



UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO
DIVISIÓN ACADÉMICA DE CIENCIAS BÁSICAS



**ESTIMACIÓN DE PARÁMETROS
DE UN MODELO DE VIH
POR EL MÉTODO BFGS Y LA ADJUNTA**

TESIS

PARA OBTENER EL TÍTULO DE:

LICENCIADA EN MATEMÁTICAS

PRESENTA:

SARAI MARTÍNEZ MÉNDEZ

BAJO LA DIRECCIÓN DE:

DR. JORGE LÓPEZ LÓPEZ

CUNDUACÁN, TABASCO, A: 16 de junio de 2025.

Declaración de Autoría y Originalidad

En la ciudad de Cunduacán, Tabasco, el día 20 del mes de Junio del año 2025, el que suscribe Sarai Martínez Méndez alumna del Programa de Licenciatura en Matemáticas con número de matrícula 202A31010, adscrito a la División Académica de Ciencias Básicas, de la Universidad Juárez Autónoma de Tabasco, como autora de la Tesis presentada para la obtención del título de Licenciada en Matemáticas y titulada ESTIMACIÓN DE PARÁMETROS DE UN MODELO DE VIH POR EL MÉTODO BFGS Y LA ADJUNTA dirigida por Dr. Jorge López López.

DECLARO QUE:

La Tesis es una obra original que no infringe los derechos de propiedad intelectual ni los derechos de propiedad industrial u otros, de acuerdo con el ordenamiento jurídico vigente, en particular, la LEY FEDERAL DEL DERECHO DE AUTOR (Decreto por el que se reforman y adicionan diversas disposiciones de la Ley Federal del Derecho de Autor del 01 de Julio de 2020 regularizando y aclarando y armonizando las disposiciones legales vigentes sobre la materia), en particular, las disposiciones referidas al derecho de cita.

Del mismo modo, asumo frente a la Universidad cualquier responsabilidad que pudiera derivarse de la autoría o falta de originalidad o contenido de la Tesis presentada de conformidad con el ordenamiento jurídico Vigente.

Villahermosa, Tabasco, a 20 de Junio de 2025.



Sarai Martínez Méndez

LICENCIATURA - ESTIMACIÓN DE PARÁMETROS DE UN MODELO DE VIH POR EL MÉTODO BFGS Y LA ADJUNTA

INFORME DE ORIGINALIDAD

4%

ÍNDICE DE SIMILITUD

FUENTES PRIMARIAS

1	gacetajuchiman.ujat.mx Internet	377 palabras — 2%
2	revistas.ujat.mx Internet	316 palabras — 1%
3	qdoc.tips Internet	127 palabras — 1%
4	idoc.pub Internet	57 palabras — < 1%
5	www.studocu.com Internet	44 palabras — < 1%
6	api-uilspace.unilorin.edu.ng Internet	37 palabras — < 1%
7	ir.jkuat.ac.ke Internet	26 palabras — < 1%
8	riunet.upv.es Internet	26 palabras — < 1%
9	doczz.net Internet	24 palabras — < 1%
	www.coursehero.com	



DIVISIÓN ACADÉMICA DE
CIENCIAS BÁSICAS
ESTUDIOS
TERMINALES

DIRECCIÓN

Cunduacán, Tabasco; a 16 de junio de 2025.

**C. SARAI MARTÍNEZ MÉNDEZ
PASANTE DE LA LIC. EN MATEMÁTICAS
PRESENTE**

Por medio del presente, me dirijo a usted para hacer de su conocimiento que proceda a la impresión del trabajo titulado: **“ESTIMACIÓN DE PARÁMETROS DE UN MODELO DE VIH POR EL MÉTODO BFGS Y LA ADJUNTA”** dirigido por el Dr. Jorge López López, bajo la modalidad de titulación por **TESIS**. La comisión de revisión conformada por el Dr. Miguel Ángel De la Rosa Castillo, Dr. Gamaliel Blé González, Dr. Iván Loreto Hernández y Dr. Justino Alavez Ramírez, liberó el documento en virtud de que reúne los requisitos para el **EXAMEN PROFESIONAL** correspondiente.

Sin otro particular, reciba usted un cordial saludo.

ATENTAMENTE



**DRA. HERMICENDA PÉREZ VIDAL
DIRECTORA**



C.c.p. Archivo.

DIR'DRA.HPV/kfvg 

Km.1 Carretera Cunduacán-Jalpa de Méndez, A.P. 24, C.P. 86690, Cunduacán, Tab., México.
Tel/Fax: (993) 3581500 Ext. 6702,6701 E-Mail: direccion.dacb@ujat.mx

www.ujat.mx

Carta de sesión de derechos

Villahermosa, Tabasco a 20 de junio de 2025

Por medio de la presente, manifestamos haber colaborado como AUTOR(A) y/o AUTORES(RAS) en la producción, creación y/o realización de la obra denominada ESTIMACIÓN DE PARÁMETROS DE UN MODELO DE VIH POR EL MÉTODO BFGS Y LA ADJUNTA.

Con fundamento en el artículo 83 de la Ley Federal del Derecho de autor y toda vez que, la creación y/o realización de la obra antes mencionada se realizó bajo la comisión de la Universidad Juárez Autónoma de Tabasco; entendemos y aceptamos el alcance del artículo en mención, de que tenemos el derecho al reconocimiento como autores de la obra, y la Universidad Juárez Autónoma de Tabasco mantendrá en un 100% la titularidad de los derechos patrimoniales por un periodo de 20 años sobre la obra en la que colaboramos, por lo anterior, cedemos el derecho patrimonial exclusivo en favor de la Universidad.

COLABORADORES



Sarai Martínez Méndez



Dr. Jorge López López

TESTIGOS



Lic. Iojany Abigail Valle Queb



Lic. Wendy Hernández Flores

Agradecimientos

Agradezco primeramente al Creador por darme la vida y sobretodo la paciencia para poder culminar mis estudios.

A mis padres quienes han sido mi apoyo en este trayecto, gracias por esforzarse cada día para darme lo mejor, por la confianza depositada en mí, por cada uno de sus consejos (aunque a veces no los tome) y por presionarme para acabar la tesis. Agradezco también a mi hermana, quien ha sido mi compañera de vida y de quien admiro su fortaleza y decisión.

A Iojany y Wendy, gracias por ser mis compañeras, amigas y confidentes desde el inicio de este camino, porque aunque hubieron risas, también hubieron momentos de estrés y frustración en los que nos hicimos compañía. Agradezco también a la vida por haberme hecho coincidir con personas que se convirtieron en grandes amigos y compañeros de viaje: Cindy, Saúl y Andry.

Gracias a mis profesores por las enseñanzas y por la dedicación puesta en cada clase. Hago una mención especial al Doctor Jorge, quien fue mi director de tesis y a quien agradezco por sus consejos (académicos y de vida), por su paciencia y por la confianza (y por no regañarme cuando llegaba tarde).

Por último (pero no menos importante) me agradezco a mí misma, porque a pesar de los días de estrés y las noches de insomnio, me esforcé para llegar a la meta.

Índice general

Declaración de Autoría y Originalidad	II
Autorización de impresión	IV
Carta de sesión de derechos	V
Agradecimientos	VI
Resumen	1
Abstract	1
Introducción	2
Marco Teórico	4
Justificación	6
Pregunta de investigación	7
Hipótesis	7
Glosario	7
Objetivo general	7
Objetivos específicos	8
Metodología	8

1. Un modelo del SIDA y su dinámica	9
1.1. Formulación del modelo para el VIH/SIDA	9
1.2. Equilibrios y estabilidad	11
1.2.1. Equilibrio trivial y no trivial	11
1.2.2. Estabilidad local de los estados de equilibrio	14
1.3. Solución numérica del modelo	17
1.4. Curvas de fase	20
2. El problema de estimación de parámetros como un problema de minimización	23
2.1. Definición de la función \mathcal{L} objetivo	23
2.2. Cálculo del gradiente de \mathcal{L}	24
2.2.1. Enfoque variacional	24
2.3. Método variacional para calcular sensibilidades	28
2.4. El método de la adjunta para calcular SO	29
2.5. Solución de la ecuación adjunta y cálculo del gradiente	32
2.6. Ejemplificación de las ideas expuestas al caso del modelo del SIDA	33
3. Solución numérica del problema de minimización de \mathcal{L}	37
3.1. El algoritmo BFGS	37
3.2. Problemas de búsqueda en línea	41
3.3. Detalles de programación del algoritmo BFGS	42
3.4. Experimentos numéricos considerando el modelo del SIDA	44
3.4.1. Estimación de los parámetros μ_e y k individualmente	46
3.4.2. Estimación de los parámetros k y μ_e simultáneamente	60
3.4.3. Estimación de los parámetros β, k, μ_V	67
4. Conclusiones	75
Bibliografía	78
Anexos	79
Programa genera datos.	79
Programa principal Main	81
Programa BFGS	83
Programa para calcular el gradiente	85
Programa de búsqueda en línea	87

Índice de tablas

1.	Casos reportados por año en el estado de Tabasco e incidencia (Secretaría de Salud, 2024).	4
2.	Casos reportados por año en el país, incidencia y mortalidad (Secretaría de Salud, 2024).	4
3.	Unidades en que se miden las variables de estado y los parámetros del sistema (1).	6
1.1.	Unidades en que se miden las variables de estado y los parámetros del sistema (1).	10
1.2.	Valores de los parámetros del modelo y del parámetro umbral R_0	19
3.1.	Resumen de las simulaciones al estimar el parámetro k tomando una variable observable.	46
3.2.	Resumen de las simulaciones al estimar el parámetro k tomando dos variables observables.	48
3.3.	Resumen de las simulaciones al estimar el parámetro k tomando tres variables observables.	51
3.4.	Resumen de las simulaciones al estimar el parámetro μ_e tomando una variable observable.	53
3.5.	Resumen de las simulaciones al estimar el parámetro μ_e tomando dos variables observables.	56
3.6.	Resumen de las simulaciones al estimar el parámetro μ_e tomando tres variables observables.	58
3.7.	Resumen de las simulaciones al estimar los parámetros k y μ_e tomando una variable observable.	60
3.8.	Resumen de las simulaciones al estimar los parámetros k y μ_e tomando dos variables observables.	63

3.9. Resumen de las simulaciones al estimar los parámetros k y μ_e tomando tres variables observables.	65
3.10. Resumen de las simulaciones al estimar los parámetros β , k y μ_V tomando una variable observable.	67
3.11. Resumen de las simulaciones al estimar los parámetros β , k y μ_V tomando dos variables observables.	70
3.12. Resumen de las simulaciones al estimar los parámetros β , k y μ_V tomando tres variables observables.	73

Universidad Juárez Autónoma de Tabasco.
México.

Índice de figuras

1.1. Equilibrios del individuo sano.	12
1.2. Equilibrios para los casos $R_0 < 1$ y $R_0 > 1$	14
1.3. Equilibrios para los casos: $R_0 < 1$, $R_0 = 1$ y $R_0 > 1$. En color rojo se muestran los equilibrios inestables y en azul los estables.	17
1.4. Caso 1. Dinámica del VIH/SIDA en el caso en el que teóricamente existe una cura, para el cual, $R_0 = 0.6$. La dinámica del sistema diferencial está determinada por el punto de equilibrio trivial $E_+ = (1421.7, 0, 0)$, que es localmente asintóticamente estable.	19
1.5. Caso 2. Dinámica del VIH/SIDA en el caso en el que teóricamente existe una cura, para el cual, $R_0 = 1$. La dinámica del sistema diferencial está determinada por el punto de equilibrio trivial $E_+ = (1421.7, 0, 0)$, que es localmente asintóticamente estable.	19
1.6. Caso 3. Dinámica del VIH/SIDA en el caso en el que el paciente tiene una enfermedad crónica, para el cual, $R_0 = 1.4$. Se tiene que $T_{s+} = 1421.7$ y el estado de equilibrio no trivial es $E^* = (T_s^*, T_e^*, V^*)^T = (1015.3, 71.7, 478.2)^T$, que es localmente asintóticamente estable.	20
1.7. Retrato fase para el caso 1, $R_0 = 0.6$ y dinámica local alrededor del equilibrio $E_+ = (1421.7, 0, 0)$, el cual es localmente asintóticamente estable, mientras que los equilibrios $E_- = (-87.9245, 0, 0)$ y $E^* = (2369.1, -372.5, -0.1)$ son inestables (e inadmisibles desde el punto de vista biológico). Las condiciones iniciales las representamos con el símbolo “o”.	21
1.8. Retrato fase para el caso 2, $R_0 = 1$ y dinámica local alrededor del equilibrio $E_+ = (1421.7, 0, 0)$, el cual es localmente asintóticamente estable, el equilibrio E^* , colapsa en E_+ y el equilibrio $E_- = (-87.9245, 0, 0)$ existe matemáticamente pero no es admisible biológicamente y no se muestra en esta figura. El símbolo “o”, representa las condiciones iniciales.	21

1.9. Retrato fase para el caso 3, $R_0 = 1.4$ y dinámica alrededor del equilibrio $E^* = (1015.1, 71.8, 478.4)$ el cual es localmente asintóticamente estable. En este caso, se tiene que los equilibrios $E_- = (-87.9245, 0, 0)$, $E_+ = (1421.7, 0, 0)$ son inestables. Las condiciones iniciales las representamos con el símbolo “o”. Las curvas que inician cerca de E_- pasan cerca de E_+ y siguen su camino hacia E^*	22
2.1. Solución exacta del modelo estándar del SIDA.	34
3.1. Diagrama de estructura del programa de minimización con el algoritmo BFGS.	43
3.2. Resultados numéricos al estimar el parámetro k y las condiciones iniciales, considerando la variable T_s como observable y tomando cien mil datos.	47
3.3. Solución exacta y estimada cuando se ajusta el parámetro k y las condiciones iniciales, considerando la variable T_s como observable y tomando cien mil datos, el algoritmo alcanzó convergencia en 2 iteraciones.	47
3.4. Resultados numéricos al estimar el parámetro k y las condiciones iniciales, considerando la variable V como observable y tomando diez mil datos.	47
3.5. Solución exacta y estimada cuando se ajusta el parámetro k y las condiciones iniciales, considerando la variable V como observable y tomando diez mil datos, el algoritmo alcanzó convergencia en 4 iteraciones.	48
3.6. Resultados numéricos al estimar el parámetro k y las condiciones iniciales, considerando las variables T_s y V como observables y tomando cien mil datos.	49
3.7. Solución exacta y estimada cuando se ajusta el parámetro k y las condiciones iniciales, considerando las variables T_s y V como observables y tomando cien mil datos, el algoritmo alcanzó convergencia en 5 iteraciones.	49
3.8. Resultados numéricos al estimar el parámetro k y las condiciones iniciales, considerando las variables T_s y V como observables y tomando cinco datos.	49
3.9. Solución exacta y estimada cuando se ajusta el parámetro k y las condiciones iniciales, considerando las variables T_s y V como observables y tomando cinco datos, el algoritmo alcanzó convergencia en 3 iteraciones.	50
3.10. Resultados numéricos al estimar el parámetro k y las condiciones iniciales, considerando tres variables como observables y tomando cien mil datos.	52
3.11. Solución exacta y estimada cuando se ajusta el parámetro k y las condiciones iniciales, considerando tres variables observables y tomando cien mil datos, el algoritmo alcanzó convergencia en 4 iteraciones.	52

3.12. Resultados numéricos al estimar el parámetro k y las condiciones iniciales, considerando tres variables observables y tomando cinco datos.	52
3.13. Solución exacta y estimada cuando se ajusta el parámetro k y las condiciones iniciales, considerando las tres variables observables y tomando cinco datos, el algoritmo alcanzó convergencia en 4 iteraciones.	53
3.14. Resultados numéricos al estimar el parámetro μ_e y las condiciones iniciales, considerando solo la variable T_s como observable y tomando cien mil datos, .	54
3.15. Solución exacta y estimada cuando se ajusta el parámetro μ_e y las condiciones iniciales, considerando solo la variable T_s como observable y tomando cien mil datos, el algoritmo alcanzó convergencia en 5 iteraciones.	54
3.16. Resultados numéricos al estimar el parámetro μ_e y las condiciones iniciales, considerando solo la variable V como observable y tomando cien mil datos.	54
3.17. Solución exacta y estimada cuando se ajusta el parámetro μ_e y las condiciones iniciales, considerando solo la variable V como observable y tomando cien mil datos, el algoritmo alcanzó convergencia en 4 iteraciones.	55
3.18. Resultados numéricos al estimar el parámetro μ_e y las condiciones iniciales, considerando solo la variable V como observable y tomando cien datos. .	55
3.19. Solución exacta y estimada cuando se ajusta el parámetro μ_e y las condiciones iniciales, considerando solo la variable V como observable y tomando cien datos, el algoritmo alcanzó convergencia en 7 iteraciones.	55
3.20. Resultados numéricos al estimar el parámetro μ_e y las condiciones iniciales, considerando las variables T_s y V como observables y tomando cien mil datos.	56
3.21. Solución exacta y estimada cuando se ajusta el parámetro μ_e y las condiciones iniciales, considerando las variables T_s y V como observables y tomando cien mil datos, el algoritmo alcanzó convergencia en 5 iteraciones.	57
3.22. Resultados numéricos al estimar el parámetro μ_e y las condiciones iniciales, considerando las variables T_s y V como observables y tomando veinte datos.	57
3.23. Solución exacta y estimada cuando se ajusta el parámetro μ_e y las condiciones iniciales, considerando las variables T_s y V como observables y tomando veinte datos, el algoritmo alcanzó convergencia en 9 iteraciones.	57
3.24. Resultados numéricos al estimar el parámetro μ_e y las condiciones iniciales, considerando las variables T_s , T_e y V como observables y tomando cien mil datos.	58
3.25. Solución exacta y estimada cuando se ajusta el parámetro μ_e y las condiciones iniciales, considerando las variables T_s , T_e y V como observables y tomando cien mil datos, el algoritmo alcanzó convergencia en 6 iteraciones.	59

3.26. Resultados numéricos al estimar el parámetro μ_e y las condiciones iniciales, considerando las variables T_s , T_e y V como observables y tomando diez datos.	59
3.27. Solución exacta y estimada cuando se ajusta el parámetro μ_e y las condiciones iniciales, considerando las variables T_s , T_e y V como observables y tomando diez datos, el algoritmo alcanzó convergencia en 6 iteraciones.	59
3.28. Resultados numéricos al estimar los parámetros k y μ_e y las condiciones iniciales, considerando la variable T_s como observable y tomando diez mil datos.	61
3.29. Solución exacta y estimada cuando se ajustan los parámetros k y μ_e y las condiciones iniciales, considerando la variable T_s como observable y tomando diez mil datos, el algoritmo alcanzó convergencia en 4 iteraciones.	61
3.30. Resultados numéricos al estimar los parámetros k y μ_e y las condiciones iniciales, considerando la variable V como observable y tomando cien mil datos.	61
3.31. Solución exacta y estimada cuando se ajustan los parámetros k y μ_e y las condiciones iniciales, considerando la variable V como observable y tomando cien mil datos, el algoritmo alcanzó convergencia en 15 iteraciones.	62
3.32. Resultados numéricos al estimar los parámetros k y μ_e y las condiciones iniciales, considerando la variable V como observable y tomando 280 datos.	62
3.33. Solución exacta y estimada cuando se ajustan los parámetros k y μ_e y las condiciones iniciales, considerando la variable V como observable y tomando 280 datos, el algoritmo alcanzó convergencia en 9 iteraciones.	62
3.34. Resultados numéricos al estimar los parámetros k y μ_e y las condiciones iniciales, considerando las variables T_s y V como observables y tomando cien mil datos.	64
3.35. Solución exacta y estimada cuando se ajustan los parámetros k y μ_e y las condiciones iniciales, considerando las variables T_s y V como observables y tomando cien mil datos, el algoritmo alcanzó convergencia en 8 iteraciones.	64
3.36. Resultados numéricos al estimar los parámetros k y μ_e y las condiciones iniciales, considerando las variables T_s y V como observables y tomando veinte datos.	64
3.37. Solución exacta y estimada cuando se ajustan los parámetros k y μ_e y las condiciones iniciales, considerando las variables T_s y V como observables y tomando veinte datos, el algoritmo alcanzó convergencia en 4 iteraciones.	65
3.38. Resultados numéricos al estimar los parámetros k y μ_e y las condiciones iniciales, considerando las variables T_s , T_e y V como observables y tomando cien mil datos.	66

3.39. Solución exacta y estimada cuando se ajustan los parámetros k y μ_e y las condiciones iniciales, considerando las variables T_s , T_e y V como observables y tomando cien mil datos, el algoritmo alcanzó convergencia en 6 iteraciones.	66
3.40. Resultados numéricos al estimar los parámetros k y μ_e y las condiciones iniciales, considerando las variables T_s , T_e y V como observables y tomando veinte datos.	66
3.41. Solución exacta y estimada cuando se ajustan los parámetros k y μ_e y las condiciones iniciales, considerando las variables T_s , T_e y V como observables y tomando veinte datos, el algoritmo alcanzó convergencia en 4 iteraciones.	67
3.42. Resultados numéricos al estimar los parámetros β , k y μ_V y las condiciones iniciales, considerando la variable T_s como observable y tomando cien mil datos.	68
3.43. Solución exacta y estimada cuando se ajustan los parámetros β , k y μ_V y las condiciones iniciales, considerando la variable T_s como observable y tomando cien mil datos, el algoritmo alcanzó convergencia en 11 iteraciones.	68
3.44. Resultados numéricos al estimar los parámetros β , k y μ_V y las condiciones iniciales, considerando la variable V como observable y tomando cien mil datos.	68
3.45. Solución exacta y estimada cuando se ajustan los parámetros β , k y μ_V y las condiciones iniciales, considerando la variable V como observable y tomando cien mil datos, el algoritmo alcanzó convergencia en 20 iteraciones.	69
3.46. Resultados numéricos al estimar los parámetros β , k y μ_V y las condiciones iniciales, considerando la variable V como observable y tomando trescientos datos.	69
3.47. Solución exacta y estimada cuando se ajustan los parámetros β , k y μ_V y las condiciones iniciales, considerando la variable V como observable y tomando trescientos datos, el algoritmo alcanzó convergencia en 12 iteraciones.	69
3.48. Resultados numéricos al estimar los parámetros β , k y μ_V y las condiciones iniciales, considerando las variables T_s y V como observables y tomando cien mil datos.	70
3.49. Solución exacta y estimada cuando se ajustan los parámetros β , k y μ_V y las condiciones iniciales, considerando las variables T_s y V como observables y tomando cien mil datos, el algoritmo alcanzó convergencia en 14 iteraciones.	71

3.50. Resultados numéricos al estimar los parámetros β , k y μ_V y las condiciones iniciales, considerando la variables T_s y V como observables y tomando quinientos datos.	71
3.51. Solución exacta y estimada cuando se ajustan los parámetros β , k y μ_V y las condiciones iniciales, considerando las variables T_s y V como observables y tomando quinientos datos, el algoritmo alcanzó convergencia en 9 iteraciones. 71	71
3.52. Resultados numéricos al estimar los parámetros β , k y μ_V y las condiciones iniciales, considerando la variables T_s y V como observables y tomando cien datos.	72
3.53. Solución exacta y estimada cuando se ajustan los parámetros β , k y μ_V y las condiciones iniciales, considerando las variables T_s y V como observables y tomando cien datos, el algoritmo alcanzó convergencia en 9 iteraciones. 72	72
3.54. Resultados numéricos al estimar los parámetros β , k y μ_V y las condiciones iniciales, considerando las variables T_s , T_e y V como observables y tomando cien mil datos.	73
3.55. Solución exacta y estimada cuando se ajustan los parámetros β , k y μ_V y las condiciones iniciales, considerando las variables T_s , T_e y V como observables y tomando cien mil datos, el algoritmo alcanzó convergencia en 11 iteraciones.	74
3.56. Resultados numéricos al estimar los parámetros β , k y μ_V y las condiciones iniciales, considerando las variables T_s , T_e y V como observables y tomando cien datos.	74
3.57. Solución exacta y estimada cuando se ajustan los parámetros β , k y μ_V y las condiciones iniciales, considerando las variables T_s , T_e y V como observables y tomando cien datos, el algoritmo alcanzó convergencia en 5 iteraciones.	74

Resumen

En este trabajo se realizó estimación de parámetros de un modelo de VIH/SIDA mediante el algoritmo de minimización Broyden- Fletcher- Goldfarb- Shanno (BFGS), que no solo toma información del gradiente sino también de la Hessiana. Para calcular el gradiente se utilizó un método variacional que implica resolver el sistema adjunto del sistema de estado. El sistema de estado y el sistema adjunto se resolvieron en cada iteración usando el método Runge-Kutta de cuarto orden. Para perturbaciones en los datos del 10 %, se obtuvieron estimaciones con un error máximo en promedio del 4 %.

Palabras clave: Estimación de parámetros, sistema adjunto, algoritmo BFGS, Runge-Kutta, Gradiente, Hessiana.

Abstract

In this work, parameter estimation of an HIV/AIDS model was performed using the Broyden-Fletcher-Goldfarb-Shanno (BFGS) minimization algorithm, which utilizes not only gradient information but also the Hessian. To compute the gradient, a variational method was used, which involves solving the adjoint system of the state system. Both the state system and the adjoint system were solved at each iteration using the fourth-order Runge-Kutta method. For data perturbations of 10 %, the estimated parameters had a maximum average error of 4 %

Keywords: Parameter estimation, adjoint system, BFGS algorithm, Runge-Kutta, Gradient, Hessian.

Introducción

De acuerdo con (Secretaría de Salud, 2020), el Virus de Inmunodeficiencia Humana (VIH) es descrito como un lentivirus de la familia Retroviridae, este virus tiene un periodo de incubación prolongado, de hecho la enfermedad se presenta después de varios años. Existen dos clases de VIH, el VIH-1 y el VIH-2; el primero ha resultado ser el más contagioso y es responsable de la mayoría de las infecciones por VIH en todo el mundo, convirtiéndolo en un gran problema de salud global; la segunda clase de virus, el VIH-2, es menos contagioso y se encuentra casi exclusivamente en África occidental.

Cuando no se cuenta con el tratamiento antirretroviral, el virus evoluciona de la siguiente manera: se tiene la fase de infección aguda, en ella el virus comienza a replicarse con incremento de la carga viral y un descenso de linfocitos TCD4+, los cuales son parte importante de nuestro sistema inmunológico; posteriormente se llega a la fase crónica, aunque la persona puede no presentar síntomas por años, el virus se multiplica de manera importante; por último se tiene el Síndrome de Inmunodeficiencia Adquirida (SIDA), que es la etapa crítica del VIH .

Para el año 2010 alrededor de 34 millones de personas vivían con SIDA, de las cuales el 68 % se encontraba en Asia Subsahariana, seguida de Asia Oriental y Sudoriental con alrededor de 4 millones de infectados; en tercer lugar, se encontraba América Latina, Europa Oriental y Asia Central con alrededor de 1.5 millones de personas. En este mismo año, la situación de la epidemia mejoró, sin embargo, la cantidad de muertes ocasionadas por el virus seguía siendo preocupante (Secretaría de Salud, 2012).

Como se detalla en el documento (ONUSIDA, 2024), el número de nuevas infecciones a nivel mundial ha ido descendiendo desde 2010 alrededor de un 39%, considerando que en ese año se reportaron 2.1 millones [1.7 millones-2.7 millones] de nuevas infecciones en comparación con las 1.3 millones [1 millón-1.7 millones] reportadas en 2023. Aunque esto refleja un importante avance, el objetivo para 2025 era estar por debajo de 370 000, lo cual está muy lejos de ocurrir.

En América Latina se estima que en el periodo 2010-2022 el número de nuevas infecciones ha aumentado un 8%, teniendo alrededor de 110 mil infecciones en 2022. Además, en este mismo año la prevalencia del VIH en adultos de 15 a 49 años fue del 0.5% (0.3-0.6%). Sin embargo, el número de personas fallecidas ha disminuido pasando de 42 mil muertes en 2010 a 27 mil en 2022, lo que se atribuye al uso del tratamiento Antirretroviral (CEN-SIDA, 2022).

En México, fue en el año de 1983 cuando se reportaron los primeros casos de VIH, para los siguientes años el número de infectados fue en aumento, ya que no se contaba con un tratamiento eficaz y no todos los pacientes eran atendidos, de hecho generalmente la

atención era brindada en hospitales de tercer nivel. Para finales de los años ochenta, en nuestro país se comenzó a utilizar el AZT (Zidovudina, Azidotimidina), esto como parte de los ensayos clínicos con el fármaco, pero fue hasta 1991 que el tratamiento estuvo totalmente disponible; a pesar de esto, la situación no mejoró ya que el AZT no era suficientemente eficaz (CIENI, 2011).

Para 1996, se presentó un tratamiento con resultados extraordinarios, se trató de un Tratamiento Antirretrovírico (TAR), el cual combinaba tres fármacos. En ese momento, México tuvo la oportunidad de ser el primer país en vías de desarrollo que permitiera el acceso a dicho tratamiento de manera inmediata, sin embargo, las autoridades impidieron que esto ocurriera y consecuentemente los resultados fueron desastrosos, debido a la cantidad de personas que murieron por no contar con un tratamiento (CIENI, 2011).

El activismo jugó un papel importante para que el TAR pudiera llegar a nuestro país, a tal grado que entre 1997 y 1998 se aprobó la compra de fármacos antirretrovirales. En un inicio, éstos fueron ofrecidos a las personas que tenían seguridad social (IMSS o ISSSTE). Cabe mencionar que no hubo una correcta organización, coordinación y regulación del uso del TAR, pues se usaban inadecuadamente las dosis y los intervalos de aplicación, sin regulación en el uso y distribución de fármacos. Así, a pesar de que hubo cierta mejoría, no fue la esperada en comparación de países como Brasil o Estados Unidos (CIENI, 2011).

En 1997, surge el FONSIDA (Fondo Nacional para Atención a Personas con VIH/SIDA), con el cual se buscaba facilitar el acceso a los antirretrovirales de las personas con VIH que no contaban con seguridad social. La falta de interés por parte de las autoridades llevó a que este proyecto fracasara y la consecuencia de esto fue la cancelación del fondo, dejando a más de treinta mil personas sin tratamiento (CIENI, 2011).

Desde el inicio de la pandemia de VIH/SIDA y hasta el 15 de abril del 2024, en México se han reportado en total 375 296 casos de VIH, de los cuales 10 892 se encuentran en el estado de Tabasco. En la tabla 2 se muestra el número de casos nuevos de VIH en el país, desde 1983 hasta el primer trimestre de 2024, junto con otros indicadores. Por otro lado, en la tabla 1 se muestran los datos para el estado de Tabasco en los mismos periodos. En estas tablas se consideran las siguientes definiciones:

$$\text{Incidencia} = \frac{\text{Número de casos nuevos por año}}{\text{No. total de personas en riesgo al comienzo del periodo}} \times 100000.$$

$$\text{Mortalidad} = \frac{\text{Número de defunciones por año}}{\text{Población total}} \times 100000.$$

Tabasco								
Indicador/Periodo	1983-2017	2018	2019	2020	2021	2022	2023	2024
Casos	7277	618	594	382	656	670	575	120
Incidencia		25.2	24.0	15.3	26.0	25.5	21.7	4.5

Tabla 1: Casos reportados por año en el estado de Tabasco e incidencia (Secretaría de Salud, 2024).

México					
Indicador/Periodo	1983-2017	2018	2019	2020	2021
Casos	272972	17608	17618	10575	16241
Incidencia		14.1	14.0	8.3	12.7
Mortalidad		4.04	4.2	3.63	3.62

Indicador/Periodo	2022	2023	2024
Casos	18,012	17,739	4531
Incidencia	14.0	13.5	3.4
Mortalidad	3.71		

Tabla 2: Casos reportados por año en el país, incidencia y mortalidad (Secretaría de Salud, 2024).

Muchos aspectos que ayudan a comprender y analizar enfermedades, incluido el VIH/SIDA, han sido obtenidos del uso de modelos matemáticos y las herramientas de modelación correspondientes, incluidos los aspectos teóricos de existencia y unicidad de las soluciones de los modelos, hasta la solución numérica para un estudio eficiente y económico de escenarios, pasando por la estimación de parámetros de los modelos (Barberis et al., 2021; Olarte & Muñoz, 2021; Peregrino et al., 2023; Sungchakit et al., 2022).

Marco Teórico

A finales de la década de 1980 surgieron los primeros modelos de infección de VIH, mismos que tomaron como base modelos epidemiológicos existentes, como el modelo SIS o el modelo SIR diseñados para describir la propagación de infecciones entre individuos en una población. En los casos como el VIH, se requería no solo modelar cómo se transmitía entre individuos, sino también modelar la dinámica viral dentro del individuo y para ello surgieron los modelos de *dinámica viral*, con los cuales se buscaba describir cómo se propagaba un virus entre las células sanas dentro del cuerpo de un solo individuo. Este tipo de modelos también han sido utilizados para estudiar otras infecciones virales humanas, como la hepatitis B (Nowak et al., 1996), hepatitis C (Alavez-Ramírez et al., 2011), influenza (de la Pava Salgado, 2010; Murillo et al., 2013), el dengue (Nuraini et al.,

2009; Olarte & Muñoz, 2021), entre otras (Hill, 2018).

En el caso del SIDA, se cuenta con modelos sofisticados que constan de 6 ecuaciones o más y consideran entre sus variables las terapias antiretrovirales (Hill, 2018) mientras que los modelos más simples solo estudian la relación entre carga viral V , la cantidad de células linfáticas sanas T_s y la cantidad de células linfáticas infectadas T_e , dando lugar a modelos de solo tres ecuaciones con tres incógnitas. Un ejemplo de estos últimos es el siguiente modelo diferencial

$$\begin{cases} \dot{T}_s = q + \beta T_s \left(1 - \frac{T_s}{T_{Max}}\right) - \mu_s T_s - k T_s V, \\ \dot{T}_e = k T_s V - \mu_e T_e, \\ \dot{V} = \rho T_e - \mu_V V. \end{cases} \quad (1)$$

Las variables y parámetros involucrados en el sistema (1) son:

- $T_s(t)$: densidad de población de células linfáticas sanas en el tiempo t . Esta variable se considera medible u observable.
- $T_e(t)$: densidad de población de células linfáticas $TCD4^+$ infectadas productivamente al tiempo t .
- $V(t)$: densidad de carga viral del VIH-1 en el tiempo t . Esta también se considera una variable observable o medible.
- q : tasa de creación de células T_s desde fuentes del cuerpo.
- β : tasa de proliferación máxima de células T_s . La proliferación se considera de tipo logístico.
- T_{max} : valor de saturación de la densidad de la población T_s , este es un dato medible, cuyo rango es 200 – 1500.
- μ_s : tasa per cápita de mortalidad de las células T_s .
- k : tasa de infección de células T_s por el VIH-1.
- μ_e : tasa per cápita de mortalidad de las células T_e .
- ρ : tasa de producción de virus por células infectadas productivamente.
- μ_V : tasa per cápita de mortalidad del virus.

Cada parámetro del sistema (1) es estrictamente positivo, a menos que el virus esté ausente. En la tabla 3 se especifican las unidades en que se deben medir las variables de estado y los parámetros. Nuestro interés en este trabajo de tesis es dentro del campo de la estimación de parámetros, la cual consiste en minimizar un funcional que depende de mediciones de algunas de las variables del modelo, las diferentes metodologías difieren básicamente en cuál es el funcional y cómo lo minimizan. Nuestro interés es mostrar la aplicación de una metodología que para minimizar el funcional utiliza un método quasi Newton, el algoritmo de minimización Broyden- Fletcher- Goldfarb- Shanno (BFGS), que como usa información no sólo del gradiente sino también de la matriz Hessiana o de la concavidad de la función, requiere de menos iteraciones para dar resultados satisfactorios, aunque esto implica que se tenga una manera eficiente de calcular el diferencial del funcional (Alavez-Ramírez, 2007; Juárez et al., 2024; Juárez Valencia & Rojas, 2022). Siendo este nuestro interés principal, trabajaremos con el modelo (1).

Parámetros	Unidades
T_s	células/ mm^3
T_e	células/ mm^3
V	copias/ ml (10^{-3} copias/ μl)
q	(células/ mm^3) \times día $^{-1}$
β	células/ mm^3
T_{max}	células/ mm^3
μ_s	día $^{-1}$
k	copias/ μl \times día $^{-1} \equiv \mu l / (\text{copias} \times \text{día})$
μ_e	día $^{-1}$
p	(copias/ μl) \times (células/ mm^3) \times día $^{-1} \equiv \text{copias} / (\text{células} \times \text{día})$
μ_V	día $^{-1}$

Tabla 3: Unidades en que se miden las variables de estado y los parámetros del sistema (1).

Justificación

Aunque actualmente se puede notar una evolución en el tratamiento del VIH, las estadísticas muestran que este virus continúa siendo causa de preocupación a nivel mundial. Para estudiar, con modelos del tipo (1), la relación entre la carga viral y las células inmunitarias TCD4+, es necesario conocer los valores de los parámetros correspondientes a cada situación específica, lo cual normalmente no se tiene disponible en la práctica y deben estimarse estos valores. En este contexto, en el presente trabajo de tesis nos enfocamos en el ajuste de parámetros a partir de observaciones sintéticas, que corresponderían a

mediciones de campo en la práctica.

Pregunta de investigación

¿Es posible estimar parámetros del modelo del SIDA conociendo mediciones de algunas de sus variables y minimizando un funcional con el algoritmo BFGS?

Hipótesis

Los parámetros del modelo (1) se pueden estimar minimizando un funcional con el método BFGS y calculando el gradiente con el método de la adjunta, tomando como datos las mediciones de algunas de sus variables.

Glosario

Término	Definición
Prevalencia	Proporción de la población que padece la enfermedad en estudio en un momento dado, y se denomina únicamente como prevalencia.
Incidencia	Volumen de casos nuevos que aparecen en un periodo determinado en proporción a la población en riesgo; también puede manifestar la velocidad con que aparecen tales casos.
Mortalidad	Número de defunciones que ocurren por cada 1000, 10000 o 100000 habitantes en un año determinado. Es un indicador que depende de la estructura por edad de la población, por lo que para hacerlo comparable con el mismo indicador de otras poblaciones, se debe de estandarizar (tipificar).

Objetivo general

Diseñar un programa para implementar el método BFGS y minimizar el funcional de ajuste de parámetros, calculando su gradiente por el método de la adjunta.

Objetivos específicos

1. Entender conceptualmente el modelo del SIDA (1).
2. Estudiar los resultados sobre la estabilidad local y global de los equilibrios del modelo (1).
3. Hacer un programa para resolver numéricamente el modelo (1) y visualizar algunas dinámicas.
4. Estudiar el método de la adjunta para calcular gradientes de funcionales y calcular el gradiente del funcional \mathcal{L} de estimación de parámetros.
5. Hacer un programa para resolver numéricamente el modelo adjunto de (1).
6. Hacer un programa con el algoritmo BFGS para minimizar el funcional \mathcal{L} de estimación de parámetros.

Metodología

1. Revisar (Alavez-Ramírez, 2007) para estudiar conceptualmente el modelo del SIDA (1).
2. Revisar (Alavez-Ramírez, 2007) para estudiar los resultados sobre la estabilidad local y global de los equilibrios del modelo (1).
3. Hacer un programa en MatLab para resolver numéricamente el modelo (1) y su sistema adjunto.
4. Revisar (Juárez Valencia & Rojas, 2022) para estudiar el problema de minimización asociado con la estimación de parámetros del modelo (1).
5. Aplicar la metodología de estimación de parámetros para recuperar los parámetros del modelo a partir de datos sintéticos.

Un modelo del SIDA y su dinámica

En este capítulo se presenta un modelo estándar del VIH/SIDA propuesto por Perelson y Nelson (Perelson & Nelson, 1999) que estudia la relación entre el virus y el sistema inmune. Dicho modelo consta de tres variables de estado: la cantidad de células sanas (T_s), la cantidad de células infectadas (T_e) y la carga viral o la cantidad de virus (V). Además, se muestran en este capítulo algunos resultados teóricos sobre la dinámica del sistema de EDO's asociado a dicho modelo.

De igual forma, se incluyen simulaciones numéricas para visualizar el comportamiento, a través del tiempo, de las variables de estado, para algunos escenarios (condiciones iniciales) que ilustran los resultados teóricos mencionados.

1.1. Formulación del modelo para el VIH/SIDA

En este trabajo se estudia el siguiente modelo matemático (Perelson & Nelson, 1999) para la dinámica del VIH/SIDA:

$$\begin{aligned} \dot{T}_s &= q + \beta T_s \left(1 - \frac{T_s}{T_{max}}\right) - \mu_s T_s - k T_s V, \\ \dot{T}_e &= k T_s V - \mu_e T_e, \\ \dot{V} &= \rho T_e - \mu_V V. \end{aligned} \tag{1.1}$$

Teóricamente, la existencia y unicidad de las curvas solución de este sistema de EDO's, dependen de un valor inicial para cada una de la variables de estado. Las variables y parámetros de este sistema diferencial, se definen a continuación.

- $T_s(t)$: densidad de población de células linfáticas sanas en el tiempo t . Esta variable se considera medible u observable.

- $T_e(t)$: densidad población de células linfáticas $TCD4^+$ infectadas productivamente al tiempo t .
- $V(t)$: densidad de carga viral del VIH-1 en el tiempo t . Esta también se considera una variable observable o medible.
- q : tasa de creación de células T_s desde fuentes del cuerpo.
- β : tasa de proliferación máxima de células T_s . La proliferación se considera de tipo logístico.
- T_{max} : valor de saturación de la densidad de la población T_s , este es un dato medible, cuyo rango es 200 – 1500.
- μ_s : tasa per cápita de mortalidad de las células T_s .
- k : tasa de infección de células T_s por el VIH-1.
- μ_e : tasa per cápita de mortalidad de las células T_e .
- ρ : tasa de producción de virus por células infectadas productivamente.
- μ_V : tasa per cápita de mortalidad del virus.

En este contexto, la proliferación se refiere al proceso por el cual las células inmunitarias se dividen y aumentan en número dentro del organismo.

Cada parámetro del sistema (1.1) es estrictamente positivo, a menos que el virus esté ausente. En la tabla 1.1 se especifican las unidades de cada variable de estado (T_s , T_e , V) y cada uno de los parámetros.

Parámetros	Unidades
T_s	células/ mm^3
T_e	células/ mm^3
V	copias/ ml (10^{-3} copias/ μl)
q	(células/ mm^3) \times día $^{-1}$
β	células/ mm^3
T_{max}	células/ mm^3
μ_s	día $^{-1}$
k	copias/ $\mu l \times$ día $^{-1} \equiv \mu l / (\text{copias} \times \text{día})$
μ_e	día $^{-1}$
ρ	(copias/ μl) \times (células/ mm^3) \times día $^{-1} \equiv$ copias/ (células \times día)
μ_V	día $^{-1}$

Tabla 1.1: Unidades en que se miden las variables de estado y los parámetros del sistema (1).

1.2. Equilibrios y estabilidad

1.2.1. Equilibrio trivial y no trivial

Para obtener los estados de equilibrio del sistema (1.1), se debe hacer cero cada derivada del sistema, es decir los equilibrios del sistema son los valores de las variables de estado que satisfacen el sistema algebraico

$$\begin{aligned} q + \beta T_s \left(1 - \frac{T_s}{T_{max}}\right) - \mu_s T_s - k T_s V &= 0, \\ k T_s V - \mu_e T_e &= 0, \\ \rho T_e - \mu_V V &= 0. \end{aligned} \quad (1.2)$$

Para encontrar los equilibrios se consideran dos casos: cuando el virus está ausente y cuando este se encuentra presente. En el primer caso, se tiene que $V = 0$, de lo cual se obtiene que $T_e = 0$, por lo que, para encontrar los equilibrios en este caso, basta con encontrar las raíces de la ecuación cuadrática

$$q + T_s(\beta - \mu_s) - \frac{\beta}{T_{max}}(T_s)^2 = 0. \quad (1.3)$$

Las raíces de la ecuación cuadrática (1.3) son

$$\begin{aligned} T_{s\pm} &= \frac{-(\beta - \mu_s) \pm \sqrt{(\beta - \mu_s)^2 + 4\frac{\beta q}{T_{max}}}}{-2\frac{\beta}{T_{max}}} \\ &= -\frac{T_{max}}{2\beta} \left[-(\beta - \mu_s) \pm \sqrt{(\beta - \mu_s)^2 + 4\frac{\beta q}{T_{max}}} \right]. \end{aligned}$$

Nótese que

$$-\frac{T_{max}}{2\beta} \left[-(\beta - \mu_s) + \sqrt{(\beta - \mu_s)^2 + 4\frac{\beta q}{T_{max}}} \right] = \frac{T_{max}}{2\beta} \left[(\beta - \mu_s) - \sqrt{(\beta - \mu_s)^2 + 4\frac{\beta q}{T_{max}}} \right] < 0$$

y que

$$-\frac{T_{max}}{2\beta} \left[-(\beta - \mu_s) - \sqrt{(\beta - \mu_s)^2 + 4\frac{\beta q}{T_{max}}} \right] = \frac{T_{max}}{2\beta} \left[(\beta - \mu_s) + \sqrt{(\beta - \mu_s)^2 + 4\frac{\beta q}{T_{max}}} \right] > 0.$$

Por lo que, la raíz positiva está dada por

$$T_{s+} = \frac{T_{max}}{2\beta} \left[(\beta - \mu_s) + \sqrt{(\beta - \mu_s)^2 + 4\frac{\beta q}{T_{max}}} \right] \quad (1.4)$$

y la negativa por

$$T_{s-} = \frac{T_{max}}{2\beta} \left[(\beta - \mu_s) - \sqrt{(\beta - \mu_s)^2 + 4\frac{\beta q}{T_{max}}} \right] \quad (1.5)$$

Por lo que se tienen dos equilibrios para el caso $V = 0$ pero, como se aclarará más adelante, solo uno tiene sentido biológico (ver figura 1.1):

$$E_+ = (T_{s+}, 0, 0)^T \text{ y } E_- = (T_{s-}, 0, 0)^T.$$

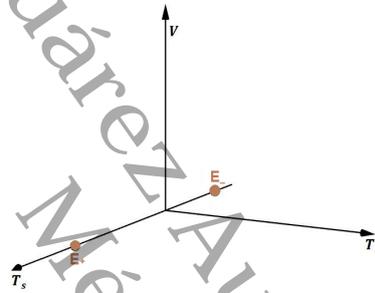


Figura 1.1: Equilibrios del individuo sano.

Para el segundo caso, en el que se tiene un enfermo crónico, se consideran los estados de equilibrio no triviales, los cuales se representan como $E^* = (T_s^*, T_e^*, V^*)$. Para estos casos $V^* > 0$, así que de (1.2), se obtiene:

$$T_e^* = \frac{\mu_V}{\rho} \cdot V^*, \quad (1.6)$$

$$T_s^* = \frac{\mu_e \mu_V}{k\rho}, \quad (1.7)$$

$$V^* = \frac{1}{k} \left[(\beta - \mu_s) + \frac{q}{T_s^*} - \beta \frac{T_s^*}{T_{max}} \right]. \quad (1.8)$$

De estas expresiones se ve que hay a lo mas un equilibrio no trivial E^* . Para expresar los valores de las variables del equilibrio no trivial E^* , se hará uso del parámetro umbral R_0 que se define como

$$R_0 \equiv \frac{k\rho}{\mu_e \mu_V} T_{s+}. \quad (1.9)$$

Este parámetro es fundamental para determinar cómo evoluciona la enfermedad, ya sea

hacia la cura o hacia la enfermedad crónica.

De (1.7), resulta que

$$R_0 = \frac{T_{s+}}{T_s^*} > 0. \quad (1.10)$$

Por otro lado, se sabe que T_{s+} satisface (1.3), es decir,

$$q + T_{s+}(\beta - \mu_s) - \frac{\beta}{T_{max}}(T_{s+})^2 = 0,$$

de donde se sigue que

$$\beta - \mu_s = \frac{\beta T_{s+}}{T_{max}} - \frac{q}{T_{s+}}. \quad (1.11)$$

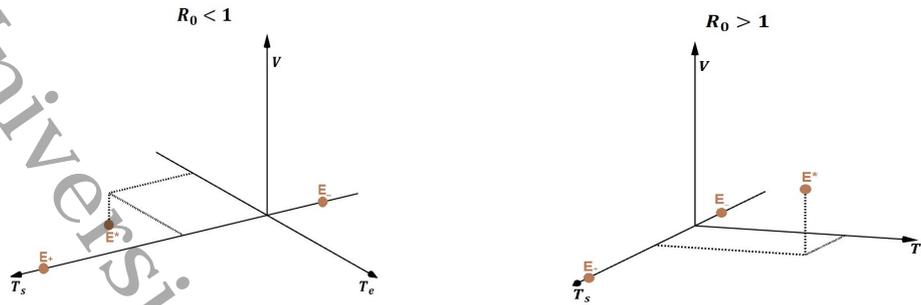
Sustituyendo (1.11) en (1.8) y haciendo uso de (1.10), se tiene que

$$\begin{aligned} V^* &= \frac{1}{k} \left[\frac{\beta T_{s+}}{T_{max}} - \frac{q}{T_{s+}} + \frac{q}{T_s^*} - \beta \frac{T_s^*}{T_{max}} \right] \\ &= \frac{1}{k} \left[\frac{\beta T_{s+}}{T_{max}} - \frac{q}{T_{s+}} + \frac{q}{T_s^*} - \beta \frac{T_s^*}{T_{max}} \right] \frac{T_{s+} T_s^*}{T_{s+} T_s^*} \\ &= \frac{1}{k} \left[\frac{\beta T_{s+}}{T_{max}} \frac{1}{R_0} - \frac{q}{T_{s+}} + \frac{q}{T_s^*} \frac{T_{s+} T_s^*}{T_{s+} T_s^*} - \beta \frac{T_s^*}{T_{max}} \frac{1}{R_0} \frac{T_{s+}}{T_s^*} \right] \\ &= \frac{1}{k} \left[\frac{\beta T_{s+}}{R_0 T_{max}} R_0 - \frac{q}{T_{s+}} + \frac{q}{T_{s+}} \frac{T_{s+} T_s^*}{(T_s^*)^2} - \beta \frac{T_{s+}}{R_0 T_{max}} \frac{T_s^*}{T_s^*} \right] \\ &= \frac{1}{k} \left[\frac{\beta T_{s+}}{R_0 T_{max}} R_0 - \frac{q}{T_{s+}} + \frac{q}{T_{s+}} R_0 - \beta \frac{T_{s+}}{R_0 T_{max}} \right] \\ &= \frac{1}{k} \left[\frac{\beta T_{s+}}{R_0 T_{max}} (R_0 - 1) + \frac{q}{T_{s+}} (R_0 - 1) \right], \end{aligned}$$

por lo que

$$V^* = \frac{1}{k} \left[\frac{q}{T_{s+}} + \frac{\beta T_{s+}}{R_0 T_{max}} \right] (R_0 - 1).$$

Nótese que si $R_0 = 1$, $V^* = 0$ y $T_e^* = 0$, entonces $E^* = (T_s^*, 0, 0)$. De (1.10), si $R_0 = 1$, entonces $T_s^* = T_{s+}$, por lo que en este caso $E^* = E_+$ y solo existen los dos equilibrios libres de enfermedad. Por otro lado, para cada uno de los casos $R_0 < 1$ y $R_0 > 1$ existe un equilibrio no trivial, los cuales se ilustran en la figura 1.2.

Figura 1.2: Equilibrios para los casos $R_0 < 1$ y $R_0 > 1$.

Lo anterior se puede resumir en el siguiente teorema.

Teorema 1.2.1. El modelo (1.1) tiene los siguientes equilibrios con respecto al parámetro R_0 .

- i) Si $R_0 < 1$, tiene tres equilibrios E_- , E_+ y E^* , pero E_- y E^* no tienen sentido biológico.
- ii) Si $R_0 = 1$, solo tiene dos equilibrios E_- y E_+ , pero E_- no tiene sentido biológico.
- iii) Si $R_0 > 1$, tiene tres equilibrios E_- , E_+ y E^* , pero solo E_+ y E^* tienen sentido biológico.

De esta manera, se tienen los dos teoremas siguientes.

Teorema 1.2.2 (Alavez-Ramírez, 2007). Si $R_0 \leq 1$, entonces $E_+ = (T_{s+}, 0, 0)^T$ es el único estado de equilibrio admisible para el sistema (1.1).

Teorema 1.2.3 (Alavez-Ramírez, 2007). El estado de equilibrio $E^* = (T_s^*, T_e^*, V^*)$ es admisible si y sólo si $R_0 > 1$. En este caso, también se tiene que $0 < T_s^* < T_{s+} (< T_{max})$.

La prueba de los teoremas 1.2.2 y 1.2.3 se puede consultar en (Alavez-Ramírez, 2007).

1.2.2. Estabilidad local de los estados de equilibrio

Para realizar el estudio de la estabilidad de los estados de equilibrio, es necesario obtener la matriz jacobiana del sistema (1.1) en todos los puntos del espacio fase, la cual está dada de la siguiente manera

$$J(T_s, T_e, V) = \begin{pmatrix} \beta \left(1 - 2\frac{T_s}{T_{max}}\right) - \mu_s - kV & 0 & -kT_s \\ kV & -\mu_e & kT_s \\ 0 & \rho & -\mu_V \end{pmatrix}. \quad (1.12)$$

Considérese el caso del estado de equilibrio trivial, para ver la estabilidad local se evalúa la matriz jacobiana del sistema (1.1) en $E_+ = (T_{s+}, 0, 0)^T$, de donde se obtiene

$$J(E_+) = \begin{pmatrix} \beta \left(1 - 2\frac{T_{s+}}{T_{max}}\right) - \mu_s & 0 & -kT_{s+} \\ 0 & -\mu_e & kT_{s+} \\ 0 & \rho & -\mu_V \end{pmatrix}. \quad (1.13)$$

De la expresión en (1.4), se tiene que

$$\begin{aligned} T_{s+} &= \frac{T_{max}}{2\beta} \left[(\beta - \mu_s) + \sqrt{(\beta - \mu_s)^2 + 4\frac{\beta q}{T_{max}}} \right] \\ \frac{2\beta}{T_{max}} T_{s+} &= (\beta - \mu_s) + \sqrt{(\beta - \mu_s)^2 + 4\frac{\beta q}{T_{max}}} \\ \frac{2\beta}{T_{max}} T_{s+} - \beta + \mu_s &= \sqrt{(\beta - \mu_s)^2 + 4\frac{\beta q}{T_{max}}} \\ \beta \left(2\frac{T_{s+}}{T_{max}} - 1\right) + \mu_s &= \sqrt{(\beta - \mu_s)^2 + 4\frac{\beta q}{T_{max}}}, \\ -\beta \left(2\frac{T_{s+}}{T_{max}} - 1\right) - \mu_s &= -\sqrt{(\beta - \mu_s)^2 + 4\frac{\beta q}{T_{max}}}, \\ \beta \left(1 - 2\frac{T_{s+}}{T_{max}}\right) - \mu_s &= -\sqrt{(\beta - \mu_s)^2 + 4\frac{\beta q}{T_{max}}}, \\ \beta \left(1 - 2\frac{T_{s+}}{T_{max}}\right) - \mu_s &= -\sqrt{\Delta}. \end{aligned}$$

Donde $\Delta \equiv (\beta - \mu_s)^2 + 4\frac{\beta q}{T_{max}} > 0$, de esta manera se reescribe (1.13) como

$$J(E_+) = \begin{pmatrix} -\sqrt{\Delta} & 0 & -kT_{s+} \\ 0 & -\mu_e & kT_{s+} \\ 0 & \rho & -\mu_V \end{pmatrix}. \quad (1.14)$$

Para obtener el polinomio característico de (1.13), notemos que de (1.9) se sigue que $T_{s+} = \frac{\mu_e \mu_V}{k\rho} R_0$, de esta manera

$$\begin{aligned} p(\lambda) &= \det(J(E_+) - \lambda I) \\ &= (-\sqrt{\Delta} - \lambda)[(-\mu_e - \lambda)(-\mu_V - \lambda) - \rho k T_{s+}] \\ &= (-\sqrt{\Delta} - \lambda)[\mu_e \mu_V + (\mu_e + \mu_V)\lambda + \lambda^2 - \rho k T_{s+}] \\ &= (-\sqrt{\Delta} - \lambda) \left[\mu_e \mu_V + (\mu_e + \mu_V)\lambda + \lambda^2 - \rho k \left(\frac{\mu_e \mu_V}{k\rho} R_0 \right) \right] \\ &= (-\sqrt{\Delta} - \lambda) [(\mu_e + \mu_V)\lambda + \lambda^2 - \mu_e \mu_V R_0 + \mu_e \mu_V]. \end{aligned}$$

Por lo que

$$p(\lambda) = (-\sqrt{\Delta} - \lambda)[\lambda^2 + (\mu_e + \mu_V)\lambda + (1 - R_0)\mu_e\mu_V].$$

Así uno de los autovalores de (1.13) es $-\sqrt{\Delta} < 0$, mientras que los otros se obtienen calculando las raíces del polinomio cuadrático

$$\pi(\lambda) = \lambda^2 + (\mu_e + \mu_V)\lambda + (1 - R_0)\mu_e\mu_V. \quad (1.15)$$

Para que el equilibrio sea localmente asintóticamente estable, todos sus valores propios deben tener parte real negativa. Las raíces de la ecuación anterior tendrán parte real negativa si y sólo si todos sus coeficientes son positivos. Nótese que para que se cumpla $(1 - R_0)\mu_e\mu_V > 0$ se debe tener $R_0 < 1$, y por lo tanto, se concluye que si $R_0 < 1$, el estado de equilibrio trivial $E_+ = (T_{s+}, 0, 0)^T$ es localmente asintóticamente estable.

Nota 1.2.4. Para el caso en el que $R_0 = 1$, hay al menos una raíz de la ecuación cuadrática (1.15) igual a cero, por lo cual no se puede dar una conclusión inmediata con respecto a la estabilidad, pero en (Alavez-Ramírez, 2007) se dan condiciones bajo las cuales cuando $R_0 = 1$, el equilibrio E_+ es localmente asintóticamente estable.

Para estudiar la estabilidad del equilibrio no trivial, primero se evalúa la matriz jacobiana (1.12) en $E^* = (T_s^*, T_e^*, V^*)^T$:

$$J(E^*) = \begin{pmatrix} \beta \left(1 - 2\frac{T_s^*}{T_{max}}\right) - \mu_s - kV^* & 0 & -kT_s^* \\ kV^* & -\mu_e & kT_s^* \\ 0 & \rho & -\mu_V \end{pmatrix}. \quad (1.16)$$

Con fines de simplificar la expresión para la entrada (1, 1) de $J(E^*)$ se considera la primera ecuación del sistema (1.2), tomando $T_s = T_s^*$ y $V = V^*$:

$$\begin{aligned} q + \beta T_s^* \left(1 - \frac{T_s^*}{T_{max}}\right) - \mu_s T_s^* - kT_s^* V^* &= 0; \\ q + T_s^* \left[\beta \left(1 - 2\frac{T_s^*}{T_{max}} + \frac{T_s^*}{T_{max}}\right) - \mu_s - kV^* \right] &= 0; \\ T_s^* \left[\beta \left(1 - 2\frac{T_s^*}{T_{max}}\right) + \beta \frac{T_s^*}{T_{max}} - \mu_s - kV^* \right] &= -q; \\ \beta \left(1 - 2\frac{T_s^*}{T_{max}}\right) + \beta \frac{T_s^*}{T_{max}} - \mu_s - kV^* &= -\frac{q}{T_s^*}; \\ \beta \left(1 - 2\frac{T_s^*}{T_{max}}\right) - \mu_s - kV^* &= -\frac{q}{T_s^*} - \beta \frac{T_s^*}{T_{max}}; \end{aligned}$$

$$\beta \left(1 - 2 \frac{T_s^*}{T_{max}} \right) - \mu_s - kV^* = -\frac{1}{T_s^*} \left(q + \beta \frac{(T_s^*)^2}{T_{max}} \right).$$

La última ecuación nos da una expresión alternativa para la entrada (1, 1) de $J(E^*)$, que si la llamamos b , entonces $b = -\frac{1}{T_s^*} \left(q + \beta \frac{(T_s^*)^2}{T_{max}} \right)$ y (1.16) se reescribe como

$$J(E^*) = \begin{pmatrix} b & 0 & -kT_s^* \\ kV^* & -\mu_e & kT_s^* \\ 0 & \rho & -\mu_V \end{pmatrix} \quad (1.17)$$

Para este caso, el polinomio característico es

$$\sigma(\lambda) = (\lambda + b)[\lambda^2 + (\mu_e + \mu_V)\lambda] + k^2 p T_s^* V^*.$$

En (Alavez-Ramírez, 2007) se puede ver que si $R_0 > 1$ entonces todos los autovalores de (1.17) tendrán parte real negativa y entonces el estado de equilibrio no trivial será localmente asintóticamente estable. Además, allí también se muestra que en este caso, el estado de equilibrio E_+ es inestable.

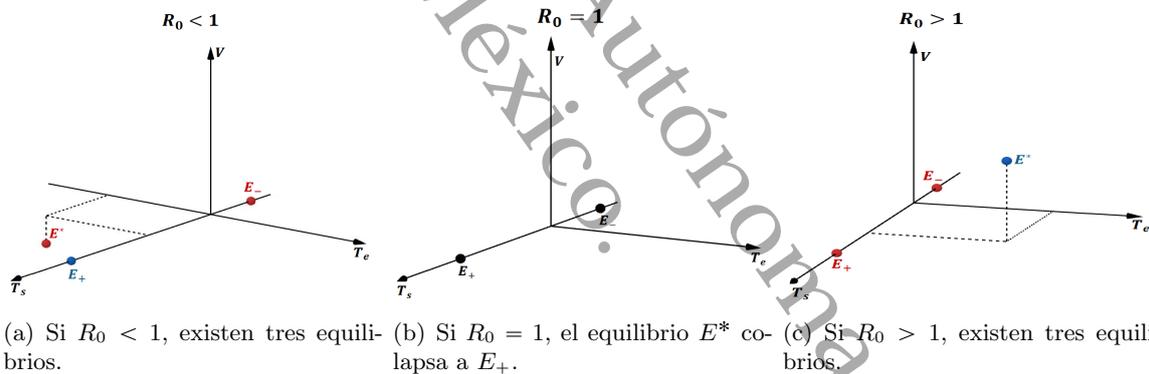


Figura 1.3: Equilibrios para los casos: $R_0 < 1$, $R_0 = 1$ y $R_0 > 1$. En color rojo se muestran los equilibrios inestables y en azul los estables.

1.3. Solución numérica del modelo

En esta sección se tratará la solución numérica del modelo del VIH/SIDA que nos ocupa. Será resuelto por un método de Runge-Kutta. Los métodos de Runge Kutta son métodos iterativos que son utilizados con frecuencia para aproximar las soluciones de sistemas de Ecuaciones Diferenciales Ordinarias con condiciones iniciales, siendo el más popular el Runge-Kutta de cuarto orden (RK4), el cual, si el sistema a resolver es

$$\begin{aligned}y'(t) &= f(t, y), \quad t \in (0, T) \\y(0) &= \alpha,\end{aligned}$$

entonces se hace una partición de $[0, T]$ definiendo $N + 1$ nodos como $t_i = ih$, para $i = 0, 1, \dots, N$ donde $h = T/N$. Con esto, el método RK4 proporciona las aproximaciones y_i para $y(t_i)$ de la siguiente manera:

$$\begin{aligned}y_{i+1} &= y_i + \frac{h}{6} (k_1 + 2k_2 + 2k_3 + k_4), \\y_0 &= \alpha,\end{aligned}$$

donde

$$\begin{aligned}k_1 &= f(t_i, y_i), \\k_2 &= f\left(t_i + \frac{h}{2}, y_i + \frac{h}{2}k_1\right), \\k_3 &= f\left(t_i + \frac{h}{2}, y_i + \frac{h}{2}k_2\right), \\k_4 &= f(t_{i+1}, y_i + k_3),\end{aligned}$$

para $i = 0, 1, \dots, N - 1$. El término $\frac{1}{6} (k_1 + 2k_2 + 2k_3 + k_4)$, puede ser considerado como una "pendiente promedio" de la solución en el intervalo $[t_i, t_{i+1}]$.

Este método tiene un error local de truncamiento $\mathcal{O}(h^4)$ siempre que la solución tenga cinco derivadas continuas (Burden & Faires, 2002). En general, los métodos del tipo Runge-Kutta son adecuados para problemas no rígidos (como en el caso del modelo del SIDA que nos ocupa) en los que se requiere una precisión moderada, pero no son adecuados para problemas rígidos (Burden & Faires, 2002).

Utilizando este método (RK4), se resolvió numéricamente el modelo (1.1) para ilustrar la dinámica del VIH. Se muestran dos escenarios que corresponden a: un enfermo que se cura (caso 1 y caso 2) y un enfermo crónico (caso 3). Los valores utilizados de los parámetros se especifican en la tabla 1.2; las condiciones iniciales para los tres casos fueron $T_{s0} = 70$, $T_{e0} = 30$ y $V_0 = 200$. De acuerdo a las fórmulas obtenidas para los equilibrios en la sección 2.2: en el caso 1 se tiene que $E_+ = (1421.7, 0, 0)$, $E_- = (-87.9245, 0, 0)$, $E^* = (T_s^*, T_e^*, V^*)^T = (2369.1, -372.46, -0.13036)^T$; en el caso 2, $E_+ = (1421.7, 0, 0)$,

$E_- = (-87.9245, 0, 0)$ y E^* colapsa en E_+ ; por último, para el caso 3, $E_+ = (1421.7, 0, 0)$, $E_- = (-87.9245, 0, 0)$ y $E^* = (T_s^*, T_e^*, V^*)^T = (1015.1, 71.8, 478.4)^T$.

El comportamiento de las variables para los casos 1, 2 y 3 se muestran en las figuras 1.4, 1.5 y 1.6, respectivamente.

	q	β	T_{max}	μ_s	μ_e	k	μ_V	ρ	R_0
Caso 1	10	0.12	1500	1.33×10^{-2}	0.5	6.03×10^{-1}	0.4	1.4×10^{-4}	0.6
Caso 2	10	0.12	1500	1.33×10^{-2}	0.5	5.2754×10^{-5}	0.4	2.6667	1.0
Caso 3	10	0.12	1500	1.33×10^{-2}	0.5	7.3885×10^{-5}	0.4	2.6667	1.4

Tabla 1.2: Valores de los parámetros del modelo y del parámetro umbral R_0 .

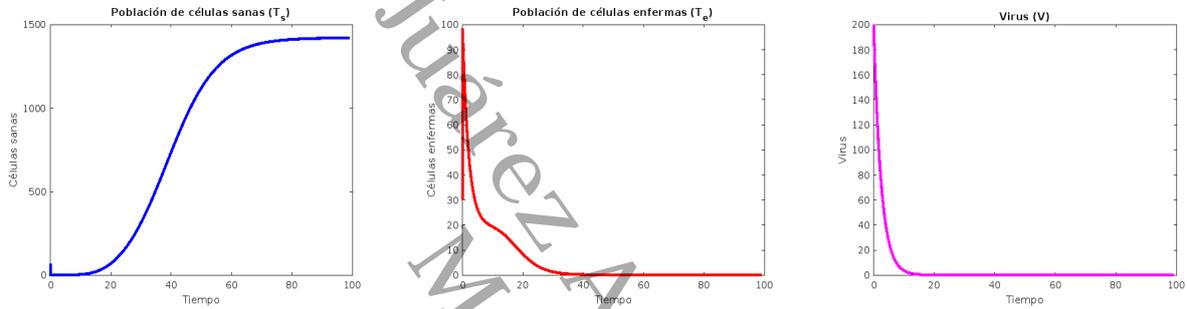


Figura 1.4: Caso 1. Dinámica del VIH/SIDA en el caso en el que teóricamente existe una cura, para el cual, $R_0 = 0.6$. La dinámica del sistema diferencial está determinada por el punto de equilibrio trivial $E_+ = (1421.7, 0, 0)$, que es localmente asintóticamente estable.

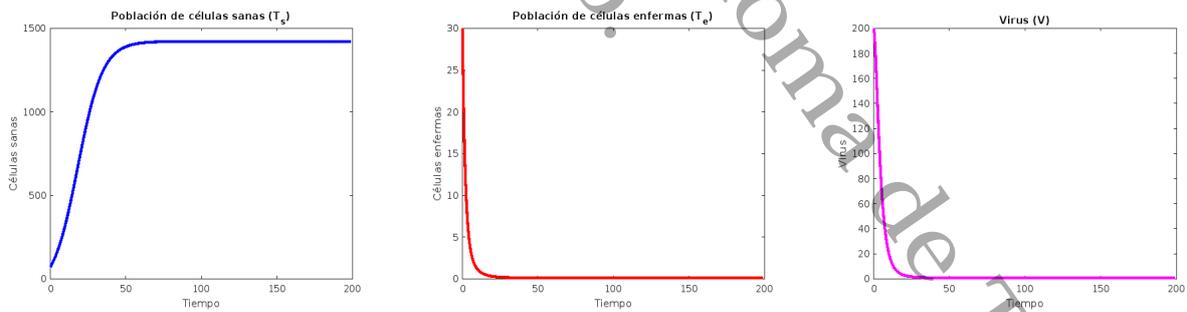


Figura 1.5: Caso 2. Dinámica del VIH/SIDA en el caso en el que teóricamente existe una cura, para el cual, $R_0 = 1$. La dinámica del sistema diferencial está determinada por el punto de equilibrio trivial $E_+ = (1421.7, 0, 0)$, que es localmente asintóticamente estable.

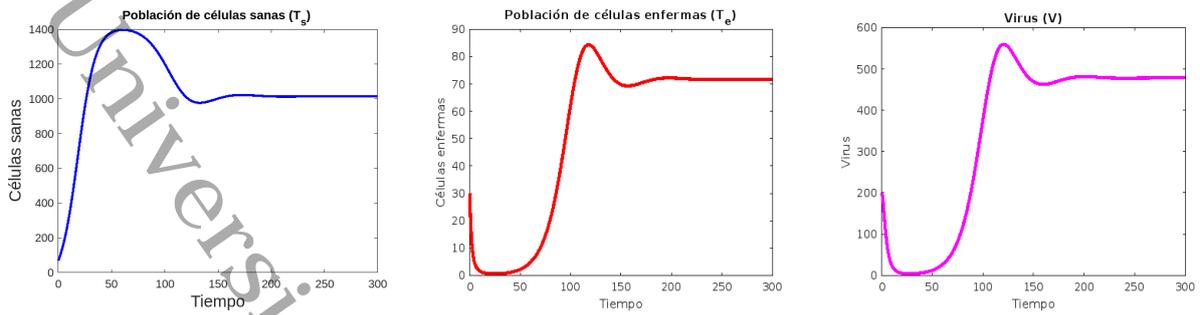


Figura 1.6: Caso 3. Dinámica del VIH/SIDA en el caso en el que el paciente tiene una enfermedad crónica, para el cual, $R_0 = 1.4$. Se tiene que $T_{s+} = 1421.7$ y el estado de equilibrio no trivial es $E^* = (T_s^*, T_e^*, V^*)^T = (1015.3, 71.7, 478.2)^T$, que es localmente asintóticamente estable.

1.4. Curvas de fase

En esta sección se presentan las curvas de fase alrededor de los puntos de equilibrio de acuerdo con los casos presentados en la sección anterior, con el fin de ilustrar y verificar los resultados teóricos obtenidos para la estabilidad de estos equilibrios.

En las figuras que se presentarán, los puntos de color azul, representan equilibrios estables, mientras que los puntos de color rojo representan equilibrios inestables.

En la figura 1.7 se presenta el diagrama de fase para el caso 1, en el que $R_0 = 0.6$, es decir, $R_0 < 1$. Se tienen tres equilibrios E_- , E_+ y E^* , para los cuales con los parámetros utilizados, se obtiene $E_- = (-87.9245, 0, 0)$, $E_+ = (1421.7, 0, 0)$ y $E^* = (2369.1, -372.5, -0.1)$, siendo E_- y E^* inestables. En el sentido biológico, solo se admite E_+ , el cual, de acuerdo con los resultados de estabilidad, es localmente asintóticamente estable, lo que se refleja en la figura 1.7.

En la figura 1.8 se muestra el diagrama de fase para el caso 2, en el que $R_0 = 1$. En dicha figura se puede observar el equilibrio E_+ , y la dinámica alrededor del mismo. En este caso, para los valores que se asignaron a los parámetros, se observa que este equilibrio es estable. En esta figura no se muestra el equilibrio E_- y el equilibrio E^* no existe. Con los parámetros utilizados, se tiene que $E_- = (-87.9245, 0, 0)$ y $E_+ = (1421.7, 0, 0)$.

Por último, la figura 1.9 corresponde al caso 3, en el que $R_0 = 1.4 > 1$. Para este caso, se admiten matemáticamente tres equilibrios, que con los valores asignados a los parámetros son $E_- = (-87.9245, 0, 0)$, $E_+ = (1421.7, 0, 0)$ y $E^* = (1015.1, 71.8, 478.4)$. En este caso, los equilibrios E_- y E_+ son inestables pero E^* es localmente asintóticamente estable, tal

como se logra ver en la figura.

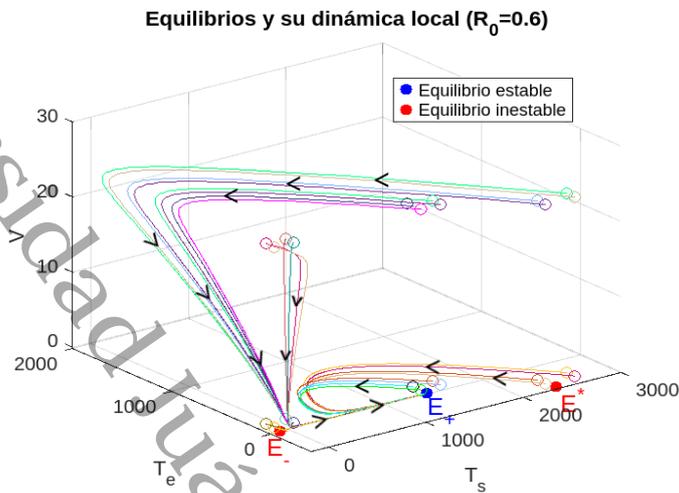


Figura 1.7: Retrato fase para el caso 1, $R_0 = 0.6$ y dinámica local alrededor del equilibrio $E_+ = (1421.7, 0, 0)$, el cual es localmente asintóticamente estable, mientras que los equilibrios $E_- = (-87.9245, 0, 0)$ y $E^* = (2369.1, -372.5, -0.1)$ son inestables (e inadmisibles desde el punto de vista biológico). Las condiciones iniciales las representamos con el símbolo “o”.

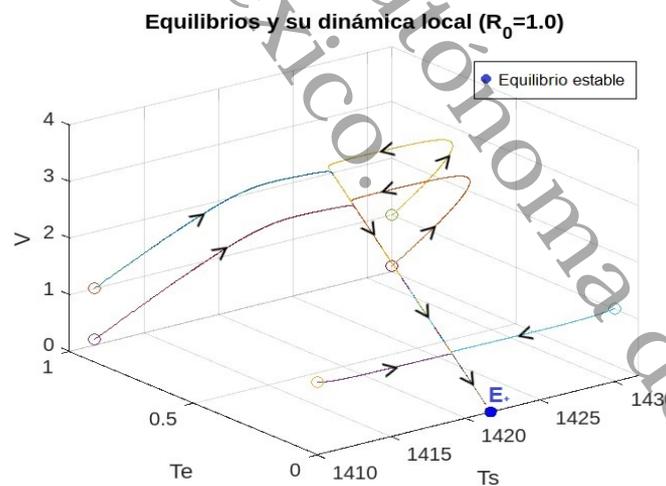
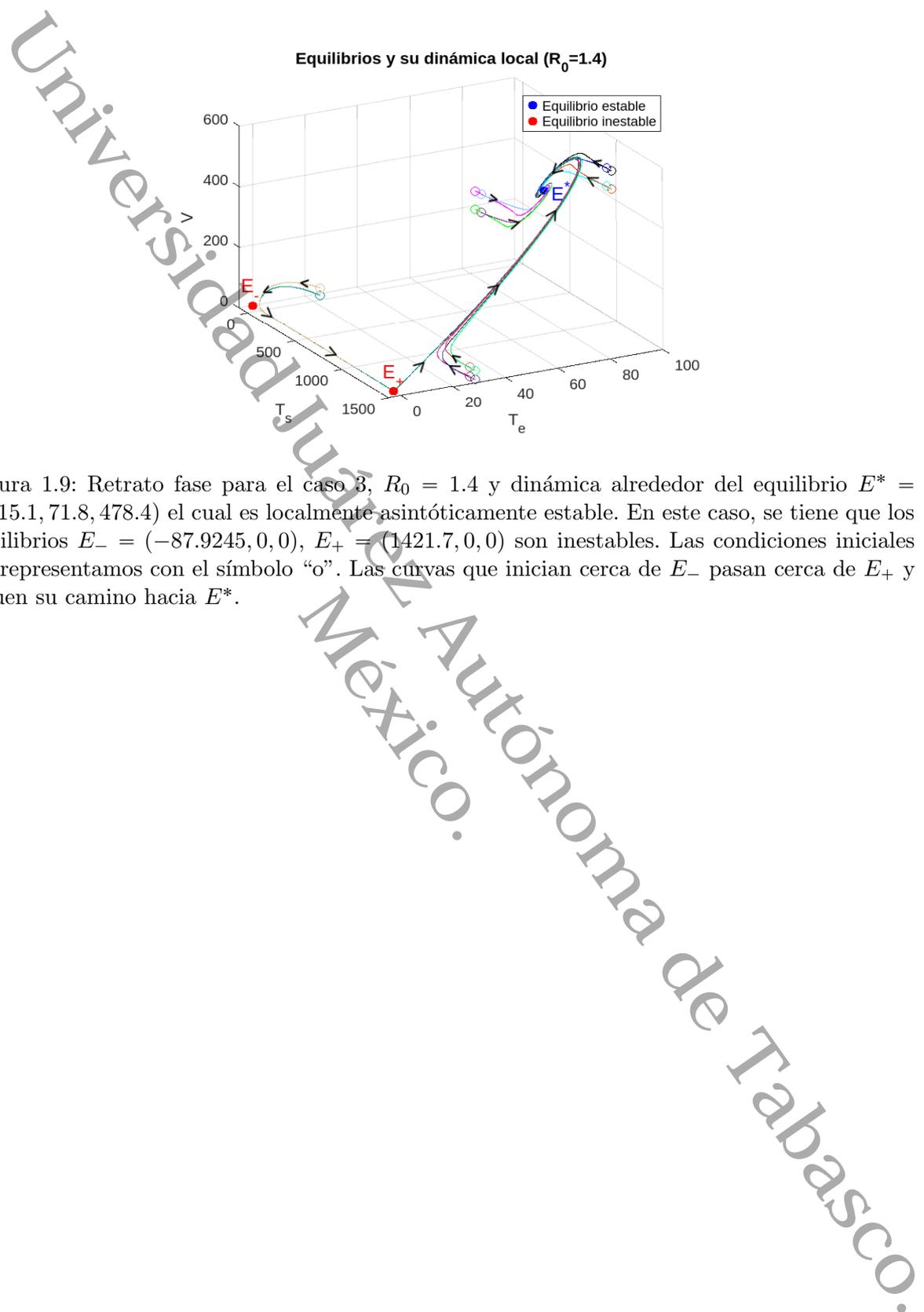


Figura 1.8: Retrato fase para el caso 2, $R_0 = 1$ y dinámica local alrededor del equilibrio $E_+ = (1421.7, 0, 0)$, el cual es localmente asintóticamente estable, el equilibrio E^* , colapsa en E_+ y el equilibrio $E_- = (-87.9245, 0, 0)$ existe matemáticamente pero no es admisible biológicamente y no se muestra en esta figura. El símbolo “o”, representa las condiciones iniciales.



El problema de estimación de parámetros como un problema de minimización

2.1. Definición de la función \mathcal{L} objetivo

Considérese $x(t) \in \mathbb{R}^d$ como una variable de estado en el tiempo $t \in I = [t_0, t_f]$, que cumple el problema de valores iniciales

$$\dot{\mathbf{x}}(t) = \mathbf{f}(t, \mathbf{x}(t), \boldsymbol{\theta}), \quad \mathbf{x}(t_0) = \mathbf{x}_0. \quad (2.1)$$

La función vectorial $\mathbf{f} : \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^{np} \mapsto \mathbb{R}^d$, se supone en $C^1(\mathbb{R}^{d+np})$ (derivable con respecto a las variables \mathbf{x} y $\boldsymbol{\theta}$). Aquí np representa la dimensión del vector de parámetros $\boldsymbol{\theta}$. Se denotará la solución de este problema como $\mathbf{x}(t; \mathbf{x}_0, \boldsymbol{\theta})$ cuando sea explícita su dependencia de \mathbf{x}_0 y $\boldsymbol{\theta}$. Además, se supone que se dispone de m mediciones vectoriales en los tiempos $t_0 \leq t_1 \leq \dots \leq t_m \leq t_f$ en I :

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

En este caso, $\boldsymbol{\epsilon}_i \in \mathbb{R}^d$ son vectores aleatorios independientes que representan errores de medición y siguen una distribución Gaussiana multivariada con media cero y varianzas vectoriales $\sigma_i^2 \in \mathbb{R}^d$. Comúnmente, se usa una perturbación Gaussiana porque muchos fenómenos siguen ese comportamiento en sus datos y en sus perturbaciones, y además por sus propiedades matemáticas de linealidad y estabilidad. En general, no todas las componentes de $\mathbf{x}(t)$ son observables, por tal motivo, se consideran variables observables y no observables, y por tanto,

$$\bar{\mathbf{x}}_{ij} = \begin{cases} 0, & \text{para toda } i = 1, \dots, m, \text{ si la variable } j \text{ es no observable} \\ \bar{\mathbf{x}}_{ij}, & \text{si la variable } j \text{ es observable} \end{cases} \quad (2.3)$$

donde $\bar{\mathbf{x}}_{ij}$ es el valor observado de la variable j en el tiempo de observación i .

El objetivo del problema de estimación de parámetros es encontrar aproximaciones satisfactorias del vector desconocido $\boldsymbol{\theta}$ y de las condiciones iniciales \mathbf{x}_0 , a partir de las mediciones, para lo cual se minimiza la siguiente función objetivo de mínimos cuadrados

$$\mathcal{L}(\mathbf{x}_0, \boldsymbol{\theta}) = \frac{1}{2} \left\| \frac{\mathbf{x}_0 - \mathbf{s}_0}{\boldsymbol{\sigma}_0} \right\|_d^2 + \frac{1}{2} \sum_{i=1}^m \left\| \frac{\bar{\mathbf{x}}(t_i; \mathbf{x}_0, \boldsymbol{\theta}) - \bar{\mathbf{x}}_i}{\bar{\boldsymbol{\sigma}}_i} \right\|_d^2, \quad (2.4)$$

donde $\mathbf{x}(t)$ debe satisfacer la ODE (2.1).

El vector fijo \mathbf{s}_0 denota la medición experimental de las condiciones iniciales. Se asume que todas las componentes de la variable de estado \mathbf{x} son observables en el tiempo inicial t_0 .

2.2. Cálculo del gradiente de \mathcal{L}

En esta sección el objetivo es calcular de manera eficiente el gradiente de \mathcal{L} con respecto a \mathbf{x}_0 y $\boldsymbol{\theta}$, con el fin de minimizar el funcional \mathcal{L} de (2.4) con un método tipo Newton o que use derivadas, es decir, minimizar \mathcal{L} para encontrar las condiciones iniciales $\mathbf{x}_0 \in \mathbb{R}^d$ y los parámetros $\boldsymbol{\theta} \in \mathbb{R}^{np}$ que resuelven el problema de ajuste. Debido a que en este trabajo de tesis se supone que la variable de estado \mathbf{x} es C^1 y satisface un modelo determinista, se pueden usar métodos de descenso por gradientes o algoritmos cuasi-Newton y sus variantes. Uno de los aspectos más costosos al usar estos métodos es calcular el gradiente o jacobiano de la función objetivo en cada iteración. En seguida, se describen algunos enfoques que se usan para calcular este gradiente.

2.2.1. Enfoque variacional

Sea \mathbf{z} definida por $\mathbf{z} = (\mathbf{x}_0, \boldsymbol{\theta})^T \in \mathbb{R}^{d+np}$, que contiene las condiciones iniciales y el vector de parámetros desconocidos para el sistema dinámico de (2.1). Entonces se puede escribir, con una precisión de primer orden,

$$\delta \mathcal{L} = \mathcal{L}(\mathbf{z} + \delta \mathbf{z}) - \mathcal{L}(\mathbf{z}) \approx \nabla \mathcal{L}(\mathbf{z}) \cdot \delta \mathbf{z}, \quad (2.5)$$

donde $\delta \mathbf{z} = (\delta \mathbf{x}_0, \delta \boldsymbol{\theta})^T$ es una perturbación para \mathbf{z} . Lo anterior se puede reescribir con más detalle como

$$\delta \mathcal{L} = \mathcal{L}(\mathbf{x}_0 + \delta \mathbf{x}_0, \boldsymbol{\theta} + \delta \boldsymbol{\theta}) - \mathcal{L}(\mathbf{x}_0, \boldsymbol{\theta}) \approx \nabla_{\mathbf{x}_0} \mathcal{L}(\mathbf{x}_0, \boldsymbol{\theta}) \cdot \delta \mathbf{x}_0 + \nabla_{\boldsymbol{\theta}} \mathcal{L}(\mathbf{x}_0, \boldsymbol{\theta}) \cdot \delta \boldsymbol{\theta}. \quad (2.6)$$

Con $\nabla_{\mathbf{x}_0}$ y $\nabla_{\boldsymbol{\theta}}$ se denotan los gradientes con respecto a \mathbf{x}_0 y $\boldsymbol{\theta}$, respectivamente. El lado derecho de (2.6), se puede desarrollar de acuerdo a la definición de \mathcal{L} , como sigue:

$$\begin{aligned} \mathcal{L}(\mathbf{x}_0 + \delta \mathbf{x}_0, \boldsymbol{\theta} + \delta \boldsymbol{\theta}) - \mathcal{L}(\mathbf{x}_0, \boldsymbol{\theta}) &= \frac{1}{2} \left(\left\| \frac{\mathbf{x}_0 + \delta \mathbf{x}_0 - \mathbf{s}_0}{\boldsymbol{\sigma}_0} \right\|_d^2 + \sum_{i=1}^m \left\| \frac{\bar{\mathbf{x}}(t_i; \mathbf{x}_0 + \delta \mathbf{x}_0, \boldsymbol{\theta} + \delta \boldsymbol{\theta}) - \bar{\mathbf{x}}_i}{\bar{\boldsymbol{\sigma}}_i} \right\|_d^2 \right) \\ &\quad - \frac{1}{2} \left(\left\| \frac{\mathbf{x}_0 - \mathbf{s}_0}{\boldsymbol{\sigma}_0} \right\|_d^2 + \sum_{i=1}^m \left\| \frac{\bar{\mathbf{x}}(t_i; \mathbf{x}_0, \boldsymbol{\theta}) - \bar{\mathbf{x}}_i}{\bar{\boldsymbol{\sigma}}_i} \right\|_d^2 \right) \\ &= \frac{1}{2} \left(\left\| \frac{\mathbf{x}_0 + \delta \mathbf{x}_0 - \mathbf{s}_0}{\boldsymbol{\sigma}_0} \right\|_d^2 - \left\| \frac{\mathbf{x}_0 - \mathbf{s}_0}{\boldsymbol{\sigma}_0} \right\|_d^2 \right) \\ &\quad + \frac{1}{2} \left(\sum_{i=1}^m \left(\left\| \frac{\bar{\mathbf{x}}(t_i; \mathbf{x}_0 + \delta \mathbf{x}_0, \boldsymbol{\theta} + \delta \boldsymbol{\theta}) - \bar{\mathbf{x}}_i}{\bar{\boldsymbol{\sigma}}_i} \right\|_d^2 - \left\| \frac{\bar{\mathbf{x}}(t_i; \mathbf{x}_0, \boldsymbol{\theta}) - \bar{\mathbf{x}}_i}{\bar{\boldsymbol{\sigma}}_i} \right\|_d^2 \right) \right). \end{aligned} \quad (2.7)$$

Nótese que

$$\begin{aligned} \left\| \frac{\mathbf{x}_0 + \delta \mathbf{x}_0 - \mathbf{s}_0}{\boldsymbol{\sigma}_0} \right\|_d^2 &= \left\| \frac{(\mathbf{x}_0 - \mathbf{s}_0) + \delta \mathbf{x}_0}{\boldsymbol{\sigma}_0} \right\|_d^2 = \frac{((\mathbf{x}_0 - \mathbf{s}_0) + \delta \mathbf{x}_0)^2}{\boldsymbol{\sigma}_0^2} \\ &= \frac{(\mathbf{x}_0 - \mathbf{s}_0)^2 + 2(\mathbf{x}_0 - \mathbf{s}_0) \cdot \delta \mathbf{x}_0 + (\delta \mathbf{x}_0)^2}{\boldsymbol{\sigma}_0^2}, \end{aligned}$$

donde se está usando la notación $\mathbf{x}^2 \equiv \mathbf{x} \cdot \mathbf{x}$, es decir, el producto punto de \mathbf{x} con él mismo. Despreciando el término cuadrático $(\delta \mathbf{x}_0)^2$ resulta que

$$\left\| \frac{\mathbf{x}_0 + \delta \mathbf{x}_0 - \mathbf{s}_0}{\boldsymbol{\sigma}_0} \right\|_d^2 \approx \frac{(\mathbf{x}_0 - \mathbf{s}_0)^2 + 2(\mathbf{x}_0 - \mathbf{s}_0) \cdot \delta \mathbf{x}_0}{\boldsymbol{\sigma}_0^2}. \quad (2.8)$$

Además,

$$\left\| \frac{\mathbf{x}_0 - \mathbf{s}_0}{\boldsymbol{\sigma}_0} \right\|_d^2 = \frac{(\mathbf{x}_0 - \mathbf{s}_0)^2}{\boldsymbol{\sigma}_0^2}. \quad (2.9)$$

Haciendo la diferencia de (2.8) y (2.9)

$$\left\| \frac{\mathbf{x}_0 + \delta \mathbf{x}_0 - \mathbf{s}_0}{\boldsymbol{\sigma}_0} \right\|_d^2 - \left\| \frac{\mathbf{x}_0 - \mathbf{s}_0}{\boldsymbol{\sigma}_0} \right\|_d^2 \approx \frac{2(\mathbf{x}_0 - \mathbf{s}_0) \cdot \delta \mathbf{x}_0}{\boldsymbol{\sigma}_0^2}. \quad (2.10)$$

Para los términos dentro de la sumatoria de (2.7) se considera la expansión en serie de Taylor de primer orden alrededor del punto $(\mathbf{x}_0, \boldsymbol{\theta})$ para $\bar{\mathbf{x}}(t_i; \mathbf{x}_0 + \delta \mathbf{x}_0, \boldsymbol{\theta} + \delta \boldsymbol{\theta})$. De esta manera,

$$\bar{\mathbf{x}}(t_i; \mathbf{x}_0 + \delta \mathbf{x}_0, \boldsymbol{\theta} + \delta \boldsymbol{\theta}) \approx \bar{\mathbf{x}}(t_i; \mathbf{x}_0, \boldsymbol{\theta}) + \frac{\partial \bar{\mathbf{x}}(t_i; \mathbf{x}_0, \boldsymbol{\theta})}{\partial \mathbf{x}_0} \delta \mathbf{x}_0 + \frac{\partial \bar{\mathbf{x}}(t_i; \mathbf{x}_0, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \delta \boldsymbol{\theta}. \quad (2.11)$$

Por lo que

$$\begin{aligned} \left\| \frac{\bar{\mathbf{x}}(t_i; \mathbf{x}_0 + \delta \mathbf{x}_0, \boldsymbol{\theta} + \delta \boldsymbol{\theta}) - \bar{\mathbf{x}}_i}{\bar{\boldsymbol{\sigma}}_i} \right\|_{no}^2 &\approx \left\| \frac{\bar{\mathbf{x}}(t_i; \mathbf{x}_0, \boldsymbol{\theta}) + \frac{\partial \bar{\mathbf{x}}(t_i; \mathbf{x}_0, \boldsymbol{\theta})}{\partial \mathbf{x}_0} \delta \mathbf{x}_0 + \frac{\partial \bar{\mathbf{x}}(t_i; \mathbf{x}_0, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \delta \boldsymbol{\theta} - \bar{\mathbf{x}}_i}{\bar{\boldsymbol{\sigma}}_i} \right\|_d^2 \\ &= \left\| \frac{(\bar{\mathbf{x}}(t_i; \mathbf{x}_0, \boldsymbol{\theta}) - \bar{\mathbf{x}}_i) + \left(\frac{\partial \bar{\mathbf{x}}(t_i; \mathbf{x}_0, \boldsymbol{\theta})}{\partial \mathbf{x}_0} \delta \mathbf{x}_0 + \frac{\partial \bar{\mathbf{x}}(t_i; \mathbf{x}_0, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \delta \boldsymbol{\theta} \right)}{\bar{\boldsymbol{\sigma}}_i} \right\|_d^2 \\ &= \frac{\left((\bar{\mathbf{x}}(t_i; \mathbf{x}_0, \boldsymbol{\theta}) - \bar{\mathbf{x}}_i) + \left(\frac{\partial \bar{\mathbf{x}}(t_i; \mathbf{x}_0, \boldsymbol{\theta})}{\partial \mathbf{x}_0} \delta \mathbf{x}_0 + \frac{\partial \bar{\mathbf{x}}(t_i; \mathbf{x}_0, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \delta \boldsymbol{\theta} \right) \right)^2}{\bar{\boldsymbol{\sigma}}_i^2}. \end{aligned}$$

Se debe desarrollar el binomio cuadrado del numerador, pero antes, para simplificar se usará la notación $D_{\mathbf{x}_0} \bar{\mathbf{x}} = \frac{\partial \bar{\mathbf{x}}(t_i; \mathbf{x}_0, \boldsymbol{\theta})}{\partial \mathbf{x}_0}$, $D_{\boldsymbol{\theta}} \bar{\mathbf{x}} = \frac{\partial \bar{\mathbf{x}}(t_i; \mathbf{x}_0, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$ y $\bar{\mathbf{x}} = \bar{\mathbf{x}}(t_i; \mathbf{x}_0, \boldsymbol{\theta})$. Con esto se obtiene

$$\begin{aligned} ((\bar{\mathbf{x}} - \bar{\mathbf{x}}_i) + ((D_{\mathbf{x}_0} \bar{\mathbf{x}}) \delta \mathbf{x}_0 + (D_{\boldsymbol{\theta}} \bar{\mathbf{x}}) \delta \boldsymbol{\theta}))^2 &= (\bar{\mathbf{x}} - \bar{\mathbf{x}}_i)^2 + 2(\bar{\mathbf{x}} - \bar{\mathbf{x}}_i) \cdot ((D_{\mathbf{x}_0} \bar{\mathbf{x}}) \delta \mathbf{x}_0 + (D_{\boldsymbol{\theta}} \bar{\mathbf{x}}) \delta \boldsymbol{\theta}) \\ &\quad + ((D_{\mathbf{x}_0} \bar{\mathbf{x}}) \delta \mathbf{x}_0 + (D_{\boldsymbol{\theta}} \bar{\mathbf{x}}) \delta \boldsymbol{\theta})^2. \end{aligned}$$

Despreciando el término cuadrático $((D_{\mathbf{x}_0} \bar{\mathbf{x}}) \delta \mathbf{x}_0 + (D_{\boldsymbol{\theta}} \bar{\mathbf{x}}) \delta \boldsymbol{\theta})^2$, se tendrá

$$\left\| \frac{\bar{\mathbf{x}}(t_i; \mathbf{x}_0 + \delta \mathbf{x}_0, \boldsymbol{\theta} + \delta \boldsymbol{\theta}) - \bar{\mathbf{x}}_i}{\bar{\boldsymbol{\sigma}}_i} \right\|_d^2 \approx \frac{(\bar{\mathbf{x}} - \bar{\mathbf{x}}_i)^2 + 2(\bar{\mathbf{x}} - \bar{\mathbf{x}}_i) \cdot ((D_{\mathbf{x}_0} \bar{\mathbf{x}}) \delta \mathbf{x}_0 + (D_{\boldsymbol{\theta}} \bar{\mathbf{x}}) \delta \boldsymbol{\theta})}{\bar{\boldsymbol{\sigma}}_i^2}. \quad (2.12)$$

Por otro lado,

$$\left\| \frac{\bar{\mathbf{x}}(t_i; \mathbf{x}_0, \boldsymbol{\theta}) - \bar{\mathbf{x}}_i}{\bar{\boldsymbol{\sigma}}_i} \right\|_d^2 = \frac{(\bar{\mathbf{x}}(t_i; \mathbf{x}_0, \boldsymbol{\theta}) - \bar{\mathbf{x}}_i)^2}{\bar{\boldsymbol{\sigma}}_i^2}. \quad (2.13)$$

Así, de restar (2.13) de (2.12) se obtiene que

$$\left\| \frac{\bar{\mathbf{x}}(t_i; \mathbf{x}_0 + \delta \mathbf{x}_0, \boldsymbol{\theta} + \delta \boldsymbol{\theta}) - \bar{\mathbf{x}}_i}{\bar{\boldsymbol{\sigma}}_i} \right\|_d^2 - \left\| \frac{\bar{\mathbf{x}}(t_i; \mathbf{x}_0, \boldsymbol{\theta}) - \bar{\mathbf{x}}_i}{\bar{\boldsymbol{\sigma}}_i} \right\|_d^2 \approx \frac{2(\bar{\mathbf{x}} - \bar{\mathbf{x}}_i) \cdot ((D_{\mathbf{x}_0} \bar{\mathbf{x}}) \delta \mathbf{x}_0 + (D_{\boldsymbol{\theta}} \bar{\mathbf{x}}) \delta \boldsymbol{\theta})}{\bar{\boldsymbol{\sigma}}_i^2}. \quad (2.14)$$

Sustituyendo (2.10) y (2.14) en (2.7) resulta

$$\delta \mathcal{L} = \mathcal{L}(\mathbf{x}_0 + \delta \mathbf{x}_0, \boldsymbol{\theta} + \delta \boldsymbol{\theta}) - \mathcal{L}(\mathbf{x}_0, \boldsymbol{\theta}) \approx \frac{(\mathbf{x}_0 - \mathbf{s}_0)}{\boldsymbol{\sigma}_0^2} \cdot \delta \mathbf{x}_0 + \sum_{i=1}^m \frac{(\bar{\mathbf{x}} - \bar{\mathbf{x}}_i)}{\bar{\boldsymbol{\sigma}}_i^2} \cdot ((D_{\mathbf{x}_0} \bar{\mathbf{x}}) \delta \mathbf{x}_0 + (D_{\boldsymbol{\theta}} \bar{\mathbf{x}}) \delta \boldsymbol{\theta}). \quad (2.15)$$

De forma equivalente,

$$\delta \mathcal{L} \approx \frac{\mathbf{x}_0 - \mathbf{s}_0}{\boldsymbol{\sigma}_0^2} \cdot \delta \mathbf{x}_0 + \sum_{i=1}^m \frac{\bar{\mathbf{x}}(t_i; \mathbf{x}_0, \boldsymbol{\theta}) - \bar{\mathbf{x}}_i}{\bar{\boldsymbol{\sigma}}_i^2} \cdot \left(\frac{\partial \bar{\mathbf{x}}(t_i; \mathbf{x}_0, \boldsymbol{\theta})}{\partial \mathbf{x}_0} \delta \mathbf{x}_0 + \frac{\partial \bar{\mathbf{x}}(t_i; \mathbf{x}_0, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \delta \boldsymbol{\theta} \right), \quad (2.16)$$

Asimismo, se tiene que

$$\delta\mathcal{L} \approx \frac{(\mathbf{x}_0 - \mathbf{s}_0)}{\sigma_0^2} \cdot \delta\mathbf{x}_0 + \sum_{i=1}^m \left((D_{\mathbf{x}_0}\bar{\mathbf{x}})^T \frac{(\bar{\mathbf{x}} - \bar{\mathbf{x}}_i)}{\bar{\sigma}_i^2} \cdot \delta\mathbf{x}_0 + (D_{\theta}\bar{\mathbf{x}})^T \frac{(\bar{\mathbf{x}} - \bar{\mathbf{x}}_i)}{\bar{\sigma}_i^2} \cdot \delta\theta \right)$$

o equivalentemente

$$\delta\mathcal{L} \approx \left(\frac{(\mathbf{x}_0 - \mathbf{s}_0)}{\sigma_0^2} + \sum_{i=1}^m (D_{\mathbf{x}_0}\bar{\mathbf{x}})^T \frac{(\bar{\mathbf{x}} - \bar{\mathbf{x}}_i)}{\bar{\sigma}_i^2} \right) \cdot \delta\mathbf{x}_0 + \sum_{i=1}^m (D_{\theta}\bar{\mathbf{x}})^T \frac{(\bar{\mathbf{x}} - \bar{\mathbf{x}}_i)}{\bar{\sigma}_i^2} \cdot \delta\theta. \quad (2.17)$$

La expresión (2.17) está prácticamente en la forma

$$\delta\mathcal{L} \approx \text{Algo} \cdot \delta\mathbf{z},$$

salvo términos de segundo orden, con lo que se concluye que $\nabla\mathcal{L}(\mathbf{z})$ es igual a la expresión *Algo*, es decir

$$\nabla\mathcal{L}(\mathbf{z}) = \left(\frac{(\mathbf{x}_0 - \mathbf{s}_0)}{\sigma_0^2} + \sum_{i=1}^m (D_{\mathbf{x}_0}\bar{\mathbf{x}})^T \frac{(\bar{\mathbf{x}} - \bar{\mathbf{x}}_i)}{\bar{\sigma}_i^2}, \sum_{i=1}^m (D_{\theta}\bar{\mathbf{x}})^T \frac{(\bar{\mathbf{x}} - \bar{\mathbf{x}}_i)}{\bar{\sigma}_i^2} \right).$$

El problema con esta expresión del gradiente es que involucra las derivadas $D_{\mathbf{x}_0}\bar{\mathbf{x}}$ y $D_{\theta}\bar{\mathbf{x}}$. Las matrices $\partial\bar{\mathbf{x}}(t_i; \mathbf{x}_0, \theta)/\partial\mathbf{x}_0$ y $\partial\bar{\mathbf{x}}(t_i; \mathbf{x}_0, \theta)/\partial\theta$ de (2.16) son los jacobianos de las variables observables $\bar{\mathbf{x}}(t_i; \mathbf{x}_0, \theta)$ en con respecto a \mathbf{x}_0 y θ , respectivamente. Estas derivadas parciales, comúnmente son conocidas como “sensibilidades de las observables” y su evaluación puede requerir de un esfuerzo computacional grande. La primera matriz tiene dimensiones $d \times d$ y la segunda $d \times np$, sin embargo, algunos de los renglones de estas matrices son 0, por lo que el número total de derivadas parciales a calcular en (2.16) es $m \times n_o(d + np)$, donde n_o representa el número de variables observables.

Nota 2.2.1. En la deducción de (2.17) se utiliza la siguiente propiedad matricial en combinación con el producto escalar:

Si M es una matriz de $m \times n$, v un vector de $m \times 1$ y w un vector de $n \times 1$, se cumple que $v \cdot Mw = (M^T v) \cdot w$.

Una justificación de esto se sigue de considerar x e y vectores columna de dimensión $n \times 1$ y recordar que $x \cdot y = x^T y$; así

$$v \cdot (Mw) = v^T (Mw) = v^T (M^T)^T w = (M^T v)^T w = (M^T v) \cdot w.$$

2.3. Método variacional para calcular sensibilidades

Como se explica en (Juárez Valencia & Rojas, 2022), el método variacional para calcular sensibilidades (las matrices $D_{\mathbf{x}_0}\bar{\mathbf{x}}$ y $D_{\boldsymbol{\theta}}\bar{\mathbf{x}}$) es también conocido como el “enfoque directo”, debido a que para calcularlas se debe resolver un problema de valor inicial para el cual las variables dependientes son las derivadas parciales que definen las sensibilidades. Para ver esto, lo primero a considerar es el cálculo de las sensibilidades con respecto al parámetro $\boldsymbol{\theta}$. Para ello el jacobiano $\partial\mathbf{x}(t; \mathbf{x}_0, \boldsymbol{\theta})/\partial\boldsymbol{\theta}$ se denotará como $J(t; \mathbf{x}_0, \boldsymbol{\theta})$, para simplificar la notación. De aquí:

$$J(t; \mathbf{x}_0, \boldsymbol{\theta}) \delta\boldsymbol{\theta} \approx \mathbf{x}(t; \mathbf{x}_0, \boldsymbol{\theta} + \delta\boldsymbol{\theta}) - \mathbf{x}(t; \mathbf{x}_0, \boldsymbol{\theta}).$$

Derivando con respecto a t ,

$$\begin{aligned} \dot{J}(t; \boldsymbol{\theta})\delta\boldsymbol{\theta} &\approx \dot{\mathbf{x}}(t; \mathbf{x}_0, \boldsymbol{\theta} + \delta\boldsymbol{\theta}) - \dot{\mathbf{x}}(t; \mathbf{x}_0, \boldsymbol{\theta}) \\ &= \mathbf{f}(\mathbf{x}(t; \mathbf{x}_0, \boldsymbol{\theta} + \delta\boldsymbol{\theta}), \boldsymbol{\theta} + \delta\boldsymbol{\theta}) - \mathbf{f}(\mathbf{x}(t; \mathbf{x}_0, \boldsymbol{\theta}), \boldsymbol{\theta}) \\ &\approx \mathbf{f}_{\mathbf{x}}(\mathbf{x}(t; \mathbf{x}_0, \boldsymbol{\theta}), \boldsymbol{\theta}) J(t; \mathbf{x}_0, \boldsymbol{\theta}) \delta\boldsymbol{\theta} + \mathbf{f}_{\boldsymbol{\theta}}(\mathbf{x}(t; \mathbf{x}_0, \boldsymbol{\theta}), \boldsymbol{\theta}) \delta\boldsymbol{\theta}. \end{aligned}$$

Entonces, $\mathbf{x}(t) = \mathbf{x}(t; \mathbf{x}_0, \boldsymbol{\theta})$ y $J(t) = J(t; \mathbf{x}_0, \boldsymbol{\theta})$ satisfacen las siguientes ecuaciones:

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{f}(\mathbf{x}(t), \boldsymbol{\theta}), \quad t_0 < t \leq t_f, \\ \mathbf{x}(t_0) &= \mathbf{x}_0; \\ \dot{J}(t) &= \mathbf{f}_{\mathbf{x}}(\mathbf{x}(t), \boldsymbol{\theta})J(t) + \mathbf{f}_{\boldsymbol{\theta}}(\mathbf{x}(t), \boldsymbol{\theta}), \quad t_0 < t \leq t_f, \\ J(t_0) &= \mathbb{O}. \end{aligned} \tag{2.18}$$

Las ecuaciones para el jacobiano J forman un sistema matricial de ecuaciones diferenciales y la condición inicial para J resulta de

$$J(t_0) = J(t_0; \mathbf{x}_0, \boldsymbol{\theta}) = \frac{\partial\mathbf{x}(t_0; \mathbf{x}_0, \boldsymbol{\theta})}{\partial\boldsymbol{\theta}} = \frac{\partial\mathbf{x}_0}{\partial\boldsymbol{\theta}} = \mathbb{O} \quad (\text{matriz nula}).$$

La condición inicial para J está resaltando el hecho de que \mathbf{x}_0 no depende de $\boldsymbol{\theta}$. De manera similar, se puede concluir para la matriz de sensibilidades con respecto a \mathbf{x}_0 , que satisface un sistema de ecuaciones diferenciales similar al anterior. Así que, computacionalmente, se debe encontrar la solución de sistemas matriciales de ecuaciones diferenciales que deben ser evaluadas en cada tiempo experimental $t_i, 1 \leq i \leq m$. Con esto queda claro que el trabajo computacional requerido se incrementa en cada iteración, por lo que en algunos casos llega a ser excesivo. Para más detalles ver (Juárez Valencia & Rojas, 2022). Esto motiva el estudio de métodos alternativos para calcular las sensibilidades observables (SO); en la siguiente sección se desarrollará el método de la adjunta. •

2.4. El método de la adjunta para calcular SO

Se desea calcular el gradiente de \mathcal{L} sin calcular directamente las matrices jacobianas, con el fin de evitar el trabajo computacional que esto implica. Para esto se utilizará una metodología que requiere de los siguientes espacios de Hilbert:

$$V = L^2([t_0, t_f]; \mathbb{R}^d) = \left\{ \mathbf{v} : [t_0, t_f] \rightarrow \mathbb{R}^d \mid \int_{t_0}^{t_f} \|\mathbf{v}(t)\|^2 dt < \infty \right\},$$

$$H = H^1([t_0, t_f]; \mathbb{R}^d) = \{ \mathbf{v} \in V \mid d\mathbf{v}/dt \in V \}.$$

El producto interno considerado en V está dado por $\langle \mathbf{u}, \mathbf{v} \rangle = \int_{t_0}^{t_f} \mathbf{u}(t) \cdot \mathbf{v}(t) dt$ que induce la norma $\|\mathbf{u}\|_V = \langle \mathbf{u}, \mathbf{u} \rangle^{1/2}$. Para ver que $\dot{\mathbf{x}}(t) \in V$, se debe verificar que $f(t, \mathbf{x}(t), \boldsymbol{\theta}) \in V$; esto se cumple ya que $f \in C^1(\mathbb{R}^{d+np})$. De esta manera, se concluye que $\mathbf{x} \in H$.

Del sistema de estado (2.1), se tiene

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \boldsymbol{\theta}).$$

Diferenciando ambos lados de la igualdad anterior

$$\delta \dot{\mathbf{x}}(t) = \mathbf{f}_{\mathbf{x}}(\mathbf{x}(t), \mathbf{x}_0, \boldsymbol{\theta}) \delta \mathbf{x}(t) + \mathbf{f}_{\mathbf{x}_0}(\mathbf{x}(t), \mathbf{x}_0, \boldsymbol{\theta}) \delta \mathbf{x}_0 + \mathbf{f}_{\boldsymbol{\theta}}(\mathbf{x}(t), \mathbf{x}_0, \boldsymbol{\theta}) \delta \boldsymbol{\theta}.$$

Y multiplicando esta igualdad por una función \mathbf{p} (arbitraria en H), se tendrá

$$\delta \dot{\mathbf{x}}(t) \cdot \mathbf{p}(t) = [\mathbf{f}_{\mathbf{x}}(\mathbf{x}(t), \mathbf{x}_0, \boldsymbol{\theta}) \delta \mathbf{x}(t) + \mathbf{f}_{\mathbf{x}_0}(\mathbf{x}(t), \mathbf{x}_0, \boldsymbol{\theta}) \delta \mathbf{x}_0 + \mathbf{f}_{\boldsymbol{\theta}}(\mathbf{x}(t), \mathbf{x}_0, \boldsymbol{\theta}) \delta \boldsymbol{\theta}] \cdot \mathbf{p}(t), \forall \mathbf{p} \in H.$$

Integrando se obtiene lo siguiente:

$$\int_{t_0}^{t_f} \delta \dot{\mathbf{x}}(t) \cdot \mathbf{p}(t) dt = \int_{t_0}^{t_f} [\mathbf{f}_{\mathbf{x}}(\mathbf{x}(t), \mathbf{x}_0, \boldsymbol{\theta}) \delta \mathbf{x}(t) + \mathbf{f}_{\mathbf{x}_0}(\mathbf{x}(t), \mathbf{x}_0, \boldsymbol{\theta}) \delta \mathbf{x}_0 + \mathbf{f}_{\boldsymbol{\theta}}(\mathbf{x}(t), \mathbf{x}_0, \boldsymbol{\theta}) \delta \boldsymbol{\theta}] \cdot \mathbf{p}(t) dt, \quad (2.19)$$

donde $\mathbf{f}_{\mathbf{x}}$, $\mathbf{f}_{\mathbf{x}_0}$ y $\mathbf{f}_{\boldsymbol{\theta}}$ son los jacobianos de \mathbf{f} con respecto a \mathbf{x} , \mathbf{x}_0 y $\boldsymbol{\theta}$, respectivamente.

La función \mathbf{p} es arbitraria en H , pero si se escoge adecuadamente, aportará información sobre las variaciones de la solución $\mathbf{x}(t)$ con respecto a \mathbf{x}_0 y $\boldsymbol{\theta}$, evitando el cálculo de las matrices jacobianas.

Analizando el lado izquierdo de la igualdad (2.19) y suponiendo para fines prácticos, que

usaremos mas adelante, que $\delta \mathbf{x} = (\delta x_1, \delta x_2, \delta x_3)^T$ y $\mathbf{p} = (p_1, p_2, p_3)^T$, entonces

$$\begin{aligned} \int_{t_0}^{t_f} \delta \dot{\mathbf{x}}(t) \cdot \mathbf{p}(t) dt &= \int_{t_0}^{t_f} [\delta \dot{x}_1(t)p_1(t) + \delta \dot{x}_2(t)p_2(t) + \delta \dot{x}_3(t)p_3(t)] dt \\ &= \int_{t_0}^{t_f} \delta \dot{x}_1(t)p_1(t) dt + \int_{t_0}^{t_f} \delta \dot{x}_2(t)p_2(t) dt + \int_{t_0}^{t_f} \delta \dot{x}_3(t)p_3(t) dt. \end{aligned}$$

Aplicando integración por partes a cada sumando de la expresión anterior, se tiene que

$$\begin{aligned} \int_{t_0}^{t_f} \delta \dot{\mathbf{x}}(t) \cdot \mathbf{p}(t) dt &= p_1(t)\delta x_1(t)|_{t_0}^{t_f} - \int_{t_0}^{t_f} \delta x_1(t)\dot{p}_1(t) dt + p_2(t)\delta x_2(t)|_{t_0}^{t_f} \\ &\quad - \int_{t_0}^{t_f} \delta x_2(t)\dot{p}_2(t) dt + p_3(t)\delta x_3(t)|_{t_0}^{t_f} - \int_{t_0}^{t_f} \delta x_3(t)\dot{p}_3(t) dt, \\ &= [p_1(t)\delta x_1(t) + p_2(t)\delta x_2(t) + p_3(t)\delta x_3(t)]_{t_0}^{t_f} \\ &\quad - \int_{t_0}^{t_f} [\delta x_1(t)\dot{p}_1(t) + \delta x_2(t)\dot{p}_2(t) + \delta x_3(t)\dot{p}_3(t)] dt, \\ &= [\mathbf{p}(t) \cdot \delta \mathbf{x}(t)]_{t_0}^{t_f} - \int_{t_0}^{t_f} \delta \mathbf{x}(t) \cdot \dot{\mathbf{p}}(t) dt. \end{aligned}$$

Sustituyendo esta expresión en el lado izquierdo de (2.19) se obtiene

$$\mathbf{p}(t) \cdot \delta \mathbf{x}(t)|_{t_0}^{t_f} - \int_{t_0}^{t_f} \dot{\mathbf{p}}(t) \cdot \delta \mathbf{x}(t) dt = \int_{t_0}^{t_f} [\mathbf{f}_x \delta \mathbf{x}(t) + \mathbf{f}_{x_0} \delta \mathbf{x}_0 + \mathbf{f}_\theta \delta \boldsymbol{\theta}] \cdot \mathbf{p}(t) dt,$$

y desarrollando el lado derecho de la igualdad queda:

$$\mathbf{p}(t) \cdot \delta \mathbf{x}(t)|_{t_0}^{t_f} - \int_{t_0}^{t_f} \dot{\mathbf{p}}(t) \cdot \delta \mathbf{x}(t) dt = \int_{t_0}^{t_f} [\mathbf{f}_x \delta \mathbf{x}(t) \cdot \mathbf{p}(t) + \mathbf{f}_{x_0} \delta \mathbf{x}_0 \cdot \mathbf{p}(t) + \mathbf{f}_\theta \delta \boldsymbol{\theta} \cdot \mathbf{p}(t)] dt.$$

Aplicando la propiedad de la nota 2.2.1, se sigue que

$$\mathbf{p}(t) \cdot \delta \mathbf{x}(t)|_{t_0}^{t_f} - \int_{t_0}^{t_f} \dot{\mathbf{p}}(t) \cdot \delta \mathbf{x}(t) dt = \int_{t_0}^{t_f} \{ [\mathbf{f}_x^T \mathbf{p}] (t) \cdot \delta \mathbf{x}(t) + [\mathbf{f}_{x_0}^T \mathbf{p}] (t) \cdot \delta \mathbf{x}_0(t) + [\mathbf{f}_\theta^T \mathbf{p}] (t) \cdot \delta \boldsymbol{\theta} \} dt.$$

Tomando en cuenta que \mathbf{f} no depende explícitamente de \mathbf{x}_0 , se tiene lo siguiente:

$$\mathbf{p}(t) \cdot \delta \mathbf{x}(t)|_{t_0}^{t_f} - \int_{t_0}^{t_f} \dot{\mathbf{p}}(t) \cdot \delta \mathbf{x}(t) dt = \int_{t_0}^{t_f} [\mathbf{f}_x^T \mathbf{p}] (t) \cdot \delta \mathbf{x}(t) + \int_{t_0}^{t_f} [\mathbf{f}_\theta^T \mathbf{p}] (t) \cdot \delta \boldsymbol{\theta} dt.$$

Agrupando términos resulta que

$$\mathbf{p}(t) \cdot \delta \mathbf{x}(t)|_{t_0}^{t_f} - \int_{t_0}^{t_f} \{ \dot{\mathbf{p}}(t) + [\mathbf{f}_x^T \mathbf{p}] (t) \} \cdot \delta \mathbf{x}(t) dt = \int_{t_0}^{t_f} [\mathbf{f}_\theta^T \mathbf{p}] (t) \cdot \delta \boldsymbol{\theta} dt. \quad (2.20)$$

En este punto vale la pena recordar que lo que se quiere hacer es calcular el gradiente de \mathcal{L} sin calcular directamente las SO. Para esto se tomará como referencia la ecuación (2.16). Es decir, se desea calcular la variación total de $\mathbf{x}(t; \mathbf{x}_0, \boldsymbol{\theta})$ con respecto a \mathbf{x}_0 y $\boldsymbol{\theta}$, que está dada de la siguiente manera:

$$\begin{aligned} \delta \mathbf{x}(t; \mathbf{x}_0, \boldsymbol{\theta}) &= \mathbf{x}(t; \mathbf{x}_0 + \delta \mathbf{x}_0, \boldsymbol{\theta} + \delta \boldsymbol{\theta}) - \mathbf{x}(t; \mathbf{x}_0, \boldsymbol{\theta}) \\ &\approx \frac{\partial \mathbf{x}(t; \mathbf{x}_0, \boldsymbol{\theta})}{\partial \mathbf{x}_0} \delta \mathbf{x}_0 + \frac{\partial \mathbf{x}(t; \mathbf{x}_0, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \delta \boldsymbol{\theta}. \end{aligned} \quad (2.21)$$

Para evitar el cálculo explícito de δx como en (2.21), se escogerá una \mathbf{p} de tal manera que a partir de (2.20) se obtenga una expresión práctica para $\delta \mathbf{x}(t)$. Para esto se escoge a \mathbf{p} como la solución del siguiente problema:

$$-\dot{\mathbf{p}}(t) = \mathbf{f}_{\mathbf{x}}(\mathbf{x}(t), \boldsymbol{\theta})^T \mathbf{p}(t) + \sum_{i=1}^m \frac{(\bar{\mathbf{x}}(t_i; \mathbf{x}_0, \boldsymbol{\theta}) - \bar{\mathbf{x}}_i)}{\sigma_i^2} \boldsymbol{\delta}_{\mathbf{D}}(t - t_i), t_f > t \geq t_0, \quad (2.22)$$

$$\mathbf{p}(t_f) = \mathbf{0}. \quad (2.23)$$

Aquí, m es el número de observaciones y $\boldsymbol{\delta}_{\mathbf{D}}(t - t_i)$ es la función de la medida de Dirac centrada en t_i . Al sistema (2.22) se le conoce como ecuación adjunta al sistema de estado o a la ecuación (2.1). La formulación variacional de este sistema adjunto es obtenida al multiplicar por una función de prueba diferenciable $\phi(t)$ e integrando:

$$-\int_{t_0}^{t_f} \{ \dot{\mathbf{p}}(t) + [\mathbf{f}_{\mathbf{x}}^T \mathbf{p}](t) \} \cdot \phi(t) dt = \int_{t_0}^{t_f} \sum_{i=1}^m \frac{(\bar{\mathbf{x}}(t_i; \mathbf{x}_0, \boldsymbol{\theta}) - \bar{\mathbf{x}}_i)}{\sigma_i^2} \boldsymbol{\delta}_{\mathbf{D}}(t - t_i) \cdot \phi(t) dt. \quad (2.24)$$

La integral del lado derecho puede ser escrita como

$$\sum_{i=1}^m \frac{(\bar{\mathbf{x}}(t_i; \mathbf{x}_0, \boldsymbol{\theta}) - \bar{\mathbf{x}}_i)}{\sigma_i^2} \cdot \phi(t_i).$$

Entonces (2.24) se convierte en

$$-\int_{t_0}^{t_f} \{ \dot{\mathbf{p}}(t) + [\mathbf{f}_{\mathbf{x}}^T \mathbf{p}](t) \} \cdot \phi(t) dt = \sum_{i=1}^m \frac{(\bar{\mathbf{x}}(t_i; \mathbf{x}_0, \boldsymbol{\theta}) - \bar{\mathbf{x}}_i)}{\sigma_i^2} \cdot \phi(t_i). \quad (2.25)$$

Al elegir $\phi(t) = \delta \bar{\mathbf{x}}(t)$ en la ecuación (2.25) y sustituir en la ecuación (2.20), se obtiene

$$\mathbf{p}(t) \cdot \delta \mathbf{x}(t) \Big|_{t_0}^{t_f} + \sum_{i=1}^m \frac{(\bar{\mathbf{x}}(t_i; \mathbf{x}_0, \boldsymbol{\theta}) - \bar{\mathbf{x}}_i)}{\sigma_i^2} \cdot \delta \bar{\mathbf{x}}(t_i) = \int_{t_0}^{t_f} [\mathbf{f}_{\boldsymbol{\theta}}^T \mathbf{p}](t) \cdot \delta \boldsymbol{\theta} dt. \quad (2.26)$$

Evaluando el primer término en los límites de integración y considerando que \mathbf{p} satisface

la condición final $\mathbf{p}(t_f) = 0$, se obtiene

$$-\mathbf{p}(t_0) \cdot \delta \mathbf{x}(t_0) + \sum_{i=1}^m \frac{(\bar{\mathbf{x}}(t_i; \mathbf{x}_0, \boldsymbol{\theta}) - \bar{\mathbf{x}}_i)}{\bar{\sigma}_i^2} \cdot \delta \bar{\mathbf{x}}(t_i) = \int_{t_0}^{t_f} [\mathbf{f}_\theta^T \mathbf{p}](t) \cdot \delta \boldsymbol{\theta} dt. \quad (2.27)$$

Que se puede reescribir como

$$\sum_{i=1}^m \frac{(\bar{\mathbf{x}}(t_i; \mathbf{x}_0, \boldsymbol{\theta}) - \bar{\mathbf{x}}_i)}{\bar{\sigma}_i^2} \cdot \delta \bar{\mathbf{x}}(t_i) = \mathbf{p}(t_0) \cdot \delta \mathbf{x}(t_0) + \int_{t_0}^{t_f} [\mathbf{f}_\theta^T \mathbf{p}](t) \cdot \delta \boldsymbol{\theta} dt. \quad (2.28)$$

Sustituyendo en (2.16)

$$\delta \mathcal{L} \approx \frac{\mathbf{x}_0 - \mathbf{s}_0}{\sigma_0^2} \cdot \delta \mathbf{x}_0 + \mathbf{p}(t_0) \cdot \delta \mathbf{x}(t_0) + \int_{t_0}^{t_f} [\mathbf{f}_\theta^T \mathbf{p}](t) \cdot \delta \boldsymbol{\theta} dt, \quad (2.29)$$

de donde resulta que

$$\begin{aligned} & \nabla_{\mathbf{x}_0} \mathcal{L}(\mathbf{x}_0, \boldsymbol{\theta}) \cdot \delta \mathbf{x}_0 + \nabla_{\boldsymbol{\theta}} \mathcal{L}(\mathbf{x}_0, \boldsymbol{\theta}) \cdot \delta \boldsymbol{\theta} \\ &= \frac{\mathbf{x}_0 - \mathbf{s}_0}{\sigma_0^2} \cdot \delta \mathbf{x}_0 + \mathbf{p}(t_0) \cdot \delta \mathbf{x}_0 + \left[\int_{t_0}^{t_f} [\mathbf{f}_\theta^T \mathbf{p}](t) dt \right] \cdot \delta \boldsymbol{\theta}. \end{aligned} \quad (2.30)$$

Así, el gradiente de la función objetivo es $\nabla \mathcal{L}(\mathbf{z}) = (\nabla_{\mathbf{x}_0} \mathcal{L}(\mathbf{x}_0, \boldsymbol{\theta}), \nabla_{\boldsymbol{\theta}} \mathcal{L}(\mathbf{x}_0, \boldsymbol{\theta}))^T$ y queda expresado por

$$\nabla_{\mathbf{x}_0} \mathcal{L}(\mathbf{x}_0, \boldsymbol{\theta}) = \frac{\mathbf{x}_0 - \mathbf{s}_0}{\sigma_0^2} + \mathbf{p}(t_0), \quad (2.31)$$

$$\nabla_{\boldsymbol{\theta}} \mathcal{L}(\mathbf{x}_0, \boldsymbol{\theta}) = \int_{t_0}^{t_f} \mathbf{f}_\theta(\mathbf{x}(t), \mathbf{x}_0, \boldsymbol{\theta})^T \mathbf{p}(t) dt. \quad (2.32)$$

Observación 2.4.1. En esta sección el problema abordado fue obtener una fórmula para el gradiente, que a diferencia de la fórmula (2.17), evite el cálculo de las matrices Jacobianas, con el fin de eficientar los cálculos. Al introducir la ecuación adjunta, se llega a la expresión (2.31) y (2.32) para el gradiente. Nótese que estas expresiones sólo requieren de la solución de la ecuación de estado (2.1) y la solución del sistema adjunto (2.22), lo que disminuye considerablemente los cálculos con respecto a la expresión (2.17).

2.5. Solución de la ecuación adjunta y cálculo del gradiente

El sistema adjunto (2.22) es un sistema para el cual se conoce una condición final en t_f y se debe resolver hacia atrás en el tiempo hasta t_0 , que es el valor que se utiliza en el cálculo del gradiente. Este sistema se puede resolver numéricamente hacia atrás en el tiempo pero en esta sección se muestra cómo transformarlo en un problema de valor inicial. Para eso

se introduce un cambio de variable adecuado y se aplica el método numérico RK4, que ya se describió en el Capítulo I. Se denota la nueva variable independiente como τ , y la nueva variable dependiente como \mathbf{p}_A . El cambio de variable se obtiene a partir de la relación $t_f - \tau = t - t_0$, con lo cual se tiene

$$\mathbf{p}(t) = \mathbf{p}(t_f - (\tau - t_0)) \equiv \mathbf{p}_A(\tau) \implies \dot{\mathbf{p}}(t) = -\dot{\mathbf{p}}_A(\tau), \quad (2.33)$$

obteniendo el sistema dinámico continuo (con dinámica hacia adelante en el tiempo)

$$\begin{aligned} \dot{\mathbf{p}}_A(\tau) &= \mathbf{f}_x(\mathbf{x}_A(\tau), \boldsymbol{\theta})^T \mathbf{p}_A(\tau) + \sum_{i=1}^m \frac{\bar{\mathbf{x}}_A(\tau_i, \mathbf{x}_0, \boldsymbol{\theta}) - \bar{\mathbf{x}}_{m+1-i}}{\sigma_{m+1-i}^2} \delta_D(\tau_i - \tau), \quad t_0 \leq \tau < t_f, \\ \mathbf{p}_A(\tau_0) &= \mathbf{0}. \end{aligned} \quad (2.34)$$

Este sistema tampoco es fácil de resolver de manera analítica, por lo cual se resuelve numéricamente. Una vez obtenido \mathbf{p}_A se utiliza el cambio de variable para obtener \mathbf{p} , y una vez obtenida \mathbf{p} se calcula la integral que define a la parte correspondiente del gradiente, dada por (2.32), la cual se puede resolver numéricamente por ejemplo con el método de Simpson.

Hasta este momento se ha definido el problema de ajuste mediante la minimización de un funcional, para el cual se cuenta con una fórmula eficiente para calcular su gradiente. Para minimizar numéricamente dicho funcional existen varios algoritmos, como el de descenso por el gradiente, gradiente conjugado, método de Newton, o el algoritmo BFGS, entre otros. En la siguiente sección ilustramos las ideas expuestas para el caso en el que el modelo corresponde al del SIDA que nos ocupa en este trabajo de tesis.

2.6. Ejemplificación de las ideas expuestas al caso del modelo del SIDA

Definición de la función objetivo. Para ejemplificar las ideas de la sección 2.1 consideramos el modelo que es de interés en este trabajo, en el que la variable de estado \mathbf{x} satisface el modelo estándar del SIDA (1.1):

$$\begin{aligned} \dot{T}_s &= q + \beta T_s \left(1 - \frac{T_s}{T_{max}}\right) - \mu_s T_s - k T_s V, \\ \dot{T}_e &= k T_s V - \mu_e T_e, \\ \dot{V} &= \rho T_e - \mu_V V, \end{aligned}$$

que se complementa con condiciones iniciales $T_s(t_0) = T_{s0}$, $T_e(t_0) = T_{e0}$, $V(t_0) = V_0$. En este caso, el número de ecuaciones del modelo es $d = 3$ y la variable de estado es el vector $\mathbf{x}(t) = (T_s(t), T_e(t), V(t))^T$. El número de parámetros es $np = 8$ (aunque no todos podrían ser desconocidos). En la figura 2.1 se muestra la solución $\mathbf{x}(t)$ usando el algoritmo *RK4*, en el intervalo de tiempo $[t_0, t_f] = [0, 30]$, usando los valores de los parámetros siguientes $\boldsymbol{\theta} = (q, \beta, T_{max}, \mu_s, k, \mu_e, \rho, \mu_v)^T = (10, 0.12, 1500, 1.33 \times 10^{-2}, 3.1662 \times 10^{-3}, 0.5, 2.6667, 0.4)^T$, que serán llamados valores exactos, y condiciones iniciales $x_0 = (70, 30, 200)^T$ que serán llamadas exactas. Estos valores se tomaron de la referencia (Perelson & Nelson, 1999; Perelson et al., 1993) ya que en esa referencia se muestran también series de tiempo asociadas con estos valores. En la figura 2.1 se muestra el comportamiento de las variables para estos valores de parámetros que corresponden a un caso en el que $R_0 = 60$ y el equilibrio no trivial $E^* = (23.6874, 24.9651, 166.4362)$ es localmente asintóticamente estable.

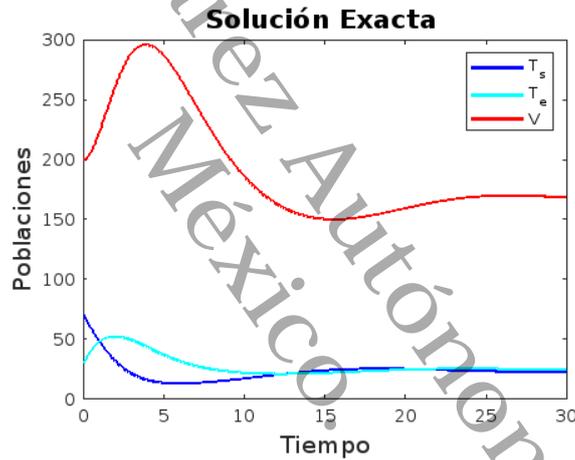


Figura 2.1: Solución exacta del modelo estándar del SIDA.

Cálculo del gradiente de \mathcal{L} con sensibilidades. En el caso del modelo del SIDA (1.1) de interés y suponiendo que el número de variables observables fuera $n_o = 3$ y si hay $m = 13$ mediciones experimentales, se tendrían que calcular $13 \times 3(3 + 8) = 429$ sensibilidades para obtener el gradiente en cada iteración de un método del tipo gradiente o cuasi-Newton.

El método de la adjunta para calcular SO. Nótese que para el modelo que se está trabajando, $\mathbf{x}(t) = \mathbf{x}(t; \mathbf{x}_0, \boldsymbol{\theta}) \in H$. Para verificar esto, primero se verá que $\mathbf{x}(t) \in V$; en nuestro caso $\mathbf{x}(t) = (T_s(t), T_e(t), V(t))$ y de acuerdo con los resultados sobre la estabilidad estudiados en el capítulo 1, se sabe que $\mathbf{x}(t)$ se va a un equilibrio, por lo que no toma valores infinitos y de esta manera la integral de su norma al cuadrado es finita.

Cálculo del gradiente y solución de la ecuación adjunta. Para el modelo del SIDA que se

está trabajando, si suponemos que $\bar{\mathbf{x}} = (T_s(t), T_e(t), V(t))^T$, entonces la ecuación adjunta (2.34) toma la siguiente forma:

$$\begin{bmatrix} \dot{p}_{A_1}(\tau) \\ \dot{p}_{A_2}(\tau) \\ \dot{p}_{A_3}(\tau) \end{bmatrix} = \begin{bmatrix} \beta \left(1 - 2\frac{T_s}{T_{max}}\right) - \mu_s - kV & kV & 0 \\ 0 & -\mu_e & \rho \\ -kT_s & kT_s & -\mu_V \end{bmatrix} \begin{bmatrix} p_{A_1}(\tau) \\ p_{A_2}(\tau) \\ p_{A_3}(\tau) \end{bmatrix} + \begin{bmatrix} \sum_{i=1}^m \frac{T_s(\tau_i) - T_{s_{m+1-i}}}{\sigma_{i1}^2} \delta_D(\tau_i - \tau) \\ \sum_{i=1}^m \frac{T_e(\tau_i) - T_{e_{m+1-i}}}{\sigma_{i2}^2} \delta_D(\tau_i - \tau) \\ \sum_{i=1}^m \frac{V(\tau_i) - V_{m+1-i}}{\sigma_{i3}^2} \delta_D(\tau_i - \tau) \end{bmatrix} \quad (2.35)$$

Una vez que se resuelve esta ecuación adjunta hacia adelante, se recupera la solución \mathbf{p} de la ecuación adjunta hacia atrás (2.22) tomando $\mathbf{p}(t_i) = \mathbf{p}_A(\tau_{m+1-i})$, $i = 1, \dots, m$. Para concluir con el cálculo del gradiente se debe calcular la integral en (2.32) cuyo integrando es

$$\begin{aligned} \mathbf{f}_\theta(\mathbf{x}(t), \theta)^T \mathbf{p}(t) &= \begin{bmatrix} 1 & 0 & 0 \\ T_s(t) \left(1 - \frac{T_s(t)}{T_{max}}\right) & 0 & 0 \\ \beta \left(\frac{T_s(t)}{T_{max}}\right)^2 & 0 & 0 \\ -T_s(t) & 0 & 0 \\ -T_s(t)V(t) & T_s(t)V(t) & 0 \\ 0 & -T_e(t) & 0 \\ 0 & 0 & T_e(t) \\ 0 & 0 & -V(t) \end{bmatrix} \begin{bmatrix} p_1(t) \\ p_2(t) \\ p_3(t) \end{bmatrix} \\ &= \begin{bmatrix} p_1(t) \\ T_s(t) \left(1 - \frac{T_s(t)}{T_{max}}\right) p_1(t) \\ \beta \left(\frac{T_s(t)}{T_{max}}\right)^2 p_1(t) \\ -T_s(t)p_1(t) \\ T_s(t)V(t)(p_2(t) - p_1(t)) \\ -T_e(t)p_2(t) \\ T_e(t)p_3(t) \\ -V(t)p_3(t) \end{bmatrix} \equiv \mathbf{I}(t). \end{aligned} \quad (2.36)$$

Nota 2.6.1. En la expresión (2.36) se están involucrando todos los parámetros, pero en el caso de que no todos se tengan que estimar, es necesario eliminar (no es lo mismo que poner un cero en esa posición) los respectivos renglones. Por ejemplo, si el parámetro 2 y el parámetro 5 ya están definidos y no se deben identificar, entonces se deben eliminar

las filas 2 y 5 de (2.36).

Finalmente, haciendo uso de la regla de Simpson en el conjunto de nodos de la malla temporal t_j entre t_0 y t_f , se tiene

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

donde \mathbf{I} se calcula como en la ecuación (2.36) de arriba. Con esto tendríamos disponible el gradiente y se podría usar cualquier método del tipo cuasi-Newton para minimizar \mathcal{L} . En el siguiente capítulo veremos cómo minimizar \mathcal{L} , para un modelo en general y para el modelo del SIDA en particular, utilizando el método cuasi-Newton BFGS.

Universidad Juárez Autónoma de Tabasco.
México.

Solución numérica del problema de minimización de \mathcal{L}

Los métodos cuasi-Newton son una familia de algoritmos iterativos para aproximar la solución de problemas de optimización no lineal como el que nos ocupa. Se derivan del método de Newton. Una de las principales ventajas de estos métodos es que mientras el método de Newton requiere del cálculo directo de la matriz Hessiana, los métodos cuasi-Newton construyen solo una aproximación de esta en cada iteración, usando información del gradiente previamente calculado. Al no requerir de las segundas derivadas, estos métodos son más eficientes que el método de Newton en términos de operaciones. En las siguientes secciones se explorará uno de los algoritmos cuasi-Newton más conocidos: el algoritmo BFGS, que implementaremos para resolver el problema de minimización (2.4), cuando el sistema de estado es el modelo del SIDA (1.1), para los cuales ya calculamos su gradiente en el capítulo 2.

3.1. El algoritmo BFGS

Siguiendo la descripción de (Nocedal & Wright, 1999), el algoritmo BFGS, nombrado así por sus creadores (Broyden, Fletcher, Goldfarb y Shanno) es el algoritmo cuasi-Newton más popular. Como todos los algoritmos de búsqueda y suponiendo que se tiene ya la aproximación \mathbf{z}^k , utiliza el gradiente en \mathbf{z}^k para definir una dirección de búsqueda \mathbf{d}^k y una vez definida esta, se minimiza la función objetivo restringida a la recta

$$\mathbf{z} = \mathbf{z}^k + \alpha \mathbf{d}^k.$$

Esto dará un valor óptimo para α , que si se denota como α_k , se toma la nueva aproxima-

ción como

$$\mathbf{z}^{k+1} = \mathbf{z}^k + \alpha_k \mathbf{d}^k.$$

Para calcular la dirección \mathbf{d}_k , el algoritmo BFGS, además del gradiente, utiliza una aproximación para la inversa de la Hessiana. Específicamente, el método BFGS utiliza la dirección de búsqueda

$$\mathbf{d}^k = -H^k \nabla f^k,$$

donde ∇f^k es el gradiente de f evaluado en \mathbf{z}^k y H^k es la aproximación a la inversa de la Hessiana en \mathbf{z}^k , que se calcula con la fórmula

$$H^{k+1} = (I - \gamma_k s_k y_k^T) H^k (I - \gamma_k y_k s_k^T) + \gamma_k s_k s_k^T, \quad (3.1)$$

donde

$$\gamma_k = \frac{1}{y_k^T s_k}, \quad (3.2)$$

y además, $s_k = \mathbf{z}^{k+1} - \mathbf{z}^k = \alpha_k \mathbf{d}^k$ y $y_k = \nabla f^{k+1} - \nabla f^k$.

Para la elección de la matriz inicial H^0 no existe una fórmula única, sin embargo, en algunos casos se opta por tomarla como la matriz identidad o un múltiplo de la misma. Estos pasos están resumidos y sistematizados en el siguiente algoritmo BFGS, tomado de (Nocedal & Wright, 1999).

Algoritmo BFGS

1. **Inicialización:** Dar un punto inicial \mathbf{z}^0 , una aproximación inicial a la inversa de la matriz Hessiana H^0 y una tolerancia $\varepsilon > 0$ para decidir cuando detener el proceso iterativo.

Tomar $k = 0$

2. **Descenso:** Mientras $\|\nabla f^k\| > \varepsilon$, hacer:

Calcular la dirección de búsqueda: $\mathbf{d}^k = -H^k \nabla f^k$;

Actualizar la aproximación: $\mathbf{z}^{k+1} = \mathbf{z}^k + \alpha_k \mathbf{d}^k$, donde α_k se calcula mediante el procedimiento de búsqueda en línea [$\alpha_k = \arg \min_{\alpha \geq 0} \varphi(\alpha) = f(\mathbf{z}^k + \alpha \mathbf{d}^k)$].

Calcular las diferencias entre las aproximaciones más recientes y los gradientes más recientes:

$s^k = \mathbf{z}^{k+1} - \mathbf{z}^k$; y $y_k = \nabla f^{k+1} - \nabla f^k$;

Actualizar la aproximación de la inversa de la Hessiana: Calcular H^{k+1}

usando la fórmula (3.1);

Hacer $k = k + 1$

Terminar el ciclo *Mientras*.

3. **Termina** el algoritmo

Nota 3.1.1. Para un modelo en particular, el algoritmo podría tener problemas durante la solución del problema de búsqueda en línea que se requiere para actualizar la aproximación.

En (Nocedal & Wright, 1999) se puede ver que este algoritmo tiene una tasa de convergencia súper lineal y las condiciones bajo las cuales converge local y globalmente a un mínimo de f .

Para la convergencia global, la función objetivo f debe cumplir las siguientes condiciones

- i) f es dos veces continuamente diferenciable.
- ii) El conjunto $\mathcal{N} = \{x \in \mathbb{R}^n | f(x) \leq f(x_0)\}$ es convexo y existen constantes positivas m y M tales que

$$m\|z\|^2 \leq z^T G(x)z \leq M\|z\|^2,$$

para todo $z \in \mathbb{R}^n$ y $x \in \mathcal{N}$. Aquí $G(x)$ representa la matriz Hessiana $\nabla^2 f$.

Esta condición implica que $G(x)$ es definida positiva en \mathcal{N} y que f tiene un único mínimo x^* en \mathcal{N} .

De aquí, se tiene el siguiente teorema.

Teorema 3.1.2 (Nocedal y Wright, 1999). Sea B_0 una matriz inicial simétrica y definida positiva y sea x_0 un punto de partida que cumple las condiciones dadas anteriormente. Entonces, la sucesión $\{x_k\}$ generada por el algoritmo BFGS (con $\epsilon = 0$) converge al mínimo x^* de f .

Para la tasa de convergencia superlineal se debe suponer que se cumple lo siguiente:

- iii) La matriz Hessiana G es Lipschitz continua en x^* , esto es

$$\|G(x) - G(x^*)\| \leq L\|x - x^*\|,$$

para todo x cerca de x^* , donde L es una constante positiva.

Considerando la afirmación anterior se tiene el siguiente teorema.

Teorema 3.1.3 (Nocedal y Wright, 1999). Sea f dos veces continuamente diferenciable. Supóngase que las iteraciones generadas por el algoritmo BFGS convergen a un mínimo x^* que cumple iii) y que se cumple además

$$\sum_{k=1}^{\infty} \|x^k - x^*\| < \infty.$$

Entonces, x^k converge a x^* con una tasa superlineal.

Las demostraciones de este teorema se pueden encontrar en (Nocedal & Wright, 1999). Note que en los teoremas la matriz inicial B_0 es una aproximación a la Hessiana, mientras que la matriz inicial H^0 en el algoritmo es una aproximación a la inversa de esa Hessiana.

En la práctica comúnmente se toma a H^0 como la matriz identidad. Enseguida, se da el algoritmo BFGS pero sustituyendo f por la función \mathcal{L} del problema (2.4) y agregando más detalles en cada paso.

Algoritmo BFGS

Inicialización:

1. Dar \mathbf{z}^0 y $H^0 = I$.
2. Calcular el gradiente inicial: $\mathbf{g}^0 = \nabla \mathcal{L}(\mathbf{z}^0)$.
3. Dar la dirección inicial $\mathbf{d}^0 = -H^0 \mathbf{g}^0 = -\mathbf{g}^0$.

Descenso: para $k \geq 0$, dados $\mathbf{z}^k, \mathbf{g}^k, \mathbf{d}^k, H^k$, determinar $\mathbf{z}^{k+1}, \mathbf{g}^{k+1}, \mathbf{d}^{k+1}, H^{k+1}$:

4. Encontrar $\alpha_k = \arg \min_{\alpha \geq 0} \varphi(\alpha) = \mathcal{L}(\mathbf{z}^k + \alpha \mathbf{d}^k)$.
5. Hacer $\mathbf{z}^{k+1} = \mathbf{z}^k + \alpha_k \mathbf{d}^k$.
6. Hacer $\mathbf{g}^{k+1} = \nabla \mathcal{L}(\mathbf{z}^{k+1})$.

Prueba de convergencia y nueva dirección:

Si $|\mathbf{g}^{k+1}| \leq \epsilon |\mathbf{g}^0|$, hacer el paso 7; en caso contrario, continuar en el paso 8.

7. Tomar $\mathbf{z}^* = \mathbf{z}^{k+1}$. Detenerse y salir.
8. Evaluar $\mathbf{u} = \Delta \mathbf{z}^k = \mathbf{z}^{k+1} - \mathbf{z}^k$ y $\mathbf{v} = \Delta \mathbf{g}^k = \mathbf{g}^{k+1} - \mathbf{g}^k$.
9. Actualizar $H^{k+1} = \left(I - \frac{\mathbf{v}\mathbf{u}^T}{\mathbf{u}^T \mathbf{v}} \right) H^k \left(I - \frac{\mathbf{u}\mathbf{v}^T}{\mathbf{u}^T \mathbf{v}} \right) + \frac{\mathbf{v}\mathbf{v}^T}{\mathbf{u}^T \mathbf{v}}$.
10. Hacer $\mathbf{d}^{k+1} = -H^{k+1} \mathbf{g}^{k+1}$.

11. Hacer $k = k + 1$ y regresar al paso 4.
12. Terminar el algoritmo.

3.2. Problemas de búsqueda en línea

Los problemas de búsqueda en línea aparecen comúnmente en los algoritmos iterativos de optimización, en particular, representa un paso clave en el algoritmo BFGS que se está trabajando (paso 4 de descenso), ya que este permite encontrar el tamaño de paso adecuado en la dirección de descenso. Para resolver este problema, se define $\varphi(\alpha) = \mathcal{L}(\mathbf{z}^k + \alpha \mathbf{d}^k)$ y el objetivo es minimizar φ :

$$\min_{\alpha > 0} \varphi(\alpha) = \min_{\alpha > 0} \mathcal{L}(\mathbf{z}^k + \alpha \mathbf{d}^k), \quad (3.3)$$

donde α es el tamaño de paso y $\mathcal{L}(\mathbf{z}^k + \alpha \mathbf{d}^k)$ es la función \mathcal{L} restringida a la recta $\mathbf{z}^k + \alpha \mathbf{d}^k$. El problema (3.3) se resuelve numéricamente con un método iterativo y se obtiene la nueva aproximación como

$$\mathbf{z}^{k+1} = \mathbf{z}^k + \alpha_k \mathbf{d}^k.$$

Para que un método de línea sea exitoso, se debe elegir de forma efectiva la dirección \mathbf{d}^k y el tamaño de paso α_k . Aquí, se aproximará el tamaño de paso α_k vía el método de la secante.

En la sección 2.3, se obtuvo una fórmula para calcular $\nabla \mathcal{L}(\mathbf{z})$, así que se puede calcular de manera eficiente la derivada $\varphi'(\alpha) = \nabla \mathcal{L}(\mathbf{z}^k + \alpha \mathbf{d}^k) \cdot \mathbf{d}^k$, lo que permitirá hacer uso del método de la secante. La fórmula de iteración del método de la secante es

$$\alpha_{k+1} = \alpha_k - \frac{\varphi'(\alpha_k)(\alpha_k - \alpha_{k-1})}{\varphi'(\alpha_k) - \varphi'(\alpha_{k-1})} = \frac{\alpha_{k-1}\varphi'(\alpha_k) - \alpha_k\varphi'(\alpha_{k-1})}{\varphi'(\alpha_k) - \varphi'(\alpha_{k-1})}, \quad k = 1, 2, \dots \quad (3.4)$$

Esta fórmula de iteración considera dos valores iniciales α_0 y α_1 . Para calcular φ' , se necesita resolver primero la ecuación de estado (2.1) tomando como condición inicial $\mathbf{x}_0^k + \alpha \mathbf{d}_0^k$ y como vector de parámetros $\boldsymbol{\theta}^k + \alpha \mathbf{d}_\theta^k$. La solución obtenida se denotará como \mathbf{x}_α^k y posteriormente se resuelve la ecuación adjunta (2.22) con $\mathbf{x} = \mathbf{x}_\alpha^k$ y $\boldsymbol{\theta} = \boldsymbol{\theta}^k + \alpha \mathbf{d}_\theta^k$, de donde se obtiene la solución \mathbf{p}_α^k . De las fórmulas (2.31) y (2.32) para el cálculo del gradiente se llega a

$$\varphi'(\alpha) = \begin{bmatrix} \mathbf{p}_\alpha^k(t_0) + \frac{\mathbf{x}_0^k + \alpha \mathbf{d}_0^k - \mathbf{s}_0}{\sigma_0^2} \\ \int_{t_0}^{t_f} \mathbf{f}_\theta(\mathbf{x}_\alpha^k, \boldsymbol{\theta}^k + \alpha \mathbf{d}_\theta^k)^T \mathbf{p}_\alpha^k dt \end{bmatrix} \cdot \begin{bmatrix} \mathbf{d}_0^k \\ \mathbf{d}_\theta^k \end{bmatrix}.$$

3.3. Detalles de programación del algoritmo BFGS

A continuación se presentan algunos detalles de los pasos más importantes sobre la programación del algoritmo BFGS.

1. En la inicialización del BFGS, el valor inicial de \mathbf{z} , \mathbf{z}^0 se obtiene perturbando los valores exactos con un mismo orden α como se perturbaron los datos. La matriz inicial H^0 no tiene una fórmula específica, por lo que en los experimentos, se tomará H^0 como la matriz identidad.
2. Tanto en la inicialización como en el descenso, es necesario calcular el gradiente de \mathcal{L} en \mathbf{z} . Para esto, se considera que $\nabla \mathcal{L}(\mathbf{z}^0)$ está compuesto por los gradientes de la función objetivo con respecto a \mathbf{x}_0 y con respecto a $\boldsymbol{\theta}$. Estos subgradiientes siguen las estructuras dadas en (2.31) y (2.32), respectivamente.
Para calcular los dos subgradiientes se necesita conocer la solución del sistema adjunto en \mathbf{x}_0^k y $\boldsymbol{\theta}^k$; aunque para la ecuación (2.31) se requiere la solución del sistema adjunto \mathbf{p} en t_0 , y para (2.32) la solución \mathbf{p} global, en todo el intervalo de tiempo. El integrando es un vector columna de dimensión igual que el número de parámetros a estimar, que se obtiene de multiplicar el jacobiano transpuesto de f con respecto a los parámetros a estimar, multiplicado por el vector solución \mathbf{p} que es siempre de dimensión 3. Como ya se describió en la sección 2.5, la integral de este vector se calcula usando el método numérico de Simpson.
Las componentes del gradiente con respecto a \mathbf{x}_0 se guardan siempre en las tres primeras componentes del gradiente general.
3. Cálculo de $\varphi'(\alpha)$. Este cálculo interviene en la solución del problema de búsqueda en línea por secante. Dados \mathbf{z}^k y \mathbf{d}^k se tiene que $\varphi'(\alpha) = \nabla \mathcal{L}(\mathbf{z}^k + \alpha \mathbf{d}^k) \cdot \mathbf{d}^k$. Así para obtener este valor, lo más interesante es calcular el gradiente de \mathcal{L} , en $\mathbf{z}^k + \alpha \mathbf{d}^k$, lo cual se hace de acuerdo al apartado anterior.

En la figura 3.1 se muestra un diagrama de estructura del programa que minimiza a \mathcal{L} , siguiendo el algoritmo BFGS. Una breve descripción de lo que hace cada módulo es:

1. **Genera datos (genera-datos):** Define el sistema de estado y se resuelve usando RK-4, posteriormente la solución se perturba añadiendo ruido aleatorio. Los datos perturbados se guardan en memoria.
2. **Archivo principal (función Main):** Considera los datos perturbados generados en Genera datos. También aquí se define el sistema adjunto y las matriz jacobiana con respecto a los parámetros, además se define nuevamente el sistema de estado. De

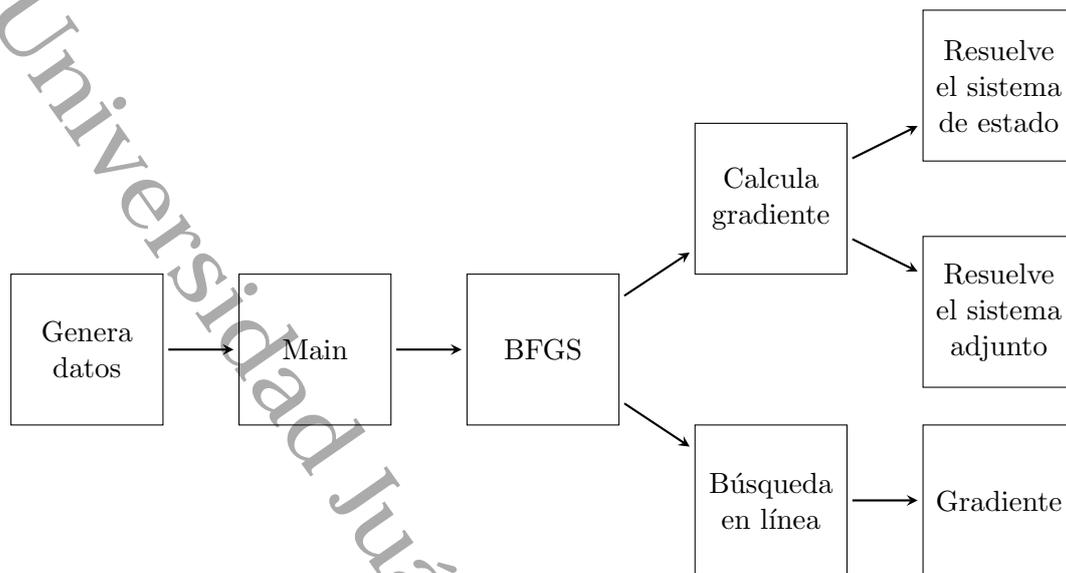


Figura 3.1: Diagrama de estructura del programa de minimización con el algoritmo BFGS.

igual manera, define aproximaciones iniciales de los parámetros y de las condiciones iniciales para finalmente llamar al algoritmo BFGS.

3. **BFGS (BFGS):** Contiene los pasos del algoritmo BFGS, mismo que fue descrito en secciones anteriores y que en cada paso del proceso iterativo requiere del cálculo del gradiente y de la solución del problema de búsqueda en línea:
 - * **Gradiente:** Aquí se calcula el gradiente con respecto a las condiciones iniciales y a los parámetros por cada iteración del algoritmo BFGS, además grafica la solución del sistema de estado y el sistema adjunto en cada iteración.
 - * **Búsqueda en línea:** Calcula el tamaño de paso por cada iteración del algoritmo BFGS, vía el método de la secante explicado en la sección anterior (este a su vez, requiere del cálculo del gradiente en la subiteración actual).

Nota 3.3.1. El sistema de estado se define nuevamente en el programa principal, esto se debe a que el programa “Genera datos” es un archivo tipo script, que genera datos resolviendo el sistema de estado y mantiene en memoria los valores generados para cada una de las variables, lo que se hace con el fin de poder repetir experimentos, si así se desea, con los mismos datos. Al ser el programa “Genera datos” un archivo tipo script no pueden enviar las ecuaciones del sistema como argumento.

Nota 3.3.2. A pesar de que se realizan varios cálculos en BFGS, al archivo principal solo se transfieren los valores aproximados de los parámetros, de las condiciones iniciales

y la norma del gradiente en cada iteración. Estos datos son utilizados para el cálculo de los errores y realizar las respectivas gráficas.

3.4. Experimentos numéricos considerando el modelo del SIDA

En esta sección se muestran los resultados obtenidos al aplicar al modelo matemático (1.1) del VIH/SIDA la metodología descrita en los capítulos 2 y 3. El modelo del SIDA (1.1) tiene ocho parámetros y de acuerdo con la literatura, se conocen valores para seis de ellos: $q, \beta, T_{max}, \mu_s, \rho, \mu_V$ (Culshaw et al., 2004; Perelson & Nelson, 1999; Perelson et al., 1993; Wein et al., 1997). Además, a través de estudios clínicos se pueden conocer datos para dos de las variables: T_s y V . Normalmente las condiciones iniciales del VIH/SIDA no se conocen, así que en todos los experimentos se consideran desconocidas las condiciones iniciales y fueron estimadas en todas las simulaciones. Tomando en cuenta todas las afirmaciones anteriores, los experimentos se organizaron en tres grandes grupos:

1. Estimar individualmente cada uno de los parámetros para los que es difícil clínicamente conocer su valor, k y μ_e y considerar los subcasos de una, dos o tres variables observables, es decir, $n_o = 1, 2, 3$.
2. Estimar los dos parámetros, k y μ_e simultáneamente y considerar los subcasos de una, dos o tres variables observables, es decir, $n_o = 1, 2, 3$.
3. Estimar los tres parámetros, β, k y μ_V simultáneamente y considerar los subcasos de una, dos o tres variables observables, es decir, $n_o = 1, 2, 3$. Los experimentos de este grupo no corresponden a un caso práctico y se hizo solamente con fines de validación de la metodología.

En todas las simulaciones se consideró al vector de valores $[q, \beta, T_{max}, \mu_s, k, \mu_e, \rho, \mu_V]^T = (10, 0.12, 1500, 1.33 \times 10^{-2}, 3.1662 \times 10^{-3}, 0.5, 2.6667, 0.4)^T$, como el vector de valores exactos. En todas las simulaciones el intervalo de tiempo fue $[t_0, t_f] = [0, 30]$. Las condiciones iniciales exactas siempre fueron $T_{s0} = 70, T_{e0} = 30$ y $V_0 = 200$. Con estos valores, en cada simulación se resolvió el modelo obteniendo datos numéricos para las tres variables de estado. De estos datos se escogieron m tiempos y los valores de las variables de estado en estos m tiempos fueron perturbados con un ruido gaussiano del 10% (esto implica $\sigma_i^2 = 0.1 * |x|$) y media 0. Estos datos perturbados son los que se toman como datos para el ajuste de parámetros. Para el criterio de paro en las iteraciones del algoritmo BFGS se estableció lo siguiente

$$\frac{\|\nabla \mathcal{L}(\mathbf{z}^{k+1})\|}{\|\nabla \mathcal{L}(\mathbf{z}^0)\|} < 10^{-4}.$$

Como ya se describió en los capítulos anteriores, para resolver el sistema de estado y el sistema adjunto se usó el método RK4. Para resolver los problemas de búsqueda en línea dentro del algoritmo BFGS se usó el método de la secante.

De todas las simulaciones un resumen de los resultados es que:

1. En general los parámetros se estiman bien usando pocos datos y aunque se hicieron experimentos usando una gran cantidad de datos, estos muestran que, a parte de ser caros e imprácticos, no se obtienen beneficios en cuanto a la estimación.
2. Para el criterio de paro dado en el algoritmo BFGS, convergió en pocas iteraciones.

En seguida se dará un resumen de las simulaciones mostrando para cada grupo de experimentos una tabla donde se muestran las variables observables consideradas, los parámetros a estimar (las tres condiciones iniciales siempre se estiman) y los valores de inicio en BFGS tanto para las condiciones iniciales como para los parámetros a estimar, los valores estimados, el error relativo para cada estimación, las iteraciones y el número de tiempos en los que se dispone de datos. Asociado con cada tabla se incluyen gráficas de resultados para algunas simulaciones: una gráfica que muestra los valores estimados de las condiciones iniciales por iteración de BFGS, una gráfica donde se muestra la evolución del valor estimado para los parámetros con respecto a cada iteración, otra gráfica donde se muestra la evolución de la norma del gradiente por cada iteración, aunque medido en escala logarítmica y también se incluye una gráfica donde se muestra el comportamiento de las variables de estado al tomar tanto los valores verdaderos de los parámetros y las condiciones iniciales, como los valores estimados.

3.4.1. Estimación de los parámetros μ_e y k individualmente

Variables observables.	Condiciones iniciales y parámetro a estimar.	Valores exactos.	Valores iniciales en BFGS.	Iteraciones en BFGS.	Núm. de tiempos observables.	Valores estimados.	Error relativo.
T_s	T_{s0}	70	64.76	2	100000	69.36	0.009091
	T_{e0}	30	25.98			28.03	0.065695
	V_0	200	203.49			203.65	0.018257
	k	3.1662×10^{-3}	1×10^{-3}			3.1834×10^{-3}	0.005433
T_s	T_{s0}	70	64.76	4	100	64.79	0.074473
	T_{e0}	30	25.98			25.99	0.133645
	V_0	200	203.49			203.50	0.017475
	k	3.1662×10^{-3}	1×10^{-3}			3.2665×10^{-3}	0.031673
V	T_{s0}	70	73.76	4	100000	71.68	0.023998
	T_{e0}	30	26.02			27.20	0.093379
	V_0	200	206.23			207.00	0.035023
	k	3.1662×10^{-3}	1.662×10^{-4}			3.1558×10^{-3}	0.003271
V	T_{s0}	70	73.76	4	10000	73.73	0.053332
	T_{e0}	30	26.02			26.31	0.123143
	V_0	200	206.23			206.34	0.031718
	k	3.1662×10^{-3}	1.662×10^{-4}			3.1674×10^{-3}	0.000373
V	T_{s0}	70	73.76	6	100	73.78	0.053983
	T_{e0}	30	26.02			25.64	0.145448
	V_0	200	206.23			205.96	0.029823
	k	3.1662×10^{-3}	1.662×10^{-4}			-0.07384	24.32278

Tabla 3.1: Resumen de las simulaciones al estimar el parámetro k tomando una variable observable.

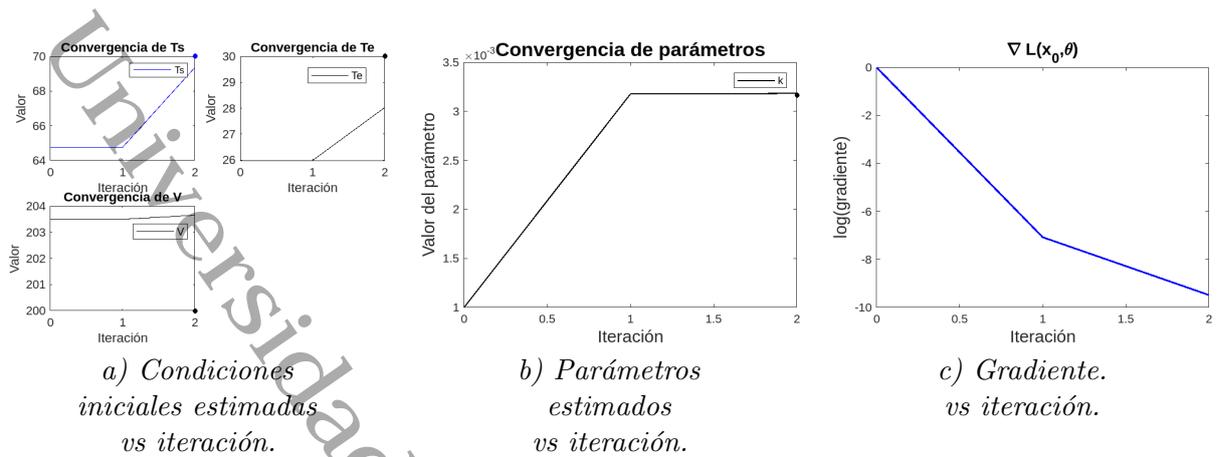


Figura 3.2: Resultados numéricos al estimar el parámetro k y las condiciones iniciales, considerando la variable T_s como observable y tomando cien mil datos.

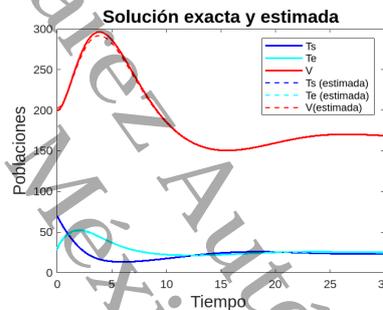


Figura 3.3: Solución exacta y estimada cuando se ajusta el parámetro k y las condiciones iniciales, considerando la variable T_s como observable y tomando cien mil datos, el algoritmo alcanzó convergencia en 2 iteraciones.

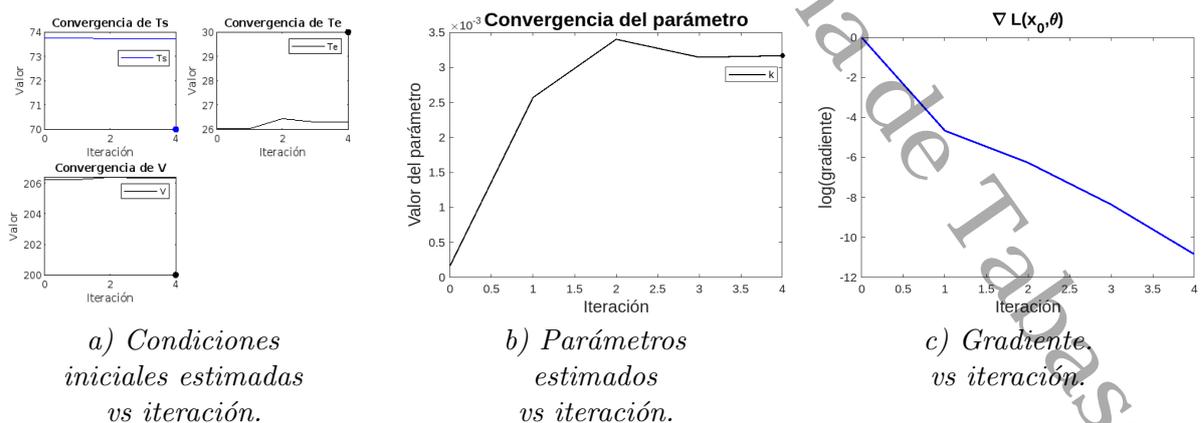


Figura 3.4: Resultados numéricos al estimar el parámetro k y las condiciones iniciales, considerando la variable V como observable y tomando diez mil datos.

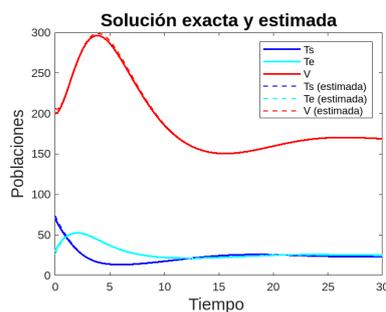


Figura 3.5: Solución exacta y estimada cuando se ajusta el parámetro k y las condiciones iniciales, considerando la variable V como observable y tomando diez mil datos, el algoritmo alcanzó convergencia en 4 iteraciones.

Variab- les ob- ser- va- bles.	Condi- ciones ini- ciales y paráme- tro a esti- mar.	Valores exactos.	Valores iniciales en BFGS.	Itera- cio- nes en BFGS.	Núm. de tiem- pos obser- vables.	Valores estimados.	Error relativo.
T_s V	T_{s0} T_{e0} V_0 k	70 30 200 3.1662×10^{-3}	64.63 31.05 219.44 1×10^{-3}	5	100000	70.11 26.12 212.69 3.1761×10^{-3}	0.001640 0.129323 0.063463 0.003122
T_s V	T_{s0} T_{e0} V_0 k	70 30 200 3.1662×10^{-3}	64.63 31.05 219.44 1×10^{-3}	3	100	64.65 31.05 219.44 3.1113×10^{-3}	0.076404 0.034917 0.097192 0.017340
T_s V	T_{s0} T_{e0} V_0 k	70 30 200 3.1662×10^{-3}	64.63 31.05 219.44 1×10^{-3}	3	50	64.64 31.05 219.44 3.1388×10^{-3}	0.076545 0.034889 0.097203 0.008648
T_s V	T_{s0} T_{e0} V_0 k	70 30 200 3.1662×10^{-3}	64.63 31.05 219.44 1×10^{-3}	3	5	64.63 31.05 219.44 3.0559×10^{-3}	0.076742 0.034908 0.097217 0.034826

Tabla 3.2: Resumen de las simulaciones al estimar el parámetro k tomando dos variables observables.

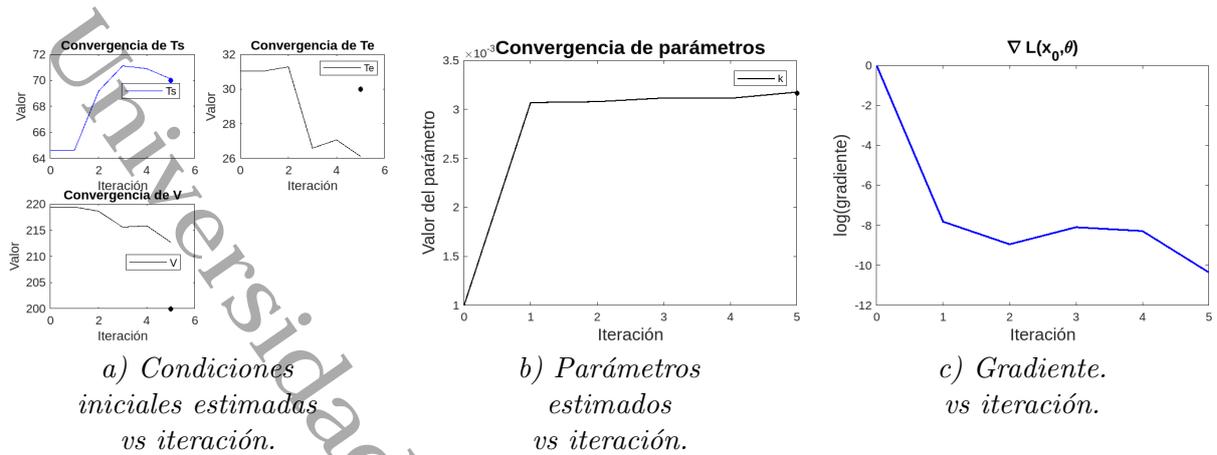


Figura 3.6: Resultados numéricos al estimar el parámetro k y las condiciones iniciales, considerando las variables T_s y V como observables y tomando cien mil datos.

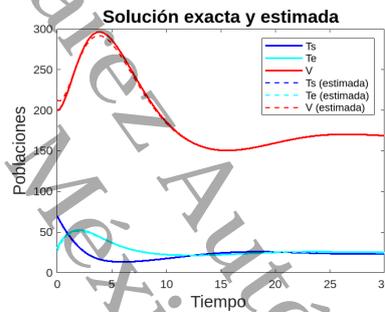


Figura 3.7: Solución exacta y estimada cuando se ajusta el parámetro k y las condiciones iniciales, considerando las variables T_s y V como observables y tomando cien mil datos, el algoritmo alcanzó convergencia en 5 iteraciones.

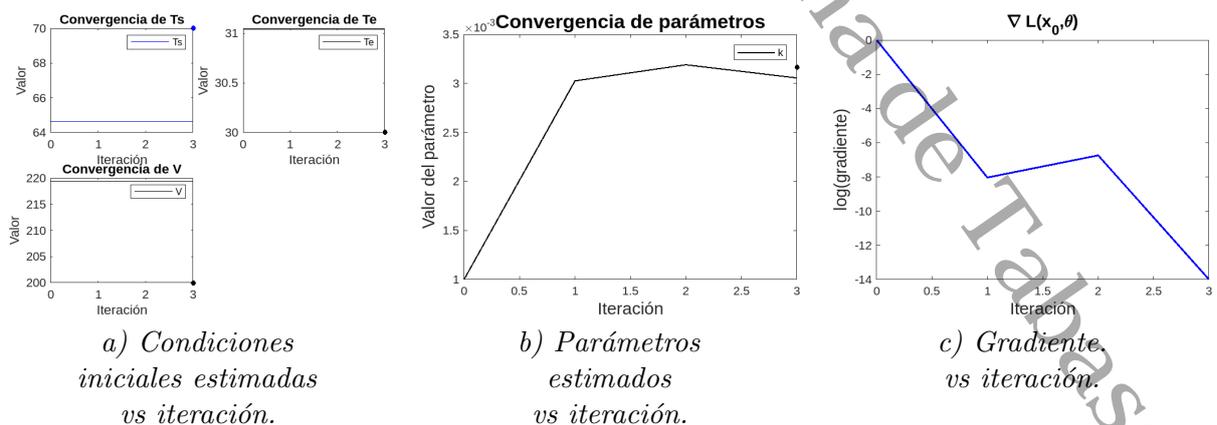


Figura 3.8: Resultados numéricos al estimar el parámetro k y las condiciones iniciales, considerando las variables T_s y V como observables y tomando cinco datos.

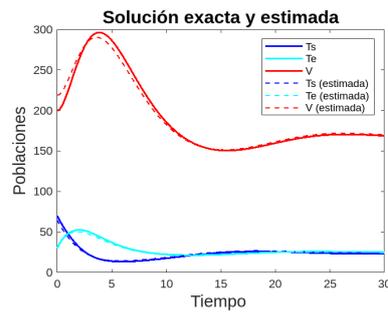


Figura 3.9: Solución exacta y estimada cuando se ajusta el parámetro k y las condiciones iniciales, considerando las variables T_s y V como observables y tomando cinco datos, el algoritmo alcanzó convergencia en 3 iteraciones.

Variables observables.	Condiciones iniciales y parámetro a estimar.	Valores exactos.	Valores iniciales en BFGS.	Iteraciones en BFGS.	Núm. de tiempos observables.	Valores estimados.	Error relativo.
T_s T_e V	T_{s0} T_{e0} V_0 k	70 30 200 3.1662×10^{-3}	73.76 26.02 206.23 1×10^{-3}	4	100000	72.18 27.00 208.79 3.1656×10^{-3}	0.031167 0.100129 0.043968 0.000179
T_s T_e V	T_{s0} T_{e0} V_0 k	70 30 200 3.1662×10^{-3}	73.76 26.02 206.23 1×10^{-3}	4	100	73.76 26.03 206.23 3.2521×10^{-3}	0.053724 0.132361 0.031160 0.027125
T_s T_e V	T_{s0} T_{e0} V_0 k	70 30 200 3.1662×10^{-3}	73.76 26.02 206.23 1×10^{-3}	4	50	73.77 26.02 206.23 3.1944×10^{-3}	0.053854 0.132568 0.031131 0.008908
T_s T_e V	T_{s0} T_{e0} V_0 k	70 30 200 3.1662×10^{-3}	73.76 26.02 206.23 1×10^{-3}	4	20	73.76 26.02 206.23 3.1910×10^{-3}	0.053774 0.132741 0.031128 0.007822
T_s T_e V	T_{s0} T_{e0} V_0 k	70 30 200 3.1662×10^{-3}	73.76 26.02 206.23 1×10^{-3}	4	10	73.76 26.02 206.23 3.200×10^{-3}	0.053746 0.132764 0.031129 0.010815
T_s T_e V	T_{s0} T_{e0} V_0 k	70 30 200 3.1662×10^{-3}	73.76 26.02 206.23 1×10^{-3}	4	5	73.77 26.02 206.23 3.2014×10^{-3}	0.053791 0.132729 0.031128 0.011122

Tabla 3.3: Resumen de las simulaciones al estimar el parámetro k tomando tres variables observables.

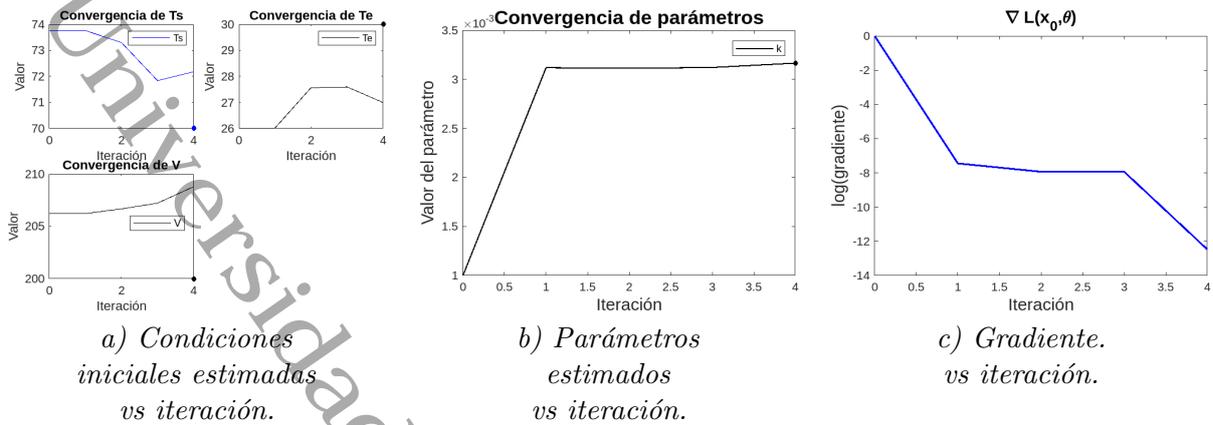


Figura 3.10: Resultados numéricos al estimar el parámetro k y las condiciones iniciales, considerando tres variables como observables y tomando cien mil datos.

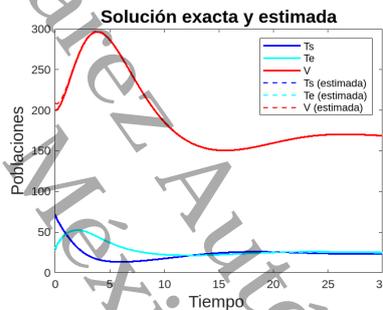


Figura 3.11: Solución exacta y estimada cuando se ajusta el parámetro k y las condiciones iniciales, considerando tres variables observables y tomando cien mil datos, el algoritmo alcanzó convergencia en 4 iteraciones.

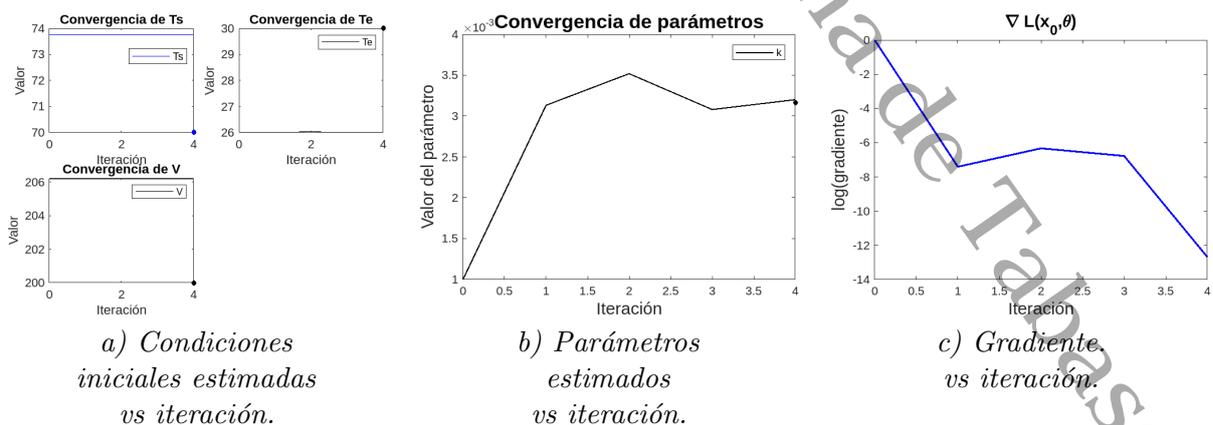


Figura 3.12: Resultados numéricos al estimar el parámetro k y las condiciones iniciales, considerando tres variables observables y tomando cinco datos.

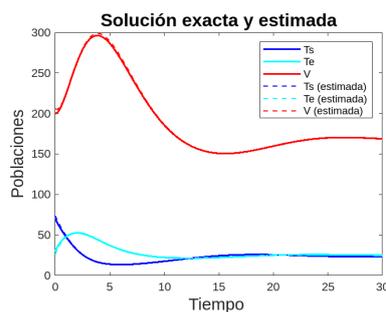


Figura 3.13: Solución exacta y estimada cuando se ajusta el parámetro k y las condiciones iniciales, considerando las tres variables observables y tomando cinco datos, el algoritmo alcanzó convergencia en 4 iteraciones.

Variables observables.	Condiciones iniciales y parámetro a estimar.	Valores exactos.	Valores iniciales en BFGS.	Iteraciones en BFGS.	Núm. de tiempos observables.	Valores estimados.	Error relativo.
T_s	T_{s0}	70	63.99	5	100000	67.63	0.03381
	T_{e0}	30	29.18			29.43	0.01890
	V_0	200	201.77			201.09	0.00546
	μ_e	0.5	0.4			0.4903	0.01934
T_s	T_{s0}	70	63.99	11	500	64.00	0.08569
	T_{e0}	30	29.18			43.45	0.44835
	V_0	200	201.77			159.14	0.20431
	μ_e	0.5	0.4			0.5121	0.02413
V	T_{s0}	70	63.99	4	10000	65.80	0.05997
	T_{e0}	30	29.18			31.84	0.06140
	V_0	200	201.77			202.37	0.01184
	μ_e	0.5	0.4			0.4899	0.02016
V	T_{s0}	70	63.99	7	500	65.31	0.06696
	T_{e0}	30	29.18			30.91	0.03028
	V_0	200	201.77			201.75	0.00874
	μ_e	0.5	0.4			0.4885	0.02294
V	T_{s0}	70	63.99	7	100	64.67	0.07609
	T_{e0}	30	29.18			31.55	0.05158
	V_0	200	201.77			201.78	0.00892
	μ_e	0.5	0.4			0.4911	0.01773

Tabla 3.4: Resumen de las simulaciones al estimar el parámetro μ_e tomando una variable observable.

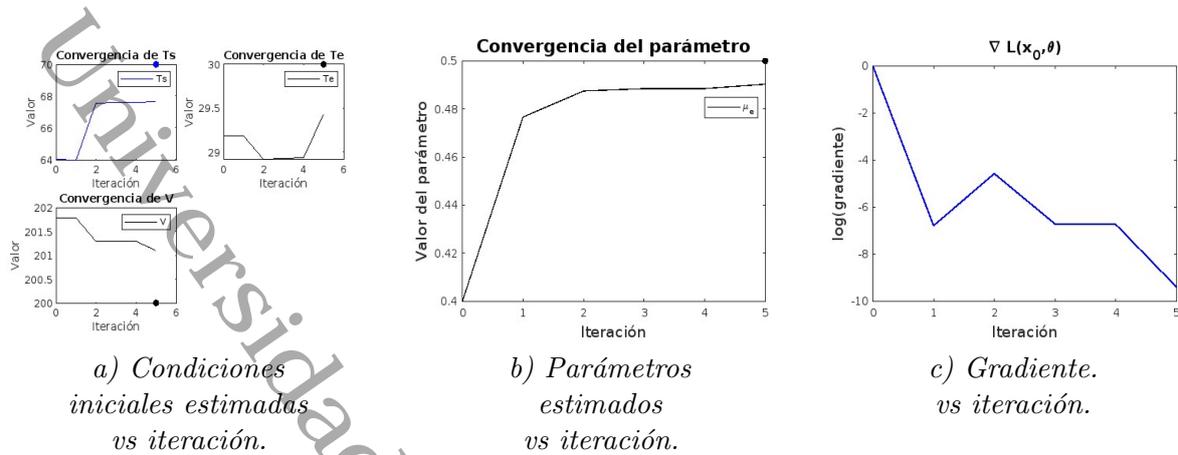


Figura 3.14: Resultados numéricos al estimar el parámetro μ_e y las condiciones iniciales, considerando solo la variable T_s como observable y tomando cien mil datos, .

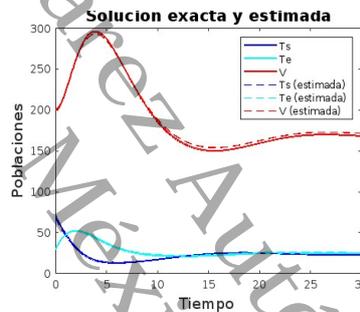


Figura 3.15: Solución exacta y estimada cuando se ajusta el parámetro μ_e y las condiciones iniciales, considerando solo la variable T_s como observable y tomando cien mil datos, el algoritmo alcanzó convergencia en 5 iteraciones.

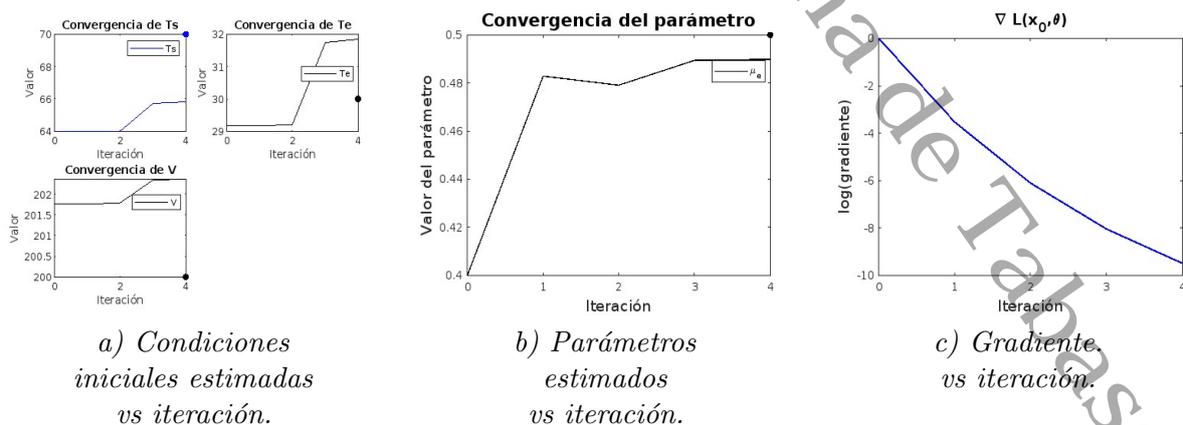


Figura 3.16: Resultados numéricos al estimar el parámetro μ_e y las condiciones iniciales, considerando solo la variable V como observable y tomando cien mil datos.

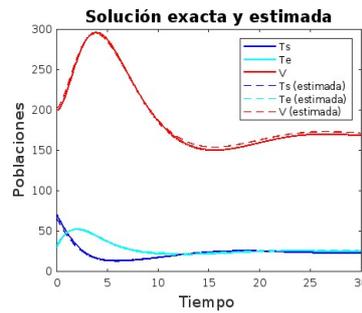


Figura 3.17: Solución exacta y estimada cuando se ajusta el parámetro μ_e y las condiciones iniciales, considerando solo la variable V como observable y tomando cien mil datos, el algoritmo alcanzó convergencia en 4 iteraciones.

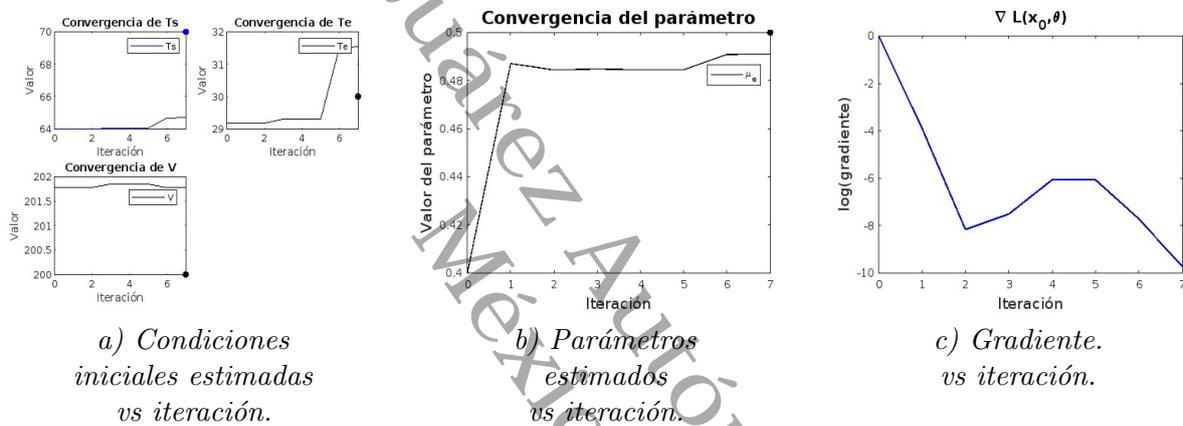


Figura 3.18: Resultados numéricos al estimar el parámetro μ_e y las condiciones iniciales, considerando solo la variable V como observable y tomando cien datos.

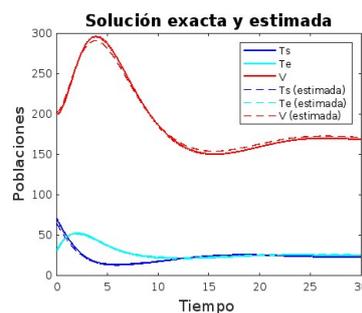


Figura 3.19: Solución exacta y estimada cuando se ajusta el parámetro μ_e y las condiciones iniciales, considerando solo la variable V como observable y tomando cien datos, el algoritmo alcanzó convergencia en 7 iteraciones.

Variables observables.	Condiciones iniciales y parámetro a estimar.	Valores exactos.	Valores iniciales en BFGS.	Iteraciones en BFGS.	Núm. de tiempos observables.	Valores estimados.	Error relativo.
T_s V	T_{s0} T_{e0} V_0 μ_e	70 30 200 0.5	70.72 32.63 194.51 0.4	5	100000	69.64 31.87 194.75 0.5026	0.005204 0.062316 0.026239 0.005211
T_s V	T_{s0} T_{e0} V_0 μ_e	70 30 200 0.5	70.72 32.63 194.51 0.4	1	100	70.72 32.63 194.51 0.5081	0.010347 0.087787 0.027461 0.016106
T_s V	T_{s0} T_{e0} V_0 μ_e	70 30 200 0.5	70.72 32.63 194.51 0.4	9	20	70.72 31.54 194.51 0.5084	0.010333 0.051242 0.027449 0.016790
T_s V	T_{s0} T_{e0} V_0 μ_e	70 30 200 0.5	70.72 32.63 194.51 0.4	10	10	70.72 27.81 194.51 0.4655	0.010346 0.072931 0.027458 0.069013

Tabla 3.5: Resumen de las simulaciones al estimar el parámetro μ_e tomando dos variables observables.

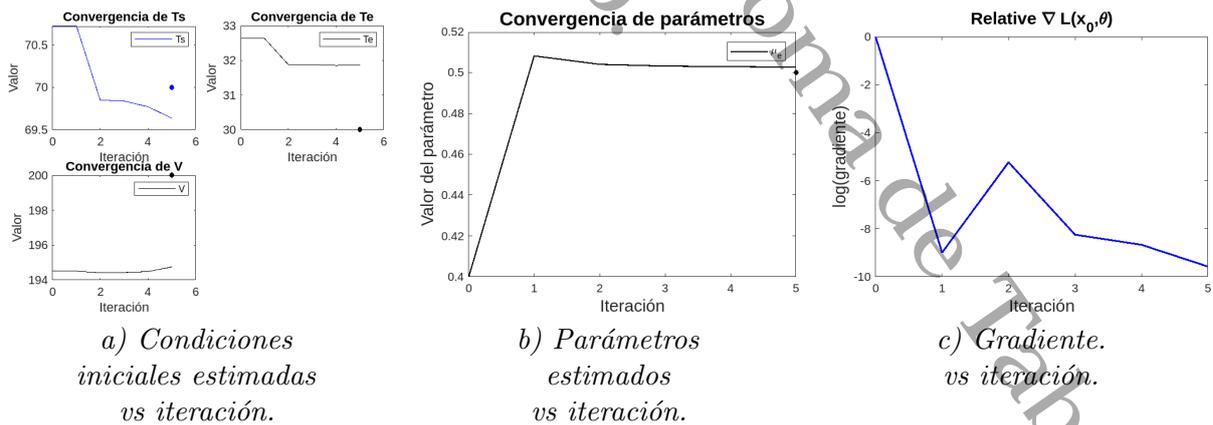


Figura 3.20: Resultados numéricos al estimar el parámetro μ_e y las condiciones iniciales, considerando las variables T_s y V como observables y tomando cien mil datos.

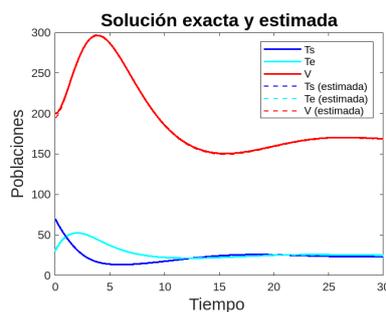


Figura 3.21: Solución exacta y estimada cuando se ajusta el parámetro μ_e y las condiciones iniciales, considerando las variables T_s y V como observables y tomando cien mil datos, el algoritmo alcanzó convergencia en 5 iteraciones.



Figura 3.22: Resultados numéricos al estimar el parámetro μ_e y las condiciones iniciales, considerando las variables T_s y V como observables y tomando veinte datos.

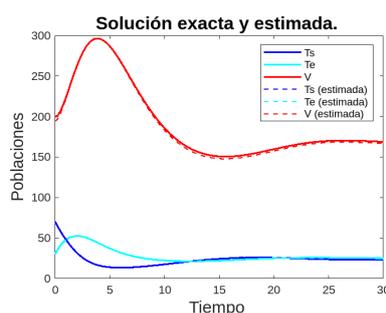


Figura 3.23: Solución exacta y estimada cuando se ajusta el parámetro μ_e y las condiciones iniciales, considerando las variables T_s y V como observables y tomando veinte datos, el algoritmo alcanzó convergencia en 9 iteraciones.

Variab-les observa-bles.	Condicio-nes iniciales y paráme-tro a estimar.	Valores exactos.	Valores inicia-les en BFGS.	Iteracio-nes en BFGS.	Núm. de tiempos observa-bles.	Valores estima-dos.	Error relativo.
T_s T_e V	T_{s0} T_{e0} V_0 μ_e	70 30 200 0.5	74.24 28.97 207.05 0.4	6	100000	71.45 28.57 206.01 0.5014	0.020783 0.047649 0.030063 0.002776
T_s T_e V	T_{s0} T_{e0} V_0 μ_e	70 30 200 0.5	74.24 28.97 207.05 0.4	6	100	74.22 28.96 207.04 0.5124	0.060344 0.034544 0.035219 0.024777
T_s T_e V	T_{s0} T_{e0} V_0 μ_e	70 30 200 0.5	74.24 28.97 207.05 0.4	4	50	74.24 28.97 207.05 0.5041	0.060530 0.034417 0.035244 0.008120
T_s T_e V	T_{s0} T_{e0} V_0 μ_e	70 30 200 0.5	74.24 28.97 207.05 0.4	6	10	74.24 28.97 207.05 0.5238	0.060578 0.034422 0.035233 0.047660
T_s T_e V	T_{s0} T_{e0} V_0 μ_e	70 30 200 0.5	74.24 28.97 207.05 0.4	2	2	74.24 28.97 207.05 0.5216	0.060585 0.034415 0.035246 0.043275

Tabla 3.6: Resumen de las simulaciones al estimar el parámetro μ_e tomando tres variables observables.

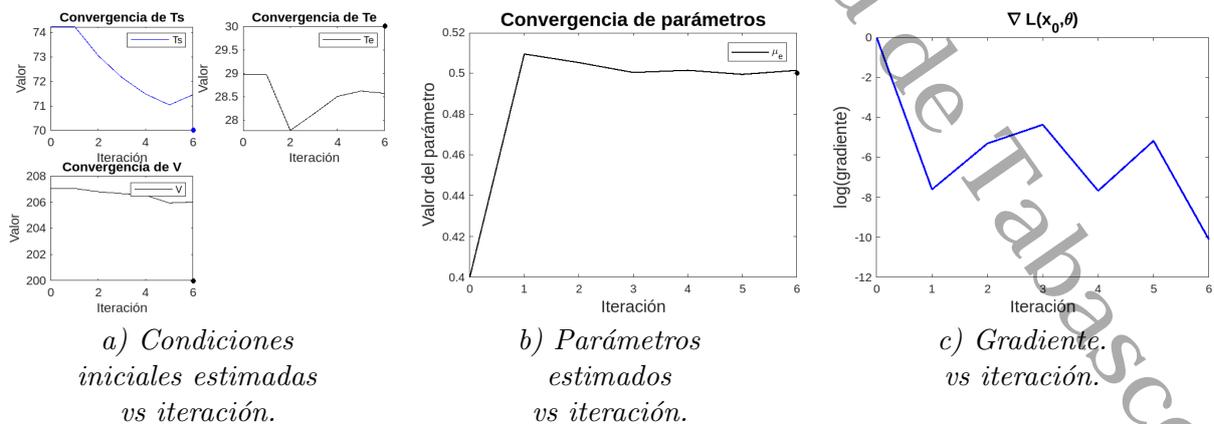


Figura 3.24: Resultados numéricos al estimar el parámetro μ_e y las condiciones iniciales, considerando las variables T_s , T_e y V como observables y tomando cien mil datos.

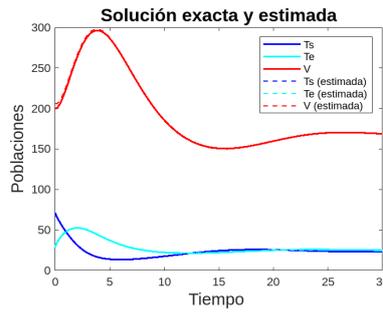


Figura 3.25: Solución exacta y estimada cuando se ajusta el parámetro μ_e y las condiciones iniciales, considerando las variables T_s , T_e y V como observables y tomando cien mil datos, el algoritmo alcanzó convergencia en 6 iteraciones.

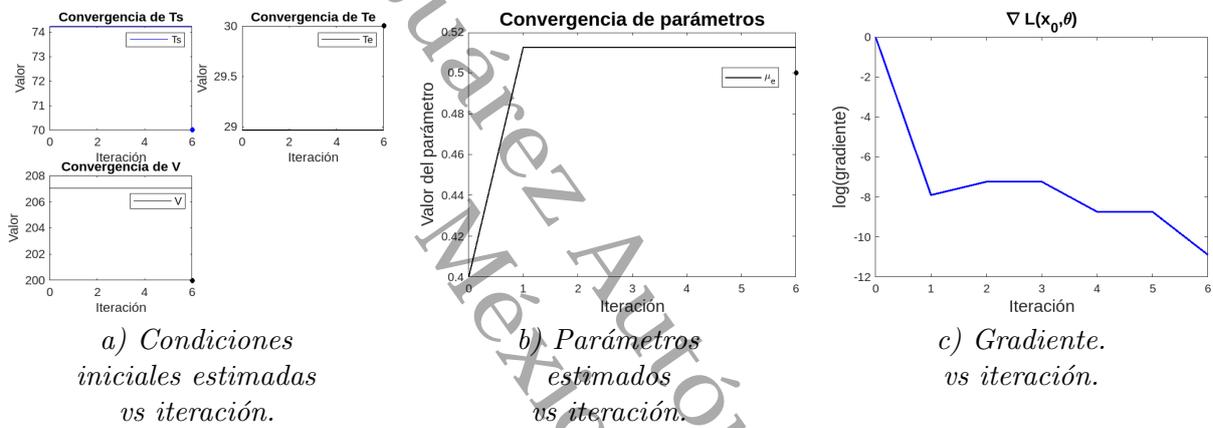


Figura 3.26: Resultados numéricos al estimar el parámetro μ_e y las condiciones iniciales, considerando las variables T_s , T_e y V como observables y tomando diez datos.

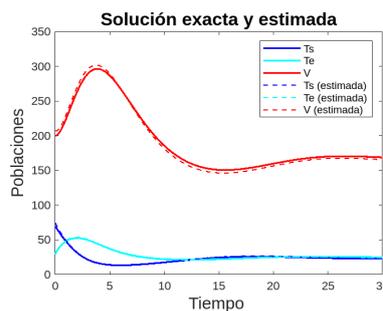


Figura 3.27: Solución exacta y estimada cuando se ajusta el parámetro μ_e y las condiciones iniciales, considerando las variables T_s , T_e y V como observables y tomando diez datos, el algoritmo alcanzó convergencia en 6 iteraciones.

3.4.2. Estimación de los parámetros k y μ_e simultáneamente

Variables observables.	Condiciones iniciales y parámetros a estimar.	Valores exactos.	Valores iniciales en BFGS.	Iteraciones en BFGS.	Núm. de tiempos observables.	Valores estimados.	Error relativo.
T_s	T_{s0}	70	73.76	4	10000	72.97	0.042464
	T_{e0}	30	26.02			29.58	0.013992
	V_0	200	206.23			207.75	0.038743
	k	3.1662×10^{-3}	1×10^{-3}			3.1206×10^{-3}	0.014387
	μ_e	0.5	0.4			0.5036	0.007300
T_s	T_{s0}	70	73.76	6	100	73.75	0.053602
	T_{e0}	30	26.02			26.07	0.131106
	V_0	200	206.23			206.25	0.031236
	k	3.1662×10^{-3}	1×10^{-3}			3.3140×10^{-3}	0.046674
	μ_e	0.5	0.4			0.5262	0.052301
V	T_{s0}	70	58.90	15	100000	79.32	0.133182
	T_{e0}	30	29.39			24.75	0.174972
	V_0	200	212.17			209.76	0.048801
	k	3.1662×10^{-3}	1.662×10^{-4}			2.9152×10^{-3}	0.079274
	μ_e	0.5	0.4			0.5245	0.048973
V	T_{s0}	70	58.90	9	1000	59.08	0.156013
	T_{e0}	30	29.39			29.87	0.004414
	V_0	200	212.17			212.26	0.061319
	k	3.1662×10^{-3}	1.662×10^{-4}			3.5484×10^{-3}	0.120705
	μ_e	0.5	0.4			0.4597	0.080603
V	T_{s0}	70	58.90	8	300	58.95	0.157824
	T_{e0}	30	29.39			29.53	0.015760
	V_0	200	212.17			212.20	0.060984
	k	3.1662×10^{-3}	1.662×10^{-4}			3.5855×10^{-3}	0.132424
	μ_e	0.5	0.4			0.4540	0.092140
V	T_{s0}	70	58.90	9	280	58.95	0.157854
	T_{e0}	30	29.39			29.52	0.016135
	V_0	200	212.17			212.19	0.060974
	k	3.1662×10^{-3}	1.662×10^{-4}			3.6380×10^{-3}	0.148900
	μ_e	0.5	0.4			0.4508	0.098341

Tabla 3.7: Resumen de las simulaciones al estimar los parámetros k y μ_e tomando una variable observable.

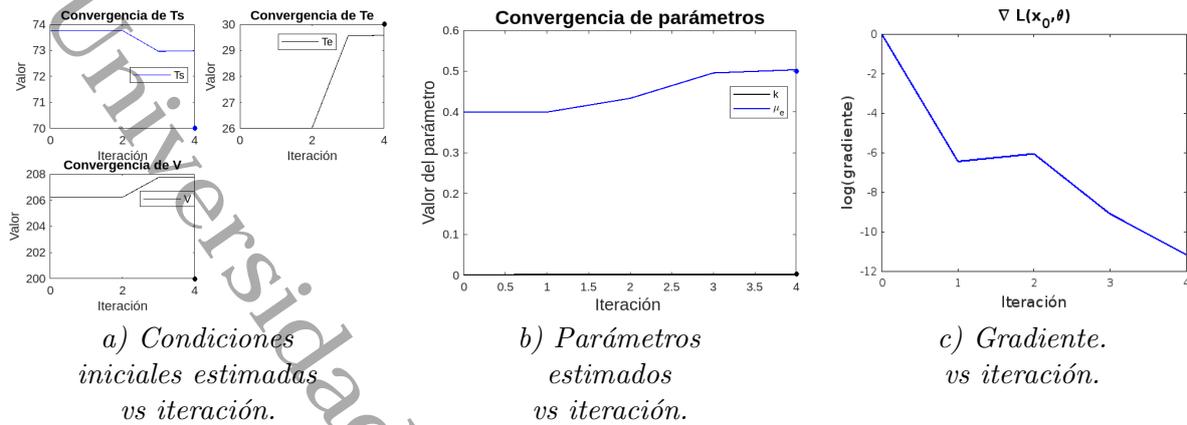


Figura 3.28: Resultados numéricos al estimar los parámetros k y μ_e y las condiciones iniciales, considerando la variable T_s como observable y tomando diez mil datos.

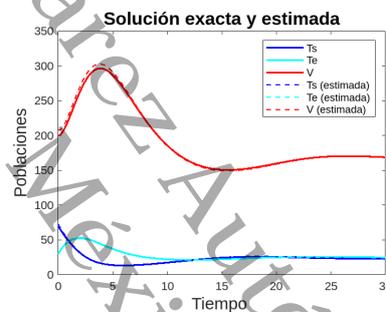


Figura 3.29: Solución exacta y estimada cuando se ajustan los parámetros k y μ_e y las condiciones iniciales, considerando la variable T_s como observable y tomando diez mil datos, el algoritmo alcanzó convergencia en 4 iteraciones.

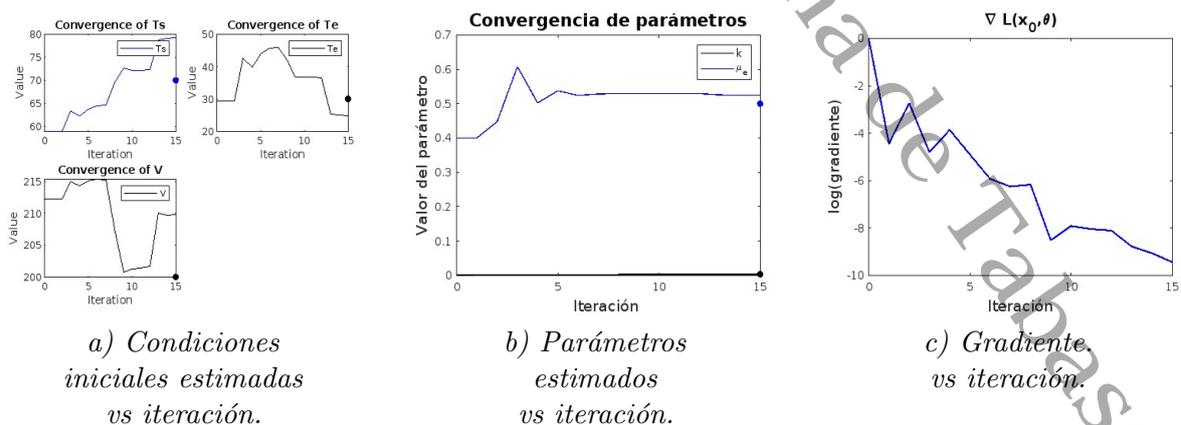


Figura 3.30: Resultados numéricos al estimar los parámetros k y μ_e y las condiciones iniciales, considerando la variable V como observable y tomando cien mil datos.

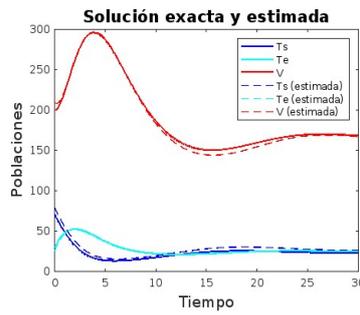


Figura 3.31: Solución exacta y estimada cuando se ajustan los parámetros k y μ_e y las condiciones iniciales, considerando la variable V como observable y tomando cien mil datos, el algoritmo alcanzó convergencia en 15 iteraciones.

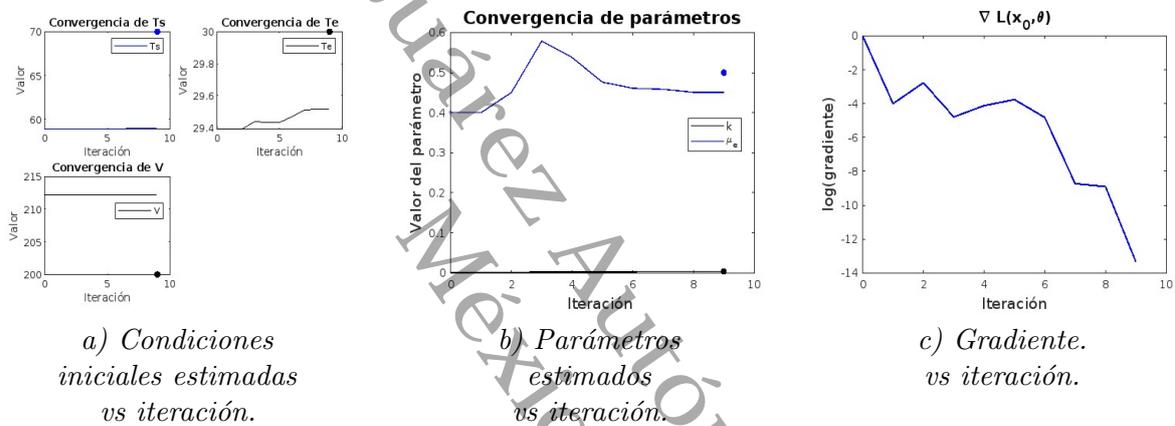


Figura 3.32: Resultados numéricos al estimar los parámetros k y μ_e y las condiciones iniciales, considerando la variable V como observable y tomando 280 datos.

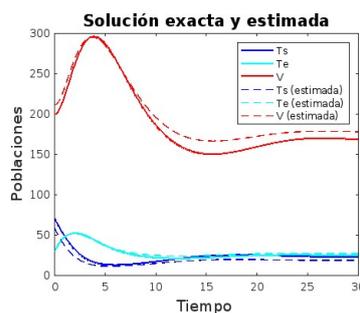


Figura 3.33: Solución exacta y estimada cuando se ajustan los parámetros k y μ_e y las condiciones iniciales, considerando la variable V como observable y tomando 280 datos, el algoritmo alcanzó convergencia en 9 iteraciones.

Variables observables.	Condiciones iniciales y parámetros a estimar.	Valores exactos.	Valores iniciales en BFGS.	Iteraciones en BFGS.	Núm. de tiempos observables.	Valores estimados.	Error relativo.
T_s V	T_{s0} T_{e0} V_0 k μ_e	70 30 200 3.1662×10^{-3} 0.5	70.72 32.63 194.50 1×10^{-3} 0.4	8	100000	69.98 32.74 196.08 3.1416×10^{-3} 0.5031	0.000352 0.091298 0.019605 0.007772 0.006294
T_s V	T_{s0} T_{e0} V_0 k μ_e	70 30 200 3.1662×10^{-3} 0.5	70.72 32.63 194.50 1×10^{-3} 0.4	4	100	70.73 32.66 194.52 3.1415×10^{-3} 0.5049	0.010457 0.088549 0.027412 0.007798 0.009712
T_s V	T_{s0} T_{e0} V_0 k μ_e	70 30 200 3.1662×10^{-3} 0.5	70.72 32.63 194.50 1×10^{-3} 0.4	4	20	70.73 32.66 194.52 3.1559×10^{-3} 0.5094	0.010348 0.087920 0.027451 0.003263 0.018768
T_s V	T_{s0} T_{e0} V_0 k μ_e	70 30 200 3.1662×10^{-3} 0.5	70.72 32.63 194.50 1×10^{-3} 0.4	4	15	70.73 32.64 194.51 3.1610×10^{-3} 0.5178	0.070096 0.263818 0.049664 0.050145 0.013339

Tabla 3.8: Resumen de las simulaciones al estimar los parámetros k y μ_e tomando dos variables observables.

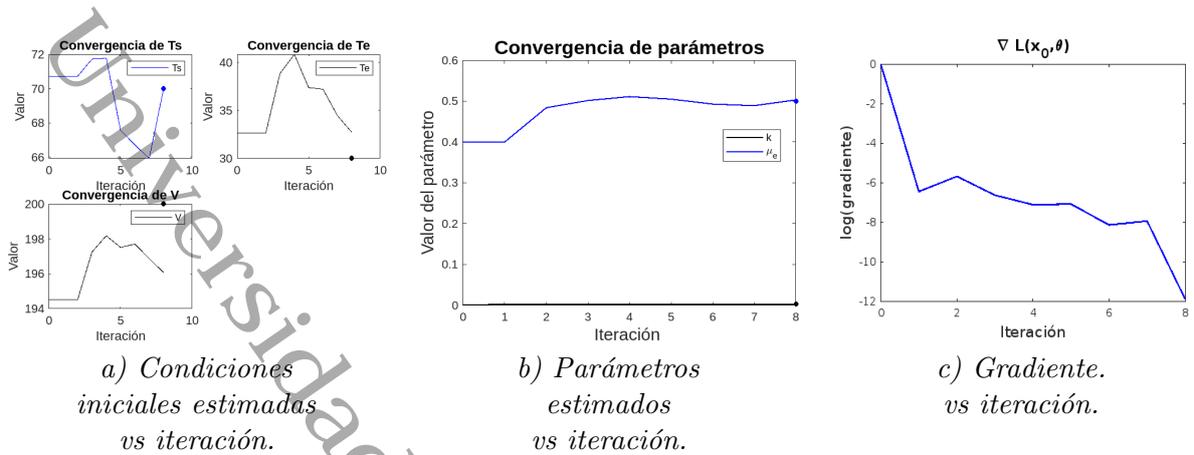


Figura 3.34: Resultados numéricos al estimar los parámetros k y μ_e y las condiciones iniciales, considerando las variables T_s y V como observables y tomando cien mil datos.

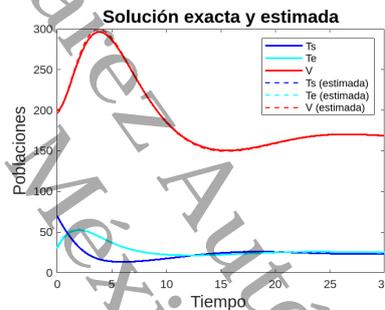


Figura 3.35: Solución exacta y estimada cuando se ajustan los parámetros k y μ_e y las condiciones iniciales, considerando las variables T_s y V como observables y tomando cien mil datos, el algoritmo alcanzó convergencia en 8 iteraciones.

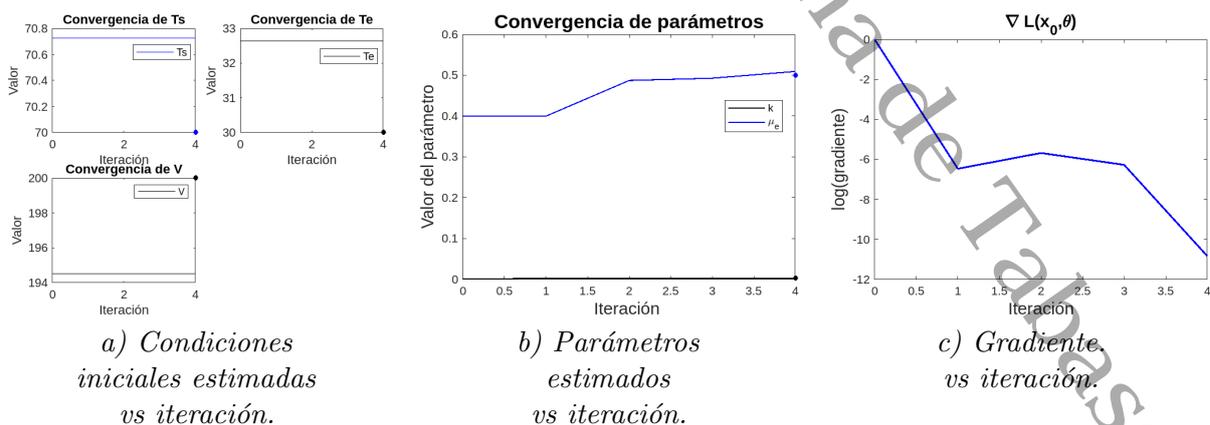


Figura 3.36: Resultados numéricos al estimar los parámetros k y μ_e y las condiciones iniciales, considerando las variables T_s y V como observables y tomando veinte datos.

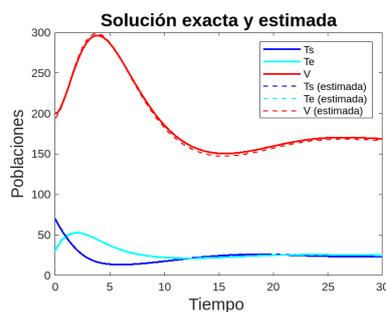


Figura 3.37: Solución exacta y estimada cuando se ajustan los parámetros k y μ_e y las condiciones iniciales, considerando las variables T_s y V como observables y tomando veinte datos, el algoritmo alcanzó convergencia en 4 iteraciones.

Variables observables.	Condiciones iniciales y parámetros a estimar.	Valores exactos.	Valores iniciales en BFGS.	Iteraciones en BFGS.	Núm. de tiempos observables.	Valores estimados.	Error relativo.
T_s T_e V	T_{s0} T_{e0} V_0 k μ_e	70 30 200 3.1662×10^{-3} 0.5	71.95 24.80 180.32 1×10^{-3} 0.4	6	100000	70.80 27.26 194.99 3.2554×10^{-3} 0.5017	0.011394 0.091344 0.025061 0.028188 0.003389
T_s T_e V	T_{s0} T_{e0} V_0 k μ_e	70 30 200 3.1662×10^{-3} 0.5	71.95 24.80 180.32 1×10^{-3} 0.4	4	100	71.97 24.86 180.34 3.3883×10^{-3} 0.5097	0.028209 0.171425 0.098307 0.070162 0.019304
T_s T_e V	T_{s0} T_{e0} V_0 k μ_e	70 30 200 3.1662×10^{-3} 0.5	71.95 24.80 180.32 1×10^{-3} 0.4	4	50	71.96 24.82 180.33 3.3005×10^{-3} 0.5021	0.027950 0.172599 0.098362 0.042408 0.004210
T_s T_e V	T_{s0} T_{e0} V_0 k μ_e	70 30 200 3.1662×10^{-3} 0.5	71.95 24.80 180.32 1×10^{-3} 0.4	4	20	71.96 24.81 180.32 3.450×10^{-3} 0.5052	0.027944 0.173025 0.098391 0.089543 0.010429

Tabla 3.9: Resumen de las simulaciones al estimar los parámetros k y μ_e tomando tres variables observables.

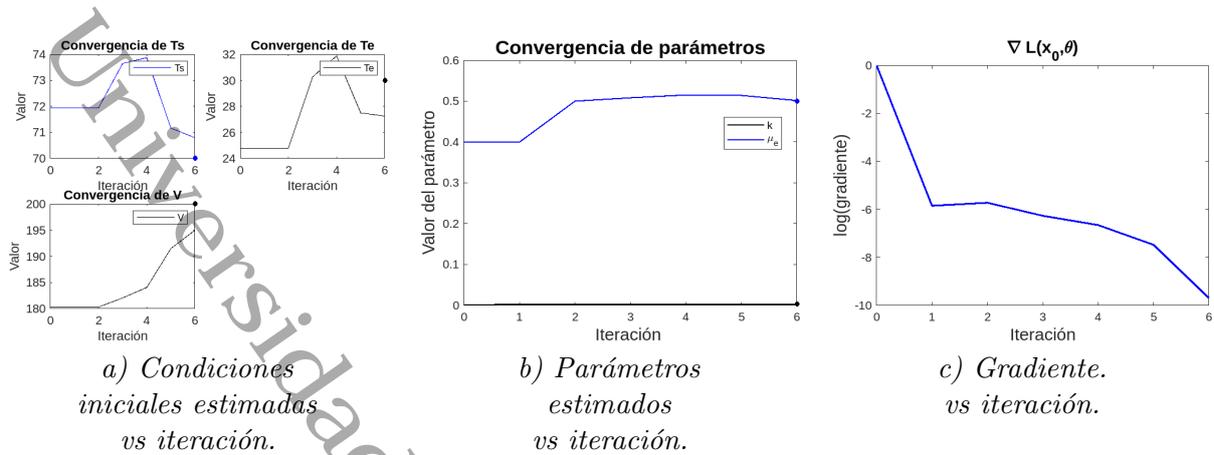


Figura 3.38: Resultados numéricos al estimar los parámetros k y μ_e y las condiciones iniciales, considerando las variables T_s , T_e y V como observables y tomando cien mil datos.

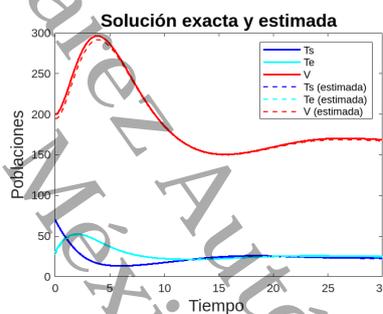


Figura 3.39: Solución exacta y estimada cuando se ajustan los parámetros k y μ_e y las condiciones iniciales, considerando las variables T_s , T_e y V como observables y tomando cien mil datos, el algoritmo alcanzó convergencia en 6 iteraciones.

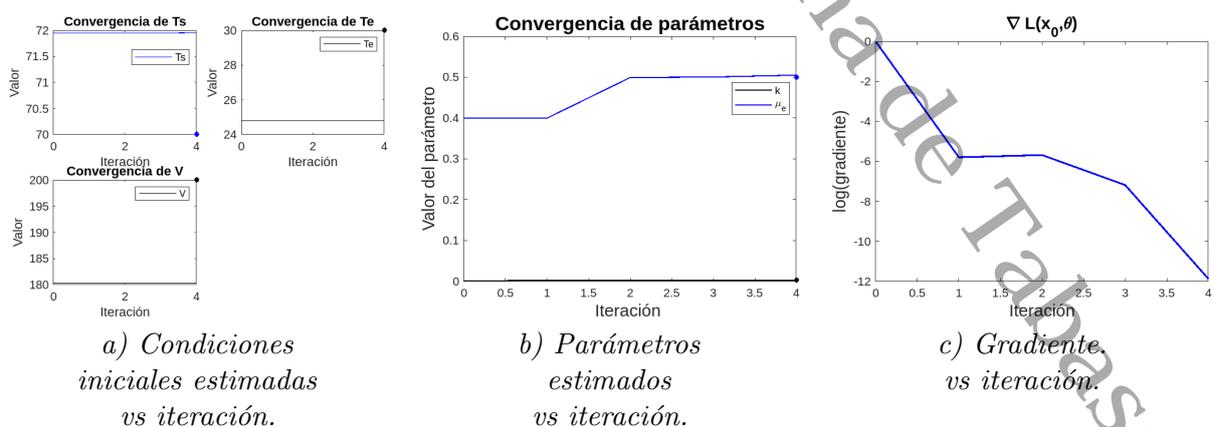


Figura 3.40: Resultados numéricos al estimar los parámetros k y μ_e y las condiciones iniciales, considerando las variables T_s , T_e y V como observables y tomando veinte datos.

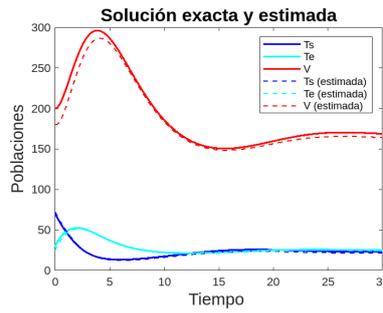


Figura 3.41: Solución exacta y estimada cuando se ajustan los parámetros k y μ_e y las condiciones iniciales, considerando las variables T_s , T_e y V como observables y tomando veinte datos, el algoritmo alcanzó convergencia en 4 iteraciones.

3.4.3. Estimación de los parámetros β , k , μ_V .

Variab-les ob-serva-bles.	Condi-ciones ini-ciales y paráme-tros a esti-mar.	Valores exactos.	Valores iniciales en BFGS.	Itera-ciones en BFGS.	Núm. de tiempos obser-vables.	Valores estimados.	Error relativo.
T_s	T_{s0}	70	66.85	11	100000	67.71	0.032650
	T_{e0}	30	31.67			39.02	0.300529
	V_0	200	178.57			180.31	0.098411
	β	0.12	0.1			0.1011	0.157555
	k	3.1662×10^{-3}	1.3×10^{-4}			3.0533×10^{-3}	0.035656
	μ_V	0.4	0.3			0.3955	0.011113
V	T_{s0}	70	66.26	20	100000	68.35	0.023517
	T_{e0}	30	27.05			29.63	0.012198
	V_0	200	193.41			196.50	0.017524
	β	0.12	0.09			0.9342	0.221476
	k	3.1662×10^{-3}	1×10^{-3}			3.2878×10^{-3}	0.038411
	μ_V	0.4	0.33			0.3832	0.041815
V	T_{s0}	70	66.26	12	300	66.27	0.053350
	T_{e0}	30	27.05			27.05	0.098209
	V_0	200	193.41			193.41	0.032928
	β	0.12	0.09			0.1298	0.082321
	k	3.1662×10^{-3}	1×10^{-3}			3.8354×10^{-3}	0.211361
	μ_V	0.4	0.33			0.3832	0.042113

Tabla 3.10: Resumen de las simulaciones al estimar los parámetros β , k y μ_V tomando una variable observable.

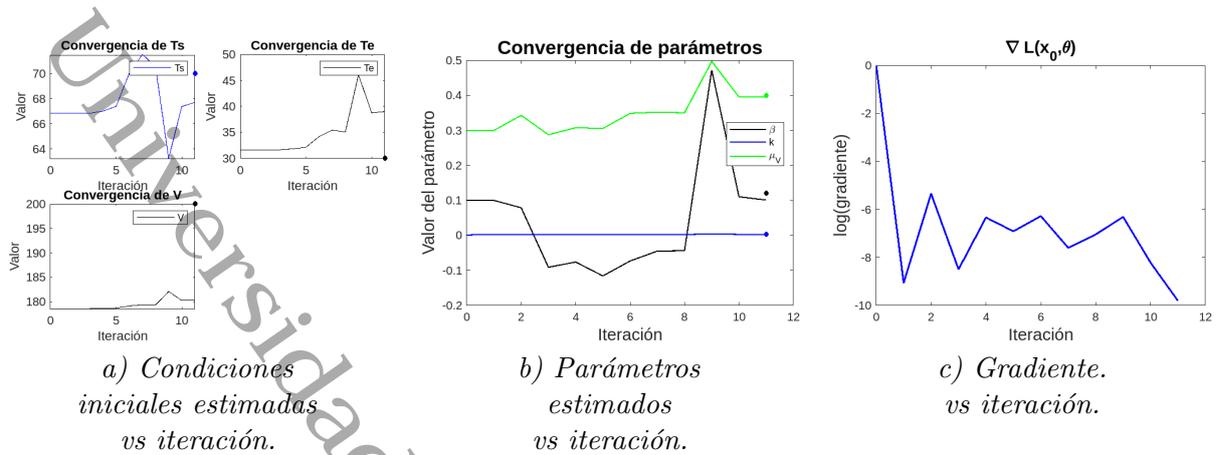


Figura 3.42: Resultados numéricos al estimar los parámetros β , k y μ_V y las condiciones iniciales, considerando la variable T_s como observable y tomando cien mil datos.

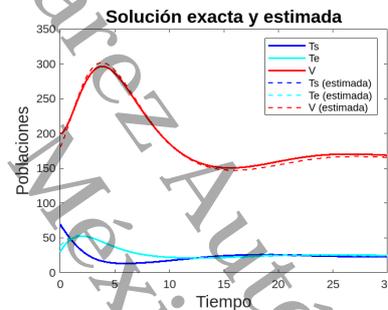


Figura 3.43: Solución exacta y estimada cuando se ajustan los parámetros β , k y μ_V y las condiciones iniciales, considerando la variable T_s como observable y tomando cien mil datos, el algoritmo alcanzó convergencia en 11 iteraciones.

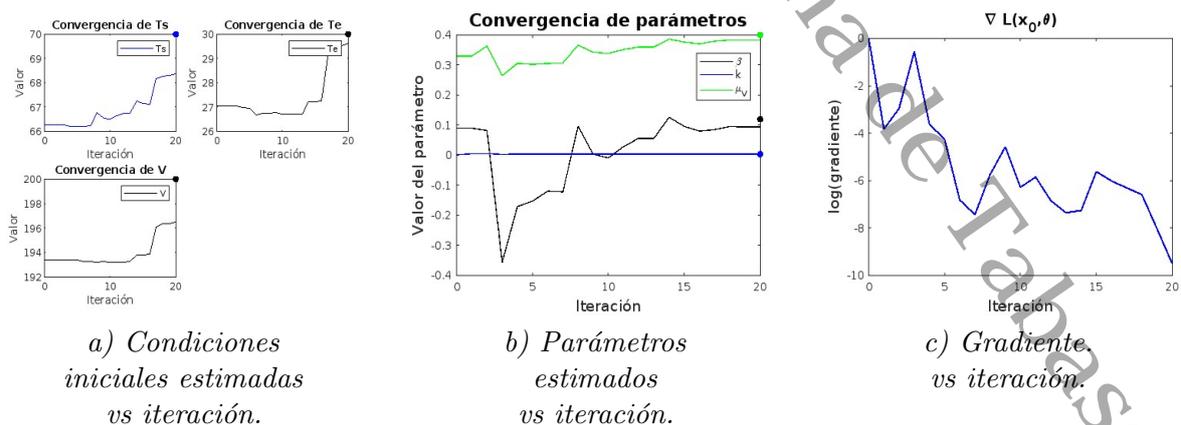


Figura 3.44: Resultados numéricos al estimar los parámetros β , k y μ_V y las condiciones iniciales, considerando la variable V como observable y tomando cien mil datos.

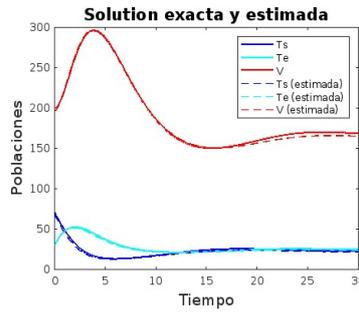
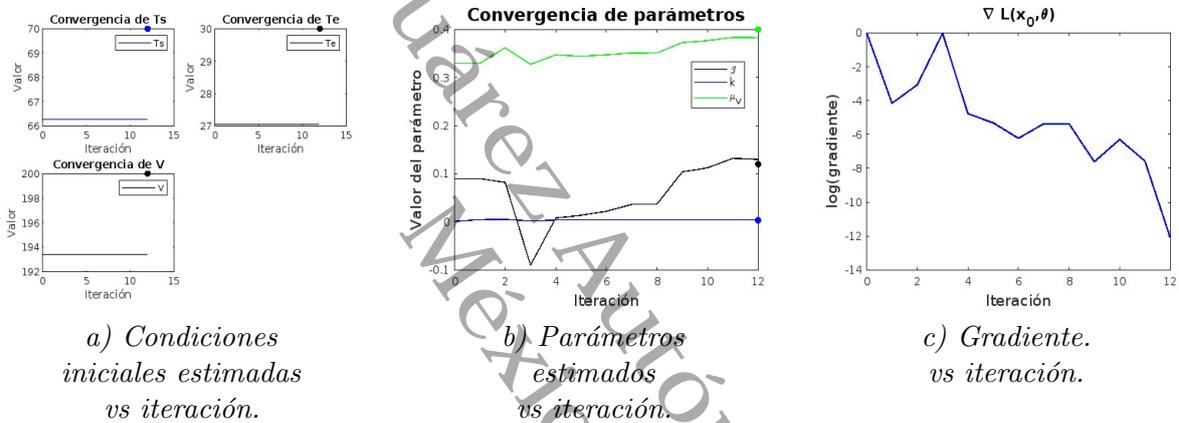


Figura 3.45: Solución exacta y estimada cuando se ajustan los parámetros β , k y μ_V y las condiciones iniciales, considerando la variable V como observable y tomando cien mil datos, el algoritmo alcanzó convergencia en 20 iteraciones.



a) Condiciones iniciales estimadas vs iteración.

b) Parámetros estimados vs iteración.

c) Gradiente vs iteración.

Figura 3.46: Resultados numéricos al estimar los parámetros β , k y μ_V y las condiciones iniciales, considerando la variable V como observable y tomando trescientos datos.

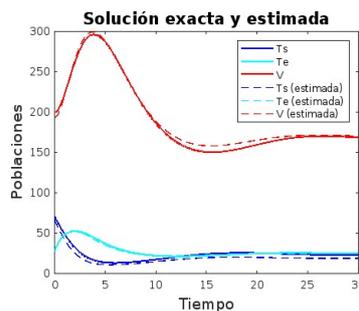


Figura 3.47: Solución exacta y estimada cuando se ajustan los parámetros β , k y μ_V y las condiciones iniciales, considerando la variable V como observable y tomando trescientos datos, el algoritmo alcanzó convergencia en 12 iteraciones.

Variables observables.	Condiciones iniciales y parámetros a estimar.	Valores exactos.	Valores iniciales en BFGS.	Iteraciones en BFGS.	Núm. de tiempos observables.	Valores estimados.	Error relativo.
T_s V	T_{s0}	70	70.72	14	100000	70.22	0.003209
	T_{e0}	30	32.63			31.69	0.056322
	V_0	200	194.51			195.95	0.020261
	β	0.12	0.1			0.0980	0.183058
	k	3.0833×10^{-3}	2.9×10^{-3}			3.939×10^{-3}	0.026173
	μ_V	0.4	0.35			0.3877	0.015204
T_s V	T_{s0}	70	70.72	9	500	70.72	0.010267
	T_{e0}	30	32.63			32.63	0.087735
	V_0	200	194.51			194.51	0.027453
	β	0.12	0.1			0.0739	0.384048
	k	3.1662×10^{-3}	2.9×10^{-3}			2.9978×10^{-3}	0.053181
	μ_V	0.4	0.35			0.3877	0.030857
T_s V	T_{s0}	70	70.72	9	100	70.72	0.010329
	T_{e0}	30	32.63			32.64	0.087936
	V_0	200	194.51			194.51	0.027440
	β	0.12	0.1			0.1395	0.162197
	k	3.1662×10^{-3}	2.9×10^{-3}			3.2162×10^{-3}	0.015767
	μ_V	0.4	0.35			0.4119	0.029708

Tabla 3.11: Resumen de las simulaciones al estimar los parámetros β , k y μ_V tomando dos variables observables.

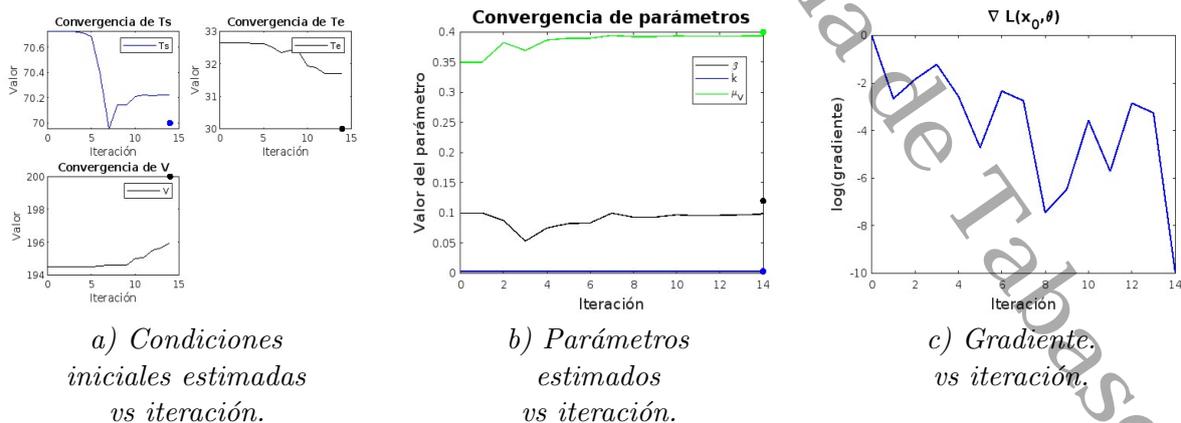


Figura 3.48: Resultados numéricos al estimar los parámetros β , k y μ_V y las condiciones iniciales considerando la variables T_s y V como observables y tomando cien mil datos.

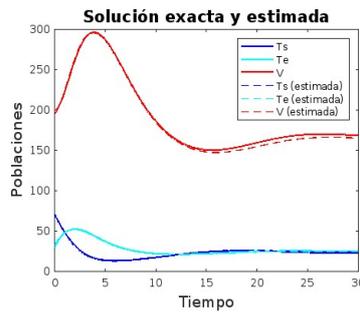


Figura 3.49: Solución exacta y estimada cuando se ajustan los parámetros β , k y μ_V y las condiciones iniciales, considerando las variables T_s y V como observables y tomando cien mil datos, el algoritmo alcanzó convergencia en 14 iteraciones.

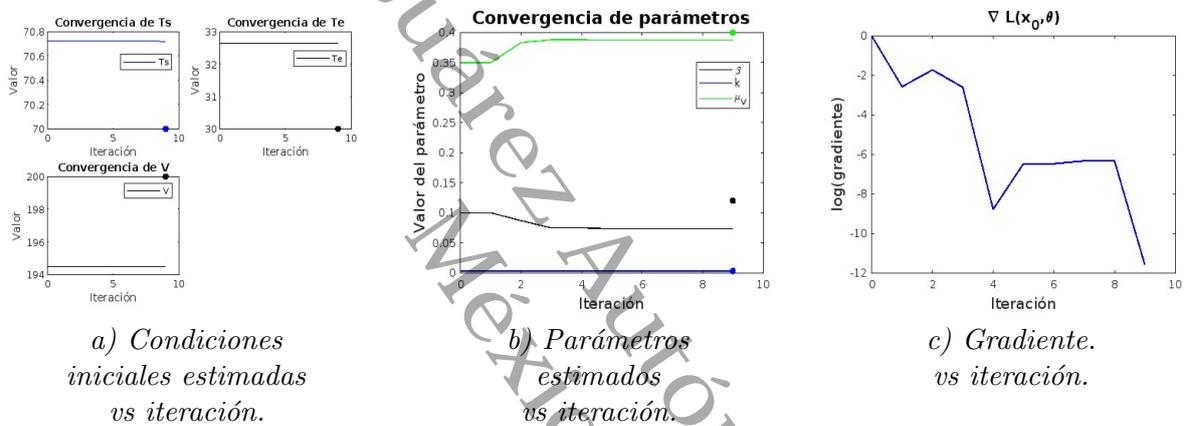


Figura 3.50: Resultados numéricos al estimar los parámetros β , k y μ_V y las condiciones iniciales, considerando la variables T_s y V como observables y tomando quinientos datos.

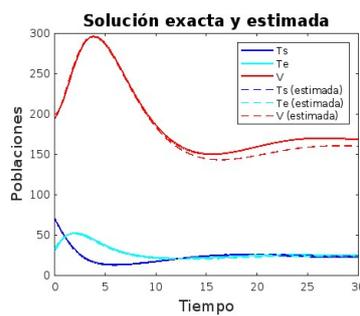


Figura 3.51: Solución exacta y estimada cuando se ajustan los parámetros β , k y μ_V y las condiciones iniciales, considerando las variables T_s y V como observables y tomando quinientos datos, el algoritmo alcanzó convergencia en 9 iteraciones.

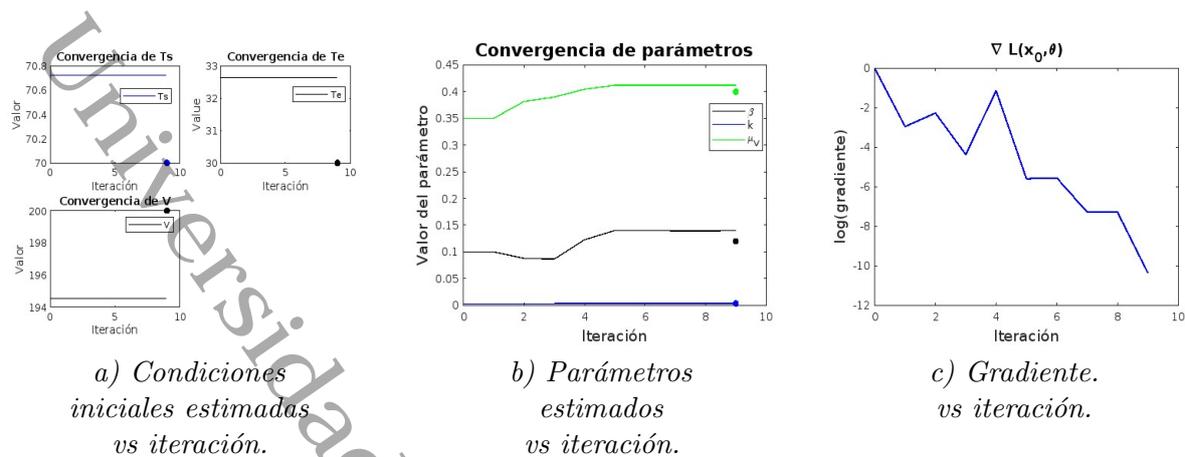


Figura 3.52: Resultados numéricos al estimar los parámetros β , k y μ_V y las condiciones iniciales, considerando las variables T_s y V como observables y tomando cien datos.

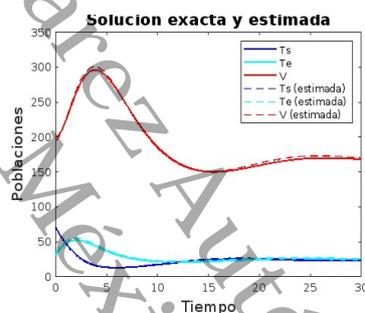


Figura 3.53: Solución exacta y estimada cuando se ajustan los parámetros β , k y μ_V y las condiciones iniciales, considerando las variables T_s y V como observables y tomando cien datos, el algoritmo alcanzó convergencia en 9 iteraciones.

Vari- bles ob- serva- bles.	Condi- ciones ini- ciales y paráme- tros a esti- mar.	Valores exactos.	Valores iniciales en BFGS.	Itera- cio- nes en BFGS.	Núm. de tiem- pos obser- vables.	Valores estimados.	Error relativo.
T_s T_e V	T_{s0} T_{e0} V_0 β k μ_V	70 30 200 0.12 3.1662×10^{-3} 0.4	63.99 29.18 201.78 0.1 1×10^{-3} 0.3	11	100000	68.54 29.21 207.21 0.1306 3.1726×10^{-3} 0.3986	0.020818 0.026352 0.036066 0.088095 0.002032 0.003492
T_s T_e V	T_{s0} T_{e0} V_0 β k μ_V	70 30 200 0.12 3.1662×10^{-3} 0.4	63.99 29.18 201.78 0.1 1×10^{-3} 0.3	5	100	64.02 29.22 201.79 0.1532 3.3094×10^{-3} 0.4030	0.085413 0.026026 0.008934 0.277299 0.045234 0.007504

Tabla 3.12: Resumen de las simulaciones al estimar los parámetros β , k y μ_V tomando tres variables observables.

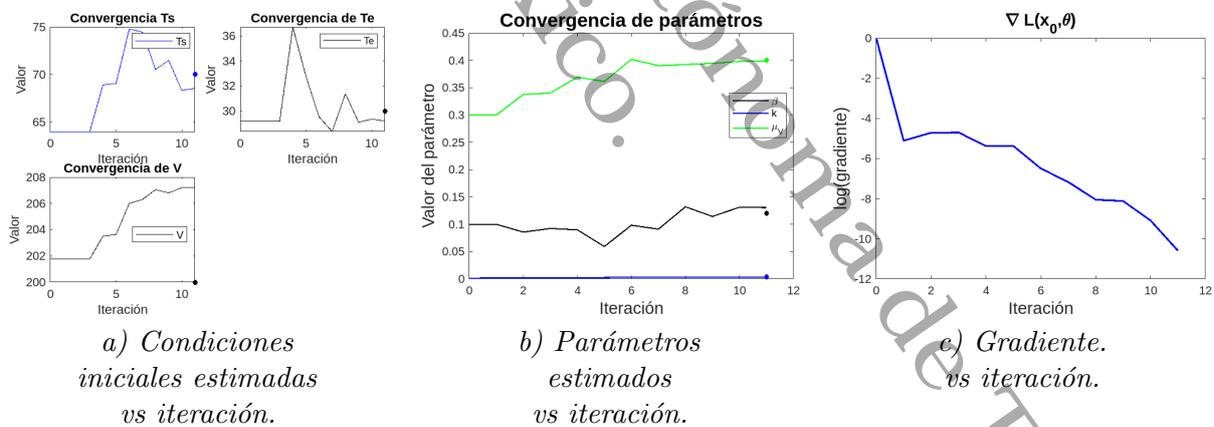


Figura 3.54: Resultados numéricos al estimar los parámetros β , k y μ_V y las condiciones iniciales, considerando las variables T_s , T_e y V como observables y tomando cien mil datos.

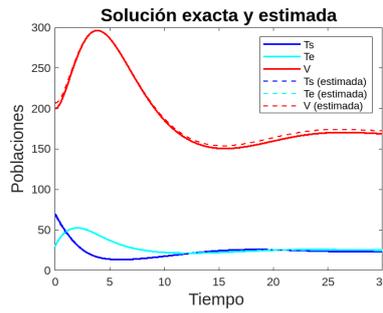


Figura 3.55: Solución exacta y estimada cuando se ajustan los parámetros β , k y μ_V y las condiciones iniciales, considerando las variables T_s , T_e y V como observables y tomando cien mil datos, el algoritmo alcanzó convergencia en 11 iteraciones.

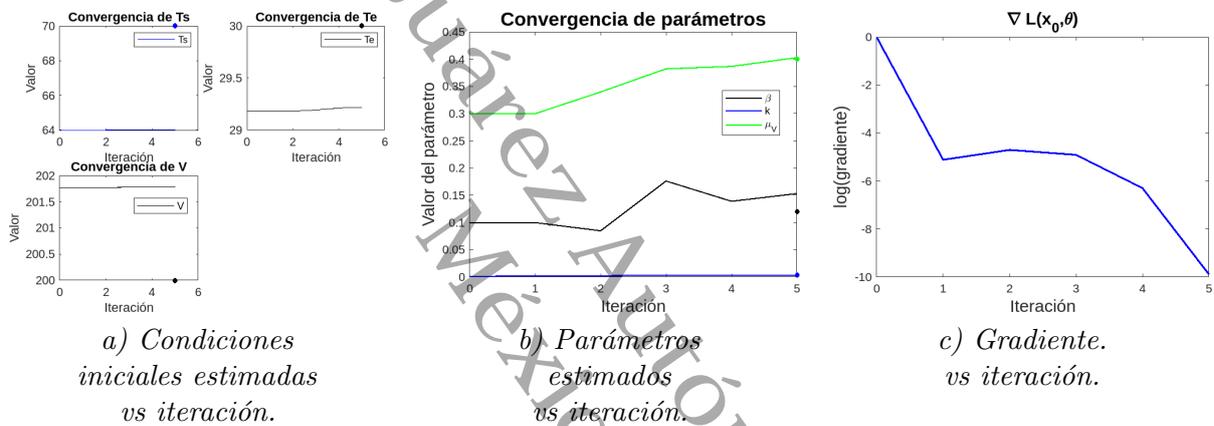


Figura 3.56: Resultados numéricos al estimar los parámetros β , k y μ_V y las condiciones iniciales, considerando las variables T_s , T_e y V como observables y tomando cien datos.

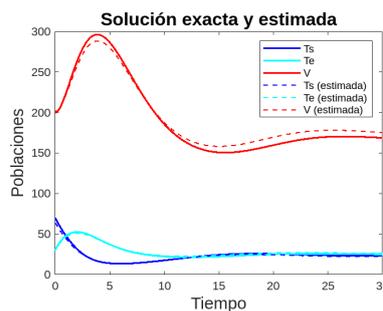


Figura 3.57: Solución exacta y estimada cuando se ajustan los parámetros β , k y μ_V y las condiciones iniciales, considerando las variables T_s , T_e y V como observables y tomando cien datos, el algoritmo alcanzó convergencia en 5 iteraciones.

Conclusiones

En este trabajo se estudió el problema de estimar parámetros de un modelo del VIH/SIDA, el cual considera la interacción a través del tiempo, entre la población de células linfáticas sanas, la población de células linfáticas infectadas y la carga viral del VIH-1.

La interacción entre estas variables involucra ocho parámetros, de los cuales en este trabajo se estimaron dos: la tasa de infección de células linfáticas sanas por el VIH-1, k , y la tasa per cápita de mortalidad de las células sanas, μ_e , ya que son los que se consideran más difíciles de obtener.

Para resolver el problema de estimación de parámetros, se considera un problema de minimización de una función \mathcal{L} que en esencia es una función de mínimos cuadrados. Esta función \mathcal{L} va de \mathbb{R}^{d+np} a \mathbb{R} , donde d es la dimensión del modelo y np el número de parámetros a estimar y para minimizarla se siguen los pasos típicos de minimización en \mathbb{R}^n , que son: encontrar los puntos críticos, es decir encontrar los puntos que hacen que el gradiente sea cero, que sería suficiente si la función \mathcal{L} es convexa.

En este trabajo no se tocó el tema de verificar que la función \mathcal{L} es convexa, pero normalmente lo es o casi es convexa. Usualmente, para una función de \mathbb{R}^n a \mathbb{R} es fácil calcular el gradiente y también es relativamente fácil encontrar o aproximar los puntos en los que el gradiente se hace cero; sin embargo en el caso \mathcal{L} , su valor depende de la solución de un sistema diferencial, lo que hace difícil calcular el gradiente. Dependiendo de cómo se calcula este gradiente, es la metodología que resulta. En algunos casos, para calcular el gradiente se tienen que calcular lo que se conoce como las matrices de sensibilidades, lo que resulta en un trabajo computacional muy costoso porque no solo depende del número de parámetros a estimar, también depende del número de datos observables.

En este trabajo, se hace el cálculo del gradiente considerando un enfoque variacional o de perturbación, que permite calcular el gradiente de manera eficiente y el precio que se tiene que pagar es resolver no solamente el sistema de estado sino un sistema adjunto asociado, todo esto permitió obtener una fórmula eficiente para el gradiente. Una vez que se tiene el gradiente se pueden aplicar una variedad de métodos para minimizar la función \mathcal{L} , por ejemplo: descenso por el gradiente, gradiente conjugado, el método de Newton o un método cuasi Newton como BFGS. En este trabajo se decidió trabajar con BFGS con el fin de conocer y explorar sus alcances y porque aunque en los problemas de baja escala el trabajo involucrado en este método puede ser demasiado, comparado con la ganancia en la aproximación de la solución, en teoría debe ser mucho más eficiente que los demás en los problemas de gran escala, de grandes dimensiones o los que requieran de muchas iteraciones, ideas que se podrían explorar en trabajos futuros.

Con las ideas expuestas en este trabajo se adaptó, para el caso del modelo del VIH/SIDA considerado en esta tesis, un programa que resuelve problemas de minimización asociados con la estimación de parámetros como los que se consideran en esta tesis. Un resumen de los resultados de las simulaciones es:

1. En general los parámetros se estiman bien usando pocos datos y aunque se hicieron experimentos usando una gran cantidad de datos, estos muestran que, aparte de ser caro e impráctico, no se obtienen beneficios en cuanto a la mejora de la estimación.
2. Para el criterio de paro dado en el algoritmo, $tol = 10^{-4}$ este convergió en promedio en 6 iteraciones.
3. Añadiendo ruido aleatorio del 10 % se obtienen buenas estimaciones para los parámetros.
4. Dentro de todo el desarrollo del programa se implementó una función que resuelve tanto el sistema de estado como el sistema adjunto.

Trabajo futuro:

En un trabajo futuro se espera comparar el desempeño entre BFGS y los métodos mencionados anteriormente como por ejemplo gradiente conjugado. Además se espera hacer simulaciones con perturbaciones mayores en los datos.

También se espera comparar formalmente con algún otro software, por ejemplo, DIFFPAR y estudiar lo referente a identificabilidad del problema de estimación de parámetros.

Bibliografía

- Alavez-Ramírez, J. (2007). *Estimación de Parámetros en Ecuaciones Diferenciales Ordinarias: Identificabilidad y Aplicaciones a Medicina* [Tesis doctoral, Universidad Nacional Autónoma de México].
- Alavez-Ramírez, J., Fuentes-Allen, J., & López-Estrada, J. (2011). Noninvasive monitoring of hepatic damage from hepatitis C virus infection. *Computational and mathematical methods in medicine*, 2011(1), 325470.
- Barberis, L. M., Benítez, L., & Condat, C. (2021). Elucidating the role played by cancer stem cells in cancer growth.
- Burden, R. L., & Faires, J. D. (2002). Análisis numérico. *Editorial Thomson*.
- Calver, J., & Enright, W. (2017). Numerical methods for computing sensitivities for ODEs and DDEs. *Numerical Algorithms*, 74(4), 1101-1117.
- Cao, Y., Li, S., Petzold, L., & Serban, R. (2003). Adjoint sensitivity analysis for differential-algebraic equations: The adjoint DAE system and its numerical solution. *SIAM journal on scientific computing*, 24(3), 1076-1089.
- CENSIDA. (2022). Boletín de Atención Integral de Personas con VIH.
- CIENI. (2011). 30 años del VIH-SIDA: Perspectivas desde México.
- Culshaw, R. V., Ruan, S., & Spiteri, R. J. (2004). Optimal HIV treatment by maximising immune response. *Journal of mathematical biology*, 48(5), 545-562.
- de la Pava Salgado, E. (2010). Modelado matemático de la transmisión de la gripe AH1N1. *Matemáticas: Enseñanza Universitaria*, 18(1), 1-10.
- Hill, A. L. (2018). Mathematical models of HIV latency. *HIV-1 latency*, 131-156.
- 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.
- Juárez Valencia, L. H., & Rojas, J. (2022). Parameter estimation in ODEs. Modelling and computational issues. *Boletín de la Sociedad Mexicana de Computación Científica y sus Aplicaciones*, (8), 34-49.

- Murillo, L. N., Murillo, M. S., & Perelson, A. S. (2013). Towards multiscale modeling of influenza infection. *Journal of theoretical biology*, 332, 267-290.
- Nocedal, J., & Wright, S. J. (1999). *Numerical optimization*. Springer.
- Nowak, M. A., Bonhoeffer, S., Hill, A. M., Boehme, R., Thomas, H. C., & McDade, H. (1996). Viral dynamics in hepatitis B virus infection. *Proceedings of the National Academy of Sciences*, 93(9), 4398-4402.
- Nuraini, N., Tasman, H., Soewono, E., & Sidarto, K. A. (2009). A with-in host dengue infection model with immune response. *Mathematical and Computer Modelling*, 49(5-6), 1148-1155.
- Olarte, J. A., & Muñoz, A. (2021). Modeling the dengue fever transmission in a periodic environment. *Revista Colombiana de Matemáticas*, 55(1), 71-107.
- ONUSIDA. (2024). *La urgencia del ahora. El SIDA frente a una encrucijada*.
- Peregrino, A., González, D., & López, J. (2023). A Mathematical Model to Estimate the COVID-19 Pandemic in Red Lights of Mexico. *Boletín de la Sociedad Matemática Mexicana de Computación Científica y sus Aplicaciones*, 9(9), 77-91.
- Perelson, A. S., Kirschner, D. E., & De Boer, R. (1993). Dynamics of HIV infection of CD4+ T cells. *Mathematical biosciences*, 114(1), 81-125.
- Perelson, A. S., & Nelson, P. W. (1999). Mathematical analysis of HIV-1 dynamics in vivo. *SIAM review*, 41(1), 3-44.
- Secretaría de Salud. (2012). Informe Nacional en la lucha contra el SIDA.
- Secretaría de Salud. (2020). *Manual de procedimientos estandarizados para la vigilancia epidemiológica de la infección por el Virus de la Inmunodeficiencia Humana*.
- Secretaría de Salud. (2024). *Vigilancia Epidemiológica de casos de VIH/SIDA en México. Registro Nacional de Casos de Sida Actualización 1er. Trimestre 2024*.
- Sungchait, R., Tang, I.-M., & Pongsumpun, P. (2022). Mathematical modeling: Global stability analysis of super spreading transmission of respiratory syncytial virus (RSV) disease. *Computation*, 10(7), 120.
- Wein, L. M., Zenios, S. A., & Nowak, M. A. (1997). Dynamic multidrug therapies for HIV: a control theoretic approach. *Journal of Theoretical Biology*, 185(1), 15-29.

Anexos

Programa genera datos.

```
%%% SOLUCION DE LA ECUACIÓN DE ESTADO EN UN INTERVALO INICIAL [t0,tf]
clearvars
%%%theta(1)=k
%%%theta(2)=miue
%%Se define el sistema diferencial
f = @(t,x,theta)[10+0.12*x(1)*(1-x(1)/1500)-1.33e-2*x(1)-theta(1)*x(1)*x(3);
theta(1)*x(1)*x(3)-theta(2)*x(2);2.6667*x(2)-0.4*x(3)];
%%%Se definen las condiciones iniciales, tiempos y parametros de
%%%discretizacion
x0 = [70; 30; 200];
t0 = 0;
tf = 30;
h = 0.0001; %es necesario un tamaño de paso muuy pequeño para no tener
%problemas al momento de resolver por rk el sistema de estado y el
%sistema adjunto, a lo largo de las iteraciones de BFGS.
thetaS = [10;0.12;1500;1.33e-2;3.1662e-3;0.5;2.6667;0.4];
T_s0=(thetaS(3)/(2*thetaS(2)))*(thetaS(2)-thetaS(4)
+sqrt((thetaS(2)-thetaS(4)).^2+4*thetaS(2)*thetaS(1)/thetaS(3)));
%%Calculamos el número reproductivo básico (R0),
%%el equilibrio trivial y el no trivial
R_0=(thetaS(5)*thetaS(7))/(thetaS(6)*thetaS(8))*T_s0;
E0=[T_s0,0,0]
V_a=1/thetaS(5)*(thetaS(1)/T_s0+thetaS(2)*T_s0/(R_0*thetaS(3)))*(R_0-1);
T_ea=thetaS(8)/thetaS(7)*V_a;
```

```

T_sa=thetaS(6)*thetaS(8)/(thetaS(5)*thetaS(7));
E_estrella=[T_sa, T_ea, V_a]
noise_level = 0.1; % Se escoge segun el nivel del ruido deseado
theta = [3.1662e-3;0.5];

%% Método de RK-4 para resolver de 0 a tf
[t,x] = rk4(f,theta,x0,t0,tf,h);
%%grafica las soluciones de 0 a tf
figure(1)
plot(t,x(1,:), 'b',t,x(2,:), 'c',t,x(3,:), 'r', 'LineWidth',2)
hold on
title('Exact state solution ', 'FontSize',15)
xlabel('Time (days)', 'FontSize',14)
ylabel('Populations', 'FontSize',14)
legend('T_s', 'T_e', 'V')

% PERTURBACIÓN DE Ts, Te y V PARA GENERAR DATOS SINTÉTICOS CON RUIDO
%GAUSIANO
n = length(t);
sigma = noise_level*abs(x);
dim=3;
xdatos = zeros(dim,n);
%Generamos ruido
xdatos(1,:) = x(1,)+ random('Normal',0,sigma(1,:),1,n);
xdatos(2,:) = x(2,)+ random('Normal',0,sigma(2,:),1,n);
xdatos(3,:) = x(3,)+ random('Normal',0,sigma(3,:),1,n);
%%grafica las soluciones de t0 a tf y
%%datos con ruido:
figure(2)
plot(t,x(1,:), 'b',t,x(2,:), 'c',t,x(3,:), 'r', 'LineWidth',2)
hold on
plot(t,xdatos(1,:), '. ',t,xdatos(2,:), '. ',t,xdatos(3,:), '. ')
title('Solución exacta y datos con ruido', 'FontSize',15)
xlabel('Tiempo (días)', 'FontSize',14)
ylabel('Poblaciones', 'FontSize',14)
hold off
legend('T_s', 'T_e', 'V')

```

Programa principal Main

```

%% Programa original: Agosto 12 de 2022
% Actualizacion: 15/08/2022
% Actualizacion: 19/10/2022
% Actualización: 21/06/2023
% Se utilizan los datos generados sinteticamente con el programa
% 'genera_datos'
% Si ya están generados se cargan desde el archivo:
% 1. 'xdatos', datos sin filtrar, con load('xdatos', '-ascii')
% 2. 'sigma', con load('sigma', '-ascii')
%% Parámetros del modelo de VIH
%% theta(1)=q
%% theta(2)=beta
%% theta(3)=Tmax
%% theta(4)=mius
%% theta(5)=k
%% theta(6)=miue
%% theta(7)=p
%% theta(8)=miuv
clc
format long
%% S: Modelo del VIH y matrices jacobianas transpuestas
%size(theta)
% Flujo f y transpuestas de Jacobianas f_x y f_theta tomando como
% parámetros a estimar k y mue
f = @(t,x,theta)[10+0.12*x(1)*(1-x(1)/1500)-1.33e-2*x(1)-theta(1)*x(1)*x(3);
theta(1)*x(1)*x(3)-theta(2)*x(2);2.6667*x(2)-0.4*x(3)];
fthT = @(x,y,theta)[x(1)*x(3)*(y(2)-y(1));-x(2)*y(2)];
fxT = @(x,y,source,theta)[(0.12*(1-2*x(1)/1500)-1.33e-2-theta(1)*x(3))*y(1)
+theta(1)*x(3)*y(2)+source(1);-theta(2)*y(2)+2.6667*y(3)+source(2);
-theta(1)*x(1)*y(1)+theta(1)*x(1)*y(2)-0.4*y(3)+source(3)];
dim = 3; % Dimensión de la ODE

% Parametros para las iteraciones
tol = 10^(-4); % Tolerancia para BFGS
max_iters = 50; % No maximo iters para BFGS

```

```

dindex = [3]; % Datos observables
%Subrutina para seleccionar datos en el intervalo [0,10]
index=round(linspace(1,100000,100));
m = length(index); % Numero de datos

t_dat = zeros(1,m); % Arreglo para guardar tiempos en la ventana de datos
x_dat = zeros(dim,m); % Arreglo para guardar datos experimentales
sigma_dat = zeros(dim,m); % Arreglo para las desviaciones estandar de los
%datos
j = 1; % Indices para t_dat, x_dat y sigma_dat, i.e 1:m
for i = index
    t_dat(j) = t(i); % Dado i se recupera j con j = i - index(1) + 1
    x_dat(:,j) = xdatos(:,i);
    s_dat(:,j) = sigma(:,i).^2;
    j = j+1;
end

% Valores iniciales para las iteraciones BFGS
theta_guess =[1e-3;0.4];
ntheta = 2;

x0_guess = xdatos(:,1);
p_guess = [x0_guess;theta_guess]; % Vector de parametros: x0 y theta

% Se llama al algoritmo BFGS
[ps,err,niter,normg] = BFGS(f,fxT,fthT,p_guess,index,dindex,x_dat,s_dat,
t_dat,h,tol,max_iters,m,t0);
% Arreglos en donde se guardan las iteraciones para x0 y theta
x0s = ps(1:dim,:);
thetas = ps(dim+1:dim+ntheta,:);

error_relativo_x0 = abs(x0s(:,end) - x0) ./ abs(x0);
disp(error_relativo_x0')
error_relativo_th = abs(thetas(:,end) - theta) ./ abs(theta);
disp(error_relativo_th')
disp(niter)

```

```

[x0s(1:dim,1), x0s(:,end) ; thetas(1:nthetas,1), thetas(:,end)]

% Graficacion de resultados

figure(6)
fig_ode(f,ps(:,end),index,dindex,t_dat,x_dat,h,niter)

figure(7)
fig_s(x0s,x0)

%figure(3)
figure(8)
fig_th(thetas,theta)
legend('\theta_1','\theta_2')%%,'FontSize',12
title('Convergencia de parámetros','FontSize',16)
figure(9)
fig_norm(normg)

```

Programa BFGS

```

function [ps,err,iter,normg] = BFGS(f,fxT,fthT,p,index,dindex,x_dat,s_dat,
t_dat,h,tol,max_iters,m,t0)
% Inicializacion
iter = 0;
H = eye(length(p));
g = GradJ(f,fxT,fthT,p,N,index,dindex,x_dat,s_dat,t_dat,h,iter);
d = -H*g;
rho1 = g'*p/(g'*d+eps);

% Arreglos para guardar iteraciones de p y de la norma de g
ps = zeros(length(p),max_iters);
normg = zeros(1,max_iters);
ps(:,1) = p;
normgref = norm(g);
normg(1) = normgref;

if normgref < tol

```

```

    err = tol*1e-1;
else
    err = 10^10;
end
err = 10^10;
i1=length(index);
t1 = t_dat(i1); % tiempo final de la ventana de dato

while err > tol && iter < max_iters
% BUSQUEDA DE LINEA: Metodo de la secante con 5 iters max
    fprintf('minimizando y calculando el gradiente en la iteracion
n de bfgs: %d\n', iter);
    rho = minJ_sec(f,fxT,fthT,p,d,N,index,dindex,x_dat,s_dat,
t_dat,h,rho1,10^-4,4,m,t0,t1,iter+1);
    dp = rho*d;
    p = p + dp; % p = (p >= 0).*p;
    gn = GradJ(f,fxT,fthT,p,N,index,dindex,x_dat,s_dat,t_dat,h,iter+1);
    dg = gn - g;
    u = dp;
    v = H*dg;
    denom = 1/(u'*dg);
    c = denom*(1 + denom*(v'*dg));
    H = H + c*(u*u') - denom*(u*v' + v*u');
    % Actualizacion para la siguiente iteracion
    g = gn;
    d = -H*g;
    rho1 = g'*dp/(g'*d+eps);
    iter = iter+1;
    ps(:,iter+1) = p;
    normg(iter+1) = norm(g);
    err = normg(iter+1)/normgref;
end
ps = ps(:,1:iter+1);
normg = normg(1:iter+1);
end

```

Programa para calcular el gradiente

```

function g =GradJ(f,fxT,fthT,p,index,dindex,x_dat,s_dat
,t_dat,h,iteration)
% CALCULO DEL GRADIENTE
np = length(p);
dim = size(x_dat); % Dimension de la edo y numero de datos
ntheta = np - dim(1); % longitud de theta
g = zeros(np,1);
s0 = p(1:dim(1));
theta = p(dim(1)+1:np);

% 1. Se resuelve la ecuacion de estado
i0 = index(1);
i1 = index(end);
t0 = t_dat(1); % tiempo inicial de la ventana de datos
t1 = t_dat(dim(2)); % tiempo final de la ventana de datos
[t,x] = rk4(f,theta,s0,t0,t1,h);
n = length(t);

figure(3)
plot(t,x(1,:), 'b',t,x(2,:), 'c',t,x(3,:), 'r', 'LineWidth',2)
title(['Solución de estado. ITERACIÓN ', num2str(iteration)], 'FontSize',15)
%title('state solution it: n', 'FontSize',15)
xlabel('Tiempo ', 'FontSize',14)
ylabel('Poblaciones', 'FontSize',14)
legend('T_s', 'T_e', 'V')

% 2. Se resuelve la ecuacion adjunta
m = length(index);
source = zeros(dim(1),m);
% Introduciendo las varianzas
source(dindex,:) = (x(dindex,index) - x_dat(dindex,1:m))
./ s_dat(dindex,1:m);
yf = zeros(dim(1),1); % Condicion final para la adjunta
y = rk4adj(fxT,theta,x,index,source,yf,h);

```

```

figure(4)
plot(t,y(1,:), 'b',t,y(2,:), 'c',t,y(3,:), 'r', 'LineWidth',2)
title(['Solución adjunta. ITERACIÓN', num2str(iteration)], 'FontSize',15)
xlabel('Tiempo ', 'FontSize',14)
ylabel('p', 'FontSize',14)
legend('p_1', 'p_2', 'p_3')%, 'FontSize',14

% 3. Se calcula el gradiente con respecto a x0

ds0 = zeros(dim(1),1);
ds0(dindex) = ( s0(dindex) - x_dat(dindex,i0) ) ./ s_dat(dindex,i0);
onev = ones(length(s0),1);
%g(1:5) = y(:,1) + ds0 + (s0'*onev - N)*onev;
g(1:dim(1)) = y(:,1) + ds0;

% 4. Se calcula la integral para el gradiente respecto a theta

H = zeros(ntheta,n);
for i = 1:n
    H(:,i) = fthT(x(:,i),y(:,i),theta);
end
% Metodo de Simpson
sum = zeros(np-dim(1),1); %JLL
for i = 2:2:n-1
    sum = sum + H(:,i-1) + 4*H(:,i) + H(:,i+1);
end
integral = h*sum/3; % El paso es en realidad 2*h y 2*h/6 = h/3
g(dim(1)+1:np) = integral;
g = g + (eps)*p; % Regularizacion, ayuda a estabilizar en casos
dificiles JLL
end

```

Programa de Búsqueda en línea

```

function rho = minJ_sec_3p(f,fxT,fthT,p,d,index,dindex,x_dat,s_dat,
t_dat,h,rho1,tol,max_iters,m,t0,t1,iter)

g0 = GradJ(f,fxT,fthT,p+rho0*d,index,dindex,x_dat,s_dat,
t_dat,h,iter) '*d;
g1 = GradJ(f,fxT,fthT,p+rho1*d,N,index,dindex,x_dat,s_dat,
t_dat,h,iter) '*d;
gref = g0;
err = 10^10;
iter = 0;

while err > tol && iter < max_iters
    drho = rho1 - rho0;
    dg = g1 - g0;
    err = min( abs(drho)/abs(rho1), norm(dg)/norm(gref) );
    g0 = g1;
    rho0 = rho1;
    rho1 = rho0 - g0*drho/(dg+eps);
    g1 = GradJ(f,fxT,fthT,p+rho1*d,N,index,dindex,x_dat,s_dat,
t_dat,h,iter) '*d;
    iter = iter + 1;
end
rho = rho1;
end

```

Alojamiento de la Tesis en el Repositorio Institucional	
Título de Tesis:	Estimación de parámetros de un modelo de VIH por el método BFGS y la adjunta.
Autor de la Tesis:	Sarai Martínez Méndez
ORCID:	0009-0002-5514-256
Resumen de la Tesis:	En este trabajo se realizó estimación de parámetros de un modelo de VIH/SIDA mediante el algoritmo de minimización Broyden- Fletcher- Goldfarb- Shanno (BFGS), que no solo toma información del gradiente sino también de la Hessiana. Para calcular el gradiente se utilizó un método variacional que implica resolver el sistema adjunto del sistema de estado. El sistema de estado y el sistema adjunto se resolvieron en cada iteración usando el método Runge-Kutta de cuarto orden. Para perturbaciones en los datos del 10 %, se obtuvieron estimaciones con un error máximo en promedio del 4 %.
Palabras claves de la Tesis:	Estimación de parámetros, sistema adjunto, algoritmo BFGS, Runge-Kutta, Gradiente, Hessiana.
Referencias citadas:	<p>Alavez-Ramírez, J. (2007). Estimación de Parámetros en Ecuaciones Diferenciales Ordinarias: Identificabilidad y Aplicaciones a Medicina [Tesis doctoral, Universidad Nacional Autónoma de México].</p> <p>Alavez-Ramírez, J., Fuentes-Allen, J., & López-Estrada, J. (2011). Noninvasive monitoring of hepatic damage from hepatitis C virus infection. <i>Computational and mathematical methods in medicine</i>, 2011(1), 325470.</p> <p>Barberis, L. M., Benítez, L., & Condat, C. (2021). Elucidating the role played by cancer stem cells in cancer growth.</p> <p>Burden, R. L., & Faires, J. D. (2002). <i>Análisis numérico</i>. Editorial Thomson.</p> <p>Calver, J., & Enright, W. (2017). Numerical methods for computing sensitivities for ODEs and DDEs. <i>Numerical Algorithms</i>, 74 (4), 1101-1117.</p>

- Cao, Y., Li, S., Petzold, L., & Serban, R. (2003). Adjoint sensitivity analysis for differential algebraic equations: The adjoint DAE system and its numerical solution. *SIAM journal on scientific computing*, 24 (3), 1076-1089.
- CENSIDA. (2022). Boletín de Atención Integral de Personas con VIH.
- CIENI. (2011). 30 años del VIH-SIDA: Perspectivas desde México. Culshaw, R. V., Ruan, S., & Spiteri, R. J. (2004). Optimal HIV treatment by maximising immune response. *Journal of mathematical biology*, 48(5), 545-562.
- de la Pava Salgado, E. (2010). Modelado matemático de la transmisión de la gripe AH1N1. *Matemáticas: Enseñanza Universitaria*, 18(1), 1-10.
- Hill, A. L. (2018). Mathematical models of HIV latency. *HIV-1 latency*, 131-156.
- 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.
- Juárez Valencia, L. H., & Rojas, J. (2022). Parameter estimation in ODEs. Modelling and computational issues. *Boletín de la Sociedad Mexicana de Computación Científica y sus Aplicaciones*, (8), 34-49.
- Murillo, L. N., Murillo, M. S., & Perelson, A. S. (2013). Towards multiscale modeling of influenza infection. *Journal of theoretical biology*, 332, 267-290.
- Nocedal, J., & Wright, S. J. (1999). *Numerical optimization*. Springer.
- Nowak, M. A., Bonhoeffer, S., Hill, A. M., Boehme, R., Thomas, H. C., & McDade, H. (1996). Viral dynamics in hepatitis B virus infection. *Proceedings of the National Academy of Sciences*, 93 (9), 4398-4402.
- Nuraini, N., Tasman, H., Soewono, E., & Sidarto, K. A. (2009). A with-in host dengue infection model with immune response. *Mathematical and Computer Modelling*, 49(5-6), 1148-1155.

Olarte, J. A., & Muñoz, A. (2021). Modeling the dengue fever transmission in a periodic environment. *Revista Colombiana de Matemáticas*, 55(1), 71-107.

ONUSIDA. (2024). La urgencia del ahora. El SIDA frente a una encrucijada.

Peregrino, A., González, D., & López, J. (2023). A Mathematical Model to Estimate the COVID-19 Pandemic in Red Lights of Mexico. *Boletín de la Sociedad Matemática Mexicana de Computación Científica y sus Aplicaciones*, 9(9), 77-91.

Perelson, A. S., Kirschner, D. E., & De Boer, R. (1993). Dynamics of HIV infection of CD4+ T cells. *Mathematical biosciences*, 114(1), 81-125.

Perelson, A. S., & Nelson, P. W. (1999). Mathematical analysis of HIV-1 dynamics in vivo. *SIAM review*, 41(1), 3-44.

Secretaría de Salud. (2012). Informe Nacional en la lucha contra el SIDA.

Secretaría de Salud. (2020). Manual de procedimientos estandarizados para la vigilancia epidemiológica de la infección por el Virus de la Inmunodeficiencia Humana.

Secretaría de Salud. (2024). Vigilancia Epidemiológica de casos de VIH/SIDA en México. Registro Nacional de Casos de Sida Actualización 1er. Trimestre 2024.

Sungchait, R., Tang, L.-M., & Pongsumpun, P. (2022). Mathematical modeling: Global stability analysis of super spreading transmission of respiratory syncytial virus (RSV) disease. *Computation*, 10(7), 120.

Wein, L. M., Zenios, S. A., & Nowak, M. A. (1997). Dynamic multidrug therapies for HIV: a control theoretic approach. *Journal of Theoretical Biology*, 185(1), 15-29.