



UNIVERSITÀ
DEGLI STUDI
FIRENZE

Scuola di Scienze Matematiche, Fisiche e Naturali

Corso di Laurea in Fisica e Astrofisica

Tesi di Laurea

TITOLO ITALIANO

ENGLISH TITLE

GIULIA LEONE

Relatore: Prof. *Franco Bagnoli*

Anno Accademico 2019-2020

*A <Nome>,
frase di dedica.*

"Citazione coltissima" Autore, "Titolo", 19xx, rif bibliografia.

INDICE

Introduzione	1
1 TITOLO DEL PRIMO CAPITOLO	2
2 TITOLO DEL SECONDO CAPITOLO	7
3 TITOLO DEL TERZO CAPITOLO	13
4 CONCLUSIONI	14
A CODICE SVILUPPATO: COMPLETO E COMMENTATO	15
Acronimi	16
Bibliografia	17

ELENCO DELLE FIGURE

Figura 1	dipendenza temporale delle dimensioni relative dei compartimenti rispetto alla totalità della popo- lazione; $\gamma = 1$ e $\beta = 1.6$. [1]	6
Figura 2	(a) Un grafo semplice, cioè privo di loop o link multipli. (b) Un grafo che presenta entrambi. [2] .	8
Figura 3	Un digrafo. [2]	9
Figura 4	[2]	10

ELENCO DEI CODICI

ELENCO DELLE TABELLE

*"Citazione colta,
ma colta colta,
in italiano."*

*"Citazione colta,
ma colta colta,
in inglese."
— Opera, Autore*

INTRODUZIONE

PROVA DI CITAZIONE

[3]

TITOLO DEL PRIMO CAPITOLO

L'obiettivo dietro alla creazione di un modello matematico di una malattia infettiva è quello di arrivare a comprenderne e a descriverne il processo di trasmissione. In linea del tutto generale, possiamo andare a semplificarlo come segue:

- a. in primo luogo, uno o più soggetti infetti vengono introdotti in una popolazione di individui suscettibili (a rischio, cioè, di contrarre la malattia);
- b. un individuo che viene infettato può inizialmente rimanere asintomatico, per poi mostrare i sintomi; può guarire, sia grazie all'assunzione di medicinali che all'azione del sistema immunitario, ed acquisire così una protezione nei confronti di una possibile reinfezione;
- c. quando il bacino dei potenziali suscettibili viene sufficientemente svuotato, la diffusione inizia a rallentare fino a fermarsi; se vengono aggiunti nuovi soggetti alla popolazione, che sia a seguito di flussi migratori o di nascite, l'epidemia può persistere per un lungo periodo di tempo e diventare così endemica.

La modellazione matematica si è rivelata di centrale importanza nel saper rispondere alle domande che possono sorgere all'alba di quella che, a tutti gli effetti, potrebbe rivelarsi una nuova epidemia (se non una pandemia), come, ad esempio, quale possa essere il numero di persone bisognose di cure ospedaliere o quali effetti possa sortire l'imposizione di una quarantena. La sua forza sta anche nel fatto che gli approcci tradizionali, quello statistico e quello sperimentale, in questo frangente non si rivelano altrettanto utili: se da una parte diventa complicato riprodurre in laboratorio il comportamento su grande scala di una malattia infettiva,

che può coinvolgere un gran numero di persone distribuite in aree geografiche spazialmente estese, dall'altra è difficile fare affidamento su di un'analisi statistica se i dati raccolti non sono completi o accurati (basti pensare alla difficoltà di reperire informazioni su soggetti asintomatici). Lo scopo della modellazione si fa, dunque, triplice [4]:

1. come già detto, consentire una migliore comprensione dei meccanismi di trasmissione dell'infezione;
2. riuscire, di conseguenza, a predirne l'andamento futuro;
3. infine, individuare delle modalità di contenimento per tenere sotto controllo la diffusione.

Il processo che si mette in atto consiste di una serie di passi, che vanno dalla formulazione di assunzioni sulla trasmissione della malattia, a partire dalle quali si può costruire un primo modello, alla validazione dello stesso mediante i dati raccolti. C'è, tuttavia, da tenere conto del fatto che un modello matematico non è che un'approssimazione, basata sulle ipotesi che facciamo a seguito di quanto siamo riusciti ad osservare; ciò si traduce nella necessità, imposta anche da una non adeguata conoscenza della malattia in questione, di fare ricorso a delle semplificazioni: risulta, pertanto, chiara l'esigenza di andare a confermare coi dati, qualora possibile, il modello che si è messo in piedi.

Ci sono generalmente tre approcci che si possono seguire:

- quello dei modelli statistici, ampiamente usati in epidemiologia, ma con lo svantaggio, come abbiamo già sottolineato, di necessitare di grandi campioni di dati;
- quello dei modelli deterministici, retti dall'assunzione secondo la quale la dimensione delle popolazioni dei suscettibili e degli infetti sia una funzione continua del tempo; risultano meno affidabili se queste ultime constano di pochi individui, tuttavia sono matematicamente maturi e meno dipendenti dai dati;
- infine, quello dei modelli stocastici, che ben si adattano ad essere impiegati nel caso in cui si abbia a che fare con gruppi ristretti, ma che, al contempo, hanno bisogno di un gran numero di simulazioni numeriche.

[5]

In primo luogo, si va a suddividere la popolazione sotto esame in gruppi

mutualmente esclusivi - come se fossero, per l'appunto, *compartimenti* stagni - così che questi possano riflettere caratteristiche osservabili del processo di infezione. Come evidenziato in [1], quello che ci si propone di fare è descrivere nel modo più formale possibile le transizioni da una classe all'altra, così che si possa tener traccia del numero di individui che vi appartengono; detto risultato può essere raggiunto in modo diverso a seconda di quale tra i già citati approcci si sta seguendo: nella fattispecie, quello deterministico, che richiede la scrittura di un sistema di equazioni differenziali ordinarie ove le variabili in gioco sono le dimensioni dei vari compartimenti, consente di conoscere in modo esaustivo il comportamento di una popolazione a partire dalle sue condizioni iniziali, mentre quello stocastico si basa sull'esistenza implicita di una rete di connettività fra gli individui e, pertanto, tratta le informazioni relative all'infezione in funzione del rate di spostamento fra gli elementi di un reticolo.

Sulla falsariga del modello più noto e forse storicamente più rilevante, quello delineato da Kermack e McKendrick nel 1927 [6], possiamo ora introdurre il modello SIR: esso comporta la ripartizione del campione in esame in tre gruppi, ovvero

- S , i *suscettibili*
- I , gli *infetti*
- R , interpretabile sia come *guariti* (recovered) che come *rimossi* (removed)

la cui evoluzione temporale - messa in evidenza nella figura 1 è ben descritta dalle seguenti

$$\begin{cases} \frac{dS(t)}{dt} = -\beta I(t) \frac{S(t)}{N} \\ \frac{dI(t)}{dt} = \beta I(t) \frac{S(t)}{N} - \gamma I(t) \\ \frac{dR(t)}{dt} = \gamma I(t) \end{cases} \quad (1)$$

laddove β e γ , entrambe costanti > 0 , sono rispettivamente il rate d'infezione e di guarigione/rimozione, mentre $S(t) + I(t) + R(t) = N$ fornisce la dimensione della popolazione, anch'essa una quantità costante. Al sistema soprastante vengono aggiunte anche le opportune condizioni iniziali: $S(0) = S_0 > 0$, $I(0) = I_0 > 0$, $R(0) = 0$.

Osserviamo che il modello non perde di senso fintanto che $S(t)$ e $I(t)$ rimangono non negativi; è, pertanto, sufficiente che anche solo uno dei due diventi pari a zero perché ciò accada. Notiamo, inoltre, che $\frac{dS(t)}{dt} < 0$

per ogni possibile valore di t , mentre $\frac{dI(t)}{dt} > 0$ se e solo se $\frac{\beta S(t)}{\gamma} > 1$: ciò significa che, fin quando quest'ultima disuguaglianza rimane vera, $I(t)$ cresce, ma in seguito, in virtù del fatto che $S(t)$ diminuisce a prescindere da t , inverte il suo comportamento e finisce per tendere a zero.

Quanto appena detto ci dà il via per andare a definire una quantità fondamentale nello studio della diffusione di una malattia infettiva, il *numero di riproduzione di base* \mathcal{R}_0

$$\mathcal{R}_0 = \frac{\beta N}{\gamma} \quad (2)$$

che conta quanti individui, in media, vengono contagiati da un singolo infetto in una popolazione che altrimenti sarebbe totalmente sana. In aggiunta, è un buon indicatore dell'eventuale esplodere o meno dell'epidemia; più precisamente

$$\text{se } \mathcal{R}_0 \begin{cases} < 1, & \text{allora } I < 0 \text{ e il processo d'infezione si arresta} \\ > 1, & \text{allora } I \text{ cresce e l'epidemia prende piede} \end{cases}$$

La transizione fra uno scenario e l'altro si ha per $\mathcal{R}_0 = 1$, punto che prende il nome di *soglia epidemica*.

Combinando le prime due equazioni [7] che compaiono in (1), si ottiene

$$\begin{aligned} \log \frac{S_0}{S_\infty} &= \frac{\beta}{\gamma} [N - S_\infty] \\ &= \mathcal{R}_0 \left[1 - \frac{S_\infty}{N} \right] \end{aligned}$$

che mette in luce la relazione fra il numero di riproduzione di base e la dimensione finale dell'epidemia, $[N - S_\infty]$.¹

¹ Qui abbiamo sottinteso di star cercando delle soluzioni costanti della coppia di equazioni differenziali, (S_∞, I_∞) ; partendo dal presupposto che, a $t = 0$, $S_0 + I_0 = N$ e ricordando che $\lim_{t \rightarrow \infty} I(t) = I_\infty = 0$, risulta che $\lim_{t \rightarrow \infty} S(t) + I(t) = S_\infty$.

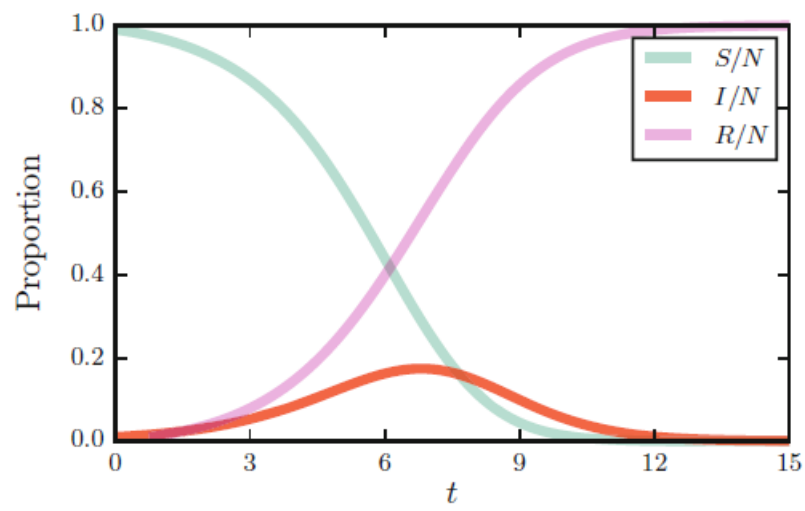


Figura 1: dipendenza temporale delle dimensioni relative dei compartimenti rispetto alla totalità della popolazione; $\gamma = 1$ e $\beta = 1.6$. [1]

TITOLO DEL SECONDO CAPITOLO

Le assunzioni secondo cui ogni individuo possa entrare in contatto con chiunque (mixing omogeneo) e il numero di interazioni di ciascun soggetto sia confrontabile con quello degli altri non sono realistiche: anzi, la probabilità che si verifichi un incontro fra due individui presi a caso è praticamente infinitesima. Di norma, ognuno ha una serie di contatti regolari con un numero ristretto di persone (familiari, colleghi, etc), mentre ignora tutto il resto della popolazione; questo li rende particolarmente adatti ad essere rappresentati tramite una rete. Andiamo ad introdurre una serie di definizioni che in seguito ci risulteranno utili.

Definizione 2.1 (*Grafo*). Un *grafo* (o una *rete*) è un insieme di elementi detti vertici o *nodi* che possono essere collegati fra loro da segmenti detti archi o *link*.

Consideriamo una rete non orientata - cioè una rete in cui i link possono essere percorsi indistintamente in un verso e nell'altro - con n vertici, che andiamo ad etichettare da 1 a n . Se indichiamo con (i, j) l'arco fra i nodi i e j , allora l'intera rete può essere descritta in funzione della

Definizione 2.2 (*Matrice di adiacenza I*). La matrice di adiacenza \mathbf{A} relativa ad un grafo semplice è una matrice i cui elementi sono così definiti:

$$A_{ij} = \begin{cases} 1, & \text{se esiste un arco fra } i \text{ e } j \\ 0, & \text{altrimenti} \end{cases}$$

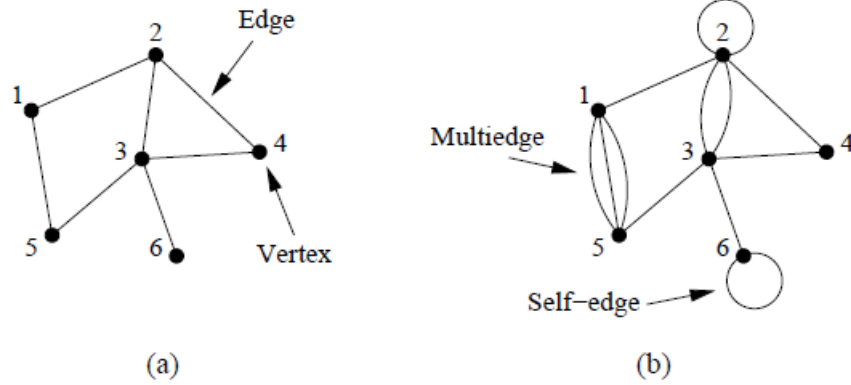


Figura 2: (a) Un grafo semplice, cioè privo di loop o link multipli. (b) Un grafo che presenta entrambi. [2]

Se dunque prendiamo, ad esempio, la rete (a) in figura 2, assumerà la seguente forma:

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix} \quad (3)$$

Possiamo osservare che risulta essere una matrice simmetrica con tutti 0 sulla diagonale.

Qualora invece ne avessimo sotto esame una più simile alla (b) della medesima figura, dovremmo tener conto del fatto che sono presenti link multipli e loop; si stabilisce di assegnare ai primi un numero pari alla loro molteplicità e ai secondi il valore 2, così che

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & 3 & 0 \\ 1 & 2 & 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 3 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 2 \end{pmatrix} \quad (4)$$

matrice che rimane, ad ogni modo, simmetrica.

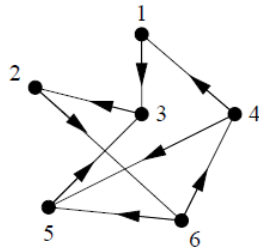


Figura 3: Un digrafo. [2]

La questione cambia leggermente se si va a considerare una rete diretta o *digrafo*.

Definizione 2.3 (*Digrafo*). Un *digrafo* è un tipo di grafo in cui ogni arco ha una direzione, punta cioè da un vertice ad un altro.

Ciò ci porta a rivedere quanto detto per la matrice di adiacenza: affermiamo, quindi, che

Definizione 2.4 (*Matrice di adiacenza II*). La matrice di adiacenza **A** relativa ad un grafo orientato è una matrice i cui elementi sono così definiti:

$$A_{ij} = \begin{cases} 1, & \text{se esiste un arco da } j \text{ a } i \\ 0, & \text{altrimenti} \end{cases}$$

Relativamente alla figura 3, si avrà pertanto

$$A = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (5)$$

Ci soffermiamo poi su un particolare tipo di grafo, che ci risulterà più utile in seguito.

Definizione 2.5 (*Grafo bipartito*). [8] Un grafo si dice *bipartito* se i suoi vertici possono essere suddivisi in due sottoinsiemi X e Y tali che ogni link ha un'estremità in X ed una in Y .

In modo del tutto equivalente a quanto abbiamo già fatto, possiamo andare a definire per un grafo siffatto una matrice che lo va a descrivere.

Definizione 2.6 (*Matrice di incidenza*). La matrice di incidenza **B** è una matrice $g \times n$, dove g è il numero di sottoinsiemi e n quello dei vertici che ne fanno parte, i cui elementi sono definiti come segue:

$$B_{ij} = \begin{cases} 1, & \text{se il vertice } j \text{ appartiene al gruppo } i \\ 0, & \text{altrimenti} \end{cases}$$

Uno dei pregi di una rete bipartita è quello di consentire in modo fluido il passaggio da una rappresentazione in cui compaiono sia vertici che gruppi ad una in cui figurano, ad esempio, soltanto i vertici; questa possibilità permette di visualizzare il formarsi di quelle che prendono il nome di cricche (*cliques*).

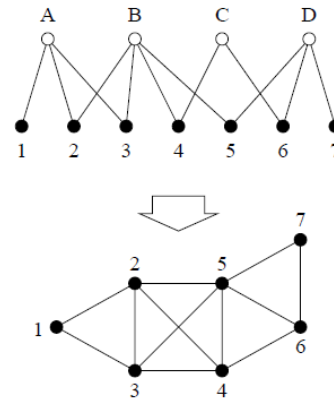


Figura 4: [2]

Definizione 2.7 (Cricca). [9] Una *cricca* è un sottografo completo, cioè un sottografo in cui ogni vertice è connesso a tutti gli altri.

L'emergere di questi raggruppamenti, di conseguenza, mette in rilievo l'esistenza di gruppi fortemente coesi all'interno della rete; questo aspetto assume ulteriore valore, in particolare, quando cerchiamo di rispondere ad una domanda che ci siamo posti nel precedente capitolo, ovvero come poter rappresentare una popolazione nella quale stia prendendo piede un certo tipo di malattia infettiva, perché ci consente non solo di dare forma agli individui che la costituiscono, ma anche all'intensità dei legami che intercorrono fra di essi: ci aspettiamo, infatti, che membri di una famiglia o colleghi di lavoro vadano a costituire una cricca.

Definizione 2.8 (Grado di un vertice). Diciamo *grado* di un vertice il numero di archi ad esso connessi.

Se abbiamo un grafo non orientato con n vertici, possiamo esprimere il grado dell' i -esimo vertice come

$$k_i = \sum_{j=1}^n A_{ij} \quad (6)$$

Osserviamo adesso che, se una rete consta di m link, il numero totale di estremità che possiamo contare - pari a $2m$ - altro non è se non la somma dei gradi di tutti i vertici:

$$m = \frac{1}{2} \sum_{i=1}^n k_i \quad (7)$$

il che ci porta ad esprimere il grado *medio* di un vertice come

$$c = \frac{1}{2} \sum_{i=1}^n k_i = \frac{2m}{n} \quad (8)$$

Indichiamo poi la *densità* di un grafo semplice come la frazione di archi effettivamente presenti:

$$\rho = \frac{m}{\binom{n}{2}} = \frac{2m}{n(n-1)} = \frac{c}{n-1} \quad (9)$$

con $0 \leq \rho \leq 1$ e $\binom{n}{2}$ numero massimo di link che è possibile tracciare. Ci soffermiamo brevemente sulle stesse grandezze nel caso in cui si stia esaminando una rete orientata:

- si può distinguere fra il grado esterno (relativo agli archi uscenti da un vertice)

$$k_i^{\text{out}} = \sum_{j=1}^n A_{ij} \quad (10)$$

e quello interno (relativo agli archi entranti)

$$k_j^{\text{in}} = \sum_{i=1}^n A_{ij} \quad (11)$$

- il numero di link presenti è pari a

$$m = \sum_{i=1}^n k_i^{\text{in}} = \sum_{j=1}^n k_j^{\text{out}} = \sum_{ij} A_{ij} \quad (12)$$

- il grado medio esterno è equivalente a quello interno

$$c^{\text{in}} = \frac{1}{n} \sum_{i=1}^n k_i^{\text{in}} = \frac{1}{n} \sum_{j=1}^n k_j^{\text{out}} = c^{\text{out}} \quad (13)$$

Il concetto di grado ci è immediatamente utile per definirne un altro: quello di *assortatività* o di *mixing assortativo*. Con questo termine indichiamo una tendenza, da parte dei vertici, a legarsi con altri nodi che, in qualche modo, somigliano loro; è piuttosto immediato pensare che un aspetto che possono avere in comune sia proprio il grado e, in effetti, in questo contesto si parla di mixing assortativo per grado: è possibile andarlo a quantificare mediante un coefficiente [2]

$$r = \frac{\sum_{ij} \left(A_{ij} - \frac{k_i k_j}{2m} \right) k_i k_j}{\sum_{ij} \left(k_i \delta_{ij} - \frac{k_i k_j}{2m} \right) k_i k_j} \quad (14)$$

3

TITOLO DEL TERZO CAPITOLO

CONCLUSIONI

A

CODICE SVILUPPATO: COMPLETO E COMMENTATO

ACRONIMI

BIBLIOGRAFIA

- [1] István Z. Kiss, Joel C. Miller e Péter L. Simon: *Mathematics of Epidemics on Network: From Exact to Approximate Models*. Springer, 2017. (Cited on pages v, 4, and 6.)
- [2] Mark E. J. Newman: *Networks: An Introduction*. Oxford University Press, 2010. (Cited on pages v, 8, 9, 10, and 12.)
- [3] Franco Bagnoli, Daniele Lorini e Pietro Lió: *Modeling Social Groups, Policies and Cognitive Behavior in COVID-19 Epidemic Phases. Basic Scenarios*. Substantia, vol. 4(no. 1): 914, Jun. 2020. <https://doi.org/10.13128/Substantia-914>. (Cited on page 1.)
- [4] Daryl J. Daley e Joseph Gani: *Epidemic Modelling: An Introduction*. Cambridge University Press, 2001. (Cited on page 3.)
- [5] Michael Y. Li: *An Introduction to Mathematical Modeling of Infectious Diseases*. Springer, 2018. (Cited on page 3.)
- [6] William O. Kermack e Anderson G. McKendrick: *A contribution to the mathematical theory of epidemics*. Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character, 115(772):700–721, Aug. 1927. <https://doi.org/10.1098/rspa.1927.0118>. (Cited on page 4.)
- [7] Fred Brauer, Carlos Castillo-Chavez e Zhilan Feng: *Mathematical Models in Epidemiology*. Springer, 2019. (Cited on page 5.)
- [8] John A. Bondy e Uppaluri S. R. Murty: *Graph Theory*. Springer, 2008. (Cited on page 9.)
- [9] Allan Bickle: *Fundamentals of Graph Theory*. American Mathematical Society, 2020. (Cited on page 10.)

RINGRAZIAMENTI

Ringraziamento strappalacrime 1.

Ringraziamento strappalacrime 2.