



UNIVERSIDAD AUTÓNOMA
METROPOLITANA
UNIDAD IZTAPALAPA

MAESTRÍA EN CIENCIAS MATEMÁTICAS
APLICADAS E INDUSTRIALES

MÉTODOS VARIACIONALES Y DE
DISPARO MÚLTIPLE PARA
DETERMINACIÓN DE PARÁMETROS EN
ECUACIONES DIFERENCIALES

T E S I S

QUE PARA OBTENER EL GRADO DE:

Maestro en Ciencias

PRESENTA:

JORGE GABRIEL LEYVA DURÁN

DIRECTOR DE TESIS:

DR. LORENZO HÉCTOR JUÁREZ
VALENCIA

Asesor



Ciudad de México, Enero 2026

Resumen

En el presente trabajo se estudia el método de disparo múltiple y su aplicación a la estimación de parámetros en dos modelos: el modelo SEIRD y el modelo de síntesis enzimática inducible. Se explorará dicho método de forma numérica a través del Lagrangiano aumentado y modelos penalizados para estimar los parámetros usando datos con ruido previamente ajustados.

A diferencia de otros enfoques presentes en la literatura, que emplean algoritmos como Gauss-Newton o SQP, en este estudio se implementa el algoritmo BFGS con búsqueda de línea inexacta. Asimismo, se incorpora el método de la ecuación adjunta para el cálculo eficiente de derivadas y gradientes.

Finalmente, se realizan experimentos utilizando los datos sintéticos con ruido ajustados previamente generados, y se presentan los resultados junto con un análisis comparativo bajo el desempeño realizado por el método de resolución de la ecuación adjunta, en particular el método de Runge-Kutta de orden 2, usando el método del trapecio.

Palabras clave: Estimación de parámetros, disparo múltiple, lagrangiano aumentado, método cuasi-Newton, búsqueda lineal, ecuación adjunta, enfoque variacional.

Abstract

In this work, the multiple shooting method is studied and applied to parameter estimation in two models: the SEIRD model and the inducible enzyme synthesis model. This method will be explored numerically using the augmented Lagrangian and penalized models to estimate parameters using previously adjusted noisy data.

Unlike other approaches in the literature, which use algorithms such as Gauss-Newton or SQP, this study implements the BFGS algorithm with inexact line search. Furthermore, the adjoint equation method is incorporated for the efficient calculation of derivatives and gradients.

Finally, experiments are performed using the previously generated adjusted noisy synthetic data, and the results are presented along with a comparative analysis of the performance of the adjoint equation method, specifically the second order Runge-Kutta method, using the trapezoidal rule.

Keywords: Parameter estimation, multiple shooting, augmented Lagrangian, quasi-Newton method, line search, adjoint equation, variational approach.

Agradecimientos

Este trabajo está dedicado a las personas e instituciones que estuvieron brindándome su apoyo en todo momento. Quiero expresar mi más sincero agradecimiento a mi director de tesis, el Dr. Lorenzo Héctor Juárez Valencia, por su guía, apoyo, dedicación y valiosas observaciones a lo largo del desarrollo y evolución de este trabajo. Su apoyo académico y humano fue invaluable para la culminación de este trabajo.

Agradezco también a los miembros del jurado y a todos mis profesores tanto de la Licenciatura en Matemáticas como de la Maestría en Ciencias Matemáticas Aplicadas e Industriales, que estuvieron brindándome el conocimiento y apoyo necesario para que este trabajo pudiera ser posible. Sin su significativa contribución a mi formación académica este trabajo no hubiera sido posible.

De igual manera, agradezco a la Universidad Autónoma Metropolitana, plantel Iztapalapa (UAM-I) por haber sido mi casa durante mi licenciatura y mi posgrado. También agradezco a CONAHCYT/SECIHTI por haberme brindado el apoyo financiero que hizo posible la realización de este trabajo.

Finalmente, agradezco profundamente a mi familia. A mi madre Alicia de la Luz Durán González y a mi padre Jorge Enrique Leyva Martínez, quienes han estado siempre presentes en mi vida y a lo largo de toda mi trayectoria académica. Gracias por la comprensión en todo momento y por el apoyo en mis momentos de mayor angustia y tranquilidad y por enseñarme a perseverar incluso en los momentos más difíciles y ser siempre mi motivación. A mi pareja, Claudia, le agradezco el haberme apoyado incondicionalmente en todo momento y por haber sido siempre mi calma en mis momentos más complicados. Su paciencia, comprensión y apoyo constante brindado en esta etapa ha sido clave y gran parte de mi motivación.

Índice general

1. Introducción	1
2. Modelo cuadrático, cálculo variacional y disparo simple	3
2.1. Modelo cuadrático para la determinación de parámetros a partir de datos experimentales	3
2.2. Enfoque Variacional	5
2.3. La variación respecto de los parámetros y la ecuación adjunta	7
2.3.1. La función Lagrangiana y la derivada de $J(\mathbf{s}_0, \boldsymbol{\theta})$	9
2.4. Solución de la ecuación adjunta y cálculo del gradiente	11
2.5. Método de disparo simple	13
2.5.1. El Lagrangiano aumentado y el cálculo del gradiente	14
3. Método de disparo múltiple para la estimación de parámetros a partir de datos con ruido	17
3.1. Lagrangiano aumentado	18
3.2. Análisis de perturbación	19
3.3. Ecuación adjunta	20
3.4. Cálculo del gradiente	20
3.5. Algoritmos de optimización	21
3.6. Problema de Mínimos Cuadrados	22
3.6.1. Factorización QR	24
4. Modelo SEIRD	31
4.1. Resultados numéricos para el modelo SEIRD	32
5. Modelo de la síntesis enzimática inducible	37
5.1. Descripción del Modelo	38
5.2. Resultados numéricos para el modelo de la síntesis enzimática inducible . .	41
5.2.1. Resultados numéricos con condiciones iniciales de Heinmets.	42
5.2.2. Resultados numéricos con condiciones iniciales de Milstein.	47
6. Conclusiones del trabajo	53

Capítulo 1

Introducción

Los sistemas de ecuaciones diferenciales ordinarias (y parciales) son una herramienta importante para la modelación de fenómenos físicos que ocurren en el mundo real. La constante evolución de los métodos de solución de estos sistemas de ecuaciones ha generado métodos que permiten hacer predicciones a futuro, o bien controlar estos procesos. Para que dicha predicción sea útil, no solamente se requiere una descripción certera del sistema sino también encontrar o mejorar los parámetros a calcular. La estimación de parámetros inexactos, o bien desconocidos, requiere un ajuste parcial de los datos (con ruido) o datos experimentales al modelo.

En el contexto estadístico y de *machine learning*, se han implementado varios métodos para ajustar los sistemas de ecuaciones diferenciales a los datos. Desde el enfoque de máxima verosimilitud [12], algoritmos de muestreo Bayesiano [13], hasta aproximaciones determinísticas tradicionales. Por lo tanto, la estimación de parámetros requiere de un *solver* eficiente de sistemas de ecuaciones diferenciales, rutinas de optimización, así como también procedimientos estadísticos y estocásticos. A diferencia de los algoritmos estocásticos, los deterministas son computacionalmente eficientes, pero tienden a converger a mínimos locales. A pesar de esta desventaja, estos últimos algoritmos son el punto de partida en diversas aplicaciones y también para el diseño de procedimientos más robustos y eficientes. En particular para estos métodos, el gradiente se combina con información de la búsqueda de línea o bien métodos que utilicen estimadores Newton, cuasi-Newton o Fisher que permitan actualizar los parámetros del modelo [22].

Una de las mayores dificultades computacionales en estos algoritmos es el cálculo de los gradientes de la función objetivo. Por lo tanto, generar métodos eficientes para el cálculo de los gradientes o para el análisis de la sensibilidad paramétrica de modelos diferenciales-algebraicos es muy importante, no solamente para la determinación de parámetros, sino también para otras áreas de aplicación tales como la simplificación de modelos, asimilación de datos, control óptimo, entre otros.

Hoy en día el disparo múltiple, introducido décadas atrás [14, 15, 16, 17], es un méto-

do bastante conocido. Desde su implementación ha sido estudiado y aplicado con mayor frecuencia cuando se utilizan datos con algún nivel de ruido, mediciones experimentales para sistemas de ecuaciones diferenciales ordinarias [23, 18, 6, 1], parciales [20, 21, 24, 25], así como también para resolver problemas de control óptimo [26, 27, 28, 29]. Dos de las características principales del método de disparo múltiple son la estabilidad y la convergencia mejorada, las cuáles se obtienen de la partición de la función objetivo en conjunto con el número de subintervalos en el intervalo de tiempo, de modo que la minimización de la función objetivo es muy eficiente, ya que mitiga el posible estancamiento en los mínimos locales. Estas ventajas mencionadas pueden ser empleadas para diseñar procesos de optimización mucho más rápidos y robustos.

El método de disparo múltiple ha mostrado ser muy efectivo en la estimación de parámetros y en los problemas de control óptimo, cuando se utiliza programación no lineal con métodos de descenso basados en direcciones del gradiente. Por lo general se utilizan métodos de Gauss-Newton [7, 18, 1], programación secuencial cuadrática (SQP) [19, 20, 21] en conjunto con los métodos Newton o cuasi-Newton. Esta clase de métodos requieren que se calculen o estimen las derivadas de las variables de estado respecto a los parámetros, para encontrar el gradiente de la función objetivo. Además, los métodos iterativos anteriores requieren de la solución de sistemas de ecuaciones lineales en cada iteración, por lo que su costo computacional es elevado.

Por otro lado, el método de la ecuación adjunta, en conjunto con el método de disparo múltiple rara vez se usa para la estimación de parámetros en sistemas de ecuaciones diferenciales [1], principalmente porque ya existen métodos de optimización como los mencionados anteriormente. Otra de las razones es que la ecuación adjunta introduce discontinuidades asociadas a los datos con ruido, que aparecen como saltos puntuales en la ecuación diferencial. En este trabajo se propone suavizar esos términos discontinuos mediante el ajuste lineal de los datos a curvas suaves.

En el presente trabajo, se mostrará que el método de la ecuación adjunta puede ser implementado de manera muy eficiente. Se utiliza el cálculo de variaciones y análisis de perturbaciones para el cálculo del gradiente de la función objetivo. Esta estrategia permite métodos iterativos simples cuasi-Newton, como BFGS, evitando la solución de sistemas de ecuaciones lineales en cada iteración. Primeramente, se explicará el método de disparo simple y se extenderá al método de disparo múltiple. Este enfoque da como resultado un algoritmo simple y efectivo, el cual se usará en dos modelos diferentes. El modelo epidemiológico SEIRD, y el modelo de la reacción enzimática inducible. Las aproximaciones variacionales tradicionales que usan modelos cuadráticos penalizados son muy sensibles a las condiciones iniciales y suelen estancarse en regiones con mínimos locales [3]. Se mostrará que el método de disparo múltiple, en conjunto con un preprocesamiento de datos (con ruido) adecuado supera los obstáculos mencionados y además ofrece buenos resultados.

Capítulo 2

Modelo cuadrático, cálculo variacional y disparo simple

2.1. Modelo cuadrático para la determinación de parámetros a partir de datos experimentales

Sea $\mathbf{x}(t) \in \mathbb{R}^d$ una variable de estado en el tiempo $t \in I = [t_0, t_f]$ de una ecuación diferencial ordinaria que satisface el problema de valor inicial siguiente:

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{f}(t, \mathbf{x}(t), \boldsymbol{\theta}), \quad \mathbf{x}(t_0) = \mathbf{s}_0 \quad t_0 < t \leq t_f \quad (2.1)$$

donde $\mathbf{f} : \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^{np} \rightarrow \mathbb{R}^d$ puede no depender de t , pero claramente depende de $\mathbf{x}(t)$ y del vector de parámetros $\boldsymbol{\theta} \in \mathbb{R}^{np}$, donde np es el número de parámetros. Para simplificar la notación, haremos $\mathbf{x}(t) = \mathbf{x}(t; \mathbf{s}_0, \boldsymbol{\theta})$ cuando sea conveniente.

En muchos modelos, no todas las variables de estado son observables, por lo que se denotará con $\bar{\mathbf{x}}$ a las variables observables, y $\underline{\mathbf{x}}$ a las variables no observables. Supóngase, además, que se tienen datos experimentales con ruido incluido en los tiempos, $t_0 \leq t_1 < t_2 < \dots < t_m \leq t_f$, de modo que:

$$\bar{\mathbf{x}}_i = \bar{\mathbf{x}}(t_i; \mathbf{s}_0, \boldsymbol{\theta}) + \bar{\epsilon}_i, \quad i = 1, 2, \dots, m \quad (2.2)$$

donde $\bar{\epsilon} \in \mathbb{R}^d$ son vectores aleatorios independientes y representan mediciones del error con una distribución Gaussiana multivariada con media cero y vector de varianzas $\bar{\sigma}_i^2 \in \mathbb{R}^d$. En ocasiones también puede utilizarse otra distribución de probabilidad, por ejemplo, la distribución uniforme, entre otras.

Un modelo, introducido en [3], para estimar los parámetros de (2.1) consiste en identificar

$\mathbf{s}_0$ y $\boldsymbol{\theta}$ a través del problema de minimización de la función objetivo siguiente:

$$J(\mathbf{s}_0, \boldsymbol{\theta}) = \frac{1}{2} \left\| \frac{\mathbf{x}_0 - \mathbf{s}_0}{\boldsymbol{\sigma}_0} \right\|_{\mathbb{R}^d}^2 + \frac{1}{2} \sum_{i=1}^m \left\| \frac{\bar{\mathbf{x}}(t_i; \mathbf{s}_0, \boldsymbol{\theta}) - \bar{\mathbf{x}}_i}{\bar{\boldsymbol{\sigma}}_i} \right\|_{\mathbb{R}^{no}}^2 \quad (2.3)$$

sujeta a la restricción de que la variable de estado $\mathbf{x}(t) = \mathbf{x}(t, \mathbf{s}_0, \boldsymbol{\theta})$ satisfaga la EDO (2.1). La primera norma es la norma euclidiana en $\mathbb{R}^d$ y las normas dentro de la suma son las normas euclidianas en $\mathbb{R}^{no}$, donde no es el número de variables observables. El vector $\mathbf{s}_0$ denota una medida experimental de las condiciones iniciales.

Observación 1. *El cociente en (2.3) es calculado componente a componente. en particular: $\left\| \frac{\bar{\mathbf{x}}(t_i; \mathbf{s}_0, \boldsymbol{\theta}) - \bar{\mathbf{x}}_i}{\bar{\boldsymbol{\sigma}}_i} \right\|_{\mathbb{R}^{no}}^2 = (\bar{\mathbf{x}}(t_i; \mathbf{s}_0, \boldsymbol{\theta}) - \bar{\mathbf{x}}_i)^T \bar{W}_i (\bar{\mathbf{x}}(t_i; \mathbf{s}_0, \boldsymbol{\theta}) - \bar{\mathbf{x}}_i)$ donde $\bar{W}_i = \boldsymbol{\sigma}_i^{-1}$ es la matriz de precisión. Si una matriz simétrica W es positivo-definida, entonces define una norma en el subespacio correspondiente de variables observable como $\|\bar{\mathbf{x}}\|_W = \bar{\mathbf{x}}^T W \bar{\mathbf{x}}$.*

Para ilustrar mejor los conceptos y la notación mostrada anteriormente, considérese el modelo epidemiológico SEIRD. Este es un modelo determinista que describe la propagación de enfermedades infecciosas (como el COVID-19). Sean N la población total constante y $S(t), E(t), I(t), R(t), D(t)$ las variables. El sistema de ecuaciones de dicho modelo es:

$$\begin{aligned} \frac{dS}{dt} &= -\frac{\alpha}{N} SI \\ \frac{dE}{dt} &= \frac{\alpha}{N} SI - \beta E \\ \frac{dI}{dt} &= \beta E - \frac{1}{T_I} I \\ \frac{dR}{dt} &= \frac{1-f}{T_I} I \\ \frac{dD}{dt} &= \frac{f}{T_I} I \end{aligned} \quad (2.4)$$

donde las variables a considerar son:

- $S(t)$: Población susceptible.
- $E(t)$: Población expuesta.
- $I(t)$: Población infectada.
- $R(t)$: Población recuperada.
- $D(t)$: Población fallecida.

mientras que los parámetros son:

- α es la tasa de infección.
- β es la tasa de incubación.
- T_I es el promedio de periodo de infección.
- f es la fracción de individuos que mueren.

Este sistema tiene sus condiciones iniciales respectivas $S(t_0), E(t_0), I(t_0), R(t_0), D(t_0)$. Aquí, la dimensión del sistema es $d = 5$, el vector de variables de estado es $\mathbf{x}(t) = (S(t), E(t), I(t), R(t), D(t))^T$, el número de parámetros es $np = 4$ y $\boldsymbol{\theta} = (\alpha, \beta, T_I, f)^T \approx (\alpha, \beta, \gamma, \mu)^T$ con $\gamma = 1/T_I$ y $\mu = f/T_I$. Para los resultados que se mostrarán más adelante, se usarán las soluciones de $\mathbf{x}(t)$ obtenidas con los métodos de Runge-Kutta de ordenes 2 y 4 en el intervalo de tiempo $[t_0, t_f] = [40, 80]$, respectivamente, con los parámetros exactos $\boldsymbol{\theta} = (1, 1/7, 1/5, 1/70)^T$, con condición inicial $\mathbf{x}_0 = (91647, 4853, 1755, 1620, 125)^T$ y población total de $N = 10^5$. Las mediciones experimentales se considerarán en los tiempos $t_i = 40 + i$, $1 \leq i \leq 14$ ($m = 14$). Luego, las variables en este modelo serán $\bar{\mathbf{x}}_i = (I_i, R_i, D_i)$ en los tiempos t_i , de modo que las variables no observables serán $\underline{\mathbf{x}}_i = (S_i, E_i)$.

Suponiendo que la población total se mantiene constante, o bien que su cambio es imperceptible en el transcurso de tiempo dado, el sistema cumple lo siguiente:

$$S(t) + E(t) + I(t) + R(t) + D(t) = N, \quad \forall t \in [t_0, t_f] \quad (2.5)$$

de modo que (2.4)-(2.5) describen un sistema diferencial-algebraico. De esta manera, se modifica el modelo (2.3), añadiendo el término penalizado, al modelo extendido:

$$J(\mathbf{s}_0, \boldsymbol{\theta}) = \frac{1}{2} \left\| \frac{\mathbf{x}_0 - \mathbf{s}_0}{\boldsymbol{\sigma}_0} \right\|_{\mathbb{R}^5}^2 + \frac{k}{2} (\mathbf{x}_0 \cdot \mathbf{1} - N)^2 + \frac{1}{2} \sum_{i=1}^m \left\| \frac{\bar{\mathbf{x}}(t_i; \mathbf{s}_0, \boldsymbol{\theta}) - \bar{\mathbf{x}}_i}{\bar{\boldsymbol{\sigma}}_i} \right\|_{\mathbb{R}^3}^2 \quad (2.6)$$

donde el término de penalización añadido ayuda a estabilizar el proceso de optimización para estimar las condiciones iniciales. El vector constante $\mathbf{1} \in \mathbb{R}^d$ tiene componentes igual a uno, de modo forma que el producto punto $\mathbf{x}_0 \cdot \mathbf{1}$ es igual a la suma de las componentes de $\mathbf{x}_0$.

2.2. Enfoque Variacional

En este trabajo, el interés es encontrar la condición inicial $\mathbf{s}_0 \in \mathbb{R}^d$ y el vector de parámetros $\boldsymbol{\theta} \in \mathbb{R}^{np}$ que minimicen la función objetivo (2.6) sujeta a la restricción (2.1) y otras

posibles como la condición algebraica en el modelo SEIRD, $N = S(t) + E(t) + I(t) + R(t) + D(t)$ (véase capítulo 4 para más información), suponiendo que se cuenta con un conjunto de datos experimentales (2.2) en diferentes tiempos. Este modelo en particular es determinista y todas las variables y funciones que se presenten más adelante son continuas y suaves, de modo que se podrán emplear métodos de descenso de gradiente o algoritmos cuasi-Newton. Sin embargo, hay que resaltar que la tarea más costosa y delicada al momento de aplicar estos métodos, es el cálculo del gradiente, o bien las matrices Jacobianas de la función objetivo en cada iteración. Se tienen opciones sobre cómo calcular estas cantidades, tal y como se muestra en [4, 5, 11]. Uno de los objetivos del presente trabajo consiste en la aproximación de las derivadas mediante cálculo variacional y el método de la ecuación adjunta.

Sea $\mathbf{p} = (\mathbf{s}_0, \boldsymbol{\theta})^T \in \mathbb{R}^{d+np}$ el vector con las condiciones iniciales $\mathbf{s}_0$ y el vector de parámetros desconocidos $\boldsymbol{\theta}$ del sistema dinámico (2.1), entonces se tiene la aproximación a primer orden:

$$\nabla J(\mathbf{p}) \cdot \delta \mathbf{p} \approx J(\mathbf{p} + \delta \mathbf{p}) - J(\mathbf{p}) \quad (2.7)$$

donde $\delta \mathbf{p} = (\delta \mathbf{s}_0, \delta \boldsymbol{\theta})^T$ es un pequeño incremento de $\mathbf{p}$. La expresión anterior en forma detallada se puede escribir como:

$$J(\mathbf{s}_0 + \delta \mathbf{s}_0, \boldsymbol{\theta} + \delta \boldsymbol{\theta}) - J(\mathbf{s}_0, \boldsymbol{\theta}) \approx \nabla_{\mathbf{s}_0} J(\mathbf{s}_0, \boldsymbol{\theta}) \cdot \delta \mathbf{s}_0 + \nabla_{\boldsymbol{\theta}} J(\mathbf{s}_0, \boldsymbol{\theta}) \cdot \delta \boldsymbol{\theta} \quad (2.8)$$

donde $\nabla_{\mathbf{s}_0}$ y $\nabla_{\boldsymbol{\theta}}$ son los gradientes respecto $\mathbf{s}_0$ y $\boldsymbol{\theta}$. Evaluando directamente $J(\mathbf{s}_0 + \delta \mathbf{s}_0, \boldsymbol{\theta} + \delta \boldsymbol{\theta})$ de (2.6) y simplificando, se obtiene

$$\begin{aligned} J(\mathbf{s}_0 + \delta \mathbf{s}_0, \boldsymbol{\theta} + \delta \boldsymbol{\theta}) - J(\mathbf{s}_0, \boldsymbol{\theta}) &\approx \frac{\mathbf{x}_0 - \mathbf{s}_0}{\sigma_0^2} \cdot \delta \mathbf{s}_0 + k(\mathbf{x}_0 \cdot \mathbf{1} - N)\mathbf{1} \cdot \delta \mathbf{s}_0 \\ &+ \sum_{i=1}^m \frac{\bar{\mathbf{x}}(t_i; \mathbf{s}_0, \boldsymbol{\theta}) - \bar{\mathbf{x}}_i}{\bar{\sigma}_i^2} \cdot \left(\frac{\partial \bar{\mathbf{x}}(t_i; \mathbf{s}_0, \boldsymbol{\theta})}{\partial \mathbf{s}_0} \delta \mathbf{s}_0 + \frac{\partial \bar{\mathbf{x}}(t_i; \mathbf{s}_0, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \delta \boldsymbol{\theta} \right) \end{aligned} \quad (2.9)$$

Nótese que $\frac{\partial \bar{\mathbf{x}}(t_i; \mathbf{s}_0, \boldsymbol{\theta})}{\partial \mathbf{s}_0}$ y $\frac{\partial \bar{\mathbf{x}}(t_i; \mathbf{s}_0, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$ son las matrices Jacobianas de las variables observables $\bar{\mathbf{x}}(t_i; \mathbf{x}_0, \boldsymbol{\theta})$ con respecto a $\mathbf{s}_0$ y $\boldsymbol{\theta}$, respectivamente. Estas matrices de tamaño $no \times d$ y $no \times np$, respectivamente, de modo que el número de derivadas parciales a calcular es de $m \times no(d + np)$. Dichas matrices son conocidas como las sensibilidades de las variables observables. Por ejemplo, si en el modelo SEIRD se considerara que el número de variables observables es $no = 3$, y que hay $m = 13$ mediciones experimentales, se deberían calcular $13 \times 3(5 + 4) = 351$ sensibilidades en cada iteración de un gradiente o método cuasi-Newton. El método más básico para calcular estas derivadas es el de diferencias finitas, sin embargo, tiene algunas desventajas. Tal y como se estableció anteriormente, el enfoque

es solamente en los métodos variacionales, los cuáles pueden ser mucho más sofisticados, pero en general, son más eficientes.

2.3. La variación respecto de los parámetros y la ecuación adjunta

La variación total de $\mathbf{x}(t; \mathbf{s}_0, \boldsymbol{\theta})$ respecto $\mathbf{s}_0$ y $\boldsymbol{\theta}$ está dada por

$$\delta \mathbf{x}(t; \mathbf{s}_0, \boldsymbol{\theta}) = \mathbf{x}(t; \mathbf{s}_0 + \delta \mathbf{s}_0, \boldsymbol{\theta} + \delta \boldsymbol{\theta}) - \mathbf{x}(t; \mathbf{s}_0, \boldsymbol{\theta}) = \frac{\partial \mathbf{x}(t; \mathbf{s}_0, \boldsymbol{\theta})}{\partial \mathbf{s}_0} \delta \mathbf{s}_0 + \frac{\partial \mathbf{x}(t; \mathbf{s}_0, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \delta \boldsymbol{\theta} \quad (2.10)$$

Dado que los cálculos de las matrices Jacobianas son complicados y costosos, se propondrá un método para evitar este cálculo, el cuál consiste en resolver solamente las ecuaciones de estado y la *ecuación adjunta*.

Antes de proceder con dicho cálculo, se introducirá la siguiente definición:

Definición 1. Sea V un espacio lineal (o vectorial) con un producto escalar o interior $\langle \cdot, \cdot \rangle$ con correspondiente norma $\| \cdot \|$, entonces V será un espacio de Hilbert si V es completo. Equivalentemente, V será un espacio de Hilbert si cada sucesión de Cauchy con respecto a $\| \cdot \|$ es convergente.

Ahora, introducimos los siguientes espacios de Hilbert:

$$\begin{aligned} V &= L_2([t_0, t_f]; \mathbb{R}^d) = \left\{ \mathbf{v} : [t_0, t_f] \rightarrow \mathbb{R}^d \mid \int_{t_0}^{t_f} \|\mathbf{v}(t)\|^2 dt < \infty \right\} \\ H &= H^1([t_0, t_f]; \mathbb{R}^d) = \left\{ \mathbf{v} \in L_2([t_0, t_f]; \mathbb{R}^d) \mid \frac{\partial \mathbf{v}}{\partial t} \in L_2([t_0, t_f]) \right\} \end{aligned}$$

donde el espacio V tiene producto interior o escalar $\langle \mathbf{v}, \mathbf{w} \rangle = \int_{t_0}^{t_f} \mathbf{v}(t) \cdot \mathbf{w}(t) dt$ con norma inducida $\|\mathbf{v}(t)\|^2 = \langle \mathbf{v}, \mathbf{v} \rangle$.

Si $\mathbf{x}(t) = \mathbf{x}(t; \mathbf{s}_0, \boldsymbol{\theta}) \in H$ satisface la ecuación (2.1), entonces el siguiente producto interior es nulo

$$\int_{t_0}^{t_f} \left(\frac{d\mathbf{x}(t)}{dt} - \mathbf{f}(t, \mathbf{x}(t), \boldsymbol{\theta}) \right) \cdot \mathbf{y}(t) dt = 0, \quad \forall \mathbf{y} \in H$$

Por simplicidad, hagamos $\dot{\mathbf{x}}(t) = \frac{d\mathbf{x}(t)}{dt}$. Diferenciando la expresión anterior, se obtiene:

$$\int_{t_0}^{t_f} \delta \dot{\mathbf{x}}(t) \cdot \mathbf{y}(t) dt = \int_{t_0}^{t_f} [\mathbf{f}_{\mathbf{x}}(\mathbf{x}(t), \boldsymbol{\theta}) \delta \mathbf{x}(t) + \mathbf{f}_{\mathbf{s}_0}(\mathbf{x}(t), \boldsymbol{\theta}) \delta \mathbf{s}_0(t) + \mathbf{f}_{\boldsymbol{\theta}}(\mathbf{x}(t), \boldsymbol{\theta}) \delta \boldsymbol{\theta}] \cdot \mathbf{y}(t) dt$$

donde $\mathbf{f}_{\mathbf{x}}$ y $\mathbf{f}_{\boldsymbol{\theta}}$ son las matrices Jacobianas de $\mathbf{f}$ respecto $\mathbf{x}$ y $\boldsymbol{\theta}$, respectivamente. Ahora, integrando por partes, y observando que $[\mathbf{f}_{\mathbf{x}} \delta \mathbf{x}] \cdot \mathbf{y} = [\mathbf{f}_{\mathbf{x}}^T \mathbf{y}] \cdot \delta \mathbf{x}$, $[\mathbf{f}_{\boldsymbol{\theta}} \delta \boldsymbol{\theta}] \cdot \mathbf{y} = [\mathbf{f}_{\boldsymbol{\theta}}^T \mathbf{y}] \cdot \delta \boldsymbol{\theta}$, y $[\mathbf{f}_{\mathbf{s}_0} \delta \mathbf{s}_0] \cdot \mathbf{y} = [\mathbf{f}_{\mathbf{s}_0}^T \mathbf{y}] \cdot \delta \mathbf{s}_0$, se obtiene

$$\begin{aligned} \mathbf{y}(t) \cdot \delta \mathbf{x}(t) \Big|_{t_0}^{t_f} - \int_{t_0}^{t_f} \dot{\mathbf{y}}(t) \cdot \delta \mathbf{x}(t) dt &= \int_{t_0}^{t_f} \{ [\mathbf{f}_{\mathbf{x}}^T \mathbf{y}](t) \cdot \delta \mathbf{x}(t) + [\mathbf{f}_{\mathbf{s}_0}^T \mathbf{y}](t) \cdot \delta \mathbf{s}_0(t) \\ &+ [\mathbf{f}_{\boldsymbol{\theta}}^T \mathbf{y}](t) \cdot \delta \boldsymbol{\theta} \} dt. \end{aligned}$$

Evaluando y asumiendo que $\mathbf{f}$ no depende explícitamente de $\mathbf{s}_0$, pues $\delta \mathbf{x}(t_0) = \delta \mathbf{s}_0$, obtenemos

$$\mathbf{y}(t_f) \cdot \delta \mathbf{x}(t_f) - \mathbf{y}(t_0) \cdot \delta \mathbf{x}(t_0) - \int_{t_0}^{t_f} \{ \dot{\mathbf{y}}(t) + [\mathbf{f}_{\mathbf{x}}^T \mathbf{y}](t) \} \cdot \delta \mathbf{x}(t) dt = \int_{t_0}^{t_f} [\mathbf{f}_{\boldsymbol{\theta}}^T \mathbf{y}](t) \cdot \delta \boldsymbol{\theta} dt \quad (2.11)$$

donde además $\delta \mathbf{x}$ está determinado por (2.9) y (2.10). Luego, una forma de calcular explícitamente (2.10) es estableciendo una relación con (2.11), introduciendo la ecuación adjunta siguiente:

$$\begin{aligned} -\dot{\mathbf{y}}(t) &= \mathbf{f}_{\mathbf{x}}(\mathbf{x}(t), \boldsymbol{\theta})^T \mathbf{y}(t) + \sum_{i=1}^m \frac{(\bar{\mathbf{x}}(t_i; \mathbf{s}_0, \boldsymbol{\theta}) - \bar{\mathbf{x}}_i)}{\bar{\sigma}_i^2} \delta_D(t - t_i), \quad t_f > t \geq t_0 \\ \mathbf{y}(t_f) &= 0 \end{aligned} \quad (2.12)$$

donde δ_D es llamada la *medida de Dirac* centrada 0, la cual induce una clase especial de distribución que satisface:

$$\int_{-\infty}^{\infty} \phi(t) \delta(t - a) dt = \phi(a)$$

La ecuación adjunta (2.12) es una ecuación diferencial con retroceso en el tiempo, y contiene toda la información acerca de los datos experimentales $\{\bar{\mathbf{x}}_i\}_{i=1}^m$, y su formulación variacional se obtiene multiplicando por una función de prueba $\phi(t)$ y luego se integra:

$$-\int_{t_0}^{t_f} \{\dot{\mathbf{y}}(t) + [\mathbf{f}_{\mathbf{x}}^T \mathbf{y}](t)\} \cdot \boldsymbol{\phi}(t) dt = \int_{t_0}^{t_f} \sum_{i=1}^m \frac{(\bar{\mathbf{x}}(t_i; \mathbf{s}_0, \boldsymbol{\theta}) - \bar{\mathbf{x}}_i)}{\bar{\boldsymbol{\sigma}}_i^2} \delta_D(t - t_i) \cdot \boldsymbol{\phi}(t) dt \quad (2.13)$$

Ahora, la integral de lado derecho es igual a $\sum_{i=1}^m \frac{(\bar{\mathbf{x}}(t_i; \mathbf{s}_0, \boldsymbol{\theta}) - \bar{\mathbf{x}}_i)}{\bar{\boldsymbol{\sigma}}_i^2} \cdot \boldsymbol{\phi}(t_i)$. Si se utiliza $\boldsymbol{\phi}(t) = \delta \mathbf{x}(t)$ en (2.13) y se sustituye dentro de (2.11) se obtiene lo siguiente:

$$-\mathbf{y}(t_0) \cdot \delta \mathbf{s}_0 + \sum_{i=1}^m \frac{(\bar{\mathbf{x}}(t_i; \mathbf{s}_0, \boldsymbol{\theta}) - \bar{\mathbf{x}}_i)}{\bar{\boldsymbol{\sigma}}_i^2} \cdot \delta \mathbf{x}(t_i) = \int_{t_0}^{t_f} [\mathbf{f}_{\boldsymbol{\theta}}^T \mathbf{y}](t) \cdot \delta \boldsymbol{\theta} dt \quad (2.14)$$

donde $\mathbf{y}(t)$ es la solución de la ecuación adjunta. La sustitución de (2.14) dentro de (2.9) proporciona

$$\begin{aligned} & \nabla_{\mathbf{s}_0} J(\mathbf{s}_0, \boldsymbol{\theta}) \cdot \delta \mathbf{s}_0 + \nabla_{\boldsymbol{\theta}} J(\mathbf{s}_0, \boldsymbol{\theta}) \cdot \delta \boldsymbol{\theta} \\ & \approx \frac{\mathbf{x}_0 - \mathbf{s}_0}{\boldsymbol{\sigma}_0^2} \cdot \delta \mathbf{s}_0 + k(\mathbf{s}_0 \cdot \mathbf{1} - N)\mathbf{1} \cdot \delta \mathbf{s}_0 + \mathbf{y}(t_0) \cdot \delta \mathbf{s}_0 + \left[\int_{t_0}^{t_f} [\mathbf{f}_{\boldsymbol{\theta}}^T \mathbf{y}](t) dt \right] \cdot \delta \boldsymbol{\theta} \end{aligned} \quad (2.15)$$

De modo que el gradiente de la función objetivo (2.6) es $\nabla J(\mathbf{p}) = (\nabla_{\mathbf{s}_0} J(\mathbf{s}_0, \boldsymbol{\theta}), \nabla_{\boldsymbol{\theta}} J(\mathbf{s}_0, \boldsymbol{\theta}))^T$ con

$$\nabla_{\mathbf{s}_0} J(\mathbf{s}_0, \boldsymbol{\theta}) = \frac{\mathbf{x}_0 - \mathbf{s}_0}{\boldsymbol{\sigma}_0^2} + k(\mathbf{x}_0 \cdot \mathbf{1} - N)\mathbf{1} + \mathbf{y}(t_0) \quad (2.16)$$

$$\nabla_{\boldsymbol{\theta}} J(\mathbf{s}_0, \boldsymbol{\theta}) = \int_{t_0}^{t_f} \mathbf{f}_{\boldsymbol{\theta}}(\mathbf{x}(t), \boldsymbol{\theta})^T \mathbf{y}(t) dt \quad (2.17)$$

donde $\mathbf{x}(t)$ resuelve la ecuación de estado $\frac{d\mathbf{x}(t)}{dt} = \mathbf{f}(t, \mathbf{x}(t), \boldsymbol{\theta})$, $\mathbf{x}(t_0) = \mathbf{s}_0$ y $\mathbf{y}(t)$ es la solución de la ecuación adjunta (2.12).

2.3.1. La función Lagrangiana y la derivada de $J(\mathbf{s}_0, \boldsymbol{\theta})$

La solución obtenida de la ecuación adjunta es el multiplicador de Lagrange del problema de optimización (2.6) con restricción (2.1). Para demostrar esta afirmación, primero se escribe la función Lagrangiana asociada:

$$\mathcal{L}(\mathbf{p}, \mathbf{y}) = J(\mathbf{p}) + \int_{t_0}^{t_f} \mathbf{y}(t)^T \cdot \{\mathbf{f}(\mathbf{x}(t), \boldsymbol{\theta}) - \dot{\mathbf{x}}(t)\} dt \quad (2.18)$$

y recordando que $\mathbf{p} = (\mathbf{s}_0, \boldsymbol{\theta})^T$, calculando ahora la derivada respecto $\boldsymbol{\theta}$ utilizando en el proceso integración por partes en la última integral, se obtiene:

$$\begin{aligned}
\frac{\partial \mathcal{L}(\mathbf{s}_0, \boldsymbol{\theta}, y)}{\partial \boldsymbol{\theta}} &= \frac{\partial}{\partial \boldsymbol{\theta}} \left\{ J(\mathbf{s}_0, \boldsymbol{\theta}) + \int_{t_0}^{t_f} \mathbf{y}(t)^T [\mathbf{f}(\mathbf{x}(t), \boldsymbol{\theta})] - \int_{t_0}^{t_f} \mathbf{y}(t)^T [\dot{\mathbf{x}}(t)] dt \right\} \\
&= \frac{\partial J(\mathbf{s}_0, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} + \int_{t_0}^{t_f} \mathbf{y}^T \left[\mathbf{f}_x \frac{\partial \mathbf{x}}{\partial \boldsymbol{\theta}} + \mathbf{f}_\theta \right] dt - \mathbf{y}^T \frac{\partial \mathbf{x}}{\partial \boldsymbol{\theta}} \Big|_{t_0}^{t_f} + \int_{t_0}^{t_f} \dot{\mathbf{y}}(t)^T \frac{\partial \mathbf{x}}{\partial \boldsymbol{\theta}} dt \\
&= \sum_{i=1}^m \left(\frac{(\bar{\mathbf{x}}(t_i; \mathbf{s}_0, \boldsymbol{\theta}) - \bar{\mathbf{x}}_i)}{\bar{\sigma}_i^2} \right)^T \frac{\partial \bar{\mathbf{x}}}{\partial \boldsymbol{\theta}} - \mathbf{y}^T \frac{\partial \mathbf{x}}{\partial \boldsymbol{\theta}} \Big|_{t_0}^{t_f} + \int_{t_0}^{t_f} (\dot{\mathbf{y}} + \mathbf{f}_x^T \mathbf{y})^T \frac{\partial \mathbf{x}}{\partial \boldsymbol{\theta}} dt \\
&\quad + \int_{t_0}^{t_f} \mathbf{y}^T \mathbf{f}_\theta dt
\end{aligned}$$

El término de la suma se obtiene de (2.6) y puede ser expresado como

$$\sum_{i=1}^m \left(\frac{(\bar{\mathbf{x}}(t_i; \mathbf{s}_0, \boldsymbol{\theta}) - \bar{\mathbf{x}}_i)}{\bar{\sigma}_i^2} \right)^T \frac{\partial \bar{\mathbf{x}}}{\partial \boldsymbol{\theta}} = \int_{t_0}^{t_f} \sum_{i=1}^m \left(\frac{(\bar{\mathbf{x}}(t_i; \mathbf{s}_0, \boldsymbol{\theta}) - \bar{\mathbf{x}}_i)}{\bar{\sigma}_i^2} \right)^T \frac{\partial \bar{\mathbf{x}}}{\partial \boldsymbol{\theta}} \delta_D(t - t_i) dt$$

y debido a que la función $\mathbf{y}(t)$ satisface la ecuación adjunta (2.12) se obtiene

$$\frac{\partial \mathcal{L}}{\partial \boldsymbol{\theta}} = -\mathbf{y}^T \frac{\partial \mathbf{x}}{\partial \boldsymbol{\theta}} \Big|_{t_0}^{t_f} + \int_{t_0}^{t_f} \mathbf{y}^T \mathbf{f}_\theta dt$$

en donde el término de frontera $\mathbf{y}^T \frac{\partial \mathbf{x}}{\partial \boldsymbol{\theta}} \Big|_{t_0}^{t_f}$ se anula, dado que la condición $\mathbf{y}(t_f) = 0$ y $\frac{\partial \mathbf{s}(t_0)}{\partial \boldsymbol{\theta}} = \mathbf{0}$ (matriz nula), pues $\mathbf{s}_0$ no depende de $\boldsymbol{\theta}$. De modo que

$$\frac{\partial \mathcal{L}}{\partial \boldsymbol{\theta}} = \int_{t_0}^{t_f} \mathbf{y}^T \mathbf{f}_\theta dt$$

Y así se obtiene

$$\nabla_{\boldsymbol{\theta}} \mathcal{L}(\mathbf{s}_0, \boldsymbol{\theta}) = \int_{t_0}^{t_f} \mathbf{f}_\theta(\mathbf{x}(t), \boldsymbol{\theta})^T \mathbf{y}(t) dt \quad (2.19)$$

En forma análoga, se calcula fácilmente el gradiente del lagrangiano respecto a $\mathbf{s}_0$, obteniendo:

$$\nabla_{\mathbf{s}_0} \mathcal{L}(\mathbf{s}_0, \boldsymbol{\theta}) = \frac{\mathbf{x}_0 - \mathbf{s}_0}{\sigma_0^2} + k(\mathbf{x}_0 \cdot \mathbf{1} - N)\mathbf{1} + \mathbf{y}(t_0) \quad (2.20)$$

El gradiente obtenido del lagrangiano $\mathcal{L}(\mathbf{p}, \mathbf{y})$ es entonces:

$$\nabla_{\mathbf{p}} \mathcal{L}(\mathbf{p}, \mathbf{y}) = \nabla J(\mathbf{p}) = \begin{bmatrix} \nabla_{\mathbf{s}_0} J(\mathbf{p}) \\ \nabla_{\boldsymbol{\theta}} J(\mathbf{p}) \end{bmatrix} = \begin{bmatrix} \frac{\mathbf{x}_0 - \mathbf{s}_0}{\sigma_0^2} + k(\mathbf{x}_0 \cdot \mathbf{1} - N)\mathbf{1} + \mathbf{y}(t_0) \\ \int_{t_0}^{t_f} \mathbf{f}_{\boldsymbol{\theta}}(\mathbf{x}(t), \boldsymbol{\theta})^T \mathbf{y}(t) dt \end{bmatrix} \quad (2.21)$$

donde si comparamos las ecuaciones (2.16), (2.17) con los términos de (2.21), se nota que ambos métodos resuelven la ecuación adjunta (2.12).

2.4. Solución de la ecuación adjunta y cálculo del gradiente

Como se estableció anteriormente, la ecuación adjunta (2.12) es un sistema de ecuaciones diferenciales con retroceso en el tiempo. Sin embargo, la mayoría de los métodos numéricos están diseñados para resolver sistemas dinámicos con avance en el tiempo. Para obtener una ecuación equivalente con avance en el tiempo, se aplica un cambio de variable, utilizando la relación simétrica $t_f - \tau = \tau - t_0$, la cual proporciona la relación

$$\mathbf{y}(t) = \mathbf{y}(t_f - (\tau - t_0)) \equiv \mathbf{y}_A(\tau)$$

por lo que

$$\dot{\mathbf{y}}(t) = -\dot{\mathbf{y}}_A(\tau).$$

Realizando las sustituciones necesarias en (2.12) se obtiene el sistema dinámico con avance en el tiempo

$$\dot{\mathbf{y}}_A(\tau) = \mathbf{f}_{\mathbf{x}}(\mathbf{x}_A(\tau), \boldsymbol{\theta})^T \mathbf{y}_A(\tau) + \sum_{i=1}^m \frac{\bar{\mathbf{x}}_A(\tau_i; \mathbf{s}_0, \boldsymbol{\theta}) - \bar{\mathbf{x}}_{m+1-i}}{\sigma_{m+1-i}^2} \delta_D(\tau_i - \tau), \quad t_0 \leq \tau < t_f \quad (2.22)$$

$$\mathbf{y}_A(\tau_0) = 0 \quad (2.23)$$

Este sistema se puede resolver por medio del método de Runge-Kutta. En este trabajo se emplearán los métodos de Runge-Kutta de órdenes 4 y 2.

Dado que el modelo SEIRD considera una población total, dividida en cinco apartados (población susceptible, expuesta, infectada, recuperada y muerta), se sabe que las variables que se pueden observar por las agencias gubernamentales, y organismos de salud, son generalmente $\bar{\mathbf{x}}(\tau) = (I(\tau), R(\tau), D(\tau))$. Si se considera $\mathbf{x} = [S(\tau), E(\tau), I(\tau), R(\tau), D(\tau)] = [x_1, x_2, x_3, x_4, x_5]$ y además $\boldsymbol{\theta} = [\alpha, \beta, \gamma, \mu] = [\theta_1, \theta_2, \theta_3, \theta_4]$, en general, la ecuación adjunta (2.22) se ve de la siguiente forma:

$$\begin{bmatrix} \dot{y}_{A1}(\tau) \\ \dot{y}_{A2}(\tau) \\ \dot{y}_{A3}(\tau) \\ \dot{y}_{A4}(\tau) \\ \dot{y}_{A5}(\tau) \end{bmatrix} = \begin{bmatrix} -\theta_1 x_3 & \theta_1 x_3 & 0 & 0 & 0 \\ 0 & -\theta_2 & \theta_2 & 0 & 0 \\ -\theta_1 x_1 & \theta_1 x_1 & -\theta_3 & \theta_3 - \theta_4 & \theta_4 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} y_{A1}(\tau) \\ y_{A2}(\tau) \\ y_{A3}(\tau) \\ y_{A4}(\tau) \\ y_{A5}(\tau) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \sum_{i=1}^m \frac{\mathbf{x}_3(\tau_i) - x_{3m+1-i}}{\bar{\sigma}_{i3}^2} \delta_D(\tau_i - \tau) \\ \sum_{i=1}^m \frac{\mathbf{x}_4(\tau_i) - x_{4m+1-i}}{\bar{\sigma}_{i4}^2} \delta_D(\tau_i - \tau) \\ \sum_{i=1}^m \frac{\mathbf{x}_5(\tau_i) - x_{5m+1-i}}{\bar{\sigma}_{i5}^2} \delta_D(\tau_i - \tau) \end{bmatrix}$$

o equivalentemente

$$\begin{bmatrix} \dot{y}_{A1}(\tau) \\ \dot{y}_{A2}(\tau) \\ \dot{y}_{A3}(\tau) \\ \dot{y}_{A4}(\tau) \\ \dot{y}_{A5}(\tau) \end{bmatrix} = \begin{bmatrix} -\frac{\alpha I(\tau)}{N} & \frac{\alpha I(\tau)}{N} & 0 & 0 & 0 \\ 0 & -\beta & \beta & 0 & 0 \\ -\frac{\alpha S(\tau)}{N} & -\frac{\alpha S(\tau)}{N} & -\gamma & \gamma - \mu & \mu \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} y_{A1}(\tau) \\ y_{A2}(\tau) \\ y_{A3}(\tau) \\ y_{A4}(\tau) \\ y_{A5}(\tau) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \sum_{i=1}^m \frac{I(\tau_i) - I_{m+1-i}}{\bar{\sigma}_{i3}^2} \delta_D(\tau_i - \tau) \\ \sum_{i=1}^m \frac{R(\tau_i) - R_{m+1-i}}{\bar{\sigma}_{i4}^2} \delta_D(\tau_i - \tau) \\ \sum_{i=1}^m \frac{D(\tau_i) - D_{m+1-i}}{\bar{\sigma}_{i5}^2} \delta_D(\tau_i - \tau) \end{bmatrix}$$

donde $\gamma = \frac{1}{T_1}$, $\mu = \frac{f}{T_1}$ de acuerdo con el inicio del trabajo. Una vez resuelta la ecuación adjunta con avance de tiempo, se encuentra la solución de $\mathbf{y}$ de la ecuación adjunta con retroceso en el tiempo haciendo $\mathbf{y}(t_i) = \mathbf{y}_A(\tau_{m+1-i})$ con $i = 1, 2, \dots, m$.

Para calcular la otra parte del gradiente en la integral (2.19), nótese que

$$\begin{aligned}
\mathbf{f}_\theta(\mathbf{x}(t), \boldsymbol{\theta})^T \mathbf{y}(t) &= \begin{bmatrix} -\frac{S(t)I(t)}{N} & \frac{S(t)I(t)}{N} & 0 & 0 \\ 0 & -E(t) & E(t) & 0 \\ 0 & 0 & -I(t) & I(t) \\ 0 & 0 & 0 & -I(t) \end{bmatrix} \begin{bmatrix} y_1(t) \\ y_2(t) \\ y_3(t) \\ y_4(t) \end{bmatrix} \\
&= \begin{bmatrix} \frac{S(t)I(t)}{N}(y_2(t) - y_1(t)) \\ E(t)(y_3(t) - y_2(t)) \\ I(t)(y_4(t) - y_3(t)) \\ -I(t)(y_4(t)) \end{bmatrix} \equiv \mathbf{H}(t)
\end{aligned}$$

Ahora, usando la regla de Simpson en un conjunto de tiempos determinados se tiene que:

$$\int_{t_0}^{t_f} \mathbf{f}_\theta(\mathbf{x}(t), \boldsymbol{\theta})^T \mathbf{y}(t) dt = \frac{\Delta t}{6} \sum_{j=1}^{N-1} \{ \mathbf{H}(t_{j-1}) + 4\mathbf{H}(t_j) + \mathbf{H}(t_{j+1}) \} \quad (2.24)$$

En algunos casos, cuando el sistema dinámico tiene soluciones suaves, es decir en problemas no rígidos y sin soluciones caóticas, basta con utilizar la regla del trapecio, que es más simple.

2.5. Método de disparo simple

En este método, se suponen desconocidos la condición inicial $\mathbf{s}_0 = \mathbf{x}(t_0)$, el vector de parámetros $\boldsymbol{\theta}$, y además el valor final $\mathbf{s}_1 = \mathbf{x}(t_f)$, los cuáles son determinados a partir de datos con ruido, definidos en (2.2). La Figura 2.1 ilustra la idea del disparo simple.

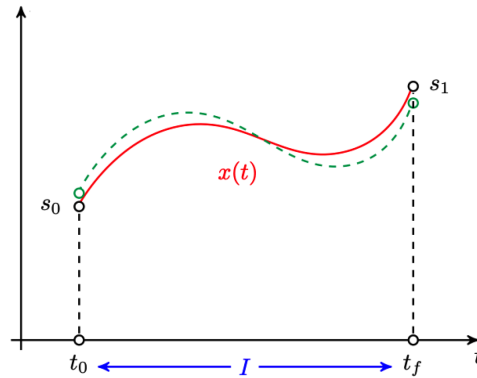


Figura 2.1: Método de disparo simple.

El modelo de optimización en este caso deberá incorporar la nueva restricción asociada a la restricción $\mathbf{s}_1 = \mathbf{x}(t_f)$. Una forma sencilla (partiendo de ideas del control óptimo)

es agregando a la función objetivo un término penalizado, lo cual resulta en el modelo cuadrático siguiente:

$$\min J_k(\mathbf{s}_0, \mathbf{s}_1, \boldsymbol{\theta}) = \frac{1}{2} \sum_{i=1}^m \left\| \frac{\bar{\mathbf{x}}(t_i; \mathbf{s}_0, \boldsymbol{\theta}) - \bar{\mathbf{x}}_i}{\bar{\boldsymbol{\sigma}}_i} \right\|_{\mathbb{R}^{n_o}}^2 + \frac{k}{2} \|\mathbf{x}(t_f; \mathbf{s}_0, \boldsymbol{\theta}) - \mathbf{s}_1\|_{\mathbb{R}^d}^2 \quad (2.25)$$

donde k es un parámetro no negativo que penaliza la discrepancia $\mathbf{x}(t_f) - \mathbf{s}_1$, por lo que es llamado *parámetro de penalización*. La variable de estado $\mathbf{x}(t) = \mathbf{x}(t; \mathbf{s}_0, \boldsymbol{\theta})$ satisface

$$\begin{cases} \dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \boldsymbol{\theta}), & t_0 < t \leq t_f, \\ \mathbf{x}(t_0) = \mathbf{s}_0. \end{cases}$$

2.5.1. El Lagrangiano aumentado y el cálculo del gradiente

Sea $\mathbf{p} = (\mathbf{s}_0, \mathbf{s}_1, \boldsymbol{\theta})$, entonces el Lagrangiano asociado al modelo (2.25) es:

$$\mathcal{L}_k(\mathbf{p}, \mathbf{y}, \boldsymbol{\lambda}) = J_k(\mathbf{p}) + \int_{t_0}^{t_f} \mathbf{y} \cdot (\mathbf{f}(\mathbf{x}, \boldsymbol{\theta}) - \dot{\mathbf{x}}) dt + \boldsymbol{\lambda} \cdot (\mathbf{x}(t_f) - \mathbf{s}_1) \quad (2.26)$$

donde $\mathbf{y} \in H$ es el multiplicador asociado a las restricciones anteriores y $\boldsymbol{\lambda} \in \mathbb{R}^d$ es el multiplicador asociado a la condición de continuidad en el tiempo final, es decir, $\mathbf{x}(t_f) - \mathbf{s}_1 = 0$. Se le llama el Lagrangiano aumentado, debido al término adicional asociado al término de penalización en el modelo (2.25). Luego, la variación de $\mathcal{L}_k$ respecto a $\mathbf{p}$ es:

$$\nabla_{\mathbf{p}} \mathcal{L}_k \cdot \delta \mathbf{p} = \nabla J_k(\mathbf{p}) \cdot \delta \mathbf{p} + \int_{t_0}^{t_f} \mathbf{y} \cdot (\delta \mathbf{f}(\mathbf{x}, \boldsymbol{\theta}) - \delta \dot{\mathbf{x}}) dt + \boldsymbol{\lambda} (\delta \mathbf{x}(t_f) - \delta \mathbf{s}_1) \quad (2.27)$$

Usando regla de la cadena, se tiene que

$$\nabla J_k(\mathbf{p}) \cdot \delta \mathbf{p} = \sum_{i=1}^m \frac{\bar{\mathbf{x}}(t_i) - \bar{\mathbf{x}}_i}{\bar{\boldsymbol{\sigma}}_i^2} \cdot \delta \bar{\mathbf{x}}(t_i) + k(\mathbf{x}(t_f) - \mathbf{s}_1) \cdot (\delta \mathbf{x}(t_f) - \delta \mathbf{s}_1)$$

con $\delta \mathbf{x}(t) = \frac{\partial \mathbf{x}(t; \mathbf{s}_0; \boldsymbol{\theta})}{\partial \mathbf{s}_0} \delta \mathbf{s}_0 + \frac{\partial \mathbf{x}(t; \mathbf{s}_0; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \delta \boldsymbol{\theta}$.

Por otra parte:

$$\int_{t_0}^{t_f} \mathbf{y} \cdot \delta \mathbf{f} dt = \int_{t_0}^{t_f} \mathbf{y} \cdot (\mathbf{f}_{\mathbf{x}} \delta \mathbf{x} + \mathbf{f}_{\boldsymbol{\theta}} \delta \boldsymbol{\theta}) dt = \int_{t_0}^{t_f} \mathbf{f}_{\mathbf{x}}^T \mathbf{y} \cdot \delta \mathbf{x} dt + \int_{t_0}^{t_f} \mathbf{f}_{\boldsymbol{\theta}}^T \mathbf{y} \cdot \delta \boldsymbol{\theta} dt$$

y también:

$$\int_{t_0}^{t_f} \mathbf{y} \cdot \delta \dot{\mathbf{x}} dt = \mathbf{y}(t_f) \cdot \delta \mathbf{x}(t_f) - \mathbf{y}(t_0) \cdot \delta \mathbf{x}_0 - \int_{t_0}^{t_f} \dot{\mathbf{y}} \cdot \delta \mathbf{x} dt$$

sustituyendo los tres valores previos en (2.27) y simplificando términos, se obtiene que

$$\begin{aligned} \nabla_{\mathbf{p}} \mathcal{L}_k \cdot \delta_{\mathbf{p}} &= \sum_{i=1}^m \frac{\bar{\mathbf{x}}(t_i) - \bar{\mathbf{x}}_i}{\bar{\sigma}_i^2} \cdot \delta \bar{\mathbf{x}}(t_i) + \int_{t_0}^{t_f} \mathbf{f}_{\mathbf{x}}^T \mathbf{y} \cdot \delta \mathbf{x} dt + \int_{t_0}^{t_f} \mathbf{f}_{\boldsymbol{\theta}}^T \mathbf{y} \cdot \delta \boldsymbol{\theta} dt \\ &\quad - \mathbf{y}(t_f) \cdot \delta \mathbf{x}(t_f) + \mathbf{y}(t_0) \cdot \delta \mathbf{x}_0 + \int_{t_0}^{t_f} \dot{\mathbf{y}} \cdot \delta \mathbf{x} dt + [\boldsymbol{\lambda} + k(\mathbf{x}(t_f) - \mathbf{s}_1)] \cdot (\delta \mathbf{x}(t_f) - \delta \mathbf{s}_1) \end{aligned}$$

$\forall \mathbf{y} \in H$. Si se toma $\mathbf{y}$ como la solución de la ecuación adjunta con retroceso en el tiempo

$$\begin{cases} -\dot{\mathbf{y}}(t) = \mathbf{f}_{\mathbf{x}}(\mathbf{x}(t), \boldsymbol{\theta})^T \mathbf{y}(t) + \sum_{i=1}^m \frac{\bar{\mathbf{x}}(t_i) - \bar{\mathbf{x}}_i}{\bar{\sigma}_i^2} \delta_D(t - t_i), & t_f > t \geq t_0 \\ \mathbf{y}(t_f) = \boldsymbol{\lambda} + k(\mathbf{x}(t_f) - \mathbf{s}_1) \end{cases}$$

finalmente se obtiene

$$\nabla_{\mathbf{p}} \mathcal{L}_k(\mathbf{p}, \mathbf{y}, \boldsymbol{\lambda}) \cdot \delta_{\mathbf{p}} = \mathbf{y}(t_0) \cdot \delta \mathbf{s}_0 - [\boldsymbol{\lambda} + k(\mathbf{x}(t_f) - \mathbf{s}_1)] \cdot \delta \mathbf{s}_1 + \left(\int_{t_0}^{t_f} \mathbf{f}_{\boldsymbol{\theta}}^T \mathbf{y} dt \right) \cdot \delta \boldsymbol{\theta} \quad (2.28)$$

de modo que si $\mathbf{x}(t)$ resuelve la ecuación de estado $\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \boldsymbol{\theta})$ con $t_0 < t \leq t_f$ y condición $\mathbf{x}(t_0) = \mathbf{s}_0$, y además $\mathbf{y}(t)$ resuelve la ecuación adjunta, entonces, como $\nabla J_k(\mathbf{p}) = \nabla_{\mathbf{p}} \mathcal{L}_k(\mathbf{p}, \mathbf{y}, \boldsymbol{\lambda})$, debido a que la integral y el último término en (2.26) se anulan cuando se satisfacen las restricciones (es decir, cuando $\mathbf{p}$ es factible). Por lo tanto, de acuerdo con (2.28), el gradiente de la función objetivo del modelo cuadrático (2.25) es:

$$\nabla J_k(\mathbf{p}) = \begin{bmatrix} \partial J_k / \partial \mathbf{s}_0 \\ \partial J_k / \partial \mathbf{s}_1 \\ \partial J_k / \partial \boldsymbol{\theta} \end{bmatrix} = \begin{bmatrix} \mathbf{y}(t_0) \\ -\boldsymbol{\lambda} - k(\mathbf{x}(t_f) - \mathbf{s}_1) \\ \int_{t_0}^{t_f} \mathbf{f}_{\boldsymbol{\theta}}(\mathbf{x}(t), \boldsymbol{\theta})^T \mathbf{y}(t) dt \end{bmatrix} \quad (2.29)$$

Capítulo 3

Método de disparo múltiple para la estimación de parámetros a partir de datos con ruido

La inestabilidad y dificultad en la determinación de las condiciones iniciales, así como también el efecto del uso de datos con ruido en intervalos de tiempo muy largos para la determinación de los parámetros, puede reducirse empleando el método de disparo múltiple [1, 6].

Este método es una modificación del método de disparo simple. La idea en general es dividir el subintervalo de tiempo $[t_0, t_f]$ en ns diferentes subintervalos más pequeños, con $I_j = [\tau_j, \tau_{j+1}]$ y $j = 1, 2, \dots, ns$ y además dividir la función objetivo en cada subintervalo. Se adivina una *condición inicial* $\mathbf{s}_j$ desconocida y un *objetivo* $\mathbf{s}_{j+1}$ en cada subintervalo, como se ilustra en la Figura 3.1.

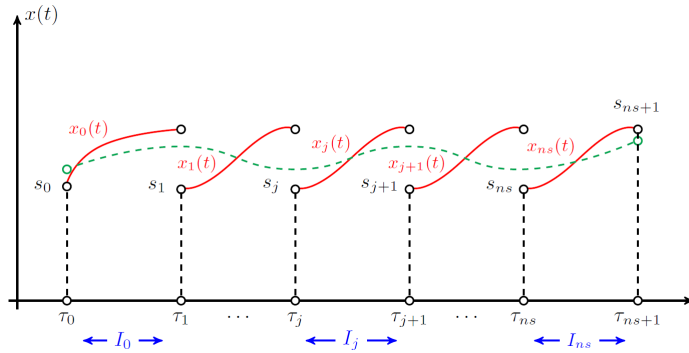


Figura 3.1: Método de disparo múltiple con ns subintervalos de $[t_0, t_f]$.

Con este método, se introducen los *tiempos de disparo* $\tau_1, \tau_2, \dots, \tau_{ns+1}$ y además el vector de parámetros de disparo asociado $\mathbf{p} = (\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_{ns+1}, \boldsymbol{\theta})^T$ en $\mathbb{R}^{d \times (ns+1)np}$. Para fines de este trabajo, se asumirá que cada subintervalo de disparo $I_j = [\tau_j, \tau_{j+1}]$ contiene al menos un tiempo experimental t_i , con t_i no necesariamente igual al tiempo de disparo τ_j . Si $\mathbf{p}_j = (\mathbf{s}_j, \mathbf{s}_{j+1}, \boldsymbol{\theta})$ para cada $j = 1, \dots, ns$, el modelo a considerar es el siguiente:

$$J(\mathbf{p}) = \sum_{j=1}^{ns} J_{k_j}(\mathbf{p}_j) \text{ con } J_{k_j}(\mathbf{p}_j) = \frac{1}{2} \sum_{t_i \in I_j} \left\| \frac{\bar{\mathbf{x}}_j(t_i; \mathbf{s}_j, \boldsymbol{\theta}) - \bar{\mathbf{x}}_i}{\bar{\boldsymbol{\sigma}}_i} \right\|_{\mathbb{R}^{no}}^2 + \frac{k_j}{2} \|\mathbf{x}_j(\tau_{j+1}) - \mathbf{s}_{j+1}\|_{\mathbb{R}^d}^2 \quad (3.1)$$

sujeto a las condiciones siguientes:

$$\begin{cases} \dot{\mathbf{x}}_j(t) = \mathbf{f}(\mathbf{x}_j(t), \boldsymbol{\theta}), & \tau_j < t \leq \tau_{j+1}, \\ \mathbf{x}_j(\tau_j) = \mathbf{s}_j. \end{cases} \quad (3.2)$$

$$\mathbf{x}_j(\tau_{j+1}) = \mathbf{s}_{j+1} \quad (3.3)$$

La restricción (3.2) garantiza que $\mathbf{x}_j$ satisface la ecuación de estado en el intervalo de tiempo local I_j , mientras que el término penalizado en (3.1) impone continuidad entre las soluciones en los intervalos consecutivos, de manera aproximada, penalizando por la posible discrepancia. El último término de (3.1) fuerza la tercera restricción (3.3). Por otra parte, la constante $k_j > 0$ es llamado el parámetro de penalización.

Para los problemas de optimización mostrados anteriormente, sean $\mathbf{y}_j \in (H^1(I_j))^d$ y $\boldsymbol{\lambda}_j \in \mathbb{R}^d$, los multiplicadores de Lagrange asociados a las restricciones (3.2) y (3.3) con $j = 1, \dots, ns$.

3.1. Lagrangiano aumentado

Usando la notación $\mathbf{y} = (\mathbf{y}_1, \dots, \mathbf{y}_{ns})^T$, $\boldsymbol{\lambda} = (\boldsymbol{\lambda}_1, \dots, \boldsymbol{\lambda}_{ns})$, $\mathbf{k} = (\mathbf{k}_1, \dots, \mathbf{k}_{ns})^T$ y $\mathbf{p} = (\mathbf{s}_j, \mathbf{s}_{j+1}, \boldsymbol{\theta})$ el Lagrangiano aumentado es

$$\mathcal{L}_{k_j}(\mathbf{p}, \mathbf{y}, \boldsymbol{\lambda}) = \sum_{i=1}^{ns} \left\{ J_{k_j}(\mathbf{p}_j) + \int_{I_j} \mathbf{y}_j \cdot [\mathbf{f}(\mathbf{x}_j, \boldsymbol{\theta}) - \dot{\mathbf{x}}_j] dt + \boldsymbol{\lambda}_j \cdot [\mathbf{x}_j(\tau_{j+1}) - \mathbf{s}_{j+1}] \right\} \quad (3.4)$$

donde los parámetros de penalización k_j son escalares positivos.

3.2. Análisis de perturbación

Haciendo una perturbación a la expresión anterior, se obtiene que

$$\nabla_{\mathbf{p}_j} \mathcal{L}_{k_j} \cdot \delta \mathbf{p}_j = \sum_{i=1}^{ns} \left\{ \nabla J_{k_j}(\mathbf{p}_j) \cdot \delta \mathbf{p}_j + \int_{I_j} \mathbf{y}_j \cdot [\delta \mathbf{f}(\mathbf{x}_j, \boldsymbol{\theta}) - \delta \dot{\mathbf{x}}_j] dt + \boldsymbol{\lambda}_j \cdot (\delta \mathbf{x}_j(\tau_{j+1}) - \delta \mathbf{s}_{j+1}) \right\} \quad (3.5)$$

donde

$$\nabla J_{k_j}(\mathbf{p}_j) \cdot \delta \mathbf{p}_j = \sum_{t_i \in I_j} \frac{\bar{\mathbf{x}}_j(t_j) - \bar{\mathbf{x}}_i}{\bar{\sigma}_i^2} \cdot \delta \bar{\mathbf{x}}_j(t_i) + k_j [\mathbf{x}_j(\tau_{j+1}) - \mathbf{s}_{j+1}] \cdot (\delta \mathbf{x}_j(\tau_{j+1}) - \delta \mathbf{s}_{j+1})$$

con

$$\delta \bar{\mathbf{x}}_j(t_i) = \frac{\partial \bar{\mathbf{x}}_j(t_i)}{\partial \mathbf{s}_j} \delta \mathbf{s}_j + \frac{\partial \bar{\mathbf{x}}_j(t_i)}{\partial \mathbf{s}_{j+1}} \delta \mathbf{s}_{j+1} + \frac{\partial \bar{\mathbf{x}}_j(t_i)}{\partial \boldsymbol{\theta}} \delta \boldsymbol{\theta}$$

Por otra parte, se tiene que

$$\int_{I_j} \mathbf{y}_j \cdot \delta \mathbf{f} dt = \int_{I_j} \mathbf{y}_j \cdot (\mathbf{f}_x \delta \mathbf{x}_j + \mathbf{f}_\theta \delta \boldsymbol{\theta}) dt = \int_{I_j} \mathbf{f}_x^T \mathbf{y}_j \cdot \delta \mathbf{x}_j dt + \int_{I_j} \mathbf{f}_\theta^T \mathbf{y}_j \cdot \delta \boldsymbol{\theta} dt$$

además de la siguiente igualdad, obtenida por integración por partes sobre $I_j = [\tau_j, \tau_{j+1}]$

$$\int_{I_j} \mathbf{y}_j \cdot \delta \dot{\mathbf{x}}_j dt = \mathbf{y}_j(\tau_{j+1}) \cdot \delta \mathbf{x}_j(\tau_{j+1}) - \mathbf{y}_j(\tau_j) \cdot \delta \mathbf{x}_j(\tau_j) - \int_{I_j} \dot{\mathbf{y}}_j \cdot \delta \mathbf{x}_j dt$$

Sustituyendo las expresiones anteriores en (3.5) se obtiene

$$\begin{aligned} \nabla_{\mathbf{p}_j} \mathcal{L}_{k_j} \cdot \delta \mathbf{p}_j &= \sum_{t_i \in I_j} \frac{\bar{\mathbf{x}}_j(t_i) - \bar{\mathbf{x}}_i}{\bar{\sigma}_i^2} \cdot \delta \bar{\mathbf{x}}_j(t_i) + \int_{I_j} \mathbf{f}_x^T \mathbf{y}_j \cdot \delta \mathbf{x}_j dt + \int_{I_j} \mathbf{f}_\theta^T \mathbf{y}_j \cdot \delta \boldsymbol{\theta} dt \\ &\quad - \mathbf{y}_j(\tau_{j+1}) \cdot \delta \mathbf{x}_j(\tau_{j+1}) + \mathbf{y}_j(\tau_j) \cdot \delta \mathbf{s}_j \\ &\quad + \int_{I_j} \dot{\mathbf{y}}_j \cdot \delta \mathbf{x}_j dt + k_j (\mathbf{x}_j(\tau_{j+1}) - \mathbf{s}_{j+1}) \cdot (\delta \mathbf{x}_j(\tau_{j+1}) - \delta \mathbf{s}_{j+1}) \end{aligned} \quad (3.6)$$

para toda $\mathbf{y}_j \in H_j$.

3.3. Ecuación adjunta

Si se elige $\mathbf{y}_j$ como la solución de la ecuación adjunta con retroceso en el tiempo en cada subintervalo I_j :

$$\begin{cases} -\dot{\mathbf{y}}_j(t) = \mathbf{f}_{\mathbf{x}}(\mathbf{x}_j(t), \boldsymbol{\theta})^T \mathbf{y}_j(t) + \sum_{t_i \in I_j} \frac{\bar{\mathbf{x}}_j(t_i) - \bar{\mathbf{x}}_i}{\bar{\boldsymbol{\sigma}}_i^2} \delta_D(t - t_i), & \tau_{j+1} > t \geq \tau_j \\ \mathbf{y}_j(\tau_{j+1}) = \boldsymbol{\lambda}_j + k_j[\mathbf{x}_j(\tau_{j+1}) - \mathbf{s}_{j+1}] \end{cases} \quad (3.7)$$

se obtiene de (3.6) la siguiente expresión simplificada:

$$\nabla_{\mathbf{p}_j} \mathcal{L}_{k_j} \cdot \delta \mathbf{p}_j = \mathbf{y}_j(\tau_j) \cdot \delta \mathbf{s}_j - [\boldsymbol{\lambda}_j + k_j(\mathbf{x}_j(\tau_{j+1}) - \mathbf{s}_{j+1}) \cdot \delta \mathbf{s}_{j+1} + \left(\int_{I_j} \mathbf{f}_{\boldsymbol{\theta}}^T \mathbf{y}_j dt \right) \delta \boldsymbol{\theta}] \quad (3.8)$$

3.4. Cálculo del gradiente

El gradiente obtenido de las derivadas parciales en cada subintervalo I_j , tal y como sucede en disparo simple es

$$\nabla_{\mathbf{p}_j} J_{k_j}(\mathbf{p}_j) = \begin{bmatrix} \mathbf{y}_j(\tau_j) \\ -\boldsymbol{\lambda}_j - k_j[\mathbf{x}_j(\tau_{j+1}) - \mathbf{s}_{j+1}] \\ \int_{\tau_j}^{\tau_{j+1}} \mathbf{f}_{\boldsymbol{\theta}}(\mathbf{x}_j(t), \boldsymbol{\theta})^T \mathbf{y}_j(t) dt \end{bmatrix}, \quad j = 1, 2, \dots, ns$$

Por lo tanto, después de acumular todos los gradientes asociados a cada subintervalo se obtiene el gradiente completo:

$$\nabla_{\mathbf{p}} \mathcal{L}_k = \nabla J(\mathbf{p}) = \begin{bmatrix} \mathbf{y}_1(\tau_1) \\ \mathbf{y}_2(\tau_2) - \mathbf{r}_1 \\ \vdots \\ \mathbf{y}_{j+1}(\tau_{j+1}) - \mathbf{r}_j \\ \vdots \\ \mathbf{y}_{ns}(\tau_{ns}) - \mathbf{r}_{ns-1} \\ -\mathbf{r}_{ns} \\ \sum_{j=1}^{ns} \int_{\tau_j}^{\tau_{j+1}} \mathbf{f}_{\boldsymbol{\theta}}(\mathbf{x}_j(t), \boldsymbol{\theta})^T \mathbf{y}_j(t) dt \end{bmatrix} \quad (3.9)$$

donde $\mathbf{r}_j = \boldsymbol{\lambda}_j + k_j[\mathbf{x}_j(\tau_{j+1}) - \mathbf{s}_{j+1}]$ para $j = 1, \dots, ns$. Claramente, cada $\mathbf{x}_j(t)$ satisface la ecuación de estado (3.2) en el subintervalo I_j , y cada $\mathbf{y}_j(t)$ satisface la ecuación adjunta correspondiente (3.7) en el mismo subintervalo.

3.5. Algoritmos de optimización

Las variables de estado y adjuntas usadas en el cálculo del gradiente se calculan independientemente en cada ventana de tiempo I_j , permitiendo así una implementación en paralelo. Si se hace $\boldsymbol{\lambda}_j = 0$ para $j = 1, \dots, ns$, el algoritmo BFGS puede ser aplicado al modelo penalizado resultante con parámetros de penalización constantes.

El algoritmo cuasi-Newton más popular es el método BFGS (Broyden, Fletcher, Goldfarb, Shanno). Es un algoritmo de optimización iterativo utilizado para encontrar el mínimo de una función no lineal, además de que es uno de los métodos más eficaces para la optimización de problemas de gran escala, especialmente cuando no se dispone de información explícita sobre la segunda derivada (Hessiana) de la función objetivo. El algoritmo es el siguiente:

Algoritmo BFGS

Inicialización

1. $\mathbf{p}^0 = (\mathbf{s}_1^0, \dots, \mathbf{s}_{ns+1}^0, \boldsymbol{\theta}^0)^T$, matriz Hessiana $H^0 = I$, gradiente $\mathbf{g}^0 = \nabla J(\mathbf{p}^0)$, dirección $\mathbf{d}^0 = -\mathbf{g}^0$.

Descenso

Para $l \geq 0$, dados $\mathbf{p}^l, \mathbf{g}^l, \mathbf{d}^l, H^l$, encontrar $\mathbf{p}^{l+1}, \mathbf{g}^{l+1}, \mathbf{d}^{l+1}, H^{l+1}$ haciendo:

2. Encontrar $\rho_l = \arg \min_{\rho \geq 0} \phi(\rho) = J(\mathbf{p}^l + \rho \mathbf{d}^l)$.
3. Actualizar $\mathbf{p}^{l+1} = \mathbf{p}^l + \rho_l \mathbf{d}^l$ y $\mathbf{g}^{l+1} = \nabla J(\mathbf{p}^{l+1})$

Convergencia y dirección actualizada

Si $\|\mathbf{g}^{l+1}\| \leq \epsilon \|\mathbf{g}^0\|$, entonces

4. Hacer $\mathbf{p}^* = \mathbf{p}^{l+1}$, parar y salir.
5. Actualizar $H^{l+1} = \left(I - \frac{\mathbf{v}\mathbf{u}^T}{\mathbf{u}^T\mathbf{v}}\right) H^l \left(I - \frac{\mathbf{u}\mathbf{v}^T}{\mathbf{u}^T\mathbf{v}}\right) + \frac{\mathbf{v}\mathbf{v}^T}{\mathbf{u}^T\mathbf{v}}$ con $\mathbf{u} = \mathbf{p}^{l+1} - \mathbf{p}^l$ y $\mathbf{v} = \mathbf{g}^{l+1} - \mathbf{g}^l$.
6. Actualizar $\mathbf{d}^{l+1} = -H^{l+1} \mathbf{g}^{l+1}$.
7. Hacer $l = l + 1$ y regresar a 2.

Salir

Con este algoritmo, el paso más importante es el de la búsqueda de línea (paso 2), para hallar el tamaño de paso adecuado. Se busca elegir un ρ que reduzca J , es decir, se trata de encontrar un minimizador global de la función $\phi(\rho) = J(\mathbf{p}^l + \rho \mathbf{d}^l)$, con $\rho > 0$. El éxito de la búsqueda de línea depende en la efectividad de las elecciones tanto de la dirección $\mathbf{d}^l$ como el tamaño de paso ρ . Dado que se tiene una forma eficiente de calcular las derivadas $\phi'(\rho) = \nabla J(\mathbf{p}^l + \rho \mathbf{d}^l) \cdot \mathbf{d}^l$, para este trabajo, se utilizará el método de la secante para la búsqueda de línea, el cual se ve de la siguiente forma:

$$\rho_{k+1} = \rho_k - \frac{\phi'(\rho_k)(\rho_k - \rho_{k-1})}{\phi'(\rho_k) - \phi'(\rho_{k-1})} = \frac{\rho_{k-1}\phi'(\rho_k) - \rho_k\phi'(\rho_{k-1})}{\phi'(\rho_k) - \phi'(\rho_{k-1})}, \quad k = 1, 2, \dots \quad (3.10)$$

Este método requiere de dos valores iniciales ρ_0 y ρ_1 . Para $l \geq 1$, se inicializa $\rho_0 = 0$ y $\rho_1 = \mathbf{g}^l \cdot (\mathbf{p}^l - \mathbf{p}^{l-1}) / \mathbf{g}^l \cdot \mathbf{d}^l$, mientras que para $l = 0$, se utiliza $\rho_0 = 0$ y $\rho_1 = -0.1(\mathbf{g}^0 \cdot \mathbf{p}^0 / \mathbf{g}^0 \cdot \mathbf{d}^0)$. Estas elecciones hechas consideran una escala adecuada (véase paso 3). Cabe destacar que en los experimentos y resultados que se mostrarán en los capítulos siguientes se utilizarán a lo más tres iteraciones del método de la secante. Tras realizar varios experimentos, se aprecia que los resultados obtenidos entre los métodos de Runge-Kutta de órdenes 2 y 4 no varían significativamente, por lo que los resultados que se muestren a continuación serán aquellos utilizados con el método de orden 2.

La solución teórica se considerará usando el intervalo de tiempo $[t_0, t_f] = [40, 80]$ (medido en días), con parámetros $\boldsymbol{\theta} = (\alpha, \beta, \gamma, \mu)^T = (1/N, 1/7, 1/5, 1/70)^T$, con condición inicial $\mathbf{x}_0 = [91647, 4853, 1755, 1620, 125]^T$ y con población $N = 10^5$. Los datos usados en este trabajo serán generados sintéticamente agregando ruido Gaussiano con media cero y desviación estándar $\sigma(t_i) = \text{noise.level} * \mathbf{x}(t_i)$ en los tiempos t_i con $i = 1, \dots, m$. Adicionalmente, se considerarán ruidos del 5 %, 10 %, 20 %, donde se hará la estimación del vector de parámetros $\boldsymbol{\theta}$.

3.6. Problema de Mínimos Cuadrados

Una vez hecha la generación de datos con ruido Gaussiano, estos datos se necesitan escoger de manera que mejor se asemejen a la tendencia de la solución teórica del problema original. Recordando la ecuación adjunta (3.7), esta presenta discontinuidades en el término δ_D , y aparecen como saltos puntuales en la ecuación diferencial. Para ajustar un modelo a los datos con ruido generados, se buscará minimizar la suma de los cuadrados de los residuales. Es decir, se suavizarán los términos discontinuos, a través del ajuste lineal de los datos a curvas suaves.

Sean x una variable independiente y $y(x)$ una función desconocida que depende de x la cual se desea aproximar.

Supongamos entonces, que tenemos m observaciones $(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)$, donde la idea es modelar $y(x)$ como combinación lineal de funciones base $\phi_1(x), \phi_2(x), \dots, \phi_n(x)$:

$$y(x) \approx c_1\phi_1(x) + c_2\phi_2(x) + \dots + c_n\phi_n(x)$$

donde $m > n$, de modo que la matriz de diseño A con coeficientes $a_{ij} = \phi_j(x_i)$ tiene más renglones que columnas. Luego, los datos deben satisfacer:

$$y(x_i) \approx c_1\phi_1(x_i) + c_2\phi_2(x_i) + \dots + c_n\phi_n(x_i), \quad i = 1, 2, \dots, m$$

que resultan en el sistema lineal

$$\begin{pmatrix} \phi_1(x_1) & \phi_2(x_1) & \dots & \phi_n(x_1) \\ \phi_1(x_2) & \phi_2(x_2) & \dots & \phi_n(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1(x_m) & \phi_2(x_m) & \dots & \phi_n(x_m) \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} \approx \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}$$

Las funciones base $\phi_j(x)$ pueden ser polinomiales, exponenciales, Gaussianas, funciones racionales, entre otras. Haciendo énfasis en las primeras mencionadas, si se escoge la base de funciones polinomiales canónica $\phi_j(x) = x^{j-1}$, con $j = 1, 2, \dots, n$ y además, si $m = n$ el polinomio de grado $m - 1$ que ajustará los m datos está definido por:

$$p(x) = c_1 + c_2x + \dots + c_nx^{n-1}$$

donde $p(x)$ es de grado $n - 1$ y además se cumple que $p(x_i) \approx y(x_i) = y_i$ para $i = 1, 2, \dots, m$, y c_j con $j = 1, 2, \dots, n$ son los coeficientes por calcular del sistema

$$\begin{pmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_m & x_m^2 & \dots & x_m^{n-1} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} \approx \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}$$

con $A \in \mathbb{R}^{m \times n}$ la matriz de Vandermonde. El problema de mínimos cuadrados se plantea, suponiendo que la aproximación $p(x_i)$ está lo más cerca posible del dato y_i en la norma euclidiana, es decir

$$\min_{\mathbf{c} \in \mathbb{R}^n} \|\mathbf{y} - A\mathbf{c}\|_2^2$$

Si se hace $\mathbf{c} = \mathbf{x}$ y $\mathbf{y} = \mathbf{b}$, se puede reescribir el problema de mínimos cuadrados como:

Dado $A \in \mathbb{R}^{m \times n}$ y $\mathbf{b} \in \mathbb{R}^m$, encontrar $\hat{\mathbf{x}} \in \mathbb{R}^n$ tal que

$$\|\mathbf{b} - A\hat{\mathbf{x}}\|_2 = \min_{\mathbf{x} \in \mathbb{R}^n} \|\mathbf{b} - A\mathbf{x}\|_2$$

donde los residuales son las diferencias entre las observaciones y los valores del modelo, es decir $\mathbf{r} = \mathbf{b} - A\mathbf{x}$, por lo que el objetivo será encontrar $\hat{\mathbf{x}} \in \mathbb{R}^n$ tal que satisfaga la ecuación anterior.

Existen distintos métodos para resolver el problema de mínimos cuadrados, entre ellos el de ecuaciones normales y la factorización QR. El primero es uno de los más utilizados; sin embargo, cuando la matriz de diseño A está mal condicionada —es decir, una pequeña perturbación en A produce variaciones muy grandes en la solución del sistema $A\mathbf{x} = \mathbf{b}$ — la solución del problema de mínimos cuadrados se ve afectada de manera significativa. Para enfrentar estos casos de mal condicionamiento, se recurre a la factorización QR.

3.6.1. Factorización QR

Para remediar el mal condicionamiento de las matrices de diseño mencionadas anteriormente, otra forma de resolver el problema de mínimos cuadrados es a través de la factorización QR.

Teorema 1. *Toda matriz $A \in \mathbb{C}^{m \times n}$ con $(m \geq n)$ de rango completo tiene una factorización QR única y reducida, donde $A = \hat{Q}\hat{R}$, con $r_{jj} > 0$.*

La demostración de este teorema se puede encontrar en [30].

Por simplicidad, se hará $A \in \mathbb{R}^{m \times n}$. En este caso $\hat{Q}$ tiene el mismo tamaño de A y $\hat{R} \in \mathbb{R}^{m \times n}$ es una matriz triangular superior. Esta factorización es la versión matricial del proceso de ortogonalización de Gram-Schmidt. Sean $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n \in \mathbb{R}^m$ vectores columna linealmente independientes de A

$$A = \begin{bmatrix} | & | & & | \\ \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_n \\ | & | & & | \end{bmatrix}$$

luego, los vectores ortonormales $\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_n \in \mathbb{R}^m$ obtenidos del proceso de ortogonalización de Gram-Schmidt dan la siguiente matriz

$$\hat{Q} = \begin{bmatrix} | & | & & | \\ \mathbf{q}_1 & \mathbf{q}_2 & \dots & \mathbf{q}_n \\ | & | & & | \end{bmatrix}$$

que tienen el mismo espacio columna de A .

Definición 2. Sea $A \in \mathbb{R}^{m \times n}$. El espacio columna, $\mathcal{C}(A)$, es el conjunto de todos los vectores que son combinaciones lineales de las columnas de A .

$$\mathcal{C}(A) = \{\mathbf{y} \in \mathbb{R}^m | \mathbf{y} = A\mathbf{x}\}$$

Estos vectores son construidos secuencialmente, empezando con $\mathbf{q}_1 = \mathbf{a}_1 / \|\mathbf{a}_1\|$ y que satisfacen

$$\mathbf{q}_j = \frac{\mathbf{v}_j}{\|\mathbf{v}_j\|_2} \quad \text{con} \quad \mathbf{v}_j = \mathbf{a}_j - (\mathbf{q}_1^T \mathbf{a}_j) \mathbf{q}_1 - \dots - (\mathbf{q}_{j-1}^T \mathbf{a}_j) \mathbf{q}_{j-1} = \mathbf{a}_j - \sum_{i=1}^{j-1} (\mathbf{q}_i^T \mathbf{a}_j) \mathbf{q}_i \quad 1 \leq j \leq n.$$

Usando la notación $r_{ij} \equiv \mathbf{q}_i^T \mathbf{a}_j$ para $i > j$ y $r_{jj} = \|\mathbf{v}_j\|_2$, se obtiene

$$\begin{aligned} \mathbf{q}_1 &= \frac{\mathbf{a}_1}{r_{11}}, \\ \mathbf{q}_2 &= \frac{\mathbf{a}_2 - r_{12} \mathbf{q}_1}{r_{22}}, \\ &\vdots \\ \mathbf{q}_n &= \frac{\mathbf{a}_n - \sum_{i=1}^{n-1} r_{in} \mathbf{q}_i}{r_{nn}}. \end{aligned} \quad \Longrightarrow \quad \begin{aligned} \mathbf{a}_1 &= r_{11} \mathbf{q}_1, \\ \mathbf{a}_2 &= r_{12} \mathbf{q}_1 + r_{22} \mathbf{q}_2, \\ &\vdots \\ \mathbf{a}_n &= r_{1n} \mathbf{q}_1 + r_{2n} \mathbf{q}_2 + \dots + r_{nn} \mathbf{q}_n. \end{aligned}$$

Este conjunto de ecuaciones se convierte en la factorización QR reducida:

$$A = \begin{bmatrix} | & | & \dots & | \\ \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & \dots & | \\ \mathbf{q}_1 & \mathbf{q}_2 & \dots & \mathbf{q}_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1n} \\ 0 & r_{22} & \dots & r_{2n} \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & r_{nn} \end{bmatrix} = \hat{Q} \hat{R}$$

Esta factorización es otra forma de resolver el sistema sobredeterminado $A\mathbf{x} = \mathbf{b}$, que surge del problema de mínimos cuadrados. La propiedad clave es que $\hat{Q}^T \hat{Q} = I_n$, donde I_n es la matriz identidad de tamaño $n \times n$. Luego,

$$\hat{Q} \hat{R} \mathbf{x} = \mathbf{b} \quad \Longrightarrow \quad \hat{R} \mathbf{x} = \hat{Q}^T \mathbf{b},$$

y este sistema triangular es resuelto usando sustitución hacia atrás. Más aún, la solución obtenida es la solución para el problema de mínimos cuadrados, ya que

$$\hat{R}\mathbf{x} = \hat{Q}^T \mathbf{b} \iff \hat{Q}\hat{R}\mathbf{x} = \hat{Q}\hat{Q}^T \mathbf{b} \iff A\mathbf{x} = P_{\hat{Q}} \mathbf{b} \iff A\mathbf{x} = P_A \mathbf{b},$$

donde las matrices de proyección satisfacen $P_{\hat{Q}} \mathbf{b} = P_A \mathbf{b}$ ya que el espacio columna de A es igual al espacio columna de $\hat{Q}$.

Definición 3. Una proyección en $\mathbb{R}^n$ es una matriz cuadrada P tal que $P^2 = P$, es decir, es una matriz idempotente.

Definición 4. Una proyección P es ortogonal si P es simétrica, es decir, si $P^T = P$.

Una factorización QR completa de $A \in \mathbb{R}^{m \times n}$, con $m \geq n$, consiste en agregar $m - n$ columnas ortogonales a $\hat{Q}$ y $m - n$ renglones de ceros a $\hat{R}$, de tal forma que se obtenga una matriz ortogonal $Q \in \mathbb{R}^{m \times m}$ y otra matriz $R \in \mathbb{R}^{m \times n}$ triangular superior.

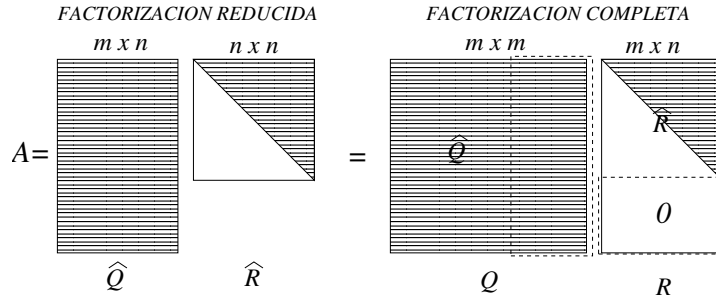


Figura 3.2: Factorizaciones QR reducida y completa de A

Teorema 2. Toda matriz $A \in \mathbb{R}^{m \times n}$ ($m \geq n$) tiene una factorización QR completa, dada por $A = QR$ con $Q \in \mathbb{R}^{m \times m}$ una matriz ortogonal y $R \in \mathbb{R}^{m \times n}$ una matriz triangular superior.

El algoritmo de Gram-Schmidt es numéricamente inestable, es decir, es sensible a errores de redondeo. Es posible utilizar métodos que lo vuelvan estable, cambiando el orden de las operaciones, sin embargo, el algoritmo que resuelve el problema de mínimos cuadrados y que es numéricamente estable es la factorización QR de Householder.

Factorización QR de Householder

Este método usa un proceso de triangularización ortogonal llamado *triangularización de Householder*, que consiste en realizar n iteraciones, donde en cada etapa se multiplica por la izquierda una matriz ortogonal (o unitaria) Q_k , de forma que al final se genere una matriz triangular superior R , es decir,

$$Q_n \dots Q_2 Q_1 A = R$$

donde las Q_k con $k = 1, \dots, n$ son matrices ortogonales. Si A fuese una matriz de 5×3 , el proceso de triangularización se ve de la siguiente forma:

$$\underbrace{\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \\ a_{51} & a_{52} & a_{53} \end{bmatrix}}_{A^{(0)} \atop A} \rightarrow \underbrace{\begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} \\ 0 & a_{22}^{(1)} & a_{23}^{(1)} \\ 0 & a_{32}^{(1)} & a_{33}^{(1)} \\ 0 & a_{42}^{(1)} & a_{43}^{(1)} \\ 0 & a_{52}^{(1)} & a_{53}^{(1)} \end{bmatrix}}_{A^{(1)} \atop Q_1 A} \rightarrow \underbrace{\begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} \\ 0 & 0 & a_{33}^{(2)} \\ 0 & 0 & a_{43}^{(2)} \\ 0 & 0 & a_{53}^{(2)} \end{bmatrix}}_{A^{(2)} \atop Q_2 Q_1 A} \rightarrow \underbrace{\begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} \\ 0 & 0 & a_{33}^{(3)} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}}_{A^{(3)} \atop Q_3 Q_2 Q_1 A} = R,$$

donde se puede apreciar que $A^{(1)} = Q_1 A$, $A^{(2)} = Q_2 Q_1 A$ y $A^{(3)} = Q_3 Q_2 Q_1 A$.

Reflexiones de Householder

Si $A \in \mathbb{R}^{m \times n}$, donde $m > n$, cada matriz ortogonal Q_k debe tener la estructura siguiente:

$$Q_k = \begin{pmatrix} I & 0 \\ 0 & H \end{pmatrix}$$

donde I es la matriz identidad de orden $(k-1) \times (k-1)$ y H es una matriz ortogonal de orden $(m-(k-1)) \times (m-(k-1))$. En la k -ésima iteración, se debe calcular $Q_k A^{k-1}$ para obtener A^k , de modo que A^k deberá tener ceros debajo de la diagonal hasta la columna k , entonces:

$$Q_k = \begin{pmatrix} I & 0 \\ 0 & H \end{pmatrix}$$

$$A^{k-1} = \begin{pmatrix} a_{1,1} & \dots & a_{1,k-1} & a_{1,k} & \dots & a_{1,n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & a_{k-1,k-1} & a_{k-1,k} & \dots & a_{k-1,n} \\ 0 & \dots & 0 & a_{k,k} & \dots & a_{k,n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & a_{m,k} & \dots & a_{m,n} \end{pmatrix}$$

por lo que se busca que

$$H \cdot \begin{pmatrix} a_{k,k} & \dots & a_{k,n} \\ \vdots & \ddots & \vdots \\ a_{m,k} & \dots & a_{m,n} \end{pmatrix} = \begin{pmatrix} a'_{k,k} & a'_{k,k+1} & \dots & a'_{k,n} \\ 0 & a'_{k+1,k+1} & \dots & a'_{k+1,n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a'_{m,k+1} & \dots & a'_{m,n} \end{pmatrix}$$

es decir, se debe escoger una matriz tal que al multiplicarla por el vector $(a_{kk}, a_{k+1k}, \dots, a_{mk})^T$, introduzca ceros bajo la diagonal de la primera columna:

$$H \begin{pmatrix} a_{k,k} \\ a_{k+1,k} \\ \vdots \\ a_{mk} \end{pmatrix} = \begin{pmatrix} a'_{kk} \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

donde $a'_{kk} = \|\mathbf{x}\|_2$ con $\mathbf{x} = (a_{kk}, a_{k+1k}, \dots, a_{mk})^T$. H debe satisfacer que:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{m-k+1} \end{bmatrix} \rightarrow H\mathbf{x} = \begin{bmatrix} \|\mathbf{x}\| \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \|\mathbf{x}\| \mathbf{e}_1$$

con $\mathbf{e}_1 = (1, 0, \dots, 0)^T$ de tamaño $m - (k - 1)$. La matriz H se conoce como la *reflexión de Householder*, y se define como

$$H = I - 2 \frac{\mathbf{v}\mathbf{v}^T}{\mathbf{v}^T\mathbf{v}}, \quad \text{con } \mathbf{v} = \|\mathbf{x}\|_2 \mathbf{e}_1 - \mathbf{x} \quad (3.11)$$

En términos generales, la reflexión de Householder es una transformación lineal con matriz H , la cual es construida dado un vector $\mathbf{x}$ en un espacio euclidiano $\mathbb{R}^p$ ($p \geq 2$), que busca que su vector reflejado sea $H\mathbf{x} = \|\mathbf{x}\| \mathbf{e}_1 = (\|\mathbf{x}\|, 0, \dots, 0)^T$ tal como se muestra en la siguiente Figura:

Para cualquier vector $\mathbf{x}$ existen dos reflexiones de Householder, tal como se muestra en la Figura 3.4, y cada matriz H se construye con la elección de $\mathbf{v} = \text{signo}(\mathbf{x}_1) \|\mathbf{x}\|_2 \mathbf{e}_1 + \mathbf{x}$. Claramente, esta elección permite que $\|\mathbf{v}\|$ nunca sea más pequeño que $\|\mathbf{x}\|$, evitando la cancelación por la sustracción cuando se divide entre $\|\mathbf{v}\|$ para encontrar $\mathbf{u}$ en (3.11), y así asegurar la estabilidad del método.

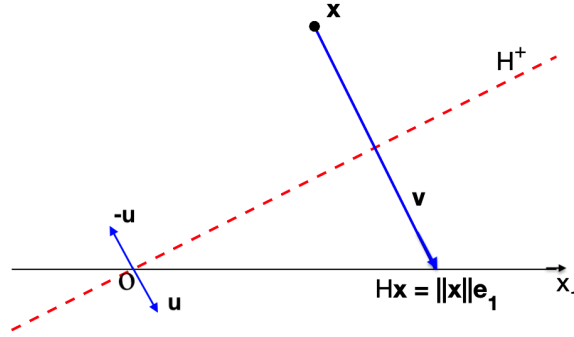


Figura 3.3: Reflexión de Householder con el vector $\mathbf{v} = \|\mathbf{x}\| \mathbf{e}_1 - \mathbf{x}$.

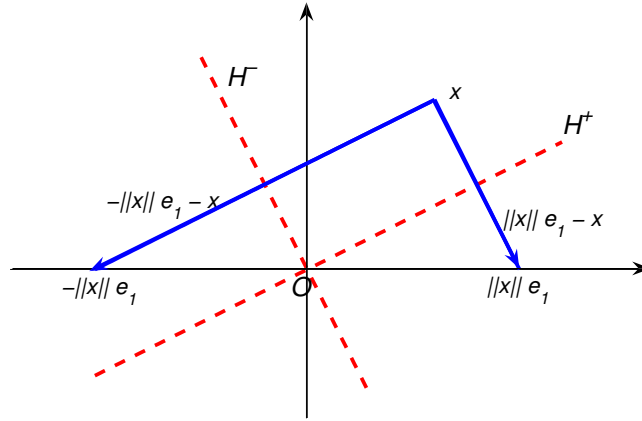


Figura 3.4: Reflexiones de Householder, construidas a partir de $\mathbf{x}$.

Se puede calcular la factorización $A = \hat{Q}\hat{R}$, con $\hat{Q} = (Q_n \dots Q_2 Q_1)^T$. Sin embargo, si el interés es solamente en la solución del problema de mínimos cuadrados, no es necesario calcular explícitamente las matrices Q_k o $\hat{Q}$. Solamente se necesita encontrar el factor R y almacenarlo en el mismo espacio de memoria que A , y $\hat{Q}\mathbf{b}$ y almacenarlo en el mismo espacio de memoria que $\mathbf{b}$. Al final, se está resolviendo el sistema triangular usando sustitución hacia atrás. El algoritmo para la triangularización de Householder para resolver $A\mathbf{x} = \mathbf{b}$ con factorización QR se ve de la siguiente manera:

Triangularización

Para $k = 1, \dots, n$

- $\mathbf{x} = A(k : m, k)$
- $\mathbf{v} = \text{sign}(x_1) \|\mathbf{x}\| \mathbf{e}_1 + \mathbf{x}$
- $\mathbf{v} = \mathbf{v} / \|\mathbf{v}\|$
- $A(k : m, k : n) = A(k : m, k : n) - 2\mathbf{v} (\mathbf{v}^T A(k : m, k : n))$

- $\mathbf{b}(k : m) = \mathbf{b}(k : m) - 2\mathbf{v}(\mathbf{v}^T \mathbf{b}(k : m))$

Termina para

Sustitución hacia atrás

- $\mathbf{x}(n) = \mathbf{b}(n)/A(n, n)$

Para $k = n - 1 : -1 : 1$

- $\mathbf{x}(k) = (\mathbf{b}(k) - A(k, k + 1 : n) \cdot \mathbf{b}(k + 1 : n)) / A(k, k)$

Termina para

Capítulo 4

Modelo SEIRD

Uno de los sistemas de interés en este trabajo es el modelo epidemiológico SEIRD. Este modelo es evolución del primer modelo epidemiológico establecido en la historia, el modelo SIR. El sistema de ecuaciones del modelo SEIRD está definido por (2.4).

Los datos sintéticos serán generados a partir de la solución teórica usando ruido Gaussiano establecido por MATLAB, con media igual a cero y desviaciones estándar $\sigma(t_i) = \text{noise_level} * \mathbf{x}(t_i)$, para $i = 1, \dots, m$, donde se generaron $m = 15$ medidas experimentales con las variables observables $\bar{\mathbf{x}}_i = (I_i, R_i, D_i)$ en los tiempos $t_i = 40 + i$, $0 \leq i \leq 14$ (días). Para esto, a la solución teórica se le estará implementando tres niveles de ruido del 5 %, 10 %, 20 %. Bajo estas condiciones, se espera que se pueda recuperar la aproximación de las condiciones iniciales $\mathbf{s}_0$ y los valores verdaderos del parámetro $\boldsymbol{\theta}$ con el algoritmo numérico propuesto.

Cabe resaltar que dependiendo de la información con la que se cuente, esto puede restringir el número de subintervalos en el método de disparo múltiple. Mientras haya más información disponible, el desempeño del método es mejor, por lo tanto, a los 15 puntos con los que se estará trabajando se les deberá ajustar un polinomio de tercer grado (o segundo grado si la tendencia de la curva así lo permite), usando el algoritmo QR (véase subsección 3.6.1). Con la curva resultante de dicho ajuste, se generan datos adicionales con espaciado de $h = 0.1$, de modo que ahora se estarán trabajando en el intervalo $[40, 54]$ con $m = (54 - 40)/h = 141$ datos. Dado que estos puntos pertenecen a una curva suave, los pulsos en las series del lado derecho de las ecuaciones adjuntas (3.7) serán suavizadas, obteniendo una mejor estabilidad que la obtenida con el método de disparo simple. Ahora, como la dinámica del modelo SEIRD no es caótica, se empleará el método de Runge-Kutta de orden 2 para encontrar la solución numérica de las ecuaciones de estado (3.2) y las ecuaciones adjuntas (3.7) en cada iteración del algoritmo BFGS.

Tabla 4.1: Estimación de θ con diferentes valores de ns y distintos niveles de ruido.

Nivel de ruido	ns	Valor estimado de θ	Error relativo	iter
5 %	7	$[1.0250, 0.1510, 0.2074, 0.0138]^T$	$[0.0250, 0.0570, 0.0372, 0.0349]^T$	40
	14	$[1.0125, 0.1508, 0.2087, 0.0138]^T$	$[0.0125, 0.0558, 0.0436, 0.0366]^T$	66
	28	$[1.0045, 0.1504, 0.2083, 0.0120]^T$	$[0.0045, 0.0525, 0.0416, 0.1578]^T$	66
	35	$[1.0034, 0.1504, 0.2088, 0.0125]^T$	$[0.0034, 0.0530, 0.0441, 0.1274]^T$	95
	70	$[0.9978, 0.1502, 0.2087, 0.0118]^T$	$[0.0002, 0.0517, 0.0433, 0.1734]^T$	92
10 %	7	$[1.0438, 0.1467, 0.2107, 0.0148]^T$	$[0.0438, 0.0272, 0.0536, 0.0370]^T$	109
	14	$[1.0510, 0.1439, 0.2137, 0.0150]^T$	$[0.0510, 0.0074, 0.0685, 0.0521]^T$	90
	28	$[1.0086, 0.1421, 0.2112, 0.0144]^T$	$[0.0086, 0.0052, 0.0561, 0.0080]^T$	73
	35	$[1.0075, 0.1421, 0.2114, 0.0145]^T$	$[0.0075, 0.0051, 0.0571, 0.0118]^T$	84
	70	$[1.0054, 0.1421, 0.2118, 0.0146]^T$	$[0.0054, 0.0051, 0.0591, 0.0197]^T$	179
20 %	7	$[1.0564, 0.1505, 0.2027, 0.0117]^T$	$[0.0564, 0.0533, 0.0134, 0.1829]^T$	25
	14	$[1.0404, 0.1518, 0.2049, 0.0136]^T$	$[0.0404, 0.0627, 0.0245, 0.0512]^T$	60
	28	$[1.0316, 0.1523, 0.2056, 0.0137]^T$	$[0.0316, 0.0663, 0.0278, 0.0432]^T$	93
	35	$[1.0289, 0.1521, 0.2046, 0.0123]^T$	$[0.0289, 0.0647, 0.0232, 0.1406]^T$	84
	70	$[1.0255, 0.1524, 0.2051, 0.0125]^T$	$[0.0255, 0.0666, 0.0254, 0.1269]^T$	157

4.1. Resultados numéricos para el modelo SEIRD

Para este conjunto de resultados, se tomará como propuesta (adivinanza) $\theta_0 = (1.4, 0.09, 0.3, 0.005)^T$, se tomarán 15 datos generados y suavizados para cada nivel de ruido, y se estará considerando el método de Runge-Kutta de orden 2, pues la diferencia de resultados entre el método de orden 2 y de orden 4 no varía tanto. Adicionalmente, se estará considerando en el algoritmo de optimización $\rho_0 = 0$ y $\rho_1 = \mathbf{g}^l \cdot (\mathbf{p}^l - \mathbf{p}^{l-1}) / \mathbf{g}^l \cdot \mathbf{d}^l$ para $l \geq 1$, mientras que para $l = 0$, se utilizarán $\rho_0 = 0$ y $\rho_1 = -0.1(\mathbf{g}^0 \cdot \mathbf{p}^0 / \mathbf{g}^0 \cdot \mathbf{d}^0)$, considerando $\epsilon = 10^{-7}$ como criterio de parada para el algoritmo BFGS. Los nuevos datos generados se ven de la siguiente manera:

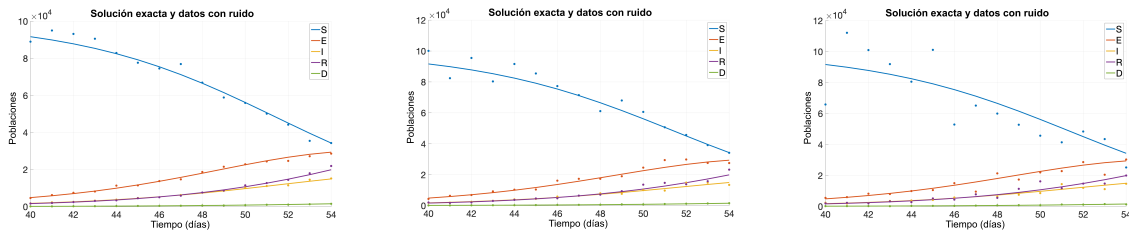


Figura 4.1: Datos con ruido. De izquierda a derecha: niveles de ruido = 0.05, 0.10, 0.20

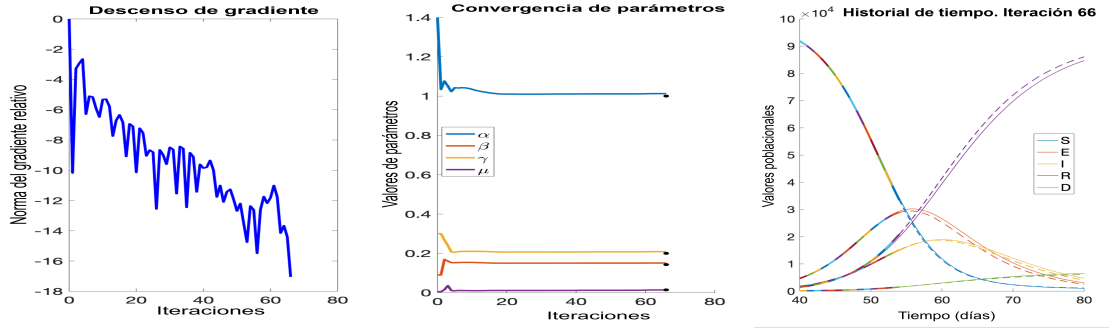
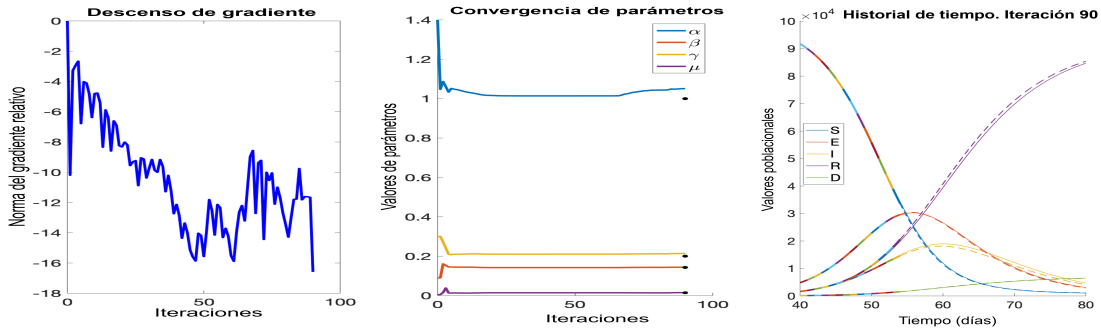
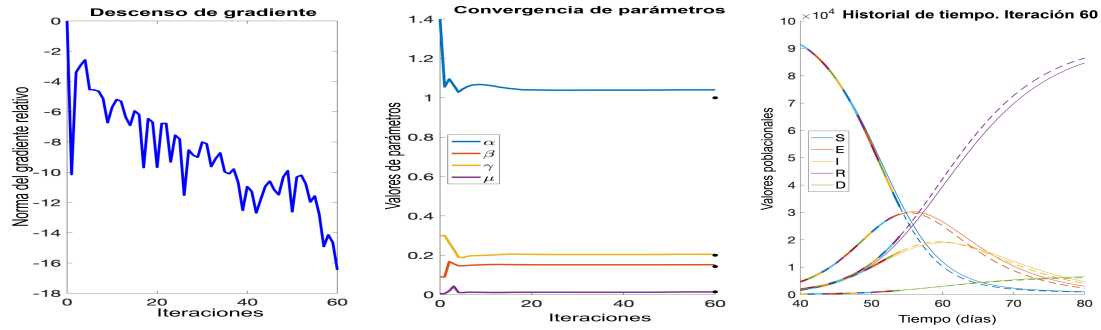
Sintetizando la información para los tres niveles de ruido, algunas de las aproximaciones del vector de parámetros θ y la condición inicial $\mathbf{s}_0$ se muestran en las tablas 4.1 y 4.2, respectivamente:

Tabla 4.2: Estimación de $\mathbf{s}_0$ con diferentes valores de ns y distintos niveles de ruido.

Nivel de ruido	ns	Valor estimado de $\mathbf{s}_0$	Error relativo	iter
5 %	7	$[92103, 4496, 1733, 1542, 126]^T$	$[0.0050, 0.0738, 0.0131, 0.0481, 0.0080]^T$	40
	14	$[92015, 4689, 1739, 1431, 126]^T$	$[0.0040, 0.0340, 0.0097, 0.1167, 0.0080]^T$	66
	28	$[91965, 4791, 1746, 1340, 158]^T$	$[0.0035, 0.0130, 0.0057, 0.1728, 0.2640]^T$	66
	35	$[91970, 4821, 1745, 1330, 134]^T$	$[0.0035, 0.0068, 0.0063, 0.1790, 0.0720]^T$	95
	70	$[91952, 4861, 1743, 1281, 163]^T$	$[0.0033, 0.0014, 0.0074, 0.2093, 0.3040]^T$	92
10 %	7	$[91827, 5119, 1485, 1453, 116]^T$	$[0.0020, 0.0546, 0.1543, 0.1031, 0.0720]^T$	109
	14	$[91794, 4835, 1641, 1616, 114]^T$	$[0.0016, 0.0039, 0.0655, 0.0025, 0.0880]^T$	90
	28	$[92499, 3869, 2363, 1164, 105]^T$	$[0.0093, 0.2029, 0.3457, 0.2815, 0.1600]^T$	73
	35	$[92496, 3881, 2368, 1152, 103]^T$	$[0.0093, 0.2005, 0.3485, 0.2889, 0.1760]^T$	84
	70	$[92489, 3909, 2375, 1128, 99]^T$	$[0.0092, 0.1947, 0.3525, 0.3037, 0.2080]^T$	179
20 %	7	$[91579, 4509, 1575, 2069, 268]^T$	$[0.0007, 0.0711, 0.1031, 0.2772, 1.1440]^T$	25
	14	$[91580, 4650, 1603, 2035, 132]^T$	$[0.0007, 0.0420, 0.0871, 0.2562, 0.0560]^T$	60
	28	$[91520, 4730, 1614, 2003, 133]^T$	$[0.0014, 0.0255, 0.0809, 0.2364, 0.0640]^T$	93
	35	$[91434, 4733, 1617, 1984, 232]^T$	$[0.0023, 0.0249, 0.0792, 0.2247, 0.8560]^T$	84
	70	$[91416, 4773, 1619, 1972, 220]^T$	$[0.0025, 0.0167, 0.0780, 0.2173, 0.7600]^T$	157

De los resultados mostrados se puede apreciar que las mejores estimaciones de $\boldsymbol{\theta}$ se obtienen con $ns = 7, 14$ subintervalos para ruido del 5 %, luego para el nivel de ruido del 10 %, en general, para todos los números de subintervalos se obtienen buenos resultados, mientras que para ruido del 20 %, es con $ns = 14, 28$ subintervalos con los que se obtienen los mejores resultados. En general, el método de disparo múltiple con $ns = 14$ es la mejor opción para los tres niveles de ruido diferentes. Por otra parte, las estimaciones de las condiciones iniciales $\mathbf{s}_0$ son menos precisas, especialmente en la cuarta componente (población recuperada) cuando se consideran niveles de ruido más altos.

Si se grafica $\log(\|\mathbf{g}^l\|/\|\mathbf{g}\|^0)$ respecto al número de iteraciones se obtienen las Figuras de lado izquierdo, mientras que la convergencia de los parámetros con este conjunto de datos utilizados se aprecia en las imágenes en medio. En dichas imágenes se muestra cómo $(\alpha^l, \beta^l, \gamma^l, \mu^l)$ convergen al valor óptimo. El valor del parámetro de la solución teórica se muestra por medio de los puntos negros al final de cada curva. Finalmente, las imágenes de lado derecho corresponden al comportamiento de las variables de estado en el transcurso de tiempo. Las líneas de varios colores representan la dinámica reconstruida con la condición inicial calculada $\mathbf{s}_0$ y los parámetros estimados $\boldsymbol{\theta}$ en el intervalo de tiempo $[40, 54]$, luego las líneas punteadas muestran la predicción a futuro en el intervalo $[54, 80]$ en conjunto con la solución exacta (líneas continuas).

(a) $ns = 14$ and $noise_level = 0.05$.(b) $ns = 14$ and $noise_level = 0.10$.(c) $ns = 14$ and $noise_level = 0.20$.Figura 4.2: Resultados para $ns = 14$ con distintos niveles de ruido.

Cabe resaltar que en general, es posible utilizar más o menos variable observable, y gracias al método de disparo múltiple en conjunto con el algoritmo BFGS, se puede resolver el problema. En particular, si se llegasen a usar distintas combinaciones de variables observables, se obtienen las siguientes aproximaciones, condensadas en las tablas 4.3 y 4.4:

Tabla 4.3: Estimación de θ con $ns = 14$ y nivel de ruido de 20 %.

Variables observables	Valor estimado de θ	Error relativo	iter
I	$[1.0392, 0.1514, 0.2037, 0.0119]^T$	$[0.0392, 0.0600, 0.0184, 0.1653]^T$	48
R	$[1.0393, 0.1514, 0.2037, 0.0119]^T$	$[0.0393, 0.0600, 0.0184, 0.1652]^T$	47
D	$[1.0405, 0.1518, 0.2049, 0.0136]^T$	$[0.0405, 0.0627, 0.0245, 0.0504]^T$	60
I, R	$[1.0393, 0.1514, 0.2037, 0.0119]^T$	$[0.0393, 0.0600, 0.0184, 0.1652]^T$	47
I, D	$[1.0404, 0.1518, 0.2049, 0.0136]^T$	$[0.0404, 0.0627, 0.0245, 0.0504]^T$	60
R, D	$[1.0404, 0.1518, 0.2049, 0.0136]^T$	$[0.0404, 0.0627, 0.0245, 0.0512]^T$	60
I, R, D	$[1.0404, 0.1518, 0.2049, 0.0136]^T$	$[0.0404, 0.0627, 0.0245, 0.0512]^T$	60

Tabla 4.4: Estimación de s_0 con $ns = 14$ y nivel de ruido de 20 %.

Variables observables	Valor estimado de s_0	Error relativo	iter
I	$[91489, 4633, 1605, 2020, 253]^T$	$[0.0017, 0.0455, 0.0860, 0.2469, 1.0240]^T$	48
R	$[91489, 4633, 1605, 2020, 253]^T$	$[0.0017, 0.0455, 0.0860, 0.2469, 1.0240]^T$	47
D	$[91579, 4650, 1602, 2038, 131]^T$	$[0.0007, 0.0420, 0.0877, 0.2580, 0.0480]^T$	60
I, R	$[91489, 4633, 1605, 2020, 253]^T$	$[0.0017, 0.0455, 0.0860, 0.2469, 1.0240]^T$	47
I, D	$[91579, 4650, 1602, 2038, 131]^T$	$[0.0007, 0.0420, 0.0877, 0.2580, 0.0480]^T$	60
R, D	$[91580, 4650, 1603, 2035, 132]^T$	$[0.0007, 0.0420, 0.0871, 0.2562, 0.0560]^T$	60
I, R, D	$[91580, 4650, 1603, 2035, 132]^T$	$[0.0007, 0.0420, 0.0871, 0.2562, 0.0560]^T$	60

De este último par de tablas, se aprecia que se obtienen buenos resultados cuando la población muerta (D) es parte de las variables observables, y también se nota que el número de subintervalos óptimo para el método de disparo múltiple tiene correlación con el número de datos con ruido originales.

Por último, se considerará un último conjunto de datos con ruido del 20 % en distintos intervalos de tiempo, de modo que, naturalmente, el número de observaciones cambiará. Para los casos y resultados siguientes que se mostrarán a continuación (tablas 4.5 y 4.6), se estará haciendo el ajuste con polinomios de segundo grado para los primeros tres casos, mientras que para los últimos dos casos se estarán considerando polinomios de grado tres. De esta manera, se estarán generando 71, 101, 141, 171, 201 puntos en las curvas ajustadas, respectivamente.

Para estos casos, se aprecia que los parámetros calculados continúan siendo buenos y se mantienen estables, ya que la diferencia relativa de los parámetros calculados de θ se mantienen por debajo del 20 % en cada caso. A continuación, la Figura 4.3 muestra la evolución de las variables de estado y su predicción respecto al tiempo. Tal y como

Tabla 4.5: Estimación de θ con diferente número de datos con ruido.

$t_i = 40 + i$	ns	Valor estimado de θ	Error relativo	iter
$i = 0 : 7$	7	$[0.9916, 0.1619, 0.2340, 0.0144]^T$	$[0.0084, 0.1332, 0.1698, 0.0062]^T$	33
$i = 0 : 10$	5	$[0.9622, 0.1587, 0.2219, 0.0149]^T$	$[0.0378, 0.1110, 0.1096, 0.0428]^T$	33
$i = 0 : 14$	14	$[1.0404, 0.1518, 0.2049, 0.0136]^T$	$[0.0404, 0.0627, 0.0245, 0.0512]^T$	60
$i = 0 : 17$	17	$[1.1009, 0.1369, 0.2099, 0.0163]^T$	$[0.1009, 0.0419, 0.0496, 0.1419]^T$	72
$i = 0 : 20$	25	$[1.0539, 0.1297, 0.2075, 0.0150]^T$	$[0.0539, 0.0922, 0.0374, 0.0488]^T$	214

Tabla 4.6: Estimación de s_0 con diferente número de datos con ruido.

$t_i = 40 + i$	ns	Valor estimado de θ	Error relativo	iter
$i = 0 : 7$	7	$[91729, 4895, 1962, 1299, 115]^T$	$[0.0009, 0.0084, 0.1173, 0.1981, 0.0800]^T$	33
$i = 0 : 10$	5	$[91635, 4850, 2042, 1363, 110]^T$	$[0.0001, 0.0008, 0.1629, 0.1586, 0.1200]^T$	33
$i = 0 : 14$	14	$[91580, 4650, 1603, 2035, 132]^T$	$[0.0007, 0.0420, 0.0871, 0.2562, 0.0560]^T$	60
$i = 0 : 17$	17	$[90994, 5289, 1308, 2289, 120]^T$	$[0.0071, 0.0896, 0.2551, 0.4130, 0.0400]^T$	72
$i = 0 : 20$	25	$[91168, 4868, 1884, 1964, 116]^T$	$[0.0052, 0.0029, 0.0729, 0.2123, 0.0720]^T$	214

anteriormente se mencionó, la solución por medio del método de disparo múltiple está representado por las líneas de colores (cada color es un subintervalo del método) en cada intervalo de tiempo correspondiente. La predicción a futuro se muestra con las líneas punteadas, mientras que la solución exacta está representada por las líneas continuas. Tal como se esperaba, los resultados son mejores cuando se tienen más datos experimentales en consideración.

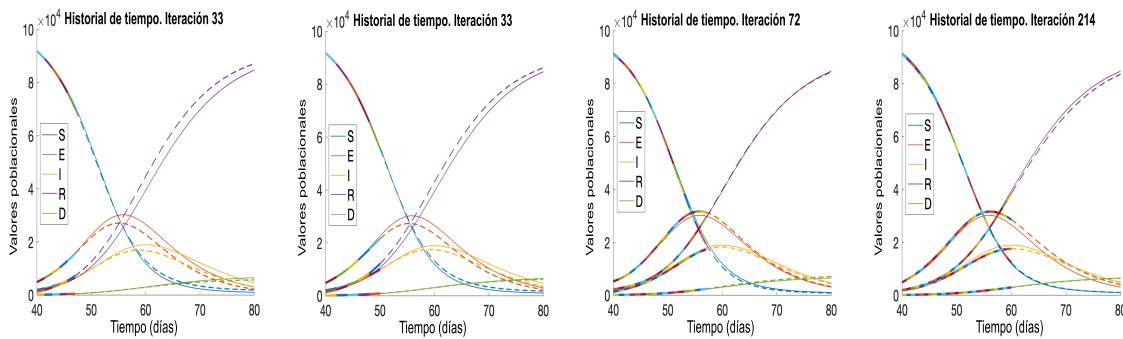


Figura 4.3: Solución calculada respecto al tiempo y predicción, en conjunto con la solución exacta. De izquierda a derecha: $i = 0 : 7$, $i = 0 : 10$, $i = 0 : 17$, $i = 0 : 20$

A continuación, se mostrará el otro modelo de interés, al cual se le podrá hacer la estimación de parámetros con la misma metodología presentada a lo largo de este trabajo. Estos son: los sistemas de reacciones químicas.

Capítulo 5

Modelo de la síntesis enzimática inducible

Antes de dar inicio al segundo modelo, se desarrollarán brevemente algunos conceptos para familiarizarse con la terminología que se usará.

Las reacciones químicas son procesos que transforman materias primas en productos manufacturados. Para que este proceso se pueda llevar a cabo se necesitan realizar maniobras tales como: la medición de las concentraciones durante la reacción, monitorización del proceso, control y optimización del proceso, entre otras. Estos últimos dan pauta a modelos contruidos utilizando la cinética química. Esta área de estudio se encarga de la rapidez de reacción y los factores que la afectan.

Ahora bien, los sistemas de reacciones químicas dan origen a sistemas de ecuaciones no lineales complejas que no poseen una solución analítica. Por esto, utilizaremos métodos numéricos para determinar soluciones del modelo y las técnicas introducidas previamente para la determinación de las tasas de cambio de los parámetros en estos sistemas. En esta sección del trabajo, el enfoque principal estará en el problema de la síntesis enzimática inducible. Gran parte de las reacciones químicas a nivel celular en las plantas y animales dependen directamente de la presencia de enzimas específicas y en la disponibilidad de una fuente de energía externa. A las sustancias que son introducidas en un medio de crecimiento y que incrementan la cantidad de enzimas son conocidas como inductores, y sus enzimas correspondientes son llamadas enzimas inducibles. Dichas enzimas están presentes cuando se necesitan para sintetizar constituyentes celulares.

La figura 5.1 muestra una descripción gráfica del modelo de síntesis enzimática inducible. Las variables x_i , $i = 1, \dots, 12$ representan las concentraciones de los intermedios involucrados en la síntesis enzimática inducible; Los constantes k_j , $j = 1, \dots, 15$ son los parámetros cinéticos del sistema. En este trabajo, $k_j = a_j$. Las flechas indican la dirección de la reacción.

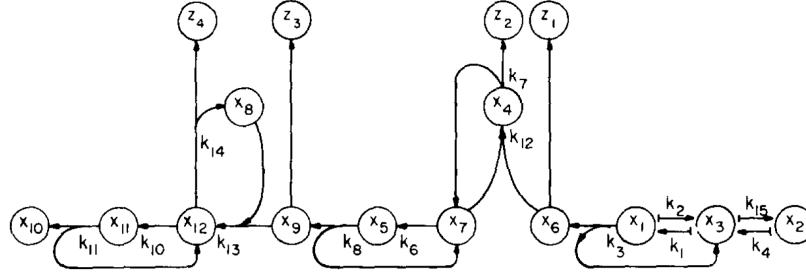


Figura 5.1: Modelo de la síntesis enzimática inducible.

Observación 2. Las componentes z_1, z_2, z_3 y z_4 no aparecen en el sistema de ecuaciones que se presentará más adelante. Sin embargo, describimos brevemente su significado:

- z_1 : Productos de descomposición de la molécula represora.
- z_2 : Otro producto de descomposición de la molécula represora.
- z_3 : Producto de descomposición del mensajero ARN.
- z_4 : Producto de descomposición de ribosomas.

5.1. Descripción del Modelo

El modelo de la síntesis enzimática inducible, originalmente, fue propuesta y publicada en 1964 por F. Heinmets [9]. Luego, en 1969, R.S.Roth y M.M.Roth [10] hicieron modificaciones en el trabajo original, donde se modificó la notación en sistema y actualmente el sistema de ecuaciones a considerar será:

$$\begin{aligned}
 \dot{x}_1 &= a_1 x_3 - a_2 x_1 - a_3 x_1 \\
 \dot{x}_2 &= a_{15} x_3 - a_4 x_2 \\
 \dot{x}_3 &= -a_1 x_3 + a_2 x_1 + a_3 x_1 - a_{15} x_3 + a_4 x_2 \\
 \dot{x}_4 &= a_{12} x_6 x_7 - a_7 x_4 \\
 \dot{x}_5 &= a_6 x_7 - a_8 x_5 \\
 \dot{x}_6 &= a_3 x_1 - a_5 x_6 - a_{12} x_6 x_7 \\
 \dot{x}_7 &= -a_{12} x_6 x_7 + a_7 x_4 - a_6 x_7 + a_8 x_5 \\
 \dot{x}_8 &= a_{14} x_{12} - a_{13} x_8 x_9 \\
 \dot{x}_9 &= a_8 x_5 - a_9 x_9 - a_{13} x_8 x_9 \\
 \dot{x}_{10} &= a_{11} x_{11} \\
 \dot{x}_{11} &= a_{10} x_{12} - a_{11} x_{11} \\
 \dot{x}_{12} &= a_{13} x_8 x_9 - a_{10} x_{12} + a_{11} x_{11} - a_{14} x_{12}
 \end{aligned} \tag{5.1}$$

donde las variables son:

- x_1 : Producto complejo del gen regulador y producto metabólico del metabolito represor.
- x_2 : Producto complejo del gen regulador y producto metabólico del inductor.
- x_3 : Gen regulador.
- x_4 : Gen funcional reprimido.
- x_5 : Un conjunto integrado de precursores de ARN mensajero en el gen funcional.
- x_6 : Molécula represora, sintetizada por el gen regulador y el complejo S_R . Este último complejo es un producto metabólico del metabolito represor.
- x_7 : Gen funcional para la síntesis de ARN mensajero; tiene un sitio para interactuar con el represor, lo que conduce al gen funcional reprimido (x_4).
- x_8 : Ribosomas.
- x_9 : ARN mensajero específico del gen funcional (x_7).
- x_{10} : Enzima recién formada separada de la plantilla.
- x_{11} : Un conjunto integrado de aminoácidos en la plantilla.
- x_{12} : Plantilla para la síntesis de la enzima.

y los parámetros a estimar serán:

- a_1 : Aumento de la concentración de producto complejo x_1 por unidad de actividad del gen regulador x_3 .
- a_2 : Disminución de la concentración de producto complejo x_1 por unidad de concentración.
- a_3 : Tasa a la cual una concentración unitaria de producto complejo x_1 disminuye debido a su reacción con un grupo constante para la formación del represor.
- a_4 : Tasa a la que se descompone la concentración unitaria del producto complejo x_2 .
- a_5 : Tasa de disminución de una concentración unitaria de la molécula represora x_6 debido a la descomposición.
- a_6 : Tasa de aumento de los precursores x_5 debido a una actividad unitaria del gen funcional x_7 .

- a_7 : Tasa de disminución de la actividad del gen reprimido no-funcional recién formado x_4 .
- a_8 : Tasa de disminución de una concentración unitaria de precursores x_5 que forman ARN mensajero x_9 .
- a_9 : Tasa de disminución de la concentración de ARN mensajero x_9 por disociación.
- a_{10} : Tasa de aumento de la concentración de la plantilla del aminoácido complejo x_{11} debido a una concentración unitaria de la plantilla x_{12} .
- a_{11} : Tasa de aumento de la concentración de la enzima x_{10} debido a una concentración unitaria de la plantilla x_{11} .
- a_{12} : Tasa de aumento de la actividad del gen funcional reprimido x_4 debido a la reacción combinada de una concentración unitaria de la molécula represora x_6 y una concentración unitaria del gen funcional activo x_7 .
- a_{13} : Tasa de aumento de la concentración de la plantilla x_{12} debido a las reacciones combinadas de una concentración unitaria de ribosomas x_8 y una concentración unitaria de ARN mensajero x_9 .
- a_{14} : Tasa de aumento de la concentración de ribosomas x_8 debido a la descomposición de una concentración unitaria de la plantilla x_{12} .
- a_{15} : Tasa de aumento de la concentración de producto complejante x_2 cuando la actividad del gen regulador x_3 tiene un valor de la unidad (concentración del inductor externo).

A grandes rasgos, tal como se estableció anteriormente, este modelo establece que, a nivel celular, tanto plantas como animales, tienen la capacidad de “encender” o “apagar” partes de las funciones metabólicas de acuerdo con las necesidades. Para lograr esto se producen enzimas que son las responsables de crear productos metabólicos importantes para la célula. Las enzimas con las características previamente mencionadas son las *enzimas inducibles*. Para la producción de estas enzimas se necesitan células de ADN y sus mecanismos para sintetizar proteínas.

El proceso que se realiza en el sistema presentado establece que para cada enzima inducible existe un gen regulador (x_3) y un gen funcional (x_6). El gen funcional contiene la información genética para la construcción de la enzima. El gen regulador se encarga de encender o apagar el gen funcional.

Para cada enzima inducible, en el cromosoma bacteriano hay un gen regulador (x_3) y un gen funcional (x_7), los cuáles interactúan. Los complejos reversibles (x_1 y x_2) se forman entre el gen regulador (x_3) y entre S_R (producto metabólico de s_x , donde s_x es el metabolito represor) e i (producto metabólico de S_i , donde S_i es el inductor) respectivamente.

Solo x_1 es funcionalmente activo e interactúa con el grupo p_I (grupo de formación represora) dando como resultado una molécula represora (x_6). Los represores tienen un tiempo de vida limitado y se descompone en productos desconocidos (llamado z_1). Las moléculas represoras (x_6) forman ahora un complejo irreversible con el gen funcional (x_7), y el nuevo gen reprimido (x_4) el cuál es no funcional. La represión es liberada por disociación del represor del complejo (x_4) hacia un producto no-funcional (llamado z_2). El gen funcional x_7 interactúa con el grupo p_M (grupo de formación mensajera) y da como resultado un mensajero de ARN (x_9) el cuál es específico para el gen funcional (x_7). Ahora, x_9 tiene también un tiempo de vida finito y se descompone en z_3 . También interactúa con los ribosomas (x_8) dando espacio a una plantilla x_{12} , cuya especificidad es determinada por el gen funcional x_7 . Después se transportan los aminoácidos y su interacción con la enzima activadora, sin embargo en el modelo de interés este proceso no está modelado. Finalmente, el conjunto integrado de aminoácidos en la plantilla x_{11} disociará en la plantilla x_{12} y también en una enzima unitaria x_{10} . La polimerización en varias unidades de x_{10} podrán ser requeridas para la producción de una enzima activa. La plantilla x_{12} tiene un tiempo de vida finito y se descompone en los ribosomas x_8 y en un producto de alguna reacción z_4 (producto de descomposición de ribosomas). Del esquema del proceso descrito se desarrolla el sistema de ecuaciones y por tanto el sistema de interés.

5.2. Resultados numéricos para el modelo de la síntesis enzimática inducible

Los resultados numéricos obtenidos para este modelo se obtienen sobre la base de las siguientes consideraciones:

1. En los trabajos mencionados [8, 9, 10] no se hace una distinción entre las variables observables y no observables. Se muestran algunas gráficas en ciertos artículos donde las variables a las cuáles le hacen énfasis particular son x_6, x_9, x_{10} . Sin embargo, para fines de este trabajo, se estarán alternando las variables observables utilizadas y se mostrarán dichos resultados.
2. Los parámetros presentados en el artículo de F. Heinmets [9] se presentarán como la solución teórica. Es decir, se considerarán:

Valores exactos de las tasas a_i en el modelo dinámico								
Componente	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8
Valor teórico	5.0	0.9	0.1	0.1	0.035	0.006	0.1	0.1
Componente	a_9	a_{10}	a_{11}	a_{12}	a_{13}	a_{14}	a_{15}	
Valor teórico	0.005	0.1	0.05	0.2	0.005	0.02	1.0	

3. Las condiciones iniciales presentadas en los artículos de F. Heinmets [9] y J. Milstein [8] son las siguientes:

Condiciones iniciales exactas $x_i(0)$ en el modelo dinámico												
Variable	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}
Heinmets	0.0	0.0	1.0	0.0	0.0	0.0	1.0	1.0	0.0	0.0	0.0	0.0
Milstein	0.5	0.1	0.25	0.75	0.0	0.5	0.1	1.0	0.75	0.0	1.0	0.0

Por lo tanto, este capítulo se incluyen dos bloques de resultados, uno por cada conjunto de condiciones iniciales. La dinámica de las soluciones ‘exactas’ sobre el intervalo de tiempo $0 \leq t \leq 20$ para las condiciones iniciales de Heinmets y Milstein se muestran en la Figura 5.2.

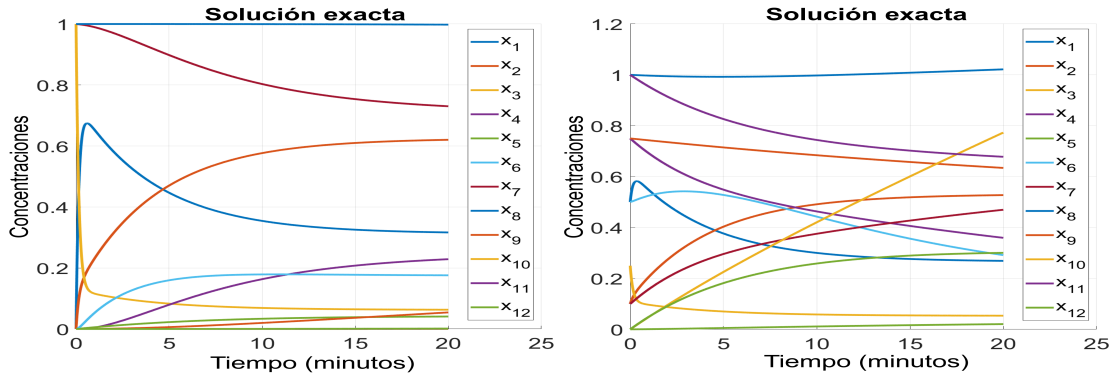


Figura 5.2: Soluciones con condiciones iniciales de Heinmets (izquierda) y de Milstein (derecha).

Para validar la metodología propuesta en el capítulo 3 se presentarán resultados numéricos con datos sintéticos (‘medidas experimentales’). Los datos sintéticos se generan perturbando aleatoriamente las soluciones exactas con ruido del 5 %, utilizando la distribución uniforme, como se reporta en la referencia [8]. Se generaron $m = 51$ mediciones experimentales con variables observables en los tiempos $t_i = 0 + i h$, $0 \leq i \leq 10$ (minutos), con $h = 0.2$. El objetivo es recuperar, a partir de los datos perturbados los 12 valores de condiciones iniciales $\mathbf{s}_0$ y los 15 valores del vector de parámetros $\boldsymbol{\theta}$ en forma razonable.

5.2.1. Resultados numéricos con condiciones iniciales de Heinmets.

A partir de los $m = 51$ datos experimentales se ajustan curvas polinomiales con el objeto de extraer datos adicionales sobre una malla más fina, con paso de tiempo $h = 0.01$. Hay razones importantes para introducir datos adicionales sobre las curvas ajustadas:

1. Los datos sobre curvas suaves ajustadas permiten evitar los saltos puntuales en el término (discontinuo), asociado a la sumatoria, en la ecuación adjunta (ver ecuación (3.7)).

2. Datos adicionales sobre una malla fina permiten mayor precisión en la solución numérica de las ecuaciones de estado y adjunta.
3. Es posible generar más subintervalos para aplicar el método de disparo múltiple.

Los grados de los polinomios ajustados a los datos de cada variable son:

Grados utilizados para el ajuste de polinomios												
Variable	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}
Grado	19	9	19	3	2	2	3	1	1	1	1	1

Las curvas ajustadas, junto a los datos experimentales originales (con ruido) se muestran en la Figura 5.3.

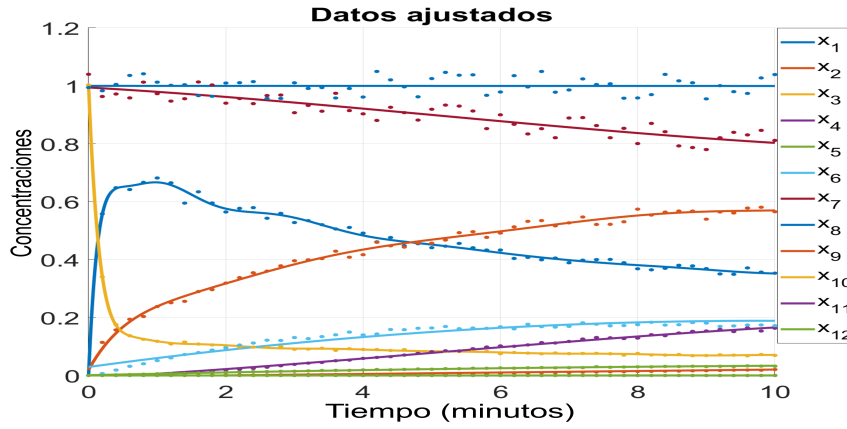


Figura 5.3: Datos con nivel de 5 % de ruido y las curvas ajustadas.

Sobre el intervalo de tiempo $[0, 10]$, con la malla fina con paso $h = 0.01$, se obtiene un conjunto de $m = 10/h = 1001$ datos para aplicar el método de disparo múltiple en la estimación de las condiciones iniciales y los parámetros del sistema de ecuaciones. Para el algoritmo de optimización BFGS se proponen los siguientes valores del vector θ^0 para iniciar las iteraciones. Recordar que $\theta_i^0 = a_i$, con $i = 1, \dots, 15$.

θ_1^0	θ_2^0	θ_3^0	θ_4^0	θ_5^0	θ_6^0	θ_7^0	θ_8^0
4.9661	0.7674	0.1112	0.0957	0.0278	0.0080	0.0848	0.1470
θ_9^0	θ_{10}^0	θ_{11}^0	θ_{12}^0	θ_{13}^0	θ_{14}^0	θ_{15}^0	
0.0045	0.1024	0.0505	0.2145	0.0070	0.0175	1.4680	

Por otro lado, los valores iniciales de los parámetros de disparo, $s_1^0, \dots, s_{ns+1}^0$, en el algoritmo de optimización BFGS, se toman de los datos experimentales con ruido (en los tiempos de disparo correspondientes $\tau_i, \dots, \tau_{ns+1}$). Para el método de la secante en la búsqueda

de línea (3.10) se consideran los mismos valores iniciales ρ_0 y ρ_1 tomados en la sección 4.1. La tolerancia, como criterio de paro del algoritmo BFGS, se fija en $\epsilon = 10^{-4}$. Para resolver las ecuaciones de estado (3.2)-(3.3) en cada subintervalo, así como las ecuaciones adjuntas correspondientes (3.7), es suficiente con utilizar el método de Runge-Kutta de orden 2 (método de Heun o del trapecio).

Las tablas 5.1 y 5.2 muestran las estimaciones obtenidas para θ y $\mathbf{s}_0$, cuando se consideran dos variables observables, x_7, x_8 , y diferente número de subintervalos de disparo ns en el intervalo de tiempo $[0, 10]$.

Tabla 5.1: Estimación de θ con diferente ns , 5 % de ruido, y x_7, x_8 como observables.

ns	θ			Error relativo calculado		
	10	25	100	10	25	100
	4.9737	4.9779	4.9663	0.0053	0.0044	0.0067
	0.7405	0.7358	0.7987	0.1772	0.1824	0.1126
	0.0827	0.0766	0.0841	0.1727	0.2336	0.1592
	0.1441	0.1366	0.1464	0.4408	0.3655	0.4636
	0.0096	0	0.0193	0.7264	1.0000	0.4493
	0.0071	0.0070	0.0075	0.1886	0.1742	0.2548
	0.0808	0.0665	0.0970	0.1923	0.3353	0.0301
	0.1481	0.1493	0.1464	0.4812	0.4929	0.4641
	0.0050	0.0061	0.0050	0.0023	0.2119	0.0025
	0.1024	0.1024	0.1024	0.0244	0.0244	0.0244
	0.0505	0.0505	0.0505	0.0104	0.0104	0.0104
	0.1868	0.1751	0.1909	0.0661	0.1244	0.0455
	0.0061	0.0059	0.0073	0.2148	0.1818	0.4687
	0.0175	0.0175	0.0175	0.1238	0.1226	0.1250
	1.4315	1.3950	1.3662	0.4315	0.3950	0.3662
iters	26	100	100	-	-	-

Tabla 5.2: Estimación de $\mathbf{s}_0$ con diferentes ns , 5 % de ruido, x_7, x_8 variables observables.

ns	$\mathbf{s}_0$			Error relativo calculado		
	10	25	100	10	25	100
	0.0160	0.0279	0.0442	0.0160	0.0279	0.0442
	0	0.0002	0	0	0.0002	0
	0.9907	0.9704	0.9778	0.0093	0.0296	0.0222
	0.0023	0.0034	0	0.0023	0.0034	0
	0	0	0.0001	0	0	0.0001
	0.0267	0.0320	0.0202	0.0267	0.0320	0.0202
	0.9936	0.9940	0.9941	0.0064	0.0060	0.0059
	0.9997	0.9998	0.9998	0.0003	0.0002	0.0002
	0	0	0	0	0	0
	0	0	0	0	0	0
	0	0	0	0	0	0
	0	0	0	0	0	0
iters	26	100	100	-	-	-

El mejor resultado simultaneo para ambos, θ y $\mathbf{s}_0$, se obtiene cuando se utilizan $ns = 100$ subintervalos. La estimación es mejor para las condiciones iniciales $\mathbf{s}_0$ que para los parámetros θ . Para apreciar mejor los resultados se incluye la figura 5.4 para el caso $ns = 10$. La gráfica izquierda muestra el descenso del gradiente relativo, $\log(\|\mathbf{g}^l\|/\|\mathbf{g}\|^0)$ vs las iteraciones; la gráfica al centro muestra las curvas de convergencia de los parámetros

vs las iteraciones; la gráfica derecha muestra la solución exacta (en línea continua) junto a la solución aproximada (en colores sólidos) en el intervalo $0 \leq t \leq 10$, y el pronóstico en el intervalo $10 \leq t \leq 20$ (en línea punteada).

Las gráficas análogas obtenidas cuando se utilizan $ns = 25$ y 100 subintervalos se muestran en las figuras 5.5, 5.6, respectivamente.

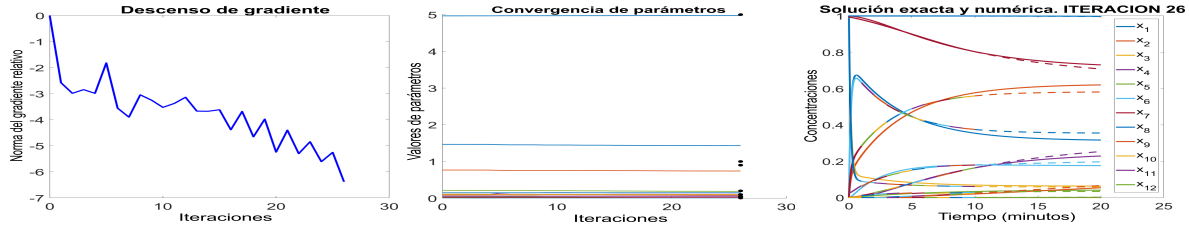


Figura 5.4: Resultados con $ns = 10$. Nivel de ruido 5 %, variables observables x_7, x_8 .

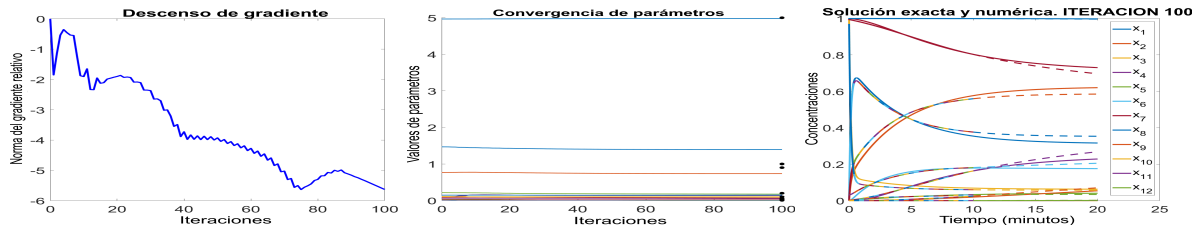


Figura 5.5: Resultados con $ns = 25$. Nivel de ruido 5 %, variables observables x_7, x_8 .

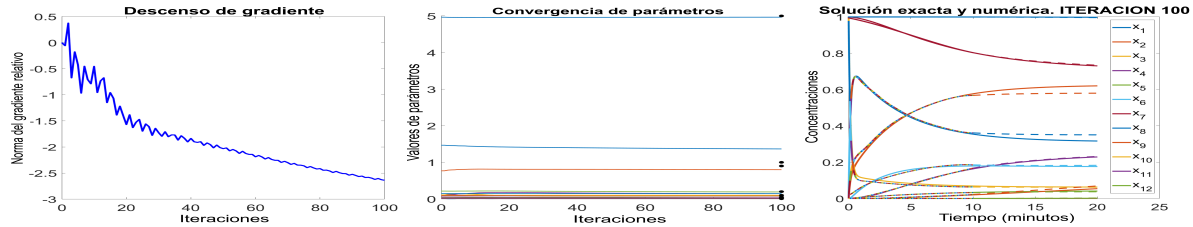


Figura 5.6: Resultados con $ns = 100$. Nivel de ruido 5 %, variables observables x_7, x_8 .

Es posible utilizar diferentes conjuntos de variables observables, como se muestra en las tablas 5.3 y 5.4. Se muestran resultados para cuatro conjuntos diferentes de variables observables. Los resultados en dicha tabla muestran que se obtienen resultados satisfactorios incluso con una sola variable observable, siendo x_8 aquella que proporciona los mejores resultados. De manera similar al caso del modelo SEIRD, la variable x_8 juega un papel preponderante en la determinación de los parámetros θ y de la condición inicial s_0 , al menos en este caso.

Finalmente, tomando x_8 como única variable observable, se realizan diferentes experimentos numéricos considerando ‘datos experimentales’ en intervalos de tiempo crecientes. En las tablas 5.5 y 5.6 se reportan las estimaciones obtenidas para los parámetros θ , y las condiciones iniciales s_0 , en los subintervalos de tiempo $[0, 5]$, $[0, 7]$, $[0, 11]$, $[0, 14]$ con 501, 701, 1101 y 1401 ‘datos experimentales’, respectivamente. El número de subintervalos

Tabla 5.3: Estimación de θ usando distintas variables observables. $ns = 100$.

V. Obs	θ				Error relativo calculado			
	x_7	x_8	x_3, x_7, x_8	x_2, x_3, x_7, x_8	x_7	x_8	x_3, x_7, x_8	x_2, x_3, x_7, x_8
—	4.9663	4.9673	4.9200	4.9278	0.0067	0.0065	0.0160	0.0144
	0.7986	0.7950	0.8650	0.8092	0.1126	0.1166	0.0388	0.1009
	0.0841	0.0844	0.1739	0.1286	0.1593	0.1561	0.7391	0.2859
	0.1464	0.1450	0.1202	0.0407	0.4636	0.4502	0.2024	0.5932
	0.0192	0.0155	0.0360	0.0328	0.4506	0.5584	0.0278	0.0641
	0.0075	0.0070	0.0038	0.0003	0.2533	0.1695	0.3692	0.9467
	0.0969	0.0948	0.0864	0.0878	0.0307	0.0521	0.1358	0.1220
	0.1464	0.1455	0.1472	0.1475	0.4640	0.4553	0.4716	0.4747
	0.0050	0.0051	0.0045	0.0045	0.0027	0.0181	0.0938	0.0921
	0.1024	0.1024	0.1024	0.1024	0.0244	0.0244	0.0244	0.0244
	0.0505	0.0505	0.0505	0.0505	0.0104	0.0104	0.0104	0.0104
	0.1909	0.1913	0.2197	0.2150	0.0454	0.0437	0.0987	0.0749
	0.0074	0.0074	0.0070	0.0070	0.4869	0.4790	0.3963	0.3972
	0.0175	0.0175	0.0175	0.0175	0.1251	0.1250	0.1251	0.1251
	1.3661	1.3595	1.4019	0.5608	0.3661	0.3595	0.4019	0.4392
iters	100	100	59	99	-	-	-	-

Tabla 5.4: Estimación de s_0 con distintas variables observables. $ns = 100$.

V. Obs	s_0				Error relativo calculado			
	x_7	x_8	x_3, x_7, x_8	x_2, x_3, x_7, x_8	x_7	x_8	x_3, x_7, x_8	x_2, x_3, x_7, x_8
—	0.0442	0.0455	0.0406	0.0063	0.0442	0.0455	0.0406	0.0063
	0	0	0	0.0174	0	0	0	0.0174
	0.9778	0.9755	1.0216	0.9811	0.0222	0.0245	0.0216	0.0189
	0	0	0	0	0	0	0	0
	0.0001	0.0006	0.0013	0.0011	0.0001	0.0006	0.0013	0.0011
	0.0202	0.0186	0.0270	0.0277	0.0202	0.0186	0.0270	0.0277
	0.9941	0.9932	0.9939	0.9921	0.0059	0.0068	0.0061	0.0079
	0.9997	0.9998	0.9998	0.9997	0.0003	0.0002	0.0002	0.0003
	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0
iters	100	100	59	99	-	-	-	-

ns para aplicar el método de disparo múltiple se incrementa con el objeto de que cada subintervalo de disparo tenga el mismo número de datos y la comparación sea justa.

De nuevo, las estimaciones obtenidas para las condiciones iniciales s_0 son mejores que para los parámetros θ . Esta discrepancia, al final, produce un balance en el resultado global de la dinámica, al resultar en ajuste satisfactorio, sobre todo cuando se cuenta con suficientes datos experimentales, como se muestra en la última gráfica de la figura 5.7.

Tabla 5.5: Estimación de θ vs número de datos. Variable observable, x_8 .

Intervalo	θ				Error relativo calculado			
	[0, 5]	[0, 7]	[0, 11]	[0, 14]	[0, 5]	[0, 7]	[0, 11]	[0, 14]
<i>ns</i>	10	14	11	20	10	14	11	20
—	4.9753	4.9758	4.9744	4.9772	0.0049	0.0048	0.0051	0.0046
	0.7399	0.7404	0.7376	0.7284	0.1779	0.1774	0.1804	0.1906
	0.0830	0.0871	0.0909	0.0910	0.1698	0.1294	0.0910	0.0903
	0.1681	0.1579	0.1429	0.1323	0.6808	0.5789	0.4294	0.3230
	0.0260	0.0205	0.0155	0.0187	0.2564	0.4138	0.5560	0.4671
	0.0056	0.0059	0.0062	0.0059	0.0735	0.0162	0.0291	0.0084
	0.0863	0.0880	0.0891	0.0960	0.1367	0.1196	0.1085	0.0395
	0.1469	0.1464	0.1428	0.1345	0.4690	0.4637	0.4282	0.3449
	0.0045	0.0047	0.0062	0.0105	0.0901	0.0551	0.2451	1.0969
	0.1024	0.1024	0.1024	0.1024	0.0244	0.0244	0.0243	0.0242
	0.0505	0.0505	0.0505	0.0505	0.0104	0.0104	0.0104	0.0104
	0.2084	0.2009	0.1966	0.1981	0.0419	0.0043	0.0171	0.0095
	0.0070	0.0064	0.0020	0.0010	0.3982	0.2851	0.6057	0.7924
	0.0175	0.0175	0.0177	0.0178	0.1251	0.1245	0.1162	0.1095
	1.4195	1.4153	1.4290	1.4102	0.4195	0.4153	0.4290	0.4102
iters	100	100	88	100	-	-	-	-

Tabla 5.6: Estimación de s_0 vs número de datos. Variable observable x_8 .

Intervalo	s_0				Error relativo calculado			
	[0, 5]	[0, 7]	[0, 11]	[0, 14]	[0, 5]	[0, 7]	[0, 11]	[0, 14]
<i>ns</i>	10	14	11	20	10	14	11	20
—	0.0208	0.0220	0.0137	0.0183	0.0208	0.0220	0.0137	0.0183
	0	0	0.0001	0	0	0	0.0001	0
	0.9830	0.9797	0.9865	0.9773	0.0170	0.0203	0.0135	0.0227
	0	0.0005	0.0012	0.0022	0	0.0005	0.0012	0.0022
	0.0021	0.0018	0	0	0.0021	0.0018	0	0
	0.0223	0.0180	0.0148	0.0173	0.0223	0.0180	0.0148	0.0173
	0.9943	0.9947	0.9959	0.9954	0.0057	0.0053	0.0041	0.0046
	0.9998	0.9997	0.9997	0.9997	0.0002	0.0003	0.0003	0.0003
	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0
	0	0	0.0001	0	0	0	0.0001	0
	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0
iters	100	100	88	100	-	-	-	-

5.2.2. Resultados numéricos con condiciones iniciales de Milstein.

En esta subsección se tomarán en cuenta las condiciones iniciales propuestas por Milstein. A partir del conjunto de $m = 51$ datos experimentales, se ajustan curvas polinomiales para nuevamente extraer datos adicionales sobre una malla más fina, con paso de tiempo de $h = 0.01$. Los grados de dichos polinomios ajustados a los datos de cada variable se muestran a continuación.

Grados utilizados para el ajuste de polinomios												
Variable	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}
Grado	17	9	19	3	4	3	3	1	1	1	2	2

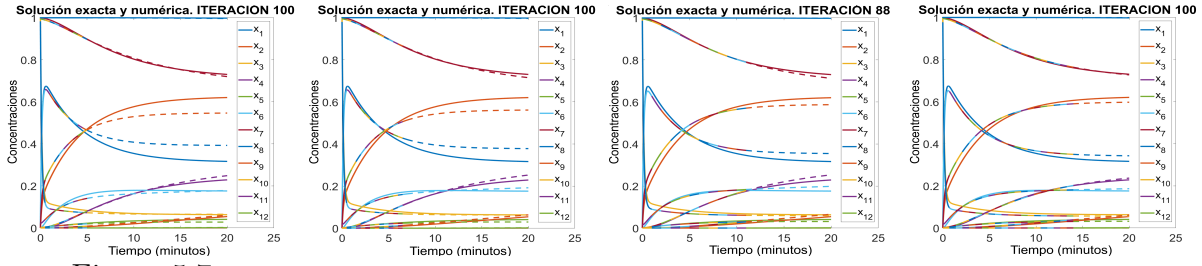


Figura 5.7: Solución exacta, calculada y pronóstico. $i = 0 : 5, i = 0 : 7, i = 0 : 11, i = 0 : 14$.

Las curvas ajustadas, junto a los datos experimentales originales (con ruido) se muestran en la Figura 5.8.

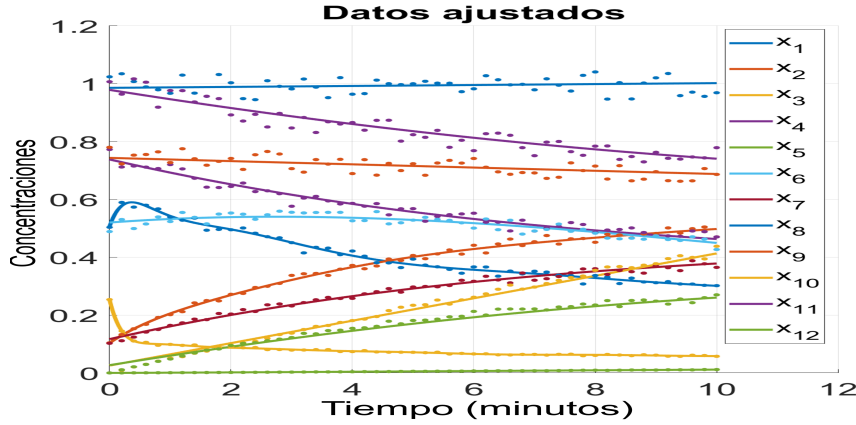


Figura 5.8: Datos con nivel de 5 % de ruido y las curvas ajustadas.

Tal como se procedió anteriormente, se generan datos adicionales con mismo espaciado de $h = 0.01$, de modo que también se estarán trabajando con $m = 1001$ datos para poder aplicar el método de disparo múltiple en la estimación de θ y s_0 . En esta ocasión se tomará el valor inicial θ^0 , dado a continuación, para inicializar las iteraciones en el algoritmo BFGS.

θ_1^0	θ_2^0	θ_3^0	θ_4^0	θ_5^0	θ_6^0	θ_7^0	θ_8^0
5.2909	0.7063	0.0943	0.1250	0.0399	0.0081	0.1479	0.1330
θ_9^0	θ_{10}^0	θ_{11}^0	θ_{12}^0	θ_{13}^0	θ_{14}^0	θ_{15}^0	
0.0062	0.0581	0.0315	0.2674	0.0034	0.0205	1.4344	

Bajo los mismos supuestos establecidos en la sección 5.2.1, pero ahora tomando como variables observables x_1, x_2, x_3 , la tabla 5.7 muestra las aproximaciones obtenidas para θ , mientras que la tabla 5.8 muestra las aproximaciones obtenidas para la condición inicial s_0 .

Se observa que el número de subintervalos que ofrece el error relativo menor en la aproximación de θ y de s_0 de forma simultánea es $ns = 50$. De nuevo la aproximación de la

Tabla 5.7: Estimación de θ con diferentes ns . Valores observables: x_1, x_2, x_3 .

ns	θ			Error relativo calculado		
	10	25	50	10	25	50
	5.0988	5.0171	4.7201	0.0198	0.0034	0.0560
	0.8311	0.8301	0.7953	0.0766	0.0776	0.1164
	0.1807	0.1641	0.1404	0.8066	0.6407	0.4043
	0.0921	0.0850	0.0826	0.0790	0.1498	0.1739
	0.0794	0.0714	0.0562	1.2697	1.0390	0.6053
	0.0095	0.0107	0.0072	0.5810	0.7769	0.2065
	0.1168	0.1111	0.1083	0.1676	0.1107	0.0828
	0.1334	0.1336	0.1339	0.3342	0.3358	0.3387
	0.0067	0.0075	0.0047	0.3465	0.4967	0.0676
	0.0827	0.0737	0.0697	0.1726	0.2633	0.3026
	0.0453	0.0443	0.0436	0.0937	0.1141	0.1271
	0.2688	0.2435	0.2377	0.3442	0.2173	0.1884
	0.0023	0.0009	0.0037	0.5306	0.8184	0.2515
	0.0177	0.0151	0.0245	0.1159	0.2459	0.2259
	0.9656	0.9100	0.8791	0.0344	0.0900	0.1209
iters	47	89	135	-	-	-

Tabla 5.8: Estimación de s_0 con diferentes ns . Variables observables x_1, x_2, x_3 .

ns	s_0			Error relativo calculado		
	10	25	50	10	25	50
	0.4932	0.4967	0.5014	0.0136	0.0066	0.0028
	0.0979	0.0988	0.0999	0.0213	0.0117	0.0009
	0.2596	0.2573	0.2550	0.0383	0.0291	0.0200
	0.7488	0.7577	0.7546	0.0015	0.0102	0.0061
	0	0	0	0	0	0
	0.4893	0.4840	0.4942	0.0215	0.0320	0.0117
	0.1023	0.1011	0.1025	0.0233	0.0112	0.0249
	0.9883	0.9891	0.9898	0.0117	0.0109	0.0102
	0.7442	0.7450	0.7444	0.0077	0.0067	0.0074
	0.0266	0.0222	0.0232	0.0266	0.0222	0.0232
	0.9986	0.9891	0.9851	0.0014	0.0109	0.0149
	0.0142	0.0172	0.0186	0.0142	0.0172	0.0186
iters	47	89	135	-	-	-

condición inicial s_0 es mejor que la aproximación a los valores de los parámetros. Con el objeto de obtener una mejor apreciación global del efecto en la dinámica del modelo incluimos la figura 5.9. Se muestran las gráficas de $\log(\|\mathbf{g}^l\|/\|\mathbf{g}\|^0)$ vs número de iteraciones (izquierda), la convergencia de los parámetros (centro) y el comportamiento de las variables de estado en el transcurso de tiempo (derecha). Nuevamente, las líneas de varios colores representan la dinámica reconstruida con la condición inicial calculada s_0 y los parámetros estimados θ en el intervalo de tiempo $[0, 10]$, mientras que las líneas punteadas muestran la predicción a futuro en el intervalo $[10, 20]$ en conjunto con la solución exacta (líneas continuas).

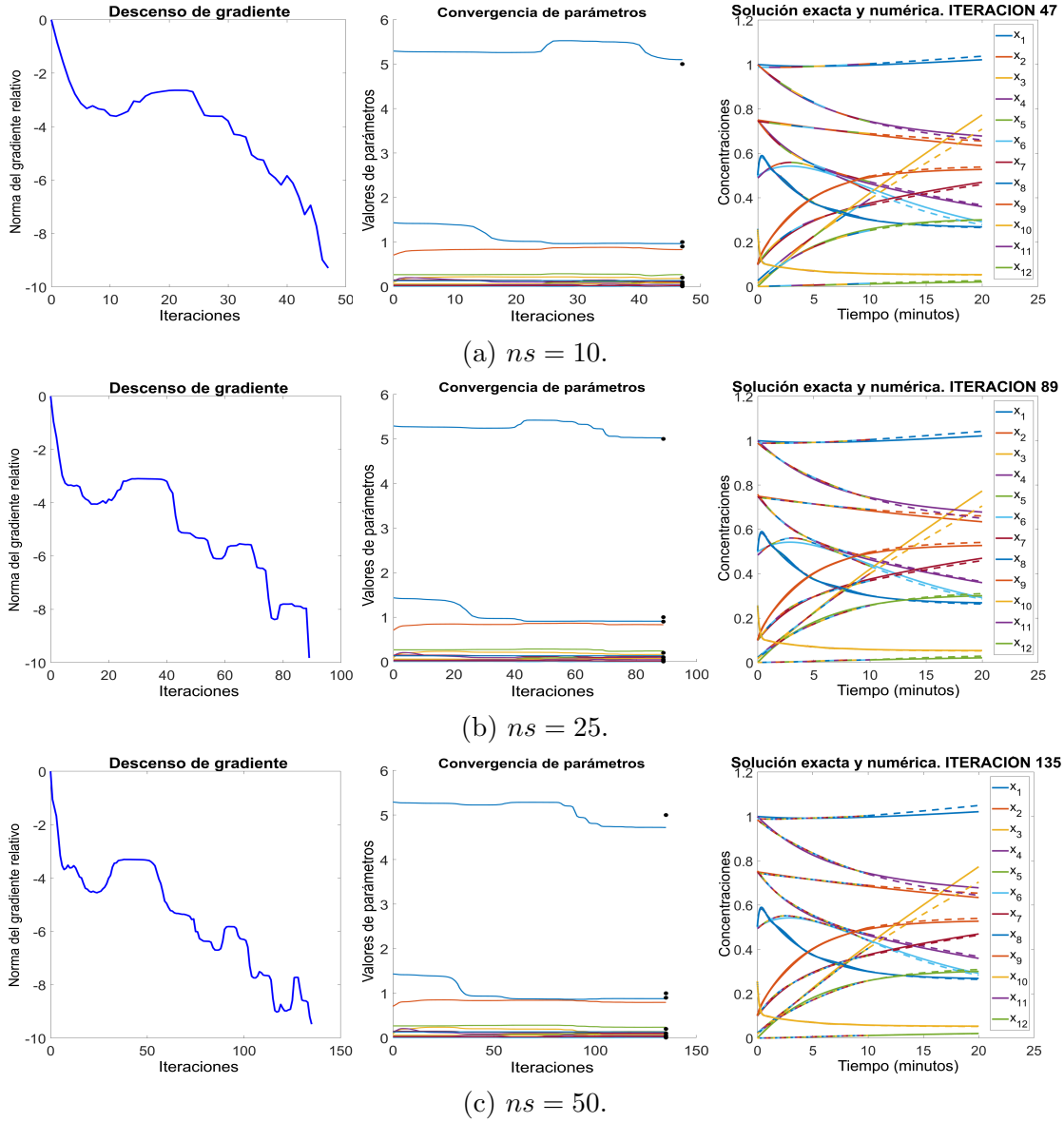


Figura 5.9: Descenso del gradiente (izquierda), convergencia de los parámetros (centro), y comparación de la dinámica exacta con la calculada (derecha). Casos: $ns = 10, 25, 50$.

Realizaremos experimentos numéricos similares a los del caso estudiado en la sección 5.2.1. Para ello primero consideramos el caso con $ns = 50$ subintervalos y distintos conjuntos de variables observables. Los resultados obtenidos se resumen en las tablas 5.9 y 5.10. Los resultados obtenidos muestran que la mejor estimación conjunta de θ y s_0 se obtiene cuando las variables observables son $x_1, x_2, x_3, x_9, x_{11}$. Hay una mejora en este caso con respecto a los resultados mostrados en las tablas 5.7 y 5.8, donde las variables observables son x_1, x_2, x_3 . Los resultados para estimar s_0 son muy satisfactorios, pero la estimación de θ sigue siendo regular, a pesar de la pequeña mejora. Por otro lado, los resultados obtenidos con una sola variable observable, aunque no son los mejores, sí muestran potencial

Tabla 5.9: Estimación de θ con distintas variables observables. $ns = 50$.

V. Obs	θ				Error relativo calculado			
	x_1	x_3	x_2, x_3	$x_1, x_2, x_3, x_9, x_{11}$	x_1	x_3	x_2, x_3	$x_1, x_2, x_3, x_9, x_{11}$
—	5.2474	4.2727	4.3491	4.7263	0.0495	0.1455	0.1302	0.0547
	0.6592	0.8169	0.8290	0.7956	0.2676	0.0923	0.0789	0.1160
	0.0591	0.1504	0.1459	0.1413	0.4089	0.5040	0.4591	0.4128
	0.0854	0.0593	0.0601	0.0827	0.1465	0.4073	0.3993	0.1733
	0.0004	0.0544	0.0605	0.0553	0.9894	0.5533	0.7291	0.5790
	0.0060	0.0063	0.0122	0.0018	0.0004	0.0436	1.0298	0.6959
	0.1021	0.1041	0.0975	0.1089	0.0206	0.0409	0.0247	0.0885
	0.1339	0.1334	0.1338	0.1326	0.3390	0.3343	0.3380	0.3264
	0.0090	0.0075	0.0085	0.0073	0.8069	0.4965	0.7089	0.4597
	0.0800	0.0792	0.0874	0.0537	0.2004	0.2077	0.1262	0.4631
	0.0450	0.0450	0.0458	0.0391	0.0993	0.0997	0.0834	0.2185
	0.2151	0.2270	0.1952	0.2373	0.0755	0.1352	0.0240	0.1866
	0	0.0014	0	0.0015	1.0000	0.7227	1.0000	0.6982
	0.0072	0.0145	0.0085	0.0135	0.6415	0.2731	0.5732	0.3256
	1.3304	0.6950	0.8214	0.8795	0.3304	0.3050	0.1786	0.1205
iters	115	156	175	140	-	-	-	-

Tabla 5.10: Estimación de s_0 con distintas variables observables. $ns = 50$.

V. Obs	s_0				Error relativo calculado			
	x_1	x_3	x_2, x_3	$x_1, x_2, x_3, x_9, x_{11}$	x_1	x_3	x_2, x_3	$x_1, x_2, x_3, x_9, x_{11}$
—	0.5028	0.3543	0.3937	0.5013	0.0057	0.2914	0.2126	0.0026
	0.1107	0.1409	0.1008	0.0999	0.1068	0.4089	0.0077	0.0010
	0.2109	0.2548	0.2550	0.2551	0.1564	0.0191	0.0200	0.0204
	0.7544	0.7491	0.7514	0.7572	0.0059	0.0012	0.0018	0.0097
	0	0	0	0.0006	0	0	0	0.0006
	0.5121	0.5084	0.4967	0.4915	0.0242	0.0169	0.0066	0.0171
	0.1012	0.1056	0.1098	0.0998	0.0118	0.0555	0.0976	0.0021
	0.9877	0.9899	0.9892	0.9889	0.0123	0.0101	0.0108	0.0111
	0.7452	0.7448	0.7459	0.7435	0.0064	0.0069	0.0055	0.0087
	0.0200	0.0213	0.0173	0.0267	0.0200	0.0213	0.0173	0.0267
	0.9918	0.9933	0.9978	0.9790	0.0082	0.0067	0.0022	0.0210
	0.0168	0.0155	0.0118	0.0238	0.0168	0.0155	0.0118	0.0238
iters	115	156	175	140	-	-	-	-

para obtener resultados aceptables.

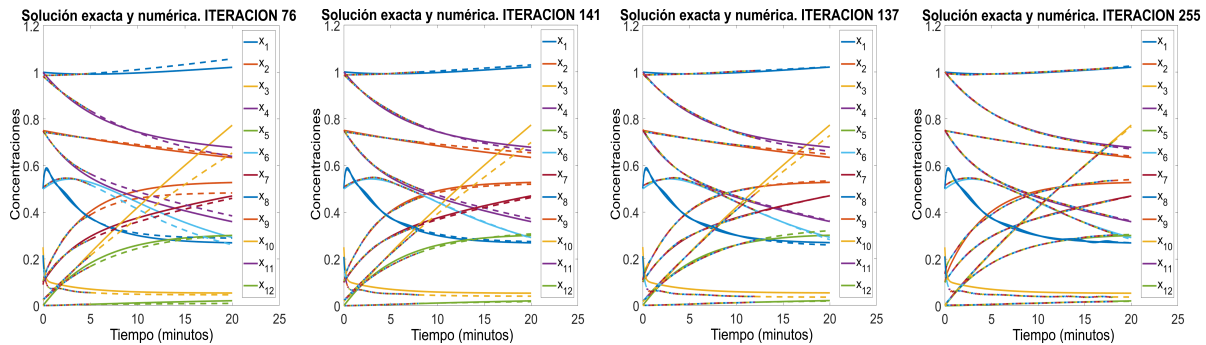
Por lo tanto, con el objeto de verificar como afecta el número de datos, consideramos una sola variable observable, digamos x_1 , y cinco subintervalos de tiempo crecientes $[0, 5]$, $[0, 8]$, $[0, 12]$, $[0, 18]$, de tal manera que el número de datos considerado es 101, 801, 1201 y 1801, respectivamente. Los resultados numéricos obtenidos para θ y s_0 se muestran en las tablas 5.11 y 5.12, respectivamente. Se observa que se puede mantener una buena aproximación con una sola variable observable, x_1 y que mientras más datos en intervalos crecientes se incorporen a la información mejor será el resultado. Para ilustrar los resultados, se muestra la figura 5.10 con gráficas de la dinámica del modelo en cada caso. Se puede observar claramente la mejora de la aproximación y el pronóstico conforme aumenta el número de datos y el horizonte de tiempo.

Tabla 5.11: Estimación de θ con número de datos creciente. Variable observable x_1

Intervalo	θ				Error relativo calculado			
	[0, 5]	[0, 8]	[0, 12]	[0, 18]	[0, 5]	[0, 8]	[0, 12]	[0, 18]
ns	50	80	60	90	50	80	60	90
—	5.2703	5.2487	5.2507	5.1429	0.0541	0.0497	0.0501	0.0286
	0.7329	0.7050	0.6614	0.6394	0.1857	0.2166	0.2651	0.2895
	0.1081	0.0772	0.0588	0.0763	0.0810	0.2280	0.4122	0.2368
	0.1316	0.1021	0.0850	0.0814	0.3158	0.0206	0.1495	0.1863
	0.0387	0.0160	0.0024	0.0159	0.1046	0.5431	0.9326	0.5465
	0.0032	0.0062	0.0064	0.0067	0.4745	0.0380	0.0666	0.1100
	0.1036	0.0996	0.1016	0.1025	0.0362	0.0044	0.0163	0.0249
	0.1332	0.1338	0.1340	0.1309	0.3319	0.3381	0.3397	0.3089
	0.0078	0.0084	0.0092	0.0037	0.5509	0.6872	0.8430	0.2524
	0.0633	0.0781	0.0797	0.0936	0.3674	0.2189	0.2029	0.0644
	0.0408	0.0438	0.0458	0.0492	0.1833	0.1250	0.0838	0.0159
	0.2587	0.2113	0.2083	0.2117	0.2934	0.0564	0.0417	0.0585
	0.0007	0	0	0.0061	0.8516	1.0000	1.0000	0.2269
	0.0183	0.0091	0.0067	0.0254	0.0836	0.5451	0.6633	0.2676
	1.3892	1.3245	1.3272	1.2287	0.3892	0.3245	0.3272	0.2287
iters	76	141	137	255	-	-	-	-

Tabla 5.12: Estimación de s_0 con número de datos creciente. Variable observable x_1 .

Intervalo	s_0				Error relativo calculado			
	[0, 5]	[0, 8]	[0, 12]	[0, 18]	[0, 5]	[0, 8]	[0, 12]	[0, 18]
ns	50	80	60	90	50	80	60	90
—	0.5041	0.5043	0.5028	0.5033	0.0083	0.0086	0.0057	0.0065
	0.0893	0.0994	0.1087	0.1368	0.1069	0.0060	0.0865	0.3680
	0.2156	0.2091	0.2110	0.2115	0.1376	0.1637	0.1561	0.1538
	0.7488	0.7513	0.7541	0.7534	0.0016	0.0018	0.0055	0.0045
	0	0	0	0.0024	0	0	0	0.0024
	0.5077	0.5085	0.5129	0.5155	0.0155	0.0170	0.0258	0.0310
	0.1066	0.1044	0.1017	0.0993	0.0661	0.0443	0.0173	0.0070
	0.9869	0.9874	0.9877	0.9950	0.0131	0.0126	0.0123	0.0050
	0.7438	0.7445	0.7457	0.7489	0.0083	0.0073	0.0058	0.0015
	0.0255	0.0225	0.0177	0.0040	0.0255	0.0225	0.0177	0.0040
	0.9835	0.9883	0.9927	1.0012	0.0165	0.0117	0.0073	0.0012
	0.0234	0.0196	0.0156	0.0032	0.0234	0.0196	0.0156	0.0032
iters	76	141	137	255	-	-	-	-

Figura 5.10: Solución exacta, aproximada y pronóstico vs número de datos creciente. De izquierda a derecha, los intervalos con datos son: $[0, 5]$, $[0, 8]$, $[0, 12]$, $[0, 18]$.

Capítulo 6

Conclusiones del trabajo

En este trabajo se presenta la técnica de disparo múltiple que, junto con métodos clásicos de optimización, se utiliza para la estimación de parámetros en ecuaciones diferenciales, partiendo de un conjunto finito de datos, o mediciones experimentales con ruido y el modelo dinámico asociado. El modelo de optimización para resolver el problema se construye a partir del Lagrangiano aumentado, que se simplifica proponiendo una variante de penalización para reforzar las restricciones de igualdad, y así evitar el cálculo de los multiplicadores de Lagrange.

El modelo de optimización obtenido es más simple y puede resolverse por medio de algoritmos iterativos de tipo cuasi-Newton clásicos, como BFGS con búsqueda de línea inexacta. De esta manera evitamos ‘resolvedores’ no lineales más sofisticados, tales como Gauss-Newton y Programación Cuadrática Secuencial, utilizados ampliamente para este tipo de problemas inversos, pero que involucran la solución de un sistema lineal en cada iteración y, por lo tanto, requieren más recursos computacionales. Por el contrario, el modelo de penalización relaja las restricciones y es suficiente con calcular de manera eficiente el gradiente de la función objetivo, lo cual se logra utilizando el método de la ecuación adjunta.

Es conveniente mencionar que el método de la ecuación adjunta, por lo general, no es utilizado en la determinación de parámetros a partir de datos con ruido, debido a dos factores principales:

1. Se cuenta con métodos clásicos conocidos como Gauss-Newton y programación cuadrática secuencial, los cuales generalmente producen resultados satisfactorios.
2. La discrepancia entre datos con ruido y los valores del modelo dinámico genera pulsos instantáneos, en los tiempos experimentales, que aparecen como término fuente en la ecuación adjunta (ver ecuación 3.7), lo cual puede ocasionar discontinuidades y/o inestabilidad numérica en el método iterativo.

En este trabajo incorporamos un método de suavización o filtrado por medio del ajuste de una curva a los datos con ruido, suprimiendo de esta manera los pulsos instantáneos. Una ventaja adicional de ‘suavizado’ es que se pueden generar más ‘datos’ a partir de la curva ajustada, lo cual además permite más flexibilidad en el número de subintervalos en el método de disparo múltiple, logrando mejores resultados de aproximación, como lo muestran los resultados obtenidos.

Para validar la estrategia y métodos propuestos, se consideraron dos modelos: el modelo SEIRD y un modelo de síntesis enzimática. Los datos ‘experimentales’ se generaron de manera sintética, con ruido gaussiano para el primer modelo y ruido distribuido uniformemente para el segundo.

Los resultados para el modelo SEIRD resultaron satisfactorios, incluso utilizando una variable observable y pocos datos experimentales. La variable de estado correspondiente al deceso de la población (D) es más significativa que otras para obtener buenos resultados. Debido a que la dinámica para esta variable es más simple que las demás, las posibles perturbaciones son menos significativas. Para este caso, la metodología propuesta estima bien θ por sobre las condiciones iniciales $\mathbf{s}_0$. Además, pronostica muy bien la dinámica en tiempos posteriores.

Por otra parte, para el modelo de la síntesis enzimática inducible los resultados son menos precisos, aunque aceptables para la complejidad del problema, que contiene 12 variables y 15 parámetros a determinar. Por los trabajos previos [7, 8, 10] se podría anticipar la dificultad para obtener resultados precisos como los obtenidos en otro tipo de ecuaciones diferenciable con menos parámetros y variables de estado. En este caso la metodología propuesta estima mejor $\mathbf{s}_0$ que los mismos parámetros, debido al número elevado de estos últimos. Afortunadamente, al igual que en el caso anterior, es posible obtener resultados adecuados con una sola variable observable, siendo la variable x_1 la más significativa. Creemos que es posible mejorar los resultados, introduciendo un modelo de optimización que considere explícitamente algunas restricciones algebraicas que ocurren en la ecuación diferencial que modela el problema.

En resumen, el método propuesto produce resultados satisfactorios, para muchos modelos, incluido el considerado en la referencia [2] para la ecuación de Lorenz en régimen caótico y otro tipo de problemas más simples. En general, el método produce mejores resultados cuando las condiciones iniciales $\mathbf{s}_0$ son conocidas o muy cercanas a las reales. Por supuesto, la naturaleza del problema también indica que algunas variables observables juegan un papel más importante que otras en la determinación de los parámetros.

Finalmente, queremos mencionar algunas perspectivas y posible trabajo futuro relacionadas con este trabajo de tesis. Algunas de ellas se incluyen con el objeto de explotar el método de disparo múltiple en combinación con métodos y problemas más generales.

- Introducir ‘suavizadores’ más elaborados que permitan ajustar curvas similares a los ‘splines’.

- Combinar varios conjuntos o muestras datos, que sirvan para ‘entrenar’ el algoritmo, tal como se hace con las redes neuronales, o bien buscar generar algoritmos híbridos. Similarmente, buscar una combinación con métodos bayesianos. Ver por ejemplo las referencias [31, 32].
- Introducir técnicas de aceleración en algoritmos de optimización, tales como descenso estocástico y ADMM (alternating direction method of multipliers).
- Aplicar a problemas modelados por ecuaciones diferenciales parciales.
- Extender la aplicación de disparo múltiple a problemas de control óptimo (ver referencia [2]).

Bibliografía

- [1] O. Aydogmus, Ali Hakan TOR, *A Modified Multiple Shooting Algorithm for Parameter Estimation in ODEs Using Adjoint Sensitivity Analysis*, Applied Mathematics and Computation Volume 390 (2021) 125644.
- [2] Juárez, L. H., López, J., & Rojas, J. T. (2024). *Parameter estimation and control by penalized multiple shooting*. Annals Of Mathematical Sciences And Applications, 9(1), 237-279. <https://doi.org/10.4310/amsa.2024.v9.n1.a7>
- [3] Juárez, L. H., & Rojas, J. T. (2022). *Parameter estimation in ODEs. Modelling and computational issues*. Boletín de la Sociedad Mexicana de Computación Científica y Sus Aplicaciones, 8, 34-49. <https://smcca.org.mx/documentos/publicaciones/Boletin-SMCCA-No-8.pdf>
- [4] J. Calver and W. Enright, *Numerical methods for computing sensitivities for ODEs and DDEs*, Numerical Algorithms 74(4) (2017) 1101–1117.
- [5] Y. Cao, S. Li, L. Petzold, R. Serban, *Adjoint sensitivity analysis for differential algebraic equations: the adjoint DAE system and its numerical solution*. SIAM J. Sci. Comput. 3(24) (2003), 1076–1089.
- [6] F. Carbonell, Y. Iturria-Medina, J.C. Jimenez, *Multiple Shooting-Local Linearization method for the identification of dynamical systems*, Communications in Nonlinear Science and Numerical Simulation, 37 C (2016) 292-304.
- [7] Bock, H. G. (1981). *Numerical Treatment of Inverse Problems in Chemical Reaction Kinetics*. En Springer series in chemical physics (pp. 102-125). https://doi.org/10.1007/978-3-642-68220-9_8
- [8] Milstein, J. (1978). *Fitting multiple trajectories simultaneously to a model of inducible enzyme synthesis*. Mathematical Biosciences, 40(3-4), 175-184. [https://doi.org/10.1016/0025-5564\(78\)90082-2](https://doi.org/10.1016/0025-5564(78)90082-2)
- [9] Heinmets, F. (1964). *Analog computer analysis of a model-system for the induced enzyme synthesis*. Journal Of Theoretical Biology, 6(1), 60-75. [https://doi.org/10.1016/0022-5193\(64\)90066-9](https://doi.org/10.1016/0022-5193(64)90066-9)

-
- [10] Roth, R. S., & Roth, M. M. (1969). *Data unscrambling and the analysis of inducible enzyme synthesis*. Mathematical Biosciences, 5(1-2), 57-92. [https://doi.org/10.1016/0025-5564\(69\)90035-2](https://doi.org/10.1016/0025-5564(69)90035-2)
- [11] B. Sengupta, K.J. Friston, W.D. Penny, *Efficient gradient computation for dynamical models*, NeuroImage 98 (2014) 521–527.
- [12] Ramsay, J., Hooker, H., Campbell, D., Cao, J. *Parameter estimation for differential equations: a generalized smoothing approach*. J. R. Stat. Soc. Ser. B 69 (5), (2007) 741–796.
- [13] Calderhead, B., Girolami, M. *Estimating Bayes factors via thermodynamic integration and population MCMC*. Comput. Stat. Data Anal. 53 (12), (2009), 4028–4045.
- [14] B. van Domselaar and P. Hemker, *Nonlinear parameter estimation in initial value problems*, Technical Report NW 18/75, Mathematical Centre Amsterdam (1975).
- [15] H. B. Keller, *Numerical Methods for Two-Point Boundary Value Problems*, London: Blaisdell (1968).
- [16] M. R. Osborne, *On shooting methods for boundary value problems*, J. Math. Anal Appl, 27 (1969), 417-433.
- [17] J. Stoer, R. Bulirsch, *Introduction to Numerical Analysis*, Springer-Verlag, New York (1980) (The original German version was published by Springer-Verlag Berlin Heidelberg, 1972, 1976).
- [18] H.G. Bock, E. Kostina, J.P. Schlöder, *Direct multiple shooting and generalized Gauss-Newton method for parameter estimation problems in ODE models*, in: *Multiple Shooting and Time Domain Decomposition Methods*, Springer, 2015, pp. 1–34.
- [19] P. Drag, K. Styczén, *Multiple shooting SQP algorithm for optimal control of DAE systems with inconsistent initial conditions*, in: *Recent Advances in Computational Optimization* (2015), Springer, pp. 53–65.
- [20] A. Potschka, *Direct multiple shooting for parabolic PDE constrained optimization*, *Contributions in Mathematical and Computational Sciences*, Springer International Publishing (2015), 159–181.
- [21] L. Fang, S. Vandewalle, J. Meyers, *An SQP-based multiple shooting algorithm for large-scale PDE-constrained optimal control problems*, Journal of Computational Physics 477 (2023) 111927.
- [22] Jorge Nocedal, Stephen J. Wright, *Numerical Optimization*, New York: Springer, 1999.

-
- [23] M. Peifer, J. Timmer, *Parameter estimation in ordinary differential equations using the method of multiple shooting — a review*, Freiburg Centre for Data Analysis and Modelling, 1, 79104, Freiburg, Germany (2005).
- [24] T. Carraro, M. Geiger, *Direct and indirect multiple shooting for parabolic optimal control problems*, *Contributions in Mathematical and Computational Sciences*, Springer International Publishing (2015) 35–67.
- [25] L. Fang, S. Vandewalle, J. Meyers, *A parallel-in-time multiple shooting algorithm for large-scale PDE-constrained optimal control problems*, *Journal of Computational Physics*, Volume 452, (2022) 110926.
- [26] H. Diedam, S. Sager, *Global optimal control with the direct multiple shooting method*, *Optim. Control Appl. Methods* 39(2) (2017) 449–470.
- [27] M. Diehl, H. Bock, H. Diedam, P.B. Wieber, *Fast direct multiple shooting algorithms for optimal robot control*, *Lecture Notes in Control and Information Sciences*, Springer Berlin Heidelberg (2006) 65–93.
- [28] C. Kirches, L. Wirsching, H. Bock, J. Schlöder, *Efficient direct multiple shooting for nonlinear model predictive control on long horizons*, *J. Process Control* 22 (3) (2012) 540–550.
- [29] S. Subchan, *A direct multiple shooting for the optimal trajectory of missile guidance*, 2008 IEEE International Conference on Control Applications (2008).
- [30] Lloyd N. Trefethen, David Bau III, *Numerical Linear Algebra*, SIAM, 1997.
- [31] Ricky T. Q. Chen, Yulia Rubanova, Jesse Bettencourt, and David K Duvenaud. Neural ordinary differential equations. In S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett (eds.), *Advances in Neural Information Processing Systems*, volume 31. Curran Associates, Inc., 2018. URL <https://proceedings.neurips.cc/paper/2018/file/69386f6bb1dfed68692a24c8686939b9-Paper.pdf>.
- [32] Pashupati Hegde et. al. Variational multiple shooting for bayesian ODEs with gaussian processes. In *The 38th Conference on Uncertainty in Artificial Intelligence*, 2022. URL <https://openreview.net/forum?id=r2NuhIUoecq>.