



Universidad Autónoma Metropolitana - Iztapalapa
División de Ciencias Básicas e Ingeniería

Modelación de Enfermedades Infecciosas con Información Geográfica

Tesis que presenta
Luis Alberto Zarate Siordia
Para obtener el grado de
Maestro en Ciencias
(Matemáticas Aplicadas e Industriales)

Asesores: Dr. Joaquín Delgado Fernández
Dr. Jorge X. Velasco Hernández

México D.F. Marzo 2012

A Jennifer

Agradecimientos

Agradezco a mi esposa por darme su apoyo y paciencia, a mi familia por el apoyo brindado para la realización de este proyecto, a CONACYT por brindarme los recursos financieros, a la UAM-I por abrirme sus puertas, a mis profesores por formarme, a mis asesores Joaquín Delgado y Jorge Velasco por darme su tiempo y apoyo, a mis sinodales Marcos Capistrán y Baltazar Aguirre por sus valiosos comentarios que ayudaron a mejorar el trabajo, a mis amigos que hicieron mas ameno mi paso por la UAM-I y a todos los que trabajaron directa e indirectamente para ayudarme a concluir este proyecto.

Índice general

Dedicatoria	III
Agradecimientos	V
Lista de figuras	IX
Introducción	1
1. Modelos epidemiológicos	3
1.1. Preliminares	3
1.2. Número reproductivo básico	4
1.3. El modelo epidémico SIR clásico	5
1.4. El modelo endémico SIR clásico	8
1.5. El modelo SIR(S) normalizado	10
1.5.1. Adaptación espacial del modelo SIR	11
1.5.2. Poblaciones vecinas infecciosas para el modelo SIR	12
2. Modelador espacio-temporal de epidemias STEM	13
2.1. Preliminares	13
2.2. Modelación en STEM	15
2.2.1. Ejemplo del modelo SIR(S) utilizando STEM	15
3. Estimación de parámetros para el modelo SIR endémico	21
3.1. Estimación de parámetros por máxima verosimilitud	21
3.2. Estimación de parámetros utilizando ji cuadrada χ^2	24
3.3. Solución al problema de minimización	25
3.3.1. Método de gradiente conjugado no-lineal GC	25
3.3.2. Sobrerelajación sucesiva SOR	26
3.3.3. Regularización de Tikhonov	26
3.3.4. Intervalo de confianza para los parámetros estimados	27
3.4. Implementación computacional	28
3.4.1. Implementación computacional del método de máxima verosimilitud	28
3.4.2. Implementación computacional de la distribución Ji-cuadrada χ^2	29

3.4.3. Resultados de las simulaciones	29
3.5. Estimación de parámetros para Bélgica y Francia	37
4. Estimación de parámetros para el modelo SIR epidémico	39
4.1. Estimación de parámetros utilizando evolución diferencial	39
4.2. Implementación computacional	41
4.2.1. Resultados de la simulación	42
5. Adaptación espacial del modelo	45
5.1. Adaptación espacial del modelo	46
5.2. Poblaciones vecinas infecciosas	47
5.2.1. Ejemplo de la adaptación espacial para una estructura de tres vecinos	48
5.3. Adaptación espacial para la estructura de vecinos de la Zona Metropoli- tana de Guadalajara	50
5.4. Implementación computacional de la estructura de vecinos de la Zona Metropolitana de Guadalajara	53
Conclusiones	61
A. Estimación de parámetros para un modelo epidemiológico SIRQ	63
A.1. Modelo SIQR	63
A.2. Resultados	64
A.3. Conclusiones del apéndice	66
Bibliografía	69

Índice de figuras

1.1.	Solución del modelo SIR epidémico clásico con $\sigma = 3$ y $1/\gamma = 3$ días	7
1.2.	Plano fase para el modelo SIR epidémico clásico con $\sigma = 3$	7
1.3.	Plano fase para el modelo SIR endémico clásico con $R_0 = \sigma = 0.05$	9
1.4.	Plano fase para el modelo SIR endémico clásico con $R_0 = \sigma = 3$	10
2.1.	Red Aérea de México	16
2.2.	Red Carretera de México	16
2.3.	Gráfica de incidencia para el modelo SIR	17
2.4.	Serie de tiempo para el modelo SIR	17
2.5.	Propagación en 40 días con el modelo SIR	18
2.6.	Propagación en 207 días con el modelo SIR	18
2.7.	Propagación en 494 días con el modelo SIR	19
3.1.	Infectados en Jalisco	30
3.2.	Propagación durante 79 semanas utilizando máxima verosimilitud y GC no-lineal	31
3.3.	Propagación durante 79 semanas utilizando máxima verosimilitud, GC no-lineal y μ fija	32
3.4.	Propagación en las primeras 6 semanas utilizando máxima verosimilitud, GC no-lineal y μ fija	33
3.5.	Propagación de las 79 semanas con los parámetros de las primeras 6 utilizando máxima verosimilitud, GC no-lineal y μ fija	33
3.6.	Propagación en las primeras 6 semanas utilizando el método SOR y μ fija	34
3.7.	Propagación durante 79 semanas con los parámetros de 6, utilizando el método SOR	35
3.8.	Propagación en las primeras 6 semanas utilizando máxima verosimilitud, GC no-lineal y μ fija	36
3.9.	Propagación de las 79 semanas con los parámetros de las primeras 6 utilizando máxima verosimilitud, GC no-lineal y μ fija	36
3.10.	Propagación para 91 semanas para Bélgica con todos los parámetros utilizando máxima verosimilitud, GC no-lineal	37
3.11.	Propagación para 101 semanas para Francia con todos los parámetros utilizando máxima verosimilitud, GC no-lineal	38

4.1. Infectados en Jalisco	42
4.2. Infectados en EUA	43
5.1. Zona Metropolitana de Guadalajara	50
5.2. Ambiente Gráfico	55
5.3. Propagación de la enfermedad para la clase susceptible de la semana 10 a la 80	56
5.4. Propagación de la enfermedad para la clase infecciosa de la semana 10 a la 80	57
5.5. Propagación de la enfermedad para la clase de removidos de la semana 10 a la 80	58
A.1. Propagación de la semana 6 a la 10 para el modelo SIQR	64
A.2. Propagación de la semana 6 a la 47 para el modelo SIQR	65
A.3. Propagación de la semana 10 a la 22 para el modelo SIQR	65
A.4. Propagación de la semana 22 a la 47 para el modelo SIQR	66
A.5. Propagación de la semana 1 a la 47 para el modelo SIR/SIQR	66

Introducción

Debido al brote del virus de la influenza A(H1N1) en México tuvimos la inquietud de implementar un modelo que describiera la propagación de la enfermedad. Nos concentramos en la propagación en el estado de Jalisco. En la literatura encontramos que un sistema de ecuaciones diferenciales puede describir procesos epidemiológicos por lo que decidimos proponer que la propagación del virus de la influenza A(H1N1) puede ser descrita con un modelo SIR debido a la naturaleza de la enfermedad.

El modelo SIR divide a la población en diferentes clases de acuerdo al estado de la enfermedad, estas son, cuando la población se encuentra susceptible a enfermarse, cuando se encuentran en estado infeccioso y por último cuando se han recuperado de la enfermedad, es decir, fueron removidos de la clase de infectados.

Una vez con el modelo, buscamos estimar los parámetros del sistema considerando el número reproductivo básico de la infección, que es el promedio de infecciones secundarias que ocurren cuando un individuo infectado se introduce en una población que es completamente susceptible a la enfermedad. Utilizamos los métodos de estimación de parámetros en sistemas dinámicos por máxima verosimilitud, el de la distribución Ji-cuadrada χ^2 y el de evolución diferencial, tomando como referencia los datos proporcionados por la Secretaria de Salud Jalisco (SSJ) y el Consejo Estatal de Población del estado de Jalisco (COEPO) en el período de abril del 2009 a agosto del 2010.

Utilizamos una adaptación espacial del modelo para implementar datos geográficos y estadísticos de la población en una estructura de vecinos, debido a que una persona no se encuentra en un lugar en específico, si no que tiene que trasladarse a centros de estudio o de trabajo, que en ocasiones se encuentran en un lugar diferente al lugar en donde se vive. Propusimos modelar el brote en la Zona Metropolitana de Guadalajara, que esta dividida en 8 municipios y por lo regular las personas viajan entre estos para realizar sus actividades diarias.

Además vamos a usar una aplicación de código abierto que contiene modelos espaciales y temporales de enfermedades infecciosas STEM, este programa contiene datos de la Republica Mexicana y lo utilizamos con parámetros que nos muestren propagaciones muy marcadas a lo largo del territorio nacional.

Capítulo 1

Modelos epidemiológicos

Los modelos epidemiológicos nos permiten describir la propagación de enfermedades que atacan a una población, en este trabajo nos concentraremos en los que se utiliza un sistema de ecuaciones diferenciales, donde cada ecuación nos proporciona información sobre las diferentes etapas que tiene la enfermedad y su distribución en la población.

Los modelos epidemiológicos son utilizados para describir brotes repentinos de enfermedades que infectan a una porción sustancial de la población, que por lo regular se dan en períodos cortos de tiempo, mientras que los modelos endémicos se utilizan para describir enfermedades que se encuentran permanentemente en la población, es decir, cuando hay persistencia de la enfermedad y por lo regular se estudian por periodos largos de tiempo [19].

En este capítulo vamos a presentar como se formulan los modelos SI y SIR, vamos a dar una explicación mas detallada para el modelo SIR, debido a que es nuestra propuesta para explicar la propagación del virus de la influenza A(H1N1) y el desarrollo de este trabajo se basa en su teoría.

1.1. Preliminares

Para formular el modelo necesitamos hacer una serie de suposiciones que nos permitan establecer las bases teóricas para así poder describir la propagación de la enfermedad que vamos a estudiar. Las suposiciones para el modelo son las siguientes;

La población se divide en clases disjuntas que cambian conforme pasa el tiempo, al que vamos a denotar con t . La clase susceptible consiste de todos aquellos individuos que pueden adquirir la enfermedad pero no están infectados. La clase de infectados consiste de todos los individuos que adquirieron la enfermedad y a su vez, pueden transmitirla a otros que están en la clase de los susceptibles. La clase de removidos o recuperados consiste de todos los individuos que son removidos de la fracción

susceptible-infecciosa ya sea por muerte, inmunidad o aislamiento. Estas clases se denotan por $S(t)$, $I(t)$ y $R(t)$ respectivamente, además, si denotamos por N al número total de individuos en la población, tenemos la siguiente relación,

$$N = S + I + R$$

Consideremos una población suficientemente grande para que cada clase se pueda considerar como una variable continua. Esta población se mezcla de forma homogénea, es decir, los individuos tienen la misma probabilidad de encontrarse e infectarse. Si el modelo incluye dinámica vital, es decir, ocurren nacimientos y muertes en el periodo de estudio, suponemos que los nacimientos y las muertes por causa natural ocurren a una misma tasa, de tal manera que todos los que nacen son susceptibles. Los individuos son removidos de cada clase a una tasa que es proporcional al tamaño de la clase con constante de proporcionalidad μ la cual llamaremos tasa de mortalidad, que corresponde al inverso proporcional del promedio de vida, es decir, la esperanza de vida de la población es igual a $\frac{1}{\mu}$.

La tasa de contacto diario que denotamos por β es el promedio de contactos adecuados de individuos infecciosos por cada unidad de tiempo. Un contacto adecuado de un individuo infeccioso es una interacción que da como resultado la infección de un individuo que es susceptible. El tipo contacto adecuado directo o indirecto depende de una enfermedad específica.

Los individuos se recuperan y son removidos de la clase de infectados a una tasa proporcional al número de infectados, con una tasa de proporcionalidad constante γ , que llamaremos tasa de recuperación por unidad de tiempo. El período latente es cero y se define como el período entre el tiempo de exposición, y el tiempo en que comienza la infecciosidad. El promedio del período infeccioso es $\frac{1}{\gamma}$.

Antes de la infección todos los individuos de la población se encuentran en la clase susceptible S . Si $S(t)$ es el número de susceptibles al tiempo t , $I(t)$ es el número de infectados y N es el total de la población, entonces $s(t) = S(t)/N$ e $i(t) = I(t)/N$ son las fracciones de susceptibles e infectados respectivamente. Si β es el número promedio de contactos adecuados o tasa de transmisión de la enfermedad, es decir, contactos suficientes para la transmisión de una persona por unidad de tiempo, entonces $\beta I/N = \beta i$ es el número promedio de contactos infectados por unidad de tiempo de un susceptible, y $(\beta I/N)S = \beta N i s$ es el número de casos nuevos por unidad de tiempo debido a $S = N s$ susceptibles.

1.2. Número reproductivo básico

El número reproductivo básico o tasa de reproducción básica de la infección R_0 es el promedio de infecciones secundarias que ocurren cuando un individuo infectado se

introduce en una población totalmente susceptible. Cuando el R_0 es mayor que uno, significa que en promedio cada individuo infeccioso, al inicio del proceso de contagio, infecta a más de un nuevo individuo y entonces se producirá un brote epidémico. En caso de que R_0 sea menor que uno, cada individuo infectara en promedio a menos de una persona, y por lo tanto no habrá brote epidémico y el virus desaparecerá de la población.

El número de contacto σ se define como el promedio de contactos adecuados de la infección típica durante el período infeccioso. Un contacto adecuado es aquel que es suficiente para la transmisión, es decir, si el individuo en la clase de susceptibles esta ahora en la clase de infectado.

El número de reemplazamiento R se define como el promedio de infecciones secundarias producidas por infección típica durante todo el período de infecciosidad. Notemos que R_0 , σ y R son iguales al inicio de la propagación de la enfermedad cuando toda la población es susceptible excepto el invasor infectado, debido a que el promedio de infecciones secundarias y el promedio de contactos adecuados es el mismo al tiempo que ingresa el invasor.

En la literatura reciente de los modelos epidemiológicos R_0 es frecuentemente utilizada como la cantidad que determina si una enfermedad puede invadir a la población [19]. R_0 se define solamente al principio de la invasión, σ y R se definen durante todo el tiempo. Para muchos modelos, el número de contacto σ permanece constante durante toda la propagación y es igual siempre al número de reproducción básica R_0 , en las secciones siguientes daremos una explicación mas detallada de como se relaciona el número de contacto, con el estado libre de enfermedad y el estado de equilibrio endémico. El número de reemplazamiento R es el número actual de casos secundarios de una infección típica, es decir, después de que la infección ha invadido a la población y cada individuo no es susceptible, R es siempre menor que el número de reproducción básica R_0 . También después de la invasión, la fracción de susceptibles es menor que uno, es decir, que no todos los contactos resultan en un nuevo caso. Así, el número de reemplazamiento R es siempre menor que el número de contacto σ después de la invasión. Combinando estos resultados, tenemos que

$$R_0 \geq \sigma \geq R$$

con las tres cantidades iguales al tiempo de la invasión [19]. Note que $R_0 = \sigma$ para muchos modelos, y $\sigma > R$ después de la invasión para todos los modelos.

1.3. El modelo epidémico SIR clásico

En el modelo epidémico SIR clásico consideramos un período de tiempo muy corto, por lo que no se incluyen las muertes y nacimientos en la población, es decir no

incluimos dinámica vital, consideremos entonces el problema de valor inicial

$$\frac{dS}{dt} = -\frac{\beta SI}{N} \quad (1.1a)$$

$$\frac{dI}{dt} = \frac{\beta SI}{N} - \gamma I \quad (1.1b)$$

$$\frac{dR}{dt} = \gamma I \quad (1.1c)$$

al inicio de la infección tenemos $S(0) = S_0 \geq 0$, $I(0) = I_0 \geq 0$ y $R(0) = R_0 \geq 0$.

Este modelo utiliza la incidencia estándar y tiene tasa de recuperación γI .

Si dividimos las ecuaciones (1.1) por la población total N , es decir, tomamos las fracciones $s = S/N$, $i = I/N$ y $r = R/N$ del modelo, obtenemos

$$\frac{ds}{dt} = -\beta si \quad (1.2a)$$

$$\frac{di}{dt} = \beta si - \gamma i \quad (1.2b)$$

con $r(t) = 1 - s(t) - i(t)$ donde $s(t)$, $i(t)$, $r(t)$ son fracciones de las clases, el triángulo T en el plano si se define como,

$$T = \{(s, i) | s \geq 0, i \geq 0, s + i \leq 1\} \quad (1.3)$$

es positivamente invariante y existe una solución de tal manera que el modelo esta matemáticamente y epidemiológicamente bien planteado[19]. Aquí el número de contacto $\sigma = \frac{\beta}{\gamma}$ se define como la tasa de contacto β por unidad de tiempo, multiplicada por el promedio del período infeccioso $\frac{1}{\gamma}$ y se interpreta como el promedio de contactos adecuados de los infectados durante el período de infección. El número de reemplazamiento al tiempo de inicio de la epidemia es σs_0 , que es el producto del número de contacto σ y la fracción inicial de susceptibles s_0 .

Teorema 1. *Sea $(s(t), i(t))$ una solución de (1.2) en T . Si $\sigma s_0 \leq 1$, entonces $i(t)$ decrece a cero cuando $t \rightarrow \infty$. Si $\sigma s_0 > 1$, entonces $i(t)$ primero se incrementa hasta llegar a un valor máximo $i_{max} = i_0 + s_0 - \frac{1}{\sigma} - \frac{\ln(\sigma s_0)}{\sigma}$ y después decrece a cero cuando $t \rightarrow \infty$. La fracción susceptible $s(t)$ es una función decreciente y el límite s_∞ es la única raíz en $(0, \frac{1}{\sigma})$ de la ecuación*

$$i_0 + s_0 - s_\infty + \frac{\ln(s_\infty/s_0)}{\sigma} = 0 \quad (1.4)$$

Una de las características de un brote epidémico, es la curva de individuos infectados que se incrementa desde un punto de inicio i_0 cercano a cero, llega a un punto máximo y empieza a descender y se va acercando mucho al cero conforme pasa el tiempo, figura(1.1). La fracción susceptible $s(t)$ siempre decrece, pero al final, la fracción susceptible s_∞ será positiva. La epidemia tiende a desaparecer, porque cuando

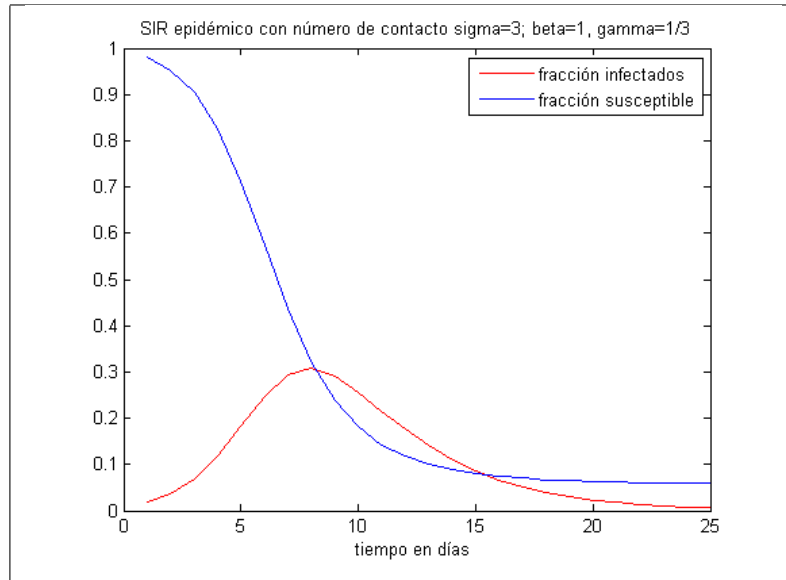


Figura 1.1: Solución del modelo SIR epidémico clásico con $\sigma = 3$ y $1/\gamma = 3$ días

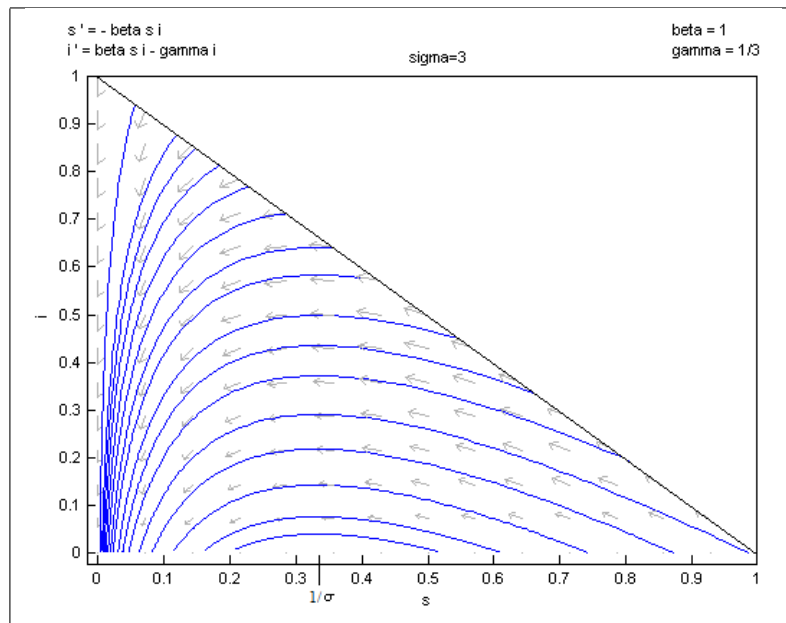


Figura 1.2: Plano fase para el modelo SIR epidémico clásico con $\sigma = 3$

la fracción susceptible $s(t)$ es menor que $1/\sigma$, el número de reemplazamiento $\sigma s(t)$ es menor que 1.

En la figura (1.2) los puntos de equilibrio alrededor del eje s son neutros inestables para $s > 1/\sigma$ y son neutros estables para $s < 1/\sigma$. Una aproximación clásica [18] dice que para un i_0 pequeño y s_0 ligeramente más grande que $s_{max} = 1/\sigma$, la diferencia $s_{max} - s(\infty)$ es igual a $s_0 - s_{max}$, donde s_{max} es el valor máximo que alcanza la fracción susceptible, observe la figura (1.1). Podemos ver que los resultados principales, envuelven el número de reemplazamiento σs_0 , aunque se pueden expresar en términos del número de reproducción básica R_0 .

1.4. El modelo endémico SIR clásico

El modelo SIR endémico clásico a diferencia del modelo epidémico, incluye dinámica vital, es decir, contempla las muertes y los nacimientos.

$$\frac{dS}{dt} = \mu N - \frac{\beta SI}{N} - \mu S \quad (1.5a)$$

$$\frac{dI}{dt} = \frac{\beta SI}{N} - \gamma I - \mu I \quad (1.5b)$$

$$\frac{dR}{dt} = \gamma I - \mu R \quad (1.5c)$$

al inicio de la infección tenemos $S(0) = S_0 \geq 0$, $I(0) = I_0 \geq 0$ y $R(0) = R_0 \geq 0$.

Si dividimos las ecuaciones (1.5) por la población total N , al igual que en el modelo epidémico, obtenemos

$$\frac{ds}{dt} = \mu - \beta si - \mu s \quad (1.6a)$$

$$\frac{di}{dt} = \beta si - (\gamma + \mu)i \quad (1.6b)$$

con $r(t) = 1 - s(t) - i(t)$ donde $s(t)$, $i(t)$, $r(t)$ son fracciones de las clases, el triángulo T en el plano si se define como,

$$T = \{(s, i) | s \geq 0, i \geq 0, s + i \leq 1\} \quad (1.7)$$

es positivamente invariante y existe una solución de tal manera que el modelo esta matemáticamente y epidemiológicamente bien planteado [19]. El número de contacto σ es igual al número reproductivo básico de la epidemia, para todo el período de tiempo, debido a que no ocurren nuevas clases de susceptibles o infecciosos después de la invasión, para este modelo la cantidad principal es

$$R_0 = \sigma = \frac{\beta}{\gamma + \mu} \quad (1.8)$$

la cual es la tasa de contacto β por unidad de tiempo, multiplicada por el número de muertes ajustadas al período de infección $\frac{1}{\gamma + \mu}$

Teorema 2. Sean $(s(t), i(t))$ una solución de (1.6) en T . Si $\sigma \leq 1$ o $i_0 = 0$ entonces las trayectorias de la solución empiezan en T y convergen a un estado de equilibrio libre de enfermedad, dado por, $s = 1$ e $i = 0$. Si $\sigma > 1$ entonces todas las trayectorias de la solución con $i_0 > 0$ convergen a un estado de equilibrio endémico, dado por, $s_e = \frac{1}{\sigma}$ e $i_e = \frac{\mu(\sigma-1)}{\beta}$.

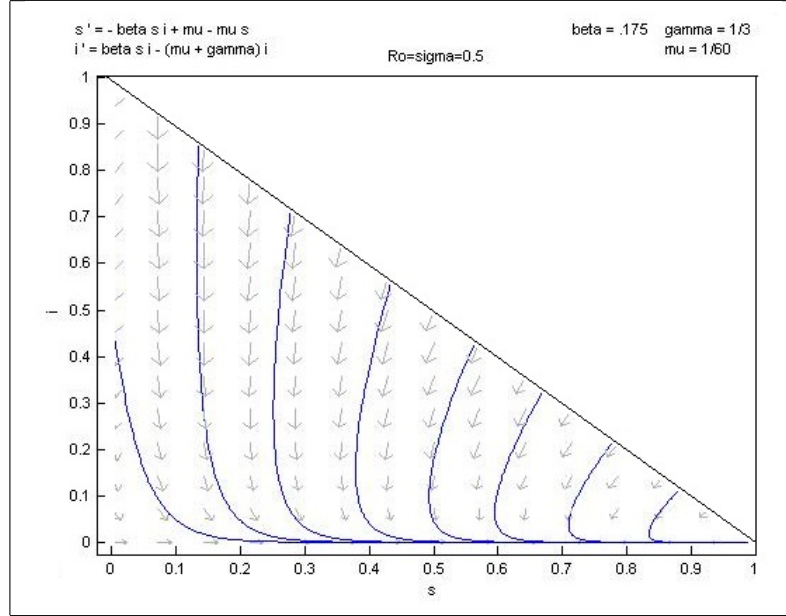


Figura 1.3: Plano fase para el modelo SIR endémico clásico con $R_0 = \sigma = 0.05$

Si $R_0 = \sigma \leq 1$ entonces el número de reemplazo σs es menor que 1 cuando $i_e > 0$, y la infección decrece hasta llegar a cero. Además la velocidad del movimiento alrededor de las trayectorias no es aparente figura(1.3), la fracción infecciosa decrece muy rápidamente muy cerca del cero y entonces como para 100 años o mas, la población recuperada muere muy lentamente y el proceso de nacimientos incrementa muy lentamente a los susceptibles, hasta que eventualmente cada uno es susceptible en el estado de equilibrio libre de enfermedad con $s = 1$ e $i = 0$.

Si $R_0 = \sigma > 1$, i_0 es pequeño y s_0 es grande con $\sigma s_0 > 1$ entonces $s(t)$ decrece e $i(t)$ se incrementa hasta llegar a un punto máximo y entonces empieza a decrecer, justo como podría pasar en una epidemia, observe la figura(1.4). Sin embargo, después de que la fracción infecciosa decrece a un nivel muy bajo, el lento proceso de muerte de gente recuperada y los nacimientos de nuevos individuos susceptibles (alrededor de 10 a 20 años) gradualmente incrementa la fracción susceptible hasta que $\sigma s(t)$ es suficientemente grande para que otra pequeña epidemia ocurra. Este proceso de rápida alternancia epidémica y lenta recuperación de susceptibles continúa, como las trayectorias se acercan al estado de equilibrio endémico descrito en el teorema 2. En este equilibrio endémico el número de reemplazamiento σs_e es 1, lo cual es plausible

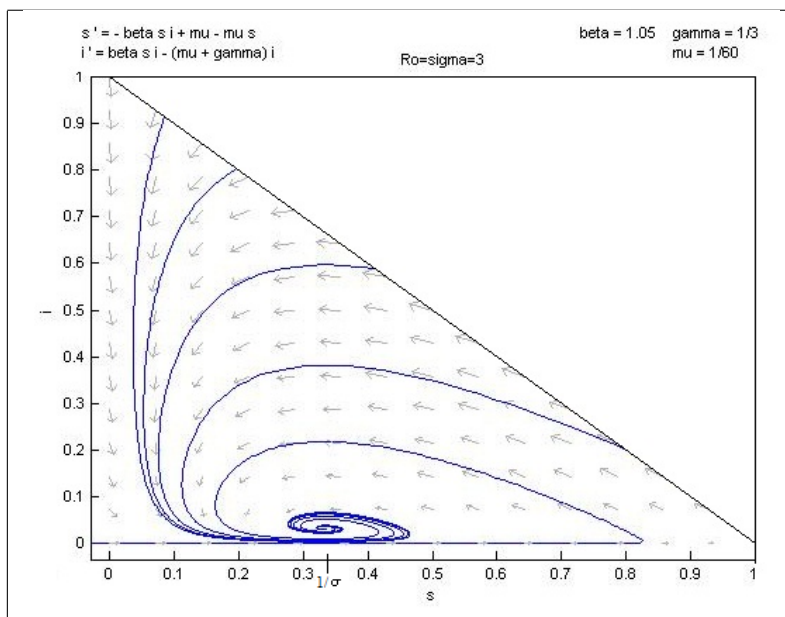


Figura 1.4: Plano fase para el modelo SIR endémico clásico con $R_0 = \sigma = 3$

si el número de reemplazamiento es mayor o menor a uno, la fracción infecciosa $i(t)$ puede incrementarse o disminuir respectivamente.

1.5. El modelo SIR(S) normalizado

Los modelos SIR (susceptible, infeccioso, recuperado) y SIRS (susceptible, infeccioso, recuperado, susceptible), para una población normalizada, se definen por el sistema de ecuaciones diferenciales ordinarias [35]

$$\frac{ds}{dt} = \mu - \beta si + \delta r - \mu s \quad (1.9a)$$

$$\frac{di}{dt} = \beta si - \gamma i - \mu i \quad (1.9b)$$

$$\frac{dr}{dt} = \gamma i - \delta r - \mu r \quad (1.9c)$$

donde

- s es la proporción de la población que es susceptible
- i es la proporción de la población que es infecciosa
- r es la proporción de la población que esta recuperada de la infección, ya no es susceptible y por tanto no puede ser infectada

- μ es tanto la tasa de inmigración (i.e. nacimiento) y emigración (i.e. muerte) de la población, y se suponen constantes durante el período de tiempo de interés
- β es la tasa de transmisión de la enfermedad. Es la tasa con la que los individuos infecciosos infectan a los individuos susceptibles. Una vez infectados, los individuos susceptibles entran a ser expuestos.
- γ es la tasa a la cual los individuos se curan. En este modelo los individuos curados no podrán ser reinfectados por un tiempo después de la infección (siendo removidos de la población o adquiriendo inmunidad)
- δ es la tasa de pérdida de la inmunidad. Este coeficiente determina la tasa a la cual los miembros recuperados de la población pierden su inmunidad a la enfermedad y son susceptibles de nuevo. Para un modelo SIR esta tasa es 0.

Sea μ_i la tasa de mortalidad infecciosa, es decir, la tasa a la cual la población infectada muere específicamente de la enfermedad, si incluimos esta tasa al modelo, obtenemos

$$\frac{ds}{dt} = \mu - \beta si + \delta r - \mu s \quad (1.10a)$$

$$\frac{di}{dt} = \beta si - \gamma i - (\mu + \mu_i)i \quad (1.10b)$$

$$\frac{dr}{dt} = \gamma i - \delta r - \mu r \quad (1.10c)$$

donde $(\mu + \mu_i)$ es la tasa neta de mortalidad y $p = s + i + r$ es la proporción de población total.

1.5.1. Adaptación espacial del modelo SIR

Tomemos una adaptación espacial del modelo que se desarrolla con detalle en el capítulo 5.

El modelo SIR se puede adaptar a poblaciones espacialmente distribuidas. Esto relaja la hipótesis de que las poblaciones están en un solo lugar y abre la posibilidad de que diferentes lugares tengan diferentes tamaño y número de habitantes, es decir, diferente densidad de población .

Para implementar esta situación se utilizan versiones no normalizadas del sistema (1.10). Sea P_l el número de individuos de la población en el lugar l y $P_l = S_l + I_l + R_l$. Además, se necesita tomar en cuenta la variabilidad de la tasa de transmisión β debido a las posibles diferencias en la densidad de población, por lo que se supone una tasa específica que se denota por β_l .

$$\frac{dS_l}{dt} = \mu P_l - (\beta_l/P_l)S_l I_l + \delta R_l - \mu S_l \quad (1.11a)$$

$$\frac{dI_l}{dt} = (\beta_l/P_l)S_l I_l - \gamma I_l - (\mu + \mu_i)I_l \quad (1.11b)$$

$$\frac{dR_l}{dt} = \gamma I_l - \delta R_l - \mu R_l \quad (1.11c)$$

1.5.2. Poblaciones vecinas infecciosas para el modelo SIR

La adyacencia física de las poblaciones conduce naturalmente al contacto de población a población y eventualmente a la diseminación de la enfermedad. Podemos extender el modelo incorporando este aspecto de una población espacialmente distribuida como sigue:

Dos lugares están relacionados positivamente cuando es posible el intercambio de miembros de las poblaciones mediante algún sistema de transporte como caminos, carreteras, trenes o aviones. En este caso es posible que los miembros de las poblaciones de infecciosos visiten a miembros susceptibles de otras poblaciones. Se supone que los visitantes tienen el mismo nivel de contacto infeccioso que los miembros de la población visitada, es decir, que se pueden contar como miembros de la población en el lugar en cuestión.

Tomamos el siguiente sistema de ecuaciones, que se describe a detalle en el capítulo 5.

$$\frac{dS_l}{dt} = \mu P_l - \beta_l S_l I_{vecinos} + \delta R_l - \mu S_l \quad (1.12a)$$

$$\frac{dI_l}{dt} = \beta_l S_l I_{vecinos} - (\gamma + \mu)I_l \quad (1.12b)$$

$$\frac{dR_l}{dt} = \gamma I_l - \delta R_l - \mu R_l \quad (1.12c)$$

donde

- $I_{vecinos} = \frac{\sum_{k=1}^K m_{lk} I_k}{\sum_{k=1}^K m_{lk} P_k}$ es el número de visitantes infecciosos del vecino k .
- m_{lk} es una constante que depende del tipo de relación positiva; si las poblaciones son vecinas, i.e. comparten una frontera, si están conectados por carreteras, entre otros.
- P_k es el número de individuos en la población vecina k .
- I_k es el número de individuos infecciosos de la población vecina k .

Capítulo 2

Modelador espacio-temporal de epidemias STEM

STEM (Modelador Espacio-temporal de Epidemias) es una aplicación de código abierto creado en la plataforma ECLIPSE. Es una infraestructura y herramienta de desarrollo para la composición de modelos espaciales y temporales de enfermedades infecciosas. Utiliza la arquitectura de software de componentes más recientes basadas en un estándar llamado la iniciativa OSGI, la plataforma Equinox de Eclipse es la implementación de referencia del estándar. Al utilizar la arquitectura de software de componentes STEM permite la composición de modelos, todos los componentes de STEM, incluidos los códigos y datos son reutilizables y están disponibles como componentes y paquetes. STEM viene con un gran número de complementos como son, los datos globales, geográficos, poblacionales, demográficos, información sobre transporte, como caminos y rutas aéreas y modelos básicos de enfermedades.[35]

Esta arquitectura es compatible con la colaboración, ya que los usuarios pueden no solo crear nuevos modelos y componer nuevos escenarios, si no también intercambiarlos como componentes reutilizables y por lo tanto basarse en el trabajo de los demás.

Uno de los objetivos principales al crear STEM es facilitar la creación y utilización de modelos espacio temporales de enfermedades infecciosas que ayuden a entender y prevenir la dispersión de la enfermedad.

2.1. Preliminares

Supongamos que una enfermedad se esta propagando por algún país, por ejemplo México, uno puede estimar cuanto tiempo después de que aparezcan 10 casos en la ciudad de Monterrey, se observará el pico del brote ocurrido en la ciudad de Guadalajara, así como el número de casos probables que ocurrirán durante el brote, para responder estas preguntas se requiere un modelo predictivo.

Modelar la propagación de una enfermedad en una población, es en muchos aspec-

tos similar al modelado del clima, sin embargo a diferencia del clima, la política de salud pública puede cambiar el resultado de una epidemia, por ejemplo, como cambiara el número total de casos, si implementamos políticas de distanciamiento social, añadimos vacunas o cerramos escuelas.

Los datos requeridos para un modelo de enfermedad infecciosa son llamados, datos denominadores [35], el tipo de datos denominadores que se necesitan depende del tipo de enfermedad que se quiere modelar, para nuestro ejemplo, ciertamente necesitamos saber el número de personas que viven en la ciudad de Monterrey y en la ciudad de Guadalajara, también como las personas se desplazan entre estas ciudades incluyendo no solo los transportes terrestres, sino también el transporte aéreo.

Algunas de las características de la enfermedad a modelar son;

- La velocidad de transmisión o tasa de reproducción básica del virus,
- El tiempo en promedio será infecciosa una persona, (esto se llama tasa de recuperación),
- El tiempo en promedio son inmunes las personas una vez que se recuperan, (esta es la tasa de pérdida de la inmunidad),
- La tasa de mortandad de la enfermedad,
- Que fracción de la población se espera sea resistente o inmune a este tipo de virus,

Estas características se conocen como los parámetros básicos de la enfermedad, los cuales se utilizan como coeficientes en los modelos matemáticos de la enfermedad misma. Los usuarios pueden crear sus propios modelos matemáticos, o elegir algunos modelos que se encuentran disponibles como complementos en STEM, uno de los modelos es el SIR que es un modelo compartamental.

Un modelo compartamental divide la población o poblaciones en estudio en varios estados de la enfermedad o compartimientos diferentes. Podemos pensar en cada compartimiento o estado, como en un cuarto de una casa, una persona puede moverse entre los diferentes cuartos, pero solo puede existir en un cuarto a la vez, la población completa se distribuirá entre los cuartos, o compartimientos definidos por el modelo, la población total puede aumentar o disminuir según las tasas de nacimientos y defunciones.

En un modelo SIR la personas susceptibles dejan el estado susceptible, interactuando con personas infecciosas, esta interacción define el proceso de infección. Se requieren modelos matemáticos diferentes para modelar una enfermedad infecciosa, debido a que cada enfermedad tiene un comportamiento diferente, por ejemplo, para infecciones que confieren inmunidad de por vida en el estado recuperado (sarampión, paperas) un modelo SIR es apropiado, debido a que, las personas nunca dejan el estado recuperado mientras vivan.

Los modelos SIRS (involucran virus de influenza) debido a que la inmunidad no es de por vida y la inmunidad disminuye con el tiempo. Algunas enfermedades infecciosas se caracterizan por un período de incubación entre la exposición a un patógeno y el desarrollo de síntomas clínicos, si el individuo expuesto no es portador del virus durante este período de incubación, puede ser importante modelar el tiempo de incubación explícitamente, por ejemplo, en el virus de la viruela es importante modelar este estado expuesto, ya que existe un retraso de 7 a 14 días entre el momento en que el individuo se infecta al momento en que llega a ser infeccioso, así pues, el modelado de enfermedades requiere de muchos modelos, también de una gran cantidad de datos denominadores.

2.2. Modelación en STEM

2.2.1. Ejemplo del modelo SIR(S) utilizando STEM

El modelo SIR(S) es útil para modelar enfermedades con un período corto de tiempo, entre que un individuo es infectado y comienza a ser infeccioso, cuando el individuo se recupera, adquiere inmunidad definitiva después de la infección (modelo SIR) o inmunidad temporal (modelo SIRS). Ejemplos de enfermedades que siguen estos modelos son la varicela y la influenza. Los individuos susceptibles son infectados con una tasa de βI y en un tiempo pierden infecciosidad en una tasa γ y llegan a ser inmunes, la tasa de pérdida de la inmunidad se denota como δ .

Daremos un ejemplo ilustrativo de como trabaja STEM:

- I. Diseñamos un nuevo proyecto al que llamaremos SIR.

- II. Ahora creamos un nuevo modelo al cual llamaremos ModeloSIR, y aquí incluiremos un modelo dado en STEM llamado MEX Human Population (Levels:0,1), donde el nivel 0 representa la Republica Mexicana y el 1 la subdivisión de los Estados, también incluiremos las redes de transporte de las principales autopistas entre los estados y las redes aéreas principales entre los estados, nombradas en STEM como Air Transport between MEX(0) and MEX(1) y Road Transportation Network (2001) For MEX(1).



Figura 2.1: Red Aérea de México



Figura 2.2: Red Carretera de México

- III. Definimos una enfermedad a la cual llamaremos EnfermedadSIR, en la cual utilizaremos el modelo determinístico que viene dado en STEM llamada DeterministicSIRDiseaseModelImpl buscando su solución mediante Finite Difference, ahora definimos los parámetros de la enfermedad, primero que afecte a los humanos, la tasa de transmisión $\beta = 1.5$, la tasa de recuperación de la infección $\gamma = 1$ y la tasa de mortalidad $\mu_i = 0.003$ y para nuestra tasa de perdida de inmunidad $\delta = 0.01$
- IV. Ahora determinemos el período de tiempo en que modelaremos la enfermedad en STEM lo llaman Sequencer y la fecha es del 4 de mayo del 2009 al 4 de mayo del 2011
- v. Ahora definiremos un escenario llamado MexicoSIR

vi. Diseñemos el infector principal el cual iniciará con un número de 10 personas infectadas en el DF

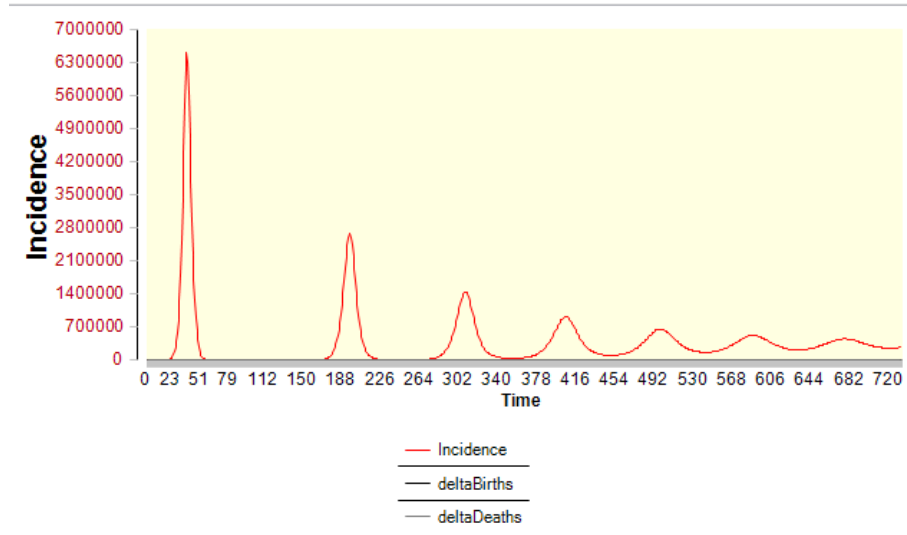


Figura 2.3: Gráfica de incidencia para el modelo SIR

Esta es la gráfica de incidencia, figura (2.3), muestra como se propaga la enfermedad, primero infecta a un número alto de susceptibles, luego baja el número de infectados, y este fenómeno se presenta pero conforme avanza el tiempo las oleadas presentan un número menor de infectados.

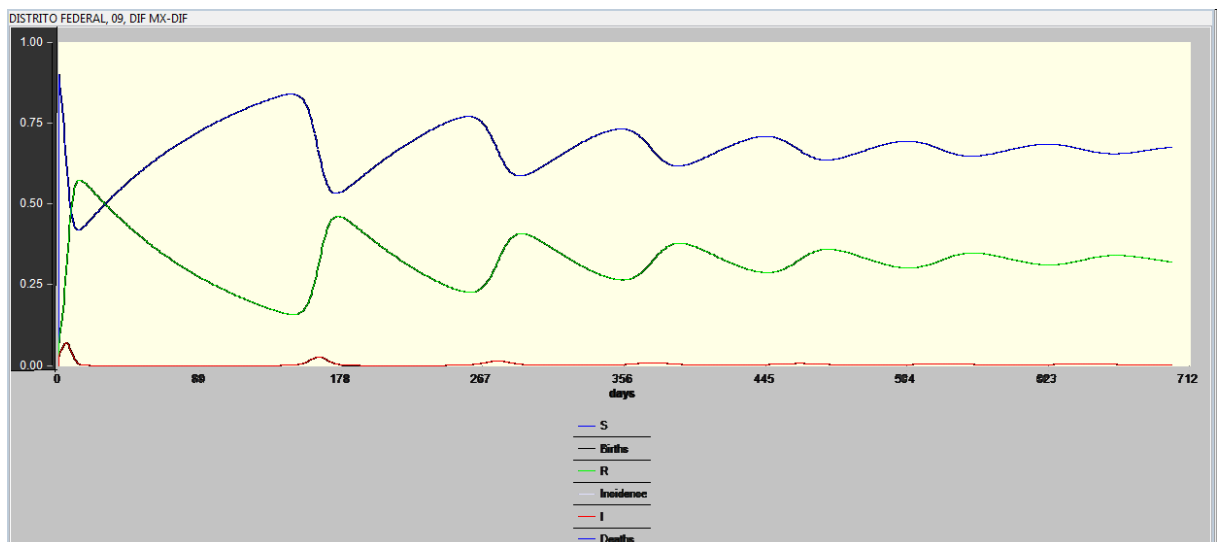


Figura 2.4: Serie de tiempo para el modelo SIR

La gráfica de serie de tiempo, figura (2.4) muestra como población infecciosa muestra altibajos y conforme las personas pierden el estado infectado, pasan al estado recuperado y el estado susceptible aumenta.

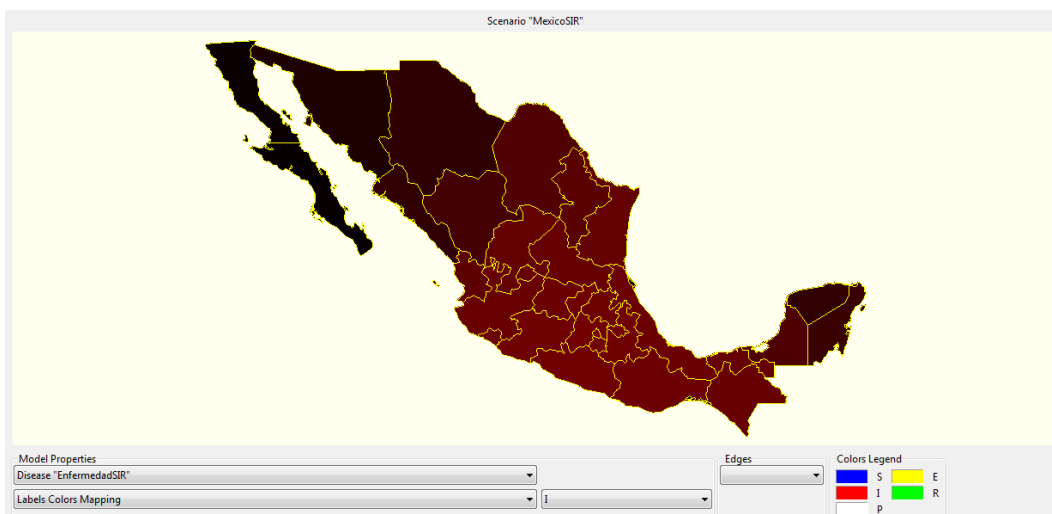


Figura 2.5: Propagación en 40 días con el modelo SIR



Figura 2.6: Propagación en 207 días con el modelo SIR



Figura 2.7: Propagación en 494 días con el modelo SIR

En diferentes días podemos ver como se va propagando la enfermedad conforme pasa el tiempo en 40, 207 y 494 días, figuras (2.5, 2.6, 2.7)

Capítulo 3

Estimación de parámetros para el modelo SIR endémico

Debido a la naturaleza del virus de la influenza A(H1N1) proponemos un modelo endémico SIR para modelar su propagación en el estado de Jalisco.

Para resolver nuestro problema tomamos los reportes de la propagación y estado del virus de la influenza A(H1N1) del período de abril del 2009 a agosto del 2010, debido a que con estos datos vamos a estimar los parámetros del modelo.

3.1. Estimación de parámetros por máxima verosimilitud

Para estimar los parámetros por este método, vamos a utilizar una versión normalizada del modelo SIR y vamos a tomar la tasa de pérdida de inmunidad igual a cero.

$$\frac{ds}{dt} = \mu - \beta si - \mu s \quad (3.1a)$$

$$\frac{di}{dt} = \beta si - \gamma i - \mu i \quad (3.1b)$$

$$\frac{dr}{dt} = \gamma i - \mu r \quad (3.1c)$$

Partiendo de la estimación de parámetros en sistemas dinámicos, podemos escribir el modelo (3.1) como una representación diferencial paramétrica de espacio-estado como

$$\dot{x} = f(x, \theta, u), \quad x \in \mathbb{R}^n, \quad u \in \mathbb{R}^m, \quad \theta \in \mathbb{R}^P \quad (3.2)$$

donde $x = [x_1, \dots, x_n]^T$ es el vector de estado, $u = [u_1, \dots, u_m]^T$ es el vector de entradas y $\theta = [\theta_1, \dots, \theta_p]^T$ es el vector de parámetros.

Supongamos que conocemos el vector de entradas y que están uniformemente distribuidas en los instantes de tiempo t_1, \dots, t_N , alternativamente estas proporciones pueden ser escritas como

$$y(t_1), \dots, y(t_N)$$

con

$$y(t_j) = [y_1(t_j), \dots, y_n(t_j)]^T$$

Supongamos que conocemos el valor inicial del vector de estado $x(t_0)$.

Para un vector de entradas $u(t)$ y dado un estado inicial $x(t_0)$, la solución al sistema diferencial (3.12) es parametrizada por el vector θ y denotada por $x(t, \theta)$. Existe una desviación entre el estado $x(t_0)$ calculado con el modelo y las entradas $y(t_j)$, esto se debe a diferentes factores, tales como, errores de modelado, de entradas, procesos ruidosos o entradas ruidosas, entre otros. Esta desviación usualmente es llamada error de salida y se denota como;

$$w(t_j, \theta) = y(t_j) - x(t_j, \theta), \quad j = 1, \dots, N$$

El método de máxima verosimilitud sigue de una representación de una sucesión de errores de salida $w(t_j, \theta)$ en los instantes del muestreo como una realización de un proceso estocástico. El parámetro óptimo a estimar $\hat{\theta}$ se define como el vector de parámetros que maximiza la probabilidad de que $y(t_j)$ sea una medida de $x(t_j, \theta)$, vista como la realización de un vector aleatorio. Esto nos permite definir la matriz de covarianza del estimador como,

$$C_{\hat{\theta}} = E\{(\hat{\theta} - E(\hat{\theta}))(\hat{\theta} - E(\hat{\theta}))^T\}$$

que puede ser usado para estimar intervalos de confianza de los parámetros estimados.

En la practica el vector $w(t_j, \theta)$ puede estar altamente correlacionado, esto es una evidencia de que es posible la aproximación por máxima verosimilitud para sistemas no lineales, para explicar la correlación de los errores de salida. **Sin embargo, la dificultad principal aparente, esta en la necesidad de el cálculo de la inversa de una matriz de covarianza muy grande.**

La estimación del vector de parámetros θ resulta de la maximización de la función que describe la probabilidad de que n vectores y_j sean medida de n vectores $x_j(\theta)$. Para construir esta función de probabilidad, suponemos que $w_j(\theta)$ es un vector aleatorio que toma valores en \mathbb{R}^N con media cero y una función de densidad de probabilidad dada por;

$$f_{w_j} = \frac{1}{\sqrt{(2\pi)^N |\Sigma_j|}} e^{-1/2 w_j(\theta)^T \Sigma_j w_j(\theta)} \quad (3.3)$$

donde Σ_j es la matriz de covarianza del vector aleatorio $w_j(\theta)$ denotada como:

$$\Sigma_j = E\{(w_j - E(w_j))(w_j - E(w_j))^T\} \quad (3.4)$$

Debido a que θ es desconocido, consideramos a (3.3) como una función del parámetro θ que será estimado.

Supongamos que todos los vectores w_j son independientes uno del otro, es decir,

$$E\{w_{j_1}(t_{k_1})w_{j_2}(t_{k_2})\} = 0 \quad \forall k_1, k_2 \quad \text{and} \quad \forall j_1 \neq j_2$$

Esto significa que la medida de cada variable de estado puede ser alterada por diferentes e independientes procesos estocásticos.

Cada proceso puede estar altamente correlacionado en el tiempo, esta correlación se puede jerarquizar en las matrices Σ_j .

Esta suposición nos permite calcular la densidad de probabilidad conjunta para todos los vectores y_j , como una función de θ :

$$f_{y_1, \dots, y_n}(\theta) = \prod_{j=1}^n \frac{1}{\sqrt{(2\pi)^N |\Sigma_j|}} e^{-1/2 w_j(\theta)^T \Sigma_j^{-1} w_j(\theta)} \quad (3.5)$$

Esta función determinística de θ es llamada la *función de verosimilitud* que es proporcional a la probabilidad de que n vectores y_j sean medidos de n vectores $x_j(\theta)$. La estimación del vector de parámetros θ resulta de la maximización de esta función. Considerando su logaritmo, podemos replantear este problema como un problema de minimización:

$$\hat{\theta}^{ML} = \arg \min_{\theta} J(\theta) \quad (3.6)$$

donde

$$J(\theta) = \sum_{j=1}^n w_j(\theta)^T \Sigma_j^{-1} w_j(\theta) \quad (3.7)$$

El gradiente de la función de costo $J(\theta)$ esta dado por

$$\frac{\partial J(\theta)}{\partial \theta} = -2 \sum_{j=1}^n w_j(\theta)^T \Sigma_j^{-1} G_j \quad (3.8)$$

donde G_i son las matrices de sensibilidad de estado y están dadas por

$$G_j = \frac{\partial x_j(\theta)}{\partial \theta} \in \mathbb{R}^{N \times P} \quad (3.9)$$

las cuales pueden ser calculadas por la integración de la siguiente ecuación diferencial matricial.

$$\frac{d}{dt}G_j(t) = \frac{\partial f(x, \theta)}{\partial x}G_j(t) + \frac{\partial f(x, \theta)}{\partial \theta}, \quad G_j(0) = 0 \quad (3.10)$$

3.2. Estimación de parámetros utilizando ji cuadrada χ^2

Para estimar los parámetros por este método vamos a utilizar una versión no normalizada del modelo SIR

$$S' = \mu N - \frac{\beta SI}{N} - \mu S \quad (3.11a)$$

$$I' = \frac{\beta SI}{N} - (\gamma + \mu)I \quad (3.11b)$$

$$R' = \gamma I - \mu R \quad (3.11c)$$

Aplicando la misma metodología para la estimación por máxima verosimilitud, si partimos de la estimación de parámetros en sistemas dinámicos, podemos escribir el modelo (3.11) como una representación diferencial paramétrica de espacio-estado

$$\dot{x} = f(x, \theta), \quad x \in \mathbb{R}^n, \quad \theta \in \mathbb{R}^P \quad (3.12)$$

donde $x = [x_1, \dots, x_n]^T$ es el vector de estado y $\theta = [\theta_1, \dots, \theta_p]^T$ es el vector de parámetros.

El vector de parámetros dependiente de la variable de estado $x_j(t, \theta)$ evaluado en los instantes del tiempo $t_k = t_1, \dots, t_N$ se denota por $x_j(\theta) = [x_j(t_1, \theta), \dots, x_j(t_N, \theta)]^T$, $j = 1, \dots, N$.

El vector de las medidas de la variable de estado x_j al paso del tiempo se denota por $y_j = [y_j(t_1), \dots, y_j(t_N)]^T$

A la desviación entre el estado $x_j(\theta)$ calculado con el modelo y las entradas y_j se le conoce como el vector de errores de salida $w_j(\theta) = y_j - x_j(\theta)$.

Si tenemos un estado inicial $x(t_0)$, la solución para el sistema diferencial (3.11) es parametrizada por el vector θ y se denota por $x(t, \theta)$

Para estimar el vector de parámetros θ debemos maximizar la función que describe la verosimilitud de n vectores y_j sean medidos de n vectores $x_j(\theta)$.

Utilizando el vector de errores de salida, necesitamos resolver el siguiente problema de minimización

$$\hat{\theta} = \arg \min_{\theta} J(\theta), \quad \text{donde} \quad J(\theta) = \sum_{j=1}^n w_j(\theta)^T w_j(\theta) \quad (3.13)$$

El gradiente de la función de costo $J(\theta)$ esta dado por

$$\frac{\partial J(\theta)}{\partial \theta} = -2 \sum_{j=1}^n \left(\frac{\partial x_j(\theta)}{\partial \theta} \right)^T w_j(\theta) \quad (3.14)$$

el cual se puede calcular de las ecuaciones variacionales.

3.3. Solución al problema de minimización

Para estimar los parámetros del modelo, necesitamos minimizar la función de costo y debido a la naturaleza del problema no es recomendable utilizar métodos directos, por ejemplo, factorización de Cholesky, por lo que utilizaremos los métodos de optimización no cuadrática de gradiente conjugado en el caso no-lineal y por el método de sobrerelajación sucesiva SOR.

Otro problema en la estimación por máxima verosimilitud es el de invertir la matriz de covarianza, debido a que es una matriz muy cercana a ser singular, por lo que necesitamos un método de regularización para dar una aproximación de la matriz inversa.

3.3.1. Método de gradiente conjugado no-lineal GC

Este método es iterativo y a partir de un valor de inicio va calculando valores que se van acercando a la solución exacta del problema utilizando direcciones conjugadas y búsqueda de línea. En el algoritmo usamos la formula de descenso conjugado de Polak-Riviere [15].

- Dado x^0
- $g^0 = \nabla f(x^0)$
- $d^0 = g^0$
- mientras $norm(g^k) > tol$ y $k < kmax$
- $\alpha_k = argmin f(x^k - \alpha d^k)$
- $x^{k+1} = x^k - \alpha_k d^k$
- $g^{k+1} = \nabla f(x^{k+1})$
- $B_k = -\frac{(g^{k+1})^T (g^{k+1} - g^k)}{(g^k)^T (g^k)}$
- $d^{k+1} = g^{k+1} - B_k d^k$
- fin mientras

Cuando el criterio de paro se cumple, esta solucionado el problema de minimización.

3.3.2. Sobrerrelajación sucesiva SOR

Este método es una variante del método de Gauss-Seidel que resuelve sistemas de ecuaciones, pero con la ventaja de que converge más rápido, su principio es que puede predecir la nueva solución, extrapolando entre la iteración anterior y la iteración calculada con Gauss-Seidel para cada componente. Vamos a resolver el problema de minimización utilizando el método SOR [3].

Si la iteración original tiene la forma

$$x_{n+1} = f(x_n)$$

entonces podemos utilizar su versión modificada que es

$$x_{n+1}^{SOR} = (1 - \omega)x_n^{SOR} + \omega f(x_n^{SOR})$$

que es equivalente a

$$x_n = \omega x_{n-1} + (1 - \omega)x_{n-2}$$

claro que para un valor conveniente $0 < \omega < 1$. Entonces llegamos al mínimo de la función objetivo.

3.3.3. Regularización de Tikhonov

La regularización de Tikhonov es un método comúnmente utilizado para regularizar problemas bien planteados (existe solución, es única, dependen continuamente de los datos). En estadística, este método se conoce como regresión lineal. Se relaciona con el algoritmo de mínimos cuadrados para minimizar problemas no lineales

La forma estándar para resolver un problema de regresión lineal, está dado por,

$$Ax = b$$

que se conoce como mínimos cuadrados lineales donde se busca minimizar el residual

$$\|Ax - b\|^2$$

donde $\|\cdot\|$ es la norma Euclidiana. Sin embargo, la matriz A es singular o muy cercana a serlo. Para obtener una solución con propiedades deseables, se incluye un término de regularización en la minimización:

$$\|Ax - b\|^2 + \|\alpha x\|^2$$

para alguna matriz α elegida a conveniencia. Por ejemplo, si elegimos α como la matriz identidad $\alpha = I$, obtenemos resultados con normas pequeñas, es decir con un error muy pequeño. Al utilizar este método de regularización, se mejora el

condicionamiento del problema, permitiendo una buena aproximación numérica a la solución x , la cual denotamos como:

$$\hat{x} = (A^T A + \alpha^T \alpha)^{-1} A^T b$$

Elección del pámetro de regularización por el método de la curva L

La idea del método de la curva L es elegir una sucesión de candidatos del parámetro de regularización α , de manera que para cada α_m desde $1 \leq m \leq M$ la sucesión tenga la siguiente forma,

$$0 < \alpha_1 < \alpha_2 < \dots < \alpha_M < \infty$$

para cada valor de α_m , se calcula la aproximación a la solución del problema $\|Ax - b\|^2 + \|\alpha x\|^2$, es decir, la \hat{x}_{α_m} solución. Una vez que tenemos la solución regularizada para cada α_m , calculamos su residual, y se define una sucesión de coordenadas de la siguiente manera,

$$(\log \|x - \hat{x}_{\alpha_m}\|, \log \|\hat{x}_{\alpha_m}\|) \in \mathbb{R}^2$$

de tal manera que al graficar estos puntos en un plano, debe formarse una curva suave que por lo regular tiene la forma de una L. El valor óptimo para α es el que se encuentra mas a la esquina de la curva [31].

Máxima curvatura de la curva L

Se puede encontrar el punto de máxima curvatura de la curva L de la siguiente manera; sea \hat{x}_α la solución regularizada y $r_\alpha = x - \hat{x}_\alpha$ el residual regularizado. Se define,

$$R(\alpha) = \|r_\alpha\|, \quad X(\alpha) = \log R(\alpha) \quad (3.15)$$

$$S(\alpha) = \|\hat{x}_\alpha\|, \quad Y(\alpha) = \log \hat{x}_\alpha \quad (3.16)$$

Suponemos que $X(\alpha)$ y $Y(\alpha)$ varían suavemente con respecto a α . Se selecciona el valor que maximiza la función de curvatura, de la siguiente manera,

$$\kappa(\alpha) = \frac{X''(\alpha)Y'(\alpha) - X'(\alpha)Y''(\alpha)}{(X'(\alpha)^2 - Y'(\alpha)^2)^{3/2}} \quad (3.17)$$

donde la prima (\prime) denota diferenciación con respecto de α [44].

3.3.4. Intervalo de confianza para los parámetros estimados

Para calcular el intervalo de confianza de los parámetros estimados, utilizamos la matriz de información de Fisher, que se determina con la siguiente ecuación [24],

$$F = G_j^T \Sigma^{-1} G_j \quad (3.18)$$

donde G_j es la matriz de sensibilidad que se calculo anteriormente y Σ es la matriz de covarianza.

La desviación estándar de los parámetros es entonces la raíz cuadrada de los elementos de la diagonal de la matriz F^{-1} .

$$\sigma_i \geq \sqrt{F_{ii}^{-1}} \quad (3.19)$$

El intervalo de confianza del 95 % puede ser aproximado por [24],

$$\hat{\theta} - 2\sigma_i \leq \theta \leq \hat{\theta} + 2\sigma_i \quad (3.20)$$

3.4. Implementación computacional

Consideremos el sistema diferencial (3.12) del modelo SIR definimos los siguientes vectores $x = [s, i, r]^T \in \mathbb{R}^3$ es el vector de estado, $u = [PI]^T$ es el vector de entradas donde PI es la proporción de infectados suponiendo toda la población de Jalisco como susceptible a enfermar con el virus y $\theta = [\mu, \beta, \gamma]^T \in \mathbb{R}^3$ es el vector de parámetros.

Tomamos diferentes tamaños de vector, con $N = 6$ y $N = 79$ para $t_k = t_1, \dots, t_N$, así, podemos tomar el vector parámetro-dependiente de la variable de estado $x_j(t_k, (\mu, \beta, \gamma))$ como

$$x_i((\mu, \beta, \gamma)) = [x_j(t_1, (\mu, \beta, \gamma)), \dots, x_j(t_N, (\mu, \beta, \gamma))]^T, \quad j = 1, \dots, 3$$

además el vector de medida de la variable de estado x_i al paso del tiempo es denotado por

$$y_j = [(PI)_j(t_1), \dots, (PI)_j(t_N)]^T$$

y el vector de errores de salida $w_j(\mu, \beta, \gamma)$ se define por,

$$w_j(\mu, \beta, \gamma) = y_j - x_j(\mu, \beta, \gamma). \quad (3.21)$$

una vez que encontramos el error del vector de salida, utilizamos dos caminos para estimar los parámetros.

3.4.1. Implementación computacional del método de máxima verosimilitud

Utilizamos el cálculo del error de salida (3.21) procedemos a calcular la matriz de covarianza de la siguiente manera,

$$\Sigma_j = E\{(w_j - E(w_j))(w_j - E(w_j))^T\} \quad (3.22)$$

Debido a que la matriz de covarianza se forma con el vector de errores de salida es muy cercana a ser singular y es necesario aplicar un método de regularización,

elegimos la regularización de Tikhonov. Otra dificultad del método consiste en invertir la matriz de covarianza en cada paso del tiempo, por lo que vamos a utilizar el método de gradiente conjugado no-lineal con la variante Polak-Riviere.

Una vez invertida la matriz de covarianza, calculamos la función objetivo como

$$J(\mu, \beta, \gamma) = \sum_{j=1}^3 w_j(\mu, \beta, \gamma)^T \Sigma_j^{-1} w_j(\mu, \beta, \gamma) \quad (3.23)$$

y al integrar la ecuación diferencial matricial

$$\frac{d}{dt} G_j(t) = \frac{\partial f(x, (\mu, \beta, \gamma))}{\partial x} G_j(t) + \frac{\partial f(x, (\mu, \beta, \gamma))}{\partial (\mu, \beta, \gamma)}, \quad G_j(0) = 0 \quad (3.24)$$

obtenemos

$$G_j = \frac{\partial x_j}{\partial (\mu, \beta, \gamma)} \in \mathbb{R}^{3 \times 3} \quad (3.25)$$

con esta matriz ahora ya podemos calcular el gradiente de la función de costo

$$\frac{\partial J(\mu, \beta, \gamma)}{\partial (\mu, \beta, \gamma)} = -2 \sum_{j=1}^3 w_j(\mu, \beta, \gamma)^T \Sigma_j^{-1} G_j \quad (3.26)$$

3.4.2. Implementación computacional de la distribución Ji-cuadrada χ^2

Una vez que tenemos el vector de errores de salida (3.21) calculamos la función objetivo que necesitamos minimizar,

$$J(\mu, \beta, \gamma) = \sum_{i=1}^3 w_i(\mu, \beta, \gamma)^T w_i(\mu, \beta, \gamma) \quad (3.27)$$

El gradiente de la función de costo $J(\mu, \beta, \gamma)$ esta dado por

$$\frac{\partial J(\mu, \beta, \gamma)}{\partial (\mu, \beta, \gamma)} = -2 \sum_{i=1}^3 \left(\frac{\partial w_i(\mu, \beta, \gamma)}{\partial (\mu, \beta, \gamma)} \right)^T w_i(\mu, \beta, \gamma) \quad (3.28)$$

el cual se puede calcular de las ecuaciones variacionales. Hay que notar que para este método utilizamos una versión no normalizada del modelo SIR, por lo que los casos de las estadísticas de la SSJ.

3.4.3. Resultados de las simulaciones

Para hacer nuestras simulaciones tomamos $R_0 = 1.58$ que es tomado de Fraser [16] para calcular nuestros parámetros de inicio. Como los datos vienen dados semanalmente, calculamos μ como el inverso de la esperanza de vida en Jalisco que es de 76.86

años entonces $\mu = \frac{1}{76.86 \cdot 52} = 0.00025$, decimos que el promedio infeccioso del virus es de 5 días [43] por lo que $\gamma = \frac{7}{5}$ y si sustituimos estos resultados en la ecuación (1.8) encontramos que la tasa de transmisión es $\beta = 2.212395$.

Además, tomamos los datos estadísticos proporcionados por la SSJ figura (3.1) para 6 y 79 semanas respectivamente.

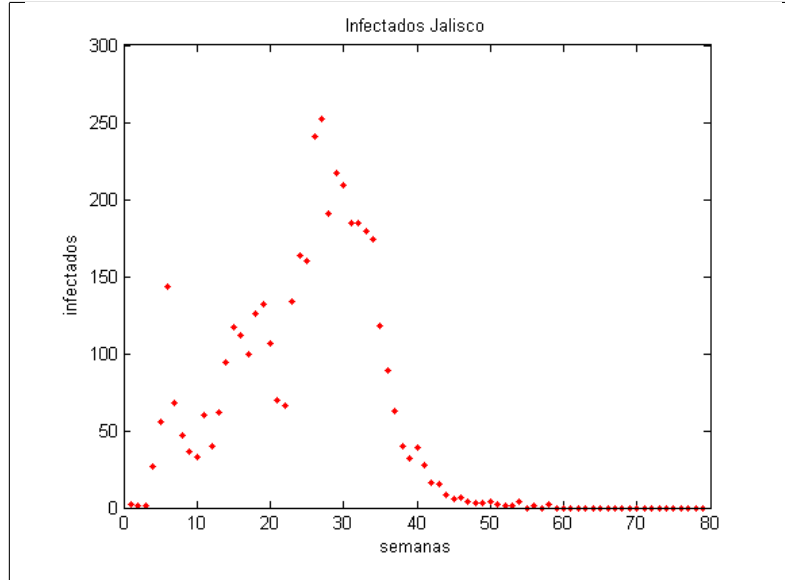


Figura 3.1: Infectados en Jalisco

Simulaciones utilizando el método de máxima verosimilitud

En esta parte del trabajo, vamos a estimar μ , β y γ del modelo normalizado (3.1), tomando los datos para las 79 semanas que fue todo el período de estudio. Tomamos como parámetros iniciales los que calculamos con $R_0 = 1.58$ y minimizamos el problema utilizando GC no-lineal y regularización de Tikhonov.

Encontramos que al minimizar la función objetivo, los valores para los parámetros son los que se presentan en el cuadro (3.1):

μ	β	γ
0.012307998965084	11.074366171596264	10.993436181379380
R_0	Intervalo de confianza 95 %	error relativo
1.006235106879402	(0.920818197131399, 1.099611693191965)	1.199211906516079

Cuadro 3.1: Parámetros estimados para la propagación durante 79 semanas utilizando máxima verosimilitud y GC no-lineal

Si hacemos la gráfica de incidencia

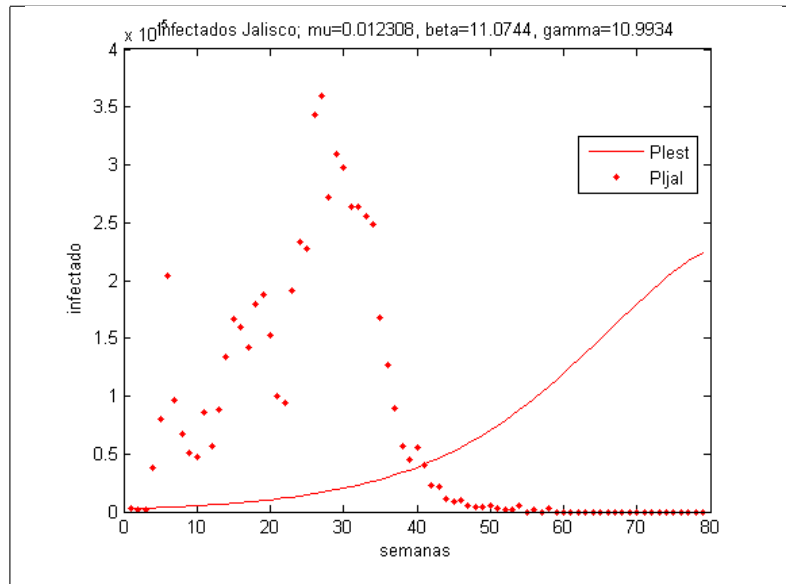


Figura 3.2: Propagación durante 79 semanas utilizando máxima verosimilitud y GC no-lineal

Si ahora tomamos a $\mu = 0.00025$ como una constante, y solo estimamos β y γ , nuestro algoritmo se vuelve mas rápido y obtenemos los siguientes valores de los parámetros, cuadro (3.2).

μ	β	γ
0.00025	11.083855209831800	10.984000677203728
R_0	Intervalo de confianza 95 %	error relativo
1.009067940595601	(0.871937096634555,1.168012851868845)	1.957572392914411

Cuadro 3.2: Parámetros estimados para la propagación durante 79 semanas con μ fija y utilizando máxima verosimilitud y GC no-lineal

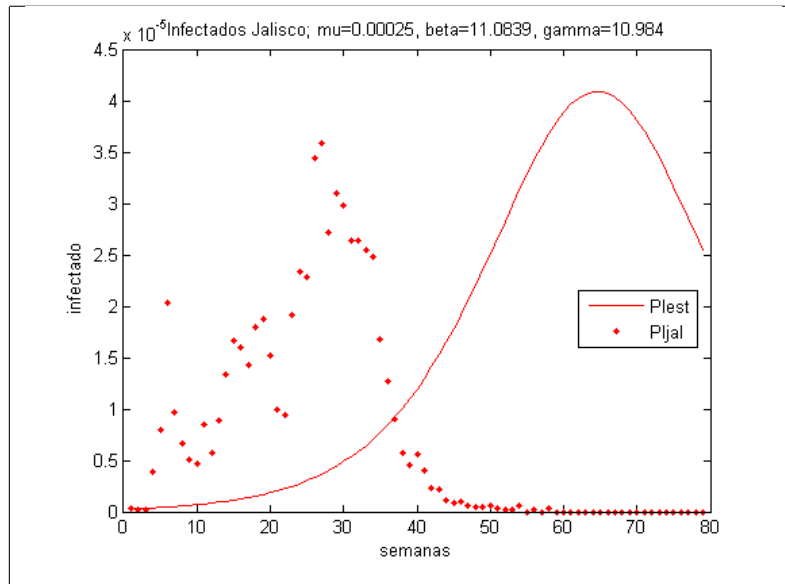


Figura 3.3: Propagación durante 79 semanas utilizando máxima verosimilitud, GC no-lineal y μ fija

Al observar los datos proporcionados por la SSJ se puede apreciar que en el periodo de estudio sucedieron 3 brotes de la enfermedad, figura (3.1), al ver los datos de las primeras seis semanas observamos que el número de infectados empieza a crecer por lo que se tomaron medidas sanitarias precautorias en las que la mas notable fue el aislamiento social de la semana 4 a la 6, para disminuir el número de personas infectadas por el virus. Al estimar los paraámetros para las primeras 6 semanas, con μ fija, obtuvimos los resultados que se muestran en el cuadro (3.3) y la gráfica de incidencia (3.4)

μ	β	γ
0.00025	2.041370324078564	1.227584844548659
R_0	Intervalo de confianza 95 %	error relativo
1.662577286466368	(1.413977384520947, 1.968259586983927)	0.183106920291212

Cuadro 3.3: Parámetros estimados para la propagación durante 6 semanas con μ fija y utilizando máxima verosimilitud y GC no-lineal

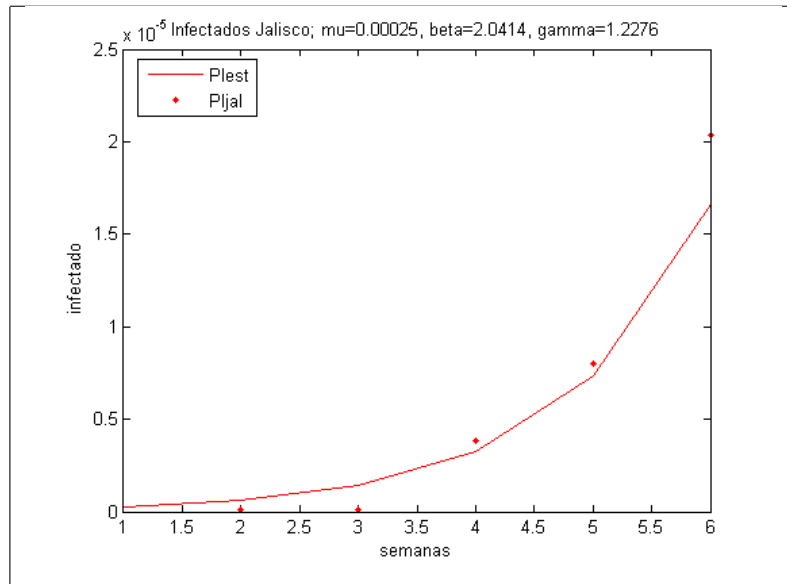


Figura 3.4: Propagación en las primeras 6 semanas utilizando máxima verosimilitud, GC no-lineal y μ fija

Si extendemos el período de propagación a 79 semanas pero utilizamos los parámetros estimados para las primeras 6, suponiendo que no se hubieran tomado medidas de prevención figura (3.5)

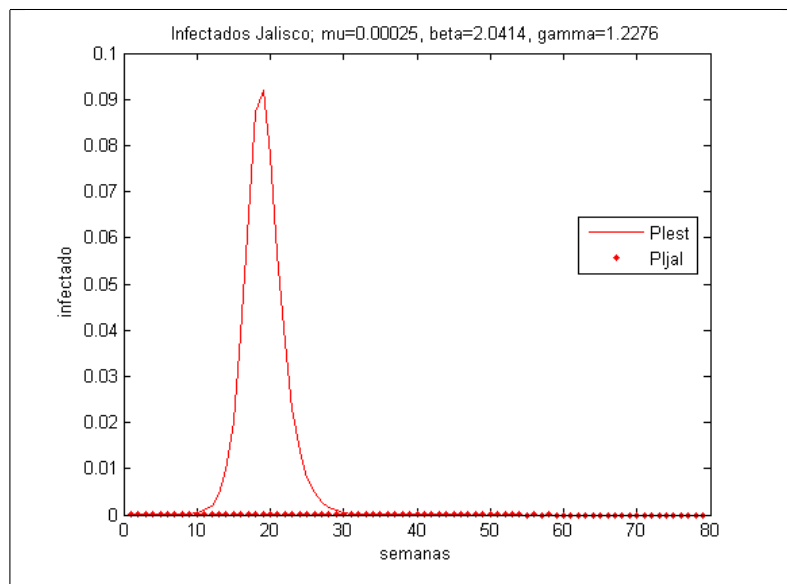


Figura 3.5: Propagación de las 79 semanas con los parámetros de las primeras 6 utilizando máxima verosimilitud, GC no-lineal y μ fija

Si ahora estimamos los parámetros utilizando el método SOR para las primeras 6 semanas y μ fija, podemos observar en la figura (3.6) que la solución que obtuvimos con los parámetros del cuadro (3.4) es buena y se adapta a la naturaleza de la epidemia para este periodo de tiempo tan corto,

μ	β	γ
0.00025	1.681739997636880	0.827565183705449
R_0	Intervalo de confianza 95 %	error relativo
2.031540409912646	(1.696429799154473,2.461766925664745)	0.075680671609574

Cuadro 3.4: Parámetros estimados para la propagación en las primeras 6 semanas utilizando el método SOR y μ fija

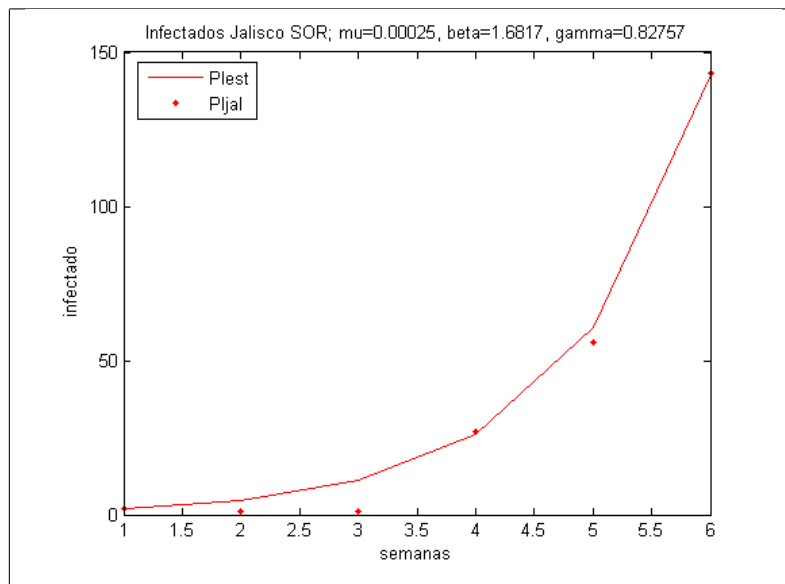


Figura 3.6: Propagación en las primeras 6 semanas utilizando el método SOR y μ fija

Si extendemos el periodo de propagación a 79 semanas pero utilizamos los parámetros estimados para las primeras 6, suponiendo que no se hubieran tomado medidas de prevención, figura (3.7),

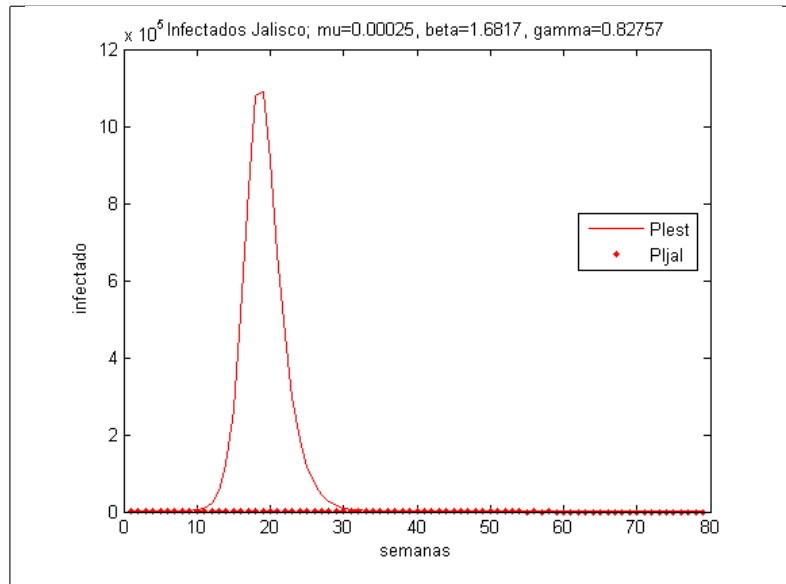


Figura 3.7: Propagación durante 79 semanas con los parámetros de 6, utilizando el método SOR

Si ahora utilizamos el modelo sin normalizar (3.11) y estimamos los parámetros de igual manera que para el modelo normalizado. Al estimar los parámetros para las primeras 6 semanas, con μ fija, obtuvimos los resultados del cuadro (3.5) y la figura (3.8)

μ	β	γ
0.00025	2.064878623642913	1.209792232622020
R_0	Intervalo de confianza 95 %	error relativo
1.706451698936625	(1.437364264820157, 2.042328442707254)	0.076461042083420

Cuadro 3.5: Parámetros estimados para la propagación durante 6 semanas con μ fija y utilizando máxima verosimilitud y GC no-lineal

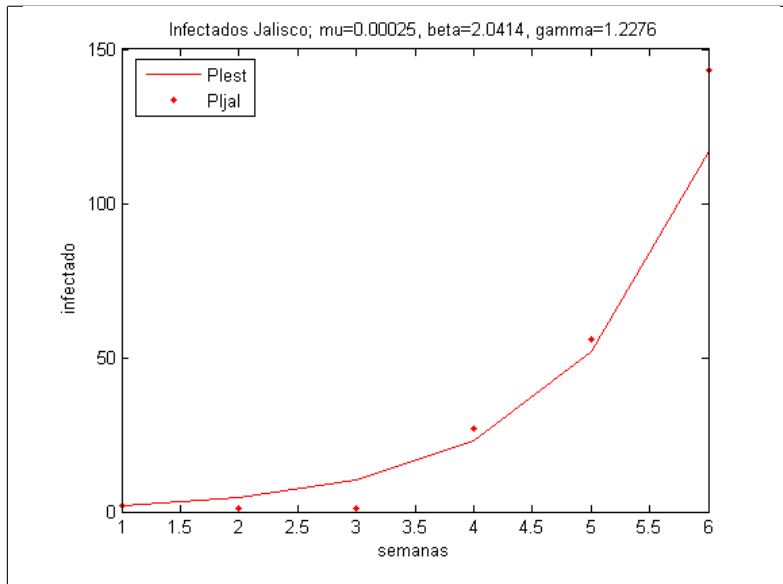


Figura 3.8: Propagación en las primeras 6 semanas utilizando máxima verosimilitud, GC no-lineal y μ fija

Si extendemos el periodo de propagación a 79 semanas pero utilizamos los parámetros estimados para las primeras 6, suponiendo que no se hubieran tomado medidas de prevención, figura (3.9)

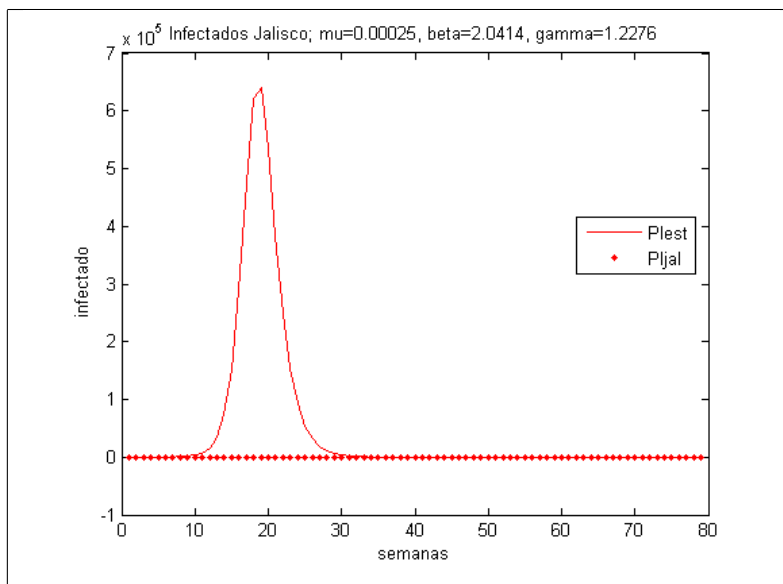


Figura 3.9: Propagación de las 79 semanas con los parámetros de las primeras 6 utilizando máxima verosimilitud, GC no-lineal y μ fija

3.5. Estimación de parámetros para Bélgica y Francia

Utilizamos el método de máxima verosimilitud descrito anteriormente, para estimar los parámetros de infecciones de enfermedades respiratorias para Bélgica y Francia.

Vamos a considerar las estimaciones de los casos positivos en estos países para una población de 100000 habitantes, estos datos se encuentran disponibles en Google flu trends [34].

Utilizamos la serie de datos para Bélgica tomados semanalmente desde el 23 de agosto del 2009 hasta el 22 de mayo del 2011, y estimamos los parámetros μ , β y γ obteniendo los siguientes resultados, cuadro (3.6).

μ	β	γ
0.024989840045384	3.398537516045367	2.701850997336147
R_0	Intervalo de confianza 95 %	error relativo
1.246327790553716	(1.016447908320612, 1.542949650460891)	2.107831170126933

Cuadro 3.6: Parámetros estimados para Bélgica durante 91 semanas utilizando máxima verosimilitud y GC no-lineal

La figura (3.10) muestra la incidencia para los parámetros estimados para Bélgica,

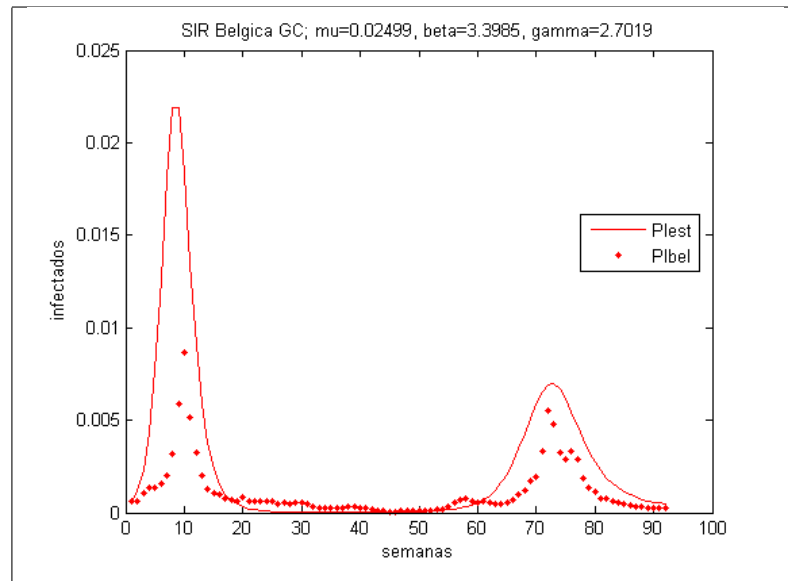


Figura 3.10: Propagación para 91 semanas para Bélgica con todos los parámetros utilizando máxima verosimilitud, GC no-lineal

Utilizamos la serie de datos para Francia tomados semanalmente desde el 8 de

agosto del 2004 hasta el 9 de julio del 2006, y estimamos los parámetros μ , β y γ obteniendo los siguientes resultados, cuadro (3.7).

μ	β	γ
0.040062534893096	2.197182339489372	1.803239906099932
R_0	Intervalo de confianza 95 %	error relativo
1.191981462524240	(0.914466372015850,1.405866431613108)	2.555540173677056

Cuadro 3.7: Parámetros estimados para Francia durante 101 semanas utilizando máxima verosimilitud y GC no-lineal

La figura (3.11) muestra la incidencia para los parámetros estimados para Francia,

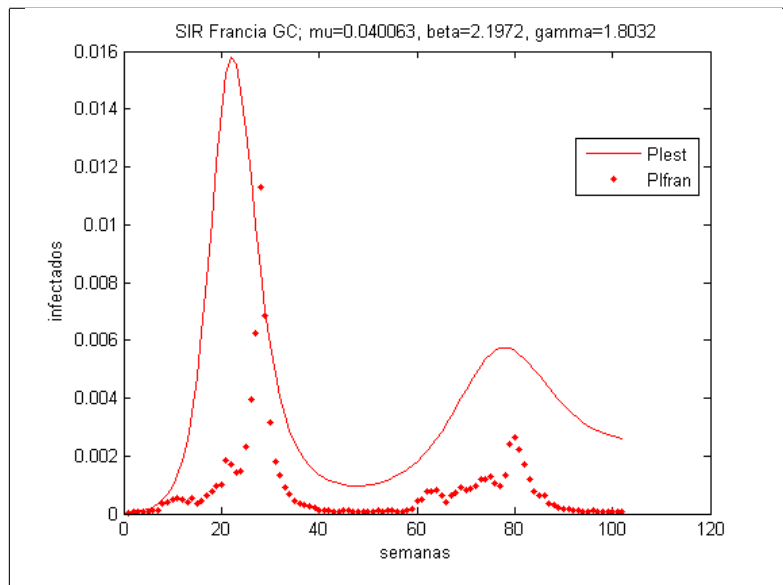


Figura 3.11: Propagación para 101 semanas para Francia con todos los parámetros utilizando máxima verosimilitud, GC no-lineal

Capítulo 4

Estimación de parámetros para el modelo SIR epidémico

En este capítulo, proponemos un modelo epidémico SIR para modelar la propagación de la influenza A(H1N1) en el estado de Jalisco. En este modelo, consideramos un período de tiempo muy corto, por lo que no se incluyen las muertes y nacimientos en la población, es decir no incluimos dinámica vital, consideremos entonces el modelo normalizado,

$$\frac{ds}{dt} = -\beta si \quad (4.1a)$$

$$\frac{di}{dt} = \beta si - \gamma i \quad (4.1b)$$

$$\frac{dr}{dt} = \gamma i \quad (4.1c)$$

4.1. Estimación de parámetros utilizando evolución diferencial

La evolución diferencial o DE, es una rama de la computación evolutiva para optimización en espacios continuos, es una técnica de búsqueda paralela, directa y estocástica, que maneja funciones objetivo no diferenciables y no lineales [25].

Utiliza NP , D – dimensional vectores de parámetros [32],

$$x_{i,G} \quad \text{para} \quad i = 1, 2, \dots, NP$$

como población en cada generación G , y este número queda fijo durante todo el proceso. El vector de población inicial, se elige aleatoriamente del espacio que cubre a todos los posibles parámetros. Como una regla, se supone una distribución de probabilidad uniforme, para todas las decisiones, a menos que se indique lo contrario. En caso de disponer de una solución preliminar, la población inicial puede ser generada

agregando desviaciones estándar aleatorias normalmente distribuidas, a la solución nominal $x_{nom,0}$. DE genera un nuevo vector de parámetros agregando la diferencia de pesos entre dos vectores de la población a un tercer vector. Esta operación es llamada mutación. El vector de parámetros que ha sido transformado, vector mutante, se mezcla con los parámetros de otro vector predeterminado, el vector nuevo, es llamado vector de ensayo. La mezcla de estos parámetros es llamada frecuentemente cruza.

El vector de ensayo, es aceptado para reemplazar la siguiente generación, si y solo si, existe una reducción en el valor de la función objetivo, de otra forma nos quedamos con el vector de ensayo anterior. Esta operación es conocida como selección. Cada vector de la población sirve una vez como vector de ensayo, hasta que las NP competiciones tengan lugar en una generación.

Se puede describir esta estrategia mas específicamente como sigue;

- **Mutación**

Para cada vector objetivo, $x_{i,G}$, el vector mutante es generado de acuerdo a;

$$v_{i,G+1} = x_{r_1,G} + F \cdot (x_{r_2,G} - x_{r_3,G})$$

con índices aleatorios enteros $x_{r_1,G}, x_{r_2,G}, x_{r_3,G} \in 1, 2, \dots, NP$, mutuamente diferentes y $F > 0$. Los enteros elegidos aleatoriamente $x_{r_1,G}, x_{r_2,G}$ y $x_{r_3,G}$ se eligen de manera diferente a el índice i de tal manera que NP debe ser mayor o igual a cuatro para que se pueda cumplir esta condición. El factor F es una constante real que se encuentra entre 0 y 2, que controla la amplificación de la variación diferencial $(x_{r_2,G} - x_{r_3,G})$.

- **Cruza**

Se introduce la cruza, para incrementar la diversidad del vector de parámetros perturbados. Con este fin, el vector de ensayo,

$$u_{i,G+1} = (u_{1i,G+1}, u_{2i,G+1}, \dots, u_{Di,G+1})$$

se forma de la siguiente manera,

$$u_{ji,G+1} = \begin{cases} v_{ji,G+1} & \text{si } (randb(j) \leq CR) \text{ ó } j = nbrn(i) \\ x_{ji,G} & \text{si } (randb(j) > CR) \text{ ó } j \neq nbrn(i), \end{cases} \quad \text{para } j = 1, 2, \dots, D \quad (4.2)$$

$randb(j)$ es la j -ésima evaluación de un número aleatoriamente generado entre 0 y 1. CR es la constante de cruza que se encuentra entre cero y uno y es generada por el usuario. $nbrn(i)$ es un índice aleatorio $\in (1, 2, \dots, D)$ que asegura que $u_{i,G+1}$ obtendrá al menos un parámetro de $v_{i,G+1}$

■ Selección

Para decidir independientemente si el vector de ensayo $u_{i,G+1}$ debe o no pertenecer a la generación $G + 1$, se compara con el vector objetivo $x_{i,G}$ utilizando el siguiente criterio. Si el vector $u_{i,G+1}$ tiene menor valor de la función de costo que $x_{i,G}$, entonces $x_{i,G+1}$ será $u_{i,G+1}$, de otra manera, se retiene el valor de $x_{i,G}$.

■ Algoritmo

Sea $x_{i,G}$ el vector candidato solución, en la población. El algoritmo básico es el siguiente;

-Inicializar todos los vectores $x_{i,G}$ con posiciones aleatorias en el espacio de búsqueda.

-Repetir lo siguiente hasta que se cumpla un criterio de paro.

+ Para cada $x_{i,G}$ en la población, hacer;

◇ Escoger aleatoriamente tres vectores $x_{r_i,G}, x_{r_j,G}, x_{r_k,G} \in 1, 2, \dots, NP$ de la población, sin que se repitan.

◇ Escoger aleatoriamente un índice $i \in 1, 2, \dots, NP$ donde el más grande valor de NP es la dimensionalidad del problema.

◇ Calcular el vector de prueba en la nueva posición $u_{i,G+1}$ iterandolo como sigue;

★ Escoger $randb(j) \sim U(0, 1)$

★ Si $j = nbrn(i)$ ó $randb(j) \leq CR$ entonces $u_{i,G+1} = x_{r_1,G} + F \cdot (x_{r_2,G} - x_{r_3,G})$

◇ Si $f(u_{i,G+1}) < f(x_{i,G+1})$ entonces se reemplaza el vector de prueba y $u_{i,G+1} = x_{i,G}$

+Elegir el vector de prueba como el que tiene la menor función de costo y regresarlo como el candidato solución.

4.2. Implementación computacional

La implementación en MATLAB, permite la ejecución de funciones de optimización, de un modo simple y seguro; toda vez que los clásicos métodos de optimización son inconvenientes al resolver problemas de optimización con múltiples objetivos [25].

Se tomo el algoritmo de uso libre implementado en MATLAB que se encuentra en [5], utilizando los siguientes criterios como condiciones iniciales para estimar los parámetros, que deben estar acotados en los valores probables que pueden tomar los parámetros. Las poblaciones de infectados, susceptibles y recuperados se encuentran acotados entre 0 y 1 porque estamos normalizando a las poblaciones. Los parámetros β y γ se encuentran acotados entre 0 y 100.

El factor de peso aplicado a la diferencia aleatoria, F es un factor de escala entre 0 y 2, lo vamos a tomar de 0.85. Tomamos la constante de cruza como 0.9, que es la probabilidad que tienen los individuos de ser cruzados y esta entre 0 a 1. Tomamos el tamaño de la población NP como diez veces el número de parámetros, aunque en nuestro caso se obtuvieron mejores resultados aumentando a 20 veces el número de parámetros. Iteramos 10000 veces, utilizando la estrategia basada en el vector a perturbar DE/best/1 with jitter, que corresponde a utilizar, cruza exponencial, perturbando un solo vector que debe ser el mejor.

4.2.1. Resultados de la simulación

Para resolver nuestro problema tomamos los reportes de la propagación y estado del virus de la influenza A(H1N1) del período de abril del 2009 a febrero del 2010, debido a que con estos datos estadísticos, vamos a estimar los parámetros del modelo, para las primeras 46 semanas de la enfermedad.

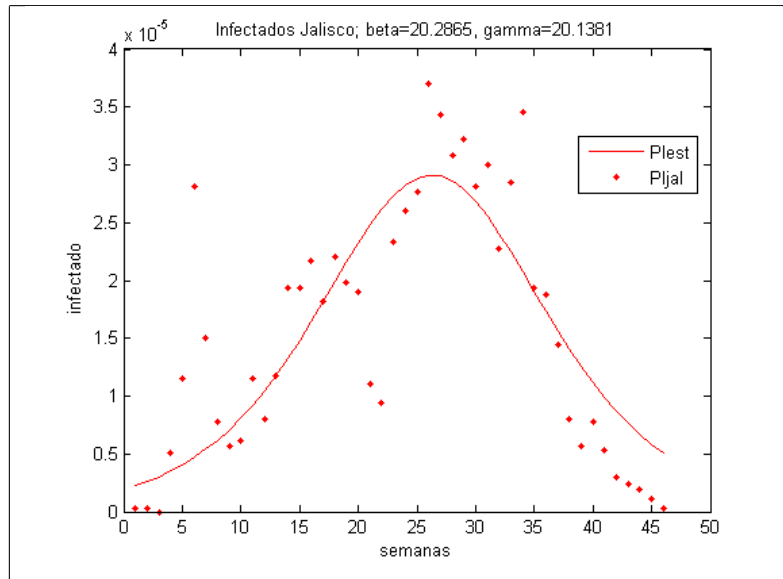


Figura 4.1: Infectados en Jalisco

β	γ	R_0
20.286545077602579	20.138063987550360	1.007373156135764
Intervalo de confianza 95 %	error absoluto	error relativo
(1.007361536328892,1.007384776210068)	4.426772577887333e-005	0.344356611228278

Cuadro 4.1: Parámetros estimados para la propagación durante 46 semanas utilizando evolución diferencial

Para estimar los parámetros β y γ del modelo SIR epidémico, que se propone para modelar las infecciones de enfermedades respiratorias en Estados Unidos, del 7 de noviembre del 2004 al 25 de septiembre del 2005, es decir. 46 semanas. Vamos a considerar las estimaciones de los casos positivos en estos países para una población de 100000 habitantes, estos datos se encuentran disponibles en Google flu trends [34]. Obteniendo los siguientes resultados, cuadro (4.2).

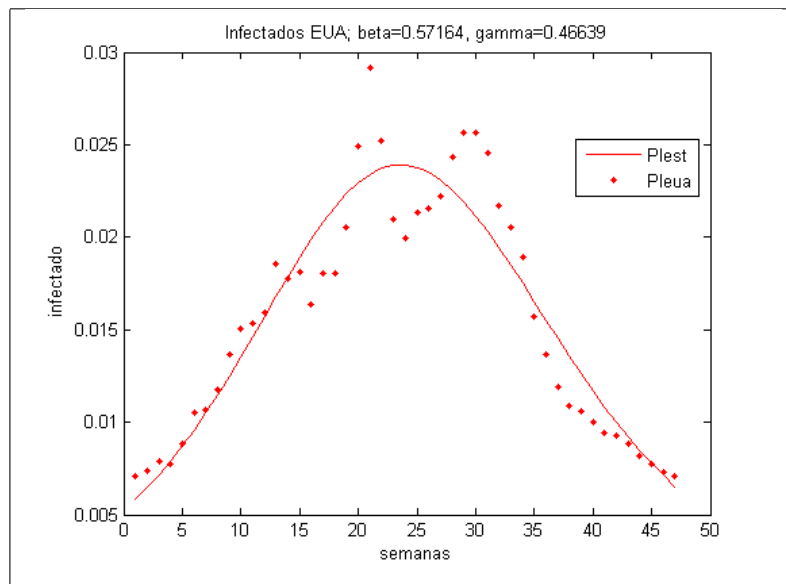


Figura 4.2: Infectados en EUA

β	γ	R_0
0.571641744500185	0.466385978965574	1.225683811867724
Intervalo de confianza 95 %	error absoluto	error relativo
(0.993320021164155, 1.598000450127583)	0.114554623447735	8.911151688262218e+002

Cuadro 4.2: Parámetros estimados para la propagación durante 46 semanas utilizando evolución diferencial

Al estimar los parámetros utilizando evolución diferencial en el modelo epidémico, observamos un mejor ajuste en los datos en comparación con el modelo y los métodos de estimación utilizados en el capítulo anterior.

Capítulo 5

Adaptación espacial del modelo

Para un gran número de personas infecciosas, porque no toda la población es susceptible a la infección. Una vez que no hay gente susceptible que se puede infectar, no puede haber nuevas infecciones, esta expresión es

$$\Delta I \propto \frac{\beta}{P} SI \quad (5.1)$$

donde P es la población total y $\frac{S}{P}$ es la fracción de la población susceptible a la infección. La ecuación (5.1) es un término de interacción o producto de interacción, porque representa la interacción entre la gente infectada y la gente susceptible como un simple producto. Este producto de interacción determina la nueva incidencia de la infección.

Si la ecuación (5.1) se utiliza para definir la tasa a la cual la gente infectada en una población particular j causa que la gente susceptible en la misma población j sea infectada, entonces podemos escribir (5.1) como,

$$\Delta I_j \propto \beta \frac{S_j}{P_j} I_j \quad (5.2)$$

Sea la tasa de infección en la población j mezclando con la población k

$$m_{jk} \beta \frac{I_k}{\sum_{k=1}^K m_{jk} P_k} \quad (5.3)$$

Entonces la tasa de nuevas infecciones (tasa de incidencia) en la población j es

$$S \frac{\beta}{\sum_{k=1}^K m_{jk} P_k} \sum_{k=1}^K m_{jk} I_k \quad (5.4)$$

donde

- m_{jk} es la tasa de mezcla entre j y k (note que $m_{jj} = 1$)
- I_k es el número de infectados en la población k

- k es el número total de poblaciones
- P_k es la población en el lugar k
- β es la tasa de transmisión

5.1. Adaptación espacial del modelo

El modelo SIR se puede adaptar a poblaciones espacialmente distribuidas. Esto relaja la hipótesis de que las poblaciones están en un solo lugar y abre la posibilidad de que diferentes lugares tengan diferentes áreas, número de pobladores y tasas de transmisión.

Para dar solución a esta situación se usan versiones no normalizadas del modelo SIR (1.10) tal como sigue. Sea P_l el número de individuos en la población en el lugar l (Nótese que $P_l = S_l + I_l + R_l$). Es necesario tomar en cuenta la variabilidad de la tasa de transmisión β debida a las posibles diferencias en la densidad poblacional. Para ello se supone una tasa específica que denotamos por β_l .

Sustituyendo β por β_l multiplicando ambos miembros de el sistema (1.10) por P_l , y tomando la tasa de perdida de inmunidad $\delta = 0$ y la tasa de muertes debido a la enfermedad $\mu_i = 0$ obtenemos

$$\frac{ds}{dt}P_l = \mu P_l - \beta_l s i P_l - \mu s P_l \quad (5.5a)$$

$$\frac{di}{dt}P_l = \beta_l s i P_l - (\gamma + \mu i) P_l \quad (5.5b)$$

$$\frac{dr}{dt}P_l = \gamma i P_l - \mu r P_l \quad (5.5c)$$

Se utiliza el valor de β_l como β escalándolo por la razón entre la densidad de la población en el lugar l y el promedio de la densidad de población de todas las locaciones [35]:

$$\beta_l = \beta d_l / APD \quad (5.6)$$

donde d_l es la densidad local de la población en el lugar l y $APD = \frac{\sum_{l=1}^L d_l}{L}$ es el promedio de la densidad de población de todos los lugares, L es el total de lugares.

Sea $S_l = sP_l$ el número de miembros susceptibles de la población en el lugar l , sea $I_l = iP_l$ el número de miembros infecciosos de la población en el lugar l y sea $R_l = rP_l$ la población recuperada, al sustituir estas expresiones en el sistema (5.5) se obtiene,

$$\frac{dS_l}{dt} = \mu P_l - \beta_l S_l i - \mu S_l \quad (5.7a)$$

$$\frac{dI_l}{dt} = \beta_l S_l i - (\gamma + \mu) I_l \quad (5.7b)$$

$$\frac{dR_l}{dt} = \gamma I_l - \mu R_l \quad (5.7c)$$

sustituyendo $i = I/P_l$, en (5.7) obtenemos

$$\frac{dS_l}{dt} = \mu P_l - (\beta_l/P_l) S_l I_l - \mu S_l \quad (5.8a)$$

$$\frac{dI_l}{dt} = (\beta_l/P_l) S_l I_l - (\gamma + \mu) I_l \quad (5.8b)$$

$$\frac{dR_l}{dt} = \gamma I_l - \mu R_l \quad (5.8c)$$

Sustituyendo la tasa de nuevas infecciones, es decir, tasa de incidencia en la población l , ecuación (5.4) en (5.9),

$$\frac{dS_l}{dt} = \mu P_l - \beta_l S_l \frac{\sum_{k=1}^K m_{lk} I_k}{\sum_{k=1}^K m_{lk} P_k} - \mu S_l \quad (5.9a)$$

$$\frac{dI_l}{dt} = \beta_l S_l \frac{\sum_{k=1}^K m_{lk} I_k}{\sum_{k=1}^K m_{lk} P_k} - \gamma I_l - \mu I_l \quad (5.9b)$$

$$\frac{dR_l}{dt} = \gamma I_l - \mu R_l \quad (5.9c)$$

5.2. Poblaciones vecinas infecciosas

La adyacencia física de las poblaciones conduce naturalmente a contacto población a población y eventualmente a la diseminación de la enfermedad. Podemos extender el modelo incorporando este aspecto de una población espacialmente distribuida como sigue:

Dos lugares están relacionados positivamente cuando es posible el intercambio de miembros de las poblaciones mediante algún sistema de transporte como caminos, carreteras, trenes o aviones. En este caso es posible que los miembros de las poblaciones de infecciosos visiten a miembros susceptibles de otras poblaciones. Se supone que los visitantes tienen el mismo nivel de contacto infeccioso que los miembros de la población visitada, es decir, que se pueden contar como miembros de la población en el lugar en cuestión.

Las ecuaciones se modifican de la siguiente manera:

$$\frac{dS_l}{dt} = \mu P_l - \beta_l S_l I_{vecinos} - \mu S_l \quad (5.10a)$$

$$\frac{dI_l}{dt} = \beta_l S_l I_{vecinos} - (\gamma + \mu) I_l \quad (5.10b)$$

$$\frac{dR_l}{dt} = \gamma I_l - \mu R_l \quad (5.10c)$$

donde

- $I_{vecinos} = \frac{\sum_{k=1}^K m_{lk} I_k}{\sum_{k=1}^K m_{lk} P_k}$ es el número de visitantes infecciosos del vecino k .
- m_{lk} es una constante que depende del tipo de relación positiva; si las poblaciones son vecinas, i.e. comparten una frontera, si están conectados por carreteras, entre otros.
- P_k es el número de individuos en la población vecina k .
- I_k es el número de individuos infecciosos de la población vecina k .

Modificando este modelo con términos apropiados, se pueden calcular estadísticas específicas del total de nacimientos, muertes y muertes debidas a la enfermedad, por ejemplo:

- $B = \mu(S + I + R)$, el número de nacimientos.
- $D = \mu S + (\mu + \mu_i)I + \mu R$, el número de muertes.
- $DD = \mu_i I$, el número de muertes debidas a la enfermedad.

5.2.1. Ejemplo de la adaptación espacial para una estructura de tres vecinos

Tomemos como ejemplo una estructura de tres vecinos, denotados por $l = 1, 2, 3$ con $k = 1, 2, 3$. En donde $P_k^t = S_k^t + I_k^t + R_k^t$ es la población en cada vecindario en el paso del tiempo, así, formemos el modelo con la siguiente estructura,

$$\frac{dS_l}{dt} = \mu P_l - \beta_l S_l I_{vecinos} - \mu S_l \quad (5.11a)$$

$$\frac{dI_l}{dt} = \beta_l S_l I_{vecinos} - (\gamma + \mu) I_l \quad (5.11b)$$

$$\frac{dR_l}{dt} = \gamma I_l - \mu R_l \quad (5.11c)$$

Si tenemos que $I_k = P_k i_k$ y al denotar $I_{vecinos} = \frac{\sum_{k=1}^K m_{lk} I_k}{\sum_{k=1}^K m_{lk} P_k}$ como

$$I_{vecinos} = \frac{m_{l1}I_1 + m_{l2}I_2 + m_{l3}I_3}{m_{l1}P_1 + m_{l2}P_2 + m_{l3}P_3}$$

Hay que notar que cuando $l = k$ entonces m_{ll} es igual a 1 y si denotamos

$$a_{l1} = \frac{m_{l1}}{m_{l1}P_1 + m_{l2}P_2 + m_{l3}P_3}$$

$$a_{l2} = \frac{m_{l2}}{m_{l1}P_1 + m_{l2}P_2 + m_{l3}P_3}$$

y

$$a_{l3} = \frac{m_{l3}}{m_{l1}P_1 + m_{l2}P_2 + m_{l3}P_3}$$

llegamos a la siguiente expresión:

$$I_{vecinos} = I_1a_{l1} + I_2a_{l2} + I_3a_{l3}$$

sustituyendo en el sistema (5.23) obtenemos

$$S'_l = \mu P_l - \beta_l S_l (I_1 a_{l1} + I_2 a_{l2} + I_3 a_{l3}) - \mu S_l \quad (5.12a)$$

$$I'_l = \beta_l S_l (I_1 a_{l1} + I_2 a_{l2} + I_3 a_{l3}) - (\gamma + \mu) I_l \quad (5.12b)$$

$$R'_l = \gamma I_l - \mu R_l \quad (5.12c)$$

Si tomamos $l = 1, 2, 3$ el sistema (5.12) se convierte en tres sistemas, uno para cada población l

Proposición 1.

$$P'_k = S'_k + I'_k + R'_k = 0$$

Demostración: Tomando la tasa neta de mortalidad en la ecuación (5.12b) es decir, $\mu = (\mu + \mu_i)$

$$\begin{aligned} S'_k + I'_k + R'_k &= \mu P_l - \beta_k^* S_k (I_k + I_l a_{kl} + I_m a_{km}) - \mu S_k \\ &\quad + \beta_k^* S_k (I_k + I_n a_{kl} + I_m a_{km}) - \gamma I_k - \mu I_k \\ &\quad + \gamma I_k - \mu R_k \\ &= \mu P_l - \mu S_k - \mu I_k - \mu R_k \\ &= \mu P_l - (\mu S_k + \mu I_k + \mu R_k) \\ &= 0 \end{aligned}$$

□

Corolario 1. $P_k = S_k + I_k + R_k = ctte.$

Este hecho muestra que las poblaciones vecinas permanecerán constantes al paso del tiempo.

5.3. Adaptación espacial para la estructura de vecinos de la Zona Metropolitana de Guadalajara

Tomemos una estructura de vecinos para los 8 municipios de la ZMG

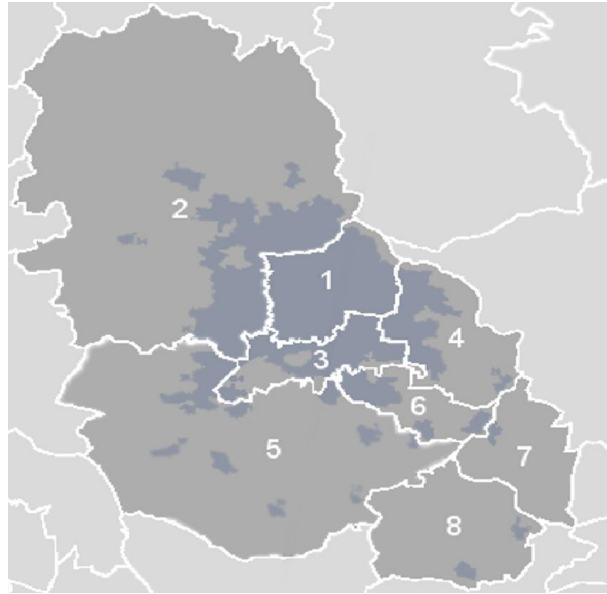


Figura 5.1: Zona Metropolitana de Guadalajara

en la figura (5.1) los números representan respectivamente a los municipios

- I. Guadalajara
- II. Zapopan
- III. Tlaquepaque
- IV. Tonalá
- V. Tlajomulco
- VI. El Salto
- VII. Juanacatlán
- VIII. Ixtlahuacán de los Membrillos

Para fines prácticos reordenaremos la lista en orden alfabético

- I. Guadalajara
- II. Ixtlahuacán de los Membrillos

- III. Juanacatlán
- IV. Salto El
- V. Tlajomulco
- VI. Tlaquepaque
- VII. Tonalá
- VIII. Zapopan

Definamos las áreas de cada municipio en kilómetros cuadrados como A_k

$$A_{Gdl}, A_{Ixt}, A_{Jua}, A_{Sal}, A_{Tlaj}, A_{Tlaq}, A_{Ton}, A_{Zap}$$

y el área total de la ZMG como la suma de las áreas de todos los municipios como

$$A_{Total} = A_{Gdl} + A_{Ixt} + A_{Jua} + A_{Sal} + A_{Tlaj} + A_{Tlaq} + A_{Ton} + A_{Zap}$$

El número de habitantes de cada municipio como P_k

$$P_{Gdl}, P_{Ixt}, P_{Jua}, P_{Sal}, P_{Tlaj}, P_{Tlaq}, P_{Ton}, P_{Zap}$$

Y el número total de habitantes de la ZMG como

$$P_{Total} = P_{Gdl} + P_{Ixt} + P_{Jua} + P_{Sal} + P_{Tlaj} + P_{Tlaq} + P_{Ton} + P_{Zap}$$

La densidad de población para cada municipio es el cociente entre el número de habitantes y el área del municipio, $d_k = \frac{P_k}{A_k}$

El factor de mezcla m_{lk} se toma como la diferencia entre viajes atraídos en cada municipio menos los viajes generados, tomado como base los desplazamientos de la ZMG [39].

Tomando como base el ejemplo de estructura para tres vecinos, formemos la estructura de vecinos para la ZMG, denotando por

$$l = Gdl, Ixt, Jua, Sal, Tlaj, Tlaq, Ton, Zap$$

y

$$k = Gdl, Ixt, Jua, Sal, Tlaj, Tlaq, Ton, Zap$$

a cada municipio, en donde k corresponde a los vecinos del municipio y $P_k^t = S_k^t + I_k^t + R_k^t$ es la población en cada vecindario en el paso del tiempo, así, formemos el modelo con la siguiente estructura,

$$I_{vecinos} = \frac{m_{lGdl}I_{Gdl} + m_{lIxt}I_{Ixt} + m_{lJua}I_{Jua} + m_{lSal}I_{Sal} + m_{lTlaj}I_{Tlaj} + m_{lTlaq}I_{Tlaq} + m_{lTon}I_{Ton} + m_{lZap}I_{Zap}}{m_{lGdl}P_{Gdl} + m_{lIxt}P_{Ixt} + m_{lJua}P_{Jua} + m_{lSal}P_{Sal} + m_{lTlaj}P_{Tlaj} + m_{lTlaq}P_{Tlaq} + m_{lTon}P_{Ton} + m_{lZap}P_{Zap}} \quad (5.13)$$

Hay que notar que cuando $l = k$ entonces m_{ll} es igual a 1 y si denotamos

$$a_{lGdl} = \frac{m_{lGdl}}{m_{lGdl}P_{Gdl} + m_{lIxt}P_{Ixt} + m_{lJua}P_{Jua} + m_{lSal}P_{Sal} + m_{lTlaj}P_{Tlaj} + m_{lTlaq}P_{Tlaq} + m_{lTon}P_{Ton} + m_{lZap}P_{Zap}} \quad (5.14)$$

$$a_{lIxt} = \frac{m_{lIxt}}{m_{lGdl}P_{Gdl} + m_{lIxt}P_{Ixt} + m_{lJua}P_{Jua} + m_{lSal}P_{Sal} + m_{lTlaj}P_{Tlaj} + m_{lTlaq}P_{Tlaq} + m_{lTon}P_{Ton} + m_{lZap}P_{Zap}} \quad (5.15)$$

$$a_{lJua} = \frac{m_{lJua}}{m_{lGdl}P_{Gdl} + m_{lIxt}P_{Ixt} + m_{lJua}P_{Jua} + m_{lSal}P_{Sal} + m_{lTlaj}P_{Tlaj} + m_{lTlaq}P_{Tlaq} + m_{lTon}P_{Ton} + m_{lZap}P_{Zap}} \quad (5.16)$$

$$a_{lSal} = \frac{m_{lSal}}{m_{lGdl}P_{Gdl} + m_{lIxt}P_{Ixt} + m_{lJua}P_{Jua} + m_{lSal}P_{Sal} + m_{lTlaj}P_{Tlaj} + m_{lTlaq}P_{Tlaq} + m_{lTon}P_{Ton} + m_{lZap}P_{Zap}} \quad (5.17)$$

$$a_{lTlaj} = \frac{m_{lTlaj}}{m_{lGdl}P_{Gdl} + m_{lIxt}P_{Ixt} + m_{lJua}P_{Jua} + m_{lSal}P_{Sal} + m_{lTlaj}P_{Tlaj} + m_{lTlaq}P_{Tlaq} + m_{lTon}P_{Ton} + m_{lZap}P_{Zap}} \quad (5.18)$$

$$a_{lTlaq} = \frac{m_{lTlaq}}{m_{lGdl}P_{Gdl} + m_{lIxt}P_{Ixt} + m_{lJua}P_{Jua} + m_{lSal}P_{Sal} + m_{lTlaj}P_{Tlaj} + m_{lTlaq}P_{Tlaq} + m_{lTon}P_{Ton} + m_{lZap}P_{Zap}} \quad (5.19)$$

$$a_{lTon} = \frac{m_{lTon}}{m_{lGdl}P_{Gdl} + m_{lIxt}P_{Ixt} + m_{lJua}P_{Jua} + m_{lSal}P_{Sal} + m_{lTlaj}P_{Tlaj} + m_{lTlaq}P_{Tlaq} + m_{lTon}P_{Ton} + m_{lZap}P_{Zap}} \quad (5.20)$$

$$a_{lZap} = \frac{m_{lZap}}{m_{lGdl}P_{Gdl} + m_{lIxt}P_{Ixt} + m_{lJua}P_{Jua} + m_{lSal}P_{Sal} + m_{lTlaj}P_{Tlaj} + m_{lTlaq}P_{Tlaq} + m_{lTon}P_{Ton} + m_{lZap}P_{Zap}} \quad (5.21)$$

llegamos a la siguiente expresión:

$$I_{vecinos} = I_{Gdl}a_{lGdl} + I_{Ixt}a_{lIxt} + I_{Jua}a_{lJua} + I_{Sal}a_{lSal} + I_{Tlaj}a_{lTlaj} + I_{Tlaq}a_{lTlaq} + I_{Ton}a_{lTon} + I_{Zap}a_{lZap} \quad (5.22)$$

que podemos sustituir en la siguiente ecuación

$$\frac{dS_l}{dt} = \mu P_l - \beta_l S_l I_{vecinos} - \mu S_l \quad (5.23a)$$

$$\frac{dI_l}{dt} = \beta_l S_l I_{vecinos} - (\gamma + \mu) I_l \quad (5.23b)$$

$$\frac{dR_l}{dt} = \gamma I_l - \mu R_l \quad (5.23c)$$

Estructura de vecinos para el municipio de Guadalajara

$$\frac{dS_{Gdl}}{dt} = \mu P_{Gdl} - \beta_{Gdl} S_{Gdl} I_{vecinosGdl} - \mu S_{Gdl} \quad (5.24a)$$

$$\frac{dI_{Gdl}}{dt} = \beta_{Gdl} S_{Gdl} I_{vecinosGdl} - (\gamma + \mu) I_{Gdl} \quad (5.24b)$$

$$\frac{dR_{Gdl}}{dt} = \gamma I_{Gdl} - \mu R_{Gdl} \quad (5.24c)$$

donde, la tasa de transmisión es

$$\beta_{Gdl} = \frac{P_{Gdl}}{A_{Gdl}}$$

e

$$I_{vecinosGdl} = I_{Gdl}^{\alpha_{GdlGdl}} + I_{Ixt}^{\alpha_{GdlIxt}} + I_{Jua}^{\alpha_{GdlJua}} + I_{Sal}^{\alpha_{GdlSal}} + I_{Tlaj}^{\alpha_{GdlTlaj}} + I_{Tlaq}^{\alpha_{GdlTlaq}} + I_{Ton}^{\alpha_{GdlTon}} + I_{Zap}^{\alpha_{GdlZap}} \quad (5.25)$$

Siguiendo esta metodología, formamos la estructura de vecinos para cada municipio de la ZMG

5.4. Implementación computacional de la estructura de vecinos de la Zona Metropolitana de Guadalajara

En esta sección implementaremos las estructuras de vecinos de todos los municipios de la ZMG en un mismo algoritmo para simular la propagación de la enfermedad al paso del tiempo, tomando en cuenta periodos semanales, además de considerar los parámetros β y γ que estimamos en el capítulo anterior.

Tomamos en cuenta las estadísticas el consejo estatal de población COEPO del estado del Jalisco [40] y los desplazamientos entre los municipios [39] para adecuar nuestro modelo a la ZMG.

Tomemos el área en kilómetros cuadrados para cada municipio de la ZMG

- I. Guadalajara 187.91
- II. Ixtlahuacán de los Membrillos 184.25
- III. Juanacatlán 89.08
- IV. Salto El 41.5
- V. Tlajomulco 636.93
- VI. Tlaquepaque 270.88
- VII. Tonalá 119.58
- VIII. Zapopan 893.15

que en suma nos dan un total de 2423.3 kilómetros cuadrados. También necesitamos el número de habitantes de cada municipio

- I. Guadalajara 1 600 940

- II. Ixtlahuacán de los Membrillos 23 420

- III. Juanacatlán 11 902

- IV. Salto El 111 436

- V. Tlajomulco 220 630

- VI. Tlaquepaque 563 006

- VII. Tonalá 408 729

- VIII. Zapopan 1 155 790

que en total son 4 095 853 habitantes

Estos datos nos permitirán calcular el parámetro β_l para cada municipio, debido a la variabilidad de la densidad de población para cada uno de ellos.

Para hacer la simulación utilizamos MATLAB y creamos un ambiente gráfico de modo que al dar valores a los parámetros, el número de semanas de la epidemia que deseamos modelar así como la clase ya sea susceptible, infeccioso o removido, obtenemos los resultados conforme se avanza la enfermedad en el tiempo. La ventaja de hacerlo en un ambiente gráfico es que podemos variar los parámetros con los que se desea modelar como se muestra en la figura (5.2)

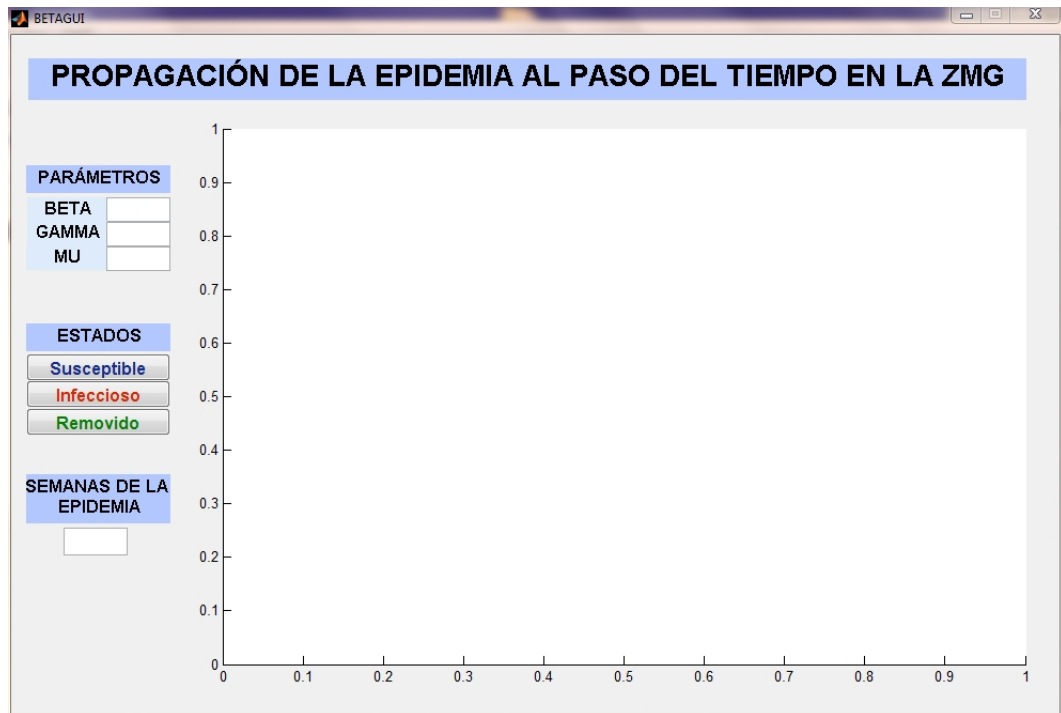


Figura 5.2: Ambiente Gráfico

para ejemplificar tomemos los parámetros que estimamos en el capítulo anterior de la figura (3.4), $\beta = 2.0414$, $\gamma = 1.2276$, $\mu = 0.00025$ y veamos la propagación a las 10, 20, 40 y 80 semanas.

La clase susceptible se muestra en la figura (5.3).

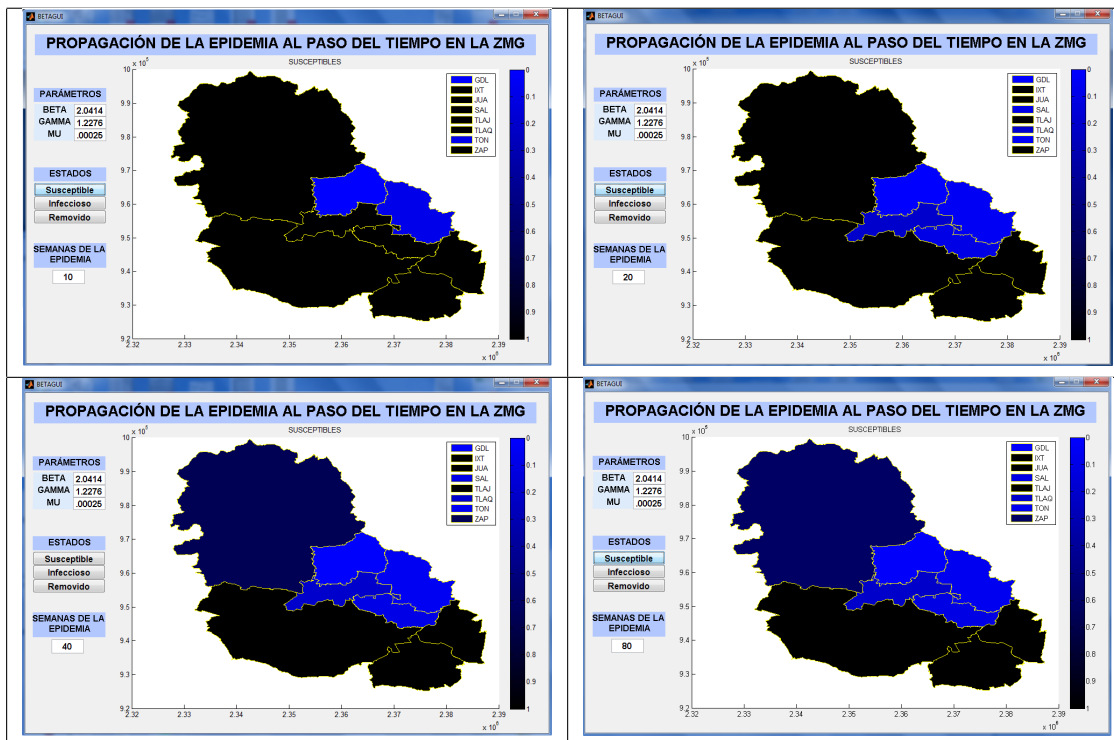


Figura 5.3: Propagación de la enfermedad para la clase susceptible de la semana 10 a la 80

La clase de infectados en la figura (5.4).

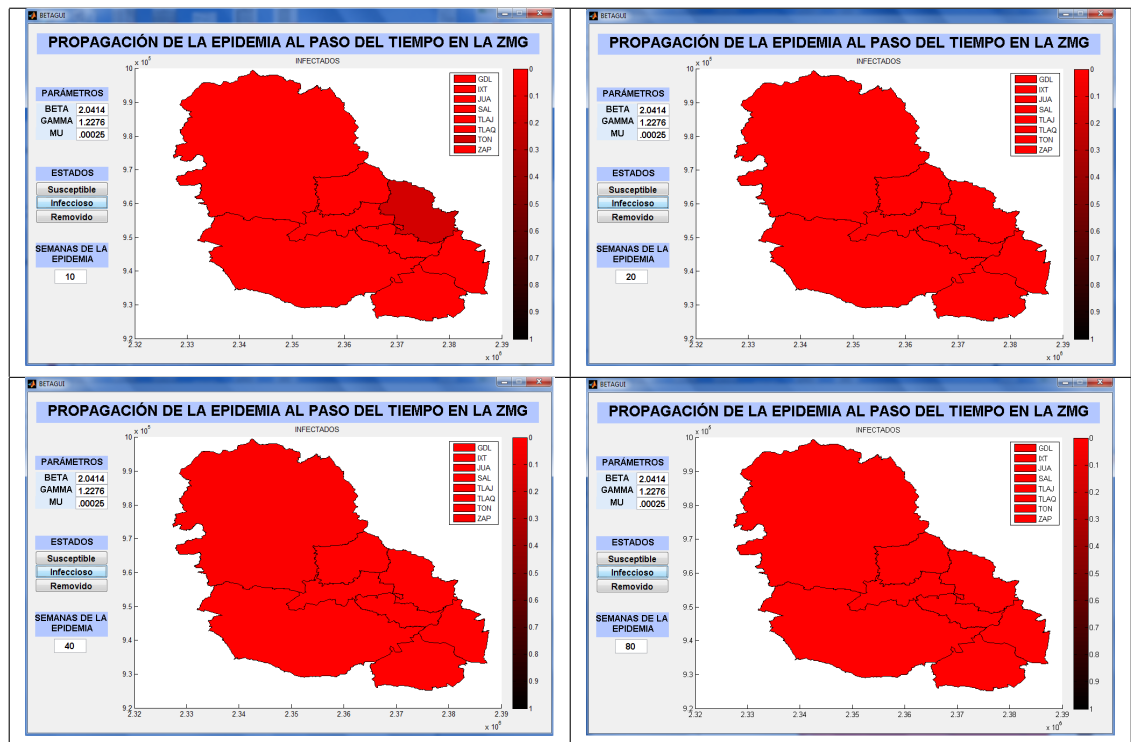


Figura 5.4: Propagación de la enfermedad para la clase infecciosa de la semana 10 a la 80

Y las personas que se han recuperado de la enfermedad, figura (5.5).

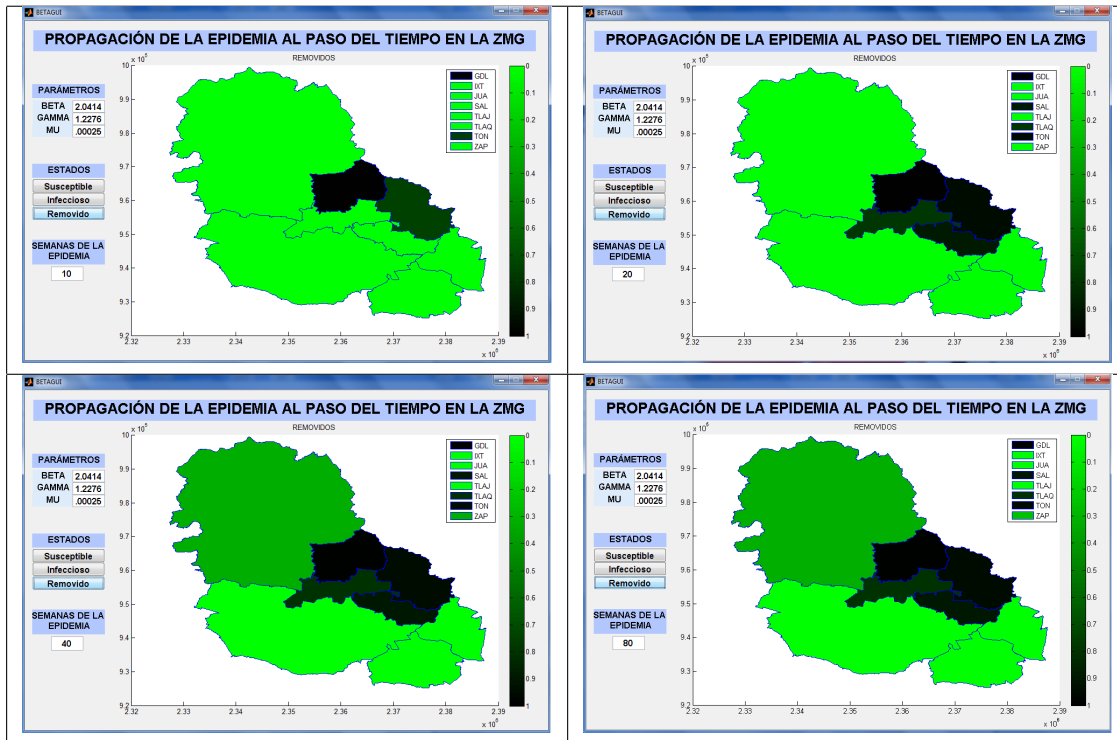


Figura 5.5: Propagación de la enfermedad para la clase de removidos de la semana 10 a la 80

Este ambiente gráfico nos permite observar como se propaga la epidemia semana a semana, debido a que podemos ver como cambian los tonos de los colores acorde a la clase que representan, el verde a las personas propensas a contraer la enfermedad susceptibles, el rojo a las personas que han adquirido la enfermedad, los infectados y el azul a las personas que se han recuperado después de contraer la enfermedad, los removidos o recuperados.

Los diferentes tonos nos representan el número de personas que se encuentran semana a semana en cada clase, por eso podemos ver en las figuras que no todos los municipios tienen la misma tonalidad pero si el mismo color, entre más oscuro sea el color significa que más personas están en la clase de la enfermedad, por ejemplo, en la gráfica para las personas infectadas que esta en color rojo, podemos ver como los municipios de la parte inferior que representan a Ixtlahuacán de los Membrillos y a Juanacatlán se encuentran en un tono rojizo muy claro, esto significa que pocas personas han adquirido la enfermedad, en cambio el municipio del centro que es Guadalajara tiene un tono rojo oscuro y esto significa que un mayor número de habitantes ha contraído la enfermedad.

En resumen para las tres clases, si el tono del color que corresponde a cada uno de ellos es más oscuro, significa que hay más personas en la clase que cuando la tonalidad

es más clara, esto se cumple para todos los municipios y para todas las clases .

Conclusiones

Podemos utilizar STEM para darnos una idea de como se propaga una enfermedad por el territorio nacional y hacer comparaciones entre diferentes escenarios y diferentes enfermedades, tenemos que utilizar parámetros conocidos que se encuentran en base a la experiencia o con una estimación.

Cuando utilizamos estimación de parámetros por máxima verosimilitud una de las dificultades principales fue que la matriz de covarianza es muy cercana a ser singular, por lo que para invertirla necesitamos un método de regularización, la desventaja de utilizar regularización es que se introduce ruido, que al ser muy pequeño obtenemos buenos resultados. Utilizamos un algoritmo para obtener el mejor parámetro de regularización, pero aun así, necesitamos más tiempo maquina porque se debe invertir la matriz cada vez que se calculan los gradientes y la función de costo.

Al utilizar la distribución χ^2 , solo utilizamos el vector de errores por lo que se disminuye la complejidad de invertir la matriz de covarianza. Cuando utilizamos el método SOR la desventaja es que necesitamos estar muy cerca del mínimo para que funcione el algoritmo y la convergencia es muy lenta, al usar GC no lineal obtenemos mejores resultados y una convergencia mas rápida.

Si usamos un modelo SIR sin dinámica vital y el algoritmo de evolución diferencial, los datos tienen mejor ajuste que con los modelos descritos anteriormente, mas no muestran de manera precisa el comportamiento del virus en el estado de Jalisco en el tiempo de estudio.

En cuanto al modelo SIR con y sin dinámica vital, obtenemos mejores aproximaciones al modelar la influenza por períodos largos de tiempo. Al observar los datos originales y compararlos con los resultados obtenidos, podemos ver que el modelo no muestra un comportamiento típico como se puede observar en los datos proporcionados por la SSJ, una explicación es que la enfermedad fue inesperada y hubo tres brotes significativos en un período corto de tiempo, mientras que los modelos SIR se utilizan regularmente por períodos largos. En el artículo [10] menciona, que para obtener un mejor ajuste a los datos de la influenza A(H1N1), se necesitan separar las personas infectadas por cohortes de edad, siendo contrario a lo que se pensaba, que se necesita aislar como medida preventiva a los niños y ancianos porque son los

mas vulnerables, las personas en edad productiva son las que están mas expuestas a infectarse con el virus y así mismo a propagarlo. Debido a la escases de vacunas, se tomo como prioridad vacunar a niños y ancianos, que en la practica, seria mejor vacunar a las poblaciones jóvenes que tienen mayor movilidad.

Podemos utilizar el modelo SIR para describir enfermedades respiratorias que tienen un comportamiento típico, en donde el invierno juega un papel importante en las personas que se pueden enfermar, como sucede en los ejemplos para Bélgica y Francia, donde es más sencillo y observamos mejores resultados al estimar los parámetros que para la influenza en México.

Al implementar el modelo espacial SIR en conjunto con el demo de la ZMG, observamos como se propaga el virus de la influenza A(H1N1) suponiendo que la enfermedad se adapta a un modelo SIR, para adaptar este modelo de una manera mas precisa, necesitamos un modelo que se aproxime a los datos reales tal como se propone en la sección de estimación de parámetros.

Apéndice A

Estimación de parámetros para un modelo epidemiológico SIQR

Otra manera de abordar el problema de la propagación del virus de la influenza A(H1N1) en la Zona Metropolitana de Guadalajara, es utilizando un modelo SIR para las primeras seis semanas de la enfermedad y un modelo SIQR propuesto en [29] para las semanas siguientes, para explicar el brote durante las primeras 47 semanas de la epidemia. Vamos a utilizar la distribución χ^2 y el método de sobrerelajación sucesiva SOR que vimos en el capítulo 3.

A.1. Modelo SIQR

El modelo SIQR es útil para modelar enfermedades cuando algunos individuos se encuentran en la clase de personas aisladas.

$$S' = \mu N - \beta S \frac{I}{N - Q} - \mu S \quad (\text{A.1a})$$

$$I' = \beta S \frac{I}{N - Q} - (\gamma + \mu)I \quad (\text{A.1b})$$

$$Q' = \gamma I - (\mu + \delta)Q \quad (\text{A.1c})$$

$$R' = -\mu R + \delta Q \quad (\text{A.1d})$$

donde

- S es la población susceptible
- I es la población infectada
- Q es la población en cuarentena
- R es la población removida
- μ es la tasa de mortalidad per-cápita

– β es la tasa de infección per-cápita, el promedio de personas que susceptibles que se infectan por cada infectado.

– γ es la tasa a la cual los individuos se recuperan de la infección

– δ es la tasa a la cual los individuos entran a la clase de aislados

A.2. Resultados

Para estimar los parámetros del modelo SIQR, utilizamos la distribución χ^2 y el método SOR, pero vamos a dejar fijos los valores para de $\mu = 0.00015$, $\beta = 1.692349916183763$ y $\gamma = 0.838567691345682$ estimados para el modelo SIR y solo vamos a calcular δ para los intervalos semanales de 6 a 10, 10 a 22 y 22 a 47 semanas.

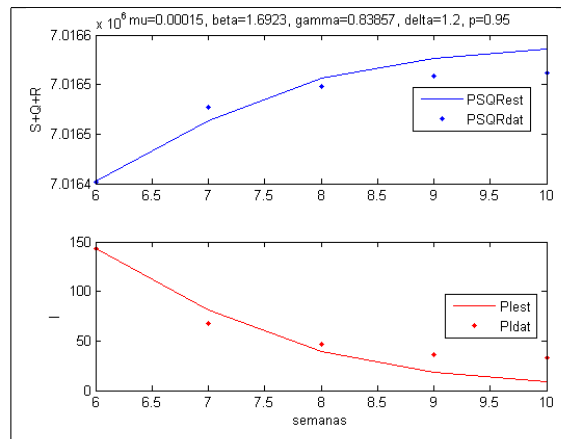


Figura A.1: Propagación de la semana 6 a la 10 para el modelo SIQR

Debido al paro de labores y actividades básicas en la ciudad y al período de vacaciones de semana santa, suponemos que la proporción de la población que estuvo aislada durante 4 semanas en sus hogares fue del 95 %, al tomar esta proporción y los parámetros $\mu = 0.00015$, $\beta = 1.692349916183763$ y $\gamma = 0.838567691345682$, estimamos $\delta = 1.199999564613630$. Al tomar estos parámetros observamos que hubo una disminución en el número de personas que fueron infectadas con el virus, figura (A.1).

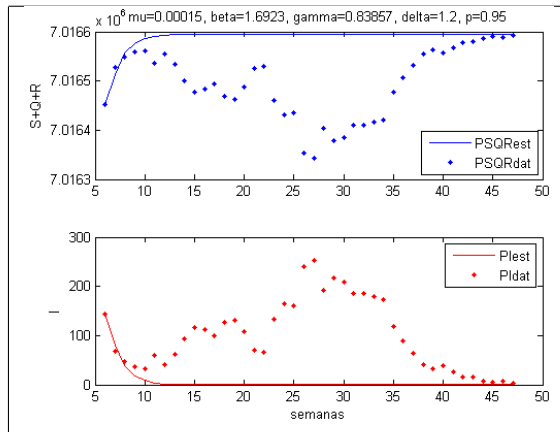


Figura A.2: Propagación de la semana 6 a la 47 para el modelo SIQR

Si suponemos que toda la población está en cuarentena de la semana 6 a la 47, observamos que la enfermedad tiende a desaparecer conforme pasa el tiempo, pero es imposible que se paren las actividades por tan largo período de tiempo para que ocurra este fenómeno, figura (A.2).

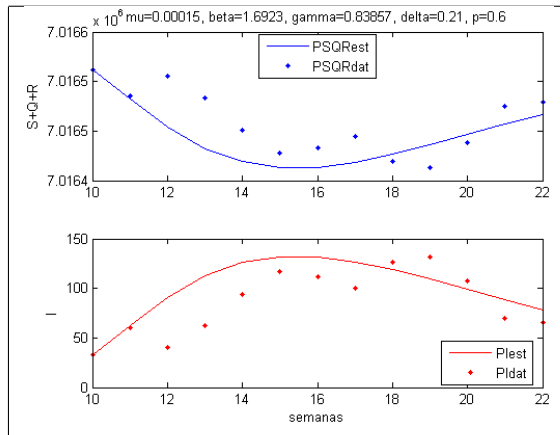


Figura A.3: Propagación de la semana 10 a la 22 para el modelo SIQR

Cuando paso la alerta epidemiológica, las personas regresaron a sus labores normales por lo que la proporción de personas aisladas disminuyó. Al modelar la propagación de la semana 10 a la 22 que comprende el periodo de vacaciones de verano, suponemos que la proporción de personas aisladas es del 60%, obteniendo la estimación del parámetro $\delta=2.1$, figura (A.3).

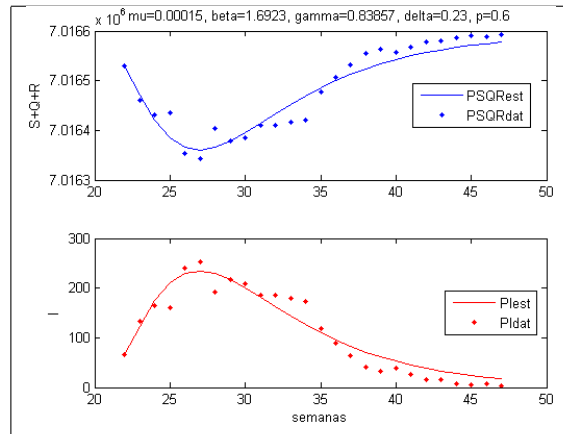


Figura A.4: Propagación de la semana 22 a la 47 para el modelo SIQR

Si modelamos la propagación de la semana 22 a la 47 que comprende el período invernal y sus vacaciones, suponemos que la proporción de personas aisladas es del 60 %, obteniendo el parámetro $\delta = 2.3$, figura (A.4).

A.3. Conclusiones del apéndice

Para estimar los parámetros este modelo, nuestros parámetros iniciales necesitan estar cerca del mínimo de la función de costo para poder hacer uso del método de sobrerelajación sucesiva SOR.

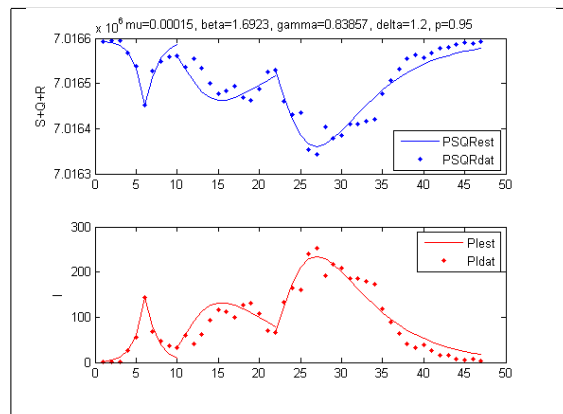


Figura A.5: Propagación de la semana 1 a la 47 para el modelo SIR/SIQR

Para modelar la propagación del virus de la influenza A(H1N1) en la ZMG y comparar con los datos reales proporcionados por la secretaria de salud Jalisco, necesitamos utilizar un modelo SIR para las primeras 6 semanas de la enfermedad y algunos modelos SIQR para las siguientes, debido a los diversos episodios que sucedieron, que fueron la alerta epidemiológica, vacaciones y actividades normales figura (A.5). El

modelo por partes muestra una buena aproximación a los datos reales. Hay que notar que en [29] proponen el modelo SIQR para modelar el virus de la influenza, pero no necesariamente es el mejor modelo para explicar la propagación de este tipo de virus.

Bibliografía

- [1] W. Ahrens, K. Krickeberg, I. Pidgeot, *An Introduction to Epidemiology*, Handbook of Epidemiology, Springer Berlin / Heidelberg 2005.
- [2] J.R. Banga, C.G. Moles, P. Mendes *Parameter estimation in biochemical pathways: a comparison of global optimization methods*, Genome Res. 2003 13: 2467-2474.
- [3] R. L. Burden and J. D. Faires, *Análisis numérico*, 7a. ed. Thomson Learning, 2002.
- [4] V. Capasso, *Mathematical structures of epidemic systems*, Springer Verlag, 1993.
- [5] <http://www.icsi.berkeley.edu/storn/code.html>
- [6] H. W. Engl, C. Flamm, P. Kuger, J. Lu, S. Muller, P. Schuster, *Inverse problems in systems biology*, Inverse Problems, 25(2009) 123014 (51pp).
- [7] G. Cruz-Pacheco, L. Duran, L. Esteva, A. A. Minzoni, M. López-Cervantes, P. Panayotaros, A. Ahued, I. Villaseñor, *Modelling of the influenza A(H1N1)v outbreak in Mexico city april-may 2009, with control sanitary measures*, Euro Surveill. 2009 Jul 2–14(26). pii: 19254
- [8] G. Cruz-Pacheco, L. Esteva, A.A. Minzoni, P. Panayotaros, N.F. Smyth, *Modelación matemática de la epidemia*, <http://www.fenomec.unam.mx/publicaciones/invest/panos/EpidemiaN09.pdf>
- [9] E. Chong and S. Zack, *An introduction to optimization*, 2nd Edition, John Wiley and Sons, Inc, 1996.
- [10] G Chowel, S. Bertozzi, M. A. Colchero, H. Lopez-Gatell, C. Alpuche-Aranda, M. Hernandez, M. A. Miller, *Severe respiratory disease concurrent with the circulation of H1N1 influenza*, N Engl J Med, 361, 2009.
- [11] B. David and G. Bastin, *An estimator of inverse covariance matrix and its application to ML parameter estimation in dynamical system*, Automatica, (37)99–106, 2001.

- [12] S. Edlund, M. Bromberg, G. Chodick, J. Douglas, D. Ford, Z. Kaufman, J. Lessler, R. Marom, Y. Mesika, R. Ram, V. Shalev, and J. Kaufman, *A Spatio Temporal Model for Influenza*, HIC 2009, Frontiers of Health Informatics, HIC.
- [13] J. J. Esqueda Elizondo, *Matlab e Interfaces Gráficas*, CONATEC 2002, Instituto Tecnológico de ciudad Madero.
- [14] D.A. Ford, J. H. Kaufman and I. Eiron, *An extensible spatial and temporal epidemiological modelling system*, International Journal of Health Geographics 2006, 5–4.
- [15] R. Fletcher, *Practical methods of optimization*, John Wiley and Sons, Inc, 1980.
- [16] C. Fraser, *Pandemic Potential of a Strain of Influenza A(H1N1): Early Findings*, Science 324, 1557–1561 (2009).
- [17] F. Brauer, P. V. Driessche, J. Wu, *Mathematical epidemiology*, Springer 2008.
- [18] P. C. Hansen, *The L-curve and its use in the numerical treatment of inverse problems*, DK-2800 Lyngby, Denmark.
- [19] H. W. Hethcote, *The Mathematics of Infectious Diseases*, SIAM 2000, vol. 42, no. 4, pp. 599–653.
- [20] H. W. Hethcote, *Qualitative analyses of communicable disease models*, Math. Bio, 28 (1976), pp. 335–356.
- [21] H. W. Hethcote, *Three basic epidemiological models*, In Applied Mathematical Ecology, Springer, 1989, pp. 119–144.
- [22] W. O. Kermack and A. G. McKendrick, *A Contribution to the Mathematical Theory of Epidemics*, Mathematical and Physical Character, Vol. 115, No. 772 (Aug. 1, 1927), pp. 700–721.
- [23] W. O. Kermack and A. G. McKendrick, *Contributions to the Mathematical Theory of Epidemics II problem of endemicity*, Bulletin of Mathematical Biology Vol. 53, No. 1/2, pp. 57–87, 1991.
- [24] A. Kremling, S. Fischer, K. Gadkar, F. J. Doyle, T. Sauter, E. Bullinger, F. Allgöwer and E. D. Gilles, *A Benchmark for Methods in Reverse Engineering and Model Discrimination: Problem Formulation and Solutions*, Genome Res. 14, 1773-1785, 2004.
- [25] E. R. Lecca, E.R. Lizama *Optimización por computación evolucionaria* .
- [26] Leon S. Lasdon, *Optimization theory for large systems* , McMillan, New York, 1970.

- [27] S. Ma, J Xia, *Mathematical understanding of infectious disease dynamics*, World Scientific Publishing, 2009.
- [28] J. Nocedal, S. J. Wright, *Numerical Optimization*, Springer Verlag, 1999.
- [29] M. Nuño, C. Castillo-Chavez, Z. Feng and M. Martcheva, *Mathematical Models of Influenza: The Role of Cross-Immunity, Quarantine and Age-Structure*, Lecture Notes in Mathematics, Springer 1945/2008, 349–364, 2008.
- [30] G. Rodriguez, *An algorithm for estimating the optimal regularization parameter by the L-curve*, Rendiconti di Matematica, Serie VII Volume 25, Roma (2005), 69-84.
- [31] S. Siltanen, *Inverse Problems*, February 20, 2009, <http://matwww.ee.tut.fi/courses/MAT-52500/IPnotes12.pdf>
- [32] R. Storn, K. Price, *Differential evolution-a simple and efficient heuristic for global optimization over continuous spaces*, Journal of Global Optimization 11, 341–359, 1997.
- [33] SOR method, <http://en.wikipedia.org/wiki/Successiveoverrelaxation>
- [34] <http://www.google.org/flutrends/>
- [35] *The Spatiotemporal Epidemiological Modeler (STEM) Project*, <http://www.eclipse.org/stem/>
- [36] *Tutorials STEM Project*, <http://wiki.eclipse.org/TutorialsforDevelopers/>
- [37] *Estadísticas de influenza en Jalisco*, <http://informacioninfluenza.jalisco.gob.mx/>
- [38] *Polígono de la ZMG, INEGI Guadalajara*
- [39] *Estudio de demanda multimodal de desplazamientos de la Zona Metropolitana de Guadalajara*, actualización 2007, SEDEUR.
- [40] *Estadísticas de población del estado de Jalisco*, <http://coepo.jalisco.gob.mx/>
- [41] *Segunda batería de preguntas-respuestas sobre la influenza AH1N1*, Respuestas de la ciencia y acciones UNAM, 7 octubre 2009.
- [42] J. X. Velasco, *Modelos matemáticos en epidemiología enfoques y alcances*, Miscelánea Matemática 44 (2007) 11–27, SMM.
- [43] J. X. Velasco, M. Conceicao, *A model for the A(H1N1) epidemic in Mexico, including social isolation*, Salud Pública de México vol. 53, no. 1, enero-febrero de 2011.
- [44] Curtis R. Vogel, *Computational methods for inverse problems*, SIAM 2002.

- [45] Lih-Ing Wu, Zhilan Feng, *Homoclinic Bifurcation in an SIQR Model for Childhood Diseases*, *Journal of Differential Equations* 168, 150–167, 2000.



Universidad Autónoma Metropolitana - Iztapalapa
División de Ciencias Básicas e Ingeniería

Modelación de Enfermedades Infecciosas con Información Geográfica

Tesis que presenta
Luis Alberto Zarate Siordia
Para obtener el grado de
Maestro en Ciencias
(Matemáticas Aplicadas e Industriales)

Asesores: Dr. Joaquín Delgado Fernández
Dr. Jorge X. Velasco Hernández

México D.F. Marzo 2012