Apuntes de Optimización no lineal

Jose S. Cánovas
Departamento de Matemática Aplicada y Estdistica,
Universidad Politécnica de Cartagena

4 de noviembre de 2019

Capítulo 1

Optimización no lineal

Para la redacción de estos apuntes me he basado bastante en los de mi compañero Silvestre Paredes, con el que compartí esta asignatura varios años y del que he tomado prestado algún ejemplo que otro, así como la lista de problemas del final. Si quiere consultar su versión mucho más extensa y con bastante más información, se puede hacer en su web http://www.dmae.upct.es/~paredes/

1.1. Introducción

Este tema está dedicado a estudiar los elementos esenciales de la optimiación no lineal desde un punto de vista práctico. Aunque la convexidad juega un papel muy importante dentro de la teoría de optimización, en la práctica para los problemas eminentemente prácticos que vamos a abordar en este tema, dicha convexidad va a quedar enmascarada en los resultados principales que vamos a desarrollar.

A modo de ejemplo, consideremos el problema bien conocido: "De entre todos los rectángulos de lados x e y y perímetro $100cm$. indicar el que tiene área máxima". Para analizar este problema hemos de poner los datos de los que disponemos:

1. El área del rectángulo $A = x \cdot y$.

2. El perímetro del rectángulo en $p = 2x + 2y$.

De esta forma, podemos plantear el problema de la forma

$$\left\{ \begin{array}{c} \text{Maximizar } A(x,y) = x \cdot y, \\ \text{Sujeto a } 100 = 2x + 2y. \end{array} \right.$$

Este problema es bastante sencillo de resolver despejando la variable $y = 50 - x$ de la segunda igualdad y sustituyendo en la función área $A(x, y)$ de modo que pasa a ser una función unidimensional, es decir, $A(x) = 50x - x^2$. Como sabemos del bachillerato, para obtener los extremos relativos de la función $A(x)$ debemos plantear la ecuacion

$$A'(x) = 50 - 2x = 0,$$

de donde inmediatamente tenemos $x = 25$. Como además la segunda derivada $A''(x) = -2 < 0$, tenemos que el punto que hemos hallado es la solución de dicha ecuación, por lo que el área máxima que queríamos calcular es

$$A(25, 25) = 625.$$

Otro ejemplo bien conocido que entronca con los contenidos de bachillerato es del de calcular los máximos y mínimos de la función $f(x) = x^2 + 1$ cuando $x \in [-1, 2]$. Este problema lo podemos escribir de la forma

$$\begin{cases} \text{Optimizar } f(x) = x^2 + 1, \\ \text{Sujeto a } \begin{cases} x \geq -1, \\ x \leq 2. \end{cases} \end{cases}$$

Este problema también hemos de ser capaces de resolverlo a partir de los contenidos cursados por el alumno en el bachillerato. Calculamos en primer lugar los extremos locales en el intervalo $(-1, 2)$ que es dónde la función es derivable. Igualando

$$f'(x) = 2x = 0,$$

obtenemos que $x = 0$ es un mínimo local dado que la segunda derivada $f''(x) = 2 > 0$. Este mínimo local daría el valor $f(0) = 1$. En los puntos de no derivabilidad calculamos el valor de la función $f(-1) = 2$ y $f(2) = 5$. Como resumen, el problema tiene un mínimo global en $f(0) = 1$ y máximo global $f(2) = 5$.

En general, todos los problemas de optimización que vamos a tratar tienen una función

$$f(x_1, ..., x_n),$$

que supondremos siempre suficientemente regular, de la que pretendemos obtener su máximo o su mínimo, o de forma más general, optimizar dicha función f, es decir, decidir cuáles son sus máximos o mínimos. Dicha función f se llama función objetivo del problema de optimización.

Normalmente, dicha optimización puede estar sujeta a dos tipos de ligaduras. que vienen impuestas normalmente por la condiciones del problema. Las ligaduras tipo pueden ser de igualdad, dadas por una función

$$g(x_1, ...x_n) = 0,$$

o bien de desigualdad, cuando lo que tenemos es una expresión de la forma

$$h(x_1, ..., x_n) \leq 0.$$

En un problema de optimización podemos tener varias ligaduras de igualdad, y varias de desigualdad, siendo la forma general de un problema de optimización de la forma

$$\begin{cases} \text{Optimizar } f(x_1, ..., x_n) \\ \text{Sujeto a } \begin{cases} g_i(x_1, ..., x_n) = 0, \; i = 1, ..., k, \\ h_j(x_1, ..., x_n) \leq 0, \; j = 1, ..., l, \end{cases} \end{cases}$$

donde las funciones implicadas f, g_i y h_j poseen el grado de regularidad que deseemos, normalmente siendo funciones de clase C^2. Claramente, estos problemas de optimización pueden abarcar varios tipos según el tipo de restricciones que aparezcan en los mismos. Grosso modo, podemos distinguir cuatro tipos de problemas de optimización en función de sus ligaduras:

- Optimización sin ligaduras, donde se trata el problema

$$\{\text{Optimizar } f(x_1, ..., x_n).$$

  Este tipo de problemas ya se analizó en cursos anteriores y no hay muchos comentarios que hacer al respecto.

- Optimización con ligaduras de igualdad. Se trata de problemas de la forma

$$\begin{cases} \text{Optimizar } f(x_1, ..., x_n) \\ \text{Sujeto a } g_i(x_1, ..., x_n) = 0, \ i = 1, ..., k, \end{cases}$$

  también tratados someramente en cursos anteriores, básicamente para funciones definidas en conjuntos abiertos del plano y del espacio.

- Optimización con ligaduras de desigualdad. Son problemas de la forma

$$\begin{cases} \text{Optimizar } f(x_1, ..., x_n) \\ \text{Sujeto a } h_i(x_1, ..., x_n) \leq 0, \ i = 1, ..., l, \end{cases}$$

  que aunque podrían admitir un tratamiento combinando los dos anteriores, vamos a ver como hacer un tratamiento general de los mismos. Sirva esta sección para indicar el sentido de la desigualdad. En efecto, los problemas

$$\begin{cases} \text{Optimizar } f(x_1, ..., x_n) \\ \text{Sujeto a } h_i(x_1, ..., x_n) \leq 0, \ i = 1, ..., l, \end{cases}$$

  y

$$\begin{cases} \text{Optimizar } f(x_1, ..., x_n) \\ \text{Sujeto a } \widetilde{h}_i(x_1, ..., x_n) \geq 0, \ i = 1, ..., l, \end{cases}$$

  no son iguales a la forma de resolverlos. Este último problema deberíamos transformarlo en

$$\begin{cases} \text{Optimizar } f(x_1, ..., x_n) \\ \text{Sujeto a } -\widetilde{h}_i(x_1, ..., x_n) \leq 0, \ i = 1, ..., l, \end{cases}$$

  para resolverlo tal y como vamos a explicar a continuación. Obviamente, lo mismo se aplica si sólo una parte de restricciones de desigualdad vienen afectados del signo $\geq$. Ahora sólo éstas tendrían que cambiarse a tal y como indicamos anteriormente.

- Finalmente, tenemos el problema completo, de la forma general

$$\left\{\begin{array}{l}\text{Optimizar } f(x_1,...,x_n)\\ \text{Sujeto a } \left\{\begin{array}{l} g_i(x_1,...,x_n)=0,\ i=1,...,k,\\ h_j(x_1,...,x_n)\le 0,\ j=1,...,l,\end{array}\right.\end{array}\right.$$

  tal y como explicamos anteriormente. En general, vamos a desarrollar la teoría que nos va a permitir dar la solución de los problemas anteriores partiendo de la forma general, aunque iremos haciendo comentarios sobre los casos particulares cuando sea preciso.

1.2. Solución de problemas de optimización

Antes de resolver los problemas de optimización, debemos entender qué es una solución de los mismos. Dado el problema en forma general

$$\left\{\begin{array}{l}\text{Optimizar } f(x_1,...,x_n)\\ \text{Sujeto a } \left\{\begin{array}{l} g_i(x_1,...,x_n)=0,\ i=1,...,k,\\ h_j(x_1,...,x_n)\le 0,\ j=1,...,l,\end{array}\right.\end{array}\right.$$

una solución del mismo es un vector $(x_1^*,...,x_n^*)$ de manera que verifique que $f(x_1^*,...,x_n^*)$ es máximo o mínimo del problema, con lo cual debe de pertenecer al dominio de la función f, y además verifique todas las restricciones del problema, esto es, $g_i(x_1,...,x_n)=0,\ i=1,...,k$, y $h_j(x_1,...,x_n)\le 0,\ j=1,...,l$. Es lo que llamaremos un punto factible del problema.

De hecho, definimos el conjunto factible del problema anterior como

$$\mathcal{F}:=\left\{(x_1,...,x_n)\in\mathcal{D}_f:\ \begin{array}{l} g_i(x_1,...,x_n)=0,\ i=1,...,k,\\ h_j(x_1,...,x_n)\le 0,\ j=1,...,l,\end{array}\right\}$$

donde $\mathcal{D}_f$ representa el dominio de la función f.

Debemos también de distinguir entre soluciones globales y locales. Por ejemplo, consideremos que $(x_1^*,...,x_n^*)$ es un mínimo (respectivamente máximo), de un problema de optimización. Si tal mínimo verifica que $f(x_1^*,...,x_n^*)\le f(x_1,...x_n)$ para todo $(x_1,...,x_n)\in\mathcal{F}$, diremos que $(x_1^*,...,x_n^*)$ es un mínimo global para el problema de minimización. Si además, se verifica que $f(x_1^*,...,x_n^*)< f(x_1,...x_n)$, diremos que se trata de un mínimo global estricto. Si por el contrario, la desigualdades anteriores $f(x_1^*,...,x_n^*)\le f(x_1,...x_n)$ y $f(x_1^*,...,x_n^*)< f(x_1,...x_n)$ son válidas para un entorno U del punto $(x_1^*,...,x_n^*)$, estaremos hablando de un mínimo local y un mínimo local estricto, respectivamente. Las mismas definiciones son validas si de lo que estamos hablando es de máximos de la función f.

Ahora bien, no todo problema de optimización tiene porqué tener solución. Es fácil darse cuenta que la función objetivo

$$f(x,y)=x^2-y^2$$

sin añadir ninguna restricción adicional, no posee ni máximo ni mínimo. En efecto, basta considerar puntos de la forma $(x, 0)$ para ver que sobre estos puntos la función objetivo se dirige hacia el infinito, mientras que puntos de la forma $(0, y)$ se dirigen hacie menos infinito. Así, pues, dicha función no tiene extremos globales. Pero tampoco los tiene locales, ya que como conocemos de años anteriores, los extremos locales de dicha función deben verificar la condición

$$\begin{aligned} \frac{\partial f}{\partial x} &= 2x = 0, \\ \frac{\partial f}{\partial y} &= -2y = 0, \end{aligned}$$

de donde necesariamente $x = y = 0$ deberían ser el extremo local. Pero, como sabemos su matriz Hessiana

$$\mathbf{H}f = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$$

no es ni definida positiva ni negativa, por lo que se trata de un punto de silla. Volveremos posteriormente sobre estas nociones.

Para garantizar la existencia de soluciones globales a problemas de optimización, debe de verificarse el Teorema de Weiertrass sobre el conjunto factible de nuestro problema, es decir, éste debe ser compacto, o lo que es lo mismo en $\mathbb{R}^n$, cerrado y acotado. Como veremos en la práctica, no siempre es fácil ver que un conjunto es compacto, y en ocasiones tendremos que conformarnos con verificar que los extremos que estamos calculando son al menos locales.

A continuación, vamos a describir de una forma sistemática cómo obtener las soluciones de un problema de optimización.

1.3. Puntos regulares y singulares

Antes de exponer la condiciones de Karush–Kuhn–Tucker, a partir de ahora KKT para simplificar, necesitamos algunas definiciones previas. Tomemos el problema general

$$\left\{ \begin{array}{l} \text{Optimizar } f(x_1, ..., x_n) \\ \text{Sujeto a } \left\{ \begin{array}{l} g_i(x_1, ..., x_n) = 0, \ i = 1, ..., k, \\ h_j(x_1, ..., x_n) \leq 0, \ j = 1, ..., l. \end{array} \right. \end{array} \right.$$

Como recordamos, tenemos una serie de condiciones que son de desigualdad, de la forma $h_j(x_1, ..., x_n) \leq 0, \ j = 1, ..., l$. Dado un elemento del conjunto factible $(x_1, ..., x_n)$ y fijado j_0, diremos que la restricción $h_{j_0}(x_1, ..., x_n) \leq 0$ está activa si se verifica la igualdad $h_{j_0}(x_1, ..., x_n) = 0$. Así, denotamos por

$$\Omega(x_1, ..., x_n) = \{j_0 \in \{1, ..., j\} : h_{j_0}(x_1, ..., x_n) = 0\},$$

es decir, el conjunto de restricciones de desigualdad activas para el punto factible $(x_1, ..., x_n)$.

Diremos que un punto del conjunto factible $(x_1, ..., x_n)$ es regular si no hay restricción alguna o si el conjunto de vectores

$$\{\nabla g_1(x_1, ..., x_n), ..., \nabla g_k(x_1, ..., x_n)\} \cup \{\nabla h_j(x_1, ..., x_n) : j \in \Omega(x_1, ..., x_n)\}$$

es linealmente independiente. En caso contrario, se dira que el punto es no regular o singular. Veámoslo con un ejemplo.

$$\left\{ \begin{array}{l} \text{Optimizar } \ x+y+z^2 \\ \text{Sujeto a } \left\{ \begin{array}{l} x+y+z=0 \\ x^2+y^2+z^2 \leq 4. \end{array} \right. \end{array} \right.$$

Como vemos, tenemos una restricción de igualdad, cuyo gradiente es $(1,1,1)$ y otra de desigualdad cuyo gradiente es $(2x, 2y, 2z)$. Así, hemos de distinguir dos casos. Si la segunda restricción no está activa, esto es, $x^2 + y^2 + z^2 < 4$, entonces el único vector que hemos de considerar es el $(1,1,1)$, que obviamente el linealmente independiente por ser no nulo. Así, todos los puntos del conjunto factible que no activan la segunda restricción son regulares. El segundo caso que hemos de considerar es cuando la segunda restricción está activa, en cuyo caso hemos de verificar que los vectores $(1,1,1)$ y $(2x, 2y, 2z)$ sean linealmente independientes. Para ello realizamos las operaciones con determinantes

$$\left| \begin{array}{cc} 1 & 1 \\ 2x & 2y \end{array} \right| = 2(y-x),$$

$$\left| \begin{array}{cc} 1 & 1 \\ 2x & 2z \end{array} \right| = 2(z-x),$$

que se anulan cuando $x = y = z$, en caso de ser linealmente dependientes. Por la restricción de igualdad, necesariamente se tendría que $x = y = z = 0$, lo que implicaría que $x^2 + y^2 + z^2 = 0$, y por lo tanto la segunda restricción no sería activa. Así, los vectores sería linealmente independientes y por tanto todos los puntos del ejemplo serían regulares.

No es díficil construir un ejemplo donde tengamos puntos singulares. Por ejemplo

$$\left\{ \begin{array}{l} \text{Optimizar } \ x^2+y \\ \text{Sujeto a } y^3 \leq 0. \end{array} \right.$$

En este caso, el conjunto factible sería

$$\mathcal{F} = \{(x,y) : y \leq 0\},$$

y distinguimos dos casos como antes. Si $y^3 < 0$, entonces los puntos son regulares al no aplicarsele ninguna restricción. En el caso $y^3 = 0$, tenemos que el vector gradiente es $(0, 3y^2)$, que como vemos es linealmente dependiente cuando $y = 0$. Así, los punto de la forma $(x, 0)$ son puntos singulares del problema de optimización.

La relevancia de los puntos singulares no es menor. A continuación vamos a explicar unos métodos que nos permiten encontrar los máximos o mínimos de problemas de optimización para aquellos punto que son regulares. Cuando tengamos puntos singulares, tendremos que tratar éstos aparte, ya que pueden ser soluciones del problema de optimización, pero no tendremos una teoría sistemática como vamos a desarrollar ahora para comprobarlo.

1.4. Condiciones de Karush-Kuhn-Tucker

A continuación, vamos a exponer la condiciones de Karush–Kuhn–Tucker, a partir de ahora KKT para simplificar. Para ello y como de costumbre, partimos del problema general de optimización en el que supondremos que todos los puntos del conjunto factible son regulares.

$$\left\{\begin{array}{l} \text{Optimizar } f(x_1,...,x_n) \\ \text{Sujeto a } \left\{\begin{array}{l} g_i(x_1,...,x_n)=0,\ i=1,...,k, \\ h_j(x_1,...,x_n)\leq 0,\ j=1,...,l. \end{array}\right. \end{array}\right.$$

Construimos el Lagrangiano del problema, que es la función

$$L(x_1,...,x_n,\lambda_1,...,\lambda_k,\mu_1,...,\mu_l)=f(x_1,...,x_n)+\sum_{i=1}^{k}\lambda_i g_i(x_1,....,x_n)+\sum_{j=1}^{l}\mu_j h_j(x_1,...,x_n),$$

donde los números λ_i, $i=1,....,k$, se llama multiplicadores de Lagrange y μ_j, $j=1,...,l$, se llaman multiplicadores de KKT. Esta función lagrangiana es la función básica respecto de la cual se construyen las condiciones de KKT, siendo posible que sólo haya restricciones de igualdad, de desigualdad, o incluso ninguna restricción en cuyo caso la función L coincide con f. Dicho ésto, las condiciones de KKT son las siguientes:

1. Condición estacionaria. Dada por las ecuaciones
$$\frac{\partial L}{\partial x_m}=0,\quad m=1,2,...,m.$$
2. Condición de factibilidad, si se verifican las restricciones
$$\left\{\begin{array}{l} g_i(x_1,...,x_n)=0,\ i=1,...,k, \\ h_j(x_1,...,x_n)\leq 0,\ j=1,...,l. \end{array}\right.$$
3. Condición de holgura, para la restricciones de desigualdad
$$\mu_j h_j(x_1,...,x_n)=0,\quad j=1,...,l.$$
4. Condición de signo, también para las restricciones de desigualdad, y que afirman que si $\mu_j\geq 0$ para todo $j=1,...,l$, entonces tenemos la posibilidad de mínimo, mientras que si $\mu_j\leq 0$ para todo $j=1,...,l$, lo que podemos encontrarnos sea un máximo.

Hemos de advertir aquí que todas estas condiciones nos permiten encontrar unos puntos que son los posibles candidatos a máximo o minimo local para nuestro problema, pero de ninguna manera garantizan que estos puntos son lo que pensamos que son. En todo caso, un extremo debe de cumplir las condiciones KKT, junto con la condición de segundo orden que descrbiremos posteriormente, por lo que hablamos de las condiciones necesarias para la existencia de extremo. Veamos a continuación un par de ejemplos tomado de los apuntes de Silvestre Paredes.

$$\left\{\begin{array}{l}\text{Optimizar } \ x+y+z \\ \text{Sujeto a } \left\{\begin{array}{l}(y-1)^2+z^2 \leq 1 \\ x^2+(y-1)^2+z^2 \leq 3\end{array}\right.\end{array}\right.$$

En primer lugar, démonos cuenta que tenemos dos restricciones de desigualdad, y antes de nada, hemos de comprobar que los puntos son regulares. Si ninguna restricción activa, obviamente todos los puntos son regulares. Supongamos ahora que únicamente la primera es activa. Su gradiente es $(0, 2(y-1), 2z)$, que es linealmente independiente salvo si $y = 1$ y $z = 0$, valores que harían que la restricción no estuviese activa, y por lo tanto todos los puntos que activan sólo la primera restricción son regulares. De forma análoga se comprueba que todos los puntos que sólo activan la segunda restricción son regulares, así que supongamos que ambas restricciones son activas, siendo los gradientes $(0, 2(y-1), 2z)$ y $(2x, 2(y-1), 2z)$. De

$$\left|\begin{array}{cc}0 & 2(y-1) \\ 2x & 2(y-1)\end{array}\right| = 4x(y-1),$$

$$\left|\begin{array}{cc}0 & 2z \\ 2x & 2z\end{array}\right| = 4xz,$$

vemos que estos determinantes son ambos nulos si $x = 0$, y entonces $x^2 + (y-1)^2 + z^2 = (y-1)^2 + z^2 = 1$, por lo que la segunda restricción no sería activa. Otra contradicción obtenemos si $y = 1$ y $z = 0$. De estas concluimos que todos los puntos son regulares.

Planteamos ahora el Lagrangiano

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

Explícitemos ahora las condiciones KKT.

1. Condición estacionaria.

$$\frac{\partial L}{\partial x} = 1 + 2x\mu_2 = 0,$$

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

$$\frac{\partial L}{\partial x} = 1 + 2z(\mu_1 + \mu_2) = 0,$$

2. Condición de factibilidad
$$\begin{cases} (y-1)^2+z^2 \leq 1 \\ x^2+(y-1)^2+z^2 \leq 3. \end{cases}$$
3. Condición de holgura
$$\mu_1\left((y-1)^2+z^2-1\right)=0,$$
$$\mu_2\left(x^2+(y-1)^2+z^2-3\right)=0.$$
4. Condición de signo. $\mu_1, \mu_2 \geq 0$ para mínimo y $\mu_1, \mu_2 \leq 0$ para máximo.

Unas vez escritas las condiciones KKT, debemos de encontrar los puntos KKT, es decir, resolver el sistema anterior. En primer lugar, nos centramos en las condiciones de holgura que nos dan lugar a 4 casos: $\mu_1=\mu_2=0$ (caso 1), $\mu_1=0$, $\mu_2 \neq 0$ (caso 2), $\mu_1 \neq 0, \mu_2=0$ (caso 3) y $\mu_1 \neq 0, \mu_2 \neq 0$ (caso 4). Vamos a analizar cada uno de los cuatro casos anteriores.

Casos 1 y 3. De la primera condición estacionaria tenemos que
$$\frac{\partial L}{\partial x}=1=0,$$
lo cual es imposible y por tanto no hay soluciones para los casos 1 y 2.

Caso 2. Ahora $\mu_2 \neq 0$, por lo que de la condición de holgura, necesariamente $x^2+(y-1)^2+z^2=3$. Por otra parte la condición estacionaria queda de la forma
$$\frac{\partial L}{\partial x}=1+2x\mu_2=0,$$
$$\frac{\partial L}{\partial y}=1+2(y-1)\mu_2=0,$$
$$\frac{\partial L}{\partial x}=1+2z\mu_2=0,$$
de donde obtenemos
$$x=y-1=z=-\frac{1}{2\mu_2},$$
y sustituyendo en la segunda restricción, que es activa, obtenemos
$$\left(-\frac{1}{2\mu_2}\right)^2+\left(-\frac{1}{2\mu_2}\right)^2+\left(-\frac{1}{2\mu_2}\right)^2=3,$$
de donde
$$\frac{3}{4\mu_2^2}=3,$$
y así
$$\mu_2=\pm\frac{1}{2}.$$
Obtenemos entonces los primeros puntos KKT sustituyendo directamente en las ecuaciones anteriores:

1. $x = -1$, $y = 0$, $z = -1$, $\mu_1 = 0$, $\mu_2 = \frac{1}{2}$.

2. $x = 1$, $y = 2$, $z = 1$, $\mu_1 = 0$, $\mu_2 = -\frac{1}{2}$.

Tendríamos ahora que comprobar la condición de factibilidad, ya que una de las restricciones no estaba activa y comprobamos que ninguno de los puntos que hemos obtenido satisface la desigualdad $(y-1)^2 + z^2 \leq 1$, por lo que no serían puntos KKT.

Caso 3. Ahora μ_1 y μ_2 no son nulos, por lo que nos queda un sistema de 5 ecuaciones con 5 incógnitas dado por

$$1 + 2x\mu_2 = 0,$$

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

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

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

$$x^2 + (y-1)^2 + z^2 = 3.$$

Ahora

$$x = -\frac{1}{2\mu_2}, \quad y - 1 = z = -\frac{1}{2(\mu_1 + \mu_2)}.$$

Sustituyendo $y - 1$ y z en la primera restrición tenemos que

$$\left(-\frac{1}{2(\mu_1 + \mu_2)}\right)^2 + \left(-\frac{1}{2(\mu_1 + \mu_2)}\right)^2 = 1,$$

de donde

$$\frac{1}{2(\mu_1 + \mu_2)^2} = 1,$$

o equivalentemente

$$\mu_1 + \mu_2 = \pm\frac{\sqrt{2}}{2}.$$

Sustituyendo en la tercera ecuación obtenemos que

$$\left(-\frac{1}{2\mu_2}\right)^2 + \left(-\frac{1}{2(\mu_1 + \mu_2)}\right)^2 + \left(-\frac{1}{2(\mu_1 + \mu_2)}\right)^2 = 3,$$

de donde

$$\frac{1}{4\mu_2^2} = 4,$$

de donde

$$\mu_2 = \pm\frac{\sqrt{2}}{4},$$

y obtenemos así los puntos:

1. $x = -\sqrt{2}$, $y = 1 - \frac{\sqrt{2}}{2}$, $z = -\frac{\sqrt{2}}{2}$, $\mu_1 = \frac{\sqrt{2}}{4}$, $\mu_2 = \frac{\sqrt{2}}{4}$.

2. $x = -\sqrt{2}$, $y = 1 + \frac{\sqrt{2}}{2}$, $z = \frac{\sqrt{2}}{2}$, $\mu_1 = -\frac{3\sqrt{2}}{2}$, $\mu_2 = \frac{\sqrt{2}}{4}$.

3. $x = \sqrt{2}$, $y = 1 - \frac{\sqrt{2}}{2}$, $z = -\frac{\sqrt{2}}{2}$, $\mu_1 = \frac{3\sqrt{2}}{2}$, $\mu_2 = -\frac{\sqrt{2}}{4}$.

4. $x = \sqrt{2}$, $y = 1 + \frac{\sqrt{2}}{2}$, $z = \frac{\sqrt{2}}{2}$, $\mu_1 = -\frac{\sqrt{2}}{4}$, $\mu_2 = -\frac{\sqrt{2}}{4}$.

Si analizamos los puntos que nos han salido, nos damos cuenta que tanto el 2 como el 3 no cumplen la condición de signo. Por otra parte, al tener las dos restricciones activas, siempre se cumple la condición de factibilidad. En consecuencia los puntos 1 y 4 son los puntos KKT para nuestro problema, siendo el primero un posible mínimo y el segundo un posible máximo.

Aalicemos a continuacuón otro ejemplo proveniente de los apuntes de Silvestre Paredes.

$$\begin{cases} \text{Optimizar } x^2 + y^2 \\ \text{Sujeto a } 2x + y - 2 = 0. \end{cases}$$

En este caso se trata de un problema donde no hay restricciones de desigualdad y el gradiente de la restricción es $(2, 1)$, obviamente linealmente independiente por lo que todos los puntos son regulares. El Lagrangiano es en este caso

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

Las condiciones KKT serán ahora.

1. Condición estacionaria.

$$\frac{\partial L}{\partial x} = 2x + 2\lambda = 0,$$

$$\frac{\partial L}{\partial y} = 2y + \lambda = 0,$$

2. Condición de factibilidad

$$2x + y - 2 = 0.$$

3. Condición de holgura. No hay.

4. Condición de signo. No hay.

Resolver el problema ahora es mucho más fácil técnicamente ya que

$$x = -\lambda, \quad y = -\lambda/2,$$

y sustituyendo tenemos

$$-2\lambda - \frac{\lambda}{2} = 2,$$

de donde $x = \frac{4}{5}$, $y = \frac{2}{5}$, $\lambda = -\frac{4}{5}$ que es el punto KKT del problema.

Es buen momento para volver al problema

$$\begin{cases} \text{Optimizar } x^2 + y \\ \text{Sujeto a } y^3 \leq 0. \end{cases}$$

que como sabíamos, todos los puntos de la forma $(x, 0)$ son irregulares. El lagrangiano es ahora

$$L(x, y, \mu) = x^2 + y + \mu y^3,$$

y las condiciones de KKT serán

1. Condición estacionaria.
$$\frac{\partial L}{\partial x} = 2x = 0,$$
$$\frac{\partial L}{\partial y} = 1 + 3\mu y^2 = 0,$$
2. Condición de factibilidad
$$y^3 \leq 0.$$
3. Condición de holgura.
$$\mu y^3 = 0.$$
4. Condición de signo. Si $\mu \geq 0$ será mínimo, siendo máximo si $\mu \leq 0$.

Como vemos, tenemos dos casos dados por la condición de holgura. Si $\mu = 0$, de la segunda condición estacionaria tenemos que $1 = 0$, por lo que ho habría solución. Si $\mu > 0$, y entonces $y^3 = y = 0$, tenemos de nuevo que no hay solución. Asi, tenemos un problema que aparentemente no tiene solución. Sin embargo, hemos olvidado que tiene puntos singulares, por lo que las condiciones KKT no nos dirían nada a priori sobre ellos. Si pensamos un poco y simplificamos el problema tomando $y = 0$, es decir, tomando los puntos no regulares, vemos que el problema se reduce a optimizar x^2, que obviamente tiene un mínimo local en 0. Así, vemos como los puntos singulares pueden esconder extremos y soluciones que a priori no pueden encontrarse con los puntos KKT. El problema, no obstante, es que estos casos singulares hay que estudiarlos uno a uno, de forma particularizada y no hay reglas generales para ello.

1.5. Condiciones de segundo orden

Como hemos visto, los puntos de KKT, bajo condiciones de regularidad, proporcionan posibles soluciones a problemas de optimización con restricciones en el caso en que las funciones implicadas sean de clase C^1. Si además disponemos de funciones de clase C^2, disponemos además de otra condición necesaria basada en la matriz Hessiana.

Como siempre, consideremos un problema de optimización de la forma

$$\begin{cases} \text{Optimizar } f(x_1,...,x_n) \\ \text{Sujeto a } \begin{cases} g_i(x_1,...,x_n)=0, \; i=1,...,k, \\ h_j(x_1,...,x_n)\le 0, \; j=1,...,l. \end{cases} \end{cases}$$

Dado un punto de KKT $(x_1,...,x_n)$ del mismo. Si las funciones f, g_i y h_j son de clase C^2 podemos construir, a partir del Lagrangiano

$$L(x_1,...,x_n,\lambda_1,...,\lambda_k,\mu_1,...,\mu_l)=f(x_1,...,x_n)+\sum_{i=1}^{k}\lambda_i g_i(x_1,...,x_n)+\sum_{j=1}^{l}\mu_j h_j(x_1,...,x_n),$$

la matriz Hessiana

$$HL(x_1,...,x_n,\lambda_1,...,\lambda_k,\mu_1,...,\mu_l)=\begin{pmatrix} \frac{\partial^2 L}{\partial x_1^2} & \frac{\partial^2 L}{\partial x_1 \partial x_2} & ... & \frac{\partial^2 L}{\partial x_1 \partial x_n} \\ \frac{\partial^2 L}{\partial x_2 \partial x_1} & \frac{\partial^2 L}{\partial x_2^2} & ... & \frac{\partial^2 L}{\partial x_2 \partial x_n} \\ ... & ... & ... & ... \\ \frac{\partial^2 L}{\partial x_n \partial x_1} & \frac{\partial^2 L}{\partial x_n \partial x_2} & ... & \frac{\partial^2 L}{\partial x_n^2} \end{pmatrix}$$

donde hemos omitido los argumentos de las derivadas segundas por simplicidad. Esta matriz Hessiana es simétrica por el Teorema de Schwarz, que afirma para funciones de clase C^2 la igualdad $\frac{\partial^2 L}{\partial x_1 \partial x_2}=\frac{\partial^2 L}{\partial x_2 \partial x_1}$. Dado el mencionado punto de KKT dado por $(x_1,...,x_n)$, se define el espacio tangente como[1]

$$\mathcal{M}(x_1,...,x_n)=\left\{ \begin{array}{c} (y_1,...,y_n)\in\mathbb{R}^n : \bigtriangledown g_i(x_1,...,x_n)\cdot(y_1,...,y_n)^t, \; i=1,...,k, \\ \bigtriangledown h_j(x_1,...,x_n)\cdot(y_1,...,y_n)^t, \; j\in\Omega(x_1,...,x_n), \end{array} \right\}$$

donde recordamos que $\Omega(x_1,...,x_n)$ era el conjunto de restricciones de desigualdad que eran activas para $(x_1,...,x_n)$, y el símbolo τ denota la traspuesta. Una condición necesaria para mínimo es que

$$(y_1,...,y_n)\cdot HL(x_1,...,x_n,\lambda_1,...,\lambda_k,\mu_1,...,\mu_l)\cdot(y_1,...,y_n)^t\ge 0$$

para todo $(y_1,...,y_n)\in\mathcal{M}(x_1,...,x_n)$ distinto de $(0,0,...,0)$, a lo que llamamos ser semidefinido positivo (definido positivo si la desigualdad es estricta) mientras que la condición análoga para máximo es de la forma

$$(y_1,...,y_n)\cdot HL(x_1,...,x_n,\lambda_1,...,\lambda_k,\mu_1,...,\mu_l)\cdot(y_1,...,y_n)^t\le 0$$

para todo $(y_1,...,y_n)\in\mathcal{M}(x_1,...,x_n)$ distinto de $(0,0,...,0)$, a lo que llamamos ser semidefinido negativo (definido negativo si la desigualdad es estricta).

A modo de ejemplo, tomemos el problema

$$\begin{cases} \text{Optimizar } x+y+z \\ \text{Sujeto a } \begin{cases} (y-1)^2+z^2\le 1 \\ x^2+(y-1)^2+z^2\le 3 \end{cases} \end{cases}$$

que como sabíamos, tenía como puntos de KKT:

[1] Por $(y_1,...,y_n)^t$ denotamos la matriz transpuesta del vector $(y_1,...,y_n)$, es decir, dicho vector en forma de columna.

1. $x = -\sqrt{2}$, $y = 1 - \frac{\sqrt{2}}{2}$, $z = -\frac{\sqrt{2}}{2}$, $\mu_1 = \frac{\sqrt{2}}{4}$, $\mu_2 = \frac{\sqrt{2}}{4}$.
2. $x = \sqrt{2}$, $y = 1 + \frac{\sqrt{2}}{2}$, $z = \frac{\sqrt{2}}{2}$, $\mu_1 = -\frac{\sqrt{2}}{4}$, $\mu_2 = -\frac{\sqrt{2}}{4}$.

Recordemos que el Lagrangiano era

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

por lo que la matriz Hessiana es

$$HL(x,y,z,\mu_1,\mu_2) = \begin{pmatrix} 2\mu_2 & 0 & 0 \\ 0 & 2(\mu_1+\mu_2) & 0 \\ 0 & 0 & 2(\mu_1+\mu_2) \end{pmatrix}.$$

Si tomamos los puntos KKT, vemos que están activas las dos restricciones de desigualdad. Los gradientes de éstas son $(0, 2(y-1), 2z)$ y $(2x, 2(y-1), 2z)$. Particularizando en el primer punto tenemos que es espacio tangente debe cumplir las ecuaciones

$$\begin{cases} (0, -\sqrt{2}, -\sqrt{2}) \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0, \\ (-2\sqrt{2}, -\sqrt{2}, -\sqrt{2}) \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0, \end{cases}$$

que simplifican a

$$\begin{cases} y+z=0, \\ 2x+y+z=0, \end{cases}$$

que en coordenadas paramétricas se escrbiría como

$$\begin{cases} x=0, \\ y=t, \\ z=-t, \end{cases}$$

por lo que el espacio tangente es

$$\mathcal{M}\left(-\sqrt{2}, 1-\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right) = \{(x,y,z) : y+z=0, x=0\}.$$

Como el Hessiano es

$$HL\left(-\sqrt{2}, 1-\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{4}, \frac{\sqrt{2}}{4}\right) = \begin{pmatrix} \frac{\sqrt{2}}{2} & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix},$$

y calculando

$$\begin{aligned}
(x,y,z)\cdot\begin{pmatrix} \frac{\sqrt{2}}{2} & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix}\cdot\begin{pmatrix} x \\ y \\ z \end{pmatrix} &= (0,t,-t)\cdot\begin{pmatrix} \frac{\sqrt{2}}{2} & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix}\cdot\begin{pmatrix} 0 \\ t \\ -t \end{pmatrix} \\
&= (0,t,-t)\cdot\begin{pmatrix} -0 \\ \sqrt{2}t \\ -\sqrt{2}t \end{pmatrix} \\
&= 2\sqrt{2}t^2 > 0
\end{aligned}$$

salvo si $t = 0$, por lo que el Hessiano es definido positivo, y el punto es candidado a mínimo.

Análogamente se comprueba que el otro punto es candidato a máximo.

1.6. Condiciones suficientes

En general es complicado dar condiciones suficientes para los puntos KKT. No obstante, definimos el espacio tangente ampliado

$$\mathcal{M}^*(x_1,...,x_n) = \left\{ \begin{array}{c} (y_1,...,y_n) \in \mathbb{R}^n : \bigtriangledown g_i(x_1,...,x_n)\cdot(y_1,...,y_n)^t,\ i = 1,...,k, \\ \bigtriangledown h_j(x_1,...,x_n)\cdot(y_1,...,y_n)^t,\ j \in \Omega(x_1,...,x_n) : \mu_j \neq 0. \end{array} \right\}$$

Claramente, $\mathcal{M}(x_1,...,x_n)$ es un subespacio vectorial de $\mathcal{M}^*(x_1,...,x_n)$, ya que en el segundo espacio eliminamos ecuaciones. Entonces una condición suficiente para mínimo es que

$$(y_1,...,y_n)\cdot HL(x_1,...,x_n,\lambda_1,...,\lambda_k,\mu_1,...,\mu_l)\cdot(y_1,...,y_n)^t > 0$$

para todo $(y_1,...,y_n) \in \mathcal{M}^*(x_1,...,x_n)$ distinto de $(0,0,...,0)$, mientras que la condición análoga para máximo es de la forma

$$(y_1,...,y_n)\cdot HL(x_1,...,x_n,\lambda_1,...,\lambda_k,\mu_1,...,\mu_l)\cdot(y_1,...,y_n)^t < 0$$

para todo $(y_1,...,y_n) \in \mathcal{M}^*(x_1,...,x_n)$ distinto de $(0,0,...,0)$.

Para finalizar con el problema

$$\left\{ \begin{array}{l} \text{Optimizar } x+y+z \\ \text{Sujeto a } \left\{ \begin{array}{l} (y-1)^2+z^2 \leq 1 \\ x^2+(y-1)^2+z^2 \leq 3 \end{array} \right. \end{array} \right.$$

que como sabíamos, tenía como puntos de KKT:

1. $x = -\sqrt{2}$, $y = 1 - \frac{\sqrt{2}}{2}$, $z = -\frac{\sqrt{2}}{2}$, $\mu_1 = \frac{\sqrt{2}}{4}$, $\mu_2 = \frac{\sqrt{2}}{4}$.
2. $x = \sqrt{2}$, $y = 1 + \frac{\sqrt{2}}{2}$, $z = \frac{\sqrt{2}}{2}$, $\mu_1 = -\frac{\sqrt{2}}{4}$, $\mu_2 = -\frac{\sqrt{2}}{4}$.

Es fácil darse cuenta que los espacios tangentes y los espacios tangentes ampliados son indénticos, y al ser los Hessianos definidos positivos y negativos, respectivamente, tenemos que efectivamente nuestros candidatos a extremos locales lo son de facto al verificarse las condiciones necesarias y suficientes.

En los ejercicios prácticos, resulta difícil encontrar ejemplos en los que el espacio tangente y el ampliado difieran, por lo que en la práctica, si las condiciones de segundo orden se dan para el caso semidefinido, es decir menor o igual o mayor o igual, no va a ser posible verificar las condiciones suficientes de esxistencia de máximo. A modo de ejemplo, la función $f(x) = x^4$ presenta un mínimo global en cero que no es detectable con las condiciones de segundo orden, sino con el uso de la derivada cuarta. La extensión de este hecho es bastante complicado en dimensiones mayores como las que estamos tratando.

Por otra parte, al ser la matriz Hessiana simétrica, ésta tiene todos sus valores propios reales. Precisamente estos valores propios son los que dan lugar a los casos semidefinido o definido positivo y negativo. Por ejemplo, si todos los valores propios son positivos tendremos un caso definido positivo, mientras que si son negativos, tendremos definido negativo, independientemente de cual sea el espacio tangente ampliado. Los casos conflictivos o problemáticos son aquellos en los que aparecen mezclados valores propios positivos o negativos o nulos. Es en este caso cuando los espacios tangentes pueden ser de utilidad a la hora de determinar el carácter del extremo que estamos estudiando.

1.7. Ejercicios resueltos

Presentamos a continuación algunos ejercicios que han formado parte de exámenes de la asignatura y de la asignatura correspondiente del grado de Tecnologías Industriales.

1. Resolver el siguiente problema

$$\begin{array}{cc} \text{Maximizar} & x+y+z \\ \text{Sujeto a} & x^2+y^2+z^2=1, \end{array}$$

sabiendo que el conjunto en que está definido es compacto.

Solución. Dado que existe una única ligadura de igualdad, tenemos que todos los puntos son regulares. Contruimos el lagrangiano

$$L(x,y,z,\lambda) = x+y+z+\lambda(x^2+y^2+z^2-1),$$

y tomamos la condición estacionaria

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

junto con la ligadura

$$x^2 + y^2 + z^2 = 1.$$

De las tres primeras condiciones tenemos que

$$x = y = z = -\frac{1}{2\lambda},$$

y sustituyendo en la expresión anterior tenemos que

$$\frac{1}{4\lambda^2} + \frac{1}{4\lambda^2} + \frac{1}{4\lambda^2} = \frac{3}{4\lambda^2} = 1,$$

de donde

$$\lambda = \pm\frac{\sqrt{3}}{2}.$$

Obtenemos entonces los candidatos

$$x = y = z = -\frac{\sqrt{3}}{3},\ \lambda = \frac{\sqrt{3}}{2}$$

y

$$x = y = z = \frac{\sqrt{3}}{3},\ \lambda = -\frac{\sqrt{3}}{2},$$

que obviamente son factibles. Calculamos la matriz Hessiana

$$HL(x, y, z, \lambda) = \begin{pmatrix} 2\lambda & 0 & 0 \\ 0 & 2\lambda & 0 \\ 0 & 0 & 2\lambda \end{pmatrix}.$$

Obviamente

$$HL\left(\frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3}, -\frac{\sqrt{3}}{2}\right) = \begin{pmatrix} -\sqrt{3} & 0 & 0 \\ 0 & -\sqrt{3} & 0 \\ 0 & 0 & -\sqrt{3} \end{pmatrix}$$

es definida negativa al tener todos sus valores propios negativos y por tanto $\left(\frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3}\right)$ es el máximo pedido, que tomará el valor $\sqrt{3}$ en su función objetivo.

2. Resolver el siguiente problema

$$\begin{array}{cc} \text{Maximizar} & x + y + z \\ \text{Sujeto a} & x^2 + y^2 + z^2 \leq 1, \end{array}$$

sabiendo que el conjunto en que está definido es compacto.

Solución. En primer lugar, el problema es regular ya que cuando la restricción está activa, el gradiente queda como $(2x, 2y, 2z) \neq (0, 0, 0)$ para aquellos valores que activan la desigualdad. Tomamos el Lagrangiano

$$L(x, y, z, \mu) = x + y + z + \mu(x^2 + y^2 + z^2 - 1).$$

Distinguimos dos casos: $\mu = 0$ y $\mu \neq 0$. En el primer caso es fácil darse cuenta de que si planteamos la condición estacionaria

$$\frac{\partial L}{\partial x} = 1 = 0,$$

nos lleva a una contradicción, por lo que las soluciones deben de estar para el caso $\mu \neq 0$. En este caso, planteamos las condiciones estacionarias

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

junto con la condición

$$x^2 + y^2 + z^2 = 1.$$

De las tres primeras obtenemos

$$x = y = z = \frac{-1}{2\mu},$$

que sustituyendo nos da

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

de donde

$$\frac{3}{4\mu^2} = 1,$$

de donde

$$\mu = \pm\frac{\sqrt{3}}{2}.$$

Como buscamos un máximo, hemos de tomar $\mu = -\frac{\sqrt{3}}{2}$, que da lugar a los valores

$$x = y = z = \frac{\sqrt{3}}{3}.$$

Como

$$\left(\frac{\sqrt{3}}{3}\right)^2 + \left(\frac{\sqrt{3}}{3}\right)^2 + \left(\frac{\sqrt{3}}{3}\right)^2 = 1,$$

se trata de un punto factible. Calculamos el Hessiano

$$H(x,y,z,\mu) = \begin{pmatrix} 2\mu & 0 & 0 \\ 0 & 2\mu & 0 \\ 0 & 0 & 2\mu \end{pmatrix},$$

que particularizando

$$H(x,y,z,\mu) = \begin{pmatrix} -\sqrt{3} & 0 & 0 \\ 0 & -\sqrt{3} & 0 \\ 0 & 0 & -\sqrt{3} \end{pmatrix},$$

que como vemos es definida negativa y por tanto se trata de un máximo.

3. Resolver el siguiente problema

$$\begin{array}{cc} \text{Minimizar} & xy + z \\ \text{Sujeto a} & x^2 + y^2 + z^2 = 5, \\ & x + 1 \leq 0, \end{array}$$

sabiendo que el conjunto en que está definido es compacto.

Solución. En primer lugar, comprobamos que todos los puntos son regulares. Si $x+1 < 0$, entonces el gradiente $(2x, 2y, 2z)$ es linealmente independiente salvo si $x = y = z = 0$, lo cual es imposible porque implicaría que $x^2+y^2+z^2 = 0$. Si $x+1 = 0$, entonces los vectore gradiente $(2x, 2y, 2z)$ y $(1, 0, 0)$ son linealmente independientes salvo que $y = z = 0$. En este caso $x^2 = 5$, lo que implicaría que $x = \pm\sqrt{5}$, que no puede verificarse al ser $x = -1$.

Consideramos el Lagrangiano

$$L(x,y,z,\lambda,\mu) = xy + z + \lambda(x^2 + y^2 + z^2 - 5) + \mu(x + 1).$$

De la condición de holgura $\mu(x + 1) = 0$ disitinguimos dos casos a) $\mu = 0$ y b) $x = -1$.

Caso a) $\mu = 0$. El Lagrangiano es $L(x,y,z,\lambda) = xy + z + \lambda(x^2 + y^2 + z^2 - 5)$ y los puntos de KKT se calculan a partir de las ecuaciones

$$\left\{ \begin{array}{l} \frac{\partial}{\partial x} = y + 2\lambda x = 0, \\ \frac{\partial}{\partial y} = x + 2\lambda y = 0, \\ \frac{\partial}{\partial z} = 1 + 2\lambda z = 0, \\ x^2 + y^2 + z^2 = 5. \end{array} \right.$$

De las dos primeras obtenemos que $x(1 - 4\lambda^2) = 0$, y como $x \leq -1$, necesariamente $\lambda = \pm 1/2$. Si $\lambda = 1/2$, entonces $x = -y$, $z = -1$ y por tanto $x = \pm\sqrt{2}$. Como $x \leq -1$, la única solución posible es

$$x = -\sqrt{2},\ y = \sqrt{2},\ z = -1,\ \lambda = 1/2,\ \mu = 0.\ (KKT1)$$

Si $\lambda = -1/2$, obtenemos como única solución

$$x = -\sqrt{2},\ y = -\sqrt{2},\ z = 1,\ \lambda = -1/2,\ \mu = 0.\ (KKT2)$$

Caso b) $x = -1$. Los puntos de KKT se calculan a partir de las ecuaciones

$$\begin{cases} \frac{\partial}{\partial x} = y + 2\lambda x + \mu = 0, \\ \frac{\partial}{\partial y} = x + 2\lambda y = 0, \\ \frac{\partial}{\partial z} = 1 + 2\lambda z = 0, \\ x^2 + y^2 + z^2 = 5, \\ x = -1. \end{cases}$$

Nótese que $\lambda \neq 0$ (en otro caso 1=0 en la tercera ecuación). De la segunda y tercera ecuaciones obtenemos $y = \frac{1}{2\lambda} = -z$, por lo que de la cuarta ecuación $y = \pm\sqrt{2}$. Si $y = \sqrt{2}$ obtenemos la solución

$$x = -1,\ y = \sqrt{2},\ z = -\sqrt{2},\ \lambda = \sqrt{2}/4,\ \mu = -\sqrt{2}/2,$$

que no es posible candidato a mínimo al ser μ negativo. Si $y = -\sqrt{2}$ tenemos

$$x = -1,\ y = -\sqrt{2},\ z = \sqrt{2},\ \lambda = -\sqrt{2}/4,\ \mu = \sqrt{2}/2.\ (KKT3)$$

que es el tercer posible candidato a mínimo. Calculamos ahora el Hessiano

$$HL(x, y, z, \lambda, \mu) = \begin{pmatrix} 2\lambda & 1 & 0 \\ 1 & 2\lambda & 0 \\ 0 & 0 & 2\lambda \end{pmatrix}.$$

Para el punto (KKT1) tenemos que el Hessiano

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

cuyo espacio tangente es

$$\begin{aligned} M(-\sqrt{2}, \sqrt{2}, -1) &= \{(x, y, z) : -\sqrt{2}x + \sqrt{2}y - z = 0\} \\ &= \left\{ \begin{array}{c} x = u, \\ y = v, \\ z = -\sqrt{2}u + \sqrt{2}v, \end{array} \right. \quad u, v \in \mathbb{R}, \end{aligned}$$

Calculamos

$$(u, v, -\sqrt{2}u + \sqrt{2}v) \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} u \\ v \\ -\sqrt{2}u + \sqrt{2}v \end{pmatrix}$$

$$= \left(u + v, u + v, -\sqrt{2}u + \sqrt{2}v\right) \begin{pmatrix} u \\ v \\ -\sqrt{2}u + \sqrt{2}v \end{pmatrix} = (u + v)^2 + (-\sqrt{2}u + \sqrt{2}v)^2,$$

que es estrictamente positivo si $u, v \neq 0$, por lo que es definida positiva y por tanto mínimo local.

Para el punto (KKT2) tenemos que el Hessiano

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

cuyo espacio tangente es

$$\begin{aligned} M(-\sqrt{2}, -\sqrt{2}, 1) &= \{(x, y, z) : -\sqrt{2}x - \sqrt{2}y + z = 0\} \\ &= \left\{ \begin{array}{c} x = u, \\ y = v, \\ z = \sqrt{2}u + \sqrt{2}v, \end{array} \right. \quad u, v \in \mathbb{R}, \end{aligned}$$

Calculamos

$$\begin{aligned} & (u, v, \sqrt{2}u + \sqrt{2}v) \begin{pmatrix} -1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} u \\ v \\ \sqrt{2}u + \sqrt{2}v \end{pmatrix} \\ = & \left(-u + v, u - v, -\sqrt{2}u - \sqrt{2}v\right) \begin{pmatrix} u \\ v \\ \sqrt{2}u + \sqrt{2}v \end{pmatrix} = (u - v)^2 - (\sqrt{2}u + \sqrt{2}v)^2, \end{aligned}$$

que es estrictamente negativo si $u, v \neq 0$, por lo que es definida positiva y por tanto máximo local.

Finalmente, para el punto (KKT3) tenemos que el Hessiano

$$HL(-1, -\sqrt{2}, \sqrt{2}, -\sqrt{2}/4, \sqrt{2}/2) = \begin{pmatrix} -\sqrt{2}/2 & 1 & 0 \\ 1 & -\sqrt{2}/2 & 0 \\ 0 & 0 & -\sqrt{2}/2 \end{pmatrix},$$

cuyo espacio tangente es

$$\begin{aligned} M(-1, -\sqrt{2}, \sqrt{2}) &= \{(x, y, z) : x = 0; z = y\} \\ &= \left\{ \begin{array}{l} x = 0, \\ y = u, \quad u \in \mathbb{R}, \\ z = u, \end{array} \right. \end{aligned}$$

Calculamos

$$\begin{aligned} & (0, u, u) \begin{pmatrix} -\sqrt{2}/2 & 1 & 0 \\ 1 & -\sqrt{2}/2 & 0 \\ 0 & 0 & -\sqrt{2}/2 \end{pmatrix} \begin{pmatrix} 0 \\ u \\ u \end{pmatrix} \\ = & \left(u, -u\sqrt{2}/2, -u\sqrt{2}/2\right) \begin{pmatrix} 0 \\ u \\ u \end{pmatrix} = -\sqrt{2}u^2, \end{aligned}$$

que es menor que cero si $u \neq 0$, por lo que es definida negativa, por lo que no puede ser mínimo. Por lo tanto $KKT1$ es el mínimo pedido.

4. Resolver el siguiente problema

$$\begin{array}{cc} \text{Minimizar} & x^2 + y^2 + z^2 \\ \text{Sujeto a} & x + y + z \leq -1. \end{array}$$

Solución. Planteamos el Lagrangiano

$$L(x, y, z, \mu) = x^2 + y^2 + z^2 + \mu(x + y + z - 1),$$

y las ecuaciones

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

y la condición de holgura

$$\mu(x + y + z + 1) = 0.$$

De esta última condición, podemos concluir que $\mu = 0$, que implicaría que $x = y = z = 0$, incompatible con la ligadura por lo que la descartamos. Nos queda entonces la ecuación

$$x + y + z = -1.$$

Despejando en las tres primeras ecuaciones

$$x = y = z = -\frac{1}{2\mu},$$

y sustituyendo en la cuarta ecuación obtenemos

$$\frac{-3}{2\mu} = -1,$$

con lo que $\mu = 3/2$ y $x = y = z = -\frac{1}{3}$, que verifican la ligadura. Todos los puntos son regulares al ser $(1, 1, 1)$ linealmente independiente para todos los posibles valores de x, y y z. Calculamos el Hessiano

$$HL = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

que es siempre deifinido positivo, por lo que $(-1/3, -1/3, -1/3)$ es el mínimo pedido.

5. Resolver el siguiente problema

$$\begin{array}{cc} \text{Maximizar} & x+y+z \\ \text{Sujeto a} & x^2+y^2=1, \\ & 3x+2z\leq 1. \end{array}$$

Solución: En primer lugar, el problema no tiene puntos sigulares ya que los vectores $(2x, 2y, 0)$ y $(3, 0, 2)$ son siempre linealmente independientes. Planteamos $L(x, y, z, \lambda, \mu) = x + y + z + \lambda(x^2 + y^2 - 1) + \mu(3x + 2z - 1)$ y escribimos las ecuaciones

$$\begin{aligned} \frac{\partial L}{\partial x} &= 1+2x\lambda+3\mu=0, \\ \frac{\partial L}{\partial y} &= 1+2y\lambda=0, \\ \frac{\partial L}{\partial z} &= 1+2\mu=0, \\ x^2+y^2 &= 1, \\ \mu(3x+2z-1) &= 0. \end{aligned}$$

Como por la tercera ecuación $\mu = -1/2$, la condición $\mu = 0$ no es posible y la condición de holgura se reduce a

$$3x+2z=1.$$

Despejamos en las dos primeras ecuaciones

$$x=\frac{1}{4\lambda} \text{ e } y=-\frac{1}{2\lambda},$$

y sustituimos es la cuarta

$$\frac{1}{16\lambda^2}+\frac{1}{4\lambda^2}=1,$$

de donde

$$16\lambda^2=5$$

y por tanto

$$\lambda=\pm\frac{\sqrt{5}}{4},$$

de donde obtenemos los posibles extremos

$$(x,y,z,\lambda,\mu)=\left(\frac{\sqrt{5}}{5},-2\frac{\sqrt{5}}{5},\frac{5-3\sqrt{5}}{10},\frac{\sqrt{5}}{4},-\frac{1}{2}\right)$$

y

$$(x,y,z,\lambda,\mu)=\left(-\frac{\sqrt{5}}{5},2\frac{\sqrt{5}}{5},\frac{5+3\sqrt{5}}{10},-\frac{\sqrt{5}}{4},-\frac{1}{2}\right).$$

La primera solución no es posible para un máximo, así que nos quedamos con la segunda como candidato a máximo y calculamos el Hessiano

$$HL = \begin{pmatrix} 2\lambda & 0 & 0 \\ 0 & 2\lambda & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

que particularizada nos queda

$$HL\left(-\frac{\sqrt{5}}{5}, 2\frac{\sqrt{5}}{5}, \frac{5+3\sqrt{5}}{10}, -\frac{\sqrt{5}}{4}, -\frac{1}{2}\right) = \begin{pmatrix} -2\frac{\sqrt{5}}{5} & 0 & 0 \\ 0 & -4\frac{\sqrt{5}}{5} & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Es espacio tangente es

$$\begin{aligned} M\left(-\frac{\sqrt{5}}{5}, 2\frac{\sqrt{5}}{5}, \frac{5+3\sqrt{5}}{10}, -\frac{\sqrt{5}}{4}, -\frac{1}{2}\right) &= \{(x,y,z) : -\frac{2\sqrt{5}}{5}x + \frac{4\sqrt{5}}{5}y = 0, 3x+2z=0\} \\ &= \{(x,y,z) : x = 2y, z = -3/2x\} \\ &= \{(2y, y, -3y) : y \in \mathbb{R}\}. \end{aligned}$$

Evaluamos

$$(2y,y,-3y)\cdot\begin{pmatrix} -2\frac{\sqrt{5}}{5} & 0 & 0 \\ 0 & -4\frac{\sqrt{5}}{5} & 0 \\ 0 & 0 & 0 \end{pmatrix}\cdot\begin{pmatrix} 2y \\ y \\ -3y \end{pmatrix} = (2y,y,-3y)\cdot\begin{pmatrix} -4y\frac{\sqrt{5}}{5} \\ -4y\frac{\sqrt{5}}{5} \\ 0 \end{pmatrix} = -12y^2\frac{\sqrt{5}}{5} < 0,$$

por lo que $\left(-\frac{\sqrt{5}}{5}, 2\frac{\sqrt{5}}{5}, \frac{5+3\sqrt{5}}{10}\right)$ es el máximo buscado.

1.8. Ejercicios

1. Encuentra sobre $\mathbb{R}$, los extremos locales y globales de las siguientes funciones:

   a) $f(x) = x^5 + x^4 - \frac{x^3}{3} + 2$ b) $g(x) = (2x+1)^2(x-4)$

2. Halla los extremos locales y globales de $f(x) = x^3 - 12x + 3$ en el intervalo $[-4, 4]$.

3. 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 $(0,0)$.

4. Estudia los extremos relativos y absolutos de las siguientes funciones

$$\begin{array}{l} f(x,y) = x^2 + 3xy - y^2 \\ f(x,y) = x^3 + 3xy^2 - 15x - 12y \\ f(x,y) = \operatorname{sen} x + \operatorname{sen} y + \cos(x+y) \quad 0 < x, y < 2\pi \end{array}$$

5. 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?

6. Encuentra los extremos de las siguientes funciones sobre los conjuntos correspondientes:

$$\begin{array}{ll} f_1(x,y) = xy\left(1 - x^2 - y^2\right) & \text{en } \Omega_1 = [0,1] \times [0,1] \\ f_2(x,y) = xy & \text{en } \Omega_2 = \text{ Triángulo de vértices } (0,0) - (1,0) - (0,1) \end{array}$$

7. Determina, si existen, los extremos relativos y absolutos de la función $f(x,y,z) = y$ sobre el conjunto

$$F = \left\{x^2 + z^2 = 9; x + y + z = 1\right\}$$

8. Resuelve el problema

$$\begin{array}{cc} \text{Optimizar} & f(x,y,z) = x^2 + y^2 + z^2 \\ \text{Sujeto a} & x^2 + y^2 + z^2 + xy + yz + zx - 1 = 0 \\ & x + 2y - 3z = 0 \end{array}$$

9. Halla la mínima distancia entre la recta $x + y = 4$ y la circunferencia $x^2 + y^2 = 1$

10. Maximiza la función $f(x,y,z) = 3 + x^2 + 2y^2 + 4y - 2x + (z-2)^2$, sujeta a la restricción $2x + 4y + z = 0$.

11. Resuelve el problema

$$\begin{array}{cc} \text{Optimizar} & x^2 y \\ \text{Sujeto a} & x^2 + y^2 = 1 \end{array}$$

12. Determina los óptimos de $f(x,y,z) = xy + yz + zx$, sujeto a la restricción $x + y + z = 120$

13. Resuelve el siguiente problema

$$\begin{array}{cc} \text{Optimizar} & -x^2 + 2y + z \\ \text{Sujeto a} & x + 2y - z = 0 \\ & 2y + z = 0 \end{array}$$

14. 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.

15. Halla los óptimos relativos y absolutos, si los hay, que alcanza $f(x, y, z) = z$, sobre el conjunto $F = \left\{x^2 + y^2 \leq 4; x + y + z = 5\right\}$

16. Resuelve el problema

$$\begin{array}{cc} \text{Optimizar} & x^2 + 2(y-1) \\ \text{Sujeto a} & x + y^2 \leq 1 \end{array}$$

17. 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?

18. Halla, si existen, los extremos relativos y absolutos de la función $f(x, y) = x^2 y$ sobre el conjunto de los puntos que cumplen $x^2 + y^2 \leq 1$.

19. 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.

20. 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.

21. 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.

22. 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.