UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO

FACULTAD DE CIENCIAS

AUMENTABILIDAD EN CÁLCULO DE VARIACIONES

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

M A T E M Á T I C A

P R E S E N T A:

Z A Z I L S A N T I Z O H U E R T A

DIRECTOR DE TESIS:

DR. JAVIER FERNANDO ROSENBLUETH
LAGUETTE

2012

1. Datos del alumno:
Santizo
Huerta
Zazil
55547087, jbjzash@hotmail.com
Universidad Nacional Autónoma de México
Facultad de Ciencias
Matemáticas
407008786

2. Datos del tutor:
Dr.
Javier Fernando
Rosenblueth
Laguette

3. Datos del sinodal 1:
Dr.
Ricardo
Berlanga
Zubiaga

4. Datos del sinodal 2:
Dr.
Luis Bernardo
Morales
Mendoza

5. Datos del sinodal 3:
Dr.
Luis Octavio
Silva
Pereyra

6. Datos del sinodal 4:
Dr.
Juan Manuel
García
Islas

7. Datos del trabajo escrito:
"Aumentabilidad en Cálculo de Variaciones"
47 p.
2012

AGRADECIMIENTOS

Al Dr. Rosenblueth, por aceptar ser mi tutor, por apoyarme incondicionalmente hasta que este proyecto se vio culminado, por alentarme y darme bríos a lo largo del camino.

Muy especialmente a mis padres Reyna y Mario, por haberme proporcionado las bases necesarias para llegar a este punto, por todo su amor, comprensión y apoyo.

A mi hermana Itzel que ha sido mi ejemplo a seguir.

A mi novio Alan por haber compartido los momentos más difíciles y los más alegres junto a mi y por todo su amor incondicional.

Especialmente a Rocío Ortega y a la Familia Ortega Vélez que me ha brindado su cariño y me han hecho sentir parte de su familia.

A mis abuelos y tíos paternos y maternos, que me han dado sabios e innumerables consejos.

A mis grandes amigos, Emiliano, Manuel, Fernando, Raybel , José Alberto, Susana, Daniel y Tania, con los cuales he pasado inigualables momentos.

Y finalmente a mi fiel compañera Ekishi que me ha traído grandes momentos de alegría.

CONTENIDO

Introducción

Existen varios enfoques en el estudio de problemas de optimización. En el caso de querer minimizar una función con valores reales definida en un espacio de dimensión finita sujeta a restricciones con igualdades y desiguladades, el enfoque más frecuente consiste en la derivación de las reglas de multiplicadores de Lagrange de primer y segundo orden bajo la hipótesis de normalidad o regularidad.

Un enfoque alternativo consiste en utilizar cierto tipo de *aumentabilidad*, basado en el lagrangiano generalizado, el cual permite establecer condiciones necesarias de primer y segundo orden para un mínimo de problemas con restricciones. De acuerdo con Hestenes [3], dicho enfoque no ha recibido la atención necesaria en el desarrollo de la teoría de optimización.

El principal objetivo de este trabajo es mostrar cómo se puede aplicar la teoría de aumentabilidad a problemas de optimización que involucran restricciones con igualdades. El tipo de problemas que estudiaremos no se reduce a problemas en espacios de dimensión finita. Veremos cómo es posible generalizar la teoría de aumentabilidad a ciertos problemas en cálculo de variaciones que incluyen restricciones con igualdades en las trayectorias admisibles.

Con la idea de situar claramente la contribución de este trabajo veremos primero, con cierto detalle, algunos de los principales resultados de la teoría clásica de optimización en espacios de dimensión finita. Derivaremos condiciones necesarias y suficientes para problemas sin restricciones y para problemas con restricciones expresadas en términos de igualdades. La idea central en el segundo caso consistirá en la derivación de las reglas de multiplicadores de Lagrange de primer y segundo orden suponiendo que la solución del problema es regular. Dado que no es fácil verificar las condiciones que definen regularidad, introduciremos el concepto de normalidad, fácilmente verificable, el cual implica regularidad y nos asegura la existencia de multiplicadores adecuados en un punto mínimo. Por otra parte, suficiencia se obtendrá restringiendo ligeramente las condiciones necesarias de optimalidad.

Sin embargo, hay una gran cantidad de problemas para los cuales la solución no es regular pero sí satisface las reglas de multiplicadores de Lagrange. La teoría de aumentabilidad juega en estos casos un papel crucial.

Ocasionalmente, uno encuentra la siguiente derivación de la regla de multiplicadores de primer orden. Para minimizar una función f sujeta a la restricción $g = 0$, minimizar la función F de la forma $F(x) = f(x) + \lambda g(x)$ sin restricciones. Este procedimiento implica la regla de multiplicadores $F'(x) = f'(x) + \lambda g'(x) = 0$ en el punto mínimo, junto con la condición original $g(x) = 0$. Sin embargo, aunque este procedimiento puede ser una herramienta conveniente para recordar la regla de multiplicadores de primer orden, no es muy satisfactorio debido a que la clase de problemas en donde se puede aplicar es bastante limitada.

Veamos un ejemplo. Consideremos el problema de minimizar la función $f\colon \mathbf{R}^2 \to \mathbf{R}$ dada por

$$f(x,y) = x^2 - 3y - y^2$$

en el conjunto de restricciones

$$S = \{(x,y) \mid g(x,y) = y = 0\}.$$

El origen es la solución de este problema. El lagrangiano correspondiente,

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

no tiene un mínimo sin restricciones, sin importar qué λ se escoja. Sin embargo, $F'(0,0) = 0$ cuando λ es el multiplicador correcto de Lagrange, es decir, $\lambda = 3$. Este procedimiento falla ya que no toma en cuenta los términos de segundo orden que podrían convexificar la función F. En general, la convexidad está determinada por los términos de segundo orden. Introduzcamos para este ejemplo la función aumentada

$$H = f + \lambda g + \left(\frac{\sigma}{2}\right) g^2$$

dada en este caso por

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

La elección $\lambda = 3$ garantiza la anulación del gradiente en el origen y, si $\sigma > 2$, la función aumentada se convexifica.

Este ejemplo sugiere que, si un punto x_0 es una solución local del problema de minimizar una función f sujeta a la restricción $g = 0$, entonces existen constantes λ y σ tales que x_0 es una solución local del problema de minimizar H sin restricciones. En este caso diremos que el problema con restricciones es *aumentable*. Es importante mencionar que todo problema es aumentable en un punto x_0 que satisface las condiciones clásicas de suficiencia. Veremos cómo las reglas de multiplicadores de Lagrange son una consecuencia de aumentabilidad y que aumentabilidad es una consecuencia de la regla de multiplicadores restringida. Además, como quedará claro a lo largo del trabajo, la derivación de las reglas de multiplicadores suponiendo aumentabilidad resultará mucho más sencilla que su derivación suponiendo regularidad.

Por otro lado, veremos cómo un problema puede ser regular sin ser aumentable, y un problema puede ser aumentable sin ser regular. En consecuencia aumentabilidad y regularidad son condiciones alternativas pero no equivalentes para la existencia de multiplicadores de Lagrange.

En la parte final del trabajo analizaremos problemas de optimización de dimensión infinta. Nos concentraremos en el problema de Lagrange que involucra restricciones diferenciales en forma de igualdades. Veremos inicialmente cómo obtener condiciones necesarias de primer y segundo orden para el problema de puntos fijos sin restricciones. Posteriormente plantearemos el problema de Lagrange con restricciones diferenciales y enunciaremos condiciones necesarias para trayectorias normales, a través del principio máximo de Pontryagin, en términos tanto del hamiltoniano como del lagrangiano. Finalmente introduciremos un posible concepto de aumentabilidad, semejante al que utilizamos para problemas de dimensión finita, que permite estudiar soluciones a problemas que no son normales. Como veremos, esta noción permite derivar de manera muy sencilla las condiciones necesarias clásicas del problema de Lagrange.

Capítulo I

Conceptos Básicos

Con la idea de presentar un resumen de los principales conceptos y resultados sobre problemas de optimización que se utilizarán a lo largo del trabajo, resulta conveniente analizar las condiciones necesarias y suficientes que aseguran la existencia de un óptimo.

En este capítulo repasaremos los conceptos de máximos y mínimos de una función con valores reales definida en un subconjunto de $\mathbf{R}^n$, así como los resultados clásicos de condiciones necesarias y suficientes para extremos de funciones diferenciables en un intervalo compacto de $\mathbf{R}$. Posteriormente repasaremos algunos resultados básicos relacionados con diferenciales de funciones de valor real definidas en un subconjunto de $\mathbf{R}^n$.

El problema de optimización a estudiar en este capítulo , que denotamos por $\mathrm{P}(S)$, es "Dados un conjunto $\emptyset \neq S \subset \mathbf{R}^n$ y una función $f\colon \mathbf{R}^n \to \mathbf{R}$, minimizar a f en S". Si S es un conjunto abierto de $\mathbf{R}^n$, diremos que es un problema *sin restricciones.*

1 PUNTOS EXTREMOS DE FUNCIONES

1.1 Definición. El mayor número m (con $m = -\infty$ admitido) tal que $f(x) \geq m$ para toda $x \in S$ se llama la *máxima cota inferior de* f *en* S o el *ínfimo de* f *en* S y se denota por "inf $f(x)$ en S". Si existe un punto x_0 en S tal que $f(x_0) = \inf f(x)$ en S, escribimos

$f(x_0) = \min f(x)$ en S y se dice que dicho punto *minimiza a f en S* y $f(x_0)$ es el *mínimo de f en S*.

Análogamente, la *mínima cota superior* o el *supremo de f en S* es el mínimo valor M (con $M = +\infty$ admitido) tal que $f(x) \leq M$ para toda $x \in S$ y se denota por "$\sup f(x)$ en S". Si existe un punto x_1 en S tal que $f(x_1) = \sup f(x)$ en S, escribimos $f(x_1) = \max f(x)$ en S y se dice que dicho punto *maximiza a f en S* y $f(x_1)$ es el *máximo de f en S*.

Claramente se tiene en S que

$$\sup[-f(x)] = -\inf f(x), \quad \inf[-f(x)] = -\sup f(x)$$

y, por lo tanto, un punto $x_1 \in S$ maximiza a f en S si y solo si x_1 minimiza a $-f$ en S, y un punto $x_0 \in S$ minimiza a f en S si y solo si x_0 maximiza a $-f$ en S. Un punto que maximiza o minimiza a f se llama un *punto extremo* de f.

Comencemos repasando bajo qué condiciones impuestas en el conjunto S y la función f se puede asegurar la existencia de máximos y mínimos.

1.2 Teorema. *Si $\emptyset \neq S \subset \mathbf{R}^n$ es compacto y $f\colon S \to \mathbf{R}$ continua, entonces f alcanza su máximo y mínimo en S.*

Demostración: Sea $m := \inf f(x)$ en S y sea $\{x_q\}$ una sucesión de puntos en S tales que $m = \lim_{q\to\infty} f(x_q)$. Como S está acotado, por el teorema de Bolzano-Weierstrass (toda sucesión acotada de puntos en $\mathbf{R}^n$ tiene una subsecesión convergente) $\{x_q\}$ tiene una subsucesión $\{y_q\}$ convergente y, por ser S cerrado, el límite $x_0 := \lim_{q\to\infty} y_q$ está en S. Por continuidad,

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

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. ▮

Como consecuencia inmediata de este resultado obtenemos el siguiente corolario para funciones continuas en un intervalo compacto.

1.3 Corolario. *Si $f\colon [a,b] \to \mathbf{R}$ es continua y $f(a) = f(b) = 0$ entonces existe un punto extremo de f en (a,b).*

Demostración: Sin pérdida de generalidad, $f \not\equiv 0$. Si $x \in [a,b]$ es tal que $f(x) < 0$ entonces $m = \inf f(x) < 0$. Por el Teorema 1.2, existe un punto mínimo x_0 en $[a,b]$. Como $f(x_0) = m < 0$ y $f(a) = f(b) = 0$, se sigue que $a < x_0 < b$. Por otro lado, si $f(x) > 0$, el máximo x_0 de f está en (a,b). ▮

El concepto de extremo es de naturaleza global dado que se toman en cuenta todos los puntos de S. De acuerdo con la definición, 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.

1.4 Definición. 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*.

2 EXTREMOS EN UN INTERVALO COMPACTO

En esta sección estudiaremos condiciones necesarias y suficientes en el caso en que la función f está definida en un intervalo compacto $S := [a, b]$ y veremos con detalle los resultados para mínimos y máximos locales. Notemos que, para el caso de máximos locales, estos se pueden hallar encontrando los mínimos correspondientes de la función $-f$.

Consideremos entonces el problema $\mathrm{P}(S)$ de minimizar a f en S, en el caso particular en el que $S = [a, b]$ y $n = 1$.

• 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 izquierda o derecha respectivamente.

• **Condiciones suficientes en a y b**

2.1 Teorema. *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

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

Por lo tanto

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

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. ∎

• **Condiciones necesarias en a y b**

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

Demostración: Por el Teorema 2.1, $f(a)$ 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 $f'(a) \geq 0$. Anlogamente, $f'(b) \leq 0$ si $f(b)$ es un mínimo local de f en $[a,b]$ Las afirmaciones entre paréntesis se obtienen remplazando f por $-f$. ▮

• **Condiciones necesarias en (a,b)**

2.3 Teorema. *Supongamos que 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 2.2 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 $x = x_0$. ▮

2.4 Definición. Si f es diferenciable en $x = c$ y $f'(c) = 0$, el punto c se llama un *punto crítico de f* y $f(c)$ es un *valor crítico de f*.

• **Teorema del valor medio**

2.5 Teorema. *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: La función $g\colon [a,b] \to \mathbf{R}$ dada por

$$g(x) := [f(b) - f(a)](x-a) - [f(x) - f(a)](b-a)$$

es continua en $[a,b]$ y $g(a) = g(b) = 0$. Por el Corolario 1.3, existe $c \in (a,b)$ punto extremo de g. Por el Teorema 2.3,

$$0 = g'(c) = f(b) - f(a) - (b-a)f'(c). \ ▮$$

• Si f es diferenciable en $[a,b]$ y $f''(x_0)$ existe entonces

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

donde la función r es tal que

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

• **Condiciones suficientes para un extremo local**

2.6 Teorema. *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

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

Por lo tanto, si $|x-x_0|<\delta$, entonces

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

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

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

• **Condiciones necesarias para un extremo local**

2.7 Teorema. *Si 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* $f''(x_0)\geq 0$.

Demostración: La demostración es inmediata de los Teoremas 2.2 y 2.6. $\blacksquare$

3 FUNCIONES DE VARIAS VARIABLES

En esta sección haremos un repaso de resultados importantes en el estudio de funciones reales definidas en un subconjunto de $\mathbf{R}^n$ y su respectiva diferenciación. La mayoría de los restulados se puede ver con detalle en [2, 5].

3.1 Definición. Sean X y Y espacios vectoriales. Una función A que mapea X en Y se dice que es una *transformación lineal* si

$$A(x_1 + x_2) = Ax_1 + Ax_2, \qquad A(cx) = cAx$$

para toda x, x_1, $x_2 \in X$ y c escalar. Es frecuente utilizar la notación Ax en vez de $A(x)$ si A es lineal. Denotamos por $L(X,Y)$ al conjunto de todas las transformaciones lineales de X en Y y, en lugar de $L(X,\mathbf{R})$, escribimos $L(X)$. Si $A_1, A_2 \in L(X,Y)$ y c_1, c_2 son escalares, definimos $c_1A_1 + c_2A_2$ como

$$(c_1A_1 + c_2A_2)x = c_1A_1x + c_2A_2x \qquad (x \in X).$$

Es claro que $c_1A_1+c_2A_2 \in L(X,Y)$. Si $(X, \|\cdot\|_1)$ y $(Y, \|\cdot\|_2)$ son espacios lineales normados, decimos que $A \in L(X,Y)$ está *acotada* si existe una constante M tal que para toda $x \in X$ se tiene $\|Ax\|_2 \leq M\|x\|_1$. La mínima M con esta propiedad se llama la *norma* de A y la denotamos por $\|A\|$.

3.2 Definición. 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

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

3.3 Definición. 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

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

siempre que el límite exista, y llamamos a D_if una *derivada parcial*.

3.4 Teorema. *Supongamos que f mapea $E \subset \mathbf{R}^n$ en $\mathbf{R}$, E es abierto y f es diferenciable en $x \in E$. Entonces las derivadas parciales $(D_if)(x)$ existen y $f'(x)e_i = (D_if)(x)$.*

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,

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

3.5 Definición. 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 $f \in C^1(E)$.

El siguiente resultado nos da una caracterización de las funciones continuamente difierenciales en términos de las derivadas parciales. Para la demostración, véase [5].

3.6 Teorema. *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 $1 \leq i \leq n$.*

• Si las derivadas parciales $D_i f$ existen para toda $1 \leq i \leq n$, denotamos por $(\nabla f)(x)$ al vector

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

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 3.4, 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 3.2 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 $L(\mathbf{R}^n)$, 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.

3.7 Definición. 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.

Una aplicación inmediata del Teorema 1.2 es el siguiente resultado de existencia para formas cuadráticas.

3.8 Teorema. *Dada una forma cuadrática Q en $\mathbf{R}^n$, existen vectores unitarios $x_0, x_1 \in \mathbf{R}^n$ tales que*

$$Q(x_0)|x|^2 \leq Q(x) \leq Q(x_1)|x|^2 \quad (x \in \mathbf{R}^n).$$

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 $x \in \mathbf{R}^n$ con $x \neq 0$,

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

y el resultado se sigue. ∎

3.9 Definición. Sean $E \subset \mathbf{R}^n$ abierto, f una función que mapea E en $\mathbf{R}$ y $x_0 \in E$. Decimos 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 $\langle f''(x_0)h,h\rangle = f''(x_0;h)$ para toda $h \in \mathbf{R}^n$ se llama el *Hessiano de* f *en* x_0. Nótese que, si f tiene una segunda diferencial en x_0, entonces

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

donde r es tal que

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

3.10 Definición. 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 todas las 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.

El Teorema de Taylor nos resultará crucial a lo largo del trabajo. Para una demostración referimos a [2].

- **Taylor**

3.11 Teorema. *Sean S un subconjunto de $\mathbf{R}^n$, f una función que mapea S en $\mathbf{R}$ y supongamos que $x+th \in S$ para toda $t \in [0,1]$.*

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

$$\begin{aligned} f(x+h) &= f(x)+f'(x+t_1h;h) \\ &= f(x)+\int_0^1 f'(x+th;h)\,dt. \end{aligned}$$

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

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

4 S ABIERTO

En esta sección estudiaremos condiciones necesarias y suficientes de optimalidad para el caso en el que S es un conjunto abierto.

Denotaremos a un mínimo local que es solución de P(S) como una solución local de P(S).

A lo largo de la sección supondremos la hipótesis, que denotamos por (H), de acuerdo con la cual $f \in C^1(S)$ (para condiciones de primer orden) y $f \in C^2(S)$ (de segundo orden).

• Condiciones necesarias para la existencia de una solución local en S

4.1 Teorema. *Si* x_0 *es solución local de* P(S) *entonces* $f'(x_0) = 0$ *y* $f''(x_0) \geq 0$.

Demostración: Sea $\delta > 0$ tal que $f(x) \geq f(x_0)$ para toda x con $|x - x_0| < \delta$. Sean $h \neq 0$, $\epsilon := \delta / |h|$ y definamos $z(t) := x_0 + th$ y $g(t) := f(z(t))$ para toda $t \in (-\epsilon, \epsilon)$. Como

$$|z(t) - x_0| = |t||h| < \epsilon |h| = \delta \quad \text{para toda } t \in (-\epsilon, \epsilon),$$

se tiene que $f(z(t)) \geq f(x_0)$, o sea, $g(t) \geq g(0)$ para toda $t \in (-\epsilon, \epsilon)$. Por lo tanto $g\colon (-\epsilon, \epsilon) \to \mathbf{R}$ tiene un mínimo en $t = 0$. Por el Teorema 2.7(a), $0 = g'(0) = f'(x_0; h)$ y $0 \leq g''(0) = f''(x_0; h)$. ∎

• Condiciones suficientes para la existencia de una solucin local en S

4.2 Teorema. *Si* $x_0 \in S$, $f'(x_0) = 0$ *y* $f''(x_0; h) > 0$ *para toda* $h \neq 0$, *entonces existen* $\delta, m > 0$ *tales que*

$$|x - x_0| < \delta \Rightarrow f(x) \geq f(x_0) + m|x - x_0|^2.$$

Demostración: Por la Definición 3.9, sabemos que

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

donde la función r es tal que

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

Como la forma cuadrática $f''(x_0;h)$ es positiva en h, por el Teorema 3.8 existe $m>0$ tal que, para toda $h\in\mathbf{R}^n$, $f''(x_0;h)\geq 4m|h|^2$. Sea $\delta>0$ tal que $|h|<\delta\Rightarrow|r(h)|\leq m|h|^2$. Por lo tanto, si $|x-x_0|<\delta$ entonces

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

Veamos una aplicación de estos resultados.

4.3 Ejemplo. Consideremos la función $f\colon\mathbf{R}^2\to\mathbf{R}$ dada por

$$f(x,y)=5x^2+4xy+y^2-6x-2y+6.$$

Tenemos que

$$f'(x,y)=(f_x(x,y),f_y(x,y))=(10x+4y-6,4x+2y-2)$$

Si (x_0,y_0) minimiza a f entonces $f'(x_0,y_0)=0$, lo cual se satisface en $(x_0,y_0)=(1,-1)$. Ahora

$$f''(x,y)=\begin{pmatrix} f_{xx}(x,y) & f_{xy}(x,y)\\ f_{yx}(x,y) & f_{yy}(x,y)\end{pmatrix}=\begin{pmatrix}10 & 4\\ 4 & 2\end{pmatrix}$$

y por lo tanto, para toda $(h,k)\neq(0,0)$,

$$f''(x_0,y_0;h,k)=(h,k)\begin{pmatrix}10 & 4\\ 4 & 2\end{pmatrix}\begin{pmatrix}h\\ k\end{pmatrix}=10h^2+8hk+2k^2>0.$$

Esto último se puede ver de la siguiente forma. La desigualdad se satisface si $h=0$ y $k\neq 0$. Si $h\neq 0$, su mínimo relativo a k se encuentra al diferenciar con respecto a k. El resultado es

$$8h+4k=0 \quad \text{o bien} \quad k=-2h.$$

En este caso tenemos

$$10h^2+8hk+2k^2=10h^2-16h^2+8h^2=2h^2>0.$$

Por el Teorema 4.2, $(1,-1)$ minimiza localmente a f.

4.4 Proposición. *Sea A una matriz simétrica no negativa. Entonces son equivalentes:*

a. *A es positiva.*

b. *A es no singular, i.e.,* $\det A \neq 0$.

Demostración: Sea Q la forma cuadrática asociada a A.

(a) $\Rightarrow$ (b): Supongamos que A es singular. Por lo tanto existe $x_0 \neq 0$ tal que $Ax_0 = 0$ lo cual implica que $Q(x_0) = 0$ y A es no positiva.

(b) $\Rightarrow$ (a): Supongamos que existe $x_0 \neq 0$ tal que $Q(x_0) = 0$. Por lo tanto x_0 es un mínimo de Q. Por el Teorema 4.1, $Q'(x_0) = 2Ax_0 = 0$ y A es singular. ∎

4.5 Definición. Un punto $x = c$ es un *punto crítico de* f si $f'(c) = 0$. Un punto crítico c es *no degenerado* si $f''(c)$ es no singular, esto es, $|f''(c)| \neq 0$.

4.6 Corolario. Si c es un punto crítico no degenerado de f entonces $f(c)$ es un mínimo local estricto de f en $S \Leftrightarrow f''(c) \geq 0$.

5 FUNCIONALES LINEALES

En esta sección usaremos el concepto de funcionales lineales para derivar un resultado de suma importancia en la determinación de multiplicadores de Lagrange, que utilizaremos en el capítulo siguiente para derivar condiciones de optimalidad para problemas con igualdades.

5.1 Definición. Sean X un espacio vectorial y $L_i : X \to \mathbf{R}$ $(i = 1, \ldots, m)$ funcionales lineales. Decimos que el conjunto $\{L_i\}_1^m$ es *linealmente independiente* si $\sum_1^m a_i L_i(x) = 0$ $(x \in X) \Rightarrow a_i = 0$ $(i = 1, \ldots, m)$.

5.2 Proposición. *Sean X un espacio vectorial y $\{L_i\}_1^m$ funcionales lineales en X. Entonces 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: Sea $F(x) := (L_1(x), \ldots, L_m(x))$ $(x \in X)$. Por lo tanto F mapea X en un subespacio $Y := F(X)$ de $\mathbf{R}^m$. Como la relación $\sum_1^m a_i L_i(x) = 0$ $(x \in X)$ se satisface $\Leftrightarrow$ $\sum_1^m a_i y^i = 0$ $(y \in Y)$, el conjunto $\{L_i\}$ es linealmente independiente $\Leftrightarrow$ no existe ningún vector en $\mathbf{R}^m$ ortogonal a $Y \Leftrightarrow Y = \mathbf{R}^m \Leftrightarrow$ existen m vectores linealmente independientes $y_1, \ldots, y_m \in Y \Leftrightarrow$ existen $x_1, \ldots, x_m \in X$ tales que $y_j = (L_1(x_j), \ldots, L_m(x_j))$ son linealmente independientes $\Leftrightarrow$ (b). ∎

5.3 Observación. Si $X = \mathbf{R}^n$, un funcional lineal es de la forma $L(x) = \sum_1^n a_j x^j =$

$\langle a, x\rangle$. Si

$$L_i(x) = \langle a_i, x\rangle \quad (i = 1, \ldots, m)$$

entonces $\{L_i\}_1^m$ es linealmente independiente $\Leftrightarrow A = (a_{ij})$ es de rango m.

5.4 Teorema. *Sean X un espacio vectorial, L, L_i $(i = 1, \ldots, m)$ funcionales lineales en X,*

$$R = \{x \in X \mid L_i(x) = 0 \ (i = 1, \ldots, m)\},$$

y supongamos que $L(x) = 0$ para toda $x \in R$. Entonces existen multiplicadores $\{\lambda_i\}_1^m$ tales que

$$L(x) = \sum_1^m \lambda_i L_i(x) \quad (x \in X).$$

Si $\{L_i\}_1^m$ es linealmente independiente entonces los multiplicadores son únicos.

Demostración: Sin pérdida de generalidad, $\{L_i\}_1^m$ es linealmente independiente. Por la Proposición 5.2, existen $x_1, \ldots, x_m \in X$ tales que si

$$A := \begin{pmatrix} L_1(x_1) & \cdots & L_m(x_1) \\ \vdots & & \vdots \\ L_1(x_m) & \cdots & L_m(x_m) \end{pmatrix}$$

entonces $|A| \neq 0$. Sea $\lambda = (\lambda_1, \ldots, \lambda_m)^*$ (que es único) tal que $A\lambda = (L(x_1), \ldots, L(x_m))^*$, de manera que

$$L(x_j) = \sum_{i=1}^m \lambda_i L_i(x_j) \quad (j = 1, \ldots, m).$$

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

$$0 = L_i(x) - \sum_{j=1}^m L_i(x_j)b_j = L_i\Big(x - \sum_1^m x_j b_j\Big) \qquad (i = 1, \ldots, m).$$

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

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

Capítulo II

El Caso de Dimensión Finita

En este capítulo estudiaremos problemas de optimización de dimensión finita con restricciones en forma de igualdades. Buscamos minimizar una función f en el conjunto S definido por ecuaciones de la forma

$$S = \{x \in \mathbf{R}^n \mid g_\alpha(x) = 0 \ (\alpha \in A)\} \text{ donde } A = \{1, \ldots, m\} \ (m < n) \text{ y } g_\alpha\colon \mathbf{R}^n \to \mathbf{R}$$

El problema de minimizar f en S sujeto a la restricción $g = 0$ es equivalente al de minimizar $F = f + \lambda g$ sin restricciones. Sin embargo un mejor procedimiento es utilizar la regla de multiplicadores de Lagrange.

La existencia de adecuados multiplicadores es consecuencia del concepto de regularidad. Sin embargo puede resultar extremadamente difícil, en ciertos casos, verificar dicha propiedad por lo que también estudiaremos el concepto de normalidad en las restricciones. El concepto de normalidad implica a su vez regularidad así como la unicidad de los multiplicadores y dicha implicación se puede probar a través del teorema de la función implícita.

1 RESTRICCIONES CON IGUALDADES

Supongamos la hipótesis (H)(ver página 13, Capítulo 1, Sección 4) y sean $A = \{1, \ldots, m\}$

$(m < n)$, $g_\alpha\colon \mathbf{R}^n \to \mathbf{R}$ $(\alpha \in A)$ funciones de clase C^1 y

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

1.1 Definición. Para toda $x_0 \in S$ definimos

a. *el espacio de vectores tangentes curvilíneos a S en x_0*

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

b. *el espacio de restricciones tangenciales de S en x_0*

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

La notación $g'_\alpha(x_0; h)$, de igual manera que en el capítulo I página 13, hace referancia al producto interno de $g'_\alpha(x_0)$ con h. Es decir $\langle g'_\alpha(x_0), h\rangle = g'_\alpha(x_0; h)$

1.2 Proposición. *Para toda $x_0 \in S$, $C_S(x_0) \subset R_S(x_0)$.*

Demostración: Sea $h \in C_S(x_0)$ y sean $\epsilon > 0$ y $x\colon(-\epsilon, \epsilon) \to S$ tales que $x(0) = x_0$, $\dot{x}(0) = h$. Como $g_\alpha(x(t)) = 0$ para toda $t \in (-\epsilon, \epsilon)$, $\alpha \in A$, diferenciando esta identidad en $t = 0$ se tiene

$$0 = g'_\alpha(x(0); \dot{x}(0)) = g'_\alpha(x_0; h). \ \blacksquare$$

Veamos un ejemplo en el que el converso no necesariamente es cierto.

1.3 Ejemplo. Consideremos la función

$$g(x, y) = (x^2 + y^2 - 1)^2$$

para toda $(x, y) \in \mathbf{R}^2$ de manera que $m = 1$, $n = 2$ y

$$S = \{(x, y) \in \mathbf{R}^2 \mid g(x, y) = 0\}.$$

Como

$$g'(x, y; h, k) = 2(x^2 + y^2 - 1)(2xh + 2yk),$$

tenemos que $g'(0, 1; h, k) = 0$ para toda $(h, k) \in \mathbf{R}^2$ lo cual implica que $R_S(0, 1) = \mathbf{R}^2$. Por otro lado, $(h, k) \in \mathbf{R}^2$ pertenece a $C_S(0, 1)$ si y solo si existen $\epsilon > 0$ y $(x, y)\colon(-\epsilon, \epsilon) \to \mathbf{R}^2$ tales que

i. $(x^2(t) + y^2(t) - 1)^2 = 0$ $(t \in (-\epsilon, \epsilon))$;

ii. $(x(0), y(0)) = (0, 1)$;

iii. $(\dot{x}(0), \dot{y}(0)) = (h, k)$.

Por (i) se tiene que $x^2(t) = 1 - y^2(t)$ para toda $t \in (-\epsilon, \epsilon)$ y por lo tanto

$$2x(t)\dot{x}(t) = -2y(t)\dot{y}(t) \quad (t \in (-\epsilon, \epsilon)).$$

Por (ii), $0 = -2\dot{y}(0)$ y por lo tanto, por (iii), $k = \dot{y}(0) = 0$. Como h es arbitraria, se sigue que

$$C_S(0,1) = \{(h,k) \in \mathbf{R}^2 \mid k = 0\}.$$

1.4 Definición. Decimos que $x_0 \in S$ es un *punto regular de* S si $C_S(x_0) = R_S(x_0)$.

1.5 Definición. Dada $\lambda = (\lambda_1, \ldots, \lambda_m) \in \mathbf{R}^m$, definamos (*función de Lagrange*)

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

1.6 Teorema. Regla de Multiplicadores de Lagrange

Supongamos que un punto (x_0, y_0) es un mínimo local para $f(x, y)$ sujeto a la restricción $g(x, y) = 0$. Supongamos, además que $(g_x, g_y) \neq (0, 0)$ en (x_0, y_0), entonces existe un multiplicador λ tal que, si establecemos

$$F(x, y) = f(x, y) + \lambda g(x, y)$$

entonces

$$F_x(x_0, y_0) = 0 \quad , \quad F_y(x_0, y_0) = 0$$

además

$$F''(x_0, y_0; h, k) = F_{xx}(x_0, y_0)h^2 + 2F_{xy}(x_0, y_0)hk + F_{yy}(x_0, y_0)k^2 \geq 0$$

para toda $(h, k) \neq (0, 0)$ tal que

$$g'(x_0, y_0; h, k) = g_x(x_0, y_0)h + g_y(x_0, y_0)k = 0$$

.

- **Condiciones necesarias para que un punto regular sea solución de** $\mathrm{P}(S)$

1.7 Teorema. *Si x_0 es solución local de $\mathrm{P}(S)$ y un punto regular de S entonces existe $\lambda \in \mathbf{R}^m$ tal que $F'(x_0) = 0$ y $F''(x_0; h) \geq 0$ 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 $\varphi(t) := f(x(t))$ $(-\epsilon < t < \epsilon)$ 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 5.4. Ahora, como $\varphi(t) = F(x(t))$, por el Teorema de Taylor 3.11(b),

$$\frac{\varphi(t) - \varphi(0)}{t^2} = \frac{F(x(t)) - F(x_0)}{t^2} = \frac{1}{2} F''\left(\bar{x}(t); \frac{x(t) - x_0}{t}\right)$$

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

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

• Condiciones suficientes para que un punto regular sea solución de P(S)

1.8 Teorema. *Supongamos que $x_0 \in S$ y que existe $\lambda \in \mathbf{R}^m$ tal que $F'(x_0) = 0$ y $F''(x_0; h) > 0$ para toda $h \in R_S(x_0)$, $h \neq 0$. Entonces existen $m > 0$ y una vecindad N de x_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, reemplacemos $\{x_q\}$ 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 3.11(b),

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

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

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

Enunciaremos ahora una versión del Teorema de la función implícita. Para la demostración referimos al lector a [2].

- **Teorema de la función implícita**

1.9 Teorema. *Sean $S \subset \mathbf{R}^m \times \mathbf{R}^n$ abierto y $f\colon S \to \mathbf{R}^n$ 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*

i. $(t, x_0(t)) \in S$ $(t \in T_0)$.

ii. $f(t, x_0(t)) = 0$ *y* $|f_x(t, x_0(t))| \neq 0$ $(t \in T_0)$.

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. $x(t) = x_0(t)$ $(t \in T_0)$.

b. $f(t, x(t)) = 0$ $(t \in T)$.

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

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

Antes de introducir el concepto de normalidad, probemos el siguiente resultado auxiliar que nos permitirá imponer condiciones que implican regularidad.

1.10 Teorema. *Supongamos que $g\colon \mathbf{R}^n \to \mathbf{R}^m$ $(m < n)$ es de clase C^k $(k \geq 1)$ en una vecindad de un punto x_0 y $g'(x_0)$ es de rango m. Entonces para toda $h \in \mathbf{R}^n$ existen $\delta > 0$ y $x : [-\delta, \delta] \to \mathbf{R}^n$ de clase C^k tales que*

$$x(0) = x_0, \quad \dot{x}(0) = h \quad y \quad g(x(\epsilon)) = g(x_0) + \epsilon g'(x_0; h) \quad (|\epsilon| \leq \delta).$$

Demostración: Sea $h \in \mathbf{R}^n$ y definamos, para toda $(\epsilon, b) \in \mathbf{R} \times \mathbf{R}^m$,

$$\bar{x}(\epsilon, b) := x_0 + \epsilon h + g'^*(x_0)b \quad \text{y} \quad G(\epsilon, b) := g(\bar{x}(\epsilon, b)) - g(x_0) - \epsilon g'(x_0; h).$$

Como $G(0,0) = 0$ y $|G_b(0,0)| = |g'(x_0)g'^*(x_0)| \neq 0$, por el Teorema de la función implícita existen $\delta > 0$ y $b : [-\delta, \delta] \to \mathbf{R}^m$ de clase C^k tales que $b(0) = 0$ y $G(\epsilon, b(\epsilon)) = 0$ $(\epsilon \in [-\delta, \delta])$. Diferenciando esta última identidad con respecto a ϵ en $\epsilon = 0$ obtenemos

$$0 = G_\epsilon(0,0) + G_b(0,0)b'(0) = G_b(0,0)b'(0)$$

y, como $|G_b(0,0)| \neq 0$, se sigue que $b'(0) = 0$. Sea $x(\epsilon) := \bar{x}(\epsilon, b(\epsilon))$ para toda $|\epsilon| \leq \delta$. Entonces $x(0) = x_0$,

$$\dot{x}(0) = \bar{x}_\epsilon(0,0) + \bar{x}_b(0,0)b'(0) = h$$

y se cumple

$$g(x(\epsilon)) - g(x_0) - \epsilon g'(x_0; h) = G(\epsilon, b(\epsilon)) = 0 \quad (|\epsilon| \leq \delta). \blacksquare$$

• Normalidad

1.11 Definición. Decimos que $x_0 \in S$ es un *punto normal de* S si las ecuaciones lineales $g'_\alpha(x_0; h) = 0$ $(\alpha \in A)$ en h 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.

Veamos ahora por qué normalidad implica regularidad.

1.12 Teorema. *Sea x_0 un punto normal de S y sea $h \in R_S(x_0)$. Entonces existen $\epsilon > 0$ y $x\colon (-\epsilon, \epsilon) \to S$ tales que $x(0) = x_0$, $\dot{x}(0) = h$, o sea, $h \in C_S(x_0)$. Si g_α son de clase C^k en una vecindad de x_0, entonces también x es de clase C^k.*

Demostración: El resultado es inmediato del Teorema 1.10 con h satisfaciendo

$$g'_\alpha(x_0; h) = 0 \quad (\alpha \in A). \blacksquare$$

• Condiciones necesarias para que un punto normal sea solución de $\mathrm{P}(S)$

1.13 Teorema. *Si x_0 es solución local de $\mathrm{P}(S)$ y un punto normal de S entonces existe $\lambda \in \mathbf{R}^m$ único tal que $F'(x_0) = 0$ y $F''(x_0; h) \geq 0$ para toda $h \in R_S(x_0)$.*

Demostración: La existencia de $\lambda \in \mathbf{R}^m$ se sigue de los Teoremas 1.7 y 1.12 y la unicidad del Teorema I.5.4. $\blacksquare$

2 FUNCIONES DE PENALIZACION

Normalmente una solución del problema de minimizar una función sujeta a restricciones puede ser obtenida como un conjunto de soluciones elegidas convenientemente del problema sin restricciones. El procedimiento mostrado en esta sección consiste en construir una función $G(x)$ no negativa tal que los puntos que satisfacen nuestras restricciones estén dados por la solución de $G(x) = 0$. Después agregaremos el término de penalización $\sigma G(x)$ a la función $F(x)$ para ser minimizada bajo las restricciones dadas. Finalmente obtendremos el punto mínimo $x(\sigma)$ de la función aumentada $H(x, \sigma) = F(x) + \sigma G(x)$

2.1 Teorema. *Sean F y G funciones continuas 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 punto mínimo de $H(\cdot, \sigma)$ sobre N. Supongamos que existe un único punto x_0 el cual minimza a F en 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 ∞ 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 entonces

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

Ya que G es no negativa en N, se sigue que

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

Por lo tanto, si $\bar{x}$ es el límite de una subsucesión convergente de $\{x_q\}$,

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

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

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

De lo anterior se sigue que $x(\sigma) \to x_0, \sigma \to \infty$. ∎

Veamos ahora un resultado crucial, en términos de formas cuadráticas, que nos permitirá derivar suficiencia para el problema con restricciones.

2.2 Teorema. *Sea C un cono cerrado en $\mathbf{R}^n$ y sean P y Q dos formas cuadráticas con la propiedad de que si $Q(x) = 0$ entonces $P(x) > 0$ para toda $x \neq 0$ en C. Supongamos que $Q(x) \geq 0$ para toda x en C. Entonces existe $\sigma_0 > 0$ tal que $P(x) + \sigma Q(x) > 0$ para toda $x \neq 0$ en C y toda $\sigma \geq \sigma_0$.*

Demostración: Sea $N = \{x \in C : |x| \leq 1\}$, el cual es compacto. Definimos

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

y sea $x(\sigma)$ un punto mínimo de $H(\cdot, \sigma)$ sobre N. En particular, para todo $\sigma \in \mathbf{R}$ se cumple

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

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

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

por lo que $0 < b(\sigma) \leq 1$. Se sigue que si $H(x(\sigma), \sigma) < 0$ entonces $b(\sigma) = 1$. Si $H(x(\sigma), \sigma) = 0$ y $b(\sigma) < 1$, podemos sustituir $x(\sigma)$ por $x(\sigma)/b(\sigma)$. Por lo tanto, o $x(\sigma)$ es un vector unitario, o bien $x(\sigma) = 0$. Por 2.1, tenemos $x(\sigma) \to 0, \sigma \to \infty$. Como $|x(\sigma)| = 1$ a menos que $x(\sigma) = 0$, esto es posible si y sólo si $x(\sigma) = 0$ para valores grandes de σ. Existe, pues, 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

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

Por lo tanto, para cualquier $x \neq 0$ en C y $\sigma \geq \sigma_0$,

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

y se sigue el resultado. ∎

3 AUMENTABILIDAD

El concepto de aumentabilidad ha estado presente en la teoría variacional por un largo tiempo. Hestenes [2], fue el primero en utilizarlo entre 1946 y 1947 en el estudio de problemas de Bolza en cálculo de variaciones.

En esta sección mostraremos un enfoque alternativo de la regla de multiplicadores de Lagrange para el caso de restricciones con igualdades. Mostraremos que la regla de multiplicadores de Lagrange es una consecuencia de aumentabilidad y que aumentabilidad es una consecuencia de la regla de multiplicadores restringida.

• **3.1** Supongamos la hipótesis (H)(ver capítulo 1, página 13, sección 4) y sean $A = \{1, \ldots, m\}$ $(m < n)$ y $g_\alpha\colon \mathbf{R}^n \to \mathbf{R}$ $(\alpha \in A)$ funciones de clase C^2. Consideremos el problema $\mathrm{P}(S)$ de minimizar a f en el conjunto

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

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

$$H(x) = f(x) + \sum_{\alpha=1}^{m} \lambda_\alpha g_\alpha(x) + \frac{\sigma}{2} \sum_{\alpha=1}^{m} g_\alpha(x)^2.$$

Decimos que $\mathrm{P}(S)$ es *aumentable en* x_0 si $x_0 \in S$ es un mínimo local para H, para algún $(\lambda, \sigma) \in \mathbf{R}^m \times \mathbf{R}$. Notemos que $H(x) = f(x)$ para toda $x \in S$ y así, si $\mathrm{P}(S)$ es aumentable en x_0, entonces x_0 es un mínimo local para f en S.

3.2 Teorema. *Si* $\mathrm{P}(S)$ *es aumentable en* x_0 *entonces la regla de multiplicadores de Lagrange se cumple para* x_0*. Además,* x_0 *es un mínimo local de* f *en* S*.*

Demostración: Sea x_0 un mínimo local para H y notemos que $H(x) = F(x) + \sigma G(x)$ con

$$G(x) = \frac{1}{2}\sum_{\alpha=1}^{m} g_\alpha(x)^2, \quad F = f + \lambda_1 g_1 + \ldots + \lambda_m g_m$$

y $S = \{x \in \mathbf{R}^n \mid G(x) = 0\}$. Como x_0 es un mínimo local sin restricciones de H, por el Teorema 2.1 tenemos que $H'(x_0) = 0$ y $H''(x_0) \geq 0$. Como $G(x_0) = 0$ y $G(x) \geq 0$ $(x \in \mathbf{R}^n)$, x_0 minimiza a G, También por el Teorema 2.1 tenemos que $G'(x_0) = 0$. Por lo tanto

$$0 = H'(x_0) = F'(x_0) + \sigma G'(x_0) = F'(x_0),$$

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

por lo que $F''(x_0; h) \geq 0$ siempre que $G''(x_0; h) = 0$. Como $G''(x_0; h) = \sum_1^m g'_\alpha(x_0; h)^2$, tenemos que $F''(x_0; h) \geq 0$ siempre que $g'_\alpha(x_0; h) = 0$ $(\alpha \in A)$, y por lo tanto x_0 satisface la regla de los multiplicadores de Lagrange (ver Teoremas 1.6 y 1.12). ∎

3.3 Teorema. *Supongamos que* $x_0 \in S$ *y, existe* $\lambda \in \mathbf{R}^m$ *tal que* $F'(x_0) = 0$ *y* $F''(x_0) > 0$ *para toda* $h \in R_S(x_0)$ *con* $h \neq 0$*. Entonces existen* σ_0*,* $k > 0$ *y una vecindad*

N de x_0 tales que, si $\sigma \geq \sigma_0$ y $x \in N$, entonces

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

En particular, $\mathrm{P}(S)$ *es aumentable en x_0 y, para toda $x \in N \cap S$,*

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

Demostración: Como x_0 minimiza a G, $G''(x_0) \geq 0$. Además las formas cuadráticas $P = F''(x_0)$ y $Q = G''(x_0)$ satisfacen las hipótesis del Teorema 2.2. Por lo tanto existe $\sigma_0 > 0$ tal que, para toda $h \neq 0$,

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

En consecuencia, $H'(x) = 0$ y, para $h \neq 0$, $H''(x_0; h) > 0$. Por el Teorema 4.2 existen una constante $k > 0$ y una vecindad N de x_0 tales 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 sustituimos σ_0 por cualquier número mayor σ. La última desigualdad se sigue de la anterior cuando $G(x) = 0$. ∎

De acuerdo con los resultados de los últimos dos teoremas, concluimos que la hipótesis de aumentabilidad implica la regla de los multiplicadores de Lagrange y que la regla de los multiplicadores de Lagrange restringida implica aumentabilidad, sin el supuesto del criterio de regularidad. En consecuencia ambos enfoques son alternativos pero no equivalentes para la existencia de adecuados multiplicadores de Lagrange.

4 EJEMPLOS

Consideremos el problema (P) de minimizar $f: \mathbf{R}^2 \to \mathbf{R}$ en el conjunto

$$S = \{(x, y) \in R^2 \mid g(x, y) = 0\}$$

4.1 Ejemplo. (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. 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 local estricto en $(0,0)$ si $\lambda = 4$ y $\sigma > 2$.

4.2 Ejemplo. (P) no es regular ni aumentable en $(0,0)$, pero $(0,0)$ minimiza a f en 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 $T((0,0))$ coincide con la línea $y = 0$, mientras $R_S((0,0))$ es $\mathbf{R}^2$. No es aumentable en $(0,0)$ ya que

$$H(x,y) = x^2 - 4y + (\lambda - 1)y^2 + \frac{\sigma}{2} y^4$$

falla en tener un mínimo en $(0,0)$ para cualesquiera λ y σ.

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

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

De la misma forma que en 2, (P) no es regular en $(0,0)$. Sin embargo, es aumentable en $(0,0)$ ya que

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

se minimiza en $(0,0)$ cuando $\lambda = 4$ y $\sigma > 2$.

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

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

Es claro que f tiene un mínimo local estricto en $(0,0)$ sobre S. Como $g'(0,0) \neq (0,0)$, es un punto normal y por lo tanto también regular para (P). Sin embargo, (P) no es aumentable ya que consideremos la función

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

Si $H'(0,0) = (0,0)$, debemos tener $\lambda = 2$. Por lo tanto,

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

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

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

y por lo tanto (P) no es aumentable en $(0,0)$.

Capítulo III

El Problema Básico de Cálculo de Variaciones

En este capítulo estudiaremos problemas de optimización de dimensión infinta. De manera similar al caso de dimensión finita, comenzaremos estudiando un problema de optimización sin restricciones conocido como el problema básico de puntos fijos en cálculo de variaciones y obtendremos condiciones necesarias de primer y segundo orden para dicho problema.

En las secciones siguientes, nos concentraremos en el problema de Lagrange que involucra restricciones diferenciales en forma de igualdades y enunciaremos condiciones necesarias para trayectorias normales en términos de la formulación Lagrangiana y Hamiltoniana.

De manera semejante al caso de dimensión finita se mostrará un posible concepto de aumentabilidad. Este permitirá derivar de manera muy sencilla las condiciones clásicas del problema de Lagrange, con el fin de estudiar problemas que no son normales.

1 EL PROBLEMA, SOLUCIONES Y CONDICIONES NECESARIAS

Empezaremos definiendo el espacio de funciones en el cual está definida el funcional que se desea minimizar.

1.1 Definición. Una función $x\colon [a,b] \to \mathbf{R}$ se dice que es *continua a trozos en* $[a,b]$ si existen

$$a = t_1 < t_2 < \cdots < t_N < t_{N+1} = b$$

tales que x es continua en cada intervalo $[t_i, t_{i+1}]$ $(i = 1, ..., N)$ y x tiene límites (finitos) por la derecha $x(t_i + 0)$ y por la izquierda $x(t_i - 0)$ para toda $i = 1, ..., N$. Usamos el símbolo $x(t_i)$ para denotar $x(t_i - 0)$ si se trata del intervalo $[a, t_i]$ y $x(t_i + 0)$ si se trata de $[t_i, b]$. En otro caso, $x(t_i)$ denota alguno de los dos valores $x(t_i - 0)$ o $x(t_i + 0)$.

Decimos que una función $x\colon [a, b] \to \mathbf{R}^n$, dada por $x = (x_1, ..., x_n)$, es *continua a trozos en* $[a, b]$ si lo es x_i para toda $i = 1, ..., n$.

Decimos que una función x es C^1 *a trozos*, si x es continua y su derivada $\dot{x}$ es continua a trozos. Los puntos en los cuales $\dot{x}$ es discontinua se llaman *puntos esquina* de x.

• **Planteamiento del problema**

Tenemos dados un intervalo $T := [t_0, t_1]$ en $\mathbf{R}$, dos puntos ξ_0, ξ_1 en $\mathbf{R}^n$, un conjunto $\mathcal{A}$ abierto (relativo) en $T \times \mathbf{R}^n \times \mathbf{R}^n$ y una función L (*lagrangiano*) que mapea $T \times \mathbf{R}^n \times \mathbf{R}^n$ en $\mathbf{R}$. Sea X el espacio vectorial de las funciones C^1 a trozos que mapean T en $\mathbf{R}^n$, definamos

$$X(\mathcal{A}) := \{x \in X \mid (t, x(t), \dot{x}(t)) \in \mathcal{A}\ (t \in T)\},$$

$$X_e(\mathcal{A}) := \{x \in X(\mathcal{A}) \mid x(t_0) = \xi_0,\ x(t_1) = \xi_1\},$$

y consideremos el funcional $I\colon X \to \mathbf{R}$ dada por

$$I(x) := \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t))dt \quad (x \in X).$$

El *problema básico de puntos fijos en cálculo de variaciones*, que denotamos por $\mathrm{P}(\mathcal{A})$, consiste en minimizar I en $X_e(\mathcal{A})$.

El problema depende del lagrangiano, las restricciones y el conjunto $\mathcal{A}$, pero escribiremos explícitamente solo su dependencia de $\mathcal{A}$. Los elementos of X son *arcos* o *trayectorias*, y son *admisibles* si pertenecen a $X_e(\mathcal{A})$. Si $x \in X$, usamos la notación $(\tilde{x}(t))$ para representar $(t, x(t), \dot{x}(t))$ $(t \in T)$. Denotamos los valores de L por $L(t, x, \dot{x})$ y, para u igual a x, $\dot{x}$ o $(x, \dot{x})$, diremos que $L \in C^i(\mathcal{A}; u)$ si L es continua en $\mathcal{A}$ y C^i con respecto a u en $\mathcal{A}$.

1.2 Definición. Una trayectoria x es una *solución* de $\mathrm{P}(\mathcal{A})$ si pertenece a

$$\mathcal{S}(\mathcal{A}) := \{x \in X_e(\mathcal{A}) \mid I(x) \le I(y) \text{ para toda } y \in X_e(\mathcal{A})\}.$$

Para mínimos locales, consideremos las siguientes normas en X:

$$\|x\|_0 := \sup\{|x(t)| : t \in T\} \quad \text{(norma fuerte)},$$

$$\|x\|_1 := \|x\|_0 + \|\dot{x}\|_0 \quad \text{(norma débil).}$$

Decimos que $x \in X_e(\mathcal{A})$ es un *mínimo fuerte* de $\mathrm{P}(\mathcal{A})$ si es un mínimo local de I en $X_e(\mathcal{A})$ con respecto a la norma fuerte en X, o sea, si existe $\epsilon > 0$ tal que $I(x) \leq I(y)$ para toda $y \in X_e(\mathcal{A}) \cap N_0(x;\epsilon)$, donde

$$N_0(x;\epsilon) = \{y \in X : \|x - y\|_0 < \epsilon\}.$$

Un *mínimo débil* de $\mathrm{P}(\mathcal{A})$ corresponde a reemplazar $N_0(x;\epsilon)$ por

$$N_1(x;\epsilon) = \{y \in X : \|x - y\|_1 < \epsilon\}$$

en la definición anterior. Nótese que, si definimos para toda $x \in X$ y $\epsilon > 0$,

$$T_0(x;\epsilon) := \{(t,y) \in T \times \mathbf{R}^n : |x(t) - y| < \epsilon\} \quad \text{(tubo alrededor de } x\text{)}$$

entonces son equivalentes las siguientes afirmaciones:

a. x es un mínimo fuerte de $\mathrm{P}(\mathcal{A})$.

b. Existe $\epsilon > 0$ tal que $x \in \mathcal{S}((T_0(x;\epsilon) \times \mathbf{R}^n) \cap \mathcal{A})$.

c. $x \in X_e(\mathcal{A})$ y existe $\epsilon > 0$ tal que $I(x) \leq I(y)$ para toda $y \in X_e(\mathcal{A})$ con $(t, y(t)) \in T_0(x;\epsilon)$ $(t \in T)$.

Análogamente, si definimos para toda $x \in X$ y $\epsilon > 0$,

$$T_1(x;\epsilon) := \{(t,y,v) \in T_0(x;\epsilon) \times \mathbf{R}^n : |\dot{x}(t) - v| < \epsilon\} \quad \text{(tubo restringido alrededor de } x\text{)}$$

entonces son equivalentes las siguientes afirmaciones:

a. x es un mínimo débil de $\mathrm{P}(\mathcal{A})$.

b. Existe $\epsilon > 0$ tal que $x \in \mathcal{S}(T_1(x;\epsilon) \cap \mathcal{A})$.

c. $x \in X_e(\mathcal{A})$ y existe $\epsilon > 0$ tal que $I(x) \leq I(y)$ para toda $y \in X_e(\mathcal{A})$ con $(t, y(t), \dot{y}(t)) \in T_1(x;\epsilon)$ $(t \in T)$.

1.3 Definición. Para toda $x \in X$ definimos la *primera variación de I con respecto a x* como

$$I'(x;y) := \int_{t_0}^{t_1} \{L_x(\tilde{x}(t))y(t) + L_{\dot{x}}(\tilde{x}(t))\dot{y}(t)\}dt \qquad (y \in X),$$

y la *segunda variación de I con respecto a x* como

$$I''(x;y) := \int_{t_0}^{t_1} 2\Omega(t, y(t), \dot{y}(t))dt \quad (y \in X)$$

donde, para toda $(t, y, \dot{y}) \in T \times \mathbf{R}^{2n}$,

$$2\Omega(t, y, \dot{y}) := \langle y, L_{xx}(\tilde{x}(t))y\rangle + 2\langle y, L_{x\dot{x}}(\tilde{x}(t))\dot{y}\rangle + \langle \dot{y}, L_{\dot{x}\dot{x}}(\tilde{x}(t))\dot{y}\rangle.$$

1.4 Proposición. *Sea $x \in X(\mathcal{A})$. Entonces, para toda $y \in X$, $I'(x;y)$ y $I''(x;y)$ coinciden con la primera y segunda derivada de Gâteux (restringidas a $X(\mathcal{A})$) de I en x en la dirección y (suponiendo que L y sus derivadas involucradas son continuas en $\mathcal{A}$).*

Demostración: Sea $y \in X$. Como $\mathcal{A}$ es abierto, existe $\delta > 0$ tal que $x + \epsilon y \in X(\mathcal{A})$ para toda $|\epsilon| < \delta$. Sea

$$F(t, \epsilon) := L(t, x(t) + \epsilon y(t), \dot{x}(t) + \epsilon \dot{y}(t)) \qquad (t \in T, |\epsilon| < \delta).$$

Sean $t'_1, \dots\ t'_m$ aquellos puntos en donde $\dot{x}$ o $\dot{y}$ es discontinua, $t'_0 := t_0$, $t'_{m+1} := t_1$. Por continuidad en cada intervalo $[t'_i, t'_{i+1}]$ $(i = 0, \dots, m)$ se tiene, sumando de 0 a m,

$$\frac{d}{d\epsilon} I(x + \epsilon y)\ |_{\epsilon=0} = \int_{t_0}^{t_1} F_\epsilon(t, 0)dt = I'(x; y),$$

$$\frac{d^2}{d\epsilon^2} I(x + \epsilon y)\ |_{\epsilon=0} = \int_{t_0}^{t_1} F_{\epsilon\epsilon}(t, 0)dt = I''(x; y). \ \blacksquare$$

1.5 Definición. Sea $Y := \{y \in X \mid y(t_0) = y(t_1) = 0\}$ (*variaciones admisibles*) y consideremos los siguientes conjuntos:

$$\begin{aligned}
\mathcal{E} &:= \{x \in X \mid I'(x; y) = 0 \text{ para toda } y \in Y\}, \\
\mathcal{H} &:= \{x \in X \mid I''(x; y) \geq 0 \text{ para toda } y \in Y\}, \\
\mathcal{L} &:= \{x \in X \mid L_{\dot{x}\dot{x}}(\tilde{x}(t)) \geq 0 \text{ para toda } t \in T\}, \\
\mathcal{W}(\mathcal{A}) &:= \{x \in X(\mathcal{A}) \mid E(t, x(t), \dot{x}(t), u) \geq 0 \text{ para toda } (t, u) \in T \times \mathbf{R}^n \\
&\qquad\qquad \text{con } (t, x(t), u) \in \mathcal{A}\}
\end{aligned}$$

donde la "función exceso de Weierstrass" $E: T \times \mathbf{R}^{3n} \to \mathbf{R}$ está dada por

$$E(t, x, \dot{x}, u) := L(t, x, u) - L(t, x, \dot{x}) - L_{\dot{x}}(t, x, \dot{x})(u - \dot{x}).$$

Los elementos de $\mathcal{E} \cap C^1$ son *extremos*, de $\mathcal{L}$ se dice que satisfacen la *condición de Legendre*, y de $\mathcal{W}(\mathcal{A})$ la *condición de Weierstrass.*

1.6 Teorema. $\mathcal{S}(\mathcal{A}) \subset \mathcal{E} \cap \mathcal{H} \cap \mathcal{L} \cap \mathcal{W}(\mathcal{A})$ *si las derivadas de L en cada caso son continuas.*

Demostración:

a. "$L \in C^1(\mathcal{A}; x, \dot{x}) \Rightarrow \mathcal{S}(\mathcal{A}) \subset \mathcal{E}$". Sean $x \in \mathcal{S}(\mathcal{A})$, $y \in Y$. Como $\mathcal{A}$ es abierto, existe $\delta > 0$ tal que $x + \epsilon y \in X_e(\mathcal{A})$ para toda $|\epsilon| < \delta$. Sea $f\colon(-\delta, \delta) \to \mathbf{R}$ dada por $f(\epsilon) := I(x + \epsilon y)$. Como $x \in \mathcal{S}(\mathcal{A})$, f tiene un mínimo en $\epsilon = 0$. Por lo tanto, $I'(x; y) = f'(0) = 0$.

b. "$L \in C^2(\mathcal{A}; x, \dot{x}) \Rightarrow \mathcal{S}(\mathcal{A}) \subset \mathcal{H}$". Sean $x \in \mathcal{S}(\mathcal{A})$, $y \in Y$. Como en (a), $I''(x; y) = f''(0) \geq 0$.

c. "$L \in C^2(\mathcal{A}; x) \Rightarrow \mathcal{W}(\mathcal{A}) \subset \mathcal{L}$". Sean $x \in \mathcal{W}(\mathcal{A})$, $t \in T$, y definamos

$$G(u) := E(t, x(t), \dot{x}(t), u) \quad (u \in \mathbf{R}^n).$$

Como $G(\dot{x}(t)) = 0$, G tiene un mínimo local en $\dot{x}(t)$. Por lo tanto, para toda $c \in \mathbf{R}^n$,

$$0 \leq \langle c, G''(\dot{x}(t))c \rangle = \langle c, L_{\dot{x}\dot{x}}(\tilde{x}(t))c \rangle.$$

d. "$L \in C^1(\mathcal{A}; \dot{x}) \Rightarrow \mathcal{S}(\mathcal{A}) \subset \mathcal{W}(\mathcal{A})$". Sea $x_0 \in \mathcal{S}(\mathcal{A})$. Sean $s \in (t_0, t_1)$ tal que $\dot{x}_0$ es continua en s y $v \in \mathbf{R}^n$ tal que $(s, x_0(s), v) \in A$. Sea $\mu \in (0, t_1 - s)$ y, para toda $\delta \in [0, \mu]$ y $\epsilon \in [0, 1)$, definamos

$$x(t; \epsilon, \delta) := \begin{cases} x_0(t) & t \in [t_0, s] \cup [s + \delta, t_1] \\ x_0(t) + (t - s)[v - \dot{x}_0(s)] & t \in [s, s + \epsilon\delta] \\ x_0(t) + \lambda(\epsilon)(s + \delta - t)[v - \dot{x}_0(s)] & t \in [s + \epsilon\delta, s + \delta] \end{cases}$$

donde $\lambda(\epsilon) = \epsilon/(1 - \epsilon)$. Como $\mathcal{A}$ es abierto, exis te $\eta > 0$ tal que, para toda $\delta \in [0, \mu]$ y $\epsilon \in [0, \eta]$, $x(\cdot; \epsilon, \delta) \in X_e(\mathcal{A})$. Por lo tanto

$$0 \leq I(x(\cdot; \epsilon, \delta)) - I(x_0) = F(\epsilon, \delta) + G(\epsilon, \delta)$$

donde

$$F(\epsilon, \delta) = \int_s^{s+\epsilon\delta} \{L(t, x(t; \epsilon, \delta), \dot{x}_0(t) + v - \dot{x}_0(s)) - L(\tilde{x}_0(t))\} dt,$$

$$G(\epsilon, \delta) = \int_{s+\epsilon\delta}^{s+\delta} \{L(t, x(t; \epsilon, \delta), \dot{x}_0(t) - \lambda(\epsilon)[v - \dot{x}_0(s)]) - L(\tilde{x}_0(t))\} dt.$$

Como L es continua en $\mathcal{A}$ se tiene, para toda $0 < \epsilon \leq \eta$,

$$\lim_{\delta \to 0} \frac{F(\epsilon, \delta)}{\epsilon\delta} = L(s, x_0(s), v) - L(\tilde{x}_0(s)),$$

$$\lim_{\delta \to 0} \frac{G(\epsilon, \delta)}{\epsilon\delta} = \frac{1}{\lambda(\epsilon)} \{L(s, x_0(s), \dot{x}_0(s) - \lambda(\epsilon)[v - \dot{x}_0(s)]) - L(\tilde{x}_0(s))\}.$$

Por lo tanto

$$0 \leq \lim_{\epsilon \to 0} \left[\lim_{\delta \to 0} \frac{F(\epsilon, \delta)}{\epsilon \delta} + \lim_{\delta \to 0} \frac{G(\epsilon, \delta)}{\epsilon \delta} \right] = E(s, x_0(s), \dot{x}_0(s), v).$$

Finalmente, por continuidad de L y $L_{\dot{x}}$, esta relación se satisface en los puntos finales y esquina de x_0. ∎

1.7 Observación. Nótese que este último resultado da condiciones necesarias para una solución de P($\mathcal{A}$) donde $\mathcal{A}$ es cualquier abierto en $T \times \mathbf{R}^n \times \mathbf{R}^n$. Por lo tanto, condiciones necesarias para un mínimo débil o fuerte se derivan reemplazando $\mathcal{A}$ por $T_1(x;\epsilon) \cap \mathcal{A}$ o $(T_0(x;\epsilon) \times \mathbf{R}^n) \cap \mathcal{A}$ respectivamente.

En particular, si x es un mínimo débil de P($\mathcal{A}$) entonces $x \in \mathcal{E} \cap \mathcal{H} \cap \mathcal{L} \cap \mathcal{W}(T_1(x;\epsilon) \cap \mathcal{A})$ para alguna $\epsilon > 0$ y, si x es un mínimo fuerte de P($\mathcal{A}$), entonces $x \in \mathcal{E} \cap \mathcal{H} \cap \mathcal{L} \cap \mathcal{W}(\mathcal{A})$ ya que, para toda $\epsilon > 0$, $x \in \mathcal{W}((T_0(x;\epsilon) \times \mathbf{R}^n) \cap \mathcal{A}) \Leftrightarrow x \in \mathcal{W}(\mathcal{A})$.

2 LA ECUACION DE EULER

En esta sección derivaremos la ecuación de Euler y algunas de sus consecuencias. En particular veremos con detalle la condición de esquina de Weierstrass-Erdmann y el Teorema de Hilbert.

2.1 Teorema. *Si $L \in C^1(\mathcal{A}; x, \dot{x})$ y $x \in X(\mathcal{A})$ entonces son equivalentes las siguientes afirmaciones:*

a. $x \in \mathcal{E}$.

b. *Existe $c \in \mathbf{R}^n$ tal que*

$$L_{\dot{x}}(\tilde{x}(t)) = \int_{t_0}^{t} L_x(\tilde{x}(s))ds + c \quad (t \in T).$$

Demostración: Sea $P(t) := \int_{t_0}^{t} L_x(\tilde{x}(s))ds$ $(t \in T)$.

(a) $\Rightarrow$ (b): Definamos

$$c := \frac{1}{t_1 - t_0} \int_{t_0}^{t_1} \{L_{\dot{x}}(\tilde{x}(t)) - P(t)\}dt,$$

$$z(t) := \int_{t_0}^{t} \{L_{\dot{x}}(\tilde{x}(s)) - P(s) - c\}^* ds \quad (t \in T).$$

Nótese que $z \in Y$ y $\dot{z}(t)^* = L_{\dot{x}}(\tilde{x}(t)) - P(t) - c$. Por lo tanto

$$\begin{aligned} 0 &= \int_{t_0}^{t_1} \frac{d}{dt}\langle z(t), (P(t)+c)^*\rangle dt \\ &= \int_{t_0}^{t_1} \{\langle z(t), L_x(\tilde{x}(t))^*\rangle + \langle \dot{z}(t), L_{\dot{x}}(\tilde{x}(t))^* - \dot{z}(t)\rangle\} dt \\ &= I'(x;z) - \int_{t_0}^{t_1} |\dot{z}(t)|^2 dt. \end{aligned}$$

Por (a), $\dot{z}(t) = 0$ $(t \in T)$ y por lo tanto se cumple (b).

(b) $\Rightarrow$ (a): Por definición de P se tiene, para toda $y \in Y$,

$$\begin{aligned} 0 &= \int_{t_0}^{t_1} \frac{d}{dt}\langle y(t), (P(t)+c)^*\rangle dt \\ &= \int_{t_0}^{t_1} \{\langle y(t), \dot{P}(t)^*\rangle + \langle \dot{y}(t), (P(t)+c)^*\rangle\} dt \\ &= I'(x;y). \ \blacksquare \end{aligned}$$

2.2 Observación. La ecuación en el Teorema 2.1 es la forma integral de la *ecuación de Euler*

$$\frac{d}{dt}L_{\dot{x}}(\tilde{x}(t)) = L_x(\tilde{x}(t)) \qquad (t \in T).$$

Si $\dot{x}$ tiene una discontinuidad, la derivada d/dt se interpreta como una derivada por la izquierda o la derecha y la ecuación se satisface incluso si x no tiene una segunda derivada.

Aunque (con las hipótesis del Teorema 2.1) $x \in \mathcal{E} \Rightarrow x$ satisface la ecuación de Euler, el converso no necesariamente es cierto. Por ejemplo, sean $T = [0, 2\pi]$,

$$L(t, x, \dot{x}) = (\dot{x}^2 - x^2)/2,$$

$$x_0(t) := \begin{cases} \operatorname{sen} t & t \in [0, \pi] \\ 0 & t \in [\pi, 2\pi]. \end{cases}$$

Entonces x_0 satisface la ecuación de Euler, pero $x_0 \notin \mathcal{E}$.

Si x satisface la ecuación de Euler entonces $x \in \mathcal{E}$ si la función $t \mapsto L_{\dot{x}}(\tilde{x}(t))$ $(t \in T)$ pertenece a X.

2.3 Observación. Para toda $(t, x, u, p) \in T \times \mathbf{R}^{3n}$ sea

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

y definamos

$$M(x) := \{p \in X \mid \dot{p}(t) = -H_x^*(\tilde{x}(t), p(t)),\ H_u(\tilde{x}(t), p(t)) = 0\ (t \in T)\} \quad (x \in X).$$

Entonces (con las hipótesis del Teorema 2.1) $x \in \mathcal{E} \Leftrightarrow M(x) \neq \emptyset$.

• **Definiciones**

• Llamamos *extremos* a los elementos de $\mathcal{E} \cap C^1$.
• Una trayectoria $x \in X$ es *no singular* si $|L_{\dot{x}\dot{x}}(\tilde{x}(t))| \neq 0$ $(t \in T)$.
• El integrando L es *positivo regular en* $\mathcal{A}$ si $\mathcal{A}$ es convexo en $\dot{x}$ y $L_{\dot{x}\dot{x}}(t, x, \dot{x}) > 0$ para toda $(t, x, \dot{x}) \in \mathcal{A}$.

2.4 Proposición. *Supongamos que $L \in C^1(\mathcal{A}; x, \dot{x})$ y $x \in X(\mathcal{A}) \cap \mathcal{E}$. Entonces*

a. *La función $t \mapsto L_{\dot{x}}(t, x(t), \dot{x}(t))$ es continua. En particular, si t es un punto de discontinuidad de $\dot{x}$, se satisface la condición de esquina de Weierstrass-Erdmann:*

$$L_{\dot{x}}(t, x(t), \dot{x}(t-0)) = L_{\dot{x}}(t, x(t), \dot{x}(t+0)).$$

b. *Si $x \in \mathcal{W}(\mathcal{A})$ y $F(t, u) := L(t, x(t), u) - L_{\dot{x}}(\tilde{x}(t))u$ $((t, u) \in T \times \mathbf{R}^n)$ entonces la función $t \mapsto F(t, \dot{x}(t))$ es continua y*

$$E(t, x(t), \dot{x}(t-0), \dot{x}(t+0)) = E(t, x(t), \dot{x}(t+0), \dot{x}(t-0)) = 0.$$

Demostración:
(a): Se sigue por ser continua la función $P(t) := \int_{t_0}^{t} L_x(\tilde{x}(s))ds$ del Teorema 2.1. ∎
(b): Nótese que

$$\begin{aligned}
0 &\leq E(t, x(t), \dot{x}(t-0), \dot{x}(t+0)) \\
&= L(t, x(t), \dot{x}(t+0)) - L(t, x(t), \dot{x}(t-0)) - L_{\dot{x}}(t, x(t), \dot{x}(t-0))(\dot{x}(t+0) - \dot{x}(t-0)) \\
&= F(t, \dot{x}(t+0)) - F(t, \dot{x}(t-0)).
\end{aligned}$$

Análogamente,

$$0 \leq E(t, x(t), \dot{x}(t+0), \dot{x}(t-0)) = F(t, \dot{x}(t-0)) - F(t, \dot{x}(t+0))$$

y por lo tanto $F(t, \dot{x}(t+0)) = F(t, \dot{x}(t-0))$. ∎

• **Teorema de Hilbert**

2.5 Teorema. *Si $x \in X(\mathcal{A})$ es un extremo no singular y $L \in C^r(\mathcal{A})$ entonces $x \in C^r(T)$ $(r \geq 2)$.*

Demostración: Sea $c \in \mathbf{R}^n$ tal que

$$L_{\dot{x}}(\tilde{x}(t)) = P(t) + c \quad (t \in T), \qquad P(t) := \int_{t_0}^{t} L_x(\tilde{x}(s))ds$$

y sea $G(t,v) := L_{\dot{x}}(t,x(t),v) - P(t) - c$. Nótese que $G(t,\dot{x}(t)) = 0$ y

$$|G_v(t,\dot{x}(t))| = |L_{\dot{x}\dot{x}}(\tilde{x}(t))| \neq 0.$$

Supongamos que $L \in C^r(\mathcal{A})$ y $x \in C^m(T)$ para $1 \leq m < r$. Por hipótesis, $m \geq 1$ y, como también $G \in C^m$, por el teorema de la función implícita $\dot{x} \in C^m(T)$. Por lo tanto $x \in C^{m+1}(T)$. ∎

2.6 Proposición. *Sea L positivo regular en $\mathcal{A}$ y $L \in C^2(\mathcal{A};\dot{x})$. Entonces*

$$E(t,x,\dot{x},u) > 0$$

para toda $(t,x,\dot{x}) \in \mathcal{A}$ y $(t,x,u) \in \mathcal{A}$ con $u \neq \dot{x}$. En particular, si $x \in X(\mathcal{A}) \cap \mathcal{E}$ entonces x es un extremo no singular.

Demostración: Por Taylor, dados $(t,x,\dot{x})$ y (t,x,u) en $\mathcal{A}$ con $u \neq \dot{x}$, existe $\lambda \in (0,1)$ tal que

$$E(t,x,\dot{x},u) = \tfrac{1}{2}\langle u - \dot{x}, L_{\dot{x}\dot{x}}(t,x,\dot{x} + \lambda(u-\dot{x}))(u-\dot{x})\rangle > 0.$$

La segunda afirmación se sigue de la Proposición 2.4(b). ∎

2.7 Teorema. *Sea $x \in X(\mathcal{A})$ y definamos*

$$F(t,u) := L(t,x(t),u) - p(t)u \quad ((t,u) \in T \times \mathbf{R}^n)$$

donde

$$p(t) = L_{\dot{x}}(\tilde{x}(t_0)) + \int_{t_0}^{t} L_x(\tilde{x}(s))ds.$$

Entonces son equivalentes:

a. $x \in \mathcal{E} \cap \mathcal{W}(\mathcal{A})$.

b. *$F(t,u) \geq F(t,\dot{x}(t))$ para toda $(t,u) \in T \times \mathbf{R}^n$ con $(t,x(t),u) \in \mathcal{A}$.*

Demostración: Nótese que, si $p(t) = L_{\dot{x}}(\tilde{x}(t))$ $(t \in T)$, entonces

$$F(t,u) - F(t,\dot{x}(t)) = E(t,x(t),\dot{x}(t),u).$$

(a) $\Rightarrow$ (b): $x \in \mathcal{E} \Rightarrow p(t) = L_{\dot{x}}(\tilde{x}(t))$ y (b) se sigue ya que $x \in \mathcal{W}(\mathcal{A})$.

(b) $\Rightarrow$ (a): Por (b),

$$0 = F_u(t,\dot{x}(t)) = L_{\dot{x}}(\tilde{x}(t)) - p(t).$$

Por lo tanto $x \in \mathcal{E} \cap \mathcal{W}(\mathcal{A})$. ∎

3 LA CONDICION DE JACOBI

La teoría de Jacobi en cálculo de variaciones caracteriza el signo de la segunda variación en términos de puntos conjugados. Veremos en esta sección algunos de sus resultados principales.

Para toda $x \in X$ denotemos por J_x la segunda variación de I con respecto a x, i.e.,

$$J_x(y) := I''(x; y) = \int_{t_0}^{t_1} 2\Omega(t, y(t), \dot{y}(t))dt \qquad (y \in X)$$

donde, para toda $(t, y, \dot{y}) \in T \times \mathbf{R}^{2n}$,

$$2\Omega(t, y, \dot{y}) := \langle y, L_{xx}(\tilde{x}(t))y\rangle + 2\langle y, L_{x\dot{x}}(\tilde{x}(t))\dot{y}\rangle + \langle \dot{y}, L_{\dot{x}\dot{x}}(\tilde{x}(t))\dot{y}\rangle.$$

Sea $\mathcal{E}_x$ el conjunto de trayectorias (llamadas *accesorias* o *extremos secundarios con respecto a x*) que satisfacen la ecuación de Euler con respecto al integrando Ω, i.e.,

$$\begin{aligned}\mathcal{E}_x &:= \{y \in X \mid J'_x(y; z) = 0 \text{ para toda } z \in Y\}\\ &= \{y \in X \mid \text{existe } c \in \mathbf{R}^n \text{ tal que } \Omega_{\dot{y}}(\tilde{y}(t)) = \int_{t_0}^{t} \Omega_y(\tilde{y}(s))ds + c \ \ (t \in T)\}.\end{aligned}$$

Nótese que

$$\Omega^*_y(t, y, \dot{y}) = L_{xx}(\tilde{x}(t))y + L_{x\dot{x}}(\tilde{x}(t))\dot{y}, \quad \Omega^*_{\dot{y}}(t, y, \dot{y}) = L_{\dot{x}x}(\tilde{x}(t))y + L_{\dot{x}\dot{x}}(\tilde{x}(t))\dot{y}$$

y por lo tanto $y \in \mathcal{E}_x \Rightarrow y$ satisface la ecuación diferencial

$$\frac{d}{dt}\Big[L_{\dot{x}x}(\tilde{x}(t))y(t) + L_{\dot{x}\dot{x}}(\tilde{x}(t))\dot{y}(t)\Big] = L_{xx}(\tilde{x}(t))y(t) + L_{x\dot{x}}(\tilde{x}(t))\dot{y}(t) \qquad (t \in T)$$

que denotamos por (J), llamada la *ecuación de Jacobi*.

3.1 Proposición. *Supongamos que $L \in C^2(A; \dot{x})$ y $x \in X(A) \cap \mathcal{L}' \cap C^1$. Entonces $\mathcal{E}_x \subset C^1$.*

Demostración: El resultado se sigue de la Proposición 2.6 ya que

$$\Omega_{\dot{y}\dot{y}}(t, y, \dot{y}) = L_{\dot{x}\dot{x}}(\tilde{x}(t)) \quad (t \in T). \ \blacksquare$$

Veamos ahora cómo se pueden caracterizar los extremos secundarios en términos de un sistema lineal de ecuaciones diferenciales. La demostración es inmediata de las definiciones.

3.2 Teorema. *Supongamos que* $L \in C^2(A; x, \dot{x})$ *y* $x \in X(A) \cap \mathcal{L}' \cap C^1$. *Entonces, para toda* (y, q) *en* $X \times X$, *son equivalentes:*

a. $y \in \mathcal{E}_x$ *y* $q(t) = L_{\dot{x}x}(\tilde{x}(t))y(t) + L_{\dot{x}\dot{x}}(\tilde{x}(t))\dot{y}(t)$.

b. (y, q) *satisface, para toda* $t \in T$, *el sistema lineal*

$$\begin{aligned} \dot{y}(t) &= A(t)y(t) + B(t)q(t) \\ \dot{q}(t) &= C(t)y(t) - A^*(t)q(t) \end{aligned}$$

que denotamos por $(\mathrm{J})'$, *donde*

$$\begin{aligned} A(t) &= -L_{\dot{x}\dot{x}}^{-1}(\tilde{x}(t))L_{\dot{x}x}(\tilde{x}(t)) \\ B(t) &= L_{\dot{x}\dot{x}}^{-1}(\tilde{x}(t)) \\ C(t) &= L_{xx}(\tilde{x}(t)) - L_{x\dot{x}}(\tilde{x}(t))L_{\dot{x}\dot{x}}^{-1}(\tilde{x}(t))L_{\dot{x}x}(\tilde{x}(t)). \end{aligned}$$

3.3 Definición. Dado $x \in X$, un punto $s \in (t_0, t_1]$ es *conjugado a* t_0 *en* x si existe $y \in \mathcal{E}_x$ tal que $y(t_0) = y(s) = 0$ y $y \not\equiv 0$ en (t_0, s). Denotamos por $\mathcal{C}(x)$ el conjunto de puntos conjugados a t_0 en x.

3.4 Observación. Si $L \in C^2(\mathcal{A}; x, \dot{x})$ y $x \in X(\mathcal{A}) \cap \mathcal{L}' \cap C^1$, de la teoría de ecuaciones diferenciales lineales aplicada a $(\mathrm{J})'$ se sigue que un extremo secundario y está determinado de manera única por los valores de $y(a)$, $\dot{y}(a)$ o, equivalentemente, por los valores $y(a)$, $q(a)$ en un punto $t = a$ en T. En particular, si $y(a) = \dot{y}(a) = 0$, entonces $y \equiv 0$.

3.5 Teorema. *Si* $L \in C^2(\mathcal{A}; x, \dot{x})$ *y* $x \in X(\mathcal{A}) \cap \mathcal{L}' \cap C^1$, *entonces*

$$x \in \mathcal{H} \Rightarrow \mathcal{C}(x) \cap (t_0, t_1) = \emptyset.$$

Demostración: Supongamos que existe $s \in \mathcal{C}(x) \cap (t_0, t_1)$. Sea y como en la Definición 3.3 y definamos $z(t) := y(t)$ si $t \in [t_0, s]$, $z(t) := 0$ si $t \in [s, t_1]$. Claramente $z \in Y$ y

$$\begin{aligned} J_x(z) &= \int_{t_0}^{s} \{\Omega_y(\tilde{z}(t))z(t) + \Omega_{\dot{y}}(\tilde{z}(t))\dot{z}(t)\}dt \\ &= \int_{t_0}^{s} \frac{d}{dt}\Omega_{\dot{y}}(\tilde{z}(t))z(t)dt \\ &= 0. \end{aligned}$$

Como $x \in \mathcal{H}$, z minimiza J_x en Y. Como x es C^1, las funciones Ω, Ω_y, $\Omega_{\dot{y}}$, son continuas y, por lo tanto, $z \in \mathcal{E}_x$. Como $z(t) = \dot{z}(t) = 0$ $(t \in (s, t_1))$ se sigue por la Observación 3.4 que $z \equiv 0$, lo cual contradice que y no se anula. ∎

El siguiente resultado caracteriza los elementos de $\mathcal{H}$ en términos de la condición de Legendre y puntos conjuntados. Para la demostración, véase [1].

3.6 Teorema. *Si $L \in C^2(\mathcal{A})$ y x es un extremo no singular, entonces*

$$x \in \mathcal{H} \Leftrightarrow x \in \mathcal{L} \quad y \quad \mathcal{C}(x) \cap (t_0, t_1) = \emptyset.$$

4 EL PROBLEMA DE LAGRANGE

Supongamos la información anterior pero ahora también consideremos una función φ que mapea $T \times \mathbf{R}^n \times \mathbf{R}^n$ en $\mathbf{R}^q$. Sea

$$\mathcal{B} := \{(t, x, \dot{x}) \in \mathcal{A} \mid \varphi(t, x, \dot{x}) = 0\}$$

y consideremos el problema P($\mathcal{B}$) (el *problema clásico de Lagrange con restricciones con igualdades*).

El problema es, pues, el de minimizar

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

sujeto a

a. $x: T \to \mathbf{R}^n$ es C^1 a trozos;
b. $x(t_0) = \xi_0$, $x(t_1) = \xi_1$;
c. $(t, x(t), \dot{x}(t)) \in \mathcal{A}$ $(t \in T)$;
d. $\varphi(t, x(t), \dot{x}(t)) = 0$ $(t \in T)$.

Asumimos que $\varphi \in C^2(\mathcal{A})$ y que la matriz $\varphi_{\dot{x}}(t, x, \dot{x})$ de $q \times n$, tiene rango q en $\mathcal{B}$.

Normalidad

• Formulación Hamiltoniana

Para toda $(t, x, \dot{x}, p, \mu, \lambda) \in T \times \mathbf{R}^n \times \mathbf{R}^n \times \mathbf{R}^n \times \mathbf{R}^q \times \mathbf{R}$ sea

$$H(t, x, \dot{x}, p, \mu, \lambda) := \langle p, \dot{x} \rangle - \lambda L(t, x, \dot{x}) - \langle \mu, \varphi(t, x, \dot{x}) \rangle$$

y denotemos por $\mathcal{U}_q$ el espacio de las funciones continuas a trozos que mapean T en $\mathbf{R}^q$.

El siguiente resultado es conocido como el Prinicpio Máximo de Pontryagin para el problema de Lagrange con restricciones con igualdades y referimos a [1] para su demostración.

4.1 Teorema. *Si x_0 es solución de P($\mathcal{B}$) entonces existen $\lambda_0 \geq 0$, $p \in X$, y $\mu \in \mathcal{U}_q$ continuas en cada intervalo de continuidad de $\dot{x}_0$, no simultaneamente iguales a cero en T, tal que*

a. *$\dot{p}(t) = -H_x^*(\tilde{x}_0(t), p(t), \mu(t), \lambda_0)$ y $H_{\dot{x}}(\tilde{x}_0(t), p(t), \mu(t), \lambda_0) = 0$ en cada intervalo de continuidad de $\dot{x}_0$.*

b. *$H(t, x_0(t), u, p(t), \mu(t), \lambda_0) \leq H(\tilde{x}_0(t), p(t), \mu(t), \lambda_0)$ para toda $(t, u) \in T \times \mathbf{R}^n$ con $(t, x_0(t), u) \in \mathcal{B}$.*

4.2 Observación. Supongamos que x_0 es solución de P($\mathcal{B}$). Sea (p, μ, λ_0) como en el Teorema 4.1. Entonces

$$\langle h, H_{\dot{x}\dot{x}}(\tilde{x}_0(t), p(t), \mu(t), \lambda_0)h\rangle \leq 0 \quad \textit{para toda } h \in \mathbf{R}^n \textit{ tal que } \varphi_{\dot{x}}(\tilde{x}_0(t))h = 0.$$

4.3 Diremos que una trayectoria $x \in X(\mathcal{B})$ es *normal* en P($\mathcal{B}$) si, dado $(p, \mu) \in X \times \mathcal{U}_q$ tal que, para toda $t \in T$,

$$\dot{p}(t) = \varphi_x^*(\tilde{x}(t))\mu(t) \quad [\, = -H_x^*(\tilde{x}(t), p(t), \mu(t), 0)\,]$$

$$0 = p(t) - \varphi_{\dot{x}}^*(\tilde{x}(t))\mu(t) \quad [\, = H_{\dot{x}}^*(\tilde{x}(t), p(t), \mu(t), 0)\,]$$

entonces $p \equiv 0$. En este caso, claramente, también $\mu \equiv 0$, ya que

$$\mu^*(t) = p^*(t)\varphi_{\dot{x}}^*(\tilde{x}(t))[\varphi_{\dot{x}}(\tilde{x}(t))\varphi_{\dot{x}}^*(\tilde{x}(t))]^{-1}.$$

Por lo tanto x es normal en P($\mathcal{B}$) si no existe una solución no nula de

$$\dot{p}^*(t) = p^*(t)\varphi_{\dot{x}}^*(\tilde{x}(t))[\varphi_{\dot{x}}(\tilde{x}(t))\varphi_{\dot{x}}^*(\tilde{x}(t))]^{-1}\varphi_x(\tilde{x}(t)).$$

Notemos que, si x_0 es una solución normal de P($\mathcal{B}$) entonces, en el Teorema 4.1, $\lambda_0 > 0$. En este caso, los multiplicadores λ_0, p, μ pueden ser elegidos tal que $\lambda_0 = 1$ y, cuando son elegidos así, son únicos.

4.4 Definición. Dado $(x, p, \mu) \in X \times X \times \mathcal{U}_q$, definimos la *segunda variación* (con respecto a $H(t, 1)$) como

$$K(x, p, \mu; y) = \int_{t_0}^{t_1} 2\tilde{\Omega}(t, y(t), \dot{y}(t))dt \qquad (y \in X)$$

donde, para toda $(t, y, \dot{y}) \in T \times \mathbf{R}^n \times \mathbf{R}^n$,

$$2\tilde{\Omega}(t, y, \dot{y}) := -\big[\langle y, H_{xx}(t,1)y\rangle + 2\langle y, H_{x\dot{x}}(t,1)\dot{y}\rangle + \langle \dot{y}, H_{\dot{x}\dot{x}}(t,1)\dot{y}\rangle\big]$$

y $H(t,1)$ denota $H(\tilde{x}(t), p(t), \mu(t), 1)$.

4.5 Definición. Para toda $x \in X$ definimos el conjunto $Y(\mathcal{B}, x)$ de *$\mathcal{B}$-variaciones admisibles a lo largo de x* como el conjunto de todas las $y \in X$ que satisfacen

$$y(t_0) = y(t_1) = 0 \quad \text{y} \quad \varphi_x(\tilde{x}(t))y(t) + \varphi_{\dot{x}}(\tilde{x}(t))\dot{y}(t) = 0 \quad (t \in T).$$

Veamos ahora condiciones de primer y segundo orden para soluciones normales del problema. La demostración se puede ver en [1].

4.6 Teorema. *Supongamos x_0 es solución de P($\mathcal{B}$) y es normal en P($\mathcal{B}$). Entonces existen únicas $p \in X$ y $\mu \in \mathcal{U}_q$ continuas en cada intervalo de continuidad de $\dot{x}_0$ tal que, si $H(t,1)$ denota $H(\tilde{x}_0(t), p(t), \mu(t), 1)$, entonces*

a. *$\dot{p}(t) = -H_x^*(t,1)$ y $H_{\dot{x}}(t,1) = 0$ en cada intervalo de continuidad de $\dot{x}_0$;*

b. *$H(t, x_0(t), \dot{x}, p(t), \mu(t), 1) \le H(t,1)$ para toda $(t, \dot{x}) \in T \times \mathbf{R}^n$ con $(t, x_0(t), \dot{x}) \in \mathcal{B}$;*

c. *$\langle h, H_{\dot{x}\dot{x}}(t,1)h\rangle \le 0$ para toda $h \in \mathbf{R}^n$ tal que $\varphi_{\dot{x}}(\tilde{x}_0(t))h = 0$;*

d. *$K(x_0, p, \mu; y) \ge 0$ para toda $y \in Y(\mathcal{B}, x_0)$.*

• Formulación Lagrangiana

Para toda $(t, x, \dot{x}, \mu, \lambda) \in T \times \mathbf{R}^n \times \mathbf{R}^n \times \mathbf{R}^q \times \mathbf{R}$ Sea

$$F(t, x, \dot{x}, \mu, \lambda) := \lambda L(t, x, \dot{x}) + \langle \mu, \varphi(t, x, \dot{x})\rangle$$

de manera que

$$H(t, x, \dot{x}, p, \mu, \lambda) = \langle p, \dot{x}\rangle - F(t, x, \dot{x}, \mu, \lambda).$$

Notemos que, si x_0 es solución de P($\mathcal{B}$) y λ_0, p, μ son como en el Teorema 4.1, entonces

$$0 = H_{\dot{x}}^*(\tilde{x}_0(t), p(t), \mu(t), \lambda_0) = p(t) - F_{\dot{x}}^*(\tilde{x}_0(t), \mu(t), \lambda_0)$$

y además la función $t \mapsto F_{\dot{x}}^*(\tilde{x}_0(t), \mu(t), \lambda_0)$ $(t \in T)$ pertenece a X. Como $H_x = -F_x$, obtenemos

$$\frac{d}{dt} F_{\dot{x}}(\tilde{x}_0(t), \mu(t), \lambda_0) = F_x(\tilde{x}_0(t), \mu(t), \lambda_0) \quad (t \in T).$$

Observemos también que si E_F es la función E de Weierstrass

$$E_F(t,x,\dot{x},u,\mu,\lambda) := F(t,x,u,\mu,\lambda) - F(t,x,\dot{x},\mu,\lambda) - F_{\dot{x}}(t,x,\dot{x},\mu,\lambda)(u-\dot{x})$$

entonces tenemos

$$E_F(\tilde{x}_0(t),u,\mu(t),\lambda) = H(\tilde{x}_0(t),p(t),\mu(t),\lambda) - H(t,x_0(t),u,p(t),\mu(t),\lambda).$$

Esta observación implica, por el Teorema 4.1, el siguiente resultado.

4.7 Teorema. *Si x_0 es solución de $P(\mathcal{B})$ entonces existen $\lambda_0 \geq 0$ y $\mu \in \mathcal{U}_q$ continua en cada intervalo de continuidad de $\dot{x}_0$, no simultaneamente cero en T, tales que*

a. *Existe $c \in \mathbf{R}^n$ tal que*

$$F_{\dot{x}}(\tilde{x}_0(t),\mu(t),\lambda_0) = \int_{t_0}^{t} F_x(\tilde{x}_0(s),\mu(s),\lambda_0)ds + c \quad (t \in T);$$

b. *$E_F(\tilde{x}_0(t),u,\mu(t),\lambda_0) \geq 0$ para toda $(t,u) \in T \times \mathbf{R}^n$ con $(t,x_0(t),u) \in \mathcal{B}$.*

Notemos que $H_{xx} = -F_{xx}$, $H_{x\dot{x}} = -F_{x\dot{x}}$, y $H_{\dot{x}\dot{x}} = -F_{\dot{x}\dot{x}}$. Por la Observación 4.2, si x_0 es solución de P($\mathcal{B}$) y (μ,λ_0) como en el Teorema 4.7, entonces

$$\langle h, F_{\dot{x}\dot{x}}(\tilde{x}_0(t),\mu(t),\lambda_0)h\rangle \geq 0 \quad \textit{para toda } h \in \mathbf{R}^n \textit{ tal que } \varphi_{\dot{x}}(\tilde{x}_0(t))h = 0 \quad (t \in T).$$

Además, $K(x,p,\mu;y)$ coincide con $J''(x,\mu;y)$, la segunda variación de la funcional

$$J(x,\mu) := \int_{t_0}^{t_1} F(t,x(t),\dot{x}(t),\mu(t),1)dt \qquad ((x,\mu) \in X \times \mathcal{U}_q),$$

dada por

$$J'(x,\mu;y) = \int_{t_0}^{t_1} 2\Omega_\mu(t,y(t),\dot{y}(t))dt \quad (y \in X)$$

donde

$$2\Omega_\mu(t,y,\dot{y}) := \langle y, F_{xx}(\tilde{x}(t),\mu(t),1)y\rangle + 2\langle y, F_{x\dot{x}}(\tilde{x}(t),\mu(t),1)\dot{y}\rangle + \langle \dot{y}, F_{\dot{x}\dot{x}}(\tilde{x}(t),\mu(t),1)\dot{y}\rangle.$$

Para expresar el resultado correspondiente al Teorema 4.6 (para soluciones normales) de manera similar al Teorema 1.6 definamos, para toda $\mu \in \mathcal{U}_q$,

$$\mathcal{E}(\mu) := \{x \in X \mid \text{existe } c \in \mathbf{R}^n \text{ tal que}$$
$$F_{\dot{x}}(\tilde{x}(t),\mu(t),1) = \textstyle\int_{t_0}^{t} F_x(\tilde{x}(s),\mu(s),1)ds + c \ (t \in T)\},$$

$$\mathcal{H}(\mu) := \{x \in X \mid J'(x,\mu;y) \geq 0 \text{ para toda } y \in Y(\mathcal{B},x)\},$$

$$\mathcal{L}(\mu) := \{x \in X \mid \langle h, F_{\dot{x}\dot{x}}(\tilde{x}(t),\mu(t),1)h\rangle \geq 0 \text{ para toda } h \in \mathbf{R}^n \text{ tal que } \varphi_{\dot{x}}(\tilde{x}(t))h = 0 \ (t \in T)\},$$

$$\mathcal{W}(\mathcal{B},\mu) := \{x \in X(\mathcal{B}) \mid E_F(\tilde{x}(t),u,\mu(t),1) \geq 0 \text{ para toda } (t,u) \in T \times \mathbf{R}^n \text{ con } (t,x(t),u) \in \mathcal{B}\}.$$

4.8 Teorema. *Si x_0 es una solución normal de P($\mathcal{B}$) entonces existe un único $\mu \in \mathcal{U}_q$ tal que x_0 pertenece a $\mathcal{E}(\mu)$, $\mathcal{H}(\mu)$, $\mathcal{L}(\mu)$ y $\mathcal{W}(\mathcal{B},\mu)$.*

Aumentabilidad

Para una función dada $\sigma: T \times \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$ y para todo $(t,x,\dot{x},\mu) \in T \times \mathbf{R}^n \times \mathbf{R}^n \times \mathbf{R}^q$, definimos

$$\tilde{F}(t,x,\dot{x},\mu) := L(t,x,\dot{x}) + \langle \mu, \varphi(t,x,\dot{x})\rangle + \sigma(t,x,\dot{x})G(t,x,\dot{x})$$

donde

$$G(t,x,\dot{x}) := \frac{1}{2}\sum_1^q \varphi_\alpha(t,x,\dot{x})^2.$$

Notemos que

$$\tilde{F}(t,x,\dot{x},\mu) = F(t,x,\dot{x},\mu,1) + \frac{\sigma(t,x,\dot{x})}{2}|\varphi(t,x,\dot{x})|^2.$$

Asociada a la integral I, consideremos la integral aumentada

$$\tilde{J}(x,\mu) := \int_{t_0}^{t_1} \tilde{F}(t,x(t),\dot{x}(t),\mu(t))dt \qquad ((x,\mu) \in X \times \mathcal{U}_q)$$

y denotemos por Q($\mathcal{A},\mu,\sigma$) el problema de minimizar sin restricciones a $\tilde{J}(\cdot,\mu)$ en $X_e(\mathcal{A})$.

4.9 Definición. Para $x_0 \in X_e(\mathcal{B})$ diremos que P($\mathcal{B}$) es *aumentable en* x_0 si existen σ y μ tales que x_0 es solución de Q($\mathcal{A},\mu,\sigma$). Notemos que, en este caso, x_0 es solución de P($\mathcal{B}$) ya que, para $x \in X_e(\mathcal{B})$, tenemos

$$I(x_0) = \tilde{J}(x_0,\mu) \leq \tilde{J}(x,\mu) = I(x).$$

Como en el caso de dimensión finita, veamos ahora cómo las condiciones de primer y segundo orden obtenidas anteriormente en el Teorema 4.8 bajo la hipótesis de normalidad, se pueden obtener de manera sencilla bajo la hipótesis de aumentabilidad.

4.10 Teorema. *Sea $x_0 \in X_e(\mathcal{B})$ y supongamos que* $\mathrm{P}(\mathcal{B})$ *es aumentable en x_0. Entonces existe $\mu \in \mathcal{U}_q$ tal que x_0 pertenece a $\mathcal{E}(\mu)$, $\mathcal{H}(\mu)$, $\mathcal{L}(\mu)$ y $\mathcal{W}(\mathcal{B}, \mu)$. Además, x_0 es solución de P($\mathcal{B}$).*

Demostración: Sean σ y μ tales que x_0 es solución de $\mathrm{Q}(\mathcal{A}, \mu, \sigma)$. Aplicando el Teorema 1.6 que establece que x_0 pertenece a $\mathcal{E}$, $\mathcal{H}$, $\mathcal{L}$ y $\mathcal{W}(\mathcal{A})$ con respecto a la integral $\tilde{J}(\cdot, \mu)$, se tiene que x_0 pertenece a

$$\tilde{\mathcal{E}}(\mu, \sigma) := \{x \in X \mid \text{existe } c \in \mathbf{R}^n \text{ tal que}$$
$$\tilde{F}_{\dot{x}}(\tilde{x}(t), \mu(t)) = \textstyle\int_{t_0}^{t} \tilde{F}_x(\tilde{x}(s), \mu(s))ds + c \ (t \in T)\},$$

$$\tilde{\mathcal{H}}(\mu, \sigma) := \{x \in X \mid \tilde{J}''(x, \mu; y) \geq 0 \text{ para toda } y \in Y\},$$

$$\tilde{\mathcal{L}}(\mu, \sigma) := \{x \in X \mid \tilde{F}_{\dot{x}\dot{x}}(\tilde{x}(t), \mu(t)) \geq 0 \ (t \in T)\},$$

$$\tilde{\mathcal{W}}(\mathcal{A}, \mu, \sigma) := \{x \in X(\mathcal{A}) \mid E_{\tilde{F}}(t, x(t), \dot{x}(t), u, \mu(t)) \geq 0$$
$$\text{para toda } (t, u) \in T \times \mathbf{R}^n \text{ con } (t, x(t), u) \in \mathcal{A}\}.$$

Por la primera afirmación, existe $c \in \mathbf{R}^n$ tal que

$$\tilde{F}_{\dot{x}}(\tilde{x}_0(t), \mu(t)) = \int_{t_0}^{t} \tilde{F}_x(\tilde{x}_0(s), \mu(s))ds + c \quad (t \in T)$$

y por lo tanto

$$F_{\dot{x}}(\tilde{x}_0(t), \mu(t), 1) = \int_{t_0}^{t} F_x(\tilde{x}_0(s), \mu(s), 1)ds + c \quad (t \in T)$$

demostrando que $x_0 \in \mathcal{E}(\mu)$.

Por la segunda afirmación, tenemos

$$\tilde{J}''(x_0, \mu; y) \geq 0 \quad \text{para toda } y \in X \text{ tal que } y(t_0) = y(t_1) = 0$$

donde

$$\tilde{J}''(x_0, \mu; y) = \int_{t_0}^{t_1} \{\langle y(t), \tilde{F}_{xx}(t)y(t)\rangle + 2\langle y(t), \tilde{F}_{x\dot{x}}(t)\dot{y}(t)\rangle + \langle \dot{y}(t), \tilde{F}_{\dot{x}\dot{x}}(t)\dot{y}(t)\rangle\}dt$$

y $\tilde{F}(t)$ denota $\tilde{F}(\tilde{x}_0(t), \mu(t))$. Como uno prueba fácilmentte,

$$\tilde{J}''(x_0, \mu; y) = J'(x_0, \mu; y) + \int_{t_0}^{t_1} \sigma(\tilde{x}_0(t))|\varphi_x(\tilde{x}_0(t)y(t) + \varphi_{\dot{x}}(\tilde{x}_0(t))\dot{y}(t)|^2 dt$$

y por lo tanto $J''(x_0, \mu; y) \geq 0$ para toda $y \in Y(\mathcal{B}, x_0)$, demostrando que $x_0 \in \mathcal{H}(\mu)$.

Por la tercer afirmación, tenemos $\tilde{F}_{\dot{x}\dot{x}}(\tilde{x}_0(t), \mu(t)) \geq 0$ $(t \in T)$. Observemos que

$$\tilde{F}_{\dot{x}\dot{x}}(\tilde{x}_0(t), \mu(t)) = F_{\dot{x}\dot{x}}(\tilde{x}_0(t), \mu(t), 1) + \sigma(\tilde{x}_0(t)) G_{\dot{x}\dot{x}}(\tilde{x}_0(t)).$$

Por otro lado, para toda $x \in X$,

$$G_{\dot{x}}(\tilde{x}(t)) = \sum_1^q \varphi_\alpha(\tilde{x}(t)) \varphi_{\alpha\dot{x}}(\tilde{x}(t))$$

y por lo tanto, para toda $h \in \mathbf{R}^n$,

$$\begin{aligned} \langle h, G_{\dot{x}\dot{x}}(\tilde{x}_0(t))h\rangle &= h^* \Big(\sum_1^q \varphi^*_{\alpha\dot{x}}(\tilde{x}_0(t)) \varphi_{\alpha\dot{x}}(\tilde{x}_0(t))\Big) h \\ &= \sum_1^q \big(\varphi_{\alpha\dot{x}}(\tilde{x}_0(t))h\big)^2 = |\varphi_{\dot{x}}(\tilde{x}_0(t))h|^2. \end{aligned}$$

Por consiguiente

$$\langle h, F_{\dot{x}\dot{x}}(\tilde{x}_0(t), \mu(t), 1)h\rangle + \sigma(\tilde{x}_0(t)) |\varphi_{\dot{x}}(\tilde{x}_0(t))h|^2 \geq 0 \quad (h \in \mathbf{R}^n,\ t \in T)$$

lo que implica que $x_0 \in \mathcal{L}(\mu)$.

Finalmente, por la cuarta afirmación, tenemos

$$\tilde{F}(t, x_0(t), u, \mu(t)) - \tilde{F}(\tilde{x}_0(t), \mu(t)) - \tilde{F}_{\dot{x}}(\tilde{x}_0(t)), \mu(t))(u - \dot{x}_0(t)) \geq 0$$

para toda $(t, u) \in T \times \mathbf{R}^n$ con $(t, x_0(t), u) \in \mathcal{A}$. Esto implica que

$$E_F(\tilde{x}_0(t), u, \mu(t), 1) + \frac{\sigma(t, x_0(t), u)}{2} |\varphi(t, x_0(t), u)|^2 \geq 0$$

y por lo tanto x_0 pertenece a $\mathcal{W}(\mathcal{B}, \mu)$. ∎

5 EJEMPLO

Terminaremos el trabajo dando un ejemplo de un problema de Lagrange con restricciones para el cual la solución no es normal, pero el problema es aumentable en esa solución.

5.1 Ejemplo. Sea $T = [0, \pi]$ y conisderemos el problema de minimizar

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

sobre todas las funciones $x\colon T \to \mathbf{R}$ de clase C^1 a trozos que satisfacen $x(0) = x(\pi) = 0$ y $\operatorname{sen} \dot{x}(t) = 0$ $(t \in T)$.

Para este caso tenemos $L(t, x, \dot{x}) = \dot{x}^2 - x^2$ y $\varphi(t, x, \dot{x}) = \operatorname{sen} \dot{x}$.

Por un lado, sabemos que $x_0 \equiv 0$ es solución del problema (ver [1]). Para ver si es normal, consideremos el sistema

$$\dot{p}(t) = \varphi_x(\tilde{x}_0(t))\mu(t), \quad 0 = p(t) - \varphi_{\dot{x}}(\tilde{x}_0(t))\mu(t)$$

que en este caso corresponde a $\dot{p}(t) = 0$ y $p(t) = \mu(t)$ por lo que p no necesariamente es cero. Por lo tanto x_0 no es normal.

Por otro lado, la función

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

está dada en este caso por

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

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

$$\tilde{J}(x, \mu) = \int_0^\pi F(t, x(t), \dot{x}(t), \mu(t))dt = \int_0^\pi \{\dot{x}^2(t) - x^2(t)\}dt$$

sobre las funciones $x\colon T \to \mathbf{R}$ de clase C^1 por fragmentos que satisfacen $x(0) = x(\pi) = 0$. Esto implica que (P) es aumentable en x_0.

Bibliografía

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

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

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

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

[5] Rudin W (1976) *Principles of Mathematical Analysis*, 3a edición, McGraw-Hill, New York

[6] Rupp RD (1973) *A nonlinear optimal control minimization technique*, Transactions of the American Mathematical Society, **178**: 357-381