UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

POSGRADO EN CIENCIAS MATEMÁTICAS

FACULTAD DE CIENCIAS

REGULARIDAD Y AUMENTABILIDAD EN PROBLEMAS DE OPTIMIZACIÓN

T E S I S

QUE PARA OBTENER EL GRADO DE
MAESTRA EN CIENCIAS

P R E S E N T A

ELIZABETH GARCILAZO BOTELLO

DIRECTOR: DR. JAVIER FERNANDO ROSENBLUETH LAGUETTE

MÉXICO, D.F. FEBRERO 2012

Índice general

Introducción

El objetivo principal de este trabajo consiste en, por un lado, presentar un resumen detallado de las teorías de regularidad y aumentabilidad para problemas con restricciones en espacios de dimensión finita y, por otro, proponer una posible generalización del concepto de aumentabilidad para cierto tipo de problemas en control óptimo.

Con la idea de entender este objetivo, consideremos el problema de minimizar una función $f : \mathbf{R}^n \to \mathbf{R}$ sujeta a restricciones del tipo $g(x) = 0$. Una posible derivación de la regla de multiplicadores de Lagrange consiste en minimizar una función F de la forma $F(x) = f(x) + \lambda g(x)$ sin restricciones. El nuevo problema, sin restricciones, implica la regla de multiplicadores de Lagrange $F'(x) = f'(x) + \lambda g'(x) = 0$ en el punto mínimo junto con la condición original $g(x) = 0$.

Aunque este procedimiento puede ser útil para recordar la regla de multiplicadores de primer orden, es poco satisfactorio ya que la clase de problemas en los que se puede aplicar es limitado. Por ejemplo, no es aplicable al problema de minimizar la función $f(x, y) = x^2 - 3y - y^2$ sujeta a la restricción $g(x, y) = y = 0$. El Lagrangiano está dado por

$$F(x, y) = f(x, y) + \lambda g(x, y) = x^2 + (\lambda - 3)y - y^2$$

el cual, sin importar el valor de λ, no tiene un mínimo sin restricciones. Sin embargo, $F'(0, 0) = (0, 0)$ cuando λ es el multiplicador correcto de Lagrange, $\lambda = 3$. Por otro lado, si utilizamos la función aumentada $H = f+\lambda g+(\sigma/2)g^2$, en este caso dada por

$$H(x, y) = x^2 + (\lambda - 3)y + \left(\frac{\sigma - 2}{2}\right) y^2,$$

entonces la elección $\lambda = 3$ y $\sigma > 2$ implica que H tiene un mínimo local sin restricciones en el origen y, en ese punto, el gradiente se anula.

Este ejemplo sugiere que, si x_0 es una solución local del problema de minimizar f sujeta a $g = 0$, entonces existen constantes λ y σ tales que x_0 proporciona un mínimo local sin restricciones de la función $H = f + \lambda g + (\frac{\sigma}{2})g^2$. En este caso diremos que el problema original con restricciones es *aumentable.* Veremos cómo la regla de multiplicadores de Lagrange es consecuencia de aumentabilidad y que, a su vez, aumentabilidad es consecuencia de la regla reforzada de multiplicadores de Lagrange.

A través de varios ejemplos veremos cómo un problema puede ser regular sin ser aumentable o aumentable sin ser regular, por lo que ambas hipótesis son alternativas, pero no equivalentes para la existencia de multiplicadores apropiados de Lagrange.

Veremos primeramente algunos conceptos y resultados básicos que se estarán utilizando a lo largo de todo el trabajo, luego presentaremos la teoría de regularidad y posteriormente estudiaremos el método de aumentabilidad. Finalmente veremos cómo aplicar este método al caso con igualdades en control óptimo y algunos ejemplos de ello.

Capítulo 1

Definiciones y conceptos básicos

Comenzaremos en este primer capítulo repasando algunos conceptos y resultados básicos que estaremos utilizando a lo largo de todo el trabajo.

Vamos a suponer dados un subconjunto S de $\mathbf{R}^n$ y una función f que mapea S en $\mathbf{R}$. Nuestro propósito será minimizar la función f sobre el conjunto S.

Veremos algunos resultados básicos relacionados con diferenciales de funciones de valor real definidas en un subconjunto de $\mathbf{R}^n$ así como una versión del teorema de la función implícita.

Posteriormente derivaremos de manera sencilla un resultado crucial (Teorema 1.23) en la determinación de multiplicadores de Lagrange que utilizaremos para derivar condiciones de optimalidad para problemas con igualdades. Para obtener un resultado equivalente aplicable a problemas que involucran desigualdades (Teorema 1.29) introduciremos una breve teoría sobre conos convexos. Por último, para la derivación de condiciones de optimalidad para problemas con desigualdades presentaremos los conceptos de conos tangentes y funciones soporte.

1.1. Mínimos y máximos

Supongamos dados un subconjunto S de $\mathbf{R}^n$ y una función f que mapea S en $\mathbf{R}$.

Definiciones 1.1. A un punto de S que maximiza o minimiza a f le llamamos *punto extremo* de f sobre S.

Una condición que asegura la existencia de un punto extremo de una función está dada por el siguiente teorema. El resultado es una consecuencia inmediata del Teorema de Bolzano-Weierstrass el cual afirma que toda sucesión acotada de puntos en $\mathbf{R}^n$ tiene una subsucesión convergente.

Teorema 1.2. *Si S es un compacto no vacío de $\mathbf{R}^n$ y $f: S \to \mathbf{R}$ es continua entonces f alcanza su máximo y mínimo en S.*

Demostración: Sea $m := \text{ínf}\, f(x)$ en S y sea $\{x_q\}$ una sucesión de puntos en S tales que $m = \text{lím}_{q\to\infty} f(x_q)$. Como S está acotado, por el teorema de Bolzano-Weierstrass $\{x_q\}$ tiene una subsucesión $\{y_q\}$ convergente y, por ser S cerrado, el límite $x_0 := \text{lím}_{q\to\infty} y_q$ está en S. Por continuidad,

$$f(x_0) = \lim_{q\to\infty} f(y_q) = m = \text{ínf}\, f(x) \text{ en } S.$$

Por lo tanto x_0 minimiza a f en S. Aplicando este resultado a $-f$, se prueba que existe un punto x_1 en S que minimiza a $-f$ en S y, por lo tanto, x_1 maximiza a f en S. ∎

El concepto de extremo es de naturaleza global dado que se toman en cuenta todos los puntos de S, x_0 minimiza a f en S si $x_0 \in S$ y $f(x_0) \leq f(x)$ para toda $x \in S$. En contraste, extremos locales toman en cuenta el comportamiento de la función en vecindades alrededor de x_0.

Definiciones 1.3. Diremos que $x_0 \in S$ *minimiza localmente a f en S* si existe una vecindad N de x_0 tal que x_0 minimiza a f en $S \cap N$, es decir, $f(x_0) \leq f(x)$ para toda $x \in S \cap N$. Si la desigualdad es estricta para toda $x \neq x_0$ en $S \cap N$, diremos que $f(x_0)$ es un *mínimo local estricto de f en S*.

1.2. Extremos en un intervalo compacto

En esta sección repasaremos condiciones necesarias y suficientes de optimalidad para el caso en que la función f está definida en un intervalo compacto $S = [a, b]$ en $\mathbf{R}$.

Recordemos que la derivada $f'(x_0)$ de f en x_0, cuando existe, está dada por

$$f'(x_0) = \lim_{x\to x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{h\to 0} \frac{f(x_0 + h) - f(x_0)}{h}.$$

Si $x_0 = a$ o $x_0 = b$, este límite es un límite por la derecha o izquierda respectivamente. Suponiendo la existencia de derivadas, tenemos las siguientes condiciones suficientes para un mínimo local estricto en los puntos inicial y final de $[a, b]$.

Teorema 1.4. *Si $f'(x_0) > 0$ y $x_0 = a$, o $f'(x_0) < 0$ y $x_0 = b$, entonces $f(x_0)$ es un mínimo local estricto de f en $[a, b]$. Análogamente, si $f'(x_0) < 0$ y $x_0 = a$, o $f'(x_0) > 0$ y $x_0 = b$, entonces $f(x_0)$ es un máximo local estricto de f en $[a, b]$.*

Demostración: Supongamos que $f'(a) > 0$. Sea $\epsilon := f'(a)/2$ y sea $\delta > 0$ tal que

$$a < x < a + \delta \ \Rightarrow \ \left| \frac{f(x) - f(a)}{x - a} - f'(a) \right| < \epsilon.$$

Entonces

$$\frac{f(x) - f(a)}{x - a} > f'(a) - \epsilon = \frac{f'(a)}{2} > 0 \quad (a < x < a + \delta).$$

Por lo tanto $f(x) > f(a)$ para toda $x \in (a, a + \delta)$, o sea, $f(a)$ es un mínimo local estricto de f en $[a, b]$. Las afirmaciones restantes se prueban de manera semejante. ∎

Este resultado implica de manera inmediata las siguientes condiciones necesarias para mínimos locales en los puntos inicial y final de $[a, b]$.

Teorema 1.5. *Si $f'(a)$ existe y $f(a)$ es un mínimo local de f en $[a, b]$ entonces $f'(a) \geq 0$. Si $f'(b)$ existe y $f(b)$ es un mínimo local de f en $[a, b]$ entonces $f'(b) \leq 0$.*

Demostración: Por el Teorema 1.4, $f(a)$ es un máximo local estricto de f en $[a, b]$ si $f'(a) < 0$. Por lo tanto, si $f(a)$ es un mínimo local de f en $[a, b]$, se tiene que $f'(a) \geq 0$. Análogamente, $f'(b) \leq 0$ si $f(b)$ es un mínimo local de f en $[a, b]$. ∎

Para puntos interiores del intervalo $[a, b]$, el teorema anterior implica el siguiente resultado.

Teorema 1.6. *Si f tiene un extremo local en $x_0 \in (a, b)$ y $f'(x_0)$ existe entonces $f'(x_0) = 0$.*

Demostración: Supongamos que f tiene un mínimo local en x_0. Por el Teorema 1.5 aplicado en el intervalo $[x_0, b]$ se tiene que $f'(x_0) \geq 0$. Aplicando el mismo teorema en el intervalo $[a, x_0]$ vemos que $f'(x_0) \leq 0$. Por lo tanto $f'(x_0) = 0$. Análogamente $f'(x_0) = 0$ si f tiene un máximo local en $x = x_0$. ∎

El siguiente teorema del valor medio es básico.

Teorema 1.7. *Si f es continua en $[a, b]$ y tiene derivada en todo punto de (a, b) entonces existe $c \in (a, b)$ tal que*

$$f(b) = f(a) + (b - a)f'(c).$$

Demostración: Sea

$$h(x) := f(x) - \left[\frac{f(b) - f(a)}{b - a}\right](x - a) \quad (x \in [a, b]).$$

Claramente h es continua en $[a, b]$, derivable en (a, b), $h(a) = f(a)$ y

$$h(b) = f(b) - \left[\frac{f(b) - f(a)}{b - a}\right](b - a) = f(a).$$

Por el Teorema 1.2 existe c en (a, b) punto extremo de h. Por el Teorema 1.6,

$$0 = h'(c) = f'(c) - \frac{f(b) - f(a)}{b - a}$$

y, por lo tanto,

$$f'(c) = \frac{f(b) - f(a)}{b - a}. \blacksquare$$

Recordemos que si f es diferenciable en $[a, b]$ y $f''(x_0)$ existe, el teorema de Taylor asegura que

$$f(x) - f(x_0) = f'(x_0)(x - x_0) + \tfrac{1}{2}f''(x_0)(x - x_0)^2 + r(x - x_0)$$

donde

$$\lim_{h \to 0} \frac{r(h)}{h^2} = 0.$$

El siguiente teorema nos da condiciones suficientes para un extremo local.

Teorema 1.8. *Supongamos que f es diferenciable en una vecindad de x_0, $f''(x_0)$ existe y $f'(x_0) = 0$. Entonces*

a. $f''(x_0) > 0 \Rightarrow f(x_0)$ *es un mínimo local estricto de f.*

b. $f''(x_0) < 0 \Rightarrow f(x_0)$ *es un máximo local estricto de f.*

Demostración:

(a): Sea m tal que $0 < 2m < f''(x_0)$ y sea $\delta > 0$ tal que

$$\left|\frac{r(h)}{h^2}\right| \leq \frac{f''(x_0)}{2} - m \quad (0 < |h| < \delta).$$

Si $|x - x_0| < \delta$, entonces

$$f(x) - f(x_0) = \tfrac{1}{2} f''(x_0)(x - x_0)^2 + r(x - x_0) \geq m(x - x_0)^2.$$

(b): De manera análoga obtenemos que existe $\delta > 0$ tal que, si $|x - x_0| < \delta$, entonces

$$f(x) \leq f(x_0) - m(x - x_0)^2. \blacksquare$$

Una consecuencia inmediata de los Teoremas 1.4 y 1.8 son las siguientes condiciones necesarias para un extremo local.

Teorema 1.9. *Supongamos que f es diferenciable en una vecindad de x_0, $f''(x_0)$ existe y f tiene un mínimo local en x_0. Entonces*
a. $x_0 \in (a, b) \Rightarrow f'(x_0) = 0$ *y* $f''(x_0) \geq 0$.
b. $x_0 = a \Rightarrow f'(x_0) > 0$ *o bien* $f'(x_0) = 0$ *y* $f''(x_0) \geq 0$.
c. $x_0 = b \Rightarrow f'(x_0) < 0$ *o bien* $f'(x_0) = 0$ *y* $f''(x_0) \geq 0$.

1.3. Funciones de varias variables

En esta sección repasaremos algunos resultados básicos relacionados con diferenciales de funciones con valores reales definidas en un subconjunto de $\mathbf{R}^n$.

Dados X, Y espacios vectoriales sobre $\mathbf{R}$, denotamos por $L(X, Y)$ al conjunto de todas las transformaciones lineales de X en Y y cuando $Y = \mathbf{R}$, escribiremos simplemente $L(X)$.

Definición 1.10. Sean $E \subset \mathbf{R}^n$ abierto, f una función que mapea E en $\mathbf{R}$ y $x_0 \in E$. Decimos que f es *diferenciable en* x_0 si existe $A \in L(\mathbf{R}^n)$ tal que

$$\lim_{x \to x_0} \frac{f(x) - f(x_0) - A(x - x_0)}{|x - x_0|} = 0.$$

En este caso escribimos $f'(x_0) = A$ y llamamos a $f'(x_0)$ la *diferencial de f en* x_0. Si f es diferenciable en x para toda $x \in E$ decimos que f es diferenciable en E. Nótese que, si f es diferenciable en x_0, entonces

$$f(x) - f(x_0) = f'(x_0)(x - x_0) + r(x - x_0) \qquad \text{donde} \quad \lim_{h \to 0} \frac{r(h)}{|h|} = 0.$$

Definición 1.11. Sean $f\colon E \subset \mathbf{R}^n \to \mathbf{R}$ con E abierto y $\{e_1, \ldots, e_n\}$ la base canónica de $\mathbf{R}^n$. Para $x \in E$, $1 \leq i \leq n$, definimos

$$(D_i f)(x) = \frac{\partial f(x)}{\partial x_i} := \lim_{t \to 0} \frac{f(x + te_i) - f(x)}{t}$$

siempre que el límite exista, y llamamos a $D_i f$ una *derivada parcial*.

Teorema 1.12. *Supongamos que f mapea $E \subset \mathbf{R}^n$ en $\mathbf{R}$ y f es diferenciable en $x \in E$. Entonces las derivadas parciales $(D_i f)(x)$ existen y*

$$f'(x)e_i = (D_i f)(x).$$

Demostración: Sea $1 \leq i \leq n$. Como f es diferenciable en x,

$$f(x + te_i) - f(x) = f'(x)(te_i) + r(te_i) \qquad \text{donde} \quad \lim_{t \to 0} \frac{r(te_i)}{t} = 0.$$

Como $f'(x)$ es lineal,

$$f'(x)e_i = \lim_{t \to 0} \frac{f(x + te_i) - f(x)}{t} = (D_i f)(x). \ ∎$$

Definición 1.13. Sean $E \subset \mathbf{R}^n$ abierto y $f\colon E \to \mathbf{R}$ diferenciable. Decimos que f es *continuamente diferenciable en E* si f' es una función continua de E en $L(\mathbf{R}^n)$. Explícitamente requerimos que, para cada $x \in E$ y $\epsilon > 0$, exista $\delta > 0$ tal que

$$y \in E \text{ y } |x - y| < \delta \Rightarrow \|f'(y) - f'(x)\| < \epsilon.$$

En este caso escribimos $f \in C^1(E)$.

Teorema 1.14. *Supongamos que f mapea un abierto $E \subset \mathbf{R}^n$ en $\mathbf{R}$. Entonces son equivalentes:*

a. $f \in C^1(E)$.

b. *Las derivadas parciales $D_i f$ existen y son continuas en E para toda $1 \leq i \leq n$.*

Observación 1.15. Si para toda $1 \leq i \leq n$ las derivadas parciales $D_i f$ existen, denotamos al vector $((D_1 f)(x), \ldots, (D_n f)(x))$ como

$$((D_1 f)(x), \ldots, (D_n f)(x)) = (\nabla f)(x)$$

y lo llamamos el *gradiente* de f en x, i.e.,

$$(\nabla f)(x) = \sum_{i=1}^{n} (D_i f)(x) e_i.$$

Ahora, dada $A \in L(\mathbf{R}^n)$, el vector $a = (Ae_1, \ldots, Ae_n)$ satisface $Ah = \langle a, h \rangle$ para toda $h \in \mathbf{R}^n$ donde $\langle \cdot, \cdot \rangle$ es el producto interno en $\mathbf{R}^n$. Por otro lado, si $a \in \mathbf{R}^n$, definimos $A \colon \mathbf{R}^n \to \mathbf{R}$ como $Ah := \langle a, h \rangle$ y, claramente, $A \in L(\mathbf{R}^n)$. Por lo tanto existe una correspondencia uno-a-uno entre $\mathbf{R}^n$ y $L(\mathbf{R}^n)$. Por el Teorema 1.12, si f es diferenciable en x, entonces

$$f'(x)h = \langle (\nabla f)(x), h \rangle \quad \text{para toda } h \in \mathbf{R}^n.$$

Debido a esta igualdad, es usual utilizar la notación $f'(x)$ para el gradiente de f en x.

Visto de otra manera, si en la Definición 1.10 escribimos $f'(x_0; \cdot) := A(\cdot)$ y llamamos a $f'(x_0; \cdot)$ la diferencial de f en x_0 entonces, con la correspondencia uno-a-uno entre $\mathbf{R}^n$ y $L(\mathbf{R}^n)$ de arriba, le asociamos a A el vector $f'(x_0)$ tal que $\langle f'(x_0), h \rangle = f'(x_0; h)$ para toda $h \in \mathbf{R}^n$ y lo llamamos el gradiente de f en x_0.

Definición 1.16. Una *forma cuadrática* Q en $\mathbf{R}^n$ es una función $Q \colon \mathbf{R}^n \to \mathbf{R}$ tal que, para toda $x = (x_1, \ldots, x_n) \in \mathbf{R}^n$,

$$Q(x) = \sum_{i,j=1}^{n} a_{ij} x_i x_j$$

donde $a_{ij} \in \mathbf{R}$ y $A = (a_{ij})$ es una matriz simétrica. Nótese que

$$Q(x) = \langle Ax, x \rangle = \langle x, Ax \rangle = x^* Ax$$

donde x es un vector columna y x^* su transpuesta. Decimos que Q es *positiva*, *negativa*, etc., si lo es $Q(x)$ para toda $x \neq 0$. Análogamente, a toda matriz simétrica $A = (a_{ij})$ le asociamos la forma cuadrática $Q(x) := \langle Ax, x \rangle$ para toda $x \in \mathbf{R}^n$. La matriz A es *positiva*, *negativa*, etc., si lo es su forma cuadrática asociada.

Observación 1.17. Dada una forma cuadrática Q en $\mathbf{R}^n$, existen vectores unitarios $x_0, x_1 \in \mathbf{R}^n$ tales que, para toda $x \in \mathbf{R}^n$,

$$Q(x_0)|x|^2 \leq Q(x) \leq Q(x_1)|x|^2.$$

Demostración: Sea $E := \{x \in \mathbf{R}^n : |x| = 1\}$ la $(n-1)$-esfera. Como E es compacto, existen puntos $x_0, x_1 \in E$ tales que, para toda $x \in E$,

$$Q(x_0) \leq Q(x) \leq Q(x_1).$$

Por lo tanto, para toda $x \in \mathbf{R}^n - \{0\}$,

$$Q(x_0) \leq Q\left(\frac{x}{|x|}\right) = \frac{Q(x)}{|x|^2} \leq Q(x_1)$$

y el resultado se sigue. ∎

Definición 1.18. Sean $E \subset \mathbf{R}^n$ abierto, f una función que mapea E en $\mathbf{R}$ y $x_0 \in E$. Decimos que f tiene una *segunda diferencial en* x_0 si f tiene una diferencial $f'(x_0;\cdot)$ en x_0 y existe una forma cuadrática Q en $\mathbf{R}^n$ tal que

$$\lim_{x \to x_0} \frac{f(x) - f(x_0) - f'(x_0; x - x_0) - \frac{1}{2}Q(x - x_0)}{|x - x_0|^2} = 0.$$

En este caso escribimos $f''(x_0;\cdot) := Q(\cdot)$ y llamamos a $f''(x_0;\cdot)$ la *segunda diferencial de* f *en* x_0. La matriz simétrica $f''(x_0)$ tal que para toda $h \in \mathbf{R}^n$

$$\langle f''(x_0)h, h\rangle = f''(x_0; h)$$

se llama el *Hessiano de* f *en* x_0. Nótese que, si f tiene una segunda diferencial en x_0, entonces

$$f(x) - f(x_0) = f'(x_0; x - x_0) + \tfrac{1}{2}f''(x_0; x - x_0) + r(x - x_0)$$

donde

$$\lim_{h \to 0} \frac{r(h)}{|h|^2} = 0.$$

Definición 1.19. Supongamos que f es una función real definida en un abierto $E \subset \mathbf{R}^n$, con derivadas parciales $D_1 f, \ldots, D_n f$. Si las funciones $D_i f$ son a su vez diferenciables, las *derivadas parciales de segundo orden* de f están definidas como

$$D_{ij} f := D_i D_j f \quad (i, j = 1, \ldots, n).$$

Si estas funciones $D_{ij} f$ son continuas en E, decimos que f es de clase C^2 en E y escribimos $f \in C^2(E)$. Análogamente, $f \in C^m(E)$ si f es continua y posee derivadas parciales continuas de grado menor o igual que m en E. Diremos que f es de clase C^m en un conjunto $E \subset \mathbf{R}^n$ arbitrario si f es de clase C^m en una vecindad de E.

Teorema 1.20. (Taylor) *Dados* $S \subset \mathbf{R}^n$ *y* $f\colon S \to \mathbf{R}$ *supongamos que* $x + th \in S$ *para toda* $t \in [0, 1]$. *Entonces:*

a. $f \in C^1(S) \Rightarrow$ *existe* $t_1 \in (0, 1)$ *tal que*

$$\begin{aligned} f(x+h) &= f(x) + f'(x + t_1 h; h) \\ &= f(x) + \int_0^1 f'(x + th; h)dt. \end{aligned}$$

b. $f \in C^2(S) \Rightarrow$ *existe* $t_2 \in (0, 1)$ *tal que*

$$\begin{aligned} f(x+h) &= f(x) + f'(x; h) + \tfrac{1}{2} f''(x + t_2 h; h) \\ &= f(x) + f'(x; h) + \int_0^1 (1-t) f''(x + th; h)dt. \end{aligned}$$

El teorema de la función implícita juega un papel fundamental en la teoría que presentaremos en el siguiente capítulo.

Teorema 1.21. (Función implícita) *Sea* S *un abierto de* $\mathbf{R}^m \times \mathbf{R}^n$ *y* f *una función que mapea* S *en* $\mathbf{R}^n$ *con* $f(t, x)$ *y* $f_x(t, x)$ *continuas en* S. *Supongamos que existen un compacto* $T_0 \subset \mathbf{R}^m$ *y una función continua* $x_0\colon T_0 \to \mathbf{R}^n$ *tales que, para toda* $t \in T_0$,

i. $(t, x_0(t)) \in S$.

ii. $f(t, x_0(t)) = 0$ *y* $|f_x(t, x_0(t))| \neq 0$.

Entonces existen una vecindad T *de* T_0, *una función continua* $x\colon T \to \mathbf{R}^n$ *y una constante* $\epsilon > 0$ *tales que*

a. $x = x_0$ *en* T_0.

b. $f(t, x(t)) = 0$ *para toda* t *en* T.

c. $(t \in T,\ f(t, x) = 0$ *y* $|x - x(t)| < \epsilon) \Rightarrow x = x(t)$.

d. $f \in C^m(S) \Rightarrow x \in C^m(T)$.

1.4. Funcionales lineales

Teorema 1.22. *Sean* X *un espacio vectorial y* $\{L_i\}_1^m$ *funcionales lineales en* X. *Entonces las siguientes propiedades son equivalentes:*

a. $\{L_i\}$ *es linealmente independiente.*

b. *Existen* $x_1, \ldots, x_m \in X$ *tales que* $|L_i(x_j)| \neq 0$ $(i, j = 1, \ldots, m)$.

Demostración: Sean $F(x) := (L_1(x), \ldots, L_m(x))$ $(x \in X)$ y $Y := F(X)$. Entonces (a) $\Leftrightarrow$ no existe ningún vector en $\mathbf{R}^m$ ortogonal a $Y \Leftrightarrow Y = \mathbf{R}^m$

(ya que Y es un subespacio de $\mathbf{R}^m$) $\Leftrightarrow$ existen m vectores linealmente independientes $y_1, \ldots, y_m \in Y \Leftrightarrow$ existen $x_1, \ldots, x_m \in X$ tales que $y_i = (L_1(x_i), \ldots, L_m(x_i))$ son linealmente independientes $\Leftrightarrow$ (b). ∎

En particular nótese que, si $X = \mathbf{R}^n$, las funcionales lineales $L_i(x) = \langle a_i, x\rangle$ $(i = 1, \ldots, m)$ son linealmente independientes $\Leftrightarrow A = (a_{ij})$ es de rango m.

Teorema 1.23. *Sean X un espacio vectorial, L, L_i $(i \in A := \{1, \ldots, m\})$ funcionales lineales en X,*

$$R = \{x \in X \mid L_i(x) = 0 \ (i \in A)\},$$

y supongamos que $L(x) = 0$ para toda $x \in R$. Entonces existen multiplicadores $\{\lambda_i\}_1^m$ tales que $L(x) = \sum_1^m \lambda_i L_i(x)$ $(x \in X)$. Si $\{L_i\}_1^m$ es linealmente independiente entonces los multiplicadores son únicos.

Demostración: Sin pérdida de generalidad supongamos que $\{L_i\}_1^m$ es linealmente independiente. Por el teorema 1.22, existen $x_1, \ldots, x_m \in X$ tales que $|L_i(x_j)| \neq 0$ $(i, j \in A)$. Sean

$$y_i := \begin{pmatrix} L_i(x_1) \\ \vdots \\ L_i(x_m) \end{pmatrix} \quad (i \in A),$$

$$C := (y_1, \ldots, y_m) = \begin{pmatrix} L_1(x_1) & \cdots & L_m(x_1) \\ \vdots & & \vdots \\ L_1(x_m) & \cdots & L_m(x_m) \end{pmatrix}$$

Sea $\lambda = (\lambda_1, \ldots, \lambda_m)^*$ tal que $C\lambda = (L(x_1), \ldots, L(x_m))^*$, por lo que

$$L(x_j) = \sum_1^m \lambda_i L_i(x_j) \quad (j \in A)$$

Sea $x \in X$, definamos $u := (L_1(x), \ldots, L_m(x))^*$ y sea $b = (b_1, \ldots, b_m)^*$ tal que $C^* b = u$, de manera que

$$0 = L_i(x) - \sum_{j=1}^m L_i(x_j) b_j = L_i\left(x - \sum_1^m x_j b_j\right) \qquad (i \in A).$$

Como $x - \sum_1^m x_j b_j \in R$, se tiene $L(x - \sum_1^m x_j b_j) = 0$. Por lo tanto

$$\begin{aligned} 0 &= L(x) - \sum_1^m L(x_j) b_j = L(x) - \sum_{i,j=1}^m \lambda_i L_i(x_j) b_j \\ &= L(x) - \sum_1^m \lambda_i L_i(x). \blacksquare \end{aligned}$$

1.5. Conos convexos

Definiciones 1.24. Sea C un subconjunto de $\mathbf{R}^n$. Decimos que
i. C es *convexo* si para toda $x, y \in C$ y $\lambda \in [0,1]$, $(1-\lambda)x + \lambda y \in C$.
ii. C es un *cono* si $x \in C \Rightarrow \alpha x \in C$ para toda $\alpha \geq 0$.

Observación 1.25. Sea C un cono en $\mathbf{R}^n$. Entonces son equivalentes:
a. C es convexo.
b. $x, y \in C \Rightarrow x + y \in C$.
Demostración:
(a) $\Rightarrow$ (b): Sean $x, y \in C$. Por (a), $z := \frac{1}{2}(x+y) \in C$. Como C es un cono, $2z = x + y \in C$.
(b) $\Rightarrow$ (a): Sean $x, y \in C$ y $0 \leq \lambda \leq 1$. Como C es un cono, $(1-\lambda)x$ y λy están en C y, por (b), $(1-\lambda)x + \lambda y \in C$. $\blacksquare$

Definición 1.26. Dado $B \subset \mathbf{R}^n$ definimos a $C(B)$ como el conjunto de todas las combinaciones lineales no negativas finitas de elementos de B, es decir,

$$C(B) := \left\{ \sum_J a_i x_i \mid a_i \geq 0, x_i \in B, J \text{ finito} \right\}.$$

Claramente $C(B)$ es un cono convexo y, de hecho, es el mínimo cono convexo que contiene a B. Llamamos a $C(B)$ el *cono convexo generado por* B.

Definición 1.27. Dado $B \subset \mathbf{R}^n$ definimos

$$B^* := \{ z \in \mathbf{R}^n \mid \langle y, z \rangle \leq 0 \text{ para toda } y \in B \}$$

y lo llamamos el *cono dual de* B.

Mencionaremos ahora algunas propiedades básicas de conos convexos.

Proposición 1.28. *Las siguientes propiedades se cumplen:*
a. *Si $B \subset \mathbf{R}^n$ entonces B^* es un cono convexo cerrado y*

$$B^* = (C(B))^* = (\overline{C(B)})^*.$$

b. *Si C es un cono convexo entonces $C^{**} = \overline{C}$.*
c. *Un cono convexo con un número finito de generadores es cerrado.*

Como se verá en el Capítulo 2, el siguiente resultado es crucial en la obtención de condiciones necesarias de optimalidad para problemas con igualdades y desigualdades.

Teorema 1.29. *Sean $A = \{1, \ldots, p\}$, $B = \{p+1, \ldots, m\}$ y $b_\alpha \in \mathbf{R}^n$ $(\alpha \in A \cup B)$. Entonces*

$$R := \{h \in \mathbf{R}^n \mid \langle b_\alpha, h\rangle \leq 0 \ (\alpha \in A),\ \langle b_\beta, h\rangle = 0 \ (\beta \in B)\}$$

es un cono convexo cerrado. Su dual R^* está generado por los vectores

$$b_1, \ldots, b_m, -b_{p+1}, \ldots, -b_m.$$

Más aún, para toda $b \in \mathbf{R}^n$, las siguientes afirmaciones son equivalentes:
a. *$\langle b, h\rangle \leq 0$ para toda $h \in \mathbf{R}^n$.*
b. *Existen $\lambda_1, \ldots, \lambda_m$ en $\mathbf{R}$ con $\lambda_\alpha \geq 0$ $(\alpha \in A)$ tales que $b = \sum_1^m \lambda_i b_i$.*
Demostración: Nótese primero que

$$R = \{h \in \mathbf{R}^n \mid \langle b_\alpha, h\rangle \leq 0 \ (\alpha \in A \cup B),\ \langle -b_\beta, h\rangle \leq 0 \ (\beta \in B)\}.$$

Por lo tanto $R = C(B)^*$ donde

$$B = \{b_1, \ldots, b_p, b_{p+1}, \ldots, b_m, -b_{p+1}, \ldots, -b_m\}.$$

Por la Proposición 1.28 (a), R es un cono convexo cerrado. Ahora, $C(B)$ es cerrado por 1.28 (c) y por lo tanto, por 1.28 (b), $R^* = C(B)^{**} = C(B)$. ∎

1.6. Conos tangentes

Definición 1.30. Una sucesión $\{x_m\} \subset \mathbf{R}^n$ *converge a x_0 en la dirección h* si h es un vector unitario, $x_m \neq x_0$, y

$$\lim_{m\to\infty} |x_m - x_0| = 0, \qquad \lim_{m\to\infty} \frac{x_m - x_0}{|x_m - x_0|} = h.$$

Claramente, si una sucesión $\{x_m\}$ converge a x_0 y $x_m \neq x_0$ para un número infinito de m's, entonces $\{x_m\}$ tiene una subsucesión que converge direccionalmente.

Definición 1.31. Dada $x_0 \in S \subset \mathbf{R}^n$, el *cono tangente de S en x_0*, denotado por $T_S(x_0)$, es el cono (cerrado) determinado por los vectores unitarios h para los cuales existe una sucesión $\{x_m\}$ en S convergente a x_0 en la dirección h.

Nótese que, si $\{x_m\}$ converge a x_0 en la dirección h y f es diferenciable en x_0 entonces

$$\lim_{m\to\infty} \frac{f(x_m) - f(x_0)}{|x_m - x_0|} = f'(x_0; h).$$

Si f tiene segunda diferencial en x_0 entonces

$$\lim_{m\to\infty} \frac{f(x_m) - f(x_0) - f'(x_0; x_m - x_0)}{|x_m - x_0|^2} = \frac{1}{2} f''(x_0; h).$$

Proposición 1.32. *Dada $x_0 \in S \subset \mathbf{R}^n$ sea $C_S^+(x_0)$ el conjunto de todas las $h \in \mathbf{R}^n$ para las cuales existen $\epsilon > 0$ y $x\colon [0, \epsilon) \to S$ tales que $x(0) = x_0$ y $\dot{x}(0) = h$. Entonces $C_S^+(x_0) \subset T_S(x_0)$.*

Demostración: Sea $h \in C_S^+(x_0)$ y sean $\epsilon > 0$ y $x : [0, \epsilon) \to S$ tales que $x(0) = x_0$ y $\dot{x}(0) = h$. Definimos para toda $m \in \mathbf{N}$, $t_m := \epsilon/m$ y $x_m := x(t_m)$. Entonces

$$\lim_{m\to\infty} \frac{x_m - x_0}{t_m} = \lim_{m\to\infty} \frac{x(t_m) - x(0)}{t_m} = \dot{x}(0) = h.$$

Así $x_m \to x_0$, $m \to \infty$ y, si $h \neq 0$, entonces $x_m \neq x_0$ para m suficientemente grande y

$$\lim_{m\to\infty} \frac{x_m - x_0}{|x_m - x_0|} = \lim_{m\to\infty} \frac{x_m - x_0}{t_m} \lim_{m\to\infty} \frac{t_m}{|x_m - x_0|} = \frac{h}{|h|}$$

lo cual implica que $\{x_m\}$ converge a x_0 en la dirección $h/|h|$. ∎

Definición 1.33. Sean $x_0 \in S \subset \mathbf{R}^n$ y $f\colon \mathbf{R}^n \to \mathbf{R}$. Decimos que $F\colon \mathbf{R}^n \to \mathbf{R}$ es una *función de soporte inferior para f en (S, x_0)* si F tiene las mismas propiedades de diferenciabilidad que f en x_0, $F(x) \leq f(x)$ $(x \in S)$, $F(x_0) = f(x_0)$, y $F'(x_0) = 0$.

Observación 1.34. Sean $x_0 \in S \subset \mathbf{R}^n$ y $f\colon \mathbf{R}^n \to \mathbf{R}$. Si f tiene una función de soporte inferior F en (S, x_0) entonces $f'(x_0; h) \geq 0$ para toda $h \in T_S(x_0)$.

Demostración: Sea $h \in T_S(x_0)$ unitario y sea $\{x_m\} \subset S$ convergente a x_0 en la dirección h. Si $g := f - F$ entonces $g(x) \geq 0 = g(x_0)$ para toda $x \in S$, lo cual implica que

$$0 \leq \lim_{m\to\infty} \frac{g(x_m)}{|x_m - x_0|} = g'(x_0; h) = f'(x_0; h). \blacksquare$$

Capítulo 2

Regularidad

Estudiaremos primeramente el caso de restricciones con igualdades y posteriormente extenderemos los resultados al caso de restricciones con desigualdades.

2.1. Restricciones con igualdades

A lo largo de este capítulo, supongamos dadas funciones f, g_α que mapean $\mathbf{R}^n$ en $\mathbf{R}$, con $\alpha \in A = \{1, \ldots, m\}$ y $m < n$. Denotaremos por $P(S)$ al problema de minimizar f sobre el conjunto

$$S = \{x \in \mathbf{R}^n \mid g_\alpha(x) = 0 \ (\alpha \in A)\}.$$

Para condiciones de primer orden supondremos que las funciones involucradas son de clase C^1 y, para condiciones de segundo orden, de clase C^2.

Para toda $x_0 \in S$ definamos los conjuntos

$$\begin{aligned} C_S(x_0) \ &:= \ \{h \in \mathbf{R}^n \mid \text{existen } \epsilon > 0 \text{ y } x\colon (-\epsilon, \epsilon) \to S \text{ tales que} \\ &\qquad\qquad\qquad\qquad\qquad x(0) = x_0 \text{ y } \dot{x}(0) = h\} \end{aligned}$$

llamado el conjunto de *vectores tangentes curvilíneos de* S *en* x_0, y

$$R_S(x_0) := \{h \in \mathbf{R}^n \mid g'_\alpha(x_0; h) = 0 \ (\alpha \in A)\}$$

que corresponde al conjunto de vectores que satisfacen las *restricciones tangenciales de* S *en* x_0. Claramente $C_S(x_0) \subset R_S(x_0)$ para toda $x_0 \in S$, pero el converso no necesariamente es cierto.

En general resulta difícil determinar cuándo ambos conjuntos coinciden. Un criterio ampliamente utilizado en la literatura es el de normalidad.

Definición 2.1. Sea $x_0 \in S$. Decimos que x_0 es un

a. *punto regular de* S si $C_S(x_0) = R_S(x_0)$.

b. *punto normal de* S si las ecuaciones lineales $g'_\alpha(x_0; h) = 0$ $(\alpha \in A)$ con $h \in R_S(x_0)$ son linealmente independientes, esto es, si los gradientes $g'_1(x_0), \ldots, g'_m(x_0)$ son linealmente independientes, lo cual es equivalente a requerir que la matriz

$$\left(\frac{\partial g_\alpha(x_0)}{\partial x^i}\right) \qquad (\alpha = 1, \ldots, m;\ i = 1, \ldots, n)$$

sea de rango m.

Como veremos al final de esta sección, si x_0 es un punto es normal de S, entonces x_0 es también un punto regular de S.

Para toda $(x, \lambda) \in \mathbf{R}^n \times \mathbf{R}^m$ sea

$$F(x, \lambda) := f(x) + \sum_1^m \lambda_\alpha g_\alpha(x).$$

Teorema 2.2. *Supongamos que x_0 es un punto regular de S y una solución local de* $\mathrm{P}(S)$*. Entonces existe $\lambda \in \mathbf{R}^m$ tal que $F_x(x_0, \lambda) = 0$ y $F_{xx}(x_0, \lambda; h) \geq 0$ para toda $h \in R_S(x_0)$.*

Demostración: Sea $h \in R_S(x_0) = C_S(x_0)$ y sean $\epsilon > 0$ y $x \colon (-\epsilon, \epsilon) \to S$ tales que $x(0) = x_0$ y $\dot{x}(0) = h$. Entonces $\varphi := f \circ x$ tiene un mínimo local en $t = 0$ y por lo tanto

$$0 = \varphi'(0) = f'(x_0; h) \quad \text{y} \quad 0 \leq \varphi''(0) = f''(x_0; h).$$

La primera conclusión se sigue del Teorema 1.23. Como $\varphi(t) = F(x(t), \lambda)$, por Taylor tenemos

$$\frac{\varphi(t) - \varphi(0)}{t^2} = \frac{F(x(t), \lambda) - F(x_0, \lambda)}{t^2} = \frac{1}{2} F_{xx}\left((\bar{x}(t), \lambda); \frac{x(t) - x_0}{t}\right)$$

donde $\bar{x}(t) = x_0 + \theta(t)[x(t) - x_0]$ con $0 < \theta(t) < 1$. Por lo tanto

$$0 \leq \frac{1}{2}\varphi''(0) = \lim_{t \to 0} \frac{\varphi(t) - \varphi(0)}{t^2} = \frac{1}{2} F_{xx}((x_0, \lambda); h). \ \blacksquare$$

Teorema 2.3. *Supongamos que $x_0 \in S$ y que existe $\lambda \in \mathbf{R}^m$ tal que $F_x(x_0, \lambda) = 0$ y $F_{xx}(x_0, \lambda; h) > 0$ para toda $h \in R_S(x_0)$, $h \neq 0$. Entonces existen una vecindad N de x_0 y $m > 0$ tales que*

$$f(x) \geq f(x_0) + m|x - x_0|^2 \quad \text{para toda } x \in S \cap N.$$

Demostración: Supongamos lo contrario. Entonces, para toda $q \in \mathbf{N}$ existe $x_q \in S$ tal que

$$t_q := |x_q - x_0| < \frac{1}{q}, \qquad f(x_q) < f(x_0) + \frac{t_q^2}{q}.$$

Claramente $t_q > 0$. Si es necesario, remplacemos $\{x_q\}$ por una subsucesión (denotada otra vez por $\{x_q\}$) tal que la sucesión de vectores unitarios

$$h_q := \frac{x_q - x_0}{t_q} = \frac{x_q - x_0}{|x_q - x_0|}$$

converge al vector unitario h. Entonces

$$0 = \lim_{q \to \infty} \frac{g_\alpha(x_q) - g_\alpha(x_0)}{t_q} = g'_\alpha(x_0; h)$$

y por lo tanto $h \in R_S(x_0)$, $h \neq 0$. Ahora, por Taylor,

$$\frac{1}{q} > \frac{F(x_q, \lambda) - F(x_0, \lambda)}{t_q^2} = \frac{1}{2} F_{xx}\left((\bar{x}_q, \lambda); \frac{x_q - x_0}{t_q}\right) = \frac{1}{2} F_{xx}((\bar{x}_q, \lambda); h_q)$$

donde $\bar{x}_q = x_0 + \theta_q[x_q - x_0]$ con $0 < \theta_q < 1$. Por lo tanto

$$0 \geq \lim_{q \to \infty} \frac{F(x_q, \lambda) - F(x_0, \lambda)}{t_q^2} = \frac{1}{2} F_{xx}((x_0, \lambda); h). \ \blacksquare$$

Como consecuencia del teorema de la función implícita tenemos el siguiente resultado.

Lema 2.4. *Supongamos que tenemos funciones $g_1, \ldots, g_m$ que mapean $\mathbf{R}^n$ en $\mathbf{R}$, de clase C^2 en una vecindad de un punto x_0 en $\mathbf{R}^n$, y tal que el conjunto $\{g'_\alpha(x_0) \mid \alpha = 1, \ldots, m\}$ es linealmente independiente. Entonces, para cada $h \in \mathbf{R}^n$, existen $\epsilon > 0$ y una función $x \colon [-\epsilon, \epsilon] \to \mathbf{R}^n$ de clase C^2 tal que $x(0) = x_0$, $\dot{x}(0) = h$ y, para toda $t \in [-\epsilon, \epsilon]$ y $\alpha \in \{1, \ldots, m\}$,*

$$g_\alpha(x(t)) = g_\alpha(x_0) + t g'_\alpha(x_0; h).$$

Demostración: Sea h_0 un vector arbitrario en $\mathbf{R}^n$ y, para toda $\alpha = 1, \dots, m$, denotemos por h_α el gradiente $g'_\alpha(x_0)$. Como los vectores $h_1, \dots, h_m$ son linealmente independientes,

$$|g'_\alpha(x_0; h_\beta)| = |\langle h_\alpha, h_\beta \rangle| \neq 0.$$

Sea H la matriz $n \times m$-dimensional dada por $(h_1, \dots, h_m)$ y definamos, para toda $(t, b) \in \mathbf{R} \times \mathbf{R}^m$ y $\alpha \in \{1, \dots, m\}$,

$$G_\alpha(t, b) = g_\alpha(x_0 + Hb + th_0) - g_\alpha(x_0) - tg'_\alpha(x_0; h_0),$$

$$G(t, b) = (G_1(t, b), \dots, G_m(t, b)).$$

Como $G(0, 0) = 0$ y $|G_b(0, 0)| = |g'_\alpha(x_0; h_\beta)| \neq 0$, se sigue del teorema de la función implícita que existen $\epsilon > 0$ y $b\colon [-\epsilon, \epsilon] \to \mathbf{R}^m$ de clase C^2 tales que, para toda $t \in [-\epsilon, \epsilon]$, $b(0) = 0$ y $G(t, b(t)) = 0$. Diferenciando la identidad anterior en $t = 0$ encontramos que $b'(0) = 0$. Por lo tanto, la función $x\colon [-\epsilon, \epsilon] \to \mathbf{R}^n$ definida por $x(t) = x_0 + Hb(t) + th_0$ satisface las condiciones requeridas. ∎

Si en el lema anterior elegimos h de tal manera que $g'(x_0, h) = 0$, se prueba que si x_0 es un punto normal de S, entonces es un punto regular de S.

2.2. Restricciones con desigualdades

Supongamos ahora que el conjunto de restricciones sobre el cual queremos minimizar a la función f está dado por

$$S = \{x \in \mathbf{R}^n \mid g_\alpha(x) \leq 0 \ (\alpha \in A),\ g_\beta(x) = 0 \ (\beta \in B)\}$$

donde $A = \{1, \dots, p\}$, $B = \{p + 1, \dots, m\}$, y se tiene dada una función $g = (g_1, \dots, g_m)$ que mapea $\mathbf{R}^n$ a $\mathbf{R}^m$.

Seguimos denotando por $P(S)$ al problema de minimización.

Para toda $x_0 \in S$ definamos

$$I(x_0) := \{\alpha \in A \mid g_\alpha(x_0) = 0\}$$

y consideremos el conjunto

$$S(x_0) := \{x \in \mathbf{R}^n \mid g_\alpha(x) = 0 \ (\alpha \in I(x_0)),\ g_\beta(x) = 0 \ (\beta \in B)\}$$

junto con su correspondiente conjunto de restricciones tangenciales en x_0, es decir,

$$R_{S(x_0)}(x_0) := \{h \in \mathbf{R}^n \mid g_i'(x_0; h) = 0 \ (i \in I(x_0) \cup B)\}.$$

Supongamos que x_0 es una solución local de P(S). Si $g_\alpha(x_0) < 0$, sea $\epsilon_\alpha > 0$ tal que $|x - x_0| < \epsilon_\alpha \Rightarrow g_\alpha(x) < 0$ y sea

$$N(x_0) := \{x \in \mathbf{R}^n : |x - x_0| < \epsilon\}$$

donde $\epsilon = \text{mín}\{\epsilon_\alpha \mid g_\alpha(x_0) < 0\}$.

Si $A = I(x_0)$, definimos $N(x_0) := \mathbf{R}^n$. Como $S(x_0) \cap N(x_0) \subset S$, x_0 también es una solución local de P$(S(x_0))$. Por el Teorema 2.2, si x_0 es un punto normal de $S(x_0)$ (las ecuaciones lineales $g_i'(x_0; h) = 0$ $(i \in I(x_0) \cup B)$ en h son linealmente independientes), entonces existe un único $\lambda \in \mathbf{R}^q$ (q denota la cardinalidad de $I(x_0) \cup B$) tal que $G_x(x_0, \lambda) = 0$, donde

$$G(x, \lambda) := f(x) + \sum_{i \in I(x_0) \cup B} \lambda_i g_i(x) \quad ((x, \lambda) \in \mathbf{R}^n \times \mathbf{R}^q).$$

Además, $\langle h, G_{xx}(x_0, \lambda)h \rangle \geq 0$ para toda $h \in R_{S(x_0)}(x_0)$. Se puede probar que, en este caso, $\lambda_\alpha \geq 0$ para toda $\alpha \in I(x_0)$ y por lo tanto, si

$$P(x_0) := \{\lambda \in \mathbf{R}^m \mid \lambda_\alpha \geq 0 \ (\alpha \in I(x_0)), \ \lambda_\alpha = 0 \ (\alpha \in A \setminus I(x_0))\}$$

y definimos F como

$$F(x, \lambda) := f(x) + \sum_1^m \lambda_\alpha g_\alpha(x),$$

entonces obtenemos el siguiente conjunto de condiciones necesarias de primer y segundo orden.

Teorema 2.5. *Supongamos que x_0 es una solución local de* P(S). *Si x_0 es un punto normal de $S(x_0)$ entonces existe un único $\lambda \in P(x_0)$ tal que $F_x(x_0, \lambda) = 0$. Más aún, $\langle h, F_{xx}(x_0, \lambda)h \rangle \geq 0$ para toda $h \in R_{S(x_0)}(x_0)$.*

Este resultado nos proporciona condiciones necesarias de segundo orden, sin embargo dichas condiciones se pueden mejorar considerablemente. Para hacerlo consideraremos no los vectores tangentes curvilíneos $C_S(x_0)$ sino el cono tangente $T_S(x_0)$ de S en x_0.

Definamos el conjunto de vectores que satisfacen las *restricciones tangenciales de* S *en* x_0 por

$$R_S(x_0) := \{h \in \mathbf{R}^n \mid g'_\alpha(x_0;h) \leq 0 \ (\alpha \in I(x_0)),\ g'_\beta(x_0;h) = 0 \ (\beta \in B)\}.$$

Como en el caso de igualdades, tenemos que $T_S(x_0) \subset R_S(x_0)$ para toda $x_0 \in S$, pero el converso no necesariamente es cierto.

Definición 2.6. Decimos que $x_0 \in S$ es un:
a. *punto regular de* S si $T_S(x_0) = R_S(x_0)$.
b. *punto normal de* S si las relaciones $\sum_1^m \lambda_i g'_i(x_0) = 0$, $(\lambda \in P(x_0))$ implican que $\lambda = 0$.

Como en el caso con igualdades, se tiene que normalidad implica regularidad.

Consideremos el conjunto

$$\mathcal{L} := \{x_0 \in S \mid \text{existe } \lambda \in P(x_0) \text{ tal que } F_x(x_0,\lambda) = 0\}$$

cuyos elementos se dice que satisfacen la *regla de multiplicadores de Lagrange* de primer orden. La siguiente proposición nos da condiciones necesarias y suficientes para que un punto satisfaga dichas condiciones.

Proposición 2.7. *Sea* $x_0 \in S$. *Entonces se satisfacen:*
a. *Si* $x_0 \in \mathcal{L}$ *entonces* $F(\cdot,\lambda)$ *es una función de soporte inferior para* f *en* (S, x_0) *y, por lo tanto,* $f'(x_0;h) \geq 0$ *para toda* $h \in T_S(x_0)$.
b. $x_0 \in \mathcal{L}$ *si y sólo si* $f'(x_0;h) \geq 0$ *para toda* $h \in R_S(x_0)$.
Demostración:
(a): La primera afirmación es clara y la segunda se sigue de la Observación 1.34 o directamente de (b) ya que $T_S(x_0) \subset R_S(x_0)$.
(b): Sea $x_0 \in \mathcal{L}$ y sea $\lambda \in P(x_0)$ tal que $F_x(x_0,\lambda) = 0$. Por lo tanto, para toda $h \in R_S(x_0)$,

$$0 = \langle F_x(x_0,\lambda), h\rangle = f'(x_0;h) + \sum_1^m \lambda_\alpha g'_\alpha(x_0;h) \leq f'(x_0;h).$$

Para el converso, si $\langle -f'(x_0);h\rangle \leq 0$ para toda $h \in R_S(x_0)$ entonces, por el Teorema 1.29, existe $\{\lambda_i\}_{i\in J} \subset \mathbf{R}$ donde $J = I(x_0) \cup B$ tal que $\lambda_\alpha \geq 0$ para toda $\alpha \in I(x_0)$ y $f'(x_0) + \sum_J \lambda_i g'_i(x_0) = 0$. El resultado se sigue definiendo $\lambda_\alpha = 0$ para toda $\alpha \in A \setminus I(x_0)$. ∎

Los siguientes resultados corresponden a condiciones suficientes para un mínimo local estricto de f en S.

Comenzaremos con un lema que caracteriza los elementos h de $R_S(x_0)$ para los cuales $f'(x_0; h) = 0$.

Lema 2.8. *Supongamos que* $x_0 \in \mathcal{L}$ *y* $\lambda \in P(x_0)$ *es tal que* $F_x(x_0, \lambda) = 0$. *Sea* $\Gamma := \{\alpha \in A \mid \lambda_\alpha > 0\}$. *Entonces, para toda* $h \in R_S(x_0)$, $h \neq 0$,

$$f'(x_0; h) = 0 \Leftrightarrow g'_\alpha(x_0; h) = 0 \quad (\alpha \in \Gamma).$$

Demostración: Sea $h \in R_S(x_0)$, $h \neq 0$. Entonces

$$0 = F_x(x_0, \lambda; h) = f'(x_0; h) + \sum_{\alpha \in \Gamma} \lambda_\alpha g'_\alpha(x_0; h).$$

Por lo tanto, si $g'_\alpha(x_0; h) = 0$ $(\alpha \in \Gamma)$, tenemos $f'(x_0; h) = 0$. Para el converso, si $f'(x_0; h) = 0$, se sigue de las relaciones $\lambda_\alpha > 0$ y $g'_\alpha(x_0; h) \leq 0$ que $g'_\alpha(x_0; h) = 0$ $(\alpha \in \Gamma)$. ∎

Teorema 2.9. *Si* $f'(x_0; h) > 0$ *para toda* $h \in T_S(x_0)$, $h \neq 0$, *entonces existen una vecindad* N *de* x_0 *y* $m > 0$ *tales que*

$$f(x) \geq f(x_0) + m|x - x_0| \quad (x \in S \cap N).$$

Demostración: Supongamos lo contrario. Existe entonces una sucesión $\{x_m\}$ de puntos de S tal que

$$|x_m - x_0| < \frac{1}{m} \quad \text{y} \quad f(x_m) - f(x_0) < \frac{1}{m}|x_m - x_0|.$$

Claramente $x_m \neq x_0$. Por lo tanto $\{x_m\}$ tiene una subsucesión (denotada otra vez por $\{x_m\}$) que converge a x_0 en la dirección h. Esto es una contradicción ya que

$$f'(x_0; h) = \lim \frac{f(x_m) - f(x_0)}{|x_m - x_0|} \leq 0. \ ∎$$

Como una consecuencia del teorema anterior tenemos:

Teorema 2.10. *Si* $x_0 \in \mathcal{L}$ *y* $\{h \in R_S(x_0) \mid f'(x_0; h) = 0\} = \{0\}$ *entonces* x_0 *es un mínimo local estricto de* f *en* S.

Demostración: Como $f'(x_0; h) > 0$ para toda $h \in T_S(x_0) \subset R_S(x_0)$, $h \neq 0$, el resultado se sigue del Teorema anterior. ∎

Teorema 2.11. *Sea $R \subset \mathbf{R}^n$ tal que $T_S(x_0) \subset R$ y sea*

$$\mathcal{A} := \{h \in R \setminus \{0\} \mid f'(x_0; h) = 0\}.$$

Supongamos que $\mathcal{A} \neq \emptyset$ y, para toda $h \in \mathcal{A}$, existe una función de soporte inferior F para f en (S, x_0) tal que $F''(x_0; h) > 0$. Entonces existen una vecindad N de x_0 y $m > 0$ tales que

$$f(x) \geq f(x_0) + m|x - x_0|^2 \quad (x \in S \cap N).$$

Demostración: Supongamos lo contrario. Entonces, para toda $m \in \mathbf{N}$ existe $x_m \in S$ tal que

$$|x_m - x_0| < \frac{1}{m} \quad \text{y} \quad f(x_m) - f(x_0) < \frac{1}{m}|x_m - x_0|^2.$$

Remplacemos $\{x_m\}$ por una subsucesión convergente a x_0 en una dirección h_0. Entonces

$$f'(x_0; h_0) = \lim \frac{f(x_m) - f(x_0)}{|x_m - x_0|} \leq \lim \frac{1}{m}|x_m - x_0| \leq \lim \frac{1}{m^2} = 0.$$

Por otra parte, $f'(x_0; h_0) \geq 0$ ya que $\mathcal{A} \neq \emptyset$ y, por hipótesis, existe una función de soporte inferior para f en (S, x_0). Por lo tanto, $f'(x_0; h_0) = 0$ y, como $h_0 \in T_S(x_0) \subset R$ es un vector unitario, $h_0 \in \mathcal{A}$. Sea F una función de soporte inferior para f en (S, x_0) que satisface $F''(x_0; h_0) > 0$ y definamos $G := f - F$. Tenemos

$$\frac{F(x_m) - F(x_0)}{|x_m - x_0|^2} + \frac{G(x_m)}{|x_m - x_0|^2} < \frac{1}{m}.$$

Sin embargo, $G(x_m) \geq 0$ para toda $m \in \mathbf{N}$ y $F'(x_0) = 0$. Esto implica que

$$\frac{1}{2}F''(x_0; h_0) = \lim \frac{F(x_m) - F(x_0)}{|x_m - x_0|^2} \leq 0$$

y llegamos a una contradicción. ∎

Teorema 2.12. *Sea $x_0 \in S$ y supongamos que existe $h \in R_S(x_0)$, $h \neq 0$ tal que $f'(x_0; h) = 0$. Si, para cada $h \neq 0$ de este tipo, existe $\lambda \in P(x_0)$ tal que $F_x(x_0, \lambda) = 0$ y $F_{xx}(x_0, \lambda; h) > 0$, entonces existen una vecindad N de x_0 y $m > 0$ tales que*

$$f(x) \geq f(x_0) + m|x - x_0|^2 \quad (x \in S \cap N).$$

Demostración: Como $F(\cdot, \lambda)$ es una función de soporte inferior para f en (S, x_0), el resultado se sigue del Teorema anterior. ∎

Teorema 2.13. *Supongamos que* $x_0 \in \mathcal{L}$, $\lambda \in P(x_0)$ *y* $F_x(x_0, \lambda) = 0$. *Sea* $\Gamma := \{\alpha \in A \mid \lambda_\alpha > 0\}$ *y consideremos el conjunto de restricciones tangenciales modificadas* $\tilde{R}_S(x_0; \lambda)$ *definido como*

$$\{h \in \mathbf{R}^n \mid g'_\alpha(x_0; h) \leq 0 \ (\alpha \in I(x_0),\ \lambda_\alpha = 0),\ g'_\beta(x_0; h) = 0 \ (\beta \in \Gamma \cup B)\}$$

el cual satisface

$$\begin{aligned} \tilde{R}_S(x_0; \lambda) &= \{h \in R_S(x_0) \mid g'_\alpha(x_0; h) = 0 \ (\alpha \in \Gamma)\} \\ &= \{h \in R_S(x_0) \mid f'(x_0; h) = 0\}. \end{aligned}$$

Supongamos que $F_{xx}(x, \lambda; h) > 0$ *para toda* $h \in \tilde{R}_S(x_0; \lambda)$, $h \neq 0$. *Entonces existen una vecindad* N *de* x_0 *y* $m > 0$ *tales que*

$$f(x) \geq f(x_0) + m|x - x_0|^2 \quad (x \in S \cap N).$$

Demostración: Las igualdades se satisfacen por definición de $R_S(x_0)$ y por el Lema 2.8. Ahora, si existe $h \in \tilde{R}_S(x_0; \lambda) \subset R_S(x_0)$, $h \neq 0$, tal que $f'(x_0; h) = 0$, la conclusión se sigue del teorema anterior. De otro modo, se tiene que $f'(x_0; h) > 0$ para toda $h \in \tilde{R}_S(x_0; \lambda)$, $h \neq 0$. En este caso existen, por el Teorema 2.9, una vecindad N de x_0 y $\tilde{m} > 0$ tales que $f(x) - f(x_0) \geq \tilde{m}|x - x_0|$ en $S \cap N$. Seleccionando $m > 0$ tal que $\tilde{m}|x - x_0| \geq m|x - x_0|^2$ en N, se obtiene el resultado. ∎

El siguiente resultado nos proporciona condiciones necesarias de optimalidad en términos del cono tangente, y lo usaremos para obtener condiciones necesarias de optimalidad en términos del conjunto de restricciones tangenciales modificadas.

Teorema 2.14. *Supongamos que* x_0 *es una solución local de* $\mathrm{P}(S)$. *Entonces* $f'(x_0; h) \geq 0$ *para toda* $h \in T_S(x_0)$, *y si* $f'(x_0) = 0$, *entonces* $f''(x_0; h) \geq 0$ *para toda* $h \in T_S(x_0)$.

Demostración: Sea $h \in T_S(x_0)$ un vector unitario y $\{x_m\} \subset S$ una sucesión convergente a x_0 en la dirección h. Para valores grandes de m tenemos $f(x_m) \geq f(x_0)$ y, por lo tanto,

$$0 \leq \lim \frac{f(x_m) - f(x_0)}{|x_m - x_0|} = f'(x_0; h).$$

Si además $f'(x_0) = 0$, entonces

$$0 \leq \lim \frac{f(x_m) - f(x_0)}{|x_m - x_0|^2} = \frac{1}{2} f''(x_0; h). \ ∎$$

Teorema 2.15. *Supongamos que x_0 es una solución de* P(S) *y* $x_0 \in \mathcal{L}$. *Sea* $\lambda \in P(x_0)$ *tal que* $F_x(x_0, \lambda) = 0$ *y sea* $\Gamma := \{\alpha \in A \mid \lambda_\alpha > 0\}$. *Consideremos el conjunto de restricciones modificadas*

$$\tilde{S}_\lambda := \{x \in \mathbf{R}^n \mid g_\alpha(x) \leq 0 \ (\alpha \in A,\ \ \lambda_\alpha = 0),\ \ g_\beta(x) = 0 \ (\beta \in \Gamma \cup B)\}$$

el cual satiface

$$\tilde{S}_\lambda = \{x \in S \mid g_\alpha(x) = 0 \ (\alpha \in \Gamma)\} = \{x \in S \mid F(x, \lambda) = f(x)\}.$$

Si x_0 es un punto regular de $\tilde{S}_\lambda$ entonces $F_{xx}(x_0, \lambda; h) \geq 0$ para toda $h \in \tilde{R}_S(x_0; \lambda)$.

Demostración: Claramente $\Gamma \subset I(x_0)$ y, para toda $x \in S$,

$$F(x, \lambda) - f(x) = \sum_{\alpha \in \Gamma} \lambda_\alpha g_\alpha(x).$$

Como $g_\alpha(x) \leq 0$ para toda $x \in S$ y $\alpha \in \Gamma$, se cumple la primera afirmación. Por lo tanto $\tilde{R}_S(x_0; \lambda)$ corresponde al conjunto de restricciones tangenciales de $\tilde{S}_\lambda$ en x_0. Si x_0 es un punto regular de $\tilde{S}_\lambda$, entonces $T_{\tilde{S}_\lambda}(x_0) = \tilde{R}_S(x_0; \lambda)$. Como $f(x) = F(x, \lambda)$ en $\tilde{S}_\lambda$, el punto x_0 minimiza $F(\cdot, \lambda)$ en $\tilde{S}_\lambda$. Como $F_x(x_0, \lambda) = 0$ tenemos, por el Teorema 2.14, $F_{xx}(x_0, \lambda; h) \geq 0$ para toda $h \in T_{\tilde{S}_\lambda}(x_0)$. ∎

Es importante resaltar el hecho de que el conjunto

$$R_{S(x_0)}(x_0) := \{h \in \mathbf{R}^n \mid g_i'(x_0; h) = 0 \ (i \in I(x_0) \cup B)\}$$

en el cual se basan las condiciones necesarias de segundo orden obtenidas en el Teorema 2.5, está contenido en el conjunto $\tilde{R}_S(x_0; \lambda)$ para toda $\lambda \in P(x_0)$ y la contención es generalmente propia.

Capítulo 3

Aumentabilidad

Continuamos con el problema de minimizar a una función $f\colon \mathbf{R}^n \to \mathbf{R}$ sobre un conjunto de restricciones al que denotamos S. En el capítulo anterior obtuvimos la regla de multiplicadores de Lagrange bajo una hipótesis de regularidad y aunque este procedimiento resulta útil para recordar la regla de multiplicadores de primer orden, la clase de problemas a los que puede aplicarse es limitado.

Introduciremos en este capítulo la función aumentada $H(x) = F(x) + \sigma G$, donde G es una función no negativa. Nuestro problema ahora consistirá en minimizar H sin restricciones y veremos en varios ejemplos que, aunque es un procedimiento alternativo al de regularidad, no es equivalente para la existencia de multiplicadores apropiados de Lagrange.

Comenzaremos con un resultado auxiliar relacionado con funciones de penalización, el cual será fundamental en la derivación de suficiencia para problemas con restricciones. Posteriormente estudiaremos el caso de restricciones con igualdades para después extender los resultados al caso de restricciones con desigualdades.

3.1. Funciones de penalización

Normalmente una solución de un problema de minimización con restricciones se puede obtener como un límite de soluciones de problemas apropiados de minimización sin restricciones. El procedimiento que veremos en esta sección se puede resumir de la siguiente manera. Supongamos que queremos minimizar una función F bajo ciertas restricciones. Definimos primero una función no negativa G tal que los puntos que satisfacen las restricciones están dados por las soluciones de $G(x) = 0$. Agregando un *término de penalización* σG a la función F, obtenemos el punto mínimo $x(\sigma)$ de la función aumentada $H(x,\sigma) = F(x)+\sigma G(x)$ de tal manera que el límite $x_0 = \lim_{\sigma\to\infty} x(\sigma)$ existe y es el mínimo de F sujeto a las restricciones dadas.

Ejemplo 3.1. Supongamos que queremos minimizar a la función $F(x,y) = x^2 - y^2 - 4y$ sujeta a las resricciones $g(x,y) = y = 0$. Siguiendo el procedimiento que describimos arriba, consideramos a la función

$$G(x,y) = \frac{1}{2}[g(x,y)]^2 = \frac{1}{2}y^2$$

y a la función F le sumamos el término de penalización σG para obtener

$$H(x,y;\sigma) = F(x,y) + \sigma G(x,y) = x^2 - 4y + \left(\frac{\sigma-2}{2}\right)y^2.$$

Así, para cada $\sigma > 2$, el punto $[x(\sigma), y(\sigma)] = [0{,}4/(\sigma-2)]$ es el mínimo de H, y cuando σ tiende a infinito ese mínimo de H converge al mínimo de F, que es $(0,0)$, sujeto a la restricción $g = 0$.

Teorema 3.2. *Sean F y G funcionales continuos en un conjunto compacto $N \subset \mathbf{R}^n$ con $G(x) \geq 0$ para toda $x \in N$. Definimos*

$$H(x,\sigma) = F(x) + \sigma G(x) \qquad (x \in N,\ \sigma \in \mathbf{R})$$

y sea $x(\sigma)$ un mínimo de $H(\cdot,\sigma)$ sobre N. Supongamos que existe un único punto x_0 que minimiza a F sobre el conjunto

$$S = \{x \in N \mid G(x) = 0\}.$$

Entonces $x(\sigma) \to x_0,\ \sigma \to \infty$.

Demostración: Sea $\{\sigma_q\} \subset \mathbf{R}$ una sucesión que tiende a infinito y sea $x_q = x(\sigma_q)$. Como $G(x_0) = 0$ y x_q minimiza a $H(\cdot,\sigma_q)$, tenemos

$$F(x_q) + \sigma_q G(x_q) = H(x_q,\sigma_q) \leq H(x_0,\sigma_q) = F(x_0)$$

y así

$$\limsup_{q\to\infty}\{F(x_q) + \sigma_q G(x_q)\} \leq F(x_0).$$

Como G es no negativa en N, se sigue que

$$\limsup_{q\to\infty} F(x_q) \leq F(x_0) \quad y \quad \lim_{q\to\infty} G(x_q) = 0.$$

Y de aquí, si $\bar{x}$ es cualquier límite de una subsucesión convergente de $\{x_q\}$, se tiene que

$$F(\bar{x}) \leq F(x_0) \quad y \quad G(\bar{x}) = 0.$$

Por la compacidad de N, $\bar{x}$ pertenece a N y, por tanto, a S. Por la unicidad de x_0, tenemos que $\bar{x} = x_0$. Así, cada subsucesión convergente de $\{x_q\}$ converge a x_0. Como $\{x_q\}$ es acotada, esto es posible sólo si

$$\lim_{q\to\infty} x_q = \lim_{q\to\infty} x(\sigma_q) = x_0.$$

Y esto implica que $x(\sigma) \to x_0, \quad \sigma \to \infty$. ∎

Lema 3.3. *Sea $\mathcal{C}$ un cono cerrado y sean $P(x)$ y $Q(x)$ dos formas cuadráticas con la propiedad de que $P(x) > 0$ para toda $x \neq 0$ en $\mathcal{C}$ tal que $Q(x) = 0$. Supongamos que $Q(x) \geq 0$ para toda x en $\mathcal{C}$. Entonces existe $\sigma_0 > 0$ tal que $P(x) + \sigma Q(x) > 0$ para toda $x \neq 0$ en $\mathcal{C}$ y toda $\sigma \geq \sigma_0$.*

Demostración: Sea $N = \{x \in \mathcal{C} : |x| \leq 1\}$. Definamos

$$H(x,\sigma) = P(x) + \sigma Q(x) \qquad (x \in \mathcal{C},\ \sigma \in \mathbf{R})$$

y sea $x(\sigma)$ un mínimo de $H(\cdot,\sigma)$ en N. Para toda $\sigma \in \mathbf{R}$ tenemos que

$$H(x(\sigma),\sigma) \leq H(0,\sigma) = 0.$$

Supongamos que $b := |x(\sigma)| \neq 0$. Entonces $x(\sigma)/b$ es un vector unitario en N y

$$H(x(\sigma),\sigma) \leq H\left(\frac{x(\sigma)}{b},\sigma\right) = \frac{H(x(\sigma),\sigma)}{b^2} \leq 0$$

con $0 < b \leq 1$. Se sigue entonces que si $H(x(\sigma),\sigma) < 0$ entonces $b = 1$. Si $H(x(\sigma),\sigma) = 0$ y $b < 1$, podemos remplazar $x(\sigma)$ por $x(\sigma)/b$. Así, $x(\sigma)$ es un vector unitario, o bien, $x(\sigma) = 0$. Por el teorema 3.2, tenemos que $x(\sigma) \to 0$ cuando $\sigma \to \infty$. Debido a que $|x(\sigma)| = 1$, a menos que $x(\sigma) = 0$, esto es posible solamente si $x(\sigma) = 0$ para valores grandes de σ, y existe una constante $\sigma_1 > 0$ tal que $x(\sigma) = 0$ para toda $\sigma \geq \sigma_1$. Ahora, sea $\sigma_0 > \sigma_1$ y tomemos $\sigma \geq \sigma_0$ y $x \neq 0$ en N. Si $Q(x) = 0$ entonces $H(x,\sigma) = P(x) > 0$ y, si $Q(x) > 0$, entonces

$$H(x,\sigma) > H(x,\sigma_1) \geq H(x(\sigma_1),\sigma_1) = 0.$$

Así, para toda $x \neq 0$ en $\mathcal{C}$ y $\sigma \geq \sigma_0$, se tiene

$$H(x,\sigma) = |x|^2 H\left(\frac{x}{|x|},\sigma\right) > 0$$

y se tiene el resultado. ▮

3.2. Restricciones con igualdades

Sea $A = \{1,\ldots,m\}$ con $m < n$ y supongamos dadas funciones $f,\ g_\alpha$ que mapean $\mathbf{R}^n$ en $\mathbf{R}$ con $\alpha \in A$ y

$$S = \{x \in \mathbf{R}^n \mid g_\alpha(x) = 0 \quad (\alpha \in A)\}.$$

Supongamos que f, g_α son de clase C^2 sobre S.

Consideremos el problema $P(S)$ que consiste en minimizar a la función f sobre el conjunto S y definamos el conjunto de *restricciones tangenciales* análogo al definido en la sección *2.1.*

- Para todo $x_0 \in S$ definimos el conjunto de *restricciones tangenciales* por
$$R_S(x_0) := \{h \in \mathbf{R}^n \mid g'_\alpha(x_0;h) = 0 \quad (\alpha \in A)\}.$$

- Se dice que un punto x_0 satisface la *regla de multiplicadores de Lagrange* si pertenece a
$$\mathcal{L}(\lambda) := \{x_0 \in S \mid F'(x_0) = 0 \ y \ F''(x_0;h) \geq 0 \ \forall\ h \in R_S(x_0)\}$$

- Y que satisface la *regla de multiplicadores de Lagrange reforzada* si pertenece a
$$\mathcal{L}'(\lambda) := \{x_0 \in S \mid F'(x_0) = 0 \ y \ F''(x_0;h) > 0 \ \forall\ h \in R_S(x_0),\ h \neq 0\}$$

donde $F(x) = f(x) + \langle\lambda, g(x)\rangle$.

Dado $(\lambda, \sigma) \in \mathbf{R}^m \times \mathbf{R}$, definimos

$$H(x) = f(x) + \langle\lambda, g(x)\rangle + \sigma G(x), \quad G(x) = \frac{1}{2}\sum_1^m g_\alpha(x)^2$$

y consideremos el conjunto

$$\mathcal{A}(\lambda, \sigma) := \{x_0 \in S \mid x_0 \text{ es un mínimo local de } H\}.$$

Definición 3.4. Decimos que $P(S)$ es aumentable en x_0 si $x_0 \in \mathcal{A}(\lambda, \sigma)$ para algún $(\lambda, \sigma) \in \mathbf{R}^m \times \mathbf{R}$.

Como $H(x) = f(x)$ para toda $x \in S$, si $P(S)$ es aumentable en x_0, entonces x_0 es un mínimo local de f en S.

Teorema 3.5. *Para todo $(\lambda, \sigma) \in \mathbf{R}^m \times \mathbf{R}$, $\mathcal{A}(\lambda, \sigma) \subset \mathcal{L}(\lambda)$. Es decir, si $P(S)$ es aumentable en x_0 entonces la regla de multiplicadores de Lagrange se cumple en x_0. Además, x_0 minimiza localmente a $f(x)$ sobre S.*

Demostración: Sea $x_0 \in \mathcal{A}(\lambda, \sigma)$. Como x_0 minimiza localmente a H, entonces $H'(x_0) = 0$ y $H''(x_0) \geq 0$. Como $G(x_0) = 0$ y $G(x) \geq 0$ para toda $x \in \mathbf{R}^n$, entonces x_0 minimiza a G, así que $G'(x_0) = 0$. Tenemos entonces

$$0 = H'(x_0) = F'(x_0) + \sigma G'(x_0) = F'(x_0) \quad y \quad 0 \leq H''(x_0) = F''(x_0) + \sigma G''(x_0)$$

Así que $F''(x_0; h) \geq 0$ cuando $G''(x_0; h) = 0$. Como $G''(x_0; h) = \sum_1^m g'_\alpha(x_0; h)^2$, tenemos $F''(x_0; h) \geq 0$ cuando $g'_\alpha(x_0; h) = 0 \quad (\alpha \in A)$, así que $x_0 \in \mathcal{L}(\lambda)$ ∎

Teorema 3.6. *Supongamos que $x_0 \in \mathcal{L}'(\lambda)$ para algún $\lambda \in \mathbf{R}^m$. Entonces existen σ_0, $k > 0$ y una vecindad N de x_0 tales que si $x \in N$ y $\sigma \geq \sigma_0$ se tiene*

$$H(x) \geq H(x_0) + k|x - x_0|^2$$

En particular, $P(S)$ es aumentable en x_0, y para toda $x \in N \cap S$ se tiene que

$$f(x) \geq f(x_0) + k|x - x_0|^2.$$

Demostración: Como $x_0 \in \mathcal{L}'(\lambda)$ tenemos que $F'(x_0) = 0$ y $F''(x_0; h) > 0$ si $h \neq 0$ y $G''(x_0; h) = 0$. Como x_0 minimiza a G, $G''(x_0) \geq 0$. Por lo tanto, haciendo $P = F''(x_0)$ y $Q = G''(x_0)$ tenemos que se cumplen las hipótesis del lema 3.3. Así que existe $\sigma_0 > 0$ tal que, para toda $h \neq 0$,

$$F''(x_0; h) + \sigma_0 G''(x_0; h) > 0.$$

En consecuencia, $H'(x_0) = 0$ y para toda $h \neq 0$, $H''(x_0; h) > 0$. Por el teorema 1.8, existe una constante $k > 0$ y una vecindad N de x_0 tal que, para toda $x \in N$,

$$H(x) \geq H(x_0) + k|x - x_0|^2.$$

Como $G(x) \geq 0$, esta desigualdad también es válida si reemplazamos σ_0 por $\sigma \geq \sigma_0$. Y de aquí, cuando $G(x) = 0$ se tiene que $f(x) \geq f(x_0) + k|x - x_0|^2$. ∎

3.3. Restricciones con desigualdades

Supongamos ahora que tenemos funciones f, g_α, $(\alpha = 1, \ldots, m)$ que mapean $\mathbf{R}^n$ en $\mathbf{R}$ y nuestro problema, al que seguimos llamando P(S), ahora consistirá en minimizar a la función f sobre el conjunto

$$S = \{x \in \mathbf{R}^n \mid g_\alpha(x) \leq 0, \ (\alpha \in A); \ g_\beta(x) = 0, \ (\beta \in B)\}.$$

donde $A = \{1, \ldots, p\}$ y $B = \{p+1, \ldots, m\}$. Suponemos que f y g son de clase C^2 sobre S.

Como en el caso con igualdades, tenemos una función Lagrangiana (con respecto a $\lambda \in \mathbf{R}^m$) asociada a nuestro problema dada por

$$F(x) = f(x) + \langle \lambda, g(x) \rangle \quad (x \in \mathbf{R}^n).$$

Consideremos también al conjunto de multiplicadores de Lagrange en el punto $x \in S$ dado por

$$P(x) = \{\lambda \in \mathbf{R}^m \mid \lambda_\alpha \geq 0 \ (\alpha \in A), \lambda_\alpha = 0 \ si \ g_\alpha(x) < 0\},$$

y al conjunto de restricciones tangenciales en $x \in S$:

$$R_S(x) = \{h \in \mathbf{R}^n \mid g'_\alpha(x; h) \leq 0 \quad (\alpha \in A, \ g_\alpha(x) = 0), \ g'_\beta(x; h) = 0 \ (\beta \in B)\}.$$

Al conjunto de *restricciones tangenciales modificadas* lo denotamos por $\tilde{R}_S(x,\lambda)$ y está dado por

$$\{h \in R_S(x) \mid g'_\alpha(x;h) = 0 \ (\alpha \in A,\ \lambda_\alpha > 0)\}.$$

Además, decimos que un punto satisface:

- la regla de multiplicadores de Lagrange de primer orden si pertenece a

$$M(\lambda) = \{x_0 \in S \mid \lambda \in P(x_0) \ y \ F'(x_0) = 0\};$$

- la regla de multiplicadores de Lagrange si pertenece a

$$\mathcal{L}(\lambda) = \{x_0 \in M(\lambda) \mid F''(x_0;h) \geq 0 \ \forall \ h \in \tilde{R}_S(x_0,\lambda)\};$$

- la regla de multiplicadores de Lagrange restringida si pertenece a

$$\mathcal{L}'(\lambda) = \{x_0 \in M(\lambda) \mid F''(x_0;h) > 0 \ \forall \ h \neq 0 \ en \ \tilde{R}_S(x_0,\lambda)\}.$$

Como en la sección anterior, lo que queremos hacer es simplificar el problema quitando restricciones. Para esto, estudiaremos primeramente el problema de minimizar una función $H\colon \mathbf{R}^n \times \mathbf{R}^p \to \mathbf{R}$ sobre el conjunto

$$K = \{(x,y) \in \mathbf{R}^n \times \mathbf{R}^p \mid y^\alpha \leq 0 \ (\alpha \in A)\}$$

donde $A = \{1,\ldots,p\}$, $y = (y^1,\ldots,y^p)$.

Supongamos que H es de clase C^2 y para todo $(x,y) \in K$ definamos el cono

$$C_H(x,y) = \{(h,k) \in \mathbf{R}^n \times \mathbf{R}^p \mid k^\alpha \leq 0 \text{ si } y^\alpha = 0 \text{ y } k^\alpha = 0 \text{ si } H_{y^\alpha}(x,y) < 0\}$$

para enunciar el siguiente teorema que nos da condiciones necesarias para un mínimo local de H.

Teorema 3.7. *Si (x_0,y_0) es un mínimo local de H sobre K, entonces*

1. $H_x(x_0,y_0) = 0,\ H_{y^\alpha}(x_0,y_0) \leq 0$ *con igualdad si* $y_0^\alpha < 0$.

2. $H''((x_0,y_0);(h,k)) \geq 0$ *si* $(h,k) \in C_H(x_0,y_0)$.

Demostración: Para las condiciones en (1) basta verificar cada componente por separado. Para probar (2), sean $(h,k) \in C_H(x_0,y_0)$ y $\delta > 0$ tales que

$$g(t) := (x_0 + th, y_0 + tk) \in K \quad (0 \leq t \leq \delta).$$

Así, $w := H \circ g$ tiene un mínimo local en $t = 0$. Como $k^\alpha = 0$ si $H_{y^\alpha}(x_0,y_0) < 0$, de (1) tenemos que

$$w'(0) = H'((x_0,y_0);(h,k)) = 0.$$

De modo que $H''((x_0,y_0);(h,k)) = w''(0) \geq 0$. ∎

Teorema 3.8. *Sea $(x_0,y_0) \in K$ y supongamos que*

1. $H_x(x_0,y_0) = 0$, $H_{y^\alpha}(x_0,y_0) \leq 0$ *con igualdad si* $y_0^\alpha < 0$.
2. $H''((x_0,y_0);(h,k)) > 0$ *para todo* $(h,k) \neq (0,0)$ *en* $C_H(x_0,y_0)$.

Existe entonces una vecindad N de (x_0,y_0) y $k > 0$ tal que, para todo $(h,k) \in N \cap K$ se tiene que

$$H(x,y) \geq H(x_0,y_0) + k(|x-x_0|^2 + |y-y_0|^2).$$

Demostración: Supongamos lo contrario. Entonces, para toda $q \in \mathbf{N}$, existe $(x_q,y_q) \in K$ tal que

$$t_q := (|x-x_0|^2 + |y-y_0|^2)^{\frac{1}{2}} < \frac{1}{q},$$

$$H(x_q,y_q) - H(x_0,y_0) < \frac{t_q^2}{q}.$$

Así, $(x_q,y_q) \neq (x_0,y_0)$, y $t_q > 0$ para toda $q \in \mathbf{N}$. Remplazamos la sucesión original por una subsucesión, denotada también por $\{(x_q,y_q)\}$, tal que los vectores unitarios (h_q,k_q) definidos por la relación

$$h_q = \frac{x_q - x_0}{t_q}, \qquad k_q = \frac{y_q - y_0}{t_q}$$

converjan al vector (h,k) que también es unitario. Sea

$$I = \{\alpha \in A \mid H_{y^\alpha}(x_0,y_0) < 0\}.$$

Observemos que $k_q^\alpha = \frac{y_q^\alpha}{t_q} \leq 0$ si $y_0^\alpha = 0$. Así que

$$H'((x_0, y_0); (h_q, k_q)) = \sum_{\alpha \in I} H_{y^\alpha}(x_0, y_0) k_q^\alpha \geq 0.$$

Además

$$t_q H'((x_0, y_0); (h_q, k_q)) + \frac{t_q^2}{2} H''((x_0, y_0); (h_q, k_q)) + R_q < \frac{t_q^2}{q}$$

donde R_q es tal que $R_q / t_q^2 \to 0,\ q \to \infty$. Así,

$$\limsup_{q \to \infty} \frac{H'((x_0, y_0); (h_q, k_q))}{t_q} + \frac{1}{2} \lim_{q \to \infty} H''((x_0, y_0); (h_q, k_q) \leq 0.$$

El primer término es no negativo, así que

$$H''((x_0, y_0); (h, k)) = \lim_{q \to \infty} H''((x_0, y_0); (h_q, k_q)) \leq 0.$$

Además, como $t_q \to 0,\ q \to \infty$, tenemos que

$$0 = \lim_{q \to \infty} H'((x_0, y_0); (h_q, k_q)) = \sum_{\alpha \in I} H_{y^\alpha}(x_0, y_0) k^\alpha.$$

Como $H_{y^\alpha}(x_0, y_0) < 0$ y $k^\alpha \leq 0$ para $\alpha \in I$, se tiene que $k^\alpha = 0$ $(\alpha \in I)$ y así (h, k) es un vector unitario en $C_H(x_0, y_0)$ que satisface $H''((x_0, y_0); (h_q, k_q)) \leq 0$. ∎

Consideremos ahora el conjunto

$$\hat{S} = \{(x, y) \in \mathbf{R}^n \times \mathbf{R}^p \mid y^\alpha \leq 0,\ g_\alpha(x) - y^\alpha = 0\ (\alpha \in A),\ g_\beta(x) = 0\ (\beta \in B)\}$$

y definamos la función Lagrangiana aumentada con respecto a f, g y $\lambda \in \mathbf{R}^m$ por

$$\hat{F}(x, y) = f(x) + \langle \lambda, g(x) \rangle - \sum_1^p \lambda_\alpha y^\alpha.$$

Para todo $(\lambda, \sigma) \in \mathbf{R}^m \times \mathbf{R}$, sea

$$H(x, y) = f(x) + \sum_1^p \lambda_\alpha (g_\alpha(x) - y^\alpha) + \sum_{p+1}^m \lambda_\beta g_\beta(x) + \sigma G(x, y)$$

donde

$$G(x,y) = \frac{1}{2}\left(\sum_{1}^{p}\{g_\alpha(x) - y^\alpha\}^2 + \sum_{p+1}^{m} g_\beta(x)^2\right).$$

de modo que también podemos expresar a $\hat{S}$ de la forma

$$\hat{S} = \{(x,y) : G(x,y) = 0,\ y^\alpha \leq 0\ (\alpha \in A)\}.$$

Observemos que

$$H(x,y) = \hat{F}(x,y) + \sigma G(x,y)$$

y que $H(x,y) = \hat{F}(x,y) = f(x)$ para toda $(x,y) \in \hat{S}$. Por lo tanto, (x_0, y_0) con $y_0^\alpha = g_\alpha(x_0)$ es solución local del problema de minimizar H en $\hat{S}$ si y sólo si (x_0, y_0) es solución local del problema de minimizar $\hat{F}$ en $\hat{S}$ si y sólo si x_0 es una solución local del problema original P(S).

Definición 3.9. Dado $x_0 \in S$, sea $y_0^\alpha = g_\alpha(x_0)$ para todo $\alpha \in A$. Decimos que el problema P(S) es *aumentable* en x_0 si existe $(\lambda, \sigma) \in \mathbf{R}^m \times \mathbf{R}$ tal que (x_0, y_0) es un mínimo local de H sobre el conjunto

$$K = \{(x,y) \in \mathbf{R} \times \mathbf{R}^p \mid y^\alpha \leq 0\ (\alpha \in A)\}.$$

Notemos que, por la observación anterior, como $\hat{S} \subset K$, si P(S) es aumentable en x_0, entonces x_0 es un mínimo local de f sobre S.

Consideremos el cono

$$\begin{aligned} C_{\hat{F}}(x,y) \ = \ & \{(h,k) \in \mathbf{R}^n \times \mathbf{R}^p \mid k^\alpha \leq 0 \ si\ y^\alpha = 0,\ y \\ & k^\alpha = 0 \ si\ \hat{F}_{y^\alpha}(x,y) = -\lambda_\alpha < 0\}. \end{aligned}$$

Teorema 3.10. *Sea $(x_0, y_0) \in \hat{S}$ y supongamos que, para algún $\sigma \in \mathbf{R}$, (x_0, y_0) es un mínimo local de H sobre K. Entonces (x_0, y_0) es un mínimo local de $\hat{F}$ sobre $\hat{S}$ y, además, cumple lo siguiente:*

1. *$\hat{F}_x(x_0, y_0) = 0$, $\hat{F}_{y^\alpha}(x_0, y_0) \leq 0$ con igualdad si $y_0^\alpha < 0$.*

2. *$\hat{F}''((x_0, y_0); (h,k)) \geq 0$ para todo $(h,k) \in C_{\hat{F}}(x_0, y_0)$ que satisface $G''((x_0, y_0); (h,k)) = 0$.*

Demostración: La primera conclusión se sigue de la definición de H. Luego, como (x_0, y_0) minimiza a G, se tiene que $G_x(x_0, y_0) = G_y(x_0, y_0) = 0$ y por (1) del Teorema 3.7 tenemos (1). Además, esto implica que

$$C_{\hat{F}}(x_0, y_0) = C_H(x_0, y_0)$$

y por (2) del mismo teorema tenemos la última conclusión. ∎

Teorema 3.11. *Sea $(x_0, y_0) \in \hat{S}$ y supongamos lo siguiente:*

1. $\hat{F}_x(x_0, y_0) = 0, \hat{F}_{y^\alpha}(x_0, y_0) \leq 0$ *con igualdad si* $y_0^\alpha < 0$.

2. $\hat{F}''((x_0, y_0); (h, k)) > 0$ *para todo* $(h, k) \neq (0, 0)$ *en* $C_{\hat{F}}(x_0, y_0)$ *tal que* $G''((x_0, y_0); (h, k)) = 0$.

Entonces existen $\sigma_0, k > 0$ y una vecindad N de (x_0, y_0) tales que, para todo $(x, y) \in N \cap K$ y toda $\sigma \geq \sigma_0$, se tiene que

$$H(x, y) \geq H(x_0, y_0) + k(|x - x_0|^2 + |y - y_0|^2).$$

En particular, para todo $(x, y) \in N \cap \hat{S}$ se tiene

$$\hat{F}(x, y) \geq \hat{F}(x_0, y_0) + k(|x - x_0|^2 + |y - y_0|^2).$$

Demostración: Sean $P = \hat{F}''(x_0, y_0)$ y $Q = G''(x_0, y_0)$. Como (x_0, y_0) minimiza a G, la forma cuadrática Q es no negativa. Aplicando el lema 3.3 tenemos que existe $\sigma_0 > 0$ tal que, para todo $(h, k) \neq (0, 0)$ en $C_{\hat{F}}(x_0, y_0)$,

$$\hat{F}''((x_0, y_0); (h, k)) + \sigma_0 G''((x_0, y_0); (h, k)) > 0.$$

Como $C_{\hat{F}}(x_0, y_0) = C_H(x_0, y_0)$, H satisface las condiciones del teorema 3.8, así que existe una vecindad N de (x_0, y_0) y $k > 0$ tal que, para todo $(x, y) \in N \cap K$,

$$H(x, y) \geq H(x_0, y_0) + k(|x - x_0|^2 + |y - y_0|^2).$$

Finalmente, como $G(x, y) \geq 0$, esta desigualdad también es válida para todo $\sigma \geq \sigma_0$. ∎

Definamos para $\lambda \in \mathbf{R}^m$,

$$\begin{aligned} N(\lambda) \;=\; & \{(x_0, y_0) \in \mathbf{R}^n \times \mathbf{R}^p \mid \hat{F}_x(x_0, y_0) = 0,\ \hat{F}_{y^\alpha}(x_0, y_0) = -\lambda_\alpha \leq 0, \\ & \text{con igualdad si } y_0^\alpha < 0\}, \end{aligned}$$

y

$$
\begin{aligned}
\mathcal{A}(\lambda) &= \{(x_0, y_0) \in N(\lambda) \mid \hat{F}''((x_0, y_0); (h, k)) \geq 0 \ \forall (h, k) \in C_{\hat{F}}(x_0, y_0) \\
&\qquad \text{con } G''((x_0, y_0); (h, k)) = 0\}
\end{aligned}
$$

y $\mathcal{A}'(\lambda)$ como antes, pero con la desigualdad estricta para vectores (h, k) distintos de cero.

El siguiente resultado muestra la relación entre estos conjuntos y la regla de multiplicadores de Lagrange.

Lema 3.12. *Sean $x_0 \in S$, $y_0^\alpha = g_\alpha(x_0)$ para toda $\alpha \in A$. Entonces, para toda $\lambda \in \mathbf{R}^m$, $x_0 \in \mathcal{L}(\lambda) \Leftrightarrow (x_0, y_0) \in \mathcal{A}(\lambda)$. Se tiene el mismo resultado remplazando $\mathcal{L}$ y $\mathcal{A}$ por $\mathcal{L}'$ y $\mathcal{A}'$ respectivamente.*

Demostración: Notemos primeramente que $(x_0, y_0) \in N(\lambda) \Leftrightarrow F'(x_0) = 0$ y $\lambda_\alpha \geq 0$ $(\alpha \in A)$ con $\lambda_\alpha = 0$ si $g_\alpha(x_0) < 0 \quad \Leftrightarrow \quad x_0 \in M(\lambda)$. Ahora, $G''((x_0, y_0); (h, k))$ corresponde a

$$\sum_1^p (g'_\alpha(x_0; h) - k^\alpha)^2 + \sum_{p+1}^m g'_\beta(x_0; h)^2$$

y así la condición $G''((x_0, y_0); (h, k)) = 0$ se cumple si y sólo si

$$k^\alpha = g'_\alpha(x_0; h) \ (\alpha \in A), \quad g'_\beta(x_0; h) = 0 \ (\beta \in B).$$

De modo que si $(h, k) \in C_{\hat{F}}(x_0, y_0)$ con $G''((x_0, y_0); (h, k)) = 0$, entonces $h \in \tilde{R}_S(x_0, \lambda)$. Recíprocamente, si $h \in \tilde{R}_S(x_0, \lambda)$ y hacemos $k^\alpha = g'_\alpha(x_0; h)$, entonces $(h, k) \in C_{\hat{F}}(x_0, y_0)$ con $G''((x_0, y_0); (h, k)) = 0$. Como $\hat{F}''((x_0, y_0); (h, k)) = F''(x_0; h)$, el resultado se sigue. ∎

El siguiente resultado establece que la regla de multiplicadores de Lagrange es consecuencia de la aumentabilidad. Y el siguiente que la aumentabilidad es consecuencia de la regla de multiplicadores de Lagrange restringida.

Teorema 3.13. *Sea $x_0 \in S$ y supongamos que P(S) es aumentable en x_0. Entonces la regla de multiplicadores de Lagrange se cumple en x_0. Más aun, x_0 es un mínimo local de f sobre S.*

Demostración: Sea $y_0 := (g_1(x_0), \ldots, g_p(x_0))$. Como P(S) es aumentable en x_0, existe $(\lambda, \sigma) \in \mathbf{R}^m \times \mathbf{R}$ tal que (x_0, y_0) es un mínimo local de H sobre K. Por el teorema 3.10, $(x_0, y_0) \in \mathcal{A}(\lambda)$ y por el lema 3.12, $x_0 \in \mathcal{L}(\lambda)$. ∎

Teorema 3.14. *Sea $x_0 \in S$ y supongamos que la regla de multiplicadores de Lagrange restringida se cumple en x_0 para alguna λ. Sea*

$$y_0 = (g_1(x_0), \ldots, g_p(x_0)).$$

Existe entonces $\sigma_0, k > 0$ y una vecindad N de x_0 tal que, si $\sigma \geq \sigma_0$, $x \in N$ y $(x, y) \in K$, entonces

$$H(x, y) \geq H(x_0, y_0) + k|x - x_0|^2.$$

En particular, P(S) es aumentable en x_0 y, para toda $x \in N \cap S$ se cumple que

$$f(x) \geq f(x_0) + k|x - x_0|^2.$$

Demostración: Supongamos que $x_0 \in \mathcal{L}'(\lambda)$. Por el lema 3.12, $(x_0, y_0) \in \mathcal{A}'(\lambda)$. Por el teorema 3.11, existen $\sigma_0,\ k > 0$ y una vecindad $\hat{N}$ de (x_0, y_0) tal que si

$$H((x, y), \sigma) = \hat{F}(x, y) + \sigma G(x, y)$$

entonces, para todo $(x, y) \in \hat{N} \cap K$ y todo $\sigma \geq \sigma_0$,

$$H((x, y), \sigma) \geq H((x_0, y_0), \sigma) + k(|x - x_0|^2 + |y - y_0|^2).$$

Ahora, tomemos r y $\delta > 0$ tales que el conjunto

$$\hat{B} = \{(x, y) \in \mathbf{R}^n \times \mathbf{R}^p \mid x \in \bar{B}(x_0; r),\ |g_\alpha(x) - y^\alpha| \leq \sqrt{2\delta}\ (\alpha \in A)\}$$

está contenido en $\hat{N}$. Sea $R = \min\{F(x) \mid x \in \bar{B}(x_0; r)\}$ y tomemos σ_0 de tal manera que

$$R + \sigma_0 \delta \geq f(x_0) + kr^2.$$

Sea (x, y) con $x \in \bar{B}(x_0; r)$ y $y^\alpha \leq 0\ (\alpha \in A), \sigma \geq \sigma_0$. Si $(x, y) \in \hat{B}$, entonces $(x, y) \in \hat{N} \cap K$ y así

$$H((x, y), \sigma) \geq H((x_0, y_0), \sigma) + k|x - x_0|^2.$$

Si $(x, y) \notin \hat{B}$, entonces $G(x, y) \geq \delta$, $y^\alpha \leq 0\ (\alpha \in A)$, y

$$H((x, y), \sigma) \geq R + \sigma_0 \delta \geq f(x_0) + kr^2.$$

Como $|x - x_0| \leq r$, también en este caso se tiene

$$H((x, y), \sigma) \geq H((x_0, y_0), \sigma) + k|x - x_0|^2. \ ∎$$

3.4. Ejemplos

Consideremos el problema (P) que consiste en minimizar $f : \mathbf{R}^2 \to \mathbf{R}$ sobre el conjunto
$$S = \{(x, y) \in \mathbf{R}^2 \mid g(x, y) = 0\}.$$

1. (P) es regular y aumentable en $(0, 0)$. Sea
$$f(x, y) = x^2 - y^2 - 4y, \quad g(x, y) = y.$$
Como $g'(0, 0) \neq (0, 0)$, (P) es regular en $(0, 0)$. Es aumentable en $(0, 0)$ ya que
$$H(x, y) = x^2 + \left(\frac{\sigma - 2}{2}\right) y^2 + (\lambda - 4)y$$
tiene un mínimo estricto en $(0, 0)$ si $\lambda = 4$ y $\sigma > 2$.

2. (P) no es regular ni aumentable en $(0, 0)$, pero este punto minimiza f sobre S. Sea
$$f(x, y) = x^2 - y^2 - 4y, \quad g(x, y) = y^2.$$
Este problema no es regular en $(0, 0)$ ya que $T_S((0, 0))$ coincide con la recta $y = 0$, mientras que $R_S((0, 0))$ es $\mathbf{R}^2$. No es aumentable en $(0, 0)$ porque
$$H(x, y) = x^2 - 4y + (\lambda - 1)y^2 + \frac{\sigma}{2}y^4$$
no tiene un mínimo en $(0, 0)$ para ningún λ y σ.

3. (P) no es regular en $(0, 0)$, pero sí es aumentable en ese punto. Sea
$$f(x, y) = x^2 - y^4 - 4y^2, \quad g(x, y) = y^2.$$
(P) no es regular en $(0, 0)$ como se vio en (2). Sin embargo, es aumentable en $(0, 0)$ porque con $\lambda = 4$ y $\sigma > 2$, la función
$$H(x, y) = x^2 + \left(\frac{\sigma - 2}{2}\right) y^4 + (\lambda - 4)y^2$$
tiene un mínimo en $(0, 0)$.

4. (P) es regular en $(0,0)$, pero no es aumentable en ese punto. Sea

$$f(x,y) = x^2 + 2x + y^4, \quad g(x,y) = xy - x.$$

f tiene un mínimo local estricto en $(0,0)$ sobre S. Como $g'(0,0) \neq (0,0)$, es un punto normal y, por tanto, regular para (P), sin embargo, (P) no es aumentable. Consideremos

$$H(x,y) = x^2\left[1 + \frac{\sigma}{2}(y-1)^2\right] + (2-\lambda)x + \lambda xy + y^4.$$

Si $H'(0,0) = (0,0)$ se debe tener $\lambda = 2$. Así que,

$$H(x,y) = x^2\left[1 + \frac{\sigma}{2}(y-1)^2\right] + 2xy + y^4.$$

Podemos suponer que $\sigma \geq 0$. Si $(y-1)^2 \leq 4$, $y^2 < 1/(1+2\sigma)$ y $x = -y/(1+2\sigma)$, vemos que

$$H(x,y) \leq y^2\left(y^2 - \frac{1}{1+2\sigma}\right) < 0 = H(0,0).$$

Capítulo 4

Problemas de control óptimo que involucran restricciones con igualdades mixtas

Veremos primeramente condiciones de optimalidad para problemas sin restricciones.

4.1. Planteamiento del problema sin restricciones

Supongamos que tenemos un intervalo $T := [t_0, t_1]$ en $\mathbf{R}$, un conjunto abierto $\mathcal{A} \subset T \times \mathbf{R}^n \times \mathbf{R}^m$, dos puntos $\xi_0,\ \xi_1$ en $\mathbf{R}^n$, y funciones L, f que mapean $T \times \mathbf{R}^n \times \mathbf{R}^m$ en $\mathbf{R}$ y $\mathbf{R}^n$ respectivamente. Denotemos por X el espacio de funciones de clase C^1 por fragmentos que mapean T en $\mathbf{R}^n$, por $\mathcal{U}$ al espacio de funciones continuas por fragmentos que mapean T en $\mathbf{R}^m$, y sean $Z := X \times \mathcal{U}$,

$$D := \{(x,u) \in Z \mid \dot{x}(t) = f(t, x(t), u(t))\ (t \in T)\},$$

$$Z_e(\mathcal{A}) := \{(x,u) \in D \mid (t, x(t), u(t)) \in \mathcal{A}\ (t \in T),\ x(t_0) = \xi_0,\ x(t_1) = \xi_1\},$$

y

$$I(x,u) := \int_{t_0}^{t_1} L(t, x(t), u(t))dt.$$

Consideremos el problema (P_u) de minimizar I sobre $Z_e(\mathcal{A})$.

Los elementos de Z se llaman *procesos*, los de $Z_e(\mathcal{A})$ *procesos admisibles* y un proceso (x,u) resuelve (P_u) si (x,u) es admisible e $I(x,u) \leq I(y,v)$ para todo proceso admisible (y,v). Supongamos que f, L son de clase C^2 sobre $\mathcal{A}$ y denotemos por "*" la transpuesta.

- Para todo (t,x,u,p) en $T \times \mathbf{R}^n \times \mathbf{R}^m \times \mathbf{R}^n$ sea

$$H(t,x,u,p) := \langle p, f(t,x,u)\rangle - L(t,x,u).$$

- Para todo $(x,u,p) \in Z \times X$ y $(y,v) \in Z$ sea

$$J((x,u,p);(y,v)) := \int_{t_0}^{t_1} 2\Omega(t,y(t),v(t))dt$$

  donde, para todo $(t,y,v) \in T \times \mathbf{R}^n \times \mathbf{R}^m$,

$$2\Omega(t,y,v) := -[\langle y, H_{xx}(t)y\rangle + 2\langle y, H_{xu}(t)v\rangle + \langle v, H_{uu}(t)v\rangle]$$

  y $H(t)$ denota $H(t,x(t),u(t),p(t))$.

Definición 4.1. Un proceso (x,u) se dice que es *normal* si existe una solución no nula p en T del sistema

$$\dot{p}(t) = -f_x^*(t,x(t),u(t))p(t), \quad f_u^*(t,x(t),u(t))p(t) = 0.$$

El siguiente resultado es bien conocido y con él obtenemos condiciones necesarias de primer y segundo orden bajo suposiciones de normalidad.

Teorema 4.2. *Supongamos que (x_0,u_0) es una solución normal de (P_u). Existe entonces un único $p \in X$ tal que*

$$\dot{p}(t) = -H_x^*(t,x_0(t),u_0(t),p(t)), \quad H_u(t,x_0(t),u_0(t),p(t)) = 0 \quad (t \in T).$$

Más aún, $J((x_0,u_0,p);(y,v)) \geq 0$ para todo $(y,v) \in Z$ que satisface $y(t_0) = y(t_1) = 0$ y

$$\dot{y}(t) = f_x(t,x_0(t),u_0(t))y(t) + f_u(t,x_0(t),u_0(t))v(t) \quad (t \in T).$$

4.2. Planteamimento del problema con restricciones

Supongamos que los datos del problema son como antes excepto el conjunto $\mathcal{A}$ que ahora estará dado por

$$\mathcal{A} := \{(t, x, u) \in T \times \mathbf{R}^n \times \mathbf{R}^m \mid \varphi(t, x, u) = 0\},$$

y φ es una función que mapea $T \times \mathbf{R}^n \times \mathbf{R}^m$ en $\mathbf{R}^q$ $(q \leq m)$.

El problema del que ahora nos ocuparemos, al cual llamaremos (P_c), consiste en minimizar I sobre $Z_e(\mathcal{A})$.

Suponemos que φ es de clase C^2 sobre $\mathcal{A}$ y la matriz $\varphi_u(t, x, u)$ es de rango q sobre $\mathcal{A}$. Denotamos por $\mathcal{U}_q$ al espacio de funciones continuas por fragmentos que mapean T en $\mathbf{R}^q$.

- Para todo $(t, x, u, p, \mu) \in T \times \mathbf{R}^n \times \mathbf{R}^m \times \mathbf{R}^n \times \mathbf{R}^q$ sea

$$H(t, x, u, p, \mu) := \langle p, f(t, x, u)\rangle - L(t, x, u) - \langle \mu, \varphi(t, x, u)\rangle.$$

- Para todo $(x, u, p, \mu) \in Z \times X \times \mathcal{U}_q$ y $(y, v) \in Z$ sea

$$J((x, u, p, \mu); (y, v)) := \int_{t_0}^{t_1} 2\Omega(t, y(t), v(t))dt$$

donde $2\Omega(t, y, v)$ es como antes, pero $H(t)$ denota $H(t, x(t), u(t), p(t), \mu(t))$.

• Normalidad

Definición 4.3. Dado $(x, u) \in Z$ definimos $A(t) := f_x(t, x(t), u(t))$, $B(t) := f_u(t, x(t), u(t))$ $(t \in T)$. Decimos que un proceso (x, u) es *normal* (con respecto a (P_c)) si, dados $p \in X$ y $\mu \in \mathcal{U}_q$ que satisfacen

$$\dot{p}(t) = -A^*(t)p(t) + \varphi_x^*(t, x(t), u(t))\mu(t),$$
$$0 = B^*(t)p(t) - \varphi_u^*(t, x(t), u(t))\mu(t)$$

entonces $p \equiv 0$.

Una derivación del siguiente conjunto de condiciones necesarias, suponiendo normalidad, se puede encontrar en [2].

Teorema 4.4. *Supongamos que* (x_0, u_0) *es una solución normal de* (P_c). *Entonces existe un único* $(p, \mu) \in X \times \mathcal{U}_q$ *tal que*

$$\begin{aligned} \dot{p}(t) &= -H_x^*(t, x_0(t), u_0(t), p(t), \mu(t)), \\ 0 &= H_u(t, x_0(t), u_0(t), p(t), \mu(t)) \quad (t \in T). \end{aligned}$$

Más aún, $J((x_0, u_0, p, \mu); (y, v)) \geq 0$ *para todo* $(y, v) \in Z$ *que satisface*

1. $y(t_0) = y(t_1) = 0$;
2. $\dot{y}(t) = f_x(t, x_0(t), u_0(t))y(t) + f_u(t, x_0(t), u_0(t))v(t)$ $(t \in T)$;
3. $\varphi_x(t, x_0(t), u_0(t))y(t) + \varphi_u(t, x_0(t), u_0(t))v(t) = 0$ $(t \in T)$.

- **Aumentabilidad**

Asociada con la integral I consideremos la integral aumentada

$$K(x, u) = \int_{t_0}^{t_1} F(t, x(t), u(t))dt$$

donde

$$F(t, x, u) = L(t, x, u) + \langle \mu(t), \varphi(t, x, u) \rangle + \sigma(t, x, u)G(t, x, u),$$

$$G(t, x, u) = \frac{1}{2} \sum_{\alpha=1}^{q} \varphi_\alpha(t, x, u)^2$$

y denotemos por (A) el problema sin restricciones de minimizar $K(x, u)$ sobre todo $(x, u) \in D$ con $x(t_0) = \xi_0$, $x(t_1) = \xi_1$.

Diremos que el problema (P_c) es *aumentable* en (x_0, u_0) (con respecto a (μ, σ)) si existe una integral aumentada tal que (x_0, u_0) resuelve (A). Note que, en este caso, (x_0, u_0) resuelve (P_c) ya que para cualquier proceso admisible para (P_c), (x, u), tenemos

$$I(x_0, u_0) = K(x_0, u_0) \leq K(x, u) = I(x, u).$$

Teorema 4.5. *Supongamos que* (P_c) *es aumentable en* (x_0, u_0) *con respecto a* (μ, σ), *y* (x_0, u_0) *es normal con respecto a (A), esto es,* $z \equiv 0$ *es la única solución del sistema*

$$\dot{z}(t) = -f_x^*(t, x_0(t), u_0(t))z(t), \quad f_u^*(t, x_0(t), u_0(t))z(t) = 0.$$

Entonces existe $p \in X$ *tal que*

$$\begin{aligned} \dot{p}(t) &= -H_x^*(t, x_0(t), u_0(t), p(t), \mu(t)), \\ 0 &= H_u(t, x_0(t), u_0(t), p(t), \mu(t)) \ (t \in T). \end{aligned}$$

Más aún, $J((x_0, u_0, p, \mu); (y, v)) \geq 0$ *para todo* $(y, v) \in Z$ *que satisface*

1. $y(t_0) = y(t_1) = 0$;
2. $\dot{y}(t) = f_x(t, x_0(t), u_0(t))y(t) + f_u(t, x_0(t), u_0(t))v(t)$ $(t \in T)$;
3. $\varphi_x(t, x_0(t), u_0(t))y(t) + \varphi_u(t, x_0(t), u_0(t))v(t) = 0$ $(t \in T)$.

Demostración: Definimos

$$\tilde{H}(t, x, u, p) := \langle p, f(t, x, u)\rangle - F(t, x, u)$$

y sea

$$\tilde{J}((x, u, p); (y, v)) := \int_{t_0}^{t_1} 2\tilde{\Omega}(t, y(t), v(t))dt$$

donde, para todo $(t, y, v) \in T \times \mathbf{R}^n \times \mathbf{R}^m$,

$$2\tilde{\Omega}(t, y, v) := -[\langle y, \tilde{H}_{xx}(t)y\rangle + 2\langle y, \tilde{H}_{xu}(t)v\rangle + \langle v, \tilde{H}_{uu}(t)v\rangle]$$

y $\tilde{H}(t)$ denota $\tilde{H}(t, x_0(t), u_0(t), p(t))$. Como (x_0, u_0) resuelve (A), por el teorema 4.2 tenemos que existe un único $p \in X$ tal que

$$\dot{p}(t) = -\tilde{H}_x^*(t, x_0(t), u_0(t), p(t)), \quad \tilde{H}_u(t, x_0(t), u_0(t), p(t)) = 0 \quad (t \in T).$$

Además $\tilde{J}((x_0, u_0, p); (y, v)) \geq 0$ para todo $(y, v) \in Z$ que satisface $y(t_0) = y(t_1) = 0$ y

$$\dot{y}(t) = f_x(t, x_0(t), u_0(t))y(t) + f_u(t, x_0(t), u_0(t))v(t) \quad (t \in T).$$

Observemos que

$$\tilde{H}(t, x, u, p) = H(t, x, u, p, \mu) - \sigma(t, x, u)G(t, x, u).$$

Como $\varphi(t, x_0(t), u_0(t)) = 0$ para toda $t \in T$, vemos que $\tilde{H}_x(t) = H_x(t)$ y $\tilde{H}_u(t) = H_u(t)$, y se tiene la primera conclusión. Por otro lado,

$$\tilde{J}((x_0, u_0, p); (y, v)) = J((x_0, u_0, p, \mu)(y, v))$$

$$+ \int_{t_0}^{t_1} \sigma(t, x_0(t), u_0(t)) |\varphi_x(t, x_0(t), u_0(t)) y(t) + \varphi_u(t, x_0(t), u_0(t)) v(t)|^2 dt$$

y tenemos la segunda conclusión. ∎

4.3. Ejemplo

A continuación mostraremos un ejemplo de un problema de control óptimo para el cual la solución no es normal, sin embargo el problema es aumentable en dicha solución.

Ejemplo 4.6. Sean $T = [0, \pi]$, X el espacio de funciones de clase C^1 por fragmentos que mapean T en $\mathbf{R}$, $\mathcal{U}$ el espacio de funciones continuas por fragmentos que mapean T en $\mathbf{R}$, y consideremos el problema (P) de minimizar

$$I(x, u) = \int_0^\pi \{u^2(t) - x^2(t)\} dt$$

sobre todas las funciones $(x, u) \in X \times \mathcal{U}$ que satisfacen

$$\dot{x}(t) = u(t) \ (t \in T), \quad x(0) = x(\pi) = 0, \quad \operatorname{sen} u(t) = 0 \ (t \in T).$$

Para este caso tenemos $n = m = q = 1$, $\xi_0 = \xi_1 = 0$,

$$f(t, x, u) = u, \quad L(t, x, u) = u^2 - x^2, \quad \varphi(t, x, u) = \operatorname{sen} u.$$

Notemos primero que $(x_0, u_0) \equiv (0, 0)$ es solución del problema. Esto se sigue ya que, como es sabido (ver [2, p 64]), $x_0 \equiv 0$ es solución del problema de minimizar

$$\int_0^\pi \{\dot{x}^2(t) - x^2(t)\} dt$$

sobre las funciones $x \in X$ que satisfacen $x(0) = x(\pi) = 0$.

Ahora, de la definición de A y B, tenemos que

$$A(t) = f_x(t, x(t), u(t)) = 0, \quad B(t) = f_u(t, x(t), u(t)) = 1.$$

Claramente (x_0, u_0) no es normal con respecto a (P) ya que, si (p, μ) satisface

$$\dot{p}(t) = -A(t)p(t) + \varphi_x(t, x_0(t), u_0(t))\mu(t) = 0,$$

$$0 = B(t)p(t) - \varphi_u(t, x_0(t), u_0(t))\mu(t) = p(t) - \cos u_0(t)\mu(t) = p(t) - \mu(t)$$

se tendría que $\dot{p}(t) = 0$ y $p(t) = \mu(t)$, o sea, p no necesariamente es cero. Por lo tanto no podemos aplicar el Teorema 4.2.

Por otro lado, la función

$$F(t, x, u) = L(t, x, u) + \mu(t)\varphi(t, x, u) + \frac{1}{2}\sigma(t, x, u)\varphi(t, x, u)^2$$

está dada en este caso por

$$F(t, x, u) = u^2 - x^2 + \mu(t) \operatorname{sen} u + \frac{1}{2}\sigma(t, x, u) \operatorname{sen}^2 u.$$

Por lo tanto, si $\mu \equiv \sigma \equiv 0$, entonces (x_0, u_0) es solución del problema (sin restricciones) de minimizar

$$K(x, u) = \int_0^\pi F(t, x(t), u(t))dt = \int_0^\pi \{u^2(t) - x^2(t)\}dt$$

sobre las funciones $(x, u) \in X \times \mathcal{U}$ que satisfacen

$$\dot{x}(t) = u(t) \ (t \in T), \quad x(0) = x(\pi) = 0.$$

Esto implica que (P) es aumentable en (x_0, u_0).

Además, $z \equiv 0$ es la única solución del sistema

$$\dot{z}(t) = -A(t)z(t) = 0, \quad B(t)z(t) = z(t) = 0$$

y por lo tanto, por el Teorema 4.5, existe $p \in X$ tal que

$$\begin{aligned} \dot{p}(t) &= -H_x(t, x_0(t), u_0(t), p(t), \mu(t)), \\ 0 &= H_u(t, x_0(t), u_0(t), p(t), \mu(t)) \quad (t \in T). \end{aligned}$$

De la definición de H, tenemos que

$$H(t, x, u, p, \mu) = pu - u^2 + x^2 - \mu \operatorname{sen} u$$

y por lo tanto

$$H_x(t,x,u,p,\mu) = 2x, \quad H_u(t,x,u,p,\mu) = p - 2u - \mu \cos u.$$

Nótese que, si $\mu \equiv 0$, se tiene

$$\dot{p}(t) = -2x_0(t) = 0, \quad p(t) - 2u_0(t) - \mu(t)\cos u_0(t) = p(t) = 0$$

y por lo tanto $p \equiv 0$.

Por último, como

$$H_{xx} \equiv 2, \quad H_{xu} \equiv 0, \quad H_{uu}(t,x,u,p,\mu) = -2 + \mu \operatorname{sen} u$$

se sigue que

$$J((x,u,p,\mu);(y,v)) = \int_0^\pi \{2v^2(t) - v^2(t)\mu(t)\operatorname{sen} u(t) - 2y^2(t)\}dt.$$

Por el Teorema 4.5,

$$J((x_0,u_0,p,\mu);(y,v) = 2\int_0^\pi \{v^2(t) - y^2(t)\}dt \geq 0$$

para toda $(y,v) \in X \times \mathcal{U}$ que satisface

$$y(0) = y(\pi) = 0, \quad \dot{y}(t) = v(t) \ (t \in T), \quad v(t) = 0 \ (t \in T).$$

Bibliografía

[1] de Pinho MR, Rosenblueth JF (2006) *Mixed constraints in optimal control: an implicit function theorem approach*, IMA Journal of Mathematical Control and Information **24**: 197-218

[2] Hestenes MR (1966) *Calculus of Variations and Optimal Control Theory*, John Wiley & Sons, New York

[3] Hestenes MR (1975) *Optimization Theory, The Finite Dimensional Case*, John Wiley & Sons, New York

[4] Hestenes MR (1980) *Augmentability in optimization theory*, Journal of Optimization Theory and Applications **32**: 427-440

[5] Rosenblueth, JF (2009) *Augmentability in optimal control*, International Journal of Mathematics and Statistics **57**: 102-109

[6] Rosenblueth, JF (2009) *Equality-inequality mixed constraints in optimal control*, International Journal of Mathematical Analysis **28**: 1369-1387

[7] Marsden J, Tromba A (1998) *Cálculo Vectorial*, Pearson Addison-Wesley