

ALMA MATER STUDIORUM · UNIVERSITÀ DI BOLOGNA

FACOLTÀ DI SCIENZE MATEMATICHE, FISICHE E NATURALI
Corso di Laurea Triennale in Matematica

**TRASMISSIONE DI MALATTIE
INFETTIVE E DIFFUSIONE DI
EPIDEMIE SU NETWORK:
MODELLI MATEMATICI**

Tesi di Laurea in Biomatemática

Relatore:
Chiar.mo Prof.
MIRKO DEGLI ESPOSTI

Presentata da:
SARA ANDRAGHETTI

Sessione III
Anno Accademico 2010/2011

Introduzione

La dinamica di malattie infettive rappresenta una delle più antiche e ricche aree di studio della Biomatemática; infatti da sempre malattie ed epidemie sono una delle maggiori cause di morte nel mondo, anche più di guerre e carestie. Gli agenti patogeni si adattano e si evolvono nel corso del tempo, determinando così la comparsa di nuove malattie o il rigenerarsi di antiche forme di contagio che si pensavano ormai estinti. È chiaro che l'invasione umana di nuovi ecosistemi, fattori quali il surriscaldamento globale, il degrado ambientale, il potenziamento degli spostamenti e comunicazioni su scala internazionale, e i cambiamenti in campo economico, hanno cominciato a fornire sempre più terreno fertile per lo sviluppo di nuove malattie.

Si è ipotizzato che per apprendere a fondo questo complesso fenomeno occorresse il contributo di diverse e molteplici discipline: ecco che modelli matematici sono risultati estremamente adatti a questo scopo e hanno iniziato a giocare un ruolo incredibilmente significativo. Innanzi tutto il principale fine di questi studi risiede nel tentativo di sviluppare modelli che possano aiutare a valutare i possibili effetti su popolazioni, a seguito della diffusione di un particolare microorganismo, e possano di conseguenza fornire gli strumenti necessari per controllare e prevedere la diffusione di malattie infettive. Queste analisi consentono anche di capire quali siano le migliori strategie di cura e rendono possibili importanti scelte riguardo programmi di immunizzazione (vaccinazioni, terapie, prevenzione, programmi di controllo) e distribuzione di risorse monetarie e tecniche. Una trattazione matematica risulta quindi calzante e necessaria, se le dinamiche di tali eventi richiedono un'analisi quantitativa del fenomeno, specialmente quando, come in questi casi, gli agenti che causano le epidemie possono essere dei più svariati tipi.

Problema cruciale nel definire tali modelli di dinamica di malattie infettive è la rappresentazione matematica del meccanismo di trasmissione del contagio. Si parla di epidemia quando il ricorrere della malattia raggiunge tempi e proporzioni che sono oltre le normali aspettative, fino a superare un punto di *soglia*,

ovvero un valore critico che indica il passaggio da una situazione di contagio ristretta, a una diffusione su larga scala. Il concetto di *forza di infezione* viene quindi ad essere una delle parole chiave in questa trattazione.

Indice

Introduzione	i
1 Modelli matematici di diffusione di epidemie	1
1.1 Assunzioni e ipotesi iniziali	1
1.2 Legge di massa-azione	3
1.3 Teoria classica di modelli lineari	3
1.3.1 Modello SIR	3
1.3.2 Modello SIR con immunizzazione di suscettibili e caso della vaccinazione	8
1.3.3 Caso della vaccinazione	9
1.3.4 Modello SIR con immissione di suscettibili che dà luogo a epidemie ricorrenti, che intervengono in maniera periodica	10
1.3.5 Modello SIRS	11
1.3.6 Modello SIS	13
1.3.7 Modello SIR con dinamica vitale	15
2 Diffusione di epidemie su networks	21
2.1 Considerazioni preliminari: networks e grafi	21
2.2 Grado e misure di centralità	23
2.2.1 Degree centrality	23
2.2.2 Closeness centrality	23
2.2.3 Betweenness centrality	23
2.3 Clustering	23
2.4 Degree distribution	24
2.5 Betweenness distribution	24
2.6 Network e il mondo reale	25
2.6.1 Social networks	25
2.6.2 Small-world	25
2.6.3 Eterogeneità e code lunghe	26

2.7	Diffusione di epidemie all'interno di network di popolazioni . . .	28
2.7.1	Random networks	29
2.7.2	Modelli su reticolo	30
2.7.3	Small-world networks	31
2.7.4	Spatial networks	31
2.7.5	Scale-free networks	31
2.8	Richiami	33
2.9	Epidemie su network eterogenei	34
2.9.1	Modello SI	34
2.9.2	Modelli SIR e SIS	37
2.9.3	L'effetto di modelli di mescolamento	37
2.10	Diffusione epidemica negli scale-free networks: assenza del punto di equilibrio	39
A	Appendice	43
A.1	Sistemi lineari omogenei	43
B	Appendice	51
B.1	Grafi random	51
	Bibliografia	53

Elenco delle figure

1.1	Piano delle fasi SI per il classico modello SIR nel caso epidemico.	7
1.2	Andamento di una epidemia descrivibile da un modello SIS. Se $S = \frac{\gamma}{r} = \frac{\lambda}{\beta}$ allora $I = N - \frac{\gamma}{r}$	14
1.3	Piano delle fasi SI per il classico modello SIR endemico con tasso di riproduttività $\rho_0 = 0.5 < 1$	17
1.4	Piano delle fasi SI per il classico modello SIR endemico con tasso di riproduttività di base $\rho_0 = 3$, in periodo medio di infezione di 3 giorni e durata media della vita di 60 giorni. Questa lunghezza di vita irrealisticamente breve è stata appositamente scelta in modo che l'equilibrio endemico sia chiaramente raggiunto al di sopra dell'asse orizzontale e possano essere osservate le orbite che raggiungono il punto di equilibrio endemico, avvolgendosi a spirale.	17
1.5	Diagramma di biforcazione per il modello SIR nel caso endemico, che mostra che il punto di equilibrio in assenza della malattia e l'equilibrio allo stato endemico, cambiano stabilità quando il valore di contatto ρ_0 (qui rappresentato come σ) è 1	19
2.1	Cinque tipi diversi di networks, da sinistra a destra e dall'alto in basso: random, lattice, small-world, spatial e scale-free.	32
2.2	Questa figura mostra il tipico modello SIR di diffusione di un'epidemia, all'interno delle cinque tipologie di networks, che sono da sinistra a destra: random, lattice, small world, e (al di sotto) spatial e scale-free. Ogni grafo mostra 100 curve per l'epidemia (in grigio), assieme con il valore medio dei valori più alti dell'epidemia (in nero)	33
A.1	Nodo stabile	44
A.2	Nodo instabile	45
A.3	Punto sella	45

A.4	Fuoco (spirale) stabilea	45
A.5	Fuoco (spirale) instabile	46
A.6	Centro	47
A.7	(a sinistra) Nodo stella attrattivo; (a destra) Nodo stella repulsivo	47
A.8	Nodo stabile di seconda specie	48
A.9	Nodo instabile di seconda specie	48
A.10	Diagramma che riassume e mostra come $\text{tr}A$ e $\text{det}A$, determinano il tipo di singolarità nel piano delle fasi. Quando la traccia è negativa e il determinante è positivo, abbiamo il nodo stabile al di sotto della parabola, e una spirale stabile al di sopra. Quando la traccia è nulla e il determinante è positivo, si ha il centro. Se il determinante e la traccia sono entrambi positivi, troviamo la spirale instabile e il nodo instabile, rispettivamente al di sopra e al di sotto della parabola. I punti di sella si trovano invece nel semipiano in cui il determinante è negativo.	49

Capitolo 1

Modelli matematici di diffusione di epidemie

1.1 Assunzioni e ipotesi iniziali

Al fine di consentire l'esplicitazione di tali modelli, occorre spiegare che quelli presentati ed analizzati in questo lavoro, si fondano su un'organizzazione a compartimenti. Intendiamo cioè che il totale della popolazione o della comunità presa in considerazione come soggetto del dilagarsi dell'epidemia, viene suddivisa in un numero (solitamente piccolo) di categorie discrete. La classe dei *Suscettibili* S , consiste in quegli individui che possono incorrere nella malattia ma non ne sono ancora stati contagiati. La classe degli *Infettivi* I , è costituita da coloro che sono ammalati e che stanno trasmettendo a loro volta la malattia ad altri, e quella dei *Rimossi* R , contiene coloro che sono stati eliminati dalla cerchia dei suscettibili in quanto o totalmente guariti (e quindi resisi immuni), o perché stati isolati o deceduti (sia per cause naturali sia a seguito del contagio stesso). Queste elencate sono solo le principali e più comuni categorie che si ripresentano trasversalmente in tutti i modelli, ma molte altre suddivisioni sono possibili, a seconda del tipo di contagio preso in considerazione.

Oltre a questa prima caratterizzazione, vediamo dunque gli altri aspetti peculiari dei modelli per epidemie in Biomatemática.

1. Nei modelli classici viene fatta l'assunzione che il totale N della popolazione sia costante e anche sufficientemente grande da poter considerare l'ampiezza di ogni classe come variabile continua invece che discreta.

2. I modelli possono distinguersi in quelli con *dinamica vitale* o *senza dinamica vitale*. Nel primo caso si assume che le nascite e le morti di individui si susseguano con uguale peso all'interno del processo vitale, e che tutti i nuovi nati rientrino totalmente fra i suscettibili. Gli individui vengono poi rimossi in caso di morte da ogni classe con una stessa costante di proporzionalità, detta *coefficiente di mortalità* (δ). La durata della vita di ciascun individuo sarà dunque l'inverso di tale coefficiente. Possiamo semplificare il concetto parlando di modelli con dinamica vitale quando la durata della malattia è tale da influenzare in modo rilevante l'andamento della vita della popolazione, a differenza di quello che succede nei casi di modelli senza dinamica vitale.
3. La popolazione è uniformemente e omogeneamente miscelata all'interno di ogni classe. Un *fattore di contatto giornaliero* λ indica il valore medio di contatti per infettivo al giorno. Il contatto di un infettivo è un'interazione che porta al contagio di un altro individuo, appartenente alla categoria dei suscettibili. λ è un valore fissato e non varia stagionalmente. Da questa definizione segue che il numero medio di suscettibili contagiati da un infettivo giornalmente è dato da λS , e il numero medio di contagiati dall'intera classe degli infettivi, di dimensione (NI) , è λSIN .
4. Gli individui guariscono e vengono rimossi dalla classe I a un tasso proporzionale agli infettivi (chiamiamo la costante di proporzionalità γ). Il periodo di latenza (ovvero il tempo che intercorre fra il momento di esposizione al contagio e il momento in cui comincia a manifestarsi la malattia) è nullo. Quindi la funzione che indica la porzione di individui esposti (e immediatamente capaci di trasmettere l'infezione) all'istante iniziale t_0 , e che sono ancora infettivi all'istante di tempo successivo $t_0 + t$, è $e^{-\gamma t}$. Il periodo medio in cui l'individuo è contagioso risulta essere l'inverso della costante di proporzionalità $\frac{1}{\gamma}$.

Dalle considerazioni fatte si evince che il *tasso di rimozione* dalla categoria degli infettivi sia per morte che per guarigione è dato da $\gamma + \delta$, da cui segue immediatamente che il periodo di infezione è $\frac{1}{\gamma + \delta}$. Analogamente possiamo calcolare il numero medio di contatti con suscettibili, durante il periodo di infettività di un individuo, come $\rho = \frac{\lambda}{\gamma + \delta}$; mentre il numero medio di suscettibili contagiati da un infettivo durante il suo periodo di malattia è dato da ρS .

1.2 Legge di massa-azione

Tipicamente alla base dei modelli di epidemie c'è l'idea di voler stabilire il tasso con cui i suscettibili diventano infettivi, basato sul rapporto

$$\frac{dS}{dt} = -\lambda SI$$

dove λ rappresenta il *valore di contatto* tramite cui l'infettivo contagia altri individui.

Questa legge è stata stimata sulla base delle seguenti ipotesi. Supponiamo che p rappresenti la probabilità con cui un infettivo, situato in un intorno di un individuo suscettibile alla malattia, entri in contatto con tale individuo e lo contagi. Ne segue che $1 - p$ indica la probabilità che questo non avvenga. Il valore assegnato a p è dato da $p = \lambda \Delta t$, dove Δt esprime l'intervallo di tempo con cui l'infettivo resta nella cerchia del suscettibile. Estendiamo il concetto al caso in cui I individui infettivi entrano in contatto con S : la probabilità che il contagio non avvenga è data da $(1 - p)^I$, se queste infezioni vengono trattate come eventi indipendenti l'una dall'altra. La probabilità che un infettivo sul totale degli I contagi un suscettibile è data da

$$1 - (1 - p)^I = 1 - \left[1 + \binom{I}{1}p + \binom{I}{2}p^2 \dots \right] = \binom{I}{1}p + \binom{I}{2}p^2 + \dots \approx Ip = \lambda \Delta t I$$

(per un valore di $p \ll 1$).

Riscriviamo quest'ultimo risultato nei seguenti termini

$$I(t + \Delta t) - I(t) = \lambda \Delta t I$$

ottenendo una relazione che vale quando il contatto avviene con un solo suscettibile. Nel caso invece si considerino S suscettibili si ha che: $I(t + \Delta t) - I(t) = \lambda \Delta t IS$ da cui, passando al continuo:

$$\lim_{\Delta t \rightarrow 0} \frac{I(t + \Delta t) - I(t)}{\Delta t} = \lambda IS$$

da cui il risultato atteso $\frac{dI}{dt} = \lambda SI$, che rappresenta il *tasso con cui I infettivi contagiano S suscettibili*. Segue allora

$$\frac{dS}{dt} = -\frac{dI}{dt}$$

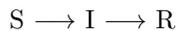
1.3 Teoria classica di modelli lineari

1.3.1 Modello SIR

Consideriamo l'evoluzione di una epidemia all'interno di una popolazione ospite, di cui il totale di individui sia preso costante, di valore N . Uno dei più

elementari modelli, ma ancora estremamente rilevante e particolarmente appropriato per descrivere quelle malattie infettive che conferiscono una immunità duratura, è il così detto modello SIR, la cui prima formulazione, attorno al 1927, è merito dei due scienziati *Kermack* e *McKendrick*. Questo è un modello che si basa sulle tre categorie principali S,I,R e prende in analisi il caso in cui un piccolo gruppo di infettivi viene introdotto in una vasta popolazione: un problema classico potrebbe essere quello di voler descrivere la diffusione della malattia all'interno della popolazione, in funzione del tempo. Chiaramente questo fenomeno dipende da una grande varietà di circostanze, fra cui anche la tipologia stessa di malattia considerata, ma come primo tentativo di formulazione di un modello, faremo le seguenti assunzioni generali.

Consideriamo una malattia che conferisca totale immunità all'individuo, una volta raggiunta la guarigione, oppure porti direttamente alla morte se letale. In entrambi i casi questi individui rientreranno a far parte dei rimossi, poiché in ogni caso, non potranno più essere contagiati. Il processo che regola la diffusione di tale malattia può essere schematizzato in questo modo



Al fine di scrivere una formulazione matematica di tale fenomeno, dobbiamo introdurre una serie di equazioni differenziali, in modo da indicare i tassi di trasferimento di individui, da un compartimento all'altro:

$$\frac{dS}{dt} = f_1(I, S, R)$$

$$\frac{dI}{dt} = f_2(I, S, R)$$

$$\frac{dR}{dt} = f_3(I, S, R)$$

Per questo tipo di processo infettivo, nel passaggio dalla classe S alla classe I viene tipicamente assunta valida la *legge di massa-azione*, mentre per definire il trasferimento dalle classe I ad R, viene considerata una funzione di decrescita esponenziale. Il modo più semplice per definire le tre funzioni $f_i, i = 1, 2, 3$, è il seguente:

$$f_1 = -\lambda IS$$

$$f_2 = \lambda IS - \gamma I$$

$$f_3 = \gamma I$$

dove γ e λ sono costanti positive che rappresentano rispettivamente il *tasso di rimozione* dalla categoria degli infettivi e il *tasso di infezione* della malattia.

Il modello SIR così descritto ha validità se si assume che, quando un suscettibile viene contagiato, questo divenga immediatamente infettivo, ovvero non si considera alcun periodo di latenza della malattia. Nel caso in cui invece sia presente un tempo di incubazione dovrà essere inclusa nel modello un'ulteriore classe contenente quegli individui soggetti a latenza.

Il sistema di equazioni differenziali così ottenuto è il seguente:

$$\begin{cases} \frac{dS}{dt} = -\lambda IS \\ \frac{dI}{dt} = \lambda IS - \gamma I \\ \frac{dR}{dt} = \gamma I \end{cases}$$

con condizioni iniziali $S(0) = S_0 > 0, I(0) = I_0 > 0, R(0) = 0$. Siamo interessati a risolvere questo sistema completo, cercando unicamente soluzioni che siano positive per S,I,R, in quanto altre soluzioni non avrebbero senso nell'ottica di meccanismi biologici. Dalla somma delle equazioni

$$\frac{dS}{dt} + \frac{dI}{dt} + \frac{dR}{dt} = 0$$

segue che

$$S(t) + I(t) + R(t) = N,$$

dove N è il totale della popolazione.

Il valore R(t) può sempre essere ricavato da S(t) e da I(t) usando la relazione $R(t) = N - S(t) - I(t)$, e per questo motivo possiamo andare a studiare le soluzioni del sistema nel piano delle fasi SI. Poichè $S + I + R = S_0 + I_0 + R_0 = N$ e $R_0 = 0$, segue che nel piano delle fasi S,I si ottiene la seguente relazione:

$$S + I = S_0 + I_0 = N$$

ovvero la conservazione della somma dei suscettibili e infettivi.

Problema cruciale in una qualunque situazione epidemica è cercare di determinare se, dati λ, γ e S_0 e il valore iniziale di infettivi I_0 , l'infezione sia nelle condizioni tali da diffondersi o se resterà invece un fenomeno ristretto di piccola portata, e a quale valore inizierà il suo declino. Dalla relazione

$$\frac{dI}{dt} = (\lambda S - \gamma)I$$

segue che

$$\frac{dI}{dt} > 0 \Leftrightarrow (\lambda S - \gamma) > 0$$

in particolare

$$\frac{dI}{dt} > 0 \Leftrightarrow (\lambda S_0 - \gamma) > 0$$

Questa condizione è soddisfatta se $S_0 > \frac{\gamma}{\lambda}$, il che significa che l'epidemia si propagherà. Per un valore di $S_0 < \frac{\gamma}{\lambda}$ l'epidemia invece si estingue. Il rapporto $\rho = \frac{\gamma}{\lambda}$ assume quindi un *valore di soglia* rispetto al quale si verifica un fenomeno di *biforcazione*. Scriviamo

$$R_0 = \frac{\lambda}{\gamma} S_0$$

dove R_0 è il *tasso di riproduttività di base* della malattia che rappresenta il numero di individui che hanno contratto la malattia da parte di coloro che erano già infettivi, all'interno di una popolazione totalmente suscettibile, (il numero dei suscettibili è uguale al totale della popolazione stessa). Il valore $\frac{1}{\lambda}$, rappresenta invece la *durata media della malattia*. Questo valore di R_0 è estremamente utile per caratterizzare l'andamento dell'epidemia. Rispetto a questo tasso possiamo riformulare il teorema di soglia in questi termini:

se $R_0 > 1$, l'epidemia si propaga;

se $R_0 < 1$, l'epidemia si estingue.

Un modo per ridurre il valore di R_0 , onde evitare che la malattia si diffonda ulteriormente, se il suo valore tende ad essere troppo elevato, sarebbe quello di ridurre il numero dei suscettibili S_0 . La vaccinazione, ad esempio, è uno dei metodi più efficaci che consentono questo, ma può avere successo se applicata a comunità stabili e non di dimensioni eccessivamente ampie.

La cosa importante da tenere in considerazione è che una epidemia può comunque iniziare e diffondersi molto rapidamente nonostante il valore di riproduttività cresca al di sotto della soglia di criticità.

Possiamo ricavare interessanti risultati analitici sull'andamento dell'epidemia, dallo studio delle *orbite* del sistema. Per ricavare queste traiettorie, dividiamo la seconda equazione del sistema per la prima

$$\frac{dI}{dS} = -\frac{(\lambda S - \gamma)I}{\lambda S I} = -1 + \frac{\gamma}{\lambda S} = -1 + \frac{\rho}{S}, (I \neq 0)$$

Separando le variabili, $dI = (-1 + \frac{\gamma}{\lambda S})dS$ e successivamente integrando

$$I = \oint (-1 + \frac{\gamma}{\lambda S})dS = -S + \frac{\gamma}{\lambda} \ln S + c.$$

Imponiamo le condizioni iniziali c.i. per calcolare c , e rispetto a un istante iniziale $t = 0$ vale,

$$I_0 = -S_0 + \frac{\gamma}{\lambda} \ln S_0 + c, c = I_0 + S_0 - \frac{\gamma}{\lambda} \ln S_0.$$

Otteniamo quindi la seguente condizione per I ,

$$I = N - S + \rho \ln \frac{S}{S_0}, \quad \text{dove} \quad N = S_0 + I_0.$$

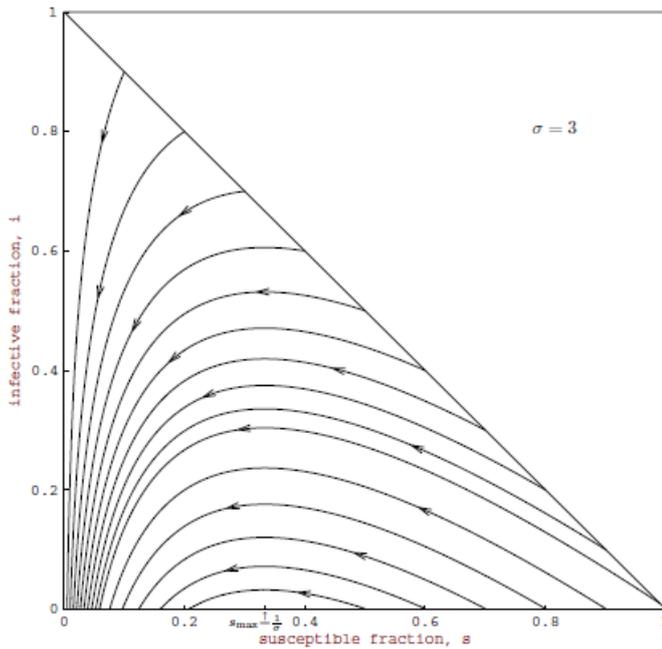


Figura 1.1: Piano delle fasi SI per il classico modello SIR nel caso epidemico.

Le traiettorie nel piano delle fasi SI sono rappresentate nel grafico soprastante, da cui si evincono alcuni interessanti risultati.

Le curve sono determinate dalle condizioni iniziali $I(0) = I_0$ e $S(0) = S_0$. Ponendo $R(0) = R_0$, tutte le traiettorie hanno punto di inizio sulla retta $S + I = N$ e rimangono all'interno del triangolo formato dalle intersezioni degli assi coordinati con tale retta, e la retta stessa, finché $0 < S + I < N$ per qualunque valore di t . Troviamo una situazione epidemica quando $I(t) > I_0$ per ogni $t > 0$; questo succede sempre quando $S_0 > \frac{1}{\sigma}$, con $\frac{1}{\sigma} = \rho$.

Si può osservare che I è crescente, da un valore iniziale I_0 , ovvero da questo istante in poi si può riscontrare l'evolversi di una epidemia, che raggiunge un picco, per poi scendere a zero ed estinguersi. Potrebbe tuttavia anche presentarsi un caso differente, quando ad esempio I_0 è vicino al valore I_{max} (ordinata del punto ad ascissa S_{max} , nell'orbita più vicina all'asse delle ascisse). In questa circostanza, come si vede dalla figura, se $S_0 < \rho$, allora I decresce da un valore iniziale I_0 e quindi non da inizio ad alcuna epidemia. Studiamo il comportamento asintotico, da cui si ottiene:

$$\lim_{t \rightarrow \infty} I(t) = 0 \quad \lim_{t \rightarrow \infty} S(t) = S_\infty$$

dove il valore S_∞ indica un residuo di suscettibili che non si sono estinti al passaggio della malattia. Questo comportamento lo si nota dal fatto che

non tutte le curve tendono a 0, ma molte raggiungono valori maggiori di 0. Per dimostrare questo risultato consideriamo nel piano delle fasi SR il rapporto $\frac{dS}{dR} = -S\sigma$. Integrando per separazione di variabili si ottiene

$$S(R) = S_0 e^{-R\sigma}.$$

Poiché $R \leq N$, segue che:

$$S(R) = S_0 e^{-R\sigma} \geq S_0 e^{-N\sigma} > 0.$$

Il valore di S decresce, mentre il

$$\lim_{t \rightarrow \infty} S(t) \geq S_0 e^{-N\sigma} > 0.$$

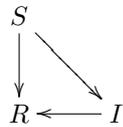
Questo significa che asintoticamente la popolazione dei suscettibili non si estingue. Per un tempo $t \rightarrow \infty$ (ovvero verso il termine della malattia), il valore degli infettivi tende ad estinguersi, e possiamo dunque stimare come

$$S_\infty - N = \frac{1}{\sigma} \ln\left(\frac{S_\infty}{S_0}\right)$$

dove $I_{tot} = N - S_\infty$, il numero totale di coloro che hanno contratto la malattia. I punti di equilibrio si ottengono determinando le soluzioni del sistema di equazioni differenziali e in particolare, in questo caso, sono tutti i punti appartenenti alla retta $I = 0$. Vale quindi che per il modello SIR senza dinamica vitale, esistono infiniti punti di equilibrio.

1.3.2 Modello SIR con immunizzazione di suscettibili e caso della vaccinazione

Si tratta di un modello SIR senza dinamica vitale, per cui vale la legge di massa-azione, in cui una parte di suscettibili passa direttamente nella categoria dei rimossi a un tasso proporzionale al numero S stesso.



Tale modello può essere quindi espresso dal sistema:

$$\begin{cases} \frac{dS}{dt} = -\lambda SI - \mu S \\ \frac{dI}{dt} = \lambda SI - \gamma I \\ \frac{dR}{dt} = \gamma I + \mu S \end{cases}$$

con condizioni iniziali $S(0) = S_0 > 0$, $I(0) = I_0 > 0$, $R(0) = 0$. Vale che $\frac{dS}{dt} + \frac{dI}{dt} + \frac{dR}{dt} = 0 \Rightarrow S + I + R = cost = N$.

Il sistema sopra può quindi essere riscritto nei seguenti termini:

$$\begin{cases} \frac{dS}{dt} = (-\lambda I - \mu)S \\ \frac{dI}{dt} = (\lambda S - \gamma)I \end{cases}$$

Per la validità del teorema di soglia, se $S > \frac{\gamma}{\lambda}$, l'epidemia si propaga ($\frac{dI}{dt} > 0$); se $S < \frac{\gamma}{\lambda}$, l'epidemia si estingue ($\frac{dI}{dt} < 0$).

Nel piano delle fasi SI, possiamo osservare che l'unico punto di equilibrio è l'origine $0 = (0, 0)$, la cui stabilità è ricavabile dallo studio della matrice Jacobiana:

$$J = \begin{bmatrix} -\lambda I - \mu & -\lambda S \\ \lambda I & -\gamma \end{bmatrix}$$

$$J(0, 0) = \begin{bmatrix} -\mu & 0 \\ 0 & -\gamma \end{bmatrix}$$

i cui autovalori sono $\lambda_1 = -\mu$, $\lambda_2 = -\gamma$. $(0, 0)$ rappresenta dunque un nodo di tipo stabile. Questo indica la tendenza sia dei suscettibili che degli infettivi a diventare rimossi, determinando così l'estinzione della malattia. In particolare $(0, 0)$ è un punto globalmente stabile, infatti:

$$\frac{dS}{dt} + \frac{dI}{dt} = -\mu S - \gamma I \leq 0 \implies S + I$$

è sempre decrescente. Poiché per $S < \frac{\gamma}{\lambda}$, $\frac{dI}{dt} < 0$, quindi $I \rightarrow 0$ per $t \rightarrow \infty \implies S + I \rightarrow 0$, per $t \rightarrow \infty$. Dal rapporto

$$\frac{dI}{dS} = \frac{(\lambda S - \gamma)I}{(-\lambda I - \mu)S}$$

è possibile ricavare le orbite nel piano delle fasi SI, per separazione di variabili

$$\frac{(-\lambda I - \mu)}{I} dI = \frac{(\lambda S - \gamma)}{S} dS$$

e integrando da ambo le parti $-\lambda I - \mu \ln I = \lambda S - \gamma \ln S + c$. L'equazione delle orbite sarà dunque:

$$-\lambda I - \mu \ln I - \lambda S + \gamma \ln S = c$$

1.3.3 Caso della vaccinazione

Sia p il valore indicante la frazione di individui di una popolazione N , sottoposta a vaccinazione (si suppone che tale vaccinazione risulti efficace al 100%). Ne segue che questa porzione di individui non apparterrà più alla categoria dei suscettibili, in quanto non più esposti alla possibilità di contrarre la malattia. La componente di popolazione che costituisce la classe dei suscettibili sarà data

da $S = (1 - p)N$. Sostituendo questa espressione nella seconda equazione del sistema del modello con immunizzazione di suscettibili, ne viene la relazione

$$\frac{dI}{dt} = (\lambda(1 - p)N - \gamma)I$$

Il problema che ci si pone è quello di determinare a quanto deve ammontare il valore di p affinché l'epidemia si estingua completamente. La soluzione a questo quesito viene fornita dal teorema di soglia, in tal modo: se

$$\frac{dI}{dt} < 0 \implies \lambda(1 - p)N - \gamma < 0,$$

$$\lambda(1 - p)N < \gamma,$$

$$(1 - p) < \frac{\gamma}{N\lambda}$$

$$p > 1 - \frac{\gamma}{N\lambda}$$

dove p è il più piccolo naturale che soddisfi questa relazione.

Poiché il tasso di riproduttività di base vale $R_0 = \frac{N\lambda}{\gamma}$ (quando S coincide con il totale della popolazione N), se $R_0 > 1$, l'epidemia si propaga, mentre se $R_0 < 1$, l'epidemia si estingue. Da queste considerazioni, segue che posso riscrivere il valore assunto da p rispetto a R_0 , cioè $p > 1 - \frac{1}{R_0}$.

1.3.4 Modello SIR con immissione di suscettibili che dà luogo a epidemie ricorrenti, che intervengono in maniera periodica

$$\longrightarrow S \longrightarrow I \longrightarrow R$$

Supponiamo che l'immissione di individui nella categoria dei suscettibili avvenga a un tasso costante μ . Per comodità non andremo a considerare la mortalità all'interno delle classi, ovvero ci si pone nelle condizioni in cui la malattia abbia una durata breve rispetto alla vita media della popolazione.

Le equazioni di tale modello sono:

$$\begin{cases} \frac{dS}{dt} = -\lambda SI + \mu \\ \frac{dI}{dt} = \lambda SI - \gamma I \\ \frac{dR}{dt} = \gamma I \end{cases}$$

di condizioni iniziali $S(0) = S_0 > 0$, $I(0) = I_0 > 0$, $R(0) = 0$.

$$J_{S,I} = \begin{bmatrix} -\lambda I & -\lambda S \\ \lambda I & \lambda S - \gamma \end{bmatrix}$$

$$J(E) = \begin{bmatrix} -\frac{\lambda\mu}{\gamma} & -\gamma \\ \frac{\lambda\mu}{\gamma} & 0 \end{bmatrix}$$

dove $E = (\frac{\gamma}{\lambda}, \frac{\mu}{\gamma})$ è il punto di equilibrio di tale modello.

$Tr(E) = -\frac{\lambda\mu}{\gamma} < 0$, $detJ(E) = \lambda\mu > 0 \implies E$ è un punto asintoticamente stabile.

Dall'equazione caratteristica $t^2 + \frac{\lambda\mu}{\gamma}t + \lambda\mu = 0$ (calcolata rispetto all'incognita t), $\Delta = (-\frac{\lambda\mu}{\gamma})^2 - 4\lambda\mu$ ne segue che:

- se $\mu < \frac{4\gamma}{\lambda}$ esistono due radici complesse coniugate e dunque, in questo caso E è un fuoco stabile e si hanno oscillazioni smorzate;
- se $\mu > \frac{4\gamma}{\lambda}$ E è un nodo;
- se $\mu = \frac{4\gamma}{\lambda}$, è il caso in cui $\Delta = 0$ e dunque E è un nodo stellato o nodo di seconda specie.

Supponiamo ora che il tasso di immissione di individui nei suscettibili sia proporzionale al numero di suscettibili S stesso.

$$\begin{cases} \frac{dS}{dt} = -\lambda SI + \mu S \\ \frac{dI}{dt} = \lambda SI - \gamma I \\ \frac{dR}{dt} = \gamma I \end{cases}$$

di condizioni iniziali $S(0) = S_0 = 0$, $I(0) = I_0 = 0$, $R(0) = 0$.

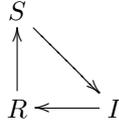
Punti di equilibrio $E_0 = (0, 0)$ e $E_1 = (\frac{\gamma}{\lambda}, \frac{\mu}{\lambda})$, sono rispettivamente punto di sella e centro. Questo risultato lo si può desumere dallo studio delle orbite nel piano delle fasi:

$$\frac{dS}{dI} = \frac{(-\lambda I + \mu)S}{(\lambda S - \gamma)I} \implies \frac{(\lambda S - \gamma)}{S} dS = \frac{(-\lambda I + \mu)}{I} dI$$

Le orbite ottenute hanno equazione $\mu \ln I - \lambda I + \gamma \ln S - \lambda S = c$, sono cioè orbite chiuse attorno al punto E_1 indicanti l'esistenza di soluzioni periodiche, ovvero epidemie ricorrenti.

1.3.5 Modello SIRS

Si immagini il caso in cui gli individui siano solo temporaneamente immunitizzati, ovvero in cui i rimossi possono tornare ad essere suscettibili, e quindi esposti nuovamente al contagio. Questa situazione può essere descritta in questo modo:



e prende il nome di modello SIRS. Il sistema per questo modello

$$\begin{cases} \frac{dS}{dt} = -\lambda IS + aR \\ \frac{dI}{dt} = \lambda IS - \gamma I \\ \frac{dR}{dt} = \gamma I - aR \end{cases}$$

presenta un coefficiente a proporzionale al tasso dei rimossi che rientrano nella categoria dei suscettibili. Dalla conservazione della popolazione totale N , si ricava $R = N - S - I$, e quindi è possibile considerare il piano delle fasi SI in cui definire un sistema equivalente a quello dato

$$\begin{cases} \frac{dS}{dt} = -\lambda IS + a(N - S - I) \\ \frac{dI}{dt} = \lambda IS - \gamma I \end{cases}$$

Dallo studio dei punti di equilibrio e della loro stabilità, è possibile capire l'evoluzione della malattia. Costruiamo quindi la matrice Jacobiana del sistema

$$J = \begin{bmatrix} -\lambda I - a & -\lambda S - a \\ \lambda I & \lambda S - a \end{bmatrix}$$

I punti di equilibrio $E_1 = (N, 0)$ e $E_2 = (\frac{\gamma}{\lambda}, \frac{a(N - \frac{\gamma}{\lambda})}{a + \gamma})$, si calcolano come soluzioni del sistema ponendo $\frac{dS}{dt} = 0$ e $\frac{dI}{dt} = 0$.

Osserviamo che E_2 esiste se e solo se $N > \frac{\gamma}{\lambda}$, ovvero $\frac{\lambda N}{\gamma} > 1$.

$$J(E_1) = \begin{bmatrix} -a & -\lambda N - a \\ 0 & \lambda N - \gamma \end{bmatrix}$$

i cui autovalori $\lambda_1 = -a < 0$, $\lambda_2 = \lambda N - \gamma$, risulta essere < 0 se il punto E_2 non esiste, e > 0 se esiste E_2 (dall'osservazione precedente).

Poiché il determinante $\det J(E_1) = -a(\lambda N - \gamma)$, ne viene che E_1 è un nodo stabile se non esiste il punto E_2 (caso in cui λ_1 e λ_2 sono autovalori negativi reali); E_1 è un punto di sella se esiste E_2 (caso in cui $\lambda_1 < 0$ e $\lambda_2 > 0$).

Da un punto di vista intuitivo questa analisi corrisponde a dire che se E_1 è un nodo stabile, si giunge all'estinzione della malattia. Mentre se esiste il valore E_2 , questo corrisponde al punto in cui la malattia è endemica, cioè resta sempre un residuo di infettivi che non guariscono del tutto.

$$J(E_2) = \begin{bmatrix} \frac{-\lambda a(N - \frac{\gamma}{\lambda})}{a + \gamma} - a & -\gamma - a \\ \frac{\lambda a(N - \frac{\gamma}{\lambda})}{a + \gamma} & 0 \end{bmatrix}$$

$$\begin{aligned} \text{Tr}J(E_2) &= \frac{-\lambda a(N - \frac{\gamma}{\lambda})}{a + \gamma} < 0 \\ \text{det}J(E_2) &> 0, \end{aligned}$$

per cui E_2 è un punto asintoticamente stabile, e in particolare:

se $\Delta = \frac{[(\lambda N - \gamma) - a(\gamma + a)]^2}{(a + \gamma)^2} - 4(\lambda N - \gamma) > 0$ allora ne segue che E_2 è un nodo stabile, mentre se < 0 , E_2 è un fuoco stabile. In generale E_2 rappresenta un punto dove asintoticamente si arriva nel tempo a una situazione endemica della malattia.

Per questo modello il tasso di riproduttività di base in assenza di suscettibili lo otteniamo da: $\frac{dI}{dt} = -\gamma I$, da cui segue $I(t) = I_0 e^{(-\gamma t)}$, ovvero una soluzione con decadimento esponenziale in quanto $I(t) \rightarrow 0$, per $t \rightarrow \infty$.

Da $\frac{dI}{dt} = \lambda I S - \gamma I$, otteniamo che $I(\Delta t) - I_0 \cong \lambda S_0 I_0 \Delta t$.

Posto $\Delta t = \frac{1}{\gamma}$, dove $\frac{1}{\gamma}$ rappresenta la durata media della malattia, quando $S_0 \cong N$ (ovvero considerando il valore iniziale degli individui che possono contrarre la malattia pari alla quasi totalità della popolazione) e $I_0 \ll N$, otteniamo $R_0 = \frac{N\lambda}{\gamma}$, il *tasso di riproduttività di base*.

- se $R_0 > 1$ la malattia si propaga fino a raggiungere uno stato endemico.
- se $R_0 < 1$ la malattia giunge al punto E_1 e si estingue.

1.3.6 Modello SIS



In virtù della legge di massa-azione, valida anche per questo modello, le equazioni del moto sono:

$$\begin{cases} \frac{dS}{dt} = -\lambda I S + \beta I \\ \frac{dI}{dt} = \lambda I S - \beta I \end{cases}$$

dove β rappresenta il tasso di immissione di infettivi fra i suscettibili. Poiché $\frac{dS}{dt} + \frac{dI}{dt} = 0$ vale se e solo se $S + I = S_0 + I_0 = N$, questo è il caso di un sistema isolato dove si mantiene la totalità della popolazione, ovvero la malattia ha una mortalità trascurabile ($R = 0$). Poniamo $I = N - S$, e sostituiamo questa relazione nella prima equazione del sistema, per cui

- $\frac{dS}{dt} = (-\lambda S + \beta)(N - S)$ è un'equazione di tipo logistico
- $\frac{dS}{dt} > 0 \implies -\lambda S + \beta > 0 \implies S < \frac{\beta}{\lambda}$
- $\frac{dS}{dt} < 0 \implies S > \frac{\beta}{\lambda}$

Quando $S = \frac{\beta}{\lambda}$, otteniamo il valore di equilibrio.

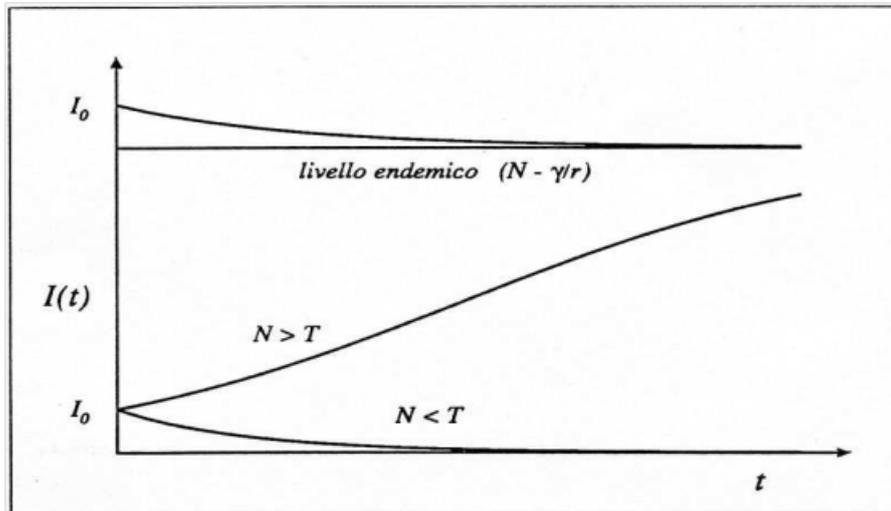
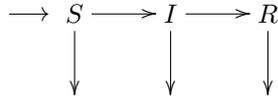


Figura 1.2: Andamento di una epidemia descrivibile da un modello SIS. Se $S = \frac{\gamma}{r} = \frac{\lambda}{\beta}$ allora $I = N - \frac{\gamma}{r}$

Se $S_0 < \frac{\beta}{\lambda}$, si ha una crescita della funzione e quindi la malattia tenderà a propagarsi. Se $S_0 > \frac{\beta}{\lambda}$, si ottiene una decrescita. Il valore di $S_0 = \frac{\beta}{\lambda}$, rappresenta il valore di soglia, rappresentante lo stato endemico dell'epidemia.

1.3.7 Modello SIR con dinamica vitale



Si parla di modello con dinamica vitale, quando la durata della malattia è tale da influenzare l'andamento della vita della popolazione colpita. La prima freccia orizzontale a sinistra indica l'immissione di nuovi nati nella classe dei suscettibili, a un tasso costante $B = \mu N$, proporzionale al totale della popolazione N . Le altre frecce verticali indicano un tasso di mortalità μ , uguale per tutte e tre le classi (e quindi si suppone che la malattia non sia effettivamente letale poiché la mortalità viene compensata da un uguale valore di input di natalità all'interno dei suscettibili).

Osserviamo che i due modelli SIR, il classico senza dinamica vitale, e quello con dinamica vitale, risultano molto diversi fra loro non solo da un punto di vista matematico, ma anche da un punto di vista qualitativo. Infatti, mentre il modello senza dinamica vitale può essere appropriato per descrivere la diffusione di un'epidemia in un breve periodo di tempo, il modello SIR con dinamica vitale si applica bene ai casi in cui questa esplosione perdura per tempi più lunghi.

Per la validità della legge di massa azione, le equazioni di tale modello si presentano in questo modo:

$$\begin{cases} \frac{dS}{dt} = -\lambda SI + \mu N - \mu S \\ \frac{dI}{dt} = \lambda SI - \gamma I - \mu I \\ \frac{dR}{dt} = \gamma I - \mu R \end{cases}$$

Nonostante l'introduzione di una dinamica vitale all'interno del modello, il totale della popolazione $S(t) + I(t) + R(t) = N(t)$ viene comunque mantenuto nel tempo.

Annulliamo le equazioni del sistema:

$$\begin{cases} -\lambda SI + \mu N - \mu S = 0 \\ (\lambda S - \gamma - \mu)I = 0 \\ \gamma I - \mu R = 0 \implies R = \frac{\gamma I}{\mu} \end{cases}$$

$I = 0, S = N$.

Quindi possiamo scrivere i punti di equilibrio del sistema come $E_1 = (N, 0)$, $E_2 = (\tilde{S}, \tilde{I})$, dove $\tilde{S} = \frac{\gamma + \mu}{\lambda}$ e $\tilde{I} = \frac{\mu(N - \tilde{S})}{\lambda \tilde{S}}$.

Vediamo la stabilità di tali punti di equilibrio, introducendo prima alcuni parametri. Sia $\rho_0 = \frac{N\lambda}{\gamma + \mu}$; segue che $\tilde{I} = \frac{\mu N}{\gamma + \mu} \left(1 - \frac{1}{\rho_0}\right)$ e $\tilde{S} = \frac{N}{\rho_0}$. In questo modello

il valore ρ_0 ha la stessa funzione del tasso di riproduzione R_0 , e rappresenta un valore di soglia. Si ha infatti che se $\rho_0 = R_0 > 1$, la malattia si propaga. Se $\rho_0 = R_0 < 1$, la malattia tenderà invece a scomparire. Questo lo si può ottenere dall'analisi degli autovalori della matrice Jacobiana

$$J = \begin{bmatrix} -\lambda I - \mu & -\lambda S \\ \lambda I & \lambda S - (\gamma + \mu) \end{bmatrix}$$

nei punti di equilibrio E_1 ed E_2 .

$$J(E_1) = \begin{bmatrix} -\mu & -\lambda N \\ 0 & \lambda N - (\gamma + \mu) \end{bmatrix}$$

$\lambda_1 = -\mu < 0$. Se $\lambda_2 = \lambda N - (\gamma + \mu) < 0$ (cioè $\rho_0 < 1$) $\implies E_1$ è un nodo asintoticamente stabile \implies l'epidemia si estingue.

$$J(E_2) = \begin{bmatrix} -\lambda \tilde{I} - \mu & -\lambda \tilde{S} \\ \lambda \tilde{I} & 0 \end{bmatrix}$$

$Tr J(E_2) < 0$ e $det J(E_2) > 0 \implies E_2$ è un punto asintoticamente stabile \implies malattia raggiunge uno stato endemico in cui persiste un valore non nullo di infettivi. E_2 può essere un nodo o fuoco, a seconda che il $\Delta = \frac{\lambda^2 \mu^2 N^2}{(\gamma + \mu)^2} - 4\lambda\mu N(1 - \frac{\gamma + \mu}{\lambda N}) = \frac{\lambda^2 \mu^2 N^2}{(\gamma + \mu)^2} - 4\lambda\mu N(1 - \frac{1}{\rho_0})$ sia positivo o negativo.

Osserviamo che se considero $\frac{1}{\mu}$ (l'inverso del coefficiente di mortalità) come valore della vita media della popolazione, $\frac{1}{\gamma + \mu}$ corrisponde alla durata media della malattia. Se $\mu \ll \gamma + \mu$, cioè la vita media è molto più lunga rispetto alla durata della malattia, segue che $\Delta \approx -\lambda\mu N(1 - \frac{1}{\rho_0}) < 0$ e quindi E_2 è un fuoco, ovvero si raggiunge lo stato endemico con oscillazioni smorzate.

I grafici sottostanti rappresentano la diffusione dell'epidemia nei due casi appena discussi, e illustrano le due seguenti possibilità.

Sia $(s(t), i(t))$ una soluzione del sistema all'interno del triangolo T nel piano delle fasi SI, visibile nelle Figure 1.3 e 1.4. Se $\rho_0 \leq 1$ oppure $i_0 = 0$, allora le traiettorie delle soluzioni partono da T e si avvicinano al valore di equilibrio in cui è debellata la malattia, dato da $s = 1$ e $i = 0$. Se $\rho_0 > 1$, allora tutti i cammini con $i_0 > 0$ tendono all'equilibrio endemico dato da $s_e = \frac{1}{\rho_0}$ e da $i_e = \frac{\mu(\rho_0 - 1)}{N\lambda}$.

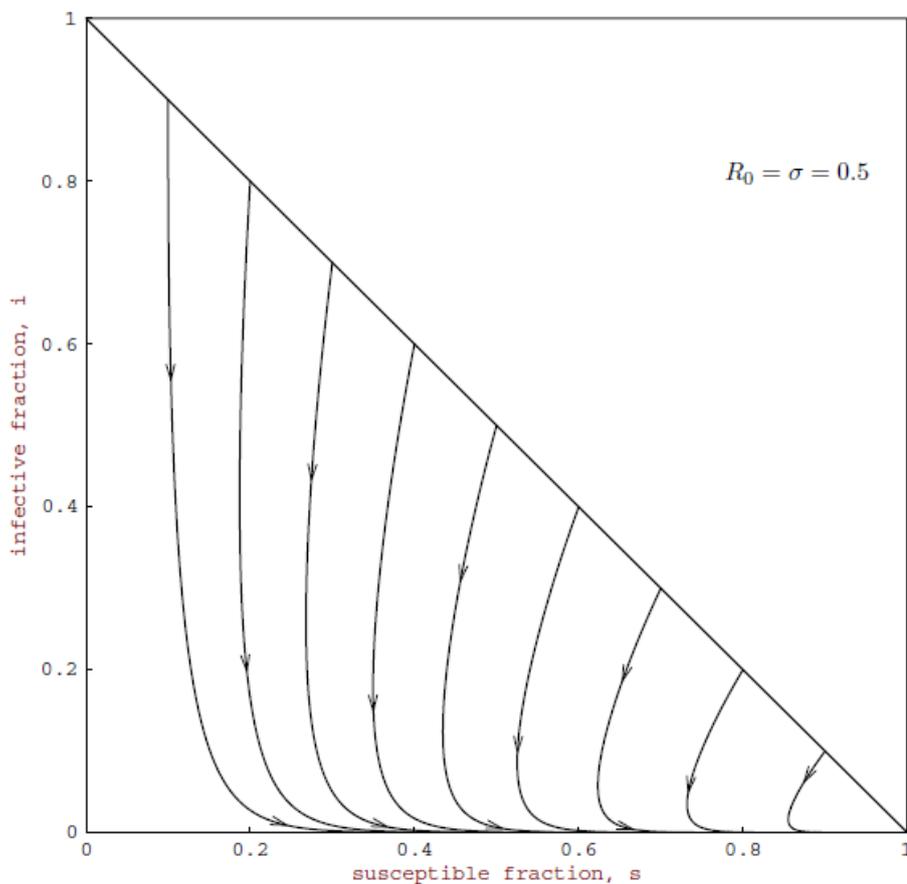


Figura 1.3: Piano delle fasi SI per il classico modello SIR endemico con tasso di riproduttività $\rho_0 = 0.5 < 1$

Quando $R_0 = \rho_0 < 1$, anche $\rho_0 s$ è minore di 1 quando $i_0 > 0$, quindi il valore degli infettivi decresce a zero. Questi scendono rapidamente molto vicino allo zero, e dopo oltre 100 o più anni, gli individui che erano guariti muoiono un dopo l'altro, mentre il processo di nascita lentamente aumenta il numero dei suscettibili fino al punto di equilibrio, in mancanza totale della malattia con $s = 1$ e $i = 0$, in cui eventualmente ogni individuo è suscettibile.

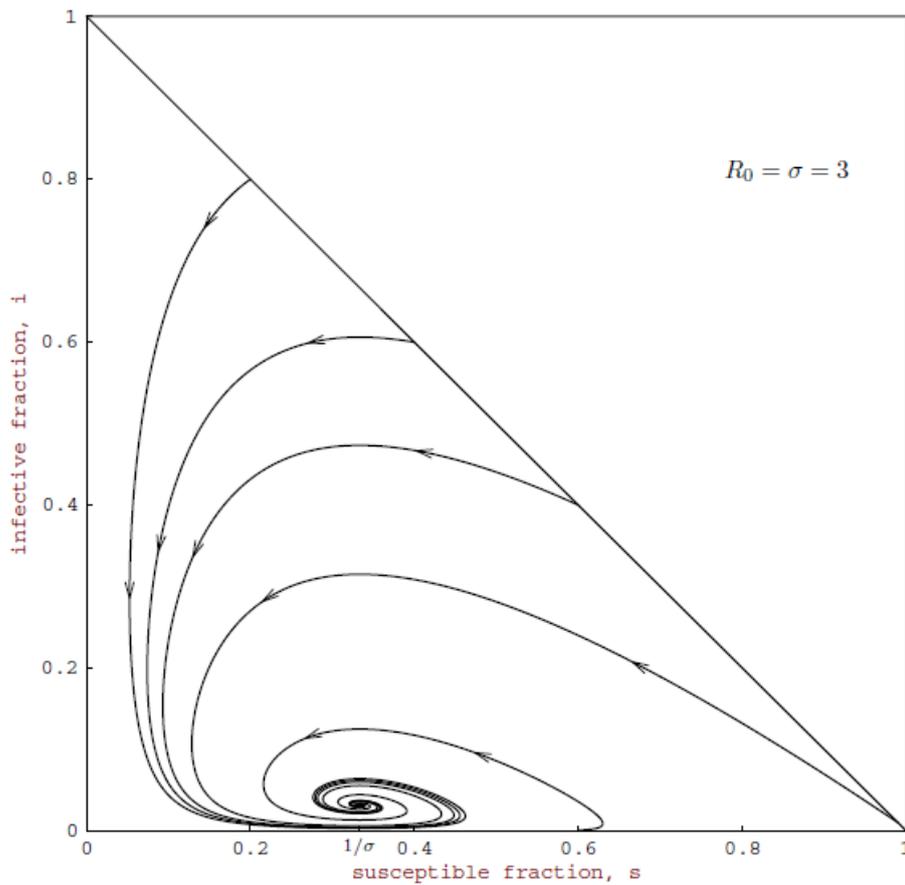


Figura 1.4: Piano delle fasi SI per il classico modello SIR endemico con tasso di riproduttività di base $\rho_0 = 3$, in periodo medio di infezione di 3 giorni e durata media della vita di 60 giorni. Questa lunghezza di vita irrealisticamente breve è stata appositamente scelta in modo che l'equilibrio endemico sia chiaramente raggiunto al di sopra dell'asse orizzontale e possano essere osservate le orbite che raggiungono il punto di equilibrio endemico, avvolgendosi a spirale.

Se $R_0 = \rho_0 > 1$, i_0 piccolo e s_0 grande, con $\rho_0 s_0 > 1$, allora $s(t)$ decresce e $i(t)$ aumenta fino al raggiungimento di un picco, per poi decrescere. Dopo che gli infettivi scendono a un numero più basso, il lento processo di morte degli individui guariti e le nascite di nuovi suscettibili gradualmente (oltre circa 10 o 20 anni) aumenta la frazione dei suscettibili, fino a che $\rho_0 s$ è grande a sufficienza da consentire il ricorrere di una nuova epidemia. Questo processo, in cui si alternano epidemie molto rapide e lenti meccanismi di rigenerazione di suscettibili, continua finché le orbite raggiungono l'equilibrio endemico. In questo punto di equilibrio il valore $\rho_0 s$ è 1, il che è plausibile poiché se il valore di ripristino $\rho_0 s$ è maggiore o minore di 1, la frazione di infettivi $i(t)$ aumenta o diminuisce rispettivamente.

Riassumendo, da un punto di vista intuitivo, se il tasso di riproduttività di base R_0 (uguale al valore ρ_0 quando l'intera popolazione è suscettibile) è minore di 1, allora il punto di equilibrio in assenza della malattia, è localmente asintoticamente stabile, ovvero l'epidemia può invadere la popolazione. Ma se $R_0 > 1$, allora il punto di equilibrio è instabile di tipo repulsivo. Le orbite si muovono nel quadrante SI in modo che l'epidemia si può diffondere, nel senso che ogni traiettoria che parte con un valore piccolo di i_0 , resta nel quadrante positivo dove la malattia persiste. Il comportamento di questo modello dipende completamente dal valore di soglia R_0 che determina non solo quando la stabilità locale del punto di equilibrio varia, ma anche quando l'equilibrio endemico è raggiunto nella regione in cui è presente una frazione positiva di infettivi.

Per questo modello SIR è possibile definire dunque un valore di biforcazione $\sigma = \rho_0 = 1$, rispetto al quale cambia la stabilità del punto di equilibrio, come mostrato nella figura sottostante.

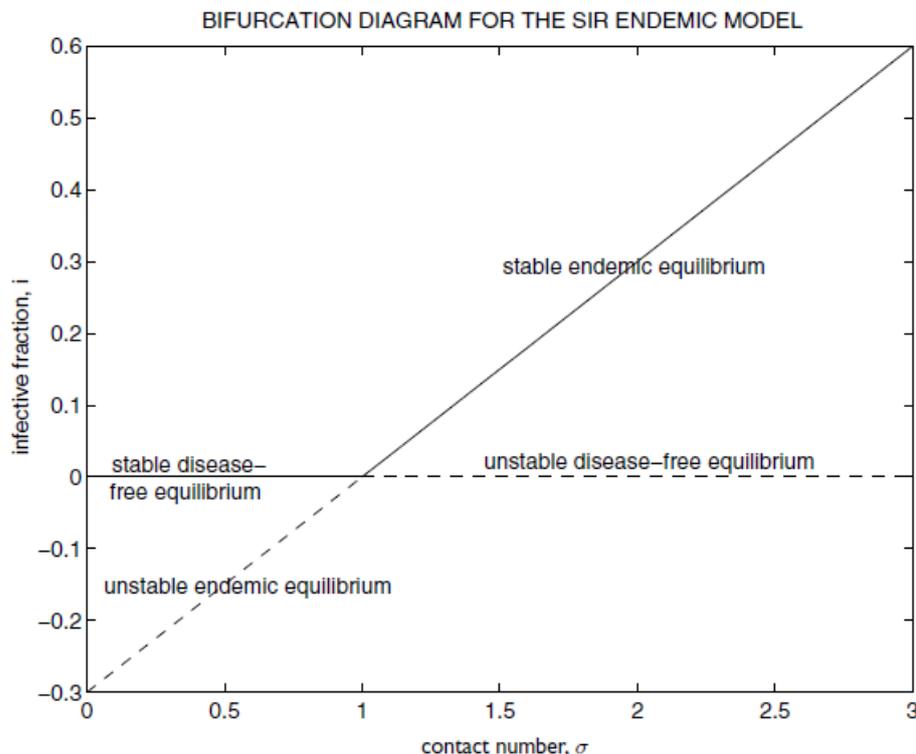


Figura 1.5: Diagramma di biforcazione per il modello SIR nel caso endemico, che mostra che il punto di equilibrio in assenza della malattia e l'equilibrio allo stato endemico, cambiano stabilità quando il valore di contatto ρ_0 (qui rappresentato come σ) è 1

Notiamo che i_e , coordinata del punto di equilibrio endemico, è negativa quando $\sigma < 1$; coincide con il valore nullo dell'equilibrio in assenza della malattia quando $\sigma = 1$ e diventa positivo quando $\sigma > 1$. Tale equilibrio, dato da $s_e = \frac{1}{\rho_0}$ e da $i_e = \frac{\mu(\rho_0 - 1)}{N\lambda}$ è instabile quando $\rho_0 < 1$ ed è localmente asintoticamente stabile quando $\rho_0 > 1$; mentre l'equilibrio in assenza della malattia dato da $s = 1$ e da $i = 0$ è localmente stabile quando $\rho_0 < 1$ e instabile quando $\rho_0 > 1$. Questi due valori di equilibrio si scambiano di stabilità, quando l'equilibrio endemico passa attraverso il valore di equilibrio in assenza di malattia per $\rho_0 = 1$ e diventa punto di equilibrio localmente asintoticamente stabile quando $\rho_0 > 1$. Le tipologie di modelli SIR finora descritti, hanno chiaramente delle limitazioni.

Esse infatti assumono che la popolazione sia uniforme e miscelata in modo omogeneo al suo interno, mentre è invece risaputo che tale commistione è dovuta a molti fattori (i bambini ad esempio hanno più contatti in un giorno che un

adulto). Inoltre la percentuale di contatti varia anche in funzione di differenti aree geografiche e del tipo di organizzazione socio-economica di quella particolare popolazione. Tuttavia, nonostante le loro approssimazioni, i modelli SIR classici possono essere utilizzati per ottenere stime e fare paragoni molto utili e interessanti anche nei contesti più attuali.

Capitolo 2

Diffusione di epidemie su networks

2.1 Considerazioni preliminari: networks e grafi

Possiamo definire in termini generali un *network* come un qualunque sistema che ammetta una rappresentazione matematica tramite un grafo, i cui *nodi* (o *vertici*) identifichino gli elementi del sistema e in cui i segmenti di congiunzione rappresentino una relazione o interazione fra questi elementi. Un *grafo* è definito tramite una coppia di insiemi $G=(\nu,\epsilon)$, dove ν è un insieme non vuoto e misurabile, i cui elementi sono i nodi del sistema, ed ϵ è la collezione dei segmenti di congiunzione fra coppie non ordinate di differenti vertici, ovvero i rami. Ci si può riferire a un determinato vertice tramite un indice i che lo ordina all'interno di ν . Il ramo (i,j) congiunge i vertici i e j , che vengono detti *connessi*. È inoltre convenzionale chiamare i vertici connessi, *vicini* (*neighbors* o *nearest neighbors*). Il numero totale di vertici del grafo (ovvero la cardinalità dell'insieme ν) è denotata con N e definisce l'ordine del grafo. La dimensione del grafo invece, è identificato dal numero totale dei rami E . Tuttavia è consueto trovare in molti contesti fisici e biologici, riferimento ad N come dimensione del grafo, in quanto identifica il numero di distinti elementi che compongono il sistema. Per un grafo di dimensione N , il massimo numero di rami è $\binom{N}{2}$. Un grafo in cui $E=\binom{N}{2}$, ovvero in cui tutte le possibili coppie di vertici sono collegate, è chiamato *grafo completo*. Si parla invece di *grafo diretto* quando i segmenti di congiunzione dei vertici sono orientati, e quindi l'insieme delle coppie di distinti vertici appartenenti a ϵ , è ordinato. Graficamente questo può essere rappresentato da frecce indicanti il verso di percorrenza del ramo. La

principale differenza fra *grafi indiretti* e *diretti* è che, nel primo caso, la presenza di un ramo fra i e j collega i vertici in entrambe le direzioni mentre, nel secondo caso, non si presuppone la presenza dell'inverso fra i due nodi. Da un punto di vista matematico, è conveniente definire un grafo nei termini di una matrice $X=x_{ij}$, ovvero una matrice $N \times N$ definita come

$$x_{ij} = \begin{cases} 1 & se(i, j) \in \epsilon \\ 0 & se(i, j) \notin \epsilon \end{cases}$$

Per grafi indiretti questa matrice è simmetrica ($x_{ij} = x_{ji}$), e quindi contiene spesso informazioni ridondanti. Per grafi diretti invece la matrice X non è simmetrica.

Punto cardine nella caratterizzazione di grafi è il concetto di 'raggiungibilità' dei vertici, ovvero la possibilità di andare da un vertice all'altro seguendo le connessioni tracciate dai rami. In un network connesso, ogni vertice è raggiungibile da un qualunque altro vertice. Le componenti connesse di un grafo definiscono dunque molte delle sue proprietà fisiche. Un cammino P_{i_0, i_n} in un grafo $G=(\nu, \epsilon)$ può essere definito come una collezione ordinata di $n + 1$ vertici $\nu_p = i_0, i_1, \dots, i_n$ e dei rami che collegano ciascuno di questi vertici. La lunghezza di tale cammino sarà dunque n .

Un *ciclo*, è un cammino chiuso ($i_0 = i_n$) in cui tutti i vertici e tutti i rami sono distinti. Un grafo è detto *connesso* se esiste un cammino che colleghi ogni coppia di vertici.

Una *componente* C è un sotto grafo connesso se non è possibile costruire un cammino $P_{i, j}$ che colleghi il vertice i della prima componente, col vertice j della seconda componente e se risulta inoltre disconnesso da un qualunque altro di tali sotto grafi.

Il concetto di *cammino* pone le basi per la definizione di distanza fra i vertici, di cui una misurazione naturale è definita come il numero di rami che il cammino più breve deve percorrere per collegare due vertici. La lunghezza di tale cammino viene denotata come l_{ij} . Quando due vertici appartengono a componenti disconnesse di un grafo, si usa definire $l_{ij}=\infty$; è invece una quantità simmetrica per grafi indiretti, mentre in generale l_{ij} non coincide con l_{ji} per grafi diretti. Il valore medio della lunghezza del cammino più breve, è definito come la media del valore l_{ij} fra tutte le possibili coppie di vertici presenti nel network

$$\langle l \rangle = \frac{1}{N(N-1)} \sum_{ij} l_{ij}$$

Per un grafo completo vale $\langle l \rangle = 1$.

2.2 Grado e misure di centralità

2.2.1 Degree centrality

Il grado k_i di un vertice è definito come il numero di rami del grafo incidenti al vertice i . Possiamo dunque distinguere fra *in-degree* $k_{in,i}$ del vertice i , come il numero di linee che arrivano ad esso, e *out-degree* $k_{out,i}$, definito come il numero di rami uscenti da i . Il grado di un vertice in un grafo diretto è definito come la somma di questi due gradi, mentre invece per un grafo indiretto con una matrice di rappresentazione simmetrica vale $k_{in,i} = k_{out,i}$.

2.2.2 Closeness centrality

Questo valore esprime la distanza media di un vertice da tutti gli altri, come

$$g_i = \frac{1}{\sum_{i \neq j} l_{ij}}$$

Questa misura assegna un alto valore della centralità a quei nodi che hanno distanze minori rispetto ad altri.

2.2.3 Betweenness centrality

Viene definito come il numero dei cammini più brevi fra coppie di vertici, che passano attraverso un determinato vertice. Più precisamente, se σ_{hj} è il numero totale dei cammini più brevi da h a j e $\sigma_{hj}(i)$ è il numero di questi cammini che passano proprio attraverso il vertice i , il valore di *betweenness* di i è definito come

$$b_i = \sum_{h \neq j \neq i} \frac{\sigma_{hj}(i)}{\sigma_{hj}}$$

2.3 Clustering

Accanto alle misure di centralità, i vertici sono caratterizzati dalla struttura locale dei loro 'vicini'. Il concetto del *clustering* di un grafo si riferisce alla tendenza osservata in molti networks, di formare cicli nell'intorno di un dato vertice. Più precisamente questo concetto implica la proprietà che, se un vertice i è connesso al vertice j , e allo stesso tempo j è connesso a l , allora con elevata probabilità, il vertice i risulterà anche connesso a l . Dato un vertice i , il valore di *clustering* $C(i)$ di un nodo i è definito come il rapporto del numero di collegamenti che esistono fra i vicini di i e il numero massimo di questi collegamenti.

Se il grado del nodo i è k_i e se questi nodi hanno e_i cammini fra loro, vale

$$C(i) = \frac{e_i}{\frac{k_i(k_i-1)}{2}}.$$

$C(i)$ assume valori compresi fra 0 e 1 e ha significato solo per $k_i > 1$. Per $k_i \leq 1$, definiamo $C(i) \equiv 0$. Il valore medio del coefficiente di clustering di un grafo, è dato semplicemente da

$$\langle C \rangle = \frac{1}{N} \sum_i C(i)$$

2.4 Degree distribution

La *degree distribution* $P(k)$ di grafi indiretti è definita come la probabilità che un qualunque vertice scelto casualmente abbia grado k . Nel caso di grafi diretti, occorre considerare invece due distribuzioni, una sarà di *in-degree* e l'altra di *out-degree*, rispettivamente per il grado k_{in} e il gradi k_{out} di ciascun vertice. Il valore medio di un grafo indiretto è definito come il valore medio di k fra tutti i vertici nel network,

$$\langle k \rangle = \frac{1}{N} \sum_i k_i = \sum_k k P(k) \equiv \frac{2E}{N}$$

in quanto ogni estremo di un ramo contribuisce al grado di un vertice. Per un grafo diretto, vale la relazione

$$\langle k_{in} \rangle = \langle k_{out} \rangle = \frac{\langle k \rangle}{2}$$

in quanto ogni ramo uscente da un vertice dovrà necessariamente arrivare a un altro vertice. Da tali considerazioni segue la definizione di momento n-esimo per una distribuzione del grado come,

$$\langle k^n \rangle = \sum_k k^n P(k)$$

2.5 Betweenness distribution

È possibile introdurre la distribuzione di probabilità $P_{b'}(b)$ che un vertice abbia *betweenness* b' assegnato, e valore medio di *betweenness* $\langle b \rangle$ definito come

$$\langle b \rangle = \sum_b b P_{b'}(b) \equiv \frac{1}{N} \sum_i b_i$$

Possiamo quindi caratterizzare $P_{b'}(b)$ come la cardinalità dell'insieme dei vertici i con valore di *betweenness* b_i uguale a b' , divisa N . Per questa quantità è importante mostrare la relazione con il valore medio della lunghezza del cammino più

breve $\langle l \rangle$. Semplicemente riordinando le somme nella definizione di betweenness otteniamo:

$$\sum_i b_i = \sum_{h \neq j} \frac{1}{\sigma_{hj}} \sum_{i \neq h, j} \sigma_{hj}(i).$$

Poichè per motivi puramente topologici vale $\sum_{i \neq h, j} \sigma_{hj}(i) = \sigma_{hj}(l_{hj} - 1)$, che sostituita nella relazione precedente, deriva che $\sum_i b_i = N(N - 1)(\langle l \rangle - 1)$. Da cui

$$\langle b \rangle = (N - 1)(\langle l \rangle - 1)$$

2.6 Network e il mondo reale

Nei tempi recenti, l'aumentato potere di computer e la rivoluzione informatica hanno reso possibile la raccolta e il trattamento di dati da diversi networks su larga scala. Due sono le classi principali ovvero sistemi di infrastrutture, e sistemi naturali o viventi. Di particolare interesse per la nostra trattazione sono i *social networks*, che descrivono relazioni fra individui, quali ad esempio rapporti familiari o di lavoro, amicizie.

2.6.1 Social networks

I networks in ambito sociale possono essere descritti da nodi che rappresentano gli individui di una popolazione, e le relazioni sociali fra di essi sono i cammini che collegano questi nodi. Questi collegamenti possono dunque fare riferimento a ragioni molto diverse come ad esempio legami di amicizia fra compagni, o interazioni sessuali fra adulti, o semplicemente l'appartenenza a uno stesso gruppo o team di lavoro. L'importanza di questi network va al di là delle scienze sociali e hanno effetti sulla possibilità di comprensione di una grande varietà di processi, dalla diffusione di malattie sessualmente trasmissibili alla diffusione di conoscenze in diversi tipi di organizzazioni e strutture sociali.

2.6.2 Small-world

Si fa riferimento a una struttura di questo tipo quando è possibile muoversi da un vertice a un qualunque altro punto, passando attraverso un numero molto piccolo di vertici intermedi. Situazioni di questo tipo sono state riscontrate in una vasta gamma di network, in particolare in quelli riguardanti le infrastrutture dove un valore piccolo per le distanze è di cruciale importanza per trasmettere celermente informazioni.

Per essere più precisi, le proprietà di uno *small-world network* si riferiscono a

network in cui $\langle l \rangle$ varia in modo logaritmico, o più lentamente, rispetto al numero di vertici. Analogamente il valore medio del numero di nodi compresi per una distanza minore o uguale a l , segue un andamento esponenziale o di crescita anche più veloce.

La casualità appare essere invece la caratteristica di maggiore importanza. Infatti, data l'intricata e apparente assenza di ordine di questi tipi di network, si è portati a fare come assunzione più semplice ed immediata che la creazione di una connessione fra due elementi sia un evento casuale, determinato dalla somma di un numero molto vasto di circostanze imprevedibili. In breve questi modelli presuppongono che la presenza di un cammino fra due vertici sia un processo casuale che ricorre con stessa probabilità indipendentemente dalle caratteristiche di un vertice.

I risultati ottenuti da una situazione descrivibile come *small-world*, possono dunque essere spiegati dalla semplice inclusione della casualità e vanno a pari passo con un alto livello di *clustering*. Il coefficiente di clustering caratterizza la locale coesione dei network, come la tendenza a formare un gruppo di elementi interconnessi. In grafi casuali di grandi dimensioni, il coefficiente di clustering risulta infatti molto piccolo. Grafi casuali descrivono bene il caso di *small-world*, cosa che invece non fanno delle strutture a griglie regolari.

2.6.3 Eterogeneità e code lunghe

Quando si parla di *network omogenei* si fa riferimento a distribuzioni, caratterizzanti il grado, betweenness e quantità pesate, aventi espressioni con andamenti simili a decadimenti veloci o, come ad esempio distribuzioni Gaussiane o di Poisson. Un'altra possibile classe raccoglie i network con connettività *eterogenea*, le cui distribuzioni sono non limitate esponenzialmente e dissimmetriche. L'evidenza di un alto livello di eterogeneità di molti network è semplicemente fornito dal fatto che possono essere presenti anche molti vertici con poche connessioni. La distribuzione è dissimmetrica e altamente variabile nel senso che i gradi variano in un intervallo molto ampio, abbracciando vari ordini di grandezza. Questo comportamento è molto diverso da quello di una distribuzione con decadimento esponenziale, e in molti casi può essere approssimato con un decadimento secondo legge di potenza, del tipo $p(k) \sim k^{-\gamma}$. Questa proprietà si traduce in fluttuazioni molto ampie del grado.

È possibile calcolare facilmente il valore medio che il grado assume in un network come

$$\langle k \rangle = \int_m^{\infty} kP(k)dk,$$

dove $m \geq 1$ è il grado più basso possibile del network.

Una stima dell'errore che si commette, assumendo che $\langle k \rangle$ sia il valore del grado di un vertice è dato dalla varianza normalizzata della distribuzione $\frac{\sigma^2}{\langle k \rangle^2}$ che esprime le fluttuazioni statistiche presenti nel nostro sistema. La varianza $\sigma^2 = \langle k^2 \rangle - \langle k \rangle^2$ è dominata dal secondo momento della distribuzione:

$$\langle k^2 \rangle \sim \int_m^\infty k^2 P(k) dk \sim \int_m^{k_c} k^{2-\gamma} dk.$$

Nel caso di network finiti, k_c è un valore finito che corrisponde al più grande grado assumibile. Per un network di dimensione infinita possiamo dare una stima di questo integrale con il limite asintotico per $k_c \rightarrow \infty$, in modo che anche $\langle k^2 \rangle \rightarrow \infty$. L'assenza di una qualsiasi scala dei tempi intrinseca a causa delle fluttuazioni del grado in network a coda lunga implica che il valore medio non sia una scala temporale rappresentativa per il sistema. In altre parole, possiamo più semplicemente osservare un network di tipo *scale-free* ogni volta che si prende in considerazione il grado.

È inoltre possibile fornire una caratterizzazione del livello di eterogeneità dei network dalla definizione del seguente parametro

$$\kappa = \frac{\langle k^2 \rangle}{\langle k \rangle}$$

Gli *scale-free networks* sono caratterizzati da $\kappa \rightarrow \infty$, mentre per i network omogenei vale $\kappa \sim \langle k \rangle$. Per questa ragione, si fa generalmente riferimento a *scale-free networks* come a tutti quei networks eterogenei con parametro $\kappa \gg \langle k \rangle$. Sottolineiamo il fatto che questo parametro κ è di cruciale importanza per derivare tutte le proprietà e i processi fisici che sono affetti da fluttuazioni di grado.

Sebbene la presenza o l'assenza di forme che seguono un andamento descrivibile da una legge di potenza, sono problemi ben posti da un punto di vista statistico, è chiaro che nella trattazione di sistemi reali, in molti casi, le proprietà statistiche risulteranno di dimensione finita, o affette da rumori, o soggette ad altri effetti di troncamento. Quindi in molte situazioni il comportamento risulterà ben più complicato di una pura legge di potenza. *Scale-free networks* fanno dunque riferimento a tutti quei sistemi in cui le fluttuazioni hanno ordini di grandezza più grandi dei valori attesi.

2.7 Diffusione di epidemie all'interno di network di popolazioni

Networks e modelli di diffusione di epidemie sono ambiti strettamente connessi. Possiamo immaginare che questi sistemi siano basati su una popolazione in cui ogni individuo rappresenti un nodo e i rami esprimano le interazioni fra di essi, ovvero i contatti attraverso i quali si possono trasmettere le infezioni. L'insieme di tutti questi contatti forma un *'mixing network'*. La conoscenza della struttura del network e delle proprietà definenti il tipo di contatto fra gli individui, consente ai modelli matematici di ridisegnare le dinamiche epidemiche di diffusione di un contagio. Si possono perciò analizzare modelli di trasmissione, quali SIR, appropriato per malattie infettive che conferiscono immunità di lunga durata e SIS, usato molto nei casi di malattie sessualmente trasmissibili e per malattie ricorrenti.

Come studiato nel capitolo precedente, un parametro comune a tutti i modelli è la forza di infezione, che rappresenta il tasso con cui individui suscettibili divengono infetti, ed è una funzione del numero degli infettivi; questo parametro contiene informazioni sulle interazioni fra gli individui e si riconduce alla trasmissione dell'infezione. Quando la popolazione interagisce in modo casuale, così che ogni individuo abbia una piccola ed uguale probabilità di entrare in contatto con un qualunque altro individuo, la forza di infezione può essere calcolata come:

$$\lambda = (\text{tasso di trasmissione}) \times (\text{numero effettivo di contatti nell'unità di tempo}) \times (\text{proporzione di contatti infettivi}) = \tau \times m \times \frac{I}{N} = \frac{\beta I}{N}.$$

Spesso inoltre ci si può trovare davanti all'esigenza di considerare diversi tassi di interazione fra sottogruppi della popolazione (ad esempio i bambini entrano in contatto fra loro molto più facilmente che con gli adulti), e questo può essere tradotto in un modello in cui al posto del singolo parametro β si sostituisce una matrice B di parametri di trasmissione, e che descriva la trasmissione dell'infezione fra i diversi gruppi.

Nei modelli *random-mixing* gli individui possono sì avere lo stesso numero di contatti effettivi nell'unità di tempo, ma questo valore cambia continuamente. La presenza di relazioni fra individui, fa sì che si possa definire *'grado'* degli individui, l'insieme dei contatti di un individuo e la dimensione di questo insieme. Particolare attenzione viene data alla natura delle connessioni e in particolare alle proprietà di simmetria (una relazione fra A e B implica una relazione anche fra B e A) e di transitività (del tipo l'amico di un amico è esso stesso un

amico), che sono le misure caratteristiche per determinare la coesione sociale e rispecchiano le proprietà caratteristiche dei network viste in precedenza.

La ricerca nella teoria dei grafi ha fornito una vasta gamma di strumenti e meccanismi per la descrizione di networks, molte dei quali hanno applicazioni epidemiologiche. Possiamo utilizzare una 'socio matrice', A , per descrivere le connessioni all'interno della popolazione. Se $A_{ij} = 1$, esiste una connessione attraverso la quale un'infezione può essere trasmessa dall'individuo i all'individuo j ; in caso contrario vale $A_{ij} = 0$. A riassume quindi tutte le connessioni all'interno del network. Per una popolazione di dimensione N , il numero medio di contatti per individuo è $\bar{n} = \frac{1}{N} \sum_{ij} A_{ij}$. La matrice A^m contiene invece informazioni riguardo i cammini di lunghezza m all'interno del network, e le potenze della matrice possono essere utilizzate per dare una stima della transitività e del clustering. In ultimo, diciamo che un grafo è connesso se ogni individuo (nodo) può essere raggiunto da ogni altro tramite i rami del network; da un punto di vista epidemiologico, questo è equivalente al caso di una infezione capace di raggiungere l'intera popolazione da un qualunque punto di partenza, cosa che si verifica se $\sum_{m=1}^{\infty} A^m$ non ha termini nulli. Gli zeri di queste matrici dimostrano che il network è suddiviso in due o più componenti separate, nessuna delle quali è collegata ad altre.

Supposto che p rappresenti la probabilità per cui i nodi vicini sono connessi casualmente, quando questa probabilità è sufficientemente grande, è possibile trovare un cammino da un punto all'altro del grafo. Se i networks sono utilizzati per scopi epidemiologici, allora le connessioni devono essere incluse solo se descrivono relazioni capaci di consentire la diffusione dell'epidemia.

2.7.1 Random networks

Nei random networks, la posizione spaziale degli individui è irrilevante, e le connessioni vengono stabilite in modo casuale, in quanto c'è un'assoluta mancanza di conoscenza dei principi che guidano la creazione di un ramo fra nodi. In mancanza di queste informazioni, l'assunzione più semplice che si possa fare è di connettere coppie di nodi casualmente con una data probabilità p . Ne consegue che ogni individuo ha un numero fissato di contatti attraverso i quali l'infezione può dilagare.

Si può costruire il grafo partendo da N diversi vertici in cui ognuno degli $\frac{N(N-1)}{2}$ possibili rami è presente con probabilità p e assente con probabilità $(1-p)$.

Otteniamo una distribuzione di probabilità di tipo binomiale, dove il numero medio di rami generati nella costruzione del grafo è $\langle E \rangle = \frac{1}{2} N(N-1)p$. Poiché ogni ramo contribuisce al grado di due vertici, si ottiene $\langle k \rangle = \frac{2\langle E \rangle}{N} \simeq Np$.

Questi tipi di network sono caratterizzati dalla mancanza del clustering e di omogeneità, da un tasso di crescita iniziale della malattia e da una dimensione finale dell'epidemia più ridotti rispetto ai modelli random-mixing.

$$\text{Tasso di crescita nei random network} = \tau(n - 1) - g$$

$$\text{tasso di crescita con random-mixing} = \beta - g = \tau m - g$$

dove τ è il tasso di trasmissione per contatto, n è il numero di contatti nel network, ed m è l'effettivo numero di contatti per unità di tempo in un modello random-mixing e g il tasso di guarigione dall'infezione. La riduzione del tasso di crescita è conseguenza di due fattori: prima di tutto ogni individuo infetto è stato a sua volta infettato da uno dei suoi contatti, riducendo il numero dei contatti dei suscettibili a $n - 1$; in secondo luogo, non appena un individuo infettato inizia a infettare i suoi contatti suscettibili, limita il tasso di diffusione della malattia.

Una formulazione alternativa di un random network può esser fatta quando N è grande per cui si ottiene una distribuzione di grado approssimabile da una funzione di Poisson, dove il numero medio di contatti per nodo varrà $\bar{n} = p(N - 1)$, (N è il numero totale di nodi). In un siffatto network, il tasso di crescita risulterà comunque ancora ridotto:

$$\text{tasso di crescita in un random network di tipo poissoniano} = \tau \frac{(\bar{n} - 1)\bar{n}}{(\bar{n} + 1)} - g.$$

2.7.2 Modelli su reticolo

Gli individui sono posizionati su una griglia regolare di punti, solitamente in due dimensioni, e connessi agli individui adiacenti; ne segue che i contatti sono localizzati nello spazio. Questi tipi di networks sono omogenei e possiedono un alto valore di clustering. Anche questi modelli mostrano un ridotto tasso di crescita iniziale dell'infezione se posti a confronto con modelli random-mixing, nonostante questo effetto sia molto più rilevante che nei random networks perché il clustering spaziale dei contatti causa una più rapida saturazione dell'ambiente locale. In generale i modelli lattices mostrano una diffusione dell'infezione a onda, ovvero da una fonte iniziale l'infezione si propaga in un modo approssimativamente circolare.

Per questo tipo di modelli possiamo trovare una distribuzione approssimabile da una legge di potenza $\propto s^{-\alpha}$, dove α varia a seconda della malattia.

2.7.3 Small-world networks

I modelli su reticolo mostrano un alto livello di clustering ma anche cammini lunghi, che significa che si necessita di molti rami per muoversi fra due individui scelti casualmente. Gli small-world networks offrono in effetti un mezzo di spostamento, in quanto possono essere formati dall'aggiunta di un piccolo numero di connessioni casuali a un lattice network, consentendo quindi alla malattia di raggiungere tutte le parti del grafo, relativamente in fretta. Tuttavia, poiché queste connessioni a lungo raggio sono rare, la trasmissione dell'infezione rimane predominantemente localizzata, così che sono ancora osservabili effetti di forte saturazione e di diffusione di epidemie a onda.

Gli small-world networks sono caratterizzati da un alto valore del clustering e da cammini di piccola lunghezza. L'alto livello di clustering significa che la maggioranza delle infezioni intervengono localmente, ma i cammini di piccola lunghezza implicano che la diffusione dell'epidemia attraverso il network sia rapida e che la malattia rimanga difficilmente contenuta all'interno di piccole regioni della popolazione. Connessioni casuali ad ampio raggio all'interno del network possono drammaticamente aumentare la probabilità di nascita di una vera e propria epidemia. L'evoluzione virulenta di un agente patogeno nel caso di small-world networks, la presenza di collegamenti anche a lunga distanza e cammini di breve lunghezza suggeriscono che la diffusione spaziale della malattia sia complessivamente estremamente rapida.

2.7.4 Spatial networks

Questi tipi di networks sono fra le forme più flessibili conosciute. Gli spatial networks sono caratterizzati da un ragionevolmente alto livello di eterogeneità, con un grado di distribuzione solitamente approssimabile da una distribuzione di Poisson.

2.7.5 Scale-free networks

Gli scale-free networks forniscono gli strumenti per realizzare livelli elevati di eterogeneità. Possono essere costruiti aggiungendo nuovi individui a un network preesistente, con un meccanismo di connessione che imiti la naturale formazione dei contatti sociali. Ogni nuovo individuo aggiunto alla popolazione si metterà in contatto più preferibilmente a quegli individui che possiedono già un vasto numero di rapporti (corrisponde al caso di soggetti che vogliono divenire amici della persona più popolare). Questo porta come conseguenza ad avere un numero di contatti per individuo che assume una distribuzione di

potenza. La presenza di *super-spreaders* e di gruppi centrali ricoprono un ruolo cardine nella diffusione dell'infezione, la cui estrema eterogeneità nel numero dei contatti mostrato da questo tipo di network, ha due effetti principali: tali individui sono esposti a un più elevato rischio di infezione, e una volta infettati possono trasmettere la malattia a molti altri. Inoltre l'infezione è concentrata fra gli individui con il grado più alto.

L'esistenza di individui con un grado arbitrariamente elevato significa che non c'è una percentuale di vaccinazioni sufficientemente alta da poter prevenire l'epidemia. Al contrario, quando è fissato un limite superiore per il grado degli individui, o quando uno scale-free network è generato dall'aggiunta del vicino più prossimo all'interno di una struttura a reticolo, diviene possibile controllare il meccanismo infettivo tramite la vaccinazione random. Una vaccinazione mirata per gli scale-free network è estremamente efficiente: il ruolo dominante degli super-spreaders implica che la vaccinazione anche solo di pochi di questi individui possa essere sufficiente a prevenire l'epidemia.

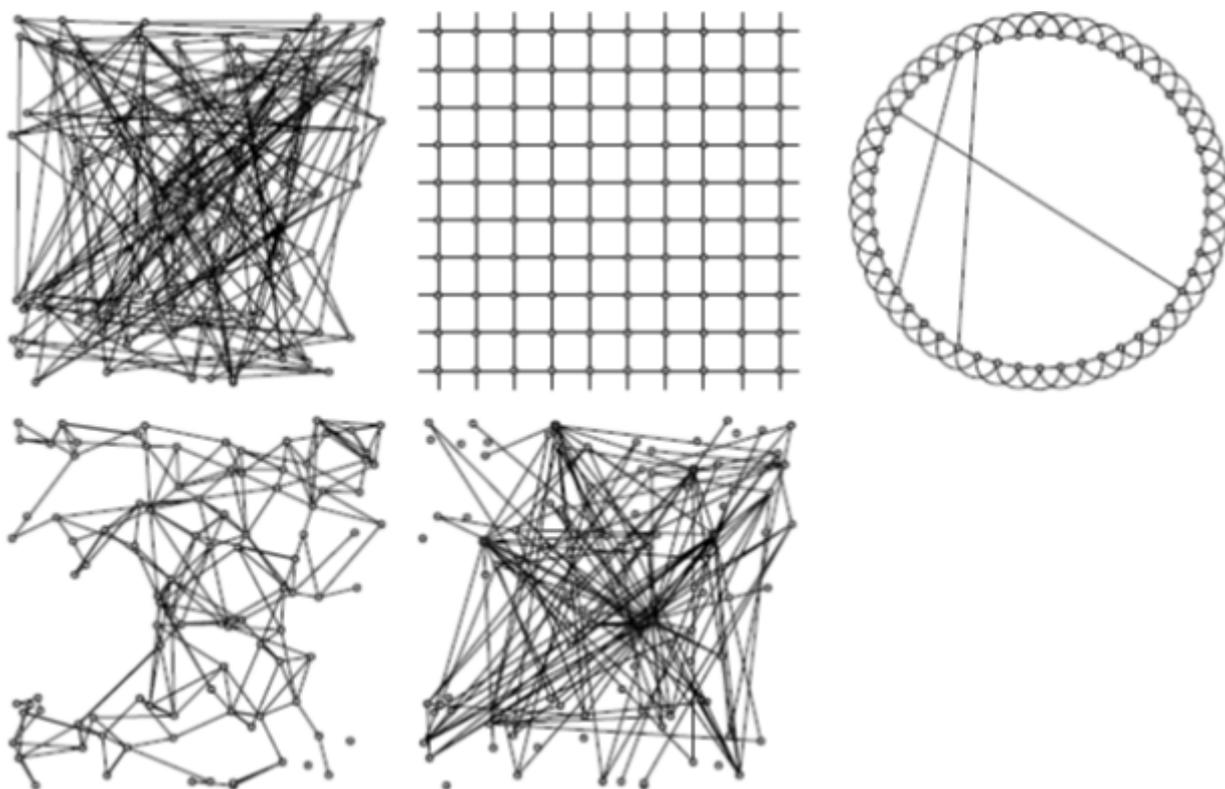


Figura 2.1: Cinque tipi diversi di networks, da sinistra a destra e dall'alto in basso: random, lattice, small-world, spatial e scale-free.

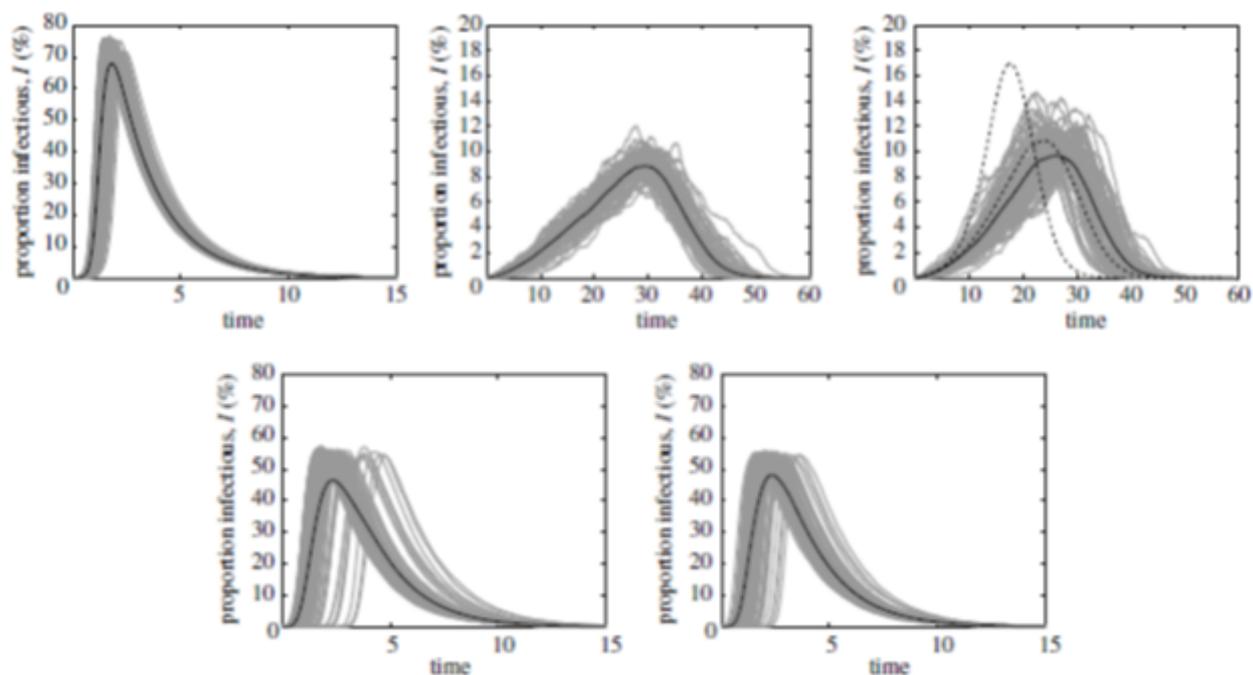


Figura 2.2: Questa figura mostra il tipico modello SIR di diffusione di un'epidemia, all'interno delle cinque tipologie di networks, che sono da sinistra a destra: random, lattice, small world, e (al di sotto) spatial e scale-free. Ogni grafo mostra 100 curve per l'epidemia (in grigio), assieme con il valore medio dei valori più alti dell'epidemia (in nero)

2.8 Richiami

Per quanto visto nel primo capitolo sui modelli di dinamica di epidemie, questi ultimi sono generalizzabili in tre categorie principali quali il modello SI (il più semplice ed elementare), il modello SIR e SIS. Definiamo con β il tasso di diffusione della malattia, e cerchiamo di definire esplicitamente la dipendenza di questo tasso con il numero di contatti che ogni individuo può avere nella popolazione.

Come prima approssimazione, si consideri che ogni individuo/vertice abbia lo stesso numero di contatti/rami, $k \cong \langle k \rangle$. L'equazione che descrive l'evoluzione del modello SI, può essere espressa come

$$\frac{di(t)}{dt} = \beta \langle k \rangle i(t) [1 - i(t)]$$

Questa equazione stabilisce che il tasso di crescita di individui infettati è proporzionale al tasso di diffusione $\beta\langle k \rangle$, alla densità dei vertici suscettibili che possono divenire infetti $s(t) = 1 - i(t)$, (dove $s(t) = \frac{S(t)}{N}$), e al numero di individui infettati che sono in contatto con individui suscettibili. Questo modello di base può essere facilmente risolvibile alle prime fasi dell'insorgere dell'epidemia, quando possiamo cioè assumere che il numero di individui infettati sia una frazione molto piccola della popolazione totale ($i(t) \ll 1$).

In questo caso l'equazione per il modello SI può essere riscritta come $\frac{di(t)}{dt} = \beta\langle k \rangle i(t)$, la cui soluzione è data $i(t) \cong i_0 e^{\frac{t}{\tau}}$, con i_0 la densità iniziale degli individui infettati e dove τ è il momento di massima esplosione dell'infezione

$$\tau^{-1} = \beta\langle k \rangle - \mu$$

con μ , tasso di guarigione. Il tempo caratteristico in forma esponenziale è combinazione di due termini e può assumere valori negativi se il tasso di guarigione è sufficientemente grande. In questo caso l'epidemia non si propagherà nella popolazione, mentre invece svanirà gradualmente nel tempo con andamento $|\tau|$. Queste considerazioni portano alla definizione di un concetto cruciale da un punto di vista epidemiologico, ovvero quello di equilibrio. La condizione di equilibrio per l'epidemia può essere scritto nella forma

$$\tau^{-1} = \mu(R_0 - 1) > 0$$

dove $R_0 = \frac{\beta\langle k \rangle}{\mu}$ identifica il tasso di riproduttività di base nei modelli SIR e SIS, e deve essere maggiore di 1 affinché si verifichi la diffusione dell'epidemia. Infatti: quando $\tau^{-1} = \mu(R_0 - 1) > 0$, $R_0 > 1$, segue che il valore di $e^{\frac{t}{\tau}} \rightarrow \infty$; quando $\tau^{-1} < 0$, $R_0 < 1$ segue $e^{\frac{t}{\tau}} \rightarrow 0$.

Si ha il valore di equilibrio per $e^{\frac{t}{\tau}} = 0$, per cui $R_0 = 1$.

2.9 Epidemie su network eterogenei

La situazione generale presentata nella sezione precedente è stata ottenuta ponendo delle ipotesi di omogeneità. Tuttavia, molti dei networks che descrivono sistemi di epidemiologica rilevanza esibiscono una struttura estremamente eterogenea. La trattazione di casi che presentano queste fluttuazioni può essere fatta considerando una approssimazione a blocchi rispetto al grado, ovvero assumendo che tutti i nodi con lo stesso grado siano statisticamente equivalenti. Questa assunzione ci consente di raggruppare i nodi con lo stesso grado in una classe k , portando a una conveniente rappresentazione del sistema tramite quantità quali la densità dei nodi infettati e suscettibili all'interno della classe

di grado k

$$i_k = \frac{I_k}{N_k}; \quad s_k = \frac{S_k}{N_k}$$

dove N_k è il numero di nodi di grado k e I_k e S_k sono rispettivamente il numero di nodi infettati e suscettibili nella classe. Chiaramente, le medie globali sono date dalle espressioni

$$i = \sum_k P(k) i_k; \quad s = \sum_k P(k) s_k;$$

2.9.1 Modello SI

Per questo modello ($S \rightarrow I$) sappiamo che è possibile che si verifichi il caso in cui il sistema sia totalmente infettato, indipendentemente dal tasso di propagazione dell'infezione, ma è tuttavia interessante vedere gli effetti di fluttuazioni topologiche sulla velocità di diffusione. Per il modello SI l'equazione di evoluzione può essere espressa come

$$\frac{di_k(t)}{dt} = \beta[1 - i_k(t)]k\Phi_k(t)$$

dove il termine di creazione è proporzionale al tasso di diffusione β , al grado k , alla probabilità $1 - i_k$ che ogni vertice con grado k non sia infettato, e la densità Φ_k dei vicini infettati di vertici con grado k . L'ultimo termine esprime quindi la probabilità media che ogni dato vicino di un vertice di grado k sia infettato. Questa è una nuova e sconosciuta quantità che nel caso di assunzione di omogeneità, è uguale alla densità dei nodi infettati. Per un network eterogeneo invece, questa è un'espressione complicata che deve tener conto delle differenti classi di grado e delle loro connessioni.

La situazione più semplice che possiamo trovare, corrisponde al caso in cui ci sia una mancanza totale di correlazioni dei gradi. Si parla di network che non hanno correlazioni di grado, quando la probabilità che un ramo uscente da un vertice di grado k arrivi a un vertice di grado k' è indipendente dal grado k del vertice iniziale. In questo caso, la probabilità condizionata non dipende dal nodo di partenza ed è possibile mostrare che $P(k|k') = k' \frac{P(k')}{\langle k \rangle}$. Questa relazione stabilisce semplicemente che ogni ramo ha una probabilità di puntare a un nodo di grado k' che è proporzionale a k' stesso. Considerando che almeno uno dei rami uscenti da ogni vertice infettato punta a un altro vertice infettato dal quale l'infezione è stata trasmessa, si ottiene

$$\Phi_k(t) = \Phi(t) = \frac{\sum_{k'} (k' - 1) P(k') i_{k'}(t)}{\langle k \rangle}$$

dove $\langle k \rangle = \sum_{k'} k' P(k')$. Notiamo che $\Phi_k(t)$ è indipendente da k per la mancanza di correlazione fra gradi dei vertici vicini. Combinando le precedenti due equa-

zioni, omettendo i termini di ordine $O(i^2)$ e derivando l'espressione di $\Phi(t)$, si ottiene

$$\begin{aligned}\frac{di_k(t)}{dt} &= \beta k \Phi(t) \\ \frac{d\Phi(t)}{dt} &= \beta \left(\frac{\langle k^2 \rangle}{\langle k \rangle} - 1 \right) \Phi(t)\end{aligned}$$

Infatti derivando l'espressione di $\Phi(t)$:

$$\frac{d\Phi(t)}{dt} = \frac{1}{\langle k \rangle} \sum_{k'} (k' - 1) P(k') \frac{di_{k'}(t)}{dt},$$

in cui possiamo sostituire il valore delle prima delle due equazioni, e

$$\begin{aligned}\frac{d\Phi(t)}{dt} &= \frac{1}{\langle k \rangle} \sum_{k'} (k' - 1) P(k') \beta k' \Phi(t) \\ &= \frac{1}{\langle k \rangle} \beta \sum_{k'} ((k')^2 P(k') \Phi(t) - P(k') k' \Phi(t)) \\ &= \frac{1}{\langle k \rangle} \beta [\sum_{k'} k'^2 P(k') - k' P(k')] \Phi(t).\end{aligned}$$

Utilizzando la definizione di momento $\langle k^n \rangle = \sum k^n P(k)$, riscriviamo l'ultimo membro come

$$\frac{1}{\langle k \rangle} \beta \left[\sum_{k'} k'^2 P(k') - k' P(k') \right] \Phi(t) = \beta \left[\frac{\langle (k')^2 \rangle}{\langle k \rangle} - \frac{\langle k' \rangle}{\langle k \rangle} \right] \Phi(t).$$

Queste equazioni possono essere risolte e nel caso di una condizione iniziale uniforme $i_k(t=0) = i_0$, ottenendo un valore di *diffusione* di nodi di grado k

$$i_k(t) = i_0 \left[1 + \frac{k(\langle k \rangle - 1)}{\langle k^2 \rangle - \langle k \rangle} (e^{\frac{t}{\tau}} - 1) \right]$$

con

$$\tau = \frac{\langle k \rangle}{\beta(\langle k^2 \rangle - \langle k \rangle)}$$

La prevalenza di questi nodi ha una crescita con velocità esponenziale, e i nodi di grado più alto mostrano valori più elevati di distribuzione. Il valore medio della prevalenza è anche ottenuto come $i(t) = \sum_k P(k) i_k(t)$,

$$i(t) = i_0 \left[1 + \frac{\langle k \rangle^2 - \langle k \rangle}{\langle k^2 \rangle - \langle k \rangle} (e^{\frac{t}{\tau}} - 1) \right]$$

Il valore τ per networks non correlati implica che la scala dei tempi per l'insorgere dell'epidemia, è legata all'eterogeneità del grafo e al rapporto $\kappa = \frac{\langle k^2 \rangle}{\langle k \rangle}$. Nei networks omogenei con una distribuzione di grado di Poisson, in cui $\kappa = \langle k \rangle + 1$, ritroviamo il risultato $\tau = (\beta \langle k \rangle)^{-1}$, corrispondente all'ipotesi di mescolamento omogeneo. Nei network in cui invece sono presenti casi in cui la connettività

è molto eterogenea, κ assume un valore molto elevato e la scala dei tempi τ è molto piccola, il che segnala una diffusione molto veloce dell'infezione.

In particolare negli scale-free networks, caratterizzati da un esponente del grado $2 < \gamma \leq 3$ abbiamo che $\kappa \sim \langle k^2 \rangle \rightarrow \infty$ per i network in cui la dimensione $N \rightarrow \infty$. Ne consegue che per i network scale-free non correlati, possiamo assistere a un aumento quasi istantaneo dell'incidenza dell'epidemia. La ragione fisica di questo fenomeno risiede nel fatto che quando la malattia ha raggiunto i centri del network, questa può propagarsi molto rapidamente.

2.9.2 Modelli SIR e SIS

Nel caso di networks non correlati, l'equazione $\frac{di_k(t)}{dt} = \beta[1 - i_k(t)]k\Phi_k(t)$ conterrà, sia per il modello SIR che per il modello SIS, il termine aggiuntivo $-\mu i_k(t)$ che definisce il tasso con cui gli individui infettati di grado k guariscono e divengono nuovamente suscettibili o permanentemente immuni, rispettivamente:

$$\frac{di_k(t)}{dt} = \beta k s_k(t) \Phi_k(t) - \mu i_k(t)$$

Per il modello SIS vale come al solito, $s_k(t) = 1 - i_k(t)$. Nel modello SIR si ha $s_k(t) = 1 - i_k(t) - r_k(t)$, dove $r_k(t)$ è la densità degli individui rimossi di grado k .

I tempi nel caso SIR seguono un andamento del tipo

$$\tau = \frac{\langle k \rangle}{\beta \langle k^2 \rangle - (\mu + \beta) \langle k \rangle}$$

Nel caso di fluttuazioni divergenti, i valori del tempo sono dominati dal termine $\langle k^2 \rangle$ e la diffusione è tanto più veloce quanto maggiore è il livello di eterogeneità del network. Per quanto riguarda il valore di equilibrio vale

$$\frac{\beta}{\mu} \geq \frac{\langle k \rangle}{\langle k^2 \rangle - \langle k \rangle}$$

Questo fatto implica che nei networks in cui $\langle k^2 \rangle \rightarrow \infty$, quando $N \rightarrow \infty$, abbiamo un valore di equilibrio per l'epidemia nullo. Questo fatto ci mostra che i networks eterogenei si comportano in modo molto diverso rispetto a quelli omogenei. In particolare, l'assenza di un punto di equilibrio rende gli scale-free networks una sorta di ambiente ideale per la propagazione di virus, che sono quindi capaci di pervadere l'intero network anche nel caso di capacità di diffusione molto deboli.

2.9.3 L'effetto di modelli di mescolamento

Abbiamo fino ad ora ristretto il nostro studio al caso di networks non correlati, ma non tutti i networks reali soddisfano questa assunzione. Per semplicità, consideriamo le equazioni di evoluzione per un modello SI, che in queste diverse circostanze, valgono

$$\begin{aligned}\frac{di_k(t)}{dt} &= \beta[1 - i_k(t)]k\Phi_k(t) \\ \Phi_k &= \sum_{k'} i_{k'} \frac{k' - 1}{k'} P(k'|k)\end{aligned}$$

La funzione Φ_k tiene conto esplicitamente della probabilità condizionata con cui un vertice infettato di grado k' punta a un vertice di grado k , tramite ciascuno dei suoi $k' - 1$ rami liberi (escludendo il caso in cui punti alla sorgente della sua infezione). Se il network presenta correlazioni, misurate dal termine $P(k'|k)$, la situazione è leggermente più complessa. L'equazione di evoluzione per $i_k(t)$ può essere scritta come

$$\frac{di_k(t)}{dt} = \sum_{k'} \beta k \frac{k' - 1}{k'} P(k'|k) i_{k'}(t) \equiv \sum_{k'} C_{k,k'} i_{k'}(t)$$

che è un sistema di equazioni differenziali dato dalla matrice $C \equiv \{C_{k,k'}\}$ di elementi

$$C_{k,k'} = \beta k \frac{k' - 1}{k'} P(k'|k)$$

Il comportamento di $i_k(t)$ è dato da una combinazione lineare di funzioni esponenziali della forma $\exp(\Lambda_i t)$, dove Λ_i sono gli autovalori della matrice C . Ne consegue che il comportamento dell'infezione avrà andamento del tipo:

$$i(t) \sim e^{\Lambda_m t}$$

dove Λ_m è l'autovalore più grande della matrice C . L'andamento del tempo che governa l'aumentare del valore di diffusione è dato da $\tau \sim \frac{1}{\Lambda_m}$.

Nel caso di network non correlati, la matrice C , i cui elementi sono $C_{k,k'}^{nc} = \beta k (k' - 1) \frac{P(k')}{\langle k \rangle}$, ha un unico autovalore soddisfacente

$$\sum_{k'} C_{k,k'} \Psi_{k'} = \Lambda_m^{nc} \Psi_k$$

dove $\Lambda_m^{nc} = \beta \frac{\langle k^2 \rangle}{\langle k \rangle - 1}$, e dove il corrispondente autovettore è $\Psi_k = k$. Nel caso di networks correlati, si può vedere che l'autovalore più grande di C è inferiormente limitato

$$\Lambda_m^2 \geq \min_k \sum_{k'} \sum_l (k' - 1)(l - 1) P(l|k) P(k'|l)$$

Questa equazione risulta molto interessante se viene scritta come

$$\Lambda_m^2 \geq \min_k \sum_l (l - 1) P(l|k) (k_{nn}(l) - 1)$$

Nel caso di scale-networks con distribuzione $P(k) \sim k^{-\gamma}$ con $2 \leq \gamma \leq 3$, il valore medio del grado dei vicini più prossimi $k_{nn}(l)$ diverge per $N \rightarrow \infty$, il che implica che anche Λ_m diverge. La scala dei tempi $\tau \sim \frac{1}{\Lambda_m}$ tende a zero per un qualunque scale-free network con $2 < \gamma \leq 3$. Questo fatto sottolinea anche la rilevanza della quantità k_{nn} , che rappresenta l'estremo inferiore di Λ_m in networks finiti. Questi stessi risultati possono chiaramente essere generalizzati sia per il modello SIR che per il modello SIS.

2.10 Diffusione epidemica negli scale-free networks: assenza del punto di equilibrio

In molti casi reali, come per Internet e il mondo del Web, la dinamica dei rapporti sono assimilabili a una struttura di tipo scale-free.

Per questi tipi di networks, la probabilità che uno dei loro nodi abbia k connessioni segue una distribuzione di tipo $P(k) \sim k^{-\gamma}$ con l'esponente γ che varia fra 2 e 3. La presenza di nodi con un ampio numero di connessioni è un elemento chiave nella determinazione di un modello per questi networks, e si può vedere che per network non strutturati con *assortative* o *dissortative mixing* è una condizione sufficiente per annullare il punto di equilibrio.

Ricordiamo che in un network, quando i nodi con un alto valore di connettività si relazionano più facilmente a nodi, anch'essi altamente connessi, si parla di *assortative mixing*. Quando invece, nodi altamente connessi di grado elevato si relazionano a nodi con bassa connettività, si parla di *dissortative mixing*.

In questo paragrafo, siamo interessati a studiare nel dettaglio il caso di network indiretti in cui l'unica proprietà rilevante dei nodi sia la connettività, e con distribuzione $P(k) \sim Ck^{-\gamma}$, con $2 < \gamma \leq 3$. In questo caso, il network presenta fluttuazioni di connettività non limitate, segnalate dal fatto che il secondo momento diverge, $\langle k^2 \rangle \rightarrow \infty$, per $k_c \rightarrow \infty$.

Questa caratteristica degli scale-free networks li rende terreno fertile per la diffusione di epidemie a prescindere da qualunque tasso di propagazione che gli agenti patogeni possiedono, cambiando così radicalmente molte delle conclusioni raggiunte dallo studio del modello classico.

Consideriamo in questa trattazione il caso del modello SIS, ponendo la condizione che gli individui (o nodi del network) assumano unicamente due possibili stati: sani o malati. Al susseguirsi degli intervalli di tempo, i suscettibili sono infettati in modo proporzionale a un tasso λ se connessi con uno o più nodi infetti. Allo stesso tempo gli infettivi guariscono e diventano nuovamente suscettibili con probabilità 1.

Nei network omogenei e regolari in cui ogni nodo ha più o meno lo stesso numero di connessioni, $k \simeq \langle k \rangle$, è possibile capire il comportamento del modello guardando la densità media $\rho(t)$ degli individui infettati. In tali circostanze, se il valore di λ è al di sopra del punto di equilibrio λ_c , $\lambda \geq \lambda_c$, l'infezione si propaga e raggiunge uno stato endemico con densità finita ρ . Se $\lambda < \lambda_c$, l'infezione tende a scomparire ($\rho = 0$) con una velocità di decadimento esponenziale, in un tempo finito.

Poichè negli scale-free networks le fluttuazioni della connettività sono non limitate ($\langle k^2 \rangle = \infty$), l'approssimazione $k \simeq \langle k \rangle$, è del tutto inadeguata. Al fine di tenere in considerazione queste fluttuazioni è appropriato considerare la densità $\rho_k(t)$ dei nodi infettati all'interno di ogni classe di grado k . L'equazione che descrive il sistema può essere scritta come:

$$\frac{\partial \rho_k(t)}{\partial t} = -\rho_k(t) + \lambda k [1 - \rho_k(t)] \Phi.$$

Il primo termine a destra rappresenta l'estinzione degli individui infetti per guarigione, con tasso unitario. L'equazione di creazione è proporzionale alla densità degli individui suscettibili, $1 - \rho_k$, al tasso di propagazione λ , al numero di nodi vicini k e alla probabilità che ogni ramo punti a un nodo vicino infettato. Quest'ultimo termine è espresso nella formula da Φ , che può essere esplicitato come $\sum_{k'} P(k'|k) \rho_{k'}(t)$.

Ricordiamo che per network non correlati, la probabilità condizionata $P(k'|k)$ indica la probabilità che un ramo uscente da un nodo con grado k punti a un nodo di grado k' . In tali network inoltre ogni ramo punta, con probabilità proporzionale a $k'P(k')$, a un nodo di grado k' . Imponendo la condizione di stazionarietà $\frac{\partial \rho_k(t)}{\partial t} = 0$, $\sum_{k'} P(k'|k) \rho_{k'}(t)$ assume un valore costante indipendente da k e t e il sistema può essere risolto ottenendo le soluzioni

$$\rho_k = \frac{k \lambda \Phi}{1 + k \lambda \Phi}$$

le quali denotano che, tanto più alto è il valore di connettività di un nodo, tanto maggiore è la probabilità di essere infettato. Inoltre da queste si deriva che il valore medio della densità dei nodi infettati nel sistema è dato da $\rho = \sum_k k P(k) \rho_k$. Per questo sistema, il valore di equilibrio è dato da $\lambda_c = \frac{\langle k \rangle}{\langle k^2 \rangle}$.

Per network SF infiniti con $\gamma \leq 3$ vale $\langle k^2 \rangle = \infty$ e conseguentemente $\lambda_c = 0$; questo significa che per network non correlati si ottiene una distribuzione di infettivi per qualunque tasso di infezione λ .

Nel caso di network correlati non è possibile in generale dare un valore esplicito della soluzione del sistema. E' comunque dimostrabile che il punto di equilibrio è dato da $\lambda_c = \frac{1}{\Lambda_m}$, dove Λ_m è l'autovalore della matrice C , definita da

$C_{kk'} = kP(k'|k)$. L'assenza di un punto di equilibrio corrisponde alla divergenza dell'autovalore più grande della matrice C, per network di dimensione infinita ($N \rightarrow \infty$).

Da quanto visto nel paragrafo 2.9.3, è possibile trovare un estremo inferiore e superiore di Λ_m . In particolare, possiamo scrivere, come già visto,

$$\Lambda_m^2 \geq \min_k \sum_{k'} \sum_l k' l P(l|k) P(k'|l).$$

Questa disuguaglianza ci consente di trovare una condizione sufficiente per l'assenza del punto di equilibrio.

Come si può vedere dall'Appendice B, vale questa relazione

$$kP(k'|k)P(k) = k'P(k|k')P(k'),$$

e moltiplicando per un fattore k entrambi i termini e sommando rispetto a k e k', otteniamo

$$\langle k^2 \rangle = \sum_{k'} k' P(k') \sum_k k P(k|k'),$$

dove abbiamo utilizzato la condizione $\sum_k P(k) = \sum_{k'} P(k'|k) = 1$.

Il termine $\bar{k}_{nn}(k', k_c) = \sum_k k P(k|k')$ definisce il grado medio di connettività dei vicini di nodi di grado k' .

In scale-free networks con $2 < \gamma < 3$, il secondo momento della distribuzione del grado diverge; in altre parole, la funzione $\bar{k}_{nn}(k', k_c) \rightarrow \infty$ per $k_c \rightarrow \infty$.

Mettendo in evidenza la funzione di questo nuovo termine, possiamo scrivere che

$$\Lambda_m^2 \geq \min_k \sum_l l P(l|k) \bar{k}_{nn}(l, k_c).$$

Il termine alla destra di questa espressione è una somma di termini positivi e diverge con k_c almeno quanto $\bar{k}_{nn}(l, k_c)$, sia nel caso *assortative* che in quello *dissortative*. Questo implica immediatamente che $\Lambda_m \leq \infty$ per tutti i network con $\langle k^2 \rangle$ divergente.

Infine, l'equazione $\lambda_c = \frac{1}{\Lambda_m}$ determina che il punto di equilibrio si annulla andando al limite, in tutti gli SF networks, se la distribuzione del grado ha un momento secondo divergente; per esempio, una distribuzione di connettività scale-free con esponente $2 < \gamma \leq 3$ è una condizione sufficiente per l'assenza di un punto di equilibrio in network non strutturati.

Definendo in questo modo

$$\langle \bar{k}_{nn} \rangle_N = \sum_k P(k) \bar{k}_{nn}(k, k_c),$$

abbiamo considerato esplicitamente k_c come una funzione crescente rispetto alla dimensione del network N. Per quanto visto, segue che $\langle \bar{k}_{nn} \rangle_N \rightarrow \infty$ quando

$N \rightarrow \infty$.

Negli SF networks, questo parametro prende in considerazione il livello di fluttuazione del grado e appare come responsabile delle dinamiche di diffusione dell'epidemia. In un certo qual modo il numero di vicini che possono essere infettati negli intervalli successivi è una quantità estremamente rilevante: il comportamento degli SF networks, che segue la legge di potenza, impone una divergenza di $\langle \bar{k}_{nn} \rangle_N$ per $N \rightarrow \infty$, indipendentemente dal livello e dalla struttura delle correlazioni presenti nel network. Questo porta a una decrescita a zero del punto di equilibrio: le infezioni possono proliferare a prescindere dal tasso di diffusione che esse possiedono.

Appendice A

Appendice

A.1 Sistemi lineari omogenei

Sia A un operatore di \mathfrak{R}^n . Siamo interessati ad esprimere le soluzioni dell'equazione $x' = Ax$ nei termini di operatori esponenziali.

Teorema A.1.1 (Teorema).

Sia A un operatore in \mathfrak{R}^n . Allora la soluzione a $x' = Ax$, di condizione iniziale $x(0)=K, K \in \mathfrak{R}^n$, è $e^{tA}K$ e non vi sono altre soluzioni.

Anche senza calcolare le soluzioni esplicite del sistema di equazioni differenziali, possiamo comunque ottenere importanti informazioni qualitative sulle soluzioni, dallo studio degli autovalori di A .

Si ha infatti che se $\det A \neq 0$, allora $x' = Ax$ ammette come unico punto di equilibrio l'origine, e la sua soluzione generale è data da $x(t) = Ke^{\lambda t} = (K_1e^{t\lambda}, K_2e^{t\lambda})$,
 $\lambda Ke^{t\lambda} = AKe^{\lambda t}$,
 $\lambda K = AK \implies (A - \lambda I)K = 0$.

λ è un autovalore della matrice A e K è il corrispondente autovettore. Possiamo distinguere vari casi, di cui considereremo i più importanti.

Caso I. Gli autovalori della matrice A hanno entrambi parti reali negative. In questo caso l'origine si dice nodo stabile, in quanto

$$\lim_{t \rightarrow +\infty} x(t) = 0$$

e l'origine è un punto asintoticamente stabile. Gli autovettori K_1 e K_2 corrispondenti ai due autovalori distinti, sono dunque linearmente indipendenti, e la soluzione generale del sistema la si può scrivere come loro combinazione lineare:

$$x(t) = c_1 K_1 e^{(\lambda_1 t)} + c_2 K_2 e^{(\lambda_2 t)}$$

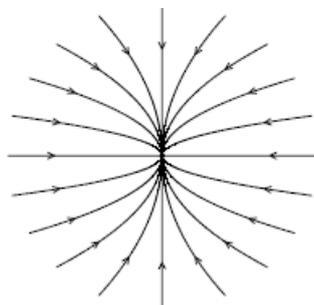


Figura A.1: Nodo stabile

Caso II. Gli autovalori della matrice A sono reali positivi \implies l'origine è un punto instabile repulsivo, in quanto la soluzione $x(t)$ nel tempo si allontana dall'origine, la quale viene detta, per questo motivo, nodo instabile.

$$\lim_{t \rightarrow +\infty} |x(t)| = \infty$$

$$\lim_{t \rightarrow -\infty} |x(t)| = 0$$

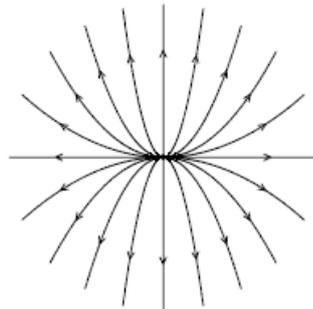


Figura A.2: Nodo instabile

Caso III. A ha autovalori reali di segno opposto. In questo caso l'origine è chiamata punto di *sella*, e risulterà essere di tipo instabile.

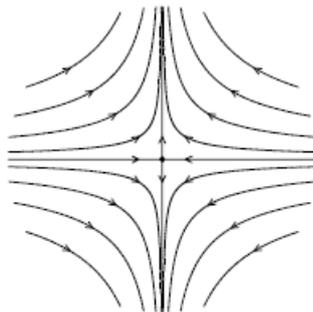


Figura A.3: Punto sella

Caso IV. Gli autovalori sono complessi coniugati del tipo $\lambda_1 = a + ib$ e $\lambda_2 = a - ib$; dunque la soluzione di $x' = Ax$ sta nel piano complesso. Esistono due autovettori complessi coniugati K_1 e K_2 , tali per cui la soluzione generale può essere espressa come

$$x(t) = e^{at}(K_1 \cos bt - K_2 \sin bt, K_2 \cos bt + K_1 \sin bt).$$

Ne segue che:

1. Se $a < 0$ e $|\cos bt| \leq 1$ e $|\sin bt| \leq 1$,
 $\lim_{t \rightarrow \infty} x(t) = 0$, e l'origine è un punto asintoticamente stabile, detto *fuoco stabile* (o *spirale*). A seconda che $b < 0$ oppure $b > 0$, la spirale viene percorsa in senso orario o antiorario rispettivamente.
2. Se $a > 0$, l'origine è instabile repulsiva, detta *fuoco instabile*

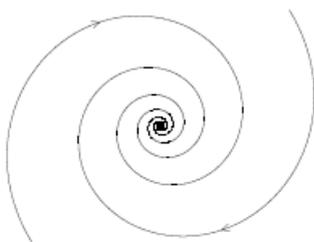


Figura A.4: Fuoco (spirale) stabilea

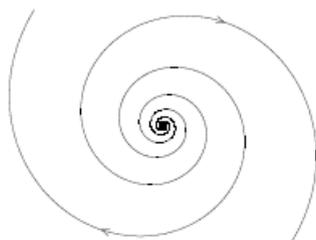


Figura A.5: Fuoco (spirale) instabile

3. Se $a = 0$, gli autovalori ottenuti sono immaginari puri, $e^{at} = 1$, e l'origine viene detta *centro*. È un punto stabile, ma non asintoticamente stabile, in quanto è caratterizzato dalla proprietà che tutte le soluzioni $x(t)$ sono periodiche, di uguale periodo $\frac{2\pi}{b}$.

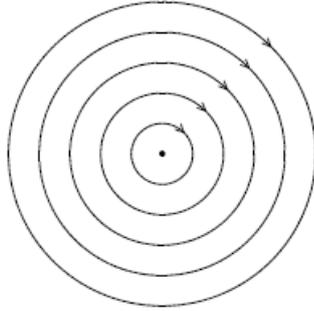


Figura A.6: Centro

Caso V. Gli autovalori sono uguali e negativi ($\lambda_1 = \lambda_2 = \lambda < 0$)

1. A λ corrispondono due autovettori linearmente indipendenti K_1 e K_2 . L'origine è un punto asintoticamente stabile detto *nodo stella*, e le soluzioni possono essere scritte come nel *Caso I*.

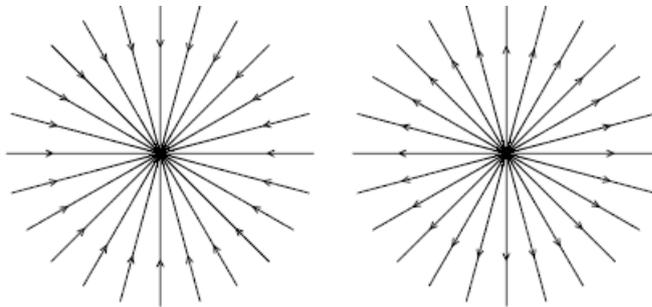


Figura A.7: (a sinistra) Nodo stella attrattivo; (a destra) Nodo stella repulsivo

2. A λ corrisponde un solo autovettore K . Posso quindi trovare un altro autovettore U , che sia linearmente indipendente con K , e la soluzione generale sarà data da $x(t) = c_1 K e^{\lambda t} + c_2 (U + \mu K) e^{\lambda t}$. L'origine è un *nodo stabile di seconda specie*.

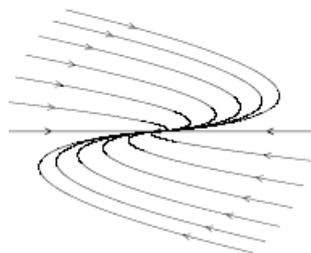


Figura A.8: Nodo stabile di seconda specie

*Caso VI. Gli autovalori sono uguali ma positivi ($\lambda_1 = \lambda_2 = \lambda > 0$) L'origine è di tipo repulsivo, detto *nodo di 2° specie instabile*.*

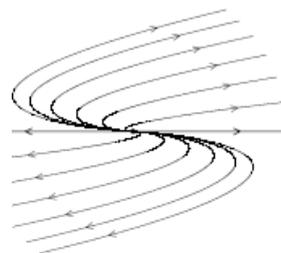


Figura A.9: Nodo instabile di seconda specie

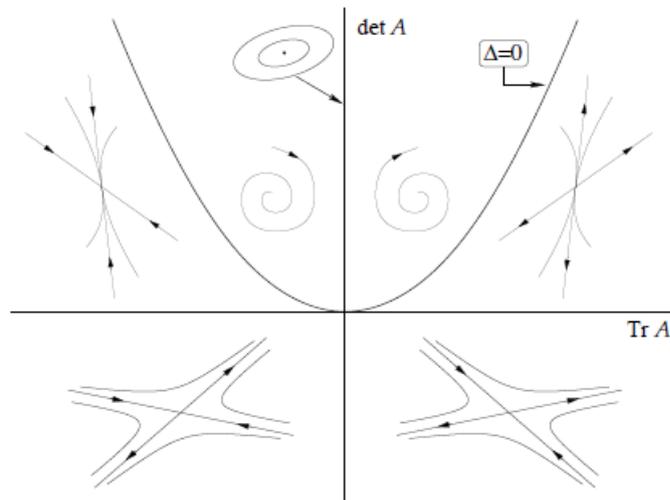


Figura A.10: Diagramma che riassume e mostra come trA e $detA$, determinano il tipo di singolarità nel piano delle fasi. Quando la traccia è negativa e il determinante è positivo, abbiamo il nodo stabile al di sotto della parabola, e una spirale stabile al di sopra. Quando la traccia è nulla e il determinante è positivo, si ha il centro. Se il determinante e la traccia sono entrambi positivi, troviamo la spirale instabile e il nodo instabile, rispettivamente al di sopra e al di sotto della parabola. I punti di sella si trovano invece nel semipiano in cui il determinante è negativo.

La figura riassume le informazioni geometriche delle orbite del sistema $x' = Ax$ che possono essere desunte dal polinomio caratteristico di A . Scriviamo questo polinomio come

$$\lambda^2 - (TrA)\lambda + DetA$$

Il discriminante Δ è definito come

$$\Delta = (TrA)^2 - 4DetA.$$

Gli autovalori sono

$$\frac{1}{2}(TrA \pm \sqrt{\Delta})$$

Quindi gli autovalori reali corrisponderanno al caso di $\Delta \leq 0$; gli autovalori con parte reale negativa li possiamo ottenere quando $TrA < 0$; e così via si deducono tutti gli altri casi.

Appendice B

Appendice

B.1 Grafi random

Un network è detto random quando la probabilità dell'esistenza di un ramo fra due nodi è completamente indipendente dalle caratteristiche dei nodi. In altre parole, l'unica funzione rilevante è la distribuzione di grado $P(k)$.

La prima relazione che queste quantità devono soddisfare è che

$$\sum_k P(k) = 1.$$

Una seconda condizione è imposta dal fatto che tutti i rami devono puntare da un vertice a un altro. Questo fa sì che il numero totale di rami che puntano da vertici di grado k a vertici di grado k' , sia uguale al numero di rami che puntano da vertici di grado k' a vertici di grado k . Denotiamo con N_k il numero di vertici di grado k . Poiché $\sum_k N_k = N$, possiamo definire la distribuzione del grado come

$$P(k) = \lim_{k \rightarrow \infty} \frac{N_k}{N}.$$

Per definire completamente un network, a parte il numero di vertici di un dato grado, occorre specificare come questi vertici sono connessi fra loro, tramite una matrice simmetrica $E_{kk'}$ del numero totale di rami fra vertici di grado k e vertici di grado k' , per $k \neq k'$. I valori diagonali E_{kk} sono uguali a due volte il numero di connessioni fra vertici della stessa classe, $k = k'$ da cui si ottengono le seguenti identità:

$$\sum_{k'} E_{kk'} = kN_k,$$

$$\sum_{k,k'} E_{kk'} = \langle k \rangle N = 2E,$$

dove $\langle k \rangle$ è il grado medio ed E il numero totale di rami nel network. Difatti, la prima relazione ci dice che il numero di rami uscenti da tutti i vertici di grado k è semplicemente kN_k , mentre il secondo indica che la somma di tutti i gradi dei vertici è uguale a due volte il numero di rami. La prima identità ci consente di scrivere la probabilità condizionata $P(k'|k)$ definita come la probabilità che un ramo da un vertice k punti a un vertice k' , come

$$P(k'|k) = \frac{E_{kk'}}{kN_k}.$$

La seconda identità consente invece di dare una definizione della distribuzione congiunta come

$$P(k, k') = \frac{E_{kk'}}{\langle k \rangle N},$$

rappresentante la probabilità che un ramo scelto casualmente connetta due vertici di gradi k e k' . Da questo segue che

$$P(k'|k) = \frac{E_{kk'}}{kN_k} \equiv \frac{\langle k \rangle P(k, k')}{kP(k)},$$

da cui segue immediatamente, come conseguenza della simmetria di $P(k, k')$ che

$$kP(k'|k)P(k) = k'P(k|k')P(k').$$

Per grafi non correlati in cui $P(k'|k)$ non dipende da k , l'imposizione della condizione $\sum_k P(k|k') = 1$, all'interno di quest'ultima equazione, porta alla formula

$$P(k'|k) = \frac{1}{\langle k \rangle} k' P(k').$$

Bibliografia

- [1] A. Barrat, M. Barthélemy, A. Vespignani, Dynamical processes on complex networks, Cambridge university Press, Cambridge, UK, 2008.
- [2] Herbert W. Hethcote, The Mathematics of Infectious Diseases, SIAM Review, Vol. 42, No. 4, pp.599-653, 2000.
- [3] Herbert W. Hethcote, Qualitative Analyses of Communicable Disease Models, Mathematical Biosciences, Vol. 28, 335-356, 1976.
- [4] R. Pastor-Satorras, A. Vespignani, Epidemic Spreading in Scale-Free Networks, Physical Review Letters 86, No. 14, 2001.
- [5] Matt J. Keeling, Ken T. D. Eames, Networks and epidemic models, J. R. Soc. Interface 2, 295-307, 2005.
- [6] V. Capasso, Mathematical Structures of Epidemic Systems, Lecture Notes in Biomathematics, Springer-Verlag, Berlin Heidelberg, 2008.
- [7] J. D. Murray, Mathematical Biology, Springer-Verlag, 2002.
- [8] M.Boguna, R.Pastor-Satorras, A.Vespignani, Absence of epidemic threshold in scale-free networks with connectivity correlations, 2008.

