

Métodos Dinámicos en Economía.
Otra Búsqueda del Tiempo Perdido

Héctor Lomelí & Beatriz Rumbos
Instituto Tecnológico Autónomo de México
Río Hondo #1
01000 México DF

27 de septiembre de 2001

Prefacio

Acerca de este Libro

La enseñanza de métodos cuantitativos para economistas ha sufrido grandes cambios durante la última década. En el pasado era posible, para un estudiante de licenciatura en economía, tener una idea clara de su profesión con un mínimo de conocimientos cuantitativos. Éstos no pasaban más allá del cálculo diferencial e integral y algunos rudimentos de métodos dinámicos. Más importante aún, era el que esta herramienta cuantitativa se adquiría con poca formalidad: lo importante era el saber utilizarla para los fines requeridos (por ejemplo, para resolver un problema de optimización) sin profundizar mucho en su origen.

La labor del economista es, y debe seguir siendo, la del estudio y creación del conocimiento económico; sin embargo, dado que el lenguaje utilizado para la descripción de este conocimiento es el de las matemáticas, no queda más remedio que conocerlo a fondo para evitar su uso incorrecto. Es en este sentido que han cambiado los cursos de matemáticas para economistas. De manera dual, dentro del ámbito de las matemáticas aplicadas, la enseñanza de la economía se ha vuelto sumamente importante en ciertas áreas. Una gran cantidad de problemas a resolver por el matemático aplicado tienen su origen en algún fenómeno económico. Los casos más claros son el de la teoría de juegos y las matemáticas financieras.

Durante los últimos años hemos estado enseñando diversos cursos de matemáticas para economistas, a nivel de licenciatura y posgrado, mismos que son frecuentados también por matemáticos y actuarios. En el caso de los cursos de cálculo y de álgebra lineal, existe una gran variedad de textos que ofrecen los niveles requeridos de profundidad matemática y aplicaciones económicas. El problema surge para cursos más avanzados como son los relacionados con métodos dinámicos y optimización. Existe una gran cantidad de textos, en su mayoría escritos por economistas, que cubren éstos y otros temas. Por mencionar sólo algunos, tenemos, [BS94], [Chi84], [Chi92], [dlF00], [KS91], [Sho97], [SS87] y [Tak85]. Debemos dar crédito a todos éstos ya que fueron un punto de partida y fuente de inspiración para este libro. El inconveniente que encontramos fue que ninguno de estos libros cubría exactamente nuestras necesidades.

Los problemas iban desde la falta hasta el exceso de formalidad, combinado con la ausencia de algunos temas y carencia de ejemplos suficientes. Así, la teoría se tomaba de un texto, los ejemplos de otro, la parte de sistemas dinámicos continuos de otro, los discretos de uno más, etc. Aunado a esto, algunos temas, como las ecuaciones en diferencia estocásticas, se tenían que buscar en apéndices de textos de macroeconomía avanzada (como [BF92] o [Sar87], por ejemplo).

Motivados por la necesidad, comenzamos la escritura de nuestras propias notas para estos cursos, inicialmente, cada quien por su lado. Después de un par de semestres, en aras de la búsqueda de eficiencia, decidimos cooperar en la escritura de las notas. Al poco tiempo surgió la idea de que valía la pena convertir estos apuntes en un texto formal. A partir de entonces repartimos las notas entre los profesores que imparten estos cursos para detectar errores, encontrar omisiones y obtener sugerencias. Ahora, más de un año después, consideramos que el texto ha llegado a un punto en el cual puede ser publicado.

La diferencia de este texto con respecto a otros, es que está diseñado tanto para economistas que deben aprender métodos dinámicos como para matemáticos que desean aprender algo de economía. La mayoría de los resultados se presentan con sus demostraciones correspondientes con el propósito de tener un mínimo de formalidad. Incluimos una gran cantidad de aplicaciones a la economía, tratando de que éstas puedan ser entendidas tanto por matemáticos como economistas. Se hace énfasis en el análisis cualitativo de los sistemas dinámicos por ser éste de gran utilidad en el estudio de un gran número de modelos económicos. Se incluye un capítulo de análisis convexo y optimización estática con el fin de que el texto sea lo más autocontenido posible. El capítulo referente a “La Mano Invisible” es novedoso, en cuanto a que lleva al lector a darse cuenta de la complejidad existente en modelos económicos de aparente sencillez.

Los tópicos incluidos son aquellos que consideramos importantes por su utilización en la economía. Por poner un ejemplo, el equilibrio de punto silla no es particularmente interesante para el especialista en sistemas dinámicos; sin embargo, en economía es un concepto sumamente útil. Este texto no pretende ser un texto exhaustivo en ninguno de los temas que trata, y el estudiante de posgrado puede requerir de un tratamiento más profundo en algún tema específico. Para este propósito, tratamos de dar las referencias adecuadas para aquél que lo requiera. El libro no es un texto tradicional de sistemas dinámicos, de teoría de control o de programación dinámica; se puede pensar como un libro escrito por matemáticos para economistas o bien como un libro escrito por economistas para matemáticos.

Requisitos y Lectura del Libro

Estrictamente hablando, suponemos que al adquirir este texto el estudiante habrá cursado todos sus cursos de cálculo (diferencial, integral y vectorial) y preferentemente algún curso básico de álgebra lineal. Adicionalmente, esperamos que se tengan conocimientos elementales de economía, o al menos una idea de los conceptos básicos junto con la terminología asociada. En la medida de lo posible, tratamos de desarrollar todos los modelos económicos en detalle, para hacerlos accesibles al público no especialista en economía. Lo análogo sucede con los resultados matemáticos del texto. Los apéndices del libro incluyen, ya sea un relleno a ciertos “huecos” en el conocimiento como números complejos y elementos de espacios vectoriales, así como temas más avanzados que complementan los temas tratados en el libro como teoremas de separación y topología elemental.

El texto fue escrito pensando en que puede leerse a varios niveles de profundidad. Estamos conscientes de que no todos los estudiantes de economía se dedicarán a la teoría económica y, por lo tanto, quizás no requieran de mucha formalidad. Por ejemplo, un curso de un semestre de métodos dinámicos puede consistir en los siguientes capítulos y secciones, *omitiendo las demostraciones formales y dando únicamente la intuición detrás de los resultados*.

- Capítulo 1.
- Capítulo 2, omitiendo la sección 2.3.4.1.
- Capítulo 3 en su totalidad.
- Capítulos 4 y 5, restringiéndose únicamente el caso de sistemas de 2×2 .
- Secciones 6.1 y 6.2 del capítulo 6. Adicionalmente, analizar el ejemplo 6.3.2 de la sección 6.3.
- Capítulo 7, omitiendo la sección 7.5.1.1.
- Capítulo 11, omitiendo la sección 11.4.
- Capítulo 12, omitiendo el ejemplo 12.4.1.
- Capítulo 14, únicamente las secciones 14.1, 14.2 y 14.3.

Un curso de sistemas dinámicos de un semestre puede consistir de los capítulos 1 al 9. La cantidad de demostraciones que se incluyan depende del nivel del curso. Para un curso de posgrado se sugiere, no sólo incluir todas las demostraciones sino los apéndices B y C.

Un curso de optimización para estudiantes que ya hayan cursado sistemas dinámicos, puede consistir del capítulo 10, suponiendo que este material no se ha visto en un curso anterior, junto con los capítulos 11 a 14. Una vez más, para un curso avanzado, se sugiere ver todas las demostraciones y los apéndices D y E.

Adicionalmente, consideramos que el texto también cumple las funciones de un libro de referencia para profesionistas en las áreas de economía y matemáticas que hayan estados alejados de estos temas. Debemos aclarar que no se trata de una enciclopedia de métodos cuantitativos para economistas. No sólo damos por hecho que el estudiante llega con conocimientos sólidos de cálculo, sino que omitimos temas de extrema importancia en economía como lo son los teoremas del punto fijo, dualidad y teoría de juegos, entre otros. No obstante, tratamos de incluir todos los temas relevantes de métodos dinámicos utilizados en economía. Cabe mencionar algunas excepciones: métodos numéricos, ecuaciones diferenciales parciales y cálculo estocástico. Consideramos que estos temas, por su importancia, ameritan un libro de texto cada uno; además, su utilidad en economía es dentro de temas muy específicos por lo que abandonamos al lector interesado en lo que respecta al aprendizaje de los mismos.

Agradecimientos

Varios de nuestros colegas han utilizado este libro en sus cursos y lo han enriquecido con sus sugerencias y comentarios. Quisiéramos mencionar de manera muy especial a la Dra. Lorena Zogaib, cuya lectura

cuidadosa del manuscrito fue invaluable. Gracias a ella se localizaron y corrigieron un gran número de errores y se realizaron cambios de orden didáctico la adición de ejercicios y gráficas y el esclarecer ciertos conceptos tanto matemáticos como económicos.

Por azares del destino, tuvimos la suerte y el honor de tener a un corrector de estilo altamente profesional. Fue un placer contar con la colaboración de Tomás Granados Salinas para lograr que la lectura de este texto fuese lo más amena posible. Su labor fue mucho más allá de una simple labor de redacción, su formación de matemático y su experiencia editorial permitieron una cuidadosa revisión del manuscrito. Cualquier error restante es responsabilidad nuestra y agradecemos al amable lector hacernoslo saber de manera que en una edición futura pueda ser corregido.

ÍNDICE GENERAL

Prefacio	i
I Introducción	1
1 En defensa de los modelos matemáticos	3
1.1 Introducción	3
1.2 ¿Qué es un modelo matemático?	5
1.2.1 Abstracción del mundo	6
1.2.2 Deducción a partir del modelo	9
1.2.3 Verificación, predicción y usos	10
1.3 El uso del tiempo en economía	13
1.4 Un ejemplo	14
1.4.1 Abstracción del mundo	15
1.4.2 Deducción a partir del modelo	16
1.4.3 Verificación y conclusiones	16
Ejercicios	17
II Ecuaciones Diferenciales y Dinámica Continua	19
2 Ecuaciones diferenciales lineales	21
2.1 Introducción	21
2.2 Ecuaciones de primer orden	23
2.2.1 Caso autónomo	23
2.2.2 Caso no autónomo	27
2.2.3 Bonos y tasas de interés	28

2.2.4	Caso no autónomo y no homogéneo	31
2.3	Ecuaciones de segundo orden	33
2.3.1	Raíces reales distintas	33
2.3.2	Una raíz real doble	34
2.3.3	Raíces complejas	35
2.3.4	Caso no homogéneo	36
2.3.5	Aversión absoluta al riesgo constante	39
	Ejercicios	41
3	Ecuaciones no lineales de primer orden	45
3.1	Ecuaciones separables	45
3.1.1	Aversión relativa al riesgo constante	46
3.1.2	Modelo logístico	47
3.2	Ecuación de Bernoulli	48
3.3	Diagramas de fase y estabilidad	50
3.3.1	Modelo de Solow-Swan	52
3.4	Expectativas y estabilidad: un modelo monetario	56
3.4.1	Expectativas adaptativas	56
3.4.2	Expectativas racionales	58
	Ejercicios	60
4	Sistemas de ecuaciones diferenciales lineales	63
4.1	Introducción	63
4.2	Método de valores propios	66
4.2.1	Valores propios reales distintos	69
4.2.2	Valores propios complejos	74
4.2.3	Valores propios reales repetidos	79
4.3	El caso no homogéneo	81
4.4	Ecuaciones lineales de orden superior	83
	Ejercicios	92
5	Análisis cualitativo	95
5.1	Puntos de equilibrio	95
5.2	Clasificación de puntos de equilibrio	97
5.3	Diagramas de fase	100
5.4	Clasificación de sistemas lineales de 2×2	104
5.4.1	Región I	105
5.4.2	Región II	106
5.4.3	Región III	106
5.4.4	Región IV	107
5.4.5	Región V	108
5.4.6	Casos degenerados	109

5.5	Linealización de sistemas no lineales	109
5.6	Análisis de puntos silla	114
5.7	Dos aplicaciones	118
5.7.1	Sobreajuste del tipo de cambio	118
5.7.2	Devaluación: niveles <i>vs.</i> tasas	123
	Ejercicios	127
III	Ecuaciones en Diferencias y Dinámica Discreta	131
6	Conceptos básicos de dinámica discreta	133
6.1	Introducción	133
6.2	Solución de ecuaciones lineales	135
6.3	Análisis cualitativo	137
6.4	Caos	145
6.5	Modelo logístico discreto	148
	Ejercicios	156
7	Sistemas de ecuaciones en diferencias lineales	161
7.1	Introducción	161
7.2	La solución general	162
7.3	Caso no homogéneo	162
7.4	Análisis cualitativo	163
7.5	Ecuaciones lineales de segundo orden	167
7.5.1	Caso no homogéneo	168
	Ejercicios	171
8	La mano invisible	173
8.1	Introducción	173
8.2	El modelo básico	174
8.3	Dinámica de precios	177
8.4	Caos en el proceso de precios	180
8.5	Funciones de Liapunov	181
8.6	Axioma de la preferencia revelada	182
8.7	Otras posibilidades	184
	Ejercicios	185
9	Ecuaciones en diferencias estocásticas	187
9.1	Iteración hacia adelante	187
9.2	Valor esperado y expectativas racionales	190
9.3	Algunos procesos estocásticos	193
9.4	Burbujas	196

9.5	Forma reducida	197
9.5.1	Método de iteración	198
9.5.2	Coeficientes indeterminados	201
	Ejercicios	205
IV	Intermezzo	207
10	Optimización estática	209
10.1	Análisis convexo	209
10.1.1	Caracterización de las funciones cóncavas y convexas	214
10.1.2	Matrices definidas	216
10.1.3	Funciones cuasi cóncavas y cuasi convexas	218
10.2	Optimización estática.	221
10.2.1	Restricciones de igualdad	223
10.2.2	Dos aplicaciones microeconómicas.	227
10.2.3	Condiciones de Kuhn-Tucker	230
10.2.4	Teorema de la envolvente	234
	Ejercicios	238
V	Optimización Dinámica	241
11	Introducción al cálculo en variaciones	243
11.1	Preliminares	243
11.2	Ecuación de Euler	249
11.3	Modelo de Ramsey	254
11.4	Extensiones a la ecuación de Euler	260
11.4.1	Varias variables	260
11.4.2	Derivadas de orden superior	261
11.5	Condiciones de segundo orden	263
11.6	Condiciones de transversalidad	265
11.7	Problemas con horizonte infinito	269
11.7.1	Un modelo de inversión	271
	Ejercicios	273
12	Teoría de control	275
12.1	Introducción	275
12.2	Planteamiento del problema	276
12.3	Otras condiciones de transversalidad	282
12.4	Problemas con horizonte infinito	286
12.5	Hamiltoniano en tiempo corriente	288

12.6 Problemas con más de una variable	292
12.7 Interpretación económica del problema de control	294
12.8 Dos aplicaciones	298
12.8.1 Un modelo monetario	298
12.8.2 Una economía pequeña y abierta	301
Ejercicios	307
13 Problemas con restricciones	313
13.1 Restricciones de igualdad sobre las variables de control	313
13.2 Restricciones integrales	316
13.3 Restricciones de desigualdad sobre las variables de control	319
13.3.1 Otro modelo de inversión	320
13.4 Restricción sobre el tiempo terminal	322
13.5 Restricciones en el espacio de estados	322
Ejercicios	328
14 Elementos de programación dinámica	331
14.1 Introducción	331
14.2 Estructura del problema	332
14.3 Problemas con descuento temporal	337
14.4 Problemas con horizonte infinito	338
14.5 Modelo de Ramsey discreto	339
14.6 Encontrando la función valor	342
14.6.1 El viñedo de Weitzman	347
14.7 Programación dinámica estocástica	349
14.7.1 Introducción	349
14.7.2 Planteamiento del problema	350
14.8 Algunas aplicaciones	351
14.8.1 El consumo como martingala	351
14.8.2 Consumo de bienes duraderos	353
14.8.3 Ciclos económicos	356
14.8.4 Rendimientos de activos (CAPM)	358
Ejercicios	361
VI Apéndices	365
A Números complejos	367
A.1 Propiedades básicas	367
A.2 Representación polar y fórmula de Euler	368
B Existencia y unicidad de soluciones	371

C Raíces repetidas	379
C.1 Multiplicidad y defecto de un valor propio	379
C.2 Vector propio generalizado	380
C.3 Construcción de soluciones linealmente independientes	381
D Elementos básicos de topología en \mathbb{R}^n	387
D.1 Algo sobre números reales	387
D.2 Abiertos, cerrados y algo más.	389
D.3 Sucesiones	394
E Espacios vectoriales y teoremas de separación	401
E.1 Propiedades básicas de los espacios vectoriales	401
E.2 Hiperplanos y teoremas de separación	406
Bibliografía	414

Parte I

Introducción

CAPÍTULO 1

En defensa de los modelos matemáticos

§1.1 Introducción

*La profesión de economía se divide en macroeconomía, que observa lo que no puede explicarse, y microeconomía, que explica lo que no puede observarse.*¹

El uso de las matemáticas en economía es vista, con frecuencia, como el origen de un sinnúmero de problemas terribles en el mundo. Los enemigos más vociferantes de esta matematización consideran a las matemáticas como invasoras en un campo que no les corresponde, restringiendo el desarrollo de la disciplina económica y atrapándola dentro de un marco de formalidad innecesaria. Por otro lado, la economía moderna, al menos desde un punto de vista académico, no existiría sin el uso sistemático del lenguaje matemático. Para corroborar esto, invitamos al lector a tomar cualquier revista académica especializada en economía, y después abrirla en una página al azar: imaginemos que desaparece todo el lenguaje matemático, ¿qué queda? Muy poco.

El estudiante de economía tiende a sentir ansiedad por saber si todos esos modelos llenos de matemáticas le serán útiles algún día. Dado que esa aplicación esperada generalmente no es inmediata, es necesario aclarar el contexto en el que se justifica el uso de la sofisticación teórica del lenguaje matemático. En este capítulo nos proponemos defender la presencia de las matemáticas en la ciencias económicas, aunque con plena conciencia de que pueden existir posibles abusos. Esencialmente, el problema que tenemos enfrente es el de entender cuál es el origen, la pretensión y el objetivo de los modelos matemáticos y justificar su uso en la economía. Intentaremos dar los conceptos básicos de lo que se entiende por modelo matemático; sin embargo, dada nuestra incapacidad de entrar en una plena argumentación filosófica, hacemos este ejercicio desde el limitado entendimiento que los autores pueden tener de su propia labor.

¹Nota anónima tomada de [CH97].

En este libro nos proponemos estudiar la herramienta matemática utilizada en los modelos en los que el tiempo es una variable. Nos parece pertinente hacer notar, que el uso del tiempo en los modelos económicos es un hecho reciente, no obstante que desde el inicio de la economía se consideró que los modelos debían ser esencialmente dinámicos, es decir, tenían que involucrar una evolución temporal. Baste recordar a la célebre “mano invisible” propuesta por Adam Smith: el tiempo se encargará de llevar los sistemas económicos al equilibrio óptimo. El interés tardío en la variable temporal se pueda explicar, acaso, por el rechazo, dentro de ciertos círculos, a todo aquello que pudiera sonar a marxismo, es decir, la evolución en el tiempo y en particular el pretendido uso de métodos de la dialéctica.

Los modelos dinámicos han existido en las ciencias exactas casi desde su inicio. Junto con ellos, las matemáticas también han evolucionado. El mejor ejemplo de esto es la relación entre la llamada mecánica clásica y la teoría de sistemas dinámicos o teoría de los sistemas no lineales. La pregunta principal que la teoría de sistemas dinámicos pretende contestar es acerca del comportamiento futuro de un sistema que evoluciona. Los sistemas dinámicos siguen, de este modo, la vieja tradición gitana de querer predecir aquello que va a ocurrir en el futuro. Tradicionalmente se piensa que es esta capacidad de predicción lo que distingue a una verdadera ciencia exacta; son estas predicciones las que verifican, de forma inequívoca, los modelos teóricos de la realidad.

En contraste con esto, más allá de una simple verificación, las matemáticas intentan moverse en el ámbito de las demostraciones y de la plena certeza. Aunque Gödel demostró que esta certeza no es del todo posible, esta distinción hace que las matemáticas sean una ciencia distinta a las demás. Las matemáticas no comparten la manera de proceder de las ciencias naturales. Son independientes a ellas, tienen existencia propia. Sin embargo, paradójicamente, son las mejores servidoras de las ciencias exactas y sociales. Las matemáticas no son animales ni vegetales, son un ente extraño que nos proporciona, al mismo tiempo, un lenguaje universal y una estructura lógica impecable. El lenguaje natural, junto con el discurso literario que lo acompaña, puede, si es suficientemente persuasivo, tener eslabones faltantes en el razonamiento. El lenguaje matemático previene contra esta falla y pone por buen camino al esfuerzo creativo dentro de un marco teórico con estructura propia, evitando que éste divage sin sentido en el mundo de la fantasía.

Recordemos que Bertrand Russell decía que “Cabe definir la Matemática como la materia en la que nunca sabemos de qué estamos hablando, o si lo que decimos es verdad.” Con mejor tino Einstein dijo “En la medida en que las leyes de las Matemáticas conciernen a la realidad, no son ciertas; y en cuanto que son ciertas, no conciernen a la realidad”.²

Más allá de la capacidad de manipulación matemática, un científico debería ser apreciado por su capacidad de percibir la realidad e intuir lo que es ésta. Albert Einstein no era el mejor matemático de su época.[Fra89] Cuando era joven ni siquiera fue reconocido como un estudiante brillante. Sin embargo, le bastó un año –el *annus mirabilis* 1905– para revolucionar la física. Los principios en los que basó su teorías, eran generalmente simples. Su genialidad no consistía tanto en su capacidad matemática sino en su inmensa capacidad como modelador. Otros³ fueron los que aportaron las matemáticas que le eran necesarias.

A continuación intentaremos responder a la pregunta de qué es un modelo matemático. El estudio

²Ambas citas se pueden leer en [Col99].

³Sobre todo Minkowski.

de la economía como ciencia, el uso de las matemáticas y en general la crítica a la profesión se han aglutinado en algo que es conocido como metodología de la economía. Existen muchas obras que se podrían consultar para profundizar sobre el tema, en particular, remitimos al lector interesado a [Bla92].

§1.2 ¿Qué es un modelo matemático?

El por qué y cómo conocemos la realidad es, acaso, el problema filosófico de mayor trascendencia, quizás el único. La realidad nos rodea, pero para conocerla plenamente sería necesario tomar en cuenta todos sus aspectos. Desde un punto de vista puramente científico, esto es imposible. Una de las ideas de la ciencia moderna es estudiar la realidad en pequeñas porciones, limitando la pretensión del conocimiento a una área limitada. Como consecuencia, vemos cómo, dentro de las disciplinas académicas, existe cada vez más una enorme especialización y es muy difícil pretender conocer algo en más de un aspecto.

El método de estudio de la realidad necesariamente viene dado por el objeto mismo. Ante un fenómeno de tipo religioso no se pueden utilizar métodos de la física, y no se puede juzgar con justicia una obra literaria con métodos científicos. Esto limita aún más el campo de acción de las ciencias exactas, pues casi, por definición, éstas tienen como único lenguaje a las matemáticas.

Es a partir de Galileo que las matemáticas tienen un papel central en la física y, en general, en toda la ciencia. Con una reducción consciente o inconsciente de lo que se entiende por realidad, se han ido aproximando distintas “leyes de la naturaleza”. Estamos ante una limitación muy grande, pues en el pensamiento contemporáneo se ha identificado lo real con lo científico y aunque la ciencia es un proyecto que evoluciona, pareciera que dentro de nuestro pensamiento moderno, sólo estas “leyes de la naturaleza” pueden identificarse plenamente con lo real.

La economía no escapa a esta manera de ver el mundo, y a partir del siglo XIX se da a la tarea de “matematizarse”, es decir, a utilizar cada vez más las matemáticas como lenguaje de expresión. Dicen los detractores del lenguaje matemático, que esto fue motivado por “envidia” a la física, y la mayor crítica es que el edificio teórico resultante no ha tenido la capacidad de predicción de ésta última. Por lo tanto, para muchos la economía tiene la pretensión inalcanzable de ser una ciencia exacta.

Por un lado, es un error el querer comparar a la economía con la física. Un economista tiene más en común con un meteorólogo que con un físico ya que ambos comparten la frustración de poder predecir el futuro de manera muy limitada. Por otro lado, desde su nacimiento, la economía trata con conceptos cuantitativos: cantidades de bienes, precios, medidas de riqueza, etc., de manera que el lenguaje matemático es el candidato ideal para la descripción de los fenómenos económicos.

Para responder a la pregunta de ¿qué es un modelo matemático?, primero debemos ponernos de acuerdo en lo que entenderemos por modelo. Jesús Mosterín [Mos00] nos indica que el término “modelo”, se utiliza comúnmente en al menos dos sentidos, en ocasiones contrapuestos. Por un lado, este autor identifica un “modelo” con el objeto representado, en el sentido en el que un pintor hace una pintura de un modelo. La representación de este “modelo” constituye una teoría. Lo importante es que se puede conservar la abstracción de la teoría y así aplicarla a distintos “modelos” de ella. Esto es, la mecánica de Kepler es una teoría para la cual nuestro sistema solar es un simple “modelo”. Por otro lado, Mosterín reconoce que se le puede dar otro significado a la palabra, y esto es identificarla con la teoría misma. Este es el sentido que le daremos aquí a la palabra modelo, pues nos interesa poder añadir el adjetivo

“matemático”. Nuestro sentido es el que comúnmente se utiliza en el ámbito matemático y económico. Esto es, cuando usemos la palabra modelo, hablaremos generalmente de una teoría que, en principio, es una representación abstracta de la realidad.

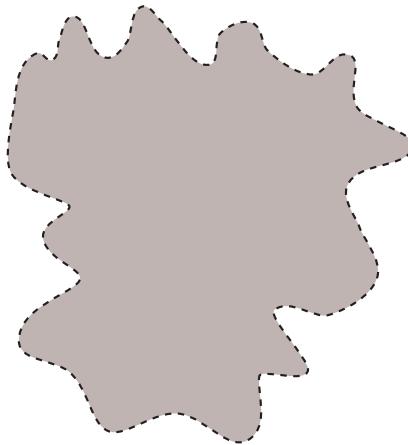


Figura 1.1: Una representación completamente inútil de la realidad.

El proceso de modelación matemática se puede entender a partir de una abstracción de los elementos en juego. Haremos a continuación una abstracción de esto. (¿Acaso un meta-modelo?) Por un lado tenemos a la realidad, a la cual consideramos infinita y sólo parcialmente accesible. Un modelo de la realidad puede verse en la figura 1.1. En dicha figura, la representación no proporciona ninguna información. Quisiéramos tener algo que nos fuera útil. La introducción de medidas en el modelo nos permite poder usar números; de este modo aparecen las matemáticas en juego. Reconocemos en nuestra pequeña abstracción tres pasos que en mayor o menor grado determinan al proceso de **modelación**:

- abstracción del mundo,
- deducción a partir del modelo ,
- verificación, predicción y usos.

Éstos pasos, que más adelante explicaremos con mayor detalle, se ilustran en la figura 1.2.

§1.2.1 Abstracción del mundo

Ante la complejidad del mundo, en particular del mundo económico, podemos responder con un principio utilizado frecuentemente. Hablamos de la idea de *ceteris paribus*, frase del latín que literalmente quiere decir “todo lo demás igual”. Es decir, que sólo se consideran ciertos factores y todo lo demás se mantiene constante. En realidad, este principio no tiene nada nuevo dentro de la ciencia, pues todo el conocimiento científico está basado en él. Sin embargo, bien ha hecho la economía en insistir y recalcar

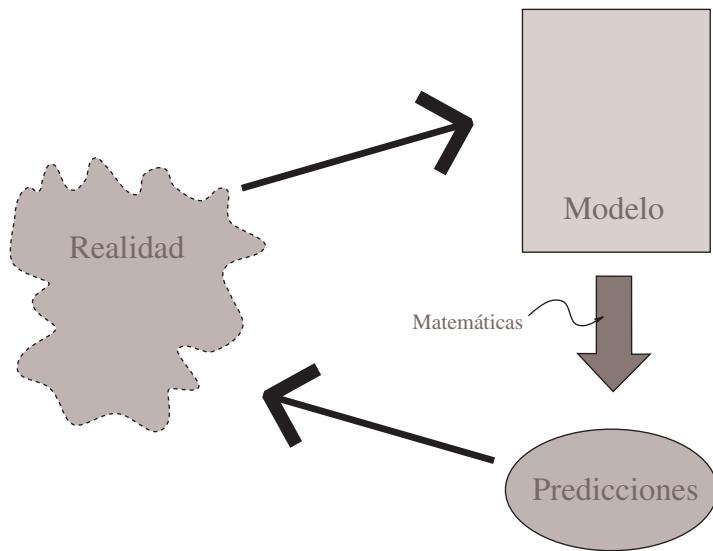


Figura 1.2: Las matemáticas nos permiten sacar conclusiones a partir de los modelos.

su importancia. Dada la imposibilidad de tomar en cuenta todos los aspectos de un objeto, se asume que ciertas variables son irrelevantes y se intenta identificar sólo las que producen el fenómeno que desea estudiarse. Esto es totalmente claro en ciertas ramas de las ciencias exactas, sobre todo aquéllas en las que ciertos fenómenos pueden ser descritos por una teoría simple. La labor del modelador es, en estos casos, escoger las variables relevantes y deducir de lo general a lo particular⁴.

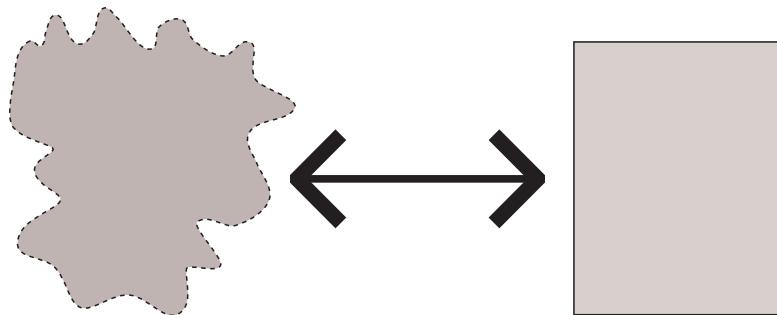


Figura 1.3: ¿Podemos representar a la realidad?

⁴Como un ejemplo clásico de esto, tenemos la teoría de la mecánica de fluidos. Las ecuaciones de los fluidos llevan ya muchos años de estar bien establecidas (ecuaciones de Navier-Stokes). El mayor problema es *resolverlas*, es decir sacar conclusiones a partir de ellas. En la práctica, lo que generalmente se hace es *desechar* la mayor parte de las variables y usar el resto en una versión simplificada de la teoría. Esta sería una típica aplicación del principio *ceteris paribus* en física. Desgraciadamente, en economía aún no existen teorías con la generalidad de las ecuaciones de Navier-Stokes. Por ahora, cada modelo en economía parecería que se hace *ad hoc*.

Así pues, el primer paso en el proceso de modelación matemática consiste en proponer una representación de esta realidad (véase la figura 1.3). Esta propuesta es arbitraria, pero generalmente (a partir de Galileo) se basa en la experiencia y la observación. La presunción de que podamos representar la realidad es una hipótesis fundamental y acaso atrevida. Einstein decía que lo más misterioso del mundo es que fuera inteligible.

En nuestro caso, los modelos económicos se dedican a representar algunos aspectos de lo que podríamos llamar la realidad económica. ¿En qué consiste ésta? Dejando todo lo demás igual, tradicionalmente se piensa que un modelo económico es una representación del mundo de las relaciones entre agentes que poseen bienes y preferencias, buscan un mayor bienestar y están dispuestos al intercambio. Los modelos son creados a partir de la especulación acerca de las causas y procesos que pudieron haber producido los efectos observados.⁵

Quizás quepa recordar el cuento de Borges⁶ acerca del imperio en el que la ciencia de la Cartografía avanzó tanto que los mapas eran casi perfectos.

Del Rigor en la Ciencia

... En aquel Imperio, el Arte de la Cartografía logró tal perfección que el mapa de una sola provincia ocupaba toda una ciudad, y el mapa del Imperio, toda una provincia. Con el tiempo, estos mapas desmesurados no satisficieron y los Colegios de Cartógrafos levantaron un mapa del Imperio, que tenía el tamaño del Imperio y coincidía puntualmente con él. Menos adictas al estudio de la Cartografía, las generaciones siguientes entendieron que ese dilatado mapa era inútil y no sin impiedad lo entregaron a las inclemencias del Sol y de los inviernos. En los desiertos del Oeste perduran despedazadas ruinas del mapa, habitadas por animales y por mendigos; en todo el país no hay otra reliquia de las Disciplinas Geográficas.

(Suárez Miranda: Viajes de Varones Prudentes, libro cuarto, cap. XIV, Lérida, 1658)

Cuando llegó un punto en el que se dieron a la tarea de crear el mapa perfecto, el resultado fue una réplica exacta: el mapa era el Imperio. Lo que es relevante de esta metáfora para nuestro problema, es el hecho de que todo mapa es imperfecto, o dicho de otro modo toda representación de la realidad necesariamente no puede tener todos los detalles. En particular, todo modelo matemático es imperfecto. Además una representación perfecta sería completamente inútil: la más inútil posible. El contraste con la figura 1.1 no podría ser más interesante. La representación de esta figura es inútil por no tener información. La representación de la realidad del mapa de Borges es inútil por tener toda la información.

Otra consecuencia de estas ideas es que, dado que un modelo sólo representa algunos aspectos del objeto, pueden existir distintos modelos de un mismo objeto. Pensemos en una vaca. Mucha gente puede estar interesada en estudiar este animal. Sin embargo, puede verse que según el punto de vista del sujeto que estudia a la vaca, se tendrán abstracciones distintas (véase la discusión en [CM85]). Por ejemplo, el economista la verá como una fuente de riqueza y el biólogo como un mamífero bovino. Aunque parezca absurdo, el estudio de las vacas ha tenido recientemente mucha relevancia desde el punto de vista de la economía (recordemos el problema de las vacas locas).

⁵Véase [LM75].

⁶Este cuento se encuentra en [Bor97].

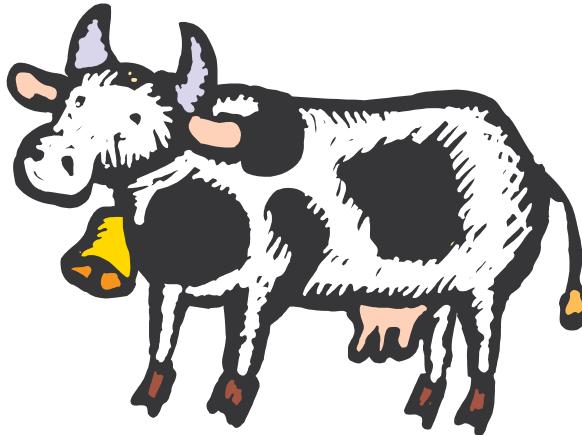


Figura 1.4: “Esto no es una vaca” diría René Magritte. Pueden utilizarse distintos puntos de vista para abstraer el concepto de “vaca”, según los objetivos del artista.

En un modelo matemático, la abstracción adecuada del objeto de estudio incluye la identificación y medición de los factores considerados relevantes. Asimismo, se tiene una conjetura de las relaciones que existen entre ellos y su comportamiento. Mediante el análisis matemático el modelo debe ser ahora capaz de producir conclusiones, y pasamos así al proceso deductivo.

§1.2.2 Deducción a partir del modelo

Lo primero que se debe tener en mente al obtener conclusiones referentes a nuestro modelo, es que éstas se refieren al modelo y no al fenómeno real que lo originó. El modelo es una representación simplificada del objeto de estudio y está construido con base en observaciones de ciertos factores. Éstas nunca son perfectas y normalmente sufren de errores de medición en mayor o menor grado. Es importante tomar en cuenta estas limitaciones antes de inferir cualquier cosa acerca del comportamiento de la realidad.

John Maynard Smith, en su libro “Mathematical Ideas in Biology” [MS71], comienza la introducción diciendo que el propósito de su libro es mostrar que “el razonamiento matemático es, en ocasiones, iluminador...”. A continuación inicia una defensa del uso de las matemáticas en biología, más allá de la estadística. Nos dice:

Es asumido en general -y en especial por los estadísticos- que la única rama de las matemáticas necesaria para la biología es la estadística. No comparto esta visión. La estadística es necesaria en biología, pues no existen dos organismos idénticos. Pero tengo la impresión de que la estadística, y en particular la rama que trata con las pruebas de significancia estadística, ha sido sobreestimada...

La misma opinión puede darse con respecto a la economía. Durante mucho tiempo se ha visto a la econometría como la rama de las matemáticas más directamente relacionada con los fenómenos

económicos. Reconocemos el inmenso valor de esta herramienta, pero quisieramos insistir que la cosa no termina ahí. Otras ramas de las matemáticas pueden utilizarse como herramienta para iluminar los procesos económicos, sobre todo aquéllos en los que el tiempo entra explícitamente como variable.

Dentro del esquema propuesto para los modelos matemáticos, de una manera burda podríamos localizar la econometría como la herramienta que nos permite medir las variables propuestas por el modelo. De este modo, se sitúa con mayor proximidad a la realidad y es parte tanto del proceso de modelación como del proceso de donde se obtienen predicciones. El resto de las matemáticas sirve para obtener conclusiones y predicciones del modelo en cuestión. La naturaleza de dichas conclusiones está limitada al modelo de donde surgen. Las conclusiones, no necesariamente son adecuadas a la realidad. De hecho, la calidad de éstas ha servido como parámetro para juzgar al modelo: mientras mejor sea el modelo mejor serán las conclusiones.

En principio, la deducción a partir del modelo podría ser hecha con otras herramientas. Algunos autores han tratado de minimizar la importancia de las matemáticas, relegándolas a un papel puramente sintáctico. Esto es, se pueden usar como un lenguaje abreviado, más no como un motor para la investigación. Puede pensarse que las ecuaciones no son más que abreviaciones de algo que ya se sabía de manera intuitiva.

Lo que está en discusión en este punto es, si las matemáticas pueden generar algo nuevo. El famoso economista Paul Krugman⁷ nos aporta algunas ideas a este respecto. Primero, reconoce que más allá de servir de lenguaje y prueba de consistencia, las matemáticas sirven, en muchos casos, como “fuentes de intuición”. Es decir, la investigación económica se lleva a cabo a través de una serie de modelos, cada vez más complicados, en los que las matemáticas van aportando claves e “intuiciones” cada vez más complejas. Sin embargo, nos advierte que esto algunas veces tiene un costo: el modelador puede convertirse en sirviente del modelo, y empezar a ver la realidad sólo con los ojos que convienen a la teoría. La crítica, en este caso, está en el hecho de que fácilmente nos podemos sesgar en favor de la simplicidad y elegancia teóricas.⁸

En conclusión, las matemáticas nos permiten ver lo que naturalmente no es obvio. Esto lleva ya varios siglos demostrándose en la práctica, sobre todo en las ciencias exactas. Pensemos, por ejemplo, en la teoría cuántica. Nuestra intuición no nos permite de ningún modo aproximarnos a la realidad descrita por esa teoría. Simplemente, debemos de creer en las matemáticas y darles una interpretación. Sin entrar en mayores detalles, podemos recordar que las diversas interpretaciones de la teoría produjeron un inmenso debate, al cual Einstein aportó la frase: “...no puedo creer que Dios juegue a los dados con el mundo”.

§1.2.3 Verificación, predicción y usos

En el paso definitivo del proceso de modelación, se confronta la conclusión puramente matemática con la realidad que se pretendía estudiar en un principio. Esto se puede hacer de distintas maneras. En el caso de las ciencias exactas, la verificación se da con base, ya sea en la predicción de comportamientos que pueden ser observados a través de experimentos controlados, o bien en la explicación de fenómenos observados para los cuales no existía tal explicación. De este modo, tradicionalmente se evalúa la calidad

⁷Véase, por ejemplo, el artículo [Kru98].

⁸¿No es un ejemplo de esto la insistencia en utilizar con abrumadora frecuencia las funciones de tipo Cobb-Douglas?

de un modelo según su habilidad de predecir y explicar correctamente otros hechos. Decimos que un modelo es **robusto**, si las conclusiones de éste no dependen del cumplimiento exacto de los supuestos, de no ser así se dice que el modelo es **frágil**.

Pensemos en las primeras aplicaciones de las teorías de Galileo. Las observaciones del sabio italiano le llevaron a concluir que un buen modelo para el movimiento de una bala de cañón implicaba necesariamente que la trayectoria de ésta fuera una curva parabólica. Esto tuvo aplicaciones inmediatas en el arte de la guerra: se podía predecir con cierta exactitud en dónde caería una bala. Otras conclusiones implícitas del modelo, imponen un alcance máximo a cada arma. Aunque estos hechos nos parecen obvios, no lo eran del todo en la época de Galileo.

Isaac Newton fue realmente el iniciador de la Física como la conocemos hoy en día. A partir de principios básicos que son bastante cercanos a nuestra intuición –las tres leyes de movimiento y la ley de gravedad– logró crear una teoría del universo que durante muchos años se pensó definitiva. El momento cumbre de esta nueva teoría, el punto en el que todos quedaron convencidos de su valor, quizás haya sido aquel en el que las leyes de Kepler surgieron como una consecuencia. Kepler había pasado muchos años intentando describir las trayectorias de los planetas observables en su época. Un colega suyo, el astrónomo Tycho Brahe le facilitó los datos de las observaciones que había realizado durante años. La conclusión de Kepler, de que todas esas observaciones tuvieran algo en común, es un hecho sorprendente. Más sorprendente aún es el hecho de que Las leyes de Kepler estén implícitas en las leyes de Newton.

Se podrían dar muchos otros ejemplos del éxito del poder de predicción de las ciencias exactas. ¿Qué se puede decir de la Economía? La Economía ha procedido de una manera similar a la de las ciencias exactas. La literatura está llena de modelos que pretenden representar la realidad económica. Sin embargo, si nos basamos únicamente en la capacidad de predicción, los modelos económicos aún se encuentran en su infancia pues su capacidad predictiva es limitada; sin embargo, un buen modelo debe poder, sino predecir, al menos explicar adecuadamente el fenómeno en cuestión. En este punto podemos recordar las críticas de Velasco [Vel96] a los economistas. Para él, la profesión es vista bajo sospecha y tiene muchos aspectos aún sin resolver, desde los ideológicos hasta los prácticos y éticos. No obstante, insiste en el hecho de que la Economía –a pesar de su inadecuación con la realidad– es de suma importancia y está en plena construcción.

Las razones por las cuales los modelos económicos no tienen la objetividad y capacidad de predicción que aquéllos en ciencias exactas son claras: los fenómenos que se estudian son, no sólo complejos, sino muy difíciles de aislar. La actividad económica se desarrolla dentro de un marco legal, técnico, social y político que evoluciona constantemente. No puede ignorarse los efectos que esto tiene sobre los fenómenos económicos, sin embargo, difícilmente podemos cuantificar estos efectos. En las ciencias exactas existe la capacidad de llevar a cabo experimentos para verificar los modelos propuestos. En economía casi nunca se pueden realizar experimentos. Cabe mencionar que la excepción es dentro de la microeconomía, específicamente dentro del área del comportamiento humano. Desde hace unos veinte años, la llamada **economía experimental** ha puesto a prueba el comportamiento del *homo economicus* en condiciones controladas que simulan situaciones reales. Se ha concluido que, por ejemplo, los modelos clásicos de oferta y demanda son sumamente robustos, más no así el modelo de la utilidad esperada.

En economía se corre el grave peligro de creer que los modelos matemáticos necesariamente funcionan

como si fueran leyes de la física. Como mencionamos antes, la economía está lejos de ser una ciencia exacta y su capacidad de predicción es limitada. Por esta razón debe tenerse mucha cautela al tratar de imponer a la realidad las ideas surgidas a partir de los modelos económicos.

El uso de las conclusiones de un modelo ciertamente constituye un **problema ético**. ¿Está la economía más allá de la ética? ¿Cómo se debe usar lo que nos dicen los modelos económicos? En un contexto distinto, Feynman se hacía la pregunta de si la Ciencia –y en nuestro caso la economía– tiene algún valor.

...una vez en Hawaii fui llevado a un templo budista. En el templo, un hombre me dijo: “te voy a decir algo que nunca olvidarás”. A continuación me dijo lo siguiente: “a cada hombre se le dan las llaves de la puerta del Cielo. Las mismas llaves abren el Infierno.”

Es lo mismo con la Ciencia. Es, en cierto modo, una llave de las puertas del Cielo y la misma llave abre el Infierno y no tenemos las instrucciones de qué puerta es cuál. ¿Debemos desechar la llave y nunca entrar al Cielo? O mas bien, ¿no deberíamos luchar con el problema de cuál es el mejor modo de usar la llave? Esta es, claro está, una pregunta bastante seria, pero pienso que no podemos despreciar el valor de las llaves de las puertas del Cielo.

Los distintos pasos que quedan incluidos dentro de la modelación matemática, también expresan la idea de que el científico recibe algo llamado *su formación profesional*. La idea de la formación profesional está plenamente vinculada a los pasos del proceso de modelación. La formación pretende hacer que el individuo pueda ver el mundo de una cierta manera. Constituye, en definitiva, la capacidad de ver lo esencial de la realidad desde el punto de vista de su propia disciplina.⁹ Por eso, sin temor a equivocarnos, podemos decir que la mejor manera de entender lo que es un modelo matemático es en la práctica. La modelación matemática, más que una ciencia, es un arte y podemos afirmar que en un futuro se reconocerá que las matemáticas utilizadas en economía son tan o más complicadas que las utilizadas en las ciencias exactas.

Quisiéramos mencionar, para terminar, un uso poco explorado del proceso de modelación. Este es el poder usar el modelo y, en particular las conclusiones de éste, de una manera más que nada cualitativa.

Pensemos por un momento en la teoría de la evolución. Es ya clásica la aventura que trajo como consecuencia la genial teoría de Darwin. Nos sería quizás muy difícil poder ver la innovación de Darwin con suficiente perspectiva. Por primera vez se podía ver al mundo biológico como algo en desarrollo. No sólo eso, el mecanismo mediante el cual esto ocurre se explica de manera simple: la lucha por la vida y la supervivencia del más fuerte. Si nos detenemos a pensar un momento, la teoría de Darwin no predice nada en concreto. No nos dice, por ejemplo, cuando aparecerá una nueva especie de hormiga o cuanto tiempo pasó para que un humano dejara de parecerse a sus ancestros. Lo importante de la teoría de Darwin es que propuso un mecanismo y una explicación cualitativa para un sinnúmero de fenómenos.

La aportación no termina ahí. La verdadera importancia de Darwin fue que la idea de evolución se filtró en casi todos los ámbitos de la vida. Baste recordar que dentro del mundo de los negocios, se utiliza precisamente el lenguaje evolutivo: adaptación de la empresa a nuevos retos, supervivencia del más fuerte, cambio de estructuras, etc.

En cuanto a lo que nos interesa podríamos decir, acaso de manera atrevida, que el verdadero uso de los modelos es el de servir de metáforas de fenómenos de los que se puede decir poco. Esto es, el objetivo es

⁹Como antídoto a esto, muchos han propuesto recurrir a la “interdisciplinariedad”.

crear conceptos con tal fuerza que puede explicar muchos fenómenos, aunque esto no se haga de manera cuantitativamente exacta. Esto es de particular importancia en economía, pues la mayor parte de los modelos son de índole cualitativa.

§1.3 El uso del tiempo en economía

El proceso de matematización de la economía inicia con modelos estáticos. No fue sino hasta los últimos 50 años que se comenzaron a utilizar, de manera estándar, los modelos en los que el tiempo es una variable. Naturalmente, los modelos más interesantes necesariamente incluyen la variable temporal. La variable tiempo ha sido utilizada de distintas maneras y podemos identificar dos posibilidades: tiempo discreto y tiempo continuo. El tiempo discreto corresponde a la idea de períodos: primer año, segundo año, etc. El tiempo continuo incluye al tiempo discreto y a todos los instantes intermedios.

El tiempo discreto se ajusta de una mejor manera a la realidad económica, pues en general se ha utilizado la idea de períodos económicos. Sin embargo, el tiempo continuo es una abstracción que ha demostrado su utilidad aproximando procesos en donde la evolución de una situación a la siguiente no es muy brusca, y además se tiene la posibilidad de hablar de los estados intermedios. Dentro de la ciencia, esta idea tiene una larga tradición. Ya dentro de los trabajos de Newton, se encuentran muchas de las bases de la teoría que describen la evolución temporal. Incluso la notación, la manera “arbitraria” de escribir matemáticas, es una herencia de los primeros trabajos de Newton y de Leibniz en el siglo XVII. De hecho, algunos de los problemas que hoy en día se estudian, son problemas que vienen desde ese inicio. El uso generalizado de tiempo discreto no fue sino hasta el advenimiento de la tecnología computacional, la razón es que involucran procesos de iteración que pueden ser sumamente tediosos.

Con riesgo de caer en una sobre simplificación, podríamos decir que los modelos de tiempo continuo, y por lo tanto las matemáticas que los describen, están mejor adaptados a procesos con una evolución regular (como la mayoría de los procesos físicos). Pensemos, por ejemplo, en el recorrido de un planeta. La órbita se puede describir con una trayectoria continua que a cada instante de tiempo se encuentra en una posición determinada. Cabe reconocer ésta como una de las ideas más antiguas de la ciencia. De hecho, el desarrollo del cálculo y del análisis matemático tiene su origen, entre otros, en el problema de cuantificar el grado de variación de las variables de los distintos modelos físicos que iniciaron con la astronomía. De ahí surgieron aplicaciones a toda la ciencia.

En el caso de los modelos surgidos de las ciencias económicas, el uso del tiempo continuo es, en general, una abstracción que no surge de manera natural. Quizás por ello, se pospuso el uso del tiempo como variable y se inició con modelos estáticos. El modelo clásico de la microeconomía, basado en agentes, preferencias y alocaciones, es un modelo esencialmente estático. Con un conjunto de hipótesis básicas sobre los agentes, se llega a la conclusión de la existencia de un equilibrio entre las fuerzas económicas. Sin embargo, la conclusión fundamental de este modelo tiene la grave deficiencia de no describir el cómo se llega al equilibrio, es decir, podría darse el caso de que el modelo prediga la existencia de un equilibrio inestable al que nunca se llegará.

La rama de las matemáticas que describe la evolución temporal de las variables es la teoría de los **sistemas dinámicos**. Esta teoría toca una gran diversidad de aspectos de las matemáticas, unidos por un común denominador: la idea de cambio o evolución. Este cambio, generalmente se piensa como un

cambio temporal y siempre se puede conceptualizar como un proceso que evoluciona desde un pasado y hacia un futuro. La pregunta, hablando de una manera muy general, es acerca de lo que sucederá en el futuro.

La teoría de los sistemas dinámicos tiene su origen en los trabajos del matemático francés Henri Poincaré, que en 1890 publica una primera versión de su artículo “Sur le Problème des Trois Corps”. No es sino hasta los años sesenta del siglo pasado, que varios autores vuelven a descubrir muchas de las ideas de Poincaré y con ellas se crea formalmente el área de los sistemas dinámicos dentro de las matemáticas. En el transcurso de la siguiente década, estas ideas se popularizan entre una gran cantidad de científicos y matemáticos. El advenimiento de las computadoras ha permitido explorar más profundamente los conceptos teóricos y las observaciones numéricas producen, a su vez, nuevos resultados teóricos.

La popularización de algunos de estos resultados originó la **teoría del caos**. La llamada teoría del caos es, en realidad, sólo un subconjunto de resultados de la teoría de sistemas dinámicos. En nuestra opinión, no es una verdadera revolución científica ni un cambio de paradigma. El concepto de “caos” ha sido utilizado erróneamente por muchos autores y esto ha llegado a filtrarse en las ciencias sociales. En particular, en muchos casos se ha utilizado el término “caos” de manera poco seria. El “caos matemático” no se refiere a la acepción común, de caos en la vida cotidiana. Recientemente, Sokal y Bricmont [SB99] han denunciado ésta y otras confusiones que intentan crear o justificar una visión equivocada de la ciencia.

Un sinnúmero de aplicaciones han demostrado la utilidad de la teoría; sin embargo, al popularizarse las ideas relacionadas con lo caótico también se ha dado origen a una serie de abusos y confusiones que han perjudicado la difusión de las verdaderas ideas. Dada la pretensión de certeza de las matemáticas, el teorema de Gödel y el descubrimiento de algunos fenómenos matemáticos que de cierto modo son impredecibles ocasionó gran malestar dentro de la comunidad que los descubrió. La falta de capacidad de predicción viene, en este caso, no de la inhabilidad del modelador matemático, sino de las ecuaciones mismas.

Es de suma importancia que el modelador económico tenga plena conciencia de que la falta de predictibilidad de sus modelos puede tener dos orígenes muy distintos. El primero, que las variables del modelo sean intrínsecamente estocásticas y, por lo tanto, su evolución es impredecible hasta cierto punto. El segundo, que las variables involucradas sean de origen determinista, pero su evolución está descrita por un sistema dinámico no lineal que presenta comportamiento caótico. En el capítulo 6 damos una breve introducción a este concepto.

§1.4 Un ejemplo

El siguiente ejemplo de modelo económico no dinámico es sumamente sencillo desde el punto de vista matemático; sin embargo, muestra los pasos principales del proceso típico de modelación. Seguiremos los pasos esbozados en las secciones anteriores teniendo en cuenta que, en general, no todos los modelos se hacen del mismo modo. Sería un error, por lo tanto, insistir en ser exhaustivos en la descripción del proceso de modelación. Sin embargo, podemos reconocer que casi siempre se hace algo parecido a lo que haremos como ejemplo. Recordemos que no siempre las matemáticas son números o fórmulas: el razonamiento matemático consiste más que nada en conceptos.

Un modelo clásico en economía es el que describe los efectos del salario mínimo en el mercado laboral.

El salario mínimo es una cota inferior que se impone al salario de manera que, legalmente, el empleador no puede pagar un salario menor al mínimo. En la práctica existen salarios mínimos diferenciados, es decir, los salarios mínimos se asignan según el tipo de actividad productiva, región geográfica, etc. El problema que se quiere tratar es el de analizar los efectos de la existencia del salario mínimo sobre los trabajadores (empleados y desempleados). El propósito del modelo no es hacer predicciones cuantitativas, sino describir el comportamiento cualitativo de las variables relevantes.

§1.4.1 Abstracción del mundo

Este proceso lo vamos a subdividir en varios pasos que no necesariamente siguen un orden cronológico. Lo importante es que entendamos que, en la abstracción de la realidad, necesariamente hacemos simplificaciones. Para abstraer, se hacen supuestos acerca del fenómeno a estudiar que capturen, de la manera más simple posible, el comportamiento observado o la hipótesis de trabajo.

Variables relevantes y notación

Haciendo una gran simplificación, se identifican como variables relevantes en el mercado laboral, el salario (el que existiría de no existir el salario mínimo), el salario mínimo, el número de trabajadores que están dispuestos a trabajar por un salario determinado (oferta laboral) y el número de trabajadores que las empresas están dispuestas a contratar a un salario dado (demanda laboral). Nótese que, como en todos los modelos económicos, estamos utilizando el principio *ceteris paribus*, pues se mantienen muchas otras variables como constantes. En particular, el tiempo no transcurre.

Utilizamos la siguiente notación:

- w = salario,
- w_{\min} = salario mínimo y
- L = número de trabajadores.

Supuestos

- Tomamos un sector “representativo” de la economía, digamos el sector manufacturero.
- El mercado es competitivo en el siguiente sentido: tanto trabajadores como empleadores toman al salario como dado.
- El salario es la variable independiente, la oferta y demanda laboral las dependientes.
- El salario mínimo está dado de forma externa (por alguna autoridad gubernamental, por ejemplo).
- Existe un equilibrio en el mercado de trabajo: cuando la oferta y la demanda laborales coinciden.

Relaciones entre las variables

- La oferta laboral se puede representar por una función del salario, continua y creciente, digamos $L = f(w)$, y
- la demanda laboral por una función continua y decreciente, por ejemplo, $L = g(w)$.
- Si w^* es el salario de equilibrio existente, se cumple $w^* > w_{\min}$ ya que el salario mínimo es una cota inferior al salario.

Nótese que el buen comportamiento de las funciones es una enorme simplificación. En muchas ocasiones estas simplificaciones (por ejemplo de continuidad, diferenciabilidad o convexidad) son naturales dentro de una familia de modelos, aunque en realidad se hacen con el objetivo de hacer al problema manejable.

§1.4.2 Deducción a partir del modelo

El equilibrio en el mercado laboral se obtiene al intersecciar las funciones de oferta y demanda laborales. Las cantidades de equilibrio, obtenidas cuando $f(w) = g(w)$, son w^* y L^* . En este equilibrio no existe desempleo involuntario (ningún trabajador que deseé trabajar a este salario está desempleado). Al imponer un salario mínimo mayor a w^* , la oferta laboral aumenta a $L_1 > L^*$ y la demanda disminuye a $L_2 < L^*$. La cantidad $D = L_1 - L_2$ de trabajadores serán desempleados involuntarios. De estos, $L^* - L_2$ perdieron su empleo y $L_1 - L^*$ ahora desean trabajar por el salario mínimo. La cantidad L_2 de trabajadores que conservan su empleo incrementa su salario a w_{\min} . La figura 1.5 representa estas consideraciones. Esta será la única ocasión en la cual el precio (el salario, en este caso) se presente en el eje horizontal y la cantidad en el vertical. En economía normalmente se consideran las funciones inversas de manera que los ejes se invierten.

§1.4.3 Verificación y conclusiones

El modelo propuesto indica que la imposición de un salario mínimo aumenta las percepciones de algunos trabajadores a costa de un incremento en el desempleo. Una observación importante es que la magnitud de estos efectos depende de las pendientes de las curvas de oferta y demanda. La situación presentada hace una gran cantidad de simplificaciones, entre otras cosas, supone que el mercado de trabajo es como el de cualquier otro bien, que carece de imperfecciones, que existe el equilibrio, las ganancias y pérdidas no están cuantificadas, no nos ocupamos de la obtención formal de las funciones de oferta y demanda, etc. A pesar de que existe un gran número de críticos al modelo y a sus conclusiones, la mayor parte de los economistas lo aceptan como una representación adecuada de la realidad.¹⁰

¹⁰Como ejemplo, recordemos que en 1995 causó gran revuelo una publicación de los economistas David Card y Alan Krueger (véase [CK95]). En ella se realizó un estudio en las cadenas de restaurantes de comida rápida en New Jersey y Pennsylvania, concluyendo que el aumento al salario mínimo no tenía efecto en el empleo. El estudio fue utilizado con fines políticos por la administración del presidente Clinton. Las críticas no se hicieron esperar y los métodos estadísticos utilizados en el estudio han sido puestos bajo sospecha.

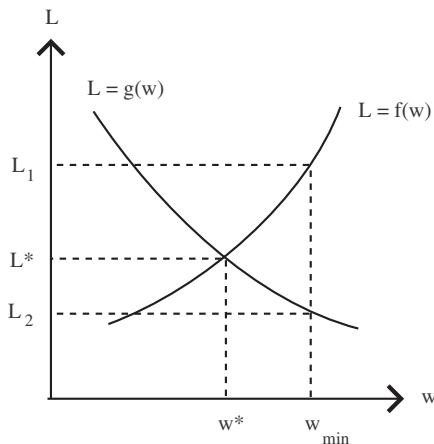


Figura 1.5: Las cantidades w^* y L^* representan el equilibrio del mercado de trabajo. Al imponer el salario mínimo $w_{\min} > w^*$, se crea una diferencia entre la cantidad ofrecida y demandada de trabajo.

Ejercicios

▷ 1.1 Los siguientes escenarios describen vagamente fenómenos hipotéticos que se desean estudiar. Para cada caso, identificar un problema específico, las variables relevantes, proponer relaciones entre ellas y justificar. (Evidentemente las respuestas pueden variar.)

- ¿Es el momento de comprar un auto nuevo?
- La población de perros callejeros en la ciudad de México es alarmante.
- Se desea sustituir los “microbuses” urbanos por autobuses eficientes operados por el gobierno de la ciudad.
- La Secretaría de Salud se interesa por conocer la eficiencia de una nueva vacuna contra el sarampión.
- Nos preguntamos si debemos ahorrar una porción de nuestro ingreso.

▷ 1.2 El fenómeno de **difusión social** (*o proceso del chisme*) describe la difusión de información dentro de una sociedad. Los miembros de la población se dividen en dos clases: los que poseen la información y los que no la poseen. En una población conocida (digamos de N individuos), es razonable suponer que la tasa de difusión de la información es proporcional al producto del número de individuos que poseen la información por aquéllos que no la tienen. Nombrar las variables relevantes y describir un modelo que proporcione el número de individuos que poseen la información después de n días.

▷ 1.3 Se ha observado que el ahorro de los individuos depende de su ingreso y de la tasa de interés del mercado. Asimismo, la cantidad de ahorro es sumamente sensible a los cambios en el ingreso, pero cambia poco ante variaciones de la tasa de interés. Describe un modelo que represente estas consideraciones.

Parte II

Ecuaciones Diferenciales y Dinámica Continua

CAPÍTULO 2

Ecuaciones diferenciales lineales

§2.1 Introducción

Dada una función f que toma valores reales, una expresión de la forma $f(x) = 0$ es lo que conocemos como una ecuación. Resolverla implica encontrar un valor de la variable x en el dominio de la función f , normalmente un número real o complejo, de tal forma que se satisfaga dicha ecuación. La función f puede tomar distintas formas. Si f es de la forma $f(x) = mx + b$, decimos que la ecuación es lineal, si f es un polinomio cuadrático decimos que la ecuación es cuadrática, etc.

Si f_1, \dots, f_m es una colección de funciones reales cuyo dominio es un subconjunto de \mathbb{R}^n , se pueden considerar también sistemas de ecuaciones; es decir, una lista de ecuaciones

$$\begin{aligned}f_1(x_1, x_2, \dots, x_n) &= 0, \\f_2(x_1, x_2, \dots, x_n) &= 0, \\&\vdots \\f_m(x_1, x_2, \dots, x_n) &= 0.\end{aligned}$$

Una solución a este sistema es un vector

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

que satisface las m ecuaciones simultáneamente. Una vez más, las funciones f_i , $i = 1, 2, \dots, m$, pueden tomar diversas formas. La solución puede no existir y puede no ser única.

Sobra decir que el problema de resolver ecuaciones surge constantemente en casi cualquier área del conocimiento en la que se empleen matemáticas. En economía, encontrar la solución a un sistema de ecuaciones (en particular puede ser un sistema de una sola ecuación con una sola variable) está frecuentemente asociado con encontrar puntos de equilibrio o bien puntos que resuelven algún problema de optimización.

mización. Se obtiene así, por ejemplo, la cantidad de cierto artículo que iguala la oferta con la demanda, o bien se puede obtener la cantidad óptima de insumos que maximiza alguna función de beneficios.

Asimismo, mediante el uso del teorema de la función implícita se puede determinar cómo cambian estos equilibrios, o valores óptimos, al modificar ciertos parámetros exógenos. De esta forma, puesto que el mundo del economista está lleno de puntos de equilibrio y valores óptimos, esto sería suficiente para cualquier tipo de análisis si tales procesos no se vieran afectados por el paso del tiempo.

Cuando un sistema económico pasa de un punto de equilibrio a otro normalmente no lo hace de forma instantánea. El tiempo transcurre y existe una trayectoria definida entre ambos puntos de equilibrio. El equilibrio puede también cambiar de manera definida a lo largo del tiempo; por ejemplo, ciertas variables económicas pueden crecer exponencialmente en el tiempo a una tasa constante originando el crecimiento económico. Una implicación inmediata es que la forma en la que un sistema se aproxima a un equilibrio resulta de suma importancia.

El estudio de variables que son dependientes del tiempo es algo fundamental en este libro. Así, el consumo c puede considerarse como una trayectoria de consumo $c(t)$ a lo largo del tiempo y lo mismo sucede con otras variables. Se consideran ecuaciones que, no sólo contienen las variables, sino también su tasa de cambio (instantánea) en el tiempo y/o la tasa de cambio de esta última y así sucesivamente. A este tipo de ecuaciones se las llama **ecuaciones diferenciales**.

De aquí en adelante, dada cualquier variable que es función del tiempo, digamos $x(t)$, denotaremos a su derivada $x'(t)$ o $\frac{dx(t)}{dt}$, por $\dot{x}(t)$. Abusando de la notación, omitiremos al tiempo como variable independiente, por lo que utilizaremos simplemente x , \dot{x} , \ddot{x} , etc. Obviamente, el tiempo podría ser cualquier otra variable independiente, como se verá en algunos ejemplos, en cuyo caso utilizaremos la notación usual para las derivadas.

Definición 2.1.1 Una **ecuación diferencial ordinaria de primer orden** es una ecuación de la forma $F(\dot{x}, x, t) = 0$. Si la ecuación no involucra explícitamente al tiempo t , decimos que es una **ecuación autónoma**. En general, una **ecuación diferencial ordinaria de orden n** es una ecuación de la forma $G(x^{(n)}, \dots, \ddot{x}, \dot{x}, x, t) = 0$ que involucra a las n primeras derivadas de la función x . Finalmente, si la ecuación no posee término independiente se dice que es una **ecuación homogénea**.

Si x es una función de varias variables, por ejemplo $x(s, t)$, entonces la ecuación diferencial involucra a las derivadas parciales de x . En este caso, se dice que tenemos una **ecuación diferencial parcial** (en vez de ordinaria). Este tipo de ecuaciones no será abordado en este texto, aunque son de gran importancia, especialmente en algunas áreas de finanzas y economía financiera.

Ejemplos

Las siguientes son ecuaciones diferenciales.

Ej 2.1.1 $\dot{x} = 5x + t$ (lineal de primer orden, no autónoma, no homogénea).

Ej 2.1.2 $\dot{x} = 2x$ (lineal de primer orden, autónoma y homogénea).

Ej 2.1.3 $\dot{x} = t^2x$ (lineal de primer orden, no autónoma y homogénea).

Ej 2.1.4 $\ddot{x} - 3\dot{x} + 2x - 2 = 0$ (lineal de segundo orden, autónoma y no homogénea).

Ej 2.1.5 $\dot{x} - x^2 - t^2 = 0$ (no lineal, de primer orden, no autónoma, no homogénea).



Se puede observar que en los cuatro primeros ejemplos la función x y sus derivadas aparecen en forma lineal. A este tipo de ecuaciones se las llama **ecuaciones diferenciales lineales**. Estas ecuaciones siempre tienen una solución que podemos encontrar de forma explícita. El último ejemplo es de una ecuación de primer orden no lineal, que a pesar de su apariencia inofensiva no tiene solución explícita.

§2.2 Ecuaciones de primer orden

Las ecuaciones diferenciales más simples son las lineales de primer orden. A continuación procedemos a encontrar soluciones para los diversos tipos de estas ecuaciones.

§2.2.1 Caso autónomo

La ecuación diferencial autónoma más sencilla que podemos encontrar es de la forma

$$\dot{x} = ax + b, \quad (2.1)$$

donde a y b son constantes con $a \neq 0$. Si $b = 0$, decimos que la ecuación es homogénea. Sea x_h la solución general de la ecuación homogénea asociada a (2.1), es decir, la ecuación que obtendríamos al asumir $b = 0$, y x_p cualquier solución particular de (2.1). La siguiente proposición nos dice que cualquier solución es la suma de estas dos.

Proposición 2.2.1 *Consideremos la ecuación de orden n con coeficientes constantes*

$$a_n x^{(n)} + a_{n-1} x^{(n-1)} + \dots + a_0 x = h(t),$$

entonces x es una solución a esta ecuación si y sólo si $x = x_h + x_p$, en donde x_h es la solución general de la ecuación homogénea asociada y x_p es cualquier solución particular de la ecuación original.

Demostración

Por un lado, si x es solución, entonces $x_p = x - x_h$ es una solución particular. Simplemente hay que ver que satisface la ecuación como sigue:

$$\begin{aligned} a_n x_p^{(n)} + a_{n-1} x_p^{(n-1)} + \dots + a_0 x_p &= \\ a_n(x^{(n)} - x_h^{(n)}) + a_{n-1}(x^{(n-1)} - x_h^{(n-1)}) + \dots + a_0(x - x_h) &= \\ (a_n x^{(n)} + a_{n-1} x^{(n-1)} + \dots + a_0 x) - (a_n x_h^{(n)} + a_{n-1} x_h^{(n-1)} + \dots + a_0 x_h) &= \\ h(t) - 0 &= h(t). \end{aligned}$$

De aquí que $x_p = x - x_h$ sea una solución particular. Por otro lado, si $x = x_h + x_p$, entonces es solución de la ecuación puesto que

$$\begin{aligned} a_n x^{(n)} + a_{n-1} x^{(n-1)} + \dots + a_0 x &= \\ a_n(x_h^{(n)} + x_p^{(n)}) + a_{n-1}(x_h^{(n-1)} + x_p^{(n-1)}) + \dots + a_0(x_h + x_p) &= \\ (a_n x_h^{(n)} + a_{n-1} x_h^{(n-1)} + \dots + a_0 x_h) + (a_n x_p^{(n)} + a_{n-1} x_p^{(n-1)} + \dots + a_0 x_p) &= \\ 0 + h(t) &= h(t). \end{aligned}$$

Con esto queda demostrada la proposición. ■

Es claro que podemos aplicar la proposición anterior a la ecuación (2.1) ya que se trata de un caso particular en donde $h(t) = b$, $n = 1$, $a_0 = -a$ y $a_1 = 1$.

Ejemplos

Ej 2.2.1 Considérese la ecuación dada por $\dot{x} = 2x$. Es posible que el lector recuerde este ejemplo de algún curso anterior. Se trata de una ecuación homogénea y su solución se obtiene fácilmente de la siguiente forma: se reescribe (asumiendo $x \neq 0$) como

$$\frac{\dot{x}}{x} = 2.$$

Al integrar ambos lados con respecto a t , tenemos

$$\int \frac{\dot{x}}{x} dt = \int 2 dt,$$

y de aquí

$$\ln x = 2t + k.$$

Aplicando la exponencial se tiene

$$x = e^k e^{2t},$$

y definiendo $K \equiv e^k$ queda finalmente

$$x(t) = K e^{2t}.$$

Notemos que $K > 0$; sin embargo, $K e^{2t}$ es solución aun cuando $K < 0$. La constante K queda determinada si algún valor inicial de x está dado. Típicamente conocemos el valor $x(0) = x_0$, así que poniendo $t = 0$ tenemos que $K = x_0$ y por lo tanto $x(t) = x_0 e^{2t}$. Este es el modelo clásico de crecimiento a una tasa constante, igual a 2 en este ejemplo.

Ej 2.2.2 Tomemos la ecuación, $\dot{x} = 2x + 4$. La solución es la suma de la solución general de la ecuación homogénea asociada que obtuvimos en el ejemplo anterior y de cualquier solución particular de la ecuación dada. La solución particular más sencilla es una función constante, es decir, una en la que se cumple $\dot{x} = 0$, con lo cual se tiene $x_p = -2$. Así pues, la solución general de la ecuación es

$$x(t) = -2 + K e^{2t}.$$

De nuevo, podemos obtener el valor de K si conocemos un valor inicial, por ejemplo $x(0) = x_0$. De esta forma, $K = x_0 + 2$ y $x(t) = -2 + (x_0 + 2)e^{2t}$.

Es claro de los dos ejemplos anteriores que la solución general a la ecuación $\dot{x} = ax + b$ está dada por

$$x(t) = -\frac{b}{a} + Ke^{at} \quad (2.2)$$

y K puede ser determinada con un valor inicial, $x(0) = x_0$, como $K = x_0 + \frac{b}{a}$. Si $a = 0$, la solución particular no puede ser calculada por este método, pero en tal caso la ecuación sería simplemente $\dot{x} = b$ que se resuelve por integración directa.

Ej 2.2.3 Consideremos el siguiente modelo de crecimiento económico. Como en todo buen modelo, debemos escoger un conjunto de variables adecuadas y un conjunto de relaciones entre éstas. La idea es utilizar el menor número de variables, de tal manera que el fenómeno que se quiere estudiar pueda ser representado adecuadamente. La utilización de modelos se convierte de este modo en un proceso de abstracción de la realidad. En el siguiente modelo se describe la relación entre el ahorro la inversión y el ingreso, llegando a una ecuación que describe la evolución del ingreso en el tiempo.

Sean S = ahorro, Y = ingreso, I = inversión y K = capital y s y v dos constantes positivas menores que 1. Supongamos que el ahorro y la inversión son proporciones fijas del ingreso y de su tasa de cambio, respectivamente, de manera que se cumplen las siguientes ecuaciones:

$$\begin{aligned} S &= sY, \\ I &= \dot{K} = v\dot{Y}. \end{aligned}$$

En equilibrio, $S = I$, por lo que se tiene

$$v\dot{Y} = sY,$$

o bien

$$\dot{Y} = \frac{s}{v}Y,$$

y por lo tanto

$$Y = Y_0 e^{(\frac{s}{v})t}.$$

Es decir, dadas las hipótesis del modelo, concluimos que el ingreso crece de manera exponencial y la tasa de crecimiento es la razón $\frac{s}{v}$. A la constante s se la llama **propensión marginal a ahorrar** y a v se la llama **coeficiente de inversión**. En este modelo, estos coeficientes son constantes y no existe fundamento microeconómico alguno que justifique por qué los individuos se comportan de esta manera. Más adelante veremos modelos más elaborados.

Ej 2.2.4 Supongamos que se tienen ingresos constantes y se quiere ajustar el consumo a un nivel C^* . No queremos tener “saltos” y $C(t)$ representa el nivel de consumo al tiempo t . Supongamos que se tiene un mecanismo mediante el cual se puede controlar el cambio en el consumo, es decir, se puede escoger una función $f(C)$ tal que

$$\dot{C} = \frac{dC}{dt} = f(C).$$

Imponemos el siguiente modelo: el cambio instantáneo $\frac{dC}{dt}$ es positivo si $C < C^*$ y negativo si $C > C^*$. Esto se puede lograr si escogemos a f como una función continua y decreciente, tal que $f(C^*) = 0$. Una función así tiene una gráfica como la que se ilustra en la figura 2.1.

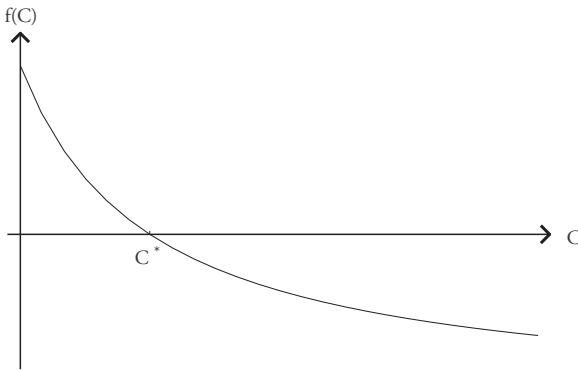


Figura 2.1: Una posible función de consumo.

En particular, podemos utilizar una función proporcional a la diferencia $C^* - C$. Por ejemplo,

$$\dot{C} = f(C) = \alpha(C^* - C),$$

en donde $\alpha > 0$. La ecuación diferencial correspondiente es la ecuación diferencial lineal

$$\dot{C} = -\alpha C + \alpha C^*.$$

La solución, según lo visto anteriormente (ver ecuación 2.2), está dada por

$$C(t) = C^* + K e^{-\alpha t} = C^* + (C_0 - C^*) e^{-\alpha t}.$$

En este momento vale la pena verificar si al escoger f como lo hemos hecho, llegamos a la situación deseada. Tomando en cuenta que α es positiva, se obtiene

$$\lim_{t \rightarrow \infty} C(t) = C^*,$$

que es lo que se quiere. La figura 2.2 ilustra la trayectoria del consumo cuando $C_0 < C^*$.

◊◊◊◊◊◊◊◊◊◊◊◊◊

En este último ejemplo el sistema converge a un valor, C^* , para el cual se tiene que $\dot{C} = 0$, o bien, en donde ya no existe ningún ajuste dinámico. La siguiente definición precisa este concepto.

Definición 2.2.2 *Dada la ecuación*

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

un valor x^ que satisface $f(x^*) = 0$ se llama **punto fijo**, **estado estacionario** o **punto de equilibrio**.*

Adicionalmente, si existe un número $\varepsilon > 0$ de tal forma que se cumplen,

- a) *$x(t)$ está definida para todo $t \geq 0$, y*

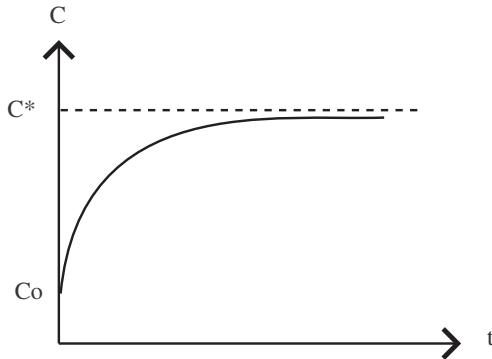


Figura 2.2: Trayectoria para el consumo.

- b) $|x_0 - x^*| < \varepsilon \Rightarrow \lim_{t \rightarrow \infty} x(t) = x^*$, es decir, la solución tiende al estado estacionario cuando $t \rightarrow \infty$ siempre que el valor inicial esté suficientemente cerca¹ de x^* , entonces se dice que el equilibrio es **asintóticamente estable** o simplemente **estable**. De no ser así, decimos que es **inestable**. Es claro que no existe ningún punto fijo si la ecuación $f(x) = 0$ no tiene solución.

Para la ecuación (2.1) es fácil verificar que el estado estacionario es igual a la solución particular, $x_p = -\frac{b}{a}$; nótese por otro lado que $\lim_{t \rightarrow \infty} x = -\frac{b}{a}$ si y sólo si $\lim_{t \rightarrow \infty} K e^{at} = 0$, lo cual sucede si $a < 0$.

§2.2.2 Caso no autónomo

El caso no autónomo de las ecuaciones anteriores se presenta cuando los coeficientes no son constantes. El caso homogéneo es sencillo de resolver por simple integración. Dada la ecuación

$$\dot{x} = a(t)x, \quad (2.3)$$

se tiene (si $x \neq 0$)

$$\frac{\dot{x}}{x} = a(t),$$

e integrando ambos lados

$$\int \frac{\dot{x}}{x} dt = \int_{t_0}^t a(s) ds.$$

En el paso anterior tomamos la antiderivada de $a(t)$, que por el teorema fundamental del cálculo es simplemente $\int_{t_0}^t a(s) ds$. De aquí tenemos que,

$$\ln x = \int_{t_0}^t a(s) ds + k$$

¹O bien, en los términos del apéndice D, existe $\varepsilon > 0$ tal que $x_0 \in B_\varepsilon(x^*)$.

y por lo tanto si $K = e^k$,

$$\begin{aligned} x(t) &= K \left(e^{f(t)} \right), \\ \text{con } f(t) &= \int_{t_0}^t a(s) ds. \end{aligned} \tag{2.4}$$

Por supuesto que la solución explícita se puede obtener siempre y cuando la función $a(t)$ se pueda integrar.² La constante K se obtiene si se conoce un valor inicial, $x(t_0)$. Como se vio anteriormente, (2.4) es una solución aun si K es negativa.

Ejemplo

Ej 2.2.5 Consideremos la ecuación $\dot{x} = 2tx$. Se tiene que, de acuerdo con la fórmula (2.4),

$$x(t) = Ke^{(t^2 - t_0^2)}.$$

◇◇◇◇◇◇◇◇◇

§2.2.3 Bonos y tasas de interés

Un bono es una promesa de pago en el futuro. Distintas entidades emiten bonos para financiar sus actividades; en particular, los gobiernos emiten diversos bonos (en México los más comunes son los **Cetes** o Certificados de la Tesorería). Una propiedad que caracteriza a los bonos es que en todo momento se conoce su valor final. Digamos que tenemos un periodo de tiempo $[0, T]$ y al final de éste el bono vale B_T ; sin embargo, no se conoce el valor del bono en instantes intermedios. En principio, el valor del bono al tiempo t depende del mercado, es decir, el intercambio de bonos entre los participantes en el mercado determina el precio, $B(t)$, para todo $t \in [0, T]$. Un concepto muy útil es el de rendimiento o tasa de interés.

Definición 2.2.3 Sea $B(t)$, el precio de un bono al tiempo t . El **rendimiento** del bono en el periodo de tiempo $[t, s]$ se define como

$$\gamma(t, s) = \frac{1}{s - t} \left[\frac{B(s) - B(t)}{B(t)} \right]$$

donde $t < s$, es decir, es la tasa porcentual de cambio en el valor del bono en el periodo $[t, s]$. El **rendimiento instantáneo** al tiempo t , $r(t)$, es el límite de los rendimientos cuando $s \rightarrow t$; esto es,

$$\begin{aligned} r(t) &= \lim_{s \rightarrow t} \gamma(t, s) \\ &= \lim_{s \rightarrow t} \frac{1}{s - t} \left[\frac{B(s) - B(t)}{s - t} \right] \\ &= \frac{1}{B(t)} \frac{dB}{dt} = \frac{\dot{B}}{B}. \end{aligned}$$

La función $r(t)$ también es llamada **tasa instantánea**³ la cual representa la tasa porcentual de cambio (instantánea) en el valor del bono en el “instante” t .

²Nótese que si $a(t) = a$, una constante, se obtiene el caso autónomo si se define $K \equiv e^{k - at_0}$.

³Es común denotar la tasa instantánea de cambio, \dot{x}/x , por \hat{x} .

Notemos que B satisface

$$\dot{B} - r(t)B = 0,$$

que es una de las ecuaciones que hemos estudiado. Procedamos ahora a resolverla junto con la condición final $B(T) = B_T$ (nótese que en este caso $t_0 = T$). Por (2.4) la solución es

$$B(t) = K \left(e^{\int_t^T r(s) ds} \right).$$

Si sustituimos $t = T$, obtenemos que $K = B_T$ y por lo tanto podemos escribir

$$B(t) = B_T \left(e^{\int_t^T r(s) ds} \right) = B_T \left(e^{-\int_t^T r(s) ds} \right). \quad (2.5)$$

Los límites de integración se invierten puesto que $t < T$ y de esta forma el resultado queda expresado como “el valor presente (en t) de B_T ”. De este modo se ha demostrado que existe una función creciente,

$$\beta(t) = e^{\int_t^T r(s) ds}, \quad (2.6)$$

tal que el valor del bono al tiempo t es $B(t) = \beta(t)B_T$. La función $\beta(t)$ es llamada el **descuento** del bono y claramente tiene que satisfacer $\beta(T) = 1$. El caso más simple se presenta cuando $r(t) = r$, donde r es una constante, ya que entonces se tiene $B(t) = B_T e^{-r(T-t)}$.

La fórmula (2.5) nos permite dar varias interpretaciones. En primer lugar, vemos que el valor del bono al tiempo t depende esencialmente de los rendimientos en el intervalo $[t, T]$, es decir, el valor actual es el valor futuro descontado mediante los rendimientos del periodo. Además, la relación entre el valor actual y el final es lineal. En otras palabras, si se duplica el precio B_T , necesariamente se duplicará $B(t)$. Puede decirse que en este caso la ecuación diferencial se resuelve “hacia atrás” ya que se obtiene el valor presente del bono a partir de un valor futuro conocido. La otra posibilidad es resolver “hacia adelante”, es decir, el valor actual queda expresado en términos de un valor conocido en el pasado, que normalmente es el valor inicial. Cuando analicemos el tiempo discreto, este concepto de resolver hacia adelante o hacia atrás será de suma importancia.

Por otra parte, si tenemos dos instrumentos, $B_1(t)$ y $B_2(t)$, con tasas $r_1(t)$ y $r_2(t)$ tales que $r_1(t) > r_2(t)$, entonces los precios de los bonos satisfacen $B_1(t) < B_2(t)$. Esto puede verificarse de la siguiente manera: dado que $r_1(t) > r_2(t)$, se sabe que

$$\begin{aligned} \int_t^T r_1(s) ds &> \int_t^T r_2(s) ds, \\ - \int_t^T r_1(s) ds &< - \int_t^T r_2(s) ds, \\ e^{-\int_t^T r_1(s) ds} &< e^{-\int_t^T r_2(s) ds}, \end{aligned}$$

y al multiplicar por B_T nos queda

$$B_T \left(e^{-\int_t^T r_1(s) ds} \right) < B_T \left(e^{-\int_t^T r_2(s) ds} \right),$$

es decir, $B_1(t) < B_2(t)$.

El modelo de tasas de interés se puede mejorar; por ejemplo, podemos considerar que a lo largo del periodo de vida del bono se efectúan varios depósitos o retiros. Sea $Y(t)$ el valor de la inversión en bonos al tiempo t . Para simplificar los cálculos, escribiremos Y en términos del valor de un bono fijo, $B(t)$,

$$Y(t) = Z(t)B(t). \quad (2.7)$$

La variable $Z(t)$ representa la cantidad de bonos que se tienen en la inversión. A la cantidad $Z(t)$ se la conoce, en ocasiones, como la **posición** y en principio puede ser positiva o negativa (una posición negativa es posible cuando se debe dinero). Se supone que el cambio marginal en la cantidad $Z(t)$ está dado por

$$\dot{Z}(t) = \delta(t). \quad (2.8)$$

Nos proponemos ahora encontrar una ecuación que sea satisfecha por Y , para después proceder a resolverla. De las ecuaciones (2.7) y (2.8) obtenemos

$$\begin{aligned} \dot{Y}(t) &= Z(t)\dot{B}(t) + \dot{Z}(t)B(t) \\ &= r(t)Z(t)B(t) + \delta(t)B(t). \end{aligned}$$

Nótese que la ecuación anterior dice que el cambio en el valor de una inversión proviene de dos fuentes: del cambio en el valor de un bono y del cambio en la posición. De este modo, concluimos que los bonos, pueden modelarse con la siguiente ecuación:

$$\dot{Y}(t) - r(t)Y(t) = \delta(t)B(t), \quad (2.9)$$

donde $\delta(t)$ es una función que representa los depósitos (o retiros) y $Y(T) = Y_T$. Nótese que la anterior es una ecuación lineal no autónoma y no homogénea. En la siguiente sección estudiaremos este tipo de ecuaciones. La solución de (2.9) está dada por

$$Y(t) = \left(\frac{Y_T}{B_T} + \int_T^t \delta(s)ds \right) B(t). \quad (2.10)$$

y se deja como ejercicio para el lector (ejercicio 2.10).

§2.2.3.1 Una curiosidad: la transformada de Laplace

La llamada transformada de Laplace se utiliza constantemente en economía, no obstante que la mayoría de los usuarios desconocen su existencia. Hacemos aquí este pequeño paréntesis pues en la sección anterior acabamos de utilizar de manera inadvertida esta famosa transformada.

En general una **transformada integral** es una relación de la forma

$$F(s) = \int_0^\infty K(s, t)f(t)dt,$$

en donde una función f es transformada en otra función F por medio de una integral. La función F se conoce como la **trasformada de** f y la función K es el **kernel** de la transformación. Las transformadas integrales se utilizan para convertir algún problema que involucra a la función F en otro problema más sencillo que involucra a F .

La transformada de Laplace⁴ $\mathcal{L}\{f(t)\}$ es una transformada integral en donde el kernel está dado por

⁴Nombrada así en honor del matemático francés del siglo XVIII Pierre S. Laplace.

e^{-st} de manera que

$$\begin{aligned}\mathcal{L}\{f(t)\} &= \\ F(s) &= \int_0^\infty e^{-st} f(t) dt.\end{aligned}$$

La transformada de Laplace es útil en la resolución de algunos tipos de ecuaciones diferenciales; sin embargo en este libro no será utilizada con este fin. Simplemente nos interesa hacer notar que la transformada de Laplace de una función tiene una interpretación económica evidente: $\mathcal{L}\{f(t)\}$ es el valor presente de la función f durante el periodo $[0, \infty)$ y con un factor de descuento igual a s .

§2.2.4 Caso no autónomo y no homogéneo

El caso no homogéneo resulta un poco más complicado y utiliza la técnica de buscar un factor de integración como sigue. Dada una ecuación del tipo

$$\dot{x} = r(t)x + b(t), \quad (2.11)$$

que se reescribe como

$$\dot{x} - r(t)x = b(t), \quad (2.12)$$

notamos que al multiplicar ambos lados de (2.12) por el factor $e^{-\int_{t_0}^t r(s)ds}$ obtenemos una expresión integrable ya que la ecuación se convierte en

$$\dot{x} \left(e^{-\int_{t_0}^t r(s)ds} \right) - r(t)x \left(e^{-\int_{t_0}^t r(s)ds} \right) = b(t) \left(e^{-\int_{t_0}^t r(s)ds} \right). \quad (2.13)$$

El lado izquierdo de (2.13) no es más que la derivada con respecto a t del producto

$$x(t) \left(e^{-\int_{t_0}^t r(s)ds} \right).$$

La justificación intuitiva es que para convertir el lado izquierdo de la ecuación en la derivada de un producto, hay que multiplicarlo por el factor e^{-u} , donde u es una antiderivada de r , por lo que si $r(t)$ es simplemente una constante r , el factor de integración es $e^{-r(t-t_0)}$ y si $r(t)$ es función de t entonces, por el teorema fundamental del cálculo, la expresión $\int_{t_0}^t r(s)ds$ es la que tiene como derivada a $r(t)$.

Integrando ambos lados de (2.13), obtenemos

$$x(t) \left(e^{-\int_{t_0}^t r(s)ds} \right) = C + \int b(t) \left(e^{-\int_{t_0}^t r(s)ds} \right) dt,$$

por lo que,

$$x(t) = \left(e^{\int_{t_0}^t r(s)ds} \right) \left[C + \int b(t) \left(e^{-\int_{t_0}^t r(s)ds} \right) dt \right].$$

Dado que sería ridículo memorizar esta expresión para la solución, sugerimos simplemente encontrar el factor de integración y realizar todos los pasos pertinentes para encontrar la solución como se muestra en el siguiente ejemplo.

Este resultado se puede interpretar en términos de bonos y tasas de interés. Supongamos que $x(t)$ es el valor de una cartera de bonos. El término $B(t) = e^{\int_{t_0}^t r(s)ds}$ representa (ver ecuación (2.5)) el valor de un bono con tasa de interés instantánea $r(t)$ y valor $B(t_0) = 1$. La ecuación (2.11) se interpreta como sigue: el cambio en el valor de la cartera, \dot{x} , es igual a los rendimientos de la misma, $r(t)x$, más las ganancias (o pérdidas) por el cambio (instantáneo) en el valor de los bonos, $b(t)$. Éste último se puede expresar como $b(t) = \delta(t)B(t)$ donde

$$\delta(t) = b(t) \left(e^{-\int_{t_0}^t r(s)ds} \right)$$

representa el cambio instantáneo en el número de bonos. De este modo, podemos interpretar la solución

$$x(t) = B(t) \left[C + \int \delta(t)dt \right] \quad (2.14)$$

como: el valor de la inversión es el valor de un bono por el número total de los mismos.

Ejemplo

Ej 2.2.6 Consideremos la ecuación $\dot{x} = x - t$, que puede reescribirse como

$$\dot{x} - x = -t.$$

El factor de integración es $e^{-(t-t_0)}$ y por lo tanto la ecuación se transforma en

$$\dot{x}e^{-(t-t_0)} - xe^{-(t-t_0)} = -te^{-(t-t_0)},$$

o bien en

$$\frac{d}{dt}(xe^{-(t-t_0)}) = -te^{-(t-t_0)}.$$

Integrando ambos lados obtenemos

$$xe^{-(t-t_0)} = - \int te^{-(t-t_0)} dt = te^{-(t-t_0)} + e^{-(t-t_0)} + K,$$

donde K es una constante indeterminada. Notemos que en el lado derecho se utilizó la técnica de integración por partes. Despejando para x , se tiene finalmente

$$x(t) = Ke^{(t-t_0)} + t + 1.$$

Si $x(t_0)$ es conocido, entonces K queda determinada por

$$K = x(t_0) - t_0 - 1.$$

Una solución alternativa a este problema se basa en la interpretación de la ecuación lineal con base en bonos y tasas de interés. Claramente, se puede ver que en este problema la tasa de interés es $r(t) = 1$ y por lo tanto el factor $B(t) = e^{(t-t_0)}$ representa el valor de un bono tal que $B(t_0) = 1$. De este modo, podemos identificar el cambio marginal en el número de bonos en la inversión como la función $\delta(t)$ tal que

$$\delta(t)B(t) = \delta(t)e^{(t-t_0)} = -t,$$

es decir, $\delta(t) = -te^{-(t-t_0)}$. Usando la ecuación (2.14), obtenemos que

$$\begin{aligned} x(t) &= B(t) \left[K + \int \delta(t) dt \right] \\ &= e^{(t-t_0)} \left[K + te^{-(t-t_0)} + e^{-(t-t_0)} \right] \\ &= Ke^{-(t-t_0)} + t + 1. \end{aligned}$$

§2.3 Ecuaciones de segundo orden

Para ecuaciones de segundo orden estudiamos únicamente el caso con coeficientes constantes. Al igual que para las ecuaciones de primer orden, la solución es casi inmediata una vez resuelto el caso homogéneo debido a la proposición 2.2.1. La forma general de la ecuación homogénea es

$$a\ddot{x} + b\dot{x} + cx = 0. \quad (2.15)$$

Si $a = 0$, se trata de una ecuación de primer orden como en la sección anterior y la solución es de la forma general $x = Ke^{rt}$. La pregunta natural ahora es si seguimos teniendo este tipo de soluciones. Una manera de verificarlo es simplemente sustituir $x = Ke^{rt}$ en la ecuación (2.15) y simplificar la expresión resultante. Al hacer esto, obtenemos $Ke^{rt}(ar^2 + br + c) = 0$ que se satisface si y sólo si se cumple

$$ar^2 + br + c = 0. \quad (2.16)$$

A esta última ecuación se la denomina **ecuación característica** de la ecuación (2.15). Así, $x = Ke^{rt}$ es una solución siempre y cuando r satisfaga (2.16). Esto nos lleva a tener tres casos distintos, según si las raíces de (2.16) son ambas reales y distintas o si obtenemos una raíz real doble o dos raíces complejas.

§2.3.1 Raíces reales distintas

Supongamos que la ecuación (2.16) tiene dos raíces reales distintas r_1 y r_2 . Del párrafo anterior se tiene que $K_1e^{r_1t}$ y $K_2e^{r_2t}$ son soluciones de la ecuación (2.15). Dado que esta ecuación es lineal y homogénea, es fácil ver que la suma de esas soluciones,

$$x(t) = K_1e^{r_1t} + K_2e^{r_2t}, \quad (2.17)$$

también es solución (a esto se le llama **principio de superposición** y su prueba se pide en el ejercicio 2.18). Ésta es la solución general ya que, dada cualquier pareja de condiciones iniciales, $x(t_0)$ y $\dot{x}(t_0)$, siempre pueden encontrarse constantes únicas K_1 y K_2 que resuelven (2.15) dadas las condiciones iniciales. Esto último puede verificarse sustituyendo las condiciones iniciales en x y \dot{x} .

Para ilustrar este resultado, supongamos que $t_0 = 0$ y se conocen $x(0)$ y $\dot{x}(0)$; entonces, sustituyendo estos valores en (2.17), se obtiene el siguiente sistema de ecuaciones en K_1 y K_2 :

$$\begin{aligned} K_1 + K_2 &= x(0), \\ r_1K_1 + r_2K_2 &= \dot{x}(0), \end{aligned}$$

cuyo determinante es $r_2 - r_1 \neq 0$, por hipótesis. Por lo tanto, el sistema posee una solución única.

En conclusión, dadas r_1 y r_2 raíces reales distintas, la solución general de la ecuación (2.15) está dada por (2.17).

Ejemplo

Ej 2.3.1 Consideremos la ecuación $\ddot{x} - \dot{x} - 2x = 0$, con $x(0) = 3$ y $\dot{x}(0) = 0$. La ecuación característica es

$$r^2 - r - 2 = 0$$

y tiene como raíces a 2 y -1 por lo que la solución general está dada por

$$x(t) = K_1 e^{2t} + K_2 e^{-t}.$$

Sustituyendo los valores iniciales para $x(0)$ y $\dot{x}(0)$ se obtiene que $K_1 = 1$ y $K_2 = 2$ y por lo tanto la solución final queda como

$$x(t) = e^{2t} + 2e^{-t}.$$

◇◇◇◇◇◇◇◇◇

§2.3.2 Una raíz real doble

Supongamos ahora que el polinomio característico tiene una raíz real doble, r . La expresión $K_1 e^{rt}$ es una solución del sistema pero necesitamos otra solución (independiente de ésta para generar la solución general). Es fácil ver que la solución $K_2 t e^{rt}$ también satisface la ecuación (2.15), tomando en cuenta que $r = \frac{-b}{2a}$. La suma de estas dos soluciones también es una solución de (2.15) y de manera análoga al caso anterior se puede ver que es la solución general. En resumen, si r es una raíz real doble del polinomio (2.16), entonces la solución general de la ecuación (2.15) está dada por

$$x(t) = K_1 e^{rt} + K_2 t e^{rt}. \quad (2.18)$$

Ejemplo

Ej 2.3.2 Considérese $\ddot{x} - 6\dot{x} + 9x = 0$, con las condiciones iniciales $x(0) = 1$ y $\dot{x}(0) = 0$. La ecuación característica es

$$r^2 - 6r + 9 = 0,$$

que tiene la raíz doble $r = 3$. La solución está dada, entonces, por

$$x(t) = K_1 e^{3t} + K_2 t e^{3t}.$$

Sustituyendo las condiciones iniciales para $x(0)$ y $\dot{x}(0)$ se tiene el sistema

$$\begin{aligned} K_1 &= 1, \\ 3K_1 + K_2 &= 0, \end{aligned}$$

y se obtienen así los valores $K_1 = 1$ y $K_2 = -3$.

◇◇◇◇◇◇◇◇◇

§2.3.3 Raíces complejas

Tomemos ahora el caso en que el polinomio (2.16) tiene raíces complejas. Dado que se trata de un polinomio con coeficientes reales, las raíces necesariamente son conjugadas así que podemos suponer que son de la forma $r_1 = \alpha + \beta i$ y $r_2 = \alpha - \beta i$ (la prueba de esto es sumamente sencilla y aparece como ejercicio para el lector en el apéndice A). Podemos proponer entonces, al igual que en el caso de dos raíces reales, que la función

$$x(t) = K_1 e^{(\alpha+\beta i)t} + K_2 e^{(\alpha-\beta i)t} \quad (2.19)$$

es una solución general de la ecuación (2.15). El problema es que ésta es una solución compleja y lo que se requiere es una solución real, la cual puede obtenerse como sigue. Utilizando la fórmula de Euler para números complejos, reescribimos

$$\begin{aligned} e^{(\alpha+\beta i)t} &= e^{\alpha t} (\cos \beta t + i \sin \beta t), \\ e^{(\alpha-\beta i)t} &= e^{\alpha t} (\cos \beta t - i \sin \beta t). \end{aligned}$$

Si en (2.19) se escogen K_1 y K_2 como complejos conjugados, por ejemplo $K_1 = c_1 + ic_2$ y $K_2 = c_1 - ic_2$, puede obtenerse una solución real ya que la suma

$$(c_1 + ic_2)e^{\alpha t}(\cos \beta t + i \sin \beta t) + (c_1 - ic_2)e^{\alpha t}(\cos \beta t - i \sin \beta t),$$

es simplemente igual a

$$e^{\alpha t}(2c_1 \cos \beta t - 2c_2 \sin \beta t).$$

Renombrando las constantes como $C_i = 2c_i$, $i = 1, 2$, podemos escribir la solución general como,

$$x(t) = e^{\alpha t}(C_1 \cos \beta t - C_2 \sin \beta t), \quad (2.20)$$

donde α y β son la parte real e imaginaria de la pareja de raíces complejas conjugadas.

Ejemplo

Ej 2.3.3 Consideremos la ecuación $\ddot{x} + 2\dot{x} + 2x = 0$, con $x(0) = 1$ y $\dot{x}(0) = 0$. La ecuación característica es

$$r^2 + 2r + 2 = 0$$

que tiene como raíces $-1 \pm i$. La solución general está dada, de acuerdo con (2.20), por

$$x(t) = e^{-t}(C_1 \cos t - C_2 \sin t).$$

Sustituyendo las condiciones iniciales, y tomando en cuenta que

$$\dot{x} = -e^{-t}[(C_1 + C_2) \cos t + (C_1 - C_2) \sin t],$$

se obtiene el sistema

$$C_1 = 1,$$

$$C_1 + C_2 = 0,$$

y por lo tanto $C_1 = 1$ y $C_2 = -1$.



§2.3.4 Caso no homogéneo

Al igual que en las ecuaciones de primer orden con coeficientes constantes, la proposición 2.2.1 implica que la solución del caso no homogéneo es la suma de una solución particular, x_p , de la ecuación

$$a\ddot{x} + b\dot{x} + cx = d, \quad (2.21)$$

más la solución general de la ecuación homogénea asociada dada por (2.15).

La solución particular más simple es una función constante, de manera que se cumple $\dot{x} = \ddot{x} = 0$. Esto es posible siempre y cuando $c \neq 0$ en (2.21), ya que entonces se tiene $x_p = \frac{d}{c}$. Si $c = 0$ pero $b \neq 0$, entonces simplemente se toma como solución particular una función lineal, $x = At + B$, y por lo tanto se cumple $\ddot{x} = 0$. Resolvemos $b\dot{x} = d$ por integración, de manera que $x_p = \frac{d}{b}t$ es una solución particular. Finalmente, si $c = b = 0$, pero $a \neq 0$, entonces la ecuación original es $a\ddot{x} = d$, cuya solución, tras integrar dos veces, es

$$x_p(t) = \frac{d}{2a}t^2 + K_1t + K_2.$$

Ejemplo

Ej 2.3.4 Considérese la ecuación $\ddot{x} - \dot{x} - 2x = -4$. La solución de la ecuación homogénea asociada se calculó en el ejemplo 2.3.1:

$$x_h = K_1 e^{2t} + K_2 e^{-t}.$$

Una solución particular se obtiene poniendo $\dot{x} = \ddot{x} = 0$, por lo que $x_p = 2$. Así, la solución general queda dada por

$$x(t) = K_1 e^{2t} + K_2 e^{-t} + 2.$$

Como no tenemos condiciones iniciales, las constantes quedan indeterminadas.

Ej 2.3.5 (Inflación y desempleo) Denotemos por $u(t)$ a la tasa de desempleo y por $p(t)$ al logaritmo del nivel de precios, de manera que \dot{p} representa la tasa de inflación⁵. Las constantes \bar{m} y \bar{u} representan, respectivamente, el logaritmo de la cantidad nominal de dinero y la tasa natural de desempleo.⁶ Supongamos que la evolución de u y p está dada por las ecuaciones

$$\dot{u} = -\alpha(\bar{m} - p) + \beta(\bar{u} - u), \quad (2.22)$$

$$\dot{p} = \gamma(\bar{u} - u), \quad (2.23)$$

en donde α, β y γ son constantes positivas tales que $\beta^2 > 4\alpha\gamma$. La primera ecuación nos dice que el desempleo responde a la política monetaria (primer término) y que tiende hacia su tasa natural (segundo

⁵Esto debe ser evidente ya que si $P(t)$ es el nivel de precios, entonces $p(t) = \ln P(t)$ y $\dot{p} = \frac{\dot{P}}{P}$ es la tasa instantánea de cambio en el nivel de precios o inflación.

⁶ \bar{m} se puede pensar como la oferta monetaria o circulante que ofrece el Banco Central. La tasa natural de desempleo existe ya que es imposible que en una economía toda la fuerza laboral esté empleada. Siempre habrá aquellos que recién ingresen al mercado de trabajo o bien que estén “entre trabajos” y aparezcan estadísticamente como desempleados.

término). La segunda ecuación es lo que se conoce como una relación de Phillips.⁷ De 2.23 obtenemos las siguientes relaciones

$$\begin{aligned}\bar{u} - u &= \frac{\dot{p}}{\gamma}, \\ \dot{u} &= \frac{-\ddot{p}}{\gamma}.\end{aligned}$$

Sustituyendo para $\bar{u} - u$ y \dot{u} en 2.22 y reescribiendo se tiene

$$\ddot{p} + \beta\dot{p} + \alpha\gamma p = \alpha\gamma\bar{m}. \quad (2.24)$$

Ésta última es una ecuación diferencial de segundo orden, con coeficientes constantes, en p . La ecuación característica es

$$\lambda^2 + \beta\lambda + \alpha\gamma = 0,$$

con raíces

$$\lambda_{1,2} = \frac{-\beta \pm \sqrt{\beta^2 - 4\alpha\gamma}}{2}.$$

La condición $\beta^2 > 4\alpha\gamma$ implica que ambas raíces son reales y negativas. Una solución particular a 2.24 se obtiene al considerar $\dot{p} = \ddot{p} = 0$, de manera que $p = \bar{m}$. Así, la solución general queda dada por

$$p(t) = \bar{m} + Ae^{\lambda_1 t} + Be^{\lambda_2 t},$$

con lo cual

$$\begin{aligned}u(t) &= \bar{u} + \frac{\dot{p}}{\gamma} \\ &= \bar{u} + \frac{1}{\gamma} (\lambda_1 A e^{\lambda_1 t} + \lambda_2 B e^{\lambda_2 t}).\end{aligned}$$

Dadas las condiciones iniciales $p(0) = p_0$ y $u(0) = u_0$ se pueden determinar A y B como

$$\begin{aligned}A &= \frac{\lambda_2(p_0 - \bar{m}) - \gamma(u_0 - \bar{u})}{\lambda_2 - \lambda_1}, \\ B &= \frac{\gamma(u_0 - \bar{u}) - \lambda_1(p_0 - \bar{m})}{\lambda_2 - \lambda_1}\end{aligned}$$

Nótese que se cumplen

$$\begin{aligned}\lim_{t \rightarrow \infty} p(t) &= \bar{m}, \\ \lim_{t \rightarrow \infty} u(t) &= \bar{u}.\end{aligned}$$

◇◇◇◇◇◇◇◇◇

⁷Véase [Phi58]. Hoy en día se conoce como relación o **curva de Phillips** a la relación positiva que existe entre el ingreso (o equivalentemente el empleo) y la inflación esperada. En este caso, lo que se tiene es una relación negativa entre la inflación y el desempleo. Podríamos decir que la inflación esperada es igual a la inflación real suponiendo que los agentes tienen “previsión perfecta”. Esto se analizará con detalle en la sección 3.4.

§2.3.4.1 Coeficientes indeterminados

Si ahora consideramos una ecuación no autónoma del tipo

$$a\ddot{x} + b\dot{x} + cx = f(t), \quad (2.25)$$

también tenemos, por la proposición 2.2.1, que la solución general es de la forma $x_h + x_p$; sin embargo, aquí el problema es que puede no ser claro cómo encontrar una solución particular x_p . El método más utilizado es el llamado **método de coeficientes indeterminados**. La idea es “adivinar” una solución particular que tenga la misma forma general de la función $f(t)$, sustituirla en (2.25) y determinar los coeficientes de manera que la ecuación sea válida. Este método puede aplicarse para la mayoría de la funciones $f(t)$ que aparecen comúnmente y cuyas derivadas son combinaciones lineales de un número finito de funciones posibles. Entre éstas se encuentran los polinomios, las funciones trigonométricas y las exponenciales.

Ejemplo

Ej 2.3.6 Se quiere encontrar la solución general de la ecuación $\ddot{x} - \dot{x} - 2x = 4t^2 + 1$. La solución de la ecuación homogénea asociada está dada, como en el ejemplo 2.3.1, por $x_h = K_1e^{2t} + K_2e^{-t}$. Para encontrar la solución particular suponemos que ésta tiene la forma de la función $f(t)$, que en este caso es una función cuadrática. De esta forma adivinamos una solución particular dada por las combinaciones lineales del conjunto $\{t^2, t, 1\}$, es decir,

$$x_p = At^2 + Bt + C,$$

en donde se deben determinar los valores específicos de las constantes A, B y C . Obtenemos \dot{x}_p y \ddot{x}_p , y sustituimos en la ecuación original:

$$\begin{aligned} 2A - (2At + B) - 2(At^2 + Bt + C) &= \\ -2At^2 - 2(A + B)t + (2A - B - 2C) &= 4t^2 + 1. \end{aligned}$$

Los términos en ambos lados de la ecuación, polinomios en este caso, deben ser idénticos, de manera que debe cumplirse,

$$\begin{aligned} -2A &= 4, \\ 2(A + B) &= 0, \\ 2A - B - 2C &= 1. \end{aligned}$$

Resolviendo, se obtiene que $A = -2, B = 2$ y $C = -\frac{7}{2}$, con lo cual la solución particular queda determinada; por lo tanto, la solución general es

$$x(t) = K_1e^{2t} + K_2e^{-t} - 2t^2 + 2t - \frac{7}{2}.$$



$f(t)$	$x_p(t)$
$a_n t^n + \dots + a_1 t + a_0$	$t^k (A_n t^n + \dots + A_1 t + A_0)$
$a e^{\alpha t}$	$t^k A e^{\alpha t}$
$a \cos \beta t + b \sin \beta t$	$t^k (A \cos \beta t + B \sin \beta t)$
$e^{\alpha t} (a_n t^n + \dots + a_1 t + a_0)$	$e^{\alpha t} t^k (A_n t^n + \dots + A_1 t + A_0)$
$a e^{\alpha t} \cos \beta t + b e^{\alpha t} \sin \beta t$	$t^k (A e^{\alpha t} \cos \beta t + B e^{\alpha t} \sin \beta t)$

Tabla 2.1: Soluciones particulares comunes.

El método tiene que modificarse ligeramente si la función $f(t)$ o alguna de sus derivadas es solución de la ecuación homogénea. En este caso, la solución particular es una combinación lineal del conjunto,

$$\left\{ t^k f(t), t^k f'(t), \dots, t^k f^{(n)}(t) \right\},$$

donde k es el mínimo exponente tal que ningún elemento del conjunto es solución de la ecuación homogénea. La tabla 2.1 muestra la forma de una solución particular para algunas formas funcionales de f .

El siguiente ejemplo muestra estas consideraciones.

Ejemplo

Ej 2.3.7 Encontrar la solución general de $\ddot{x} + 16x = \cos 4t$. En este caso, la ecuación homogénea asociada es $\ddot{x} + 16x = 0$. La ecuación característica tiene como raíces $\lambda_1 = 4i$ y $\lambda_2 = -4i$, y por lo tanto la solución general de la ecuación homogénea es de la forma

$$x_h(t) = K_1 \cos 4t - K_2 \sin 4t. \quad (2.26)$$

Como primera opción, la solución particular debería ser una combinación de las funciones $\cos 4t$ y $\sin 4t$. Sin embargo, la forma de la solución homogénea dada por (2.26) implica que las combinaciones lineales de $\{\cos 4t, \sin 4t\}$ no pueden ser usadas para obtener una solución particular. Se debe intentar, por tanto, con soluciones particulares de la forma

$$x_p(t) = At \cos 4t + Bt \sin 4t.$$

Al sustituir en la ecuación original se obtiene

$$8B \cos 4t - 8A \sin 4t = \cos 4t,$$

de manera que $B = 1/8$ y $A = 0$. Concluimos de este modo que la solución general al problema es

$$x(t) = K_1 \cos 4t - K_2 \sin 4t + \frac{1}{8}t \sin 4t.$$

§2.3.5 Aversión absoluta al riesgo constante

Con el afán de modelar la toma de decisiones de los agentes económicos la teoría económica supone que éstos tratan de maximizar alguna función de beneficios. Dichas funciones deben cumplir ciertas

propiedades para representar el comportamiento de los agentes; en particular, pensemos en un individuo que obtiene “utilidad” (medida en “útiles” o unidades de placer) a partir de su riqueza. Sean w la riqueza y $u(w)$ la utilidad proporcionada por la misma, en donde la función u es conocida como **función de utilidad** y tiene al menos dos derivadas continuas. Esta función es creciente en w y adicionalmente es cóncava, es decir cumple $u' > 0$ y $u'' < 0$ (esta última propiedad representa “rendimientos decrecientes” en w). Se define el **coeficiente de aversión absoluta al riesgo** como

$$s(w) = -\frac{u''(w)}{u'(w)}.$$

Éste representa el porcentaje de cambio de la utilidad marginal para un nivel de riqueza dado, w . Por su uso frecuente es útil caracterizar aquellas funciones de utilidad que poseen **aversión absoluta al riesgo constante** (conocidas como **funciones CARA**, por sus siglas en inglés). El problema se reduce a resolver la ecuación diferencial

$$-\frac{u''}{u'} = s,$$

donde s es una constante. Reescribiendo, tenemos la ecuación

$$u'' + su' = 0,$$

cuyo polinomio característico es simplemente,

$$r^2 + sr = 0,$$

con raíces $r_1 = 0$ y $r_2 = -s$. Por lo tanto, la solución general está dada por

$$u(w) = A + Be^{-sw},$$

donde A y B se pueden determinar con las condiciones iniciales de u y u' .

Ejercicios

▷ **2.1** Probar que:

- a) $x(t) = t + \frac{A}{t}$, con A constante, es solución de

$$\dot{x} + \frac{x}{t} = 2.$$

- b) $x(t) = A \cos 3t + B \sin 3t + \frac{1}{8} \cos t$, con A y B constantes, es solución de

$$\ddot{x} + 9x - \cos t = 0.$$

- c) $x(t) = \frac{t^3(1+t)}{4} + \frac{1}{t} + 1$ es solución de

$$\dot{x} = \frac{x}{1+t} - \frac{x}{t} + t^2 + t^3,$$

$$\text{con } x(1) = \frac{5}{2}.$$

▷ **2.2** La ecuación $\dot{x} + bx + c = 0$, con $x(0) = x_0$, tiene como solución a $x(t) = 3e^{3t} + 2$. Encontrar los coeficientes b y c y el valor inicial x_0 .

▷ **2.3** Si todas las soluciones de la ecuación $\ddot{x} - 9x = 1$ son de la forma

$$x(t) = \beta + Ae^{\alpha t} + Be^{-\alpha t},$$

encontrar los valores de α y β y posteriormente los valores de A y B tales que se satisfagan las condiciones iniciales $x(0) = \dot{x}(0) = 0$.

▷ **2.4** Considerar la ecuación $y''(v) + 2y'(v) + 2y(v) = 0$. Verificar que para toda α y β la siguiente es una solución

$$y(v) = e^{-v}(\alpha \cos v + \beta \sin v).$$

Posteriormente encontrar los valores de α y β para los cuales se satisfacen las condiciones iniciales $y(0) = 0$ y $y'(0) = 7$.

▷ **2.5** Resolver las siguientes ecuaciones y determinar si la solución converge o no a su único estado estacionario.

a) $\dot{x} = 5x.$

b) $2\dot{x} - x = 0.$

c) $\dot{x} = 8 - x.$

d) $\dot{x} + 5x - 10 = 0.$

▷ **2.6** Considerar las siguientes funciones de oferta y demanda de un bien en donde P denota el precio y Q la cantidad del bien:

$$Q^o = P - 4,$$

$$Q^d = 11 - 2P.$$

Supongamos que $P = P(t)$ es una función del tiempo y por lo tanto también Q lo es. El precio cambia en el tiempo de acuerdo con la ecuación diferencial

$$\dot{P} = 2(Q^d - Q^o) = 30 - 6P.$$

Resolver para $P(t)$, encontrar el estado estacionario y determinar si el sistema converge a este último.

▷ **2.7 (Modelo de Malthus)** El modelo de Malthus de crecimiento para una población propone que ésta evoluciona de acuerdo con la ecuación

$$\dot{P} = aP,$$

en donde P representa el número de individuos de la población y $a > 0$ es la tasa de crecimiento. Contestar las siguientes preguntas:

- a) ¿Cuál es la población al tiempo t si inicialmente es $P(0) = P_0$?

- b) ¿Cuánto tiempo se necesita para que se duplique la población?

- c) Si se supone ahora que $a < 0$, calcular $\lim_{t \rightarrow \infty} P(t)$.

▷ **2.8** Supongamos que ahora la población evoluciona de acuerdo con la ecuación

$$\dot{P} = (\alpha - \beta)P,$$

en donde α representa la tasa de natalidad y β la tasa de mortalidad. Resolver para $P(t)$ dado que $P(0) = P_0$ y analizar el comportamiento de $\lim_{t \rightarrow \infty} P(t)$ dependiendo de las magnitudes de α y β .

▷ **2.9** Supongamos que ahora se tiene que la población evoluciona de acuerdo con la ecuación

$$\dot{P} = aP - E,$$

en donde $a > 0$ es la tasa de crecimiento y E representa un número fijo de personas que emigra cada año. Resolver para $P(t)$ dado P_0 y analizar qué pasa con $\lim_{t \rightarrow \infty} P(t)$ para distintos valores de los parámetros.

▷ **2.10** Obtener la solución dada por (2.10) de la ecuación (2.9), usando el método propuesto en la sección 2.2.4. Interpretar el resultado.

▷ **2.11** Consideremos la siguiente ecuación:

$$\dot{Y} = \left(r_0 - \frac{1}{t+1}\right)Y - \frac{T+1}{t+1}$$

con condición $Y(T) = 1$. Resolver la ecuación usando la interpretación financiera de la sección 2.2.3.

- a)** Identificar la tasa de interés $r(t)$.
- b)** Encontrar el valor de un bono $B(t)$ tal que $B(T) = 1$.
- c)** Identificar el cambio marginal $\delta(t)$ en la posición de la inversión.
- d)** Encontrar la posición $Z(t)$ si se supone que $Y(T) = 1$.

e) Encontrar la solución $Y(t)$ al problema y verificar la solución.

▷ **2.12** Sea $r > 0$ una constante. La ecuación lineal

$$\dot{Y} = rY - X(t) \quad (2.27)$$

relaciona el valor de una inversión Y al tiempo t con la función $X(t)$, que representa los flujos de dicha inversión. Se supone que el valor presente de la inversión satisface

$$Y(0) = \int_0^\infty e^{-r\tau} X(\tau) d\tau.$$

- a)** Dar una justificación intuitiva para la ecuación (2.27). (Sugerencia: si h es un número pequeño y positivo, entonces $Y(t) \approx e^{-rh} Y(t+h) + \int_{t-h}^t X(\tau) d\tau$.)
- b)** Resolver la ecuación diferencial usando el método de la sección 2.2.4.
- c)** Demostrar que la solución Y se puede reescribir como

$$Y(t) = \int_0^\infty e^{-r\tau} X(\tau+t) d\tau.$$

- d)** Interpretar el resultado.

▷ **2.13** Sean f y g dos funciones cuyas transformadas de Laplace existen para toda $s > 0$. Probar que para todo $a, b \in \mathbb{R}$ se cumple

$$\mathcal{L}\{af(t) + bg(t)\} = a\mathcal{L}\{f(t)\} + b\mathcal{L}\{g(t)\},$$

es decir, la transformada de Laplace es un operador lineal.

▷ **2.14** Resolver las siguientes ecuaciones:

- a)** $\dot{x} + (\cos t)2x = \cos t$.
- b)** $\dot{x} + 2tx = t$.
- c)** $\dot{x} + t^2x = 5t^2$, con la condición inicial $x(0) = 6$.

d) $2\dot{x} + 12x + 2e^t = 0$, con la condición inicial $x(0) = \frac{6}{7}$.

e) $y' + 3u^2y = u^2$, con la condición inicial $y(0) = 1$ (y es una función de la variable u).

▷ **2.15** Sea p el precio de un activo, p^e el precio esperado del mismo, d los dividendos constantes y r la tasa de un instrumento alternativo, por ejemplo un bono gubernamental, también constante. Asumamos que se cumple la siguiente condición de no arbitraje:

$$pr = d + \dot{p}^e. \quad (2.28)$$

Ésta nos dice que es equivalente invertir una cantidad p en un bono a una tasa r que quedarnos con el activo y obtener los dividendos más las ganancias (pérdidas) dadas por el cambio esperado en su valor (\dot{p}^e). Por otro lado, las expectativas acerca del precio esperado p^e evolucionan de acuerdo con

$$\dot{p}^e = \alpha(p - p^e), \quad (2.29)$$

en donde $r > \alpha$ (en la sección 3.4 veremos más acerca de las expectativas de los agentes económicos).

a) Eliminar p de las ecuaciones dadas, obtener una ecuación diferencial para \dot{p}^e y resolver la ecuación suponiendo que $p^e(0) = p_0^e$.

b) Probar que $\int_0^\infty de^{-rt} dt = \frac{d}{r}$. Éste es simplemente el valor presente de los dividendos futuros del activo, descontados a una tasa r .

c) Definir el **valor fundamental del activo** como $p^* = \frac{d}{r}$ y probar que se cumple $\lim_{t \rightarrow \infty} p^e(t) = p^*$.

d) Probar que $p = p^* + \frac{\alpha}{r}(p - p^e)$, sustituir la solución para p^e y demostrar que $\lim_{t \rightarrow \infty} p(t) = p^*$, es decir, tanto el precio esperado como el precio convergen al valor fundamental del activo.

e) Graficar, cualitativamente, las trayectorias de $p(t)$ y $p^e(t)$.

▷ **2.16** Verificar que la sustitución $v = \log(y)$ transforma la ecuación

$$y' + P(t)y = Q(t)y \log(y)$$

en la ecuación $v' - Q(t)v = -P(t)$.

▷ **2.17** Utilizar el método del problema anterior para resolver la ecuación

$$ty' + t^3y + y \log(y) = 0.$$

▷ **2.18** Probar el siguiente **principio de superposición**: dada una ecuación diferencial homogénea, lineal y con coeficientes constantes, si x_1 y x_2 son soluciones, entonces $x_3 = ax_1 + bx_2$ es solución. En otras palabras, cualquier combinación lineal de soluciones también es solución.

▷ **2.19** Encontrar una ecuación diferencial lineal de segundo orden cuya solución sea:

a) $x(t) = \frac{-t^2}{2} + 4t + 2$.

b) $x(t) = Ae^{2t} + Be^t + 3t + 1$.

c) $x(t) = e^{-2t}(A \cos t - B \sin t)$.

▷ **2.20** Resolver las ecuaciones dadas las condiciones iniciales:

a) $\ddot{x} - x = 0$, $x(0) = \dot{x}(0) = 1$.

b) $2\ddot{x} - 5\dot{x} + 6x = 0$, $x(0) = 3$, $\dot{x}(0) = 7$.

c) $\ddot{x} + 2\dot{x} + 2x = 0$, $x(0) = 1$, $\dot{x}(0) = 0$.

d) $\ddot{x} - 6\dot{x} + 9x = 0$, $x(0) = 1$, $\dot{x}(0) = 0$.

▷ **2.21** Encontrar la solución general de las siguientes ecuaciones:

a) $\ddot{x} - 2\dot{x} - x = 7.$

b) $\ddot{x} - 2\dot{x} - 3x = 9t^2.$

b) $\ddot{x} + x = 1.$

c) $\ddot{x} - \dot{x} - 2x = 4e^{-t}.$

c) $2\ddot{x} - 5\dot{x} + 6x = 18.$

d) $\ddot{x} - 2\dot{x} - 3x = 2e^t - 10 \sin t.$ (Sugerencia: considerar una solución particular de la forma $Ae^t + B \cos t + C \sin t.$)

▷ 2.22 Resolver el ejemplo 2.3.5 para el caso $\beta^2 \leq$

$4\alpha\gamma$ y calcular $\lim_{t \rightarrow \infty} p(t)$ y $\lim_{t \rightarrow \infty} u(t).$

▷ 2.23 Encontrar la solución general de las siguientes ecuaciones:

a) $\ddot{x} + 4x = \sin 2t.$

$\ddot{x} - 2\ddot{x} - \dot{x} + 2x = 0.$

▷ 2.24 Extender el método de solución para las ecuaciones de segundo orden con coeficientes constantes y resolver la ecuación

CAPÍTULO 3

Ecuaciones no lineales de primer orden

§3.1 Ecuaciones separables

Un tipo de ecuación diferencial de primer orden que surge con cierta frecuencia es el de las llamadas **ecuaciones diferenciales separables**, las cuales tienen la forma general:

$$\dot{x} = f(x)h(t). \quad (3.1)$$

Su nombre proviene del hecho de que pueden agruparse las funciones que contienen a x en un lado de la ecuación y las que contienen a t en el otro, de manera que se puede integrar directamente. La ecuación (3.1) se reescribe como

$$\frac{dx}{dt} = f(x)h(t),$$

y separando términos tenemos¹

$$\frac{dx}{f(x)} = h(t)dt,$$

con lo cual, si tenemos un poco de suerte, podríamos integrar ambos lados para obtener la solución.

Ejemplos

Ej 3.1.1 Consideremos la ecuación $\dot{x} = \sin t$. Éste es el caso más sencillo pues se tiene una ecuación de la forma $\dot{x} = h(t)$, con lo cual simplemente se integran ambos lados de la ecuación para obtener la solución. En nuestro ejemplo se obtiene

$$x(t) = -\cos t + k.$$

¹Simplemente utilizamos el concepto de **diferencial de una función**: si $x(t)$ es diferenciable, entonces $dx = \dot{x}dt$.

Ej 3.1.2 Dada la ecuación $\dot{x} = xt^2$, ésta se reescribe como

$$\frac{dx}{x} = t^2 dt,$$

que se integra fácilmente para obtener

$$\ln x = \frac{t^3}{3} + k.$$

Tomando la exponencial y renombrando la constante, la solución resulta finalmente

$$x(t) = Ke^{\frac{t^3}{3}}.$$

Ej 3.1.3 Dada la ecuación $\dot{x} = -\frac{x^3}{2t}$ con condición inicial $x(1) = 0$, se tiene

$$-\frac{2dx}{x^3} = \frac{dt}{t}.$$

Integrando ambos lados obtenemos

$$\frac{1}{x^2} = \ln t + k,$$

de manera que, despejando x ,

$$x(t) = \left(\frac{1}{\ln t + k} \right).$$

Sustituyendo la condición inicial $x(0) = 1$ se tiene que $k = 1$. Nótese que aquí no podemos tener una condición inicial $t = 0$ ya que la función $\ln t$ no está definida en $t = 0$.

◇◇◇◇◇◇◇◇◇

§3.1.1 Aversión relativa al riesgo constante

La notación de esta sección es la misma que la utilizada en la sección 2.3.5 en donde se estudiaron las funciones de utilidad con aversión absoluta al riesgo constante. Aquí se caracterizan las funciones de utilidad que tienen un coeficiente de **aversión relativa al riesgo constante** (conocidas por funciones **CRRA** por sus siglas en inglés). La **medida de aversión relativa al riesgo de Arrow-Pratt** se define por el cociente

$$r(w) \equiv -\frac{u''(w)w}{u'(w)}.$$

¿Qué significa esta expresión? Si la reescribimos como

$$r(w) = -\frac{\frac{u''}{u'}}{\frac{1}{w}},$$

vemos que indica la razón del cambio porcentual instantáneo en la utilidad marginal entre el cambio porcentual de la riqueza, cuando ésta se incrementa en una unidad. A este cociente se lo conoce normalmente como la elasticidad de la utilidad marginal con respecto a la riqueza. El signo negativo es simplemente para que el coeficiente $r(w)$ sea positivo. Si $r(w) = r$, una constante, el cambio porcentual en la utilidad

marginal es siempre el mismo cuando la riqueza aumenta en una unidad, independientemente del nivel de ésta.

Para ver qué tipo de funciones poseen $r(w)$ constante, se resuelve la ecuación diferencial dada por

$$-\frac{u''(w)w}{u'(w)} = r.$$

Aquí la variable independiente no es el tiempo sino la riqueza, w , y se puede interpretar como una ecuación de primer orden si la pensamos como ecuación de u' y su derivada u'' . Esta ecuación puede reescribirse como

$$u'' = -\frac{u'r}{w}$$

o bien,

$$\frac{du'}{dw} = -\frac{u'r}{w}.$$

Al separar variables se obtiene

$$\frac{du'}{u'} = -\frac{r}{w} dw,$$

e integrando ambos lados de la ecuación queda

$$\ln u' = -r \ln w + a.$$

Tomando la exponencial y renombrando la constante se llega finalmente a

$$u' = Ae^{-r \ln w} = Aw^{-r}.$$

Ésta es la utilidad marginal. Para obtener la función u simplemente se toma la integral con respecto a w , lo que conduce a

$$u = \frac{Aw^{1-r}}{1-r} + B, \text{ si } r \neq 1,$$

o bien,

$$u = A \ln w + B, \text{ si } r = 1.$$

Las constantes A y B pueden determinarse con las condiciones iniciales para u y u' .

§3.1.2 Modelo logístico

En el ejercicio 2.7 vimos el modelo malthusiano para crecimiento de una población. Otro modelo sumamente utilizado, dado que representa varios fenómenos reales, es el **modelo de crecimiento logístico o de Verhulst**. Éste incluye la idea de que a mayor población, P , menor tasa de crecimiento, lo cual puede expresarse con una ecuación del tipo

$$\frac{\dot{P}}{P} = a - bP, \tag{3.2}$$

donde $a \neq 0$ y $b > 0$. Podemos reescribir (3.2) como

$$\frac{dP}{P(a - bP)} = dt$$

y resolver integrando directamente ambos lados. La integral del lado izquierdo de la ecuación se resuelve por el método de fracciones parciales como sigue:

$$\begin{aligned}\frac{1}{P(a-bP)} &= \frac{A}{P} + \frac{B}{a-bP} = \frac{(B-bA)P+aA}{P(a-bP)} \\ \Rightarrow aA &= 1 \text{ y } B-bA = 0 \\ \Rightarrow A &= \frac{1}{a} \text{ y } B = \frac{b}{a}.\end{aligned}$$

De aquí que

$$\begin{aligned}\int \frac{dP}{P(a-bP)} &= \\ \int \frac{1}{aP} dP + \int \frac{b}{a(a-bP)} dP &= \frac{1}{a} \ln \frac{P}{a-bP},\end{aligned}$$

por lo que

$$\begin{aligned}\frac{P}{a-bP} &= Ke^{at} \Rightarrow \\ P &= \frac{aKe^{at}}{bKe^{at}+1} = \frac{a}{b+ke^{-at}},\end{aligned}$$

donde $k = \frac{1}{K}$. Dada una población inicial $P(0) = P_0$, podemos entonces expresar la solución como

$$P(t) = \frac{a}{b + \left(\frac{a}{P_0} - b\right)e^{-at}}.$$

Podemos ahora ver qué pasa cuando $t \rightarrow \infty$ para distintos valores de los parámetros. Es fácil calcular que,

$$\lim_{t \rightarrow \infty} P(t) = \begin{cases} 0 & \text{si } a \leq 0, \\ \frac{a}{b} & \text{si } a > 0. \end{cases}$$

Finalmente, observamos que si $P_0 = \frac{a}{b}$ entonces $P(t)$ es una función constante igual a $\frac{a}{b}$ en todo momento. La figura 3.1 ilustra algunas de las consideraciones anteriores.

Esta sencilla modificación al modelo malthusiano produce cambios cualitativos enormes, que en cierto modo lo hacen superior. Ilustra, en particular, el principio fundamental de que todo modelo no es más que una representación incompleta de la realidad.

§3.2 Ecuación de Bernoulli

Una ecuación de la forma

$$\dot{x} + P(t)x = Q(t)x^n \quad (3.3)$$

se llama ecuación de Bernoulli. Nótese que para $n = 0, 1$, (3.3) se reduce a la ecuación lineal del capítulo 2.

Supongamos que $n \neq 0, 1$. En este caso utilizamos la siguiente sustitución,

$$w = x^{1-n}$$

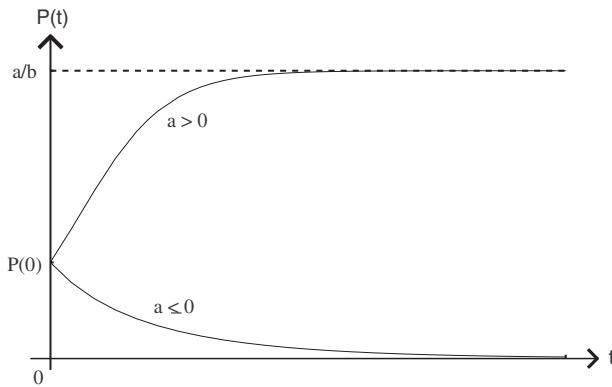


Figura 3.1: Crecimiento logístico.

y transformamos la ecuación en otra ecuación en w . Bajo este cambio de variable se obtiene

$$\begin{aligned}\dot{w} &= (1 - n)x^{-n}\dot{x} \\ &= (1 - n)x^{-n}[Q(t)x^n - P(t)x] \\ &= (1 - n)[Q(t) - P(t)x^{1-n}] \\ &= (1 - n)[Q(t) - P(t)w].\end{aligned}$$

Al reescribir la ecuación, obtenemos

$$\dot{w} + (1 - n)P(t)w = (1 - n)Q(t).$$

Observemos que esta ecuación es lineal en w , por lo que ahora el problema se reduce al caso estudiado con anterioridad.

Ejemplos

Ej 3.2.1 Resolveremos de nuevo el problema del modelo logístico. Tenemos el siguiente modelo de población:

$$\begin{aligned}\dot{P} &= aP - bP^2, \\ P(0) &= P_0.\end{aligned}$$

Ésta es una ecuación de Bernoulli de la forma (3.3) con $n = 2$ y para resolverla utilizaremos $w = P^{-1}$. La nueva ecuación en w es

$$\dot{w} = -aw + b,$$

que es del tipo (2.1) y su solución está dada por

$$w(t) = Ke^{-at} + \frac{b}{a},$$

o bien, en términos de P ,

$$P(t) = \frac{a}{b + ke^{-at}},$$

donde $k \equiv aK$. Sustituyendo la condición inicial $P(0) = P_0$, obtenemos de nuevo una función equivalente a la encontrada en la sección 3.1.2:

$$P(t) = \frac{a}{b + (\frac{a}{P_0} - b)e^{-at}}.$$

Ej 3.2.2 Se quiere resolver la ecuación

$$\dot{x} + \tan(t)x = \sec(t)x^2, \text{ con } x(0) = x_0.$$

Podemos identificar ésta como una ecuación de Bernoulli con $n = 2$. Sustituimos $w = x^{-1}$ y obtenemos la siguiente ecuación en w :

$$\dot{w} - \tan(t)w = -\sec(t).$$

Para resolverla, debemos multiplicar por el factor de integración $\rho(t) = e^{-\int_0^t \tan(s) ds}$. Sabemos que

$$\int_0^t \tan(s) ds = -\ln |\cos(t)|,$$

por lo que, el factor de integración es $\rho(t) = e^{\ln|\cos(t)|} = \cos(t)$. Al multiplicar ambos lados de la ecuación, obtenemos

$$\rho\dot{w} - \rho \tan(t)w = -\rho \sec(t)$$

$$\dot{w} \cos t - w \sin t = -1$$

$$\frac{d}{dt} [w \cos t] = -1$$

$$w \cos t = -t + c$$

$$w(t) = \frac{c - t}{\cos(t)}.$$

Puesto que $x(t) = \frac{1}{w(t)}$,

$$x(t) = \frac{\cos(t)}{c - t}$$

y como $x(0) = x_0$,

$$x(t) = \frac{x_0 \cos(t)}{1 - x_0 t}.$$

◇◇◇◇◇◇◇◇◇

§3.3 Diagramas de fase y estabilidad

En muchos casos uno se encuentra con ecuaciones diferenciales que no es posible resolver explícitamente. El análisis cualitativo es una herramienta que nos ayuda a estudiar el comportamiento del sistema a pesar

de que no tengamos una solución analítica explícita. Para llevar a cabo este tipo de análisis se utilizan los **diagramas de fase**. En el caso más sencillo, estos consisten en la gráfica de sistemas autónomos del tipo

$$\dot{x} = f(x), \quad (3.4)$$

donde los estados estacionarios, o puntos fijos, de acuerdo con la definición 2.2.2 están representados por los puntos para los cuales $\dot{x} = 0$, es decir, en donde $f(x) = 0$.

La evolución del sistema (3.4), también llamado **sistema dinámico**, depende de la condición inicial $x(0) = x_0$. La representación visual se hace en el plano $X\dot{X}$ donde las flechas indican la evolución del sistema en el tiempo, como se indica en la figura 3.2 correspondiente a la ecuación

$$\dot{x} = -x + 2. \quad (3.5)$$

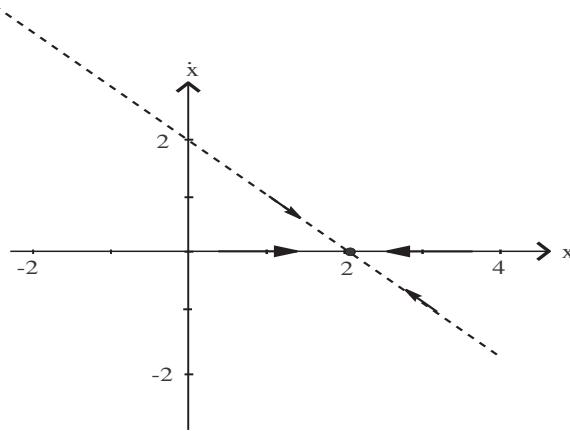


Figura 3.2: Diagrama de fase para la ecuación (3.5).

El punto $(2, 0)$ representa un equilibrio asintóticamente estable. Dado cualquier valor inicial $x(0)$, el sistema se mueve sobre el eje horizontal en la dirección de las flechas. Éstas se obtienen notando que $\dot{x} > 0$ arriba del eje x y $\dot{x} < 0$ por debajo del mismo, lo cual nos determina si la función $x(t)$ crece o decrece conforme transcurre el tiempo. Una situación más interesante se presenta al graficar

$$\dot{x} = x^3 - 2x^2 - x + 2. \quad (3.6)$$

como puede verse en la figura 3.3.

En este caso existen tres equilibrios: en $x_1 = -1$, $x_2 = 1$ y $x_3 = 2$. El primero y el tercero son inestables pero el segundo es asintóticamente estable. Se puede observar que la pendiente de la gráfica alrededor del punto de equilibrio determina la estabilidad del mismo. El equilibrio es estable si la pendiente es negativa e inestable si es positiva. La figura 3.4 muestra la trayectoria para $x(t)$ cuando $x(0) = 0$. Observamos que $\lim_{t \rightarrow -\infty} x(t) = -1$ y $\lim_{t \rightarrow \infty} x(t) = 1$.

El siguiente teorema resume estas consideraciones.

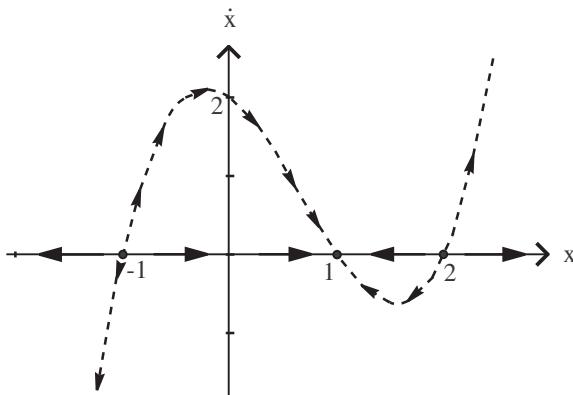


Figura 3.3: Diagrama de fase de la ecuación (3.6).

Teorema 3.3.1

Dados el sistema dinámico $\dot{x} = f(x)$ y un punto de equilibrio x^* tal que $f'(x^*) \neq 0$, se cumple

$$f'(x^*) < 0 \Leftrightarrow x^* \text{ es asintóticamente estable.}$$

Demostración (esbozo)

Sean x^* un punto fijo del sistema dinámico y $\eta(t) = x(t) - x^*$ una pequeña perturbación alrededor de x^* ; se tiene entonces que $\dot{\eta} = \dot{x} = f(x)$. Usando la aproximación lineal de Taylor de $\dot{\eta}$ alrededor de x^* tenemos que

$$\dot{\eta} \simeq f'(x^*)(x - x^*) = \eta f'(x^*). \quad (3.7)$$

Notamos que (3.7) tiene como solución $\eta = K e^{f'(x^*)t}$, con lo cual tenemos que la perturbación tiende a cero si $f'(x^*) < 0$ y diverge si $f'(x^*) > 0$. Por lo tanto si $f'(x^*) < 0$,

$$\lim_{t \rightarrow \infty} \eta(t) = \lim_{t \rightarrow \infty} (x - x^*) = 0,$$

con lo cual,

$$\lim_{t \rightarrow \infty} x(t) = x^*$$

y x^* es asintóticamente estable si $f'(x^*) < 0$. De manera análoga, si $f'(x^*) > 0$,

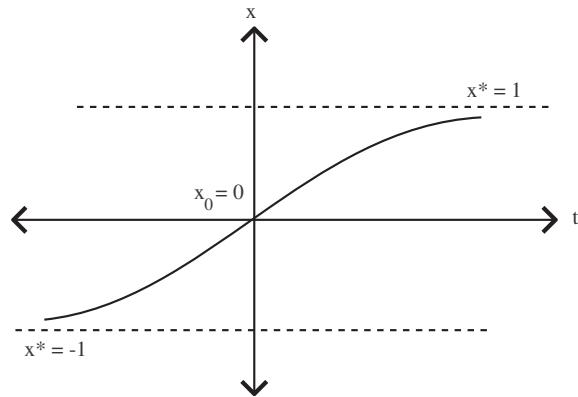
$$\lim_{t \rightarrow \infty} \eta(t) = \lim_{t \rightarrow \infty} (x - x^*) = \pm\infty$$

con lo cual x^* es inestable. ■

A continuación realizamos el análisis cualitativo de algunos modelos económicos.

§3.3.1 Modelo de Solow-Swan

El modelo de crecimiento de Solow y Swan (véase [Sol56] y [Swa56]) extiende el modelo visto en el ejemplo 2.2.3 incorporando una función de producción. Una vez más, se trata de explicar el crecimiento

Figura 3.4: Trayectoria de $x(t)$ si $x(0) = 0$.

de la producción por medio de un modelo plausible. Al igual que antes, se asume que la propensión marginal a ahorrar está dada de manera exógena. Se utiliza la siguiente notación:

- Y = producción,
- K = capital,
- L = fuerza laboral,
- S = ahorro,
- I = inversión bruta,
- δ = tasa de depreciación del capital,
- s = propensión marginal a ahorrar (constante),
- n = tasa constante de crecimiento de L , suponiendo crecimiento exponencial,
- $y = \frac{Y}{L}$ = producción per cápita, suponiendo que todos trabajan,
- $k = \frac{K}{L}$ = capital per cápita.

Se asume que $Y = F(K, L)$, donde F es una función de producción linealmente homogénea² y con las propiedades usuales: $F_K, F_L > 0$, $F_{KK}, F_{LL} < 0$, $F_{KL} > 0$, $F_K(0, L) > \delta$ y $F(0, L) = F(K, 0) = 0$.

²Una función $f(x_1, x_2, \dots, x_n)$ es homogénea de grado k si para todo $\lambda \neq 0$ se tiene

$$f(\lambda x_1, \lambda x_2, \dots, \lambda x_n) = \lambda^k f(x_1, x_2, \dots, x_n).$$

En particular, si $k = 1$ decimos que la función es linealmente homogénea.

Estas propiedades garantizan que se tiene una función creciente en cada factor y estrictamente cóncava (véase capítulo 10).

La población es igual a la fuerza laboral y crece a la tasa constante n y el ahorro es una proporción fija, s , del ingreso, de manera que se cumplen

$$\begin{aligned}\dot{L} &= nL, \\ S &= sY.\end{aligned}$$

Por definición, la inversión bruta está dada por

$$I = \dot{K} + \delta K,$$

es decir, está compuesta de la inversión neta, \dot{K} , más la inversión requerida para reponer el capital depreciado, δK . Dado que la función de producción es linealmente homogénea podemos expresar la producción per cápita como sigue:

$$y = \frac{Y}{L} = F\left(\frac{K}{L}, \frac{L}{L}\right) = F(k, 1) = f(k),$$

donde $f' > 0$, $f'' < 0$, $f'(0) > \delta$ y $f(0) = 0$. Al igual que en el modelo de Harrod-Domar se tiene que, en equilibrio, $I = S$ y por lo tanto

$$sY = \dot{K} + \delta K.$$

De aquí se obtiene $\dot{K} = sY - \delta K$ y dividiendo ambos lados entre K

$$\frac{\dot{K}}{K} = \frac{sY - \delta K}{K} = \frac{sY}{K} - \frac{\delta K}{K} = \frac{sY}{K} - \delta = \frac{sf(k)}{k} - \delta.$$

No es difícil verificar que

$$\frac{\dot{K}}{K} = \frac{\dot{k}}{k} + \frac{\dot{L}}{L} = \frac{\dot{k}}{k} + n,$$

así que sustituyendo en la ecuación anterior y reacomodando términos se obtiene finalmente

$$\dot{k} = sf(k) - (n + \delta)k. \quad (3.8)$$

El análisis cualitativo se puede realizar, a pesar de no conocer la forma específica de la función f . Nótese que la gráfica en el plano KK pasa por el origen, es cóncava y tiene un máximo en k_{\max} , donde k_{\max} es tal que $f'(k_{\max}) = \frac{n+\delta}{s}$. Asimismo, existen dos equilibrios: uno en el origen (inestable) y otro en k^* (estable), donde k^* satisface $sf(k^*) - (n + \delta)k^* = 0$. Podemos utilizar el teorema 3.3.1 para verificar este hecho. Lo importante es notar que la función $sf(k) - (n + \delta)k$ cruza al eje horizontal en dos puntos, uno con pendiente positiva y otro con pendiente negativa. Estas consideraciones pueden observarse en la figura 3.5.

La trayectoria del capital se ilustra en la figura 3.6, donde puede observarse la convergencia asintótica a k^* . Como consecuencia de esto, la producción per cápita converge a un estado estacionario dado por $y^* = f(k^*)$.

Antes de la aparición del modelo de Solow-Swan el modelo más utilizado era aquel el propuesto por Harrod y Domar (véase [Har39] y [Dom46]). Influenciados por la gran depresión de los años treinta del siglo pasado, la finalidad de este modelo era el concluir que el sistema capitalista era intrínsecamente

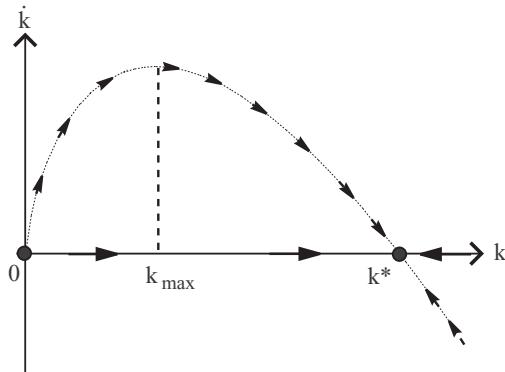


Figura 3.5: Diagrama de fase para el modelo de Solow.

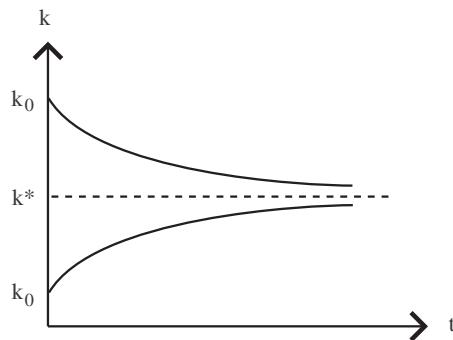


Figura 3.6: Trayectoria del capital.

inestable. Una característica de este modelo es la relación $\dot{K} = v\dot{Y}$ entre el capital y la producción utilizada en el ejemplo 2.2.3. A diferencia del modelo visto en el ejemplo mencionado, Harrod y Domar tenían en mente una función de producción con dos insumos: K y L en donde la razón capital-trabajo era constante.

El modelo de Harrod-Domar puede plantearse básicamente igual al de Solow-Swan con excepción de la función de producción, que tomamos tipo Leontief, es decir, de la forma

$$Y = F(K, L) = \min\{AK, BL\}, \quad A, B > 0.$$

Esta especificación de la función de producción implica que no hay capacidad de sustitución entre el capital y el trabajo, de manera que estos factores siempre se utilizan en las mismas proporciones. El problema radica en que, salvo para ciertos valores de los parámetros (concretamente si $sA = n + \delta$), siempre existe desempleo en alguno de los factores de producción (K o L). En el estado estacionario existirán máquinas ociosas o trabajadores desempleados cuyo número aumentará de manera sostenida. El modelo de

Solow-Swan soluciona este problema al considerar una función de producción doblemente diferenciable, en donde los productos marginales se ajustan continuamente a las cantidades de los insumos. Remitimos al lector a [SiM94] para mayor detalle sobre el modelo de Harrod-Domar.

§3.4 Expectativas y estabilidad: un modelo monetario

En un mundo dinámico con individuos racionales, se sabe que las acciones presentes tienen consecuencias futuras. Los individuos deben, de alguna forma, tratar de predecir lo que deparará el futuro. Es de esperarse que las expectativas acerca del valor futuro de variables, o parámetros, se formen de acuerdo con alguna regla “sensata”.

Los primeros modelos que incorporan expectativas suponen que éstas se forman de manera **adaptativa**, es decir, los individuos toman el valor pasado de la variable y lo corrigen de alguna forma para obtener el valor actual. Esta manera de comportarse asume que los individuos no actualizan sus conocimientos utilizando toda la información a su alcance.

Para corregir este supuesto ilógico, surgen los modelos de **expectativas racionales**. En éstos, los individuos usan toda la información pertinente y actualizan sus conocimientos utilizando la esperanza matemática (valor esperado) condicionada a la información existente. Un caso particular es cuando los individuos tienen toda la información necesaria y por lo tanto el valor esperado coincide con el valor real. Cuando esto sucede decimos que los individuos poseen **previsión perfecta**. Si el esfuerzo de formar las expectativas futuras no conlleva ningún costo, se dice que las expectativas son **racionales y acotadas**.

En esta sección vamos a analizar dos modelos monetarios: uno con expectativas adaptativas y el otro con expectativas racionales y previsión perfecta. El primero se basa en un modelo de P. Cagan [Cag56] y el segundo en uno de T. Sargent y N. Wallace [SW73].

§3.4.1 Expectativas adaptativas

Todas las variables del modelo están expresadas en logaritmos, es decir son el logaritmo (natural) de las variables originales. Abusando de la notación omitimos la palabra logaritmo del nombre de las variables. Se utiliza la siguiente notación:

- m = oferta monetaria,
- p = índice de precios,
- $\pi = \dot{p}$ = inflación,
- π^e = inflación esperada.

Suponemos que la demanda de dinero (o más precisamente, los balances nominales) por parte de los individuos es proporcional al índice de precios e inversamente proporcional a la inflación esperada; asimismo, en equilibrio, la oferta y la demanda monetaria son iguales. Dado que las variables están en logaritmos esto lo podemos expresar por la relación

$$m = p - \lambda\pi^e, \quad (3.9)$$

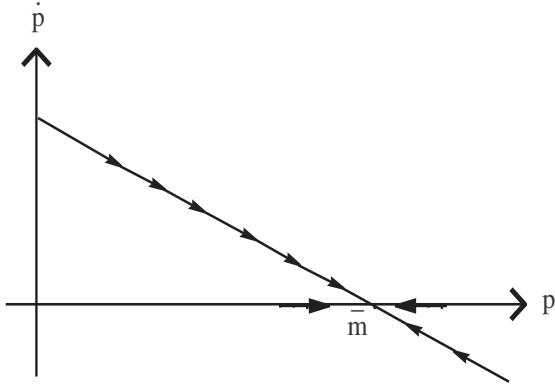


Figura 3.7: Equilibrio con expectativas adaptativas

en donde $\lambda > 0$ es una constante.

Los individuos forman sus expectativas acerca del valor de la inflación esperada de manera adaptativa de acuerdo con la siguiente relación:

$$\dot{\pi}^e = \alpha(\pi - \pi^e) = \alpha(\dot{p} - \pi^e), \quad (3.10)$$

con α una constante positiva tal que $\alpha\lambda < 1$. Es decir, actualizan el valor de la inflación esperada de acuerdo a la diferencia entre la inflación real y la inflación esperada.

Despejando π^e de (3.9) se obtiene

$$\pi^e = \frac{p - m}{\lambda}, \quad (3.11)$$

y derivando con respecto al tiempo

$$\dot{\pi}^e = \frac{\dot{p} - \dot{m}}{\lambda}. \quad (3.12)$$

Sustituyendo (3.11) y (3.12) en (3.10), despejando \dot{p} y reacomodando términos se obtiene

$$\dot{p} = \frac{1}{1 - \alpha\lambda} (\alpha(m - p) + \dot{m}). \quad (3.13)$$

Supongamos que la oferta monetaria se mantiene constante en un nivel \bar{m} , por lo que $\dot{m} = 0$. El sistema (3.13) queda entonces descrito por la ecuación diferencial

$$\dot{p} = \frac{\alpha}{1 - \alpha\lambda} (\bar{m} - p) = f_0(p).$$

El diagrama de fase en el plano $P\dot{P}$ es simplemente una recta con pendiente negativa y con un punto de equilibrio estable en $p = \bar{m}$. Esto se puede observar en la figura 3.7.

¿Qué pasaría si la autoridad monetaria sorpresivamente aumentara la oferta de dinero a $\tilde{m} > \bar{m}$? Suponemos que el índice de precios no puede cambiar discretamente (es decir, la función $p(t)$ es continua

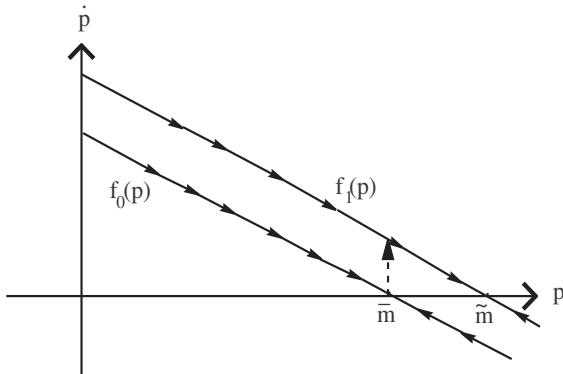


Figura 3.8: Aumento en la oferta monetaria (expectativas adaptativas).

al cambiar la oferta monetaria) pero la inflación sí puede hacerlo ($\dot{p}(t)$ no es continua). Una vez realizado el cambio en la oferta monetaria, la dinámica del sistema queda dada por la siguiente ecuación:

$$\dot{p} = \frac{\alpha}{1 - \alpha\lambda}(\tilde{m} - p) = f_1(p),$$

cuyo punto de equilibrio se encuentra en $p = \tilde{m}$, a la derecha del equilibrio anterior. En el momento del cambio, la inflación $\pi = \dot{p}$ pasa del valor 0 a un valor positivo $\bar{\pi}$ correspondiente a la condición inicial $p = \bar{m}$. Después, la inflación disminuye en el tiempo y el sistema se approxima asintóticamente hacia su nuevo valor de equilibrio $p = \tilde{m}$. Estas consideraciones pueden verse en la figura 3.8.

Notamos que la inflación esperada es negativa durante el periodo de transición, ya que $p < \tilde{m}$, a pesar de que la inflación real siempre es positiva. Los agentes, en este modelo adaptativo, consistentemente esperan una tasa de inflación menor a la inflación real.

El análisis de este modelo es típico del que se realiza con frecuencia en modelos económicos dinámicos: primero se considera algún cambio en un parámetro y se determina el nuevo equilibrio (estática comparativa); posteriormente se determina el ajuste dinámico durante el periodo de transición.

§3.4.2 Expectativas racionales

En este modelo todo es igual que en el anterior, excepto que ahora los individuos tienen expectativas racionales con previsión perfecta y por lo tanto el valor de la inflación esperada siempre coincide con el de la inflación real. De esta forma se tiene que $\pi^e = \pi$ y $\dot{\pi}^e = 0$, por lo que el modelo anterior se transforma en

$$\dot{p} = \frac{1}{\lambda}(p - m).$$

Realizando el mismo análisis que en la sección anterior y denotando por \bar{m} la oferta monetaria constante, notamos que ahora la recta tiene pendiente positiva y el equilibrio en $p = \bar{m}$ es inestable.

Si la autoridad monetaria aumenta sorpresivamente la oferta de dinero a un nivel $\tilde{m} > \bar{m}$, ¿cómo se puede llegar al nuevo nivel de equilibrio $p = \tilde{m}$? La única posibilidad es que $p(t)$ tenga una discontinuidad

dad en el momento del cambio y tome su nuevo valor de manera instantánea; es decir, no hay dinámica. La ecuación diferencial lineal con coeficientes constantes se resuelve fácilmente para obtener

$$p(t) = (p_0 - \tilde{m})e^{\frac{t}{\lambda}} + \tilde{m}.$$

El nivel de precios diverge, a menos que $p_0 = \tilde{m}$. La figura 3.9 muestra las consideraciones anteriores.

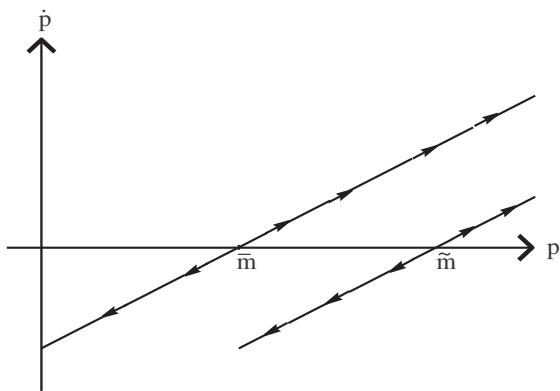


Figura 3.9: Aumento en la oferta monetaria (expectativas racionales)

En este caso, dado que el nivel de precios se ajusta en forma instantánea, no existe ningún efecto real sobre la economía: los cambios son simplemente nominales. Podemos concluir que en un mundo con individuos racionales y previsión perfecta, la política monetaria, por lo menos en este modelo, es totalmente intrascendente para la economía real.

Ejercicios

▷ **3.1** Resolver las siguientes ecuaciones con la condición inicial $x(1) = 1$:

a) $\dot{x} = \frac{t^3}{x^3}$.

b) $\dot{x} = \frac{x^3}{t^3}$.

c) $\dot{x} = x^2$.

d) $\dot{x} = 1 + x^2$.

e) $\dot{x} = \frac{\sqrt{t+1}}{x^2}$, con condición inicial $x(0) = \frac{5}{3}$.

▷ **3.2** Resolver:

a) $y'(x^2 + 1) \cos y = -x \sin y$ con $y(1) = \frac{\pi}{2}$.

b) $y'y(x+4) + y + 2 = 0$ con $y(-3) = -1$.
(Sugerencia: expresar la solución en forma implícita.)

c) $2yy' - y^2 = e^{3t}$, con $y(0) = -1$.

▷ **3.3** En un pueblito del estado de Cuévano se inicia un rumor. La población total del pueblito es de N^* personas. La tasa de cambio en el número de personas N que ha oído el rumor es proporcional al número de personas que no lo ha oido. Un modelo razonable para esta situación es $\dot{N} = kN(N^* - N)$ en donde $k > 0$.

a) Resolver el problema de valores iniciales con $N(0) = N_0$.

b) ¿Qué predice el modelo que sucederá en el futuro?

c) ¿Cuál es el nombre del pueblito?

▷ **3.4** Resolver la ecuación diferencial (3.8) del modelo de Solow, si se supone que la función de producción es del tipo Cobb-Douglas y está dada por $F(K, L) = K^\alpha L^{1-\alpha}$, y por lo tanto $f(k) = k^\alpha$, con $0 < \alpha < 1$. Encontrar $\lim_{t \rightarrow \infty} k(t)$.

▷ **3.5** El siguiente es el modelo de empleo de Haavelmo.³ Sea K el nivel de capital en una industria y L el nivel de empleo. Supongamos que se tiene una función de producción $Y = K^\gamma L^{1-\gamma}$, con $0 < \gamma < 1$. Se supone, además, que el crecimiento en el cambio en la tasa de empleo está dado por

$$\frac{\dot{L}}{L} = \alpha - \beta \frac{L}{Y} = \alpha - \beta \frac{1}{Y/L},$$

donde $\alpha, \beta > 0$. Esto es, el cambio en la tasa de empleo crece cuando la producción per cápita crece. El nivel de capital se supone constante.

a) Demostrar que, bajo las hipótesis anteriores, la ecuación para el empleo se puede escribir como

$$\dot{L} = \alpha L - \beta \frac{L^{1+\gamma}}{K^\gamma}.$$

b) Resolver la ecuación anterior si se supone que en un principio se tiene $L(0) = L_0$.

c) Encontrar $\lim_{t \rightarrow \infty} L(t)$.

▷ **3.6** En este problema resolveremos la ecuación logística con emigración dada por

$$\dot{P} = rP \left(1 - \frac{P}{C}\right) - E$$

$$\text{con } P(0) = P_0,$$

donde se supone que $E, r, C > 0$ y que $C > 4E/r$. En este caso E representa un coeficiente de emigración constante.

a) Se definen $L = \frac{1}{2} \left(C + \sqrt{C^2 - 4EC/r} \right)$ y $\alpha = \sqrt{r^2 - 4Er/C}$.

Sea $y = \frac{r}{C}(P - L)$. Demostrar que la ecuación diferencial en términos de y es de la forma $\dot{y} + \alpha y = -y^2$. (Sugerencia: nótese que $L = \frac{1}{2} \left(C + \frac{C\alpha}{r} \right)$.)

³Véase [Haa56].

- b)** Resolver la ecuación anterior para $y(t)$.
Encontrar la solución correspondiente para $P(t)$.
- c)** Encontrar $\lim_{t \rightarrow \infty} P(t)$, si se supone que $P_0 \neq C - L$.

▷ **3.7** Resolver las siguientes ecuaciones:

- a)** $\dot{x} + 2tx^2 = x$.
- b)** $\dot{x} + x = x^7$ con condición inicial $x(0) = 1$.
- c)** $t^2\dot{x} - 11tx = x^5$ con condición inicial $x(1) = 1$.
- d)** $4t^2y' - ty = y^3$, con condición inicial $y(1) = 1$.
- e)** $y' + y = xy^3$.
- f)** $y' - \frac{y}{x} = \frac{-y^2}{x}$.
- g)** $xy' + y = -2x^6y^4$.

▷ **3.8** Realizar el diagrama de fase para cada una de las siguientes ecuaciones, identificar los puntos de equilibrio y decir si son estables e inestables:

- a)** $\dot{x} = -4x^2 + 8x$.
- b)** $\dot{x} = x^3 - 15x^2 + 36x$.
- c)** $\dot{x} = e^x \sin x$.
- d)** $\dot{x} = 2x \ln \frac{K}{x}$ con $x > 0$ y $K > 0$. Este sistema se conoce como **modelo de crecimiento de Gompertz**, en donde x representa una población.

▷ **3.9** Realizar el diagrama de fase para cada una de las siguientes ecuaciones, identificar los puntos de equilibrio y analizar cómo se comporta el sistema alrededor de éstos. ¿Qué tienen de particular estos ejemplos?

- a)** $\dot{x} = (x - 2)^2$.

b) $\dot{x} = x^2(1 - x)$.

c) $\dot{x} = \frac{x^2}{2+x^2}$.

▷ **3.10 (Funciones HARA)** Sea w la riqueza. Considérese una función de utilidad $u(w)$ triplemente diferenciable, con $u' > 0$, $u'' < 0$ tal que se cumple,

$$\frac{\frac{u'''}{u''}}{\frac{u''}{u'}} = k,$$

es decir $\frac{u'''u'}{(u'')^2} = k$, en donde k es una constante.

a) Probar que se cumple

$$\frac{-u''(w)}{u'(w)} = \frac{1}{A + (k - 1)w}$$

con A una constante. Debido a esta propiedad, estas funciones se conocen como funciones de utilidad con **aversión absoluta al riesgo hiperbólica** (o bien **funciones HARA**, por sus siglas en inglés). Comparar con las funciones CARA vistas en la sección 2.3.5.

- b)** Caracterizar a las funciones de utilidad tipo HARA y dar condiciones sobre los parámetros para que estén bien definidas y cumplan con las propiedades dadas.
- c)** Probar que si $k = 0$ se obtiene una función de utilidad cuadrática, si $k = 1$ una función CARA y las funciones CRRA cumplen con $k > 1$.

▷ **3.11** Resolver los modelos de la sección 3.4 suponiendo que la autoridad monetaria tiene como política mantener una tasa constante de crecimiento para la base monetaria, es decir, $\dot{m} = \mu$ es constante.

▷ **3.12** En el ejercicio 2.15 del capítulo anterior, se estudió la evolución del precio real y del precio esperado de un activo. La ecuación (2.29), dice simplemente que los agentes forman sus expectativas acerca del precio de forma adaptativa.

- a) Suponer ahora que existe un impuesto τ sobre los dividendos, de manera que éstos se transforman en $(1 - \tau)d$. Reescribir (2.28) y (2.29) en este caso y resolver para $p(t)$ y $p^e(t)$.
- b) Realizar el análisis cualitativo si inesperadamente, en $t = T$, el impuesto aumenta a $\bar{\tau} > \tau$. Interpretar los resultados.
- c) Aparte de tener un impuesto τ sobre los dividendos, suponer ahora que las expectativas son racionales y los agentes tienen previsión perfecta. Utilizar (2.28) y (2.29) para resolver el modelo en este caso y obtener $p(t)$.
- d) Realizar de nuevo el análisis cualitativo si inesperadamente, en $t = T$, el impuesto aumenta a $\bar{\tau} > \tau$. Interpretar.

CAPÍTULO 4

Sistemas de ecuaciones diferenciales lineales

§4.1 Introducción

En este capítulo se dará una introducción general a la teoría de los sistemas de ecuaciones diferenciales lineales con más de una variable dependiente. Asumimos que el lector a llevado un curso previo de álgebra lineal; sin embargo, incluimos los conceptos básicos de valores y vectores propios a manera de repaso.

Sea t la variable independiente y $X(t) = \begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix}$ un vector de variables dependientes. Un sistema de ecuaciones diferenciales autónomo de primer orden es una expresión del tipo

$$\dot{X} = f(X), \quad (4.1)$$

donde la función $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ es continuamente diferenciable. A la función f se la conoce comúnmente como **campo vectorial**. Además, si escribimos la función como

$$f \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} f_1(x_1, \dots, x_n) \\ \vdots \\ f_n(x_1, \dots, x_n) \end{pmatrix},$$

entonces a las funciones f_1, \dots, f_n se las llama funciones coordenadas de f .

Ejemplo

Ej 4.1.1 Dado el sistema de ecuaciones

$$\dot{k} = k^\alpha - c - nk,$$

$$\dot{c} = \frac{1}{\theta}(\alpha k^{\alpha-1} - \rho)c,$$

podemos escribirlo como $\dot{X} = f(X)$, donde X es el vector $X = \begin{pmatrix} k \\ c \end{pmatrix}$ y

$$f(X) = f\left(\begin{array}{c} k \\ c \end{array}\right) = \left(\begin{array}{c} k^\alpha - c - nk \\ \frac{1}{\theta}(\alpha k^{\alpha-1} - \rho)c \end{array}\right).$$

Este sistema de ecuaciones es el llamado modelo de Ramsey de crecimiento económico. Se verá posteriormente que k representa el capital per cápita, c el consumo per cápita y α, n, θ, ρ son parámetros del sistema.



El sistema de ecuaciones diferenciales de primer orden (4.1) se puede expresar como

$$\dot{x}_1 = f_1(x_1, \dots, x_n),$$

⋮

$$\dot{x}_n = f_n(x_1, \dots, x_n).$$

Si las funciones $f_i, i = 1, 2, \dots, n$, son lineales, decimos que se trata de un sistema lineal el cual se puede escribir, en el caso homogéneo, como

$$\dot{x}_1 = a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n,$$

⋮

$$\dot{x}_n = a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n,$$

donde todos los coeficientes a_{ij} son reales.

Al considerar el caso específico de un sistema lineal podemos utilizar la notación matricial. Sea

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

la matriz de coeficientes. Un sistema lineal de ecuaciones diferenciales se escribe entonces como

$$\dot{X} = AX, \tag{4.2}$$

al cual llamaremos *el problema lineal*. En este capítulo nos dedicaremos al estudio de estos sistemas.

Ejemplos

Ej 4.1.2 Sea

$$\begin{aligned}\dot{x}_1 &= 2x_1 - 3x_2 + 5x_3, \\ \dot{x}_2 &= x_1 + 4x_2 + 7x_3, \\ \dot{x}_3 &= -x_2 + 15x_3,\end{aligned}$$

un sistema lineal de ecuaciones diferenciales. Vamos a reescribir el sistema usando la notación vectorial.

Sea $X(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix}$; entonces el sistema de ecuaciones se escribe de la forma $\dot{X} = AX$, donde

A es la siguiente matriz de coeficientes de 3×3 :

$$A = \begin{pmatrix} 2 & -3 & 5 \\ 1 & 4 & 7 \\ 0 & -1 & 15 \end{pmatrix}.$$

Ej 4.1.3 Sea $A = \begin{pmatrix} 0 & 3 \\ -3 & 0 \end{pmatrix}$. Encontrar una solución al problema lineal dado por la ecuación (4.2). Queremos encontrar una función vectorial

$$X(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix},$$

tal que $\dot{X}(t) = AX$; es decir, necesitamos dos funciones $x_1(t)$ y $x_2(t)$ tales que

$$\begin{aligned}\dot{x}_1 &= 3x_2, \\ \dot{x}_2 &= -3x_1.\end{aligned}$$

Es fácil ver que $\ddot{x}_1 = 3\dot{x}_2 = -9x_1$, de manera que buscamos una función cuya segunda derivada es igual a -9 veces ella misma. Por ejemplo, podemos proponer una solución de la forma $x_1(t) = \sin(\omega t)$, en donde ω es una constante por determinar. Es claro que $\ddot{x}_1 = -\omega^2 \sin(\omega t) = -\omega^2 x_1(t)$. Por lo tanto, podemos deducir que $\omega^2 = 9$. Escojamos $\omega = 3$. Esto implica que

$$\begin{aligned}x_1(t) &= \sin(3t), \\ x_2(t) &= \frac{1}{3}\dot{x}_1 = \cos(3t),\end{aligned}$$

son una solución. Este método es esencialmente equivalente al que se estudió anteriormente para ecuaciones lineales de una variable.



El procedimiento anterior tiene varias deficiencias. En primer lugar, hemos obtenido sólo *una* solución, aunque, como se verá más adelante, un sistema lineal tiene un número infinito de soluciones. Por

otro lado, el método que seguimos se basa en proponer la forma adecuada de la solución, es decir, en “adivinar” cómo se ve la solución, lo cual puede no ser obvio. Es claro que necesitamos una mejor manera de proceder.

Consideremos las funciones vectoriales $X_1(t), \dots, X_m(t)$ que satisfacen

$$\begin{aligned}\dot{X}_1 &= AX_1, \\ &\vdots \\ \dot{X}_m &= AX_m,\end{aligned}$$

y sea

$$X(t) = c_1X_1(t) + c_2X_2(t) + \dots + c_mX_m(t)$$

una combinación lineal cualquiera de X_1, \dots, X_m . Vamos a demostrar que $X(t)$ es solución del problema (4.2). Esto es llamado el **principio de superposición** visto con anterioridad en el capítulo 2. Tenemos que verificar que $\dot{X} = AX$. Haciendo los cálculos obtenemos lo siguiente,

$$\begin{aligned}\dot{X} &= \frac{d}{dt}(c_1X_1 + \dots + c_mX_m) \\ &= c_1\dot{X}_1 + \dots + c_m\dot{X}_m \\ &= c_1AX_1 + \dots + c_mAX_m \\ &= A(c_1X_1 + \dots + c_mX_m) \\ &= AX,\end{aligned}$$

por lo tanto $\dot{X} = AX$, es decir, $X(t)$ es solución.

§4.2 Método de valores propios

Consideremos nuevamente la ecuación (4.2). Supongamos que podemos encontrar n soluciones linealmente independientes, es decir,

$$X_1(t), \dots, X_n(t)$$

son linealmente independientes en \mathbb{R}^n para toda t y por lo tanto forman una base de ese espacio. Por el principio de superposición, cualquier combinación lineal es solución y adicionalmente, por tratarse de una base, toda solución X puede escribirse como

$$X = c_1X_1 + \dots + c_nX_n.$$

Por otro lado, dada una condición inicial $X(t_0) = X_0$, las constantes c_1, \dots, c_n pueden determinarse de manera única tal que se satisfaga,

$$c_1X_1(t_0) + \dots + c_nX_n(t_0) = X_0.$$

El esquema anterior nos da la estrategia general para resolver un sistema lineal: primero se encuentran n soluciones linealmente independientes (i.e. que formen una base en el espacio de soluciones) y luego se ajustan las constantes para satisfacer la condición inicial, si ésta existe. En esta sección se describe un método basado en las propiedades de la matriz de coeficientes, cuyo objetivo es encontrar suficientes soluciones linealmente independientes.

Definición 4.2.1 Sea A una matriz cuadrada de $n \times n$. Un **vector propio** (o **vector característico o eigenvector**) \mathbf{v} , con **valor propio** λ , es un vector, diferente del cero que satisface

$$A\mathbf{v} = \lambda\mathbf{v}.$$

Nótese que \mathbf{v} debe ser tal que $A\mathbf{v} - \lambda\mathbf{v} = \mathbf{0}$, o bien, $(A - \lambda I)\mathbf{v} = \mathbf{0}$, donde I es la matriz identidad. Se trata, entonces, de encontrar una solución no trivial (o sea, distinta de $\mathbf{0}$) para un sistema homogéneo de $n \times n$. Recordando un resultado básico de álgebra lineal, sabemos que existe una solución no trivial siempre y cuando se cumpla $\det(A - \lambda I) = 0$. A esta última ecuación se la denomina **ecuación característica asociada a la matriz A** y al polinomio $\det(A - \lambda I)$ se lo llama **polinomio característico asociado a la matriz A** y se denota por $p_A(\lambda)$. Los valores propios no son más que las raíces de este polinomio y una vez obtenidas se puede proceder a encontrar los vectores propios. El siguiente ejemplo ilustra este procedimiento.

Ejemplo

Ej 4.2.1 Encontrar los valores propios y vectores propios de la matriz

$$A = \begin{pmatrix} -2 & 1 \\ 4 & 1 \end{pmatrix}.$$

Primero calculamos el polinomio característico $p_A(\lambda)$ para luego encontrar sus raíces. Se obtiene que,

$$\begin{aligned} p_A(\lambda) &= \det \left(\begin{pmatrix} -2 & 1 \\ 4 & 1 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right) \\ &= \det \begin{pmatrix} -2 - \lambda & 1 \\ 4 & 1 - \lambda \end{pmatrix} \\ &= (-2 - \lambda)(1 - \lambda) - 4 \\ &= \lambda^2 + \lambda - 6 \end{aligned}$$

y, factorizando, nos queda $(\lambda + 3)(\lambda - 2) = 0$, con lo cual las raíces son $\lambda_1 = -3$ y $\lambda_2 = 2$. Para encontrar el vector propio correspondiente a cada valor propio λ_i , se debe encontrar un vector $\mathbf{v}_i \neq \mathbf{0}$ tal que

$$(A - \lambda_i I)\mathbf{v}_i = \mathbf{0}.$$

Si hacemos $\mathbf{v}_1 = \begin{pmatrix} a \\ b \end{pmatrix}$, entonces

$$\begin{aligned} (A - \lambda_1 I)\mathbf{v}_1 &= \left(\begin{pmatrix} -2 & 1 \\ 4 & 1 \end{pmatrix} + 3 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right) \begin{pmatrix} a \\ b \end{pmatrix} \\ &= \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \end{aligned}$$

es decir,

$$\begin{pmatrix} 1 & 1 \\ 4 & 4 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

que, por construcción, es un sistema homogéneo de ecuaciones dependientes entre sí. Concluimos que las entradas del vector \mathbf{v}_1 satisfacen la ecuación $a + b = 0$. Al escoger el valor de una de estas incógnitas, la otra queda determinada. Por ejemplo, si $a = 1$ entonces $b = -1$ y por lo tanto

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

es un vector propio con valor propio $\lambda_1 = -3$. El vector \mathbf{v}_1 es un generador del espacio de soluciones del sistema $(A - \lambda_1 I)\mathbf{v} = \mathbf{0}$; por lo tanto, es evidente que cualquier múltiplo de \mathbf{v}_1 también es un vector propio con valor propio λ_1 . De manera análoga obtenemos que

$$\mathbf{v}_2 = \begin{pmatrix} 1 \\ 4 \end{pmatrix}$$

es un vector propio con valor propio $\lambda_2 = 2$.

◇◇◇◇◇◇◇◇◇

Una pregunta natural es si existe una manera simple de encontrar el polinomio característico de una matriz. Para el caso de una matriz de 2×2 la respuesta es afirmativa y el polinomio característico puede ser calculado con base en la traza (suma de los elementos de la diagonal) y el determinante.

Dada la matriz

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

es inmediato ver que

$$\begin{aligned} p_A(\lambda) &= \det(A - \lambda I) \\ &= \det \begin{pmatrix} a - \lambda & b \\ c & d - \lambda \end{pmatrix} \end{aligned} \tag{4.3}$$

$$= \lambda^2 - \lambda \operatorname{tr}(A) + \det(A), \tag{4.4}$$

por lo que el polinomio característico queda expresado en términos de la traza y el determinante de A .

Los vectores propios también poseen una representación sencilla en el caso de una matriz A de 2×2 como la anterior. Para encontrar el vector propio $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ asociado al valor propio λ , se resuelve el sistema

$$ax_1 + bx_2 = \lambda x_1,$$

$$cx_1 + dx_2 = \lambda x_2.$$

Por construcción, las ecuaciones son dependientes y si el sistema no es originalmente diagonal tenemos que alguno de los coeficientes b o c es diferente de cero. Supongamos que $b \neq 0$; poniendo $x_1 = b$ en la primera ecuación, es fácil ver que $x_2 = \lambda - a$, de manera que $\mathbf{v} = \begin{pmatrix} b \\ \lambda - a \end{pmatrix}$ es un vector propio con valor propio λ . En el caso en que $b = 0$ y $c \neq 0$, utilizamos la segunda ecuación y obtenemos que $\begin{pmatrix} \lambda - d \\ c \end{pmatrix}$ es el vector propio buscado.

Encontrar valores y vectores propios parece ser un proceso sumamente entretenido pero, ¿cuál es su relación con la solución al problema (4.2)? El siguiente teorema contesta esta pregunta.

Teorema 4.2.2

Si \mathbf{v} es un vector propio de A con valor propio λ , entonces $X(t) = \mathbf{v}e^{\lambda t}$ es una solución al sistema dado por (4.2).

Demostración:

Tenemos que verificar que se cumple $\dot{X}(t) = AX(t)$. Por un lado, derivando $X(t) = \mathbf{v}e^{\lambda t}$, tenemos que,

$$\dot{X}(t) = \lambda e^{\lambda t} \mathbf{v}$$

y por el otro

$$AX(t) = A e^{\lambda t} \mathbf{v} = e^{\lambda t} A \mathbf{v} = e^{\lambda t} \lambda \mathbf{v},$$

y por lo tanto se cumple $\dot{X}(t) = AX(t)$. ■

Como en el caso de las ecuaciones lineales de orden mayor que 1 con coeficientes constantes, existen varios casos para las raíces del polinomio característico, o equivalentemente, para los valores propios. La similitud con las ecuaciones lineales con coeficientes constantes no es casual y más adelante precisaremos esta conexión. Las posibilidades que se tienen para los valores propios son las siguientes:

- valores propios reales y distintos,
- valores propios complejos y
- valores propios reales repetidos.

§4.2.1 Valores propios reales distintos

De la sección anterior sabemos que dados $\mathbf{v}_1, \dots, \mathbf{v}_n$ vectores propios linealmente independientes de la matriz A de $n \times n$, con valores propios $\lambda_1, \dots, \lambda_n$, entonces

$$\begin{aligned} X_1(t) &= e^{\lambda_1 t} \mathbf{v}_1, \\ &\vdots \\ X_n(t) &= e^{\lambda_n t} \mathbf{v}_n, \end{aligned}$$

son soluciones linealmente independientes, y toda solución del problema (4.2) es combinación lineal de éstas. Si la matriz A es diagonalizable, siempre puede encontrarse una base de \mathbb{R}^n de vectores propios de A y por lo tanto puede encontrarse la solución como sigue.

Dado $\dot{X} = AX$, sea $P = \begin{bmatrix} \mathbf{v}_1 & \dots & \mathbf{v}_n \end{bmatrix}$ la matriz cuyas columnas son los vectores propios de A ; entonces, si $Y = P^{-1}X$ se tiene que $X = PY$ y, por lo tanto, $\dot{Y} = P^{-1}\dot{X} = P^{-1}AX = P^{-1}APY = DY$, donde $D = P^{-1}AP$ es la matriz diagonal de valores propios.¹ Este último sistema se resuelve fácilmente como,

$$Y = \begin{pmatrix} c_1 e^{\lambda_1 t} \\ \vdots \\ c_n e^{\lambda_n t} \end{pmatrix},$$

y por lo tanto se tiene que

$$X = PY = c_1 e^{\lambda_1 t} \mathbf{v}_1 + \dots + c_n e^{\lambda_n t} \mathbf{v}_n$$

es la solución general del problema original.

Ejemplos

Ej 4.2.2 Encontremos la solución general del siguiente problema lineal:

$$\dot{X} = \begin{pmatrix} 1 & 12 \\ 3 & 1 \end{pmatrix} X.$$

Primero obtenemos los valores propios de $A = \begin{pmatrix} 1 & 12 \\ 3 & 1 \end{pmatrix}$. El polinomio característico de A es $\det(A - \lambda I) = \lambda^2 - 2\lambda - 35$. Al resolver $\lambda^2 - 2\lambda - 35 = 0$, obtenemos $\lambda_1 = 7$ y $\lambda_2 = -5$. Corresponde ahora encontrar los vectores propios. Para $\lambda_1 = 7$, resolvemos $(A - \lambda_1 I)\mathbf{v}_1 = \mathbf{0}$ para algún $\mathbf{v}_1 \neq \mathbf{0}$. Si hacemos $\mathbf{v}_1 = \begin{pmatrix} a \\ b \end{pmatrix}$, entonces

$$\begin{pmatrix} -6 & 12 \\ 3 & -6 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Obtenemos las ecuaciones $-6a + 12b = 0$ y $3a - 6b = 0$, que por construcción, son linealmente dependientes. Escogemos $b = 1$ y, por lo tanto, $a = 2$. De aquí que $\mathbf{v}_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ es vector propio con valor propio $\lambda_1 = 7$. De modo análogo, para $\lambda_2 = -5$ obtenemos $\mathbf{v}_2 = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$. Obsérvese que podríamos haber tomado

$$\bar{\mathbf{v}}_1 = \begin{pmatrix} 12 \\ 7 - 1 \end{pmatrix} = \begin{pmatrix} 12 \\ 6 \end{pmatrix}$$

y

$$\bar{\mathbf{v}}_2 = \begin{pmatrix} 12 \\ -5 - 1 \end{pmatrix} = \begin{pmatrix} 12 \\ -6 \end{pmatrix}$$

¹ El hecho de que $P^{-1}AP = D$, con D la matriz diagonal de vectores propios es un resultado elemental de álgebra lineal y remitimos al lector a cualquier texto de álgebra lineal para su demostración.

como vectores propios, de acuerdo con la simplificación que vimos para el caso de matrices de 2×2 . En conclusión,

$$X_1(t) = e^{7t} \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

y

$$X_2(t) = e^{-5t} \begin{pmatrix} -2 \\ 1 \end{pmatrix}$$

son dos soluciones linealmente independientes y cualquier otra solución es de la forma,

$$X(t) = c_1 X_1(t) + c_2 X_2(t),$$

esto es,

$$X(t) = c_1 e^{7t} \begin{pmatrix} 2 \\ 1 \end{pmatrix} + c_2 e^{-5t} \begin{pmatrix} -2 \\ 1 \end{pmatrix}.$$

Si además pedimos que se satisfaga la condición inicial

$$X(0) = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

se pueden encontrar c_1 y c_2 de la siguiente manera:

$$X(0) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = c_1 \begin{pmatrix} 2 \\ 1 \end{pmatrix} + c_2 \begin{pmatrix} -2 \\ 1 \end{pmatrix},$$

con lo cual se tiene el sistema

$$2c_1 - 2c_2 = 0,$$

$$c_1 + c_2 = 1,$$

y por lo tanto $c_1 = c_2 = 1/2$. La solución es entonces

$$\begin{aligned} X(t) &= \frac{1}{2} e^{7t} \begin{pmatrix} 2 \\ 1 \end{pmatrix} + \frac{1}{2} e^{-5t} \begin{pmatrix} -2 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} e^{7t} - e^{-5t} \\ \frac{1}{2}(e^{7t} + e^{-5t}) \end{pmatrix}. \end{aligned}$$

Si $X(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$, entonces el sistema original está dado por

$$\dot{x} = x + 12y,$$

$$\dot{y} = 3x + y,$$

y la solución se puede expresar alternativamente como

$$x(t) = e^{7t} - e^{-5t},$$

$$y(t) = \frac{1}{2}(e^{7t} + e^{-5t}).$$

Ej 4.2.3 Queremos encontrar tres soluciones linealmente independientes del sistema,

$$\dot{X} = \begin{pmatrix} 1 & -1 & 4 \\ 3 & 2 & -1 \\ 2 & 1 & -1 \end{pmatrix} X.$$

Sea

$$A = \begin{pmatrix} 1 & -1 & 4 \\ 3 & 2 & -1 \\ 2 & 1 & -1 \end{pmatrix}.$$

Como primer paso, encontramos las raíces del polinomio $p_A(\lambda)$. Desarrollaremos el determinante por cofactores. Para hacer esto, escogemos un renglón o columna y reducimos usando los signos adecuados. Obtenemos así que

$$\begin{aligned} p_A(\lambda) &= \det \begin{pmatrix} 1-\lambda & -1 & 4 \\ 3 & 2-\lambda & -1 \\ 2 & 1 & -1-\lambda \end{pmatrix} \\ &= (1-\lambda) \det \begin{pmatrix} 2-\lambda & -1 \\ 1 & -1-\lambda \end{pmatrix} - \\ &\quad (-1) \det \begin{pmatrix} 3 & -1 \\ 2 & -1-\lambda \end{pmatrix} + 4 \det \begin{pmatrix} 3 & 2-\lambda \\ 2 & 1 \end{pmatrix} \\ &= (1-\lambda)(\lambda^2 - \lambda - 1) + (-3 - 3\lambda + 2) + 4(3 + 4 + 2\lambda) \\ &= (1-\lambda)(\lambda^2 - \lambda - 1) + 5\lambda - 5 \\ &= (1-\lambda)(\lambda^2 - \lambda - 1) - 5(1-\lambda) \\ &= (1-\lambda)(\lambda^2 - \lambda - 6) \\ &= (1-\lambda)(\lambda - 3)(\lambda + 2). \end{aligned}$$

En conclusión, el polinomio característico de A es

$$p_A(\lambda) = (1-\lambda)(\lambda - 3)(\lambda + 2)$$

y por lo tanto los valores propios de A son $\lambda_1 = 1$, $\lambda_2 = 3$ y $\lambda_3 = -2$. Como segundo paso, encontramos los vectores propios correspondientes. Si $\lambda_1 = 1$, resolvemos el sistema $(A - \lambda I)\mathbf{v} = \mathbf{0}$. Esto es, si

$$\mathbf{v} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}, \text{ entonces}$$

$$\begin{pmatrix} 0 & -1 & 4 \\ 3 & 1 & -1 \\ 2 & 1 & -2 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Al resolver estas ecuaciones y escoger $a = -1$, obtenemos $b = 4$ y $c = 1$. En conclusión,

$$\mathbf{v}_1 = \begin{pmatrix} -1 \\ 4 \\ 1 \end{pmatrix}$$

es vector propio con valor propio $\lambda_1 = 1$. Para verificarlo consideramos

$$A\mathbf{v}_1 = \begin{pmatrix} 1 & -1 & 4 \\ 3 & 2 & -1 \\ 2 & 1 & -1 \end{pmatrix} \begin{pmatrix} -1 \\ 4 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 4 \\ 1 \end{pmatrix} = \lambda_1 \mathbf{v}_1.$$

Del modo semejante se obtiene que $\mathbf{v}_2 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$ es vector propio de $\lambda_2 = 3$ y $\mathbf{v}_3 = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$ es un vector propio correspondiente a $\lambda_3 = -2$. Podemos ahora obtener tres soluciones linealmente independientes del problema dadas por

$$\begin{aligned} X_1(t) &= e^t \begin{pmatrix} -1 \\ 4 \\ 1 \end{pmatrix}, \\ X_2(t) &= e^{3t} \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \\ X_3(t) &= e^{-2t} \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}. \end{aligned}$$

Finalmente, por el principio de superposición, cualquier solución se puede escribir como

$$X(t) = c_1 X_1(t) + c_2 X_2(t) + c_3 X_3(t).$$

Ej 4.2.4 Consideremos el sistema

$$\dot{\mathbf{X}} = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} \mathbf{X}.$$

Notemos que la matriz del sistema es singular y que por lo tanto $\lambda = 0$ es una raíz del polinomio característico. En efecto, el polinomio está dado por

$$p_A(\lambda) = \lambda(\lambda - 3),$$

con raíces $\lambda_1 = 3$ y $\lambda_2 = 0$. Es fácil ver que $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ es un vector propio asociado al valor propio $\lambda_1 = 3$ y que $\mathbf{v}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ es un vector no nulo que satisface la ecuación $A\mathbf{v} = \mathbf{0}$, o sea, es un vector propio con valor propio $\lambda = 0$. La solución queda entonces como

$$X(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = K_1 \begin{pmatrix} 1 \\ 2 \end{pmatrix} e^{3t} + K_2 \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$



§4.2.2 Valores propios complejos

Dado que el polinomio característico tiene coeficientes reales, los valores propios complejos siempre aparecen en parejas de complejos conjugados, $\lambda = \alpha + \beta i$ y $\bar{\lambda} = \alpha - \beta i$. Los vectores propios correspondientes son también conjugados en el siguiente sentido: si $\mathbf{v} = \mathbf{r} + \mathbf{s}i$ es el vector propio asociado a λ , entonces $\bar{\mathbf{v}} = \mathbf{r} - \mathbf{s}i$ es el vector propio asociado a $\bar{\lambda}$, donde \mathbf{r} y \mathbf{s} denotan vectores en \mathbb{R}^n . Por analogía con el caso escalar llamamos a \mathbf{r} y \mathbf{s} las partes real e imaginaria del vector \mathbf{v} y las denotamos por $\text{Re}(\mathbf{v})$ e $\text{Im}(\mathbf{v})$, respectivamente.

Al igual que en la sección anterior, sabemos que $e^{\lambda t}\mathbf{v}$ y $e^{\bar{\lambda}t}\bar{\mathbf{v}}$ son soluciones linealmente independientes; sin embargo, al igual que en (2.19), queremos soluciones reales. El siguiente teorema nos indica cómo encontrar las soluciones reales.

Teorema 4.2.3

Sea $\mathbf{v} = \mathbf{r} + \mathbf{s}i$ un vector propio de la matriz A con valor propio $\lambda = \alpha + \beta i$ y sean

$$\begin{aligned}\mathbf{w}_1(t) &= \text{Re}(e^{\lambda t}\mathbf{v}), \\ \mathbf{w}_2(t) &= \text{Im}(e^{\lambda t}\mathbf{v}).\end{aligned}$$

Entonces, \mathbf{w}_1 y \mathbf{w}_2 son soluciones linealmente independientes del problema $\dot{\mathbf{X}} = A\mathbf{X}$ y se pueden expresar como

$$\mathbf{w}_1(t) = e^{\alpha t}(\mathbf{r} \cos \beta t - \mathbf{s} \sin \beta t), \quad (4.5)$$

$$\mathbf{w}_2(t) = e^{\alpha t}(\mathbf{r} \sin \beta t + \mathbf{s} \cos \beta t). \quad (4.6)$$

Demostración

Sabemos que $\mathbf{w}(t) = e^{\lambda t}\mathbf{v}$ es una solución del problema lineal (4.2), de manera que se cumple $\dot{\mathbf{w}} = A\mathbf{w}$. Podemos escribir $\mathbf{w} = \mathbf{w}_1 + i\mathbf{w}_2$ y claramente se tiene que $\dot{\mathbf{w}} = \dot{\mathbf{w}}_1 + i\dot{\mathbf{w}}_2$, así como también $A\mathbf{w} = A\mathbf{w}_1 + A_i\mathbf{w}_2$. Comparando ambos lados se tiene que $\dot{\mathbf{w}}_1 = A\mathbf{w}_1$ y $\dot{\mathbf{w}}_2 = A\mathbf{w}_2$, es decir, \mathbf{w}_1 y \mathbf{w}_2 son soluciones. Asimismo,

$$\begin{aligned}e^{\lambda t}\mathbf{v} &= e^{\alpha t}(\cos \beta t + i \sin \beta t)(\mathbf{r} + i\mathbf{s}) \\ &= e^{\alpha t}(\mathbf{r} \cos \beta t + i\mathbf{s} \cos \beta t + i\mathbf{r} \sin \beta t - \mathbf{s} \sin \beta t) \\ &= e^{\alpha t}[(\mathbf{r} \cos \beta t - \mathbf{s} \sin \beta t) + i(\mathbf{r} \sin \beta t + \mathbf{s} \cos \beta t)],\end{aligned}$$

con lo cual se tiene que

$$\begin{aligned}\text{Re}(e^{\lambda t}\mathbf{v}) &= e^{\alpha t}(\mathbf{r} \cos \beta t - \mathbf{s} \sin \beta t) = \mathbf{w}_1(t), \\ \text{Im}(e^{\lambda t}\mathbf{v}) &= e^{\alpha t}(\mathbf{r} \sin \beta t + \mathbf{s} \cos \beta t) = \mathbf{w}_2(t),\end{aligned}$$

que es lo que se quería demostrar. ■

Obsérvese que en el teorema anterior el resultado sería idéntico si se tomaran $\bar{\mathbf{v}}$ y $\bar{\lambda}$ en lugar de \mathbf{v} y λ . Concluimos que a cada pareja de valores y vectores propios, que son complejos conjugados,

le corresponden dos soluciones reales y linealmente independientes de la ecuación $\dot{X} = AX$. De esta forma, se pueden encontrar suficientes soluciones linealmente independientes que generen la solución general.

En el caso particular de un sistema de 2×2 , la solución general está dada por,

$$X = C_1 \mathbf{w}_1 + C_2 \mathbf{w}_2,$$

que se puede reescribir como,

$$X = e^{\alpha t} [(C_1 \cos \beta t + C_2 \sin \beta t) \mathbf{r} + (C_2 \cos \beta t - C_1 \sin \beta t) \mathbf{s}]. \quad (4.7)$$

Ejemplos

Ej 4.2.5 Queremos resolver el sistema $\dot{X} = AX$, donde A es la matriz $\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$. Obsérvese que esto es equivalente a resolver el sistema

$$\begin{aligned} \dot{x} &= x - y, \\ \dot{y} &= x + y, \end{aligned}$$

con

$$X(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}.$$

El polinomio característico de A es,

$$p_A(\lambda) = \det(A - \lambda I) = \lambda^2 - 2\lambda + 2.$$

Las raíces de $p_A(\lambda)$ (i.e. los valores propios de A) son

$$\begin{aligned} \lambda_{1,2} &= \frac{2 \pm \sqrt{4 - 8}}{2} \\ &= \frac{2 \pm 2i}{2} = 1 \pm i, \end{aligned}$$

por lo que $\lambda_1 = \lambda = 1 + i$ y $\lambda_2 = \bar{\lambda} = 1 - i$ y tenemos que $\alpha = \beta = 1$. Para $\lambda = 1 + i$, el vector propio correspondiente se encuentra resolviendo

$$(A - \lambda I)\mathbf{v} = \mathbf{0},$$

para algún $\mathbf{v} = \begin{pmatrix} a \\ b \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$; es decir,

$$\begin{pmatrix} -i & -1 \\ 1 & -i \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Esto corresponde al sistema de ecuaciones (dependiente, por construcción)

$$-ia - b = 0,$$

$$a - ib = 0.$$

Si escogemos $b = 1$, entonces $a = i$ y concluimos que \mathbf{v} es un vector propio con valor propio $\lambda = 1 + i$, donde \mathbf{v} se puede escribir como

$$\mathbf{v} = \begin{pmatrix} i \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} + i \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

En consecuencia, para $\bar{\lambda} = 1 - i$ se tiene que el vector propio correspondiente está dado por,

$$\bar{\mathbf{v}} = \begin{pmatrix} -i \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} - i \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

En este caso, $\mathbf{r} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ y $\mathbf{s} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, con lo que, de acuerdo a (4.5) y (4.6) dos soluciones independientes, \mathbf{w}_1 y \mathbf{w}_2 , están dadas por

$$\begin{aligned} \mathbf{w}_1 &= e^t \left(\begin{pmatrix} 0 \\ 1 \end{pmatrix} \cos t - \begin{pmatrix} 1 \\ 0 \end{pmatrix} \sin t \right) = \begin{pmatrix} -e^t \sin t \\ e^t \cos t \end{pmatrix}, \\ \mathbf{w}_2 &= e^t \left(\begin{pmatrix} 0 \\ 1 \end{pmatrix} \sin t + \begin{pmatrix} 1 \\ 0 \end{pmatrix} \cos t \right) = \begin{pmatrix} e^t \cos t \\ e^t \sin t \end{pmatrix}, \end{aligned}$$

y la solución general del sistema queda expresada como

$$X(t) = C_1 \begin{pmatrix} -e^t \sin t \\ e^t \cos t \end{pmatrix} + C_2 \begin{pmatrix} e^t \cos t \\ e^t \sin t \end{pmatrix},$$

o bien, si $X(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$,

$$\begin{aligned} x(t) &= e^t(-C_1 \sin t + C_2 \cos t), \\ y(t) &= e^t(C_1 \cos t + C_2 \sin t). \end{aligned}$$

Si tenemos la condición inicial $X(0) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, entonces, sustituyendo en la solución general se obtiene $C_2 = 0$ y $C_1 = 1$.

Ej 4.2.6 Encontremos la solución general del siguiente problema:

$$\dot{X} = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 13 & 0 \\ 1 & 0 & 1 \end{pmatrix} X.$$

La matriz de coeficientes es

$$A = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 13 & 0 \\ 1 & 0 & 1 \end{pmatrix},$$

cuyo polinomio característico es

$$\begin{aligned} p_A(\lambda) &= \det(A - \lambda I) \\ &= \det \begin{pmatrix} 1 - \lambda & 0 & -1 \\ 0 & 13 - \lambda & 0 \\ 1 & 0 & 1 - \lambda \end{pmatrix} \\ &= (13 - \lambda) \det \begin{pmatrix} 1 - \lambda & -1 \\ 1 & 1 - \lambda \end{pmatrix} \\ &= (13 - \lambda)(\lambda^2 - 2\lambda + 2). \end{aligned}$$

Las raíces de $p_A(\lambda)$ son $\lambda_1 = \lambda = 1 + i$, $\lambda_2 = \bar{\lambda} = 1 - i$ y $\lambda_3 = 13$. Para $\lambda = 1 + i$, deseamos encontrar un vector propio $\mathbf{v} \neq \mathbf{0}$. Para hacer esto, resolvemos $(A - \lambda I)\mathbf{v} = \mathbf{0}$, es decir,

$$\begin{pmatrix} -i & 0 & -1 \\ 0 & 12 - i & 0 \\ 1 & 0 & -i \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

que equivale al sistema

$$\begin{aligned} -ia - c &= 0, \\ (12 - i)b &= 0, \\ a - ic &= 0. \end{aligned}$$

Obviamente, $b = 0$. Escogemos $a = 1$ y por consiguiente $c = -i$. Entonces

$$\mathbf{v} = \begin{pmatrix} 1 \\ 0 \\ -i \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + i \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}$$

es un vector propio con valor propio $\lambda = 1 + i$. De lo anterior tenemos que $\alpha = \beta = 1$, $\mathbf{r} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ y $\mathbf{s} = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}$. A partir de esto, (4.5) y (4.6) dan dos soluciones linealmente independientes,

$$\mathbf{w}_1(t) = e^{\alpha t} (\mathbf{r} \cos \beta t - \mathbf{s} \sin \beta t),$$

$$\mathbf{w}_2(t) = e^{\alpha t} \mathbf{r} (\sin \beta t + \mathbf{s} \cos \beta t),$$

que explícitamente son

$$\begin{aligned}\mathbf{w}_1(t) &= e^t \left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \cos t - \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} \sin t \right), \\ \mathbf{w}_2(t) &= e^t \left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \sin t + \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} \cos t \right).\end{aligned}$$

Simplificando, llegamos a que

$$\mathbf{w}_1(t) = \begin{pmatrix} e^t \cos t \\ 0 \\ e^t \sin t \end{pmatrix}$$

y

$$\mathbf{w}_2(t) = \begin{pmatrix} e^t \sin t \\ 0 \\ -e^t \cos t \end{pmatrix}.$$

Para $\lambda_3 = 13$, encontramos un vector $\mathbf{v}_3 \neq \mathbf{0}$, tal que se satisfaga

$$(A - \lambda_3 I)\mathbf{v}_3 = \mathbf{0},$$

es decir,

$$\begin{pmatrix} -12 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & -12 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Resolviendo, encontramos $\mathbf{v}_3 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ y esto nos da una solución de la forma

$$\mathbf{w}_3(t) = e^{13t} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

Por lo tanto, la solución general para este sistema es

$$X(t) = C_1 \begin{pmatrix} e^t \cos t \\ 0 \\ e^t \sin t \end{pmatrix} + C_2 \begin{pmatrix} e^t \sin t \\ 0 \\ -e^t \cos t \end{pmatrix} + C_3 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$



§4.2.3 Valores propios reales repetidos

El caso general de una matriz A de $n \times n$ se puede ver en el apéndice C; aquí se analiza únicamente el caso de una matriz de 2×2 . El problema, en el caso de raíces repetidas, es que no se tiene una base de vectores propios y por lo tanto la matriz no puede ser diagonalizada. Sin embargo, se puede obtener una matriz triangular de la forma

$$T = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}, \quad (4.8)$$

con la cual el problema queda expresado como,

$$\begin{aligned} \dot{x} &= \lambda x + y, \\ \dot{y} &= \lambda y. \end{aligned}$$

Podemos resolver la segunda de estas ecuaciones para posteriormente sustituir en la primera y llegar a

$$\begin{aligned} y &= c_2 e^{\lambda t}, \\ x &= c_1 e^{\lambda t} + c_2 t e^{\lambda t}. \end{aligned}$$

Para obtener la matriz T , lo que se necesita es el vector propio \mathbf{v} correspondiente al valor propio λ y otro vector \mathbf{w} tal que satisface que la matriz

$$P = \begin{bmatrix} \mathbf{v} & \mathbf{w} \end{bmatrix}$$

cumple $P^{-1}AP = T$, donde T es la matriz triangular dada en (4.8). Para obtener el vector \mathbf{w} se procede como sigue.

Definición 4.2.4 *Sea \mathbf{v} un vector propio con valor propio, r . Se dice que \mathbf{w} es un vector propio generalizado si satisface*

$$(A - \lambda I)\mathbf{w} = \mathbf{v}.$$

Dados \mathbf{v} y \mathbf{w} como arriba, podemos probar la siguiente proposición.

Proposición 4.2.5 *$X(t) = e^{\lambda t}(t\mathbf{v} + \mathbf{w})$ es solución del sistema $\dot{X} = AX$.*

Demostración

Por definición, tenemos que

$$\begin{aligned} \dot{X} &= e^{\lambda t}\mathbf{v} + re^{\lambda t}(t\mathbf{v} + \mathbf{w}) \\ &= e^{\lambda t}(t\lambda\mathbf{v} + \mathbf{v} + \lambda\mathbf{w}) \\ &= e^{\lambda t}(tA\mathbf{v} + A\mathbf{w}) \\ &= Ae^{\lambda t}(t\mathbf{v} + \mathbf{w}) \\ &= AX \end{aligned}$$

y por lo tanto, X es una solución. ■

De esta forma se tienen dos soluciones independientes y se concluye que la solución general del sistema $\dot{X} = AX$, en donde A es una matriz de 2×2 con un valor propio repetido, puede expresarse en términos de \mathbf{v} y \mathbf{w} como sigue:

$$X(t) = c_1 \mathbf{v} e^{\lambda t} + c_2(t\mathbf{v} + \mathbf{w})e^{\lambda t}. \quad (4.9)$$

El siguiente paso es obtener un vector propio generalizado. Se puede, por supuesto, simplemente utilizar la definición y resolver el sistema $(A - \lambda I)\mathbf{w} = \mathbf{v}$; sin embargo, el caso de una matriz de 2×2 se puede simplificar aún más. Recordemos que si la matriz A está dada por

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

y $b \neq 0$, entonces un vector propio asociado al valor propio λ está dado por $\mathbf{v} = \begin{pmatrix} b \\ \lambda - a \end{pmatrix}$. El valor propio es una raíz del polinomio $\lambda^2 - \lambda \text{tr}(A) + \det(A)$ y si es una raíz doble debe ser de la forma $\lambda = \frac{\text{tr}(A)}{2} = \frac{a+d}{2}$. Resolvemos ahora el sistema $(A - \lambda I)\mathbf{w} = \mathbf{v}$ con estos valores específicos de \mathbf{v} y λ . Si $\mathbf{w} = \begin{pmatrix} x \\ y \end{pmatrix}$, se tiene que

$$\begin{pmatrix} \frac{a-d}{2} & b \\ c & \frac{d-a}{2} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} b \\ \frac{d-a}{2} \end{pmatrix}$$

y por lo tanto es fácil ver que $\mathbf{w} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ siempre es solución. Si en la matriz A , $b = 0$ pero $c \neq 0$, entonces se utiliza $\mathbf{v} = \begin{pmatrix} \lambda - d \\ c \end{pmatrix}$ como vector propio y procediendo de manera análoga tenemos que $\mathbf{w} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ es un vector propio generalizado correspondiente.

Ejemplo

Ej 4.2.7 Resolvamos el sistema

$$\dot{X} = \begin{pmatrix} 3 & -2 \\ 2 & -1 \end{pmatrix} X.$$

La ecuación característica es $\lambda^2 - 2\lambda + 1 = 0$, cuya única raíz es $\lambda = 1$. El vector propio correspondiente es $\mathbf{v} = \begin{pmatrix} -2 \\ -2 \end{pmatrix}$ y el vector propio generalizado, que se encuentra resolviendo

$$\begin{pmatrix} 2 & -2 \\ 2 & -2 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} -2 \\ -2 \end{pmatrix},$$

es simplemente $\mathbf{w} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Por lo tanto, la solución general queda dada por

$$X(t) = C_1 e^t \begin{pmatrix} -2 \\ -2 \end{pmatrix} + C_2 e^t \begin{pmatrix} -2t \\ -2t+1 \end{pmatrix}.$$

Si se tiene la condición inicial $X(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, entonces, resolviendo el sistema

$$\begin{aligned} -2C_1 &= 1, \\ -2C_1 + C_2 &= 0, \end{aligned}$$

llegamos a $C_1 = -\frac{1}{2}$ y $C_2 = -1$, y por lo tanto la solución final es

$$\begin{aligned} X(t) &= e^t \begin{pmatrix} 1 \\ 1 \end{pmatrix} + e^t \begin{pmatrix} 2t \\ 2t-1 \end{pmatrix} \\ &= e^t \begin{pmatrix} 1+2t \\ 2t \end{pmatrix}. \end{aligned}$$

◇◇◇◇◇◇◇◇◇

§4.3 El caso no homogéneo

Consideremos ahora un sistema de la forma

$$\dot{X} = AX + B, \quad (4.10)$$

en donde

$$B = \begin{pmatrix} k_1 \\ \vdots \\ k_n \end{pmatrix}$$

es un vector fijo en \mathbb{R}^n . Sabemos que la solución a este sistema es de la forma

$$X = X_h + X_p,$$

en donde X_h es la solución general del sistema homogéneo asociado (es trivial ver que la proposición 2.2.1 se aplica también a sistemas de ecuaciones), que se obtiene con los métodos de las secciones anteriores, y X_p es una solución particular del sistema no homogéneo. La solución particular puede obtenerse fácilmente suponiendo que X_p es constante, de manera que $\dot{X}_p = 0$, para después resolver el sistema dado por

$$AX_p = -B.$$

La solución puede no existir ya que A puede ser una matriz singular. Véase el ejercicio 4.9.

Ejemplo

Ej 4.3.1 Queremos encontrar la solución general del siguiente problema lineal:

$$\dot{X} = \begin{pmatrix} 1 & 12 \\ 3 & 1 \end{pmatrix} X + \begin{pmatrix} -25 \\ -5 \end{pmatrix}.$$

La solución del caso homogéneo se obtuvo en el ejemplo 4.2.2 como

$$X_h = c_1 e^{7t} \begin{pmatrix} 2 \\ 1 \end{pmatrix} + c_2 e^{-5t} \begin{pmatrix} -2 \\ 1 \end{pmatrix}.$$

Para la solución particular se resuelve

$$\begin{pmatrix} 1 & 12 \\ 3 & 1 \end{pmatrix} X_p = \begin{pmatrix} 25 \\ 5 \end{pmatrix},$$

obteniendo $X_p = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, de manera que la solución general queda expresada como

$$X(t) = c_1 e^{7t} \begin{pmatrix} 2 \\ 1 \end{pmatrix} + c_2 e^{-5t} \begin{pmatrix} -2 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

◇◇◇◇◇◇◇◇◇

También es posible resolver sistemas no homogéneos del tipo

$$\dot{X}(t) = AX(t) + Y(t),$$

en donde $Y(t) = y(t) \begin{pmatrix} a \\ b \end{pmatrix}$ y la función $y(t)$ es polinomial, exponencial o trigonométrica. La solución es de la forma $X = X_h + X_p$, al igual que antes, y la solución particular X_p se obtiene utilizando la técnica de coeficientes indeterminados que vimos en la sección 2.3.4.1. El siguiente ejemplo ilustra el método.

Ejemplo

Ej 4.3.2 Consideremos la ecuación $\dot{X} = \begin{pmatrix} 1 & 12 \\ 3 & 1 \end{pmatrix} X + \begin{pmatrix} 2t \\ t \end{pmatrix}$, en donde la ecuación homogénea asociada es la misma del ejemplo anterior. La solución general de la ecuación homogénea es

$$X_h = c_1 e^{7t} \begin{pmatrix} 2 \\ 1 \end{pmatrix} + c_2 e^{-5t} \begin{pmatrix} -2 \\ 1 \end{pmatrix}.$$

Proponemos una solución particular de la misma forma que $Y(t)$, es decir, $X_p = \begin{pmatrix} A_1 \\ B_1 \end{pmatrix} t + \begin{pmatrix} A_0 \\ B_0 \end{pmatrix}$. Dado que X_p satisface la ecuación original, tenemos que debe cumplirse

$$\begin{aligned} \begin{pmatrix} A_1 \\ B_1 \end{pmatrix} &= \begin{pmatrix} 1 & 12 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} A_1 t + A_0 \\ B_1 t + B_0 \end{pmatrix} + \begin{pmatrix} 2t \\ t \end{pmatrix} \\ &= \begin{pmatrix} (A_1 + 12B_1 + 2)t + A_0 + 12B_0 \\ (3A_1 + B_1 + 1)t + 3A_0 + B_0 \end{pmatrix}. \end{aligned}$$

Igualando coeficientes en ambos lados se tiene

$$\begin{aligned} A_1 &= A_0 + 12B_0, \\ B_1 &= 3A_0 + B_0, \\ A_1 + 12B_1 + 2 &= 0, \\ 3A_1 + B_1 + 1 &= 0 \end{aligned}$$

y resolviendo obtenemos

$$\begin{aligned} A_0 &= -\frac{2}{49}, \quad A_1 = -\frac{2}{7}, \\ B_0 &= -\frac{5}{49}, \quad B_1 = -\frac{1}{7}, \end{aligned}$$

con lo cual la solución final es

$$X(t) = c_1 e^{7t} \begin{pmatrix} 2 \\ 1 \end{pmatrix} + c_2 e^{-5t} \begin{pmatrix} -2 \\ 1 \end{pmatrix} + \begin{pmatrix} -\frac{2}{7} \\ -\frac{1}{7} \end{pmatrix} t + \begin{pmatrix} -\frac{2}{49} \\ -\frac{5}{49} \end{pmatrix}.$$

◇◇◇◇◇◇◇◇◇

En el caso anterior, si $y(t)$ no es polinomial, exponencial ni trigonométrica, entonces el problema se complica. Para una solución general del problema no homogéneo, véase el ejercicio 4.9.

§4.4 Ecuaciones lineales de orden superior

Un sistema de ecuaciones lineales puede convertirse en una ecuación de orden superior con coeficientes constantes mediante sustituciones sucesivas. Para ilustrar esto, veamos el siguiente ejemplo de un sistema de 2×2 .

Ej 4.4.1 Consideremos el sistema

$$\begin{aligned} \dot{x} &= ax + by, \\ \dot{y} &= cx + dy. \end{aligned}$$

Despejando y de la primera ecuación, obtenemos

$$y = \frac{1}{b}\dot{x} - \frac{a}{b}x; \tag{4.11}$$

derivando esta expresión con respecto al tiempo, se tiene

$$\dot{y} = \frac{1}{b}\ddot{x} - \frac{a}{b}\dot{x} \quad (4.12)$$

y sustituyendo (4.11) y (4.12) en la segunda ecuación del sistema se obtiene

$$\frac{1}{b}\ddot{x} - \frac{a}{b}\dot{x} = cx + d\left(\frac{1}{b}\dot{x} - \frac{a}{b}x\right),$$

o bien, reescribiendo,

$$\ddot{x} - (a + d)\dot{x} + (ad - bc)x = 0.$$

Nótese que ésta es una ecuación de segundo orden con coeficientes constantes cuya ecuación característica es precisamente $p_A(\lambda) = 0$, con $p_A(\lambda)$ el polinomio característico de la matriz del sistema, $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

Supongamos, por simplicidad, que se tienen dos raíces reales distintas, λ_1 y λ_2 , las cuales, por construcción, coinciden con los valores propios de la matriz. Podemos así resolver para x y posteriormente sustituir en (4.11) para obtener y . La solución queda como

$$\begin{aligned} x &= K_1 e^{\lambda_1 t} + K_2 e^{\lambda_2 t} \\ y &= K_1 \left(\frac{\lambda_1 - a}{b} \right) e^{\lambda_1 t} + K_2 \left(\frac{\lambda_2 - a}{b} \right) e^{\lambda_2 t} \end{aligned}$$

o bien, en forma vectorial

$$\begin{pmatrix} x \\ y \end{pmatrix} = K_1 \begin{pmatrix} 1 \\ \frac{\lambda_1 - a}{b} \end{pmatrix} e^{\lambda_1 t} + K_2 \begin{pmatrix} 1 \\ \frac{\lambda_2 - a}{b} \end{pmatrix} e^{\lambda_2 t}.$$

El vector $\begin{pmatrix} 1 \\ \frac{\lambda_i - a}{b} \end{pmatrix}$ es simplemente el vector propio correspondiente al valor propio λ_i . En resumen, dado el sistema de dos ecuaciones, éste se puede resolver transformándolo en una ecuación de segundo orden. En general, para sistemas de orden mayor este método de transformar el sistema en una ecuación de orden superior puede resultar sumamente tedioso por lo cual simplemente se resuelve el sistema extendiendo el método del capítulo anterior.



A continuación veremos que la situación inversa también es válida, o sea que una ecuación de orden superior con coeficientes constantes puede transformarse en un sistema de ecuaciones lineales de primer orden. Extenderemos así la teoría del capítulo 2 para ecuaciones de orden superior.

Una ecuación de orden m con coeficientes constantes es de la forma

$$y^{(m)} + a_{m-1} y^{(m-1)} + \dots + a_1 \dot{y} + a_0 y = 0.$$

Para resolverla, se transforma en un sistema lineal:

$$z_0(t) = y(t),$$

$$z_1(t) = \dot{y}(t),$$

⋮

$$z_{m-1}(t) = y^{(m-1)}(t).$$

Usando estas variables, definimos una función vectorial del siguiente modo:

$$Z(t) = \begin{pmatrix} z_0(t) \\ z_1(t) \\ \vdots \\ z_{m-1}(t) \end{pmatrix} = \begin{pmatrix} y(t) \\ \dot{y}(t) \\ \vdots \\ y^{(m-1)}(t) \end{pmatrix}.$$

Veremos que $Z(t)$ satisface una ecuación lineal. Su derivada es

$$\dot{Z}(t) = \begin{pmatrix} \dot{z}_0(t) \\ \vdots \\ \dot{z}_{m-1}(t) \end{pmatrix}.$$

Obsérvese que

$$\begin{aligned} \dot{z}_0(t) &= \dot{y}(t) = z_1(t), \\ \dot{z}_1(t) &= \ddot{y}(t) = z_2(t), \\ &\vdots \\ \dot{z}_{m-1}(t) &= y^{(m)}(t) = -a_{m-1} y^{(m-1)}(t) - \dots - a_1 \dot{y}(t) - a_0 y(t) \\ &= -a_{m-1} z_{m-1}(t) - \dots - a_1 z_1(t) - a_0 z_0(t). \end{aligned}$$

Notamos que, usando $\dot{Z}(t)$, podemos escribir las ecuaciones anteriores como un sistema lineal. La matriz correspondiente a dicho sistema es

$$A = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & & 0 \\ \vdots & \vdots & \vdots & \vdots & 1 \\ -a_0 & -a_1 & -a_2 & & -a_{m-1} \end{pmatrix} \quad (4.13)$$

es decir, $\dot{Z}(t) = AZ(t)$.

Ejemplo

Ej 4.4.2 Transformar la ecuación lineal de tercer orden

$$\ddot{y} - 3\ddot{y} + 2\dot{y} + 7y = 0$$

en un sistema lineal de primer orden.

Como lo hicimos anteriormente, definimos

$$X(t) = \begin{pmatrix} y(t) \\ \dot{y}(t) \\ \ddot{y}(t) \end{pmatrix}.$$

Entonces

$$\begin{aligned}\dot{X}(t) &= \begin{pmatrix} \dot{y}(t) \\ \ddot{y}(t) \\ \dddot{y}(t) \end{pmatrix} = \begin{pmatrix} \dot{y}(t) \\ \ddot{y}(t) \\ 3\ddot{y}(t) - 2\dot{y}(t) - 7y(t) \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -7 & -2 & 3 \end{pmatrix} \begin{pmatrix} y(t) \\ \dot{y}(t) \\ \ddot{y}(t) \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -7 & -2 & 3 \end{pmatrix} X(t).\end{aligned}$$

◇◇◇◇◇◇◇◇◇

En la siguiente proposición, veremos que, de hecho, es fácil encontrar el polinomio característico del sistema lineal correspondiente a la matriz (4.13).

Proposición 4.4.1 *Sea $y^{(m)} + a_{m-1}y^{(m-1)} + \dots + a_1\dot{y} + a_0y = 0$ una ecuación lineal de orden m .*

Sea

$$A = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & & 0 \\ \vdots & \vdots & \vdots & \vdots & 1 \\ -a_0 & -a_1 & -a_2 & & -a_{m-1} \end{pmatrix}$$

la matriz del sistema lineal de primer orden correspondiente, $\dot{X} = AX$. Entonces, la ecuación característica de la matriz A es

$$\lambda^m + a_{m-1}\lambda^{m-1} + \dots + a_1\lambda + a_0 = 0.$$

La demostración se deja como ejercicio para el lector.

Ejemplos

Ej 4.4.3 Se desea resolver el siguiente problema con valores iniciales:

$$\ddot{y} - \dot{y} - 6y = 0,$$

$$y(0) = 3,$$

$$\dot{y}(0) = 0.$$

Lo transformaremos en un sistema lineal de primer orden. Sea

$$X(t) = \begin{pmatrix} y(t) \\ \dot{y}(t) \end{pmatrix}.$$

El sistema lineal, en este caso, es

$$\begin{aligned}\dot{X}(t) &= \begin{pmatrix} \dot{y}(t) \\ \ddot{y}(t) \end{pmatrix} = \begin{pmatrix} \dot{y}(t) \\ 6y(t) + \dot{y}(t) \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 \\ 6 & 1 \end{pmatrix} X(t).\end{aligned}$$

Utilizando la proposición 4.4.1, el polinomio característico es

$$p_A(\lambda) = \lambda^2 - \lambda - 6 = (\lambda - 3)(\lambda + 2),$$

y los valores propios son $\lambda_1 = 3$, y $\lambda_2 = -2$. Siguiendo el método usual, se puede ver que los siguientes son vectores propios que corresponden a los valores propios:

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 3 \end{pmatrix},$$

$$\mathbf{v}_2 = \begin{pmatrix} 1 \\ -2 \end{pmatrix}.$$

Concluimos que la solución general para el problema lineal es

$$X(t) = c_1 e^{3t} \begin{pmatrix} 1 \\ 3 \end{pmatrix} + c_2 e^{-2t} \begin{pmatrix} 1 \\ -2 \end{pmatrix}.$$

Para encontrar las constantes c_1 y c_2 usamos las condiciones iniciales:

$$X(0) = \begin{pmatrix} y(0) \\ y'(0) \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \end{pmatrix}.$$

Es fácil ver que $c_1 = 6/5$ y $c_2 = 9/5$; por lo tanto,

$$X(t) = \begin{pmatrix} y(t) \\ y'(t) \end{pmatrix} = \frac{6}{5} e^{3t} \begin{pmatrix} 1 \\ 3 \end{pmatrix} + \frac{9}{5} e^{-2t} \begin{pmatrix} 1 \\ -2 \end{pmatrix}.$$

Finalmente, se tiene que $y(t) = \frac{6}{5}e^{3t} + \frac{9}{5}e^{-2t}$.

Ej 4.4.4 Deseamos resolver el siguiente problema de segundo orden, usando el método de vectores propios:

$$\begin{aligned} \ddot{y} - 2\dot{y} + 2y &= 0, \\ y(0) &= 0, \\ \dot{y}(0) &= 1. \end{aligned}$$

Resolvamos el sistema lineal $\dot{X} = AX$, en donde

$$X(t) = \begin{pmatrix} y(t) \\ \dot{y}(t) \end{pmatrix}$$

y

$$A = \begin{pmatrix} 0 & 1 \\ -2 & 2 \end{pmatrix}.$$

El polinomio característico de la ecuación es $p_A(\lambda) = \lambda^2 - 2\lambda + 2$, por lo que los valores propios son

$$\lambda_{1,2} = \frac{2 \pm \sqrt{4-8}}{2} = \frac{2 \pm \sqrt{-4}}{2} = 1 \pm i,$$

es decir,

$$\begin{aligned}\lambda_1 &= \lambda = 1 + i, \\ \lambda_2 &= \bar{\lambda} = 1 - i.\end{aligned}$$

Tomemos $\lambda = 1 + i$. Encontremos vectores propios de la manera usual; esto es, si

$$\mathbf{v} = \begin{pmatrix} a \\ b \end{pmatrix},$$

entonces,

$$(A - \lambda I)\mathbf{v} = \begin{pmatrix} -1 - i & 1 \\ -2 & 1 - i \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Si escogemos $a = 1$ obtenemos que $b = 1 + i$, y por lo tanto

$$\mathbf{v} = \begin{pmatrix} 1 \\ 1 + i \end{pmatrix}$$

es vector propio con valor propio $1 + i$. Al descomponerlo en parte real y parte imaginaria se tiene

$$\begin{aligned}\lambda &= \alpha + i\beta = 1 + i, \\ \mathbf{v} &= \mathbf{r} + i\mathbf{s} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} + i \begin{pmatrix} 0 \\ 1 \end{pmatrix}.\end{aligned}$$

Utilizando las fórmulas (4.5) y (4.6)

$$\begin{aligned}\mathbf{w}_1(t) &= e^t (\mathbf{r} \cos t - \mathbf{s} \sin t), \\ \mathbf{w}_2(t) &= e^t (\mathbf{r} \sin t + \mathbf{s} \cos t),\end{aligned}\tag{4.14}$$

llegamos a,

$$\begin{aligned}\mathbf{w}_1(t) &= e^t \left[\begin{pmatrix} 1 \\ 1 \end{pmatrix} \cos t - \begin{pmatrix} 0 \\ 1 \end{pmatrix} \sin t \right] = \begin{pmatrix} e^t \cos t \\ e^t \cos t - e^t \sin t \end{pmatrix}, \\ \mathbf{w}_2(t) &= e^t \left[\begin{pmatrix} 1 \\ 1 \end{pmatrix} \sin t + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \cos t \right] = \begin{pmatrix} e^t \sin t \\ e^t \sin t + e^t \cos t \end{pmatrix}.\end{aligned}$$

La solución general es $X(t) = c_1 \mathbf{w}_1(t) + c_2 \mathbf{w}_2(t)$; usando la condición inicial, obtenemos que $c_1 = 0$ y $c_2 = 1$. La solución al problema de valores iniciales es, por lo tanto,

$$X(t) = \begin{pmatrix} e^t \sin t \\ e^t \sin t + e^t \cos t \end{pmatrix},$$

y concluimos que $y(t) = e^t \sin t$.



El método de valores propios no sería totalmente adecuado para el caso que nos ocupa, si es que encontrar los vectores propios fuera muy complicado. En el siguiente teorema se presenta una manera simple de calcularlos. Advertimos que este teorema es aplicable sólo a un tipo de matrices. No se debe intentar aplicarlo al caso general.

Teorema 4.4.2

Sea $y^{(m)} + a_{m-1}y^{(m-1)} + \dots + a_1\dot{y} + a_0y = 0$ una ecuación lineal de orden m . Sea

$$A = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & & 0 \\ \vdots & \vdots & \vdots & \vdots & 1 \\ -a_0 & -a_1 & -a_2 & & -a_{m-1} \end{pmatrix}$$

la matriz del sistema lineal de primer orden correspondiente, $\dot{X} = AX$. Entonces, para cada λ_k , el siguiente es un vector propio, con valor propio λ_k :

$$\mathbf{v}_k = \begin{pmatrix} 1 \\ \lambda_k \\ (\lambda_k)^2 \\ \vdots \\ (\lambda_k)^{m-1} \end{pmatrix}.$$

La demostración se deja como ejercicio al lector.

Ejemplos

Ej 4.4.5 Se desea resolver el siguiente problema:

$$\ddot{y} + y = 0,$$

$$y(0) = 6,$$

$$\dot{y}(0) = -1.$$

Reescribimos la ecuación de segundo orden como un sistema lineal de primer orden. Sea

$$Y(t) = \begin{pmatrix} y(t) \\ \dot{y}(t) \end{pmatrix}.$$

El sistema lineal correspondiente es,

$$\dot{Y}(t) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} Y(t).$$

El polinomio característico es

$$p_A(\lambda) = \lambda^2 + 1.$$

Las raíces de p_A son $\lambda = i$ y $\bar{\lambda} = -i$. Encontremos un vector propio \mathbf{v} para el valor propio $\lambda = i$ usando el teorema 4.4.2. Así,

$$\mathbf{v} = \begin{pmatrix} 1 \\ i \end{pmatrix}$$

es un posible vector propio con valor propio λ . Entonces $\alpha = 0$, $\beta = 1$ y

$$\mathbf{r} = \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

$$\mathbf{s} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Usando las fórmulas (4.5) y (4.6), obtenemos dos soluciones linealmente independientes, $\mathbf{w}_1(t)$ y $\mathbf{w}_2(t)$. Después de simplificar se tiene

$$\mathbf{w}_1(t) = \begin{pmatrix} \cos t \\ -\sin t \end{pmatrix},$$

$$\mathbf{w}_2(t) = \begin{pmatrix} \sin t \\ \cos t \end{pmatrix}.$$

La solución general es $Y(t) = c_1 \mathbf{w}_1(t) + c_2 \mathbf{w}_2(t)$. De las condiciones iniciales, concluimos que $c_1 = 6$ y $c_2 = -1$. Por lo tanto, la solución al problema de valores iniciales es

$$Y(t) = 6 \begin{pmatrix} \cos(t) \\ -\sin(t) \end{pmatrix} - \begin{pmatrix} \sin(t) \\ \cos(t) \end{pmatrix},$$

y por lo tanto, $y(t) = 6 \cos(t) - \sin(t)$. Es claro que este método coincide con lo visto en el capítulo 2.

Ej 4.4.6 Vamos a resolver la ecuación $\ddot{y} - \dot{y} = 0$ con condiciones iniciales $y(0) = 5$, $\dot{y}(0) = 2$ y $\ddot{y}(0) = 0$. Reescribamos la ecuación anterior como un sistema de primer orden, usando $Y(t) = \begin{pmatrix} y(t) \\ \dot{y}(t) \\ \ddot{y}(t) \end{pmatrix}$. El sistema asociado es,

$$\dot{Y}(t) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} Y(t),$$

con lo cual la ecuación característica es $\lambda^3 - \lambda = 0$ y los valores propios son $\lambda_1 = 0$, $\lambda_2 = 1$ y $\lambda_3 = -1$.

Del teorema 4.4.2, concluimos que, para $\lambda_1 = 0$,

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ \lambda_1 \\ \lambda_1^2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

es vector propio. Del mismo modo,

$$\mathbf{v}_2 = \begin{pmatrix} 1 \\ \lambda_2 \\ \lambda_2^2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

es vector propio con valor propio $\lambda_2 = 1$. Finalmente,

$$\mathbf{v}_3 = \begin{pmatrix} 1 \\ \lambda_3 \\ \lambda_3^2 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$$

es vector propio con valor propio $\lambda_3 = -1$. De aquí que la solución general sea de la forma

$$X(t) = C_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + C_2 e^t \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + C_3 e^{-t} \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}.$$

Para encontrar las constantes, usamos las condiciones iniciales,

$$X(0) = \begin{pmatrix} 5 \\ 2 \\ 0 \end{pmatrix} = C_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + C_2 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + C_3 \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}.$$

y al resolver, obtenemos $C_1 = 5$, $C_2 = 1$ y $C_3 = -1$; es decir, la solución es

$$X(t) = \begin{pmatrix} 5 \\ 0 \\ 0 \end{pmatrix} + e^t \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - e^{-t} \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}.$$

y concluimos que $y(t) = 5 + e^t - e^{-t}$.



Ejercicios

▷ 4.1 Resolver los siguientes sistemas de ecuaciones:

a) $\dot{X} = \begin{pmatrix} 2 & 4 \\ 1 & -3 \end{pmatrix} X.$

b) $\dot{X} = \begin{pmatrix} -3 & 1 \\ 1 & -3 \end{pmatrix} X.$

c) $\dot{X} = \begin{pmatrix} 2 & 1 \\ 3 & 2 \end{pmatrix} X.$

d) $\dot{X} = \begin{pmatrix} 0 & 1 \\ 0 & -4 \end{pmatrix} X.$

e) $\dot{X} = \begin{pmatrix} -1 & 1 \\ -1 & -1 \end{pmatrix} X.$

f) $\dot{X} = \begin{pmatrix} 5 & 4 \\ -1 & 1 \end{pmatrix} X.$

g) $\dot{X} = \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix} X.$

h) $\dot{X} = \begin{pmatrix} 3 & -1 \\ 4 & -1 \end{pmatrix} X.$

i) $\dot{X} = \begin{pmatrix} 3 & -6 \\ 1 & -2 \end{pmatrix} X.$

▷ 4.2 Resolver los siguientes sistemas de ecuaciones:

a) $\dot{X} = \begin{pmatrix} -2 & -1 \\ 1 & -1 \end{pmatrix} X + \begin{pmatrix} 9 \\ 3 \end{pmatrix}.$

b) $\dot{X} = \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix} X + \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$

c) $\dot{X} = \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix} X + \begin{pmatrix} 2e^t \\ e^t \end{pmatrix}.$

▷ 4.3 Resolver los siguientes problemas con valores iniciales:

a) $\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} 2 & \frac{1}{4} \\ 25 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix},$ con condición inicial $(x(0), y(0)) = \begin{pmatrix} -1 \\ 10 \end{pmatrix}.$

Encontrar $\lim_{t \rightarrow \infty} (x(t), y(t)),$ si éste existe.

b) $\dot{X} = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix} X, X(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ (suponer que $\beta \neq 0$).

c) $\dot{X} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & -2 & 2 \end{pmatrix} X,$ con la condición inicial $X(0) = \begin{pmatrix} 10 \\ 20 \\ 30 \end{pmatrix}.$

▷ 4.4 Consideremos el siguiente problema con valores iniciales:

$$\begin{aligned} \dot{X} &= \begin{pmatrix} 0 & 1 \\ 2 & -1 \end{pmatrix} X, \\ X(0) &= \begin{pmatrix} 3 \\ 3w \end{pmatrix}, \end{aligned}$$

donde w es una constante real.

a) Resolver el problema.

b) ¿Cuánto debe valer w si deseamos que $\lim_{t \rightarrow \infty} X(t) = \mathbf{0}?$

▷ 4.5 (Curva Catenaria) Considerar el siguiente problema lineal:

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

a) Determinar las condiciones sobre a, b, c y d de manera que la matriz tenga como valores propios a λ y $-\lambda$ con λ real.

b) Escribir la solución para el caso del inciso anterior. Las curvas solución $x(t)$ y $y(t)$ se llaman **curvas catenarias**.

▷ **4.6** Considerar el siguiente problema lineal.

$$\dot{X} = \begin{pmatrix} -1 & 3 \\ 1 & -3 \end{pmatrix} X.$$

a) Encontrar la solución general $X(t)$.

b) Calcular $\lim_{t \rightarrow \infty} X(t)$.

▷ **4.7** Resolver el siguiente sistema:

$$\dot{X} = \begin{pmatrix} -2 & -1 & 7 \\ -10 & 0 & 16 \\ -4 & 0 & 6 \end{pmatrix} X.$$

(Sugerencia: $\lambda = 2$ es valor propio.)

▷ **4.8** Las siguientes son tres soluciones *linealmente independientes* de un problema lineal $\dot{X} = AX$:

$$\mathbf{w}_1(t) = \begin{pmatrix} 3 - 2e^t \\ 0 \\ -3 + 3e^t \end{pmatrix},$$

$$\mathbf{w}_2(t) = \begin{pmatrix} 7 - 7e^t \\ e^t \\ -7 + 7e^t \end{pmatrix},$$

$$\mathbf{w}_3(t) = \begin{pmatrix} 2 - 2e^t \\ 0 \\ -2 + 3e^t \end{pmatrix}.$$

Encontrar la matriz A . (Sugerencia: usar $\Phi(t) = [\mathbf{w}_1(t) \ \mathbf{w}_2(t) \ \mathbf{w}_3(t)]$.)

▷ **4.9** Sea A una matriz de $n \times n$. Se sabe que el problema lineal $\dot{X} = AX$ tiene n soluciones linealmente independientes, digamos $\{\mathbf{w}_1(t), \dots, \mathbf{w}_n(t)\}$. Sea $\Phi(t) = [\mathbf{w}_1(t) \ \dots \ \mathbf{w}_n(t)]$ la matriz cuyas columnas son estas soluciones. (Esta matriz es llamada matriz fundamental.)

a) Explicar por qué, para cada tiempo t , la matriz $\Phi(t)$ es invertible.

b) Demostrar que $\dot{\Phi}(t) = A\Phi(t)$.

c) Sea $f(t)$ una función cualquiera. Sea $Y(t) = \Phi(t) \int_0^t \Phi(s)^{-1} f(s) ds$. Mostrar que Y es una solución particular del problema $\dot{X} = AX + f(t)$.

▷ **4.10** Demostrar la proposición 4.4.1.

▷ **4.11** Demostrar el teorema 4.4.2.

▷ **4.12** En una economía se producen únicamente los bienes $x(t)$ y $y(t)$ y éstos son utilizados tanto como insumos para la producción como para consumo final. La producción se ajusta en el tiempo de manera proporcional al exceso de demanda. De esta forma, un modelo que describe las trayectorias $x(t)$ y $y(t)$ es el siguiente:

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} a_{xx} & a_{xy} \\ a_{yx} & a_{yy} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} d_x(t) \\ d_y(t) \end{pmatrix} - \begin{pmatrix} x \\ y \end{pmatrix},$$

en donde, d_x y d_y son las demandas de consumo final y a_{ij} representa la cantidad del bien i que es necesaria para producir una unidad del bien j . El lector reconocerá la matriz como una típica matriz de insumo-producto.

a) Resolver el modelo descrito cuando la matriz está dada por

$$\begin{pmatrix} \frac{3}{10} & \frac{4}{10} \\ \frac{3}{10} & \frac{2}{10} \end{pmatrix}$$

y el vector de demanda final por $\begin{pmatrix} 100 \\ 50 \end{pmatrix}$.

b) Resolver el modelo descrito cuando la matriz está dada por

$$\begin{pmatrix} \frac{3}{10} & \frac{4}{10} \\ \frac{3}{10} & \frac{2}{10} \end{pmatrix}$$

y el vector de demanda final por

$$\begin{pmatrix} e^{\frac{t}{10}} \\ 2e^{\frac{t}{10}} \end{pmatrix}.$$

▷ **4.13** Considerar la ecuación no autónoma

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

- a) Describir cómo se puede transformar esta ecuación en un sistema de ecuaciones autónomas.
- b) Utilizar el método anterior para resolver la ecuación $\dot{x} = 2x + t$.

▷ **4.14** Considerar el sistema dado por

$$t\dot{x} = a_1x + b_1y,$$

$$t\dot{y} = a_2x + b_2y.$$

Probar que la transformación $t = e^w$ convierte al sistema en un sistema lineal con coeficientes constantes.

▷ **4.15** Utilizar el resultado del ejercicio anterior para resolver

$$t\dot{x} = x + y,$$

$$t\dot{y} = -3x + 5y.$$

▷ **4.16** Probar los siguientes resultados.

- a) Si \mathbf{v}_1 y \mathbf{v}_2 son vectores propios con valores propios $\lambda_1 \neq \lambda_2$, respectivamente, entonces \mathbf{v}_1 y \mathbf{v}_2 son linealmente independientes.
- b) Si $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ es un conjunto de vectores propios de una matriz de $n \times n$, $k \leq n$, con valores propios $\{\lambda_1, \dots, \lambda_k\}$, respectivamente, en donde $\lambda_i \neq \lambda_j$ si $i \neq j$; entonces, $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ es un conjunto linealmente independiente de vectores de \mathbb{R}^n .

▷ **4.17** Resolver los siguientes problemas con valores iniciales. Graficar la solución. ¿Qué pasa cuando $t \rightarrow \infty$?

a) $\ddot{x} + 2\dot{x} + 2x = 0$, $x(0) = 1$, $\dot{x}(0) = 0$.

b) $\ddot{x} + 5\dot{x} - 6 = 0$, $x(0) = -1$, $\dot{x}(0) = 10$.

c) $\ddot{x} - 3\dot{x} = 0$, $x(0) = 2$, $\dot{x}(0) = 5$.

d) $\ddot{x} + 2\dot{x} + x = 0$, $x(0) = 1$, $\dot{x}(0) = 1$.

e) $\ddot{x} - 4\dot{x} + 4x = 0$, $x(0) = 1$, $\dot{x}(0) = 0$.

f) $\ddot{x} + 5\dot{x} + 6 = 2$, $x(0) = -1$, $\dot{x}(0) = 0$.

g) $y''' - 2y' + 4y = 0$, $y(0) = 1$, $y'(0) = 0$, $y''(0) = 0$.

▷ **4.18** Consideremos el siguiente problema de tercer orden:

$$y''' + y'' + y' + y = 0.$$

- a) Resolver el problema usando las siguientes condiciones iniciales:

$$y(0) = u, \quad y'(0) = v, \quad y''(0) = w.$$

- b) Identificar el conjunto de condiciones iniciales para las cuales se cumple

$$\lim_{t \rightarrow \infty} y(t) = 0.$$

▷ **4.19** Resolver el siguiente problema de tercer orden:

$$y''' - 2y' + 4y = 0, \quad y(0) = 1,$$

$$y'(0) = 0, \quad y''(0) = 0.$$

▷ **4.20** Consideremos el siguiente problema de tercer orden:

$$y''' - y'' + 2y = 0.$$

- a) Resolver el problema, usando las condiciones iniciales,

$$y(0) = u, \quad y'(0) = v, \quad y''(0) = 1.$$

- b) ¿Cuánto deben valer u y v si deseamos que $\lim_{t \rightarrow \infty} y(t) = 0$?

▷ **4.21** Resolver el siguiente problema de cuarto orden.

$$y^{(4)} - y = 0,$$

con condiciones iniciales dadas por $y(0) = 0$, $y'(0) = 0$, $y''(0) = 0$ y $y'''(0) = 1$.

CAPÍTULO 5

Análisis cualitativo

§5.1 Puntos de equilibrio

En este capítulo se presenta el método de análisis cualitativo de los sistemas lineales. Extenderemos la definición 2.2.2 de punto de equilibrio para sistemas de ecuaciones como sigue:

Definición 5.1.1 *Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ un campo vectorial. Un sistema de ecuaciones diferenciales autónomo de primer orden $\dot{X} = f(X)$ tiene un **punto de equilibrio** (o **punto fijo** o **estado estacionario**) en \mathbf{p} si $f(\mathbf{p}) = \mathbf{0}$.*

Ejemplo

Ej 5.1.1 Considerar el sistema de ecuaciones

$$\begin{aligned}\dot{y}_1 &= y_1^2 - y_2 + 1, \\ \dot{y}_2 &= y_1(y_2 + 3),\end{aligned}$$

y encontrar los puntos de equilibrio del sistema. Al igualar ambas ecuaciones a cero, obtenemos lo siguiente:

$$\begin{aligned}y_2 &= y_1^2 + 1, \\ y_1(y_2 + 3) &= 0.\end{aligned}$$

De aquí se tiene que $y_1 = 0$ y $y_2 = 1$. Por lo tanto,

$$\mathbf{p}^* = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

es el único punto fijo o punto de equilibrio del sistema de ecuaciones dado.



En el caso de un sistema lineal homogéneo, tenemos $f(X) = AX$ donde A es una matriz de $n \times n$. Si $\det(A) \neq 0$, entonces existe un único punto de equilibrio $\mathbf{p}^* = \mathbf{0}$, pues dado cualquier otro punto de equilibrio, digamos \mathbf{q}^* , entonces $A\mathbf{q}^* = \mathbf{0}$, y esto implica $\mathbf{q}^* = A^{-1}\mathbf{0} = \mathbf{0}$. Si $\det(A) = 0$ entonces A^{-1} no existe y se tiene una infinidad de puntos de equilibrio; éste es el caso en el cual algún valor propio es nulo.

Ejemplo

Ej 5.1.2 Vamos a encontrar los puntos de equilibrio del siguiente sistema

$$\begin{aligned}\dot{y}_1 &= y_1 - y_2, \\ \dot{y}_2 &= 4y_1 - 4y_2.\end{aligned}$$

Este es un problema lineal del tipo $\dot{X} = AX$, donde $X = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ y

$$A = \begin{pmatrix} 1 & -1 \\ 4 & -4 \end{pmatrix}.$$

Observamos que $\det(A) = 0$. Si $\mathbf{p}^* = \begin{pmatrix} a \\ b \end{pmatrix}$ es un punto de equilibrio, entonces se cumple

$$A\mathbf{p}^* = \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 4 & -4 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix},$$

es decir, necesariamente $a - b = 0$, o bien $a = b$, de manera que el conjunto de puntos de equilibrio está generado por el vector $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, es decir, se trata de la recta $y_1 = y_2$. Una representación geométrica de esta situación está dada en la figura 5.1.

Ej 5.1.3 Consideremos el siguiente sistema:

$$\begin{aligned}\dot{x} &= y + x - x^3, \\ \dot{y} &= -y + 3x^2y.\end{aligned}$$

Vamos a encontrar los puntos fijos o de equilibrio. El campo vectorial es, en este caso,

$$f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} y + x - x^3 \\ -y + 3x^2y \end{pmatrix}$$

y los puntos de equilibrio satisfacen $f\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{0}$. Tenemos que resolver simultáneamente el siguiente par de ecuaciones:

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

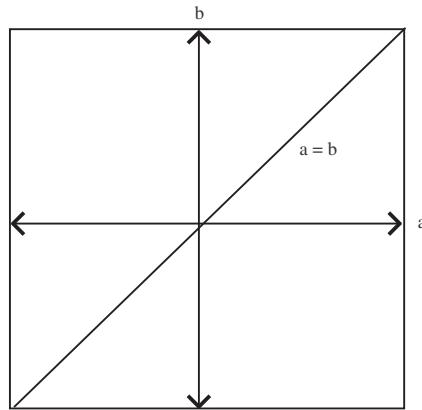


Figura 5.1: Conjunto de Puntos de Equilibrio para el ejemplo 5.1.2.

Al resolver este sistema, tenemos los siguientes puntos de equilibrio:

$$\begin{aligned} p_1^* &= \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad p_2^* = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad p_3^* = \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \\ p_4^* &= \begin{pmatrix} \sqrt{1/3} \\ (-2/3)\sqrt{1/3} \end{pmatrix}, \quad p_5^* = \begin{pmatrix} -\sqrt{1/3} \\ (2/3)\sqrt{1/3} \end{pmatrix}. \end{aligned}$$

§5.2 Clasificación de puntos de equilibrio

Para el caso lineal tenemos una clasificación de los puntos de equilibrio según el comportamiento dinámico de la soluciones. Para empezar, clasificamos los puntos de acuerdo con si son degenerados o no lo son.

Definición 5.2.1 Un punto fijo o de equilibrio p^* de una matriz A es **degenerado** si se cumple alguna de las siguientes condiciones:

- $\det(A) = 0$,
- A tiene valores propios repetidos,
- A tiene un valor propio complejo con parte real igual a 0.

Ejemplo

Ej 5.2.1 Consideremos el siguiente sistema lineal:

$$\dot{y}_1 = -y_2,$$

$$\dot{y}_2 = y_1.$$

Queremos ver si los puntos de equilibrio son o no degenerados. Este sistema corresponde a

$$\dot{X} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} X.$$

El polinomio característico de la matriz es

$$\lambda^2 - \text{tr}(A)\lambda + \det(A) = \lambda^2 + 1.$$

En este caso, es fácil ver que los valores propios son $\pm i$. La parte real de cada uno es cero y tenemos un solo punto de equilibrio: $y_1 = 0$ y $y_2 = 0$. Decimos por lo tanto que $(0, 0)$ es un punto de equilibrio degenerado.

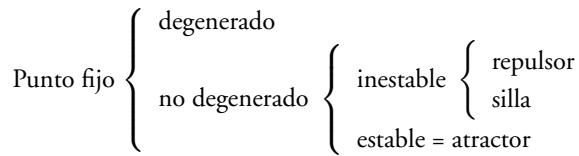


Si el sistema lineal es homogéneo, el origen siempre es un punto fijo, como queda claro del ejemplo anterior. De manera análoga, si el sistema no es homogéneo, cualquier función constante que sea una solución particular del sistema es un punto fijo.

Los puntos no degenerados se clasifican según su estabilidad. Este concepto es de suma importancia en el desarrollo posterior de la teoría cualitativa. De hecho, podemos obtener gran cantidad de información de un sistema dando simplemente una clasificación de sus puntos fijos.

Definición 5.2.2 Un punto de equilibrio no degenerado es **estable** si todos los valores propios de la matriz tienen parte real negativa. Un punto de equilibrio es **inestable** si algún valor propio tiene parte real positiva. Un punto fijo inestable es un **repulsor** si todos los valores propios tienen parte real positiva. Un punto fijo inestable es **silla** si la matriz tiene tanto valores propios con parte real positiva como valores propios con parte real negativa. A los puntos estables también se les llama **atractores**.

El siguiente diagrama ilustra los casos posibles.:



Ejemplo

Ej 5.2.2 Vamos a clasificar el punto fijo $\mathbf{p}^* = \mathbf{0}$ del siguiente sistema:

$$\dot{X} = \begin{pmatrix} 3 & -1 & 0 \\ 1 & 3 & 0 \\ 0 & 0 & a \end{pmatrix} X.$$

La ecuación característica es $(a - \lambda)(\lambda^2 - 6\lambda + 10) = 0$ y los valores propios son

$$\begin{aligned}\lambda_1 &= a, \\ \lambda_{2,3} &= \frac{6 \pm \sqrt{-4}}{2} = 3 \pm i.\end{aligned}$$

Tenemos las siguientes posibilidades para el punto fijo:

- a) $a > 0$. El punto es no degenerado, inestable y repulsor, pues los valores propios tienen parte real positiva.
- b) $a = 0$. Esto implica que el determinante es cero y por lo tanto tenemos un caso degenerado.
- c) $a < 0$. El punto es no degenerado, inestable y silla.



Una característica de los atractores es que si $X(t)$ es una solución del sistema lineal entonces converge al atractor; es decir, si \mathbf{p}^* es un atractor, entonces se cumple $\lim_{t \rightarrow \infty} X(t) = \mathbf{p}^*$. Esta propiedad es de utilidad en otras situaciones, no necesariamente lineales, en las que la condición inicial se encuentra “cerca” del punto de equilibrio.

Ejemplo

Ej 5.2.3 Considérese el sistema

$$\dot{X} = \begin{pmatrix} 0 & 1 \\ -3 & -2w \end{pmatrix} X,$$

con punto fijo

$$\mathbf{p}^* = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Queremos obtener condiciones sobre w para que se cumpla lo siguiente:

$$\lim_{t \rightarrow \infty} X(t) = \mathbf{p}^* = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Esto sucede sólo si \mathbf{p}^* es atractor, por lo que necesitamos que todos los valores propios tengan parte real negativa. La ecuación característica es

$$\lambda^2 + 2w\lambda + 3 = 0,$$

y los valores propios son

$$\lambda_{1,2} = \frac{-2w \pm \sqrt{4w^2 - 12}}{2} = -w \pm \sqrt{w^2 - 3}.$$

De aquí que, si $w > 0$ y $w^2 - 3 > 0$ o $w^2 - 3 < 0$, las partes reales de λ_1 y λ_2 son negativas. Si $w^2 - 3 = 0$, entonces se trata de un caso degenerado. Si $w \leq 0$ entonces \mathbf{p}^* no es atractor; por lo tanto, deseamos que $w > 0$ con $w^2 \neq 3$.



§5.3 Diagramas de fase

Las soluciones a un sistema de ecuaciones del tipo

$$\dot{X} = f(X),$$

en donde $X(t) = \begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix}$, se pueden representar de manera gráfica de distintas formas. Una posibilidad es simplemente graficar las soluciones $x_i(t)$, para $i = 1, \dots, n$. El problema de esto es que es necesario conocer, si no la solución explícita, al menos su comportamiento cualitativo. Otra posibilidad es que, dada una solución

$$X(t) = \begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix},$$

ésta puede interpretarse como un conjunto de ecuaciones paramétricas de manera que para cada valor de t se tiene un punto $X(t)$ en \mathbb{R}^n .

Dado un punto inicial $(x_1(0), \dots, x_n(0))$ y un intervalo de valores para t , los puntos $X(t)$ de \mathbb{R}^n corresponden a una trayectoria solución. Para el caso $n = 2$, esta representación se puede hacer en el plano y a los diagramas correspondientes se los conoce como **diagramas de fase** (recuérdese lo visto en la sección 3.3). Debemos hacer notar que dado que se trata de un sistema autónomo, dos trayectorias solución distintas nunca se intersectan.

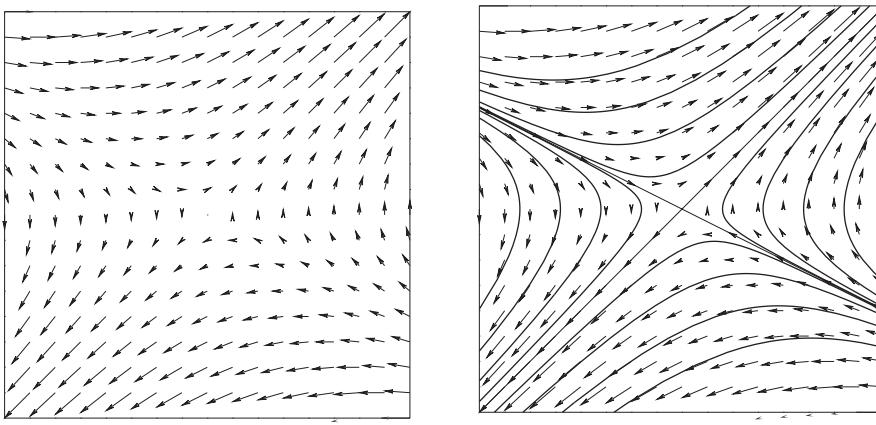
El **campo vectorial** o **campo de dirección** para el sistema $\dot{X} = f(X)$ es una función $g : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ dada por $g(x, y) = (\dot{x}, \dot{y})$; es decir, es un conjunto de vectores (flechas) en el plano XY , tal que la pendiente del vector en el punto (x, y) está dada por $\frac{\dot{y}}{\dot{x}}$. La trayectoria solución que pasa por el punto (x, y) es tangente al vector ya que $\frac{dy}{dx} = \frac{dy}{dt}/\frac{dx}{dt} = \frac{\dot{y}}{\dot{x}}$. La dirección del vector es la dirección del flujo, es decir, nos dice hacia dónde se mueven las variables cuando t avanza.

La figura 5.2.a ilustra un campo de dirección alrededor de un punto silla. Si añadimos algunas trayectorias solución obtenemos la figura 5.2.b.

Definición 5.3.1 Sea $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$. Dado el sistema $\dot{X} = F(X)$, donde F es una función no necesariamente lineal, los lugares geométricos de \mathbb{R}^n que satisfacen $\dot{x}_i = a$ para alguna $i = 1, \dots, n$, con a una constante, se llaman **isoclinas**.

En particular las isoclinas de la forma $\dot{x}_i = 0$ son muy informativas pues nos dan los puntos para los cuales ya no hay ajuste dinámico para x_i . Las isoclinas son útiles para obtener la dirección de los vectores del campo de dirección y por lo tanto sirven para esbozar las curvas solución.

En el caso de un sistema de dos dimensiones, con $X = \begin{pmatrix} x \\ y \end{pmatrix}$, la isocrina $\dot{x} = 0$ contiene los puntos para los cuales las curvas solución tienen pendiente vertical. Análogamente, la isocrina $\dot{y} = 0$ está



a) Campo de Dirección

b) Algunas Trayectorias

Figura 5.2: Diagrama de fase en torno de un punto silla.

formada por los puntos para los cuales las soluciones tienen pendiente horizontal. Esto se debe a que la pendiente de los vectores del campo de dirección (tangentes a las curvas solución) está dada por $\frac{\dot{y}}{\dot{x}}$. Recordemos lo siguiente:

- $\dot{x} > 0 \Rightarrow x(t)$ crece.
- $\dot{x} < 0 \Rightarrow x(t)$ decrece.
- $\dot{y} > 0 \Rightarrow y(t)$ crece.
- $\dot{y} < 0 \Rightarrow y(t)$ decrece.

La isoclina $\dot{x} = 0$ divide al plano en dos regiones, una con $\dot{x} > 0$ y la otra con $\dot{x} < 0$; la isoclina $\dot{y} = 0$ hace lo análogo. Al graficar ambas isoclinas, el plano queda dividido en cuatro tipos de región y es posible dar el flujo de x y y en cada uno.

Dada una función continua y diferenciable $f : D \rightarrow \mathbb{R}$, con $D \subset \mathbb{R}^2$ y 0 en la imagen de f , considérese la curva de nivel $f(x, y) = 0$; ésta divide la región D en dos subregiones: $f^+ = \{(x, y) \in D \mid f(x, y) > 0\}$ y $f^- = \{(x, y) \in D \mid f(x, y) < 0\}$. Dado cualquier punto (x, y) sobre esta curva de nivel anterior, sabemos que la función crece en la dirección del gradiente $\nabla f(x, y)$ (en particular, ésta es la dirección en la que la función crece con mayor rapidez); entonces, el campo vectorial de los gradientes sobre la curva de nivel apunta siempre hacia la región f^+ .

Los siguientes ejemplos ilustran estas consideraciones.

Ejemplos

Ej 5.3.1 Considerar la función $f(x, y) = 3x - 2y$. La recta $3x - 2y = 0$ divide al plano en dos regiones. El gradiente sobre cualquier punto de esta curva de nivel es el vector $\begin{pmatrix} 3 \\ -2 \end{pmatrix}$ que apunta hacia la región

$$f^+ = \{(x, y) \mid 3x - 2y > 0\},$$

como se muestra en la figura 5.3.

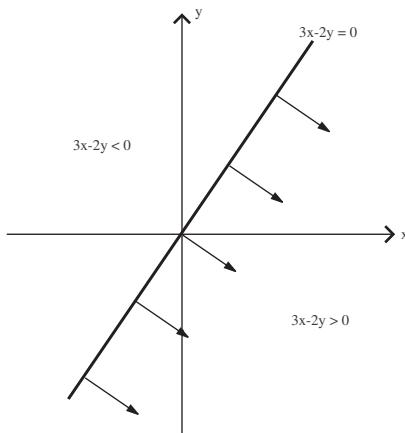


Figura 5.3: Regiones en que se divide el dominio en el ejemplo 5.3.1. Las flechas sobre la curva de nivel $3x - 2y = 0$ representan el gradiente.

Ej 5.3.2 Considerar la función $f(x, y) = y + (x - 1)^2 - 1$. La curva de nivel $y + (x - 1)^2 - 1 = 0$ divide al plano en dos regiones. El gradiente sobre cualquier punto de la curva está dado por $\nabla f(x, y) = \begin{pmatrix} 2(x - 1) \\ 1 \end{pmatrix}$ que apunta hacia la región

$$f^+ = \{(x, y) \mid y + (x - 1)^2 - 1 > 0\},$$

como se ilustra en la figura 5.4.



Dado un sistema lineal de 2×2 , es relativamente sencillo realizar el diagrama de fase correspondiente. Debido a que el comportamiento cualitativo de la solución depende exclusivamente de la matriz del sistema, sin pérdida de generalidad, podemos analizar sólo sistemas homogéneos. Así, el origen siempre es un punto de equilibrio. Es claro que si un sistema lineal no es homogéneo simplemente se traslada el punto de equilibrio a la solución particular.

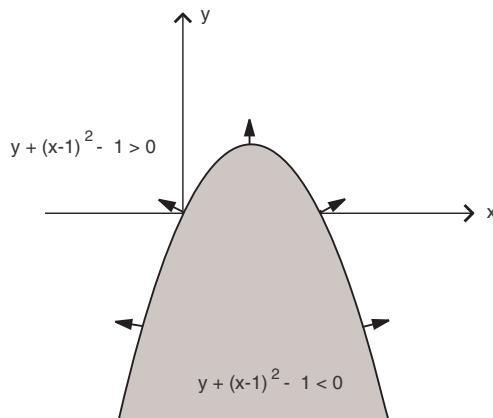


Figura 5.4: Regiones en las que se divide el dominio y vectores gradiente sobre la curva de nivel $y + (x - 1)^2 = 0$ para el ejemplo 5.3.2.

Consideremos el sistema dado por

$$\begin{aligned}\dot{x} &= f(x, y) = ax + by, \\ \dot{y} &= g(x, y) = cx + dy.\end{aligned}\tag{5.1}$$

Las isoclinas están dadas por las rectas

$$\begin{aligned}ax + by &= 0, \\ cx + dy &= 0.\end{aligned}$$

Estas rectas dividen el plano en cuatro regiones: las direcciones cualitativas de flujo para cada una de estas regiones se pueden determinar como sigue. Sabemos que la dirección $\nabla f(x, y) = \begin{pmatrix} a \\ b \end{pmatrix}$ apunta hacia la región donde $\dot{x} = f(x, y) > 0$ y la dirección $\nabla g(x, y) = \begin{pmatrix} c \\ d \end{pmatrix}$ apunta hacia la región donde $\dot{y} = g(x, y) > 0$. Esto se muestra en el siguiente ejemplo.

Ejemplo

Ej 5.3.3 Se desea ilustrar las isoclinas y determinar las direcciones de flujo para el sistema

$$\begin{aligned}\dot{x} &= x - y, \\ \dot{y} &= x + 3y.\end{aligned}$$

En el plano XY dibujamos las isoclinas correspondientes; escogemos un punto en cada región y ponemos flechas horizontales a la derecha o a la izquierda para indicar si x crece o decrece, y verticales para indicar

si y crece o decrece. Asimismo, y sólo con propósitos ilustrativos, dibujamos cuatro posibles soluciones. Nótese que las soluciones cortan a $\dot{x} = 0$ verticalmente y a $\dot{y} = 0$ horizontalmente. La figura 5.5 muestra estas consideraciones.

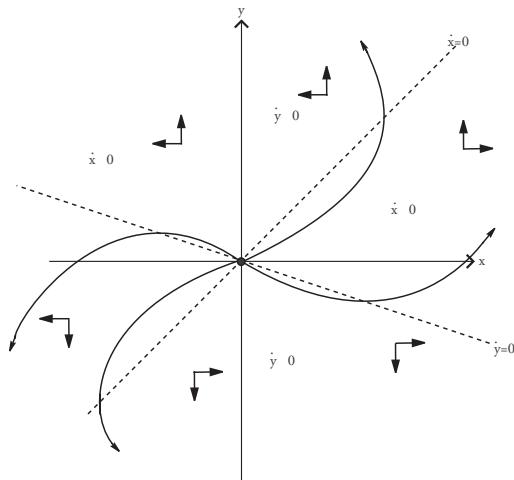


Figura 5.5: Las líneas punteadas corresponden a las isoclinas del ejemplo 5.3. Las curvas sólidas son ejemplos de soluciones (nótese que la solución corta verticalmente a la isocrina $\dot{x} = 0$ y horizontalmente a la isocrina $\dot{y} = 0$).



§5.4 Clasificación de sistemas lineales de 2×2

En esta sección haremos un examen más detallado de los sistemas lineales de 2×2 . Sea A una matriz de 2×2 cuyo polinomio característico está dado por (4.4), por lo que los valores propios de A son

$$\lambda_{1,2} = \frac{\text{tr}(A) \pm \sqrt{\text{tr}^2(A) - 4 \det(A)}}{2}.$$

Definición 5.4.1 *El discriminante de una matriz A de 2×2 es*

$$D(A) = \text{tr}^2(A) - 4 \det(A).$$

Queremos identificar los casos degenerados de las matrices de 2×2 . Recordemos que el sistema determinado por A tiene puntos fijos degenerados si se cumple cualquiera de las siguientes condiciones (cf. definición 5.2.1):

- a) $\det(A) = 0$,
- b) A tiene raíces repetidas,

- c) los valores propios imaginarios de A tienen parte real igual a 0.

Si $D(A) > 0$, tenemos dos raíces reales distintas. Entonces, la única manera de tener un caso degenerado es si uno de los valores propios es 0 y esto ocurre sólo si $\det(A) = 0$, obteniéndose el caso a. El caso b, cuando las raíces son repetidas, se cumple sólo si el discriminante se anula, es decir, si $D(A) = \text{tr}^2(A) - 4\det(A) = 0$. Para el caso c, si $D(A) < 0$, entonces las partes reales son

$$\text{Re}(\lambda_1) = \text{Re}(\lambda_2) = \frac{\text{tr}(A)}{2}.$$

En esta situación se tiene un punto degenerado si $\text{tr}(A) = 0$. Si $D(A) < 0$ y $\text{tr}(A) = 0$, entonces $-4\det(A) < 0$ y por lo tanto $\det(A) > 0$. En conclusión, el caso c ocurre si y sólo si $\det(A) > 0$ y $\text{tr}(A) = 0$. Podemos graficar los diversos casos en un diagrama que tenga a $\det(A)$ en el eje vertical y a $\text{tr}(A)$ en el horizontal (véase la figura 5.6).

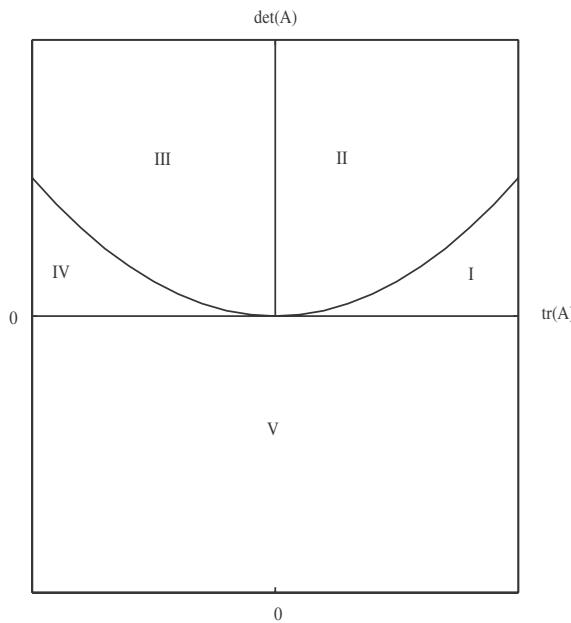


Figura 5.6: Regiones del plano definido por la traza y el determinante de una matriz A , separadas por casos degenerados.

Nótese que los casos degenerados dividen el plano en cinco regiones. Estudiaremos el comportamiento dinámico en cada una de ellas, y para hacerlo numeraremos las regiones como en la figura 5.6.

§5.4.1 Región I

Esta región se define como el conjunto de matrices para las cuales $\det(A) > 0$, $\text{tr}(A) > 0$ y $D(A) > 0$. Es fácil ver que una matriz en esta región tiene dos valores propios reales y distintos, ambos positivos. Si

λ_1 y λ_2 son los valores propios con vectores propios \mathbf{v}_1 y \mathbf{v}_2 , entonces la solución general del problema está dada por

$$X(t) = C_1 e^{\lambda_1 t} \mathbf{v}_1 + C_2 e^{\lambda_2 t} \mathbf{v}_2.$$

En este caso, se tiene un punto de equilibrio inestable, y $\lim_{t \rightarrow -\infty} X(t) = \mathbf{0}$. Se dice que $\mathbf{0}$ es un **nodo repulsor**. En la figura 5.7 se ilustra el aspecto típico de las soluciones cercanas a un nodo inestable.

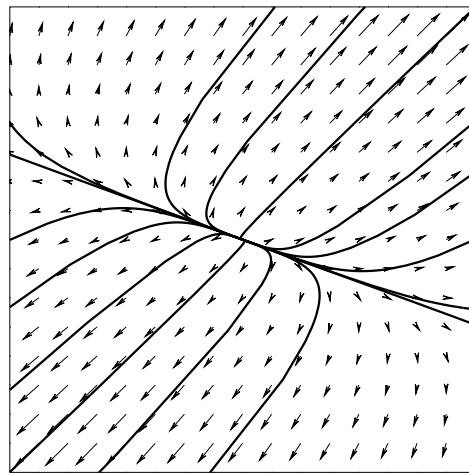


Figura 5.7: Algunas curvas solución alrededor de un nodo repulsor, correspondiente a la región I.

§5.4.2 Región II

Esta región está caracterizada por las siguientes condiciones: $\det(A) > 0$, $\text{tr}(A) > 0$ y $D(A) < 0$. Si A es una matriz que corresponde a esta región, entonces sus valores propios son números complejos con parte real positiva. Ya se vio con anterioridad que las soluciones a este problema son combinación de funciones de la forma $e^{\alpha t} \sin \beta t$ y $e^{\alpha t} \cos \beta t$, con $\alpha > 0$. Por lo tanto, la solución general oscilará con el tiempo. Resulta que la forma típica en este caso es la de una espiral. Además, si $X(0) \neq \mathbf{0}$, la solución satisface lo siguiente: $\lim_{t \rightarrow -\infty} X(t) = \mathbf{0}$ y $\lim_{t \rightarrow \infty} |X(t)| = \infty$. Decimos que el punto fijo $\mathbf{0}$ es una **espiral repulsora**. La diferencia esencial con la región I consiste en que las funciones oscilan con el tiempo. En la figura 5.8 se ilustra el aspecto típico de las soluciones cercanas una espiral repulsora.

§5.4.3 Región III

En este caso, la matriz satisface lo siguiente: $\det(A) > 0$, $\text{tr}(A) < 0$ y $D(A) < 0$. Se puede ver fácilmente que los valores propios de una matriz correspondiente a esta región son complejos con parte real negativa. Nuevamente, las soluciones a este problema son combinación de funciones de la forma $e^{\alpha t} \sin \beta t$ y $e^{\alpha t} \cos \beta t$, con $\alpha < 0$, y la forma de las trayectorias es la de una espiral. Al punto fijo se lo denomina **espiral atractora**. Además, si $X(0) \neq \mathbf{0}$, entonces la solución satisface lo siguiente:

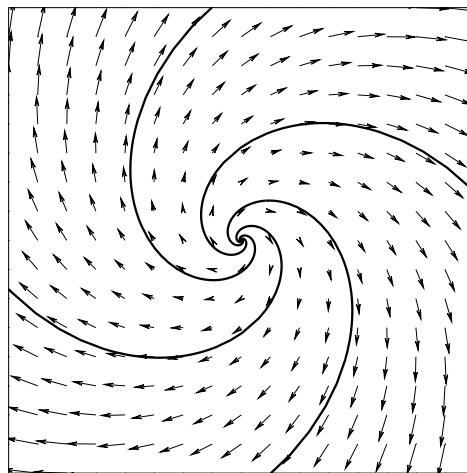


Figura 5.8: Algunas curvas solución alrededor de una espiral repulsora, correspondiente a la región II.

$\lim_{t \rightarrow \infty} X(t) = \mathbf{0}$ y $\lim_{t \rightarrow -\infty} |X(t)| = \infty$. En la figura 5.9 se ilustra el aspecto típico de las soluciones cercanas a una espiral atractora.

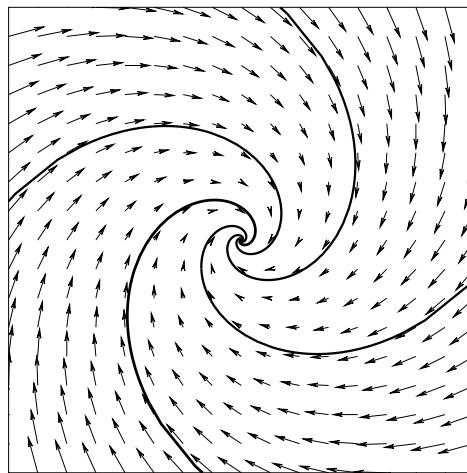


Figura 5.9: Algunas curvas solución alrededor de una espiral atractora, correspondiente a la región III.

§5.4.4 Región IV

Esta región está caracterizada por las siguientes condiciones: $\det(A) > 0$, $\text{tr}(A) < 0$ y $D(A) > 0$. Entonces tenemos dos raíces reales negativas. La dinámica es similar a la de la región I, salvo la dirección en el tiempo. Se dice que el origen es un punto fijo estable o **nodo atractor** y se cumple $\lim_{t \rightarrow \infty} X(t) = \mathbf{0}$. En la figura 5.10 se ilustra el aspecto típico de las soluciones cercanas a un nodo atractor.

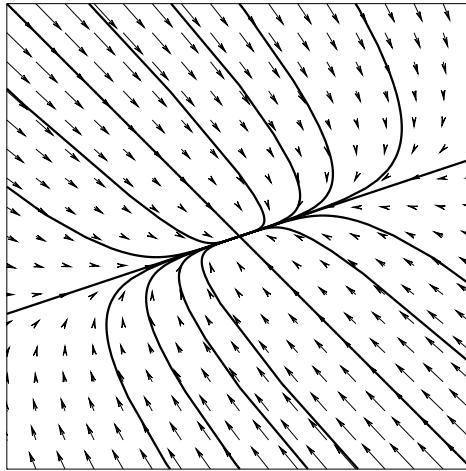


Figura 5.10: Algunas curvas solución alrededor de un nodo atractor, correspondiente a la región IV.

§5.4.5 Región V

Esta región está caracterizada por la condición $\det(A) < 0$, que es independiente del valor de la traza. Se puede ver que la matriz tiene dos raíces reales distintas, con $\lambda_1 < 0$ y $\lambda_2 > 0$. El punto fijo es llamado **punto silla**. Estos puntos son de gran importancia en aplicaciones económicas pues surgen repetidamente en diversos modelos. La forma general de la soluciones está dada por

$$X(t) = C_1 e^{\lambda_1 t} \mathbf{v}_1 + C_2 e^{\lambda_2 t} \mathbf{v}_2.$$

El límite al infinito en esta clase de puntos es más complicado que en los casos anteriores. De hecho, aunque el punto es inestable, existen algunas soluciones que en el futuro se acercan asintóticamente al punto. Del mismo modo, se puede ver que sólo algunas soluciones provienen del punto fijo en el pasado. Por su importancia en economía, estudiaremos con detalle este tipo de solución en la sección 5.6.

El aspecto típico de las soluciones cercanas a un punto silla está dado en la figura 5.2 de la sección anterior. Destacan dos trayectorias: una que converge asintóticamente al estado estacionario y otra que diverge de él. Éstas se conocen como **variedad estable** e **inestable**, respectivamente, y es fácil ver de dónde provienen. Si λ_1 denota al valor propio negativo con vector propio \mathbf{v}_1 y λ_2 es el valor propio positivo con vector propio \mathbf{v}_2 , entonces la variedad estable surge de las condiciones iniciales $\begin{pmatrix} x(0) \\ y(0) \end{pmatrix}$ que implican $c_2 = 0$ en la solución general; análogamente, la variedad inestable surge cuando las condiciones iniciales implican $c_1 = 0$.

De esta forma, si expresamos a \mathbf{v}_1 y a \mathbf{v}_2 como $\begin{pmatrix} 1 \\ \frac{b_1}{a_1} \end{pmatrix}$ y $\begin{pmatrix} 1 \\ \frac{b_2}{a_2} \end{pmatrix}$, es decir, normalizamos la

primera coordenada a la unidad, las ecuaciones de estas trayectorias quedan dadas, respectivamente, por

$$\begin{aligned} \text{variedad estable} &= \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} \\ &= (x(0) - x_p)e^{\lambda_1 t} \mathbf{v}_1 + \begin{pmatrix} x_p \\ y_p \end{pmatrix} \text{ y} \\ \text{variedad inestable} &= \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} \\ &= (x(0) - x_p)e^{\lambda_2 t} \mathbf{v}_2 + \begin{pmatrix} x_p \\ y_p \end{pmatrix}, \end{aligned}$$

en donde $y(0)$ es tal que $c_2 = 0$, en el primer caso, y $c_1 = 0$ en el segundo. $\begin{pmatrix} x_p \\ y_p \end{pmatrix}$ se refiere a la solución particular que resuelve el sistema lineal

$$\begin{aligned} \dot{x} &= 0, \\ \dot{y} &= 0. \end{aligned}$$

Nótese que los vectores propios nos indican la dirección de estas trayectorias. Esto se ilustra en la figura 5.11.

§5.4.6 Casos degenerados

De acuerdo con la definición 5.2.1, los casos degenerados de la figura 5.6 se localizan sobre el eje horizontal, sobre la parte positiva del eje vertical y sobre la parábola determinada por $D(A) = 0$. En el primer caso, $\det(A) = 0$ y se tiene que uno de los valores propios es cero y el conjunto de puntos de equilibrio es una recta. El diagrama de fase correspondiente queda como la figura 5.12.

El segundo caso corresponde a $4\det(A) = \text{tr}^2(A)$, y se tiene un valor propio real repetido. Los puntos de equilibrio son nodos estables o inestables, dependiendo del signo del valor propio y los diagramas de fase son como los de las figuras 5.7 y 5.10.

Finalmente, si $\text{tr}(A) = 0$ y $\det(A) > 0$, se tienen dos valores propios complejos con parte real igual a cero; el punto de equilibrio se llama **centro** y el diagrama de fase es como se muestra en la figura 5.13.

§5.5 Linealización de sistemas no lineales

En esta sección analizaremos la dinámica de sistemas no lineales y autónomos del tipo $\dot{X} = f(X)$, con f un campo vectorial de \mathbb{R}^2 en \mathbb{R}^2 . Los resultados son fácilmente generalizables a \mathbb{R}^n , sin embargo; nos restringiremos a \mathbb{R}^2 para poder utilizar el análisis gráfico.

Consideremos un sistema de ecuaciones no lineales y autónomo de la forma

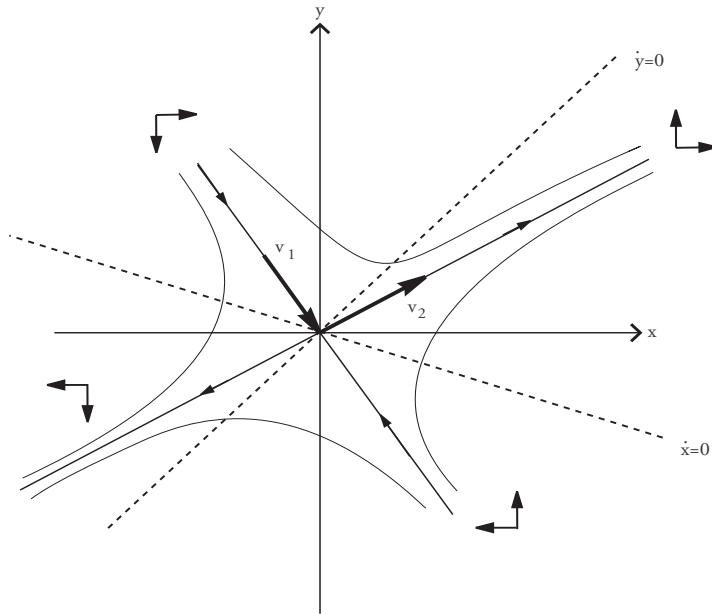


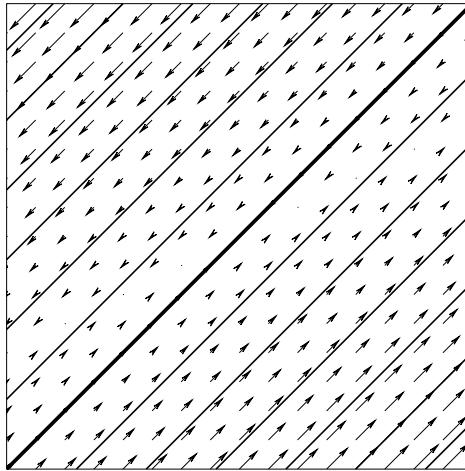
Figura 5.11: Variedades estable e inestable; nótese que la primera está en la dirección de v_1 y la segunda en la de v_2 (aquí la solución particular es el origen).

$$\begin{aligned} \dot{x} &= f(x, y), \\ \dot{y} &= g(x, y). \end{aligned} \tag{5.2}$$

Entonces, los puntos fijos se obtienen resolviendo el sistema,

$$\begin{aligned} f(x, y) &= 0, \\ g(x, y) &= 0. \end{aligned} \tag{5.3}$$

En ocasiones desconocemos la forma específica de las funciones f y g , pero sabemos algunas de sus propiedades, como el que sean doblemente diferenciables o el signo de las derivadas parciales. Esta situación es sumamente común en economía, ya que frecuentemente se tienen funciones “genéricas” (típicas) y se desean conclusiones generales de tipo cualitativo. La forma de proceder es realizando el análisis local alrededor de los puntos fijos (x^*, y^*) que satisfacen el sistema (5.3). Para lograr esto último, se linealiza el sistema (5.2) alrededor de alguno de sus puntos de equilibrio y se analiza el sistema lineal resultante.

Figura 5.12: Diagrama de fase para el caso degenerado con $\det(A) = 0$.

La aproximación lineal de (5.2) alrededor de (x^*, y^*) está dada por

$$\begin{aligned}\dot{x} &\simeq f(x^*, y^*) + f_x(x^*, y^*)(x - x^*) + f_y(x^*, y^*)(y - y^*), \\ \dot{y} &\simeq g(x^*, y^*) + g_x(x^*, y^*)(x - x^*) + g_y(x^*, y^*)(y - y^*).\end{aligned}\quad (5.4)$$

Este es un sistema lineal cuya matriz, conocida como **matriz jacobiana**, está dada por

$$J(x^*, y^*) = \begin{pmatrix} f_x(x^*, y^*) & f_y(x^*, y^*) \\ g_x(x^*, y^*) & g_y(x^*, y^*) \end{pmatrix}.$$

Si tenemos suficiente información acerca de las funciones f y g , entonces podemos determinar el tipo de valores propios que se pueden obtener y por lo tanto las características del punto de equilibrio. Al realizar esta clase de análisis hay que tener en cuenta que se refiere al ámbito “local” y que cualquier cambio en los parámetros que se llegue a considerar debe ser “pequeño”. A continuación se analizan dos ejemplos típicos de linealización.

Ejemplos

Ej 5.5.1 Se desea linealizar el sistema

$$\begin{aligned}\dot{x} &= x + y^2 - 1, \\ \dot{y} &= xy + x^2\end{aligned}$$

y clasificar su puntos de equilibrio. Sea (x^*, y^*) algún punto de equilibrio; entonces la aproximación lineal del sistema alrededor de este punto es,

$$\begin{aligned}\dot{x} &\simeq (x - x^*) + 2y^*(y - y^*), \\ \dot{y} &\simeq (y^* + 2x^*)(x - x^*) + x^*(y - y^*).\end{aligned}$$

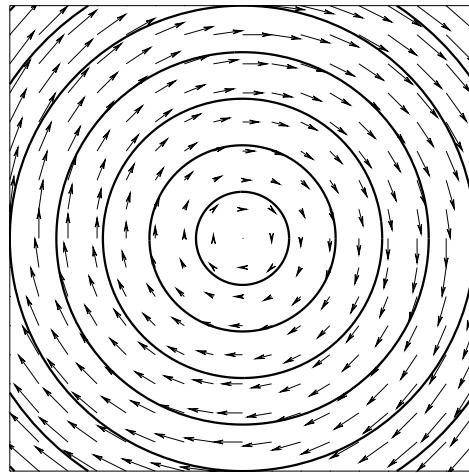


Figura 5.13: Algunas curvas solución alrededor de un centro.

La matriz jacobiana asociada al sistema linealizado alrededor de (x^*, y^*) es

$$J(x^*, y^*) = \begin{pmatrix} 1 & 2y^* \\ y^* + 2x^* & x^* \end{pmatrix}.$$

Los puntos de equilibrio se obtienen resolviendo el sistema

$$\begin{aligned} 0 &= x + y^2 - 1, \\ 0 &= xy + x^2. \end{aligned}$$

Obtenemos así los cuatro puntos $(0, 1)$, $(0, -1)$, $(-1.62, 1.62)$ y $(0.62, -0.62)$. De esta forma, para los primeros dos puntos de equilibrio tenemos

$$\begin{aligned} J(0, 1) &= \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix}, \\ J(0, -1) &= \begin{pmatrix} 1 & -2 \\ -1 & 0 \end{pmatrix}, \end{aligned}$$

ambas matrices con valores propios $\lambda_1 = -1$ y $\lambda_2 = 2$. La matriz

$$J(-1.62, 1.62) = \begin{pmatrix} 1 & 3.24 \\ -1.62 & -1.62 \end{pmatrix}$$

tiene valores propios $\lambda_1 = -0.31 + 1.88i$ y $\lambda_2 = -0.31 - 1.88i$, y finalmente

$$J(0.62, -0.62) = \begin{pmatrix} 1 & -1.24 \\ 0.62 & 0.62 \end{pmatrix}$$

posee valores propios $\lambda_1 = 0.81 + 0.856i$ y $\lambda_2 = 0.81 - 0.856i$.

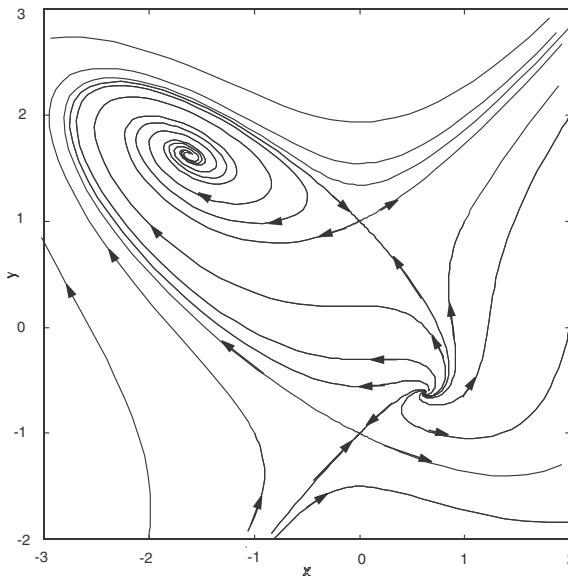


Figura 5.14: Diagrama de fase correspondiente al ejemplo 5.5.1.

Vemos entonces que $(0, 1)$ y $(0, -1)$ son puntos silla $(-1.62, 1.62)$ es un punto espiral estable y $(0.62, -0.62)$ un punto espiral inestable. La figura 5.14 ilustra estos resultados.

Ej 5.5.2 Ahora vamos a encontrar y clasificar los puntos fijos del siguiente sistema:

$$\begin{aligned}\dot{y}_1 &= y_1^2 + y_2^2 - 2, \\ \dot{y}_2 &= y_1^2 - y_2^2.\end{aligned}$$

Los puntos fijos satisfacen

$$\begin{aligned}y_1^2 + y_2^2 - 2 &= 0, \\ y_1^2 - y_2^2 &= 0.\end{aligned}$$

Por lo tanto éstos son $(1, 1)$, $(1, -1)$, $(-1, 1)$ y $(-1, -1)$.

La matriz jacobiana asociada al sistema lineal, alrededor de (y_1, y_2) , está dada por

$$J(y_1, y_2) = \begin{pmatrix} 2y_1 & 2y_2 \\ 2y_1 & -2y_2 \end{pmatrix}.$$

Entonces,

$$J(1, 1) = \begin{pmatrix} 2 & 2 \\ 2 & -2 \end{pmatrix}.$$

Como el determinante de la matriz es negativo, se tiene que $(1, 1)$ es punto silla. Por otra parte,

$$J(1, -1) = \begin{pmatrix} 2 & -2 \\ 2 & 2 \end{pmatrix}.$$

En este caso $\text{tr}(J(1, -1)) = 4$, $\det(J(1, -1)) = 8$ y $D(J(1, -1)) = -16$, por lo que se trata de una espiral repulsora. Para el punto $(-1, 1)$, la matriz jacobiana es

$$J(-1, 1) = \begin{pmatrix} -2 & 2 \\ -2 & -2 \end{pmatrix},$$

ahora se tiene que $\text{tr}(J(-1, 1)) = -4$, $\det(J(-1, 1)) = 8$ y $D(J(-1, 1)) = -16$ por lo que se trata de una espiral atractora. Finalmente, para $(-1, -1)$

$$J(-1, -1) = \begin{pmatrix} -2 & -2 \\ -2 & 2 \end{pmatrix}.$$

Basta notar que el determinante es negativo, por lo tanto se trata de un punto silla. En la figura 5.15 se puede apreciar el comportamiento global de las soluciones.

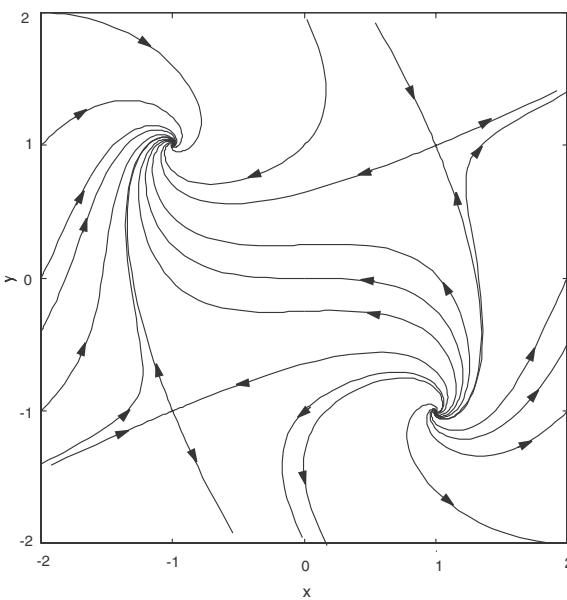


Figura 5.15: Diagrama de fase para el ejemplo 5.5.2.



§5.6 Análisis de puntos silla

Los sistemas que poseen puntos silla son los más utilizados en economía, especialmente en modelos macroeconómicos. La idea detrás de esto es que, dado un valor inicial de alguna de las variables, existe un valor inicial único de las variables restantes tal que el sistema se ubica en la variedad estable. Así, por ejemplo, si comenzamos con un valor inicial de capital, existe un único valor inicial para el consumo

que nos coloca sobre la trayectoria que converge al estado estacionario, o bien, dado un nivel inicial de la producción, existe un único nivel inicial de precios que nos pone en la trayectoria hacia el estado estacionario.

Tener sistemas con equilibrios de tipo atractor implica que cualquier condición inicial nos conduce al estado estacionario, y un equilibrio de tipo repulsor nos lleva al otro extremo, en el cual ninguna condición inicial nos conduce al equilibrio. El mundo económico sería entonces uno dentro del cual nuestras decisiones iniciales serían intrascendentes para lograr la estabilidad,¹ lo que se puede interpretar como dirigirse asintóticamente hacia un estado estacionario.

Consideremos un sistema lineal $\dot{X} = AX$ de 2×2 , donde \mathbf{v}_1 y \mathbf{v}_2 son vectores propios linealmente independientes, correspondientes al par de valores propios $\lambda_1 < 0 < \lambda_2$. Para todas las condiciones iniciales de la forma $c_1\mathbf{v}_1$, la trayectoria solución converge a cero. La dirección determinada por el vector \mathbf{v}_1 es llamada **dirección estable**. Del mismo modo, si iniciamos en la recta generada por \mathbf{v}_2 , es decir si $X(\mathbf{0}) = c_2\mathbf{v}_2$, entonces el límite $\lim_{t \rightarrow \infty} X(t) = \lim_{t \rightarrow \infty} c_2 e^{\lambda_2 t} \mathbf{v}_2$ no existe y la trayectoria diverge. La dirección determinada por \mathbf{v}_2 es llamada **dirección inestable**.

Si $X(t) = c_1 e^{\lambda_1 t} \mathbf{v}_1 + c_2 e^{\lambda_2 t} \mathbf{v}_2$, c_1 y $c_2 \neq 0$, entonces en el futuro ($t \rightarrow \infty$) la solución se aproxima a $c_2 e^{\lambda_2 t} \mathbf{v}_2$ (dirección inestable) y en el pasado ($t \rightarrow -\infty$) la solución se aproxima a $c_1 e^{\lambda_1 t} \mathbf{v}_1$ (dirección estable). En general, tenemos que la figura que corresponde a la dinámica cercana a un punto silla lineal es similar a la mostrada en 5.11. Al valor propio negativo λ_1 se le conoce en economía como **tasa de convergencia**. La razón es evidente, ya que efectivamente representa la tasa a la que la solución converge hacia el equilibrio sobre la variedad estable.

Las direcciones estable e inestable, y sus variedades correspondientes, son sumamente sencillas de calcular, como se ilustra en el siguiente ejemplo.

Ejemplos

Ej 5.6.1 Se desea encontrar las variedades estable e inestable y la tasa de convergencia para el sistema dado por

$$\dot{x} = -3x + 2y,$$

$$\dot{y} = -2x + 2y.$$

Éste es un sistema homogéneo y el origen es su único punto de equilibrio. La ecuación característica está dada por $\lambda^2 + \lambda - 2 = 0$, que tiene como raíces $\lambda_1 = -2$ y $\lambda_2 = 1$, de manera que, en efecto, el equilibrio es un punto silla con tasa de convergencia igual a -2 . Los vectores propios correspondientes están dados por $\mathbf{v}_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ y $\mathbf{v}_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, y la solución general queda dada por

$$x(t) = 2C_1 e^{-2t} + C_2 e^t,$$

$$y(t) = C_1 e^{-2t} + 2C_2 e^t.$$

¹Por ejemplo, sería irrelevante el cumplir o no con las restricciones presupuestales intertemporales pues a fin de cuentas lograremos la estabilidad.

Sobre la variedad estable, la condición inicial es tal que $C_2 = 0$, de manera que se tiene

$$\begin{aligned}x(t) &= 2C_1e^{-2t}, \\y(t) &= C_1e^{-2t},\end{aligned}$$

o bien, eliminando C_1e^{-2t} , se obtiene la recta o variedad estable

$$y(t) = \frac{1}{2}x(t).$$

Nótese que, efectivamente, se trata de la recta que pasa por el punto de equilibrio, $(0, 0)$ en este caso, y cuya dirección está dada por el vector

$$\mathbf{v}_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

De forma análoga, la variedad inestable es la recta, $y(t) = 2x(t)$.

Ej 5.6.2 Utilizando el sistema del ejemplo anterior, calculamos $y(0)$ si $x(0) = x_0$ es conocida y el sistema debe estar sobre la variedad estable. Puesto que se debe tener $C_2 = 0$, sustituyendo la condición inicial se tiene el siguiente sistema:

$$\begin{aligned}x(0) &= x_0 = 2C_1, \\y(0) &= C_1.\end{aligned}$$

De aquí se obtiene

$$\begin{aligned}C_1 &= \frac{x_0}{2}, \\y(0) &= \frac{x_0}{2};\end{aligned}$$

esto determina el único valor de $y(0)$ que coloca al sistema sobre la variedad estable.



En general, para sistemas lineales de más de dos dimensiones, existen puntos silla si el conjunto de valores propios reales distintos contiene valores tanto positivos como negativos. Para este caso, se tiene que las direcciones estables son aquellas determinadas por los vectores propios con valor propio real negativo. De forma similar, las direcciones inestables son las determinadas por aquellos vectores propios con valor propio real positivo. El análisis de puntos silla se puede extender también a modelos no lineales como se muestra a continuación.

Definición 5.6.1 Sea \mathbf{p}^* un punto silla del sistema no lineal $\dot{\mathbf{X}} = f(\mathbf{X})$. El conjunto de condiciones iniciales tales que la solución correspondiente converge a \mathbf{p}^* cuando $t \rightarrow \infty$ es llamado **variedad estable**. Análogamente, el conjunto de condiciones iniciales tales que la solución converge a \mathbf{p}^* cuando $t \rightarrow -\infty$ es llamado **variedad inestable**. Se denota por $W^s(\mathbf{p}^*)$ a la variedad estable y por $W^u(\mathbf{p}^*)$ a la inestable.²

²Se utilizan los superíndices s y u del inglés “stable” y “unstable”.

La definición anterior es simplemente la generalización de la definición de variedad estable e inestable vista en la sección 5.4.5.

Definición 5.6.2 *Dados cualquier subespacio $V \subset \mathbb{R}^n$ y un elemento $\mathbf{p} \notin V$, en \mathbb{R}^n , el conjunto*

$$\mathbf{p} + V = \{\mathbf{q} \in \mathbb{R}^n \mid \mathbf{p} - \mathbf{q} \in V\}$$

*es llamado **subespacio afín**. Intuitivamente, éste se puede imaginar como el subespacio V pero con el origen trasladado al punto \mathbf{p} .*

Podemos dar ahora la siguiente definición.

Definición 5.6.3 *Sea \mathbf{p}^* un punto silla del sistema no lineal $\dot{X} = f(X)$. Considérese el sistema linealizado alrededor de \mathbf{p}^* . Sea $V^s = \langle \mathbf{v}_1, \dots, \mathbf{v}_k \rangle$ el subespacio generado por los vectores propios con valores propios reales negativos. El subespacio afín $\mathbf{p}^* + V^s$ se denota por $E^s(\mathbf{p}^*)$ y se llama **espacio estable**. Asimismo, sea $V^u = \langle \mathbf{v}_{k+1}, \mathbf{v}_2, \dots, \mathbf{v}_m \rangle$ el subespacio generado por los vectores propios con valores propios reales positivos. El subespacio afín $\mathbf{p}^* + V^u$ se denota por $E^u(\mathbf{p}^*)$ y es llamado **espacio inestable**.*

En el caso de dos dimensiones, que es el que podemos analizar gráficamente, $W^s(\mathbf{p}^*)$ y $W^u(\mathbf{p}^*)$ son curvas en el plano, y $E^s(\mathbf{p}^*)$ y $E^u(\mathbf{p}^*)$ son rectas que pasan por \mathbf{p}^* . La relación geométrica existente entre estas variedades y los subespacios afines está dada por el siguiente teorema que se enuncia sin demostración.

Teorema 5.6.4

Sean $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ un campo vectorial, $\dot{X} = f(X)$ un sistema no lineal y \mathbf{p}^* un punto silla del sistema. $W^s(\mathbf{p}^*)$ y $W^u(\mathbf{p}^*)$ son curvas diferenciables que pasan por \mathbf{p}^* y además son tangentes a los espacios estable $E^s(\mathbf{p}^*)$ e inestable $E^u(\mathbf{p}^*)$.

La figura 5.16 ilustra el teorema 5.6.4.

Ejemplo

Ej 5.6.3 Analicemos el punto silla del siguiente sistema:

$$\begin{aligned}\dot{y}_1 &= y_2 + 1 - e^{y_1}, \\ \dot{y}_2 &= y_2 e^{y_1}.\end{aligned}$$

Observemos que $\mathbf{p}^* = (0, 0)$ es el único punto de equilibrio del sistema. Las isoclinas del sistema son curvas que satisfacen $\dot{y}_1 = 0$ y $\dot{y}_2 = 0$. En este caso, si $\dot{y}_1 = 0$, entonces $y_2 = e^{y_1} - 1$. Para y_2 tenemos que $\dot{y}_2 = 0$ implica $y_2 e^{y_1} = 0$ y por lo tanto $y_2 = 0$. Nótese que el gradiente sobre la curva $\dot{y}_1 = 0$ apunta hacia arriba y a la izquierda. Análogamente el gradiente sobre la curva $\dot{y}_2 = 0$ apunta hacia arriba, con lo cual se determinan las regiones en donde $\dot{y}_1 > 0$ y $\dot{y}_2 > 0$. Las isoclinas, junto con las flechas de flujo, se observan en la figura 5.17.

La matriz jacobiana que aproxima linealmente al sistema está dada por

$$J(y_1, y_2) = \begin{pmatrix} -e^{y_1} & 1 \\ y_2 e^{y_1} & e^{y_1} \end{pmatrix},$$

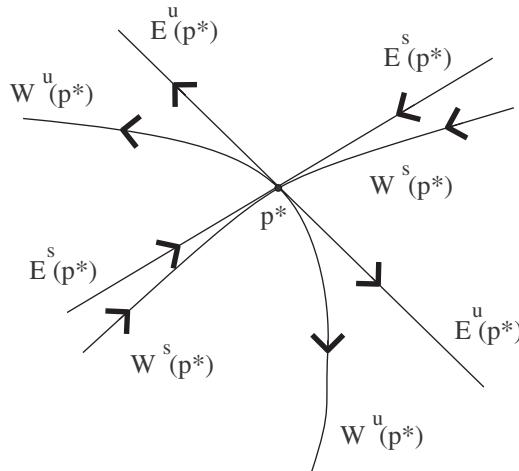


Figura 5.16: Variedades estable, $W^s(p^*)$, e inestable $W^u(p^*)$, y espacios estable, $E^s(p^*)$, e inestable, $E^u(p^*)$. Obsérvese que los espacios son tangentes a las variedades en p^* .

que evaluada en $\mathbf{p}^* = (0, 0)$ queda como

$$J(0, 0) = \begin{pmatrix} -1 & 1 \\ 0 & 1 \end{pmatrix}.$$

El polinomio característico es $\lambda^2 - 1$, con raíces $\lambda_1 = -1$ y $\lambda_2 = 1$. Para $\lambda_1 = -1$, el vector propio es $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, que corresponde a la dirección estable. Para $\lambda_2 = 1$, el vector propio es $\mathbf{v}_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, correspondiente a la dirección inestable. En $\mathbf{p}^* = (0, 0)$, la variedad estable $W^s(\mathbf{p}^*)$ tiene una tangente con la misma pendiente que \mathbf{v}_1 , es decir, una pendiente igual a 0. La línea tangente que aproxima a $W^s(\mathbf{p}^*)$ es $y_2 = 0$. Del mismo modo, $W^u(\mathbf{p}^*)$ se aproxima con la dirección inestable. La aproximación lineal de $W^u(\mathbf{p}^*)$ cerca de \mathbf{p}^* es $y_2 = 2y_1$. La figura 5.18 muestra estas consideraciones, en donde $\langle v_i \rangle$ denota el espacio generado por el vector v_i , $i = 1, 2$.

◇◇◇◇◇◇◇◇◇◇

§5.7 Dos aplicaciones

§5.7.1 Sobreajuste del tipo de cambio

Este modelo es conocido comúnmente como “modelo de overshooting”. Antes de presentarlo,³ daremos algunas consideraciones económicas. Se tienen dos países: uno pequeño, que es el país de casa, digamos

³Modelo basado en [Dor76]

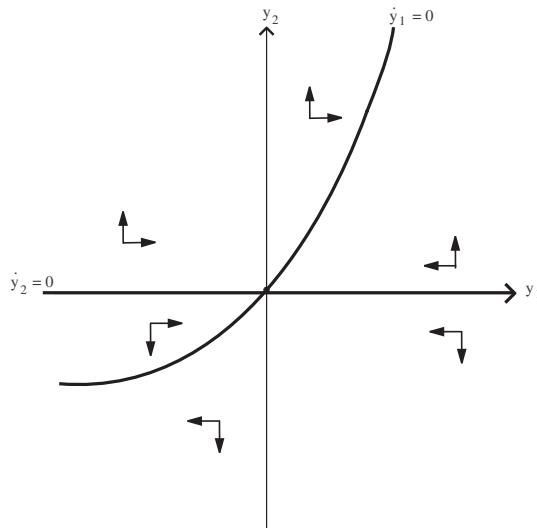


Figura 5.17: Diagrama de fase para el ejemplo 5.6.3.

México, y el otro grande, que es el país extranjero, por ejemplo Estados Unidos. Ambos países están poblados por agentes racionales con previsión perfecta. Se coloca un asterisco a las variables cuando éstas corresponden al país extranjero. Si M y P se refieren a la cantidad nominal de dinero, medida en pesos, y al índice de precios, medida en $\left(\frac{\text{pesos}}{\text{consumible}}\right)$, entonces la cantidad real de dinero, o balances reales, es el cociente M/P , cuyas unidades son simplemente consumibles, es decir, unidades reales de consumo (de ahí el calificativo de *real*).

Pensemos en la tasa nominal de interés como aquella que nos proporciona un instrumento gubernamental local, como un Cete (certificado de la tesorería). Si la tasa nominal de interés aumenta, los agentes económicos tienen un incentivo para comprar Cetes por lo que se deshacen de sus balances reales y adquieren esos instrumentos. Este comportamiento lo podemos expresar simplemente diciendo que los balances reales son inversamente proporcionales a la tasa nominal de interés.

El **tipo de cambio** F está medido en $\left(\frac{\text{unidades de moneda local}}{\text{unidad de moneda extranjera}}\right)$, que en nuestro ejemplo sería $\left(\frac{\text{pesos}}{\text{dólar}}\right)$. El **tipo de cambio real** se define como el cociente FP^*/P , cuyas unidades quedan dadas por

$$\frac{\left(\frac{\text{pesos}}{\text{dólar}}\right) \left(\frac{\text{dólar}}{\text{consumible}^*}\right)}{\left(\frac{\text{peso}}{\text{consumible}}\right)} = \left(\frac{\text{consumible}}{\text{consumible}^*}\right).$$

¿Qué significa esto? Podemos pensar en este cociente como

$$\left(\frac{\text{tortillas}}{\text{oreos}}\right), \left(\frac{\text{tacos al pastor}}{\text{Big Mac}}\right) \text{ o } \left(\frac{\text{Coronas}}{\text{Budweiser}}\right),$$

es decir, nos da el valor de los consumibles o bienes extranjeros en términos de los bienes locales. De

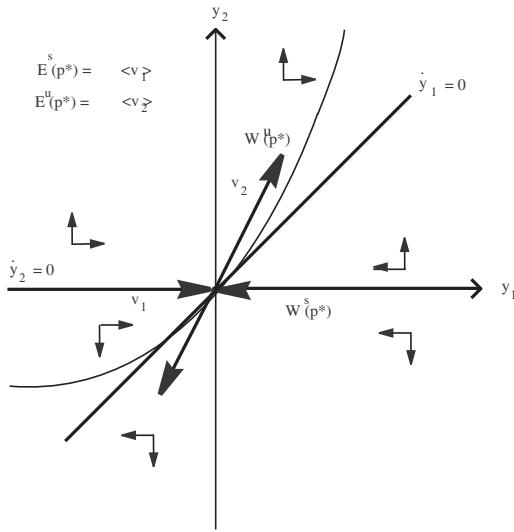


Figura 5.18: Variedades y direcciones estable e inestable del ejemplo 5.6.3. OJO CHECAR !!! caption más explicativo...

ahí viene el nombre de “tipo de cambio real”. Si el tipo de cambio real aumenta, entonces los bienes extranjeros son más “caros” con respecto a los bienes locales o, equivalentemente, los bienes locales son más “baratos” con respecto a los extranjeros, lo que promueve el aumento en las exportaciones nacionales. Debido a esta observación, es natural suponer que la producción de bienes nacionales o ingreso Y es proporcional al tipo de cambio real.

Para convertir el modelo en uno lineal, todas las variables se expresan en logaritmos, con la siguiente notación:

- $m = \ln M$,
- $p = \ln P$,
- $f = \ln F$,
- $y = \ln Y$.

Si además i e i^* denotan, respectivamente, las tasas de interés nominal del país casa y del país extranjero, entonces las observaciones hechas anteriormente implican las siguientes relaciones:

$$m - p = -\mu i, \quad (5.5)$$

$$y = \beta(f + p^* - p). \quad (5.6)$$

Aquí μ y β son constantes positivas, m es la oferta nominal de dinero que suponemos constante y la expresión $f + p^* - p$ en la segunda ecuación representa el logaritmo del tipo de cambio real.

Asumamos que i^* es constante; esto sucede si \dot{p}^* , la inflación en el país extranjero, es cercana a 0 y por lo tanto la tasa de interés nominal es igual a la tasa real. Adicionalmente, asumimos que la economía es totalmente abierta en ambos países. De esta forma, la inflación en el país casa está dada por el cambio porcentual en el tipo de cambio:

$$\dot{f} = \frac{d \ln F}{dt}.$$

La condición de paridad en la tasa de interés implica la relación

$$i = i^* + \dot{f}. \quad (5.7)$$

Esto simplemente nos dice que no existe arbitraje en el mercado de bonos de ambos países. Finalmente, tenemos una relación tipo curva de Phillips⁴ en el sentido de que la inflación (esperada) es proporcional al ingreso

$$\dot{p} = \theta y. \quad (5.8)$$

Sin pérdida de generalidad, las constantes p^* e i^* las normalizamos a cero (recuérdese que están expresados en logaritmos). De esta forma las ecuaciones (5.5) y (5.7) implican

$$\dot{f} = \frac{p - m}{\mu} \quad (5.9)$$

y las ecuaciones (5.6) y (5.8) nos dan

$$\dot{p} = \beta \theta (f - p). \quad (5.10)$$

Estas últimas dos ecuaciones forman un sistema lineal en las variables f y p , que puede expresarse en forma matricial como

$$\begin{pmatrix} \dot{f} \\ \dot{p} \end{pmatrix} = \begin{pmatrix} 0 & \frac{1}{\mu} \\ \beta \theta & -\beta \theta \end{pmatrix} \begin{pmatrix} f \\ p \end{pmatrix} + \begin{pmatrix} \frac{-m}{\mu} \\ 0 \end{pmatrix}.$$

El determinante del sistema es negativo, por lo que se tiene que los valores propios son reales y de signos diferentes. Explícitamente, los valores propios están dados por

$$\lambda_{1,2} = \frac{1}{2} \left(-\beta \theta \pm \sqrt{(\beta \theta)^2 + \frac{4\beta \theta}{\mu}} \right). \quad (5.11)$$

El estado estacionario se obtiene resolviendo para $\dot{f} = \dot{p} = 0$ y está dado por, $(p^*, f^*) = (m, m)$. Cuando el sistema se ajusta a lo largo de la variedad estable, el valor absoluto del valor propio negativo, digamos λ_1 , nos da la tasa de convergencia al estado estacionario. En este caso, teniendo la forma explícita para λ_2 , puede verse cómo afectan los diversos parámetros a esta tasa de convergencia. Por ejemplo, si la demanda por balances reales es más sensible a la tasa nominal de interés, de manera que μ es mayor, entonces $|\lambda_2|$ disminuye y el sistema converge a una tasa menor (más lentamente). El diagrama de fase correspondiente es como en la figura 5.19, en donde se ve que el estado estacionario es un punto silla.

Supongamos que el sistema se encuentra en su estado estacionario, (m, m) , y que el índice de precios p en el país casa es “pegajoso”, en el sentido de que los precios no se ajustan instantáneamente en respuesta

⁴Los agentes poseen previsión perfecta, de manera que la inflación esperada es igual a la inflación real.

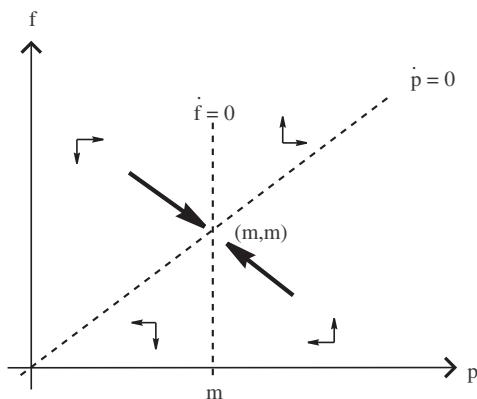


Figura 5.19: Variedad estable para el modelo de sobreajuste del tipo de cambio. La recta $\dot{p} = 0$ es simplemente la identidad y el estado estacionario es el punto (m, m) .

a cambios en los parámetros. Si hay un cambio en la política monetaria del país casa y la oferta nominal de dinero aumenta inesperadamente a un nivel \bar{m} , el nuevo estado estacionario, (\bar{m}, \bar{m}) , está arriba y a la derecha del anterior. Dada la condición inicial $p_0 = m$, para colocarse en la variedad estable el tipo de cambio f_0 debe ser como en la figura 5.20.

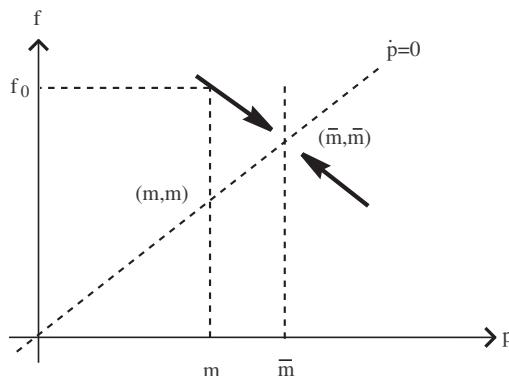


Figura 5.20: Aumento en balances nominales ante un aumento en la oferta nominal de dinero de un valor inicial m a otro $\bar{m} > m$.

Notamos que f_0 no sólo es mayor que su valor anterior, $f = m$, sino que también es mayor que su nuevo valor de equilibrio \bar{m} . A este fenómeno se le conoce como **“sobreajuste” del tipo de cambio**. Las trayectorias que siguen p y f en el tiempo están dadas en las figuras 5.21 y 5.22 que representan el

siguiente modelo numérico:

$$\begin{aligned}\dot{p} &= f - p, \\ \dot{f} &= 2p - 1, \\ p(0) &= \frac{1}{4}, \\ f(0) &= \frac{3}{4},\end{aligned}$$

con soluciones $p(t) = -\frac{1}{4}e^{-2t} + \frac{1}{2}$ y $f(t) = \frac{1}{4}e^{-2t} + \frac{1}{2}$.

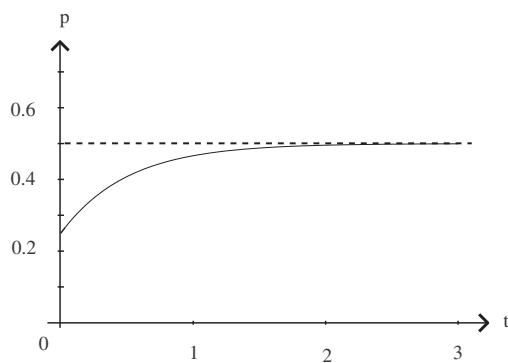


Figura 5.21: Trayectoria de p .

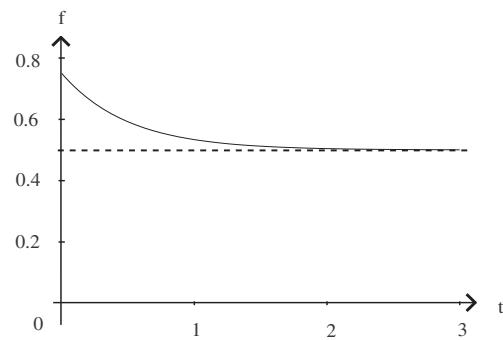


Figura 5.22: Trayectoria de f .

§5.7.2 Devaluación: niveles vs. tasas

Como en el ejemplo anterior, se tiene un país pequeño (país casa) con una economía abierta⁵ y un país grande (país extranjero), llamado “resto del mundo”. Supongamos que sólo existe un bien consumible

⁵Adaptado de [Cal81]

que se produce tanto en el país casa como en el resto del mundo y que la población del país pequeño consiste en una familia y un Banco Central, ambos con previsión perfecta. El peso es la moneda local y el dólar la del resto del mundo. Con la notación del ejemplo anterior se puede pensar al tipo de cambio como el cociente $F = \frac{P}{P^*}$, ya que esto no es más que

$$\frac{\frac{\text{pesos}}{\text{consumible}}}{\frac{\text{dólares}}{\text{consumible}}} = \frac{\text{pesos}}{\text{dólar}}.$$

Normalizando $P^* = 1$ y suponiendo que no hay inflación en el resto del mundo, se puede identificar al tipo de cambio con el índice de precios P . De esta forma, la tasa de inflación $\pi = \frac{\dot{P}}{P}$ es igual a la tasa de depreciación del tipo de cambio $\frac{\dot{F}}{F}$.

La cantidad real de dinero, o balances reales, está dada por $m = \frac{M}{P}$, en donde M es la cantidad nominal de dinero o balances nominales. (Nótese que en el ejemplo anterior m denotaba a $\ln M$ y aquí representa a los balances reales.) La producción es un flujo exógeno, el maná venido del cielo por ejemplo, que denotaremos por y , y la trayectoria de consumo de la familia está dada por c . La inflación se puede pensar como la “tasa de depreciación” de los balances reales.

Suponemos que el Banco Central tiene la posibilidad de imprimir cualquier cantidad de balances nominales sin costo (**dinero fiduciario**). La familia posee una función de utilidad $u(c, m)$ que satisface $u_c, u_m > 0$; $u_{cc}, u_{mm} < 0$, y $u_{cm} = 0$. La restricción presupuestal de la familia, válida para toda t , está dada por

$$y + g = c + \dot{m} + m\pi,$$

en donde, del lado izquierdo tenemos los ingresos dados por la producción y y una transferencia (“lump sum”) g hecha por el Banco Central a la familia. Esta última proviene de los ingresos que el banco obtiene a través de la inflación; concretamente, el banco imprime dinero nominal a costo 0, obtiene consumibles reales a cambio y éstos los transfiere de regreso a la familia. Asimismo, el banco fija la tasa de devaluación, igual a la inflación, a un nivel dado. Del lado derecho de la ecuación se tienen los egresos dados por el consumo c , la inversión en balances reales \dot{m} , y la depreciación de los mismos $m\pi$ (nótese la analogía con la inversión bruta en capital físico dada por $\dot{k} + \delta k$, con δ la tasa de depreciación del capital).

Dada una cantidad inicial de balances reales m_0 , la familia encuentra una trayectoria óptima para el consumo tal que maximiza el valor presente de su utilidad tomando en cuenta su restricción presupuestal. El valor presente se toma descontando a una tasa ρ que es la tasa subjetiva de descuento de la familia (esta tasa representa qué tanto se descuenta el futuro). Una vez determinada la trayectoria de consumo, los balances reales m , quedan dados por la restricción presupuestal y la condición inicial m_0 . Posteriormente, en el capítulo 12, se resolverá este problema de optimización dinámica; por lo pronto, damos la solución como

$$\frac{\dot{u}_c}{u_c} = \pi + \rho - \frac{u_m}{u_c}. \quad (5.12)$$

Esta última ecuación se puede interpretar como una función de demanda de dinero.

En el problema agregado debe tomarse en cuenta la restricción presupuestal del Banco Central dada

por $m\pi = g$. La restricción presupuestal agregada queda entonces como,

$$\dot{m} = y - c. \quad (5.13)$$

Asimismo, notando que $\dot{u}_c = u_{cc}\dot{c}$, reescribimos la función de demanda de dinero (5.12) como

$$\dot{c} = \frac{1}{u_{cc}} [-u_m + u_c(\pi + \rho)]. \quad (5.14)$$

Las ecuaciones (5.13) y (5.14) forman un sistema no lineal cuyo estado estacionario, (m^*, c^*) , satisface

$$\begin{aligned} y - c^* &= 0, \\ -u_m^* + u_c^*(\pi + \rho) &= 0 \end{aligned}$$

$(u_m^* \text{ y } u_c^*$ significa que evaluamos las funciones marginales en el punto de equilibrio (m^*, c^*)).

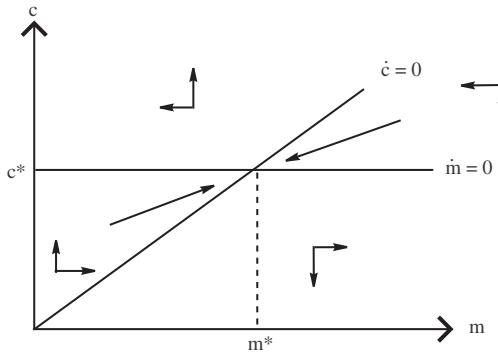


Figura 5.23: Diagrama de fase y trayectoria estable hacia el estado estacionario (m^*, c^*) .

Linealizando alrededor del estado estacionario se obtiene⁶

$$\begin{aligned} \dot{m} &\simeq -(c - c^*), \\ \dot{c} &\simeq \frac{-u_{mm}^*}{u_{cc}^*} (m - m^*) + (\pi + \rho)(c - c^*). \end{aligned}$$

La matriz de este sistema está dada por

$$A = \begin{pmatrix} 0 & -1 \\ -\frac{u_{mm}^*}{u_{cc}^*} & \pi + \rho \end{pmatrix}.$$

Nótese que $\det(A) = -\frac{u_{mm}^*}{u_{cc}^*} < 0$, lo que implica que se tienen dos valores propios reales de signos opuestos y por lo tanto el estado estacionario es un punto silla. El diagrama de fase queda como en la figura 5.23, en donde se muestra la trayectoria estable.

⁶La única parte que no es obvia es $\frac{d\dot{c}}{dc} |_{(c^*, m^*)} = \frac{(u_{cc}^*)^2(\pi + \rho) - [-u_m^* + u_c^*(\pi + \rho)]u_{cc}^*}{(u_{cc}^*)^2} = \pi + \rho$, dado que en el estado estable se tiene que $-u_m^* + u_c^*(\pi + \rho) = 0$.

Vamos a analizar qué pasa cuando el Banco Central devalúa la moneda local (peso) de manera inesperada y el sistema se encuentra inicialmente en el estado estacionario. Dado que $F = P$, una devaluación equivale a aumentar el nivel de precios y por lo tanto a una “destrucción” de balances reales $m = \frac{M}{P}$. Si en $t = 0$ ocurre la devaluación y ahora se tiene un nivel de precios \bar{P} , igual al tipo de cambio \bar{F} , mayor que el nivel inicial, entonces la condición inicial se convierte en $m_0 = M/\bar{P} < m^*$. El consumo baja a un nivel c_0 y el sistema se mueve a lo largo de la variedad estable hacia el equilibrio anterior (m^*, c^*). Durante la transición, la baja en el consumo implica una mejora en la balanza de pagos (la producción no cambia). Esto se ilustra en la figura 5.24.

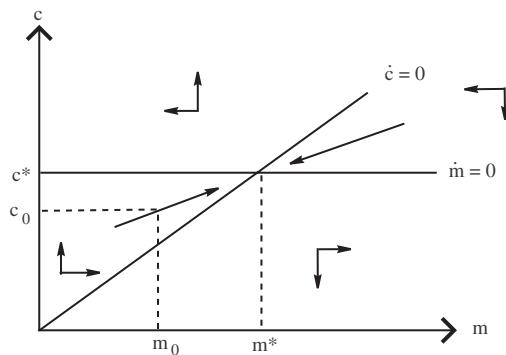


Figura 5.24: Efecto de una devaluación. Los balances reales toman un valor inicial de m_0 . El consumo debe ser c_0 para colocarse en la trayectoria estable rumbo al estado estacionario (m^*, c^*) .

Si en lugar de devaluar la moneda inesperadamente, el Banco Central cambia la tasa de devaluación (cambio en la tasa de inflación, no en el nivel de precios), entonces la recta $\dot{c} = 0$ se mueve hacia la izquierda y se tiene un nuevo estado estacionario (m^{**}, c^*) , con el mismo nivel de consumo que antes pero menor cantidad de balances reales. Dada la condición inicial $m_0 = m^*$, el consumo inicial es tal que el sistema se coloca sobre la nueva variedad estable, como se ve en la figura 5.25. Nótese que, durante el periodo de ajuste, el consumo es mayor y, por lo tanto, hay un deterioro en la balanza de pagos.

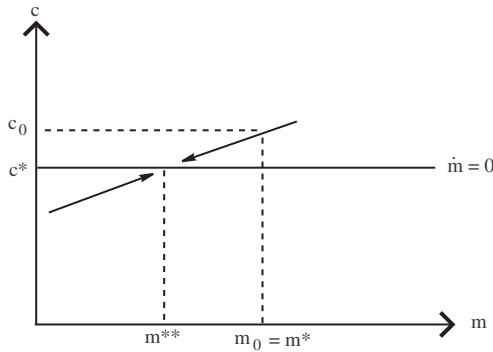


Figura 5.25: Efecto de un aumento en la tasa de devaluación. Se observa el nivel de consumo inicial c_0 que debe tenerse para llegar al nuevo estado estacionario, (m^{**}, c^*) .

Ejercicios

▷ 5.1 Encontrar y clasificar los puntos fijos de los siguientes sistemas y realizar el diagrama de fase correspondiente:

$$\text{a) } \dot{X} = \begin{pmatrix} -1 & 3 \\ -2 & 1 \end{pmatrix} X.$$

$$\text{b) } \dot{X} = \begin{pmatrix} 0 & 1 \\ -2 & -2 \end{pmatrix} X.$$

$$\text{c) } \dot{X} = \begin{pmatrix} 1 & 3 \\ -2 & -6 \end{pmatrix} X.$$

$$\text{d) } \dot{X} = \begin{pmatrix} 1 & 3 \\ -2 & 11 \end{pmatrix} X.$$

$$\text{e) } \dot{X} = \begin{pmatrix} 4 & -1 \\ 1 & 2 \end{pmatrix} X.$$

$$\text{f) } \dot{X} = \begin{pmatrix} 0 & 3 \\ 2 & 100 \end{pmatrix} X.$$

$$\text{g) } \dot{X} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} X.$$

$$\text{h) } \dot{X} = \begin{pmatrix} 3 & 3 \\ -1 & 2 \end{pmatrix} X.$$

$$\text{i) } \dot{X} = \begin{pmatrix} -1 & 2 \\ -2 & -1 \end{pmatrix} X.$$

$$\text{j) } \dot{X} = \begin{pmatrix} 2 & 3 \\ -2 & 1 \end{pmatrix} X + \begin{pmatrix} 9 \\ 9 \end{pmatrix}.$$

▷ 5.2 Encontrar y clasificar los puntos fijos de los siguientes sistemas:

$$\text{a) } \dot{X} = \begin{pmatrix} 3 & 1 & -1 \\ 1 & 3 & -1 \\ 3 & 3 & -1 \end{pmatrix} X.$$

$$\text{b) } \dot{X} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 7 & 1 \\ 2 & 1 & 7 \end{pmatrix} X.$$

$$\text{c) } \dot{X} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} X.$$

$$\text{d) } \dot{X} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 7 & 1 \\ 2 & 1 & 7 \end{pmatrix} X + \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix}.$$

▷ **5.3** En los siguientes problemas, x y y representan poblaciones de especies distintas que compiten entre sí. El sistema dinámico que describe su evolución es la versión vectorial del modelo logístico de crecimiento. Describir los diagramas de fase correspondientes y analizar el tipo de equilibrio y la evolución de las poblaciones.

a) $\dot{x} = x(2 - x - y)$, $\dot{y} = y(6 - 2x - y)$.

b) $\dot{x} = x(6 - 2x - y)$, $\dot{y} = y(2 - x - y)$.

c) $\dot{x} = x(2 - 2x + y)$, $\dot{y} = y(6 + x - 2y)$. Este último caso no describe un sistema en competencia sino uno en simbiosis. Explicarlo.

▷ **5.4** Considerar el siguiente sistema:

$$\dot{P} = P(1 - P) - \gamma PN,$$

$$\dot{N} = N + \gamma PN,$$

donde γ es un parámetro positivo. Encontrar las direcciones estable e inestable del punto silla $(P^*, N^*) = (1, 0)$.

▷ **5.5** El ingreso y y el índice de precios p se relacionan de acuerdo con el siguiente sistema lineal:

$$\begin{aligned}\dot{y} &= ay - p, \\ \dot{p} &= y - bp + ab - 1,\end{aligned}$$

donde $a, b > 0$.

a) ¿Para qué valores de a y b se tiene un comportamiento “cíclico” de las variables?

b) ¿Para qué valores de a y b es estable el sistema?

▷ **5.6** Sea p el nivel de precios y w el salario nominal. El cambio en el salario está dado por

$$\dot{w} = A(w - ap),$$

y la inflación \dot{p} está determinada por el cambio en el salario y por la presión en la demanda, de manera que satisface la ecuación

$$\dot{p} = B\dot{w} + C(w - ap).$$

Adicionalmente se tiene que se cumplen $A, B, C, a > 0$ y $a(AB + C) > A$. Resolver el sistema y analizar el tipo de equilibrio que se obtiene.

▷ **5.7** Considerar el siguiente sistema lineal:

$$\dot{X} = \begin{pmatrix} 2\alpha & -1 \\ \alpha^2 - \beta^2 & 0 \end{pmatrix} X,$$

donde se supone que $\beta > \alpha > 0$.

a) Explicar por qué $p^* = (0, 0)$ es un punto silla.

b) Encontrar los espacios lineales estable e inestable, E^s y E^u , correspondientes a este punto.

▷ **5.8** En este ejercicio se introduce el concepto de estabilidad asintótica.

a) Sea \mathbf{p}^* un punto fijo atractor de una ecuación diferencial. Demostrar que se satisface la siguiente condición: existe un número $\delta > 0$ tal que, si una solución $\mathbf{p}(t)$ cumple que $\mathbf{p}(0) = \mathbf{p}_0$ y $|\mathbf{p}_0 - \mathbf{p}^*| < \delta$, entonces $\lim_{t \rightarrow \infty} \mathbf{p}(t) = \mathbf{p}^*$. Es decir, la solución tiende al equilibrio cuando $t \rightarrow \infty$ siempre y cuando el valor inicial esté suficientemente cerca de \mathbf{p}^* . Se dice, entonces, que el equilibrio es **asintóticamente estable**.

b) Considerar el sistema

$$\begin{aligned}\dot{x} &= -y - x(x^2 + y^2), \\ \dot{y} &= x - y(x^2 + y^2).\end{aligned}$$

Demostrar que $(0, 0)$ es un punto fijo degenerado, y sin embargo es asintóticamente estable. (sugerencia: usar coordenadas polares.)

▷ **5.9** Considerar el siguiente modelo de crecimiento económico. (Modelo de Ramsey-Cass-Koopmans.)

$$\begin{aligned}\dot{k} &= f(k) - nk - c, \\ \dot{c} &= \frac{(f'(k) - \rho)c}{\theta}.\end{aligned}$$

- a) Asumiendo que $f(k) = k^{0.5}$, $\rho = 0.5$, $n = 0.2$ y $\theta = 0.5$. Encontrar un punto de equilibrio (k^*, c^*) tal que $k^* > 0$ y $c^* > 0$. Explicar por qué (k^*, c^*) es un punto silla.
- b) Encontrar las ecuaciones para la aproximación tangente de las variedades estable e inestable cerca de (k^*, c^*) .
- c) Suponer que el capital per cápita inicial es

$$k_0 = 0.1 + k^*.$$

Aproximadamente, ¿qué valor debe tener el consumo per cápita inicial c_0 para que el sistema converja al punto de equilibrio (k^*, c^*) ?

▷ **5.10** Considerar la ecuación no lineal $\ddot{y} = 2 \sin 2y$, con condiciones iniciales $y(0) = u$, $\dot{y}(0) = v$. Encontrar condiciones necesarias para u y v si deseamos que la solución satisfaga,

$$\begin{aligned}\lim_{t \rightarrow \infty} y(t) &= 0, \\ \lim_{t \rightarrow \infty} \dot{y}(t) &= 0.\end{aligned}$$

(Sugerencia: escribir la ecuación como un sistema no lineal de primer orden.)

▷ **5.11** En general, es muy difícil encontrar expresiones analíticas para las curvas estables e inestables; sin embargo, esto es posible en algunos casos. Consideremos el siguiente sistema no lineal:

$$\begin{aligned}\dot{x} &= 2xy, \\ \dot{y} &= 1 - 3x^2 - y^2.\end{aligned}$$

a) Encontrar los puntos silla del sistema y calcular los espacios estable e inestable de los sistemas lineales asociados.

b) Explicar por qué, si $(x(t), y(t))$ es una solución que inicia en (x_0, y_0) , entonces la pendiente de la curva solución debe satisfacer la siguiente ecuación:

$$\frac{dy}{dx} = \frac{\dot{y}}{\dot{x}} = \frac{1 - 3x^2}{2x} y^{-1} - \frac{1}{2x} y.$$

- c) Obsérvese que la ecuación anterior es de Bernoulli. Resolverla considerando a x como la variable independiente. (Respuesta: $x(y^2 + x^2 - 1) = C = x_0(y_0^2 + x_0^2 - 1)$.)
- d) Encontrar una expresión analítica para las curvas estable e inestable de todos los puntos silla.

▷ **5.12** Linealizar cada uno de los siguientes sistemas alrededor de su(s) punto(s) de equilibrio y realizar los diagramas de fase correspondientes:

- a) $\dot{x} = 4x - 3xy$, $\dot{y} = 3y - xy$.
- b) $\dot{x} = 3y^2 + x$, $y = -(3x^2 + y)$.

▷ **5.13** Linealizar el siguiente sistema alrededor de su(s) punto(s) de equilibrio y realizar los diagramas de fase correspondientes:

$$\begin{aligned}\dot{k} &= I(p) - \delta k, \\ \dot{p} &= p(r + \delta) - f'(k).\end{aligned}$$

Aquí k representa el capital, p el precio del mismo, δ su tasa de depreciación, r la tasa real de interés, $I(p)$ es la función de oferta de inversión bruta que cumple con $I'(p) > 0$; finalmente, f es la función de producción de la firma que satisface $f'(k) > 0$ y $f''(k) < 0$.

- a) Describir el ajuste del sistema cuando $k_0 < k^*$ y $k_0 > k^*$, es decir, describir lo que pasa con las variables k , p , $I(p)$ cuando el sistema avanza por la variedad estable.

- b)** ¿Cuál es la tasa de convergencia a (k^*, p^*) a lo largo de la variedad estable? (Calcular el valor propio negativo.) ¿Qué pasa con esta tasa si aumenta r ?
- ▷ **5.14** Considérese el siguiente sistema no lineal:
 $\dot{x} = x - 1$, $\dot{y} = xe^x - y$.
- a)** Encontrar los puntos fijos del sistema y clasificarlos.
- b)** Para cada punto silla \mathbf{p} , encontrar los espacios $E^s(\mathbf{p})$ y $E^u(\mathbf{p})$ de la aproximación lineal.
- c)** Usando la regla de la cadena, sabemos que, si $(x(t), y(t))$ es una solución del sistema, entonces la pendiente de la curva debe satisfacer
- $$\frac{dy}{dx} = \frac{\dot{y}}{\dot{x}} = \frac{xe^x - y}{x - 1}.$$
- Resolver la ecuación anterior, considerando a x como la variable independiente.
- d)** Encontrar, para cada punto silla \mathbf{p} , expresiones analíticas para las curvas estable $W^s(\mathbf{p})$ e inestable $W^u(\mathbf{p})$.
- e)** En cada caso, graficar $W^s(\mathbf{p})$ y $W^u(\mathbf{p})$.

Parte III

Ecuaciones en Diferencias y Dinámica Discreta

CAPÍTULO 6

Conceptos básicos de dinámica discreta

§6.1 Introducción

Cuando el tiempo t es una variable continua una relación del tipo

$$\dot{X} = f(X), \quad (6.1)$$

que describe la evolución de una función X (que toma valores en \mathbb{R}^n) se la denomina un **sistema dinámico**. En general, la variable temporal puede ser continua o discreta y también se denomina como sistema dinámico a la descripción de la evolución de una función X con valores en \mathbb{R} y dominio \mathbb{N} . En la primera parte del libro nos dedicamos al estudio de los sistemas dinámicos continuos y en esta segunda parte daremos una introducción a los sistemas dinámicos discretos.

Existen diversos modelos económicos en donde el tiempo está expresado de manera discreta. Cualquier simulación numérica de un proceso continuo es, en realidad, un proceso discreto, pues una computadora sólo puede actuar en el mundo de tales procesos. Por otra parte, se ha visto que en los procesos discretos no es tan fácil obtener soluciones explícitas a las ecuaciones. Tradicionalmente, los economistas han utilizado los modelos discretos directamente y se ha mostrado que aun los más sencillos poseen gran complejidad y riqueza. En ocasiones, estos modelos pueden representar más acertadamente la realidad económica. Adicionalmente, al “discretizar” el tiempo es posible introducir elementos estocásticos de manera relativamente sencilla, como veremos en el capítulo 9.

En general, en un modelo discreto se utilizan períodos: primer periodo, segundo periodo, etc. Si consideramos al tiempo como una variable discreta que puede tomar valores $t = 0, 1, 2, \dots$, entonces una función que depende de esta variable, $X : \mathbb{N} \rightarrow \mathbb{R}^n$, es simplemente una sucesión de vectores X_0, X_1, X_2, \dots Supongamos que cada vector está relacionado con el vector previo por medio de una relación dada de manera general por

$$X_{t+1} = f(X_t), \quad (6.2)$$

donde $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ es un campo vectorial. La imagen de la función X es un subconjunto de \mathbb{R}^n , al cual denominamos **conjunto de estados**, y la relación (6.2) nos proporciona una forma de pasar de un periodo al siguiente. Esta situación es bastante común en economía, pues en muchos modelos el valor de las variables en un momento dado depende sólo del valor de las variables en el periodo anterior. En general, podemos decir que los vectores X se toman dentro de un conjunto de estados, E , que es un subconjunto del espacio euclídeo \mathbb{R}^n . De este modo, un sistema dinámico discreto puede ser considerado como una función $f : E \rightarrow E$.

A continuación se dan algunas definiciones que son análogas a las dadas para ecuaciones diferenciales.

Definición 6.1.1 *Un sistema dinámico discreto de primer orden* es una ecuación de la forma $X_{t+1} = f(X_t, t)$. Si la variable t no aparece explícitamente, se dice que el sistema es **autónomo**. El sistema es de **orden m** si es de la forma

$$X_{t+m} = f(X_{t+m-1}, X_{t+m-2}, \dots, X_t, t),$$

es decir, el estado de la variable depende, no sólo del estado en el periodo anterior, sino de los estados en los m periodos anteriores. Es común denotar la función no autónoma $f(X_{t+m-1}, X_{t+m-2}, \dots, X_t, t)$ por $f_t(X_{t+m-1}, X_{t+m-2}, \dots, X_t)$. Si la función f es lineal, decimos que el sistema es **lineal**.

Definición 6.1.2 Denotamos por Δ al operador **diferencia**, definido por $\Delta X_t \equiv X_{t+1} - X_t$. Si aplicamos este operador consecutivamente m veces, tenemos la diferencia de orden m de X_t dada por $\Delta^m X_t \equiv \Delta(\Delta(\dots(\Delta X_t)))$.

Ejemplo

Ej 6.1.1 Si $m = 2$, la diferencia de orden dos es

$$\Delta^2 X_t = \Delta(X_{t+1} - X_t) = (X_{t+2} - X_{t+1}) - (X_{t+1} - X_t) = X_{t+2} - 2X_{t+1} + X_t.$$

Definición 6.1.3 Asociada a un sistema dinámico discreto de orden m se tiene una **ecuación en diferencias de orden m** (si la imagen de X está en \mathbb{R}^n , en realidad se tiene un sistema de n ecuaciones en diferencias); ésta es una ecuación del tipo

$$\Delta^m X_t = g(X_{t+m-1}, X_{t+m-2}, \dots, X_t, t).$$

Lo anterior es claro, ya que dado el sistema $X_{t+m} = f(X_{t+m-1}, X_{t+m-2}, \dots, X_t, t)$, la función g se obtiene de la función f sumando o restando los términos correspondientes para obtener $\Delta^m X_t$ del lado izquierdo de la ecuación. Abusando de la notación, normalmente nos referimos al sistema dinámico original como una ecuación en diferencias de orden m , a pesar de que las diferencias no aparezcan explícitamente. Claramente, las ecuaciones en diferencias son el equivalente discreto de las ecuaciones diferenciales.

Ejemplos

En todos los ejemplos siguientes, x es una función que toma valores en la recta real \mathbb{R} .

Ej 6.1.2 El sistema dinámico descrito por $x_{t+1} = x_t + t - 1$ es lineal, no autónomo y de primer orden. La ecuación en diferencias asociada es $\Delta x_t = x_{t+1} - x_t = t - 1$.

Ej 6.1.3 El sistema dinámico descrito por $x_{t+2} = 4x_{t+1} + 4x_t + 5$ es lineal, autónomo y de segundo orden. La ecuación en diferencias de segundo orden asociada es $\Delta^2 x_t = 2x_{t+1} + 5x_t + 5$.

Ej 6.1.4 El sistema dinámico descrito por $x_{t+1} = x_t^2 + 1$ no es lineal (es cuadrático), es autónomo y de primer orden. La ecuación en diferencias asociada es $\Delta x_t = x_t^2 - x_t + 1$.



Se puede ver que, aumentando el número de las variables, todos los sistemas dinámicos pueden reescribirse como sistemas autónomos de primer orden. Véase el ejercicio 6.1.

En este texto sólo obtendremos soluciones analíticas explícitas para sistemas discretos lineales. (El caso no lineal es mucho más complicado, y aun en dimensiones pequeñas puede producir fenómenos muy interesantes.) Dado un sistema dinámico autónomo de primer orden $X_{t+1} = f(X_t)$ y un valor inicial X_0 , podemos determinar el valor de X_t mediante el siguiente proceso de iteración en donde f^t representa la composición t veces de la función f consigo misma:

$$\begin{aligned} X_1 &= f(X_0), \\ X_2 &= f(X_1) = f(f(X_0)) = f^2(X_0), \\ &\vdots \\ X_t &= f(f(\cdots f(X_0) \cdots)) = f^t(X_0). \end{aligned}$$

A este proceso de iteración lo llamamos **iterar hacia el pasado** puesto que el “presente”, X_t , queda expresado en términos del valor inicial X_0 . Al igual que en el caso continuo, queremos predecir qué sucederá en el futuro, por lo que es de interés el saber qué pasa con el límite

$$\lim_{t \rightarrow \infty} f^t(X_0).$$

Como veremos a continuación, las técnicas que se usan para analizar la solución de las ecuaciones o sistemas lineales son análogas a las del caso continuo. En realidad, como en la primera parte del texto, sólo nos interesa el comportamiento de las soluciones más que la solución analítica en sí.

§6.2 Solución de ecuaciones lineales

La ecuación lineal más sencilla es de la forma

$$x_{t+1} = ax_t + b, \tag{6.3}$$

donde x_t es una variable real y a, b son constantes reales. Tomemos primero el caso homogéneo con $b = 0$. Se tiene que $x_{t+1} = ax_t$, así que dada una condición inicial x_0 obtenemos, iterando,

$$\begin{aligned}x_1 &= ax_0, \\x_2 &= ax_1 = a^2x_0, \\&\vdots \\x_t &= a^tx_0.\end{aligned}$$

Nótese que, en este caso, la función $f : \mathbb{R} \rightarrow \mathbb{R}$ que define al sistema dinámico está dada por $f(x) = ax$ y la t -ésima iteración es $f^t(x) = a^tx$.

Ejemplo

Ej 6.2.1 La solución de la ecuación $x_{t+1} = 2x_t$ con condición inicial $x_0 = 7$ es $x_t = 2^t(7)$.



Definición 6.2.1 Dado el sistema dinámico descrito por $X_{t+1} = f(X_t)$, decimos que X^* es un **punto fijo** o **punto de equilibrio** si se cumple $X^* = f(X^*)$.

Consideremos ahora la ecuación (6.3) con $a \neq 1$ y $b \neq 0$. Podemos obtener un punto fijo de dicha ecuación resolviendo

$$x^* = ax^* + b, \quad (6.4)$$

con lo cual se tiene que

$$x^* = \frac{b}{1-a}. \quad (6.5)$$

Si definimos $y_t = x_t - x^*$, restando (6.4) de (6.3) y reescribiendo en términos de y_t , se obtiene

$$y_{t+1} = ay_t,$$

que se puede resolver mediante el proceso de iteración para llegar a

$$y_t = a^t y_0.$$

Tomando en cuenta la definición de y_t reescribimos la solución como

$$x_t = x^* + a^t(x_0 - x^*)$$

y finalmente, sustituimos (6.5) para obtener la solución general de (6.3):

$$x_t = \frac{b}{1-a} + a^t \left(x_0 - \frac{b}{1-a} \right). \quad (6.6)$$

Como era de esperarse, esta solución es simplemente la suma de la solución general a la ecuación homogénea asociada y la solución particular que se obtiene al poner $x_t = x_{t+1} = x^*$.

Ejemplo

Ej 6.2.2 Encontrar la solución de $x_{t+1} = \frac{1}{2}x_t - 1$, con condición inicial $x_0 = 1$. En este caso, $a = \frac{1}{2}$ y $b = -1$, de manera que (6.6) nos da la solución como

$$x_t = -2 + 3\left(\frac{1}{2}\right)^t.$$

◇◇◇◇◇◇◇◇◇

Dado el sistema dinámico descrito por (6.3) la pregunta obligada es: ¿cómo evoluciona el sistema cuando $t \rightarrow \infty$? Esto se puede contestar fácilmente observando la forma de la solución dada en (6.6).

Notemos que $\lim_{t \rightarrow \infty} x_t$ existe si y sólo si $\lim_{t \rightarrow \infty} a^t$ existe, y esto último sucede si y sólo si $|a| < 1$. En este caso se tiene que $\lim_{t \rightarrow \infty} a^t = 0$ y por lo tanto $\lim_{t \rightarrow \infty} x_t = \frac{b}{1-a} = x^*$; es decir, el sistema converge a su punto fijo. Obsérvese que esta convergencia puede ser monótona, cuando $0 < a < 1$, o bien alternante, si $-1 < a < 0$. Las figuras 6.1 a 6.4 nos muestran los diversos casos para $a = 0.5, -0.5, 2, -2, b = -1$ y $x_0 = 0.25$. Los puntos se han unido para mayor claridad, sin embargo, se debe recordar que estas líneas no son parte de la gráfica.

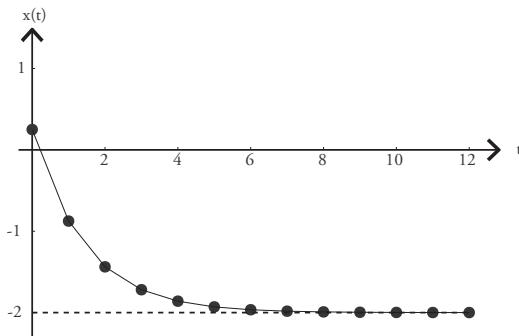


Figura 6.1: Valores de $x(t)$ para $a = 0.5, b = -1$ y $x_0 = 0.25$.

§6.3 Análisis cualitativo

En muchas ocasiones, tenemos sistemas discretos autónomos del tipo $x_{t+1} = f(x_t)$ que no pueden resolverse de forma explícita. Al igual que en el caso continuo, el análisis cualitativo nos da una idea de la evolución del sistema y de la estabilidad de los puntos fijos, que en este caso identificamos con los puntos de equilibrio. En el ámbito discreto, existen resultados análogos a los estudiados en el caso continuo. Por ejemplo, se tiene una definición equivalente a 2.2.2 para sistemas discretos.

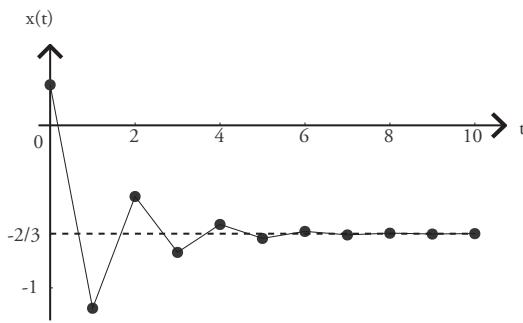


Figura 6.2: Valores de $x(t)$ para $a = -0.5$, $b = -1$ y condición inicial $x_0 = 0.25$.

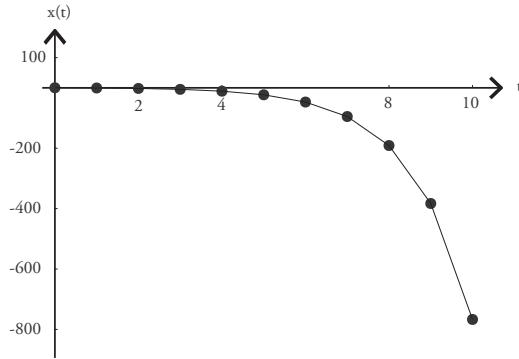


Figura 6.3: Valores de $x(t)$ para $a = 2$, $b = -1$ y condición inicial $x_0 = 0.25$.

Definición 6.3.1 Supongamos que el punto fijo x^* de un sistema dinámico discreto dado por $x_{t+1} = f(x_t)$ satisface la siguiente condición: existe un número $\delta > 0$ tal que se cumple

$$|x_0 - x^*| < \delta \Rightarrow \lim_{t \rightarrow \infty} x_t = x^*,$$

es decir, la solución converge al equilibrio cuando $t \rightarrow \infty$ siempre y cuando el valor inicial esté suficientemente cerca de x^* . Se dice entonces que el equilibrio es **asintóticamente estable**.

De manera análoga, se puede definir inestabilidad para un punto fijo.

Definición 6.3.2 Supongamos que el punto fijo x^* de un sistema dinámico discreto dado por $x_{t+1} = f(x_t)$ satisface la siguiente condición: para todo $\delta > 0$ existen un punto x_1 y un número natural n_0 tales que $|x_0 - x^*| < \delta$ y

$$|x_t - x^*| \geq \delta$$

para todo $t \geq n_0$. Es decir, todos los puntos cercanos a x^* se alejan de x^* . Se dice, entonces, que el equilibrio es **inestable**.

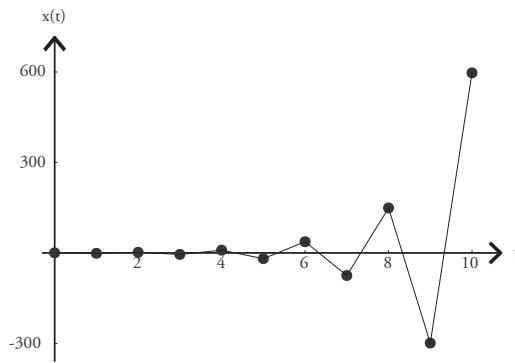


Figura 6.4: Valores de $x(t)$ para $a = -2$, $b = -1$ y condición inicial $x_0 = 0.25$.

Nótese que existen puntos fijos que no son asintóticamente estable ni inestables.

Ejemplo

Ej 6.3.1 Sea $f : \mathbb{R} \rightarrow \mathbb{R}$ el sistema dinámico dado por $f(x) = x + x^2$, es decir, $x_{t+1} = x_t + x_t^2$. El punto $x^* = 0$ es un punto de equilibrio que no es asintóticamente estable ni inestable.

◇◇◇◇◇◇◇◇

A los puntos de equilibrio que no satisfacen ninguna de las condiciones dadas en las anteriores definiciones se les llama **puntos degenerados**.

Con objeto de determinar la estabilidad de un punto fijo, se puede utilizar un método gráfico. Éste consiste en lo siguiente. En el plano $X_t X_{t+1}$ graficamos la función $x_{t+1} = f(x_t)$. Si trazamos la recta de 45° , $x_{t+1} = x_t$, observamos que los puntos fijos o equilibrios del sistema, si existen, están dados por las intersecciones de la gráfica de f con dicha recta, ya que son puntos para los que se cumple $x^* = f(x^*)$. Por analogía con el caso continuo llamamos **diagrama de fase** a este tipo de herramienta gráfica.

La estabilidad, o su ausencia, de acuerdo con la definición 6.3.1 puede encontrarse considerando un punto inicial x_0 alrededor de x^* y analizando la evolución del sistema con la sucesión de puntos $\{x_0, x_1, x_2, \dots\}$. El equilibrio es asintóticamente estable si y sólo si ésta converge a x^* .

Tomenos x_0 suficientemente cerca de x^* , de manera que la aproximación lineal de Taylor sea una buena aproximación a la función, es decir,

$$f(x_0) \simeq f(x^*) + f'(x^*)(x_0 - x^*).$$

Reescribimos esta expresión como

$$x_1 - x^* \simeq f'(x^*)(x_0 - x^*).$$

Nótese que podemos seguir aplicando este procedimiento para x_1, x_2, \dots si y sólo si $|f'(x^*)| < 1$ (así el error en la aproximación lineal cada vez es más pequeño). Procediendo de esta forma se tiene

$$\begin{aligned} f(x_1) - f(x^*) &\simeq f'(x^*)(x_1 - x^*) = [f'(x^*)]^2(x_0 - x^*), \\ \text{por lo que } x_2 - x^* &\simeq [f'(x^*)]^2(x_0 - x^*) \\ &\vdots \\ x_t - x^* &\simeq [f'(x^*)]^t(x_0 - x^*), \end{aligned}$$

con lo cual vemos que

$$\lim_{t \rightarrow \infty} x_t = x^*.$$

De este modo, hemos llegado a un resultado de estabilidad basado en la derivada, tal como lo teníamos para el caso continuo. Es necesario, sin embargo, ser más cuidadosos en cuanto a las aproximaciones anteriores. El siguiente teorema es una expresión formal de lo anterior y es equivalente al teorema 3.3.1.

Teorema 6.3.3

Sea f una función diferenciable. Sea x^* un punto fijo del sistema $x_{t+1} = f(x_t)$ tal que $|f'(x^*)| \neq 1$. Entonces se cumple

- a) Si $|f'(x^*)| < 1$, entonces x^* es asintóticamente estable.
- b) Si $|f'(x^*)| > 1$, entonces x^* es inestable.

Demostración

Se deja como ejercicio para el lector. El ejercicio 6.18 propone los pasos a seguir en la demostración de la parte a. ■

Definición 6.3.4 Dados el sistema dinámico discreto $x_{t+1} = f(x_t)$ y un valor inicial x_0 , la **órbita de x_0** es el conjunto de puntos

$$orb(x_0) = \{x_0, x_1, \dots\}$$

Las figuras 6.5 y 6.6 ilustran el teorema 6.3.3, mostrando dos puntos de equilibrio: uno de ellos asintóticamente estable y el otro inestable, junto con las órbitas que se obtienen del proceso de iteración.

Ejemplo

Ej 6.3.2 (La telaraña) Supongamos que S_t , D_t y p_t denotan la oferta la demanda y el precio del maíz en el periodo t . El precio inicial p_0 es conocido. Se tienen las siguientes relaciones:

$$D_t = -ap_t + b,$$

$$S_t = cp_{t-1} + d,$$

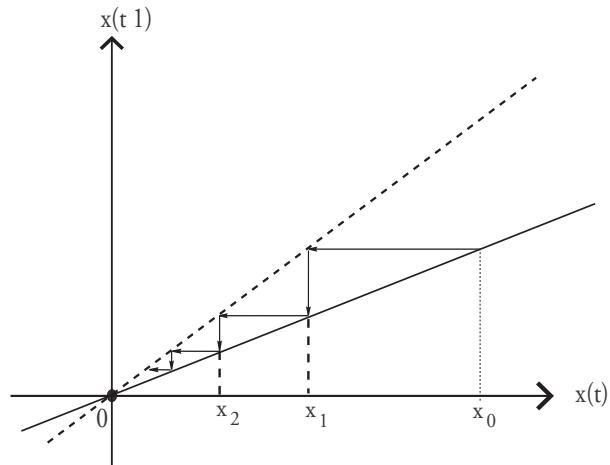


Figura 6.5: Punto fijo estable.

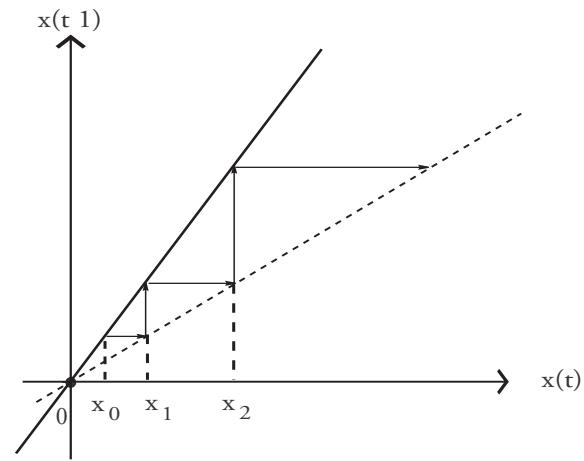


Figura 6.6: Punto fijo inestable.

en donde todos los coeficientes son positivos y cumplen $a > c$ y $b > d$. Notemos que la oferta de maíz en el periodo t depende del precio en el periodo inmediato anterior. En cada periodo se tiene que $D_t = S_t$, por lo que el precio del maíz evoluciona de acuerdo con la ecuación

$$p_t = \left(-\frac{c}{a}\right) p_{t-1} + \frac{b-d}{a},$$

para $t = 1, 2, \dots$

Por (6.6) sabemos que la solución está dada explícitamente por

$$p_t = \left(p_0 - \frac{b-d}{a+c}\right) \left(-\frac{c}{a}\right)^t + \frac{b-d}{a+c}$$

y, dado que $a > c$, ésta converge a su punto fijo $p^* = \frac{b-d}{a+c}$. En el diagrama de fase que se muestra en la figura 6.7 queda claro el porqué del nombre de “telaraña” para este tipo de diagramas.

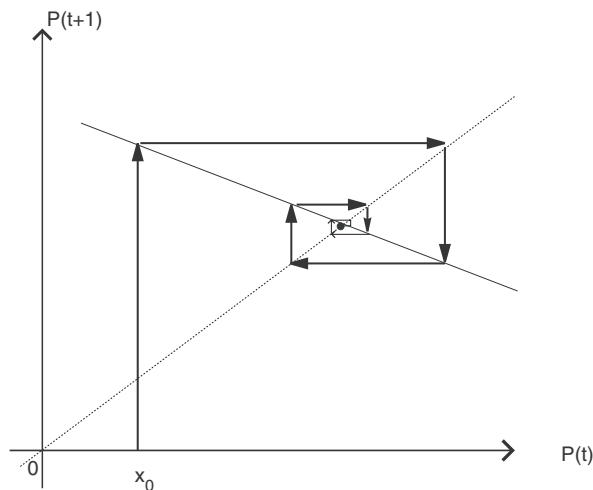


Figura 6.7: Diagrama de telaraña.

En este caso, los granjeros deciden la cantidad de maíz que se va a plantar en $t = 0$ de acuerdo con precio inicial; esta cantidad es la oferta en el periodo siguiente. Si ésta excede la demanda, entonces en $t = 1$ el precio cae y los granjeros disminuyen la oferta para $t = 2$. Ahora existe un exceso de demanda y el precio aumenta en $t = 2$. Los granjeros aumentan la oferta para $t = 3$, etc. El proceso continua así y converge a p^* . Nótese que la convergencia se da siempre que la pendiente de la recta de demanda sea mayor (en valor absoluto) que la de la oferta. Esto se observa en el diagrama 6.8, en donde se grafican cantidades y precios para cada periodo.



Definición 6.3.5 Dados el sistema dinámico discreto $x_{t+1} = f(x_t)$ y un valor inicial x_0 , decimos que la órbita de x_0 tiene **periodo n**, si n es el número más pequeño para el cual se cumple $f(x_{n-1}) = x_0$,

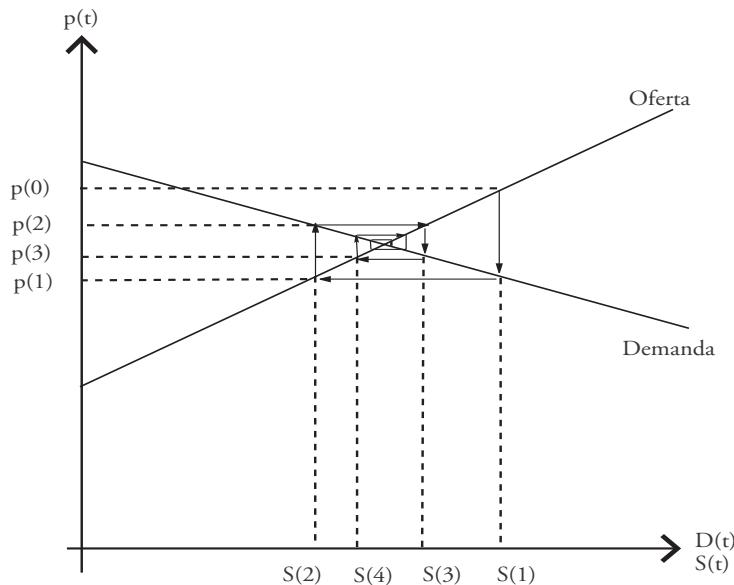


Figura 6.8: Ajuste de la oferta y demanda para el ejemplo 6.3.2.

equivalentemente, $f^n(x_0) = x_0$. En particular, si $\text{orb}(x_0)$ tiene periodo n , la órbita es un conjunto finito con cardinalidad n . Si x_0 tiene una órbita de periodo n , decimos que x_0 es un **punto periódico** de periodo n y que el sistema posee **n -ciclos**.

La definición implica que una condición necesaria, mas no suficiente, para que x_0 tenga periodo n es que sea un punto fijo de f^n . Es claro, también, que si $x_{n-k} \in \text{orb}(x_0)$, entonces x_{n-k} también tiene periodo n y por lo tanto es un punto fijo de f^n . Dado que se tienen puntos de equilibrio, se puede definir la estabilidad de la órbita a partir de lo que ya se sabe para puntos fijos.

Definición 6.3.6 Sea x^* un punto de periodo n del sistema $x_{t+1} = f(x_t)$. Decimos que x^* es **asintóticamente estable** si es estable como punto fijo de f^n . Equivalentemente, se dice que $\text{orb}(x_0)$ es asintóticamente estable, ya que cualquier elemento de la órbita es asintóticamente estable como punto fijo de f^n .

También se tiene lo equivalente para el concepto de inestabilidad.

Definición 6.3.7 Sea x^* un punto de periodo n del sistema $x_{t+1} = f(x_t)$, decimos que x^* es **inestable** si es inestable como punto fijo de f^n . Equivalentemente se dice que $\text{orb}(x_0)$ es inestable ya que cualquier elemento de la órbita es inestable como punto fijo de f^n .

En otro contexto, a las órbitas asintóticamente estables se las llama **órbitas atractoras** y a las inestables **repulsoras**. Cabe la aclaración de que una órbita periódica puede no ser atractora ni repulsora como se muestra en el siguiente ejemplo.

Ejemplo

Ej 6.3.3 Sea $f : \mathbb{R} \rightarrow \mathbb{R}$ el sistema dinámico dado por $x_{t+1} = f(x_t) = x_t^3 - 2x_t$. El 2-ciclo $\{-1, 1\}$ no es asintóticamente estable ni inestable. La demostración de este resultado se deja como ejercicio para el lector (véase ejercicio 6.7).



La siguiente proposición nos da un método para clasificar la estabilidad de los puntos de periodo n .

Proposición 6.3.8 *Sea $x^* = x_0$ un punto de periodo n del sistema $x_{t+1} = f(x_t)$ tal que $\left| \frac{df^n(x_0)}{dx} \right| \neq 1$. Entonces, $orb(x_0)$ es estable si y sólo si*

$$|f'(x_0)f'(x_1)f'(x_2)\dots f'(x_{n-1})| < 1.$$

Al producto $|f'(x_0)f'(x_1)f'(x_2)\dots f'(x_{n-1})|$ se lo llama **multiplicador de la órbita** y lo denotamos por Λ .

Demostración

La demostración se sigue del teorema 6.3.3, ya que $orb(x_0)$ es estable si y sólo si $\left| \frac{df^n(x_0)}{dx} \right| < 1$ y utilizando la regla de la cadena tenemos que

$$\begin{aligned} \left| \frac{df^n(x_0)}{dx} \right| &= |f'(f^{n-1}(x_0))f'(f^{n-2}(x_0))\dots f'(f(x_0))f'(x_0)| \\ &= |f'(x_0)f'(x_1)f'(x_2)\dots f'(x_{n-1})|, \end{aligned}$$

con lo cual concluimos la demostración. ■

Ejemplos

Ej 6.3.4 En particular, si el sistema tiene un punto fijo entonces posee un 1-ciclo ya que un punto fijo tiene periodo 1.

Ej 6.3.5 Por definición, f tiene un 2-ciclo si cumple con

$$\begin{aligned} x_1 &= f(x_0), \\ x_0 &= f(x_1), \end{aligned}$$

y $x_0 \neq x_1$. Notemos que en este caso se tiene que $x_0 = x_2 = x_{2k}$ y $x_1 = x_3 = x_{2k+1}$ para $k = 0, 1, \dots$. Observemos también que $f^2(x_0) = x_0$ y $f^2(x_1) = x_1$.

Ej 6.3.6 Supongamos que se tiene un sistema de la forma

$$x_{t+1} = -x_t + a.$$

Notamos que, dado un valor inicial x_0 , este sistema queda dado por

$$x_0 = x_2 = x_4 = x_{2k},$$

$$a - x_0 = x_1 = x_3 = x_{2k+1}.$$

para $k = 0, 1, \dots$. De esta forma se tiene que el sistema tiene un 2-ciclo. El punto fijo del sistema se obtiene al resolver $x^* = -x^* + a$ y queda dado por $x^* = \frac{a}{2}$; sin embargo, el sistema no se aleja ni se acerca a éste sino simplemente oscila a su alrededor. Notamos que el teorema 6.3.3 no se aplica en este caso ya que $|f'(x^*)| = 1$. La figura 6.9 muestra el diagrama de telaraña resultante con puntos de periodo 2, A y B . Notemos que este ejemplo no cae dentro de las hipótesis de la proposición 6.3.8 pues el multiplicador es $\Lambda = 1$.

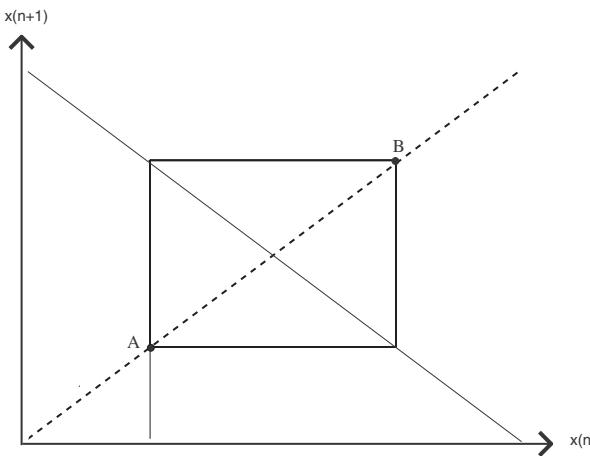


Figura 6.9: 2-Ciclos para el ejemplo 6.3.6.

Ej 6.3.7 La figura 6.10 ilustra otro tipo de 2-ciclo: se trata de un diagrama de telaraña en donde la solución converge asintóticamente a los puntos de periodo 2.

Ej 6.3.8 De igual manera se pueden estudiar órbitas de periodo 3. La figura 6.11 ilustra un 3-ciclo en un diagrama de telaraña en donde la solución converge asintóticamente a los puntos de periodo 3.



§6.4 Caos

Una de las características más interesantes de los sistemas dinámicos es que pueden tener comportamiento caótico. Aunque una definición formal de caos es bastante complicada, se puede identificar herméticamente como un comportamiento netamente desordenado. El primero en descubrir tal comportamiento

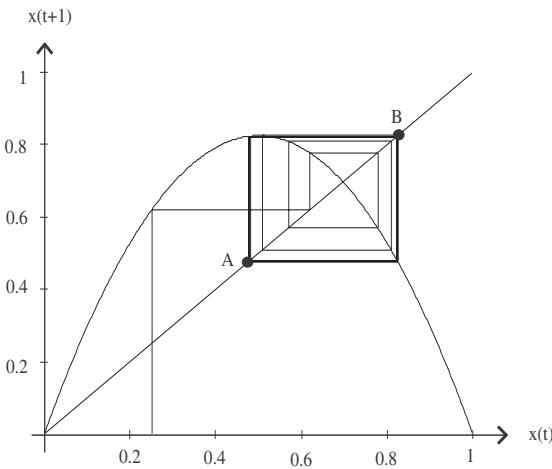


Figura 6.10: 2-Ciclos para el ejemplo 6.3.7.

fue el gran matemático francés Henri Poincaré a finales del siglo XIX. Posteriormente, muchos científicos redescubrieron el concepto y lo aplicaron a diversas ramas del conocimiento.

Recordemos que una ecuación diferencial ordinaria puede ser vista como un sistema dinámico. Sin embargo, no es sencillo observar comportamiento caótico en las soluciones de una ecuación diferencial: se necesita tener al menos un sistema autónomo no lineal de dimensión tres. En franco contraste con esto, existen sistemas discretos relativamente sencillos de formular que presentan este tipo de comportamiento caótico o desordenado.

Aquí daremos únicamente una breve introducción; referimos al lector interesado al libro de Goodwin [Goo91] para una descripción del caos en fenómenos económicos y al de Devaney [Dev89] para el estudio del caos en general. Para tener una idea general sin tecnicismos, se puede ver el libro de Ruelle [Rue91]. Existen diversas aplicaciones en economía y se ha encontrado caos en los más diversos modelos. Los siguientes textos son también útiles: [Dra92], [Rob99], [ASY96], [Str94] y [Sch96].

Querríamos poder medir qué tan inestable es un sistema y demostrar que éste posee un comportamiento caótico. Empecemos por considerar una órbita periódica. Dados un sistema $x_{t+1} = f(x_t)$ y una condición inicial x_0 , consideremos un punto cercano $y_0 = x_0 + \delta_0$ en donde $|\delta_0|$, que mide la separación entre ambos puntos, es “pequeño”. Iteramos repetidamente n veces comenzando con x_0 y y_0 , y definimos $\delta_n \equiv y_n - x_n$, es decir, la separación que existe después de n iteraciones. Si $|\delta_n| \simeq |\delta_0| e^{n\lambda}$, llamamos a λ el **exponente de Liapunov**. Se puede definir, por tanto, el exponente de Liapunov de una órbita periódica como

$$\lambda = \lim_{\delta_0 \rightarrow 0} \frac{1}{n} \ln \left| \frac{\delta_n}{\delta_0} \right|.$$

La siguiente proposición nos da una forma explícita de calcularlo.

Proposición 6.4.1 *Sea $x^* = x_0$ un punto de periodo n del sistema $x_{t+1} = f(x_t)$ con multiplicador de*

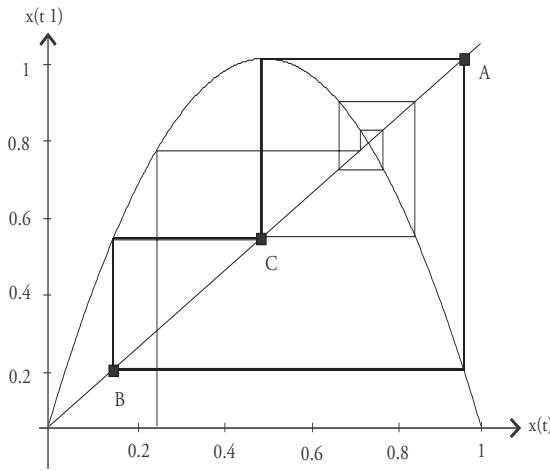


Figura 6.11: 2-Ciclos para el ejemplo 6.3.8.

la órbita dado por Λ ; entonces, podemos expresar al exponente de Liapunov como,

$$\lambda = \frac{1}{n} \ln \Lambda = \frac{1}{n} \sum_{k=0}^{n-1} \ln |f'(x_k)|.$$

Demostración (esbozo)

Tomamos logaritmos a la expresión $|\delta_n| \simeq |\delta_0| e^{n\lambda}$ para obtener

$$\lambda = \frac{1}{n} \ln \left| \frac{\delta_n}{\delta_0} \right|;$$

por definición, tenemos que

$$\delta_n = f^n(x_0 + \delta_0) - f^n(x_0),$$

por lo que reescribimos

$$\frac{1}{n} \ln \left| \frac{\delta_n}{\delta_0} \right| = \frac{1}{n} \ln \left| \frac{f^n(x_0 + \delta_0) - f^n(x_0)}{\delta_0} \right|.$$

Tomando el límite cuando $\delta_0 \rightarrow 0$ esta expresión es igual a

$$\lambda = \frac{1}{n} \ln \frac{df^n(x_0)}{dx} = \frac{1}{n} \ln \Lambda$$

donde Λ es el multiplicador de la órbita. Esto es lo que se quería demostrar. ■

La definición se extiende para el caso de órbitas infinitas siempre y que exista el límite cuando $n \rightarrow \infty$. Tenemos la siguiente definición.

Definición 6.4.2 El exponente de Liapunov de una órbita se define como

$$\lambda = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} \ln |f'(x_k)|,$$

si el límite existe.

La definición anterior propone un método para encontrar evidencia de caos en muchas situaciones. Basta calcular numéricamente el exponente de Liapunov y concluir que, si $\lambda > 0$ existe caos. Se puede ver (cf. ejercicio 6.12) que si una órbita es asintóticamente estable, entonces su exponente de Liapunov es negativo. De este modo, un requisito para que la órbita exhiba lo que llamamos comportamiento caótico es que $\lambda > 0$, ya que esto implica que es “extremadamente” sensible a las condiciones iniciales.

Otra manera de crear un comportamiento caótico está basada en el célebre teorema de Li y Yorke [LY75], el cual exponemos a continuación.

Teorema 6.4.3 (Li y Yorke)

Sean I un intervalo y $f : I \rightarrow I$ una función continua. Supongamos que existe un punto a para el cual los puntos $b = f(a)$, $c = f(b) = f^2(a)$ y $d = f(c) = f^3(a)$ satisfacen

$$d \leq a < b < c$$

o

$$d \geq a > b > c.$$

Entonces, f tiene órbitas periódicas de todos los períodos y además existe un conjunto no numerable $S \subset I$, sin puntos periódicos, tal que

- a) Para cada $p, q \in S$ con $p \neq q$

$$\begin{aligned} \limsup_{n \rightarrow \infty} |f^n(p) - f^n(q)| &> 0, \\ \liminf_{n \rightarrow \infty} |f^n(p) - f^n(q)| &= 0. \end{aligned}$$

- b) Para cada $p \in S$ y $q \in I$ periódico,

$$\limsup_{n \rightarrow \infty} |f^n(p) - f^n(q)| > 0.$$

Lo que dice el teorema, en realidad, es que la dinámica generada por f es muy complicada, en el sentido de que existen muchas órbitas periódicas distintas y además existe dependencia de las condiciones iniciales. Esto es, sin importar lo cercano que inicien dos puntos, en el futuro se separarán. Por otro lado, cada órbita no periódica se mueve arbitrariamente cerca de cualquier órbita. Finalmente, si una órbita no periódica se aproxima a una órbita periódica, después de algún tiempo se separará. En esencia, el comportamiento es impredecible y muy complicado.

§6.5 Modelo logístico discreto

La versión discreta del modelo logístico es el mejor ejemplo de un sistema discreto no lineal sencillo que posee un comportamiento sumamente interesante.¹ Consideraremos el siguiente sistema:

$$x_{t+1} = f(x_t) = cx_t(1 - x_t).$$

¹ Robert May enfatizó este punto en [May76].

Restringimos el valor del parámetro, c , al intervalo $[0, 4]$ de manera que el sistema lleva al intervalo $[0, 1]$ en sí mismo (cf. ejercicio 6.9). Los puntos fijos están dados al resolver,

$$x^* = cx^*(1 - x^*)$$

y en este caso son $x_1^* = 0$ y $x_2^* = \frac{c-1}{c}$. Supongamos que $0 < c < 1$, entonces $x_1^* = 0$ es el único punto fijo dentro de $[0, 1]$. Notemos que

$$f'(x_t) = c(1 - 2x_t),$$

de manera que $f'(0) = c \in (0, 1)$ y el teorema 6.3.3 implica que $x_1^* = 0$ es un punto fijo asintóticamente estable. Es fácil verificar que dada cualquier condición inicial $x_0 \in [0, 1]$ su órbita converge a 0.

Si $1 < c < 3$, entonces $x_2^* = \frac{c-1}{c}$ es otro punto fijo en $[0, 1]$; en este caso, $f'(0) = c > 1$ y $f'(x_2^*) = 2 - c \in (-1, 1)$. De esta forma, $x_1^* = 0$ es inestable y $x_2^* = \frac{c-1}{c}$ es asintóticamente estable. Dado $x_0 \in [0, 1]$, su órbita converge a $\frac{c-1}{c}$.

Supongamos que ahora $3 < c < 4$, vamos a ver que, en este caso, existen 2-ciclos. Necesitamos encontrar los puntos de periodo 2, es decir, los puntos x que satisfacen $f^2(x) = x$. Para este propósito resolvemos la ecuación $f(f(x)) - x = 0$. Esto es,

$$\begin{aligned} c^2x(1-x)[1-cx(1-x)]-x=0, \\ c^3x^4-2c^3x^3+(c^3+c^2)x^2-(c^2-1)x=0. \end{aligned}$$

Tenemos un polinomio de cuarto grado del cual conocemos dos raíces ya que los puntos fijos $x_1^* = 0$ y $x_2^* = \frac{c-1}{c}$ están en el intervalo $[0, 1]$ y necesariamente son raíces dado que $f(x) = x$ implica $f(f(x)) = x$. Factorizamos ambas raíces del polinomio utilizando el algoritmo de la división y obtenemos la siguiente factorización:

$$\begin{aligned} c^3x^4-2c^3x^3+(c^3+c^2)x^2-(c^2-1)x= \\ x\left(x-\frac{c-1}{c}\right)\left(x-\frac{c+1+\sqrt{(c-3)(c+1)}}{2c}\right)\left(x-\frac{c+1-\sqrt{(c-3)(c+1)}}{2c}\right). \end{aligned}$$

Así, los puntos de periodo 2 están dados por

$$\begin{aligned} x_3^* &= \frac{(c+1)+\sqrt{(c-3)(c+1)}}{2c}, \\ x_4^* &= \frac{(c+1)-\sqrt{(c-3)(c+1)}}{2c}. \end{aligned}$$

Notamos que estos no existían cuando $c < 3$ ya que en este caso las raíces no son reales. Para $c = 3$ se tiene que $x_2^* = x_3^* = x_4^* = \frac{2}{3}$ y finalmente, si $c > 3$ se tiene que x_3^* y x_4^* son puntos de periodo 2 y por lo tanto existe un 2-ciclo. Decimos que $c = 3$ es un **punto de bifurcación**, para este valor del parámetro c el sistema sufre un cambio cualitativo en su comportamiento. Para analizar la estabilidad de x_3^* y x_4^* simplemente utilizamos la proposición 6.3.8, de manera que calculamos el multiplicador de la órbita $\{x_3^*, x_4^*\}$ obteniendo,

$$\begin{aligned} \Lambda &= f'(x_3^*)f'(x_4^*) = c(1 - 2x_3^*)c(1 - 2x_4^*) \\ &= c^2[1 - 2(x_3^* + x_4^*) + 4x_3^*x_4^*] \\ &= |4 + 2c - c^2|. \end{aligned}$$

De aquí que la órbita es estable para valores de c que satisfacen $|4 + 2c - c^2| < 1$ o bien si $3 < c < 1 + \sqrt{6} = 3.449\dots$

En las figuras 6.12 a 6.14 mostramos las órbitas correspondientes (los puntos se han unido para mayor claridad), $\{x_n\}$, con $x_0 = 0.25$ para $c = \frac{1}{2}, 2$ y 3.3 .

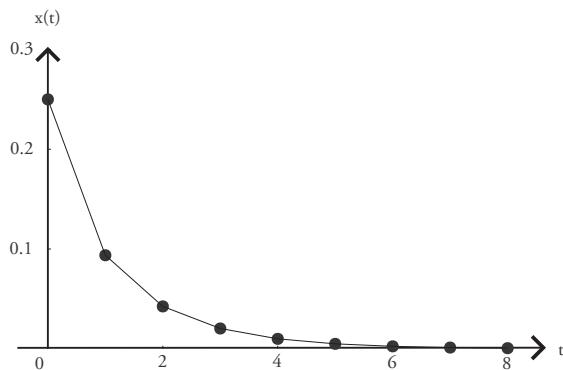


Figura 6.12: Orbita para para $c = 0.5$ y condición inicial $x_0 = 0.25$.

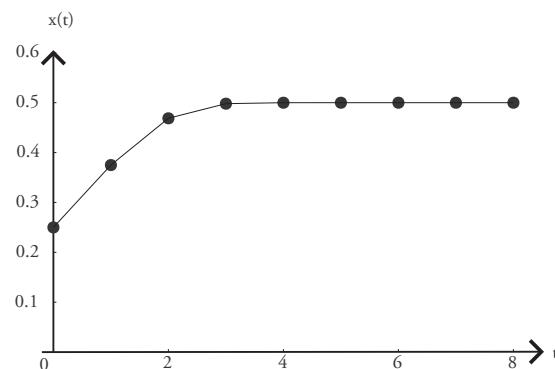


Figura 6.13: Orbita para $c = 2$ y Condición Inicial $x_0 = 0.25$.

La historia del modelo logístico no termina aquí, si $c > 3.449\dots$ se puede ver que aparecen 4-ciclos para $c > 3.54409\dots$ aparecen 8-ciclos, para $c > 3.5644$ aparecen 16-ciclos, para $c > 3.569946$ aparecen 32-ciclos y observamos que las bifurcaciones aparecen cada vez más frecuentemente. Existe un valor límite $c_\infty = 3.569946\dots$ para el cual el periodo es infinito. A continuación ilustramos estas consideraciones graficando en el plano c vs. x^* . A este tipo de diagrama se le conoce como **diagrama de bifurcación**.

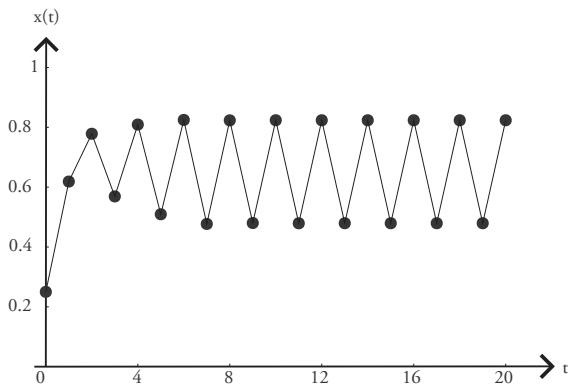


Figura 6.14: Órbita para $c = 3.3$ y con Condición Inicial $x_0 = 0.25$.

Para construir el diagrama se omiten algunas de las primeras iteraciones, digamos cincuenta, de manera que lo que muestra es efectivamente una buena aproximación al comportamiento a largo plazo. Si el diagrama muestra un solo punto asociado a un valor particular del parámetro, c , esto significa que es un punto fijo estable o atractor. Cuando el diagrama se bifurca en dos ramas, quiere decir que ahora existen órbitas de periodo 2 y los puntos asociados a los valores del parámetro son órbitas atractoras de periodo 2. Así sucesivamente se construye el resto del diagrama. Por ejemplo, en la figura 6.15, se muestra la primera bifurcación del modelo logístico discreto. En la figura 6.17 se puede ver la aparición de los 4-ciclos atractores descritos anteriormente.

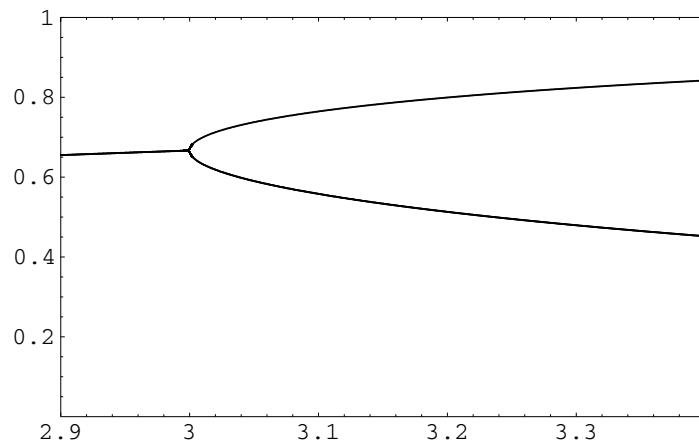


Figura 6.15: Diagrama de bifurcación correspondiente al modelo logístico para valores del parámetro $2.9 \leq c \leq 3.4$. La coordenada horizontal representa al parámetro c y en la vertical se han graficado 100 iteraciones del sistema dinámico.

Es necesario aclarar que en el diagrama solo se muestra el comportamiento asintótico observable numéricamente. Esto es, aunque siguen existiendo órbitas de distintos períodos sus puntos no son observados si éstos resultan, por ejemplo repulsores. En resumen, numéricamente solo se observa el comportamiento estable o caótico, más no el repulsor.

Una pregunta natural es, ¿qué sucede para valores de c mayores a c_∞ ? Por ejemplo, la figura 6.16 muestra la serie de tiempo $\{x_n\}$ cuando $c = 3.9$. Se observa que no existen ciclos definidos, es decir, el comportamiento de la órbita es aperiódico. Adicionalmente, si tomamos órbitas cercanas, éstas se separan con rapidez exponencial, decimos entonces, que el sistema exhibe **comportamiento caótico**.

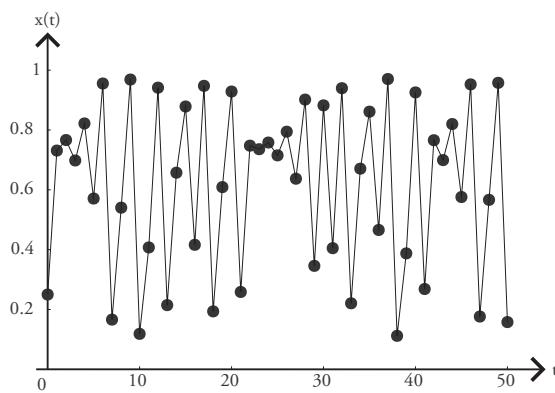


Figura 6.16: Comportamiento caótico.

El diagrama de bifurcación asociado para valores de c cercanos a 3.9 está dado por 6.18.

Notamos que mezclados entre los ciclos de periodo muy grande o infinito existen ciclos pequeños, como por ejemplo, si $c = 3.8358$ se tiene una órbita atractora de periodo 3. En el diagrama se distinguen estas áreas como “lagunas” en las cuales no hay puntos. Los diagramas de telaraña que resultan para $c = 3.8358$ y $c = 3.9$ se ven como en las figuras 6.19 y 6.20.

Como vimos en la sección 6.4, un exponente de Liapunov positivo es un indicador y evidencia de comportamiento caótico, la figura 6.21 muestra estos exponentes para el modelo logístico que estamos estudiando cuando $3 \leq c \leq 4$. Cuando $c < c_\infty$ el exponente siempre es negativo, sin embargo para $c > c_\infty$ oscila entre valores positivos y negativos de manera que tenemos n -ciclos atractores mezclados con períodos de comportamiento caótico.

Ejemplos

Ej 6.5.1 Supongamos que se introduce un nuevo producto en el mercado, éste será adoptado de manera acelerada inicialmente para posteriormente converger a un nivel máximo \bar{x} . Si denotamos por x_t a la cantidad del producto que ha sido adoptada hasta el periodo t , la ecuación logística,

$$x_{t+1} = ax_t \left(1 - \frac{x_t}{\bar{x}}\right)$$

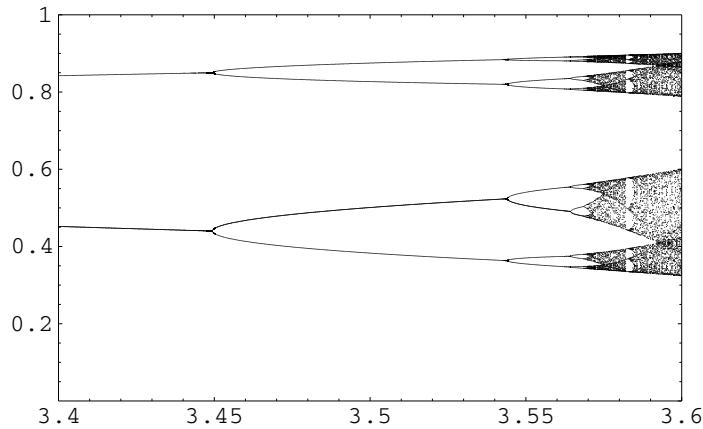


Figura 6.17: Diagrama de bifurcación correspondiente al modelo logístico para valores del parámetro $3.4 \leq c \leq 3.6$. La coordenada horizontal representa al parámetro c y en la vertical se han graficado 100 iteraciones del sistema dinámico.

representa el comportamiento mencionado. Si hacemos el cambio de variable $w_t \equiv \frac{x_t}{\bar{x}}$, esta ecuación se reescribe como,

$$w_{t+1} = aw_t(1 - w_t),$$

que es exactamente el modelo logístico estudiado anteriormente. Notemos que w_t representa el porcentaje adoptado con respecto al nivel máximo. Se concluye de esta manera que para ciertos valores del parámetro, a , se tiene comportamiento caótico.

Ej 6.5.2 Análogamente, podemos pensar en que el precio de este producto crece a alguna tasa inicial la cual disminuye conforme el precio aumenta y éste no puede ser mayor que algún máximo \bar{p} . Así el precio se modela con la ecuación,

$$p_{t+1} = ap_t \left(1 - \frac{p_t}{\bar{p}} \right)$$

al igual que en el ejemplo anterior, se tiene comportamiento caótico para ciertos valores del parámetro a .

Ej 6.5.3 (Modelo de Hicks) El siguiente modelo está tomado de [Puu00]. En la mayoría de los modelos de inversión y crecimiento económico, se establece la hipótesis de que la inversión es proporcional al cambio del ingreso. Esta suposición ha sido criticada en base a hechos empíricos y se ha buscado sustituirla por otra. Sir John Hicks sugirió desechar la hipótesis de linealidad.

Por ejemplo, debe ser claro que, dado que el ingreso no es el único factor en la producción, si el ingreso crece desproporcionadamente entonces los demás factores no se podrán ajustar en la misma manera. Es decir, no existe ningún incentivo para invertir si el cambio en el ingreso de un periodo a otro es muy grande. Del mismo modo, si el cambio en el ingreso es negativo, entonces no necesariamente la inversión tiene que reducirse.

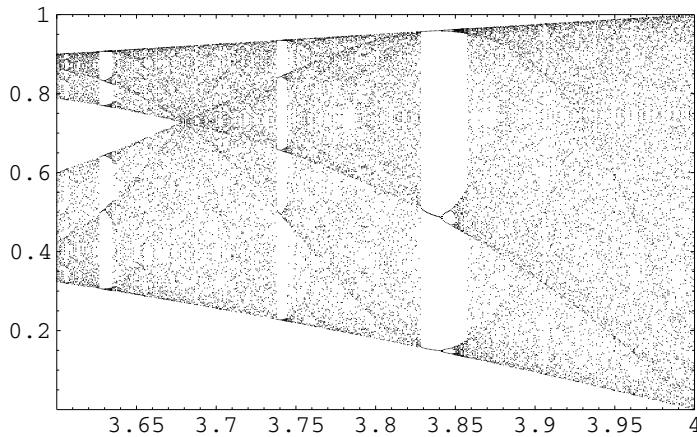


Figura 6.18: Diagrama de bifurcación correspondiente al modelo logístico para valores del parámetro $3.6 \leq c \leq 4.0$. La coordenada horizontal representa al parámetro c y en la vertical se han graficado 100 iteraciones del sistema dinámico.

La crítica de Hicks abre las puertas a nuevas posibilidades. Por ejemplo, supongamos que Y_t representa el ingreso de una fábrica en el periodo t , cantidad que se divide entre consumo C_t e inversión I_t . Esto es,

$$Y_t = C_t + I_t.$$

Supongamos que la relación entre el ingreso y la inversión está dada por una función cúbica, de la siguiente manera,

$$I_t = \nu(Y_{t-1} - Y_{t-2}) - \nu(Y_{t-1} - Y_{t-2})^3 = \nu(\Delta Y_{t-1} - (\Delta Y_{t-1})^3).$$

donde ν es una constante positiva. Por otro lado, supongamos que existe un desfasamiento en el consumo de un periodo. Además, supongamos que existe un propensión al ahorro, $s > 0$, de manera que en cada periodo se ahorra una cantidad s del ingreso para consumirla en el periodo inmediato posterior. Esto se expresa de la siguiente manera,

$$I_t = (1 - s)Y_{t-1} + sY_{t-2}.$$

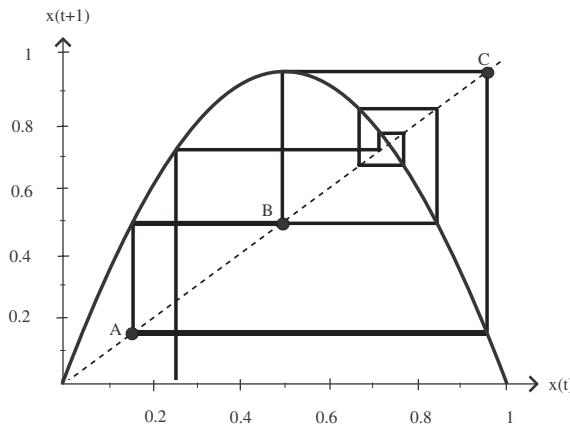
Para analizar el modelo, introducimos la variable $Z_t = \alpha(Y_{t+1} - Y_t)$ donde α es una constante positiva a determinar. Después de varios pasos, se obtiene,

$$Z_t = \lambda Z_{t-1} - (\lambda + 1)Z_{t-1}^3 \quad (6.7)$$

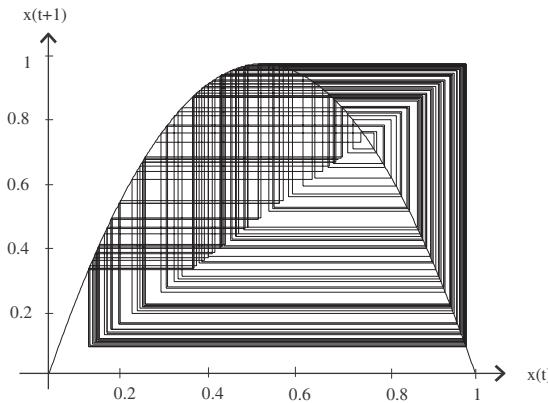
donde $\lambda = \nu - s$. El sistema dinámico resultante puede restringirse al intervalo $[-1, 1]$ si $0 \leq \lambda \leq 3$. El análisis dinámico se deja como ejercicio para el lector. Ver ejercicio 6.15.



Es común, en Economía, pasar de modelos en tiempo continuo a modelos en tiempo discreto o viceversa; adicionalmente, el estudio de las series de tiempo -a fin de cuentas sucesiones $\{x_t\}$ - se realiza

Figura 6.19: Diagrama de telaraña para $c = 3.8358$.

con modelos que en ocasiones no son lineales. El modelador debe tener en mente que el discretizar el tiempo tiene implicaciones profundas para la dinámica del sistema y puede surgir el comportamiento caótico en donde antes no había tal. Un modelo, como el logístico, que en tiempo continuo es totalmente “inofensivo”, se convierte en un modelo con un comportamiento sumamente complejo al tomar periodos discretos de tiempo. El modelador debe tener en cuenta si esta complejidad adicional es o no deseable, ¿es ésta endógena al sistema o un problema de modelación? Creemos que el sentido común es quizás el mejor indicador. El análisis gráfico de los sistemas discretos no lineales debe, en consecuencia, ser tratado con extremo cuidado.

Figura 6.20: Diagrama de telaraña para $c = 3.9$.

Ejercicios

▷ **6.1** Verificar que, aumentando el número de las variables, todos los sistemas dinámicos se pueden reescribir como sistemas autónomos de primer orden.

▷ **6.2** Demostrar la afirmación del ejemplo 6.3.1.

▷ **6.3** Comprobar que las siguientes sucesiones dadas son soluciones de las ecuaciones.

a) $x_t = \cos \pi t; x_{t+1} = -x_t$.

b) $x_t = 2^t; x_{t+1} = x_t + 2x_{t-1}$.

c) $x_t = \frac{t(t+1)}{2}; x_{t+1} = x_t + t + 1$.

▷ **6.4** Resolver las siguientes ecuaciones y analizar la convergencia del sistema. Realizar la gráfica de la función x_t en todos los casos.

a) $x_{t+1} = -0.5x_t + 3$

b) $2x_{t+1} = 3x_t + 4$

c) $x_{t+1} = -x_t + 5$

d) $x_{t+1} - x_t = -\frac{4}{3}x_t$

▷ **6.5** El ingreso, Y_t , evoluciona de acuerdo a la siguiente ecuación:

$$Y_{t+1} = C_t + I_t,$$

en donde I_t es inversión y C_t es el consumo. Si $C_t = mY_t + c$ con $m < 1$ y $c > 0$ y la inversión es constante de manera que $I_t = I$, obtener una ecuación en diferencias para el ingreso y resolver. Analizar la convergencia del modelo.

▷ **6.6** Sea P_t el valor de una fábrica al tiempo t . Sea Y_t el valor de la ganancia de la fábrica al tiempo t . Suponer que el valor inicial de la fábrica se estima como el valor presente de todas las ganancias a lo largo de la vida de ésta. Es decir, se supone

$$P_0 = \sum_{k=0}^{\infty} \beta^k Y_k, \quad (6.8)$$

donde β es una tasa de descuento $0 < \beta < 1$. Deseamos encontrar una expresión para P_t .

a) Justificar la siguiente ecuación $P_t = \beta P_{t+1} + Y_t$. Resolveremos, por lo tanto

$$P_{t+1} = \beta^{-1} (P_t - Y_t) \quad (6.9)$$

con condición inicial dada por (6.8).

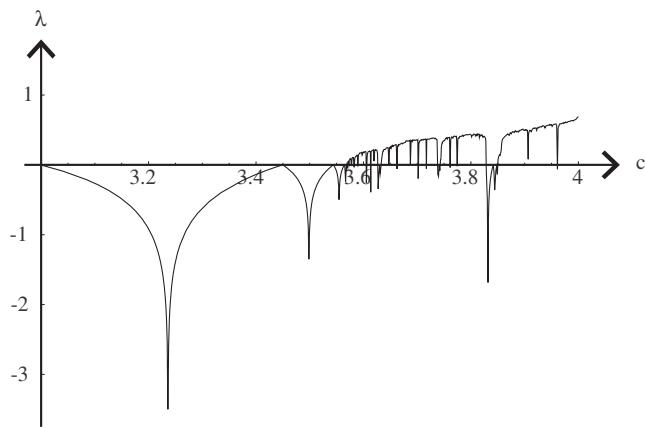


Figura 6.21: Exponentes de Liapunov del modelo logístico para valores $3 \leq c \leq 4$.

- b) Sea $Z_t = \beta^t P_t$. Demostrar que $Z_{t+1} - Z_t = -\beta^t Y_t$.
- c) Demostrar que $Z_t = Z_0 - \sum_{k=0}^{t-1} \beta^k Y_k$.
- d) Concluir $P_t = \sum_{k=0}^{\infty} \beta^k Y_{k+t}$.
- ▷ 6.7 Demostrar la afirmación del ejemplo 6.3.3.
- ▷ 6.8 Realizar los diagramas de fase y determinar los puntos fijos y tipos de equilibrio de las siguientes.

 - a) $x_{t+1} = x_t^3 - x_t^2 + 1$
 - b) $x_{t+1} = \sqrt{4x_t - 3}$

- ▷ 6.9 Verificar que si $c \in [0, 4]$ entonces el sistema dinámico logístico $x_{t+1} = f(x_t) = cx_t(1 - x_t)$ se puede restringir al intervalo $[0, 1]$. Es decir, demostrar que $f[0, 1] \subset [0, 1]$.
- ▷ 6.10 Para los siguientes sistemas de oferta y demanda encontrar el precio, p_t , de equilibrio y determinar el tipo de equilibrio que se tiene.

 - a) $D_t = -3p_t + 10, S_t = p_{t-1} + 2$.
 - b) $D_t = -4p_t + 25, S_t = 4p_{t-1} + 3$.
 - c) $D_t = -\frac{5}{2}p_t + 45, S_t = \frac{15}{2}p_{t-1} + 5$.

- ▷ 6.11 Dado que encontrar la solución de una ecuación en diferencia es un proceso iterativo se puede utilizar una hoja de cálculo para obtener la sucesión $\{x_t\}$. Dada cualquier ecuación $x_{t+1} = f(x_t)$, simplemente comenzamos con un valor inicial $x(0)$ y de ahí se pueden obtener los términos subsecuentes. Para los ejemplos de oferta y demanda del ejercicio anterior, utilizar una hoja de cálculo para graficar la sucesión de precios $\{p_t\}$.
- ▷ 6.12 Verificar que si $\{x_0, x_1, \dots, x_{n-1}\}$ es una órbita asintóticamente estable de un sistema dinámico discreto y diferenciable, entonces su exponente de Liapunov es negativo.
- ▷ 6.13 Considerar la siguiente versión discreta del modelo logístico de crecimiento

$$x_{t+1} = 3.2x_t(1 - 0.25x_t).$$

Calcular las primeras cien iteraciones de la órbita de x_0 . Esto es, obtener numéricamente

$\{x_0, x_1, x_2, \dots, x_{100}\}$ y graficar tomando las siguientes condiciones iniciales.

a) $x_0 = 2.75$,

b) $x_0 = 2.5$,

c) $x_0 = 3$.

d) ¿Qué tipo de equilibrio tiene este sistema?
¿Existen 2-ciclos? Justificar

▷ **6.14** Considerar el sistema $x_{t+1} = 1 - x_t^2$.

a) Encontrar los puntos fijos y las órbitas de periodo 2.

b) Analizar la estabilidad de los puntos encontrados en el inciso anterior.

▷ **6.15** Este ejercicio se refiere a cuestiones surgidas en el ejemplo 6.5.3.

a) Verificar que si $0 \leq \lambda \leq 3$ entonces el sistema dinámico dado por (6.7) puede restringirse al intervalo $[-1, 1]$.

b) Verificar que si $\lambda > 1$ entonces el sistema dinámico tiene dos puntos fijos. ¿Para qué valores de λ son éstos asintóticamente estables? ¿Para qué valores de λ son inestables?

c) Verificar que si $\lambda > 2$ entonces el sistema dinámico tiene órbitas de periodo 2 y analizar su estabilidad.

d) Aproximar el diagrama de bifurcación del sistema dinámico para $0 \leq \lambda \leq 3$.

▷ **6.16** Obtener diagramas de bifurcación para los siguientes sistemas:

a) $x_{t+1} = x_t^2 + k$.

b) $x_{t+1} = T(x_t)$ donde

$$T(x_t) = \begin{cases} rx_t, & 0 \leq x \leq \frac{1}{2} \\ r(1 - x_t), & \frac{1}{2} \leq x \leq 1 \end{cases}$$

con $0 \leq r \leq 2$ y $0 \leq x_t \leq 1$. A la función $T(x)$ se le conoce como la función “tienda de campaña”.

▷ **6.17** Demostrar que si $k \geq 1/4$. Entonces el sistema dinámico $x_{t+1} = x_t^2 + k$ puede ser convertido en $x_{t+1} = cx_t(1 - x_t)$ con un cambio de variable continuo.

▷ **6.18** Demostrar formalmente la parte a) del teorema 6.3.3 siguiendo los siguientes pasos.

a) Sin pérdida de generalidad suponer que $x^* = 0$. Por lo tanto

$$0 \leq |f'(0)| = |f'(x^*)| < 1.$$

b) Sea λ un número positivo tal que

$$|f'(0)| < \lambda < 1.$$

Demostrar que existe un intervalo $(-\delta, \delta)$ tal que si $x \in (-\delta, \delta)$ entonces $|f'(x)| < \lambda < 1$.

c) Usando el teorema del valor medio demostrar que si $x_0 \in (-\delta, \delta)$ entonces $|f(x_0)| < \lambda |x_0| < \delta$, y por lo tanto $f(x_0) \in (-\delta, \delta)$.

d) Demostrar, por inducción, que si $x_0 \in (-\delta, \delta)$ entonces $|f^t(x_0)| < \lambda^t |x_0|$, para todo $t \geq 0$.

e) Concluir que $\lim_{t \rightarrow \infty} f^t(x_0) = 0 = x^*$.

▷ **6.19** Considérese un intervalo $I = [\alpha, \beta]$ y una función continua $f : I \rightarrow I$ que es suprayectiva. Suponer que existe un punto γ tal que $f(\gamma) = \beta$. Suponer además, que $f(\alpha) = \alpha = f(\beta)$.

a) Hacer un diagrama que represente esta situación.

- b) Demostrar que existe δ tal que $f(\delta) = \gamma$ y
 $\alpha < \delta < \gamma$. Yorke.
- c) Demostrar que $f^3(\delta) < \delta < f(\delta) < f^2(\delta)$ por lo que f cumple con las hipótesis del teorema 6.4.3 y por lo tanto existe caos según d) Verificar que el modelo logístico cumple con las condiciones anteriores si $c = 4$. Verificar que la tienda de campaña también cumple con las condiciones.

CAPÍTULO 7

Sistemas de ecuaciones en diferencias lineales

§7.1 Introducción

Supongamos que ahora X_t toma valores en \mathbb{R}^n . Un sistema dinámico discreto lineal está dado por

$$X_{t+1} = AX_t, \quad (7.1)$$

en donde A es una matriz de $n \times n$. La solución de (7.1) se obtiene de manera semejante al caso continuo como veremos a continuación.

Proposición 7.1.1 *Si \mathbf{v} es un vector propio de la matriz A con valor propio λ , entonces $X_t = \lambda^t \mathbf{v}$ es una solución de (7.1).*

Demostración

Si $X_t = \lambda^t \mathbf{v}$, entonces se tiene

$$\begin{aligned} AX_t &= A(\lambda^t \mathbf{v}) = \lambda^t A\mathbf{v} \\ &= \lambda^t \lambda \mathbf{v} = \lambda^{t+1} \mathbf{v} \\ &= X_{t+1}, \end{aligned}$$

con lo cual se demuestra que efectivamente $X_t = \lambda^t \mathbf{v}$ es solución de (7.1). ■

Al igual que en el caso continuo, en el discreto se cumple el principio de superposición y la solución general de (7.1) es una combinación lineal de n soluciones linealmente independientes. Podemos analizar los distintos casos dependiendo de las raíces del polinomio característico $p_A(\lambda)$.

§7.2 La solución general

Omitiremos todos los detalles ya que los procedimientos son prácticamente idénticos al caso de sistemas lineales continuos estudiados en el capítulo 4. El caso más simple es cuando se tienen n raíces reales distintas ya que tenemos directamente una base de vectores propios reales. Sean $\lambda_1, \lambda_2, \dots, \lambda_n$ valores propios reales distintos de la matriz A , con vectores propios dados, respectivamente, por $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$. Entonces la solución general se puede expresar como

$$X_t = K_1 \mathbf{v}_1 \lambda_1^t + K_2 \mathbf{v}_2 \lambda_2^t + \dots + K_n \mathbf{v}_n \lambda_n^t. \quad (7.2)$$

Supongamos ahora que se tiene una pareja de valores propios complejos dados por las raíces del polinomio:

$$\begin{aligned}\lambda &= \|\lambda\| (\cos \theta + i \sin \theta), \\ \bar{\lambda} &= \|\lambda\| (\cos \theta - i \sin \theta)\end{aligned}$$

con vectores propios asociados,

$$\begin{aligned}\mathbf{v} &= \mathbf{r} + \mathbf{s}i, \\ \bar{\mathbf{v}} &= \mathbf{r} - \mathbf{s}i.\end{aligned}$$

Sabemos que¹

$$\begin{aligned}\lambda^t \mathbf{v} &= \|\lambda\|^t (\cos \theta t + i \sin \theta t)(\mathbf{r} + \mathbf{s}i) \\ &= \|\lambda\|^t (\mathbf{r} \cos \theta t - \mathbf{s} \sin \theta t) + i \|\lambda\|^t (\mathbf{s} \cos \theta t + \mathbf{r} \sin \theta t)\end{aligned}$$

es solución y es fácil verificar que la parte real y la parte imaginaria son dos soluciones independientes, por lo que la solución asociada a la pareja de raíces complejas es

$$X_t = \|\lambda\|^t [(K_1 \cos \theta t + K_2 \sin \theta t)\mathbf{r} + (K_2 \cos \theta t - K_1 \sin \theta t)\mathbf{s}], \quad (7.3)$$

que es simplemente el análogo del teorema 4.2.3.

Finalmente, analizaremos el caso en que se tiene una raíz real doble (el caso general es similar al del apéndice C). Se tiene entonces que λ es una raíz real repetida con vector propio \mathbf{v} y en este caso $\lambda^t \mathbf{v}$ y $\lambda^t t \mathbf{v}$ son dos soluciones independientes de manera que la solución general asociada a esta raíz doble está dada por

$$X_t = K_1 \lambda^t \mathbf{v} + K_2 \lambda^t t \mathbf{v}. \quad (7.4)$$

§7.3 Caso no homogéneo

Dado el sistema

$$X_{t+1} = AX_t + B, \quad (7.5)$$

¹Véase el ejercicio A.7 del apéndice A.

con $B \in \mathbb{R}^n$, la solución general es simplemente la suma de la solución al sistema homogéneo asociado X_h , dada por alguno de los casos de la sección anterior, más una solución particular, X^p . Ésta se obtiene considerando $X_t = X^p$, una función constante, de manera que $X_t = X_{t+1} = X^p$ es la solución de

$$(I - A)X^p = B.$$

Nótese que la solución puede no existir ya que $I - A$ puede ser una matriz singular.

§7.4 Análisis cualitativo

Consideremos un sistema de 2×2 como (7.5). El punto fijo o equilibrio se obtiene al resolver $X^* = AX^* + B$ y coincide con la solución particular del sistema. Al igual que en el caso continuo nos interesa saber cuándo existe convergencia hacia el punto fijo.

Cuando ambas raíces son reales y menores que 1 en valor absoluto, es claro que $\lim_{t \rightarrow \infty} X_t = X^*$, independientemente de las condiciones iniciales. Si ambas raíces son mayores que 1 en valor absoluto y $X_0 \neq X^*$, entonces el sistema se aleja de X^* ; finalmente, si $|\lambda_1| > 1$ y $|\lambda_2| < 1$ se tiene que existe una sucesión (estable) que converge hacia X^* si las condiciones iniciales son tales que $K_1 = 0$ y $K_2 \neq 0$. Análogamente, la sucesión que diverge más rápidamente de X^* se obtiene si las condiciones iniciales son tales que $K_1 \neq 0$ y $K_2 = 0$. Los equilibrios descritos se conocen, al igual que en el caso continuo, como nodos atractores, nodos repulsores y puntos silla.

Si la raíz real es doble, la convergencia se obtiene cuando ésta es menor que 1 en valor absoluto y el punto de equilibrio es un nodo atractor. Si la raíz es mayor que 1 en valor absoluto, el sistema diverge y el punto fijo es un nodo repulsor.

El último caso es cuando se tiene una pareja de raíces complejas. La solución dada por (7.3) nos indica que el sistema converge hacia su punto fijo siempre y cuando $\|\lambda\| < 1$, es decir, la raíz compleja tiene norma compleja menor que uno. De otra forma, la solución diverge. Debido a la periodicidad de las funciones trigonométricas tanto la convergencia como la divergencia son periódicas.

En analogía con el caso continuo, los casos degenerados se presentan cuando

- a) $|\lambda| = 1$ o $\|\lambda\| = 1$,
- b) las raíces son repetidas (se cumple $tr^2(A) = 4 \det(A)$),
- c) los valores propios imaginarios tienen parte real igual a cero (se cumple $tr(A) = 0$).

Nótese que el primero de estos casos equivale a $\det(A) = 0$ en el caso continuo, o bien a tener un valor propio igual a 0. Aquí lo equivalente es tener valores propios cuyo valor absoluto, o norma en el caso complejo, es igual a 1. Se puede realizar una figura análoga a la figura 5.6 en la cual los casos degenerados dividen al plano definido por $tr(A)$ y $\det(A)$ en distintas regiones de acuerdo con el comportamiento dinámico del sistema. Para este propósito tenemos que determinar cuándo se tiene $|\lambda| = 1$ o $\|\lambda\| = 1$, lo cual se realiza a continuación.

Si el discriminante $tr^2(A) - 4 \det(A) \geq 0$, las raíces λ_1 y λ_2 del polinomio característico son reales, nos interesa determinar cuándo se cumplen los siguientes casos:

- a) $|\lambda_1|, |\lambda_2| > 1$,
- b) $|\lambda_1|, |\lambda_2| < 1$,
- c) $|\lambda_1| > 1$ y $|\lambda_2| < 1$.

Si $p_A(\lambda)$ es el polinomio característico, entonces podemos factorizarlo como

$$p_A(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2).$$

En particular, si $\lambda = 1$ se tiene que $p_A(1) = (1 - \lambda_1)(1 - \lambda_2)$. De aquí que $p_A(1) > 0$ si y sólo si $\lambda_1, \lambda_2 > 1$ o $\lambda_1, \lambda_2 < 1$; por otra parte, $p_A(1) < 0$ si y sólo si $\lambda_1 > 1, \lambda_2 < 1$ o $\lambda_1 < 1, \lambda_2 > 1$. De forma análoga se tiene que $p_A(-1) > 0$ si y sólo si $\lambda_1, \lambda_2 > -1$ o $\lambda_1, \lambda_2 < -1$ o bien, $p_A(-1) < 0$ si y sólo si $\lambda_1 > -1, \lambda_2 < -1$ o $\lambda_1 < -1, \lambda_2 > -1$. Dado que $p_A(1) = 1 - \text{tr}(A) + \det(A)$ y $p_A(-1) = 1 + \text{tr}(A) + \det(A)$, se pueden representar estas regiones en el plano definido por la traza y el determinante como las comprendidas entre las rectas $p_A(1) = 0$ y $p_A(-1) = 0$.

Si el discriminante es negativo y se tienen raíces complejas, entonces $\text{Re}(\lambda) = \frac{1}{2}\text{tr}A$, $\text{Im}(\lambda) = \frac{1}{2}\sqrt{4\det(A) - \text{tr}^2(A)}$ y por lo tanto $\|\lambda\| = \det(A)$ de manera que se tiene

$$\|\lambda\| < 1 \Leftrightarrow \det(A) < 1.$$

De esta forma, las distintas regiones quedan como en la figura 7.1.

La numeración de las regiones corresponde a la del caso continuo visto en la sección 5.4. Así, la región I corresponde a un nodo inestable, la II a una espiral inestable, la III a una espiral estable, la IV a un nodo estable y la V a un punto silla. Asimismo, los casos degenerados corresponden primero, a los puntos con discriminante positivo, sobre las rectas $p_A(1) = 0$ y $p_A(-1) = 0$, que satisfacen $|\text{tr}(A)| = \det(A) + 1$, en donde $|\lambda| = 1$ y el conjunto de puntos fijos es una recta; y, segundo, al conjunto de puntos con discriminante negativo sobre la recta $\det(A) = 1$, en donde $\|\lambda\| = 1$ y los puntos fijos son centros.

En algunos libros se presentan diagramas de fase de los sistemas discretos, tal como se hace con los sistemas continuos. Esto no es del todo correcto ya que en realidad aquí no se tienen trayectorias sino sucesiones de puntos, por eso, no tiene sentido dibujar trayectorias continuas. Asimismo, linealizar sistemas discretos no es enteramente correcto ya que, como se vio en la sección 6.5, los sistemas discretos no lineales pueden presentar comportamientos sumamente complejos. En este caso, el proceso de linealización cambiaría totalmente la dinámica del sistema con lo cual las conclusiones que extrajéramos serían del todo inútiles.

Ejemplos

Ej 7.4.1 Resolvamos el sistema $X_{t+1} = \begin{pmatrix} 4 & -1 \\ 2 & 1 \end{pmatrix} X_t$. La ecuación característica de la matriz está dada por

$$\lambda^2 - 5\lambda + 6 = 0,$$

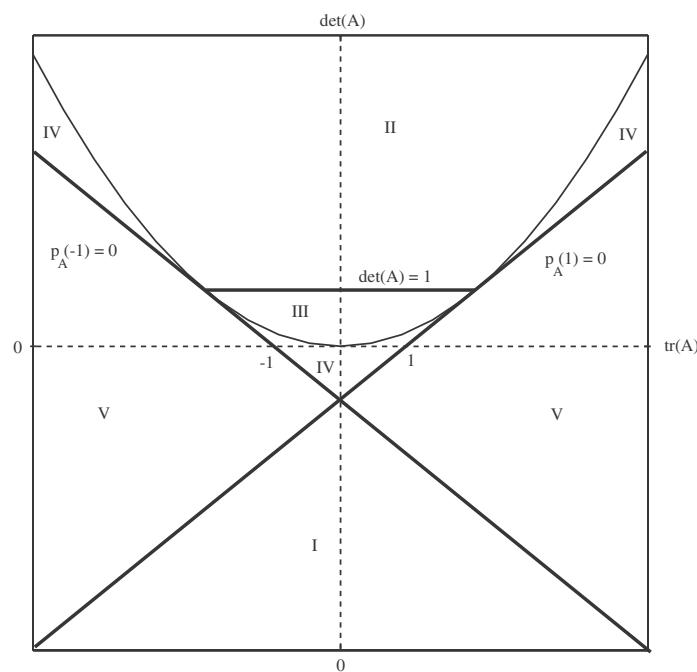


Figura 7.1: Regiones del plano definido por la traza y el determinante de una matriz A .

con raíces $\lambda_1 = 3$ y $\lambda_2 = 2$. Estos valores propios tienen, respectivamente, asociados los vectores propios $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ y $\mathbf{v}_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$. La solución queda dada de acuerdo a (7.2) como,

$$\begin{aligned} X_t &= \begin{pmatrix} x_t \\ y_t \end{pmatrix} = K_1 3^t \begin{pmatrix} 1 \\ 1 \end{pmatrix} + K_2 2^t \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \\ &\quad \begin{pmatrix} K_1 3^t + K_2 2^t \\ K_1 3^t + 2K_2 2^t \end{pmatrix}. \end{aligned}$$

Si se tiene la condición inicial $X_0 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, entonces las constantes quedan determinadas por el sistema

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} K_1 + K_2 \\ K_1 + 2K_2 \end{pmatrix},$$

con lo que finalmente $K_1 = -1$, $K_2 = 1$. El origen es un punto fijo que en este caso es un repulsor puesto que en valor absoluto los valores propios son mayores que 1.

Ej 7.4.2 Se quiere resolver el sistema, $X_{t+1} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} X_t$. La ecuación característica de la matriz está dada por

$$\lambda^2 - 2\lambda + 2 = 0,$$

con raíces complejas $\lambda = 1 + i$ y $\bar{\lambda} = 1 - i$. Estos valores propios tienen asociados, respectivamente, los vectores propios $\mathbf{v} = \begin{pmatrix} 1 \\ -i \end{pmatrix}$ y $\bar{\mathbf{v}} = \begin{pmatrix} 1 \\ i \end{pmatrix}$. Reescribimos a λ en su forma polar como,

$$\begin{aligned} \lambda &= \sqrt{2} \left(\frac{1}{\sqrt{2}} + i \frac{1}{\sqrt{2}} \right) \\ &= \sqrt{2} \left(\cos \frac{\pi}{4} + i \sin \frac{\pi}{4} \right) \end{aligned}$$

y al vector \mathbf{v} como $\mathbf{v} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} + i \begin{pmatrix} 0 \\ -1 \end{pmatrix}$, la solución queda, de acuerdo con (7.3), como

$$\begin{aligned} X_t &= (\sqrt{2})^t \left[(K_1 \cos \frac{\pi}{4} t + K_2 \sin \frac{\pi}{4} t) \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \right. \\ &\quad \left. (K_2 \cos \frac{\pi}{4} t - K_1 \sin \frac{\pi}{4} t) \begin{pmatrix} 0 \\ -1 \end{pmatrix} \right] \\ &\Rightarrow \begin{pmatrix} x_t \\ y_t \end{pmatrix} = (\sqrt{2})^t \begin{pmatrix} K_1 \cos \frac{\pi}{4} t + K_2 \sin \frac{\pi}{4} t \\ -K_2 \cos \frac{\pi}{4} t + K_1 \sin \frac{\pi}{4} t \end{pmatrix}. \end{aligned}$$

Si se tiene la condición inicial $X_0 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, entonces las constantes se obtienen resolviendo

$$\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} K_1 \\ -K_2 \end{pmatrix}$$

es decir, $K_1 = 1$ y $K_2 = -2$. El origen es una espiral inestable, dado que $\|\lambda\| = \sqrt{2} > 1$.

Ej 7.4.3 Ahora resolvamos el sistema $X_{t+1} = \begin{pmatrix} 3 & -1 \\ 1 & 1 \end{pmatrix} X_t + \begin{pmatrix} -4 \\ 1 \end{pmatrix}$. La ecuación característica de la matriz del sistema está dada por

$$\lambda^2 - 4\lambda + 4 = 0,$$

que tiene una raíz real repetida $\lambda = 2$ con vector propio asociado $\mathbf{v} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. En este caso, (7.4) nos da la solución a la ecuación homogénea asociada:

$$\begin{aligned} X_t^h &= \begin{pmatrix} x_t^h \\ y_t^h \end{pmatrix} = K_1 2^t \begin{pmatrix} 1 \\ 1 \end{pmatrix} + K_2 2^t t \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} K_1 2^t + K_2 t 2^t \\ K_1 2^t + K_2 t 2^t \end{pmatrix}. \end{aligned}$$

Una solución particular se obtiene al resolver

$$\begin{pmatrix} -2 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} x^* \\ y^* \end{pmatrix} = \begin{pmatrix} -4 \\ 1 \end{pmatrix},$$

con lo que se tiene que $x^* = -1$ y $y^* = -6$. Finalmente, la solución general del sistema queda dada por

$$\begin{pmatrix} x_t \\ y_t \end{pmatrix} = \begin{pmatrix} K_1 2^t + K_2 t 2^t - 1 \\ K_1 2^t + K_2 t 2^t - 6 \end{pmatrix}.$$

El punto fijo $(-1, -6)$ es un nodo inestable ya que $|\lambda| = 2 > 1$.



§7.5 Ecuaciones lineales de segundo orden

En esta sección se estudiará únicamente el caso de ecuaciones de segundo orden con coeficientes constantes, tomando en cuenta que las ecuaciones de orden superior se resuelven de forma análoga.

Consideremos la siguiente ecuación:

$$ax_{t+2} + bx_{t+1} + cx_t = 0 \quad (7.6)$$

con valores iniciales x_0, x_1 dados. Definimos $y_t = x_{t+1}$ y reescribimos (7.6) como el sistema,

$$\begin{aligned} x_{t+1} &= y_t, \\ y_{t+1} &= -\frac{c}{a}x_t - \frac{b}{a}y_t, \end{aligned}$$

cuya matriz asociada es $\begin{pmatrix} 0 & 1 \\ -\frac{c}{a} & -\frac{b}{a} \end{pmatrix}$, con ecuación característica $a\lambda^2 + b\lambda + c = 0$ y vectores propios

de la forma $\begin{pmatrix} 1 \\ \lambda \end{pmatrix}$. De esta manera, al igual que en el caso continuo, existe una equivalencia entre las

ecuaciones de segundo orden con coeficientes constantes y los sistemas lineales con dos incógnitas. En este caso es fácil ver que las soluciones están dadas por

$$x_t = K_1 \lambda_1^t + K_2 \lambda_2^t,$$

si λ_1 y λ_2 son reales y distintos;

$$x_t = K_1 \lambda^t + K_2 t \lambda^t,$$

si $\lambda_1 = \lambda_2 = \lambda$ es una raíz real repetida, y finalmente

$$x_t = \|\lambda\|^t [(K_1 \cos \theta t + K_2 \sin \theta t)],$$

si $\lambda = \|\lambda\| (\cos \theta + i \sin \theta)$ y $\bar{\lambda} = \|\lambda\| (\cos \theta - i \sin \theta)$ son raíces complejas.

§7.5.1 Caso no homogéneo

Considérese el caso no homogéneo con un término constante dado por

$$ax_{t+2} + bx_{t+1} + cx_t = d.$$

Éste se resuelve, como de costumbre, sumando una solución particular a la solución general de la ecuación homogénea asociada. Si $a + b + c \neq 0$, entonces una solución particular es la constante

$$x^p = \frac{d}{a + b + c}. \quad (7.7)$$

En el caso en que $a + b + c = 0$ pero $2a + b \neq 0$, se puede ver que una solución particular está dada por

$$x^p = \frac{d}{2a + b} t. \quad (7.8)$$

Si $a + b + c = 0$ y $2a + b = 0$, entonces $a \neq 0$ ya que de otra forma todos los coeficientes serían nulos. Se puede entonces ver que una solución particular es

$$x^p = \frac{d}{2a} t^2. \quad (7.9)$$

Ejemplo

Ej 7.5.1 Consideremos la ecuación $x_{t+2} - 3x_{t+1} + 2x_t = -1$ con $x_0 = 0$ y $x_1 = 1$. La ecuación característica asociada es

$$\lambda^2 - 3\lambda + 2 = 0,$$

con raíces reales distintas $\lambda_1 = 2$ y $\lambda_2 = 1$. Se obtiene así la solución de la ecuación homogénea asociada:

$$x_t = K_1 2^t + K_2.$$

Sustituyendo los valores iniciales, obtenemos que $K_1 = 1$ y $K_2 = -1$. Finalmente, una solución particular se obtiene al poner $x^p = t$, de acuerdo con (7.8). Obsérvese que no podemos usar (7.7) pues $a = 1, b = -3$ y $c = 2$ implican $a + b + c = 0$. De esta forma, la solución general está dada por

$$x_t = 2^t - 1 + t.$$

◇◇◇◇◇◇◇◇◇

§7.5.1.1 Coeficientes indeterminados

Al igual que en el caso continuo, cuando se tiene una ecuación no-homogénea del tipo

$$ax_{t+2} + bx_{t+1} + cx_t = f(t)$$

puede utilizarse el método de coeficientes indeterminados, visto en la sección 2.3.4.1, proponiendo una solución particular que tenga la misma forma que la función $f(t)$.

Ejemplos

Ej 7.5.2 Consideremos la ecuación $x_{t+2} - 3x_{t+1} + 2x_t = 3^t$. La solución de la ecuación homogénea asociada se obtuvo en el ejemplo anterior. Para obtener una solución particular, proponemos $x_t^p = A3^t$, una solución que tiene la forma general de la función $f(t) = 3^t$. Encontramos que $x_{t+1}^p = A3^{t+1}$ y $x_{t+2}^p = A3^{t+2}$, y sustituyendo en la ecuación original se tiene

$$\begin{aligned} A3^{t+2} - 3A3^{t+1} + 2A3^t &= 3^t, \\ 3^t(9A - 9A + 2A) &= 3^t, \\ 3^t(2A) &= 3^t, \end{aligned}$$

por lo que $A = \frac{1}{2}$. La solución particular es entonces $x_t^p = \frac{1}{2}3^t$ y por lo tanto la solución general a la ecuación es

$$x_t = K_1 2^t + K_2 + \frac{1}{2}3^t.$$

Ej 7.5.3 Considérese el sistema

$$\begin{pmatrix} x_{t+1} \\ y_{t+1} \end{pmatrix} = \begin{pmatrix} 1 & 3 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x_t \\ y_t \end{pmatrix} + \begin{pmatrix} e^t \\ e^t \end{pmatrix}.$$

Podemos convertir este sistema en una ecuación de segundo orden como sigue: iteramos un periodo al futuro la primera ecuación, con lo que obtenemos

$$x_{t+2} = x_{t+1} + 3y_{t+1} + e^{t+1};$$

sustituimos y_{t+1} de la segunda ecuación:

$$x_{t+2} = x_{t+1} + 3(x_t - y_t + e^t) + e^{t+1},$$

y finalmente sustituimos y_t de la primera ecuación y simplificamos para obtener

$$x_{t+2} - 4x_t = (4 + e)e^t.$$

Se puede entonces proceder a resolver esta ecuación.

El método de coeficientes indeterminados funciona de la misma forma para sistemas de ecuaciones, así que lo utilizaremos directamente. Los valores propios del sistema original son $\lambda_1 = 2$ y $\lambda_2 = -2$,

con vectores propios asociados $\mathbf{v}_1 = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$ y $\mathbf{v}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$. La solución de la ecuación homogénea asociada queda entonces expresada como

$$\begin{pmatrix} x_t \\ y_t \end{pmatrix} = K_1 2^t \begin{pmatrix} 3 \\ 1 \end{pmatrix} + K_2 (-2)^t \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Proponemos una solución particular de la forma

$$\begin{pmatrix} x_t^p \\ y_t^p \end{pmatrix} = \begin{pmatrix} A \\ B \end{pmatrix} e^t.$$

Sustituyendo en la ecuación original, tenemos

$$\begin{pmatrix} Ae^{t+1} \\ Be^{t+1} \end{pmatrix} = \begin{pmatrix} 1 & 3 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} Ae^t \\ Be^t \end{pmatrix} + \begin{pmatrix} e^t \\ e^t \end{pmatrix},$$

de manera que

$$\begin{pmatrix} Ae \\ Be \end{pmatrix} e^t = \begin{pmatrix} A + 3B + 1 \\ A - B + 1 \end{pmatrix} e^t.$$

Igualando coeficientes se tiene el siguiente sistema para A y B :

$$Ae = A + 3B + 1,$$

$$Be = A - B + 1,$$

por lo que $A = \frac{e+4}{e^2-4}$ y $B = \frac{e}{e^2-4}$. Así, la solución final es

$$\begin{pmatrix} x_t \\ y_t \end{pmatrix} = K_1 2^t \begin{pmatrix} 3 \\ 1 \end{pmatrix} + K_2 (-2)^t \begin{pmatrix} 1 \\ -1 \end{pmatrix} + \begin{pmatrix} \frac{e+4}{e^2-4} \\ \frac{e}{e^2-4} \end{pmatrix} e^t.$$

◇◇◇◇◇◇◇◇◇

Ejercicios

▷ **7.1** Resolver los siguientes sistemas y determinar el tipo de equilibrio:

a) $X_{t+1} = \begin{pmatrix} 6 & 9 \\ -1 & 0 \end{pmatrix} X_t.$

b) $X_{t+1} = \begin{pmatrix} -1 & -2 \\ -2 & 2 \end{pmatrix} X_t.$

c) $X_{t+1} = \begin{pmatrix} 3 & -2 \\ 1 & 0 \end{pmatrix} X_t$ con la condición inicial $X_0 = (1, 0).$

d) $X_{t+1} = \begin{pmatrix} -2 & -1 \\ 1 & -1 \end{pmatrix} X_t + \begin{pmatrix} 9 \\ 3 \end{pmatrix}.$

▷ **7.2** Resolver el siguiente sistema y determinar el tipo de equilibrio:

$$\begin{aligned} x_{t+1} - x_t - \frac{1}{3}y_t &= -1, \\ x_{t+1} + y_{t+1} - \frac{1}{6}y_t &= \frac{17}{2}. \end{aligned}$$

Usar las condiciones iniciales $x(0) = 5$ y $y(0) = 4.$

▷ **7.3 (Sucesión de Fibonacci)** La sucesión de números de Fibonacci² está dada por la fórmula recursiva

$$x_{t+1} = x_t + x_{t-1}.$$

Encontrar la solución con las condiciones iniciales $x_0 = 0$ y $x_1 = 1.$ Calcular $\gamma = \lim_{t \rightarrow \infty} \frac{x_{t+1}}{x_t}$, al número γ se lo llama radio dorado o razón áurea y constituía el fundamento de la belleza entre los griegos.

▷ **7.4** Suponer que x_t representa el número de personas empleadas en el periodo t y y_t el número de desempleados en el mismo periodo. La fuerza laboral $L = x_t + y_t$ se mantiene constante. Sea p la probabilidad de que una persona desempleada encuentre trabajo en cualquier periodo dado y q la

probabilidad de que una persona empleada continúe empleada. El siguiente modelo describe la dinámica del empleo:

$$\begin{aligned} x_{t+1} &= qx_t + py_t, \\ y_{t+1} &= (1 - q)x_t + (1 - p)y_t. \end{aligned}$$

Resolver en términos de p y q e interpretar los resultados.

▷ **7.5 (Modelo de Crecimiento de Leslie)**

Considerar un organismo que vive durante dos periodos. Sean n_1 y n_2 las tasas de natalidad en cada periodo, y sea m_1 la tasa de mortalidad de los individuos cuya edad es de un periodo, es decir, $(1 - m_1)$ individuos nacidos en el primer periodo llegan al segundo. Sean x_t y y_t las poblaciones de individuos de primer periodo y segundo periodo, respectivamente, que coexisten en el periodo $t.$ La dinámica de estas poblaciones está descrita por el sistema

$$\begin{aligned} x_{t+1} &= n_1 x_t + n_2 y_t, \\ y_{t+1} &= (1 - m_1)x_t. \end{aligned}$$

Resolver el modelo en general y para el caso particular $n_1 = 1, n_2 = 4$ y $m_1 = 0.5.$

▷ **7.6** Supongamos que los ingresos nacionales de dos países que son socios comerciales están dados por Y_t^1 y Y_t^2 , respectivamente; asimismo el consumo, las importaciones y las exportaciones están dados por $C_t^1, C_t^2, M_t^1, M_t^2, X_t^1$ y $X_t^2.$ La inversión no depende del tiempo y está dada simplemente por I^1 y $I^2.$ El consumo de productos nacionales y las importaciones son proporcionales al ingreso del periodo anterior de manera que para $i, j = 1, 2,$

$$\begin{aligned} C_t^i - M_t^i &= a_{ii} Y_{t-1}^i \\ M_t^i &= a_{ji} Y_{t-1}^j, \quad i \neq j, \end{aligned} \tag{7.10}$$

²Leonardo Fibonacci fue un matemático italiano del siglo XIII que descubrió esta sucesión al estudiar un modelo de reproducción de conejos.

en donde $a_{ij} > 0 \forall i, j$. La restricción presupuestal de cada país está dada por

$$Y_t^i = C_t^i + X_t^i - M_t^i + I^i \quad (7.11)$$

y se tiene también que en cada periodo las importaciones de un país son iguales a las exportaciones del otro, es decir,

$$\begin{aligned} M_t^1 &= X_t^2 \\ M_t^2 &= X_t^1 \end{aligned} \quad (7.12)$$

- a)** Combinar (7.10), (7.11) y (7.12) para obtener el sistema,

$$\begin{aligned} Y_t^1 &= a_{11}Y_{t-1}^1 + a_{12}Y_{t-1}^2 + I^1, \\ Y_t^2 &= a_{21}Y_{t-1}^1 + a_{22}Y_{t-1}^2 + I^2. \end{aligned} \quad (7.13)$$

- b)** Los coeficientes del sistema, a_{ii} , se pueden interpretar como las propensiones marginales a consumir bienes domésticos y análogamente, a_{ij} , $i \neq j$, como las propensiones marginales a importar. Suponemos que todos estos coeficientes son positivos y satisfacen $a_{11} + a_{22} < 1$ y $a_{11}a_{22} - a_{21}a_{12} > 0$. Encontrar la solución del sistema (7.13) y analizar que pasa cuando $t \rightarrow \infty$.

▷ **7.7** Denotamos por x_t , y_t a la producción de dos industrias en el periodo t . La demanda final de ambos productos en el mismo periodo está dada por d_t^x y d_t^y , respectivamente. Supongamos que la producción de cada industria se utiliza como insumo de la producción en ambas, aparte de satisfacer la demanda final. Si la producción en el periodo $t+1$ se determina de acuerdo a las demandas del periodo anterior, se debe satisfacer entonces un sistema del tipo

$$\begin{aligned} x_{t+1} &= a_{11}x_t + a_{12}y_t + d_t^x, \\ y_{t+1} &= a_{21}x_t + a_{22}y_t + d_t^y, \end{aligned}$$

en donde $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ es una matriz de insumo-producto. Resolver el sistema para los siguientes casos utilizando el método de coeficientes indeterminados dado en el ejemplo 7.5.3 para determinar una solución particular. En ambos casos usar la matriz

$$A = \begin{pmatrix} \frac{3}{10} & \frac{4}{10} \\ \frac{3}{10} & \frac{2}{10} \end{pmatrix}.$$

- a)** $d_t^x = d_t^y = \left(\frac{12}{10}\right)^t$, $x(0) = \frac{187}{39}$ y $y(0) = \frac{72}{13}$
- b)** $d_t^x = e^{\frac{t}{10}}$, $d_t^y = 2e^{\frac{t}{10}}$, $x(0) = \frac{53}{6}$ y $y(0) = \frac{25}{6}$

▷ **7.8** Resolver las siguientes ecuaciones dadas las condiciones iniciales dadas por $x_0 = 1$ y $x_1 = 2$:

- a)** $x_{t+2} - x_t = 0$.
- b)** $2x_{t+2} - 5x_{t+1} + 2x_t = 0$.
- c)** $9x_{t+2} - 6x_{t+1} + x_t = 0$.
- d)** $x_{t+2} + 6x_{t+1} + 25x_t = 0$.
- e)** $x_{t+2} - 3x_{t+1} + 2x_t = 1$.
- f)** $4x_{t+2} - 4x_{t+1} + x_t = 2$.
- g)** $2x_{t+2} - 5x_{t+1} + 2x_t = t^2 + 1$.
- h)** $x_{t+2} - 3x_{t+1} + 2x_t = 2^t$. En este problema no podemos usar una solución particular de la forma $A2^t$ ¿por qué? Se propone utilizar una solución del tipo $At2^t$.

CAPÍTULO 8

La mano invisible

§8.1 Introducción

En este capítulo se hace un análisis crítico de la llamada “mano invisible” de Adam Smith. Recordemos que uno de los primeros conceptos de las ciencias económicas fue la de punto de equilibrio. La genial idea de Adam Smith fue reconocer que un sistema económico puede converger en el largo plazo a ese punto de inmovilidad; su argumento es suficientemente convincente como para considerarlo una de las principales leyes económicas.

En la versión más simple de esta idea, se considera un sistema económico cuyos participantes son los miembros de un conjunto de individuos: la sociedad. Cada uno de ellos tiene una sola motivación: su propio interés. Sin embargo, dice Smith, el intercambio de bienes hace que todos los participantes adquieran mayor bienestar. Existe una “mano invisible” que guía a la sociedad a mejorar su situación a través del intercambio de bienes. Cada agente económico sólo ve por sí mismo pero, dado que el intercambio se da únicamente si cada individuo mejora su situación, a la larga se llega a un **punto de equilibrio** en el cual ya nadie puede mejorar. De forma sorprendente, la “mano invisible” conduce a una multitud de egoístas a la mejor situación colectiva, o sea que el egoísmo de los individuos hace que las naciones progresen. El argumento de Smith tiene como corolario que lo mejor que se puede hacer por un sistema económico es no interferir en él.

A continuación nos proponemos demostrar que, desde un punto de vista puramente matemático, las cosas no son tan sencillas. Matemáticamente se puede probar que, bajo hipótesis razonables,¹ debe existir un equilibrio económico. La dificultad surge al querer introducir la variable temporal y esperar que el equilibrio surja como consecuencia de la dinámica natural del sistema. En este sentido, se esperaría que cualquier dinámica que se introduzca debe poseer, al menos, un punto de equilibrio dinámicamente

¹Estamos hablando de las hipótesis del modelo de Arrow-Debreu. Véase, por ejemplo, la exposición que se hace en [Deb86, Deb73].

estable. Propondremos algunas posibilidades y estudiaremos, en cada caso, las consecuencias correspondientes. Hemos de advertir que existen diversas versiones de cómo hacer esto y que sólo analizaremos algunos casos simples.

Recordemos el modelo microeconómico de Arrow-Debreu. Los individuos o “agentes” poseen diversas capacidades y habilidades. La división del trabajo hace que cada uno de ellos se especialice en producir un bien o un servicio para así poder comerciar con otros y satisfacer sus necesidades. Lo anterior es claramente una enorme simplificación de la realidad, pero es un buen punto de partida para el desarrollo de la teoría.

Lo primero que nos preguntamos es si existe un estado en donde todos los agentes estén plenamente satisfechos con lo que producen, bienes o servicios, y lo que reciben a cambio. Este problema es, esencialmente, el de existencia de un punto de equilibrio. Como ya hemos visto con anterioridad, la existencia de un punto de equilibrio sólo constituye un primer paso. Desde el punto de vista dinámico, el asunto va más allá, o sea que tenemos que estudiar la estabilidad de ese punto de equilibrio.

Existen varias maneras de introducir el aspecto dinámico en el modelo y aquí describiremos sólo una de ellas. La inquietante conclusión a la que se llega es que la dinámica en un sistema económico puede ser tan complicada como se deseé. La moraleja inevitable es que el desarrollo futuro de la economía conducirá a leyes y modelos mucho más complejos que los de la física.

§8.2 El modelo básico

En todo lo que sigue se asume en el lector un conocimiento básico de análisis convexo y optimización estática. El capítulo 10 contiene el material requerido, por lo que sugerimos en este punto, cubrir dicho material si es que el lector lo desconoce o lo ha relegado al olvido.

Sean $n \geq 2$ el número de bienes de una economía y m el número de agentes. Asumamos que inicialmente cada uno de los agentes posee ciertas cantidades de los n bienes; cada agente intuye que estaría mejor con otras cantidades y observa que lo mismo ocurre con los demás. En este punto, surge naturalmente la idea de hacer un “trueque” de bienes. El intercambio a través de instrumentos monetarios puede incluirse dentro de este modelo, si se piensa en el dinero como un bien adicional.

Las cantidades de cada uno de los bienes pueden incluirse en un vector de tamaño n ; definimos, de este modo, el **vector de bienes** como

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}_+^n,$$

en donde $\mathbb{R}_+^n = \{(x_1, \dots, x_n) \in \mathbb{R}^n : x_1 \geq 0, \dots, x_n \geq 0\}$ corresponde a los vectores reales no negativos. A este vector se lo llama comúnmente **canasta de bienes**. El intercambio entre los agentes depende del valor relativo entre los bienes a intercambiar. Estos valores relativos quedan determinados por un

vector de precios $\mathbf{p} = \begin{pmatrix} p_1 \\ \vdots \\ p_n \end{pmatrix} \in \mathbb{R}_+^n$. (La introducción de precios en la situación que nos concierne es

una idea, en cierto modo artificial, que necesariamente es introducida *a posteriori*.)

De ahora en adelante se utilizarán superíndices para indicar el agente en cuestión y subíndices para indicar el bien. Sean w_i^k la dotación inicial del bien i que tiene el agente k y $\mathbf{w}_k = \begin{pmatrix} w_1^k \\ \vdots \\ w_n^k \end{pmatrix} \in \mathbb{R}_+^n$ el **vector de dotación inicial** del agente k . Entonces, $\mathbf{p} \cdot \mathbf{w}^k$ es una medida de la riqueza inicial de dicho agente. Este número es una restricción para el agente, pues no puede intercambiar bienes para obtener una mayor riqueza.

El comportamiento de cada agente, dado un vector de precios \mathbf{p} , es consecuencia de la solución de un problema de optimización. Es decir, dada su riqueza $\mathbf{p} \cdot \mathbf{w}^k$, el individuo desea maximizar su **función de utilidad** $u^k(\mathbf{x})$ (la función de utilidad es un concepto que también se introduce *a posteriori*). La idea detrás de la función de utilidad es que los agentes tienen preferencias sobre las diferentes canastas de bienes. Bajo hipótesis aparentemente naturales acerca de esta relación de preferencia, se puede demostrar que existe una función de utilidad que refleja el siguiente hecho: el agente k prefiere (estrictamente) la canasta de bienes \mathbf{x} sobre la canasta de bienes \mathbf{y} si y sólo si $u^k(\mathbf{x}) > u^k(\mathbf{y})$. Del mismo modo, existen canastas que tienen la misma utilidad: éstas son canastas que le son indiferentes al individuo pues le proporcionan la misma utilidad.

No obstante que aquí se trabajará con la función de utilidad, es común utilizar directamente la relación de preferencia sobre las canastas de bienes en lugar de la función de utilidad que las representa. Si \preceq_k es la relación de preferencia del agente k y \mathbf{x} y \mathbf{y} son dos canastas, entonces $\mathbf{x} \preceq_k \mathbf{y}$ significa que la canasta \mathbf{y} es al menos tan buena como la \mathbf{x} . Se dice que las preferencias son estrictamente convexas si los conjuntos

$$V_{\mathbf{x}} = \{\mathbf{y} \mid \mathbf{x} \preceq_k \mathbf{y}\}$$

son estrictamente convexos en \mathbb{R}^n . Asimismo, la relación de preferencia es estrictamente monótona si

dados $\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ y $\mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$, $\mathbf{x} \neq \mathbf{y}$, con $x_i \geq y_i$, $i = 1, \dots, n$, se tiene que $\mathbf{x} \succ \mathbf{y}$, es decir, \mathbf{x}

es *estrictamente preferido* a \mathbf{y} (mayor cantidad de algún bien siempre es mejor). En este caso, la función de utilidad u_k que representa a las preferencias es estrictamente cuasicóncava y estrictamente creciente.

Dada la función de utilidad y un vector de precios $\mathbf{p} \in \mathbb{R}_{++}^n$, en donde

$$\mathbb{R}_{++}^n = \{(x_1, \dots, x_n) \in \mathbb{R}^n : x_1 > 0, \dots, x_n > 0\}$$

corresponde a los vectores positivos, el problema de optimización del agente k se puede formular de la siguiente manera:

$$\begin{aligned} & \max u^k(\mathbf{x}), \\ & \text{sujeto a } \mathbf{p} \cdot \mathbf{x} \leq \mathbf{p} \cdot \mathbf{w}^k. \end{aligned}$$

La desigualdad $\mathbf{p} \cdot \mathbf{x} \leq \mathbf{p} \cdot \mathbf{w}^k$ se conoce como **restricción presupuestal**, y al conjunto de canastas que satisfacen la restricción se lo llama **conjunto presupuestal**.

Para garantizar la solución a este problema de maximización, se supone que las funciones de utilidad son estrictamente cuasicónicas, diferenciables² y estrictamente crecientes. Dadas estas hipótesis, se encuentra que el problema tiene una solución única dada por $\mathbf{x}^k(\mathbf{p})$. Nótese que, en realidad, tenemos una familia de problemas de optimización, uno por agente, con una familia de soluciones que dependen de el vector de precios \mathbf{p} . A la función $\mathbf{x}^k(\mathbf{p})$ se la llama **función de demanda** dado el nivel de precios \mathbf{p} y es fácil ver que ésta es homogénea de grado cero.³ Se define la función de **exceso de demanda** para el agente k como

$$\xi^k(\mathbf{p}) = \mathbf{x}^k(\mathbf{p}) - \mathbf{w}^k,$$

la cual representa la diferencia entre la dotación inicial del agente y su demanda final. Una de las conclusiones del modelo clásico de Arrow-Debreu es que existe al menos un precio en donde todos los agentes están satisfechos con lo que poseen.

Vale la pena describir con mayor detalle lo que esto significa. Si a un cierto precio se tuviera que $\xi^k(\mathbf{p}) = \mathbf{0}$, entonces el agente está satisfecho con su dotación inicial y no es necesario hacer intercambio alguno. Sin embargo, generalmente esto no ocurre así. Notemos que dada la concavidad de la función de utilidad, se tiene que, de hecho, la restricción presupuestal en el máximo se cumple como igualdad. Esto implica que $\mathbf{p} \cdot \mathbf{x}^k(\mathbf{p}) = \mathbf{p} \cdot \mathbf{w}^k$ y por lo tanto $\mathbf{p} \cdot \xi^k(\mathbf{p}) = 0$. Dado que $\mathbf{p} \in \mathbb{R}_{++}^n$, algunas entradas del vector $\xi^k(\mathbf{p})$ deben ser positivas y otras negativas. Esto quiere decir que, a un precio determinado, necesariamente se tiene más demanda que dotación para unos bienes y menos demanda que dotación para otros.

Sea

$$\xi(\mathbf{p}) = \sum_{k=1}^m \xi^k(\mathbf{p})$$

la suma de todas las funciones de exceso de demanda de los agentes. A $\xi(\mathbf{p})$ se lo conoce como **exceso de demanda agregada** y es una función vectorial, $\xi(\mathbf{p}) = (\xi_1(\mathbf{p}), \dots, \xi_n(\mathbf{p}))$, en donde $\xi_i(\mathbf{p})$, $i = 1, \dots, n$ representa el exceso de demanda en el mercado del bien i . En general, con las hipótesis dadas, se cumple el siguiente resultado

Proposición 8.2.1 *La función de demanda agregada $\xi : \mathbb{R}_{++}^n \rightarrow \mathbb{R}^n$ satisface las siguientes propiedades:*

Definición 8.2.2 a) ξ es una función continua.

b) ξ es homogénea de grado cero.

c) Para todo $\mathbf{p} \in \mathbb{R}_{++}^n$ se cumple $\mathbf{p} \cdot \xi(\mathbf{p}) = 0$. (A esta propiedad se la conoce como **ley de Walras**.)

El resultado principal del modelo es que si una función cumple con las propiedades de la proposición 8.2.1, entonces existe al menos un precio \mathbf{p}^* tal que $\xi(\mathbf{p}^*) = \mathbf{0}$. A este precio se lo conoce como precio de equilibrio o **equilibrio walrasiano** y, en principio, no es único. Deteniéndonos a pensar en esto por

²En realidad la existencia y unicidad de la solución queda garantizada si la función de utilidad es continua y no necesariamente diferenciable.

³De esta forma, si todos los precios aumentan en la misma proporción la cantidad óptima demandada es la misma. La función de demanda también es continua en \mathbb{R}_{++}^n ; la demostración no es trivial (remitimos al lector interesado a [Sun96], por ejemplo). La demanda $x_i(\mathbf{p})$ puede ser infinita si $p_i = 0$ para algún $i = 1, \dots, n$, de aquí que los precios deban ser positivos.

un segundo, nos damos cuenta de que este resultado es precisamente el deseado: la demanda de todos los agentes se satisface si se intercambian dotaciones entre ellos. Esto es, en equilibrio,

$$\sum_{k=1}^m \mathbf{x}^k(\mathbf{p}^*) = \sum_{k=1}^m \mathbf{w}^k.$$

Además, se puede demostrar que, a ese nivel de precios, el equilibrio satisface la optimalidad de Pareto, la cual, en esencia, dice que ningún agente puede mejorar sin perjudicar a otro.

Desgraciadamente, este modelo no sugiere cómo se lleva a cabo el intercambio entre los agentes. Para ello, se han propuesto distintos esquemas dinámicos de ajuste de precios con la esperanza de encontrar un precio al cual se converja en el largo plazo. Esto es, desde el punto de vista de los sistemas dinámicos, lo que se busca demostrar es que puede obtenerse un precio de equilibrio \mathbf{p}^* como un punto fijo para el cual las soluciones que inician cerca de \mathbf{p}^* convergen a él en el futuro. Como hemos visto antes, este tipo de comportamiento se lo llama asintóticamente estable. Recordemos que un punto asintóticamente estable de un sistema dinámico es un punto fijo al cual convergen, en el largo plazo, todas las órbitas o trayectorias que inician suficientemente cerca de éste. Véanse al respecto las definiciones 6.3.1 5.2.2 y el ejercicio 5.8.

§8.3 Dinámica de precios

A continuación, se enunciará un conjunto de propiedades que es razonable esperar en un sistema dinámico que sea consistente con el modelo de Arrow-Debreu y con la intuición. Primero, queremos que los puntos fijos del sistema dinámico correspondan exactamente con los puntos de equilibrio de la demanda agregada. Segundo, el ajuste en los precios debe ser “natural”, en el sentido de que, si la demanda agregada de un bien es positiva, entonces el precio debe aumentar y, análogamente, si la demanda agregada es negativa, el precio debe disminuir. Por último, se espera que al menos uno de los puntos de equilibrio sea un atractor.

Existen varias maneras de modelar el ajuste de precios. Una de ellas consiste en la introducción de un “subastador” ficticio, encargado de anunciar precios en cada periodo, hasta que, después de varias iteraciones, el precio lleve a un equilibrio. De este modo, los precios serán ajustados por el subastador de acuerdo con el comportamiento de la demanda agregada: si ésta es positiva, entonces se reduce la demanda y aumenta el precio de ese bien, y se procede de forma análoga si la demanda agregada es negativa.

Si se utiliza una dinámica de precios en tiempo continuo, entonces debe encontrarse una ecuación diferencial con las propiedades descritas. El caso más simple queda descrito por la ecuación

$$\dot{\mathbf{p}} = \xi(\mathbf{p}), \quad (8.1)$$

donde $\xi(\mathbf{p})$ es el exceso de demanda agregada. La idea detrás de esta ecuación es que el *signo* en el cambio de precios debe ser el mismo que en el exceso de demanda. En otras palabras, si existe un exceso de demanda positiva en algún bien, entonces se espera que su precio aumente, es decir, que la derivada sea positiva. Un mecanismo de ajuste más general que también cumple con los requerimientos mínimos enunciados al principio de esta sección es

$$\dot{\mathbf{p}} = f(\xi(\mathbf{p})), \quad (8.2)$$

con f una función diferenciable y $f' > 0$. En general, a este tipo de mecanismo se lo denomina *el subastador walrasiano*.⁴

La versión en tiempo discreto del subastador walrasiano queda dada por

$$\mathbf{p}_{t+1} = f(\xi(\mathbf{p}_t)). \quad (8.3)$$

Concretamente, diversos autores han propuesto el siguiente ajuste para los precios:⁵

$$\mathbf{p}_{t+1} = \mathbf{p}_t + h\xi(\mathbf{p}_t), \quad (8.4)$$

con $h > 0$ una constante arbitraria. Este mecanismo tiene un par de problemas obvios: no se tiene una función homogénea y, dado que $\xi(\mathbf{p}_t)$ puede ser negativa, \mathbf{p}_{t+1} también podría serlo. Quizás una manera más adecuada de proceder sea considerar el sistema dinámico

$$\mathbf{p}_{t+1} = f(\mathbf{p}_t) = M(\mathbf{p}_t)\mathbf{p}_t, \quad (8.5)$$

donde M es una matriz cuadrada que ajusta los precios en la misma dirección del signo de la demanda agregada. Una opción natural es utilizar *factores de descuento* como relación entre los vectores de precios en tiempos sucesivos. Es decir, podríamos usar la siguiente regla de ajuste:

$$M(\mathbf{p}_t) = \begin{pmatrix} \beta_1(\mathbf{p}_t) & 0 & \cdots & 0 \\ 0 & \beta_2(\mathbf{p}_t) & 0 & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & \cdots & 0 & \beta_n(\mathbf{p}_t) \end{pmatrix}, \quad (8.6)$$

donde los elementos de la diagonal son los factores de descuento⁶ que hacen variar el precio según el signo de la demanda agregada. De manera adicional, se debe suponer que $M(\mathbf{p}_t)$ es una función homogénea de grado cero en \mathbf{p}_t .

Supongamos que el exceso de demanda agregada $\xi(\mathbf{p}_t)$ está dado, como antes, por

$$\xi(\mathbf{p}_t) = (\xi_1(\mathbf{p}_t), \dots, \xi_n(\mathbf{p}_t)).$$

Entonces, se esperaría que para cada entrada i , el factor $\beta_i(\mathbf{p}_t)$ fuera homogéneo de grado 0 y que, además, cumpliera con lo siguiente:

- a) $\beta_i(\mathbf{p}_t) > 1$, si $\xi_i(\mathbf{p}_t) > 0$,
- b) $\beta_i(\mathbf{p}_t) = 1$, si $\xi_i(\mathbf{p}_t) = 0$,
- c) $\beta_i(\mathbf{p}_t) < 1$, si $\xi_i(\mathbf{p}_t) < 0$.

⁴Al subastador walrasiano también se le conoce como *tâtonnement*.

⁵Ver por ejemplo [Lor93, Saa96, Day94].

⁶Quizá sería mejor llamarlos **factores de rendimiento**, ya que representan el factor por el cual cambia p_t para obtener p_{t+1} .

Una opción es hacer que el factor de descuento dependa directamente de la demanda agregada de la siguiente forma:

$$\beta_i(\mathbf{p}_t) = \rho_i^{\xi_i(\mathbf{p}_t)}, \quad (8.7)$$

donde ρ_i es una constante mayor que 1. Claramente, el factor de descuento definido en (8.7) satisface las condiciones impuestas. Obviamente, existen otras posibilidades; por ejemplo, utilizando el ajuste propuesto en (8.4) se tendría que $\beta_i(\mathbf{p}_t) = \max\{0, 1 + h_i \xi_i(\mathbf{p}_t)\}$ (esto evita que \mathbf{p}_{t+1} tenga componentes negativas), pero la definición propuesta en (8.7) nos parece más sencilla.

Notemos que la dinámica discreta definida por las ecuaciones (8.5), (8.6) y (8.7) es una función $f : \mathbb{R}_{++}^n \rightarrow \mathbb{R}_{++}^n$, que es homogénea de grado cero. Esto implica que, en realidad, la dinámica depende únicamente de los precios relativos. De este modo, podemos “normalizar” el vector de precios multiplicándolo por una constante, restringiendo así la dinámica a un conjunto de dimensión menor. Por ejemplo, en cada paso puede escogerse el vector de precios que pertenezca al hiperplano $\mathbf{p} \cdot \mathbf{1} = 1$ o a la semiesfera $\mathbf{p} \cdot \mathbf{p} = 1$. Aquí $\mathbf{1} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$. En la figura 8.1 se ilustran estas dos posibilidades para el caso $n = 2$.

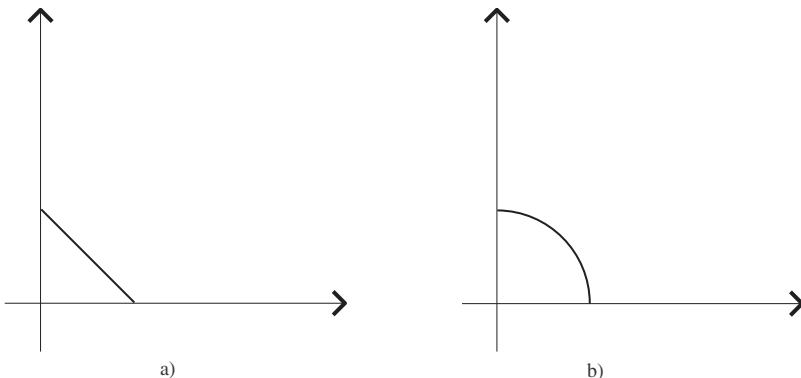


Figura 8.1: En la figura se representan los conjuntos a) $p_1 + p_2 = 1$ y b) $p_1^2 + p_2^2 = 1$. Éstas son dos posibles normalizaciones de precios.

La pregunta que surge ahora es qué tan complicada puede ser la dinámica generada por sistemas de la forma (8.5), donde ξ es una función de exceso de demanda que satisface la proposición 8.2.1. En otras palabras, nos interesa saber si el modelo microeconómico de Arrow-Debreu implícitamente contiene la “mano invisible” de Adam Smith.

Idealmente, esperaríamos que toda dinámica generada por estas funciones de exceso de demanda poseyera un punto atractor. Desgraciadamente, se ha demostrado que en general esto no ocurre. Steve Smale, uno de los matemáticos que más ha contribuido al estudio de los sistemas dinámicos, trabajó varios años en el tema y ha señalado [Sma98, Sma76] que éste es uno de los principales problemas de la economía teórica.

El problema planteado por Smale consiste en la generalización del modelo de Arrow-Debreu de tal modo que la dinámica de precios surja de las decisiones individuales de los agentes. En las ecuaciones (8.1) a (8.5) la dinámica de precios surge a partir de la suposición de que existe un subastador, es decir, un agente externo que va cambiando los precios hasta que todos los agentes satisfacen sus demandas, lo cual, en principio, podría nunca ocurrir. A continuación damos un ejemplo de lo complicado que puede resultar la dinámica generada por una función de exceso de demanda.

§8.4 Caos en el proceso de precios

Si no se introducen hipótesis adicionales y se utilizan los mecanismos de ajuste de precios propuestos, la dinámica de precios resultante puede ser tan complicada como se quiera. Una discusión bastante accesible de esto puede verse en [Saa95]. En esta sección simplemente daremos algunos ejemplos de comportamientos posibles en la dinámica de precios.

Como punto de partida tenemos cualquier función de exceso de demanda que cumpla con la proposición 8.2.1. Es bien sabido, dentro de la literatura económica, que cualquier función de exceso de demanda puede ser generada de la manera que hemos descrito al principio de este capítulo. Es decir, para cada función $\xi : \mathbb{R}_{++}^n \rightarrow \mathbb{R}^n$ que satisface la proposición 8.2.1, existen agentes, preferencias y dotaciones iniciales que generan ξ a través de la maximización de la utilidad.⁷

Considérese una economía con dos bienes y escojamos, por ejemplo, la función $\xi : \mathbb{R}_{++}^2 \rightarrow \mathbb{R}^2$ dada por

$$\begin{aligned}\xi_1(\mathbf{p}) &= \log\left(m \frac{p_2}{p_1 + p_2}\right), \\ \xi_2(\mathbf{p}) &= -\frac{p_1}{p_2} \log\left(m \frac{p_2}{p_1 + p_2}\right),\end{aligned}\tag{8.8}$$

donde \mathbf{p} es el vector de precios $\mathbf{p} = \begin{pmatrix} p_1 \\ p_2 \end{pmatrix}$ y m es una constante positiva. Es fácil verificar que la función ξ satisface la proposición 8.2.1 y la prueba se deja como ejercicio al lector (véase ejercicio 8.3).

A partir de la definición de ξ , puede definirse la dinámica de precios de acuerdo con la idea del subastador walrasiano. Se utiliza la regla de ajuste dada por la ecuación (8.5) y la matriz

$$M(\mathbf{p}) = \begin{pmatrix} \beta_1(\mathbf{p}) & 0 \\ 0 & \beta_2(\mathbf{p}) \end{pmatrix},$$

donde los factores $\beta_1(\mathbf{p})$ y $\beta_2(\mathbf{p})$ están dados por

$$\begin{aligned}\beta_1(\mathbf{p}) &= e^{\xi_1(\mathbf{p})}, \\ \beta_2(\mathbf{p}) &= e^{h \xi_2(\mathbf{p})}.\end{aligned}$$

Los precios se normalizan de tal manera que $p_1 + p_2 = 1$. De este modo, podemos asumir que el vector de precios es de la forma $\mathbf{p} = (p, 1-p)$, con $p \in (0, 1)$. Al hacer esto, la dinámica también queda

⁷Véase por ejemplo [Deb86, Deb73].

normalizada y obtenemos el mapeo $f : [0, 1] \rightarrow [0, 1]$ dado por

$$f(p) = \frac{m p(m - mp)^{\frac{h}{1-p}}}{1 + m p(m - mp)^{\frac{h}{1-p}}}, \quad (8.9)$$

en donde se define $f(0) = f(1) = 0$. La forma de la gráfica de f varía con los parámetros h y m . No pretendemos estudiar a profundidad la dinámica dada por el sistema $p_{t+1} = f(p_t)$. Baste decir que puede ser muy complicada. Una observación simple es que f tiene un punto fijo en $p^* = 1 - 1/m$. Se deja como ejercicio mostrar que la derivada de f en p^* está dada por

$$f'(p^*) = \frac{1 - h(m - 1)^2}{m}.$$

Es claro que existen valores de h y m para los cuales $f'(p^*) > 1$ y p^* es *inestable* (véase ejercicio 8.4).

Se puede observar numéricamente que, escogiendo los parámetros h y m de manera adecuada, la dinámica dada por f es caótica. De esto podemos inferir que, en definitiva, la dinámica de precios dada por el subastador walrasiano no incluye el esperado comportamiento de la mano invisible. Por ejemplo, si escogemos $h = 0.5$ y $m = 8$, entonces se tiene que el sistema $p_{t+1} = f(p_t)$ presenta evidencia de comportamiento caótico, tal como se ilustra en la figura 8.2.

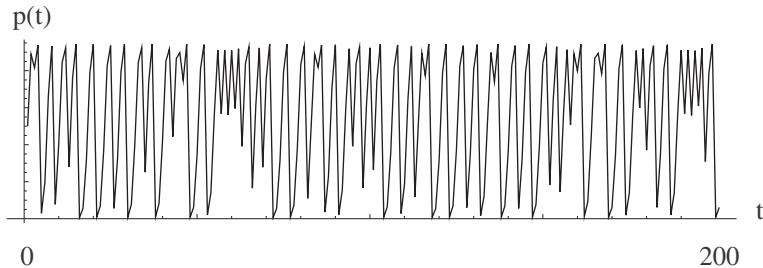


Figura 8.2: Serie de tiempo de las primeras 200 iteraciones del sistema dinámico generado por (8.9), con $h = 0.5$ y $m = 8$.

§8.5 Funciones de Liapunov

Recordemos la definición de estabilidad asintótica.

Definición 8.5.1 Supongamos que el punto fijo \mathbf{p}^* de un sistema dinámico (discreto o continuo) satisface la siguiente condición: existe un número $\delta > 0$ tal que, para cada órbita \mathbf{p}_t , se cumple

$$|\mathbf{p}_0 - \mathbf{p}^*| < \delta \Rightarrow \lim_{t \rightarrow \infty} \mathbf{p}_t = \mathbf{p}^*.$$

Es decir, la solución tiende al equilibrio cuando $t \rightarrow \infty$, siempre y cuando el valor inicial esté suficientemente cerca de \mathbf{p}^* . Se dice, entonces, que el equilibrio es **asintóticamente estable**. Si el límite $\lim_{t \rightarrow \infty} \mathbf{p}_t = \mathbf{p}^*$ es válido para todas las condiciones iniciales, entonces se dirá que \mathbf{p}^* es **globalmente asintóticamente estable**.

Una manera de demostrar que un punto es asintóticamente estable es utilizar los métodos de linearización propuestos en el capítulo 5. El problema con este procedimiento es que, en ocasiones, *existen puntos degenerados que pueden ser asintóticamente estables*.⁸

Por esta razón, se utiliza frecuentemente un método de Liapunov, basado en encontrar una función con ciertas propiedades. En general no es fácil encontrar estas funciones y en ocasiones se requiere de mucho ingenio. Introduciremos a continuación una versión simplificada de la definición de función de Liapunov y sus consecuencias.

Definición 8.5.2 *Sea \mathbf{p}^* un punto fijo de un sistema dinámico (discreto o continuo). Sea V una función continua de variable real definida en una vecindad de \mathbf{p}^* . Se dice que V es una función de Liapunov para \mathbf{p}^* si*

- a) $V(\mathbf{p}^*) = 0$,
- b) $V(\mathbf{p}) > 0$, si $\mathbf{p} \neq \mathbf{p}^*$.
- c) Si \mathbf{p}_t es una órbita del sistema dinámico y $\mathbf{p}_0 \neq \mathbf{p}^*$ entonces $V(\mathbf{p}_t)$ es estrictamente decreciente en t .

En el caso de sistemas dinámicos generados por ecuaciones diferenciales, la última condición puede ser sustituida por una ecuación que es más fácil de verificar. Supongamos que \mathbf{p} cumple

$$\dot{\mathbf{p}} = f(\mathbf{p})$$

y que la función de Liapunov es diferenciable. Entonces, la condición c es equivalente a $V'(\mathbf{p})f(\mathbf{p}) < 0$ para todo $\mathbf{p} \neq \mathbf{p}^*$. Esto equivale a decir $\dot{V} < 0$. Una vez que se tiene una función de Liapunov para un punto dado, se puede utilizar el siguiente teorema, que enunciamos sin demostrar, para probar la estabilidad.

Teorema 8.5.3 (Liapunov)

Si \mathbf{p}^ posee una función de Liapunov entonces es globalmente asintóticamente estable.*

§8.6 Axioma de la preferencia revelada

Anteriormente vimos que si sólo se toman como hipótesis las leyes de la proposición 8.2.1, no puede utilizarse la idea del subastador para demostrar que por lo menos se obtiene un equilibrio atractor de manera dinámica. Por ello, se han propuesto distintas alternativas para resolver este problema, que van desde hacer una seria crítica a la idea de la mano invisible [Saa95] hasta incluir hipótesis adicionales en el modelo. Una de las hipótesis más socorridas es la del axioma de preferencia revelada (véase por ejemplo [MCWG95] o [Var92]).

Este axioma está basado en algunos supuestos acerca del comportamiento de los agentes. Supongamos que se observa que un agente, limitado por su conjunto presupuestal, adquiere una canasta de bienes. Este simple hecho “revela” que el agente la prefiere sobre todas las que puede adquirir dentro de su conjunto

⁸Un ejemplo de esto se da en el ejercicio 5.8.

presupuestal. Se dice entonces que ha “revelado” la preferencia de la canasta adquirida sobre todas las del conjunto presupuestal.

El axioma consiste en la suposición de que si un agente ha revelado su preferencia de una canasta x sobre una canasta y , entonces no podrá suceder subsecuentemente que el agente revele una preferencia de y sobre x . Es decir, supongamos que a un cierto precio p el agente k demanda $x^k(p)$ y a otro precio p' su demanda es $x^k(p')$. Si $x^k(p')$ está dentro del conjunto presupuestal al precio p , entonces el agente “reveló” su preferencia de $x^k(p)$ sobre $x^k(p')$. Por lo tanto, si el agente escoge $x^k(p')$ cuando se anuncia el precio p' es porque su presupuesto en este nivel de precios no le alcanza para adquirir $x^k(p)$. Daremos ahora una definición formal de este concepto.

Definición 8.6.1 (Axioma débil de preferencia revelada) *Se dice que una función de demanda $x(p)$ satisface el **axioma de preferencia revelada** si $p \cdot x(p) \geq p' \cdot x(p')$ implica $p' \cdot x(p) > p' \cdot x(p')$.*

Teorema 8.6.2

Considérese una economía en donde cada uno de los agentes tiene funciones de demanda que satisfacen el axioma de preferencia revelada y sea $\xi(p)$ la función de exceso de demanda agregada. Entonces, para cualquiera pareja de precios p y p' , $p \neq p'$, se cumple que $\xi(p) \neq \xi(p')$ y $p \cdot \xi(p) \leq 0$ implican $p' \cdot \xi(p) > 0$. En este caso, se dice también que la función de demanda agregada satisface el axioma de preferencia revelada.

La demostración del teorema y del siguiente corolario queda como ejercicio para el lector (véanse los ejercicios 8.7 y 8.8).

Corolario 8.6.3 *Si la demanda agregada satisface el axioma de preferencia revelada, entonces el conjunto de puntos de equilibrio es un conjunto convexo.*

Lo que nos interesa ahora es estudiar si existe un punto atractor global luego de añadir el axioma de preferencia revelada a la dinámica determinada por un subastador walrasiano. Al respecto se tiene el siguiente resultado.

Teorema 8.6.4

Supongamos que el exceso de demanda $\xi(p)$ tiene un único punto de equilibrio p^* y satisface el axioma de preferencia revelada. Entonces p^* es un punto globalmente asintóticamente estable para el sistema dinámico

$$\dot{p} = \xi(p).$$

Esto es, para toda solución $p(t)$ de la ecuación se tiene $\lim_{t \rightarrow \infty} p(t) = p^*$.

Demostración

El axioma de preferencia revelada implica que para todo $p \neq p^*$ se cumple que $p^* \cdot \xi(p) > 0$. Sea

$$V(p) = \frac{1}{2} [(p - p^*) \cdot (p - p^*)].$$

Se demostrará que V es una función de Liapunov. Esta función mide la distancia de un vector \mathbf{p} al equilibrio \mathbf{p}^* y claramente se tiene que $V(\mathbf{p}^*) = 0$ y $V(\mathbf{p}) > 0$ si $\mathbf{p} \neq \mathbf{p}^*$. Supongamos que $\mathbf{p}(t)$ es solución de la ecuación $\dot{\mathbf{p}} = \xi(\mathbf{p})$. Entonces,

$$\begin{aligned}\dot{V} &= \frac{d}{dt} V(\mathbf{p}(t)) = \\ &= (\mathbf{p} - \mathbf{p}^*) \frac{d}{dt} (\mathbf{p} - \mathbf{p}^*) \\ &= (\mathbf{p} - \mathbf{p}^*) \cdot \xi(\mathbf{p}) = -\mathbf{p}^* \cdot \xi(\mathbf{p}) < 0.\end{aligned}$$

Nótese que en la última ecuación se utilizó la ley de Walras: $\xi(\mathbf{p}) \cdot \mathbf{p} = 0$. Esto implica que $V(\mathbf{p}(t))$ decrece a lo largo de las órbitas que no inician en \mathbf{p}^* y se concluye que es una función de Liapunov para \mathbf{p}^* . Por el teorema de Liapunov, \mathbf{p}^* es asintóticamente estable.

§8.7 Otras posibilidades

La dinámica de precios no es necesariamente la dinámica surgida del subastador walrasiano. La dinámica se puede ajustar de una manera diferente. Por ejemplo, los agentes pueden ir cambiando, en cada periodo, sus vectores de demanda y sus dotaciones. Los agentes, mediante el comercio, pueden ir mejorando su situación paso a paso. En los modelos que incluyen al subastador walrasiano lo único que se ajusta es el vector de precios y se supone que no se lleva a cabo el intercambio si no se ha llegado al equilibrio. Normalmente, no existe tal subastador y los agentes económicos van ajustando sus dotaciones poco a poco.

Algunas de estas consideraciones se pueden analizar matemáticamente. Baste decir que la herramienta que se utiliza frecuentemente es la que hemos desarrollado aquí, sobre todo lo relacionado con las funciones de Liapunov. Por ejemplo, en [Var92] se mencionan dos maneras en las que la mano invisible podría siempre funcionar: el proceso de Edgeworth y el proceso de Hahn. En ambos casos se demuestra que alguna cantidad siempre decrece si el estado inicial está fuera de equilibrio.

Finalmente se debe mencionar que puede realizarse un análisis parecido tomando las cantidades a producir como aquellas que se ajustan. A este tipo de ajuste se lo denomina **ajuste marshalliano**. En este caso conviene pensar en agentes productores que ajustan la cantidad a producir a la alta o a la baja, dependiendo de si el precio de demanda es mayor o menor que el precio de oferta.

Ejercicios

▷ **8.1** Sea $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{x} \neq \mathbf{0}$. Encontrar constantes positivas α y β tales que $\alpha\mathbf{x} \cdot \mathbf{1} = 1$ y $\beta\mathbf{x} \cdot \beta\mathbf{x} = 1$.

▷ **8.2** Demostrar que si el ajuste de precios $\mathbf{p}(t)$ satisface la ecuación (8.4), entonces la función $\mathbf{p}(t) \cdot \mathbf{p}(t)$ es constante. (Sugerencia: verificar que $\frac{d}{dt}\mathbf{p}(t) \cdot \mathbf{p}(t) = 0$.) Esto implica que se puede restringir la dinámica al conjunto $\mathbf{p}(t) \cdot \mathbf{p}(t) = 1$, que es una porción de esfera en \mathbb{R}_+^n .

▷ **8.3** Demostrar que la función de exceso de demanda dada por (8.8) satisface la proposición 8.2.1.

▷ **8.4** Considerar la función f dada por 8.9.

- Demostrar que si $m > 0$ entonces es un sistema dinámico en $[0, 1]$. Es decir, se cumple que $f([0, 1]) \subset [0, 1]$.
- Demostrar que $p^* = 1 - 1/m$ es un punto fijo del sistema dinámico.
- Demostrar que $f'(p^*) = \frac{1-h(m-1)^2}{m}$. ¿En qué casos p^* es un punto fijo estable?

▷ **8.5** Considerar el sistema

$$\begin{aligned}\dot{x} &= -y - x(x^2 + y^2), \\ \dot{y} &= x - y(x^2 + y^2).\end{aligned}$$

Demostrar que $V(x, y) = x^2 + y^2$ es una función de Liapunov para ese sistema y por lo tanto $(0, 0)$ es asintóticamente estable.

▷ **8.6** Supongamos que la demanda agregada $\xi(\mathbf{p})$ posee un único punto fijo \mathbf{p}^* y la matriz jacobiana $\xi'(\mathbf{p})$ es invertible para cada \mathbf{p} . Un subastador sigue la siguiente regla de ajuste de precios:

$$\dot{\mathbf{p}} = -(\xi'(\mathbf{p}))^{-1} \xi(\mathbf{p}).$$

Demostrar que $V(\mathbf{p}) = \xi(\mathbf{p}) \cdot \xi(\mathbf{p})$ es una función de Liapunov y por lo tanto \mathbf{p}^* es un punto fijo asintóticamente estable.

▷ **8.7** Demostrar el teorema 8.6.2.

▷ **8.8** Demostrar el corolario 8.6.3.

▷ **8.9** Demostrar que, si la demanda agregada satisface el axioma de preferencia revelada y tiene un número finito de puntos de equilibrio, entonces existe un único punto de equilibrio.

CAPÍTULO 9

Ecuaciones en diferencias estocásticas

§9.1 Iteración hacia adelante

En la sección 3.4 de la primera parte de este libro vimos dos modelos monetarios: uno con expectativas adaptativas y el otro con expectativas racionales. Vamos a analizar este concepto con más detalle. Supongamos que $\{x_t\}$ es una sucesión de variables relacionadas por la siguiente ecuación en diferencias:

$$x_t = \alpha x_{t+1} + \beta. \quad (9.1)$$

El lector notará que (9.1) es simplemente una reescritura de (6.3) con $\alpha = \frac{1}{a}$ y $\beta = -\frac{b}{a}$. ¿Cuál es el sentido de hacer esto?

Con frecuencia el valor de una variable en el presente depende, no de los valores que tomó en el pasado, sino de sus valores futuros o más precisamente de lo que esperamos que sean sus valores futuros. El mejor ejemplo es quizás el precio p_t de algún activo; recordemos, por ejemplo, el cálculo del precio de un bono visto en la sección 2.2.3.

Para estos casos es conveniente iterar (9.1) hacia el futuro de la siguiente forma:

$$\begin{aligned} x_t &= \alpha(\alpha x_{t+2} + \beta) + \beta, \\ x_t &= \alpha(\alpha(\alpha x_{t+3} + \beta) + \beta) + \beta, \\ &\vdots \\ x_t &= \alpha^n x_{t+n} + \beta \sum_{k=0}^{n-1} \alpha^k, \end{aligned} \quad (9.2)$$

con lo cual la siguiente proposición es inmediata.

Proposición 9.1.1 *La solución a (9.1) que se obtiene iterando hacia el futuro indefinidamente existe, siempre y cuando se cumplan,*

- a) $|\alpha| < 1$,
- b) $\lim_{n \rightarrow \infty} \alpha^n x_{t+n}$ existe.

En particular, si $\lim_{n \rightarrow \infty} \alpha^n x_{t+n} = 0$, la solución está dada por

$$x_t = \frac{\beta}{1 - \alpha}.$$

Demostración

Reescribamos (9.2) en el límite como

$$x_t = \lim_{n \rightarrow \infty} \alpha^n x_{t+n} + \beta \sum_{k=0}^{\infty} \alpha^k. \quad (9.3)$$

Recordemos que el sistema (6.3) es asintóticamente estable si y sólo si $|a| < 1$. Notemos que (9.3) converge cuando $|\alpha| = |\frac{1}{a}| < 1$ (equivalentemente si $|a| > 1$) y adicionalmente $\lim_{n \rightarrow \infty} \alpha^n x_{t+n}$ existe. Puesto que el segundo término del lado derecho de (9.3) es una serie geométrica convergente, la solución queda expresada como

$$x_t = \lim_{n \rightarrow \infty} \alpha^n x_{t+n} + \frac{\beta}{1 - \alpha},$$

de manera que, si $\lim_{n \rightarrow \infty} \alpha^n x_{t+n} = 0$,

$$x_t = \frac{\beta}{1 - \alpha},$$

lo que da por terminada la demostración. ■

Definición 9.1.2 *Dado el sistema (9.1) que satisface las hipótesis de la proposición 9.1.1 se dice que el término*

$$\beta \sum_{k=0}^{\infty} \alpha^k = \frac{\beta}{1 - \alpha},$$

es la **parte fundamental** de la solución. Por su parte, el término,

$$\lim_{n \rightarrow \infty} \alpha^n x_{t+n}$$

es la parte de **burbuja** de la solución. Si

$$\lim_{n \rightarrow \infty} \alpha^n x_{t+n} \neq 0,$$

entonces la solución no es igual a su parte fundamental y decimos que existen **burbujas especulativas**.

El ejemplo 9.1.1 ilustra estas definiciones para el caso concreto del precio de un activo.

Un caso más general es el siguiente. Supongamos que se tiene la ecuación

$$x_t = \alpha x_{t+1} + \beta y_t. \quad (9.4)$$

Iterando hacia el futuro n veces, obtenemos

$$x_t = \alpha^n x_{t+n} + \beta \sum_{k=0}^{n-1} \alpha^k y_{t+k}. \quad (9.5)$$

Siguiendo el mismo procedimiento de antes, se obtiene la siguiente proposición cuya demostración se deja a cargo del lector por ser casi idéntica a la que acabamos de hacer.

Proposición 9.1.3 *Supongamos que se cumplen $|\alpha| < 1$ y $\lim_{n \rightarrow \infty} \alpha^n x_{t+n} = 0$; en este caso la solución a (9.4) que se obtiene iterando hacia el futuro indefinidamente se puede expresar como*

$$x_t = \beta \sum_{k=0}^{\infty} \alpha^k y_{t+k}.$$

Por supuesto no sabemos si la serie converge o no, a menos que tengamos más información sobre la sucesión $\{y_{t+k}\}_{k=0}^{\infty}$.

Ejemplo

Ej 9.1.1 Sea p_t el precio de un activo con dividendos variables d_t , es decir, dependientes del tiempo. Los dividendos d_t se obtienen al final del periodo t . Sea r la tasa de un activo sin riesgo que suponemos constante en el tiempo. Se debe cumplir entonces la siguiente **condición de no arbitraje**:

$$d_t + (p_{t+1} - p_t) = p_t r.$$

Esta condición simplemente nos dice que en el periodo t es equivalente invertir una cantidad p_t en el activo obteniendo rendimientos de $d_t + (p_{t+1} - p_t)$ en el siguiente periodo, o bien, invertir la misma cantidad en el activo sin riesgo obteniendo rendimientos de $p_t r$. Podemos definir

$$R_t = \frac{d_t + (p_{t+1} - p_t)}{p_t},$$

en donde R_t denota la tasa de rendimiento del activo original. La condición de no arbitraje queda entonces como

$$R_t = r,$$

es decir la tasa R_t debe ser constante e igual a r . Suponemos que se tiene previsión perfecta en el sentido de que se conocen todos los valores de las sucesiones $\{p_t\}$ y $\{d_t\}$. Escribamos la condición de no arbitraje como

$$p_t = \frac{1}{1+r} p_{t+1} + \frac{1}{1+r} d_t,$$

que podemos iterar hacia adelante para obtener,

$$\begin{aligned} p_t &= \left(\frac{1}{1+r} \right)^n p_{t+n} + \left(\frac{1}{1+r} \right) \sum_{k=0}^{n-1} \left(\frac{1}{1+r} \right)^k d_{t+k} \\ &= \left(\frac{1}{1+r} \right)^n p_{t+n} + \sum_{k=0}^{n-1} \left(\frac{1}{1+r} \right)^{k+1} d_{t+k}. \end{aligned}$$

Si no hay burbujas especulativas entonces $\lim_{n \rightarrow \infty} \left(\frac{1}{1+r} \right)^n p_{t+n} = 0$. Tomando el límite cuando $n \rightarrow \infty$ reescribimos la solución como,

$$p_t = \sum_{k=0}^{\infty} \left(\frac{1}{1+r} \right)^{k+1} d_{t+k},$$

es decir, el precio hoy es simplemente el valor presente de los dividendos futuros que es igual a su parte fundamental. Remitimos al lector al ejercicio 2.15, en el cual se realiza un análisis parecido para el caso continuo.

§9.2 Valor esperado y expectativas racionales

Suponer que los individuos tienen previsión perfecta acerca del futuro es una condición sumamente irreal. Por ello es conveniente dar a la ecuación (9.4) un sentido aleatorio y reescribirla como

$$x_t = \alpha E\text{sp}(x_{t+1}) + \beta y_t, \quad (9.6)$$

en donde $E\text{sp}(x_{t+1})$ denota el valor que en el periodo t , el individuo espera que tome x en el siguiente periodo ($t+1$). La forma de calcular $E\text{sp}(x_{t+1})$ puede variar.

Como se mencionó en la sección 3.4, pueden tenerse expectativas adaptativas de manera que el pasado sea el mejor indicador para el futuro. Se mencionó también la posibilidad de tener expectativas racionales, o sea que el individuo toma la esperanza matemática condicionada por toda la información disponible en el periodo. De esta forma, \mathcal{I}_t denota el conjunto de toda la información disponible en t . Éste contiene al conjunto $\{x_{t-i}, y_{t-i} \mid i = 0, 1, \dots, t\}$, es decir, todos los valores pasados de las variables, así como también al conjunto de valores de todos los parámetros del modelo. Denotemos por $E(\cdot \mid \mathcal{I})$ la esperanza matemática dado el conjunto de información \mathcal{I} . En general, éste es un concepto que es difícil de formalizar; sin embargo, intuitivamente tiene sentido. En el caso particular en que $\mathcal{I} = \mathcal{I}_t$, esta esperanza se denota por $E(\cdot \mid \mathcal{I}_t) = E_t(\cdot)$. La siguiente observación es inmediata:

$$E_t(x_{t-i}) = x_{t-i}, \quad i = 0, 1, \dots, t.$$

Estamos suponiendo que los individuos tienen toda la información pasada y ésta es idéntica para todos, de manera que tiene sentido hablar de la esperanza matemática. Esta esperanza se toma con respecto a alguna medida de probabilidad común para todos los agentes. Esta suposición es bastante aventurada y no es trivial, ya que en la mayor parte de los casos tratamos con probabilidades subjetivas. El valor futuro de x_t es desconocido y puede tomar toda una gama de valores, cada uno de ellos con cierta

probabilidad (x_t es una **variable aleatoria** para cada periodo). De alguna manera, los agentes utilizan toda su información disponible para encontrar esta distribución de probabilidad y es con respecto a ésta que se toma el valor esperado o esperanza.

A continuación se dan algunas propiedades de la esperanza condicionada que serán de gran utilidad. (Remitimos al lector a cualquier libro de probabilidad elemental para los detalles.) La primera parte de la siguiente proposición es inmediata; la segunda es conocida como la **ley de esperanzas iteradas**.

Proposición 9.2.1 *Sean \mathcal{I} y \mathcal{J} dos conjuntos de información tales que $\mathcal{I} \subset \mathcal{J}$. Entonces, dada una variable aleatoria x , se cumple*

$$\begin{aligned} E(E(x | \mathcal{I}) | \mathcal{J}) &= E(x | \mathcal{I}), \\ E(E(x | \mathcal{J}) | \mathcal{I}) &= E(x | \mathcal{I}). \end{aligned}$$

Corolario 9.2.2 *Si \mathcal{I}_t denota la información disponible en el periodo t y $k \geq 0$, entonces se cumple*

$$\begin{aligned} E_t(E_{t+k}(\cdot)) &= E_t(\cdot), \\ E_{t+k}(E_t(\cdot)) &= E_t(\cdot). \end{aligned}$$

La demostración es trivial dado que $\mathcal{I}_t \subset \mathcal{I}_{t+k}$, pues nuestra información aumenta con el tiempo.

Corolario 9.2.3 *Si denotamos por $E(\cdot)$ a la esperanza no-condicionada, es decir al valor esperado, se tiene que*

$$E(E_t(\cdot)) = E(\cdot).$$

El resultado anterior es claro ya que podemos pensar que la esperanza no condicionada corresponde al conjunto de información que es vacío.

Una propiedad de la esperanza condicionada (y no condicionada) sumamente útil es que es lineal, o sea, satisface

$$E(ax + by | \mathcal{I}) = aE(x | \mathcal{I}) + bE(y | \mathcal{I}),$$

para cualesquiera x, y variables aleatorias y a, b constantes reales. Recordemos que, dadas dos variables aleatorias, x y y , la covarianza y la varianza están definidas como,

$$\begin{aligned} Cov(x, y) &= E[(x - E(x))(y - E(y))], \\ Var(x) &= Cov(x, x) = E[(x - E(x))^2]. \end{aligned}$$

Decimos que x y y no están correlacionadas si $Cov(x, y) = 0$; en particular, si x y y son independientes (su distribución conjunta es igual al producto de sus distribuciones) se puede ver que $Cov(x, y) = 0$. Las siguientes propiedades se derivan directamente de la definición y son muy útiles.

Sean x y y dos variables aleatorias. Entonces

$$Cov(x, y) = E(xy) - E(x)E(y),$$

en particular, si $E(x) = 0$ o $E(y) = 0$, se tiene que

$$\text{Cov}(x, y) = E(xy).$$

$$\text{Si } x = \sum_{i=1}^n a_i y_i,$$

$$\text{Var}(x) = \sum_{i=1}^n a_i^2 \text{Var}(y_i) + 2 \sum_{i < j} a_i a_j \text{Cov}(y_i, y_j);$$

en particular, si $\text{Cov}(y_i, y_j) = 0, j \neq i$ se tiene que

$$\text{Var}(x) = \sum_{i=1}^n a_i^2 \text{Var}(y_i).$$

Ahora nos dedicaremos a resolver la siguiente versión de la ecuación (9.6):

$$x_t = \alpha E_t(x_{t+1}) + \beta y_t \quad (9.7)$$

con $|\alpha| < 1$. La idea es iterar hacia el futuro, como en la sección anterior.

Para poder resolver iteramos un periodo hacia el futuro y tomamos la esperanza condicionada a la información en t para obtener,

$$E_t(x_{t+1}) = \alpha E_t(E_{t+1}(x_{t+2})) + \beta E_t(y_{t+1}).$$

Utilizando la ley de esperanzas iteradas esta expresión se reescribe como

$$E_t(x_{t+1}) = \alpha E_t(x_{t+2}) + \beta E_t(y_{t+1}), \quad (9.8)$$

y sustituyendo (9.8) en (9.7) obtenemos

$$x_t = \alpha[\alpha E_t(x_{t+2}) + \beta E_t(y_{t+1})] + \beta y_t \quad (9.9)$$

$$= \alpha^2 E_t(x_{t+2}) + \alpha\beta E_t(y_{t+1}) + \beta y_t. \quad (9.10)$$

Si ahora iteramos (9.7) dos periodos hacia el futuro, tomamos la esperanza condicionada en t y aplicamos una vez más la ley de esperanzas iteradas se obtiene

$$E_t(x_{t+2}) = \alpha E_t(x_{t+3}) + \beta E_t(y_{t+2}),$$

y sustituyendo en (9.9) resulta que

$$\begin{aligned} x_t &= \alpha^2[\alpha E_t(x_{t+3}) + \beta E_t(y_{t+2})] + \alpha\beta E_t(y_{t+1}) + \beta y_t \\ &= \alpha^3 E_t(x_{t+3}) + \alpha^2\beta E_t(y_{t+2}) + \alpha\beta E_t(y_{t+1}) + \beta y_t. \end{aligned} \quad (9.11)$$

Después de n iteraciones obtenemos

$$x_t = \alpha^n E_t(x_{t+n}) + \beta \sum_{k=0}^{n-1} \alpha^k E_t(y_{t+k}), \quad (9.12)$$

que es el equivalente estocástico de (9.5). Podemos también formular una proposición similar a 9.1.3.

Proposición 9.2.4 *Supongamos que se cumplen*

- a) $|\alpha| < 1$,
- b) $\lim_{n \rightarrow \infty} \alpha^n E_t(x_{t+n}) = 0$ (*no hay burbujas especulativas*).

En este caso, la solución a (9.7) que se obtiene iterando hacia el futuro indefinidamente puede expresarse como

$$x_t^* = \beta \sum_{k=0}^{\infty} \alpha^k E_t(y_{t+k}).$$

*Ésta es llamada la parte **fundamental** de la solución. Análogamente, $\lim_{n \rightarrow \infty} \alpha^n E_t(x_{t+n})$ es la **burbuja** de la solución.*

Igual que antes, no sabemos si la serie converge o no a menos de que tengamos más información sobre la sucesión $\{E_t(y_{t+k})\}_{k=0}^{\infty}$.

Ejemplo

Ej 9.2.1 Usando el ejemplo 9.1.1 pero en su versión estocástica, tenemos que el precio satisface la ecuación

$$p_t = \frac{1}{1+r} E_t(p_{t+1}) + \frac{1}{1+r} d_t. \quad (9.13)$$

Suponiendo que no hay burbujas especulativas, es decir, $\lim_{n \rightarrow \infty} \left(\frac{1}{1+r}\right)^n E_t(p_{t+n}) = 0$, y tomando el límite cuando $t \rightarrow \infty$ se tiene que

$$p_t = \sum_{k=0}^{\infty} \left(\frac{1}{1+r}\right)^{k+1} E_t(d_{t+k}), \quad (9.14)$$

o sea, el precio hoy es simplemente el valor presente de los dividendos futuros esperados.



§9.3 Algunos procesos estocásticos

En el ejemplo 9.2.1, en cada periodo t el dividendo d_t es una variable aleatoria que sigue alguna distribución de probabilidad. Se dice entonces que la sucesión $\{d_t\}$ forma un **proceso estocástico discreto**. Existen también **procesos estocásticos continuos**, cuando el tiempo se mide en forma continua; sin embargo, en este libro nos restringimos al caso discreto.

Definición 9.3.1 *Sean $\{x_t\}$ un proceso estocástico discreto y*

$$\begin{aligned} \mu_t &= E(x_t), \\ \sigma_{ts} &= E[(x_t - \mu_t)(x_s - \mu_s)], \end{aligned}$$

*su valor esperado y su covarianza. Se dice que $\{x_t\}$ es **estacionario**, si $\mu_t = \mu$ para toda t y σ_{ts} depende únicamente de $|t - s|$.*

Normalmente estaremos interesados en procesos estacionarios debido a que sus propiedades estadísticas son invariantes en el tiempo.

Definición 9.3.2 Se dice que un proceso estocástico $\{x_t\}$ no está **serialmente correlacionado** si $Cov(x_t x_{t-s}) = 0$ para todo $t \geq s > 0$. De otra manera se dice que el proceso está serialmente correlacionado.

En ocasiones suele especificarse el tipo de proceso estocástico que sigue alguna variable. Algunos procesos comunes son los siguientes.

Definición 9.3.3 Sea $\{x_t\}$ un proceso estocástico. Si x_t sigue la misma distribución para toda t decimos que el proceso está **idénticamente distribuido** y denotamos esto por **ID**; si adicionalmente $\{x_t\}$ no está serialmente correlacionado, entonces decimos que es **independiente e idénticamente distribuido** y esto lo denotamos por **IID**.

Definición 9.3.4 Sea $\{x_t\}$ un proceso estocástico IID distribuido normalmente y que satisface

$$\begin{aligned} E(x_t) &= 0 \text{ para toda } t, \\ E(x_t^2) &= \sigma^2 \text{ para toda } t. \end{aligned}$$

Se dice, entonces, que se trata de un proceso de **ruido blanco**. Si la distribución no es normal, decimos que se trata simplemente de un proceso de **ruido**.

Definición 9.3.5 Un proceso estocástico $\{x_t\}$ se llama **martingala** si se cumple $E_t(x_{t+1}) = x_t$.

Definición 9.3.6 Un proceso estocástico $\{x_t\}$ se llama **autorregresivo** si x_t es función de x_{t-s} , con $s \neq 0$; un proceso autorregresivo está serialmente correlacionado.

En los ejemplos que se presentan a continuación, se dan algunos procesos estocásticos comunes.

Ejemplos

Ej 9.3.1 x_t es IID distribuido de acuerdo con una distribución normal con media μ y varianza σ^2 .

Ej 9.3.2 $x_t = \rho x_{t-1} + \delta + \varepsilon_t$, en donde δ y ρ son constantes y ε_t es un proceso de ruido con $Cov(x_t, \varepsilon_t) = 0$, es un proceso autorregresivo llamado **AR1**. Si $|\rho| < 1$, entonces el proceso es estacionario.

Ej 9.3.3 $x_t = x_{t-1} + \varepsilon_t$, en donde ε_t es un proceso de ruido y $Cov(x_t, \varepsilon_t) = 0$, es una martingala. En particular a este tipo de proceso se lo llama **caminata aleatoria**.

Ej 9.3.4 $x_t = x_{t-1} + \delta + \varepsilon_t$, en donde δ es una constante y ε_t es un proceso de ruido con $Cov(x_t, \varepsilon_t) = 0$, es una **caminata aleatoria con deriva**.

Ej 9.3.5 $x_t = \left(\sum_{i=1}^n \rho_i x_{t-i} \right) + \delta + \varepsilon_t$, en donde $\rho_i, i = 1, \dots, n$ y δ son constantes y ε_t es un proceso de ruido con $Cov(x_t, \varepsilon_t) = 0$, es un proceso autorregresivo llamado **ARn**. Si para toda $i | \rho_i | < 1$, entonces el proceso es estacionario.



Los ejemplos anteriores sugieren una clasificación de los procesos estocásticos de acuerdo con su “memoria”. Los procesos más simples son aquellos que carecen de memoria, es decir, cada evento ocurre de manera independiente de sus predecesores, como sucede al lanzar una misma moneda en forma consecutiva. A este tipo de procesos se lo denomina **proceso Bernoulli**. El proceso puede poseer memoria a “corto plazo”, por ejemplo si lo que ocurre en un periodo dado depende de lo que ocurrió en el periodo inmediato anterior pero es independiente de los periodos anteriores a éste. A estos procesos con memoria de un periodo se los llama **procesos de Markov**. Finalmente, la generalización natural es un proceso que tiene memoria de un cierto número finito de periodos en el pasado.

Ejemplo

Ej 9.3.6 Utilizando el ejemplo 9.2.1, supongamos que los dividendos siguen el siguiente proceso estocástico:

$$d_t - \bar{d} = \rho(d_{t-1} - \bar{d}) + \varepsilon_t$$

en donde $E_{t-1}(\varepsilon_t) = 0$, $\bar{d} > 0$ y $\rho < 1 + r$. Iterando k pasos al futuro tenemos que

$$d_{t+k} = \bar{d} + \rho^k(d_t - \bar{d}) + \sum_{i=0}^k \rho^i \varepsilon_{t+i}$$

y tomando la esperanza dado el conjunto de información \mathcal{I}_t , obtenemos

$$E_t(d_{t+k}) = \bar{d} + \rho^k(d_t - \bar{d}).$$

Sustituyendo esta expresión en (9.14), se llega a ¹,

$$\begin{aligned} p_t &= \sum_{k=0}^{\infty} \left\{ \left(\frac{1}{1+r} \right)^{k+1} [\bar{d} + \rho^k(d_t - \bar{d})] \right\} \\ &= \left(\frac{1}{1+r} \right) \sum_{k=0}^{\infty} \left\{ \left(\frac{1}{1+r} \right)^k \bar{d} + \left(\frac{\rho}{1+r} \right)^k (d_t - \bar{d}) \right\} \\ &= \frac{1}{r} \bar{d} + \frac{1}{1+r-\rho} (d_t - \bar{d}). \end{aligned}$$

Si definimos $\bar{p} \equiv \frac{1}{r} \bar{d}$ tenemos,

$$p_t - \bar{p} = \frac{1}{1+r-\rho} (d_t - \bar{d}).$$

Notamos que si se cumple $\rho < r$, entonces $\frac{1}{1+r-\rho} < 1$; de esta forma el precio depende solamente de los dividendos en el mismo periodo y varía proporcionalmente menos que los dividendos.

¹La condición $\rho < 1 + r$ es para garantizar la convergencia de la serie geométrica.

§9.4 Burbujas

Dada la ecuación (9.7), la proposición 9.2.4 nos da la solución fundamental para el caso en que no existen burbujas especulativas; sin embargo, existen otras soluciones que admiten burbujas. Supongamos que x_t^* es la solución fundamental de (9.7) dada en la proposición 9.2.4 y sea

$$x_t = x_t^* + \gamma_t \quad (9.15)$$

cualquier otra solución. Se tiene, entonces, la siguiente proposición acerca del proceso γ_t .

Proposición 9.4.1 *Si $x_t = x_t^* + \gamma_t$ es una solución de (9.7), entonces*

$$\gamma_t = \alpha E_t(\gamma_{t+1}).$$

Demostración

Iterando la expresión (9.15) un periodo hacia el futuro, se tiene que $x_{t+1} = x_{t+1}^* + \gamma_{t+1}$ y por lo tanto, tomando la esperanza E_t ,

$$E_t(x_{t+1}) = E_t(x_{t+1}^*) + E_t(\gamma_{t+1}).$$

Sustituyendo x_t y $E_t(x_{t+1})$ en (9.7), obtenemos

$$x_t^* + \gamma_t = \alpha E_t(x_{t+1}^*) + \alpha E_t(\gamma_{t+1}) + \beta y_t.$$

Como $x_t^* = \beta \sum_{k=0}^{\infty} \alpha^k E_t(y_{t+k})$, iteramos un periodo y tomamos la esperanza E_t para obtener

$$E_t(x_{t+1}^*) = \beta \sum_{k=0}^{\infty} \alpha^k E_t(y_{t+k+1}).$$

Multiplicando esta expresión por α y sustituyendo para x_t^* se tiene que, $x_t^* = \alpha E_t(x_{t+1}^*) + \beta y_t$; por lo tanto $\gamma_t = \alpha E_t(\gamma_{t+1})$, que es el resultado deseado. ■

Dado que $E_t(\gamma_{t+1}) = \gamma_t \alpha^{-1}$, no es difícil verificar que $E_t(\gamma_{t+k}) = \alpha^{-k} \gamma_t$, por lo tanto, dado que $|\alpha| < 1$, se obtiene que el límite

$$\lim_{k \rightarrow \infty} E_t(\gamma_{t+k}) = \lim_{k \rightarrow \infty} \alpha^{-k} \gamma_t,$$

diverge, es decir, la esperanza de γ_t es “explosiva”. El término γ_t representa la parte de burbuja de la solución (9.15).

Ejemplos

Ej 9.4.1 Refiriéndonos al ejemplo 9.2.1 consideramos la solución dada por

$$p_t = \sum_{k=0}^{\infty} \left(\frac{1}{1+r} \right)^{k+1} E_t(d_{t+k}) + \gamma_t.$$

Sea $\gamma_t = \gamma_0(1+r)^t$, con γ_0 dado. Si $d_t = d$ es una constante, entonces la parte fundamental de la solución es

$$p^* = \frac{d}{r}$$

y p_t se reescribe como

$$p_t = \frac{d}{r} + \gamma_0(1+r)^t.$$

Vemos que el precio crece de forma exponencial a pesar de que los dividendos son constantes.

Ej 9.4.2 Consideremos la solución para p_t del ejemplo anterior, pero ahora consideremos la siguiente descripción de γ_t :

$$\gamma_t = \begin{cases} \frac{1+r}{q}\gamma_{t-1} + \varepsilon_t & \text{con probabilidad } q \\ \varepsilon_t & \text{con probabilidad } (1-q) \end{cases}$$

en donde $q \in [0, 1]$ y $E_{t-1}(\varepsilon_t) = 0$. El precio está dado por

$$p_t = \begin{cases} \frac{d}{r} + \frac{1+r}{q}\gamma_{t-1} + \varepsilon_t & \text{con probabilidad } q \\ \frac{d}{r} + \varepsilon_t & \text{con probabilidad } (1-q) \end{cases}$$

De acuerdo con este modelo, en cada periodo la burbuja se “revienta” con probabilidad $1 - q$ y continúa con probabilidad q .



Es común que en los mercados financieros existan períodos en los cuales los precios de algunos activos excedan cualquier predicción acerca del valor presente de los dividendos futuros; en tales casos, existen burbujas especulativas. Históricamente, los ejemplos más famosos de burbujas especulativas son la “manía del tulipán” de 1637 en Holanda, la “burbuja de los Mares del Sur” en 1720 y el alza y la caída de los mercados en 1929. Todos los mercados de capital padecen, de forma intermitente, de burbujas en mayor o menor grado.

Notemos que si x_t está sujeta a una condición terminal en el futuro, digamos en $t = T$, entonces en T se debe tener que $\gamma_T = 0$. Iterando hacia el pasado a partir de $t = T$ se tiene que $\gamma_t = 0$ para todo $t = 0, \dots, T$, de manera que no hay burbujas. Una consecuencia de esta observación es que el valor de un bono no puede tener burbujas especulativas ya que el valor final del bono está dado.

§9.5 Forma reducida

Es común encontrar en economía ecuaciones del tipo

$$x_t = a_0 + a_1 E_{t-1}(x_t) + a_2 E_{t-2}(x_t) + y_t, \quad (9.16)$$

en donde $\{y_t\}$ está determinado de manera exógena y $a_i \in (0, 1)$. Se tiene, entonces que x_t depende de un parámetro y_t y de los valores esperados de x_t que se tuvieron en períodos anteriores.

Definición 9.5.1 Decimos que x_t está expresado en forma reducida si puede expresarse como función de $\{y_t\}$.

§9.5.1 Método de iteración

Es relativamente sencillo expresar x_t en su forma reducida a partir de (9.16) mediante un proceso de iteración. Tomamos la esperanza más antigua, E_{t-2} en este caso, a toda la ecuación y utilizamos la ley de esperanzas iteradas para obtener

$$E_{t-2}(x_t) = a_0 + a_1 E_{t-2}(x_t) + a_2 E_{t-2}(x_t) + E_{t-2}(y_t).$$

Despejamos $E_{t-2}(x_t)$:

$$E_{t-2}(x_t) = \frac{1}{1 - a_1 - a_2} [E_{t-2}(y_t) + a_0]$$

y sustituimos este valor en (9.16) para llegar a

$$x_t = a_0 + a_1 E_{t-1}(x_t) + a_2 \left[\frac{1}{1 - a_1 - a_2} (E_{t-2}(y_t) + a_0) \right] + y_t. \quad (9.17)$$

Tomamos ahora la esperanza siguiente, es decir E_{t-1} , y utilizamos una vez más la ley de esperanzas iteradas para obtener

$$E_{t-1}(x_t) = a_0 + a_1 E_{t-1}(x_t) + a_2 \left[\frac{1}{1 - a_1 - a_2} (E_{t-2}(y_t) + a_0) \right] + E_{t-1}(y_t).$$

De aquí, despejamos $E_{t-1}(x_t)$ y sustituimos en (9.17) para llegar finalmente a la forma reducida para x_t ,

$$\begin{aligned} x_t &= a_0 + a_1 \left[\frac{1}{1 - a_1} \left(a_0 + a_2 \left[\frac{1}{1 - a_1 - a_2} (E_{t-2}(y_t) + a_0) \right] + E_{t-1}(y_t) \right) \right] \\ &\quad + a_2 \left[\frac{1}{1 - a_1 - a_2} (E_{t-2}(y_t) + a_0) \right] + y_t \\ &= y_t + \frac{a_1}{1 - a_1} E_{t-1}(y_t) + \frac{a_2}{(1 - a_1 - a_2)(1 - a_1)} E_{t-2}(y_t) + K, \end{aligned}$$

en donde $K = a_0 \left(1 + \frac{a_1}{1 - a_1} + \frac{a_1 a_2}{1 - a_1} + \frac{a_2}{1 - a_1 - a_2} \right)$.

Es claro que y_t , en general podría consistir en todo un vector de variables exógenas que pueden incluir shocks aleatorios. El siguiente ejemplo ilustra este caso.

Ejemplo

Ej 9.5.1 Consideremos el siguiente modelo de economía agregada:

$$\begin{aligned} y_t^d &= m_t - p_t + v_t, \\ y_t^o &= p_t - E_{t-1}(p_t) + u_t, \end{aligned}$$

en donde y_t^d es la demanda agregada, y_t^o la oferta agregada, m_t los balances nominales, p_t el índice de precios y u_t y v_t son shocks estocásticos. La primera ecuación da a la demanda agregada como los balances reales más un shock estocástico, y la segunda ecuación dice que la oferta agregada está dada por una relación de Phillips más otro shock estocástico. El conjunto de información en cada periodo contiene

todos los valores pasados de las variables y de los shocks. Igualando la oferta y la demanda agregadas se tiene,

$$p_t = \frac{1}{2} (E_{t-1}(p_t) + m_t + v_t - u_t). \quad (9.18)$$

Aplicando E_{t-1} se tiene que

$$E_{t-1}(p_t) = \frac{1}{2} [E_{t-1}(p_t) + E_{t-1}(m_t) + E_{t-1}(v_t) - E_{t-1}(u_t)],$$

de manera que, despejando $E_{t-1}(p_t)$,

$$E_{t-1}(p_t) = E_{t-1}(m_t) + E_{t-1}(v_t) - E_{t-1}(u_t).$$

Sustituimos en (9.18) para obtener

$$p_t = \frac{1}{2} [E_{t-1}(m_t) + E_{t-1}(v_t) - E_{t-1}(u_t) + m_t + v_t - u_t]. \quad (9.19)$$

Ésta es la forma reducida para p_t , puesto que m_t es exógena, pues está dada por la política monetaria. Sustituyendo en la ecuación de oferta se tiene a su vez que

$$\begin{aligned} y_t &= \frac{1}{2} [E_{t-1}(m_t) + E_{t-1}(v_t) - E_{t-1}(u_t) + m_t + v_t - u_t] - \\ &\quad [E_{t-1}(m_t) + E_{t-1}(v_t) - E_{t-1}(u_t)] + u_t, \end{aligned}$$

o bien,

$$y_t = \frac{1}{2} [m_t - E_{t-1}(m_t) + v_t - E_{t-1}(v_t) + u_t + E_{t-1}(u_t)]. \quad (9.20)$$

Supongamos que los shocks u_t y v_t siguen algún tipo de proceso conocido y que en t se conocen todos los valores pasados de los shocks. El Banco Central decide que su política monetaria estará dada de acuerdo con

$$m_t = Au_{t-1} + Bv_{t-1}, \quad (9.21)$$

es decir, ajusta la oferta monetaria de acuerdo con los shocks del periodo anterior. Los coeficientes A y B se escogen para optimizar algún objetivo. Observemos que, aplicando la esperanza E_{t-1} a (9.21), obtenemos

$$E_{t-1}(m_t) = AE_{t-1}(u_{t-1}) + BE_{t-1}(v_{t-1}) = m_t. \quad (9.22)$$

Si sustituimos (9.22) en (9.20) se tiene que m_t no influye en lo absoluto y el ingreso es afectado únicamente por los shocks estocásticos; en otras palabras, la política monetaria no tiene ningún efecto real. Notemos, sin embargo, que al sustituir (9.21) en (9.19) el nivel de precios es proporcional a m_t , de manera que los valores de A y B sí afectan el nivel de precios. Un posible objetivo del Banco Central es minimizar la varianza en el nivel de precios (en aras de la estabilidad). Veamos cómo se lograría esto para el caso específico en que los procesos $\{u_t\}$ y $\{v_t\}$ están serialmente correlacionados. Específicamente, tenemos

$$u_t = \rho_1 u_{t-1} + \eta_t,$$

$$v_t = \rho_2 v_{t-1} + \varepsilon_t,$$

en donde $\rho_1, \rho_2 \in (0, 1)$ y

$$E(\eta_t) = E(\varepsilon_t) = E(\eta_t \eta_{t-i}) = E(\varepsilon_t \varepsilon_{t-i}) = E(\eta_t \varepsilon_s) = 0,$$

de manera que u_t y v_t son procesos independientes. Con esta especificación para los procesos, el precio está dado por

$$p_t = \frac{1}{2} [2(A - \rho_1)u_{t-1} + 2(B + \rho_2)v_{t-1} + \varepsilon_t - \eta_t].$$

La varianza de p_t es

$$Var(p_t) = (A - \rho_1)^2 Var(u_{t-1}) + (B + \rho_2)^2 Var(v_{t-1}) + \frac{1}{4} Var(\varepsilon_t) + \frac{1}{4} Var(\eta_t),$$

y ésta se minimiza escogiendo $A = \rho_1$ y $B = -\rho_2$.

◊◊◊◊◊◊◊◊◊◊◊

Otra ecuación común es la que puede expresarse como

$$x_t = a_0 + a_1 E_{t-1}(x_t) + a_2 E_t(x_{t+1}) + y_t, \quad (9.23)$$

con $a_i \in (0, 1)$. Aquí x_t depende de la esperanza que se tuvo de x_t en el periodo anterior, la esperanza del valor futuro y algún parámetro y_t . La forma reducida se obtiene una vez más iterando de la siguiente forma: tomamos la esperanza E_{t-1} en (9.23) para obtener,

$$E_{t-1}(x_t) = a_0 + a_1 E_{t-1}(x_t) + a_2 E_{t-1}(x_{t+1}) + E_{t-1}(y_t);$$

de aquí despejamos $E_{t-1}(x_t)$:

$$E_{t-1}(x_t) = \frac{1}{1 - a_1} [a_0 + a_2 E_{t-1}(x_{t+1}) + E_{t-1}(y_t)]. \quad (9.24)$$

Iteramos (9.24) un periodo hacia el futuro y tomamos nuevamente la esperanza E_{t-1} :

$$E_{t-1}(x_{t+1}) = \frac{1}{1 - a_1} [a_0 + a_2 E_{t-1}(x_{t+2}) + E_{t-1}(y_{t+1})]. \quad (9.25)$$

Seguimos este proceso para obtener

$$E_{t-1}(x_{t+k}) = \frac{1}{1 - a_1} [a_0 + a_2 E_{t-1}(x_{t+k+1}) + E_{t-1}(y_{t+k})].$$

Sustituimos $E_{t-1}(x_{t+k})$ en $E_{t-1}(x_{t+k-1})$, $E_{t-1}(x_{t+k-1})$ en $E_{t-1}(x_{t+k-2})$, ..., $E_{t-1}(x_{t+2})$ en (9.25) y $E_{t-1}(x_{t+1})$ en (9.24), con lo que obtenemos

$$\begin{aligned} E_{t-1}(x_t) &= \frac{a_0}{1 - a_1} \sum_{i=0}^{k-1} \left(\frac{a_2}{1 - a_1} \right)^i + \frac{1}{1 - a_1} \sum_{i=0}^{k-1} \left(\frac{a_2}{1 - a_1} \right)^i E_{t-1}(y_{t+i}) \\ &\quad + \frac{a_2^k}{(1 - a_1)^k} E_{t-1}(x_{t+k}). \end{aligned} \quad (9.26)$$

Suponiendo que $\frac{a_2}{1-a_1} < 1$ y que $\lim_{k \rightarrow \infty} \frac{a_2^k}{(1-a_1)^k} E_{t-1}(x_{t+k}) = 0$ (esto es necesario para obtener una solución estacionaria), tomamos el límite de (9.26) cuando $k \rightarrow \infty$ para obtener

$$E_{t-1}(x_t) = \frac{a_0}{1-a_1-a_2} + \frac{1}{1-a_1} \sum_{i=0}^{\infty} \left(\frac{a_2}{1-a_1} \right)^i E_{t-1}(y_{t+i}). \quad (9.27)$$

Iteramos (9.27) un periodo hacia el futuro y obtenemos

$$E_t(x_{t+1}) = \frac{a_0}{1-a_1-a_2} + \frac{1}{1-a_1} \sum_{i=0}^{\infty} \left(\frac{a_2}{1-a_1} \right)^i E_t(y_{t+1+i}). \quad (9.28)$$

Sustituyendo (9.27) y (9.28) en (9.23) se llega finalmente a la forma reducida para x_t :

$$\begin{aligned} x_t &= \frac{a_0}{1-a_1-a_2} + \frac{a_1}{1-a_1} \sum_{i=0}^{\infty} \left(\frac{a_2}{1-a_1} \right)^i E_{t-1}(y_{t+i}) \\ &\quad + \frac{a_2}{1-a_1} \sum_{i=0}^{\infty} \left(\frac{a_2}{1-a_1} \right)^i E_t(y_{t+1+i}) + y_t. \end{aligned} \quad (9.29)$$

Esta solución se puede simplificar si conocemos específicamente $\{y_t\}$. Existen diversos métodos de solución para este tipo de ecuaciones; remitimos al lector a [Bla85] para una presentación detallada de éstos. A continuación, se analiza un método común que utiliza la técnica de coeficientes indeterminados.

§9.5.2 Coeficientes indeterminados

Para simplificar la exposición ligeramente, consideremos la ecuación (9.23) con $a_0 = 0$, es decir,

$$x_t = a_1 E_{t-1}(x_t) + a_2 E_t(x_{t+1}) + y_t \quad (9.30)$$

con $a_1, a_2 > 0$ y $a_1 + a_2 < 1$. Conociendo (9.29), proponemos una solución estacionaria de la forma

$$x_t = \sum_{k=0}^{\infty} B_k E_t(y_{t+k}) + \sum_{k=0}^{\infty} C_k E_{t-1}(y_{t-1+k}). \quad (9.31)$$

Calculamos $E_t(x_{t+1})$ y $E_{t-1}(x_t)$, a partir de esta solución, utilizando la ley de esperanzas iteradas:

$$\begin{aligned} E_t(x_{t+1}) &= \sum_{k=0}^{\infty} B_k E_t(y_{t+1+k}) + \sum_{k=0}^{\infty} C_k E_t(y_{t+k}), \\ E_{t-1}(x_t) &= \sum_{k=0}^{\infty} B_k E_{t-1}(y_{t+k}) + \sum_{k=0}^{\infty} C_k E_{t-1}(y_{t-1+k}). \end{aligned}$$

Sustituyendo estas expresiones en (9.30) se obtiene

$$\begin{aligned} x_t &= a_1 \left(\sum_{k=0}^{\infty} B_k E_{t-1}(y_{t+k}) + \sum_{k=0}^{\infty} C_k E_{t-1}(y_{t-1+k}) \right) \\ &\quad + a_2 \left(\sum_{k=0}^{\infty} B_k E_t(y_{t+1+k}) + \sum_{k=0}^{\infty} C_k E_t(y_{t+k}) \right) + y_t. \end{aligned} \quad (9.32)$$

Para que esta expresión satisfaga (9.30) se debe tener la igualdad entre los coeficientes de (9.31) y (9.32); en particular, para el coeficiente de y_t se tiene

$$B_0 = a_2 C_0 + 1;$$

para el de $E_t(y_{t+1})$,

$$B_1 = a_2(B_0 + C_1),$$

y así sucesivamente, de manera que para el coeficiente de $E_t(y_{t+k})$ se tiene

$$B_k = a_2(B_{k-1} + C_k);$$

por otro lado, para el de y_{t-1} se tiene

$$C_0 = a_1 C_0;$$

para el de $E_{t-1}(y_t)$,

$$C_1 = a_1(B_0 + C_1),$$

y así sucesivamente, de manera que para el coeficiente de $E_{t-1}(y_{t+k})$ se tiene

$$C_{k+1} = a_1(B_k + C_{k+1}).$$

Puesto que $a_1 < 1$, $C_0 = 0$, lo que implica

$$\begin{aligned} B_0 &= 1, \\ C_1 &= \frac{a_1}{1 - a_1}, \\ B_k &= \frac{a_2^k}{(1 - a_1)^k}, \\ C_{k+1} &= \frac{a_1 a_2^k}{(1 - a_1)^{k+1}}. \end{aligned}$$

Se puede observar que estos coeficientes son idénticos a los obtenidos con el método de iteración y se requiere la condición $\frac{a_2}{1-a_1} < 1$ para que éstos converjan a 0. En el siguiente ejemplo aplicamos el método de coeficientes indeterminados a ecuaciones del tipo (9.16).

Ejemplo

Ej 9.5.2 (Contratos atrasados)² Consideremos una economía en la cual, en cada periodo t la mitad de la fuerza de trabajo determina sus salarios nominales para el periodo en curso y el periodo siguiente con la información disponible hasta $t - 1$. Podemos pensar que al final del periodo $t - 1$ se especifican los salarios w_t y w_{t+1} . El salario nominal se determina de manera que la esperanza del salario real es la misma para los dos periodos. La economía se describe mediante las siguientes ecuaciones de demanda y

²Basado en [Fis77]

oferta agregadas:

$$y_t^d = m_t - p_t + v_t, \quad (9.33)$$

$$y_t^s = A \left[p_t - \frac{1}{2}(w_t + w_{t-1}) \right] + u_t, \quad (9.34)$$

$$w_t = E_{t-1}(p_t), \quad w_{t-1} = E_{t-2}(p_t), \quad (9.35)$$

en donde todas las variables están en logaritmos, $A > 0$ y los procesos $\{v_t\}$ y $\{u_t\}$ son iguales que en el ejemplo 9.5.1. La ecuación de oferta no es enteramente trivial y se obtiene de la siguiente manera. Si pensamos en una función de producción del tipo $Y_t = N_t^\alpha e^{\gamma_t}$, en donde N es la demanda de trabajadores, $\{\gamma_t\}$ es un proceso estocástico y $\alpha \in (0, 1)$, entonces el problema de optimización para la producción implica que $\frac{\bar{W}_t}{P_t} = \alpha N_t^{\alpha-1} e^{\gamma_t}$ (el salario real es igual al producto marginal del trabajo). Si expresamos todas las variables en logaritmos (representadas por letras minúsculas), se tiene que la demanda por trabajadores está dada por

$$n_t = \frac{p_t - \bar{w}_t + \ln \alpha + \gamma_t}{1 - \alpha},$$

que es igual a la oferta de trabajadores, la cual que suponemos constante. Al sustituir el óptimo en la función de producción, la oferta agregada (en logaritmos) resulta $y_t = \frac{\alpha(p_t - \bar{w}_t + \ln \alpha + \gamma_t)}{1 - \alpha} + \gamma_t$, en donde $1 - \alpha > 0$. Si pensamos en \bar{w}_t como el salario promedio de los dos sectores de trabajadores, entonces $\bar{w}_t = \frac{1}{2}(w_t + w_{t-1})$; definamos $A \equiv \frac{\alpha}{1-\alpha}$ y $u_t \equiv \frac{\alpha}{1-\alpha}(\ln \alpha + \gamma_t) + \gamma_t$, entonces obtenemos la ecuación de oferta agregada dada por (9.34). En un abuso de simplificación podemos asumir $A = 1$, dado que las conclusiones son las mismas y la notación se facilita enormemente

Sustituyendo (9.35) en (9.34) (con $A = 1$) e igualando la oferta y la demanda agregadas se obtiene la siguiente ecuación para p_t :

$$p_t = \frac{1}{4}E_{t-1}(p_t) + \frac{1}{4}E_{t-2}(p_t) + \frac{1}{2}(m_t + v_t - u_t). \quad (9.36)$$

Definamos $y_t \equiv m_t + v_t - u_t$ y propongamos como solución

$$p_t = AE_{t-1}(y_t) + BE_{t-2}(y_t) + Cy_t.$$

Tomando las esperanzas E_{t-1} y E_{t-2} se tiene

$$\begin{aligned} E_{t-1}(p_t) &= AE_{t-1}(y_t) + BE_{t-2}(y_t) + CE_{t-1}(y_t) \\ &= (A + C)E_{t-1}(y_t) + BE_{t-2}(y_t), \\ E_{t-2}(p_t) &= AE_{t-2}(y_t) + BE_{t-2}(y_t) + CE_{t-2}(y_t) \\ &= (A + B + C)E_{t-2}(y_t). \end{aligned}$$

Sustituyendo en (9.36),

$$\begin{aligned} &AE_{t-1}(y_t) + BE_{t-2}(y_t) + Cy_t \\ &= \frac{1}{4}[(A + C)E_{t-1}(y_t) + BE_{t-2}(y_t)] + \frac{1}{4}[(A + B + C)E_{t-2}(y_t)] + \frac{1}{2}y_t \\ &= \frac{1}{4}(A + C)E_{t-1}(y_t) + \frac{1}{4}(A + 2B + C)E_{t-2}(y_t) + \frac{1}{2}y_t \end{aligned}$$

e igualando coeficientes

$$\begin{aligned} A &= \frac{1}{4}(A + C), \\ B &= \frac{1}{4}(A + 2B + C), \\ C &= \frac{1}{2}, \end{aligned}$$

por lo que se tiene $A = \frac{1}{6}$, $B = \frac{1}{3}$ y $C = \frac{1}{2}$, y la forma reducida para p_t es

$$p_t = \frac{1}{6}E_{t-1}(m_t + v_t - u_t) + \frac{1}{3}E_{t-2}(m_t + v_t - u_t) + \frac{1}{2}(m_t + v_t - u_t).$$

Sustituyendo en la ecuación de demanda tenemos que la forma reducida para el ingreso es $y_t =$

$$m_t - \left[\frac{1}{6}E_{t-1}(m_t + v_t - u_t) + \frac{1}{3}E_{t-2}(m_t + v_t - u_t) + \frac{1}{2}(m_t + v_t - u_t) \right] + v_t.$$

◇◇◇◇◇◇◇◇◇

Ejercicios

▷ **9.1** La siguiente es la versión discreta del modelo monetario dado en la sección 3.4; la notación es la misma que antes con las variables en logaritmos.

$$m_t = p_t - \lambda(E_t(p_{t+1}) - p_t)$$

en donde $E_t(p_{t+1} - p_t) = E_t(p_{t+1}) - p_t$ representa la inflación esperada y $\lambda > 0$. Reescribimos como

$$p_t = \frac{\lambda}{1 + \lambda} E_t(p_{t+1}) + \frac{1}{1 + \lambda} m_t$$

- a) Suponiendo que no existen burbujas, resolver para p_t en términos de los balances nominales esperados.
- b) El Banco Central tiene como política fijar los balances nominales en una cantidad constante $m_t = \bar{m}$; aplicar el resultado obtenido en el inciso anterior para este caso.
- c) Inesperadamente, el Banco Central aumenta la oferta monetaria a un nivel $\tilde{m} > \bar{m}$. Describir qué pasa con la trayectoria de p_t . Comparar estos resultados con lo que se obtuvo en 3.4.

▷ **9.2** Probar que la ecuación (9.13) proviene de la condición de no arbitraje: $E_{t-1}(R_t) = r$.

▷ **9.3** En 1937 Hicks [Hic37] inventó el modelo conocido por **IS-LM** con el fin de resumir los contenidos de la teoría keynesiana; a partir de entonces se han hecho cientos de variantes al modelo original. Este ejercicio es una de ellas que incorpora expectativas racionales. Como siempre sean y_t, p_t, m_t e i_t el ingreso, el nivel de precios, los balances nominales y la tasa nominal de interés; asimismo, y^* denota el nivel natural del ingreso, g_t el gasto gubernamental y u_t un shock estocástico. Todas las variables están en logaritmos. La economía queda descrita por las siguientes ecuaciones:

$$y_t^o = y^* + [p_t - E_{t-1}(p_t)], \quad (\text{AS})$$

$$y_t^d = (a - i_t) + g_t + u_t, \quad (\text{IS})$$

$$m_t - p_t = y_t - i_t. \quad (\text{LM})$$

La ecuación AS (“aggregate supply”) representa la oferta agregada mediante una relación de Phillips.

La ecuación IS (“investment = saving”) representa el equilibrio en el mercado de bienes; la cantidad $a - i_t$ es simplemente la inversión, que es inversamente proporcional a la tasa nominal. Finalmente, la ecuación LM (“liquidity preference = money”) iguala la oferta con la demanda de balances reales.

- a) Combinar IS y LM para obtener la demanda agregada (AD)
- b) Combinar AD con AS para obtener la forma reducida de y_t ; concretamente, obtener a y_t como función de $g_t, m_t, E_{t-1}(g_t)$ y $E_{t-1}(m_t)$.
- c) Si los individuos tienen previsión perfecta ($E_{t-1}(x_t) = x_t$), ¿qué pasa con y_t ?

▷ **9.4** Considerar ahora el siguiente modelo con las variables en logaritmos, en donde la demanda simplemente queda dada por

$$y_t = m_t - p_t \quad (\text{AD})$$

que se conoce como una ecuación de velocidad, ya que da una idea de cuantas veces se utiliza una unidad nominal de dinero en la economía (velocidad de circulación, en este caso igual a uno). La oferta agregada está dada por una relación de Phillips:

$$y_t = p_t - E_{t-1}(p_t) + u_t, \quad (\text{AS})$$

y finalmente la autoridad monetaria se apega a la siguiente regla para determinar los balances nominales:

$$m_t = m + aut. \quad (\text{MS})$$

Observar que la oferta es dependiente del shock a la oferta u_t . Definimos el nivel natural de ingreso y_t^* como el que resulta de tener previsión perfecta, en este caso $y_t^* = u_t$.

- a) Utilizar AD, AS y MS para probar que la forma reducida de y_t es función únicamente de u_t .
- b) ¿Cómo se debe escoger el parámetro a para que la varianza $\text{var}(y_t - y_t^*)$ sea mínima?
- ▷ 9.5 Resolver la ecuación estocástica del ejemplo 9.5.1 por medio del método de coeficientes indeterminados.
- ▷ 9.6 En el ejemplo 9.5.2 suponer que los procesos son autorregresivos dados por

$$u_t = \rho_1 u_{t-1} + \eta_t,$$

$$v_t = \rho_2 v_{t-1} + \varepsilon_t,$$

en donde $\rho_1, \rho_2 \in (0, 1)$, $E(\eta_t) = E(\varepsilon_t) = E(\eta_t \eta_{t-i}) = E(\varepsilon_t \varepsilon_{t-i}) = E(\eta_t \varepsilon_s) = 0$. La au-

toridad monetaria determina la oferta de balances nominales como

$$m_t = K_1 u_{t-1} + K_2 v_{t-1}.$$

- a) Reescribir p_t y y_t en términos de los procesos u_t y v_t . ¿Se puede afectar el valor esperado $E(y_t)$ mediante los valores de K_1 y K_2 ?
- b) Determinar K_1 y K_2 de manera que se minimice la varianza de y_t .
- c) Determinar K_1 y K_2 de manera que se minimice la varianza de p_t .
- d) Comparar con el ejemplo 9.5.1. ¿Por qué aquí sí tiene efecto la política monetaria sobre el ingreso?

Parte IV

Intermezzo

CAPÍTULO 10

Optimización estática

La mayor parte de los problemas que se presentan en economía involucran a individuos racionales que resuelven algún tipo de problema de optimización. En este capítulo se recordarán algunos conceptos básicos de análisis convexo y optimización estática que servirán para el desarrollo posterior de los conceptos fundamentales de optimización dinámica. En realidad, nos desviaremos brevemente del tema fundamental de este libro, que es el estudio de las matemáticas de los procesos que cambian con el tiempo. Incluimos, es este punto este repaso ya que es de suma importancia que se tengan claros los conceptos de optimización estática antes de comenzar con la optimización dinámica. Que sirva este capítulo como un *intermezzo*.

§10.1 Análisis convexo

La noción de convexidad es crucial dentro de la teoría de optimización. Por una parte se tiene la definición general de un conjunto convexo y partiendo de ella se definen a las funciones cóncavas, convexas, cuasicónicas, cuasiconvexas, etc. La razón para definir estos conceptos es que, bajo ciertas condiciones de convexidad, las condiciones necesarias para un óptimo local son también suficientes para un óptimo global.

Definición 10.1.1 *Sea $X \subset \mathbb{R}^n$. Se dice que X es **convexo** si para todo $\mathbf{x}, \mathbf{y} \in X$ y para toda $\lambda \in (0, 1)$ se cumple $\lambda\mathbf{x} + (1 - \lambda)\mathbf{y} \in X$.*

Notemos que todos los puntos del segmento entre \mathbf{x} y \mathbf{y} tienen la forma $\lambda\mathbf{x} + (1 - \lambda)\mathbf{y}$, donde $0 \leq \lambda \leq 1$. Por lo tanto, la definición anterior en realidad dice que dados cualquier par de puntos de un conjunto convexo, todo el segmento que los une está totalmente contenido en el conjunto. La figura 10.1 ilustra este concepto.

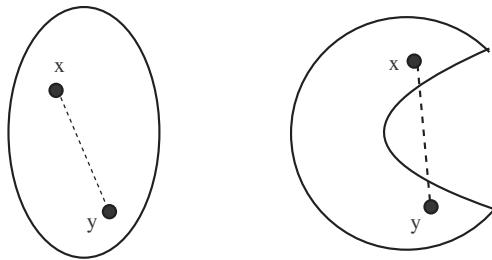


Figura 10.1: Un conjunto convexo y otro no convexo

Proposición 10.1.2 Sean A y B dos subconjuntos convexos de \mathbb{R}^n . Entonces

- a) $A \cap B$ es convexo,
- b) $A + B = \{a + b : a \in A, b \in B\}$ es convexo,
- c) Para todo $k \in \mathbb{R}$, el conjunto $kA = \{ka : a \in A\}$ es convexo.

Demostración:

- a) Sean $x, y \in A \cap B$ y $\lambda \in (0, 1)$. Queremos demostrar que

$$\lambda x + (1 - \lambda)y \in A \cap B.$$

Esto es, queremos verificar que $\lambda x + (1 - \lambda)y \in A$ y $\lambda x + (1 - \lambda)y \in B$. Puesto que $x, y \in A \cap B$, sabemos que $x \in A$, $x \in B$, $y \in A$, $y \in B$. Dado que A es convexo, y dado que $x, y \in A$, concluimos que

$$\lambda x + (1 - \lambda)y \in A.$$

Del mismo modo se puede argumentar que $\lambda x + (1 - \lambda)y \in B$. Por lo tanto,

$$\lambda x + (1 - \lambda)y \in A \cap B.$$

Los otros dos incisos quedan como ejercicio para el lector. ■

Definición 10.1.3 Sea $X \subset \mathbb{R}^n$ un conjunto convexo. $f : X \rightarrow \mathbb{R}$ es una **función convexa** si para todos $x_1 \neq x_2 \in X$ y $\lambda \in (0, 1)$ se tiene

$$f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2).$$

Si la desigualdad es estricta se dice que la función es **estRICTAMENTE CONVEXA**.

Definición 10.1.4 Sea $X \subset \mathbb{R}^n$ un conjunto convexo. $f : X \rightarrow \mathbb{R}$ es una **función cóncava** si para todos $x_1 \neq x_2 \in X$ y $\lambda \in (0, 1)$ se tiene

$$f(\lambda x_1 + (1 - \lambda)x_2) \geq \lambda f(x_1) + (1 - \lambda)f(x_2).$$

Si la desigualdad es estricta se dice que la función es **estRICTAMENTE CÓNCAVA**.

En la primera gráfica de la figura 10.2 se muestra una función cóncava que no es estrictamente cóncava; esto se debe a la parte “plana” de la gráfica. También se muestran las gráficas de funciones estrictamente cóncavas y convexas. Notamos que la definición nos dice que el segmento que une los puntos $(x_1, f(x_1))$ y $(x_2, f(x_2))$ está por debajo de la gráfica de la función si ésta es cóncava y por encima si es convexa.

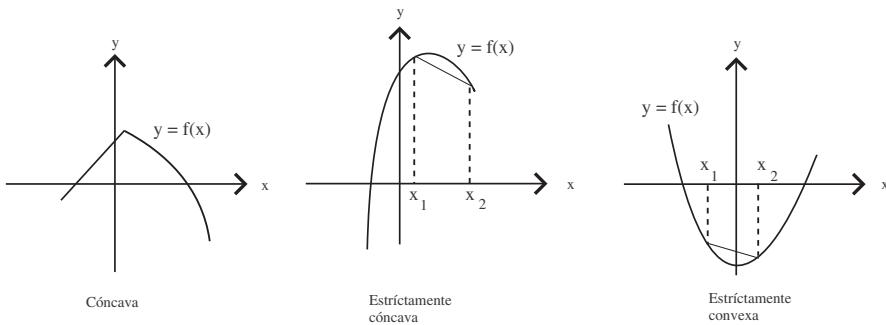


Figura 10.2: Ejemplos de función cóncava, estrictamente cóncava y estrictamente convexa.

Definición 10.1.5 Sea $X \subset \mathbb{R}^n$ y $f : X \rightarrow \mathbb{R}$ una función.

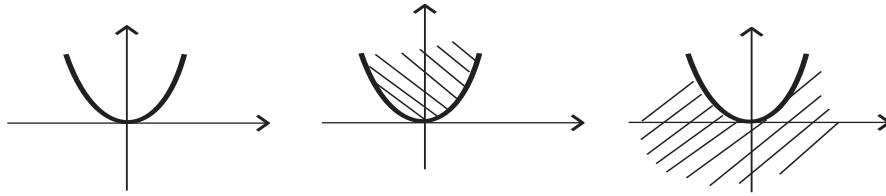
- La **gráfica** de f es el conjunto $G_f = \{(\mathbf{x}, r) \in X \times \mathbb{R} : f(\mathbf{x}) = r\}$.
- El **epígrafo** de f es el conjunto $E_f = \{(\mathbf{x}, r) \in X \times \mathbb{R} : f(\mathbf{x}) \leq r\}$.
- El **hipógrafo** de f es el conjunto $H_f = \{(\mathbf{x}, r) \in X \times \mathbb{R} : f(\mathbf{x}) \geq r\}$.

La figura 10.3 ilustra estos conceptos si f es la función $f(x) = x^2$.

Teorema 10.1.6

Sean $X \subset \mathbb{R}^n$ un conjunto convexo.

- Una función $f : X \rightarrow \mathbb{R}$ es convexa si y sólo si E_f es un conjunto convexo de \mathbb{R}^{n+1} .
- Una función $f : X \rightarrow \mathbb{R}$ es cóncava si y sólo si H_f es un conjunto convexo de \mathbb{R}^{n+1} .

Figura 10.3: Gráfica, epígrafo e hipógrafo de $f(x) = x^2$.

Demostración

Se demostrará el inciso **b**.

- Necesidad.

Supongamos que f es cóncava veamos que H_f es convexo. Sean $(\mathbf{x}, r), (\mathbf{y}, s)$ dos elementos de H_f . Sea $\lambda \in (0, 1)$. Entonces

$$\lambda(\mathbf{x}, r) + (1 - \lambda)(\mathbf{y}, s) = (\lambda\mathbf{x} + (1 - \lambda)\mathbf{y}, \lambda r + (1 - \lambda)s).$$

Sabemos que

$$\begin{aligned} f(\lambda\mathbf{x} + (1 - \lambda)\mathbf{y}) &\geq \lambda f(\mathbf{x}) + (1 - \lambda)f(\mathbf{y}) \\ &\geq \lambda r + (1 - \lambda)s, \end{aligned}$$

pues $(\mathbf{x}, r) \in H_f$ es equivalente a decir $f(\mathbf{x}) \geq r$. Por lo tanto,

$$(\lambda\mathbf{x} + (1 - \lambda)\mathbf{y}, \lambda r + (1 - \lambda)s) \in H_f.$$

- Suficiencia.

Supongamos que H_f es convexo. Notemos que para todo $\mathbf{x}, \mathbf{y} \in X$ se tiene que $(\mathbf{x}, f(\mathbf{x})) \in H_f$ y $(\mathbf{y}, f(\mathbf{y})) \in H_f$. Dado que H_f es convexo, entonces para toda $\lambda \in (0, 1)$

$$\lambda(\mathbf{x}, f(\mathbf{x})) + (1 - \lambda)(\mathbf{y}, f(\mathbf{y})) \in H_f.$$

Esto implica que $(\lambda\mathbf{x} + (1 - \lambda)\mathbf{y}, \lambda f(\mathbf{x}) + (1 - \lambda)f(\mathbf{y})) \in H_f$. Esto es equivalente a decir que

$$f(\lambda\mathbf{x} + (1 - \lambda)\mathbf{y}) \geq \lambda f(\mathbf{x}) + (1 - \lambda)f(\mathbf{y})$$

para todo $\lambda \in (0, 1)$. Concluimos que f es cóncava. ■

Proposición 10.1.7 *Sea $X \subset \mathbb{R}^n$ un conjunto convexo. Sean $f : X \rightarrow \mathbb{R}$ y $g : X \rightarrow \mathbb{R}$ dos funciones cóncavas y $\alpha \in \mathbb{R}$.*

- a) Si $\alpha > 0$, entonces αf es cóncava.
- b) Si $\alpha < 0$, entonces αf es convexa.
- c) $f + g$ es cóncava.
- d) Sea $h : Y \rightarrow \mathbb{R}$ una función cóncava y creciente tal que $g(X) \subset Y \subset \mathbb{R}$. Entonces, $h \circ g$ es cóncava.

La demostración de la proposición anterior se deja como ejercicio para el lector.

Una pregunta natural acerca de las funciones cóncavas (convexas) es si son continuas; el siguiente ejemplo nos da la respuesta:

Ejemplo

Ej 10.1.1 Sea $f : [0, 1] \rightarrow \mathbb{R}$ la función dada por

$$f = \begin{cases} 0 & \text{si } x \in [0, 1], \\ -1 & \text{si } x = 1. \end{cases}$$

Claramente esta función tiene una discontinuidad en $x = 1$, sin embargo, es cóncava pues su hipógrafo es un conjunto convexo como se ilustra en la figura 10.4.

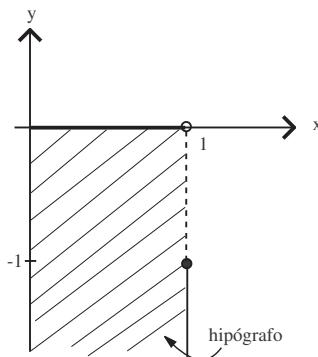


Figura 10.4: El área sombreada representa el hipógrafo de la función.



En el ejemplo anterior el dominio de la función es un conjunto cerrado. Si el dominio es un conjunto abierto, entonces la concavidad (convexidad) de la función implica continuidad. Para una demostración de este resultado remitimos al lector a [PM92]. El caso particular en el cual el dominio de la función es un intervalo (a, b) en \mathbb{R} se deja como ejercicio al lector (véase ejercicio 10.3).

§10.1.1 Caracterización de las funciones cóncavas y convexas

Sea $f \in \mathcal{C}^1(X)$, es decir, $f : X \rightarrow \mathbb{R}$ es una función diferenciable con $X \subset \mathbb{R}^n$. Recordemos que el **gradiente** de f es el vector de las derivadas parciales de f dado por

$$\nabla f(\mathbf{x}) = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix}.$$

Si adicionalmente, $f \in \mathcal{C}^2(X)$, es decir, f es doblemente diferenciable, entonces podemos definir el **hessiano** de f como la matriz de las segundas derivadas dada por

$$Hf(\mathbf{x}) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix}.$$

Nótese que el hessiano es una matriz simétrica.

Ejemplo

Ej 10.1.2 Sea $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ dada por $f(x, y, z) = 15x + xy - 4x^2 - 2y^2 - z^2 + 2yz + 7$. Encontrar la matriz hessiana de f . Las segundas derivadas parciales están dadas por $f_{xx} = -8$, $f_{yy} = -4$, $f_{zz} = -2$, $f_{xy} = f_{yx} = 1$, $f_{xz} = f_{zx} = 0$ y $f_{yz} = f_{zy} = 2$. Por ende, el hessiano está dado por

$$Hf \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -8 & 1 & 0 \\ 1 & -4 & 2 \\ 0 & 2 & -2 \end{pmatrix}.$$



El siguiente teorema da la relación que existe entre funciones convexas (cóncavas) y sus planos (o hiperplanos si $n > 2$) tangentes. Geométricamente dice que la función es convexa si y sólo si el plano tangente está por “debajo” de la gráfica y cóncava si está por “arriba” de la gráfica. El caso más sencillo, cuando $n = 1$, se ilustra en la figura 10.5.

Teorema 10.1.8

Sea $X \subset \mathbb{R}^n$ un conjunto convexo. Suponer que $f \in \mathcal{C}^1(X)$. Entonces

- a) f es convexa si y sólo si para todo $\mathbf{x}, \mathbf{y} \in X$, $f(\mathbf{y}) \geq f(\mathbf{x}) + (\mathbf{y} - \mathbf{x})^T \nabla f(\mathbf{x})$.
- b) f es cóncava si y sólo si para todo $\mathbf{x}, \mathbf{y} \in X$, $f(\mathbf{y}) \leq f(\mathbf{x}) + (\mathbf{y} - \mathbf{x})^T \nabla f(\mathbf{x})$.

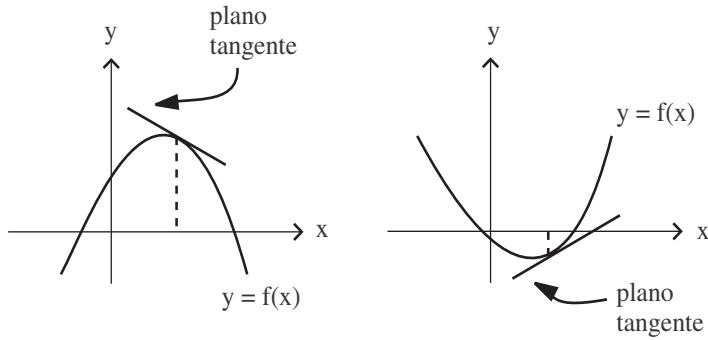


Figura 10.5: Plano tangente (en este caso recta tangente) para una función cóncava y una función convexa.

Demostración:

Se demostrará el inciso **b**.

- Necesidad.

Supongamos que para todo $\mathbf{x}, \mathbf{y} \in X$,

$$f(\mathbf{y}) \leq f(\mathbf{x}) + (\mathbf{y} - \mathbf{x})^T \nabla f(\mathbf{x}).$$

Sean $\mathbf{y}, \mathbf{z} \in X$ y $\lambda \in (0, 1)$. Demostraremos que

$$\lambda f(\mathbf{y}) + (1 - \lambda) f(\mathbf{z}) \leq f(\lambda \mathbf{y} + (1 - \lambda) \mathbf{z}).$$

Definimos $\mathbf{x} = \lambda \mathbf{y} + (1 - \lambda) \mathbf{z}$. Entonces

$$\begin{aligned} \lambda f(\mathbf{y}) + (1 - \lambda) f(\mathbf{z}) &\leq \lambda \left[f(\mathbf{x}) + (\mathbf{y} - \mathbf{x})^T \nabla f(\mathbf{x}) \right] \\ &\quad + (1 - \lambda) \left[f(\mathbf{x}) + (\mathbf{z} - \mathbf{x})^T \nabla f(\mathbf{x}) \right] \\ &= f(\mathbf{x}) + \left[\lambda (\mathbf{y} - \mathbf{x})^T + (1 - \lambda) (\mathbf{z} - \mathbf{x})^T \right] \nabla f(\mathbf{x}) \\ &= f(\mathbf{x}) + (\lambda \mathbf{y} + (1 - \lambda) \mathbf{z} - \lambda \mathbf{x} - (1 - \lambda) \mathbf{x})^T \nabla f(\mathbf{x}) \\ &= f(\mathbf{x}) + (\mathbf{x} - \mathbf{x})^T \nabla f(\mathbf{x}) \\ &= f(\lambda \mathbf{y} + (1 - \lambda) \mathbf{z}). \end{aligned}$$

- Suficiencia.

Supongamos que f es cóncava. Sean $\mathbf{x}, \mathbf{y} \in X$. Definimos para todo $\lambda \in [0, 1]$ la función

$$\varphi(\lambda) = \lambda f(\mathbf{x}) + (1 - \lambda) f(\mathbf{y}) - f(\lambda \mathbf{x} + (1 - \lambda) \mathbf{y}).$$

La concavidad de f implica que $\varphi(\lambda) \leq 0$ para todo $\lambda \in [0, 1]$. Además

$$\begin{aligned}\varphi(0) &= 0, \\ \varphi(1) &= 0.\end{aligned}$$

Las condiciones en φ implican que $\varphi'(0) \leq 0$ y $\varphi'(1) \geq 0$. Dado que

$$\varphi'(\lambda) = f(\mathbf{x}) - f(\mathbf{y}) - (\mathbf{x} - \mathbf{y})^T \nabla f(\lambda \mathbf{x} + (1 - \lambda) \mathbf{y}),$$

para $\lambda = 0$ tenemos

$$\varphi'(0) = f(\mathbf{x}) - f(\mathbf{y}) + (\mathbf{x} - \mathbf{y})^T \nabla f(\mathbf{y}) \leq 0$$

y, análogamente para $\lambda = 1$

$$\varphi'(1) = f(\mathbf{x}) - f(\mathbf{y}) - (\mathbf{x} - \mathbf{y})^T \nabla f(\mathbf{x}) \geq 0.$$

Por lo tanto, de esta última desigualdad se tiene que $f(\mathbf{x}) + (\mathbf{y} - \mathbf{x})^T \nabla f(\mathbf{x}) \geq f(\mathbf{y})$. ■

§10.1.2 Matrices definidas

Definición 10.1.9 *Sea A una matriz simétrica. Se dice que*

- a) *A es **positiva definida** si para todo $\mathbf{x} \neq \mathbf{0}$ se tiene que $\mathbf{x}^T A \mathbf{x} > 0$.*
- b) *A es **negativa definida** si $-A$ es positiva definida, es decir, si para todo $\mathbf{x} \neq \mathbf{0}$ se tiene que $\mathbf{x}^T A \mathbf{x} < 0$.*
- c) *A es **positiva semidefinida** si para todo \mathbf{x} se tiene que $\mathbf{x}^T A \mathbf{x} \geq 0$.*
- d) *A es **negativa semidefinida** si $-A$ es positiva semidefinida, es decir, si para todo \mathbf{x} se tiene que $\mathbf{x}^T A \mathbf{x} \leq 0$.*

Ejemplo

Ej 10.1.3 La siguiente matriz es positiva semidefinida, pero no positiva definida:

$$A = \begin{pmatrix} 3 & 0 \\ 0 & 0 \end{pmatrix}.$$

Para verificarlo, sea $\mathbf{x}_0 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, es decir $\mathbf{x}_0 \neq \mathbf{0}$, y sin embargo $\mathbf{x}_0^T A \mathbf{x}_0 = 0$. Verifiquemos que A es positiva semidefinida. Sea $\mathbf{x} = \begin{pmatrix} a \\ b \end{pmatrix}$, entonces

$$\mathbf{x}^T A \mathbf{x} = (a \quad b) \begin{pmatrix} 3 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = 3a^2 \geq 0.$$



Toda matriz positiva definida es positiva semidefinida y toda matriz negativa definida es negativa semidefinida.

Definición 10.1.10 Sea A una matriz de $n \times n$. La matriz B es una **submatriz principal**¹ de A si B se obtiene a partir de A eliminando los últimos k renglones y columnas, $k = 0, 1, \dots, n$.

Definición 10.1.11 Sea A una matriz de $n \times n$. Los **menores principales** de A son los determinantes de las submatrices principales.

Ejemplo

Ej 10.1.4 Encontrar las submatrices y menores principales de la matriz

$$B = \begin{pmatrix} 5 & 0 & 7 \\ 3 & -2 & 2 \\ 1 & 10 & -3 \end{pmatrix}.$$

Usando la definición 10.1.10, podemos ver que las submatrices principales son B , $\begin{pmatrix} 5 & 0 \\ 3 & -2 \end{pmatrix}$ y (5) . los menores principales son los determinantes correspondientes, es decir, 154, -10 y 5.



El siguiente teorema da una manera fácil de verificar que una matriz es positiva definida, negativa definida, etc. La demostración se puede encontrar en [BS94].

Teorema 10.1.12

Sea A una matriz simétrica.

- a) A es positiva definida si y sólo si todos sus menores principales son positivos.
- b) A es positiva semidefinida si y sólo si todos sus menores principales son no negativos.

Finalmente, el siguiente teorema nos da una forma de determinar la concavidad de una función mediante su matriz hessiana.

Teorema 10.1.13

Sea $X \subset \mathbb{R}^n$ un conjunto convexo. Sea $f : X \rightarrow \mathbb{R}$ una función de clase C^2 . Entonces,

- a) f es cóncava si y sólo si para todo $\mathbf{x} \in X$ el hessiano $Hf(\mathbf{x})$ es negativo semidefinido.
- b) f es convexa si y sólo si para todo $\mathbf{x} \in X$ el hessiano $Hf(\mathbf{x})$ es positivo semidefinido.
- c) Si para todo $\mathbf{x} \in X$ $Hf(\mathbf{x})$ es negativa definida, entonces f es estrictamente cóncava.
- d) Si para todo $\mathbf{x} \in X$ $Hf(\mathbf{x})$ es positiva definida, entonces f es estrictamente convexa.

La demostración de este teorema se deja como ejercicio.

¹En ocasiones se define **submatriz principal** de una matriz A de $n \times n$ como aquella que se obtiene de eliminar cualesquiera k renglones y sus columnas correspondientes. Si se trata de los últimos k renglones y columnas entonces se denomina **submatriz principal dominante**. Véase por ejemplo [BS94].

Ejemplo

Ej 10.1.5 Verificar que la función del ejemplo 10.1.1 es estrictamente cóncava. Sabemos que el hessiano es

$$Hf \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -8 & 1 & 0 \\ 1 & -4 & 2 \\ 0 & 2 & -2 \end{pmatrix}.$$

Veremos que $-H$ es una matriz positiva definida.

$$-H = \begin{pmatrix} 8 & -1 & 0 \\ -1 & 4 & -2 \\ 0 & -2 & 2 \end{pmatrix}.$$

Las submatrices principales son

$$(8), \quad \begin{pmatrix} 8 & -1 \\ -1 & 4 \end{pmatrix} \text{ y } -H,$$

y los menores principales correspondientes: 8, 31 y 30. Por lo tanto, H es negativa definida y la función es estrictamente cóncava.



§10.1.3 Funciones cuasi cóncavas y cuasi convexas

Definición 10.1.14 Sean $X \subset \mathbb{R}^n$ un conjunto convexo y $f : X \rightarrow \mathbb{R}$.

- a) El **contorno de f en k** es el conjunto $C_f(k) = \{\mathbf{x} \in X : f(\mathbf{x}) = k\}$.
- b) El **contorno superior de f en k** es $CS_f(k) = \{\mathbf{x} \in X : f(\mathbf{x}) \geq k\}$.
- c) El **contorno inferior de f en k** es $CI_f(k) = \{\mathbf{x} \in X : f(\mathbf{x}) \leq k\}$.

Los contornos pueden ser conjuntos vacíos.

Ejemplo

Ej 10.1.6 Encontrar el contorno superior e inferior de la función $f(x, y) = xy$ en $k = 1$. La figura 10.6 ilustra los contornos correspondientes, nótese que el contorno superior es convexo.



Proposición 10.1.15 Para una función cóncava f el contorno superior $CS_f(k)$ siempre es convexo.

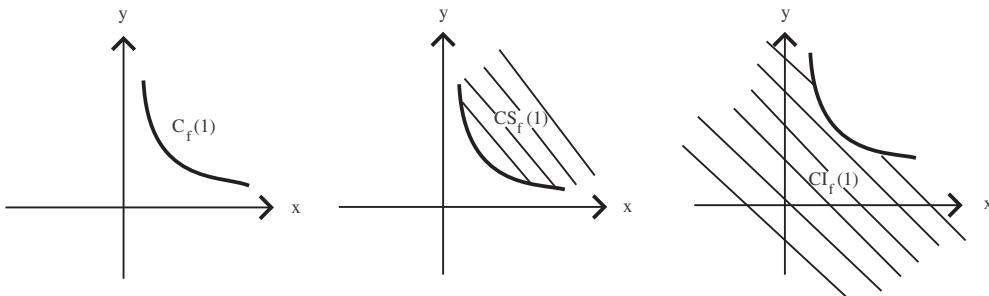


Figura 10.6: Contorno, contorno superior y contorno inferior de $f(x, y) = xy$ en $k = 1$.

Demostración

Sea $f : X \rightarrow \mathbb{R}$ una función cóncava. Sea k en la imagen de f . Por definición,

$$CS_f(k) = \{\mathbf{x} \in X : f(\mathbf{x}) \geq k\}.$$

Demostraremos que $CS_f(k)$ es un conjunto convexo. Sean $\mathbf{a}, \mathbf{b} \in CS_f(k)$ y $\lambda \in (0, 1)$. Sabemos que $f(\mathbf{a}) \geq k$, $f(\mathbf{b}) \geq k$ y que f es cóncava. Entonces

$$\begin{aligned} f(\lambda\mathbf{a} + (1 - \lambda)\mathbf{b}) &\geq \lambda f(\mathbf{a}) + (1 - \lambda)f(\mathbf{b}) \\ &\geq \lambda k + (1 - \lambda)k = k. \end{aligned}$$

Por lo tanto, $\lambda\mathbf{a} + (1 - \lambda)\mathbf{b} \in CS_f(k)$, lo que implica que $CS_f(k)$ es convexo. ■

De igual modo podemos relacionar el contorno inferior con las funciones convexas.

Proposición 10.1.16 Si f es una función convexa, $CI_f(k)$ es un conjunto convexo para todo k en la imagen de f .

Las anteriores son condiciones necesarias más no suficientes. Si f es cóncava (convexa), entonces el conjunto $CS_f(k)$ ($CI_f(k)$) es convexo, pero si el conjunto es convexo, no necesariamente f es cóncava (convexa). Es decir, existen funciones que cumplen con la condición de que todos los contornos superiores (inferiores) son conjuntos convexos y, sin embargo, no son funciones convexas (cóncavas).

Ejemplo

Ej 10.1.7 Sea $f : \mathbb{R} \rightarrow \mathbb{R}$ la función dada por $f(x) = x^3$. Notemos que $CI_f(k) = (-\infty, \sqrt[3]{k}]$. Claramente f no es convexa; sin embargo, $CI_f(k)$ es un conjunto convexo para toda k .



El ejemplo anterior motiva la siguiente definición.

Definición 10.1.17 Sea $X \subset \mathbb{R}^n$ un conjunto convexo. Se dice que una función $f : X \rightarrow \mathbb{R}$ es **cuasi cóncava** si, para todo k en la imagen de f , el conjunto $CS_f(k)$ es convexo. Se dice que una función $f : X \rightarrow \mathbb{R}$ es **cuasi convexa** si, para todo k en la imagen de f , el conjunto $CI_f(k)$ es convexo.

Ejemplos

Ej 10.1.8 La función de la figura 10.7 claramente no es cóncava. El contorno superior $CS_f(k)$ siempre es convexo por lo que se trata de una función cuasi cóncava.

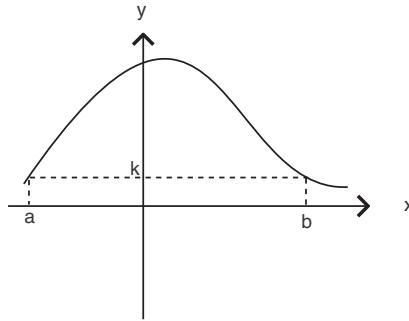


Figura 10.7: Función cuasi cóncava con $CS_f(k) = [a, b]$.

Ej 10.1.9 La figura 10.8 representa una función Cobb-Douglas dada por $f(x, y) = x^{\frac{1}{3}}y^{\frac{5}{3}}$. Una curva de nivel típica se muestra en la figura 10.9, en donde se aprecia que el contorno superior dado por $CS_f(k) = \{(x, y) \in \mathbb{R}^2 \mid x^{\frac{1}{3}}y^{\frac{5}{3}} \geq k\}$ es convexo.



Podríamos establecer una especie de equivalente del teorema 10.1.13 definiendo la matriz conocida como **hessiano orlado**; esto nos daría otro criterio para determinar la cuasi concavidad (cuasi convexidad) de una función. Aquí simplemente utilizaremos el criterio de la convexidad de los contornos superiores (inferiores), puesto que la mayoría de las funciones consideradas tienen dominios en \mathbb{R} o en \mathbb{R}^2 , con lo cual este criterio es fácil de visualizar geométricamente.

El siguiente teorema, que se enuncia sin demostración, nos da una forma de obtener funciones cuasi cóncavas.

Teorema 10.1.18

Sean $X \subset \mathbb{R}^n$ un conjunto convexo, $g : X \rightarrow \mathbb{R}$ una función cuasi cóncava y $h : Y \subset \mathbb{R} \rightarrow \mathbb{R}$ una función creciente donde $g(X) \subset Y$. Entonces $h \circ g$ es cuasi cóncava.

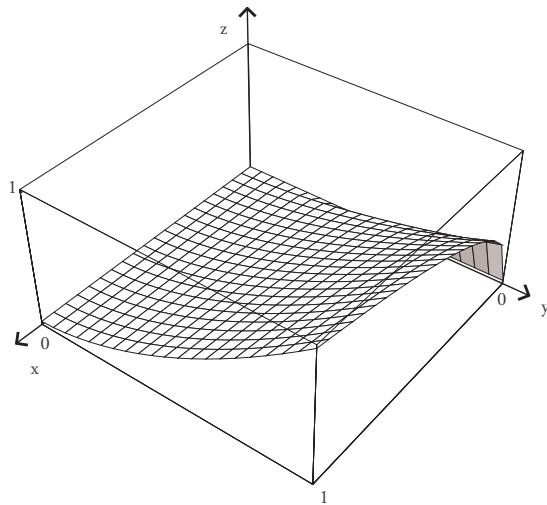


Figura 10.8: Función Cobb-Douglas

Ejemplo

Ej 10.1.10 Sea g cuasi cóncava. Entonces $\log(g)$ y $\exp(g)$ son cuasi cóncavas. En particular, $\log(xy)$ es cuasi cóncava.

◇◇◇◇◇◇◇◇◇

§10.2 Optimización estática.

El problema general de optimización consiste en maximizar (minimizar) una función en un conjunto X , cumpliendo adicionalmente ciertas restricciones. Es decir, deseamos encontrar un punto $\mathbf{x}^* \in X \subset \mathbb{R}^n$

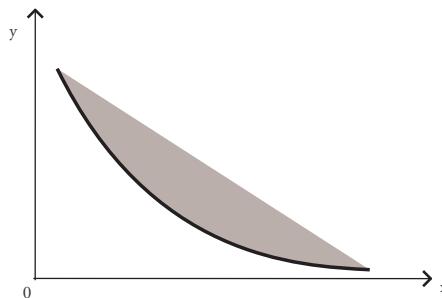


Figura 10.9: Contorno superior de una función Cobb-Douglas

que maximice una función f y que además cumpla con restricciones del tipo

$$g_1(\mathbf{x}^*) \leq 0, \dots, g_m(\mathbf{x}^*) \leq 0.$$

donde g_1, \dots, g_m son funciones de clase \mathcal{C}^1 en X . Si las restricciones son de la forma

$$\begin{aligned} g_1(\mathbf{x}^*) &= 0, \\ &\vdots \\ g_m(\mathbf{x}^*) &= 0, \end{aligned} \tag{10.1}$$

con $m < n$, se dice que tenemos restricciones de igualdad. Generalmente se piensa que X es un subconjunto convexo² de \mathbb{R}^n .

El problema clásico de **programación no lineal** es el siguiente:

$$\begin{aligned} \max \quad & f(\mathbf{x}) \\ \text{sujeto a} \quad & \mathbf{x} \in X, \\ & g_1(\mathbf{x}) \leq 0, \\ & \vdots \\ & g_m(\mathbf{x}) \leq 0. \end{aligned}$$

En caso de que las funciones f y g_k , $k = 1, \dots, m$ sean lineales, decimos que se trata de un problema de **programación lineal**.³

Sea $X \subset \mathbb{R}^n$ y $f : X \rightarrow \mathbb{R}$ una función de clase \mathcal{C}^2 . Recordemos que si $\mathbf{x}^* \in X$ es un máximo (mínimo) local de f entonces se tiene que $\nabla f(\mathbf{x}^*) = 0$. Adicionalmente, si \mathbf{h} es un vector de dirección arbitrario en \mathbb{R}^n , al efectuar el desarrollo en serie de Taylor alrededor de \mathbf{x}^* se obtiene

$$\begin{aligned} f(\mathbf{x}^* + \alpha\mathbf{h}) &\simeq f(\mathbf{x}^*) + \alpha\mathbf{h}\nabla f(\mathbf{x}^*) + \frac{1}{2}\alpha^2\mathbf{h}^T Hf(\mathbf{x}^* + \alpha\mathbf{h})\mathbf{h} \\ &= f(\mathbf{x}^*) + \frac{1}{2}\alpha^2\mathbf{h}^T Hf(\mathbf{x}^* + \alpha\mathbf{h})\mathbf{h}, \end{aligned}$$

en donde α es “suficientemente” pequeño. De esta forma, si la matriz hessiana $Hf(\mathbf{x}^* + \alpha\mathbf{h})$ es negativa definida (positiva definida), entonces se tiene que $f(\mathbf{x}^* + \alpha\mathbf{h}) < f(\mathbf{x}^*)$ ($f(\mathbf{x}^* + \alpha\mathbf{h}) > f(\mathbf{x}^*)$), por lo cual \mathbf{x}^* es un máximo (mínimo) local. En particular, el teorema 10.1.13 implica que si f es una función cóncava (convexa) entonces \mathbf{x}^* es un máximo (mínimo) global de f . Aún más, es posible generalizar este resultado cuando X es convexo y f es cuasi cóncava (cuasi convexa); en este caso, todo extremo local es un extremo global. Finalmente, si X es un conjunto convexo y la función es estrictamente cóncava, entonces, el punto óptimo es único.

²Formalmente, el conjunto X debe ser también lo que se conoce como un conjunto abierto; sin embargo, aquí no tomaremos en cuenta estas consideraciones topológicas.

³En este caso es común también incluir las condiciones de no negatividad, es decir, $x_i > 0 \forall i = 1, \dots, n$.

§10.2.1 Restricciones de igualdad

El problema de programación no lineal (PNL) con restricciones de igualdad es

$$\begin{aligned} \max \quad & f(\mathbf{x}) \\ \text{sujeto a} \quad & \mathbf{x} \in X, \\ & g_1(\mathbf{x}) = 0, \\ & \vdots \\ & g_m(\mathbf{x}) = 0. \end{aligned}$$

Para resolver PNL con restricciones de igualdad, definiremos una nueva función llamada lagrangiano.

Definición 10.2.1 Sea $X \subset \mathbb{R}^n$. Sean $f : X \rightarrow \mathbb{R}$ de clase C^1 y g_1, \dots, g_m , también de clase C^1 en X . El **lagrangiano** asociado al problema de programación no lineal

$$\begin{aligned} \max \quad & f(\mathbf{x}) \\ \text{sujeto a} \quad & \mathbf{x} \in X, \\ & g_1(\mathbf{x}) = 0, \\ & \vdots \\ & g_m(\mathbf{x}) = 0, \end{aligned}$$

es la función $\mathcal{L} : X \times \mathbb{R}^m \rightarrow \mathbb{R}$, dada por,

$$\mathcal{L}(\mathbf{x}, \lambda) = f(\mathbf{x}) - \sum_{k=1}^m \lambda_k g_k(\mathbf{x}).$$

En términos económicos, la función f podría ser una función de beneficios netos dependiente de \mathbf{x} y las g_k podrían representar restricciones de recursos. En el diagrama 10.10, el máximo no está en la curva $f = f_1$, pues con la misma restricción, $g = 0$, se puede mejorar la utilidad; tampoco está en $f = f_3$ pues estaríamos fuera del presupuesto. La utilidad, por lo tanto, se maximiza en donde la función es tangente a la curva presupuestal, en $f = f_2$.

Notamos que en el máximo, el gradiente de la función a maximizar es paralelo al gradiente de la restricción; es decir, si estamos considerando

$$\begin{aligned} \max \quad & f(\mathbf{x}) \\ \text{sujeto a} \quad & g(\mathbf{x}) = 0; \end{aligned}$$

entonces, en el máximo \mathbf{x}^* , se tiene $\nabla f(\mathbf{x}^*) = \lambda^* \nabla g(\mathbf{x}^*)$ para alguna $\lambda^* \in \mathbb{R}$. Tenemos así dos condiciones necesarias, siempre y cuando $\nabla g(\mathbf{x}^*) \neq \mathbf{0}$:

$$\begin{aligned} g(\mathbf{x}^*) &= 0, \\ \nabla f(\mathbf{x}^*) &= \lambda^* \nabla g(\mathbf{x}^*), \end{aligned}$$

para alguna λ^* . El lagrangiano es en este caso

$$\mathcal{L}(\mathbf{x}, \lambda) = f(\mathbf{x}) - \lambda g(\mathbf{x}),$$

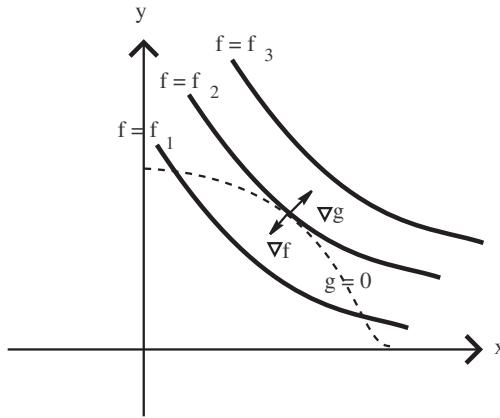


Figura 10.10: Solución al problema de maximización.

y su gradiente es

$$\nabla \mathcal{L}(\mathbf{x}, \lambda) = \begin{pmatrix} \frac{\partial f}{\partial x_1} - \lambda \frac{\partial g}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} - \lambda \frac{\partial g}{\partial x_n} \\ -g(\mathbf{x}) \end{pmatrix} = \begin{pmatrix} \nabla f(\mathbf{x}) - \lambda \nabla g(\mathbf{x}) \\ -g(\mathbf{x}) \end{pmatrix}.$$

Por lo tanto, si \mathbf{x}^* es solución del PNL, entonces también existe λ^* tal que $\nabla \mathcal{L}(\mathbf{x}^*, \lambda^*) = 0$. La siguiente proposición generaliza estas consideraciones.

Proposición 10.2.2 (Condiciones necesarias) *Sea $X \subset \mathbb{R}^n$. Sean $f : X \rightarrow \mathbb{R}$ y $g_1, \dots, g_m : X \rightarrow \mathbb{R}$ funciones de clase C^1 con $m < n$. Supongamos que \mathbf{x}^* es una solución del problema*

$$\begin{aligned} \max \quad & f(\mathbf{x}) \\ \text{sujeto a} \quad & \mathbf{x} \in X, \\ & g_1(\mathbf{x}) = 0, \\ & \vdots \\ & g_m(\mathbf{x}) = 0. \end{aligned}$$

Si suponemos que en \mathbf{x}^ el conjunto de vectores $\{\nabla g_k(\mathbf{x}^*)\}_{k=1}^m$ es linealmente independiente, entonces se puede escribir $\nabla f(\mathbf{x}^*)$ como combinación lineal de estos vectores, es decir, existen $\lambda_1^*, \dots, \lambda_m^* \in \mathbb{R}$ tales que el gradiente del lagrangiano es cero, o bien*

$$\begin{aligned} \nabla f(\mathbf{x}^*) &= \sum_{k=1}^m \lambda_k^* \nabla g_k(\mathbf{x}^*), \\ g_1(\mathbf{x}^*) &= \dots = g_m(\mathbf{x}^*) = 0. \end{aligned}$$

En los ejercicios de este capítulo se pide la demostración formal para el caso particular de $n = 2$ y $m = 1$. Para una prueba del caso general se remite al lector a [BS94].

En analogía al caso de optimización libre se tiene la siguiente proposición.

Proposición 10.2.3 (Condiciones suficientes) *Con la misma notación que antes, supongamos que $X \subset \mathbb{R}^n$ es convexo y sea $(\mathbf{x}^*, \lambda^*)$ tal que $\nabla \mathcal{L}(\mathbf{x}^*, \lambda^*) = \mathbf{0}$; entonces, si $\mathcal{L}(\mathbf{x}^*, \lambda^*)$ es cuasi cóncava, \mathbf{x}^* es un máximo global para el problema de la proposición 10.2.2.*

Como caso particular se tiene el siguiente corolario:

Corolario 10.2.4 *Si f es cuasi-cóncava y g_k es lineal para toda $k = 1, \dots, m$, entonces \mathbf{x}^* es un máximo global para el problema de la proposición 10.2.2.*

Es claro que se cumplen los resultados análogos para el caso de un problema de minimización.

Ejemplos

Ej 10.2.1 Encontrar la distancia máxima entre la circunferencia $x^2 + y^2 = 1$ y el punto $(3, 4)$. Tenemos

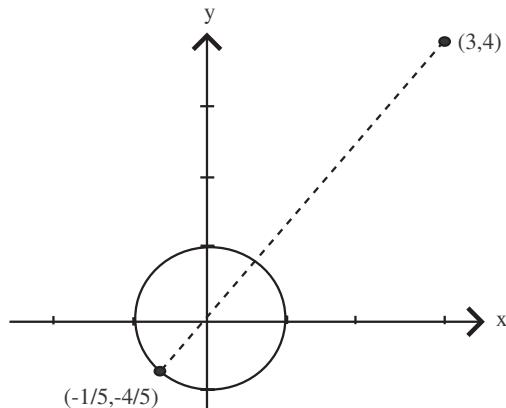


Figura 10.11: Distancia máxima entre $(3, 4)$ y la circunferencia $x^2 + y^2 = 1$.

que elegir la función a maximizar y la restricción. En este caso, tenemos la siguiente restricción de igualdad:

$$g(x, y) = x^2 + y^2 - 1$$

y la función a maximizar es $f(x, y) = (x - 3)^2 + (y - 4)^2 = (d[(x, y), (3, 4)])^2$. El lagrangiano es, por ende,

$$\begin{aligned}\mathcal{L}(x, y, \lambda) &= f(x, y) - \lambda g(x, y) \\ &= (x - 3)^2 + (y - 4)^2 - \lambda(x^2 + y^2 - 1)\end{aligned}$$

El gradiente del lagrangiano es

$$\nabla \mathcal{L}(x, y, \lambda) = \begin{pmatrix} 2(x - 3) - 2\lambda x \\ 2(y - 4) - 2\lambda y \\ -(x^2 + y^2 - 1) \end{pmatrix}.$$

Queremos resolver el siguiente sistema de ecuaciones:

$$\begin{aligned} 2(x - 3) - 2\lambda x &= 0, \\ 2(y - 4) - 2\lambda y &= 0, \\ x^2 + y^2 &= 1. \end{aligned}$$

Existen dos soluciones que satisfacen este sistema de ecuaciones, concretamente

$$\begin{aligned} (x_1, y_1, \lambda_1) &= \left(\frac{3}{5}, \frac{4}{5}, -4\right) \text{ y} \\ (x_2, y_2, \lambda_2) &= \left(-\frac{3}{5}, -\frac{4}{5}, 6\right). \end{aligned}$$

Sabemos, por la figura 10.11, que el máximo se satisface en $x = -\frac{3}{5}$ y $y = -\frac{4}{5}$, por lo tanto la distancia máxima es de 6.

Ej 10.2.2 Consideremos el siguiente problema:

$$\begin{aligned} \max f(x, y) &= -y \\ \text{sujeto a } y^3 - x^2 &= 0. \end{aligned}$$

De la figura 10.12 se infiere que el máximo de la función se obtiene en el punto $(0, 0)$; sin embargo,

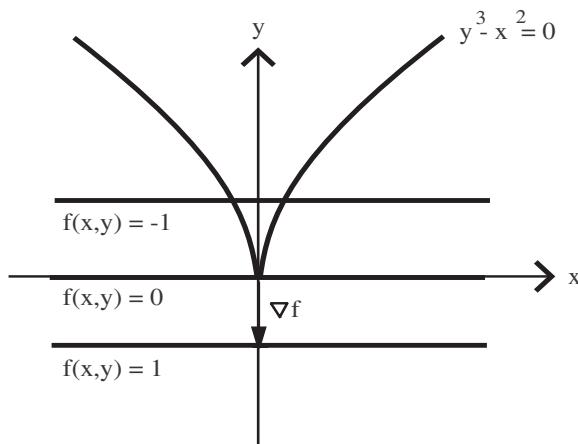


Figura 10.12: El vector $\nabla g(0, 0)$ es nulo.

observemos que $\nabla f(x, y) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$ y $\nabla g(x, y) = \begin{pmatrix} -2x \\ 3y^2 \end{pmatrix}$, de manera que en $(0, 0)$ se tiene $\nabla g(0, 0) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ y no puede existir λ^* tal que $\nabla f(0, 0) = \lambda^* \nabla g(0, 0)$. Aquí el punto óptimo no satisface la condición necesaria de la proposición 10.2.2, pero no la contradice puesto que el conjunto de vectores $\{\nabla g_k(\mathbf{x}^*)\}_{k=1}^m$ no es linealmente independiente; concretamente, aquí se tiene $k = 1$ y el vector $\nabla g(0, 0)$ es nulo.



§10.2.2 Dos aplicaciones microeconómicas.

§10.2.2.1 Problema del consumidor

Consideremos n bienes x_1, \dots, x_n . Supongamos que se tiene una función de utilidad que es el logaritmo de una función tipo Cobb-Douglas, es decir,

$$U(x_1, \dots, x_n) = \sum_{k=1}^n \alpha_k \log(x_k),$$

donde $\alpha_1 + \dots + \alpha_n = 1$. Suponemos que la utilidad se mide en “unidades de utilidad”, normalmente llamadas **útiles**. Supongamos, además, que en el mercado los bienes tienen precios p_1, \dots, p_n . También tenemos la siguiente restricción presupuestal, basada en la dotación inicial, c ,

$$\sum_{k=1}^n p_k x_k = c.$$

El problema es encontrar la “canasta óptima” $\mathbf{x}^* = (x_1^*, \dots, x_n^*)$ que maximiza U sujeta a la restricción presupuestal.

El espacio en donde se trabaja es $X = \mathbb{R}_{++}^n$, es decir,

$$X = \{(x_1, \dots, x_n) \in \mathbb{R}^n \mid x_1 > 0, \dots, x_n > 0\}.$$

La función a maximizar es $U : X \rightarrow \mathbb{R}$, donde

$$U(x_1, \dots, x_n) = \sum_{k=1}^n \alpha_k \log(x_k).$$

Si $g : X \rightarrow \mathbb{R}$ es $g(x_1, \dots, x_n) = \sum_{k=1}^n p_k x_k - c$, entonces la restricción es $g(\mathbf{x}) = 0$. Resolvamos ahora el problema

$$\begin{aligned} \max \quad & U(\mathbf{x}) \\ \text{sujeto a} \quad & \mathbf{x} \in X, \\ & g(\mathbf{x}) = 0. \end{aligned}$$

El lagrangiano es $\mathcal{L}(x_1, \dots, x_n, \lambda) = U(x_1, \dots, x_n) - \lambda g(x_1, \dots, x_n)$. Una condición necesaria dada por la proposición 10.2.2 es $\nabla \mathcal{L}(x_1, \dots, x_n, \lambda) = 0$. En este caso

$$\nabla \mathcal{L}(x_1, \dots, x_n, \lambda) = \begin{pmatrix} \frac{\alpha_1}{x_1} - \lambda p_1 \\ \vdots \\ \frac{\alpha_n}{x_n} - \lambda p_n \\ -g(x_1, \dots, x_n) \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}.$$

Obtenemos, de este modo, el sistema de ecuaciones

$$\begin{aligned} \alpha_1 &= \lambda p_1 x_1, \\ &\vdots \\ \alpha_n &= \lambda p_n x_n, \\ c &= \sum_{k=1}^n p_k x_k. \end{aligned}$$

Si sumamos las primeras n ecuaciones, tenemos que

$$1 = \lambda \sum_{k=1}^n p_k x_k = \lambda c,$$

por lo que concluimos que $\lambda^* = \frac{1}{c}$ y, por lo tanto, la canasta óptima está dada por

$$\begin{aligned} x_1^* &= \frac{\alpha_1}{\lambda p_1} = \frac{\alpha_1 c}{p_1}, \\ &\vdots \\ x_n^* &= \frac{\alpha_n}{\lambda p_n} = \frac{\alpha_n c}{p_n}. \end{aligned}$$

§10.2.2.2 Problema del productor

Se tiene una fábrica y se supone que el nivel de producción se modela con una función Cobb-Douglas $f(K, L) = AK^\alpha L^\beta$. Se necesita producir q unidades y los precios de L y K son w y r , respectivamente. Se desea obtener una función que represente el costo de producir bajo estas condiciones. Es decir, $C = C(w, r, q)$. Esta función es la solución del siguiente problema de optimización:

$$\begin{aligned} \min \quad & wL + rK \\ \text{sujeto a} \quad & f(K, L) = q. \end{aligned}$$

Definimos el siguiente lagrangiano:

$$\mathcal{L}(K, L, \lambda) = wL + rK - \lambda (AK^\alpha L^\beta - q).$$

Las condiciones de primer orden son

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial K} &= r - \lambda A\alpha K^{\alpha-1} L^\beta = 0, \\ \frac{\partial \mathcal{L}}{\partial L} &= w - \lambda A\beta K^\alpha L^{\beta-1} = 0, \\ \frac{\partial \mathcal{L}}{\partial \lambda} &= -(AK^\alpha L^\beta - q) = 0. \end{aligned}$$

Entonces

$$\begin{aligned} r &= \lambda A \alpha K^{\alpha-1} L^\beta, \\ w &= \lambda A \beta K^\alpha L^{\beta-1}, \\ \frac{r}{w} &= \left(\frac{\alpha}{\beta}\right) \left(\frac{L}{K}\right) \end{aligned}$$

y por lo tanto

$$L = \left(\frac{\beta}{\alpha}\right) \left(\frac{r}{w}\right) K.$$

Esta última expresión implica

$$q = AK^\alpha L^\beta = AK^\alpha \left(\frac{\beta}{\alpha}\right)^\beta \left(\frac{r}{w}\right)^\beta K^\beta$$

y de aquí se obtiene

$$K^{\alpha+\beta} = \left(\frac{q}{A}\right) \left(\frac{\alpha}{\beta}\right)^\beta \left(\frac{w}{r}\right)^\beta,$$

con lo cual obtenemos finalmente los valores óptimos

$$\begin{aligned} K^* &= \left(\frac{q}{A}\right)^{\frac{1}{\alpha+\beta}} \left(\frac{\alpha}{\beta}\right)^{\frac{\beta}{\alpha+\beta}} \left(\frac{w}{r}\right)^{\frac{\beta}{\alpha+\beta}}, \\ L^* &= \left(\frac{q}{A}\right)^{\frac{1}{\alpha+\beta}} \left(\frac{\beta}{\alpha}\right)^{\frac{\alpha}{\alpha+\beta}} \left(\frac{r}{w}\right)^{\frac{\alpha}{\alpha+\beta}}. \end{aligned}$$

Concluimos que $C = C(w, r, q) = rK^* + wL^*$, esto es, el costo de producir q unidades usando los valores óptimos, K^* y L^* . Al simplificar se obtiene

$$C(w, r, q) = \left[\left(\frac{\alpha}{\beta}\right)^{\frac{\beta}{\alpha+\beta}} + \left(\frac{\beta}{\alpha}\right)^{\frac{\alpha}{\alpha+\beta}} \right] \left[\left(\frac{q}{A}\right)^{\frac{1}{\alpha+\beta}} w^{\frac{\beta}{\alpha+\beta}} r^{\frac{\alpha}{\alpha+\beta}} \right].$$

Observemos que el cambio en el costo con respecto al precio de los insumos es

$$\begin{aligned} \frac{\partial C}{\partial w} &= \left[\left(\frac{\alpha}{\beta}\right)^{\frac{\beta}{\alpha+\beta}} + \left(\frac{\beta}{\alpha}\right)^{\frac{\alpha}{\alpha+\beta}} \right] \left[\left(\frac{q}{A}\right)^{\frac{1}{\alpha+\beta}} \left(\frac{\beta}{\alpha+\beta}\right) w^{-\frac{\alpha}{\alpha+\beta}} r^{\frac{\alpha}{\alpha+\beta}} \right] \\ &= \left(\frac{\beta}{\alpha}\right)^{\frac{\alpha}{\alpha+\beta}} \left(\frac{q}{A}\right)^{\frac{1}{\alpha+\beta}} \left(\frac{r}{w}\right)^{\frac{\alpha}{\alpha+\beta}} = L^*, \end{aligned}$$

y del mismo modo

$$\frac{\partial C}{\partial r} = K^*.$$

Posteriormente, daremos una explicación de este hecho (cf. lema 10.2.6).

§10.2.3 Condiciones de Kuhn-Tucker

Queremos ahora resolver el siguiente problema con restricciones de desigualdad:

$$\begin{aligned} \max \quad & f(\mathbf{x}) \\ \text{sujeto a} \quad & g_1(\mathbf{x}) \leq 0, \\ & \vdots \\ & g_m(\mathbf{x}) \leq 0, \end{aligned}$$

donde $\mathbf{x} \in X \subset \mathbb{R}^n$, con X convexo, y las funciones f, g_1, \dots, g_m son de clase C^1 . En este caso no es necesario pedir $m < n$ dado que se trata de restricciones de desigualdad. Al igual que antes, se define el lagrangiano como

$$\mathcal{L}(\mathbf{x}, \lambda_1, \dots, \lambda_m) = f(\mathbf{x}) - \sum_{k=1}^m \lambda_k g_k(\mathbf{x}).$$

Sea $(\mathbf{x}^*, \lambda_1^*, \dots, \lambda_m^*)$ un máximo de f . Las restricciones pueden satisfacerse de dos maneras: con la igualdad, y entonces decimos que la restricción está **activa**, o con la desigualdad estricta, y entonces decimos que está **inactiva**. Sean g_1, \dots, g_h , con $h \leq m$, las restricciones que están activas en el punto óptimo y supongamos que el conjunto de vectores $\{\nabla g_k(\mathbf{x}^*)\}_{k=1}^h$ es linealmente independiente (ésta se conoce como condición de **cualificación de las restricciones** y puede ser reemplazada por otras condiciones; remitimos al lector a [Tak85]). En este caso, unas condiciones necesarias para que $(\mathbf{x}^*, \lambda_1^*, \dots, \lambda_m^*)$ sea máximo son las llamadas **condiciones de Kuhn-Tucker** (CKT):

$$\frac{\partial \mathcal{L}}{\partial x_j} = 0, \quad j = 1, \dots, n, \quad (10.2)$$

$$g_k(\mathbf{x}^*) \leq 0, \quad (10.3)$$

$$\lambda_k^* \geq 0, \quad (10.4)$$

$$\lambda_k^* g_k(\mathbf{x}^*) = 0, \quad k = 1, \dots, m. \quad (10.5)$$

Las condiciones (10.5) son conocidas como **condiciones de holgura** y nos dicen que si $\lambda_k^* > 0$, entonces $g_k(\mathbf{x}^*) = 0$ (la restricción está activa) y $\lambda_k^* = 0$ cuando $g_k(\mathbf{x}^*) < 0$ (la restricción no está activa).

Intuitivamente, CKT generaliza el caso en el cual todas las restricciones son de igualdad dado por la proposición 10.2.2, ya que $\nabla f(\mathbf{x}^*)$ es una combinación lineal del conjunto de gradientes $\{\nabla g_j(\mathbf{x}^*)\}_A$, en donde A es el conjunto de restricciones activas, es decir, se cumple $g_j(\mathbf{x}) = 0 \iff j \in A$. Sin embargo, aquí se tiene que la combinación lineal es positiva puesto que $\lambda_k \geq 0$ para $k = 1, \dots, m$, o bien que $\nabla f(\mathbf{x}^*)$ pertenece al “cono positivo” generado por el conjunto $\{\nabla g_j(\mathbf{x}^*)\}_A$. La demostración de este hecho, aparentemente inocuo, no es trivial y remitimos al lector a [Lue84].

Las condiciones de suficiencia son un poco más elaboradas para este caso y quedan resumidas por la siguiente proposición.

Proposición 10.2.5 (Condiciones de suficiencia) *Sean f, g_1, \dots, g_m funciones de clase C^1 con dominio convexo, $X \subset \mathbb{R}^n$. Sea*

$$\mathcal{D} = \{\mathbf{x} \in X \mid g_i(\mathbf{x}) \leq 0, i = 1, \dots, m\},$$

entonces, si se satisface CKT para $\mathbf{x}^* \in \mathcal{D}$, f es cuasicóncava y las funciones g_1, \dots, g_m son cuasiconvexas, se tiene que \mathbf{x}^* es un máximo global en \mathcal{D} si al menos una de las siguientes se cumple:

- a) $\nabla f(x^*) \neq 0$,
- b) f es cóncava.

Ejemplos

Ej 10.2.3

$$\begin{array}{ll} \max & f(x, y) = x - \frac{x^2}{2} + y^2 \\ \text{sujeto a} & \frac{x^2}{2} + y^2 \leq \frac{9}{8}, \\ & -y \leq 0. \end{array}$$

Escribimos primero el lagrangiano correspondiente:

$$\mathcal{L}(x, y, \lambda, \mu) = x - \frac{x^2}{2} + y^2 - \lambda \left(\frac{x^2}{2} + y^2 - \frac{9}{8} \right) - \mu(-y).$$

Las condiciones de Kuhn-Tucker son

$$\begin{array}{ll} \frac{\partial \mathcal{L}}{\partial x} = 1 - x - \lambda x = 0, & \frac{\partial \mathcal{L}}{\partial y} = 2y - 2\lambda y + \mu = 0, \\ \frac{x^2}{2} + y^2 \leq \frac{9}{8}, & -y \leq 0, \\ \lambda \geq 0, & \mu \geq 0, \\ \lambda \left(\frac{x^2}{2} + y^2 - \frac{9}{8} \right) = 0, & \mu(-y) = 0. \end{array}$$

Consideremos cuatro casos

- a) $\lambda = 0, \mu = 0$; esto implica

$$\begin{aligned} 1 - x &= 0 \\ 2y &= 0 \end{aligned}$$

Por lo que $x = 1$ y $y = 0$. Tenemos que verificar si $x = 1, y = 0$ satisface las otras CKT. Sustituyendo, es fácil ver que, en efecto, las satisface.

- b) $\lambda > 0, \mu = 0$; al ser $\lambda > 0$, la restricción correspondiente se activa, es decir,

$$\lambda \left(\frac{x^2}{2} + y^2 - \frac{9}{8} \right) = 0,$$

lo que implica

$$\frac{x^2}{2} + y^2 - \frac{9}{8} = 0.$$

Si $\mu = 0$, entonces

$$2y - 2\lambda y = 0,$$

es decir, $(1 - \lambda)y = 0$. Resolvemos el sistema

$$\begin{aligned} 1 - x - \lambda x &= 0, \\ (1 - \lambda)y &= 0, \\ \frac{x^2}{2} + y^2 &= \frac{9}{8}. \end{aligned}$$

Si $\lambda \neq 1$, entonces $y = 0$ y $x = \pm\frac{3}{2}$. Si $x = \frac{3}{2}$ entonces

$$1 - \frac{3}{2} - \lambda \frac{3}{2} = 0,$$

lo que implica $\lambda = -\frac{1}{3}$; además, si $x = -\frac{3}{2}$, entonces

$$1 + \frac{3}{2} + \lambda \frac{3}{2} = 0,$$

lo que implica $\lambda = -\frac{5}{3}$. Dado que se supuso $\lambda > 0$, el caso $\lambda \neq 1$ no cumple con CKT, entonces sólo es posible $\lambda = 1$, con $x = \frac{1}{2}$ y $y = \pm 1$. La igualdad $y = -1$ viola CKT, por lo tanto, sólo se considerará $y = 1$. Concluimos así que sólo el punto $x = \frac{1}{2}$, $y = 1$ cumple con CKT.

- c) $\lambda = 0, \mu > 0$; si $\lambda = 0$, entonces $x = 1$, y dado que $\mu > 0$, la restricción correspondiente es activa, es decir,

$$-y = 0,$$

o sea, $x = 1, y = 0$, posiblemente satisface CKT. Para ello, debe satisfacer $2y - 2\lambda y + \mu = 0$ y por lo tanto $\mu = 0$. Esto contradice la suposición de que $\mu > 0$.

- d) $\lambda > 0, \mu > 0$; de igual modo, $\mu > 0$ nos lleva a una contradicción. Por lo tanto, este caso también es imposible.

Consideraremos sólo las siguientes posibilidades:

$$\begin{aligned} x_1 &= 1, y_1 = 0; \\ x_2 &= \frac{1}{2}, y_2 = 1. \end{aligned}$$

Comparando el valor de $f(x, y)$ en $(1, 0)$ y $(\frac{1}{2}, 1)$, concluimos que el máximo ocurre en $(x^*, y^*) = (\frac{1}{2}, 1)$, para el cual $f(\frac{1}{2}, 1) = \frac{11}{8}$. La figura 10.13 ilustra este resultado.



La restricción $-y \leq 0$ dada en el ejemplo anterior aparece en la mayor parte de los problemas de economía ya que es simplemente una restricción de no negatividad: $y \geq 0$. Es fácil ver que CKT para las restricciones de no negatividad dadas por $x_1, \dots, x_n \geq 0$ equivale a las siguientes:

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial x_j} &\leq 0, \quad j = 1, \dots, n, \\ x_j \frac{\partial \mathcal{L}}{\partial x_j} &= 0, \quad j = 1, \dots, n. \end{aligned} \tag{10.6}$$

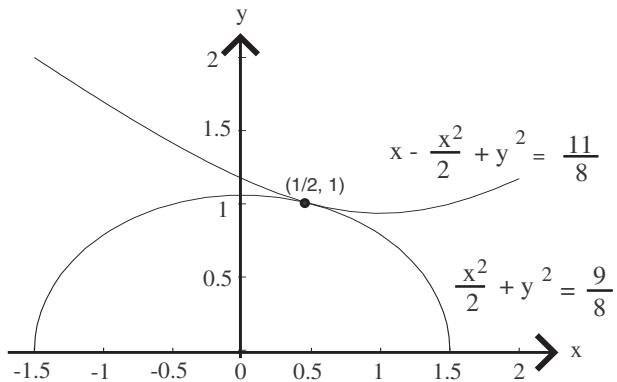


Figura 10.13: Solución al problema de optimización 10.2.3.

Ej 10.2.4 La mayoría de los problemas de optimización que surgen en economía requieren de las condiciones de no negatividad para las variables. Si las aplicamos, por ejemplo, al problema del consumidor descrito en 10.2.2.1 se tiene que las condiciones necesarias de primer orden deben ser sustituidas por

$$u_{x_i} = \frac{\alpha_i}{x_i} \leq \lambda p_i,$$

$$\left(\frac{\alpha_i}{x_i} - \lambda p_i \right) x_i = 0,$$

para toda $i = 1, \dots, n$. Si pensamos a u_{x_i} como la valuación marginal subjetiva del bien i y a λp_i como la valuación real del mismo bien, ambas medidas en útiles, entonces estas condiciones nos dicen lo siguiente: si para un bien la valuación subjetiva es estrictamente menor que la valuación real, el óptimo es consumir cero unidades de dicho bien (en este caso se dice que tenemos una **solución de esquina**); análogamente, si en el óptimo se consume una cantidad positiva del bien, entonces la valuación subjetiva es igual a la valuación real. Notemos que implícitamente estamos pensando al multiplicador λ como el equivalente, en útiles, del valor marginal de una unidad de riqueza nominal (i.e. pesos). Esta interpretación se aclara en el ejercicio 10.2.7 de la siguiente sección. La figura 10.14 ilustra estas consideraciones. Observemos que el punto A sobre la curva de indiferencia, U_1 , es el punto óptimo sin la restricción de no negatividad; dado que ahí $x_2 < 0$, este punto no resuelve el problema de optimización. La mejor opción para el consumidor es consumir $x_2 = 0$ en el punto B , sobre la curva de indiferencia U_2 .

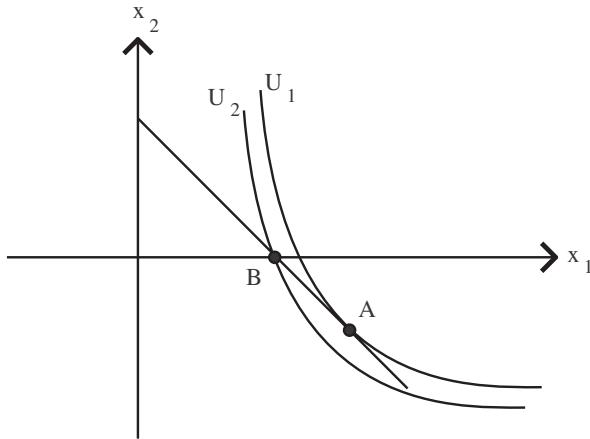


Figura 10.14: Solución de esquina.

§10.2.4 Teorema de la envolvente

En muchas situaciones, las funciones a maximizar y las restricciones dependen de parámetros $\mathbf{a} = (a_1, \dots, a_k)$. Sea $V(\mathbf{a})$ el valor del máximo del problema de optimización. Sea $x^*(\mathbf{a})$ el punto que resuelve el problema

$$\begin{aligned} \max \quad & f(\mathbf{x}, \mathbf{a}) \\ \text{sujeto a} \quad & g_1(\mathbf{x}, \mathbf{a}) \leq 0 \\ & \vdots \\ & g_n(\mathbf{x}, \mathbf{a}) \leq 0. \end{aligned}$$

Entonces se define la función valor como $V(\mathbf{a}) = f(x^*(\mathbf{a}), \mathbf{a})$. Un resultado no trivial, cuya demostración omitiremos es que, con las hipótesis dadas, los óptimos, \mathbf{x}^* , son funciones continuas y diferenciables en los parámetros.⁴ Cuando esto se cumple se tiene el siguiente teorema:

Teorema 10.2.6 (Envolvente)

Sea $\mathbf{x}^*(\mathbf{a})$ el punto que maximiza el problema de optimización con parámetros \mathbf{a} y supongamos que todas las restricciones están activas. Si $\mathcal{L}(\mathbf{x}^*(\mathbf{a}), \lambda^*(\mathbf{a}), \mathbf{a})$ es el lagrangiano en el óptimo, entonces se cumplen

a) $V(\mathbf{a}) = \mathcal{L}(\mathbf{x}^*(\mathbf{a}), \lambda^*(\mathbf{a}), \mathbf{a})$,

b) $\frac{\partial V}{\partial a_j}(\mathbf{a}) = \frac{\partial \mathcal{L}}{\partial a_j}(\mathbf{x}^*(\mathbf{a}), \lambda^*(\mathbf{a}), \mathbf{a})$.

⁴Se refiere al lector a [MCWG95] y a [Sun96] para profundizar sobre este tema.

Demostración:

La primera parte del teorema es inmediata puesto que estamos evaluando en el óptimo así que probaremos únicamente la segunda parte. De las condiciones de primer orden sabemos que

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial x_i}(\mathbf{x}^*(\mathbf{a}), \lambda^*(\mathbf{a}), \mathbf{a}) &= 0, i = 1, \dots, n, \\ \frac{\partial \mathcal{L}}{\partial \lambda_k}(\mathbf{x}^*(\mathbf{a}), \lambda^*(\mathbf{a}), \mathbf{a}) &= 0, k = 1, \dots, m.\end{aligned}$$

Por la regla de la cadena

$$\frac{\partial V}{\partial a_j} = \sum_{i=1}^n \left(\frac{\partial \mathcal{L}}{\partial x_i} \right) \left(\frac{\partial x_i^*}{\partial a_j} \right) + \sum_{k=1}^m \left(\frac{\partial \mathcal{L}}{\partial \lambda_k} \right) \left(\frac{\partial \lambda_k^*}{\partial a_j} \right) + \frac{\partial \mathcal{L}}{\partial a_j},$$

donde todas las funciones están evaluadas en $(\mathbf{x}^*(\mathbf{a}), \lambda^*(\mathbf{a}), \mathbf{a})$. Por las condiciones de primer orden, los primeros términos desaparecen y nos queda

$$\frac{\partial V}{\partial a_j} = \frac{\partial \mathcal{L}}{\partial a_j}.$$

■

La figura 10.15 ilustra el teorema de la envolvente para el caso de un parámetro a y $V(a) = \max f(x, a)$.

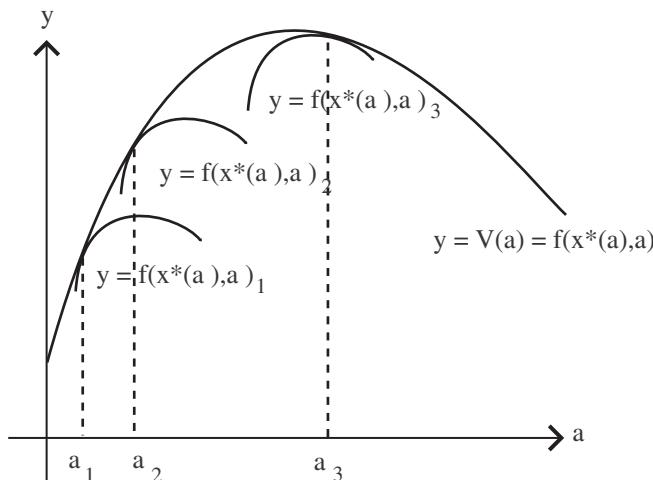


Figura 10.15: Teorema de la envolvente.

Ejemplos

Ej 10.2.5 Nos interesa aplicar el teorema de la envolvente 10.2.6 al problema de costos. En este caso,

$$\mathcal{L}(K, L, \lambda, w, r, q) = wL + rK - \lambda(AK^\alpha L^\beta - q).$$

Identificando las variables del teorema, se tiene que $\mathbf{x} = \begin{pmatrix} K \\ L \end{pmatrix}$ y $a = \begin{pmatrix} w \\ r \\ q \end{pmatrix}$. Sea (K^*, L^*, λ^*) el punto donde se cumplen las condiciones de primer orden. Entonces

$$\begin{aligned} C(w, r, q) &= \mathcal{L}(K^*, L^*, \lambda^*, w, r, q) \\ &= rK^* + wL^* \\ &= rK^*(w, r, q) + wL^*(w, r, q). \end{aligned}$$

Por el teorema de la envolvente,

$$\begin{aligned} \frac{\partial C}{\partial w} &= \frac{\partial \mathcal{L}}{\partial w} = L^* \\ \frac{\partial C}{\partial r} &= \frac{\partial \mathcal{L}}{\partial r} = K^* \\ \frac{\partial C}{\partial q} &= \frac{\partial \mathcal{L}}{\partial q} = \lambda^* \end{aligned}$$

Ej 10.2.6 (Lema de Shepard) Sea $f : \mathbb{R}_{++}^n \rightarrow \mathbb{R}$ una función de producción (f es homogénea de grado menor que 1, continua y cuasi cóncava) y $x^*(\mathbf{w}, q)$ una solución al problema

$$\begin{aligned} \min \quad & \mathbf{w}^T \mathbf{x} \\ \text{sujeto a} \quad & f(\mathbf{x}) = q. \end{aligned}$$

Sea $C(\mathbf{w}, q) = \mathbf{w}^T \mathbf{x}^*(\mathbf{w}, q)$ el costo mínimo; entonces se cumple

$$\frac{\partial C}{\partial w_j} = x_j^*(\mathbf{w}, q) \quad \forall j = 1, \dots, n.$$

Usaremos el teorema de la envolvente para probar esta igualdad. En este caso

$$\mathcal{L}(\mathbf{x}, \mathbf{w}, q, \lambda) = \mathbf{w}^T \mathbf{x} - \lambda(f(\mathbf{x}) - q).$$

Sabemos que si $\mathbf{x}^*(\mathbf{w}, q)$ es la solución, entonces

$$C(\mathbf{w}, q) = \mathcal{L}(\mathbf{x}^*(\mathbf{w}, q), \mathbf{w}, q, \lambda^*(\mathbf{w}, q)).$$

Por el teorema de la envolvente, $\frac{\partial C}{\partial w_j} = \frac{\partial \mathcal{L}}{\partial w_j}(\mathbf{x}^*(\mathbf{w}, q), \mathbf{w}, q, \lambda^*(\mathbf{w}, q)) = x_j^*(\mathbf{w}, q)$.

Ej 10.2.7 (Identidad de Roy) Sea $\mathbb{R}_+^n = \{(x_1, x_2, \dots, x_n) \in \mathbb{R}^n \mid x_i \geq 0 \ \forall i\}$. Supongamos que un consumidor tiene una función de utilidad $U(\mathbf{x})$, donde $\mathbf{x} \in \mathbb{R}_+^n$. El consumidor tiene la siguiente restricción presupuestal:

$$\mathbf{p}^T \mathbf{x} = m,$$

donde \mathbf{p} es un vector de precios y m es el ingreso o presupuesto. Si suponemos que U es cuasi cóncava, el siguiente problema tiene solución:

$$\begin{aligned} \max \quad & U(\mathbf{x}) \\ \text{sujeto a} \quad & \mathbf{p}^T \mathbf{x} = m. \end{aligned}$$

Supongamos que U es diferenciable y que el problema de maximización tiene una única solución $\mathbf{x}^*(\mathbf{p}, m)$ (**demandas marshalianas o de Marshall**). La utilidad máxima la denotaremos $V(\mathbf{p}, m)$:

$$V(\mathbf{p}, m) = U(\mathbf{x}^*(\mathbf{p}, m)).$$

Esta función se conoce como la **función de utilidad máxima** o función de **utilidad indirecta**. Bajo las condiciones anteriores, se tiene el siguiente resultado conocido como la **identidad de Roy**:

$$x_j^*(\mathbf{p}, m) = \frac{-\frac{\partial V}{\partial p_j}}{\frac{\partial V}{\partial m}}.$$

Usaremos el teorema de la envolvente para probar esta identidad. El lagrangiano, en este caso, es

$$\mathcal{L}(\mathbf{x}, \lambda, \mathbf{p}, m) = U(\mathbf{x}) - \lambda (\mathbf{p}^T \mathbf{x} - m).$$

Sabemos que existe $\lambda^*(\mathbf{p}, m)$ tal que

$$\mathcal{L}(\mathbf{x}^*(\mathbf{p}, m), \lambda^*(\mathbf{p}, m), \mathbf{p}, m) = V(\mathbf{p}, m)$$

y que $\mathbf{x}^*(\mathbf{p}, m)$ y $\lambda^*(\mathbf{p}, m)$ cumplen las condiciones de primer orden. El teorema de la envolvente dice que

$$\begin{aligned} \frac{\partial V}{\partial p_j} &= \frac{\partial \mathcal{L}}{\partial p_j}(\mathbf{x}^*(\mathbf{p}, m), \lambda^*(\mathbf{p}, m), \mathbf{p}, m) \\ &= -\lambda^*(\mathbf{p}, m) x_j^*(\mathbf{p}, m). \end{aligned}$$

Por otro lado

$$\frac{\partial V}{\partial m} = \frac{\partial \mathcal{L}}{\partial m}(\mathbf{x}^*(\mathbf{p}, m), \lambda^*(\mathbf{p}, m), \mathbf{p}, m) = \lambda^*(\mathbf{p}, m).$$

Esta última relación dice que, en el óptimo, el multiplicador λ representa el valor (en útiles) de la unidad marginal de riqueza. Podemos concluir entonces que $x_j^* = -\frac{\partial V}{\partial p_j} / \frac{\partial V}{\partial m}$.



Ejercicios

▷ **10.1** Demostrar los incisos **b** y **c** de la proposición 10.1.2.

▷ **10.2** Demostrar la proposición 10.1.7.

▷ **10.3** Sea $f : (a, b) \rightarrow \mathbb{R}$ una función cóncava (convexa).

- a) Demostrar que si $a < z_1 < z_2 < z_3 < b$ entonces,

$$\frac{f(z_2) - f(z_1)}{z_2 - z_1} \leq \frac{f(z_3) - f(z_2)}{z_3 - z_2}.$$

- b) Sea $x \in (a, b)$. Explicar porqué se pueden escoger números r, s, t, u tales que

$$a < r < s < x < t < u < b.$$

- c) Si x, r, s, t y u están dados como en el inciso anterior y y es tal que $x < y < t$ entonces,

$$C_1 \leq \frac{f(y) - f(x)}{y - x} \leq C_2,$$

donde $C_1 = \frac{f(s) - f(r)}{s - r}$ y $C_2 = \frac{f(u) - f(t)}{u - t}$.

- d) Concluir, del inciso anterior, que se tiene el siguiente límite:

$$\lim_{y \rightarrow x^+} f(y) = f(x).$$

Del mismo modo se puede demostrar que $\lim_{y \rightarrow x^-} f(y) = f(x)$ y, por lo tanto, f es continua en x . Dado que x es cualquier punto en (a, b) , f es continua en todo el intervalo.

▷ **10.4** Demostrar el teorema 10.1.13 (sugerencia: utilizar el desarrollo de Taylor de segundo orden junto con el teorema 10.1.8).

▷ **10.5** Probar las siguientes afirmaciones:

- a) $f(x, y) = xy$ con $x, y \geq 0$, es cuasicónica.

- b) $f(x, y) = (x + y)^2$ es convexa.

- c) $f(x, y) = x^2 + y^2$ es estrictamente convexa

▷ **10.6** Demostrar el teorema 10.2.2 para el caso $n = 2, m = 1$. (Sugerencia: $\nabla g(x^*, y^*) \neq 0 \Rightarrow g_x(x^*, y^*) \neq 0$ o bien $g_y(x^*, y^*) \neq 0$. Sin pérdida de generalidad suponer que $g_x(x^*, y^*) \neq 0$, por lo que se puede aplicar el teorema de la función implícita alrededor de (x^*, y^*) para obtener $x = x(y)$ y $\frac{dx}{dy} = -\frac{g_y}{g_x}$. Considerar $f(y) = f(x(y), y)$ y resolver el problema como un problema de optimización sin restricciones.)

▷ **10.7** Considerar la función de costos $C(\mathbf{w}, q)$ definida en el ejemplo 10.2.6. Probar que se cumplen las siguientes propiedades de la función de costos:

- a) Si $w^1 > w^2$, entonces $C(w^1, q) > C(w^2, q)$. (Aquí $w^j = (w_1^j, \dots, w_k^j)$, $j = 1, 2$, y $w^1 > w^2$ significa que $w_i^1 > w_i^2$, $i = 1, \dots, k$.)

- b) Si $t > 0$, entonces $C(t\mathbf{w}, q) = tC(\mathbf{w}, q)$. es decir, C es homogénea de grado 1 en \mathbf{w} .
- c) C es cóncava en \mathbf{w} .

▷ **10.8** Resolver los siguientes problemas de optimización:

- a) $\max(50x^{\frac{1}{2}}y^2)$ sujeto a $x + y = 80000$.

- b) Encontrar la distancia mínima de la elipse $x^2 + xy + y^2 = 3.3$ al origen.

- c) $\max(\ln x + \ln(y + 5))$ sujeto a $x + y \leq 4$, $x > 0, y \geq 0$.

- d) $\min(x^2 + y^2)$ sujeto a $xy \geq 25, x, y \geq 0$.

- e) $\max(xyz)$ sujeto a $x + y + z \leq 4$, $x \leq 1$, $y \geq 2$, $x, y, z \geq 0$.

- ▷ **10.9** Considera el siguiente problema de maximización de utilidad de un individuo:

$$\max\left(\frac{1}{3} \ln x + \frac{1}{3} \ln y\right)$$

sujeto a

$$3x + y \leq A,$$

$$x + y \leq 40,$$

$$x, y \geq 0,$$

donde $A \in (40, 120)$ representa el ingreso.

- a) Escribir las condiciones de Kuhn-Tucker para este problema. ¿Por qué se debe restringir el ingreso al intervalo $(40, 120)$?
- b) Dar las soluciones para los siguientes casos y representar gráficamente la solución para cada uno:
 - i) $A \in (40, 60)$,
 - ii) $A \in [60, 80)$,
 - iii) $A \in [80, 120)$.

- ▷ **10.10** Sea $f : \mathbb{R}_+^n \rightarrow \mathbb{R}$ una función de producción. Consideremos el problema de maximizar ganancias dado un nivel de producción, q , si el precio del producto es p y los precios de los insumos $\mathbf{x} = (x_1, \dots, x_n)$ están representados por un vector $w = (w_1, \dots, w_n)$:

$$\begin{aligned} \max & \quad pq - \mathbf{w}^T \mathbf{x} \\ \text{sujeto a} & \quad f(\mathbf{x}) = q \\ & \quad x_i \geq 0, i = 1, \dots, n. \end{aligned}$$

Supongamos que $\mathbf{x}^*(\mathbf{w}, p)$ es la solución del problema y el nivel de producción correspondiente está dado por $q^*(\mathbf{w}, p) = f(\mathbf{x}^*(\mathbf{w}, p))$. Entonces definimos la función de máxima ganancia como:

$$\Pi(\mathbf{w}, p) = pq^*(\mathbf{w}, p) - \mathbf{w}^T \mathbf{x}^*(\mathbf{w}, p).$$

Probar que se cumplen,

$$\frac{\partial \Pi}{\partial p}(\mathbf{w}, p) = q^*(\mathbf{w}, p),$$

$$\frac{\partial \Pi}{\partial w_j}(\mathbf{w}, p) = -x_j^*(\mathbf{w}, p) \quad \forall j = 1, \dots, n.$$

A este resultado se lo conoce como **lema de Hotelling**.

- ▷ **10.11** Considerar la misma notación que en el ejemplo 10.2.7 y el problema relacionado de minimizar el gasto sujeto a un nivel fijo de utilidad \bar{U} .

$$\begin{aligned} \min & \quad \mathbf{p}^T \mathbf{x} \\ \text{sujeto a} & \quad U(\mathbf{x}) = \bar{U}. \end{aligned}$$

Sea $\mathbf{x}^h(\mathbf{p}, \bar{U})$ la solución del problema de minimización conocida como **demandas hicksiana o de Hicks**. A la función definida por $E(\mathbf{p}, \bar{U}) = \mathbf{p}^T \mathbf{x}^h(\mathbf{p}, \bar{U})$ se la llama **función de gasto**. Probar que se cumple

$$x_j^h(p, \bar{U}) = \frac{\partial E}{\partial p_j}(\mathbf{p}, \bar{U}).$$

- ▷ **10.12** Sean \mathbf{x}^* y \mathbf{x}^h las demandas de Marshall y de Hicks, y V y E las funciones de utilidad máxima y de gasto. Probar que se cumplen las siguientes afirmaciones:

$$\mathbf{x}^h(\mathbf{p}, V(\mathbf{p}, m)) = \mathbf{x}^*(\mathbf{p}, m), \quad (\text{D1})$$

$$\mathbf{x}^*(\mathbf{p}, E(\mathbf{p}, \bar{U})) = \mathbf{x}^h(\mathbf{p}, \bar{U}), \quad (\text{D2})$$

$$V(\mathbf{p}, E(\mathbf{p}, \bar{U})) = \bar{U}, \quad (\text{D3})$$

$$E(\mathbf{p}, V(\mathbf{p}, m)) = m. \quad (\text{D4})$$

- ▷ **10.13** Sean $\alpha + \beta = 1$ y sea $U(x, y) = \alpha \log x + \beta \log y$.

- a) Obtener las demandas de Marshall resolviendo el problema de maximización

$$\begin{aligned} \max & \quad U(x, y) \\ \text{sujeto a} & \quad p_1 x + p_2 y = m. \end{aligned}$$

- b) Obtener la función de utilidad máxima, $V(\mathbf{p}, m) = U(x^*(\mathbf{p}, m), y^*(\mathbf{p}, m))$.

- c) Obtener la función de gasto, $E(\mathbf{p}, \bar{U})$.
- d) Obtener las demandas de Hicks, $x^h(\mathbf{p}, \bar{U})$ y $y^h(\mathbf{p}, \bar{U})$.
- ▷ **10.14** Sea $r < 0$ y $U(x, y) = (x^r + y^r)^{\frac{1}{r}}$.
- Obtener las demandas de Marshall.
 - Obtener la función de utilidad máxima.
 - Obtener la función de gasto.
 - Obtener las demandas de Hicks.
- ▷ **10.15** Probar la siguiente relación entre las demandas, en donde $j, k = 1, \dots, n$:
- $$\frac{\partial x_k^*(\mathbf{p}, m)}{\partial p_j} = \underbrace{\frac{\partial x_k^h(\mathbf{p}, V(\mathbf{p}, m)) - x_j^*(\mathbf{p}, m)}{\partial p_j}}_{\text{Efecto sustitución}} - \underbrace{x_j^*(\mathbf{p}, m) \frac{\partial x_k^*(\mathbf{p}, m)}{\partial m}}_{\text{Efecto ingreso}}$$

A esta relación se la conoce como **ecuación de Slutsky**. (Sugerencia: considerar (D1) o (D2) del ejercicio 10.12.)

Parte V

Optimización Dinámica

CAPÍTULO 11

Introducción al cálculo en variaciones

§11.1 Preliminares

El problema de optimización típico, en un entorno estático, está dado como

$$\max f(\mathbf{x})$$

sujeto a un conjunto de condiciones de la forma

$$g_1(\mathbf{x}) \leq 0,$$

⋮

$$g_m(\mathbf{x}) \leq 0,$$

en donde $\mathbf{x} \in \mathbb{R}^n$ y f y g_i son funciones “bien portadas” con valores en \mathbb{R} . Con frecuencia, se incluye también la restricción de no negatividad $\mathbf{x} \geq \mathbf{0}$. Si todo funciona adecuadamente, se encuentra $\mathbf{x}^* = (x_1^*, \dots, x_n^*)$ que resuelve el problema, es decir, hay un punto óptimo en \mathbb{R}^n . Como ejemplo podemos tener que \mathbf{x} es una canasta de n bienes, f una función de utilidad, una restricción presupuestal $g(\mathbf{x}) \leq 0$ y la condición de no negatividad $\mathbf{x} \geq \mathbf{0}$. Así el problema consiste en escoger la canasta óptima que maximiza la utilidad satisfaciendo la restricción presupuestal.

Si los individuos sólo existieran durante un instante en el tiempo no habría nada más que hacer. Pero los agentes económicos existen a lo largo de un periodo de tiempo y en cada instante deben resolver éste o algún otro problema de optimización. La solución no es únicamente una canasta de consumo para el día de hoy, sino toda una trayectoria de consumo a lo largo de una vida: $\mathbf{x}^*(t) = (x_1^*(t), \dots, x_n^*(t))$. Podría pensarse que si un problema satisface las condiciones de optimalidad estática en cada instante entonces tendría que ser óptimo desde el punto de vista dinámico; sin embargo, éste no es el caso: existen soluciones que son óptimas desde el punto de vista estático cuya trayectoria intertemporal no es eficiente.

Es inevitable, entonces, el estudio de técnicas de optimización dinámica para la resolución de un gran número de problemas en economía.

Un ejemplo sencillo es aquel en el cual sólo hay un bien en la economía, el cual puede tomar la forma de capital para la producción o de consumible, y una sola familia (pensemos en la “familia Robinson”, que vive en una isla desierta), que lleva a cabo la producción, la inversión y el consumo. Sea $c(t)$ el consumo de la familia en el tiempo t y supongamos que los Robinson poseen una función de utilidad doblemente diferenciable, $u(c(t))$, con $u' > 0$ y $u'' < 0$. La producción se lleva a cabo por medio de una función f cuyo único insumo es el capital, $k(t)$. Esta función es doblemente diferenciable, estrictamente cóncava y cumple con $f(0) = 0$. Al quedar varados en la isla, la familia cuenta con una cantidad inicial de capital k_0 que se deprecia a una tasa constante, δ . Si la familia vive durante un intervalo de tiempo $[0, T]$ y no descuenta el futuro, su problema de optimización se puede plantear como sigue:

$$\max \int_0^T u(c(t))dt$$

sujeto a

$$\begin{aligned} f(k(t)) &= c(t) + \dot{k}(t) + \delta k(t), \\ k(0) &= k_0. \end{aligned}$$

La integral representa la utilidad acumulada a lo largo del periodo de tiempo dado; la restricción presupuestal simplemente nos dice que para todo t se debe cumplir que la producción es igual al consumo más la inversión bruta dada por $\dot{k}(t) + \delta k(t)$; la condición inicial es que hay una cantidad inicial de capital especificada. Si la familia Robinson descuenta el futuro a una tasa ρ (tasa subjetiva de descuento), entonces el problema cambia ligeramente. Si expresamos la utilidad de cada periodo en unidades de utilidad del periodo inicial, $t = 0$, entonces la integral representa el valor presente de la utilidad acumulada a lo largo del periodo de tiempo y el problema se reescribe como

$$\max \int_0^T u(c(t))e^{-\rho t}dt$$

sujeto a

$$\begin{aligned} f(k(t)) &= c(t) + \dot{k}(t) + \delta k(t), \\ k(0) &= k_0. \end{aligned}$$

El término $u(c(t))e^{-\rho t}$ representa la utilidad que proporciona el consumo en el tiempo t , descontada a $t = 0$ o el valor presente de dicha utilidad. La solución al problema, en ambos casos, es un par de trayectorias óptimas para el consumo y el capital, $c^*(t)$ y $k^*(t)$, es decir, se encuentran funciones tales que la integral dada tome su valor máximo. El problema más antiguo de este tipo se remonta al año 850 aC, cuando, según cuenta Virgilio, la reina Dido de Cartago preguntó por la forma geométrica que debe

tener un terreno con un perímetro dado de manera que contenga la mayor cantidad de tierra posible. Este problema se conoce como problema isoperímetrónico y regresaremos a él más adelante.

Con el objeto de estudiar este tipo de problemas de optimización, primero se verán los fundamentos del cálculo en variaciones y posteriormente analizaremos los principios de la teoría de control.

El problema más sencillo del cálculo en variaciones es aquel que nos pide encontrar la curva de longitud mínima que une dos puntos dados \mathbf{A} y \mathbf{B} en \mathbb{R}^2 ; es decir, se busca la curva (diferenciable) $x : [a, b] \rightarrow \mathbb{R}$ tal que $(a, x(a)) = \mathbf{A}$ y $(b, x(b)) = \mathbf{B}$, y cuya longitud, dada por $\int_a^b \sqrt{1 + x'(t)^2} dt$ sea mínima. Por supuesto, todos sabemos que la respuesta es el segmento de recta que une los puntos \mathbf{A} y \mathbf{B} ; sin embargo, lo que nos importa aquí es la estructura del problema, que puede resumirse como sigue: encontrar la función que optimiza una integral dadas ciertas condiciones iniciales y finales.

Así como el cálculo tiene por objeto de estudio a las funciones, definidas sobre subconjuntos de \mathbb{R}^n , el objeto de estudio del cálculo en variaciones son las funciones definidas sobre subconjuntos de espacios vectoriales reales más generales, concretamente espacios de funciones.

Definición 11.1.1 Sea V un espacio vectorial sobre \mathbb{R} . Una **funcional** (lineal) F es una función (lineal) $F : V \rightarrow \mathbb{R}$. En particular, V puede ser un espacio de funciones.

Ejemplos

Ej 11.1.1 Sea $V = \mathbb{R}^n$. Cualquier función $F : \mathbb{R}^n \rightarrow \mathbb{R}$ se puede pensar como una funcional.

Ej 11.1.2 Sea $V = \{x : [a, b] \rightarrow \mathbb{R} \mid x \text{ es continua}\}$. Se puede verificar fácilmente que éste es un espacio vectorial sobre \mathbb{R} , con la suma de funciones y el producto por escalares definidos de la manera usual. Podemos definir la siguiente funcional:

$$J : V \rightarrow \mathbb{R} \text{ tal que } J[x] = \sup_{t \in [a, b]} |x(t)|.$$

La notación $J[x]$ es común para denotar a las funcionales.

Ej 11.1.3 Tomando V igual que en el ejemplo anterior podemos definir

$$J : V \rightarrow \mathbb{R} \text{ tal que } J[x] = \int_a^b x(t) dt.$$

Ej 11.1.4 Si $V = \{x : [a, b] \rightarrow \mathbb{R} \mid x \text{ es diferenciable en } (a, b)\}$ y $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ es una función diferenciable, entonces podemos definir $J : V \rightarrow \mathbb{R}$ como la siguiente funcional:

$$J[x] = \int_a^b f(x, \dot{x}, t) dt.$$

Esta funcional va a ser importante en el desarrollo posterior.



En los cursos de cálculo se optimizan funciones $f : V \rightarrow \mathbb{R}$ sujetas o no a restricciones sobre sus dominios, en donde el espacio vectorial V es el espacio de dimensión finita \mathbb{R}^n . La diferencia ahora es que queremos sustituir a \mathbb{R}^n por un espacio vectorial V de dimensión infinita, como lo son los espacios de funciones. El cálculo en variaciones se encarga de resolver este tipo de problemas de optimización, concretamente, de la optimización de funcionales. Para hacer la exposición lo más accesible posible la desarrollaremos de manera semejante al caso conocido en cálculo. Lo primero que se tiene que hacer es definir el concepto de diferenciabilidad para el caso de una funcional.

Dada una función $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$, donde Ω es un conjunto abierto (véase el apéndice D), podemos pensar que es diferenciable en un punto \mathbf{x}_0 si “alrededor” de éste la función se puede aproximar por una función lineal: la recta tangente en el punto si $n = 1$, el plano tangente si $n = 2$, etc. En general se puede dar la siguiente definición.

Definición 11.1.2 *Dada una función $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$ con Ω abierto en \mathbb{R}^n , decimos que f es diferenciable en $\mathbf{x}_0 \in \Omega$ si existe una función lineal continua $Df(\mathbf{x}_0)$ tal que*

$$f(\mathbf{x}_0 + \mathbf{h}) - f(\mathbf{x}_0) = Df(\mathbf{x}_0)(\mathbf{h}) + o(\|\mathbf{h}\|),$$

en donde $\mathbf{h} = (h_1, \dots, h_n)$ y $o(\|\mathbf{h}\|)$ es un residuo¹ que satisface $\lim_{\mathbf{h} \rightarrow \mathbf{0}} \frac{o(\|\mathbf{h}\|)}{\|\mathbf{h}\|} = 0$ donde $\|\mathbf{h}\| = \sqrt{\sum_{i=1}^n h_i^2}$ es la norma euclídea (véase el apéndice D).

Dado que $Df(\mathbf{x}_0)$ es una función lineal de un subconjunto de \mathbb{R}^n a \mathbb{R} , se puede expresar como una matriz de $1 \times n$, es decir, un vector renglón. De aquí que se tengan las notaciones alternativas $Df(\mathbf{x}_0)(\mathbf{h}) = Df(\mathbf{x}_0)\mathbf{h}^T = Df(\mathbf{x}_0) \cdot \mathbf{h}$.

Definición 11.1.3 *Análogamente, decimos que la función f es doblemente diferenciable si existe una función lineal $Df(\mathbf{x}_0)$ y una forma cuadrática $D^2f(\mathbf{x}_0)$, inducida por una matriz simétrica H (es decir, $D^2f(\mathbf{x}_0)(\mathbf{h}, \mathbf{h}) = \mathbf{h}H\mathbf{h}^T$), tal que*

$$f(\mathbf{x}_0 + \mathbf{h}) - f(\mathbf{x}_0) = Df(\mathbf{x}_0)(\mathbf{h}) + \frac{1}{2}\mathbf{h}H\mathbf{h}^T + o(\|\mathbf{h}\|^2),$$

en donde el residuo $o(\|\mathbf{h}\|^2)$ satisface $\lim_{\mathbf{h} \rightarrow \mathbf{0}} \frac{o(\|\mathbf{h}\|^2)}{\|\mathbf{h}\|^2} = 0$.

Observamos que puede definirse la diferenciabilidad de f en \mathbf{x}_0 como la existencia de una función lineal continua $Df(\mathbf{x}_0)$ tal que se cumple

$$\lim_{\mathbf{h} \rightarrow \mathbf{0}} \frac{|f(\mathbf{x}_0 + \mathbf{h}) - f(\mathbf{x}_0) - Df(\mathbf{x}_0)(\mathbf{h})|}{\|\mathbf{h}\|} = 0.$$

Esta forma de la definición se extiende para campos vectoriales en el apéndice B.

¹ La “pequeña o” de Landau, denotada por o , se utiliza para denotar cantidades “infinitesimales” de orden mayor a a^k . Concretamente, dado $a > 0$, $o(a^k)$ es tal que $\lim_{a \rightarrow 0} \frac{o(a^k)}{a^k} = 0$, es decir, $o(a^k)$ se acerca al 0 más rápido que $\frac{1}{a^k}$.

Ejemplos

Ej 11.1.5 Si $n = 1$, la función lineal está dada por

$$Df(x_0)(h) = f'(x_0)h$$

y la matriz H es de 1×1 dada por $f''(x_0)$.

Ej 11.1.6 Si $n = 2$ y $\mathbf{h} = (h_1, h_2)$, se tiene que

$$Df(\mathbf{x}_0)(\mathbf{h}) = (f_1(\mathbf{x}_0), f_2(\mathbf{x}_0)) \cdot (h_1, h_2),$$

en donde f_1 y f_2 denotan las derivadas parciales. La matriz simétrica H es la matriz hessiana

$$\begin{pmatrix} f_{11}(\mathbf{x}_0) & f_{12}(\mathbf{x}_0) \\ f_{12}(\mathbf{x}_0) & f_{22}(\mathbf{x}_0) \end{pmatrix}$$

y así sucesivamente. El lector observador identificará los términos lineal y cuadrático del desarrollo de Taylor de la función f .



Lo único que necesitamos para extender este concepto a cualquier funcional es aclarar qué queremos decir con una funcional “continua”. Para esto es necesario tener el concepto de “tamaño” o norma para cualquier espacio vectorial real, que generalice naturalmente el concepto de norma euclídea en \mathbb{R}^n .

Definición 11.1.4 Decimos que un espacio vectorial V (sobre \mathbb{R}) es **normado**, si existe una función (norma) $\|\cdot\| : V \rightarrow \mathbb{R}^+ \cup \{0\}$ tal que se satisfacen

- i) Dado $\mathbf{v} \in V$, $\|\mathbf{v}\| = 0 \Leftrightarrow \mathbf{v} = \mathbf{0}$.
- ii) $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$ para cualesquiera \mathbf{u}, \mathbf{v} en V (desigualdad del triángulo).
- iii) $\|a\mathbf{v}\| = |a| \|\mathbf{v}\|$ para cualquier $a \in \mathbb{R}$ y $\mathbf{v} \in V$.

Ejemplos

Ej 11.1.7 Sea $V = \mathbb{R}$, entonces dado $r \in \mathbb{R}$ podemos definir $\|r\| = |r|$.

Ej 11.1.8 Sea $V = \mathbb{R}^n$, entonces dado $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ podemos definir

$$\|\mathbf{x}\| = \sqrt{\sum_{i=1}^n x_i^2}.$$

Ésta es la definición D.2.1 del apéndice D.

Ej 11.1.9 Sea $V = \mathbb{R}^n$, entonces dado $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ podemos definir

$$\|\mathbf{x}\| = \max_i \{|x_i|\}.$$

Ej 11.1.10 Nótese que la funcional $J[x] = \sup_{t \in [a,b]} |x(t)|$ del ejemplo 11.1.2 es una norma. A esta norma se la conoce como **norma del supremo**.



Ahora podemos definir el concepto de continuidad para una funcional preservando la idea clásica de continuidad.

Definición 11.1.5 Sea V un espacio vectorial normado. Una funcional $J : V \rightarrow \mathbb{R}$, es **continua en x_0** si para toda $\varepsilon > 0$ existe $\delta > 0$ tal que

$$\|x - x_0\| < \delta \Rightarrow |J[x] - J[x_0]| < \varepsilon.$$

En particular decimos simplemente que J es continua si es continua en todo V .

Notemos que la norma es necesaria para tener el concepto de “distancia” entre x y x_0 , expresada por $\|x - x_0\|$. En el apéndice D se definió el concepto de “bola” para espacios \mathbb{R}^n . La siguiente definición extiende esta idea a espacios más generales.

Definición 11.1.6 Sea V un espacio vectorial normado y sean $x_0 \in V$, $\varepsilon > 0$. Se define la **bola abierta de radio ε alrededor de x_0** como

$$B_\varepsilon(x_0) = \{x \in V \mid \|x - x_0\| < \varepsilon\}.$$

Estamos listos para extender el concepto de diferenciabilidad a funcionales.

Definición 11.1.7 Sea V un espacio vectorial normado. Una funcional $J : V \rightarrow \mathbb{R}$ es **diferenciable en x_0** si existe una funcional lineal continua $DJ[x_0]$ tal que para todo $h \in V$ se cumple

$$J[x_0 + h] - J[x_0] = DJ[x_0](h) + o(\|h\|),$$

en donde $o(\|h\|)$ es un residuo que satisface $\lim_{h \rightarrow 0} \frac{o(\|h\|)}{\|h\|} = 0$. Se dice que J es **dblemente diferenciable en x_0** si, aparte de $DJ[x_0]$, existe una funcional bilineal, simétrica² y continua

$$D^2J[x_0] : V \times V \rightarrow \mathbb{R}$$

tal que

$$J[x_0 + h] - J[x_0] = DJ[x_0](h) + \frac{1}{2}D^2J[x_0](h, h) + o(\|h\|^2),$$

en donde el residuo $o(\|h\|^2)$ satisface $\lim_{h \rightarrow 0} \frac{o(\|h\|^2)}{\|h\|^2} = 0$. Es común llamar a $DJ[x_0]$ y a $D^2J[x_0]$ la **primera y segunda derivadas de Fréchet**, respectivamente; también a $DJ[x_0](h)$ y $D^2J[x_0](h, h)$ se las conoce como **primera y segunda variaciones**.

Podemos ahora precisar lo que se quiere decir con puntos máximos (mínimos); la definición coincide, como era de esperarse, con el concepto usual.

²Bilineal significa que es lineal en cada variable y simétrica que $D^2J[x_0](h, k) = D^2J[x_0](k, h)$.

Definición 11.1.8 Sea V un espacio vectorial normado. Una funcional $J : V \rightarrow \mathbb{R}$ tiene un **máximo (mínimo) local** en $x^* \in V$ si existe una bola abierta de radio ε alrededor de x^* tal que

$$J[x^*] \geq J[x] \quad (J[x^*] \leq J[x]) \text{ para toda } x \in B_\varepsilon(x^*).$$

Análogamente tiene un **máximo (mínimo) global** si

$$J[x^*] \geq J[x] \quad (J[x^*] \leq J[x]) \text{ para toda } x \in V.$$

En general nos referimos a **extremos locales o extremos globales**.

La siguiente proposición es la extensión del resultado aprendido en cálculo.

Proposición 11.1.9 Sea V un espacio vectorial normado. Una condición necesaria para que la funcional diferenciable $J : V \rightarrow \mathbb{R}$ tenga un extremo local (o global) en $x^* \in V$ es que $DJ[x^*] = 0$.

Demostración

Defínase la función $\phi(\omega) = J[x^* + \omega h]$ con $\omega \in \mathbb{R}$ y $h \in V$. Por construcción, ϕ tiene un extremo en $\omega = 0$. Dado que es simplemente una función de \mathbb{R} en \mathbb{R} se tiene que $\phi'(0) = 0$. Utilizando la regla de la cadena podemos reescribir esta condición como $DJ[x^*] \cdot h = 0$, por lo que se debe tener $DJ[x^*] = 0$.

■

Como siguiente paso nos vamos a concentrar en cierto tipo de funcionales y veremos cómo se puede escribir la condición dada en la proposición anterior.

§11.2 Ecuación de Euler

En cálculo, es relativamente sencillo calcular la derivada de *cualquier* función diferenciable. Lo único que necesitamos es conocer las derivadas de las funciones comunes y las reglas para la derivada de la suma, producto y cociente, y la regla de la cadena. Cuando tenemos funcionales con dominio en algún espacio de funciones tenemos el concepto de derivada pero no existen reglas generales para calcularla. En general, se tiene que utilizar directamente la definición de derivada y calcular la funcional lineal correspondiente. A continuación realizamos este proceso para cierto tipo de funcional que nos será de gran utilidad.

Recordemos que si f es una función definida en $[a, b]$, se dice que $f \in \mathcal{C}^n[a, b]$ si f posee n derivadas continuas en $[a, b]$. Consideremos el espacio vectorial

$$V = \{x : [a, b] \rightarrow \mathbb{R} \mid x \in \mathcal{C}^2[a, b]\}$$

con la norma del supremo (véase el ejemplo 11.1.10). Sean $D = \{x \in V \mid x(a) = A, x(b) = B\} \subset V$ y $f : \Omega \rightarrow \mathbb{R}$, $f \in \mathcal{C}^2(\Omega)$, una función cuyo dominio $\Omega \subset \mathbb{R}^3$ contiene al conjunto

$$\{(x(t), \dot{x}(t), t) \in \mathbb{R}^3 \mid t \in [a, b], x \in D\}.$$

Definimos entonces la siguiente funcional, $J : D \rightarrow \mathbb{R}$, como

$$J[x] = \int_a^b f(x(t), \dot{x}(t), t) dt \quad (11.1)$$

y queremos calcular $DJ[x]$. Para este propósito desarrollamos la expresión

$$\begin{aligned} J[x(t) + h(t)] - J[x(t)] &= \\ \int_a^b f(x(t) + h(t), \dot{x}(t) + \dot{h}(t), t) dt - \int_a^b f(x(t), \dot{x}(t), t) dt, \end{aligned} \quad (11.2)$$

en donde $h : [a, b] \rightarrow \mathbb{R}$ es diferenciable y satisface $h(a) = h(b) = 0$. (Esta última condición sobre la función h es para garantizar que la suma $x+h$ sea un elemento de D . Notemos que $x+h$ es simplemente un desplazamiento de la función x como se muestra en la figura 11.1. Dado un punto $(x(t), \dot{x}(t), t)$,

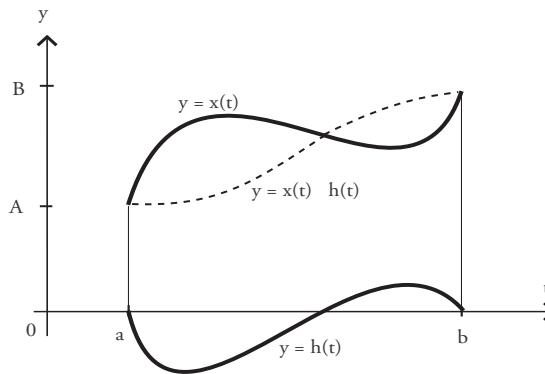


Figura 11.1: Desplazamiento de la función x .

consideraremos la expansión de Taylor de primer orden (o sea la aproximación lineal) de la función f alrededor de este punto. Si, abusando de la notación, omitimos la variable t como argumento de las funciones x y \dot{x} , se obtiene

$$\begin{aligned} f(x + h, \dot{x} + \dot{h}, t) &= \\ f(x, \dot{x}, t) + (f_x(x, \dot{x}, t), f_{\dot{x}}(x, \dot{x}, t), f_t(x, \dot{x}, t)) \cdot (h, \dot{h}, 0) + o\left(\|(h, \dot{h}, 0)\|\right), \end{aligned}$$

en donde f_x , $f_{\dot{x}}$ y f_t denotan las derivadas parciales. Sustituyendo en (11.2) se tiene que si $h \rightarrow 0$

$$\begin{aligned} J[x + h] - J[x] &\simeq \int_a^b [(f_x(x, \dot{x}, t), f_{\dot{x}}(x, \dot{x}, t), f_t(x, \dot{x}, t)) \cdot (h, \dot{h}, 0)] dt \\ &= \int_a^b (f_x(x, \dot{x}, t)h + f_{\dot{x}}(x, \dot{x}, t)\dot{h}) dt. \end{aligned}$$

En esta expresión, el lado derecho es la funcional lineal y continua que aproxima mejor a $J[x+h] - J[x]$, por lo que la definición 11.1.7 nos dice que

$$DJ[x](h) = \int_a^b (f_x(x, \dot{x}, t)h + f_{\dot{x}}(x, \dot{x}, t)\dot{h})dt. \quad (11.3)$$

Integrando el término $\int_a^b f_{\dot{x}}(x, \dot{x}, t)\dot{h}dt$ por partes, obtenemos

$$\begin{aligned} \int_a^b f_{\dot{x}}\dot{h}dt &= (hf_{\dot{x}}) \Big|_a^b - \int_a^b hdf_{\dot{x}} \\ &= - \int_a^b hdf_{\dot{x}}, \end{aligned}$$

pues $h(a) = h(b) = 0 \Rightarrow (hf_{\dot{x}}) \Big|_a^b = 0$. Dado que $-\int_a^b hdf_{\dot{x}} = -\int_a^b h \frac{d}{dt} f_{\dot{x}} dt$, sustituimos en (11.3) para obtener finalmente,

$$DJ[x](h) = \int_a^b h(f_x - \frac{d}{dt} f_{\dot{x}})dt. \quad (11.4)$$

La proposición 11.1.9 dice que la condición necesaria para que $x(t)$ sea un extremo de la funcional 11.1 es que se cumpla $DJ[x](h) = 0$ para toda h . Así, el siguiente teorema es inmediato.

Teorema 11.2.1

Sean

$$D = \{x : [a, b] \rightarrow \mathbb{R} \mid x \in C^2([a, b]); x(a) = A; x(b) = B\}$$

y $f : \Omega \rightarrow \mathbb{R}$ una función de clase C^2 , cuyo dominio $\Omega \subset \mathbb{R}^3$ contiene al conjunto

$$\{(x(t), \dot{x}(t), t) \in \mathbb{R}^3 \mid t \in [a, b], x \in D\}.$$

Una condición necesaria para que la funcional

$$J[x] = \int_a^b f(x, \dot{x}, t)dt$$

tenga un extremo en x es que se cumpla

$$f_x - \frac{d}{dt} f_{\dot{x}} = 0. \quad (11.5)$$

La condición (11.5) es conocida como la **ecuación de Euler**, cuya solución implica resolver una ecuación diferencial. De esta forma se encuentra la función x que “optimiza” la funcional $J[x]$. x está dentro del conjunto V , es decir, estamos restringidos a buscar dentro de las funciones doblemente diferenciables que empiezan y terminan en un punto dado. Esto se ilustra en la figura 11.2.

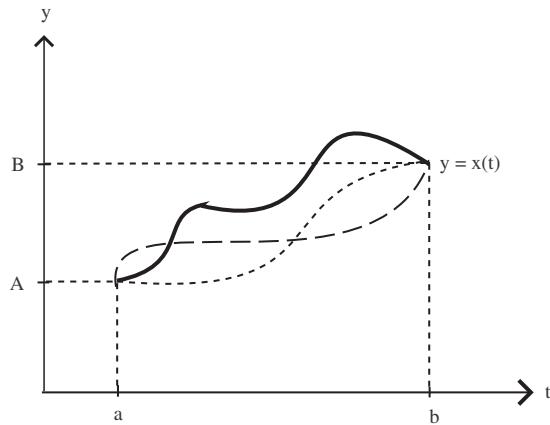


Figura 11.2: Trayectorias posibles entre los puntos (a, A) y (b, B) .

Ejemplos

Ej 11.2.1 Deseamos encontrar un extremo para la funcional

$$J[x] = \int_0^1 (5\dot{x}^2 + 2t) dt,$$

en donde $x(0) = 0$ y $x(1) = 2$. En este caso $f(x, \dot{x}, t) = 5\dot{x}^2 + 2t$ y la ecuación de Euler es

$$0 - \frac{d}{dt}(10\dot{x}) = 0,$$

que se transforma en $10\ddot{x} = 0$, o bien, $\ddot{x} = 0$. La solución se puede obtener fácilmente como $x(t) = At + B$. Las constantes se determinan con las condiciones iniciales y finales: $x(0) = 0 = B$ y $x(1) = 2 = A$, por lo que tenemos que la solución al problema es $x(t) = 2t$. Observemos que f podía haber tomado la forma $f(x, \dot{x}, t) = a\dot{x}^2 + bt$ para cualesquiera constantes a, b con $a \neq 0$.

Ej 11.2.2 Encontrar la trayectoria de producción óptima $x(t)$ en el periodo $[0, 1]$ para una empresa cuya función de costos está dada por

$$c(x, \dot{x}) = \dot{x}^2 + x^2.$$

El precio del producto se mantiene fijo e igual a \$4.00 y la tasa de descuento es igual a 0.1. La producción inicial es $x(0) = 0$ y la producción al final del periodo es $x(1) = 10$ (en miles de unidades). La producción óptima es aquella que resuelve el problema

$$\max \int_0^1 (4x - \dot{x}^2 - x^2) e^{-0.1t} dt,$$

en donde $x(0) = 0$ y $x(1) = 10$. Aquí la función $f(x, \dot{x}, t) = 4x - \dot{x}^2 - x^2$ es simplemente la función de ganancias netas de la empresa y la multiplicación por $e^{-0.1t}$ nos proporciona el valor presente de $f(x, \dot{x}, t)$. Así, el problema de la empresa es el de maximizar el valor presente de las ganancias netas acumuladas en el periodo, dadas condiciones iniciales y finales. La ecuación de Euler es

$$e^{-0.1t}(4 - 2x) - \frac{d}{dt} [(-2\dot{x})e^{-0.1t}] = 0,$$

que podemos transformar en

$$\ddot{x} - 0.1\dot{x} - x = -2.$$

La ecuación característica es

$$\lambda^2 - 0.1\lambda - 1 = 0,$$

con soluciones $\lambda_1 = -0.951$ y $\lambda_2 = 1.051$. Una solución particular es $x_p = 2$, de aquí que la solución general esté dada por

$$x(t) = 2 + Ae^{-0.951t} + Be^{1.051t}.$$

Sustituyendo las condiciones iniciales se obtienen los valores de A y B ; finalmente la trayectoria óptima de producción es

$$x(t) = 2 + 0.92e^{-0.951t} - 2.92e^{1.051t}.$$

Por el momento, suponemos que este extremo es un máximo, aunque todavía no tenemos elementos para probar que en efecto lo es.

Ej 11.2.3 Vamos a resolver uno de los problemas planteados en la introducción de este capítulo. Sean (t_1, y_1) y (t_2, y_2) dos puntos en el plano. Se quiere encontrar la curva (diferenciable) de longitud mínima entre ellos. El problema se puede plantear como

$$\min \int_{t_1}^{t_2} \sqrt{1 + \dot{x}^2} dt,$$

dado que $x(t_1) = y_1$ y $x(t_2) = y_2$, donde $x(t)$ es la curva buscada. La integral representa la longitud de la curva entre los dos puntos dados ya que $\sqrt{1 + \dot{x}^2}$ es la longitud del vector tangente a la curva en el punto $(t, x(t))$ y la integral es simplemente el límite de la aproximación de la longitud de la curva por medio de rectas secantes a ésta. La ecuación de Euler queda como

$$0 - \frac{d}{dt} \frac{\dot{x}}{\sqrt{1 + \dot{x}^2}} = 0$$

o bien,

$$\frac{\dot{x}}{\sqrt{1 + \dot{x}^2}} = K,$$

en donde K es una constante indeterminada. Elevando al cuadrado y reescribiendo llegamos a

$$\dot{x}^2 = \frac{K^2}{1 - K^2} \equiv a^2.$$

Así, finalmente se obtiene

$$\dot{x} = a,$$

La solución está dada por, $x(t) = at + b$. Las constantes se pueden obtener en términos de los puntos inicial y final para obtener la expresión

$$x(t) = \left(\frac{y_2 - y_1}{t_2 - t_1} \right) t + \left(\frac{y_1 t_2 - y_2 t_1}{t_2 - t_1} \right).$$

Al igual que en el ejemplo anterior, suponemos que este extremo es efectivamente un mínimo. Como era de esperarse, se obtuvo como respuesta la recta que pasa por los puntos dados.



§11.3 Modelo de Ramsey

Consideremos el problema planteado al principio del capítulo, en el cual la familia Robinson resuelve el siguiente problema de optimización:

$$\max \int_0^T u(c) dt$$

sujeto a

$$\begin{aligned} f(k) &= c + \dot{k} + \delta k, \\ k(0) &= k_0, \quad k(T) = k_T. \end{aligned}$$

Las funciones satisfacen $u' > 0$, $u'' < 0$, $f(0) = 0$, $f'(0) > \delta$, $f' > 0$ y $f'' < 0$. Ésta es una variante de un problema similar planteado por Ramsey en 1928 [Ram28]. A diferencia de los modelos de crecimiento de Harrod-Domar (véase la sección 3.3.1) y Solow (véase el ejemplo 2.2.3), aquí se toman en consideración fundamentos microeconómicos (la resolución del problema de maximización) para obtener la trayectoria de consumo de la familia, que no es simplemente un porcentaje fijo de la producción.

Para resolver el problema, sustituimos el consumo de la restricción presupuestal y lo reexpresamos como

$$\max \int_0^T u(f(k) - \dot{k} - \delta k) dt$$

con $k(0) = k_0$ y $k(T) = k_T$.

De esta forma, la ecuación de Euler está dada por $u_k - \frac{d}{dt}u_{\dot{k}} = 0$ y esto implica

$$\begin{aligned} u'(f' - \delta) + \frac{d}{dt}u' &= 0 \Leftrightarrow \\ \frac{\dot{u}'}{u'} &= \frac{u''\dot{c}}{u'} = \delta - f'. \end{aligned}$$

Tomemos una forma específica de la función de utilidad, por ejemplo con aversión relativa al riesgo constante, como se vio en la sección 2.3.5, de manera que $u(c) = \frac{c^{1-\theta}}{1-\theta}$, en donde θ representa el

coeficiente de aversión relativa al riesgo. Sustituyendo $u' = c^{-\theta}$ y $u'' = -\theta c^{-\theta-1}$ reescribimos la ecuación de Euler como $-\frac{\theta \dot{c}}{c} = \delta - f'$ o bien

$$\frac{\dot{c}}{c} = \frac{1}{\theta}(f' - \delta). \quad (11.6)$$

El problema de la familia Robinson queda resuelto por el sistema de ecuaciones dado por (11.6), la restricción presupuestal

$$\dot{k} = f(k) - \delta k - c \quad (11.7)$$

y las condiciones inicial y final $k(0) = k_0$ y $k(T) = k_T$.

La ecuación (11.6) se puede interpretar de la siguiente forma: $f' - \delta$ es la tasa neta de rendimiento del capital (tasa real neta), por lo que si ésta es positiva el consumo aumenta en el tiempo y si es negativa disminuye. La tasa a la cual el consumo crece o decrece es proporcional a la tasa real neta. Las ecuaciones (11.7) y (11.6) forman un sistema no lineal de ecuaciones diferenciales en k y c . Para encontrar los puntos de equilibrio se resuelve

$$\begin{aligned}\dot{c} &= 0 = \frac{c}{\theta}(f' - \delta) \\ \dot{k} &= 0 = f(k) - \delta k - c,\end{aligned}$$

y se obtienen dos puntos: $(0, 0)$ y (k^*, c^*) , tal que $f'(k^*) - \delta = 0$ y $c^* = f(k^*) - \delta k^*$. Linealizando el sistema original alrededor de (k, c) se obtiene que la matriz Jacobiana asociada al sistema lineal está dada por

$$J(k, c) = \begin{pmatrix} f'(k) - \delta & -1 \\ \left(\frac{c}{\theta}\right) f''(k) & \frac{f'(k) - \delta}{\theta} \end{pmatrix}.$$

De esta forma, la matriz que representa al sistema linealizado alrededor de $(0, 0)$ es,

$$J(0, 0) = \begin{pmatrix} f'(0) - \delta & -1 \\ 0 & \frac{f'(0) - \delta}{\theta} \end{pmatrix}.$$

Dado que $f'(0) > \delta$, tanto la traza como el determinante de esta matriz son positivos, de aquí que $(0, 0)$ sea un repulsor. Este equilibrio no es de interés ya que la única forma de llegar a él es comenzando ahí mismo, es decir, con una cantidad nula de capital y de consumo. Si consideramos ahora (k^*, c^*) , la matriz correspondiente es

$$J(k^*, c^*) = \begin{pmatrix} 0 & -1 \\ \left(\frac{c^*}{\theta}\right) f''(k^*) & 0 \end{pmatrix}.$$

El determinante de esta matriz es $\left(\frac{c^*}{\theta}\right) f''(k^*) < 0$ y por lo tanto (k^*, c^*) es un punto silla. El diagrama de fase correspondiente queda como en la figura 11.3.

Dada la condición inicial $k(0) = k_0$, la condición final $k(T) = k_T$ y un tiempo final T , existe una única trayectoria que las satisface. Sin embargo, existen una infinidad de trayectorias que nos llevan de un valor inicial del capital a un valor final, dependiendo del tiempo final. En la figura 11.4 se muestran dos trayectorias posibles con condiciones iniciales y finales $k(0) = k_0$ y $k(T_1) = k(T_2) = k_F$: La trayectoria

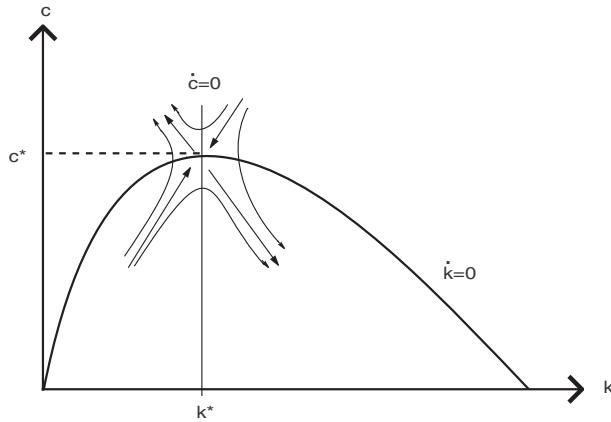


Figura 11.3: Modelo de Ramsey.

I (entre A_1 y A_2), con duración T_1 y la II (entre B_1 y B_2), con duración T_2 , respectivamente. La primera de éstas muestra el comportamiento de **trébol**, del inglés turnpike (a veces llamado comportamiento de **autopista**) en el cual la ruta entre los puntos inicial y final no es la más directa: en este caso, para llegar a la cantidad final $k_F > k_0$, primero se desacumula capital para después acumularlo. Este tipo de comportamiento es característico de horizontes de tiempo finitos.

Cuando la familia no descuenta el futuro, el estado estacionario del sistema (k^*, c^*) se encuentra en el punto máximo de la curva $\dot{k} = 0$. A este nivel de consumo se lo denomina **regla de oro del consumo** y es aquel que satisface la ecuación $c^g = f(k^g) - \delta k^g$ en donde k^g es tal que $f'(k^g) = \delta$. En este caso, los niveles de consumo y capital de la regla de oro coinciden con el estado estacionario del sistema; veremos que la situación cambia cuando la familia descuenta el futuro.³

Consideraremos el problema anterior con la variante de que ahora la familia Robinson, consciente de su incierto futuro en la isla, tiene un factor subjetivo de descuento temporal que denotamos por $\rho > 0$. El problema anterior se convierte en

$$\max \int_0^T u(c) e^{-\rho t} dt$$

sujeto a

$$\begin{aligned} f(k) &= c + \dot{k} + \delta k, \\ k(0) &= k_0, k(T) = k_T. \end{aligned}$$

Es relativamente fácil resolver la ecuación de Euler correspondiente para obtener

$$\frac{u'' \dot{c}}{u'} = \rho + \delta - f'$$

³El superíndice g utilizado para el consumo y capital de la regla de oro proviene del inglés: golden rule.

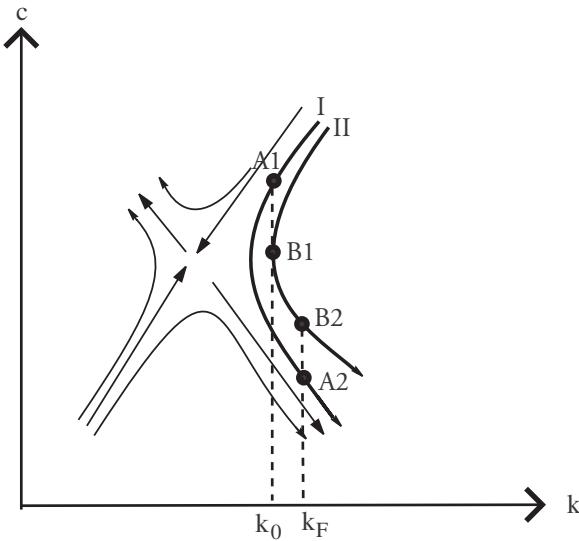


Figura 11.4: La trayectoria I ilustra el comportamiento de trébol: para pasar de un nivel de capital inicial k_0 a un nivel final $k_F > k_0$, el capital debe disminuir, inicialmente para luego aumentar.

y considerando, como antes, la función de utilidad $u(c) = \frac{c^{1-\theta}}{1-\theta}$, se tiene una condición análoga a (11.6), que en este caso es

$$\frac{\dot{c}}{c} = \frac{1}{\theta}[(f' - \delta) - \rho]. \quad (11.8)$$

La restricción presupuestal es la misma, por lo que la ecuación (11.7) sigue siendo válida. Los puntos de equilibrio se obtienen al resolver

$$\begin{aligned}\dot{k} &= 0 = f(k) - \delta k - c, \\ \dot{c} &= 0 = \frac{c}{\theta}[(f' - \delta) - \rho]\end{aligned}$$

y obtenemos, una vez más, dos puntos: $(0, 0)$ y (k^{**}, c^{**}) , en donde solo el segundo es de interés. La matriz jacobiana asociada al sistema lineal alrededor de este punto es

$$\begin{aligned}J(k^{**}, c^{**}) &= \begin{pmatrix} f'(k^{**}) - \delta & -1 \\ \left(\frac{c^{**}}{\theta}\right) f''(k^{**}) & 0 \end{pmatrix} \\ &= \begin{pmatrix} \rho & -1 \\ \left(\frac{c^{**}}{\theta}\right) f''(k^{**}) & 0 \end{pmatrix},\end{aligned} \quad (11.9)$$

cuyo determinante es negativo y por lo tanto el equilibrio es, una vez más, un punto silla. El diagrama de fase queda ahora como en la figura 11.5.

Se puede observar que el estado estacionario del sistema (k^{**}, c^{**}) satisface las ecuaciones $c^{**} = f(k^{**}) - \delta k^{**}$ y $f'(k^{**}) = \delta + \rho$. Éste corresponde a una cantidad menor tanto de capital como de

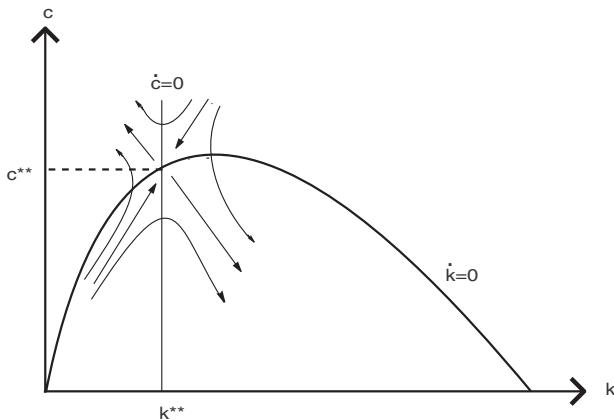


Figura 11.5: Modelo de Ramsey con descuento temporal.

consumo que el dado anteriormente por la regla de oro. La razón es que la tasa subjetiva de descuento temporal actúa como una tasa de depreciación del capital, por lo que es parte de su costo (de oportunidad) marginal. La ecuación (11.8) nos dice que la tasa a la cual la familia cambia su consumo es proporcional a la diferencia entre la tasa real neta $f' - \delta$ y la tasa subjetiva ρ . De esta forma, si la familia es más impaciente que el mercado (expresado por $\rho > f' - \delta$) prefiere consumir en el presente en vez de en el futuro, por lo que el consumo cae en el tiempo y la tasa de cambio del consumo es negativa, es decir, $\frac{\dot{c}}{c} < 0$. Análogamente si el mercado es más impaciente que la familia, esta última estará dispuesta a ceder consumo presente e invertirlo en capital, a cambio de consumo en el futuro. En este caso se tiene $\frac{\dot{c}}{c} > 0$. Posteriormente regresaremos a este modelo cuando el horizonte de tiempo es infinito.

Utilizamos ahora la matriz (11.9) con los siguientes valores de los parámetros y formas funcionales para obtener una solución explícita:

$$\begin{aligned}f(k) &= 3k^{\frac{1}{3}}, \\ \theta &= 0.5, \\ \rho &= 0.05, \\ \delta &= 0.05.\end{aligned}$$

De aquí se obtiene $k^{**} = 10\sqrt[3]{10}$ y $c^{**} = 2.5\sqrt[3]{10}$, y sustituyendo en (11.9) se obtiene la matriz

$$\begin{pmatrix} 0.05 & -1 \\ -\frac{1}{30} & 0 \end{pmatrix},$$

con valores propios $\lambda_1 = 0.2$ y $\lambda_2 = -0.16$, y vectores propios $\mathbf{v}_1 = \begin{pmatrix} 1 \\ -0.15 \end{pmatrix}$ y $\mathbf{v}_2 = \begin{pmatrix} 1 \\ 0.21 \end{pmatrix}$,

respectivamente. La solución general está dada por

$$\begin{aligned} k(t) &= Ae^{0.2t} + Be^{-0.16t} + 10\sqrt{10}, \\ c(t) &= -0.15Ae^{0.2t} + 0.21Be^{-0.16t} + 2.5\sqrt{10}. \end{aligned} \quad (11.10)$$

Si $T = \ln 2$ y las condiciones inicial y final son $k(0) = 35$ y $k(\ln 2) = 40$, entonces se obtiene que

$$\begin{aligned} A &= 21.1, \\ B &= -17.7, \\ c(0) &= 1.02, \\ c(\ln 2) &= 0.94. \end{aligned}$$

En este tipo de modelos es posible inferir que el equilibrio es un punto silla simplemente realizando el diagrama de fase, como en el ejemplo 5.6.3. En cualquiera de los dos casos expuestos arriba, el gradiente sobre la isoclina $\dot{k} = 0$ está dado por $(f'(k) - \delta, -1)$ y el gradiente sobre la isoclina $\dot{c} = 0$ es $(f''(k), 0)$. El primero apunta hacia la región donde $\dot{k} > 0$ (interior de la curva) y por lo tanto k crece y el segundo apunta hacia la región donde $\dot{c} > 0$ (a la izquierda de la recta) y por lo tanto c crece. Los diagramas de la figura 11.6 ilustran los pasos anteriores; en particular, se observa que el diagrama de flujo implica que el equilibrio distinto del origen es un punto silla.

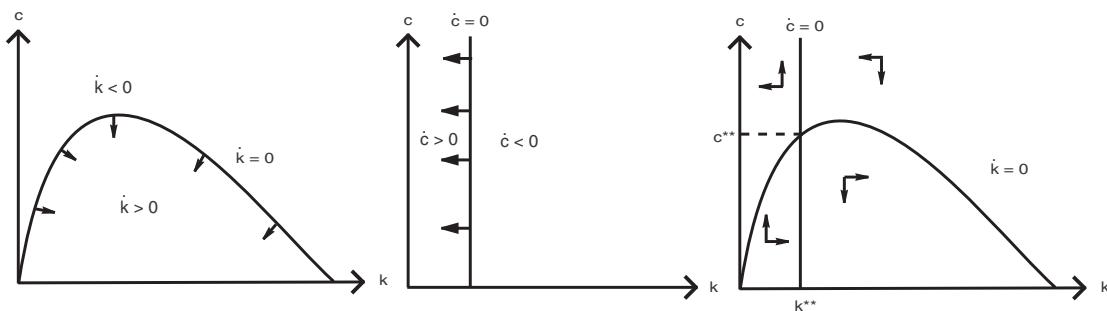


Figura 11.6: Análisis cualitativo del modelo de Ramsey.

§11.4 Extensiones a la ecuación de Euler

§11.4.1 Varias variables

La primera extensión inmediata de la ecuación de Euler es cuando se tiene más de una función en el dominio, es decir $\mathbf{x} = (x_1, \dots, x_n)$, y la funcional que se quiere optimizar es de la forma

$$J[\mathbf{x}] = \int_a^b f(\mathbf{x}, \dot{\mathbf{x}}, t) dt, \quad (11.11)$$

en donde la función f es, como antes, doblemente diferenciable, pero ahora su dominio está en \mathbb{R}^{2n+1} . Dado un vector $\mathbf{v} = (v_1, \dots, v_n) \in \mathbb{R}^n$, denotamos, para toda $i = 1, \dots, n$, por \mathbf{v}_i al vector

$$\mathbf{v}_i = (0, \dots, 0, v_i, 0, \dots, 0).$$

La diferencial de la funcional (11.11) está dada simplemente por

$$DJ[\mathbf{x}](\mathbf{h}) = (DJ[\mathbf{x}_1](\mathbf{h}_1), \dots, DJ[\mathbf{x}_n](\mathbf{h}_n)),$$

en donde $\mathbf{h} = (h_1, \dots, h_n)$. Ésta será igual a 0 si y sólo si cada una de sus componentes lo son y utilizando (11.4) sabemos que

$$DJ[\mathbf{x}_i](\mathbf{h}_i) = \int_a^b h_i(f_{x_i} - \frac{d}{dt}f_{\dot{x}_i}) dt.$$

De lo anterior, es inmediato que una condición necesaria para que la funcional (11.11) posea un extremo es que se cumplan

$$f_{x_i} - \frac{d}{dt}f_{\dot{x}_i} = 0 \quad (11.12)$$

para toda $i = 1, \dots, n$, es decir, si se satisfacen las n ecuaciones de Euler, correspondientes a cada una de las variables.

Ejemplo

Ej 11.4.1 Encontrar el extremo de la funcional

$$J[x, y] = \int_0^2 (x + y + \dot{x}^2 + \dot{y}^2) dt,$$

dadas las condiciones iniciales y finales $x(0) = 0$, $y(0) = 1$, $x(2) = 5$ y $y(2) = 8$. Las ecuaciones de Euler están dadas por

$$1 - 2\ddot{x} = 0,$$

$$1 - 2\ddot{y} = 0,$$

con lo cual se tiene que la solución está dada por

$$\begin{aligned}x(t) &= \frac{1}{4}t^2 + At + B, \\y(t) &= \frac{1}{4}t^2 + Ct + D.\end{aligned}$$

Las constantes se determinan con las condiciones iniciales y finales y se obtiene $A = 2$, $B = 0$, $C = 3$ y $D = 1$.

◇◇◇◇◇◇◇◇◇

§11.4.2 Derivadas de orden superior

Supongamos que ahora queremos una condición necesaria para obtener el extremo de una funcional del siguiente tipo

$$J[x] = \int_a^b f(x, \dot{x}, \ddot{x}, \ddot{\ddot{x}}, \dots, x^{(m)}, t) dt, \quad (11.13)$$

es decir, tenemos derivadas de orden superior de la función $x(t)$. Las condiciones iniciales y finales están dadas por

$$\begin{aligned}x(a) &= x_a, \quad x(b) = x_b; \\ \dot{x}(a) &= \dot{x}_a, \quad \dot{x}(b) = \dot{x}_b, \\ &\vdots \\ x^{(m-1)}(a) &= x_a^{(m-1)}, \quad x(b) = x_b^{(m-1)}.\end{aligned}$$

Para resolver este problema se procede a calcular $DJ[x]$ directamente, al igual que en la sección 11.2. Sea la función

$$h : [a, b] \rightarrow \mathbb{R},$$

de clase $\mathcal{C}^m[a, b]$, que satisface $h(a) = h(b) = \dot{h}(a) = \dot{h}(b) = \dots = h^{(m-1)}(a) = h^{(m-1)}(b) = 0$; de esta forma, la función desplazada $x + h$ pertenece al dominio de la funcional dado por aquellas funciones que satisfacen las condiciones iniciales y finales. Dado un punto $\mathbf{x} = (x, \dot{x}, \ddot{x}, \dots, x^{(m)}, t)$, consideremos la expansión lineal de la función f alrededor de este punto para obtener

$$\begin{aligned}f(\mathbf{x} + h) - f(\mathbf{x}) &\simeq \\ \simeq (f_x(\mathbf{x}), f_{\dot{x}}(\mathbf{x}), f_{\ddot{x}}(\mathbf{x}), \dots, f_{x^{(m)}}(\mathbf{x}), t) \cdot (h(t), \dot{h}(t), \ddot{h}(t), \dots, h^{(m)}(t), 0),\end{aligned}$$

por lo tanto, como en (11.3), se obtiene que

$$DJ[x](h) = \int_a^b \left(f_x h + f_{\dot{x}} \dot{h} + f_{\ddot{x}} \ddot{h} + \dots + f_{x^{(m)}} h^{(m)} \right) dt.$$

Integrando por partes repetidamente los términos $f_{\dot{x}} \dot{h}$, $f_{\ddot{x}} \ddot{h}$, ..., $f_{x^{(m)}} h^{(m)}$ (se deja al lector como ejercicio), llegamos a la expresión

$$DJ[x](h) = \int_a^b h \left(f_x - \frac{d}{dt} f_{\dot{x}} + \frac{d^2}{dt^2} f_{\ddot{x}} + \dots + (-1)^m \frac{d^m}{dt^m} f_{x^{(m)}} \right) dt, \quad (11.14)$$

que es una expresión análoga a (11.4). De aquí que la condición necesaria para que la funcional (11.13) tenga un extremo está dada por

$$f_x - \frac{d}{dt}f_{\dot{x}} + \frac{d^2}{dt^2}f_{\ddot{x}} + \dots + (-1)^m \frac{d^m}{dt^m}f_{x^{(m)}} = 0, \quad (11.15)$$

que es conocida como la **ecuación de Euler-Poisson**.

Ejemplos

Ej 11.4.2 Encontremos un extremo para la funcional

$$J[x] = \int_0^1 (\ddot{x}^2 + \dot{x} + 2t^2) dt$$

dadas las condiciones iniciales y finales $x(0) = 0$, $\dot{x}(0) = 1$, $x(1) = 1$ y $\dot{x}(1) = 1$. La ecuación de Euler-Poisson está dada por

$$0 - \frac{d}{dt}1 + \frac{d^2}{dt^2}2\ddot{x} = 0,$$

o bien, por

$$\frac{d^4x}{dt^4} = 0,$$

cuya solución es

$$x(t) = \frac{At^3}{6} + \frac{Bt^2}{2} + Ct + D.$$

Las condiciones iniciales y finales implican $A = B = D = 0$ y $C = 1$.

Ej 11.4.3 Encontremos un extremo para la funcional

$$J[x] = \int_0^1 (\ddot{x}^2 + x\dot{x} - x^2) dt.$$

La ecuación de Euler-Poisson es

$$-2x + \dot{x} - \frac{d}{dt}x + \frac{d^2}{dt^2}2\ddot{x} = 0,$$

o bien,

$$\frac{d^4x}{dt^4} - x = 0,$$

con polinomio característico $\lambda^4 - 1 = 0$. Las cuatro raíces distintas están dadas por $\lambda = 1, -1, i, -i$, con lo cual la solución queda como

$$x(t) = C_1 \cos t + C_2 \sin t + C_3 e^t + C_4 e^{-t}.$$



§11.5 Condiciones de segundo orden

Hasta este momento no se ha podido determinar si efectivamente el punto extremo que se ha encontrado es un máximo o un mínimo. Al igual que en otras situaciones más simples, es necesario recurrir a alguna condición de suficiencia de segundo orden. Si x^* es tal que $DJ[x^*] = 0$, entonces se satisface

$$J[x^* + h] - J[x^*] = \frac{1}{2} D^2 J[x^*](h, h) + o(\|h\|^2),$$

puesto que el término lineal es igual a cero. Tenemos que si dada $\varepsilon > 0$ se cumple $D^2 J[x^*](h, h) < 0$ ($D^2 J[x^*](h, h) > 0$) para toda h tal que $\|h\| < \varepsilon$, entonces $J[x^*] > J[x^* + h]$ ($J[x^*] < J[x^* + h]$) y, de acuerdo con la definición 11.1.8, x^* es un máximo (mínimo) local. Análogamente, si se cumple $D^2 J[x^*](h, h) < 0$ ($D^2 J[x^*](h, h) > 0$) para toda h , entonces x^* es un máximo (mínimo) global. En otras palabras, el signo de la segunda derivada de Fréchet es suficiente para determinar si x^* es máximo o mínimo.

Como antes, estamos interesados en las funcionales del tipo

$$J[x] = \int_a^b f(x, \dot{x}, t) dt,$$

cuyo dominio es el conjunto de funciones

$$D = \{x : [a, b] \rightarrow \mathbb{R} \mid x \in C^2([a, b]), x(a) = A, x(b) = B\}.$$

Procediendo como en 11.2, en donde llegamos a la ecuación de Euler, tomamos la expansión de Taylor de segundo orden de la función f para así obtener

$$J[x + h] - J[x] \simeq \int_a^b (f_x h + f_{\dot{x}} \dot{h}) dt + \frac{1}{2} \int_a^b (f_{xx} h^2 + 2f_{x\dot{x}} h \dot{h} + f_{\dot{x}\dot{x}} \dot{h}^2) dt.$$

El primer término del lado derecho es $DJ[x](h)$ y la segunda integral es la correspondiente a $D^2 J[x](h, h)$. La siguiente proposición es ahora inmediata.

Proposición 11.5.1 *Si x^* es tal que $DJ[x^*] = 0$ y denotamos por f^* a las segundas parciales de la función f evaluadas en (x^*, \dot{x}^*, t) , entonces una condición suficiente para que x^* sea un máximo (mínimo) local es que se cumpla*

$$D^2 J[x^*](h, h) = \int_a^b (f_{xx}^* h^2 + 2f_{x\dot{x}}^* h \dot{h} + f_{\dot{x}\dot{x}}^* \dot{h}^2) dt < 0 \quad (> 0),$$

para toda h .

Notemos que el integrando es simplemente la forma cuadrática con matriz asociada

$$\begin{pmatrix} f_{xx}^* & f_{x\dot{x}}^* \\ f_{\dot{x}x}^* & f_{\dot{x}\dot{x}}^* \end{pmatrix};$$

por lo tanto, si esta forma cuadrática es negativa definida para toda $t \in [a, b]$ entonces se cumple $D^2J[x^*] < 0$ y x^* es un máximo. Análogamente si es positiva definida para toda $t \in [a, b]$, x^* es un mínimo.

Otra manera de obtener una condición de suficiencia para que x^* sea un máximo (mínimo) es examinando la concavidad (convexidad) de la función f .

Proposición 11.5.2 *Sea $J[x] = \int_a^b f(x, \dot{x}, t) dt$ una funcional con f cóncava (convexa) como función de x y \dot{x} , entonces x^* es un máximo global si y sólo si $DJ[x^*] = 0$. Además, si f es estrictamente cóncava (convexa), el extremo x^* es único.*

Demostración

Basta probar la suficiencia.

La concavidad de f implica que se cumple

$$f(x, \dot{x}, t) \leq f(x^*, \dot{x}^*, t) + f_x(x^*, \dot{x}^*, t)h + f_{\dot{x}}(x^*, \dot{x}^*, t)\dot{h},$$

para todo $x = x^* + h$. Por lo tanto,

$$J[x] \leq J[x^*] + \int_a^b (f_x(x^*, \dot{x}^*, t)h + f_{\dot{x}}(x^*, \dot{x}^*, t)\dot{h})dt,$$

o bien,

$$J[x] \leq J[x^*] + DJ[x^*].$$

Pero $DJ[x^*] = 0$ implica $J[x] \leq J[x^*]$, con lo cual x^* es un máximo global. La demostración es obviamente análoga para el caso de un mínimo global. Si f es estrictamente cóncava, la desigualdad es estricta y por lo tanto x^* es el máximo (mínimo) global único. ■

Ejemplos

Ej 11.5.1 El problema de encontrar la trayectoria óptima de producción en el ejemplo 11.2.2, se refería a maximizar la funcional

$$\int_0^1 (4x - \dot{x}^2 - x^2)e^{-0.1t} dt.$$

Efectivamente, la trayectoria obtenida maximiza el problema ya que la función

$$(4x - \dot{x}^2 - x^2)e^{-0.1t}$$

es estrictamente cóncava como función de (x, \dot{x}) . La matriz hessiana está dada por

$$\begin{pmatrix} -2e^{-0.1t} & 0 \\ 0 & -2e^{-0.1t} \end{pmatrix},$$

en donde $-2e^{-0.1t} < 0$ para toda t y el determinante es igual a $4e^{-0.2t} > 0$ para toda t . Por lo tanto, siempre es negativa definida y la trayectoria obtenida es un máximo global único.

Ej 11.5.2 En el ejemplo 11.2.3, se encontró la trayectoria de longitud mínima entre dos puntos dados. En este caso se obtuvo un extremo para la funcional

$$\int_{t_1}^{t_2} \sqrt{1 + \dot{x}^2} dt.$$

Dado que $z = f(x, \dot{x})$ es simplemente una hoja de hiperboloide como se indica en la figura 11.7, la convexidad es inmediata y efectivamente el extremo obtenido es un mínimo.

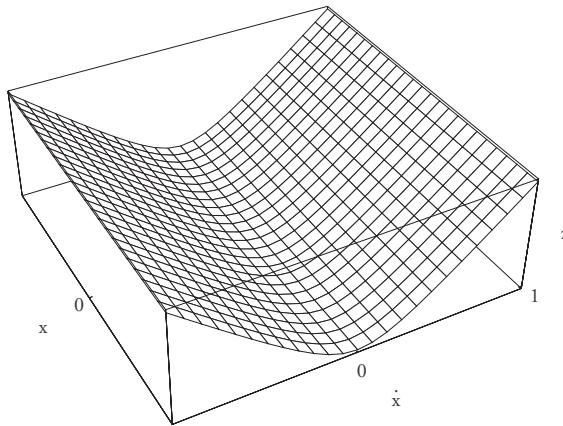


Figura 11.7: Gráfica del hiperboloide correspondiente al ejemplo 11.5.2



§11.6 Condiciones de transversalidad

Hasta este punto hemos considerado funcionales del tipo

$$J[x] = \int_a^b f(x, \dot{x}, t) dt,$$

cuyo dominio está dado por el conjunto de funciones,

$$D_0 = \{x : [a, b] \rightarrow \mathbb{R} \mid x \in \mathcal{C}^2([a, b]), x(a) = A, x(b) = B\}.$$

Por conveniencia, tomamos $a = 0$ y $b = T$ o bien $b = \infty$, según sea necesario. Si se piensa en el problema de maximizar, por ejemplo, la trayectoria de producción de una empresa, podemos pensar que la producción final $x(T)$ no está dada, sino que se va a determinar de manera óptima. Asimismo, puede ser que el tiempo final T tampoco esté dado y se determine de forma endógena.

Si T está dado pero $x(T)$ está libre, el dominio D_0 se transforma en

$$D_1 = \{x : [0, T] \rightarrow \mathbb{R} \mid x \in C^2([0, T]), x(0) = x_0, T \text{ dado}, x(T) \text{ libre}\}.$$

Análogamente, si $x(T)$ está dado pero $T \in [0, \infty)$ está libre, se tiene

$$D_2 = \{x : [0, \infty) \rightarrow \mathbb{R} \mid x \in C^2([0, \infty)), x(0) = x_0, x(T) = x_T, T \text{ libre}\}.$$

Finalmente, podemos dejar que tanto T como $x(T)$ estén libres y así

$$D_3 = \{x : [0, \infty) \rightarrow \mathbb{R} \mid x \in C^2([0, \infty)), x(0) = x_0, x(T) \text{ y } T \text{ libres}\}.$$

Geométricamente, se pueden visualizar estos distintos dominios como en la figura 11.8. Estos tres casos los denominamos, respectivamente, problema con recta vertical final, problema con recta horizontal final y problema libre (la razón debería ser evidente dada la figura 11.8).

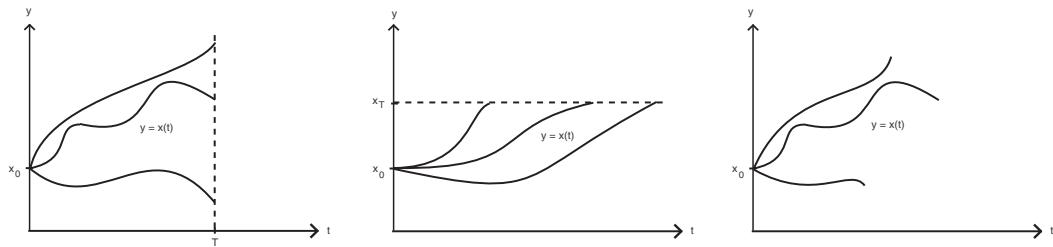


Figura 11.8: Algunos elementos de los dominios D_1 , D_2 y D_3 .

Cuando desarrollamos la ecuación de Euler, en el desplazamiento $x + h$, se tomó h de tal forma que $h(0) = h(T) = 0$ de manera que la trayectoria $x + h$ perteneciera al dominio V . Ahora seguimos teniendo $h(0) = 0$, pero $h(T)$ puede ser distinto de 0 y también el tiempo final T puede ser variable. Por lo tanto, podemos expresar la funcional como

$$J[x, T] = \int_0^T f(x, \dot{x}, t) dt$$

y se tiene que

$$\begin{aligned}
 J[x + h, T + \Delta T] - J[x, T] &= \int_0^{T+\Delta T} f(x + h, \dot{x} + \dot{h}, t) dt - \int_0^T f(x, \dot{x}, t) dt \\
 &= \int_0^T [f(x + h, \dot{x} + \dot{h}, t) - f(x, \dot{x}, t)] dt + \int_T^{T+\Delta T} f(x + h, \dot{x} + \dot{h}, t) dt \\
 &= \int_0^T (f_x h + f_{\dot{x}} \dot{h}) dt + (J[x + h, T + \Delta T] - J[x + h, T]) \\
 &= \int_0^T (f_x h + f_{\dot{x}} \dot{h}) dt + f(x(T), \dot{x}(T), T) \Delta T.
 \end{aligned}$$

El primer término del lado derecho, dado por la integral, representa el cambio en la funcional inducido por el desplazamiento h . Éste se obtiene de manera idéntica que en la sección 11.2. El segundo término representa el cambio inducido por T y se obtiene aplicando el teorema fundamental del cálculo para derivar $J[x, T]$ con respecto a T . La derivada de Fréchet queda ahora dada por

$$DJ[x, T](h, \Delta T) = \int_0^T (f_x h + f_{\dot{x}} \dot{h}) dt + f(x(T), \dot{x}(T), T) \Delta T.$$

Al integrar por partes el segundo término de la integral, nos queda

$$\int_0^T f_{\dot{x}} \dot{h} dt = (h f_{\dot{x}}) |_0^T - \int_0^T h df_{\dot{x}} = h(T)(f_{\dot{x}})_{t=T} - \int_0^T h df_{\dot{x}},$$

por lo que

$$DJ[x, T](h, \Delta T) = \int_0^T h(f_x - \frac{d}{dt} f_{\dot{x}}) dt + f(x(T), \dot{x}(T), T) \Delta T + h(T)(f_{\dot{x}})_{t=T}. \quad (11.16)$$

De la figura 11.9 se observa que $h(T) = \Delta x(T) - \dot{x}(T) \Delta T$ y sustituyendo en (11.16) se tiene

$$DJ[x, T](h, \Delta T) = \int_0^T h(f_x - \frac{d}{dt} f_{\dot{x}}) dt + (f_{\dot{x}})_{t=T} \Delta x(T) + (f - \dot{x} f_{\dot{x}})_{t=T} \Delta T. \quad (11.17)$$

De (11.17) se infiere que las condiciones necesarias para un extremo son la ecuación de Euler

$$f_x - \frac{d}{dt} f_{\dot{x}} = 0$$

y alguna de las condiciones siguientes, llamadas de transversalidad.

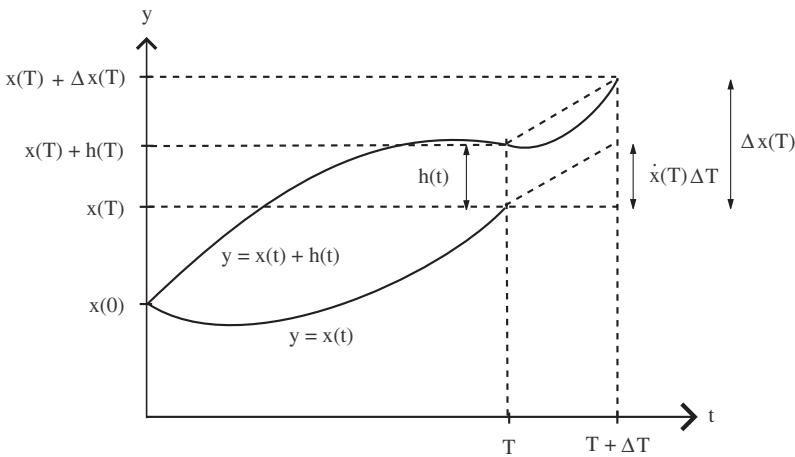


Figura 11.9: Nótese la igualdad $h(T) = \Delta x(T) - \dot{x}(T)\Delta T$.

Definición 11.6.1 (Condiciones de transversalidad) *Las siguientes son llamadas condiciones de transversalidad*

- a) $(f_{\dot{x}})_{t=T} = 0$ si T está dado (por lo que $\Delta T = 0$) y $x(T)$ está libre.
- b) $(f - \dot{x}f_{\dot{x}})_{t=T} = 0$ si $x(T)$ está dado (por lo que $\Delta x(T) = 0$) y T está libre.
- c) $(f_{\dot{x}})_{t=T} = 0$ y $(f - \dot{x}f_{\dot{x}})_{t=T} = 0$ si T y $x(T)$ están libres.

Ejemplos

Ej 11.6.1 Encontrar un extremo para la funcional

$$\int_0^2 (t^2 + \dot{x}^2 + x^2) dt,$$

dado que $x(0) = 1$ y $x(2)$ está libre. La ecuación de Euler está dada por

$$2x - \frac{d}{dt} 2\dot{x} = 0,$$

por lo que se tiene $x - \ddot{x} = 0$ y por lo tanto la solución es

$$x(t) = Ae^t + Be^{-t}.$$

La condición inicial implica

$$A + B = 1$$

y por la primera condición de transversalidad en (11.6.1) se tiene que $2\dot{x}(2) = 0$ o bien

$$Ae^2 - Be^{-2} = 0.$$

Resolviendo para A y B tenemos finalmente $A = \frac{1}{1+e^4}$ y $B = \frac{e^4}{1+e^4}$. El extremo obtenido es un mínimo ya que la función $z = \dot{x}^2 + x^2$ es un paraboloide convexo con vértice en el origen.

Ej 11.6.2 Encontrar un extremo de la funcional

$$\int_0^T (t\dot{x} + \dot{x}^2) dt,$$

dado que $x(0) = 1$, $x(T) = 10$ y T está libre. La ecuación de Euler queda dada por

$$-\frac{d}{dt}(t + 2\dot{x}) = 0,$$

o bien $1 + 2\ddot{x} = 0$ cuya solución es

$$x(t) = -\frac{1}{4}t^2 + At + B.$$

La condición inicial implica $B = 1$ y utilizando la segunda condición de transversalidad en (11.6.1) se obtiene

$$[t\dot{x} + \dot{x}^2 - \dot{x}(t + 2\dot{x})]_{t=T} = 0,$$

que se reduce a $\dot{x}(T) = 0$ o bien

$$-\frac{1}{2}T + A = 0.$$

Por otro lado, $x(T) = 10$ implica

$$-\frac{1}{4}T^2 + AT + 1 = 10,$$

al resolver para A y T , tomando en cuenta que $T \geq 0$, obtenemos

$$A = 3,$$

$$T = 6.$$

Es inmediato verificar que el integrando es una función convexa y por lo tanto el extremo es un mínimo.

◇◇◇◇◇◇◇◇

§11.7 Problemas con horizonte infinito

Un supuesto común en economía es considerar que los agentes son inmortales, o más bien, que se comportan como si lo fueran. Las causas de este comportamiento pueden ser diversas pero entre las más citadas y sensatas están: primero, los agentes se preocupan por sus descendientes y, segundo, a pesar de que el periodo de vida de los agentes es finito existe incertidumbre acerca de la duración de este periodo. Independientemente de las razones que justifiquen este comportamiento, modelarlo es relativamente sencillo ya que estamos buscando el extremo de una funcional del tipo

$$J[x] = \int_0^\infty f(x, \dot{x}, t) dt.$$

Lo primero que notamos es que ahora se tiene una integral impropia y por lo tanto para que efectivamente la funcional esté definida, la integral debe converger. Un caso en donde la convergencia es inmediata es aquel en donde la funcional es del tipo

$$J[x] = \int_0^\infty g(x, \dot{x}, t) e^{-\rho t} dt,$$

con $\rho > 0$ una tasa de descuento temporal. Esta integral converge si, por ejemplo, la función g es positiva y acotada, **i.e.** existe $M \geq 0$ tal que $0 \leq g \leq M$ para toda t , o bien, g puede ser incluso creciente pero a una tasa menor que ρ . Casi todas las funciones que se usan en economía caen en alguna de estas categorías; sin embargo, se sugiere que cuando haya duda se verifique la convergencia explícitamente.

Cuando el horizonte es infinito es evidente que no existe un tiempo final específico y por lo tanto la segunda condición de transversalidad en (11.6.1) se debe verificar en el límite, es decir,

$$\lim_{T \rightarrow \infty} (f - \dot{x} f_{\dot{x}})_{t=T} = 0.$$

Asimismo, debe satisfacerse alguna de las dos siguientes condiciones

$$\lim_{T \rightarrow \infty} x(T) = x^*,$$

o bien,

$$\lim_{T \rightarrow \infty} (f_{\dot{x}})_{t=T} = 0.$$

Es decir, $x(T)$ converge asintóticamente a un valor x^* o en el límite se cumple la primera condición de transversalidad de (11.6.1).

Ejemplo

Ej 11.7.1 En el modelo de crecimiento de Ramsey que se estudió en la sección 11.3, las condiciones de transversalidad para un horizonte de tiempo infinito están dadas por

$$\begin{aligned} \lim_{T \rightarrow \infty} u' e^{-\rho T} &= 0, \\ \lim_{T \rightarrow \infty} (u + \dot{k} u') e^{-\rho T} &= 0 \end{aligned}$$

y éstas se satisfacen si la solución converge a su estado estacionario cuando $T \rightarrow \infty$, es decir, la solución del sistema converge a lo largo de la variedad estable. Para el ejemplo numérico cuya solución está dada por (11.10) podemos poner la condición inicial $k(0) = 35$ y tomar un horizonte infinito, es decir, $T = \infty$. La única trayectoria que satisface las condiciones de transversalidad es la variedad estable; de esta forma, el consumo inicial $c(0)$ debe ser tal que $A = 0$. Así, debemos resolver el sistema dado por

$$\begin{aligned} k(0) &= B + 10\sqrt{10} = 35, \\ c(0) &= 0.21B + 2.5\sqrt{10}. \end{aligned}$$

Se obtiene entonces que

$$B = 3.38,$$

$$c(0) = 8.6,$$

y la solución está dada por

$$\begin{aligned} k(t) &= 3.38e^{-0.16t} + 10\sqrt{10}, \\ c(t) &= 0.71e^{-0.16t} + 2.5\sqrt{10}. \end{aligned}$$

La ecuación de la variedad estable, sobre la cual se mueve el sistema, es entonces

$$c(t) = 0.21k(t) + 0.4\sqrt{10}.$$



Parecería que las condiciones de transversalidad con horizonte infinito no presentan problema alguno; sin embargo, esto no es así. Por definición, las condiciones de transversalidad requieren que la trayectoria $x(t)$ esté en algún punto final (óptimo) en un tiempo terminal (también óptimo). El límite de las condiciones de transversalidad cuando $T \rightarrow \infty$ puede no existir en algunos casos y es posible tener trayectorias óptimas que no satisfagan las condiciones de transversalidad en horizonte infinito.⁴ En general, no se tienen condiciones de transversalidad apropiadas para el problema de horizonte infinito; no obstante, en los problemas de economía siempre se puede apelar al sentido común para determinar el estado terminal del sistema.

§11.7.1 Un modelo de inversión

Vamos a suponer que una empresa tiene como único insumo el capital. Sea $\Pi(K) = AK - BK^2$, su función de ganancias, y sea $C(\dot{K}) = \alpha\dot{K}^2 + \beta\dot{K}$, su función de costos de inversión. Se tiene también que $A - \beta\rho > 0$. No hay depreciación del capital. El problema de la empresa es el siguiente:⁵

$$\max \int_0^\infty [\Pi(K) - C(\dot{K})]e^{-\rho t} dt,$$

dado $K(0) = K_0$.

La ecuación de Euler queda dada por

$$\Pi'e^{-\rho t} - \frac{d}{dt}(-C'e^{-\rho t}) = 0,$$

o bien,

$$\Pi' - C'\rho + C''\ddot{K} = 0.$$

Sustituyendo $\Pi' = A - 2BK$, $C' = 2\alpha\dot{K} + \beta$ y $C'' = 2\alpha$, reescribimos esta última ecuación como

$$\ddot{K} - \rho\dot{K} - \frac{B}{\alpha}K = \frac{\beta\rho - A}{2\alpha},$$

cuya solución está dada por

$$K(t) = \kappa_1 e^{r_1 t} + \kappa_2 e^{r_2 t} + K_p,$$

⁴Véase por ejemplo [Tak85].

⁵Basado en [ES63].

en donde,

$$r_i = \frac{1}{2} \left(\rho \pm \sqrt{\rho^2 + \frac{4B}{\alpha}} \right), \quad i = 1, 2,$$

y

$$K_p = \frac{A - \beta\rho}{2B}.$$

Notemos que ambas raíces son reales y de signos opuestos (supongamos $r_1 > 0$ y $r_2 < 0$) y que la condición $A - \beta\rho > 0$ implica que la solución particular K_p tiene sentido, es decir, $K_p > 0$. La condición inicial implica $\kappa_1 + \kappa_2 + K_p = K_0$ y la única forma de que $K(t)$ converja a un valor finito es poniendo $\kappa_1 = 0$, por lo que entonces $\kappa_2 = K_0 - K_p$ y la solución queda finalmente como

$$K(t) = (K_0 - K_p)e^{r_2 t} + K_p$$

y $\lim_{t \rightarrow \infty} K(t) = K_p$. La trayectoria de inversión óptima queda especificada por

$$\begin{aligned} I(t) &= \dot{K}(t) \\ &= r_2(K_0 - K_p)e^{r_2 t} \\ &= -r_2(K_p - K(t)). \end{aligned}$$

Las condiciones de transversalidad son

$$\begin{aligned} \lim_{T \rightarrow \infty} C'(\dot{K})e^{-\rho T} &= 0 \text{ y} \\ \lim_{T \rightarrow \infty} [\Pi(K) - C(\dot{K}) + \dot{K}C'(\dot{K})]e^{-\rho T} &= 0. \end{aligned}$$

Es fácil ver que la solución encontrada satisface estas condiciones. A este tipo de ajuste se lo denomina **modelo del acelerador flexible**: la inversión, en este caso, es proporcional a la diferencia entre el tamaño “estable” de la empresa K_p y el tamaño real $K(t)$. Esta situación se presentará siempre que tengamos dos raíces reales y de signos distintos en un modelo de horizonte infinito, puesto que la convergencia es sobre la variedad estable. Normalmente se pide que $-r_2 \in (0, 1)$ y esto puede lograrse escogiendo adecuadamente los parámetros del problema.

Si la función de costos es lineal, es decir supongamos $\alpha = 0$, entonces la ecuación de Euler se reduce a

$$A - 2BK - \beta\rho = 0,$$

por lo que

$$K^* = \frac{A - \beta\rho}{2B}.$$

Es decir, no hay dinámica, sólo existe un valor del capital dado por K^* , que es el tamaño óptimo de la empresa. El ajuste tiene que ser instantáneo en $t = 0$. La intuición detrás de esto es que si pensamos en el costo marginal C' como “el precio de la unidad marginal de capital”, al tener una función lineal este precio es constante en el tiempo por lo que no existe ningún incentivo para posponer la acumulación o desacumulación del capital y todo el ajuste se realiza en el instante inicial. El primero en hacer esta observación fue Jorgenson en 1967 [Jor67].

Ejercicios

▷ **11.1** Demostrar lo siguiente:

a) Si $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$, se define

$$\|\mathbf{x}\|_1 = |x_1| + \dots + |x_n|;$$

probar que $\|\cdot\|_1$ es una norma en \mathbb{R}^n .

b) Si $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$, se define

$$\|\mathbf{x}\|_\infty = \sup \{|x_1|, \dots, |x_n|\};$$

probar que $\|\cdot\|_\infty$ es una norma en \mathbb{R}^n .

▷ **11.2** Sea

$$V = \{f : [a, b] \rightarrow \mathbb{R} \mid f \text{ es continua}\}.$$

Se define

$$\|f\|_p = \left\{ \int_a^b |f|^p dt \right\}^{\frac{1}{p}};$$

probar que $\|f\|_p$ es una norma en V para $p = 1, 2$.

▷ **11.3** Sea $V = \{f : [0, 1] \rightarrow \mathbb{R} \mid f \text{ es continua}\}$.

a) Probar que V es un espacio vectorial sobre \mathbb{R} .

b) Dar dos ejemplos de funcionales lineales definidas sobre V .

c) Dar dos ejemplos de funcionales no lineales definidas sobre V .

▷ **11.4** Encontrar los extremos de las siguientes funcionales.

a) $J[x] = \int_0^{40} -\frac{\dot{x}^2}{2} dt$ con $x(0) = 20$ y $x(40) = 0$.

b) $J[x] = \int_0^{10} -(2x\dot{x} + \dot{x}^2) dt$ con $x(0) = 10$ y $x(10) = 100$.

c) $J[x] = \int_0^2 (12tx + \dot{x}^2) dt$ con $x(0) = 1$ y $x(2) = 17$.

d) $J[x] = \int_0^2 (x + \dot{x}^2) dt$ con $x(0) = 1$ y $x(2) = 10$.

e) $J[x] = \int_0^2 (x^2 + t^2 \dot{x}) dt$ con $x(0) = 0$ y $x(2) = 2$.

f) $J[x, y] = \int_0^T (2xy - 2x^2 - \dot{x}^2 + \dot{y}^2) dt$, encontrar únicamente la solución general.

g) $J[x, y] = \int_0^{10} (\dot{x}^2 + \dot{y}^2 + e^t) dt$ dadas $x(0) = 0$, $y(0) = 2$, $x(10) = 11$ y $y(10) = 6$.

h) $J[x, y] = \int_0^{\frac{\pi}{2}} (\dot{x}^2 + \dot{y}^2 + 2xy) dt$ dadas $x(0) = y(0) = 0$ y $x(\frac{\pi}{2}) = y(\frac{\pi}{2}) = 1$.

i) $J[x] = \int_0^1 (1 + \ddot{x}^2) dt$ con $x(0) = 0$ y $\dot{x}(0) = x(1) = \dot{x}(1) = 1$.

▷ **11.5** Una empresa tiene un pedido de N unidades que debe surtir en $T = 1$. Si $x(t)$ denota el número de unidades producidas en $[0, t]$ (se puede interpretar como el inventario acumulado en t), el costo en t está dado por $c(x, \dot{x}) = 2x + \dot{x}^2$. Resolver el problema de minimización de costos de la empresa:

$$\min \int_0^1 (2x + \dot{x}^2) dt \text{ con } x(0) = 0 \text{ y } x(1) = N.$$

▷ **11.6** Encontrar el extremo de

$$\int_0^T (t^2 + \dot{x}^2) dt$$

para los siguientes casos y determinar si es un máximo o un mínimo:

a) $x(0) = 4, T = 2$ y $x(T)$ libre.

b) $x(0) = 4, x(T) = 5$ y T libre.

▷ 11.7 Encontrar el extremo de

$$\int_0^T (x + \dot{x}^2 + t) dt$$

para los siguientes casos y determinar si es un máximo o un mínimo.

a) $x(0) = 1, T = 2$ y $x(2)$ libre.

b) $x(0) = x(T) = 0$ y $T > 0$ libre.

▷ 11.8 En el modelo de inversión de la sección 11.7.1, utilizar las condiciones de transversalidad para llegar a $\kappa_1 = 0$.

▷ 11.9 Resolver el problema de la empresa dado en el ejercicio 11.5 con la siguiente variante:

$$\min \int_0^T (2x + \dot{x}^2) dt$$

con $x(0) = 0, x(T) = N$ y T libre.

▷ 11.10 Considerar la siguiente variante del modelo de Ramsey:

$$\max \int_0^\infty (\ln c) e^{-\rho t} dt$$

sujeto a

$$Ak = c + \dot{k} + \delta k,$$

$$k(0) = k_0 > 0,$$

$$A - \delta > 0, A - \delta - \rho < 0.$$

Es decir, la función de utilidad está dada por $u(c) = \ln c$ y la función de producción es $f(k) = Ak$.

- a) Resolver el problema, incluyendo la condición de transversalidad.
- b) Realizar un diagrama de fase del sistema de ecuaciones resultante en k y c .
- c) Encontrar el punto de equilibrio y probar que es un punto silla.
- d) Encontrar la ecuación de la variedad estable; ¿cómo se describiría el futuro de esta economía? Explicar.

CAPÍTULO 12

Teoría de control

§12.1 Introducción

El cálculo en variaciones es una herramienta muy útil que, sin embargo, no es suficientemente poderosa para resolver muchos de los problemas que se presentan en las aplicaciones. Nos gustaría poder considerar casos en donde la función $f(x, \dot{x}, t)$ es lineal, casos en donde las trayectorias $x(t)$ son funciones más generales y no son necesariamente doblemente diferenciables, y casos con restricciones sobre las trayectorias y otras generalizaciones.

Los primeros trabajos al respecto fueron efectuados por Valentine en 1937, McShane en 1939 y Hestenes en 1947. El tema no atrajo una gran atención sino hasta mediados del siglo veinte cuando el matemático ruso Pontryagin y sus colaboradores, Boltyanskii, Gamkrelidze y Mishchenko, desarrollaron la llamada teoría de control óptimo.¹ El resultado básico que presentaron es conocido como *el principio del máximo de Pontryagin*, que se refiere a las condiciones necesarias para la optimalidad del problema que plantearemos a continuación.

Supongamos que por economía se entiende un sistema descrito en el tiempo t por ciertas variables, denominadas de **estado**, dadas por $x_1(t), \dots, x_n(t)$, que podrían ser capital, inflación, cantidad de algún activo u otras. El problema que se plantea es obtener trayectorias óptimas para estas variables de manera que se maximice o minimice algún objetivo dado. Este objetivo puede ser el valor presente del bienestar social, o el valor presente de la deuda pública, o el valor presente de la riqueza de los servidores públicos, o cualquier otra cosa pertinente. Las variables $x_i(t)$ pueden ser “controladas” por otras variables, denominadas de **control**, como podrían ser el consumo, los balances nominales (política monetaria), los impuestos (política fiscal), etc., denotadas por $u_1(t), \dots, u_m(t)$.

El problema general de obtener una trayectoria para las variables de estado $x_i(t)$ escogiendo adecuadamente los controles $u_j(t)$ de manera que se optimice algún objetivo es conocido como problema de

¹La traducción del ruso al inglés fue publicada en [PBGM62]

control óptimo. Este problema puede ser tan sencillo como minimizar el tiempo en el que se llena una tina si la variable de estado es la cantidad de agua dentro de ella y el control es el flujo del agua de la llave, o bien puede ser algo complicado como minimizar el gasto de combustible de una nave espacial.

§12.2 Planteamiento del problema

La exposición rigurosa del problema general de teoría de control es bastante complicada;² sin embargo, podemos dar una versión simplificada que sea suficiente para los problemas que se tratan en el texto. Para mayor claridad de la exposición, supongamos que se tienen únicamente un estado y un control.

Definición 12.2.1 *Se dice que una función $\alpha(t)$ tiene una **discontinuidad del primer tipo** en t_0 , si los límites*

$$\lim_{t \rightarrow t_0^+} \alpha(t) \text{ y } \lim_{t \rightarrow t_0^-} \alpha(t)$$

ambos existen pero son distintos. Esto es lo que se conoce por discontinuidad de tipo “brinco” o de primer tipo.

Definición 12.2.2 *Se dice que una función $\alpha(t)$ es continua por pedazos si es continua en todo su dominio, excepto en un número finito de puntos en los que tiene discontinuidades del primer tipo.*

Denotemos a la variable de estado por $x(t)$ y supongamos que es continua y sus derivadas son continuas por pedazos. La variable de control $u(t)$ puede escogerse de un conjunto de funciones U llamado el **conjunto de controles admisibles**, que son funciones continuas por pedazos. Al conjunto $\bar{U} = \{u(t) \mid u \in U\}$, de imágenes de los controles admisibles, se lo llama **región de control**. Se tienen dos funciones $f(x, u, t)$ y $g(x, u, t)$ con derivadas parciales continuas. En ecuaciones diferenciales existe un teorema (no trivial) conocido como el teorema de Peano (véase el apéndice B), que garantiza que dada una trayectoria $u(t) = u^*(t)$ y la condición inicial $x(t_0) = x_0$, existe (localmente) una solución única de la ecuación diferencial

$$\dot{x} = g(x, u^*, t).$$

Nuestro problema será escoger el control, u^* , de manera que se optimice el valor de la funcional (conocida como **función objetivo**) dada por

$$\int_0^T f(x, u, t) dt$$

o, en otras palabras, se debe encontrar una trayectoria $u^* \in U$ que resuelva el siguiente problema:

$$\begin{aligned} & \max \int_0^T f(x, u, t) dt \\ & \text{sujeto a } \dot{x} = g(x, u, t), \\ & x(0) = x_0, T \text{ dado y } x(T) \text{ libre.} \end{aligned} \tag{12.1}$$

El principio del máximo nos dice cómo hacer esto y establece lo siguiente.

²Remitimos al lector interesado a [PBGM62]

Teorema 12.2.3

Supongamos que u^* y x^* resuelven el problema (12.1), entonces existe $\lambda(t)$ continua, llamada **variable de coestado**, tal que el **hamiltoniano**, definido por

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

posee un máximo (mínimo) en u^* , es decir,

$$H(x^*, u^*, \lambda, t) \geq H(x^*, u, \lambda, t) \text{ para todas } u \in U, t \in [0, T] \quad (12.2)$$

(o bien \leq si se trata de un mínimo). Adicionalmente, $\lambda(t)$, $u^*(t)$ y $x^*(t)$ resuelven el sistema dinámico

$$\begin{aligned} \dot{\lambda} &= -H_x, \\ \dot{x} &= H_\lambda, \end{aligned} \quad (12.3)$$

y se cumple la condición de transversalidad

$$\lambda(T) = 0. \quad (12.4)$$

Nótese que la segunda condición en (12.3) es simplemente la restricción $\dot{x} = g(x, u, t)$.

De esta forma, las condiciones (12.2)-(12.4) forman un conjunto de condiciones necesarias para que u^* y x^* sean solución al problema (12.1). El nombre del principio del máximo proviene del hecho de que una de las condiciones necesarias para maximizar la función objetivo es la maximización del hamiltoniano. Observemos que si la región de control \bar{U} es un conjunto compacto, el hamiltoniano siempre tiene un extremo e incluso podría ser lineal en el control u (una función continua siempre alcanza sus valores máximo y mínimo si está definida sobre un conjunto compacto). En el caso del cálculo en variaciones, esta posibilidad queda excluida.

Demostración

Sea u^* el control óptimo y x^* la trayectoria del estado óptimo correspondiente. Supongamos, sin pérdida de generalidad, que estas trayectorias maximizan la funcional dada en 11.1. Sean f_x^* y g_x^* las derivadas parciales correspondientes de f y g evaluadas en u^* , x^* . La demostración se hará para el caso particular en que $f_x^* = f_x(x^*, u^* + \Delta u, t)$ y $g_x^* = g_x(x^*, u^* + \Delta u, t)$, por ejemplo, cuando f y g son separables en x y u .

Definimos $\lambda(t)$ como la solución a la ecuación diferencial

$$\begin{aligned} \dot{\lambda} &= -f_x^* - \lambda g_x^*, \\ \lambda(T) &= 0 \end{aligned} \quad (12.5)$$

y definimos al hamiltoniano como

$$H(x, u, \lambda, t) = f(x, u, t) + \lambda(t)g(x, u, t). \quad (12.6)$$

La pareja (x^*, u^*) es un máximo (local) para la funcional

$$J[x, u] = \int_0^T f(x, u, t) dt,$$

si para todo punto $(x^* + \Delta x, u^* + \Delta u)$ en una vecindad³ de (x^*, u^*) se cumple

$$J[x^*, u^*] \geq J[x^* + \Delta x, u^* + \Delta u].$$

Esta desigualdad se puede reescribir como $\Delta J^* \leq 0$, o bien

$$\int_0^T [f(x^* + \Delta x, u^* + \Delta u, t) - f(x^*, u^*, t)] dt \leq 0. \quad (12.7)$$

Vamos a demostrar que una condición necesaria para que (12.7) se cumpla es que se verifique

$$H(x^*, u^* + \Delta u, \lambda, t) - H(x^*, u^*, \lambda, t) \leq 0, \quad (12.8)$$

es decir, que u^* maximice el hamiltoniano.

Por conveniencia, denotamos $\Delta x = h$, en donde, como en la sección 11.2, $h(t)$ es una curva de desplazamiento con $h(0) = 0$. Denotemos por Δf^* , Δg^* y ΔH^* el cambio total de las funciones con respecto a x y u . Explícitamente, se tiene

$$\begin{aligned} \Delta f^* &= f(x^* + h, u^* + \Delta u, t) - f(x^*, u^*, t), \\ \Delta g^* &= g(x^* + h, u^* + \Delta u, t) - g(x^*, u^*, t), \\ \Delta H^* &= H(x^* + h, u^* + \Delta u, \lambda, t) - H(x^*, u^*, \lambda, t). \end{aligned}$$

Por $\Delta^u f^*$, $\Delta^u g^*$ y $\Delta^u H^*$ denotamos los cambios con respecto a u , o bien

$$\begin{aligned} \Delta^u f^* &= f(x^*, u^* + \Delta u, t) - f(x^*, u^*, t), \\ \Delta^u g^* &= g(x^*, u^* + \Delta u, t) - g(x^*, u^*, t), \\ \Delta^u H^* &= H(x^*, u^* + \Delta u, \lambda, t) - H(x^*, u^*, \lambda, t). \end{aligned}$$

Consideremos la siguiente expresión:

$$\int_0^T (\Delta f^* + \lambda \Delta \dot{x}) dt = \int_0^T (\Delta f^* + \lambda \dot{h}) dt. \quad (12.9)$$

Por un lado, $\dot{x} = g(x, u, t)$ implica que (12.9) se puede reescribir como

$$\int_0^T (\Delta f^* + \lambda \Delta g^*) dt = \int_0^T \Delta H^* dt \quad (12.10)$$

³Podemos pensar en una bola abierta $B_\varepsilon(x^*, u^*)$, para alguna $\varepsilon > 0$.

y por otro lado, integrando el término $\int_0^T \lambda \dot{h} dt$ por partes, se tiene

$$\int_0^T \lambda \dot{h} dt = (\lambda h) \Big|_0^T - \int_0^T \lambda \dot{h} dt.$$

Dado que $\lambda(T) = 0$ y $h(0) = 0$, se tiene que $(\lambda h) \Big|_0^T = 0$ y por lo tanto reescribimos (12.9) como

$$\int_0^T (\Delta f^* + \lambda \dot{h}) dt = \int_0^T (\Delta f^* - \lambda \dot{h}) dt. \quad (12.11)$$

De (12.5), (12.10), (12.11) y recordando que $h = \Delta x$ y $\dot{h} = \Delta \dot{x} = \Delta g^*$, se tiene que

$$\Delta J^* = \int_0^T \Delta f^* dt = \int_0^T (\Delta f^* + \lambda \Delta g^* + (-f_x^* - \lambda g_x^*) \Delta x) dt. \quad (12.12)$$

Consideremos el desarrollo de Taylor de primer orden para f y g alrededor de $(x^*, u^* + \Delta u, t)$ para obtener

$$\begin{aligned} f(x^* + \Delta x, u^* + \Delta u, t) &\simeq f(x^*, u^* + \Delta u, t) + f_x^* \Delta x, \\ g(x^* + \Delta x, u^* + \Delta u, t) &\simeq g(x^*, u^* + \Delta u, t) + g_x^* \Delta x. \end{aligned}$$

Notemos que aquí se usa el supuesto inicial $f_x^* = f_x(x^*, u^* + \Delta u, t)$ y $g_x^* = g_x(x^*, u^* + \Delta u, t)$. Restando $f(x^*, u^*, t)$ y $g(x^*, u^*, t)$, respectivamente, de ambos lados de estas igualdades se tiene,

$$\begin{aligned} \Delta f^* &\simeq \Delta^u f^* + f_x^* \Delta x, \\ \Delta g^* &\simeq \Delta^u g^* + g_x^* \Delta x, \end{aligned} \quad (12.13)$$

y sustituyendo (12.13) en (12.12) obtenemos

$$\Delta J^* = \int_0^T \Delta f^* dt = \int_0^T (\Delta^u f^* + \lambda \Delta^u g^*) dt = \int_0^T \Delta^u H^* dt. \quad (12.14)$$

Supongamos que en algún intervalo $(a, b) \subset [0, T]$ se tiene que $\Delta^u H^* > 0$. Podemos tomar $\Delta u \neq 0$ en (a, b) y $\Delta u = 0$ en $[0, T] \cap (a, b)^c$ para obtener

$$\int_0^T \Delta^u H^* dt > 0$$

pero en este caso $\Delta J > 0$, lo cual sería una contradicción ya que (x^*, u^*) es un máximo, por lo que no existe intervalo alguno en el que pueda darse $\Delta^u H^* > 0$. Con esto terminamos la demostración. Obviamente la demostración es análoga para el caso de un mínimo. ■

El teorema anterior nos proporciona también una condición de suficiencia dada en el siguiente corolario.

Corolario 12.2.4 La condición $\Delta^u H^* < 0$ ($\Delta^u H^* > 0$) es suficiente para que $J[x^*, u^*]$ sea un máximo (mínimo).

Demostración

Es inmediato de (12.14) ya que en este caso $\int_0^T \Delta^u H^* dt < 0$ ($\int_0^T \Delta^u H^* dt > 0$). ■

El corolario 12.2.4 implica que, en el caso de que el extremo sea un punto interior, podemos aplicar el criterio de la segunda derivada para H .

Observemos que si adicionalmente el valor final de la variable de estado es conocido, i.e. T y $x(T)$ están dados, entonces se puede prescindir de la condición de transversalidad $\lambda(T) = 0$.

Ejemplos

Ej 12.2.1 Notemos que el siguiente problema de cálculo en variaciones se puede traducir al lenguaje de teoría de control: encontrar el extremo de

$$J[x] = \int_0^T f(x, \dot{x}, t) dt$$

con $x(0) = x_0$, T dado y $x(T)$ libre. Simplemente definimos como variable de control $u = \dot{x}$, de manera que la función g es $g(x, u, t) = u$. El hamiltoniano queda como

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

y las condiciones necesarias de primer orden, suponiendo que f no es lineal, están dadas por

$$\begin{aligned} H_u &= f_u + \lambda = 0, \\ \dot{\lambda} &= -H_x = -f_x, \\ \dot{x} &= H_\lambda = u. \end{aligned}$$

La primera de estas condiciones implica $\lambda = -f_u$; reescribimos la segunda condición como,

$$f_x - \frac{d}{dt} f_u = f_x - \frac{d}{dt} (-f_u) = 0,$$

que no es más que la ecuación de Euler. Se cumple también la condición inicial $x(0) = x_0$ y la condición de transversalidad queda dada por $\lambda(T) = 0 = (-f_u)_{t=T}$, o bien, $(f_{\dot{x}})_{t=T} = 0$ como antes.

Ej 12.2.2 Se desea encontrar un extremo para la funcional

$$\int_0^1 (x + u^2) dt$$

sujeto a $\dot{x} = -u$; $x(0) = 0$ y $x(1)$ libre. En este caso, el hamiltoniano es

$$H = x + u^2 - \lambda u$$

por lo que las condiciones de primer orden son,

$$\begin{aligned} H_u &= 2u - \lambda = 0, \\ \dot{\lambda} &= -H_x = -1, \\ \dot{x} &= H_\lambda = -u. \end{aligned}$$

De la segunda condición se tiene $\lambda(t) = -t + A$; sustituyendo en la primera, obtenemos $u(t) = \frac{1}{2}(-t+A)$ y finalmente de la tercera condición llegamos a $\dot{x} = \frac{1}{2}(t-A)$ con lo cual $x(t) = \frac{1}{4}t^2 - \frac{At}{2} + B$. Los valores de A y B se obtienen con la condición de transversalidad $\lambda(1) = 0$, que implica, $A = 1$ y la condición inicial $x(0) = 0$, que implica $B = 0$. Notemos que, dado que $H_{uu} = 2 > 0$, el extremo obtenido es un mínimo.

Ej 12.2.3 Se quiere resolver el siguiente problema:

$$\max \int_0^2 (2x - 3u) dt$$

sujeto a $\dot{x} = x + u$, $x(0) = 4$, $x(2)$ libre y $u(t) \in [0, 2]$. El hamiltoniano está dado por

$$H = 2x - 3u + \lambda(x + u),$$

o bien,

$$H = (2 + \lambda)x + (\lambda - 3)u.$$

Observemos que se trata de una función lineal en el control u ; dado que la región de control es el conjunto compacto $[0, 2]$, el valor u^* que maximiza H es

$$u^*(t) = \begin{cases} 2 & \text{si } \lambda > 3, \\ 0 & \text{si } \lambda < 3. \end{cases}$$

Si $\lambda = 3$, entonces u^* puede tomar cualquier valor en $[0, 2]$ pero vamos a ver que esto sólo sucede en un “instante” \bar{t} . Si existiera un intervalo de tiempo para el cual $\lambda = 3$, entonces tendríamos lo que se denomina un **problema de control singular** (más adelante se ve un ejemplo de esto). Asimismo, se debe satisfacer

$$\begin{aligned} \dot{\lambda} &= -H_x = -(2 + \lambda) \\ \dot{x} &= H_\lambda = x + u \end{aligned}$$

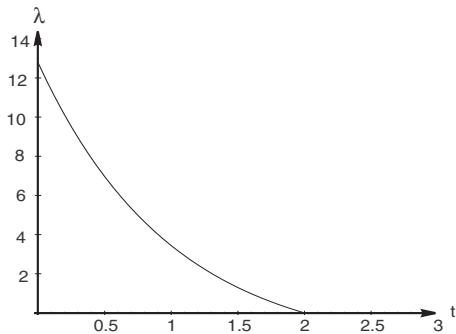
La primera de estas ecuaciones, junto con la condición de transversalidad $\lambda(2) = 0$, se resuelve fácilmente para obtener

$$\lambda(t) = 2e^{2-t} - 2,$$

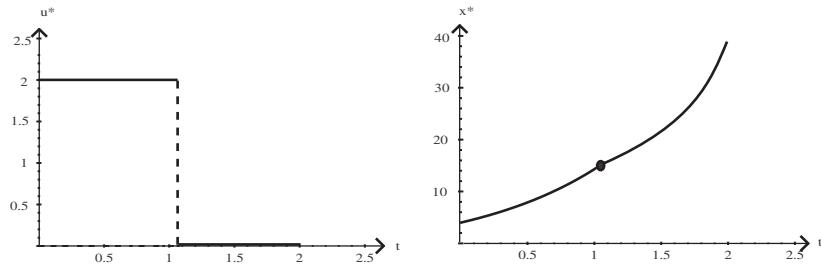
cuya gráfica se muestra en la figura 12.1.

El valor $\lambda = 3$ se alcanza cuando $\bar{t} = 2 - \ln \frac{5}{2} \simeq 1.08$. De esta forma, si $t \in [0, 1.08)$, entonces $u^* = 2$ y por lo tanto la ecuación de evolución del estado queda como

$$\dot{x} = x + 2,$$

Figura 12.1: Trayectoria de λ .

cuya solución, dada la condición inicial $x(0) = 4$, es $x^*(t) = 6e^t - 2$. De la misma forma, si $t \in [1.08, 2]$, se tiene que $u^* = 0$ y el estado óptimo queda como $x^*(t) = Ae^t$. Aquí no podemos usar la condición inicial dado que $0 \notin [1.08, 2]$; usamos, en cambio, para determinar el valor de A , la continuidad de la función x en $t = 1.08$, es decir, $x^*(1.08) = 6e^{1.08} - 2 = 15.668 = Ae^{1.08}$, y por lo tanto $A = 5.32$. Las gráficas de las trayectorias óptimas $u^*(t)$ y $x^*(t)$ se pueden ver en la figura 12.2.

Figura 12.2: Trayectorias para u y x .

A este tipo de problemas se los denomina problemas de “bang-bang”, dado que la variable de control solo toma (quizá de manera intermitente) sus valores extremos.



§12.3 Otras condiciones de transversalidad

La demostración del teorema 12.2.3 puede modificarse fácilmente para incluir el caso en el cual el tiempo final T está libre y en donde $x(T)$ puede o no estar especificada. La integral por partes del término

$\int_0^T \lambda \Delta \dot{x} dt = \int_0^T \lambda \dot{h} dt$ está dada por

$$\int_0^T \lambda \dot{h} dt = (\lambda h) |_0^T - \int_0^T \lambda \dot{h} dt,$$

o bien, tomando en cuenta que $h(0) = 0$,

$$\int_0^T \lambda \dot{h} dt = \lambda(T)h(T) - \int_0^T \lambda \dot{h} dt.$$

Recordando la figura 11.9, tenemos que $h(T) = \Delta x(T) - \dot{x}(T)\Delta T$ y, por lo tanto, el término $\lambda(T)h(T)$ puede ser reescrito como

$$\lambda(T)h(T) = \lambda(T)\Delta x(T) - \lambda(T)\dot{x}(T)\Delta T. \quad (12.15)$$

Utilicemos (12.15) y el teorema fundamental del cálculo para reescribir (12.12) como

$$\begin{aligned} \Delta J^* &= \int_0^T \Delta f^* dt + f^*(T)\Delta T \\ &= \int_0^T (\Delta f^* + \lambda \Delta g^* + (-f_x^* - \lambda g_x^*)\Delta x) dt + (f^*(T) + \lambda(T)\dot{x}^*(T))\Delta T - \lambda(T)\Delta x(T) \\ &= \int_0^T \Delta^u H^* dt + H^*(T)\Delta T - \lambda(T)\Delta x(T). \end{aligned} \quad (12.16)$$

Notemos que las condiciones de transversalidad que se deben aplicar son

$$\lambda(T) = 0 \text{ si } x(T) \text{ libre y } T \text{ dado}, \quad (12.17)$$

$$H^*(T) = 0 \text{ si } T \text{ libre y } x(T) \text{ dado}, \quad (12.18)$$

y se aplican ambas, en el caso en que tanto T como $x(T)$ están libres. La condición (12.17) es simplemente la condición terminal para $\lambda(T)$ dada en (12.5) (nótese que no es necesario especificarla ex ante ya que se obtiene de manera natural a partir de (12.16)).

Ejemplo

Ej 12.3.1 (Problema de tiempo óptimo) En ocasiones, el problema que se tiene puede resumirse como aquel de llegar de un estado inicial a uno final lo más pronto posible; en otras palabras, queremos resolver el siguiente problema:

$$\min T$$

sujeto a

$$\dot{x} = g(x, u, t), \quad x(0) = x_0, \quad x(T) = x_T, \quad u \in \overline{U} \text{ y } T \text{ libre.}$$

Notemos que $T = \int_0^T 1 dt$, de manera que el problema equivale a minimizar una funcional. Un ejemplo específico es el siguiente:

$$\min \int_0^T 1 dt$$

sujeto a

$$\dot{x} = x + u, \quad x(0) = 5, \quad x(T) = 11, \quad u \in [-1, 1] \text{ y } T \text{ libre.}$$

El hamiltoniano correspondiente es

$$H = 1 + \lambda(x + u),$$

y dado que éste es lineal en u se tiene,

$$u^*(t) = \begin{cases} -1 & \text{si } \lambda > 0, \\ 1 & \text{si } \lambda < 0. \end{cases} \quad (12.19)$$

Adicionalmente se cumplen

$$\dot{\lambda} = -H_x = -\lambda, \quad (12.20)$$

$$\dot{x}^* = x^* + u^* \quad (12.21)$$

y la condición de transversalidad

$$H^*(T) = 1 + \lambda(T)(x^*(T) + u^*(T)) = 0. \quad (12.22)$$

Para poder proseguir se necesita determinar λ . De 12.20 se obtiene

$$\lambda = Ae^{-t}, \quad (12.23)$$

y sustituyendo la condición final $x(T) = 11$ y (12.23) en (12.22) se llega a

$$1 + Ae^{-T}[11 + u^*(T)] = 0,$$

con lo cual,

$$A = -\frac{e^T}{11 + u^*(T)} < 0, \quad (12.24)$$

puesto que los únicos valores posibles para $u^*(T)$ son 1 y -1. De esta forma, se tiene que $\lambda(t) < 0$ en todo momento y por lo tanto (12.19) implica que $u^*(t) = 1$. Utilizando (12.21) y la condición inicial $x(0) = 5$ llegamos a que $x^*(t) = 6e^t - 1$ y por lo tanto $x^*(T) = 6e^T - 1 = 11$. De aquí que $T = \ln 2$. Finalmente, sustituyendo estos valores en (12.24), se tiene que $A = -\frac{1}{6}$ y de aquí que $\lambda(t) = -\frac{1}{6}e^{-t}$, con lo cual queda resuelto el problema.



Una restricción usada con frecuencia es que el valor final de la variable de estado tenga una cota inferior dada por x_{\min} (por supuesto podríamos también considerar una cota superior). En este caso, si el tiempo terminal, T , está dado se tiene que una condición necesaria para que se verifique $\Delta J^* \leq 0$ y tengamos un máximo es que

$$-\lambda(T)\Delta x(T) \leq 0. \quad (12.25)$$

En este caso, $x^*(T)$ no está libre y debe satisfacer la restricción $x^*(T) - x_{\min} = \Delta x(T) \geq 0$, por lo que (12.25) implica que $\lambda(T) \geq 0$. Si $x^*(T) > x_{\min}$, entonces la restricción no está activa y el problema es igual al caso en que $x(T)$ está libre y se cumple $\lambda(T) = 0$. En resumen, la condición apropiada de transversalidad en este caso se puede enunciar como⁴

$$\begin{aligned} \lambda(T) &\geq 0, \quad x^*(T) \geq x_{\min}, \\ \lambda(T)[x^*(T) - x_{\min}] &= 0. \end{aligned} \quad (12.26)$$

En economía, esta restricción es sumamente utilizada para el caso $x_{\min} = 0$, ya que la variable de estado puede estar restringida a ser no negativa.

Ejemplo

Ej 12.3.2 Consideremos el siguiente problema:

$$\max \int_0^1 -u^2 dt$$

sujeto a

$$\dot{x} = x + u; \quad x(0) = 1; \quad x(1) \geq 3.$$

El hamiltoniano es

$$H = -u^2 + \lambda(x + u).$$

Las condiciones de primer orden quedan dadas por

$$\begin{aligned} H_u &= -2u^* + \lambda = 0, \\ \dot{\lambda} &= -H_x = -\lambda, \\ \dot{x}^* &= x^* + u^*, \end{aligned}$$

y las condiciones de transversalidad son

$$\begin{aligned} \lambda(1) &\geq 0, \quad x^*(1) \geq 3, \\ \lambda(1)[x^*(1) - 3] &= 0. \end{aligned}$$

⁴El lector observador reconocerá esta condición como una de holgura al igual que las que aparecen en las condiciones de Kuhn-Tucker

Resolviendo las ecuaciones tenemos que

$$\begin{aligned}\lambda(t) &= Ae^{-t}, \\ u^*(t) &= \frac{1}{2}Ae^{-t}, \\ x^*(t) &= -\frac{A}{4}e^{-t} + Be^t.\end{aligned}$$

La condición inicial $x(0) = 1$ implica $-\frac{A}{4} + B = 1$. Si $x^*(1) > 3$, entonces se debe tener que $\lambda(1) = 0$ y por lo tanto $A = 0$ y $B = 1$. Sin embargo, de ser así tendríamos que $x^*(1) = e < 3$, violando la condición $x^*(1) > 3$. Se debe tener, entonces, que $x^*(1) = 3$ y el problema se resuelve como uno en donde los valores iniciales y finales, así como el tiempo final, están dados. Más aún, $x^*(1) = 3 = -\frac{A}{4}e^{-1} + Be$ y $x^*(0) = -\frac{A}{4} + B = 1$; resolviendo para A y B , se tiene que $A = 0.48$ y $B = 1.12$.

◇◇◇◇◇◇◇◇◇◇

§12.4 Problemas con horizonte infinito

Al igual que en el cálculo en variaciones, los problemas con horizonte infinito presentan la dificultad de que la integral impropia puede no converger. Suponiendo que este problema está resuelto, las condiciones de transversalidad, en la mayoría de los casos, se obtienen simplemente tomando el límite de las condiciones usuales cuando $T \rightarrow \infty$. De esta forma, si el valor terminal de la variable de estado está dado, la condición pertinente es

$$\lim_{t \rightarrow \infty} H(t) = 0. \quad (12.27)$$

De manera similar, si el valor terminal de la variable de estado está libre, entonces adicionalmente se debe cumplir

$$\lim_{t \rightarrow \infty} \lambda(t) = 0, \quad (12.28)$$

o bien, si $\lim_{t \rightarrow \infty} x(t) \geq x_{\min}$, entonces la condición debe ser

$$\begin{aligned}\lim_{t \rightarrow \infty} \lambda(t) &\geq 0, \\ \lim_{t \rightarrow \infty} x(t) &\geq x_{\min}, \\ \lim_{t \rightarrow \infty} \lambda(t)[x(t) - x_{\min}] &= 0.\end{aligned} \quad (12.29)$$

La siguiente proposición es útil cuando se combina con (12.18) o (12.27).

Proposición 12.4.1 *Si H^* representa, como es usual, al hamiltoniano evaluado en el punto (x^*, u^*, λ, t) , entonces*

$$\frac{dH^*}{dt} = \frac{\partial H^*}{\partial t}.$$

Demostración

Simplemente desarrollemos la derivada total $\frac{dH^*}{dt}$ para obtener

$$\frac{dH^*}{dt} = \frac{\partial H^*}{\partial t} + \frac{\partial H^*}{\partial x} \dot{x}^* + \frac{\partial H^*}{\partial u} \dot{u}^* + \frac{\partial H^*}{\partial \lambda} \dot{\lambda}$$

y notemos que $\frac{\partial H^*}{\partial u} = 0$; $\dot{\lambda} = -\frac{\partial H^*}{\partial x}$ y $\dot{x}^* = \frac{\partial H^*}{\partial \lambda}$, con lo cual se tiene que $\frac{dH^*}{dt} = \frac{\partial H^*}{\partial t}$, que es el resultado deseado. ■

El siguiente corolario es inmediato, pues si el problema es autónomo $\frac{\partial H^*}{\partial t} = 0$.

Corolario 12.4.2 *Si el problema es autónomo, entonces*

$$H(t) = c = \text{constante}.$$

En particular, si se trata de un problema en donde las condiciones de transversalidad son (12.18) o (12.27), entonces $H(t) = 0$ en todo momento.

Ejemplo

Ej 12.4.1 (Halkin) Como ya se mencionó, existe cierto grado de controversia en lo que respecta a las condiciones de transversalidad con horizontes infinitos, particularmente en lo que se refiere a la condición (12.28) en la cual el estado terminal está libre. El “contraejemplo” más conocido es aquel construido por Halkin,⁵ el cual exponemos a continuación.

Considérese el problema

$$\max \int_0^\infty (1-x)udt$$

sujeto a

$$\dot{x} = (1-x)u, \quad x(0) = 0 \quad \text{y} \quad u(t) \in [0, 1].$$

Nótese que la función en el integrando es idéntica a la de la ecuación de evolución del estado, con lo cual se puede reescribir la integral como

$$\int_0^\infty \dot{x}dt = \lim_{t \rightarrow \infty} x(t) - x(0) = \lim_{t \rightarrow \infty} x(t),$$

de manera que la función objetivo depende únicamente del estado terminal. A este tipo de problemas se los denomina **problemas de control terminal**. Resolviendo la ecuación de evolución del estado⁶ se obtiene que

$$x(t) = Ae^{-G(t)} + 1,$$

⁵Ver [Hal74].

⁶Notemos que es una ecuación lineal no homogénea y con coeficientes variables. El método utilizado es con un factor de integración como se describe en la sección 2.2.4.

en donde $G(t) = \int_0^t u(s)ds..$ La condición inicial implica $A = -1$ y por lo tanto

$$x(t) = -e^{-G(t)} + 1.$$

Dado que $u(t) \in [0, 1]$, $G(t) > 0$ y $e^{-G(t)} \in (0, 1]$, con lo cual $x(t) \in [0, 1]$. Se observa que el máximo se obtiene cuando $\lim_{t \rightarrow \infty} x(t) = 1$ y esto sucede siempre y cuando $\lim_{t \rightarrow \infty} G(t) \rightarrow \infty$. Existe una infinidad de funciones de control que logran este objetivo, por ejemplo $u^* = c$, en donde $c \in (0, 1]$. Si tomamos a $c \in (0, 1)$ entonces se trata de una solución interior y debe cumplirse la condición $H_u = 0$. Alternativamente, como

$$H = (1 - x)u + \lambda(1 - x)u,$$

se tiene que

$$H_u = (1 + \lambda)(1 - x) = 0,$$

y como $x < 1$, se concluye que $\lambda(t) = -1$. Esto último viola la condición de transversalidad (12.28). Se puede argumentar que el error está en suponer que el estado terminal está libre, ya que implícitamente está determinado como su cota superior. Se puede resolver el problema si, suponiendo esto, tomamos como la condición de transversalidad pertinente $\lim_{t \rightarrow \infty} H^* = 0$ y dado que el problema es autónomo suponemos $H^*(t) = 0$. El hamiltoniano es lineal en u y consecuentemente escogemos alguno de los valores extremos como solución. La solución dada por $u^* = 0$ se puede descartar con base en que tendríamos que $\dot{x} = 0$ y el estado sería $x = 0$ en todo momento. Escogemos pues $u^* = 1$, con lo cual se tienen las ecuaciones

$$\begin{aligned}\dot{\lambda} &= 1 + \lambda, \\ \dot{x} &= 1 - x\end{aligned}$$

y por lo tanto

$$\begin{aligned}\lambda(t) &= Ae^t - 1, \\ x^*(t) &= Be^{-t} + 1.\end{aligned}$$

La condición inicial $x(0) = 0$ implica $B = -1$. Observemos que la condición $H^*(t) = 0$ se satisface únicamente si $\lambda = -1$, por lo que se debe tener $A = 0$. Así, el estado x^* efectivamente converge a 1 y $\lambda = -1$ no representa una violación de la condición de transversalidad.



§12.5 Hamiltoniano en tiempo corriente

Hemos observado que, en los problemas de economía, con frecuencia la funcional objetivo es del tipo

$$\int_0^T f(x, u)e^{-\rho t} dt,$$

en donde $e^{-\rho t}$ representa un factor de descuento y la ecuación de evolución está dada por

$$\dot{x} = g(x, u).$$

En otras palabras, el problema es “casi autónomo” y el tiempo aparece de forma independiente sólo mediante el factor de descuento. En estos casos, es conveniente definir el *hamiltoniano en tiempo corriente* para resolver el problema como un problema autónomo.

Definición 12.5.1 *Supongamos que se tiene el siguiente problema de control:*

$$\max \int_0^T f(x, u) e^{-\rho t} dt,$$

sujeto a

$$\dot{x} = g(x, u), \quad x(0) = x_0,$$

y alguna condición terminal para T o $x(T)$. Si H denota al hamiltoniano y λ a la variable de coestado asociados, se define el **hamiltoniano en tiempo corriente** \bar{H} como

$$\bar{H} = H e^{\rho t}$$

y la variable de coestado en tiempo corriente como

$$\bar{\lambda} = \lambda e^{\rho t}.$$

Las condiciones dadas por el principio del máximo se pueden reexpresar en términos de \bar{H} como sigue.

Proposición 12.5.2 *Las condiciones (12.3) y (12.4) del teorema 12.2.3 se pueden reescribir como*

$$\frac{d\bar{\lambda}}{dt} = -\bar{H}_x + \bar{\lambda}\rho, \tag{12.30}$$

$$\dot{x} = \bar{H}_{\bar{\lambda}} \tag{12.31}$$

y

$$\bar{\lambda}(T)e^{-\rho T} = 0 \text{ (equivalentemente } \bar{\lambda}(T) = 0\text{).} \tag{12.32}$$

Adicionalmente,

$$H_u = 0 \iff \bar{H}_u = 0. \tag{12.33}$$

Demostración

El hamiltoniano asociado al problema está dado por

$$H = f(x, u)e^{-\rho t} + \lambda g(x, u),$$

con lo cual, suponiendo una solución interior, las condiciones de primer orden son

$$\begin{aligned} H_u &= 0, \\ \dot{\lambda} &= -H_x = -f_x e^{-\rho t} - \lambda g_x, \\ \dot{x} &= H_\lambda = g \end{aligned}$$

y la condición de transversalidad, si T está dado y $x(T)$ está libre es

$$\lambda(T) = 0.$$

Es claro, de la definición de \bar{H} , que (12.33) se cumple; asimismo,

$$\begin{aligned} -H_x &= -\bar{H}_x e^{-\rho t}, \\ \dot{\lambda} &= \frac{d\bar{\lambda}}{dt} e^{-\rho t} - \bar{\lambda} \rho e^{-\rho t}, \end{aligned}$$

por lo que (12.30) y (12.31) también son inmediatas. Finalmente, (12.32) se sigue de la definición de $\bar{\lambda}$. La condición de transversalidad análoga cuando $x(T)$ está dado y T está libre es simplemente

$$\bar{H}(T)e^{-\rho T} = 0. \quad (12.34)$$

Con esto concluimos la demostración. ■

El uso del hamiltoniano y la variable de coestado en tiempo corriente es tan frecuente en economía que en la mayoría de los casos ni siquiera se aclara que se están utilizando. Para simplificar la notación, de aquí en adelante simplemente diremos que se trata del hamiltoniano en tiempo corriente y lo denotaremos por H y a la variable de coestado correspondiente por λ .

Ejemplos

Ej 12.5.1 En el modelo de Ramsey,

$$\max \int_0^\infty u(c) e^{-\rho t} dt$$

sujeto a

$$\dot{k} = f(k) - \delta k - c, \quad k(0) = k_0, \quad k \geq 0,$$

el hamiltoniano en tiempo corriente es

$$H = u(c) + \lambda(f(k) - \delta k - c).$$

Las condiciones de primer orden son

$$\begin{aligned} H_c &= u' - \lambda = 0, \\ \dot{\lambda} &= -H_k + \lambda \rho = -\lambda(f' - \delta - \rho), \\ \dot{k} &= H_\lambda = f(k) - \delta k - c, \end{aligned}$$

y las condiciones de transversalidad son

$$\lim_{t \rightarrow \infty} \lambda(t)k(t)e^{-\rho t} = 0,$$

$$\lim_{t \rightarrow \infty} H(t)e^{-\rho t} = 0.$$

Observamos que de aquí se obtiene el sistema

$$\dot{k} = f(k) - \delta k - c,$$

$$\dot{c} = -\frac{u''}{u'}(f' - \delta - \rho),$$

que es el mismo que se tenía anteriormente. La diferencia es que aquí, el valor sombra del capital $\lambda(t)$ está dado en unidades del tiempo t .

Ej 12.5.2 Consideremos el siguiente problema

$$\max \int_0^{10} (x - u^2)e^{-t} dt$$

sujeto a

$$\dot{x} = x + u, \quad x(0) = \frac{1}{2}.$$

El hamiltoniano normal (en tiempo presente) es

$$H = (x - u^2)e^{-t} + \lambda(x + u),$$

por lo que se tienen las condiciones de primer orden

$$H_u = -2ue^{-t} + \lambda = 0,$$

$$\dot{\lambda} = -H_x = -e^{-t} - \lambda,$$

$$\dot{x} = x + u,$$

y la condición de transversalidad $\lambda(10) = 0$. La ecuación

$$\dot{\lambda} + \lambda = -e^{-t}$$

se resuelve multiplicando ambos lados por el factor de integración e^t e integrando, de manera que

$$\lambda e^t = -t + A$$

y por lo tanto

$$\lambda^* = Ae^{-t} - te^{-t}.$$

La condición de transversalidad $\lambda(10) = 0$ implica $A = 10$. Ahora se puede obtener

$$u^* = \frac{\lambda}{2}e^t$$

$$= \frac{1}{2}(10 - t),$$

y finalmente se resuelve la ecuación

$$\dot{x} - x = \frac{1}{2}(10 - t)$$

multiplicando por el factor de integración e^{-t} de ambos lados e integrando para obtener

$$x^* = -\frac{9}{2} + \frac{t}{2} + Be^t.$$

De la condición inicial $x(0) = \frac{1}{2}$ se tiene $B = 5$. Puesto que $H_{uu} = -2e^{-t} < 0$, la solución es, en efecto, un máximo.

Resolvamos ahora el mismo problema con el hamiltoniano en tiempo corriente, teniendo en cuenta que el factor de descuento es $\rho = 1$. Se tiene que

$$\bar{H} = x - u^2 + \bar{\lambda}(x + u)$$

y las condiciones de primer orden son

$$\begin{aligned}\bar{H}_u &= -2u + \bar{\lambda} = 0, \\ \dot{\bar{\lambda}} &= -\bar{H}_x + \bar{\lambda} = -1, \\ \dot{x} &= x + u.\end{aligned}$$

La condición de transversalidad ahora es $\bar{\lambda}(10)e^{-10} = 0$ o simplemente $\bar{\lambda}(10) = 0$. Ahora es trivial resolver para $\bar{\lambda}$ y se tiene

$$\bar{\lambda}^* = -t + a.$$

La condición de transversalidad implica $a = 10$. Podemos obtener el control u como

$$u^* = \frac{\bar{\lambda}}{2} = \frac{1}{2}(10 - t),$$

e igual que antes se resuelve la ecuación

$$\dot{x} - x = \frac{1}{2}(10 - t),$$

para obtener

$$x^* = -\frac{9}{2} + \frac{t}{2} + 5e^t.$$

Como se puede observar, las trayectorias del control y del estado son las mismas y $\bar{\lambda}e^{-t} = \lambda$; sin embargo, la resolución del problema es más sencilla utilizando el hamiltoniano en tiempo corriente.



§12.6 Problemas con más de una variable

En general podemos tener un vector $\mathbf{x} = (x_1, \dots, x_n)$ de variables de estado, diferenciables por pedazos, y otro, $\mathbf{u} = (u_1, \dots, u_m)$ de variables de control, continuas por pedazos, con lo cual el problema es

$$\max \int_0^T f(\mathbf{x}, \mathbf{u}, t) dt$$

sujeto a

$$\begin{aligned}\dot{x}_1 &= g_1(\mathbf{x}, \mathbf{u}, t), \\ &\vdots \\ \dot{x}_n &= g_n(\mathbf{x}, \mathbf{u}, t), \\ \mathbf{x}(0) &= \mathbf{x}_0, \quad T \text{ dado y } \mathbf{x}(T) \text{ libre,}\end{aligned}\tag{12.35}$$

en donde f y g_i son funciones con dominio en $\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}$ con todas sus derivadas continuas.

Salvo por algunas variantes notacionales, el principio del máximo se aplica de manera semejante al teorema 12.2.3. Supongamos que \mathbf{u}^* y \mathbf{x}^* resuelven el problema (12.35), entonces existe $\lambda(t) = (\lambda_1(t), \dots, \lambda_n(t))$ continua tal que la función

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

donde $\mathbf{g} = (g_1, \dots, g_n)$ posee un máximo en \mathbf{u}^* . Adicionalmente, $\lambda(t)$, $\mathbf{u}^*(t)$ y $\mathbf{x}^*(t)$ resuelven el sistema,

$$\begin{aligned}\dot{\lambda}_i &= -H_{x_i}, \\ \dot{x}_i &= H_{\lambda_i},\end{aligned}\tag{12.36}$$

para $i = 1, \dots, n$, y se cumple la condición de transversalidad

$$\lambda(T) = \mathbf{0}.\tag{12.37}$$

En general, la condición (12.37) se puede extender a $x_i(T)$ dado o bien $\lambda_i(T) = 0$ si $x_i(T)$ está libre para $i = 1, \dots, n$.

Ejemplo

Ej 12.6.1 Consideremos el siguiente problema:

$$\min \int_0^1 \frac{u^2}{2} dt$$

sujeto a

$$\begin{aligned}\dot{x}_1 &= x_2, \\ \dot{x}_2 &= u, \\ x_1(0) &= x_2(0) = 0, \\ x_1(1) &= 2, x_2(1) = 3.\end{aligned}$$

En este caso hay dos variables de estado y por lo tanto tendremos dos variables de coestado λ_1 y λ_2 . El hamiltoniano está dado por

$$H = \frac{u^2}{2} + \lambda_1 x_2 + \lambda_2 u$$

y las condiciones de primer orden quedan como,

$$\begin{aligned} H_u &= u + \lambda_2 = 0, \\ \dot{\lambda}_1 &= -H_{x_1} = 0, \\ \dot{\lambda}_2 &= -H_{x_2} = -\lambda_1, \\ \dot{x}_1 &= H_{\lambda_1} = x_2, \\ \dot{x}_2 &= H_{\lambda_2} = u. \end{aligned}$$

No es necesaria ninguna condición de transversalidad dado que tenemos todos los valores iniciales y finales especificados. Resolviendo, obtenemos $\lambda_1 = A$, $\lambda_2 = -At + B$, $u = At - B$, $x_2 = \frac{At^2}{2} - Bt + C$ y $x_1 = \frac{At^3}{6} - \frac{Bt^2}{2} + Ct + D$. Los valores iniciales y finales determinan las constantes: $A = B = -6$ y $C = D = 0$. Notemos que, en efecto, se trata de un mínimo puesto que $H_{uu} = 1 > 0$.

◇◇◇◇◇◇◇◇◇

§12.7 Interpretación económica del problema de control

Esta sección está basada en [Dor69], en donde las condiciones necesarias para resolver el problema de control óptimo se derivan de la llamada función valor. Este método se basa en el trabajo de Bellman ([Bel57],[BD62] y [BK65]) que paralelamente a Pontryagin desarrolló los principios de la teoría de control. Mostraremos con cierto detalle cómo se obtienen las condiciones de optimalidad pues éste es el desarrollo que usaremos posteriormente, al considerar el caso en el cual el tiempo es una variable discreta y no continua.

Consideremos el siguiente problema:

$$\max \int_0^T b(x, u, \tau) d\tau$$

sujeto a

$$\dot{x} = g(x, u, t), \quad (12.38)$$

con $x(0)$ dado y $x(T)$ libre. Como de costumbre, x es la variable de estado y u la de control. A la selección óptima del control la llamamos una **política óptima**.

El integrando representa una función de beneficios como las ganancias de una empresa o una función de utilidad. Podemos suponer que los beneficios se miden en unidades del tiempo inicial $t = 0$, es decir, los evaluamos en tiempo presente. Podríamos pensar que la función de beneficios es de la forma $b(x, u, \tau) = B(x, u)e^{-\rho\tau}$ para alguna tasa de descuento temporal ρ , pero el análisis del problema no cambiaría, de manera que trabajaremos con la función b .

Definamos la función valor $V(x, t)$ como sigue

$$V(x, t) = \max \int_t^T b(x, u, \tau) d\tau, \quad (12.39)$$

es decir, es el valor (presente) de los beneficios máximos a partir del tiempo t . Consideremos un intervalo de tiempo Δt , suficientemente pequeño como para que el control óptimo no cambie durante ese intervalo. Una observación no trivial, llamada el **principio de optimalidad de Bellman**, dice que

Una política óptima tiene la propiedad de que, cualquiera que sea la acción inicial, las posibles elecciones restantes constituyen una política óptima con respecto al subproblema que comienza en el estado determinado por las acciones iniciales.

Así, por ejemplo, si se ha determinado que la trayectoria óptima de México a Guadalajara es vía Toluca y Morelia, entonces la trayectoria óptima de Toluca a Guadalajara también pasa por Morelia.

En nuestro caso, podemos expresar el principio de optimalidad en términos de la función valor como⁷

$$\begin{aligned} V(x(t), t) &= \max \left\{ \int_t^{t+\Delta t} b(x, u, \tau) d\tau + \int_{t+\Delta t}^T b(x, u, \tau) d\tau \right\} \\ &= \max \left\{ b(x(t), u(t), t) \Delta t + \underbrace{V(x(t + \Delta t), t + \Delta t)}_{\text{óptimo del subproblema}} \right\}. \end{aligned} \quad (12.40)$$

Una solución interior óptima u^* de este problema de maximización debe satisfacer

$$\frac{\partial}{\partial u} [b(x(t), u(t), t) \Delta t + V(x(t + \Delta t), t + \Delta t)] = 0. \quad (12.41)$$

Para evaluar esta derivada es necesario conocer $\frac{\partial V[x(t+\Delta t), t+\Delta t]}{\partial u}$. Dado que $x(t + \Delta t) \simeq x(t) + \dot{x}(t)\Delta t$ y $\dot{x} = g(x, u, t)$ se tiene que

$$\frac{\partial x(t + \Delta t)}{\partial u} = \frac{\partial g}{\partial u} \Delta t$$

y por lo tanto

⁷El segundo término de la expresión es claro pues representa el óptimo del subproblema a partir de $t + \Delta t$. El primero se obtiene si interpretamos a

$$\int_t^{t+\Delta t} b(x, u, \tau) d\tau$$

como $\gamma(t + \Delta t) - \gamma(t)$, en donde

$$\gamma(t) = \int_a^t b(x, u, \tau) d\tau.$$

De aquí, la aproximación lineal de Taylor nos da

$$\gamma(t + \Delta t) - \gamma(t) \simeq \gamma'(t) \Delta t,$$

y por el teorema fundamental del cálculo

$$\gamma'(t) = b(x(t), u(t), t),$$

con lo cual se obtiene el resultado deseado.

$$\begin{aligned}\frac{\partial V[x(t + \Delta t), t + \Delta t]}{\partial u} &= \frac{\partial V[x(t + \Delta t), t + \Delta t]}{\partial x(t + \Delta t)} \frac{\partial x(t + \Delta t)}{\partial u} \\ &= \frac{\partial V[x(t + \Delta t), t + \Delta t]}{\partial x(t + \Delta t)} \frac{\partial g}{\partial u} \Delta t.\end{aligned}\quad (12.42)$$

Sustituyendo (12.42) en (12.41) y reescribiendo llegamos a que

$$\left[b_u + \frac{\partial V[x(t + \Delta t), t + \Delta t]}{\partial x(t + \Delta t)} g_u \right] \Delta t = 0,$$

con lo cual, definiendo λ como

$$\lambda(t) = \frac{\partial V(x, t)}{\partial x}, \quad (12.43)$$

se concluye

$$b_u + \lambda(t + \Delta t) g_u = 0. \quad (12.44)$$

A la variable $\lambda(t)$ se la conoce también como **valor o precio sombra** del estado. Se puede pensar como el valor, o contribución marginal, a la función valor de la unidad adicional de x en el tiempo t , en unidades del tiempo inicial $t = 0$. La ecuación (12.44) puede interpretarse como sigue: a lo largo de la trayectoria óptima del control se tiene que, en cada instante, la contribución marginal de un cambio en el control se contrarresta exactamente con el efecto de dicho cambio en el valor del estado al instante siguiente.

Supongamos que se encuentra u^* que satisface (12.44). Sustituyendo en (12.40) se tiene que

$$V(x, t) = b(x(t), u^*(t), t)\Delta t + V(x^*(t + \Delta t), t + \Delta t), \quad (12.45)$$

en donde

$$x^*(t + \Delta t) \simeq x(t) + \dot{x}(t)\Delta t = x(t) + g(x, u^*, t)\Delta t.$$

Derivemos (12.45) con respecto a x y utilicemos (12.43) y la aproximación de primer orden $\lambda(t + \Delta t) \simeq \lambda(t) + \dot{\lambda}\Delta t$ para obtener

$$\begin{aligned}\lambda(t) &= b_x \Delta t + \lambda(t + \Delta t)(1 + g_x \Delta t) \\ &\simeq b_x \Delta t + (\lambda(t) + \dot{\lambda}\Delta t)(1 + g_x \Delta t).\end{aligned}$$

Despreciando el término de segundo orden dado por $(\Delta t)^2$, dividiendo entre Δt y reescribiendo se tiene

$$\dot{\lambda} = -(b_x + \lambda g_x). \quad (12.46)$$

La interpretación de esta ecuación⁸ es la siguiente: $\dot{\lambda}$ es la tasa de cambio del valor del estado, así que, a lo largo de la trayectoria óptima, la disminución en el valor de una unidad de x entre t y $t + \Delta t$ es la suma del beneficio marginal b_x en t y su contribución al valor del estado al final del periodo dada por λg_x . Si, por ejemplo, x es capital que se usa para la producción, entonces su valor cae a la misma tasa a la cual genera producción, ya sea ésta utilizable de inmediato o en forma de inversión para su uso futuro.

⁸Nótese que es una ecuación diferencial parcial para la función valor V , ya que $\lambda = \frac{\partial V}{\partial x}$.

Las ecuaciones (12.44) y (12.46) se pueden obtener, como ya vimos antes, construyendo el hamiltoniano

$$H = b(x, u, t) + \lambda(t)g(x, u, t),$$

en donde λ tiene la interpretación dada arriba, y aplicando el principio del máximo. La ecuación (12.44) es simplemente $H_u = 0$ y (12.46) es la ecuación de movimiento de la variable de coestado.

El primer término en H representa los beneficios en el instante t y el segundo término, el valor del capital acumulado en dicho instante, que producirá beneficios en el futuro inmediato. La trayectoria óptima del estado y del control deben ser tales que maximicen no sólo los beneficios representados en H , sino también aquellos que surgen por el cambio en el valor del estado. Esto podemos representarlo como la elección óptima de x y u de manera que se maximiza el **hamiltoniano modificado**⁹ dado por

$$\mathcal{H} = H + \dot{\lambda}x. \quad (12.47)$$

Esto se logra si $\mathcal{H}_u = 0$ y $\mathcal{H}_x = 0$, que equivalen a las ecuaciones (12.44) y (12.46); adicionalmente debe tenerse la ecuación de evolución del estado dada por (12.38). La variable de estado no está sujeta a elección pues está determinada por su valor inicial junto con la ecuación de evolución; sin embargo, el control u y el valor sombra λ se escogen de manera que el problema sea equivalente a maximizar \mathcal{H} con respecto a u y x .

La condición de transversalidad si $x(T)$ está libre es $\lambda(T) = 0$; es decir, que el valor terminal del estado, llevado a tiempo presente, debe ser nulo en la trayectoria óptima: no sería óptimo quedarnos con algo de valor cuando se “acabara el mundo”, por lo menos en términos del problema, en el tiempo T . Si se tiene por ejemplo, x es capital y se debe cumplir $x(T) \geq 0$, entonces la condición de transversalidad es

$$\lambda(T) \geq 0, \lambda(T)x(T) = 0.$$

Esto es, el valor terminal total en tiempo presente del capital (estado) debe ser nulo y su valor sombra no puede ser negativo. La conclusión es intuitivamente clara ya que el capital siempre tiene un valor no negativo y en el tiempo terminal podemos, asignarle un valor igual a 0 o bien terminar con una cantidad nula de capital.

Si el capital (estado) final está especificado pero el tiempo en el cual llegar a él está libre, entonces la condición de transversalidad es $H(T) = 0$. Esto nos dice que el tiempo óptimo T para llegar a nuestra meta debe ser tal que los beneficios en T y en el futuro inmediato sean nulos. De no ser así, sería óptimo tomar más tiempo para llegar al estado final.

Notamos que si se trata de un hamiltoniano en tiempo corriente, simplemente interpretamos $\lambda(t)$ como el valor sombra del estado en unidades del tiempo t . El hamiltoniano representa, como antes, la suma de los beneficios en t más aquellos del futuro inmediato pero ahora éstos se miden en unidades del tiempo corriente t .

⁹Notemos que

$$\mathcal{H} = b(x, u, t) + \frac{d\lambda k}{dt},$$

es decir, se trata de los beneficios “estáticos” en t , o sea $b(x, u, t)$, más los beneficios “dinámicos” dados por el cambio en el valor (sombra) total del capital, es decir $\frac{d\lambda k}{dt}$.

§12.8 Dos aplicaciones

§12.8.1 Un modelo monetario

Existen varias maneras de obtener la función de demanda de dinero con base en fundamentos microeconómicos. Una que resulta analíticamente sencilla es considerar como bienes a los servicios que ofrecen los balances reales, para así introducirlos en la función de utilidad. El siguiente es un modelo basado en uno sumamente popular introducido por Sidrauski [Sid67], que es una simple extensión del modelo de Ramsey. Definamos las siguientes variables y funciones:

- c = consumo,
- M = balances nominales,
- p = índice de precios,
- $m = \frac{M}{p}$ = balances reales,
- $\pi = \frac{\dot{p}}{p}$ = tasa de inflación,
- w = ingreso laboral en términos reales (salario real),
- r = tasa real,
- k = capital por familia,
- T = transferencia gubernamental a través del Banco Central,
- $u(c, m)$ función de utilidad con $u_c, u_m > 0; u_{cc}, u_{mm} < 0; u_{cm} = 0$,
- $f(k)$ función de producción con $f' > 0; f'' < 0$.

Supongamos que se tiene una economía poblada de familias idénticas e inmortales. Cada familia provee una unidad de trabajo recibiendo a cambio un salario w ; asimismo, la riqueza se puede tener en forma de capital (que no se deprecia) o balances reales. Adicionalmente el Banco Central imprime dinero fiduciario (con un costo nulo) y los bienes (reales) que obtiene a cambio los regresa a las familias en forma de una transferencia T .

En términos reales, la restricción presupuestal de la familia representativa está dada por

$$w + rk + T = c + \dot{k} + \frac{\dot{M}}{p}, \quad (12.48)$$

en donde $\frac{\dot{M}}{p}$ representa la “inversión” en balances nominales expresada en términos reales.

Notemos que

$$\dot{m} = \frac{\dot{M}}{p} - \frac{M}{p} \frac{\dot{p}}{p} = \frac{\dot{M}}{p} - m\pi,$$

de manera que sustituyendo para $\frac{\dot{M}}{p}$ en (12.48) la restricción presupuestal se puede reescribir como

$$w + rk + T = c + \dot{k} + \dot{m} + m\pi. \quad (12.49)$$

Es claro de (12.49) que los balances reales se pueden interpretar como un activo que se deprecia a una tasa igual a la inflación π .

El problema que resuelve la familia representativa en $t = 0$, suponiendo que ρ es la tasa subjetiva de descuento, es el siguiente:

$$\max \int_0^{\infty} u(c, m) e^{-\rho t} dt \quad (12.50)$$

sujeto a (12.49) con $k(0) = k_0$ y $m(0) = m_0$.

La variable de control es el consumo c y las variables de estado son k y m . Para expresar el problema en la forma estándar se define un control auxiliar μ , de manera que $\dot{m} = \mu$. Así el problema consiste en maximizar (12.50) sujeto a

$$\begin{aligned} \dot{k} &= w + rk + T - c - \mu - m\pi, \\ \dot{m} &= \mu, \\ k(0) &= k, \quad m(0) = m_0. \end{aligned} \quad (12.51)$$

El hamiltoniano asociado en tiempo corriente es

$$H = u(c, m) + \lambda_1(w + rk + T - c - \mu - m\pi) + \lambda_2\mu.$$

Por lo tanto las condiciones de primer orden son

$$\begin{aligned} H_c &= u_c - \lambda_1 = 0, \\ H_{\mu} &= -\lambda_1 + \lambda_2 = 0, \\ \dot{\lambda}_1 &= -H_k + \lambda_1\rho = -\lambda_1r + \lambda_1\rho, \\ \dot{\lambda}_2 &= -H_m + \lambda_2\rho = -u_m + \lambda_1\pi + \lambda_2\rho, \end{aligned}$$

junto con la ecuación de evolución (12.51) para k . Dado que $\lambda_1 = \lambda_2 = u_c$, podemos reescribir las ecuaciones de evolución de las variables de coestado como

$$\frac{\dot{u}_c}{u_c} = \rho - r, \quad (12.52)$$

$$\frac{\dot{u}_c}{u_c} = -\frac{u_m}{u_c} + \pi + \rho. \quad (12.53)$$

El lector reconocerá (12.53) como la ecuación (5.12) dada en la sección 5.7.2, en donde se propuso de manera intuitiva, sin fundamentos microeconómicos.

El problema del gobierno es simplemente satisfacer su restricción presupuestal suponiendo que transfiere todos los recursos obtenidos por la creación de dinero. Se debe cumplir entonces

$$T = \dot{m} + m\pi, \quad (12.54)$$

donde T es la transferencia que se otorga al hogar representativo. De esta forma, la restricción presupuestal de la economía agregada se obtiene sustituyendo (12.54) en (12.49) y queda como

$$\dot{k} = w + rk - c. \quad (12.55)$$

La ecuación (12.52) junto con (12.55) forman un sistema en k y c , puesto que por hipótesis u_c es independiente de m . Si r y w se obtienen mediante la resolución del problema de maximización de la empresa, se puede llegar al modelo de Ramsey. En este caso se dice que el dinero es **superneutral** ya que no influye en la parte real de la economía. Para recuperar el modelo de Ramsey, supongamos que se tiene una función de producción $F(K, L)$ linealmente homogénea, tal que $F_K, F_L > 0$; $F_{KK}, F_{LL} < 0$, y $F_{KL} > 0$. Aquí L representa el número de familias y $K = kL$ es el capital total. Aplicando a la función F el teorema de Euler¹⁰ para funciones homogéneas, se tiene que

$$F_L L + F_K K = F,$$

y por lo tanto la producción por familia f está dada por

$$F_L + F_K k = f.$$

Como $f(k) = \frac{F(K, L)}{L} = F(k, 1)$, derivando con respecto a K obtenemos que $f' \frac{1}{L} = F_K \frac{1}{L}$ de manera que $f' = F_K$. La producción de la firma optimizadora cumple con $F_L = w$ y $F_K = f' = r$. De esta forma se tiene que $w = f - f'k$ y sustituyendo para w y r en (12.55) se obtiene

$$\dot{k} = f(k) - c.$$

Esta última ecuación, junto con la ecuación (12.52), son exactamente las mismas que las obtenidas en el modelo de Ramsey suponiendo que el capital no se deprecia ($\delta = 0$).

Por otro lado, la función de demanda de dinero se obtiene combinando las ecuaciones (12.52) y (12.53) para obtener

$$u_m = u_c(r + \pi). \quad (12.56)$$

Esta ecuación nos dice que la demanda de balances reales aumenta al incrementarse el consumo c y disminuye al aumentar la tasa de interés nominal $r + \pi$. Vistos los balances reales como un bien, la tasa marginal de substitución $\frac{u_m}{u_c}$ es igual al costo de oportunidad relativo de los bienes, que en este caso es $\frac{r+\pi}{1} = r + \pi$, que es la condición usual de optimalidad.¹¹

Las condiciones de transversalidad, suponiendo que los valores terminales de k y m son no negativos, resultan

$$\begin{aligned} \lim_{t \rightarrow \infty} \lambda_1(t) e^{-\rho t} k &= 0, \\ \lim_{t \rightarrow \infty} \lambda_2(t) e^{-\rho t} m &= 0, \\ \lim_{t \rightarrow \infty} H(t) e^{-\rho t} &= 0, \end{aligned}$$

¹⁰El teorema de Euler dice que si una función $F : \mathbb{R}^n \rightarrow \mathbb{R}$ de clase C^1 es homogénea de grado n , entonces se cumple $\mathbf{x} \cdot \nabla F(\mathbf{x}) = nF(\mathbf{x})$.

¹¹La tasa nominal $r + \pi$ puede interpretarse como el costo de los balances reales y dividirse en dos partes: π es el costo debido a la depreciación (inflación) y r es el costo de oportunidad de tener balances reales m en lugar de capital, k .

y éstas se satisfacen cuando k y c convergen a lo largo de la variedad estable hacia su punto de equilibrio.

§12.8.2 Una economía pequeña y abierta

En el modelo de Ramsey, el consumo más la inversión siempre deben ser iguales a la producción doméstica total. Aun en el caso en que permitiéramos que el agente representativo tuviera algún otro activo o pasivo, digamos a , en el agregado la deuda total interna debería ser nula y por lo tanto, en equilibrio, el agente representativo poseería $a = 0$. Las cosas cambian, sin embargo, si a es un activo o pasivo extranjero ya que se pueden tener déficits o superávits temporales en la balanza de pagos.

Supongamos que el país casa es pequeño y, dado el proceso de globalización, posee una economía abierta. Se considera también un país extranjero, grande y estable, que está en equilibrio y posee un activo a con tasa de rendimiento real r . El país casa tiene acceso al activo extranjero que puede ser positivo (activo) o negativo (pasivo). La tasa de cambio \dot{a} es la inversión en este activo y, finalmente, I es la inversión bruta en capital doméstico k .

El problema que debe resolver la familia representativa, suponiendo un horizonte infinito, es el siguiente:

$$\max \int_0^\infty u(c)e^{-\rho t} dt$$

sujeto a la restricción presupuestal

$$f(k) + ar = c + I + \dot{a},$$

donde $\dot{k} + \delta k = I$, $k(0)$ y $a(0)$ están dados. La notación y las funciones son iguales que en el modelo de Ramsey.

Así las cosas nada limita a los agentes domésticos en su acumulación o desacumulación de a . Esto no es razonable ya que, por ejemplo, las familias podrían vivir eternamente excediendo sus posibilidades con un déficit en la balanza de pagos y simplemente pagarán su deuda con nuevos préstamos. La forma de evitar esta situación es limitando la acumulación o desacumulación del activo a que sea menor que la tasa real, es decir, debe cumplirse

$$\lim_{t \rightarrow \infty} a(t)e^{-rt} = 0.$$

Esta condición es conocida como **no admitir juegos de Ponzi**. El nombre proviene de Charles Ponzi que vivió en Boston durante los años veinte del siglo pasado y fue condenado por el fraude de organizar esquemas de inversión tipo “pirámide”.¹²

Los controles son c e I y los estados a y k . El hamiltoniano en tiempo corriente asociado al problema es

$$H = u(c) + \lambda_1(f(k) + ar - c - I) + \lambda_2(I - \delta k).$$

¹²Probablemente el lector ha sido alguna vez víctima de un juego de Ponzi. El que comienza el juego pide dinero, digamos \$100, a n personas, cada una de las cuales hace lo mismo con otros n individuos y así sucesivamente. De esta forma funcionan las cadenas de cartas y las llamadas pirámides. En particular, Charles Ponzi emitió un tipo de bono llamado cupón postal y pagaba los rendimientos de los mismos con nueva deuda, es decir, con la venta de más cupones.

Las condiciones de primer orden son

$$\begin{aligned} H_c &= u' - \lambda_1 = 0, \\ H_I &= -\lambda_1 + \lambda_2 = 0, \\ \dot{\lambda}_1 &= -H_a + \lambda_1\rho = \lambda_1(\rho - r), \\ \dot{\lambda}_2 &= -H_k + \lambda_2\rho = -\lambda_1 f' + \lambda_2(\delta + \rho), \\ \dot{a} &= f(k) + ar - c - I, \\ \dot{k} &= I - \delta k. \end{aligned}$$

De aquí se concluye que $\lambda_1 = \lambda_2 = u'$ y por lo tanto las ecuaciones de evolución de ambos coestados se reescriben como

$$\begin{aligned} \frac{\dot{u}_c}{u_c} &= \rho - r, \\ \frac{\dot{u}_c}{u_c} &= \rho + \delta - f', \end{aligned}$$

con lo cual concluimos que $r = f' - \delta$, es decir, la tasa real neta del capital doméstico $f' - \delta$ debe ser igual a la tasa real del activo extranjero.

El resultado anterior no debe sorprendernos, ya que los precios sombra de los activos k y a son iguales. Surge, sin embargo, un problema con nuestras hipótesis: la condición $\frac{\dot{u}_c}{u_c} = \rho - r$ no es sostenible si $\rho \neq r$. La razón es simple: si $\rho > r$, el consumo decrece en el tiempo y, por lo tanto, el país casa tiende asintóticamente a un estado de inanición: la deuda externa se acumula a tal punto que su pago absorbe todos los recursos domésticos. Si $\rho < r$, la situación es inversa y el consumo crece en el tiempo: el país casa acumula los activos del resto del mundo. Así, en algún momento el país deja de ser una economía pequeña y la hipótesis de que r es constante ya no es sostenible. La única solución es suponer $\rho = r$. La justificación es la siguiente: el país extranjero (grande) está en equilibrio y por lo tanto la tasa real r es igual a la tasa subjetiva de descuento de sus familias. Las familias del país casa son esencialmente iguales a las extranjeras y por lo tanto su tasa subjetiva de descuento también debe ser r .

Si ahora tomamos $\rho = r$, entonces concluimos que $\frac{\dot{u}_c}{u_c} = 0$ y $f' = r + \delta$ a lo largo de la trayectoria óptima. Esto implica que no hay dinámica, ya que $f' = r + \delta$ determina un acervo de capital k igual a una constante k^* en todo momento. En el instante inicial, si $k(0) \neq k^*$, la familia instantáneamente reacomoda su cartera, consistente en capital y activos, comprando o vendiendo el activo a para obtener la cantidad óptima de capital k^* . Esto es posible ya que las familias son indiferentes entre k y a , pues sus precios sombra son iguales. La cantidad de activos en equilibrio está determinada por $k(0) + a(0) = k^* + a^*$.

La condición $\frac{\dot{u}_c}{u_c} = 0$ implica un nivel constante de consumo. Se tiene además que $\dot{a} = \dot{k} = 0$ en la restricción presupuestal; sustituyendo $a = a^*$, $k = k^*$ e $I = \delta k^*$ en la misma restricción, se determina el nivel de consumo c^* en equilibrio. Notemos que esto es algo parecido a lo que sucedía con el problema de la empresa, cuando no había costos de ajuste para la inversión, tal como se vio en la sección 11.7.1.

Para obtener una dinámica no trivial podemos también introducir una función de costos de ajuste para la inversión bruta I . Sea $h(I)$ la función de costos de ajuste de la inversión. Ésta es doblemente

diferenciable y satisface

$$\begin{aligned} h(0) &= 0, \\ h'(0) &= 0, \\ h''(I) &> 0. \end{aligned}$$

Estas condiciones implican que el costo de ajuste es convexo, con un mínimo igual a 0 en $I = 0$ y que es costoso tanto invertir como “desinvertir”.

El problema se reformula como

$$\max \int_0^\infty u(c)e^{-rt}dt$$

sujeto a

$$\begin{aligned} f(k) + ar &= c + h(I) + \dot{a}, \\ \dot{k} + \delta k &= I, \quad k(0) \text{ y } a(0) \text{ dados.} \end{aligned}$$

El hamiltoniano es ahora

$$H = u(c) + \lambda_1(f(k) + ar - c - h(I)) + \lambda_2(I - \delta k).$$

Las condiciones de primer orden son

$$H_c = u'(c) - \lambda_1 = 0, \tag{12.57}$$

$$H_I = -\lambda_1 h'(I) + \lambda_2 = 0, \tag{12.58}$$

$$\dot{\lambda}_1 = -H_a + \lambda_1 r = -\lambda_1 r + \lambda_1 r = 0, \tag{12.59}$$

$$\dot{\lambda}_2 = -H_k + \lambda_2 r = -\lambda_1 f'(k) + \lambda_2(r + \delta), \tag{12.60}$$

$$\dot{a} = f(k) + ar - c - h(I), \tag{12.61}$$

$$\dot{k} = I - \delta k \tag{12.62}$$

y las condiciones de transversalidad son

$$\lim_{t \rightarrow \infty} \lambda_1 e^{-rt} a = 0,$$

$$\lim_{t \rightarrow \infty} \lambda_2 e^{-rt} k = 0,$$

$$\lim_{t \rightarrow \infty} H(t) e^{-rt} = 0.$$

Definimos el precio sombra real del capital como el costo marginal de la inversión:

$$p \equiv \frac{\lambda_2}{\lambda_1} = h'(I).$$

Esta ecuación expresa implícitamente a la inversión I como función de p , es decir, $I = \varphi(p)$, y por el teorema de la función implícita tenemos que

$$\varphi'(p) = -\frac{1}{h'} > 0.$$

φ puede interpretarse como una función de oferta de inversión. La condición (12.59) implica que λ_1 es constante, por lo que $\dot{p} = \frac{\lambda_2}{\lambda_1}$. La ecuación (12.60), que da la evolución para λ_2 , puede reescribirse como

$$\dot{p} = p(r + \delta) - f'(k), \quad (12.63)$$

y combinando $I = \varphi(p)$ con (12.62) se tiene

$$\dot{k} = \varphi(p) - \delta k. \quad (12.64)$$

Estas dos ecuaciones forman un sistema de ecuaciones diferenciales en k y p .

Linealizamos el sistema alrededor del estado estacionario (k^*, p^*) , distinto del origen, que se obtiene de igualar (12.63) y (12.64) a 0, con lo que se llega a

$$\begin{pmatrix} \dot{k} \\ \dot{p} \end{pmatrix} = \begin{pmatrix} -\delta & \varphi'(p^*) \\ -f''(k^*) & r + \delta \end{pmatrix} \begin{pmatrix} k - k^* \\ p - p^* \end{pmatrix}. \quad (12.65)$$

La matriz del sistema tiene como determinante

$$-\delta(r + \delta) + f'(k^*)\varphi'(p^*) < 0$$

y por lo tanto existen dos valores propios reales de signos opuestos de manera que (k^*, p^*) es un punto silla. Es claro que las condiciones de transversalidad se satisfacen si la trayectoria de equilibrio coincide con la variedad estable del sistema.

En el estado estacionario la inversión es simplemente para reponer el capital depreciado: $I^* = \delta k^*$, y el precio sombra real del capital es tal que el costo marginal de éste (que incluye el costo de oportunidad pr y el costo por depreciación $p\delta$) es igual a su beneficio marginal: $p^*(r + \delta) = f'(k^*)$.

El diagrama de fase del sistema (12.65) se muestra en la figura 12.3.

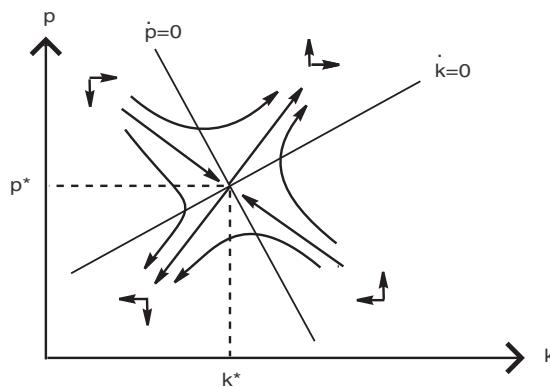


Figura 12.3: Diagrama de fase del sistema 12.65.

¿Cómo se determina el consumo? La relación (12.57), en donde λ_1 es igual a una constante, implica que se mantiene el mismo nivel de consumo c^* a lo largo de la trayectoria óptima. Para determinar este nivel reescribimos la restricción presupuestal como

$$\dot{a} - ar = f(k) - h(I) - c^*,$$

en donde las trayectorias para k e I son las óptimas, obtenidas al resolver el sistema dado por (12.63) y (12.64). Multiplicamos por el factor de integración e^{-rt} ambos lados de la restricción presupuestal e integramos de 0 a ∞ para obtener

$$(ae^{-rt})_{t=0}^{\infty} = \int_0^{\infty} [f(k) - h(I)]e^{-rt}dt - c^* \int_0^{\infty} e^{-rt}.$$

Utilizando la condición de no juego de Ponzi y simplificando se obtiene finalmente

$$c^* = r \int_0^{\infty} [f(k) - h(I)]e^{-rt}dt + ra_0, \quad (12.66)$$

con lo cual queda determinado el nivel de consumo. Esta ecuación establece que el valor presente del consumo acumulado en el periodo $[0, \infty)$ debe ser igual a la suma del valor presente de la producción neta en el mismo periodo más la cantidad inicial de activos a_0 , es decir, el valor presente de la riqueza total en $t = 0$.

Si, por ejemplo, $a_0 = 0$ y la trayectoria óptima de la producción neta está dada como en la figura 12.4, entonces el nivel de consumo c^* es tal que el valor presente del área sombreada es nulo. Notemos que cuando $c^* > f(k) - h(I)$ se tiene un déficit en la balanza de pagos y se acumula deuda ($a < 0$) en dicho periodo. Posteriormente, cuando $c^* < f(k) - h(I)$ se tiene un superávit en la balanza de pagos y en ese periodo se paga la deuda adquirida anteriormente. La situación, por supuesto, podría ser a la inversa.

Cuando no hay costos de ajuste, o sea $h = 0$ y $p = 1$, se tiene el caso en donde la dinámica es trivial. La dinámica aparece cuando $p \neq 1$, es decir, cuando el capital y los activos extranjeros tienen un precio sombra distinto y la movilidad del capital no es perfecta.

Este tipo de modelos es de utilidad para el análisis, tanto de estática comparativa como de ajuste dinámico, con cambios en los parámetros. Es tradicional denominar *shocks*¹³ a tales cambios; sin embargo, debe quedar claro que éstos nada tienen que ver con shocks estocásticos o aleatorios; simplemente se refieren a cambios, anticipados o no por los agentes, en algún parámetro.

Como ejemplo consideremos que la función de producción es de la forma $zf(k)$, con $z \in [1, \infty)$. Un cambio en el parámetro z se denomina usualmente como un shock tecnológico a la producción. Supongamos que originalmente el sistema se encuentra descansando en su estado estacionario (k_1, p_1) , para algún valor z_1 , e inesperadamente en $t = \tau$ se tiene que z_1 cambia a $z_2 > z_1$ (“shock positivo” a la producción); entonces el nuevo estado estacionario es (k_2, p_2) (estática comparativa). El sistema

¹³Así se hace en textos de uso común en economía. Véase por ejemplo [BF89].

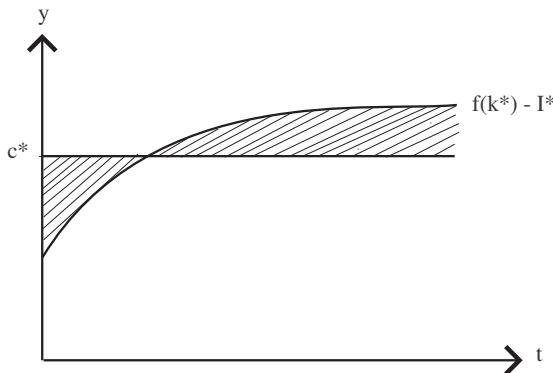


Figura 12.4: Restricción presupuestal intertemporal

(12.65) se transforma como se muestra en la figura 12.5, en donde puede verse el nuevo sistema (el lugar geométrico $\dot{p} = 0$ se recorre hacia arriba) junto con el estado estacionario anterior. En $t = \tau$, la condición inicial del nuevo sistema es $k(\tau) = k_1$ y por lo tanto $p(\tau) = p_\tau$ debe ser tal que el sistema esté sobre la nueva variedad estable.

Aquí se tiene también el fenómeno de sobreajuste en p , igual que en el modelo del tipo de cambio visto en la sección 5.7. Debe quedar claro que todo el análisis que se hace en estos casos es de tipo local ya que se está aproximando el modelo por uno lineal. Esto es importante ya que el cambio en los parámetros o shock debe ser pequeño con respecto al valor de las variables en el estado estacionario. Es un error común perder de vista este punto y encontrar, en ocasiones, análisis de shocks que son más bien terremotos, ¡como la destrucción de la mitad del acervo de capital!

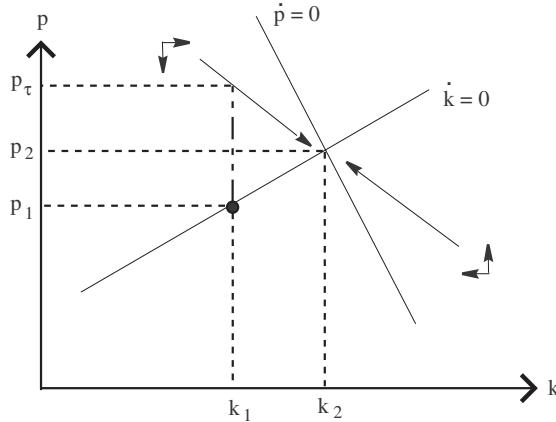


Figura 12.5: Dinámica para un shock positivo a la producción.

Ejercicios

▷ **12.1** Encontrar los extremos de las siguientes funcionales y determinar si son máximos o mínimos:

a) $J[x, u] = \int_0^{40} -\frac{u^2}{2} dt$ sujeto a $\dot{x} = u$,
 $x(0) = 20$ y $x(40) = 0$.

b) $J[x, u] = \int_0^1 u^2 dt$ sujeto a $\dot{x} = x + u$,
 $x(0) = 1$ y $x(1) = 0$.

c) $J[x, u] = \int_0^{10} (2xu + u^2) dt$ sujeto a $\dot{x} = u$,
 $x(0) = 10$ y $x(10)$ libre.

d) $J[x, u] = \int_0^2 (x - u^2) dt$ sujeto a $\dot{x} = u$,
 $x(0) = 0$ y $x(2)$ libre.

e) $J[x, u] = \int_0^{40} -\frac{1}{2}(x^2 + u^2) dt$ sujeto a $\dot{x} = u - x$,
 $x(0) = 1$ y $x(40)$ libre.

f) $J[x, u] = \int_0^2 (2x - 3u - \frac{u^2}{2}) dt$ sujeto a
 $\dot{x} = x + u$, $x(0) = 5$ y $x(2)$ libre.

g) $J[x, u] = \int_0^1 \ln u dt$ sujeto a $\dot{x} = -3u$,
 $x(0) = 0$ y $x(1) = 30$.

▷ **12.2** Encontrar el máximo de las siguientes funcionales:

a) $J[x, u] = \int_0^4 (x - u) dt$ sujeto a $\dot{x} = x + u$,
 $x(0) = 5$, $x(4)$ libre y $u(t) \in [0, 2]$.

b) $J[x, u] = \int_0^1 (x - 2u) dt$ sujeto a $\dot{x} = x + u$,
 $x(0) = 4$, $x(1)$ libre y $u(t) \in [0, 1]$.

c) $J[x, u] = \int_0^4 3x dt$ sujeto a $\dot{x} = x + u$,
 $x(0) = 5$, $x(4)$ libre y $u(t) \in [0, 2]$.

▷ **12.3** Una firma de microchips desea maximizar la siguiente función de ganancias

$$\int_0^5 \left(K - K^2 - \frac{I^2}{2} \right) dt,$$

en donde K es el capital e I la inversión bruta. Dada la naturaleza de la firma, el capital tiene una alta tasa de depreciación igual a 0.5 y su ecuación de evolución está dada por

$$\dot{K} = I - 0.5K.$$

Se tiene que inicialmente $K(0) = 0$ y se desea que $K(5) = 10$. Encontrar las trayectorias óptimas para el capital y para la inversión. Posteriormente, resolver el problema si $K(5)$ está libre. Probar que el punto de equilibrio es un punto silla.

▷ **12.4** Un monopolio tiene una función de costos dada por

$$C(x) = \alpha x^2 + \beta,$$

en donde x es el nivel de producción. La función de demanda p es dinámica¹⁴ y evoluciona de acuerdo con la ecuación

$$\dot{p} = Ap + x - B.$$

Todos los parámetros son positivos. Encontrar el máximo de

$$\int_0^T (px - C(x)) dt$$

si $p(0) = p_0$ y $p(T) = p_T$.

¹⁴El primer ejemplo de una función de demanda dinámica fue proporcionado por Evans [Eva24].

▷ **12.5** Resolver el siguiente problema:

$$\max \int_0^4 3x dt$$

sujeto a

$$\dot{x} = x + u, \quad x(0) = 5, \quad x(4) \geq 300 \text{ y } u(t) \in [0, 2].$$

▷ **12.10** La población de salmón en determinada región se denota por $x(t)$ y su función de crecimiento es $f(x)$. La pesca del salmón está dada por $h(t)$, en donde $h(t) \in [0, p_{\max}]$. Así, la población de salmón evoluciona de acuerdo con

$$\dot{x} = f(x) - h.$$

▷ **12.6** Resolver el siguiente problema:

$$\min \int_0^1 (x^2 + u^2) dt$$

sujeto a

$$\dot{x} = u, \quad x(0) = 0 \text{ y } x(1) \geq 1.$$

▷ **12.7** Encontrar el tiempo mínimo para pasar de $x(0) = 8$ a $x(T) = 0$ suponiendo que el estado x evoluciona de acuerdo con la ecuación $\dot{x} = 2u$ y que $u(t) \in [-1, 1]$.

▷ **12.8** Resolver el ejemplo 12.6.1 con las siguientes condiciones iniciales y finales: $x_1(0) = x_2(0) = 0, x_1(1)$ y $x_2(1)$ libres.

▷ **12.9** El acervo de plata de una mina en el tiempo t está dado por $A(t)$ y la tasa de extracción es $\dot{A} = -x$ con $x \in [0, x_{\max}]$. El precio de la plata está dado de forma exógena y sigue una trayectoria $p(t)$. No hay ningún costo de extracción. El acervo inicial está dado por $A(0) = A_0 > x_{\max}$ y el acervo final es $A(T) = 0$, el tiempo final T está libre y debe ser determinado. Si el tiempo se descuenta a una tasa r , resolver el problema de maximizar el valor presente de las ganancias, es decir,

$$\max \int_0^T p x e^{-rt} dt$$

sujeto a

$$\dot{A} = -x,$$

$$A(0) = A_0, \quad A(T) = 0, \quad T \text{ libre y } x \in [0, x_{\max}].$$

La función de crecimiento satisface las siguientes propiedades: existe un nivel de saturación K , de tal forma que $f(0) = 0 = f(K)$; $f(x) > 0$ si $x \in (0, K)$; $f(x) < 0$ si $x > K$, f es doblemente diferenciable y $f'' < 0$. Resolver el siguiente problema de optimización:

$$\max \int_0^\infty (p - c(x)) h(t) e^{-\rho t} dt$$

sujeto a $\dot{x} = f(x) - h$ y con x_0 dado. Se supone que $c(x)$ es el costo unitario de la pesca, que es diferenciable con $c' < 0$; p es el precio de mercado del salmón y ρ la tasa de descuento temporal.

▷ **12.11** Después de años de noviazgo, Miguel y Rosita se unen en feliz matrimonio. Deciden llevar una vida poco ortodoxa en la cual planean nunca tener hijos y sólo vivir hasta cumplir 50 años de matrimonio. La función de utilidad de este matrimonio está dada por $u(c) = 2\sqrt{c}$, donde c representa, como es usual, el consumo. Asimismo, poseen una tasa de descuento temporal dada por ρ . Sus ingresos reales provienen de dos fuentes: w el ingreso laboral y ar que corresponde a los rendimientos reales de una cantidad a de activos y una tasa de interés r . Sus egresos reales son el consumo c y la acumulación de activos \dot{a} . Las condiciones iniciales y finales son $a(0) = 0$ y $a(50) = 0$.

- a)** Resolver el problema de maximización de Miguel y Rosita al inicio de su vida juntos. Ilustrar con un diagrama de fase.

- b)** Integrando la restricción presupuestal con la condición de *no juego de Ponzi*, probar que el valor presente de su ingreso laboral a lo largo de toda su vida es igual al valor presente del consumo durante el mismo periodo.
- c)** Resolver el problema de maximización suponiendo que Miguel y Rosita son una familia común que decide tener hijos y preocuparse por el bienestar de todos sus descendientes. En otras palabras, suponer que ahora el horizonte de tiempo es infinito. Representar la solución en un diagrama de fase.

▷ **12.12** Plantear las condiciones de primer orden y resolver el siguiente problema. La autoridad monetaria controla la cantidad nominal de dinero mediante un variable que denotamos por $u(t)$. Por razones institucionales, esta variable sólo puede tomar valores en el intervalo $[-0.1, 0.1]$. La tasa de crecimiento del ingreso nominal se denota por $x(t)$ y satisface la siguiente ecuación diferencial:

$$\frac{a^2}{2} \ddot{x} + a\dot{x} + x = u.$$

El objetivo de la autoridad monetaria es llegar a una tasa constante de crecimiento del ingreso nominal en un mínimo de tiempo. El problema se puede plantear como,

$$\min \int_0^T dt$$

sujeto a

$$\frac{a^2}{2} \ddot{x} + a\dot{x} + x = u$$

y a las siguientes condiciones de frontera: $x(0) = x_0$, $\dot{x}(0) = \dot{x}_0$, $x(T) = x_T$, $\dot{x}(T) = 0$ y $u \in [-0.1, 0.1]$. (Sugerencia: convertir el problema en uno con dos variables de estado definiendo $y(t) = \dot{x}(t)$.)

▷ **12.13** En 5.7.2 se mostró un ejemplo de los efectos de una devaluación y se utilizó una ecuación de demanda de dinero. En este ejercicio se

obtiene dicha ecuación a partir de un problema de optimización.

Considerar el siguiente problema:

$$\max \int_0^\infty u(c, m) e^{-\rho t} dt$$

sujeto a $y + T = c + \dot{m} + m\pi$ y con m_0 dado. La notación es la misma que en la sección 12.8.1 salvo por y , que representa un ingreso exógeno ya que en este modelo no hay producción. La transferencia T que hace el Banco Central al agente representativo (“la sociedad”) es financiada mediante la inflación y está dada por

$$T = m\pi.$$

- a)** Resolver el problema de optimización y obtener la función de demanda de dinero.
- b)** Obtener un sistema de ecuaciones diferenciales en m y c . Éste debe ser el mismo sistema que se utilizó en la sección 5.7.2.
- c)** Resolver el problema si ahora el Banco Central, en lugar de transferir al agente los ingresos que obtiene por la inflación, los desperdicia en actividades inútiles (por ejemplo, los tira al golfo de México).

▷ **12.14** Para el modelo de una economía abierta pequeña realizar los siguientes ejercicios:

- a)** Dibujar las trayectorias para k , p , I y c en el caso del “shock positivo” en la producción.
- b)** Analizar el caso en que, en $t = \tau$, inesperadamente la tasa de interés extranjera r cambia a $\bar{r} > r$. Representar el diagrama de fase y dibujar las trayectorias de todas las variables como en el inciso anterior.

- ▷ **12.15** Considerar la siguiente variante del modelo de Ramsey en donde el hogar-empresa representativo resuelve el problema de maximización

$$\max \int_0^\infty u(c)e^{-\rho t} dt$$

sujeto a

$$f(k) + T = c(1 + \tau) + \dot{k},$$

en donde u y f son funciones de utilidad y producción, respectivamente, con las propiedades usuales. Adicionalmente, ρ es la tasa subjetiva de descuento, c el consumo, w el salario, k el capital, τ un impuesto al consumo y T una transferencia del gobierno a los hogares (igual para todos), financiada por medio del impuesto. Como siempre, c es la variable de control y k la variable de estado; todas las demás variables son exógenas para el optimizador.

- a) Resolver el problema de optimización, incorporar la restricción presupuestal del gobierno dada por $T = \tau c$ y obtener un sistema de ecuaciones diferenciales en c y k . Realizar un diagrama de fase para el sistema de ecuaciones resultante. Comprobar que el resultado es idéntico al modelo de Ramsey sin depreciación del capital ($\delta = 0$). La conclusión es que este tipo de impuesto al consumo no afecta el estado estacionario ni la trayectoria hacia éste.
- b) Ahora cambiemos la restricción presupuestal como sigue.

$$(1 + \gamma)f(k) = c + \dot{k} - \chi,$$

en donde γ es un subsidio a la producción, financiado por un impuesto χ (igual para todos). Resolver el problema de optimización y probar que la tasa real que observan los hogares está dada por $(1 + \gamma)f'(k)$, que es mayor que $f'(k)$ en el inciso anterior. Incorporar

la restricción presupuestal del gobierno dada por $\gamma f(k) = \chi$ y obtener un sistema de ecuaciones en c y k . Realizar un diagrama de fase. ¿Cómo se compara el estado estacionario con aquél en el cual no hay ningún impuesto (modelo usual de Ramsey)?

- c) Resolver el problema si la función de utilidad es lineal, dada por $u(c) = c$ y la restricción presupuestal está dada por

$$f(k) = c + \dot{k}.$$

Ilustrar con un diagrama de fase.

- ▷ **12.16 (Crecimiento endógeno)** Considerar el mismo modelo del ejercicio anterior, pero con $u(c) = \ln c$ y $f(k) = Ae^{\mu t}k^\alpha$, con $\mu > 0$ un coeficiente de desarrollo tecnológico (exógeno) y $\alpha \in (0, 1)$. No hay ningún tipo de impuesto, así que la restricción presupuestal queda dada por

$$Ae^{\mu t}k^\alpha = c + \dot{k}.$$

- a) Resolver el problema de optimización.
- b) Probar que $\frac{\dot{k}}{k} = \frac{\dot{c}}{c} = \frac{\dot{y}}{y} = \frac{\mu}{1-\alpha}$. A esto se le llama crecimiento balanceado (las variables crecen a la misma tasa).

- ▷ **12.17** En el artículo original de Frank Ramsey [Ram28], los individuos no descuentan el futuro ya que Ramsey considera que es éticamente insostenible el “descontar” a las generaciones futuras. Para poder acotar la integral impropia considera un nivel de utilidad máximo, que llamamos la *dicha absoluta* y denotamos por D . El problema de optimización a resolver es, entonces,

$$\min \int_0^\infty (D - \ln c) dt$$

sujeto a $k^\alpha = c + \dot{k}$ y con k_0 dado. Se supone, además, que $u(c) = \ln c$ y $f(k) = k^\alpha$, $0 < \alpha < 1$, son las funciones de utilidad y producción, respectivamente.

- a) Obtener las condiciones de primer orden para el problema de optimización ¿Cómo evoluciona el consumo?
- b) Comparar los resultados del inciso anterior con el modelo usual, en donde hay una tasa subjetiva de descuento.

▷ **12.18 (Gasto gubernamental)**¹⁵ Supongamos que el gobierno provee servicios G que son parte de la función de producción (pueden pensarse como infraestructura básica necesaria para la producción). Estos servicios se pagan por medio de un impuesto τ a la producción. Si c y k denotan las trayectorias de consumo y capital, ρ es la tasa subjetiva de descuento temporal, la función de producción es $f(k, G) = G^\alpha k^{1-\alpha}$, $0 < \alpha < 1$ y la función de utilidad es $u(c) = \frac{c^{1-\sigma}}{1-\sigma}$ con $\sigma \neq 1$ el coeficiente positivo de aversión relativa al riesgo¹⁶, el problema del hogar representativo es,

$$\max \int_0^{\infty} \frac{c^{1-\sigma}}{1-\sigma} e^{-\rho t} dt,$$

sujeto a $(1 - \tau)G^\alpha k^{1-\alpha} = c + \dot{k}$, con k_0 y G dados. Nótese que la trayectoria de G es paramétrica (exógena) para el optimizador.

- a) Encontrar las condiciones de primer orden y obtener una ecuación que describa la evolución del consumo.
- b) Incorporar la restricción presupuestal del gobierno $G = \tau f(k, G)$. Obtener G como función de τ y k , y sustituir en la ecuación de evolución del consumo. ¿Cuál es la tasa de crecimiento $\gamma = \frac{\dot{c}}{c}$?
- c) Encontrar τ que maximiza γ .
- d) El planeador central toma la economía en sus manos y resuelve el problema de optimización del hogar representativo. La diferencia es que G no es paramétrico para el planeador y por lo tanto incorpora la restricción presupuestal del gobierno en el problema de optimización. La restricción presupuestal resultante es

$$G^\alpha k^{1-\alpha} = c + G + \dot{k}.$$

Resolver el problema de optimización en este caso.

- e) Obtener, como antes, $\gamma^p = \frac{\dot{c}}{c}$. ¿Es ésta menor o mayor que la tasa γ obtenida anteriormente? Interpretar.

¹⁵Basado en [Bar90].

¹⁶Véase la sección 3.1.1.

CAPÍTULO 13

Problemas con restricciones

§13.1 Restricciones de igualdad sobre las variables de control

El principio del máximo facilita la introducción de cierto tipo de restricciones. Las más sencillas son aquellas que involucran a las variables de control ya que el problema es equivalente a optimizar el hamiltoniano sujeto a las restricciones dadas. En este caso, simplemente utilizamos las técnicas conocidas de lagrangianos y las condiciones de Kuhn-Tucker expuestas en el capítulo 10.

Consideremos el caso más simple de restricciones de igualdad. Para facilitar la exposición sólo tomamos una restricción, pero el caso general, con un número de restricciones de igualdad menor que el número de variables de control, es análogo. El problema que se quiere resolver es el siguiente:

$$\max \int_0^T f(x, u_1, u_2, t) dt$$

sujeto a

$$\begin{aligned}\dot{x} &= g(x, u_1, u_2, t), \\ h(x, u_1, u_2, t) &= c,\end{aligned}$$

donde x es la variable de estado, u_1 y u_2 las de control, c es una constante y se tienen ciertas condiciones de frontera.

El principio del máximo nos dice que para todo tiempo t se debe maximizar el hamiltoniano,

$$H = f(x, u_1, u_2, t) + \lambda(t)g(x, u_1, u_2, t)$$

con respecto al control u , sujeto a la restricción

$$h(x, u_1, u_2, t) = c.$$

Esto se logra considerando el lagrangiano

$$\mathcal{L} = f(x, u_1, u_2, t) + \lambda(t)g(x, u_1, u_2, t) + \mu(t)(c - h(x, u_1, u_2, t)),$$

Nótese que el multiplicador $\mu(t)$ también es función del tiempo.

Aplicamos el teorema 12.2.3 al lagrangiano para obtener las condiciones de primer orden:

$$\begin{aligned}\mathcal{L}_{u_i} &= f_{u_i} + \lambda g_{u_i} - \mu h_{u_i} = 0, \quad i = 1, 2, \\ \dot{\lambda} &= -\mathcal{L}_x = -f_x - \lambda g_x + \mu h_x, \\ \dot{x} &= g(x, u_1, u_2, t), \\ h(x, u_1, u_2, t) &= c\end{aligned}$$

y las condiciones de transversalidad apropiadas.

Ejemplos

Ej 13.1.1 Se desea resolver el siguiente problema:

$$\max \int_0^1 (2x^2 - 2u^2 - 2v^2) dt$$

sujeto a

$$\dot{x} = -x - 4u - 4v,$$

$$1 - x + u + v = 0,$$

$$x(0) = 1.$$

El lagrangiano asociado es

$$\mathcal{L} = 2x^2 - 2u^2 - 2v^2 + \lambda(-x - 4u - 4v) + \mu(1 - x + u + v).$$

Puede verse que ésta es una función cóncava en u y v y por lo tanto el extremo será un máximo. Las condiciones de primer orden están dadas por

$$\begin{aligned}\mathcal{L}_u &= -4u - 4\lambda + \mu = 0, \\ \mathcal{L}_v &= -4v - 4\lambda + \mu = 0, \\ \dot{\lambda} &= -\mathcal{L}_x = -4x + \lambda + \mu, \\ \dot{x} &= -x - 4u - 4v, \\ 1 - x + u + v &= 0.\end{aligned}$$

Dado que $u + v = x - 1$, se tiene que $\dot{x} = -x - 4(x - 1) = -5x + 4$ y por lo tanto $x^*(t) = Ae^{-5t} + \frac{4}{5}$. La condición inicial $x(0) = 1$ determina la constante $A = \frac{1}{5}$. Asimismo, las primeras dos ecuaciones implican $u = v$, con lo cual $u^*(t) = v^*(t) = \frac{x^*(t)-1}{2} = \frac{1}{10}e^{-5t} - \frac{1}{10}$.

Ej 13.1.2 Encontrar la curva de longitud mínima en \mathbb{R}^3 que va de $(0, 1, \sqrt{2})$ a $(1, 1, 1)$ y que está sobre la esfera $t^2 + x^2 + y^2 = 3$. Cualquier curva de longitud mínima entre dos puntos de una superficie se denomina **curva geodésica**. En el ejemplo 11.2.3 demostramos que las curvas geodésicas sobre un plano son líneas rectas. Si $(t, x(t), y(t))$ es un punto sobre la curva, el vector tangente está dado por $(1, \dot{x}, \dot{y})$, cuya longitud es $\sqrt{1 + \dot{x}^2 + \dot{y}^2}$. La longitud de la curva, al igual que en el caso bidimensional, es la integral de las longitudes de las tangentes en el intervalo dado, es decir,

$$\text{longitud} = \int_0^1 \sqrt{1 + \dot{x}^2 + \dot{y}^2} dt.$$

El problema se puede plantear como

$$\min \int_0^1 \sqrt{1 + u^2 + v^2} dt$$

sujeto a

$$\begin{aligned} \dot{x} &= u, \\ \dot{y} &= v, \\ t^2 + x^2 + y^2 &= 3, \\ x(0) &= 1, \quad x(1) = 1, \\ y(0) &= \sqrt{2}, \quad y(1) = 1. \end{aligned}$$

El lagrangiano asociado es

$$\mathcal{L} = \sqrt{1 + u^2 + v^2} + \lambda_1 u + \lambda_2 v + \mu(3 - t^2 - x^2 - y^2),$$

de forma que las condiciones de primer orden son

$$\begin{aligned} \mathcal{L}_u &= \frac{u}{\sqrt{1 + u^2 + v^2}} + \lambda_1 = 0, \\ \mathcal{L}_v &= \frac{v}{\sqrt{1 + u^2 + v^2}} + \lambda_2 = 0, \\ \dot{\lambda}_1 &= -\mathcal{L}_x = 2x\mu, \\ \dot{\lambda}_2 &= -\mathcal{L}_y = 2y\mu, \\ \dot{x} &= u, \\ \dot{y} &= v, \\ t^2 + x^2 + y^2 &= 3. \end{aligned}$$



§13.2 Restricciones integrales

Este tipo de restricción se conoce también como restricción isoperimétrica ya que es la que aparece cuando se busca la figura geométrica con un perímetro especificado y área máxima.¹ El problema a resolver es el siguiente:

$$\max \int_0^T f(x, u, t) dt$$

sujeto a

$$\begin{aligned} \dot{x} &= g(x, u, t), \\ \int_0^T G(x, u, t) dt &= K, \end{aligned}$$

donde K es una constante, y dadas ciertas condiciones de frontera.

Este problema se puede reformular como un problema sin restricciones y con una variable de estado adicional como sigue.

Definimos

$$y(t) = \int_0^t G(x, u, \tau) d\tau$$

y por el teorema fundamental del cálculo

$$\dot{y} = G(x, u, t).$$

De esta forma, el problema se puede reescribir como

$$\max \int_0^T f(x, u, t) dt$$

sujeto a

$$\begin{aligned} \dot{x} &= g(x, u, t), \\ \dot{y} &= G(x, u, t), \\ y(0) &= 0, \quad y(T) = k \end{aligned}$$

y condiciones de frontera para x .

El problema se puede resolver ahora como un problema normal, sin restricciones. Una observación pertinente es que la trayectoria de la variable de estado auxiliar y no es de interés para el problema, por lo que puede omitirse su ecuación de evolución.

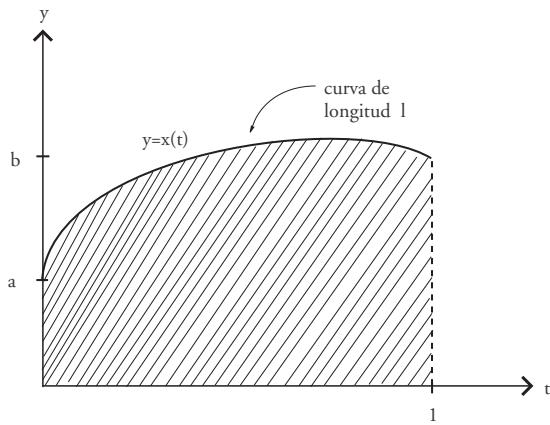


Figura 13.1: Restricción integral para el ejemplo 13.2.1.

Ejemplo

Ej 13.2.1 Encuentra la curva que va de $(0, a)$ a $(1, b)$ con longitud l dada, de manera que el área entre la curva y el eje t sea máxima. La figura 13.1 muestra el área a maximizar. El problema se puede plantear como

$$\max \int_0^1 x dt$$

sujeto a

$$\begin{aligned} \dot{x} &= u, \\ \int_0^1 \sqrt{1+u^2} dt &= l, \\ x(0) &= a, x(1) = b. \end{aligned}$$

Definimos

$$y(t) = \int_0^t \sqrt{1+u^2} d\tau$$

y el problema es ahora

$$\max \int_0^1 x dt$$

¹Es el problema mencionado en la sección 11.1 propuesto por Dido de Cartago.

sujeto a

$$\begin{aligned}\dot{x} &= u, \\ \dot{y} &= \sqrt{1+u^2}, \\ x(0) &= a, \quad x(1) = b; \\ y(0) &= 0, \quad y(1) = l.\end{aligned}$$

El hamiltoniano asociado es

$$H = x + \lambda_1 u + \lambda_2 \sqrt{1+u^2}$$

y por lo tanto las condiciones relevantes de primer orden son

$$\begin{aligned}H_u &= \lambda_1 + \frac{\lambda_2 u}{\sqrt{1+u^2}} = 0, \\ \dot{\lambda}_1 &= -H_x = -1, \\ \dot{\lambda}_2 &= -H_y = 0, \\ \dot{x} &= u.\end{aligned}$$

Resolviendo se tiene que

$$\begin{aligned}\lambda_1 &= A - t, \\ \lambda_2 &= B, \\ A - t + \frac{Bu}{\sqrt{1+u^2}} &= 0,\end{aligned}$$

con A y B constantes indeterminadas. Se resuelve para u y obtenemos

$$u^*(t) = \frac{(t-A)}{\sqrt{B^2 - (t-A)^2}}.$$

Asimismo,

$$\dot{x}^*(t) = u^*(t) = \frac{(t-A)}{\sqrt{B^2 - (t-A)^2}}$$

por lo que integrando se tiene

$$x^*(t) = \int \frac{(t-A)}{\sqrt{B^2 - (t-A)^2}} dt = -\sqrt{B^2 - (t-A)^2} + C,$$

donde C es una constante indeterminada. Ésta es la ecuación de un semicírculo. Las constantes se determinan con las condiciones iniciales y finales junto con la restricción integral.²



²El sistema de ecuaciones resultante es bastante tedioso de resolver ya que la restricción integral se convierte en

$$\int_0^1 \frac{B}{\sqrt{B^2 - (t-A)^2}} dt = l = B \left(\sin^{-1}\left(\frac{1-A}{B}\right) + \sin^{-1}\left(\frac{A}{B}\right) \right).$$

§13.3 Restricciones de desigualdad sobre las variables de control

El problema de la sección 13.1 se transforma ahora en el siguiente:

$$\max \int_0^T f(x, u, t) dt$$

sujeto a

$$\begin{aligned}\dot{x} &= g(x, u, t), \\ h(x, u, t) &\leq c,\end{aligned}$$

donde c es una constante, dadas ciertas condiciones de frontera.

El principio del máximo, una vez más, nos da una forma de encontrar la solución. La diferencia es que ahora debemos aplicar las condiciones de Kuhn-Tucker³ al problema de maximización del hamiltoniano.

Se debe maximizar el hamiltoniano

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

con respecto al control u y sujeto a la restricción,

$$h(x, u, t) \leq c.$$

Esto se logra considerando el lagrangiano,

$$\mathcal{L} = f(x, u, t) + \lambda(t)g(x, u, t) + \mu(t)(c - h(x, u, t)),$$

con las siguientes condiciones de primer orden:

$$\begin{aligned}\mathcal{L}_u &= f_u + \lambda g_u - \mu h_u = 0, \\ \dot{\lambda} &= -\mathcal{L}_x = -f_x - \lambda g_x + \mu h_x, \\ \dot{x} &= g(x, u, t), \\ h(x, u, t) - c &\leq 0, \quad \mu \geq 0, \quad \mu[h(x, u, t) - c] = 0,\end{aligned}$$

y las condiciones de transversalidad pertinentes. Hacemos las siguientes observaciones importantes:

- Si adicionalmente se tiene la restricción de no-negatividad $u \geq 0$ para la variable de control, entonces la condición $\mathcal{L}_u = 0$ debe ser reemplazada por

$$\mathcal{L}_u = f_u + \lambda g_u - \mu h_u \leq 0, \quad u \geq 0, \quad u[f_u + \lambda g_u - \mu h_u] = 0.$$

³En el capítulo 10 se da un breve resumen de estas condiciones.

- En el problema general se tiene un vector de estado $\mathbf{x} = (x_1, \dots, x_n)$ y un vector de control $\mathbf{u} = (u_1, \dots, u_m)$. Recordemos que puede haber cualquier número de restricciones de desigualdad, $g_i(\mathbf{x}, \mathbf{u}, t) \leq c_i, i = 1, \dots, k$, aun si es mayor que el número de controles, siempre y cuando el número de restricciones activas no sea mayor que el número de controles. Adicionalmente las restricciones deben satisfacer la condición de cualificación o alguno de sus equivalentes. En el capítulo 10 se aclara este punto.

NOTA. Es sumamente importante que cualquier problema de optimización con restricciones de desigualdad sea planteado siempre de la misma forma. Es decir, se debe expresar como un problema de maximización con restricciones de la forma $h(x, u, t) \leq c$; claramente, un problema de minimización de la funcional F puede convertirse en un problema de maximización de $-F$. Esta observación es importante ya que las condiciones de primer orden contienen desigualdades que dependen del planteamiento adecuado del problema.

§13.3.1 Otro modelo de inversión

Supongamos que los ingresos de una empresa están dados por la función

$$\Pi(K) = \frac{1}{\alpha} K^\alpha, 0 < \alpha < 1,$$

en donde K es capital. Podemos pensar que la función de producción es $\frac{1}{\alpha} K^\alpha$ y el precio del producto es igual a 1. Si I es la inversión bruta, denotemos por $C(I) = qI$ donde $q > 0$ representa los costos asociados a ésta. El capital se deprecia a una tasa δ . La inversión bruta es no negativa y debe estar financiada por los ingresos de la empresa. El problema de optimización, si se descuenta a una tasa r , queda dado por

$$\max \int_0^{\infty} \left(\frac{1}{\alpha} K^\alpha - qI \right) e^{-rt} dt$$

sujeto a

$$\dot{K} = I - \delta K,$$

$$0 \leq I \leq \frac{1}{\alpha} K^\alpha,$$

$$K(0) = K_0 \text{ dado.}$$

El problema se puede ver como un problema con una restricción, $I \leq \frac{1}{\alpha} K^\alpha$ más la condición de no negatividad sobre I . El lagrangiano asociado (en tiempo corriente) es

$$\mathcal{L} = \frac{1}{\alpha} K^\alpha - qI + \lambda(I - \delta K) + \mu(\frac{1}{\alpha} K^\alpha - I)$$

y las condiciones de primer orden son

$$\mathcal{L}_I = -q + \lambda - \mu \leq 0, \quad I \geq 0, \quad I[-q + \lambda - \mu] = 0,$$

$$\dot{\lambda} = -\mathcal{L}_K + \lambda r = -K^{\alpha-1} + \lambda\delta - \mu K^{\alpha-1} + \lambda r,$$

$$\mu \geq 0, \quad I \leq \frac{1}{\alpha} K^\alpha, \quad \mu[\frac{1}{\alpha} K^\alpha - I] = 0,$$

más la ecuación de evolución. Tenemos entonces que, analizando la relación entre λ y q ,

$$\begin{aligned} \lambda < q \Rightarrow \\ I = 0, \mu = 0, \dot{K} = -\delta K, \dot{\lambda} = \lambda(r + \delta) - K^{\alpha-1}, \end{aligned} \quad (13.1)$$

$$\begin{aligned} \lambda = q \text{ y } 0 < I < \frac{1}{\alpha} K^\alpha \Rightarrow \\ \mu = 0, \dot{K} = I - \delta K, \dot{\lambda} = \lambda(r + \delta) - K^{\alpha-1}, \end{aligned} \quad (13.2)$$

$$\begin{aligned} \mu \neq 0 \Rightarrow \\ \lambda = q + \mu > q, I = \frac{1}{\alpha} K^\alpha, \dot{K} = \frac{1}{\alpha} K^\alpha - \delta K, \\ \dot{\lambda} = \lambda(r + \delta) - K^{\alpha-1}(1 + \mu). \end{aligned} \quad (13.3)$$

Podemos realizar el diagrama de fase si tomamos como variables a K y a λ , ya que las ecuaciones (13.1), (13.2) y (13.3) definen un sistema en estas variables. Graficamos las isoclinas $\dot{\lambda} = 0 = \dot{K}$ considerando los tres casos: $\lambda < q$, $\lambda = q$ y $\lambda > q$. El equilibrio del sistema se encuentra como la intersección de ambas isoclinas y está dado por $\lambda^* = q$ y $K^* = [\lambda^*(r + \delta)]^{\frac{1}{\alpha-1}}$. Analizando la dinámica del flujo notamos la existencia de dos trayectorias: una estable, que converge a (K^*, λ^*) , y otra inestable que se aleja del equilibrio, de manera que éste se comporta como un punto silla. Esto se muestra en la figura 13.2.

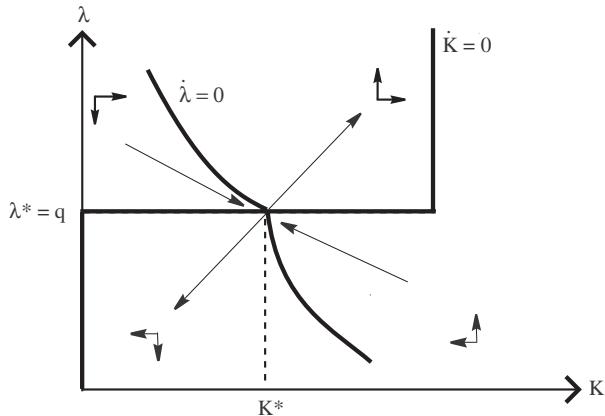


Figura 13.2: Diagrama de fase cuando la inversión no puede ser negativa.

Se pueden calcular las trayectorias de las variables a lo largo de la variedad estable como sigue. Si $K_0 < K^*$, entonces $\lambda = q + \mu$, $\mu > 0$, y por lo tanto se cumple (13.3). La ecuación $\dot{K} = \frac{1}{\alpha} K^\alpha - \delta K$ es una ecuación del tipo Bernoulli y se resuelve realizando el cambio de variable $w = K^{1-\alpha}$. La solución queda como

$$K(t) = \left(A e^{-(1-\alpha)\delta t} + \frac{1}{\alpha \delta} \right)^{\frac{1}{1-\alpha}}. \quad (13.4)$$

Notemos que, dados $K(0) = K_0$ y $K(T^*) = K^*$, podemos determinar la constante A y el tiempo terminal T^* . Observemos que este sistema difiere de los analizados anteriormente en que llega a su estado

estacionario en un tiempo finito. La inversión durante el ajuste es siempre igual al máximo posible: $I = \frac{1}{\alpha} K^\alpha$. En este caso decimos que el ajuste se da por medio de la **trayectoria de aproximación más rápida (most rapid approach path)**. Veamos ahora qué pasa si $K_0 > K^*$. En esta situación se aplica (13.1) y tenemos que $\dot{K} = -\delta K$, de manera que

$$K = Be^{-\delta t}; \quad (13.5)$$

al igual que antes pueden determinarse $B = K_0$ y el tiempo terminal T^* , lo que en este caso es sumamente sencillo ya que $K^* = K_0 e^{-\delta T^*}$ implica que $T^* = \frac{1}{\delta} \ln \frac{K_0}{K^*}$. La inversión en este caso es la mínima posible, $I = 0$, y el ajuste se da, una vez más, de la manera más rápida posible.

Para finalizar, recordemos que en la sección 11.7.1 se estudió un modelo semejante pero con costos $c(I)$ no lineales. Con costos lineales no existe dinámica alguna y el ajuste hacia el estado estacionario es instantáneo. Aquí, la dinámica se logra acotando la inversión, a pesar de que los costos $c(I) = qI$ son lineales. La idea es que, para lograr el ajuste instantáneo, la inversión estaría indeterminada (sería $\pm\infty$) por un instante; al acotarla, evitamos que esto suceda.

§13.4 Restricción sobre el tiempo terminal

Si el tiempo terminal T del problema no está especificado pero se requiere que esté en algún intervalo específico, el problema se plantea entonces como

$$\max \int_0^T f(x, u, t) dt$$

sujeto a

$$\begin{aligned} \dot{x} &= g(x, u, t), \\ x(0) &= x_0 \text{ y } x(T) = x_T \text{ dados, } t_1 \leq T \leq t_2. \end{aligned}$$

Recordemos que la condición de transversalidad correspondiente a $x(T)$ dado y T libre está dada por $H(T) = 0$; la restricción $t_1 \leq T \leq t_2$ obliga a modificar esta condición para incorporar las condiciones de holgura. Tenemos entonces que la condición de transversalidad correspondiente es

$$\begin{aligned} H(T) &\geq 0, \quad T \leq t_2, \quad (T - t_2)H(T) = 0, \\ H(T) &\leq 0, \quad T \geq t_1, \quad (T - t_1)H(T) = 0. \end{aligned} \quad (13.6)$$

El hamiltoniano puede ser sustituido por un lagrangiano si el problema se combina con alguna otra restricción. En el ejemplo 13.5.1 se ilustra este tipo de restricción sobre el tiempo terminal junto con una restricción adicional sobre el estado.

§13.5 Restricciones en el espacio de estados

Cuando no aparece ninguna variable de control en las restricciones no podemos invocar el principio del máximo para obtener la solución. Lo que se está restringiendo ahora es el dominio de las funciones de

estado. Si pensamos al problema como uno de cálculo en variaciones, la restricción es sobre el dominio de la funcional. El problema que se quiere resolver es del tipo

$$\max \int_0^T f(x, u, t) dt$$

sujeto a

$$\begin{aligned} \dot{x} &= g(x, u, t), \\ h(x, t) &\leq c, \end{aligned}$$

más las condiciones de frontera.

Observamos que no tendría sentido pedir que la restricción estuviese activa en todo momento pues esto determinaría automáticamente al estado x . El primer método que seguimos se basa en el introducido por Hestenes [Hes66]. La idea es que mientras x sea tal que la restricción se cumple no hay ningún problema, pero en el momento en que la restricción se activa, es decir $h(x, t) = c$, entonces no debemos dejar que la función h continúe creciendo. Esto último se puede lograr si se impone la condición,

$$\dot{h} = \frac{dh}{dt} \leq 0 \text{ si } h(x, t) = c. \quad (13.7)$$

La derivada total $\frac{dh}{dt}$ involucra a \dot{x} y por lo tanto al control u a través de la ecuación de evolución del estado. De esta forma, la restricción se transforma en una que contiene a la variable de control y el problema se reescribe como

$$\max \int_0^T f(x, u, t) dt$$

sujeto a

$$\begin{aligned} \dot{x} &= g(x, u, t), \\ \dot{h} &= \frac{dh}{dt} \leq 0 \text{ si } h(x, t) = c \end{aligned}$$

y las condiciones de frontera. Ahora sí podemos invocar el principio del máximo para obtener la solución por medio del lagrangiano

$$\mathcal{L} = f(x, u, t) + \lambda g(x, u, t) - \mu \dot{h}.$$

Las condiciones de primer orden son

$$\begin{aligned} \mathcal{L}_u &= f_u + \lambda g_u - \mu \dot{h}_u = 0, \\ \dot{\lambda} &= -\mathcal{L}_x = -f_x - \lambda g_x + \mu \dot{h}_x, \\ \dot{h} &\leq 0, \quad \mu \geq 0, \quad \mu \dot{h} = 0, \\ h(x, t) - c &\leq 0, \quad \mu [h(x, t) - c] = 0, \\ \dot{\mu} &\leq 0, \quad \dot{\mu} [h(x, t) - c] = 0, \\ \dot{x} &= g(x, u, t). \end{aligned}$$

Afortunadamente, obtener la solución de esta forma es, en la mayor parte de los casos, equivalente a considerar simplemente el lagrangiano

$$\mathcal{L} = f(x, u, t) + \lambda g(x, u, t) + \mu(c - h),$$

es decir, de manera idéntica a cuando la restricción contiene variables de control. Las condiciones de primer orden son entonces

$$\begin{aligned}\mathcal{L}_u &= f_u + \lambda g_u = 0, \\ \dot{\lambda} &= -\mathcal{L}_x = -f_x - \lambda g_x + \mu h_x, \\ c - h(x, t) &\geq 0, \quad \mu \geq 0, \quad \mu[h(x, t) - c] = 0, \\ \dot{x} &= g(x, u, t).\end{aligned}$$

El problema técnico es que λ puede ser discontinua en los puntos donde la restricción pasa de inactiva a activa o viceversa. Si τ denota uno de estos puntos, existe una condición de brinco⁴ dada por

$$\begin{aligned}\lim_{t \rightarrow \tau^+} \lambda(t) &= \lim_{t \rightarrow \tau^-} \lambda(t) + b h_x(x(\tau), \tau), \\ b &\geq 0.\end{aligned}\tag{13.8}$$

Esta condición no especifica el valor de b , así que sólo da la dirección del brinco y no su magnitud, que puede ser cero si $b = 0$. Sugerimos simplemente estar alertas a esta posibilidad y utilizar este segundo método más simple. Intuitivamente, podemos suponer la continuidad de λ si en el punto donde la restricción se vuelve activa las pendientes de x y la restricción coinciden.

Ejemplo

Ej 13.5.1 Se quiere resolver el siguiente problema de optimización:

$$\max \int_0^T (-x^2 - u^2) dt$$

sujeto a

$$\begin{aligned}\dot{x} &= u, \\ -t + 2 &\leq x \leq -t + 3, \\ x(0) &= 2.5, \quad x(T) = 0 \text{ y } T \text{ libre.}\end{aligned}$$

En este caso, las restricciones sobre la variable de estado se pueden reescribir como

$$\begin{aligned}-x - t &\leq -2 \\ x + t &\leq 3.\end{aligned}$$

Esto se puede interpretar gráficamente como el que la trayectoria $y = x(t)$ debe estar siempre dentro de la región que se muestra en la figura 13.3. Así, el tiempo final no es del todo libre sino que debe cumplir $2 \leq T \leq 3$, de manera que hay también una restricción sobre el tiempo terminal.

⁴A esta condición se la llama, en ocasiones, de Erdmann-Weierstrass. Para más detalles véase el libro de Sydsaeter [SS87].

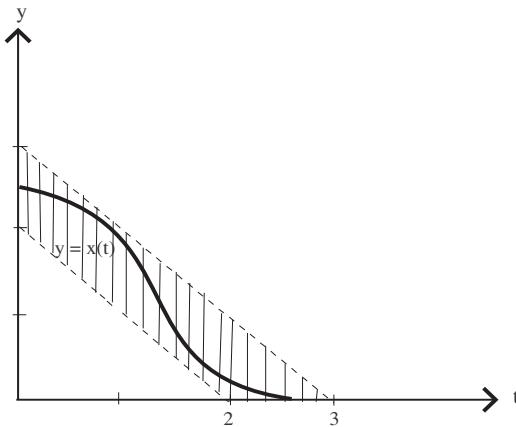


Figura 13.3: Región permitida del ejemplo 13.5.1.

El lagrangiano asociado al sistema es

$$\mathcal{L} = -x^2 - u^2 + \lambda u + \mu_1(x + t - 2) + \mu_2(-x - t + 3)$$

y las condiciones de primer orden están dadas por

$$\begin{aligned}\mathcal{L}_u &= -2u + \lambda = 0, \\ \dot{\lambda} &= -\mathcal{L}_x = 2x - \mu_1 + \mu_2, \\ x + t - 2 &\geq 0, \quad \mu_1 \geq 0, \quad \mu_1(x + t - 2) = 0, \\ -x - t + 3 &\geq 0, \quad \mu_2 \geq 0, \quad \mu_2(-x - t + 3) = 0,\end{aligned}$$

junto con la ecuación de evolución. En cualquier intervalo en donde las restricciones no están activas se tiene que $\mu_1 = \mu_2 = 0$ y por lo tanto $\lambda = 2u$ y $\dot{\lambda} = 2\dot{u} = 2x$. Dado que $\dot{x} = u$ se tiene $\ddot{x} - x = 0$, cuya solución es

$$x(t) = Ae^t + Be^{-t}$$

y esto implica

$$u(t) = Ae^t - Be^{-t}.$$

Observemos que si ninguna restricción está activa en T , entonces $2 < T < 3$ y la condición de transversalidad pertinente es $\mathcal{L}(T) = 0$, la cual, utilizando $x(T) = 0$ y $\lambda = 2u$, se transforma en

$$\begin{aligned}\mathcal{L}(T) &= -u^2(T) + \lambda(T)u(T) \\ &= -(Ae^T - Be^{-T})^2 + 2(Ae^T - Be^{-T})^2 \\ &= Ae^T - Be^{-T} \\ &= 0.\end{aligned}$$

De aquí que $u(T) = 0$ y junto con $x(T) = 0$ se tiene $A = B = 0$. Esto es absurdo puesto que $x(0) = 2.5 \neq 0$. Se debe tener entonces, que en T alguna restricción está activa y por lo tanto $T = 2$ o bien $T = 3$. El lagrangiano es positivo puesto que $\mathcal{L}(T) = -u^2(T) + 2u^2(T) = u^2(T) > 0$. Apliquemos la condición de transversalidad (13.6) para obtener $T = 3$. Si $\mu_1 > 0$ y la primera restricción está activa, entonces $x = -t + 2$ y $\dot{x} = u = -1$; lo mismo pasa si $\mu_2 > 0$.

La solución, en el caso sin restricciones y tiempo final $T = 3$, se obtiene poniendo $x = Ae^t + Be^{-t}$ y determinando A y B de manera que se cumpla $x(0) = 2.5$ y $x(3) = 0$. Al hacer esto, notamos que queda la trayectoria $x(t) = -0.0062123e^t + 2.5062e^{-t}$, que se sale de la región permitida, como se ilustra en la figura 13.4. ¿Qué hacer entonces? La solución es que x debe seguir inicialmente la trayectoria

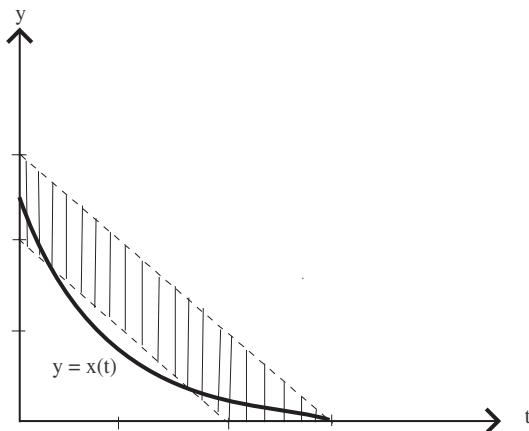


Figura 13.4: Solución sin restricción para el ejemplo 13.5.1.

$x(t) = -0.0062123e^t + 2.5062e^{-t}$ hasta un punto $(t_1, x(t_1))$ en el cual la restricción $x = -t + 2$ se activa; x continúa ahora sobre esta restricción hasta el punto $(t_2, x(t_2))$ en el cual vuelve a seguir la trayectoria $x(t) = -0.0062123e^t + 2.5062e^{-t}$ hasta llegar al punto final $(0, 3)$. La figura 13.5 ilustra estas consideraciones.

Para determinar los valores de t_1 y t_2 se utiliza la continuidad de la trayectoria x , de manera que deben satisfacer la ecuación

$$-0.0062123e^t + 2.5062e^{-t} = -t + 2.$$

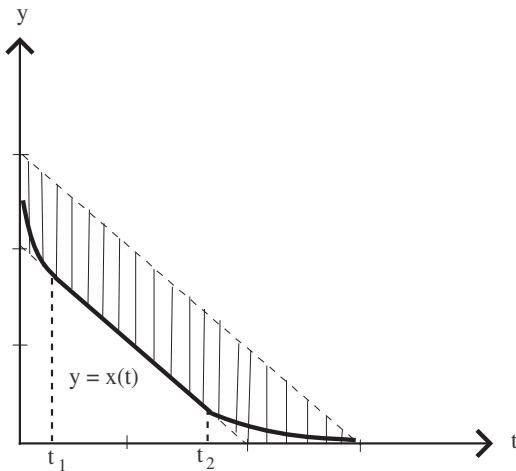


Figura 13.5: Solución con restricción para el ejemplo 13.5.1.

De esta forma se obtienen $t_1 = 0.52$ y $t_2 = 1.42$. La solución final queda entonces dada por

$$\begin{aligned}x^*(t) &= \begin{cases} -0.0062123e^t + 2.5062e^{-t} & \text{si } t \in [0, 0.52) \\ -t + 2 & \text{si } t \in [0.52, 1.42) \\ -0.0062123e^t + 2.5062e^{-t} & \text{si } t \in [1.42, 3] \end{cases}, \\u^*(t) &= \begin{cases} -0.046643e^t - 2.5466e^{-t} & \text{si } t \in [0, 0.52) \\ -1 & \text{si } t \in [0.52, 1.42) \\ -0.046643e^t - 2.5466e^{-t} & \text{si } t \in [1.42, 3] \end{cases}, \\ \lambda(t) &= \begin{cases} -0.093286e^t - 5.0932e^{-t} & \text{si } t \in [0, 0.52) \\ -2 & \text{si } t \in [0.52, 1.42) \\ -0.093286e^t - 5.0932e^{-t} & \text{si } t \in [1.42, 3] \end{cases}.\end{aligned}$$

Notemos que efectivamente λ (y en consecuencia u) es discontinua en $t = 0.52$ y $t = 1.42$ ya que

$$\begin{aligned}\lim_{t \rightarrow 0.52^-} \lambda(t) &= -3.1849, \\ \lim_{t \rightarrow 0.52^+} \lambda(t) &= -2 = \lim_{t \rightarrow 1.42^-} \lambda(t), \\ \lim_{t \rightarrow 1.42^+} \lambda(t) &= -1.617.\end{aligned}$$

Dado que la restricción en este caso es $h(x, t) = -x - t$, se tiene que $h_x = -1$ y por lo tanto los brincos van en la dirección indicada por la condición (13.8).



Ejercicios

- ▷ **13.1** Encontrar la solución general del problema

$$\max \int_0^T (u^2 + v^2) dt$$

sujeto a

$$\begin{aligned}\dot{x} &= u, \\ \dot{y} &= v,\end{aligned}$$

$$x - v = 0.$$

- ▷ **13.2** Resolver el siguiente dual al problema isoperimétrico del ejemplo 13.2.1. Encontrar la solución general para la curva de longitud mínima que pasa por $(0, 1)$ y $(1, 2)$, y que determina un área A entre la curva y el eje t .

- ▷ **13.3** Encontrar la solución general al problema

$$\max \int_0^T u^2 dt$$

sujeto a

$$\dot{x} = u,$$

$$\int_0^T x dt = A.$$

con A una constante.

- ▷ **13.4** Sea S la cantidad total de algún recurso natural, sea Q la tasa de extracción de este recurso. Suponer que este recurso es esencial para la producción de todos los bienes y servicios que proporcionan utilidad a los individuos. De esta forma, la función de utilidad del individuo representativo queda dada por $u(Q) = \frac{Q^{1-\gamma}}{1-\gamma}$, con $\gamma > 0, \gamma \neq 1$ (nótese que se trata de una función tipo CRRA como en la sección 3.1.1). Los individuos descuentan el tiempo a una tasa ρ . Resolver el siguiente problema:

$$\max \int_0^\infty u(Q) e^{-\rho t} dt$$

sujeto a

$$\int_0^\infty Q dt = S,$$

en donde S es una constante. Resolver el problema definiendo $x(t) = S - \int_0^t Q ds$, el acervo restante del recurso en el tiempo t (nótese que $x(0) = S$ y $x(T) = 0$).

- ▷ **13.5** Sea S la cantidad total de algún recurso natural, sea Q la tasa de extracción de este recurso, $C(Q)$ una función convexa y doblemente diferenciable de costos de extracción, r la tasa real y P una trayectoria de precios dada de manera exógena. Si la empresa tiene un horizonte infinito, su problema es

$$\max \int_0^\infty (PQ - C(Q)) e^{-rt} dt$$

sujeto a

$$\int_0^\infty Q dt = S.$$

Plantear el problema como un problema de control y resolverlo.

- ▷ **13.6** Resolver el siguiente problema:

$$\max \int_0^\infty \left(\frac{1}{\alpha} K^\alpha - qI \right) e^{-rt} dt$$

sujeto a

$$\begin{aligned}\dot{K} &= I - \delta K, \\ I_{\min} &\leq I \leq I_{\max}, \\ K(0) &= K_0 \text{ dado.}\end{aligned}$$

En donde I_{\min} e I_{\max} son constantes que acotan la inversión. El resto de la notación es como en la sección 13.3.1.

- a) Resolver el modelo obteniendo un sistema de ecuaciones en K y λ como en 13.3.1.
- b) Realizar el diagrama de fase correspondiente y calcular el tiempo terminal T^* al igual que en 13.3.1.

▷ 13.7 Considerar el siguiente problema

$$\max \int_0^3 (4-t)udt$$

sujeto a

$$\begin{aligned}\dot{x} &= u, \\ x - t &\leq 1, \\ x(0) &= 0, \\ x(3) &= 3, \\ u &\in [0, 2]\end{aligned}$$

y con lagrangiano asociado

$$\mathcal{L} = (4-t)u + \lambda u + \mu(1-x+t).$$

- a) Dado que $x(0) = 0$, inicialmente la restricción no está activa y $\mu = 0$. Probar que

en este caso λ es constante. Puesto que el lagrangiano es lineal en u , argumentar por qué se debe tener inicialmente $u = 2$. (Sugerencia: ver que $u = 0, u \in (0, 2)$ no son posibles.) Resolver para x si $0 < t < 1$.

- b) Probar que la restricción se activa en $t = 1$ y utilizar que $\dot{x} = u \geq 0$ para probar que se desactiva en $t = 2$ y que $u = 0$ si $t > 2$. Encontrar x si $t > 2$.
- c) Probar que mientras la restricción está activa se tiene un problema de control singular con $u = 1$ y λ lineal.
- d) Graficar las soluciones $u(t)$ y $x(t)$ con $t \in [0, 3]$.

▷ 13.8 Resolver los siguientes problemas:

- a) $\min \int_0^5 (4x + u^2)dt$ sujeto a $\dot{x} = u$, $x(0) = 10$, $x(5) = 0$ y $x \geq 6 - 2t$.
- b) $\min \int_0^T (2x + u^2)dt$ sujeto a $\dot{x} = u$, $x(0) = 2$, $x(T) = 0$ y $x \geq 1 - \frac{1}{2}t$.

CAPÍTULO 14

Elementos de programación dinámica

§14.1 Introducción

En la sección 12.7 se utilizó el concepto de función valor introducido por Bellman. De hecho, toda la teoría de control se puede desarrollar a partir del principio de optimalidad de Bellman en lugar del principio del máximo de Pontryagin. La diferencia es que, por un lado, el método de Bellman es intrínsecamente recursivo y su finalidad es encontrar la forma del valor óptimo o función valor. El método de Pontryagin, por otro lado, pone el énfasis en encontrar las trayectorias de las variables de control que conducen a este valor óptimo.¹

La teoría de control en tiempo continuo puede derivarse a partir del principio de optimalidad; sin embargo, la ecuación diferencial que se debe resolver para obtener la función valor involucra sus derivadas parciales. Esto es lo que se conoce como una ecuación diferencial parcial cuya solución está más allá de lo que podemos ofrecer en este texto. Esta es la razón principal por la que el desarrollo en tiempo continuo se hizo a partir del principio del máximo.

Si consideramos al tiempo como variable discreta, o sea t toma valores en el conjunto $\{0, 1, \dots, n, \dots\}$, la estructura recursiva del principio de optimalidad resulta sumamente útil. La idea es resolver para el último periodo, después para los dos últimos periodos, luego para los tres últimos y así sucesivamente. Cuando el número de periodos es grande, este mecanismo resulta extremadamente complejo, pero hoy en día una gran cantidad de problemas pueden resolverse mediante el uso de computadoras. Una ventaja adicional de utilizar el método de Bellman es que resulta relativamente simple introducir variables estocásticas en el problema de control. Si el tiempo es una variable continua se requieren técnicas de cálculo estocástico, las cuales están fuera del alcance de este texto. Por el contrario, cuando el tiempo es discreto, el problema se simplifica enormemente, por lo que consideraremos ese caso.

¹Como una tercera opción, el cálculo en variaciones enfatiza las trayectorias de las variables de estado.

§14.2 Estructura del problema

Para simplificar la exposición, consideremos únicamente un estado x y un control u . Sean

$$\begin{aligned}\mathcal{F} &= \{f_t : D \rightarrow \mathbb{R} \mid D \subseteq \mathbb{R}^2, t = 0, \dots, T\}, \\ \mathcal{G} &= \{g_t : E \rightarrow \mathbb{R} \mid E \subseteq \mathbb{R}^2, t = 0, \dots, T\},\end{aligned}$$

dos familias de funciones de clase \mathcal{C}^2 , y sean

$$\begin{aligned}x &: \{0, \dots, T+1\} \rightarrow \mathbb{R}, \\ u &: \{0, \dots, T\} \rightarrow \mathbb{R},\end{aligned}$$

dos funciones. Denotemos por $x(t) = x_t$ y $u(t) = u_t$ y, como antes, decimos que x es la variable de estado y u la de control. Finalmente sea V_{T+1} una función con dominio e imagen en \mathbb{R} , de clase \mathcal{C}^2 . La estructura general del problema de programación dinámica es escoger u y x que resuelvan

$$\begin{aligned}&\max \sum_{k=0}^T f_k(x_k, u_k) \\ \text{sujeto a } &x_{k+1} = g_k(x_k, u_k), k = 0, \dots, T, \\ &x_0 \text{ y } x_{T+1} \text{ dados.}\end{aligned}\tag{14.1}$$

El equivalente de la función valor dada en (12.39) es ahora

$$V_t(x_t) = \max \sum_{k=t}^T f_k(x_k, u_k),\tag{14.2}$$

que representa, igual que antes, el máximo a partir del periodo $t \in \{0, \dots, T\}$. El principio de optimalidad de Bellman dado en la sección 12.7 se expresa con lo que usualmente se conoce como **ecuación de Bellman**, que es

$$V_t(x_t) = \max\{f_t(x_t, u_t) + V_{t+1}(x_{t+1})\}.\tag{14.3}$$

Aquí el máximo es con respecto al control u_t sujeto a $x_{t+1} = g_t(x_t, u_t)$ y x_t dado. Con esta ecuación se reduce el problema original de T periodos a una sucesión de problemas de dos periodos. Este enfoque permite resolver el problema comenzando por el último periodo y luego procediendo recursivamente hacia atrás. La ecuación es válida a pesar de que se carezca de soluciones interiores; sin embargo, si éstas existen, se tiene el siguiente resultado.

Teorema 14.2.1

Si u^ y x^* resuelven el problema (14.1) y suponemos una solución interior, entonces se satisfacen*

$$\frac{\partial f_t}{\partial u_t} + \frac{dV_{t+1}}{dx_{t+1}} \frac{\partial g_t}{\partial u_t} = 0,\tag{14.4}$$

$$\frac{dV_t}{dx_t} = \frac{\partial f_t}{\partial x_t} + \frac{dV_{t+1}}{dx_{t+1}} \frac{\partial g_t}{\partial x_t},\tag{14.5}$$

$$x_{t+1} = g_t(x_t, u_t),\tag{14.6}$$

en donde $t \in \{0, \dots, T-1\}$.

A continuación se da un esbozo de la demostración.²

²Un desarrollo formal de la teoría de programación dinámica se puede encontrar en [SL89] y en [Sun96], entre otros.

Demostración

La primera ecuación es inmediata al maximizar (14.3) (suponiendo la existencia de una solución interior) con respecto a u_t , dado x_t e incluyendo la restricción $x_{t+1} = g_t(x_t, u_t)$. La segunda condición es la versión discreta de (12.46) y se obtiene al sustituir el óptimo u_t^* , que satisface (14.4), en (14.3) y después derivando con respecto a x_t . La ecuación (14.6) es simplemente la restricción original conocida como **ecuación de transición** para la variable de estado.

Como en (12.43), puede definirse a la variable de coestado como

$$\lambda_t = \frac{dV_t}{dx_t};$$

notemos que la condición (14.5) nos da la evolución de esta variable. ■

Una observación importante es que este resultado se extiende fácilmente a varias variables. De esta forma se tiene que dados $\mathbf{x}_t = (x_{1t}, \dots, x_{nt})$ y $\mathbf{u}_t = (u_{1t}, \dots, u_{mt})$, junto con las ecuaciones de movimiento $x_{j(t+1)} = g_{jt}(\mathbf{x}_t, \mathbf{u}_t)$, $j = 1, \dots, n$, y $\mathbf{x}_0, \mathbf{x}_{T+1}$ dados, la ecuación de Bellman es

$$V_t(\mathbf{x}_t) = \max\{f_t(\mathbf{x}_t, \mathbf{u}_t) + V_{t+1}(\mathbf{x}_{t+1})\},$$

con las restricciones correspondientes, y se aplican las condiciones del teorema 14.2.1 a cada una de las variables, es decir, se tienen

$$\begin{aligned} \frac{\partial f_t}{\partial u_{it}} + \sum_{j=1}^n \frac{\partial V_{t+1}}{\partial x_{j(t+1)}} \frac{\partial g_{jt}}{\partial u_{it}} &= 0, \quad i = 1, \dots, m, \\ \frac{dV_t}{dx_{jt}} &= \frac{\partial f_t}{\partial x_{jt}} + \sum_{k=1}^n \frac{\partial V_{t+1}}{\partial x_{k(t+1)}} \frac{\partial g_{kt}}{\partial x_{kt}}, \quad j = 1, \dots, n, \\ x_{j(t+1)} &= g_{jt}(\mathbf{x}_t, \mathbf{u}_t). \end{aligned}$$

Finalmente, en lo que respecta a condiciones de suficiencia, es claro que basta con que la función valor sea cóncava. No probaremos nada al respecto pero, en analogía con el caso continuo, la concavidad de las funciones f_t y la convexidad de las funciones g_t garantizan la concavidad de V_t y por lo tanto es una condición de suficiencia para obtener un máximo. Todos los ejemplos que se estudian tienen $f_t = f$ cóncava y $g_t = g$ lineal, por lo que automáticamente se tiene la suficiencia.

Ejemplo

Ej 14.2.1 Se desean encontrar u_0, \dots, u_T números no negativos cuya suma sea igual a una constante $C > 0$, de manera que la suma de sus cuadrados sea mínima. El problema se puede expresar como

$$\begin{aligned} \max & \left(- \sum_{k=0}^T u_k^2 \right) \\ \text{sujeto a } & \sum_{k=0}^T u_k = C, \quad u_k \geq 0, \quad k = 0, \dots, T. \end{aligned}$$

Para expresar el problema de la forma (14.1), definamos

$$x_t = \sum_{k=t}^T u_k,$$

de manera que $x_0 = C$ y $x_{T+1} = 0$. Adicionalmente, se cumple la ecuación de transición

$$x_{t+1} = x_t - u_t.$$

La ecuación de Bellman es

$$V_t(x_t) = \max\{-u_t^2 + V_{t+1}(x_{t+1})\}$$

sujeto a $x_{t+1} = x_t - u_t$, x_t dado.

Las condiciones de primer orden son

$$\begin{aligned} -2u_t - V'_{t+1}(x_{t+1}) &= 0, \\ V'_t(x_t) &= V'_{t+1}(x_{t+1}), \\ x_{t+1} &= x_t - u_t. \end{aligned}$$

Así, $V'_t(x_t) = -2u_t$ e, iterando un periodo al futuro, $V'_{t+1}(x_{t+1}) = -2u_{t+1}$. De aquí se concluye que $u_{t+1} = u_t$. Se tiene entonces el siguiente sistema de ecuaciones en diferencia:

$$\begin{aligned} u_{t+1} &= u_t, \\ x_{t+1} &= x_t - u_t. \end{aligned}$$

Éste se puede reescribir como una ecuación de segundo orden en x_t dada por

$$x_{t+2} - 2x_{t+1} + x_t = 0.$$

La ecuación característica tiene $\lambda = 1$ como raíz real doble, con lo cual la solución general es

$$x_t = K_1 + K_2 t.$$

Las condiciones iniciales y finales implican $K_1 = C$ y $K_2 = -\frac{C}{T+1}$. Asimismo, $0 = x_T - u_T = C - \frac{CT}{T+1} - u_T$, de manera que

$$u_T = u_{T-1} = \dots = u_0 = \frac{C}{T+1}.$$

◇◇◇◇◇◇◇◇◇

La condición de que el estado final x_{T+1} esté especificado puede relajarse de manera que se tenga un premio o valor terminal que denotaremos por $V_{T+1}(x_{T+1})$, el cual, en este caso, no está especificado pues el estado final x_{T+1} está libre. El teorema 14.2.1 cambia ligeramente ya que ahora las ecuaciones

(14.4) y (14.5) también se aplican para $t = T$, que es la resolución del problema de maximización en el estado final.

Cabe aclarar que el problema dado por (14.1) podía haberse resuelto con las técnicas de optimización estática dadas en el capítulo 10. Simplemente se piensa como un problema de optimización con restricciones en donde debemos escoger u_0, \dots, u_T y x_1, \dots, x_T (y adicionalmente x_{T+1} si no está especificada) para maximizar

$$\sum_{k=0}^T f_k(x_k, u_k)$$

sujeto a las restricciones $x_{k+1} = g_k(x_k, u_k)$, $k = 0, \dots, T$. Para resolverlo se construye el lagrangiano como

$$\mathcal{L} = \sum_{k=0}^T f_k(x_k, u_k) + \sum_{k=0}^T \lambda_{k+1}(g_k(x_k, u_k) - x_{k+1}). \quad (14.7)$$

Se deja como ejercicio al lector probar que se obtienen condiciones necesarias equivalentes a las del teorema 14.2.1. La razón para proceder utilizando el principio recursivo de Bellman es que nos permite extender la solución a horizontes de tiempo infinitos.

Ejemplo

Ej 14.2.2 Una compañía minera desea maximizar el valor presente de sus ganancias netas a lo largo del periodo de tiempo $t = 0, \dots, T + 1$. El precio de mercado del mineral extraído está dado por p . Denotemos por y_t la producción (extracción) y x_t las reservas restantes en el periodo t . El costo de extracción está dado por

$$c_t = \frac{2y_t^2}{x_t}$$

y las reservas iniciales son $x_0 = 600$ toneladas. El problema de maximización de la empresa, suponiendo que no hay descuento temporal, es

$$\max \sum_{t=0}^T \left(py_t - \frac{2y_t^2}{x_t} \right)$$

sujeto a

$$x_{t+1} = x_t - y_t,$$

$$x_0 = 600.$$

La variable de control es y y x la de estado. La ecuación de Bellman está dada por

$$V_t(x_t) = \max \left\{ \left(py_t - \frac{2y_t^2}{x_t} \right) + V_{t+1}(x_{t+1}) \right\}$$

$$x_{t+1} = x_t - y_t, \quad x_t \text{ dado},$$

y las condiciones de primer orden son

$$p - \frac{4y_t}{x_t} - V'_{t+1} = 0, \quad (14.8)$$

$$V'_t = \frac{2y_t^2}{x_t^2} + V'_{t+1}, \quad (14.9)$$

$$x_{t+1} = x_t - y_t, \quad (14.10)$$

$$V'_{T+1}(x_{T+1}) = 0. \quad (14.11)$$

La condición (14.11) se obtiene al resolver el problema en el último periodo, dado por

$$\begin{aligned} & \max_{y_T} V_{T+1}(x_{T+1}) \\ & \text{sujeto a } x_{T+1} = x_T - y_T. \end{aligned}$$

Vamos a suponer $T = 2$ y que la firma suspende su producción en $T + 1 = 3$. La ecuación (14.11) nos dice que $V'_3 = 0$ y, sustituyendo en (14.8) cuando $t = 2$, se tiene que

$$p = \frac{4y_2}{x_2} \Rightarrow \frac{y_2}{x_2} = \frac{p}{4}.$$

Sustituyendo $\frac{y_2}{x_2}$ en (14.9) para $t = 2$, se tiene

$$V'_2 = \frac{p^2}{8}.$$

Sustituyendo V'_2 en (14.8) para $t = 1$, obtenemos

$$p - \frac{4y_1}{x_1} - \frac{p^2}{8} = 0 \Rightarrow \frac{y_1}{x_1} = \frac{8p - p^2}{32}.$$

Sustituyendo $\frac{y_1}{x_1}$ en (14.9) para $t = 1$, queda

$$V'_1 = 2 \left(\frac{8p - p^2}{32} \right)^2 + \frac{p^2}{8}.$$

Sustituyendo V'_1 en (14.8) para $t = 0$, se tiene

$$p - \frac{4y_0}{x_0} - \left[2 \left(\frac{8p - p^2}{32} \right)^2 + \frac{p^2}{8} \right] = 0.$$

Poniendo el valor inicial $x_0 = 600$, se llega a

$$y_0 = 150 \left[p - 2 \left(\frac{8p - p^2}{32} \right)^2 - \frac{p^2}{8} \right]$$

y, sustituyendo en (14.10) para $t = 0$,

$$x_1 = 600 - 150 \left[p - 2 \left(\frac{8p - p^2}{32} \right)^2 - \frac{p^2}{8} \right].$$

Dado que $\frac{y_1}{x_1} = \frac{8p-p^2}{32}$ se obtiene

$$y_1 = \left[\frac{8p-p^2}{32} \right] \left\{ 600 - 150 \left[p - 2 \left(\frac{8p-p^2}{32} \right)^2 - \frac{p^2}{8} \right] \right\}.$$

Sustituyendo en (14.10) para $t = 1$,

$$x_2 = \left\{ 600 - 150 \left[p - 2 \left(\frac{8p-p^2}{32} \right)^2 - \frac{p^2}{8} \right] \right\} \left[1 - \frac{8p-p^2}{32} \right]$$

y dado que $\frac{y_2}{x_2} = \frac{p}{4}$, entonces

$$y_2 = \frac{p}{4} \left\{ 600 - 150 \left[p - 2 \left(\frac{8p-p^2}{32} \right)^2 - \frac{p^2}{8} \right] \right\} \left[1 - \frac{8p-p^2}{32} \right].$$

Sustituyendo en (14.10) para $t = 2$,

$$x_3 = \left\{ 600 - 150 \left[p - 2 \left(\frac{8p-p^2}{32} \right)^2 - \frac{p^2}{8} \right] \right\} \left[1 - \frac{8p-p^2}{32} \right] \left[1 - \frac{p}{4} \right].$$

Finalmente, dado que la firma cierra en $t = 3$ se tiene que $y_3 = 0$.

◇◇◇◇◇◇◇◇◇

Es claro que este proceso es algebraicamente tedioso, sin embargo, el algoritmo recursivo es evidente y es sumamente sencillo realizar este proceso en alguna hoja de cálculo. Esto se deja al lector como ejercicio.

§14.3 Problemas con descuento temporal

Sea ρ la tasa de descuento temporal. El problema general (14.1), con el estado final libre, se puede plantear ahora como

$$\begin{aligned} & \max \sum_{k=0}^T \left(\frac{1}{1+\rho} \right)^k f_k(x_k, u_k) \\ & \text{sujeto a } x_{k+1} = g_k(x_k, u_k), \\ & \quad x_0 \text{ dado, } x_{T+1} \text{ libre.} \end{aligned} \tag{14.12}$$

De manera que la ecuación de Bellman es

$$\begin{aligned} V_t(x_t) &= \max \{ \beta^t f_t(x_t, u_t) + V_{t+1}(x_{t+1}) \}, \\ x_{t+1} &= g_t(x_t, u_t), \quad x_t \text{ dado.} \end{aligned}$$

en donde la función valor $V_t(x_t)$ se mide en unidades al tiempo $t = 0$ y $\beta \equiv \frac{1}{1+\rho}$ es el **factor de descuento temporal**. Si expresamos a la función valor en unidades del periodo t , esta ecuación queda como

$$\bar{V}_t(x_t) = \max \{ f_t(x_t, u_t) + \beta \bar{V}_{t+1}(x_{t+1}) \},$$

La función \bar{V}_t es la función valor en tiempo corriente t y se relaciona con la función valor en tiempo presente V_t (donde V_t se mide en unidades del tiempo inicial $t = 0$) mediante $\bar{V}_t \beta^t = V_t$. Siempre que exista el factor de descuento consideraremos la función valor y la ecuación de Bellman en tiempo corriente. Abusando de la notación, al igual que en el caso de los hamiltonianos en tiempo corriente y en tiempo presente, la denotaremos simplemente por V_t .

Las condiciones de primer orden se reescriben ahora como

$$\begin{aligned} \frac{\partial f_t}{\partial u_t} + \beta \frac{dV_{t+1}}{dx_{t+1}} \frac{\partial g_t}{\partial u_t} &= 0, \\ \frac{dV_t}{dx_t} &= \frac{\partial f_t}{\partial x_t} + \beta \frac{dV_{t+1}}{dx_{t+1}} \frac{\partial g_t}{\partial x_t}, \\ x_{t+1} &= g_t(x_t, u_t). \end{aligned} \tag{14.13}$$

De esta forma se introduce el descuento temporal de manera trivial. Normalmente tendremos que ρ es la tasa subjetiva de descuento o bien la tasa real r . En economía, es común que las funciones f_k y g_k no dependan explícitamente del tiempo, de manera que el problema (14.12) es “casi autónomo” pues el tiempo sólo aparece como potencia en el factor de descuento. Al igual que en el caso continuo, la función valor en tiempo corriente transforma al problema en uno autónomo (en la sección 14.6 se da la definición formal de un problema autónomo). Por esta razón, de aquí en adelante omitiremos el subíndice t en la función valor en tiempo corriente.

§14.4 Problemas con horizonte infinito

Al igual que en el caso de tiempo continuo, cuando el horizonte es infinito se debe poder garantizar la convergencia de la serie $\sum_{k=0}^{\infty} f_t(x_t, u_t)$. Si existe un factor de descuento $\beta \in (0, 1)$, la convergencia se da si f_t es acotada (misma cota para toda t) o creciente, siempre y cuando $\sum_{t=0}^{\infty} \beta^t f_t(x_t, u_t)$ converja. Dado que en economía casi siempre aparece el factor de descuento, suponemos que el problema general está dado por

$$\begin{aligned} \max & \sum_{t=0}^{\infty} \beta^t f_t(x_t, u_t) \\ \text{sujeto a } & x_{t+1} = g_t(x_t, u_t), \quad x_0 \text{ dado.} \end{aligned} \tag{14.14}$$

No tiene sentido empezar con el último periodo y proceder recursivamente hacia $t = 0$ dado que no hay un periodo terminal. No obstante, la función valor se define, igual que antes, como el máximo a partir del periodo t , es decir,

$$V_t(x_t) = \max \sum_{k=t}^{\infty} f_k(x_k, u_k).$$

La ecuación de Bellman es, una vez mas, (14.3). Como ésta sólo involucra dos períodos de tiempo, las condiciones de primer orden son las mismas que en el teorema 14.2.1. De las dos primeras condiciones se puede eliminar V'_t para obtener una expresión análoga a la ecuación de Euler utilizada en cálculo

en variaciones (véase el capítulo 11). El procedimiento es el siguiente: se despeja $\frac{dV_{t+1}}{dx_{t+1}}$ de la primera condición y se sustituye en la segunda para obtener

$$\frac{dV_t}{dx_t} = \frac{\partial f_t}{\partial x_t} - \frac{\frac{\partial f_t}{\partial u_t}}{\frac{\partial g_t}{\partial u_t}} \frac{\partial g_t}{\partial x_t}.$$

Iterando un periodo y sustituyendo $\frac{dV_{t+1}}{dx_{t+1}}$ una vez mas en la primera condición se llega a

$$\frac{\partial f_t}{\partial u_t} + \left(\frac{\partial f_{t+1}}{\partial x_{t+1}} - \frac{\frac{\partial f_{t+1}}{\partial u_{t+1}}}{\frac{\partial g_{t+1}}{\partial u_{t+1}}} \frac{\partial g_{t+1}}{\partial x_{t+1}} \right) \frac{\partial g_t}{\partial u_t} = 0. \quad (14.15)$$

A esta última ecuación (en diferencias) se la conoce como **ecuación de Euler** y junto con la ecuación de transición $x_{t+1} = g_t(x_t, u_t)$ forma un sistema dinámico discreto en x_t y u_t , el cual debemos resolver para encontrar la respuesta al problema original. Dicho sistema es, en general, muy complicado de resolver explícitamente ya que se trata de un sistema, probablemente no lineal, de ecuaciones en diferencias. No obstante, el análisis cualitativo puede ser útil como se ilustra a continuación.

§14.5 Modelo de Ramsey discreto

Esta sección tiene como propósito no sólo exponer la versión discreta del modelo de Ramsey, sino ilustrar la importancia de la secuencia de eventos cuando se trabaja con períodos de tiempo discretos con cierta duración.

La familia Robinson ahora mide el tiempo en forma discreta y debe resolver el siguiente problema:

$$\max \sum_{t=1}^{\infty} \beta^t u(c_t)$$

sujeto a la restricción presupuestal

$$f(k_{t-1}) = c_t + k_t + \delta k_{t-1}, \quad k_0 > 0 \text{ dado}, \quad (14.16)$$

donde las funciones de utilidad y producción, $u(c_t)$ y $f(k_t)$, satisfacen las propiedades usuales y el capital se deprecia a la tasa δ . La variable de control es el consumo c_t y la de estado el capital k_t . El problema de la familia es maximizar el valor presente de su utilidad sujeto a la restricción presupuestal. Se trata, pues, de una versión discreta del modelo de Ramsey.

La restricción presupuestal requiere cierto análisis debido a los índices de tiempo. Los períodos están medidos en alguna unidad, por ejemplo en semanas. De esta forma, cada periodo tiene como duración una semana. La producción se realiza al final de cada semana, digamos los domingos. Parte del bien producido se consume durante la siguiente semana y el resto se utiliza en inversión bruta. Ésta se divide en capital nuevo y en sustituir el capital que se depreció en la semana anterior. Lo que es crucial aquí, y es reflejado por la restricción (14.16), es que lo que se consume e invierte en t , fue producido en $t-1$.

La ecuación de Bellman en tiempo corriente $t-1$ es

$$V(k_{t-1}) = \max \{u(c_t) + \beta V(k_t)\}$$

sujeto a $k_t = f(k_{t-1}) - c_t - \delta k_{t-1}$, k_{t-1} dado.

Las condiciones de primer orden son

$$u'(c_t) - \beta V'(k_t) = 0, \quad (14.17)$$

$$V'(k_{t-1}) = \beta V'(k_t)(f'(k_{t-1}) - \delta), \quad (14.18)$$

$$k_t = f(k_{t-1}) - c_t - \delta k_{t-1}. \quad (14.19)$$

El proceso para obtener la ecuación de Euler es el siguiente: de la ecuación (14.17) se tiene

$$\beta V'(k_t) = u'(c_t)$$

y sustituimos esta expresión en (14.18) para obtener

$$V'(k_{t-1}) = u'(c_t)(f'(k_{t-1}) - \delta).$$

Iterando un periodo llegamos a

$$V'(k_t) = u'(c_{t+1})(f'(k_t) - \delta);$$

sustituyendo esta expresión en (14.17) y reacomodando términos se tiene

$$\frac{u'(c_t)}{u'(c_{t+1})} = \beta(f'(k_t) - \delta).$$

Si reescribimos $\beta = \frac{1}{1+\rho}$ y $f'(k_t) = R_t = 1 + r_t$, en donde ρ es la tasa de descuento temporal, R_t denota los rendimientos del capital y r_t la tasa real, la ecuación de Euler queda como

$$\frac{u'(c_t)}{u'(c_{t+1})} = \frac{1 + r_t - \delta}{1 + \rho}. \quad (14.20)$$

La ecuación (14.20) es el equivalente en tiempo discreto de la ecuación

$$\frac{\dot{u}'}{u'} = \rho - (r - \delta),$$

obtenida en el modelo de Ramsey. Su interpretación es clara: dice que la tasa marginal de sustitución del consumo en t y el consumo en $t + 1$ es igual al cociente $\frac{1+r_t-\delta}{1+\rho}$, en donde el numerador representa el costo de oportunidad marginal de consumir en t (igual a los rendimientos netos de depreciación) y el denominador representa el costo de oportunidad marginal de esperar un periodo y consumir en $t + 1$. Si $r_t - \delta > \rho$, la familia es “menos impaciente” que el mercado y prefiere posponer el consumo en t por el de $t + 1$; esto es claro, ya que

$$\frac{u'(c_t)}{u'(c_{t+1})} = \frac{1 + r_t - \delta}{1 + \rho} > 1$$

implica $c_{t+1} > c_t$. Análogamente, si $\rho > r_t - \delta$ se tiene que la familia es “más impaciente” que el mercado y $c_t > c_{t+1}$. Finalmente si r_t es constante y $\rho = r - \delta$ (tasa real neta igual a tasa subjetiva), el consumo se mantiene constante a lo largo del tiempo. Como era de esperarse, éste es un resultado idéntico al obtenido anteriormente con el modelo de Ramsey (en aquel caso tomamos $\delta = 0$).

Consideremos ahora la función de utilidad $u(c_t) = \ln c_t$. Como en el caso continuo, haremos el análisis del sistema de ecuaciones en diferencias dado por la ecuación de Euler (14.20) y la restricción presupuestal (14.19), llevada un periodo al futuro:

$$\begin{aligned} k_{t+1} &= f(k_t) - c_{t+1} - \delta k_t, \\ \frac{u'(c_t)}{u'(c_{t+1})} &= \beta(f'(k_t) - \delta), \end{aligned}$$

o bien,

$$k_{t+1} = f(k_t) - \delta k_t - \beta(f'(k_t) - \delta)c_t, \quad (14.21)$$

$$c_{t+1} = \beta(f'(k_t) - \delta)c_t. \quad (14.22)$$

El estado estacionario, diferente del $(0, 0)$, se denota por (k^*, c^*) y éste satisface

$$\begin{aligned} k^* &= f(k^*) - \delta k^* - \beta(f'(k^*) - \delta)c^*, \\ 1 &= \beta(f'(k^*) - \delta). \end{aligned}$$

La matriz jacobiana del sistema dado por las ecuaciones (14.21) y (14.22) es

$$J(k^*, c^*) = \begin{pmatrix} \frac{1}{\beta} - \beta f''(k^*)c^* & -1 \\ \beta f''(k^*)c^* & 1 \end{pmatrix}.$$

No es difícil ver que el discriminante es positivo, de manera que los valores propios son reales. Adicionalmente, el polinomio característico es

$$p(\lambda) = \lambda^2 - (1 + \frac{1}{\beta} - \beta f''(k^*)c^*)\lambda + \frac{1}{\beta},$$

con lo cual se tiene

$$\begin{aligned} p(1) &= \beta f''(k^*)c^* < 0, \\ p(-1) &= 2 + \frac{2}{\beta} - \beta f''(k^*)c^* > 0. \end{aligned}$$

Recordando lo visto en la sección 7.4 sobre sistemas lineales discretos, se tiene que los valores propios λ_1 y λ_2 son ambos reales y cumplen $|\lambda_1| > 1$ y $\lambda_2 < 1$, con lo cual (k^*, c^*) es un punto silla, al igual que en la versión continua del modelo.

Considérese ahora la siguiente variante del problema anterior. La producción se realiza al principio de la semana, digamos el lunes, y durante el resto de la semana se consume o se invierte. La inversión se divide en capital que se utilizará para la producción de la semana siguiente y en reposición del capital que se deprecie durante la semana en curso. La diferencia es que ahora lo consumido en t se produce en t y no en $t - 1$, como en el caso anterior. El problema queda expresado ahora como

$$\max \sum_{t=0}^{\infty} \beta^t u(c_t)$$

sujeto a la restricción presupuestal

$$f(k_t) = c_t + k_{t+1} + \delta k_t, \quad k_0 > 0 \text{ dado.} \quad (14.23)$$

El lector debe observar la diferencia sutil que existe entre las expresiones (14.16) y (14.23). Se deja como ejercicio mostrar que la ecuación de Euler ahora está dada por

$$\frac{u'(c_t)}{u'(c_{t+1})} = \beta(f'(k_{t+1}) - \delta).$$

§14.6 Encontrando la función valor

Bajo ciertos supuestos generales, es posible proceder a encontrar la función valor explícitamente, para luego determinar las trayectorias óptimas sin utilizar las condiciones de Euler. En esta sección describiremos dos métodos para encontrar la función valor. La siguiente definición precisa la última observación que se hizo en la sección 14.3.

Definición 14.6.1 Decimos que el problema (14.1) es **autónomo** si $f_t = f$ y $g_t = g$ para todo t , es decir, si las formas funcionales no cambian entre períodos. Es claro que ésta es la versión discreta de un problema autónomo cuando la variable t es continua. En el caso autónomo, la función valor en tiempo corriente depende del estado x_t pero no del índice de tiempo, por lo que la denotaremos simplemente por V .

Cuando el problema es autónomo y el factor de descuento temporal β no es trivial, i.e. $\beta \in (0, 1)$, es posible encontrar soluciones analíticas al problema (14.12) debido al siguiente teorema.

Teorema 14.6.2

Consideremos la versión autónoma del problema de maximización (14.14) con $\beta \in (0, 1)$. Sea D la imagen de x , y supongamos que D es un conjunto no vacío y compacto. La función valor V (en tiempo corriente) es una función continua con dominio D y es la única función continua que satisface la ecuación de Bellman.

Demostración

Éste es sólo un esbozo de la demostración.³ Sea F una función continua que satisface la ecuación de Bellman, es decir,

$$F(x_t) = \max\{f(x_t, u_t) + \beta F(x_{t+1})\}.$$

Aplicando esta ecuación de manera recursiva, llegamos a que para toda T se cumple

$$F(x_t) = \max\left\{\sum_{k=t}^{T-1} \beta^k f(x_k, u_k) + \beta^T F(x_T)\right\}.$$

En el límite, cuando $T \rightarrow \infty$, se tiene que $\beta^T F(x_T) \rightarrow 0$ ya que $\beta \in (0, 1)$ y por lo tanto $F = V$. ■

³Una demostración formal se puede encontrar en [SL89].

El teorema anterior nos proporciona un método para encontrar la función valor. Buscamos una función V tal que se cumpla la ecuación de Bellman (en tiempo corriente)

$$V(x_t) = \max\{f(x_t, u_t) + \beta V(x_{t+1})\}$$

sujeto a $x_{t+1} = g_t(x_t, u_t)$, x_t dado.

Definición 14.6.3 *Dados dos espacios vectoriales W_1 y W_2 , un **operador** es una función*

$$\mathcal{O} : W_1 \rightarrow W_2.$$

*Si \mathcal{O} es lineal se dice que es un **operador lineal**. Si $W_2 = \mathbb{F}$, en donde \mathbb{F} es el campo escalar (siempre igual a \mathbb{R} en este texto), se dice que \mathcal{O} es una funcional (véase definición 11.1.1).*

Ejemplos

Ej 14.6.1 Sea $W = C[a, b] = \{f : [a, b] \rightarrow \mathbb{R} \mid f \text{ es continua}\}$. Entonces $I : W \rightarrow W$ definido por

$$I[f](x) = \int_a^x f(t) dt$$

es un operador lineal.

Ej 14.6.2 $\mathcal{O} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ definido por $\mathcal{O}(x, y) = (1 + x, y)$ es un operador no lineal, pues $\mathcal{O}(0, 0) \neq (0, 0)$.

Ej 14.6.3 Sea $W_n = C^n(a, b) = \{f : (a, b) \rightarrow \mathbb{R} \mid f, f', \dots, f^{(n)} \text{ son continuas}\}$, $n \in \mathbb{N}$. Entonces $D : C^n(a, b) \rightarrow C^{n-1}(a, b)$ definido por

$$Df(x) = f'(x)$$

es un operador lineal.

Al igual que en el teorema 14.6.2, sea D la imagen de x . Se define $W = \{V : D \rightarrow \mathbb{R} \mid V \text{ es continua}\}$. Podemos pensar la ecuación de Bellman como una ecuación funcional de la forma

$$V = \mathcal{O}(V),$$

en donde $\mathcal{O} : W \rightarrow W$ es un operador del espacio vectorial de funciones en sí mismo. Resolver la ecuación equivale a encontrar un **punto fijo** del operador, es decir, una función V tal que $\mathcal{O}(V) = V$.

Para encontrar esta función comenzamos con cualquier función inicial F_0 , continua y acotada, después aplicamos el operador, esto es, se elige u_t de tal manera que resuelva el problema

$$\max\{f(x_t, u_t) + \beta F_0(x_{t+1})\}$$

sujeto a

$$x_{t+1} = g_t(x_t, u_t).$$

Si u_t^* es solución a este problema, entonces definimos

$$F_1(x_t) = f(x_t, u_t^*) + \beta F_0(x_{t+1}).$$

Ahora aplicamos el operador a F_1 y llamamos al resultado F_2 , y así sucesivamente, de manera que

$$F_{n+1}(x_t) = \max\{f(x_t, u_t) + \beta F_n(x_{t+1})\}.$$

Podemos ahora probar la siguiente proposición.

Proposición 14.6.4 *Si se tiene una sucesión de funciones $\{F_n\}$ como arriba, entonces $\lim_{n \rightarrow \infty} F_n = V$, es decir, el límite existe y es un punto fijo del operador, el cual además es único.*

Solo daremos la demostración para el caso en que las funciones iniciales difieren por una constante y remitimos al lector al texto citado anteriormente para que vea todos los detalles.

Demostración (esbozo)

Si $G_0 = F_0 + A$ con $A > 0$, es otra función inicial, construyamos la sucesión $\{G_n\}$ al igual que arriba. Es relativamente sencillo ver que

$$G_n(x_t) = F_n(x_t) + \beta^n A,$$

de manera que el límite de las dos sucesiones es el mismo pues $\lim_{n \rightarrow \infty} \beta^n A = 0$. Se puede probar que esta propiedad se cumple para cualquier función inicial G_0 . Llaremos a este límite V ; entonces V es solución de la ecuación de Bellman y el teorema 14.6.2 nos garantiza que es única. ■

Así, para encontrar la solución podemos empezar con cualquier función inicial (continua y acotada), por ejemplo $F_0 = 0$, y construir la sucesión anterior para obtener como límite la función valor .

Dado que la solución es única, otro método para encontrarla es “adivinar” una solución (la forma funcional de la función valor) y probar que ésta satisface la ecuación de Bellman. El problema puede simplificarse aún más cuando la función f es de la clase **HARA** (véase el ejercicio 3.10). Este tipo de funciones incluye las de aversión absoluta al riesgo constante (CARA) y de aversión relativa al riesgo constante (CRRA), en particular las funciones logarítmicas y exponenciales, y también incluye las funciones cuadráticas. Es conocido que cuando f es de este tipo, la función valor tiene la misma forma funcional [Mer69]. A continuación resolveremos dos ejemplos utilizando cada uno de estos métodos para ilustrar los pasos a seguir.

Ejemplos

Ej 14.6.4 (Brock-Mirman) ⁴ Consideremos el siguiente problema del agente representativo:

$$\max \sum_{t=0}^{\infty} \beta^t \ln c_t$$

sujeto a la restricción presupuestal

$$Ak_t^{\alpha} = c_t + k_{t+1}, \quad k_0 > 0 \text{ dado,}$$

⁴Véase [BM72].

en donde c_t (la variable de control) y k_t (la de estado) representan el consumo y el capital en el periodo t , β es el factor de descuento temporal y $A > 0$, $\alpha \in (0, 1)$, son constantes. La cantidad Ak_t^α representa la producción en el periodo t , que en este caso es igual al consumo c_t en el mismo periodo mas la inversión k_{t+1} , y $\ln c_t$ es la utilidad proporcionada por el consumo en t .

Sea $F_0 = 0$ (continua y acotada) y constrúyase la sucesión

$$F_{n+1}(k_t) = \max\{\ln c_t + \beta F_n(k_{t+1})\}$$

sujeto a

$$Ak_t^\alpha = c_t + k_{t+1}, \quad k_t > 0 \text{ dado.}$$

Se tiene entonces que

$$F_1(k_t) = \max\{\ln c_t\},$$

sujeto a la restricción presupuestal. El máximo se obtiene cuando c_t es lo mas grande posible, o sea, explícitamente, cuando $k_{t+1} = 0$ y $c_t = Ak_t^\alpha$ por lo que

$$F_1(k_t) = \ln Ak_t^\alpha = \ln A + \alpha \ln k_t.$$

El siguiente término en la sucesión es

$$F_2(k_t) = \max\{\ln c_t + \beta F_1(k_{t+1})\}$$

dada la restricción presupuestal. Sustituyendo para F_1 se obtiene

$$F_2(k_t) = \max\{\ln c_t + \beta(\ln A + \alpha \ln k_{t+1})\},$$

sujeto a la restricción presupuestal. Resolviendo el problema de maximización se llega a

$$\begin{aligned} c_t &= \frac{1}{1 + \alpha\beta} Ak_t^\alpha, \\ k_{t+1} &= \frac{\alpha\beta}{1 + \alpha\beta} Ak_t^\alpha. \end{aligned}$$

De manera que

$$\begin{aligned} F_2(k_t) &= \ln \left(\frac{A}{1 + \alpha\beta} \right) + \alpha \ln k_t + \beta \ln A + \alpha\beta \ln \left(\frac{\alpha\beta A}{1 + \alpha\beta} \right) + \alpha^2\beta \ln k_t \\ &= \left[\ln \left(\frac{A}{1 + \alpha\beta} \right) + \beta \ln A + \alpha\beta \ln \left(\frac{\alpha\beta A}{1 + \alpha\beta} \right) \right] + \alpha(1 + \alpha\beta) \ln k_t. \end{aligned}$$

Análogamente se tiene

$$\begin{aligned} F_3(k_t) &= \left\{ \ln \left(\frac{A}{1 + \alpha\beta + \alpha^2\beta^2} \right) + \beta \ln \left(\frac{A}{1 + \alpha\beta} \right) + \beta^2 \ln A \right. \\ &\quad \left. + \alpha\beta(1 + \alpha\beta) \ln \left(\frac{\alpha\beta A(1 + \alpha\beta)}{1 + \alpha\beta + \alpha^2\beta^2} \right) + \beta \left[\alpha\beta \ln \left(\frac{\alpha\beta A}{1 + \alpha\beta} \right) \right] \right\} \\ &\quad + \alpha(1 + \alpha\beta + \alpha^2\beta^2) \ln k_t \end{aligned}$$

y

$$\begin{aligned}
 F_4(k_t) = & \left\{ \ln \left(\frac{A}{1 + \alpha\beta + \alpha^2\beta^2 + \alpha^3\beta^3} \right) + \beta \ln \left(\frac{A}{1 + \alpha\beta + \alpha^2\beta^2} \right) \right. \\
 & + \beta^2 \ln \left(\frac{A}{1 + \alpha\beta} \right) + \beta^3 \ln A \\
 & + \alpha\beta(1 + \alpha\beta + \alpha^2\beta^2) \ln \left(\frac{\alpha\beta A(1 + \alpha\beta)}{1 + \alpha\beta + \alpha^2\beta^2 + \alpha^3\beta^3} \right) \\
 & + \beta \left[\alpha\beta(1 + \alpha\beta) \ln \left(\frac{\alpha\beta A(1 + \alpha\beta)}{1 + \alpha\beta + \alpha^2\beta^2} \right) \right] \\
 & \left. + \beta^2 \left[\alpha\beta \ln \left(\frac{\alpha\beta A}{1 + \alpha\beta} \right) \right] \right\} + \alpha(1 + \alpha\beta + \alpha^2\beta^2 + \alpha^3\beta^3) \ln k_t.
 \end{aligned}$$

Siguiendo este proceso y tomando el límite de F_n cuando $n \rightarrow \infty$ se tiene⁵

$$\begin{aligned}
 \lim_{n \rightarrow \infty} F_n(k_t) &= V(k_t) \\
 &= \frac{1}{1 - \beta} \left\{ \ln[A(1 - \alpha\beta)] + \frac{\alpha\beta}{1 - \alpha\beta} \ln \alpha\beta A \right\} + \frac{\alpha}{1 - \alpha\beta} \ln k_t.
 \end{aligned}$$

Ej 14.6.5 (Coeficientes indeterminados) Como se vio en el ejemplo anterior, encontrar directamente el límite de la sucesión $\{F_n\}$ puede ser un proceso muy tedioso. Vamos a usar ahora el hecho de que, dada una función de utilidad logarítmica, la función valor debe tener la misma forma. Es decir, se propone una función valor (en tiempo corriente) de la forma

$$V(k_t) = E + F \ln k_t,$$

donde el problema es ahora encontrar los coeficientes indeterminados E y F . Sustituyendo esta función valor en la ecuación de Bellman se tiene,

$$E + F \ln k_t = \max \{ \ln c_t + \beta(E + F \ln k_{t+1}) \}$$

sujeto a

$$Ak_t^\alpha = c_t + k_{t+1}, \quad k_t \text{ dado.}$$

Resolviendo el problema de maximización se obtiene que el valor óptimo del consumo está dado por

$$c_t = \frac{1}{1 + \beta F} Ak_t^\alpha \tag{14.24}$$

y por lo tanto los valores óptimos del capital evolucionan de acuerdo con

$$k_{t+1} = \frac{\beta F}{1 + \beta F} Ak_t^\alpha. \tag{14.25}$$

Sustituyendo estos valores en la ecuación de Bellman se tiene que

$$\begin{aligned}
 E + F \ln k_t &= \ln \frac{A}{1 + \beta F} + \alpha \ln k_t + \beta E + \beta F \ln \frac{A\beta F}{1 + \beta F} + \alpha\beta F \ln k_t \\
 &= \left[\beta E + \ln \frac{A}{1 + \beta F} + \beta F \ln \frac{A\beta F}{1 + \beta F} \right] + \alpha(1 + \beta F) \ln k_t.
 \end{aligned}$$

⁵Porque mostrarlo no es enteramente trivial, se dejan los detalles al lector interesado.

Igualando los coeficientes de ambos lados de la ecuación se tiene

$$\begin{aligned} E &= \beta E + \ln \frac{A}{1 + \beta F} + \beta F \ln \frac{A\beta F}{1 + \beta F}, \\ F &= \alpha(1 + \beta F). \end{aligned}$$

Resolviendo para E y F , tenemos

$$\begin{aligned} F &= \frac{\alpha}{1 - \alpha\beta}, \\ E &= \frac{1}{1 - \beta} \left[\ln A(1 - \alpha\beta) + \frac{\alpha\beta}{1 - \alpha\beta} \ln \alpha\beta A \right], \end{aligned}$$

de manera que se obtiene el mismo resultado que en el ejemplo anterior.



§14.6.1 El viñedo de Weitzman

Éste es un modelo⁶ que se puede resolver de forma analítica y en donde el sistema dinámico resultante posee 2-ciclos como los estudiados en el capítulo 6. La historia va como sigue. Lorenzo es un vinatero que lleva una vida bastante sencilla y frugal. Su dieta cotidiana consiste de pan y vino, ambos bienes producidos por él mismo. Diariamente, nuestro vinatero debe asignar sus horas de trabajo entre pisar uvas para obtener jugo de uva, que con el tiempo se transforma en vino, y hornear hogazas de pan. Una amable vecina le proporciona la harina, la levadura y las uvas requeridas para la fabricación del pan y del jugo de uva, de manera que el único insumo que Lorenzo debe proporcionar es su trabajo.

Considérese la jornada laboral diaria como una unidad de trabajo y el tiempo t se mide en días. Lorenzo no tiene gran interés en ser particularmente productivo y su tecnología de producción es la siguiente: una unidad de trabajo produce una hogaza de pan, h_t , o alternativamente un galón de jugo de uva, j_t . De esta forma, en cualquier periodo t , Lorenzo produce una cantidad h_t de pan y una cantidad $1 - h_t$ de jugo. Para obtener vino v_{t+1} en el siguiente periodo, simplemente se deja reposar el jugo de uva de manera que $v_{t+1} = j_t = 1 - h_t$.

Como es común, se descuenta el futuro con un factor de descuento $\beta \in (0, 1)$. La utilidad de Lorenzo proviene del consumo de pan y vino,⁷ y la función de utilidad es la misma en cada periodo y está dada por

$$u(h_t, v_t) = h_t^{\frac{1}{2}} v_t^{\frac{1}{2}}.$$

Lorenzo está convencido de que su dieta y estilo de vida son ideales y espera mantenerse vivo indefinidamente. Su problema de maximización puede plantearse como sigue:

$$\max \sum_{t=0}^{\infty} \beta^t \left(h_t^{\frac{1}{2}} v_t^{\frac{1}{2}} \right)$$

sujeto a $v_{t+1} = 1 - h_t$, v_0 dado y $h_t \in [0, 1]$. Tomemos h_t como control y v_t como estado.

⁶Este modelo está basado en un ejemplo atribuido, en el folclor de la teoría del capital, a Martin Weitzman.

⁷Notemos que Lorenzo es un hombre sensato que no obtiene utilidad alguna del jugo de uva.

La ecuación de Bellman (en tiempo corriente) queda dada por

$$V(v_t) = \max \left\{ h_t^{\frac{1}{2}} v_t^{\frac{1}{2}} + \beta V(v_{t+1}) \right\}$$

con $v_{t+1} = 1 - h_t$, v_t dado.

La función de utilidad es de tipo HARA por lo que proponemos la misma forma funcional para la función valor, en este caso

$$V(v_t) = A(Bv_t + C)^{\frac{1}{2}}.$$

Sustituyendo en la ecuación de Bellman obtenemos

$$A(Bv_t + C)^{\frac{1}{2}} = \max \left\{ h_t^{\frac{1}{2}} v_t^{\frac{1}{2}} + \beta A(Bv_{t+1} + C)^{\frac{1}{2}} \right\}, \quad (14.26)$$

con $v_{t+1} = 1 - h_t$, v_t dado. Resolviendo el problema de maximización del lado derecho de esta ecuación se obtiene

$$h_t = \frac{(B+C)v_t}{(\beta AB)^2 + Bv_t}, \quad (14.27)$$

y sustituyendo este valor en (14.26) se llega a

$$A(Bv_t + C)^{\frac{1}{2}} = \frac{(B+C)^{\frac{1}{2}}(v_t + (\beta A)^2 B)}{((\beta AB)^2 + Bv_t)^{\frac{1}{2}}}. \quad (14.28)$$

Para obtener la misma forma funcional de ambos lados se requiere $B = 1$, con lo que se tiene

$$A(v_t + C)^{\frac{1}{2}} = (1 + C)^{\frac{1}{2}}(v_t + (\beta A)^2)^{\frac{1}{2}}. \quad (14.29)$$

Para tener la igualdad entre ambos lados se requiere que se cumplan

$$\begin{aligned} A &= (1 + C)^{\frac{1}{2}}, \\ C &= (\beta A)^2. \end{aligned}$$

Resolviendo para A y C , obtenemos

$$A = \left(\frac{1}{1 - \beta^2} \right)^{\frac{1}{2}}, \quad C = \frac{\beta^2}{1 - \beta^2},$$

con lo cual, sustituyendo en (14.27), se llega a

$$\begin{aligned} h_t &= \frac{v_t}{\beta^2 + (1 - \beta^2)v_t} \\ &= \frac{1 - h_{t-1}}{1 - (1 - \beta^2)h_{t-1}} \end{aligned} \quad (14.30)$$

y, por lo tanto,

$$v_t = \frac{\beta^2(1 - v_{t-1})}{\beta^2(1 - v_{t-1}) + v_{t-1}}. \quad (14.31)$$

Claramente se tiene $h_t, v_t \in [0, 1]$.

Podemos realizar el análisis cualitativo de estos sistemas discretos; tomemos por ejemplo el sistema descrito por (14.30). Un punto fijo h^* se obtiene al resolver

$$h^* = \frac{1 - h^*}{1 - (1 - \beta^2)h^*};$$

por lo tanto, $h^* = \frac{1}{1+\beta}$. Asimismo, si pensamos que el sistema (14.30) está dado por $h_t = f(h_{t-1})$, entonces $f^{-1} = f$, de manera que $f(f(h)) = h$ para toda h , es decir, con excepción del punto fijo h^* dado arriba, todos los puntos $h_t \in [0, 1]$, son puntos de periodo 2, de manera que el sistema dado por (14.30) posee 2-ciclos. Esto puede observarse en el diagrama de la figura 14.1. (El análisis del sistema dado por 14.31 es semejante.) Lorenzo tiene dos opciones: si la cantidad inicial de vino está dada por $v_0 = v^* = \frac{\beta}{1+\beta}$, entonces el sistema se queda en su estado estacionario dado por (v^*, h^*) en todos los periodos subsecuentes; en cambio, si $v_0 \neq v^*$, el sistema oscila entre los valores (v_0, h_0) en los períodos pares y (v_1, h_1) en los impares.

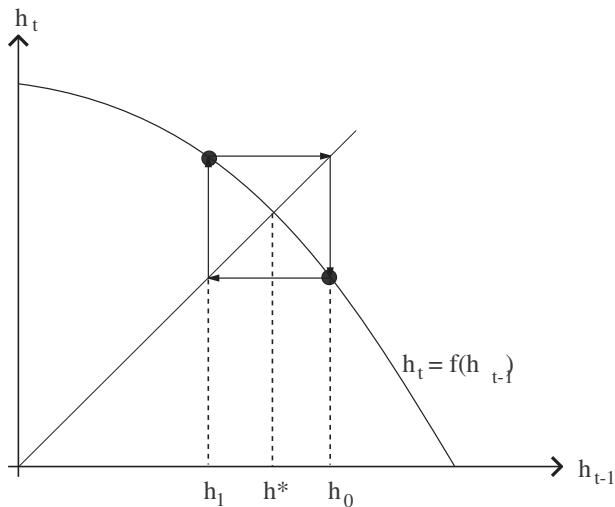


Figura 14.1: Comportamiento cíclico de Lorenzo.

§14.7 Programación dinámica estocástica

§14.7.1 Introducción

En el capítulo anterior vimos cómo se puede utilizar la programación dinámica para resolver problemas de optimización dinámica en tiempo discreto. Como ejemplo se vio el equivalente discreto al modelo de Ramsey cuya solución y análisis nos dio la misma información que en el caso continuo. Podríamos preguntarnos cuál es el sentido de aprender esta técnica nueva, no trivial, si nos sirve para resolver el mismo tipo de problemas, simplemente trasladados a un entorno discreto.

Al igual que en el caso de ecuaciones en diferencias, la introducción de ciertos tipos de incertidumbre es relativamente sencilla en el caso discreto. Si tomamos al tiempo como una variable continua, necesitamos técnicas de cálculo estocástico para poder incorporar incertidumbre en nuestros modelos; sin embargo, en tiempo discreto el problema se simplifica enormemente. Ésta es la razón fundamental del uso de programación dinámica en economía; si el lector hojea cualquier libro de macroeconomía avanzada, notará que la mayoría de los problemas que utilizan esta técnica incluyen incertidumbre.

§14.7.2 Planteamiento del problema

El problema que vamos a considerar es el siguiente:

$$\max E_0 \left(\sum_{t=0}^{\infty} \beta^t f(x_t, u_t) \right)$$

sujeto a

$$x_{t+1} = g(x_t, u_t, \varepsilon_t), \quad x_0 \text{ dado},$$

en donde $\{\varepsilon_t\}$ es una sucesión de variables aleatorias independientes, idénticamente distribuidas (IID) y, como se vio en la segunda parte del libro, $E_t(\cdot)$ denota la esperanza matemática dada la información en el periodo t . Es decir, si \mathcal{I}_t es el conjunto de información disponible al **principio del periodo t** , entonces

$$E_t(X) = E(X \mid \mathcal{I}_t).$$

La incertidumbre entra en el problema anterior mediante la ecuación estocástica en diferencias, $x_{t+1} = g(x_t, u_t, \varepsilon_t)$. El lector observador habrá notado la inusual especificación de que \mathcal{I}_t se refiere a la información disponible al principio del periodo. Debe tenerse cuidado cuando suceden varios eventos en un mismo periodo ya que, si y_t es una variable que se realiza después del inicio del periodo, entonces $E_t(y_t) \neq y_t$. Este tipo de problemas nunca surgen cuando el tiempo es una variable continua ya que en ese caso no tiene sentido hablar de la duración o longitud de un periodo de tiempo. Se sugiere que, en problemas donde el tiempo es discreto, se tenga clara la sucesión de los eventos y el conjunto de información disponible.

La versión estocástica de la ecuación de Bellman dada en (14.3) es simplemente

$$V(x_t) = \max\{f(x_t, u_t) + \beta E_t[V(x_{t+1})]\},$$

con lo cual las condiciones de primer orden quedan dadas por

$$\begin{aligned} \frac{\partial f}{\partial u_t} + \beta E_t[V'(x_{t+1})] \frac{\partial g}{\partial u_t} &= 0, \\ V'(x_t) &= \frac{\partial f}{\partial x_t} + \beta E_t[V'(x_{t+1})] \frac{\partial g}{\partial x_t}, \\ x_{t+1} &= g(x_t, u_t, \varepsilon_t). \end{aligned}$$

§14.8 Algunas aplicaciones

§14.8.1 El consumo como martingala

Consideremos el siguiente problema:

$$\max E_0 \left(\sum_{t=0}^{\infty} \beta^t u(c_t) \right)$$

sujeto a

$$\begin{aligned} a_{t+1} &= R_t(a_t + y_t - c_t), \\ \{R_t\}, \{y_t\}, a_0 &\text{ dados.} \end{aligned}$$

Aquí u es una función de utilidad con las propiedades usuales, a_{t+1} es un activo con rendimiento R_t y y_t es una fuente de ingreso exógena que se materializa al inicio de cada periodo. La incertidumbre proviene de R_t y de y_t . Una observación importante es que, a pesar de que $\{R_t\}$ y $\{y_t\}$ están dados de forma exógena, esto no implica que se conozcan todos sus valores en cada periodo. Concretamente, en el periodo t , el conjunto de información \mathcal{I}_t únicamente contiene los valores pasados de estas variables y el valor presente y_t . La razón de que $R_t \notin \mathcal{I}_t$ es simple: \mathcal{I}_t siempre se refiere al conjunto de información disponible al principio del periodo t y los rendimientos se producen al final del mismo periodo. El consumo c_t es la variable de control y el activo a_t la de estado.

La ecuación de Bellman (en tiempo corriente) está dada por

$$\begin{aligned} V(a_t) &= \max\{u(c_t) + \beta E_t(V(a_{t+1}))\}, \\ a_{t+1} &= R_t(a_t + y_t - c_t), \{R_t\}, \{y_t\}, a_t \text{ dados,} \end{aligned}$$

de aquí que las condiciones de primer orden sean

$$u'(c_t) - \beta E_t(V'(a_{t+1})R_t) = 0, \quad (14.32)$$

$$V'(a_t) = \beta E_t(V'(a_{t+1})R_t). \quad (14.33)$$

Combinando (14.32) y (14.33) mediante el procedimiento usual, obtenemos la ecuación de Euler

$$u'(c_t) = \beta E_t(u'(c_{t+1})R_t).$$

Si los rendimientos no son estocásticos, entonces podemos expresar esta última ecuación como

$$u'(c_t) = \beta R_t E_t(u'(c_{t+1})). \quad (14.34)$$

Si expresamos $\beta = \frac{1}{1+\rho}$ y $R_t = 1 + r_t$, en donde ρ es la tasa subjetiva de descuento y r_t la tasa real del activo a_t , entonces tenemos

$$\frac{u'(c_t)}{E_t(u'(c_{t+1}))} = \frac{1 + r_t}{1 + \rho}, \quad (14.35)$$

que es simplemente la versión estocástica de (14.20). Este tipo de ecuaciones no lineales en diferencias es, en general sumamente complicado de resolver en forma explícita.

Para conocer el tipo de proceso estocástico que siguen los controles, en este caso el consumo, considérese una función u tal que la ecuación sea lineal. Por ejemplo, supongamos que la función de utilidad está dada por

$$u(c_t) = c_t - \frac{\alpha}{2}c_t^2.$$

Entonces, la ecuación (14.34) puede reescribirse como

$$\begin{aligned} 1 - \alpha c_t &= \beta R_t E_t(1 - \alpha c_{t+1}) \\ &= \beta R_t - \alpha \beta R_t E_t(c_{t+1}). \end{aligned} \quad (14.36)$$

Si quisieramos saber qué pasa con el consumo presente c_t en términos de los consumos esperados futuros $E_t(c_{t+n})$, simplemente notamos que se trata de una ecuación estocástica en diferencias de la forma (9.7) vista en el capítulo 9, cuya solución es

$$c_t = \lim_{n \rightarrow \infty} (\beta R)^n E_t(c_{t+n}) + \frac{1 - \beta R}{\alpha} \sum_{n=0}^{\infty} (\beta R)^n,$$

en donde, a manera de simplificación, se supone que $R_t = R$ es una constante. La convergencia se da únicamente si $\beta R < 1$, es decir si $r < \rho$, en cuyo caso $c_t = \frac{1}{\alpha}$.

En el caso del consumo, lo que interesa es tener una idea de su evolución en el tiempo y determinar qué tipo de proceso estocástico sigue. Para este propósito se procede de la siguiente forma: sea ε_t el proceso definido por $\varepsilon_{t+1} = c_{t+1} - E_t(c_{t+1})$; nótese que éste es un proceso de ruido, definido en el capítulo 9, que satisface $E_t(\varepsilon_{t+1}) = 0$. Tenemos que $E_t(c_{t+1}) = c_{t+1} - \varepsilon_{t+1}$ y sustituimos en (14.36) para obtener

$$1 - \alpha c_t = \beta R_t - \alpha \beta R_t (c_{t+1} - \varepsilon_{t+1}),$$

o bien, despejando c_{t+1} ,

$$c_{t+1} = \frac{1}{\beta R_t} c_t + \frac{\beta R_t - 1}{\alpha \beta R_t} + \varepsilon_{t+1}. \quad (14.37)$$

Notemos que dado que la esperanza no condicionada E satisface $EE_t = E$ (véase la sección sobre esperanzas iteradas en el capítulo 9) y que $E_t(\varepsilon_{t+1}) = E(\varepsilon_{t+1}) = 0$, se cumple

$$\begin{aligned} cov(c_t, \varepsilon_{t+1}) &= E[(c_t - E(c_t))(\varepsilon_{t+1} - E(\varepsilon_{t+1}))] \\ &= E\{E_t[(c_t - E(c_t))(\varepsilon_{t+1} - E(\varepsilon_{t+1}))]\} \\ &= E[(c_t - E(c_t))E_t(\varepsilon_{t+1} - E(\varepsilon_{t+1}))] \\ &= E(c_t - E(c_t)) \\ &= 0. \end{aligned}$$

Se concluye que el proceso dado por 14.37 es AR1.

Para el caso en que $R_t = R$, una constante igual a $1 + \rho$ (de manera que la tasa subjetiva ρ es igual a la tasa real r), tenemos que $\beta R_t = \beta R = 1$ y el consumo es una caminata aleatoria dada por

$$c_{t+1} = c_t + \varepsilon_{t+1};$$

en particular, $E_t(c_{t+1}) = c_t$, de manera que se trata de una martingala. Este resultado fue por primera vez resaltado por Hall [Hal78].

Finalmente, notemos que a_t puede obtenerse de la restricción presupuestal iterando hacia el futuro de manera que

$$a_t = \frac{a_{t+n}}{R_t} + \sum_{k=0}^{n-1} \frac{1}{R_t^k} c_{t+k} - \sum_{k=0}^{n-1} \frac{1}{R_t^k} y_{t+k}$$

y tomando el límite cuando $n \rightarrow \infty$ se tiene

$$a_t = \lim_{n \rightarrow \infty} \frac{a_{t+n}}{R_t^n} + \sum_{k=0}^{\infty} \frac{1}{R_t^k} c_{t+k} - \sum_{k=0}^{\infty} \frac{1}{R_t^k} y_{t+k}. \quad (14.38)$$

El término de burbuja $\lim_{n \rightarrow \infty} \frac{a_{t+n}}{R_t^n}$ es cero si, por ejemplo, se supone que no hay juegos de Ponzi en su versión discreta (véase la sección 12.8.2). Esta expresión se simplifica en el caso de rendimientos e ingreso constantes, digamos $R_t = R$ y $y_t = y$, ya que entonces, despejando para el valor presente del consumo,⁸ se tiene

$$\sum_{k=0}^{\infty} \frac{1}{R^k} c_{t+k} = a_t + \frac{yR}{R-1}, \quad (14.39)$$

o bien, tomando esperanzas E_t ,

$$E_t \left(\sum_{k=0}^{\infty} \frac{1}{R^k} c_{t+k} \right) = a_t + \frac{yR}{R-1} = a_t + y \left(\frac{1+r}{r} \right).$$

Tomando $t = 0$, esta última expresión simplemente dice que el valor presente esperado de todo el consumo futuro es igual al valor presente esperado de la riqueza total.

§14.8.2 Consumo de bienes duraderos

Consideremos el modelo de la sección 14.8.1 con la siguiente variante: adicionalmente se consumen los servicios de un bien duradero, que denotamos por D_t . La utilidad que proporciona este bien es proporcional al acervo del bien; sin pérdida de generalidad supongamos que la constante de proporcionalidad es la unidad. Asimismo, D_t se deprecia a una tasa constante δ y los rendimientos del activo son constantes e iguales a $R = 1 + r$. La inversión en bienes duraderos se denota por χ_t . El problema del hogar representativo es

$$\max E_0 \left(\sum_{t=0}^{\infty} \beta^t u(c_t, D_t) \right)$$

sujeto a

$$\begin{aligned} a_{t+1} &= R(a_t + y_t - c_t - \chi_t), \\ D_t &= (1 - \delta)D_{t-1} + \chi_t, \\ \{y_t\}, D_0, a_0 &\text{ dados.} \end{aligned}$$

⁸Esta expresión es la versión discreta de la expresión 12.66.

Los controles son c_t y χ_t , y los estados a_t y D_t . Deseamos saber qué tipo de proceso siguen las variables c_t , D_t y χ_t . La ecuación de Bellman (en tiempo corriente) es

$$\begin{aligned} V(a_t, D_{t-1}) &= \max\{u(c_t, D_t) + \beta E_t(V(a_{t+1}, D_t))\} \\ \text{con } D_t &= (1 - \delta)D_{t-1} + \chi_t, \\ a_{t+1} &= R(a_t + y_t - c_t - \chi_t), \\ \{y_t\}, D_t \text{ y } a_t &\text{ dados.} \end{aligned}$$

Las condiciones de primer orden son

$$u_{c_t} + \beta E_t(-RV_{a_{t+1}}) = 0, \quad (14.40)$$

$$u_{D_t} + \beta E_t(-RV_{a_{t+1}} + V_{D_t}) = 0, \quad (14.41)$$

$$V_{a_t} = \beta E_t(RV_{a_{t+1}}), \quad (14.42)$$

$$V_{D_{t-1}} = (1 - \delta)u_{D_t} + \beta E_t((1 - \delta)V_{D_t}). \quad (14.43)$$

Combinando (14.40) y (14.42) obtenemos la relación usual

$$u_{c_t} = \beta R E_t(u_{c_{t+1}}). \quad (14.44)$$

Combinando (14.40) y (14.41) tenemos que

$$\beta E_t(V_{D_t}) = u_{c_t} - u_{D_t}. \quad (14.45)$$

Sustituyendo esta última relación en (14.43) se tiene

$$\begin{aligned} V_{D_{t-1}} &= (1 - \delta)u_{D_t} + (1 - \delta)(u_{c_t} - u_{D_t}) \\ &= (1 - \delta)u_{c_t} \end{aligned}$$

e iteramos un periodo para llegar a

$$V_{D_t} = (1 - \delta)u_{c_{t+1}}. \quad (14.46)$$

Sustituyendo (14.46) en (14.45) tenemos

$$u_{c_t} - u_{D_t} = \beta E_t((1 - \delta)u_{c_{t+1}}), \quad (14.47)$$

que combinamos con (14.44) para obtener

$$\frac{u_{c_t}}{u_{D_t}} = \frac{R}{R - (1 - \delta)} = \frac{1 + r}{r + \delta}. \quad (14.48)$$

Observemos que el cociente del lado derecho es simplemente el de los costos de oportunidad de consumir c_t y D_t , respectivamente. Como siempre, este cociente es igual a la tasa marginal de sustitución $\frac{u_{c_t}}{u_{D_t}}$.

Para obtener la relación equivalente a (14.44) para el consumo de bienes duraderos se procede como sigue: iteramos un periodo la relación (14.48) y tomamos la esperanza E_t obteniendo

$$E_t(u_{c_{t+1}}) = \frac{1 + r}{r + \delta} E_t(u_{D_{t+1}}),$$

eliminamos $E_t(u_{c_{t+1}})$ mediante (14.44) llegando a

$$\frac{u_{c_t}}{\beta R} = \frac{1+r}{r+\delta} E_t(u_{D_{t+1}}),$$

y finalmente eliminamos u_{c_t} mediante (14.48) obteniendo

$$u_{D_t} = \beta R E_t(u_{D_{t+1}}). \quad (14.49)$$

Las ecuaciones (14.44), (14.48) y (14.49) nos proporcionan la relación entre las utilidades marginales en un periodo dado.

Supongamos ahora que se tiene la función de utilidad dada por

$$u(c_t, D_t) = \left(c_t - \frac{\alpha c_t^2}{2} \right) + \left(D_t - \frac{\alpha D_t^2}{2} \right).$$

Podemos reescribir (14.44) como

$$\begin{aligned} 1 - \alpha c_t &= \beta R E_t(1 - \alpha c_{t+1}) \\ &= \beta R - \alpha \beta R E_t(c_{t+1}), \end{aligned} \quad (14.50)$$

de manera que si definimos el proceso $\varepsilon_{t+1} = c_{t+1} - E_t(c_{t+1})$ (14.50) se reescribe como

$$c_{t+1} = \frac{1}{\beta R} c_t + \frac{\beta R - 1}{\alpha \beta R} + \varepsilon_{t+1}. \quad (14.51)$$

Al igual que en la sección anterior se puede ver que $cov(c_t, \varepsilon_{t+1}) = 0$; de esta forma, la ecuación (14.51) nos dice que el consumo sigue un proceso AR1. De manera análoga, definimos $\gamma_{t+1} = D_{t+1} - E_t(D_{t+1})$ y utilizamos la relación (14.49), junto con la función de utilidad dada, para obtener

$$D_{t+1} = \frac{1}{\beta R} D_t + \frac{\beta R - 1}{\alpha \beta R} + \gamma_{t+1}, \quad (14.52)$$

en donde $E_t(\gamma_{t+1}) = 0$ y $cov(D_t, \gamma_{t+1}) = 0$. De aquí que el acervo de bienes duraderos también siga un proceso AR1.

Para determinar el tipo de proceso que sigue la inversión en bienes duraderos primero, resolvemos, iterando hacia atrás, la ecuación en diferencias dada por la restricción $D_t = (1 - \delta)D_{t-1} + \chi_t$. Así se obtiene

$$D_t = (1 - \delta)^t D_0 + \sum_{k=1}^t (1 - \delta)^{t-k} \chi_k. \quad (14.53)$$

Después, iteramos la restricción un periodo y despejamos χ_{t+1} para llegar a

$$\chi_{t+1} = D_{t+1} - (1 - \delta)D_t. \quad (14.54)$$

Sustituyendo (14.52) en (14.54) se tiene que

$$\chi_{t+1} = \left(\frac{1}{\beta R} - (1 - \delta) \right) D_t + \left(\frac{\beta R - 1}{\alpha \beta R} \right) + \gamma_{t+1}. \quad (14.55)$$

La ecuación (14.55) nos indica el tipo de proceso que sigue la inversión: puesto que D_t es una función $D_t(D_0, \chi_1, \dots, \chi_t)$ y se cumplen $E_t(\gamma_t) = 0$ y $\text{cov}(\chi_k, \gamma_t) = 0$ para $k = 1, \dots, t$, χ_{t+1} sigue un proceso autorregresivo ARt. En particular, los resultados anteriores se simplifican si la tasa subjetiva es igual a la tasa real, de manera que $\beta R = 1$; si adicionalmente suponemos que los bienes duraderos no se deprecian, es decir, $\delta = 0$, entonces

$$\begin{aligned} c_{t+1} &= c_t + \varepsilon_{t+1}, \\ D_{t+1} &= D_t + \gamma_{t+1}, \\ \chi_{t+1} &= \gamma_{t+1}. \end{aligned}$$

De esta forma, el consumo de ambos bienes es simplemente una caminata aleatoria y la inversión en bienes duraderos es un proceso de ruido.

§14.8.3 Ciclos económicos

El comportamiento de las variables económicas agregadas puede separarse en dos partes: la primera es una tendencia determinista y la segunda está formada por fluctuaciones alrededor de dicha tendencia. La teoría keynesiana se desarrolló con el fin principal de reducir, y hasta eliminar, estas fluctuaciones con base en políticas adecuadas. Esta forma de pensar supone que el gobierno (o la autoridad monetaria) tiene la capacidad de anticipar estas fluctuaciones. La teoría de los ciclos económicos se desarrolló en los años setenta del siglo pasado, e introdujo shocks (normalmente IID) que afectan la producción y otras variables mediante algún mecanismo de propagación.⁹ A continuación presentamos un modelo sencillo, en el cual los shocks estocásticos aparecen como shocks “tecnológicos” en la función de producción.

Considérese el siguiente problema del hogar representativo:

$$\max E_0 \left(\sum_{t=0}^{\infty} \beta^t \ln c_t \right)$$

sujeto a

$$\begin{aligned} z_t k_t^\alpha &= k_{t+1} + c_t, \\ \{z_t\} &\text{ y } k_0 \text{ dados.} \end{aligned} \tag{14.56}$$

Aquí, $\ln c_t$ es la función de utilidad, k_t^α la función de producción y z_t es un proceso estocástico IID con $E_t(z_{t+1}) = 0$, que se realiza al principio de cada periodo, de manera que $E_t(z_t) = z_t$. El consumo es la variable de control. Dado que la función de utilidad es del tipo HARA, podemos utilizar el método de coeficientes indeterminados para obtener la función valor (en tiempo corriente) de forma explícita.

Propongamos la siguiente forma de la función valor (logarítmica, de la misma forma que la función de utilidad tipo HARA):

$$V(k_t) = a + b \ln k_t + c \ln z_t.$$

Iterando un periodo, sustituyendo $k_{t+1} = z_t k_t^\alpha - c_t$ y tomando la esperanza E_t ,

$$E_t(V(k_{t+1})) = a + b \ln(z_t k_t^\alpha - c_t).$$

⁹Véase, por ejemplo, [Luc77].

La ecuación de Bellman nos da la función valor como

$$V(k_t) = \max\{\ln c_t + \beta E_t(V(k_{t+1}))\}$$

con $k_{t+1} = z_t k_t^\alpha - c_t$, $\{z_t\}$ y k_t dados.

Sustituyendo para $E_t(V(k_{t+1}))$ y maximizando la expresión del lado derecho de la ecuación con respecto a c_t , se obtiene

$$\begin{aligned} \frac{1}{c_t} &= \frac{b\beta}{z_t k_t^\alpha - c_t} \Rightarrow \\ c_t &= \frac{z_t k_t^\alpha}{1 + b\beta}. \end{aligned} \quad (14.57)$$

Sustituyendo la forma propuesta para $V(k_t)$ del lado izquierdo de la ecuación de Bellman y del lado derecho los valores óptimos, se llega a

$$\begin{aligned} a + b \ln k_t + c \ln z_t &= \ln \frac{z_t k_t^\alpha}{1 + b\beta} + \beta \left(a + b \ln(z_t k_t^\alpha - \frac{z_t k_t^\alpha}{1 + b\beta}) \right) \\ &= \alpha(1 + b\beta) \ln k_t + (1 + b\beta) \ln z_t \\ &\quad + [a\beta + b\beta \ln b\beta - (1 + b\beta) \ln(1 + b\beta)]. \end{aligned}$$

Igualando los coeficientes de ambos lados se tienen las ecuaciones

$$\begin{aligned} a &= a\beta + b\beta \ln b\beta - (1 + b\beta) \ln(1 + b\beta), \\ b &= \alpha(1 + b\beta), \\ c &= 1 + b\beta. \end{aligned}$$

Resolviendo, tenemos que $b = \frac{\alpha}{1-\alpha\beta}$ que es el único valor que necesitamos pues podemos sustituir en (14.57) para llegar a

$$c_t = (1 - \alpha\beta)k_t^\alpha z_t.$$

Sustituimos esta expresión para c_t en (14.56), con lo que obtenemos

$$k_{t+1} = \alpha\beta k_t^\alpha z_t.$$

Tomando logaritmos de esta última expresión se tiene

$$\ln k_{t+1} = \ln \alpha + \ln \beta + \alpha \ln k_t + \ln z_t,$$

de manera que $\ln k_t$ sigue un proceso autorregresivo (AR1), que depende del proceso $\ln z_t$; por ejemplo, $\ln z_t$ puede tomarse como un proceso que sigue una distribución normal con $E_t(\ln z_{t+1}) = 0$ o bien puede ser que sea un proceso autorregresivo de la forma $\ln z_{t+1} = \ln z_t + \mu + \varepsilon_{t+1}$, con ε_t IID y $E_t(\varepsilon_{t+1}) = 0$. Con esto logramos fluctuaciones estocásticas en el capital y por lo tanto en la producción y el consumo.

§14.8.4 Rendimientos de activos (CAPM)

Existen numerosos modelos que tienen como finalidad determinar los precios de los activos. Los más simples, como los modelos tradicionales de **CAPM** (capital asset pricing model) de Sharpe [Sha64] y Lintner [Lin65], únicamente toman en cuenta dos periodos. Estos modelos pueden refinarse, como se ve, entre otros, en los trabajos de Merton [Mer73], Lucas [Luc78], Breeden [Bre79] y Mankiw y Shapiro [MS86]. Aquí expondremos una versión llamada *CAPM basada en el consumo*, que aparece en Blanchard y Fischer [BF89].

Supongamos que w_t representa la riqueza del hogar representativo en términos reales. En cada periodo este hogar consume cierta cantidad c_t y el resto lo invierte en $n + 1$ activos con rendimientos R_t^0, \dots, R_t^n . Los rendimientos R_t^0 son deterministas y por lo tanto conocidos en t , mientras que los rendimientos R_t^1, \dots, R_t^n son estocásticos y sólo sus valores pasados son conocidos en t , es decir, el activo 0 es un activo sin riesgo y los activos 1, ..., n son activos con riesgo. Sean θ_t^j , $j = 0, \dots, n$, las proporciones que se invierten en cada activo, de tal suerte que $\sum_{j=0}^n \theta_t^j = 1$. El resto de la notación es como antes. El problema de optimización es escoger $\{c_t\}$ y $\{\theta_t^j\}$ para toda $j = 0, \dots, n$, de manera que se resuelva,

$$\max E_0 \left(\sum_{t=0}^{\infty} \beta^t u(c_t) \right)$$

sujeto a

$$w_{t+1} = (w_t - c_t)(\theta_t^0 R_t^0 + \sum_{j=1}^n \theta_t^j R_t^j),$$

$$\sum_{j=0}^n \theta_t^j = 1 \text{ y } w_0 \text{ dado.}$$

La ecuación de Bellman asociada (en tiempo corriente) es

$$V(w_t) = \max\{u(c_t) + \beta E_t(V(w_{t+1}))\},$$

con las restricciones correspondientes. Así, las condiciones de primer orden (tomando en cuenta que $\theta_t^0 = 1 - \sum_{j=1}^n \theta_t^j$) son

$$u'(c_t) - \beta E_t[V'(w_{t+1})(\theta_t^0 R_t^0 + \sum_{j=1}^n \theta_t^j R_t^j)] = 0, \quad (14.58)$$

$$\beta E_t[V'(w_{t+1})(-R_t^0 + R_t^j)(w_t - c_t)] = 0, \quad j = 1, \dots, n, \quad (14.59)$$

$$\beta E_t[V'(w_{t+1})(\theta_t^0 R_t^0 + \sum_{j=1}^n \theta_t^j R_t^j)] = V'(w_t). \quad (14.60)$$

De (14.58) y (14.60) tenemos $u'(c_t) = V'(w_t)$ y, por lo tanto,

$$u'(c_t) = \beta E_t[u'(c_{t+1})(\theta_t^0 R_t^0 + \sum_{j=1}^n \theta_t^j R_t^j)]. \quad (14.61)$$

Utilizando que w_t , c_t y R_t^0 son conocidos en el periodo t , es decir, son parte del conjunto de información \mathcal{I}_t , reescribimos (14.59) como

$$E_t[u'(c_{t+1})R_t^j] = R_t^0 E_t[u'(c_{t+1})], \quad j = 1, \dots, n. \quad (14.62)$$

Finalmente, (14.61), (14.62) y el hecho de que $\sum_{j=0}^n \theta_t^j = 1$ implican las siguientes igualdades:

$$u'(c_t) = \beta R_t^0 E_t[u'(c_{t+1})], \quad (14.63)$$

$$u'(c_t) = \beta E_t[u'(c_{t+1})R_t^j], \quad j = 1, \dots, n. \quad (14.64)$$

Notemos que estas ecuaciones son versiones de (14.34) para cada uno de los activos. El optimizador escoge su trayectoria de consumo de manera que la utilidad marginal del consumo en t es igual al valor presente de la utilidad marginal esperada del consumo en $t + 1$, independientemente del activo en el cual se está invirtiendo. Las ecuaciones (14.63) y (14.64) imponen restricciones sobre los procesos que siguen el consumo y los rendimientos. Igualando (14.63) y (14.64) tenemos

$$E_t[u'(c_{t+1})(R_t^j - R_t^0)] = 0, \quad j = 1, \dots, n. \quad (14.65)$$

Tomando la esperanza no condicionada de esta ecuación y recordando que para cualquier pareja x y y de variables aleatorias se cumple $E(xy) = E(x)E(y) + cov(x, y)$, reescribimos (14.65) como

$$E(u'(c_{t+1}))E(R_t^j - R_t^0) + cov(u'(c_{t+1}), R_t^j) = 0, \quad j = 1, \dots, n$$

y despejando para $E(R_t^j)$ se tiene

$$E(R_t^j) = R_t^0 - \frac{cov(u'(c_{t+1}), R_t^j)}{E(u'(c_{t+1}))}, \quad j = 1, \dots, n. \quad (14.66)$$

Esta última ecuación nos dice que, entre más alta sea la covarianza entre los rendimientos del activo y la utilidad marginal del consumo, menores serán los rendimientos esperados del activo. Si suponemos, como es usual, que $u'' < 0$, entonces esto quiere decir que los individuos están dispuestos a aceptar rendimientos menores de un activo que ofrece cobertura cuando el consumo es bajo (por ejemplo, durante una recesión).

Supongamos ahora que uno de los activos con riesgo, digamos el activo k , representa al “mercado” (se puede pensar como un conglomerado de activos); lo que se busca con esto es que ese activo esté perfectamente correlacionado con el consumo, o bien negativamente correlacionado con la utilidad marginal del mismo. Así, $u'(c_{t+1}) = -\gamma R_t^k$ y (14.66) se puede reescribir para el activo k como

$$E(R_t^k) = R_t^0 + \frac{\gamma var(R_t^k)}{E(u'(c_{t+1}))},$$

de manera que

$$\frac{1}{E(u'(c_{t+1}))} = \frac{E(R_t^k) - R_t^0}{\gamma var(R_t^k)}. \quad (14.67)$$

Adicionalmente, dado cualquier otro activo j , distinto de los activos 0 y k , se tiene $\text{cov}(u'(c_{t+1}), R_t^j) = -\gamma \text{cov}(R_t^j, R_t^k)$, por lo que utilizando (14.67), (14.66) se reescribe como

$$E(R_t^j) - R_t^0 = \left(\frac{\text{cov}(R_t^j, R_t^k)}{\text{var}(R_t^k)} \right) (E(R_t^k) - R_t^0),$$

o bien, definiendo $\beta_j = \frac{\text{cov}(R_t^j, R_t^k)}{\text{var}(R_t^k)}$, se tiene

$$E(R_t^j) - R_t^0 = \beta_j (E(R_t^k) - R_t^0). \quad (14.68)$$

A esta última ecuación se la conoce como “línea del mercado de valores” (**security market line**) en finanzas.

¿Cómo se interpreta la expresión (14.68)? Directamente vemos que el valor esperado del rendimiento de un activo, por arriba del que daría un activo sin riesgo (lado izquierdo de (14.68)), es proporcional a la diferencia entre el valor esperado del mercado y el activo sin riesgo. Los coeficientes β_j son los que aparecen al realizar la regresión lineal de los R_t^j en R_t^k . La prima de riesgo para el activo j puede pensarse como la diferencia $E(R_t^j) - R_t^0$ y vemos que ésta disminuye si el activo tiene baja correlación con el mercado, pues en ese caso el coeficiente β_j sería pequeño; la razón es que el activo nos ofrecería cobertura contra posibles bajas en el mercado.

En el ejercicio 14.13 se ve cómo determinar el precio de un activo. Una observación acerca de estos modelos de hogares representativos es que los precios obtenidos son, de cierta forma, “virtuales”, pues si todos los agentes son idénticos no puede haber compradores y vendedores simultáneamente; entonces, todos quieren vender (forzando los precios a la baja) o comprar (forzando los precios a la alta) o, en el caso de equilibrio, los agentes están conformes con la cantidad de activos (o cualquier otro ente que pueda tener un precio) en su poder.

Ejercicios

▷ **14.1** Considerar el problema de maximización especificado por (14.1).

- a) Utilizar la función lagrangiana (14.7) para obtener condiciones necesarias sobre las variables para la existencia de un máximo.
- b) Definir la función valor e interpretar los multiplicadores utilizando el teorema 10.2.6. Comparar los resultados con los obtenidos en el teorema 14.2.1.

▷ **14.2** Reproducir el ejemplo 14.2.2 con alguna hoja de cálculo para los siguientes casos:

- a) $p = 3$ y $T = 9$.
- b) $p = 3$, $T = 9$ y un factor de descuento temporal $\beta = 0.9$.
- c) $p = 3$, $T = 9$ y un factor de descuento temporal $\beta = 0.5$.

▷ **14.3** Considerar el siguiente problema de maximización de una empresa

$$\max \sum_{t=0}^{T-1} \left(\frac{1}{1+r} \right)^t \Pi_t(k_t, l_t, I_t)$$

sujeto a

$$\begin{aligned} k_{t+1} &= (1 - \delta)k_t + I_t, \\ k(0) &= k_0 \text{ y } k(T) = k_T \text{ dados,} \end{aligned}$$

en donde r es la tasa real de descuento, k_t el capital, l_t la mano de obra, I_t la inversión, δ la tasa de depreciación del capital y $\Pi_t(k_t, l_t, I_t)$ la función de ganancias netas.

- a) Encontrar e interpretar las condiciones de primer orden si

$$\Pi_t(k_t, l_t, I_t) = F(k_t, l_t) - w_t l_t - I_t,$$

en donde w_t es la trayectoria exógena del salario y F es una función de producción con las propiedades usuales.

- b) Resolver el mismo problema si el costo de instalar el capital es $\frac{\gamma I_t^2}{2k_t}$ y por lo tanto la función de ganancias queda dada por

$$\Pi_t(k_t, l_t, I_t) = F(k_t, l_t) - w_t l_t - I_t - \frac{\gamma I_t^2}{2k_t}.$$

En este caso, asumir que el capital no se deprecia, es decir, $\delta = 0$. Comparar con el modelo de inversión en tiempo continuo dado en la sección 11.7.1

▷ **14.4** Sean w_0, \dots, w_T y C números positivos dados. Expresar el siguiente problema como un problema de programación dinámica y resolverlo:

$$\begin{aligned} \max & \sum_{t=0}^T w_t u_t^2 \\ \text{sujeto a } & \sum_{t=0}^T u_t = C, \quad u_t \geq 0 \text{ para toda } t. \end{aligned}$$

▷ **14.5** Sean p_0, \dots, p_T y s_0, \dots, s_T números no negativos tales que $\sum_{k=0}^T p_k = 1$. Expresar el siguiente problema como un problema de programación dinámica y resolverlo:

$$\begin{aligned} \max & \sum_{t=0}^T \frac{p_t s_t}{s_t + u_t} \\ \text{sujeto a } & \sum_{t=0}^T u_t = C, \quad u_t \geq 0 \text{ para toda } t. \end{aligned}$$

▷ **14.6** Resolver el modelo discreto de Ramsey de la sección 14.5 con la restricción presupuestal dada por (14.23).

▷ **14.7** La familia Robinson tiene una cantidad inicial de riqueza w_0 . En cada periodo de tiempo, la familia consume parte de esta riqueza e invierte la restante a una tasa r . La función de utilidad de la

familia es $u(c_t) = \sqrt{c_t}$, de manera que el problema a resolver, suponiendo un factor de descuento igual a β y un horizonte infinito, es

$$\max \sum_{t=0}^{\infty} \beta^t \sqrt{c_t}$$

sujeto a $w_{t+1} = (1+r)(w_t - c_t)$ y w_0 dado.

- a) Encontrar la ecuación de Euler correspondiente.
- b) Encontrar explícitamente la función valor.
- c) Hacer los dos incisos anteriores si $u(c_t) = c^\alpha$, $\alpha \in (0, 1)$.

▷ **14.8** En el ejemplo 14.6.5,

- a) sustituir el valor de F en las ecuaciones (14.24) y (14.25) y posteriormente realizar el análisis cualitativo de la ecuación (14.25),
- b) encontrar el valor óptimo $V(k_0)$ para los siguientes valores de los parámetros: $\alpha = \beta = 0.5$, $A = 1$ y $k_0 = 0.2$.

▷ **14.9** Considerar el ejemplo dado en la sección 14.8.1 con las siguientes variantes: la función de utilidad es $u(c_t) = \ln c_t$, $R_t = R$, $y_t = y$ constantes y no hay incertidumbre. Realizar las siguientes demostraciones:

- a) La condición (14.34) implica que

$$c_t = \frac{1}{(\beta R)^n} c_{t+n}.$$

- b) Utilizando la expresión del inciso anterior, probar que (14.39) se puede reescribir como

$$c_t = (1 - \beta) \left(a_t + \frac{yR}{R-1} \right).$$

Observar que, si definimos la riqueza total en t como $w_t = a_t + \frac{yR}{R-1}$, la ecuación anterior nos dice que el consumo en t es una proporción constante de esta riqueza.

- c) Definir el ingreso permanente como

$$y_t^p = \frac{R-1}{R} a_t + y$$

y probar que y_t^p satisface

$$\sum_{k=0}^{\infty} \frac{y_t^p}{R^k} = a_t + \frac{yR}{R-1} = w_t.$$

Es decir, el ingreso permanente es aquel cuyo valor presente, a lo largo de toda la vida (infinita, en este caso), es igual a la riqueza total.

- d) Combinar los dos incisos anteriores y probar que

$$c_t = \left[1 - \left(\frac{\beta R - 1}{R - 1} \right) \right] y_t^p.$$

A esto se lo conoce como la **hipótesis del ingreso permanente**.¹⁰ El consumo no es simplemente una proporción fija del ingreso ($c_t = (1-s)y_t$) sino una proporción fija del ingreso permanente. Aquí la tasa de ahorro está dada por $s = \frac{\beta R - 1}{R - 1}$.

▷ **14.10** Realizar el análisis cualitativo del sistema dinámico dado por la ecuación (14.31) de la sección 14.6.1.

▷ **14.11** Considerar el siguiente problema del hogar representativo:

$$\max E_0 \left(\sum_{t=0}^{\infty} \beta^t u(c_t) \right)$$

sujeto a

$$a_{t+1} = R(a_t + y_t - c_t), \\ \{y_t\} \text{ y } a_0 \text{ dados,}$$

en donde toda la notación es como en el ejemplo 14.8.1 y los rendimientos R son constantes.

¹⁰Milton Friedman menciona la hipótesis del ingreso permanente en [Fri56].

a) Encontrar las condiciones de primer orden.

b) Si la función de utilidad es de la forma

$$u(t) = (B + \varepsilon_t)c_t - \frac{D}{2}c_t^2,$$

en donde ε_t es un shock aleatorio a las preferencias y B, D son constantes conocidas, describir el tipo de proceso que sigue el consumo si se tiene que $\varepsilon_{t+1} = \alpha\varepsilon_t + \vartheta_{t+1}$ con $E_t(\vartheta_{t+1}) = 0$, $\gamma_{t+1} = c_{t+1} - E_t(c_{t+1})$ y la tasa subjetiva es igual a la tasa real, de manera que $\beta R = 1$.

c) Describir el proceso que sigue el consumo si ε_t es una caminata aleatoria.

▷ 14.12 Considerar el siguiente problema del hogar representativo:

$$\max E_0 \left(\sum_{t=0}^{T-1} \frac{e^{-\alpha c_t}}{-\alpha} \right)$$

sujeto a

$$a_{t+1} = a_t + y_t - c_t,$$

$$y_t = y_{t-1} + \varepsilon_t,$$

a_0 dado,

en donde la función de utilidad es de aversión absoluta al riesgo constante $u(c_t) = \frac{e^{-\alpha c_t}}{-\alpha}$. Toda la notación es como antes y ε_t es un proceso IID, distribuido como una normal con media 0 y varianza σ^2 .

a) Obtener las condiciones de primer orden y utilizar el hecho de que, dada x distribuida como una normal con varianza σ^2 se tiene que $E(e^x) = e^{E(x) + \frac{\sigma^2}{2}}$ para probar que

$$c_{t+1} = c_t + \frac{\alpha\sigma^2}{2} + \varepsilon_t.$$

b) Probar que

$$c_t = \left(\frac{1}{T-t} \right) a_t + y_t - \frac{\alpha(T-t-1)\sigma^2}{4}.$$

▷ 14.13 (Basado en el modelo de activos de Lucas citado en la sección 14.8.4) Considerar el siguiente problema del hogar representativo:

$$\max E_0 \left(\sum_{t=0}^{\infty} \beta^t u(c_t) \right)$$

sujeto a

$$y_t + (d_t + p_t)s_t = c_t + p_t s_{t+1},$$

$\{y_t\}, \{d_t\}, \{p_t\}$ y s_0 dados.

Aquí, s_t denota el número de acciones de un activo al principio del periodo t , d_t son los dividendos del activo generados al principio del periodo, p_t es el precio del activo (en términos reales) una vez producidos los dividendos al principio del periodo. Las demás variables son iguales que en el problema anterior. Como es costumbre, la restricción es simplemente una restricción presupuestal en donde $y_t + (d_t + p_t)s_t$ son los ingresos totales del periodo y $c_t + p_t s_{t+1}$ los egresos totales.

a) Obtener las condiciones de primer orden.

b) Por definición de los rendimientos R_t , en el periodo t éstos quedan dados por el cociente

$$R_t = \frac{d_{t+1} + p_{t+1}}{p_t}.$$

Probar que se obtiene entonces la condición (14.35).

c) Suponer que la función de utilidad está dada por $u(c_t) = c_t$. Probar entonces que

$$p_{t+1} = \beta E_t(p_{t+1} + d_{t+1}).$$

Resolver esta ecuación iterando hacia el futuro y dividir la solución en la parte fundamental y la parte de burbuja.

Parte VI

Apéndices

APÉNDICE A

Números complejos

§A.1 Propiedades básicas

Definimos $i = \sqrt{-1}$, de manera que $i^2 = -1$. Recordemos que el campo de los **números complejos** es el conjunto $\mathbb{C} = \{\alpha + \beta i \mid \alpha, \beta \in \mathbb{R}\}$, con la suma y el producto definidos como sigue:

$$\begin{aligned}(\alpha_1 + \beta_1 i) + (\alpha_2 + \beta_2 i) &= (\alpha_1 + \alpha_2) + (\beta_1 + \beta_2)i, \\(\alpha_1 + \beta_1 i)(\alpha_2 + \beta_2 i) &= \alpha_1 \alpha_2 + \alpha_1 \beta_2 i + \alpha_2 \beta_1 i + \beta_1 \beta_2 i^2 \\&= (\alpha_1 \alpha_2 - \beta_1 \beta_2) + (\alpha_1 \beta_2 + \alpha_2 \beta_1)i.\end{aligned}$$

Nótese que $\mathbb{R} \subset \mathbb{C}$ ya que cualquier real es un complejo con $\beta = 0$.

Dado un número complejo $z = \alpha + \beta i$, decimos que α es la **parte real** de z y la denotamos por $\alpha = \operatorname{Re} z$. De forma análoga, β es la **parte imaginaria** de z y se denota por $\beta = \operatorname{Im} z$. Es relativamente sencillo probar que los números complejos forman un campo con la suma y el producto definidos arriba. Los números reales 0 y 1 son, respectivamente, los **elementos identidad** para la suma y el producto y se conocen como **idéntico aditivo** e **idéntico multiplicativo**. Concretamente, se satisfacen las siguientes propiedades para todo $z_1, z_2, z_3 \in \mathbb{C}$:

Propiedades de un campo

- a) $z_1 + z_2, z_1 z_2 \in \mathbb{C}$ (**cerradura**),
- b) $z_1 + z_2 = z_2 + z_1, z_1 z_2 = z_2 z_1$ (**commutatividad**),
- c) $(z_1 + z_2) + z_3 = z_1 + (z_2 + z_3), (z_1 z_2) z_3 = z_1 (z_2 z_3)$ (**asociatividad**),
- d) existen $0, 1 \in \mathbb{C}$, únicos, tal que, $z_1 + 0 = z_1 = 0 + z_1, z_1 1 = z_1 = 1 z_1$ (**idéntico aditivo y multiplicativo**).

- e) $z_1(z_2 + z_3) = z_1z_2 + z_1z_3$ (**distributividad**).
f) Dado $z \in \mathbb{C}$, existe $-z$, único, tal que $z + (-z) = 0$ (**inverso aditivo**).
g) Dado $z \in \mathbb{C}, z \neq 0$, existe z^{-1} , único, tal que $zz^{-1} = 1$ (**inverso multiplicativo**).

§A.2 Representación polar y fórmula de Euler

Definición A.2.1 Dado $z \in \mathbb{C}$, con $z = \alpha + \beta i$, se define la norma de z como $\|z\| = \sqrt{\alpha^2 + \beta^2}$.

Es inmediato ver que $\|z\| \geq 0$ y que $\|z\| = 0$ si y sólo si $z = 0$.

Definición A.2.2 Dado $z \in \mathbb{C}$, con $z = \alpha + \beta i$, se define el conjugado de z como $\bar{z} = \alpha - \beta i$.

Podemos verificar fácilmente que $z\bar{z} = \|z\|^2$.

Los números complejos se pueden representar geométricamente en el plano complejo, con la parte real en el eje horizontal y la parte imaginaria en el eje vertical. Asimismo, dado $z = \alpha + \beta i$, se puede escribir en forma polar como

$$z = \|z\| (\cos \theta + i \sin \theta),$$

en donde $\theta = \arctan \frac{\beta}{\alpha}$. Estas consideraciones se muestran en la figura A.1.

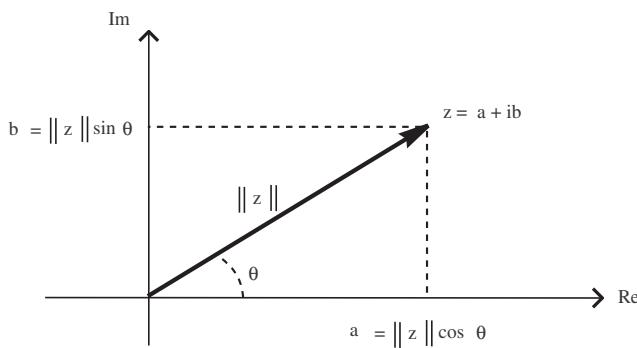


Figura A.1: Representación geométrica y polar para números complejos.

Decimos que un complejo z es unitario si $\|z\| = 1$. Es claro que si z es unitario entonces su representación polar es de la forma,

$$z = \cos \theta + i \sin \theta.$$

La función exponencial $e^z = \exp z$ se puede definir para cualquier $z \in \mathbb{C}$ de manera que, si $z = \alpha + \beta i$, entonces $e^z = e^\alpha e^{i\beta}$. La fórmula de Euler, cuya prueba se deja como ejercicio, dice que para cualquier complejo unitario, $z = \cos \theta + i \sin \theta$, se tiene,

$$\cos \theta + i \sin \theta = e^{i\theta}.$$

Es claro que dado cualquier $z \in \mathbb{C}$, éste se puede expresar como

$$z = \|z\| (\cos \theta + i \sin \theta) = \|z\| e^{i\theta}.$$

◇◇◇◇◇◇◇◇◇

Ejercicios

▷ **A.1** Probar que se cumple la siguiente propiedad para $z \in \mathbb{C}$, $z \neq 0$: existe z^{-1} tal que $zz^{-1} = 1$ (inverso multiplicativo). (Sugerencia: utilizar la igualdad $z\bar{z} = \|z\|^2$.)

▷ **A.2** Sean $z_1 = 2 + i$ y $z_2 = -3 + 5i$. Encuentra los siguientes números:

- a) $z_1 + z_2$.
- b) $z_1 z_2$.
- c) z_1^{-1}, z_2^{-1} .
- d) $\frac{z_1}{z_2}$.
- e) $\|z_1\|, \|z_2\|, \|z_1 + z_2\|$.
- f) \bar{z}_1, \bar{z}_2 .

▷ **A.3** Sean $z = \alpha + \beta i \in \mathbb{C}$ y \bar{z} su conjugado. Demostrar las siguientes aseveraciones si $z_1, z_2 \in \mathbb{C}$, $r \in \mathbb{R}$ y $p(x)$ es un polinomio con coeficientes reales:

- a) $\overline{(z_1 + z_2)} = \bar{z}_1 + \bar{z}_2$.
- b) $\overline{(z_1 z_2)} = \bar{z}_1 \bar{z}_2$.
- c) $\bar{r} = r$.
- d) $p(z) = 0 \Leftrightarrow p(\bar{z}) = 0$ para todo $z \in \mathbb{C}$. Es decir, las raíces complejas siempre aparecen por parejas conjugadas.

▷ **A.4** Probar la fórmula de Euler, $\cos \theta + i \sin \theta = e^{i\theta}$. (Sugerencia: considerar el desarrollo en serie de Taylor, alrededor de $\theta = 0$, para la función $e^{i\theta}$. Posteriormente, separar la serie en la parte real e imaginaria y notar que éstas son las series de Taylor para $\cos \theta$ y $\sin \theta$, respectivamente.)

▷ **A.5** Probar que, dados $z_1 = r_1(\cos \theta_1 + i \sin \theta_1)$ y $z_2 = r_2(\cos \theta_2 + i \sin \theta_2)$, $r_1, r_2 \geq 0$, se tiene que $z_1 z_2 = r_1 r_2 [\cos(\theta_1 + \theta_2) + i \sin(\theta_1 + \theta_2)]$.

▷ **A.6** Utilizar el ejercicio anterior para probar que $\cos(\theta_1 + \theta_2) = \cos \theta_1 \cos \theta_2 - \sin \theta_1 \sin \theta_2$ y $\sin(\theta_1 + \theta_2) = \sin \theta_1 \cos \theta_2 + \cos \theta_1 \sin \theta_2$.

▷ **A.7** Sea $z = r(\cos \theta + i \sin \theta)$, $r \geq 0$. Probar que $z^n = r^n(\cos n\theta + i \sin n\theta)$.

APÉNDICE B

Existencia y unicidad de soluciones

En ocasiones podemos calcular explícitamente la solución de una ecuación diferencial o bien realizar algún tipo de análisis cualitativo; sin embargo, a veces es imposible encontrar una solución a pesar de que, teóricamente, se sepaa que ésta existe y es única. En este apéndice enunciamos el teorema fundamental de existencia y unicidad de soluciones, omitiendo su demostración ya que ésta requiere de conceptos que están más allá del alcance de este libro.

Consideremos el siguiente problema de valores iniciales.

$$\begin{aligned}y' &= f(y), \\y(t_0) &= y_0.\end{aligned}\tag{B.1}$$

En general, se espera que la solución $y(t)$ sea una función de una variable cuya imagen sea un vector de \mathbb{R}^n . Dicho de otra manera, nos gustaría encontrar una función diferenciable $y : I \rightarrow \mathbb{R}^n$, en donde I es un intervalo alrededor de t_0 de la forma $I = (\alpha, \beta)$ y tal que cumpla las dos condiciones dadas por (B.1). La función f es un campo vectorial en el sentido de la definición 4.1 del capítulo 4. Como hipótesis, la función f está definida en un subconjunto $\Omega \subset \mathbb{R}^n$ y se supone que $(t_0, y_0) \in \Omega$.

Recordemos que existen dos tipos de ecuaciones,

$$\left\{ \begin{array}{l} \text{autónomas} \\ \text{no autónomas} \end{array} \right.$$

En las autónomas, f no depende de t tal como en (B.1). En las no autónomas f tiene dependencia explícita del tiempo t y, en principio tienen la forma

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

En este caso, si hacemos $y = \begin{pmatrix} x \\ t \end{pmatrix}$, entonces

$$\dot{y} = \begin{pmatrix} \dot{x} \\ 1 \end{pmatrix} = \begin{pmatrix} f(x, t) \\ 1 \end{pmatrix}$$

Por lo tanto, si

$$y = \begin{pmatrix} x \\ t \end{pmatrix} \text{ y } F(y) = \begin{pmatrix} f(y) \\ 1 \end{pmatrix}$$

entonces $\dot{y} = F(y)$ es autónoma (recordar ejercicio 4.13 del capítulo 4). De este modo, toda ecuación no autónoma se reduce a una autónoma.

Lo siguiente es un breve repaso de Cálculo de varias variables. Si $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ es un campo vectorial, éste se puede expresar en términos de sus funciones coordenadas como,

$$f(\mathbf{x}) = \begin{pmatrix} f_1(\mathbf{x}) \\ \vdots \\ f_m(\mathbf{x}) \end{pmatrix},$$

con $\mathbf{x} \in \mathbb{R}^n$, $f_i : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$, $i = 1, 2, \dots, m$. Recordamos la observación a la definición 11.1.2 del capítulo 11 para extender el concepto de diferenciabilidad a campos vectoriales como sigue.

Definición B.0.3 *Sea Ω un abierto de \mathbb{R}^n . Se dice que $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ es diferenciable en $\mathbf{x}_0 \in \Omega$, si existe una matriz de $m \times n$, $Df(\mathbf{x}_0)$, tal que*

$$\lim_{\mathbf{h} \rightarrow 0} \frac{\|f(\mathbf{x}_0 + \mathbf{h}) - f(\mathbf{x}_0) - Df(\mathbf{x}_0)(\mathbf{h})\|}{\|\mathbf{h}\|} = 0,$$

en donde $\|\cdot\|$ es la norma Euclídea de la definición D.2.1 dada en el apéndice D. La matriz $Df(\mathbf{x}_0)$ es llamada **matriz jacobiana** y se escribe $Df(\mathbf{x}_0) = f'(\mathbf{x}_0)$.

Notemos que $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ es diferenciable en todo Ω , entonces f' es una función definida en Ω con valores en el conjunto de matrices de $m \times n$. Al igual que en el caso de las funciones con imagen en \mathbb{R} ($m = 1$), la matriz $Df(\mathbf{x}_0) = f'(\mathbf{x}_0)$ representa la “mejor” aproximación lineal de la función alrededor de \mathbf{x}_0 .

Podemos hablar de la continuidad de esta derivada f' en Ω . Aunque la definición parezca algo complicada, el siguiente teorema nos permite saber cuando un campo vectorial es diferenciable y calcular su derivada.

Definición B.0.4 *Sea $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ una función dada por*

$$f = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix}.$$

Entonces, se dice que la función f es de **clase C^1** en \mathbf{x}_0 si todas las derivadas parciales de f son continuas en \mathbf{x}_0 ; es decir, si $\left(\frac{\partial f_i}{\partial x_i}\right)_{\mathbf{x}=\mathbf{x}_0}$ es continua para todo $i = 1, \dots, m$ y $j = 1, \dots, n$. En este caso se tiene

$$f'(\mathbf{x}_0) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}_{\mathbf{x}=\mathbf{x}_0}.$$

Si f es de clase C^1 en el abierto Ω escribimos $f \in C^1(\Omega)$. Para el caso en que f toma valores reales, es decir, $m = 1$, decimos que f es de clase C^k en \mathbf{x}_0 si existen las primeras k derivadas parciales y son continuas; análogamente, si f es de **clase C^k** en el abierto Ω escribimos $f \in C^k(\Omega)$.

Ejemplo

Ej B.0.1 Considérese el campo vectorial $f : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ dado por

$$f(x, y, z) = \begin{pmatrix} x^2 - 2xy + z \\ xyz \end{pmatrix}.$$

Dado que se cumplen las hipótesis del teorema anterior se tiene que la matriz jacobiana en cualquier punto es

$$f'(x, y, z) = \begin{pmatrix} 2x - y & -2x & 1 \\ yz & xz & xy \end{pmatrix}.$$

Existen varios teoremas que nos permiten afirmar la existencia de una solución al problema (B.1). Sin embargo, enunciamos uno para el cual el conjunto de hipótesis es relativamente fácil de comprobar.

Teorema B.0.5 (Existencia y unicidad. Versión débil)

Sean Ω un abierto de \mathbb{R}^n , $y_0 \in \Omega$ y $f \in C^1(\Omega)$. Entonces existe un intervalo $(t_0 - a, t_0 + a)$ tal que el problema (B.1) tiene una solución única en $(t_0 - a, t_0 + a)$.

Notemos que lo único que el teorema nos dice es que existe una única solución. En principio, no es posible encontrar una solución explícita de esta solución y el intervalo que es dominio de la función podría ser muy pequeño.

Ejemplos

Ej B.0.2 Si f no es C^1 en todo el plano no podemos garantizar la unicidad. Consideremos el siguiente problema

$$\begin{aligned} \dot{x} &= 3x^{\frac{2}{3}}, \\ x(0) &= x_0. \end{aligned}$$

Se puede ver que $f(x) = 3x^{\frac{2}{3}}$ es de clase C^1 en $\Omega = \mathbb{R} \setminus \{0\}$ y, por tanto, dicha solución es única si $x_0 \neq 0$. Si $x_0 = 0$, se puede verificar directamente que $x(t) = 0$ y $x(t) = t^3$ son dos soluciones posibles. Esto no viola el resultado del teorema B.0.5 pues la función f no es diferenciable en 0, por lo que no se puede garantizar unicidad.

Ej B.0.3 El problema de valores iniciales

$$\begin{aligned} y' &= \tan(e^y - \arctan y) \\ y(t_0) &= y_0 \end{aligned}$$

tiene una solución única para todo t_0 y todo y_0 pues la función f que define a la ecuación es de clase C^1 en un abierto de \mathbb{R}^n . Esto se sabe a pesar de que, en general, la solución no se puede calcular explícitamente.

Ej B.0.4 Considérese la ecuación

$$\dot{x} = \frac{x}{t}.$$

En el plano TX dibujamos las curvas $\frac{x}{t} = K$ para diversos valores de K . Sobre cada una de estas curvas, el valor de K nos indica el valor de \dot{x} sobre dicha curva, es decir, la pendiente que debe tener una solución posible. Dibujamos pequeños segmentos de dirección $\dot{x} = K$ sobre los puntos $\frac{x}{t} = K$ (el conjunto de estos segmentos se conoce como **campo vectorial**, **campo integral** o **campo de dirección** del sistema dinámico). Se pueden obtener así las trayectorias para las posibles soluciones. En este caso podemos aplicar el teorema si $t \neq 0$, puesto que la función, $f(x, t) = \frac{x}{t}$, es de clase C^1 y por lo tanto, dada una condición inicial la solución debe existir y ser única. En la figura B.1 se presentan tres ejemplos, cada uno con una condición inicial diferente: en el caso (a), $t_0 = 2$ y $x(2) = 5$; en (b), $t_0 = 1$ y $x(1) = 1$ y en (c), $t_0 = 1.5$ y $x(1.5) = -1$. Como puede verse dada una condición inicial, existe una única curva solución.

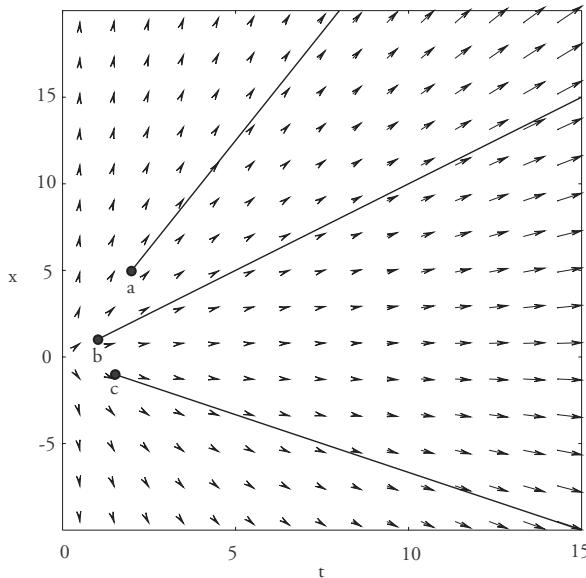


Figura B.1: Campo vectorial y algunas trayectorias solución para el ejemplo B.0.4.

Observamos que si el sistema es autónomo, entonces los lugares geométricos $f(x) = K$ son simplemente rectas horizontales en el plano TX . El siguiente ejemplo ilustra esto.

Ej B.0.5 Deseamos ilustrar el campo de dirección y algunas curvas solución de la ecuación

$$\dot{x} = x(2 - x).$$

Los campos de dirección los dibujamos sobre la familia de curvas $x(2 - x) = K$ que son simplemente rectas horizontales. El resultado se muestra en la figura B.2 en donde se esbozan algunas curvas solución dados los siguientes puntos iniciales en $t_0 = 0$: en (a), $x(0) = 3$; en (b), $x(0) = 2$; en (c), $x(0) = 0.2$; en (d), $x(0) = 0$; en (e), $x(0) = -0.2$.

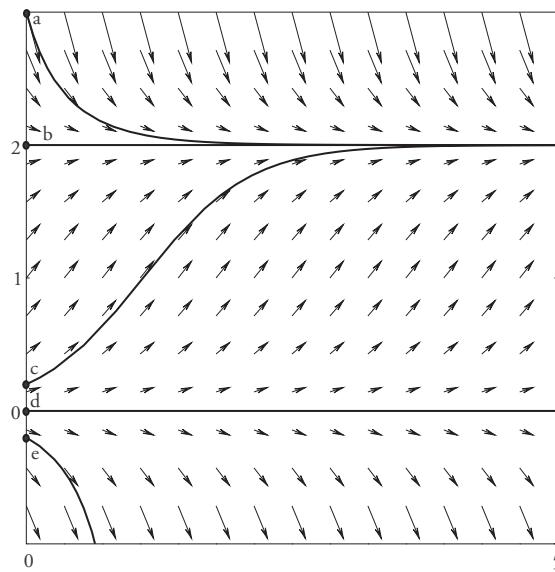


Figura B.2: Campo vectorial y algunas trayectorias solución para el ejemplo B.0.5.

La unicidad de las soluciones es, a veces, más importante de lo que a simple vista podría parecer. En el plano, tenemos la propiedad de que las curvas “separan” a éste en regiones. Una manera geométrica de entender la unicidad es pensar que una curva que es solución no puede ser cruzada por *otra* solución. Si tal fuera el caso, se tendrían dos soluciones distintas que iniciarían en el mismo punto: el punto de intersección. Estas consideraciones se ilustran en la figura B.3. En el plano \mathbb{R}^2 , la unicidad trae consecuencias importantes pues en ocasiones una curva solución divide al plano en dos regiones y así las demás soluciones no pueden cruzar de una estas regiones a otra.

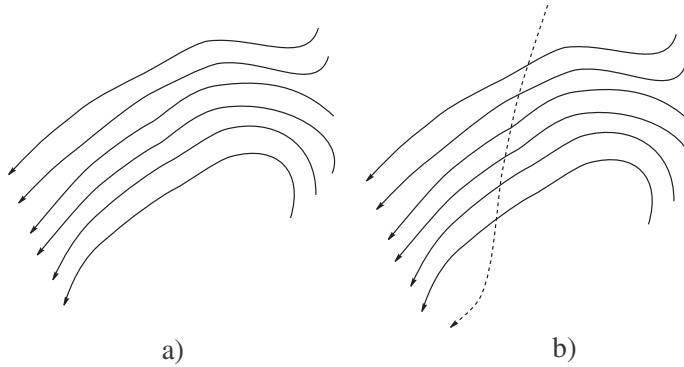


Figura B.3: El teorema de existencia y unicidad no permite que las soluciones se crucen. En la figura *a*) se representa la situación típica en la que se ven las soluciones de una ecuación diferencial en el plano. En la figura *b*) se da una situación imposible si se tiene unicidad: una solución no puede cruzar otras.

Ej B.0.6 Consideremos el siguiente problema

$$\begin{aligned} \frac{d}{dt} \begin{pmatrix} u \\ v \end{pmatrix} &= \begin{pmatrix} u + v \sin u \\ -1 + uv + \cos v \end{pmatrix}, \\ u(0) &= u_0, \\ v(0) &= v_0. \end{aligned}$$

Se puede demostrar que cualquier curva solución que inicie en el primer cuadrante $\mathbb{R}_{++}^2 = \{(u, v) : u > 0, v > 0\}$, permanecerá ahí para todo tiempo en el intervalo de definición de la solución. Es decir, si $(u_0, v_0) \in \mathbb{R}_{++}^2$, y $(u(t), v(t))$ es la solución que inicia en (u_0, v_0) entonces para todo tiempo en el intervalo de definición se tiene que

$$(u(t), v(t)) \in \mathbb{R}_{++}^2.$$

Esto se puede comprobar de la siguiente manera. Sea $(u_0, 0)$ un punto en el eje horizontal. Se puede ver que la función $(u(t), v(t)) = (u_0 e^t, 0)$ es una solución de la ecuación para todo tiempo. Por lo tanto, si $(u(t), v(t))$ es la solución que inicia en el primer cuadrante, no podrá cruzar *nunca* el eje horizontal puesto que éste es la imagen de una solución. Del mismo modo se puede argumentar con respecto al eje vertical. En esencia, esta pareja de soluciones “atrapa” al resto.

Por último quisieramos mencionar que otras versiones del teorema contemplan hipótesis distintas. Por ejemplo, si sólo pedimos continuidad –una condición bastante general– entonces sólo podemos afirmar que la solución existe.

Teorema B.0.6 (Peano)

Si f es continua en Ω , entonces el problema (B.1) posee al menos una solución definida en una vecindad de t_0 .

Para mayor información acerca del teorema de Peano y demás teoremas de existencia y unicidad, consultar algún libro avanzado de Sistemas Dinámicos, Análisis o ecuaciones diferenciales, recomendamos ver, por ejemplo, el libro de Robinson [Rob99].

APÉNDICE C

Raíces repetidas

§C.1 Multiplicidad y defecto de un valor propio

Muchas de las matrices que nos hemos de encontrar tienen buenas propiedades, en el sentido de que los valores propios son distintos y por lo tanto los vectores propios correspondientes son linealmente independientes, sean éstos reales o complejos. Desgraciadamente, no todas las matrices reales de $n \times n$ tienen necesariamente n vectores propios. Existen casos degenerados en los que las raíces están repetidas y no hay un número suficiente de vectores propios linealmente independientes.

Ejemplo

Ej C.1.1 Sea $A = \begin{pmatrix} 7 & 1 \\ 0 & 7 \end{pmatrix}$. Encontrar los vectores propios de A .

El polinomio característico de A es $\lambda^2 - 14\lambda + 49 = (\lambda - 7)^2$. Las raíces de este polinomio son $\lambda_1 = 7$ y $\lambda_2 = 7$, es decir, tenemos un caso de raíces repetidas. Para encontrar un vector propio debemos resolver

$$(A - 7I) \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Esto implica, forzosamente, que $b = 0$. Escogemos $a = 1$ y obtenemos el siguiente vector propio

$$\mathbf{v} = \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

con valor propio 7. Este ejemplo muestra que existen matrices con un solo vector propio.



Definición C.1.1 La **multiplicidad de un valor propio** ω es el número de factores $(\lambda - \omega)$ que aparecen en el polinomio característico. El **defecto de un valor propio** es la diferencia entre la multiplicidad y el máximo número de vectores propios linealmente independientes que corresponden a ω . Si k es la multiplicidad y p es este número de vectores propios, entonces el defecto es $d = k - p$.

Ejemplo

Ej C.1.2 Encontrar la multiplicidad y el defecto de cada una de las siguientes matrices.

$$A = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{pmatrix}, B = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 4 & 1 \\ 0 & 0 & 4 \end{pmatrix}, C = \begin{pmatrix} 4 & 1 & 0 \\ 0 & 4 & 1 \\ 0 & 0 & 4 \end{pmatrix}.$$

En este caso es fácil ver que las tres matrices tienen el mismo polinomio característico, es decir,

$$p_A(\lambda) = p_B(\lambda) = p_C(\lambda) = (\lambda - 4)^3.$$

Por lo tanto 4 tiene multiplicidad $k = 3$, para cada una de las matrices A , B y C . Recordemos que si k es la multiplicidad de λ y λ tiene p vectores propios, el defecto es $d = k - p$.

Para la matriz A los siguientes son vectores propios linealmente independientes:

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

Por lo tanto $p = 3$ y el defecto es $d_A = 3 - 3 = 0$.

Por otro lado, la matriz B tiene sólo 2 vectores propios linealmente independientes:

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\}.$$

En este caso el defecto es $d_B = 3 - 2 = 1$.

Finalmente, C tiene un solo vector propio linealmente independiente: $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$. Concluimos que el defecto es $d_C = 3 - 1 = 2$.



§C.2 Vector propio generalizado

Definición C.2.1 Sea λ un valor propio de una matriz A . Un **vector propio generalizado de rango r** es un vector \mathbf{w} tal que

$$(A - \lambda I)^r \mathbf{w} = 0, \\ (A - \lambda I)^{r-1} \mathbf{w} \neq 0,$$

donde r es un entero $r \geq 1$. Dado un vector propio generalizado \mathbf{w} de rango r , se define la cadena asociada $\{\mathbf{v}_1, \dots, \mathbf{v}_r\}$ como

$$\begin{aligned}\mathbf{v}_r &= \mathbf{w}, \\ \mathbf{v}_{r-1} &= (A - \lambda I)\mathbf{v}_r, \\ &\vdots \\ \mathbf{v}_2 &= (A - \lambda I)\mathbf{v}_3 = (A - \lambda I)^{r-2}\mathbf{w}, \\ \mathbf{v}_1 &= (A - \lambda I)\mathbf{v}_2 = (A - \lambda I)^{r-1}\mathbf{w}.\end{aligned}$$

En este punto, caben algunas observaciones. Por ser \mathbf{w} un vector propio generalizado de rango r , tenemos que $\mathbf{v}_1 = (A - \lambda I)^{r-1}\mathbf{w} \neq 0$. Por otro lado, $(A - \lambda I)\mathbf{v}_1 = (A - \lambda I)^r\mathbf{w} = 0$. Concluimos de este modo que \mathbf{v}_1 es siempre un vector propio. El lector deberá notar que cualquier vector propio es vector propio generalizado de rango 1.

El siguiente resultado es un complemento a la definición anterior. Sirve sobre todo en dimensiones mayores, cuando la multiplicidad puede ser grande. La demostración es demasiado complicada para lo que se pretende en este libro, por lo que será omitida.

Teorema C.2.2

Dado el valor propio λ de multiplicidad k , existen $\mathbf{w}_1, \dots, \mathbf{w}_p$ vectores propios generalizados de rango r_1, \dots, r_p , respectivamente, tales que

- a) p es el número de vectores propios linealmente independientes,
- b) los vectores propios generalizados $\mathbf{w}_1, \dots, \mathbf{w}_p$ son linealmente independientes,
- c) $r_1 + r_2 + \dots + r_p = k$.

Ejemplo

Ej C.2.1 Sea λ un valor propio con multiplicidad 6 y defecto 2. ¿Qué posibles rangos tienen los vectores propios generalizados del teorema C.2.2?

Sabemos que $d = k - p$. Por lo tanto $p = 4$. Tenemos que encontrar r_1, r_2, r_3, r_4 tales que $r_1 + r_2 + r_3 + r_4 = 6$. Observemos que, puesto que $r_i \geq 1, i = 1, \dots, 4$, tenemos dos posibilidades

- a) $r_1 = 3, r_2 = 1, r_3 = 1, r_4 = 1$,
- b) $r_1 = 2, r_2 = 2, r_3 = 1, r_4 = 1$.



§C.3 Construcción de soluciones linealmente independientes

Ahora podemos utilizar vectores propios generalizados para construir un número suficiente de soluciones linealmente independientes del problema lineal.

Proposición C.3.1 Sean λ un valor propio de A y \mathbf{w} un vector propio generalizado de rango r . Sea $\{\mathbf{v}_1, \dots, \mathbf{v}_r\}$ la cadena generada por \mathbf{w} . Entonces las siguientes son r soluciones linealmente independientes del problema lineal

$$\begin{aligned} X_1(t) &= e^{\lambda t} \mathbf{v}_1, \\ X_2(t) &= e^{\lambda t} (\mathbf{v}_2 + t\mathbf{v}_1), \\ X_3(t) &= e^{\lambda t} (\mathbf{v}_3 + t\mathbf{v}_2 + \frac{t^2}{2}\mathbf{v}_1), \\ &\vdots \\ X_r(t) &= e^{\lambda t} (\mathbf{v}_r + t\mathbf{v}_{r-1} + \dots + \frac{t^{r-1}}{(r-1)!} \mathbf{v}_1). \end{aligned}$$

Ejemplo

Ej C.3.1 Resolver

$$\dot{X} = \begin{pmatrix} 1 & 0 & 0 \\ -4 & 1 & 0 \\ 3 & 6 & 2 \end{pmatrix} X.$$

La matriz tiene el siguiente polinomio característico:

$$p_A(\lambda) = (1 - \lambda)^2(2 - \lambda).$$

El valor propio $\lambda_1 = 1$ tiene multiplicidad $k = 2$. Resolvemos

$$\begin{aligned} (A - \lambda_1 I) \begin{pmatrix} a \\ b \\ c \end{pmatrix} &= \\ \begin{pmatrix} 0 & 0 & 0 \\ -4 & 0 & 0 \\ 3 & 6 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \end{aligned}$$

Esto implica que $a = 0$ y $6b + c = 0$. Si escogemos $b = 1$ entonces $c = -6$. De este modo, encontramos un solo vector propio $\begin{pmatrix} 0 \\ 1 \\ -6 \end{pmatrix}$.

Concluimos que $p = 1$ y por lo tanto debemos encontrar un vector propio generalizado de rango 2 (cf. Teorema C.2.2). Tenemos que encontrar un vector \mathbf{w} tal que $(A - I)^2 \mathbf{w} = \mathbf{0}$ y $(A - I)^1 \mathbf{w} \neq \mathbf{0}$.

Si hacemos $\mathbf{w} = \begin{pmatrix} A \\ B \\ C \end{pmatrix}$, el vector debe satisfacer

$$\begin{pmatrix} 0 & 0 & 0 \\ -4 & 0 & 0 \\ 3 & 6 & 1 \end{pmatrix} \begin{pmatrix} A \\ B \\ C \end{pmatrix} = \begin{pmatrix} 0 \\ -4A \\ 3A + 6B + C \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

y

$$(A - I)^2 \begin{pmatrix} A \\ B \\ C \end{pmatrix} =$$

$$\begin{pmatrix} 0 & 0 & 0 \\ -4 & 0 & 0 \\ 3 & 6 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ -4 & 0 & 0 \\ 3 & 6 & 1 \end{pmatrix} \begin{pmatrix} A \\ B \\ C \end{pmatrix} =$$

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ -21 & 6 & 1 \end{pmatrix} \begin{pmatrix} A \\ B \\ C \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Si $a = 0$, esto no es posible, por lo que tenemos forzosamente que $A \neq 0$. Si escogemos $A = 1$ entonces $B = 0$ y $C = 21$. Concluimos, por lo tanto, que

$$\mathbf{w} = \begin{pmatrix} 1 \\ 0 \\ 21 \end{pmatrix}$$

es vector propio generalizado de rango 2. La cadena correspondiente a \mathbf{w} es $\{\mathbf{v}_1, \mathbf{v}_2\}$ donde $\mathbf{v}_2 = \mathbf{w} = \begin{pmatrix} 1 \\ 0 \\ 21 \end{pmatrix}$ y $\mathbf{v}_1 = (A - I)\mathbf{v}_2 = (A - I)\mathbf{w} = \begin{pmatrix} 0 \\ -4 \\ 24 \end{pmatrix}$. De esto, obtenemos las siguientes soluciones linealmente independientes al problema lineal:

$$X_1(t) = e^{\lambda t}\mathbf{v}_1 = e^t \begin{pmatrix} 0 \\ -4 \\ 24 \end{pmatrix},$$

$$X_2(t) = e^{\lambda t}(\mathbf{v}_2 + t\mathbf{v}_1) = e^t \left(\begin{pmatrix} 1 \\ 0 \\ 21 \end{pmatrix} + t \begin{pmatrix} 0 \\ -4 \\ 24 \end{pmatrix} \right)$$

$$= \begin{pmatrix} e^t \\ -4te^t \\ 21e^t + 24te^t \end{pmatrix}.$$

Por otro lado, para $\lambda_2 = 2$, se encuentra que $X_3(t) = e^{2t} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$. La solución general se escribe como $X(t) = c_1 X_1(t) + c_2 X_2(t) + c_3 X_3(t)$.

Ej C.3.2 Encontrar la solución general de $\dot{X} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} X$.

El valor propio $\lambda = 1$ tiene multiplicidad 4. Es fácil ver que $p = 2$. Tenemos que encontrar dos vectores propios generalizados de rangos r_1 y r_2 tales que $r_1 + r_2 = 4$, es decir tenemos las siguientes posibilidades:

$$\begin{aligned}r_1 &= 3 \text{ y } r_2 = 1, \\r_1 &= 2 \text{ y } r_2 = 2.\end{aligned}$$

Por otra parte,

$$(A - I)^2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Si \mathbf{w} fuera de rango 3, entonces $(A - I)^2\mathbf{w} \neq 0$. Por lo tanto, no existen vectores propios generalizados de rango 3. La única posibilidad es encontrar dos vectores propios generalizados de rango 2, es decir, vectores \mathbf{w}_1 y \mathbf{w}_2 tales que

$$\begin{aligned}(A - I)^2\mathbf{w}_i &= 0, \\(A - I)\mathbf{w}_i &\neq 0,\end{aligned}$$

para $i = 1, 2$. En este problema,

$$(A - I) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Escogemos $\mathbf{w}_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$ y $\mathbf{w}_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$. Usando estos dos vectores propios generalizados (cf. Proposición C.3.1), obtenemos las siguientes soluciones linealmente independientes:

$$X_1(t) = e^t \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} e^t \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

$$X_2(t) = e^t \left(\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + t \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right) = \begin{pmatrix} te^t \\ e^t \\ 0 \\ 0 \end{pmatrix},$$

$$X_3(t) = e^t \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ e^t \end{pmatrix},$$

$$X_4(t) = e^t \left(\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right) = \begin{pmatrix} 0 \\ 0 \\ e^t \\ te^t \end{pmatrix}.$$

La solución general es $X(t) = c_1 X_1(t) + c_2 X_2(t) + c_3 X_3(t) + c_4 X_4(t)$.



Ejercicios

▷ **C.1** Encontrar la solución del siguiente sistema lineal:

$$\dot{X} = \begin{pmatrix} 0 & 0 & 0 \\ -1 & -1 & -1 \\ 1 & 1 & 1 \end{pmatrix} X,$$

si suponemos que $X(0) = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$.

▷ **C.2** Considerar el sistema lineal $\dot{X} = AX$, donde

$$A = \begin{pmatrix} -3 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & -1 & 2 \end{pmatrix}.$$

No todas las condiciones iniciales corresponden a soluciones que converjan al origen. Suponer que $X(0) = (u, v, w)$. Dar condiciones necesarias y suficientes para u, v y w si se desea que

a) $\lim_{t \rightarrow \infty} X(t) = \mathbf{0}$.

b) $\lim_{t \rightarrow -\infty} X(t) = \mathbf{0}$.

APÉNDICE D

Elementos básicos de topología en \mathbb{R}^n

En este apéndice se dan algunos conceptos elementales de topología en el espacio \mathbb{R}^n . Se utiliza la notación usual de conjuntos; así \subset , \cap y \cup denotan, respectivamente, la contención, la intersección y la unión de conjuntos y dado un conjunto $X \subset \mathbb{R}^n$, su complemento se denota por X^c . Se sugiere seguir las demostraciones con atención ya que uno de los propósitos de este apéndice es el tratar de introducir el tipo de demostración formal que se encuentra con frecuencia en cursos avanzados.

§D.1 Algo sobre números reales

Para comenzar, recordemos algunas propiedades de los números reales. Los reales son un campo ordenado con las operaciones de suma y producto; poseen la propiedad de ser completos, es decir, dados A y B conjuntos no vacíos de reales con la propiedad

$$a \leq b \quad \forall a \in A, b \in B,$$

entonces se tiene

$$\exists r \in \mathbb{R} \text{ tal que } \forall a \in A, b \in B \text{ se cumple } a \leq r \leq b.$$

Los números reales se pueden identificar con los puntos de una recta, es decir, a cada número real le corresponde un punto de la recta y viceversa. Es común, entonces, hablar de la recta de números reales o recta real. Los números racionales, \mathbb{Q} , son algebraicamente similares a los reales ya que también forman un campo ordenado con las operaciones de suma y producto. Sin embargo, no podemos identificar los racionales con los puntos de una recta puesto que en ésta existen puntos que no corresponden a ningún racional. Intuitivamente, si concebimos a los racionales como puntos de la recta real, existe una infinidad de “huecos”; así, tenemos que los puntos que corresponden, por ejemplo, a \sqrt{p} , para cualquier número primo p , o bien los correspondientes a números como π y e , no se pueden identificar con racional

alguno. A este tipo de números se los llama **irracionales**. Existen otras propiedades, aparte de la *completez* mencionada arriba, que determinan el comportamiento de los números reales como se ve a continuación.

Definición D.1.1 *Dado $X \subset \mathbb{R}$, $X \neq \emptyset$, se dice que s es una **cota superior** de X si $s \geq x$ para todo $x \in X$. De manera análoga se define una **cota inferior**. Si X posee una cota superior (inferior) se dice que X está **acotado superiormente (inferiormente)** o **acotado por arriba (por abajo)**.*

Definición D.1.2 *Dado $X \subset \mathbb{R}$, se define el **supremo** de X que se denota por $\sup X$, como la mínima cota superior del conjunto, es decir, $\sup X \geq x$, $\forall x \in X$, y si $c \geq x$, $\forall x \in X$, entonces $\sup X \leq c$. De manera análoga se define el **ínfimo** de X , que se denota como $\inf X$.*

Una propiedad sumamente importante de los números reales, conocida como el **axioma del supremo**, es que todo conjunto de reales que está acotado por arriba posee un supremo. Análogamente, si el conjunto está acotado inferiormente entonces posee un ínfimo. El axioma del supremo es equivalente a la propiedad de completez mencionada arriba. Otra propiedad que caracteriza a los números reales es la **propiedad arquimediana**, que simplemente dice que los naturales, \mathbb{N} , no están acotados superiormente de manera que dado $r \in \mathbb{R}$, existe $n \in \mathbb{N}$ tal que $n > r$.

Ejemplos

Ej D.1.1 Sea $X \subset \mathbb{R}$. Si s es cota superior de X , entonces $s + n$ también es cota superior para toda $n \in \mathbb{N}$; en particular, concluimos que si el conjunto de cotas superiores es distinto del vacío, entonces es infinito. Se tiene el resultado análogo para cotas inferiores.

Ej D.1.2 Sea $X = \{1, 2, \dots, n\}$ el elemento máximo de X , denotado por $\max X$, es un elemento de X tal que $x \leq \max X$ para toda $x \in X$. En este ejemplo $\max X = n = \sup X$. En general si existe $\max X$, éste coincide con $\sup X$, sin embargo el conjunto X puede tener un supremo y no poseer un elemento máximo. Por ejemplo, sea $X = (0, 1) = \{x \in \mathbb{R} \mid 0 < x < 1\}$, se tiene que $\sup X = 1$ pero $\max X$ no existe.

Ej D.1.3 ($\sqrt{2}$ es un número real.) Vamos a ver que una solución a la ecuación $x^2 = 2$ es un número real positivo. En efecto, consideremos el conjunto

$$S = \{y \in \mathbb{R} \mid y \geq 0, y^2 \leq 2\}.$$

Es fácil ver que 2 es una cota superior de S . Por el axioma del supremo, existe $x = \sup S$ y claramente $x \geq 0$. Supongamos que $x^2 > 2$, entonces, la propiedad arquimediana implica (ver ejercicio D.3) que existe $n \in \mathbb{N}$ tal que

$$0 < \frac{1}{n} < \frac{x^2 - 2}{2x}. \quad (\text{D.1})$$

Como $x = \sup S$, se tiene que existe $x_0 \in S$ tal que $x_0 > x - \frac{1}{n}$ (de otra forma $x - \frac{1}{n}$ sería una cota superior más pequeña que x) y por lo tanto

$$x_0^2 > \left(x - \frac{1}{n}\right)^2.$$

De D.1 se obtiene

$$\begin{aligned} 2 &< x^2 - \frac{2x}{n} \Rightarrow \\ 2 &< x^2 - \frac{2x}{n} + \left(\frac{1}{n}\right)^2 \Rightarrow \\ 2 &< \left(x - \frac{1}{n}\right)^2 \Rightarrow \\ 2 &< x_0^2. \end{aligned}$$

Esta última desigualdad es una contradicción ya que por hipótesis $x_0 \in S$, de manera que $x_0^2 \leq 2$. Tenemos así que $x^2 \leq 2$. Para concluir, supongamos que $x^2 < 2$. Utilizando nuevamente la propiedad arquimediana se escoge $m \in \mathbb{N}$ tal que

$$0 < \frac{1}{m} < \frac{2-x^2}{2x+1},$$

lo cual implica

$$\frac{2x+1}{m} < 2 - x^2. \quad (\text{D.2})$$

Se tiene entonces que

$$\begin{aligned} \left(x + \frac{1}{m}\right)^2 &= x^2 + \frac{2x}{m} + \left(\frac{1}{m}\right)^2 \\ &\leq x^2 + \frac{2x}{m} + \frac{1}{m} = x^2 + \frac{2x+1}{m} \\ &< x^2 + (2 - x^2) = 2, \end{aligned}$$

en donde la última desigualdad es consecuencia directa de D.2. Se concluye que $\left(x + \frac{1}{m}\right)^2 < 2$ y por lo tanto $\left(x + \frac{1}{m}\right)^2 \in S$, lo cual es imposible pues x es cota superior de S . De lo anterior se tiene que la única posibilidad es tener $x^2 = 2$. En consecuencia, $x = \sqrt{2}$.

◇◇◇◇◇◇◇◇

§D.2 Abiertos, cerrados y algo más.

El espacio $\mathbb{R}^n = \{(x_1, x_2, \dots, x_n) \mid x_i \in \mathbb{R} \forall i = 1, 2, \dots, n\}$ es un espacio vectorial sobre \mathbb{R} y es el espacio sobre el cual se desarrollaron los cursos de cálculo, que el lector debe haber tomado, y, ahora, los conceptos elementales de topología que damos a continuación. Los elementos de \mathbb{R}^n se conocen como **vectores** y, como sabemos, éstos se pueden sumar entre sí y multiplicar por elementos de \mathbb{R} , conocidos como **escalares**.

Definición D.2.1 La **norma (euclíadiana)** de un vector $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ es

$$\|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}.$$

Es fácil ver que la norma posee las siguientes propiedades para todo par de vectores $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ y cualquier escalar $a \in \mathbb{R}$:

- a) $\|\mathbf{x}\| \geq 0$,
- b) $\|\mathbf{x}\| = 0 \Leftrightarrow \mathbf{x} = 0$,
- c) $\|a\mathbf{x}\| = |a| \|\mathbf{x}\|$,
- d) $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$.

Una definición más general es la siguiente.

Definición D.2.2 Sea X un conjunto. Una función $d : X \times X \rightarrow \mathbb{R}_+$ es llamada una **distancia** si, dados $\mathbf{x}, \mathbf{y}, \mathbf{z} \in X$, se cumple:

- a) $d(\mathbf{x}, \mathbf{y}) \geq 0$,
- b) $d(\mathbf{x}, \mathbf{y}) = 0 \Leftrightarrow \mathbf{x} = \mathbf{y}$,
- c) $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$,
- d) $d(\mathbf{x}, \mathbf{y}) \leq d(\mathbf{x}, \mathbf{z}) + d(\mathbf{z}, \mathbf{y})$. A esta propiedad se la conoce como **desigualdad del triángulo**.

En general, cualquier conjunto X que posee una función distancia $d : X \times X \rightarrow \mathbb{R}_+$ como la de arriba, se denomina un **espacio métrico**. La siguiente proposición es consecuencia inmediata de la definición de distancia. Conviene notar que con base en la norma, puede definirse una distancia.

Proposición D.2.3 Sea $X = \mathbb{R}^n$. Para $\mathbf{x}, \mathbf{y} \in X$, se define

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|.$$

Entonces d es una distancia en X .

Definición D.2.4 La **bola abierta** con centro en \mathbf{x}_0 y radio r es el conjunto

$$B_r(\mathbf{x}_0) = \{\mathbf{z} \in \mathbb{R}^n : \|\mathbf{z} - \mathbf{x}_0\| < r\},$$

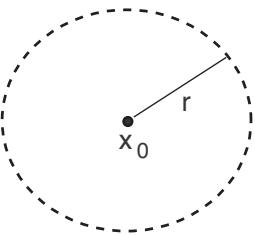
es decir, el conjunto de puntos que distan de \mathbf{x}_0 en menos de r .

Ejemplos

Ej D.2.1 En \mathbb{R} , la norma $\|x\|$ de un real x es simplemente su valor absoluto, $|x|$.

Ej D.2.2 Si el espacio es \mathbb{R} , la bola abierta con centro en x_0 y radio r es

$$B_r(x_0) = (x_0 - r, x_0 + r).$$

Figura D.1: Bola abierta en \mathbb{R}^2 .

Ej D.2.3 Si el espacio es \mathbb{R}^2 , la bola abierta con centro en x_0 y radio r es $B_r(x_0) = \{\mathbf{z} \in \mathbb{R}^2 : d(\mathbf{z}, x_0) < r\}$, como se muestra en la figura D.1. Nótese que no se incluyen los puntos de la circunferencia.

◇◇◇◇◇◇◇◇

Definición D.2.5 Decimos que $\mathbf{p} \in G \subset \mathbb{R}^n$ es un **punto interior** de G si

$$\exists r > 0 \quad \text{tal que } B_r(\mathbf{p}) \subset G.$$

Definición D.2.6 Un conjunto $G \subset \mathbb{R}^n$ es **abierto** si

$$\forall \mathbf{p} \in G \quad \exists r > 0 \quad \text{tal que } B_r(\mathbf{p}) \subset G,$$

es decir, todos los puntos de G son interiores, no existen puntos “fronterizos” (esto quedará claro más adelante).

Ejemplo

Ej D.2.4 El conjunto $X = \{(x, y) \mid x > 1; y < -1\}$ es un abierto en \mathbb{R}^2 y éste se ilustra en la figura D.2.

◇◇◇◇◇◇◇◇

Proposición D.2.7 La bola abierta $B_r(x_0)$ es un conjunto abierto.

Demostración

Sea $\mathbf{y} \in B_r(x_0)$. Por definición $\|\mathbf{y} - x_0\| = d(x_0, \mathbf{y}) < r$ de manera que $r - d(x_0, \mathbf{y}) > 0$. Sea $s = \frac{1}{2}[r - d(x_0, \mathbf{y})]$; veamos que $B_s(\mathbf{y}) \subset B_r(x_0)$, como se ilustra en la figura D.3. Sea $\mathbf{z} \in B_s(\mathbf{y})$, entonces se cumple $d(\mathbf{y}, \mathbf{z}) < s = \frac{1}{2}[r - d(x_0, \mathbf{y})]$. De aquí que $2d(\mathbf{y}, \mathbf{z}) + d(x_0, \mathbf{y}) < r$ y por lo tanto $d(\mathbf{y}, \mathbf{z}) + d(x_0, \mathbf{y}) < r$. Utilizando la desigualdad del triángulo tenemos que $d(x_0, \mathbf{z}) < d(\mathbf{y}, \mathbf{z}) + d(x_0, \mathbf{y}) < r$ y por lo tanto $\mathbf{z} \in B_r(x_0)$. Como \mathbf{z} es un elemento arbitrario de $B_s(\mathbf{y})$, se tiene que $B_s(\mathbf{y}) \subset B_r(x_0)$, con lo que concluimos la demostración. ■

Definición D.2.8 Un conjunto $G \subset \mathbb{R}^n$ es **cerrado** si su complemento, G^c , es abierto.

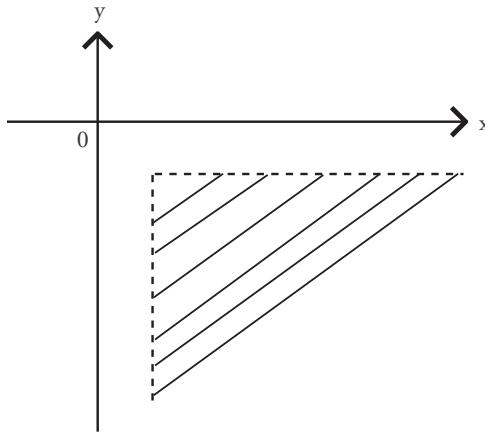
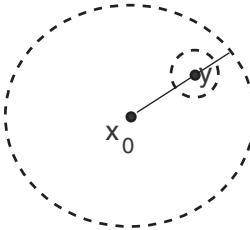
Figura D.2: Un abierto en \mathbb{R}^2 .

Figura D.3: La bola Abierta es Abierta...

Definición D.2.9 *Un conjunto $G \subset \mathbb{R}^n$ es acotado si*

$$\exists r \in \mathbb{R} \text{ tal que } \|\mathbf{x}\| < r \quad \forall \mathbf{x} \in G.$$

Definición D.2.10 *Un conjunto G es compacto si es cerrado y acotado.*

Definición D.2.11 *Sean $G \subset \mathbb{R}^n$ y $\mathbf{p} \in \mathbb{R}^n$. Decimos que \mathbf{p} es un punto frontera de G si*

$$\forall r > 0 \text{ se cumplen } B_r(\mathbf{p}) \cap G \neq \emptyset \text{ y } B_r(\mathbf{p}) \cap G^c \neq \emptyset.$$

Esta última definición aclara el concepto de punto “fronterizo” expuesto de forma intuitiva en la definición D.2.6.

Proposición D.2.12 *El conjunto $G \subset \mathbb{R}^n$ es cerrado si y sólo si contiene todos sus puntos frontera.*

Demostración

- Necesidad.

Supongamos primero que G es cerrado, es decir G^c es abierto, y sea \mathbf{p} un punto frontera arbitrario de G . Si $\mathbf{p} \in G^c$, entonces G^c abierto implica que existe $r > 0$ tal que $B_r(\mathbf{p}) \subset G^c$. Se tiene entonces que $B_r(\mathbf{p}) \cap G = \emptyset$ lo cual es imposible dada la definición de punto frontera. De esta forma debemos tener que $\mathbf{p} \in G$, y como \mathbf{p} es arbitrario, G contiene a todos sus puntos frontera.

- Suficiencia.

Supongamos ahora que G contiene a todos sus puntos frontera; vamos a demostrar que G^c es abierto. Sea $\mathbf{p} \in G^c$. Como G contiene a todos sus puntos frontera, \mathbf{p} no es punto frontera de G , es decir, existe $r > 0$ tal que $B_r(\mathbf{p}) \cap G = \emptyset$ o $B_r(\mathbf{p}) \cap G^c = \emptyset$. Como $\mathbf{p} \in B_r(\mathbf{p}) \cap G^c$, es decir, $B_r(\mathbf{p}) \cap G^c \neq \emptyset$ debe cumplirse, entonces $B_r(\mathbf{p}) \cap G = \emptyset$, lo que implica que $B_r(\mathbf{p}) \subset G^c$. Concluimos que G^c es abierto y en consecuencia G es cerrado. ■

Ejemplos

Ej D.2.5 El intervalo $[0, 1]$ es acotado pero no es ni abierto ni cerrado ni compacto.

Ej D.2.6 \mathbb{R}^n y \emptyset son abiertos y cerrados.

Ej D.2.7 El conjunto que consta de un solo punto, $\{\mathbf{p}\}$, es cerrado.

Ej D.2.8 El intervalo $[0, \infty)$ es cerrado pero no es acotado y por lo tanto no es compacto.

Ej D.2.9 Cualquier intervalo de la forma $[a, b]$ es compacto; la unión de un número finito de estos intervalos también es compacta.



Proposición D.2.13 *Las siguientes propiedades se cumplen:*

- a) *la unión arbitraria de conjuntos abiertos es abierta y la intersección de un número finito de abiertos es abierta.*
- b) *la intersección arbitraria de conjuntos cerrados es cerrada y la unión de un número finito de cerrados es cerrada.*

Demostración

Se demostrará sólo el inciso a y se dejará la demostración del b para el ejercicio D.7. Para la primera parte, sea $\{G_i\}_{i \in I}$ un conjunto arbitrario de abiertos en \mathbb{R}^n . Debemos probar que $G = \bigcup_{i \in I} G_i$ es abierto. Sea $\mathbf{p} \in G$, entonces $\mathbf{p} \in G_i$ para algún índice i . Como G_i es abierto, existe $r > 0$ tal que $B_r(\mathbf{p}) \subset G_i$ y por lo tanto $B_r(\mathbf{p}) \subset G$, con lo cual queda demostrado que G es abierto.

Para la segunda parte sea, $\{H_i\}_{i=1}^k$ un conjunto finito de abiertos en \mathbb{R}^n . Debemos ahora probar que $H = \bigcap_{i=1}^k H_i$ es abierto. Sea $\mathbf{p} \in H$, entonces $\mathbf{p} \in H_i$ para toda $i = 1, 2, \dots, k$; de aquí que existen r_1, r_2, \dots, r_k tales que $B_{r_i}(\mathbf{p}) \subset H_i$. Sea $r = \min\{r_i\}_{i=1}^k$, entonces $B_r(\mathbf{p}) \subset \bigcap_{i=1}^k H_i = H$ y por lo tanto H es abierto, lo que da por terminada la demostración. ■

La proposición anterior nos permite definir los siguientes conceptos.

Definición D.2.14 *Sea $G \subset \mathbb{R}^n$.*

- a) *El **interior** de G es el conjunto $G^\circ \equiv \bigcup_{i \in I} G_i$, en donde $\{G_i\}_{i \in I}$ es el conjunto de todos los abiertos G_i tales que $G_i \subset G$. El interior de G se puede interpretar como el abierto más grande contenido en G . Otra notación común es $G^\circ = \text{int } G$.*
- b) *La **cerradura** de G es el conjunto $\bar{G} = \bigcap_{i \in I} G_i$, en donde $\{G_i\}_{i \in I}$ es el conjunto de todos los cerrados G_i tales que $G \subset G_i$. La cerradura de G se puede interpretar como el cerrado más pequeño que contiene a G . Otra notación común es $\bar{G} = \text{cerr } G$ (o $\text{cl } G$, del inglés closure).*

Ejemplos

Ej D.2.10 Los intervalos $(-\frac{1}{n}, \frac{1}{n})$ son abiertos para toda $n \in \mathbb{N}$; nótese, sin embargo, que la intersección $\bigcap_{n \in \mathbb{N}} (-\frac{1}{n}, \frac{1}{n}) = \{0\}$ no es abierta. Esto no contradice la proposición anterior ya que se trata de una intersección infinita.

Ej D.2.11 $(-1, 5)^\circ = (-1, 5)$, $\overline{(-1, 5)} = [-1, 5]$.

Ej D.2.12 Sean $\mathbf{p} \in \mathbb{R}^n$ y $r > 0$, entonces $\overline{B_r(\mathbf{p})} = \{\mathbf{x} \in \mathbb{R}^n \mid d(\mathbf{x}, \mathbf{p}) \leq r\}$. El conjunto $\overline{B_r(\mathbf{p})}$ es conocido como la **bola cerrada** con centro en \mathbf{p} y radio r .



§D.3 Sucesiones

Definición D.3.1 *Una sucesión de puntos en \mathbb{R}^m es una función $f : \mathbb{N} \rightarrow \mathbb{R}^m$.*

La notación usual para los elementos de la imagen de f es $f(n) = \mathbf{x}_n$ y normalmente la sucesión se expresa como el conjunto de estos elementos, es decir, $\{\mathbf{x}_n\}_{n \in \mathbb{N}}$ o simplemente por $\{\mathbf{x}_n\}$.

Definición D.3.2 *Sea $\{\mathbf{x}_n\} \subset \mathbb{R}$, una sucesión correspondiente a la función $f : \mathbb{N} \rightarrow \mathbb{R}$. Decimos que $\{\mathbf{x}_n\}$ es **creciente (decreciente)** si la función f es **creciente (decreciente)**. En general, si una sucesión es creciente o decreciente, decimos que es **monótona**.*

Definición D.3.3 *Sean $G \subset \mathbb{R}^m$ y $\mathbf{p} \in \mathbb{R}^m$. Decimos que \mathbf{p} es un **punto límite o punto de acumulación** de G si*

$$\forall r > 0 \quad \exists \mathbf{q} \in G, \mathbf{q} \neq \mathbf{p}, \text{ tal que } \mathbf{q} \in B_r(\mathbf{p}) \cap G.$$

Es importante no confundir puntos límite con puntos frontera: si $G = \{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_m\} \subset \mathbb{R}^n$, es decir un conjunto finito de puntos, entonces todos ellos son puntos frontera de G , sin embargo, G no tiene puntos límite. De aquí podemos concluir que solo los conjuntos infinitos pueden poseer puntos límite.

Definición D.3.4 Decimos que la sucesión $\{\mathbf{x}_n\}$ converge a $\mathbf{x} \in \mathbb{R}^m$ si

$$\forall \varepsilon > 0, \exists N \in \mathbb{N} \text{ tal que } n > N \Rightarrow d(\mathbf{x}_n, \mathbf{x}) < \varepsilon.$$

En este caso decimos que \mathbf{x} es el **límite de la sucesión** lo que se denota por $\lim_{n \rightarrow \infty} \mathbf{x}_n = \mathbf{x}$.

La idea intuitiva detrás de esta definición es clara: podemos acercarnos arbitrariamente al límite siempre y cuando avancemos suficientemente “lejos” en la sucesión. La siguiente proposición dice que un punto límite de G es el límite de una sucesión de puntos de G .

Proposición D.3.5 Sea $G \subset \mathbb{R}^m$. \mathbf{p} es un punto límite de G si y sólo si existe una sucesión $\{\mathbf{p}_n\} \subset G$ tal que $\lim_{n \rightarrow \infty} \mathbf{p}_n = \mathbf{p}$.

Demostración

- Necesidad.

Supongamos primero que \mathbf{p} es punto límite de G . Vamos a construir una sucesión que converja a \mathbf{p} . Dado $n \in \mathbb{N}$, existe $\mathbf{p}_n \in G$, $\mathbf{p}_n \neq \mathbf{p}$, tal que $\mathbf{p}_n \in B_{\frac{1}{n}}(\mathbf{p})$. La sucesión $\{\mathbf{p}_n\}$ es un sucesión de puntos de G que satisface $\lim_{n \rightarrow \infty} \mathbf{p}_n = \mathbf{p}$.

- Suficiencia.

Ahora consideremos la sucesión $\{\mathbf{p}_n\} \subset G$ tal que $\lim_{n \rightarrow \infty} \mathbf{p}_n = \mathbf{p}$. Vamos a ver que \mathbf{p} es punto límite de G . Sea $r > 0$, entonces $\lim_{n \rightarrow \infty} \mathbf{p}_n = \mathbf{p}$ implica que existe $N \in \mathbb{N}$ tal que, si $n > N$, entonces $\mathbf{p}_n \in B_r(\mathbf{p})$. Como $\mathbf{p}_n \in G$ se tiene que \mathbf{p} es punto límite de G , con lo que terminamos la demostración. ■

Ejemplos

Ej D.3.1 La sucesión $\{n\}$ es una sucesión creciente y la sucesión $\{\frac{1}{n}\}$ es decreciente.

Ej D.3.2 Los naturales \mathbb{N} son un subconjunto infinito de \mathbb{R} ; sin embargo, no poseen ningún punto límite. De aquí se infiere que ser infinito es una condición necesaria, más no suficiente, para que un conjunto tenga puntos límite.

Ej D.3.3 La sucesión $\{(\frac{2n^2}{n^2-1}, \frac{3n}{n+5})\}$ de puntos de \mathbb{R}^2 converge a $(2, 3)$.

Ej D.3.4 Si $p \in [0, 1]$, entonces p es punto límite de $[0, 1]$. Para verificarlo, consideraremos dos casos. Si p coincide con uno de los extremos, es decir $p = 0$ o $p = 1$, entonces la sucesión $\{\frac{1}{n}\}$ converge a $p = 0$ o bien $\{1 - \frac{1}{n}\}$ converge a $p = 1$. Si $p \in (0, 1)$, entonces, como este intervalo es abierto, existe $r > 0$ tal que $B_r(p) = (p - r, p + r) \subset (0, 1)$; si definimos, para toda $n \in \mathbb{N}$, $p_n = p - \frac{r}{n}$, entonces $\lim_{n \rightarrow \infty} p_n = p$. Concluimos que en cualquiera de los casos p es punto límite de $[0, 1]$.



Los conjuntos cerrados pueden caracterizarse en términos de sucesiones como sigue.

Proposición D.3.6 *El conjunto $G \subset \mathbb{R}^m$ es cerrado si y sólo si*

$$\begin{aligned}\forall \{\mathbf{x}_n\} \subset G \text{ se tiene que} \\ \lim_{n \rightarrow \infty} \mathbf{x}_n = \mathbf{x} \Rightarrow \mathbf{x} \in G.\end{aligned}$$

Es decir, G es cerrado si y sólo si contiene a todos sus puntos límite.

Demostración

- Necesidad.

Sea $G \subset \mathbb{R}^m$ cerrado y sea $\{\mathbf{x}_n\} \subset G$ una sucesión de puntos de G tal que $\lim_{n \rightarrow \infty} \mathbf{x}_n = \mathbf{x}$; debemos probar que $\mathbf{x} \in G$. Supongamos, por el contrario, que $\mathbf{x} \in G^c$. Dado que G^c es abierto, existe $r > 0$ tal que $B_r(\mathbf{x}) \subset G^c$; sin embargo, esto no es posible puesto que \mathbf{x} es punto límite de G y por lo tanto existe $N \in \mathbb{N}$ tal que $\mathbf{x}_n \in B_r(\mathbf{x})$ para toda $n > N$, de manera que $B_r(\mathbf{x}) \cap G \neq \emptyset$. De aquí se debe tener que $\mathbf{x} \notin G^c$, es decir, $\mathbf{x} \in G$.

- Suficiencia.

Supongamos ahora que G contiene a todos sus puntos límite; debemos probar que G^c es abierto. Sea $\mathbf{x} \in G^c$. Entonces \mathbf{x} no es punto límite de G y existe $r > 0$ tal que $B_r(\mathbf{x}) \cap G = \emptyset$, es decir, $B_r(\mathbf{x}) \subset G^c$ con lo cual se tiene que G^c es abierto y por lo tanto G es cerrado. ■

Proposición D.3.7 *Sea $\{x_n\} \subset \mathbb{R}$ una sucesión creciente (decreciente) y acotada por arriba (por abajo), entonces $\lim_{n \rightarrow \infty} x_n = \sup\{x_n\}$ ($\lim_{n \rightarrow \infty} x_n = \inf\{x_n\}$).*

Demostración

Supongamos que $\{x_n\} \subset \mathbb{R}$ es una sucesión creciente y acotada por arriba. Sabemos, por el axioma del supremo, que existe $s = \sup\{x_n\}$; falta ver que éste efectivamente es el límite de la sucesión. Sea $\varepsilon > 0$. Dado que $s = \sup\{x_n\}$, tenemos que $s - \varepsilon$ no puede ser cota superior y por lo tanto existe $N \in \mathbb{N}$ tal que $s \geq x_N > s - \varepsilon$. Como la sucesión es creciente se cumple $s \geq x_n > s - \varepsilon$ para toda $n \geq N$, y por lo tanto

$$s - \varepsilon < x_n < s + \varepsilon, \quad \forall n > N,$$

o bien, $d(x_n, s) < \varepsilon, \forall n > N$. Así, dada $\varepsilon > 0$, existe $N \in \mathbb{N}$ tal que

$$n > N \Rightarrow d(x_n, s) < \varepsilon,$$

es decir, $\lim_{n \rightarrow \infty} x_n = s$. ■

Definición D.3.8 Sea $f : \mathbb{N} \rightarrow \mathbb{R}^m$ una sucesión y sea $g : \mathbb{N} \rightarrow \mathbb{N}$ una función estrictamente creciente, es decir, $n_1 < n_2 \Rightarrow g(n_1) < g(n_2)$. La composición $f \circ g : \mathbb{N} \rightarrow \mathbb{R}^m$ se conoce como una **subsucesión** de f . Con la notación usual, la sucesión f está dada por $\{\mathbf{x}_n\}$ y la subsucesión $f \circ g$ como $\{\mathbf{x}_{n_k}\}$, en donde $n_k = g(n)$.

Ejemplos

Ej D.3.5 $\{2, 4, \dots, 2n, \dots\} = \{2n\}$ es una subsucesión de $\{1, 2, \dots, n, \dots\} = \{n\}$.

Ej D.3.6 En general, dada una sucesión $\{\mathbf{x}_n\}$ en \mathbb{R}^m , podemos obtener la subsucesión de términos pares, $\{\mathbf{x}_{2n}\}$, o bien la sucesión de términos impares, $\{\mathbf{x}_{2n-1}\}$.

◇◇◇◇◇◇◇◇◇

Proposición D.3.9 Toda sucesión $\{x_n\} \subset \mathbb{R}$ posee una subsucesión monótona.

Demostración

Sea $\{x_n\} \subset \mathbb{R}$. Consideremos el siguiente conjunto:

$$M = \{m \in \mathbb{N} \mid x_m > x_n \ \forall n > m\}.$$

Si M es un conjunto infinito, dados m_1 y m_2 en M

$$m_1 < m_2 \Rightarrow x_{m_1} > x_{m_2},$$

y por lo tanto $\{x_m\}_{m \in M}$ es una subsucesión decreciente de $\{x_n\}$.

Si $M = \emptyset$, entonces $M^c = \mathbb{N}$, de manera que se cumple

$$\forall m \in \mathbb{N}, \exists n \in \mathbb{N}, n > m \text{ tal que } x_m \leq x_n.$$

Podemos escoger así la subsucesión

$$\begin{aligned} m_1 &= 1, \\ m_2 &> m_1 \text{ tal que } x_{m_2} \geq x_{m_1}, \\ m_3 &> m_2 \text{ tal que } x_{m_3} \geq x_{m_2}, \\ &\vdots \end{aligned}$$

De esta forma, la subsucesión $\{x_{m_k}\}$ es una subsucesión creciente de $\{x_n\}$.

Finalmente, si M es un conjunto finito sea \bar{m} su elemento máximo. Nótese que $\{\bar{m}+1, \bar{m}+2, \dots\} \subset M^c$. Construimos ahora la subsucesión como

$$\begin{aligned} m_1 &= \bar{m} + 1, \\ m_2 &> m_1 \text{ tal que } x_{m_2} \geq x_{m_1}, \\ m_3 &> m_2 \text{ tal que } x_{m_3} \geq x_{m_2}, \\ &\vdots \end{aligned}$$

La subsucesión $\{x_{m_k}\}$ es, una vez más, una subsucesión creciente de $\{x_n\}$. Se concluye así la demostración. ■

Teorema D.3.10 (Bolzano-Weierstrass)

Sea $\{\mathbf{x}_n\}$ una sucesión acotada en \mathbb{R}^m . Entonces ésta posee una subsucesión convergente, $\{\mathbf{x}_{n_k}\}$.

Demostración

Por el ejercicio D.18, es suficiente demostrar el teorema para el caso $m = 1$. La proposición D.3.7 dice que existe una subsucesión monótona. Por hipótesis esta subsucesión es acotada y la proposición D.3.7 nos dice que ésta converge ya sea a su supremo o ínfimo, según sea el caso, con lo cual queda demostrado el teorema. ■

Ejercicios

▷ **D.1** Sean $A, B \subset \mathbb{R}$ distintos del vacío y acotados por arriba. Si $A + B = \{a + b \mid a \in A, b \in B\}$, probar que

$$\sup(A + B) = \sup A + \sup B.$$

▷ **D.2** Sea $A \subset \mathbb{R}$, probar que si $\sup A$ existe, es único.

▷ **D.3** Sean $y, z > 0$.

a) Probar que existe $n \in \mathbb{N}$ tal que $0 < \frac{1}{n} < z$.

b) Probar que existe $m \in \mathbb{N}$ tal que $my > z$.

c) Probar que existe $k \in \mathbb{N}$ tal que $k - 1 < y \leq k$.

▷ **D.4** Demostrar la proposición D.2.3 utilizando las propiedades de la norma.

▷ **D.5** Sea X cualquier conjunto. Se define la siguiente función en $X \times X$:

$$d(x, y) = \begin{cases} 0 & \text{si } x = y \\ 1 & \text{si } x \neq y \end{cases}$$

a) Probar que esta función satisface la definición D.2.2.

b) Describir $B_r(x)$ para $r \leq 1$ y para $r > 1$ utilizando la función distancia definida arriba.

▷ **D.6** Sabemos que si $\|\cdot\|$ es una norma en X , entonces la función $d(x, y) = \|x - y\|$, definida en $X \times X$ es una función distancia. ¿Es válido el recíproco de este resultado? Concretamente, ¿se puede definir una norma a partir de una función distancia? Explicar.

▷ **D.7** Demostrar el inciso b de la proposición D.2.13.

▷ **D.8** Probar que \mathbb{Z} es cerrado en \mathbb{R} . ¿Es compacto? Justificar.

- ▷ **D.9** Probar que si A es cerrado y B es compacto, entonces $A \cap B$ es compacto.
- ▷ **D.10** Probar que el conjunto $\{(x, y) \in \mathbb{R}^2 \mid x, y \geq 0\}$ es cerrado en \mathbb{R}^2 .
- ▷ **D.11** Dar un ejemplo en donde la unión infinita de cerrados no es cerrada.
- ▷ **D.12** Probar la siguiente afirmación: sea $G \subset \mathbb{R}^n$ y sea \mathbf{p} un punto límite de G , entonces para toda $r > 0$, la bola $B_r(\mathbf{p})$ contiene un número infinito de puntos de G . (Sugerencia: suponer que existe $s > 0$ tal que $B_s(\mathbf{p})$ solo contiene un número finito de puntos de G , digamos $\{\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_k\}$. Definir $r \equiv \min_{1 \leq i \leq k} d(\mathbf{p}, \mathbf{q}_i)$ y probar que $B_r(\mathbf{p})$ no contiene ningún punto de G distinto de \mathbf{p} .)
- ▷ **D.13** Probar que $G \subset \mathbb{R}^n$ es abierto si y sólo si $G = G^\circ$.
- ▷ **D.14** Probar que $G \subset \mathbb{R}^n$ es cerrado si y sólo si $G = \overline{G}$.
- ▷ **D.15** Sea $G \subset \mathbb{R}^n$. Denotemos por G' al conjunto de todos los puntos límite de G . Probar que $\overline{G} = G \cup G'$. (Sugerencia: probar que para todo cerrado F tal que $G \subset F$, se tiene que $G \cup G' \subset F$ y utilizar el ejercicio anterior.)
- ▷ **D.16** Sean $G \subset \mathbb{R}^n$ y \mathbf{p} un punto frontera de G tal que $\mathbf{p} \notin G$. Demostrar que \mathbf{p} es un punto límite de G .
- ▷ **D.17** Demostrar formalmente que el conjunto $\mathbb{N} \subset \mathbb{R}$ no tiene puntos límite.
- ▷ **D.18** Demostrar que una sucesión $\{\mathbf{x}_n\} \subset \mathbb{R}^m$ converge a

$$\mathbf{x} = (x^1, x^2, \dots, x^m)$$
 si y sólo si cada coordenada converge, es decir, si $\mathbf{x}_n = (x_n^1, x_n^2, \dots, x_n^m)$, entonces $\{x_n^k\}$ converge a $\{x^k\}$ para toda $k = 1, 2, \dots, m$.
- ▷ **D.19** Probar que si $\{\mathbf{x}_n\} \subset \mathbb{R}^m$ converge, entonces el límite es único. (Sugerencia: suponer que existen dos límites, $\mathbf{L}_1 \neq \mathbf{L}_2$, y probar que los puntos de la sucesión no pueden estar “arbitrariamente cerca” de ambos.)

APÉNDICE E

Espacios vectoriales y teoremas de separación

§E.1 Propiedades básicas de los espacios vectoriales

Para entender el concepto de espacio vectorial es necesario entender antes el concepto de campo. Un

campo es un conjunto \mathbb{F} con las operaciones de suma y producto y dos elementos denotados por 0 y 1 llamados idéntico aditivo y multiplicativo, respectivamente. El conjunto \mathbb{F} , junto con las operaciones y elementos dados, satisface las propiedades de campo A.1 dadas en el apéndice A. Podemos definir ahora el concepto de espacio vectorial sobre \mathbb{F} .

Definición E.1.1 *Sea V un conjunto. Decimos que V es un **espacio vectorial sobre el campo \mathbb{F}** , si se satisfacen las siguientes propiedades:*

- a) *Existe una operación binaria de suma (+), entre los elementos de V . La suma satisface las propiedades de*
 - i) *cerradura,*
 - ii) *commutatividad,*
 - iii) *asociatividad.*
- b) *Existe $\mathbf{0} \in V$, único, tal que $\mathbf{0} + \mathbf{v} = \mathbf{v} + \mathbf{0} = \mathbf{v}, \forall \mathbf{v} \in V$.*
- c) *Existe una operación binaria de producto en $\mathbb{F} \times V$ tal que cumple las siguientes propiedades:*
 - i) *Dados $r \in \mathbb{F}, \mathbf{v} \in V$, se tiene que $r\mathbf{v} \in V$.*

- ii) $r\mathbf{v} = \mathbf{v}r, \forall r \in \mathbb{F} \text{ y } \mathbf{v} \in V.$
- iii) $r(s\mathbf{v}) = s(r\mathbf{v}) = (rs)\mathbf{v}.$
- iv) $(r+s)\mathbf{v} = \mathbf{v}(r+s) = r\mathbf{v}+s\mathbf{v} \text{ y } r(\mathbf{u}+\mathbf{v}) = (\mathbf{u}+\mathbf{v})r = r\mathbf{u}+r\mathbf{v}, \forall r, s \in \mathbb{F} \text{ y } \mathbf{u}, \mathbf{v} \in V$
- v) $\mathbf{0}\mathbf{v} = \mathbf{0}, 1\mathbf{v} = \mathbf{v}, \forall \mathbf{v} \in V.$

Los elementos del conjunto V son llamados **vectores** y los elementos del campo \mathbb{F} son llamados **escalares**.

Ejemplos

Ej E.1.1 El conjunto $V = \mathbb{R}^n = \{(x_1, \dots, x_n) \mid x_i \in \mathbb{R} \forall i = 1, \dots, n\}$ es un espacio vectorial sobre el campo \mathbb{R} .

Ej E.1.2 El conjunto $V = \mathbb{C}^n = \{(x_1, \dots, x_n) \mid x_i \in \mathbb{C} \forall i = 1, \dots, n\}$ es un espacio vectorial sobre el campo \mathbb{C} .

Ej E.1.3 Sea V el conjunto de matrices de $m \times n$ con coeficientes en un campo \mathbb{F} . V es un espacio vectorial con la suma de matrices y el producto por escalares definidos coordenada a coordenada.

Ej E.1.4 El conjunto $V = \{f : \mathbb{R} \rightarrow \mathbb{R} \mid f \text{ es continua}\}$ es un espacio vectorial sobre \mathbb{R} con $(f + g)(x) = f(x) + g(x)$ y $(rf)(x) = rf(x), \forall f, g \in V \text{ y } r \in \mathbb{R}$.

◇◇◇◇◇◇◇◇◇

En este texto se utilizan espacios vectoriales sobre \mathbb{R} , ya sea \mathbb{R}^n o espacios de funciones parecidos al ejemplo E.1.4.

Los vectores en \mathbb{R}^n pueden denotarse como vectores columna $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$, o bien como vectores renglón $\mathbf{x}^t = (x_1, x_2, \dots, x_n)$. Por analogía con el caso de matrices, decimos que \mathbf{x}^t es el vector transpuesto de \mathbf{x} . Es claro que $(\mathbf{x}^t)^t = \mathbf{x}$. En este apéndice se utiliza la notación de vectores renglón, sin embargo se alerta al lector que los vectores en el texto son denotados como vectores columna ya que esta notación es más apropiada para tratar con sistemas de ecuaciones.

Algunos espacios vectoriales, incluyendo \mathbb{R}^n y algunos espacios de funciones, poseen un **producto escalar** o **producto punto**. Éste es una operación binaria $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{F}$, que es bilineal, es decir, cumple con

$$\begin{aligned} \langle r\mathbf{u}_1 + s\mathbf{u}_2, \mathbf{v} \rangle &= r \langle \mathbf{u}_1, \mathbf{v} \rangle + s \langle \mathbf{u}_2, \mathbf{v} \rangle, \\ \langle \mathbf{u}, r\mathbf{v}_1 + s\mathbf{v}_2 \rangle &= r \langle \mathbf{u}, \mathbf{v}_1 \rangle + s \langle \mathbf{u}, \mathbf{v}_2 \rangle, \end{aligned}$$

para todo $\mathbf{u}, \mathbf{v} \in V$ y $r, s \in \mathbb{F}$. Este producto asocia un escalar con cada pareja de elementos del espacio V . La notación $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u} \cdot \mathbf{v}$ es también usual, de ahí el nombre de *producto punto*.

Ejemplos

Ej E.1.5 Sea $V = \mathbb{R}^n$. Si $\mathbf{x}^t = (x_1, \dots, x_n)$ y $\mathbf{y}^t = (y_1, \dots, y_n)$ se puede definir $\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^n x_i y_i \in \mathbb{R}$. Observemos que para este caso el producto escalar puede definirse en términos de un producto de matrices, es decir, $\mathbf{x} \cdot \mathbf{y} = \mathbf{x}^t \mathbf{y}$.

Ej E.1.6 Sea $V = \{f : [0, 1] \rightarrow \mathbb{R} \mid f \text{ es continua}\}$. V es un espacio vectorial sobre \mathbb{R} , a semejanza del ejemplo E.1.4. Entonces se puede definir $f \cdot g = \int_0^1 f(x)g(x)dx$.



Una propiedad interesante es que, si el espacio vectorial V posee un producto escalar, entonces es un espacio vectorial normado en el sentido de la definición 11.1.4, o bien de la definición de norma en \mathbb{R}^n dada en el apéndice D. Esto puede verse con la siguiente proposición, cuya demostración se deja al lector.

Proposición E.1.2 *Sea V un espacio vectorial sobre \mathbb{R} con un producto escalar. Entonces, dado $\mathbf{x} \in V$,*

$$\|\mathbf{x}\| = \sqrt{(\mathbf{x} \cdot \mathbf{x})}$$

es una norma en V .

La norma nos permite tener la noción de “tamaño” de un vector \mathbf{v} y por lo tanto se tiene también el concepto de distancia entre dos vectores \mathbf{u} y \mathbf{v} , como se vio en el apéndice D.

El producto punto nos da una forma de definir el concepto de “ortogonalidad” ya que decimos que dos vectores \mathbf{v} y \mathbf{w} , diferentes del vector $\mathbf{0}$, son **ortogonales** si y sólo si $\mathbf{v} \cdot \mathbf{w} = 0$. Para el caso del producto punto definido en \mathbb{R}^n como en el ejemplo E.1.5, se puede ver que $\mathbf{v} \cdot \mathbf{w} = \|\mathbf{v}\| \|\mathbf{w}\| \cos \theta$, donde θ es el ángulo entre \mathbf{v} y \mathbf{w} , de manera que este concepto de ortogonalidad coincide con nuestra intuición geométrica ya que $\mathbf{v} \cdot \mathbf{w} = 0$ si y sólo si $\cos \theta = 0$.

Si pensamos en el espacio vectorial \mathbb{R}^n , intuitivamente se incluyen en éste, de manera natural, los espacios \mathbb{R}^m con $m < n$. Este concepto se formaliza con la siguiente definición.

Definición E.1.3 *Sea V un espacio vectorial sobre un campo \mathbb{F} . Decimos que un subconjunto S de V es un **subespacio vectorial** de V si S es un espacio vectorial sobre \mathbb{F} con las operaciones de suma y multiplicación por escalares de V .*

Es relativamente fácil verificar que un subconjunto S de V es un subespacio vectorial si se cumplen,

- $\mathbf{u}, \mathbf{v} \in S \Rightarrow \mathbf{u} + \mathbf{v} \in S$,
- $r \in \mathbb{F}, \mathbf{v} \in S \Rightarrow r\mathbf{v} \in S$.

Ejemplos

En los siguientes ejemplos, V es cualquier espacio vectorial sobre un campo \mathbb{F} .

Ej E.1.7 $\{\mathbf{0}\}$ y V son subespacios vectoriales de V .

Ej E.1.8 Dado $\mathbf{v} \in V$, el conjunto $S = \{r\mathbf{v} \mid r \in \mathbb{F}\}$ es un subespacio vectorial de V .

Ej E.1.9 Dados $\mathbf{v}_1, \dots, \mathbf{v}_n \in V$, el conjunto

$$S = \left\{ \sum_{i=1}^n r_i \mathbf{v}_i \mid r_i \in \mathbb{F} \forall i \right\}$$

es un subespacio vectorial de V . Decimos que S es el subespacio generado por los vectores $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$.

Ej E.1.10 En este ejemplo, $V = \mathbb{R}^n$ y A es una matriz de $m \times n$. El conjunto

$$S = \{\mathbf{x} = (x_1, \dots, x_n)^t \mid A\mathbf{x} = \mathbf{0}\}$$

es un subespacio vectorial de \mathbb{R}^n .

◇◇◇◇◇◇◇◇◇

Definición E.1.4 Sea V un espacio vectorial sobre \mathbb{F} . Decimos que un conjunto de vectores $\{\mathbf{v}_1, \dots, \mathbf{v}_n\} \subset V$ es **linealmente independiente** si

$$\sum_{i=1}^n r_i \mathbf{v}_i = \mathbf{0} \Leftrightarrow r_i = 0, \forall i = 1, \dots, n.$$

Si el conjunto no es linealmente independiente decimos que es **linealmente dependiente**.

Ejemplos

Ej E.1.11 Los vectores $(1, 2, 3)$ y $(1, 1, 0)$ son linealmente independientes en \mathbb{R}^3 .

Ej E.1.12 Sea $V = \{f : \mathbb{R} \rightarrow \mathbb{R} \mid f \text{ es continua}\}$. Los polinomios $p_1(x) = 1, p_2(x) = x$ y $p_3(x) = x^2$ son linealmente independientes en V .

◇◇◇◇◇◇◇◇◇

Definición E.1.5 Sea V un espacio vectorial sobre \mathbb{F} . Decimos que un conjunto de vectores $\{\mathbf{v}_1, \dots, \mathbf{v}_n\} \subset V$ es una **base** de V si

- $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ es linealmente independiente.

- Dado $\mathbf{u} \in V$, existen $r_1, \dots, r_n \in \mathbb{F}$ tales que $\mathbf{u} = \sum_{i=1}^n r_i \mathbf{v}_i$, es decir, el conjunto $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ genera a V .

En este caso, decimos que n , el número de elementos de la base, es la **dimensión** del espacio V . La dimensión de un espacio puede ser infinita (éste es normalmente el caso para espacios de funciones, véase el ejemplo E.1.14). Si, adicionalmente, el espacio tiene un producto escalar y se cumple que

$$\mathbf{v}_i \cdot \mathbf{v}_j = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases},$$

decimos que se trata de una **base ortogonal** de V . Finalmente, si una base ortogonal cumple con $\|\mathbf{v}_i\| = 1, \forall i$, se dice que la base es **ortonormal**.

Ejemplos

Ej E.1.13 Sea $V = \mathbb{R}^n$ y sea $e_i = (0, \dots, 0, 1, 0, \dots, 0)$ donde el 1 aparece en el i -ésimo lugar y todas las otras coordenadas son 0. El conjunto $\{e_i\}_{i=1}^n$ es una base ortonormal de \mathbb{R}^n , llamada la **base canónica**.

Ej E.1.14 Sea $V = \{p : \mathbb{R} \rightarrow \mathbb{R} \mid p \text{ es un polinomio}\}$. Es fácil verificar que V es un espacio vectorial sobre \mathbb{R} . El conjunto infinito $\{1, x, x^2, \dots, x^n, \dots\}$ es una base de V .



Las funciones entre espacios vectoriales deben preservar la estructura de los mismos, es decir la suma y la multiplicación por escalares. Surge así el concepto de función lineal.

Definición E.1.6 Sean V y W dos espacios vectoriales sobre un campo \mathbb{F} . Una **función (o transformación u operador) lineal** $F : V \rightarrow W$ es una función que cumple

- $F(\mathbf{u} + \mathbf{v}) = F(\mathbf{u}) + F(\mathbf{v}), \forall \mathbf{u}, \mathbf{v} \in V,$
- $F(r\mathbf{v}) = rF(\mathbf{v}) \quad \forall r \in \mathbb{F} \text{ y } \mathbf{v} \in V.$

Ejemplos

Ej E.1.15 Si $V = \mathbb{R}^n$ y $W = \mathbb{R}^m$ las funciones lineales entre V y W corresponden a las matrices A de $m \times n$.

Ej E.1.16 Si $V = \{f : [0, 1] \rightarrow \mathbb{R} \mid f \text{ es continua}\}$ y $W = \mathbb{R}$, entonces $J : V \rightarrow \mathbb{R}$ definida como $J[f] = \int_0^1 f(x)dx$ es una función lineal.



§E.2 Hiperplanos y teoremas de separación

Dados $\mathbf{x}^t = (x_1, \dots, x_n)$, $\mathbf{y}^t = (y_1, \dots, y_n)$ en \mathbb{R}^n , consideremos el producto punto dado por

$$\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^n x_i y_i.$$

Recordemos que $\mathbf{x} \perp \mathbf{y}$ (\mathbf{x} es ortogonal a \mathbf{y}) si y sólo si $\mathbf{x} \cdot \mathbf{y} = 0$.

Definición E.2.1 Sean $\mathbf{n} \in \mathbb{R}^m$, $\mathbf{n} \neq 0$ y $c \in \mathbb{R}$. Entonces, el **hiperplano** generado por \mathbf{n} y c es el conjunto de puntos en \mathbb{R}^m dado por

$$H(\mathbf{n}, c) = \{\mathbf{x} \in \mathbb{R}^m \mid \mathbf{n} \cdot \mathbf{x} = c\}.$$

Al vector \mathbf{n} se lo conoce como la **normal al hiperplano**. Dado un hiperplano, éste divide al espacio \mathbb{R}^m en dos semiespacios en el sentido de la siguiente definición.

Definición E.2.2 Dado el hiperplano $H(\mathbf{n}, c)$ en \mathbb{R}^m , se definen el semiespacio superior como

$$\{\mathbf{x} \in \mathbb{R}^m \mid \mathbf{n} \cdot \mathbf{x} \geq c\}$$

y el semiespacio inferior como

$$\{\mathbf{x} \in \mathbb{R}^m \mid \mathbf{n} \cdot \mathbf{x} \leq c\}.$$

No es difícil demostrar que tanto el hiperplano $H(\mathbf{n}, c)$ como los semiespacios asociados son conjuntos convexos en \mathbb{R}^n . A continuación veremos algunos ejemplos.

Ejemplos

Ej E.2.1 Si $m = 1$, el hiperplano $H(n, c) = \{x \in \mathbb{R} \mid nx = c\}$ es simplemente el punto $x = \frac{c}{n}$. El semiespacio superior es el intervalo $[\frac{c}{n}, \infty)$ y el semiespacio inferior es $(-\infty, \frac{c}{n}]$.

Ej E.2.2 Si $m = 2$, $\mathbf{n} = (n_1, n_2) \in \mathbb{R}^2$ y $c \in \mathbb{R}$, entonces $H(\mathbf{n}, c) = \{(x, y) \in \mathbb{R}^2 \mid (x, y) \cdot (n_1, n_2) = c\}$, es decir, el hiperplano es el conjunto de puntos que satisfacen

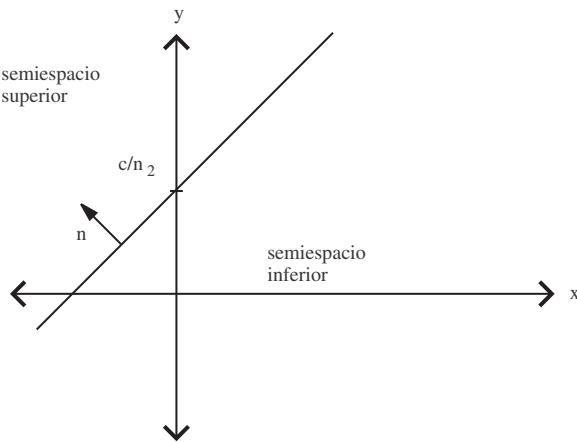
$$n_1 x + n_2 y = c.$$

Esto es simplemente la ecuación de la recta con pendiente $-\frac{n_1}{n_2}$ que pasa por el punto $(0, \frac{c}{n_2})$ (si $n_2 \neq 0$). El semiespacio superior está dado por $\{(x, y) \in \mathbb{R}^2 \mid n_1 x + n_2 y \geq c\}$ y el semiespacio inferior es $\{(x, y) \in \mathbb{R}^2 \mid n_1 x + n_2 y \leq c\}$. La figura E.1 ilustra estas consideraciones.

Ej E.2.3 Si $m = 3$, $\mathbf{n}^t = (n_1, n_2, n_3) \in \mathbb{R}^3$ y $c \in \mathbb{R}$, entonces $H(\mathbf{n}, c) = \{(x, y, z) \in \mathbb{R}^3 \mid (x, y, z) \cdot (n_1, n_2, n_3) = c\}$, es decir, el hiperplano es el conjunto de puntos que satisfacen

$$n_1 x + n_2 y + n_3 z = c.$$

Esto es simplemente el plano con normal \mathbf{n} que pasa por el punto $(0, 0, \frac{c}{n_3})$ (si $n_3 \neq 0$). Los semiespacios superior e inferior están dados por $\{(x, y, z) \in \mathbb{R}^3 \mid n_1 x + n_2 y + n_3 z \geq c\}$ y $\{(x, y, z) \in \mathbb{R}^3 \mid n_1 x + n_2 y + n_3 z \leq c\}$, respectivamente. La figura E.2 ilustra estas consideraciones.

Figura E.1: Hiperplano en \mathbb{R}^2 y semiespacios inferior y superior.

◊◊◊◊◊◊◊◊◊◊

Definición E.2.3 Sea $H(\mathbf{n}, c) \subset \mathbb{R}^m$ un hiperplano y $X, Y \subset \mathbb{R}^m$ dos conjuntos cualesquiera. Decimos que el hiperplano separa a los conjuntos X y Y si X está contenido en el semiespacio superior y Y en el semiespacio inferior, o viceversa. Adicionalmente, decimos que la separación es estricta, si ningún punto de X o Y está en el hiperplano.

El siguiente ejemplo ilustra esta definición.

Ejemplo

Ej E.2.4 Consideremos los siguientes conjuntos en \mathbb{R}^2 : $X = \{(x, y) \in \mathbb{R}^2 \mid y \geq x^2\}$ y $Y = \{(x, y) \in \mathbb{R}^2 \mid y \leq -x^2 - 1\}$. Entonces un hiperplano que separa estrictamente a estos conjuntos es $H(\mathbf{n}, c)$ con $\mathbf{n}^t = (0, 1)$ y $c = -\frac{1}{2}$, es decir, la recta $y = -\frac{1}{2}$. Es claro que en este caso hay una infinidad de hiperplanos que separan estrictamente a los conjuntos dados. La figura E.3 ilustra lo anterior; las líneas punteadas representan algunos otros hiperplanos separadores diferentes de $y = -\frac{1}{2}$.

◊◊◊◊◊◊◊◊◊◊

Lema E.2.4 Sea $A \subset \mathbb{R}^n$ un conjunto convexo. Entonces \bar{A} y A° son convexos.

La demostración del lema E.2.4 se deja como ejercicio al lector.

Recordemos que siempre se puede definir una distancia a partir de una norma (véase la proposición D.2.3). En particular, dada la norma inducida por el producto punto en \mathbb{R}^n se tiene que para todo $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$,

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\| = \sqrt{(\mathbf{x} - \mathbf{y}) \cdot (\mathbf{x} - \mathbf{y})}.$$

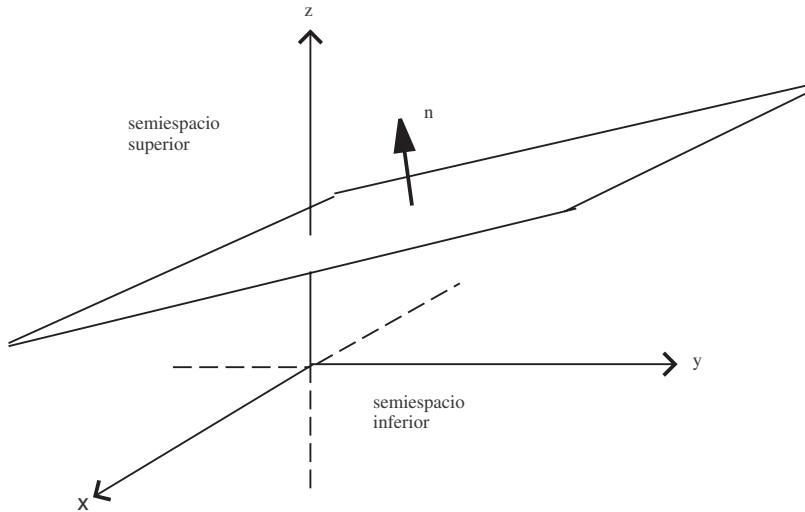


Figura E.2: Hiperplano en \mathbb{R}^3 y semiespacios inferior y superior.

En todo lo que sigue se utiliza este concepto de distancia, que coincide con la métrica euclídea; sin embargo, todos los resultados pueden extenderse a espacios y métricas más generales.

Lema E.2.5 *Sea $A \subset \mathbb{R}^m$ cerrado y convexo y sea $\mathbf{b} \notin A$. Entonces existe $\mathbf{a} \in A$ tal que*

$$d(\mathbf{a}, \mathbf{b}) = \min\{d(\mathbf{x}, \mathbf{b}) \mid \mathbf{x} \in A\}.$$

Podemos definir así la distancia entre A y \mathbf{b} como $d(A, \mathbf{b})$.

Demostración

Sean $\mathbf{x} \in A$ y $r = d(\mathbf{x}, \mathbf{b})$. Recordemos que $\overline{B_r(\mathbf{b})} = \{\mathbf{y} \mid d(\mathbf{y}, \mathbf{b}) \leq r\}$ es un conjunto cerrado y acotado, de manera que el conjunto

$$B = \overline{B_r(\mathbf{b})} \cap A$$

es cerrado y acotado (recordar el ejercicio D.9). La función distancia $d(\cdot, \mathbf{b}) : B \rightarrow \mathbb{R}_+$ es una función continua sobre el compacto B , de manera que existe $\mathbf{a} \in B$ tal que $d(\mathbf{a}, \mathbf{b}) = \min\{d(\mathbf{y}, \mathbf{b}) \mid \mathbf{y} \in B\}$ ya que, por el teorema del valor extremo, la función alcanza su máximo y mínimo en B . Dado $\mathbf{y} \in A \cap B^c$ se tiene que $d(\mathbf{y}, \mathbf{b}) > r > d(\mathbf{a}, \mathbf{b})$ y por lo tanto $d(\mathbf{a}, \mathbf{b}) = \min\{d(\mathbf{y}, \mathbf{b}) \mid \mathbf{y} \in A\}$. Finalmente, falta probar que el punto \mathbf{a} es único. Supongamos que existe \mathbf{a}' de manera que $d(\mathbf{a}', \mathbf{b}) = d(\mathbf{a}, \mathbf{b})$; como A es convexo, se tiene que, dado $\lambda \in (0, 1)$, el punto $\mathbf{p} = \lambda\mathbf{a} + (1 - \lambda)\mathbf{a}'$ pertenece a A . En este caso, se tiene que $d(\mathbf{p}, \mathbf{b}) < d(\mathbf{a}, \mathbf{b})$ (esta desigualdad se deja como ejercicio al lector, véase el ejercicio E.10) lo cual es una contradicción y por lo tanto \mathbf{a} es único, con lo que queda demostrado el lema. ■

Teorema E.2.6

Sea $A \subset \mathbb{R}^m$ cerrado y convexo, $A \neq \emptyset$, y sea $\mathbf{b} \notin A$ un punto en A^c . Entonces se cumplen:

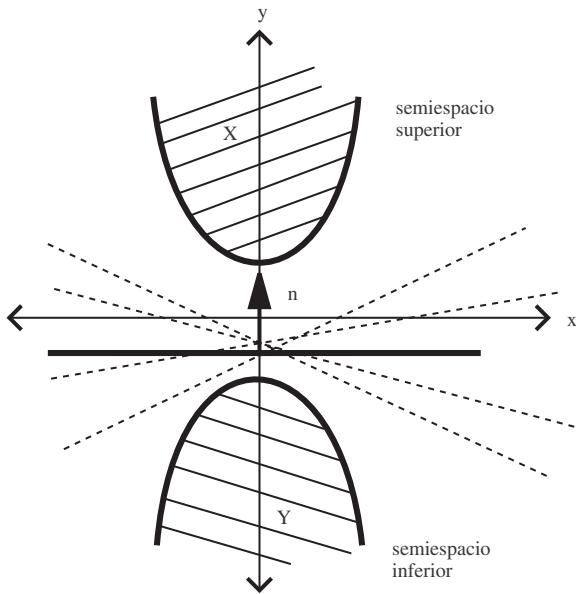


Figura E.3: Hiperplano $y = -\frac{1}{2}$ separando a X y Y . Las líneas punzadas muestran otros hiperplanos separadores posibles.

- a) Existen un punto $\mathbf{a} \in A$ y un hiperplano $H(\mathbf{n}, c)$ que pasa por \mathbf{a} tales que

$$\begin{aligned}\mathbf{n} \cdot \mathbf{b} &< c, \\ \mathbf{n} \cdot \mathbf{a} &= c = \inf\{\mathbf{n} \cdot \mathbf{x} \mid \mathbf{x} \in A\}.\end{aligned}$$

En este caso decimos que $H(\mathbf{n}, c)$ es un **hiperplano soporte** del conjunto A .

- b) Existe un segundo hiperplano $H(\mathbf{n}, d)$, con la misma normal, que separa estrictamente a A y $\{\mathbf{b}\}$, es decir, se cumple

$$\mathbf{n} \cdot \mathbf{b} < d < \mathbf{n} \cdot \mathbf{x}, \forall \mathbf{x} \in A.$$

Demostración

- a) Sea $\mathbf{a} \in A$ como en el lema E.2.5. Definamos $\mathbf{n} = \mathbf{a} - \mathbf{b}$ y $c = \mathbf{n} \cdot \mathbf{a}$; demostraremos que $H(\mathbf{n}, c)$ es el hiperplano buscado. Notemos que este hiperplano es normal al segmento que une a \mathbf{a} y \mathbf{b} . Observemos que $\mathbf{n} \cdot \mathbf{b} = \mathbf{a} \cdot \mathbf{b} - \mathbf{b} \cdot \mathbf{b}$ y que $c = \mathbf{a} \cdot \mathbf{a} - \mathbf{b} \cdot \mathbf{a}$, de manera que $\mathbf{n} \cdot \mathbf{b} - c = -(\mathbf{a} - \mathbf{b}) \cdot (\mathbf{a} - \mathbf{b}) = -\|\mathbf{a} - \mathbf{b}\|^2 < 0$ y por lo tanto $\mathbf{n} \cdot \mathbf{b} < c$. Por definición de \mathbf{n} y c , se tiene que $\mathbf{n} \cdot \mathbf{a} = c$. Finalmente, tenemos que ver que, para todo $\mathbf{x} \in A$, se cumple $\mathbf{n} \cdot \mathbf{x} \geq \mathbf{n} \cdot \mathbf{a}$ y que $\mathbf{n} \cdot \mathbf{a}$ es la máxima cota inferior, de manera que $\mathbf{n} \cdot \mathbf{a} = \inf\{\mathbf{n} \cdot \mathbf{x} \mid \mathbf{x} \in A\}$. En efecto, supongamos que para algún $\mathbf{x} \in A$ se tiene $\mathbf{n} \cdot \mathbf{x} < \mathbf{n} \cdot \mathbf{a}$ y sea $\mathbf{z} = \lambda \mathbf{x} + (1 - \lambda) \mathbf{a}$ con $\lambda \in (0, 1)$.

Dado que A es convexo se tiene que $\mathbf{z} \in A$; adicionalmente puede verse que

$$\begin{aligned}\|\mathbf{a} - \mathbf{b}\|^2 - \|\mathbf{b} - \mathbf{z}\|^2 &= \|\mathbf{a} - \mathbf{b}\|^2 - (\mathbf{b} - \lambda\mathbf{x} - \mathbf{a} + \lambda\mathbf{a}) \cdot (\mathbf{b} - \lambda\mathbf{x} - \mathbf{a} + \lambda\mathbf{a}) \\ &= \|\mathbf{a} - \mathbf{b}\|^2 - \|\mathbf{a} - \mathbf{b}\|^2 - 2\mathbf{n} \cdot (\mathbf{x} - \mathbf{a}) - \lambda^2 \|\mathbf{x} - \mathbf{a}\|^2 \\ &= -\lambda[2\mathbf{n} \cdot (\mathbf{x} - \mathbf{a}) + \lambda \|\mathbf{x} - \mathbf{a}\|^2] \\ &= -\lambda[2(\mathbf{n} \cdot \mathbf{x} - \mathbf{n} \cdot \mathbf{a}) + \lambda \|\mathbf{x} - \mathbf{a}\|^2].\end{aligned}$$

Utilizando la relación $\mathbf{n} \cdot \mathbf{x} < \mathbf{n} \cdot \mathbf{a}$ se obtiene que $(\mathbf{n} \cdot \mathbf{x} - \mathbf{n} \cdot \mathbf{a}) < 0$ de manera que si λ es suficientemente pequeña, o bien, concretamente,

$$\lambda < \frac{-2(\mathbf{n} \cdot \mathbf{x} - \mathbf{n} \cdot \mathbf{a})}{\|\mathbf{y} - \mathbf{a}\|^2},$$

la última expresión es positiva y tenemos que se cumple

$$\begin{aligned}\|\mathbf{a} - \mathbf{b}\|^2 &> \|\mathbf{b} - \mathbf{z}\|^2 \Rightarrow \\ d(\mathbf{a}, \mathbf{b}) &> d(\mathbf{b}, \mathbf{z}).\end{aligned}$$

Esto último es una contradicción, ya que $d(\mathbf{a}, \mathbf{b}) = \min\{d(\mathbf{x}, \mathbf{b}) \mid \mathbf{x} \in A\}$ y por lo tanto para todo $\mathbf{x} \in A$ se debe cumplir $\mathbf{n} \cdot \mathbf{x} \geq \mathbf{n} \cdot \mathbf{a}$. Se deja al lector verificar que efectivamente $\mathbf{n} \cdot \mathbf{a}$ es la máxima cota superior, de manera que $\mathbf{n} \cdot \mathbf{a} = \inf\{\mathbf{n} \cdot \mathbf{x} \mid \mathbf{x} \in A\}$, con lo que se concluye la demostración. La figura E.4 ilustra esta parte del teorema.

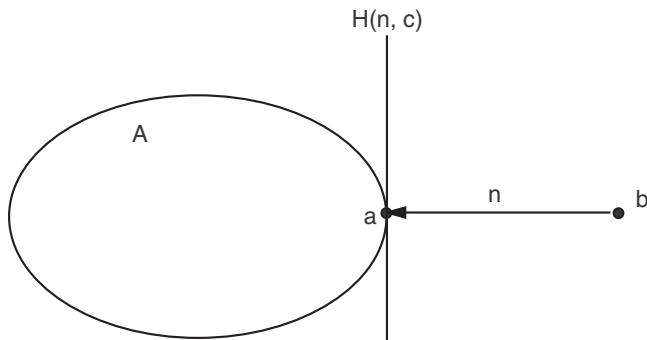


Figura E.4: $H(\mathbf{n}, c)$ es el hiperplano soporte del conjunto A que pasa por el punto \mathbf{a} .

- b) Sea \mathbf{n} como en el inciso anterior y sea $d = \mathbf{n} \cdot (\frac{1}{2}\mathbf{a} + \frac{1}{2}\mathbf{b})$. Entonces el hiperplano $H(\mathbf{n}, d)$ separa estrictamente a A y $\{\mathbf{b}\}$, ya que es el hiperplano con normal \mathbf{n} que pasa por el punto medio entre \mathbf{a} y \mathbf{b} .

Concluimos así la demostración. ■

Teorema E.2.7

Sea $A \subset \mathbb{R}^m$ convexo, $A \neq \emptyset$. Entonces

- a) si $\mathbf{b} \in A^c$, existe un hiperplano $H(\mathbf{p}, c)$ que separa a A y $\{\mathbf{b}\}$.
- b) Si \mathbf{b} es un punto frontera de A , entonces existe un hiperplano soporte de A que pasa por \mathbf{b} .

Demostración

- a) Sea $\mathbf{b} \in A^c$. Por el lema E.2.4, el conjunto \bar{A} es convexo. Entonces, si $\mathbf{b} \notin \bar{A}$, aplicamos el teorema anterior para \bar{A} y $\{\mathbf{b}\}$. Si $\mathbf{b} \in \bar{A}$, entonces \mathbf{b} es un punto frontera de A . Se deja al lector probar que A y \bar{A} tienen los mismos puntos frontera, de forma que \mathbf{b} es punto frontera de \bar{A} . Para cualquier $n > 0$, se tiene que $B_{\frac{1}{n}}(\mathbf{b}) \cap \bar{A}^c \neq \emptyset$. Sea $\mathbf{b}_n \in B_{\frac{1}{n}}(\mathbf{b})$. Entonces la sucesión $\{\mathbf{b}_n\}$ es una sucesión de puntos en \bar{A}^c que converge a \mathbf{b} . Por el teorema anterior se cumple

$$\begin{aligned} \mathbf{b}_n \notin \bar{A} \Rightarrow \\ \exists \mathbf{p}_n \text{ tal que } \mathbf{p}_n \cdot \mathbf{b}_n < \mathbf{p}_n \cdot \mathbf{x}, \forall \mathbf{x} \in \bar{A}. \end{aligned}$$

Sin pérdida de generalidad, podemos tomar $\|\mathbf{p}_n\| = 1$ para toda n . La sucesión $\{\mathbf{p}_n\}$ está acotada y, por lo tanto, el teorema D.3.10, nos dice que posee una subsucesión convergente, digamos $\{\mathbf{p}_{n_k}\}$. Sea $\mathbf{p} = \lim_{n_k \rightarrow \infty} \mathbf{p}_{n_k}$. Dado que para esta subsucesión se cumple

$$\mathbf{p}_{n_k} \cdot \mathbf{b}_{n_k} < \mathbf{p}_{n_k} \cdot \mathbf{x}, \forall \mathbf{x} \in \bar{A},$$

tomando límites, definiendo $c = \mathbf{p} \cdot \mathbf{b}$ y dado que $A \subset \bar{A}$, se tiene

$$c = \mathbf{p} \cdot \mathbf{b} \leq \mathbf{p} \cdot \mathbf{x}, \forall \mathbf{x} \in A,$$

de manera que el hiperplano $H(\mathbf{p}, c)$ separa a A y $\{\mathbf{b}\}$.

- b) Sea \mathbf{b} un punto frontera de A . Si $\mathbf{b} \notin A$, entonces el resultado se obtiene directamente por la primera parte del teorema. Si $\mathbf{b} \in A$, entonces sabemos que también es punto frontera de \bar{A} y utilizamos un argumento idéntico al del inciso anterior, construyendo la sucesión correspondiente. Se deja esta parte como ejercicio al lector. ■

Teorema E.2.8 (Del hiperplano separador)

Sean $A, B \subset \mathbb{R}^m$, no vacíos y convexos, tales que $A \cap B = \emptyset$. Entonces existen $\mathbf{n} \in \mathbb{R}^m$, $\mathbf{n} \neq \mathbf{0}$, y $c \in \mathbb{R}$ tales que se cumplen

$$\mathbf{n} \cdot \mathbf{y} \leq c \leq \mathbf{n} \cdot \mathbf{x}, \forall \mathbf{x} \in A, \forall \mathbf{y} \in B.$$

Es decir, existe un hiperplano $H(\mathbf{n}, c)$ que separa a A y B .

Demostración

Por la proposición 10.1.2, el conjunto $C = A + (-B) = A - B$, es convexo y $\mathbf{0} \notin C$ puesto que $A \cap B = \emptyset$. Utilizando el teorema anterior, existe un hiperplano $H(\mathbf{n}, c)$ que separa a C y $\{\mathbf{0}\}$, de manera que se cumple

$$0 = \mathbf{n} \cdot \mathbf{0} \leq \mathbf{n} \cdot \mathbf{z}, \forall \mathbf{z} \in C.$$

De aquí se tiene que

$$0 \leq \mathbf{n} \cdot (\mathbf{x} - \mathbf{y}), \forall \mathbf{x} \in A, \mathbf{y} \in B,$$

es decir,

$$\mathbf{n} \cdot \mathbf{y} \leq \mathbf{n} \cdot \mathbf{x}, \forall \mathbf{x} \in A, \mathbf{y} \in B.$$

El conjunto de números reales $\{\mathbf{n} \cdot \mathbf{y} \mid \mathbf{y} \in B\}$ está acotado por arriba (por $\mathbf{n} \cdot \mathbf{x}$, para cualquier $\mathbf{x} \in A$) de manera que posee un supremo, digamos c . Se tiene así finalmente que

$$\mathbf{n} \cdot \mathbf{y} \leq c \leq \mathbf{n} \cdot \mathbf{x}, \forall \mathbf{x} \in A, \mathbf{y} \in B,$$

o bien, el hiperplano $H(\mathbf{n}, c)$ separa a A y B , con lo que damos por terminada la demostración. ■

Ejemplo

Ej E.2.5 Sean x y y dos bienes, $y u(x, y)$ la función de utilidad, creciente y cuasicóncava, de la sociedad. Sean $x = f(K_x, L_x)$ y $y = g(K_y, L_y)$ las funciones de producción, crecientes y cuasicóncavas. Aquí K_x y K_y son el capital utilizado en la industria correspondiente, y L_x y L_y se refieren a la mano de obra. Si \bar{K} y \bar{L} son las cantidades totales de estos recursos, entonces obtenemos la frontera de posibilidades de producción como

$$\{(f(K_x, L_x), g(K_y, L_y)) \mid K_x + K_y = \bar{K}, L_x + L_y = \bar{L}\}.$$

Sea R el conjunto de posibilidades de producción delimitado por esta frontera y los ejes coordenados. Tanto el conjunto R , como los contornos superiores $CS_u(\bar{u})$ de la función u para cualquier nivel de utilidad fijo \bar{u} , son conjuntos convexos y cerrados. Notemos que el punto óptimo (x^*, y^*) es un punto frontera de R y de $CS_u(\bar{u})$ y, si p_x y p_y son los precios de equilibrio, $\mathbf{n} = (p_x, p_y)$ y $c = \mathbf{n} \cdot (x^*, y^*)$, el hiperplano $H(\mathbf{n}, c)$ es un hiperplano soporte de ambos conjuntos. La existencia del vector de precios equivale a la existencia del hiperplano soporte. La figura E.5 ilustra estas consideraciones.



El ejemplo anterior ilustra la importancia que tiene la hipótesis de cuasiconcavidad para las funciones de utilidad y de producción. Remitimos al lector a la obra clásica de Debreu [Deb86] para un tratamiento cuidadoso sobre este tema.

Ejercicios

▷ **E.1** Probar que los ejemplos E.1.1, E.1.2, E.1.3 y E.1.4 efectivamente cumplen las propiedades de espacio vectorial.

▷ **E.2** Sean U_1 y U_2 subespacios del espacio vectorial V .

a) Probar que $U_1 \cap U_2$ también es un subespacio de V .

b) Mostrar con un contraejemplo que $U_1 \cup U_2$ no necesariamente es subespacio de V .

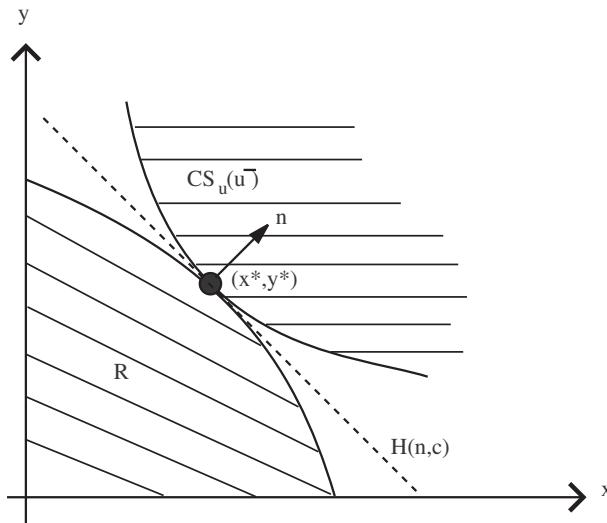


Figura E.5: Hiperplano soporte determinado por el vector de precios.

▷ **E.3** Sea V el conjunto de matrices reales de $n \times n$. Probar que el conjunto $U = \{A \in V \mid AB = BA$ para una matriz B dada} es un subespacio de V .

▷ **E.4** Probar que el conjunto S dado en el ejemplo E.1.10 es un subespacio de \mathbb{R}^n .

▷ **E.5** Sea $B = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ una base de un espacio vectorial V . Probar que:

a) Dado $\mathbf{w} \in V$, $\mathbf{w} \notin B$, entonces $\{\mathbf{v}_1, \dots, \mathbf{v}_n, \mathbf{w}\}$ es linealmente dependiente.

b) Los conjuntos $B_{-i} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\} - \{\mathbf{v}_i\}$, $i = 1, \dots, n$ no generan V .

▷ **E.6** Considerar el producto $f \cdot g = \int_0^1 f(x)g(x)dx$ definido para el espacio del ejemplo E.1.6.

a) Probar que este producto es bilineal.

b) Probar que las funciones $f(x) = x^2$ y $g(x) = 4x - 3$ son ortogonales.

▷ **E.7** Probar el lema E.2.4.

▷ **E.8** Probar que, dado $A \in \mathbb{R}^n$, A y \bar{A} tienen los mismos puntos frontera.

▷ **E.9** Probar la segunda parte del teorema E.2.7.

▷ **E.10** Probar la desigualdad $d(\mathbf{p}, \mathbf{b}) < d(\mathbf{a}, \mathbf{b})$ del lema E.2.5. (Sugerencia: utilizar el hecho de que se trata de la distancia euclídea y considerar el triángulo isósceles con lados $A = B$ de longitud $d(\mathbf{a}, \mathbf{b})$ y C de longitud $d(\mathbf{a}, \mathbf{a}')$.)

▷ **E.11 (Kuhn-Tucker)** Recordemos del capítulo 10 que, dado el problema de optimización

$$\begin{aligned} \max \quad & f(\mathbf{x}) \\ \text{sujeto a} \quad & g_1(\mathbf{x}) \leq 0, \\ & \vdots \\ & g_m(\mathbf{x}) \leq 0, \end{aligned}$$

donde $x \in X \subset \mathbb{R}^n$, X convexo, y las funciones f, g_1, \dots, g_m son de clase \mathcal{C}^1 . Si \mathbf{x}^* resuelve este problema de maximización, entonces satisface las condiciones 10.2 a 10.5, conocidas como las condiciones de Kuhn-Tucker. Si A es el conjunto de restricciones activas en \mathbf{x}^* , es decir, $g_j(\mathbf{x}^*) = 0 \iff j \in A$, las condiciones de K-T nos dicen esencialmente que el vector gradiente $\nabla f(\mathbf{x}^*)$ es una combinación lineal no negativa del conjunto de vectores gradiente $\{\nabla g_j(\mathbf{x}^*)\}_A$. Podemos expresar esto último como

$$\nabla f(\mathbf{x}^*) \in \text{cone}^+ \{\nabla g_j(\mathbf{x}^*)\}_A,$$

donde $\text{cone}^+ \{v_j\}_A$ es el conjunto de combinaciones lineales positivas del conjunto de vectores $\{v_j : j \in A\}$. Proceder de la siguiente manera para probar esta aseveración:

- a) Probar que $\text{cone}^+ \{v_j\}_A$ es un conjunto convexo.
- b) Suponer que $\nabla f(\mathbf{x}^*) \notin \text{cone}^+ \{\nabla g_j(\mathbf{x}^*)\}_A$ y probar que existen $\mathbf{n} \in \mathbb{R}^n$ y $c \in \mathbb{R}$ tales que se cumplen

$$\begin{aligned} \nabla f(\mathbf{x}^*) \cdot \mathbf{n} &> c, \\ \mathbf{y} \cdot \mathbf{n} &< c, \quad \forall \mathbf{y} \in \text{cone}^+ \{\nabla g_j(\mathbf{x}^*)\}_A. \end{aligned}$$

En otras palabras probar que existe un hiperplano que separa estrictamente a $\nabla f(\mathbf{x}^*)$ y a $\text{cone}^+ \{\nabla g_j(\mathbf{x}^*)\}_A$. Verificar que, sin pérdida de generalidad, se puede tomar $\|\mathbf{n}\| = 1$.

- c) Probar que $\nabla f(\mathbf{x}^*) \cdot \mathbf{n} > 0$. (Sugerencia: probar que $c > 0$.)
- d) Probar que $\nabla g_j(\mathbf{x}^*) \cdot \mathbf{n} < 0$ para toda $j = 1, \dots, m$. (Sugerencia: utilizar la siguiente propiedad del cono: $r\mathbf{y} \cdot \mathbf{n} < c \quad \forall \mathbf{y} \in \text{cone}^+ \{\nabla g_j(\mathbf{x}^*)\}_A, r \in \mathbb{R}_+$.)
- e) Probar que, evaluando en \mathbf{x}^* , en la dirección del vector \mathbf{n} , la función f es creciente y las restricciones g_j son decrecientes. Llegar con esto a una contradicción y concluir que $\nabla f(\mathbf{x}^*) \in \text{cone}^+ \{\nabla g_j(\mathbf{x}^*)\}_A$, que es lo que se quería demostrar. (Sugerencia: recordar que $\nabla f(\mathbf{x}^*) \cdot \mathbf{n}$ es la derivada direccional de f en la dirección \mathbf{n} en el punto dado y que $\nabla g_j(\mathbf{x}^*) \cdot \mathbf{n}$ lo es para las restricciones.)

BIBLIOGRAFÍA

- [ASY96] K. Aligood, T. Sauer, and J. Yorke. *Chaos*. Springer Verlag, New York, NY, 1996.
- [Bar90] R. Barro. Government spending in a simple model of endogenous growth. *Journal of Political Economy*, 98:103–125, 1990.
- [BD62] Richard E. Bellman and Stuart E. Dreyfus. *Applied dynamic programming*. Princeton University Press, Princeton, N.J., 1962.
- [Bel57] Richard Bellman. *Dynamic programming*. Princeton Univeristy Press, Princeton, N. J., 1957.
- [BF89] O. Blanchard and S. Fischer. *Lectures on Macroeconomics*. MIT Press, Cambridge, Massachusetts, 1st edition, 1989.
- [BF92] Olivier J. Blanchard and Stanley Fischer. *Lectures on Macroeconomics*. MIT Press, Cambridge, Ma, 1992.
- [BK65] Richard Bellman and Robert Kalaba. *Dynamic programming and modern control theory*. Academic Press, New York, 1965.
- [Bla85] O. Blanchard. Methods of solution for dynamic rational survey. *Mathematical Programming Study*, 23:210–225, 1985.
- [Bla92] Mark Blaug. *The Methodology of Economics or how economists explain*. Cambridge University Press, New York, NY, 1992.
- [BM72] William A. Brock and Leonard J. Mirman. Optimal economic growth and uncertainty: the discounted case. *J. Econom. Theory*, 4(3), 1972.
- [Bor97] J. L. Borges. *El hacedor*. Alianza Editorial, Madrid, 1997.

- [Bre79] D. Breeden. An intertemporal asset pricing model with stochastic consumption and investment opportunities. *Journal of Financial Economics*, 7:265–296, 1979.
- [BS94] L. Blume and C. Simon. *Mathematics for Economists*. W.W. Norton & Company Inc., New York, NY, 1st edition, 1994.
- [Cag56] Phillip Cagan. The monetary dynamics of hyperinflation. In Milton Friedman, editor, *Studies in the quantity Theory of Money*. University of Chicago Press, Chicago, IL, 1956.
- [Cal81] Guillermo Calvo. Devaluation: Levels vs rates. *Journal of International Economics*, 11:165–172, 1981.
- [CH97] R. Clower and P. Howitt. Foundations of economics. In A. D'Autume and J. Cartelier, editors, *Is Economics Becoming a Hard Science?*, pages 17–34. Edward Elgar, Cheltenham, UK, 1997.
- [Chi84] Alpha C. Chiang. *Fundamental Methods of Mathematical Economics*. McGraw-Hill, New York, NY, 1984.
- [Chi92] Alpha C. Chiang. *Elements of Dynamic Optimization*. McGraw-Hill, New York, NY, 1992.
- [CK95] David E. Card and B. Krueger, Alan. *Myth and Measurement: The New Economics of Minimum Wage*. Princeton University Press, Princeton, NJ, 1995.
- [CM85] Mark Cross and A. O. Moscardini. *Learning the art of Mathematical modelling*. Ellis Horwood Ltd., Chichester, 1985.
- [Col99] K. C. Cole. *El universo y la taza de té*. SineQuaNon, ediciones grupo Zeta, Barcelona, 1999. Título original: The universe and the teacup.
- [Day94] Richard H. Day. *Complex economic dynamics. Vol. I*. MIT Press, Cambridge, MA, 1994. An introduction to dynamical systems and market mechanisms, With a foreword by Paul A. Samuelson.
- [Deb73] G. Debreu. *Teoría del Valor*. Bosch, Casa Editorial, Barcelona, 1973. Del original de 1959 editado por la Cowles Foundation, Yale University.
- [Deb86] G. Debreu. *Theory of Value: an Axiomatic Analysis of Economic Equilibrium*. Yale University Press, New Haven, CT, 1986.
- [Dev89] R. Devaney. *An Introduction to Chaotic Dynamical Systems*. Addison Wesley, 1989.
- [dlF00] Angel de la Fuente. *Mathematical Methods and models for Economists*. Cambridge University Press, Cambridge, UK, 2000.
- [Dom46] Evsey D. Domar. Capital expansion, rate of growth and employment. *Econometrica*, 14:137–147, April 1946.

- [Dor69] R. Dorfman. An economic interpretation of optimal control theory. *American Economic Review*, 59:817–831, 1969.
- [Dor76] Rudiger Dornbusch. Expectations and exchange rate dynamics. *Journal of Political Economy*, pages 1161–1176, 1976.
- [Dra92] P. G. Drazin. *Nonlinear systems*. Cambridge University Press, Cambridge, 1992.
- [ES63] R. Eisner and R. Strotz. Determinants of business investment. In *Impacts of Monetary policy*, pages 60–233. Prentice-Hall, Englewood Cliffs, NJ, 1963.
- [Eva24] G. Evans. The dynamics of monopoly. *American Mathematical Monthly*, 31(2):77–83, 1924.
- [Fis77] S. Fischer. Long term contracts, rational expectations, and the optimal money supply rule. *Journal of Political Economy*, 85:191–205, 1977.
- [Fra89] Philipp Frank. *Einstein : his life and times*. Da Capo Press, New York, NY, 1989. Reimpresión del original de Knopf, 1947.
- [Fri56] M. Friedman. *Studies in the Quantity Theory of Money*. University of Chicago Press, Chicago, IL, 1956.
- [Goo91] R. Goodwin. *Chaotic Economic Dynamics*. Oxford University Press, Oxford, UK, 1991.
- [Haa56] T. Haavelmo. *A Study in the theory of Economic evolution*. North-Holland, Amsterdam, 2nd edition, 1956.
- [Hal74] H. Halkin. Necessary conditions for optimal control problems with infinite horizons. *Econometrica*, pages 267–272, 1974.
- [Hal78] R. Hall. Stochastic implications of the life cycle permanent income hypothesis: Theory and evidence. *Journal of Political Economy*, 86(5):971–987, 1978.
- [Har39] Roy F. Harrod. An essay in dynamic theory. *Economic Journal*, 49:14–33, June 1939.
- [Hes66] Magnus R. Hestenes. *Calculus of variations and optimal control theory*. John Wiley & Sons Inc., New York, 1966.
- [Hic37] J. Hicks. Mr. Keynes and the classics: a suggested interpretation. *Econometrica*, 1937. reprinted in Hicks, J. *Critical essays in Monetary Theory*, Oxford University press, 1967.
- [Jor67] D.W. Jorgenson. The theory of investment behavior. In R. Ferber, editor, *Determinants of Investment Behavior*, pages 129–155. Columbia University Press, New York, NY, 1967.
- [Kru98] Paul Krugman. Dos aplausos para el formalismo. *Gaceta de Economía*, 4(7):235–246, 1998.

- [KS91] Morton I. Kamien and Nancy L. Schwartz. *Dynamic Optimization: The Calculus of Variations and Optimal Control in Economics and Management*. Elsevier, New York, NY, 1991.
- [Lin65] J. Lintner. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics*, 47:13–37, 1965.
- [LM75] Charles Lave and James March. *An Introduction to Models in the Social Sciences*. Harper & Row, New York, NY, 1975.
- [Lor93] Hans-Walter Lorenz. *Nonlinear dynamical economics and chaotic motion*. Springer-Verlag, Berlin, second edition, 1993.
- [Luc77] R. Lucas. Understanding business cycles. In K. Brunner and A. Meltzer, editors, *Stabilization of the Domestic and International Economy*, pages 7–29. North Holland, Amsterdam, 1977.
- [Luc78] Robert E. Lucas, Jr. Asset prices in an exchange economy. *Econometrica*, 46(6):1429–1445, 1978.
- [Lue84] D.G. Luenberger. *Linear and Nonlinear Programming*. Addison Wesley, Reading, MA, second edition, 1984.
- [LY75] Tien Yien Li and James A. Yorke. Period three implies chaos. *Amer. Math. Monthly*, 82(10):985–992, 1975.
- [May76] R. May. Simple mathematical models with very complicated dynamics. *Nature*, 261:459, 1976.
- [MCWG95] Andreu Mas-Colell, Michael D. Whinston, and Jerry R. Green. *Microeconomic Theory*. Oxford University Press, Oxford, UK, 1995.
- [Mer69] R. Merton. Lifetime portfolio selection under uncertainty: The continuous time case. *Review of Economics and Statistics*, 51:247–257, 1969.
- [Mer73] Robert C. Merton. An intertemporal capital asset pricing model. *Econometrica*, 41(5):867–887, 1973.
- [Mos00] Jesús Mosterín. *Conceptos y Teorías en la ciencia*. Alianza, Madrid, 2000.
- [MS71] John Maynard Smith. *Mathematical Ideas in Biology*. Cambridge University Press, Cambridge, UK, 1971.
- [MS86] N. Mankiw and M. Shapiro. Risk and return, consumption beta versus market beta. *Review of Economics and Statistics*, 68(3):452–459, 1986.
- [PBGM62] L.S. Pontryagin, V.G. Boltyansky, R.V. Gamkrelidze, and E.F. Mishchenko. *The Mathematical Theory of Optimal Processes*. Wiley-Interscience, New York, NY, 1962.

- [Phi58] A.W. Phillips. The relation between employment and the rate of change of money wages in the united kingdom 1861-1957. *Economica*, 25:283–299, 1958.
- [PM92] M.H. Protter and C.B. Morrey. *A First Course in Real Analysis*. Undergraduate Texts in Mathematics, Springer Verlag, New York, NY, 1992.
- [Puu00] Tonu Puu. *Attractors, bifurcations and Chaos. Nonlinear Phenomena in Economics*. Springer-Verlag, 2000.
- [Ram28] Frank Ramsey. A mathematical theory of savings. *Economic Journal*, 38:543–559, 1928.
- [Rob99] Clark Robinson. *Dynamical systems*. CRC Press, Boca Raton, FL, second edition, 1999. Stability, symbolic dynamics, and chaos.
- [Rue91] David Ruelle. *Chance and chaos*. Princeton University Press, Princeton, NJ, 1991.
- [Saa95] Donald Saari. Mathematical complexity of simple economics. *Notices Amer. Math. Soc.*, 42(2):222–230, 1995.
- [Saa96] Donald G. Saari. The ease of generating chaotic behavior in economics. *Chaos Solitons Fractals*, 7(12):2267–2278, 1996. Complex dynamics in economic and social systems (Umeå, 1995).
- [Sar87] T. Sargent. *Macroeconomic Theory*. Academic Press, San Diego, Ca, 1987.
- [SB99] Alan Sokal and Jean Bricmont. *Imposturas Intelectuales*. Paidós, Barcelona, 1999.
- [Sch96] Edward R. Scheinerman. *Invitation to Dynamical Systems*. Prentice-Hall, first edition, 1996.
- [Sha64] W. Sharpe. Capital asset prices: a theory of market equilibrium under conditions of risk. *Journal of Finance*, 19:425–442, 1964.
- [Sho97] Ronald Shone. *Economic Dynamics*. Cambridge University Press, Cambridge, UK, 1997.
- [Sid67] M. Sidrauski. Rational choice and patterns of growth in a monetary economy. *American Economic Review*, 57(2):534–544, 1967.
- [SiM94] Xavier Sala-i Martín. *Apuntes de Crecimiento Económico*. Antoni Bosch, Barcelona, 1994.
- [SL89] Nancy L. Stokey and Robert E. Lucas, Jr. *Recursive methods in economic dynamics*. Harvard University Press, Cambridge, MA, 1989. With the collaboration of Edward C. Prescott.
- [Sma76] Steve Smale. Dynamics in general equilibrium theory. *Amer. Economic Review*, 66:288–294, 1976.
- [Sma98] Steve Smale. Mathematical problems for the next century. *Mathematical Intelligencer*, 20(2):7–15, 1998.

- [Sol56] Robert M. Solow. A contribution to the theory of economic growth. *Quarterly Journal of Economics*, 70:65–94, February 1956.
- [SS87] Atle Seierstad and Knut Sydsæter. *Optimal control theory with economic applications*. North-Holland Publishing Co., Amsterdam, 1987.
- [Str94] S. H. Strogatz. *Nonlinear dynamics and chaos: With applications to biology, chemistry, and engineering*. Addison-Wesley, Reading, MA, 1994.
- [Sun96] K. Sundaram, Rangarajan. *A First Course in Optimization Theory*. Cambridge University Press, Cambridge, UK, 1996.
- [SW73] T. Sargent and N. Wallace. The stability of models of money and perfect foresight. *Econometrica*, 41:1043–1048, 1973.
- [Swa56] Trevor W. Swan. Economic growth and capital accumulation. *Economic Record*, 32:334–361, November 1956.
- [Tak85] A. Takayama. *Mathematical Economics*. Cambridge University Press, Cambridge, UK, second edition, 1985.
- [Var92] Hal R. Varian. *Microeconomic Analysis*. W.W. Norton & Company, New York, NY, 1992.
- [Vel96] Roberto Velasco. *Los economistas en su laberinto*. Taurus, Madrid, 1996.