



GRADO EN ESTADÍSTICA

TRABAJO FIN DE GRADO

Programación no lineal

Paula Gutiérrez Sánchez

Sevilla, junio de 2022

Índice general

Prólogo	III
Resumen	V
Abstract	VI
Índice de Figuras	VII
Índice de Tablas	IX
1. Introducción	1
1.1. Investigación operativa	1
1.2. Programación no lineal	1
1.3. Conceptos basicos de programación no lineal	2
1.4. Planteamientos de problemas	3
2. Conjuntos convexos	5
2.1. Convexidad	5
2.1.1. Definición conjunto convexo	5
2.1.2. Definición punto extremo	6
2.1.3. Definición de dirección	7
2.1.4. Definición dirección extrema	7
2.1.5. Teorema de caracterización de conjuntos poliédricos	7
2.1.6. Definición cono convexo	7
2.1.7. Ejemplos de conjuntos convexos	7
2.2. Envolverte convexos	8
2.2.1. Teorema de Carathéodory	9
2.2.2. Definición conjunto arbitrario	10
2.2.3. Definición poliedro y simplex	11
2.2.4. Teoremas tipo Helly	12
2.2.4.1. Teorema de Radon	13
2.2.4.2. Teorema de Helly	13

3. Funciones convexas y generalización	15
3.1. Funciones convexas y cóncavas	15
3.1.1. Continuidad de las funciones convexas	18
3.1.2. Derivada direccional de funciones convexas	19
3.2. Subgradientes de funciones convexas	20
3.2.1. Epígrafe e hipógrafo de una función	20
3.3. Funciones convexas diferenciables	24
3.3.1. Funciones convexas y cóncavas doblemente diferenciables	25
3.4. Mínimos y maximos de funciones convexas	26
3.4.1. Minimización de una función convexa	26
3.4.2. Maximización de una función convexa	29
3.4.3. Funciones continuas	30
3.4.4. Funciones convexas	30
3.5. Generalización de una función convexa	31
3.5.1. Funciones cuasiconvexas	31
3.5.2. Funciones cuasiconvexas diferenciables	32
3.5.3. Funciones estrictamente cuasiconvexas	32
3.5.4. Funciones fuertemente cuasiconvexas	33
3.5.5. Funciones pseudoconvexas	34
3.6. Convexidad en un punto	36
3.7. Ejercicios	37
4. Aplicaciones de programación no lineal	45
4.1. Conceptos de optimización	45
4.1.1. Elementos de la optimización	45
4.2. Métodos Univariable	47
4.2.1. Métodos con derivadas	47
4.2.1.1. Método que usan intervalos	48
4.2.1.2. Métodos de búsqueda incremental	48
4.2.1.3. Método de bisección	49
4.2.1.4. Métodos abiertos	64
4.2.1.4.1. Método de Newton-Raphson	64
4.2.2. Métodos de comparación de la función objetivo	69
4.2.2.1. Método de la sección áurea - Golden Ratio	70
4.2.2.2. Método de búsqueda dicotómica	77

4.3.	Métodos Multivariables sin usar derivadas	81
4.3.1.	Optimización de funciones diferenciables multidimensionales	81
4.3.1.1.	Método de Coordenadas Cíclicas	81
4.3.1.2.	Método de Hooke y Jeeves	87
4.4.	Método de búsqueda multivariante usando derivadas	93
4.4.0.1.	Método de paso descendente	93
4.4.0.2.	Método de Newton	96
4.4.1.	Optimización multidimensional con restricciones	100
4.4.1.1.	Métodos de penalizaciones	100
4.4.1.2.	Métodos de barrera	103
4.5.	Fundamentos de Optimización Restringida	105
4.5.1.	Condiciones de Karush-Kuhn y Tucker y optimalidad	105
4.5.1.1.	Problemas de Optimización con Restricciones	105
4.5.1.1.1.	Teorema Condiciones Necesaria de Karush, Kuhn y Tucker (K-K-T)	106
4.5.1.1.2.	Teorema Condiciones Suficiente de Karush, Kuhn y Tucker (K-K-T)	111
4.5.1.2.	Problemas de Optimización con Restricciones de desigual- dad e igualdad	112
4.5.1.2.1.	Teorema Condición Necesaria de Karush, Kuhn y Tucker (K-K-T)	112
4.5.1.2.2.	Teorema Condición Suficiente de Karush, Kuhn y Tucker (K-K-T)	112
4.5.1.3.	Ejercicios	112
5.	Conclusiones	127
	Bibliografía	129

Prólogo

Me gustaría agradecer a todo mi familia pero en especial a mis padres, mis hermanas y mi pareja, el apoyo incondicional durante estos cuatro años, lo cual ha sido primordial para que hoy pueda entregar este trabajo. Gracias, por vuestra paciencia, por calmar mis agobios y no dejar que me rinda nunca.

También dar las gracias a la gran familia que me ha dado esta etapa académica, que han estado apoyándome en todo momento y animándome a seguir adelante.

Por último, agradecer a mi tutor, Antonio Rufian Lizana, su dedicación, entrega y apoyo para realizar este trabajo.

Resumen

El presente trabajo de Fin de Grado tiene como objetivo abordar el tema sobre “*Programación no lineal*”, en el cuál, se muestra un análisis detallado sobre programación no lineal, conjuntos convexos, funciones y aplicaciones de determinados algoritmos. El propósito de este proyecto es que cualquier persona con algunos conceptos de programación lineal pueda entender las técnicas desarrolladas de programación no lineal, enfocándose más en la aplicación de determinadas técnicas.

Se definirá el modelo matemático de programación no lineal discutiendo las diferentes características de este tipo de problema e incluyendo comparaciones de modelos lineales. Adentrándonos en la programación no lineal se analizará los conjuntos convexos y la funciones convexas junto con determinados teoremas y ejemplos de gran importancia.

Dedicaremos una sección a resolver problema de programación no lineal utilizando métodos univariantes y multivariantes, además, presentaremos distintas conclusiones y reflexiones de los resultados obtenidos con ayuda del programa informático R y la aplicación Excel.

Por último, veremos los fundamentos de optimización restringida desarrollando los teoremas de condiciones necesarias y suficientes de Karush-Kuhn y Tucker incorporando ejemplos prácticos.

Abstract

The objective of this Final Degree project is to address the topic of “*Non-linear programming*”, in which, a detailed analysis of nonlinear programming, convex sets, functions and applications of certain algorithms is shown. The purpose of this project is that anyone with some concepts of linear programming can understand the developed techniques of nonlinear programming, focusing more on the application of certain techniques.

The mathematical model of nonlinear programming will be defined by discussing the different characteristics of this type of problem and including comparisons of linear models. Going deeper into nonlinear programming, convex sets and convex functions will be analyzed along with certain theorems and examples of great importance.

We will dedicate a section to solve nonlinear programming problem using univariate and multivariable methods, in addition, we will present different conclusions and reflections of the results obtained with the help of the computer program R and the Excel application.

Finally, we will see the fundamentals of constrained optimization by developing the necessary and sufficient conditions theorems of Karush-Kuhn and Tucker incorporating practical examples.

Índice de figuras

2.1. Envolvente convexa	9
2.2. Cocos convexos	11
2.3. Politopo y simplex	12
3.1. Funciones convexas y cóncavas	16
3.2. Epígrafes e hipógrafos	21
3.3. Interpretación geométrica de los subgradientes	22
3.4. Ejemplo	23
3.5. Mínimos globales y locales	27
3.6. Ejemplo	28
3.7. Funciones cuasiconvexas y cuasicóncavas	31
3.8. Funciones estrictamente cuasiconvexas y estrictamente cuasicóncavas: (a) estrictamente cuasiconvexas, (b) estrictamente cuasiconvexas, (c) estrictamente cuasicóncavas, (d) ni estrictamente cuasiconvexa ni cuasicóncava. . .	33
3.9. Funciones pseudoconvexas y pseudocóncavas: (a) pseudoconvexas, (b) tanto pseudoconvexas como pseudocóncavas, (c) ni pseudoconvexas ni pseudocóncavas	35
4.1. La región factible es la línea negra contenida en la región de color amarillo a lo largo de la restricción $h(x)$	47
4.2. El punto óptimo x^* en la región factible	47
4.3. Existe un mínimo si ambas pendientes en a y b satisfacen que: $f'_a < 0$ y $f'_b > 0$	49
4.4. Existe un mínimo si en los punto a y b se satisface que: $f_a > 0$, $f_b > 0$ y $f_b < f_a$	49
4.5. Gráfico - Método bisección ejercicio 2	55
4.6. Excel - Método bisección ejercicio 2	55
4.7. Excel - Método bisección ejercicio 3	59
4.8. Gráfico - Método bisección ejercicio 3	60
4.9. Excel - Método de bisección ejercicio 4	63
4.10. Gráfico - Método bisección ejercicio 4	64

4.11. Gráfico - Método Newton Raphson ejercicio 1	68
4.12. Gráfico - Método Newton Raphson ejercicio 1	69
4.13. Excel - Método sección áurea ejercicio 1	74
4.14. Gráfico - Método sección áurea ejercicio 1	74
4.15. Gráfico - Método sección áurea ejercicio 1 evaluando una vez la función . .	77
4.16. Excel - Método búsqueda dicotómica ejercicio 1	79
4.17. Gráfico - Método búsqueda dicotómica ejercicio 2	80
4.18. Gráfico en 3D - Método coordenadas cíclicas ejercicio 1	85
4.19. Gráfico en 2D - Método coordenadas cíclicas ejercicio 1	85
4.20. Gráfico - Método coordenadas cíclicas ejercicio 1	86
4.21. Gráfico en XY - Método coordenadas cíclicas ejercicio 1	87
4.22. Búsqueda exploratoria y patrón de búsqueda	87
4.23. Gráfico en 3D - Método Hooke y Jeeves ejercicio 1	91
4.24. Gráfico en 2D - Método Hooke y Jeeves ejercicio 1	91
4.25. Gráfico - Método Hooke y Jeeves ejercicio 1	92
4.26. Gráfico en XY - Método Hooke y Jeeves ejercicio 1	92
4.27. Gráfico en 2D - Método paso descendente ejercicio 1	95
4.28. Gráfico - Método paso descendente ejercicio 1	96
4.29. Gráfico en 3D - Método Newton Raphson ejercicio 1	99
4.30. Gráfico en 2D - Método Newton Raphson ejercicio 1	99
4.31. Gráfico - Método Newton Raphson ejercicio 1	100
4.32. Ejemplo condiciones necesarias de Karush, Kuhn y Tucker	108
4.33. Ejercicio 1	115
4.34. Ejercicio 2	120
4.35. Ejercicio 3	121
4.36. Ejercicio 4	123
4.37. Ejercicio 5	125

Índice de tablas

4.1. Ejercicio 2 - Método de bisección	53
4.2. Ejercicio 2 - Método de bisección solución	54
4.3. Ejercicio 3 - Método de bisección	57
4.4. Ejercicio 3 - Método de bisección solución	58
4.5. Ejercicio 4 - Método de bisección	62
4.6. Ejercicio 4 - Método de bisección solución	62
4.7. Ejercicio 1 - Método de Newton Raphson	67
4.8. Ejercicio 2 - Método Newton Raphson	68
4.9. Ejercicio 1 - Método sección áurea	73
4.10. Ejercicio 1 - Método sección áurea solución	73
4.11. Ejercicio 1 - Método sección áurea evaluando una vez la función	76
4.12. Ejercicio 1 - Método búsqueda dicotómica	79
4.13. Ejercicio 2 - Método búsqueda dicotómica	80
4.14. Ejercicio 1 - Método coordenadas cíclicas	84
4.15. Ejercicio 1 - Método coordenadas cíclicas solución	84
4.16. Ejercicio 1 - Método Hooke y Jeeves	90
4.17. Ejercicio 1 - Método paso descendente	95
4.18. Ejercicio 1 - Método Newton Raphson	98
4.19. Ejercicio 1 - Método de penalización	103

Capítulo 1

Introducción

1.1. Investigación operativa

La investigación de operaciones, también llamada investigación operativa, es una disciplina que se ocupa de la aplicación de métodos analíticos avanzados para ayudar a tomar mejores decisiones.

La programación no lineal en investigación de operaciones es un método matemático de optimización que permite la representación de modelos no lineales para reducir costos, o maximizar ganancias en diferentes áreas de una organización o empresa.

Un problema general se define como:

$$\begin{aligned} \text{Opt } & f(x) \\ \text{s.a. } & q_i \leq 0 \\ & x \in \bar{x} \end{aligned}$$

Si f , q_i son lineales el problema es lineal, en otro caso es no lineal.

1.2. Programación no lineal

La **programación no lineal** forma parte de la investigación de operaciones y también, como la programación lineal, tiene como finalidad proporcionar los elementos para encontrar los puntos óptimos para una función objetivo.

La programación no lineal es un proceso de resolución de un sistema de igualdades y desigualdades sujetas a un conjunto de restricciones sobre un conjunto de variables reales desconocidas, con una función objetivo a optimizar, cuando alguna de las restricciones o la función objetivo no son lineales.

La función objetivo en la programación no lineal, puede ser cóncava, convexa o ni cóncava, ni convexa.

Los problemas de programación no lineal se presentan de muchas formas distintas. Al contrario del método simplex para programación lineal, no se dispone aún de un algoritmo que resuelva todos estos tipos de problemas de forma eficiente.

1.3. Conceptos basicos de programación no lineal

Estos conceptos vienen recogidos en los libros [4] y [2].

Se considera programación no lineal, al conjunto de métodos utilizados para optimizar una función objetivo, sujeta a una serie de restricciones en los que una o más de las funciones incluidas es no lineal.

Por ejemplo:

$$\begin{aligned} \text{Minimizar : } & f(x) \\ \text{Sujeto : } & g_i(x) \leq 0 \quad i = 1, \dots, m \\ & h_j(x) = 0 \quad j = 1, \dots, l \end{aligned}$$

Donde $f, g_1, g_2, \dots, g_m, h_1, h_2, \dots, h_l$ son funciones definidas en el espacio euclidiano de n dimensiones. x es un subconjunto de E_m y x es un vector de componentes x_1, x_2, \dots, x_n .

El problema anterior puede ser resuelto para los valores de la variable x_1, x_2, \dots, x_n que satisfacen las relaciones y minimicen la función f .

La función f se llama, **función objetivo o función criterio**. Cada una de las restricciones $g_i(x)$ para $i = 1, 2, \dots, m$ es llamada restricción de desigualdad y cada una de las restricciones $h_j(x)$ para $j = 1, 2, \dots, l$ es llamada restricción de igualdad.

Un vector \bar{x} que satisface todas las restricciones es llamado una **solución factible** al problema.

La colección de todas las posibles soluciones forman la **región factible**. Los problemas de programación no lineal consisten en encontrar un punto factible x tal que $f(x) \geq f(\bar{x})$ para cada punto factible x . Un punto tal \bar{x} es llamado una solución óptima o simplemente una solución al problema. Si existe más de un punto óptimo, estos son referidos como **soluciones alternativas óptimas**.

Un **problema de programación no lineal** puede expresarse como un problema de maximización y las restricciones de desigualdad pueden estar escritas en la forma $g_i(x) \geq 0$ para $i = 1, 2, \dots, m$. En el caso especial cuando la función objetivo es lineal y cuando todas las restricciones, incluyendo al conjunto X , puede ser representado por desigualdades lineales y/o ecuaciones lineales, el problema anterior es llamado un problema lineal.

Esto indica que la programación lineal es un caso particular de la programación no lineal. Por tal motivo, los problemas no lineales son más difíciles de resolver que los de programación lineal.

A continuación, se indicaran algunas características de los problemas lineales y no lineales:

Programación lineal	Programación no lineal
La solución óptima se encuentra en un punto extremo de la región de factibilidad.	No siempre la solución óptima se encuentra en un punto extremo de la región de factibilidad.
El punto óptimo nunca esta en el interior de la región de factibilidad.	Hay casos donde el punto óptimo esta en el interior de la región factible.

Programación lineal	Programación no lineal
Sus métodos de optimización generan óptimos absolutos ó globales.	Generalmente se encuentra un óptimo local ó relativo, más no el óptimo global ó absoluto.
La región de factibilidad es un conjunto convexo.	Se pueden generar regiones de factibilidad que no son necesariamente convexas.
Sus funciones objetivos y restricciones son lineales.	La función objetivo y las restricciones ambas pueden ser no lineal.

1.4. Planteamientos de problemas

Ejemplos seleccionados de [4].

■ **Ejemplo 1.1** A una compañía de teléfono le cuesta c euros por unidad, cada vez que fabrica un teléfono. Si la compañía cobra x euros por unidad de producto, los clientes pedirán $D(x)$ unidades. Para maximizar las ganancias, ¿Qué precio tendrá que poner la compañía?

Variable

x

Función objetivo

Maximizar $z = (x - c)D(x)$

Sin restricciones.

■

■ **Ejemplo 1.2** Si se utilizan x_1 unidades de capital y x_2 unidades de trabajos, una compañía puede producir x_1x_2 unidades de un bien manufacturado. Se puede conseguir el capital a 30€ la unidad y el trabajo a 7€ la unidad. Se dispone de un total de 600€ para conseguir capital y trabajo. ¿Cómo puede la compañía maximizar la cantidad de bienes que se pueden fabricar?

Variables

x_1 : Unidades de capital

x_2 : Unidades de trabajo

Función objetivo

Maximizar $z = x_1x_2$

Restricciones

$30x_1 + 7x_2 \leq 600$

$x_1 \geq 0 \quad x_2 \geq 0$

■

Capítulo 2

Conjuntos convexos

2.1. Convexidad

Antes de definir propiamente lo que es un *conjunto convexo*, resulta conveniente recordar algunos conceptos previos.

Definición 2.1.1 Se dice que un vector x de \mathbb{R}^n es una *combinación lineal* de los vectores x_1, \dots, x_k si existen $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ adecuados tales que $x = \lambda_1 x_1 + \dots + \lambda_k x_k$. Además:

- Si los λ_i verifican $\lambda_1 + \dots + \lambda_k = 1$, entonces se dice que x es una *combinación afín* de los x_i .
- Si los λ_i verifican $\lambda_i \geq 0$, $i = 1, \dots, k$, entonces x es una *combinación positiva* de los x_i .
- Finalmente, si se verifican ambas condiciones a la vez, esto es, si $\lambda_1 + \dots + \lambda_k = 1$ y $\lambda_i \geq 0$, $i = 1, \dots, k$, entonces se dice que x es una *combinación convexa* de los x_i .

2.1.1. Definición conjunto convexo

Definición 2.1.2 Un **conjunto** $C \subseteq \mathbb{R}^n$ se dice que es **convexo** si el segmento que une un par cualesquiera de puntos del conjunto pertenece al conjunto, o sea

$$\text{si } x_1, x_2 \in C, \text{ entonces } [x_1, x_2] \subseteq C$$

De otra forma,

$$\text{si } x_1, x_2 \in C, \text{ entonces } \lambda x_1 + (1 - \lambda)x_2 \in C, \forall \lambda \in [0, 1]$$

Proposición 1 Un *conjunto* C es convexo si, y sólo si, cualquier combinación convexa de puntos de C está en C .

Lema 1 Si $C \subseteq \mathbb{R}^n$ es convexo, entonces contiene a todas las combinaciones lineales de la forma

$$\sum_{j=1}^k \lambda_j x_j, \quad x_j \in C, \quad \text{con } \lambda_j \geq 0 \quad \text{y} \quad \sum_{j=1}^k \lambda_j = 1$$

Lema 2 Sean $C_1, C_2 \subseteq \mathbb{R}^n$, convexos, entonces

1. $C_1 \cap C_2$ es convexo.
2. $C_1 + C_2 = \{x_1 + x_2 : x_1 \in C_1, x_2 \in C_2\}$ es convexo.
3. $C_1 - C_2 = \{x_1 - x_2 : x_1 \in C_1, x_2 \in C_2\}$ es convexo.

Como consecuencia de la primera propiedad enunciada, se tiene que:

■ **Ejemplo 2.1** $P = \{x \in \mathbb{R}^n : Ax \leq b\}$ donde $A \in M_{m \times n}$ y $b \in \mathbb{R}^m$ es convexo.

A este tipo de conjuntos se les llama **poliédricos o poliedros** y son intersección de semiespacios. ■

■ **Ejemplo 2.2**

$$3x_1 + 2x_2 + x_3 \leq 65x_1 - 6x_2 \leq 7$$

que escrito en forma matricial queda

$$\begin{pmatrix} 3 & 2 & 1 \\ 5 & -6 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \leq \begin{pmatrix} 6 \\ 7 \end{pmatrix}$$

■ **Observación 1** Las columnas y filas de la matriz anterior se pueden reordenar sin que el poliedro cambie. ■

$$\begin{pmatrix} -6 & 5 & 0 \\ 2 & 3 & 1 \end{pmatrix} \begin{pmatrix} x_2 \\ x_1 \\ x_3 \end{pmatrix} \leq \begin{pmatrix} 7 \\ 6 \end{pmatrix}$$

2.1.2. Definición punto extremo

Definición 2.1.3 Sea C un conjunto convexo cualquiera, se dice que $\bar{x} \in C$ es un

punto extremo $\iff \bar{x} = \lambda x_1 + (1 - \lambda)x_2$, con $x_1, x_2 \in C$ y $\lambda \in [0, 1] \implies \lambda = 0$ ó $\lambda = 1$.

O lo que es lo mismo, no existen $x_1, x_2 \in C$, tal que $\bar{x} = \lambda x_1 + (1 - \lambda)x_2$, con $\lambda \in (0, 1)$. Por tanto, no existe una combinación lineal convexa estricta para \bar{x} .

2.1.3. Definición de dirección

Definición 2.1.4 Sea $C \in \mathbb{R}^n$ un conjunto convexo no vacío. Se dice que $d \in \mathbb{R}^n$ es una **dirección** $\iff \forall x \in C$ se verifica que $x + \mu d \in C, \forall \mu > 0$.

Observación 2 Para que un conjunto convexo tenga al menos una dirección, es necesario que sea no acotado. El recíproco también es cierto, como se verá posteriormente.

Observación 3 Si d es una dirección, también lo es $\alpha d, \forall \alpha > 0, \alpha \in \mathbb{R}$.

2.1.4. Definición dirección extrema

Definición 2.1.5 Una dirección $d \in C$ se dice que es una **dirección extrema** \iff dadas dos direcciones $d_1, d_2 \in C$, verificando $d = \alpha d_1 + \beta d_2$, con α y $\beta \geq 0 \implies \alpha = 0$ ó $\beta = 0$. O sea, una dirección extrema no se puede poner como combinación lineal estricta de otras direcciones.

2.1.5. Teorema de caracterización de conjuntos poliédricos

Teorema 2.1.1 Se considera el conjunto $\mathbb{P} = \{x \in \mathbb{R}^n : Ax \leq b\}$ donde $A \in M_{m \times n}$ y $b \in \mathbb{R}^m$. Un punto $\bar{x} \in \mathbb{P} \implies$ se puede expresar de la forma:

$$\bar{x} = \sum_{i=1}^k \lambda_i x_i + \sum_{j=1}^l \mu_j d_j$$

donde x_1, \dots, x_k son los puntos extremos de \mathbb{P} , y d_1, \dots, d_l son las direcciones extremas de \mathbb{P} y además $\lambda_i \geq 0, \forall i = 1, \dots, k, \mu_j \geq 0, \forall j = 1, \dots, l$ y $\sum_{i=1}^k \lambda_i = 1$.

Corolario 2.1.2 $C \subseteq \mathbb{R}^n$ tiene al menos una dirección extrema $\iff C$ es no acotado.

Otro concepto interesante es el de *cono convexo*, que se puede definir de la siguiente forma.

2.1.6. Definición cono convexo

Definición 2.1.6 Un **cono convexo** es un subconjunto A de \mathbb{R}^n que es convexo, no vacío y tal que si x está en A , entonces λx también está en A para todo $\lambda \geq 0$.

En definitiva, un subconjunto no vacío A de \mathbb{R}^n es un cono convexo si es cerrado para la suma y el producto por números reales no negativos.

2.1.7. Ejemplos de conjuntos convexos

■ **Ejemplo 2.3** 1. $S = \{(x_1, x_2, x_3) : x_1 + 2x_2 - x_3 = 4\} \subset \mathbb{R}^3$

Esta es una ecuación de un plano de \mathbb{R}^3 . En general $S = \{x : p^t x = \alpha\}$ se llama un hiperplano en \mathbb{R}^n , donde p es un vector no nulo en \mathbb{R}^n , usualmente llamado vector no nulo en \mathbb{R}^n , se denomina gradiente o normal, al hiperplano, y α es un escalar. Nótese

que si $\bar{x} \in S$, se tiene $p^t \bar{x} = \alpha$, por lo que se puede escribir equivalentemente $S = \{x : p^t(x - \bar{x}) = 0\}$. Por lo tanto, el vector p es ortogonal a todos los vectores $(x - \bar{x})$ para $x \in S$, por lo que es perpendicular a la superficie del hiperplano S . ■

■ **Ejemplo 2.4** 2. $S = \{(x_1, x_2, x_3) : x_1 + 2x_2 - x_3 \leq 4\} \subset R^3$

Se trata de puntos situados en un lado del hiperplano definido anteriormente. Estos puntos forman un semiespacio. En general, un semiespacio $S = \{x : p^t x \leq \alpha\}$ en R^n es un conjunto convexo. ■

■ **Ejemplo 2.5** 3. $S = \{(x_1, x_2, x_3) : x_1 + 2x_2 - x_3 \leq 4, 2x_1 - x_2 + x_3 \leq 6\} \subset R^3$

Este conjunto es la intersección de dos semiespacios. En general, el conjunto $S = \{x : Ax \leq b\}$ es un conjunto convexo, donde A es una matriz $m \times n$ y b es un m -vector. Este conjunto es la intersección de m semiespacios y suele llamarse conjunto poliédrico. ■

■ **Ejemplo 2.6** 4. $S = \{(x_1, x_2) : x_1^2 + x_2^2 \leq 4\} \subset R^2$

Este conjunto representa puntos sobre y dentro de un círculo con centro $(0, 0)$ y radio 2. ■

■ **Ejemplo 2.7** 5. $S = \{x : x \text{ resuelve el problema } P \text{ de abajo}\}$

P: Minimizar $c^t x$ sujeto a $Ax = b, x \geq 0$

Aquí c , es un n -vector, b es un m -vector, A es una matriz $m \times n$, y x es un n -vector. El conjunto S da todas las soluciones óptimas al problema de programación lineal de minimizar la función lineal $c^t x$ sobre la región poliédrica definida por $Ax = b, x \geq 0$. Este conjunto es a su vez un conjunto poliédrico, siendo la intersección de $c^t x = v^*$ con $Ax = b, x \geq 0$, donde v^* es el valor óptimo de P . ■

2.2. Envolverte convexos

En esta sección introducimos primero la nociones de envolvente convexa.

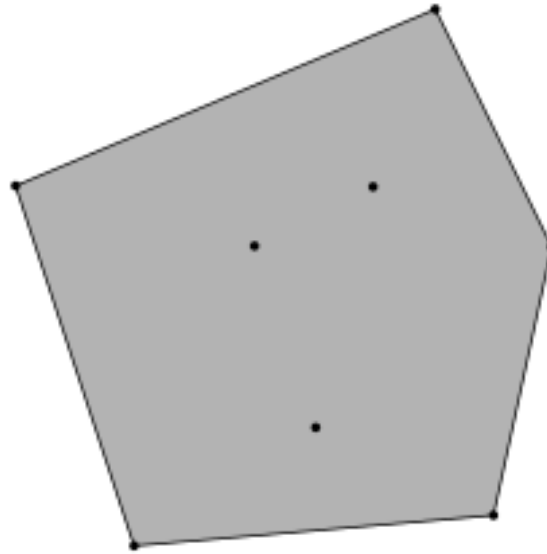
Definición 2.2.1 La **envolvente convexa, envoltura convexo o cápsula convexa** de un conjunto de puntos X de dimensión n como la intersección de todos los conjuntos convexos que contienen a X .

Dados k puntos x_1, x_2, \dots, x_k su envolvente convexa C viene dada por la expresión:

$$C(X) = \left\{ \sum_{i=1}^k \alpha_i x_i \mid x_i \in X, \alpha_i \in \mathbb{R}, \alpha_i \geq 0, \sum_{i=1}^k \alpha_i = 1 \right\}$$

En el caso particular de puntos en un plano, si no todos los puntos están alineados, entonces su envolvente convexa corresponde a un polígono convexo cuyos vértices son algunos de los puntos del conjunto inicial de puntos.

Una forma intuitiva de ver la envolvente convexa de un conjunto de puntos en el plano, es imaginar una banda elástica estirada que los encierra a todos. Cuando se libere la banda elástica tomará la forma de la envolvente convexa. Esta figura seleccionada del libro [2] nos muestra un ejemplo.



Envoltura convexa de un conjunto de 8 puntos en el plano.

Figura 2.1: Envoltente convexa

La unión de todas las combinaciones convexas de conjuntos finitos de puntos de $A \subset \mathbb{R}^n$ se denomina **cápsula convexa de A** .

Análogamente se define la **envoltura afín** y la **envoltura positiva** de A , y se representa por $aff A$ y $pos A$, respectivamente, como la intersección de todos los subespacios afines, en el primer caso, o de todos los conos convexos, en el segundo, de \mathbb{R}^n que contiene a A .

En definitiva, $con A$ (respectivamente, $aff A$ o $pos A$) no es otra cosa que el menor conjunto convexo (respectivamente, subespacio afín o cono positivo) que contiene a A .

Además:

Proposición 2 Dada un subconjunto cualquiera A de \mathbb{R}^n , su envoltura convexa es, precisamente el conjunto de todas las combinaciones convexas de una cantidad finita de elementos de A .

Resultados análogos se demuestran para las envolturas afín y convexa de un conjunto A ; por lo tanto, $aff A$ y $pos A$ coinciden con los conjuntos de todas las combinaciones afines y positivas, respectivamente, de cualquier cantidad finita de elementos de A .

Uno de los resultados fundamentales sobre la generación de las envolturas convexas es el *Teorema de Carathéodory*.

2.2.1. Teorema de Carathéodory

El siguiente teorema asegura que se puede expresar un punto de la envoltura convexa de un conjunto como combinación lineal convexa de un máximo de $n+1$ puntos del conjunto.

Teorema 2.2.1 Sea $X \subseteq \mathbb{R}^n$ no vacío y $x \in \text{conv}(X)$. Entonces x se puede expresar como combinación lineal convexa de un máximo de $n + 1$ puntos de X .

Demostración

Sea $x \in \text{conv}(X)$. Entonces existen escalares $\lambda_1, \dots, \lambda_m \geq 0$ y puntos de X , x^1, \dots, x^m , tal que $x = \lambda_1 x^1 + \dots + \lambda_m x^m$ y $\lambda_1 + \dots + \lambda_m = 1$. Supongamos que m es el mínimo número natural tal que x puede ser expresado de esta forma. Así, todos los puntos x^i son distintos y $\lambda_i > 0$, para todo $i = 1, \dots, m$.

Supongamos, por reducción al absurdo, que $m > n + 1$. El método de la reducción al absurdo consiste en demostrar que un resultado es cierto demostrando que NO puede ser de otra forma. Consiste en suponer que el resultado a demostrar es falso y llegar, a partir de ahí, a una contradicción.

Entonces x^1, \dots, x^m son puntos afínmente dependientes, por lo que existen $\mu_1, \dots, \mu_m \in \mathbb{R}$, no todos nulos, tales que $\mu_1 x^1 + \dots + \mu_m x^m = 0$ y $\mu_1 + \dots + \mu_m = 0$. Como la suma de los μ_i es igual a 0 y no todos ellos son nulos, habrá positivos y negativos. Además, como los λ son positivos, existirá $t > 0$ tal que $\lambda_i + t\mu_i \geq 0$ para todo $i = 1, \dots, m$, con al menos uno de ellos en que se alcanza la igualdad a cero. Así,

$$x = \sum_{i=1}^m (\lambda_i + t\mu_i) x^i$$

y se expresa x como combinación lineal convexa de $m - 1$ puntos de X a lo sumo, una vez eliminados los términos con coeficientes igual a 0. Sin embargo, habíamos supuesto que m era el mínimo natural para el que existía una combinación lineal convexa, por lo que llegamos a una contradicción. Así pues, se tiene que $m \leq n + 1$.

El siguiente lema es una consecuencia inmediata de la definición de convexidad. Afirma que la intersección de dos conjuntos convexos es convexa y que la suma algebraica de dos conjuntos convexos es también convexa.

Especial interés presentan los conjuntos obtenidos como al envoltura convexa de puntos:

Definición 2.2.2 La envoltura convexa de un número finito de puntos se denomina **politopo (convexo)**, en general, o **polígono (convexo)**, en el caso del plano euclídeo. En particular, un **k-símplice** es a la envoltura convexa de $k + 1$ puntos afínmente independientes.

Es sencillo demostrar que todo punto de un k -símplice tiene una única representación como combinación convexa de sus vértices. Además, el Teorema de Carathéodory establece entonces que la envoltura convexa de un conjunto A es la unión de todos los símplices con vértices en A .

2.2.2. Definición conjunto arbitrario

Definición 2.2.3 Sea S un conjunto arbitrario en \mathbb{R}^n . La envoltura convexa de S , denotado $\text{conv}(S)$, es la colección de todas las combinaciones convexas de S . En otras palabras, $x \in \text{conv}(S)$ si y sólo si x puede representarse como

- $x = \sum_{j=1}^k \lambda_j x_j$

- $\sum_{j=1}^k \lambda_j = 1$
- $\lambda_j \geq 0 \quad j = 1, \dots, k$

donde k es un número entero positivo y $x_1, x_2, \dots, x_k \in S$

La figura 2.2 muestra algunos ejemplos de envolventes convexas y esta seleccionada del libro [2]. En realidad, tal y como se ve en cada caso, $\text{conv}(S)$ es el conjunto convexo mínimo (más envolvente) que contiene a S . Este es el caso en general, como se indica en el lema siguiente.

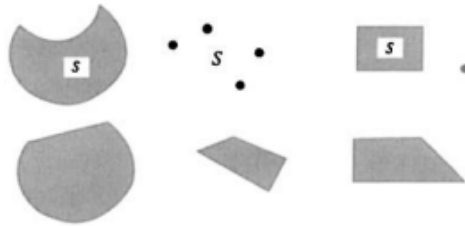


Figura 2.2: Cocos convexos

Lema 3 Sea S un conjunto arbitrario en R^n . Entonces, $\text{conv}(S)$ es el conjunto convexo más pequeño que contiene a S . De hecho, $\text{conv}(S)$ es la intersección de todos los conjuntos convexos que contienen a S .

De forma similar a la discusión anterior, se puede definir la envoltura afín de S como la colección de todas las combinaciones afines de puntos en S . Este es el subespacio afín más pequeño.

Por ejemplo, la envoltura afín de dos puntos distintos es el subespacio unidimensional que contiene a puntos distintos es la línea unidimensional que contiene estos dos puntos. Del mismo modo, la envoltura lineal de S es la colección de todas las combinaciones lineales de puntos en S .

La envoltura convexa de un número finito de puntos conduce a las definiciones de un poliedro y de simplex.

2.2.3. Definición poliedro y simplex

Definición 2.2.4 La envoltura convexa de un número finito de puntos x_1, \dots, x_{k+1} en R^n se denomina **poliedro**. Si x_1, x_2, \dots, x_k y x_{k+1} son independientes, lo que significa que $x_2 - x_1, x_3 - x_1, \dots, x_{k+1} - x_1$ son linealmente independientes entonces $\text{cov}(x_1, \dots, x_{k+1})$, el convexo de x_1, \dots, x_{k+1} , se denomina **simplex** con vértices x_1, \dots, x_{k+1} .

La figura 2.3, seleccionada de [2], muestra ejemplos de un poliedro y un simplex en R^n . Obsérvese que el número máximo de vectores linealmente independientes en R^n es n , y por tanto no puede haber ningún simplex en R^n que tenga más de $n + 1$ vértices.

Un **politopo** es la generalización del concepto polígono (2D), o *poliedro* (3D) a cualquier otra dimensión. Geométricamente un politopo es una región finita de un espacio n -dimensional encerrado por un número finito de hiperplanos.

En geometría, un **simplex** o **n-simplex** es el análogo n-dimensional del triángulo. Debe su nombre a que es el objeto geométrico cerrado “más simple” que tiene n dimensiones.

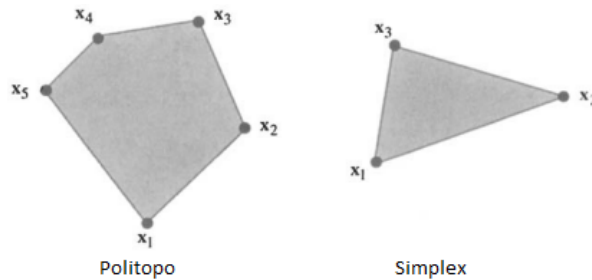


Figura 2.3: Politopo y simplex

Definición 2.2.5 La dimensión de un conjunto cualquiera A de \mathbb{R}^n no es más que la dimensión del menor subespacio afín que lo contiene, y se denota por $\dim A$.

Observación 4 Dada un conjunto convexo de \mathbb{R}^n , entonces $\text{int } K$ e $y \in \text{cl } K$, entonces el segmento semiabierto $[x, y) \subset \text{int } K$. Esto permite demostrar, entre otros, los siguientes resultados:

- i) Si K es un conjunto convexo de \mathbb{R}^n , entonces $\text{int } K$ y $\text{cl } K$ son también convexos.
- ii) Si $K = \text{con}\{x_1, \dots, x_{k+1}\}$ es un simplexe k-dimensional, entonces $\text{relint } K \neq \emptyset$.
- iii) Si K es un conjunto convexo y no vacío de \mathbb{R}^n , entonces su interior relativo $\text{relint } K \neq \emptyset$.
- iv) Si K es un conjunto convexo de \mathbb{R}^n , entonces $\text{relint } K = \text{relint } \text{cl } K$ y $\text{cl } K = \text{cl } \text{relint } K$.
- v) Si A es un conjunto abierto de \mathbb{R}^n , entonces $\text{conv } A$ también es abierto.
- vi) La envoltura convexa de un conjunto compacto es compacta.

2.2.4. Teoremas tipo Helly

El Teorema de Carthéodory establecía cuál es el número máximo de puntos de un conjunto A necesarios para construir, mediante una combinación convexa, un punto cualquiera de $\text{conv } A$. otro resultado igualmente sencillo e importante sobre envolturas convexas es el siguiente:

2.2.4.1. Teorema de Radon

Teorema 2.2.2 Todo subconjunto de puntos afinmente dependientes de \mathbb{R}^n (en particular, cualquier conjunto con, al menos, $n+2$ puntos) puede expresarse como la unión de dos conjuntos disjuntos cuyas envolturas convexas tienen un punto en común.

A partir del Teorema de Radon puede deducirse fácilmente el conocido Teorema de Helly, un resultado fundamental y típico de la geometría combinatoria de conjuntos convexos.

2.2.4.2. Teorema de Helly

Teorema 2.2.3 Sea K_1, \dots, K_r conjuntos convexos del espacio euclídeo \mathbb{R}^n . Si cualquier colección de $n + 1$ de tales subconjuntos tienen un punto en común, entonces todos los K_i tienen un punto en común.

Este teorema es el más representativo de toda una clase de resultados conocidos como los Teoremas de Tipo Helly, los cuales responden a la siguiente formulación general:

Sea \mathcal{F} una familia de conjuntos y sea r un entero positivo. Si r conjuntos cualesquiera de \mathcal{F} satisfacen la propiedad P , entonces toda la familia \mathcal{F} verifica la propiedad Q .

Capítulo 3

Funciones convexas y generalización

3.1. Funciones convexas y cóncavas

Las funciones cóncavas y convexas son un tipo especialmente importantes de funciones para la programación matemática. Ahora se verán algunas definiciones, así como algunas de sus propiedades más importantes, sobre todo aquellas que pueden utilizarse para resolver problemas de optimización. Estas propiedades serán utilizadas para desarrollar las condiciones de optimalidad y los algoritmos de resolución de los problemas no lineales.

Al contrario que lo que ocurría para los problemas lineales, no existe ningún método eficiente para resolver los problemas de programación no lineal en general, ni siquiera una colección de métodos tales que al menos uno de ellos pudiera ser utilizado para resolver un problema dado. Existen, sin embargo, métodos más o menos eficientes para resolver problemas no lineales donde las restricciones y la función objetivo verifican algunas “buenas” propiedades, que se pueden dividir en dos categorías:

-*Condiciones de uniformidad*: continuidad, diferenciabilidad.

-*Condiciones de convexidad*: convexidad, concavidad, convexidad y concavidad generalizada, monotonicidad.

En general, las condiciones de uniformidad son fáciles de verificar por la forma de las funciones. Las condiciones de convexidad en muchos casos son difíciles de estudiar de una forma eficiente.

Por ello en esta sección se centrará en el estudio de la convexidad.

Definición 3.1.1 Sea $S \subseteq \mathbb{R}^n$ se dice que es **convexo** si el segmento lineal que une dos puntos cualesquiera de S también pertenece al conjunto, es decir:

$$x, y \in S \Rightarrow \lambda x + (1 - \lambda)y \in S, \quad \forall \lambda \in [0, 1]$$

A las medias ponderadas de la forma $\lambda x + (1 - \lambda)y$, donde $\lambda \in [0, 1]$ se les llama **combinaciones convexas** de x, y .

Nota 1 El conjunto \mathbb{R}^n , el vacío y los conjuntos con un único punto son conjuntos convexos.

Sea $f : S \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$, donde S es un conjunto convexo no vacío. Se dice que f es

Definición 3.1.2 Convexa en S si

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y), \quad \forall x, y \in S, \quad \forall \lambda \in [0, 1].$$

Definición 3.1.3 Estrictamente convexa en S si

$$f(\lambda x + (1 - \lambda)y) < \lambda f(x) + (1 - \lambda)f(y), \quad \forall x, y \in S, \quad \forall \lambda \in [0, 1].$$

Definición 3.1.4 Cóncava en S si

$$f(\lambda x + (1 - \lambda)y) \geq \lambda f(x) + (1 - \lambda)f(y), \quad \forall x, y \in S, \quad \forall \lambda \in [0, 1].$$

Definición 3.1.5 Estrictamente cóncava en S si

$$f(\lambda x + (1 - \lambda)y) > \lambda f(x) + (1 - \lambda)f(y), \quad \forall x, y \in S, \quad \forall \lambda \in [0, 1].$$

Nota 2 Es inmediato que

- Estrictamente convexa (cóncava) \Rightarrow convexa (cóncava), pero el recíproco no es cierto.
- f es (estrictamente) convexa si y sólo si $-f$ es (estrictamente) cóncava.

Se considera ahora la interpretación geométrica de las funciones convexas y cóncavas.

Sea x e y dos puntos distintos en el dominio de f ; y se considera el punto $\lambda x + (1 - \lambda)y$, con $\lambda \in (0, 1)$.

Obsérvese que $\lambda f(x) + (1 - \lambda)f(y)$ da la media ponderada de $f(x)$ y $f(y)$, mientras que $f[\lambda x + (1 - \lambda)y]$ da el valor de f en el punto $\lambda x + (1 - \lambda)y$. Por lo tanto, para una función convexa f , el valor de f en puntos del segmento de línea $\lambda x + (1 - \lambda)y$ es menor o igual que la altura de la cuerda que une los puntos $[x, f(x)]$ y $[y, f(y)]$. Para una función cóncava la cuerda está por debajo de la propia función. Por tanto, una función es convexa y cóncava si y sólo si es afín.

La figura extraída del libro [2], muestra algunos ejemplos de funciones convexas y cóncavas.

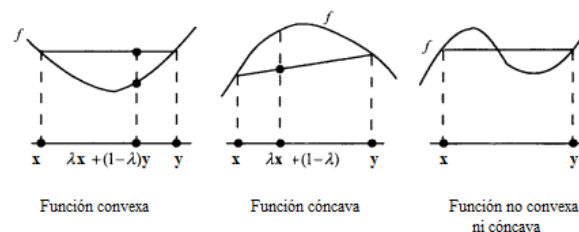


Figura 3.1: Funciones convexas y cóncavas

A continuación se muestran algunos ejemplos de funciones convexas. Tomando los negativos de esta funciones, se obtiene algunos ejemplos de funciones cóncavas:

1. $f(x) = 3x + 4$
2. $f(x) = |x|$
3. $f(x) = x^2 - 2x$
4. $f(x) = -x^{1/2} \quad x \geq 0$
5. $f(x_1, x_2) = 2x_1^2 + x_2^2 - 2x_1x_2$
6. $f(x_1, x_2, x_3) = x_1^4 + 2x_2^2 + 3x_3^2 - 4x_1 - 4x_2x_3$

Nótese que en cada uno de los ejemplos anteriores, excepto en el ejemplo 4, la función f es convexa sobre R^n . En el ejemplo 4, la función no está definida para $x < 0$. Se pueden construir fácilmente ejemplos de funciones que son convexas sobre una región pero no sobre R^n . Por ejemplo, $f(x) = x^3$ no es convexa sobre R pero es convexa sobre $S = \{x : x \geq 0\}$.

Los ejemplos anteriores citan algunos casos ilustrativos de funciones convexas. En cambio, a continuación damos algunos casos particulares importantes de funciones convexas que surgen muy a menudo en la práctica y que es útil recordar.

1.- Sea $f_1, f_2, \dots, f_k : R^n \rightarrow R$ sea funciones convexas. Entonces:

- (a) $f(x) = \sum_{j=1}^k \alpha_j f_j(x)$, donde $\alpha_j > 0$ para $j = 1, 2, \dots, k$ es una función convexa.
- (b) $f(x) = \max\{f_1(x), \dots, f_k(x)\}$ es una función convexa

2.- Supongamos que $g : R^n \rightarrow R$ es una función cóncava. Sea $S = \{x : g(x) > 0\}$, y definimos $f : S \rightarrow R$ como $f(x) = 1/g(x)$. Entonces f es convexa sobre S .

3.- Sea $g : R \rightarrow R$ es una función no decreciente, univariante y convexa y sea $h : R^n \rightarrow R$ una función convexa. Entonces la función compuesta $f : R^n \rightarrow R$ definida como $f(x) = [g(x)]$ es una función convexa.

4.- Sea $g : R^n \rightarrow R$ una función convexa, y sea $h : R^n \rightarrow R$ sea una función afín de la forma $h(x) = Ax + b$, donde A es una matriz $m \times n$ y b es un vector $m \times 1$. Entonces la función compuesta $f : R^n \rightarrow R$ definida como $f(x) = [g(x)]$ es una función convexa.

Definición 3.1.6 Sea $f : S \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ y $\alpha \in \mathbb{R}$. Al conjunto $S_\alpha = \{x \in S / f(x) \leq \alpha\}$ se le llama **conjunto de nivel inferior** y $S^\alpha = \{x \in S / f(x) \geq \alpha\}$ se le llama **conjunto superior** de orden o nivel *alpha*.

Sea S un conjunto convexo no vacío en R^n , y sea $f : S \rightarrow R$ una función convexa. Entonces el conjunto de nivel $S_\alpha = \{x \in S : f(x) \leq \alpha\}$ es convexo para todo $\alpha \in \mathbb{R}$.

Demostración

Sea $x_1, x_2 \in S_\alpha$. Por tanto, $x_1, x_2 \in S$ y $f(x_1) \leq \alpha$ y $f(x_2) \leq \alpha$. Sea ahora $\lambda \in (0, 1)$ y $x = \lambda x_1 + (1 - \lambda)x_2$. Por la convexidad de S , se tiene que $x \in S$. Además, por la convexidad de f ,

$$f(x) \leq \lambda f(x_1) + (1 - \lambda)f(x_2) \leq \lambda\alpha + (1 - \lambda)\alpha = \alpha$$

Por lo tanto, $x \in S_\alpha$, y por lo tanto, S_α es convexo.

Nota 3 Si f es cóncava en S convexo no vacío entonces S^α es convexo para todo $\alpha \in \mathbb{R}$.

El recíproco no es, en general, cierto.

3.1.1. Continuidad de las funciones convexas

Una propiedad importante de las funciones convexas y cóncavas es que son continuas en el interior de su dominio.

Una función $f(x)$ es continua en el punto $x = x_0$ si:

- 1) $f(x_0)$ está definida.
- 2) $\lim_{x \rightarrow x_0} f(x)$ existe.
- 3) $\lim_{x \rightarrow x_0} f(x) = f(x_0)$.

Teorema 3.1.1 Sea S un conjunto convexo no vacío en R^n , y sea $f : S \rightarrow R$ convexo. Entonces f es continua en el interior de S .

Demostración

Sea $\bar{x} \in \text{int } S$. Para demostrar la continuidad de \bar{x} , necesitamos mostrar que dado $\epsilon > 0$, existe un $\delta > 0$ tal que $\|x - \bar{x}\| \leq \delta$ por lo que implica que $|f(x) - f(\bar{x})| \leq \epsilon$. Ya que $\bar{x} \in \text{int } S$, existe un $\delta' > 0$ tal que $\|x - \bar{x}\| \leq \delta'$ implica que $x \in S$. Siendo θ de la siguiente forma:

$$\theta = \max_{1 \leq i \leq n} \{ \max[f(\bar{x} + \delta' e_i) - f(\bar{x}), f(\bar{x} - \delta' e_i) - f(\bar{x})] \} \quad (3.1)$$

donde e_i , es un vector de ceros excepto un 1 en la posición de la i . Nótese que $0 \leq \theta \leq \infty$.

$$\delta = \min\left(\frac{\delta'}{n}, \frac{\epsilon \delta'}{n\theta}\right) \quad (3.2)$$

Elija una x con $\|x - \bar{x}\| \leq \delta$. Si $x_i - \bar{x}_i \geq 0$, $z_i = \delta' e_i$; en caso contrario, $z_i = -\delta' e_i$. Entonces $x - \bar{x} = \sum_{i=1}^n \alpha_i z_i$, donde $\alpha_i \geq 0$ para $i = 1, \dots, n$. Además,

$$\|x - \bar{x}\| = \delta' \left(\sum_{i=1}^n \alpha_i^2 \right)^{1/2} \quad (3.3)$$

A partir de (3.2), se deduce que $\alpha_i \leq 1/n$ para $i = 1, \dots, n$. Por tanto, por la convexidad de f , y dado que $0 \leq n\alpha_i \leq 1$, se obtiene

$$\begin{aligned} f(x) &= f\left(\bar{x} + \sum_{i=1}^n \alpha_i z_i\right) = f\left[\frac{1}{n} \sum_{i=1}^n (\bar{x} + n\alpha_i z_i)\right] \leq \frac{1}{n} \sum_{i=1}^n f(\bar{x} + n\alpha_i z_i) = \\ &= \frac{1}{n} \sum_{i=1}^n f[(1 - n\alpha_i)\bar{x} + n\alpha_i(\bar{x} + z_i)] \leq \frac{1}{n} \sum_{i=1}^n [(1 - n\alpha_i)f(\bar{x}) + n\alpha_i f(\bar{x} + z_i)] \end{aligned}$$

Por tanto, $f(x) - f(\bar{x}) \leq \sum_{i=1}^n \alpha_i [f(\bar{x} + z_i) - f(\bar{x})]$. De (3.1) es evidente que $f(\bar{x} + z_i) - f(\bar{x}) \leq \theta$, para cada i ; y como $\alpha_i \geq 0$, se deduce que

$$f(x) - f(\bar{x}) \leq \theta \sum_{i=1}^n \alpha_i \quad (3.4)$$

Observando (3.3) y (3.2), resulta que $\alpha_i \leq \epsilon/n\theta$, y (3.4) implica que $f(x) - f(\bar{x}) \leq \epsilon$. Hasta ahora, se ha demostrado que $\|x - \bar{x}\| \leq \delta$ implica que $f(x) - f(\bar{x}) \leq \epsilon$. Por definición, esto establece la semicontinuidad superior de f en \bar{x} . Para completar la prueba, necesitamos establecer también la semicontinuidad inferior de f en \bar{x} , es decir, para demostrar que $f(x) - f(\bar{x}) \leq \epsilon$. Sea $y = 2\bar{x} - x$ y observe que $\|y - \bar{x}\| \leq \delta$. Por lo tanto, como arriba,

$$f(y) - f(\bar{x}) \leq \epsilon \quad (3.5)$$

Pero $\bar{x} = (1/2)y + (1/2)x$, y por la convexidad de f , se tiene

$$f(\bar{x}) \leq (1/2)f(y) + (1/2)f(x) \quad (3.6)$$

Combinando (3.5) y (3.6) anteriores, se deduce que $f(\bar{x}) - f(x) \leq \epsilon$, y la prueba es completa. Nótese que las funciones convexas y cóncavas pueden no ser continuas en todas partes. Sin embargo, por este último teorema, sólo se permiten puntos de discontinuidad en la frontera de S , como se ilustra con la siguiente función convexa definida en $S = \{x : -1 \leq x \leq 1\}$:

$$f(x) = \begin{cases} x^2 & |x| < 1 \\ 2 & |x| = 1 \end{cases}$$

Teorema 3.1.2 Si f es convexa en S convexo no vacío entonces f es continua en $\text{int}(S)$.
El recíproco no es cierto.

Teorema 3.1.3 Sea $f_1, \dots, f_m : S \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ convexas en S convexo no vacío. Entonces:

1. $\sum_{i=1}^m \alpha_i f_i$ con $\alpha_i \geq 0$ es una función convexa.
2. $\max_{i=1, \dots, m} \{f_i\}$ es una función convexa.

Nota 4 La combinación lineal positiva de funciones estrictamente convexas es estrictamente convexa.

3.1.2. Derivada direccional de funciones convexas

El concepto de derivadas direccionales es particularmente útil en la motivación y desarrollo de algunos criterios de optimalidad y procedimientos computacionales en programación no lineal, donde uno está interesado en encontrar una dirección a lo largo de la función disminuye o aumenta.

Definición 3.1.7 Sea S un conjunto no vacío en \mathbb{R}^n , y sea $f : S \rightarrow \mathbb{R}$. Sea $\bar{x} \in S$ y d un vector no nulo tal que $\bar{x} + \lambda d \in S$ para $\lambda > 0$ y suficientemente pequeño. La derivada direccional de f en \bar{x} a lo largo del vector d , denotada por $f'(\bar{x}; d)$, está dada por el siguiente límite si existe:

$$f'(\bar{x}; d) = \lim_{\lambda \rightarrow 0^+} \frac{f(\bar{x} + \lambda d) - f(\bar{x})}{\lambda}$$

En particular, el límite de la definición existe para funciones globalmente definidas como funciones convexas y cóncavas.

Lema 1 Sea $f : R^n \rightarrow R$ una función convexa. Se considera cualquier punto $\bar{x} \in R^n$ y una dirección distinta de 0 $d \in R^n$. Entonces existe la derivada direccional $f'(\bar{x}; d)$, de f en \bar{x} en la dirección d .

Demostración

Sea $\lambda_2 > \lambda_1 > 0$. Observando la convexidad de f se tiene:

$$f(\bar{x} + \lambda_1 d) = f\left[\frac{\lambda_1}{\lambda_2}(\bar{x} + \lambda_2 d) + \left(1 - \frac{\lambda_1}{\lambda_2}\right)\bar{x}\right] \leq \frac{\lambda_1}{\lambda_2} f(\bar{x} + \lambda_2 d) + \left(1 - \frac{\lambda_1}{\lambda_2}\right) f(\bar{x})$$

Esta desigualdad implica que

$$\frac{f(\bar{x} + \lambda_1 d) - f(\bar{x})}{\lambda_1} \leq \frac{f(\bar{x} + \lambda_2 d) - f(\bar{x})}{\lambda_2}$$

Así, el cociente de diferencias $[f(\bar{x} + \lambda d) - f(\bar{x})]/\lambda$ es monótono decreciente (no creciente) a medida que $\lambda \rightarrow 0^+$.

Ahora, dado cualquier $\lambda \geq 0$, también se tiene, por la convexidad de f , que

$$f(\bar{x}) = f\left[\frac{\lambda}{1 + \lambda}(\bar{x} - d) + \frac{1}{1 + \lambda}(\bar{x} + \lambda d)\right] \leq \frac{\lambda}{1 + \lambda} f(\bar{x} - d) + \frac{1}{1 + \lambda} f(\bar{x} + \lambda d)$$

Así que

$$\frac{f(\bar{x} + \lambda d) - f(\bar{x})}{\lambda} \geq f(\bar{x}) - f(\bar{x} - d)$$

Por tanto, la secuencia monótona decreciente de valores $[f(\bar{x} + \lambda d) - f(\bar{x})]/\lambda$, como $\lambda \rightarrow 0^+$, está limitada desde abajo por la constante $f(\bar{x}) - f(\bar{x} - d)$. Por tanto, el límite del teorema existe y viene dado por

$$\lim_{\lambda \rightarrow 0^+} \frac{f(\bar{x} + \lambda d) - f(\bar{x})}{\lambda} = \inf_{\lambda > 0} \frac{f(\bar{x} + \lambda d) - f(\bar{x})}{\lambda}$$

3.2. Subgradientes de funciones convexas

3.2.1. Epígrafe e hipógrafe de una función

Una función f sobre S puede ser descrita completamente por el conjunto $\{[x, f(x)]\} \subset \mathbb{R}^{n+1}$, que se denomina gráfico de la función. Se pueden construir dos conjuntos que están relacionados con la gráfica de f : el epígrafe, que consta de puntos por encima del gráfica de f , y el hipógrafe que consiste en los puntos por debajo del gráfico de f .

Sea $f : S \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$.

Definición 3.2.1 Se llamará **epígrafo** de f al conjunto

$$Epi(f) = \{(x, y) \in \mathbb{R}^{n+1} / x \in S, y \in \mathbb{R} \quad f(x) \leq y\}$$

Definición 3.2.2 El **hipógrafe** de f

$$Hip(f) = \{(x, y) \in \mathbb{R}^{n+1} / x \in S, y \in \mathbb{R} \quad f(x) \geq y\}$$

La figura 3.2 ilustra los epígrafes e hipógrafos de varias funciones, seleccionadas del libro [2].

En la figura 3.2 a, ni epígrafe ni el hipógrafe son conjuntos convexos. Pero en la figura 3.2 b y c, respectivamente, el epígrafe y el hipógrafe de f son conjuntos convexos.

Una función es convexa si y sólo si su epígrafe es un conjunto convexo y, equivalentemente que una función es cóncava si y sólo si su hipógrafe es un conjunto convexo.

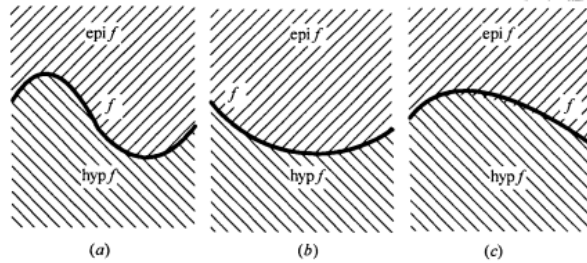


Figura 3.2: Epígrafes e hipógrafos

Teorema 3.2.1 Sea S un conjunto convexo no vacío en \mathbb{R}^n , y sea $f : S \rightarrow \mathbb{R}$. Entonces f es convexo si y sólo si $Epi(f)$ es convexo.

Demostración

Se supone que f es convexa, y se deja que (x_1, y_1) y $(x_2, y_2) \in epi f$, es decir, $x_1, x_2 \in S$, $y_1 \geq f(x_1)$, y $y_2 \geq f(x_2)$. Sea $\lambda \in (0, 1)$. Entonces

$$\lambda y_1 + (1 - \lambda)y_2 \geq \lambda f(x_1) + (1 - \lambda)f(x_2) \geq f(\lambda x_1 + (1 - \lambda)x_2)$$

donde la última desigualdad se deduce por la convexidad de f . Nótese que $\lambda x_1 + (1 - \lambda)x_2 \in S$. Por tanto, $[\lambda x_1 + (1 - \lambda)x_2, \lambda y_1 + (1 - \lambda)y_2] \in epi f$, y por tanto $epi f$ es convexo. A la inversa, supongamos que $epi f$ es convexo, y se deja que $x_1, x_2 \in S$. Entonces $[x_1, f(x_1)]$ y $[x_2, f(x_2)]$ pertenecen a $epi f$, y por la convexidad de $epi f$, se debe tener

$$[\lambda x_1 + (1 - \lambda)x_2, \lambda f(x_1) + (1 - \lambda)f(x_2)] \in epi f \quad \text{para } \lambda \in (0, 1)$$

En otras palabras, $\lambda f(x_1) + (1 - \lambda)f(x_2) \geq f[\lambda x_1 + (1 - \lambda)x_2]$ para cada $\lambda \in (0, 1)$, es decir, f es convexo. Esto completa la demostración.

Este teorema puede utilizarse para verificar la convexidad o concavidad dada un función f . Haciendo uso de este resultado, está claro que las funciones ilustradas en la figura 3.2 (a) no es ni convexa ni cóncava, (b) convexa, y (c) cóncava.

Como el epígrafe de una función convexa y el hipógrafo de una función cóncava son conjuntos convexos, tienen hiperplanos de apoyo en puntos de su límite. Estos hiperplanos de apoyo conducen a a la noción de subgradietes, que se define a continuación.

Definición 3.2.3 Sea S un conjunto convexo no vacío en R^n , y sea $f : S \rightarrow R$ convexo. Entonces ξ se llama **subgradiente** de f en $\bar{x} \in S$ si

$$f(x) \geq f(\bar{x}) + \xi^t(x - \bar{x}) \quad \text{para todo } x \in S$$

Análogamente, sea $f : S \rightarrow R$ cóncavo. Entonces ξ se llama subgradiente de f en $\bar{x} \in S$ si

$$f(x) \leq f(\bar{x}) + \xi^t(x - \bar{x}) \quad \text{para todo } x \in S$$

De la definición anterior se deduce inmediatamente que la colección de subgradietes de f en \bar{x} (conocido como el subdiferencial de f en \bar{x}) es un conjunto convexo.

La figura 3.3 seleccionad de [2], muestra ejemplos de subgradietes de funciones convexas y cóncavas. En la figura se ve que la función $f(\bar{x}) + \xi^t(x - \bar{x})$ corresponde a un hiperplano de soporte del epígrafe o hipógrafo de la función f . El vector subgradiente ξ corresponde a al pendiente del hiperplano de apoyo.

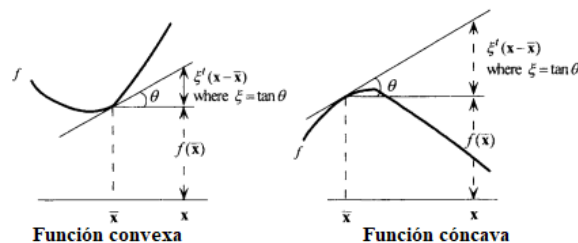


Figura 3.3: Interpretación geométrica de los subgradietes

■ **Ejemplo 3.1** Sea $f(x) = \min\{f_1(x), f_2(x)\}$, donde f_1 y f_2 son:

$$f_1(x) = 4 - |x|, \quad x \in R \quad f_2(x) = 4 - (x - 2)^2, \quad x \in R$$

Dado que $f_2(x) \geq f_1(x)$ para $1 \leq x \leq 4$, f puede representarse como:

$$f(x) = \begin{cases} 4 - x & 1 \leq x \leq 4 \\ 4 - (x - 2)^2 & c.c \end{cases}$$

En la figura 3.4 la función cóncava f se muestra en líneas oscuras. Nótese que $\xi = -1$ es la pendiente y, por tanto, el subgradiente de f en cualquier punto x del intervalo abierto $(1, 4)$. Si $x < 1$ o $x > 4$, $\xi = -2(x - 2)$ es el único subgradiente de f . En los puntos $x = 1$

y $x = 4$, los subgradienates no son únicos porque existen muchos hiperplanos de apoyo. En $x = 1$, la familia de subgradienates se caracteriza por

$$\lambda \nabla f_1(1) + (1 - \lambda) \nabla f_2(1) = \lambda(-1) + (1 - \lambda)(2) = 2 - 3\lambda \quad \text{para } \lambda \in [0, 1]$$

En otras palabras, cualquier ξ en el intervalo $[-1, 2]$ es un subgradiente de f en $x = 1$, y esto corresponde a las pendientes de la familia de hiperplanos de apoyo en $x = 1$. En $x = 4$, la familia de subgradienates se caracteriza por

$$\lambda \nabla f_1(4) + (1 - \lambda) \nabla f_2(4) = \lambda(-1) + (1 - \lambda)(-4) = -4 + 3\lambda \quad \text{para } \lambda \in [0, 1]$$

En otras palabras, cualquier ξ en el intervalo $[-4, -1]$ es un subgradiente de f en $x = 4$. El ejemplo aborda la caracterización general de subgradiente de funciones de la forma $f(x) = \min\{f_1(x), f_2(x)\}$.

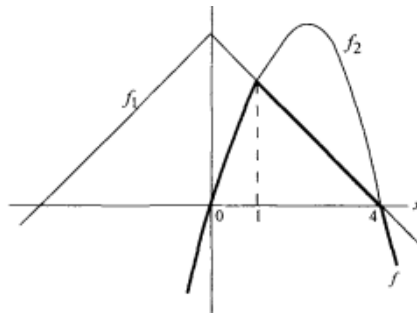


Figura 3.4: Ejemplo

■

El siguiente teorema muestra que toda función convexa o cóncava tiene al menos un subgradiente en puntos del interior de su dominio. La prueba se basa en el hecho de que un conjunto convexo tiene un hiperplano de apoyo en los puntos de la frontera.

Teorema 3.2.2 Sea S un conjunto no convexo en R^n , y sea $f : S \rightarrow R$ convexa. Entonces para $\bar{x} \in \text{int } S$, existe un vector ξ tal que el hiperplano

$$H = \{(x, y) : y = f(\bar{x}) + \xi^t(x - \bar{x})\}$$

soporta *epi* f en $[\bar{x}, f(\bar{x})]$. En particular,

$$f(x) \geq f(\bar{x}) + \xi^t(x - \bar{x}) \quad \text{para todo } x \in S$$

es decir, ξ es un subgradiente de f en \bar{x}

Corolario 3.2.3 Sea S un conjunto convexo no vacío en R^n , y sea $f : S \rightarrow R$ estrictamente convexo. Entonces para $\bar{x} \in \text{int } S$ existe un vector ξ tal que

$$f(x) > f(\bar{x}) + \xi^t(x - \bar{x}) \quad \text{para todo } x \in S, x \neq \bar{x}$$

Teorema 3.2.4 Sea S un conjunto convexo no vacío en R^n , y sea $f : S \rightarrow R$. Supongamos que para cada punto $\bar{x} \in \text{int } S$ existe un vector subgradiente ξ tal que

$$f(x) \geq f(\bar{x}) + \xi^t(x - \bar{x}) \quad \text{para todo } x \in S$$

Entonces, f es convexa en $\text{int } S$.

3.3. Funciones convexas diferenciables

La caracterización de funciones convexas mediante su definición es, en general, muy difícil de aplicar en la práctica. Para comprobar si una función es o no convexa es necesario encontrar otras caracterizaciones más sencillas de aplicar. En particular, en el caso de funciones diferenciables, existen resultados que proporcionan esas caracterizaciones en términos del gradiente y del Hessiano y que serán útiles en los problemas de programación no lineal con diferenciables.

Definición 3.3.1 Sea S un conjunto no vacío en R^n , y sea $f : S \rightarrow R$. Entonces se dice que f es diferenciable en $\bar{x} \in \text{int } S$ si existe un vector $\nabla f(\bar{x})$, llamado **vector gradiente** y una función $\alpha : R^n \rightarrow R$ tal que

$$f(x) = f(\bar{x}) + \nabla f(\bar{x})^t(x - \bar{x}) + \|x - \bar{x}\|\alpha(\bar{x}; x - \bar{x}) \quad \text{para cada } x \in S$$

donde $\lim_{x \rightarrow \bar{x}} \alpha(\bar{x}; x - \bar{x}) = 0$. Se dice que la función f es diferenciable en el conjunto abierto $S' \subseteq S$ si es diferenciable en cada punto de S' . La representación de f anterior se denomina expansión de primer orden (serie de Taylor) de f el punto \bar{x} , y sin el término de resto definido implícitamente que implica la función α , la representación resultante se denomina aproximación de primer orden (serie de Taylor) de f en (o alrededor de) el punto \bar{x} .

Nótese que si f es diferenciable en \bar{x} , sólo puede haber un **vector gradiente** y este vector viene dado por

$$\nabla f(\bar{x}) = \left(\frac{\partial f(\bar{x})}{\partial x_1}, \dots, \frac{\partial f(\bar{x})}{\partial x_n} \right)^t = (f_1(\bar{x}), \dots, f_n(\bar{x}))^t$$

donde, $f_i(x) = \partial f(\bar{x})/\partial x_i$ es la derivada parcial de f con respecto a x_i en \bar{x} .

Lema

Sea S un conjunto convexo no vacío en R^n , y sea $f : S \rightarrow R$ es convexo. Supongamos que f es diferenciable en $\bar{x} \in \text{int } S$. Entonces la colección de subgradiente de f en \bar{x} es el conjunto único $\{\nabla f(\bar{x})\}$.

Teorema 3.3.1 Sea $f : C \subseteq R^n \rightarrow R$ diferenciable definida en un conjunto C convexo no vacío, entonces

$$f \text{ es convexa} \iff \forall \bar{x} \in C \text{ se verifica } f(x) \geq f(\bar{x}) + \nabla f(\bar{x})^t(x - \bar{x}), \forall x \in C.$$

Teorema 3.3.2 Sea $f : S \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ diferenciable en S convexo y abierto, entonces

1. f es convexa si y solo si $f(x) - f(y) \geq \nabla f(y)^T(x - y), \forall x, y \in S$.
2. f es estrictamente convexa si y solo si $f(x) - f(y) > \nabla f(y)^T(x - y), \forall x \neq y \in S$.

Teorema 3.3.3 Sea $f : S \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ diferenciable en S convexo y abierto, entonces

1. f es convexa si y solo si $(\nabla f(x) - \nabla f(y))^T(x - y) \geq 0, \forall x, y \in S$.
2. f es estrictamente convexa si y solo si $(\nabla f(x) - \nabla f(y))^T(x - y) > 0, \forall x \neq y \in S$.

3.3.1. Funciones convexas y cóncavas doblemente diferenciables

Definición 3.3.2 Sea S un conjunto no vacío en \mathbb{R}^n , y sea $f : S \rightarrow \mathbb{R}$. Entonces se dice que f es dos veces diferenciable en $\bar{x} \in \text{int } S$ si existe un vector $\nabla f(\bar{x})$, y una matriz simétrica $n \times n$, $H(\bar{x})$, llamada matriz hessiana, y una función $\alpha : \mathbb{R}^n \rightarrow \mathbb{R}$ tal que

$$f(x) = f(\bar{x}) + \nabla f(\bar{x})^T(x - \bar{x}) + \frac{1}{2}(x - \bar{x})^T H(\bar{x})(x - \bar{x}) + \|x - \bar{x}\|^2 \alpha(\bar{x}; x - \bar{x})$$

para cada $x \in S$, donde $\lim_{x \rightarrow \bar{x}} \alpha(\bar{x}; x - \bar{x}) = 0$. Se dice que la función f es dos veces diferenciable en el conjunto abierto $S' \subseteq S$ si es dos veces diferenciable en cada punto de S' .

Si $f : X \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ es dos veces diferenciable en X no vacío, llamamos **Hessiano** de f en $\bar{x} \in \text{int}(X)$ y se denota por $H_f(\bar{x})$ a la matriz $n \times n$. Esta compuesta por las derivadas parciales de segundo orden $f_{ij}(\bar{x}) \equiv \partial^2 f(\bar{x}) / \partial x_i \partial x_j$ para $i = 1, \dots, n, j = 1, \dots, n$ y viene dada como

$$H_f(\bar{x}) = \begin{pmatrix} f_{11}(\bar{x}) & f_{12}(\bar{x}) & \dots & f_{1n}(\bar{x}) \\ f_{21}(\bar{x}) & f_{22}(\bar{x}) & \dots & f_{2n}(\bar{x}) \\ \dots & \dots & \dots & \dots \\ f_{n1}(\bar{x}) & f_{n2}(\bar{x}) & \dots & f_{nn}(\bar{x}) \end{pmatrix}$$

En forma expandida, la representación anterior puede escribirse como

$$f(x) = f(\bar{x}) + \sum_{j=1}^n f_j(\bar{x})(x_j - \bar{x}_j) + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (x_i - \bar{x}_i)(x_j - \bar{x}_j) f_{ij}(\bar{x}) + \|x - \bar{x}\|^2 \alpha(\bar{x}; x - \bar{x})$$

De nuevo, sin el término del resto asociado a la función α , esta representación se conoce como una aproximación de segundo orden (serie de Taylor) en (o sobre) el punto \bar{x} .

Teorema 3.3.4 Sea $f : C \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ dos veces diferenciable en un conjunto C convexo no vacío. Entonces f es convexa \iff el Hessiano es semidefinido positivo $\forall x \in C$.

Lema 2 Considérese el siguiente Hessiano

$$H = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

Entonces H es semidefinido positivo $\iff a \geq 0, c \geq 0$ y $ac - b^2 \geq 0$.

Teorema 3.3.5 Sea $f : S \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ dos veces diferenciable en S abierto convexo. Entonces

1. f es convexa en S si solo si $H_f(x)$ es semidefinido positivo en cada $x \in S$.
2. Si $H_f(x)$ es definido positivo en cada $x \in S$, entonces f es estrictamente convexa.

■ **Ejemplo 3.2** Sea $f(x_1, x_2) = 2x_1 + 6x_2 - 2x_1^2 - 3x_2^2 + 4x_1x_2$. Se tiene

$$\nabla f(\bar{x}) = \begin{bmatrix} 2 - 4\bar{x}_1 + 4\bar{x}_2 \\ 6 - 6\bar{x}_2 + 4\bar{x}_1 \end{bmatrix}$$

$$H(\bar{x}) = \begin{bmatrix} -4 & 4 \\ 4 & -6 \end{bmatrix}$$

Por ejemplo, tomando $\bar{x} = (0, 0)^t$, la expansión de segundo orden de esta función es dada por

$$f(x_1, x_2) = (2, 6) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \frac{1}{2} (x_1, x_2) \begin{bmatrix} -4 & 4 \\ 4 & -6 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Obsérvese que aquí no hay término de resto ya que la función dada es cuadrática, por lo que la representación anterior es exacta.

■

3.4. Mínimos y máximos de funciones convexas

En esta sección consideramos los problemas de minimización y maximización de una función convexa sobre un conjunto convexo y desarrollamos condiciones necesarias y/o suficientes para la optimización.

3.4.1. Minimización de una función convexa

El caso de la maximización de una función cóncava es similar al de la minimización de una función convexa.

Sea $f : S \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ definida en un conjunto S no vacío y se considera el Problema de Programación Matemática:

$$\begin{aligned} \text{Min } & f(x) \\ \text{s.a } & x \in S \\ & (PM) \end{aligned}$$

Definición 3.4.1 Un punto $x \in S$ se dice **solución factible** de (PM).

Definición 3.4.2 Si $\bar{x} \in S$ verifica que $f(\bar{x}) \leq f(x) \forall x \in S$, entonces \bar{x} es la **solución óptima** de (PM). También es llamado **solución global** o simplemente **solución**.

Definición 3.4.3 $\bar{x} \in S$ es una **solución local** de (PM), si existe $B(\bar{x}, r)$ tal que $f(\bar{x}) \leq f(x) \forall x \in B(\bar{x}, r)$.

Definición 3.4.4 Se considera (PM). Si el conjunto factible es convexo y la función f es convexa, se dice que el **problema es convexo**.

Observación 1 Los problemas de programación lineal son problemas convexos.

Teorema 3.4.1 Sea $f : C \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ convexa definida en un conjunto C convexo no vacío. Se considera el problema (PM) y supongamos que \bar{x} es una solución local del problema. Entonces

1. \bar{x} es una solución global.
2. Si f es estrictamente convexa, entonces \bar{x} es solución única.

Teorema 3.4.2 Sea $f : C \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ convexa diferenciable definida en un conjunto C convexo no vacío, entonces

$$\bar{x} \text{ es óptimo global de (PM)} \iff \nabla f(\bar{x})^t(x - \bar{x}) \geq 0 \forall x \in C.$$

Corolario 3.4.3 Sea $f : C \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ convexa diferenciable definida en un conjunto C convexo no vacío, entonces

$$\bar{x} \text{ es mínimo global de (PM)} \iff \nabla f(\bar{x}) = 0$$

Todos estos tipos de óptimos o mínimos locales se denominan a veces mínimos relativos. La figura 3.5 ilustra casos de mínimos locales y globales para el problema de minimizar $f(x)$ sujeto a $x \in S$, donde f y S se muestran en la figura.

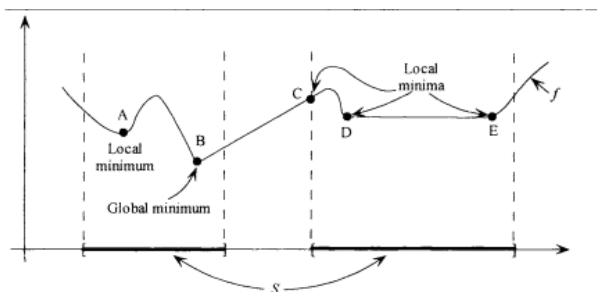


Figura 3.5: Mínimos globales y locales

Los puntos de S correspondientes a A, B y C son también mínimos locales estrictos y son mínimos locales estrictos y fuertes, mientras que los correspondientes al segmento plano entre D y E son mínimos locales que no son ni estrictos ni fuertes.

Teorema 3.4.4 Se considera el problema de minimizar $f(x)$ sujeto a $x \in S$, donde f es una función convexa y dos veces diferenciable y S es conjunto convexo, y supongamos que existe una solución óptima \bar{x} . Entonces el conjunto de soluciones óptimas alternativas está caracterizado por el conjunto

$$S^* = \{x \in S : \nabla f(x)^t(x - \bar{x}) \leq 0 \quad y \quad \nabla f(x) = \nabla f(\bar{x})\}$$

Corolario 3.4.5 El conjunto S^* de soluciones óptimas alternativas puede definirse de forma equivalente como

$$S^* = \{x \in S : \nabla f(x)^t(x - \bar{x}) = 0 \quad y \quad \nabla f(x) = \nabla f(\bar{x})\}$$

Corolario 3.4.6 Supongamos que f es una función cuadrática dada por $f(x) = c^t x + (1/2)x^t H x$ y que S es poliédrico. Entonces S^* es un conjunto poliédrico dado por

$$S^* = \{x \in S : c^t(x - \bar{x}) \leq 0, H(x - \bar{x}) = 0\} = \{x \in S : c^t(x - \bar{x}) = 0, H(x - \bar{x}) = 0\}$$

■ **Ejemplo 3.3** Ejemplo procedente del libro [2].

$$\begin{aligned} \text{Minimiza} \quad & (x_1 - \frac{3}{2})^2 + (x_2 - 5)^2 \\ \text{s.a.} \quad & -x_1 + x_2 \leq 2 \\ & 2x_1 + 3x_2 \leq 11 \\ & -x_1 \leq 0 \\ & -x_2 \leq 0 \end{aligned}$$

Claramente, $f(x_1, x_2) = (x_1 - \frac{3}{2})^2 + (x_2 - 5)^2$ es una función convexa, que da el cuadrado de la distancia al punto $(\frac{3}{2}, 5)$. El conjunto poliédrico convexo S está representado por las cuatro desigualdades anteriores. El problema se representa en la figura 3.6. De la figura se observa que el punto óptimo es $(1, 3)$. El vector gradiente en el punto $(1, 3)$ es $\nabla f(1, 3) = (-1, -4)^t$. Se ve geoméricamente que el vector $(-1, -4)$ forma un ángulo de $< 90^\circ$ con cada vector de la forma $(x_1 - 1, x_2 - 3)$, donde $(x_1, x_2) \in S$. Así, la condición de optimalidad se verifica y, el punto $(1, 3)$ es el único óptimo.

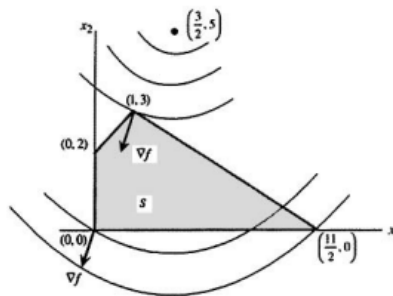


Figura 3.6: Ejemplo

3.4.2. Maximización de una función convexa

Ahora desarrollamos una condición necesaria para un máximo de una función convexa sobre un conjunto convexo. Por desgracia, esta condición no es suficiente. Por lo tanto, es posible, y en realidad no es improbable, que existan varios máximos locales. A diferencia del caso de minimización, no existe información local en dichas soluciones que pueda llevarnos a puntos mejores. Por lo tanto, maximizar una función convexa suele ser una tarea mucho más difícil que minimizar una función convexa. De nuevo, minimizar una función cóncava es similar a maximizar una función convexa.

Teorema 3.4.7 Sea $f : R^n \rightarrow R$ una función convexa, y sea S un conjunto convexo no vacío en R^n . Se considera el problema de maximizar $f(x)$ sujeto a $x \in S$. Si $\bar{x} \in S$ es una solución óptima local, $\xi^t(x - \bar{x}) \leq 0$ para cada $x \in S$, donde ξ es cualquier subgradiente de f en \bar{x} .

Demostración

Supongamos que $\bar{x} \in S$ es una solución óptima local. Entonces existe una vecindad $N_\epsilon(\bar{x})$ existe tal que $f(x) \leq f(\bar{x})$ para cada $x \in S \cap N_\epsilon(\bar{x})$. Sea $x \in S$, y observe que $\bar{x} + \lambda(x - \bar{x}) \in S \cap N_\epsilon(\bar{x})$ para $\lambda > 0$ y suficientemente pequeño. Por lo tanto,

$$f[\bar{x} + \lambda(x - \bar{x})] \leq f(\bar{x})$$

Sea ξ un subgradiente de f en \bar{x} . Por la convexidad de f , se tiene

$$f[\bar{x} + \lambda(x - \bar{x})] - f(\bar{x}) \geq \lambda \xi^t(x - \bar{x})$$

Las dos desigualdades anteriores, implica que $\lambda \xi^t(x - \bar{x}) \leq 0$, y dividiendo por $\lambda > 0$, el resultado se deduce.

Corolario 3.4.8 Además de los supuestos del teorema, supongamos que f es diferenciable. Si $\bar{x} \in S$ es una solución óptima local, $\nabla f(\bar{x})^t(x - \bar{x}) \leq 0$ para todo $x \in S$.

Nótese, que el resultado anterior del ejemplo 3.3 es, en general, necesario pero no suficiente para optimalidad. Para ilustrar, se deja que $f(x) = x^2$ y $S = \{x : -1 \leq x \leq 2\}$. El máximo de f sobre S es igual a 4 y se alcanza en $x = 2$. Sin embargo, en $\bar{x} = 0$, se tiene $\nabla f(\bar{x}) = 0$ y por tanto $\nabla f(\bar{x})^t(x - \bar{x}) = 0$ para cada $x \in S$. En relación al ejemplo anterior, se tiene dos máximos locales $(0, 0)$ y $(\frac{11}{2}, 0)$. Ambos puntos satisfacen la condición necesaria. Si nos encontramos en el punto óptimo local $(0, 0)$ desgraciadamente no existe ninguna información local que nos conduzca hacia el punto máximo global $(\frac{11}{2}, 0)$ no hay ningún criterio local conveniente que nos diga que estamos en el punto óptimo.

Teorema 3.4.9 Sea $f : R^n \rightarrow R$ sea una función convexa, y sea S un conjunto poliédrico compacto no vacío en R^n . Se considera el problema de maximizar $f(x)$ sujeto a $x \in S$. Existe entonces una solución óptima \bar{x} al problema, donde \bar{x} es un punto extremo de S .

3.4.3. Funciones continuas

Entre las propiedades deseables de las funciones a optimizar se encuentra también la continuidad, que se ve que está relacionada con la convexidad. Con esta propiedad se tiene el **teorema de Weierstrass**.

Teorema 3.4.10 Sea $f : K \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$, donde K es un conjunto compacto (cerrado y acotado) no vacío. Si f una función continua, entonces existe x_1 y $x_2 \in K$, tales que

$$f(x_1) \leq f(x) \leq f(x_2) \quad \forall x \in K$$

O sea, se tiene máximo y mínimo.

3.4.4. Funciones convexas

Teorema 3.4.11 Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $S \subseteq \mathbb{R}^n$ convexo no vacío. Si \bar{x} es un **mínimo local** de f en S y

- f es convexa, entonces \bar{x} es un **mínimo global**.
- f es estrictamente convexa, entonces \bar{x} es el **único mínimo global**.

Teorema 3.4.12 Condición necesaria de optimalidad

Sea f diferenciable en S abierto no vacío. Si \bar{x} es un mínimo local entonces $\nabla f(\bar{x}) = 0$. Si f es dos veces diferenciable en S y \bar{x} es un mínimo local entonces $\nabla f(\bar{x}) = 0$ y $H_f(\bar{x})$ es **semidefinido positivo**.

Teorema 3.4.13 Condición suficiente de optimalidad global

Sea $f : S \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ diferenciable en S abierto, convexo.

1. Si f es convexo en S ,
 \bar{x} es un *mínimo (máximo) global* si y sólo si $\nabla f(\bar{x}) = 0$.
2. Si f es estrictamente convexa en S ,
 \bar{x} es un *el único mínimo global* si y sólo si $\nabla f(\bar{x}) = 0$.

Teorema 3.4.14 Condición suficiente de optimalidad local

Sea $f : S \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ dos veces diferenciable en S abierto. Sea $\bar{x} \in S$ tal que $\nabla f(\bar{x}) = 0$. Se verifica que:

1. Si $H_f(\bar{x})$ es definida *positiva*, entonces \bar{x} es un **mínimo local estricto** de f .
2. Si $H_f(\bar{x})$ es *indefinida*, entonces \bar{x} es un **punto de silla** (la función pasa de cóncava a convexa) de f .
3. Si $H_f(\bar{x})$ es *semidefinida* positiva, entonces \bar{x} es un **mínimo local o punto de silla** (la función pasa de cóncava a convexa) de f .

Teorema 3.4.15 Sea $f : S \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ convexa en S un **poliedro** compacto no vacío. Entonces existe $\bar{x} \in S$ donde f alcanza el máximo en S y además \bar{x} es un punto extremo de S .

3.5. Generalización de una función convexa

En esta sección se verán varios tipos de funciones que son similares a las funciones convexas y cóncavas, pero que sólo comparten algunas de sus propiedades deseables de las funciones convexas, pero que son más generales, ya que en algunos de los resultados vistos anteriormente la convexidad es una condición suficiente, pero no necesaria.

3.5.1. Funciones cuasiconvexas

Definición 3.5.1 Sea $f : S \subseteq \mathbb{R}^n \Rightarrow \mathbb{R}$, donde S es un conjunto convexo no vacío en \mathbb{R}^n . Se dice que la función f es **cuasiconvexa** si para cada x y $y \in S$, se cumple la siguiente desigualdad:

$$f(\lambda x + (1 - \lambda)y) \leq \max\{f(x), f(y)\} \quad \text{para cada } \lambda \in (0, 1)$$

Se dice que f es cuasicóncava si $-f$ es cuasiconvexa.

A partir de esta definición, una función f es cuasiconvexa si siempre que $f(y) \geq f(x)$, $f(y)$ es mayor o igual que todas las combinaciones convexas de x y y . Por lo tanto, si f aumenta su valor en un punto a lo largo de cualquier dirección, debe seguir siendo no decreciente en esa dirección. Por lo tanto su sección transversal univariante es monótona o unimodal. Una función f es cuasicóncava siempre que $f(y) \geq f(x)$, f en todas las combinaciones convexas de x y y es mayor o igual a $f(x)$.

Teorema 3.5.1 Sea $f : C \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ definida en un conjunto C convexo no vacío. La función

$$f \text{ es cuasiconvexa} \iff C_\alpha = \{x \in C : f(x) \leq \alpha\} \text{ es convexa, } \forall \alpha \in \mathbb{R}.$$

La figura 3.7 extraída de [2], muestra algunos ejemplos de funciones cuasiconvexas y funciones cuasicóncavas. Nos centraremos en las funciones cuasiconvexas. A la función que es a la vez cuasiconvexa y cuasicóncava se llama *cuasimonótona* (d), la figura (a) es una función cuasiconvexa, (b) es cuasicóncava y (c) no es ni cuasiconvexa, ni cuasicóncava.



Figura 3.7: Funciones cuasiconvexas y cuasicóncavas

Definición 3.5.2 Sea $f : C \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ definida en un conjunto C convexo no vacío. La función f se dice cuasicóncava si y solo si $-f$ es cuasiconvexa.

Teorema 3.5.2 Sea $f : S \rightarrow \mathbb{R}$ donde S es un conjunto convexo no vacío en \mathbb{R}^n . La función f es cuasiconvexa si y sólo si $S_\alpha = \{x \in S : f(x) \leq \alpha\}$ es convexa para cada número real α .

Teorema 3.5.3 Sea S un conjunto poliédrico compacto no vacío en \mathbb{R}^n , y sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$ cuasiconvexo y continuo en S . Se considera el problema de maximizar $f(x)$ sujeto a $x \in S$. Entonces existe una solución óptima \bar{x} del problema, donde \bar{x} es un punto extremo de S .

3.5.2. Funciones cuasiconvexas diferenciables

El siguiente teorema da una caracterización necesaria y suficiente de una función cuasiconvexa diferenciable.

Teorema 3.5.4 Sea $f : C \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ diferenciable definida en C convexo no vacío.

Entonces f es cuasiconvexa \iff se da algunas de las siguientes condiciones equivalentes

1. — Si $x_1, x_2 \in C$ y $f(x_1) \leq f(x_2)$, entonces $\nabla f(x_2)^t(x_1 - x_2) \leq 0$
2. — Si $x_1, x_2 \in C$ y $\nabla f(x_2)^t(x_1 - x_2) > 0$, entonces $f(x_1) > f(x_2)$

3.5.3. Funciones estrictamente cuasiconvexas

La función estrictamente cuasiconvexas y estrictamente cuasicóncavas son especialmente importantes en la programación no lineal porque garantizan que un mínimo local sobre un conjunto convexo son respectivamente, un mínimo y un máximo global.

Definición 3.5.3 Sea $f : S \rightarrow \mathbb{R}$, donde S es un conjunto convexo no vacío en \mathbb{R}^n . Se dice que la función f es **estrictamente cuasiconvexa** si para cada $x, y \in S$ con $f(x) \neq f(y)$, se tiene

$$f(\lambda x + (1 - \lambda)y) < \max \{f(x), f(y)\} \quad \text{para cada } \lambda \in (0, 1)$$

La función f se llama estrictamente cuasicóncava si $-f$ es estrictamente cuasiconvexa.

En general no es cierto que estrictamente cuasiconvexa \Rightarrow cuasiconvexa.

Obsérvese que en esta definición toda función convexa es estrictamente cuasiconvexa. En la figura 3.8 extraída de [2], se dan ejemplos de funciones estrictamente cuasiconvexas y estrictamente cuasicóncava. Además, la definición impide que se produzcan “puntos planos” en cualquier lugar, excepto en los puntos extremos.

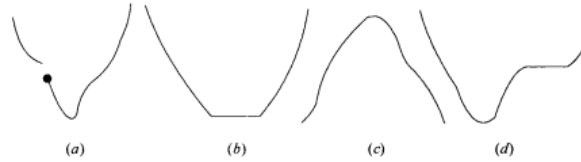


Figura 3.8: Funciones estrictamente cuasiconvexas y estrictamente cuasicóncavas: (a) estrictamente cuasiconvexas, (b) estrictamente cuasiconvexas, (c) estrictamente cuasicóncavas, (d) ni estrictamente cuasiconvexa ni cuasicóncava.

Definición 3.5.4 Sea $f : S \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ con S convexo. Se dice que f es **explícitamente cuasiconvexa** en S si f es cuasiconvexa y estrictamente cuasiconvexa en S .

Es inmediato que convexa \Rightarrow explícitamente cuasiconvexa.

Teorema 3.5.5 Sea $f : S \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$, S convexa. Si f es explícitamente cuasiconvexa en S todo mínimo local de f en S es mínimo global.

Teorema 3.5.6 Sea $f : R^n \rightarrow R$ estrictamente cuasiconvexo. Se considera el problema de minimizar $f(x)$ sujeto a $x \in S$, donde S es un conjunto convexo no vacío en R^n . Si \bar{x} es una solución óptima local, \bar{x} es también una solución óptima global.

Teorema 3.5.7 Sea $f : C \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ diferenciable definida en un conjunto C convexo no vacío, entonces f es estrictamente convexa \iff

$$\forall \bar{x} \in C \text{ se verifica } f(x) > f(\bar{x}) + \nabla f(\bar{x})^t(x - \bar{x}), \forall x \in C.$$

3.5.4. Funciones fuertemente cuasiconvexas

La cuasiconvexidad estricta no asegura la unicidad de la solución óptima global. Aquí otra versión de la cuasiconvexidad, llamada cuasiconvexidad fuerte, que asegura la unicidad del mínimo global cuando éste existe.

Definición 3.5.5 Sea S un conjunto convexo no vacío en \mathbb{R}^n , y sea $f : S \rightarrow \mathbb{R}$. Se dice que la función f es **fuertemente cuasiconvexa** si para $x, y \in S$, con $x \neq y$, se tiene

$$f(\lambda x + (1 - \lambda)y) < \max\{f(x), f(y)\} \quad \forall \lambda \in (0, 1)$$

Se dice que la función f es fuertemente cuasicóncava si $-f$ es fuertemente cuasiconvexa.

Se tiene que: estrictamente convexa \Rightarrow fuertemente convexa.

Teorema 3.5.8 Sea $f : S \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$, S convexo. Si f es fuertemente cuasiconvexa en S y \bar{x} es un mínimo local de f en S entonces \bar{x} es el único mínimo global.

Teorema 3.5.9 Sea $f : S \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$, S un conjunto poliédrico compacto no vacío. Si f es cuasiconvexa y continua en S entonces f alcanza el máximo en S y este punto es punto extremo de S .

Teorema 3.5.10 Sea $f : S \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ diferenciable en S abierto convexo. Entonces f es cuasiconvexa en S si y solo si $\forall x, y \in S$

$$f(x) - f(y) \leq 0 \Rightarrow \nabla f(y)^T(x - y) \leq 0$$

Teorema 3.5.11 Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$ fuertemente cuasiconvexo. Se considera el problema de minimizar $f(x)$ sujeto a $x \in S$, donde S es un conjunto convexo no vacío en \mathbb{R}^n . Si \bar{x} es un óptimo local, \bar{x} es la única solución óptima global.

3.5.5. Funciones pseudoconvexas

Las funciones diferenciables fuertemente (o estrictamente) no comparten la propiedad particular de las funciones convexas, que dice que si $\nabla f(\bar{x}) = 0$ en algún punto \bar{x} , \bar{x} es un mínimo global de f .

Definición 3.5.6 Sea $f : S \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ diferenciable en S abierto. Se dice que f es **pseudoconvexa** en S si $\forall x, y \in S$

$$\nabla f(y)^T(x - y) \geq 0 \Rightarrow f(x) - f(y) \geq 0$$

Lo que es equivalente a:

$$f(x) - f(y) < 0 \Rightarrow \nabla f(y)^T(x - y) < 0$$

Analogamente, f es **estrictamente pseudoconvexa** si $\forall x, y \in S$ con $x \neq y$

$$\nabla f(y)^T(x - y) \geq 0 \Rightarrow f(x) - f(y) > 0$$

Lo que es equivalente a:

$$f(x) - f(y) \leq 0 \Rightarrow \nabla f(y)^T(x - y) < 0$$

Para las funciones diferenciables se tiene que:

estrictamente convexas \Rightarrow convexas \Rightarrow pseudoconvexas \Rightarrow explícitamente cuasiconvexas \Rightarrow cuasiconvexas.

Y los recíprocos no son en general ciertos.

Teorema 3.5.12 Sea $f : S \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ pseudoconvexa en S abierto. Entonces todo mínimo local es mínimo global.

Teorema 3.5.13 Sea $f : S \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ pseudoconvexa en S abierto. Entonces si

$$\nabla f(\bar{x})^T(x - \bar{x}) \geq 0 \quad \forall x \in S \Rightarrow \bar{x} \text{ es un mínimo global de } f \text{ en } S.$$

Teorema 3.5.14 Sea $f : S \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ pseudoconvexa en S abierto. Entonces,

$$\bar{x} \text{ es un m\u00ednimo global de } f \text{ en } S \Leftrightarrow \nabla f(\bar{x}) = 0$$

Definici\u00f3n 3.5.7 f es **invex** si y solo si

$$f(x) - f(y) \geq \nabla f(y)^T \eta(x, y), \quad \forall x, y \in S$$

Definici\u00f3n 3.5.8 f es **pseudoinvex** si y solo si

$$f(x) - f(y) < 0 \Rightarrow \nabla f(y)^T \eta(x, y) < 0$$

Teorema 3.5.15 Sea $f : S \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$, f es invex si y solo si f es pseudoinvex.

Teorema 3.5.16 Sea $f : S \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ en S abierto. Entonces, f es *invex* si y solo si

$$\bar{x} \text{ es un m\u00ednimo global de } f \text{ en } S \Leftrightarrow \nabla f(\bar{x}) = 0$$

La figura 3.9 procedente de [2], muestra en la figura (a) una funci\u00f3n pseudoconvexa. A partir de la definici\u00f3n de pseudoconvexidad est\u00e1 claro que si $\nabla f(\bar{x}) = 0$ en cualquier punto \bar{x} , $f(x) \geq f(\bar{x})$ para todo \bar{x} , por lo tanto, \bar{x} es un m\u00ednimo global para f . Por tanto, la funci\u00f3n (c) de la figura no es ni pseudoconvexa ni pseudoc\u00f3ncava. De hecho, la definici\u00f3n afirma que si la derivada direccional de f en cualquier punto x en la direcci\u00f3n $(y - x)$ es no negativa, los valores de la funci\u00f3n no son decrecientes en esa direcci\u00f3n. Adem\u00e1s, la funci\u00f3n pseudoconvexas mostrada en la figura son tambi\u00e9n estrictamente cuasiconvexas, lo que es cierto. La figura (c) no es pseudoconvexa, pero es estrictamente cuasiconvexa.

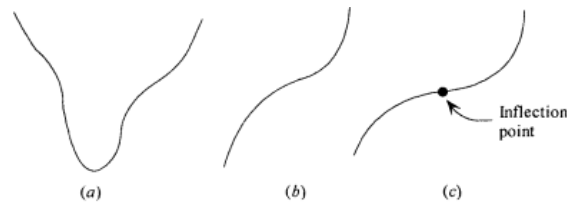


Figura 3.9: Funciones pseudoconvexas y pseudoc\u00f3ncavas: (a) pseudoconvexas, (b) tanto pseudoconvexas como pseudoc\u00f3ncavas, (c) ni pseudoconvexas ni pseudoc\u00f3ncavas

Teorema 3.5.17 Sea S un conjunto convexo abierto no vac\u00edo en \mathbb{R}^n , y sea $f : S \rightarrow \mathbb{R}$ una funci\u00f3n diferenciable sobre S . Entonces f es estrictamente cuasiconvexa y cuasiconvexa.

Teorema 3.5.18 Sea S un conjunto convexo abierto no vac\u00edo en \mathbb{R}^n , y sea $f : S \rightarrow \mathbb{R}$ una funci\u00f3n diferenciable estrictamente pseudoconvexa. Entonces f es fuertemente cuasiconvexa.

3.6. Convexidad en un punto

Otro concepto útil en la optimización la convexidad o concavidad en un punto. En algunos casos, el requisito de una función convexa o cóncava puede ser demasiado fuerte y realmente no es esencial. En cambio, la convexidad o concavidad en un punto puede ser todo lo que se necesita.

Sea S un conjunto convexo no vacío en \mathbb{R}^n , y sea $f : S \rightarrow \mathbb{R}$.

Definición 3.6.1 *Convexidad en \bar{x}* Se dice que la función f es convexa en $\bar{x} \in S$ si

$$f[\lambda\bar{x} + (1 - \lambda)x] \leq \lambda f(\bar{x}) + (1 - \lambda)f(x)$$

para cada $\lambda \in (0, 1)$ y cada $x \in S$.

Definición 3.6.2 *Convexidad estricta en \bar{x}* Se dice que la función f es estrictamente convexa en $\bar{x} \in S$ si

$$f[\lambda\bar{x} + (1 - \lambda)x] < \lambda f(\bar{x}) + (1 - \lambda)f(x)$$

para cada $\lambda \in (0, 1)$ y cada $x \in S$, $x \neq \bar{x}$.

Definición 3.6.3 *Cuasiconvexidad en \bar{x}* Se dice que la función f es cuasiconvexa en $\bar{x} \in S$ si

$$f[\lambda\bar{x} + (1 - \lambda)x] \leq \max\{f(x), f(\bar{x})\}$$

para cada $\lambda \in (0, 1)$ y cada $x \in S$.

Definición 3.6.4 *Cuasiconvexidad estricta en \bar{x}* Se dice que la función es estrictamente cuasiconvexa en $\bar{x} \in S$ si

$$f[\lambda\bar{x} + (1 - \lambda)x] < \max\{f(x), f(\bar{x})\}$$

para cada $\lambda \in (0, 1)$ y cada $x \in S$, $x \neq \bar{x}$.

Definición 3.6.5 *Pseudoconvexidad en \bar{x}* Se dice que la función f es pseudoconvexa en $\bar{x} \in S$ si

$$\nabla f(\bar{x})^t(x - \bar{x}) \geq 0$$

para $x \in S$ implica que $f(x) \geq f(\bar{x})$.

3.7. Ejercicios

■ **Ejemplo 3.4** Estudiar la convexidad de las siguientes funciones.

1. $f(x) = \frac{1}{x}$ para $x > 0$

Solución

f es dos veces derivable en intervalo $(0, +\infty)$

$f'(x) = -\frac{1}{x^2}$, $f''(x) = \frac{2}{x^3}$ con $x > 0 \Rightarrow f''(x) > 0$ en $(0, +\infty) \Rightarrow f$ es **estrictamente convexa** en el dominio.

2. $f(x_1, x_2) = 2x_1^2 + 3x_2^2$ en \mathbb{R}^2

Solución

f es dos veces derivable en el dominio.

$$\nabla f(x_1, x_2) = \begin{pmatrix} 4x_1 \\ 6x_2 \end{pmatrix}$$

$$H_f(x_1, x_2) = \begin{pmatrix} 4 & 0 \\ 0 & 6 \end{pmatrix}$$

$$\Delta_1 = 4 > 0$$

$$\Delta_2 = \begin{vmatrix} 4 & 0 \\ 0 & 6 \end{vmatrix} = 24 > 0$$

$H_f(x_1, x_2)$ es definido positivo en $\mathbb{R}^2 \Rightarrow f$ es estrictamente convexa en el dominio.

3. $f(x_1, x_2) = e^{ax_1+bx_2}$ en \mathbb{R}^2 , $a, b \in \mathbb{R}$

Solución

f es dos veces derivable en el dominio.

$$\nabla f(x_1, x_2) = \begin{pmatrix} ae^{ax_1+bx_2} \\ be^{ax_1+bx_2} \end{pmatrix}$$

$$H_f(x_1, x_2) = \begin{pmatrix} a^2e^{ax_1+bx_2} & abe^{ax_1+bx_2} \\ abe^{ax_1+bx_2} & b^2e^{ax_1+bx_2} \end{pmatrix}$$

Menores principales fundamentales

$$\Delta_1 = a^2e^{ax_1+bx_2} > 0$$

$$\Delta_2 = \det(H_f(x_1, x_2)) = a^2b^2e^{2ax_1+2bx_2} - a^2b^2e^{2ax_1+2bx_2} = 0$$

Por tanto no es ni definida positiva ni negativa.

Estudiamos los menores principales:

$$D_1 = \begin{cases} a^2 e^{ax_1+bx_2} > 0 \\ b^2 e^{ax_1+bx_2} > 0 \end{cases}$$

$$D_2 = \det(H_f(x_1, x_2)) = 0$$

Como $D_1 \geq 0$ y $D_2 \geq 0 \Rightarrow$ es semidefinida positiva $\Rightarrow f$ es convexa en el dominio.

4. $f(x_1, x_2, x_3) = -x_1^2 + 3x_2^2 + x_1x_3$ con $(x_1, x_2, x_3) \in \mathbb{R}^3$.

Solución

f es dos veces derivable en el dominio.

$$\nabla f(x_1, x_2, x_3) = \begin{pmatrix} -2x_1 + x_3 \\ 6x_2 \\ x_1 \end{pmatrix}$$

$$H_f(x_1, x_2, x_3) = \begin{pmatrix} -2 & 0 & 1 \\ 0 & 6 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

Menores principales fundamentales:

$$\Delta_1 = -2$$

$$\Delta_2 = -12$$

$$\Delta_3 = -6$$

Ni definida positiva ni negativa.

Menores principales:

$$D_1 = \begin{cases} -2 \\ 6 \\ 0 \end{cases}$$

Tampoco es semidefinida negativa o positiva, por lo que f no es ni cóncava ni convexa. ■

■ **Ejemplo 3.5** Estudiar la convexidad de los siguientes conjuntos.

1. $S = \{(x_1, x_2) \in \mathbb{R}_+^2 / x_2x_1 \geq 1\}$

Solución

Sea $S = \text{Epi}(f)$, donde $f(x) = \frac{1}{x} \Rightarrow f'(x) = \frac{-1}{x^2} \Rightarrow f''(x) = \frac{2}{x^3} > 0 \Rightarrow f$ convexa $\Rightarrow \text{Epi}(f)$ convexo

2. $S = \{(x, y) \in \mathbb{R}_+^2 / 3x^2 + 2xy + 2y^2 \leq 30\}$

Solución

Entonces $S = S_{30}(f)$, con $f(x, y) = 3x^2 + 2xy + 2y^2$.

$$\nabla f(x, y) = \begin{pmatrix} 6x + 2y \\ 4y + 2x \end{pmatrix}$$

$$H_f(x, y) = \begin{pmatrix} 6 & 2 \\ 2 & 4 \end{pmatrix}$$

$$\Delta_1 = 6 > 0$$

$$\Delta_2 = 24 - 4 = 20 > 0$$

Así $H_f(x, y)$ es definida positiva $\Rightarrow f$ es estrictamente convexa $\Rightarrow f$ convexa \Rightarrow El conjunto de nivel $S_3(f)$ es convexo. ■

■ **Ejemplo 3.6** Encontrar los valores de a tales que el siguiente conjunto sea convexo:

$$S = \{(x, y) \in \mathbb{R}^2 : ax^2 + 2y^2 - x \leq 4, x^4 - 2ay^2 \leq 3\}$$

Solución

Definimos $S = S_4(f) \cap S_3(g)$, $f(x, y) = ax^2 + 2y^2 - x$ y $g(x, y) = x^4 - 2ay^2$

$$\nabla f(x, y) = \begin{pmatrix} 2ax - 1 \\ 4y \end{pmatrix}$$

$$H_f(x, y) = \begin{pmatrix} 2a & 0 \\ 0 & 4 \end{pmatrix}$$

Menores principales fundamentales:

$$\Delta_1 = 2a > 0 \text{ si } a > 0$$

$$\Delta_2 = 8a > 0 \text{ si } a > 0$$

$$\nabla g(x, y) = \begin{pmatrix} 4x^3 \\ -4ay \end{pmatrix}$$

$$H_f(x, y) = \begin{pmatrix} 12x^2 & 0 \\ 0 & -4a \end{pmatrix}$$

Menores principales fundamentales:

$$\Delta_1 = 12x^2 > 0 \text{ si } x \neq 0$$

$$\Delta_2 = -48ax^2 > 0 \text{ si } a < 0$$

Por tanto no se puede afirmar que sean estrictamente convexas para algún valor de a , pero en realidad solo necesitamos que sean convexas, así que calculamos los menores principal.

Menores principales f :

$$D_1 = \begin{cases} 2a \geq 0 \Rightarrow a \geq 0 \\ 4 \geq 0 \end{cases}$$

$$D_2 = 8a \geq 0 \Rightarrow a \geq 0$$

Menores principales g :

$$D_1 = \begin{cases} 12x^2 \geq 0 \\ -4a \geq 0 \Rightarrow a \leq 0 \end{cases}$$

$$D_2 = -48ax^2 \geq 0 \Rightarrow a \leq 0$$

Por tanto, si $a \geq 0$, f es convexa y si $a \leq 0$, g es convexa \Rightarrow si $a = 0$, entonces $S_4(f)$ y $S_3(g)$ son convexos y por tanto su intersección. ■

■ **Ejemplo 3.7** Calcular los máximos y mínimos de las siguientes funciones:

1.- $f(x_1, x_2) = x_1^4 + 8x_1^2 - 4x_2 + x_2^2$

Solución

Calculamos el gradiente y se verán los puntos donde se anula.

$$\begin{aligned} \nabla f(x_1, x_2) &= \begin{pmatrix} 4x_1^3 + 16x_1 \\ 2x_2 - 4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow \\ \begin{cases} 4x_1^3 + 16x_1 = 0 \\ 2x_2 - 4 = 0 \end{cases} &\Rightarrow \begin{cases} 4x_1(x_1^2 + 16) = 0 \\ x_2 = 2 \end{cases} \end{aligned}$$

$$4x_1(x_1^2 + 16) = 0 \Rightarrow x_1 = 0 \Rightarrow \text{Puntos críticos} = \{(0, 2)\}$$

Veamos H_f

$$H_f(x_1, x_2) = \begin{pmatrix} 12x_1^2 + 16 & 0 \\ 0 & 2 \end{pmatrix}$$

$$H_f(0, 2) = \begin{pmatrix} 16 & 0 \\ 0 & 2 \end{pmatrix}$$

$H_f(0, 2)$ definida positiva $\Rightarrow (0, 2)$ es mínimo local estricto.

Por otro lado $H_f(x_1, x_2)$ es definida positiva en \mathbb{R}^2 , lo que significa que f es estrictamente convexa, por lo que $(0, 2)$ es mínimo global estricto.

2.- $f(x_1, x_2) = x_1^2x_2 - x_1x_2$

Solución

$$\begin{aligned} \nabla f(x_1, x_2) &= \begin{pmatrix} 2x_1x_2 - x_2 \\ x_1^2 - x_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow \\ \begin{cases} x_2(2x_1 - 1) = 0 \\ x_1(x_1 - 1) = 0 \end{cases} \end{aligned}$$

De la segunda ecuación se tiene $x_1 = 0$ o $x_1 = 1$.

Si $x_1 = 0 \Rightarrow -x_2 = 0 \Rightarrow x_2 = 0$

Si $x_1 = 1 \Rightarrow x_2 = 0$

Puntos críticos = $\{(0, 0), (1, 0)\}$.

Veamos H_f :

$$H_f(x_1, x_2) = \begin{pmatrix} 2x_2 & 2x_1 - 1 \\ 2x_1 - 1 & 0 \end{pmatrix}$$

Por lo que

$$H_f(0, 0) = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$$

$D_1 = 0$ y $D_2 = -1$, por lo que en ese punto es indefinida y por tanto $(0, 0)$ es un punto de silla.

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

$D_1 = 0$ y $D_2 = -1$, por lo que en ese punto es indefinida y por tanto $(1, 0)$ es un punto de silla.

No existen óptimos.

$$3.- f(x_1, x_2, x_3) = 3x_1^2 + 3x_1x_3 - x_1^2 - x_2^2 - 3x_3^2$$

Solución

$$\nabla f(x_1, x_2, x_3) = \begin{pmatrix} -2x_1 + 3x_3 + 3 \\ -2x_2 \\ -6x_3 + 3x_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow \begin{cases} -2x_1 + 3x_3 + 3 = 0 \\ -2x_2 = 0 \\ -6x_3 + 3x_1 \end{cases}$$

$$x_2 = 0, x_3 = 3, x_1 = 6$$

Punto crítico $(6, 0, 3)$

$$H_f(x, y) = \begin{pmatrix} -2 & 0 & 3 \\ 0 & -2 & 0 \\ 3 & 0 & -6 \end{pmatrix}$$

Menores principales fundamentales:

$$\Delta_1 = -2 < 0$$

$$\Delta_2 = \begin{vmatrix} -2 & 0 \\ 0 & -2 \end{vmatrix} = 4 > 0$$

$$\Delta_3 = \begin{vmatrix} -2 & 0 & 3 \\ 0 & -2 & 0 \\ 3 & 0 & -6 \end{vmatrix} = -24 + 18 - 6 < 0$$

Hessiano definido negativo $\Rightarrow (6, 0, 3)$ es máximo global estricto.

4.- $f(x_1, x_2) = (x_1 - x_2)^4 + (x_2 - 1)^4$

Solución

$$\nabla f(x_1, x_2) = \begin{pmatrix} 4(x_1 - x_2)^3 \\ -(x_1 - x_2)^3 + 4(x_2 - 1)^3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow$$

$$\begin{cases} 4(x_1 - x_2)^3 = 0 \\ -(x_1 - x_2)^3 + 4(x_2 - 1)^3 = 0 \end{cases}$$

$$x_2 = x_1 \text{ y } (x_2 - 1)^3 = 0 \Rightarrow x_2 = 1 \Rightarrow x_1 = 1$$

Punto crítico $(1, 1)$

$$H_f(x_1, x_2) = \begin{pmatrix} 12(x_1 - x_2)^2 & -12(x_1 - x_2)^2 \\ -12(x_1 - x_2)^2 & 12(x_1 - x_2)^2 + 12(x_1 - 1)^2 \end{pmatrix}$$

$$D_1 \begin{cases} 12(x_1 - x_2)^2 \geq 0 \\ -12(x_1 - x_2)^2 + 12(x_1 - 1)^2 \geq 0 \end{cases}$$

$$D_2 = \det(H_f(x_1, x_2)) = 12^2(x_1 - x_2)^4 + (x_1 - x_2)^2(x_2 - 1)^2 - 12^2(x_1 - x_2)^4 = (x_1 - x_2)^2(x_2 - 1)^2 \geq 0$$

Por tanto, el Hessiano es semidefinido positivo en el dominio $\Rightarrow f$ es convexo $\Rightarrow (1, 1)$ es un mínimo global. ■

■ **Ejemplo 3.8** Resolver los siguientes problemas:

1.

$$\text{Max} : x^2 - xy - y^2$$

$$\text{s.a} : x - y = 1$$

Solución

$$x - y = 1 \Leftrightarrow y = x - 1$$

$$\text{Max} \ x^2 - x(x - 1) - (x - 1)^2 \Rightarrow \text{Max} \ x - x^2 - 1 + 2x = -x^2 + 3x - 1$$

$$\nabla f(x) = f'(x) = -2x + 3 = 0 \Rightarrow x = \frac{3}{2} \text{ es el punto crítico.}$$

$$H_f = f''(x) = -2 < 0 \Rightarrow f \text{ es estrictamente cóncava} \Rightarrow \frac{3}{2} \text{ es máximo global estricto.}$$

2.

$$\begin{aligned} \text{Min} : (y - x)^3 + 2(x + y) \\ \text{s.a} : x + y = 2 \end{aligned}$$

Solución

$$y = 2 - x \Rightarrow \text{Min} (2 - x - x)^3 + 2(x + 2 - x) \Rightarrow \text{Min} 2(1 - x)^3 + 4$$

El problema es equivalente a $\text{Min} (1 - x)^3 = g(x)$

$$g'(x) = -3(1 - x)^2 = 0 \Rightarrow 1 - x = 0 \Rightarrow x = 1$$

$g''(x) = -6(1 - x) \Rightarrow g''(1) = 0$, por lo que 1 puede ser punto de silla, máximo o mínimo.

$g'''(x) = -6 \Rightarrow g'''(1) = -6 \neq 0 \Rightarrow 1$ es un punto de silla, pues la primera derivada distinta de cero es impar.

3.

$$\begin{aligned} \text{Min} : x_1^2 - x_2^2 + 2x_2 + x_3^2 \\ \text{s.a} : x_1 + x_2 = 1 \end{aligned}$$

Solución

$$x_1 = 1 - x_2$$

$$f(x_2, x_3) = (1 - x_2)^2 - x_2^2 + 2x_2 + x_3^2 = 1 + x_3^2$$

$$\nabla f(x_2, x_3) = \begin{pmatrix} 0 \\ 2x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow x_3 = 0$$

$$\text{Puntos críticos} = \{(x_2, 0) : x_2 \in \mathbb{R}\}$$

$$H_f(x_2, x_3) = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}$$

$$D_1 \begin{cases} 0 \geq 0 \\ 2 \geq 0 \end{cases}$$

$D_2 = 0 \geq 0 \Rightarrow H_f$ es semidefinida positiva $\Rightarrow f$ es convexa $\Rightarrow \{(x_2, 0) : x_2 \in \mathbb{R}\}$ son mínimos globales. Pero no se deben olvidar las restricciones y se sabe que $x_1 + x_2 = 1$, por lo que $\{(1 - \alpha, \alpha, 0) : \alpha \in \mathbb{R}\}$ son mínimos globales.

4.

$$\begin{aligned} \text{Min} : & x_1^2 + x_2^2 + x_3^2 + x_4^2 \\ \text{s.a} : & x_1 + x_2 + x_3 + x_4 = 0 \end{aligned}$$

$$x_4 = -x_1 - x_2 - x_3$$

$$f(x_1, x_2, x_3, x_4) = x_1^2 + x_2^2 + x_3^2 + (x_1 + x_2 + x_3)^2$$

$$\begin{aligned} \nabla f(x_1, x_2, x_3) &= \begin{pmatrix} 2x_1 + 2(x_1 + x_2 + x_3) \\ 2x_2 + 2(x_1 + x_2 + x_3) \\ 2x_3 + 2(x_1 + x_2 + x_3) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow \\ &\begin{cases} 2x_1 + x_2 + x_3 = 0 \\ x_1 + 2x_2 + x_3 = 0 \\ x_1 + x_2 + 2x_3 = 0 \end{cases} \end{aligned}$$

$$x_1 = x_2 = x_3 = 0$$

$$H_f(x_1, x_2, x_3) = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{pmatrix}$$

$$\Delta_1 = 4 > 0$$

$$\Delta_2 = 16 - 4 > 0$$

$$\Delta_3 = 64 + 8 + 8 - 16 - 16 - 16 = 32 > 0$$

Es definida positivo.

El Hessiano es definido positivo $\Rightarrow f$ estrictamente convexa $\Rightarrow (0, 0, 0)$ es mínimo global único.

■

Capítulo 4

Aplicaciones de programación no lineal

4.1. Conceptos de optimización

Un problema de optimización es, en general, un problema de decisión. A partir del valor de una función, que se llama la función objetivo y que se diseña para cuantificar el rendimiento y medir la calidad de la decisión, se obtendrán valores para cierto número de variables, relacionadas entre sí mediante expresiones matemáticas, de manera que minimicen o maximicen esa función objetivo y, por lo general, teniendo en cuenta una serie de restricciones que limitan la elección de esos valores.

El planteamiento general para resolver problemas de este tipo es, por tanto, el siguiente:

Se desea optimizar $f(x)$

Sujeto a: Restricciones

El empleo del término “optimizar” en la definición incluye los objetivos de minimización o maximización de la función cuando el punto cumple el conjunto de restricciones. En cualquier caso, siempre se puede transformar un problema con miras a la maximización en otro equivalente para su minimización y viceversa, simplemente multiplicando la función objetivo por -1 .

4.1.1. Elementos de la optimización

Para aplicar los conceptos matemáticos y técnicas numéricas necesarias de la teoría de optimización en problemas concretos de Ingeniería, es necesario definir previamente lo que se pretende optimizar. El enunciado general de un problema de programación matemática con restricciones podría ser:

$$\begin{aligned} & \text{Optimizar : } f(x) \\ & \text{sujeta a : } h_i(x) = 0 \quad i = 1, 2, \dots, m \\ & \quad \quad g_j(x) \leq 0 \quad j = 1, 2, \dots, l \\ & \quad \quad x \in \Omega \end{aligned}$$

donde $x^T = (x_1, x_2, \dots, x_n)$ es el vector de las variables independientes, $f(x)$ es la función objetivo, $\Omega \subseteq R^n$ (aunque puede ser cualquier espacio vectorial), $h_i(x)$ son funciones que representan las restricciones de igualdad mientras que las $g_j(x)$ representan el conjunto de las restricciones de desigualdad. El hecho de que solamente aparezcan restricciones del tipo $g_j(x) \leq 0$ y no aparezcan restricciones del tipo $g_i(x) \geq 0$ se debe a que estas últimas pueden transformarse en las primeras multiplicando la desigualdad por -1. En principio, las funciones implícitas en el problema no necesariamente tienen alguna propiedad particular, pero en nuestro caso se van a introducir hipótesis adicionales que nos ayuden a simplificar el problema. Por ejemplo, se supondrá de forma general que las funciones $f(x), h_i(x)$, $g_j(x)$ son continuas y que en la mayoría de los casos tienen derivadas primeras y segundas, también continuas. Además, el conjunto Ω será en la mayoría de los casos un conjunto convexo, aunque generalmente $\Omega = R^n$.

Definiciones seleccionadas del libro [5].

Definición 4.1.1 Una **solución factible** es un vector x que satisface todas las restricciones con el grado de precisión requerido.

Definición 4.1.2 Una **región factible** es la región del espacio n dimensional constituida por todas las soluciones factibles.

Definición 4.1.3 Una **solución óptima** x^* es aquella solución factible que produce el valor óptimo de la función objetivo.

■ **Ejemplo 4.1** Estos ejemplos aparecen en [5].

Dado el siguiente problema de optimización:

$$\text{Minimizar: } f(x) = x_1^3 - 3x_1x_2 + 4$$

$$\text{Sujeta a: } h_1(x) = -2x_1 + x_2 - 5 = 0$$

$$g_1(x) = 5x_1 + 2x_2 - 18 \leq 0$$

En la figura extraída del libro [5] muestra las curvas de nivel de la función objetivo, las líneas rectas muestran las restricciones y la región en amarillo corresponde a la región factible.

La solución óptima al problema es el vector $x^{*T} = (5, 15)$ que produce el valor óptimo de $f(x^*) = -96$.

■

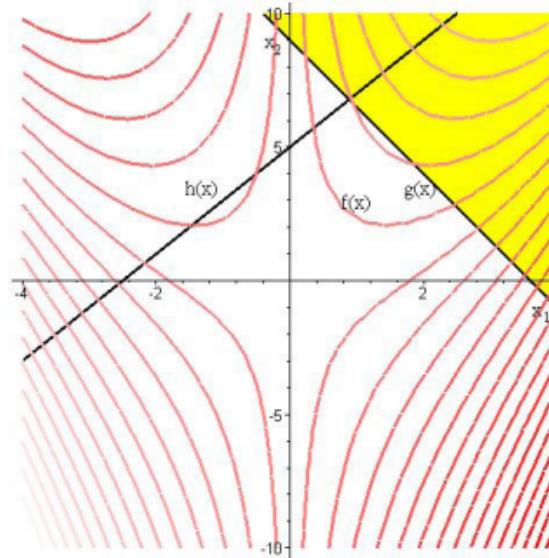


Figura 4.1: La región factible es la línea negra contenida en la región de color amarillo a lo largo de la restricción $h(x)$

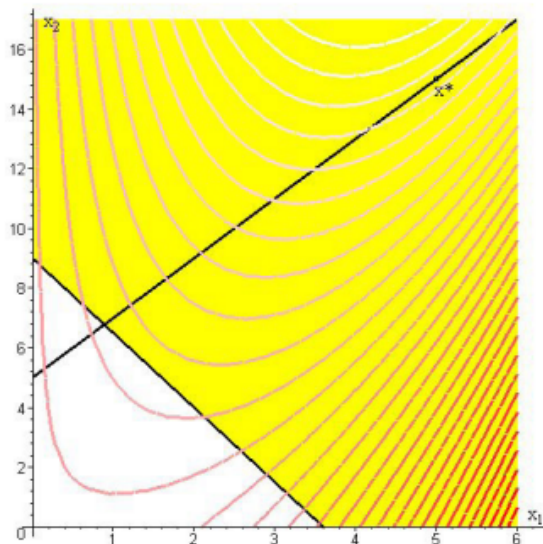


Figura 4.2: El punto óptimo x^* en la región factible

4.2. Métodos Univariable

La mayoría de los métodos seleccionados se encuentran [5].

4.2.1. Métodos con derivadas

Ahora se considerará algunos procedimientos numéricos simples que en forma indirecta localizan el mínimo de una función $f(x)$. Los métodos indirectos son aquellos que determinan los puntos extremos usando sólo condiciones necesarias, derivadas analíticas y valores de la función.

Con estos métodos, básicamente se busca el mínimo de $f(x)$ en un intervalo (a, b) en el cual se sospecha que se encuentra dicho mínimo o bien partiendo de un solo punto x_0 o dos de ellos x_0 y x_1 , lo cuales se encuentran suficientemente cercanos al punto mínimo x^* . La búsqueda se hace determinando las raíces de la función $f'(x) = 0$ en puntos elegidos dentro del intervalo, o bien mediante un punto o puntos elegidos suficientemente cerca del óptimo. Una vez encontradas las raíces de la función $f'(x)$, se usa el criterio de la segunda derivada para determinar si estos valores extremos son máximos o mínimos. En caso de no poder determinar la segunda derivada o que ésta no exista se utilizan puntos vecinos a los valores extremos y se comparan los valores de la función para determinar la naturaleza de los puntos extremos. Debe recordarse que además de la unimodalidad y continuidad en las funciones que se quieren optimizar, se requiere también la derivabilidad de las mismas. A continuación se verán algunos métodos numéricos simples para resolver la ecuación $f'(x) = 0$ y se comentará el error y/o la rapidez de convergencia de estos métodos.

4.2.1.1. Método que usan intervalos

Estos métodos aprovechan el hecho de que una función $f'(x)$ cambie de signo en la vecindad de una raíz. Éstas técnicas se conocen como métodos que usan intervalos porque se necesita de dos valores iniciales para la búsqueda subsecuente de la raíz. Como su nombre lo indica, estos valores deben encerrar a la raíz. Los métodos descritos sobre este punto emplean diferentes estrategias para reducir sistemáticamente el tamaño del intervalo también conocido como intervalo de incertidumbre y así converge en la respuesta correcta. Se verá que en el desarrollo de los métodos que usan intervalos se supuso que en un principio se conoce el intervalo que contiene a la raíz. En la práctica, dicho intervalo normalmente debe localizarse primero mediante cálculos preliminares. Entre los métodos simples disponibles para hacer esto se encuentran *métodos gráficos* y los *métodos de búsqueda incremental*. Como preámbulo a los métodos que usan intervalos, se examinará brevemente los métodos gráficos para obtener las gráficas de las funciones y estimar sus raíces.

4.2.1.2. Métodos de búsqueda incremental

Este método consiste en empezar en una zona de región de interés y realizar evaluaciones de la función con pequeños intervalos a lo largo de la región. Cuando la función cambia de signo se supone que la raíz cae dentro del incremento. Los valores de x en los extremos del intervalo pueden servir de valores iniciales para una de las técnicas que usan intervalos.

Un problema en potencia aunado a los métodos de búsqueda incremental es escoger la longitud del incremento. Si la longitud es muy pequeña, la búsqueda puede consumir demasiado tiempo; por otro lado, si la longitud es muy grande, existe la posibilidad de que las raíces muy cercanas entre sí pasen inadvertidas. El problema se combina con la existencia de raíces múltiples. Aunque el empleo de un incremento muy fino puede solucionar en parte el problema de la acotación de las raíces, debe aclararse que los métodos sencillos, como el de la búsqueda incremental, no son infalibles. Se debe ser prudente al utilizar tales técnicas automáticas con cualquier otra información que provea la visualización en la localización de las raíces (puntos extremos).

4.2.1.3. Método de bisección

Cuando la primera derivada de la función objetivo esta disponible en los puntos extremos de un intervalo de incertidumbre, es necesario evaluar la función en sólo un punto interior para reducir este intervalo. Esto se debe a que es posible decidir si un intervalo acota un mínimo simplemente al observar los valores de la función f_a , f_b y sus derivadas f'_a, f'_b en los puntos extremos a, b .

La condiciones por satisfacer en los puntos a y b son que $f'_a < 0$ y $f'_b > 0$ en la primera figura (2.3) o que $f'_a < 0$, $f_b < 0$ y $f_b > f_a$ como se muestran en la figura 2.4 (derecha o izquierda explicarlo).

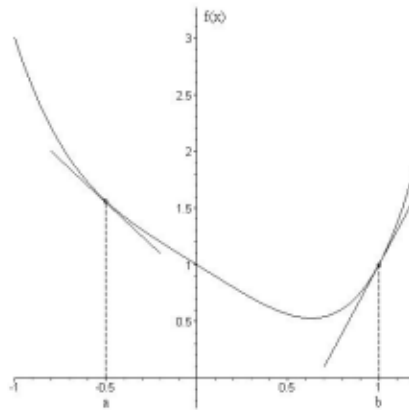


Figura 4.3: Existe un mínimo si ambas pendientes en a y b satisfacen que: $f'_a < 0$ y $f'_b > 0$

O bien que $f'_a > 0$, $f'_b > 0$ y $f_b < f_a$ como se ilustra en la siguiente figura.

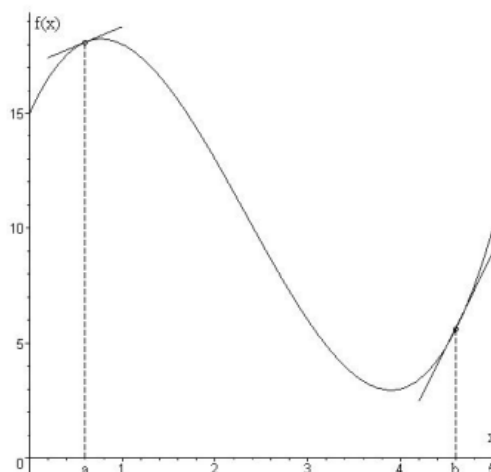


Figura 4.4: Existe un mínimo si en los punto a y b se satisface que: $f'_a > 0$, $f'_b > 0$ y $f_b < f_a$

El método de bisección parte de un intervalo inicial de incertidumbre $[a_0, b_0]$ en el cual se encuentra una de las raíces de la función $f'(x) = 0$; es decir, $f'(a_0)$ y $f'(b_0)$ tienen

signos opuestos. El método consiste en evaluar la función $f'(x)$, continua y derivable, en el punto medio del intervalo $[a_0, b_0]$. Si $f'(a_0)$ y $f'(b_0)$ tienen signos opuestos, se reducirá el intervalo de $[a_0, b_0]$ a $[a_0, x_0]$; si no se satisface la condición de los signos opuestos, entonces se reducirá el intervalo de $[a_0, b_0]$ a $[x_0, b_0]$ ya que la raíz buscada se encuentra dentro de este nuevo intervalo. Se repite este proceso hasta lograr que el intervalo sea más pequeño que una tolerancia prefijada y el último valor x_n será una buena aproximación de la raíz. Este procedimiento tiene la garantía de que converge a la raíz hasta una precisión deseada una vez que la raíz haya sido acotada y que la función sea unimodal en el intervalo de interés. Si la función no es unimodal, la condición $f'(a_0)f'(b_0) < 0$ se satisface siempre que el intervalo tenga un número impar de raíces, en este caso, el método de bisección encontrará una de las raíces separada en el intervalo dado y tal vez será no deseada.

El método de bisección también tiene problemas con las raíces dobles (o múltiples) debido a que estas funciones tocan el eje x de manera tangencial en estas raíces. Otro defecto del método de bisección es que éste puede atrapar una singularidad como si fuera una raíz, debido a que dicho método no reconoce la diferencia entre una raíz y una singularidad (un punto singular es aquel en el que el valor de la función tiende a infinito, por ejemplo que la derivada no está definida en el punto extremo de una función continua). Sin embargo, una gran ventaja es que también es útil para funciones no analíticas; en resumen, se puede decir que es un método robusto.

Algoritmo. Método de bisección

Dados ϵ y una función $f'(x)$ continua sobre el intervalo $[a_0, b_0]$ tal que

$$f'(a_0)f'(b_0) < 0$$

Para $k = 0, 1, 2, \dots$ hasta donde se satisfaga, hacer:

1. $x_k = \frac{a_k + b_k}{2}$
2. Si $|b_k - x_k| \leq \epsilon$ entonces $x_k \approx x^*$ y terminar.
3. Si $f'(a_k)f'(x_k) \leq 0$ entonces

$$a_{k+1} = a_k \text{ y } b_{k+1} = x_k$$

En otro caso,

$$a_{k+1} = x_k \text{ y } b_{k+1} = b_k$$

4. Regresar al paso 1.

Error límite en el método de bisección

Sea a_n, b_n y x_n los n -ésimos valores calculados a partir de a_0, b_0 y x respectivamente; entonces

$$b_n - a_n = \frac{1}{2}(b_{n-1} - a_{n-1})$$

luego

$$b_n - a_n = \frac{1}{2}(b_{n-1} - a_{n-1}) = \frac{1}{2^2}(b_{n-2} - a_{n-2}) = \frac{1}{2^3}(b_{n-3} - a_{n-3})$$

por lo que

$$b_n - a_n = \frac{1}{2^n}(b_0 - a_0), \quad n \geq 0$$

(4.11)

donde $(b_0 - a_0)$ representa la longitud del intervalo original con el que se inició. Ya que la raíz x^* esta entre $[a_n, x_n]$ o $[x_n, b_n]$ se sabe que el error de orden n esta dado por

$$e_n = |x_n - x^*| \leq x_n - a_n = b_n - x_n = \frac{1}{2}(b_{n-1} - a_{n-1})$$

Combinando este resultado con la ecuación 4.11, se tiene

$$e_n = |x_n - x^*| \leq \frac{1}{2^n}(b_0 - a_0)$$

Esta ecuación expresa el error límite y muestra que x_n converge a x^* si $n \rightarrow \infty$. Para mostrar cuántas iteraciones serán necesarias para que x_n se aproxime a x^* con la precisión deseada, se necesita que

$$e_n \leq \epsilon$$

donde $\epsilon > 0$ es el criterio de tolerancia deseado, entonces

$$\frac{1}{2^n}(b_0 - a_0) \leq \epsilon$$

por tanto,

$$n \leq \frac{\ln\left(\frac{b_0 - a_0}{\epsilon}\right)}{\ln 2}$$

que expresa el número máximo de iteraciones para reducir el intervalo inicial a la precisión deseada.

Ejemplo 1

¿Cuántas iteraciones se requieren para resolver un problema dónde el intervalo $[a_0, b_0] = [0, 1]$ y la tolerancia $\epsilon = 0.01$?

Solución

$$n \geq \frac{\ln\left(\frac{b_0 - a_0}{\epsilon}\right)}{\ln 2} = \frac{\ln\left(\frac{1-0}{0.01}\right)}{\ln 2} = 6.64$$

por lo que se requieren $n = 7$ iteraciones para reducir el intervalo original a menos o igual que 0.01 de precisión.

Ejemplo 2

Función que vamos a utilizar para aplicar el método de bisección.

Código R

```

optunider_nores_biseccion = function(dfunx, extizq.ini,extder.ini,
longinc=0.1,report.calculos=T) {
# funx debe ser pseudoconvexa
# longinc > 0
n = 0
an = extizq.ini
bn = extder.ini
mres = NULL
repeat {
lambda.n = (an+bn)/2
fd.lambdan = dfunx(lambda.n)
if (fd.lambdan==0) {
print("Óptimo en lambdan")
vres = c(n+1,an,lambda.n,bn,fd.lambdan,bn-an)
break
} else if (fd.lambdan>0) {
an1 = an
bn1 = lambda.n
} else {
an1 = lambda.n
bn1 = bn
}
vres = c(n+1,an,lambda.n,bn,fd.lambdan,bn-an)
if (report.calculos) {
if (is.null(mres)) {
mres = vres
} else {
mres = rbind(mres,vres)
}
}
if ((bn-an)<=longinc) {
break
}
n = n + 1
an = an1
bn = bn1
}
if (report.calculos) {
colnames(mres) = c("n","a","lambda","b","der_f_lambda","Amplitud")
rownames(mres) = NULL
dfres = as.data.frame(mres)
} else {
names(vres) = c("n","a","lambda","b","der_f_lambda","Amplitud")
}
}

```

```
dfres = as.data.frame(vres)
}
print("Óptimo en lambda")
return(dfres)
}
```

Lo utilizamos sobre la función:

$$f(x) = x^2 + 2x$$

$$f'(x) = 2x + 2$$

$$[a_0, b_0] = [-2, 1]$$

$$\epsilon = 0.1$$

Solución

Para poder realizar estas tablas con las funciones creadas se usan las siguientes librerías de R:

```
[9] [10] [11] [12]
```

En todos los métodos siguientes se usan las mismas librerías para poder construir las tablas.

```
library(kableExtra)
library(tidyr)
library(dplyr)
library(knitr)
```

```
dfunx = function(x) 2*x+2
df.res = optunider_nores_biseccion(dfunx,-2,1)
```

```
## [1] "Óptimo en lambda"
```

```
kable(df.res,digits = 6,booktabs=TRUE,longtable=TRUE,
      caption = "\\label{tabla01}Ejercicio 2 - Método de bisección") %>%
kable_styling(latex_options =c("striped","HOLD_position","repeat_header"),
              position = "center",repeat_header_text = "contin\\'ua")
```

Tabla 4.1: Ejercicio 2 - Método de bisección

n	a	lambda	b	der_f_lambda	Amplitud
1	-2.0000	-0.500000	1.00000	1.00000	3.00000
2	-2.0000	-1.250000	-0.50000	-0.50000	1.50000
3	-1.2500	-0.875000	-0.50000	0.25000	0.75000
4	-1.2500	-1.062500	-0.87500	-0.12500	0.37500
5	-1.0625	-0.968750	-0.87500	0.06250	0.18750

Tabla 4.1: Ejercicio 2 - Método de bisección continúa

n	a	lambda	b	der_f_lambda	Amplitud
6	-1.0625	-1.015625	-0.96875	-0.03125	0.09375

Ahora sin iteraciones

```
dfunx = function(x) 2*x+2
df.res2 = optunider_nores_biseccion(dfunx,-2,1,report.calculos = F)
```

```
## [1] "Óptimo en lambda"
```

```
knitr::kable(df.res2,digits = 6, booktabs=TRUE,longtable=TRUE,
             caption = "\\label{tabla01}Ejercicio 2 - Método
             de bisección solución") %>%
kable_styling(latex_options =c("striped","HOLD_position","repeat_header"),
             position = "center",repeat_header_text = "contin\\'ua")
```

Tabla 4.2: Ejercicio 2 - Método de bisección solución

	vres
n	6.000000
a	-1.062500
lambda	-1.015625
b	-0.968750
der_f_lambda	-0.031250
Amplitud	0.093750

Se puede ver como evoluciona el algoritmo gráficamente:

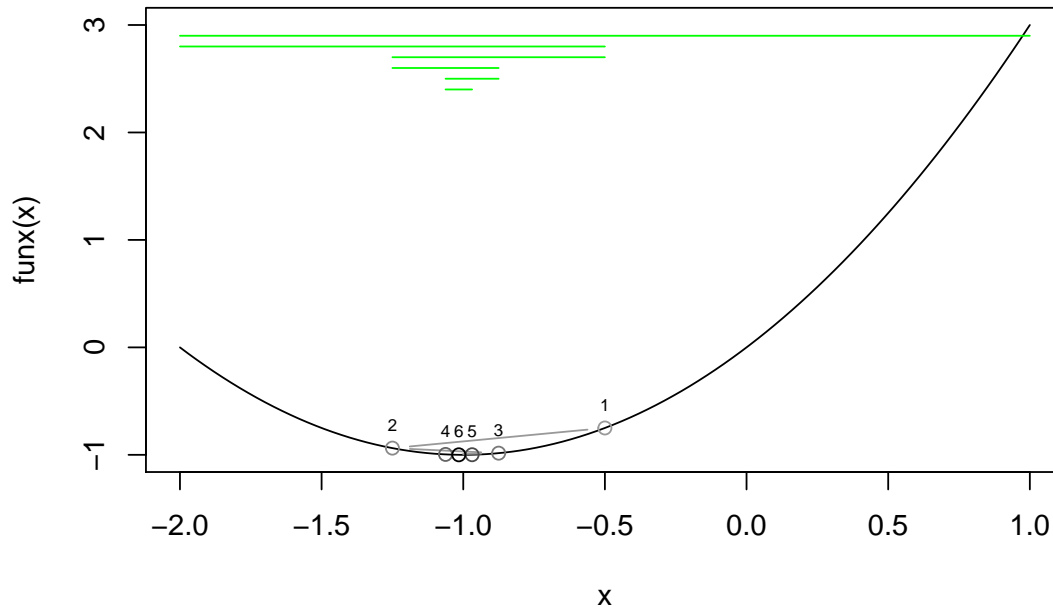


Figura 4.5: Gráfico - Método bisección ejercicio 2

Una vez hecho en el código R, vamos hacerlo en **Excel** y comparamos resultados:

	A	B	C	D	E	F	G	H	I
1	n	a_n	λ_n	b_n	$f(\lambda_n)$	longitud_intervalo	¿parar?	l	
2	1	-2	-0,5	1	1	3	FALSO	0,1	
3	2	-2	-1,25	-0,5	-0,5	1,5	FALSO		
4	3	-1,25	-0,875	-0,5	0,25	0,75	FALSO		
5	4	-1,25	-1,0625	-0,875	-0,125	0,375	FALSO		
6	5	-1,0625	-0,96875	-0,875	0,0625	0,1875	FALSO		
7	6	-1,0625	-1,015625	-0,96875	-0,03125	0,09375	VERDADERO		
8									

Figura 4.6: Excel - Método bisección ejercicio 2

Como se puede ver tanto en Excel como en R nos dan los mismo resultados por tanto se puede iniciar las conclusiones del método.

Describimos las tablas que nos aparecen en la salida R y Excel:

- **n** → Iteraciones.
- **a_n** → Valor del intervalo que se encuentra en la parte inferior.
- **b_n** → Valor del intervalo que se encuentra en la parte superior.
- λ_n → Valores para los cuales van las raíces de la ecuación.
- $f(\lambda_n)$ → valor de λ_n en la ecuación.
- **longitud_intervalo** → Longitud del intervalo.

- **¿parar?** → Valor Falso o Verdadero, nos va indicar cuando se puede parar.
- **l** → Valor del ϵ , que es un valor prefijado donde la longitud del intervalo tiene que ser menor, para poder parar.

Partimos de un intervalo inicial que es $[-2, 1]$, en el cual se encuentra una de las raíces de la función $f'(x)=0$. Vamos a evaluar la función $f'(x)$, continua y derivable, en el punto medio del intervalo. Si $f'(-2)$ y $f'(1)$ tienen signos opuestos, y se reducirá el intervalo de $[-2, 1]$ a $[-2, x_0]$. Se repetirá este proceso hasta lograr que el intervalo sea más pequeño que ϵ que es 0.1 y el último valor de λ será una buena aproximación de raíz.

Entonces, se puede concluir que el intervalo que contiene a la raíz es $[-1, 0625, -0.96875]$, la aproximación de la raíz es $\lambda = -1.015625$, en la iteración 6, con una amplitud de intervalo de 0,09375 que es menor que nuestro $\epsilon = 0.1$.

Ejemplo 3

Muestre los resultados parciales del algoritmo de la bisección sabiendo:

$$\text{Función } f(x) = x^3 + 4x^2 - 10 = 0$$

$$\text{Intervalo } [a_0, b_0] = [1, 2]$$

$$\epsilon = 0.0001$$

```

optunider_nores_biseccion = function(dfunx, extizq.ini,extder.ini,
longinc=0.0001,report.calculos=T) {
# funx debe ser pseudoconvexa
# longinc > 0
n = 0
an = extizq.ini
bn = extder.ini
mres = NULL
repeat {
lambda.n = (an+bn)/2
fd.lambdan = dfunx(lambda.n)
if (fd.lambdan==0) {
print("Óptimo en lambdan")
vres = c(n+1,an,lambda.n,bn,fd.lambdan,bn-an)
break
} else if (fd.lambdan>0) {
an1 = an
bn1 = lambda.n
} else {
an1 = lambda.n
bn1 = bn
}
vres = c(n+1,an,lambda.n,bn,fd.lambdan,bn-an)
if (report.calculos) {
if (is.null(mres)) {
mres = vres
} else {
mres = rbind(mres,vres)
}
}
}

```

```

}
}
if ((bn-an)<=longinc) {
break
}
n = n + 1
an = an1
bn = bn1
}
if (report.calculos) {
colnames(mres) = c("n", "a", "lambda", "b", "der_f_lambda", "Amplitud")
rownames(mres) = NULL
dfres = as.data.frame(mres)
} else {
names(vres) = c("n", "a", "lambda", "b", "der_f_lambda", "Amplitud")
dfres = as.data.frame(vres)
}
print("Óptimo en lambda")
return(dfres)
}

```

Solución

La ecuación mostrada tiene una raíz en $[1, 2]$, ya que $f(1) = -5$ y $f(2) = 14$ y existe cambio de signo.

$$f'(x) = 3x^2 + 8x$$

```

dfunx = function(x) 3*x^2+8*x
df.res = optunider_nores_biseccion(dfunx,1,2)

```

```
## [1] "Óptimo en lambda"
```

```

knitr::kable(df.res,digits = 6, booktabs = TRUE,longtable=TRUE,
caption = "\\label{tabla01}Ejercicio 3 - Método de bisección") %>%
kable_styling(latex_options =c("striped","HOLD_position","repeat_header"),
position = "center",repeat_header_text = "contin\\'ua")

```

Tabla 4.3: Ejercicio 3 - Método de bisección

n	a	lambda	b	der_f_lambda	Amplitud
1	1	1.500000	2.000000	18.75000	1.000000
2	1	1.250000	1.500000	14.68750	0.500000
3	1	1.125000	1.250000	12.79688	0.250000
4	1	1.062500	1.125000	11.88672	0.125000
5	1	1.031250	1.062500	11.44043	0.062500

Tabla 4.3: Ejercicio 3 - Método de bisección continúa

n	a	lambda	b	der_f_lambda	Amplitud
6	1	1.015625	1.031250	11.21948	0.031250
7	1	1.007812	1.015625	11.10956	0.015625
8	1	1.003906	1.007812	11.05473	0.007812
9	1	1.001953	1.003906	11.02736	0.003906
10	1	1.000977	1.001953	11.01367	0.001953
11	1	1.000488	1.000977	11.00684	0.000977
12	1	1.000244	1.000488	11.00342	0.000488
13	1	1.000122	1.000244	11.00171	0.000244
14	1	1.000061	1.000122	11.00085	0.000122
15	1	1.000031	1.000061	11.00043	0.000061

Ahora sin iteraciones:

```
# Sin información de las iteraciones
dfunx = function(x) 3*x^2+8*x
df.res2 = optunider_nores_biseccion(dfunx,1,2,report.calculos = F)
```

```
## [1] "Óptimo en lambda"
```

```
knitr::kable(df.res2,digits = 6 , booktabs=TRUE,longtable=TRUE,
  caption = "\\label{tabla01}Ejercicio 3
  - Método de bisección solución") %>%
kable_styling(latex_options =c("striped","HOLD_position","repeat_header"),
  position = "center",repeat_header_text = "contin\\'ua")
```

Tabla 4.4: Ejercicio 3 - Método de bisección solución

	vres
n	15.000000
a	1.000000
lambda	1.000031
b	1.000061
der_f_lambda	11.000427
Amplitud	0.000061

	A	B	C	D	E	F	G	H	I
38									
39	<i>n</i>	<i>a_n</i>	<i>λ_n</i>	<i>b_n</i>	<i>f(λ_n)</i>	<i>longitud intervalo</i>	<i>¿parar?</i>	<i>l</i>	
40	1	1	1,5	2	18,75	1	FALSO	0,0001	
41	2	1	1,25	1,5	14,6875	0,5	FALSO		
42	3	1	1,125	1,25	12,796875	0,25	FALSO		
43	4	1	1,0625	1,125	11,8867188	0,125	FALSO		
44	5	1	1,03125	1,0625	11,4404297	0,0625	FALSO		
45	6	1	1,015625	1,03125	11,2194824	0,03125	FALSO		
46	7	1	1,0078125	1,015625	11,1095581	0,015625	FALSO		
47	8	1	1,00390625	1,0078125	11,0547333	0,0078125	FALSO		
48	9	1	1,00195313	1,00390625	11,0273552	0,00390625	FALSO		
49	10	1	1,00097656	1,00195313	11,0136747	0,001953125	FALSO		
50	11	1	1,00048828	1,00097656	11,0068367	0,000976563	FALSO		
51	12	1	1,00024414	1,00048828	11,0034181	0,000488281	FALSO		
52	13	1	1,00012207	1,00024414	11,001709	0,000244141	FALSO		
53	14	1	1,00006104	1,00012207	11,0008545	0,00012207	FALSO		
54	15	1	1,00003052	1,00006104	11,0004272	6,10352E-05	VERDADERO		
55									

Figura 4.7: Excel - Método bisección ejercicio 3

Como se puede ver ambos resultados coinciden también, por lo que se pueden iniciar las conclusiones.

Partimos de un intervalo inicial que es $[1, 2]$, en el cual se encuentra una de las raíces de la función $f'(x)=0$. Vamos a evaluar la función $f'(x)$, continua y derivable, en el punto medio del intervalo. Si $f'(1)$ y $f'(2)$ como tienen el mismo signo se reducirá el intervalo de $[1, 2]$ a $[1, x_0]$. Se repetirá este proceso hasta lograr que el intervalo sea más pequeño que ϵ que es 0.0001 y el último valor de λ será una buena aproximación de raíz.

Entonces, se puede concluir que el intervalo que contiene a la raíz es $[1, 1.00006104]$, la aproximación de la raíz es $\lambda = 1.00003052$, en la iteración 15, con una amplitud de intervalo de 0.00061035 que es menor que nuestro $\epsilon = 0.0001$.

Se puede ver como evoluciona el algoritmo gráficamente:

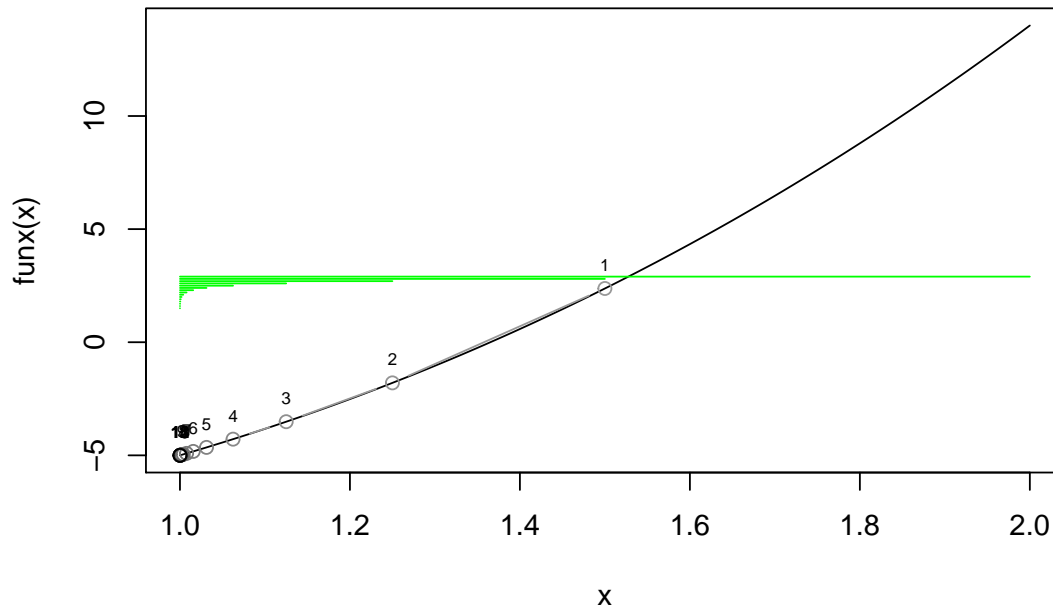


Figura 4.8: Gráfico - Método bisección ejercicio 3

Ejemplo 4

Aproximar la única raíz real de la siguiente función sabiendo $[a_0, b_0] = [-1, 1]$ con $\epsilon = 10^{-4}$

$$f(x) = x^3 - \frac{x^2}{2} - 3x + 2$$

```

optunider_nores_biseccion = function(dfunx, extizq.ini,extder.ini,
longinc=10^-4,report.calculos=T) {
# funx debe ser pseudoconvexa
# longinc > 0
n = 0
an = extizq.ini
bn = extder.ini
mres = NULL
repeat {
lambda.n = (an+bn)/2
fd.lambdan = dfunx(lambda.n)
if (fd.lambdan==0) {
print("Óptimo en lambdan")
vres = c(n+1,an,lambda.n,bn,fd.lambdan,bn-an)
break
} else if (fd.lambdan>0) {
an1 = an

```

```

bn1 = lambda.n
} else {
an1 = lambda.n
bn1 = bn
}
vres = c(n+1,an,lambda.n,bn,fd.lambdan,bn-an)
if (report.calculos) {
if (is.null(mres)) {
mres = vres
} else {
mres = rbind(mres,vres)
}
}
if ((bn-an)<=longinc) {
break
}
n = n + 1
an = an1
bn = bn1
}
if (report.calculos) {
colnames(mres) = c("n","a","lambda","b","der_f_lambda","Amplitud")
rownames(mres) = NULL
dfres = as.data.frame(mres)
} else {
names(vres) = c("n","a","lambda","b","der_f_lambda","Amplitud")
dfres = as.data.frame(vres)
}
print("Óptimo en lambda")
return(dfres)
}

```

Solución

$$f'(x) = 3x^2 - x - 3$$

```

dfunx = function(x) 3*x^2-x-3
df.res = optunider_nores_biseccion(dfunx,-1,1)

```

```
## [1] "Óptimo en lambda"
```

```

knitr::kable(df.res,digits = 6,booktabs = TRUE,longtable=TRUE,
             caption = "\\label{tabla01}Ejercicio 4 - Método de bisección") %>%
kable_styling(latex_options =c("striped","HOLD_position","repeat_header"),
             position = "center",repeat_header_text = "contin\\'ua")

```

Tabla 4.5: Ejercicio 4 - Método de bisección

n	a	lambda	b	der_f_lambda	Amplitud
1	-1.000000	0.000000	1	-3.000000	2.000000
2	0.000000	0.500000	1	-2.750000	1.000000
3	0.500000	0.750000	1	-2.062500	0.500000
4	0.750000	0.875000	1	-1.578125	0.250000
5	0.875000	0.937500	1	-1.300781	0.125000
6	0.937500	0.968750	1	-1.153320	0.062500
7	0.968750	0.984375	1	-1.077393	0.031250
8	0.984375	0.992188	1	-1.038879	0.015625
9	0.992188	0.996094	1	-1.019485	0.007812
10	0.996094	0.998047	1	-1.009754	0.003906
11	0.998047	0.999023	1	-1.004880	0.001953
12	0.999023	0.999512	1	-1.002441	0.000977
13	0.999512	0.999756	1	-1.001221	0.000488
14	0.999756	0.999878	1	-1.000610	0.000244
15	0.999878	0.999939	1	-1.000305	0.000122
16	0.999939	0.999969	1	-1.000153	0.000061

Ahora sin iteraciones

```
dfunx = function(x) 3*x^2-x-3
df.res2 = optunider_nores_biseccion(dfunx,-1,1,report.calculos = F)
```

```
## [1] "Óptimo en lambda"
```

```
knitr::kable(df.res2,digits = 6, booktabs=TRUE,longtable=TRUE,
  caption = "\\label{tabla01}Ejercicio 4
  - Método de bisección solución") %>%
kable_styling(latex_options =c("striped","HOLD_position","repeat_header"),
  position = "center",repeat_header_text = "contin\\'ua")
```

Tabla 4.6: Ejercicio 4 - Método de bisección solución

	vres
n	16.000000
a	0.999939
lambda	0.999969
b	1.000000
der_f_lambda	-1.000153
Amplitud	0.000061

Introducimos el ejemplo en Excel:

	A	B	C	D	E	F	G	H	I
17									
18									
19	<i>n</i>	<i>a_n</i>	<i>λ_n</i>	<i>b_n</i>	<i>f'(λ_n)</i>	<i>longitud intervalo</i>	<i>¿parar?</i>	<i>l</i>	
20	1	-1	0	1	-3	2	FALSO	0,0001	
21	2	0	0,5	1	-2,75	1	FALSO		
22	3	0,5	0,75	1	-2,0625	0,5	FALSO		
23	4	0,75	0,875	1	-1,578125	0,25	FALSO		
24	5	0,875	0,9375	1	-1,3007813	0,125	FALSO		
25	6	0,9375	0,96875	1	-1,1533203	0,0625	FALSO		
26	7	0,96875	0,984375	1	-1,0773926	0,03125	FALSO		
27	8	0,984375	0,9921875	1	-1,0388794	0,015625	FALSO		
28	9	0,9921875	0,99609375	1	-1,0194855	0,0078125	FALSO		
29	10	0,99609375	0,99804688	1	-1,0097542	0,00390625	FALSO		
30	11	0,99804688	0,99902344	1	-1,00488	0,001953125	FALSO		
31	12	0,99902344	0,99951172	1	-1,0024407	0,000976563	FALSO		
32	13	0,99951172	0,99975586	1	-1,0012205	0,000488281	FALSO		
33	14	0,99975586	0,99987793	1	-1,0006103	0,000244141	FALSO		
34	15	0,99987793	0,99993896	1	-1,0003052	0,00012207	FALSO		
35	16	0,99993896	0,99996948	1	-1,0001526	6,10352E-05	VERDADERO		
36									

Figura 4.9: Excel - Método de bisección ejercicio 4

Como se puede ver ambos resultados coinciden también, por tanto se pueden iniciar las conclusiones.

Partimos de un intervalo inicial que es $[-1, 1]$, en el cual se encuentra una de las raíces de la función $f'(x)=0$. Vamos a evaluar la función $f'(x)$, continua y derivable, en el punto medio del intervalo. Si $f'(-1)$ y $f'(1)$ como tienen el distinto signo se reducirá el intervalo de $[-1, 1]$ a $[x_0, 1]$. Se repetirá este proceso hasta lograr que el intervalo sea más pequeño que ϵ que es 0.0001 y el último valor de λ será una buena aproximación de raíz.

Entonces, se puede concluir que el intervalo que contiene a la raíz es $[0.999939, 1]$, la aproximación de la raíz es $\lambda = 0.999969$, en la iteración 16, con una amplitud de intervalo de 0.000061 que es menor que nuestro $\epsilon = 0.0001$.

Se puede ver como evoluciona el algoritmo gráficamente:

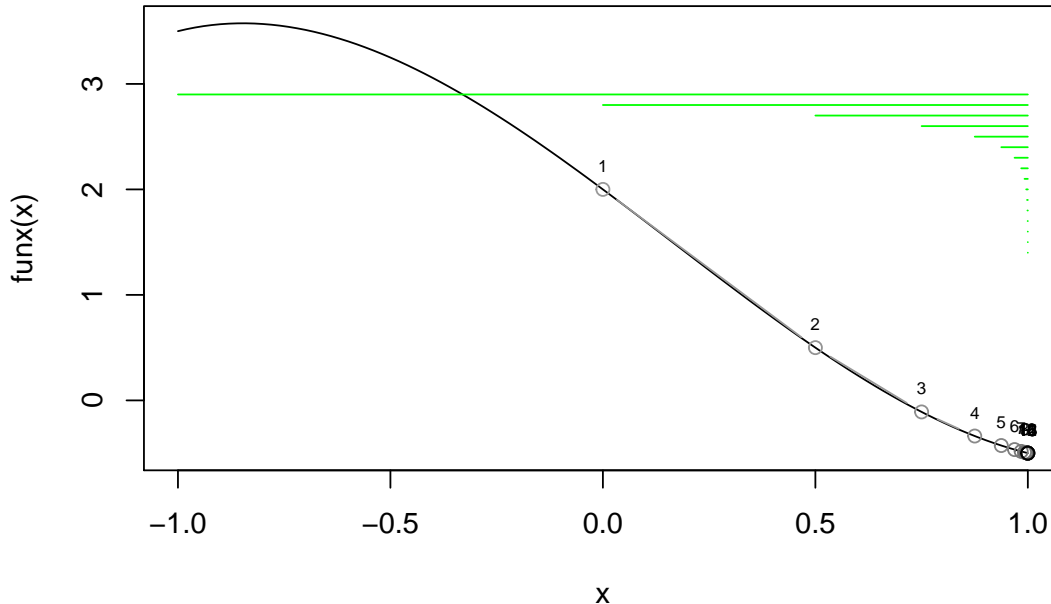


Figura 4.10: Gráfico - Método bisección ejercicio 4

4.2.1.4. Métodos abiertos

En contraste con los métodos anteriores, los métodos abiertos se basan en fórmulas que requieren de un solo valor x o un par de ellos, pero que no necesariamente encierran la raíz x^* a medida que crece el número de iteraciones; sin embargo, cuando éstos convergen, en general lo hacen mucho más rápido que los métodos que usan intervalos, entre estos métodos se encuentra el método de la secante, Newton-Raphson, Newton modificado, Halley, Chebyshev...

4.2.1.4.1. Método de Newton-Raphson

En esencia, la técnica de Newton-Raphson consiste en el uso de rectas tangentes; es por esto que al método se le conoce también como método de la tangente. Esta estrategia permite usar únicamente un punto inicial cercano a la raíz, sin requerir cosas adicionales dentro de la vecindad donde se encuentra.

Por lo general, se inicia con una estimación de x^* que se denota con x_0 , para mejorar esta estimación, considérese la línea recta que es tangente a la gráfica de $f'(x)$ en el punto $(x_0, f'(x_0))$, si x_0 está cerca de x^* , esta línea tangente casi coincidirá con la gráfica de $f'(x)$ para puntos x alrededor de x^* . Entonces si x_1 es la raíz de la línea tangente, x_1 será casi igual a x^* . En general, para determinar una fórmula para x_{n+1} , considérese la pendiente de la línea tangente. Usando la derivada $f''(x)$ de $f'(x)$, se sabe del cálculo elemental que a pendiente de la tangente en $(x_n, f'(x_n))$ es $f''(x_n)$, esto guía a la ecuación

$$\tan \theta_n = f''(x_n) = \frac{0 - f'(x_n)}{x_{n+1} - x_n}$$

y, en consecuencia,

$$x_{n+1} = x_n - \frac{f'(x)}{f''(x)}, \quad n = 0, 1, 2, \dots$$

El método de Newton-Raphson utiliza en forma iterativa las rectas tangentes que pasan por las aproximaciones consecutivas de la raíz, además, este método requiere una buena estimación inicial, pues de otro modo la solución iterativa puede divergir o converger a una solución irrelevante. La razón de convergencia iterativa del método de Newton-Raphson es alta cuando funciona. La desventaja de este método se da en el caso de abordar raíces simples se puede tener problemas cuando $f''(x_n) \rightarrow 0$, lo que implica que la raíz puede estar cerca de un punto de inflexión. En resumen, se puede decir que la convergencia del método es muy dependiente de la naturaleza de la función y del valor de la aproximación inicial.

Algoritmo. Método de Newton-Raphson

Dados ϵ , un punto inicial x_0 y una función $f'(x)$ continua y diferenciable. Para $k = 0, 1, 2, \dots$ hasta donde se satisfaga, hacer:

1. $x_{k+1} = x_k - \frac{f'(x_k)}{f''(x_k)}$
2. Si $|x_{k+1} - x| \leq \epsilon$ o $|f'(x_{k+1})| \leq \epsilon$ entonces $x_{k+1} \approx x^*$ y terminar.
3. Regresar al paso 1.

Rapidez de convergencia del método de Newton-Raphson

Suponiendo que $f'(x)$ tiene derivadas continuas por todo x en algún intervalo alrededor de la raíz x^* , por el teorema de Taylor se puede escribir

$$f'(x) = f'(x_n) + (x^* - x_n)f''(x_n) + \frac{1}{2}(x^* - x_n)^2 f'''(x_n) + \dots$$

obsérvese que $f'(x^*) = 0$ por suposición, y luego se divide por $f''(x_n)$ para obtener

$$0 = \frac{f'(x_n)}{f''(x_n)} + (x^* - x_n) + (x^* - x_n)^2 \frac{f'''(x_n)}{2f''(x_n)} + \dots$$

Como

$$\frac{f'(x_n)}{f''(x_n)} = x_n - x_{n+1}$$

entonces

$$0 = x_n - x_{n+1} + x^* - x_n + (x^* - x_n)^2 \frac{f'''(x_n)}{2f''(x_n)} + \dots$$

Resolviendo para $e_{n+1} = |x_{n+1} - x^*|$ se tiene

$$e_{n+1} = |x_{n+1} - x^*| = |(x_n - x^*) \frac{f'''(x_n)}{2f''(x_n)} + \dots| \approx |(x_n - x^*)^2 \frac{f'''(x^*)}{2f''(x_n)}|$$

donde $M = \left| \frac{f'''(x^*)}{2f''(x_n)} \right|$. Tomando el límite cuando $n \rightarrow \infty$, resulta

$$\lim_{n \rightarrow \infty} \frac{e_{n+1}}{e_n^2} = \lim_{n \rightarrow \infty} \frac{|x_{n+1} - x^*|}{|x_n - x^*|^2} = \left| \frac{f'''(x^*)}{2f''(x^*)} \right| = M^*$$

en consecuencia si $M^* < \infty$ el método de Newton-Raphson converge cuadráticamente a x^* .

Código R

En este código se usan a parte de las librerías anteriormente citadas, [3].

```
library("pracma")
library("knitr")

newtonuni <-function(funcion, x0, epsilon = 0.01) {
  #Definimos la funcion con los valores: funcion, x0, epsilon.
  #funcion: sera la funcion que queremos optimizar
  #x0: sera el punto inicial
  #epsilon: definimos un epsilon que por defecto sera 0.01

  k <-0 #Primera iteracion
  mres <-NULL #Definimos una tabla nula en la que vamos a introducir
  #los valores
  repeat {
    k=k+1
    #Calculo de la primera derivada y de la segunda derivadas necesarias
    #para el calculo de los nuevos puntos que obtenemos en el m?todo
    derivada.k <-fderiv(funcion, x = x0, n = 1, method = "central")
    derivada2.k <-fderiv(funcion, x = x0, n = 2, method = "central")
    #Calculo de los puntos de cada itetacion
    x1 <- x0 - derivada.k / derivada2.k
    #Ponemos todos los datos calculados en un mismo vector
    vres <-c(k, x0, derivada.k, derivada2.k, x1, abs(x1 -x0))

    if (is.null(mres)) {
      mres <-vres
    } else {
      mres <-rbind(mres, vres)
    }
    #Fijamos cuando termina el metodo
    if (abs(x1 -x0) < epsilon) {
      break
    } else {
      x0 <-x1
    }
  }
}
```



```

    }
  }
  colnames(mres) <-c("k", "x_k", "f'(x_k)", "f''(x_k)",
                    "x_(k+1)", "|x_(k+1)-x_k|")
  rownames(mres) <-NULL
  dfres <-as.data.frame(mres)

  print("Optimo en x_(k+1)")
  return(dfres)
}

```

Ejemplo 1

$$g(x) = \frac{80}{x} + 20x + 20$$

```

g <-function(x) (80/x) + 20*x + 20
df.res <-newtonuni(g, 1, 0.00001)

```

```
## [1] "Optimo en x_(k+1)"
```

```

kable(df.res, digits = 6, booktabs =TRUE, longtable=TRUE,
      caption = "\\label{tabla01}Ejercicio 1
      - Método de Newton Raphson") %>%
kable_styling(latex_options =c("striped","HOLD_position","repeat_header"),
              position = "center",repeat_header_text = "contin\\'ua")

```

Tabla 4.7: Ejercicio 1 - Método de Newton Raphson

k	x_k	f'(x_k)	f''(x_k)	x_(k+1)	x_(k+1)-x_k
1	1.000000	-60.000000	160.00000	1.375000	0.375000
2	1.375000	-22.314050	61.54771	1.737549	0.362549
3	1.737549	-6.498175	30.50064	1.950599	0.213050
4	1.950599	-1.025865	21.55836	1.998185	0.047585
5	1.998185	-0.036355	20.05455	1.999998	0.001813
6	1.999998	-0.000049	20.00007	2.000000	0.000002

Gráficamente:

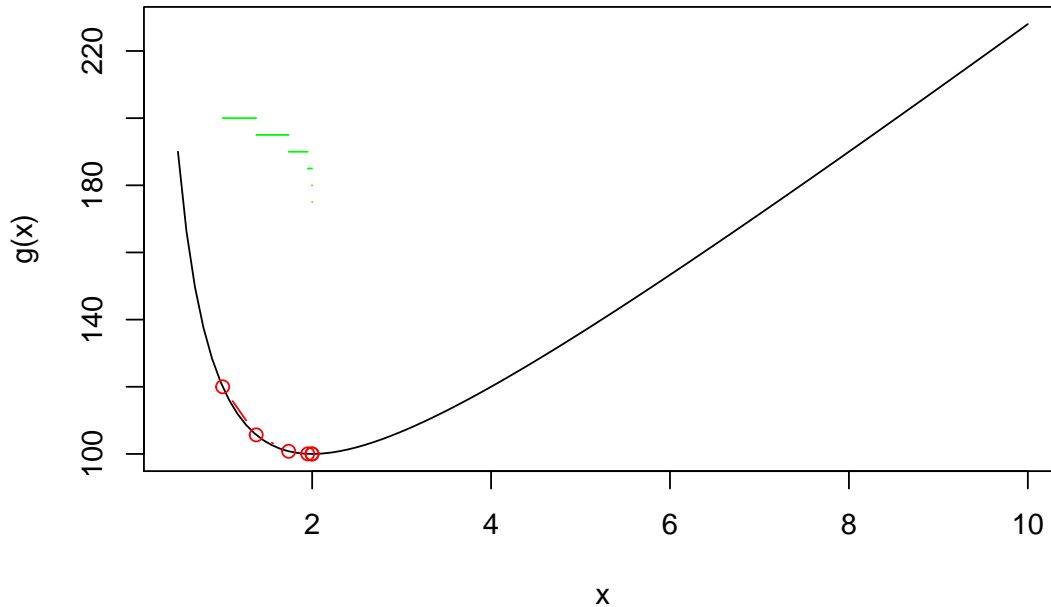


Figura 4.11: Gráfico - Método Newton Raphson ejercicio 1

Para $x = 2$, el valor óptimo de $f(x)$ será: 100.

Ejemplo 2

$$h(x) = 2x^4 - 7x^3 + 4x - 2$$

```
h <-function(x) 2*x^4 + -7*x^3+4*x-2
df.res3 <-newtonuni(h, -1, 0.00001)
```

```
## [1] "Optimo en x_(k+1)"
```

```
kable(df.res3, digits = 6, booktabs=TRUE,longtable=TRUE,
      caption = "\\label{tabla01}Ejercicio 2
      - Método Newton Raphson") %>%
kable_styling(latex_options =c("striped","HOLD_position","repeat_header"),
              position = "center",repeat_header_text = "contin\\'ua")
```

Tabla 4.8: Ejercicio 2 - Método Newton Raphson

k	x_k	$f'(x_k)$	$f''(x_k)$	$x_{(k+1)}$	$ x_{(k+1)}-x_k $
1	-1.000000	-25.000000	66.000000	-0.621212	0.378788
2	-0.621212	-6.021823	35.35262	-0.450876	0.170336

Tabla 4.8: Ejercici continúa

k	x_k	$f'(x_k)$	$f''(x_k)$	$x_{(k+1)}$	$ x_{(k+1)}-x_k $
3	-0.450876	-1.002341	23.81574	-0.408789	0.042087
4	-0.408789	-0.055770	21.17973	-0.406156	0.002633
5	-0.406156	-0.000213	21.01763	-0.406145	0.000010
6	-0.406145	0.000000	21.01701	-0.406145	0.000000

Para $x = -0.4061$, el valor óptimo de $h(x)$ sera: -3.10119

Gráficamente:

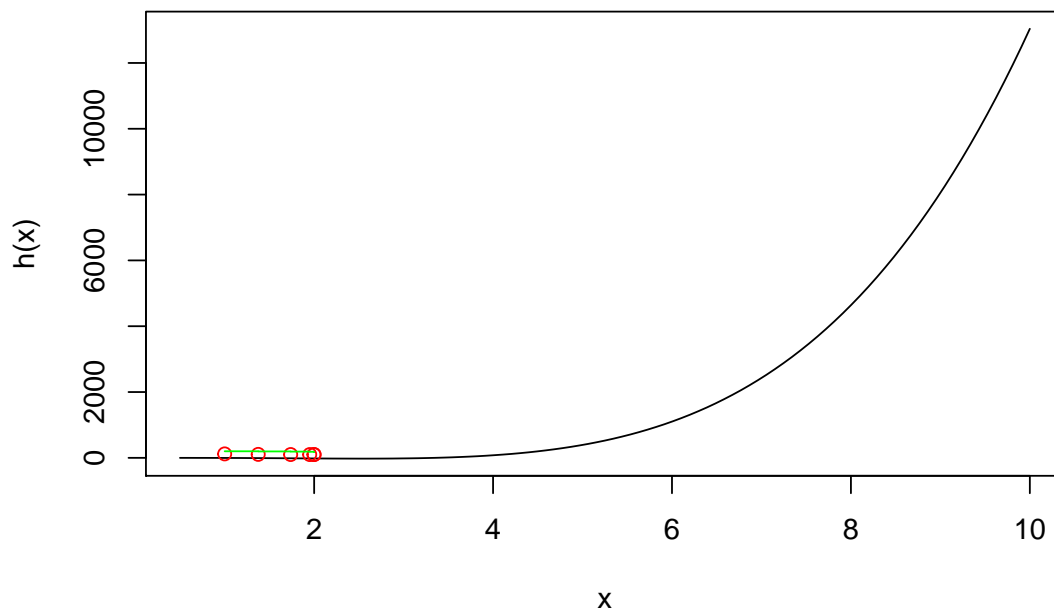


Figura 4.12: Gráfico - Método Newton Raphson ejercicio 1

4.2.2. Métodos de comparación de la función objetivo

Se considerarán ahora algunos procedimientos numéricos que localizan en forma directa el mínimo (en general, extremos) de una función $f(x)$. Básicamente, con estos métodos la búsqueda se hace evaluando la función en puntos elegidos dentro de un intervalo $[a, b]$, donde se sabe que se encuentra el punto óptimo. La idea general es obtener nuestro objetivo de la manera más eficiente posible, es decir, con el menor número de evaluaciones de la función. Una característica general de los métodos de aproximación es que el punto preciso en el cual ocurre el óptimo nunca será conocido, y lo mejor que puede obtenerse es determinar el intervalo final de incertidumbre.

Por lo tanto, nuestro objetivo primordial es obtener el intervalo final de incertidumbre para establecer la eficiencia η del método aplicado. Este intervalo de incertidumbre siempre

prevalece debido a que tales métodos calculan el valor de la función únicamente en valores discretos de la variable independientes. Para aplicar éstas técnicas, sólo se necesita conocer el intervalo inicial de incertidumbre $I_0 = b_0 - a_0$ y asegurarse de que $f(x)$ sea unimodal en el intervalo de interés. A continuación se verán los métodos más eficientes de búsqueda univariada sin derivadas.

4.2.2.1. Método de la sección áurea - Golden Ratio

Dado un intervalo (a, b) generamos inicialmente dos nuevos puntos $y_1 = a + (1 - \alpha)(b - a)$ y $z_1 = a + \alpha(b - a)$, para algún $\alpha \in (\frac{1}{2}, 1]$ convenientemente elegido. Según los valores que tomen $f(y)$ y $f(z)$, el intervalo de incertidumbre en el siguiente paso será (a, z) o (y, b) . Pues bien, la idea del método es elegir α de forma que según el caso que corresponda el valor y_k coincida con el z_{k-1} o bien z_k coincida con y_{k-1} .

Por ejemplo, si $f(y_1) \geq f(z_1)$, entonces el nuevo intervalo de incertidumbre es $(a_2 = y_1, b_2 = b)$ y α debe ser tal que $y_2 = a_2 + (1 - \alpha)(b - a_2) = z_1$. Puesto que $a_2 = y_1$, la relación anterior quedaría

$$\begin{aligned} a + (1 - \alpha)(b - a) + (1 - \alpha)(b - (a + (1 - \alpha)(b - a))) &= a + \alpha(b - a) \\ 2(1 - \alpha)(b - a) - (1 - \alpha)^2(b - a) &= \alpha(b - a) \\ 2(1 - \alpha) - (1 - \alpha)^2 &= \alpha \\ (1 - \alpha) - (2 - (1 - \alpha)) &= \alpha \\ (1 - \alpha)(1 + \alpha) &= \alpha \\ 1 - \alpha^2 &= \alpha \\ \alpha^2 + \alpha - 1 &= 0 \end{aligned}$$

cuya solución (positiva) es $\alpha = \frac{1 + \sqrt{5}}{2} \approx 1,618033989$

Obsérvese que con este método se tiene siempre que la longitud de un intervalo L_k respecto del anterior L_{k-1} es $L_k = \alpha L_{k-1} - 1$.

La denominación del método se debe a que α es el número conocido como la *sección áurea* un número descubierto en la antigüedad y que se asocia a proporciones que encontramos en la naturaleza en la morfología de diversos elementos tales como caracolas, nervaduras de las hojas de algunos árboles, el grosor de las ramas, etc. La sección áurea se obtiene dividiendo un segmento en dos partes de forma que la proporción de la parte menor frente a la parte mayor coincida con la proporción entre ésta última y la longitud total del segmento.

Algoritmo. Método de la sección áurea

◦ *Paso Inicial*

Fijar $h > 0$

$$\alpha = \frac{1 + \sqrt{5}}{2}$$

$$a_1 = a \text{ y } b_1 = b$$

Obtener $y_1 = a_1 + (1 - \alpha)(b_1 - a_1)$, $z_1 = a_1 + \alpha(b_1 - a_1)$.

Calcular $f(y_1), f(z_1)$.

Hacer $k = 1$.

◦ *Paso General*

Repetir:

• **Paso 1.-** Si $b_k - a_k < \epsilon$, paramos. $[a_k, b_k]$ es el intervalo buscado. En caso contrario, ir al paso 2.

• **Paso 2.-** Calcular $f(y_k)$ y $f(z_k)$. Si $f(y_k) \geq f(z_k)$, ir al paso 3. En caso contrario, ir al paso 4.

• **Paso 3.-** Hacer $a_{k+1} = y_k, b_{k+1} = b_k, y_{k+1} = z_k$ y $z_{k+1} = a_{k+1} + \alpha(b_{k+1} - a_{k+1})$. Calcular $f(z_{k+1})$. Hacer $k = k + 1$ e ir al paso 1.

• **Paso 4.-** Hacer $a_{k+1} = a_k, b_{k+1} = z_k, z_{k+1} = y_k$ e $y_{k+1} = a_{k+1} + (1 - \alpha)(b_{k+1} - a_{k+1})$. Calcular $f(y_{k+1})$. Hacer $k = k + 1$ e ir al paso 1.

Código R

```
optuni_nores_goldenratio = function(funx, extizq.ini,extder.ini,
                                   longinc=0.1,report.calculos=T) {
  # funx debe ser estrict. cuasiconvexa
  # longinc, epsilon > 0
  rho = ((3-sqrt(5))/2) #aproximadamente 0.381966
  n = 0
  nval = 0
  a0 = extizq.ini
  b0 = extder.ini
  a1 = a0 + rho * (b0-a0)
  b1 = a0 + (1-rho) * (b0-a0)
  f.a1 = funx(a1)
  f.b1 = funx(b1)
  mres = NULL
  repeat {
    f.a1 = funx(a1)
    f.b1 = funx(b1)

    if (f.a1>f.b1) {
      a2 = a1
      b2 = b0
      a3 = b1
      b3 = a2 + (1-rho)*(b2-a2)
    } else {
      a2 = a0
      b2 = b1
      a3 = a2 + rho*(b2-a2)
      b3 = a1
    }
  }
  vres = c(n+1,a0,b0,a1,b1,f.a1,f.b1,b0-a0)
  if (report.calculos) {
```

```

    if (is.null(mres)) {
      mres = vres
    } else {
      mres = rbind(mres,vres)
    }
  }
  if ((b0-a0)<=longinc) {
    break
  }
  n = n + 1
  nval = nval + 2
  a0 = a2
  b0 = b2
  a1 = a3
  b1 = b3
}
if (report.calculos) {
  colnames(mres) = c("n","a0","b0","a1","b1","fa1","fb1","Amplitud")
  rownames(mres) = NULL
  dfres = as.data.frame(mres)
} else {
  names(vres) = c("n","a0","b0","a1","b1","fa1","fb1","Amplitud")
  dfres = as.data.frame(vres)
}
print("Óptimo en el intervalo: (a0,b0)")
return(dfres)
}

```

Ejercicio 1

Lo utilizamos sobre la función $f(x) = x^4 - 14x^3 + 60x^2 - 70x$

```

funx = function(x) x^4-14*x^3+60*x^2-70*x
df.res = optuni_nores_goldenratio(funx,0,2)

```

```
## [1] "Óptimo en el intervalo: (a0,b0)"
```

```

knitr::kable(df.res,digits = 6, booktabs=TRUE,longtable=TRUE,
             caption = "\\label{tabla01}Ejercicio 1
             - Método sección áurea") %>%
kable_styling(latex_options =c("striped","HOLD_position","repeat_header"),
             position = "center",repeat_header_text = "contin\\'ua")

```

Tabla 4.9: Ejercicio 1 - Método sección áurea

n	a0	b0	a1	b1	fa1	fb1	Amplitud
1	0.000000	2.000000	0.763932	1.236068	-24.36068	-18.95816	2.000000
2	0.000000	1.236068	0.472136	0.763932	-21.09851	-24.36068	1.236068
3	0.472136	1.236068	0.763932	0.944272	-24.36068	-23.59246	0.763932
4	0.472136	0.944272	0.652476	0.763932	-23.83743	-24.36068	0.472136
5	0.652476	0.944272	0.763932	0.832816	-24.36068	-24.28789	0.291796
6	0.652476	0.832816	0.721360	0.763932	-24.25795	-24.36068	0.180340
7	0.721360	0.832816	0.763932	0.790243	-24.36068	-24.36691	0.111456
8	0.763932	0.832816	0.790243	0.806504	-24.36691	-24.34953	0.068884

Ahora sin iteraciones

```
funx = function(x) x^4-14*x^3+60*x^2-70*x
df.res2 = optuni_nores_goldenratio(funx,0,2,report.calculos = F)
```

```
## [1] "Óptimo en el intervalo: (a0,b0)"
```

```
knitr::kable(df.res2,digits = 6, booktabs=TRUE,longtable=TRUE,
             caption = "\\label{tabla01}Ejercicio 1 -
             Método sección áurea solución") %>%
kable_styling(latex_options =c("striped","HOLD_position","repeat_header"),
             position = "center",repeat_header_text = "contin\\'ua")
```

Tabla 4.10: Ejercicio 1 - Método sección áurea solución

	vres
n	8.000000
a0	0.763932
b0	0.832816
a1	0.790243
b1	0.806504
fa1	-24.366907
fb1	-24.349526
Amplitud	0.068884

Ahora vamos a ver si llegamos a los mismo resultado usando el método en **Excel**:

	A	B	C	D	E	F	G	H	I	J	K	L
1	n	a_{n-1}	a_n	b_n	b_{n-1}	$f(a_n)$	$f(b_n)$	$l_{intervalo}$	$\epsilon_{para?}$	$l(longitud)$		ro
2	1	0	0,76393202	1,23606798	2	-24,361	-18,958	2,000	FALSO	0,1	0,38196601	
3	2	0	0,47213595	0,76393202	1,23606798	-21,099	-24,361	1,236	FALSO			
4	3	0,47213595	0,76393202	0,94427191	1,23606798	-24,361	-23,592	0,764	FALSO			
5	4	0,47213595	0,65247584	0,76393202	0,94427191	-23,837	-24,361	0,472	FALSO			
6	5	0,65247584	0,76393202	0,83281573	0,94427191	-24,361	-24,288	0,292	FALSO			
7	6	0,65247584	0,72135955	0,76393202	0,83281573	-24,258	-24,361	0,180	FALSO			
8	7	0,72135955	0,76393202	0,79024326	0,83281573	-24,361	-24,367	0,111	FALSO			
9	8	0,76393202	0,79024326	0,8065045	0,83281573	-24,367	-24,350	0,069	VERDADERO			
10												

Figura 4.13: Excel - Método sección áurea ejercicio 1

Como se puede ver da los mismos resultados.

Lo primero de todo para poder realizar este método la función debe ser cuasiconvexa, que en este caso lo es.

Objetivo del método es buscar el mínimo en el intervalo $[0, 2]$. Después, de realizar 8 iteraciones se encuentra la sección donde está el mínimo.

La sección sería $[0,76393202, 0,83281573]$ con una amplitud de $0,069$, que es menor que $\epsilon = 0.1$.

Se puede ver como evoluciona el algoritmo gráficamente:

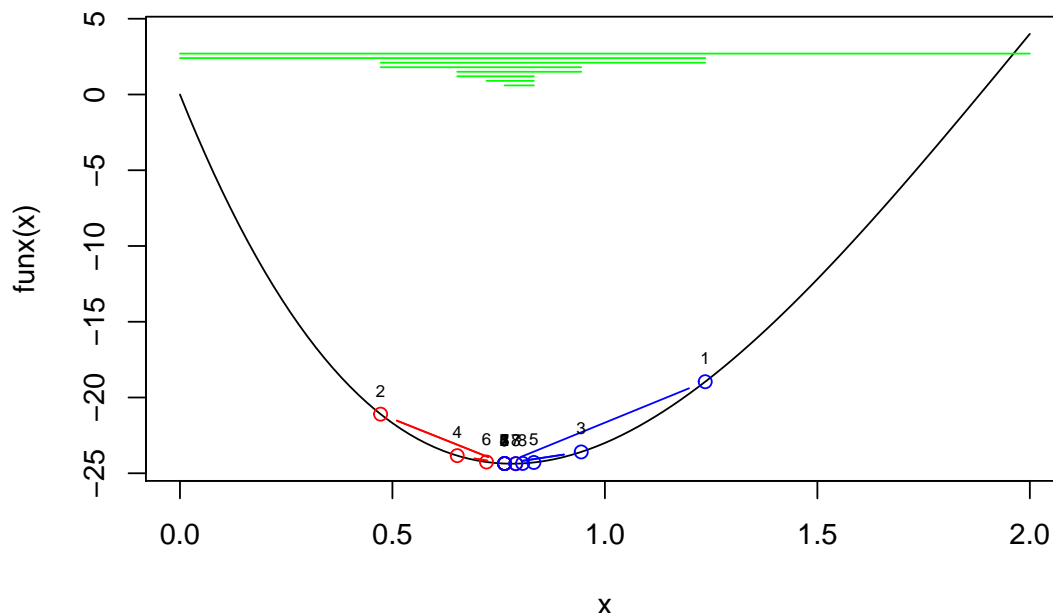


Figura 4.14: Gráfico - Método sección áurea ejercicio 1

Si queremos evaluar la función solo una vez en cada iteración es preciso hacer algunos cambios.


```

optuni_nores_goldenratio2 = function(funx, extizq.ini,extder.ini,
longinc=0.1,report.calculos=T) {
# funx debe ser estrict. cuasiconvexa
# longinc, epsilon > 0
rho = ((3-sqrt(5))/2) # 0.381966
n = 1
nval = 0
a0 = extizq.ini
b0 = extder.ini
a1 = a0 + rho * (b0-a0)
b1 = a0 + (1-rho) * (b0-a0)
f.a1 = funx(a1)
f.b1 = funx(b1)
mres = c(n,a0,b0,a1,b1,f.a1,f.b1,b0-a0)
repeat {
# f.a1 = funx(a1)
# f.b1 = funx(b1)
if (f.a1>f.b1) {
a2 = a1
b2 = b0
a3 = b1
b3 = a2 + (1-rho)*(b2-a2)
f.a1 = f.b1
f.b1 =funx(b3)
} else {
a2 = a0
b2 = b1
a3 = a2 + rho*(b2-a2)
b3 = a1
f.b1 = f.a1
f.a1 = funx(a3)
}
vres = c(n+1,a2,b2,a3,b3,f.a1,f.b1,b2-a2)
if (report.calculos) {mres = rbind(mres,vres)}
if ((b0-a0)<=longinc) {
break
}
n = n + 1
nval = nval + 2
a0 = a2
b0 = b2
a1 = a3
b1 = b3
}
if (report.calculos) {
colnames(mres) = c("n", "a0", "b0", "a1", "b1", "fa1", "fb1", "Amplitud")
rownames(mres) = NULL
dfres = as.data.frame(mres)
}
}

```

```

} else {
names(vres) = c("n", "a0", "b0", "a1", "b1", "fa1", "fb1", "Amplitud")
dfres = as.data.frame(vres)
}
print("Óptimo en el intervalo: (a0,b0)")
return(dfres)
}

```

Con información de las iteraciones.

```

funx = function(x) x^4-14*x^3+60*x^2-70*x
df.res3 = optuni_nores_goldenratio2(funx,0,2)

```

```
## [1] "Óptimo en el intervalo: (a0,b0)"
```

```

knitr::kable(df.res3,digits = 6, booktabs=TRUE,longtable=TRUE,
             caption = "\\label{tabla01}Ejercicio 1
             - Método sección áurea evaluando una vez la función ") %>%
kable_styling(latex_options = c("striped", "HOLD_position",
"repeat_header", "scale_down"),
repeat_header_text = "(continua)")

```

Tabla 4.11: Ejercicio 1 - Método sección áurea evaluando una vez la función

n	a0	b0	a1	b1	fa1	fb1	Amplitud
1	0.000000	2.000000	0.763932	1.236068	-24.36068	-18.95816	2.000000
2	0.000000	1.236068	0.472136	0.763932	-21.09851	-24.36068	1.236068
3	0.472136	1.236068	0.763932	0.944272	-24.36068	-23.59246	0.763932
4	0.472136	0.944272	0.652476	0.763932	-23.83743	-24.36068	0.472136
5	0.652476	0.944272	0.763932	0.832816	-24.36068	-24.28789	0.291796
6	0.652476	0.832816	0.721360	0.763932	-24.25795	-24.36068	0.180340
7	0.721360	0.832816	0.763932	0.790243	-24.36068	-24.36691	0.111456
8	0.763932	0.832816	0.790243	0.806504	-24.36691	-24.34953	0.068884
9	0.763932	0.806504	0.780193	0.790243	-24.36959	-24.36691	0.042572

Gráficamente:

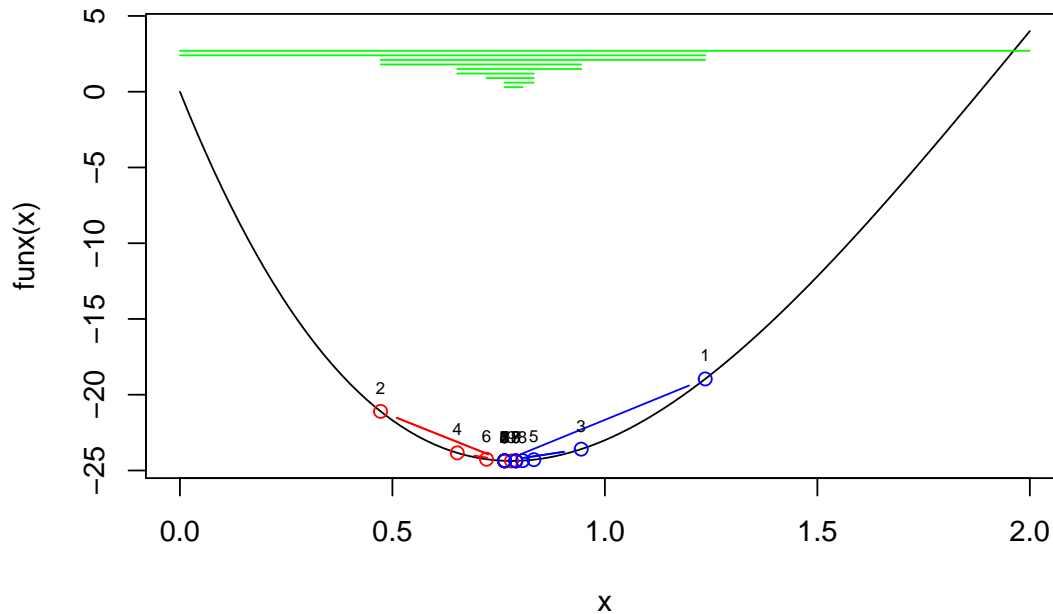


Figura 4.15: Gráfico - Método sección áurea ejercicio 1 evaluando una vez la función

4.2.2.2. Método de búsqueda dicotómica

Existe una diversidad de métodos que comparan los valores de la función objetivo y que usan el concepto de intervalo de incertidumbre como los métodos de búsqueda dicotómica.

Método de búsqueda dicotómica

Sea $f : [a_0, b_0] \subset \mathbb{R} \rightarrow \mathbb{R}$ una función estrictamente cuasiconvexa.

Se elige $\epsilon > 0$ y una amplitud de incertidumbre l , y $n = 0$.

Mientras $b_n - a_n > l$, entonces $\lambda_n = \frac{a_n + b_n}{2} - \epsilon$ y $\mu_n = \frac{a_n + b_n}{2} + \epsilon$.

Si $f(\lambda_n) < f(\mu_n)$, luego $a_{n+1} = a_n$ y $b_{n+1} = \mu_n$, en caso contrario, $a_{n+1} = \lambda_n$ y $b_{n+1} = b_n$.

$n = n + 1$

Terminamos mientras se hagan dos evaluaciones de la función en cada iteración.

Código R

```
optuni_nores_busqdicotomica = function(funx, extizq.ini,extder.ini,
                                       longinc=0.1,epsilon=0.01,
                                       report.calculos=T) {
  # funx debe ser strict. cuasiconvexa
  # longinc, epsilon > 0
```

```

n = 0
an = extizq.ini
bn = extder.ini
mres = NULL
repeat {
  lambda.n = ((an+bn)/2) - epsilon
  mu.n = ((an+bn)/2) + epsilon
  f.an = funx(an)
  f.bn = funx(bn)
  f.lambda.n = funx(lambda.n)
  f.mu.n = funx(mu.n)
  if (f.lambda.n < f.mu.n) {
    an1 = an
    bn1 = mu.n
  } else {
    an1 = lambda.n
    bn1 = bn
  }
  vres = c(n+1, an, bn, lambda.n, mu.n, f.an, f.bn, f.lambda.n,
          f.mu.n, bn-an, an1, bn1)
  if (report.calculos) {
    if (is.null(mres)) {
      mres = vres
    } else {
      mres = rbind(mres, vres)
    }
  }
  if ((bn-an) <= longinc) {
    break
  }
  n = n + 1
  an = an1
  bn = bn1
}
if (report.calculos) {
  colnames(mres) = c("n", "an", "bn", "lambdan", "mun", "fan", "fbn",
                    "flambdan", "fmun", "Amplitud", "an1", "bn1")
  rownames(mres) = NULL
  dfres = as.data.frame(mres)
} else {
  names(vres) = c("n", "an", "bn", "lambdan", "mun", "fan", "fbn",
                 "flambdan", "fmun", "Amplitud", "an1", "bn1")
  dfres = as.data.frame(vres)
}
print("Óptimo en el intervalo: (an1, bn1)")
return(dfres)
}

```

Ejercicio 1

Lo utilizamos sobre la función $f(x) = x^4 - 14x^3 + 60x^2 - 70x$

```
funx = function(x) x^4-14*x^3+60*x^2-70*x
df.res = optuni_nores_busqdicotomica(funx,-2,1)

## [1] "Óptimo en el intervalo: (an1,bn1)"

knitr::kable(df.res,digits = 4,
             caption = "\\label{tabla01}Ejercicio 1 -
             Método búsqueda dicotómica") %>%
kable_styling(latex_options = c("striped", "HOLD_position",
"repeat_header", "scale_down"),
repeat_header_text = "(continua)")
```

Tabla 4.12: Ejercicio 1 - Método búsqueda dicotómica

n	an	bn	lambdan	mun	fan	fbn	flambdan	fmun	Amplitud	an1	bn1
1	-2.0000	1.0000	-0.5100	-0.4900	508.0000	-23.0000	53.2308	50.4107	3.0000	-0.5100	1.0000
2	-0.5100	1.0000	0.2350	0.2550	53.2308	-23.0000	-13.3151	-14.1764	1.5100	0.2350	1.0000
3	0.2350	1.0000	0.6075	0.6275	-13.3151	-23.0000	-23.3842	-23.6037	0.7650	0.6075	1.0000
4	0.6075	1.0000	0.7938	0.8138	-23.3842	-23.0000	-24.3645	-24.3367	0.3925	0.6075	0.8138
5	0.6075	0.8138	0.7006	0.7206	-23.3842	-24.3367	-24.1651	-24.2551	0.2063	0.7006	0.8138
6	0.7006	0.8138	0.7472	0.7672	-24.1651	-24.3367	-24.3341	-24.3638	0.1131	0.7472	0.8138
7	0.7472	0.8138	0.7705	0.7905	-24.3341	-24.3367	-24.3662	-24.3668	0.0666	0.7705	0.8138

Ahora veremos el ejemplo en Excel.

18											
19	<i>n</i>	<i>a_n</i>	<i>λ_n</i>	<i>μ_n</i>	<i>b_n</i>	<i>f(λ_n)</i>	<i>f(μ_n)</i>	<i>l intervalo</i>	<i>¿para?</i>	<i>epsilon</i>	<i>l(longitud)</i>
20	1	-2	-0,51	-0,49	1	53,231	50,411	3,000	FALSO	0,01	0,1
21	2	-0,51	0,235	0,255	1	-13,315	-14,176	1,510	FALSO		
22	3	0,235	0,6075	0,6275	1	-23,384	-23,604	0,765	FALSO		
23	4	0,6075	0,79375	0,81375	1	-24,365	-24,337	0,393	FALSO		
24	5	0,6075	0,700625	0,720625	0,81375	-24,165	-24,255	0,206	FALSO		
25	6	0,700625	0,7471875	0,7671875	0,81375	-24,334	-24,364	0,113	FALSO		
26	7	0,7471875	0,77046875	0,79046875	0,81375	-24,366	-24,367	0,067	VERDADERO		
27											

Figura 4.16: Excel - Método búsqueda dicotómica ejercicio 1

En la iteración 7, se ha logrado que la longitud del intervalo sea menor que 0,1. El intervalo óptimo sería $[0,7471875, 0,81375]$, partiendo de $[-2, 1]$. En este método se hacen dos evaluaciones de la función en cada iteración con $f(\lambda_n)$ y $f(\mu_n)$.

Ejercicio 2

```
f2<-function(x){sin(x)-0.1*x}
bd.sin<-optuni_nores_busqdicotomica(f2,-2*pi,2*pi)
```

```
## [1] "Óptimo en el intervalo: (an1,bn1)"
```

```
knitr::kable(bd.sin,digits = 4,
             caption = "\\label{tabla01}Ejercicio 2 -
             Método búsqueda dicotómica") %>%
kable_styling(latex_options = c("striped", "HOLD_position",
"repeat_header", "scale_down"),
repeat_header_text = "(continua)")
```

Tabla 4.13: Ejercicio 2 - Método búsqueda dicotómica

n	an	bn	lambdan	mun	fan	fbn	flambdan	fmun	Amplitud	an1	bn1
1	-6.2832	6.2832	-0.0100	0.0100	0.6283	-0.6283	-0.0090	0.0090	12.5664	-6.2832	0.0100
2	-6.2832	0.0100	-3.1466	-3.1266	0.6283	0.0090	0.3197	0.2977	6.2932	-3.1466	0.0100
3	-3.1466	0.0100	-1.5783	-1.5583	0.3197	0.0090	-0.8421	-0.8441	3.1566	-1.5783	0.0100
4	-1.5783	0.0100	-0.7941	-0.7741	-0.8421	0.0090	-0.6339	-0.6217	1.5883	-1.5783	-0.7741
5	-1.5783	-0.7741	-1.1862	-1.1662	-0.8421	-0.6217	-0.8083	-0.8026	0.8041	-1.5783	-1.1662
6	-1.5783	-1.1662	-1.3823	-1.3623	-0.8421	-0.8026	-0.8441	-0.8421	0.4121	-1.5783	-1.3623
7	-1.5783	-1.3623	-1.4803	-1.4603	-0.8421	-0.8421	-0.8479	-0.8479	0.2160	-1.5783	-1.4603
8	-1.5783	-1.4603	-1.5293	-1.5093	-0.8421	-0.8479	-0.8462	-0.8472	0.1180	-1.5293	-1.4603
9	-1.5293	-1.4603	-1.5048	-1.4848	-0.8462	-0.8479	-0.8473	-0.8478	0.0690	-1.5048	-1.4603

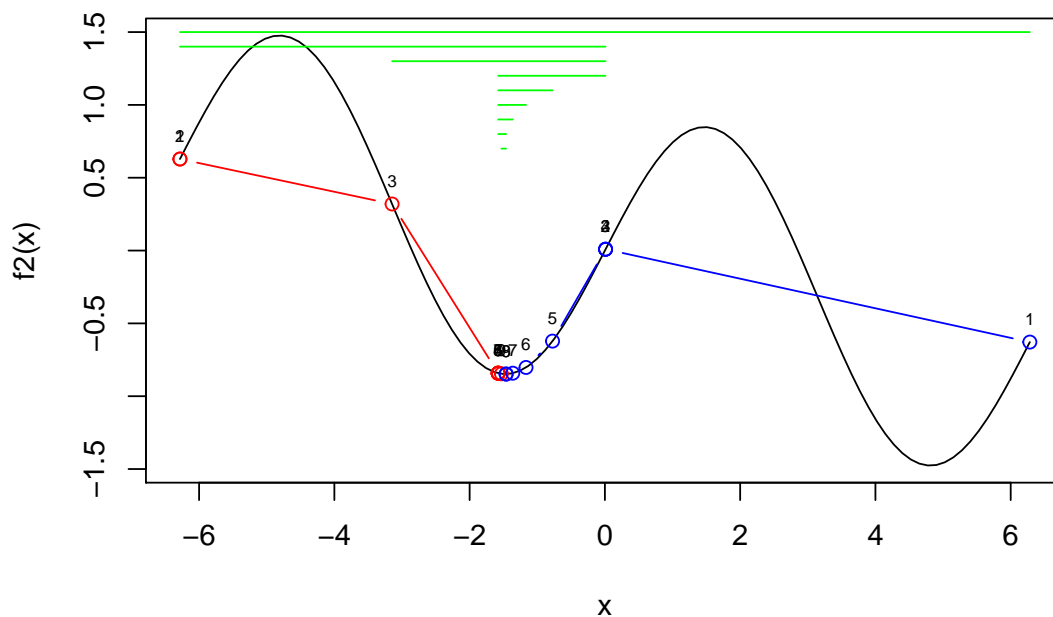


Figura 4.17: Gráfico - Método búsqueda dicotómica ejercicio 2

Después de varias iteraciones se encuentra el intervalo donde se encuentra el óptimo $[-1.5293, -1.4603]$, con una amplitud de 0.0690. En la iteración 9.

4.3. Métodos Multivariantes sin usar derivadas

La mayoría de los métodos seleccionados se encuentran [4].

4.3.1. Optimización de funciones diferenciables multidimensionales

Optimización de funciones sin usar derivadas.

Una de las grandes desventajas de los métodos que emplean derivadas para funciones de varias variables es que en estos es necesario evaluar el gradiente y en varios casos, también el Hessiano.

Otro problema en consideración, en el caso de que la función sea diferenciable, es el evaluar analíticamente en cada iteración las funciones contenidas en el gradiente y el Hessiano. Esta tarea de evaluación analítica, puede ser muy laborioso, sobre todo cuando se trata de funciones complejas con muchas variables independientes. Hay métodos numéricos que permiten expresar estas derivadas por procesos que no son analíticos, y por lo tanto se puede eliminar este problema. La gran desventaja con los métodos numéricos es el error de redondeo que introducen y, que acumulado en un gran número de operaciones e iteraciones, puede generar resultados bastante alejados de los óptimos reales.

Por este motivo, se han desarrollado una serie de métodos iterativos que, sin ser tan eficientes en cuanto al número de iteraciones necesarias para alcanzar un óptimo como los métodos basados en el gradiente, tienen las siguientes ventajas:

- No requieren de gradientes y Hessianos, por lo que sirven tanto para funciones diferenciables como no diferenciables.
- No requieren de métodos numéricos y por lo tanto su error de redondeo se reduce considerablemente.

Ahora se analizan algunos de estos métodos.

4.3.1.1. Método de Coordenadas Cíclicas

Este método usa los ejes coordenados como direcciones de búsqueda. Mas específicamente, el método busca a lo largo de las direcciones d_1, \dots, d_n , donde d_j es un vector de ceros, excepto por un 1 en la coordenada j . Así, a lo largo de la dirección de búsqueda d_j , la variable x_j es cambiada mientras todas las otras variables quedan fijas.

Si la función es diferenciable, entonces el método converge a un punto estacionario.

Procedimiento

• **Paso 0:** Elegir un escalar $\varepsilon > 0$ para ser usado para saber cuando termina el algoritmo como margen de incertidumbre y sean d_1, \dots, d_n las direcciones coordenadas. Elegir un punto inicial x_1 , sea $y_1 = x_1$, $k = j = 1$ y continuar con el paso 1.

• **Paso 1:** Sea λ_j una solución óptima a el problema de minimizar la función $f(y + \lambda d_j)$ sujeta a $\lambda \in \mathbb{R}$ y sea $y_{j+1} = y_j + \lambda_j d_j$. Si $j < n$, reemplazar j por $j + 1$ y repetir el paso 1. De otra manera, si $j = n$, continuamos con el paso 2.

• **Paso 2:** Sea $x_{k+1} - y_{n+1}$. Si $\|x_{k+1} - x_k\| < \varepsilon$ paramos. De lo contrario sea $y_1 = x_{k+1}$, sea $j = 1$, reemplazar k por $k + 1$ y repetir el paso 1.

Código R

```
#####Código método coordenadas cíclicas#####
# vx.ini = c(0,3)
# funx = function(vx) (vx[1]-2)^2+(vx[1]-2*vx[2])^2
optmul_nores_coordciclicas = function(funx, vx.ini,
                                     epsilon=0.01,
                                     nummax.iter=1000,
                                     report.calculos=T,
                                     cota.inf=-1000,
                                     cota.sup=1000) {

  # longinc, epsilon > 0
  #browser()
  mres = NULL
  k = 1
  vyi = vx.ini
  vxki = vx.ini
  ndimension = length(vx.ini)
  m.vectoresD = diag(x=1,nrow=ndimension,ncol=ndimension)
  # Paso 3
  repeat {
    j = 1
    # Paso 2
    for (j in 1:ndimension) {
      vDj = m.vectoresD[j,]
      ffunxj = function(lambda) funx(vyi+lambda*vDj)
      sol.uni = optimize(ffunxj,c(cota.inf,cota.sup),maximum = F)
      #if (j<ndimension) {
      #vDj1 = m.vectoresD[j+1,]
      vyi1 = vyi + sol.uni$minimum * vDj
      #}
      vres = c(k,j,vyi,sol.uni$minimum,vyi1)
      if (report.calculos) {
        if (is.null(mres)) {
          mres = vres
        } else {
          mres = rbind(mres,vres)
        }
      }
      vyi = vyi1
    }
    vxki1 = vyi1
    norma.l2 = sqrt(sum((vxki1-vxki)^2))
  }
}
```



```

if ((norma.l2<epsilon) | (k>=nummax.iter)) {
  break
} else {
  k = k+1
  vyi = vxki1
  vxki = vyi
}
} # final repeat
if (report.calculos) {
  colnames(mres) = c("k","j",paste0("vyi",1:ndimension),
                    "lambda",paste0("vyimas1",1:ndimension))
  rownames(mres) = NULL
  dfres = as.data.frame(mres)
} else {
  names(vres) = c("k","j",paste0("vyi",1:ndimension),
                 "lambda",paste0("vyimas1",1:ndimension))
  dfres = as.data.frame(vres)
}
print("Óptimo en vyimas1")
return(dfres)
}

```

Ejercicio 1

Minimizar $f(x_1, x_2) = (x_1 - 2)^4 + (x_1 - 2x_2)^2$ partiendo del punto inicial (0, 3).

```

vx.ini = c(0,3)
funx = function(vx) (vx[1]-2)^4+(vx[1]-2*vx[2])^2
df.res = optimul_nores_coordciclicas(funx, vx.ini, epsilon=0.0001,
                                   nummax.iter = 1000,report.calculos=T)

```

```
## [1] "Óptimo en vyimas1"
```

La tabla muestra un resumen de los cálculos para el método de coordenadas cíclicas. En cada iteración los vectores

```

knitr::kable(df.res[c(1:6,nrow(df.res)),],digits = 4,
             caption = "\\label{tabla01}Ejercicio 1 -
             Método coordenadas cíclicas") %>%
kable_styling(latex_options = c("striped", "HOLD_position",
"repeat_header", "scale_down"),
repeat_header_text = "(continua)")

```

Tabla 4.14: Ejercicio 1 - Método coordenadas cíclicas

	k	j	vyi1	vyi2	lambda	vyimas11	vyimas12
1	1	1	0.0000	3.0000	3.1282	3.1282	3.0000
2	1	2	3.1282	3.0000	-1.4359	3.1282	1.5641
3	2	1	3.1282	1.5641	-0.4987	2.6294	1.5641
4	2	2	2.6294	1.5641	-0.2494	2.6294	1.3147
5	3	1	2.6294	1.3147	-0.1807	2.4487	1.3147
6	3	2	2.4487	1.3147	-0.0904	2.4487	1.2244
374	187	2	2.0366	1.0184	0.0000	2.0366	1.0183

Ahora sin iteraciones

```
df.res2 = optimul_nores_coordciclicas(funx,vx.ini,
```

```
## [1] "Óptimo en vyimas1"
```

```
knitr::kable(df.res2,digits = 6, booktabs=TRUE,longtable=TRUE,
  caption = "\\label{tabla01}Ejercicio 1
  - Método coordenadas cíclicas solución") %>%
kable_styling(latex_options = c("striped", "HOLD_position",
"repeat_header", "scale_down"),
repeat_header_text = "(continua)")
```

Tabla 4.15: Ejercicio 1 - Método coordenadas cíclicas solución

	vres
k	187.000000
j	2.000000
vyi1	2.036647
vyi2	1.018363
lambda	-0.000039
vyimas11	2.036647
vyimas12	1.018324

El objetivo es minimizar $f(x) = (x_1 - 2)^4 + (x_1 - 2x_2)^2$ partiendo del punto inicial $(0, 3)$. La tabla muestra un resumen de los cálculos para el método de coordenadas cíclicas. En cada iteración los vectores y_1 y y_2 son obtenidos mediante la representación de una línea de búsqueda en las direcciones $(1, 0)$ y $(0, 1)$, respectivamente. Para el punto $(2.036647, 1.018363)$ la función objetivo tiene un valor de 0.00000180365 , muy cercano del valor óptimo 0 .

Se puede ver como evoluciona el algoritmo gráficamente en 3D:

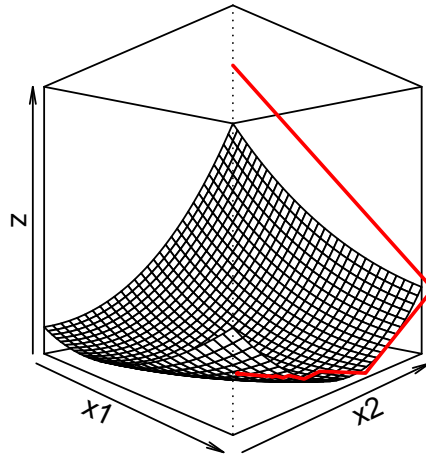


Figura 4.18: Gráfico en 3D - Método coordenadas cíclicas ejercicio 1

También se puede ver como evoluciona el algoritmo gráficamente en 2D sobre las curvas de nivel:

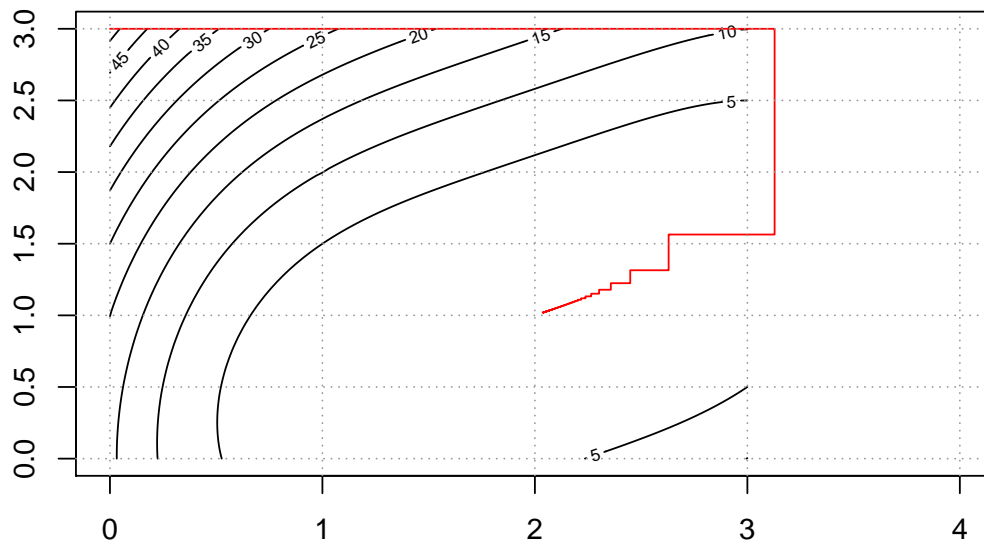


Figura 4.19: Gráfico en 2D - Método coordenadas cíclicas ejercicio 1

Aquí elegimos los niveles para los que se representan las curvas de nivel:

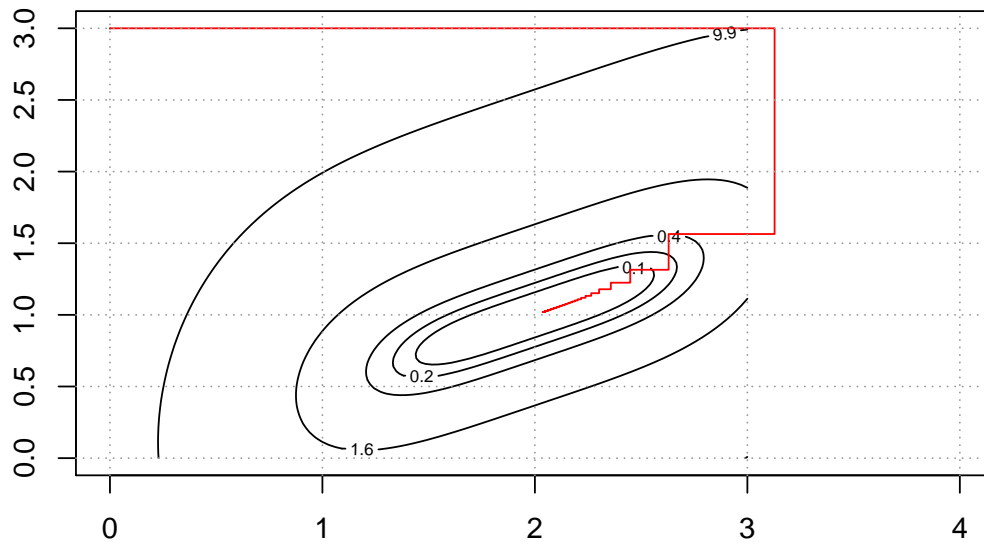


Figura 4.20: Gráfico - Método coordenadas cíclicas ejercicio 1

También puede verse como se mueve en el plano XY:

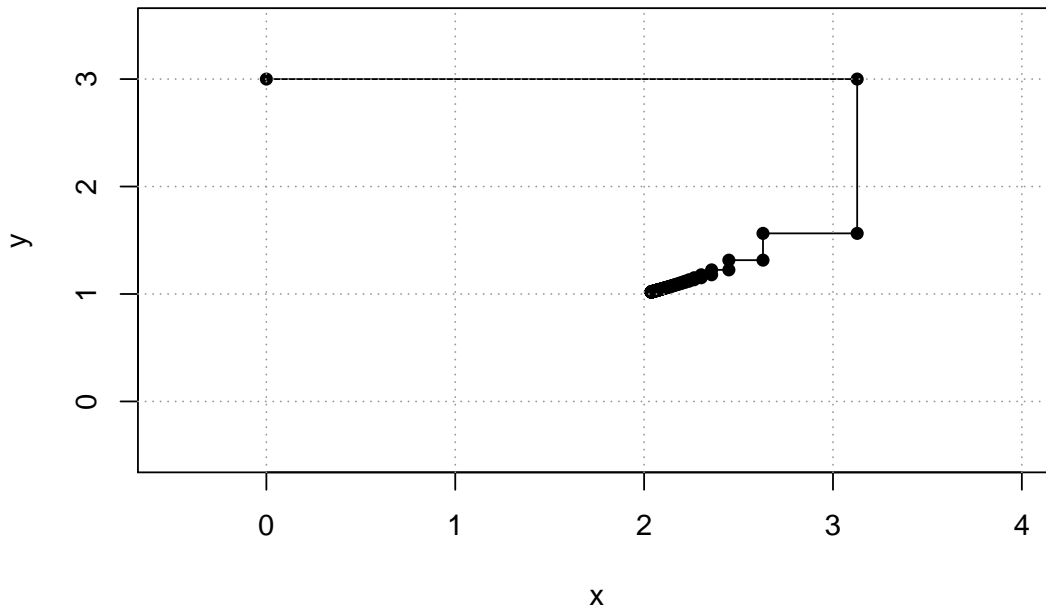


Figura 4.21: Gráfico en XY - Método coordenadas cíclicas ejercicio 1

4.3.1.2. Método de Hooke y Jeeves

El método de Hooke y Jeeves representa dos tipos de búsqueda exploratoria y patrón de búsqueda. Dado x_1 , una búsqueda exploratoria a lo largo de las direcciones coordenadas produce el punto x_2 . Un patrón de búsqueda a lo largo de la dirección $x_2 - x_1$ conduce al punto y . Otra búsqueda exploratoria a partir del punto y conduce al punto x_3 . El siguiente patrón de búsqueda es a lo largo de la dirección $x_3 - x_2$, obteniendo y' . El proceso es repetido.

La figura muestra las primeras dos iteraciones del método. Las líneas que van de x_1 a x_2 y las que van de y a x_3 representan la búsqueda exploratoria a lo largo de los ejes coordenados y la línea que une a x_2 y x_3 y la que une a x_2 e y representa el patrón de búsqueda.

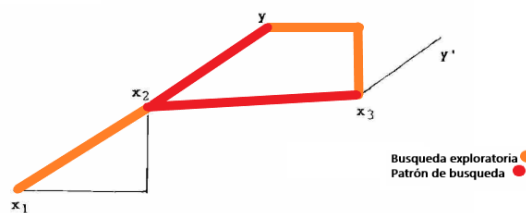


Figura 4.22: Busqueda exploratoria y patrón de búsqueda

El método de Hooke y Jeeves maneja direcciones de línea a lo largo de las direcciones coordenadas d_1, \dots, d_n y un patrón de búsqueda en su desarrollo.

Procedimiento

- **Paso 0:** Elegir un escalar $\varepsilon > 0$ como margen de incertidumbre. Elegir un punto de partida x_1 , sea $y_1 = x_1$, sea $k = j = 1$ y seguir con el paso 1.

- **Paso 1:** Sea λ_j una solución óptima a el problema de minimizar la función $f(y_j + \lambda d_j)$ sujeto a $\lambda \in \mathbb{R}$ y sea $y_{j+1} = y_j + \lambda_j d_j$. Si $j < n$, reemplazar j por $j + 1$, y repetir el paso 1. De lo contrario, si $j = n$ sea $x_{k+1} = y_{n+1}$. Si $\|x_{k+1} - x_k\| < \varepsilon$, paramos, de lo contrario, continuamos con el paso 2.

- **Paso 2:** Sea $d = x_{k+1} - x_k$ y sea λ una solución óptima al problema de minimizar $f(x_{k+1} + \lambda d)$ sujeto a $\lambda \in \mathbb{R}$. Sea $y_1 = x_{k+1} + \lambda d$, sea $j = 1$, reemplazar k por $k + 1$, y repetir el paso 1.

Código R

```
optmul_nores_HJ = function(funx, vx.ini,
                           epsilon=0.01,nummax.iter = 1000,
                           report.calculos=T,
                           cota.inf=-1000,cota.sup=1000) #Buscamos lambda
#en un área (estos valores podemos modificarlos)
{
  # longinc, epsilon > 0
  #browser()
  mres = NULL #Matriz de resultados
  k = 1
  vyi = vx.ini
  vxki = vx.ini
  ndimension = length(vx.ini)
  m.vectoresD = diag(x=1,nrow=ndimension,ncol=ndimension)
  #Vectores de direcciones
  # Paso 3
  repeat {
    j = 1
    # Paso 2
    for (j in 1:ndimension) {
      vDj = m.vectoresD[j,]
      ffunxj = function(lambda) funx(vyi+lambda*vDj)
      sol.uni = optimize(ffunxj,c(cota.inf,cota.sup),maximum = F)
      #if (j<ndimension) {
      #vDj1 = m.vectoresD[j+1,]
      vyi1 = vyi + sol.uni$minimum * vDj
      valorf=funx(vyi1)
      #}
      vres = c(k,j,vyi,sol.uni$minimum,vyi1,vDj,valorf)
      if (report.calculos) {
        if (is.null(mres)) {
          mres = vres
        }
      }
    }
  }
}
```

```

    } else {
      mres = rbind(mres,vres)
    }
  }
  vyi = vyi1
}
vxki1 = vyi1
norma.l2 = sqrt(sum((vxki1-vxki)^2))
#distancia euclidea entre los dos puntos

if ((norma.l2<epsilon) | (k>=nummax.iter)) {
  break
} else {
  #Paso 3 (Parte modificada del método de Cordenadas cíclicas)
  d=vxki1-vxki
  ffunxj=function(lambda) funx(vxki1+lambda*d) #Misma función,
  #pero d distinto
  sol.uni=optimize(ffunxj,c(cota.inf,cota.sup),maximum = F) #Optimiza
  #unidimensional
  vyi1=vxki1+sol.uni$minimum*d
  j=3 #dentro de cada etapa obtenemos 1, 2 y 3
  vres=c(k,j,vyi,sol.uni$minimum,vyi1) #vector de resultados
  mres=rbind(mres,vres)
  j=1
  k=k+1
  vyi=vyi1
  vxki=vxki1
}

} # final repeat
if (report.calculos) {
  colnames(mres) = c("k","j",paste0("vyi",1:ndimension),"lambda",
                    paste0("vyimas1",1:ndimension),
                    paste0("dj",1:ndimension),"valorf")
  rownames(mres) = NULL
  dfres = as.data.frame(mres)
} else {
  names(vres) = c("k","j",paste0("vyi",1:ndimension),"lambda",
                 paste0("vyimas1",1:ndimension),
                 paste0("dj",1:ndimension),"valorf")
  dfres = as.data.frame(vres)
}
print("Óptimo en vyimas1")
return(dfres)
}

```

Ejercicio 1

```

vx.ini = c(0,3) #coordenadas del punto en el que comienzo a iterar
funx = function(vx) ((vx[1]-2)^4+(vx[1]-2*vx[2])^2)
df.res = optimul_nores_HJ(funx, vx.ini,
                          epsilon=0.0001,nummax.iter = 1000,report.calculos=T)

```

```
## [1] "Óptimo en vyimas1"
```

```

knitr::kable(df.res[c(1:6,nrow(df.res)),c(1:7)],digits = 4,
             caption = "\\label{tabla01}Ejercicio 1
             - Método Hooke y Jeeves") %>%
kable_styling(latex_options = c("striped", "HOLD_position",
"repeat_header", "scale_down"),
             repeat_header_text = "(continua)")

```

Tabla 4.16: Ejercicio 1 - Método Hooke y Jeeves

	k	j	vyi1	vyi2	lambda	vyimas11	vyimas12
1	1	1	0.0000	3.0000	3.1282	3.1282	3.0000
2	1	2	3.1282	3.0000	-1.4359	3.1282	1.5641
3	1	3	3.1282	1.5641	-0.0972	2.8241	1.7037
4	2	1	2.8241	1.7037	-0.1187	2.7054	1.7037
5	2	2	2.7054	1.7037	-0.3510	2.7054	1.3527
6	2	3	2.7054	1.3527	1.6684	2.0000	1.0000
11	4	2	2.0000	1.0000	0.0000	2.0000	1.0000

Minimizamos la función $f(x) = (x_1 - 2)^4 + (x_1 - 2x_2)^2$, la tabla muestra los cálculos del método de Hooke y Jeeves. En cada iteración, una búsqueda exploratoria a lo largo de las direcciones coordenadas dados los puntos y_2 y y_1 y un patrón de búsqueda a lo largo de la dirección $d = x_{k+1} - x_k$ dado el punto y_1 .

En 11 iteraciones se llega al punto inicial a la solución óptima (2, 1), con valor objetivo igual a 0.

Evolución de ejemplo en 3D:

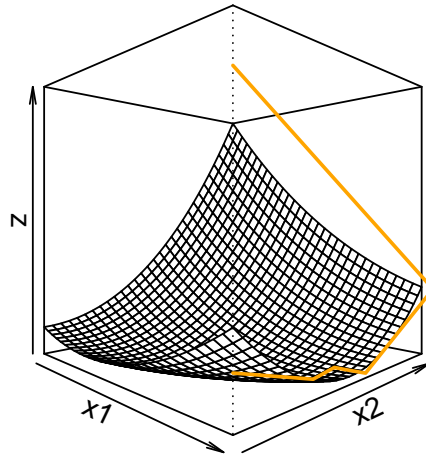


Figura 4.23: Gráfico en 3D - Método Hooke y Jeeves ejercicio 1

Evolución gráfica en 2D sobre las curvas de nivel:

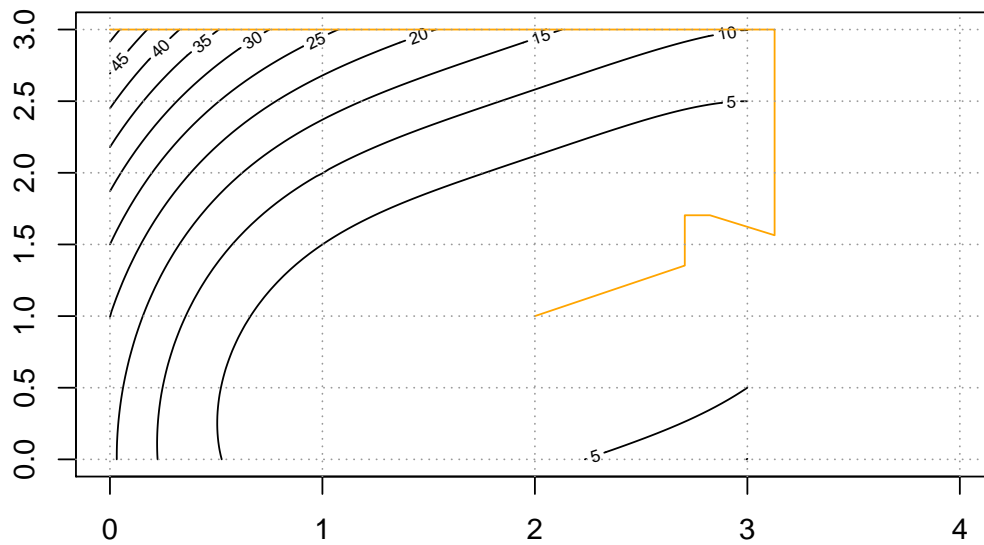


Figura 4.24: Gráfico en 2D - Método Hooke y Jeeves ejercicio 1

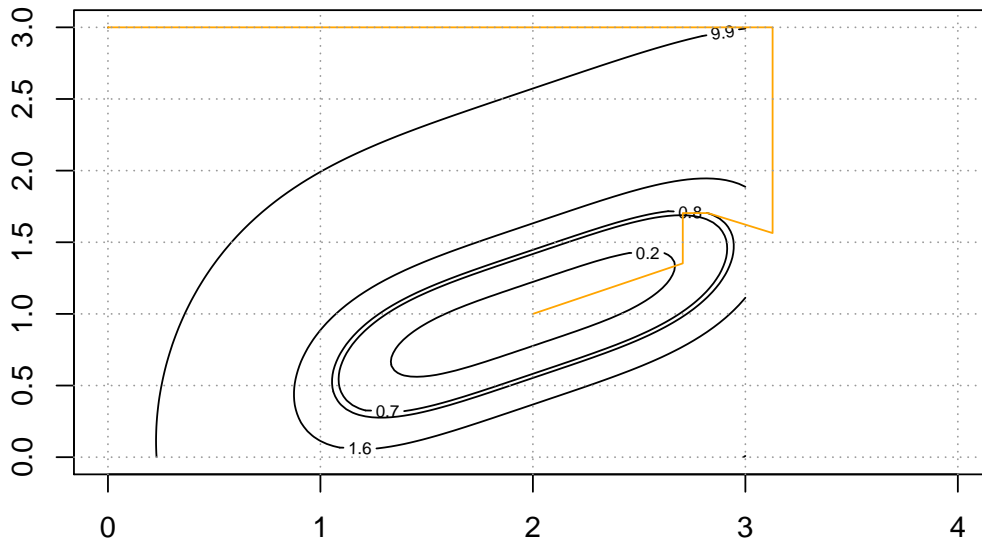


Figura 4.25: Gráfico - Método Hooke y Jeeves ejercicio 1

Evolución gráfica en el eje XY:

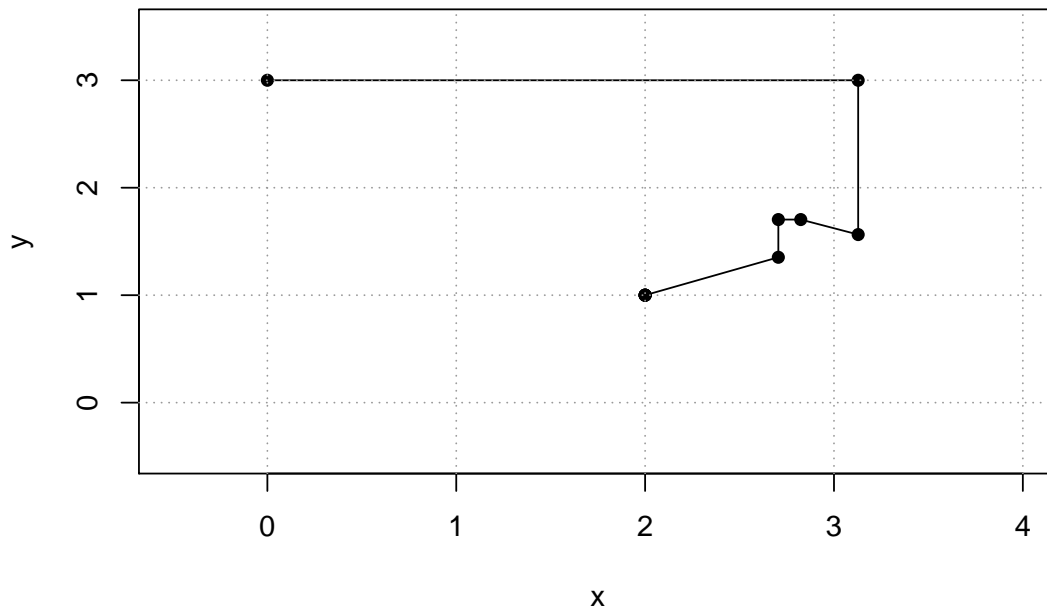


Figura 4.26: Gráfico en XY - Método Hooke y Jeeves ejercicio 1

4.4. Método de búsqueda multivariante usando derivadas

Los métodos que utilizan derivadas las cuestiones básicas de metodología son las mismas:

- 1.- ¿Hacia que dirección moverse en \mathbb{R}^k ?
- 2.- ¿Cuándo parar las iteraciones?

Obviamente la respuesta de la pregunta dos es fácil, parar las iteraciones cuando f se empiece a incrementar o disminuir.

Una elección natural direccional es determinada mediante calculo: hacer que f crezca o decrezca tan rápido como sea posible.

Una manera de lograr esto es dada por el gradiente de $f(x)$.

Los métodos de gradientes determinan la mejor dirección y la mejor longitud de recorrido en esa dirección, para poder alcanzar el máximo o mínimo, en el menor tiempo posible (menos números de iteraciones).

4.4.0.1. Método de paso descendente

Procedimiento

- **Paso 1:** Elegir el punto inicial $x_1 \in \mathbb{R}^k$ y calculamos $\nabla f(x_1)$ para $k = 1$.
- **Paso 2:** Calculamos la longitud de paso $\bar{s} > 0$ tal que $f(x_k - \bar{s}\nabla f(x_k))$ sea mínimo.
- **Paso 3:** Sea $x_{k+1} = x_k - \bar{s}\nabla f(x_k)$ calculamos $\nabla f(x_{k+1})$. Si $\nabla f(x_{k+1}) = 0$ o suficientemente pequeño paramos. En caso contrario $k = k + 1$ y vamos al paso 2.

Código R

```
optmul_nores_pasosdescendentes = function(funx,gradfunx, vx.ini,
                                           epsilon=0.01,
                                           nummax.iter=1000,
                                           report.calculos=T,
                                           cota.inf=0.00001,
                                           cota.sup=1000) {
  # longinc, epsilon > 0
  #browser()
  mres = NULL
  k = 1
  #vyi = vx.ini
  vxki = vx.ini
  ndimension = length(vx.ini)
  gradfxi = gradfunx(vxki)
  repeat {
    vD = gradfxi
    fxi = funx(vxki)
    ffunxj = function(lambda) funx(vxki-lambda*vD)
```

```

sol.uni = optimize(ffunxj,c(cota.inf,cota.sup),maximum = F)
vxki1 = vxki - sol.uni$minimum * vD
gradfxi1 = gradfunx(vxki1)
norma.l2 = sqrt(sum((gradfxi1)^2))
vres = c(k,vxki,fxi,-gradfxi,sol.uni$minimum,
        vxki1,norma.l2,norma.l2<epsilon)
if (report.calculos) {
  if (is.null(mres)) {
    mres = vres
  } else {
    mres = rbind(mres,vres)
  }
}
if ((norma.l2<epsilon) | (k>=nummax.iter)) {
  break
} else {
  k = k+1
  vxki = vxki1
  gradfxi = gradfunx(vxki)
}
}# final repeat
if (report.calculos) {
  colnames(mres) = c("k",paste0("vxki",1:ndimension),
                    "fxi",paste0("vd",1:ndimension),"lambda",
                    paste0("vxkimas1",1:ndimension),"NormaGrad","Fin")
  rownames(mres) = NULL
  dfres = as.data.frame(mres)
} else {
  names(vres) = c("k",paste0("vxki",1:ndimension),
                 "fxi",paste0("vd",1:ndimension),"lambda",
                 paste0("vxkimas1",1:ndimension),"NormaGrad","Fin")
  dfres = as.data.frame(vres)
}
print("Óptimo en vxkimas1")
return(dfres)
}

```

Ejercicio 1

Minimizar $(x_1 - 2)^4 + (x_1 - 2x_2)^2$ con un punto inicial $(0, 3)$.

```

vx.ini = c(0,3)
funx = function(vx) (vx[1]-2)^4+(vx[1]-2*vx[2])^2
gradfunx = function(vx) c(4*(vx[1]-2)^3+2*(vx[1]-2*vx[2]),-4*(vx[1]-2*vx[2]))
df.res = optmul_nores_pasosdescendentes(funx,gradfunx,vx.ini,
                                         epsilon=0.0001,nummax.iter = 1000,
                                         report.calculos=T)

```

```
## [1] "Óptimo en vxkimas1"
```

```
knitr::kable(df.res[c(1:8,nrow(df.res)),],digits = 3,
             caption = "\\label{tabla01}Ejercicio 1
             - Método paso descendente") %>%
kable_styling(latex_options = c("striped", "HOLD_position",
"repeat_header", "scale_down"),
             repeat_header_text = "(continua)")
```

Tabla 4.17: Ejercicio 1 - Método paso descendente

	k	vxki1	vxki2	fxi	vd1	vd2	lambda	vxkimas11	vxkimas12	NormaGrad	Fin
1	1	0.000	3.000	52.000	44.000	-24.000	0.062	2.708	1.523	1.544	0
2	2	2.708	1.523	0.365	-0.740	-1.355	0.231	2.537	1.210	0.969	0
3	3	2.537	1.210	0.096	-0.851	0.465	0.112	2.442	1.262	0.376	0
4	4	2.442	1.262	0.045	-0.180	-0.330	0.268	2.394	1.174	0.382	0
5	5	2.394	1.174	0.026	-0.336	0.183	0.125	2.352	1.197	0.190	0
6	6	2.352	1.197	0.017	-0.091	-0.166	0.280	2.326	1.150	0.218	0
7	7	2.326	1.150	0.012	-0.191	0.105	0.131	2.301	1.164	0.119	0
8	8	2.301	1.164	0.009	-0.057	-0.105	0.287	2.285	1.134	0.145	0
839	839	2.029	1.014	0.000	0.000	0.000	0.145	2.029	1.014	0.000	1

Minimizamos la función $f(x) = (x_1 - 2)^4 + (x_2 - 2x_2)^2$ partiendo del punto inicial $(0, 3)$. En la iteración número 839 se encuentra el óptimo y en el punto $(2.029, 1.014)$. El valor de la función en ese punto es 0.00000170728 .

También se puede ver como evoluciona el algoritmo gráficamente en 2D sobre las curvas de nivel:

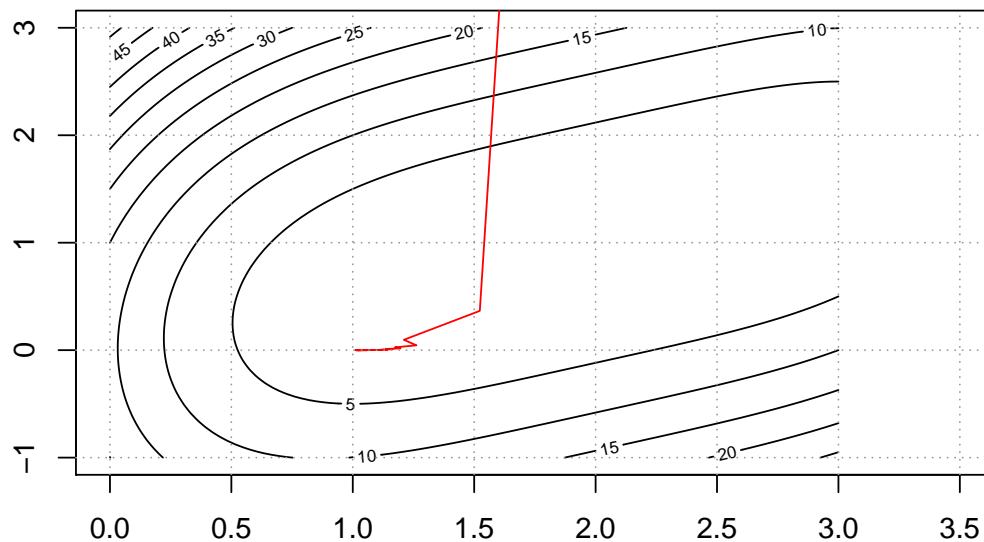


Figura 4.27: Gráfico en 2D - Método paso descendente ejercicio 1

En el siguiente código, se eligen los niveles para los que se representan las curvas de nivel:

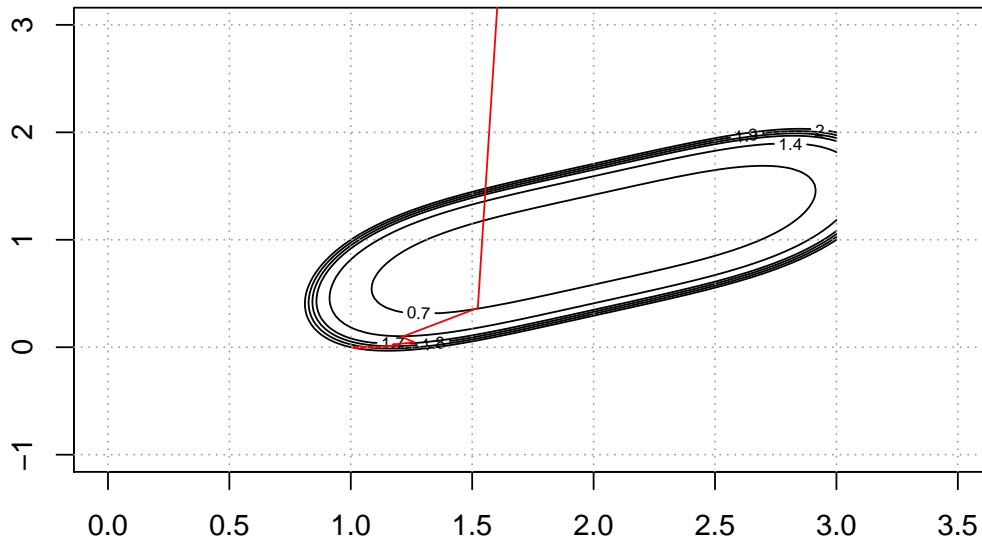


Figura 4.28: Gráfico - Método paso descendente ejercicio 1

4.4.0.2. Método de Newton

En el caso multidimensional la aproximación cuadrática de una función dos veces diferenciable viene dada por:

$$q(x) = \theta(x_k) + \nabla\theta(x_k)'(x - x_k) + \frac{1}{2}(x - x_k)'H(x_k)(x - x_k)$$

la elección del nuevo punto x_{k+1} se realiza de forma

$$\nabla q(x) = \nabla\theta(x_k) + H(x_k)(x - x_k) = 0$$

por lo que, en caso de que la matriz hessiana sea invertible,

$$\Rightarrow x_{k+1} = x_k - H(x_k)^{-1}\nabla\theta(x_k)$$

Dado un escalar positivo prefijado ε , el procedimiento termina cuando

$$\|x_{k+1} - x_k\| < \varepsilon$$

o cuando

$$\|\nabla\theta(x_k)\| < \varepsilon$$

Código R

```

newtonmultidimensional=function(fun, x, y, epsilon, D){
  #,cota.inf=0.00001, cota.sup=1000) No son necesarias ,
  #no hay etapas unidimensionales.
  #(Nos vamos a mover a base de gradientes y hessianos)

  #D=número máximo de iteraciones
  k=1 #primera iteración
  mresultados=NULL #La matriz de resultados empieza nula
  derivada1=deriv(fun, c("x","y"), hessian=TRUE, func=T)
  #Calculamos la derivada de la función con respecto a x e y
  grad=function(x,y) attr(derivada1(x,y), "gradient") #función gradiente
  hess=function(x,y) matrix(as.vector(attr(derivada1(x,y),
                                          "hessian")), nrow=2,ncol=2)

  #función hessiano (vector)
  repeat{
    vfun=derivada1(x,y)[1] #Valor de la funcion en x e y
    grad1=grad(x,y) #Gradiente
    hess1=hess(x,y) #Hessiano
    hessinversa=solve(hess(x,y))
    hessgradiente=as.vector(hessinversa%*%t(grad(x,y)))
    #Nuevos puntos para la siguiente iteración
    x1=x-hessgradiente[1]
    y1=y-hessgradiente[2]

    norma=norm(grad(x,y),type="F")

    vresultado=c(k,x,y,vfun,as.vector(grad1),as.vector(hess1),
                as.vector(hessinversa),
                as.vector(hessgradiente),x1,y1,norma)
    if(is.null(mresultados)) {
      mresultados=vresultado #igualamos la matriz de
      #resultados(nula) al primer vector de resultados
    } else {
      mresultados=rbind(mresultados,vresultado)
      #a partir de la segunda iteración vamos añadiendo
      #nuevos vectores calculados a la matriz de resultados
    }
    if (norma<epsilon || k==D){
      break #si cumple esa condición de rompe el bucle. Hemos llegado al
      #óptimo
    } else{ #si no se cumple lo
      #anterior,continua con los puntos calculados anteriormente para
      #la siguiente iteración
      x=x1
      y=y1
      k=k+1
    }
  }
}

```

```

    }
  }
  colnames(mresultados)=c("k", "x1", "x2", "valordelafun",
                          "gradx1", "gradx2", "Hess11",
                          "Hess12", "Hess21",
                          "Hess22", "InversaHess11", "InversaHess12",
                          "InversaHess21",
                          "InversaHess22", "grad1*Hessx1", "grad1*Hessx2",
                          "x1mas1", "x2mas1", "norma")
  rownames(mresultados)=NULL
  datafresultados=as.data.frame(mresultados)
  print("Óptimo en x1mas1, x2mas1")
  return(datafresultados)
}

```

Ejercicio 1

Minimizar $(x_1 - 2)^4 + (x_1 - 2x_2)^2$

```

fun=expression((x-2)^4+(x-2*y)^2)
resfin2=newtonmultidimensional(fun,0,3,epsilon=0.1,D=1000)

```

```
## [1] "Óptimo en x1mas1, x2mas1"
```

```

knitr::kable(resfin2,digits = 3,
             caption = "\\label{tabla01}Ejercicio 1
             - Método Newton Raphson") %>%
kable_styling(latex_options = c("striped", "HOLD_position",
"repeat_header", "scale_down"),
repeat_header_text = "(continua)")

```

Tabla 4.18: Ejercicio 1 - Método Newton Raphson

k	x1	x2	valordelafun	gradx1	gradx2	Hess11	Hess12	Hess21	Hess22	InversaHess11	InversaHess12	InversaHess21	InversaHess22	grad1*Hessx1	grad1*Hessx2	x1mas1	x2mas1	norma
1	0.000	3.000	52.000	-44.000	24	50.000	-4	-4	8	0.021	0.010	0.010	0.130	-0.667	2.667	0.667	0.333	50.120
2	0.667	0.333	3.160	-9.481	0	23.333	-4	-4	8	0.047	0.023	0.023	0.137	-0.444	-0.222	1.111	0.556	9.481
3	1.111	0.556	0.624	-2.809	0	11.481	-4	-4	8	0.105	0.053	0.053	0.151	-0.296	-0.148	1.407	0.704	2.809
4	1.407	0.704	0.123	-0.832	0	6.214	-4	-4	8	0.237	0.119	0.119	0.184	-0.198	-0.099	1.605	0.802	0.832
5	1.605	0.802	0.024	-0.247	0	3.873	-4	-4	8	0.534	0.267	0.267	0.258	-0.132	-0.066	1.737	0.868	0.247
6	1.737	0.868	0.005	-0.073	0	2.832	-4	-4	8	1.201	0.601	0.601	0.425	-0.088	-0.044	1.824	0.912	0.073

Minimizamos la función $f(x) = (x_1 - 2)^4 + (x_1 - 2x_2)^2$ partiendo del punto inicial $(0, 3)$. La tabla anterior muestra un resumen de los cálculos. En cada iteración $x_{k+1} = x_k - H(x_k)^{-1}f(x_k)$. Después de 6 iteraciones se obtiene el punto $(0.9122085, 1.8244170)$. El valor de la función es 0.004811659 y el procedimiento es terminado.

Evolución gráfica en 3D:

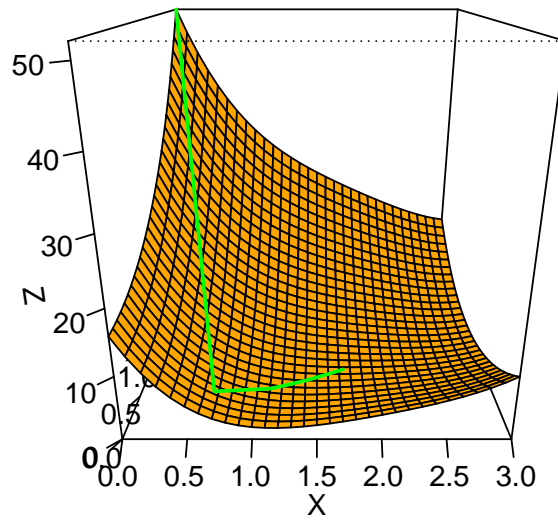


Figura 4.29: Gráfico en 3D - Método Newton Raphson ejercicio 1

Evolución gráfica en 2D sobre las curvas de nivel:

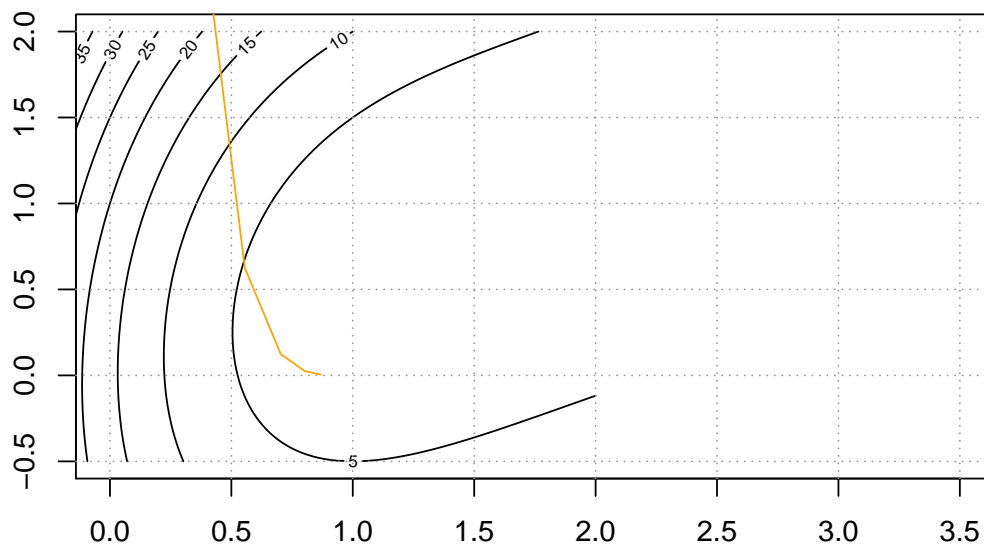


Figura 4.30: Gráfico en 2D - Método Newton Raphson ejercicio 1

Evolución gráfica en el eje XY:

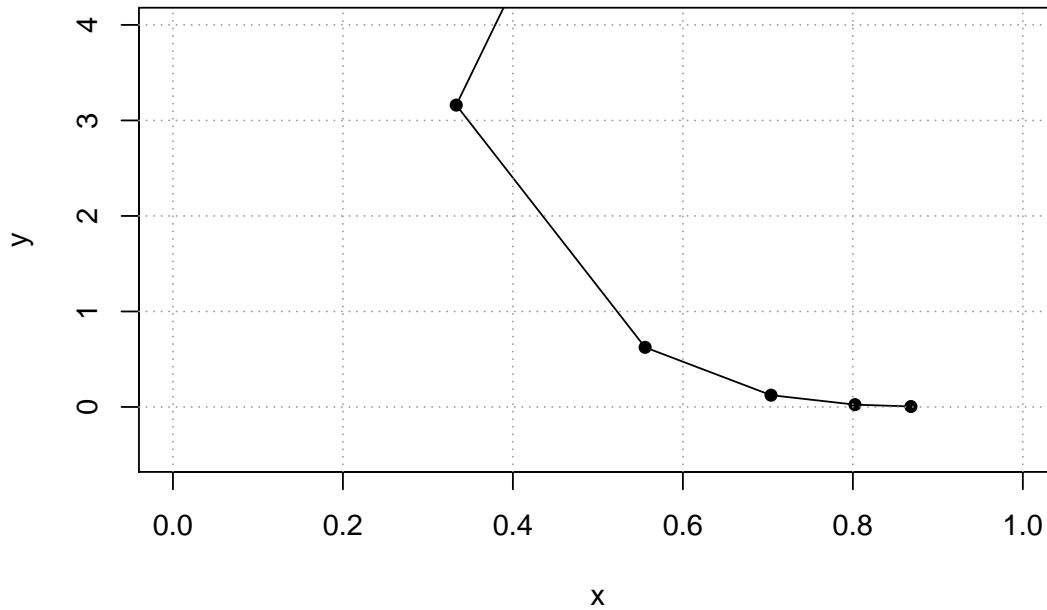


Figura 4.31: Gráfico - Método Newton Raphson ejercicio 1

4.4.1. Optimización multidimensional con restricciones

4.4.1.1. Métodos de penalizaciones

Sea $f : S \subseteq \mathbb{R}^x \rightarrow \mathbb{R}$ consideramos el Problema de Optimización Multiple con restricciones de desigualdad e igualdad

$$\begin{aligned} & \text{Min} \quad f(x) \\ \text{s.a.} \quad & g(x) \leq 0 & g : \mathbb{R}^x \rightarrow \mathbb{R}^> \\ & h(x) = 0 & h : \mathbb{R}^x \rightarrow \mathbb{R}^{\uparrow} \end{aligned}$$

Asociado con este problema está el problema penalizado:

$$\begin{aligned} & \text{Min} \quad f(x) + \mu\alpha(x) \\ \text{s.a.} \quad & x \in \mathbb{R}^x \end{aligned}$$

donde $\alpha(x)$ es la función penalizadora definida por

$$\alpha(x) = \sum_{i=1}^m \emptyset[g_i(x)] + \sum_{i=1}^k \psi[h_i(x)]$$

con \emptyset y ψ funciones continuas que verifican:

$$\phi(y) = \begin{cases} 0 & \text{si } y \leq 0 \\ > 0 & \text{si } y > 0 \end{cases}$$

$$\psi(y) = \begin{cases} 0 & \text{si } y = 0 \\ > 0 & \text{si } y \neq 0 \end{cases}$$

Las funciones más empleadas son:

$$\phi(y) = [\max\{0, y\}]^p$$

$$\psi(y) = |y|^p$$

Procedimiento

- **Paso 0** Sea $\epsilon > 0$ el escalar empleado como criterio de parada.

Elegimos:

- El punto inicial x_1 .
- Un parámetro de penalización $\mu_1 > 0$.
- Un escalar $\beta > 1$.
- $k = 1$.

• **Paso 1** Resolver el problema $\text{Min } f(x) + \mu_k \alpha(x)$ s.a : $x \in \mathbb{R}$ sea x_{k+1} la solución óptima de este problema, pasamos al paso 2.

- **Paso 2** Si $\mu_k \alpha(x_{k+1}) < \epsilon$ parar en otro caso $\mu_{k+1} = \beta \mu_k$, $k = k + 1$.

Código R

```
optmul_res_penalizaciones = function(vx.ini, funx, lrest.igualdad,
                                     phi.igualdad=function(x) abs(x)^2,
                                     mu1 = 0.1, beta=10, epsilon=0.01,
                                     nummax.iter = 1000, report.calculos=T) {
  #epsilon > 0
  #browser()
  mres = NULL
  k = 1
  vxki = vx.ini
  ndimension = length(vx.ini)
  f.alfa = function(vx, muk) {
    pri = funx(vx)
    pri = pri + muk * phi.igualdad(lrest.igualdad[[1]](vx))
    return(pri)
  }
  #mu1 = 0.1
  #beta = 10
  muk = mu1
  # Paso 3
  repeat {
    fmu.temp = function(vxki) f.alfa(vxki, muk)
```

```

solk = optim(vxki,fmu.temp)
vxki1 = solk$par
fpvxki1 = solk$value
fvxki1 = funx(vxki1)
alfavki1 = lrest.igualdad[[1]](vxki1)^2
muk.alfavki1 = muk * alfavki1
vres = c(k,muk,vxki1,fvxki1,fpvxki1,alfavki1,muk.alfavki1)
if (report.calculos) {
  if (is.null(mres)) {
    mres = vres
  } else {
    mres = rbind(mres,vres)
  }
}
if ((muk.alfavki1<epsilon) | (k>=nummax.iter)) {
  break
} else {
  k = k+1
  vxki = vxki1
  muk = muk*beta
}
} # final repeat
if (report.calculos) {
  colnames(mres) = c("k","muk",paste0("vxkmas1",1:ndimension),"fvxkmas1",
                    "fpvxkmas1",
                    "h2vxkmas1","mukporh2vxkmas1")
  rownames(mres) = NULL
  dfres = as.data.frame(mres)
} else {
  names(vres) = c("k","muk",paste0("vxkmas1",1:ndimension),
                "fvxkmas1","fpvxkmas1","h2vxkmas1","mukporh2vxkmas1")
  dfres = as.data.frame(vres)
}
print("Óptimo en vxkmas1")
return(dfres)
}

```

Ejercicio 1

Lo utilizamos sobre la función $f(x_1, x_2) = (x_1 - 2)^4 + (x_1 - 2x_2)^2$, sujeto a $x_1^2 - x_2 = 0$.

```

vx.ini = c(2,1)
funx = function(vx) (vx[1]-2)^4+(vx[1]-2*vx[2])^2
lrest.igualdad = list(function(vx) (vx[1]^2-vx[2]))
df.res = optimul_res_penalizaciones(vx.ini,funx,lrest.igualdad,
                                   phi.igualdad=function(x) abs(x)^2,mu1 = 0.1,
                                   beta=10, epsilon=0.0001,nummax.iter = 1000,
                                   report.calculos=T)

```

```
## [1] "Óptimo en vxkmas1"
```

```
knitr::kable(df.res[c(1:5,nrow(df.res)),],digits = 5, booktabs=TRUE,
             longtable=TRUE,caption = "\\label{tabla01}Ejercicio 1 -
             Método de penalización") %>%
kable_styling(latex_options = c("striped", "HOLD_position",
"repeat_header", "scale_down"),
repeat_header_text = "(continua)")
```

Tabla 4.19: Ejercicio 1 - Método de penalización

	k	muk	vxkmas11	vxkmas12	fvxkmas1	fpvxkmas1	h2vxkmas1	mukporh2vxkmas1
1	1	1e-01	1.45393	0.76077	0.09349	0.27659	1.83098	0.18310
2	2	1e+00	1.16870	0.74067	0.57529	0.96617	0.39088	0.39088
3	3	1e+01	0.99059	0.84239	1.52005	1.71295	0.01929	0.19290
4	4	1e+02	0.95077	0.88749	1.89128	1.91840	0.00027	0.02712
5	5	1e+03	0.94610	0.89342	1.94053	1.94335	0.00000	0.00282
7	7	1e+05	0.94558	0.89411	1.94613	1.94616	0.00000	0.00003

El objetivo es minimizar $(x_1 - 2)^4 + (x_1 - 2x_2)^2 + \mu_k(x_1^2 - x_2)^2$, $x_1 = (2, 1)$ $\mu_1 = 0.1$, $\beta = 10$.

Llegamos al óptimo en la iteración número 6, con $\mu = 10000$, $x_7 = (0.9454, 0.8935)$. La función en ese punto vale 1.9456.

La función penalizadora vale 0.

4.4.1.2. Métodos de barrera

Sea $f : S \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ consideramos el problema de optimización múltiple con restricciones

$$\begin{aligned} \text{Min} \quad & f(x) \\ \text{s.a} \quad & g(x) \leq 0 \\ & x \in S \quad g : \mathbb{R}^n \rightarrow \mathbb{R}^p \end{aligned}$$

Asociado con este problema está el problema de barrera

$$\text{Min} \quad \theta(\mu) = f(x) + \mu B(x) \quad \mu > 0$$

Función de Barrera

La función $B(x)$ es una función de barrera que cumple las propiedades de ser no negativa en la región S y de tender a infinito cuando se acerca a la frontera de S .

Las funciones más empleadas son

$$B(x) = \sum_{i=1}^m \frac{-1}{g_i(x)}$$

$$B(x) = - \sum_{i=1}^m \ln(-g_i(x))$$

Procedimiento

- **Paso 0** Sea $\epsilon > 0$ el escalar utilizado como criterio de parada.

Elegimos:

- El punto inicial $x_1 \in S = \{x \in \mathbb{R}^k : \delta(\curvearrowright) < \mathcal{K}\} \neq \emptyset$.

-Un parámetro de penalización $\mu_1 > 0$.

-Un escalar $0 < \beta < 1$.

$k = 1$

- **Paso 1** Resolver el problema:

$$\text{Min } f(x) + \mu_k B(x)$$

$$\text{s.a : } x \in S$$

Sea x_{k+1} la solución óptima del problema obtenido por algún método de optimización sin restricciones.

- **Paso 2** Si $u_k B(x_{k+1}) < \epsilon$ paramos, en caso contrario $\mu_{k+1} = \beta \mu_k$.

Ejercicio 1

Minimizar $(x_1 - 2)^4 + (x_1 - 2x_2)^2$ sujeto a $x_1^2 - x_2 \leq 0$. Vamos a minimizar $(x_1 - 2)^4 + (x_1 - 2x_2)^2 + \mu \frac{-1}{(x_1^2 - x_2)}$, punto inicial $(0, 1)$, $\mu_1 = 10$, $\beta = 0.1$.

```
f=function(x) {
  x1=x[1]
  x2=x[2]
  return((x1-2)^4 + (x1-2*x2)^2 - mu/(x1^2 -x2))
}
beta=0.1
mu=10
epsilon=0.001
x=c(0,1)
(soluc=optim(x,f))
```

```
## $par
## [1] 0.7080021 1.5315815
##
## $value
## [1] 18.03898
##
```

```
## $counts
## function gradient
##      61      NA
##
## $convergence
## [1] 0
##
## $message
## NULL
```

```
(vx.sol=soluc$par)
```

```
## [1] 0.7080021 1.5315815
```

```
(fvx.sol=soluc$value)
```

```
## [1] 18.03898
```

```
repeat
{
  cat("La solución es",vx.sol,"\n")
  cat("La función penalizada vale", fvx.sol,"\n")
  mu=mu*beta
  x=vx.sol
  (soluc=optim(x,f))
  (vx.sol=soluc$par)
  (fvx.sol=soluc$value)
  p=-mu/(vx.sol[1]^2 -vx.sol[2])
  if (p<epsilon) break
}
```

```
## La solución es 0.7080021 1.531581
## La función penalizada vale 18.03898
```

4.5. Fundamentos de Optimización Restringida

4.5.1. Condiciones de Karush-Kuhn y Tucker y optimalidad

4.5.1.1. Problemas de Optimización con Restricciones

Definición 4.5.1 Se considera el Problema de Optimización con restricciones, definido como sigue

$$\begin{aligned} & \text{Min} \quad f(x) \\ & \text{s.a} \quad g_i(x) \leq 0 \quad \forall i = 1, 2, \dots, m \end{aligned}$$

$$x \in X \subseteq \mathbb{R}^n$$

4.5.1.1.1. Teorema Condiciones Necesaria de Karush, Kuhn y Tucker (K-K-T)

Sea $X \subseteq \mathbb{R}^n$ un conjunto abierto no vacío. Se considera el problema de optimización con restricciones anteriormente definido. Sea \bar{x} factible y supongamos que f y g_i son diferenciables en \bar{x} . Sea $I = \{i \mid g_i(\bar{x}) = 0\}$. A este conjunto se le conoce como el conjunto de las restricciones activas en \bar{x} . Además se verifica que $\{\nabla g_i(\bar{x})\}$ para cada $i \in I$ son linealmente independientes (cualificación de restricciones). Si \bar{x} es una solución local de problema definido, entonces existen escalares μ_i , tal que se verifica:

$$\nabla f(\bar{x}) + \sum_{i=1}^m \mu_i \nabla g_i(\bar{x}) = 0 \quad (1)$$

$$\mu_i g_i(\bar{x}) = 0, \quad \forall i = 1, 2, \dots, m \quad (2)$$

$$\mu_i \geq 0, \quad \forall i = 1, 2, \dots, m \quad (3)$$

Observaciones

La ecuación (1) es una igualdad vectorial que se puede escribir como

$$\frac{\partial f(\bar{x})}{\partial x_j} + \sum_{i=1}^m \mu_i \frac{\partial g_i(\bar{x})}{\partial x_j} = 0, \quad j = 1, \dots, n$$

Observación

Si no se verifica la cualificación de restricciones, el mínimo no tiene por que ser punto K-T, como se puede ver en el siguiente ejemplo.

Ejemplo

$$\begin{aligned} \text{Min} \quad & -x_1 \\ \text{s.a} \quad & x_2 - (1 - x_1)^3 \leq 0 \\ & x_1, x_2 \geq 0 \end{aligned}$$

Buscamos los puntos K-T. Empezamos por definir la función del lagrangiano para aplicar la expresión del teorema. En este caso se tiene que

$$\begin{aligned} f(x_1, x_2) &= -x_1 \\ g(x_1, x_2) &= x_2 - (1 - x_1)^3 \\ g_2(x_1, x_2) &= -x_1 \\ g_3(x_1, x_2) &= -x_2 \end{aligned}$$

Por tanto se tiene que:

$$L(x, \lambda) = -x_1 + \lambda_1(x_2 - (1 - x_1)^3) - \lambda_2 x_1 - \lambda_3 x_2$$

Y las derivadas parciales:

$$\frac{\partial L}{\partial x_1} = -1 + 3\lambda_1(1 - x_1)^2 - \lambda_2 = 0$$

$$\frac{\partial L}{\partial x_2} = \lambda_1 - \lambda_3 = 0$$

$$\lambda_1(x_2 - (1 - x_1)^3) = 0$$

$$\lambda_2 x_1 = 0$$

$$\lambda_3 x_2 = 0$$

$$x_2 - (1 - x_1)^3 \leq 0$$

$$x_1, x_2 \geq 0$$

Veamos ahora qué valores pueden tomar los escalares λ_i

- $(0, 0, 0)$

$-1 = 0$ imposible

- $(\lambda_1, 0, 0)$

$$-1 + 3\lambda_1 x_1^2 - 3\lambda_1 x_1 + 3\lambda_1 = 0$$

$$\lambda_1 - \lambda_3 = 0$$

Imposible porque $\lambda_1 = \lambda_3$ y $\lambda_1 > 0$ y $\lambda_3 = 0$

- $(0, \lambda_2, 0)$

$\lambda_2 = -1$ No vale.

- $(0, 0, \lambda_3)$

$-1 = 0$ Imposible.

- $(\lambda_1, \lambda_2, 0)$

Imposible porque $\lambda_1 = \lambda_3$ y $\lambda_1 > 0$ y $\lambda_3 = 0$

- $(\lambda_1, 0, \lambda_3)$

$$-1 + 3\lambda_1 x_1^2 - 6\lambda_1 x_1 + 3\lambda_1 = 0$$

$$\lambda_1 - \lambda_3 = 0$$

Como $\lambda_3 > 0 \Rightarrow x_2 = 0 \Rightarrow x_1 = 1$, por ser $\lambda_1 > 0$

Así pues $-1 + 3\lambda_1 - 6\lambda_1 + 3\lambda_1 = 0 \Rightarrow -1 = 0$ Imposible

- $(0, \lambda_2, \lambda_3)$

$$-1 - \lambda_2 = 0 \Rightarrow \lambda_2 = -1$$

No vale.

- $(\lambda_1, \lambda_2, \lambda_3)$

$$x_1 = x_2 = 0$$

$$x_2 = (1 - x_1)^3 \Rightarrow x_2 = 1 \text{ Contradicción.}$$

No hay puntos K-T. Sin embargo si se resuelve el problema de manera gráfica.

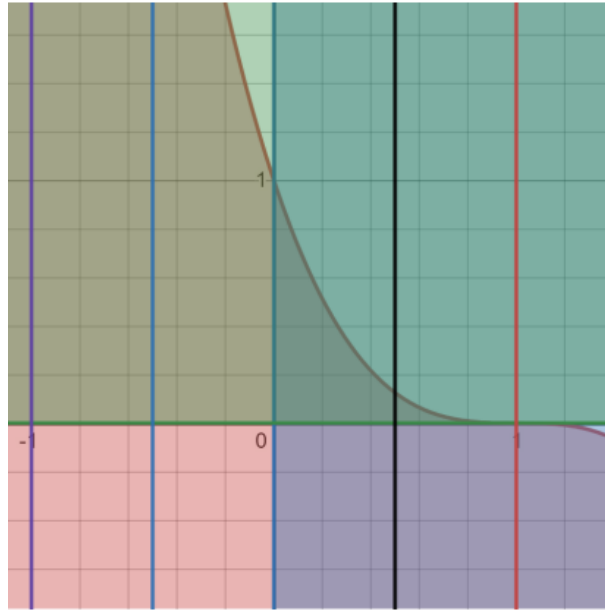


Figura 4.32: Ejemplo condiciones necesarias de Karush, Kuhn y Tucker

Se puede ver que la solución es el punto $(1, 0)$, que no es punto K-T. La razón es que no se verifica la cualificación de restricciones. Las restricciones activas en el punto $(1, 0)$ son:

$$g_1(x_1, x_2) = x_2 - (1 - x_1)^3, \text{ pues } g_1(1, 0) = 0$$

$$g_3(x_1, x_2) = -x_2, \text{ pues } g_3(1, 0) = 0$$

Y se tiene que:

$$\nabla g_1(x_1, x_2) = \begin{pmatrix} 3(1 - x_1)^2 \\ 1 \end{pmatrix}$$

$$\nabla g_3(x_1, x_2) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$$

Que evaluados en el punto $(1, 0)$ queda:

$$\nabla g_1(1, 0) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad y \quad \nabla g_3(1, 0) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$$

Obviamente, no son vectores linealmente independientes.

Observación

La condición de K-K-T es necesaria, pero no suficiente.

Ejemplo

$$\begin{aligned} \text{Min} \quad & -x_2 \\ \text{s.a} \quad & x_1^2 + x_2^2 \leq 4 \\ & -x_1^2 + x_2 \leq 0 \end{aligned}$$

$$x_1, x_2 \geq 0$$

Calculamos el lagrangiano y planteamos las ecuaciones de K-T.

$$L(x, \lambda) = -x_2 + \lambda_1(x_1^2 + x_2^2 - 4) + \lambda_2(-x_1^2 + x_2) - \lambda_3x_1 - \lambda_4x_2$$

$$\frac{\partial L}{\partial x_1} = 2\lambda_1x_1 - 2\lambda_2x_1 - \lambda_3 = 0$$

$$\frac{\partial L}{\partial x_2} = -1 + 2\lambda_1x_2 + \lambda_2 - \lambda_4 = 0$$

$$\lambda_1(x_1^2 + x_2^2 - 4) = 0$$

$$\lambda_2(-x_1^2 + x_2) = 0$$

$$\lambda_3x_1 = 0$$

$$\lambda_4x_2 = 0$$

$$x_1^2 + x_2^2 \leq 4$$

$$-x_1^2 + x_2 \leq 0$$

$$x_1, x_2 \geq 0$$

- $(0, 0, 0, 0)$
 $-1 = 0$ Imposible.
- $(\lambda_1, 0, 0, 0)$

$$\left. \begin{array}{l} 2\lambda_1x_1 = 0 \Rightarrow x_1 = 0 \\ -1 + 2\lambda_1x_2 = 0 \\ x_1^2 + x_2^2 = 4 \Rightarrow x_2^2 = 4 \Rightarrow x_2 = 2 \end{array} \right\} \Rightarrow \text{Infactible}$$

- $(0, \lambda_2, 0, 0)$

$$\left. \begin{array}{l} 2\lambda_2x_1 = 0 \Rightarrow x_1 = 0 \\ -1 + 2\lambda_2 = 0 \Rightarrow \lambda_2 = 1 \\ -x_1^2 + x_2 = 0 \Rightarrow x_2 = 0 \end{array} \right\} \Rightarrow \text{Punto } K - T$$

$(0, 0)$ y $(0, 1, 0, 0)$ cn valor 0 en función objetivo.

- $(0, 0, \lambda_3, 0)$
 $-\lambda_3 = 0$ No vale.

- $(0, 0, 0, \lambda_4)$
 $-1 - \lambda_4 = 0 \Rightarrow \lambda_4 = -1$ No vale.

$$(\lambda_1, \lambda_2, 0, 0)$$

$$2\lambda_1x_1 - 2\lambda_2x_1 = 0 \Rightarrow 2\lambda_1x_1 = 2\lambda_2x_1 \Rightarrow \lambda_1 = \lambda_2 \text{ si } x_1 \neq 0$$

$$-1 + 2\lambda_1x_2 + \lambda_2 = 0$$

$$\left. \begin{array}{l} x_1^2 + x_2^2 = 4 \\ -x_1^2 + x_2 = 0 \Rightarrow x_2 = x_1^2 \end{array} \right\} \Rightarrow x_1^2 + x_1^4 = 4$$

Se hace un cambio de variable $y = x_1^2 \Rightarrow y^2 + y - 4 = 0 \Rightarrow$

$y = \frac{-1 \pm \sqrt{1+16}}{2} = \frac{\sqrt{17}-1}{2} \Rightarrow x_1 = (\frac{\sqrt{17}-1}{2})^{1/2}$, pues la raíz negativa da una solución infactible.

$$x_2 = \frac{\sqrt{17}-1}{2}$$

$$-1 + 2\lambda_1(\frac{\sqrt{17}-1}{2}) + \lambda_1 = 0 \Rightarrow \lambda_1 = \frac{1}{\sqrt{17}} = \lambda_2$$

Punto K-T: $((\frac{\sqrt{17}-1}{2})^{1/2}, \frac{\sqrt{17}-1}{2})$ y $(\frac{1}{\sqrt{17}}, \frac{1}{\sqrt{17}}, 0, 0)$ con valor de la función objetivo $-\frac{\sqrt{17}-1}{2}$

- $(\lambda_1, 0, \lambda_3, 0)$

$$2\lambda_1 x_1 - \lambda_3 = 0 \quad -1 + 2\lambda_1 x_2 = 0$$

$$\left. \begin{array}{l} x_1^2 + x_2^2 = 4 \\ x_1 = 0 \end{array} \right\} \Rightarrow x_2 = 2 \quad \text{Infactible}$$

- $(\lambda_1, 0, 0, \lambda_4)$

$$\left. \begin{array}{l} 2\lambda_1 x_1 = 0 \Rightarrow x_1 = 0 \\ -1 + 2\lambda_1 x_2 - \lambda_4 = 0 \\ x_2 = 0 \end{array} \right\} -1 - \lambda_4 = 0 \Rightarrow \lambda_4 = -1 \quad \text{No vale.}$$

- $(0, \lambda_2, \lambda_3, 0)$

$$\left. \begin{array}{l} 2\lambda_2 x_1 - \lambda_3 = 0 \\ -1 + \lambda_2 = 0 \Rightarrow \lambda_2 = 1 \end{array} \right\} \Rightarrow \lambda_3 = -2x_1 \Rightarrow \lambda_3 \leq 0 \quad \text{No vale.}$$

- $(0, \lambda_2, 0, \lambda_4)$

$$-2\lambda_2 x_1 = 0 \Rightarrow x_1 = 0$$

$$-1 + \lambda_2 - \lambda_4 = 0 \Rightarrow \lambda_4 = \lambda_2 - 1$$

$$-x_1^2 + x_2 = 0$$

$$x_2 = 0$$

Punto K-T: $(0, 0)$ y $(0, \lambda_2, 0, \lambda_2 - 1)$

- $(0, 0, \lambda_3, \lambda_4)$

$$\lambda_3 = 0 \quad \text{No vale}$$

- $(\lambda_1, \lambda_2, \lambda_3, 0)$

$$2\lambda_1 x_1 - 2\lambda_2 x_1 - \lambda_3 = 0$$

$$-1 + 2\lambda_1 x_2 + \lambda_2 = 0$$

$$\left. \begin{array}{l} x_1^2 + x_2^2 = 4 \\ -x_1^2 + x_2 = 0 \\ x_1 = 0 \end{array} \right\} \Rightarrow x_2 = 0 \quad \text{Imposible}$$

- $(\lambda_1, 0, \lambda_3, \lambda_4)$

$$2\lambda_1 x_1 - \lambda_3 = 0$$

$$-1 + 2\lambda_1 x_2 + \lambda_2 - \lambda_4 = 0$$

$$\left. \begin{array}{l} x_1^2 + x_2^2 = 4 \\ x_1 = 0 \\ x_2 = 0 \end{array} \right\} \text{Imposible}$$

- $(\lambda_1, \lambda_2, 0, \lambda_4)$

$$2\lambda_1 x_1 - 2\lambda_2 x_1 = 0$$

$$-1 + 2\lambda_1 x_2 + \lambda_2 - \lambda_4 = 0$$

$$\left. \begin{array}{l} x_1^2 + x_2^2 = 4 \\ -x_1^2 + x_2 = 0 \\ x_2 = 0 \end{array} \right\} \Rightarrow x_1 = x_2 = 0 \quad \text{Imposible}$$

- $(0, \lambda_2, \lambda_3, \lambda_4)$

$$-2\lambda_2 x_1 - \lambda_3 = 0$$

$$-1 + \lambda_2 - \lambda_4 = 0$$

$$\left. \begin{array}{l} -x_1^2 + x_2 = 0 \\ x_1 = 0 \\ x_2 = 0 \end{array} \right\} \Rightarrow x_1 = x_2 = 0$$

$\lambda_3 = 0$ No vale

- $(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$

$$\left. \begin{array}{l} x_1^2 + x_2^2 = 4 \\ -x_1^2 + x_2 = 0 \\ x_1 = 0 \\ x_2 = 0 \end{array} \right\} \text{Imposible}$$

Candidatos a máximo

$(0, 0)$ y $(0, 1, 0, 0)$

$((\frac{\sqrt{17}-1}{2})^{1/2}, \frac{\sqrt{17}-1}{2})$ y $(\frac{1}{\sqrt{17}}, \frac{1}{\sqrt{17}}, 0, 0)$

$(0, 0)$ y $(0, \lambda_2, 0, \lambda_2 - 1)$

El punto $(0, 0)$ es punto de K-T, pero su valor en la función objetivo es mayor que el punto $((\frac{\sqrt{17}-1}{2})^{1/2}, \frac{\sqrt{17}-1}{2})$, por lo que no es el óptimo. De hecho, aplicando el teorema de Weierstrass, se sabe que el máximo es $((\frac{\sqrt{17}-1}{2})^{1/2}, \frac{\sqrt{17}-1}{2})$.

4.5.1.1.2. Teorema Condiciones Suficiente de Karush, Kuhn y Tucker (K-K-T)

Se considera el Problema de Optimización con Restricciones. Si $f(x)$ es pseudoconvexa y g_i son cuasiconvexas diferenciables $\forall i = 1, 2, \dots, m$ y \bar{x} verifica las condiciones de K-K-T, entonces \bar{x} es un mínimo para el Problema de Optimización con Restricciones.

4.5.1.2. Problemas de Optimización con Restricciones de desigualdad e igualdad

Definición 4.5.2 Se considera el Problema de Optimización con Restricciones de desigualdad e igualdad, definido como sigue

$$\begin{aligned} & \text{Min} \quad f(x) \\ & \text{s.a} \quad g_i(x) \leq 0 \quad \forall i = 1, 2, \dots, m \\ & \quad \quad h_i(x) = 0 \quad \forall i = 1, 2, \dots, l \\ & \quad \quad x \in X \subseteq \mathbb{R}^n \\ & \quad \quad (POR1) \end{aligned}$$

4.5.1.2.1. Teorema Condición Necesaria de Karush, Kuhn y Tucker (K-K-T)

Sea $X \subseteq \mathbb{R}^n$ un conjunto abierto no vacío. Se considera (POR1). Sea \bar{x} factible y supongamos que f y g_i son diferenciables en \bar{x} . Si \bar{x} es una solución local de (POR1), entonces existen escalares u_i, v_i tal que se verifica:

$$\nabla f(\bar{x}) + \sum_{i=1}^m u_i \nabla g_i(\bar{x}) + \sum_{i=1}^l v_i \nabla h_i(\bar{x}) = 0 \quad (4)$$

$$u_i g_i = 0, \quad \forall i = 1, 2, \dots, m \quad (5)$$

$$u_i \geq 0, \quad \forall i = 1, 2, \dots, m \quad (6)$$

Observación

La ecuación (4) es una igualdad vectorial que se puede escribir como

$$\frac{\partial f(\bar{x})}{\partial x_j} + \sum_{i=1}^m u_i \frac{\partial g_i(\bar{x})}{\partial x_j} + \sum_{i=1}^l v_i \frac{\partial h_i(\bar{x})}{\partial x_j} = 0, \quad j = 1, \dots, n$$

4.5.1.2.2. Teorema Condición Suficiente de Karush, Kuhn y Tucker (K-K-T)

Se considera el (POR1). Si $f(x)$ es pseudoconvexa y g_i, h_i son cuasiconvexas diferenciables $\forall i$ y \bar{x} verifica las condiciones de K-K-T para (POR1), entonces \bar{x} es un mínimo para (POR1).

4.5.1.3. Ejercicios

1.

$$\begin{aligned}
 &Max \quad x_1 + 3x_2 \\
 &s.a : \quad x_1^2 + x_2^2 \leq 5 \\
 &\quad \quad -x_1 + x_2 \leq 1 \\
 &\quad \quad x_1, x_2 \geq 0
 \end{aligned}$$

$$f(x_1, x_2) = x_1 + 3x_2$$

Es lineal y, por tanto, cóncava y convexa a la vez.

$$g_1(x_1, x_2) = x_1^2 + x_2^2 - 5$$

$$\nabla g_1(x_1, x_2) = \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix}$$

$$Hg_1(x_1, x_2) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

$$\Delta_1 > 0$$

$$\Delta_2 > 0$$

$\Rightarrow g_1$ es estrictamente convexa.

$g_2(x_1, x_2) = -x_1 + x_2 - 1$ es lineal y por tanto, cóncava y convexa. Se verifican las condiciones necesarias y suficientes de Khun y Tucker y cualquier punto crítico es máximo global.

$$L(x, \lambda) = x_1 + 3x_2 - \lambda_1(x_1^2 + x_2^2 - 5) - \lambda_2(-x_1 + x_2 - 1) + \lambda_3x_1 + \lambda_4x_2$$

$$\frac{\partial L}{\partial x_1} = 1 - 2\lambda_1x_1 + \lambda_2 + \lambda_3 = 0$$

$$\frac{\partial L}{\partial x_2} = 3 - 2\lambda_1x_2 - \lambda_2 + \lambda_4 = 0$$

$$\lambda_1(x_1^2 + x_2^2 - 5) = 0$$

$$\lambda_2(-x_1 + x_2 - 1) = 0$$

$$\lambda_3x_1 = 0$$

$$\lambda_4x_2 = 0$$

$$x_1^2 + x_2^2 \leq 5$$

$$x_1, x_2 \geq 0$$

- $(0, 0, 0, 0)$

$1 = 0$ Imposible

- $(\lambda_1, 0, 0, 0)$

$$\left. \begin{aligned}
 1 - 2\lambda_1x_1 = 0 &\Rightarrow \lambda_1 = \frac{1}{2x_1} \quad \text{si } x_1 \neq 0 \\
 3 - 2\lambda_1x_2 = 0 &\Rightarrow \lambda_1 = \frac{3}{2x_2} \quad \text{si } x_2 \neq 0
 \end{aligned} \right\} \Rightarrow \frac{1}{2x_1} = \frac{3}{2x_2} \Rightarrow x_2 = 3x_1$$

$$\left. \begin{array}{l} x_2 = 3x_1 \\ x_1^2 + x_2^2 = 5 \end{array} \right\} \Rightarrow x_1^2 + 9x_1^2 = 5 \Rightarrow 10x_1^2 = 5 \Rightarrow x_1^2 = \frac{1}{2} \Rightarrow x_1 = \frac{1}{\sqrt{2}} = \frac{\sqrt{2}}{2} \Rightarrow x_2 = \frac{3\sqrt{2}}{2}$$

Infactible.

- $(0, \lambda_2, 0, 0)$

$$1 + \lambda_2 = 0 \Rightarrow \lambda_2 = -1 \text{ No vale.}$$

- $(0, 0, \lambda_3, 0)$

$$\lambda_3 = -1 \text{ No vale.}$$

- $(0, 0, 0, \lambda_4)$

$$1 = 0 \text{ Imposible.}$$

- $(\lambda_1, \lambda_2, 0, 0)$

$$1 - 2\lambda_1 x_1 + \lambda_2 = 0$$

$$3 - 2\lambda_1 x - 2 - \lambda_2 = 0$$

$$\left. \begin{array}{l} x_1^2 + x_2^2 \\ -x_1 + x_2 = 1 \end{array} \right\} \Rightarrow \left. \begin{array}{l} x_1^2 + x_2^2 \\ x_1 = x_2 - 1 \end{array} \right\} \Rightarrow (x_2 - 1)^2 + x_2^2 = 5 \Rightarrow x_2^2 - x_2 - 2 = 0$$

$$x_2 = \frac{1 \pm \sqrt{1+8}}{2} \Rightarrow x - 2 = 2 \quad y \quad x_2 = -1$$

Si $x_2 = 2 \Rightarrow x_1 = 1$ y el punto es factible.

Si $x_2 = -1$ infactible.

$$\left. \begin{array}{l} 1 - 2\lambda_1 + \lambda_2 = 0 \\ 3 - 4\lambda_1 - \lambda_2 = 0 \end{array} \right\} \Rightarrow \lambda_1 = \frac{2}{3} > 0 \quad y \quad \lambda_2 = \frac{1}{3} > 0$$

Se tiene un punto crítico, que por el teorema es máximo $(1, 2)$ y $(\frac{2}{3}, \frac{1}{3}, 0, 0)$ con 7 de valor de la función objetivo.

Es posible que exista una solución múltiple y por eso seguimos analizando otros casos, aunque en este caso la gráfica nos dice que es un máximo único.



Figura 4.33: Ejercicio 1

- $(\lambda_1, 0, \lambda_3, 0)$

$$\left. \begin{array}{l} 1 - 2\lambda_1 x_1 + \lambda_3 = 0 \\ 3 - 2\lambda_1 x_2 = 0 \\ x_1 = 0 \end{array} \right\} \Rightarrow \lambda_3 = -1 \quad \text{No vale.}$$

- $(\lambda_1, 0, 0, \lambda_4)$

$$\left. \begin{array}{l} 1 - 2\lambda_1 x_1 = 0 \\ 3 - 2\lambda_1 x_2 + \lambda_4 = 0 \\ x_2 = 0 \end{array} \right\} \Rightarrow \lambda_4 = -3 \quad \text{No vale.}$$

- $(0, \lambda_2, \lambda_3, 0)$

$$\left. \begin{array}{l} 1 - \lambda_2 + \lambda_3 = 0 \\ 3 - \lambda_2 = 0 \\ x_1 = 0 \end{array} \right\} \Rightarrow \lambda_2 = 3 \quad y \quad \lambda_3 = -4 \quad \text{No vale.}$$

- $(0, \lambda_2, 0, \lambda_4)$

$$1 + \lambda_2 = 0 \Rightarrow \lambda_2 = -1 \quad \text{No vale.}$$

- $(0, 0, \lambda_3, \lambda_4)$

$$\lambda_3 = -1 \Rightarrow \text{No vale.}$$

- $(0, \lambda_2, \lambda_3, \lambda_4)$

$$1 + \lambda_2 + \lambda_3 = 0$$

$$1 - \lambda_2 + \lambda_4 = 0$$

$$\left. \begin{array}{l} -x_1 + x_2 = 1 \\ x_1 = 0 \\ x_2 = 0 \end{array} \right\} \Rightarrow \text{Imposible.}$$

- $(\lambda_1, 0, \lambda_3, \lambda_4)$

$$\left. \begin{array}{l} x_1^2 + x_2^2 = 5 \\ x_1 = 0 \\ x_2 = 0 \end{array} \right\} \Rightarrow \text{Imposible.}$$

- $(\lambda_1, \lambda_2, 0, \lambda_4)$

$$\left. \begin{array}{l} x_1^2 + x_2^2 = 5 \\ -x_1 + x_2 = 1 \\ x_1 = 0 \end{array} \right\} \Rightarrow \text{Imposible.}$$

- $(\lambda_1, \lambda_2, \lambda_3, 0)$

$$\left. \begin{array}{l} x_1^2 + x_2^2 = 5 \\ -x_1 + x_2 = 1 \\ x_1 = 0 \end{array} \right\} \Rightarrow \text{Imposible.}$$

- $(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$

$$\left. \begin{array}{l} x_1^2 + x_2^2 = 5 \\ -x_1 + x_2 = 1 \\ x_1 = x_2 = 0 \end{array} \right\} \Rightarrow \text{Imposible.}$$

Por tanto solo existe un máximo, $(1, 2)$ y $(\frac{2}{3}, \frac{1}{3}, 0, 0)$, con valor de la función objetivo igual a 7.

2.

$$\begin{aligned} \text{Opt } & x_1^2 + (x_2 - 1)^2 \\ \text{s.a : } & x_1 + x_2 \leq 6 \\ & 2x_1 + x_2 \geq 4 \Leftrightarrow -2x_1 - x_2 + 4 \leq 0 \\ & x_1, x_2 \geq 0 \end{aligned}$$

$$f(x_1, x_2) = x_1^2 + (x_2 - 1)^2$$

$$\nabla f(x_1, x_2) = \begin{pmatrix} 2x_1 \\ 2x_2 - 2 \end{pmatrix}$$

$$\nabla f(x_1, x_2) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

$\Delta_1 > 0$ y $\Delta_2 > 0 \Rightarrow f$ es estrictamente convexa.

Las restricciones son lineales. Por tanto se verifican las condiciones necesarias y suficientes de Khun y Tucker para un problema de minimizar y cualquier punto crítico es el único mínimo global. Como también hay que maximizar, en este caso no se verifican las condiciones del teorema. El problema se va a plantear con las ecuaciones de minimizar, por lo que si se quieren buscar máximos, se tendrá que hacer con los landas negativos. O sea, mínimo con landas positivos y máximo con landas negativos. Si hay de las dos clases, no es un punto crítico.

$$L(x, \lambda) = x_1^2 + x_2^2 + 1 - 2x_2 + \lambda_1(x_1 + x_2 - 6) + \lambda_2(-2x_1 - x_2 + 4) - \lambda_3x_1 - \lambda_4x_2$$

$$\frac{\partial L}{\partial x_1} = 2x_1 + \lambda_1 - 2\lambda_2 - \lambda_3 = 0$$

$$\frac{\partial L}{\partial x_2} = 2x_2 - 2 + \lambda_1 - \lambda_2 - \lambda_4 = 0$$

$$\lambda_1(x_1 + x_2 - 6) = 0$$

$$\lambda_2(2x_1 + x_2 - 4) = 0$$

$$\lambda_3x_1 = 0$$

$$\lambda_4x_2 = 0$$

$$x_1 + x_2 \leq 6$$

$$2x_1 + x_2 \geq 4$$

$$x_1, x_2 \geq 0$$

- $(0, 0, 0, 0)$

$$\left. \begin{array}{l} 2x_1 = 0 \\ 2x_2 - 2 = 0 \end{array} \right\} \Rightarrow \left. \begin{array}{l} x_1 = 0 \\ x_2 = 1 \end{array} \right\} \Rightarrow \text{Infactible}$$

- $(\lambda_1, 0, 0, 0)$

$$2x_1 + \lambda_1 \Rightarrow \lambda_1 = -2x_1$$

$$2x_2 - 2 + \lambda_1 = 0 \Rightarrow \lambda_1 = 2 - 2x_2$$

$$\text{Por tanto: } -2x_1 = 2 - 2x_2 \Rightarrow -2x_1 + 2x_2 = 2$$

$$\left. \begin{array}{l} -2x_1 + 2x_2 = 2 \\ x_1 + x_2 = 6 \end{array} \right\} \Rightarrow -2(6-x_2) + 2x_2 - 2 = 0 \Rightarrow -12 + 4x_2 = 2 \Rightarrow 4x_2 = 14 \Rightarrow x_2 = \frac{7}{2}, x_1 = \frac{5}{2}, \lambda_1 = -5$$

Se tiene un punto crítico candidato a máximo porque λ_1 es negativo.

$(\frac{5}{2}, \frac{7}{2})$ y $(-5, 0, 0, 0)$ con valor de la función objetivo $\frac{25}{2}$.

- $(0, \lambda_2, 0, 0)$

$$2x_1 + 2\lambda_2 = 0 \Rightarrow \lambda_2 = x_1$$

$$2x_2 - 2 + \lambda_2 = 4 \Rightarrow \lambda_2 = 2x_2 - 2$$

$$\text{Por tanto: } 2x_2 - 2 = x_1 \Rightarrow -x_1 + 2x_2 = 2$$

$$-x_1 + 2x_2 = 2 \Rightarrow -2x_1 + 4x_2 = 4$$

$$2x_1 + x_2 = 4$$

$$\text{Por tanto: } 5x_2 = 8 \Rightarrow x_2 = \frac{8}{5} \quad x_1 = \frac{6}{5} \quad \lambda_2 = \frac{6}{5}$$

Es un punto crítico para el problema de minimizar, lo que significa que es el mínimo global único, porque se cumplen las condiciones necesarias y suficientes del teorema de Kuhn y Tucker y la función objetivo es estrictamente convexa.

$$\left(\frac{6}{5}, \frac{8}{5}\right) \text{ y valor de la función objetivo } \frac{9}{5}.$$

Ya no hay que investigar más mínimos por lo que si en los próximos sistemas de ecuaciones encontramos algún lambda positivo, se puede dejar de resolver ese caso.

- $(0, 0, \lambda_3, 0)$

$$\left. \begin{array}{l} 2x_1 - \lambda_3 = 0 \\ 2x_2 = 0 \\ x_1 = 0 \end{array} \right\} \Rightarrow \lambda_3 = x_1 = x_2 = 0 \quad \text{Infeasible}$$

- $(0, 0, 0, \lambda_4)$

$$\left. \begin{array}{l} 2x_1 = 0 \\ x_2 = 0 \end{array} \right\} \Rightarrow x_1 = x_2 = 0 \quad \text{Infeasible}$$

- $(\lambda_1, \lambda_2, 0, 0)$

$$2x_1 + x_2 = 4 \Rightarrow \left. \begin{array}{l} x_1 + x_2 = 6 \\ -2x_1 - x_2 = -4 \end{array} \right\} \Rightarrow -x_1 = 2 \Rightarrow x_1 = -2 \quad \text{Infeasible}$$

- $(\lambda_1, 0, \lambda_3, 0)$

$$\left. \begin{array}{l} x_1 + x_2 = 6 \\ x_1 = 0 \end{array} \right\} \Rightarrow x_2 = 6$$

$$\left. \begin{array}{l} 2x_1 + \lambda_1 - \lambda_3 = 0 \\ 2x_2 - 2 + \lambda_1 = 0 \end{array} \right\} \Rightarrow \lambda_1 = \lambda_3 = -10$$

Punto crítico candidato a máximo.

$(0, 6)$ y $(-10, 0, -10, 0)$ con valor de la función objetivo 25.

- $(0, \lambda_2, \lambda_3, 0)$

$$\left. \begin{array}{l} 2x_1 + x_2 = 4 \\ x_1 = 0 \end{array} \right\} \Rightarrow x_2 = 4$$

$$\left. \begin{array}{l} 2x_1 + 2\lambda_2 - \lambda_3 = 0 \\ 2x_2 - 2 - \lambda_2 = 0 \end{array} \right\} \Rightarrow \lambda_3 = 12$$

No vale, pues es positivo y ya se sabe que no puede ser mínimo.

- $(0, \lambda_2, 0, \lambda_4)$

$$\left. \begin{array}{l} 2x_1 + x_2 = 4 \\ x_2 = 0 \end{array} \right\} \Rightarrow x_1 = 2$$

$$\left. \begin{array}{l} 2x_1 - 2\lambda_2 = 0 \\ 2x_2 - 2 - \lambda_2 - \lambda_4 = 0 \end{array} \right\} \Rightarrow \lambda_2 = 2, \lambda_4 = -4$$

No vale, son de distintos signos.

- $(0, 0, \lambda_3, \lambda_4)$

$$\left. \begin{array}{l} x_1 = 0 \\ x_2 = 0 \end{array} \right\} \Rightarrow \text{Infactible.}$$

- $(\lambda_1, \lambda_2, \lambda_3, 0)$

$$\left. \begin{array}{l} 2x_1 + x_2 = 4 \\ x_1 + x_2 = 6 \\ x_1 = 0 \end{array} \right\} \Rightarrow \text{Imposible.}$$

- $(\lambda_1, \lambda_2, 0, \lambda_4)$

$$\left. \begin{array}{l} 2x_1 + x_2 = 4 \\ x_1 + x_2 = 6 \\ x_2 = 0 \end{array} \right\} \Rightarrow \text{Imposible.}$$

- $(\lambda_1, 0, \lambda_3, \lambda_4)$

$$\left. \begin{array}{l} x_1 = 0 \\ x_2 = 0 \end{array} \right\} \Rightarrow \text{Infactible.}$$

- $(0, \lambda_2, \lambda_3, \lambda_4)$

$$\left. \begin{array}{l} x_1 = 0 \\ x_2 = 0 \end{array} \right\} \Rightarrow \text{Infactible.}$$

- $(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$

$$\left. \begin{array}{l} x_1 = 0 \\ x_2 = 0 \end{array} \right\} \Rightarrow \text{Infactible.}$$

El problema tiene un mínimo global único $(\frac{6}{5}, \frac{8}{5})$ y valor de la función objetivo $\frac{9}{5}$. Y se tiene que la función está definida en un compacto por lo que tiene máximo, que ha de ser el que tiene mayor valor de la función objetivo entre los puntos críticos candidatos a máximos $(6, 0)$ con 36 de valor de la función objetivo.

Gráficamente:

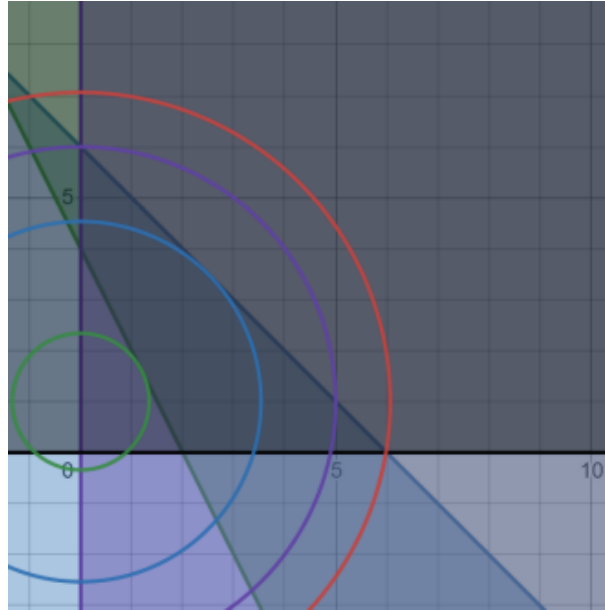


Figura 4.34: Ejercicio 2

3.

$$\begin{aligned} \text{Min} \quad & (x_1 - 3)^2 + (x_2 - 2)^2 \\ \text{s.a:} \quad & x_1^2 + x_2^2 \leq 5 \\ & x_1 + x_2 \leq 3 \\ & x_1, x_2 \geq 0 \end{aligned}$$

$$L(x, \lambda) = x_1^2 + 9 - 6x_1 + x_2^2 + 4 - 4x_2 + \lambda_1(x_1^2 + x_2^2 - 5) + \lambda_2(x_1 + x_2 - 3) - \lambda_3x_1 - \lambda_4x_2$$

$$\frac{\partial L}{\partial x_1} = 2x_1 - 6 + 2\lambda_1x_1 + \lambda_2 - \lambda_3 = 0$$

$$\frac{\partial L}{\partial x_2} = 2x_2 - 4 + 2\lambda_1x_2 + \lambda_2 - \lambda_4 = 0$$

$$\lambda_1(x_1^2 + x_2^2 - 5) = 0$$

$$\lambda_2(x_1 + x_2 - 3) = 0$$

$$\lambda_3x_1 = 0$$

$$\lambda_4x_2 = 0$$

$$x_1^2 + x_2^2 \leq 5$$

$$x_1 + x_2 \leq 3$$

$$x_1, x_2 \geq 0$$

Probamos con el punto $(2, 1)$ para ver si es crítico. Si sustituimos en las ecuaciones se tiene que $\lambda_3 = \lambda_4 = 0$

$$\left. \begin{array}{l} 4 - 6 + 4\lambda_1 + \lambda_2 = 0 \Rightarrow 4\lambda_1 + \lambda_2 = 2 \\ 2 - 4 + 2\lambda_1 + \lambda_2 = 0 \Rightarrow 2\lambda_1 + \lambda_2 = 2 \end{array} \right\} \lambda_1 = 0 \quad y \quad \lambda_2 = 2$$

Por tanto se tiene un punto crítico. Como el problema cumple las condiciones de Kuhn y Tucker, entonces también es un mínimo único con valor 2 en la función objetivo.

Gráficamente:

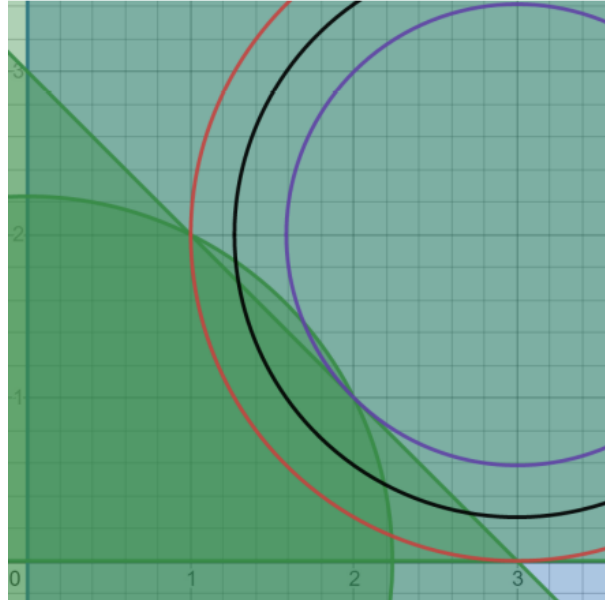


Figura 4.35: Ejercicio 3

4.

$$\begin{array}{ll} \text{Min} & 1 - x_1 + x_2^2 \\ \text{s.a} & x_1^2 + x_2^2 \leq 1 \end{array}$$

La función objetivo se puede cambiarla por $f(x_1, x_2) = x_2^2 - x_1$ porque sumar 1 no hace que cambie el valor donde se encuentra el óptimo.

$$\begin{aligned} \nabla f(x, y) &= \begin{pmatrix} -1 \\ 2x_2 \end{pmatrix} \\ H_f(x, y) &= \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix} \end{aligned}$$

$\Delta_1 = 0$ y $\Delta_2 = 0$, por lo que recurrimos a los menores principales.

Menores principales:

$$D_1 = \begin{cases} 0 \\ 2 \end{cases} \Rightarrow D_1 \geq 0$$

$$D_2 = 0 \geq 0$$

La función es convexa.

$$g_1(x_1, x_2) = x_1^2 + x_2^2 - 1$$

$$\nabla g_1(x_1, x_2) = \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix}$$

$$H_f(x, y) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

$\Delta_1 > 0$ y $\Delta_2 > 0$, por lo que es convexa.

El problema cumple las condiciones del teorema, por lo que los puntos críticos son mínimos.

$$L(x, \lambda) = x_2^2 - x_1 + \lambda(x_1^2 + x_2^2 - 1)$$

$$\frac{\partial L}{\partial x} = -1 + 2\lambda x_1 = 0$$

$$\frac{\partial L}{\partial y} = 2x_2 + 2\lambda x_2 = 0$$

$$\lambda(x_1^2 + x_2^2 - 1) = 0$$

$$x_1^2 + x_2^2 \leq 1$$

- $\lambda = 0$

$-1 = 0 \Rightarrow$ Imposible

- $\lambda \neq 0$

$$\left. \begin{array}{l} -1 + 2\lambda x_1 = 0 \Rightarrow \lambda = \frac{1}{2x_1} \quad \text{si } x_1 \neq 0 \\ 2x_2 + 2\lambda x_2 = 0 \Rightarrow \lambda = -\frac{2x_2}{2x_2} = -1 \quad \text{si } x_2 \neq 0 \end{array} \right\} \Rightarrow \lambda_1 = -1 \Rightarrow \text{No vale.}$$

Así que $x_2 = 0$

$$\left. \begin{array}{l} x_2 = 0 \\ x_1^2 + x_2^2 = 1 \end{array} \right\} \Rightarrow x_1 = 1 \quad y \quad \lambda = \frac{1}{2}$$

Punto crítico y por tanto mínimo con valor -1 en la función objetivo, que corresponde a 0 en la función objetivo inicial.

Gráficamente:

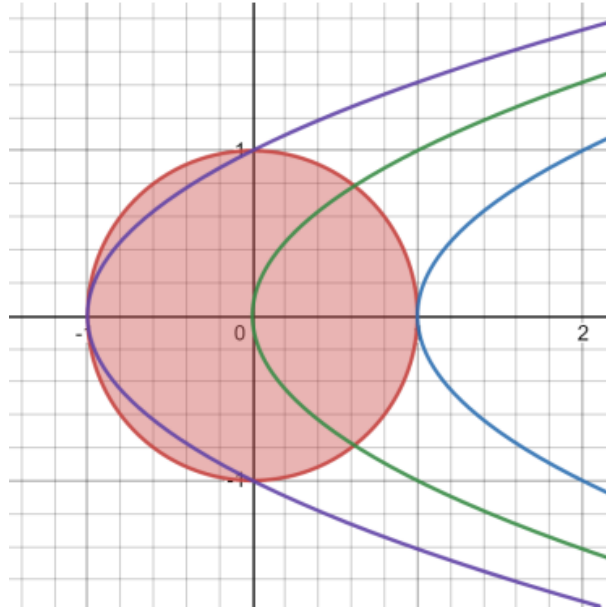


Figura 4.36: Ejercicio 4

5.

$$\begin{aligned} \text{Min} \quad & (x_1 - 3)^2 + \left(x_2 - \frac{3}{2}\right)^2 \\ \text{s.a:} \quad & x_1^2 + x_2^2 \leq 4 \\ & x_1 - x_2 \leq 0 \\ & x_1, x_2 \geq 0 \end{aligned}$$

$$f(x_1, x_2) = (x_1 - 3)^2 + \left(x_2 - \frac{3}{2}\right)^2$$

$$\nabla f(x_1, x_2) = \begin{pmatrix} 2x_1 - 6 \\ 2x_2 - 3 \end{pmatrix}$$

$$H_f(x_1, x_2) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

$\Delta_1 > 0$ y $\Delta_2 > 0 \Rightarrow f$ es estrictamente convexa.

$$g_1(x_1, x_2) = x_1^2 + x_2^2 - 5$$

$$\nabla g_1(x_1, x_2) = \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix}$$

$$H_{g_1}(x_1, x_2) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

$\Delta_1 > 0$ y $\Delta_2 > 0 \Rightarrow g_1$ es estrictamente convexa.

$g_2(x_1, x_2)$ es lineal y por tanto, cóncava y convexa. Por lo que se verifican las condiciones necesarias y suficientes de Khun y Tucker y cualquier punto crítico es el único mínimo global.

$$\frac{\partial L}{\partial x_1} = 2x_1 - 6 + 2\lambda_1 x_1 + \lambda_2 - \lambda_3 = 0$$

$$\frac{\partial L}{\partial x_2} = 2x_2 - 3 + 2\lambda_1 x_1 - \lambda_2 - \lambda_4 = 0$$

$$\lambda_1(x_1^2 + x_2^2 - 4) = 0$$

$$\lambda_2(x_1 - x_2) = 0$$

$$\lambda_3 x_4 = 0$$

$$\lambda_4 x_2 = 0$$

$$x_1^2 + x_2^2 \leq 4$$

$$x_1 - x_2 \leq 0$$

$$x_1, x_2 \geq 0$$

- $(0, 0, 0, 0)$

$$\left. \begin{array}{l} 2x_1 - 6 = 0 \\ 2x_2 - 3 = 0 \end{array} \right\} \Rightarrow \left. \begin{array}{l} x_1 = 3 \\ x_2 = \frac{3}{2} \end{array} \right\} \Rightarrow \text{Infactible}$$

- $(\lambda_1, 0, 0, 0)$

$$\left. \begin{array}{l} 2x_1 - 6 + 2\lambda_1 x_1 = 0 \Rightarrow \lambda_1 = \frac{6-2x_1}{2x_1} = \frac{3}{x_1} - 1 \quad \text{si } x_1 \neq 0 \\ 2x_2 - 3 + 2\lambda_1 x_2 = 0 \Rightarrow \lambda_1 = \frac{3-2x_2}{2x_2} = \frac{3}{2x_2} - 1 \quad \text{si } x_2 \neq 0 \end{array} \right\} \Rightarrow \frac{3}{x_1} - 1 = \frac{3}{2x_2} - 1 \Rightarrow x_1 = 2x_2$$

$$\left. \begin{array}{l} x_1 = 2x_2 \\ x_1^2 + x_2^2 = 4 \end{array} \right\} \Rightarrow 4x_2^2 + x_2^2 = 4 \Rightarrow 5x_2^2 = 4 \Rightarrow x_2^2 = \frac{4}{5} \Rightarrow x_2 = \frac{2}{\sqrt{5}} = \frac{2\sqrt{5}}{5} \Rightarrow x_1 = \frac{4\sqrt{5}}{5}$$

Infactible.

- $(0, \lambda_2, 0, 0)$

$$\left. \begin{array}{l} 2x_1 - 6 + \lambda_2 = 0 \\ 2x_2 - 3 - \lambda_2 = 0 \end{array} \right\} \Rightarrow 2x_1 + 2x_2 = 9$$

$$\left. \begin{array}{l} 2x_1 + 2x_2 = 9 \\ x_1 = x_2 \end{array} \right\} \Rightarrow 4x_1 = 9 \Rightarrow x_1 = \frac{9}{4} \Rightarrow x_2 = \frac{9}{4}$$

Infactible.

- $(0, 0, \lambda_3, 0)$

$$\left. \begin{array}{l} 2x_1 - 6 - \lambda_3 = 0 \\ 2x_2 - 3 = 0 \\ x_1 = 0 \end{array} \right\} \Rightarrow \lambda_3 = -6 \quad \text{No vale.}$$

- $(0, 0, 0, \lambda_4)$

$$\left. \begin{array}{l} 2x_1 - 6 = 0 \\ 2x_2 - 3 - \lambda_4 = 0 \\ x_2 = 0 \end{array} \right\} \Rightarrow \lambda_4 = -3 \quad \text{No vale.}$$

- $(\lambda_1, \lambda_2, 0, 0)$

$$\left. \begin{array}{l} x_1^2 + x_2^2 = 4 \\ x_1 = x_2 \end{array} \right\} \Rightarrow x_1 = 4 \Rightarrow x_1^2 = 2 \Rightarrow x_1 = \sqrt{2} \Rightarrow x_2 = \sqrt{2}$$

$$\left. \begin{array}{l} 2x_1 - 6 + 2\lambda_1 x_1 + \lambda_2 = 0 \Rightarrow 2\sqrt{2} - 6 + 2\sqrt{2}\lambda_1 + \lambda_2 = 0 \\ 2x_2 - 3 + 2\lambda_1 x_2 - \lambda_2 = 0 \Rightarrow 2\sqrt{2} - 3 + 2\sqrt{2}\lambda_1 - \lambda_2 = 0 \end{array} \right\} \Rightarrow 4\sqrt{2} - 9 + 4\sqrt{2}\lambda_1 = 0$$

$$\Rightarrow \lambda_1 = \frac{9-4\sqrt{2}}{4\sqrt{2}} = \frac{9\sqrt{2}}{8} - 1 > 0$$

$$\lambda_2 = 2\sqrt{2} - 3 + 2\sqrt{2}\left(\frac{9\sqrt{2}}{8} - 1\right) = 2\sqrt{2} - 3 + \frac{36}{8} - 2\sqrt{2} = \frac{9}{2} - 3 = \frac{3}{2}$$

Punto crítico.

Se tiene un punto crítico, que por el teorema es el único mínimo $(\sqrt{2}, \sqrt{2})$ y $(\frac{9\sqrt{2}}{8} - 1, \frac{3}{2}, 0, 0)$ con un valor de la función objetivo $\frac{61}{4} - 9\sqrt{2} \approx 2,522$.

Gráficamente:

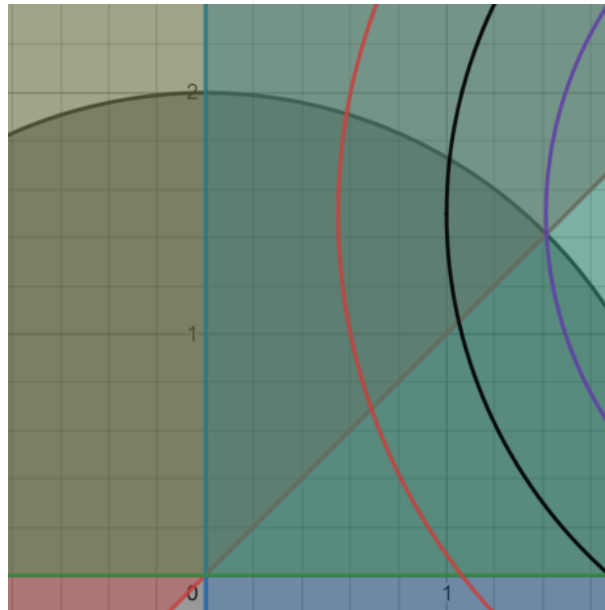


Figura 4.37: Ejercicio 5

Capítulo 5

Conclusiones

En la programación no lineal no existe un algoritmo único para cualquier problema no lineal, tal como ocurre en programación lineal, lo que hace más complicado su estudio.

Como hemos visto a lo largo del proyecto existen problemas que no pueden ser solucionados con la programación lineal debido a que posee elementos no lineales.

Primero empezamos viendo los conceptos básicos de la programación no lineal, sus restricciones, la función objetivo, conjuntos convexos y funciones convexas. Todos estos conceptos nos sirve de ayuda para comprende los métodos desarrollados.

La mayoría de los problemas de programación no lineal necesitan de aplicaciones o programas informáticos para alcanzar su solución completa. Durante este trabajo, se han desarrollado los diferentes tipos de algoritmos que tienen estos métodos de la programación no lineal, junto con ilustraciones gráficas. La ilustración gráfica nos ayudara a comprender más fácilmente la solución.

Los métodos multivariantes más prácticos son el método de Hooke y Jeeves junto con el método de Newton Raphson. De los dos métodos, el método de Newton Raphson es el más recomendado de usar, pues converge en menor número de iteraciones como vemos en el ejemplo seleccionado como modelo.

También en la programación no lineal con restricciones se han determinado condiciones que bajo ciertos requisitos se convierten en condiciones de primer orden o necesarias, e inclusive en condiciones necesarias y suficientes. Estas son las condiciones de Karush-Kuhn-Tucker, son el resultado analítico más importante en programación no lineal, por ello introducimos ejemplos con sus correspondientes soluciones y procedimientos.

Una de las causas por la que decidí realizar este trabajo fue la posibilidad de elaborar un tema completo sobre programación no lineal en castellano, ya que el inglés predomina en la mayoría de información que podemos encontrar acerca de este tema.

Bibliografía

- [1] ALLAIRE, JJ; XIE, YIHUI; MCPHERSON, JONATHAN; LURASCHI, JAVIER; USHEY, KEVIN; ATKINS, ARON; WICKHAM, HADLEY; CHENG, JOE; CHANG, WINSTON y IANNONE, RICHARD (2021). *rmarkdown: Dynamic Documents for R*. <https://CRAN.R-project.org/package=rmarkdown>. R package version 2.7.
- [2] BAZARAA, MOKHTAR S; SHERALI, HANIF D y SHETTY, CHITHARANJAN M (2013). *Nonlinear programming: theory and algorithms*. John Wiley & Sons.
- [3] BORCHERS, HANS W. (2021). *pracma: Practical Numerical Math Functions*. <https://CRAN.R-project.org/package=pracma>. R package version 2.3.3.
- [4] CANTÚ CUÉLLAR, RAMÓN (1996). *Programación no lineal*. Tesis doctoral, Universidad Autónoma de Nuevo León.
- [5] ESPINOSA-PAREDES, GILBERTO y RODRÍGUEZ, ALEJANDRO VÁZQUEZ (2016). *Aplicaciones de programación no lineal*. OmniaScience.
- [6] LOPEZ, JUAN FERNANDO; FERNÁNDEZ HENAO, SERGIO y MORALES, MARCELA MARÍA (2007). «Aplicación de la programación por metas en la distribución de servicios entre empresas operadoras del sistema de transporte masivo». *Scientia et technica*, **13**(37), pp. 339–343.
- [7] R CORE TEAM (2016). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
- [8] RSTUDIO TEAM (2015). *RStudio: Integrated Development Environment for R*. RStudio, Inc., Boston, MA. <http://www.rstudio.com/>.
- [9] WICKHAM, HADLEY (2020). *tidyr: Tidy Messy Data*. <https://CRAN.R-project.org/package=tidyr>. R package version 1.1.2.
- [10] WICKHAM, HADLEY; FRANÇOIS, ROMAIN; HENRY, LIONEL y MÜLLER, KIRILL (2021). *dplyr: A Grammar of Data Manipulation*. <https://CRAN.R-project.org/package=dplyr>. R package version 1.0.5.
- [11] XIE, YIHUI (2021). *knitr: A General-Purpose Package for Dynamic Report Generation in R*. <https://yihui.org/knitr/>. R package version 1.31.
- [12] ZHU, HAO (2021). *kableExtra: Construct Complex Table with kable and Pipe Syntax*. <https://CRAN.R-project.org/package=kableExtra>. R package version 1.3.4.