{-1} (I - B')^{-1} = (I - I + L)^{-1} (I - (I - L)')^{-1} = \\ &= (L')^{-1} (I - I + L)^{-1} = (L')^{-1} L^{-1} = (LL')^{-1} = (I - C)^{-1} = \frac{\Sigma}{\sigma^2}. \end{aligned}$$

Ma, non essendo la scomposizione di Cholesky l'unica scomposizione possibile tale che $LL' = (I - C)$ ne consegue che più matrici B implicano una sola matrice C .

Un modo alternativo di modellizzare gli elementi della matrice V è attraverso lo schema a media mobile spaziale (SMA) introdotto da Haining (1978):

$$X_i = \mu_i + \sum_{j \neq i} m_{ij} e_j + e_i \quad [8.1]$$

con e_i di nuovo un processo rumore bianco spaziale ed $\{m_{ij}\} \in \mathbf{M}$ un insieme di pesi. Dalla [8.1] discende che

$$\mathbf{V} = \sigma^2 [(\mathbf{I} + \mathbf{M})(\mathbf{I} + \mathbf{M})'].$$

Tale modello rappresenta l'analogo spaziale dei modelli a media mobile temporale. Anche in questo caso è facile verificare come semplicemente ponendo $\mathbf{C} = \mathbf{M}' \mathbf{M} + 2\mathbf{M}$, ogni processo SMA sia anche CAR.

9. Stima dei modelli

Negli anni '70, quando i modelli fin qui esposti vennero proposti, esistevano ancora grossi limiti computazionali all'uso del metodo della massima verosimiglianza per la stima dei parametri di interazione e del vettore delle medie, problemi che oggi sono in larga parte superati. Inoltre tale metodo di stima dei parametri garantisce un elevato numero di auspicabili proprietà asintotiche delle statistiche che da esso scaturiscono (Benedetti, 1994).

Per il modello CAR, ad esempio, la funzione di densità congiunta delle \mathbf{X} è

$$L = \frac{|\mathbf{I} - \mathbf{C}|^{1/2}}{(2\pi\sigma^2)^{n/2}} \exp \left[-\frac{1}{2} \sigma^{-2} (\mathbf{X} - \boldsymbol{\mu})' (\mathbf{I} - \mathbf{C}) (\mathbf{X} - \boldsymbol{\mu}) \right].$$

Pertanto, meno due volte la log-verosimiglianza sarà:

$$-2\ln(L) = n\ln(2\pi\sigma^2) - \ln|\mathbf{I} - \mathbf{C}| + \frac{(\mathbf{X} - \boldsymbol{\mu})' (\mathbf{I} - \mathbf{C}) (\mathbf{X} - \boldsymbol{\mu})}{\sigma^2}$$

che può essere facilmente massimizzata rispetto a σ^2 ottenendo:

$$\hat{\sigma}^2 = n^{-1} (\mathbf{X} - \boldsymbol{\mu})' (\mathbf{I} - \mathbf{C}) (\mathbf{X} - \boldsymbol{\mu}).$$

I valori della matrice \mathbf{C} possono essere quindi ricavati in modo da massimizzare l'espressione

$$\ln \left(\hat{\sigma}^2 \right) - n^{-1} \ln |\mathbf{I} - \mathbf{C}|.$$

Fino a qualche tempo fa il calcolo del determinante $|\mathbf{I} - \mathbf{C}|$ era considerato proibitivo già per schemi spaziali costituiti da poche unità areali. Infatti la matrice \mathbf{C} in molte applicazioni è una matrice sparsa e quasi-singolare (Arbia, 1986). Tuttavia è ormai

possibile risolvere il problema con approssimazioni numeriche in modo abbastanza rapido anche nel caso di schemi spaziali di dimensione elevata.

Nel caso in cui si possa scrivere $\mathbf{C} = \rho \mathbf{W}$ con \mathbf{W} una matrice di connettività binaria e τ_i i suoi autovalori, Ord (1975) ha suggerito di usare la scomposizione:

$$|\mathbf{I} - \mathbf{C}| = |\mathbf{I} - \rho \mathbf{W}| = \prod_{i=1}^n (1 - \rho \tau_i).$$

Un'ulteriore alternativa al metodo della massima verosimiglianza è rappresentata dal metodo della pseudo-verosimiglianza introdotto da Besag (1975 e 1977). Il metodo consiste nel definire la pseudo-verosimiglianza semplicemente come il prodotto delle varie densità condizionate

$$PL = \prod_{i=1}^n P(X_i | X_j, j \neq i)$$

e nel massimizzarla come se si trattasse di una vera e propria verosimiglianza. Per un modello CAR, ad esempio, si ha:

$$\ln(PL) = -\frac{n}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \|(\mathbf{I} - \mathbf{C})(\mathbf{X} - \mu)\|^2 = -\frac{n}{2} \ln(2\pi\sigma^2) - \frac{1}{2} \sigma^{-2} \sum_{i=1}^n [(\mathbf{I} - \mathbf{C})(\mathbf{X} - \mu)]_i^2.$$

A parte una dimostrazione della consistenza di tale metodo (Geman e Graffigne, 1986), l'elemento maggiormente a favore della pseudo-verosimiglianza sembra essere rappresentato dalla circostanza molto pragmatica che esso fornisce buoni risultati in casi reali (Ripley, 1990).

10. Simulazione di processi stocastici bidimensionali

La simulazione, o metodo di Monte Carlo, è quella tecnica che riproduce un fenomeno, date le condizioni in cui esso si sviluppa, attraverso un modello matematico. Se tale modello incorpora componenti aleatorie si parla di simulazione stocastica (Ripley, 1987).

La crescente rapidità di calcolo rapportata ai costi decrescenti dei computers e l'indubbia utilità che riveste nella soluzione di problemi statistici pratici e teorici hanno comportato negli ultimi anni un repentino sviluppo della simulazione stocastica, il cui utilizzo si è rivelato di grande utilità per la sperimentazione nell'ambito delle scienze matematiche e come metodo inferenziale. In effetti la simulazione sta assumendo sempre più i connotati di una branca della statistica inferenziale. Il vantaggio principale che deriva dal suo uso è la relativa facilità, in relazione ai metodi inferenziali classici, con cui si possono risolvere problemi pratici quali:

- i) la derivazione di distribuzioni di probabilità empiriche di statistiche complesse per le quali non sono disponibili risultati analitici;
- ii) lo studio del comportamento di procedure inferenziali;
- iii) la messa a punto di tests di verifica delle ipotesi basati sul confronto fra dati effettivi e dati provenienti da un modello simulato;
- iv) l'analisi della variabilità e/o della distorsione di uno stimatore attraverso realizzazioni di distribuzioni empiriche, come accade nei metodi di ricampionamento come il Bootstrap.

Nel caso dei processi aleatori bidimensionali i metodi di simulazione di cui è necessario verificare l'affidabilità e l'efficienza sono essenzialmente quelli relativi alla generazione dei modelli discussi nelle precedenti sezioni 7.1 e 7.2. A tale fine i due sottoparagrafi seguenti sono dedicati all'illustrazione di tecniche per la simulazione di una normale multivariata e di un campo aleatorio markoviano a stati discreti (Benedetti e Espa, 1993).

Questo paragrafo riveste una particolare importanza perchè permette di esaminare gli aspetti dinamici dei modelli esaminati precedentemente, proprio perchè la simulazione procede per stati successivi.

10.1 Generazione di un processo auto-normale in R^2

Si supponga di voler generare una realizzazione X da una distribuzione $MVN(\mu, \Sigma)$. Siano $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ delle variabili aleatorie indipendenti distribuite normalmente di media nulla e varianza unitaria ($N(0,1)$). Sotto le ipotesi formulate $X=H\varepsilon$, dove H è una matrice di dimensione $n \times n$ tale che $\Sigma = HH'$. Così per campionare dalla variabile aleatoria X è sufficiente tenere presente che:

$$x_i = \mu_i + \sum_{j=1}^n h_{ij} \varepsilon_j \quad [10.1]$$

dove μ_i è specificata per ogni sito (e può essere costante) ed il vettore ε può essere ottenuto utilizzando dei comuni generatori di numeri casuali.

Il primo passo della procedura di simulazione descritta, cioè la generazione delle ε_i per tutte le n celle per il calcolo della [10.1], deve essere eseguito con estrema cura. Infatti è raccomandabile che la sequenza dei valori generati $\{\varepsilon_i\}$ sia sottoposta a verifica per accertarsi che essa goda delle proprietà richieste ad un campione estratto da una distribuzione normale. Inoltre andrebbe studiata anche la distribuzione spaziale dei numeri generati per accertarsi che non sia affetta da autocorrelazione. Alcune procedure

per condurre tali tipi di verifiche si trovano in Haining (1990, cap. 6).

La matrice \mathbf{H} può essere ottenuta numericamente dal momento che è la matrice triangolare inferiore della nota scomposizione di Cholesky della matrice simmetrica Σ . Infatti se per Σ è possibile effettuare la semplice scomposizione $\mathbf{Q}\mathbf{A}\mathbf{Q}'$, dove le colonne della matrice \mathbf{Q} di dimensione $n \times n$ sono gli autovettori normalizzati di Σ e \mathbf{A} la matrice diagonale $n \times n$ dei corrispondenti autovalori, allora

$$\mathbf{H} = \mathbf{Q}\mathbf{A}^{1/2}$$

dove $\mathbf{A}^{1/2}$ denota le radici quadrate degli autovalori.

Il calcolo della matrice \mathbf{H} risulta comunque spesso proibitivo per grigliati di dimensioni anche non eccessive. Per questo motivo si ricorre comunemente a metodi alternativi che richiedono calcoli di tipo locale, come l'uso di medie mobili ponderate con pesi stabiliti in funzione dell'autocorrelazione spaziale del processo.

Un esempio di metodi di questa classe è la generazione di un processo a media mobile (Haining, 1978) e quindi, data l'invertibilità, anche dell'analogo a media mobile di un processo SAR (cfr. paragrafo 8.). In questo caso la matrice \mathbf{H} può essere specificata direttamente; infatti $\Sigma = \sigma^2[(\mathbf{I} - \mathbf{B})'(\mathbf{I} - \mathbf{B})]^{-1}$ ed $\mathbf{H} = \sigma(\mathbf{I} - \mathbf{B})^{-1}$ (Haining, Griffith e Bennett, 1983).

L'autocorrelazione spaziale teorica $R(r,s)$ a distanza r sull'asse delle ascisse X e a distanza s sull'asse delle ordinate Y , può essere derivata analiticamente dato il filtro di convoluzione $f(r,s)$ (cfr. tabella 10.1 e 10.2). Si supponga infatti di generare un processo $\varepsilon_{ij} \sim N(0,1)$ tale che

$$E(\varepsilon_{i+r,j+s}\varepsilon_{i,j}) = \begin{cases} 1 & \text{se } r = s \\ 0 & \text{altrimenti} \end{cases}$$

	-4	-3	-2	-1	0	1	2	3	4
-4	0,002	0,005	0,008	0,011	0,013	0,011	0,008	0,005	0,002
-3	0,005	0,010	0,019	0,030	0,035	0,030	0,019	0,010	0,005
-2	0,008	0,019	0,042	0,076	0,096	0,076	0,042	0,019	0,008
-1	0,011	0,030	0,076	0,172	0,261	0,172	0,076	0,030	0,011
0	0,013	0,035	0,096	0,261	0,708	0,261	0,096	0,035	0,013
1	0,011	0,030	0,076	0,172	0,261	0,172	0,076	0,030	0,011
2	0,008	0,019	0,042	0,076	0,096	0,076	0,042	0,019	0,008
3	0,005	0,010	0,019	0,030	0,035	0,030	0,019	0,010	0,005
4	0,002	0,005	0,008	0,011	0,013	0,011	0,008	0,005	0,002

Corr	0	1	2	3	4	5	6
0	1,000	0,723	0,429	0,226	0,106	0,038	0,012
1	0,723	0,591	0,372	0,202	0,096	0,035	0,011
2	0,429	0,372	0,254	0,147	0,073	0,028	0,009
3	0,226	0,202	0,147	0,091	0,048	0,019	0,006
4	0,106	0,096	0,073	0,048	0,026	0,011	0,004
5	0,038	0,035	0,028	0,019	0,011	0,005	0,002
6	0,012	0,011	0,009	0,006	0,004	0,002	0,001

Tab. 10.1 Filtro isotropo 9×9 e relativa funzione numerica di autocorrelazione.

	-4	-3	-2	-1	0	1	2	3	4
-4	0,009	0,011	0,012	0,013	0,014	0,013	0,012	0,011	0,009
-3	0,020	0,024	0,029	0,032	0,034	0,032	0,029	0,024	0,020
-2	0,039	0,053	0,067	0,078	0,082	0,078	0,067	0,053	0,039
-1	0,067	0,098	0,139	0,181	0,201	0,181	0,139	0,098	0,067
0	0,082	0,128	0,201	0,314	0,492	0,314	0,201	0,128	0,082
1	0,067	0,098	0,139	0,181	0,201	0,181	0,139	0,098	0,067
2	0,039	0,053	0,067	0,078	0,082	0,078	0,067	0,053	0,039
3	0,020	0,024	0,029	0,032	0,034	0,032	0,029	0,024	0,020
4	0,009	0,011	0,012	0,013	0,014	0,013	0,012	0,011	0,009

Corr	0	1	2	3	4	5	6
0	1,000	0,918	0,769	0,593	0,412	0,238	0,126
1	0,803	0,756	0,651	0,513	0,362	0,213	0,114
2	0,525	0,498	0,438	0,354	0,256	0,156	0,086
3	0,302	0,287	0,255	0,209	0,155	0,097	0,055
4	0,153	0,145	0,129	0,107	0,081	0,052	0,030
5	0,060	0,057	0,051	0,042	0,032	0,022	0,014
6	0,020	0,019	0,017	0,014	0,011	0,008	0,005

Tab. 10.2 Filtro 9×9 con effetto direzionale e relativa funzione numerica di autocorrelazione.

Il processo u_{ij} ottenuto come

$$u_{ij} = \sum_{r,s} f_{rs} \varepsilon_{i+r,j+s}$$

risulterà avere una distribuzione normale con media 0, varianza

$$\text{Var}(u_{ij}) = \sum_{r,s} f_{rs}^2 \text{E}(\varepsilon_{i+r,j+s}^2) = \sum_{r,s} f_{rs}^2$$

e covarianza

$$E(u_{ij}u_{i+m,j+n}) = E\left(\sum_{r,s} f_{rs}^2 \varepsilon_{i+r,j+s} \sum_{u,v} f_{uv} \varepsilon_{i+m+u,j+n+v}\right)$$

il cui calcolo risulta molto semplice se, nello sviluppo del prodotto fra le due sommatorie, si tiene conto dei soli termini per i quali $u=r-m$ e $v=s-n$. Infatti per questi termini $E(\varepsilon_{i+r,j+s} \varepsilon_{i+m+u,j+n+v})$ sarà pari ad uno e per tutti gli altri elementi sarà pari a zero.

Nelle tabelle 10.1 e 10.2 sono riportati due esempi di calcolo dell'autocorrelazione di un processo generato tramite convoluzione di un processo i.i.d. con distribuzione $N(0,1)$ con un filtro $f(r,s)$, rispettivamente isotropo e omogeneo ma con dipendenza differente nella direzione nord-sud da quella nella direzione est-ovest (cfr. Figura 10.1 per le realizzazioni, entrambe di media 0 e varianza 1, ottenute con questi filtri).

10.2 Generazione di un campo markoviano a stati discreti in R^2

La questione è molto più complessa per quanto riguarda la generazione di un campo markoviano a stati discreti; in questo caso i pochi algoritmi studiati sono iterativi e si basano sulla convergenza asintotica, all'aumentare del numero di iterazioni t , di una successione di configurazioni $X(t)$ ad una realizzazione del campo markoviano oggetto dell'analisi (Ripley, 1987). Tra questi il metodo di Metropolis et. al. (1953) è uno dei precursori. Sinteticamente se $X(t)$ è la configurazione alla t -ma iterazione, si sceglie casualmente un'altra configurazione, diciamo η , e definito q il rapporto tra le probabilità delle due realizzazioni:

$$q = \pi(\eta) / \pi[X(t)]$$

si pone $X(t+1) = \eta$ con probabilità pari a $\min\{1; q\}$.

La nuova configurazione η può essere scelta ad esempio estraendo casualmente una cella i di cui si intende modificare la classe di appartenenza.

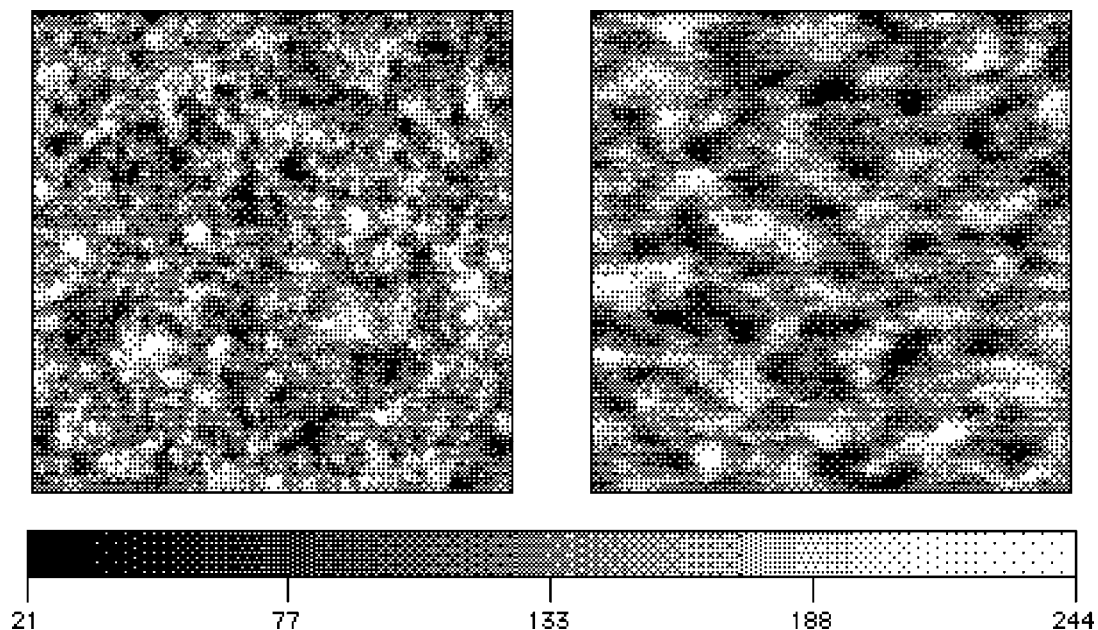


Fig. 10.1 Realizzazioni di un processo a media mobile di media 128 e deviazione standard 30, ottenute con il metodo delle medie mobili pesate, su di un grigliato 100×100 utilizzando i filtri introdotti, rispettivamente, in Tab. 10.1, isotropo, ed in Tab. 10.2, con effetto direzionale. La scala si riferisce ai livelli di grigio relativi ai valori compresi tra 0 e 255

La probabilità di passaggio dalla classe k alla classe h sarà quindi calcolata in base a $\min\{1; q\}$, dove q corrisponderà alla quantità:

$$\frac{P(X_i = h | X_j = x_j, j \neq i, j \in C\{i\})}{P(X_i = k | X_j = x_j, j \neq i, j \in C\{i\})}. \quad [10.2]$$

Un algoritmo per generare un MRF a stati discreti con il metodo di Metropolis è il seguente:

Algoritmo 1 per simulare un campo aleatorio markoviano

(1) Si inizializza il sistema spaziale di dimensioni $M \times N$ assegnando ad ogni cella una classe a caso dall'insieme $\mathcal{Q} = \{0, 1, \dots, G-1\}$ dei corrispondenti numeri interi. Si indichi con \mathbf{x} questo stato iniziale.

(2) Per l'indice s che assume valori da 1 a n

- (a) si sceglie $g \in \mathcal{Q}$ a caso e si pone $y_s = g$. Sia anche $y_t = x_t, \forall t \neq s$;
- (b) sia $p = \min\{1, P(\mathbf{X} = \mathbf{y}) / P(\mathbf{X} = \mathbf{x})\}$;

(c) si rimpiazza x con y con probabilità p .

(3) Si ripete il passo (2) un numero prefissato N_{iter} di volte.

Per quanto riguarda il passo (b) è utile mostrare con un semplice esempio relativo ad una singola cella il modo di procedere, con particolare riguardo all'applicazione della regola [10.2].

Sia lo stato iniziale x quello fornito dalla Figura 10.2 e sia inoltre $y_5=2$, in modo che il secondo stato del campo aleatorio markoviano oggetto della simulazione differisca dal primo per la classe assegnata al solo pixel in posizione $s=5$ ($x_5=1$ e $y_5=2$); nella situazione descritta le probabilità $P(\mathbf{X} = \mathbf{x})$ e $P(\mathbf{X} = \mathbf{y})$ sono computabili in base al modello scelto.

1	3	1
3	1	2
2	2	2

Fig.10.2 Stato iniziale per la simulazione di un MRF.

Se si considera il modello di Strauss (1977) (cfr. relazione [7.8]) si dovranno calcolare per il pixel in posizione $s=5$ le seguenti probabilità:

$$\gamma = P(X_5 = 2 | X_j = x_j, j \neq 5, j \in C\{5\}) \text{ e } \delta = P(X_5 = 1 | X_j = x_j, j \neq 5, j \in C\{5\}).$$

Banalmente, escludendo le costanti di normalizzazione (i denominatori della formula [7.8]) che nel calcolo della [10.2] si semplificano, si otterranno le quantità $\gamma = e^{4\beta}$ e $\delta = e^{2\beta}$. A questo punto, per ciò che concerne il passo (c), se il valore del rapporto $\gamma/\delta > 1$ allora si rimpiazzerà certamente (i.e. con probabilità pari ad 1) x_5 con y_5 e quindi x con y . Altrimenti, se $0 < \gamma/\delta < 1$, si farà ricorso al cosiddetto *rejection sampling* per sostituire x con y con probabilità $p = \gamma/\delta$. A tale proposito ci si avvarrà di un comune generatore di numeri casuali uniformemente distribuiti nell'intervallo $[0,1]$. Detto $u \sim U[0, 1]$, se $p > u$ si sostituirà x con y altrimenti si conserverà lo stato iniziale.

La convergenza dell'algoritmo 1 è assicurata se N_{iter} è abbastanza grande.

Barker (1965) ha proposto come definizione alternativa di q la quantità:

$$q = \pi(\eta) / (\pi[X(t)] + \pi(\eta))$$

più piccola della precedente e sempre inferiore ad uno.

Nel caso dei campi markoviani binari (si veda a tale proposito la Figura 10.3) quest'espressione si semplifica in:

$$\begin{aligned} & \frac{P(X_i = h | X_j = x_j, j \neq i, j \in C\{i\})}{P(X_i = k | X_j = x_j, j \neq i, j \in C\{i\}) + P(X_i = h | X_j = x_j, j \neq i, j \in C\{i\})} = \\ & = P(X_i = h | X_j = x_j, j \neq i, j \in C\{i\}) \end{aligned} \quad [10.3]$$

che equivale a scegliere la nuova classe della i -ma cella direttamente dalla distribuzione condizionata. Questa variante, denominata nella letteratura *Gibbs sampler*, può essere utilizzata anche per un numero di classi superiore a due (cfr. Figura 10.4) (Geman e Geman, 1984).

Un algoritmo per generare un MRF a stati discreti con il metodo del *Gibbs sampler* seguendo la regola decisionale [10.3] è il seguente:

Algoritmo 2 per simulare un campo aleatorio markoviano

- (1) Si inizializza il sistema spaziale di dimensioni $M \times N$ assegnando ad ogni cella una classe a caso dall'insieme $Q = \{0, 1, \dots, G-1\}$ dei corrispondenti numeri interi. Si indichi con \mathbf{x} questo stato iniziale.
- (2) Per l'indice s che assume valori da 1 a n
 - (a) si calcolano le probabilità $p_g = P(X_s = g | \mathbf{X}_{\partial s} = \mathbf{x}_{\partial s})$, per $g=0,1,\dots,G-1$, dove $\mathbf{x}_{\partial s}$ è l'insieme corrente delle classi assegnate alle celle situate nelle immediate vicinanze della cella s ;
 - (b) si assegna alla s -ma cella la classe g con probabilità p_g .
- (3) Si ripete il passo (2) un numero prefissato N_{iter} di volte.

Anche in questo caso è utile mostrare con un semplice esempio il modo di procedere, con particolare attenzione alla fase di calcolo delle probabilità di cui al passo (a). In riferimento ancora allo stato iniziale riportato in Figura 10.2, si supponga che la scansione del grigliato stesso previsto dall'Algoritmo 2 sia giunta a considerare il pixel

in posizione $s=5$. Allora le probabilità che bisognerà calcolare per assegnare una classe all' s -mo pixel saranno le seguenti:

$$\varphi = P(X_5 = 3 | \mathbf{X}_{\partial s} = \mathbf{x}_{\partial s}), \phi = P(X_5 = 1 | \mathbf{X}_{\partial s} = \mathbf{x}_{\partial s}) \text{ e } \omega = P(X_5 = 2 | \mathbf{X}_{\partial s} = \mathbf{x}_{\partial s}).$$

Se il modello di riferimento scelto è ancora quello di Strauss (1977), le quantità in base alle quali si eseguirà il passo (b) saranno le seguenti:

$$\varphi = e^{2\beta}, \phi = e^{2\beta} \text{ e } \omega = e^{4\beta}.$$

L'algoritmo 2 richiede che vengano calcolate G funzioni esponenziali e così richiederà maggior tempo di implementazione di quello che è necessario spendere per eseguire l'algoritmo 1. Entrambi gli algoritmi comunque usano uno stato iniziale scelto a caso, sebbene la teoria stabilisca che gli algoritmi convergeranno per ogni stato iniziale. Dubes e Jain (1989) sostengono che la convergenza degli algoritmi descritti avviene in un ragionevole lasso di tempo se lo stato iniziale è scelto in maniera completamente casuale.

Il parametro N_{iter} , che rappresenta il numero di scansioni dell'immagine, è un parametro cruciale. Numerose soluzioni euristiche sono state proposte in letteratura per scegliere N_{iter} . Valori molto frequenti sono 50 e 100. Comunque arrestare l'algoritmo troppo presto può condurre a realizzazioni atipiche. Un MRF è soggetto a fenomeni di *phase transition*⁸ (Pickard, 1987) quando i suoi parametri vengono specificati in modo tale che le correlazioni a corto raggio fra pixels vicini degenerino in correlazioni a lungo raggio (cfr. paragrafo 7.1 per un esempio chiarificatore di natura economica). La tendenza degli algoritmi di simulazione a produrre risultati degeneri (ovvero realizzazioni dominate da una o due classi) è stata sottolineata nella letteratura statistica (Besag, 1986) ma non in quella, per così dire, ingegneristica (Geman e Geman, 1984).

Un modo di evitare realizzazioni mono-classe è quello di imporre che il numero di pixels appartenenti ad ogni classe non cambi rispetto alla scelta casuale iniziale. Un algoritmo che sfrutta questa intuizione è quello denominato *spin exchange* (Flinn, 1974; Cross e Jain, 1983); secondo questo metodo la selezione della nuova configurazione avviene estraendo casualmente due siti e scambiandone le classi di appartenenza. Ciò conduce ad una definizione di q , nel caso i due siti non siano contigui, uguale a:

⁸ Deve essere sottolineato che in ambito economico la terminologia *phase transition* è stata utilizzata con una doppia valenza. Alcuni autori, tra cui Föllmer, utilizzano tale espressione riferendosi alla teoria dei campi di Markov infiniti, in cui *phase transition* indica l'esistenza di una molteplicità di fasi. Altri si rifanno invece al significato originale di *phase transition*, ovvero al concetto che rappresenta il superamento di una soglia critica e la modificazione qualitativa della distribuzione di probabilità. A quest'ultima interpretazione ci si riferisce in questa sede quando si parla di *phase transition*.

$$\frac{P(X_r = x_s | X_j = x_j, j \neq r, j \in C\{r\}) P(X_s = x_r | X_j = x_j, j \neq s, j \in C\{s\})}{P(X_r = x_r | X_j = x_j, j \neq r, j \in C\{r\}) P(X_s = x_s | X_j = x_j, j \neq s, j \in C\{s\})}.$$

Nel caso i due siti siano contigui deve essere derivata, a seconda del modello oggetto dell'analisi, la distribuzione condizionata congiunta delle due variabili aleatorie. Occorre sottolineare come questo algoritmo presenti il difetto di non modificare l'istogramma dei dati simulati da un'iterazione all'altra risultando quindi molto dipendente dalla configurazione iniziale (Ripley, 1987). Inoltre questo algoritmo non rispetta lo spirito della condizione di positività [6.1] che implica che tutte le realizzazioni abbiano una probabilità non nulla di presentarsi.

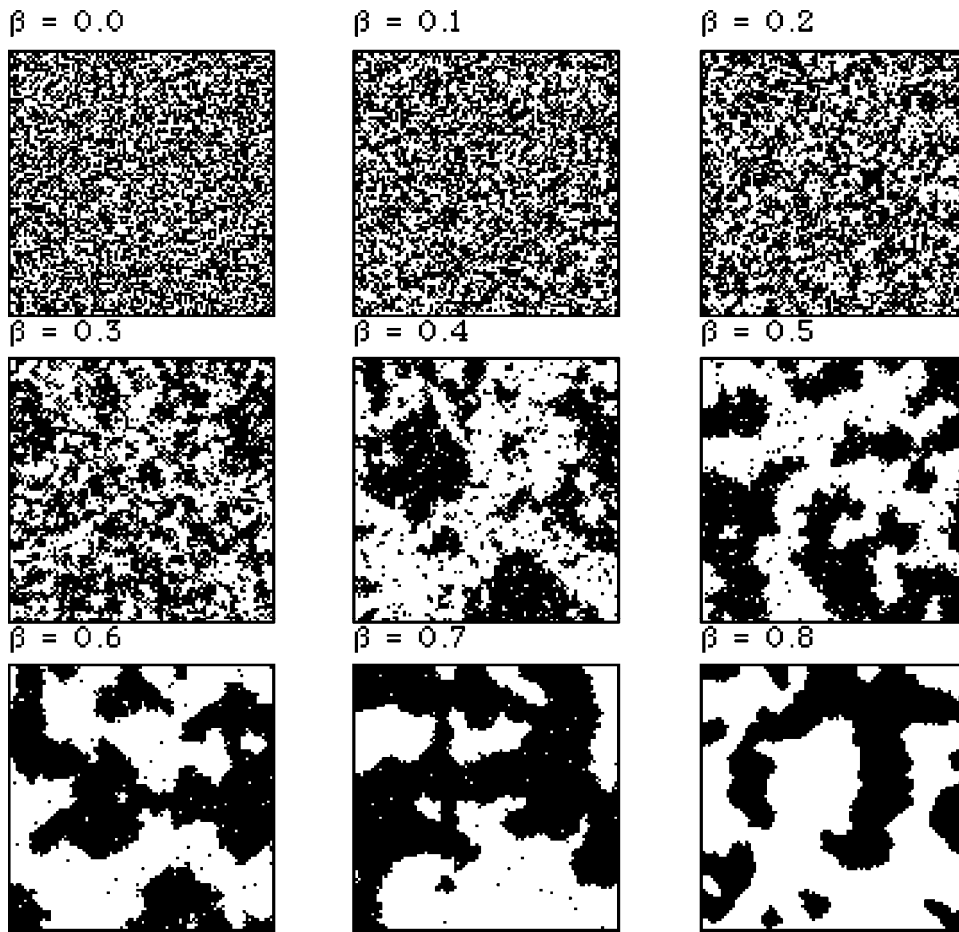


Fig.10.3 Realizzazioni ottenute con il metodo del Gibbs-sampler, di un modello auto-logistico con $\alpha=0$ e β costante su di un grigliato 100x100 per differenti valori del parametro β .

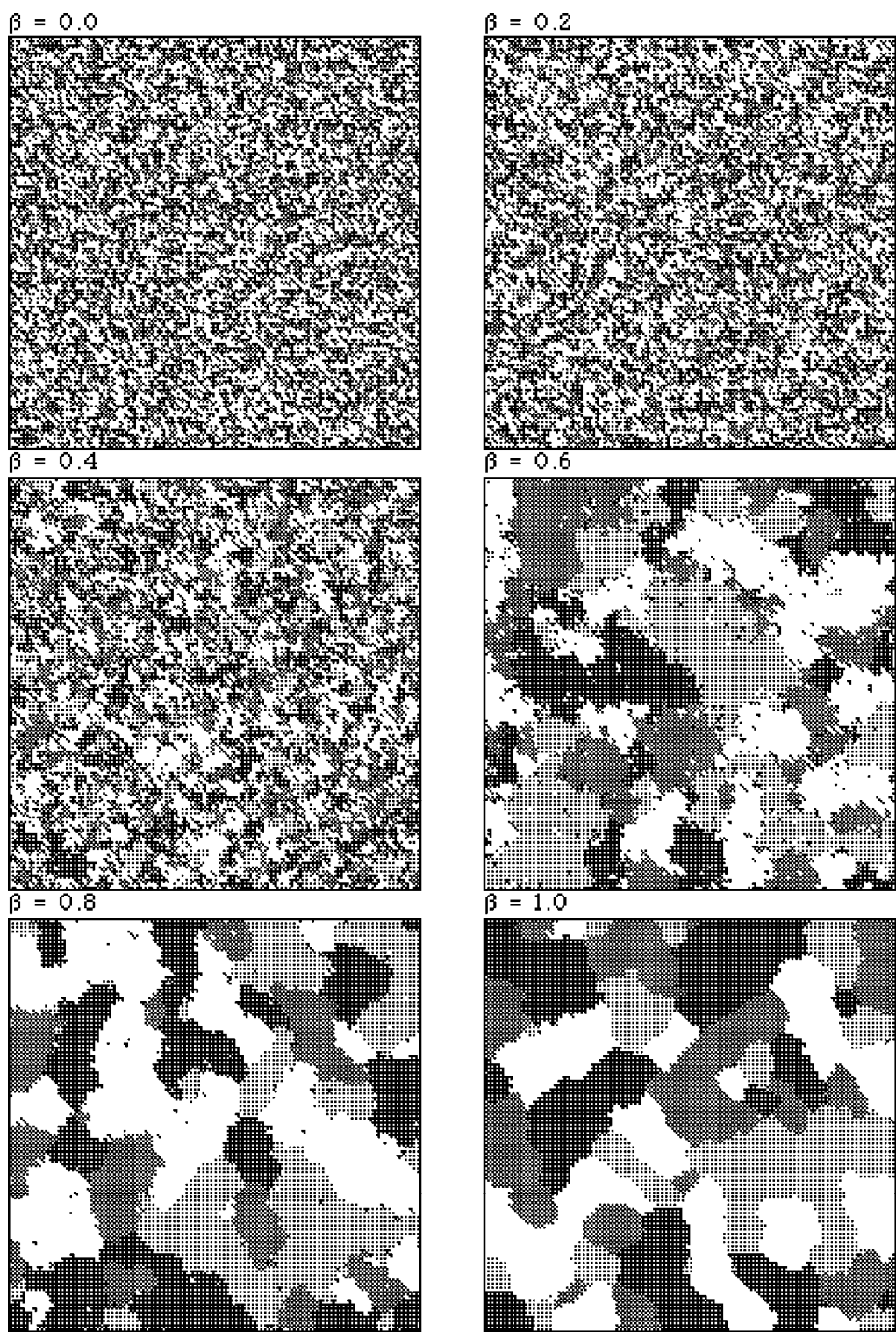


Fig. 10.4 Realizzazioni, ottenute con il metodo del Gibbs-sampler, di un modello di Strauss a 4 classi con β costante su di un grigliato 100x100 per differenti valori del parametro β .

11. Conclusioni.

Nel presente lavoro sono stati esposti i presupposti statistici su cui si fonda una classe di modelli di interazione spaziale, ovvero quei modelli basati sui campi aleatori di Gibbs e di Markov su reticoli. Inoltre si sono presentati i principali aspetti applicativi concernenti la stima e la simulazione di tali modelli. L'obiettivo è quello di fornire uno strumento analitico utilizzabile in un vasto campo di possibili applicazioni, sia teoriche sia empiriche.

La recente letteratura economica ha utilizzato tali modelli soprattutto per l'analisi di economie decentralizzate dove gli agenti, posizionati in una esplicita dimensione spaziale, interagiscono tra loro. L'applicazione dei modelli di interazione spaziale basati sui campi aleatori di Gibbs e di Markov in ambito economico presenta alcuni limiti dovuti particolarmente alle ipotesi fatte sulla struttura dell'interazione e sulle distribuzioni di probabilità che caratterizzano ogni agente.

Innanzitutto si assume che la struttura di interazione sia simmetrica, ovvero nessun agente può influenzarne un altro senza a sua volta esserne influenzato. Questo implica l'impossibilità di modellare rapporti strettamente unidirezionali, dove un agente influenza le azioni di altri agenti, ma viceversa non viene influenzato dalle azioni degli altri agenti. Se da un punto di vista puramente teorico questo può rappresentare un limite di tali modelli, poiché essi non contemplano tale possibilità, in realtà non lo costituisce, poiché è difficilmente concepibile una situazione dove un agente ne influenza un altro senza a sua volta esserne minimamente influenzato, sia che si tratti di istituzioni o di singoli individui.

Limiti assai più vincolanti derivano dalle ipotesi fatte sull'insieme dei vicini. Si assume infatti che l'insieme dei vicini sia sempre uguale per numerosità e composizione e spazialmente invariante. Inoltre ogni agente rimane collocato nella stessa posizione per tutta la durata del processo. Ciò significa ipotizzare che gli agenti si muovano sempre nel medesimo ambiente il quale a sua volta è statico, non cambia, per cui anche i vicini rilevanti con cui gli agenti interagiscono rimangono sempre gli stessi. Questo implica che questa classe di modelli sia più adatta a spiegare fenomeni circoscritti, quali possono essere la diffusione di una tecnologia o di standard utilizzati nella costruzione di beni di consumo (per esempio per studiare la battaglia tra gli standard: VCR *versus* Betamax o Pc-DOS *versus* Macintosh) dove non è inverosimile ipotizzare che gli agenti non cambino ambiente e l'insieme dei loro vicini rimanga costante durante tutto l'evolversi del fenomeno.

Inoltre deve essere evidenziato che i modelli di interazione basati sui campi aleatori di Gibbs e di Markov sono particolarmente adeguati per spiegare i processi di diffusione o di contagio perché la condizione di positività (cfr. paragrafo 6, pag.) permette che un'economia possa uscire da stati di *lock-in*. Infatti la probabilità di un agente di fare una delle scelte ammissibili è sempre positiva qualunque sia la scelta e qualsiasi sia la configurazione dell'ambiente. A livello macro ciò implica la non esistenza di stati

assorbenti e a livello micro la possibilità per ogni agente di non piegarsi alla legge della maggioranza anche se si possono avere fasi in cui una scelta può conquistare interamente l'ambiente.

Come abbiamo in precedenza evidenziato, i modelli di interazione spaziale hanno trovato un'ampia applicazione in economia poiché rappresentano uno strumento utile sia per l'analisi di problemi specifici, sia per la creazione di un paradigma teorico alternativo a quello dell'equilibrio generale. In particolare perché permettono di fondare a livello micro le dinamiche aggregate, o meglio permettono di derivare le dinamiche collettive dall'interazione di agenti eterogenei. Dati i limiti precedentemente esposti, ci sembra evidente che tali modelli non possono costituire la metodologia su cui basare un paradigma teorico alternativo all'equilibrio generale, ma rimangono strumenti che ottengono buoni risultati in contesti specifici, in particolare modo per modellare fenomeni di contagio (diffusione di una tecnologia e formazione di bolle nei mercati finanziari).

Come si è sottolineato, i modelli presi in esame in questo lavoro rappresentano solo una classe dei modelli utilizzati per spiegare l'interazione tra gli agenti. Sono stati sviluppati modelli quali i grafi stocastici e soprattutto le economie artificiali che ovviano a gran parte dei limiti sovraesposti. Tuttavia i modelli basati sui campi aleatori di Markov, , permettono di ottenere risultati analitici e possono essere stimati utilizzando dati reali, contrariamente a quanto avviene per i modelli più complessi caratterizzati da una struttura di interazione meno rigida (le economie artificiali, ad esempio, possono solo essere simulate). Pertanto si ripropone in quest'area di ricerca un "trade-off tra il desiderio di realismo delle assunzioni introdotte e la trattabilità analitica delle formalizzazioni ottenute" (Fagiolo, 1996) che ben lungi dall'essere una caratteristica peculiare di questo campo, è un elemento intrinseco e costante di tutta la ricerca in ambito economico.

Bibliografia

- Arbia G. (1986) On the Simulation of Two Dimensional Stationary Stochastic Processes: a Comparison of Different Approaches, atti di *Compstat, Short Communications and Posters*, a cura di F. De Antoni, N. Lauro e A. Rizzi, Physica, Roma.
- Arbia G. (1989) *Spatial Data Configuration in Statistical Analysis of Regional Economic and Related Problems*, Kluwer, Dordrecht.
- Arbia G. (1993a) Metodi Statistici nella Classificazione e Segmentazione di Immagini, atti del convegno *STATCHEM93 - Statistica e Chemiometria per lo studio dell'ambiente*, Venezia.
- Arbia G. (1993b) Recenti sviluppi nella modellistica spaziale, in S. Zani (ed.), *Metodi Statistici per le Analisi Territoriali*, Franco Angeli, Milano, 193-217.
- Arbia G., Benedetti R. (1990) Metodi di Ricampionamento nell'analisi delle Immagini Aeree e da Satellite, *Quaderni di Statistica ed Econometria*, 12, 143-152.
- Arbia G., Espa G. (1996) Forecasting statistical models of archaeological site location, *Archeologia e Calcolatori*, 7, 365-372.
- Arthur W.B. (1989) Competing Technologies, Increasing Returns and Lock-in by Historical Events, *Economic Journal*, 99, 116-31.
- Arthur W.B. Ermoliev Y., Kaniovski Y. (1987), Path-dependence Processes and the Emergence of Macro-structure, *European Journal of Operational Research*, 30:1.
- Barker A.A. (1965) Monte Carlo Calculations of Radial Distribution Functions for a Proton-electron Plasma, *Australian Journal of Physics*, 18, 119-133.
- Bartlett M.S. (1955) *An Introduction to Stochastic Processes*, Cambridge University Press, Cambridge.
- Benedetti R. (1994) *Analisi Statistica di Immagini Digitali*, published Ph.D. dissertation, University of Rome "La Sapienza", Dept. of Statistics, Probability and Applied Statistics.
- Benedetti R., Espa G. (1993) Alcuni Metodi per la Generazione di Processi Aleatori Bidimensionali, *Statistica Applicata*, 5, 3, 217-244.
- Bennett R.J. (1979) *Spatial Time Series*, Pion, London.

- Besag J. (1974) Spatial Interaction and the Statistical Analysis of Lattice Systems (with discussion), *Journal of the Royal Statistical Society*, B, 36, 192-236.
- Besag J. (1975) Statistical Analysis of Non-lattice Data, *The Statistician*, 24, 3, 179-195.
- Besag J. (1977) Efficiency of pseudo-likelihood estimation for simple Gaussian fields, *Biometrika*, 64, 3, 616-618.
- Besag J. (1986) On the Statistical Analysis of Dirty Pictures (with discussion), *Journal of the Royal Statistical Society*, B, 48, 259-302.
- Box G.E.P., Jenkins G.M. (1976) *Time series Analysis - forecasting and control*, Revised Edition, Holden-Day, California.
- Burrough P.A. (1986) *Principles of Geographical Information Systems for Land Resources Assesment*, Clarendon Press, Oxford.
- Cliff A.D., Ord J.K. (1981) *Spatial Processes - Models and Applications*, Pion, London.
- Cox D.R. (1970) *Analysis of Binary Data*, Methuen, London.
- Cross G.R., Jain A.K. (1983) Markov Random Field Texture Models, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-5, 25-39.
- Dalle J.M. (1994) Technological Competition, micro-decision and diversity, *EUNETIC Conference*, Strasbourg, 1059-1073.
- David P.A. (1988) *Putting the Past into the Future of Economics*, Institute for Mathematical Studies in the Social Sciences Technical Report 533, Stanford University.
- David P.A. (1992) Path-dependence and Predictability in Dynamic Systems with Local Externalities: a Paradigm fo Historical Economics, in D. Foray and C. Freeman , eds., *Technology and the Wealth of Nations*, Londra
- David P.A., Foray D. (1992) Markov Random Fields, Percolation Structures and the Economics of EDI standards diffusion, *Discussion Paper Series no.326*, CEPR, Stanford University.
- Dosi G., Ermoliev Y., Kaniovski Y. (1994), Generalized Urn Schemes and Technological Dynamics, *Journal of Mathematical Economics*, 23, 1-19.
- Derin H., Elliott H. (1987) Modeling and Segmentation of Noisy and Textured Images Using Gibbs Random Fields, *IEEE Transactions on Pattern Analysis and*

Machine Intelligence, Vol. PAMI-1, 39-55.

Diggle P.J. (1983) *Statistical Analysis of Spatial Point Patterns*, Academic Press, London.

Dubes R.C., Jain A.K. (1989) Random Field Models in Image Analysis, *Journal of Applied Statistics*, 16, 131-164.

Durlauf S.N. (1993) Nonergodic Economic Growth, *Review of Economic Studies*, 60, 349-366.

Durlauf S.N. (1994) Path-dependence in aggregate output, *Industrial and Corporate Change*, 3:1.

Fagiolo G. (1996) Modelli di Interazione Spaziale in Economie Decentralizzate: una Rassegna, Dipartimento di Scienze Economiche, Università di Roma "La Sapienza", *Discussion Paper*, 29.

Flinn P.A. (1974) Monte Carlo Calculation of Phase Separation in a 2-Dimensional Ising System, *Journal of Statistical Physics*, 10, 89-97.

Föllmer H. (1974) Random Economies with Many Interacting Agents, *Journal of Mathematical Economics*, 1, 51-62.

Geman D., Graffigne C. (1986) Markov Random Field Image Models and their Applications to Computer Vision, *Proceedings of the International Congress of Mathematics*, Gleason (ed.), American Mathematical Society, Providence, 1596-1516.

Geman S., Geman D. (1984) Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-7, 6, 721-741.

Grimmett G.R. (1973) A Theorem About Random Fields, *Bullettin of the London Mathematical Society*, 5, 81-84.

Haining R.P. (1978) The Moving Average Model for Spatial Interaction, *Transactions of the Institute of British Geographers*, New Series 3, 202-225.

Haining R.P. (1984) Testing a spatial interacting market hypothesis, *The Review of Economics and Statistics*, 66, 576-583.

Haining R.P. (1990) *Spatial data analysis in the social and environmental sciences*, Cambridge University Press, Cambridge.

- Haining R.P., Griffith D.A., Bennett R.J. (1983) Simulating two-dimensional autocorrelated surfaces, *Geographical Analysis*, 15, 247-255.
- Hammersley J.M., Clifford P. (1971) Markov Fields on Finite Graphs and Lattices, Manoscritto non pubblicato.
- Hors I., Lordon F (1995) About Some Formalism of Interaction, *EUNETIC Conference*, Strasbourg.
- Kirman A. (1991) Epidemics of Opinion and Speculative Bubbles in Financial Markets, in *Money and Financial Markets*, M.Taylor, ed., Macmillan, Londra, 354-368.
- Kirman A. (1992) Variety: the Coexistence of Techniques, *Revue d'Economie Industrielle*, 59, 62-74.
- Krugman P. (1992) A Dynamic Spatial model, *NBER Working Paper*, 4219, November.
- Mardia K.V. (1990) Maximum Likelihood Estimation for Spatial Models, in *Spatial Statistics: Past Present and Future*, D.A. Griffith (ed.), Monograph. n. 12, Institute of Mathematical Geography.
- Mardia K.V., Watkins A.J. (1989) On Multimodality of the Likelihood for the Spatial Linear Model, *Biometrika*, 76, 289-295.
- Metropolis N., Rosenbluth A.W., Rosenbluth N.M., Teller A.H., Teller E (1953) Equations of State Calculations by Fast Computing Machines, *Journal of Chemical Physics*, 21, 1087-1092.
- Morrison D.F. (1967) *Multivariate Statistical Methods*, McGraw Hill, New York.
- Ord J.K. (1975) Estimation Methods for Models of Spatial Interaction, *Journal of the American Statistical Association*, 70, 120-126.
- Orlean A.(1990) Le rôle des influences interpersonnelles dans la détermination des cours boursiers, *Revue Economique*, 41, 839-68.
- Orlean A.(1992) Contagion des opinions et fonctionnement des marchés financiers , *Revue Economique*, 43, 685-97.
- Pickard D.K. (1987) Inference for Discrete Markov Fields: the Simplest Nontrivial Case, *Journal of the American Statistical Association*, 82, 397, 90-96.
- Preston C.J. (1973) Generalized Gibbs States and Markov Random Fields, *Advances in Applied Probability*, 5, 242-261.

- Quah D. (1996) Regional Convergence Clusters across Europe, *European Economic Review*, 40, 951-958.
- Ripley B.D. (1981) *Spatial Statistics*, Wiley, New York.
- Ripley B.D. (1987) *Stochastic Simulation*, Wiley & Sons, New York.
- Ripley B.D. (1988) *Statistical Inference for Spatial Processes*, Cambridge University Press, Cambridge.
- Ripley B. D. (1990) Gibbsian Interaction Models, in *Spatial Statistics: Past Present and Future*, D.A. Griffith (ed.), Monograph. n. 12, Institute of Mathematical Geography.
- Schendel D.E., Balestra P. (1969) Retail Behavior and Gasoline Price Wars, *Applied Economics*, 89-101.
- Strauss D.J. (1977) Clustering on Coloured Lattices, *Journal of Applied Probability*, 14, 135-143.
- Tobler W.R. (1970) A computer movie simulating urban growth in the Detroit Region, *Economic Geography* (Supplement), 46, 234-240.
- Whittle P. (1954) On Stationary Processes in the Plane, *Biometrika*, 41, 434-449.