



VNiVERSIDAD D SALAMANCA

CAMPUS DE EXCELENCIA INTERNACIONAL

FACULTAD DE CIENCIAS

GRADO EN MATEMÁTICAS

Complejos Simpliciales y Ciencia de Redes.

Simplicial Complexes and Network Science.

Trabajo de Fin de Grado realizado por:
Beatriz Barbero Lucas

Tutor:
Daniel Hernández Serrano

Julio de 2021



VNiVERSIDAD D SALAMANCA

CAMPUS DE EXCELENCIA INTERNACIONAL

FACULTAD DE CIENCIAS

GRADO EN MATEMÁTICAS

Complejos Simpliciales y Ciencia de Redes.

Simplicial Complexes and Network Science.

Trabajo de Fin de Grado realizado
por:

Tutor:

Fdo: Beatriz Barbero Lucas

Fdo: Daniel Hernández Serrano

Julio de 2021

Índice general

Introducción	1
1. Teoría de Grafos	3
2. Ciencia de Redes	9
2.1. Redes aleatorias	9
2.1.1. Número de aristas	11
2.1.2. Distribución de grado de una red aleatoria	12
2.2. Redes de escala libre	14
2.2.1. El algoritmo Barabási-Albert	15
2.2.2. Variación de los grados en función del tiempo	15
2.2.3. Distribución de grados	16
2.2.4. Propiedades de las redes de escala libre	20
3. Modelos analíticos epidemiológicos	23
3.1. Modelo Susceptible-Infectado (SI)	23
3.2. Modelo Susceptible-Infectado-Susceptible (SIS)	25
3.3. Modelo Susceptible-Infectado-Recuperado (SIR)	26
3.4. Los modelos SI, SIS, SIR considerando el grado de los nodos	27
4. Complejos Simpliciales	29
4.1. Complejos simpliciales	29
4.2. Medidas de adyacencia	30
4.2.1. Para símlices de la misma dimensión	31
4.2.2. Adyacencias generales para símlices de distinta dimensión	33
5. Modelos Simpliciales en Epidemiología	39
5.1. Algoritmo RSC	39
5.1.1. Algoritmo RSC para $m=2$	40
5.2. Modelo de Contagio Simplicial (SCM)	44
5.2.1. SCM de orden D	44
5.2.2. SCM de orden $D = 2$	45

Agradecimientos: A mi tutor, Dani, por haberme ayudado a lo largo del proyecto y por todas sus sugerencias. También me gustaría agradecer a todos aquellos profesores que con su esfuerzo y dedicación, me han motivado durante todo el grado.

A Alberto, por todos los momentos en los que has sacado un ratito para ayudarme o simplemente animarme, sin ti, esto no hubiera sido posible. A mi familia, a mis padres Emi y Julio y a mi hermana Carmen, por todo vuestro apoyo incondicional y, sobre todo, por vuestra paciencia. Y finalmente, a mis compañeros, en especial a Javi, por todos los recuerdos y momentos que hemos compartido a lo largo de estos cuatro años.

Introducción

La Ciencia de Redes o Network Science surgió a finales de los noventa como un área de investigación interdisciplinaria que tiene como objetivo desarrollar enfoques y técnicas tanto teóricas como prácticas con el fin de comprender mejor las redes creadas naturalmente y las creadas artificialmente por el ser humano (véase [4]). La teoría de grafos se presenta como su punto de origen y, más concretamente, el modelo de grafos aleatorios presentado por Erdős & Renyi en 1960 [9]. El estudio de redes tiene una larga tradición tanto en el campo teórico: en teoría de grafos y matemática discreta [7]; como en el aplicado: sociología [1], investigación en comunicaciones [5], biología [15, 19] y en física. Por tanto, existen muchos estilos y enfoques diferentes a la hora de estudiar una red compleja.

Una de las aplicaciones de la ciencia de redes es estudiar la evolución de enfermedades contagiosas en un grupo de individuos, esto se hace por medio de modelos basados en clasificar la población en función de la etapa de la enfermedad en la que se encuentre cada individuo y suponiendo que todos ellos tienen la misma probabilidad de contagiarse. Estos modelos estudian la evolución a través del tiempo de la enfermedad en un grupo con un número fijo de individuos. Como veremos en este trabajo, existen modelos epidemiológicos que asumen que todos los individuos tienen un número similar de conexiones pero también hay otros, que por el contrario, estudian el avance de la enfermedad en la población en función del número de conexiones que tenga cada individuo. Ambos enfoques se pueden encontrar en [4].

Estudiar la propagación de enfermedades con las herramientas que nos da la teoría de grafos tiene la limitación de que sólo podemos considerar relaciones binarias entre individuos, es decir, no podemos considerar los casos en los que varios individuos tienen una relación común. Para solventar esta limitación es posible hacer uso de complejos simpliciales que sí permiten modelar este tipo de situaciones. Con los complejos simpliciales podemos, por ejemplo, dados tres individuos estudiar a la vez las relaciones binarias entre cada pareja de individuos de la red y al mismo tiempo la relación conjunta.

Este proyecto está dividido en 5 partes, empezaremos en el capítulo 1 con algunas definiciones de teoría de grafos que posteriormente vamos a necesitar a lo largo de todo el trabajo.

En el capítulo 2 haremos una pequeña introducción sobre ciencia de redes. En la sección 1 estudiaremos qué son las redes aleatorias y haremos un repaso de sus principales características, como la distribución de grados de estas redes y destacaremos algunas de sus propiedades. También veremos algunos algoritmos de creación de redes aleatorias, en particular, estudiaremos el algoritmo $G(n, p)$. En la sección 2 estudiaremos las redes de escala libre, su distribución de grados, clasificación, características y algoritmos de creación de redes de escala libre como el algoritmo Barabási-Albert.

El capítulo 3 está destinado a introducir algunos modelos analíticos epidemiológicos,

como son los modelos SI, SIS y SIR, siglas asociadas a los estados que pueden tener los individuos de estudio: susceptibles (S), infectados (I) o recuperados (R). Cada una de las tres primeras secciones está dedicada al estudio de uno de los modelos, bajo la hipótesis de que todos los nodos de la red tienen un grado uniforme. En la sección 4 vamos a explicar una modificación de los modelos anteriores que permite hacer un estudio en función del grado de cada nodo de la red.

El capítulo 4, consta de dos partes, en la primera veremos algunas definiciones relativas a complejos simpliciales con algunas demostraciones relativas a su construcción y propiedades y, en la sección 2, algunas medidas de adyacencia y grado en complejos simpliciales. En primer lugar estudiaremos las medidas de adyacencia superior e inferior que se pueden dar entre dos grupos de individuos de la misma dimensión, es decir, si ambos grupos están contenidos en un grupo de una dimensión mayor o contienen un grupo de una dimensión menor respectivamente. Y después consideraremos comunidades de distinta dimensión e introduciremos medidas de adyacencia y grados más generales.

Por último, en el capítulo 5 veremos el modelo de contagio simplicial (SCM), que es un modelo que estudia la evolución de una enfermedad con una construcción parecida a la del modelo SIS pero que se aplica sobre complejos simpliciales. Daremos la definición del modelo en el caso general y en el caso en el caso $D = 2$, es decir, cuando sólo hay relaciones entre tres individuos a la vez. Además, para poder recrear este modelo estudiaremos un algoritmo de creación de complejos simpliciales aleatorios, el llamado algoritmo RSC.

Capítulo 1

Teoría de Grafos

Una red es un conjunto de relaciones entre una serie de elementos. En el mundo real podemos encontrar muchos tipos de redes, por ejemplo, podemos considerar la red que forma Facebook, los agentes serían las personas que usan dicha red social y dos personas estarán relacionadas si son amigos entre sí; otro ejemplo es la red aérea mundial, donde dos ciudades están relacionadas si hay una conexión aérea entre ellas. Estas redes pueden modelizarse matemáticamente mediante un grafo, donde los vértices son los agentes y las aristas son las relaciones entre las parejas de agentes.

En este capítulo recogemos una serie de conceptos básicos sobre teoría de grafos que más adelante utilizaremos para estudiar las redes aleatorias y las redes de escala libre. Definiremos qué es un grafo dirigido y no dirigido, el grado de un vértice y el grado medio de un grafo, explicaremos como expresar el número total de aristas de un grafo en términos del grado de cada uno de sus vértices e introduciremos la distribución de grados de un grafo y el coeficiente de clustering local. Estos conceptos sobre teoría de grafos pueden encontrarse en [8] y en [22].

Definición 1.1. Un *grafo dirigido* o *digrafo*, $G = (V, A, f)$, es un par de conjuntos finitos V y A y una aplicación de conjuntos

$$\begin{aligned} f: A &\rightarrow V \times V \\ a &\mapsto (v_i, v_j). \end{aligned}$$

donde a los elementos $v \in V$ los denominaremos vértices, a los elementos $a \in A$ aristas y la aplicación f asocia a cada arista su vértice de salida v_i y su vértice de entrada v_j .

Con las mismas notaciones, un *grafo no dirigido*, $G = (V, A, f)$, es un par de conjuntos finitos V y A y una aplicación de conjuntos

$$\begin{aligned} f: A &\rightarrow S^2V \\ a &\mapsto (v_i, v_j) \end{aligned}$$

donde $S^2V = \{\text{Parejas no ordenadas de elementos de } V\}$.

Definición 1.2. Diremos que dos nodos v_i, v_j de un grafo $G = (V, A, f)$ son *adyacentes* o son *vecinos* si existe una arista $a \in A$ tal que $f(a) = (v_i, v_j)$.

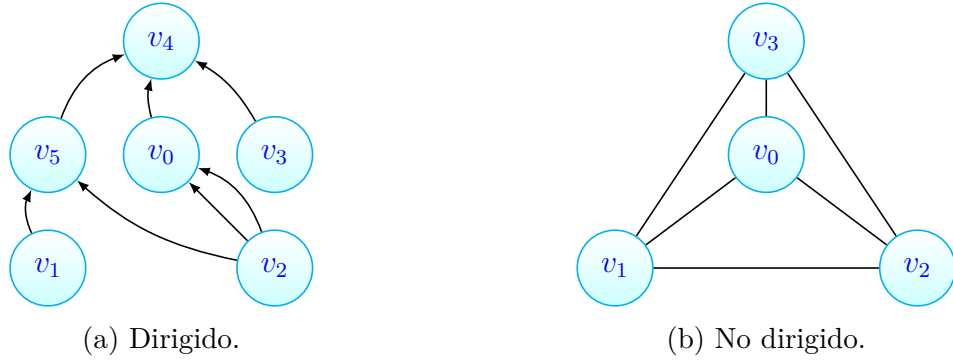


Figura 1.1: Ejemplos de grafos.

Definición 1.3. Se dice que un grafo es *simple* si entre cada par de vértices hay como mucho una arista y no tiene lazos. Se dice que es un *multigrafo* si entre dos vértices existen más de una arista que los une.

Observación 1.4. Equivalentemente, un grafo es simple si la aplicación $A \rightarrow V \times V$ o $A \rightarrow S^2V$ es inyectiva, y es un multigrafo si no es inyectiva.

Si N el número de nodos de un grafo.

Definición 1.5. Un grafo es *completo* si para todo $i, j = 1, \dots, N$ con $i \neq j$ existe $a \in A$ tal que $f(a) = (v_i, v_j)$, es decir, cuando cada uno de sus nodos está unido a todos los demás.

Ejemplo 1.6. La figura 1.1b es un grafo completo. Por otro lado, las figuras 1.1a y 1.1b son un multigrafo y un grafo simple respectivamente.

Definición 1.7. Sea G un grafo dirigido simple. Se llama *grado de entrada* de un vértice v , al número de aristas que llegan a v y se denota por $g_e(v)$.

$$g_e(v) = \#\{a \in A \mid f(a) = (_, v)\}.$$

Se llama *grado de salida* de un vértice v , al número de aristas que parten de v y se denota por $g_s(v)$.

$$g_s(v) = \#\{a \in A \mid f(a) = (v, _)\}.$$

El *grado total* de un nodo v , que denotaremos por $g(v)$, es $g(v) = g_e(v) + g_s(v)$.

Definición 1.8. Sea G un grafo no dirigido. Se llama *grado* de un vértice v , $g(v)$, al número de aristas incidente en v , es decir,

$$g(v) = \#\{a \in A \mid f(a) = (_, v) \text{ ó } f(a) = (v, _)\}.$$

Proposición 1.9. En un grafo no dirigido el número total de aristas, L , puede ser expresado como la suma de los grados de los nodos partido por dos:

$$L = \frac{1}{2} \sum_{i=1}^N g(v_i) \quad (1.1)$$

Demostración. Se obtiene de que tenemos que contar cada arista que incide en cada nodo y tenemos que dividir entre dos porque cada arista es contada dos veces, es decir, tenemos que contar la arista a una sola vez ya que $f(a) = (v_i, v_j) = (v_j, v_i)$. \square

Proposición 1.10. *El número total de aristas de un grafo dirigido es:*

$$L = \sum_{i=1}^N g_e(v_i) = \sum_{i=1}^N g_s(v_i) \quad (1.2)$$

Demostración. Todas las aristas tienen un inicio y un final. Estamos contando el número de aristas de manera distinta, por un lado las agrupamos por el vértice de entrada y, por el otro, por el vértice de salida. \square

Definición 1.11. El *grado medio* de un grafo no dirigido se define como

$$\langle g(G) \rangle = \frac{1}{N} \sum_{i=1}^N g(v_i)$$

Proposición 1.12. *Sea G un grafo no dirigido, se tiene que*

$$\langle g(G) \rangle = \frac{2L}{N}$$

Demostración. Sustituyendo en L la ecuación (1.1), obtenemos $\frac{2L}{N} = \frac{2}{N} \frac{1}{2} \sum_{i=1}^N g(v_i)$, que es precisamente la definición de grado medio. \square

Definición 1.13. El *grado medio de entrada* de un grafo dirigido se define del siguiente modo:

$$\langle g_e(G) \rangle = \frac{1}{N} \sum_{i=1}^N g_e(v_i)$$

Y el *grado medio de salida*:

$$\langle g_s(G) \rangle = \frac{1}{N} \sum_{i=1}^N g_s(v_i)$$

Proposición 1.14. *Sea G un grafo dirigido, se tiene que*

$$\langle g_e(G) \rangle = \frac{L}{N} = \langle g_s(G) \rangle$$

Demostración. Es fácil ver estas igualdades a partir de la ecuación (1.2), ya que si sustituimos en L obtenemos dichas igualdades. \square

Una vez que hemos definido el grado de un nodo es natural también preguntarse por la distribución de grados en un grafo, es decir, por la probabilidad de que un nodo cualquiera tenga un cierto grado k . La distribución de grados nos va a permitir estudiar la cantidad de conexiones de una red y cómo están repartidas las aristas para ver cómo influye en las interacciones entre nodos que queramos estudiar, es decir, nos da una idea de la estructura de la red, de la topología del grafo. Pero veamos primero qué es una distribución:

Sea Ω un conjunto. Llamaremos conjunto de partes de Ω , $\mathcal{P}(\Omega)$, al conjunto formado por todos los subconjuntos de Ω . Diremos que una clase \mathcal{G} de subconjuntos de Ω tiene estructura de σ -álgebra si es un álgebra cerrada bajo la unión numerable.

Con esto, sea Ω un conjunto, sea \mathcal{G} una σ -álgebra de $\mathcal{P}(\Omega)$ y $P : (\Omega, \mathcal{G}) \rightarrow \mathbb{R}$ una medida probabilística. Recordemos que si tenemos un espacio de probabilidad (Ω, \mathcal{G}, P) , llamamos variable aleatoria a toda función medible $X : (\Omega, \mathcal{G}) \rightarrow (\mathbb{R}, \mathbb{B})$ donde \mathbb{B} es la σ -álgebra de Borel de \mathbb{R} .

Definición 1.15. Sea (Ω, \mathcal{G}, P) un espacio de probabilidad y sea X una variable aleatoria. Podemos definir una medida $P_X : (\mathbb{R}, \mathbb{B}) \rightarrow \mathbb{R}$ por

$$\begin{array}{ccc} (\Omega, \mathcal{G}) & \xrightarrow{X} & (\mathbb{R}, \mathbb{B}) \\ & \searrow P & \downarrow P_X \\ & & \mathbb{R} \end{array}$$

donde $P_X(B) = P(X \in B) = P(X^{-1}(B))$ si B es un conjunto medible. Con esta definición P_X es una medida probabilística, a la que llamaremos *probabilidad asociada a la variable aleatoria X* .

En nuestro caso, para definir la distribución de grado, tenemos que $\Omega = \{v_1, \dots, v_n\}$, $\mathcal{G} = \mathcal{P}(\Omega)$, nuestra variable aleatoria es

$$\begin{aligned} X : \{v_1, \dots, v_n\} &\rightarrow \{1, \dots, n\} \\ v_i &\mapsto g(v_i) \end{aligned}$$

y la probabilidad

$$\begin{aligned} P : (\Omega, \mathcal{G}) &\rightarrow \mathbb{R} \\ \{v_{i_1}, \dots, v_{i_j}\} &\mapsto \frac{\#\{v_{i_1}, \dots, v_{i_j}\}}{\#\{v_1, \dots, v_n\}} = \frac{j}{n} \end{aligned}$$

Por tanto, la probabilidad asociada a esta variable aleatoria es

$$P_X(\{k\}) = P(X^{-1}(\{k\})) = \frac{\#\{\text{nodos de grado } k\}}{\#\{\text{nodos del grafo}\}}.$$

Con estas mismas notaciones:

Definición 1.16. La *distribución de grado*, p_k , es la probabilidad de que un nodo cualquiera tenga grado k y se define como $p_k = P_X(\{k\})$.

Con esto, el grado medio de un grafo no dirigido puede expresarse como

$$\langle g(G) \rangle = \sum_{k=0}^{\infty} kp_k.$$

Que es la fórmula de la esperanza matemática: la suma de la probabilidad de cada posible suceso aleatorio multiplicado por el valor de dicho suceso; en nuestro caso, la probabilidad de que un nodo tenga grado k por el grado correspondiente k .

Observación 1.17. En una red de N nodos, el número de aristas puede variar entre $L = 0$ y $L_{max} = \binom{N}{2} = \frac{N(N-1)}{2}$ donde L_{max} es el número total de aristas presentes en un grafo completo de tamaño N , que por ser completo coincide con el número de posibles parejas de nodos en un conjunto de N nodos.

Definamos ahora lo que es un camino y la longitud entre dos nodos.

Definición 1.18. Un *camino* es una sucesión alternada de vértices y aristas, es decir, $v_1 a_1 v_2 a_2 \dots v_n$ donde $f(a_i) = (v_i, v_{i+1})$. Cada camino está compuesto de $n + 1$ nodos y n aristas. La *longitud* de un camino es el número de las aristas que lo componen, es decir, si el camino es $v_1 a_1 v_2 a_2 \dots v_n$ la longitud es n .

Definición 1.19. Un grafo es *conexo* si para cada par de vértices existe un camino entre ellos.

Definición 1.20. Un grafo $G' = (V', A', f')$ se dice que es un *subgrafo* de $G = (V, A, f)$ si $V' \subseteq V$, $A' \subseteq A$ y $f'(a') = f(a')$ para todo $a' \in A'$.

Definición 1.21. Sea G un grafo. Una *componente conexa* de G es un subgrafo conexo que no está estrictamente contenido en ningún otro subgrafo conexo de G .

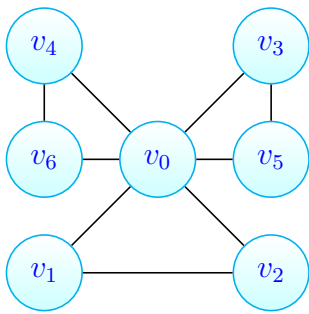
Definición 1.22. Una arista de un grafo conexo es un *punte* si el grafo obtenido al eliminar dicha arista no es conexo.

Definición 1.23.

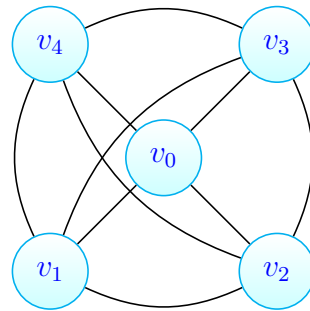
- El *coeficiente de clustering local* $C(v_i)$ de un vértice v_i puede definirse como la proporción media de pares de vecinos de un nodo que también son vecinos entre sí.
- De otro modo, para un nodo v_i con grado k_i , el *coeficiente de clustering local* está definido como

$$C(v_i) = \frac{\#\{\text{aristas (reales) entre los vecinos de } v_i\}}{\#\{\text{todas las posibles aristas entre los vecinos de } v_i\}}$$

donde todas las aristas posibles entre los vecinos de v_i son $\binom{k_i}{2} = \frac{k_i(k_i-1)}{2}$.



(a) $C_0 = \frac{2 \cdot 3}{6 \cdot 5} = \frac{1}{5}$. Ya que hay 3 aristas entre los vecinos de v_0 y todas las posibles aristas entre ellos son $\binom{6}{2}$.



(b) Grafo completo. $C_0 = \frac{2 \cdot 6}{4 \cdot 3} = 1$.

Figura 1.2: Cálculo del coeficiente de clustering de los nodos v_0 en un grafo completo y un grafo cualquiera.

Capítulo 2

Ciencia de Redes

Este capítulo tiene dos secciones claramente diferenciadas. En primer lugar estudiaremos qué son las redes aleatorias y daremos algunos ejemplos, veremos que son grafos que tienen la característica de que todos sus nodos tienen un grado similar y estudiaremos el algoritmo $G(n, p)$ que nos permitirá construirlas. Además, comprobaremos que la distribución de grados de este tipo de redes sigue una distribución de Poisson.

En segundo lugar, estudiaremos las llamadas redes de escala libre propuestas por Barabási y Albert. Veremos que la distribución de grados de este tipo de red obedece a una ley de potencia, debido a la presencia de ciertos tipos de nodos llamados hubs que se caracterizan por estar conectados a un gran número de nodos, es decir, que el grado de estos nodos es muy grande en relación a los demás nodos de la red. Este tipo de redes son libres de escala debido al cambio drástico que pueden experimentar los componentes que conforman la red al añadir continuamente nodos al sistema. Como referencia vamos a seguir fundamentalmente los textos [2, 4].

2.1. Redes aleatorias

Las redes aleatorias son unos tipos de grafos creados artificialmente con el objetivo de reflejar la topología de algunas redes reales. Algunos ejemplos de redes aleatorias son la red de carreteras de Estados Unidos y la red metabólica de la bacteria *E. Coli*. Los grafos aleatorios son grafos no dirigidos con N nodos donde cada par de nodos está conectado aleatoriamente con una probabilidad prefijada p . Su principal característica es que tienen una distribución de grado uniforme como veremos más adelante en esta sección.

Hay principalmente dos modelos de configuración para crear redes aleatorias, el modelo $G(N, p)$, también llamado de Erdős-Renyi, que toma como parámetros el número de nodos N que va a tener la red requerida y la probabilidad p de que dos nodos cualesquiera estén conectados; y el modelo $G(N, L)$, cuyos parámetros son el número de nodos de la red y el número de aristas que queremos que tenga la red. Nosotros nos vamos a centrar en el primer modelo.

Vamos a explicar el algoritmo para generar la red mediante el modelo $G(N, p)$, donde suponemos que hay N agentes, todos con igual probabilidad p de estar conectados entre sí:

1. Empezamos con N nodos aislados.

2. Seleccionamos un par de nodos y generamos un número aleatorio entre 0 y 1, si es menor que p conectamos ambos nodos con una arista.
3. Repetimos el paso (2) para cada uno de los $\frac{N(N-1)}{2}$ pares de nodos.

Observación 2.1. Cada generación de red aleatoria puede ser diferente ya que cada vez que se aplica el paso (2) del algoritmo se genera un número α_i aleatorio, luego puede tener distinto número de aristas y distinto diagrama.

Ejemplo 2.2. Vamos a utilizar el modelo $G(N, p)$ para crear un grafo aleatorio con 7 nodos y con probabilidad $p = 1/5$. En primer lugar vamos a generar $(7 \cdot 6)/2 = 21$ números aleatorios entre 0 y 1 con el programa Mathematica.

```
In [1]:= RandomReal[1, 21]

Out[1]= {0.988404, 0.186624, 0.0845713, 0.0643225, 0.0522086,
0.0872302, 0.651206, 0.406464, 0.0233091, 0.0389314,
0.443225, 0.0588116, 0.504359, 0.637999, 0.310423, 0.161128,
0.0343655, 0.223088, 0.434257, 0.626785, 0.484404}
```

Vamos a hacer una correspondencia 1 a 1 con la lista ordenada $\{(v_i, v_j)\}_{1 \leq i < j \leq 7}$, de la siguiente manera:

Número	Pareja	Nº en la lista
1	$\{v_1, v_2\}$	0,988404
2	$\{v_1, v_3\}$	0,186624
...
6	$\{v_1, v_7\}$	0,0872302
7	$\{v_2, v_3\}$	0,651206
...
20	$\{v_5, v_7\}$	0,626785
21	$\{v_6, v_7\}$	0,484404

Si el número de la lista correspondiente es menor que $1/5$ entonces pondremos una arista entre dichos vértices, en caso contrario no.

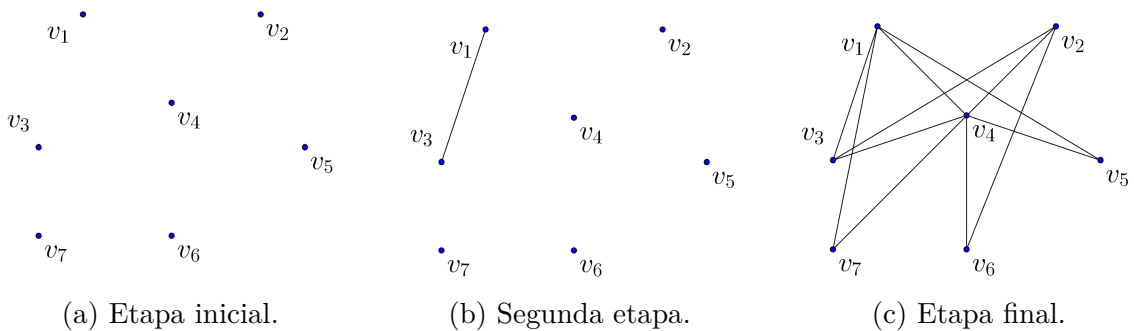


Figura 2.1: Ejemplo creado a partir del modelo $G(7, 1/5)$

2.1.1. Número de aristas

Es interesante determinar cuántas aristas podemos esperar de una generación de red aleatoria cualquiera, ya que con el mismo número de nodos y la misma probabilidad, el modelo nos puede generar redes diferentes.

La probabilidad de que una generación de la red aleatoria tenga exactamente L aristas viene dada por el producto de:

- el número de formas diferentes en las que podemos poner L aristas entre $\frac{N(N-1)}{2}$ pares de nodos,
- la probabilidad de que los intentos de conectar $\frac{N(N-1)}{2}$ pares de nodos resulten en una arista y,
- la probabilidad de que los intentos de conectar $\frac{N(N-1)}{2} - L$ restantes no resulten en una arista.

Es decir,

$$p_L = \binom{\frac{N(N-1)}{2}}{L} p^L (1-p)^{\frac{N(N-1)}{2}-L} \quad (2.1)$$

Proposición 2.3. *El número de aristas esperado en un grafo aleatorio es*

$$\langle L \rangle = \sum_{L=0}^{\frac{N(N-1)}{2}} L p_L = p \frac{N(N-1)}{2} \quad (2.2)$$

donde p es la probabilidad de que dos nodos estén conectados y p_L la probabilidad de que un nodo tenga L aristas.

Demostración. La primera igualdad es la definición de la esperanza matemática. Como (2.1) es una distribución binomial, sabemos que su esperanza es $p \frac{N(N-1)}{2}$. \square

De aquí podemos obtener el grado medio de una red aleatoria, denotando por $\langle k \rangle := \langle g(G) \rangle$ donde G es la red considerada:

$$\langle k \rangle = \frac{2\langle L \rangle}{N} = p(N-1) \quad (2.3)$$

Observación 2.4. De lo anterior tenemos que:

- A mayor probabilidad p , la red aleatoria se vuelve más densa. Si $p = 1$ el grafo es completo.
- $0 \leq \langle L \rangle \leq L_{max}$.
- $0 \leq \langle k \rangle \leq N - 1$.

2.1.2. Distribución de grado de una red aleatoria

Cuando generamos una red aleatoria cualquiera dicho grafo tiene un grado medio, sin embargo podemos también estudiar cómo se distribuyen los grados de los nodos de la red. Esta variación la mide la distribución de grado p_k , que es la probabilidad de que si elegimos un nodo cualquiera de la red tenga grado k .

Recordemos lo que es una distribución binomial y una distribución de Poisson.

Definición 2.5. A la función de probabilidad de una variable aleatoria X resultado de contar el número de éxitos al repetir n veces una experiencia aleatoria dicotómica (es decir, que tiene dos posibles resultados, A, \bar{A}) con probabilidad de éxito p la llamamos *distribución binomial* de parámetros n y p y la representamos por $\mathcal{B}(n, p)$.

Para esta distribución la variable X puede tomar los valores: $0, 1, 2, \dots, n$ y la variable aleatoria toma cada uno de estos valores con probabilidad:

$$p_k = \binom{n}{k} p^k (1-p)^{n-k}$$

Definición 2.6. Se dice que una variable aleatoria discreta X , cuyos valores posibles son: $0, 1, 2, \dots$, sigue una *distribución de Poisson* con parámetro λ y se escribe $P(k, \lambda)$, si su función de probabilidad es :

$$P(X = x) = e^{-\lambda} \frac{\lambda^x}{x!}$$

Donde $P(X = x)$ es la probabilidad de ocurrencia cuando la variable discreta X toma un valor finito x , λ es el promedio de ocurrencias en un intervalo y x es el número de ocurrencias.

Proposición 2.7. Dada una red aleatoria, la probabilidad de que un nodo v_i tenga grado $k = k_i$ sigue una distribución de probabilidad binomial de parámetros $N - 1$ y p , es decir,

$$p_k = \binom{N-1}{k} p^k (1-p)^{N-1-k}. \quad (2.4)$$

Demostración. Se sigue de la definición de distribución binomial. Siendo el suceso elemental incluir una arista entre dos vértices dados, que tiene probabilidad de éxito p . \square

Proposición 2.8. Para $N \gg 1$ y $\langle k \rangle$ fijo, la distribución Binomial se puede aproximar por la distribución de Poisson, que sólo depende del grado medio de la red $\langle k \rangle$:

$$p_k = e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!} \quad (2.5)$$

Demostración. Sabemos que

$$\binom{N-1}{k} = \frac{(N-1)(N-1-1)(N-1-2)\dots(N-1-k+1)}{k!} \approx \frac{(N-1)^k}{k!} \quad (2.6)$$

donde hemos utilizado que $N \gg k$.

Ahora podemos expresar $(1-p)^{N-1-k}$, utilizando $\langle k \rangle = p(N-1)$, del siguiente modo:

$$\log \left[(1-p)^{N-1-k} \right] = (N-1-k) \log \left(1 - \frac{\langle k \rangle}{N-1} \right)$$

y usando el desarrollo en serie de $\log(1+x) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} x^n$ tenemos que

$$\log [(1-p)^{N-1-k}] \approx (N-1-k) \frac{\langle k \rangle}{N-1} = -\langle k \rangle \left(1 - \frac{k}{N-1}\right) \approx -\langle k \rangle$$

donde las dos aproximaciones se vuelven a hacer utilizando $N \gg k$. Por tanto, tenemos que $(1-p)^{N-1-k} = e^{-\langle k \rangle}$.

Con todo esto y utilizando (2.6) en (2.4), quedaría:

$$p_k = e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!} \approx \frac{(N-1)^k}{k!} p^k e^{-\langle k \rangle} = \frac{(N-1)^k}{k!} \left(\frac{\langle k \rangle}{N-1}\right)^k e^{-\langle k \rangle}$$

Con lo cual, la distribución binomial la podemos aproximar de la siguiente forma:

$$p_k = e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!}$$

que es precisamente la distribución de Poisson. □

Comparemos las propiedades, ventajas e inconvenientes de una distribución binomial y distribución de Poisson. Podemos observar dichas propiedades en la imagen 2.2.

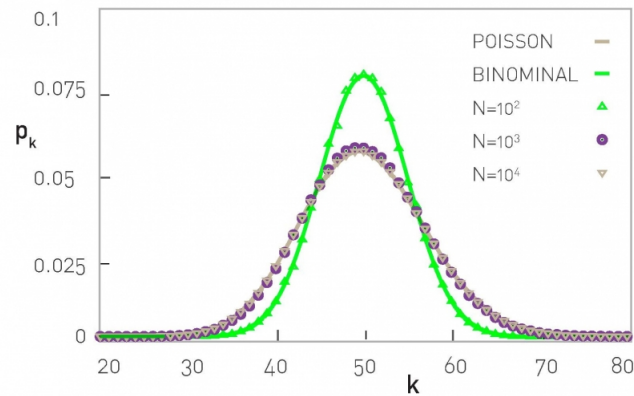


Figura 2.2: Comparación Binomial-Poisson.
Imagen procedente de [4, Figura 3.5].

- **Ambas distribuciones** tienen un pico que está centrado en torno a $k = \langle k \rangle$.
Si incrementamos p la red se vuelve más densa, si aumentamos $\langle k \rangle$ movemos el pico a la derecha.
Cuanto más densa es la red más dispersa es la distribución, por tanto mayores son las diferencias en los grados de los nodos, esto está controlado por p en el caso de la binomial y por $\langle k \rangle$ en el caso de la Poisson.
- Ventajas de utilizar la distribución de Poisson: Sólo depende de $\langle k \rangle$ y tiene una forma más simple que es útil a la hora de hacer cálculos. Además, la distribución de Poisson no diferencia el número de nodos de la red, por lo que podemos trabajar con redes de cualquier tamaño con $N \gg k$.

Por tanto, las redes aleatorias tienen la propiedad de ser homogéneas. La homogeneidad en la estructura de interacciones de estas redes significa que todos los nodos tienen un grado similar. Cada una de las $N(N - 1)/2$ posibles aristas aparece con la misma probabilidad de forma que la distribución de grados es binomial y en el límite cuando $N \rightarrow \infty$ es de Poisson.

2.2. Redes de escala libre

Al contrario que en las redes aleatorias donde asumíamos que el número de nodos, N , de la red no variaba, en las redes reales el número de nodos de la red puede crecer gracias a que se añadan nuevos nodos. Este es el caso de redes como la WWW donde los nodos son las páginas web y las aristas son las referencias entre páginas web, con el paso del tiempo se van creando nuevas páginas que se van conectando a las ya existentes; otro ejemplo es la red de citas científicas, donde los nodos son las publicaciones y las aristas son las referencias a otros artículos científicos, al igual que el caso anterior esta red cada vez se hace más grande, ya que con el paso del tiempo se publican más artículos (observemos que, en ambos ejemplos, las redes son dirigidas).

Una característica común que comparten este tipo de redes es que cada nodo nuevo que se añade a la comunidad tiende a conectarse con los nodos más populares, es decir, los que tienen un grado mayor. Este fenómeno se denomina conexión preferente. Observemos que en las redes anteriores, en el caso de la WWW, es más probable que en las páginas web creadas recientemente nos encontremos links a otra página que contengan muchos enlaces a que tengan un sólo enlace; por otro lado, en el caso de las redes de citas científicas será más probable que acabemos leyendo los artículos que tengan más citaciones.

Por tanto, las diferencias fundamentales entre las redes aleatorias y las que encontramos en el mundo real son el posible crecimiento en el número de nodos de las redes reales, al contrario que en las redes aleatorias; y la existencia del fenómeno conocido como conexión preferente en las redes reales.

En 1998, Laszlo, Jeong, Bianconi y Albert, trabajaron en la realización de un mapa de la Web, albergaban la esperanza de que el grafo obtenido en sus análisis fuera similar a una red aleatoria [3]. Para su sorpresa encontraron que la gran mayoría de las páginas de la Web tenía menos de 4 enlaces mientras que un 0,01 %, tenía mas de 1000, pero también advirtieron que estas diferencias de grado parecían ser importantes para sostener la conectividad total de Internet. La distribución de páginas con cierta cantidad de enlaces seguía una distribución en forma de “Ley de potencias”, haciendo que la probabilidad de que algún nodo esté conectado con k nodos sea proporcional a $1/k$ elevado a γ , es decir,

$$p_k \sim k^{-\gamma}.$$

Dicha distribución no tiene un pico como la curva binomial, sino que es una función continuamente decreciente y si se dibuja en un gráfico con una escala doble logarítmica aparece como una línea recta como podemos observar en la imagen 2.3.

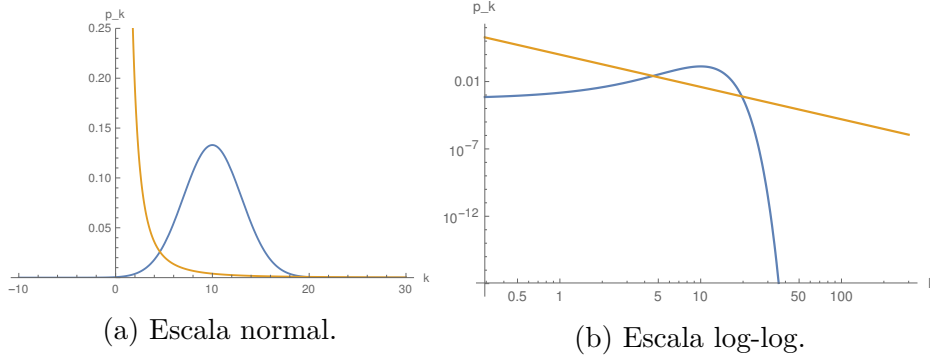


Figura 2.3: Distribución de Poisson con $\langle k \rangle = 10$ (gráfica azul) y distribución de escala libre con $\gamma = 2,4$ (gráfica naranja).

2.2.1. El algoritmo Barabási-Albert

El **algoritmo Barabási-Albert** quiere reflejar los dos fenómenos citados anteriormente, el crecimiento de las redes y la conexión preferente. Veamos en qué consiste dicho algoritmo:

- (1) Iniciamos el algoritmo con una red no dirigida con m_0 nodos con aristas entre ellos distribuidas de manera arbitraria de manera que cada nodo tenga al menos una arista.
- (2) En cada paso del algoritmo añadiremos un nodo con una cantidad de aristas prefijado, m , es decir, de grado m prefijado. Esto reflejará la condición de crecimiento de la red.

Por otro lado, para reflejar el fenómeno de la conexión preferente vamos a imponer que la probabilidad $\Pi(k)$ de que una arista del nuevo nodo añadido esté conectada con cierto nodo v_i con grado k_i sea

$$\Pi(k) = \frac{k_i}{\sum_j k_j}. \quad (2.7)$$

2.2.2. Variación de los grados en función del tiempo

Para estudiar la evolución del modelo Barabási-Albert en función del tiempo, tenemos que tener en cuenta la ecuación (2.7), que nos da la condición de la conexión preferente. Vamos a estudiar cómo varía el grado de un nodo cualquiera de la red en función del tiempo, es decir, conforme vamos añadiendo nodos a la red. Para ello aproximaremos el grado k_i de v_i mediante una función en la variable t y obtendremos el siguiente resultado:

Proposición 2.9. *La variación del grado k_i de un vértice v_i cualquiera con el modelo Barabási-Albert es:*

$$k_i(t) = m \left(\frac{t}{t_i} \right)^\beta \quad \text{donde } \beta = \frac{1}{2}. \quad (2.8)$$

Demostración. Observemos que la variación del grado de un nodo cualquiera v_i que adquiere nuevas aristas depende de los nuevos nodos que se incluyen en cada etapa con grado m prefijado. Entonces, la variación de grado de v_i se puede expresar del siguiente modo:

$$\frac{dk_i}{dt} = m\Pi(k_i) = m \frac{k_i}{\sum_{j=1}^{N-1} k_j}$$

ya que como cada nuevo nodo añadido tiene m aristas, entonces el nodo v_i tiene m oportunidades de ser elegido por la probabilidad de ser elegido (la conexión preferente).

Además, se tiene que $\sum_{j=1}^{N-1} k_j = 2mt - m$ ya que el sumatorio recorre el número de nodos de la red excepto el último añadido. Sustituyendo en la ecuación anterior y simplificando, se tiene que

$$\frac{dk_i}{dt} = \frac{k_i}{2t-1},$$

que, tomando límites cuando $t \rightarrow \infty$, se puede aproximar por

$$\frac{dk_i}{dt} = \frac{k_i}{2t} \Rightarrow \frac{dk_i}{k_i} = \frac{dt}{2t}.$$

Si integramos con la condición inicial $k_i(t_i) = m$, es decir, el nodo v_i se une a la red en tiempo t_i con m aristas. Se tiene entonces que el grado k_i de un vértice v_i cualquiera varía del siguiente modo a través del tiempo:

$$k_i(t) = m \left(\frac{t}{t_i} \right)^{\frac{1}{2}}$$

□

Observación 2.10. A partir $k_i(t)$ podemos calcular la velocidad a la que cierto nodo obtiene nuevas aristas:

$$\frac{dk_i(t)}{dt} = \frac{m}{2} \frac{1}{\sqrt{t_i t}}.$$

Observación 2.11. Los nodos que forman parte de la red en las etapas más tempranas tendrán un mayor grado conforme va aumentando el tiempo. Dichos nodos son los que tienen más posibilidades de convertirse en los llamados hubs (nodos con un número mucho mayor de aristas, los nodos más populares de la red) debido al fenómeno de la conexión preferente.

2.2.3. Distribución de grados

Vamos a calcular la distribución de grados p_k de las redes creadas a partir del algoritmo Barabási-Albert y vamos a comprobar que siguen una distribución en forma de ley de potencias.

Definición 2.12. Se dice que una variable aleatoria discreta X , cuyos valores posibles son: $0, 1, 2, \dots$, sigue una distribución en forma de ley de potencias si su función de probabilidad es de la forma

$$p_k = k^{-\gamma}.$$

Podemos calcular la distribución de grados de las redes creadas mediante el modelo Barabási-Albert mediante una aproximación discreta o mediante una aproximación continua.

Aproximación continua

Vamos a comprobar que las redes creadas a partir del modelo Barabási-Albert siguen una distribución de grados en forma de ley de potencias, para ello vamos a demostrar el siguiente resultado:

Proposición 2.13. *La distribución de grados de las redes creadas a partir del modelo Barabási-Albert siguen la distribución*

$$p(k) = 2m^{\frac{1}{\beta}}k^{\gamma} \text{ con } \gamma = \frac{1}{\beta} + 1 = 3. \quad (2.9)$$

Demostración. Como la probabilidad de que un vértice v_i tenga grado menor que k es $P[k_i(t) < k]$, podemos reescribir dicha probabilidad del siguiente modo usando (2.8):

$$P[k_i(t) < k] = P\left[t_i > t \left(\frac{m}{k}\right)^2\right]$$

Ahora, como en el modelo $t = N$ porque añadimos un nuevo nodo en cada etapa, y con esto, hay $N = m_0 + t$ nodos en la red en cada etapa. La probabilidad anterior se puede expresar del siguiente modo:

$$P\left[t_i > t \left(\frac{m}{k}\right)^2\right] = 1 - P\left[t_i \leq t \left(\frac{m}{k}\right)^2\right] = 1 - t \left(\frac{m}{k}\right)^2 \frac{1}{(t - m_0)}.$$

Como en el límite cuando $t \rightarrow \infty$ se tiene que $N \approx t$. Entonces, la probabilidad de que un nodo cualquiera tenga grado k o menor, que coincide con la distribución de grado acumulada, es

$$P[k_i(t) < k] = 1 - \left(\frac{m}{k}\right)^{\frac{1}{\beta}} \text{ con } \beta = \frac{1}{2}.$$

Finalmente, derivando esta ecuación obtenemos la distribución de grado:

$$p_k = \frac{\partial P(k)}{\partial k} = \frac{1}{\beta} \frac{m^{\frac{1}{\beta}}}{k^{\frac{1}{\beta}+1}} = 2m^2 k^{-3}.$$

□

Observación 2.14. Algunas de las ventajas de esta aproximación son que no hemos utilizado m para el cálculo de γ , además p_k no depende de t ni de N . Por otro lado, cuando $k \rightarrow \infty$ tenemos que $p_k = k^{-3}$, lo que nos da una distribución en forma de ley de potencias.

Aproximación discreta

Vamos a calcular la distribución de grados de una red creada mediante el modelo Barabási-Albert de forma recursiva y vamos a obtener el siguiente resultado:

Proposición 2.15. *La distribución de grados exacta de una red creada a partir del modelo Barabási-Albert es*

$$p_k = \frac{2m(m+1)}{k(k+1)(k+2)}, \quad (2.10)$$

Demostración. Denotemos por $N(k, t)$ el número de nodos con grado k en tiempo t de la red. La distribución de grado en tiempo t , $p_k(t)$, está relacionada con esta cantidad ya que se puede expresar como $p_k(t) = N(k, t)/N(t)$. En el modelo Barabási-Albert en cada etapa añadimos un nodo, entonces el número total de nodos es igual al de la etapa en la que estemos, luego $N = t$.

Entonces, podemos expresar la conexión preferente como:

$$\Pi(k) = \frac{k}{\sum_j k_j} = \frac{k}{2mt} \quad (2.11)$$

el término $2m$ refleja que en una red no dirigida cada arista contribuye al grado de dos nodos.

Se tiene que el número de aristas esperados a conectarse con nodos de grado k después de añadir un nuevo nodo es

$$\frac{k}{2mt} \cdot N p_k(t) \cdot m = \frac{k}{2} p_k(t), \quad (2.12)$$

donde:

- el primer término es la probabilidad de que el nuevo nodo vaya a conectarse con un nodo de grado k , es decir, refleja la conexión preferente.
- El segundo término es el número total de nodos con grado k . Además en la función refleja el hecho de que cuantos más nodos de grado k haya, mayor es la probabilidad de que un nuevo nodo se conecte a uno de ellos.
- Por último, m es el grado del nodo que vamos a añadir; en la ecuación nos dice que cuanto mayor es m , mayor es la probabilidad de que se conecte.

Veamos cómo varía el número de nodos con grado k después de que un nuevo nodo es añadido a la red. Observemos que pueden darse dos situaciones: que el nuevo nodo se conecte con un nodo de grado k , lo que convertiría este último en un nodo de grado $k + 1$; o que se conecte a un nodo de grado $k - 1$ lo que cambiaría el grado de este último a k . Veamos cómo afectan ambos casos a la ecuación (2.12): en el primer caso tendríamos que el número de nodos de grado k que cambian su grado a $k + 1$ es

$$\frac{k}{2} p_k(t)$$

y en el segundo, el número de aristas de grado $k - 1$ que cambian su grado a k es

$$\frac{k-1}{2} p_{k-1}(t).$$

Utilizando esto último podemos obtener el número de nodos de grado k en cada etapa, es decir, después de que añadimos un nodo:

$$(N + 1)p_k(t + 1) = N p_k(t) + \frac{k-1}{2} p_{k-1}(t) - \frac{k}{2} p_k(t), \quad (2.13)$$

ecuación que podemos utilizar para averiguar el número de nodos con grado $k > m$.

Como no consideramos los nodos con grados $k = 0, 1, \dots, m-1$ en la red, ya que todos los nodos tienen desde que se añaden como mínimo m aristas, necesitamos una ecuación diferente para los nodos de grado m . Siguiendo los mismos argumentos de antes, tenemos que:

$$(N+1)p_m(t+1) = Np_m(t) + 1 - \frac{m}{2}p_m(t). \quad (2.14)$$

Las ecuaciones (2.13) y (2.14) nos va a permitir calcular p_k de forma recursiva.

Pero primero, usemos el hecho de que estamos buscando una distribución de grados estática, es decir, tomando límites cuando $N = t \rightarrow \infty$ se tiene que $p_k(\infty) = p_k$. Si nos fijamos en (2.13), despejando

$$\frac{k-1}{2}p_{k-1}(t) - \frac{k}{2}p_k(t) = (N+1)p_k(t+1) - Np_k(t).$$

Y tomando el límite en la parte derecha de la ecuación:

$$(N+1)p_k(t+1) - Np_k(t) \rightarrow Np_k(\infty) + p_k(\infty) - Np_k(\infty) = p_k(\infty) = p_k.$$

Del mismo modo, para el caso m , fijándonos esta vez en (2.14):

$$(N+1)p_m(t+1) - Np_m(t) \rightarrow p_m.$$

Usando esto y despejando reescribimos (2.13) y (2.14), de la primera tenemos que:

$$\frac{k-1}{2}p_{k-1} = (N+1)p_k(t+1) - Np_k(t) + \frac{k}{2}p_k = p_k + \frac{1}{2}p_k = \frac{2+k}{2}p_k,$$

por tanto,

$$p_k = \frac{k-1}{k-2} \text{ para } k > m. \quad (2.15)$$

Y del mismo modo, para el segundo caso:

$$1 - \frac{m}{2}p_m = (N+1)p_m(t+1) - Np_m(t) = p_m$$

luego,

$$p_m = \frac{2}{m+2}. \quad (2.16)$$

Por último, vamos a aplicar la recursividad para calcular p_k . Escribimos la distribución de grado para el grado menor, es decir, cuando $k = m$ y usando las dos ecuaciones anteriores:

$$\begin{aligned} p_{m+1} &= \frac{m}{m+3}p_m = \frac{2m}{(m+2)(m+3)} \\ p_{m+2} &= \frac{m+1}{m+4}p_{m+1} = \frac{2m(m+1)}{(m+2)(m+3)(m+4)} \\ p_{m+3} &= \frac{m+2}{m+5}p_{m+2} = \frac{2m(m+1)}{(m+3)(m+4)(m+5)} \end{aligned}$$

A partir de aquí podemos obtener la fórmula recursiva, haciendo el cambio $m+3 = k$ en el denominador:

$$p_k = \frac{2m(m+1)}{k(k+1)(k+2)},$$

que nos da la probabilidad de que un nodo cualquiera tenga grado k , es decir, la fórmula exacta de la distribución de grado del modelo Barabási-Albert. \square

Corolario 2.16. *El modelo Barabási-Albert tiene una distribución en forma de ley de potencias, es decir, crea redes de escala libre.*

Demostración. A partir de la ecuación (2.13),

$$p_k = \frac{k-1}{2} p_{k-1} - \frac{k}{2} p_k,$$

podemos escribir

$$\begin{aligned} 2p_k &= (k-1)p_{k-1} - kp_k = -p_{k-1} - k(p_k - p_{k-1}) \\ 2p_k &= -p_{k-1} - k \frac{p_k - p_{k-1}}{k - (k-1)} \approx -p_{k-1} - k \frac{\partial p_k}{\partial k}. \end{aligned}$$

Obteniéndose

$$p_k = -\frac{1}{2} \frac{\partial k p_k}{\partial k}$$

que, integrando, es aproximadamente:

$$p_k \approx k^{-3}.$$

□

2.2.4. Propiedades de las redes de escala libre

Las diferencias principales entre una red aleatoria y una libre de escala se encuentran en la cola de la distribución:

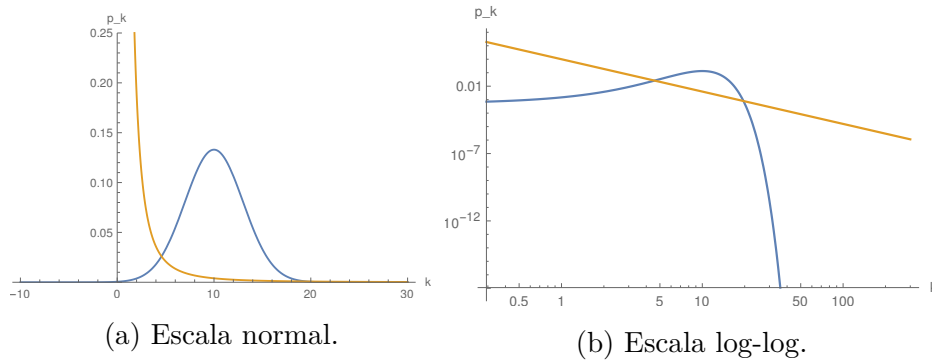


Figura 2.4: Distribución de Poisson con $\langle k \rangle = 10$ (gráfica azul) y distribución de escala libre con $\gamma = 2,4$ (gráfica naranja).

- Para valores pequeños de k , la ley de la potencia está por encima de la de Poisson: la red libre de escala tiene un número alto de nodos con grado pequeño.
- Para k cerca de $\langle k \rangle$, la distribución de Poisson está por encima de la ley de la potencia: en una red aleatoria hay muchos nodos con grado $k \approx \langle k \rangle$.

- Para valores altos de k , la ley de potencias vuelve a estar por encima de la de Poisson: la posibilidad de encontrar un nodo de grado alto, un hub, es varios órdenes de magnitud mayor en una red libre de escala.

Algunas de las **propiedades de las redes de escala libre** son:

1. Que se añadan nodos a través del tiempo en este tipo de redes permite la aparición de nodos cuyo grado supera varias veces el grado medio de la red, los llamados **hubs**. Esto se refleja en que la red sigue una distribución de grados en forma de ley de potencias, es decir, del tipo $p_k \sim k^{-\gamma}$. La cola de la distribución en forma de ley de potencias en el gráfico 2.3 nos indica que hay muy pocos hubs en comparación con los nodos con pocas conexiones al contrario que en las redes de escala libre donde todos los nodos de la red tienen un grado uniforme, el grado medio.
2. **Importancia de γ** : Mediante el valor de γ podemos clasificar las redes, sin embargo, el exponente γ en la mayoría de redes de escala libre suele estar entre 2 y 3.

Aunque no vamos a explicar concretamente los cálculos, si suponemos que en una red hay un solo hub, se puede calcular el grado de este hub en base al número de nodos de la red, obteniéndose: $k_{max} = k_{min} N^{\frac{1}{\gamma-1}}$ donde k_{min} es el grado mínimo de la red. Podemos encontrar el cálculo detallado de k_{max} en la sección 4.3 del libro [4].

Vamos a clasificar las redes de escala libre en:

- Régimen anómalo ($\gamma \leq 2$). El exponente $1/(\gamma - 1)$ en la expresión de k_{max} es mayor que 1, de esto se deduce que el número de nodos conectados al mayor hub se incrementa más rápido que el número de nodos existentes en la red, esto con el paso del tiempo hace que el hub se quede sin nodos a los que conectarse. Esta es una de las características que implican que este tipo de redes no pueden existir.
 - Régimen de escala libre ($2 < \gamma < 3$). Se tiene que el grado del hub, k_{max} , crece al mismo tiempo que el tamaño de la red con exponente $1/(\gamma - 1) < 1$. Este tipo de redes de escala libre son las más frecuentes, algunos ejemplos son la red de actores ($\gamma = 2,12$), las redes metabólicas ($\gamma = 2,2$) y las redes de interacciones de proteínas ($\gamma = 2,89$), véase [15, 21].
 - Régimen de red aleatoria ($\gamma \geq 3$). Este tipo de redes de escala libre no se diferencian mucho de las redes aleatorias. Para poder diferenciar ambas redes necesitamos que k_{min} y k_{max} se diferencien en varios ordenes de magnitud. Un ejemplo de este tipo de redes es la red de colaboración científica ($\gamma = 3,35$), véase [5].
3. Por último, las redes de escala libre tienen una topología de contacto muy heterogénea, algunos nodos están conectados a muchos vértices (hubs), mientras que otros sólo están conectados a unos pocos.

Capítulo 3

Modelos analíticos epidemiológicos

A lo largo de este capítulo veremos un resumen desde el punto de vista matemático de algunos de los modelos analíticos que se utilizan actualmente en epidemiología. Estos modelos se basan en dos hipótesis: **clasificar** cada individuo en base a la fase en la que se encuentre de una enfermedad, si el individuo es susceptible, infectado o ya ha superado la enfermedad; y que la **probabilidad de contagio es uniforme**, sólo depende de la topología de la red.

En primer lugar, vamos a clasificar a cada individuo basándonos en la etapa de la enfermedad en la que se encuentran. Así vamos a catalogar en

- **Susceptible (S)**: individuos sanos que aún no han tenido contacto con el patógeno.
- **Infectado (I)**: individuos contagiosos que han tenido contacto con el patógeno y por tanto pueden infectar a otros.
- **Recuperado (R)**: individuos que ya han sido contagiados anteriormente pero ya se han recuperado de la enfermedad, luego ya no son infectados.

La clasificación de los individuos de la red en estos estados va a dar lugar a diferentes modelos entre los que vamos a describir los modelos SI, SIS y SIR. Durante una epidemia los individuos pueden cambiar de un estado a otro, por ejemplo, inicialmente, cuando entra en contacto con el patógeno, un individuo puede pasar de susceptible a infectado y, cuando ya ha pasado la enfermedad, pasar a ser recuperado.

3.1. Modelo Susceptible-Infectado (SI)

Supongamos que una enfermedad se propaga por una población de N individuos. Denotemos por $S(t)$ el número de individuos susceptibles en tiempo t y por $I(t)$ el número de individuos infectados en tiempo t .

Como la topología de contactos es homogénea podemos suponer que un individuo cualquiera tiene de media $\langle k \rangle$ contactos y fijemos la probabilidad β de que la enfermedad se transmita de un individuo contagiado a uno susceptible.

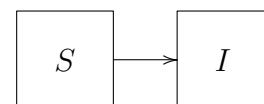


Figura 3.1: Modelo SI

Vamos a calcular el número de individuos que van a estar infectados en tiempo t suponiendo que en $t = 0$ hay un infectado, es decir, la condición inicial es $I(0) = 1$ (el estado inicial de la epidemia).

La probabilidad de que un individuo infectado se encuentre con un individuo susceptible es $S(t)/N$, entonces el individuo inicial mantendrá contacto con $\langle k \rangle S(t)/N$ individuos susceptibles en tiempo t . Como $I(t)$ individuos estarán extendiendo el patógeno con probabilidad β en tiempo t , la variación de infectados respecto al tiempo es:

$$\frac{dI(t)}{dt} = \beta \langle k \rangle \frac{S(t)I(t)}{N}$$

que denotando $s = S(t)/N$ e $i = I(t)/N$, y teniendo en cuenta que $N = S(t) + I(t)$:

$$\frac{di}{dt} = \beta \langle k \rangle si = \beta \langle k \rangle i(1 - i) \quad (3.1)$$

donde el producto $\beta \langle k \rangle$ se denomina *tasa de transmisión* o transmisibilidad. Resolvamos la ecuación diferencial (3.1) con condición inicial $I(0) = 1$:

$$\frac{di}{i} + \frac{di}{(1 - i)} = \beta \langle k \rangle dt \Rightarrow \log i - \log(1 - i) = \beta \langle k \rangle t + C.$$

Para $I(0) = 1$, es decir, $i_0 = 1/N$, se tiene $C = \log(i_0/(1 - i_0))$, con lo que, sustituyendo en la ecuación anterior y despejando, la fracción de individuos infectados a través del tiempo es

$$i = \frac{i_0 e^{\beta \langle k \rangle t}}{1 - i_0 + i_0 e^{\beta \langle k \rangle t}} \quad (3.2)$$

Observación 3.1. De esta última ecuación deducimos lo siguiente:

- Al principio, la fracción de individuos infectados se incrementa exponencialmente.
- El crecimiento de $i(t)$ disminuye cuando t crece, esto es debido a que cada vez quedan menos individuos susceptibles. Es decir, en el límite cuando t tiende a infinito todos los individuos estarán infectados, es decir, cuando $i(t \rightarrow \infty) = 1$ y $s(t \rightarrow \infty) = 0$.

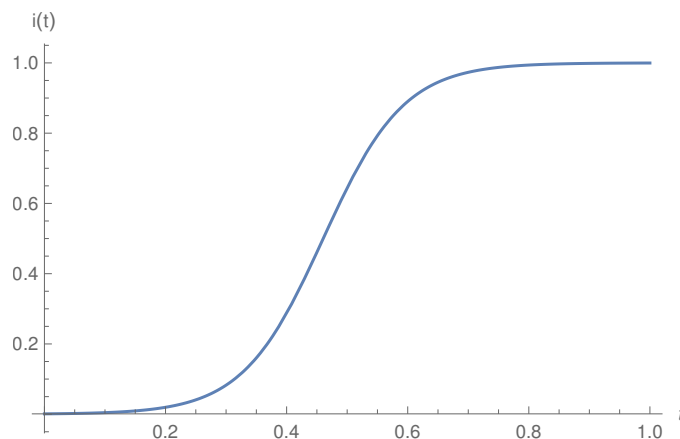


Figura 3.2: Modelo SI de parámetros $N = 1000$, $\beta = 0,5$ y $\langle k \rangle = 30$. Imagen generada en Mathematica a partir de la ecuación (3.2).

3.2. Modelo Susceptible-Infectado-Susceptible (SIS)

Mediante este modelo reflejamos el hecho de que la mayoría de los patógenos con el tiempo son anulados mediante el sistema inmune o un tratamiento, para ello permitimos que los individuos infectados que se recuperan de la enfermedad vuelvan a ser susceptibles. La diferencia con el modelo anterior se encuentra en que ahora permitimos que los individuos infectados se recuperen con una velocidad fija μ , convirtiéndose en susceptibles de nuevo. En este caso el modelo viene dado por una extensión de la ecuación (3.1):

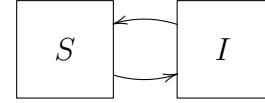


Figura 3.3: Modelo SIS

$$\frac{di}{dt} = \beta\langle k\rangle i(1-i) - \mu i, \quad (3.3)$$

donde μ es la *tasa de recuperación* y μi refleja la velocidad a la que la población se recupera de la enfermedad. Resolviendo la ecuación diferencial

$$\frac{di}{i(\beta\langle k\rangle(1-i) - \mu)} = dt \Rightarrow -\frac{\log(\mu + \beta\langle k\rangle(i-1)) - \log(i)}{\beta\langle k\rangle - \mu} = t + C$$

que junto con la condición inicial $I(0) = i_0 = 1/N$, se tiene:

$$i = \left(1 - \frac{\mu}{\beta\langle k\rangle}\right) \frac{C e^{(\beta\langle k\rangle - \mu)t}}{1 - C e^{(\beta\langle k\rangle - \mu)t}} \quad (3.4)$$

donde $C = i_0 / (1 - i_0 - \mu/\beta\langle k\rangle)$.

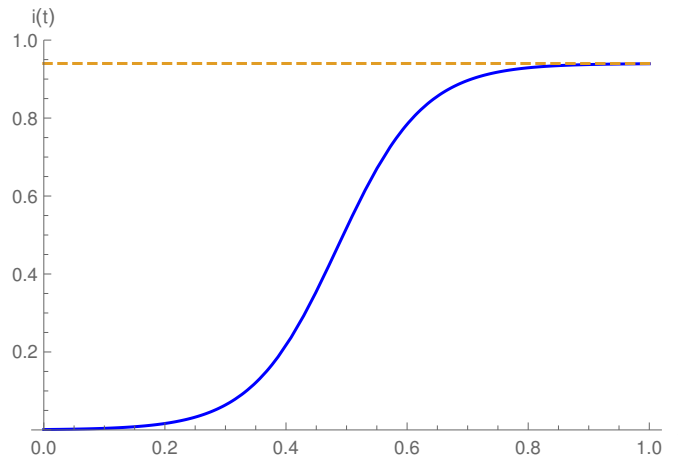


Figura 3.4: Modelo SIS de parámetros $N = 1000$, $\beta = 0,5$, $\mu = 0,9$ y $\langle k \rangle = 30$. Imagen generada en Mathematica a partir de la ecuación (3.4). En este caso $i(\infty) = 0,94$.

Mientras que en el modelo SI eventualmente todo el mundo acaba infectado, la ecuación (3.3) nos indica que en el modelo SIS la epidemia tiene dos posibles finales:

- **Estado endémico** ($\mu < \beta\langle k\rangle$). El número de nuevos individuos infectados es igual al número de individuos que se recuperan de la enfermedad, luego la porción de

individuos infectados de la sociedad no varía durante el tiempo. Esto es porque si $\mu < \beta\langle k \rangle$, el $\lim_{t \rightarrow \infty} i(t)$ es una constante menor que 1. Podemos calcular $i(\infty)$ que es el punto donde la epidemia se estabiliza, esto es cuando ya no aumenta el número de individuos infectados imponiendo $di/dt = 0$ en (3.3), obteniéndose

$$i(\infty) = 1 - \frac{\mu}{\beta\langle k \rangle}.$$

Podemos ver la gráfica de este caso en la imagen 3.4.

- **Estado libre de enfermedad** ($\mu > \beta\langle k \rangle$). Para una tasa de recuperación lo suficientemente alta el exponente $(\beta\langle k \rangle - \mu)t$ es negativo. Luego, i decrece exponencialmente a través del tiempo y por tanto la enfermedad desaparece, es decir, en este caso $i(\infty) = 0$.

Para entender cuándo tendremos una situación o la otra vamos a definir lo que es el número reproductivo básico de una enfermedad:

Definición 3.2. Denotaremos por $R_0 = \beta\langle k \rangle/\mu$ al número reproductivo básico, y se define como el número promedio de infectados que genera un individuo a lo largo de un período infeccioso.

- Si $R_0 > 1$ la epidemia se encuentra en el estado endémico. De hecho, si cada individuo infectado infecta a más de una persona sana entonces el patógeno se propagará rápidamente y persistirá entre la población. Cuanto mayor es el R_0 , más rápido es el proceso de difusión de la enfermedad.
- Si $R_0 < 1$ la epidemia se extingue, nos encontramos en el estado libre de enfermedad. Si cada individuo infecta de media a menos de una persona entonces el patógeno no puede persistir en la población.

3.3. Modelo Susceptible-Infectado-Recuperado (SIR)

Al contraer ciertos patógenos los individuos crean una inmunidad a la enfermedad a través del tiempo, el modelo SIR intenta reflejar este fenómeno. En vez de volver al estado susceptible después de superar la enfermedad, los individuos son retirados de la población. Luego el modelo está regido por las siguientes ecuaciones diferenciales:

$$\begin{cases} \frac{ds}{dt} = -\beta\langle k \rangle i(1 - r - i) \\ \frac{di}{dt} = -\mu i + \beta\langle k \rangle i(1 - r - i) \\ \frac{dr}{dt} = \mu i \end{cases} \quad (3.5)$$

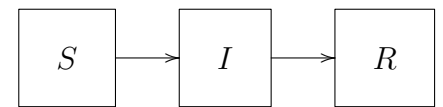


Figura 3.5: Modelo SI

que describen la evolución a través del tiempo de la fracción de individuos en los estados susceptible s , infectado i y recuperado r . Donde, como antes, los individuos infectados se recuperan con una velocidad fija μ pero ahora pasan a ser recuperados; y el factor $(1 - r - i)$ nos señala los individuos que dejan de ser susceptibles o infectados respectivamente.

En este modelo la cantidad de individuos infectados a través del tiempo tiene límite cero ya que cada vez habrá menos individuos susceptibles en la población, luego $i(\infty) = 0$.

3.4. Los modelos SI, SIS, SIR considerando el grado de los nodos

Todos estos modelos que hemos explicado anteriormente suponen que la red es completa y que la probabilidad de contagio es uniforme y, además, hemos supuesto el hecho de que cada individuo tiene $\langle k \rangle$ contactos y por esto no tienen en cuenta la existencia de nodos más populares que otros. Con esta última suposición disminuimos la complejidad del problema, ya que tenemos una misma ecuación para todos los nodos.

Para intentar predecir el comportamiento real de una epidemia con exactitud es importante estudiar el rol que la topología de contactos tiene en su expansión. No es lo mismo si el individuo que está contagiado no se relaciona con nadie o si tiene muchos contactos, en el primer caso no se extiende la enfermedad y en el segundo se pueden contagiar muchos nodos. Una posible forma de hacer esto es clasificando los nodos por grados y hacer los modelos en función del grado de cada nodo. Con esta variación del problema tenemos un sistema de ecuaciones donde cada ecuación se aplica a los nodos de cierto grado.

Los modelos siguientes difieren respecto a los que consideran el grado medio $\langle k \rangle$ en que:

- Cambian el grado medio de la red, $\langle k \rangle$, por el grado de cada nodo, k .
- En estos modelos se añade una función de densidad $\Theta_k(t)$ que representa la fracción de vecinos infectados de un nodo de grado k cualquiera, ya que esta va a depender del tiempo y del grado. Vamos a denotar $\Theta_k(t) := \Theta_k$ por simplicidad. Esto en los modelos anteriores se representaba simplemente por i , ya que estábamos asumiendo la hipótesis de que la probabilidad de contagio era uniforme.
- Ahora, al contrario que en los modelos que consideraban el grado medio de la red, vamos a tener un conjunto de ecuaciones, una para cada grado.

Podemos expresar $i = \sum_k p_k i_k$, donde $i_k = I_k/N_k$ es la fracción de individuos infectados con k contactos entre todos los individuos de grado k de la población. Por otro lado, se definen análogamente la fracción de individuos recuperados $r_k = R_k/N_k$ y la de individuos susceptibles $s_k = S_k/N_k$. Con esto, cada modelo se definirá de la siguiente manera:

1. En el modelo SI tenemos que considerar cómo varían los individuos infectados en función de su grado pero la ecuación tendrá la misma estructura que en 3.1, es decir la tasa de infección es proporcional a β y a la fracción de individuos que aún no han sido infectados, es decir $(1 - i_k)$:

$$\frac{di_k}{dt} = \beta(1 - i_k)k\Theta_k. \quad (3.6)$$

2. Análogamente, fijándonos en la ecuación (3.3), para el modelo SIS tendríamos:

$$\frac{di_k}{dt} = \beta\langle k \rangle(1 - i_k)k\Theta_k - \mu i_k. \quad (3.7)$$

3. Y, por último, para el modelo SIR, observemos que podemos obtener la fracción de individuos recuperados a partir de los infectados y recuperados. Además, este modelo tiene una ecuación parecida a (3.5), simplemente tenemos que incluir la Θ_k como en los anteriores, por lo que el modelo seguiría las siguientes ecuaciones:

$$\begin{cases} \frac{di_k}{dt} &= \beta s_k \Theta_k - \mu i_k \\ s_k &= 1 - r_k - i_k \end{cases} \quad (3.8)$$

No vamos a explicar con detalle estos modelos ya que se encuentra fuera del objetivo de este trabajo pero el lector interesado puede encontrar los modelos SI, SIS y SIR más detallados en el capítulo 10, sección 2 de [4] y los modelos que consideran el grado de los nodos en la sección 3 del capítulo 10 de [4].

Capítulo 4

Complejos Simpliciales

En el capítulo 1 hicimos una introducción sobre teoría de grafos. Observemos que en los grafos sólo había relaciones binarias, es decir, entre pares de vértices. En este capítulo vamos a introducir un nuevo concepto, los complejos simpliciales, que nos van a permitir estudiar situaciones en las que varios agentes tengan una interacción común, es decir, son una generalización de los grafos.

Por ejemplo, en el proceso de transmisión de una enfermedad, si tenemos un grupo de individuos que se relacionan continuamente, es lógico pensar que la probabilidad de que se contagie uno de ellos si un solo individuo está contagiado no será la misma que si varios de ellos están contagiados, véase [14].

Otro ejemplo es una red de colaboración científica, los nodos serían los distintos individuos y los grupos estarían formados por las personas que han trabajado en un mismo proyecto. Observemos que si tres científicos trabajan en un mismo proyecto están trabajando por parejas en él, con lo que se tendrían aristas entre cada pareja, pero también trabajan los tres conjuntamente por lo que se puede representar esta relación mediante el triángulo formado por los tres. Véase [17].

Más ejemplos de estudio de redes con complejos simpliciales son las interacciones proteína-proteína, las redes neuronales cerebrales y las redes semánticas que pueden encontrarse respectivamente en [19, 10, 18].

Para las primeras definiciones relativas a símplices y complejos simpliciales vamos a seguir mayormente los textos [11, 16].

4.1. Complejos simpliciales

Definición 4.1. El conjunto finito de puntos $\{v_0, v_1, \dots, v_q\}$ de \mathbb{R}^n se dice que están en posición general si los vectores $\{v_0 - v_1, \dots, v_0 - v_q\}$ son linealmente independientes.

Definición 4.2. Dado un conjunto finito de puntos en posición general, $\{v_0, v_1, \dots, v_q\}$ en \mathbb{R}^n , diremos que σ es un q -símplice si es el cierre convexo de dichos puntos, es decir,

$$\sigma = \left\{ \sum_{i=0}^q \lambda_i v_i : \sum_{i=0}^q \lambda_i = 1, \lambda_i \geq 0 \forall i \right\}.$$

Llamaremos vértices a los puntos $\{v_0, v_1, \dots, v_q\}$ de σ y diremos que su dimensión es q .

Proposición 4.3. Sea $q \geq 1$ y consideremos los puntos $v_0, v_1, \dots, v_q \in \mathbb{R}^n$. Las siguientes condiciones son equivalentes:

- (a) los vectores $v_0 - v_1, v_0 - v_2, \dots, v_0 - v_q$ son linealmente independientes,
- (b) si $\sum_{i=0}^q r_i v_i = \sum_{i=0}^q s_i v_i$ y $\sum_{i=0}^q r_i = \sum_{i=0}^q s_i$ entonces $r_i = s_i$ para todo $i \in \{0, 1, \dots, q\}$.

Demostración. Veamos en primer lugar (a) \Rightarrow (b). Supongamos que se cumple la hipótesis de (b), entonces si multiplicamos la segunda condición por v_0 se tiene que $\sum_{i=0}^q r_i v_0 = \sum_{i=0}^q s_i v_0$ y restando esto a la primera condición queda que $\sum_{i=1}^q r_i (v_0 - v_i) = \sum_{i=1}^q s_i (v_0 - v_i)$. Utilizando el apartado (a), se tiene que $s_i = r_i \forall i \in \{1, \dots, q\}$ entonces $\sum_{i=1}^q r_i = \sum_{i=1}^q s_i$. Y por ser $\sum_{i=0}^q r_i = \sum_{i=0}^q s_i$ se tiene $r_0 = s_0$.

Probemos ahora (b) \Rightarrow (a). Supongamos que $\sum_{i=1}^q \lambda_i (v_0 - v_i) = 0$, entonces, operando, se tiene que $\sum_{i=1}^q \lambda_i v_0 = \sum_{i=1}^q \lambda_i v_i$. Tomemos

$$\begin{cases} r_0 = 0, & r_i = \lambda_i & \forall 1 \leq i \leq q, \\ s_0 = \sum_{i=1}^q \lambda_i, & s_i = 0 & \forall 1 \leq i \leq q. \end{cases}$$

Entonces $\sum_{i=0}^q r_i v_i = \sum_{i=0}^q s_i v_i$ y usando (b) se tiene que $r_i = s_i \forall i \in \{0, \dots, q\}$. Por tanto, $\lambda_i = 0$ para todo $i \in \{1, \dots, q\}$ y los vectores $v_0 - v_1, v_0 - v_2, \dots, v_0 - v_q$ son linealmente independientes. \square

Observación 4.4. Decimos que $v_0, \dots, v_q \in \mathbb{R}^n$ están en posición general si satisfacen alguna de las condiciones de la proposición anterior. Por la condición (b), la propiedad de estar en posición general no depende del orden de los $\{v_i\}$.

Definición 4.5. Diremos que un subconjunto $\{v_{i_0}, \dots, v_{i_p}\}$ de $\sigma = \{v_0, v_1, \dots, v_q\}$ es una p -cara de σ . Si un símlice no es una cara de cualquier otro símlice, entonces decimos que es un *facet*.

Definición 4.6. Un *complejo simplicial* en \mathbb{R}^n es una colección, K , de símlices tal que

1. Si $\sigma \in K$ y τ es una cara de σ , entonces $\tau \in K$.
2. Dados dos símlices $\sigma, \sigma' \in K$, su intersección, $\tau = \sigma \cap \sigma'$, es una cara perteneciente a σ y σ' .

Observación 4.7. A cada elemento $\sigma \in K$ se le denomina q -símlice de K , siendo $q + 1$ el cardinal de σ . A la unión de los 0-símlices de K la vamos a llamar el *conjunto de los vértices de K* . A la unión de 1-símlices la vamos a denominar conjunto de aristas; a la de 2-símlices, conjunto de triángulos, etc.

La *dimensión de K* viene dada por $\dim K = \max\{\dim \sigma : \sigma \in K\}$. Vamos a denotar por $\sigma^{(q)}$ al símlice σ de dimensión q .

4.2. Medidas de adyacencia

Al igual que en los grafos tenemos la noción de grado, que nos da el número de aristas que inciden en un vértice, necesitamos una noción equivalente para símlices. Antes de

dar dicha idea tenemos que ahondar en la noción de adyacencia, que vamos a utilizar para comparar las relaciones entre dos símlices.

Hay que tener en cuenta que para definir una medida de adyacencia tenemos que expresar qué dos símlices vamos a comparar, en primer lugar, estudiaremos el caso en que los dos símlices sean de la misma dimensión. Si queremos expresar que ambos están contenidos en otro símlice de una dimensión mayor, utilizaremos la noción de adyacencia superior. Este es el caso de dos triángulos (2-símlices) que estén contenidos en un tetraedro (3-símlice). Por otro lado, si ambos contienen una misma cara de una dimensión menor, utilizaremos la adyacencia inferior. Por ejemplo, dos aristas que compartan un mismo vértice. Todas estas definiciones se pueden encontrar en [10] y en [11].

En segundo lugar, estudiaremos distintas formas de adyacencias superiores e inferiores en el caso de que los símlices que queremos comparar tengan distinta dimensión. Un ejemplo de estos casos es si consideramos un triángulo y un tetraedro y estudiamos qué tipo de relaciones inferiormente y superiormente tienen entre sí.

Una vez hecho esto, vamos a definir distintas nociones de grado en asociadas a dichas nociones de adyacencia.

4.2.1. Para símlices de la misma dimensión

Dados dos q -símlices $\sigma_i^{(q)}$, $\sigma_j^{(q)}$ de un complejo simplicial K .

Definición 4.8. Decimos que son *adyacentes inferiormente* si $\sigma_i^{(q)} \cap \sigma_j^{(q)} = \tau^{(q-1)}$ con $\tau^{(q-1)}$ $(q-1)$ -cara de ambos. A $\tau^{(q-1)}$ lo denominaremos *símlice inferior común*. Vamos a denotar a la *adyacencia inferior* como $\sigma_i^{(q)} \sim_L \sigma_j^{(q)}$.

Definición 4.9. Decimos que son *adyacentes superiormente* si ambos son caras del mismo $(q+1)$ -símlice, al que denominaremos *símlice superior común*. Vamos a denotar a la *adyacencia superior* como $\sigma_i^{(q)} \sim_U \sigma_j^{(q)}$.

Ejemplo 4.10. En la imagen 4.1, podemos observar que los 1-símlices $\{v_1, v_5\}$ y $\{v_1, v_7\}$ son **adyacentes inferiormente** ya que tienen una 0-cara o vértice en común v_1 .

Por otro lado, es fácil ver que los 2-símlices $\{v_1, v_2, v_3\}$ y $\{v_1, v_2, v_4\}$ son **adyacentes superiormente** porque ambos son caras del tetraedro $\sigma^{(3)}$.

Proposición 4.11. Si dos q -símlices son adyacentes superiormente, también lo son inferiormente.

Demostración. Sean $\sigma^{(q)}$ y $\sigma'^{(q)}$ dos símlices de dimensión q ambos caras de cierto $\sigma^{(q+1)}$. Por definición de símlice $\sigma^{(q+1)}$ es el cierre convexo de $\{v_0, \dots, v_{q+1}\}$ donde ciertos $\{v_{i_0}, \dots, v_{i_q}\}$ formarán parte de $\sigma^{(q)}$ y ciertos $\{v_{j_0}, \dots, v_{j_q}\}$ de $\sigma'^{(q)}$. La intersección de estos dos conjuntos formarán la $(q-1)$ -cara que tienen en común. \square

Proposición 4.12. Sea K un complejo simplicial finito, y sean $\sigma^{(d)}$, $\sigma'^{(d)}$ dos símlices en K . Si $\sigma^{(d)}$ y $\sigma'^{(d)}$ son adyacentes superiormente, entonces el $(d+1)$ -símlice que los contiene es único.

Demostración. Supongamos que τ y τ' son $(d+1)$ -símlices en K ambos conteniendo a $\sigma^{(d)}$ y $\sigma'^{(d)}$ como caras. Entonces $\sigma^{(d)} \cup \sigma'^{(d)} \subseteq \tau \cap \tau'$. Por la definición de complejo

simplicial, $\tau \cap \tau'$ será una cara tanto de τ como de τ' . Como τ es un $(d + 1)$ -símplice y $\sigma^{(d)}$ y $\sigma'^{(d)}$ son dos d -símplices distintos, es fácil ver que la única cara de τ conteniendo a $\sigma^{(d)}$ y $\sigma'^{(d)}$ es la misma τ . Entonces $\tau = \tau \cap \tau'$, lo que implica que τ es una cara de τ' . Como τ y τ' son ambos $(d + 1)$ -símplices, se tiene que $\tau = \tau'$. \square

Observación 4.13. En la literatura, véase [11], se puede encontrar esta otra definición de adyacencia inferior: dos q -símplices distintos $\sigma^{(q)}$, $\sigma'^{(q)}$ son inferiormente adyacentes en un complejo simplicial K , si ambos contienen un $(q - 1)$ -símplice τ en K como cara.

Ambas definiciones son equivalentes. Basta ver que si $\sigma^{(q)}$ y $\sigma'^{(q)}$ contienen a un $\tau^{(q-1)}$ entonces $\tau^{(q-1)} = \sigma^{(q)} \cap \sigma'^{(q)}$. Por hipótesis $\tau^{(q-1)} \subseteq \sigma^{(q)} \cap \sigma'^{(q)}$. Y por dimensiones $\tau^{(q-1)} = \sigma^{(q)} \cap \sigma'^{(q)}$ (por ser $\sigma^{(q)}$ y $\sigma'^{(q)}$ distintos).

Definición 4.14. El *grado inferior* de un q -símplice $\sigma^{(q)}$, al que vamos a denotar por $\deg_L(\sigma^{(q)})$, es el número de $(q - 1)$ -símplices en K que a su vez están contenidos en $\sigma^{(q)}$ (es decir, sus $(q - 1)$ -caras). El grado inferior de un q -símplice, por tanto, es siempre $\binom{q+1}{q} = q + 1$, que son las distintas combinaciones de $q + 1$ vértices tomados de q en q .

El *grado superior* de un q -símplice $\sigma^{(q)}$, al que vamos a denotar por $\deg_U(\sigma^{(q)})$, es el número de $(q + 1)$ -símplices en K tales que $\sigma^{(q)}$ es una q -cara de ellos.

Vamos a definir el *grado* de un q -símplice como:

$$\deg(\sigma^{(q)}) := \deg_L(\sigma^{(q)}) + \deg_U(\sigma^{(q)}) = \deg_U(\sigma^{(q)}) + q + 1$$

Ejemplo 4.15. En primer lugar vamos a entender mejor los distintos grados que podemos calcular en un símplice mediante un ejemplo:

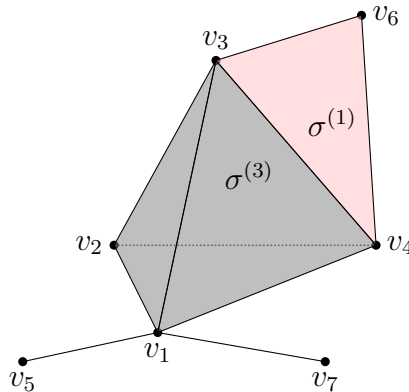


Figura 4.1: Complejo Simplicial K .

En la figura 4.1, vamos a calcular el **grado inferior** del único 3-símplice

$$\deg_L(\sigma^{(3)}) = \#\{2\text{-símplices contenidos en } \sigma^{(3)} \text{ y en } K\} = \binom{4}{3} = 4,$$

que es precisamente el número de triángulos que están contenidos en tetraedro $\sigma^{(3)}$, los formados por los conjuntos de vértices $\{v_1, v_2, v_3\}$, $\{v_1, v_2, v_4\}$, $\{v_2, v_3, v_4\}$ y $\{v_1, v_3, v_4\}$ respectivamente.

Ahora, vamos a calcular el **grado superior** de un s3mplice, por ejemplo, el de v_1 , que coincide con el grado normal de un v3rtice que vimos en el cap3tulo 1:

$$\deg_U(v_1) = \#\{1\text{-s3mplices en } K \text{ que contienen como cara a } v_1\} = 5.$$

Por 3ltimo, vamos a calcular el **grado** de un 1-s3mplice. En la figura 4.1 podemos observar que si denotamos por $\sigma^{(1)}$ a la arista que une v_3 y v_4 entonces

$$\deg(\sigma^{(1)}) = \deg_L(\sigma^{(1)}) + \deg_U(\sigma^{(1)}) = \#\{v_3, v_4\} + 3 = 5,$$

d3nde el grado superior de $\sigma^{(1)}$ es 3 porque est3 contenido en los 2-s3mplices formados por los cierres convexos de $\{v_3, v_4, v_6\}$, $\{v_1, v_3, v_4\}$ y $\{v_2, v_3, v_4\}$ respectivamente.

4.2.2. Adyacencias generales para s3mplices de distinta dimensi3n

Observemos que las anteriores definiciones de adyacencia no nos permiten comparar s3mplices de distinta dimensi3n. Por ejemplo, en el caso de que uno de ellos sea un tetraedro (3-s3mplice) y el otro sea un tri3ngulo (2-s3mplice) y ambos tuvieran una arista en com3n, necesitamos, por tanto, una nueva noci3n de adyacencia inferior; del mismo modo, si ambos estuvieran contenidos en un 5-s3mplice, necesitamos una nueva noci3n de adyacencia superior.

En esta secci3n vamos a definir distintas medidas de adyacencia y de grado para los casos en los que tenemos dos s3mplices de distinta dimensi3n. Algunas de estas nuevas medidas de adyacencia y grado se establecieron por primera vez en [11]. En [20] se incluye por primera vez una manera de contar los $(q+h)$ -s3mplices que son adyacentes superiormente a un q -s3mplice dado y una manera de contar el n3mero de facets incidentes a un v3rtice. Por 3ltimo, en [12], [13] se generalizan los grados introducidos por Moore-Drost-Swami.

Sean $\sigma_i^{(q)}$ un q -s3mplice y $\sigma_j^{(q')}$ un q' -s3mplice de un complejo simplicial K . Por simplicidad vamos a omitir los sub3ndices i, j para no recargar la notaci3n.

Definici3n 4.16. Diremos que $\sigma^{(q)}$ y $\sigma^{(q')}$ son *p-adyacentes inferiormente* si existe un p -s3mplice $\tau^{(p)}$ tal que es cara de ambos $\sigma^{(q)}$ y $\sigma^{(q')}$:

$$\sigma^{(q)} \sim_{L_p} \sigma^{(q')} \Leftrightarrow \exists \tau^{(p)} : \tau^{(p)} \subseteq \sigma^{(q)} \text{ y } \tau^{(p)} \subseteq \sigma^{(q')}.$$

Observaci3n 4.17. Si $\sigma^{(q)} \sim_{L_p} \sigma^{(q')}$ entonces $\sigma^{(q)} \sim_{L_{p'}} \sigma^{(q')}$ para todo $0 \leq p' \leq p$.

La observaci3n anterior nos da la siguiente definici3n:

Definici3n 4.18. Decimos que $\sigma^{(q)}$ y $\sigma^{(q')}$ son *estrictamente p-adyacentes inferiormente*, tambi3n denominado *p*-adyacentes inferiormente*, si $\sigma^{(q)} \sim_{L_p} \sigma^{(q')}$ pero $\sigma^{(q)} \not\sim_{L_{p+1}} \sigma^{(q')}$. Podemos denotar $\sigma^{(q)} \sim_{L_{p^*}} \sigma^{(q')}$ a la adyacencia estricta inferiormente.

Del mismo modo podemos definir los distintos tipos de adyacencia superior como:

Definici3n 4.19. Diremos que $\sigma^{(q)}$ y $\sigma^{(q')}$ son *p-adyacentes superiormente* si existe un p -s3mplice $\tau^{(p)}$ conteniendo a $\sigma^{(q)}$ y $\sigma^{(q')}$ como caras:

$$\sigma^{(q)} \sim_{U_p} \sigma^{(q')} \Leftrightarrow \exists \tau^{(p)} : \sigma^{(q)} \subseteq \tau^{(p)} \text{ y } \sigma^{(q')} \subseteq \tau^{(p)}.$$

Definición 4.20. Decimos que $\sigma^{(q)}$ y $\sigma^{(q')}$ son *estrictamente p -adyacentes superiormente*, también denominado *p^* -adyacentes superiormente*, si $\sigma^{(q)} \sim_{U_p} \sigma^{(q')}$ pero $\sigma^{(q)} \not\sim_{U_{p+1}} \sigma^{(q')}$. Podemos denotar $\sigma^{(q)} \sim_{U_{p^*}} \sigma^{(q')}$ a la adyacencia estricta superiormente.

Observación 4.21. Nótese que si $q = q'$ y $p = q - h$ se generaliza la noción de adyacencia inferior para símlices de la misma dimensión, pues para el caso en que $q = q'$ y $p = q - 1$ la noción de $(q - 1)$ -adyacencia inferior es igual a la de la adyacencia inferior de dos q -símlices de la definición 4.8.

Y del mismo modo, si $q = q'$ y $p = q + h$ se generaliza la noción de adyacencia superior para símlices de la misma dimensión, en el caso particular de $p = q + 1$, la $(q + 1)$ -adyacencia superior es igual a la definición 4.9 de adyacencia superior entre dos q -símlices.

Observación 4.22. Si $q = q'$ y $p = q + h$ con $h \geq 0$ y $\sigma^{(q)} \sim_{U_{q+h}} \sigma^{(q+h)}$, entonces $\sigma^{(q)}$ es una cara de $\sigma^{(q+h)}$ por la definición de $(q + h)$ -adyacencia superior y, por tanto, $\sigma^{(q)} \sim_{L_q} \sigma^{(q+h)}$.

Observación 4.23. Si $\sigma^{(q)} \sim_{L_p} \sigma^{(q')}$, entonces $\sigma^{(q)}$ y $\sigma^{(q')}$ comparten $p + 1$ vértices. Por tanto el menor símlice en el que estarán contenidos tendrá que tener $q+1+q'+1-(p+1) = q + q' - p + 1$ vértices, luego tendrá que ser un p' -símlice con $p' = (q + q' - p)$.

Proposición 4.24. Sea $\sigma^{(q)} \sim_{L_{p^*}} \sigma^{(q')}$ para cierto p y sea $p' = q + q' - p$. Si $\sigma^{(q)} \not\sim_{U_{p'}} \sigma^{(q')}$, entonces $\sigma^{(q)} \not\sim_{U_{p'+h}} \sigma^{(q')}$ para todo $h \geq 1$.

Demostración. Probemos el contrarrecíproco. Supongamos que $\sigma^{(q)} \sim_{U_{p'+h}} \sigma^{(q')}$ para algún $h \geq 1$. Entonces existirá algún $(p' + h)$ -símlice $\tau^{(p'+h)}$ tal que $\sigma^{(q)} \cup \sigma^{(q')} \subseteq \tau^{(p'+h)}$. En particular, $\tau^{(p'+h)}$ tiene una p' -cara conteniendo a ambos $\sigma^{(q)}$ y $\sigma^{(q')}$ como caras, entonces $\sigma^{(q)} \sim_{U_{p'}} \sigma^{(q')}$. \square

De estos resultados surge la siguiente definición:

Definición 4.25. Decimos que $\sigma^{(q)}$ y $\sigma^{(q')}$ son *p -adyacentes* si son p^* -adyacentes inferiormente pero no p' -adyacentes superiormente, para $p' = q + q' - p$:

$$\sigma^{(q)} \sim_{A_p} \sigma^{(q')} \Leftrightarrow \sigma^{(q)} \sim_{L_{p^*}} \sigma^{(q')} \text{ y } \sigma^{(q)} \not\sim_{U_{p'}} \sigma^{(q')}.$$

Decimos que $\sigma^{(q)}$ es *maximalmente p -adyacente* a $\sigma^{(q')}$ si:

$$\sigma^{(q')} \sim_{A_{p^*}} \sigma^{(q)} \Leftrightarrow \sigma^{(q')} \sim_{A_p} \sigma^{(q)} \text{ y } \sigma^{(q')} \not\sim_{A_p} \sigma^{(q'')} \quad \forall \sigma^{(q'')} : \sigma^{(q'')} \sim_{A_p} \sigma^{(q)}.$$

Ejemplo 4.26. La noción de p -adyacencia maximal nos permite ver cuáles son los símlices que se relacionan maximalmente con uno dado, es decir, de entre todos los símlices que son p -adyacentes a uno dado contar sólo los que sean facets. Por ejemplo, en la imagen 4.2, consideremos el triángulo $\sigma^{(2)}$, podemos ver que el tetraedro blanco y el triángulo azul son 1-adyacentes maximales a él. Sin embargo, los triángulos formados por los vértices $\{v_3, v_4, v_6\}$ y $\{v_3, v_4, v_5\}$ del tetraedro blanco son 1-adyacentes a $\sigma^{(2)}$ pero no son 1-adyacentes maximales por no ser facets.

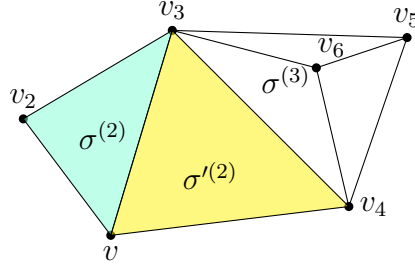


Figura 4.2: Complejo simplicial.

Definición 4.27. Podemos definir las siguientes nociones de **grado inferior**:

- el p -grado inferior de un q -símplice $\sigma^{(q)}$ como el número de q' -símplices que son p -adyacentes inferiormente a $\sigma^{(q)}$, es decir,

$$\deg_L^p(\sigma^{(q)}) := \# \{ \sigma^{(q')} : \sigma^{(q')} \sim_{L_p} \sigma^{(q)} \}.$$

- el p -grado inferior estricto de un q -símplice $\sigma^{(q)}$ como el número de q' -símplices que son p^* -adyacentes inferiormente a $\sigma^{(q)}$:

$$\deg_L^{p^*}(\sigma^{(q)}) := \# \{ \sigma^{(q')} : \sigma^{(q')} \sim_{L_p} \sigma^{(q)} \text{ y } \sigma^{(q')} \not\sim_{L_{p+1}} \sigma^{(q)} \}.$$

- el (h, p) -grado inferior de un q -símplice $\sigma^{(q)}$ como el número de $(q - h)$ -símplices que son p -adyacentes inferiormente a $\sigma^{(q)}$, es decir,

$$\deg_L^{h,p}(\sigma^{(q)}) := \# \{ \sigma^{(q-h)} : \sigma^{(q-h)} \sim_{L_p} \sigma^{(q)} \}.$$

- el (h, p) -grado inferior estricto de un q -símplice $\sigma^{(q)}$ como el número de $(q - h)$ -símplices que son p^* -adyacentes inferiormente a $\sigma^{(q)}$, es decir,

$$\deg_L^{h,p^*}(\sigma^{(q)}) := \# \{ \sigma^{(q-h)} : \sigma^{(q-h)} \sim_{L_p} \sigma^{(q)} \text{ y } \sigma^{(q-h)} \not\sim_{L_{p+1}} \sigma^{(q)} \}$$

Ejemplo 4.28. Teniendo como referencia la figura 4.3 vamos a calcular:

- el 1-grado inferior de $\sigma^{(3)}$:

$$\begin{aligned} \deg_L^1(\sigma^{(3)}) &= \# \{ \sigma^{(q)} : \sigma^{(q)} \sim_{L_1} \sigma^{(3)} \} \\ &= \# \{ \sigma^{(2)} \} + \# \{ \text{Aristas de } \sigma^{(3)} \} + \# \{ \text{Triángulos de } \sigma^{(3)} \} \\ &= 1 + 6 + 4 = 11. \end{aligned}$$

- el $(1, 1)$ -grado inferior de $\sigma^{(3)}$:

$$\begin{aligned} \deg_L^{1,1}(\sigma^{(3)}) &= \# \{ \tau^{(2)} : \tau^{(2)} \sim_{L_1} \sigma^{(3)} \} \\ &= \# \{ \sigma^{(2)} \} + \# \{ \text{Triángulos de } \sigma^{(3)} \} + \# \{ \text{Triángulos de } \sigma''^{(3)} \} \\ &\quad + \# \{ \text{Triángulos de } \sigma^{(4)} \text{ que contengan a } \{v, v_8\} \} \\ &= 1 + 4 + (4 - 1) + 3 = 11 \end{aligned}$$

donde en $\sigma''^{(3)}$ restamos 1 porque uno de los triángulos se repite (el triángulo $\{v_7, v_8, v_12\}$ está incluido en $\sigma^{(3)} \cap \sigma''^{(3)}$).

- el $(1, 1)$ -grado inferior estricto de $\sigma'^{(3)}$, es decir, $h = p = 1$:

$$\begin{aligned}
\deg_L^{1,1*}(\sigma'^{(3)}) &= \#\{\tau^{(2)} : \tau^{(2)} \sim_{L_1} \sigma'^{(3)} \text{ y } \tau^{(2)} \not\sim_{L_2} \sigma'^{(3)}\} \\
&= \#\{\sigma'^{(2)}\} + \#\{\text{triángulos de } \sigma'^{(4)} \text{ que contienen a } \{v, v_8\}\} \\
&\quad + \#\{\text{Triángulos de } \sigma''^{(3)} \text{ que sólo contienen una arista de } \sigma'^{(3)}\} \\
&= 1 + 3 + 3 = 7.
\end{aligned}$$

Observemos que con el $(1, 1)$ -grado inferior estricto estamos contando sólo las componentes de mayor dimensión que contengan sólo una arista de $\sigma'^{(3)}$.

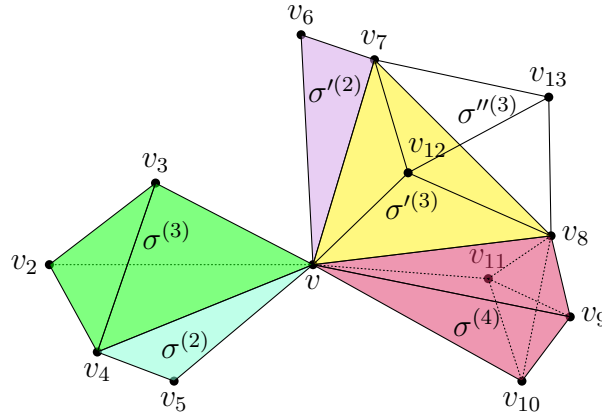


Figura 4.3: Complejo Simplicial K' .

Definición 4.29. Podemos definir las siguientes nociones de **grado superior**:

- el p -grado superior de un q -símplice $\sigma^{(q)}$ como el número de q' -símplices que son p -adyacentes superiormente a $\sigma^{(q)}$, es decir,

$$\deg_U^p(\sigma^{(q)}) := \#\{\sigma^{(q')} : \sigma^{(q)} \sim_{U_p} \sigma^{(q')}\}.$$

- el p -grado superior estricto de un q -símplice $\sigma^{(q)}$ como el número de q' -símplices que son p^* -adyacentes superiormente a $\sigma^{(q)}$:

$$\deg_U^{p*}(\sigma^{(q)}) := \#\{\sigma^{(q')} : \sigma^{(q)} \sim_{U_p} \sigma^{(q')} \text{ y } \sigma^{(q)} \not\sim_{U_{p+1}} \sigma^{(q')}\}.$$

- el (h, p) -grado superior de un q -símplice $\sigma^{(q)}$ como el número de $(q + h)$ -símplices que son p -adyacentes superiormente a $\sigma^{(q)}$, es decir,

$$\deg_U^{h,p}(\sigma^{(q)}) := \#\{\sigma^{(q+h)} : \sigma^{(q)} \sim_{U_p} \sigma^{(q+h)}\}.$$

- el (h, p) -grado superior estricto de un q -símplice $\sigma^{(q)}$ como el número de $(q + h)$ -símplices que son p^* -adyacentes superiormente a $\sigma^{(q)}$:

$$\deg_U^{h,p*}(\sigma^{(q)}) := \#\{\sigma^{(q+h)} : \sigma^{(q)} \sim_{U_p} \sigma^{(q+h)} \text{ y } \sigma^{(q)} \not\sim_{U_{p+1}} \sigma^{(q+h)}\}$$

Ejemplo 4.30. Denotemos $\tau^{(1)}$ a la arista que une $\{v, v_8\}$. En la imagen 4.3 vamos a calcular el 2-grado superior de $\tau^{(1)}$:

$$\begin{aligned}
\deg_U^2(\tau^{(1)}) &= \#\{\sigma^{(q)} : \tau^{(1)} \sim_{U_2} \sigma^{(q)}\} \\
&= \#\{\text{Triángulos de } \sigma'^{(3)} \text{ que contienen a } \tau^{(1)}\} \\
&+ \#\{\text{Aristas de } \sigma'^{(3)} \text{ que cumplen la condición}\} \\
&+ \#\{\text{Vértices de } \sigma'^{(3)} \text{ que cumplen la condición}\} \\
&+ \#\{\text{Triángulos de } \sigma^{(4)} \text{ que contienen a } \tau^{(1)}\} \\
&+ \#\{\text{Aristas de } \sigma^{(4)} \text{ que cumplen la condición}\} \\
&+ \#\{\text{Vértices de } \sigma^{(4)} \text{ que cumplen la condición}\} \\
&+ \#\{v\} + \#\{v_8\} \\
&= 2 + 4 + 2 + 3 + 6 + 3 + 1 + 1 = 22
\end{aligned}$$

Ahora, vamos a ver un ejemplo de cálculo de (h, p^*) -grado superior. Consideremos $h = 2$ y $p = 3$, calculemos el $(2, 3^*)$ -grado superior del vértice v .

$$\begin{aligned}
\deg_U^{2,3^*}(v) &= \#\{\tau^{(2)} : v \sim_{U_3} \tau^{(2)} \text{ y } v \approx_{U_4} \tau^{(2)}\} \\
&= \#\{\text{Triángulos de } \sigma'^{(3)}\} + \#\{\text{Triángulos de } \sigma^{(3)}\} \\
&= 4 + 4 = 8.
\end{aligned}$$

Definición 4.31. Definimos el p -grado adyacente de un q -símplice $\sigma^{(q)}$ como:

$$\deg_A^p(\sigma^{(q)}) := \#\{\sigma^{(q')} : \sigma^{(q)} \sim_{A_p} \sigma^{(q')}\}$$

Definición 4.32. Definimos el p -grado adyacente maximal de un q -símplice $\sigma^{(q)}$ como:

$$\deg_A^{p^*}(\sigma^{(q)}) := \#\{\sigma^{(q')} : \sigma^{(q)} \sim_{A_{p^*}} \sigma^{(q')}\}$$

Ejemplo 4.33. Vamos a calcular el 1-grado adyacente de $\sigma'^{(3)}$, en cada caso $p' = 3 + q' - 1 = q' + 2$.

$$\begin{aligned}
\deg_A^1(\sigma'^{(3)}) &= \#\{\sigma^{(q')} : \sigma'^{(3)} \sim_{A_1} \sigma^{(q')}\} \\
&= \#\{\sigma^{(q')} : \sigma'^{(3)} \sim_{L_1^*} \sigma^{(q')} \text{ y } \sigma'^{(3)} \approx_{U_{q'+2}} \sigma^{(q')}\} \\
&= \#\{\sigma'^{(2)}\} + \#\{\sigma^{(4)}\} \\
&+ \#\{\text{Tetraedros de } \sigma^{(4)} \text{ que contienen a } \{v, v_8\}\} \\
&+ \#\{\text{Triángulos de } \sigma^{(4)} \text{ que contienen alguna arista de } \sigma'^{(3)}\} \\
&+ \#\{\text{Triángulos de } \sigma''^{(3)} \text{ que contienen alguna arista de } \sigma'^{(3)}\} \\
&= 1 + 1 + 3 + 3 + 3 = 11.
\end{aligned}$$

Ahora, vamos a calcular el 1-grado adyacente maximal de $\sigma'^{(3)}$. Para ello tenemos que ver cuáles de los símlices y caras contados en el calculo del 1-grado adyacente anterior verifican la segunda condición. Es decir, tenemos que ver cuáles son facets de entre todos

los contados anteriormente. Y los únicos que no cumplen esta segunda condición son las caras de $\sigma^{(4)}$, por tanto:

$$\begin{aligned}\deg_A^{1*}(\sigma'^{(3)}) &= \#\{\sigma'^{(2)}\} + \#\{\sigma^{(4)}\} \\ &\quad + \#\{\text{Triángulos de } \sigma''^{(3)} \text{ que contienen alguna arista de } \sigma'^{(3)}\} \\ &= 1 + 1 + 3 = 5.\end{aligned}$$

Observación 4.34. El p -grado adyacente maximal de un q -símplice sólo cuenta los símlices maximales p -adyacentes de un símlice dado.

Capítulo 5

Modelos Simpliciales en Epidemiología

En el capítulo 3 vimos algunos modelos epidemiológicos que se aplicaban a redes, basados en las relaciones entre pares de individuos. Sin embargo, podemos pensar en situaciones más complejas, en las que un individuo sano tenga que estar en contacto con dos o más individuos contagiados para poder contraer la enfermedad; es decir, un contacto que sólo tenga un vecino contagiado no es suficiente para que este contraiga la enfermedad.

Vamos a estudiar un modelo basado en esta última hipótesis, el modelo SCM propuesto por Iacopini, Petri, Barrat y Latora, que puede encontrarse en el artículo [14]. Para poder entenderlo mejor, en primer lugar, veremos un algoritmo de generación de complejos simpliciales basado en el algoritmo de generación de redes aleatorias $G(n, p)$ que ya estudiamos anteriormente.

5.1. Algoritmo RSC

Una forma de generar complejos simpliciales aleatorios es mediante el algoritmo que vamos a ver a continuación, este está basado en el algoritmo $G(n, p)$ que vimos en el capítulo 2. Recordemos que el algoritmo $G(n, p)$ se basa en dar dos valores iniciales: el número de nodos n y la probabilidad p de que dos nodos cualesquiera estén unidos.

El algoritmo de generación de complejos simpliciales aleatorios se denomina algoritmo *RSC*, que son las siglas en inglés de “random simplicial complexes”. Este algoritmo nos permite mantener un (i, i) -grado superior medio $\langle \deg_U^{i,i}(v) \rangle$ entre los vértices del complejo simplicial para cada i , es decir, que todos los vértices del complejo simplicial tengan un número medio de aristas, triángulos, tetraedros, etc. Recordemos que

$$\deg_U^{i,i}(v) := \# \left\{ \sigma^{(i)} : v \sim_{U_i} \sigma^{(i)} \right\},$$

es decir, cuenta el número de i -símplices en los que está contenido v . A los distintos grados medios vamos a denotarlos por $\langle \deg_U^{i,i} \rangle$ para no recargar la notación.

Los parámetros iniciales de este algoritmo son el número de nodos n y las probabilidades $\{p_1, \dots, p_m\}$, donde p_j es la probabilidad de que dados $j + 1$ vértices, estos formen un j -símplice.

5.1.1. Algoritmo RSC para $m=2$

Para el caso $m = 2$, los parámetros iniciales son (n, p_1, p_2) y el algoritmo es el siguiente:

1. En primer lugar, vamos a crear los 1-símplices. Al igual que en el algoritmo $G(n, p_1)$ creamos una lista ordenada de parejas (v_i, v_j) con $0 \leq i < j \leq n$ y una lista de números aleatorios entre 0 y 1 tantos como parejas haya y vamos a unir dos vértices v_i y v_j si el número correspondiente es menor que p_1 .
2. Del mismo modo, para crear los 2-símplices, hacemos una lista de las posibles triplas (v_i, v_j, v_k) con $0 \leq i < j < k \leq n$ y creamos otra lista de números aleatorios. Unimos v_i, v_j y v_k mediante un triángulo relleno si el número correspondiente es menor que p_2 .

Para los casos en que $m > 2$, se puede extender el algoritmo análogamente.

Ejemplo 5.1. Vamos a generar un complejo simplicial con los parámetros $(7, 1/5, 1/10)$.

1. Podemos utilizar el primer paso del ejemplo 2.2. Entonces tenemos las aristas:

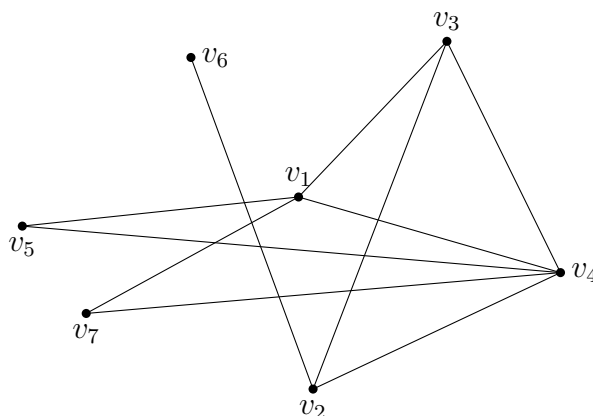


Figura 5.1: Etapa 1 algoritmo RSC.

2. Para el segundo paso, generamos una lista de $\binom{7}{3} = \frac{7!}{(7-3)!3!} = 35$ números, esto es el número de 3-tuplas sin orden ni repetición que se pueden hacer en un conjunto de 7 elementos.

```
In [2]:= RandomReal[1, 35]
```

```
Out[2]= {0.052415, 0.636277, 0.346155, 0.533427, 0.347406,
0.878891, 0.169006, 0.118943, 0.205094, 0.996771, 0.409327,
0.49475, 0.00936684, 0.0360493, 0.602858, 0.0272144,
0.785248, 0.443358, 0.589911, 0.701515, 0.808285, 0.242054,
0.729435, 0.00169908, 0.691968, 0.734377, 0.935193,
0.806095, 0.807192, 0.696801, 0.888702, 0.612262, 0.772521,
0.249138, 0.166074}
```


Y vamos a numerar las 3-tuplas de vértices $\{(v_i, v_j, v_k)\}_{1 \leq i < j < k \leq 7}$ de la siguiente manera:

Número	3-tupla	Nº en la lista
1	$\{v_1, v_2, v_3\}$	0,052415
2	$\{v_1, v_2, v_4\}$	0,636277
...
5	$\{v_1, v_2, v_7\}$	0,347406
6	$\{v_1, v_3, v_4\}$	0,878891
...
34	$\{v_4, v_6, v_7\}$	0,249138
35	$\{v_5, v_6, v_7\}$	0,166074

que a su vez vamos a hacer corresponder con los números que aparecen en la lista anterior, según orden de aparición (tercera columna).

Con la lista anterior las únicas 3-tuplas que nos van a aparecer en el complejo simplicial son: $\{v_1, v_2, v_3\}$, $\{v_1, v_5, v_6\}$, $\{v_1, v_5, v_7\}$, $\{v_2, v_3, v_4\}$ y $\{v_2, v_5, v_7\}$. Con lo cual nos quedaría el complejo simplicial 5.2:

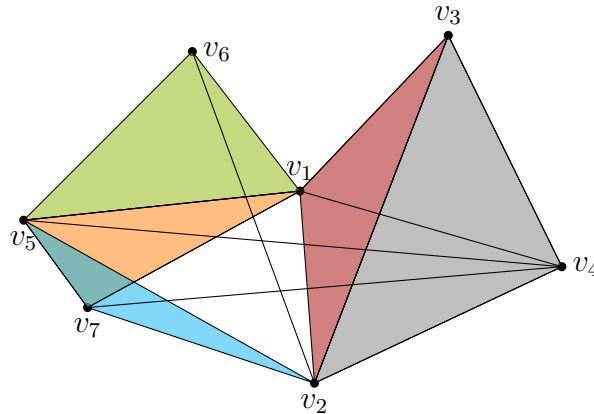


Figura 5.2: Complejo Simplicial Aleatorio generado por el algoritmo RSC. Etapa final.

Este complejo simplicial está formado por la unión de los siguientes conjuntos:

Los 0-simplices (vértices): $\{v_1, v_2, v_3, v_4, v_5, v_6, v_7\}$

Los 1-simplices (aristas): $\{\{v_1, v_4\}, \{v_2, v_6\}, \{v_4, v_5\}, \{v_4, v_7\}\}$

Los 2-simplices (triángulos): $\{\{v_1, v_2, v_3\}, \{v_1, v_5, v_6\}, \{v_1, v_5, v_7\}, \{v_2, v_3, v_4\}, \{v_2, v_5, v_7\}\}$

Observación 5.2. Los triángulos que se han incluido en la segunda etapa pueden hacer que se incremente el número de aristas del complejo simplicial, ya que algunas pueden no haberse incluido en la primera etapa del algoritmo. Por ejemplo, la arista $\{v_1, v_2\}$ ha sido añadida en la segunda etapa al añadirse el triángulo rosa.

Observación 5.3. Este modelo permite la presencia de subgrafos completos de $(k + 1)$ vértices sin que sean k -simplices. Esto se da en el ejemplo anterior, podemos observar que tenemos triángulos “reellenos” y “vacíos”, codificando las relaciones entre un grupo de 3 vértices y entre grupos de 2 vértices respectivamente.

Observación 5.4. Matemáticamente, hay otra manera de representar relaciones entre varios individuos, mediante hipergrafos. Los complejos simpliciales son un caso particular de hipergrafos ya que estos últimos consideran relaciones entre varios individuos sin tener que considerar subrelaciones entre ellos, es decir, no considera la primera condición de la definición de complejo simplicial. Más información sobre hipergrafos y modelos epidemiológicos en hipergrafos puede encontrarse en [6].

Cálculo de (i, i) -grados superiores medios de los vértices

Vamos a calcular el número esperado de i -símplices incidentes en un nodo cualquiera del complejo simplicial, al que vamos a denotar por $\langle \text{deg}_U^{i,i} \rangle$. Sólo vamos a calcular los casos en los que $i = 1$ y $i = 2$ que son los que podemos encontrar en [14]; nos vamos a centrar en calcular el número medio de aristas del complejo simplicial que corresponde al grado medio normal, el que denominábamos $\langle k \rangle$ y el número medio de triángulos en el está incluido un vértice cualquiera.

El proceso que vamos a seguir es el siguiente, como el algoritmo RSC para $m = 2$ está dividido en dos etapas (primero incluye las aristas y luego los triángulos), entonces vamos a calcular primero el número medio de aristas que tiene un vértice cualquiera en la etapa 1 y luego vamos a calcular el número de triángulos en la etapa 2. Como al añadir los triángulos en la segunda etapa del algoritmo se modifica el número de aristas en todo el complejo simplicial, tenemos que modificar el número medio de aristas que hemos calculado en la primera etapa.

Entonces, vamos a ver cómo cambia el número medio de aristas y triángulos en cada etapa para los parámetros (n, p_1, p_2) .

- **Etapa 1.** En esta etapa sólo podemos calcular el caso en que $i = 1$, dado un conjunto de n vértices con probabilidad p_1 de que dos vértices cualesquiera v_i, v_j estén unidos con una arista. Podemos calcular el número medio de aristas que sale de un vértice cualquiera, que es el mismo que cuando calculábamos el grado medio de una red aleatoria generada con el algoritmo $G(n, p_1)$. En nuestro caso tenemos que el grado medio de aristas de un nodo cualquiera del complejo simplicial $\langle \text{deg}_U^{1,1} \rangle$ es el número de nodos con los que puede estar unido mediante una arista, $N - 1$, por la probabilidad de que dos nodos cualesquiera estén unidos p_1 , es decir,

$$\langle \text{deg}_U^{1,1} \rangle = p_1(n - 1)$$

que es el $(1, 1)$ -grado superior medio de los vértices del grafo que tenemos en esta etapa. Observemos que coincide con el cálculo de $\langle k \rangle$ que hemos hecho en (2.3).

- **Etapa 2.** Calculemos primero el caso en que $i = 2$, es decir el número medio de triángulos que contienen un vértice cualquiera, y a partir de este calcularemos el caso en que $i = 1$, el número medio de aristas del complejo simplicial (ya que en este segundo paso al añadir los triángulos vamos a añadir más aristas de las que había en la primera etapa).

Como la probabilidad de que haya un 2-símplice con vértices v_i, v_j, v_k cualesquiera es p_2 . Entonces, el número medio de triángulos adyacentes a un vértice cualquiera

del complejo simplicial es el producto de la probabilidad p_2 por el número de combinaciones de los $n - 1$ vértices restantes cogidos de dos en dos, que coincide con $\binom{n-1}{2}$, luego,

$$\langle \deg_U^{2,2} \rangle = p_2 \frac{(n-1)(n-2)}{2}, \quad (5.1)$$

que es precisamente el $(2, 2)$ -grado superior medio de los vértices del complejo simplicial obtenido por el algoritmo RSC.

Calculemos ahora $\langle \deg_U^{1,1} \rangle$ del complejo simplicial. Observemos que podemos calcularlo considerando el $\langle \deg_U^{1,1} \rangle$ en la primera etapa y estudiando cómo influye a este grado los triángulos que hemos añadido en la segunda etapa.

Si consideramos un vértice cualquiera v_i , cada triángulo que sea adyacente a él $\{v_i, v_j, v_k\}$ para ciertos i, j contribuirá al $(1, 1)$ -grado superior con 2 aristas: $\{v_i, v_j\}$ y $\{v_i, v_k\}$ si ninguna de estas dos aristas estaban incluidas en la anterior etapa, esto pasará con probabilidad $(1 - p_1)^2$, que es probabilidad de la intersección de los dos sucesos, ya que ambos tienen probabilidad $(1 - p_1)$.

Análogamente, si una de las dos aristas anteriores ya estaba en la etapa 1, por ejemplo la arista $\{v_i, v_k\}$, entonces la otra arista $\{v_i, v_j\}$ contribuirá en 1 al grado, el razonamiento es análogo en el caso de que esté $\{v_i, v_j\}$ pero no $\{v_i, v_k\}$. Como cada caso ocurre con la misma probabilidad $p_1(1 - p_1)$ y la intersección de estos sucesos es vacía entonces la contribución al $(1, 1)$ -grado superior es la suma de ambas probabilidades, esto es $2p_1(1 - p_1)$. En general, el grado del vértice v_i se incrementa de media en

$$2 \cdot (1 - p_1)^2 + 2p_1(1 - p_1) = 2(1 - p_1)$$

por cada 2-símple que tenga a v_i como uno de sus vértices.

Finalmente, para $p_1, p_2 \ll 1$, podemos escribir el $(1, 1)$ -grado superior medio esperado de cualquier vértice $\langle \deg_U^{1,1} \rangle$ como la suma de las contribuciones al grado de las aristas y las contribuciones de las aristas de los 2-símplices, es decir,

$$\langle \deg_U^{1,1} \rangle \approx (n-1)p_1 + 2\langle \deg_U^{2,2} \rangle(1 - p_1), \quad (5.2)$$

donde hemos multiplicado $2(1 - p_1)$ por $\langle \deg_U^{2,2} \rangle$ porque es el número medio de triángulos que se añaden en la segunda etapa.

Observación 5.5. A partir de las expresiones (5.2) y (5.1) podemos despejar p_1 y p_2 para poder calcular las probabilidades que necesitamos introducir en el algoritmo RSC para tener ciertos (i, i) -grados superiores medios.

Se puede calcular el $\langle \deg_U^{i,i} \rangle$ de los vértices del complejo simplicial de forma análoga en las siguientes etapas del algoritmo SCM.

Para el caso $i = 3$: siguiendo un razonamiento similar, ahora como estamos añadiendo tetraedros hay que volver a calcular el número medio de triángulos y de aristas que contienen un vértice cualquiera:

$$\begin{cases} \langle \deg_U^{3,3} \rangle = p_3 \frac{(n-1)(n-2)(n-3)}{2 \cdot 3} \\ \langle \deg_U^{2,2} \rangle = p_2 \frac{(n-1)(n-2)}{2} + 3\langle \deg_U^{3,3} \rangle(1 - p_2) \\ \langle \deg_U^{1,1} \rangle = (n-1)p_1 + 2\langle \deg_U^{2,2} \rangle(1 - p_1) + 3\langle \deg_U^{3,3} \rangle(1 - p_1) \end{cases}$$

Como podemos observar, según vamos aumentando en valor de m , el valor del grado $\langle \text{deg}_U^{i,i} \rangle$ depende de todos los anteriores. En el caso general, tendremos que:

$$\begin{aligned} \langle \text{deg}_U^{i,i} \rangle &= p_i \frac{(n-1)(n-2) \cdot (n-i)}{i!} \\ &\vdots \\ \langle \text{deg}_U^{1,1} \rangle &= (n-1)p_1 + 2\langle \text{deg}_U^{2,2} \rangle(1-p_1) + \dots + i\langle \text{deg}_U^{i,i} \rangle(1-p_1) \end{aligned}$$

5.2. Modelo de Contagio Simplicial (SCM)

El modelo propuesto por Iacopini, Petri, Barrat y Latora para complejos simpliciales se denomina en inglés “Simplicial Contagion Model”, en adelante nos referiremos a él por sus siglas SCM.

El SCM de orden D es una variante de los modelos SI y SIS para complejos simpliciales. Con el objetivo de modelar un proceso de contagio en un complejo simplicial K vamos a asociar a cada vértice v_i de K una función dependiente del tiempo que represente el estado del nodo en cada instante, a esta función la vamos a denotar por $x_{v_i}(t) \in \{0, 1\}$ con $i = 1 \dots n$. Esta función será 1 si el vértice v_i está infectado y 0 si es susceptible.

En cada instante t consideremos la función:

$$\rho(t) = \frac{1}{n} \sum_{i=1}^n x_{v_i}(t), \quad (5.3)$$

que es la fracción de nodos infectados en la población en tiempo t . Observemos que esta función $\rho(t)$ la denominábamos i en el capítulo 2 cuando estábamos estudiando los modelos analíticos en epidemiología.

Consideremos además los siguientes parámetros $\{\beta_1, \dots, \beta_D\}$ donde cada β_i es la probabilidad de que un nodo susceptible v cualquiera de la red que forme parte de un símlice $\sigma^{(D)}$ sea infectado por cada una de las caras de dimensión i de $\sigma^{(D)}$ en las que esté contenido, bajo la condición de que todos los otros vértices de dichas caras están infectados.

En particular, β_1 corresponde a la probabilidad de que un nodo v_i se infecte por el nodo infectado v_j a través de la arista $\{v_i, v_j\}$. Del mismo modo β_2 es la probabilidad de que un nodo cualquiera v_i sea infectado a través de un 2-símlice $\{v_i, v_j, v_k\}$ donde v_j, v_k son nodos cualesquiera de la red y ambos están infectados.

Por otro lado, al igual que en el modelo SIS, consideraremos también el parámetro μ que es la probabilidad de recuperación de cualquier nodo v_i de K .

5.2.1. SCM de orden D

Vamos a estudiar el SCM usando la teoría del campo medio, esta consiste en reemplazar ciertos parámetros de un modelo por la media, de esta manera obviamos fluctuaciones; es una aproximación pero resulta muy efectiva para comprender mejor el comportamiento de un sistema. Estudiemos la variación del número de nodos infectados en un complejo simplicial a través del tiempo.

Sea $B = \{\beta_\omega, \omega = 1, 2, \dots, D\}$ el conjunto de las probabilidades de infección y μ la probabilidad de recuperación de un nodo cualquiera de la complejo simplicial. Vamos a asumir que entre dos nodos cualesquiera v_i, v_j no son dependientes los estados en el que se encuentre cada uno a través del tiempo: $x_{v_i}, x_{v_j} \forall i, j$ respectivamente. El SCM de orden D propuesto viene determinado por la siguiente ecuación:

$$\frac{d\rho(t)}{dt} = -\mu\rho(t) + \sum_{\omega=1}^D \beta_\omega \langle \deg_U^{\omega, \omega} \rangle \rho^\omega(t) [1 - \rho(t)] \quad (5.4)$$

que nos da la fracción de nodos infectados a través del tiempo y donde $\rho^\omega(t) := [\rho(t)]^\omega$.

Donde el primer término hace que se quite la fracción de nodos del complejo simplicial que ya se han recuperado, con probabilidad μ . El sumatorio nos añade la fracción de nuevos vértices infectados en cada instante de tiempo; y cada sumando del sumatorio nos da la fracción nueva de nodos infectados por arista ($\omega = 1$), triángulo ($\omega = 2$), tetraedro ($\omega = 3$), \dots , respectivamente.

Observemos que si en un triángulo no todos los vértices están infectados no hay que considerar la probabilidad β_2 , ya que se convierte en el caso de que una arista está contagiada. Lo mismo nos pasa en los ω -símplices con $\omega > 2$, tienen que estar los ω nodos contagiados para que el nodo que estamos estudiando tenga probabilidad β_ω de infección, el resto de casos están agrupados en los sumandos anteriores. Por tanto, para cada vértice sólo hay que considerar las caras de mayor dimensión en las que todos los vértices estén infectados, esto se puede observar en la imagen 5.3, de aquí que se multiplique por $\rho^\omega(t)$.

Observación 5.6. En el caso $D = 1$ el SCM tiene la siguiente ecuación:

$$\frac{d\rho(t)}{dt} = -\mu\rho(t) + \beta_1 \langle \deg_U^{1,1} \rangle \rho(t) [1 - \rho(t)]$$

que, si cambiamos $\rho(t)$, β_1 y $\langle \deg_U^{1,1} \rangle$ por i , β y $\langle k \rangle$ respectivamente, es el mismo que

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

que es la ecuación (3.3) que vimos en el capítulo 1 en el modelo SIS. Por tanto, el caso $D = 1$ el SCM se reduce al modelo SIS que vimos anteriormente.

Estudiemos el caso en el que $D = 2$ para entender el comportamiento de la enfermedad en función de los diferentes parámetros iniciales y comprender el primer caso puramente simplicial no trivial.

5.2.2. SCM de orden $D = 2$

Podemos ver el diagrama que nos explica qué ocurre en cada uno de los casos para el caso $D = 2$ en la figura 5.3.

La ecuación del SCM de orden 2 es

$$\frac{d\rho(t)}{dt} = -\mu\rho(t) + \beta_1 \langle \deg_U^{1,1} \rangle \rho(t) [1 - \rho(t)] + \beta_2 \langle \deg_U^{2,2} \rangle \rho^2(t) [1 - \rho(t)]. \quad (5.5)$$

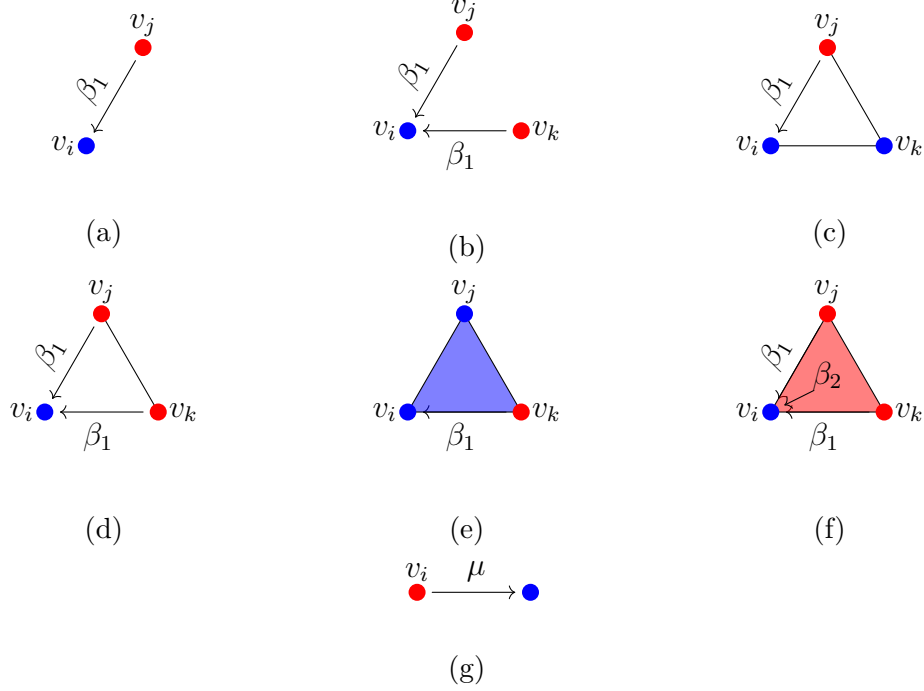


Figura 5.3: Figura creada a semejanza de la figura 1 del artículo [14].

a-f Las diferentes formas de que contraiga la enfermedad el nodo v_i mediante el SCM para el caso $D = 2$. **g** es el cambio de estado de infectado a susceptible. Los vértices en rojo son individuos infectados y los azules individuos susceptibles. β_1 y β_2 son las probabilidades de infección por arista y por un triángulo respectivamente; y μ la probabilidad de recuperación.

En **a**, **c**: v_i sólo está en contacto con un individuo infectado por una arista, mientras que en **b**, **d** está en contacto con dos individuos infectados ambos por medio de una arista. En ambos casos se tiene una probabilidad de contagio β_1 .

En **e** y **f**, el nodo v_i pertenece a un triángulo. En **e** se tiene que la infección sólo se puede contraer mediante la arista que une v_i con v_k con probabilidad de contagio β_1 ; y en **f** como sus dos vecinos están contagiados puede darse la infección por medio de las aristas correspondientes con probabilidad β_1 y mediante el triángulo con probabilidad β_2 .

Por último, **g** representa la transformación de infectado a recuperado que para cada nodo infectado v_i se da con probabilidad μ .

Como en el modelo SIS, podemos calcular $\rho(\infty) = \lim_{t \rightarrow \infty} \rho(t)$, que es el punto donde la epidemia se estabiliza, esto es cuando ya no aumenta el número de individuos infectados. Imponemos $\frac{d\rho(t)}{dt} = 0$ para obtener los puntos críticos de la función $\rho(t)$, entonces:

$$-\mu\rho(t) + \beta_1 \langle \text{deg}_U^{1,1} \rangle \rho(t) [1 - \rho(t)] + \beta_2 \langle \text{deg}_U^{2,2} \rangle \rho^2(t) [1 - \rho(t)] = 0.$$

Dividiendo por μ y sacando factor común $\rho(t)$ se tiene que

$$-\rho(t) \left[1 - \frac{\beta_1 \langle \text{deg}_U^{1,1} \rangle}{\mu} [1 - \rho(t)] - \frac{\beta_2 \langle \text{deg}_U^{2,2} \rangle}{\mu} \rho(t) [1 - \rho(t)] \right] = 0. \quad (5.6)$$

Observemos que podemos denotar por

$$\lambda_1 := \frac{\beta_1 \langle \text{deg}_U^{1,1} \rangle}{\mu} = R_0$$

que, como sabemos, es el número reproductivo básico de la definición 3.2. Por otro lado, podemos denotar por

$$\lambda_2 = \frac{\beta_2 \langle \text{deg}_U^{2,2} \rangle}{\mu},$$

que va a ser el número promedio de parejas de individuos infectadas que genera un individuo a lo largo de un periodo infeccioso. Por tanto, el número reproductivo básico que vimos en la definición 3.2 va a convertirse en un vector de probabilidades $\{\lambda_1, \lambda_2\}$, donde habrá tantos λ_i como la dimensión del complejo simplicial.

Sustituyendo los valores de λ_1 y λ_2 en (5.6), la ecuación nos queda del modo siguiente:

$$-\rho(t) [1 - \lambda_1 [1 - \rho(t)] - \lambda_2 \rho(t) [1 - \rho(t)]] = 0.$$

Las soluciones de esta ecuación $-\rho [1 - \lambda_1(1 - \rho) - \lambda_2 \rho(1 - \rho)] = 0$ son:

$$\rho_1 = 0 \quad \text{y} \quad \rho_{2\pm} = \frac{\lambda_2 - \lambda_1 \pm \sqrt{(\lambda_1 - \lambda_2)^2 - 4\lambda_2(1 - \lambda_1)}}{2\lambda_2}. \quad (5.7)$$

Por lo anterior, podemos reescribir la ecuación (5.5) del siguiente modo:

$$\frac{d\rho(t)}{dt} = -\rho(t) [\rho(t) - \rho_{2+}] [\rho(t) - \rho_{2-}]. \quad (5.8)$$

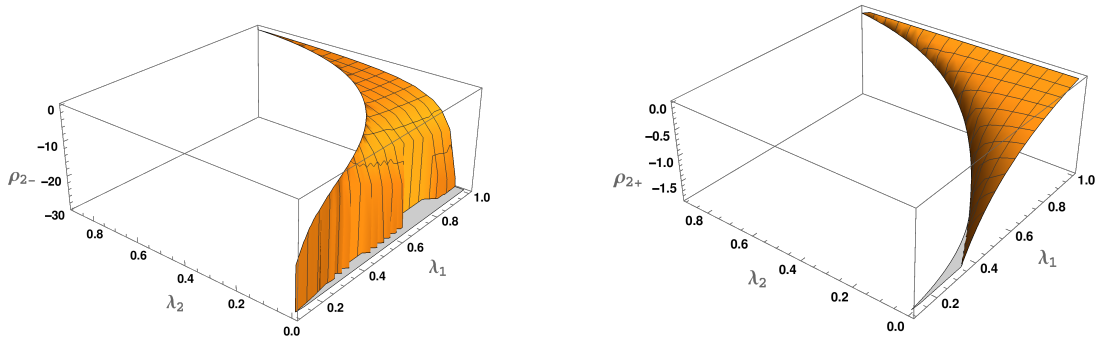
Estudiamos los posibles finales de la enfermedad dados por el modelo SCM de orden 2. De igual forma que en el modelo SIS, vamos a estudiar $\rho(\infty)$ en función de los valores de λ_1 y λ_2 . Como ya hemos visto, las soluciones de la ecuación $\frac{d\rho(t)}{dt} = 0$ son ρ_1, ρ_{2+} y ρ_{2-} ; y, por otro lado, sabemos que el valor $\rho(\infty)$, en el cual la enfermedad se estabiliza, tiene que estar en el rango $[0, 1]$.

La solución $\rho_1 = 0$ corresponde al estado libre de enfermedad en el que todos los individuos se recuperan y la enfermedad se extingue. Veamos cuándo se dan y cuáles son los valores de las otras dos soluciones ρ_{2+} y ρ_{2-} .

Como (5.7) son funciones en dos variables, a semejanza de cuando estábamos estudiando los posibles finales del modelo SIS en el capítulo 1 en función de R_0 ; vamos a estudiar dos casos diferenciados, vamos a estudiar los posibles finales de la enfermedad con modelo SCM en función de si $\lambda_2 \leq 1$ y si $\lambda_2 > 1$.

Veamos en primer lugar el caso en que $\lambda_2 \leq 1$. Para ello vamos a estudiar ρ_{2+} y ρ_{2-} en función de los posibles valores que puede tomar λ_1 . Vamos a separar en 3 casos:

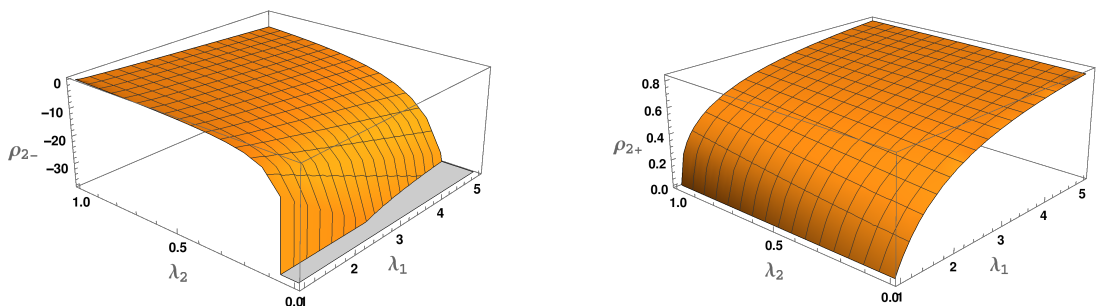
1. En el caso en que $0 \leq \lambda_1 < 1$ se tiene que cuando ρ_{2-} y ρ_{2+} son un números reales son siempre negativos, como se puede observar en la imagen 5.4. Por tanto, en este caso la única posible solución es $\rho_1 = 0$ (recordemos que las posibles soluciones siempre tienen que estar en el intervalo $[0, 1]$) y la enfermedad desaparece con el tiempo, la enfermedad se encuentra en un estado libre de enfermedad.



(a) Valores de ρ_{2-} en función de λ_1 y λ_2 . (b) Valores de ρ_{2+} en función de λ_1 y λ_2 .

Figura 5.4: Cálculo de las soluciones ρ_{2-} y ρ_{2+} en función de los valores de λ_1 y λ_2 entre $[0, 1]$. En ambos casos ρ_{2+} y ρ_{2-} son siempre negativos.

2. En el caso en que $\lambda_1 > 1$ (véase la imagen 5.5), ρ_{2-} sigue siendo negativo para todos los valores de λ_1 y λ_2 , sin embargo, ρ_{2+} pasa a ser positivo. Entonces, cuando la fracción de infectados, $\rho(t)$, es pequeña se tiene que $d\rho(t)/dt > 0$ y la enfermedad se estabiliza en la solución ρ_2 .



(a) Valores de ρ_{2-} en función de λ_1 y λ_2 . (b) Valores de ρ_{2+} en función de λ_1 y λ_2 .

Figura 5.5: Cálculo de las soluciones ρ_{2-} y ρ_{2+} en función de los valores de λ_1 entre $[1, 5]$ y de λ_2 entre $[0, 1]$. Se tiene que ρ_{2-} siempre es negativa y ρ_{2+} es mayor o igual que 0.

3. Por último, cuando $\lambda_1 = 1$ se tiene que $\rho_{2+} = 0$ y $\rho_{2-} < 0$, como refleja la imagen 5.6. Por tanto, la fracción de infectados cuando $t \rightarrow \infty$ tiende a 0, es decir, la enfermedad desaparece.

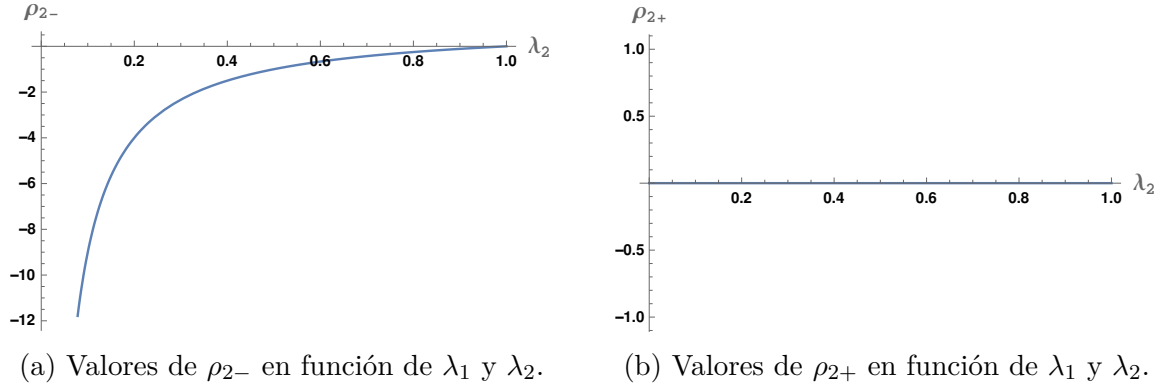


Figura 5.6: Cálculo de las soluciones ρ_{2-} y ρ_{2+} en función de los valores de $\lambda_1 = 1$ y de λ_2 entre $[0, 1]$. Se tiene que ρ_{2-} siempre es negativa y $\rho_{2+} = 0$ para todo valor de λ_2 .

Veamos ahora qué ocurre en el caso en que $\lambda_2 > 1$. De igual forma que en el caso anterior vamos a separar el estudio en función de los valores de λ_1 . Observemos que si $\lambda_1 < \lambda^c = 2\sqrt{\lambda_2} - \lambda_2$ entonces el discriminante de (5.7) es menor que cero, es decir, ρ_{2+} y ρ_{2-} no son soluciones reales. Por tanto, en este caso la única solución posible es $\rho_1 = 0$ y por ello la enfermedad desaparece cuando $t \rightarrow \infty$. Entre los demás valores de λ_1 vamos a diferenciar cuando es mayor o menor que 1.

1. Si $\lambda_1 > 1$ se tiene que $\rho_{2-} < 0 < \rho_{2+}$ como se puede observar en la imagen 5.7. Además, cuando la fracción de infectados $\rho(t)$ es pequeña se tiene que $d\rho(t)/dt$ es positiva, por esto la epidemia no puede estabilizarse en el estado absorbente y la densidad de nodos infectados cuando avanzamos en el tiempo se estabiliza en ρ_{2+} .

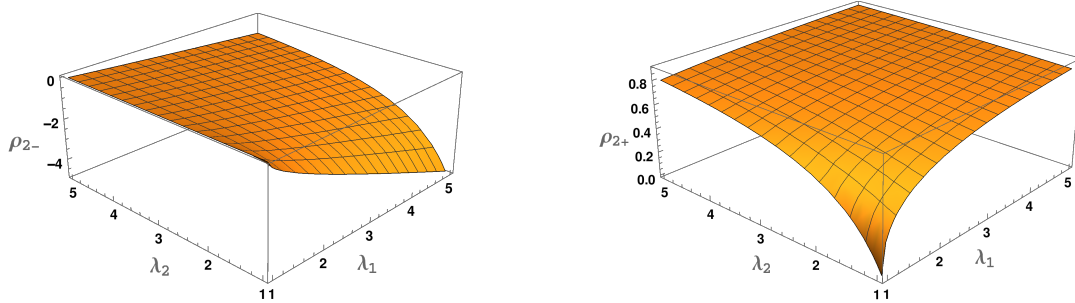


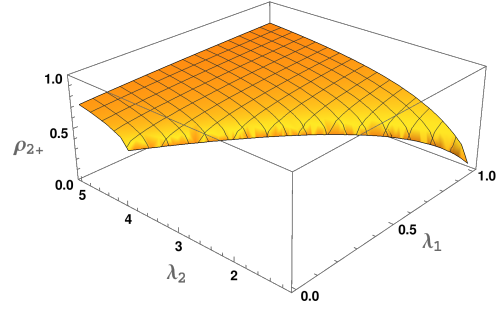
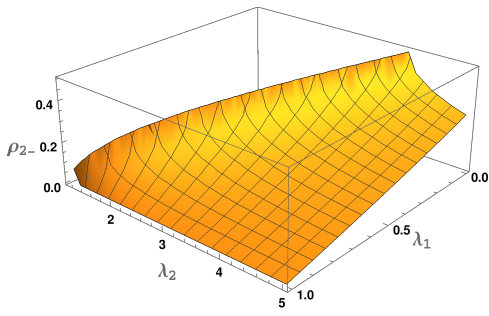
Figura 5.7: Cálculo de las soluciones ρ_{2-} y ρ_{2+} si λ_1 se encuentra en $[0, 1]$ y λ_2 en $[1, 5]$.

2. Si $\lambda^c < \lambda_1 < 1$ entonces $0 < \rho_{2-} < \rho_{2+}$ y de la ecuación (5.8) se tiene:

$$\begin{cases} \frac{d\rho(t)}{dt} < 0 & \text{para } 0 < \rho(t) < \rho_{2-} \\ \frac{d\rho(t)}{dt} > 0 & \text{para } \rho_{2-} < \rho(t) < \rho_{2+} \end{cases}$$

de aquí se tiene que ρ_1 y ρ_{2+} son soluciones estables mientras que ρ_{2-} no lo es.

Además, cuando hacemos tender $t \rightarrow \infty$ se tiene que $\rho(\infty)$ depende del estado inicial $\rho_0 = \rho(0)$. Si $\rho_0 < \rho_{2-}$ entonces $\rho(\infty) = \rho_1 = 0$; y si ρ_0 es lo suficientemente grande (mayor que ρ_{2-}) entonces $\rho(\infty) = \rho_{2+}$.



(a) Valores de ρ_{2-} en función de λ_1 y λ_2 . (b) Valores de ρ_{2+} en función de λ_1 y λ_2 .

Figura 5.8: Cálculo de las soluciones ρ_{2-} y ρ_{2+} si los valores λ_1 y λ_2 están entre $[1, 5]$.

En resumen:

- Si $\lambda_2 \leq 1$. Si ρ_{2-} es un número real entonces es siempre negativo, por lo que no vamos a considerarlo como solución. Por otro lado, ρ_{2+} es positivo cuando $\lambda_1 > 1$ y negativo para $\lambda_1 < 1$. Entonces en función de los valores de λ_1 se tiene:
 - La única posible solución para $\lambda_2 \leq 1$ y $\lambda_1 < 1$ es $\rho_1 = 0$ y la enfermedad desaparece.
 - Para el caso en que $\lambda_1 > 1$ se tiene que $\rho_{2-} < 0$ y $\rho_{2+} > 0 \Rightarrow d\rho(t)/dt = 0$ cuando $\rho(t)$ es pequeño: el estado de absorción $\rho_1 = 0$ no es estable, sin embargo, ρ_{2+} sí lo es, luego la enfermedad se encuentra en estado endémico.
 - Por último, si $\lambda_1 = 1$, entonces $\rho_{2+} = 0$. Luego la enfermedad se estabiliza en $\rho(\infty) = 0$.

En conclusión, cuando $\lambda_2 \leq 1$, la transición es similar a la del modelo SIS donde $\lambda_2 = 0$. Como podemos observar en la imagen 5.9, donde las líneas rojas es el corte dado por la ecuación (5.8) cuando $\lambda_2 = 0,8$ en la imagen 5.5b y el modelo SIS que coincide cuando $\lambda_2 = 0$. Los cuadrados y círculos azules claros coinciden con simulaciones numéricas realizadas a partir del SCM de orden 2 con $\lambda_2 = 0,8$ y $\lambda_2 = 0$ respectivamente, sobre un complejo simplicial generado a partir del algoritmo RSC de parámetros $N = 2000$ y p_1, p_2 de manera que $\langle \deg_U^{1,1} \rangle = 20$ y $\langle \deg_U^{2,2} \rangle = 6$. En ambos casos las simulaciones numéricas y el modelo coinciden.

- Si $\lambda_2 > 1$. Si $\lambda_1 < \lambda^c = 2\sqrt{\lambda_2} - \lambda_2$ entonces ρ_{2+} y ρ_{2-} no son soluciones reales, por tanto la única solución posible es $\rho_1 = 0$, por ello $\rho(\infty) = 0$ y la enfermedad desaparece.

Por otro lado, cuando $\lambda_1 > \lambda^c$ tenemos dos casos:

- Si $\lambda_1 > 1$ entonces $\rho_{2-} < 0 < \rho_{2+}$ y se tiene que para $\rho(t)$ pequeño $d\rho(t)/dt > 0$ y la densidad de nodos infectados tiende a ρ_{2+} cuando $t \rightarrow \infty$, i.e. $\rho(\infty) = \rho_{2+}$.
- Si $\lambda^c < \lambda_1 < 1$ entonces $0 < \rho_{2-} < \rho_{2+}$ y se tiene que ρ_1 y ρ_{2+} son soluciones estables mientras que ρ_{2-} no lo es. Se tiene que cuando $t \rightarrow \infty$, $\rho(\infty)$ depende del estado inicial $\rho_0 = \rho(0)$. Si $\rho_0 < \rho_{2-}$ entonces $\rho(\infty) = \rho_1 = 0$; y si ρ_0 es lo suficientemente grande entonces $\rho(\infty) = \rho_{2+}$.

Esto puede observarse en la imagen 5.9 donde los puntos azules oscuros de la imagen vienen dados por simulaciones numéricas para $\lambda_2 = 2,5$ sobre un complejo simplicial generado a partir del modelo RSC. En ellas se puede observar que el valor $\rho(\infty) = \rho^*$ es inestable en las simulaciones, depende del $\rho(0)$ inicial.

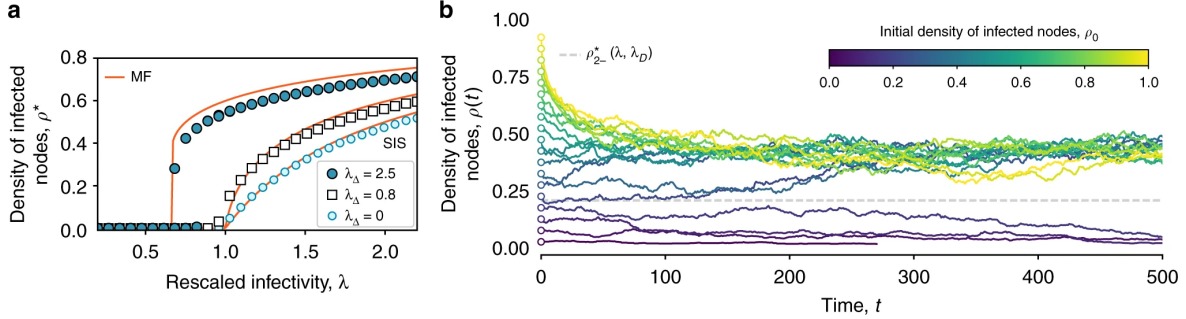


Figura 5.9: Imagen procedente de [14, Figura 3].

Gráficas generadas a partir del SCM de orden 2 sobre un complejo simplicial generado por el RSC de parámetros $N = 2000$ y p_1, p_2 elegidos de tal manera que se tenga $\langle \deg_U^{1,1} \rangle = 20$ y $\langle \deg_U^{2,2} \rangle = 6$.

En **a** está representada la fracción media de nodos infectados, ρ^* , obtenida a partir de simulaciones numéricas, frente al número reproductivo básico $\lambda = \lambda_1$; para $\lambda_2 = 0,8$ (cuadrados blancos), $\lambda_2 = 2,5$ (círculos azules oscuro). Los círculos azules claro nos dan los resultados numéricos de aplicar el modelo SIS (i.e. $\lambda_2 = 0$). Las líneas rojas corresponden a la solución dada por la ecuación (5.8).

Y en la imagen **b** está representada la fracción de infectados $\rho(t)$ en función del tiempo dependiendo del valor inicial $\rho_0 = \rho(0)$, (cada representación corresponde a un valor inicial ρ_0). Además, se ha fijado $\lambda_1 = 0,75$ y $\lambda_2 = 2,5$. La línea discontinua horizontal corresponde a la rama inestable ρ_{2-} , que viene dada por la ecuación (5.7).

Bibliografía

- [1] *Models and Methods in Social Network Analysis*. Structural Analysis in the Social Sciences. Cambridge University Press, 2005.
- [2] Albert, Réka and Barabási, Albert-László. Statistical mechanics of complex networks. *Rev. Mod. Phys.*, 74:47–97, Jan 2002.
- [3] Barabási, A.-L. and Albert, R. and Jeong, H. and Bianconi, G. Power-law distribution of the world wide web. *Science*, 287, 01 2000.
- [4] Barabási, Albert-László. *Network Science*. Cambridge University Press, 2016.
- [5] Barabási, Albert-László and Jeong, Hawoong and Nedá, Z. and Regan, Erzsébet and Schubert, Andras and Vicsek, Tamás. Evolution of the social network of scientific collaboration. *Physica A*, 311:3–4, 01 2002.
- [6] Battiston, Federico and Cencetti, Giulia and Iacopini, Iacopo and Latora, Vito and Lucas, Maxime and Patania, Alice and Young, Jean-Gabriel and Petri, Giovanni. Networks beyond pairwise interactions: Structure and dynamics. *Physics Reports*, 874:1–92, 2020. Networks beyond pairwise interactions: Structure and dynamics.
- [7] Béla Bollobás. *Modern Graph Theory*. Graduate Texts in Mathematics 184. Springer-Verlag New York, 1 edition, 1998.
- [8] Diestel, Reinhard. *Graph Theory (Graduate Texts in Mathematics)*. Springer, August 2005.
- [9] Paul Erdos and Alfred Renyi. On the evolution of random graphs. *Publ. Math. Inst. Hungary. Acad. Sci.*, 5:17–61, 1960.
- [10] Ernesto Estrada and Grant J. Ross. Centralities in simplicial complexes. applications to protein interaction networks. *Journal of Theoretical Biology*, 438:46–60, 2018.
- [11] Goldberg, Timothy E. Combinatorial Laplacians of simplicial complexes. *Senior Thesis, Bard College*, 2002.
- [12] Hernández Serrano, Daniel and Hernández-Serrano, Juan and Sánchez Gómez, Darío. Simplicial degree in complex networks. Applications of topological data analysis to network science. *Chaos, Solitons & Fractals*, 137:109839, 2020.

- [13] Hernández Serrano, Daniel and Sánchez Gómez, Darío. Centrality measures in simplicial complexes: Applications of topological data analysis to network science. *Applied Mathematics and Computation*, 382:125331, 2020.
- [14] Iacopini, Iacopo and Petri, Giovanni and Barrat, Alain and Latora, Vito . Simplicial models of social contagion. *Nature Communications*, 10(1):2485, 2019.
- [15] Hawoong Jeong, B Tombor, R Albert, Zoltan Oltvai, and Albert-Laszlo Barabasi. The large-scale organization of metabolic networks. *Nature*, 407:651–4, 11 2000.
- [16] Munkres, James R. *Elements of Algebraic Topology*. Addison Wesley Publishing Company, 1984.
- [17] Alice Patania, Giovanni Petri, and Francesco Vaccarino. The shape of collaborations. *EPJ Data Science*, 6:18, 08 2017.
- [18] Ann Sizemore, Elisabeth Karuza, Chad Giusti, and Danielle Bassett. Knowledge gaps in the early growth of semantic networks. *Nature Human Behaviour*, 2, 09 2018.
- [19] Ann E Sizemore, Chad Giusti, A. Kahn, J. Vettel, R. Betzel, and D. Bassett. Cliques and cavities in the human connectome. *Journal of Computational Neuroscience*, 44:115 – 145, 2017.
- [20] T. Moore and R. Drost and A. Swami. The Communications and Networks Collaborative Technology Alliance Publication Network: A Case Study on Graph and Simplicial Complex Analysis. 2015.
- [21] Andreas Wagner and David A. Fell. The small world inside large metabolic networks. *Proceedings: Biological Sciences*, 268(1478):1803–1810, 2001.
- [22] West, Douglas B. *Introduction to Graph Theory*. Prentice Hall, 2 edition, September 2000.