

UACM

Universidad Autónoma
de la Ciudad de México

Nada humano me es ajeno

COLEGIO DE CIENCIA Y TECNOLOGÍA

APLICACIONES DE TEORÍA DE CONTROL A UN MODELO EPIDEMIOLÓGICO

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

LICENCIADO EN MODELACIÓN MATEMÁTICA

P R E S E N T A :

DAVID FERNANDO VIZUET MORALES

DIRECTOR

MTRO. MARCO ANTONIO PRADO ZÁYAGO

CODIRECTOR

DR. PABLO PADILLA LONGORIA

CIUDAD DE MÉXICO, 2021

Índice general

1. Introducción	1
2. Preliminares	3
2.1. Sistemas Dinámicos	3
2.1.1. Sistemas lineales	3
2.1.2. Matriz exponencial y algunas propiedades	5
2.1.3. Plano fase	6
2.1.4. Linealización y estabilidad para sistemas no lineales	9
2.2. Cálculo Variacional	15
2.2.1. Formulación del problema de cálculo variacional	15
2.2.2. Condiciones necesarias de optimalidad	18
2.2.3. Formulación del problema para n funciones	22
2.2.4. Problemas variacionales con restricciones	23
2.2.5. Multiplicadores de Lagrange	26
2.3. Control Óptimo	29
2.3.1. Controlabilidad	29
2.3.2. Planteamiento del problema básico.	32
2.3.3. Diferentes formas del problema básico y un caso especial	33
2.3.4. El principio del máximo de Pontryagin	34
2.3.5. Caso particular del principio del máximo de Pontryagin	38
3. Desarrollo	41
3.1. Modelo SEIR	41
3.2. Modelo SEIR con control	46
3.2.1. Estabilidad y controlabilidad	47
3.2.2. Aplicación del principio del máximo de Pontryagin	54
4. Resultados numéricos	59
5. Conclusiones	63
Apéndice A. Algunos resultados de Sistemas dinámicos	65
Apéndice B. Algunos resultados de Cálculo Variacional	69
Apéndice C. Algunos resultados de Control Óptimo	75

Apéndice D. Implementación Numérica	83
Bibliografía	87

1 Introducción

Los modelos matemáticos son una de las herramientas utilizadas hoy en día para el estudio de problemas en epidemiología que permiten describir, explicar y predecir fenómenos y procesos en dicha área. Son muy importantes para el estudio de enfermedades, en la mayoría de los problemas de enfermedades infecciosas no es factible la experimentación por un alto costo o por un gran riesgo. Por lo tanto, es pertinente proponer un modelo que describa de manera adecuada las características básicas de la enfermedad y posteriormente usarlo para analizar cambios específicos. Además, los modelos permiten entender la dinámica de una enfermedad infecciosa bajo diferentes escenarios.

Daniel Bernoulli en 1760 desarrolló un modelo matemático explícito para una enfermedad infecciosa (en [1] se muestra el modelo con una interpretación moderna). Más tarde, en 1927, Kermack y McKendrick formularon un modelo matemático general y complejo para describir la epidemia de peste en la ciudad de Bombay, India en 1906 [2]. Ellos plantearon un modelo que divide a la población en clases según sea el estado de los individuos frente a la infección, a saber: **Susceptibles, Infectados y Recuperados (SIR)**. Este modelo básico se ha modificado para incorporar una población de expuestos o factores que intervienen en diferentes procesos infecciosos. Estos modelos evolucionan en el tiempo según un sistema de ecuaciones diferenciales no lineales acopladas. Estos sistemas no tienen solución analítica; para conocer la evolución de la epidemia a largo plazo, es necesario utilizar métodos de aproximación numérica o hacer un análisis cualitativo.

El proceso de formulación de este modelo lleva a proponer hipótesis, crear variables y añadir parámetros que tienen una interpretación física; además de un posterior análisis conceptual del número básico de reproducción \mathcal{R}_0 que determina el comportamiento de la enfermedad.

En el modelo **SIR**, cuando los individuos infectados dejan de estarlo, se vuelven inmunes a la enfermedad y pasan al estado recuperado (como la rubéola, el sarampión, etc.). En algunas enfermedades (como el VIH, la varicela, entre otras) existe un período de tiempo (período de latencia) durante el cual el individuo infectado no tiene la capacidad de transmitir la enfermedad; a lo largo de dicho período se dice que el individuo está expuesto, y la inclusión de la población expuesta describe mejor el modelo. Las siguientes hipótesis son comunes a estos modelos; (1) la enfermedad se transmite por contacto directo entre un individuo infectado y uno susceptible; (2) la enfermedad se transmite inmediatamente después de producirse el contacto directo; (3) todos los individuos susceptibles tienen la misma probabilidad de ser infectados.

El propósito de estos modelos es el de proporcionar herramientas al correspondiente gestor para que pueda adoptar las políticas sanitarias (establecer mecanismos eficientes y protocolos sanitarios de seguimiento y control) necesarias en caso de que se produzca una epidemia.

En este trabajo se estudia el progreso de una enfermedad infecciosa en una cierta población, planteado como un problema de la teoría de control. Se divide a los individuos en cuatro compartimentos diferentes relevantes a la enfermedad. Estas son **S Susceptibles**, **E Expuestos**, **I Infectados** y **R Recuperados (SEIR)**. Se propone controlar la expansión de la enfermedad con el uso de una vacuna, es decir, se estudia la política de vacunación óptima en el modelo epidemiológico **SEIR**.

En el segundo capítulo se presentan de manera preliminar algunos conceptos y resultados básicos de la teoría de los sistemas dinámicos. Conceptos fundamentales del cálculo variacional que ayudarán a introducir a la teoría de control óptimo. Posteriormente se enuncia el principio del máximo de Pontryagin y se demuestra en un caso particular.

En el tercer capítulo se estudia el modelo **SEIR**, el planteamiento de un problema de control para dicho modelo, la estabilidad, la propiedad de controlabilidad y la aplicación del principio del máximo de Pontryagin.

En el cuarto capítulo se muestran algunos resultados numéricos de lo estudiado con antelación, cuya implementación está hecha en MATLAB.

2 Preliminares

2.1. Sistemas Dinámicos

En su acepción corriente un sistema dinámico consiste en un conjunto de magnitudes medibles que evolucionan con el tiempo. Los sistemas se pueden caracterizar por ecuaciones diferenciales, las cuales describen la dinámica de ciertos subprocesos que ocurren en la realidad, por ejemplo la velocidad de crecimiento de cierta población en un intervalo de tiempo. En esta sección se estudia de manera general los sistemas de ecuaciones diferenciales, los conceptos básicos y procedimientos usuales para resolverlas o dar una descripción cualitativa del comportamiento de las soluciones. Para un tratamiento más detallado de lo descrito a continuación consultar [3–5], de donde se tomaron algunos párrafos textualmente.

2.1.1. Sistemas lineales

Un sistema de ecuaciones diferenciales de primer orden en n funciones desconocidas $x_1(t), x_2(t), \dots, x_n(t)$ puede ser escrito en términos de las derivadas de n variables dependientes como sigue

$$\begin{aligned} \dot{x}_1(t) &= f_1(t, x_1, x_2, \dots, x_n), \\ \dot{x}_2(t) &= f_2(t, x_1, x_2, \dots, x_n), \\ &\vdots \\ \dot{x}_n(t) &= f_n(t, x_1, x_2, \dots, x_n). \end{aligned} \tag{2.1.1}$$

Definición 2.1.1. Si cada una de las funciones $f_i(t, x_1, x_2, \dots, x_n)$ con $i = 1, 2, \dots, n$ en (2.1.1) es lineal en las variables dependientes x_n , es decir,

$$f_i(t, x_1, x_2, \dots, x_n) = a_{i1}(t)x_1(t) + a_{i2}(t)x_2(t) + \dots + a_{in}(t)x_n(t) + b_i(t),$$

donde a_{ij} y b_i con $i, j = 1, \dots, n$, son funciones arbitrarias de t , entonces el sistema de ecuaciones (2.1.1) se denomina un sistema lineal de la forma

$$\begin{aligned}
\dot{x}_1(t) &= a_{11}(t)x_1(t) + a_{12}(t)x_2(t) + \cdots + a_{1n}(t)x_n(t) + b_1(t), \\
\dot{x}_2(t) &= a_{21}(t)x_1(t) + a_{22}(t)x_2(t) + \cdots + a_{2n}(t)x_n(t) + b_2(t), \\
&\vdots \\
\dot{x}_n(t) &= a_{n1}(t)x_1(t) + a_{n2}(t)x_2(t) + \cdots + a_{nn}(t)x_n(t) + b_n(t).
\end{aligned} \tag{2.1.2}$$

Si las a_{ij} son constantes, se dice que el sistema (2.1.2) tiene coeficientes constantes; en caso contrario, tiene coeficientes variables. Si cada $b_i \equiv 0$ se dice que el sistema (2.1.2) es homogéneo; en caso contrario, es no homogéneo.

Cuando se trabaja con sistemas lineales es conveniente emplear la notación matricial. Si se usa la notación del sistema (2.1.2) se puede escribir en la forma:

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{b}(t), \tag{2.1.3}$$

donde

$$\dot{\mathbf{x}}(t) = \begin{pmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \vdots \\ \dot{x}_n(t) \end{pmatrix}, \quad \mathbf{x}(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{pmatrix}, \quad \mathbf{b}(t) = \begin{pmatrix} b_1(t) \\ b_2(t) \\ \vdots \\ b_n(t) \end{pmatrix} \text{ y } \mathbf{A}(t) = \begin{pmatrix} a_{11}(t) & a_{12}(t) & \cdots & a_{1n}(t) \\ a_{21}(t) & a_{22}(t) & \cdots & a_{2n}(t) \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}(t) & a_{n2}(t) & \cdots & a_{nn}(t) \end{pmatrix}.$$

Los vectores $\dot{\mathbf{x}}(t)$ y $\mathbf{x}(t)$ tienen como entradas funciones reales de una variable y la matriz cuadrada $\mathbf{A}(t)$ se llama *matriz de coeficientes*, además se asume que las $a_{ij}(t)$ y las $b_i(t)$ son funciones continuas en cierto intervalo. La ventaja de la notación matricial es que ayuda a ver similitudes entre los sistemas de primer orden y las ecuaciones lineales de primer orden. Trabajar con matrices proporciona también algunas herramientas algebraicas útiles que serán presentadas más adelante.

Una **solución del sistema** (2.1.3) en un intervalo $I = [a, b]$ es una función vectorial $\mathbf{x}(t)$ definida en I y con valores en \mathbb{R}^n que cumple

$$\dot{x}_i(t) = f_i(t, x_1, x_2, \dots, x_n),$$

para cada $t \in I$ e $i = 1, 2, \dots, n$.

Para garantizar la existencia y unicidad de las soluciones del sistema lineal, se necesita el siguiente resultado.

Teorema 2.1.1 (Existencia y unicidad). *Sea E un subconjunto abierto de \mathbb{R}^n que contiene a \mathbf{x}_0 y suponer que $\mathbf{f} \in C^1(E)$. Entonces existe $a > 0$ tal que el problema con valor inicial*

$$\begin{aligned}
\dot{\mathbf{x}} &= \mathbf{f}(\mathbf{x}) \\
\mathbf{x}(0) &= \mathbf{x}_0
\end{aligned}$$

tiene solución única $\mathbf{x}(t)$ en el intervalo $[-a, a]$.

Demostración. Se puede consultar en [4]. □

2.1.2. Matriz exponencial y algunas propiedades

Considerar la ecuación diferencial de primer orden lineal homogénea con coeficientes constantes $\dot{x}(t) = \alpha x(t)$ con $x : I \subseteq \mathbb{R} \rightarrow \mathbb{R}$ y $\alpha \in \mathbb{R}$, la solución general es la función $x(t) = ke^{\alpha t}$ con k una constante arbitraria. Cualquier solución de esta ecuación se obtiene como combinación lineal de la solución $e^{\alpha t}$. Como $e^{\alpha t} \neq 0$, $\forall t \in \mathbb{R}$ entonces $\{e^{\alpha t}\}$ es el conjunto fundamental de soluciones y la matriz fundamental es $\Phi(t) = [e^{\alpha t}]$ de tamaño 1×1 .

Un método para encontrar las soluciones de un sistema de ecuaciones diferenciales lineales con coeficientes constantes, es utilizar la *matriz exponencial*.

Definición 2.1.2 (Matriz exponencial). *La matriz exponencial de una matriz \mathbf{A} de tamaño $n \times n$ con coeficientes constantes está definida por*

$$e^{\mathbf{A}t} \equiv \mathbf{I} + \mathbf{A}t + \mathbf{A}^2 \frac{t^2}{2!} + \mathbf{A}^3 \frac{t^3}{3!} + \cdots = \sum_{k=0}^{\infty} \frac{\mathbf{A}^k t^k}{k!}. \quad (2.1.4)$$

La matriz exponencial cuenta con algunas propiedades, como las siguientes:

- $e^{\mathbf{0}} = \mathbf{I}$, donde $\mathbf{0}$ es la matriz cero de tamaño $n \times n$ e \mathbf{I} es la matriz identidad de tamaño $n \times n$. Esto es fácil de ver usando la definición 2.1.4.
- Si \mathbf{A} y \mathbf{B} conmutan, es decir, $\mathbf{AB} = \mathbf{BA}$, entonces $e^{\mathbf{A}+\mathbf{B}} = e^{\mathbf{A}}e^{\mathbf{B}}$ (se demuestra en A.0.1).
- La matriz inversa de $e^{\mathbf{A}t}$ es $e^{-\mathbf{A}t}$.

Lema 2.1.1. *Sea \mathbf{A} una matriz cuadrada, entonces*

$$\frac{d}{dt}e^{\mathbf{A}t} = \mathbf{A}e^{\mathbf{A}t}.$$

Demostración. Se puede consultar en A.1. □

Teorema 2.1.2 (Teorema fundamental para sistemas lineales). Sea \mathbf{A} una matriz de $n \times n$ con coeficientes constantes y $\mathbf{X}(t) = e^{\mathbf{A}t}$.

1. La solución única del sistema homogéneo

$$\begin{aligned}\dot{\mathbf{x}}(t) &= \mathbf{A}\mathbf{x}(t), \\ \mathbf{x}(0) &= \mathbf{x}_0,\end{aligned}\tag{2.1.5}$$

es

$$\mathbf{x}(t) = \mathbf{X}(t)\mathbf{x}_0 = e^{\mathbf{A}t}\mathbf{x}_0.\tag{2.1.6}$$

2. La solución única del sistema no homogéneo

$$\begin{aligned}\dot{\mathbf{x}}(t) &= \mathbf{A}\mathbf{x}(t) + \mathbf{b}(t), \\ \mathbf{x}(0) &= \mathbf{x}_0,\end{aligned}\tag{2.1.7}$$

es

$$\mathbf{x}(t) = \mathbf{X}(t)\mathbf{X}^{-1}(0)\mathbf{x}_0 + \int_0^t \mathbf{X}(t)\mathbf{X}^{-1}(s)\mathbf{b}(s)ds,\tag{2.1.8}$$

donde $\mathbf{X}^{-1}(0) = \mathbf{I}$ es la matriz identidad de tamaño $n \times n$.

Demostración. Se puede consultar en A.2. □

2.1.3. Plano fase

Una técnica muy usada para entender el comportamiento de un sistema dinámico es el *diagrama fase* (consultar [3] para mayor detalle), que ofrece información cualitativa acerca de los sistemas. En palabras más formales, un *diagrama fase* es una colección de trayectorias que representan las soluciones de ecuaciones en el espacio fase y proporciona información sobre los comportamientos transitorios de las soluciones.

Para analizar un sistema dinámico, es importante determinar los puntos de equilibrio.

Definición 2.1.3 (Punto de equilibrio). Un punto $\tilde{\mathbf{x}} \in \mathbb{R}^n$ es llamado punto de equilibrio de $\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t))$ si $\mathbf{f}(\tilde{\mathbf{x}}) = \mathbf{0}$.

Así pues, en cada punto del *diagrama fase*, la velocidad de fase indica la dirección de la tangente a la curva de evolución que pasa por este punto. En los puntos donde la velocidad de fase es cero, no hay curva que pase a través de ellos. En esos casos, el estado es constante.

Existen muchas ecuaciones diferenciales cuya solución no es posible obtener de manera analítica. Una forma de abordar estas ecuaciones es el uso de métodos numéricos. Otro enfoque es utilizar la geometría del campo vectorial para encontrar soluciones especiales de los sistemas lineales. Considerar el sistema más simple, es decir, un sistema lineal homogéneo con coeficientes constantes de la forma

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t),$$

donde \mathbf{A} es una matriz de tamaño 2×2 y $\mathbf{x}(t)$ un vector de 2×1 .

Se buscan soluciones de la forma $\mathbf{x}(t) = \mathbf{v}e^{\lambda t}$, es decir, buscar soluciones usando los valores y vectores propios asociados a la matriz \mathbf{A} . A partir del conocimiento de los valores propios se hace una clasificación sobre los distintos tipos de comportamiento que tienen las soluciones de los sistemas (para un tratamiento más profundo se puede consultar Differential Equations and Dynamical Systems [4]).

Ejemplo 2.1.1 (Modelo depredador-presa).

Supóngase que $x(t)$ representa el número de depredadores en el tiempo t , y $y(t)$ es el número de presas que compiten por el mismo recurso. Cada especie evoluciona por sí misma, de acuerdo con un modelo de crecimiento logístico de población. El siguiente sistemas de ecuaciones diferenciales modela este fenómeno

$$\begin{aligned}\dot{x}(t) &= 2x \left(1 - \frac{x}{2}\right) - xy, \\ \dot{y}(t) &= 3y \left(1 - \frac{y}{3}\right) - 2xy.\end{aligned}$$

La interacción de las dos especies es modelada por los términos xy . Por ejemplo, el efecto de la población $y(t)$, sobre la razón de cambio de $x(t)$, está determinado por el término $-xy$ en la ecuación $\dot{x}(t)$. Este término es negativo ya que se asume que las dos especies compiten por los mismos recursos. De manera similar, $-2xy$ define el efecto de la población $x(t)$ sobre la razón de cambio de $y(t)$. Como $x(t)$ y $y(t)$ representan poblaciones, es intuitivamente claro que las soluciones y las condiciones iniciales se encuentren en el primer cuadrante del eje coordenado.

Primero se encuentran los puntos de equilibrio igualando a cero el lado derecho de las ecuaciones diferenciales y despejando $x(t)$ y $y(t)$ en los sistemas de ecuaciones

$$\begin{aligned}2x \left(1 - \frac{x}{2}\right) - xy &= 0, \\ 3y \left(1 - \frac{y}{3}\right) - 2xy &= 0,\end{aligned}$$

o, de manera equivalente

$$\begin{aligned}x(2 - x - y) &= 0, \\ y(3 - y - 2x) &= 0.\end{aligned}$$

La primera ecuación se cumple si $x = 0$ ó si $2 - x - y = 0$, y la segunda ecuación de satisface si $y = 0$ ó $3 - y - 2x = 0$. Si $x = 0$, entonces la ecuación $y = 0$ proporciona un punto de equilibrio en el origen, y la ecuación $3 - y - 2x = 0$ lo da en el punto $(0, 3)$. Si se resuelve de manera simultánea el sistema de ecuaciones se obtiene los puntos de equilibrio $(0, 0)$, $(0, 3)$, $(2, 0)$, $(1, 1)$.

Con ayuda del campo de direcciones es posible comprender el comportamiento general de las curvas solución, si se toman suficientes soluciones es posible bosquejar el plano fase como se muestra en la figura (2.1).

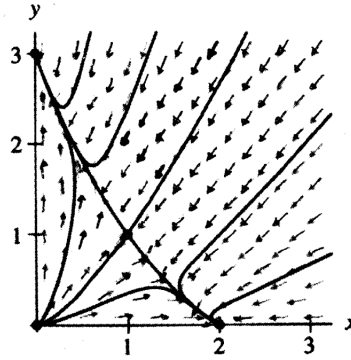


Figura 2.1: Plano fase.

Aunque el plano fase representa una manera de visualizar las soluciones de los sistemas, no se tiene información de lo que ocurre con la variable independiente, en nuestro caso el tiempo t . Las gráficas de $x(t)$ y $y(t)$ se muestran en las siguientes figuras (obtenidas en [3] p.167):

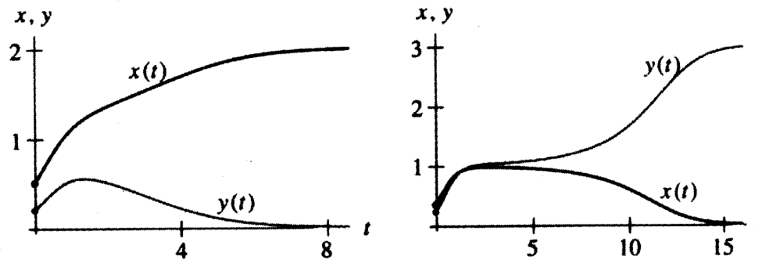


Figura 2.2: Las gráficas $x(t)$ y $y(t)$ para dos soluciones con condiciones iniciales cercanas .

2.1.4. Linealización y estabilidad para sistemas no lineales

Las herramientas de la teoría cualitativa de ecuaciones diferenciales son habitualmente utilizadas para analizar la evolución de un sistema dinámico debido a que, sin conocer la solución analítica, es posible describir el comportamiento local del sistema a partir del conocimiento de los puntos críticos. En ocasiones, cuando aparece la no linealidad en los sistemas, una aproximación lineal permite analizar a una función dada alrededor de un punto (linealización). Además es de interés que un punto de equilibrio sea estable, es decir, si se empieza cerca de un punto de equilibrio, la solución permanezca cerca de dicho punto.

Para los sistemas dinámicos se emplea un criterio de clasificación el cual establece que para las ecuaciones de evolución la variable temporal t puede aparecer o no en dichas ecuaciones. Un sistema es *autónomo* cuando sus ecuaciones de movimiento contengan implícitamente el tiempo y *no autónomo* en caso contrario.

Para los conceptos e ideas planteadas a continuación se consideran sistemas de ecuaciones diferenciales en \mathbb{R}^2 .

La idea del proceso de linealización es aproximar el sistema no lineal por medio de un sistema lineal apropiado para condiciones iniciales cerca del punto de equilibrio.

Considerar el siguiente sistema no lineal autónomo dado por:

$$\begin{aligned}\dot{x}(t) &= f(x, y), \\ \dot{y}(t) &= g(x, y),\end{aligned}\tag{2.1.9}$$

con condiciones iniciales $x(t_0) = x_0$, $y(t_0) = y_0$.

Suponer que (x_0, y_0) es un punto de equilibrio para el sistema. Para estudiar el comportamiento con las soluciones cerca de (x_0, y_0) , se linealiza el sistema alrededor de (x_0, y_0) . Considerar las siguientes variables nuevas

$$\begin{aligned}u &= x - x_0, \\ v &= y - y_0,\end{aligned}$$

que llevan el punto de equilibrio al origen. Si x y y están cerca del punto de equilibrio (x_0, y_0) , entonces u y v tienden a 0.

Como $x = u + x_0$ y $y = v + y_0$ y los valores x_0, y_0 son constantes, el sistema (2.1.9) escrito en términos de u y v es

$$\begin{aligned}\dot{u} = \dot{x}(t) &= f(x, y) = f(x_0 + u, y_0 + v), \\ \dot{v} = \dot{y}(t) &= g(x, y) = g(x_0 + u, y_0 + v).\end{aligned}\tag{2.1.10}$$

Si $u = 0 = v$, el lado derecho de este sistema desaparece, y se ha movido el punto de equilibrio al origen en el plano uv .

El cálculo diferencial permite estudiar una función analizando su “mejor aproximación lineal” dada por el plano tangente en el caso de funciones de dos variables. Aplicando la serie de Taylor para $f(x_0 + u, y_0 + v)$ y $g(x_0 + u, y_0 + v)$ alrededor del punto de equilibrio (x_0, y_0) se tiene

$$\begin{aligned} f(x_0 + u, y_0 + v) &\approx f(x_0, y_0) + \left[\frac{\partial}{\partial x} f(x_0, y_0) \right] u + \left[\frac{\partial}{\partial y} f(x_0, y_0) \right] v + O(u^2, v^2, uv), \\ g(x_0 + u, y_0 + v) &\approx g(x_0, y_0) + \left[\frac{\partial}{\partial x} g(x_0, y_0) \right] u + \left[\frac{\partial}{\partial y} g(x_0, y_0) \right] v + O(u^2, v^2, uv). \end{aligned}$$

Si los términos cuadráticos, incluidos en el término $O(u^2, v^2, uv)$, son despreciables con respecto a los términos lineales y, debido a que $f(x_0, y_0) = 0 = g(x_0, y_0)$ son puntos de equilibrio se obtiene el sistema lineal

$$\begin{aligned} \dot{u} &= \left[\frac{\partial}{\partial x} f(x_0, y_0) \right] u + \left[\frac{\partial}{\partial y} f(x_0, y_0) \right] v, \\ \dot{v} &= \left[\frac{\partial}{\partial x} g(x_0, y_0) \right] u + \left[\frac{\partial}{\partial y} g(x_0, y_0) \right] v. \end{aligned}$$

Por lo tanto, el sistema anterior se puede reescribir de manera vectorial como

$$\begin{pmatrix} \dot{u} \\ \dot{v} \end{pmatrix} = \begin{pmatrix} \frac{\partial}{\partial x} f(x_0, y_0) & \frac{\partial}{\partial y} f(x_0, y_0) \\ \frac{\partial}{\partial x} g(x_0, y_0) & \frac{\partial}{\partial y} g(x_0, y_0) \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \mathbf{J}|_{(x_0, y_0)} \begin{pmatrix} u \\ v \end{pmatrix}, \quad (2.1.11)$$

donde

$$\mathbf{J}|_{(x_0, y_0)} = \begin{pmatrix} \frac{\partial}{\partial x} f(x_0, y_0) & \frac{\partial}{\partial y} f(x_0, y_0) \\ \frac{\partial}{\partial x} g(x_0, y_0) & \frac{\partial}{\partial y} g(x_0, y_0) \end{pmatrix} \quad (2.1.12)$$

es la *matriz jacobiana* del sistema en el punto de equilibrio (x_0, y_0) .

En [3] se hace la siguiente clasificación de los puntos de equilibrio.

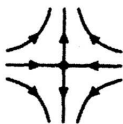




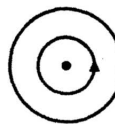
Si todos los valores propios de la *matriz jacobiana* son números reales negativos o números complejos con parte real negativa, entonces $(u, v) = (0, 0)$ es un **sumidero** para el sistema lineal y todas las soluciones se acercan a $(u, v) = (0, 0)$ cuando $t \rightarrow \infty$. Para el sistema no lineal, las soluciones que empiezan cerca del punto de equilibrio $(x, y) = (x_0, y_0)$ se acercan a éste cuando $t \rightarrow \infty$. Por lo tanto, se dice que (x_0, y_0) es un **sumidero**. Si los valores propios son complejos, entonces (x_0, y_0) es un **sumidero espiral**.

De manera similar, si la *matriz jacobiana* sólo tiene valores propios positivos o complejos con parte real positiva, entonces las soluciones con condiciones iniciales cerca del punto de equilibrio (x_0, y_0) tiende a alejarse de éste cuando t crece. En este caso, para un sistema no lineal el punto (x_0, y_0) es llamado **fuelle**. Si los valores propios son complejos, entonces (x_0, y_0) es una **fuelle espiral**.

Por último, si la *matriz jacobiana* tiene un valor propio positivo y uno negativo, entonces el punto de equilibrio (x_0, y_0) es un **punto silla**. De manera análoga, cuando se tiene un

sistema lineal con un punto silla en el origen, hay dos curvas de soluciones que se acercan al punto de equilibrio cuando t se incrementa y otras dos que también se aproximan a dicho punto conforme t decrece ([3] p.414).

La siguiente tabla (obtenida en [3] p.313) esboza la dinámica del sistema linealizado en el plano fase para los diferentes puntos mencionados a partir del conocimiento de los valores propios de la *matriz jacobiana*.

Tipo	Eigenvalores	Plano fase	Tipo	Eigenvalores	Plano fase
Punto silla	$\lambda_1 < 0 < \lambda_2$		Sumidero espiral	$\lambda = a \pm ib$ $a < 0, b \neq 0$	
Sumidero	$\lambda_1 < \lambda_2 < 0$		Fuente espiral	$\lambda = a \pm ib$ $a > 0, b \neq 0$	
Fuente	$0 < \lambda_1 < \lambda_2$		Centro	$\lambda = \pm ib$ $b \neq 0$	

El teorema de Hartman-Grobman, garantiza que, en condiciones apropiadas, en una vecindad del punto de equilibrio las propiedades cualitativas del sistema no lineal se conservan por la linealización. Las condiciones se refieren a la hiperbolicidad del punto fijo.

Definición 2.1.4. *Un punto fijo de un sistema de ecuaciones diferenciales es hiperbólico si la matriz jacobiana calculada en ese punto no tiene valores propios iguales a cero o son imaginarios puros (ningún valor tiene una parte real igual a cero).*

Formalmente el resultado de Hartman-Grobman es el siguiente:

Teorema 2.1.3. *Si $\bar{\mathbf{x}}$ es un punto fijo hiperbólico de $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^m$ entonces existe un homeomorfismo h definido en alguna vecindad N de $\bar{\mathbf{x}}$ en \mathbb{R}^m que manda órbitas locales del sistema no lineal hacia el sistema lineal. La función h preserva el sentido de las órbitas que también puede ser elegido para preservar la parametrización por el tiempo.*

La “estabilidad de Lyapunov” está relacionada con el comportamiento de las trayectorias de un sistema, cuando su estado inicial se encuentra cerca de un punto de equilibrio. Desde la perspectiva práctica este punto es importante, debido a que las perturbaciones que afectan a un sistema tienden a separarlo del equilibrio.

Los puntos de equilibrio son estables si todas las soluciones que se encuentran en sus cercanías permanecen próximas del punto, mientras que se consideran asintóticamente estables si las soluciones además de permanecer cerca, tienden hacia éste cuando $t \rightarrow \infty$. En caso contrario, el punto de equilibrio será inestable.

Definición 2.1.5 (Punto de equilibrio estable). *Sea ϕ_t la trayectoria del sistema de ecuaciones 2.1.9 definida para cada $t \in \mathbb{R}$. Un punto de equilibrio \mathbf{x}_0 de 2.1.9 es **estable** si para cada $\epsilon > 0$ existe un $\delta > 0$ tal que para cada $\mathbf{x} \in N_\delta(\mathbf{x}_0)$ y $t \geq 0$ se tiene*

$$\phi_t \in N_\epsilon(\mathbf{x}_0).$$

El punto de equilibrio \mathbf{x}_0 es **inestable** si no es estable. \mathbf{x}_0 es **asintóticamente estable** si existe un $\delta > 0$ tal que para cada $\mathbf{x} \in N_\delta(\mathbf{x}_0)$ se tiene

$$\lim_{t \rightarrow \infty} \phi_t(\mathbf{x}) = \mathbf{x}_0,$$

donde la trayectoria ϕ_t representa la solución del sistema 2.1.9 dada a partir de la condición inicial $\mathbf{x}(t_0) = \mathbf{x}_0$ en el instante de tiempo $t = t_0$. Para más detalle se puede consultar [4, 6].

El siguiente criterio permite establecer la estabilidad del sistema a partir del signo de los valores propios de la *matriz jacobiana*.

Teorema 2.1.4 (Lyapunov (1892)). *Considerar un sistema dinámico definido por*

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

donde \mathbf{f} es suave. Si se supone que el sistema tiene un punto de equilibrio \mathbf{x}_0 y se considera la matriz jacobiana \mathbf{J} de $\mathbf{f}(\mathbf{x})$ evaluada en el punto \mathbf{x}_0 . Entonces \mathbf{x}_0 es asintóticamente estable si y sólo si todos los valores $\lambda_1, \lambda_2, \dots, \lambda_n$ de \mathbf{J} tienen parte real negativa ($\text{Re } \lambda < 0$).

Las demostraciones de los teoremas (2.1.3) y (2.1.4) se omiten en este trabajo, los detalles técnicos se pueden consultar en [4–7].

Ejemplo 2.1.2. *A continuación se muestra el proceso de linealización de un sistema general para el estudio de su estabilidad. Sea el siguiente sistema no lineal y autónomo que modela el movimiento de un péndulo amortiguado.*

$$\begin{aligned}\dot{x}(t) &= y, \\ \dot{y}(t) &= -y - \sin x.\end{aligned}\tag{2.1.13}$$

Para encontrar los puntos de equilibrio se tiene que:

$$\begin{aligned}\dot{x}(t) &= y = 0, \\ \dot{y}(t) &= -y - \sin x = 0.\end{aligned}$$

Si $y = 0$, entonces se buscan los valores de x tales que $\sin x = 0$. La función $\sin x$ tiene un período de 2π y se anula en $\pm\pi, \pm2\pi, \pm3\pi, \dots$, entonces $\sin x = 0 \Leftrightarrow x = n\pi \ \forall n \in \mathbb{Z}$. Por lo tanto, los puntos de equilibrio de (2.1.13) son

$$(x, y) = (0, 0), (\pm\pi, 0), (\pm2\pi, 0), (\pm3\pi, 0), \text{ etc.}$$

Notar que se tiene una infinidad de puntos de equilibrio. Suponiendo que se quiere estudiar las soluciones cercanas al punto $(x, y) = (0, 0)$ se utiliza el método de linealización. La matriz jacobiana cerca del origen es

$$\mathbf{J}|_{(0,0)} = \left(\begin{array}{cc} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} \end{array} \right) \bigg|_{(0,0)} = \left(\begin{array}{cc} 0 & 1 \\ -\cos x & -1 \end{array} \right) \bigg|_{(0,0)} = \left(\begin{array}{cc} 0 & 1 \\ -1 & -1 \end{array} \right).$$

En este caso, el sistema linealizado es

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \mathbf{J}|_{(0,0)} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix},$$

o,

$$\begin{aligned}\dot{x}(t) &= y, \\ \dot{y}(t) &= -y - x.\end{aligned}\tag{2.1.14}$$

Calculando los valores propios de la matriz jacobiana:

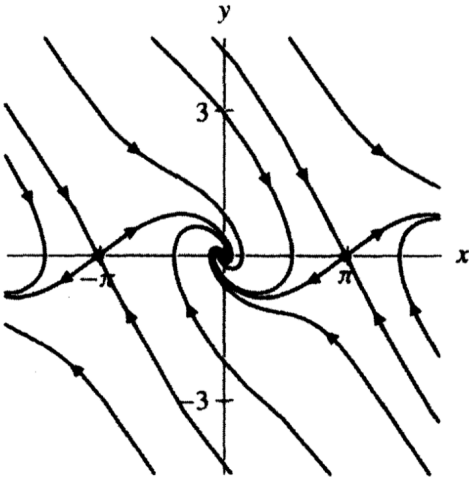
$$p(\lambda) = \det(\mathbf{J} - \lambda \mathbf{I}) = \begin{vmatrix} -\lambda & 1 \\ -1 & -1 - \lambda \end{vmatrix} = (-\lambda)(-1 - \lambda) - (1)(-1) = \lambda^2 + \lambda + 1 = 0.$$

Resolviendo $\lambda^2 + \lambda + 1 = 0$, se obtienen los siguientes valores propios

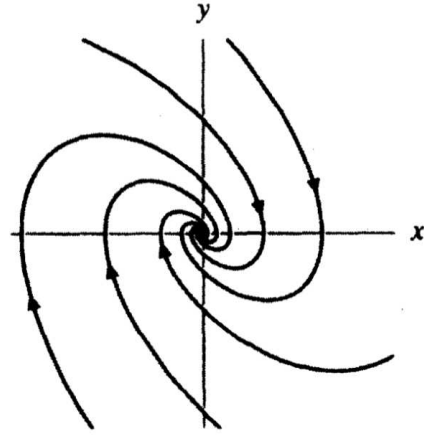
$$\lambda_1 = -\frac{1}{2} + \frac{\sqrt{3}i}{2} \text{ y } \lambda_2 = -\frac{1}{2} - \frac{\sqrt{3}i}{2}.$$

Como los valores propios λ_1 y λ_2 son números complejos con partes reales negativas el punto de equilibrio para el sistema no lineal es un sumidero espiral. Aplicando el teorema (2.1.4) se concluye que el punto $(x_0, y_0) = (0, 0)$ es asintóticamente estable.

En la figura 2.3(a) se muestra el plano fase del sistema no lineal (2.1.13), mientras que la figura 2.3(b) (obtenida en [3] p.409) muestra el plano fase del sistema linealizado (2.1.14), cerca del punto de equilibrio $(0, 0)$.



(a) Figura 1.



(b) Figura 2.

Figura 2.3: Plano fase del sistema no lineal y el sistema linealizado.

2.2. Cálculo Variacional

El propósito de esta sección es proporcionar los conceptos fundamentales del cálculo variacional, analizar su problema básico, plantear problemas con restricciones y posteriormente relacionarlo con el problema básico de la teoría de control.

El problema central del cálculo variacional es análogo a uno de los problemas más importantes del cálculo diferencial. Éste consiste en optimizar una función (encontrar máximos o mínimos), en caso de existir, a la que se le han impuesto ciertas restricciones en una región dada. Para un tratamiento más detallado de lo descrito a continuación consultar [8–10], de donde se tomaron algunos párrafos textualmente.

2.2.1. Formulación del problema de cálculo variacional

Para formular el problema de forma general, es necesario establecer algunos conceptos fundamentales. Sea t la variable independiente que representará el tiempo y x la variable dependiente.

Sean $t_0, t_1 \in \mathbb{R}$ tales que $t_0 < t_1$. Considerar el siguiente conjunto de funciones:

$$\Omega = \{x : [t_0, t_1] \rightarrow \mathbb{R} \mid x \in C^2([t_0, t_1])\}. \quad (2.2.15)$$

Definición 2.2.1 (Funcional). *Un funcional es una aplicación, cuyo dominio es un conjunto de funciones, y cuyo rango es un subconjunto de \mathbb{R} , es decir:*

$$\begin{aligned} J : \Omega &\rightarrow \mathbb{R}, \\ x &\rightarrow J(x). \end{aligned}$$

A continuación se define el problema de cálculo de variaciones para el caso escalar con extremos fijos.

Sea $F[t, x(t), \dot{x}(t)] \in C^2([t_0, t_1])$ una función tal que

$$F : [t_0, t_1] \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}.$$

Considerar el siguiente funcional:

$$J(x) = \int_{t_0}^{t_1} F[t, x(t), \dot{x}(t)] dt,$$

donde $\dot{x}(t)$ es la derivada de $x(t)$ con respecto a t .

El problema que se quiere resolver es el de encontrar una función $x^*(t)$ con primera y segunda derivada continua en $[t_0, t_1]$ tal que $x^*(t_0) = x_0$ y $x^*(t_1) = x_1$, con x_0 y x_1 dados de tal forma que el funcional $J(x)$ alcance el valor mínimo (o el valor máximo).

La función $F[t, x(t), \dot{x}(t)]$ se le conoce como Lagrangeano asociado al funcional $J(x)$.

El problema, por tanto, en el caso de minimización es:

$$\min_{x \in \Omega} J(x) = \int_{t_0}^{t_1} F[t, x(t), \dot{x}(t)] dt, \quad (2.2.16)$$

$$\text{con } x(t_0) = x_0, \quad x(t_1) = x_1,$$

donde Ω está definido en (2.2.15).

Definición 2.2.2. *Se dice que $x(t)$ es una **función admisible** para el Problema (2.2.16), si verifica que*

$$x \in \Omega, \quad x(t_0) = x_0, \quad y \quad x(t_1) = x_1.$$

Por lo tanto, para este problema, el **conjunto factible** (o el conjunto de funciones admisibles) es

$$\Psi = \{x \in \Omega \mid x(t_0) = x_0, \quad x(t_1) = x_1\}. \quad (2.2.17)$$

Considerar el mínimo (o el máximo) de la función objetivo, es decir del funcional objetivo, no supone ninguna pérdida de generalidad, debido a que

$$\min J(x), \text{ es equivalente a } \max [-J(x)],$$

el elemento x que minimiza $J(x)$ es el mismo x que maximiza $[-J(x)]$.

Ejemplo 2.2.1 (El problema de la braquistócrona).

Clásicamente, este problema se enuncia como:

“Se requiere la forma de la curva que une dos puntos fijos A y B en un plano vertical de modo que una partícula se deslice por una curva (bajo el efecto de la gravedad y sin fricción) y viaje de A a B en un tiempo mínimo”.

Este problema no tiene una solución trivial; la línea recta de A a B no es la solución (esto también es intuitivamente claro, ya que si la pendiente es alta al principio y la partícula toma una gran velocidad es razonable que el tiempo de viaje pueda reducirse). Para plantear el problema en términos matemáticos, se considera el esquema que se muestra en la Figura 2.4, de modo que A es el punto $(0, 0)$, y B corresponde a (x_0, y_0) .

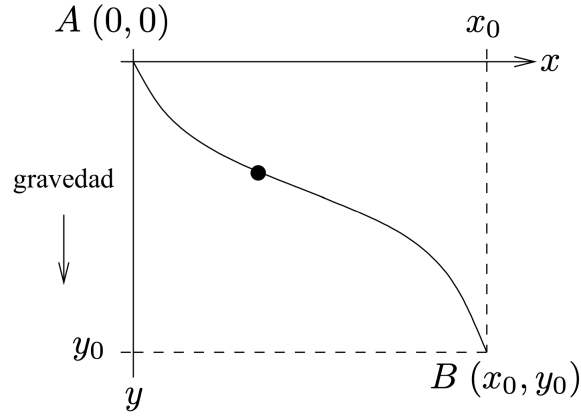


Figura 2.4: El problema de la braquistócrona

Supóngase que la partícula se libera desde el reposo en A y, por la conservación de la energía se tiene que

$$\frac{1}{2}mv^2 - mgy = 0, \quad (2.2.18)$$

donde v es la velocidad de la partícula con cierta masa m . Por lo tanto, la velocidad está dada por

$$v = \frac{ds}{dt} = \sqrt{2gy}, \quad (2.2.19)$$

donde s denota la longitud del arco a lo largo de la curva.

Por el Teorema de la función inversa se tiene que

$$\frac{1}{\dot{s}(t)} = \frac{1}{\frac{ds}{dt}} = \frac{dt}{ds},$$

además $ds = \sqrt{1 + (\dot{y})^2} dx$.

A partir de la relación entre la velocidad de la partícula y la longitud total (arco) recorrida se tiene que el tiempo total de descenso T está dado por

$$T = \int_0^{x_0} dt = \int_0^{x_0} \frac{ds}{\dot{s}(t)} = \int_0^{x_0} \frac{ds}{v} = \int_0^{x_0} \frac{1}{v} \sqrt{1 + (\dot{y})^2} dx = \int_0^{x_0} \frac{1}{\sqrt{2g}} \sqrt{\frac{1 + (\dot{y})^2}{y}} dx.$$

El problema es encontrar la trayectoria $y(x)$ tal que $y(0) = 0$ y $y(x_0) = y_0$ para la cual se minimiza el funcional $T(y(x))$. Más adelante se encuentra la solución a este problema.

2.2.2. Condiciones necesarias de optimalidad

El objetivo es plantear las condiciones necesarias sobre los funcionales para la existencia de un extremo (máximo o mínimo) similares a las del cálculo diferencial sobre la anulación de la derivada. Para resolver el Problema (2.2.16), es necesario buscar máximos y mínimos locales, por lo que se encuentran a los óptimos locales.

Definición 2.2.3 (mínimo global). *Sea x^* una función admisible para el Problema (2.2.16). Se dice que x^* es mínimo global si para cualquier función admisible x , se cumple que*

$$J(x^*) \leq J(x).$$

Definición 2.2.4 (mínimo local). *Sea x^* una función admisible para el Problema (2.2.16). Se dice que x^* es mínimo local si existe un $\epsilon > 0$, tal que para cualquier función admisible x , perteneciente a $B(x^*, \epsilon)$ se cumple que*

$$J(x^*) \leq J(x).$$

De manera similar se definen máximo global y máximo local, considerando la relación $J(x^*) \geq J(x)$.

Lema 2.2.1 (Lagrange). *Sea $g(x)$ una función continua en $[a, b]$ tal que*

$$\int_a^b g(x)h(x)dx = 0,$$

para cada $h \in C^1([a, b])$ con $h(a) = h(b) = 0$, entonces $g(x) \equiv 0$.

Demostración. Se puede consultar en B.1. □

Optimizar $J(x)$ significa buscar $x^*(t)$ tal que $J(x^*) \leq J(x)$ con $x(t_0) = x_0$ y $x(t_1) = x_1$ (en caso del mínimo). El siguiente resultado es el teorema fundamental que proporciona una condición necesaria de optimalidad local.

Teorema 2.2.1 (Ecuación de Euler). *Si $x^*(t)$ es un mínimo local de $J(x) = \int_{t_0}^{t_1} F[t, x(t), \dot{x}(t)]dt$, entonces $x^*(t)$ cumple la siguiente condición*

$$F_x - \frac{d}{dt}F_{\dot{x}} = 0, \quad \forall t \in [t_0, t_1]. \quad (2.2.20)$$

Demostración. Se puede consultar en B.2 □

En la demostración del teorema (2.2.1) se utiliza la idea de la primera derivada para buscar puntos críticos. Siguiendo este razonamiento se considera la segunda derivada del Lagrangeano F , conocida como la condición de Legendre que permite distinguir entre candidatos a mínimos y candidatos a máximos.

Teorema 2.2.2 (Condición de Legendre).

- Si $x^*(t)$ es un mínimo local del Problema (2.2.16), entonces $x^*(t)$ satisface la condición

$$F_{\dot{x}\dot{x}}[t, x^*(t), \dot{x}^*(t)] \geq 0, \quad \forall t \in [t_0, t_1].$$

- Si $x^*(t)$ es un máximo local del Problema (2.2.16), entonces $x^*(t)$ satisface la condición

$$F_{\dot{x}\dot{x}}[t, x^*(t), \dot{x}^*(t)] \leq 0, \quad \forall t \in [t_0, t_1],$$

donde $\frac{\partial^2 F}{\partial \dot{x}^2} = F_{\dot{x}\dot{x}}$.

Demostración. Se puede consultar en B.6. □

Dado el funcional

$$J(x) = \int_{t_0}^{t_1} F(t, x(t), \dot{x}(t)) dt,$$

suponer que

$$\begin{aligned} x(t) &= x^*(t) + \epsilon h(t), \\ \dot{x}(t) &= \dot{x}^*(t) + \epsilon \dot{h}(t), \end{aligned}$$

Aplicando el desarrollo de Taylor para F

$$F(t, x^*(t) + \epsilon h(t), \dot{x}^*(t) + \epsilon \dot{h}(t)) \simeq F(t, x^*(t), \dot{x}^*(t)) + \frac{\partial F}{\partial x} \epsilon h(t) + \frac{\partial F}{\partial \dot{x}} \epsilon \dot{h}(t).$$

Si se define

$$\Delta F = F(t, x^*(t) + \epsilon h(t), \dot{x}^*(t) + \epsilon \dot{h}(t)) - F(t, x^*(t), \dot{x}^*(t)),$$

entonces

$$\begin{aligned} \Delta J &= \int_{t_0}^{t_1} \left(F(t, x^*(t) + \epsilon h(t), \dot{x}^*(t) + \epsilon \dot{h}(t)) - F(t, x^*(t), \dot{x}^*(t)) \right) dt \\ &\simeq \int_{t_0}^{t_1} \left(\frac{\partial F}{\partial x} \epsilon h(t) + \frac{\partial F}{\partial \dot{x}} \epsilon \dot{h}(t) \right) dt. \end{aligned}$$

El cambio $\epsilon h(t)$ es llamado la variación de $x(t)$ y por convención se denota por δx :

$$\delta x \equiv \epsilon h,$$

además

$$\delta \dot{x} \equiv \epsilon \dot{h},$$

entonces podemos escribir

$$\Delta J \simeq \int_{t_0}^{t_1} \left(\frac{\partial F}{\partial x} \epsilon h(t) + \frac{\partial F}{\partial \dot{x}} \epsilon \dot{h}(t) \right) dt.$$

La variación de J es

$$\delta J = \int_{t_0}^{t_1} \left(\frac{\partial F}{\partial x} \epsilon h(t) + \frac{\partial F}{\partial \dot{x}} \epsilon \dot{h}(t) \right) dt.$$

Para $\epsilon = 0$ se tiene que

$$\begin{aligned} \delta J &= \left[\frac{\partial F}{\partial \dot{x}} \delta x \right]_{t_0}^{t_1} + \int_{t_0}^{t_1} \left(\frac{\partial F}{\partial x} - \frac{d}{dt} \left(\frac{\partial F}{\partial \dot{x}} \right) \right) \delta x dt \\ &= \int_{t_0}^{t_1} \left(\frac{\partial F}{\partial x} - \frac{d}{dt} \left(\frac{\partial F}{\partial \dot{x}} \right) \right) \delta x dt = 0. \end{aligned}$$

La **primera variación** del funcional $J(x)$ es

$$\delta J = \int_{t_0}^{t_1} \left(\frac{\partial F}{\partial x} - \frac{d}{dt} \left(\frac{\partial F}{\partial \dot{x}} \right) \right) \delta x dt.$$

Ejemplo 2.2.2 (Continuación del problema de la braquistocrona).

El problema de la braquistocrona es encontrar la trayectoria $y(x)$ tal que $y(0) = 0$ y $y(x_0) = y_0$ para la cual se minimiza el funcional $T(y(x))$, es decir

$$\min T(y) = \int_0^{x_0} \frac{1}{\sqrt{2g}} \sqrt{\frac{1 + (\dot{y}(x))^2}{y(x)}} dx \quad (2.2.21)$$

$$\text{sujeto a } y(0) = 0 \text{ y } y(x_0) = y_0.$$

En este caso, el Lagrangeano F asociado al funcional $T(y)$ es $F(y, \dot{y}) = \frac{1}{\sqrt{2g}} \sqrt{\frac{1 + (\dot{y})^2}{y}}$ y no depende de x .

Si el Lagrangeano F no depende explícitamente de la variable independiente, es decir, $F(x, y, \dot{y}) = F(y, \dot{y})$ entonces la ecuación de Euler (2.2.20) se puede escribir como

$$\dot{y} \frac{\partial F}{\partial x} - F = C_0,$$

con C_0 constante.

Lo anterior se conoce como la *Forma de Beltrami* o *Primera integral de la ecuación de Euler* (consultar Apéndice B.3). Usando esta ecuación para resolver el problema se tiene que

$$\frac{1}{\sqrt{2g}} \frac{(\dot{y})^2}{\sqrt{y(1 + (\dot{y})^2)}} - \frac{1}{\sqrt{2g}} \sqrt{\frac{1 + (\dot{y})^2}{y}} = C_0.$$

Al simplificar el lado izquierdo de la ecuación, y multiplicando por -1 , queda lo siguiente

$$\frac{1}{\sqrt{y(1 + (\dot{y})^2)}} = \sqrt{2g} C_0,$$

entonces

$$y(1 + (\dot{y})^2) = \frac{1}{2gC_0^2}.$$

Si definimos $k^2 = \frac{1}{2gC_0^2}$ como constante y despejando a \dot{y}

$$\dot{y}(x) = \frac{dy}{dx} = \pm \sqrt{\frac{k^2 - y}{y}}.$$

Esta ecuación diferencial de primer orden es separable y se puede resolver. Sin embargo, debido a que el eje vertical y se considera hacia abajo, se espera que la solución $y(x)$ aumente a medida que x aumenta, por lo que se toma el signo positivo en esta ecuación. Se puede demostrar que la solución de esta ecuación diferencial es

$$\begin{aligned} x(\theta) &= \frac{1}{2}k^2(2\theta - \sin 2\theta) + a \\ y(\theta) &= \frac{1}{2}k^2(1 - \cos 2\theta), \end{aligned}$$

donde a es otra constante. Las constantes se eligen de modo que la curva pase a través de los puntos $(0, 0)$ y (x_0, y_0) . Esta curva se conoce como cicloide y es la curva descrita por un punto P en un círculo que rueda sin deslizarse sobre el eje x , de tal manera que P pasa a través de (x_0, y_0) (ver la figura 2.5). Por lo tanto la curva $y(x)$ que minimiza el funcional $T(y)$ es la cicloide.

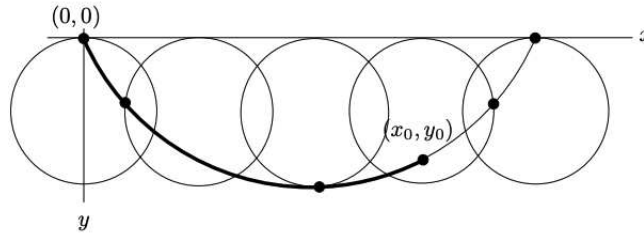


Figura 2.5: Cicloide a través de $(0, 0)$ y (x_0, y_0) .

Hasta el momento se ha considerado que la función $x^*(t)$ sea de clase $C^1([t_0, t_1])$, es decir, $x^*(t)$ tiene primera derivada continua en el intervalo $[t_0, t_1]$. Sin embargo, en algunas aplicaciones no se puede obtener como solución una función de este tipo en dicho intervalo. En lo siguiente y para solucionar los problemas de diferenciabilidad y continuidad, se define una función por pedazos con ciertas propiedades.

Definición 2.2.5. Una función $f(t)$ definida en $[a, b]$ es continua por pedazos (PWC) si es continua sobre un número finito de subintervalos (a_i, a_{i+1}) de $[a, b]$ con extremos finitos a_i y a_{i+1} para $i = 0, \dots, n$ donde $a_0 = a$ y $a_{n+1} = b$.

Definición 2.2.6. Una función $f(t)$ definida en $[a, b]$ es suave por pedazos (PWS) si es:

1. continua en $[a, b]$,
2. diferenciable en $[a, b]$ excepto probablemente en un número finito de puntos,
3. la derivada de $\dot{f}(t)$ es PWC en $[a, b]$.

Con este nuevo tipo de funciones, junto con los teoremas y lemas establecidos (que siguen siendo válidos) para la búsqueda de un extremo admisible, se puede plantear un resultado más general como condición necesaria de optimalidad. Se puede consultar [9] para revisar la prueba del siguiente teorema.

Teorema 2.2.3 (Ecuación de Euler-Lagrange). Una condición necesaria para que $x^*(t)$ (PWS) sea extremo admisible o mínimo local débil de $J(x)$ es

$$F_{\dot{x}}^*(t) = c + \int_a^t F_x^*(s) ds \quad (2.2.22)$$

con $F^* = F(t, x^*(t), \dot{x}^*(t))$, y para alguna constante c .

Observación 2.2.1. La ecuación de Euler es un caso particular de la ecuación de Euler-Lagrange, se puede probar que derivando la ecuación (2.2.22) se recupera la ecuación (2.2.20)

A menos que se diga lo contrario, y con las observaciones hechas acerca de la continuidad de las funciones diferenciables que son admisibles para resolver el problema variacional se considera como ecuación de **Euler-Lagrange** la siguiente expresión

$$F_x - \frac{d}{dt} F_{\dot{x}} = 0, \quad \forall t \in [t_0, t_1]. \quad (2.2.23)$$

2.2.3. Formulación del problema para n funciones

En la formulación del problema básico se consideró el caso escalar, es decir, obtener una función $x^*(t) \in C^2([t_0, t_1])$ definida en \mathbb{R} que verifica una condición inicial y una final, y para la cual el funcional $J(x)$ alcanza el valor óptimo (mínimo o máximo). La intención ahora es extender el problema básico al caso vectorial, en donde en lugar de calcular una única función hay que obtener n funciones $x_1^*(t), x_2^*(t), \dots, x_n^*(t)$.

Sea F una función real de $2n + 1$ variables de clase $C^2([t_0, t_1])$, es decir, con derivadas parciales continuas de orden 2. Sea el siguiente funcional

$$J(x_1, x_2, \dots, x_n) = \int_{t_0}^{t_1} F[t, x_1(t), x_2(t), \dots, x_n(t), \dot{x}_1(t), \dot{x}_2(t), \dots, \dot{x}_n(t)] dt,$$

donde $\dot{x}_i(t)$ es la derivada de $x_i(t)$ con respecto a t , $\forall i \in \{1, 2, \dots, n\}$.

Sean

$$\begin{aligned} \mathbf{x}(t) &= (x_1(t), x_2(t), \dots, x_n(t)) \in \mathbb{R}^n, \\ \dot{\mathbf{x}}(t) &= (\dot{x}_1(t), \dot{x}_2(t), \dots, \dot{x}_n(t)) \in \mathbb{R}^n, \end{aligned}$$

para cada $t \in [t_0, t_1] \subset \mathbb{R}$.

El problema, por lo tanto, en el caso de minimización es:

$$\min_{\mathbf{x}(t) \in \Omega} J(\mathbf{x}(t)) = \int_{t_0}^{t_1} F[t, \mathbf{x}(t), \dot{\mathbf{x}}(t)] dt \quad (2.2.24)$$

$$\begin{aligned} \text{con } x_i(t_0) &= x_i^0, \text{ para } i = 1, 2, \dots, n \\ x_i(t_1) &= x_i^1, \text{ para } i = 1, 2, \dots, n, \end{aligned} \quad (2.2.25)$$

donde

$$\Omega = \{(x_1, x_2, \dots, x_n) : [t_0, t_1] \rightarrow \mathbb{R}^n \mid \forall i = 1, 2, \dots, n \ x_i \in C^2([t_0, t_1])\}.$$

Similarmente al caso anterior, se obtienen n ecuaciones de Euler-Lagrange

$$\frac{\partial}{\partial x_k} F(t, x_1^*(t), \dots, x_n^*(t), \dot{x}_1^*(t), \dots, \dot{x}_n^*(t)) - \frac{d}{dt} \left(\frac{\partial}{\partial \dot{x}_k} F(t, x_1^*(t), \dots, x_n^*(t), \dot{x}_1^*(t), \dots, \dot{x}_n^*(t)) \right) = 0,$$

para cada $t \in [t_0, t_1]$, $k \in \{1, 2, \dots, n\}$. Para simplificar la notación se reescribe la expresión anterior como

$$F_{x_k} - \frac{d}{dt} (F_{\dot{x}_k}) = 0. \quad (2.2.26)$$

2.2.4. Problemas variacionales con restricciones

Sea el problema básico

$$\min_{x \in \Omega} J(x) = \int_{t_0}^{t_1} F[t, x(t), \dot{x}(t)] dt$$

$$\text{con } x(t_0) = x_0, \ x(t_1) = x_1.$$

En la formulación del problema se pide que las funciones admisibles cumplen una condición inicial $x(t_0) = x_0$, y una condición final $x(t_1) = x_1$ para t_0 y t_1 valores fijos. Estas *condiciones de frontera* sobre la función $x(t)$ son también llamadas *condiciones transversales* o de *transversalidad* y permiten determinar la solución particular del problema a partir de la solución general de la ecuación de Euler-lagrange.

Las condiciones de *transversalidad* pueden generalizarse en diferentes casos ([3] pp.45, 52, 56, 59):

■ Caso 1. Condición final t_1 dada y condición final x_1 libre.

Supóngase que se da t_1 y x_1 es libre. Serán admisibles aquellas funciones, de clase $C^2([t_0, t_1])$ que parten del punto (t_0, x_0) y que lleguen a algún punto de la recta $t = t_1$, como se representa en la Figura 2.6(a).

■ **Caso 2. Condición final t_1 libre y condición final x_1 dada.**

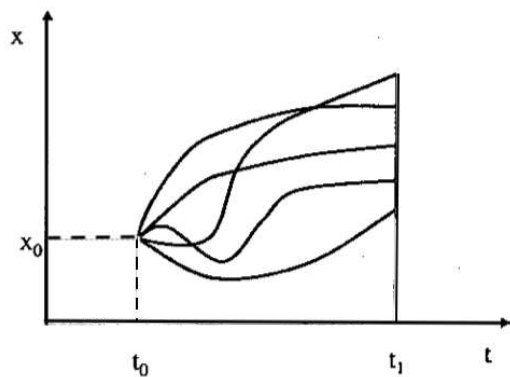
Supóngase que t_1 es un valor por determinar y $x(t_1) = x_1$ es dado. Serán admisibles aquellas funciones de clase $C^2([t_0, t_1])$ que parten del punto (t_0, x_0) y que lleguen a algún punto (t_1, x_1) , con $t_1 > t_0$ y x_1 un valor fijo como se representa en la Figura (2.6(b)).

■ **Caso 3. Condición final t_1 libre y condición final x_1 libre.**

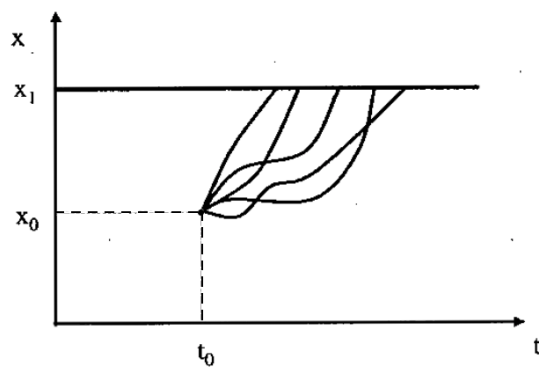
Supóngase que t_1 y x_1 son valores por determinar. Serán admisibles aquellas funciones de clase $C^2([t_0, t_1])$ que parten del punto (t_0, x_0) y que lleguen a algún punto (t_1, x_1) , con $t_1 > t_0$, $t_1 \in \mathbb{R}$ y $x_1 \in \mathbb{R}$ como se representa en la Figura (2.6(c)).

■ **Caso 4. Condición final t_1 y condición final x_1 relacionados mediante una función.**

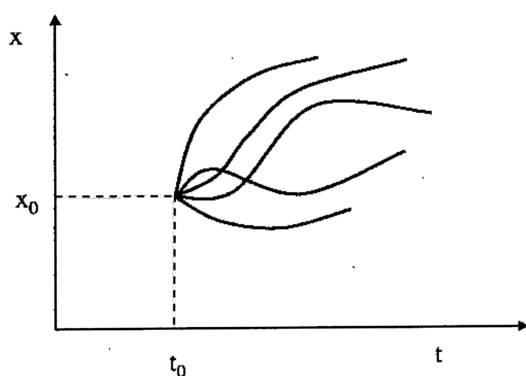
Supóngase que t_1 es una valor libre y $x_1 = \phi(t_1)$, siendo $\phi(t_1) \in C^1([t_0, t_1])$. Serán admisibles aquellas funciones de clase $C^2([t_0, t_1])$ que parten del punto (t_0, x_0) y que lleguen a algún punto (t_1, x_1) , con $t_1 > t_0$, $t_1 \in \mathbb{R}$ y $x_1 = \phi(t_1)$ como se representa en la Figura (2.6(d)).



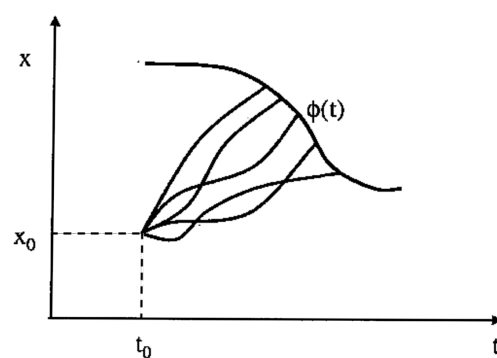
(a) Caso 1.



(b) Caso 2.



(c) Caso 3.



(d) Caso 4.

Figura 2.6: Distintos casos de las condiciones de transversalidad.

Existen problemas variacionales en los que, además de las condiciones iniciales y finales que tienen que cumplir las funciones admisibles, se consideran restricciones impuestas por la naturaleza del problema, que deben incorporarse a la formulación del mismo y que tendrán trascendencia en la solución óptima.

A continuación se describe un problema variacional cuyas restricciones son ecuaciones diferenciales, es decir, las restricciones dependen también de las derivadas de las n funciones x_i , siendo por tanto de la forma:

$$g_k(t, x_1, \dots, x_n, \dot{x}_1, \dots, \dot{x}_n) = 0,$$

para $k = 1, \dots, m$.

El problema es el siguiente:

$$\begin{aligned}
 \min_{(x_1, x_2, \dots, x_n) \in \Omega} J(x_1, x_2, \dots, x_n) &= \int_{t_0}^{t_1} F[t, x_1, x_2, \dots, x_n, \dot{x}_1, \dot{x}_2, \dots, \dot{x}_n] dt \\
 \text{sujeto a : } g_1(t, x_1, \dots, x_n, \dot{x}_1, \dots, \dot{x}_n) &= 0 \\
 &\vdots \\
 g_m(t, x_1, \dots, x_n, \dot{x}_1, \dots, \dot{x}_n) &= 0 \\
 \text{con : } x_i(t_0) = x_i^0, \quad x_i(t_1) = x_i^1 &\text{ para } i = 1, \dots, n.
 \end{aligned} \tag{2.2.27}$$

Las funciones g_k , para $k = 1, \dots, m$ son de clase $C^2([t_0, t_1])$ y $m < n$ (menos restricciones que variables). Para simplificar la notación del problema, tomar

$$\begin{aligned}
 \min_{\mathbf{x}(t) \in \Omega} J(\mathbf{x}(t)) &= \int_{t_0}^{t_1} F[t, \mathbf{x}(t), \dot{\mathbf{x}}(t)] dt \\
 \text{sujeto a : } \mathbf{g}(t, \mathbf{x}(t), \dot{\mathbf{x}}(t)) &= 0 \\
 \text{con : } \mathbf{x}(t_0) &= \mathbf{x}_0,
 \end{aligned} \tag{2.2.28}$$

para $\mathbf{x}, \dot{\mathbf{x}}, \mathbf{x}_0 \in \mathbb{R}^n$ y $\mathbf{g} \in \mathbb{R}^m$.

En la siguiente sección se describe el *Método de los Multiplicadores de Lagrange* que, adaptado a un problema variacional, resuelve el caso en que las restricciones son ecuaciones diferenciales.

2.2.5. Multiplicadores de Lagrange

Muchos problemas de optimización tienen restricciones, para los valores que pueden usarse para encontrar la solución óptima. Tales restricciones suelen complicar los problemas porque la solución puede estar en un punto frontera del dominio.

El *Método de los Multiplicadores de Lagrange* es un procedimiento que sirve para encontrar los máximos y mínimos de funciones de n variables sujetas a m restricciones. En otras palabras, el *Método de los Multiplicadores de Lagrange* permite reducir el problema restringido con n variables y m restricciones, a un problema sin restricciones de $n + m$ variables con ecuaciones que pueden ser resueltas más fácilmente. Estas nuevas variables escalares desconocidas, una para cada restricción, son llamadas *Multiplicadores de Lagrange* y suelen denotarse con la letra griega λ .

Considerar el caso de funciones de dos variables $f(x, y)$. Dos curvas son tangentes si sus vectores normales son paralelos. En consecuencia, dado que los vectores normales a dichas curvas son los vectores gradientes se tiene que, en el punto de tangencia el vector tangente $\vec{\nabla} f(x, y)$ debe ser un múltiplo escalar del vector gradiente $\vec{\nabla} g(x, y)$. Luego,

$$\vec{\nabla} f(x, y) = \lambda \vec{\nabla} g(x, y),$$

donde el escalar λ se conoce como *Multiplicadores de Lagrange*. Las condiciones necesarias para la existencia de tales multiplicadores vienen dadas por el siguiente teorema, (ver [11]).

Teorema 2.2.4 (Teorema de Lagrange). *Sean $f = f(x, y)$ y $g = g(x, y)$ funciones con primeras derivadas parciales continuas tal que f tiene un valor extremo en el punto (x_0, y_0) sobre la curva de la restricción $g(x, y) = 0$. Si $\vec{\nabla}g(x, y) \neq 0$, entonces existe un número real λ tal que*

$$\vec{\nabla}f(x, y) = \lambda \vec{\nabla}g(x, y).$$

El Teorema de Lagrange también se cumple para funciones de n variables.

El Método de los multiplicadores de Lagrange usa el Teorema anterior para hallar los extremos de una función f sujeta a la restricción $g(x, y) = 0$.

Teorema 2.2.5 (Método de los multiplicadores de Lagrange). *Si $f = f(x, y)$ y $g = g(x, y)$ satisfacen las hipótesis del Teorema 2.2.4 y f tiene un máximo o mínimo sujeto a la restricción $g(x, y) = 0$, entonces dicho extremo se produce en uno de los puntos críticos de la función L dada por,*

$$L(x, y, \lambda) = f(x, y) + \lambda g(x, y). \quad (2.2.29)$$

La función (2.2.29) recibe el nombre de *Función auxiliar*. Los puntos críticos son determinados por las soluciones del sistema

$$\begin{aligned} \frac{\partial}{\partial x}L(x, y, \lambda) = 0 &\longrightarrow \frac{\partial}{\partial x}f(x, y) + \lambda \frac{\partial}{\partial x}g(x, y) = 0, \\ \frac{\partial}{\partial y}L(x, y, \lambda) = 0 &\longrightarrow \frac{\partial}{\partial y}f(x, y) + \lambda \frac{\partial}{\partial y}g(x, y) = 0, \\ \frac{\partial}{\partial \lambda}L(x, y, \lambda) = 0 &\longrightarrow g(x, y) = 0. \end{aligned}$$

El método se puede generalizar. Para encontrar el máximo o el mínimo de una función f de n variables $x_1, x_2, x_3, \dots, x_n$,

$$f = f(x_1, x_2, x_3, \dots, x_n),$$

bajo las restricciones,

$$\begin{aligned} g_1(x_1, x_2, x_3, \dots, x_n) &= 0, \\ g_2(x_1, x_2, x_3, \dots, x_n) &= 0, \\ g_3(x_1, x_2, x_3, \dots, x_n) &= 0, \\ &\vdots \\ g_n(x_1, x_2, x_3, \dots, x_n) &= 0. \end{aligned}$$

cuya función auxiliar es

$$\begin{aligned}
L(x_1, \dots, x_n, \lambda_1, \dots, \lambda_n) &= f(x_1, \dots, x_n) + \lambda_1 g_1(x_1, \dots, x_n) \\
&\quad + \lambda_2 g_2(x_1, \dots, x_n) + \lambda_3 g_3(x_1, \dots, x_n) \\
&\quad + \dots + \lambda_n g_n(x_1, \dots, x_n),
\end{aligned}$$

o,

$$L(x_1, \dots, x_n, \lambda_1, \dots, \lambda_n) = f(x_1, \dots, x_n) + \sum_{i=1}^n \lambda_i g_i(x_1, \dots, x_n),$$

donde $\lambda_1, \dots, \lambda_n$ son los multiplicadores de Lagrange, independientes de las variables x_1, \dots, x_n . Los puntos críticos son determinados por la solución del sistema,

$$\begin{aligned}
\frac{\partial L}{\partial x_1} &= 0 \longrightarrow \frac{\partial f}{\partial x_1} + \lambda_1 \frac{\partial g_1}{\partial x_1} = 0, \\
\frac{\partial L}{\partial x_2} &= 0 \longrightarrow \frac{\partial f}{\partial x_2} + \lambda_2 \frac{\partial g_2}{\partial x_2} = 0, \\
&\vdots \\
\frac{\partial L}{\partial x_n} &= 0 \longrightarrow \frac{\partial f}{\partial x_n} + \lambda_n \frac{\partial g_n}{\partial x_n} = 0, \\
\frac{\partial L}{\partial \lambda_1} &= 0 \longrightarrow g_1 = 0, \\
\frac{\partial L}{\partial \lambda_2} &= 0 \longrightarrow g_2 = 0, \\
&\vdots \\
\frac{\partial L}{\partial \lambda_n} &= 0 \longrightarrow g_n = 0.
\end{aligned}$$

Sea el siguiente problema variacional 2.2.28. Así que, para resolver el problema cuyas restricciones son ecuaciones diferenciales se define el siguiente Lagrangeano asociado a $J(x)$

$$L(t, \mathbf{x}, \dot{\mathbf{x}}, \boldsymbol{\lambda}, \dot{\boldsymbol{\lambda}}) = F(t, \mathbf{x}, \dot{\mathbf{x}}) = \sum_{k=1}^m \lambda_k(t) g_k(t, x, \dot{x}), \quad (2.2.30)$$

donde $\dot{\mathbf{x}}$ (vector de derivadas de las variables independientes x_i , con $i = 1, \dots, n$), $\mathbf{x}, \boldsymbol{\lambda}, \dot{\boldsymbol{\lambda}} \in \mathbb{R}^n$.

Además, para la función L se tiene que cumplir las ecuaciones de Euler-Lagrange

$$\begin{aligned}
L_{x_i} - \frac{d}{dt} L_{\dot{x}_i} &= 0, \text{ para } i = 1, \dots, n, \\
L_{\lambda_k} - \frac{d}{dt} L_{\dot{\lambda}_k} &= 0, \text{ para } k = 1, \dots, m.
\end{aligned}$$

En la siguiente sección se desarrollará con más detalle el uso de los Multiplicadores de Lagrange para resolver un problema variacional con restricciones.

2.3. Control Óptimo

En esta sección se proporcionan los conceptos y algunos resultados básicos principales de la teoría de control óptimo, una teoría que se puede considerar como una extensión del cálculo variacional y que ayuda a comprender problemas más generales. Se aborda el concepto de controlabilidad en su generalidad para sistemas lineales y no lineales. Además se presenta el problema básico de control en tiempo continuo y el resultado fundamental para obtener la solución óptima del problema: el *principio del máximo de Pontryagin* (en su versión lineal). Se siguen las ideas expuestas en “An Introduction to Mathematical Optimal Control Theory” [12] y se usan las herramientas desarrolladas en [9, 10, 13], desde la perspectiva del cálculo variacional.

2.3.1. Controlabilidad

Se considera un sistema dinámico, formulado en tiempo continuo en el intervalo de tiempo $[t_0, t_1]$, cuya condición inicial está dada por el vector $\mathbf{x}(t_0) = \mathbf{x}_0 \in \mathbb{R}^n$, y que evoluciona en el tiempo. Dicha evolución depende del valor que se asigna a ciertas variables, llamadas variables de control, que permiten influir en el sistema. Sea $\mathbf{u}(t) \in \mathcal{U} \subset \mathbb{R}^m$ el vector de variables de control y $\mathbf{x}(t) \in \mathbb{R}^n$ el vector de variables de estado. Al conjunto \mathcal{U} que representa las restricciones físicas al valor de las variables de control se le conoce como *la región de control*.

La controlabilidad es uno de los conceptos fundamentales en la teoría de control moderna, esta es una propiedad cualitativa de los sistemas de control. El estudio sistemático de la controlabilidad se basa en la descripción matemática de un sistema dinámico; muchos sistemas son tales que el control no afecta el estado completo sino solo una parte del mismo. Por otro lado, muy a menudo en procesos reales es posible observar solo una cierta parte del estado completo del sistema dinámico.

Por lo tanto, es muy importante determinar si es posible o no controlar el estado completo del sistema. En términos generales, la controlabilidad significa que es posible dirigir el sistema dinámico desde un estado inicial arbitrario a un estado final, utilizando el conjunto de controles admisibles (variables de control).

Es común que en los problemas de control la propiedad de controlabilidad se estudie en sistemas dinámicos lineales. En este trabajo se enuncian las definiciones y algunas propiedades básicas de controlabilidad para el caso lineal y después enunciar resultados generales para los sistemas no lineales. Consultar [14, 15] para más detalles técnicos y profundizar en el estudio de la controlabilidad en sistemas no lineales.

Se define el siguiente sistema lineal tanto para la variable de estado \mathbf{x} como para la variable de control \mathbf{u} .

$$\begin{aligned}\dot{\mathbf{x}}(t) &= M\mathbf{x}(t) + N\mathbf{u}(t) \\ \mathbf{x}(0) &= \mathbf{x}_0,\end{aligned}\tag{2.3.31}$$

donde $M \in \mathbb{M}^{n \times n}$ y $N \in \mathbb{M}^{n \times m}$. Además se asume que los parámetros de control A es un cubo en \mathbb{R}^m :

$$A = [-1, 1]^m = \{\mathbf{a} \in \mathbb{R}^m : |a_i| \leq 1, i = 1, \dots, m\}$$

y la siguiente región de control

$$\mathcal{U} = \{\mathbf{u} : (0, \infty) \rightarrow A \mid \mathbf{u}(\cdot) \text{ es medible}\}.$$

Por el teorema (2.1.2) la solución al sistema no homogéneo (2.3.31) para un control dado $\mathbf{u}(t)$ es

$$\mathbf{x}(t) = \mathbf{X}(t)\mathbf{x}_0 + \mathbf{X}(t) \int_0^t \mathbf{X}^{-1}(s)N\mathbf{u}(s)ds,$$

donde $\mathbf{X}(t) = e^{tM}$.

Definición 2.3.1. Se define el conjunto alcanzable para el tiempo t como

$$\begin{aligned}\mathcal{C}(t) &= \text{conjunto de puntos iniciales } \mathbf{x}_0 \text{ para el} \\ &\quad \text{que existe un control tal que } \mathbf{x}(t) = \mathbf{0},\end{aligned}$$

y para cualquier instante de tiempo t

$$\begin{aligned}\mathcal{C} &= \text{conjunto de puntos iniciales } \mathbf{x}_0 \text{ para el que existe un} \\ &\quad \text{control tal que } \mathbf{x}(t) = \mathbf{0} \text{ para algún tiempo finito } t.\end{aligned}$$

Notar que

$$\mathcal{C} = \bigcup_{t \geq 0} \mathcal{C}(t).$$

Por simplicidad, se considera llevar el sistema al origen.

Definición 2.3.2. Se dice que el sistema (2.3.31) es controlable en el tiempo t_f si para cualquier estado inicial \mathbf{x}_0 y cualquier estado objetivo $\mathbf{x}_f = \mathbf{x}(t_f)$, existe un control $\mathbf{u}(t)$ que puede dirigir el sistema de \mathbf{x}_0 a \mathbf{x}_f durante el intervalo de tiempo $[0, t_f]$, (sin importar qué trayectorias siga el estado).

A continuación se establecen algunas condiciones algebraicas generales que garantizan que el conjunto \mathcal{C} contiene una vecindad en el origen.

Definición 2.3.3. *La matriz de controlabilidad es*

$$G = G(M, N) := [N, MN, M^2N, \dots, M^{n-1}N],$$

donde la matriz G es de tamaño $n \times (mn)$.

Teorema 2.3.1 (Matriz de controlabilidad). *Se tiene que*

$$\text{rank } G = n \Leftrightarrow 0 \in \mathcal{C}^\circ,$$

donde \mathcal{C}° es el interior del conjunto \mathcal{C} .

Demostración. Se puede consultar en C.1. □

El siguiente criterio es de gran importancia, ya que relaciona la propiedad de controlabilidad de un sistema con la estabilidad del mismo.

Teorema 2.3.2 (Criterio de controlabilidad). *Sea A el cubo $[-1, 1]^n$ en \mathbb{R}^n . Suponer también que $\text{rank } G = n$, y $\text{Re } \lambda < 0$ para cada valor propio λ de la matriz M . Entonces el sistema (2.3.31) es controlable.*

Demostración. Se puede consultar en (C.3). □

En las aplicaciones, muchas veces los sistemas dinámicos que se plantean son no lineales, así que se estudia la controlabilidad de manera local para dichos sistemas.

Considerar el sistema dinámico

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t)), \\ \mathbf{x}(0) &= \mathbf{x}_0, \end{aligned} \tag{2.3.32}$$

donde $\mathbf{x}_0 \in \mathbb{R}^n$, $\mathbf{f} : \mathbb{R} \times A \subset \mathbb{R}^m \rightarrow \mathbb{R}^n$, $\mathbf{u} : [0, \infty) \rightarrow A$ es la variable de control, y $\mathbf{x} : [0, \infty) \rightarrow \mathbb{R}^n$ es la respuesta del sistema.

Dado un punto $\bar{\mathbf{x}} \in \mathbb{R}^n$, se dice que el sistema es *localmente controlable* en $\bar{\mathbf{x}}$, si para cada $t > 0$, el conjunto $\mathcal{C}(t)$ contiene una vecindad de $\bar{\mathbf{x}}$. En términos generales, esto significa que el sistema se puede dirigir desde $\bar{\mathbf{x}}$ a todos los puntos cercanos, dentro de un pequeño intervalo de tiempo.

A partir de los resultados anteriores de controlabilidad para sistemas lineales y, mediante un argumento de linealización se puede deducir un resultado de controlabilidad local, válido a partir de sistemas generales no lineales.

Teorema 2.3.3 (Controlabilidad local). *Considerar el sistema de control (2.3.32) y suponer que el conjunto \mathcal{U} de controles admisibles contiene a una vecindad de el origen $\mathbf{0} \in \mathbb{R}^m$. Dado un punto $\bar{\mathbf{x}} \in \mathbb{R}^n$, suponer que*

1. $\mathbf{f}(\bar{\mathbf{x}}, 0) = 0$ (puntos de equilibrio),
2. Definir la matriz de derivadas parciales de \mathbf{f} con respecto a \mathbf{x} y \mathbf{u} calculadas en el punto de equilibrio $(\bar{\mathbf{x}}, 0)$

$$A = D_{\mathbf{x}}\mathbf{f}(\bar{\mathbf{x}}, 0), \quad B = D_{\mathbf{u}}\mathbf{f}(\bar{\mathbf{x}}, 0), \quad (2.3.33)$$

el sistema linealizado

$$\dot{\mathbf{x}} = A \cdot (\mathbf{x} - \bar{\mathbf{x}}) + B \cdot \mathbf{u} \quad (2.3.34)$$

es completamente controlable, es decir, A, B satisfacen $\text{rank } G(A, B) = n$.

Entonces el sistema (2.3.32) es localmente controlable en el punto $\bar{\mathbf{x}}$.

Demostración. La demostración se omite en este trabajo pero se puede consultar en [14]. \square

2.3.2. Planteamiento del problema básico.

Sea el siguiente sistema dinámico

$$\begin{aligned} \dot{x}_1(t) &= f_1(t, x_1, x_2, \dots, x_n, u_1, u_2, \dots, u_m), \\ \dot{x}_2(t) &= f_2(t, x_1, x_2, \dots, x_n, u_1, u_2, \dots, u_m), \\ &\vdots \\ \dot{x}_n(t) &= f_n(t, x_1, x_2, \dots, x_n, u_1, u_2, \dots, u_m), \end{aligned} \quad (2.3.35)$$

donde las funciones f_1, \dots, f_n son funciones $f_i : [t_0, \infty) \times \mathbb{R}^n \times \mathbb{R}^m \longrightarrow \mathbb{R}$ prefijadas y conocidas que describen la dinámica del sistema en términos de las variables de estado y las variables de control. Además que las derivadas parciales

$$\frac{\partial f_i}{\partial t}, \frac{\partial f_i}{\partial x_k}, \frac{\partial f_i}{\partial u_j}$$

existen y son continuas para cada $i, k = 1, \dots, n; j = 1, \dots, m$.

Asumir que el vector de variables de estado $\mathbf{x}(t)$ es una función de t , continuas en $[t_0, t_1]$, y con derivadas continuas por pedazos, tal que

$$\dot{\mathbf{x}}(t) = \mathbf{f}(t, \mathbf{x}(t), \mathbf{u}(t)), \quad (2.3.36)$$

en todos los puntos de continuidad de $\mathbf{u}(t)$. Si se toma el sistema dinámico (2.3.35), una condición inicial $\mathbf{x}(t_0) = \mathbf{x}_0 \in \mathbb{R}^n$ y se deja fija la variable de control $\mathbf{u}(t)$ definida para todo $t \geq t_0$, se tiene un sistema de n ecuaciones diferenciales de primer orden con n funciones desconocidas $x_1(t), \dots, x_n(t)$. Ya que se conoce el valor de \mathbf{x}_0 , por el Teorema fundamental para sistemas lineales (2.1.2), se asegura la unicidad de la solución $\mathbf{x}(t)$.

El objetivo es tener una medida cuantitativa del comportamiento del sistema en el tiempo para un control $\mathbf{u}(t)$.

Considerar funcionales del tipo

$$J = \int_{t_0}^{t_1} F[t, \mathbf{x}(t), \mathbf{u}(t)] dt + S[\mathbf{x}(t_1)], \quad (2.3.37)$$

donde $F : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$, $S : \mathbb{R}^n \rightarrow \mathbb{R}$ y, además, tengan primeras derivadas parciales continuas.

Observación 2.3.1. *El primer sumando del funcional J es una integral que depende de los valores que toman $\mathbf{x}(t)$ y $\mathbf{u}(t)$ a lo largo del tiempo. El segundo sumando $S[\mathbf{x}(t_1)]$ estima el estado en que queda el sistema al final del intervalo de tiempo.*

Un control óptimo se define como un control admisible que optimiza el funcional objetivo (2.3.37).

Por lo tanto, el problema de interés es el siguiente:

“Dado un sistema dinámico con condición inicial \mathbf{x}_0 , y que evoluciona de acuerdo con la ecuación de estado $\dot{\mathbf{x}}(t) = \mathbf{f}(t, \mathbf{x}(t), \mathbf{u}(t))$, se trata de encontrar el vector de control $\mathbf{u}(t)$ que sea admisible y haga que el funcional objetivo alcance el valor óptimo (máximo o mínimo)” ([10]).

En términos matemáticos el problema es:

$$\begin{aligned} \text{máx } J(\mathbf{u}) &= \int_{t_0}^{t_1} F[t, \mathbf{x}, \mathbf{u}] dt + S[\mathbf{x}(t_1)], \\ \text{sujeto a : } \dot{\mathbf{x}} &= \mathbf{f}(t, \mathbf{x}, \mathbf{u}), \\ \text{con : } \mathbf{x}(t_0) &= \mathbf{x}_0, \mathbf{u}(t) \in \mathcal{U}. \end{aligned} \quad (2.3.38)$$

Se sabe que el problema variacional de mín $J(x)$ es equivalente a máx $[-J(x)]$, entonces es posible plantear un problema de control para maximizar un funcional objetivo.

2.3.3. Diferentes formas del problema básico y un caso especial

Considerar el siguiente problema de control óptimo en tiempo continuo:

$$\begin{aligned} \text{máx } J(\mathbf{u}) &= \int_{t_0}^{t_1} F[t, \mathbf{x}, \mathbf{u}] dt + S[\mathbf{x}(t_1)], \\ \text{sujeto a : } \dot{\mathbf{x}} &= \mathbf{f}(t, \mathbf{x}, \mathbf{u}), \\ \text{con : } \mathbf{x}(t_0) &= \mathbf{x}_0, \mathbf{u}(t) \in \mathcal{U}. \end{aligned} \quad (2.3.39)$$

Se dice que el problema de control anterior está en la forma de *Bolza*.

Si el funcional objetivo es del tipo

$$J = S[\mathbf{x}(t_1)],$$

es decir, si $F = 0$, se dice que está en la forma de *Mayer*.

Si el funcional objetivo es del tipo

$$J = \int_{t_0}^{t_1} F[t, \mathbf{x}, \mathbf{u}] dt, \quad (2.3.40)$$

es decir, si $S = 0$, se dice que está en la forma de *Lagrange*.

Como caso especial, considerar el problema donde la elección de u no está restringida, y donde la ecuación de estado toma la forma más simple $\dot{x} = u$.

Entonces el problema de control óptimo se convierte en

$$\begin{aligned} \text{máx } J &= \int_{t_0}^{t_1} F[t, x, u] dt, \\ \text{sujeto a : } \dot{x} &= u, \\ \text{con : } x(t_0) &= x_0, \text{ y } x(t_1) = x_1. \end{aligned}$$

Sin embargo, al sustituir la ecuación de estado en el Lagrangeano F se puede reescribir el problema como

$$\begin{aligned} \text{máx } J &= \int_{t_0}^{t_1} F[t, x, \dot{x}] dt, \\ \text{con : } x(t_0) &= x_0, \text{ y } x(t_1) = x_1. \end{aligned}$$

Este es justo el problema del cálculo variacional. El vínculo fundamental entre el cálculo variacional y la teoría del control óptimo es así de evidente. Pero las ecuaciones de estado encontradas en los problemas de control óptimo son, en general, más complicadas.

2.3.4. El principio del máximo de Pontryagin

Para comenzar se dice que la variable de control $\mathbf{u}^*(t)$ que resuelve el problema se llama control óptimo y el vector $\mathbf{x}^*(t)$, determinado por la ecuación de estado a partir de $\mathbf{u}^*(t)$, se denomina trayectoria de estado óptima o camino óptimo.

El principio del máximo proporciona condiciones necesarias que debe cumplir el control óptimo del problema que se ha considerado.

El problema que se quiere resolver es el siguiente

$$\begin{aligned} \text{máx } J(\mathbf{u}) &= \int_{t_0}^{t_1} F[t, \mathbf{x}, \mathbf{u}] dt + S[\mathbf{x}(t_1)], \\ \text{sujeto a : } \dot{\mathbf{x}} &= \mathbf{f}(t, \mathbf{x}, \mathbf{u}), \\ \text{con : } \mathbf{x}(t_0) &= \mathbf{x}_0, \mathbf{u}(t) \in \mathcal{U}. \end{aligned}$$

La formulación de este problema es similar al problema variacional 2.2.28, en cuya solución se construye una función auxiliar $L(t, \mathbf{x}, \dot{\mathbf{x}}, \boldsymbol{\lambda}, \dot{\boldsymbol{\lambda}})$ (2.2.30), como consecuencia de aplicar el método de los multiplicadores de Lagrange.

De manera similar en el problema de control óptimo se obtiene una función auxiliar. Se define el “Hamiltoniano” asociado al problema de control óptimo de la siguiente forma:

$$H(t, \mathbf{x}, \mathbf{u}, \boldsymbol{\lambda}) = F(t, \mathbf{x}, \mathbf{u}) + \boldsymbol{\lambda} \cdot \mathbf{f}(t, \mathbf{x}, \mathbf{u}) = F(t, \mathbf{x}, \mathbf{u}) + \sum_{i=1}^n \lambda_i f_i(t, \mathbf{x}, \mathbf{u}), \quad (2.3.41)$$

donde $H : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \rightarrow \mathbb{R}$ para cada $t \in [t_0, t_1]$, $\boldsymbol{\lambda} = \boldsymbol{\lambda}(t) \in \mathbb{R}^n$ que se llama vector de variables de “coestado”.

La idea principal del principio del máximo de Pontryagin es relacionar nuestro problema de encontrar una $\mathbf{u}(t)$ que, satisfaga las restricciones impuestas y optimice el funcional objetivo J con el problema de optimizar el Hamiltoniano respecto a la variable $\mathbf{u} \in \mathcal{U} \subset \mathbb{R}^m$.

El siguiente teorema constituye el resultado fundamental para este trabajo y establece condiciones necesarias de optimalidad para el problema que nos ocupa.

Teorema 2.3.4 (Principio del máximo de Pontryagin). *Sea $\mathbf{u}^*(t)$ la trayectoria óptima de control, continua a trozos, y $\mathbf{x}^*(t)$ la trayectoria de estado óptima asociada, definidas en el intervalo $[t_0, t_1]$. Entonces existe una función vectorial $\boldsymbol{\lambda}^*(t)$ continua que posee primeras derivadas continuas por pedazos, tal que para cada $t \in [t_0, t_1]$ se cumple:*

1.

$$\dot{\lambda}_i^* = -\frac{\partial}{\partial x_i} H(t, \mathbf{x}^*(t), \mathbf{u}^*(t), \boldsymbol{\lambda}^*(t)),$$

para cada $i = 1, \dots, n$ en todos los puntos de continuidad de $\mathbf{u}^*(t)$, además

$$\lambda_1^* = \frac{\partial}{\partial x_1} S[\mathbf{x}^*(t_1)], \dots, \lambda_n^* = \frac{\partial}{\partial x_n} S[\mathbf{x}^*(t_1)],$$

donde al evaluar $t = t_1$ en cada λ_n^* se obtienen las condiciones de transversalidad.

El sistema de ecuaciones diferenciales formado por las $\dot{\lambda}_i^*$ se denomina “sistema adjunto” o “sistema de coestado”. A las funciones $\lambda_i^*(t)$ se les denomina “variables adjuntas” o “variables de coestado”.

2. Para el caso de maximizar, se cumple que

$$H(t, \mathbf{x}^*(t), \mathbf{u}^*(t), \boldsymbol{\lambda}^*(t)) \geq H(t, \mathbf{x}^*(t), \mathbf{u}(t), \boldsymbol{\lambda}^*(t)),$$

para cada $\mathbf{u}(t) \in \mathcal{U}$ (se define de manera similar cambiando la desigualdad en el caso de minimizar el funcional objetivo).

3. La trayectoria de estado óptimo $\mathbf{x}^*(t)$ es solución del sistema (2.3.36) con condición inicial \mathbf{x}_0^*

$$\dot{\mathbf{x}}(t) = \mathbf{f}^*(t, \mathbf{x}^*(t), \mathbf{u}^*(t)),$$

para cada $i = 1, \dots, n$ en todos los puntos de continuidad de $\mathbf{u}^*(t)$,

con

$$\mathbf{x}^*(t) = \begin{pmatrix} x_1^*(t) \\ x_2^*(t) \\ \vdots \\ x_n^* \end{pmatrix} \quad y \quad \mathbf{u}^*(t) = \begin{pmatrix} u_1^*(t) \\ u_2^*(t) \\ \vdots \\ u_m^*(t) \end{pmatrix}.$$

Demostración. Se puede consultar en C.4. □

Ejemplo 2.3.1. *En el siguiente problema consiste en calcular el valor de las variables de control, de estado y de coestado, para los que se verifica el principio del máximo.*

$$\begin{aligned} \text{máx } J &= \int_0^1 (x + u) dt, \\ \text{sujeto a : } \dot{x} &= 1 - u^2, \\ \text{con : } x(0) &= 1, \quad x(1) \text{ libre.} \end{aligned}$$

El Hamiltoniano asociado es

$$H(t, x, u, \lambda) = F + \lambda f = x + u + \lambda(1 - u^2).$$

Por la condición 1. del principio del máximo:

$$\dot{\lambda} = -\frac{\partial H}{\partial x} = -1.$$

En este caso notar que $\lambda(1) = 0$, ya que $\frac{d}{dx}(S[x^(1)]) = 0$ por ser $S[x(1)] = 0$.*

Derivando el Hamiltoniano con respecto a x , se obtiene

$$\dot{\lambda} = -1, \text{ con } \lambda(1) = 0,$$

que es una ecuación diferencial de primer orden con condición inicial dada. Al resolver dicha ecuación diferencial queda

$$\lambda(t) = -t + c_0, \text{ con } \lambda(1) = -1 + c_0,$$

donde $c_0 = 1$ y por lo tanto

$$\lambda^*(t) = 1 - t, \text{ para } t \in [0, 1].$$

Para la condición 2.

Se tiene que calcular el

$$\text{máx}_u H(t, x^*, u^*, \lambda^*),$$

donde u no tiene ninguna restricción en este caso. Hay que resolver el siguiente problema de optimización

$$\text{máx}_u H = x^* + u + \lambda^*(1 - u^2).$$

Sustituyendo $\lambda^* = 1 - t$ queda:

$$\max_u H = x^* + u + (1 - t)(1 - u^2).$$

Si se toma la derivada del Hamiltoniano con respecto a u y se iguala a cero, tenemos un punto critico

$$\frac{\partial H}{\partial u} = 1 - 2u(1 - t) = 0,$$

entonces

$$u = \frac{1}{2(1 - t)}, \text{ si } t \neq 1.$$

Además

$$\frac{\partial^2 H}{\partial u^2} = 2t - 2 \leq 0, \text{ para } t \in [0, 1],$$

luego corresponde a un máximo. Por lo tanto:

$$u^*(t) = \frac{1}{2(1 - t)} \text{ para } t \in [0, 1].$$

Así que, para calcular la trayectoria de estado óptima, se resuelve la ecuación de estado del problema, sustituyendo el control por el valor obtenido y haciendo uso de la condición inicial dada. Lo anterior corresponde a la condición 3 del principio del máximo:

$$\dot{x}^* = 1 - (u^*)^2 = 1 - \frac{1}{4(1 - t)^2}, \text{ con } x^*(0) = 1.$$

Al resolver la ecuación diferencial anterior se tiene que:

$$x^*(t) = t - \int \frac{1}{4(1 - t)^2} dt = t - \frac{1}{4}(1 - t)^{-1} + c_1, \text{ con } x^*(0) = 1.$$

Al imponer la condición inicial dada se obtiene el valor de c_1 :

$$1 = x^*(0) = -\frac{1}{4} + c_1, \text{ de donde } c_1 = \frac{5}{4}.$$

Por lo tanto, la trayectoria de estado es

$$x^*(t) = t - \frac{1}{4(1 - t)} + \frac{5}{4}, \text{ para } t \in [t_0, t_1].$$

2.3.5. Caso particular del principio del máximo de Pontryagin

Se ha presentado la formulación y el problema básico de control óptimo en tiempo continuo. A continuación se considera que el instante inicial t_0 y el instante final t_1 son valores dados, el estado inicial $\mathbf{x}(t_0) = \mathbf{x}_0$ es fijo y el estado final $\mathbf{x}(t_1)$ es libre. Además se considera el problema básico en la forma de *Lagrange* (2.3.40).

Las trayectorias óptimas calculadas a partir de la aplicación del principio de máximo dependen de la naturaleza de las condiciones que se especifican a los límites de los intervalos de tiempo. Siendo importante la variedad de estas condiciones y, como se menciona en la sección anterior se describen diferentes problemas impuestos por las condiciones de transversalidad. Lo cual implica que

$$\lambda_i^*(t_1) = \frac{\partial}{\partial x_i} S[\mathbf{x}^*(t_1)] = 0,$$

para cada $i = 1, 2, \dots, n$. Por lo que se plantea el siguiente problema:

$$\begin{aligned} \text{máx } J(\mathbf{u}) &= \int_{t_0}^{t_1} F[t, \mathbf{x}, \mathbf{u}] dt, \\ \text{sujeto a : } \dot{\mathbf{x}} &= \mathbf{f}(t, \mathbf{x}, \mathbf{u}), \\ \text{con : } \mathbf{x}(t_0) &= \mathbf{x}_0, \mathbf{x}(t_1) \text{ libre,} \\ \mathbf{u}(t) &\in \mathcal{U}. \end{aligned} \tag{2.3.42}$$

En esta situación el principio del máximo se enuncia de la siguiente manera.

Teorema 2.3.5 (Principio del máximo de Pontryagin para el problema (2.3.42)). *Sea $\mathbf{u}^*(t)$ la trayectoria óptima de control, continua por pedazos, y $\mathbf{x}^*(t)$ la trayectoria de estado óptima asociada, definidas en el intervalo $[t_0, t_1]$. Entonces existe una función vectorial $\boldsymbol{\lambda}^*(t)$ continua que posee primeras derivadas continuas por pedazos, tal que para cada $t \in [t_0, t_1]$ verifica:*

1.

$$\dot{\lambda}_i^*(t) = -\frac{\partial}{\partial x_i} H(t, \mathbf{x}^*(t), \mathbf{u}^*(t), \boldsymbol{\lambda}^*(t)),$$

$$\lambda_i^*(t_1) = 0$$

para cada $i = 1, \dots, n$ en todos los puntos de continuidad de $\mathbf{u}^*(t)$.

2. Para el caso de maximizar, se cumple que

$$H(t, \mathbf{x}^*(t), \mathbf{u}^*(t), \boldsymbol{\lambda}^*(t)) \geq H(t, \mathbf{x}^*(t), \mathbf{u}(t), \boldsymbol{\lambda}^*(t)),$$

para cada $\mathbf{u}(t) \in \mathcal{U}$.

3. La trayectoria de estado óptimo $\mathbf{x}^*(t)$ es solución del sistema con condición inicial \mathbf{x}_0^*

$$\dot{\mathbf{x}}(t) = \mathbf{f}_i^*(t, \mathbf{x}^*(t), \mathbf{u}^*(t)),$$

para cada $i = 1, \dots, n$ en todos los puntos de continuidad de $\mathbf{u}^*(t)$.

En la demostración del principio del máximo de Pontryagin se deduce que

$$\frac{\partial}{\partial u} H(t, \mathbf{x}^*, \mathbf{u}^*, \boldsymbol{\lambda}^*) = 0$$

es una condición necesaria para que

$$H(t, \mathbf{x}^*(t), \mathbf{u}^*(t), \boldsymbol{\lambda}^*(t)) \geq H(t, \mathbf{x}^*(t), \mathbf{u}(t), \boldsymbol{\lambda}^*(t)), \quad \forall \mathbf{u}(t) \in \mathbb{R}^m.$$

Es decir, el Hamiltoniano tiene un punto crítico, en la variable $\mathbf{u}(t)$, en $\mathbf{u}^*(t)$ para cada t .

Si el Hamiltoniano es lineal en la variable de control $\mathbf{u}(t)$, puede ser difícil resolver $\mathbf{u}^*(t)$ a partir del sistema de adjunto. De poder resolverse $\mathbf{u}^*(t)$ a partir de dicho sistema, quedan dos incógnitas $\mathbf{x}^*(t)$ y $\boldsymbol{\lambda}(t)$ que satisfacen dos ecuaciones diferenciales con dos condiciones de frontera. Se resuelve el sistema de ecuaciones diferenciales para el estado óptimo, el sistema adjunto y luego obtener el control óptimo $\mathbf{u}^*(t)$.

Una aplicación a este tipo de problema de control óptimo se plantea enseguida.

Considerar el problema:

$$\begin{aligned} \min_u J(u) &= \int_0^T AI(t) + u(t)^2 dt, \\ \text{sueto a : } \dot{S}(t) &= bN(t) - dS(t) - cS(t)I(t) - u(t)S(t), \\ \dot{E}(t) &= cS(t)I(t) - (e + d)E(t), \\ \dot{I}(t) &= eE(t) - (g + a + d)I(t), \\ \dot{N}(t) &= (b - d)N(t) - aI(t), \end{aligned}$$

$$\begin{aligned} \text{con : } S(0) &= S_0 \geq 0, \\ E(0) &= E_0 \geq 0, \\ I(0) &= I_0 \geq 0, \\ N(0) &= N_0 \geq 0, \\ u(t) &\in [0, 0.9] = \mathcal{U}. \end{aligned}$$

Más adelante se resuelve este problema usando el caso particular del principio del máximo que se ha explicado.

3 Desarrollo

3.1. Modelo SEIR

Considerar una población en la cual un número reducido de sus miembros padece alguna enfermedad infecciosa que se puede transmitir a otros de la misma población [2]. Para la construcción del modelo se necesita dividir a nuestra población en cuatro subclases, que son S , E , I , R . Los individuos susceptibles en S son aquellos que pueden contraer la enfermedad. Aquellos que ya están infectados pero que aún no pueden contagiar la enfermedad se encuentran en el grupo de los expuestos, E . Los individuos infectados con la enfermedad que son capaces de transmitirla a los susceptibles pertenecen al grupo I y aquellos quienes se han recuperado (o son inmunes) de la enfermedad y ya no son susceptibles están en el grupo R .

La cantidad de individuos en cada grupo varía con el tiempo dado que los nacimientos y las muertes afectan a la cantidad de individuos que pertenecen a cada grupo. Asumimos que la enfermedad es transmitida de un individuo a otro de manera directa o indirecta. Dicho de otra manera, se trata de una enfermedad en la cual la proximidad a un organismo ya infectado supone un aumento significativo del riesgo de resultar también infectado.

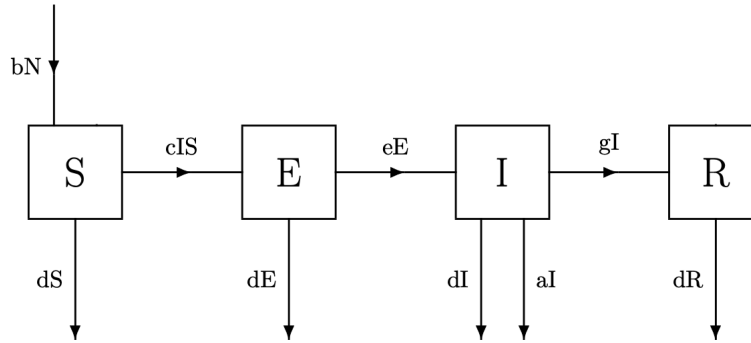
Para la formulación de nuestro modelo (ver [13]) se define lo siguiente:

- $S(t)$ = número de individuos que pertenecen al grupo de susceptibles al tiempo t .
- $E(t)$ = número de individuos que pertenecen al grupo de expuestos al tiempo t .
- $I(t)$ = número de individuos que pertenecen al grupo de infectados al tiempo t .
- $R(t)$ = número de individuos que pertenecen al grupo de recuperados al tiempo t .
- a = tasa de mortalidad por la enfermedad.
- b = tasa de natalidad de la población.
- c = coeficiente de incidencia de transmisión.
- d = tasa de mortalidad natural.
- e = tasa a la que los expuestos pasan a ser infectados.
- g = tasa a la que los infectados pasan a ser recuperados.

Sea $N(t)$ el número total de personas en la población, de modo que $N(t) = S(t) + E(t) + I(t) + R(t)$.

Para efectos de la modelación los parámetros a, b, c, d, e, g son constantes durante el tiempo que dure la medición. La tasa de transición entre susceptible e infectado es proporcional al número de susceptibles y al número de infectados, es decir, el número promedio de individuos susceptibles que pasan a ser infectados por unidad de tiempo es $cS(t)I(t)$.

El siguiente diagrama compartimental representa la transmisión de una enfermedad entre los individuos con un modelo $SEIR$:



Describiendo la razón de cambio de cada grupo como un sistema de ecuaciones diferenciales se tiene lo siguiente:

$$\begin{aligned}
 \dot{S}(t) &= bN(t) - dS(t) - cS(t)I(t), \\
 \dot{E}(t) &= cS(t)I(t) - (e + d)E(t), \\
 \dot{I}(t) &= eE(t) - (g + a + d)I(t), \\
 \dot{R}(t) &= gI(t) - dR(t),
 \end{aligned}$$

con condiciones iniciales

$$\begin{aligned}
 S(0) &= S_0 \geq 0, \\
 E(0) &= E_0 \geq 0, \\
 I(0) &= I_0 \geq 0, \\
 R(0) &= R_0 \geq 0.
 \end{aligned}$$

Observación 3.1.1. Como $N(t) = S(t) + E(t) + I(t) + R(t)$ y derivando con respecto a t , entonces

$$\dot{N}(t) = \dot{S}(t) + \dot{E}(t) + \dot{I}(t) + \dot{R}(t).$$

Sustituyendo $\dot{S}(t)$, $\dot{E}(t)$, $\dot{I}(t)$, $\dot{R}(t)$ en la ecuación anterior se obtiene

$$\begin{aligned}
 \dot{N}(t) &= bN(t) - dS(t) - cS(t)I(t) + cS(t)I(t) - (e + d)E(t) + \\
 &\quad + eE(t) - (g + a + d)I(t) + gI(t) - dR(t) \\
 &= bN(t) - dS(t) - dE(t) - aI(t) - dI(t) - dR(t) \\
 &= bN(t) - d(S(t) + E(t) + I(t) + R(t)) - aI(t) \\
 &= bN(t) - dN(t) - aI(t) \\
 &= (b - d)N(t) - aI(t) \\
 &\Rightarrow \dot{N}(t) = (b - d)N(t) - aI(t).
 \end{aligned}$$

Notar que la función $R(t)$ solo aparece en la ecuación $\dot{R}(t) = gI(t) - dR(t)$ y se puede obtener de la relación $R(t) = N(t) - S(t) - E(t) - I(t)$. Dado lo anterior, se considera el siguiente sistema de ecuaciones diferenciales:

$$\dot{S}(t) = bN(t) - dS(t) - cS(t)I(t), \quad (3.1.1)$$

$$\dot{E}(t) = cS(t)I(t) - (e + d)E(t), \quad (3.1.2)$$

$$\dot{I}(t) = eE(t) - (g + a + d)I(t), \quad (3.1.3)$$

$$\dot{N}(t) = (b - d)N(t) - aI(t), \quad (3.1.4)$$

con condiciones iniciales

$$S(0) = S_0 \geq 0,$$

$$E(0) = E_0 \geq 0,$$

$$I(0) = I_0 \geq 0,$$

$$N(0) = N_0 \geq 0.$$

En los modelos de epidemias es importante el estudio de la evolución de alguna enfermedad, así que, para conocer el desarrollo y saber si aminora o aumenta la infección. El número básico de reproducción \mathcal{R}_0 , es el número esperado de casos secundarios que se producen en una población totalmente vulnerable, que genera un primer individuo infectado. En el trabajo de David Greenhalgh [16], se estudian los puntos de equilibrio, la estabilidad y se encuentra una formulación para obtener el parámetro \mathcal{R}_0 , además, concluye que si $\mathcal{R}_0 < 1$ entonces la enfermedad tiende a desaparecer o mantenerse estable. Si $\mathcal{R}_0 > 1$ entonces la enfermedad se expande hasta producir una epidemia.

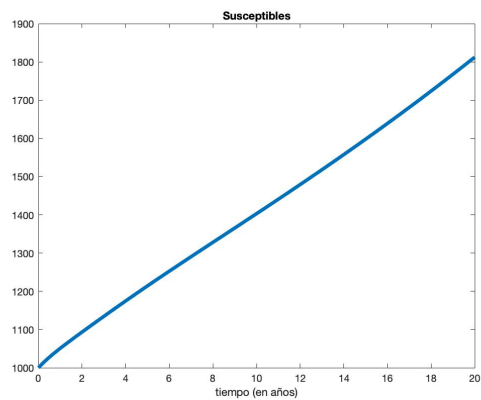
Para que la solución del sistema no se convierta en epidemia, es decir $\mathcal{R}_0 < 1$, se utilizan los siguientes parámetros (consultados en [13]):

$$a = 0.2; \quad b = 0.525; \quad c = 0.001; \quad d = 0.5; \quad e = 0.5; \quad g = 0.1,$$

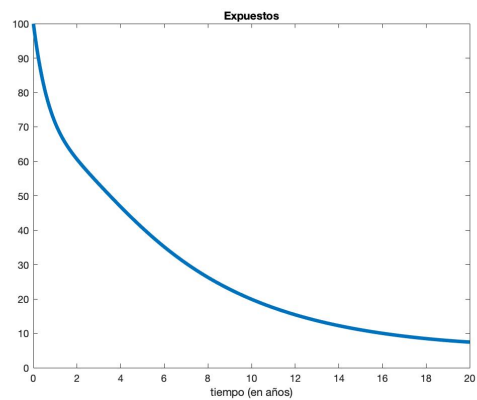
y se emplea el Método de Runge-Kutta usando MATLAB (para mayor detalle consultar [17]).

En esta simulación se consideraron las condiciones iniciales $S(0) = 1000$, $E(0) = 100$, $I(0) = 50$ y $R(0) = 15$, con los valores de las constantes a, b, c, d, e y g establecidos anteriormente. Se puede observar que la cantidad de individuos susceptibles va aumentando debido a que en la ecuación $\dot{S}(t)$ esta involucrado un término que depende de la tasa de natalidad y la población total. La gráfica de los individuos infectados muestra que en un lapso de aproximadamente 2 años (el tiempo se mide en años) se presenta un pequeño incremento de personas infectadas con la presencia del virus, que posteriormente disminuye. Para la población de expuestos se muestra que los individuos que portan el virus, pero no manifiestan síntomas, pasan a ser infectados en algún momento y la población de expuestos disminuirá. Por último, conforme evoluciona la enfermedad se tiene que tanto individuos expuestos como infectados pasan a la población de recuperados, y cuando disminuye la cantidad de individuos con la enfermedad también disminuyen los individuos que se tendrán que recuperar.

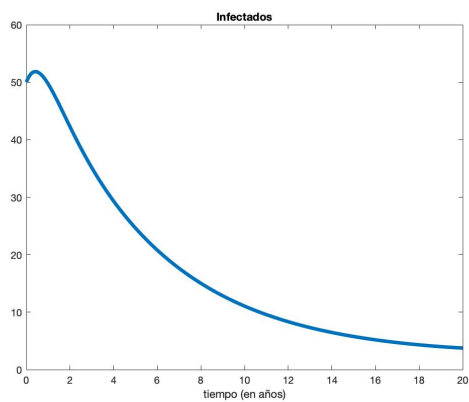
A continuación se muestran las gráficas de los comportamientos obtenidos de cada población en el modelo SEIR.



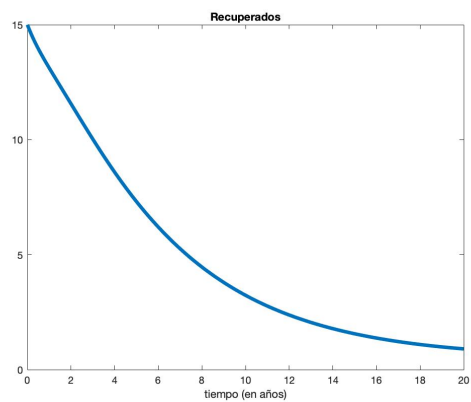
(a) Susceptibles



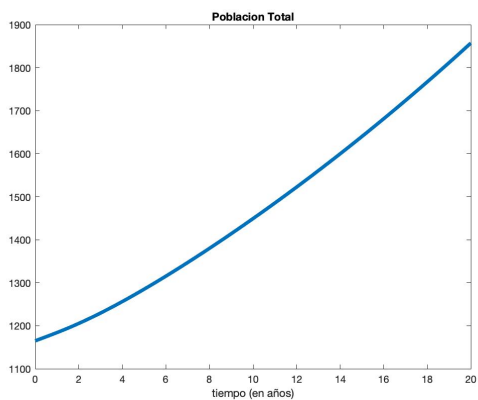
(b) Expuestos



(c) Infectados



(d) Recuperados



(e) Población Total

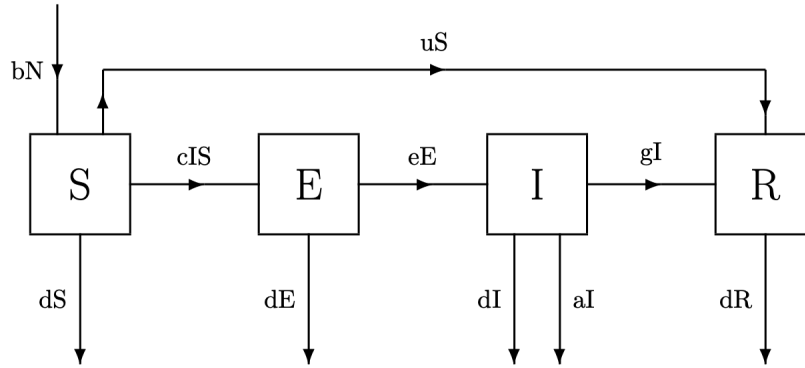
Figura 3.1: Modelo SEIR.

3.2. Modelo SEIR con control

Estudiaremos a continuación un problema de control óptimo a partir del modelo SEIR. En este caso se quiere controlar la expansión de la enfermedad infecciosa mediante el uso de una vacuna.

Como ya se mencionó, si $\mathcal{R}_0 > 1$ entonces se produce una epidemia y para evitarlo se propone controlar la enfermedad a través de la vacunación, esto es lo que se estudia en este trabajo. Suponer que la vacuna es efectiva, por lo que todo individuo susceptible que sea vacunado se vuelve inmune. Sea $u(t)$ el porcentaje de individuos susceptibles vacunados por unidad de tiempo.

El siguiente diagrama compartimental simboliza el modelo SEIR con el uso de una vacuna.



El sistema de ecuaciones diferenciales que lo representa es:

$$\begin{aligned}
 \dot{S}(t) &= bN(t) - dS(t) - cS(t)I(t) - u(t)S(t), \\
 \dot{E}(t) &= cS(t)I(t) - (e + d)E(t), \\
 \dot{I}(t) &= eE(t) - (g + a + d)I(t), \\
 \dot{R}(t) &= gI(t) - dR(t) + u(t)S(t).
 \end{aligned}$$

Además, la observación (3.1.1) se mantiene para este modelo y se obtiene de manera similar que:

$$\begin{aligned}
 N(t) &= S(t) + E(t) + I(t) + R(t), \\
 \dot{N}(t) &= (b - d)N(t) - aI(t),
 \end{aligned}$$

y

$$R(t) = N(t) - S(t) - E(t) - I(t).$$

Así que, como antes considerar el siguiente sistema de ecuaciones diferenciales:

$$\begin{aligned}
 \dot{S}(t) &= bN(t) - dS(t) - cS(t)I(t) - u(t)S(t), \\
 \dot{E}(t) &= cS(t)I(t) - (e + d)E(t), \\
 \dot{I}(t) &= eE(t) - (g + a + d)I(t), \\
 \dot{N}(t) &= (b - d)N(t) - aI(t),
 \end{aligned} \tag{3.2.5}$$

con condiciones iniciales

$$\begin{aligned} S(0) &= S_0 \geq 0, \\ E(0) &= E_0 \geq 0, \\ I(0) &= I_0 \geq 0, \\ N(0) &= N_0 \geq 0. \end{aligned}$$

Observación 3.2.1. *En la modelación, la función $u(t) \in [0, 0.9]$ actúa como la variable de control de este sistema. Si $u(t) = 0$, no se aplicó ninguna vacuna y se recupera el modelo SEIR ya analizado. Si $u(t) = 1$, indica que toda la población susceptible está vacunada.*

3.2.1. Estabilidad y controlabilidad

Como se ha mencionado en la sección anterior, la controlabilidad y la estabilidad son propiedades que describen el comportamiento de cierto proceso en un intervalo de tiempo y que están relacionadas entre si. Se estudiarán dichas propiedades en el modelo SEIR con vacunación.

Notar que el sistema (3.2.5) es no lineal con respecto a las variables de estado, así que para el estudio de la controlabilidad se emplea el teorema (2.3.3), el cual utiliza un proceso de linealización para analizar si el sistema es controlable en una vecindad de un punto de equilibrio.

En este trabajo se consideran dos puntos de equilibrio, los cuales tienen interpretaciones en términos de la evolución para cierta enfermedad. Sin embargo, en el trabajo de David Greenhalgh [16], se estudian a detalle los puntos de equilibrio que tiene el modelo SEIR.

Considerar los puntos de equilibrio:

$$\mathbf{x}_{\text{triv}} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad (3.2.6)$$

y

$$\mathbf{x}_{\text{lib}} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \quad (3.2.7)$$

donde \mathbf{x}_{triv} es el punto de equilibrio trivial y \mathbf{x}_{lib} es el punto libre de enfermedad.

Para la linealización del sistema (3.2.5) se necesitan las siguientes matrices

$$A = D_{\mathbf{x}}\mathbf{f}(\bar{\mathbf{x}}, 0) \text{ y } B = D_{\mathbf{u}}\mathbf{f}(\bar{\mathbf{x}}, 0), \quad (3.2.8)$$

en este caso

$$\mathbf{f}(\bar{\mathbf{x}}, 0) = \begin{pmatrix} bN - dS - cSI - uS \\ cSI - (e + d)E \\ eE - (g + a + d)I \\ (b - d)N - aI \end{pmatrix} = \begin{pmatrix} bN - dS - cSI \\ cSI - (e + d)E \\ eE - (g + a + d)I \\ (b - d)N - aI \end{pmatrix} + \begin{pmatrix} -uS \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

con $D_{\mathbf{x}}$ y $D_{\mathbf{u}}$ las matrices cuyas entradas son las derivadas parciales con respecto a las variables de estado y de control respectivamente.

Si se considera el punto de equilibrio \mathbf{x}_{triv} , entonces

$$\begin{aligned} A = D_{\mathbf{x}}\mathbf{f}(\mathbf{x}_{\text{triv}}, 0) &= \left[\begin{array}{cccc} -d - cI & 0 & -cS & b \\ cI & -(e + d) & cS & 0 \\ 0 & e & -(g + a + d) & 0 \\ 0 & 0 & -a & b - d \end{array} \right] \bigg|_{\mathbf{x}_{\text{triv}}}, \\ &= \begin{bmatrix} -d & 0 & 0 & b \\ 0 & -(e + d) & 0 & 0 \\ 0 & e & -(g + a + d) & 0 \\ 0 & 0 & -a & b - d \end{bmatrix} \end{aligned} \quad (3.2.9)$$

y

$$B = D_{\mathbf{u}}\mathbf{f}(\mathbf{x}_{\text{triv}}, 0) = \left[\begin{array}{c} -S \\ 0 \\ 0 \\ 0 \end{array} \right] \bigg|_{\mathbf{x}_{\text{triv}}} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}. \quad (3.2.10)$$

Ahora, se definen

$$\dot{\mathbf{x}}(t) = \begin{pmatrix} \dot{S}(t) \\ \dot{E}(t) \\ \dot{I}(t) \\ \dot{N}(t) \end{pmatrix}, \quad \mathbf{x}(t) = \begin{pmatrix} S(t) \\ E(t) \\ I(t) \\ N(t) \end{pmatrix}, \quad \mathbf{u}(t) = \begin{pmatrix} -u(t) \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

entonces la linealización del sistema del modelo SEIR con control (3.2.5) es

$$\begin{aligned} \dot{\mathbf{x}}(t) &= A \cdot \mathbf{x}(t) + B \cdot \mathbf{u}(t) \\ &= \begin{bmatrix} -d & 0 & 0 & b \\ 0 & -(e + d) & 0 & 0 \\ 0 & e & -(g + a + d) & 0 \\ 0 & 0 & -a & b - d \end{bmatrix} \cdot \mathbf{x}(t) + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \cdot \mathbf{u}(t) \\ &= \begin{bmatrix} -d & 0 & 0 & b \\ 0 & -(e + d) & 0 & 0 \\ 0 & e & -(g + a + d) & 0 \\ 0 & 0 & -a & b - d \end{bmatrix} \cdot \mathbf{x}(t). \end{aligned}$$

Por lo tanto

$$\dot{\mathbf{x}}(t) = \begin{bmatrix} -d & 0 & 0 & b \\ 0 & -(e+d) & 0 & 0 \\ 0 & e & -(g+a+d) & 0 \\ 0 & 0 & -a & b-d \end{bmatrix} \cdot \mathbf{x}(t). \quad (3.2.11)$$

Una vez linealizado el sistema (3.2.5) se desea verificar que se cumplan las hipótesis del criterio de controlabilidad (2.3.2). En este criterio se hace la suposición de que la parte real de todos los valores propios λ de la matriz asociada al sistema son negativos, esto quiere decir que se asume que el sistema es asintóticamente estable en un punto de equilibrio dado.

A continuación se emplea el método sistemático para el estudio de la estabilidad. Considerar el sistema (3.2.11) y la matriz $A = D_{\mathbf{x}}\mathbf{f}(\mathbf{x}_{\text{triv}}, 0)$. De hecho, la matriz A es la *matriz jacobiana* del sistema evaluada en el punto de equilibrio \mathbf{x}_{triv} , así que se calculan los valores propios de esta matriz

$$\begin{aligned} p(\lambda) = \det(A - \lambda I) &= \left| \begin{bmatrix} -d & 0 & 0 & b \\ 0 & -(e+d) & 0 & 0 \\ 0 & e & -(g+a+d) & 0 \\ 0 & 0 & -a & b-d \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \right| \\ &= \left| \begin{bmatrix} -d-\lambda & 0 & 0 & b \\ 0 & -(e+d)-\lambda & 0 & 0 \\ 0 & e & -(g+a+d)-\lambda & 0 \\ 0 & 0 & -a & b-d-\lambda \end{bmatrix} \right| \\ &= (-d-\lambda)(-e-d-\lambda)(-g-a-d-\lambda)(b-d-\lambda) = 0. \end{aligned}$$

Por lo que los valores propios de la matriz A son:

$$\lambda_1 = -d, \lambda_2 = -(e+d), \lambda_3 = -(g+a+d), \lambda_4 = b-d.$$

Observación 3.2.2. *Para que el sistema (3.2.11) sea asintóticamente estable se necesita que la parte real de $\lambda_1, \lambda_2, \lambda_3$ y λ_4 sean negativos. En este caso los valores a, b, d, e, g son constantes positivas, entonces*

$$\lambda_1, \lambda_2, \lambda_3 < 0.$$

Para $\lambda_4 = b-d$ se observa lo siguiente

- Si $b > d$ entonces $\lambda_4 = b-d > 0$,
- Si $b < d$ entonces $\lambda_4 = b-d < 0$,

donde b representa la tasa de natalidad de la población y d es la tasa de mortalidad de la población. Así pues, si la tasa de natalidad es menor que la tasa de mortalidad, entonces se tiene estabilidad asintótica en el punto \mathbf{x}_{triv} . En general, no se puede asegurar estabilidad asintótica.

Por otro lado, se construye la matriz de controlabilidad $G(A, B) := [B, AB, A^2B, \dots, A^{n-1}B]$ y como el dominio en el que se está trabajando es \mathbb{R}^4 se toma $n = 4$.

Sean

$$A = \begin{bmatrix} -d & 0 & 0 & b \\ 0 & -(e+d) & 0 & 0 \\ 0 & e & -(g+a+d) & 0 \\ 0 & 0 & -a & b-d \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

Se necesitan calcular A^2, A^3, AB, A^2B, A^3B , sin embargo todos elementos de B son cero y esto simplifica los cálculos. Entonces

$$\begin{aligned} AB &= \begin{bmatrix} -d & 0 & 0 & b \\ 0 & -(e+d) & 0 & 0 \\ 0 & e & -(g+a+d) & 0 \\ 0 & 0 & -a & b-d \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \\ A^2B &= \begin{bmatrix} -d & 0 & 0 & b \\ 0 & -(e+d) & 0 & 0 \\ 0 & e & -(g+a+d) & 0 \\ 0 & 0 & -a & b-d \end{bmatrix}^2 \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \\ A^3B &= \begin{bmatrix} -d & 0 & 0 & b \\ 0 & -(e+d) & 0 & 0 \\ 0 & e & -(g+a+d) & 0 \\ 0 & 0 & -a & b-d \end{bmatrix}^3 \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}. \end{aligned}$$

Así que la matriz de controlabilidad es:

$$G(A, B) = [B, AB, A^2B, A^3B] = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

y $\text{rank } G(A, B) = 0 \neq 4$.

Entonces, como no se puede garantizar la estabilidad asintótica y $\text{rank } G(A, B) \neq 4$ las hipótesis del criterio de controlabilidad (2.3.2) no se cumplen. Por lo tanto el sistema de control linealizado (3.2.11) no es controlable.

El punto de equilibrio \mathbf{x}_{lib} (3.2.7) conocido como *estado libre de la enfermedad* se interpreta como aquel donde no hay individuos infectados por alguna enfermedad, lo cual implica que los individuos no presentan síntomas, no se requiere recuperación y por lo tanto la población total es susceptible. En términos del modelo SEIR esto significa que $S(t) = 1, E(t) = 0, I(t) = 0, R(t) = 0$ y $N(t) = S(t) + E(t) + I(t) + R(t) = 1$.

Ahora se desea saber la estabilidad y controlabilidad de la linealización del sistema (3.2.5) alrededor del punto \mathbf{x}_{lib} , se emplea el proceso sistemático usado para el punto de equilibrio \mathbf{x}_{triv} .

En este caso, las matrices con derivadas parciales son

$$\begin{aligned}
 C = D_{\mathbf{x}}\mathbf{f}(\mathbf{x}_{\text{lib}}, 0) &= \left[\begin{array}{cccc} -d - cI & 0 & -cS & b \\ cI & -(e + d) & cS & 0 \\ 0 & e & -(g + a + d) & 0 \\ 0 & 0 & -a & b - d \end{array} \right] \bigg|_{\mathbf{x}_{\text{lib}}} \\
 &= \left[\begin{array}{cccc} -d & 0 & -c & b \\ 0 & -(e + d) & c & 0 \\ 0 & e & -(g + a + d) & 0 \\ 0 & 0 & -a & b - d \end{array} \right],
 \end{aligned} \tag{3.2.12}$$

y

$$D = D_{\mathbf{u}}\mathbf{f}(\mathbf{x}_{\text{lib}}, 0) = \left[\begin{array}{c} -S \\ 0 \\ 0 \\ 0 \end{array} \right] \bigg|_{\mathbf{x}_{\text{lib}}} = \left[\begin{array}{c} -1 \\ 0 \\ 0 \\ 0 \end{array} \right]. \tag{3.2.13}$$

La linealización del sistema (3.2.5) es

$$\begin{aligned}
 \dot{\mathbf{x}}(t) &= C \cdot \mathbf{x}(t) + D \cdot \mathbf{u}(t) \\
 &= \left[\begin{array}{cccc} -d & 0 & -c & b \\ 0 & -(e + d) & c & 0 \\ 0 & e & -(g + a + d) & 0 \\ 0 & 0 & -a & b - d \end{array} \right] \cdot \mathbf{x}(t) + \left[\begin{array}{c} -1 \\ 0 \\ 0 \\ 0 \end{array} \right] \cdot \mathbf{u}(t).
 \end{aligned} \tag{3.2.14}$$

Se calculan los valores propios de

$$C = \left[\begin{array}{cccc} -d & 0 & -c & b \\ 0 & -(e + d) & c & 0 \\ 0 & e & -(g + a + d) & 0 \\ 0 & 0 & -a & b - d \end{array} \right],$$

$$\begin{aligned}
 p(\lambda) = \det(C - \lambda I) &= \left| \left[\begin{array}{cccc} -d & 0 & -c & b \\ 0 & -(e + d) & c & 0 \\ 0 & e & -(g + a + d) & 0 \\ 0 & 0 & -a & b - d \end{array} \right] - \lambda \left[\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array} \right] \right| \\
 &= \left| \left[\begin{array}{cccc} -d - \lambda & 0 & -c & b \\ 0 & -(e + d) - \lambda & c & 0 \\ 0 & e & -(g + a + d) - \lambda & 0 \\ 0 & 0 & -a & b - d - \lambda \end{array} \right] \right| \\
 &= (-d - \lambda)(b - d - \lambda)[(-e - d - \lambda)(-g - a - d - \lambda) - ec] \\
 &= (-d - \lambda)(b - d - \lambda)[\lambda^2 + \lambda(e + g + a + 2d) + d^2 + eg + ea + ed + dg + da - ec] \\
 &= 0.
 \end{aligned}$$

Entonces los valores propios son los siguientes

$$\begin{aligned}\lambda_1 &= -d, \\ \lambda_2 &= b - d, \\ \lambda_3 &= \frac{1}{2} \left(-\sqrt{a^2 + e^2 + g^2 + 2ag + 4ce - 2ae - 2eg} - a - 2d - e - g \right), \\ \lambda_4 &= \frac{1}{2} \left(\sqrt{a^2 + e^2 + g^2 + 2ag + 4ce - 2ae - 2eg} - a - 2d - e - g \right),\end{aligned}$$

Observaciones 3.2.1. Los valores a, b, c, d, e, g son constantes positivas, entonces $\lambda_1 = -d < 0$ y de las observaciones en (3.2.2) se analiza cuándo puede ocurrir que $\lambda_2 = b - d < 0$. Por último $\lambda_3 = \frac{1}{2} \left(-\sqrt{a^2 + e^2 + g^2 + 2ag + 4ce - 2ae - 2eg} - a - 2d - e - g \right) < 0$ y para λ_4 notar lo siguiente:

Los parámetros e y g representan la tasa a que los expuestos pasan a ser infectados y la tasa a la que los infectados pasan a ser recuperados respectivamente, entonces resulta razonable tomar $e = 0 = g$ ya que se hizo la suposición de que se está en un estado libre de la enfermedad. Por lo que

$$\lambda_4 = \frac{1}{2} \left(\sqrt{a^2} - a - 2d \right) = \frac{1}{2} (|a| - a - 2d) = \frac{1}{2} (a - a - 2d) = \frac{1}{2} (-2d) < -d < 0.$$

Bajo estas consideraciones es posible tener estabilidad asintótica en el punto \mathbf{x}_{lib} .

Ahora se construye la matriz de controlabilidad $G(C, D) = [D, CD, C^2D, C^3D]$ a partir de las matrices

$$C = \begin{bmatrix} -d & 0 & -c & b \\ 0 & -(e+d) & c & 0 \\ 0 & e & -(g+a+d) & 0 \\ 0 & 0 & -a & b-d \end{bmatrix}, \quad D = \begin{bmatrix} -1 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

Sean

$$\begin{aligned}C^2 &= \begin{bmatrix} -d & 0 & -c & b \\ 0 & -(e+d) & c & 0 \\ 0 & e & -(g+a+d) & 0 \\ 0 & 0 & -a & b-d \end{bmatrix}^2 \\ &= \begin{bmatrix} d^2 & -ce & a(c-b) + c(2d+g) & b(b-2d) \\ 0 & e(c+e) + d^2 + 2de & -c(a+2d+e+g) & 0 \\ 0 & -e(a+2d+e+g) & (a+d+g)^2 + ce & 0 \\ 0 & -ae & a(a-b+2d+g) & (b-d)^2 \end{bmatrix}, \\ C^3 &= \begin{bmatrix} -d & 0 & -c & b \\ 0 & -(e+d) & c & 0 \\ 0 & e & -(g+a+d) & 0 \\ 0 & 0 & -a & b-d \end{bmatrix}^3 = \\ &= \begin{bmatrix} -d^3 & e(a(c-b)+c(3d+e+g)) & a^2(b-c)-a(b^2-b(3d+g)+c(3d+2g))-c(ce+3d^2+3dg+g^2) & b(b^2-3bd+3d^2) \\ 0 & -e(ac+2ce+cg+e^2)-3de(c+e)-d^3-3d^2e & c(a^2+a(3d+e+2g)+ce+3d^2+3d(e+g)+e^2+eg+g^2) & 0 \\ 0 & e(a^2+a(3d+e+2g)+ce+3d^2+3d(e+g)+e^2+eg+g^2) & -(a+d+g)(a^2+2a(d+g)+ce+d^2+2dg+g^2)-ce(a+2d+e+g) & 0 \\ 0 & ae(a-b+3d+e+g) & -a(a^2+a(-b+3d+2g)+b^2-b(3d+g)+ce+3d^2+3dg+g^2) & (b-d)^3 \end{bmatrix},\end{aligned}$$

$$CD = \begin{bmatrix} -d & 0 & -c & b \\ 0 & -(e+d) & c & 0 \\ 0 & e & -(g+a+d) & 0 \\ 0 & 0 & -a & b-d \end{bmatrix} \begin{bmatrix} -1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} d \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

$$C^2D = \begin{bmatrix} -d & 0 & -c & b \\ 0 & -(e+d) & c & 0 \\ 0 & e & -(g+a+d) & 0 \\ 0 & 0 & -a & b-d \end{bmatrix}^2 \begin{bmatrix} -1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -d^2 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

$$C^3D = \begin{bmatrix} -d & 0 & -c & b \\ 0 & -(e+d) & c & 0 \\ 0 & e & -(g+a+d) & 0 \\ 0 & 0 & -a & b-d \end{bmatrix}^3 \begin{bmatrix} -1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} d^3 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

Por lo que

$$G(C, D) = [D, CD, C^2D, C^3D] = \begin{bmatrix} -1 & d & -d^2 & d^3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

y

$$\text{rank } G(C, D) = 1 \neq 4.$$

En consecuencia no se cumplen las hipótesis del teorema (2.3.2) y por lo tanto el sistema (3.2.14) no es controlable.

Esto quiere decir que no existe una variable de control $u(t)$ que dirija la dinámica del modelo SEIR con vacunación a los puntos \mathbf{x}_{triv} y \mathbf{x}_{lib} en un tiempo finito.

Las matrices B y D tienen solo una de sus entradas distintas de cero debido a que la variable de control $u(t)$ aparece de manera única en la ecuación $\dot{S}(t)$ en el modelo y no esta involucrada en el resto de las ecuaciones, dicho de otra manera, la aplicación del término de control “vacunación” se considera solo a la población susceptible.

Además, no hay controlabilidad porque la población total crece de manera exponencial. No es difícil ver que si se resuelve la ecuación (en el supuesto que no hay infectados)

$$\dot{N}(t) = (b-d)N(t) - aI(t) = (b-d)N(t),$$

se obtiene por solución

$$N(t) = N_0 e^{(b-d)t}, \text{ para cierta condición inicial } N_0.$$

Para una modelación más realista se sugiere un problema donde se modifica la ecuación $\dot{N}(t)$ en el modelo SEIR añadiendo un término de limitación de recursos o capacidad de carga. Si se considera la ecuación logística queda

$$\dot{N}(t) = (b - d)N - BN^2 - aI,$$

donde B es un parámetro de auto limitación.

Por lo tanto, se sugiere el modelo

$$\begin{aligned}\dot{S}(t) &= bN(t) - dS(t) - cS(t)I(t) - u(t)S(t), \\ \dot{E}(t) &= cS(t)I(t) - (e + d)E(t), \\ \dot{I}(t) &= eE(t) - (g + a + d)I(t), \\ \dot{N}(t) &= (b - d)N - BN^2 - aI(t),\end{aligned}\tag{3.2.15}$$

con condiciones iniciales

$$\begin{aligned}S(0) &= S_0 \geq 0, \\ E(0) &= E_0 \geq 0, \\ I(0) &= I_0 \geq 0, \\ N(0) &= N_0 \geq 0.\end{aligned}$$

y como trabajo adicional proponer algún otro parámetro de control como distanciamiento social junto con la vacunación.

3.2.2. Aplicación del principio del máximo de Pontryagin

Utilizando el sistema de ecuaciones diferenciales (3.2.5), Neilan y Lenhart en [13] proponen un problema de control óptimo para determinar la estrategia de vacunación en un intervalo de tiempo fijo $[0, T]$. Entonces, la idea es determinar la política de vacunación $u(t)$ para minimizar el funcional

$$J(u) = \int_0^T AI(t) + u(t)^2 dt.$$

Puesto que $0 \leq u \leq 0.9$, el costo de vacunación se mantiene bajo. En cuanto al término de costo $AI(t)$ se debe tener cuidado al elegir el valor de A . Si A es igual al parámetro a (tasa de mortalidad por la enfermedad) entonces la idea es minimizar el número de muertos infectados por la enfermedad. Elegir A mucho más pequeño que a significa que se da poca importancia a los individuos que mueren como consecuencia de la enfermedad. Una A más grande que a significa que la carga de muertos por la enfermedad es más importante que el costo de la vacunación.

El objetivo es minimizar el número de personas infectadas y el costo total de la vacuna durante un período de tiempo fijo, se utilizan técnicas de control óptimo para encontrar un programa de vacunación para una enfermedad epidémica.

El problema de control óptimo es el siguiente:

$$\begin{aligned} \min_u J(u) &= \int_0^T AI(t) + u(t)^2 dt, \\ \text{sujeto a : } \dot{S}(t) &= bN(t) - dS(t) - cS(t)I(t) - u(t)S(t), \\ \dot{E}(t) &= cS(t)I(t) - (e + d)E(t), \\ \dot{I}(t) &= eE(t) - (g + a + d)I(t), \\ \dot{N}(t) &= (b - d)N(t) - aI(t), \end{aligned} \quad (3.2.16)$$

$$\begin{aligned} \text{con : } S(0) &= S_0 \geq 0, \quad S(T) \text{ libre}, \\ E(0) &= E_0 \geq 0, \quad E(T) \text{ libre}, \\ I(0) &= I_0 \geq 0, \quad I(T) \text{ libre}, \\ N(0) &= N_0 \geq 0, \quad N(T) \text{ libre}, \\ u(t) &\in [0, 0.9] = \mathcal{U}. \end{aligned}$$

El Lagrangeano asociado al funcional $J(u)$ del problema (3.2.16) es

$$F = F(S(t), E(t), I(t), N(t), u(t)) = AI(t) + u^2(t). \quad (3.2.17)$$

Sean

$$\dot{\mathbf{x}}(t) = \begin{pmatrix} \dot{S}(t) \\ \dot{E}(t) \\ \dot{I}(t) \\ \dot{N}(t) \end{pmatrix}, \quad \mathbf{x}(t) = \begin{pmatrix} S(t) \\ E(t) \\ I(t) \\ N(t) \end{pmatrix}, \quad \mathbf{x}^*(t) = \begin{pmatrix} S^*(t) \\ E^*(t) \\ I^*(t) \\ N^*(t) \end{pmatrix},$$

además

$$\mathbf{f} = \begin{pmatrix} bN(t) - dS(t) - cS(t)I(t) \\ cS(t)I(t) - (e + d)E(t) \\ eE(t) - (g + a + d)I(t) \\ (b - d)N(t) - aI(t) \end{pmatrix}, \quad \mathbf{g} = \begin{pmatrix} -S(t) \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{u}(t) = \begin{pmatrix} u(t) \\ 0 \\ 0 \\ 0 \end{pmatrix} = u(t).$$

Por lo que las restricciones se pueden reescribir como sigue

$$\dot{\mathbf{x}}(t) = \mathbf{f} + \mathbf{u}(t)\mathbf{g}.$$

Aplicando el caso particular del principio del máximo de Pontryagin (2.3.42):

Para nuestro problema el Hamiltoniano se define como:

$$H(t, \mathbf{x}, \mathbf{u}, \boldsymbol{\lambda}) = F + \boldsymbol{\lambda} \cdot (\mathbf{f} + \mathbf{u}(t)\mathbf{g}),$$

donde " \cdot " denota el producto punto entre vectores.

A saber, $\boldsymbol{\lambda}$ es el vector

$$\boldsymbol{\lambda} = \begin{pmatrix} \lambda_S \\ \lambda_E \\ \lambda_I \\ \lambda_N \end{pmatrix}.$$

Entonces el Hamiltoniano queda de la siguiente manera

$$\begin{aligned} H(t, \mathbf{x}, \mathbf{u}, \boldsymbol{\lambda}) &= F + \boldsymbol{\lambda} \cdot (\mathbf{f} + \mathbf{u}(t)\mathbf{g}) \\ &= AI(t) + u^2(t) + \begin{pmatrix} \lambda_S \\ \lambda_E \\ \lambda_I \\ \lambda_N \end{pmatrix} \cdot \left[\begin{pmatrix} bN(t) - dS(t) - cS(t)I(t) \\ cS(t)I(t) - (e + d)E(t) \\ eE(t) - (g + a + d)I(t) \\ (b - d)N(t) - aI(t) \end{pmatrix} + u(t) \begin{pmatrix} -S(t) \\ 0 \\ 0 \\ 0 \end{pmatrix} \right] \\ &= AI(t) + u^2(t) + \lambda_S(t)(bN(t) - dS(t) - cS(t)I(t) - uS(t)) + \\ &\quad + \lambda_E(t)[cS(t)I(t) - (e + d)E(t)] + \lambda_I(t)[eE(t) - (g + a + d)I(t)] + \\ &\quad + \lambda_N(t)[(b - d)N(t) - aI(t)]. \end{aligned}$$

Por lo tanto

$$\begin{aligned} H(t, \mathbf{x}, u, \boldsymbol{\lambda}) &= AI + u^2 + \lambda_S(bN - dS - cSI - uS) + \lambda_E[cSI - (e + d)E] + \\ &\quad + \lambda_I[eE - (g + a + d)I] + \lambda_N[(b - d)N - aI]. \end{aligned} \quad (3.2.18)$$

A partir del Hamiltoniano (3.2.18) se construye las ecuaciones adjuntas como en (1), donde para cada multiplicador de Lagrange $\lambda_S, \lambda_E, \lambda_I, \lambda_N$ se obtiene una ecuación diferencial

$$\dot{\lambda}_S = -\frac{\partial H}{\partial S}, \quad \dot{\lambda}_E = -\frac{\partial H}{\partial E}, \quad \dot{\lambda}_I = -\frac{\partial H}{\partial I}, \quad \dot{\lambda}_N = -\frac{\partial H}{\partial N}.$$

Entonces

$$\begin{aligned} \dot{\lambda}_S &= -[\lambda_S d - \lambda_S cI^* - \lambda_S u + \lambda_E cI^*], \\ \dot{\lambda}_E &= -[-\lambda_E(e + d) + \lambda_I e], \\ \dot{\lambda}_I &= -[A - \lambda_S cS^* + \lambda_E cS^* - \lambda_I(g + a + d) - \lambda_N a], \\ \dot{\lambda}_N &= -[\lambda_S b + \lambda_N(b - d)], \\ &\Rightarrow \\ \dot{\lambda}_S &= \lambda_S(d + cI^* + u) - \lambda_E cI^*, \\ \dot{\lambda}_E &= \lambda_E(e + d) - \lambda_I e, \\ \dot{\lambda}_I &= -A + \lambda_S cS^* - \lambda_E cS^* - \lambda_I(g + a + d) + \lambda_N a, \\ \dot{\lambda}_N &= -\lambda_S b - \lambda_N(b - d). \end{aligned}$$

Además, para el instante final dado T las condiciones de frontera $S(T), E(T), I(T), N(T)$ son valores libres, lo cual lleva a que se impongan las siguientes condiciones de transversalidad:

$$\lambda_S(T) = 0, \lambda_E(T) = 0, \lambda_I(T) = 0, \lambda_N(T) = 0.$$

De esta manera se obtiene el sistema adjunto

$$\begin{aligned}\dot{\lambda}_S &= \lambda_S (d + cI^* + u) - \lambda_E cI^*, \\ \dot{\lambda}_E &= \lambda_E (e + d) - \lambda_I e, \\ \dot{\lambda}_I &= -A + \lambda_S cS^* - \lambda_E cS^* - \lambda_I (g + a + d) + \lambda_N a, \\ \dot{\lambda}_N &= -\lambda_S b - \lambda_N (b - d),\end{aligned}$$

con :

$$\begin{aligned}\lambda_S(T) &= 0, \\ \lambda_E(T) &= 0, \\ \lambda_I(T) &= 0, \\ \lambda_N(T) &= 0.\end{aligned}$$

Resulta complicado resolver de manera analítica el sistema adjunto para determinar la variable de control $u(t)$, ya que dicho sistema esta acoplado mediante las variables adjuntas $\lambda_S, \lambda_E, \lambda_I$ y λ_N . Sin embargo, es posible caracterizar al control óptimo mediante el Hamiltoniano (3.2.18), para ello es necesario considerar que

$$\frac{\partial}{\partial u} H(t, \mathbf{x}^*, u^*, \boldsymbol{\lambda}^*) = 0$$

es una condición necesaria para que

$$H(t, \mathbf{x}^*(t), u^*(t), \boldsymbol{\lambda}^*(t)) \leq H(t, \mathbf{x}^*(t), u(t), \boldsymbol{\lambda}^*(t)), \quad \forall u(t) \in [0, 0.9],$$

o de manera equivalente

$$H(t, \mathbf{x}^*(t), u^*(t), \boldsymbol{\lambda}^*(t)) = \min_{u \in [0, 0.9]} [H(t, \mathbf{x}^*(t), u(t), \boldsymbol{\lambda}^*(t))].$$

Entonces, derivando el Hamiltoniano (3.2.18) con respecto a u e igualando a cero:

$$\frac{\partial H}{\partial u} = \frac{\partial}{\partial u} [F + \boldsymbol{\lambda} \cdot (\mathbf{f} + \mathbf{u}(t)\mathbf{g})] = 2u^* - \lambda_S S^* = 0.$$

Por lo tanto $u^* = \frac{\lambda_S S^*}{2}$ es el control óptimo.

La variable adjunta λ_S y la variable de estado S^* son necesarias para determinar a u^* , sin embargo no es sencillo de hacer de manera analítica por lo que se encuentran de forma aproximada usando una implementación numérica conocida como *Método Forward-Backward Sweep*, el cual se describe más adelante.

4 Resultados numéricos

Los resultados numéricos mostrados en esta sección corresponden al estudio del problema de control óptimo para el modelo SEIR. Se busca de manera numérica la solución del sistema de estado y el sistema adjunto para determinar los valores aproximados de la variable de control $u^*(t)$.

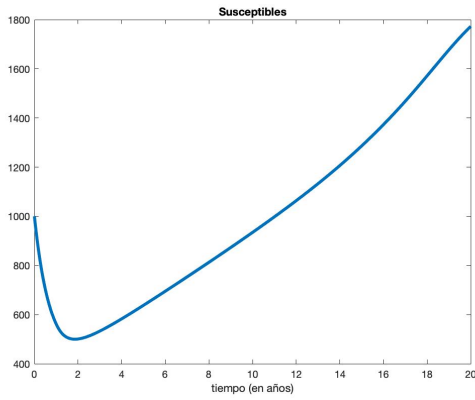
El algoritmo que se aplica se puede consultar a detalle en el artículo *Convergence of the forward-backward sweep method in optimal control* en [18]. Este algoritmo se llama *Método Forward-Backward Sweep* y es una técnica para resolver problemas de control. Es un método indirecto, es decir, aproxima las soluciones resolviendo el problema con valores acotados, que resuelve las ecuaciones diferenciales del Principio del Máximo de Pontryagin de manera numérica. Para esto se utilizan dos teoremas que no se enuncian en este trabajo, sin embargo muestran que al resolver de manera iterativa el sistema de ecuaciones diferenciales se obtiene una sucesión que converge a la solución del sistema. Y el segundo teorema asegura que al discretizar el sistema continuo, la iteración converge al aumentar el número de iteraciones y la cantidad de subintervalos.

La idea central de este algoritmo es tomar el valor inicial de la ecuación de estado y resolverla de manera *forward* (avanzando) en el tiempo, estimando el control y la variable de estado. Luego, usando los valores obtenidos, se toma el valor final de la variable adjunta y se resuelve de manera *backward* (hacia atrás) en el tiempo.

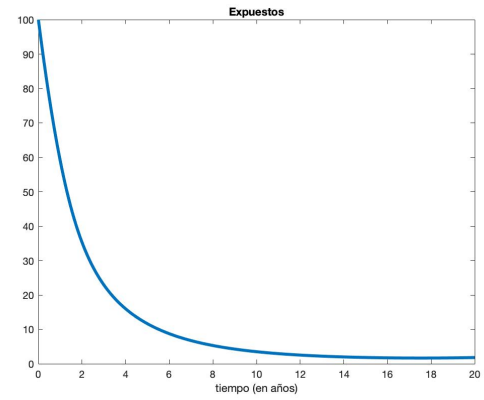
Para implementar el algoritmo primero se definen las variables de estado, las variables de coestado y las variables de control. Se hace una partición del intervalo $[t_0, T]$ en subintervalos para discretizar el tiempo y las variables se inicializan en cero. Por otro lado, se hace uso de las condiciones iniciales del sistema, es decir, se consideran los valores de S_0, E_0, I_0, N_0 en t_0 . Luego, para cada iteración del algoritmo se aplica el método de Runge-Kutta *forward* en el estado y a partir de los valores obtenidos se usa el método de Runge-Kutta *backward* en el coestado. Con la obtención de estos valores se calcula el valor de la variable de control $u(t)$.

Cuando son calculadas todas las variables, se hace una comparación con los valores de la iteración anterior con el objetivo de ir convergiendo de manera numérica a la solución. A continuación se presentan las gráficas resultantes después de aplicar la implementación numérica.

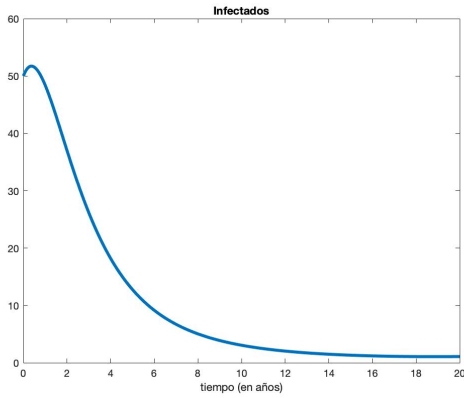
Para la interpretación de los resultados es necesario recordar que la constante c es la tasa de incidencia de la enfermedad y mide el número de casos nuevos en cierto período de tiempo. Si se considera $c = 0.001$, un valor razonablemente realista, la propagación de la enfermedad es “grave”, así que se necesitan medidas de contención. En este caso, se utiliza la política de vacunación óptima encontrada para un 90 % de la población susceptible desde un inicio, esto implica que fueron protegidos (más de la mitad) los individuos susceptibles de los expuestos e infectados, además de que se incorporan los susceptibles protegidos a la población de recuperados. Se observa que es conveniente vacunar alrededor del 40 % de la población, incluso después de que las poblaciones de expuestos e infectados hayan disminuido. Con los primeros años de evolución de la enfermedad se toman las medidas de vacunación pertinentes para contener la enfermedad e ir disminuyendo las poblaciones de expuestos e infectados.



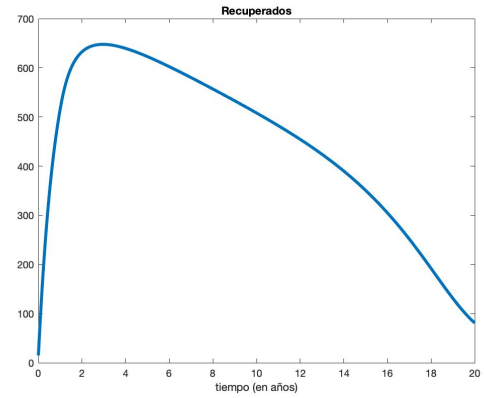
(a) Susceptibles



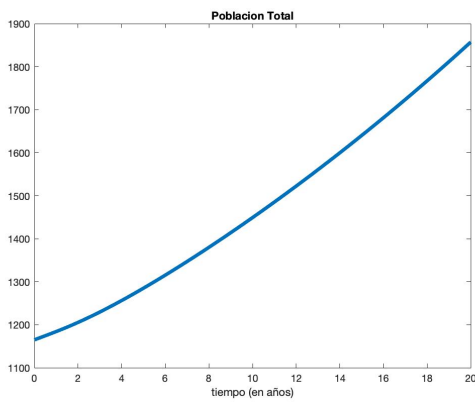
(b) Expuestos



(c) Infectados



(d) Recuperados



(e) Población Total

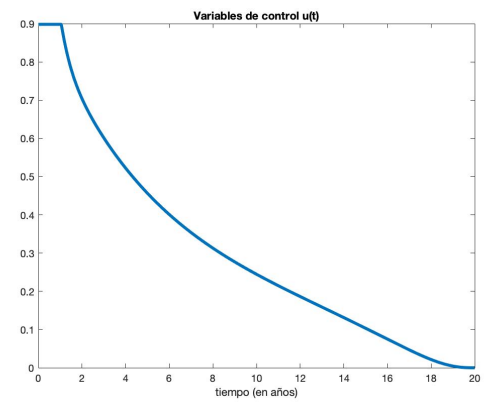
(f) Variable de control $u(t)$

Figura 4.1: Modelo SEIR con control.

5 Conclusiones

En este trabajo se presentaron en el segundo capítulo los elementos básicos de la teoría de los sistemas dinámicos para describir la evolución de un proceso a lo largo del tiempo y del cálculo variacional como una herramienta de optimización. Desde esta perspectiva se abordó el desarrollo de la teoría de control de forma básica combinando las teorías mencionadas para resolver problemas en algunas aplicaciones modernas, haciendo énfasis en dos resultados fundamentales: la Controlabilidad y el Principio del máximo de Pontryagin.

En el tercer capítulo se analizó la construcción a detalle del modelo epidemiológico SEIR, el cual se estudió desde dos enfoques. Para el primer tratamiento del modelo se usaron las técnicas básicas de los sistemas dinámicos, a través de métodos computacionales se obtuvieron gráficas de las soluciones para observar el comportamiento de las poblaciones de los individuos ante la presencia de cierta enfermedad. En el segundo criterio se formuló un problema de control usando el modelo SEIR con políticas de vacunación. El problema consistió en encontrar una política de vacunación de costo óptimo y que al mismo tiempo redujera la población de individuos infectados. En este mismo apartado se utilizó la propiedad de controlabilidad para determinar, en este planteamiento, que no hay una variable de control que dirija al sistema a un estado en particular. Además se observó que no es posible garantizar la estabilidad del modelo ni su controlabilidad. En forma general se plantean modificaciones posteriores para desarrollar un modelo más realista que sea estable y controlable. Finalmente se deja abierta la posibilidad de introducir dos controles independientes, por ejemplo la tasa de vacunación y distanciamiento social.

A Algunos resultados de Sistemas dinámicos

Proposición A.0.1. Si \mathbf{A} y \mathbf{B} son transformaciones lineales en \mathbb{R}^n que conmutan, es decir, que $\mathbf{AB} = \mathbf{BA}$, entonces $e^{\mathbf{A}+\mathbf{B}} = e^{\mathbf{A}}e^{\mathbf{B}}$.

Demostración. Si $\mathbf{AB} = \mathbf{BA}$, entonces por el teorema binomial

$$(\mathbf{A} + \mathbf{B})^n = n! \sum_{n=j+k} \frac{\mathbf{A}^j \mathbf{B}^k}{j!k!}.$$

Entonces

$$e^{\mathbf{A}+\mathbf{B}} = \sum_{n=0}^{\infty} \sum_{n=j+k} \frac{\mathbf{A}^j \mathbf{B}^k}{j!k!} = \sum_{j=0}^{\infty} \frac{\mathbf{A}^j}{j!} \sum_{k=0}^{\infty} \frac{\mathbf{B}^k}{k!} = e^{\mathbf{A}}e^{\mathbf{B}}.$$

Se ha utilizado el hecho de que el producto de dos series absolutamente convergentes es una serie absolutamente convergente (consultar [19]). \square

Lema A.1. Sea \mathbf{A} una matriz cuadrada, entonces

$$\frac{d}{dt}e^{\mathbf{A}t} = \mathbf{A}e^{\mathbf{A}t}.$$

Demostración. Ya que \mathbf{A} conmuta consigo misma, se sigue que por la Proposición A.0.1 y la Definición 2.1.4 se tiene que

$$\begin{aligned} \frac{d}{dt}e^{\mathbf{A}t} &= \lim_{h \rightarrow 0} \frac{e^{\mathbf{A}(t+h)} - e^{\mathbf{A}t}}{h} \\ &= \lim_{h \rightarrow 0} e^{\mathbf{A}t} \frac{(e^{\mathbf{A}h} - \mathbf{I})}{h} \\ &= e^{\mathbf{A}t} \lim_{h \rightarrow 0} \lim_{k \rightarrow \infty} \left(\mathbf{A} + \frac{\mathbf{A}^2 h}{2!} + \cdots + \frac{\mathbf{A}^k h^{k-1}}{k!} \right) \\ &= \mathbf{A}e^{\mathbf{A}t}. \end{aligned}$$

La última igualdad se obtiene por el hecho de que la serie que define $e^{\mathbf{A}h}$ converge uniformemente para $|h| \leq 1$ y por lo tanto se pueden intercambiar los dos límites. \square

Teorema A.2 (Teorema fundamental para sistemas lineales). *Sea \mathbf{A} una matriz de $n \times n$ con coeficientes constantes y $\mathbf{X}(t) = e^{\mathbf{A}t}$.*

1. *La solución única del sistema homogéneo*

$$\begin{aligned}\dot{\mathbf{x}}(t) &= \mathbf{A}\mathbf{x}(t), \\ \mathbf{x}(0) &= \mathbf{x}_0,\end{aligned}\tag{A.1}$$

es

$$\mathbf{x}(t) = \mathbf{X}(t)\mathbf{x}_0 = e^{\mathbf{A}t}\mathbf{x}_0.\tag{A.2}$$

2. *La solución única del sistema no homogéneo*

$$\begin{aligned}\dot{\mathbf{x}}(t) &= \mathbf{A}\mathbf{x}(t) + \mathbf{b}(t), \\ \mathbf{x}(0) &= \mathbf{x}_0,\end{aligned}\tag{A.3}$$

es

$$\mathbf{x}(t) = \mathbf{X}(t)\mathbf{X}^{-1}(0)\mathbf{x}_0 + \int_0^t \mathbf{X}(t)\mathbf{X}^{-1}(s)\mathbf{b}(s)ds,\tag{A.4}$$

donde $\mathbf{X}^{-1}(0) = \mathbf{I}$ es la matriz identidad de tamaño $n \times n$.

Demostración.

Para 1.

Por el Lema A.1, si $\mathbf{x}(t) = e^{\mathbf{A}t}\mathbf{x}_0$, entonces

$$\dot{\mathbf{x}}(t) = \frac{d}{dt}e^{\mathbf{A}t}\mathbf{x}_0 = \mathbf{A}e^{\mathbf{A}t}\mathbf{x}_0 = \mathbf{A}\mathbf{x}(t)$$

para cada $t \in \mathbb{R}$. Además, $\mathbf{x}(0) = \mathbf{I}\mathbf{x}_0 = \mathbf{x}_0$. Por lo tanto $\mathbf{x}(t) = e^{\mathbf{A}t}\mathbf{x}_0$ es una solución. Para ver que ésta es la única solución, sea \mathbf{x} cualquier solución del problema y sea

$$\mathbf{y}(t) = e^{-\mathbf{A}t}\mathbf{x}(t).$$

Entonces, por el Lema A.1 y el hecho de que $\mathbf{x}(t)$ es una solución de

$$\begin{aligned}\dot{\mathbf{y}}(t) &= -\mathbf{A}e^{-\mathbf{A}t}\mathbf{x}(t) + e^{-\mathbf{A}t}\dot{\mathbf{x}}(t) \\ &= -\mathbf{A}e^{-\mathbf{A}t}\mathbf{x}(t) + e^{-\mathbf{A}t}\mathbf{A}\mathbf{x}(t) \\ &= \mathbf{0}\end{aligned}$$

para cada $t \in \mathbb{R}$ dado que $e^{-\mathbf{A}t}$ y \mathbf{A} conmutan. Así que, $\mathbf{y}(t)$ es una constante. Si $t = 0$ muestra que $\mathbf{y}(t) = \mathbf{x}_0$ y entonces cualquier solución del problema es dada por $\mathbf{x}(t) = e^{\mathbf{A}t}\mathbf{y}(t) = e^{\mathbf{A}t}\mathbf{x}_0$.

Para 2.

Por la definición de la función $\mathbf{x}(t)$ en (A.4) se tiene que

$$\dot{\mathbf{x}} = \dot{\mathbf{X}}(t)\mathbf{X}^{-1}(0)\mathbf{x}_0 + \mathbf{X}(t)\mathbf{X}^{-1}(t)\mathbf{b}(t) + \int_0^t \dot{\mathbf{X}}(t)\mathbf{X}^{-1}(s)\mathbf{b}(s)ds. \quad (\text{A.5})$$

Como $\mathbf{X}(t) = e^{\mathbf{A}t}$ es una solución matricial fundamental del sistema homogéneo (A.1), entonces

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{A} \left[\mathbf{X}(t)\mathbf{X}^{-1}(0)\mathbf{x}_0 + \int_0^t \mathbf{X}(t)\mathbf{X}^{-1}(s)\mathbf{b}(s)ds \right] + \mathbf{b}(t) \\ &= \mathbf{A}\mathbf{x}(t) + \mathbf{b}(t), \end{aligned}$$

para cada $t \in \mathbb{R}$. Y esto completa la prueba del teorema. □

B Algunos resultados de Cálculo Variacional

Lema B.1 (Lagrange). Sea $g(x)$ una función continua en $[a, b]$ tal que

$$\int_a^b g(x)h(x)dx = 0,$$

para cada $h \in C^1([a, b])$ tal que $h(a) = h(b) = 0$, entonces $g(x) \equiv 0$.

Demostración. Suponer que existe $x_0 \in (a, b)$ tal que $g(x_0) > 0$. Por continuidad de g , debe existir un intervalo (c, d) con $x_0 \in (c, d)$ y con $(c, d) \subset (a, b)$ en que $g(x) > 0$ para cada $x \in (c, d)$. Se define para cada $x \in [a, b]$, $h(x)$ de la siguiente manera:

$$h(x) = \begin{cases} (x-c)^2(x-d)^2 & \text{si } x \in [c, d] \\ 0 & \text{si } x \notin [c, d] \end{cases}$$

$h(x) \in C^1([a, b])$, como $g(x) > 0 \forall x \in (c, d)$ y $h(x) > 0 \forall x \in (c, d)$ por definición, entonces $g(x)h(x) > 0 \forall x \in (c, d)$. Mientras que $g(x)h(x) = 0 \forall x \notin (c, d)$, entonces se tiene que

$$\int_a^b g(x)h(x)dx = \int_c^d g(x)h(x)dx > 0,$$

donde en la última desigualdad se utiliza la propiedad de monotonía de la integral.

Lo que contradice el hecho de que $\int_a^b g(x)h(x)dx = 0$. Por lo tanto, no existe $x_0 \in (a, b)$ tal que $g(x_0) > 0$. Así que $g(x) \leq 0 \forall x \in (a, b)$.

Similarmente, por el mismo razonamiento para la función $-g$ se concluye que $-g(x) \leq 0 \forall x \in (a, b)$, que es equivalente a $g(x) \geq 0 \forall x \in (a, b)$.

Entonces para cada $x \in (a, b)$, $g(x) \geq 0$ y $g(x) \leq 0$, pero solo ocurren cuando $g(x) = 0 \forall x \in (a, b)$. Por la continuidad de g , $g(a)$ y $g(b)$ también valen cero. Por lo que se concluye que $g(x) \equiv 0$ en el intervalo $[a, b]$. \square

Teorema B.2 (Ecuación de Euler). Si $x^*(t)$ es un mínimo local del Problema (2.2.16), entonces $x^*(t)$ cumple la siguiente condición

$$F_x - \frac{d}{dt}F_{\dot{x}} = 0, \quad \forall t \in [t_0, t_1].$$

Demostración. Suponer que la función $x^*(t)$ es una solución óptima al Problema (2.2.16), es decir, $J(x^*) \leq J(x)$ para toda función admisible $x(t)$.

Sea $x(t) = x^*(t) + \epsilon h(t)$ para $x^* \in C^1([t_0, t_1])$ tal que $J(x^*)$ es mínimo, $\epsilon \approx 0$ y $h(t) \in C^1([t_0, t_1])$ con $h(t_0) = h(t_1) = 0$. Entonces

$$F(t, x(t), \dot{x}(t)) = F(t, x^*(t) + \epsilon h(t), \dot{x}^*(t) + \epsilon \dot{h}(t))$$

y

$$J(x) = \int_{t_0}^{t_1} F(t, x^*(t) + \epsilon h(t), \dot{x}^*(t) + \epsilon \dot{h}(t)) dt.$$

Considerar que $F \in C^1([t_0, t_1])$, entonces

$$\begin{aligned} \dot{J}(\epsilon) \equiv \frac{dJ}{d\epsilon} &= \frac{d}{d\epsilon} \left[\int_{t_0}^{t_1} F(t, x^*(t) + \epsilon h(t), \dot{x}^*(t) + \epsilon \dot{h}(t)) dt \right] \\ &= \int_{t_0}^{t_1} \frac{d}{d\epsilon} \left[F(t, x^*(t) + \epsilon h(t), \dot{x}^*(t) + \epsilon \dot{h}(t)) \right] dt. \end{aligned}$$

Aplicando la regla de la cadena a la expresión anterior, se tiene que

$$\begin{aligned} \dot{J}(\epsilon) &= \int_{t_0}^{t_1} \left[F_t(t, x^*(t) + \epsilon h(t), \dot{x}^*(t) + \epsilon \dot{h}(t)) \frac{dt}{d\epsilon} + \right. \\ &\quad \left. + F_x(t, x^*(t) + \epsilon h(t), \dot{x}^*(t) + \epsilon \dot{h}(t)) h(t) + F_{\dot{x}}(t, x^*(t) + \epsilon h(t), \dot{x}^*(t) + \epsilon \dot{h}(t)) \dot{h}(t) \right] dt. \end{aligned}$$

Si $\epsilon = 0$, entonces

$$\begin{aligned} \dot{J}(0) &= \int_{t_0}^{t_1} F_t(t, x^*(t), \dot{x}^*(t)) \underbrace{\frac{dx}{d\epsilon}}_{=0} + F_x(t, x^*(t), \dot{x}^*(t)) h(t) + F_{\dot{x}}(t, x^*(t), \dot{x}^*(t)) \dot{h}(t) dt \\ &= \int_{t_0}^{t_1} \left[F_x(t, x^*(t), \dot{x}^*(t)) h(t) + F_{\dot{x}}(t, x^*(t), \dot{x}^*(t)) \dot{h}(t) \right] dt. \end{aligned}$$

Para simplificar la notación, se reescribe la expresión anterior como

$$\dot{J}(0) = \int_{t_0}^{t_1} \left[F_x h(t) + F_{\dot{x}} \dot{h}(t) \right] dt = \int_{t_0}^{t_1} F_x h(t) dt + \int_{t_0}^{t_1} F_{\dot{x}} \dot{h}(t) dt = 0. \quad (\text{B.1})$$

Integrando por partes el segundo sumando, se obtiene

$$\int_{t_0}^{t_1} F_{\dot{x}} \dot{h}(t) dt = [F_{\dot{x}}(t) h(t)]|_{t_0}^{t_1} - \int_{t_0}^{t_1} \left[\frac{d}{dt} F_{\dot{x}} \right] h(t) dt$$

y, como $h(t_0) = h(t_1) = 0$ entonces

$$\int_{t_0}^{t_1} F_{\dot{x}} \dot{h}(t) dt = - \int_{t_0}^{t_1} \left[\frac{d}{dt} F_{\dot{x}} \right] h(t) dt,$$

que al sustituir en (B.1) y reordenando

$$\int_{t_0}^{t_1} \left[F_x - \frac{d}{dt} F_{\dot{x}} \right] h(t) dt = 0.$$

Entonces

$$\dot{J}(0) = \int_{t_0}^{t_1} \left[F_x - \frac{d}{dt} F_{\dot{x}} \right] h(t) dt = 0$$

para cada $h \in C^1([t_0, t_1])$ con $h(t_0) = h(t_1) = 0$. Por lo tanto, por el lema B.1 se concluye que $\forall t \in [t_0, t_1]$ se cumple la siguiente condición

$$F_x - \frac{d}{dt} F_{\dot{x}} = 0.$$

□

Teorema B.3 (Primera integral). *Cuando el Lagrangeano F de $\int_{x_0}^{x_1} F(t, x(t), \dot{x}(t)) dt$ no depende explícitamente de t , entonces cualquier extremo no trivial satisface la primera integral de Euler*

$$\dot{y} \frac{\partial F}{\partial x} - F = C_0,$$

donde C_0 es una constante de integración.

Demostración. La demostración se puede consultar en [9].

□

Lema B.4 (Du Bois-Reymond). *Sean $\alpha(t)$ y $\beta(t)$ funciones continuas en $[a, b]$ tales que*

$$\int_a^b (\alpha(t)h(t) + \beta(t)\dot{h}(t)) dt = 0$$

para cada $h \in C^1([a, b])$, con $h(a) = h(b) = 0$. Entonces β es $C^1([a, b])$ con derivada $\dot{\beta}(t) = \alpha(t)$.

Demostración. Se puede consultar en [20].

□

Lema B.5. *Sean $P(t)$ y $Q(t)$ funciones continuas en $I = [a, b]$ tales que el siguiente funcional cuadrático*

$$\int_a^b \left(P(t)\dot{h}(t)^2 + Q(t)h(t)^2 \right) dt$$

esté bien definido para cada $h \in C^1([a, b])$, con $h(a) = h(b) = 0$. Una condición necesaria para que el funcional no sea positivo para toda h es que $P(t) \leq 0$ en I .

Demostración. Se puede consultar en [20].

□

Teorema B.6 (Condición de Legendre).

- Si $x^*(t)$ es un mínimo local del Problema (2.2.16), entonces $x^*(t)$ satisface la siguiente condición

$$F_{\dot{x}\dot{x}}(t, x^*(t), \dot{x}^*(t)) \geq 0, \quad \forall t \in [t_0, t_1].$$

- Si $x^*(t)$ es un máximo local del Problema (2.2.16), entonces $x^*(t)$ satisface la siguiente condición

$$F_{\dot{x}\dot{x}}(t, x^*(t), \dot{x}^*(t)) \leq 0, \quad \forall t \in [t_0, t_1].$$

donde $\frac{\partial^2 F}{\partial \dot{x}^2} = F_{\dot{x}\dot{x}}$

Demostración. Se demuestra la condición para el caso del máximo. En la demostración del Teorema B.2, en donde se obtenía la ecuación de Euler, se han aplicado condiciones de optimalidad de primer orden a la función

$$J(\epsilon) = \int_{t_0}^{t_1} F(t, x^*(t) + \epsilon h(t), \dot{x}^*(t) + \epsilon \dot{h}(t)) dt.$$

Además

$$\dot{J}(\epsilon) = \int_{t_0}^{t_1} [F_x h(t) + F_{\dot{x}} \dot{h}(t)] dt,$$

donde F_x y $F_{\dot{x}}$ dependen de $(t, x^*(t) + \epsilon h(t), \dot{x}^*(t) + \epsilon \dot{h}(t))$ y se escriben así para simplificar la notación. Derivando $\dot{J}(\epsilon)$ con respecto a ϵ se obtiene:

$$\begin{aligned} \ddot{J}(\epsilon) &= \int_{t_0}^{t_1} \frac{d}{d\epsilon} [F_x h(t) + F_{\dot{x}} \dot{h}(t)] dt \\ &= \int_{t_0}^{t_1} \left[h(t) (F_{xx} h(t) + F_{x\dot{x}} \dot{h}(t)) + \dot{h}(t) (F_{\dot{x}x} h(t) + F_{\dot{x}\dot{x}} \dot{h}(t)) \right] dt \\ &= \int_{t_0}^{t_1} [h^2(t) F_{xx} + 2h(t) \dot{h}(t) F_{x\dot{x}} + \dot{h}^2(t) F_{\dot{x}\dot{x}}] dt, \end{aligned}$$

donde $F_{xx}, F_{x\dot{x}}, F_{\dot{x}\dot{x}}$ dependen de $(t, x^*(t) + \epsilon h(t), \dot{x}^*(t) + \epsilon \dot{h}(t))$. Si $\epsilon = 0$, entonces

$$\begin{aligned} \ddot{J}(0) &= \int_{t_0}^{t_1} [h^2(t) F_{xx} + 2h(t) \dot{h}(t) F_{x\dot{x}} + \dot{h}^2(t) F_{\dot{x}\dot{x}}] dt \\ &= \int_{t_0}^{t_1} h^2(t) F_{xx} dt + 2 \int_{t_0}^{t_1} h(t) \dot{h}(t) F_{x\dot{x}} dt + \int_{t_0}^{t_1} \dot{h}^2(t) F_{\dot{x}\dot{x}} dt, \end{aligned}$$

ahora con $F_{xx}, F_{x\dot{x}}, F_{\dot{x}\dot{x}}$ que dependen de $(t, x^*(t), \dot{x}^*(t))$.

Integrando por partes la segunda integral de la expresión anterior, y ya que $h(t_0) = h(t_1) = 0$ queda

$$\begin{aligned} 2 \int_{t_0}^{t_1} h(t) \dot{h}(t) F_{x\dot{x}} dt &= 2 \left\{ \left[\frac{1}{2} h^2(t) F_{x\dot{x}} \right] \Big|_{t_0}^{t_1} - \frac{1}{2} \int_{t_0}^{t_1} h^2(t) \frac{d}{dt} F_{x\dot{x}} dt \right\} \\ &= - \int_{t_0}^{t_1} h^2(t) \frac{d}{dt} F_{x\dot{x}} dt, \end{aligned}$$

entonces

$$\begin{aligned} \ddot{J}(0) &= \int_{t_0}^{t_1} h^2(t) F_{xx} dt + 2 \int_{t_0}^{t_1} h(t) \dot{h}(t) F_{x\dot{x}} dt + \int_{t_0}^{t_1} \dot{h}^2 F_{\dot{x}\dot{x}} dt \\ &= \int_{t_0}^{t_1} \left[h^2(t) F_{xx} - h^2(t) \frac{d}{dt} F_{x\dot{x}} + \dot{h}^2 F_{\dot{x}\dot{x}} \right] dt \\ &= \int_{t_0}^{t_1} \left[F_{\dot{x}\dot{x}} \dot{h}^2 + \left(F_{xx} - \frac{d}{dt} F_{x\dot{x}} \right) h^2(t) \right] dt. \end{aligned}$$

Aplicando el Lema (B.5), con $P(t) = F_{\dot{x}\dot{x}}$ y $Q(t) = F_{xx} - \frac{d}{dt} F_{x\dot{x}}$ se concluye que $F_{\dot{x}\dot{x}} \leq 0 \quad \forall t \in [t_0, t_1]$.

□

C Algunos resultados de Control Óptimo

Teorema C.1 (Matriz de controlabilidad). *Se tiene que*

$$\text{rank } G = n \Leftrightarrow 0 \in \mathcal{C}^o,$$

donde \mathcal{C}^o es el interior del conjunto \mathcal{C} .

Demostración. 1. Suponer que $\text{rank } G < n$. Esto significa que la base de la columnas de G tiene una dimensión inferior o igual a $n - 1$. Entonces existe un vector $\mathbf{b} \in \mathbb{R}^n$, $\mathbf{b} \neq 0$, ortogonal a cada columna de G . Lo que implica que

$$\mathbf{b}^T G = 0,$$

entonces

$$\mathbf{b}^T N = \mathbf{b}^T M N = \dots = \mathbf{b}^T M^{n-1} N = 0.$$

2. Se puede afirmar que

$$\mathbf{b}^T M^k N = 0 \tag{C.1}$$

para cada valor entero positivo k .

Para demostrarlo primero se recuerda que el polinomio característico de una matriz M es:

$$p(\lambda) := \det(M - \lambda I).$$

Por el teorema de Cayley-Hamilton se establece que

$$p(M) = 0.$$

Si se escribe

$$p(\lambda) = \lambda^n + \beta_{n-1}\lambda^{n-1} + \dots + \beta_1\lambda + \beta_0,$$

entonces

$$p(M) = M^n + \beta_{n-1}M^{n-1} + \dots + \beta_1M + \beta_0I = 0.$$

Por lo tanto

$$M^n = -\beta_{n-1}M^{n-1} - \dots - \beta_1M - \beta_0I,$$

y así

$$\mathbf{b}^T M^n N = \mathbf{b}^T (-\beta_{n-1}M^{n-1} - \dots) N = 0.$$

De manera similar,

$$\mathbf{b}^T M^{n+1} N = \mathbf{b}^T (-\beta_{n-1} M^n - \dots) N = 0.$$

Por lo que (C.1) queda demostrada.

Notar que

$$\mathbf{b}^T \mathbf{X}^{-1}(s) N = \mathbf{b}^T e^{-sM} N = \mathbf{b}^T \sum_{k=0}^{\infty} \frac{(-s)^k M^k N}{k!} = \sum_{k=0}^{\infty} \frac{(-s)^k}{k!} \mathbf{b}^T M^k N = 0,$$

de acuerdo con (C.1).

3. Suponer que $\mathbf{x}_0 \in \mathcal{C}(t)$. Esto es equivalente a tener

$$\mathbf{x}_0 = - \int_0^t \mathbf{X}^{-1}(s) N \mathbf{u}(s) ds = 0$$

para algún control $\mathbf{u}(\cdot) \in \mathcal{U}$.

Entonces

$$\mathbf{b} \cdot \mathbf{x}_0 = - \int_0^t \mathbf{b}^T \mathbf{X}^{-1}(s) N \mathbf{u}(s) ds = 0.$$

Esto dice que \mathbf{b} y \mathbf{x}_0 son ortogonales. En otras palabras, \mathcal{C} debe estar en el plano ortogonal a $\mathbf{b} \neq 0$. Por lo tanto $\mathcal{C}^o = \emptyset$.

4. De manera inversa, suponer que $0 \notin \mathcal{C}^o$. Entonces $0 \notin \mathcal{C}^o(t) \forall t > 0$. Como $\mathcal{C}(t)$ es convexo, entonces existe un hiperplano soporte para $\mathcal{C}(t)$ a través de 0. Esto significa que existe $\mathbf{b} \neq 0$ tal que $\mathbf{b} \cdot \mathbf{x}_0 \leq 0, \forall \mathbf{x}_0 \in \mathcal{C}(t)$.

Elegir cualquier $\mathbf{x}_0 \in \mathcal{C}(t)$. Entonces

$$\mathbf{x}_0 = - \int_0^t \mathbf{X}^{-1}(s) N \mathbf{u}(s) ds$$

para cualquier control u , y por lo tanto

$$0 \geq \mathbf{b} \cdot \mathbf{x}_0 = - \int_0^t \mathbf{b}^T \mathbf{X}^{-1}(s) N \mathbf{u}(s) ds = 0.$$

Así

$$\int_0^t \mathbf{b}^T \mathbf{X}^{-1}(s) N \mathbf{u}(s) ds \leq 0, \forall \mathbf{u}(\cdot).$$

Se afirma que, por lo tanto

$$\mathbf{b}^T \mathbf{X}^{-1} N \equiv 0, \tag{C.2}$$

lo anterior se demuestra en el lema (C.2). Reescribiendo (C.2) como

$$\mathbf{b}^T e^{-sM} N \equiv 0. \tag{C.3}$$

Si $s = 0$ entonces $\mathbf{b}^T e^{-0M} N = \mathbf{b}^T N = 0$. Derivando (C.3) con respecto a s se tiene que

$$\mathbf{b}^T (-M) e^{-0M} N \equiv 0.$$

Para $s = 0$ en la ecuación anterior queda

$$\mathbf{b}^T M N = 0.$$

Al hacer repetidamente este proceso, se deduce

$$\mathbf{b}^T M^k N = 0 \quad \forall k = 0, 1, 2, \dots,$$

Así que $\mathbf{b}^T G = 0$. Esto implica que $\text{rank } G \leq n$, debido a que $\mathbf{b} \neq 0$. □

Lema C.2 (Desigualdades integrales). *Suponer que*

$$\int_0^t \mathbf{b}^T \mathbf{X}^{-1}(s) N \mathbf{u}(s) ds \leq 0$$

para cada $\mathbf{u}(\cdot) \in \mathcal{U}$. Entonces

$$\mathbf{b}^T \mathbf{X}^{-1} N \equiv 0$$

Demostración. La demostración se puede consultar en [12]. □

Teorema C.3 (Criterio de controlabilidad). *Sea A el cubo $[-1, 1]^n$ en \mathbb{R}^n . Suponer también que $\text{rank } G = n$, y $\text{Re } \lambda < 0$ para cada valor propio λ de la matriz M .*

Entonces el sistema (2.3.31) es controlable.

Demostración. Ya que $\text{rank } G = n$, el teorema (C.1) dice que \mathcal{C} contiene alguna bola B centrada en 0. Ahora, tomar cualquier $\mathbf{x}_0 \in \mathbb{R}^n$ y considerar la evolución

$$\begin{aligned} \dot{\mathbf{x}}(t) &= M \mathbf{x} \\ \mathbf{x}(0) &= \mathbf{x}_0; \end{aligned}$$

en otras palabras, tomar el control $\mathbf{u}(\cdot) \equiv 0$. Ya que $\text{Re } \lambda < 0$ para cada valor propio λ de la matriz M , entonces el origen es asintóticamente estable. Así que existe un tiempo T tal que $\mathbf{x}(T) \in B$. Entonces $\mathbf{x}(T) \in B \subset \mathcal{C}$; y por lo tanto existe un control $\mathbf{u}(\cdot) \in \mathcal{U}$ que dirige $\mathbf{x}(T)$ a 0 en un tiempo finito. □

Teorema C.4 (Principio del máximo de Pontryagin). *Sea $\mathbf{u}^*(t)$ la trayectoria óptima de control, continua por pedazos, y $\mathbf{x}^*(t)$ la trayectoria de estado óptima asociada, definidas en el intervalo $[t_0, t_1]$. Entonces existe una función vectorial $\boldsymbol{\lambda}^*(t)$ continua que posee primeras derivadas continuas por pedazos, tal que para cada $t \in [t_0, t_1]$ se cumple:*

1.

$$\dot{\lambda}_i^* = -\frac{\partial}{\partial x_i} H(t, \mathbf{x}^*(t), \mathbf{u}^*(t), \boldsymbol{\lambda}^*(t)),$$

para cada $i = 1, \dots, n$ en todos los puntos de continuidad de $\mathbf{u}^*(t)$, además

$$\lambda_1^* = \frac{\partial}{\partial x_1} S[\mathbf{x}^*(t_1)], \dots, \lambda_n^* = \frac{\partial}{\partial x_n} S[\mathbf{x}^*(t_1)],$$

donde al evaluar $t = t_1$ en cada λ_n^* se obtienen las condiciones de transversalidad.

El sistema de ecuaciones diferenciales formado por las $\dot{\lambda}_i^*$ se denomina “sistema adjunto” o “sistema de coestado”. A las funciones $\lambda_i^*(t)$ se les denomina “variables adjuntas” o “variables de coestado”.

2. Para el caso de maximizar, se cumple que

$$H(t, \mathbf{x}^*(t), \mathbf{u}^*(t), \boldsymbol{\lambda}^*(t)) \geq H(t, \mathbf{x}^*(t), \mathbf{u}(t), \boldsymbol{\lambda}^*(t)),$$

para cada $\mathbf{u}(t) \in \mathcal{U}$ (se define de manera similar cambiando la desigualdad en el caso de minimizar el funcional objetivo).

3. La trayectoria de estado óptimo $\mathbf{x}^*(t)$ es solución del sistema (2.3.36) con condición inicial \mathbf{x}_0^*

$$\dot{\mathbf{x}}(t) = \mathbf{f}^*(t, \mathbf{x}^*(t), \mathbf{u}^*(t)),$$

para cada $i = 1, \dots, n$ en todos los puntos de continuidad de $\mathbf{u}^*(t)$,

con

$$\mathbf{x}^*(t) = \begin{pmatrix} x_1^*(t) \\ x_2^*(t) \\ \vdots \\ x_n^*(t) \end{pmatrix} \quad y \quad \mathbf{u}^*(t) = \begin{pmatrix} u_1^*(t) \\ u_2^*(t) \\ \vdots \\ u_m^*(t) \end{pmatrix}.$$

Demostración. Considerar el siguiente problema de control óptimo:

$$\text{máx } J(\mathbf{u}) = \int_{t_0}^{t_1} F[t, \mathbf{x}, \mathbf{u}]dt + S[\mathbf{x}(t_1)] \quad (\text{C.4})$$

$$\text{sujeto a : } \dot{\mathbf{x}} = \mathbf{f}(t, \mathbf{x}, \mathbf{u})$$

$$\text{con : } \mathbf{x}(t_0) = \mathbf{x}_0,$$

donde para hacer más simple la deducción se considera que la variable de control u no esta sujeta a restricciones, es decir, $u(t) \in \mathcal{U} = \mathbb{R}$, y tanto la variable de estado como la variable de control son escalares (unidimensionales). Para esta formulación del problema de control óptimo se deducen a continuación las condiciones necesarias de optimalidad que proporcionan el principio del máximo de Pontryagin, utilizando métodos del cálculo variacional.

El Hamiltoniano asociado al problema es:

$$H(t, x, u, \lambda) = F(t, x, u) + \lambda(t)f(t, x, u).$$

De la ecuación de estado

$$\dot{x} = f(t, x, u)$$

se tiene que

$$f(t, x, u) - \dot{x} = 0, \quad \forall t \in [t_0, t_1],$$

entonces

$$\int_{t_0}^{t_1} \lambda(t)[f(t, x, u) - \dot{x}]dt = 0,$$

para cualquier función $\lambda(t)$ continua y con derivada continua por pedazos.

Sumando el valor de esta integral al funcional objetivo, se tiene

$$\begin{aligned} J &= \int_{t_0}^{t_1} \{F(t, x, u) + \lambda(t)[f(t, x, u) - \dot{x}]\} dt + S[x(t_1)] \\ &= \int_{t_0}^{t_1} [F(t, x, u) + \lambda(t)f(t, x, u) - \lambda(t)\dot{x}]dt + S[x(t_1)] \\ &= \int_{t_0}^{t_1} [F(t, x, u) + \lambda(t)f(t, x, u)]dt - \int_{t_0}^{t_1} \lambda(t)\dot{x}dt + S[x(t_1)] \\ &= \int_{t_0}^{t_1} H(t, x, u, \lambda)dt - \int_{t_0}^{t_1} \lambda(t)\dot{x}dt + S[x(t_1)]. \end{aligned}$$

Integrando por partes la segunda integral

$$\int_{t_0}^{t_1} \lambda(t)\dot{x}dt = \lambda(t)x(t)|_{t_0}^{t_1} - \int_{t_0}^{t_1} x\dot{\lambda}(t)dt = \lambda(t_1)x(t_1) - \lambda(t_0)x(t_0) - \int_{t_0}^{t_1} x\dot{\lambda}(t)dt.$$

Por lo tanto, se tiene que:

$$J = \int_{t_0}^{t_1} [H(t, x, u, \lambda) + x\dot{\lambda}(t)] dt - \lambda(t_1)x(t_1) + \lambda(t_0)x(t_0) + S[x(t_1)],$$

para cualquier trayectoria de control $u(t)$, con trayectoria de estado asociada $x(t)$, por lo que se verifica que $\dot{x} = f(t, x, u)$, con $x(t_0) = x_0$.

Sea $u^*(t)$ la trayectoria de control óptimo. Se perturba dicha trayectoria con una función arbitraria $\alpha(t)$ continua por pedazos. Sea:

$$u_\epsilon(t) = u^*(t) + \epsilon\alpha(t), \quad \forall t \in [t_0, t_1],$$

con $\alpha(t)$ fija y $\epsilon > 0$.

Sea $x^*(t)$ la trayectoria de estado óptima asociada al control $u^*(t)$. Sea $x(t, \epsilon)$ la trayectoria de estado asociada al control $u_\epsilon(t)$. Suponer que $x(t, \epsilon)$ es una función continua, con derivada parcial con respecto a ϵ continua. Es claro que si $\epsilon = 0$ entonces $x(t, 0) = x^*(t)$, $\forall t \in [t_0, t_1]$.

Dados $u^*(t)$ y $\alpha(t)$, entonces el valor del funcional objetivo asociado a $u_\epsilon(t)$ y $x(t, \epsilon)$ depende de manera única de ϵ y cuyo valor es:

$$\begin{aligned} J(\epsilon) &= \int_{t_0}^{t_1} [H(t, x(t, \epsilon), u^*(t) + \epsilon\alpha(t), \lambda(t)) + x(t, \epsilon)\dot{\lambda}(t)] dt \\ &\quad - \lambda(t_1)x(t_1, \epsilon) + \lambda(t_0)x_0 + S[x(t_1, \epsilon)]. \end{aligned}$$

Como $u^*(t)$ es el control óptimo y $x^*(t)$ es la trayectoria óptima de estado, $J(\epsilon)$ alcanza el valor máximo cuando $\epsilon = 0$, por lo que se cumplirá la condición necesaria de optimalidad, es decir $\dot{J}(0) = 0$.

Derivando $J(\epsilon)$ con respecto a ϵ se tiene que:

$$\begin{aligned} \dot{J}(\epsilon) &= \frac{dJ}{d\epsilon} = \int_{t_0}^{t_1} \left[\frac{\partial H}{\partial t} \frac{\partial t}{\partial \epsilon} + \frac{\partial H}{\partial x} \frac{\partial x}{\partial \epsilon} + \frac{\partial H}{\partial u} \frac{\partial u}{\partial \epsilon} + \frac{\partial x}{\partial \epsilon} \dot{\lambda}(t) \right] dt \\ &\quad - \lambda(t_1) \frac{\partial x(t_1, \epsilon)}{\partial \epsilon} + \frac{\partial S[x(t_1, \epsilon)]}{\partial x} \frac{\partial x}{\partial \epsilon} \\ &= \int_{t_0}^{t_1} \left[\frac{\partial H}{\partial x} \frac{\partial x}{\partial \epsilon} + \frac{\partial H}{\partial u} \alpha(t) + \frac{\partial x}{\partial \epsilon} \dot{\lambda}(t) \right] dt \\ &\quad - \lambda(t_1) \frac{\partial x}{\partial \epsilon} + \frac{\partial S[x]}{\partial x} \frac{\partial x}{\partial \epsilon}. \end{aligned}$$

Si $\epsilon = 0$ e igualando a cero se tiene que:

$$\begin{aligned} \dot{J}(0) &= \int_{t_0}^{t_1} \left[\left(\frac{\partial H^*}{\partial x} + \dot{\lambda}(t) \right) \frac{\partial x}{\partial \epsilon} + \frac{\partial H^*}{\partial u} \alpha(t) \right] dt \\ &\quad - \lambda(t_1) \frac{\partial x(t_1, 0)}{\partial \epsilon} + \frac{\partial S[x(t_1, 0)]}{\partial x} \frac{\partial x}{\partial \epsilon} \\ &= \int_{t_0}^{t_1} \left[\left(\frac{\partial H^*}{\partial x} + \dot{\lambda}(t) \right) \frac{\partial x}{\partial \epsilon} + \frac{\partial H^*}{\partial u} \alpha(t) \right] dt \\ &\quad - \lambda(t_1) \frac{\partial x(t_1)}{\partial \epsilon} + \frac{\partial S[x(t_1)]}{\partial x} \frac{\partial x}{\partial \epsilon} \\ &= \int_{t_0}^{t_1} \left[\left(\frac{\partial H^*}{\partial x} + \dot{\lambda}(t) \right) \frac{\partial x}{\partial \epsilon} + \frac{\partial H^*}{\partial u} \alpha(t) \right] dt \\ &\quad + \left(\frac{dS}{dx} - \lambda(t_1) \right) \frac{\partial x}{\partial \epsilon} = 0, \end{aligned}$$

donde

$$\frac{\partial H^*}{\partial x} = \frac{\partial}{\partial x} H(t, x^*, u^*, \lambda^*), \quad \frac{\partial H^*}{\partial u} = \frac{\partial}{\partial u} H(t, x^*, u^*, \lambda^*).$$

Como $\frac{\partial x}{\partial \epsilon}$ es difícil de determinar, se selecciona $\lambda(t)$ de manera que no haya necesidad de calcular dicha derivada parcial, por lo que $\lambda^*(t)$ se toma de manera que se cumpla:

$$\dot{\lambda}^*(t) = -\frac{\partial}{\partial x} H(t, x^*, u^*, \lambda^*), \quad \text{con } \lambda^*(t_1) = \frac{d}{dx} S[x^*(t_1)],$$

entonces

$$\dot{J}(0) = \int_{t_0}^{t_1} \left(\frac{\partial}{\partial x} H(t, x^*, u^*, \lambda^*) \right) \alpha(t) dt = 0,$$

que se debe de cumplir para cualquier función $\alpha(t)$ continua por pedazos. En particular, si se toma

$$\alpha(t) = \frac{\partial}{\partial x} H(t, x^*, u^*, \lambda^*)$$

se tiene que satisfacer lo siguiente:

$$\int_{t_0}^{t_1} \left[\frac{\partial}{\partial x} H(t, x^*, u^*, \lambda^*) \right]^2 dt = 0,$$

por lo que

$$\frac{\partial}{\partial u} H(t, x^*, u^*, \lambda^*) = 0, \forall t \in [t_0, t_1].$$

Por lo tanto, se dedujo que existe una función $\lambda^*(t)$ continua con derivada continua por pedazos, tal que para cada $t \in [t_0, t_1]$ se cumple que

1.

$$\dot{\lambda}^*(t) = -\frac{\partial}{\partial x} H(t, x^*, u^*, \lambda^*),$$

en todos los puntos de continuidad de $u^*(t)$,

$$\text{con } \lambda^*(t_1) = \frac{d}{dx} S[x^*(t_1)].$$

2.

$$\frac{\partial}{\partial u} H(t, x^*, u^*, \lambda^*) = 0,$$

que es una condición necesaria para que

$$H(t, x^*(t), u^*(t), \lambda^*(t)) \geq H(t, x^*(t), u(t), \lambda^*(t)), \quad \forall u \in \mathbb{R}.$$

3.

$$\dot{x} = f(t, x, u),$$

en todos los puntos de continuidad de $u^*(t)$, con $x^*(t_0) = x_0$.

Por lo tanto, se han obtenido las condiciones que constituyen el principio del máximo de Pontryagin. La demostración se sigue manteniendo cuando se consideran funciones vectoriales. \square

D Implementación Numérica

Se presenta el código del *Método Forward-Backward Sweep* para aproximar las soluciones al problema de control en el modelo SEIR, cuya programación se hizo en MATLAB (el código original se puede consultar en [13]). Además se añadieron algunas líneas de código para utilizar el método de Runge Kutta en el modelo SEIR.

```
1 function y = Control(b,d,c,e,g,a,S0,E0,I0,R0,A,T)
2
3 test = -1;
4
5 delta = 0.001;
6 M = 1000;
7 t=linspace(0,T,M+1);
8 h=T/M;
9 h2 = h/2;
10
11 %Variables en el modelo SEIR
12 Sn=zeros(1,M+1);
13 En=zeros(1,M+1);
14 In=zeros(1,M+1);
15 Rn=zeros(1,M+1);
16 Nn=zeros(1,M+1);
17 Sn(1)=S0;
18 En(1)=E0;
19 In(1)=I0;
20 Rn(1)=R0;
21 Nn(1)=S0+E0+I0+R0;
22
23 %Variables en el modelo SEIR con control
24 S=zeros(1,M+1);
25 E=zeros(1,M+1);
26 I=zeros(1,M+1);
27 R=zeros(1,M+1);
28 N=zeros(1,M+1);
29 S(1)=S0;
30 E(1)=E0;
31 I(1)=I0;
32 R(1)=R0;
33 N(1)=S0+E0+I0+R0;
34
35
36 %Variables adjuntas del sistema adjunto
37 lambda1=zeros(1,M+1);
```

```

38 lambda2=zeros(1,M+1);
39 lambda3=zeros(1,M+1);
40 lambda4=zeros(1,M+1);
41
42 u=zeros(1,M+1);
43
44 while(test < 0)
45
46     oldu = u;
47     oldS = S;
48     oldE = E;
49     oldI = I;
50     oldN = N;
51     oldlambda1 = lambda1;
52     oldlambda2 = lambda2;
53     oldlambda3 = lambda3;
54     oldlambda4 = lambda4;
55
56     for i = 1:M
57         % Implementacion del metodo de Runge-Kutta para el modelo SEIR
58         n11 = b*Nn(i) - d*Sn(i) - c*Sn(i)*In(i);
59         n12 = c*Sn(i)*In(i) - (e+d)*En(i);
60         n13 = e*En(i) - (g+a+d)*In(i);
61         n14 = (b-d)*Nn(i) - a*In(i);
62
63         n21 = b*(Nn(i)+h2*n14) - d*(Sn(i)+h2*n11) - c*(Sn(i)+h2*n11)*(In(i)+h2*
64         *n13);
65         n22 = c*(Sn(i)+h2*n11)*(In(i)+h2*n13) - (e+d)*(En(i)+h2*n12);
66         n23 = e*(En(i)+h2*n12) - (g+a+d)*(In(i)+h2*n13);
67         n24 = (b-d)*(Nn(i)+h2*n14) - a*(In(i)+h2*n13);
68
69         n31 = b*(Nn(i)+h2*n24) - d*(Sn(i)+h2*n21) - c*(Sn(i)+h2*n21)*(I(i)+h2*
70         n23);
71         n32 = c*(Sn(i)+h2*n21)*(I(i)+h2*n23) - (e+d)*(E(i)+h2*n22);
72         n33 = e*(E(i)+h2*n22) - (g+a+d)*(In(i)+h2*n23);
73         n34 = (b-d)*(Nn(i)+h2*n24) - a*(In(i)+h2*n23);
74
75         n41 = b*(Nn(i)+h*n34) - d*(Sn(i)+h*n31) - c*(Sn(i)+h*n31)*(In(i)+h*n33
76         );
77         n42 = c*(Sn(i)+h*n31)*(In(i)+h*n33) - (e+d)*(En(i)+h*n32);
78         n43 = e*(En(i)+h*n32) - (g+a+d)*(In(i)+h*n33);
79         n44 = (b-d)*(Nn(i)+h*n34) - a*(In(i)+h*n33);
80
81         Sn(i+1) = Sn(i) + (h/6)*(n11 + 2*n21 + 2*n31 + n41);
82         En(i+1) = En(i) + (h/6)*(n12 + 2*n22 + 2*n32 + n42);
83         In(i+1) = In(i) + (h/6)*(n13 + 2*n23 + 2*n33 + n43);
84         Nn(i+1) = Nn(i) + (h/6)*(n14 + 2*n24 + 2*n34 + n44);
85
86         %Implementacion del metodo de Runge Kutta al resolver las variables de
87         estado hacia adelante en el tiempo
88         m11 = b*N(i) - d*S(i) - c*S(i)*I(i) - u(i)*S(i);

```



```

88     m12 = c*S(i)*I(i) - (e+d)*E(i);
89     m13 = e*E(i) - (g+a+d)*I(i);
90     m14 = (b-d)*N(i) - a*I(i);
91
92     m21 = b*(N(i)+h2*m14) - d*(S(i)+h2*m11) - c*(S(i)+h2*m11)*(I(i)+h2*m13
93 ) - 0.5*(u(i)+u(i+1))*(S(i)+h2*m11);
94     m22 = c*(S(i)+h2*m11)*(I(i)+h2*m13) - (e+d)*(E(i)+h2*m12);
95     m23 = e*(E(i)+h2*m12) - (g+a+d)*(I(i)+h2*m13);
96     m24 = (b-d)*(N(i)+h2*m14) - a*(I(i)+h2*m13);
97
98     m31 = b*(N(i)+h2*m24) - d*(S(i)+h2*m21) - c*(S(i)+h2*m21)*(I(i)+h2*m23
99 ) - 0.5*(u(i)+u(i+1))*(S(i)+h2*m21);
100    m32 = c*(S(i)+h2*m21)*(I(i)+h2*m23) - (e+d)*(E(i)+h2*m22);
101    m33 = e*(E(i)+h2*m22) - (g+a+d)*(I(i)+h2*m23);
102    m34 = (b-d)*(N(i)+h2*m24) - a*(I(i)+h2*m23);
103
104    m41 = b*(N(i)+h*m34) - d*(S(i)+h*m31) - c*(S(i)+h*m31)*(I(i)+h*m33) -
105    u(i+1)*(S(i)+h*m31);
106    m42 = c*(S(i)+h*m31)*(I(i)+h*m33) - (e+d)*(E(i)+h*m32);
107    m43 = e*(E(i)+h*m32) - (g+a+d)*(I(i)+h*m33);
108    m44 = (b-d)*(N(i)+h*m34) - a*(I(i)+h*m33);
109
110    S(i+1) = S(i) + (h/6)*(m11 + 2*m21 + 2*m31 + m41);
111    E(i+1) = E(i) + (h/6)*(m12 + 2*m22 + 2*m32 + m42);
112    I(i+1) = I(i) + (h/6)*(m13 + 2*m23 + 2*m33 + m43);
113    N(i+1) = N(i) + (h/6)*(m14 + 2*m24 + 2*m34 + m44);
114
115    end
116
117    %Implementacion del metodo de Runge Kutta resolviendo el sistema adjunto (
118    sistema para las lambdas) hacia atras en el tiempo
119    for i = 1:M
120        j = M + 2 - i;
121        m11 = lambda1(j)*(d + c*I(j) + u(j)) - c*lambda2(j)*I(j);
122        m12 = lambda2(j)*(e + d) - lambda3(j)*e;
123        m13 = -A + (lambda1(j) - lambda2(j))*c*S(j) + lambda3(j)*(g+a+d) +
124        lambda4(j)*a;
125        m14 = -lambda1(j)*b - lambda4(j)*(b-d);
126
127        m21 = (lambda1(j)-h2*m11)*(d + c*0.5*(I(j) + I(j-1)) + 0.5*(u(j) + u(j
128 -1))) - c*(lambda2(j)-h2*m12)*0.5*(I(j) + I(j-1));
129        m22 = (lambda2(j)-h2*m12)*(e + d) - (lambda3(j)-h2*m13)*e;
130        m23 = -A + ((lambda1(j)-h2*m11) - (lambda2(j)-h2*m12))*c*0.5*(S(j) + S
131 (j-1)) + (lambda3(j)-h2*m13)*(g+a+d) + (lambda4(j)-h2*m14)*a;
132        m24 = -(lambda1(j)-h2*m11)*b - (lambda4(j)-h2*m14)*(b-d);
133
134        m31 = (lambda1(j)-h2*m21)*(d + c*0.5*(I(j) + I(j-1)) + 0.5*(u(j) + u(j
135 -1))) - c*(lambda2(j)-h2*m22)*0.5*(I(j) + I(j-1));
136        m32 = (lambda2(j)-h2*m22)*(e + d) - (lambda3(j)-h2*m23)*e;
137        m33 = -A + ((lambda1(j)-h2*m21) - (lambda2(j)-h2*m22))*c*0.5*(S(j) + S
138 (j-1)) + (lambda3(j)-h2*m23)*(g+a+d) + (lambda4(j)-h2*m24)*a;
139        m34 = -(lambda1(j)-h2*m21)*b - (lambda4(j)-h2*m24)*(b-d);
140
141        m41 = (lambda1(j)-h2*m31)*(d + c*0.5*(I(j) + I(j-1)) + 0.5*(u(j) + u(j
142 -1))) - c*(lambda2(j)-h2*m32)*0.5*(I(j) + I(j-1));
143        m42 = (lambda2(j)-h2*m32)*(e + d) - (lambda3(j)-h2*m33)*e;
144        m43 = -A + ((lambda1(j)-h2*m31) - (lambda2(j)-h2*m32))*c*0.5*(S(j) + S
145 (j-1)) + (lambda3(j)-h2*m33)*(g+a+d) + (lambda4(j)-h2*m34)*a;
146        m44 = -(lambda1(j)-h2*m31)*b - (lambda4(j)-h2*m34)*(b-d);
147
148        S(j) = S(j) - h*(m41);
149        E(j) = E(j) - h*(m42);
150        I(j) = I(j) - h*(m43);
151        N(j) = N(j) - h*(m44);
152    end

```

```

132     m41 = (lambda1(j)-h*m31)*(d + c*I(j-1) + u(j-1)) - c*(lambda2(j)-h*m32
) * I(j-1);
133     m42 = (lambda2(j)-h*m32)*(e + d) - (lambda3(j)-h*m33)*e;
134     m43 = -A + ((lambda1(j)-h*m31) - (lambda2(j)-h*m32))*c*S(j-1) + (
lambda3(j)-h*m33)*(g+a+d) + (lambda4(j)-h*m34)*a;
135     m44 = -(lambda1(j)-h*m31)*b - (lambda4(j)-h*m34)*(b-d);
136
137     lambda1(j-1) = lambda1(j) - (h/6)*(m11 + 2*m21 + 2*m31 + m41);
138     lambda2(j-1) = lambda2(j) - (h/6)*(m12 + 2*m22 + 2*m32 + m42);
139     lambda3(j-1) = lambda3(j) - (h/6)*(m13 + 2*m23 + 2*m33 + m43);
140     lambda4(j-1) = lambda4(j) - (h/6)*(m14 + 2*m24 + 2*m34 + m44);
141 end
142 %Representacion de las variables de control u usando los nuevos
valores para lambda
143 temp=(S.*lambda1)./2;
144 u1 = min(0.9,max(0,temp));
145 u = 0.5*(u1 + oldu);
146
147
148 %Parametros para probar la convergencia de cada variable.
149 temp1 = delta*sum(abs(u)) - sum(abs(oldu - u));
150 temp2 = delta*sum(abs(S)) - sum(abs(oldS - S));
151 temp3 = delta*sum(abs(E)) - sum(abs(oldE - E));
152 temp4 = delta*sum(abs(I)) - sum(abs(oldI - I));
153 temp5 = delta*sum(abs(N)) - sum(abs(oldN - N));
154 temp6 = delta*sum(abs(lambda1)) - sum(abs(olddlambd1 - lambda1));
155 temp7 = delta*sum(abs(lambda2)) - sum(abs(olddlambd2 - lambda2));
156 temp8 = delta*sum(abs(lambda3)) - sum(abs(olddlambd3 - lambda3));
157 temp9 = delta*sum(abs(lambda4)) - sum(abs(olddlambd4 - lambda4));
158 test = min(temp1, min(temp2, min(temp3, min(temp4, min(temp5, min(temp6,
min(temp7, min(temp8, temp9)))))))));
159 end
160
161 %Con los valores obtenidos de la variable u se implementa el metodo de Runge
Kutta para resolver la ecuacion de los recuperados
162 for i=1:M
163     n1 = g*In(i) - d*Rn(i);
164     n2 = g*0.5*(In(i)+In(i+1)) - d*(Rn(i)+h2*n1);
165     n3 = g*0.5*(In(i)+In(i+1)) - d*(Rn(i)+h2*n2);
166     n4 = g*In(i+1) - d*(Rn(i)+h*n3);
167     Rn(i+1) = Rn(i) + (h/6)*(n1 + 2*n2 + 2*n3 + n4);
168
169     m1 = g*I(i) - d*R(i) + u(i)*S(i);
170     m2 = g*0.5*(I(i)+I(i+1)) - d*(R(i)+h2*m1) + 0.5*(u(i)+u(i+1))*0.5*(S(i)+S(
i+1));
171     m3 = g*0.5*(I(i)+I(i+1)) - d*(R(i)+h2*m2) + 0.5*(u(i)+u(i+1))*0.5*(S(i)+S(
i+1));
172     m4 = g*I(i+1) - d*(R(i)+h*m3) + u(i+1)*S(i+1);
173     R(i+1) = R(i) + (h/6)*(m1 + 2*m2 + 2*m3 + m4);
174 end
175 %Una vez que se obtiene la convergencia deseada, los valores de los vectores
finales se almacenan en la matriz 'y'
176 y(1,:) = t;
177 y(2,:) = S;

```

```
178 y(3,:) = E;  
179 y(4,:) = I;  
180 y(5,:) = R;  
181  
182 y(6,:) = N;  
183 y(7,:) = u;
```


Bibliografía

- [1] K. Dietz and J. Heesterbeek, “Daniel bernoulli’s epidemiological model revisited,” *Mathematical biosciences*, vol. 180, no. 1-2, pp. 1–21, 2002.
- [2] W. O. Kermack, A. G. McKendrick, and G. T. Walker, “A contribution to the mathematical theory of epidemics,” *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, vol. 115, no. 772, pp. 700–721, 1927.
- [3] P. Blanchard, R. Devaney, and G. Hall, “Differential equations, brooks,” *Cole, Pa*, 1998.
- [4] L. Perko, *Differential equations and dynamical systems*, vol. 7. Springer Science & Business Media, 2013.
- [5] F. Brauer and J. A. Nohel, *The qualitative theory of ordinary differential equations: an introduction*. Courier Corporation, 1989.
- [6] A. Medio and M. Lines, *Nonlinear dynamics: A primer*. Cambridge University Press, 2001.
- [7] Y. A. Kuznetsov, *Elements of applied bifurcation theory*, vol. 112. Springer Science & Business Media, 2013.
- [8] J. L. Troutman, *Variational calculus and optimal control: optimization with elementary convexity*. Springer Science & Business Media, 2012.
- [9] F. Wan, *Introduction to the Calculus of Variations and its Applications*. Routledge, 2017.
- [10] C. Emilio, “Optimización dinámica,” *editorial, Prentice Hall*, 2001.
- [11] R. Larson and B. H. Edwards, *Multivariable calculus*. Cengage Learning, 2013.
- [12] L. C. Evans, “An introduction to mathematical optimal control theory version 0.2,” *Lecture notes available at <http://math.berkeley.edu/~evans/control.course.pdf>*, 1983.
- [13] S. Lenhart and J. T. Workman, *Optimal control applied to biological models*. CRC press, 2007.
- [14] A. Bressan and B. Piccoli, *Introduction to the mathematical theory of control*, vol. 1. American institute of mathematical sciences Springfield, 2007.

- [15] S. Sastry, *Nonlinear systems: analysis, stability, and control*, vol. 10. Springer Science & Business Media, 2013.
- [16] D. GREENHALGH, “Some results for an seir epidemic model with density dependence in the death rate,” *Mathematical Medicine and Biology: A Journal of the IMA*, vol. 9, no. 2, pp. 67–106, 1992.
- [17] K. Atkinson, W. Han, and D. E. Stewart, *Numerical solution of ordinary differential equations*, vol. 108. John Wiley & Sons, 2011.
- [18] M. McAsey, L. Mou, and W. Han, “Convergence of the forward-backward sweep method in optimal control,” *Computational Optimization and Applications*, vol. 53, no. 1, pp. 207–226, 2012.
- [19] W. Rudin *et al.*, *Principles of mathematical analysis*. McGraw-hill New York, 1964.
- [20] J. Ize, “Cálculo de variaciones,” *FENOMECC*, 2002.
- [21] B. Barnes and G. R. Fulford, *Mathematical modelling with case studies: a differential equations approach using Maple and MATLAB*, vol. 25. CRC Press, 2011.