Capítulo 2

Optimización no lineal

2.1 Introducción

En este tema estudiaremos la que hemos llamado cuestión estática del problema de optimización. Trataremos de desarrollar las condiciones necesarias y suficientes para encontrar de forma explícita la solución de problemas de optimización no lineal de tipo general.

Tal y como se introdujo en el tema anterior el problema general de la programación no lineal *(NLPP)* se puede expresar como

$$\begin{array}{ll} \text{Optimizar} & f(\mathbf{x}) \\ \text{Sujeto a} & \\ & h_i(\mathbf{x}) = 0 \quad \forall i = 1, \ldots, m \\ & g_j(\mathbf{x}) \leq 0 \quad \forall j = 1, \ldots, p \end{array} \tag{2.1}$$

donde $f, h_i, g_j : A \to \mathbb{R}$ son funciones reales de varias variables y $A \subseteq \mathbb{R}^n$. En general A es un conjunto abierto y además $f, h_i, g_j \in \mathcal{C}(A)$, es decir, las funciones del problema son derivables.

Si Ω es el conjunto factible del problema 2.1, resolver el problema de optimización *NLPP* 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 Ω de los puntos que cumplen todas las restricciones.

Como se comentó en el primer tema, si $m = p = 0$, entonces no hay restricciones y el problema es de optimización no lineal sin restricciones

$$\begin{array}{ll} \text{Optimizar} & f(\mathbf{x}) \\ & \mathbf{x} \in A \end{array} \tag{2.2}$$

En otro caso ($m \neq 0$ ó $p \neq 0$) el problema se dice con restricciones y a sus extremos locales o globales, se les conoce como *extremos condicionados,* para distinguirlos de los extremos de los problemas sin restricciones.

Recordemos finalmente que si $m \neq 0$ y $p = 0$, es decir, solamente hay restricciones de igualdad en el problema, entonces el problema con restricciones es un *problema de Lagrange*

$$\begin{array}{ll} \text{Optimizar} & f(\mathbf{x}) \\ \text{Sujeto a} & \\ & h_i(\mathbf{x}) = 0 \quad \forall i = 1, \ldots, m \end{array} \tag{2.3}$$

El comportamiento de las restricciones de igualdad y las de desigualdad en un problema de optimización no lineal es ligeramente distinto.

Definición 2.1 (Restricciones activas) *Dado el problema de optimización con restricciones* NLPP *descrito en 2.1; diremos que una restricción de desigualdad* $g_j(\mathbf{x}) \leq 0$ *es* activa *o* saturada *en el punto factible* $\mathbf{x}^* \in \Omega \iff g_j(\mathbf{x}^*) = 0$*; en caso contrario la restricción es* inactiva *o* no saturada *en* $\mathbf{x}^*$*.*

Las restricciones activas se comportan como las restricciones de igualdad, que por su propia naturaleza son activas en cualquier punto factible.

Definición 2.2 *Dado el problema de optimización* NLPP *con restricciones descrito en 2.1 y sea* $\mathbf{x}^* \in \Omega$*, un punto factible. Se define el* conjunto de actividad *asociado a* $\mathbf{x}^*$*,* $J(\mathbf{x}^*)$ *como*

$$J(\mathbf{x}^*) = \{j \in \{1,\ldots,p\}| \quad g_j(\mathbf{x}^*) = 0\}$$

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

Si $\mathbf{x}^*$ es una solución óptima del problema 2.1, las restricciones que sean no activas en él serán irrelevantes puesto que no se alcanza la limitación impuesta por la restricción. Sería posible eliminar las restricciones no saturadas de la formulación del problema, siempre que fueran conocidas previamente pero esto, en general, no es posible.

Definición 2.3 (Punto regular) *Dado el problema de optimización con restricciones 2.1. Sea* $\mathbf{x}^* \in \Omega$*, un punto factible. Se dice que* $\mathbf{x}^*$ *es* regular *para las restricciones si la familia de vectores*

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

es linealmente independiente.

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* $\mathbf{x}^*$ *también será regular si no hay ninguna restricción activa en él, es decir, en problemas con restricciones sólo de desigualdad* $\mathbf{x}^*$ *también es regular si* $J(\mathbf{x}^*) = \emptyset$*.*

Definición 2.4 (Espacio tangente) *Sea* $\mathbf{x}^*$ *un punto factible para el problema de optimización 2.1. Definimos el* espacio tangente *en* $\mathbf{x}^*$ *al conjunto de vectores definido como*

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

o utilizando derivadas

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

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

Caso Particular: *Si* $m = p = 0$*, es decir, si el problema no tiene restricciones, entonces el espacio tangente coincide con el espacio* $\mathbb{R}^n$*.*

2.2 Condiciones de Karush-Kuhn-Tucker

Para los problemas generales de optimización no lineal es posible obtener condiciones necesarias que deban cumplir sus posibles soluciones óptimas, estas condiciones son las llamadas condiciones de Karush-Kuhn-Tucker (CKKT).

Aunque la teoría es indentica en ambos casos, vamos a distinguir entre las condiciones de Karush-Kuhn-Tucker para un problema con objetivo de minimizar y aquellas que se deben cumplir cuando el objetivo es de maximizar.

Minimización

Definición 2.5 *Dado el problema*

$$\begin{array}{lll} \textit{Minimizar} & f(\mathbf{x}) & \\ \textit{Sujeto a} & & \\ & h_i(\mathbf{x}) = 0 & i = 1,\ldots,m \\ & g_j(\mathbf{x}) \leq 0 & j = 1,\ldots,p \end{array} \tag{2.4}$$

con $f, h_i, g_j : A \to \mathbb{R}$ *funciones de clase* $\mathcal{C}^1(A)$ *y* $A \subseteq \mathbb{R}^n$ *un conjunto abierto. Diremos que* $\mathbf{x}^* \in A$ *es un* punto de Karush-Kuhn-Tucker *o que cumple las* condiciones de Karush-Kuhn-Tucker de Mínimo (CKKTMin) *para el problema 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

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

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 positividad

$$\mu_j \geq 0, \qquad i = 1, \ldots, m$$

4. Condición de holgura

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

Los valores de $\lambda_1, \ldots, \lambda_m, \mu_1, \ldots, \mu_p$ son los llamados *multiplicadores.* Entre estos multiplicadores podemos distinguir los *multiplicadores de Lagrange* asociados a las restricciones de igualdad y los *multiplicadores de Karush-Kuhn-Tucker* $\mu_1, \ldots, \mu_p$, asociados a las restricciones de desigualdad.

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

$$L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = f(\mathbf{x}) + \lambda_1 h_1(\mathbf{x}) + \cdots + \lambda_m h_m(\mathbf{x}) + \mu_1 g_1(\mathbf{x}) + \cdots + \mu_p g_p(\mathbf{x}) \tag{2.5}$$

siendo

$$\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_m)$$

$$\boldsymbol{\mu} = (\mu_1, \ldots, \mu_p)$$

y la condición estacionaria se puede expresar como

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

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

Maximización

Se obtienen condiciones equivalentes para un problema de maximización de la forma

$$\begin{array}{ll} \text{Maximizar} & f(x) \\ \text{Sujeto a} & \\ & h_i(\mathbf{x}) = 0 \qquad i = 1, \ldots, m \\ & g_j(\mathbf{x}) \leq 0 \qquad j = 1, \ldots, p \end{array}$$

con la idea de que maximizar una función es equivalente a minimizar su opuesta. Es posible comprobar, aplicando la definición de punto de Karush-Kuhn-Tucker al problema

$$\begin{array}{ll} \text{Minimizar} & g(\mathbf{x}) = -f(x) \\ \text{Sujeto a} & \\ & h_i(\mathbf{x}) = 0 \qquad i = 1, \ldots, m \\ & g_j(\mathbf{x}) \leq 0 \qquad j = 1, \ldots, p \end{array}$$

que la única condición que cambia es la condición de positividad. Para máximo, esta condición se transforma es una *condición de negatividad.* Las condiciones y definiciones correspondientes son las siguientes:

Definición 2.6 *A partir del problema de maximización*

$$\begin{array}{ll} \textit{Maximizar} & f(\mathbf{x}) \\ \textit{Sujeto a} & \\ & h_i(\mathbf{x}) = 0 \qquad i = 1, \ldots, m \\ & g_j(\mathbf{x}) \leq 0 \qquad j = 1, \ldots, p \end{array} \tag{2.6}$$

con $f, h_i, g_j : A \to \mathbb{R}$*, funciones de clase* $\mathcal{C}^1(A)$ *y* $A \subseteq \mathbb{R}^n$ *un conjunto abierto. Diremos que* $\mathbf{x}^* \in A$ *es un* punto de Karush-Kuhn-Tucker *para este problema, o que cumple las* condiciones de Karush-Kuhn-Tucker de Máximo (CKKTMax) *para el problema 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

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

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 negatividad

$$\mu_j \leq 0 \qquad j = 1, \ldots, p$$

4. Condición de holgura

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

¿Cómo encontrar los puntos de *KKT*?

Para la búsqueda práctica de puntos que cumplan las condiciones *CKKT* o puntos de Karush-Kuhn-Tucker, ya sean de mínimo o de máximo, en primer lugar 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*

$$\frac{\partial f}{\partial x_k}(\mathbf{x}^*) + \lambda_1 \frac{\partial h_1}{\partial x_k}(\mathbf{x}^*) + \cdots + \lambda_m \frac{\partial h_m}{\partial x_k}(\mathbf{x}^*) + \mu_1 \frac{\partial g}{\partial x_k}(\mathbf{x}^*) + \cdots + \mu_p \frac{\partial g_p}{\partial x_k}(\mathbf{x}^*) \quad = \quad 0 \qquad k = 1, \ldots, n$$

$$h_i(\mathbf{x}^*) \quad = \quad 0 \qquad i = 1, \ldots, m$$

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

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

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

$$\mu_j g_j(\mathbf{x}^*) = 0 \Leftrightarrow \left\{ \begin{array}{c} \mu_j = 0 \\ \\ g_j(\mathbf{x}^*) = 0 \end{array} \right.$$

Hay que notar que en el caso de p restricciones de desigualdad tendríamo 2^p casos.

Una vez resuelto el sistema, para comprobar cuál o cuáles de las soluciones obtenidas son puntos de *KKT* hay que, por una parte comprobar que son puntos factibles ($g_j(\mathbf{x}^*) \leq 0$) y por otra que sus multiplicadores de KKT asociados tienen todos el mismo signo (≥ 0 para puntos de mínimo, o ≤ 0 para puntos de máximo).

Casos Particulares

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

Problemas sin restricciones Si el problema no tiene restricciones de ningún tipo, entonces $m = p = 0$. El planteamiento del problema sería como en 2.2 y los multiplicadores no serán necesarios, ni tampoco las condiciones relacionadas con ellos. La única condición utilizada es la estacionaria

$$\nabla f(\mathbf{x}^*) = 0 \Leftrightarrow \frac{\partial f}{\partial x_k}(\mathbf{x}^*) = 0 \qquad k = 1, \ldots, n$$

que es la misma para ambos objetivos de maximizar y minimizar.

Cabe distinguir además el caso $n = 1$, es decir, en el caso de la optimización de una función real de variable real, en el que la condición estacionaria nos conduce a un resultado bien conocido

$$x^* \text{ es un extremo local } \Rightarrow f'(x) = 0$$

Problemas de Lagrange Si el problema sólo tiene restricciones de igualdad, es decir $p = 0$ y $m \neq 0$, el problema considerado es un problema clásico de *Lagrange* representado en 2.3 y las condiciones de Karush-Kuhn-Tucker para este problema se obtienen eliminando aquellos términos relacionados con restricciones de desigualdad, para obtener el siguiente sistema

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

$$h_i(\mathbf{x}^*) = 0, \qquad i = 1, \ldots, m$$

que es el resultado que proporciona el teorema clásico de los *multiplicadores de Lagrange*, siendo estos los valores de los λ_i. Se observa como en el caso anterior sin restricciones que las condiciones de KKT para ambos objetivos de maximizar y minimizar coinciden.

Ejemplos

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

Ejemplo 2.7 *Encuentra los puntos de Karush-Kuhn-Tucker para el problema*

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

Solución: Utilizamos los multiplicadores correspondientes para construir la función Lagrangiana

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

y planteamos las condiciones de KKT, previamente hemos expresado las restricciones en la forma usual $g(x) \leq 0$.

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

$$\frac{\partial L}{\partial x} = 0 \Leftrightarrow 1 + 2\mu_2 x = 0 \tag{2.7}$$

$$\frac{\partial L}{\partial y} = 0 \Leftrightarrow 1 + 2\mu_1(y-1) + 2\mu_2(y-1) = 0 \tag{2.8}$$

$$\frac{\partial L}{\partial z} = 0 \Leftrightarrow 1 + 2\mu_1 z + 2\mu_2 z = 0 \tag{2.9}$$

2. *Condición de factibilidad*

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

3. *Condición de positividad o negatividad*

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

4. *Condición de holgura*

$$\mu_1 g_1(x) = 0 \Leftrightarrow \mu_1\left((y-1)^2+z^2-1\right)=0 \tag{2.10}$$

$$\mu_2 g_2(x) = 0 \Leftrightarrow \mu_2\left(x^2+(y-1)^2+z^2-3\right)=0 \tag{2.11}$$

El sistema que hay que resolver estará formado por las ecuaciones 2.7, 2.8, 2.9, 2.10 y 2.11.

1. Tal y como se ha sugerido, empleamos en primer lugar las condiciones de holgura 2.10 y 2.11. Se distinguen cuatro casos

$$\begin{cases} \mu_1=0 \Longrightarrow \begin{cases} \mu_2=0 & \text{Caso I} \\ x^2+(y-1)^2+z^2-3=0 & \text{Caso II} \end{cases} \\ (y-1)^2+z^2-1=0 \Longrightarrow \begin{cases} \mu_2=0 & \text{Caso III} \\ x^2+(y-1)^2+z^2-3=0 & \text{Caso IV} \end{cases} \end{cases}$$

A continuación resolvemos cada caso de forma independiente.

2. **Caso I** ($\mu_1=0,\ \mu_2=0$)**:** Este caso es imposible, puesto que sustituyendo en la ecuación 2.7 obtenemos una contradicción.

3. **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 la siguiente reducción

$$1+2\mu_2 x = 0 \tag{2.12}$$

$$1+2\mu_2(y-1) = 0 \tag{2.13}$$

$$1+2\mu_2 z = 0 \tag{2.14}$$

$$x^2+(y-1)^2+z^2-3 = 0 \tag{2.15}$$

(a) Para encontrar la solución restamos 2.12 y 2.13 para obtener

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

y como μ_2 no puede ser cero puesto que entonces la ecuación 2.7 nos daría $1=0$, se llega a la conclusión de que

$$x-y+1=0 \Rightarrow x=y-1 \tag{2.16}$$

Si ahora restamos 2.12 y 2.14, obtenemos

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

que nos proporciona (teniendo en cuenta de nuevo que $\mu_2 \neq 0$)

$$x=z \tag{2.17}$$

Las relaciones 2.16 y 2.17 llevadas a la expresión 2.15 proporciona una ecuación en la variable z

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

Y por tanto

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

Como $x \neq 0$, despejamos μ_2 de la ecuación 2.12 (o de 2.14)

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

En resumen, para este caso hemos obtenido 2 puntos

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

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

4. **Caso III** $\left((y-1)^2+z^2-1=0,\ \mu_2=0\right)$**:** Este caso también es imposible, puesto que sustituyendo $\mu_2=0$ en la ecuación 2.7 del sistema volvemos a obtener una inconsistencia del tipo $1=0$.

5. **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

$$1+2\mu_2 x = 0 \tag{2.18}$$

$$1+2\mu_1 (y-1)+2\mu_2 (y-1) = 0 \tag{2.19}$$

$$1+2\mu_1 z+2\mu_2 z = 0 \tag{2.20}$$

$$(y-1)^2+z^2-1 = 0 \tag{2.21}$$

$$x^2+(y-1)^2+z^2-3 = 0 \tag{2.22}$$

La operación que debemos hacer a continuación está muy clara, restar las ecuaciones 2.21 y 2.22 para obtener

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

Si ahora sustituimos el valor de x encontrado en 2.18 obtenemos el valor de μ_2

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

Si ahora restamos 2.19 y 2.20

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

Agrupando y sacando factor común obtenemos

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

Esta ecuación nos proporciona 2 opciones. La primera de ellas es

$$\mu_1 + \mu_2 = 0$$

pero utilizando la ecuación 2.20 (o 2.19)

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

que es obviamente imposible.
Nos queda solamente el segundo caso:

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

Sustituimos esta expresión en la ecuación 2.21

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

mientras que y será

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

Por último nos queda por determinar el valor de μ_1, y para ello utilizamos la ecuación 2.20

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

Teniendo en cuenta los distintos valores para μ_2 (2 valores) y z (otros 2) las posibles opciones son

$$\mu_2 = -\frac{1}{2\sqrt{2}} \qquad z = \frac{1}{\sqrt{2}} \Rightarrow \mu_1 = -\frac{1}{2\sqrt{2}} \tag{2.23}$$

$$\mu_2 = -\frac{1}{2\sqrt{2}} \qquad z = -\frac{1}{\sqrt{2}} \Rightarrow \mu_1 = \frac{3}{2\sqrt{2}} \tag{2.24}$$

$$\mu_2 = \frac{1}{2\sqrt{2}} \qquad z = \frac{1}{\sqrt{2}} \Rightarrow \mu_1 = -\frac{3}{2\sqrt{2}} \tag{2.25}$$

$$\mu_2 = \frac{1}{2\sqrt{2}} \qquad z = -\frac{1}{\sqrt{2}} \Rightarrow \mu_1 = \frac{1}{2\sqrt{2}} \tag{2.26}$$

Finalmente se obtienen 4 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 los puntos obtenidos es de KKT (de mínimo o de máximo) tenemos que comprobar su factiblidad y el signo de los multiplicadores. Se expone a continuación una tabla resumen con los resultados obtenidos

$\mathbf{x}$	$\boldsymbol{\mu}$	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

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

por tanto son puntos no 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 2.8 *Encuentra los puntos de Karush-Kuhn-Tucker del siguiente problema*

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

Solución: En este caso $m = 1$ y $p = 0$, es decir hay solamente una restricción de igualdad y el problema es de Lagrange, en este caso las condiciones de Karush-Kuhn-Tucker utilizadas son la condición estacionaria y la condición de factibilidad que nos proporcionan las siguientes ecuaciones

$$\begin{aligned} \nabla f(x) + \lambda_1 \nabla h_1(x) &= 0 \Rightarrow \begin{pmatrix} 2x \\ 2y \end{pmatrix} + \lambda_1 \begin{pmatrix} 2 \\ 1 \end{pmatrix} = 0 \\ h_1(x) &= 0 \Rightarrow 2x + y - 2 = 0 \end{aligned}$$

y un punto será de KKT si es solución del sistema

$$\begin{aligned} 2x + 2\lambda_1 &= 0 \Rightarrow \lambda_1 = -x \\ 2y + \lambda_1 &= 0 \Rightarrow \lambda_1 = -2y \\ 2x + y - 2 &= 0 \end{aligned}$$

que es lineal y cuya solución única es

$$\begin{aligned} \lambda_1 &= -\frac{4}{5} \\ x &= \frac{4}{5} \\ y &= \frac{2}{5} \end{aligned}$$

por tanto el único punto de KKT es

$$P = \left(\frac{4}{5}, \frac{2}{5}\right) \qquad \lambda_1 = -\frac{4}{5}$$

en este caso como el multiplicador está asociado a una restricción de igualdad y su signo no influye en el carácter del punto, por el momento no es posible distinguir si el punto P es un punto de KKT de máximo o de mínimo.

Ejemplo 2.9 *Encuentra los puntos que cumplen las condiciones de Karush-Kuhn-Tucker para el siguiente 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 usual, es decir, los términos independientes de las restricciones deben ser cero y las restricciones de desigualdad de la forma $\leq$

$$x - 1 \geq 0 \Leftrightarrow 1 - x \leq 0$$

Calculamos los vectores gradiente de todas y cada una de las funciones implicadas en el problema

$$\begin{array}{lll} f(x,y) = x^2 - y & \Rightarrow & \nabla f(x,y) = \begin{pmatrix} 2x \\ -1 \end{pmatrix} \\ h_1(x,y) = x + y - 6 & \Rightarrow & \nabla h_1(x,y) = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ g_1(x,y) = x^2 + y^2 - 26 & \Rightarrow & \nabla g_1(x,y) = \begin{pmatrix} 2x \\ 2y \end{pmatrix} \\ g_2(x,y) = 1 - x & \Rightarrow & \nabla g_2(x,y) = \begin{pmatrix} -1 \\ 0 \end{pmatrix} \end{array}$$

Las condiciones de KKT para este problema serían

1. *Condición estacionaria*

$$\nabla f(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i \nabla h_i(\mathbf{x}) + \sum_{j=1}^{p} \mu_j \nabla g_j(\mathbf{x}) = 0 \Leftrightarrow \begin{pmatrix} 2x \\ -1 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \mu_1 \begin{pmatrix} 2x \\ 2y \end{pmatrix} + \mu_2 \begin{pmatrix} -1 \\ 0 \end{pmatrix} = 0$$

que en ecuaciones equivale a

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

2. *Condición de factibilidad*

$$x + y = 6$$

$$1 - x \leq 0$$

$$(x)^2 + (y)^2 \leq 26$$

3. *Condición de positividad o negatividad.*

   Los multiplicadores asociados a desigualdades en los puntos de KKT deben tener el mismo signo.

$$\begin{array}{cccc} \mu_j & \geq & 0 & j = 1, \ldots, p \qquad (CKKTMin) \\ & \text{ó} & & \\ \mu_j & \leq & 0 & j = 1, \ldots, p \qquad (CKKTMax) \end{array}$$

4. *Condiciones de holgura.*

$$\mu_j g_j(\mathbf{x}) = 0 \Leftrightarrow \begin{cases} \mu_1 \left((x)^2 + (y)^2 - 26\right) = 0 \\ \\ \mu_2 (1 - x) = 0 \end{cases}$$

El sistema que proporciona los posibles puntos de KKT sería el siguiente

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

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

$$\begin{cases} \mu_1 = 0 \Rightarrow \begin{cases} \mu_2 = 0 & \text{Caso I} \\ \\ (1 - x) = 0 & \text{Caso II} \end{cases} \\ \\ x^2 + y^2 - 26 = 0 \Rightarrow \begin{cases} \mu_2 = 0 & \text{Caso III} \\ \\ (1 - x) = 0 & \text{Caso IV} \end{cases} \end{cases}$$

A continuación se desarrolla cada caso individualmente.

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 tiene por 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

$$P_1 = \left(\frac{-1}{2}, \frac{13}{2}\right) \quad \lambda = 1 \quad \mu_1 = \mu_2 = 0$$

Si utilizamos las condiciones de factibilidad para comprobar si el punto obtenido pertenece al conjunto factible, sustituyendo en las restricciones de desigualdad se obtiene

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

$$1 - x = 1 + \frac{1}{2} = \frac{3}{2} \nleq 0 \Rightarrow \text{No se cumple}$$

luego este punto no está en el conjunto factible y por tanto no es de KKT.

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

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

cuya solución es

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

y obtenemos otro punto

$$P_2 = (1, 5) \quad \lambda = 1 \quad \mu_1 = 0 \quad \mu_2 = 3$$

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 (*CKKTMin*).

3. *Caso III:* $x^2 + y^2 = 26$, $\mu_2 = 0$. Con estos valores 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 sustituyendo en la cuarta para obtener

$$\begin{aligned} y &= 6 - x \\ x^2 + (6 - x)^2 - 26 &= 0 \end{aligned}$$

desarrollando

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

ecuación de segundo grado con soluciones

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

A continuación se obtiene el valor de y a partir del valor de la variable x para obtener dos puntos

$$\begin{aligned} P_3 &= (5, 1) \\ &\text{y} \\ P_4 &= (1, 5) \end{aligned}$$

Comprobamos a continuación su factibilidad

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

Y por último, es necesario obtener el valor de los multiplicadores para determinar si estos puntos cumplen las condiciones de KKT de mínimo o de máximo. Utilizando las dos primeras ecuaciones, que forman un sistema lineal en λ y μ_1, y evaluando en cada punto obtenemos

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

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 $\big(\mu_j \geq 0\big)$, mientras que P_3 cumple las condiciones de KKT para máximo $(CKKTMax)$.

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 con más incógnitas que ecuaciones y por tanto será indeterminado; su solución es

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

Notar que si

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

que corresponden a los multiplicadores de los puntos P_2 y P_4. No obstante se trata en los tres casos 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.

2.3 Condiciones necesarias

Las condiciones de Karush-Kuhn-Tucker que se han presentado en la sección anterior son necesarias para la mayoría de los problemas de optimización no lineal, es decir, si $\mathbf{x}^*$ es un óptimo local para un problema no lineal, entonces tendrá que cumplir estas condiciones; no obstante, en ocasiones, para garantizar que un óptimo local tiene que cumplir estas 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 2.10 (H.C.R. Sin Restricciones) *En un problema de optimización no lineal sin restricciones, todos los puntos cumplen la primera hipótesis de cualificación de las restricciones.*

Definición 2.11 (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 Karlin.*

Definición 2.12 (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 Slater.*

Definición 2.13 (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* Fiacco-McKormik.

Condiciones necesarias de primer orden

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

Teorema 2.14 (Condiciones necesarias de primer orden) *Dado el problema general de la optimización no lineal (*NLPP*)*

$$\begin{array}{ll} \textit{Optimizar} & f(\mathbf{x}) \\ \textit{Sujeto a} & \\ & h_i(\mathbf{x}) = 0 \quad i = 1, \ldots, m \\ & g_j(\mathbf{x}) \leq 0 \quad j = 1, \ldots, p \end{array} \tag{2.27}$$

*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 [máximo] relativo $\Longrightarrow \mathbf{x}^*$ cumple las condiciones de Karush-Kuhn-Tucker de Minimización [Maximización].*

Ejemplos

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

Ejemplo 2.15 *Resuelve el problema*

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

Solución: Como solamente tiene una restricción de igualdad se trata de un problema de Lagrange; además dicha restricción es lineal, luego se cumple la hipótesis de cualificación de Karlin en todos los puntos del conjunto factible, luego cualquier mínimo local debe cumplir las condiciones de Karush-Kuhn-Tucker. Por ser un problema de Lagrange estas condiciones se reducen a la condición estacionaria y a la condición de factibilidad:

$$\nabla f(x, y, z) + \lambda \nabla h(x, y, z) = \begin{pmatrix} y + z \\ z + x \\ x + y \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = 0$$

$$h(x, y, z) = x + y + z - 3 = 0$$

Ambas condiciones se transforman en el siguiente sistema lineal

$$y + z + \lambda = 0$$

$$x + z + \lambda = 0$$

$$y + x + \lambda = 0$$

$$x + y + z - 3 = 0$$

cuya única solución es

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

y por tanto este es el único punto que cumple las condiciones de KKT.

Notar que aunque $\lambda_1 = -2 < 0$ y el objetivo sea de minimizar, el punto es de KKT puesto que λ es un multiplicador asociado a una restricción de igualdad, y por tanto no está condicionado por su signo. Por otra parte, no es posible determinar la naturaleza del punto, puesto que se cumplen las condiciones de KKT para mínimo, pero también se cumplen para máximo.

Ejemplo 2.16 *Aplica las condiciones necesarias de primer orden al problema*

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

Solución: En la sección anterior se comprobó que los únicos puntos que cumplían las condiciones de CKKT eran

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

y

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

Ahora comprobaremos que si el problema tuviera solución (máximo y/o mínimo) entonces debería ser alguno de los puntos anteriores. El razonamiento es el siguiente: *"Si el problema tiene mínimo (máximo) global, entonces también será un mínimo (máximo) local. Si además también se cumpliera alguna de las hipótesis de cualificación, entonces este mínimo (máximo) local debería ser un punto de Karush-Kuhn-Tucker y por tanto debería ser alguno de estos dos puntos".*

Es fácil comprobar que se cumple la hipótesis de cualificación de Slater, ya que el conjunto factible Ω es

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

siendo

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

Los conjuntos Ω_1 y Ω_2 son convexos puesto que son conjuntos de la forma

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

con $g(x,y,z)$una función convexa (¿porqué?).

Como Ω es intersección de dos conjuntos convexos, será él mismo un conjunto convexo. Para comprobar la condición de Slater solamente 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.

Sólo hay que comprobar que el problema tiene solución, pero esta también es directo aplicando el teorema de Weierstrass, ya que $f(x,y,z)$ es continua y el conjunto es compacto (¿porqué?).

La existencia de solución al problema implica que los puntos P_3 y P_6 son los extremos buscados. Para decidir cuál de ellos es el mínimo y cuál el máximo, lo único que queda evaluar la función objetivo en cada uno de ellos

$$f(P_3) = \left(\sqrt{2}\right) + \left(1 + \frac{1}{\sqrt{2}}\right) + \left(\frac{1}{\sqrt{2}}\right) = 1 + 2\sqrt{2}$$

y

$$f(P_6) = \left(-\sqrt{2}\right) + \left(1 - \frac{1}{\sqrt{2}}\right) + \left(-\frac{1}{\sqrt{2}}\right) = 1 - 2\sqrt{2}$$

luego P_3 es el máximo del problema y P_6 su mínimo.

La aplicación de las condiciones necesarias sobre este ejercicio permite descartar otros posibles candidatos a solución.

Ejemplo 2.17 *Plantea y resuelve el problema de construir una caja de cartón de volumen máximo y área fija A.*

Solución: Si x, y, z son las dimensiones de la caja, el problema no lineal se puede expresar como

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

siendo $A > 0$ el área fija de la caja.

La restricción no cumple ni la hipótesis de Karlin (la única restricción es no lineal), ni la de Slater (el conjunto no es convexo). Comprobaremos que se cumple la hipótesis de regularidad en todos los puntos; para ello calculamos el gradiente de la restricción

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

Como sólo hay un vector, éste será linealmente dependiente sólo cuando sea el vector nulo, es decir si ocurre

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

El único punto solución del sistema anteriore es el vector nulo

$$x = y = z = 0$$

Sin embargo, sustituyendo en la restricción comprobamos facilmente que 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. Este resultado implica que cualquier extremo local del problema debe cumplir las condiciones de Karush-Kuhn-Tucker, que expresamos a continuación:

1. *Condición estacionaria*

$$yz + \lambda(y+z) = 0$$

$$xz + \lambda(x+z) = 0$$

$$xy + \lambda(x+y) = 0$$

2. *Condición de factibilidad*

$$xy + yz + zx - \frac{A}{2} = 0$$

El sistema formado por las cuatro ecuaciones anteriores tiene como única solución a:

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

De nuevo al ser un problema solamente con restricciones de igualdad, no podemos determinar si el punto es de mínimo o de máximo.

Ejemplo 2.18 (Entropía) *Sea X una variable aleatoria discreta que toma valores dentro del conjunto $\{x_1, \ldots, x_n\}$, sabiendo que el valor medio obtenido para X es $\overline{x}$, el problema consiste en encontrar la probabilidad p_j de que X tome el valor x_j de forma que la entropía sea máxima.*

Solución: Para una variable aleatoria discreta se define la entropía de la densidad como

$$\varepsilon = -\sum_{j=1}^{m} p_j \ln p_j$$

mientras que la esperanza de X es

$$\mu = \sum_{j=1}^{m} x_j p_j$$

Si la probabilidad asociada a cada x_j es p_j, entonces debe cumplirse $p_j \geq 0$ y $\sum_{j=1}^{m} p_j = 1$. Si además sabemos cual es la media $(\overline{x})$, el argumento de entropía máxima requiere que la densidad debe tomarse como aquella que resuelve el problema

$$\begin{array}{ll} \text{Maximizar} & -\sum_{i=1}^{m} p_j \ln p_j \\ \\ \text{Sujeto a} & \sum_{i=1}^{m} p_j = 1 \\ \\ & \sum_{j=1}^{m} x_j p_j = \overline{x} \\ \\ & p_j \geq 0 \end{array}$$

Como todas las restricciones del problema son lineales se cumple la hipótesis de Karlin, luego cualquier solución óptima debe cumplir las condiciones de Karush-Kuhn-Tucker. Es decir, si el problema tiene solución, deben existir multiplicadores λ_1,λ_2,$\mu_1, \ldots, \mu_n$ cumpliendo las siguientes condiciones:

1. *Condición estacionaria:*

$$\nabla f(\mathbf{p}) + \lambda_1 \nabla h_1(\mathbf{p}) + \lambda_2 \nabla h_2(\mathbf{p}) + \sum_{j=1}^{n} \mu_j \nabla g_j(\mathbf{p}) = 0$$

En este caso

$$\begin{aligned} f(\mathbf{p}) &= -\sum_{i=1}^{n} p_i \ln p_i \Rightarrow \nabla f(\mathbf{p}) = (-\ln p_1 - 1, \ldots, -\ln p_n - 1) \\ h_1(\mathbf{p}) &= \sum_{i=1}^{n} p_i - 1 \Rightarrow \nabla h_1(\mathbf{p}) = (1, \ldots, 1) \\ h_2(\mathbf{p}) &= \sum_{i=1}^{n} x_i p_i - m \Rightarrow \nabla h_2(\mathbf{p}) = (x_1, \ldots, x_n) \\ g_j(\mathbf{p}) &= -p_j \Rightarrow \nabla g_j(\mathbf{p}) = \left(0, \ldots, \overbrace{-1}^{j}, \ldots 0_n\right) \qquad j = 1, \ldots, n \end{aligned}$$

y las ecuaciones que tienen que cumplir los extremos son

$$-\ln p_i - 1 + \lambda_1 + \lambda_2 x_i - \mu_i = 0 \qquad i = 1, \ldots, n \tag{2.28}$$

2. *Condición de factibilidad:*

$$\sum_{i=1}^{n} p_i = 1$$

$$\begin{aligned}\sum_{i=1}^{n} x_i p_i &= \overline{x}\\ p_i &\geq 0\end{aligned}$$

3. *Condición de negatividad:* Puesto que estamos buscando el máximo de la función, los multiplicadores asociados a las restricciones de desigualdad deben ser negativos, es decir, debe ocurrir

$$\mu_j \leq 0$$

4. *Condición de holgura:*

$$\mu_j g_j(\mathbf{p}) = 0 \Leftrightarrow \mu_j(-p_j) = 0$$

De la condición de holgura se deduce que

$$\mu_j = 0$$

puesto que en caso contrario, $p_j = 0$, pero para este valor la función logaritmo no está definida. Por tanto todos los multiplicadores μ_j, asociados a las restricciones de desigualdad deben ser nulos. Sustituyendo en la familia de ecuaciones 2.28 obtenemos

$$-\ln p_j - 1 + \lambda_1 + \lambda_2 x_j = 0 \qquad j = 1, \ldots, n$$

y podemos despejar p_j de cada ecuación $j = 1, \ldots, n$

$$\ln p_j = -1 + \lambda_1 + \lambda_2 x_j \Rightarrow p_j = e^{\lambda_2 x_j + (\lambda_1 - 1)}$$

Observemos que $p_j > 0$.

Los valores de λ_1 y λ_2 deben seleccionarse de forma que se cumplan las 2 restricciones

$$\begin{aligned}\sum p_j &= 1\\ \sum p_j x_i &= \overline{x}\end{aligned}$$

El teorema de las condiciones necesarias nos proporciona los requisitos que deberían 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 ejemplos.

Ejemplo 2.19 *Consideremos el problema*

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

El conjunto factible del problema es

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

y por tanto un problema equivalente al anterior viene dado por

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

cuya solución óptima podemos obtener fácilmente y es $x^* = 0$. La solución óptima del problema incial, en $\mathbb{R}^2$ será $P = (0, 0)$.

Vamos a comprobar si P es un punto de Karush-Kuhn-Tucker. Como el problema solamente tiene restricciones de igualdad, las condiciones de KKT suponen que debería existir un multiplicador λ de forma P cumplan las siguientes ecuaciones

$$\begin{aligned} \nabla f(P) + \lambda \nabla h(P) &= 0 \Rightarrow \begin{pmatrix} -2x \\ 1 \end{pmatrix} + \lambda \begin{pmatrix} 0 \\ 3y^2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ h(P) &= 0 \Rightarrow y^3 = 0 \end{aligned}$$

es decir

$$\begin{aligned} -2x &= 0 \\ 1 + 3\lambda y^2 &= 0 \\ y^3 &= 0 \end{aligned}$$

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) para el problema de optimización planteado, pero no cumple las condiciones de Karush-Kuhn-Tucker. Comprobemos que para P no se cumplen ninguna de las hipótesis de cualificación descritas.

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 no lineal.*

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

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

es decir el eje X, que en $\mathbb{R}^2$ es un conjunto con interior vacío, puesto que cualquier bola de centro un punto de Ω tiene puntos fuera de Ω y por tanto 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*

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

y al evaluar en $P = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ obtenemos

$$\{\nabla h(P)\} = \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\}$$

que por ser el vector nulo, es linealmente dependiente; por tanto $P = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ es no regular para las restricciones.

Si en el 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 tenemos que ser capaces de comprobar que el punto $P = (0,0)$ cumple las condiciones de KKT. El sistema ahora 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 2.20 *Sea el problema*

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

El conjunto factible para este problema es

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

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

Al plantear la condición estacionaria de Karush-Kuhn-Tucker para este problema

$$\nabla f(\mathbf{x}) + \lambda_1 \nabla h_1(\mathbf{x}) + \lambda_2 \nabla h_2(\mathbf{x}) = 0$$

obtenemos

$$\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \lambda_1 \begin{pmatrix} 2(x-1) \\ 2y \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 2(x+1) \\ 2y \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

y los puntos críticos deben cumplir entonces el siguiente sistema de ecuaciones:

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

sin embargo, ningún punto de la forma $(0,0,z)$ *es solución del sistema anterior, puesto que al sustituir estos puntos en la segunda ecuación obtenemos una contradicción y ninguno de los extremos locales del problema cumple las condiciones de KKT.*

Podemos comprobar, como en el caso anterior, que ninguno de ellos cumple alguna de las hipótesis de cualificación. Es un problema con restricciones no lineales donde $\overset{\circ}{\Omega} = \emptyset$ *(*Ω *es una recta), es decir, no se cumple ninguna de las tres primeras condiciones dada. Para la condición de regularidad (Fiacco-McKormik) 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 un extremo local de la función objetivo donde sí se cumplan las condiciones de Karush-Kuhn-Tucker. Consideramos, por ejemplo, las restricciones del ejemplo anterior, pero cambiamos la función objetivo por $f(x,y,z) = x$.

Para este caso, las condiciones de KKT serán

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \lambda_1 \begin{pmatrix} 2(x-1) \\ 2y \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 2(x+1) \\ 2y \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

o equivalentemente serán soluciones del sistema

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

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 alguna 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.

Observación 2.21 *Es posible comprobar que en el punto* $(5,1)$ *del ejemplo 2.9 es un punto donde no se cumple ninguna de las hipótesis de cualificación.*

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 2.22 *Los puntos 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 para obtener el siguiente resultado:

Teorema 2.23 (Condiciones necesarias de 2^o orden) *Dado el problema general de optimización*

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

*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 [máximo] relativo condicionado $\Longrightarrow \mathbf{x}^*$ es un punto que cumple las condiciones de Karush-Kuhn-Tucker para el problema de minimizar [maximizar] 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 positividad [negatividad]

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

4. Condición de holgura

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

y además se cumple la siguiente condición:

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

$$HL(\mathbf{x}^*, \boldsymbol{\lambda}, \boldsymbol{\mu}) = Hf(\mathbf{x}^*) + \sum_{i=1}^{m} \lambda_i Hh_i(\mathbf{x}^*) + \sum_{j=1}^{p} \mu_j Hg_j(\mathbf{x}^*) \tag{2.29}$$

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

Casos Particulares:

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 2.29 se convierte en*

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

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

$$Hf(\mathbf{x}^*) \text{ es semidefinida positiva [negativa respectivamente]}$$

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 2.24 *Aplica las condiciones necesarias de segundo orden al siguiente problema:*

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

Solución: Recordemos que para este problema se habían obtenido dos puntos de Karush-Kuhn-Tucker.

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

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

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 es un subespacio vectorial de $\mathbb{R}^3$, $M(P_3) \subset \mathbb{R}^3$, la matriz $HL(P_3)$ también será semidefinida negativa sobre él y por tanto se cumplen las condiciones necesarias para que P_3 pueda ser un máximo.

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 el espacio tangente es un subespacio vectorial de $\mathbb{R}^3$, $M(P_6) \subset \mathbb{R}^3$, la matriz $HL(P_6)$ también será semidefinida positiva sobre él y por tanto se cumplen las condiciones necesarias para que P_6 pueda ser un máximo.

Ejemplo 2.25 *Aplica las condiciones de segundo orden al problema de la caja de la sección anterior.*

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 como el espacio tangente en P, $M(P)$

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

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

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

Mientras que el espacio tangente es

$$M(P) = \left\{ \mathbf{d} \in \mathbb{R}^3 \middle| \nabla h(P)^T \mathbf{d} = 0 \right\}$$

Utilizando el valor de P y el gradiente de $h(\mathbf{x})$ tenemos

$$\nabla h\left(\sqrt{\frac{A}{6}}, \sqrt{\frac{A}{6}}, \sqrt{\frac{A}{6}}\right) = 2\sqrt{\frac{A}{6}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

y el espacio tangente está descrito por

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

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 al ser una diferencia de cuadrados, tomará 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.

2.4 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 2.26 *Comprueba que el punto* $P = (0, 0)$ *cumple las condiciones necesarias de primer y segundo orden para el problema*

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

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

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

que 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 evaluamos en P

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

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 esa curva sucede

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

y $(0, 0)$ 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 2.27 *Dado el problema general con restricciones*

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

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* $\mathbf{x}^* \Longleftrightarrow g_j(\mathbf{x}^*) = 0$ *y* $\mu_j = 0$.

Definición 2.28 *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 2.29 *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

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

y para un problema de Lagrange

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

Teorema 2.30 (Condiciones suficientes) *Dado el problema general de optimización*

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

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 para el problema de minimizar [maximizar], 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 positividad [negatividad]

$$\mu_j \geq 0 \quad [\mu_j \leq 0] \qquad j = 1, \ldots, p$$

4. Condición de holgura

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

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] sobre el espacio tangente ampliado $\widetilde{M}(\mathbf{x}^*)$.

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:

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 2.31 *Resuelve el problema*

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

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

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

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 $\left.\varphi\right|_{Hf(x^*)}(\mathbf{d})$ 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 2.32 (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 2.33 (Problemas con desigualdades) *Dado el problema* NLPP

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

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* NLPP.

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

Proposición 2.34 (Problemas Afines Convexos) *Dado el problema* NLPP

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

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* $h_1(\mathbf{x})$*, …,* $h_m(x)$ *son funciones afines y* $g_1(\mathbf{x})$*, …,* $g_p(\mathbf{x})$ *son funciones convexas sobre* Ω *entonces si existe un* $\mathbf{x}^* \in \Omega$ *punto* CKKTMin *[*CKKTMax*] entonces* $\mathbf{x}^*$ *es solución del problema* NLPP.

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 2.35 *Resuelve el siguiente problema* NLPP

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

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

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

$$
\begin{aligned}
\frac{\partial L}{\partial x} &= 0 \Leftrightarrow \lambda + 2\mu x = 0 \\
\frac{\partial L}{\partial y} &= 0 \Leftrightarrow 1 + \lambda = 0 \\
\frac{\partial L}{\partial z} &= 0 \Leftrightarrow \lambda + 2\mu z = 0
\end{aligned}
$$

2. *Condición de factibilidad*

$$
\begin{aligned}
x + y + z &= 1 \\
x^2 + z^2 &\leq 9
\end{aligned}
$$

3. *Condición de positividad o negatividad*

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

4. *Condición de holgura*

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

De la segunda ecuación obtenemos directamente

$$\lambda = -1$$

Utilizando ahora la condición de holgura 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 la primera de las ecuaciones estacionarias 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)

$$\begin{aligned} -1 + 2\mu x &= 0 \\ -1 + 2\mu z &= 0 \\ x + y + z &= 1 \\ x^2 + z^2 - 9 &= 0 \end{aligned}$$

Utilizando las dos primeras ecuaciones y puesto que μ no puedes ser 0 (¿porqué?) obtenemos

$$x = z$$

que sustituido en la última ecuación

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

El valor de y se obtiene de la tercera ecuación

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

y el valor de μ se obtiene de la primera ecuación

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

Hemos 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 2.36 (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.*

2.5 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

$$\begin{array}{ll} \text{Minimizar [Maximizar]} & f(\mathbf{x}) \\ \text{Sujeto a} & \\ & h_i(\mathbf{x}) = b_i \quad i = 1, \ldots, m \\ & g_j(\mathbf{x}) \leq c_j \quad j = 1, \ldots, p \end{array} \tag{2.30}$$

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 2.30 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 2.30 posee un óptimo en el punto $\mathbf{x}^*$, con multiplicadores de Karush-Kuhn-Tucker $\left(\lambda_1, \ldots, \lambda_m, \mu_1, \ldots, \mu_p\right)$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 2.30, 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 2.37 *Dado el programa de optimización con restricciones dado en la ecuación 2.30. 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 $\left(\overline{\mathbf{b}},\overline{\mathbf{c}}\right)$.

2.6 Dualidad

Para cualquier problema de optimización *NLPP*

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

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 2.38 *Para el problema de optimización* NLPP, *definimos su* problema dual (NLPP*) *como*

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

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* NLPP.

Ejercicio 2.6.1 *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 2.39 (Teorema de Dualidad) *Dado el problema de optimización* NLPP

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

con $f(x)$, $g_j(x)$ convexas y $h(x)$ afín. Entonces el problema primal NLPP *tiene solución si y sólo sí su problema dual* NLPP* *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 2.40 (Lema de Dualidad Débil) *Dado el problema de optimización* NLPP

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

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

1.

$$\max\{\theta(\boldsymbol{\lambda}, \boldsymbol{\mu}) | \boldsymbol{\mu} \geq \mathbf{0}\} \leq \min\{f(\mathbf{x}) | \mathbf{x} \in \Omega\}$$

2. Si $(\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$ *son factibles para el problema dual* NLPP* *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*

$$\max\{\theta(\boldsymbol{\lambda}, \boldsymbol{\mu}) |\, \boldsymbol{\mu} \geq \mathbf{0}\} = \min\{f(\mathbf{x}) |\, \mathbf{x} \in \Omega\}$$

Ejemplo 2.41 *Resuelve por dualidad el problema*

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

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

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

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 cumle 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

$$f(x^*, y^*) = f(2, 2) = \frac{2^2}{2} + \frac{2^2}{2} = 4$$

$$\theta(\mu_1^*, \mu_2^*) = \theta(2, 0) = -2^2 - 0^2 + 4 \cdot 2 - 4 \cdot 0 = 4$$