

UNIVERSITÀ DEGLI STUDI DI PADOVA

Facoltà di Scienze MM.FF.NN.
Corso di Laurea Magistrale in Fisica



DALLE NETWORKS E DAI POLIMERI, ALLA
COMPLESSITÀ ECONOMICA E ALLO SVILUPPO

Relatore:

Chiar.mo Prof. Attilio Stella

Correlatore:

Dr. Michele Caraglio

Laureando:

Marco Zanlorenzi

Anno Accademico 2014-2015

Indice

Introduzione	1
1 Modelli di Networks e proprietà principali	5
1.1 Proprietà principali di una network	5
1.2 Modello di Erdős-Rényi	7
1.3 Modello di Watts-Strogatz	8
1.4 Modello di Barabási	8
2 Network dei Prodotti	11
2.1 NBER dataset e classificazione dei prodotti	11
2.2 Il concetto di prossimità	12
2.3 Costruzione della network	13
2.3.1 Caratteristiche della network	15
3 Dai polimeri alla crescita globale	21
3.1 Equazione per la dinamica	21
3.1.1 Coefficienti di trasferimento	23
3.1.2 Generazione del rumore	24
3.2 Analisi della dinamica	25
3.2.1 Rumore nullo	25
3.2.2 Rumore non nullo	26
3.3 Simulazioni su network di prova	28
3.3.1 Velocità di crescita	28
3.3.2 Dinamica in reticolo unidimensionale	32
3.4 Network dei prodotti	35
4 Calibrazione della network	39
4.1 Stime di τ e σ	39
4.2 Analisi dei dati reali	40

4.3	Calibrazione in τ e σ	42
4.4	Calibrazione in μ e J^*	45
5	Estensioni del modello e considerazioni finali	49
A	Cenni di Teoria dei Grafi	53
A.1	Laplaciano discreto	53
A.2	Equazioni differenziali lineari a coefficienti costanti	55

Introduzione

La fisica dei sistemi complessi permette lo studio di fenomeni spesso apparentemente lontani da quelli che sono generalmente riconosciuti come problemi di interesse fisico. Tale branca della fisica trova applicabilità in un vastissimo range di problemi: sistemi sociali [4], ecosistemi [1], [19], studio degli esseri viventi [24], degli organi del corpo umano [11], etc. Il cardine su cui poggiano tali analisi risiede nell'individuazione di meccanismi di interazione locale del sistema determinanti comportamenti (o dinamiche) di tipo collettivo. In questa tesi ci proponiamo di analizzare l'economia globale con i metodi della fisica, cercando di individuarne le principali caratteristiche.

Per inquadrare l'ambito in cui questo lavoro va a collocarsi, descriviamo brevemente quali sono le modellizzazioni delle teorie classiche riguardo lo sviluppo economico, e le ragioni per le quali esse non sono in grado di descriverlo in modo sufficientemente adeguato. Classicamente il processo di evoluzione che una nazione intraprende al fine di migliorare il proprio status economico va sotto il nome di “*structural transformation*”, ed esso prevede lo spostarsi dalla produzione di beni semplici (tipici di nazioni poco sviluppate) verso beni più complessi. Secondo le teorie classiche, tale trasformazione avviene grazie ai “*factor endowments*”, che rappresentano la quantità di capitali (economici, lavorativi, infrastrutturali) di cui una nazione dispone e che può sfruttare per migliorare le sue esportazioni. Tuttavia, come fatto notare in [15] questa visione risulta troppo minimalista per molti motivi: possono esservi difficoltà di sviluppo di certi beni dovuti a carenze di conoscenze in settori fortemente dipendenti dal “*learning by doing*”, o ancora esternalità fra prodotti (con effetti positivi o meno) e ricadute tecnologiche che hanno ripercussioni in settori diversi da quelli in cui hanno avuto origine. Altre problematiche derivano dalle sinergie presenti fra le varie nazioni (si pensi ad esempio agli interessi delle multinazionali nello spostare interi settori produttivi nei paesi in via di sviluppo). Tutte queste dinamiche portano a rendere difficoltose previsioni di crescita (oppure crisi) attenendosi solo a criteri classici.

Risulta quindi necessario, al fine di comprendere il comportamento dell'economia globale, sia studiare con più attenzione i legami che intercorrono fra i vari paesi, che dare un'interpretazione più sottile per le trasformazioni strutturali, individuando quali siano i beni verso cui una nazione possa avere più interessi a spostare la produzione, e che

effetto questo abbia su scala globale.

Negli ultimi anni, per cercare di rispondere a queste domande, la fisica dei sistemi complessi ha cercato di caratterizzare in vari modi il mercato globale. In [25] è stato proposto un modello di network (il *World Trade Web*) in cui le varie nazioni sono messe in relazione secondo i flussi di denaro scambiato. La topologia di tale network (che risulta essere di tipo *scale-free*) ha messo in evidenza come le nazioni più influenti risiedano nelle zone centrali della rete ed esibiscano un elevato grado di interconnessione. In questo modo è stato possibile ricavare informazioni su come le eventuali crisi di certi paesi si diffondano sul mercato globale “contagiando” le altre nazioni, oltreché capire come vengano veicolati i flussi monetari fra i paesi, mettendone in luce i legami che vi intercorrono.

Altri studi sono stati rivolti a caratterizzare l’insieme dei prodotti esportati dalle nazioni. Si è così visto [12], che le nazioni piuttosto che specializzarsi su pochi beni (come previsto dalle teorie classiche), tendono a produrre qualunque bene che non sia precluso dal loro livello di sviluppo tecnologico. Da questa osservazione è stato possibile classificare i prodotti in funzione della loro complessità e determinare i “fitness” delle nazioni studiandone l’insieme dei prodotti esportati. Analizzando poi nel corso del tempo la trasformazione strutturale delle nazioni, si è indagato su quali di esse abbiano effettuato sforzi economicamente “tangibili” reinvestendo i guadagni verso nuove tecnologie.

Ancora, in un altro lavoro di grande rilievo [7] (che costituisce il punto di partenza di questa tesi) si è stabilito in che misura un determinato prodotto sia “simile” ad un altro guardando con che probabilità tali prodotti vengano contemporaneamente esportati da una stessa nazione. Attraverso questa funzione di “similarità” diventa possibile costruire una rete per lo spazio dei prodotti, i cui nodi sono i prodotti in esame, mentre i collegamenti fra essi rappresentano quanto i vari prodotti siano “prossimi” fra loro. Si è visto come per i singoli paesi risulti più facile sviluppare beni quando questi sono affini ai prodotti che essi sono già in grado di produrre. Studiando poi il ritorno per capita medio associato ai vari beni nello spazio dei prodotti e analizzando la configurazione produttiva delle singole nazioni, è stato possibile delineare quali siano le cause che portano i paesi poveri ad essere incapaci di raggiungere il benessere tipico dei paesi sviluppati.

In questa tesi ci occupiamo dello studio della network definita in [7]. L’obbiettivo è quello di introdurre una dinamica per mezzo delle equazioni che vengono generalmente impiegate nell’ambito della fisica per la risoluzione di sistemi stocastici. Come illustrato in [13] è possibile infatti usare tali equazioni in molti contesti in cui vi siano meccanismi di tipo competitivo incerniati fra la crescita di una quantità dovuta alle sue interazioni con il resto del sistema rispetto che alla sua evoluzione isolata. Tale paradigma è applicabile in svariati campi, quali ad esempio lo studio delle dinamiche di

popolazioni [22], problemi di pinning nei semiconduttori [14], strategie di reinvestimento delle risorse [10], etc. L'applicazione di tali equazioni al sistema economico ha come obiettivo quello di capire come lo sviluppo proceda e da quali parametri esso dipenda. Una volta ricostruita la rete di produzione globale, se ne studieranno dapprima le caratteristiche topologiche (confrontandole con quanto ottenuto in [7]) e successivamente si analizzerà come imporre su tale network le equazioni per la dinamica. Tali equazioni verranno simulate, oltre che sulla network economica, anche su network di prova al fine di capirne meglio le soluzioni. Una volta studiato in dettaglio il comportamento della dinamica, si cercherà di calibrare la nostra analisi sui dati reali relativi alla produzione globale, così da ottenere informazioni sulla configurazione produttiva e individuarne le caratteristiche principali. Si cercherà infine di estrarre dall'analisi fatta anche informazioni sulla configurazione produttiva delle singole nazioni, cercando di comprendere perché esse abbiano successo (o meno) nel mercato.

La struttura della tesi è la seguente: nel capitolo 1 sono riportate alcune nozioni sulle networks. Vengono esposti i modelli di grafi più usati in questo ambito, descrivendone le caratteristiche principali. Tali informazioni si riveleranno utili per la comprensione e il raffronto con l'analisi che andremo a fare nei successivi capitoli, fornendo sistemi semplici su cui mettere alla prova la dinamica da noi sviluppata. Nel capitolo 2 viene spiegato in che modo sia possibile costruire la network dei prodotti. Dapprima si spiegherà brevemente la struttura del dataset a nostra disposizione, successivamente verrà introdotto il concetto di "prossimità" e verrà costruita la network che sarà oggetto di studio nei capitoli successivi. In tale capitolo verranno anche delineate le caratteristiche topologiche principali dello spazio dei prodotti. Nel capitolo 3 tratteremo l'implementazione della dinamica in tale network. Innanzitutto descriveremo equazioni che verranno utilizzate per la simulazione, scegliendo per esse i parametri che meglio sembrano adattarsi all'analisi di un sistema economico. Successivamente studieremo i risultati delle simulazioni nelle network di prova introdotte nel capitolo 1 e sulla network di produzione globale. Nel capitolo 4 inizieremo l'analisi dei dati reali a nostra disposizione. L'obiettivo sarà anzitutto quello di capire quanto il modello scelto sia in grado di adattarsi al comportamento reale del sistema e successivamente quello di calibrare il nostro modello su tali dati. Nel capitolo 5 infine, prenderemo in considerazione lo sviluppo delle singole nazioni al fine di metterne in luce le differenze; discuteremo i limiti dell'approccio e valuteremo le possibili estensioni del modello qui analizzato. In appendice sono raccolte alcune nozioni di teoria dei grafi che verranno utilizzate nell'elaborato.

Capitolo 1

Modelli di Networks e proprietà principali

Affiancata alla teoria dei grafi si è andata sviluppando negli ultimi decenni quella che viene chiamata “Network Science”: tale branca, rispetto alla sua controparte matematica, studia principalmente l’applicabilità e le caratteristiche dei grafi derivanti da problemi reali. Per tale ragione in questa disciplina alcuni grafi giocano un ruolo centrale rispetto ad altri. In questo capitolo prendiamo in esame le network di Erdős, di Watt-Strogatz e di Barabási, che mostrano caratteristiche ben distinte fra loro (per una descrizione esaustiva delle loro proprietà si può consultare [23]). Il nostro scopo è principalmente quello di avere delle network di confronto per evidenziare differenze di comportamento rispetto alla network economica globale che andremo a costruire nei prossimi capitoli.

1.1 Proprietà principali di una network

Dato un grafo (A.1.1), la network science si interessa a capire che “forma” abbia e che ruolo abbiano i nodi al suo interno. Introduciamo qui una serie di definizioni basilari e di parametri comunemente utilizzati per la loro analisi che ci saranno utili successivamente. Indicato un grafo con $G(V, E)$, ove V è l’insieme dei suoi vertici ed E l’insieme dei link di cui esso dispone, è possibile capire molte delle sue caratteristiche studiandone diametro, coefficiente di clustering e centralità, quantità che ora illustreremo (va detto che esse non hanno sempre una definizione univoca ma tendono a variare leggermente a seconda degli autori). Per *diametro* si intende la lunghezza della più lunga geodetica (A.1.2) contenuta nel grafo e questo valore viene usato per avere un’idea di quale sia la “larghezza” della network. Il *coefficiente di clustering* (locale) stima invece quanto un

nodo sia interconnesso con i suoi vicini: dato un nodo v_i , indichiamo con S_i l'insieme dei primi vicini relativi al nodo v_i , cioè i nodi v_j per i quali esiste un link $e_{ij} = e_{i \leftrightarrow j}$. Il coefficiente di clustering è allora definito come il rapporto fra il numero di link presenti fra i nodi contenuti nell'insieme S_i e il massimo numero di link che è possibile avere con $\text{Card}(S_i) = |S_i|$ nodi, ovvero $|S_i|(|S_i| - 1)/2$,

$$C_i = \frac{2|\{e_{jk} : v_j, v_k \in S_i, e_{jk} \in E\}|}{|S_i|(|S_i| - 1)} \quad (1.1)$$

Per la centralità esistono varie definizioni. Essa mira ad identificare quali sono i nodi più importanti della network ed esponiamo qui solo alcune delle sue definizioni più importanti (fra cui quelle che useremo poi nel corso dell'elaborato). La *degree centrality* è la più semplice e forse una delle più utili fra le varie formulazioni. Per ogni nodo v_i nella rete, essa è definita come il degree (A.1.5) del nodo in esame (eventualmente normalizzato). Tuttavia tale schematizzazione non riesce a cogliere la possibilità in cui un nodo, pur avendo molti link di connessione, sia connesso con nodi aventi degree basso (risultando quindi anch'esso periferico). Un'altra definizione d'uso comune poggia sulla nozione di *distanza* (o *farness*), che per un nodo v_i è definita come la somma della lunghezza delle sue geodetiche: la *closeness centrality* è l'inverso di questa distanza e può essere interpretata come una misura di quanto velocemente le informazioni si propagano all'interno della network se trasmesse dal nodo i -esimo: dato un grafo con n nodi, indichiamo con $\ell(i, j)$ la lunghezza della geodetica dal nodo v_i a v_j . Normalizzando rispetto ai nodi presenti nel grafo la closeness-centrality del nodo i -esimo è data da

$$C_i^C = \frac{n - 1}{\sum_{i \neq j} \ell(i, j)} \quad (1.2)$$

La *betweenness centrality* misura invece quante siano le geodetiche passanti per il nodo i -esimo. Risulta

$$C_i^B = \sum_{j \neq i, k \neq i} \frac{\sigma_{jk}(v_i)}{\sigma_{jk}} \frac{2}{(n - 1)(n - 2)} \quad (1.3)$$

dove σ_{jk} è il numero di geodetiche da v_j a v_k e $\sigma_{jk}(v_i)$ sono invece le geodetiche da v_j a v_k passanti per v_i . L'ultimo fattore costituisce la normalizzazione. Descriviamo infine la *eigenvector centrality*: essa codifica nella sua formulazione il seguente ragionamento: “un nodo è importante se è collegato a altri nodi rilevanti”. Dato un grafo G consideriamo la sua matrice di adiacenza $A = [a_{ij}]$ (vedi A.1.4). La centralità x_i del nodo i -esimo è definita da

$$x_i = \frac{1}{\lambda} \sum_k a_{ki} x_k \quad (1.4)$$

con $\lambda \neq 0$ costante. In forma matriciale l'equazione corrispondente è $\lambda \vec{x} = \vec{x}A$. Al fine di ottenere valori positivi per ogni coordinata del vettore \vec{x} , sfruttando il teorema di

Perron-Frobenius [26] risulta che il valore di λ da considerare è quello relativo al più grande autovalore (in modulo) della matrice di adiacenza.

Questa piccola digressione (lontana dall'essere completa) sulle misure di centralità fa intuire quanto la scelta su quale sia la definizione più opportuna dipenda da ciò che si vuole evidenziare e dal tipo di network che si sta studiando.

1.2 Modello di Erdős-Rényi

Uno dei grafi più semplici da costruire è quello aleatorio, o di Erdős, che è per l'appunto un grafo casuale. Per costruirlo è necessario specificare il numero di nodi n e la probabilità p che vi sia un link di connessione fra i vari nodi. Dati p ed n , indichiamo uno qualunque di questi grafi con $E(n, p)$. Se E è il numero di edge, il degree medio per tale grafo vale $\langle deg(v_i) \rangle = p(n - 1) = 2 \langle E \rangle / n$. In virtù della loro struttura, in tali grafi non sono presenti nodi aventi caratteristiche distinte rispetto agli altri e risulta che la distribuzione del degree è di tipo binomiale. Ovviamente per p piccolo non tutti i nodi sono connessi fra loro; quando però $p \approx \log(n)/n$ la network contiene un'enorme componente connessa (e noi saremo interessati soprattutto a questo regime). Per $p = 1$ il grafo diventa *completo*: ogni nodo è connesso con tutti gli altri.

I grafi di Erdős hanno una scarsa applicabilità nello studio dei fenomeni reali: questi ultimi infatti esibiscono in genere elevati coefficienti di clustering (mentre in tale grafo il esso è semplicemente proporzionale alla quantità di link presenti, e quindi a p) e una distribuzione di degree a legge di potenza (cioè la probabilità di avere un nodo avente degree k , varia secondo $p \sim k^{-\gamma}$), cosa che qui non avviene. Tuttavia tali grafi vengono spesso utilizzati come metro di confronto per le altre network vista la loro imparzialità sulla distribuzione dei link, e anche nella nostra analisi verranno utilizzati con questo intento.

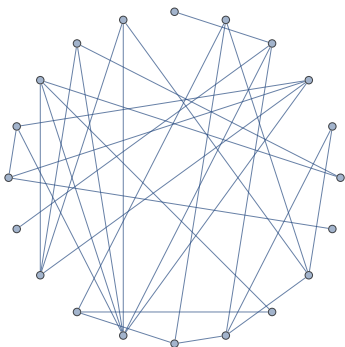


Figura 1.1: $E(20, 0.16)$

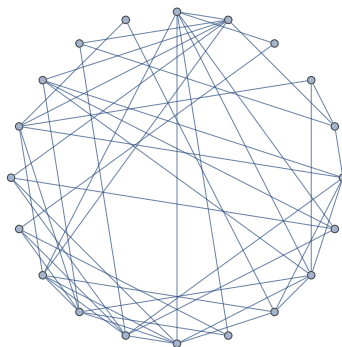


Figura 1.2: $E(20, 0.26)$

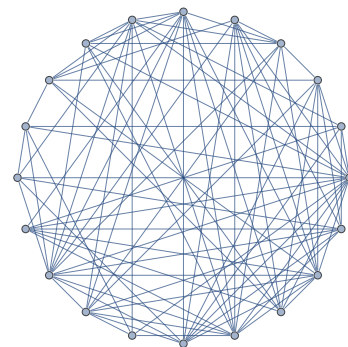
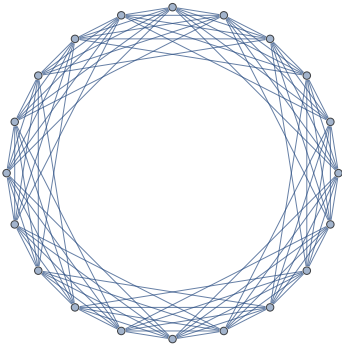
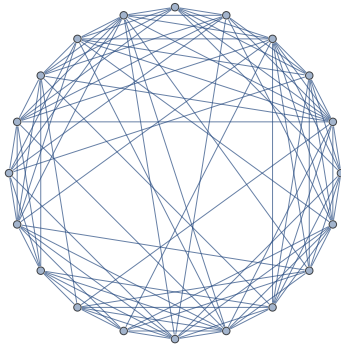
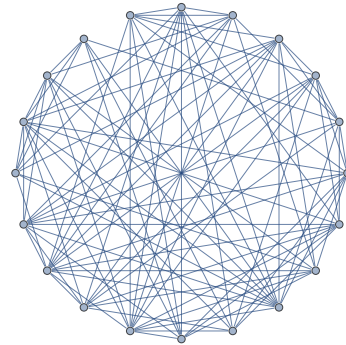


Figura 1.3: $E(20, 0.47)$

1.3 Modello di Watts-Strogatz

Per ottenere questi tipi di network si genera anzitutto un reticolo composto da n nodi in uno spazio d -dimensionale in cui ogni nodo è connesso con i suoi $2k$ primi vicini (Vedi figura 1.4, in cui $k = 5$ e $d = 1$). In questo modo si ottengono grafi (detti genericamente k -regolari) aventi un elevato coefficiente di clustering (si ha $C = 3(k - 2)/4(k - 1)$) e con cammini (A.1.2) medi grandi (proporzionali ad n/k). A questo punto la procedura di Watts-Strogatz (WS) prevede il ricablaggio con una certa probabilità p di ogni link contenuto nel grafo, che viene ricollocato in maniera casuale. È facile intuire che per $p = 1$ (ricablando quindi con certezza ogni link) il grafo viene trasformato in una network di Erdős. Fra questi due estremi (in funzione del valore di p imposto) si collocano le network di WS propriamente dette: quello che si ottiene sono grafi in cui la lunghezza delle geodetiche è simile a quella ottenibile dalle network di Erdős (con $\ell \sim \ln(n)/\ln(k)$) mantenendo però coefficienti di clustering elevati, analoghi al caso $p = 0$ (risulta $C = (3(k - 2)/4(k - 1))(1 - p)^3$, con quindi una lenta diminuzione al crescere di p).

Il modello di Watts-Strogatz riesce a cogliere molti aspetti delle network reali [27] e genera grafi aventi struttura *small-world*: con questo termine vengono generalmente chiamate tutte quelle network in cui il cammino medio è di lunghezza $\sim \log(n)$. Indichiamo tali grafici come $W(n, p, k)$ ove n è il numero di nodi, p la probabilità di ricablaggio e $2k$ è il degree di ogni nodo nella configurazione k -regolare di partenza.

Figura 1.4: $W(20, 0, 5)$ Figura 1.5: $W(20, .15, 5)$ Figura 1.6: $W(20, .65, 5)$

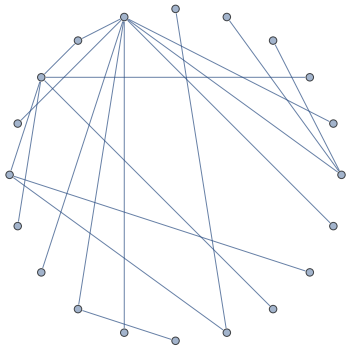
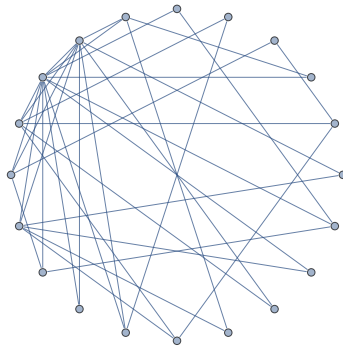
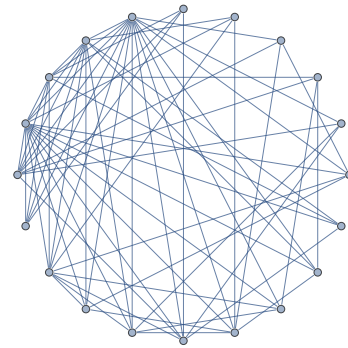
1.4 Modello di Barabási

Il modello di Barabási (detto anche ad “attaccamento preferenziale”) può essere implementato in maniera ricorsiva come segue: supponiamo di avere una rete di Barabási

con n nodi, e di voler aggiungere l' $(n + 1)$ -esimo richiedendo che esso abbia k link di connessione: il nuovo nodo verrà collegato al nodo i -esimo già presente nella network con una legge di probabilità

$$p_i = \frac{k_i}{\sum_{j=1}^n k_j} \quad (1.5)$$

ove k_i indica il degree del nodo i -esimo. Questa network fa parte delle network dette *scale-free*, per le quali il degree dei nodi segue una distribuzione a legge di potenza, ovvero la probabilità che nella network sia presente un nodo con degree k vale $P(k) \sim k^{-\gamma}$ con in genere $2 < \gamma < 3$. Questa caratteristica è riscontrabile in una grande quantità di network associate fenomeni reali (suggerendo che anch'esse possano originarsi attraverso processi analoghi a (1.5)), in cui alcuni nodi giocano un ruolo dominante all'interno del sistema in virtù della loro capacità di interagire con una grande frazione dello stesso. Il coefficiente di clustering in questo modello è dell'ordine $C \sim n^{-3/4}$ e le geodetiche hanno una lunghezza $\ell \sim \ln(n)/\ln(\ln(n))$, valore che configura questa network fra le *ultra-small newtwork*, per le quali il diametro è dell'ordine $\sim \log(\log(n))$. Indichiamo tali grafi con $B(n, k)$ ove n è il numero di nodi totali nel grafo e k è il numero di link con cui ogni nuovo nodo viene aggiunto alla network.

Figura 1.7: $B(20, 1)$ Figura 1.8: $B(20, 2)$ Figura 1.9: $B(20, 3)$

Capitolo 2

Network dei Prodotti

In questo capitolo spieghiamo in che modo sia possibile costruire una network atta a definire la prossimità fra i prodotti. Dato lo storico dei flussi di import ed export di tutte le nazioni, si può definire infatti una “affinità” fra i vari prodotti stimandone la probabilità condizionata di esportazione. Tale quantità andrà sotto il nome di “prossimità”, e rappresenterà una stima di quanto un prodotto sia correlato ad un altro.

2.1 NBER dataset e classificazione dei prodotti

I dati di import ed export delle varie nazioni vengono costantemente archiviati dall’ UN-Comtrade ¹ (*United Nations Commodity Trade Statistics Database*). Tale ente tiene traccia del 95% delle merci prodotte a livello mondiale. Il Comtrade è il punto di partenza per la costruzione dei principali dataset di import/export, che rispetto alla fonte primaria hanno il pregio di organizzare in maniera più fruibile i dati. Per il nostro studio la scelta è ricaduta sul database NBER-UN [20]. Tale database copre gli anni 1962 – 2000 e classifica i prodotti secondo la codifica SITC Rev.2 ² (*Standard International Trade Classification*): Si tratta di un codice a 4 cifre che permette di distinguere circa 1300 prodotti (o per meglio dire, categorie di prodotti: ogni codice è associato ad un insieme di merci aventi caratteristiche analoghe). Senza entrare troppo nei dettagli, per i quali rimandiamo alla documentazione ufficiale, riportiamo un piccolo esempio di questa classificazione per chiarirne la struttura.

¹<http://comtrade.un.org/>

²<http://unstats.un.org/unsd/cr/registry/regcst.asp?Cl=14>

Esempio di classificazione SITC Rev.2

1000 Beverages and tobacco
1100 Beverages
1110 Non-alcoholic beverages, n.e.s.
1120 Alcoholic beverages
1121 Wine of fresh grapes
1122 Other fermented Beverages n.e.s
1123 Beer made from malt
1124 ...

Il database è strutturato essenzialmente nel seguente modo: per ogni nazione presa in esame vengono dichiarati anno per anno gli export (verso quali nazioni) e gli import (da quali nazioni) di ogni prodotto. Per completezza facciamo presente che non sempre è possibile ricostruire esattamente il flusso di import/export (talvolta i paesi non dichiarano con precisione la merce scambiata). In tali casi i prodotti vengono indicati con codici SITC fittizi del tipo (riprendendo l'esempio in tabella): “11XX” oppure “112A”, che rappresentano rispettivamente flussi “extra” oppure “aggregati” nelle categorie “Beverages” e “Alcoholic beverages”. Tuttavia non ci soffermiamo troppo su questi aspetti perché tali quantità (dell'ordine dello 0,2% rispetto al flusso totale registrato nel dataset) verranno scartate per la successiva analisi. In ogni caso, per una spiegazione accurata di ciò che rappresentano si veda [20].

2.2 Il concetto di prossimità

In questa sezione e nella prossima ripercorriamo il lavoro [7]. Per farsi un'idea di cosa sia questa nozione di “prossimità”, che sarà il cardine della nostra rete di prodotti, facciamo prima un esempio pratico. Supponiamo che una nazione sia esportatrice di mele. Allora possiamo sospettare essa sia (o possa con facilità diventare) anche un'esportatrice di pere: sicuramente essa avrà terreno disponibile e un clima adatto, avrà infrastrutture necessarie allo stoccaggio, frigoriferi e container per il trasporto. Inoltre sarà in possesso di tutte le tecnologie necessarie alla produzione e disporrà del capitale umano per gestire il business del settore. In questo senso possiamo dire che i prodotti “mele” e “pere” hanno un coefficiente di prossimità alto. Viceversa, essere una nazione esportatrice di mele, non dà alcuna informazione (ad esempio) sulla capacità di quella nazione di produrre semiconduttori, visto che i settori in questione non hanno punti in comune. In tal caso la prossimità fra mele e semiconduttori risulterà bassa.

L'idea è quindi di studiare gli export globali, e definire più prossimi i beni che

vengono prodotti con più frequenza in tandem dalle nazioni. Il primo punto da sciogliere per l'applicazione di tale procedura è riuscire a discriminare in quali casi una nazione sia esportatrice o meno di un certo bene. A tale fine è stato utilizzato il coefficiente RCA (*Revealed Comparative Advantage*) definito come:

$$RCA(c, i) = \frac{x(c, i)}{\sum_i x(c, i)} \bigg/ \frac{\sum_c x(c, i)}{\sum_{i,c} x(c, i)} \quad (2.1)$$

$RCA(c, i)$ è il revealed comparative advantage della nazione c -esima relativo al prodotto i -esimo, di cui $x(c, i)$ ne rappresenta il valore di export associato. L'espressione quindi coglie in quali casi una nazione sia una “vera” esportatrice di un certo bene facendone un confronto con la media globale di export per quel prodotto. Nel caso in cui si abbia $RCA(c, i) > 1$ consideriamo la nazione c -esima una esportatrice del bene i -esimo. In caso contrario la sua produzione viene ritenuta marginale. A questo punto la prossimità è definita come

$$\phi_{ij} = \min \{E_{ij}, E_{ji}\} \quad (2.2)$$

ove

$$E_{ij} = \frac{\sum_c \Theta(RCA(c, j) - 1) \Theta(RCA(c, i) - 1)}{\sum_c \Theta(RCA(c, i) - 1)} \quad (2.3)$$

con $\Theta(x)$ funzione di Heaviside. E_{ij} rappresenta quindi con che probabilità una nazione che esporta il bene i -esimo esporta anche il bene j -esimo. La prossimità è poi definita come il minimo fra E_{ij} ed E_{ji} . La richiesta di minimo è necessaria per evitare di sovrastimare la prossimità nel caso in cui, ad esempio, una nazione “ c ” sia la sola produttrice di un certo bene “ i ”: si avrebbe E_{ji} molto piccolo (correttamente) mentre invece E_{ij} risulterebbe uguale a 1 per ogni bene “ j ” prodotto da “ c ”.

2.3 Costruzione della network

Sfruttando le ϕ_{ij} definite in (2.2) è possibile costruire una network pesata per lo spazio dei prodotti, ove i nodi rappresentano i beni presi in esame ed i link (pesati) sono dati dai valori della funzione di prossimità (per questa ragione in seguito confonderemo spesso i valori della funzione di prossimità ϕ_{ij} con il link ad essa associato). Per ottenere la nostra network abbiamo operato alcune restrizioni sui dati di partenza: concordemente a quanto fatto in [15] e come già accennato precedentemente, innanzitutto sono stati rimossi tutti i prodotti fittizi (cioè corrispondenti a flussi di aggregati o extra, per i quali non è possibile capire con precisione a chi siano riferiti) dal dataset; inoltre sono state scartate nazioni con meno di 2 milioni di abitanti e qualunque paese riportasse più

del 5% del proprio export sotto forma di aggregati/extra. Prendendo in considerazione il range temporale 1998-2000 è stata costruita la network, che risulta composta da 775 nodi (prodotti) e $\sim 2,7 \cdot 10^5$ link. La distribuzione dei link (cioè della funzione di prossimità) risulta approssimativamente log-normale (come riscontrato già in [7]) ed è mostrata in figura 2.1.

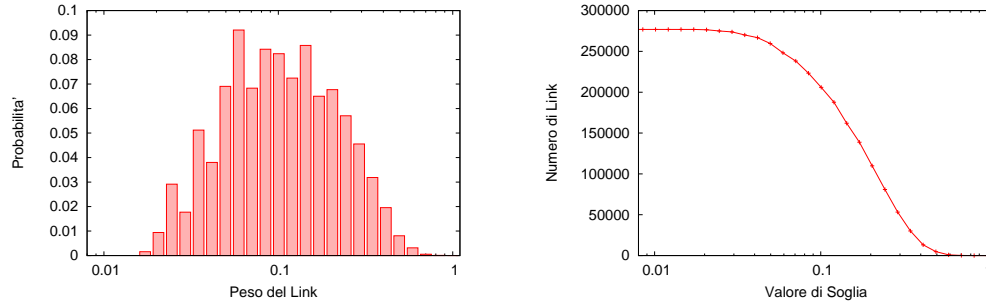


Figura 2.1: Peso dei link nella network globale. A sinistra è rappresentata la probabilità di avere un link di un dato peso: tale probabilità risulta seguire una legge di tipo log-normale. A destra: numero di link aventi peso maggiore di un dato valore di soglia.

Questa network è in grado di distinguere fra loro i diversi settori economici, attribuendo correttamente prossimità più alte a prodotti classificati come simili dai metodi classici [7]. Introdotta la nozione di “PRODY” [16], che rappresenta il ritorno per capita medio associato all’esportazione di un dato prodotto³, si evidenzia come esistano zone della network, perlopiù centrali, associate ad un ritorno economico maggiore. Infine si è studiata la capacità delle singole nazioni di sviluppare beni in funzione della loro configurazione produttiva [7]. È stato osservato come sia più probabile che una nazione sviluppi nuovi beni se questi si trovano in regioni dello spazio dei prodotti fortemente popolate (cioè circondati da prodotti con $RCA > 1$, ovvero già sviluppati). In particolare, per mettere in luce le differenze fra nazioni ricche e povere, è stata studiata attraverso una simulazione a passo singolo la velocità con cui diverse nazioni riescono a produrre nuovi beni: fissato un valore di soglia ϕ_c per le ϕ_{ij} si è supposto che una nazione riesca a sviluppare il bene j -esimo solo se esso è connesso ad un bene i già sviluppato (ovvero avente $RCA(c, i) > 1$) con un link avente peso $\phi_{ij} > \phi_c$. Si osserva come nazioni diverse abbiano velocità di crescita diverse e che il valore ϕ_c è critico per

³Più precisamente, indicato con Y_c il GDP per capita della nazione c -esima, il PRODY relativo al prodotto k è dato da

$$PRODY_k = \sum_c RCA(c, k)Y_c.$$

lo sviluppo: per $\phi_c > 0.65$ non tutte le nazioni riescono a raggiungere (dopo 20 passi di simulazione) lo sviluppo di ogni bene presente nella network e la distribuzione relativa ai beni sviluppati appare bimodale, cioè le nazioni povere tendono a rimanere confinate all'interno della network in regioni precise (ove il PRODY è più basso), incapaci di produrre i beni più centrali. Viceversa, per valori di $\phi_c < 0.55$ quasi tutte le nazioni raggiungono la produzione di tutti i beni (per una descrizione esaustiva di quanto qui riassunto, si rimanda a [7]).

2.3.1 Caratteristiche della network

In questa sezione cerchiamo di analizzare più in dettaglio le caratteristiche della network: essa è una rete connessa con 775 nodi e 276944 links aventi peso dato dalle ϕ_{ij} (vista la grande quantità di link la rete può considerarsi di fatto completa). Tuttavia, come riscontrabile nella figura 2.1, la maggior parte di questi link sono deboli (solo per l'1% di essi si ha $\phi_{ij} > 0.5$). Mostriamo qui i grafi della network (in cui i nodi sono stati colorati in funzione della loro degree centrality) al variare della soglia di cut-off per le ϕ_{ij} (scartando cioè i link aventi peso inferiore a un certo ϕ_c). Per ottenere tali grafici prima è stato ottenuto il *maximum spanning tree (MST)*, ovvero l'albero ricoprente (A.1.3) costruito selezionando i suoi $N - 1$ link in modo che la somma delle loro funzioni di prossimità ϕ_{ij} abbia il valore massimo possibile. Per la visualizzazione grafica l'algoritmo utilizzato calcola il rilassamento per un sistema carica-molla in cui ad ogni nodo viene associata una carica positiva unitaria, ed i link esercitano una forza attrattiva su di essi. I grafici ottenuti sono riportati nelle figure 2.3 - 2.6 e danno un'idea di quale sia la struttura di tale rete.

Una delle caratteristiche principali di questa network è quella di essere costituita da link pesati. Generalizzando il concetto di degree per network pesate come $deg(v_i) = \sum_{j \neq i} \phi_{ij}$, i grafici ottenuti (imponendo le soglie di cut-off sulla funzione di prossimità) si dimostrano di tipo scale-free, cioè analoghi alle network ad attaccamento preferenziale viste nella sezione 1.4. L'andamento a legge di potenza $P(k) = k^{-\gamma}$ si rivela particolarmente buono per una soglia di cut-off $\phi_c \approx 0.45$ e fornisce $\gamma = 2.17 \pm 0.07$. Nell'intorno di tale soglia è possibile individuare facilmente le zone ad elevato clustering all'interno della network, che sono state messe in evidenza in figura 2.7. Consultando la classificazione SITC dei nodi, risulta che la zona 1 è quasi interamente costituita da prodotti appartenenti ai codici 84XX, corrispondenti a "indumenti ed accessori di abbigliamento", la zona 2 è dominata da prodotti nelle categorie 64XX e 67XX (manufatti in metallo, acciaio). in 3 si trovano principalmente prodotti plastici (58XX), chimici (59XX), derivati di prodotti petroliferi (3351) e prodotti chimici organici (51XX). In 4 si hanno manifatti minerali (vetri, ceramiche, etc..) e macchinari per l'industria che coprono un grande range di applicabilità (per lavorare i metalli, turbine a vapore, for-

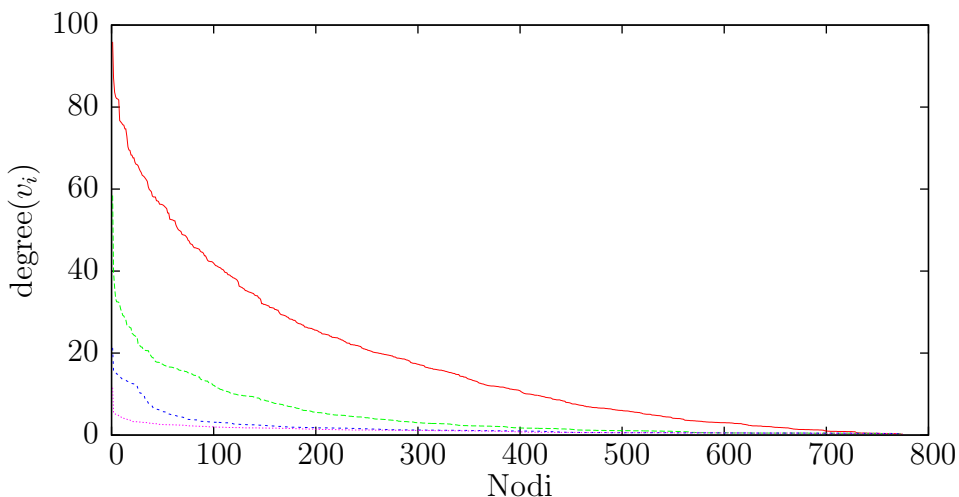


Figura 2.2: Degree dei nodi relativi alle network di figura 2.3 (viola), 2.4 (blu), 2.5 (verde), 2.6 (rosso).

naci, centrifughe, pompe, filtri, etc...) e per tale regione cercare di riassumere i codici SITC risulta un po' complicato vista l'estensione della zona stessa. La quinta zona è legata a settori d'interesse chimico, e contiene prevalentemente macchinari di stampa, per la produzione di carta, lastre fotografiche, strumenti di misura, apparecchiature mediche, prodotti chimici organici, apparecchiature per analisi chimiche e macchinari per industrie relative a settori specifici (72XX, 87XX, 7741, 51XX, 8822, 8749). Infine in 6 ricadono merci legate al settore elettronico riguardanti la produzione di semiconduttori, di valvole elettriche, tubi catodici, circuiti integrati, batterie ed accumulatori, potenziometri, insieme a settori riguardanti piccoli calcolatori, registratori, strumenti musicali (pianole e simili), microfoni, altoparlanti e fotocamere (7512, 8811, 776X, 7642, 772X, 8981, 7781). Le zone periferiche della network (non rappresentate in figura 2.7) sono invece dominate da prodotti appartenenti al settore primario.

L'analisi qui condotta mostra come la network costruita attraverso la funzione di prossimità riesca a rappresentare in maniera ragionevole lo spazio dei prodotti: i settori produttivi che condividono stesse conoscenze, materie prime e applicazioni vengono identificati correttamente come prossimi dalla rete ottenuta. In [7] si è inoltre evidenziato come tale network sia in grado di distinguere i prodotti concordemente con altre classificazioni [17] ottenute attraverso metodologie diverse, fornendo quindi una robusta prova della validità dell'approccio.



Figura 2.3: Maximum spanning tree per la network dei prodotti.

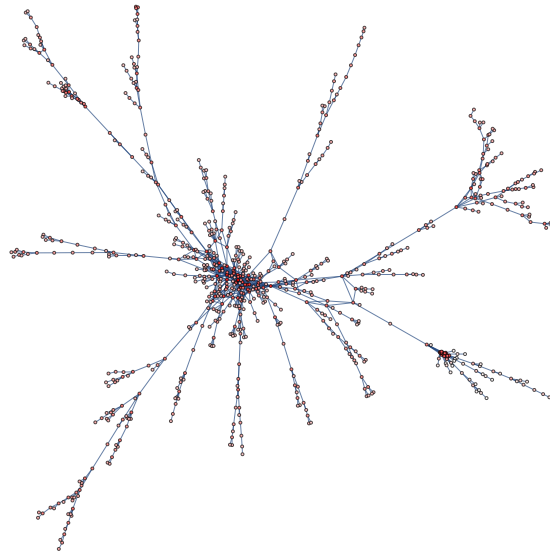


Figura 2.4: Network dei prodotti, soglia di cutoff $\phi_{ij} > 0.6$.

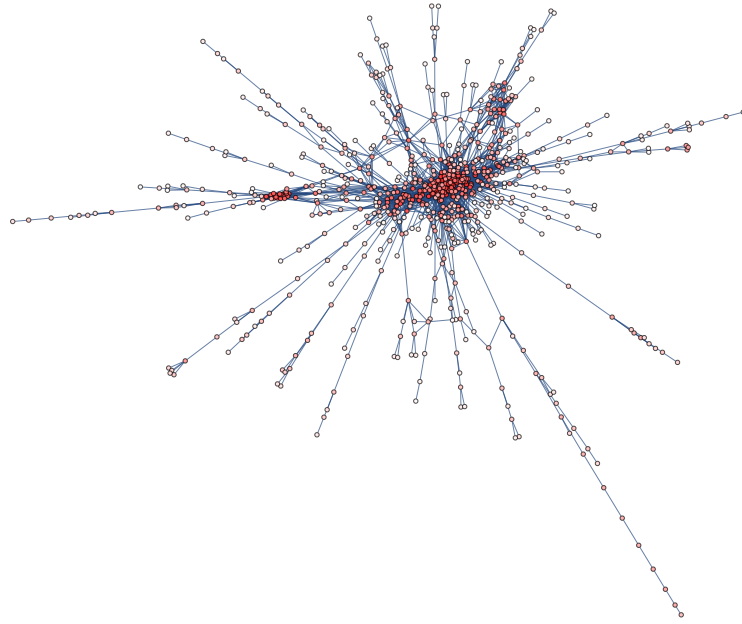


Figura 2.5: Network dei prodotti, soglia di cutoff $\phi_{ij} > 0.5$.

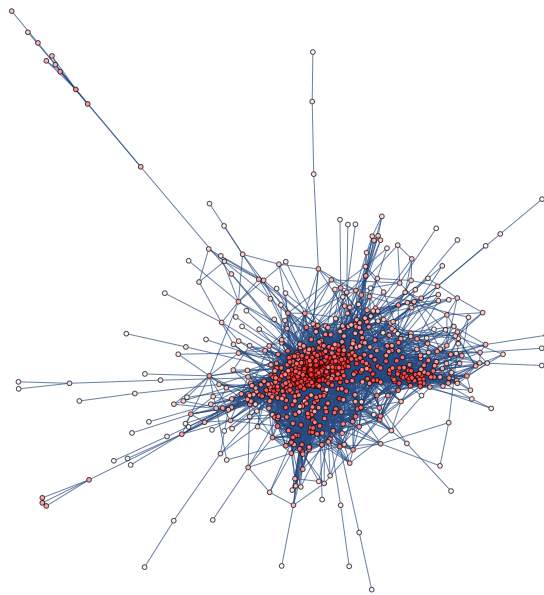


Figura 2.6: Network dei prodotti, soglia di cutoff $\phi_{ij} > 0.4$.

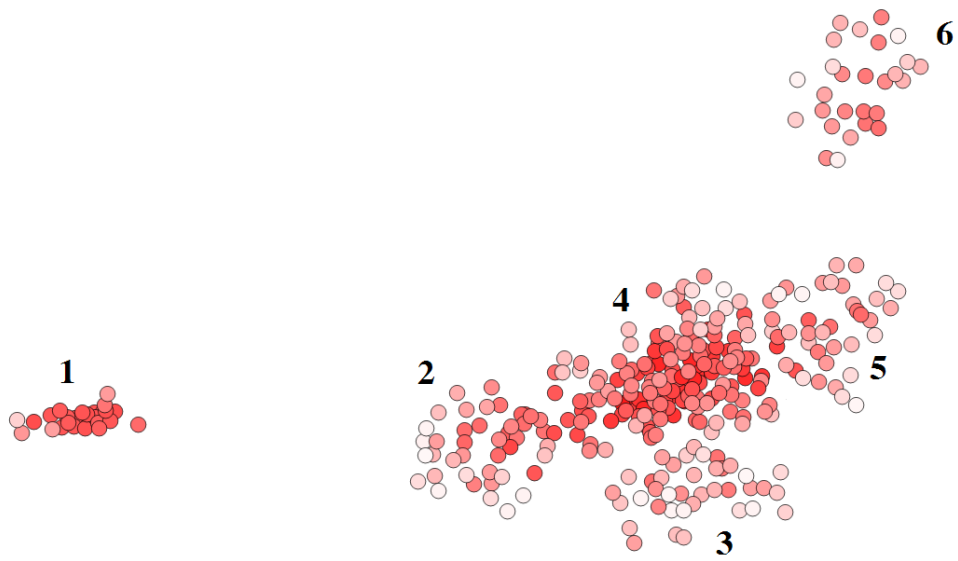


Figura 2.7: Regioni ad elevato clustering della network con soglia $\phi_{ij} > 0.5$.

Capitolo 3

Dai polimeri alla crescita globale

Un problema di interesse economico è capire fino a che punto sia conveniente investire capitali nella produzione di uno specifico bene piuttosto che cercare guadagni in nuove direzioni. Tale compromesso fra ricerca e sfruttamento delle risorse gioca un ruolo centrale in molti campi che spaziano fra i settori più vari. Parlando specificamente dell'economia globale, è facile capire come né investire sempre sullo stesso bene (andando a perdere preziose opportunità di guadagno) né spostare i capitali freneticamente in direzioni diverse (senza lasciar maturare gli investimenti) sia una strategia redditizia. In questo capitolo introduciamo e studiamo delle equazioni atte a descrivere queste dinamiche all'interno della nostra network.

3.1 Equazione per la dinamica

Come già accennato, per simulare la dinamica all'interno della network dei prodotti abbiamo bisogno di equazioni che codifichino le strategie di investimento (almeno in prima approssimazione). Vogliamo in particolare mettere in luce le differenze di guadagno che vi sono fra il continuare ad investire su una precisa produzione piuttosto che diversificare in una certa misura i propri investimenti. In [13] viene presentata un'equazione che incarna nella sua dinamica questo aspetto: prendendo come riferimento una variabile dinamica $Z_i(t)$ associata al nodo i -esimo di un generico grafo, modelliamo la sua evoluzione temporale secondo

$$\frac{\partial Z_i(t)}{\partial t} = \sum_{j \neq i} J_{ij} Z_j(t) - \sum_{j \neq i} J_{ji} Z_i(t) + \eta'_i(t) Z_i(t). \quad (3.1)$$

In questa equazione J_{ij} rappresenta un coefficiente di trasferimento dal nodo i al nodo j , mentre $\eta'_i(t)$ rappresenta un rumore agente sul nodo i -esimo. La variazione nel

tempo della quantità $Z_i(t)$ dipende quindi dal modo in cui il nodo i -esimo è connesso alla network attraverso i termini di scambio $\sum_{j \neq i} J_{ij} Z_j(t)$ e $\sum_{j \neq i} J_{ji} Z_i$ così come dal rumore $\eta'_i(t)$, che può favorire (o meno) la crescita della $Z_i(t)$. L'equazione (3.1), attraverso i coefficienti di trasferimento J_{ij} , permette di descrivere in che modo ogni nodo interagisce con il resto del sistema. Ad esempio, coefficienti J_{ij} piccoli, portano le $Z_i(t)$ ad evolvere indipendentemente (o quasi) l'una dall'altra. Viceversa, la presenza di coefficienti J_{ij} grandi fra un insieme di nodi, determinerà per essi un comportamento fortemente correlato. Nell'ambito economico, le due condizioni appena esposte corrispondono a strategie di investimento opposte fra loro: nel primo caso ($J_{ij} \rightarrow 0$) viene codificata una strategia di sfruttamento (o *exploit*) delle risorse: gli investimenti restano confinati in precisi settori (nella speranza di una loro crescita). Per J_{ij} grandi invece il regime viene detto di esplorazione (*explore*): si cerca cioè di trasferire gli eventuali guadagni (o di compensare le perdite) verso altri settori, differenziando così i propri investimenti.

Il paradigma di evoluzione codificato in (3.1) si presta alla descrizione di molti fenomeni. L'equazione (3.1), oltre ad essere già stata utilizzata in ambito economico (si veda [6], in cui ne è stato effettuato uno studio in approssimazione di campo medio) trova utilizzo in molti altri campi, fra cui citiamo lo studio di popolazioni di batteri [22] (in tal caso Z_i rappresenta il numero di batteri attorno all'habitat i -esimo e $\eta'_i(t)$ gioca il ruolo di condizioni locali, favorevoli o meno, relative al nodo quali quantità di nutrienti, clima, etc...), ed il problema del pinning [14]. Nel nostro caso $Z_i(t)$ rappresenta l'export in dollari relativo al prodotto i -esimo, in cui agisce un rate di crescita casuale $\eta'_i(t)$, corrispondente alla "fortuna" che quel determinato prodotto ha al tempo t (ad esempio un'annata climaticamente buona per i vini o, viceversa, disastrosa). Osserviamo che, posto $J_{ij} = J$ e immaginando i nodi della rete posizionati su un reticolo regolare d -dimensionale l'equazione (3.1) si traduce in una discretizzazione dell'equazione di diffusione

$$\frac{\partial Z(\vec{x}, t)}{\partial t} = J \nabla^2 Z(\vec{x}, t) + \eta'(\vec{x}, t) Z(\vec{x}, t).$$

Quest'ultima equazione, attraverso la trasformazione di Cole-Hopf $Z = \exp(h(\vec{x}, t))$ viene a sua volta mappata nell'equazione di Kardar-Parisi-Zhang

$$\frac{\partial h(\vec{x}, t)}{\partial t} = J \nabla^2 h(\vec{x}, t) + J (\nabla h(\vec{x}, t))^2 + \eta'(\vec{x}, t).$$

Tale equazione, oltre a descrivere la crescita delle interfacce [18], trova applicazioni al problema dei polimeri diretti in ambienti random [9], e in innumerevoli altri campi fra i quali i processi di ottimizzazione [10]. Per (3.1) siamo interessati soprattutto a quest'ultimo aspetto. In particolare vogliamo stabilire sotto quali condizioni le $Z_i(t)$

si sviluppano in maniera più rapida e capire quali informazioni è possibile ricavare attraverso questo modello riguardo l'economia globale e la network dei prodotti ad essa associata. Per la sua applicazione nell'ambito economico sarà necessaria una espressione per i coefficienti di trasferimento J_{ij} ed una descrizione appropriata per il rumore $\eta'_i(t)$. Nelle prossime due sezioni esponiamo quali scelte sono state operate per questi parametri.

3.1.1 Coefficienti di trasferimento

Il grafo che abbiamo ottenuto nel capitolo precedente attraverso la funzione di prossimità prevede link di tipo pesato ed è chiaro che le J_{ij} saranno funzione delle ϕ_{ij} . Facciamo anzitutto notare che, rispetto al lavoro [7] di cui si è parlato nella sezione 2.3, in cui per effettuare le simulazioni diffusive a passo singolo i link sono stati troncati ad un certo valore di soglia (senza tale accorgimento ogni nazione giungerebbe a sviluppare tutto lo spazio dei prodotti nel giro di qualche iterazione), nel nostro caso tale problema di cut-off non si pone: attraverso la network pesata il tasso di crescita dei prodotti dipenderà proprio da quanto essi siano affini. La definizione qui scelta per le J_{ij} è

$$J_{ij} = J^* \frac{\phi_{ij}}{\sum_k \phi_{kj}}, \quad \phi_{jj} = 1. \quad (3.2)$$

I coefficienti di trasferimento risultano quindi asimmetrici e in tale definizione J^* (che chiameremo *coefficiente di explore*) rappresenta, in senso medio, la propensione al reinvestimento delle risorse, mentre il secondo fattore costituisce una normalizzazione rispetto al degree del nodo associato ed ha lo scopo di mettere “sullo stesso piano” i vari nodi.

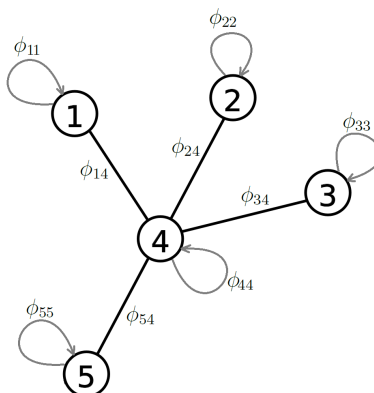


Figura 3.1: Network di esempio in cui si è aggiunto un loop per ogni nodo presente.

La richiesta¹ $\phi_{jj} = 1$ deriva invece dalla necessità di evitare situazioni patologiche per i nodi periferici. In riferimento alla figura (3.1), supponiamo infatti che ϕ_{54} sia molto piccola rispetto a $\phi_{14}, \phi_{24}, \phi_{34}$: senza l'inclusione del loop ϕ_{55} si avrebbe $J_{54} = 1$ mentre in realtà nelle ipotesi fatte tale prossimità deve essere piccola.

3.1.2 Generazione del rumore

In questa sezione vogliamo trovare una descrizione per il rumore $\eta'(t)$ da usare in (3.1). Il rumore scelto per la nostra simulazione è di tipo gaussiano, e per la sua implementazione faremo riferimento a [3]. Esso rispecchia con discreta fedeltà le caratteristiche dei trend economici senza comportare grosse difficoltà computazionali. Questo rumore è esprimibile da una funzione

$$\eta'(t) = \eta(t) + \langle \eta'_i(t) \rangle_t = \eta(t) + \mu, \quad (3.3)$$

in cui per praticità è stato isolato il suo valor medio (identificato tramite μ). Vogliamo ora esporre una procedura che permetta di costruire il rumore $\eta(t)$, caratterizzato da un valor medio nullo e covarianza secondo

$$\langle \eta(t)\eta(t') \rangle = \frac{\sigma^2}{2\tau} e^{-|t-t'|/\tau}. \quad (3.4)$$

Per la generazione di un rumore soddisfacente tali richieste è sufficiente sfruttare il seguente algoritmo: presi in considerazione un insieme G_n valori indipendenti distribuiti normalmente, si consideri la trasformazione

$$\begin{aligned} \eta(t_0) &= \frac{\sigma}{\sqrt{2\tau}} G_0 \\ \eta(t_n) &= \rho_n \eta(t_{n-1}) + \sqrt{1 - \rho_n^2} \frac{\sigma}{\sqrt{2\tau}} G_n \end{aligned} \quad (3.5)$$

con $\rho_n = \exp(-|t_n - t_{n-1}|/\tau)$. Sfruttando che $\langle G_n \rangle = 0$ e $\langle G_n G_m \rangle = \delta_{mn}$ si verifica facilmente che $\langle \eta(t_n) \rangle = 0$ e $\langle \eta(t_m)\eta(t_n) \rangle = \frac{\sigma^2}{2\tau} \exp(-|t_m - t_n|/\tau)$. Inoltre, poiché le $\eta(t_n)$ sono generate da combinazioni lineari di variabili casuali normali, anch'esse sono distribuite normalmente.

Infine, per quanto riguarda la generazione delle G_n aventi distribuzione normale, si è sfruttata la trasformazione di Box-Muller [21]: date U_1, U_2 variabili casuali indipendenti uniformemente distribuite in $]0, 1]$, la trasformazione

¹Un'altra possibile definizione, per certi versi più soddisfacente, è $\phi_{jj} = \max\{1 - \sum_k \phi_{jk}, 0\}$: in questo modo vengono aggiunti loop solo a quei nodi non sufficientemente connessi alla network.

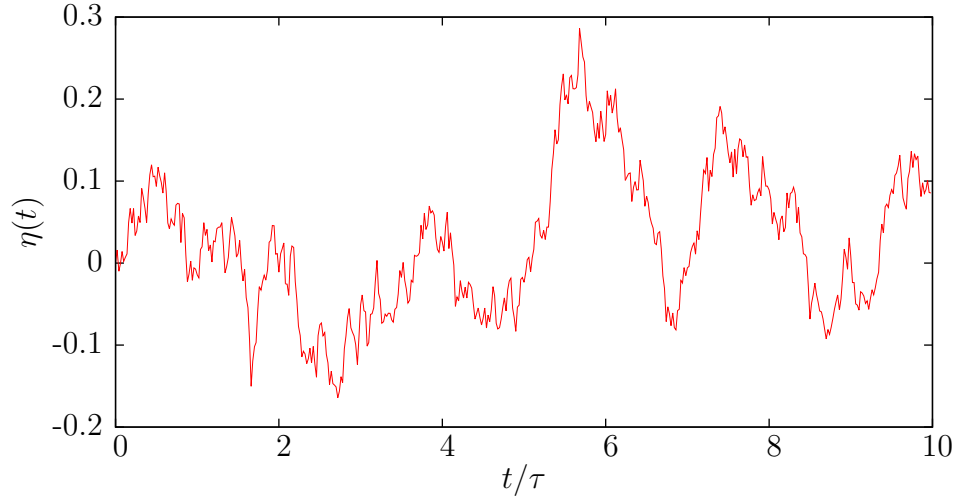


Figura 3.2: Esempio di rumore $\eta(t)$ per $\sigma = 1$ e $\tau = 50$.

$$\begin{cases} G_0 = \sqrt{-2 \ln U_1} \cos(2\pi U_2) \\ G_1 = \sqrt{-2 \ln U_1} \sin(2\pi U_2) \end{cases}$$

mappa U_1, U_2 in variabili casuali indipendenti aventi distribuzione normale con deviazione standard unitaria.

3.2 Analisi della dinamica

Ora che tutte le quantità presenti in (3.1) sono state definite, al fine di capire quali risultati ci aspettiamo e il modo in cui essi dovranno essere interpretati, studiamo l'equazione per la dinamica imposta in casi particolari di facile analisi.

3.2.1 Rumore nullo

Per prima cosa vogliamo capire che tipo di dinamica si ottenga imponendo in (3.1) $\eta'_i(t) = 0$ per ogni nodo. In tale caso l'equazione da risolvere è

$$\frac{d}{dt} Z_i(t) = \sum_{j \neq i} J_{ij} Z_j(t) - \sum_{j \neq i} J_{ji} Z_i(t). \quad (3.6)$$

Quest'equazione si presta ad essere riscritta attraverso l'uso della matrice Laplaciana (A.1.6), come ora illustreremo. Definita la matrice di degree (A.1.5) come $D_{ij} =$

$\delta_{ij} \sum_{k \neq i} \phi_{ik}$ e quella di adiacenza (A.1.4) come $A_{ij} = -\phi_{ij}$, è possibile definire la matrice Laplaciana

$$L = D - A = \begin{pmatrix} \sum_{k \neq 1} \phi_{1k} & -\phi_{12} & -\phi_{13} & \cdot & \cdot & \cdot \\ -\phi_{21} & \sum_{k \neq 2} \phi_{2k} & -\phi_{23} & \cdot & \cdot & \cdot \\ -\phi_{31} & -\phi_{32} & \sum_{k \neq 3} \phi_{3k} & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix}.$$

Posto $\tilde{Z}(t) = \left(\frac{Z_1(t)}{\sum_k \phi_{1k}}, \dots, \frac{Z_N(t)}{\sum_k \phi_{Nk}} \right)$ l'equazione (3.6) può essere riscritta come

$$\dot{\tilde{Z}}(t) = -J^* L \tilde{Z}(t). \quad (3.7)$$

Le soluzioni di (3.7) sono analoghe a quelle del sistema $Z(t) = -LZ(t)$, che è studiato in appendice, ove però le quantità $Z_i(t)$ invece che venire distribuite uniformemente su tutta la network, vengono ripartite concordemente con i pesi applicati al vettore $\tilde{Z}(t)$, ovvero secondo il degree del nodo. Per $t \rightarrow \infty$ è facile vedere che le soluzioni di (3.6) sono date da

$$\bar{Z}_i = \lim_{t \rightarrow \infty} Z_i(t) = \frac{\sum_j \phi_{ji}}{\frac{1}{2} \sum_{j \neq i} \phi_{ji} + \sum_j \phi_{jj}} \sum_j Z_j(0), \quad (3.8)$$

ove le sommatorie si intendono estese a tutti i nodi della network. Nei prossimi capitoli sarà chiaro che legami vi siano fra le \bar{Z}_i e la crescita dei prodotti quando $\eta_i(t) \neq 0$. In particolare, si vedrà come il logaritmo di tali quantità sia correlato alla velocità di crescita dei prodotti.

3.2.2 Rumore non nullo

Consideriamo adesso l'equazione (3.1) nella sua interezza, prendendo come rumore agente la funzione $\eta_i(t)$ definita nel paragrafo 3.1.2 (ponendo quindi $\mu = 0$) e cerchiamo di descriverne le caratteristiche. Osserviamo anzitutto che, se ad un tempo t_0 si ha $Z_i(t_0) > 0$ per ogni nodo i della rete, allora si avrà $Z_i(t) > 0$ per ogni tempo t . Infatti, essendo i coefficienti J_{ij} strettamente positivi si ha anche $\sum_{j \neq i} J_{ij} Z_j(t) > 0$, e

conseguentemente

$$\frac{dZ_i(t)}{dt} = \sum_{j \neq i} J_{ij} Z_j(t) + Z_i(t) \left[\eta_i(t) - \sum_{j \neq i} J_{ji} \right] > Z_i(t) \left[\eta_i(t) - \sum_{j \neq i} J_{ji} \right], \quad (3.9)$$

ove l'ultimo termine della disuguaglianza (3.9) ammette soluzioni di tipo esponenziale (e quindi strettamente positive). Lo studio generale di (3.1) da un punto di vista analitico presenta notevoli difficoltà e soluzioni approssimate sono disponibili solo nei casi più semplici (si veda ad esempio [13], in cui viene presentata una soluzione analitica per un reticolo unidimensionale in cui $J_{ij} \equiv J$). Nel nostro caso, l'interesse principale nell'analisi di (3.1) verte nel ricavare sotto quali condizioni la crescita delle $Z_i(t)$ avviene più rapidamente. Questa crescita globale può essere misurata asintoticamente [13] studiando il limite

$$c(J^*, \tau, \sigma) = \lim_{t \rightarrow \infty} \frac{1}{Nt} \sum_{i=1}^N \ln \left(\frac{Z_i(t)}{Z_i(0)} \right). \quad (3.10)$$

È facile capire perché a tempi sufficientemente lunghi, tutti i nodi della rete debbano evolvere con la stessa velocità di crescita c . Definito $c_k(t) = \frac{1}{t} \ln(Z_k(t)/Z_k(0))$, supponiamo infatti che esista un nodo k per il quale si abbia $c_k(t) - c(t) > \epsilon$ ad un tempo $t > t_M$. In tali ipotesi $Z_k(t) = \exp(c_k(t)t)$ mentre gli altri nodi mediamente evolvono secondo $Z(t) = \exp(c(t)t)$. Di conseguenza $Z_k(t) - \sum_{i \neq k} Z_i(t)$ cresce esponenzialmente all'aumentare di t e il termine di scambio $\sum_{j \neq k} [J_{kj} Z_j(t) - J_{jk} Z_k(t)]$ in (3.1), è approssimabile con $-\sum_{j \neq k} J_{jk} Z_k(t)$. Allora la dinamica per il nodo k -esimo si può scrivere come

$$\frac{d}{dt} Z_k(t) = - \sum_{j \neq k} J_{jk} Z_k(t) + \eta_k(t) Z_k(t).$$

Posto $J = \sum_{j \neq k} J_{jk}$, ricordando che $J_{ij} \geq 0$ per ogni coppia di nodi, la soluzione per il nodo k -esimo è data da un decadimento esponenziale (su cui agisce ovviamente anche il contributo dovuto ad $\eta(t)$). A conclusioni analoghe si arriva nel caso opposto, in cui $c(t) - c_k(t) > \epsilon$.

Il valore di (3.10) in funzione di J^* , τ e σ permette di capire quali possano essere le strategie da attuare al fine di aumentare la crescita della network in esame e per questa ragione la sua valutazione giocherà un ruolo importante nella nostra analisi. In [13] si mostra come questa velocità abbia un valore massimo $c_{max} = \sigma^2/2$ e uno studio perturbativo (in cui $J_{ij} = J$) permette di ricavare $c(J) \approx \sigma^2/4\sqrt{J\tau}$ per $J \rightarrow \infty$. Facciamo presente che (3.10) fornisce una stima di crescita asintotica, mentre quello che si vorrebbe conoscere è quale sia lo stato attuale dell'economia. Questa difficoltà verrà in parte superata successivamente quando cercheremo di calibrare il nostro modello sui dati a nostra disposizione.

3.3 Simulazioni su network di prova

In questa sezione presentiamo i risultati relativi alle network introdotte nelle sezioni (1.2 - 1.2). Per le simulazioni si è scelto $dt = 2 \cdot 10^{-4}$ e l'integrazione è stata protratta per $2 \cdot 10^6$ step. Il valore di dt va scelto in modo da garantire la non-negatività delle $Z_i(t)$ ed è facile stimare quanto piccolo esso debba essere: prendendo $\eta(t) = 0$ e considerando un caso limite in cui $Z_i(0) > 0$ e $Z_j(0) = 0$ per ogni $i \neq j$, allora al primo passo di integrazione si ha

$$Z_i(dt) = Z_i(0) \left[1 - dt \sum_{i \neq j} J_{ji} \right] \geq 0$$

ricordando l'espressione per J_{ij} data in (3.2) si ricava che in prima approssimazione $dt < 1/J^*$. Il valore di dt scelto per le simulazioni che seguono è ben al di sotto di tale vincolo (di circa un fattore 10). Mostriamo ora i risultati ottenuti.

3.3.1 Velocità di crescita

In questa sezione vogliamo studiare tramite simulazione l'evoluzione delle $Z_i(t)$ per l'equazione (3.1) in funzione del rumore imposto e della struttura della network considerata. Mostriamo innanzitutto il comportamento di (3.10) in funzione dei J^* definiti da (3.2) per i modelli di network di Erdős, Watts-Strogatz e Barabasi. Nelle simulazioni presenti in questo capitolo si è posto per ogni prodotto $Z_i(0) = 1$, $\tau = 0.01$ e $\sigma = 1$ (vedremo poi come tali parametri influiscano sulla dinamica).

Partendo dalle network di tipo k-regolari (vedi fig. 3.3), le curve ottenute mostrano una dinamica a doppio taglio: per coefficienti di trasferimento bassi è preferibile avere una rete con molte connessioni: queste offrono più opportunità di sfruttare eventuali occasioni favorevoli. Viceversa, quando J^* è molto grande avere meno connessioni è uno scenario preferibile (questo è in parte dovuto anche alla definizione data per le ϕ_{ii} in (3.2), che prevede un loop di peso unitario con cui il nodo re-investe su se stesso il proprio "capitale" nel caso in cui esso sia debolmente connesso alla network: quando ci sono pochi link nella rete questo contributo si fa sentire maggiormente). Osserviamo che i grafici k-regolari utilizzati nella simulazione precedente sono caratterizzati da un elevato clustering, ma a causa della loro struttura non dispongono di nodi con betweenness centrality elevata. Modificando $W(256,0,4)$ attraverso il metodo di Watts-Strogatz (ovvero imponendo una probabilità di ricablaggio $p > 0$) si ottengono grafi con coefficienti di clustering leggermente minori in favore di una closeness centrality maggiore. La velocità di crescita in funzione della probabilità di ricablaggio porta alle curve in figura 3.4.

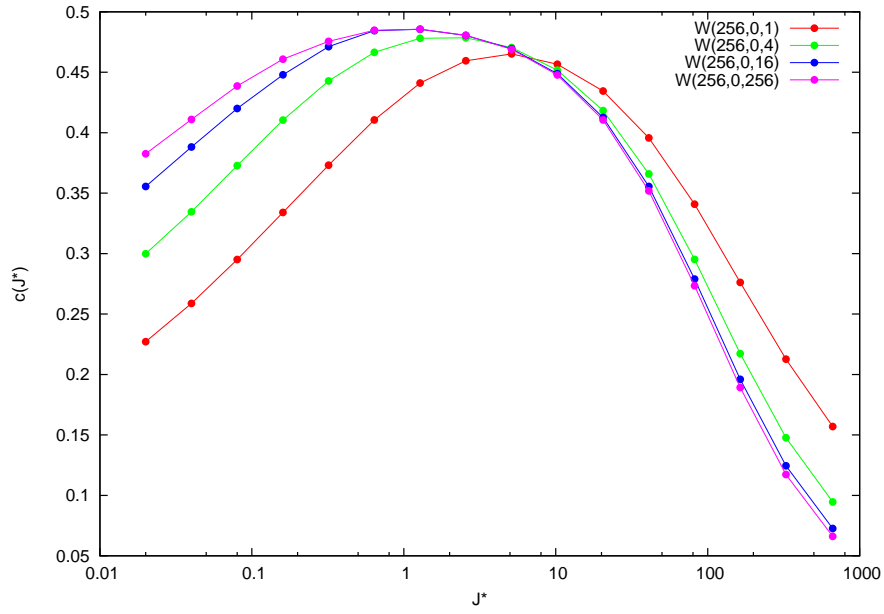


Figura 3.3: Velocità di crescita asintotica per i grafici k-regolari ($\tau = 0.01, \sigma = 1$).

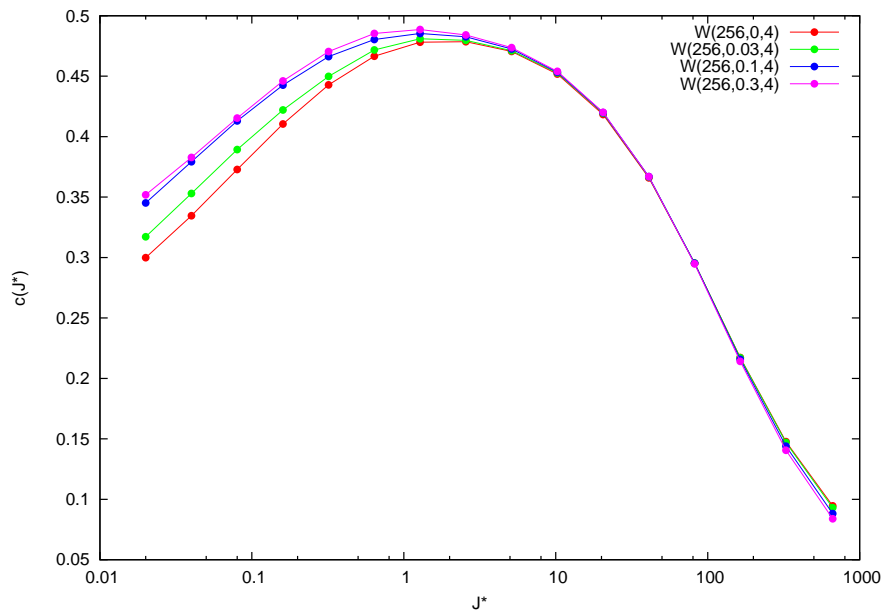


Figura 3.4: Velocità di crescita asintotica per grafi di Watts-Storgatz ($\tau = 0.01, \sigma = 1$).

Risulta che una rete small world (come lo è quella di Watts-Storgatz) è associata a tassi di crescita superiori rispetto ad un grafo k -regolare: confrontando $W(256,0,4)$ con le altre curve presenti in figura 3.4, è evidente come la velocità di crescita (specialmente quando J^* è piccolo) aumenti concordemente con la probabilità di ricablaggio. Questo è intuitivo: laddove un nodo dovesse ottenere condizioni favorevoli, esso ha la possibilità di condividere le risorse più rapidamente con le altre zone della network. Per le stesse ragioni le reti di Erdős mostrano una buona velocità di crescita (vedi figura 3.5) tant'è che $E(256,0.015)$ pur avendo ~ 500 link, ha tassi di crescita superiori a $W(256,0,4)$ che dispone del doppio delle connessioni. Per una network ad attaccamento preferenziale i risultati delle simulazioni sono invece presentati in figura 3.6: tali network, sempre in virtù della loro struttura small-world, mostrano anch'esse una buona velocità di crescita (è infatti sufficiente un numero di link molto basso per raggiungere velocità di crescita simili a quelle ottenibili su un grafo completo). In luce di queste analisi, quello che risulta è che, almeno nel caso di reti costituite da link aventi tutti lo stesso peso, la struttura della network non porta ad un cambiamento radicale della curva di crescita, che mantiene caratteristiche analoghe sulle network studiate.

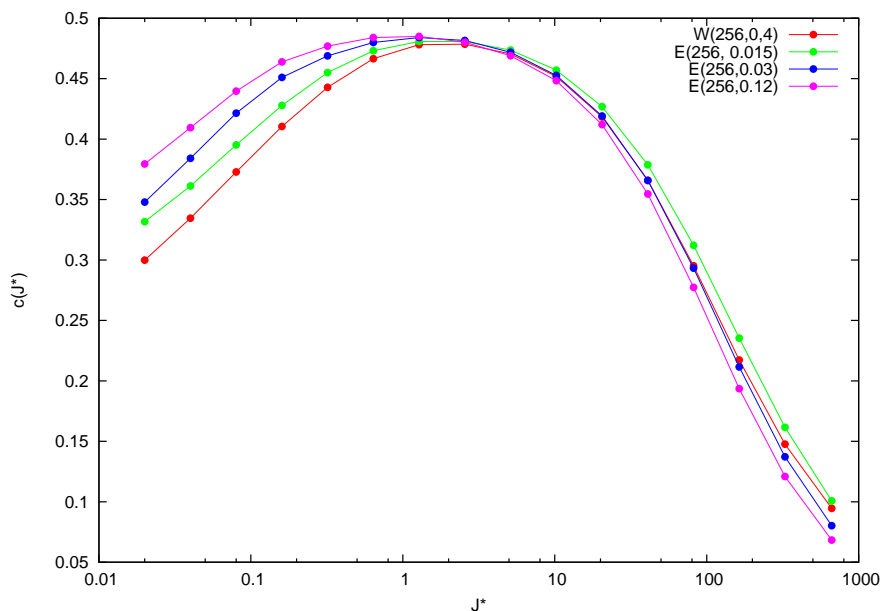


Figura 3.5: Velocità di crescita asintotica per network di Erdős ($\tau = 0.01, \sigma = 1$).

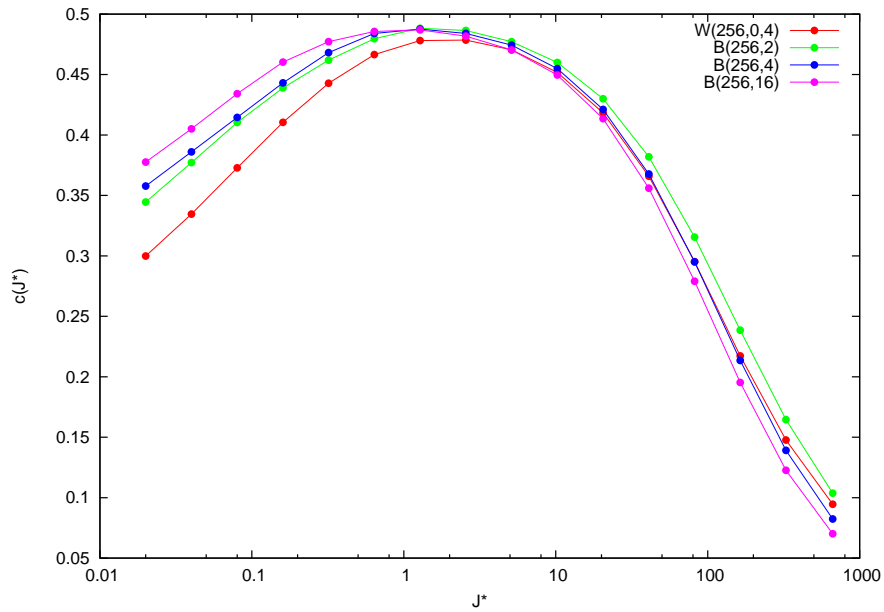


Figura 3.6: Velocità di crescita asintotica per network di Barabási ($\tau = 0.01, \sigma = 1$).

Oltre alla dipendenza dalla struttura della network, la velocità di crescita dipende dai parametri τ e σ impostati nel rumore gaussiano $\eta_i(t)$. In figura 3.7 sono riportati i risultati delle simulazioni per diversi τ e σ nella network $W(256, 0, 1)$. Come già fatto notare precedentemente, per l'equazione (3.1) risulta che la velocità di crescita per $J^* \rightarrow \infty$ è approssimata da $c(J^*, \tau, \sigma) \approx \sigma^2/4\sqrt{J^*\tau}$; questo è stato riassunto in figura 3.8, ove le curve simulate raggiungono valori di crescita analoghi (grazie al riscaldamento σ^{-2}) e collassano in un unico comportamento per valori di J^* sufficientemente grandi.

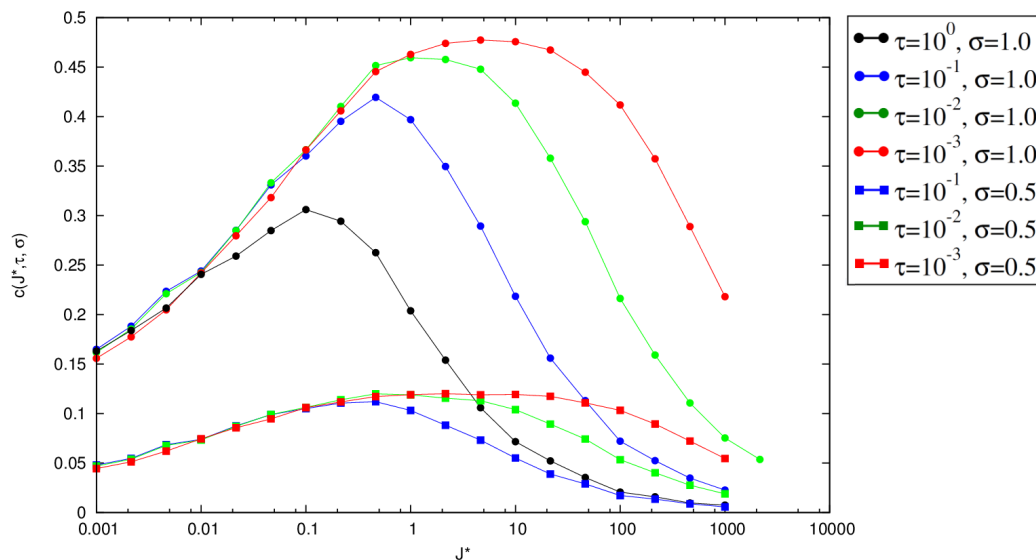


Figura 3.7: Dipendenza della velocità di crescita in funzione di τ e σ per $W(256, 0, 1)$.

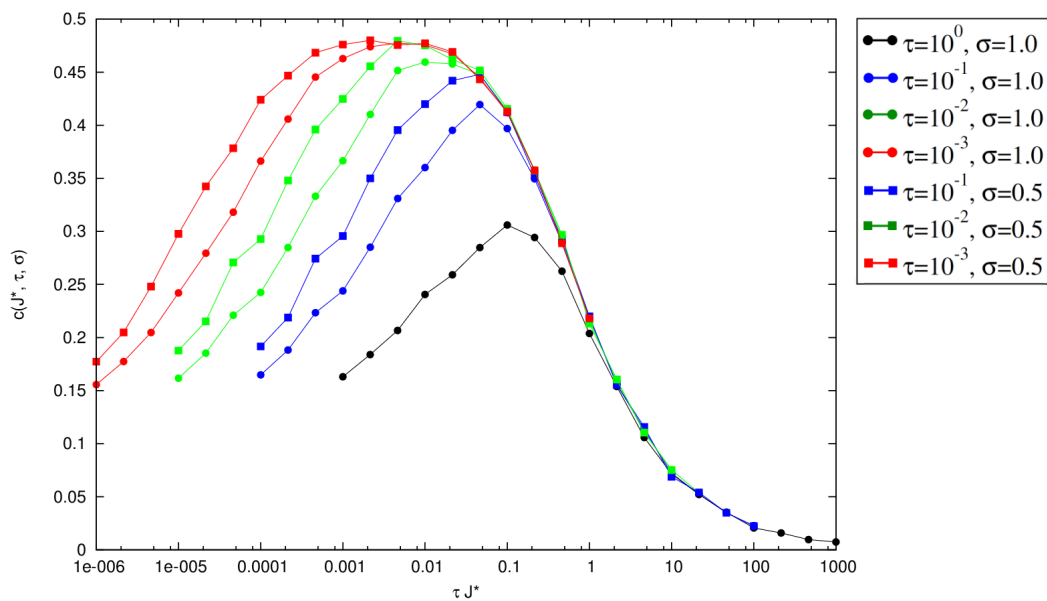


Figura 3.8: Invarianza τJ^* per la crescita della network $W(256, 0, 1)$.

3.3.2 Dinamica in reticolo unidimensionale

Uno dei casi più semplici da considerare per lo studio di (3.1) è il caso del reticolo unidimensionale e in questa sezione vogliamo analizzarne l'evoluzione delle $Z_i(t)$ al

variare del coefficiente di explore J^* . In tale network tutti i nodi sono equivalenti e non sono presenti collegamenti se non fra nodi adiacenti. Per queste ragioni è possibile, senza troppe complicazioni, osservare l'andamento delle $Z_i(t)$ nel tempo ed intuire conseguentemente il comportamento della dinamica. In figura 3.9 sono rappresentati i $c_k(t) = \ln(Z_k(t)/Z_k(0))/t$ nel caso in cui $J^* = 1, \sigma = 1, \tau = 0.01$, integrate fino a $t = 100$ con un $dt = 2 \cdot 10^{-4}$. Si osserva come, dopo un transiente iniziale della durata di circa $\sim 500\tau$, la crescita della network inizi a convergere verso il suo valore asintotico.

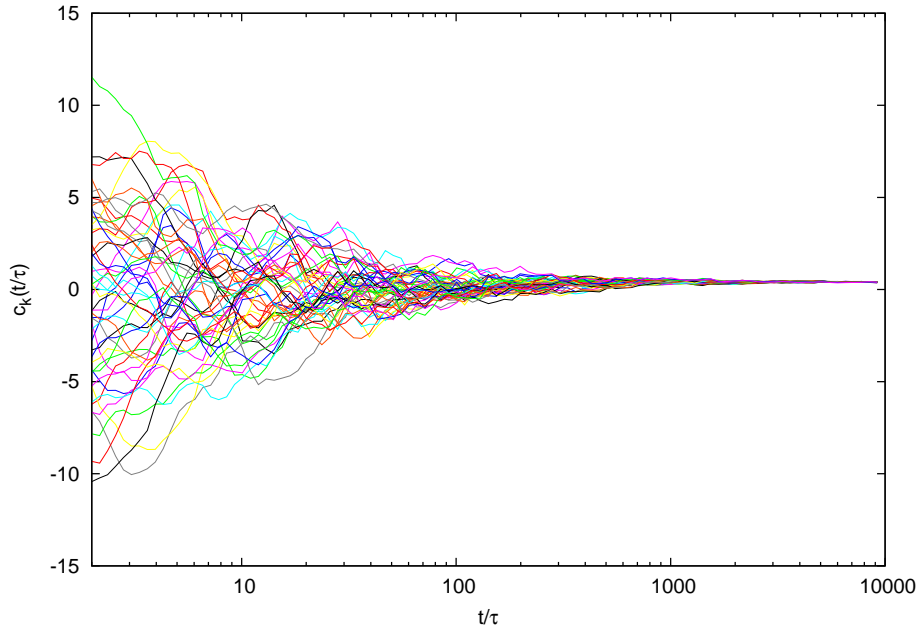


Figura 3.9: Andamento delle $c_k(t/\tau)$ per i vari nodi (solo 50 di essi sono rappresentati) per un reticolo unidimensionale costituito da 256 nodi ($\tau = 0.01, \sigma = 1, J^* = 1$).

In figura 3.10 sono invece rappresentate le quantità

$$c_k(t) - c(t) = \frac{1}{t} \left[\ln \left(\frac{Z_k(t)}{Z_k(0)} \right) - \left\langle \ln \left(\frac{Z(t)}{Z(0)} \right) \right\rangle_k \right] \quad (3.11)$$

per i nodi della rete. È ben visibile (quando il coefficiente di explore è piccolo) quali siano le zone in cui il rumore abbia avuto un ruolo favorevole (o meno) allo sviluppo delle $Z_i(t)$.

Per J^* piccoli la rete si trova evidentemente in difficoltà nel condividere le risorse (e infatti questo regime è associato a una velocità di crescita più bassa di quella ottimale,

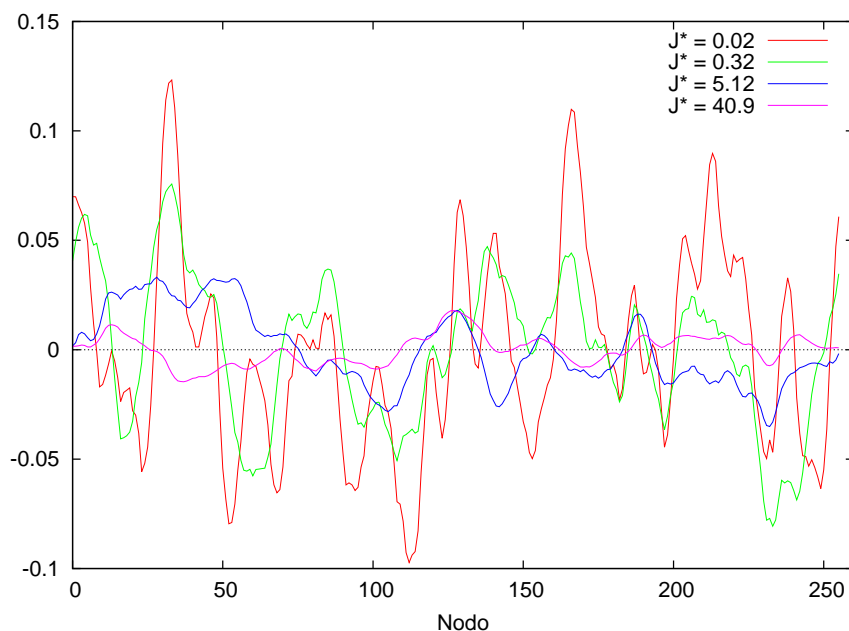


Figura 3.10: Lungo l'asse y sono rappresentate le quantità $c_k(t) - c(t)$ (al tempo $t = 1000\tau$) al variare del valore di J^* imposto nella dinamica di un reticolo unidimensionale costituito da 256 nodi. Risulta chiaro come l'effetto di J^* sia quello di ridistribuire le risorse all'interno della network. Per rendere il grafico più leggibile i valori $c_k(t) - c(t)$ sono stati mediati con i primi 5 vicini di ogni nodo ($\tau = 0.01$, $\sigma = 1$ e $Z_i(0) = 1$, $i : 1, \dots, 256$).

come visibile in figura 3.3). Via via, aumentando il coefficiente di explore ed avvicinandosi al valore di J^* associato alla velocità di crescita massima (che per $W(256,0,1)$ con le nostre definizioni si ha per $J^* \approx 5$), la crescita assume un carattere più globale e le discrepanze fra i vari nodi vengono appianate. Per valori di J^* ancora maggiori, sebbene i nodi continuino ad esibire un comportamento collettivo, la crescita della network tende ad arrestarsi: la condivisione sottrae troppe risorse al nodo, che non riesce a sfruttare in maniera ottimale i momenti in cui il rumore correlato ne favorirebbe la crescita. Questa osservazione tornerà utile nel capitolo 4, quando analizzeremo la network dei prodotti. Simulando la dinamica al variare di J^* e studiando di volta in volta la varianza dei singoli nodi sarà infatti possibile di ricavare il coefficiente di explore relativo al mercato globale.

3.4 Network dei prodotti

La crescita della network dei prodotti è riportata in figura 3.11 insieme alle simulazioni relative ad un reticolo unidimensionale e ad un grafo completo. Risulta che la network dei prodotti, per bassi valori di J^* , ha tassi di crescita maggiori rispetto ad quelli di un grafo completo. Questa differenza di comportamento è da imputare all'equivalenza in cui si trovano i vari nodi della network completa: a causa dell'uniformità di tale grafo, consegue un'uniformità nella distribuzione delle risorse nella rete, che rende più difficile lo sviluppo a J^* piccoli. Nel caso della network dei prodotti invece alcuni nodi, disponendo di un gran numero di connessioni (ed essendo con grande probabilità connessi agli altri nodi ad alto degree della network), riescono a ricevere grandi quantità di risorse nonostante il coefficiente di explore sia basso, permettendone così una crescita più sostenuta.

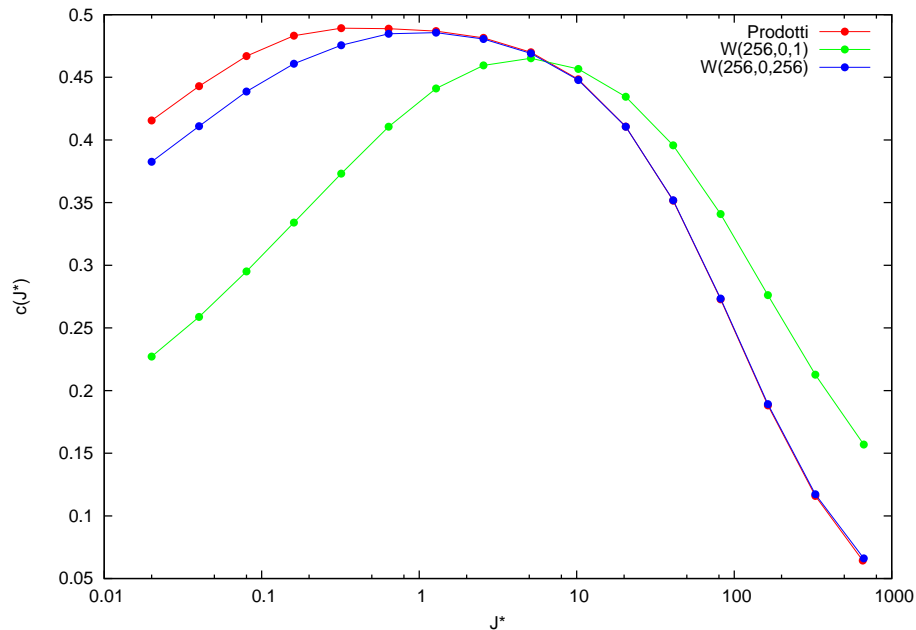


Figura 3.11: Curva di crescita per lo spazio dei prodotti ($\tau = 0.01, \sigma = 1$).

Tale osservazione suggerisce la possibilità che, nei momenti di stagnazione dei mercati, l'economia relativa a certi settori possa comunque mostrare tassi di crescita e capacità di interagire con il mercato superiori rispetto a quanto possibile per altri beni. Per capire meglio la dinamica dello spazio dei prodotti, si è studiata la velocità di crescita dei vari prodotti in funzione del loro degree. In figura 3.12, sono riportate le quantità (3.11) a vari tempi t (espressi in funzione di τ). Risulta evidente come la

velocità di crescita dei singoli prodotti dipenda anche dal degree del nodo in esame: nodi aventi degree maggiori sono associati a tassi di crescita maggiori; tale differenza è significativa nelle fasi iniziali dell'evoluzione, cioè prima che il sistema raggiunga il suo comportamento asintotico.

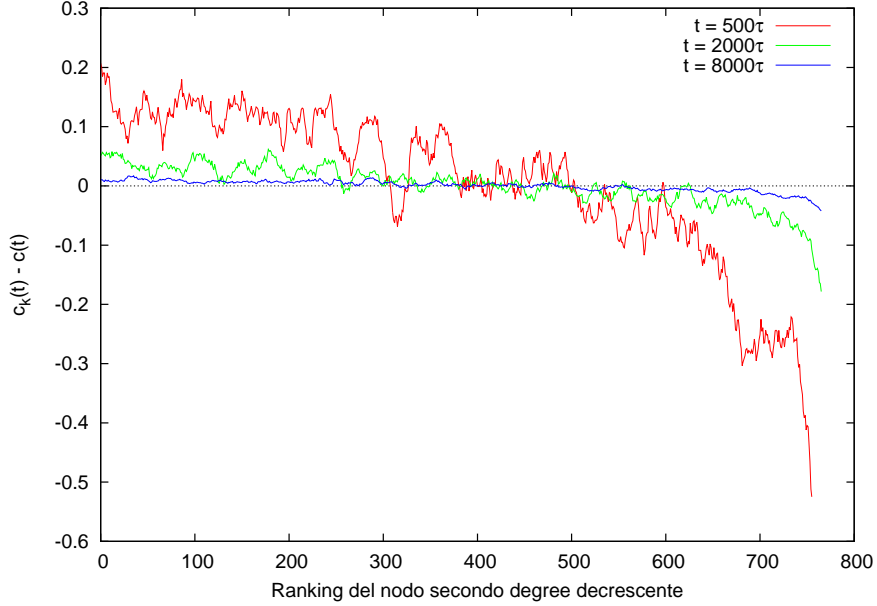


Figura 3.12: Scostamento della velocità di crescita dei prodotti rispetto al valore di crescita medio. In ascissa i prodotti sono stati ordinati concordemente con il loro degree. Per rendere il grafico più leggibile, la velocità di crescita di ogni prodotto è stata mediata con quella dei suoi primi 10 vicini ($\tau = 0.01$, $\sigma = 1$ e $Z_i(0) = 1$, $i : 1, \dots, 775$).

Questo comportamento può essere messo meglio in evidenza valutando la quantità $[c_k(t) - c(t)]t$. Tale espressione fornisce informazioni sia riguardo quale rapporto abbiano fra loro le ampiezze delle $Z_k(t)$ che sulle condizioni iniziali in cui si trovava la network al tempo $t = 0$. Infatti

$$\begin{aligned}
 [c_k(t) - c(t)]t &= \ln \left(\frac{Z_k(t)}{Z_k(0)} \right) - \frac{1}{N} \sum_{i=1}^N \ln \left(\frac{Z_i(t)}{Z_i(0)} \right) = \\
 &= \ln \left(\frac{Z_k(t)}{Z_k(0)} \right) - \ln \left(\frac{[\prod Z_i(t)]^{1/N}}{[\prod Z_i(0)]^{1/N}} \right) = \\
 &= \ln \left(\frac{Z_k(t)}{M_g(Z(t))} \right) - \ln \left(\frac{Z_k(0)}{M_g(Z(0))} \right), \tag{3.12}
 \end{aligned}$$

ove $M_g(Z(t))$ e $M_g(Z(0))$ indicano rispettivamente la media geometrica delle Z_k al tempo t e al tempo $t = 0$. Il primo termine di (3.12) fornisce quindi informazioni su quale sia l'ampiezza delle Z_i al tempo t , mentre il secondo mantiene "memoria" delle condizioni iniziali. In figura 3.13 sono rappresentate le quantità $[c_k(t) - c(t)]t$ al variare di k per la rete dei prodotti con condizioni iniziali $Z_i(0) = 1$, assieme alle soluzioni $\ln(\bar{Z}_k)$ ricavate secondo l'equazione (3.8) per la dinamica a rumore nullo. Il grafico mostra come di fatto il rapporto fra le ampiezze delle $Z_k(t)$ a rumore non nullo sia fissato (in media) dalle soluzioni (3.8): i coefficienti di trasferimento nel caso $\eta_i(t) \neq 0$ agiscono cioè cercando di ridistribuire le risorse in modo da mantenere i rapporti fra le ampiezze delle $Z_i(t)$ secondo quanto determinato dal caso a rumore nullo.

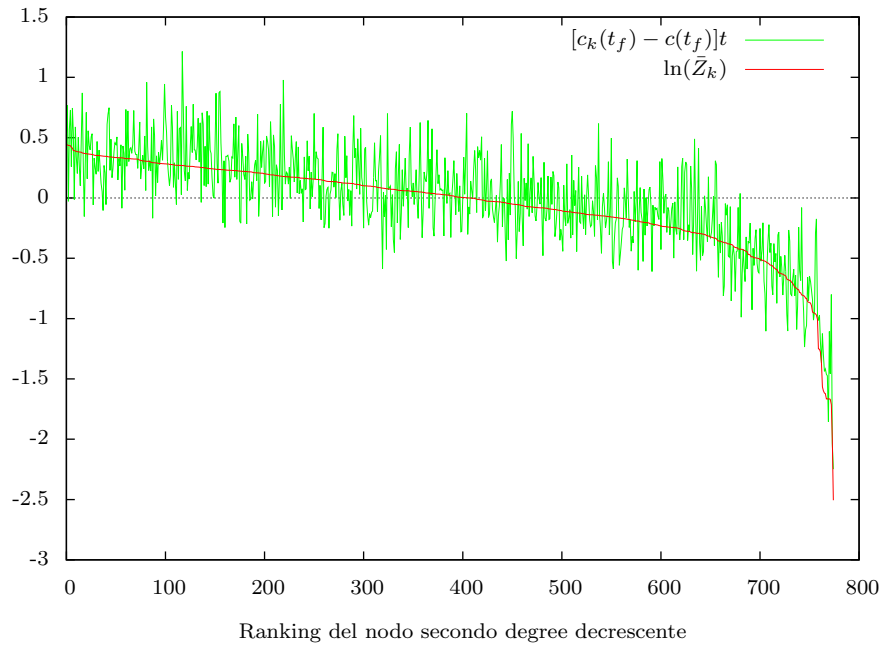


Figura 3.13: In verde è rappresentata la quantità $[c_k(t) - c(t)]t$. La curva rossa è invece il logaritmo della soluzione di (3.8), ovvero le quantità $\ln(\bar{Z}_i)$ per i vari nodi della network. ($\sigma = 1$, $\tau = 0.01$, $J^* = 5$, $t_f = 100\tau$, $Z_i(0) = 1$, $i : 1, \dots, 775$).

Capitolo 4

Calibrazione della network

Come si è visto, la velocità di crescita della network è esprimibile attraverso una funzione $c(J^*, \tau, \sigma)$. In questo capitolo ci proponiamo di individuare quali siano i valori di J^* , τ e σ che caratterizzano il mercato globale. Per raggiungere tale obiettivo sarà necessario effettuare varie restrizioni ed assunzioni sui prodotti e sulla loro evoluzione; questo è in parte dovuto alla copertura non totale dei dati a nostra disposizione (per certe merci lo storico 1962-2000 risulta incompleto) ed in parte alle difficoltà nel risalire con precisione ai parametri della dinamica avendo conoscenza soltanto delle $Z_i(t)$, che sono campionate ad intervalli di tempo annuali.

4.1 Stime di τ e σ

In questa sezione esponiamo un approccio attraverso cui è possibile stimare τ e σ . In questo capitolo, consideriamo un rumore agente nella network con media non nulla, cioè la presenza di un rumore del tipo

$$\eta'(t) = \eta(t) + \mu$$

con $\eta(t)$ rumore gaussiano a media nulla (come esposto nella sezione 3.1.2) e μ costante reale. L'equazione (3.1) si riscrive come

$$\frac{\partial Z_i(t)}{\partial t} = \sum_{j \neq i} J_{ij} Z_j(t) - \sum_{j \neq i} J_{ji} Z_i(t) + \eta_i(t) Z_i(t) + \mu Z_i(t). \quad (4.1)$$

La difficoltà principale nel risalire ai parametri associati al rumore è dovuta ai termini di trasferimento fra i vari prodotti: si tratta infatti di risolvere in $\eta_i(t)$ l'equazione

$$\frac{d}{dt} \ln Z_i(t) - \frac{1}{Z_i(t)} \left(\sum_{j \neq i} J_{ij} Z_j(t) - \sum_{j \neq i} J_{ji} Z_i(t) \right) = \eta_i(t) + \mu. \quad (4.2)$$

L'espressione (4.2) permetterebbe (senza grosse difficoltà) di risalire al rumore agente sui vari nodi del sistema. Tuttavia, applicare direttamente tale equazione ai dati in nostro possesso non porta ad alcun risultato: nel dataset NBER sono registrati i flussi di export con cadenza annuale, e tali intervalli di tempo non sono evidentemente adatti ad un calcolo differenziale. Si è così costretti a valutare l'integrale di (4.2), ovvero la quantità

$$\int_0^t \frac{d}{dt'} \ln Z_i(t') dt' - \int_0^t \frac{1}{Z_i(t')} \sum_{j \neq i} [J_{ij} Z_j(t') - J_{ji} Z_i(t')] dt = \int_0^t (\eta_i(t') + \mu) dt'. \quad (4.3)$$

Il membro a sinistra dell'uguaglianza (4.3) è ricavabile con i dati a nostra disposizione e, detta $K(t)$ tale quantità, il problema è ricondotto al determinare il tempo caratteristico e la varianza del rumore da un'equazione del tipo

$$K(t) = \int_0^t (\eta_i(t') + \mu) dt'.$$

Supponendo che il rumore agente sul mercato globale sia analogo a quello scelto nella sezione 3.1.2 è possibile risalire a τ e σ studiando la varianza di $\int_0^t (\eta_i(t') + \mu) dt'$. Infatti si ha

$$\begin{aligned} \text{var} \left\{ \int_0^t (\eta_i(t') + \mu) dt' \right\} &= \left\langle \left(\int_0^t \eta_i(t') dt' \right)^2 \right\rangle - \left\langle \int_0^t \eta_i(t') dt' \right\rangle^2 = \\ &= \left\langle \int_0^t \eta_i(t') dt' \int_0^t \eta_i(t'') dt'' \right\rangle = 2 \int_0^t dt' \int_{t'}^t \langle \eta_i(t') \eta_i(t'') \rangle dt'' = \\ &= 2 \int_0^t dt' \int_{t'}^t \frac{\sigma^2}{2\tau} e^{-\frac{t-t''}{\tau}} dt'' = \sigma^2 t + \sigma^2 \tau (e^{-t/\tau} - 1). \end{aligned} \quad (4.4)$$

Tale funzione fornisce un possibile fit a due parametri per τ e σ . Re-integrando quanto ottenuto si trova

$$\int_0^t \left[\sigma^2 t' + \sigma^2 \tau (1 - e^{-t'/\tau}) \right] dt' = \sigma^2 t \left(\frac{1}{2} t - \tau \right) - \sigma^2 \tau^2 (e^{-t/\tau} - 1). \quad (4.5)$$

Il rapporto fra (4.4) e (4.5), grazie alla semplificazione in σ^2 , permette di eseguire un fit ad un parametro su τ rendendone quindi più facile la stima.

4.2 Analisi dei dati reali

Vogliamo ora studiare quali informazioni è possibile ricavare partendo dai dati disponibili nel NBER. Anzitutto è necessaria un'ulteriore restrizione sul set di prodotti: dei

775 presi in esame, solo per 508 di essi è disponibile lo storico dal 1962 al 2000. Per tali prodotti si può calcolare direttamente la velocità di crescita semplicemente studiandone il valore nel corso degli anni, che restituisce

$$c(t_f) = \left\langle \frac{1}{t_f - t_i} \ln \frac{Z_k(t_f)}{Z_k(t_i)} \right\rangle_k \approx 0.0829 \text{ anni}^{-1} \quad t_i = 1962, \quad t_f = 2000.$$

Come prima analisi valutiamo quanto i prodotti reali seguano i trend suggeriti dalla network simulata. Come visto nella sezione 3.4, le merci associate a nodi aventi degree maggiori hanno velocità di crescita maggiori (e questa differenza è tanto più marcata tanto più ci si trova lontano dal comportamento asintotico della network).

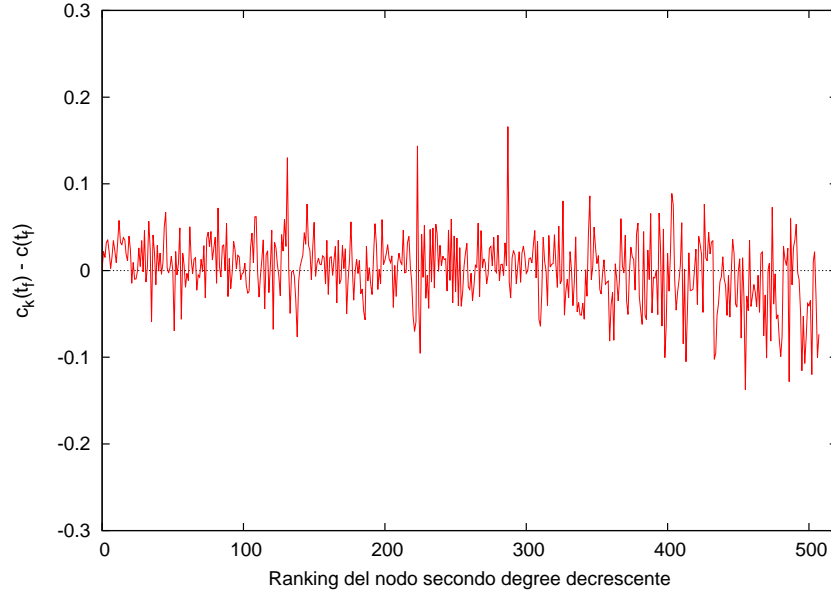


Figura 4.1: crescite $c_k(t_f) - c(t_f)$ ($t_f = 2000$) dei prodotti, ricavate dal dataset NBER per le merci coprenti gli anni 1962-2000.

In figura 4.1 sono rappresentate le quantità $c_k(2000) - c(2000)$ ricavate attraverso i dati a nostra disposizione per i 508 prodotti selezionati (che sono ordinati in funzione del loro degree lungo l'asse x). È presente una certa affinità con le curve simulate (si veda la figura 3.12 ottenuta dalle simulazioni sulla network dei prodotti). Il raffronto si fa più evidente andando a mediare $c_k(t_f) - c(t_f)$ fra primi vicini, come fatto in figura 4.2. Questo primo risultato è incoraggiante: i prodotti identificati dal nostro modello come centrali, mediamente hanno tassi di crescita maggiori rispetto a quelli periferici, come previsto.

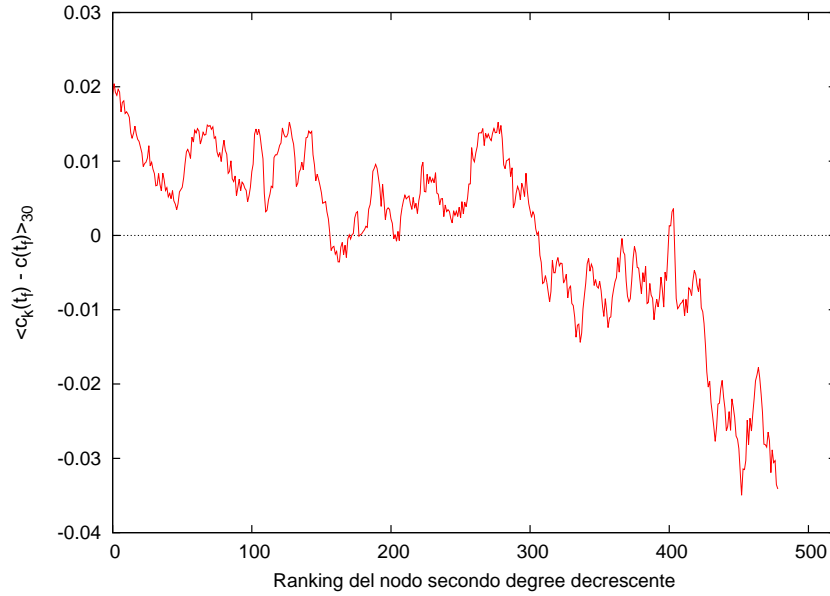


Figura 4.2: crescite $c_k(t_f) - c(t_f)$ ($t_f = 2000$) relative ai prodotti registrati nel dataset NBER coprenti gli anni 1962-2000, mediate su 30 prodotti.

4.3 Calibrazione in τ e σ

Come mostrato nella sezione 4.1, sfruttando l'equazione (4.3) è teoricamente possibile risalire ai parametri associati al rumore. Anzitutto, per ricavare τ e σ è stato necessario effettuare un'ulteriore restrizione sul set di prodotti. L'accorgimento è stato necessario per escludere dall'analisi merci per le quali l'andamento registrato nel dataset non è particolarmente convincente né concorde con il rumore che è stato introdotto nel nostro modello. A titolo di esempio riportiamo gli andamenti del prodotto 6811 (*Silver-including base metals clad with silver-, unwrought, unworked or semi-manufactured*) negli anni 1993-2000.

year	1993	1994	1995	1996	1997	1998	1999	2000
value k\$	2240392	2781378	516	3505902	1196	2920	424	3692810

Tabella 1. Andamento export mondiale per il prodotto 6811.

Le variazioni di export di tale prodotto sono incompatibili con il nostro rumore (e forse dovute a carenze del dataset). Per tale ragione questi andamenti sono stati esclusi

dall'analisi¹. In tale modo sono rimaste 346 merci d'interesse sulle quali si è valutata l'equazione (4.3) nella speranza di risalire al rumore $\eta_i(t)$. Tuttavia questo non ha portato ad alcun risultato: le restrizioni che si sono dovute fare sulla network (scartandone quasi la metà a causa delle carenze nel dataset) rendono impossibile la risoluzione rigorosa di (4.3). Per superare il seguente problema si è ragionato quindi sull'equazione (3.1). Data una $Z_i(t) = Z_i(0) \exp[c(t)t]$, si ha che $\dot{Z}_i(t) = [c_i(t) + \dot{c}_i(t)t]Z_i(t)$, e quindi l'equazione per la dinamica può essere riscritta come

$$c_i(t) + \dot{c}_i(t)t = \frac{1}{Z_i(t)} \sum_{j \neq i} [J_{ij}Z_j(t) - J_{ji}(t)Z_i(t)] + \eta_i(t). \quad (4.6)$$

Simulando la network dei prodotti, è possibile studiare i 4 termini dell'equazione sopra singolarmente: in questo modo si è evidenziato (vedi figura 4.3) come il termine di scambio abbia (il più delle volte) un ruolo marginale nella valutazione di $\text{var} \left\{ \int \eta(t) dt \right\}$. L'equazione (4.6) è stata quindi approssimata come:

$$c_i(t) + \dot{c}_i(t)t \approx \eta_i(t), \quad (4.7)$$

ovvero, integrando

$$\int_0^t \frac{d}{dt'} \ln(Z_i(t')) dt' \approx \int_0^t \eta_i(t') dt'$$

In questo modo è stato possibile fittare i dati a nostra disposizione secondo (4.4) e (4.5), ottenendo $\tau = 3.15 \pm 0.64$ anni e $\sigma = 0.346 \pm 0.005$ (vedi figura 4.4). Restano ora da individuare i parametri μ e J^* , di cui ci occuperemo nella prossima sezione.

¹Si sono esclusi dalla calibrazione tutti quei prodotti che, a distanza di un anno, mostravano cambiamenti di flusso che eccedeva un fattore 10, ovvero se $|\log_{10} Z_i(t) - \log_{10} Z_i(t-1)| > 1$, $t \in [1963, 2000]$.

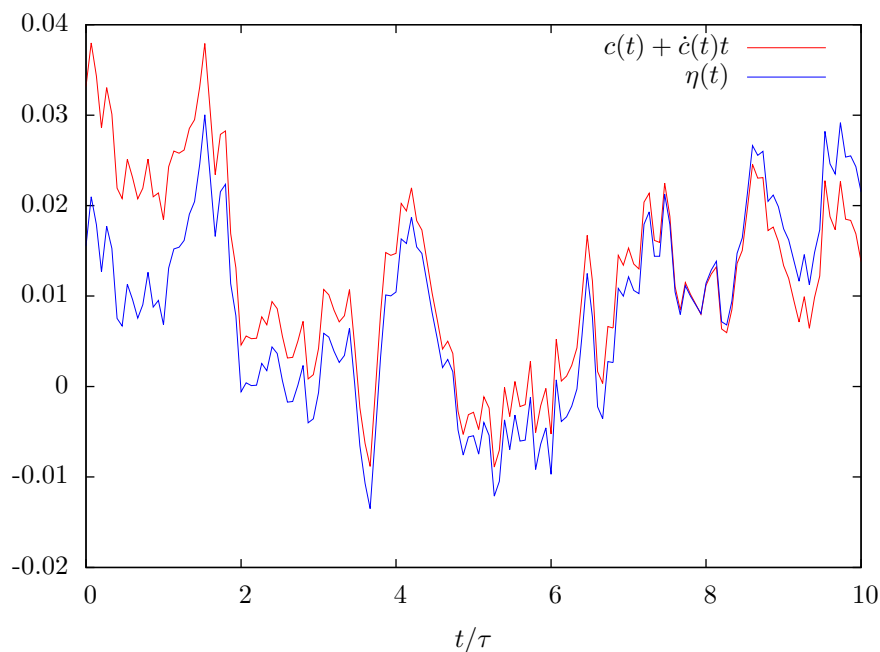


Figura 4.3: In questo grafico è mostrata, a titolo di esempio, una simulazione della dinamica sulla network ristretta a 346 prodotti al fine di valutare quanto l'approssimazione $c_i(t) + \dot{c}_i(t)t \approx \eta_i(t)$ possa ritenersi valida per i nostri propositi. La merce visualizzata è la 1123 (*beer made from malt*). Nella simulazione si sono impostati come parametri i valori $\sigma = 0.035$, $\tau = 3$ e $J^* = 0.1$ usando come condizioni iniziali $Z_i(0) = Z_i(1962)$.

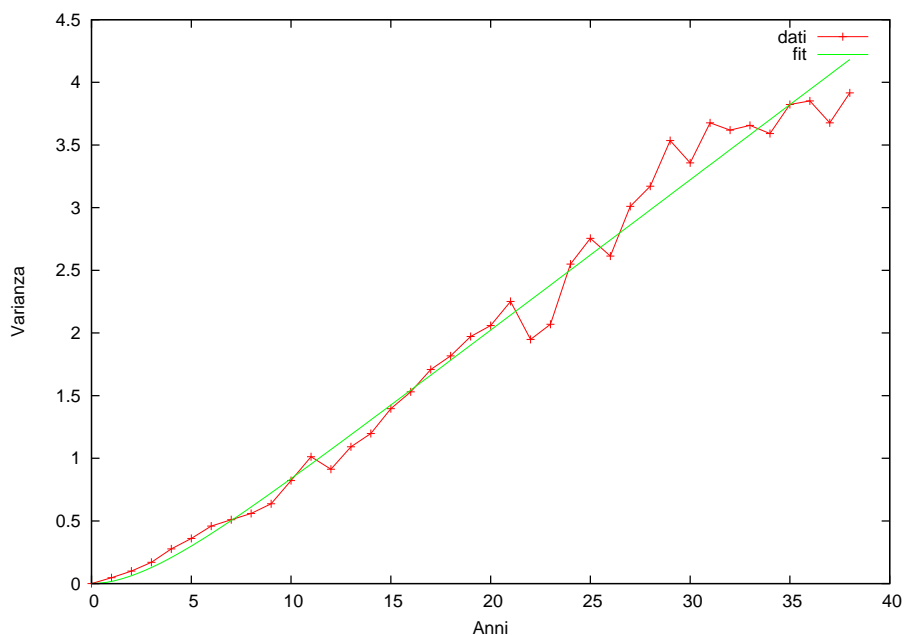


Figura 4.4: Fit dell'equazione (4.4) sui dati reali relativi alla network ristretta a 346 prodotti.

4.4 Calibrazione in μ e J^*

A questo punto restano due parametri liberi su cui fissare la dinamica, il fattore di crescita μ e il coefficiente di explore J^* . Oltre ai già noti problemi nei dati a nostra disposizione, l'individuazione di questi due parametri non è banale a causa della scala dei tempi su cui ci troviamo ad analizzare i dati, lontana dalle condizioni asintotiche di crescita. Per queste ragioni, le stime che andremo a fare avranno principalmente carattere qualitativo. Anzitutto, è necessario individuare per quali coppie di valori (μ, J^*) , la velocità di crescita della network simulata risulti analoga a quella registrata nel dataset. A tal fine si è simulata l'evoluzione di (4.1) per 38 anni al variare dei parametri (μ, J^*) , imponendo come condizioni iniziali $Z_i(0) = Z_i(1962)$ per ogni prodotto. Per individuare i valori che meglio approssimano la crescita, si sono studiati i minimi della funzione

$$f(\mu, J^*) = \sum_{i=1}^{346} [(c_i^r(t_f) - c_i^s(t_f))]^2 \quad t_f = 38 \text{ anni}, \quad (4.8)$$

in cui $c_i^r(t_f)$ e $c_i^s(t_f)$ rappresentano rispettivamente la velocità di crescita reale (cioè ricavata dai dati a nostra disposizione) e quella simulata relative al bene i -esimo dopo 38 anni. Per ogni coppia casuale di valori (μ, J^*) sono state effettuate N_s simulazioni di (4.1), con N_s sufficientemente grande da rendere l'errore sulla media $\langle f(\mu, J^*) \rangle$ minore del 10%². I risultati ottenuti da questo studio sono riportati in figura 4.5.

Fra i possibili valori di (μ, J^*) capaci di generare una velocità di crescita analoga a quella osservata per i dati reali, la scelta deve ricadere in quelli capaci di riprodurre in maniera più fedele gli andamenti delle merci, riportati in figura 4.6.

A tal fine, per ogni anno registrato nel dataset, si è studiata la varianza della quantità $\ln Z(t)$ per i vari prodotti, ottenendo $var \{\ln Z_k(t)\}_k = 2.05 \pm 0.33$. Effettuato uno studio della varianza $var \{\ln Z_k(t)\}_k$ in funzione del valore di J^* imposto nella simulazione (si veda la fig. 4.7), è stato possibile individuare il valore di J^* che meglio riproduce i dati reali. In questo modo si è ricavato $J^* \approx 0.008$. Conseguentemente, facendo riferimento alla fig. 4.5, si ha $\mu \approx 0.055$. Impostati nella dinamica i parametri $(\sigma, \tau, \mu, J^*) = (0.346, 3.15, 0.055, 0.008)$ sono stati ricostruiti gli andamenti degli export, che sono riportati in fig. 4.8. La velocità di crescita asintotica della network in funzione del coefficiente di explore J^* è invece riportata in figura 4.9.

²Questo si ottiene, in genere, effettuando circa 25 simulazioni distinte per ogni coppia di valori (μ, J^*) scelti.

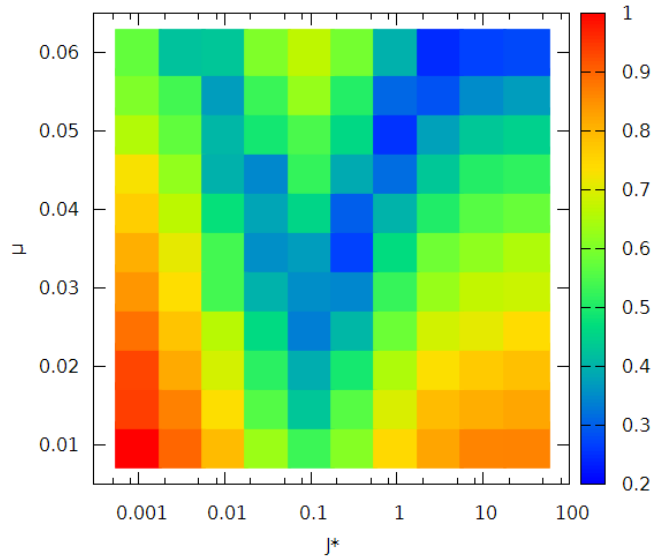


Figura 4.5: Minimizzazione della funzione (4.8) al variare dei parametri μ e J^* per la network dei prodotti.

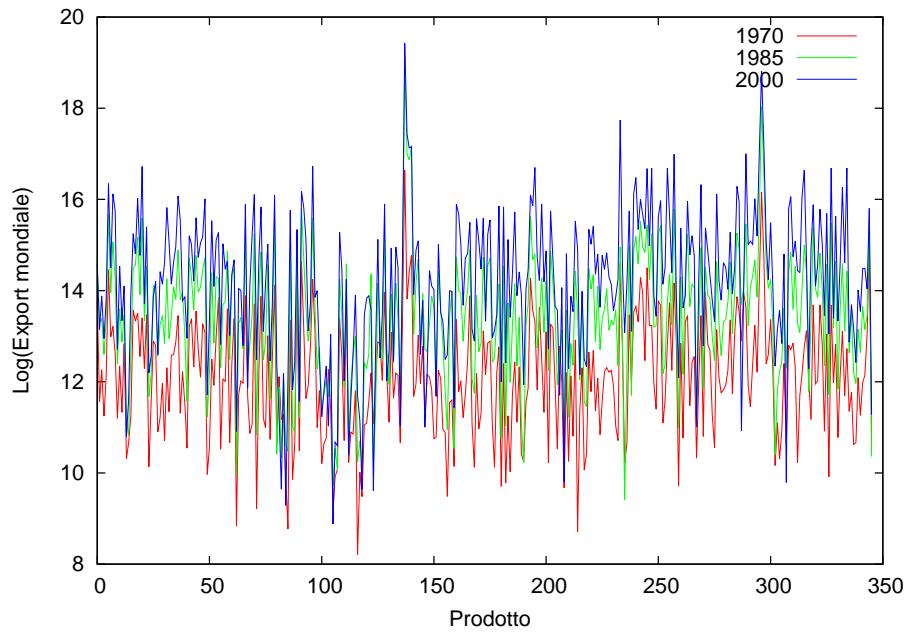


Figura 4.6: Valori di Export registrati nel dataset NBER per le 346 merci selezionate, valutati a intervalli di 15 anni.

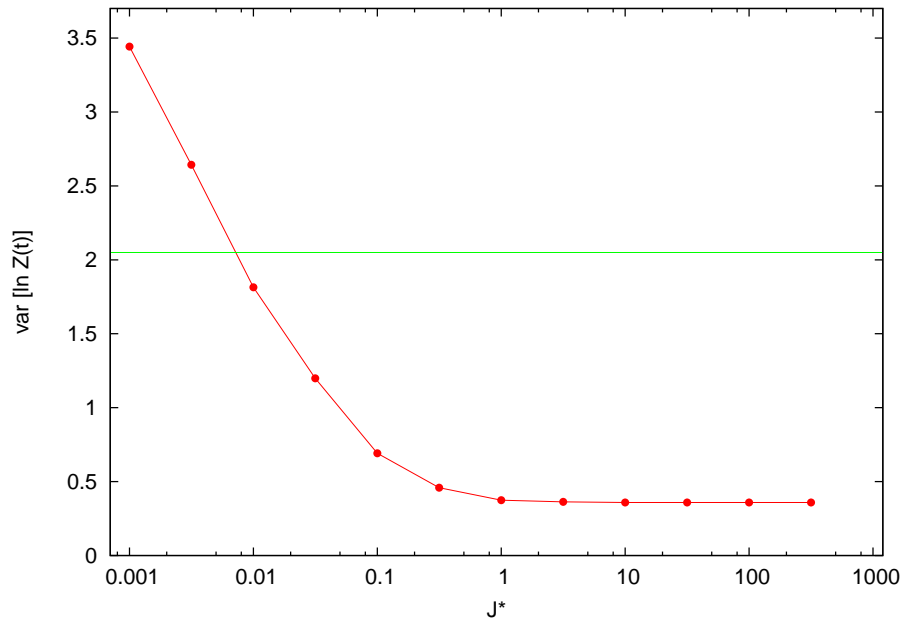


Figura 4.7: In rosso è riportata la varianza della quantità $\ln(Z(t_f))$ in funzione del valore di J^* imposto nella dinamica ($\sigma = 0.346, \tau = 3.15, t_f = 38$). In verde è invece riportata la varianza di $\ln(Z(t))$ ricavata dai dati reali per gli anni 1962-2000.

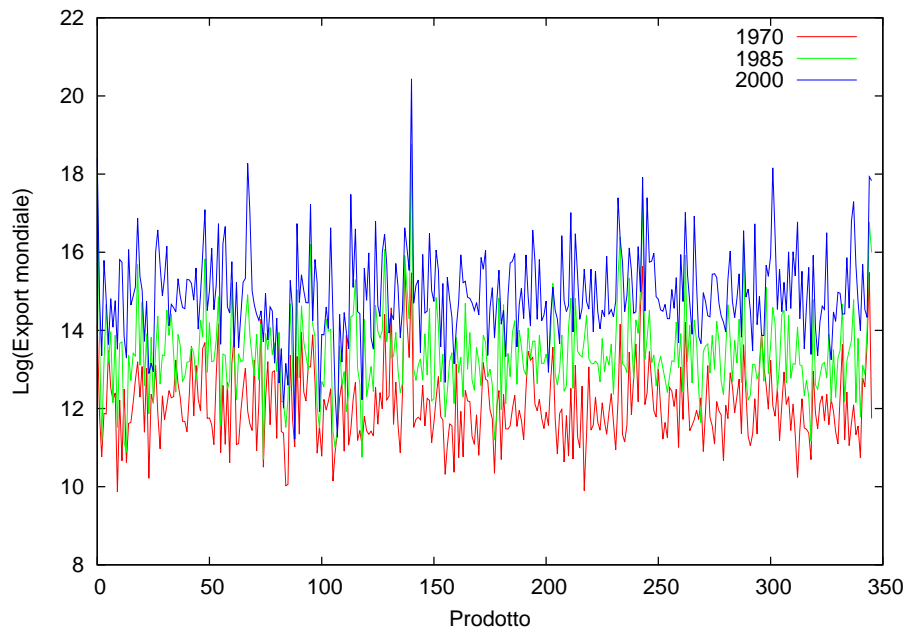


Figura 4.8: Simulazioni relative alla crescita dei 346 prodotti registrati nel dataset NBER, a intervalli di 15 anni. ($\sigma = 0.346, \tau = 3.15, \mu = 0.055, J^* = 0.008, Z_i(0) = Z_i(1962)$).

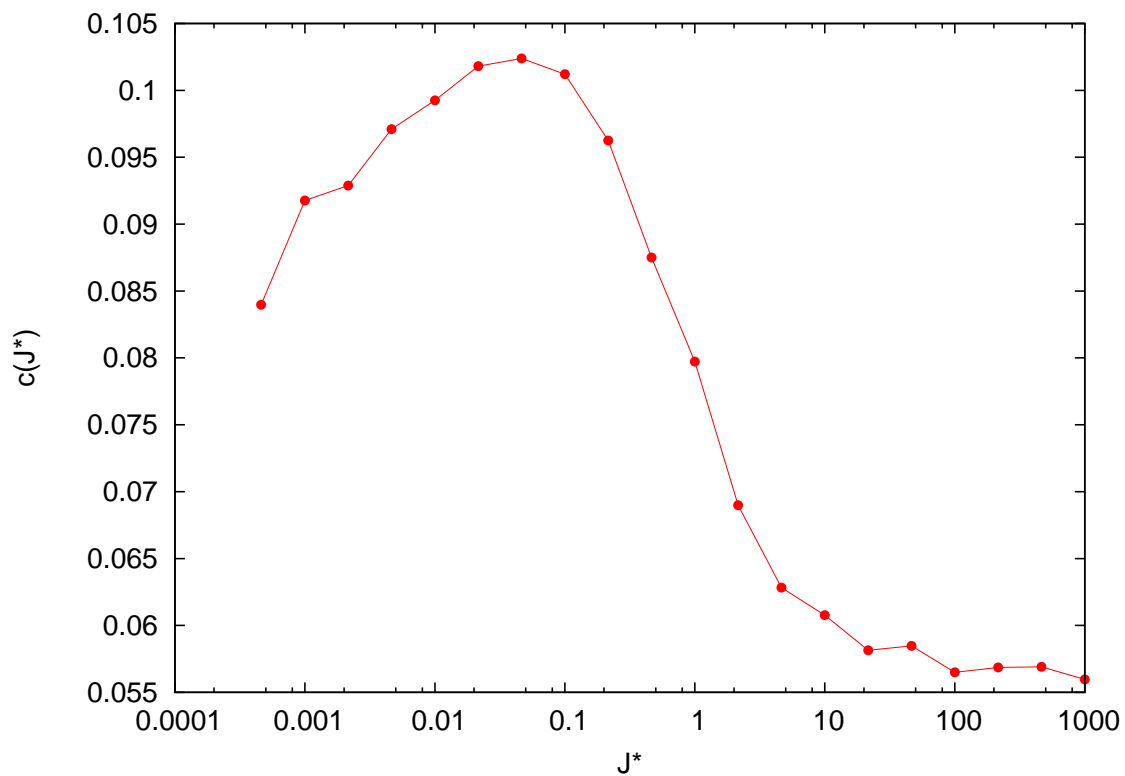


Figura 4.9: Velocità di crescita asintotica della network dei prodotti per ($\mu = 0.055, \sigma = 0.346, \tau = 3.15$). La simulazione è stata protratta per $t = 160\tau$.

Capitolo 5

Estensioni del modello e considerazioni finali

Il modello presentato permette di capire che caratteristiche abbia la crescita globale e quali siano i settori più importanti per lo sviluppo economico. Nonostante le carenze presenti nel dataset, è stato possibile fornire delle stime per i parametri di (4.1). Inoltre, rispetto al lavoro [7], in cui si era evidenziato come i beni centrali fossero quelli associati ad un ritorno economico maggiore, in questo lavoro la nozione di centralità assume un significato diverso e non necessariamente correlato al valore del bene: centrali sono i prodotti per i quali la crescita risulta superiore rispetto alla media. Una possibile estensione del modello, con implicazioni d'interesse economico, riguarda l'inserimento e l'eliminazione di nodi (cioè beni) nel corso del tempo, così da capire quali ripercussioni si abbiano sul mercato collegate alla nascita e morte dei vari prodotti. Altre osservazioni possono essere fatte sulla scelta della network. La network costruita secondo [7], pur riuscendo a distinguere i settori produttivi in maniera corretta, non sempre riesce ad attribuire la giusta centralità ai beni. Ne è un esempio il petrolio: quest'ultimo non risulta infatti essere una merce centrale (perché i paesi che lo esportano spesso non producono molti altri beni), però il petrolio e i suoi derivati sono necessari (e quindi prossimi) alla produzione di un'innumerabile quantità di beni. Tale osservazione sembra suggerire che la network scelta possa non essere la più indicata per comprendere a fondo la crescita globale e quindi la necessità di ridefinire la prossimità fra prodotti secondo nuove regole per considerare tali aspetti.

Altre limitazioni nell'analisi fatta derivano dall'aver trattato in termini di campo medio i parametri τ , σ e J^* : settori produttivi diversi possono avere comportamenti peculiari, che nell'approccio scelto (e con i dati a disposizione) non è possibile rilevare.

Nei limiti del modello vi è senza dubbio l'incapacità di trattare l'evoluzione delle nazioni singolarmente: per fare questo sarebbe necessario costruire un set di network (una per

ogni paese) collegate fra loro attraverso i flussi di import/export, valutando per altro quali prodotti una nazione abbia possibilità di sviluppare o meno. In questo modo si potrebbe studiare (nella migliore delle ipotesi) la crescita delle singole nazioni e verificare anche le ripercussioni di eventuali shock economici (un esempio su tutti: il mancato flusso di import di petrolio nei paesi sviluppati).

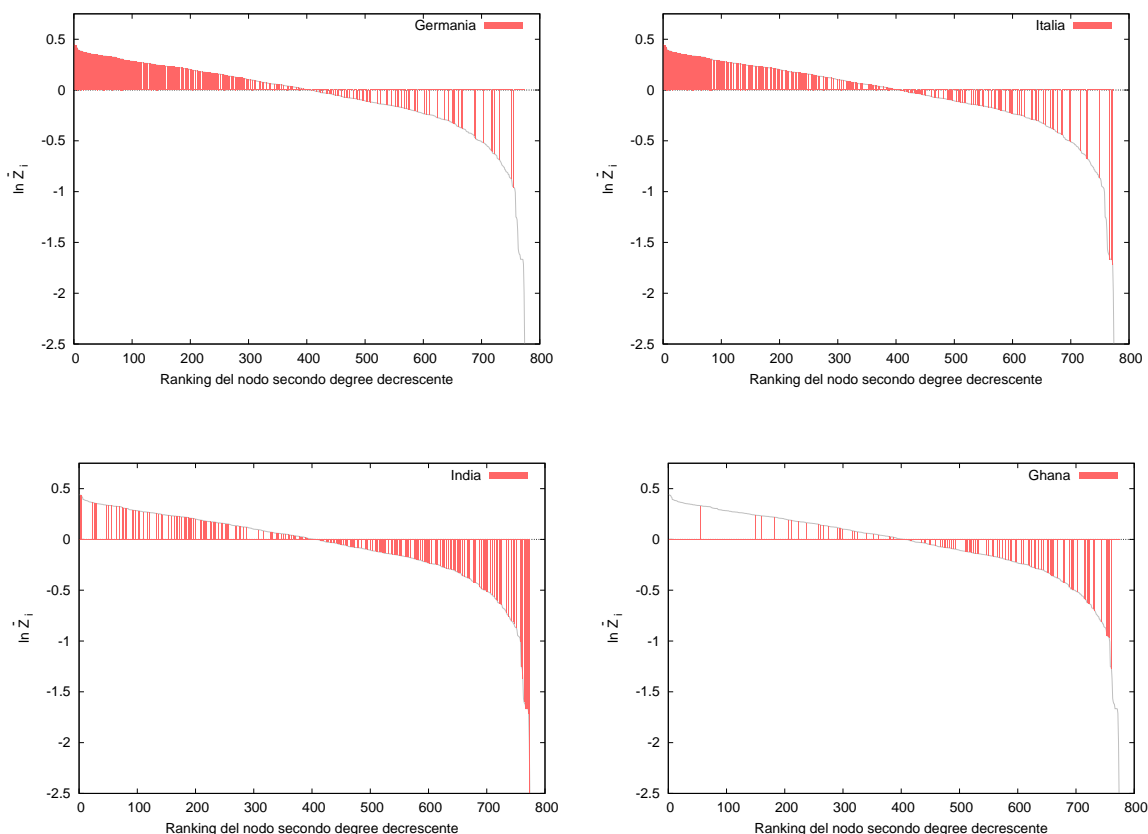


Figura 5.1: prodotti esportati per Germania, Italia, India e Ghana (anno 2000).

Detto questo, vogliamo comunque dare una stima molto rozza di quali siano le nazioni per le quali si prevede una crescita più rapida. Per farlo, si potrebbe simulare l'evoluzione della network con i parametri di τ , σ , μ , J^* ricavati e impostare le $Z_i(0)$ secondo i valori di export dichiarati dalla nazione che si vuole prendere in esame. Valutando per ogni nazione la quantità $c(t)$ a tempi di integrazione piccoli (cioè prima che la crescita asintotica si manifesti, ovvero quando le condizioni iniziali hanno ancora un

effetto sulla velocità di crescita) si potrebbero ricavare informazioni sulla crescita dei singoli paesi. Questo approccio (analogo agli step diffusivi studiati in [7]) è tuttavia insoddisfacente: in primis esso consente a una nazione di sviluppare beni che possono essergli preclusi per varie ragioni (climatiche, di materie prime, etc.); inoltre, a causa della struttura della network, beni dominanti (come il petrolio, che presenta in assoluto i valori di export più grandi fra le merci) vengono identificati erroneamente come periferici, alterando in maniera considerevole l'evoluzione dinamica del sistema, e per queste ragioni tale approccio è stato scartato. Quello che però è possibile fare è valutare il set di prodotti che ogni nazione esporta e determinare quanti di questi siano associati a tassi di crescita medi maggiori. In figura (5.1) sono riportati a titolo di esempio i prodotti esportati¹ da Germania, Italia, India e Ghana, ordinati secondo il degree associato alla merce.

Questa visualizzazione permette di capire cosa un paese sviluppi e se ci si aspetti da esso una velocità di crescita superiore rispetto ad altre nazioni. Tuttavia non mette in luce quanto grandi siano questi flussi (cioè se l'export relativo a un dato prodotto sia ingente o meno), né considera il ritorno economico legato alla vendita di un prodotto: si limita ad indicare se una nazione sviluppa prodotti da cui ci aspettiamo una velocità di crescita elevata o meno. I risultati di questa analisi sono riassunti nella figura 5.2, ove sono indicate solo alcune delle nazioni prese in esame (in modo da non rendere il grafico illeggibile). In tale grafico, la nazione c -esima è identificata da un punto (x_c, y_c) del piano in cui le due coordinate sono date da

$$\begin{aligned} x_c &= \frac{\sum_i \ln \bar{Z}_i \Theta(RCA(c, k)) \Theta(-\ln \bar{Z}_i)}{\sum_i \ln \bar{Z}_i} \\ y_c &= \frac{\sum_i \ln \bar{Z}_i \Theta(RCA(c, k)) \Theta(\ln \bar{Z}_i)}{\sum_i \ln \bar{Z}_i}, \end{aligned} \quad (5.1)$$

ove le \bar{Z}_i sono definite secondo (3.8). La figura 5.2 mostra come la discriminante principale nella differenziazione fra le nazioni sia la loro capacità di produrre beni centrali: i paesi economicamente dominanti si posizionano sulla parte alta del grafico, mentre la loro coordinata lungo l'asse x ha un ruolo più marginale: essendo i prodotti relativi a quest'asse soprattutto riguardanti il settore primario, la coordinata x di una nazione è dipendente prevalentemente dalla sua posizione geografica e dalle risorse naturali di cui dispone. Tale grafico sembra anche confermare le assunzioni fatte in [12], secondo cui un paese produce tutti i beni che gli sono consentiti dal suo livello tecnologico: non è infatti presente nessun trend evidente di spostamento verso sinistra nel grafico delle nazioni più sviluppate rispetto ai paesi poveri, sono semplicemente più a sinistra i paesi con condizioni geografiche sfavorevoli.

¹ovvero aventi $RCA > 1$.

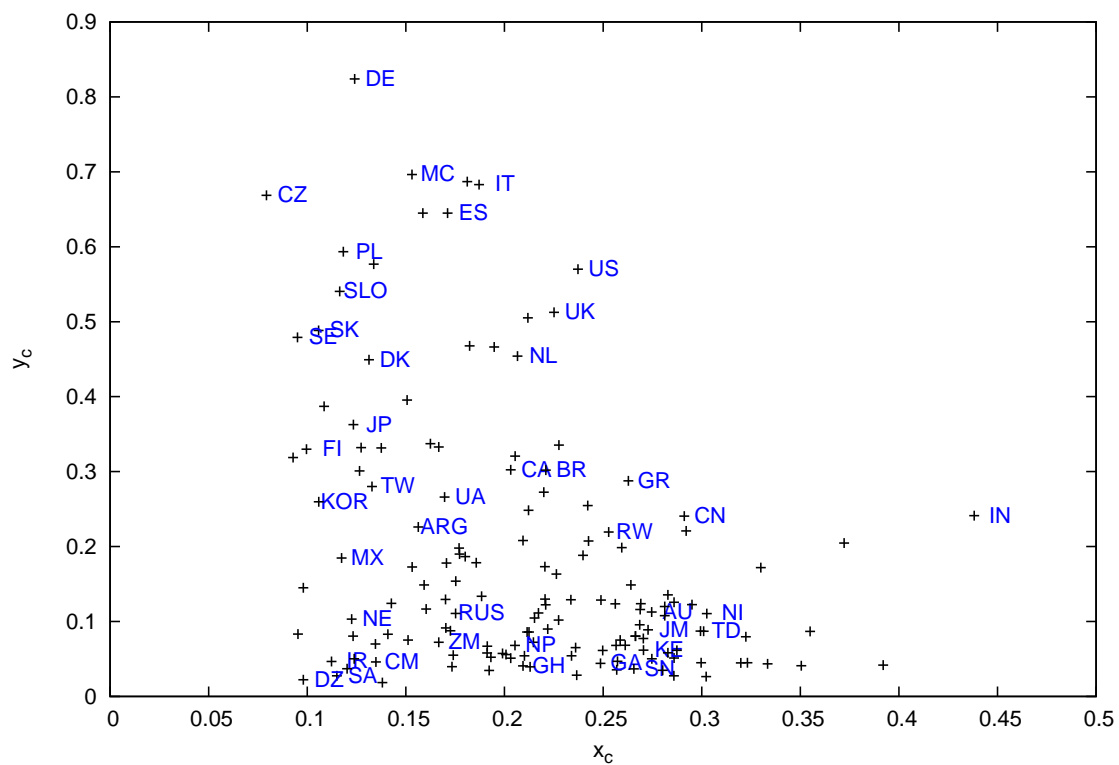


Figura 5.2: Configurazione produttiva delle nazioni (anno 2000). Ogni nazione è caratterizzata dalle sue coordinate (x_c, y_c) secondo (5.1).

Appendice A

Cenni di Teoria dei Grafi

La teoria dei grafi sta man mano entrando in ogni disciplina in virtù della sua elasticità, foriera di applicabilità in innumerevoli contesti. La materia è molto vasta e in questa appendice trattiamo solo i teoremi e le definizioni che ci saranno utili nell'analisi.

A.1 Laplaciano discreto

Prima di introdurre il Laplaciano discreto, diamo la definizione di grafo. Essa varia un po' in funzione degli autori. In [5] un grafo viene formalizzato come segue.

Definizione A.1.1. *Un grafo G è una terna ordinata $(V(G), E(G), \psi_G)$ costituita da un insieme non vuoto $V(G)$ di vertici (o nodi), un insieme $E(G)$ disgiunto da $V(G)$ di edge (o collegamenti) e una funzione di incidenza ψ_G che associa ad ogni edge una coppia di nodi (non necessariamente distinta) in $V(G)$.*

Il più delle volte in realtà non si fa distinzione fra l'insieme $E(G)$ e la funzione di incidenza ed in tal caso si definisce un grafo come la coppia ordinata $G = (V, E)$. Gli edge sono detti *link* se i nodi associati sono distinti oppure *loop* quando la funzione di incidenza associa il nodo a se stesso. In questo capitolo ci riferiamo a grafi non pesati (cioè non associamo una funzione di 'peso' per gli edge), non direzionali (non imponiamo versi di percorrenza agli edge) e connessi. Al fine di giungere rapidamente ai risultati di nostro interesse, introduciamo immediatamente gli oggetti di cui abbiamo bisogno:

Definizione A.1.2. *Un cammino dal nodo v_i al nodo v_j è una sequenza di link $(e_{i_1 i_2}, \dots, e_{i_{k-1} i_k})$ attraverso nodi distinti tali che $i_1 = i$ ed $i_k = j$. Il più breve cammino da v_i a v_j prende il nome di geodetica. Un grafo si dice connesso se per ogni coppia di nodi esiste un cammino che li collega.*

Definizione A.1.3. Un albero è un grafo in cui per ogni coppia di vertici esiste un solo cammino. Questo equivale ad affermare che non sono presenti cicli all'interno del grafo, cioè cammini con nodo iniziale e finale coincidenti.

Definizione A.1.4. La matrice di adiacenza di un grafo finito G avente n vertici è la matrice $A = [a_{ij}]_{n \times n}$ ove a_{ij} è il numero di edge dal vertice i al vertice j .

Definizione A.1.5. La matrice di degree (o grado) di un grafo finito G avente n vertici è la matrice $D = [d_{ij}]_{n \times n}$ ove $d_{ij} = \delta_{ij} \deg(v_i)$ e in cui $\deg(v_i)$ rappresenta il conteggio degli edge collegati al vertice v_i .

Definizione A.1.6. La matrice Laplaciana di un grafo finito G avente n vertici è la matrice $L = [l_{ij}]_{n \times n}$ definita secondo

$$L = D - A$$

con D matrice di degree ed A matrice di adiacenza. Si definisce inoltre la matrice Laplaciana normalizzata $\mathcal{L} = [\ell_{ij}]_{n \times n}$ come

$$\mathcal{L} = D^{-1/2} L D^{-1/2}.$$

La matrice Laplaciana normalizzata si rivela in genere più comoda per gli sviluppi teorici. Al fine di capire meglio le definizioni date, esibiamo il Laplaciano per il seguente grafo

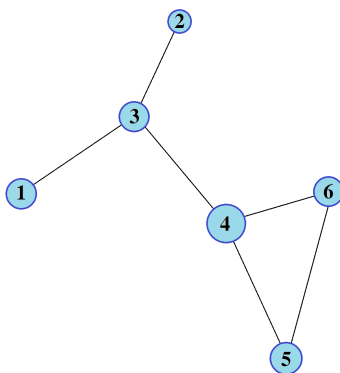


Figura A.1: Un grafo di partenza.

$$L = D - A = \begin{pmatrix} 1 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 \\ -1 & -1 & 3 & -1 & 0 & 0 \\ 0 & 0 & -1 & 3 & -1 & -1 \\ 0 & 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & -1 & -1 & 2 \end{pmatrix}.$$

La connessione con l'operatore di Laplace e quindi una giustificazione al nome di tale matrice, si può cogliere facilmente immaginando un grafo unidimensionale, in cui ogni nodo è connesso al successivo. In tale caso la matrice Laplaciana ha nella diagonale il degree dei nodi (che vale 2 in ogni nodo) e nelle entrate ad essa adiacenti il valore -1 . Supponendo di avere una funzione che assume i valori $(f_1, \dots, f_{k-1}, f_k, f_{k+1}, \dots)$ nei vari nodi, l'applicazione di L a tale vettore fornisce i termini $v_k = 2f_k - f_{k-1} - f_{k+1}$ che per l'appunto hanno la "struttura" di una derivata seconda discreta. Tornando alla definizione di Laplaciano, facciamo notare che esso può essere visto come un operatore agente nello spazio delle funzioni $g : V(G) \rightarrow \mathbb{R}$ per il quale

$$\mathcal{L}g(u) = \frac{1}{\sqrt{\deg(u)}} \sum_{\substack{v \\ u \sim v}} \left(\frac{g(u)}{\sqrt{\deg(u)}} - \frac{g(v)}{\sqrt{\deg(v)}} \right). \quad (\text{A.1})$$

ove la sommatoria si intende ristretta ai nodi v connessi con un link a u .

A.2 Equazioni differenziali lineari a coefficienti costanti

In questa sezione studiamo l'equazione $\dot{Z}(t) = -LZ(t)$, ove L è una matrice Laplaciana. A questo scopo analizziamo il grafo in figura A.2. Per esso, il sistema da considerare è dato da

$$\begin{pmatrix} \dot{z}_1(t) \\ \dot{z}_2(t) \\ \dot{z}_3(t) \end{pmatrix} = - \begin{pmatrix} 2 & -1 & -1 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} z_1(t) \\ z_2(t) \\ z_3(t) \end{pmatrix}.$$

Si tratta di risolvere un sistema di equazioni differenziali lineari a coefficienti costanti. È noto che la risolvete $\Phi(t)$ tale che $\Phi(0) = id$ è data dalla matrice e^{-tL} , ovvero

$$e^{-tL} = \frac{1}{6} \begin{bmatrix} 2 + 4e^{-3t} & 2 - 2e^{-3t} & 2 - 2e^{-3t} \\ 2 - 2e^{-3t} & 2 + e^{-3t} + 3e^{-t} & 2 + e^{-3t} - 3e^{-t} \\ 2 - 2e^{-3t} & 2 + e^{-3t} - 3e^{-t} & 2 + e^{-3t} + 3e^{-t} \end{bmatrix}.$$

Risulta chiaro che le soluzioni in questione, per tempi sufficientemente lunghi tendono a valori costanti¹. Più precisamente, si ha che per $t \rightarrow \infty$, $z_1(t) = z_2(t) = z_3(t)$, con $\lim_{t \rightarrow \infty} z_i(t) = \sum_i z_i(0)/3$. A titolo di esempio, nel caso in cui $z_1(0) = 10, z_2(0) = 2, z_3(0) = 3$, le soluzioni sono date da

$$\begin{aligned} z_1(t) &= 5 + 5e^{-3t} \\ z_2(t) &= \frac{1}{2} (10 - 5e^{-3t} - e^{-t}) \\ z_3(t) &= \frac{1}{2} (10 - 5e^{-3t} + e^{-t}) \end{aligned}$$

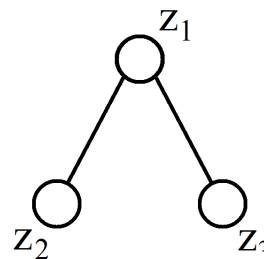


Figura A.2: Grafo di esempio.

e sono riportate in figura A.3.

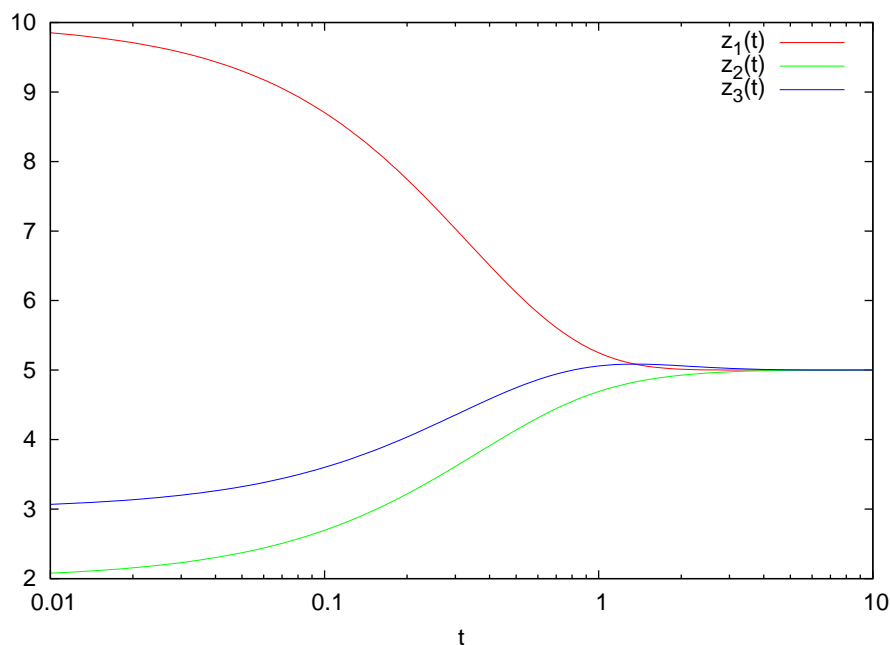


Figura A.3: evoluzione temporale per $z_i(t)$ relative all'esempio in figura A.2.

¹Questo poteva essere intuito anche ricordando che gli autovalori relativi ad una matrice laplaciana soddisfano la condizione $0 = \lambda_0 \leq \lambda_1 \leq \dots \leq \lambda_N$.

Ringraziamenti

Ringrazio *CZ*, *LDC* e \mathcal{YK} .

Bibliografía

- [1] Robert K. Adair. Stochastic contributions to global temperature changes. *Phys. Rev. Lett.*, 100:148501, Apr 2008.
- [2] R. B. Bapat. ‘The Laplacian Matrix of a Graph’. *The Mathematics Student*, 65:214–223, 1996.
- [3] Lorenz Bartosch. Generation of colored noise. *International Journal of Modern Physics C*, 12(06):851–855, 2001.
- [4] Danielle S. Bassett, David L. Alderson, and Jean M. Carlson. Collective decision dynamics in the presence of external drivers. *Phys. Rev. E*, 86:036105, Sep 2012.
- [5] John Adrian Bondy. *Graph Theory With Applications*. Elsevier Science Ltd., Oxford, UK, UK, 1976.
- [6] J.-P. Bouchaud and M. Mezard. Wealth condensation in a simple model of economy. *eprint arXiv:cond-mat/0002374*, February 2000.
- [7] A.-L. Barabási C. A. Hidalgo, B. Llinger and R. Hausmann. The product space conditions the development of nations’. *Science*, 317:482–487, 2007.
- [8] F.R.K. Chung. *Spectral Graph Theory*. Number No. 92 in CBMS Regional Conference Series. Conference Board of the Mathematical Sciences, 2006.
- [9] I. Corwin. The Kardar-Parisi-Zhang equation and universality class. *ArXiv e-prints*, June 2011.
- [10] Thomas M. Cover. Universal portfolios. *Mathematical Finance*, 1(1):1–29, 1991.
- [11] Victor M. Eguíluz, Dante R. Chialvo, Guillermo A. Cecchi, Marwan Baliki, and A. Vania Apkarian. Scale-free brain functional networks. *Phys. Rev. Lett.*, 94:018102, Jan 2005.

- [12] A. Tacchella M. Cristelli G. Caldarelli A. Gabrielli and L. Pietronero. ‘A New Metrics for Countries’ Fitness and Products’ Complexity. *Scientific Reports*, 2:723, 2012.
- [13] Thomas Gueudré, Alexander Dobrinevski, and Jean-Philippe Bouchaud. Explore or exploit? a generic model and an exactly solvable case. *Phys. Rev. Lett.*, 112:050602, Feb 2014.
- [14] T. Halpin-Healy and Y.C. Zhang. *Kinetic Roughening Phenomena, Stochastic Growth, Directed Polymers and All that: Aspects of Multidisciplinary Statistical Mechanics*. Physics reports. North-Holland Elsevier, 1995.
- [15] R. Hausmann and B. Klinger. ‘Structural Transformation and Patterns of Comparative Advantage’. *Working Papers*, 128, 2006.
- [16] Ricardo Hausmann, Jason Hwang, and Dani Rodrik. What You Export Matters. NBER Working Papers 11905, National Bureau of Economic Research, Inc, December 2005.
- [17] Elhanan Helpman. Sources of comparative advantage. *MIT Press*, 1984.
- [18] Mehran Kardar, Giorgio Parisi, and Yi-Cheng Zhang. Dynamic scaling of growing interfaces. *Phys. Rev. Lett.*, 56:889–892, Mar 1986.
- [19] A. M. Leshansky and L. M. Pismen. Do small swimmers mix the ocean? *Phys. Rev. E*, 82:025301, Aug 2010.
- [20] R. C. Feenstra, R. E. Lipsey, H. Deng, A. C. Ma, H. Mo. World trade flows : 1962-2000. *Working Paper*, 11040, 2005.
- [21] Mervin E. Muller. A note on a method for generating points uniformly on n-dimensional spheres. *Commun. ACM*, 2(4):19–20, April 1959.
- [22] David R. Nelson and Nadav M. Shnerb. Non-hermitian localization and population biology. *Phys. Rev. E*, 58:1383–1403, Aug 1998.
- [23] M. Newman. *Networks: An Introduction*. OUP Oxford, 2010.
- [24] Pawel Romanczuk, Iain D. Couzin, and Lutz Schimansky-Geier. Collective motion due to individual escape and pursuit response. *Phys. Rev. Lett.*, 102:010602, Jan 2009.
- [25] Ma Ángeles Serrano and Marián Boguñá. Topology of the world trade web. *Phys. Rev. E*, 68:015101, Jul 2003.

- [26] G. Strang. *Introduction to Linear Algebra*. Wellesley-Cambridge Press, 2009.
- [27] D.J. Watts and S.H. Strogatz. Collective dynamics of 'small-world' networks. *Nature*, pages 440–442, 1998.