Universidad de Valladolid

Facultad de Ciencias Económicas y Empresariales

Grado en ADE

Optimización y sus aplicaciones económicas

Presentado por:

Eva María Bermejo Collado

Tutelado por:

Ramón Fernández Lechón

Valladolid, 1 de Abril del 2016

ÍNDICE DE CONTENIDO

ÍNDICE DE ILUSTRACIONES

RESUMEN

Este trabajo tiene como propósito establecer la relación existente entre el concepto matemático de optimización de funciones con y sin restricciones, y la asignación de recursos en el ámbito económico. Para resolver este tipo de problemas y que el cálculo resulte más fácil, se debe intentar como primer objetivo, que el conjunto donde de encontrar el óptimo, esto es el conjunto admisible sea un conjunto convexo y posteriormente que la función a optimizar sea una función convexa o cóncava, dependiendo del tipo de óptimo a encontrar.

En la vida real, nos encontramos con limitaciones a la hora de hacer una asignación de recursos. Estas limitaciones vienen determinadas mediante restricciones a través de funciones que denominamos funciones de restricción y para poder buscar la asignación eficiente será necesario utilizar el concepto de función lagrangiana que incluye la función a optimizar y la función o funciones de restricción.

Ante el problema de asignación de recursos tanto sin restricción como con restricción el objetivo es buscar la combinación óptima para la cual, primero se buscaran los posibles puntos estacionarios y después entre ellos el óptimo.

Palabras clave: Conjunto convexo, hiperplano, función convexa, funciones de restricción, punto estacionario, función lagrangiana.

Clasificación JEL: C020, C610, C690

ABSTRACT

This work aims to establish the relationships between the mathematical concept of optimization of functions with and without restrictions, and the allocation of resources in the economic sphere. Problem solving strategies on this type of problem would make calculation easier if it firstly tried to find the set where the optimum belongs, that set would be named the admissible, which should be a convex one. Consequently, the function to optimize might be either convex or concave, regarding the type of optimal that has been found.

In real live, there are limitations when it comes to allocation of resources. These limitations are determined by means of restrictions through functions which we have named restriction functions it would be necessary to use the concept of Lagrangian Function in order to find the efficient allocation that will include, for that matter, either the function or functions to optimize.

Regarding the allocation of resources issue, in both cases, unrestricted and restricted, our target would be to search for the optimum combination, which will guarantee to find the optimum point among the possible stationary points.

Keywords: Convex set, hyperplane, convex function, constraint functions, stationary point, lagrangian function.

JEL Classification: C020, C610, C690

1. INTRODUCCIÓN

La Teoría de Optimización Matemática es una de las ramas del saber que más ha influido en otras ciencias y en concreto en la Economía, ya que un problema fundamental de la ciencia económica consiste en distribuir recursos, normalmente escasos, entre distintos objetivos y debido a esa escasez de recursos es preciso realizar elecciones racionales que permitan alcanzar los objetivos propuestos dentro de las limitaciones que establece la escasez de recursos.

Por tanto, no es aventurado señalar que la Teoría de Optimización es un instrumento fundamental en el análisis económico, pues muchos problemas económicos consisten en la determinación del valor de unas magnitudes para que otra alcance su valor óptimo.

El desarrollo de la Optimización Matemática tuvo sus aportaciones más relevantes después de la Segunda Guerra mundial, aunque algunos de los resultados se conocían desde el siglo XVIII.

En este trabajo se pretende desarrollar los diferentes modelos que pueden presentarse dentro de la Optimización Matemática, solamente se tratarán algunos problemas de optimización estática, llamada así porque la variable tiempo no interviene en la formulación del problema, mientras que si el problema depende del tiempo sería un problema de optimización dinámica.

En general la teoría de optimización estática se conoce también como Programación Matemática y normalmente en un problema de Programación Matemática se trata de encontrar el óptimo (máximo o mínimo) de una función, usualmente de varias variables, cuando éstas se encuentran sometidas o no a un conjunto de restricciones de distintos tipos.

El propósito es describir los métodos de resolución de algunos de estos programas, en particular se estudiará los denominados problemas de programación clásica, tanto sin restricciones como con restricciones y aunque es lógico pensar que desde el punto de vista económico son más realistas los programas con restricciones, es importante señalar que existen razones que aconsejan el estudio de los programas irrestrictos pues, por ejemplo, en el campo de la econometría aparecen problemas de ajustes de datos mediante modelos de regresión que requieren la solución de programas sin restricciones. Además hay programas con restricciones de igualdad que pueden resolverse como uno sin restricciones reduciendo el número de variables.

También se abordará la interpretación que en términos económicos tienen las variables, que desde el punto de vista matemático son un recurso de cálculo, y que denominamos multiplicadores de Lagrange en los problemas con restricciones y que realmente consisten en un análisis de sensibilidad del programa.

En todo momento se hablará de programas diferenciables y, por tanto, podremos emplear las técnicas del cálculo diferencial y puesto que los métodos de resolución son más fáciles de aplicar si los programas verifican ciertas condiciones de convexidad, se dedicará un epígrafe del trabajo a estos resultados de convexidad.

2. CONVEXIDAD

2.1 INTRODUCIÓN

Los conjuntos convexos y las funciones cóncavas y convexas tienen una importancia relevante dentro de la programación matemática debido a que si en un programa se cumplen determinadas condiciones de convexidad, su resolución se simplifica y además obtenemos información adicional sobre los óptimos alcanzados que serán globales y en el campo económico éste tipo de óptimos son los que normalmente interesa obtener.

H. Minkowski fue uno de los pioneros en el estudio de la convexidad y publicó, a principios del siglo XX, diversos estudios en los que ya se vislumbraba la importancia de algunos de los resultados obtenidos sobre funciones y conjuntos convexos.

Este capítulo está dividido en dos partes con sus correspondientes epígrafes. La primera trata sobre los conjuntos convexos, su definición, sus propiedades y diferentes características que pueden presentarse en este tipo de conjuntos, todo ello desarrollando ejemplos fáciles de entender que nos permitan una mejor comprensión de los diferentes conceptos y propiedades. En la segunda parte nos centraremos en el estudio de las funciones cóncavas y convexas para pasar después de definir este tipo de funciones a enumerar alguna de sus propiedades más importantes así como caracterizarlas en el caso de que sean diferenciables.

2.2 CONJUNTOS CONVEXOS

Los conjuntos convexos son una forma concreta de conjuntos, con peculiaridades que permiten enlazar y concretar propiedades asociadas a la concavidad y la convexidad de las funciones. Comenzamos definiendo el concepto de combinación convexa

Definición: Diremos que $\bar{x} \in \mathbb{R}^n$ es una combinación lineal convexa de los puntos de $\mathbb{R}^n$, $\{\bar{x}_1, \bar{x}_2, \bar{x}_3, \dots, \bar{x}_n\}$ si existen $\alpha_1, \alpha_2, \dots, \alpha_n$ que verifique:

1. $\alpha_i \geq 0$
2. $\alpha_1 + \cdots + \alpha_n = 1$
3. $\bar{x} = \alpha_1 \bar{x}_1 + \alpha_2 \bar{x}_2 + \cdots + \alpha_n \bar{x}_n$

Luego, si un punto es combinación convexa de otros es que puede expresarse como combinación lineal de ellos con los coeficientes no negativos y sumando la unidad.

Si en $\mathbb{R}^2$ consideramos el subconjunto formado por tres puntos cualesquiera, por ejemplo, el conjunto A = $\{\bar{x} = (0,2),\ \bar{y} = (1,1),\ \bar{z} = (1,3)\}$, el conjunto de las combinaciones lineales convexas de los elementos A será:

$$\left\{(x_1, x_2) \in \mathbb{R}^2 / (x_1, x_2) = \alpha_1\ (0,2) + \alpha_2\ (1,1) + \alpha_3\ (1,3), \sum_{i=1}^{3} \alpha_i = 1, \alpha_i \geq 0, i = 1,2,3 \right\}$$

$$= \{(x_1, x_2) \in \mathbb{R}^2 / 0 \leq x_1 \leq 1\ , 2 - x_1 \leq x_2 \leq 2 + x_1\}$$

Gráficamente, este conjunto se recoge en la ilustración 1.

En particular, si en la definición anterior nos referimos solamente a dos puntos $\bar{x}_1$ y $\bar{x}_2$ el conjunto de las combinaciones lineales convexas es:

$$[\bar{x}_1, \bar{x}_2] = \{\bar{x} \in \mathbb{R}^2 /\ \bar{x} = \alpha\, \bar{x}_1 + (1 - \alpha)\, \bar{x}_2\ ,\ \ \alpha \in [0,1]\}$$

que se denomina segmento cerrado de extremos $\bar{x}_1$ y $\bar{x}_2$. Hemos de notar que en $\mathbb{R}^2$ el segmento cerrado de extremos $\bar{x}_1$ y $\bar{x}_2$ es el segmento de recta que une dichos puntos. En el caso de que $\alpha \in (0,1)$ tendríamos el segmento abierto.

Definición: Se dice que un conjunto C subconjunto de $\mathbb{R}^n$ es convexo si $\forall\, \bar{x}, \bar{y} \in C$ y $\lambda \in [0,1]$ se verifica $\lambda\bar{x} + (1-\lambda)\bar{y} \in C$.

Luego teniendo en cuenta el concepto antes comentado de segmento cerrado podemos decir que una definición equivalente a la anterior será: Un conjunto es convexo si para cualquier par de puntos del conjunto, el segmento cerrado por ellos generado está contenido en el conjunto.

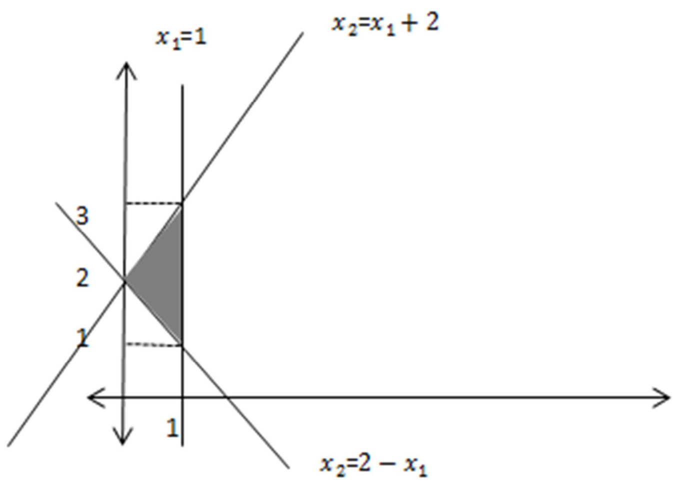

Ilustración 1: Conjunto de combinaciones lineales convexas

La ilustración 2 muestra dos conjuntos en un espacio $\mathbb{R}^2$, el conjunto A es un conjunto no convexo porque existen puntos del conjunto tales que el segmento que une dichos puntos no está contenido dentro del conjunto; sin embargo el conjunto B es un conjunto convexo porque para cualquier par de puntos del conjunto el segmento que une dichos puntos está contenido dentro del conjunto.

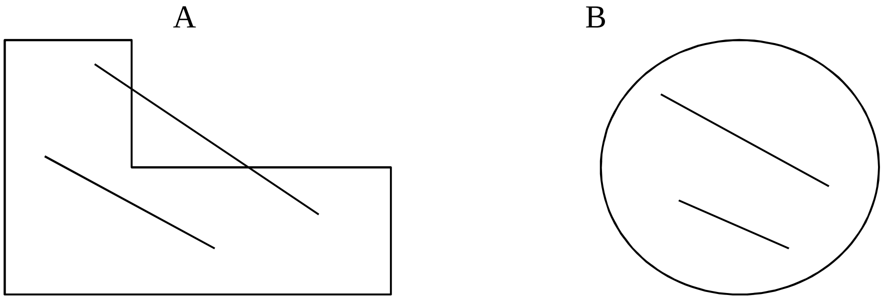

Ilustración 2: Conjunto no convexo y conjunto convexo

2.2.1 Propiedades

No pretendemos mencionar todas las propiedades de los conjuntos convexos, sino solamente algunas de las más usuales.

1. El conjunto vacío es convexo.
2. La intersección de los conjuntos convexos es un conjunto convexo.
3. Si dos conjuntos son convexos entonces el producto cartesiano y la suma de dichos conjuntos también son convexos.
4. El producto de un escalar por un conjunto convexo también es un conjuntos convexo.
5. La adherencia de un conjunto convexo es un conjunto convexo.
6. El interior de un conjunto convexo es un conjunto convexo.

2.2.2 Punto extremo, envolvente convexa e hiperplano

Los conjuntos convexos pueden tener puntos que satisfacen una propiedad particular, es lo que se conoce como puntos extremos.

*Definición: S*ea C un conjunto convexo, subconjunto de $\mathbb{R}^n$, se dice que $\bar{x} \in C$ es un punto extremo o vértice de C cuando no es posible expresarlo como combinación convexa de otros dos puntos distintos de C.

Los puntos extremos de un conjunto convexo no pueden ser puntos interiores de dicho conjunto, luego, si existen, han de ser puntos de la frontera del conjunto, pero obviamente no son puntos extremos todos los puntos frontera.

La ilustración 3 representa a la izquierda un conjunto convexo, todos los puntos comprendidos entre las dos rectas paralelas, que no tiene puntos extremos y a la derecha un conjunto convexo con un número finito de puntos extremos, los vértices del cuadrado.

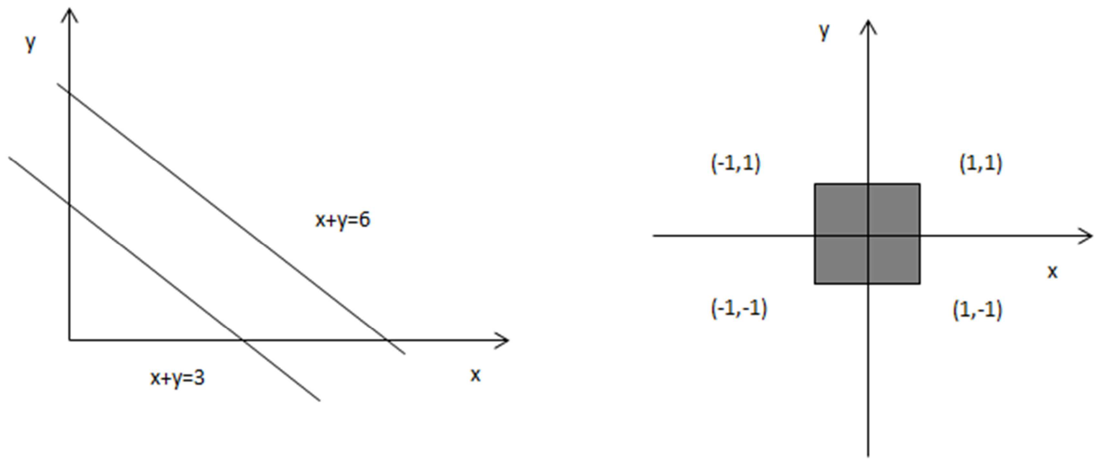

Ilustración 3: Conjunto convexo sin puntos extremos y con puntos extremos

*Definición: S*ea C un subconjunto de $\mathbb{R}^n$, llamamos envolvente convexa de C al menor conjunto convexo que lo contiene.

Por consiguiente si el conjunto ya es convexo, su envolvente convexa será el mismo y si el conjunto no es convexo, su envolvente convexa coincidirá con la intersección de todos los conjuntos convexos que contienen al conjunto o equivalentemente los elementos de la envolvente convexa son todos aquellos que pueden obtenerse como combinación convexa de puntos del conjunto.

Dado un conjunto de puntos de $\mathbb{R}^n$, P = $\{\bar{x}_1, \bar{x}_2, \bar{x}_3, \dots, \bar{x}_n\}$ (con n mayor o igual que dos), su envolvente convexa, también llamado <u>cierre convexo</u> es el menor convexo que lo contiene y está formado por todas las combinaciones convexas de esos *n* puntos. Este conjunto también se denomina poliedro convexo generado por dichos puntos y obviamente es un conjunto cerrado, convexo y tiene un número finito de puntos extremos que son precisamente algunos de los puntos que lo generan, aunque no tienen que ser todos ellos.

En la ilustración 4 se muestra un conjunto de puntos en el plano y su envolvente convexa, es un conjunto convexo, cerrado y tiene un número finito de puntos extremos, en este caso ocho, que son precisamente los vértices del poliedro generado.

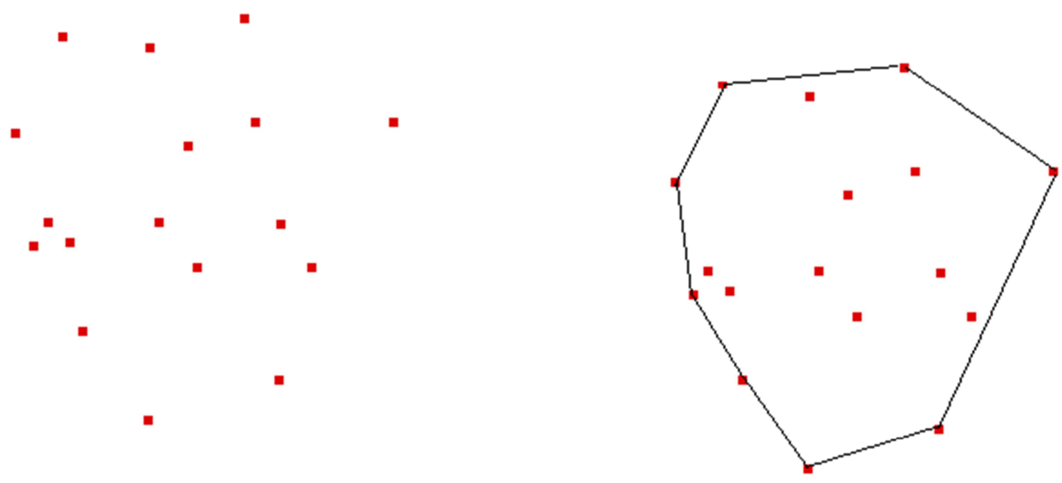

Ilustración 4: Conjunto de puntos y envolvente convexa

En la programación matemática aparecen con mucha frecuencia conjuntos que vienen definidos a través de un hiperplano y los teoremas de separación resultan necesarios para demostrar determinados resultados.

Definición: Sea $\bar{y} \in \mathbb{R}^n$ un vector no nulo y α un número real se denomina hiperplano al conjunto $H = \{\bar{x} \in \mathbb{R}^n / \bar{y}^t \bar{x} = \alpha\}$.

Un hiperplano es un conjunto convexo y divide al espacio en otros dos conjuntos $H^+ = \{\bar{x} \in \mathbb{R}^n / \bar{y}^t \bar{x} \geq \alpha\}$ y $H^- = \{\bar{x} \in \mathbb{R}^n / \bar{y}^t \bar{x} \leq \alpha\}$ que también son convexos y se denominan semiespacios cerrados; si en los conjuntos anteriores la desigualdad es en sentido estricto tendríamos definidos los semiespacios abiertos.

Los resultados relativos a los teoremas de separación pueden resumirse de la siguiente forma: si tenemos dos conjuntos convexos disjuntos, se garantiza la existencia de un hiperplano, denominado hiperplano de separación, de modo que todos los elementos de un conjunto convexo están en uno de los semiespacios cerrados que determina dicho hiperplano y el otro en el otro semiespacio; si los conjuntos convexos tienen en común solamente puntos frontera el hiperplano de separación recibe ahora el nombre de hiperplano soporte.

El teorema del hiperplano de separación demuestra que los conjuntos convexos no tienen que ser cerrados, y en el caso particular de tener un solo conjunto convexo, si tomamos un punto frontera del conjunto existirá un hiperplano, al que se denomina hiperplano soporte que pasa por dicho punto y todos los elementos del conjunto estarán en solo uno de los semiespacios que el hiperplano determina.

La ilustración 5 nos muestra un hiperplano de separación y un hiperplano soporte en $\mathbb{R}^2$. En la figura de la izquierda se observa que el punto (0,0) no pertenece al conjunto cerrado y convexo de modo que hay un hiperplano de separación. El punto (0,0) estará en uno de los semiespacios del hiperplano y los elementos del conjunto en el otro como observamos en el dibujo. En la figura de la derecha se observa el círculo con centro en el punto (0,0) y radio 1, es un conjunto convexo y abierto; el punto (1,1) pertenece a la frontera del conjunto y el hiperplano soporte en (1,1) es el indicado en la figura, estando todos los puntos del conjunto convexo en uno de los semiespacios abiertos que determina el hiperplano.

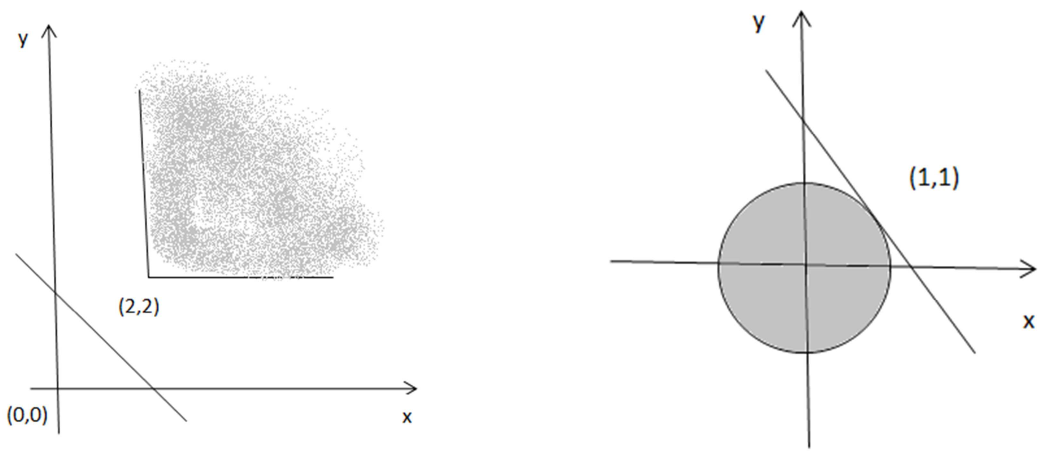

Ilustración 5: Hiperplano de separación e hiperplano de soporte

2.3 FUNCIONES CÓNCAVAS Y CONVEXAS

Para que el estudio de la convexidad quede completo es preciso extender el estudio a las funciones y como las definiciones van a venir dadas a través de desigualdades con el consiguiente problema de la dificultad de trabajas con ellas, vamos a establecer condiciones necesarias y suficientes para caracterizar a dichas funciones en el caso de que sean diferenciables.

Las definiciones de funciones cóncavas y funciones convexas guardan cierta similitud con las definiciones dadas anteriormente y presentan unas propiedades semejantes. Las definiciones pueden establecerse considerando la combinación convexa e dos o más elementos de $\mathbb{R}^n$.

*Definición: S*ea C un subconjunto de $\mathbb{R}^n$ convexo y no vacío y sea $f: C \to \mathbb{R}$

1. La función f es convexa si satisface:

$$f(\alpha \bar{x}_1 + (1-\alpha)\bar{x}_2) \leq \alpha f(\bar{x}_1) + (1-\alpha) f(\bar{x}_2)$$

para todo $\bar{x}_1$, $\bar{x}_2 \in C$ y todo $\alpha \in [0,1]$.

2. La función f es cóncava si satisface:

$$f(\alpha \bar{x}_1 + (1-\alpha)\bar{x}_2) \geq \alpha f(\bar{x}_1) + (1-\alpha) f(\bar{x}_2)$$

para todo $\bar{x}_1$, $\bar{x}_2 \in C$ y todo $\alpha \in [0,1]$.

Si las desigualdades se verifican en sentido estricto siendo $\bar{x}_1 \neq \bar{x}_2$ y $\alpha \in (0,1)$ entonces decimos que f es estrictamente convexa o cóncava respectivamente.

De la definición se desprende que una función es convexa si el segmento que une dos puntos cualesquiera de la gráfica de la función está por encima de la gráfica de la función, mientras que la función será cóncava si el segmento que une dos puntos cualesquiera de la gráfica de la función está por debajo de la gráfica de la función. Esta última idea puede utilizarse para caracterizar las funciones que se muestran en la ilustración 6 donde la función de la izquierda es convexa y además estrictamente convexa, mientras que la de la derecha es cóncava y también estrictamente cóncava.

Se destaca que en las definiciones he dicho que el conjunto donde están definidas las funciones debe ser un conjunto convexo, pues en otro caso no tendré garantizado que la combinación convexa de dos elementos del conjunto sea del conjunto y, por consiguiente, no podré conocer el valor que la función alcanza en dicho punto.

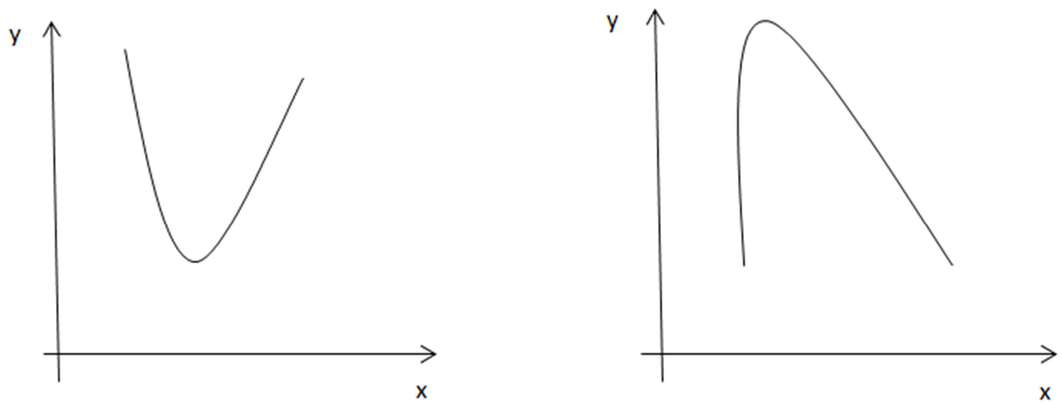

Ilustración 6: Función cóncava y función convexa

Observando las definiciones se puede confirmar que hay funciones que son convexas y cóncavas conjuntamente; funciones que no son ni cóncavas ni convexas, funciones cóncavas que no son convexas, funciones convexas que no son cóncavas, funciones cóncavas que no son estrictamente cóncavas y funciones convexas que no son estrictamente convexas.

2.3.1 Propiedades

Existen un gran número de propiedades que verifican las funciones cóncavas y convexas pero solamente se mencionan alguna de ellas.

1. Si f es una función convexa, entonces $-f$ es una función cóncava y recíprocamente.
2. Si f es una función estrictamente convexa, entonces $-f$ es una función estrictamente cóncava y recíprocamente.
3. Si *f* es una función convexa (o cóncava) y $\theta > 0$, la función θf es convexa (o cóncava) y recíprocamente.
4. Cualquier combinación lineal con coeficientes no negativos de funciones convexas (o cóncavas) es una función convexa (cóncava).
5. Si *f* es una función estrictamente convexa (o estrictamente cóncava) y $\theta > 0$, la función θf es estrcitamente convexa (o estrictamente cóncava) y recíprocamente.

6. Si f es una función convexa (o cóncava) definida en un conjunto convexo y abierto, entonces es continua en su conjunto de definición.
7. Si f es una función convexa definida en un conjunto C y α un número real, el conjunto $\{\bar{x} \in C/f(\bar{x}) \leq \alpha\}$ es un conjunto convexo.
8. Si f es una función cóncava definida en un conjunto C y α un número real, el conjunto $\{\bar{x} \in C/f(\bar{x}) \geq \alpha\}$ es un conjunto convexo.
9. Sea f es una función definida en un subconjunto convexo C de $\mathbb{R}^n$. La función f es convexa si y sólo si el conjunto

$$\{(\bar{x}, y) \in \mathbb{R}^{n+1}/\, \bar{x} \in C, y \in \mathbb{R}, f(\bar{x}) \leq y\}$$

es un conjunto convexo.
10. Sea f es una función definida en un subconjunto convexo C de $\mathbb{R}^n$. La función f es cóncava si y sólo si el conjunto

$$\{(\bar{x}, y) \in \mathbb{R}^{n+1}/\, \bar{x} \in C, y \in \mathbb{R}, f(\bar{x}) \geq y\}$$

es un conjunto convexo.
11. Si f es una función convexa que alcanza un mínimo en el punto $\bar{x}^*$, la función $-f$ alcanza un máximo en $\bar{x}^*$.
12. Si f es una función convexa (o cóncava) que alcanza un mínimo local (o máximo local) en el punto $\bar{x}^*$, dicho mínimo (o máximo) es global.
13. Si f es una función estrictamente convexa (o estrictamente cóncava) que alcanza un mínimo (o máximo) en el punto $\bar{x}^*$, dicho mínimo (o máximo) es único.
14. Si f es una función convexa (o cóncava) que alcanza mínimo (o máximo) en dos puntos distintos $\bar{x}_1$, $\bar{x}_2$, entonces también alcanza un mínimo (máximo) en cualquier punto que sea combinación convexa de $\bar{x}_1$ y $\bar{x}_2$.
15. El máximo (mínimo) de una función convexa (o cóncava) no constante debe alcanzarse en los puntos frontera de su conjunto de definición.

16. Si C es un subconjunto no vacío, convexo y compacto de $\mathbb{R}^n$ con un número finito de puntos extremos y f una función convexa (o cóncava) definida en él, entonces f posee un máximo (mínimo) global en C y se alcanza en uno de sus puntos extremos.

2.3.2 Caracterización con cálculo diferencial.

Las funciones cóncavas y convexas pueden no ser diferenciables en su dominio de definición y para poderlas caracterizar tendríamos que recurrir a alguna de las condiciones anteriormente mencionadas, resultando en algunos casos bastante complicado su caracterización; ahora bien si se puede utilizar el cálculo diferencial existen condiciones que permiten determinar de una manera sencilla si una función es cóncava o convexa sin necesidad de operar con la definición que en bastantes ocasiones resulta poco práctica. A continuación voy a comentar las condiciones de primer orden, pues la información la proporcionan las derivadas primeras y las de segundo orden, ya que la información viene dada a través de las derivadas segundas.

En todos los casos es necesario que el dominio de definición de la función sea abierto, convexo y no vacío.

Describo en primer lugar dichas condiciones para el caso de funciones de una variable.

Teorema: Sea f una función de clase C^1 definida en un intervalo abierto y no vacío I de $\mathbb{R}$, la función f es convexa si y sólo si

$$f(x_1) - f(x_2) \geq (x_1 - x_2) f'(x_2), \ \ \forall\, x_1, x_2 \in I$$

Un teorema análogo se tendría para funciones cóncavas con la desigualdad al revés.

Teorema: Sea f una función de clase C^2 definida en un intervalo abierto y no vacío I de $\mathbb{R}$, la función f es convexa si y sólo si $f''(x) \geq 0, \ \forall\, x \in I$.

De nuevo el teorema análogo para funciones cóncavas sería con la desigualdad al revés.

En el caso general de funciones de varias variables, los dos teoremas anteriores se enunciarán de la siguiente forma:

Teorema: Sea f una función de clase C^1 definida en un subconjunto A convexo, abierto y no vacío de $\mathbb{R}^n$, la función f es convexa si y sólo si se satisface alguna de las condiciones siguientes

$$f(\bar{x}_1) - f(\bar{x}_2) \geq (\bar{x}_1 - \bar{x}_2)^t \, \nabla f(\bar{x}_2), \;\; \forall \, \bar{x}_1, \bar{x}_2 \in A$$

$$(\bar{x}_1 - \bar{x}_2)^t \, (\nabla f(\bar{x}_1) - \nabla f(\bar{x}_2) \geq 0, \qquad \forall \, \bar{x}_1, \bar{x}_2 \in A$$

El teorema para funciones cóncavas sería análogo al anterior pero con las desigualdades al revés.

Teorema: Sea f una función de clase C^2 definida en un subconjunto A convexo, abierto y no vacío de $\mathbb{R}^n$, la función f es convexa si y sólo si $\forall \, \bar{x} \in A$ la forma cuadrática $\bar{h}^t H f(\bar{x}) \bar{h}$ es semidefinida positiva o definida positiva.

Podría enunciar un teorema similar para funciones cóncavas pero ahora la forma cuadrática ha de ser semidefinida negativa o definida negativa.

En el último teorema $Hf(\bar{x})$ es la matriz hessiana de la función.

En todos los teoremas anteriores si las desigualdades se verifican en sentido estricto tendría caracterizadas con condición necesaria y suficiente las funciones estrictamente convexas o bien estrictamente cóncavas. También utilizando la matriz hessiana de la función se puede establecer una condición suficiente para saber si una función es convexa (o cóncava) en sentido estricto del siguiente modo.

Teorema: Sea f una función de clase C^2 definida en un subconjunto A convexo, abierto y no vacío de $\mathbb{R}^n$. Si la forma cuadrática $\bar{h}^t H f(\bar{x}) \bar{h}$ es definida positiva la función es estrictamente convexa, mientras que si es definida negativa la función es estrictamente cóncava.

A continuación detallo algunos ejemplos para caracterizar la concavidad o convexidad de funciones aplicando los resultados anteriores.

<u>Ejemplo 1:</u> Consideremos la función real de variable real $f(x) = x^2$
Si se representa la función se observa que es una parábola con vértice en el origen, por lo que gráficamente podemos determinar su convexidad pues en este caso se verifica que el segmento que une dos puntos cualesquiera de la gráfica de la función está por encima de la gráfica de la función.

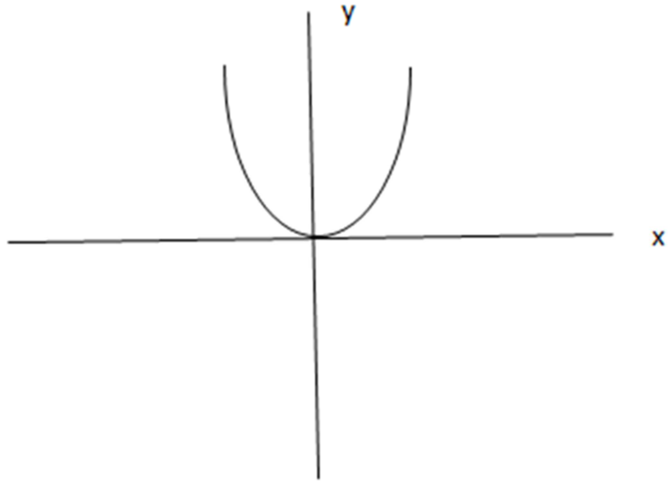

Ilustración 7: Función convexa

Luego la función es convexa, pero puesto que nunca coincide el segmento que une dos puntos distintos de la gráfica de la función con la función tenemos que se verifica

$$f(\alpha x_1 + (1-\alpha)x_2) < \alpha f(x_1) + (1-\alpha)f(x_2)$$

para todo x_1, $x_2 \in \mathbb{R}$ y todo $\alpha \in (0,1)$, con $x_1 \neq x_2$ y la función es estrictamente convexa.

También se puede llegar al mismo resultado aplicando directamente la definición, así pues hemos de ver si para todo x_1, $x_2 \in \mathbb{R}$ y $\alpha \in [0,1]$ se satisface:

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

esto es,

$$0 \leq \alpha f(x_1) + (1-\alpha)f(x_2) - f(\alpha x_1 + (1-\alpha)x_2)$$

sustituimos en el lado derecho de la desigualdad y operamos:

$$\alpha x_1^2 + (1-\alpha)x_2^2 - (\alpha x_1 + (1-\alpha)x_2)^2 =$$
$$\alpha x_1^2 + (1-\alpha)x_2^2 - \alpha^2 x_1^2 - (1-\alpha)^2 x_2^2 - 2\alpha(1-\alpha)x_1x_2 =$$
$$x_1^2(\alpha - \alpha^2) + x_2^2[(1-\alpha) - (1-\alpha)^2] - 2\alpha(1-\alpha)x_1x_2 =$$
$$\alpha(1-\alpha)x_1^2 + (1-\alpha)\alpha x_2^2 - 2\alpha(1-\alpha)x_1x_2 =$$
$$\alpha(1-\alpha)[x_1^2 + x_2^2 - 2x_1x_2] = \alpha(1-\alpha)(x_1 - x_2)^2 \geq 0$$

y se satisface la desigualdad por lo que la función es convexa, pero si $x_1 \neq x_2$ y $\alpha \in (0{,}1)$ los tres factores de la última desigualdad son positivos y, por tanto, $\alpha(1-\alpha)(x_1-x_2)^2 > 0$ y la función es estrictamente convexa.

También puesto que la función es de clase C^1, si $x_1 > x_2$ tenemos

$$f(x_1)-f(x_2)=x_1^2-x_2^2=(x_1-x_2)(x_1+x_2)>(x_1-x_2)2x_2=(x_1-x_2)f'(x_2)$$

concluyendo que la función es estrictamente convexa y también puesto que la función es clase C^2 podemos calcular su derivada segunda que en este caso tenemos $f'(x)=2x$ y $f''(x)=2$ y como la segunda derivada es mayor que cero, la función $f(x)=x^2$ es estrictamente convexa.

Queda comprobado de cuatro maneras diferentes que la función es estrictamente convexa, tanto gráficamente como analíticamente.

<u>Ejemplo 2:</u> Vamos a demostrar que el siguiente conjunto es convexo.

$$\{(x_1,x_2,x_3)\in\mathbb{R}^3/(3x_1^2+4x_2)\leq x_3\}$$

Observamos que el conjunto considerado también puede escribirse de la forma $\{(x_1,x_2,x_3)\in\mathbb{R}^3/f(x_1,x_2)\leq x_3\}$ donde $f(x_1,x_2)=(3x_1^2+4x_2)$.

Si demostramos que f es convexa, se probará que el conjunto es convexo (propiedad 9 de las funciones convexas citada anteriormente).

Para ver que f es convexa, empleamos la caracterización de las funciones de varias variables de clase C^2 cuyo teorema cita textualmente: Sea f una función de clase C^2 definida en un subconjunto A convexo, abierto y no vacío de $\mathbb{R}^n$, la función f es convexa si y sólo si $\forall\, \bar{x}\in A$ la forma cuadrática $\bar{h}^t Hf(\bar{x})\bar{h}$ es semidefinida positiva o definida positiva.

Por tanto, la matriz Hessiana en este caso quedaría:

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

que es semidefinida positiva y la forma cuadrática es semidefinida positiva y la función es convexa que era nuestro objetivo.

Otra forma de probar que el conjunto es convexo es viendo que la función

$$g(x_1,x_2,x_3)=3x_1^2+4x_2-x_3$$

es convexa en $\mathbb{R}^3$, ya que entonces el conjunto objeto de estudio es

$$\{(x_1, x_2, x_3) \in \mathbb{R}^3 / g(x_1, x_2, x_3) \leq 0\}$$

y por la propiedad 7 de las funciones convexas, se sabe que es un conjunto convexo.

En este caso, utilizando el mismo teorema que antes, hemos de estudiar la matriz hessiana

$$Hg(x_1, x_2, x_3) = \begin{pmatrix} 6 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

<u>Ejemplo 3:</u> Analicemos la concavidad o convexidad de la siguiente función:

$$f(x_1, x_2) = 5x_1^3 + 2x_2$$

De nuevo es una función de dos variables de clase C^2 y aplicamos el teorema que nos caracteriza con condición necesaria y suficiente su concavidad o convexidad.

El vector gradiente es $\nabla f(x_1, x_2) = (15x_1^2\ ,\ 2)^t$ y su matriz hessiana

$$Hf(x_1, x_2) = \begin{pmatrix} 30x_1 & 0 \\ 0 & 0 \end{pmatrix}$$

Dependiendo del signo de x_1 la forma cuadrática asociada a la matriz será una u otra forma:

- Si $x_1 > 0$ los autovalores son mayores que cero por tanto la forma cuadrática será semidefinida positiva.
- Si $x_1 < 0$ los autovalores son menores que cero por tanto la forma cuadrática será semidefinida negativa.
- Si $x_1 = 0$ la función a estudiar será $f(x_1, x_2) = 2x_2$ que es lineal.

De acuerdo con los resultados, la función no es ni cóncava ni convexa, puesto que la forma cuadrática asociada a la matriz Hessiana no mantiene siempre el mismo signo en todo el dominio de definición de la función.

Sin embargo, como cita el teorema analizado anteriormente si consideramos la función en la región $\{(x_1, x_2) \in \mathbb{R}^2 / x_1 \geq 0\}$ será convexa para $x_1 > 0$ la función lo es y para $x_1 = 0$ la función es lineal y por lo tanto cóncava y convexa. En la región $\{(x_1, x_2) \in \mathbb{R}^2 / x_1 \leq 0\}$ por un razonamiento análogo la función es cóncava.

3. OPTIMIZACIÓN SIN RESTRICCIONES

3.1 INTRODUCIÓN

En todo lo que sigue se va a suponer que los programas matemáticos que se formulan son programas diferenciables, esto es, todas las funciones que intervienen son funciones diferenciables y, por consiguiente, se podrán aplicar los resultados del cálculo diferencial.

Un problema de optimización sin restricciones también conocido como problema de programación clásica irrestricto o programación clásica libre consiste en encontrar los valores de las variables $\bar{x} \in D$ que maximicen o minimicen una función $f: \mathrm{D} \to \mathbb{R}$, donde D es un subconjunto abierto de $\mathbb{R}^n$.

Estos programas irrestrictos son los más elementales y fáciles de resolver y a veces es necesario resolver esta clase de programas para obtener las soluciones de otros programas más complejos. Este tipo de programas aparecen en economía cuando se intentan obtener las acciones que debe tomar un decisor que pretende lograr un determinado objetivo, como maximizar los beneficios de una empresa, minimizar los costes de producción, cuando se conocen sus respectivas funciones y no existen restricciones sobre las variables que intervienen en la formulación de esas funciones objetivo.

Así pues un problema de programación clásica sin restricciones se puede formular en términos matemáticos de la siguiente forma:

$$opt\ f(\bar{x}), \qquad s.a\colon\ \bar{x} \in D$$

donde D es el dominio de definición de la función.

A continuación voy a establecer unas condiciones necesarias de óptimo local, dichas condiciones son de primer orden; después estableceré unas condiciones de segundo orden que en un caso serán condiciones necesarias, mientras que en otro serán suficientes. También se explicará que en el caso de que el programa sea convexo, las condiciones necesarias de primer orden también son suficientes y además el óptimo alcanzado es global.

3.2 CONDICIONES DE PRIMER ORDEN

La proposición que se enuncian a continuación deben ser verificadas por todos los óptimos locales de los programas sin restricciones, es por tanto una condición necesaria de óptimo local, pero también veremos que bajo determinadas condiciones de convexidad es suficiente y tendremos así caracterizado un óptimo global.

Proposición: Sea $f\colon \mathrm{D} \to \mathbb{R}$ una función diferenciable, con D subconjunto abierto de $\mathbb{R}^n$, si $\bar{x}^* \in D$ es óptimo local del programa $opt\, f(\bar{x}),\ s.a\colon\ \bar{x} \in D$, entonces $\nabla f(\bar{x}^*) = \bar{0}$.

Los puntos $\bar{x} \in D$ que verifican las condiciones necesarias de óptimo se denominan puntos críticos o puntos estacionarios de la función objetivo. El conjunto de puntos críticos de una función puede ser el vacío, en este caso la función no tendrá óptimo; el conjunto de puntos críticos puede estar formado por un número finito de puntos y aun así la función puede carecer de óptimos, pero si los tiene deben ser algunos de ellos.

Por consiguiente, en el conjunto de puntos críticos de una función se busca encontrar los óptimos locales, que podrán ser globales, pero hay otros puntos que no son óptimos. A estos últimos puntos se les denomina puntos de ensilladura o puntos de silla de la función objetivo.

Si al problema se le añade la condición de que el programa sea convexo, que en este caso ha de verificar que el conjunto de definición de la función sea un conjunto convexo y que la función a optimizar sea convexa si tratamos de minimizar o sea cóncava si nuestro objetivo es maximizar, las condiciones de primer orden establecidas anteriormente son también suficientes y por una de las propiedades antes mencionadas de las funciones convexas o cóncavas, el óptimo alcanzado es global. Luego se puede enunciar la siguiente proposición:

Proposición: Sea $f\colon \mathrm{D} \to \mathbb{R}$ una función diferenciable y convexa (cóncava), con D subconjunto abierto de $\mathbb{R}^n$ convexo y no vacío.El punto $\bar{x}^* \in D$ es punto crítico o punto estacionario de f si y sólo si es mínimo (máximo) global del programa matemático $opt\, f(\bar{x}),\ s.a\colon\ \bar{x} \in D$.

Ejemplo 4: La función $f(x,y) = 2x^3 + y^3 - 6x$ admite dos puntos críticos (-1,0) y (1,0) ya que son los únicos puntos que satisfacen el sistema de ecuaciones.

$$f'_x(x,y) = 6x^2 - 6 = 0$$

$$f'_y(x,y) = 3y^2 = 0$$

Sin embargo, no se sabe si alguno de ellos es óptimo del problema y en caso de serlo de qué tipo, pues la función no es ni cóncava ni convexa y no se puede aplicar la última proposición enunciada.

Ejemplo 5: Dada la función $f(x,y) = 2x^2 + 3x + 5y^2 - 2y$ obtenemos en primer lugar los puntos críticos calculando las parciales de la función, esto es el vector gradiente e igualándolo a cero para resolver el sistema.

$$\nabla f(x,y) = (4x + 3, 10y - 2)^t$$

Igualando a cero obtenemos

$$4x + 3 = 0; \ 10y - 2 = 0$$

Al resolver ambas ecuaciones obtenemos que $x = -\frac{3}{4}, y = \frac{1}{5}$, que es un único punto crítico, pero en este caso la matriz hessiana de la función es:

$$Hf(\mathrm{x,y}) = \begin{pmatrix} 4 & 0 \\ 0 & 10 \end{pmatrix}$$

que es definida positiva, luego la función $f(x,y)$ es convexa y además estrictamente convexa y el punto crítico obtenido es único mínimo global de la función, puesto que las condiciones necesarias de primer orden son también suficientes.

3.3 CONDICIONES DE SEGUNDO ORDEN

Si un programa clásico sin restricciones no es convexo, al aplicar las condiciones necesarias de primer orden los puntos que encontramos sólo serán candidatos a óptimo local, para garantizar que realmente un punto crítico de los obtenidos es óptimo hemos de realizar un estudio de la función en un entorno del punto; sin embargo si es posible evaluar la matriz hessiana de la función en cada uno de los puntos críticos podemos indicar si es óptimo o no y además de que tipo.

A continuación voy a enunciar dos proposiciones, la primera nos establecerá unas condiciones necesarias de segundo orden mientras que la segunda nos proporcionará unas condiciones suficientes también de segundo orden.

Proposición: Sea f de clase C^2 en D abierto y no vacío. Si $\bar{x}^* \in D$ es mínimo (máximo) local del programa $opt\, f(\bar{x}),\ s.a\colon\ \bar{x} \in D$, entonces $\bar{x}^*$ es punto crítico de f y la forma cuadrática $\bar{h}^t Hf(\bar{x}^*)\bar{h}$ es definida positiva o semidefinida positiva (definida negativa o semidefinida negativa).

Precisamente la condición suficiente falla en el caso de que la forma cuadrática asociada a la matriz hessiana de la función sea semidefinida. Por eso a los puntos que verificando las condiciones de la proposición anterior la matriz hessiana en ellos es semidefinida se les denomina cuasi-optimos, sería cuasi-mínimo si la matriz hessiana es semidefinida positiva y cuasí-máximo si es semidefinida negativa, lo único que se puede asegurar es que no es máximo en el primer caso y que no es mínimo en el segundo caso.

Luego en el caso de que la matriz hessiana sea semidefinida al no poder afirmar que tipo de óptimo es, se tiene que realizar un análisis de la función en un entorno del punto o ver si existe un entorno del punto donde la forma cuadrática asociada a la matriz hessiana mantiene el mismo carácter, como indicamos en la siguiente proposición.

Proposición: Sea el programa $opt\, f(\bar{x}),\ s.a\colon\ \bar{x} \in D$, donde f de clase C^2 en D abierto y no vacío y $\bar{x}^* \in D$ un punto crítico de la función.

1. Si la forma cuadrática $\bar{h}^t Hf(\bar{x}^*)\bar{h}$ es definida positiva (definida negativa) el punto $\bar{x}^*$ es un mínimo (máximo) local en sentido estricto.
2. Si la forma cuadrática $\bar{h}^t Hf(\bar{x}^*)\bar{h}$ es indefinida el punto $\bar{x}^*$ es un punto de silla de la función.
3. Si la forma cuadrática $\bar{h}^t Hf(\bar{x}^*)\bar{h}$ es semidefinida positiva (semidefinida negativa) y existe un entorno tal que las formas cuadráticas $\bar{h}^t Hf(\bar{x})\bar{h}$ para todo punto del entorno y de D son semidefinidas positivas o definidas positivas (semidefinidas negativas o definidas negativas), el punto $\bar{x}^*$ es mínimo (máximo) del programa irrestricto.

Ejemplo 6: Tratamos de resolver el siguiente problema:

$$opt(x-2)^2+(y-3)^2$$

Aplicamos las condiciones necesarias de primer orden y tenemos:

$$\nabla f(x,y)=[2(x-2),2(y-3)]=(0,0)$$

Al resolver el sistema obtenemos un único punto crítico $x=2, y=3$ es por tanto un posible óptimo pero para caracterizarle aplicamos las condiciones suficientes de segundo orden

Hemos de calcular la matriz hesiana de la función y valorarla en el punto obtenido

$$Hf\ (x,y)=\begin{pmatrix}2 & 0\\ 0 & 2\end{pmatrix}, \qquad Hf\ (2,3)=\begin{pmatrix}2 & 0\\ 0 & 2\end{pmatrix}$$

Por tanto, Hf (2,3) corresponde a una forma cuadrática definida positiva y el punto crítico es un mínimo local estricto.

Notemos que tal y como se ha resuelto el ejemplo, sólo podemos afirmar que es un mínimo local, hemos perdido la información que podría proporcionarnos las condiciones de primer orden si nuestro programa fuese convexo y en este caso como sí es convexa la función objetivo, podemos afirmar que las condiciones necesarias de primer orden son también suficientes y el punto obtenido es un mínimo global.

Ejemplo 7: Una empresa monopolista vende sus productos en Alemania y Suecia. Sean q la cantidad total producida y q_1 , q_2 las cantidades vendidas en Alemania y Suecia respectivamente.

Suponemos que todo lo que se produce se vende y que su función de costes es

$$C(q)=2q^2+10q+120$$

y que las demandas de producción a las que se enfrenta Alemania y Suecia vienen determinadas por las funciones

$$q_1=10-1/5p_1$$
$$q_2=15-1/2p_2$$

donde p_1 y p_2 son los precios de cada uno de los países. Queremos determinar las cantidades que ha de vender en cada país para obtener el beneficio máximo.

Como la empresa todo lo que produce se vende, la cantidad total producida va a ser la suma de las cantidades vendidas en cada país, es decir, $\mathrm{q} = \mathrm{q}_1 + \mathrm{q}_2$.

Puesto que pretendemos maximizar el beneficio, hemos de determinar la función de beneficios que serán los ingresos obtenidos menos los costes. Calculamos los ingresos y los costes, ambas en función de las cantidades vendidas a cada país.

$$\mathrm{I}_1 = \mathrm{p}_1\mathrm{q}_1 = (50 - 5\mathrm{q}_1)\mathrm{q}_1 = 50\mathrm{q}_1 - 5\mathrm{q}_1^2$$

$$\mathrm{I}_2 = \mathrm{p}_2\mathrm{q}_2 = (30 - 2\mathrm{q}_2)\mathrm{q}_2 = 30\mathrm{q}_2 - 2\mathrm{q}_2^2$$

Por tanto, la función de ingresos de la empresa es I_1+I_2 y la función de costes es conocida y sustituimos q por $q_1 + q_2$. Luego, la función de beneficios será:

$$\begin{aligned} B(q_1, q_2) &= I_1 + I_2 - C(q) = \\ &= 50q_1 - 5q_1^2 + 30q_2 - 2q_2^2 - 2(q_1 + q_2)^2 - 10(q_1 + q_2) - 120 = \\ &= 40q_1 - 7q_1^2 + 20q_2 - 4q_2^2 - 4q_1q_2 - 120 \end{aligned}$$

Aplicamos las condiciones necesarias de primer orden y se resuelve el sistema:

$$\frac{\partial \mathrm{B}}{\partial q_1} = 40 - 14\mathrm{q}_1 - 4\mathrm{q}_2 = 0$$

$$\frac{\partial \mathrm{B}}{\partial \mathrm{q}_2} = 20 - 4\mathrm{q}_1 - 8\mathrm{q}_2 = 0$$

Hay un único punto crítico $\mathrm{q}_1 = \frac{5}{2};\ \mathrm{q}_2 = \frac{5}{4}$.

Mediante las condiciones de segundo orden se obtiene la matriz Hessiana:

$$\mathrm{HB}(\mathrm{q}_1, \mathrm{q}_2) = \begin{pmatrix} -14 & -4 \\ -4 & -8 \end{pmatrix}$$

que es definida negativa para todo (q_1, q_2), la función de beneficios es estrictamente cóncava y el punto crítico es el óptimo siendo un máximo global estricto, pues las condiciones necesarias de primer orden son también suficientes.

Ejemplo 8: Tratamos de obtener los óptimos de la función

$$f(x,y,z) = (x-1)^4 + y^2 + (z-1)^2$$

Hemos de aplicar primero las condiciones necesarias para obtener los puntos críticos puesto que estos son los únicos candidatos a óptimo y sin dificultad se tiene que existe un único punto crítico $x = 1, y = 0, z = 1$.

Al aplicar las condiciones de segundo orden tenemos que la matriz hessiana de la función es:

$$Hf(x,y,z) = \begin{pmatrix} 12(x-1)^2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} \;\; \rightarrow \;\; Hf(1,0,1) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

que es semidefinida positiva, luego es posible que sea un mínimo, no lo podemos afirmar, lo único que podemos decir de momento es que no es máximo. Ahora bien, si estudiamos la matriz $Hf(x,y,z)$ en un entorno del punto $(1,0,1)$ observamos que se mantiene definida positiva o semidefinida positiva y podemos entonces afirmar, sin más que tener en cuenta la proposición de las condiciones suficientes de segundo orden, que el punto es un mínimo local.

4. OPTIMIZACIÓN CON RESTRICCIONES DE IGUALDAD

4.1 INTRODUCIÓN

Como se citó en la optimización sin restricciones, se va a suponer que los programas matemáticos que se formulan son diferenciables, lo que quiere decir, que todas las funciones que intervienen se pueden derivar y por consiguiente podré aplicarles los resultados del cálculo diferencial.

En los programas sin restricciones todos los elementos que pertenecen al conjunto de definición de la función objetivo son candidatos a ser óptimos del problema, sin embargo, en los problemas con restricciones, además, se tendrán que verificar las restricciones en igualdad, por lo que sólo tendremos que trabajar con puntos del conjunto admisible, que es conjunto de puntos del dominio de definición de todas las funciones que intervienen en la formulación del problema y que verifiquen las restricciones.

La optimización con restricciones surge en la económica cuando se intenta analizar las acciones a realizar por un decisor que intenta lograr un objetivo ya sea maximizar la utilidad del consumidor, maximizar beneficios o minimizar costes de producción, cuando se dispone de cantidades fijas de uno o varios recursos, como puede ser un nivel fijo de presupuesto o de capacidades de producción de materias primas, etc. y dichas cantidades que son los recursos limitados disponibles se desean utilizar en su totalidad. Esas cantidades que vienen recogidas en los valores que alcanzan las constantes de restricciones del problema, y puesto que el decisor está obligado a utilizar los recursos en su totalidad, las restricciones estarán todas en igualdad.

Así pues un problema de programación clásica con restricciones se puede formular en términos matemáticos de la siguiente forma:

$$opt\ f(\bar{x}), \qquad s.a\colon\ \bar{h}(\bar{x}) = \bar{b},\ \ \bar{x}\ \in D$$

donde D es un subconjunto abierto de $\mathbb{R}^n$ y $f\colon \mathrm{D} \to \mathbb{R}$, $\bar{h}\colon \mathrm{D} \to \mathbb{R}^p$, con $p < n$ y al menos una de las funciones que intervienen en la formulación es no lineal.

Si en la formulación anterior llamamos $\bar{g}(\bar{x}) = \bar{h}(\bar{x}) - \bar{b}$ el problema estaría formulado

$$opt\ f(\bar{x}),\ s.a\colon\ \bar{g}(\bar{x}) = \bar{0},\ \ \bar{x}\ \in D\ ,$$

que es la formulación que utilizaremos.

Los programas sin restricciones que antes hemos estudiado son una relajación de estos programas, luego si se ha resuelto el programa irrestricto y dicha solución verifica las restricciones tendremos resuelto también el programa con restricciones.

La mayoría de los resultados que más adelante se obtendrán para estos programas con restricciones requieren que los puntos con los que se trabaja sean puntos regulares que es un concepto ligado a un sistema de ecuaciones.

Definición: Sea $\bar{g}$ con $p < n$ de clase C^1 en el conjunto abierto y no vacío D. El punto $\bar{x}^*\ \in D$ es regular del programa

$$opt\ f(\bar{x}), \qquad s.a: \ \bar{g}(\bar{x}) = \bar{0}, \ \ \bar{x} \ \in D$$

si es factible y la matriz jacobiana, de las funciones de restricción, en el punto tiene rango máximo, esto es rango p .

Luego, debemos comprobar que el rango de la matriz jacobiana de las funciones de restricción coincide con el número de restricciones del problema.

Ejemplo 9: Consideremos el programa:

$$\text{Opt: } f(x,y,z)$$
$$\text{s.a} \qquad g_1(x,y,z) = x + y + z = 0$$
$$g_2(x,y,z) = \ x^2 + \ y^2 + z^2 = 0$$

La matriz jacobiana de las dos funciones de restricción es:

$$J\bar{g}(x,y,z) = (\nabla g_1 \,, \ \nabla g_2) = \begin{pmatrix} 1 & 2x \\ 1 & 2y \\ 1 & 2z \end{pmatrix}$$

Luego el punto factible $(0{,}0{,}0)$ no es regular ya que

$$\text{rg } \left(J\bar{g}(0{,}0{,}0)\right) = \text{rg } \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \end{pmatrix} = 1 \neq 2 \quad \text{(Número de restricciones)}$$

Sin embargo, el punto factible (-1, 1, 0) sí que es punto regular pues

$$\text{rg}\left(J\bar{g}(-1{,}1{,}0)\right) = \text{rg } \begin{pmatrix} 1 & -2 \\ 1 & 2 \\ 1 & 0 \end{pmatrix} = 2$$

A continuación vamos a establecer el método para resolver este tipo de programas matemáticos que se conoce como método de sustitución de variables que nos permite obtener los óptimos de un programa con restricciones de igualdad.

4.2 METODO DE SUSTITUCION DE VARIABLES

Como consecuencia de manejar puntos regulares se puede establecer una condición necesaria y suficiente para que una solución local de un problema sin restricciones determine un óptimo local del programa de restricciones, según se establece en la siguiente proposición.

Proposición: Sea $\bar{g}$ de clase C^1 en el conjunto D abierto y no vacío y $\bar{x}^*$ un punto regular del programa $opt\ f(\bar{x}),\ s.a: \ \bar{g}(\bar{x}) = \bar{0}, \ \bar{x} \in D$. Supongamos que el primer de orden p de la matriz $J\bar{g}(\bar{x}^*)$ es no nulo y $\bar{x}_2^*$ el vector formado por las $n-p$ últimas componentes de $\bar{x}^*$. El punto $\bar{x}^*$ es máximo (resp. mínimo) del problema con restricciones sí y sólo sí $\bar{x}_2^*$ es máximo (resp. mínimo) del problema

$$opt\ f(\bar{\emptyset}(\bar{x}_2), \bar{x}_2), \qquad s.a: \bar{x}_2 \in B\ (\bar{x}_2^*, \delta)$$

donde $\bar{\emptyset}$ es la función implícita definida en el entorno B($\bar{x}_2^*$, δ), que sabemos existe, pues el punto $\bar{x}^*$ es regular.

Para poder aplicar, realmente, esta proposición al programa con restricciones necesitamos conocer la función implícita, esto es, del sistema de ecuaciones que determinan las restricciones tenemos que poder despejar p variables en función de las otras $n-p$ y así transformaríamos el problema con restricciones en uno sin restricciones de $n-p$ variables.

Ejemplo 10: Sea el programa $opt\ (x-1)^2 + y^2, \quad s.a \quad x + y = 10$

Es evidente que todos los puntos factibles son regulares puesto que la matriz jacobiana de la restricción tiene rango uno y podemos despejar de la restricción una variable en función de la otra, por ejemplo, $y = 10 - x$, y sustituyendo en la función objetivo, tendríamos el siguiente programa irrestricto

$$opt \quad (x-1)^2 + (10-x)^2$$

Si aplicamos las condiciones necesarias de primer orden, tendremos

$$f'(x) = 2(x-1) - 2(10-x) = 0$$

cuya solución es $x = \frac{11}{2}$. Aplicando las suficientes de segundo orden tenemos $f''\left(\frac{11}{2}\right) = 4 > 0$, luego $x = \frac{11}{2}$ es mínimo local del programa irrestricto y sustituyendo en la expresión $y = 10 - x$, tenemos que $x = \frac{11}{2}$, $y = \frac{9}{2}$ es mínimo local del programa con restricciones.

4.3 METODO DE LOS MULTIPLICADORES DE LAGRANE

Para aplicar este método es necesario definir una función auxiliar que se denomina función lagrangiana asociada al problema. Así pues dado el programa

$$opt\ f(\bar{x}), \qquad s.a\!: \ \bar{g}(\bar{x}) = 0, \bar{x} \in D$$

a la función $\mathcal{L}\colon D \times \mathbb{R}^P \to \mathbb{R}$ que asocia a cada $(\bar{x}, \bar{\lambda}) \in D \times \mathbb{R}^P$ el número real $\mathcal{L}(\bar{x}, \bar{\lambda}) = f(\bar{x}) + \lambda_1 g_1(\bar{x}) + \lambda_2 g_2(\bar{x}) + \cdots + \lambda_p g_p(\bar{x})$. Notemos que esta función depende de n+p variables y a las componentes del vector $\bar{\lambda}$ se las denomina multiplicadores de Lagrange, que existen tantos como funciones de restricción tiene el problema.

4.3.1 Condiciones de primer orden

La proposición que enunciamos a continuación deben verificarla todos los óptimos locales de los programas con restricciones de igualdad, es por tanto una condición necesaria de óptimo local, pero veremos, como en el caso de la optimización sin restricciones, que bajo condiciones de convexidad es suficiente de optimalidad global

4.3.1.1 Condición necesaria de óptimo local

Proposición: Sean $\bar{f}$ y $\bar{g}$ de clase C^1 en D abierto y no vacío. Si $\bar{x}^*$ es un punto regular y óptimo local del programa

$$opt\ f(\bar{x}), \qquad s.a\!: \ \bar{g}(\bar{x}) = \bar{0}, \ \bar{x} \in D$$

existe un único $\bar{\lambda}^* \in \mathbb{R}^P$ tal que $\nabla\mathcal{L}(\bar{x}^*, \bar{\lambda}^*) = \bar{0}$ siendo $\mathcal{L}$ la función lagrangiana asociada al problema.

Ejemplo 11: Aplicar la condición necesaria de óptimo local al siguiente programa: $opt\ 4x^2 + y^2$, $s.a\!: \ 2x + y = 5000$.

Todo punto factible es regular ya que $\nabla g = (2,\ 1)^t$ tiene rango 1 que es igual al número de restricciones.

Planteamos la función lagrangiana:

$$\mathcal{L}(x, y, \lambda) = 4x^2 + y^2 + \lambda(2x + y - 5000)$$

y resolvemos el sistema: $\nabla\mathcal{L}(x, y, \lambda) = \bar{0}$.

$$\frac{\partial L}{\partial x}(x, y, \lambda) = 8x + 2\lambda = 0 \iff \lambda = -4x \qquad (1)$$

$$\frac{\partial L}{\partial y}(x, y, \lambda) = 2y + \lambda \iff \lambda = -2y \qquad (2)$$

$$\frac{\partial L}{\partial \lambda}(x, y, \lambda) = 2x + y - 5000 = 0 \quad (3)$$

Igualamos las ecuaciones (1) y (2)

$$-4x = -2\text{y} \quad \Rightarrow \; y = 2x$$

y sustituyendo en la ecuación (3)

$$2x + 2x - 5000 = 0 \;\Rightarrow x = 1250$$

y, por tanto, $y = 2500, \;\; \lambda = -5000$. El punto candidato a óptimo será que tiene asociado ese valor de multiplicador de Lagrange.

4.3.1.2. Condición suficiente de óptimo global

Proposición: Sean $\bar{f}$ y $\bar{g}$ de clase C^1 en D abierto y no vacío. Si $\left(\bar{x}^*\bar{\lambda}^*\right) \in D \times \mathbb{R}^p$ es un punto crítico de la función lagrangiana asociada al programa

$$opt\; f(\bar{x}), \qquad s.a\text{:}\; \bar{g}(\bar{x}) = \bar{0}, \;\; \bar{x} \; \in D$$

con conjunto factible convexo y función objetivo convexa (resp. cóncava), entonces $\bar{x}^*$ es mínimo (resp. máximo) global del problema.

Así pues, puedo afirmar que cuando tengo un programa con restricciones de igualdad y la función objetivo es convexa y el conjunto factible convexo las condiciones necesarias de primer orden son también suficientes y el óptimo es un mínimo global y del mismo modo razonaríamos si la función objetivo es cóncava y ahora el óptimo sería un máximo local.

En el ejemplo anterior la función objetivo es convexa y el conjunto factible es convexo, pues es un hiperplano, luego podemos aplicar la última proposición y el punto obtenido $(1250,\; 2500)$, que antes era sólo candidato a óptimo, es un mínimo global, además sólo existía un posible candidato a óptimo, lo cual es obvio pues la función objetivo no sólo es convexa sino estrictamente convexa.

<u>Ejemplo</u> 12: Resolvamos el programa

$$\begin{aligned} \text{Min.} \;\; & x^2 + y^2 + z^2 \\ \text{s.a:} \quad & x + y = 1 \\ & y - z = 2 \end{aligned}$$

La función lagrangiana asociada será:

$$\mathcal{L}\left(x, y, z, \lambda_{1,} \lambda_2\right) = \ x^2 + y^2 + z^2 + \lambda_1(x + y - 1) + \lambda_2(y - z - 2)$$

Resolviendo el sistema $\nabla L\ (x, y, z, \lambda_1, \lambda_2) =\ \bar{0}$, se obtiene el único punto crítico de la lagrangiana

$$\underbrace{(0, 1, -1, 0, -2)}_{\left(x, y, z, \lambda_{1,} \lambda_2\right)}$$

La función $f(x, y, z) = x^2 + y^2 + z^2$ es estrictamente convexa (se puede comprobar con técnicas utilizadas anteriormente) y el conjunto factible es convexo, pues es la intersección de dos conjuntos convexos, entonces el punto $(0, 1, -1)$ es el único mínimo global del problema.

4.3.2 Condiciones de segundo orden

Las condiciones que se establecen a continuación requieren que las funciones que intervienen en la formulación del programa sean de clase C^2 en el dominio de definición y el punto que se pretende analizar como posible óptimo tiene que ser regular.

4.3.2.1 Condiciones necesarias de segundo orden

Proposición: Sean f y $\bar{g}$ de clase C^2 en el conjunto D abierto y no vacío. Si $\bar{x}^*$ es regular y mínimo (resp. máximo) del programa

$$opt\ f(\bar{x}), \qquad s.a\colon\ \bar{g}(\bar{x}) = \bar{0},\ \ \bar{x}\ \in D$$

entonces existe un único $\bar{\lambda}^* \epsilon\ \mathbb{R}^p$ tal que el punto $\left(\bar{x}^*, \bar{\lambda}^*\right)$ es punto crítico de la función Lagrangiana y si $\bar{x}^*$ es mínimo (resp. máximo) local, la forma cuadrática $\bar{h}^t H_{\bar{x}} L\left(\bar{x}^*, \bar{\lambda}^*\right)\bar{h}$ restringida a $\bar{h}^t J\bar{g}(\bar{x}^*) = \bar{0}^t$ es definida positiva (resp. definida negativa) o semidefinida positiva (resp. semidefinida negativa).

4.3.2.2 Condiciones suficientes de óptimo local

Como en el caso de los programas sin restricciones, la condición anterior no es suficiente cuando la forma cuadrática restringida es semidefinida, luego voy a citar una proposición donde no aparezca esa condición y tendríamos unas condiciones suficientes de optimalidad y podríamos formular la proposición siguiente.

Proposición: Sean f y $\bar{g}$ de clase C^2 en el conjunto D abierto y no vacío y $\bar{x}^*$ un punto regular del problema

$$opt\, f(\bar{x}), \qquad s.a\colon\ \bar{g}(\bar{x}) = \bar{0},\ \ \bar{x}\ \in D$$

para el que existe un $\bar{\lambda}^* \epsilon\, \mathbb{R}^p$ tal que el punto $(\bar{x}^*, \bar{\lambda}^*)$ es un punto crítico de la lagrangiana $\mathcal{L}(\bar{x}, \bar{\lambda})$ asociada al programa. Entonces si la forma cuadrática $\bar{h}^t H_{\bar{x}} L(\bar{x}^*, \bar{\lambda}^*)\bar{h}$ restringida a $\bar{h}^t J\bar{g}(\bar{x}^*) = \bar{0}^t$ es definida positiva (resp. definida negativa), el punto $\bar{x}^*$ es mínimo (resp. máximo) local del problema y si es indefinida no es máximo ni mínimo local.

Así pues, si tengo que aplicar las condiciones suficientes de segundo orden para determinar si un punto que verifica las condiciones necesarias de primer orden es máximo o mínimo, tengo que estudiar una forma cuadrática restringida y ese estudio puede realizarse despejando de las restricciones de la forma cuadrática, esto es, $\bar{h}^t J\bar{g}(\bar{x}^*) = \bar{0}^t$, p variables y sustituirlas en $\bar{h}^t H_{\bar{x}} L(\bar{x}^*, \bar{\lambda}^*)\bar{h}$ para así obtener una forma cuadrática sin restricciones y estudiar esta última. Sin embargo existe otro procedimiento para determinar si la forma cuadrática restringida es definida positiva o definida negativa y es estudiar el comportamiento de una serie de menores principales de la matriz orlada de la forma cuadrática con la matriz que define las restricciones de la forma cuadrática. En este caso sería la matriz orlada de la hessiana de la lagrangiana con la jacobiana de las restricciones, esto es la matriz:

$$\overline{H_{\bar{x}}}L(\bar{x}^*, \bar{\lambda}^*) = \begin{pmatrix} 0 & J^t\bar{g}(\bar{x}^*) \\ J\bar{g}(\bar{x}^*) & H_{\bar{x}}L(\bar{x}^*, \bar{\lambda}^*) \end{pmatrix}$$

y puesto que la función objetivo es de n variables y tengo p restricciones con $p < n$, he de estudiar los $n-p$ últimos menores principales naturalmente ordenados y si van alternando de signo, comenzando con el de $(-1)^{p+1}$ la forma cuadrática restringida es definida negativa y según la proposición, el punto $\bar{x}^*$ es un máximo local, mientras que si los $n-p$ últimos menores principales naturalmente ordenados tienen todos el mismo signo, el signo de $(-1)^p$, la forma cuadrática restringida es definida positiva y el punto $\bar{x}^*$ es un mínimo local.

<u>Ejemplo</u> 13: Tratamos de resolver el problema

$$opt\ \ 8x^3 + 2xy - 4x^2 + y^2 + 1, \qquad s.a\colon\ 25x + y = 10$$

Todos los puntos factibles son regulares pues $g(x,y) = 25x + y - 10$ y su gradiente es $\nabla g(x,y,z) = (25,\ 1)^t$ que tiene rango uno.

Sea la función $\mathcal{L}(x,y,\lambda) = 8x^3 + 2xy - 4x^2 + y^2 + 1 + \lambda(25x + y - 10)$, las condiciones necesarias de primer orden determinan los posibles óptimos del programa y se obtienen resolviendo el sistema:

(1) $\frac{\partial L}{\partial x} = 24x^2 + 2y - 8x + 25\lambda = 0$

(2) $\frac{\partial L}{\partial y} = 2x + 2y + \lambda = 0$

(3) $\frac{\partial L}{\partial \lambda} = 25x + y - 10 = 0$

Despejando de (2) tenemos $y = -\frac{1}{2}\lambda - x$ y sustituyendo en (1) y (3) tenemos:

$$24x^2 - 10x + 24\lambda = 0$$

$$50x - \lambda - 2x - 20 = 0 \quad \Rightarrow \quad \lambda = 48x - 20$$

y sustituyendo en la de arriba

$$24x^2 - 10x + 24(48x - 20) = 0$$

$$24x^2 + 1142x - 480 = 0 \ \Rightarrow \ 12x^2 + 571x - 240 = 0$$

cuya solución es:

$$x = \frac{-571 \pm \sqrt{(571)^2 + 4.12.240}}{24} = \frac{-571 \pm 581}{24} = \begin{matrix} \nearrow \frac{5}{12} \\ \searrow 48 \end{matrix}$$

- Si $x = \frac{5}{12} \Rightarrow y = -\frac{5}{12}$ y tenemos como posible óptimo el punto $P_1 = \left(\frac{5}{12}, -\frac{5}{12}\right)$ que tiene asociado un valor del multiplicador $\lambda = 0$.
- Si $x = 48 \Rightarrow y = -1190$ y tenemos como posible óptimo el punto $P_2 = (48, -1190)$ con valor del multiplicador $\lambda = 2284$.

Hemos de aplicar las condiciones suficientes de segundo orden en cada uno de los puntos que verifican las condiciones necesarias para así determinar el tipo de óptimo y aplicando lo anteriormente indicado de la matriz orlada tenemos que hemos de estudiar los $n - p = 2 - 1 = 1$ último menor principal de la matriz

- Para el punto $P_1 = \left(\frac{5}{12}, -\frac{5}{12}\right)$ con $\lambda = 0$ tendremos la matriz:

$$\overline{H_{(x,y)}}L\left(\frac{5}{12},-\frac{5}{12},0\right)=\begin{pmatrix}0 & 25 & 1\\ 25 & 48x-8 & 2\\ 1 & 2 & 2\end{pmatrix}_{\left(\frac{5}{12},-\frac{5}{12},0\right)}=\begin{pmatrix}0 & 25 & 1\\ 25 & 12 & 2\\ 1 & 2 & 2\end{pmatrix}$$

y el último menor principal toma el valor -1162 que tiene el signo de $(-1)^1$, luego el punto $P_1=\left(\frac{5}{12},-\frac{5}{12}\right)$ es un mínimo local.

- Para el punto $P_2=(48,-1190)$ con $\lambda=2284$ tendremos la matriz

$$\overline{H_{(x,y)}}L(48,-1190,2284)=\begin{pmatrix}0 & 25 & 1\\ 25 & 48x-8 & 2\\ 1 & 2 & 2\end{pmatrix}_{(48,-1190,2284)}=$$

$$=\begin{pmatrix}0 & 25 & 1\\ 25 & 2296 & 2\\ 1 & 2 & 2\end{pmatrix}$$

Su último menor principal vale -3456 que tiene el mismo signo que $(-1)^1$ y también es mínimo local el punto $P_2=(48,-1190)$.

4.4 ANÁLISIS DE SENSIBILIDAD

En economía las constantes de las restricciones de un programa

$$opt\ f(\bar{x}),\qquad s.a:\ \bar{h}(\bar{x})=\bar{b},\ \ \bar{x}\in D$$

suelen representar cantidades de ciertos recursos que un decisor tiene capacidad de utilizar en su totalidad en el proceso productivo para optimizar la función de producción o la función de utilidad o de beneficios, etc., por ejemplo, supongo que la función objetivo es una función de beneficios y las constantes de restricción cantidades de materias primas que se utilizan en la fabricación de diferentes artículos; si se ha resuelto el problema y he encontrado la cantidad que se ha de producir de los diferentes artículos para maximizar el beneficio y he utilizado las materias primas en su totalidad, podría preguntarme si es interesante acudir al mercado para comprar alguna materia prima adicional y así variar la cantidad producida.

Esta información la van a proporcionar los multiplicadores obtenidos en la resolución del programa como se indica en la siguiente proposición que indica la variación que se produce en el valor de la función objetivo en el óptimo cuando las constantes de restricción varían en un entorno del punto $\bar{b}$ fijo.

Proposición: Sean f y $\bar{h}$ de clase C^2 en el conjunto D abierto y no vacío y $\bar{x}^* \epsilon\ D$ un punto regular del programa

$$opt\ f(\bar{x}), \qquad s.a\colon\ \bar{h}(\bar{x}) = \bar{b}^*,\ \ \bar{x}\ \in D$$

que junto con $\bar{\lambda}^*$ satisfacen las condiciones necesarias de primer orden de óptimo local y las condiciones suficientes. Entonces existe un entorno $B(\bar{b}^*, \delta)$ tal que si $\bar{b}\ \epsilon B(\bar{b}^*, \delta)$ existen funciones de clase C^1 , $\bar{x}(\bar{b}),\ \bar{\lambda}(\bar{b})$ definidas en ese entorno que satisfacen las condiciones necesarias de primer orden y las suficientes de óptimo del programa

$$opt\ f(\bar{x}), \qquad s.a\colon\ \bar{h}(\bar{x}) = \bar{b},\ \ \bar{x}\ \in D$$

Además si se calcula el valor de la función objetivo en el óptimo, esto es $G(\bar{b}) = f(\bar{x}(\bar{b}))$, se tiene $\frac{\partial G(\bar{b})}{\partial b_i}|_{\bar{b}_i^*} = \pm\lambda_i^*$.

He indicado $\pm\lambda_i^*$ pues la parcial indicada, esto es la variación que experimenta el valor de la función objetivo en el óptimo cuando se realiza una variación marginal en la constante de la i-ésima restricción es el valor en el óptimo del i-ésimo multiplicador si la lagrangiana asociada al problema que se ha elegido para su resolución es $\mathcal{L}(\bar{x},\ \bar{\lambda}) = f(\bar{x}) + \bar{\lambda}^t(\bar{b} - \bar{h}(\bar{x}))$ y $-\lambda_i^*$ si la lagrangiana elegida es $\mathcal{L}(\bar{x},\ \bar{\lambda}) = f(\bar{x}) + \bar{\lambda}^t(\bar{h}(\bar{x}) - \bar{b})$.

Por consiguiente pueden interpretarse, en términos económicos, los multiplicadores como un sistema de “precios” de las materias primas que representan las constantes de restricción; no son precios de mercado, por eso también son llamados “precios marginales” o “precios de oportunidad” e indican hasta donde estaríamos dispuestos a pagar por una unidad adicional de ese recurso escaso para aumentar el valor de la función objetivo en el óptimo. Así, si por ejemplo la variación es $\lambda_i^* > 0$, pues he tomado la lagrangiana correspondiente para que suceda eso, nunca pagaría en el mercado más de dicho valor por una unidad adicional de la materia prima i-ésima pues en caso contrario el incremento que se experimentaría en el valor de la función objetivo en el óptimo si utilizo dicha unidad adicional en el proceso productivo, sería menor que el coste en conseguirla en el mercado.

5. CONCLUSIONES

El principal objetivo de este Trabajo de Fin de Grado es presentar la relación directa entre la Teoría de Optimización Matemática y la asignación de recursos en el ámbito de la Economía.

Cuando las empresas persiguen sus objetivos de maximizar beneficios y/o minimizar costes lo que hacen es optimizar funciones matemáticas y por esa razón la comprensión de lo relativo a la optimización de funciones que nos ocupa, se convierte en un punto clave para cualquier profesional.

Como he mencionado, el cálculo puede ser más sencillo si nos encontramos ante un programa convexo, porque las condiciones necesarias son también suficientes.

El objetivo principal de las empresas siempre es maximizar sus beneficios y/o minimizar sus costes, pero se encuentran sujetas a unas limitaciones, las cuales están relacionadas con la Optimización Matemática con restricciones en forma de igualdad.

Hay que destacar que resulta muy ventajoso que el programa sea convexo, es decir, que la función a optimizar sea cóncava o convexa según se quiera conseguir un máximo o mínimo y que el conjunto admisible que determinan las restricciones en forma de igualdad sea convexo, así conseguiré que las condiciones necesarias sean también suficientes y que los puntos obtenidos sean óptimos globales.

Por último hay que destacar el análisis de sensibilidad, que tiene gran aplicación en el ámbito económico y que da solución a los problemas de optimización sujetos a una restricción en forma de igualdad donde se busca la función lagrangiana y tenemos los multiplicadores de lagrange, que en el ámbito económico, representa el coste de oportunidad (que es lo que la empresa estaría dispuesta a pagar por una unidad adicional de materias primas para aumentar en el óptimo el valor de la función objetivo).

6. BIBLIOGRAFÍA

Arranz Sombría, M.R. y Pérez González, M.P. (1997): *Matemáticas para la Economía: Optimización y Operaciones Financieras.* Editorial AC, Madrid.

Fernández Lechón, R. y Castrodeza, C. (1989): *Programación Lineal.* Editorial Ariel Economía. Barcelona.

Guerrero Casas, F.M. (1994): *Curso de Optimización: Programación Matemática.* Editorial Ariel Economía. Barcelona.

Perez Grasa, I.; Minguillón, E. y Jarne, G. (2010): *Matemáticas para la Economía: Programación Matemática y Sistemas Dinámicos.* Editorial McGraw Hill. Madrid.

Soto Torres, M.D. (2007): *Métodos de Optimización.* Editorial Delta Publicaciones, Madrid.