Capítulo 5

Fundamentos de Optimización: Optimización no lineal

Sumario. Definciones básicas de optimización. Conjuntos Convexos. Funciones convexas. Optimización de funciones convexas. Condiciones de Karush-Kuhn-Tucker. Condiciones necesarias de primer orden: Ejemplos y Contraejemplos. Condiciones necesarias de segundo orden. Condiciones suficientes. Interpretación de los multiplicadores de KKT. Dualidad.

5.1. Conceptos básicos

La teoría de optimización clásica o programación matemática está constituida por un conjunto de resultados y métodos analíticos y numéricos enfocados a encontrar e identificar al mejor candidato de entre una colección de alternativas, sin tener que enumerar y evaluar explícitamente todas esas alternativas. Un problema de optimización es, en general, un problema de decisión.

Con el fin de ilustrar de forma adecuada la estructura y composición de un problema de optimización, introduciremos a continuación un sencillo ejemplo.

Ejemplo 6 ***(Construcción de una caja con volumen máximo)*** *Supongamos que queremos determinar las dimensiones de una caja rectangular de forma que contenga el mayor volumen posible, pero utilizando para ello una cantidad fija de material. El problema en forma abstracta se podría plantear en los siguientes términos*

Maximizar	*Volumen de la caja*
sujeto a	*Área lateral fija*

Con el fin de resolver este problema habrá que modelizarlo matemáticamente, es decir tendremos que expresarlo en términos matemáticos. El primer paso para modelizar un problema de optimización es identificar y definir las variables que están implicadas en dicho problema, en este caso y puesto que estamos tratando de determinar el tamaño de una caja rectangular, la opción más clara es considerar como variables sus tres dimensiones rectangulares usuales (ancho, largo, alto) y que representamos con x, y, z.

Con estas variables, la función para la que tenemos que encontrar el mejor valor será el volumen de la caja que puede expresarse como

$$V(x, y, z) = xyz$$

A continuación debemos tener en cuenta las limitaciones existentes sobre el material. Como este material se utiliza para construir las paredes de la caja, necesitaremos considerar el área lateral de la misma, y si la caja tiene tapa, dicha área será

$$A(x, y, z) = 2(xy + yz + zx)$$

Por último, teniendo en cuenta que las dimensiones de la caja no pueden ser negativas el problema puede expresarse matemáticamente como

$$\begin{array}{cc} \textit{Maximizar} & xyz \\ \textit{sujeto a} & 2(xy + yz + zx) = A \\ & x, y, z \geq 0 \end{array}$$

En este ejemplo se distinguen tres elementos fundamentales: las variables del problema, una función de esas variables y un conjunto de relaciones que deben cumplir las variables del problema. Estos elementos se repetirán en todos los problemas de optimización y se definen formalmente a continuación:

1.- *Variables de decisión:* El primer elemento clave en la formulación de problemas de optimización es la selección de las variables independientes que sean adecuadas para caracterizar los posibles diseños candidatos y las condiciones de funcionamiento del sistema. Como variables independientes se sueleng elegir aquellas que tienen un impacto significativo sobre la función objetivo.

Representaremos las variables independientes se representarán mediante vectores columna de $\mathbb{R}^n$

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

o vectores fila

$$\mathbf{x}^T = (x_1, \ldots, x_n)$$

Aunque para los casos $n = 1$, 2 y 3 se emplearán las notaciones usuales de x, (x, y) y (x, y, z) respectivamente.

2.- *Restricciones:* Una vez determinadas las variables independientes, el siguiente paso es establecer, mediante ecuaciones o inecuaciones las relaciones existentes entre las variables de decisión. Estas relaciones son debidas, entre otras razones, a limitaciones en el sistema, a leyes naturales o a limitaciones tecnológicas y son las llamadas restricciones del sistema. Podemos distinguir dos tipos de restricciones:

1. *a)* *Restricciones de igualdad:* Son ecuaciones entre las variables de la forma

$$h(\mathbf{x}) = h(x_1, \ldots, x_n) = 0$$

donde $h : A \subseteq \mathbb{R}^n \to \mathbb{R}$ es una función real de variables reales definida sobre un conjunto A de números reales.

b) *Restricciones de desigualdad:* Son inecuaciones entre las variables de la forma

$$g(\mathbf{x}) = g(x_1, \ldots, x_n) \leq 0$$

donde $g : A \subseteq \mathbb{R}^n \to \mathbb{R}$ es una función real de variables reales definida sobre un conjunto A de números reales.

Nota 1 *Solamente se han considerado restricciones de dos tipos: restricciones de igualdad del forma $h(x_1, \ldots, x_n) = 0$ y restricciones de desigualdad de la forma $g(x_1, \ldots, x_n) \leq 0$, debido a que siempre es posible, mediante una simple transformación, expresar el problema en términos de esta clase de restricciones, tal y como se puede apreciar en la siguiente tabla:*

Tipo de Restricción	*Transformación*	*Nueva Restricción*
$h(x_1, \ldots, x_n) = b$	$\widehat{h} = h - b$	$\widehat{h}(x_1, \ldots, x_n) = 0$
$g(x_1, \ldots, x_n) \leq c$	$\widehat{g} = g - c$	$\widehat{g}(x_1, \ldots, x_n) \leq 0$
$g(x_1, \ldots, x_n) \geq c$	$\widehat{g} = c - g$	$\widehat{g}(x_1, \ldots, x_n) \leq 0$

Nota 2 *Las condiciones sobre las variables del tipo $x_i \geq 0$, $x_y \leq 0$ o similares no se incluyen dentro de este conjunto de restricciones y tienen un tratamiento particular, son restricciones en las variables o llamadas de tipo cota.*

3.- *Función objetivo:* Finalmente, el último ingrediente de un problema de optimización es la función objetivo, también llamado índice de rendimiento o criterio de elección. Este es el elemento utilizado para decidir los valores adecuados de las variables de decisión que resuelven el problema de optimización. La función objetivo permite determinar los mejores valores para las variables de decisión.

Independientemente del criterio seleccionado, dentro del contexto de la optimización matemática el adjetivo "mejor" siempre indica los valores de las variables de decisión que producen el mínimo o máximo valor (según el criterio utilizado) de la función objetivo elegida.

Algunos de estos criterios pueden ser por ejemplo de tipo económico (coste total, beneficio), de tipo tecnológico (energía mínima, máxima capacidad de carga, máxima tasa de producción) o de tipo temporal (tiempo de producción mínimo) entre otros.

1. Es importante hacer notar que en esta guía se utilizará un único criterio de optimización, no estamos interesados en encontrar una solución que, por ejemplo, minimice el coste, maximice la producción y al mismo tiempo minimice la energía utilizada. Esta simplificación es muy importante, puesto que aunque en muchas situaciones prácticas sería deseable alcanzar una solución que sea la mejor con respecto a un número de criterios diferentes: la solución ideal sería maximizar beneficios al mínimo coste. No obstante una forma de tratar objetivos que chocan entre sí, es seleccionar un criterio como primario y el resto de posibles criterios como secundarios. El criterio primario se utiliza como la función objetivo del problema de optimización, mientras que los criterios secundarios serían valores mínimos y máximos aceptables que serían tratados como restricciones del problema. Los problemas que utilizan varios criterios de búsqueda entran dentro de la llamada *programación multiobjetivo.*

Con la introducción de estos tres elementos, el objetivo de los problemas de optimización matemática está claro: Un problema de optimización consiste en la búsqueda de valores para unas determinadas variables (*variables de decisión*) de forma que, cumpliendo un conjunto de requisitos representados mediante ecuaciones y/o inecuaciones algebráicas (*restricciones*) que limitarán la elección de los valores de las variables de decisión, proporcionan el mejor valor posible para una función (*función objetivo*) que es utilizada para medir el rendimiento del sistema que se estudia. Como se ha comentado previamente, con el mejor valor se quiere indicar el mayor o el menor valor posible para la función objetivo. En resumen, buscamos valores que cumplan unas condiciones y minimicen o maximicen una función que caracteriza el sistema.

El planteamiento abstracto general para resolver problemas de este tipo es el siguiente

$$\begin{array}{rl} \text{Optimizar} & f\left(x_1,\ldots,x_n\right) \\ \text{Sujeto a} & \textit{Restricciones} \end{array}$$

donde el empleo del término **Optimizar** incluirá a ambos objetivos tanto de *Minimización* como de *Maximización*. No obstante y en relación a este aspecto, la mayoría de los planteamientos pueden hacerse con uno sólo de los objetivos, por ejemplo el de minimización, ya que un problema con objetivo de maximización se puede transformar en otro equivalente con objetivo de minimización multiplicando para ello la función objetivo por (-1) tal y como podemos comprobar en la figura 5.1. El mínimo de la función $f(x) = x^2 + 1$ se alcanza en el punto $x^* = 0$. En este punto también se alcanza el máximo de la función opuesta $g(x) = -f(x) = -x^2 - 1$, notar que aunque el punto buscado en ambos casos es el mismo, los valores que cada función toma en dicho punto son justamente uno el opuesto del otro:

$$\begin{aligned} f\left(x^*\right) &= f(0) = 1 \\ g\left(x^*\right) &= g(0) = -1 \end{aligned}$$

Veamos algunos ejemplos de problemas de optimización.

Ejemplo 7 ***Distancia más corta entre dos curvas.*** *Supongamos que se quiere calcular la mínima distancia entre dos curvas de ecuaciones $C_1 \equiv y = f(x)$ y $C_2 \equiv y = g(x)$ que no se corten entre sí. El problema se resuelve considerando un punto en cada curva y utilizando la fórmula de la distancia entre dos puntos para plantear el problema como*

$$\begin{array}{rl} \textit{Minimizar} & d\left(P,Q\right) \\ \textit{sujeto a} & P \in C_1 \\ & Q \in C_2 \end{array}$$

o de forma explícita

$$\begin{array}{rc} \textit{Minimizar} & \sqrt{\left(x_2-x_1\right)^2 + \left(y_2-y_1\right)^2} \\ \textit{sujeto a} & y_1 = f\left(x_1\right) \\ & y_2 = g\left(x_2\right) \end{array}$$

siendo

$$\begin{aligned} P &= \left(x_1,y_1\right) \in C_1 \\ Q &= \left(x_2,y_2\right) \in C_2 \end{aligned}$$

las coordenadas de los dos puntos. Este problema se puede extender de forma trivial a curvas en $\mathbb{R}^n$.

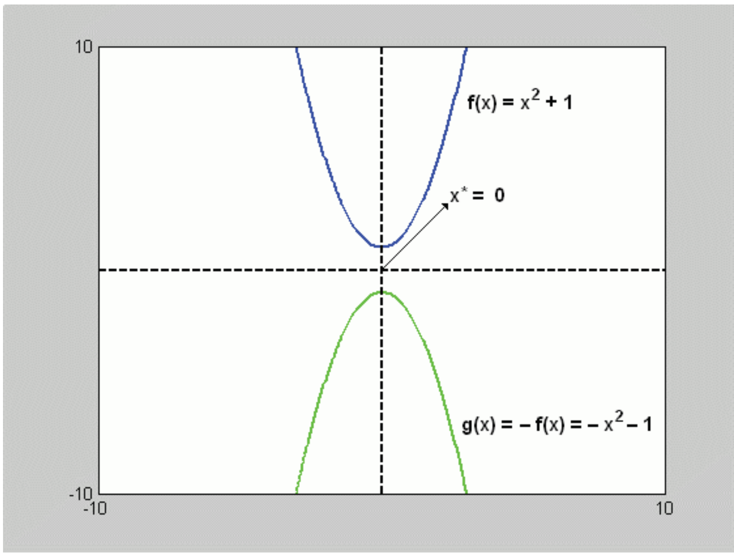


Figura 5.1: Equivalencia entre mín $f(x)$ y máx $g(x) = -f(x)$.

Ejemplo 8 ***Problema lineal.*** *Supongamos que queremos obtener el número de artículos que debemos fabricar de diferentes productos con coste fijo, teniendo para ello un presupuesto limitado y obteniendo a la misma vez el máximo beneficio posible. El problema podría plantearse como:*

$$\begin{array}{ll} \textit{Minimizar} & b_1x_1 + b_2x_2 + \cdots + b_nx_n \\ \textit{Sujeto a} & c_1x_1 + c_2x_2 + \cdots + c_n \leq P \\ & x_k \geq \quad k = 1, \ldots, n \end{array}$$

donde P es el presupuesto total disponible y los parámetros b_k y c_k para $k = 1, 2, \ldots, n$ son el beneficio y el coste, respectivamente, para cada uno de los productos y siendo x_k la cantidad de producto k que se debe fabricar.

5.2. Definiciones

En esta sección se dan las definiciones elementales relacionadas con la teoría de la optimización matemática con el objetivo de que el lector se familiarice con el lenguaje matemático utilizado.

Definición 2 (PPNL) *Se define el problema fundamental de la optimización estática o problema de programación no lineal (*PPNL*) al expresado como*

$$(\text{PPNL}) \left\{ \begin{array}{lll} \textit{Optimizar} & f(x_1, \ldots, x_n) & \\ \textit{Sujeto a} & h_i(x_1, \ldots, x_n) = 0 & i = 1, \ldots, m \\ & g_j(x_1, \ldots, x_n) \leq 0 & j = 1, \ldots, p \\ & (x_1, \ldots, x_n) \in A \subseteq \mathbb{R} & \end{array} \right. \qquad (5.1)$$

donde $f, h_i, g_j : A \to \mathbb{R}$*, o en notación vectorial como*

$$(\text{PPNL}) \left\{ \begin{array}{ll} \textit{Optimizar} & f(\mathbf{x}) \\ \\ \textit{Sujeto a} & \mathbf{h}(\mathbf{x}) = 0 \\ \\ & \mathbf{g}(\mathbf{x}) \leq 0 \\ \\ & \mathbf{x} = (x_1, \ldots, x_n) \in A \subseteq \mathbb{R} \end{array} \right.$$

donde ahora $\mathbf{h} : A \to \mathbb{R}^m$*, y* $\mathbf{g} : A \to \mathbb{R}^p$ *son funciones vectoriales.*

Las funciones implicadas en la definición del problema fundamental no tienen porque tener ninguna propiedad en particular, pero en nuestro caso vamos a introducir hipótesis adicionales que ayudarán a simplificar el problema; por ejemplo, supondremos de forma general que las funciones f, h_i y g_j son continuas y en la mayoría de los casos tendrán derivadas primeras y segundas también continuas.

La resolución del problema de optimización *PPNL* consistirá en primer lugar, en buscar valores para las variables de decisión x_i que cumplan las ecuaciones e inecuaciones que forman el sistema de las restricciones y en segundo lugar encontrar de entre estos valores aquel o aquellos que proporcionen el mayor (si el objetivo es maximizar) o menor (si el objetivo es minimizar) valor para la función real $f(x_1, \ldots, x_n)$.

Definición 3 *Se distinguen algunos casos particulares del problema general de optimización 5.1.*

1. Problemas sin restricciones: *En este tipo de problemas no hay restricciones de ningún tipo, es decir* $m = p = 0$*. La expresión general para estos problemas es*

$$(\text{PSR}) \left\{ \begin{array}{ll} \textit{Optimizar} & f(x_1, \ldots, x_n) \\ \\ \textit{Sujeto a} & (x_1, \ldots, x_n) \in A \end{array} \right. \tag{5.2}$$

   Las únicas limitaciones vienen dadas por el conjunto A *de* $\mathbb{R}^n$ *donde esté definida la función* $f(x_1, \ldots, x_n)$*.*

2. Problemas de Lagrange o problemas sólo con restricciones de igualdad: *Son problemas de optimización con restricciones donde solamente existen restricciones de igualdad, por tanto* $m \neq 0$ *y* $p = 0$*. Son problemas de la forma*

$$(\text{PRI}) \left\{ \begin{array}{ll} \textit{Optimizar} & f(x_1, \ldots, x_n) \\ \\ \textit{Sujeto a} & h_i(x_1, \ldots, x_n) = 0 \quad i = 1, \ldots, m \\ \\ & (x_1, \ldots, x_n) \in A \end{array} \right. \tag{5.3}$$

   No hay restricciones dadas por inecuaciones, sólo por ecuaciones.

3. Problemas unidimensionales o univariantes: *Este es un caso particular de los problemas sin restricciones en los que solamente hay una variable, es decir para $n = 1$, $m = 0$ y $p = 0$. El problema se expresa como*

$$\text{(P1D)} \begin{cases} \textit{Optimizar} & f(x) \\ \\ \textit{Sujeto a} & x \in I \subseteq \mathbb{R} \end{cases} \tag{5.4}$$

donde I es, en la mayoría de las ocasiones, un intervalo.

Definición 4 (Solución factible) *Diremos que $\overline{\mathbf{x}} = (\overline{x}_1, \ldots, \overline{x}_n) \in A \subseteq \mathbb{R}$ es una* solución factible *del problema* PPNL *(ecuación 5.1) si cumple todas sus restricciones, es decir*

$$\overline{\mathbf{x}} \textit{ solución factible } \Leftrightarrow \begin{cases} h_i(\overline{x}_1, \ldots, \overline{x}_n) = 0 & i = 1, \ldots, m \\ \\ g_j(\overline{x}_1, \ldots, \overline{x}_n) \leq 0 & j = 1, \ldots, p \end{cases}$$

Definición 5 (Conjunto factible) *Se define* región *o* conjunto factible Ω *del problema* PPNL *al conjunto de todas sus soluciones factibles*

$$\Omega = \{\overline{\mathbf{x}} \in A \subseteq \mathbb{R} : \ \overline{\mathbf{x}} \textit{ es una solución factible}\}$$

Nota 3 *Con estas definiciones se puede decir que el objetivo al intentar resolver el problema de optimización* PPNL *es encontrar la "mejor" de todas las soluciones factibles.*

Definición 6 (Mínimo global) *Diremos que $\mathbf{x}^* = (x_1^*, \ldots, x_n^*) \in \Omega \subseteq \mathbb{R}^n$ es un* mínimo global *del problema* PPNL *o que $f(\mathbf{x})$ tiene un mínimo global sobre Ω , el conjunto factible de* PPNL, *si*

$$\forall\, \overline{\mathbf{x}} \in \Omega; \quad \overline{\mathbf{x}} \neq \mathbf{x}^* \Rightarrow \quad f(\mathbf{x}^*) \leq f(\overline{\mathbf{x}})$$

El punto $\mathbf{x}^*$ será mínimo global estricto *si la desigualdad es estricta*

$$\forall\, \overline{\mathbf{x}} \in \Omega; \quad \overline{\mathbf{x}} \neq \mathbf{x}^* \Rightarrow \quad f(\mathbf{x}^*) < f(\overline{\mathbf{x}}).$$

Esta definición implica que no hay en Ω un punto en el que la función tome un valor menor que el que toma en $\mathbf{x}^*$.

Definición 7 (Máximo global) *Diremos que $\mathbf{x}^* = (x_1^*, \ldots, x_n^*) \in \Omega \subseteq \mathbb{R}^n$ es un* máximo global *del problema* PPNL *o que $f(x)$ tiene un máximo global sobre Ω, el conjunto factible de* PPNL, *si*

$$\forall\, \overline{\mathbf{x}} \in \Omega; \quad \overline{\mathbf{x}} \neq \mathbf{x}^* \Rightarrow \quad f(\mathbf{x}^*) \geq f(\overline{\mathbf{x}})$$

El punto $\mathbf{x}^*$ será máximo global estricto *si la desigualdad es estricta*

$$\forall\, \overline{\mathbf{x}} \in \Omega; \quad \overline{\mathbf{x}} \neq \mathbf{x}^* \Rightarrow \quad f(\mathbf{x}^*) > f(\overline{\mathbf{x}})$$

Esta definición implica que no hay en Ω un punto en el que la función tome un valor mayor que el que toma en $\mathbf{x}^*$.

Nota 4 *Los máximos y mínimos globales de un problema de optimización se denominan* extremos globales.

Definición 8 (Solución óptima) *Diremos que* $\mathbf{x}^* \in \Omega \subseteq \mathbb{R}^n$ *es una* solución óptima *del problema* PPNL *o que* $f(\mathbf{x})$ *tiene un óptimo en* $\mathbf{x}^*$ *sobre el conjunto factible* Ω *si ocurre alguna de estas dos situaciones*

1. $\mathbf{x}^*$ *es un mínimo global del problema* PPNL *y el objetivo del problema es minimizar.*

2. $\mathbf{x}^*$ *es un máximo global del problema* PPNL *y el objetivo del problema es maximizar.*

Definición 9 (Valor óptimo) *Si* $\mathbf{x}^* \in \Omega \subseteq \mathbb{R}^n$ *es una solución óptima del problema* PPNL, *entonces se define el* valor óptimo *como el valor de la función objetivo en la solución óptima, es decir, si* $\mathbf{x}^*$ *es una solución óptima del problema* PPNL, *entonces* $f\left(\mathbf{x}^{`*}\right)$ *es el valor óptimo.*

Resolver un problema de optimización es encontrar, si existen, sus soluciones óptimas, es decir los extremos globales de la función objetivo sobre el conjunto factible. Desde el punto de vista práctico y computacional en algunas ocasiones bastará con obtener los llamados extremos locales que se definen a continuación.

Definición 10 (Mínimo local) *Consideremos el problema de optimización* PPNL *y sea* Ω *su conjunto factible. Diremos que* $\mathbf{x}^* \in \Omega \subseteq \mathbb{R}^n$ *es un* mínimo local o relativo *de* $f(\mathbf{x})$ *en* Ω *si y sólo si*

$$\exists\ \epsilon > 0 \textit{ tal que } \forall\ \overline{\mathbf{x}} \in \Omega; \quad \overline{\mathbf{x}} \neq \mathbf{x}^*; \quad \|\overline{\mathbf{x}} - \mathbf{x}^*\| < \epsilon \Longrightarrow f(\mathbf{x}^*) \leq f(\overline{\mathbf{x}})$$

El punto $\mathbf{x}^*$ *será un* mínimo local estricto *de* $f(x)$ *en* Ω *si la desigualdad es estricta*

$$\exists\ \epsilon > 0 \textit{ tal que } \forall\ \overline{\mathbf{x}} \in \Omega; \quad \overline{\mathbf{x}} \neq \mathbf{x}^*; \quad \|\overline{\mathbf{x}} - \mathbf{x}^*\| < \epsilon \Longrightarrow f(\mathbf{x}^*) < f(\overline{\mathbf{x}}).$$

Definición 11 (Máximo local) *Consideremos el problema general de optimización* PPNL *y sea* Ω *su conjunto factible. Diremos que* $\mathbf{x}^* \in \Omega \subseteq \mathbb{R}^n$ *es un* máximo local o relativo *de* $f(\mathbf{x})$ *en* Ω *si y sólo si*

$$\exists\ \epsilon > 0 \textit{ tal que } \forall\ \overline{\mathbf{x}} \in \Omega; \quad \overline{\mathbf{x}} \neq \mathbf{x}^*; \quad \|\overline{\mathbf{x}} - \mathbf{x}^*\| < \epsilon \Longrightarrow f(\mathbf{x}^*) \geq f(\overline{\mathbf{x}})$$

El punto $\mathbf{x}^*$ *será un* máximo local estricto *de* $f(x)$ *en* Ω *si la desigualdad es estricta*

$$\exists\ \epsilon > 0 \textit{ tal que } \forall\ \overline{\mathbf{x}} \in \Omega; \quad \overline{\mathbf{x}} \neq \mathbf{x}^*; \quad \|\overline{\mathbf{x}} - \mathbf{x}^*\| < \epsilon \Longrightarrow f(\mathbf{x}^*) > f(\overline{\mathbf{x}}).$$

Nota 5 *Los máximos y mínimos locales de los problemas de optimización también se denominan* extremos locales o relativos.

A diferencia de los extremos globales que afectan a todo el conjunto factible Ω, los extremos locales afectan a cierto entorno a su alrededor.

La teoría inicial asociada a la optimización está orientada a la obtención de condiciones necesarias y suficientes para que un punto sea óptimo. Como veremos en los temas correspondientes, esta teoría incluye el teorema de los multiplicadores de Lagrange y el teorema de Karush-Kuhn-Tucker. Por otra parte, también es interesante conocer no sólo si un punto es o no óptimo desde el punto de vista teórico, sino también cómo encontrar esos óptimos desde el punto de vista práctico. Teniendo esto en cuenta, al considerar problemas de optimización se plantean dos cuestiones:

1. **Cuestión estática:** ¿Cómo podemos determinar si un punto $\mathbf{x}^*$ es o no la solución óptima de un problema de optimización? ¿Qué condiciones deberían cumplir las funciones $f(\mathbf{x})$, $h_i(\mathbf{x})$ y $g_j(\mathbf{x})$ para que un problema *PPNL* tenga solución? ¿Qué condiciones debe cumplir el punto $\mathbf{x}^*$?

2. **Cuestión dinámica:** Si $\overline{\mathbf{x}}$ no es el punto óptimo, entonces ¿cómo podemos encontrar una solución óptima $\mathbf{x}^*$, utilizando la información de la función en $\overline{\mathbf{x}}$?

Mientras que con la primera cuestión se trata de determinar condiciones necesarias y/o suficientes para que un punto sea o no una solución óptima, en la segunda de las cuestiones se consideran los métodos numéricos adecuados para conseguir encontrar esas soluciones óptimas.

El resultado principal utilizado para conocer si un problema de optimización tiene solución es el *teorema de Weierstrass*, que recordamos a continuación dentro del contexto de la optimización matemática.

Teorema 32 (Teorema de Weierstrass) *Sea $f(\mathbf{x})$ una función continua definida sobre un conjunto compacto (cerrado y acotado) $K \subseteq \mathbb{R}^n$. Entonces el problema de optimización*

$$\left\{\begin{array}{ll} \textit{Optimizar} & f(\mathbf{x}) \\ \\ \textit{Sujeto a} & \mathbf{x} \in K \end{array}\right.$$

tiene al menos una solución para ambos objetivos de minimización y maximización, es decir

$$\exists \quad \mathbf{x}^*_{\text{mín}}, \mathbf{x}^*_{\text{máx}} \in K : \left\{\begin{array}{l} f(\mathbf{x}^*_{\text{mín}}) = \min\limits_{\mathbf{x}\in K} f(\mathbf{x}) \\ \\ f(\mathbf{x}^*_{\text{máx}}) = \max\limits_{\mathbf{x}\in K} f(\mathbf{x}) \end{array}\right.$$

Este es un resultado importante a tener en cuenta en la resolución de problemas de optimización, sin embargo el teorema no nos proporciona un método para la localización de las soluciones, solamente de su existencia en determinadas condiciones. Desde el punto de vista de las aplicaciones, lo interesante es caracterizar los puntos solución y diseñar un método efectivo para su cálculo.

Finalmente, apuntar que un problema de optimización puede tener solución única como en el siguiente planteamiento

$$\left\{\begin{array}{ll} \text{Optimizar} & x^2 \\ \\ \text{Sujeto a} & x \in [0,1] \end{array}\right. ,$$

no tener ninguna solución como en el problema

$$\left\{\begin{array}{ll} \text{Optimizar} & \dfrac{1}{x} \\ \\ \text{Sujeto a} & x \in (0,1) \end{array}\right. ,$$

o tener más de una solución, como en el problema

$$\left\{\begin{array}{ll} \text{Optimizar} & \operatorname{sen}(x) \\ \\ \text{Sujeto a} & x \in \mathbb{R} \end{array}\right.$$

en el que hay incluso infinitas soluciones óptimas, tanto de mínimo ($\frac{3\pi}{2} + 2k\pi,\ k \in \mathbb{Z}$) como de máximo ($\frac{\pi}{2} + 2k\pi,\ k \in \mathbb{Z}$).

5.3. Conjuntos Convexos

El concepto de convexidad es de gran importancia en el estudio de los problemas de optimización desde el punto de vista de la aplicación práctica, puesto que en algunos casos, bajo condiciones de convexidad, se puede garantizar que un extremo local de un problema es realmente un extremo global y por tanto la solución óptima del problema buscada.

Se describen en esta sección algunos conceptos básicos de convexidad útiles para el desarrollo de la programación matemática y aunque es posible definirlos en el ámbito de cualquier espacio topológico, en lo sucesivo consideraremos el espacio vectorial $\mathbb{R}^n$.

Definición 12 (Segmento lineal) *Dados dos puntos* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ *el* segmento lineal cerrado *que une* $\mathbf{x}$ *con* $\mathbf{y}$ *es el conjunto definido por*

$$[\mathbf{x}, \mathbf{y}] = \{\lambda \mathbf{x} + (1-\lambda)\, \mathbf{y} \in \mathbb{R}^n : 0 \leq \lambda \leq 1\}$$

Del mismo modo, el segmento lineal abierto *que une* $\mathbf{x}$ *con* $\mathbf{y}$ *es el conjunto*

$$(\mathbf{x}, \mathbf{y}) = \{\lambda \mathbf{x} + (1-\lambda)\, \mathbf{y} \in \mathbb{R}^n : 0 < \lambda < 1\}$$

Definición 13 (Conjunto convexo) *Sea* $\Omega \subset \mathbb{R}^n$ *entonces*

$$\Omega \text{ es convexo } \Longleftrightarrow \forall\ \mathbf{x}, \mathbf{y} \in \Omega \Rightarrow [\mathbf{x}, \mathbf{y}] \subseteq \Omega$$

Esta definición se interpreta de forma que un conjunto será convexo si el segmento lineal cerrado que une cualquier par de puntos del conjunto está contenido en dicho conjunto. La figura 5.2 representa algunos conjuntos convexos y otros no convexos de $\mathbb{R}^2$.

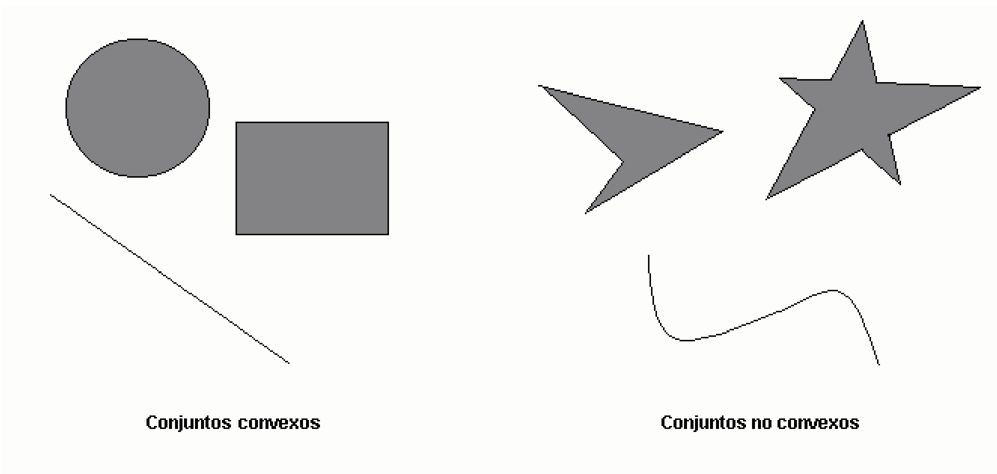


Figura 5.2: Convexidad en $\mathbb{R}^2$.

Por convenio el conjunto vacío $\emptyset$ es un conjunto convexo.

Unos de los tipos más importantes de conjuntos convexos son los *hiperplanos y los semiespacios*, que definimos a continuación.

Definición 14 (Hiperplano) *Sea* $\mathbf{a} \in \mathbb{R}^n$ *con* $\mathbf{a} \neq 0$ *y* $b \in \mathbb{R}$. *Un* hiperplano H *en* $\mathbb{R}^n$ *es un conjunto definido como*

$$H = \left\{\mathbf{x} \in \mathbb{R}^n : \ \mathbf{a}^T\mathbf{x} = b\right\}$$

El vector $\mathbf{a}$ *es el llamado vector* normal *al hiperplano. La expresión* $\mathbf{a}^T\mathbf{x}$ *es el producto escalar de ambos vectores*

$$\mathbf{a}^T\mathbf{x} = \ a_1x_1 + \cdots + a_nx_n$$

Definición 15 (Semiespacios) *Sea* $\mathbf{a} \in \mathbb{R}^n$ *con* $\mathbf{a} \neq \mathbf{0}$ *y* $b \in \mathbb{R}$. *Sea* H *el hiperplano construido a partir de* $\mathbf{a}$ *y* b *entonces definimos los* semiespacios cerrados positivos *y* negativos *asociados a* H, *respectivamente a los conjuntos*

$$H_+ \ = \ \left\{\mathbf{x} \in \mathbb{R}^n : \ \mathbf{a}^T\mathbf{x} \geq b\right\}$$

$$H_- \ = \ \left\{\mathbf{x} \in \mathbb{R}^n : \ \mathbf{a}^T\mathbf{x} \leq b\right\}$$

y semiespacios abiertos positivos *y* negativos *asociados a* H *a los conjuntos definidos como*

$$\overset{\circ}{H}_+ \ = \ \left\{\mathbf{x} \in \mathbb{R}^n : \ \mathbf{a}^T\mathbf{x} > b\right\}$$

$$\overset{\circ}{H}_- \ = \ \left\{\mathbf{x} \in \mathbb{R}^n| : \ \mathbf{a}^T\mathbf{x} < b\right\}$$

El siguiente lema es una consecuencia inmediata de la definición de convexidad y establece que la intersección de dos conjuntos convexos es convexa y que la suma algebráica de dos conjuntos convexos también es convexa. Su demostración es muy sencilla utilizando la propia definición de conjunto convexo y se deja como ejercicio.

Lema 33 *Sean* $\Omega_1, \Omega_2 \subseteq \mathbb{R}^n$ *dos conjuntos convexos entonces*

1. $\Omega_1 \cap \Omega_2$ *es convexo.*
2. $\Omega_1 + \Omega_2 = \{\mathbf{x} + \mathbf{y} \in \mathbb{R}^n : \mathbf{x} \in \Omega_1, \ \mathbf{y} \in \Omega_2\}$ *es convexo.*
3. $\Omega_1 - \Omega_2 = \{\mathbf{x} - \mathbf{y} \in \mathbb{R}^n : \mathbf{x} \in \Omega_1, \ \mathbf{y} \in \Omega_2\}$ *es convexo.*

Nota 6 *Es posible extender mediante inducción la propiedad* 1 *del lema 33 a una intersección cualquiera de conjuntos convexos, es decir si* $\{\Omega_n\}_{n=1}^{\infty} \subseteq \mathbb{R}^n$ *es una familia de conjuntos convexos de* $\mathbb{R}^n$, *entonces el conjunto intersección* $\bigcap_{n=1}^{\infty} \Omega_n$ *es de nuevo un conjunto convexo.*

Si tenemos en cuenta por una parte que los semiespacios de $\mathbb{R}^n$ son conjuntos convexos y por otra utilizamos los resultados del lema 33 se deduce fácilmente que los conjuntos definidos de la forma

$$\Omega = \left\{ (x_1, \ldots, x_n) \in \mathbb{R}^n : \quad \begin{array}{c} a_{11}x_1 + \cdots + a_{1n}x_n \leq b_1 \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \leq b_m \end{array} \right\}$$

o en de forma más compacta en notación matricial

$$\Omega = \{\mathbf{x} \in \mathbb{R}^n : \quad \mathbf{A}\mathbf{x} \leq \mathbf{b}\}$$

donde $\mathbf{A} = (a_{ij})$ es una matriz $m \times n$ en $\mathbb{R}$ y $\mathbf{b} \in \mathbb{R}^m$ es un vector; son convexos puesto que

$$\Omega = \Omega_1 \cap \cdots \cap \Omega_m$$

siendo

$$\Omega_k = \{(x_1, \ldots, x_n) \in \mathbb{R}^n : \ a_{k1}x_1 + \cdots + a_{kn}x_n \leq b_k\}\,; \quad k = 1, \ldots, m$$

semiespacios cerrados negativos, que son conjuntos convexos.

Ejemplo 9 *El conjunto definido como* $S = \{\mathbf{x} \in \mathbb{R}^n$: $\mathbf{x}$ *es solución óptima de* $P\}$ *siendo* P *el problema*

$$P = \begin{cases} \textit{Minimizar} & \mathbf{c}^T\mathbf{x} \\ \\ \textit{Sujeto a} & \mathbf{A}\mathbf{x} = \mathbf{b} \\ \\ & \mathbf{x} \geq 0 \end{cases}$$

es un conjunto convexo.

Solución: Para demostrar la convexidad de S utilizaremos la definición. Distinguimos tres caso:

1. Caso I: El problema P no tiene solución, en ese caso S es el conjunto vacío $\emptyset$, que por definición es un conjunto convexo.

2. Caso II: El problema P tiene solución única: $S = \{\mathbf{x}^*\}$. En este caso el único segmento lineal que se puede construir es
$$\lambda\mathbf{x}^* + (1-\lambda)\,\mathbf{x}^* = \mathbf{x}^* \in S$$
(en este caso incluso independientemente de si λ está en el intervalo $[0, 1]$ o no).

3. Caso III: El problema P tiene más de una solución. En este caso S tiene más de un elemento y podemos suponer que existen $\mathbf{x}^1$ y $\mathbf{x}^2 \in S$ que por ser soluciones del problema P serán puntos factibles y por tanto cumplirán las restricciones del problema
$$\begin{array}{c} \mathbf{A}\mathbf{x}^1 = \mathbf{b} \\ \mathbf{A}\mathbf{x}^2 = \mathbf{b} \\ \\ \mathbf{x}^1, \mathbf{x}^2 \geq 0 \end{array}$$
Como además son mínimos de P; si llamamos f^* al valor óptimo se tiene
$$f^* = \mathbf{c}^T\mathbf{x}^1 = \mathbf{c}^T\mathbf{x}^2$$
Ahora hay que comprobar que los elementos del segmento lineal que une $\mathbf{x}^1$ con $\mathbf{x}^2$ también están en S es decir son soluciones de P. Para ello tomamos $\lambda \in [0, 1]$ y consideremos el punto $\mathbf{x}^3$ definido por
$$\mathbf{x}^3 = \lambda\mathbf{x}^1 + (1-\lambda)\,\mathbf{x}^2$$

Comprobamos su factibilidad operando directamente. Por una parte $\mathbf{x}^3$ es factible ya que por una parte cumple el sistema de ecuaciones

$$\mathbf{A}\mathbf{x}^3 = \mathbf{A}\left(\lambda\mathbf{x}^1 + (1-\lambda)\,\mathbf{x}^2\right) = \lambda\mathbf{A}\mathbf{x}^1 + (1-\lambda)\,\mathbf{A}\mathbf{x}^2 = \lambda\mathbf{b} + (1-\lambda)\,\mathbf{b} = \mathbf{b},$$

y por otra parte, como $\lambda \in [0,1]$

$$0 \leq \lambda \leq 1 \Rightarrow \left\{\begin{array}{c} \lambda \geq 0 \\ \\ 1-\lambda \geq 0 \end{array}\right.$$

de donde al ser $\mathbf{x}^1, \mathbf{x}^2 \geq 0$ se obtiene

$$\mathbf{x}^3 = \lambda\mathbf{x}^1 + (1-\lambda)\,\mathbf{x}^2 \geq 0$$

ya que todos los sumandos son positivos. De esta forma $\mathbf{x}^3$ es factible porque cumple las restricciones del problema.

Si ahora evaluamos la función objetivo en $\mathbf{x}^3$

$$f\left(\mathbf{x}^3\right) = \mathbf{c}^T\mathbf{x}^3 = \mathbf{c}^T\left(\lambda\mathbf{x}^1 + (1-\lambda)\,\mathbf{x}^2\right) = \lambda\mathbf{c}^T\mathbf{x}^1 + (1-\lambda)\,\mathbf{c}^T\mathbf{x}^2 = \lambda f^* + (1-\lambda)\,f^* = f^*$$

y por tanto se deduce que $\mathbf{x}^3$ también es solución del problema P.

Definición 16 (Punto extremo) *Sea $\Omega \subseteq \mathbb{R}^n$ un conjunto convexo no vacío. Se dice que el punto $\mathbf{x} \in \Omega$ es un* vértice o punto extremo *de $\Omega \Longleftrightarrow$*

$$\textit{Si } \mathbf{x} = \lambda\mathbf{x}^1 + (1-\lambda)\mathbf{x}^2 \textit{ para algún } \lambda \in [0,1] \;\; \textit{y } \mathbf{x}^1, \mathbf{x}^2 \in \Omega \Longrightarrow \mathbf{x}^1 = \mathbf{x}^2 = \mathbf{x}$$

Por ejemplo el conjunto convexo

$$\Omega = \left\{\mathbf{x} = (x_1, x_2) \in \mathbb{R}^2 \middle| \quad 0 \leq x_1 \leq 2, 0 \leq x_2 \leq 2\right\}$$

tiene 4 puntos extremos dados por $P_1 = (0,0)$, $P_2 = (0,2)$, $P_3 = (2,0)$ y $P_4 = (2,2)$ (ver figura 5.3).

El conjunto de los puntos extremos de un conjunto convexo puede ser vacío, por ejemplo una bola abierta de $\mathbb{R}^n$, contener una cantidad finita de elementos, como en la figura 5.3 o tener una cantidad infinita de elementos, como una bola cerrada de $\mathbb{R}^n$.

5.4. Funciones convexas

En esta sección se proporciona la definición de función convexa y se presentan alguna de sus propiedades más importantes, sobre todo aquellas que pueden utilizarse para resolver problemas de optimización.

Definición 17 (Función convexa) *Diremos que la función $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$, con Ω un conjunto convexo no vacío, es* convexa *sobre $\Omega \Longleftrightarrow$*

$$\forall \mathbf{x}, \mathbf{y} \in \Omega, \lambda \in [0,1] \Rightarrow f\left(\lambda\mathbf{x} + (1-\lambda)\,\mathbf{y}\right) \leq \lambda f\left(\mathbf{x}\right) + (1-\lambda)\,f\left(\mathbf{y}\right)$$

Se dice que f estrictamente convexa $\Longleftrightarrow$

$$\forall \mathbf{x}, \mathbf{y} \in \Omega, \lambda \in [0,1] \Rightarrow f\left(\lambda\mathbf{x} + (1-\lambda)\,\mathbf{y}\right) < \lambda f\left(\mathbf{x}\right) + (1-\lambda)\,f\left(\mathbf{y}\right)$$

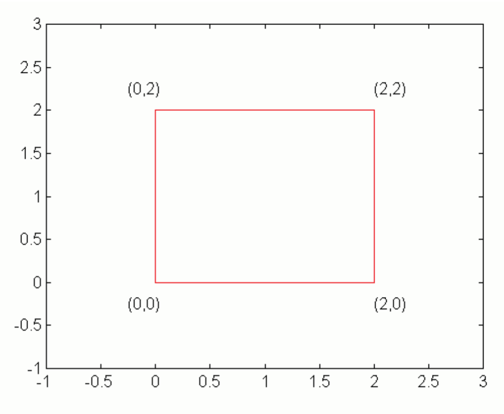


Figura 5.3: Puntos extremos.

Nota 7 *Si definimos $\mu = 1 - \lambda$, entonces la definición puede ponerse como*

$$f \text{ es convexa} \Leftrightarrow f(\lambda \mathbf{x} + \mu \mathbf{y}) \leq \lambda f(\mathbf{x}) + \mu f(\mathbf{y}), \quad \forall \lambda, \mu \geq 0 \quad \text{con } \lambda + \mu = 1$$

Y del mismo modo para función estrictamente convexa.

Definición 18 (Función cóncava) *Diremos que la función $f : \Omega \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$, con Ω conjunto convexo no vacío es* cóncava *sobre $\Omega \Longleftrightarrow g = -f$ es convexa.*

Esta definición equivale a decir que

$$f \text{ es cóncava} \Leftrightarrow \forall \mathbf{x}, \mathbf{y} \in \Omega, \lambda \in [0,1] \Rightarrow f(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}) \geq \lambda f(\mathbf{x}) + (1-\lambda) f(\mathbf{y})$$

Definición 19 *Se dice que f* estrictamente cóncava *$\Longleftrightarrow g = -f$ es estrictamente convexa.*

Proposición 34 *Si $f_1(\mathbf{x})$ y $f_2(\mathbf{x})$ son dos funciones convexas definidas sobre un conjunto convexo no vacío $\Omega \subseteq \mathbb{R}^n \Longrightarrow f(\mathbf{x}) = (f_1 + f_2)(\mathbf{x})$ es convexa sobre Ω.*

Demostración: Por definición de $f = f_1 + f_2$ se tiene

$$f(\mathbf{x}) = (f_1 + f_2)(\mathbf{x}) = f_1(\mathbf{x}) + f_2(\mathbf{x})$$

Sean $\mathbf{x}, \mathbf{y} \in \Omega$ y $\lambda \in [0,1]$

$$f(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}) = f_1(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}) + f_2(\lambda \mathbf{x} + (1-\lambda)\mathbf{y})$$

Puesto que tanto f_1 como f_2 son convexas sobre Ω se cumplen de forma simultanea las siguientes desigualdades

$$f_1(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}) \leq \lambda f_1(\mathbf{x}) + (1-\lambda) f_1(\mathbf{y})$$

$$f_2(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}) \leq \lambda f_2(\mathbf{x}) + (1-\lambda) f_2(\mathbf{y})$$

y sumando ambas expresiones

$$f_1(\lambda\mathbf{x} + (1-\lambda)\mathbf{y}) + f_2(\lambda\mathbf{x} + (1-\lambda)\mathbf{y}) \leq \lambda(f_1(\mathbf{x}) + f_2(\mathbf{x})) + (1-\lambda)(f_1(\mathbf{y}) + f_2(\mathbf{y}))$$

es decir

$$f(\lambda\mathbf{x} + (1-\lambda)\mathbf{y}) \leq \lambda f(\mathbf{x}) + (1-\lambda) f(\mathbf{y})$$

por tanto $f(\mathbf{x})$ es convexa.

Proposición 35 *Si $f(\mathbf{x})$ es convexa sobre $\Omega \subseteq \mathbb{R}^n$, siendo Ω un conjunto convexo no vacío, entonces $\forall\alpha \geq 0$, la función $(\alpha f)(\mathbf{x})$ definida por*

$$(\alpha f)(\mathbf{x}) = \alpha f(\mathbf{x})$$

es convexa sobre Ω.

Demostración: Como $f(\mathbf{x})$ es convexa sobre Ω se cumple

$$f(\lambda\mathbf{x} + (1-\lambda)\mathbf{y}) \leq \lambda f(\mathbf{x}) + (1-\lambda) f(\mathbf{y})$$

si ahora multiplicamos ambos lados de la desigualdad por $\alpha > 0$, el sentido de esta no cambia y tendremos

$$\alpha(f(\lambda\mathbf{x} + (1-\lambda)\mathbf{y})) \leq \alpha(\lambda f(\mathbf{x}) + (1-\lambda) f(\mathbf{y})) = \lambda\alpha f(\mathbf{x}) + (1-\lambda)\alpha f(\mathbf{x})$$

lo que demuestra la convexidad de αf.

Nota 8 *Si $\alpha < 0$, entonces la función αf sería cóncava sobre Ω.*

Proposición 36 *Sea f convexa en $\Omega \subseteq \mathbb{R}^n$, con Ω conjunto convexo no vacío y sea $c \in \mathbb{R}$, entonces el conjunto de nivel Γ_c definido por*

$$\Gamma_c = \{\mathbf{x} \in \Omega : f(\mathbf{x}) \leq c\}$$

es convexo $\forall c \in \mathbb{R}$.

Demostración: Hay que demostrar que $\forall\, \mathbf{x}, \mathbf{y} \in \Gamma_c$ y $\lambda \in [0,1]$, debe ocurrir $\lambda\mathbf{x} + (1-\lambda)\mathbf{y} \in \Gamma_c$. Utilizando la definición de Γ_c es equivalente a probar

$$\forall\, \mathbf{x}, \mathbf{y} \text{ con } f(\mathbf{x}) \leq c \text{ y } f(\mathbf{y}) \leq c \implies f(\lambda\mathbf{x} + (1-\lambda)\mathbf{y}) \leq c \quad \forall\lambda \in [0,1]$$

Como $f(\mathbf{x})$ es convexa se cumple

$$f(\lambda\mathbf{x} + (1-\lambda)\mathbf{y}) \leq \lambda f(\mathbf{x}) + (1-\lambda) f(\mathbf{y})$$

Por otra parte como $\lambda \in [0,1]$ se cumple

$$0 \leq \lambda \leq 1 \Longrightarrow \begin{cases} 0 \leq \lambda \\ \\ \lambda \leq 1 \Rightarrow 0 \leq 1-\lambda \end{cases}$$

y por tanto

$$f(\mathbf{x}) \leq c \Longrightarrow \lambda f(\mathbf{x}) \leq \lambda c$$

$$f(\mathbf{y}) \leq c \Longrightarrow (1-\lambda) f(\mathbf{x}) \leq (1-\lambda) c$$

Al sumar ambas desigualdades se obtiene

$$\lambda f(\mathbf{x}) + (1-\lambda) f(\mathbf{x}) \leq \lambda c + (1-\lambda) c = c$$

y por tanto

$$f(\lambda \mathbf{x} + (1-\lambda) \mathbf{y}) \leq \lambda f(\mathbf{x}) + (1-\lambda) f(\mathbf{x}) \leq c$$

como tratábamos de probar. ■

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. Los siguientes resultados proporcionan esas caracterizaciones para funciones convexas diferenciables en términos del gradiente y del Hessiano.

Proposición 37 (Caracterización de primer orden para funciones convexas) *Sea $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$, con Ω un conjunto convexo y $f(\mathbf{x}) \in \mathcal{C}^1(\Omega)$, es decir una función derivable en Ω con derivadas parciales continuas, entonces*

$$f(\mathbf{x}) \text{ es convexa sobre } \Omega \Longleftrightarrow f(\mathbf{y}) \geq f(\mathbf{x}) + \nabla f(\mathbf{x})(\mathbf{y}-\mathbf{x}) \quad \forall \mathbf{x}, \mathbf{y} \in \Omega$$

$$f(\mathbf{x}) \text{ es estrictamente convexa sobre } \Omega \Longleftrightarrow f(\mathbf{y}) > f(\mathbf{x}) + \nabla f(\mathbf{x})(\mathbf{y}-\mathbf{x}) \quad \forall \mathbf{x}, \mathbf{y} \in \Omega$$

Demostración: Demostramos la proposición para el caso en el que $f(\mathbf{x})$ sea convexa y se procedería del mismo modo si $f(\mathbf{x})$ fuera estrictamente convexa.

"$\Longrightarrow$" Supongamos en primer lugar que $f(\mathbf{x})$ es convexa, entonces tomando cualquier valor $\lambda \in [0,1]$ y $\mathbf{x}, \mathbf{y} \in \Omega$, se tiene

$$f(\lambda \mathbf{y} + (1-\lambda)\mathbf{x}) \leq \lambda f(\mathbf{y}) + (1-\lambda) f(\mathbf{x})$$

Se puede asumir el hecho de que $\lambda \neq 0$ y podremos entonces dividir la inecuación por este valor

$$\frac{f(\mathbf{x} + \lambda(\mathbf{y}-\mathbf{x})) - f(\mathbf{x})}{\lambda} \leq f(\mathbf{y}) - f(\mathbf{x}) \quad \forall \lambda \in (0,1]$$

Si ahora $\lambda \to 0$ y tenemos en cuenta que la función $f(\mathbf{x})$ es derivable en Ω, entonces

$$\lim_{\lambda \to 0} \frac{f(\mathbf{x} + \lambda(\mathbf{y}-\mathbf{x})) - f(\mathbf{x})}{\lambda} = \nabla f(\mathbf{x})(\mathbf{y}-\mathbf{x})$$

y por tanto

$$\nabla f(\mathbf{x})(\mathbf{y}-\mathbf{x}) \leq f(\mathbf{y}) - f(\mathbf{x})$$

"$\Longleftarrow$" Supongamos ahora que es cierta la siguiente desigualdad

$$\nabla f(\mathbf{x})(\mathbf{y}-\mathbf{x}) \le f(\mathbf{y}) - f(\mathbf{x}) \ \forall \mathbf{x}, \mathbf{y} \in \Omega$$

entonces por ser Ω un conjunto convexo

$$\forall \ \mathbf{x}^1, \mathbf{x}^2 \in \Omega \text{ y } \lambda \in [0,1] \Longrightarrow \mathbf{x} = \lambda \mathbf{x}^1 + (1-\lambda)\mathbf{x}^2 \in \Omega$$

Si utilizamos la desigualdad para $\mathbf{x}$ e $\mathbf{y} = \mathbf{x}^1$ por una parte y para $\mathbf{x}$ e $\mathbf{y} = \mathbf{x}^2$ por otra obtenemos

$$\begin{aligned} \mathbf{y} = \mathbf{x}^1 \quad &\Longrightarrow \quad f(\mathbf{x}^1) \ge f(\mathbf{x}) + \nabla f(\mathbf{x})(\mathbf{x}^1 - \mathbf{x}) \\ \mathbf{y} = \mathbf{x}^2 \quad &\Longrightarrow \quad f(\mathbf{x}^2) \ge f(\mathbf{x}) + \nabla f(\mathbf{x})(\mathbf{x}^2 - \mathbf{x}) \end{aligned}$$

Si multiplicamos por λ la primera desigualdad y por $(1-\lambda)$ la segunda y sumamos ambas expresiones

$$\begin{array}{rcl} \lambda f\left(\mathbf{x}^1\right) & \ge & \lambda f(\mathbf{x}) + \lambda \nabla f(\mathbf{x})(\mathbf{x}^1 - \mathbf{x}) \\ (1-\lambda)\, f\left(\mathbf{x}^2\right) & \ge & (1-\lambda) f(\mathbf{x}) + (1-\lambda)\nabla f(\mathbf{x})(\mathbf{x}^2 - \mathbf{x}) \\ \hline \lambda f(\mathbf{x}^1) + (1-\lambda) f(\mathbf{x}^2) & \ge & f(\mathbf{x}) + \nabla f(\mathbf{x})(\lambda \mathbf{x}^1 + (1-\lambda)\mathbf{x}^2 - \mathbf{x}) \\ & \ge & f(\mathbf{x}) = f(\lambda \mathbf{x}^1 + (1-\lambda)\mathbf{x}^2) \end{array}$$

lo que demuestra la convexidad de $f(\mathbf{x})$. ■

Proposición 38 (Caracterización de segundo orden para funciones convexas) *Sea $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$, con Ω un conjunto abierto convexo no vacío y $f(\mathbf{x}) \in \mathcal{C}^2(\Omega)$ entonces:*

$$f(\mathbf{x}) \ \text{ es convexa en } \Omega \Longleftrightarrow \mathbf{H}f(\mathbf{x}) \ \text{ es semidefinido positivo } \forall \mathbf{x} \in \Omega$$

siendo

$$\mathbf{H}f(\mathbf{x}) = \left[\left(\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}) \right)_{i,j=1}^{n} \right]$$

la matriz hessiana asociada a $f(\mathbf{x})$.

Demostración: Demostramos la doble implicación.

"$\Longrightarrow$" Supongamos que $f(\mathbf{x})$ es convexa en Ω y sea $\mathbf{x} \in \Omega$. Demostrar que $\mathbf{H}f(\mathbf{x})$ es semidefinida positiva en Ω equivale a demostrar que $\forall \ \mathbf{d} \in \mathbb{R}^n$ ocurre $\mathbf{d}^T \mathbf{H} f(\mathbf{x}) \mathbf{d} \ge 0$.

Puesto que $\Omega \neq \emptyset$ y abierto, entonces para $\mathbf{d} \in \mathbb{R}^n$, ocurre $\mathbf{y} = \mathbf{x} + \lambda \mathbf{d} \in \Omega$ tomando $|\lambda| \neq 0$ y suficientemente pequeño. Si utilizamos la proposición 37 de caracterización de primer orden de funciones convexas por una parte y el teorema de Taylor por otra, se obtienen las siguientes expresiones:

$$f(\mathbf{y}) \ge f(\mathbf{x}) + \nabla f(\mathbf{x})\,(\mathbf{y}-\mathbf{x}) \Leftrightarrow f(\mathbf{x} + \lambda \mathbf{d}) \ge f(\mathbf{x}) + \lambda \nabla f(\mathbf{x}) \mathbf{d}$$

$$f(\mathbf{x} + \lambda \mathbf{d}) = f(\mathbf{x}) + \lambda \nabla f(\mathbf{x})^T \mathbf{d} + \tfrac{1}{2}\lambda^2 \mathbf{d}^T \mathbf{H} f(\mathbf{x}) \mathbf{d} + o(\lambda)$$

donde $o(\lambda) = \lambda^2||\mathbf{d}||^2\alpha(\mathbf{x}, \lambda\mathbf{d})$ es una función que converge a 0 cuando λ tiende a 0.

Si se restan ambas expresiones obtenemos

$$\frac{1}{2}\lambda^2\mathbf{d}^T\mathbf{H}f(\mathbf{x})\mathbf{d} + \lambda^2||\mathbf{d}||^2\alpha(\mathbf{x}, \lambda\mathbf{d}) \geq 0$$

A continuación se divide por λ^2 y se toman límites cuando $\lambda \to 0$. El resultado que se obtiene es que $\mathbf{d}^T\mathbf{H}f(\mathbf{x})\mathbf{d} \geq 0$.

"$\Longleftarrow$" Recíprocamente si suponemos ahora que la matriz Hessiana $\mathbf{H}f(\mathbf{x})$ es semidefinida positiva en cada punto de Ω y consideramos $\mathbf{x}$ e $\mathbf{y}$ en Ω, entonces por el teorema del valor medio obtenemos:

$$f(\mathbf{y}) = f(\mathbf{x}) + \nabla f(\mathbf{x})^T(\mathbf{y} - \mathbf{x}) + \frac{1}{2}(\mathbf{y} - \mathbf{x})^T\mathbf{H}f(\mathbf{x}^*)(\mathbf{y} - \mathbf{x})$$

donde $\mathbf{x}^* = \lambda\mathbf{x} + (1 - \lambda)\mathbf{y} \in \Omega$ con $\lambda \in (0, 1)$. Como $\mathbf{H}f(\mathbf{x}^*)$ semidefinida positiva, se cumple que $(\mathbf{y} - \mathbf{x})^T\mathbf{H}f(\mathbf{x}^*)(\mathbf{y} - \mathbf{x}) \geq 0$ y concluimos que

$$f(\mathbf{y}) = f(\mathbf{x}) + \nabla f(\mathbf{x})^T(\mathbf{y} - \mathbf{x}) + \frac{1}{2}(\mathbf{y} - \mathbf{x})^T\mathbf{H}f(\mathbf{x}^*)(\mathbf{y} - \mathbf{x}) \geq f(\mathbf{x}) + \nabla f(\mathbf{x})^T(\mathbf{y} - \mathbf{x})$$

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \lambda\nabla f(\mathbf{x})\mathbf{x} \quad \forall\mathbf{x}^*, \mathbf{x} \in \Omega$$

que resulta ser la caracterización de primer orden para funciones convexas de la proposición 37 y por tanto $f(x)$ es convexa en Ω $\blacksquare$.

La matriz Hessiana de f es la generalización al espacio $\mathbb{R}^n$ del concepto de curvatura de una función y de forma análoga, la definición positiva del Hessiano es la generalización de curvatura positiva. Las funciones convexas tienen curvatura positiva (o al menos no negativa) en todas las direcciones.

5.5. Optimización de funciones convexas

En esta sección consideraremos el problema de minimizar y maximizar una función convexa sobre un conjunto convexo.

Debido a la correspondencia entre funciones cóncavas y convexas todos los resultados se presentan de forma equivalente para ambos tipos de funciones, por ello en cada teorema y entre corchetes se muestra el resultado alternativo.

Teorema 39 *Sea $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$ con Ω un conjunto convexo. Si $f(\mathbf{x})$ es convexa [cóncava] $\Longrightarrow$ El conjunto Γ donde $f(\mathbf{x})$ alcanza su mínimo [máximo] es convexo y cualquier mínimo [máximo] local de $f(\mathbf{x})$ es mínimo [máximo] global.*

Demostración: Definimos el conjunto Γ como

$$\Gamma = \{\mathbf{x}^* \in \Omega | \ f(\mathbf{x}^*) = \min_{x\in\Omega} f(x)\}$$

Si $\Gamma = \emptyset$, es decir si $f(\mathbf{x})$ no tiene mínimos relativos, entonces el teorema se cumple por la convexidad del conjunto vacío.

Supongamos ahora $\Gamma \neq \emptyset$ y por tanto existe $\mathbf{x}^* \in \Gamma$. Si $\mathbf{x}^*$ es el único elemento de Γ, entonces es convexo puesto que la única combinación con elementos de Γ que se puede hacer $\forall \lambda \in [0,1]$ es $\lambda \mathbf{x}^* + (1-\lambda)\,\mathbf{x}^* = \mathbf{x}^* \in \Gamma$.

Finalmente podemos suponer que Γ tiene al menos dos elementos $\mathbf{x}^*$, $\mathbf{y}^* \in \Gamma$. Por la definición de Γ, entonces

$$f_{\text{mín}} = \text{mín}_{x\in\Omega}\, f\,(x) = f\,(\mathbf{x}^*) = f\,(\mathbf{y}^*)$$

Si tomamos $\lambda \in [0,1]$, por convexidad de $f\,(\mathbf{x})$ se cumple

$$f\,(\lambda \mathbf{x}^* + (1-\lambda)\,\mathbf{y}^*) \leq \lambda f\,(\mathbf{x}^*) + (1-\lambda)\,f\,(\mathbf{y}^*) = \lambda f_{\text{mín}} + (1-\lambda)\,f_{\text{mín}} = f_{\text{mín}}$$

pero como $\lambda \mathbf{x}^* + (1-\lambda)\,\mathbf{y}^* \in \Omega$ y $f_{\text{mín}} = \text{mín}_{x\in\Omega}\, f\,(x)$, tendrá que ocurrir

$$f_{\text{mín}} \leq f\,(\lambda \mathbf{x}^* + (1-\lambda)\,\mathbf{y}^*)$$

de donde

$$f_{\text{mín}} \leq f\,(\lambda \mathbf{x}^* + (1-\lambda)\,\mathbf{y}^*) \leq f_{\text{mín}}$$

y por tanto

$$f\,(\lambda \mathbf{x}^* + (1-\lambda)\,\mathbf{y}^*) = f_{\text{mín}}$$

y se obtiene el resultado pedido.

Supongamos ahora que $\mathbf{x}^* \in \Omega$ es un mínimo local de $f\,(\mathbf{x})$ que no es global; es decir, que $\exists\, \mathbf{y} \in \Omega$ con $f(\mathbf{y}) < f(\mathbf{x}^*)$. Por la convexidad de $f\,(\mathbf{x})$ en el conjunto convexo Ω se obtiene la siguiente desigualdad

$$f(\lambda \mathbf{y} + (1-\lambda)\mathbf{x}^*) \leq \lambda f(\mathbf{y}) + (1-\lambda) f(\mathbf{x}^*)$$

para $\lambda \in [0,1]$. Utilizando ahora el hecho de que $f(\mathbf{y}) < f(\mathbf{x}^*)$ y como $\lambda \in [0,1]$ y por tanto tanto λ como $(1-\lambda)$ son no negativos

$$f(\lambda \mathbf{y} + (1-\lambda)\mathbf{x}^*) \leq \lambda f(\mathbf{x}^*) + (1-\lambda) f(\mathbf{x}^*) = f(\mathbf{x}^*)$$

Para un valor de λ suficientemente pequeño se contradice el hecho de que $\mathbf{x}^*$ sea un mínimo relativo, como estamos asuminedo que $\mathbf{x}^*$ es mínimo local, entonces debe cumplirse que además es global. ■.

Teorema 40 *Sea $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$ con Ω un conjunto convexo con $f \in \mathcal{C}^1(\Omega)$ una función convexa [cóncava]. Supongamos que existe un $\mathbf{x}^* \in \Omega$ que cumple la siguiente propiedad*

$$\forall \mathbf{y} \in \Omega \Rightarrow \nabla f(\mathbf{x}^*)^T(\mathbf{y} - \mathbf{x}^*) \geq 0 \quad \left[\nabla f(\mathbf{x}^*)^T(\mathbf{y} - \mathbf{x}^*) \leq 0\right]$$

Entonces $\mathbf{x}^*$ es un punto de mínimo [máximo] global de f en Ω.

Demostración: Como Ω convexo, $\mathbf{x}^* + \lambda(\mathbf{y} - \mathbf{x}^*) \in \Omega$ con $\lambda \in [0,1]$. Por la caracterización de primer orden de funciones convexas (proposición 37) y utilizando el hecho de que para $\mathbf{x}^*$ ocurre $\nabla f(\mathbf{x}^*)^T(\mathbf{y} - \mathbf{x}^*) \geq 0$ tenemos:

$$f(\mathbf{y}) \geq f(\mathbf{x}^*) + \nabla f(\mathbf{x}^*)^T(\mathbf{y} - \mathbf{x}^*) \geq f(\mathbf{x}^*)$$

Luego $\mathbf{x}^*$ es un punto de mínimo relativo de f en Ω y utilizando el resultado de la proposición anterior, también es un mínimo global ■.

Se incluyen a continuación, sin demostración, una serie de resultados muy útiles e interesantes para la resolución de problemas de optimización.

Proposición 41 *Si Ω es un subconjunto no vacío, convexo y compacto de $\mathbb{R}^n$ cuyo conjunto de puntos extremos es finito y si $f : \Omega \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ es un función convexa [cóncava] sobre $\Omega \Rightarrow f(\mathbf{x})$ posee en Ω un máximo [mínimo] global y se encuentra en uno de sus puntos extremos.*

Teorema 42 *Sea $f : \Omega \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ convexa [cóncava] en Ω convexo y compacto $\Longrightarrow$ Si $f(\mathbf{x})$ tiene máximo [mínimo] en Ω entonces lo alcanza en un punto extremo de Ω.*

Lema 43 *Si $\mathbf{x}^*, \mathbf{y}^*$ son dos puntos de mínimo [máximo] global de una función convexa [cóncava] $f(\mathbf{x})$, $f : \Omega \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ con Ω convexo $\Longrightarrow \forall\lambda \in [0,1]$ los puntos definidos como $\mathbf{z}^*(\lambda) = \lambda\mathbf{x}^* + (1-\lambda)\mathbf{y}^*$ también son mínimos [máximos] globales de $f(\mathbf{x})$.*

Demostración: Supongamos que f es convexa (para el caso de f cóncava y máximo la demostración sería equivalente), entonces

$$f(\mathbf{z}^*(\lambda)) = f(\lambda\mathbf{x}^* + (1-\lambda)\mathbf{y}^*) = f(\lambda\mathbf{x}^* + (1-\lambda)\mathbf{y}^*) \leq \lambda f(\mathbf{x}^*) + (1-\lambda) f(\mathbf{y}^*)$$

Como $\mathbf{x}^*, \mathbf{y}^*$ son mínimos globales $\Longrightarrow f(\mathbf{y}^*) = f(\mathbf{x}^*)$ y resulta

$$f(\mathbf{z}^*(\lambda)) \leq f(\mathbf{x}^*)$$

pero como $\mathbf{x}^*$ es el mínimo global de $f(\mathbf{x})$ sobre Ω, que es un conjunto convexo, ocurre

$$\mathbf{z}^*(\lambda) \in \Omega \Rightarrow f(\mathbf{z}^*(\lambda)) = f(\mathbf{x}^*)$$

y los puntos de la forma $\mathbf{z}^*(\lambda)$, con $\lambda \in [0,1]$, son todos mínimos globales de f en Ω. ■

Lema 44 *Sea $f : \Omega \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$, con Ω un conjunto convexo no vacío y $f(\mathbf{x})$ convexa [cóncava]. Sea $\mathbf{x}^*$ un mínimo [máximo] local de f. Entonces, si $\mathbf{x}^*$ es un mínimo [máximo] local estricto de $f(\mathbf{x})$ o si $f(\mathbf{x})$ es estrictamente convexa [cóncava] sobre Ω, entonces $\mathbf{x}^*$ es el único mínimo [máximo] global de $f(\mathbf{x})$ en Ω.*

Demostración: Si $\mathbf{x}^*$ es un mínimo local entonces por el teorema 39 es un mínimo global. Si suponemos ahora que existe otro punto $\mathbf{z}^*$ mínimo de $f(\mathbf{x})$, por el lema 43 el mínimo también se alcanza en todo el segmento que une ambos puntos $\mathbf{x}^*$ y $\mathbf{z}^*$ y esto contradice el hecho de que $\mathbf{x}^*$ sea estricto.

Supongamos ahora que $\mathbf{x}^*$ es un mínimo local de $f(\mathbf{x})$ y que $f(\mathbf{x})$ es estrictamente convexa. Como $f(\mathbf{x})$ es convexa, ya que es estrictamente convexa, $\mathbf{x}^*$ es un mínimo global. Si ahora suponemos que existe otro mínimo global $\mathbf{z}^*$, por la convexidad estricta de f obtenemos:

$$f(\frac{1}{2}\mathbf{x}^* + \frac{1}{2}\mathbf{z}^*) < \frac{1}{2}f(\mathbf{x}^*) + \frac{1}{2}f(\mathbf{z}^*) = f(\mathbf{x}^*)$$

y por tanto $\mathbf{x}^*$ no sería mínimo global ■.

5.6. Introducción a la optimización no lineal

En este capítulo se estudian algunos aspectos relacionados con la que hemos dado en llamar cuestión estática de los problemas de optimización. Se presentan y utilizan condiciones necesarias y suficientes para encontrar de forma explícita la solución o soluciones de un problema de optimización de tipo general.

Recordemos la formulación general, que se introdujo en el capítulo anterior, para un problema de optimización no lineal

$$\textit{(PPNL)} \left\{ \begin{array}{lll} \text{Optimizar} & f(\mathbf{x}) & \\ \text{Sujeto a} & h_i(\mathbf{x}) = 0 & \forall i = 1, \ldots, m \\ & g_j(\mathbf{x}) \leq 0 & \forall j = 1, \ldots, p \end{array} \right. \tag{5.5}$$

donde $f, h_i, g_j : A \to \mathbb{R}$ son funciones reales de varias variables reales definidas sobre los puntos $\mathbf{x} = (x_1, \ldots, x_n)$ del conjunto $A \subseteq \mathbb{R}^n$. Por regla general el conjunto A será abierto y en las mayoría de las ocasiones f, h_i, $g_j \in \mathcal{C}^1(A)$, es decir, las funciones del problema serán derivables sobre A con derivadas parciales continuas.

Recordemos que resolver el problema de optimización *PPNL* es encontrar los óptimos (máximos y/o mínimos) de la función $f(\mathbf{x})$ no sobre todo el conjunto A donde está definida, sino sobre el conjunto factible Ω de los puntos que cumplen todas las restricciones.

Como se comentó en el primer capítulo, si en el modelo general (5.5) ocurre $m = p = 0$, el problema es de *optimización sin restricciones* y se expresa como

$$\textit{(PSR)} \left\{ \begin{array}{ll} \text{Optimizar} & f(\mathbf{x}) \\ & \mathbf{x} \in A \end{array} \right| \tag{5.6}$$

Un caso particular de este tipo de problemas aparece cuando $n = 1$, es decir, cuando sólo hay una variable de decisión, es un problema de optimización univariante

$$\textit{(P1D)} \left\{ \begin{array}{ll} \text{Optimizar} & f(x) \\ & \\ \text{Sujeto a} & x \in I \subseteq \mathbb{R} \end{array} \right.$$

En otro caso (o bien m o bien p son distintos de 0) el problema se dice de *optimización con restricciones* y a sus extremos, ya sean locales o globales, se les conoce como *extremos condicionados,* para distinguirlos de los extremos de los problemas sin restricciones.

Finalmente si $m \neq 0$ y $p = 0$, es decir, solamente hay restricciones de igualdad, entonces el problema es un *problema de Lagrange*

$$\textit{(PRI)} \left\{ \begin{array}{lll} \text{Optimizar} & f(\mathbf{x}) & \\ \text{Sujeto a} & h_i(\mathbf{x}) = 0 & \forall i = 1, \ldots, m \end{array} \right| \tag{5.7}$$

En el problema general *PPNL* el comportamiento de las restricciones de igualdad y el de las de desigualdad es ligeramente distinto.

Definición 20 (Restricciones activas) *Consideremos el problema de optimización* PPNL *y su conjunto factible* Ω. *Sea* $\overline{\mathbf{x}} \in \Omega$*; entonces*

$$g_j(\mathbf{x}) \leq 0 \ \textit{es}\ \text{activa o saturada}\ \textit{en}\ \ \overline{\mathbf{x}} \Longleftrightarrow g_j(\overline{\mathbf{x}}) = 0$$

en caso contrario la restricción es inactiva *o* no saturada *en* $\overline{\mathbf{x}}$.

Nota 9 *Las restricciones activas se comportan como restricciones de igualdad; éstas por su propia naturaleza son activas en cualquier punto factible del problema.*

Definición 21 *Consideremos el problema* PPNL *y su conjunto factible* Ω. *Sea* $\overline{\mathbf{x}} \in \Omega$. *Se define el* conjunto de actividad $J(\overline{\mathbf{x}})$ *asociado a* $\overline{\mathbf{x}}$ *como*

$$J(\overline{\mathbf{x}}) = \{j \in \{1, \ldots, p\} : g_j(\overline{\mathbf{x}}) = 0\}$$

es decir, $J(\overline{\mathbf{x}})$ *es el conjunto de los índices de las restricciones que son activas en* $\overline{\mathbf{x}}$.

Si $\mathbf{x}^*$ es una solución óptima para el problema *PPNL*, entonces las restricciones no activas en él son irrelevantes puesto que no se alcanza la limitación impuesta por dicha restricción. De otro modo, si se conocieran con antelación, sería posible eliminar aquellas restricciones no saturadas de la formulación del problema, pero esto obviamente, no es posible en general.

Otra definición importante en optimización es la de punto regular, aunque para ello es necesario que las funciones del problema sean derivables. Damos a continuación dicha definición junto con la definición de espacio tangente en un punto.

Definición 22 (Punto regular) *Consideremos el problema* NLPP *con* Ω *su conjunto factible. Diremos que* $\overline{\mathbf{x}} \in \Omega$ *es* regular *para las restricciones, si y sólo si el conjunto de vectores definido por*

$$\left\{\{\bigtriangledown h_i(\overline{\mathbf{x}})\}_{i=1,\ldots,m}, \{\bigtriangledown g_j(\overline{\mathbf{x}})\}_{j\in J(\overline{\mathbf{x}})}\right\}$$

constituye una familia de vectores linealmente independientes.

Notar que en la definición sólo se tienen en cuenta las restricciones de igualdad y las de desigualdad activas.

En la definición hay que distinguir dos casos particulares:

1. ***Caso sin restricciones*** $(m = p = 0)$: *En un problema sin restricciones todos los puntos son regulares.*

2. ***Caso con restricciones de desigualdad*** $(m = 0,\ p \neq 0)$: *El punto* $\overline{\mathbf{x}}$ *también es regular si no hay ninguna restricción activa en él, es decir, en problemas que sólo contienen restricciones de desigualdad, el punto* $\overline{\mathbf{x}}$ *también será regular si* $J(\overline{\mathbf{x}}) = \emptyset$.

Definición 23 (Espacio tangente) *Consideremos el problema* PPNL *con* Ω *su conjunto factible. Sea* $\overline{\mathbf{x}} \in \Omega$. *Definimos el* espacio tangente *en* $\overline{\mathbf{x}}$ *al conjunto definido por*

$$M(\overline{\mathbf{x}}) = \left\{\mathbf{d} \in \mathbb{R}^n : \nabla^T h_i(\overline{\mathbf{x}})\,\mathbf{d} = 0, \quad i = 1, \ldots, m; \quad \nabla^T g_j(\overline{\mathbf{x}})\,\mathbf{d} = \mathbf{0}, \quad j \in J(\overline{\mathbf{x}})\right\}$$

o de forma equivalente

$$M(\overline{\mathbf{x}}) = \left\{\mathbf{d} \in \mathbb{R}^n : \sum_{k=1}^{n} \frac{\partial h_i}{\partial x_k}(\overline{\mathbf{x}})\, d_k = 0, \quad i = 1, \ldots, m; \quad \sum_{k=1}^{n} \frac{\partial g_j}{\partial x_k}(\overline{\mathbf{x}})\, d_k = 0, \quad j \in J(\overline{\mathbf{x}})\right\}$$

donde $J(\overline{\mathbf{x}})$ *es el conjunto de actividad asociado a* $\overline{\mathbf{x}}$.

Nota 10 *Como caso particular en problemas sin restricciones, el conjunto* $M(\overline{\mathbf{x}})$ *sería todo el espacio vectorial* $\mathbb{R}^n$.

5.7. Condiciones de Karush-Kuhn-Tucker

En esta sección se presentan algunas de las condiciones necesarias y suficientes que deben cumplir los candidatos a solución óptima del problema de optimización *PPNL.* Estas condiciones son las llamadas condiciones de Karush-Kuhn-Tucker (CKKT).

Para una mayor claridad, se dan a continuación las definiciones correspondientes a los problemas de minimización y maximización de forma separada.

Definición 24 *Dado el problema*

$$\left\{\begin{array}{lll} \textit{Optimizar} & f\left(x_1,\ldots,x_n\right) & \\ \textit{Sujeto a} & h_i\left(x_1,\ldots,x_n\right)=0 & i=1,\ldots,m \\ & g_j\left(x_1,\ldots,x_n\right)\leq 0 & j=1,\ldots,p \end{array}\right. \tag{5.8}$$

con $f,h_i,g_j: A\rightarrow\mathbb{R}$ *funciones de clase* $\mathcal{C}^1(A)$ *y* $A\subseteq\mathbb{R}^n$ *un conjunto abierto. Diremos que* $\mathbf{x}^*=(x_1^*,\ldots,x_n^*)\in A$ *es un* punto de Karush-Kuhn-Tucker *para el problema 5.8 si y sólo si* $\exists\ \lambda_1,\ldots,\lambda_m$, $\mu_1,\ldots,\mu_p\in\mathbb{R}$, *de forma que se cumplen las siguientes condiciones:*

1. Condición estacionaria:

$$\nabla f\left(x_1^*,\ldots,x_n^*\right)+\sum_{i=1}^{m}\lambda_i\nabla h_i\left(x_1^*,\ldots,x_n^*\right)+\sum_{j=1}^{p}\mu_j\nabla g_j\left(x_1^*,\ldots,x_n^*\right)=0$$

2. Condición de factibilidad

$$\begin{array}{rcll} h_i\left(x_1^*,\ldots,x_n^*\right) & = & 0, & i=1,\ldots,m \\ g_j\left(x_1^*,\ldots,x_n^*\right) & \leq & 0, & j=1,\ldots,p \end{array}$$

3. Condición de holgura

$$\mu_j g_j\left(x_1^*,\ldots,x_n^*\right)=0, \qquad j=1,\ldots,p$$

4. Condición de signo: Una vez que se cumplen las condiciones anteriores el punto es de

$$\textit{Mínimo (CKKTMin)}\ \Leftrightarrow\mu_j\geq 0 \qquad j=1,\ldots,p$$

o de

$$\textit{Máximo (CKKTMax)}\ \Leftrightarrow\mu_j\leq 0 \qquad j=1,\ldots,p$$

En ambos casos los valores λ_1, ..., λ_m, μ_1, ..., μ_p son los llamados *multiplicadores* y existe uno por cada restricción del problema: el multiplicador λ_i está asociado a la restricción de igualdad $h_i(x_1,\ldots,x_n)=0$ para $i=1,\ldots,m$ y el multiplicador μ_j está relacionado con la restricción de desigualdad $g_j(x_1,\ldots x_n)\leq 0$ para $j=1,\ldots,p$.

Los valores $(\lambda_1,\ldots,\lambda_m)$ son los *multiplicadores de Lagrange* y los valores $(\mu_1,\ldots,\mu_p)$ son los *multiplicadores de Karush-Kuhn-Tucker.*

De la condición de holgura se deduce que si una restricción de desigualdad es no activa en el punto solución, entonces el multiplicador de Karush-Kuhn-Tucker asociado debe tomar el valor 0.

Los puntos $\mathbf{x}^* \in A \cap \Omega$, siendo Ω el conjunto factible del problema, que cumplen la *condición estacionaria* se dice que son *puntos críticos o estacionarios.* Esta condición estacionaria suele expresarse en términos de la llamada *función Lagrangiana* definida utilizando la función objetivo y las restricciones como

$$L\left(x_1, \ldots, x_n, \lambda_1, \ldots, \lambda_m, \mu_1, \ldots, \mu_p\right) = f\left(x_1, \ldots, x_n\right) + \sum_{i=1}^{m} \lambda_i h_i\left(x_1, \ldots, x_n\right) + \sum_{j=1}^{p} \mu_j g_j\left(x_1, \ldots, x_n\right)$$

o en forma vectorial

$$L\left(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}\right) = f\left(\mathbf{x}\right) + \boldsymbol{\lambda}^T \mathbf{h}\left(\mathbf{x}\right) + \boldsymbol{\mu}^T \mathbf{g}\left(\mathbf{x}\right) \tag{5.9}$$

siendo

$$\begin{aligned}
\boldsymbol{\lambda} &= \left(\lambda_1, \ldots, \lambda_m\right) \\
\boldsymbol{\mu} &= \left(\mu_1, \ldots, \mu_p\right) \\
\mathbf{h}\left(\mathbf{x}\right) &= \left(h_1\left(x_1, \ldots, x_n\right), \ldots, h_m\left(x_1, \ldots, x_n\right)\right) \\
\mathbf{g}\left(\mathbf{x}\right) &= \left(g_1\left(x_1, \ldots, x_n\right), \ldots, g_p\left(x_1, \ldots, x_n\right)\right)
\end{aligned}$$

En forma vectorial la condición estacionaria se puede expresar de forma más compacta como

$$\nabla_{\mathbf{x}} L\left(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}\right) = 0$$

donde el subíndice indica que estamos derivando respecto a las componentes de $\mathbf{x}$.

Nota 11 *Para diferenciar los puntos estacionarios de problemas sin restricciones de los correspondientes a problemas con restricciones, a estos últimos se les suele añadir el adjetivo de* condicionados.

¿Cómo encontrar los puntos de *KKT*?

Para la búsqueda práctica de puntos que cumplan las condiciones de *CKKT* o puntos de Karush-Kuhn-Tucker, ya sean de mínimo o máximo, primero hay que resolver el sistema de ecuaciones compuesto por: *la condición estacionaria, la condición de fatibilidad para las restricciones de igualdad y la condición de holgura*

$$\frac{\partial f}{\partial x_k}\left(x_1^*, \ldots, x_n^*\right) + \sum_{i=1}^{m} \lambda_i \frac{\partial h_i}{\partial x_k}\left(x_1^*, \ldots, x_n^*\right) + \sum_{j=1}^{p} \mu_j \frac{\partial g_j}{\partial x_k}\left(x_1^*, \ldots, x_n^*\right) = 0 \qquad k = 1, \ldots, n$$

$$h_i\left(x_1^*, \ldots, x_n^*\right) = 0 \qquad i = 1, \ldots, m$$

$$\mu_j g_j\left(x_1^*, \ldots, x_n^*\right) = 0 \qquad j = 1, \ldots, p$$

Este sistema está compuesto por $(n + m + p)$ ecuaciones y $(n + m + p)$ incógnitas (las n coordenadas de $\mathbf{x}^* = (x_1^*, \ldots, x_n^*)$, los m multiplicadores de Lagrange λ_i y los p multiplicadores de Karush-Kuhn-Tucker μ_j).

La forma usual de resolver el sistema es comenzar por la condición de holgura complementaria, ya que dichas ecuaciones nos proporcionan dos opciones

$$\mu_j g_j\left(x_1^*,\ldots,x_n^*\right)=0 \Leftrightarrow \begin{cases} \mu_j=0 \\ g_j\left(x_1^*,\ldots,x_n^*\right)=0 \end{cases}$$

para p restricciones de desigualdad tendríamo 2^p posibles casos.

Una vez resuelto el sistema, para ver cuál o cuáles de las soluciones obtenidas son puntos de KKT hay que, por una parte comprobar que son puntos factibles (notar que sólo quedaría por comprobar si $g_j\left(x_1^*,\ldots,x_n^*\right)\leq 0$) y por otra que sus multiplicadores de Karush-Kuhn-Tucker asociados tienen todos el mismo signo, bien todos ≥ 0 para puntos de mínimo o bien todos ≤ 0 para puntos de máximo.

Casos Particulares

Las anteriores condiciones de Karush-Kuhn-Tucker se aplican al problema general de optimización *(NLPP)*, sin embargo se obtienen expresiones más simplificadas de las mismas cuando se aplican a problemas sin restricciones o cuando el el problema sólo tiene restricciones de igualdad.

1. *Problemas sin restricciones* ($m=p=0$)*:* Si el problema no tiene restricciones de ningún tipo (ecuación 5.6), los multiplicadores no son necesarios y tampoco las condiciones relacionadas con ellos. La única condición que queda es la estacionaria

   $$\nabla f\left(x_1^*,\ldots,x_n^*\right)=0 \Leftrightarrow \frac{\partial f}{\partial x_k}\left(x_1^*,\ldots,x_n^*\right)=0 \qquad k=1,\ldots,n$$

   que coincide para ambos objetivos de maximizar y minimizar.

   Si además $n=1$, es decir, el problema es optimizar una función real de variable real, la condición estacionaria nos conduce a un resultado bien conocido del cálculo diferencial

   $$f'\left(x^*\right)=0$$

2. *Problemas de Lagrange* ($p=0$): Si el problema sólo tiene restricciones de igualdad el problema considerado es un problema clásico de *Lagrange* (ecuación 5.7) y las condiciones de Karush-Kuhn-Tucker se obtienen eliminando aquellas ecuaciones relacionadas con restricciones de desigualdad, quedando por tanto la condición estacionaria y la condición de factibilidad

   $$\frac{\partial f}{\partial x_k}\left(x_1^*,\ldots,x_n^*\right)+\sum_{i=1}^{m}\lambda_i\frac{\partial h_i}{\partial x_k}\left(x_1^*,\ldots,x_n^*\right)=0 \qquad k=1,\ldots,n$$

   $$h_i\left(x_1^*,\ldots,x_n^*\right)=0 \qquad i=1,\ldots,m$$

   que es el resultado que proporciona el teorema clásico de los *multiplicadores de Lagrange*. De nuevo las condiciones de KKT para ambos objetivos de maximizar y minimizar coinciden.

A continuación se presentan algunos ejemplos de búsqueda de puntos de Karush-Kuhn-Tucker.

Ejemplo 10 *Encuentra los puntos de Karush-Kuhn-Tucker para el problema*

$$\left.\begin{array}{cr} \textit{Optimizar} & x+y+z \\ \textit{sujeto a} & (y-1)^2+z^2 \leq 1 \\ & x^2+(y-1)^2+z^2 \leq 3 \end{array}\right\}$$

Solución: Se utilizan los multiplicadores correspondientes para construir la función Lagrangiana, expresando previamente las restricciones en la forma $g(x) \leq 0$

$$L(x,y,z) = (x+y+z) + \mu_1 \left((y-1)^2+z^2-1\right) + \mu_2 \left(x^2+(y-1)^2+z^2-3\right)$$

y se plantean cada una de las condiciones de KKT

1. *Condición Estacionaria* ($\nabla_{\mathbf{x}} L = 0$)

$$\frac{\partial L}{\partial x} = 0 \Leftrightarrow 1+2\mu_2 x = 0 \qquad [1]$$

$$\frac{\partial L}{\partial y} = 0 \Leftrightarrow 1+2\mu_1(y-1)+2\mu_2(y-1) = 0 \qquad [2]$$

$$\frac{\partial L}{\partial z} = 0 \Leftrightarrow 1+2\mu_1 z+2\mu_2 z = 0 \qquad [3]$$

2. *Condición de factibilidad*

$$\begin{aligned} \left((y-1)^2+z^2-1\right) &\leq 0 \\ \left(x^2+(y-1)^2+z^2-3\right) &\leq 0 \end{aligned}$$

3. *Condición de holgura*

$$\mu_1 g_1(x) = 0 \Leftrightarrow \mu_1 \left((y-1)^2+z^2-1\right) = 0 \qquad [4]$$

$$\mu_2 g_2(x) = 0 \Leftrightarrow \mu_2 \left(x^2+(y-1)^2+z^2-3\right) = 0 \qquad [5]$$

4. *Condición de signo*

$$\begin{aligned} \mu_1,\mu_2 &\geq 0 \Rightarrow \text{Para mínimo} \\ \mu_1,\mu_2 &\leq 0 \Rightarrow \text{Para máximo} \end{aligned}$$

El sistema que permite localizar los puntos de KKT estará formado por las tres ecuaciones que proporciona la condición estacionaria (ecuaciones [1], [2] y [3]) y las dos de la condición de holgura (ecuaciones [4] y [5]).

A partir de las ecuaciones de la condición de holgura se obtienen cuatro casos

$$\left\{\begin{array}{lcl} \mu_1 = 0 & \Longrightarrow & \left\{\begin{array}{ll} \mu_2 = 0 & \textbf{Caso I} \\ x^2+(y-1)^2+z^2-3=0 & \textbf{Caso II} \end{array}\right. \\ (y-1)^2+z^2-1=0 & \Longrightarrow & \left\{\begin{array}{ll} \mu_2 = 0 & \textbf{Caso III} \\ x^2+(y-1)^2+z^2-3=0 & \textbf{Caso IV} \end{array}\right. \end{array}\right.$$

De la ecuación [1] de la condición estacionaria se deduce que $\mu_2 \neq 0$, ya que en caso contario se llegaría a una contradicción; por ello los casos **I** y **III** no pueden darse y quedan por comprobar los casos **II** y **IV**.

1. **Caso II** $\left(\mu_1 = 0,\, x^2 + (y-1)^2 + z^2 - 3 = 0\right)$**:** Sustituyendo el valor de $\mu_1 = 0$ en las ecuaciones del sistema se obtiene

$$
\begin{aligned}
1 + 2\mu_2 x &= 0 && [6] \\
1 + 2\mu_2 (y-1) &= 0 && [7] \\
1 + 2\mu_2 z &= 0 && [8] \\
x^2 + (y-1)^2 + z^2 - 3 &= 0 && [9]
\end{aligned}
$$

Si restamos las ecuaciones [6] y [7] se obtiene

$$(1 + 2\mu_2 x) - (1 + 2\mu_2 (y-1)) = 0 \Leftrightarrow 2\mu_2 (x - y + 1) = 0$$

y como $\mu_2 \neq 0$, se llega a la conclusión de que

$$x - y + 1 = 0 \Rightarrow x = y - 1$$

Restando las ecuaciones [6] y [8] obtenemos

$$(1 + 2\mu_2 x) - (1 + 2\mu_2 z) = 0 \Leftrightarrow 2\mu_2 (x - z) = 0$$

y como $\mu_2 \neq 0$, se obtiene

$$(x - z) = 0 \Rightarrow x = z$$

Con las relaciones que se han obtenido entre las variables x, y y z, utilizamos la ecuación [9]

$$x^2 + (y-1)^2 + z^2 - 3 = 0 \Rightarrow 3x^2 - 3 = 0 \Leftrightarrow x = \pm 1$$

y para las otras dos variables

$$
\begin{aligned}
z &= x = \pm 1 \\
y &= x + 1 \Rightarrow \begin{cases} y = 2 \\ y = 0 \end{cases}
\end{aligned}
$$

Como $x \neq 0$, despejamos μ_2 de la ecuación [6]

$$\mu_2 = -\frac{1}{2x} = \mp\frac{1}{2}$$

Finalmente y para este caso se han obtenido 2 puntos, que son junto con sus respectivos multiplicadores

$$P_1 = (1, 2, 1) \qquad \mu = \left(0, -\frac{1}{2}\right)$$
$$P_2 = (-1, 0, -1) \qquad \mu = \left(0, \frac{1}{2}\right)$$

2. **Caso IV** $\left((y-1)^2 + z^2 - 1 = 0,\ x^2 + (y-1)^2 + z^2 - 3 = 0\right)$**:** En este último caso, el sistema que hay que resolver es

$$1 + 2\mu_2 x = 0 \qquad [10]$$

$$1 + 2\mu_1 (y-1) + 2\mu_2 (y-1) = 0 \qquad [11]$$

$$1 + 2\mu_1 z + 2\mu_2 z = 0 \qquad [12]$$

$$(y-1)^2 + z^2 - 1 = 0 \qquad [13]$$

$$x^2 + (y-1)^2 + z^2 - 3 = 0 \qquad [14]$$

Restando las ecuaciones [13] y [14] se obtiene el valor de x

$$\left((y-1)^2 + z^2 - 1\right) - \left(x^2 + (y-1)^2 + z^2 - 3\right) = 0 \Rightarrow x^2 - 2 = 0 \Leftrightarrow x^2 = 2 \Leftrightarrow x = \pm\sqrt{2}$$

Se sustituye ese valor en la ecuación [10] para calcular μ_2

$$\mu_2 = -\frac{1}{2x} = -\frac{1}{2\left(\pm\sqrt{2}\right)} = \mp\frac{1}{2\sqrt{2}}$$

Si ahora se restan las ecuaciones [11] y [12] se obtiene

$$(1 + 2\mu_1 (y-1) + 2\mu_2 (y-1)) - (1 + 2\mu_1 z + 2\mu_2 z) = 0$$

y entonces podemos agrupar y sacar factor común

$$2\mu_1 (y - 1 - z) + 2\mu_2 (y - 1 - z) = 0 \Leftrightarrow 2(\mu_1 + \mu_2)(y - 1 - z) = 0$$

Esta ecuación nos proporciona dos opciones. La primera de ellas es

$$\mu_1 + \mu_2 = 0$$

pero utilizando la ecuación [12]

$$1 + 2\mu_1 z + 2\mu_2 z = 0 \Leftrightarrow 1 + 2z(\mu_1 + \mu_2) = 0 \Leftrightarrow 1 = 0$$

que obviamente es imposible. Y la otra opción es

$$y - 1 - z = 0 \Leftrightarrow y - 1 = z$$

y utilizando la ecuación [13] se obtiene el valor de z

$$(y-1)^2 + z^2 = 1 \Rightarrow 2z^2 = 1 \Leftrightarrow z = \pm\frac{1}{\sqrt{2}}$$

y por tanto el de y

$$y = 1 + z = 1 \pm \frac{1}{\sqrt{2}}$$

Queda por determinar el valor del multiplicador μ_1 y para ello utilizamos la ecuación [12]

$$1 + 2\mu_1 z + 2\mu_2 z = 0 \Rightarrow \mu_1 = -\mu_2 - \frac{1}{2z}$$

Si se consideran los distintos valores que se han encontrado para μ_2 (2 valores) y para z (otros 2 valores) tendremos 4 casos posibles

$$\begin{aligned}
\mu_2 &= -\frac{1}{2\sqrt{2}} \qquad z = \frac{1}{\sqrt{2}} \Rightarrow \mu_1 = -\frac{1}{2\sqrt{2}} \\
\mu_2 &= -\frac{1}{2\sqrt{2}} \qquad z = -\frac{1}{\sqrt{2}} \Rightarrow \mu_1 = \frac{3}{2\sqrt{2}} \\
\mu_2 &= \frac{1}{2\sqrt{2}} \qquad z = \frac{1}{\sqrt{2}} \Rightarrow \mu_1 = -\frac{3}{2\sqrt{2}} \\
\mu_2 &= \frac{1}{2\sqrt{2}} \qquad z = -\frac{1}{\sqrt{2}} \Rightarrow \mu_1 = \frac{1}{2\sqrt{2}}
\end{aligned}$$

Se han obtenido cuatro puntos

$$\begin{aligned}
P_3 &= \left(\sqrt{2}, 1+\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) \qquad \mu = \left(-\frac{1}{2\sqrt{2}}, -\frac{1}{2\sqrt{2}}\right) \\
P_4 &= \left(\sqrt{2}, 1-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right) \qquad \mu = \left(\frac{3}{2\sqrt{2}}, -\frac{1}{2\sqrt{2}}\right) \\
P_5 &= \left(-\sqrt{2}, 1+\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) \qquad \mu = \left(-\frac{3}{2\sqrt{2}}, \frac{1}{2\sqrt{2}}\right) \\
P_6 &= \left(-\sqrt{2}, 1-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right) \qquad \mu = \left(\frac{1}{2\sqrt{2}}, \frac{1}{2\sqrt{2}}\right)
\end{aligned}$$

Para determinar cuales de estos puntos son de tipo KKT, hay que comprobar su factiblidad y el signo de sus multiplicadores. Se expone a continuación una tabla resumen con los resultados obtenidos

$\mathbf{P} = (x, y, z)$	$\boldsymbol{\mu} = (\mu_1, \mu_2)$	Factibilidad	Positividad/Negatividad	CKKT
$P_1 = (1, 2, 1)$	$\mu = \left(0, -\frac{1}{2}\right)$	NO	-	-
$P_2 = (0, -1, -1)$	$\mu = \left(0, \frac{1}{2}\right)$	NO	-	-
$P_3 = \left(\sqrt{2}, 1 + \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$	$\mu = \left(-\frac{1}{2\sqrt{2}}, -\frac{1}{2\sqrt{2}}\right)$	SI	Negatividad	Máximo
$P_4 = \left(\sqrt{2}, 1 - \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$	$\mu = \left(\frac{3}{2\sqrt{2}}, -\frac{1}{2\sqrt{2}}\right)$	SI	NO	-
$P_5 = \left(-\sqrt{2}, 1 + \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$	$\mu = \left(-\frac{3}{2\sqrt{2}}, \frac{1}{2\sqrt{2}}\right)$	SI	NO	-
$P_6 = \left(-\sqrt{2}, 1 - \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$	$\mu = \left(\frac{1}{2\sqrt{2}}, \frac{1}{2\sqrt{2}}\right)$	SI	Positividad	Mínimo

Los puntos P_1 y P_2 no son factibles ya que no cumplen la primera restricción del problema

$$\begin{aligned} P_1 &= (1, 2, 1) \Rightarrow g_1(P_1) = (y-1)^2 + z^2 = (2-1)^2 + (1)^2 = 2 \nleq 1 \\ P_2 &= (0, -1, -1) \Rightarrow g_1(P_2) = (y-1)^2 + z^2 = (0-1)^2 + (-1)^2 = 2 \nleq 1 \end{aligned}$$

y por tanto no son puntos válidos para el problema.

Los puntos P_4 y P_5 sí son factibles, sin embargo no tiene multiplicadores de signo constante y por tanto tampoco son puntos de KKT.

Los puntos P_3 y P_6 son los únicos que cumplen con todas las condiciones para ser puntos de KKT; en el caso de P_3 sería un punto de máximo ya que $\mu \leq 0$, mientras que P_6 sería un punto de mínimo puesto que $\mu \geq 0$.

Ejemplo 11 *Encuentra los puntos de Karush-Kuhn-Tucker del siguiente problema*

$$\begin{array}{ll} \textit{Optimizar} & x^2 + y^2 \\ \textit{Sujeto a} & 2x + y - 2 = 0 \end{array}$$

Solución: En este caso $m = 1$ y $p = 0$, es decir hay solamente una restricción de igualdad y el problema es de Lagrange. La función Lagrangiana del problema es

$$L(x, y, \lambda) = \left(x^2 + y^2\right) + \lambda(2x + y - 2)$$

y las condiciones que debe cumplir un punto para ser de Karush-Kuhn-Tucker serán la condición estacionaria y la condición de factibilidad

$$\text{(Condición estacionaria)} \quad \Rightarrow \quad \nabla_{\mathbf{x}} L = 0 \quad \Rightarrow \quad \begin{cases} \dfrac{\partial L}{\partial x} = 0 \quad \Rightarrow \quad 2x + 2\lambda_1 = 0 \\[2ex] \dfrac{\partial L}{\partial y} = 0 \quad \Rightarrow \quad 2y + \lambda_1 = 0 \end{cases}$$

$$\text{(Condición de factibilidad)} \quad \Rightarrow \quad h\left(x, y\right) = 0 \quad \Rightarrow \quad 2x + y - 2 = 0$$

El sistema es lineal y tiene como solución única, y por tanto único punto de KKT

$$x = \frac{4}{5} \qquad y = \frac{2}{5} \qquad \lambda_1 = -\frac{4}{5}$$

Como el multiplicador está asociado a una restricción de igualdad y su signo no influye en el carácter del punto, el punto P es un punto de KKT que puede ser de máximo o de mínimo.

5.8. Condiciones necesarias

Las condiciones de Karush-Kuhn-Tucker son necesarias para la mayoría de los problemas de optimización no lineal, es decir, si $\mathbf{x}^*$ es un óptimo local del problema no lineal *(PPNL)*, entonces será un punto de Karush-Kuhn-Tucker. No obstante para que un óptimo local deba cumplir éstas u otras condiciones, las restricciones en dicho punto tienen que cumplir determinadas propiedades denominadas *Hipótesis de Cualificación de las Restricciones (H.C.R.)*. El estudio de estas hipótesis no cae dentro del ámbito de esta guía y solamente se indican algunas de ellas.

Definición 25 (H.C.R. Sin Restricciones) *En un problema de optimización no lineal sin restricciones ($m = p = 0$), todos los puntos cumplen la primera hipótesis de cualificación de las restricciones.*

Definición 26 (H.C.R. de Karlin o de Linealidad) *En un problema de optimización no lineal donde solamente hay restricciones de tipo lineal, todos los puntos factibles cumplen la hipótesis de cualificación de las restricciones de Karlin.*

Definición 27 (H.C.R. de Slater o de Convexidad) *En un problema de optimización no lineal en el que el conjunto factible, Ω, es un conjunto convexo con interior no vacío, todos los puntos factibles cumplen la hipótesis de cualificación de las restricciones de Slater.*

Definición 28 (H.C.R. de Fiacco-McKormik o de Regularidad) *En un problema de optimización no lineal, todos los puntos factibles que sean regulares cumplen la hipótesis de cualificación de las restricciones de* Fiacco-McKormik.

Condiciones necesarias de primer orden

Estamos en condiciones de establecer las llamadas condiciones necesarias de primer orden que deben cumplir los extremos locales de un problema de optimización.

Teorema 45 (Condiciones necesarias de primer orden) *Dado el problema general de la optimización no lineal (*PPNL*)*

$$\begin{array}{lll} \textit{Optimizar} & f(\mathbf{x}) & \\ \textit{Sujeto a} & h_i(\mathbf{x}) = 0 & i = 1, \ldots, m \\ & g_j(\mathbf{x}) \leq 0 & j = 1, \ldots, p \end{array} \tag{5.10}$$

*donde $f, h_i, g_j : A \to \mathbb{R}$ son funciones de clase $\mathcal{C}^1(A)$ en el conjunto abierto $A \subseteq \mathbb{R}^n$. Sea Ω su conjunto factible y $\mathbf{x}^* \in \Omega$ un punto donde las restricciones del problema cumplen alguna hipótesis de cualificación (*H.C.R.*) y en el que la función $f(\mathbf{x})$ alcanza un mínimo o máximo relativo $\Longrightarrow \mathbf{x}^*$ cumple las condiciones de Karush-Kuhn-Tucker de Mínimo o Máximo respectivamente.*

Ejemplos

Emplearemos el teorema de las condiciones necesarias de primer orden para resolver en esta sección algunos problemas de optimización con restricciones.

Ejemplo 12 *Aplica las condiciones necesarias de primer orden y resuelve el problema 10*

$$\left.\begin{array}{lr} \textit{Optimizar} & f(x, y, z) = x + y + z \\ \textit{sujeto a} & (y-1)^2 + z^2 \leq 1 \\ & x^2 + (y-1)^2 + z^2 \leq 3 \end{array}\right\}$$

Solución: Un análisis inicial permitirá deducir que el problema tiene solución para ambos objetivos de minimizar y maximizar. La función objetivo es continua y el conjunto factible Ω es compacto (Ω es cerrado porque contiene a la frontera que está expresada mediante igualdades y acotado porque es un subconjunto de una esfera de centro $(0, 1, 0)$ y radio $\sqrt{3}$), por tanto por el teorema Weierstrass existirán tanto el mínimo como el máximo de la función sobre el conjunto. Podemos ver a continuación la representación gráfica de las dos restricciones, y en la figura siguiente el conjunto factible,

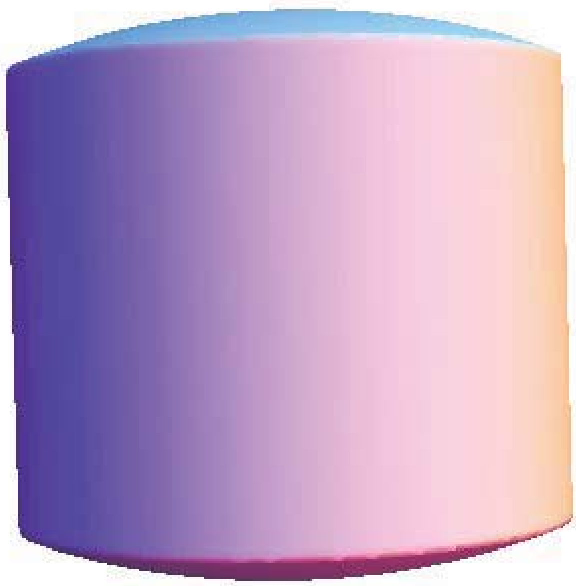
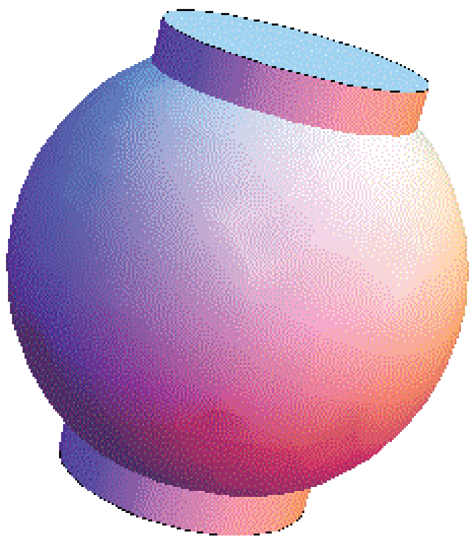


Figura 5.4:

que es la interesección de ambas.



Recordemos que para este problema, los únicos puntos que cumplían las condiciones de CKKT eran

$$P_3 = \left(\sqrt{2}, 1 + \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) \qquad \mu = \left(-\frac{1}{2\sqrt{2}}, -\frac{1}{2\sqrt{2}}\right)$$

para máximo y

$$P_6 = \left(-\sqrt{2}, 1 - \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right) \qquad \mu = \left(\frac{1}{2\sqrt{2}}, \frac{1}{2\sqrt{2}}\right)$$

para mínimo.

Deduzcamos que el problema tiene un mínimo global en P_6, el razonamiento es el siguiente: *"Como la función $f(x, y, z)$ tiene mínimo global, entonces también es local, si además se cumple alguna de las hipótesis de cualificación de las restricciones en ese punto, entonces este mínimo local debe ser un punto de Karush-Kuhn-Tucker de mínimo y por tanto debe ser el punto P_6"*. El razonamiento para máximo y el punto P_3 sería análogo.

Sólo resta por probar que se cumple alguna de las hipótesis de cualificación; en particular, es fácil comprobar que se cumple la hipótesis de cualificación de Slater ya que el conjunto factible Ω es convexo con interior no vacío.

Para comprobar la convexidad de Ω tenemos en cuenta que

$$\Omega = \left\{(x, y, z) \in \mathbb{R} : (y-1)^2 + z^2 \leq 1; \quad x^2 + (y-1)^2 + z^2 \leq 3\right\} = \Omega_1 \cap \Omega_2$$

con

$$\begin{aligned} \Omega_1 &= \left\{(x, y, z) \in \mathbb{R} : (y-1)^2 + z^2 \leq 1\right\} \\ \Omega_2 &= \left\{(x, y, z) \in \mathbb{R} : x^2 + (y-1)^2 + z^2 \leq 3\right\} \end{aligned}$$

siendo los conjuntos Ω_1 y Ω_2 convexos, puesto que son conjuntos de la forma

$$\Gamma = \{(x, y, z) \in \mathbb{R} : g(x, y, z) \leq k\}$$

con $g(x, y, z)$ una función convexa (¿por qué?). El conjunto Ω es convexo por ser intersección de dos conjuntos convexos. Queda por comprobar que el interior de Ω es no vacío, pero eso es fácil ya que al menos el punto $(0, 1, 0)$ está en su interior

$$\begin{aligned} g_1(0, 1, 0) &= (1-1)^2 + 0^2 = 0 < 1 \\ g_2(0, 1, 0) &= 0^2 + (1-1)^2 + 0^2 = 0 < 3 \end{aligned}$$

Los valores óptimos mínimo y máximo de $f(x, y, z)$ sobre Ω se obtienen al evaluar la función objetivo en cada uno de ellos

$$\text{Valor Óptimo Máximo} \Rightarrow f(P_3) = \left(\sqrt{2}\right) + \left(1 + \frac{1}{\sqrt{2}}\right) + \left(\frac{1}{\sqrt{2}}\right) = 1 + 2\sqrt{2}$$

y

$$\text{Valor Óptimo Mínimo} \Rightarrow f(P_6) = \left(-\sqrt{2}\right) + \left(1 - \frac{1}{\sqrt{2}}\right) + \left(-\frac{1}{\sqrt{2}}\right) = 1 - 2\sqrt{2}$$

Ejemplo 13 *Aplica las condiciones necesarias de primer orden al problema*

$$\begin{array}{rl} \textit{Minimizar} & xy + yz + zx \\ \textit{Sujeto a} & x + y + z = 3 \end{array}$$

Solución: Como solamente tiene una restricción de igualdad se trata de un problema de Lagrange; dicha restricción es lineal, luego se cumple la hipótesis de cualificación de Karlin en todos los puntos del conjunto factible. De este modo, si el problema tuviera solución, es decir, si existiera el mínimo global de la función sobre el conjunto factible, sería mínimo local, como debe ser factible, la consecuencia es que debe ser un punto que cumpla las condiciones Karush-Kuhn-Tucker de Mínimo. Al ser un problema de Lagrange estas condiciones se reducen a la condición estacionaria y a la condición de factibilidad.

La función Lagrangiana para este problema es

$$L(x, y, z, \lambda) = (xy + yz + zx) + \lambda (x + y + z - 3)$$

y las condiciones de Karush-Kuhn-Tucker proporcionan el siguiente sistema

$$\text{(Condición estacionaria)} \quad \Rightarrow \quad \nabla_{\mathbf{x}} L = 0 \quad \Rightarrow \quad \begin{cases} \dfrac{\partial L}{\partial x} = 0 \;\Rightarrow\; y + z + \lambda = 0 \\[2ex] \dfrac{\partial L}{\partial y} = 0 \;\Rightarrow\; z + x + \lambda = 0 \\[2ex] \dfrac{\partial L}{\partial z} = 0 \;\Rightarrow\; x + y + \lambda = 0 \end{cases}$$

$$\text{(Condición de factibilidad)} \quad \Rightarrow \quad h(x, y, z) = 0 \quad \Rightarrow \quad x + y + z - 3 = 0$$

que es lineal y con solución única

$$x = y = z = 1 \qquad \lambda = -2$$

El punto obtenido $P = (1, 1, 1)$ junto con el multiplicador asociado $\lambda = -2$ es el único que cumple las condiciones de KKT. Notar que aunque $\lambda_1 = -2 < 0$ y el objetivo sea de minimizar, el punto cumple las condiciones de KKT puesto que λ es un multiplicador asociado a una restricción de igualdad y no está condicionado por su signo. No es posible determinar la naturaleza del punto, puesto que se cumplen las condiciones de KKT tanto para mínimo, como para máximo.

Ejemplo 14 *Plantea y resuelve el problema de construir una caja de cartón rectangular de volumen máximo y área fija A.*

Solución: El planteamiento para este problema viene dado por

$$\begin{array}{rc} \text{Maximizar} & xyz \\ \text{Sujeto a} & xy + yz + zx = \dfrac{A}{2} \end{array}$$

donde x, y, z son las dimensiones de la caja y $A > 0$ su área. Se han omitido las restricciones de positividad sobre las variables ya que por la naturaleza del problema, los valores de estas variables deben > 0, es decir, serán inactivas en cualquier punto, en particular en el punto solución y por tanto los multiplicadores asociados a estas restricciones sería 0.

Comprobaremos que se cumple la hipótesis de regularidad en todos los puntos factibles, para ello calculamos el gradiente de la restricción

$$\nabla h\left(x,y,z\right)=\begin{pmatrix} y+z \\ x+z \\ x+y \end{pmatrix}$$

y estudiamos su dependencia lineal en cada punto del conjunto factible. Como sólo hay un vector, éste será linealmente dependiente cuando sea el vector nulo, es decir

$$\begin{array}{c} y+z=0 \\ x+z=0 \\ x+y=0 \end{array}$$

sistema que tiene por única solución el vector nulo

$$x=y=z=0$$

sin embargo este punto es infactible

$$0+0+0=0\neq\frac{A}{2}$$

de donde se deduce que todos los puntos de Ω (conjunto factible) son regulares. Al cumplirse una de las hipótesis de cualificación de las restricciones, cualquier extremo local del problema que existiera debería cumplir las condiciones de Karush-Kuhn-Tucker. La función Lagrangiana para este problema es

$$L\left(x,y,z,\lambda\right)=xyz+\lambda\left(xy+yz+zx-\frac{A}{2}\right)$$

aplicando las condiciones de KKT se obtiene el sistema

$$\text{(Condición estacionaria)}\quad\Rightarrow\quad\nabla_{\mathbf{x}}L=0\quad\Rightarrow\quad\left\{\begin{array}{l} \dfrac{\partial L}{\partial x}=0\ \Rightarrow\ yz+\lambda\left(y+z\right)=0 \\[2ex] \dfrac{\partial L}{\partial y}=0\ \Rightarrow\ xz+\lambda\left(x+z\right)=0 \\[2ex] \dfrac{\partial L}{\partial z}=0\ \Rightarrow\ xy+\lambda\left(x+y\right)=0 \end{array}\right.$$

$$\text{(Condición de factibilidad)}\quad\Rightarrow\quad h\left(x,y,z\right)=0\quad\Rightarrow\quad xy+yz+zx-\frac{A}{2}=0$$

El sistema anterior tiene como única solución, teniendo en cuenta que $x,y,z>0$ a

$$x=y=z=\sqrt{\frac{A}{6}}\qquad\lambda=-\sqrt{\frac{A}{24}}$$

y al ser un problema que sólo tiene restricciones de igualdad, de momento con estos datos no es posible determinar si el punto es de mínimo o de máximo.

Contraejemplos

El teorema de las condiciones necesarias de primer orden proporciona los requisitos que deben cumplir los extremos de un problema de optimización con restricciones bajo ciertas hipótesis de cualificación, sin embargo, es posible encontrar problemas en los que la solución óptima no cumple estas condiciones, y también es posible encontrar puntos que cumplen las condiciones de Karush-Kuhn-Tucker pero que no son extremos de la función. Veamos algunos de estos ejemplos "patológicos".

Ejemplo 15 *Consideremos el problema*

$$\begin{array}{ll} \textit{Maximizar} & -x^2+y \\ \textit{Sujeto a} & y^3=0 \end{array}$$

Su conjunto factible es

$$\Omega=\left\{(x,y)\in\mathbb{R}^2:(x,0)\right\}$$

El siguiente es un problema equivalente

$$\begin{array}{ll} \textit{Maximizar} & -x^2 \\ \textit{Sujeto a} & x\in\mathbb{R} \end{array}$$

y podemos obtener fácilmente que $x^*=0$ es su solución óptimaen, y por equivalencia entre los dos problemas, será también la solución óptima del problema inicial.

Vamos a comprobar si $P=(0,0)$ es un punto de Karush-Kuhn-Tucker. Si esto fuera cierto, debería existir un multiplicador λ asociado a la restricción $y^3=0$, de forma que se cumplan las ecuaciones del sistema

$$\textit{(Condición estacionaria)} \quad \Rightarrow \quad \nabla_{\mathbf{x}}L=0 \quad \Rightarrow \begin{cases} \dfrac{\partial L}{\partial x}(P)=0 \;\Rightarrow\; -2x=0 \\[2ex] \dfrac{\partial L}{\partial y}(P)=0 \;\Rightarrow\; 1+\lambda 3y^2=0 \end{cases}$$

$$\textit{(Condición de factibilidad)} \quad \Rightarrow \quad h(x,y,z)=0 \quad \Rightarrow \quad y^3=0$$

Sin embargo, este sistema no tiene solución, ya que de la última ecuación necesariamente $y=0$ y al sustituir en la segunda obtenemos una contradicción.

El punto $P=(0,0)$ es un máximo local (de hecho es global puesto que la función objetivo siempre es ≤ 0 y sólo se anula en $x=0$) para el problema de optimización planteado, sin embargo no cumple las condiciones de Karush-Kuhn-Tucker. ¿Contradice este ejemplo el teorema de las condiciones de primer orden? La respuesta es no, ya que P no cumple ninguna de las hipótesis de cualificación de las restricciones necesarias en dicho teorema.

1. Hipótesis sin restricciones: *Está claro que esta hipótesis no se cumple por la presencia de la restricción: $y^3=0$.*

2. Hipótesis de Karlin: *Esta condición tampoco se cumple, puesto que la única restricción del problema, $h(x,y)=y^3$, es claramente no lineal.*

3. Hipótesis de Slater: *El conjunto factible es*

$$\Omega = \left\{(x, y) \in \mathbb{R}^2 : (x, 0)\right\}$$

que es el eje OX, que en $\mathbb{R}^2$ es un conjunto convexo, pero su interior es vacío, puesto que cualquier bola de centro un punto de Ω tiene puntos fuera de Ω y tampoco se cumple esta hipótesis.

4. Hipótesis de Fiacco-McKormik: *En este caso tenemos que comprobar si el punto es o no regular para las restricciones del problema, es decir, habrá que comprobar si el conjunto de vectores formado por los gradientes de las restricciones activas en P está formado por vectores linealmente independientes. Como solamente tenemos una restricción activa en P (por ser de igualdad), la familia de vectores estará formada por un único vector*

$$\{\nabla h\,(x, y)\} = \left\{\left(0, 3y^2\right)\right\}$$

y al evaluar en $P = (0, 0)$ obtenemos

$$\{\nabla h\,(P)\} = \{(0, 0)\}$$

que por ser el vector nulo, es linealmente dependiente; y como consecuencia el punto $P = (0, 0)$ es no regular para las restricciones.

Si en el planteamiento del problema cambiamos la restricción $y^3 = 0$ por la restricción equivalente

$$y = 0$$

la solución del problema es la misma, pero en este caso sí se cumple la hipótesis de Karlin, ya que todas las restricciones son lineales. Ahora tendríamos que ser capaces de encontrar el valor del multiplicador λ y comprobar que el punto $P = (0, 0)$ cumple las condiciones de KKT. El sistema con este cambio es

$$\begin{aligned} -2x &= 0 \\ 1 + \lambda &= 0 \\ y &= 0 \end{aligned}$$

que tiene por solución

$$\begin{aligned} P &= (0, 0) \\ \lambda &= -1 \end{aligned}$$

y hemos encontrado el punto buscado y también su multiplicador correspondiente.

Ejemplo 16 *Sea el problema de optimización*

$$\begin{array}{ll} \textit{Maximizar} & f(x,y,z) = y \\ \textit{Sujeto a} & h_1(x,y,z) = (x-1)^2 + y^2 - 1 = 0 \\ & h_2(x,y,z) = (x+1)^2 + y^2 - 1 = 0 \end{array}$$

El conjunto factible para este problema es

$$\Omega = \left\{(x,y,z) \in \mathbb{R}^3 : (0,0,z)\right\}$$

y la solución óptima del problema es cualquier punto de Ω, puesto que $f(x,y,z)$ es constante sobre él.

La condición estacionaria para este problema nos proporciona las siguientes ecuaciones

$$\textit{(Condición estacionaria)} \quad \Rightarrow \quad \nabla_{\mathbf{x}} L = 0 \quad \Rightarrow \quad \begin{cases} \dfrac{\partial L}{\partial x} = 0 \quad \Rightarrow \quad 2\lambda_1(x-1) + 2\lambda_2(x+1) = 0 \\[2ex] \dfrac{\partial L}{\partial y} = 0 \quad \Rightarrow \quad 1 + 2\lambda_1 y + 2\lambda_2 y = 0 \\[2ex] \dfrac{\partial L}{\partial z} = 0 \quad \Rightarrow \quad 0 = 0 \end{cases}$$

siendo $L(x,y,z,\lambda_1,\lambda_2) = f + \lambda_1 h_1 + \lambda_2 h_2$ la función lagrangiana del problema.

Ninguno de los puntos de Ω, $(0,0,z)$, es solución del sistema anterior, puesto que al sustituir cualquiera de ellos en la segunda ecuación nos llevaría a una contradicción: ¡Ninguno de los extremos locales del problema cumple las condiciones de KKT!

Podemos comprobar, como en el caso anterior, que ninguno de ellos cumple ninguna de las hipótesis de cualificación. Es un problema con restricciones, ambas no lineales y donde el conjunto factible Ω tiene interior vacío por ser una recta. Para comprobar si se cumple la hipótesis de regularidad observamos que el conjunto de vectores que son gradiente de las restricciones activas (en este caso son todas puesto que es un problema con sólo igualdades) está dado por

$$\{\nabla h_1(x,y,z), \nabla h_2(x,y,z)\} = \left\{ \begin{pmatrix} 2(x-1) \\ 2y \\ 0 \end{pmatrix}, \begin{pmatrix} 2(x+1) \\ 2y \\ 0 \end{pmatrix} \right\}$$

y al evaluarlo en los puntos óptimos $(0,0,z)$ obtenemos

$$\left\{ \begin{pmatrix} -2 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} \right\}$$

que es una familia de vectores linealmente dependientes y por tanto ningún punto de la forma $(0,0,z)$ es regular.

También podría suceder que un punto donde las restricciones no cumplan ninguna de las hipótesis de cualificación, sea extremo local del problema y cumpla las condiciones de Karush-Kuhn-Tucker.

Consideremos, por ejemplo, las restricciones del ejemplo anterior, pero cambiando la función objetivo por $f(x, y, z) = x$.

En este las condición estacionaria para el problema es

$$\text{(Condición estacionaria)} \quad \Rightarrow \quad \nabla_{\mathbf{x}} L = 0 \quad \Rightarrow \quad \begin{cases} \dfrac{\partial L}{\partial x} = 0 \quad \Rightarrow \quad 1 + 2\lambda_1 (x-1) + 2\lambda_2 (x+1) = 0 \\ \\ \dfrac{\partial L}{\partial y} = 0 \quad \Rightarrow \quad 2\lambda_1 y + 2\lambda_2 y = 0 \\ \\ \dfrac{\partial L}{\partial z} = 0 \quad \Rightarrow \quad 0 = 0 \end{cases}$$

y teniendo en cuenta que el conjunto factible está formado por los puntos $(0, 0, z)$, el sistema queda como

$$\begin{aligned} 1 - 2\lambda_1 + 2\lambda_2 &= 0 \\ \\ 0 &= 0 \\ \\ 0 &= 0 \end{aligned}$$

que tiene por solución

$$\lambda_1 - \lambda_2 = \frac{1}{2}$$

Tomando ahora cualquier solución de esta ecuación, por ejemplo $\lambda = (\lambda_1, \lambda_2) = (1, 1/2)$, vemos que todos los puntos extremos cumplen las condiciones de Karush-Kuhn-Tucker, sin embargo, como se ha comprobado, en ninguno de ellos las restricciones cumplen ninguna de las hipótesis de cualificación.

Todos estos ejemplos son atípicos y en general sucederá que los extremos locales del problema tendrán que cumplir las condiciones de Karush-Kuhn-Tucker, pero ilustran la necesidad de comprobar adecuadamente los resultados obtenidos.

Por último hay que indicar que estas condiciones son necesarias, en el sentido de que bajo las hipótesis del teorema, los extremos de un problema de optimización deben ser puntos de Karush-Kuhn-Tucker. Sin embargo, las condiciones no son suficientes, ya que podemos encontrar puntos que aún cumpliendo las condiciones de Karush-Kuhn-Tucker, no son extremos, por ejemplo la función $f(x) = x^3$ tiene como único punto estacionario $x = 0$, que no es extremo puesto que la función es siempre creciente.

Definición 29 *Los puntos factibles que cumplen la condición estacionaria pero que no son extremos de la función se denominan* puntos de silla *(que son* condicionados *si hay presencia de restricciones en el problema). Para funciones reales de una variable a estos puntos se les conoce mejor por* puntos de inflexión.

Condiciones necesarias de segundo orden

Si las funciones del problema son suficientemente derivables, entonces, podemos utilizar información relativa a las segundas derivadas.

Teorema 46 (Condiciones necesarias de 2^o orden) *Dado el problema general de optimización no lineal (*PPNL*)*

$$\begin{array}{lll} \textit{Optimizar} & f(\mathbf{x}) & \\ \textit{Sujeto a} & h_i(\mathbf{x}) = 0 & i = 1, \ldots, m \\ & g_j(\mathbf{x}) \leq 0 & j = 1, \ldots, p \end{array}$$

*donde $f, h_i, g_j : A \to \mathbb{R}$ son funciones de clase $\mathcal{C}^2(A)$ en el conjunto abierto $A \subseteq \mathbb{R}^n$. Sea Ω su conjunto factible y $\mathbf{x}^* \in \Omega$ un punto donde las restricciones del problema cumplen alguna hipótesis de cualificación de las restricciones (*H.C.R.*) y en el que la función $f(\mathbf{x})$ alcanza un mínimo [respectivamente máximo] relativo condicionado $\Longrightarrow \mathbf{x}^*$ es un punto que cumple las condiciones de Karush-Kuhn-Tucker de Mínimo [respectivamente Máximo] es decir $\exists\ \lambda_1, \ldots, \lambda_m,\ \mu_1, \ldots, \mu_p \in \mathbb{R}$ de forma que se cumplen las siguientes condiciones:*

1. Condición estacionaria

$$\nabla f(\mathbf{x}^*) + \sum_{i=1}^{m} \lambda_i \nabla h_i(\mathbf{x}^*) + \sum_{j=1}^{p} \mu_j \nabla g_j(\mathbf{x}^*) = 0$$

2. Condición de factibilidad

$$\begin{array}{rcll} h_i(\mathbf{x}^*) & = & 0, & i = 1, \ldots, m \\ g_j(\mathbf{x}^*) & \leq & 0 & j = 1, \ldots, p \end{array}$$

3. Condición de holgura

$$\mu_j g_j(\mathbf{x}^*) = 0 \qquad j = 1, \ldots, p$$

4. Condición de signo

$$\mu_j \geq 0 \textit{ para mínimo} \quad \left[\mu_j \leq 0 \textit{ para máximo}\right] \qquad j = 1, \ldots, p$$

y además se cumple la siguiente condición:

5. Condición del Hessiano: *La matriz $HL(\mathbf{x}^*, \boldsymbol{\lambda}, \boldsymbol{\mu})$ definida como*

$$HL(\mathbf{x}^*, \boldsymbol{\lambda}, \boldsymbol{\mu}) = Hf(\mathbf{x}^*) + \sum_{i=1}^{m} \lambda_i Hh_i(\mathbf{x}^*) + \sum_{j=1}^{p} \mu_j Hg_j(\mathbf{x}^*) \tag{5.11}$$

es semidefinida positiva [semidefinida negativa respectivamente] sobre el espacio tangente $M(\mathbf{x}^*)$ en $\mathbf{x}^*$.

Cuando en el problema no hay restricciones o solamente hay restricciones de igualdad, las condiciones KKT tienen formas particulares que indicamos a continuación.

1. *Caso sin restricciones y una variable* ($m = p = 0$, $n = 1$): En el caso de problemas con una sola variable, el Hessiano de la función $f(x)$ es su segunda derivada y la condición 5.11 se convierte en

$$f''(x^*) \geq 0$$

2. *Sin restricciones y varias variables* ($m = p = 0$): En este caso el espacio tangente es $M(\mathbf{x}^*) = \mathbb{R}^n$, $HL = Hf$ y la condición necesaria que debe cumplirse es

$$\begin{aligned} \mathbf{x}^* \text{ es mínimo local} \ &\Rightarrow \ Hf(\mathbf{x}^*) \text{ es semidefinida positiva} \\ \mathbf{x}^* \text{ es máximo local} \ &\Rightarrow \ Hf(\mathbf{x}^*) \text{ es semidefinida negativa} \end{aligned}$$

3. *Problemas de Lagrange* ($p = 0$): Para problemas que tengan solamente restricciones de igualdad la condición del Hessiano no cambia, salvo que en este caso no hay términos asociados a restricciones de desigualdad.

Ejemplo 17 *Aplica las condiciones necesarias de segundo orden al problema 10*

$$\left.\begin{array}{rr} \textit{Optimizar} & f(x,y,z) = x + y + z \\ \textit{sujeto a} & (y-1)^2 + z^2 \leq 1 \\ & x^2 + (y-1)^2 + z^2 \leq 3 \end{array}\right\}$$

Solución: Vamos a comprobar que las soluciones que se encontraron para este problema, cumplen las condiciones de segundo orden.

$$\text{(Máximo) } P_3 = \left(\sqrt{2}, 1 + \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) \qquad \mu = \left(-\frac{1}{2\sqrt{2}}, -\frac{1}{2\sqrt{2}}\right)$$

$$\text{(Mínimo) } P_6 = \left(-\sqrt{2}, 1 - \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right) \qquad \mu = \left(\frac{1}{2\sqrt{2}}, \frac{1}{2\sqrt{2}}\right)$$

Para aplicar las condiciones de segundo orden tendremos que construir la matriz $HL(P_k)$ en cada punto y evaluarlar sobre el espacio tangente correspondiente $M(P_k)$.

Comenzamos por definir la matriz HL

$$HL(\mathbf{x}) = Hf(\mathbf{x}) + \mu_1 Hg_1(\mathbf{x}) + \mu_2 Hg_2(\mathbf{x}) = \begin{pmatrix} 2\mu_2 & 0 & 0 \\ 0 & 2(\mu_1 + \mu_2) & 0 \\ 0 & 0 & 2(\mu_1 + \mu_2) \end{pmatrix}$$

Para el punto P_3 tendremos

$$HL(P_3) = \begin{pmatrix} -\frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & -\sqrt{2} & 0 \\ 0 & 0 & -\sqrt{2} \end{pmatrix}$$

que es una matriz semidefinida negativa sobre todo $\mathbb{R}^3$ puesto que es diagonal negativa. Como el espacio tangente $M(P_3)$ es un subespacio vectorial de $\mathbb{R}^3$, la matriz $HL(P_3)$ también será semidefinida negativa sobre él y por tanto P_3 cumple las condiciones necesarias de segundo orden para máximo, como era de esperar.

Para el punto P_6 tendremos

$$HL(P_6) = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix}$$

que es una matriz semidefinida positiva sobre todo $\mathbb{R}^3$, puesto que es diagonal positiva. Como de nuevo, el espacio tangente $M(P_6)$ es un subespacio vectorial de $\mathbb{R}^3$, la matriz $HL(P_6)$ también será semidefinida positiva sobre él y por tanto P_6 cumple las condiciones necesarias de segundo orden para mínimo.

Ejemplo 18 *Aplica las condiciones de segundo orden al problema 14 de la caja.*

Solución: Recordemos que el problema de optimización podía plantearse como

$$\begin{array}{cc} \text{Maximizar} & xyz \\ \text{Sujeto a} & xy + yz + zx = \dfrac{A}{2} \end{array}$$

y que habíamos encontrado como único punto de KKT el siguiente

$$P = \left(\sqrt{\frac{A}{6}}, \sqrt{\frac{A}{6}}, \sqrt{\frac{A}{6}}\right) \qquad \lambda = -\frac{1}{2}\sqrt{\frac{A}{6}}$$

Construiremos a continuación tanto la matriz $HL(P)$ como el espacio tangente $M(P)$

$$HL = Hf + \lambda Hh = \begin{pmatrix} 0 & z & y \\ z & 0 & x \\ y & x & 0 \end{pmatrix} + \lambda \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & z+\lambda & y+\lambda \\ z+\lambda & 0 & x+\lambda \\ y+\lambda & x+\lambda & 0 \end{pmatrix}$$

evaluando en P y teniendo en cuenta que $x = y = z = -2\lambda$

$$x + \lambda = y + \lambda = z + \lambda = -\lambda = \frac{1}{2}\sqrt{\frac{A}{6}}$$

y la matriz hessiana en P es

$$HL(P) = \frac{1}{2}\sqrt{\frac{A}{6}} \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

Para determinar el espacio tangente $M(P)$ necesitamos calcular $\nabla h(P)$

$$\nabla h(x, y, z) = (y + z, z + x, x + y) \Rightarrow \nabla h(P) = 2\sqrt{\frac{A}{6}}(1, 1, 1)$$

y entonces

$$\begin{aligned} M(P) &= \left\{\mathbf{d} \in \mathbb{R}^3 : \nabla h(P)\,\mathbf{d} = 0\right\} \\ &= \left\{(d_1, d_2, d_3) \in \mathbb{R}^3 : \quad 2\sqrt{\frac{A}{6}}(1, 1, 1)\begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix} = 0\right\} = \left\{(d_1, d_2, d_3) \in \mathbb{R}^3 : \quad d_1 + d_2 + d_3 = 0\right\} \end{aligned}$$

y el espacio tangente estará descrito por

$$M(P) = \left\{(d_1, d_2, d_3) \in \mathbb{R}^3 : \quad (d_1, d_2, -(d_1 + d_2))\right\}$$

Si construimos la forma cuadrática asociada a $HL(P)$ sobre $M(P)$ tendremos

$$\begin{aligned}
\varphi_{HL(P)}(\mathbf{d})\big|_{M(P)} &= (d_1, d_2, -(d_1+d_2)) \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} d_1 \\ d_2 \\ -(d_1+d_2) \end{pmatrix} = \\
&= (d_1, d_2, -(d_1+d_2)) \begin{pmatrix} -d_1 \\ -d_2 \\ (d_1+d_2) \end{pmatrix} \\
&= -d_1^2 - d_2^2 - (d_1+d_2)^2 \\
&= -\left(d_1^2 + d_2^2 + (d_1+d_2)^2\right)
\end{aligned}$$

que claramente toma siempre valores negativos, lo que implica que $HL(P)$ es semidefinida negativa sobre $M(P)$ y el punto P cumple las condiciones necesarias de segundo orden para ser un máximo del problema.

Con las condiciones necesarias obtenemos condiciones que permiten eliminar aquellos puntos que no son candidatos a extremo de la función, sin embargo, necesitamos unas condiciones que permitan asegurar que los puntos encontrados son realmente las soluciones buscadas.

5.9. Condiciones suficientes

En el apartado anterior se han proporcionado condiciones necesarias de primer y segundo orden que permitían descartar como soluciones a aquellos puntos que no las cumplieran, sin embargo, en ocasiones, en algunos problemas, es posible encontrar puntos que cumplan tanto las primeras condiciones como las segundas, sin ser la solución al problema.

Ejemplo 19 *Comprueba que el punto $P = (0,0)$ cumple las condiciones necesarias de primer y segundo orden para el problema*

$$\begin{array}{cc} \textit{Minimizar} & (x-y^2)(x-3y^2) \\ \\ s.a. & (x,y) \in \mathbb{R}^2 \end{array}$$

pero que no es su solución.

Solución: Al tratarse de un problema sin restricciones se cumple una de las hipótesis de cualificación, por tanto cualquier mínimo debe cumplir las condiciones de Karush-Kuhn-Tucker que en este caso se reducen a la condición estacionaria

$$\nabla f(x,y) = 0 \Leftrightarrow \begin{pmatrix} 2x - 4y^2 \\ \\ 12y^3 - 8xy \end{pmatrix} = \begin{pmatrix} 0 \\ \\ 0 \end{pmatrix}$$

Este sistema tiene como única solución el punto $P = (0,0)$, es decir, P cumple las condiciones necesarias de primer orden.

Las condiciones de segundo orden para problemas sin restricciones se reducen a comprobar el Hessiano de la función $f(x,y)$ en el punto en cuestión. Si calculamos la matriz Hessiana de $f(x,y)$

$$Hf(x,y) = \begin{pmatrix} 2 & -8y \\ -8y & 36y^2 - 8x \end{pmatrix}$$

y lo evaluamos en $P = (0,0)$

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

La matriz $Hf(P)$ es semidefinida positiva, puesto que sus valores propios son $\lambda_1 = 2 \geq 0$ y $\lambda_2 = 0 \geq 0$; esto implica que el punto P también cumple las condiciones necesarias de segundo orden, concretamente las condiciones de mínimo. Sin embargo vamos a comprobar que el punto P no es un mínimo. Por una parte el valor de la función en P es nulo

$$f(0,0) = 0$$

y por otra parte si evaluamos la función sobre los puntos de la curva

$$x = 3y^2$$

obtenemos

$$f(x,y) = f\left(3y^2, y\right) = \left(3y^2 - y^2\right)\left(3y^2 - 4y^2\right) = -2y^4 \leq 0$$

es decir sobre los puntos de esa curva sucede

$$f(x,y) \leq 0 = f(0,0)$$

y P no podría ser el mínimo, puesto que hay valores cerca de él (tomando $y \to 0$) donde el valor de la función es menor.

Este tipo de problema provoca el estudio de condiciones cuyo cumplimiento garantice el hallazgo de la solución. Este tipo de condiciones son las llamadas *suficientes*.

Con el fin de dar estas condiciones de suficiencia es necesario exigir que la matriz Hessiana correspondiente sea por una parte *definida* (positiva o negativa para mínimo o máximo, respectivamente) y por otra que lo sea en un espacio mayor que el espacio tangente.

Definición 30 *Dado el problema general con restricciones*

$$\begin{array}{lll} \textit{Optimizar} & f(\mathbf{x}) & \\ \textit{Sujeto a} & h_i(\mathbf{x}) = 0 & i = 1, \ldots, m \\ & g_j(\mathbf{x}) \leq 0 & j = 1, \ldots, p \end{array}$$

donde $f, h_i, g_j : A \to \mathbb{R}$*, son funciones de clase* $\mathcal{C}^1(A)$ *en* $A \subseteq \mathbb{R}^n$ *abierto. Sea* Ω *su conjunto factible y* $\mathbf{x}^* \in \Omega$ *un punto de Karush-Kuhn-Tucker para el problema. Diremos que una restricción de desigualdad* $g_j(\mathbf{x})$ *es* degenerada *en* $\mathbf{x}^* \iff g_j(\mathbf{x}^*) = 0$ *y* $\mu_j = 0$.

Definición 31 *Definimos el* conjunto de índices de restricciones no degeneradas *en un punto* $\mathbf{x}^*$ *de KKT como*

$$\widetilde{J}(\mathbf{x}^*) = \left\{ j \in \{1,\ldots,p\} |\, g_j(\mathbf{x}^*) = 0 \text{ y } \mu_j \neq 0 \right\}$$

Notar que si en el problema no hay restricciones o son todas de igualdad entonces $\widetilde{J}(\mathbf{x}^*) = \emptyset$.

Definición 32 *Definimos el* espacio tangente ampliado *como*

$$\widetilde{M}(\mathbf{x}^*) = \left\{ \mathbf{d} \in \mathbb{R}^n | \quad \nabla^T h_i(\mathbf{x}^*)\,\mathbf{d} = 0, \quad i = 1,\ldots,m; \quad \nabla^T g_j(\mathbf{x}^*)\,\mathbf{d} = \mathbf{0}, \quad j \in \widetilde{J}(\mathbf{x}^*) \right\}$$

Notar que en el caso de un problema sin restricciones el espacio tangente y el espacio tangente ampliado coinciden:

$$\widetilde{M}(\mathbf{x}^*) = M(\mathbf{x}^*) = \mathbb{R}^n$$

y para un problema de Lagrange

$$\widetilde{M}(\mathbf{x}^*) = M(\mathbf{x}^*)$$

Teorema 47 (Condiciones suficientes) *Dado el problema general de optimización*

$$\begin{array}{lll} \textit{Optimizar} & f(\mathbf{x}) & \\ \textit{Sujeto a} & h_i(\mathbf{x}) = 0 & i = 1,\ldots,m \\ & g_j(\mathbf{x}) \leq 0 & j = 1,\ldots,p \end{array}$$

donde $f, h_i, g_j : A \to \mathbb{R}$ *son funciones de clase* $\mathcal{C}^2(A)$ *en* $A \subseteq \mathbb{R}^n$ *un conjunto abierto. Sea* Ω *su conjunto factible y* $\mathbf{x}^* \in \Omega$ *un punto donde las restricciones del problema cumplen alguna de las hipótesis de cualificación. Si* $\mathbf{x}^*$ *es un punto de Karush-Kuhn-Tucker de Mínimo [Máximo], es decir* $\exists\ \lambda_1, \ldots, \lambda_m,\ \mu_1, \ldots, \mu_p \in \mathbb{R}$ *de forma que se cumplen las siguientes condiciones:*

1. Condición estacionaria

$$\nabla f(\mathbf{x}^*) + \sum_{i=1}^{m} \lambda_i \nabla h_i(\mathbf{x}^*) + \sum_{j=1}^{p} \mu_j \nabla g_j(\mathbf{x}^*) = 0$$

2. Condición de factibilidad

$$\begin{array}{rcll} h_i(\mathbf{x}^*) & = & 0, & i = 1,\ldots,m \\ g_j(\mathbf{x}^*) & \leq & 0 & j = 1,\ldots,p \end{array}$$

3. Condición de holgura

$$\mu_j g_j(\mathbf{x}^*) = 0 \qquad j = 1,\ldots,p$$

4. Condición de signo

$$\mu_j \geq 0 \text{ para Mínimo} \quad \left[\mu_j \leq 0 \text{ para Máximo}\right] \qquad j = 1,\ldots,p$$

5. Condición del Hessiano*: La matriz $HL(\mathbf{x}^*)$ definida como*

$$HL(\mathbf{x}^*) = Hf(\mathbf{x}^*) + \sum_{i=1}^{m} \lambda_i Hh_i(\mathbf{x}^*) + \sum_{j=1}^{p} \mu_j Hg_j(\mathbf{x}^*)$$

es definida positiva [definida negativa respectivamente] sobre el espacio tangente ampliado $\widetilde{M}(\mathbf{x}^*) \Longrightarrow$ Entonces en $\mathbf{x}^*$ hay un mínimo [máximo] relativo condicionado estricto de f sobre Ω.

Si la matriz $HL(\mathbf{x}^*)$ es indefinida sobre $\widetilde{M}(\mathbf{x}^*)$ entonces en $\mathbf{x}^*$ hay un punto de silla condicionado.

Casos Particulares:

Teorema 48 *1.* Sin restricciones y una variable *($m = p = 0$, $n = 1$): En el caso de problemas con una sola variable, la condición del Hessiano se convierte en*

$$f''(x^*) > 0.$$

2. Sin restricciones y varias variables *($m = p = 0$): En este caso $M(\mathbf{x}^*) = \mathbb{R}^n$ y la condición del Hessiano es*

Si la matriz $Hf(\mathbf{x}^*)$ es definida positiva [negativa] $\Rightarrow \mathbf{x}^*$ es un mínimo [máximo] local estricto

Ejemplo 20 *Resuelve el problema*

$$\begin{array}{rl} \textit{Minimizar} & xy + yz + zx \\ \textit{Sujeto a} & x + y + z = 3 \end{array}$$

Solución: Se comprobó anteriormente que el único punto crítico obtenido era:

$$x = y = z = 1 \qquad \lambda_1 = -2$$

Si ahora tratamos de emplear las condiciones suficientes descritas en la proposición anterior, tendremos:

$$\mathbf{H}f(x, y, z) = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

que no es ni definida positiva, ni definida negativa si consideramos todos los vectores de $\mathbb{R}^3$, sin embargo si restringimos la matriz a los puntos del espacio tangente ampliado $\widetilde{M}(\mathbf{x}^*)$, que por ser un problema de Lagrange que continene solamente restricciones de igualdad, coincide con el espacio tangente $M(\mathbf{x}^*)$

$$\begin{aligned} M(\mathbf{x}^*) &= \left\{ \mathbf{d} \in \mathbb{R}^3 \middle| \nabla h^T(\mathbf{x}^*) \cdot \mathbf{d} = 0 \right\} = \left\{ \mathbf{d} \in \mathbb{R}^3 \middle| \left. (1,1,1) \right|_{(1,1,1)} \cdot \mathbf{d} = 0 \right\} = \\ &= \left\{ \mathbf{d} \in \mathbb{R}^3 \middle| (1,1,1) \begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix} = 0 \right\} = \left\{ \mathbf{d} \in \mathbb{R}^3 \middle| d_1 + d_2 + d_3 = 0 \right\} \end{aligned}$$

la forma cuadrática asociada será

$$\begin{aligned}
\varphi|_{Hf(x^*)}(\mathbf{d}) &= (d_1,d_2,d_3)\begin{bmatrix}0&1&1\\1&0&1\\1&1&0\end{bmatrix}\begin{pmatrix}d_1\\d_2\\d_3\end{pmatrix} = (d_1,d_2,-d_1-d_2)\begin{bmatrix}0&1&1\\1&0&1\\1&1&0\end{bmatrix}\begin{pmatrix}d_1\\d_2\\-d_1-d_2\end{pmatrix} = \\
&= (d_1,d_2,-d_1-d_2)\begin{pmatrix}-d_1\\-d_2\\d_1+d_2\end{pmatrix} = -d_1^2-d_2^2-(d_1+d_2)^2 = -\left(d_1^2+d_2^2+(d_1+d_2)^2\right) \geq 0
\end{aligned}$$

y solamente será 0, cuando

$$d_1 = d_2 = (d_1+d_2) = 0$$

y por tanto

$$\mathbf{d} = (0,0,0)$$

Por tanto la forma cuadrática $\varphi|_{Hf(x^*)}(\mathbf{d})$ asociada a la matriz $\mathbf{H}f(x^*)$ es definida negativa sobre el espacio tangente $M(\mathbf{x}^*)$ y por la proposición anterior el punto $\mathbf{x}^*$ será un máximo local estricto.

En el caso de los problemas convexos las condiciones necesarias de primer orden son también suficientes.

Teorema 49 (Problemas Convexos) *Dado el problema*

$$\begin{array}{lll}
\textit{Optimizar} & f(\mathbf{x}) & \\
\textit{Sujeto a} & h_i(\mathbf{x}) = 0 & i=1,\ldots,m \\
 & g_j(\mathbf{x}) \leq 0 & j=1,\ldots,p
\end{array}$$

con $f, h_i, g_j : A \to \mathbb{R}$, funciones de clase $\mathcal{C}^1(A)$, con $A \subseteq \mathbb{R}^n$ un conjunto abierto y Ω su conjunto factible. Si Ω es convexo y $f(\mathbf{x})$ es convexa [cóncava respectivamente] sobre Ω, entonces, si existe $\mathbf{x}^* \in \Omega$ y multiplicadores $\lambda_1,\ldots,\lambda_m$, $\mu_1,\ldots,\mu_p \in \mathbb{R}$ tales que se cumplen las condiciones necesarias de primer orden para mínimo [máximo] local, entonces $\mathbf{x}^*$ es un mínimo [máximo] global del problema.

Proposición 50 (Problemas con desigualdades) *Dado el problema* PPNL

$$\begin{array}{lll}
\textit{Optimizar} & f(\mathbf{x}) & \\
\textit{Sujeto a} & g_j(\mathbf{x}) \leq 0 & j=1,\ldots,p
\end{array}$$

con $f, g_j : A \to \mathbb{R}$, funciones de clase $\mathcal{C}^1(A)$, con $A \subseteq \mathbb{R}^n$ un conjunto abierto y Ω su conjunto factible. Supongamos que $f(\mathbf{x})$ es convexa [cóncava] sobre Ω y supongamos también que $g_1(\mathbf{x})$, ..., $g_p(\mathbf{x})$ son funciones convexas sobre Ω entonces si existe un $\mathbf{x}^* \in \Omega$ punto CKKTMin *[*CKKTMax*] entonces $\mathbf{x}^*$ es solución del problema* PPNL.

Además si f, g_1, ..., g_p son estrictamente convexas $\mathbf{x}^*$ es la única solución del problema PPNL.

Proposición 51 (Problemas Afines Convexos) *Dado el problema* PPNL

$$\begin{array}{lll}
\textit{Optimizar} & f(\mathbf{x}) & \\
\textit{Sujeto a} & h_i(\mathbf{x}) = 0 & i=1,\ldots,m \\
 & g_j(\mathbf{x}) \leq 0 & j=1,\ldots,p
\end{array}$$

con $f, h_i, g_j : A \to \mathbb{R}$, funciones de clase $\mathcal{C}^1(A)$, con $A \subseteq \mathbb{R}^n$ un conjunto abierto y Ω su conjunto factible. Supongamos que $f(\mathbf{x})$ es convexa [cóncava respectivamente] sobre Ω y supongamos también que $h_1(\mathbf{x}), \ldots, h_m(x)$ son funciones afines y $g_1(\mathbf{x}), \ldots, g_p(\mathbf{x})$ son funciones convexas sobre Ω entonces si existe un $\mathbf{x}^* \in \Omega$ punto CKKTMin *[*CKKTMax*] entonces $\mathbf{x}^*$ es solución del problema* PPNL.

Una función $h(\mathbf{x})$ es afín si es de la forma

$$\mathbf{h}(\mathbf{x}) = a_1 x_1 + \cdots + a_n x_n + b$$

Ejemplo 21 *Resuelve el siguiente problema* PPNL

$$\begin{array}{cc} \textit{Optimizar} & y \\ \textit{Sujeto a} & x+y+z=1 \\ & x^2+z^2 \leq 9 \end{array}$$

Solución: Planteamos las condiciones de KKT para $L(x, y, z, \lambda, \mu) = y + \lambda(x+y+z-1) + \mu(x^2+z^2-9)$

1. *Condición Estacionaria* $(\nabla_{\mathbf{x}} L = 0)$

$$\frac{\partial L}{\partial x} = 0 \Leftrightarrow \lambda + 2\mu x = 0 \tag{1}$$

$$\frac{\partial L}{\partial y} = 0 \Leftrightarrow 1 + \lambda = 0 \tag{2}$$

$$\frac{\partial L}{\partial z} = 0 \Leftrightarrow \lambda + 2\mu z = 0 \tag{3}$$

2. *Condición de factibilidad*

$$x + y + z = 1 \tag{4}$$

$$x^2 + z^2 \leq 9 \tag{5}$$

3. *Condición de holgura*

$$\mu g(x) = 0 \Leftrightarrow \mu\left(x^2 + z^2 - 9\right) = 0 \tag{6}$$

4. *Condición de positividad o negatividad*

$$\begin{aligned} \mu &\geq 0 \Rightarrow \text{ Para mínimo} \\ \mu &\leq 0 \Rightarrow \text{ Para máximo} \end{aligned}$$

De la ecuación [2] obtenemos directamente

$$\lambda = -1$$

Utilizando ahora la condición de holgura [6] obtenemos dos opciones

$$\begin{aligned} \mu &= 0 \\ x^2 + z^2 - 9 &= 0 \end{aligned}$$

pero la primera opción ($\mu = 0$) no es válida, puesto que si sustituimos en [1] obtenemos

$$\lambda = 0$$

que es una contradicción con el valor anterior que hemos obtenido para λ.

Las ecuaciones que quedan son (sustituyendo el valor de λ)

$$-1 + 2\mu x = 0 \tag{7}$$

$$-1 + 2\mu z = 0 \tag{8}$$

$$x + y + z = 1 \tag{9}$$

$$x^2 + z^2 - 9 = 0 \tag{10}$$

Utilizando [7] y [8] y puesto que $\mu \neq 0$ (¿porqué?) obtenemos

$$x = z$$

que sustituido en [10]

$$x^2 + z^2 = 9 \Leftrightarrow 2x^2 = 9 \Leftrightarrow x = \pm\frac{3}{\sqrt{2}}$$

El valor de y se obtiene de [9]

$$y = 1 - x - z = 1 - 2x = 1 - 2\left(\pm\frac{3}{\sqrt{2}}\right) = 1 \mp 3\sqrt{2}$$

y el valor de μ se obtiene de [7]

$$-1 + 2\mu x = 0 \Leftrightarrow \mu = \frac{1}{2x} = \frac{1}{2\left(\pm\frac{3}{\sqrt{2}}\right)} = \pm\frac{1}{3\sqrt{2}}$$

Se han obtenido 2 puntos

$$P = \left(\frac{3}{\sqrt{2}}, 1 - 3\sqrt{2}, \frac{3}{\sqrt{2}}\right) \qquad \lambda = -1 \qquad \mu = \frac{1}{3\sqrt{2}}$$

$$Q = \left(-\frac{3}{\sqrt{2}}, 1 + 3\sqrt{2}, -\frac{3}{\sqrt{2}}\right) \qquad \lambda = -1 \qquad \mu = -\frac{1}{3\sqrt{2}}$$

Para P obtenemos un valor de $\mu > 0$, por tanto se cumplen las condiciones de Karush-Kuhn-Tucker de mínimo, mientras que para Q obtenemos un valor $\mu < 0$ y por tanto se cumplen las condiciones de máximo.

La función objetivo $f(x, y, z) = y$, es lineal, por tanto es cóncava y convexa (¿por qué?). Por otra parte hay una restricción de igualdad que es afín ($x + y + z = 1$) y la otra restricción de desigualdad es convexa (¿por qué?), luego estamos en condiciones de aplicar el teorema anterior y podemos decir que P y Q son respectivamente el mínimo y máximo globales del problema.

Aunque es posible extender las condiciones necesarias y suficientes a órdenes superiores, en la práctica la aplicación de estas condiciones requiere de un excesivo esfuerzo y solamente tienen una utilidad práctica en el caso de funciones reales de variable real, es decir, cuando $n = 1$ y $m = p = 0$.

Teorema 52 (Condición suficiente de óptimo local) *Supongamos que, para $x^* \in I$, la función $f : I \to \mathbb{R}$ es suficientemente derivable y verifica*

$$\begin{array}{rcll} f^{(k)}(x^*) & = & 0 & \qquad k = 1, \ldots, n-1 \\ f^{(n)}(x^*) & \neq & 0 & \end{array}$$

Donde $f^{(k)}(x)$ es la derivada k–ésima de la función

1. *Si n es impar $\Longrightarrow x^*$ es un punto de inflexión.*
2. *Si n es par $\Longrightarrow x^*$ es un óptimo local. Además*
   a) *Si $f^{(n)}(x^*) > 0 \Rightarrow x^*$ es un mínimo local estricto.*
   b) *Si $f^{(n)}(x^*) < 0 \Rightarrow x^*$ es un máximo local estricto.*

5.10. Interpretación de los multiplicadores de KKT

En esta sección trataremos de explicar de forma no rigurosa el significado de los multiplicadores que aparecen en las condiciones de Karush-Kuhn-Tucker para un problema con restricciones y sus aplicaciones en el análisis de la sensibilidad de los problemas no lineales.Planteemos en primer lugar un problema no lineal con restricciones de igualdad y de desigualdad de la siguiente forma

$$\begin{array}{ll} \text{Optimizar} & f(\mathbf{x}) \\ \text{Sujeto a} & h_i(\mathbf{x}) = b_i \quad i = 1, \ldots, m \\ & g_j(\mathbf{x}) \leq c_j \quad j = 1, \ldots, p \end{array} \tag{5.12}$$

donde $f, h_i, g_j : A \to \mathbb{R}$, son funciones de clase $\mathcal{C}^2(A)$ en $A \subseteq \mathbb{R}^n$ abierto y $b_1, \ldots, b_m, c_1, \ldots, c_p \in \mathbb{R}$.

Está claro que el conjunto factible del problema 5.12 dependerá de los valores de los vectores $\mathbf{b} = (b_1, \ldots, b_m)$ y $\mathbf{c} = (c_1, \ldots, c_p)$, es decir

$$\Omega \equiv \Omega(\mathbf{b}, \mathbf{c})$$

y también es obvio que los puntos óptimos del problema, si existen, dependerán de estos valores

$$\mathbf{x}^* = \mathbf{x}^*(\mathbf{b}, \mathbf{c})$$

Supongamos que para ciertos valores de estos parámetros, $(\mathbf{b}^*, \mathbf{c}^*)$, el problema general con restricciones 5.12 posee un óptimo en el punto $\mathbf{x}^*$, con multiplicadores de Karush-Kuhn-Tucker $\lambda_1, \ldots, \lambda_m$ y $\mu_1, \ldots, \mu_p$ asociados. Podemos definir una función

$$F : U_{(\mathbf{b}^*,\mathbf{c}^*)} \longrightarrow \mathbb{R}$$

con $U_{(\mathbf{b}^*,\mathbf{c}^*)}$ un entorno de $(\mathbf{b}^*, \mathbf{c}^*) \in \mathbb{R}^{m+p}$, de forma que

$$F\left(\overline{\mathbf{b}}, \overline{\mathbf{c}}\right) = f\left(\mathbf{x}^*\left(\overline{\mathbf{b}}, \overline{\mathbf{c}}\right)\right) \qquad \forall \overline{\mathbf{b}}, \overline{\mathbf{c}} \in U_{(\mathbf{b}^*,\mathbf{c}^*)}$$

siendo $\mathbf{x}^*\left(\overline{\mathbf{b}}, \overline{\mathbf{c}}\right)$ el óptimo del programa para cuando se utilizan en el problema 5.12, los términos independientes $\left(\overline{\mathbf{b}}, \overline{\mathbf{c}}\right) \in U_{(\mathbf{b}^*,\mathbf{c}^*)}$.

El siguiente teorema da una relación entre las variaciones del término independiente y las variaciones que experimenta el valor óptimo de la función objetivo.

Teorema 53 *Dado el programa de optimización con restricciones dado en la ecuación 5.12. Si para ciertos valores de los parámetros* $\mathbf{b}$ *y* $\mathbf{c}$, $(\mathbf{b}^*, \mathbf{c}^*) = \left(b_1^*, \ldots, b_m^*, c_1^*, \ldots, c_p^*\right)$ *, el punto* $\mathbf{x}^*$ *es un punto de Karush-Kuhn-Tucker y junto con los multiplicadores asociados,* $\lambda_1, \ldots, \lambda_m$ *y* $\mu_1, \ldots, \mu_p$*; cumple las condiciones de suficiencia para que la función* $f(\mathbf{x})$ *posea en ese punto un extremo relativo sobre el conjunto* $\Omega(\mathbf{b}^*, \mathbf{c}^*)$ *y si no hay restricciones de desigualdad activas degeneradas, entonces*

$$-\lambda_i = \frac{\partial F(\mathbf{b}^*, \mathbf{c}^*)}{\partial b_i} = \frac{\partial f\left(\mathbf{x}^*_{\mathbf{b}^*,\mathbf{c}^*}\right)}{\partial b_i} \qquad i = 1, \ldots, m$$

$$-\mu_j = \frac{\partial F(\mathbf{b}^*, \mathbf{c}^*)}{\partial c_j} = \frac{\partial f\left(\mathbf{x}^*_{\mathbf{b}^*,\mathbf{c}^*}\right)}{\partial c_j} \qquad j = 1, \ldots, p$$

Los multiplicadores λ_i y μ_j, asociados a la i-ésima restricción de igualdad y a la j–ésima restricción de desigualdad respectivamente, nos mide la tasa de variación del valor de la función objetivo $f(x, y)$, en el punto óptimo respecto a la variación de su correspondiente término independiente $(b_i,\ c_j)$.

Notar finalmente que utilizando diferencias finitas obtenemos

$$\Delta F(\mathbf{b}^*, \mathbf{c}^*) = -\sum_{i=1}^{m} \lambda_i \Delta b_i - \sum_{j=1}^{p} \mu_j \Delta c_j = -\sum_{i=1}^{m} \lambda_i \left(\overline{b_i} - b_i^*\right) - \sum_{i=1}^{m} \mu_i \left(\overline{c_j} - c_j^*\right)$$

La ecuación anterior nos proporciona un valor aproximado del incremento que se producirá en el valor del objetivo óptimo al variar el término independiente de las restricciones de $(\mathbf{b}^*, \mathbf{c}^*)$ a $\left(\overline{\mathbf{b}}, \overline{\mathbf{c}}\right)$.

5.11. Dualidad

Sea el siguiente problema de optimización con objetivo de minimización

$$\left.\begin{array}{rcc} \text{Minimizar} & & f(\mathbf{x}) \\ \text{Sujeto a} & h_i(\mathbf{x}) = 0 & i = 1, \ldots, m \\ & g_j(\mathbf{x}) \leq c & j = 1, \ldots, p \end{array}\right\} \quad (P)$$

que llamaremos problema *Primal*, podemos construir el llamado problema *Dual*.

Mientras que en el problema primal los multiplicadores se utilizan como parámetros auxiliares para resolver el problema y se minimiza sobre las variables del problema, en el problema dual los multiplicadores serían las variables, mientras que las variables de decisión del problema primal harían el papel de multiplicadores.

Definición 33 *Para el problema de optimización* P, *definimos su* problema dual (P*) *como*

$$\left.\begin{array}{ll} \textit{Maximizar} & \theta(\boldsymbol{\lambda}, \boldsymbol{\mu}) \\ \textit{Sujeto a} & \boldsymbol{\mu} \geq 0 \end{array}\right\} \quad (P^*)$$

donde la función $\theta(\boldsymbol{\lambda}, \boldsymbol{\mu})$ *está definida por*

$$\theta(\boldsymbol{\lambda}, \boldsymbol{\mu}) = \theta\left(\lambda_1, \ldots \lambda_m, \mu_1 \ldots \mu_p\right) = \inf_{\mathbf{x}\in\Omega}\left\{f(\mathbf{x}) + \lambda_1 h_1(\mathbf{x}) + \ldots + \lambda_m h_m(\mathbf{x}) + \mu_1 g_1(\mathbf{x}) + \ldots + \mu_p g_p(\mathbf{x})\right\}$$

y siendo Ω *el conjunto factible del problema* P.

Ejemplo 22 *Construye el problema dual del siguiente problema de optimización*

$$\begin{array}{ll} \textit{Minimizar} & f(x, y) = \dfrac{x^2}{2} + \dfrac{y^2}{2} \\ \textit{Sujeto a} & x + y \geq 4 \\ & x - y \geq -4 \end{array}$$

Solución: La función Lagrangiana es

$$L(x, y, \mu_1, \mu_2) = \frac{x^2}{2} + \frac{y^2}{2} + \mu_1(-x - y + 4) + \mu_2(-x + y - 4)$$

Si buscamos el mínimo sobre (x, y) en Ω

$$\begin{aligned} \frac{\partial L}{\partial x} &= x - \mu_1 - \mu_2 = 0 \Leftrightarrow x = \mu_1 + \mu_2 \\ \frac{\partial L}{\partial y} &= y - \mu_1 + \mu_2 = 0 \Leftrightarrow y = \mu_1 - \mu_2 \end{aligned}$$

y la función $\theta(\mu_1, \mu_2)$ es

$$\begin{aligned} \theta(\mu_1, \mu_2) &= \frac{(\mu_1 + \mu_2)^2}{2} + \frac{(\mu_1 - \mu_2)^2}{2} + \mu_1(-(\mu_1 + \mu_2) - (\mu_1 - \mu_2) + 4) + \mu_2(-(\mu_1 + \mu_2) + (\mu_1 - \mu_2) - 4) \\ &= -\mu_1^2 - \mu_2^2 + 4\mu_1 - 4_2 \end{aligned}$$

El problema dual será entonces

$$\begin{array}{ll} \text{Maximizar} & \theta(\mu_1, \mu_2) = -\mu_1^2 - \mu_2^2 + 4\mu_1 - 4\mu_2 \\ \text{Sujeto a} & \mu_1, \mu_2 \geq 0 \end{array}$$

Teorema 54 (Teorema de Dualidad) *Dado el problema de optimización* P

$$\begin{array}{rl} \textit{Minimizar} & f(\mathbf{x}) \\ \textit{Sujeto a} & h_i(\mathbf{x}) = 0 \quad i = 1, \ldots, m \\ & g_j(\mathbf{x}) \leq c \quad j = 1, \ldots, p \end{array}$$

con $f(x)$, $g_j(x)$ *convexas y* $h(x)$ *afín. Entonces el problema primal* P *tiene solución si y sólo sí su problema dual* P* *tiene solución. Además en este caso si* $\mathbf{x}^*$ *es la solución óptima del primal y* $(\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$ *la solución óptima del dual se tiene*

$$f(\mathbf{x}^*) = \theta(\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$$

Lema 55 (Lema de Dualidad Débil) *Dado el problema de optimización* P

$$\begin{array}{rl} \textit{Minimizar} & f(\mathbf{x}) \\ \textit{Sujeto a} & h_i(\mathbf{x}) = 0 \quad i = 1, \ldots, m \\ & g_j(\mathbf{x}) \leq c \quad j = 1, \ldots, p \end{array}$$

con $f(x)$, $g_j(x)$ *y* $h(x)$ *de clase* $\mathcal{C}^1(A)$, *entonces:*

1.
$$\text{máx}\{\theta(\boldsymbol{\lambda}, \boldsymbol{\mu}) | \, \boldsymbol{\mu} \geq \mathbf{0}\} \leq \text{mín}\{f(\mathbf{x}) | \, \mathbf{x} \in \Omega\}$$

2. *Si* $(\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$ *son factibles para el problema dual* P* *y* $\mathbf{x}^* \in \Omega$, *y si además se cumple*
$$f(\mathbf{x}^*) = \theta(\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$$
entonces $(\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$ *y* $\mathbf{x}^*$ *son las soluciones óptimas del problema dual y primal respectivamente. Además*
$$\text{máx}\{\theta(\boldsymbol{\lambda}, \boldsymbol{\mu}) | \, \boldsymbol{\mu} \geq \mathbf{0}\} = \text{mín}\{f(\mathbf{x}) | \, \mathbf{x} \in \Omega\}$$

Ejemplo 23 *Resuelve por dualidad el problema*

$$\begin{array}{rc} \textit{Minimizar} & f(x, y) = \dfrac{x^2}{2} + \dfrac{y^2}{2} \\ \textit{Sujeto a} & x + y \geq 4 \\ & x - y \geq -4 \end{array}$$

Solución: Es sencillo comprobar que la función objetivo y las restricciones de desigualdad son convexas, como en este caso no hay restricciones de igualdad podemos aplicar el teorema de dualidad para indicar que este problema tendrá solución si y sólo si la tiene su problema dual, que como vimos en el ejercicio anterior viene dado por

$$\begin{array}{rc} \text{Maximizar} & \theta(\mu_1, \mu_2) = -\mu_1^2 - \mu_2^2 + 4\mu_1 - 4\mu_2 \\ \text{Sujeto a} & \mu_1, \mu_2 \geq 0 \end{array}$$

Para resolver este problema aplicaremos la teoría de KKT, ya que las restricciones son lineales (y por tanto convexas) mientras que la función objetivo es cóncava (¿por qué?), en este caso un punto que cumpla las condiciones de KKT de máximo será la solución. Si llamamos α_1 y α_2 a los multiplicadores de KKT para este problema, el Lagrangiano vendrá dado por

$$L_D(\mu_1, \mu_2, \alpha_1, \alpha_2) = -\mu_1^2 - \mu_2^2 + 4\mu_1 - 4\mu_2 - \alpha_1\mu_1 - \alpha_2\mu_2$$

y las condiciones de KKT serían

1. *Condición estacionaria:* $\nabla_\mu L_D(\mu_1, \mu_2, \alpha_1, \alpha_2)$

$$\begin{aligned}
\frac{\partial L_D}{\partial \mu_1} &= -2\mu_1 + 4 - \alpha_1 = 0 \\
\frac{\partial L_D}{\partial \mu_2} &= -2\mu_2 - 4 - \alpha_2 = 0
\end{aligned}$$

2. *Condición de holgura:*

$$\begin{aligned}
\alpha_1 \mu_1 &= 0 \\
\alpha_2 \mu_2 &= 0
\end{aligned}$$

El resto de condiciones las utilizaremos posteriormente.
De las condiciones de holgura tendremos 4 casos

$$\begin{cases}
\alpha_1 = 0 \Longrightarrow \begin{cases} \alpha_2 = 0 \Longrightarrow \text{ Caso I} \\ \mu_2 = 0 \Longrightarrow \text{ Caso II} \end{cases} \\
\mu_1 = 0 \Longrightarrow \begin{cases} \alpha_2 = 0 \Longrightarrow \text{ Caso III} \\ \mu_2 = 0 \Longrightarrow \text{ Caso IV} \end{cases}
\end{cases}$$

que resolvemos de forma independiente:

1. **Caso I:** $\alpha_1 = \alpha_2 = 0$. Sustituyendo en la primera y segunda ecuaciones, se obtiene

$$\begin{aligned}
-2\mu_1 + 4 &= 0 \Leftrightarrow \mu_1 = 2 \\
-2\mu_2 - 4 &= 0 \Leftrightarrow \mu_2 = -2
\end{aligned}$$

que no es un punto factible puesto que $\mu_2 < 0$.

2. **Caso II:** $\alpha_1 = \mu_2 = 0$. Sustituyendo

$$\begin{aligned}
-2\mu_1 + 4 &= 0 \Leftrightarrow \mu_1 = 2 \\
-4 - \alpha_2 &= 0 \Leftrightarrow \alpha_2 = -4
\end{aligned}$$

que cumples las condiciones y por tanto será el máximo buscado.

3. **Caso III:** $\mu_1 = \alpha_2 = 0$. En este caso

$$\begin{aligned}
4 - \alpha_1 &= 0 \Leftrightarrow \alpha_1 = 4 \\
-2\mu_2 - 4 &= 0 \Leftrightarrow \mu_2 = -2
\end{aligned}$$

que no es factible puesto que $\mu_2 < 0$.

4. **Caso IV:** $\mu_1 = \mu_2 = 0$

$$\begin{aligned} 4 - \alpha_1 &= 0 \Leftrightarrow \alpha_1 = 4 \\ -4 - \alpha_2 &= 0 \Leftrightarrow \alpha_2 = -4 \end{aligned}$$

que no es un punto de KKT de Máximo, puesto que $\alpha_1 > 0$.

Hemos obtenido un único punto para este problema

$$\mu = (2, 0) \qquad \alpha = (0, -4)$$

que como hemos comentado cumple las condiciones de KKT para máximo, luego por las condiciones del problema es la solución buscada.

Teniendo en cuenta ahora la relación existente entre los valores de μ_1 y μ_2 y las variables de decisión del problema primal, podemos hallar el valor de estas últimas

$$\begin{aligned} x &= \mu_1 + \mu_2 = 2 + 0 = 2 \\ y &= \mu_1 - \mu_2 = 2 - 0 = 2 \end{aligned}$$

y comprobar que se trata de un punto factible para el problema primal.

Por último es sencillo comprobar que ambos problemas coinciden en sus respectivos valores óptimos

$$f(x^*, y^*) = f(2, 2) = \frac{2^2}{2} + \frac{2^2}{2} = 4$$
$$\theta(\mu_1^*, \mu_2^*) = \theta(2, 0) = -2^2 - 0^2 + 4 \cdot 2 - 4 \cdot 0 = 4$$

5.12. Temporal

Ejemplo 24 *Encuentra los puntos que cumplen las condiciones de Karush-Kuhn-Tucker para el problema*

$$\begin{array}{ll} \textit{Optimizar} & x^2 + y^2 \\ \textit{Sujeto a} & x + y = 6 \\ & x^2 + y^2 \leq 26 \\ & x - 1 \geq 0 \end{array}$$

Solución: En primer lugar transformamos el problema en la forma general, es decir, los términos independientes de las restricciones deben ser cero y las restricciones de desigualdad de la forma $\leq$

$$\begin{array}{ll} \text{Optimizar} & x^2 + y^2 \\ \text{Sujeto a} & x + y - 6 = 0 \\ & x^2 + y^2 - 26 \leq 0 \\ & 1 - x \leq 0 \end{array}$$

La función Lagrangiana del problema será

$$L(x, y, \lambda_1, \mu_1, \mu_2) = x^2 + y^2 + \lambda(x + y - 6) + \mu_1\left(x^2 + y^2 - 26\right) + \mu_2(1 - x)$$

y planteamos las condiciones de KKT:

1. *Condición estacionaria ($\nabla_x L = 0$)*

$$
\begin{aligned}
\frac{\partial L}{\partial x} &= 0 \Leftrightarrow 2x + \lambda + 2x\mu_1 - \mu_2 = 0 && [1]\\
\frac{\partial L}{\partial y} &= 0 \Leftrightarrow -1 + \lambda + 2y\mu_1 = 0 && [2]
\end{aligned}
$$

2. *Condición de factibilidad*

$$
\begin{aligned}
x + y - 6 &= 0 && [3]\\
x^2 + y^2 - 26 &\leq 0\\
1 - x &\leq 0
\end{aligned}
$$

3. *Condición de positividad o negatividad*

$$
\begin{aligned}
\mu_1, \mu_2 &\geq 0 \Rightarrow \text{Para mínimo}\\
\mu_1, \mu_2 &\leq 0 \Rightarrow \text{Para máximo}
\end{aligned}
$$

4. *Condiciones de holgura*

$$
\begin{aligned}
\mu_1 g_1(x) &= 0 \Rightarrow \mu_1 \left(x^2 + y^2 - 26\right) = 0 && [4]\\
\mu_2 g_2(x) &= 0 \Rightarrow \mu_2 (1 - x) = 0 && ([5])
\end{aligned}
$$

El sistema que permite localizar los puntos de KKT estará formado por las dos ecuaciones que proporciona la condición estacionaria (ecuaciones [1], [2]), la restricción de igualdad ([3]) y las dos de la condición de holgura (ecuaciones [4] y [5]).

Resolvemos el sistema utilizando la condición de holgura. Este análisis produce dos opciones por cada ecuación, con un total de cuatro casos:

$$
\left\{
\begin{array}{ccl}
\mu_1 = 0 & \Longrightarrow & \left\{ \begin{array}{ll} \mu_2 = 0 & \textbf{Caso I} \\ (1 - x) = 0 & \textbf{Caso II} \end{array} \right. \\
 & & \\
x^2 + y^2 - 26 = 0 & \Longrightarrow & \left\{ \begin{array}{ll} \mu_2 = 0 & \textbf{Caso III} \\ (1 - x) = 0 & \textbf{Caso IV} \end{array} \right.
\end{array}
\right.
$$

que resolvemos de forma independiente

1. **Caso I** ($\mu_1 = \mu_2 = 0$): El sistema para estos valores queda

$$
\begin{aligned}
2x + \lambda &= 0\\
-1 + \lambda &= 0\\
x + y &= 6
\end{aligned}
$$

que es lineal y tiene como única solución

$$\begin{aligned}\lambda &= 1\\ x &= -\frac{\lambda}{2} = -\frac{1}{2}\\ y &= 6 - x = 6 + \frac{1}{2} = \frac{13}{2}\end{aligned}$$

Tenemos por tanto un punto para este caso

$$P_1 = \left(-\frac{1}{2}, \frac{13}{2}\right) \quad \lambda = 1 \quad \boldsymbol{\mu} = (0, 0)$$

Sin embargo, este punto no es factible ya que no cumple ninguna de las restricciones de desigualdad del problema

$$x^2 + y^2 = \left(\frac{1}{2}\right)^2 + \left(\frac{13}{2}\right)^2 = \frac{1}{4} + \frac{169}{4} = \frac{85}{2} \nleq 26 \Rightarrow \text{No se cumple}$$

$$1 - x = 1 - \left(-\frac{1}{2}\right) = \frac{3}{2} \nleq 0 \Rightarrow \text{No se cumple}$$

y por tanto no es de KKT.

2. **Caso II** ($\mu_1 = 0$, $x = 1$): Con estos datos el sistema queda

$$\begin{aligned}2 + \lambda - \mu_2 &= 0\\ -1 + \lambda &= 0\\ 1 + y &= 6\end{aligned}$$

que es lineal y cuya única solución es

$$\begin{aligned}y &= 5\\ \lambda &= 1\\ \mu_2 &= 2 + \lambda = 2 + 1 = 3\end{aligned}$$

Obtenemos otro punto

$$P_2 = (1, 5) \quad \lambda = 1 \quad \boldsymbol{\mu} = (0, 3)$$

Comprobamos si es un punto factible

$$\begin{aligned}x^2 + y^2 - 26 &= 1 + 25 - 26 \leq 0 \quad \text{Sí cumple la primera restricción}\\ 1 - x &= 1 - 1 = 0 \leq 0 \quad \text{Sí cumple la segunda restricción}\end{aligned}$$

como además se cumple la condición de positividad, P_2 es un punto que cumple las condiciones de KKT de mínimo.

3. **Caso III** ($x^2 + y^2 = 26$, $\mu_2 = 0$): Para este caso el sistema es

$$\begin{aligned} 2x + \lambda + \mu_1 2x &= 0 \\ -1 + \lambda + \mu_1 2y &= 0 \\ x + y &= 6 \\ x^2 + y^2 - 26 &= 0 \end{aligned}$$

cuya solución se obtiene fácilmente despejando una de las variables de la tercera ecuación, $y = 6 - x$, y sustituyendo en la cuarta para obtener una ecuación de segundo grado

$$x^2 + (6-x)^2 - 26 = 0 \Leftrightarrow 2x^2 - 12x + 10 = 0$$

con soluciones

$$x_1 = 5 \quad \text{y } x_2 = 1$$

Se obtiene un valor de y para cada valor de x

$$\begin{aligned} x_1 &= 5 \Rightarrow y_1 = 6 - x_1 = 1 \\ & \text{y} \\ x_2 &= 1 \Rightarrow y_1 = 6 - x_1 = 5 \end{aligned}$$

Se comprueba la factibilidad de estos dos puntos, $P_3 = (5,1)$ y $P_4 = (1,5)$, sustituyendo en las restricciones de desigualdad

$$\begin{aligned} P_3 &= (5,1) \Rightarrow \begin{cases} 5^2 + 1^2 - 26 = 0 \leq 0 \\ 1 - 5 = -4 \leq 0 \end{cases} \\ P_4 &= (1,5) \Rightarrow \begin{cases} 1^2 + 5^2 - 26 = 0 \leq 0 \\ 1 - 1 = 0 \leq 0 \end{cases} \end{aligned}$$

Finalmente se calculan los valores de los multiplicadores asociados a cada uno de ellos, para determinar si se cumplen algunas de las condiciones de positividad o negatividad y establecer si los puntos son de KKT. Utilizando las dos primeras ecuaciones, que forman un sistema lineal en λ y μ_1 y evaluando en cada punto obtenemos

$$\begin{aligned}
\mu_1 &= \frac{1+2x}{2(y-x)} = \begin{cases} \mu_1(P_3) = -\dfrac{11}{8} \\ \mu_1(P_4) = \dfrac{3}{8} \end{cases} \\
\lambda &= \frac{x(1+2y)}{x-y} = \begin{cases} \lambda(P_3) = \dfrac{15}{4} \\ \lambda(P_4) = -\dfrac{11}{4} \end{cases}
\end{aligned}$$

En resumen, los puntos con sus respectivos multiplicadores son:

$$P_3 = (5,1) \quad \lambda = \frac{15}{4} \quad \mu_1 = \frac{-11}{8} \leq 0 \quad \mu_2 = 0$$

y

$$P_4 = (1,5) \quad \lambda = -\frac{11}{4} \quad \mu_1 = \frac{3}{8} \geq 0 \quad \mu_2 = 0 \geq 0$$

de donde se obtiene que que P_4 es un punto de KKT para el problema de minimización $(\mu_1, \mu_2 \geq 0)$, mientras que P_3 cumple las condiciones de KKT para máximo $(\mu_1, \mu_2 \leq 0)$.

4. **Caso IV ($x^2 + y^2 = 6$, $x = 1$)**: En este último caso queda el siguiente sistema:

$$\begin{aligned}
2 + \lambda + 2\mu_1 - \mu_2 &= 0 \\
-1 + \lambda + \mu_1 2y &= 0 \\
1 + y &= 6 \\
1 + y^2 - 26 &= 0
\end{aligned}$$

De la tercera y cuarta ecuación tenemos el punto

$$P_5 = (1,5)$$

que es uno de los puntos encontrados en el apartado anterior y por tanto ya se ha discutido. Sin embargo, el cálculo de los multiplicadores se obtiene a partir de las dos primeras ecuaciones, en las que al sustituir por los valores de x e y correspondientes obtenemos el sistema

$$\begin{aligned}
2 + \lambda + 2\mu_1 - \mu_2 &= 0 \\
-1 + \lambda + \mu_1 10 &= 0
\end{aligned}$$

que es lineal y con más incógnitas que ecuaciones, por tanto será indeterminado. Su solución es en forma paramétrica

$$\lambda = t; \qquad \mu = \left(\frac{1-t}{10}, \frac{11+4t}{5}\right)$$

Notar, por ejemplo, que si

$$\begin{aligned} \text{Si } t &= 1 \Rightarrow \lambda = 1; \mu = (0,3) \\ \text{Si } t &= -\frac{11}{4} \Rightarrow \lambda = -\frac{11}{4}; \mu = \left(\frac{3}{8}, 0\right) \end{aligned}$$

que corresponden a los multiplicadores de los puntos P_2 y P_4, respectivamente. En todos los casos se trata del mismo punto. El hecho de que existan diversos multiplicadores para el mismo punto es debido, como veremos posteriormente, a que éste problema es singular.

5.13. Ejercicios

1. Demuestra que los hiperplanos de $\mathbb{R}^n$ son conjuntos convexos.

2. Demuestra que todos los semiespacios (positivos y/o negativos) asociados a un hiperplano $H \subseteq \mathbb{R}^n$ son conjuntos convexos.

3. Demuestra el lema 33. Si Ω_1 y Ω_2 son dos subconjuntos convexos no vacíos de $\mathbb{R}^n$ entonces los siguientes conjuntos son convexos:

$$\text{a) } \Omega_1 + \Omega = \{\mathbf{x}_1 + \mathbf{x}_2 \in \mathbb{R}^n : \mathbf{x}_1 \in \Omega_1; \mathbf{x}_2 \in \Omega_2\}$$

$$\text{b) } \Omega_1 - \Omega = \{\mathbf{x}_1 - \mathbf{x}_2 \in \mathbb{R}^n : \mathbf{x}_1 \in \Omega_1; \mathbf{x}_2 \in \Omega_2\}$$

$$\text{c) } \Omega_1 \cap \Omega_2$$

4. Prueba la convexidad del conjunto Ω definido por

$$\Omega = \left\{\mathbf{x} \in \mathbb{R}^n : \sum_{i=1}^{n} x_i = 1,\, x_i \geq 0; i = 1, \ldots, n\right\}$$

5. Demuestra la convexidad del conjunto

$$\Omega = \left\{(x,y) \in \mathbb{R}^2 : y \leq 3 - x^2\right\}$$

6. Comprueba la convexidad del conjunto

$$\Omega = \left\{(x,y) \in \mathbb{R}^2 : y = e^x\right\}$$

7. Siendo $r > 0$ y $\mathbf{a} \in \mathbb{R}^n$, prueba que la bola abierta de centro $\mathbf{a}$ y radio r es un conjunto convexo.

8. Demuestra que $\mathbf{z} \in \mathbb{R}^n$ es combinación lineal convexa de $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ si y sólo si $\mathbf{z} \in [\mathbf{x}, \mathbf{y}]$.

9. Prueba que los siguientes conjuntos son convexos

a) $\Omega_1 = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1 + x_2 + x_3 = 2,\ x_1 \geq 0,\ x_2 \geq 0,\ x_3 \geq 0\}$

b) $\Omega_2 = \{(x_1, x_2) \in \mathbb{R}^2 \mid 3x_1 + 2x_2 \leq 4,\ 6x_1 + x_2 \leq 3,\ x_1 \geq 0,\ x_2 \geq 0\}$

c) $\Omega_3 = \left\{(x_1, x_2) \in \mathbb{R}^2 \mid \frac{x_1^2}{2} + x_2^2 \leq 1,\ x_2 \geq 0\right\}$

d) $\Omega_4 = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1^2 + x_2^2 + x_3^2 \leq 3,\ x_1 + x_2 + x_3 \geq 1\}$

10. Estudia la concavidad o convexidad en $\mathbb{R}^2$ o $\mathbb{R}^3$ de las siguientes funciones

a) $f_1(x_1, x_2) = (x_1 - 3)^2 + (x_2 + 1)^2$

b) $f_2(x_1, x_2) = (x_1 - 2)^2 + x_2^3$

c) $f_3(x_1, x_2) = \begin{cases} \dfrac{x_1 x_2}{x_1^2 + x_2^2} & (x_1, x_2) \neq (0,0) \\ 0 & (x_1, x_2) = (0,0) \end{cases}$

d) $f_4(x_1, x_2, x_3) = x_1^2 - x_2^2 + x_3^2 + x_1 x_2$

11. ¿Será convexa la función $f_2(x_1, x_2)$ del problema anterior sobre el conjunto abierto y convexo

$$\Omega_1 = \left\{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 > 0,\ x_2 > 0\right\}?$$

¿Y $f_4(x_1, x_2)$ sobre el conjunto abierto convexo

$$\Omega_2 = \left\{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1 > 0,\ x_2 > 0,\ x_3 > 0\right\}?$$

12. Para los problemas de optimización siguientes

$$\begin{array}{lc} \text{(a) Optimizar} & x_1 + x_2 \\ \text{sujeto a} & 9x_1 + 2x_2 \leq 6 \\ & 2x_1 + 2x_2 \leq 3 \\ & x_1 \geq 0 \\ & x_2 \geq 0 \end{array} \qquad \begin{array}{lc} \text{(b) Optimizar} & (x_1 - 1)^2 + (x_2 - 1)^2 \\ \text{sujeto a} & x_1 + 2x_2 \leq 4 \\ & x_1 \geq 0 \\ & x_2 \geq 0 \end{array}$$

se pide la resolución razonada de los siguientes apartados:

a) El análisis de su convexidad.

b) ¿Qué se puede decir de sus extremos locales y globales?

c) La resolución geométrica del problema.

13. Dada la función

$$f(x) = x^p \quad 0 < p \leq 1$$

Responde de forma razonada a cada uno de los siguientes apartados

a) Estudia la concavidad o convexidad de la función $f(x)$ sobre el intervalo $I = (0, \infty)$.

b) Utilizando la información del apartado anterior, demuestra la siguiente desigualdad

$$(x+y)^p \geq 2^{p-1}(x^p + y^p) \qquad x, y > 0$$

14. Sea P_1 un problema de programación matemática y sea P_2 otro problema que resulta de anadirle a P_1 una restricción más. Supongamos que el objetivo es maximizar. Si Ω_1 y Ω_2 son respectivamente sus conjuntos factibles, resuelve los siguientes apartados:

a) Indica si existe alguna relación entre los conjuntos Ω_1y Ω_2.

b) Supongamos que $x^* \in \Omega_1$ es solución óptima de P_1 ¿será x^* solución óptima de P_2?

c) Suponiendo que ambos problemas tienen soluciones óptimas x_1^* y x_2^* respectivamente. Encuentra, si existe, la relación entre ellas.

15. Encuentra sobre $\mathbb{R}$, los extremos locales y globales de las siguientes funciones:

$$\text{a) } f(x) = x^5 + x^4 - \tfrac{x^3}{3} + 2 \qquad \text{b) } g(x) = (2x+1)^2(x-4)$$

16. Halla los extremos locales y globales de $f(x) = x^3 - 12x + 3$ en el intervalo $[-4, 4]$.

17. Se construye un barco para transportar L toneladas diarias a lo largo de una ruta de longitud D entre dos puertos. Si el coste de construcción del barco, sin los motores, varía según la capacidad de carga del bargo (l) y el coste de los motores varía como el producto de esta capacidad de carga y el cubo de la velocidad (v) que alcanza el barco, muestra que el coste total de construcción es menor cuando se gasta 2 veces más en el barco que en los motores. (Desprecia el tiempo de carga y descarga y asume que el barco se mueve de forma constante).

18. Un incendio en un bosque está quemando un estrecho valle de $2km$ de ancho a una velocidad de $32km/h$. El fuego puede contenerse mediante un cortafuegos a lo ancho del valle. Si un hombre puede limpiar $2m$ del cortafuegos en 1 minuto, el coste del transporte de cada hombre hasta el cortafuegos es de 12 euros, cada hombre cobra 6 euros/hora, por su trabajo y el valor de la madera es de 1200 euros/km^2 : ¿Cuántos hombres deben enviarse para luchar contra el fuego de manera que el coste sea mínimo?

19. Dada la función

$$f(x, y) = e^{ax+y^2} + b \operatorname{sen}\left(x^2 + y^2\right), \qquad a, b \in \mathbb{R}$$

a) Determina el valor de a y b, sabiendo que $f(x, y)$tiene un extremo relativo en $(0, 0)$ y que el polinomio de Taylor de 2^o orden de $f(x, y)$en ese punto, toma el valor 6 en el punto $(1, 2)$.

b) Indica la clase de extremo que presenta $f(x, y)$ en $(0, 0)$.

20. A partir de la siguiente función:

$$f(x,y) = (3-x)(3-y)(x+y-3)$$

   a) Dibuja el conjunto de puntos del plano en los que $f(x,y)$ es positiva.

   b) Calcula sus puntos estacionarios e indica cuáles son extremos relativos.

   c) Encuentra, si existen, los extremos absolutos o globales de f en $\mathbb{R}^2$.

21. Estudia los extremos relativos y absolutos de las siguientes funciones, en los conjuntos que se indican

$$\begin{aligned} f(x,y) &= x^2 + 3xy - y^2 \\ f(x,y) &= x^3 + 3xy^2 - 15x - 12y \\ f(x,y) &= \operatorname{sen} x + \operatorname{sen} y + \cos(x+y) \quad 0 < x, y < 2\pi \end{aligned}$$

22. Demuestra que el origen es el único punto crítico de la función

$$f(x,y,z) = xy + yz + zx$$

   a) ¿Es un punto de máximo o de mínimo relativo?

   b) Encuentra 2 rectas que pasen por el origen tal que en una de ellas sea $f > 0$ y en la otra $f < 0$. ¿Porqué es posible encontrar estas rectas?

23. **(Optimización de funciones compuestas)** Sea $f : \Omega \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ una función real, sea $h : \mathbb{R} \longrightarrow \mathbb{R}$ una función real monótona creciente. Si definimos la función compuesta como $G = h \circ f$, se pide:

   a) Demuestra que si x^* es un mínimo, máximo o punto de silla de f, también lo es de G.

   b) Si f y h son diferenciables en x^* y $c = f(x^*)$ respectivamente. Demuestra que si $h'(c) \neq 0$ entonces un punto x^* es crítico para G si y sólo si es crítico para f.

   c) Aplica los apartados anteriores para determinar los óptimos locales de la función

$$G(x,y) = \exp\left\{57x(\ln x)^2 + y^2\right\}$$

   sobre $\Omega = \{(x,y) \in \mathbb{R}^2 | \quad x > 0\}$.

24. Encuentra los extremos de las siguientes funciones sobre los conjuntos correspondientes:

$$\begin{array}{ll} f_1(x,y) = xy(1-x^2-y^2) & \text{en } \Omega_1 = [0,1] \times [0,1] \\ f_2(x,y) = xy & \text{en } \Omega_2 = \text{ Triángulo de vértices } (0,0) - (1,0) - (0,1) \end{array}$$

25. Determina, de entre todos los polígonos de n lados que se pueden inscribir en una circunferencia el de área máxima y el de perímetro máximo. (Ayuda: resuelve el problema caracterizando los polígonos por la amplitud de sus ángulos centrales).

26. Un alambre de longitud L se divide en 2 partes, con las que construímos un cuadrado y una circunferencia. ¿Cuál debe ser la longitud de cada una de las partes para que la suma de las áreas del cuadrado y del círculo sea la menor posible?

27. Determina la distancia mínima entre las gráficas de $f(x) = x^2$ y $g(x) = x - 5$.

28. (**Contraejemplo de Peano**) Para a y b constantes reales considera la función

$$f(x,y) = \left(x - a^2y^2\right)\left(x - b^2y^2\right)$$

   a) Clasifica el punto $\mathbf{x} = (0,0)^T$.

   b) Muestra que $f(x,y)$ tiene un máximo en el origen sobre la curva

$$x = \frac{1}{2}\left(a^2 + b^2\right)y^2$$

29. Dada la función

$$f(x,y) = \left[x^2 + (y+1)^2\right]\left[x^2 + (y-1)^2\right]$$

   Clasifica los siguientes puntos: $x_1 = (0,0)$, $x_2 = (0,1)$, $x_3 = (0,-1)$, $x_4 = (1,1)$.

30. Se quiere construir un contenedor para transportar material entre dos puertos. Si la cantidad de material que hay que transportar es de $400m^3$ y los costes del transporte son: 1000 euros cada viaje entre los puertos, 120 euros/m^2 el coste del material de la tapa y el fondo del contenedor y 30 euros/m^2 el material de los lados del contenedor. Resuelve el problema para que el coste sea mínimo.

31. (**Condiciones de orden superior**) ¿Es posible extender las condiciones suficientes para funciones de una variable a funciones multivariantes? es decir, ¿qué les ocurre a los extremos locales con las derivadas de orden superior para una función multivariable?. Comprueba lo qué ocurre en el punto $(0,0)$ y la función $f(x,y) = (y^2 - x)y^2 - 2x$. Aplica el resultado a la función $f(x,y) = x^3 + y^3$.

32. Determina, si existen, los extremos relativos y absolutos de la función $f(x,y,z) = y$ sobre el conjunto

$$F = \left\{x^2 + z^2 = 9; x + y + z = 1\right\}$$

   Determina los valores óptimos aproximados de $f(x,y,z)$ sobre el conjunto

$$F' = \left\{x^2 + z^2 = 9{,}25; x + y + z = 1\right\}$$

33. Determina los puntos de la elipse que se obtienen al cortar el cilindro de ecuación $x^2 + y^2 = 1$ con el plano $x + y + z = 1$, cuyas distancias al origen sean respectivamente máxima y mínima.

34. Dado el problema

$$\begin{array}{rc} \text{Optimizar} & z \\ \text{Sujeto a} & x^2 + y^2 + z^2 = 36 \\ & 2x + y - z = 2 \end{array}$$

Se sabe que los puntos

$$
\begin{aligned}
P_1 &= \left(\frac{10+\sqrt{265}}{15}, \frac{5+\sqrt{265}}{15}, \frac{-1+\sqrt{265}}{3}\right) \\
P_2 &= \left(\frac{10-2\sqrt{265}}{15}, \frac{5-\sqrt{265}}{15}, \frac{-1-\sqrt{265}}{3}\right)
\end{aligned}
$$

son los únicos puntos estacionarios de la función z dentro de la región factible.

a) Estudia si son o no extremos (relativos o globales) condicionados.

b) Encuentra, si existen, los valores óptimos aproximados del problema

$$
\begin{array}{cc}
\text{Optimizar} & z \\
\text{Sujeto a} & x^2+y^2+z^2=36{,}05 \\
 & 2x+y-z=2
\end{array}
$$

35. De entre los triángulos rectángulos de área 9, encuentra aquellos cuya hipotenusa sea mínima.

36. Resuelve el problema

$$
\begin{array}{cc}
\text{Optimizar} & f(x,y,z)=x^2+y^2+z^2 \\
\text{Sujeto a} & x^2+y^2+z^2+xy+yz+zx-1=0 \\
 & x+2y-3z=0
\end{array}
$$

37. Participas en un concurso de televisión en el que se entrega una chapa metálica de $25m^2$ de superficie y con la que debes construir una caja rectangular, que se llenará gratuitamente de gasolina. Se pide:

    a) ¿Cuáles deben ser las dimensiones de la caja que maximizan tu beneficio?

    b) Si el litro de gasolina cuesta 2 euros. ¿Cuánto estarías dispuesto a pagar por $1cm^2$ más de chapa?

38. Halla la mínima distancia entre la recta $x+y=4$ y la circunferencia $x^2+y^2=1$

39. De entre todos los paralelepípedos rectángulos cuya suma de las aristas es la misma e igual a k, determina aquel que tiene volumen máximo

40. Traza por un punto dado un plano que forme con los planos coordenados un tetraedro de volumen mínimo. Se supone el sistema de referencia cartesiano rectangular.

41. Maximiza la distancia al origen de los puntos de la elipse intersección del cilindro de ecuación $x^2+y^2=4$ con el plano $x+y+z=5$.

42. Halla la máxima distancia al origen de los puntos de la circunferencia de centro $(2,1)$ y radio 3.

43. Maximiza la función $f(x,y,z) = 3 + x^2 + 2y^2 + 4y - 2x + (z-2)^2$, sujeta a la restricción $2x + 4y + z = 0$.

44. Resuelve el problema

$$\begin{array}{cc} \text{Optimizar} & x^2y \\ \text{Sujeto a} & x^2 + y^2 = 1 \end{array}$$

45. Determina los óptimos de $f(x,y,z) = xy + yz + zx$, sujeto a la restricción $x + y + z = 120$

46. Resuelve el siguiente problema

$$\begin{array}{cc} \text{Optimizar} & -x^2 + 2y + z \\ \text{Sujeto a} & x + 2y - z = 0 \\ & 2y + z = 0 \end{array}$$

47. Dados los problemas

$$\begin{array}{cc} \text{Minimizar} & x^2 + y^2 \\ \text{Sujeto a} & (x-1)^3 - y^2 = 0 \end{array} \qquad \begin{array}{cc} \text{Minimizar} & (x-2)^2 + y^2 \\ \text{Sujeto a} & (x-1)^3 - y^2 = 0 \end{array}$$

Se pide:

a) Resuélvelos gráficamente.

b) Resuélve cada problema mediante el método de los multiplicadores.

c) Explica porqué no se puede resolver uno de ellos por este método.

48. Un meteoro se mueve a lo largo de la trayectoria de ecuación

$$y = x^2 + 3x - 6$$

Una estación espacial se encuentra en el punto $(x,y) = (2,2)$. Utiliza las condiciones de KKT para encontrar el punto más cercano entre el meteoro y la estación.

49. Se va a manufacturar una remesa de cajas de cartón rectangulares, de forma que las caras superior, inferior y frontal sean de doble peso (es decir, dos piezas de cartón) que las otras. Halla el tamaño de las cajas que maximicen el volumen, teniendo en cuenta que para cada caja se va contar con $72cm^2$ de cartón.

50. Resuelve el siguiente problema

$$\begin{array}{cc} \text{Optimizar} & \dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2} \\ \text{Sujeto a} & px + qy + rz = s \end{array}$$

Donde $a, b, c, p, q, r \in \mathbb{R}$, y $a, b, c \neq 0$.

51. Resuelve el siguiente problema

$$\begin{array}{cc} \text{Optimizar} & \dfrac{a}{x}+\dfrac{b}{y}+\dfrac{c}{z} \\ \text{Sujeto a} & px+qy+rz=s \end{array}$$

donde a, b, c, p, q, $r \in \mathbb{R}^+$.

52. Resuelve el siguiente problema utilizando el método de los multiplicadores

$$\begin{array}{cc} \text{Optimizar} & xyz \\ \text{Sujeto a} & px+qy+rz=s \end{array}$$

donde a, b, c, p, q, $r \in \mathbb{R}-\{0\}$.

53. Calcula los valores máximo y mínimo que alcanza la función $f(x,y)=y-x$ sobre el conjunto $F=\{4x^2+y^2\leq 8; e^{x-y}\leq 1\}$. Indica en qué puntos se alcanzan esos valores. ¿Cómo cambiarían los valores óptimos de la función objetivo si la región fuera $F'=\{4x^2+y^2\leq 8{,}05; e^{x-y}\leq 1{,}05\}$?

54. Un paraboloide elíptico de ecuación

$$x=\frac{y^2}{2p}+\frac{z^2}{2q},\ p>0,\ q>0$$

se corta por un plano de ecuación $x=a$. En la porción de paraboloide así determinada se inscribe un paralelepípedo recto. Determina sus dimensiones para que tenga volumen máximo.

55. Halla los óptimos relativos y absolutos, si los hay, que alcanza $f(x,y,z)=z$, sobre el conjunto $F=\{x^2+y^2\leq 4; x+y+z=5\}$

56. Resuelve el problema

$$\begin{array}{cc} \text{Optimizar} & x^2+2(y-1) \\ \text{Sujeto a} & x+y^2\leq 1 \end{array}$$

57. Realiza la descomposición del número 6 en 3 sumandos, de forma que:

a) Su producto sea máximo.

b) Su producto sea mínimo.

c) ¿Qué se puede decir si los tres números son no negativos?

58. Halla, si existen, los extremos relativos y absolutos de la función $f(x,y)=x^2y$ sobre el conjunto de los puntos que cumplen $x^2+y^2\leq 1$.

59. Resuelve el siguiente problema

$$\begin{array}{cc} \text{Maximizar} & \displaystyle\sum_{j=1}^{N} x_j \\ \text{Sujeto a} & \displaystyle\sum_{j=1}^{N}\frac{x_j^2}{K_j}\leq D \\ & x_j\geq 0;\ \forall j \end{array}$$

donde K_j, D son constantes reales positivas.

60. Sea C el arco de curva intersección de la superficie de ecuación $2z = 16 - x^2 - y^2$ con el plano $x + y = 4$, contenido en el primer octante del espacio $\{x, y, z \geq 0\}$. Encuentra, si existen, los puntos de C, cuya distancia al origen sea máxima y mínima, así como los valores de esas distancia máxima y mínima.

    Si C' es ahora el arco de curva intersección de la superficie de ecuación $2z = 17 - x^2 - y^2$ con el plano $x + y = 3{,}5$, contenido en el primer octante del espacio, calcula los valores de las distancias máxima y mínima de los puntos C' al origen.

61. Encuentra los puntos del conjunto B que están más cerca del origen de coordenada, siendo B:

$$B = \left\{(x, y) \in \mathbb{R}^2 | x + y \geq 4;\ 2x + y \geq 5\right\}$$

    a) Formula y resuelve el problema de forma gráfica.

    b) Comprueba que el punto óptimo cumple las condiciones de K.K.T. ¿Son también suficientes?

62. Resuelve mediante el método de los multiplicadores de K.K.T. el siguiente problema de programación no lineal:

$$\begin{array}{cc} \text{Optimizar} & z \\ \text{Sujeto a} & x^2 + y^2 \leq 8 \\ & x + y + z \geq 1 \\ & x + y + z \leq 4 \geq 0 \end{array}$$

63. Un comerciante se ha enterado de que en la fábrica de aceite hay una oferta de aceite de oliva extra. El aceite que compre a la fábrica lo podrá vender después por litros, ganándose 1 euro por litro. Para almacenar el aceite va a construir un déposito cilíndrico con tapa, utilizando $25\pi m^2$ de chapa. Como desea optimizar sus ganancias, pregunta a su hijo, que es matemático, cuáles son las dimensiones del depósito de máxima capacidad que se podría construir, y obtiene como respuestas $h = 10/\sqrt{6}m$ de altura y $r = 5/\sqrt{6}m$ de radio de las bases.

    Cuando va a iniciar la construcción del depósito, su amigo, el de la ferretería, necesita urgentemente $1m^2$ de chapa como la que él tiene y le propone comprársela. ¿A qué precio debería vendérselo?

64. Dado el siguiente problema de programación no lineal

$$\begin{array}{cc} \text{Optimizar} & (x-4)^2 + (y-2)^2 \\ \text{Sujeto a} & x + y - 2 \leq 0 \\ & -1 + x^2 + (y-1)^2 \end{array}$$

    a) Discute las soluciones del problema mediante los multiplicadores de K.K.T.

    b) Discute los cambios en la solución del problema si ahora la primera restricción se transforma en: $x + y - 2{,}05 \leq 0$

65. Comprueba gráficamente que el punto $(1,0)$ es una solución óptima del siguiete problema, pero que no cumple las condiciones de K.K.T. Explica el resultado:

$$\begin{array}{cc} \text{Maximizar} & x \\ \text{Sujeto a} & y-(1-x)^3 \leq 0 \\ & y \geq 0 \end{array}$$

66. Dado el siguiente problema de programación no lineal

$$\begin{array}{cc} \text{Optimizar} & (x-7)^2+(y-10)^2 \\ \text{Sujeto a} & y-8 \leq 0 \\ & (x-10)^2+(y-10)^2-36 \leq 0 \end{array}$$

a) Discute las posibles soluciones del problema utilizando el método de los multiplicadores de K.K.T.

b) ¿Qué sucede con la solución del mismo si ahora la primera restricción se transforma en: $y-8{,}05 \leq 0$?

67. Dado el siguiente problema de programación no lineal

$$\begin{array}{cc} \text{Optimizar} & x+y \\ \text{Sujeto a} & x^2+\dfrac{y^2}{4} \leq 1 \\ & y-x^2-1 \geq 0 \end{array}$$

Discute las posibles soluciones del problema utilizando el método de los multiplicadores de K.K.T.

68. Dado el siguiente problema de programación no lineal

$$\begin{array}{cc} \text{Optimizar} & (x-2)^2+(y-2)^2 \\ \text{Sujeto a} & y-5 \leq 0 \\ & (x-4)^2+(y-4)^2-16 \leq 0 \end{array}$$

Discute las posibles soluciones del problema utilizando el método de los multiplicadores de K.K.T.

69. Determinar gráficamente el máximo y mínimo absolutos de la función $f(x,y)=x^2+(y-4)^2$ sobre el conjunto $F=\{(x,y)\,|x+y\geq 3; -x+y\leq 3; x\leq 2\}$.

Plantea las condiciones de K.K.T. del problema anterior y utiliza la gráfica para deducir qué restricciones son activas y cuales inactivas en los puntos de máximo y mínimo. Calcular a partir de los datos anteriores esos valores máximo y mínimo. ¿Qué ocurrirá con los valores máximo y mínimo de $f(x,y)$ sobre el conjunto $F=\{(x,y)\}\, x+y\geq 3{,}01; -x+y\leq 3{,}02; x\leq 2{,}05$?