UNIVERSIDAD NACIONAL DE INGENIERIA

FACULTAD DE CIENCIAS

ESCUELA PROFESIONAL DE MATEMATICA

INFORME DE SUFICIENCIA PARA OPTAR EL TÍTULO PROFESIONAL DE LICENCIADO EN MATEMATICA

TITULADO:

PRINCIPIO VARIACIONAL DE EKELAND Y APLICACIONES

PRESENTADO POR:

LUIS ALBERTO IQUISE MAMANI

ASESOR:

Dra. YBOON GARCÍA RAMOS

LIMA - PERÚ

2011

Dedico este modesto trabajo a Tomasa, mi madre.

Agradecimiento

En estas lineas quiero expresar mi agradecimiento a la Dra. Yboon García Ramos, mi asesora, por sus enseñanzas, paciencia y comprensión en la realización del presente informe, asimismo a todas las personas que, de noble corazón ayudaron a la culminación de este trabajo.

Contenido

RESUMEN

Desde su publicación en 1972, el Principio Variacional de Ekeland, ha sido utilizado en múltiples aplicaciones en Optimización, Ecuaciones Diferenciales, Geometría Diferencial, Teoría de Control, entre otras. Además de proporcionar demostraciones elegantes de muchos resultados ya conocidos. El Teorema de Weiertrass es de sobra conocido, que establece la existencia de máximos y mínimos para funciones continuas sobre espacios métricos compactos, pero si se está interesado sólo en problemas de minimización, no es necesario suponer que f sea continua, bastará con una noción más débil. Surge el concepto de función semicontinua inferiormente. Entonces, en cierto modo lo que se va hacer a lo largo de este trabajo es estudiar como es que se debilitan dichas hipótesis para resolver el problema de existencia de mínimos de funciones, lo cual tiene su punto de partida en el principio ya indicado.

En el Capítulo 1 se establecen los conceptos teóricos y definiciones que se utilizarán posteriormente. Principalmente se revisará el concepto de función semicontinua inferiormente en espacios métricos, así también el concepto de función localmente Lipschitziana, el gradiente generalizado de una función en un punto determinado y el cono normal a un conjunto, todo esto sobre espacios de Banach.

En el Capítulo 2 se presenta el famoso Principio Variacional de Ekeland en sus versiones fuerte y débil, tema central del presente trabajo. Así mismo se ve la aplicación directa de la versión débil a la Teoría de Punto Fijo y a la existencia de mínimos para funciones Gâteaux diferenciables, y se obtendrán algunos resultados para funciones que además satisfacen la condición de Palais-Smale.

Finalmente, en el Capítulo 3 se dan aplicaciones de la versión fuerte del principio, una dirigida a la Teoría de Optimización, y otra aplicada al Cálculo Generalizado, donde se establece el concepto de gradiente generalizado para funciones localmente Lipschitzianas, así como una versión $n-$dimensional del Teorema del Valor Medio.

Capítulo 1

FUNDAMENTOS TEÓRICOS

1.1. Minimización de funciones semicontinuas inferiormente

Definición 1.1 *([4], [9]) Sean X un conjunto no vacío, $f : X \to \mathbb{R} \cup \{+\infty\}$ una función.*

1. *Se define el dominio de f (o dominio efectivo de f) como el siguiente conjunto*

$$dom f = \{x \in X \,/\, f(x) < +\infty\}.$$

2. *Se dice que f es propia si y sólo si $dom f \neq \emptyset$.*

Observación:

El hecho de que en la definición se haya incluido el valor de $+\infty$ a la función surge de lo siguiente:

Si nuestro objetivo original es minimizar una función f sobre un subconjunto propio K de X, consideraremos la función $\tilde{f} : X \to \mathbb{R} \cup \{+\infty\}$ definida por

$$\tilde{f}(x) = \begin{cases} f(x) & \text{si} \quad x \in K \\ +\infty & \text{si} \quad x \notin K \end{cases}$$

Así, nuestro esquema será uniforme y no dependerá exclusivamente de K, al menos en las definiciones.

Definición 1.2 *([3]) Un espacio métrico es un conjunto X en el cual se ha definido una función distancia (o métrica) $d : X \times X \to \mathbb{R}$, la cual satisface:*

i) $d(x, y) \geq 0$, *para todo* $x, y \in X$.

ii) $d(x, y) = 0$ *si y sólo si* $x = y$.

iii) $d(x, y) = d(y, x)$, *para todo* $x, y \in X$.

iv) $d(x, z) \leq d(x, y) + d(y, z)$, *para todo* $x, y, z \in X$.

La propiedad (iv) se denomina desigualdad triangular. *Denotamos el espacio métrico por* (X, d) *o simplemente* X.

En un espacio métrico X, una sucesión $(x_n)_{n\in\mathbb{N}} \subset X$ converge a $x \in X$ si y sólo si $\lim\limits_{n\to\infty} d(x_n, x) = 0$. Es decir

$$\text{dado} \quad \epsilon > 0, \quad \text{existe} \quad N \in \mathbb{N} \quad \text{tal que} \quad n \geq N \quad \text{implica que} \quad d(x_n, x) < \epsilon.$$

Diremos que una sucesión $(x_n)_{n\in\mathbb{N}} \subset X$ es de Cauchy si y sólo si

$$\text{dado} \quad \epsilon > 0, \quad \text{existe} \quad N \in \mathbb{N} \quad \text{tal que} \quad m, n \geq N \quad \text{implica que} \quad d(x_m, x_n) < \epsilon.$$

Un espacio métrico X se dice que es completo si y sólo si, toda sucesión de Cauchy en X es convergente.

Definición 1.3 *([4], [5], [9]) Sean* (X, d) *un espacio métrico,* $f : X \to \mathbb{R} \cup \{+\infty\}$ *una función.*

1. *Se dice que* f *es semicontinua inferiormente en un punto* $x_0 \in X$ *si y sólo si, para cada sucesión* (x_n) *en* X *convergente a* x_0*, se tiene que*

$$f(x_0) \leq \liminf_{n\to\infty} f(x_n).$$

2. *Se dice que* f *es semicontinua inferiormente si y sólo si* f *es semicontinua inferiormente en cada* $x \in X$.

Ejemplo 1.1 *Definamos la función* $f : \mathbb{R} \to \mathbb{R}$*, mediante* $f(x) = \lceil x \rceil$*, función menor entero de* x*. Es decir, dados* $x \in \mathbb{R}$ *y* $n \in \mathbb{Z}$ *cualesquiera*

$$\lceil x \rceil = n \quad \text{si y sólo si} \quad n - 1 < x \leq n$$

Parte de la gráfica de dicha función se muestra en la Figura 1.1. Nótese que dados $x_0 \in \mathbb{R}$ *y* $\epsilon > 0$ *cualesquiera, para cada* $(x_n) \subset \mathbb{R}$ *con* $x_n \to x_0$*, se puede encontrar un* $N \in \mathbb{N}$ *tal que*

$$\forall n \geq N : \ f(x_n) \ \geq \ f(x_0) - \epsilon$$

Lo cual, equivalentemente expresa que $\lim\limits_{n\to\infty} \inf f(x_n) \ \geq \ f(x_0)$*. Así concluimos que dicha función es semicontinua inferiormente en* $\mathbb{R}$.

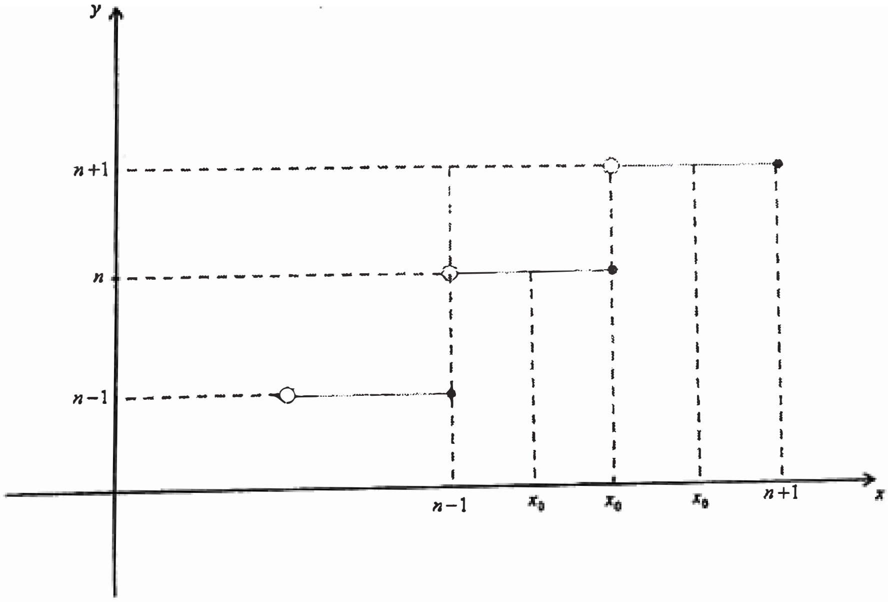

Figura 1.1: Función Menor Entero de x en $\mathbb{R}$

Teorema 1.1 *([4]) Sean (X, d) un espacio métrico y $f : X \to \mathbb{R} \cup \{+\infty\}$ una función. Luego, las siguientes afirmaciones son equivalentes*

a) f es semicontinua inferiormente.

b) Para cada $\lambda \in \mathbb{R}$, el conjunto $S_\lambda = \{x \in X \,/\, f(x) \leq \lambda\}$ es cerrado.

Demostración:

Veamos que (a) implica (b). En efecto, sea $\lambda \in \mathbb{R}$ cualquiera. Tomemos una sucesión (x_n) en $\{x \in X \,/\, f(x) \leq \lambda\}$ convergente a x_0, debemos probar que x_0 pertenece a dicho conjunto, lo cual es cierto, pues, si f es semicontinua inferiormente, lo es en x_0, entonces

$$f(x_0) \leq \liminf_{n\to\infty} f(x_n) \leq \lambda.$$

De aquí se concluye que $x_0 \in \{x \in X \,/\, f(x) \leq \lambda\}$, en conclusión, dicho conjunto es cerrado.

Veamos ahora que (b) implica (a). En efecto, tomemos una sucesión (x_n) convergente a x_0 en X.

Si $f(x_0) < +\infty$, consideremos un $\epsilon > 0$ arbitrario y $\lambda = f(x_0) - \epsilon$, luego el conjunto $\{x \in X \,/\, f(x) > f(x_0) - \epsilon\}$ será abierto debido a que su complemento es cerrado por hipótesis. Como x_0 pertenece a dicho conjunto, existe un $N \in \mathbb{N}$ tal que para todo $n \geq N$, $f(x_n) > f(x_0) - \epsilon$, entonces, $\liminf_{n\to\infty} f(x_n) \geq f(x_0) - \epsilon$. Como ϵ fue tomado arbitrariamente, se sigue que

$$f(x_0) \leq \liminf_{n\to\infty} f(x_n)$$

Si $f(x_0) = +\infty$, consideremos $\lambda > 0$ cualquiera. Se tiene que el conjunto dado por $\{x \in X \,/\, f(x) > \lambda\}$ es abierto y x_0 pertenece a este conjunto. Por lo cual se deduce que $\liminf_{n\to\infty} f(x_n) \geq \lambda$, y como λ es arbitrario, $\liminf_{n\to\infty} f(x_n) = +\infty$.
Así, se concluye que f es semicontinua inferiormente.

1.2. Espacios de Banach

Definición 1.4 *([3]) Sea X un espacio vectorial sobre $\mathbb{R}$. Se dice que X es un espacio vectorial normado o simplemente un espacio normado si existe una función $||\cdot||: X \to \mathbb{R}$, llamada* **norma**, *tal que satisface las siguientes condiciones:*

i) $||\, x \,|| \geq 0$, *para todo* $x \in X$.

ii) $||\, x \,|| = 0$, *si y sólo si* $x = 0$.

iii) $||\, \lambda x \,|| = |\lambda| \, ||\, x \,||$, *para todo* $x \in X$ *y* $\lambda \in \mathbb{R}$.

iv) $||\, x + y \,|| \leq ||\, x \,|| + ||\, y \,||$, *para todo* $x, y \in \mathbb{R}$.

Tomando, $d(x, y) = ||\, x - y \,||$, tenemos que $d : X \times X \to \mathbb{R}$ es una métrica, luego todo espacio normado es un espacio métrico.

Se dice que un espacio métrico X es de Banach si y sólo si, es un espacio métrico completo con la métrica inducida por la norma.

Dado un espacio normado X, denotaremos con X^* al espacio dual de X, es decir,

$$X^* = \{f : X \to \mathbb{R} \,:\, f \quad \text{es una funcional lineal y continua}\}$$

y $\langle \cdot, \cdot \rangle : X^* \times X \to \mathbb{R}$ al producto de dualidad, $\langle x^*, x \rangle = x^*(x)$ para todo $x^* \in X^*$ y $x \in X$.

El espacio dual X^* también es un espacio vectorial normado, con la norma definida por

$$||\, f \,|| = \sup_{\substack{x\in X \\ x\neq 0}} \frac{\langle f, x \rangle}{||\, x \,||}, \quad \text{para todo} \quad f \in X^*.$$

Los siguientes resultados, nos permiten hallar funcionales lineales acotadas con propiedades especiales, la prueba de ellos se puede encontrar en [7].

Teorema 1.2 (Hahn Banach) *Sean X un espacio vectorial real, $p : X \to \mathbb{R}$ una función que verifica*

$$p(\lambda x) = \lambda p(x), \quad \forall \lambda > 0 \quad y \quad \forall x \in X;$$
$$p(x+y) \leq p(x) + p(y), \quad \forall x, y \in X;$$

(la función p que satisface las dos propiedades anteriores, se dice que es una funcional sublineal positivamente homogénea) y T una funcional lineal definida en un espacio Y de X, la cual satisface $T(x) \leq p(x)$, $\forall x \in Y$. Entonces existe una funcional lineal T, definida en X, que satisface

$$T(x) \leq p(x), \forall x \in X;$$
$$T(x) = T(x), \forall x \in Y.$$

Corolario 1.3 *Sea X un espacio normado y $x \in X$. Entonces existe un $x^* \in X^*$, diferente de cero, tal que*

$$\langle x^*, y\rangle = \| x^* \| \| y \|$$

Dado un espacio de Banach X, se pueden definir en él dos topologías. La primera, la topología fuerte, inducida por la norma $\| \cdot \|$, y la segunda que es conocida como topología débil, denotada por $\sigma(X, X^*)$. La topología débil es la topología menos fina sobre X que hace continuas a todas las funcionales lineales $T \in X^*$. La convergencia en ésta topología esta caracterizada de la siguiente manera: una sucesión (x_n) converge débilmente a x (converge en la topología $\sigma(X, X^*)$), denotado por $x_n \overset{*}{\to} x$ si y sólo si

$$\langle T, x_n\rangle \to \langle T, x\rangle, \forall T \in X^*.$$

Todo conjunto cerrado en la topología débil (llamado débilmente cerrado) $\sigma(X, X^*)$ es cerrado en la topología fuerte. En efecto, sea $A \subset X$ débilmente cerrado y sea $(x_n)_{n\in\mathbb{N}} \subset A$ una sucesión tal que $x_n \to x$, entonces tenemos que $x_n \overset{*}{\to} x$, de donde $x \in A$; así A es fuertemente cerrado. Lo recíproco no es cierto en general, como lo muestra el siguiente ejemplo.

Ejemplo 1.2 *Sea E un espacio de Banach de dimensión infinita. La esfera unitaria $S = \{x \in E : \| x \| = 1\}$ no es cerrado en la topología débil. Más aún, la clausura débil de S es la bola unitaria $B = \{x \in B : \| x \| \leq 1\}$. (Ver [7]).*

Para conjuntos convexos, las nociones de cerradura fuerte y débil coinciden como muestra el siguiente teorema, el cual puede ser encontrado en [3].

Teorema 1.4 *Sea $C \subset X$ un conjunto convexo y cerrado. Entonces C es débilmente cerrado si y sólo si es fuertemente cerrado.*

Corolario 1.5 *Sea $f : X \to \mathbb{R}$ una función semicontinua inferiormente (respecto a la topología fuerte), para la cual el conjunto $\{x \in X \,/\, f(x) \leq \lambda\}$, es convexo para cada $\lambda \in \mathbb{R}$. Entonces f es semicontinua inferiormente en la topología débil. En particular, si $x_n \overset{*}{\to} x$, entonces*

$$f(x) \leq \liminf_{n\to\infty} f(x_n)$$

Demostración:

Se sigue de los Teoremas 1.1 y 1.4.

1.3. Derivadas Generalizadas

En esta sección seguiremos el esquema de [6], estudiaremos los gradientes generalizados, algunas de sus propiedades, que serán utilizadas más adelante cuando apliquemos el Principio Variacional de Ekeland a la optimización y al cálculo generalizado. A lo largo de esta sección, X denotará un espacio de Banach.

1.3.1. Gradiente Generalizado

Sea $f : X \to \mathbb{R}$ una función y $a \in \mathbb{R}$. Se dice que f es lipstchitziana en una vecindad de x, si existen $\epsilon > 0$ y $k > 0$ tales que

$$|| \; f(y) - f(y') \; || \leq \; k \; || \; y - y' \; ||, \quad \text{para todo } y, y' \in B_\epsilon(x),$$

donde k es llamada constante de Lipschitz.

Definición 1.5 *Sea f una función lipschitziana en una vecindad de x y $v \in X$. Definimos la derivada direccional generalizada de f en x en la dirección v como el límite*

$$f^\circ(x, v) = \limsup_{y\to x \; t\downarrow 0} \frac{f(y + tv) - f(y)}{t}$$

Notemos que este límite siempre existe en $\mathbb{R} \cup \{+\infty\}$ pues es un límite superior, pero como veremos en la siguiente proposición este límite, entre otras, tiene la propiedad de ser finito. Para la prueba ver [6].

Proposición 1.6 *Sea $f : X \to \mathbb{R}$ una función lipschitziana con constante de Lipschitz k en una vecindad de x. Entonces*

1. *La función $v \to f^\circ(x, v)$ es finita, positivamente homogénea, y subaditiva en X. Además*

$$|f^\circ(x,v)| \;\leq\; k \;||\, v \,||$$

2. *$f^\circ(x, v)$ es semicontinua superiormente como función de (x, v) y lipschitziana con constante k como función solo de v para x fijo.*

3. *$f^\circ(x, -v) = (-f)^\circ(x, v)$, $\forall\, x, v \in X$.*

Por el Teorema 1.2 (Hahn Banach), cualquier funcional subaditiva y positivamente homogénea, mayora alguna funcional lineal en X, luego, bajo las condiciones de la Proposición 1.6, tenemos que existe por lo menos una funcional lineal $x^* \in X^*$ tal que $x^*(v) \;\leq\; f^\circ(x, v)$.
Esta conclusión nos permite definir el gradiente generalizado de f en x, denotado por $\partial f(x)$, como un subconjunto de X^*, dado por

$$\partial f(x) = \{x^* \in X^* \,/\, \langle x^*, v\rangle \leq f^\circ(x, v);\;\; \forall\, v \in X\}.$$

Además, a la multifunción asociada $\partial f : X \rightrightarrows X^*$, la llamaremos subdiferencial de f. Luego, se tienen las siguientes propiedades, que se pueden ver en [6].

Proposición 1.7 *Sea $f : X \to \mathbb{R}$ una función lipschitziana, con constante de Lipschitz k en una vecindad de x. Entonces*

1. *$\partial f(x)$ es un conjunto no vacío, convexo, débilmente compacto en la topología débil estrella, y $||\, x^* \,|| \leq k$ para cada $x^* \in \partial f(x)$.*

2. *Para cada $v \in X$, se tiene que*

$$f^\circ(x, v) = max\{\langle x^*, v\rangle \,/\, x^* \in \partial f(x)\}.$$

Proposición 1.8 *Sea $f : X \to \mathbb{R}$ una función lipschitziana en una vecindad de x. Sean $\{x_n\}$ y $\{x_n^*\}$ sucesiones en X y X^* tales que $x_n^* \in \partial f(x_n)$. Si $x_n \to x$ y x^* es un punto de acumulación en la topología débil estrella de $\{x_n^*\}$, entonces $x \in \partial f(x^*)$.*

Demostración:

Sea $v \in X$, entonces existe una subsucesión de números $\langle x_n^*, v\rangle$ que convergen a $\langle x^*, v\rangle$. Esto junto con $f^\circ(x, v) \geq \langle x_n, v\rangle$ y la semicontinuidad superior de f°, implica que $f^\circ(x, v) \geq \langle x^*, v\rangle$. Como v fue tomado arbitrariamente, concluimos que $x^* \in \partial f(x)$.

Los siguientes resultados proporcionan propiedades para el gradiente generalizado, siendo el más importante el que muestra la relación existente entre el subdiferencial del análisis convexo con el que estamos estudiando, seguiremos el esquema de [6].

Proposición 1.9 *Sea $f : X \to \mathbb{R}$ una función lipschitziana en una vecindad de x, además tiene derivada según Gâteaux $f'(x) \in X^*$. Entonces $f'(x) \in \partial f(x)$.*

Recordemos que el subdiferencial para una función $f : X \to \mathbb{R}$ convexa en un punto $x \in X$, denotado por $\partial f(x)$ esta definido como

$$\partial f(x) = \{x^* \in X^* \,/\, \langle x^*, v\rangle \leq f'(x, v),\ \forall\, v \in X\}$$

Los elementos de $\partial f(x)$ son llamados subgradientes. Ahora establezcamos el siguiente resultado.

Proposición 1.10 *Sea $f : X \to \mathbb{R}$ una función convexa, lipschitziana en una vecindad U de x. Entonces $\partial f(x)$ coincide con el subdiferencial del análisis convexo, y $f^\circ(x, v)$ coincide con la derivada direccional $f'(x, v)$ para cada $v \in X$.*

Ahora, daremos unas propiedades que facilitarán el cálculo de ∂f, las pruebas se pueden ver en [6].

Proposición 1.11 *Sea $f : X \to \mathbb{R}$ una función lipschitziana en una vecindad de x. Entonces*

1. *Para cualquier $\lambda \in \mathbb{R}$: $\partial(\lambda f)(x) = \lambda \partial f(x)$.*
2. *Si f alcanza un mínimo o máximo local en x, entonces $0 \in \partial f(x)$.*
3. $\partial \left(\sum_{i=1}^{r} f_i \right)(x) \subset \sum_{i=1}^{r} \partial f_i(x)$.

1.3.2. Conos Tangente y Normal de Clarke de un conjunto

Ahora estudiaremos algunos conceptos geométricos, relacionados con las definiciones dadas hasta el momento, esto se puede ver en detalle en [6].

Definición 1.6 *Sea C un subconjunto no vacío de X. La función distancia a C es la función $d_C : X \to \mathbb{R}$, definida por*

$$d_C(x) = \text{ínf}\{\| x - c \| : c \in C\}.$$

Observermos que si C es un conjunto cerrado, entonces $x \in C$ si y sólo si $d_C(x) = 0$. Esta función no siempre es diferenciable, pero siempre es globalmente lipschitziana. Usaremos el gradiente generalizado de d_C para definir los conceptos de cono tangente de Clarke, y cono normal de Clarke a un conjunto arbitrario C. Comenzaremos con la siguiente proposición.

Proposición 1.12 *Sea C un subconjunto no vacío de X. La función distancia d_C satisface*

$$|d_C(x) - d_C(y)| \leq \| x - y \|$$

Definición 1.7 *Dados un conjunto C y punto $x \in C$. Decimos que un vector $v \in X$ es tangente a C en x, cuando $d_C^{\circ}(x, v) = 0$. Definimos el cono tangente de Clarke de C en x, denotado por $T_C(x)$ como el conjunto de los vectores tangentes a C en x.*

Notemos que de la Proposición 1.6(2), obtenemos que $T_C(x)$ es un cono convexo cerrado en X.

Definición 1.8 *Se define el cono normal de Clarke de un conjunto C en un punto x, denotado por $N_C(x)$, como el conjunto*

$$N_C(x) = \{x^* \in X^* \ : \ \langle x^*, v\rangle \leq 0, \forall\, v \in T_C(x)\}.$$

El siguiente resultado da una caracterización de $N_C(x)$ en términos de gradientes generalizado.

Proposición 1.13

$$N_C(x) = cl \left(\bigcup_{\lambda > 0} \lambda\, \partial d_C(x) \right),$$

donde cl denota la clausura en la topología débil estrella.

Demostración:

De la definición de $T_C(x)$, junto con la Proposición 1.7(2), se sigue que v pertenece a $T_C(x)$ si y sólo si $\langle x^*, v\rangle \leq 0$ para cada x^* en $\partial d_C(x)$. De esto, se tiene que $N_C(x)$ es el cono convexo, débilmente cerrado en la topología débil estrella generado por $\partial d_C(x)$, lo cual es el enunciado de la proposición.

A continuación daremos unos resultados que serán utilizados en aplicaciones a la optimización.

Proposición 1.14 *Sea f una función lipschitziana con constante de Lipschitz k en un conjunto S. Sea $x \in C \subset S$ y supongamos que f alcanza un mínimo sobre C en x. Entonces para cada $\hat{k} \geq k$, la función $g(y) = f(y) + \hat{k}d_C(y)$ alcanza su mínimo sobre S en x. Si $\hat{k} > k$ y C es cerrado, entonces cualquier otro punto que minimiza g sobre S debe estar en C.*

Corolario 1.15 *Sea f una función lipschitziana en una vecindad de x y que alcanza un mínimo sobre C en x. Entonces*

$$0 \in \partial f(x) + N_C(x).$$

Capítulo 2

EL PRINCIPIO VARIACIONAL DE EKELAND

2.1. El principio variacional de Ekeland

Teorema 2.1 (Principio variacional de Elekand:Forma Fuerte) *([8]) Sean* (X, d) *un espacio métrico completo,* $f : X \to \mathbb{R} \cup \{+\infty\}$ *una función semicontinua inferiormente, acotada inferiormente y propia,* $\epsilon > 0$ *y* $\overline{u} \in X$ *tales que*

$$f(\overline{u}) \leq \inf_X f + \epsilon.$$

Entonces, para cada $k > 0$ *existe* $u = u_k \in X$ *tal que*

$$f(u) \leq f(\overline{u}), \tag{2.1}$$

$$d(u, \overline{u}) \leq \frac{1}{k}, \tag{2.2}$$

$$\forall v \neq u : \ f(u) < f(v) + \epsilon k d(v, u). \tag{2.3}$$

Demostración:

Sean $\epsilon > 0$ y $k > 0$ cualquiera. En vista de que f es acotada inferiormente y propia, se tiene que $-\infty < \inf_X f < +\infty$. Luego, se puede asegurar que existe $\overline{u} \in X$ tal que

$$f(\overline{u}) \leq \inf_X f + \epsilon \quad \text{(definición de ínfimo).}$$

Vamos a definir de manera inductiva una sucesión (u_n) en X, que será convergente y cuyo límite será el elemento u en X que buscamos.

Tomemos $u_1 = \overline{u} \in X$ que verifica las condiciones (2.1) y (2.2). Supongamos que $u_n \in X$ es conocido, que verifica las condiciones (2.1) y (2.2). A partir de este elemento, definiremos u_{n+1} considerando el conjunto

$$S_n = \{w \in X \,/\, w \neq u_n \quad \text{y} \quad f(w) + \epsilon\, k\, d(w, u_n) \leq f(u_n)\}.$$

Distinguimos dos casos:

Si $S_n = \emptyset$, es decir, si para todo $w \neq u_n$, $f(u_n) < f(w) + \epsilon k d(w, u_n)$. Elegimos $u_{n+1} = u_n$, en cuyo caso será una sucesión constante a partir de n, luego convergente. Tomando $u = u_n$, tendremos que se verifica claramente (2.1), (2.2) y (2.3).

Si $S_n \neq \emptyset$, es decir, si existe $w \neq u_n$ tal que $f(w) + \epsilon k d(w, u_n) \leq f(u_n)$, entonces

$$\inf_{S_n} f \leq f(w) \leq f(u_n) - \epsilon\, k\, d(w, u_n) < f(u_n).$$

Luego $\inf_{S_n} f < f(u_n)$, y se sigue que $\frac{1}{2}\left[f(u_n) - \inf_{S_n} f\right] > 0$.
Así, podemos elegir $u_{n+1} \in S_n$ (debido que existe) tal que

$$\inf_{S_n} f \leq f(u_{n+1}) \leq \inf_{S_n} f + \frac{1}{2}\left[f(u_n) - \inf_{S_n} f\right]. \tag{2.4}$$

A partir de ahora, consideraremos el caso $S_n \neq \emptyset$, pues en caso contrario, como vimos (u_n) es constante a partir de algún n, por lo tanto convergente, y se verificaba el teorema.

Veamos que (u_n) en el caso a considerarse, es convergente:

En efecto, desde que $u_{n+1} \in S_n$, entonces

$$u_{n+1} \neq u_n \quad \text{y} \quad \inf_{S_n} f \leq f(u_{n+1}) < f(u_n).$$

Luego, la sucesión $(f(u_n))$ es decreciente y acotada inferiormente, por lo tanto, convergente. Sea $l = \lim_{n\to\infty} f(u_n)$, de la definición de S_n, es claro que

$$\epsilon k d(u_{n+1}, u_n) \leq f(u_n) - f(u_{n+1}), \quad \forall n \in \mathbb{N}$$

Tomemos $n, m \in \mathbb{N}$ con $n \leq m$. Luego

$$\begin{aligned}
\epsilon k d(u_n, u_m) &\leq \epsilon\, k \sum_{i=n}^{m-1} d(u_i, u_{i+1}) \\
&= \sum_{i=n}^{m-1} \epsilon k d(u_i, u_{i+1}) \\
&\leq \sum_{i=n}^{m-1} [f(u_i) - f(u_{i+1})] = f(u_n) - f(u_m)
\end{aligned} \tag{2.5}$$

y como $(f(u_n))$ es convergente, (u_n) es una sucesión de Cauchy y en vista de que X es un espacio completo, se concluye que (u_n) es convergente. Sea

$$u = \lim_{n\to\infty} u_n$$

Veamos que se cumple (2.1):

En efecto, como f es semicontinua inferiormente y $\lim_{n\to\infty} u_n = u$, entonces se tiene que $f(u) \leq \liminf_{n\to\infty} f(u_n) = \lim_{n\to\infty} f(u_n) = l$, y se sigue que

$$f(u) \leq l. \tag{2.6}$$

Por otro lado $f(\overline{u}) = f(u_1) \geq f(u_n)$, $\forall n \in \mathbb{N}$. Luego $f(\overline{u}) \geq \lim_{n\to\infty} f(u_n) = l$. De aquí y de (2.6) se concluye que

$$f(u) \leq f(\overline{u}).$$

Veamos que se cumple (2.2):

En efecto, tomemos en (2.5), n = 1:

$$\epsilon k d(u_1, u_m) \leq f(u_1) - f(u_m), \quad \forall m \in \mathbb{N}.$$

Se sigue que

$$\begin{aligned} \epsilon\, k\, d(\overline{u}, u_m) &\leq f(\overline{u}) - f(u_m), \\ &\leq f(\overline{u}) - \inf_X f, \\ &\leq \epsilon. \end{aligned}$$

Se concluye que

$$d(\overline{u}, u_m) \leq \frac{1}{k}, \quad \forall m \in \mathbb{N}.$$

Haciendo $m \to \infty$ y sabiendo que la función distancia a un punto es continua, podemos concluir que

$$d(u, \overline{u}) \leq \frac{1}{k}$$

Veamos que se cumple (2.3):

En efecto, por contradicción, supongamos que (2.3) no se cumple, es decir, existe $v \neq u$ tal que

$$f(u) \geq f(v) + \epsilon k d(v, u). \tag{2.7}$$

Luego, haciendo $m \to \infty$ en (2.5), obtenemos

$$\epsilon k d(u_n, u) \leq f(u_n) - l. \tag{2.8}$$

De este resultado, de (2.6) y (2.7) se tiene que

$$\begin{aligned} f(v) &\leq f(u) - \epsilon k d(v,u) \\ &< l - \epsilon k d(v,u) \\ &\leq f(u_n) - \epsilon k d(u_n,u) - \epsilon k d(v,u) \\ &\leq f(u_n) - \epsilon\, k\, d(u_n,v). \end{aligned}$$

Además, como $u = \lim\limits_{n\to\infty} u_n$, podemos asegurar que existe un $N \in \mathbb{N}$ suficientemente grande tal que para todo $n \geq N, v \neq u_n$. Luego, se sigue que para todo $n \geq N$, $v \in S_n$. De aquí y de (2.4) se tiene que

$$2f(u_{n+1}) - f(u_n) \leq \inf_{S_n} f \leq f(v) \leq f(u_n) - \epsilon k d(u_n, v), \quad \forall n \geq N.$$

De aquí se sigue que

$$2(f(u_{n+1}) - f(u_n)) \leq -\epsilon k d(u_n, v),$$

haciendo $n \to \infty$ obtenemos que

$$\epsilon k d(u_n, v) \leq 0,$$

lo cual es falso, pues $\epsilon\, k\, d(u_n, v) > 0$. Luego, podemos concluir que se cumple (2.3).

Corolario 2.2 *Con las mismas hipótesis del teorema anterior, se concluye que, existe $u \in X$ tal que*

$$\begin{aligned} f(u) &\leq f(\overline{u}) \\ d(u,\overline{u}) &\leq \sqrt{\epsilon} \\ \forall\, v \neq u : f(v) &> f(u) - \sqrt{\epsilon}\, d(v,u) \end{aligned}$$

Demostración:

Basta tomar en el Teorema 2.1, $k = \dfrac{1}{\sqrt{\epsilon}} > 0.$

Corolario 2.3 (Principio variacional de Ekeland: Forma Débil) *Sean (X, d) un espacio métrico completo, $f : X \to \mathbb{R} \cup \{+\infty\}$ una función semicontinua inferiormente, acotada inferiormente y propia. Entonces, para cada $\epsilon > 0$ existe $u = u_\epsilon \in X$ tal que*

$$f(u) \leq \inf_X f + \epsilon, \tag{2.9}$$

$$\forall\, v \neq u : f(u) < f(v) + \epsilon\, d(v,u). \tag{2.10}$$

Demostración:

Sea $\epsilon > 0$ cualquiera. Como f es acotada inferiormente y propia, $-\infty < \inf f < +\infty$. Luego, por la definición de ínfimo, existe algún $\overline{u} \in X$ tal que

$$f(\overline{u}) \leq \inf_X f + \epsilon$$

Tomando este $\epsilon > 0$ y haciendo $k = 1$ en el Teorema 2.1 se asegura la existencia de $u = u_\epsilon \in X$ tal que

$$f(u) \leq f(\overline{u}) \leq \inf_X f + \epsilon$$

De aquí se verifica (2.9). Además

$$\forall\, v \neq u : f(u) < f(v) + \epsilon k d(v, u).$$

Lo cual verifica (2.10).

2.2. Aplicaciones del principio variacional de Ekeland en su forma débil a la teoría de puntos fijos

Teorema 2.4 (Punto Fijo de Caristi) *([5], [8]) Sean (X, d) un espacio métrico completo, una función $\phi : X \to \mathbb{R}$ semicontinua inferiormente y acotada inferiormente, $F : X \to X$ una función que satisface*

$$\forall\, w \in X : d(w, F(w)) \leq \phi(w) - \phi(F(w)). \tag{2.11}$$

Entonces, F tiene un punto fijo, es decir,

$$\exists w_0 \in X \quad \textit{tal que} \quad F(w_0) = w_0.$$

Demostración:

Del Corolario 2.3, tomando $\epsilon = 1$ existe $u \in X$ tal que

$$\forall\, v \neq u \; : \; \phi(u) < \phi(v) + d(v, u).$$

Tomemos $w_0 = u$, entonces

$$\phi(w_0) < \phi(v) + d(v, w_0).$$

Afirmamos que $F(w_0) = w_0$, pues si no fuera cierto, se tendría

$$\phi(w_0) < \phi(F(w_0)) + d(F(w_0), w_0),$$

pues $F(w_0) \neq w_0$. Entonces

$$d(w_0, F(w_0)) > \phi(w_0) - \phi(F(w_0)).$$

Lo cual sería una contradicción con la condición (2.11). Luego, se debe cumplir que $F(w_0) = w_0$. ■

Observación:

Sean (X, d) un espacio métrico completo, $T : X \to X$ una contracción, es decir, $\exists k \in [0, 1\rangle$ tal que

$$d(T(u), T(v)) \leq kd(u, v) \quad \forall u, v \in X.$$

Probemos que T tiene un punto fijo.

En efecto, usaremos el Teorema 2.4, para esto definamos:

$$\begin{aligned} \phi : X &\to \mathbb{R} \\ x &\mapsto \phi(x) = \frac{1}{1-k} d(x, T(x)) \end{aligned}$$

Notemos que ϕ es semicontinua inferiormente y acotada, debido a que es continua y no negativa. Sea ahora un $w \in X$ cualesquiera,

$$\begin{aligned} d(w, T(w)) - \phi(w) + \phi(T(w)) &= \\ &= d(w, T(w)) - \frac{1}{1-k} d(w, T(w)) + \frac{1}{1-k} d(T(w), T^2(w)) \\ &\leq d(w, T(w)) - \frac{1}{1-k} d(w, T(w)) + \frac{k}{1-k} d(w, T(w)) = 0 \end{aligned}$$

Esto muestra que T satisface la condición (2.11), por lo tanto, existe $w_0 \in X$ tal que $T(w_0) = w_0$. ■

Teorema 2.5 (Punto fijo de Caristi para operadores multivaluados) *([5], [8]) Sean (X, d) un espacio métrico completo, $\phi : X \to \mathbb{R}$ una función semicontinua inferiormente y acotada inferiormente, $\Gamma : X \to 2^X$ un operador multivaluado tal que para cada $u \in X$ se tiene que $\Gamma(u) \neq \emptyset$ y para cada $w \in \Gamma(u)$:*

$$d(u, w) \leq \phi(u) - \phi(w). \tag{2.12}$$

Entonces, existe $u_0 \in X$ tal que $\Gamma(u_0) = \{u_0\}$.

Demostración:

Del Corolario 2.3, tomando $\epsilon = 1$, se tiene que existe $\overline{u} \in X$ tal que

$$\forall v \neq \overline{u} : \phi(\overline{u}) < \phi(v) + d(v, \overline{u}).$$

Tomemos $u_0 = \overline{u}$, entonces

$$\phi(u_0) < \phi(v) + d(v, u_0) \quad \forall\, v \neq u_0. \tag{2.13}$$

Por otro lado, como $\Gamma(u_0) \neq \emptyset$, debe existir un $w \in \Gamma(u_0)$. Supongamos $w \neq u_0$, de (2.12) se tiene que

$$d(u_0, w) \leq \phi(u_0) - \phi(w),$$

y de (2.13)

$$\phi(u_0) - \phi(w) < d(u_0, w).$$

En conclusión, $d(u_0, w) < d(u_0, w)$ lo cual es una contradicción. Por lo tanto, $w = u_0$ y como para cualquier otro $w \neq u_0$ se llega a una contradicción, luego, se debe cumplir $\Gamma(u_0) = \{u_0\}$.

Teorema 2.6 (Punto Fijo según Clarke) *([5], [8]) Sean $(X, ||\ .\ ||)$ un espacio de Banach, V un subconjunto convexo y cerrado de X y $f : V \to V$ una función continua satisfaciendo que existe $\delta \in \langle 0, 1\rangle$ tal que para todo $u \in V$, $\exists\, t \in \langle 0, 1]$:*

$$|| f(u_t) - f(u) || \leq \delta \ || u_t - u || \tag{2.14}$$

donde $u_t = tf(u) + (1 - t)u$. Entonces, f tiene un punto fijo.

Demostración:

Definamos la función $F : V \to \mathbb{R}$, dada por

$$F(w) = || w - f(w) ||$$

la cual será continua en el conjunto cerrado V pues f es continua y además acotado inferiormente. Luego, por el Corolario 2.3, tomando $0 < \epsilon < 1 - \delta$, existe $v \in V$, tal que

$$F(v) < F(w) + \epsilon d(w, v) \quad \forall\, v \neq w.$$

Equivalentemente, $\forall\, w \in V : F(v) \leq F(w) + \epsilon d(w, v)$ de donde

$$|| v - f(v) \ | \leq || w - f(w) || + \epsilon \ | \ w - v || \,.$$

Tomando $w = v_t = tf(v) + (1 - t)v \in V$, pues V es convexo, para algún $t \in \langle 0, 1]$, se sigue que

$$\begin{aligned}
\| v - f(v) \| &\leq \| v_t - f(v_t) \| + \epsilon \| v_t - v \| \\
&\leq \| v_t - f(v) \| + \| f(v) - f(v_t) \| + \epsilon \| v_t - v \| \\
&\leq \| v_t - f(v) \| + \delta \| v_t - v \| + \epsilon \| v_t - v \| \\
&= \| v_t - f(v) \| + (\delta + \epsilon) \| v_t - v \| \\
&= \| tf(v) + (1 - t)v - f(v) \| + (\delta + \epsilon) \| tf(v) + (1 - t)v - v \| \\
&= \| (1 - t)v - (1 - t)f(v) \| + (\delta + \epsilon) \| t(f(v) - v) \| \\
&= (1 - t) \| v - f(v) \| + (\delta + \epsilon)t \| v - f(v) \|
\end{aligned}$$

Entonces $\quad t \| v - f(v) \| \leq (\delta + \epsilon)t \| v - f(v) \|, \quad$ y en consecuencia

$$\| v - f(v) \| \leq (\delta + \epsilon) \| v - f(v) \|$$

Como $\delta + \epsilon < 1$, entonces $\| v - f(v) \| = 0$ y por lo tanto $v = f(v)$. Así, se concluye que f tiene un punto fijo.

2.3. Existencia de Soluciones

Definición 2.1 *Sean X un espacio de Banach, f una función. Se dice que f es Gâteaux diferenciable si para cada $v \in X$ existe una funcional lineal continua $f'(v) \in X^*$ tal que*

$$\forall u \in X \ : \ \lim_{t \to 0^+} \frac{f(v + tu) - f(v)}{t} = \langle f'(v), u \rangle.$$

Teorema 2.7 *([8]) Sean X un espacio de Banach, f una función Gâteaux diferenciable sobre X, semicontinua inferiormente y acotada inferiormente. Entonces, para cada $\epsilon > 0$, existe un punto $v \in X$ tal que*

$$f(v) \leq \inf_X f + \epsilon \qquad (2.15)$$

$$\| f'(v) \|^* \leq \epsilon. \qquad (2.16)$$

Demostración:

Sea $\epsilon > 0$ cualquiera. Del Corolario 2.3, existe $v \in X$ tal que

$$f(v) \leq \inf_X f + \epsilon,$$

lo cual prueba (2.15). Además, tomando $d(w, v) = \| w - v \|$ se tiene que

$$\forall w \neq v \ : \ f(v) < f(w) + \epsilon \ \| w - v \| \ .$$

Esta última relación también se puede escribir

$$\forall w \in X \ : \ f(v) \leq f(w) + \epsilon \ || \ w - v \ || \ . \tag{2.17}$$

Luego, tomemos $x \in X$ y $t > 0$ cualesquiera, y en particular $w = v - tx \in X$, reemplazando en (2.17)

$$f(v) \leq f(v - tx) + \epsilon \ || -tx \ ||$$

Luego, $\dfrac{f(v + t(-x)) - f(v)}{t} \geq -\epsilon \ || \ x \ ||.$

De aquí, $\displaystyle\lim_{t \to 0^+} \left[\frac{f(v + t(-x)) - f(v)}{t}\right] \geq -\epsilon \ || \ x \ ||,$ y como f es Gâteaux diferenciable

$$\langle f'(v), -x \rangle \geq -\epsilon \ || \ x \ ||, \quad \text{se sigue que}$$

$$\frac{\langle f'(v), x \rangle}{|| \ x \ ||} \leq \epsilon; \ x \neq 0,$$

luego

$$\frac{|\langle f'(v), x \rangle|}{|| \ x \ ||} \leq \epsilon,$$

de lo cual,

$$\sup_{\substack{x \in X \\ x \neq 0}} \frac{|\langle f'(v), x \rangle|}{|| \ x \ ||} \leq \epsilon$$

y se concluye que

$$|| \ f'(v) \ ||^* \leq \epsilon.$$

Finalmente, esto prueba (2.16). ∎

Definición 2.2 *([5], [8]) Sean X un espacio de Banach, $f : X \to \mathbb{R}$ una función de clase C^1. Se dice que f satisface la condición Palais-Smale si y sólo si para toda sucesión $(u_n) \subset X$ con $f'(u_n) \to 0$ en X^* y $f(u_n)$ acotada, se tiene que, ó $f'(u_n) = 0$ para algún n ó la sucesión (u_n) tiene un punto límite en X.*

Teorema 2.8 *([8]) Sea X un espacio de Banach, f una función de clase C^1, acotada inferiormente y satisface la condición de Palais-Smale. Entonces, la función f alcanza su mínimo, es decir*

$$\exists \, v \in X \quad \textit{tal que} \quad f(v) = \inf_X f \quad y \quad f'(v) = 0.$$

Demostración:

Sea $n \in \mathbb{N}$ cualquiera. Tomando $\epsilon = \frac{1}{n} > 0$, en el Teorema 2.7, existe $u_n \in X$ tal que

$$f(u_n) \leq \inf_X f + \frac{1}{n}$$

$$\| f'(u_n) \|^* \leq \frac{1}{n}$$

Luego, cuando $n \to \infty$

$$f(u_n) \to \inf_X f$$

$$\| f'(u_n) \|^* \to 0, \quad \text{de aquí} \quad f'(u_n) \to 0$$

concluimos que existe $(u_n) \subset X$ tal que $f(u_n)$ es acotado, pues es convergente y $f'(u_n) \to 0$.

Del segundo resultado, o podemos extraer una subsucesión (u_{n_k}) tal que $f'(u_{n_k}) \neq 0$ para todo k, ó $f'(u_n) = 0$ para todos excepto un número finito de n.

En el primer caso, $f'(u_{n_k}) \neq 0$ para todo $k \in \mathbb{N}$, luego por la condición de Palais-Smale, la sucesión (u_{n_k}) tiene un punto límite $v \in X$, hacia el cual una subsucesión converge. Por continuidad de f, se tiene que

$$f(v) = f(\lim_{k\to\infty} u_{n_k}) = \lim_{k\to\infty} f(u_{n_k}) = \inf_X f \quad \text{y además} \quad f'(v) = 0.$$

En el segundo caso, sea

$$\mathcal{S} = \{v \in X \,/\, f'(v) = 0\}$$

Para cada $n \in \mathbb{N}$, existe un camino C^1 en $\mathcal{S}$ que conecta cada u_n con algún punto v_n que está en la frontera de $\mathcal{S}$. Por el teorema del valor medio, $f(u_n) = f(v_n)$ y desde que v_n está en la frontera de $\mathcal{S}$, existe algún $w_n \notin \mathcal{S}$ tal que

$$|f(v_n) - f(w_n)| \leq \frac{1}{n}$$

como $f'(w_n) \neq 0$ y $f(w_n) \to \inf_X f$, esto nos conduce al primer caso. Con esto concluimos la prueba. ∎

Teorema 2.9 *([5]) Sean X un espacio de Banach, $f : X \to \mathbb{R}$ una función Gâteaux diferenciable sobre X, semicontinua inferiormente y acotada inferiormente. Además $k > 0, c \in \mathbb{R}$ son constantes tales que*

$$\forall x \in X : f(x) \geq k \| x \| + c,$$

donde

$$\| x \| \geq r \quad \textit{para algún } r > 0 \tag{2.18}$$

$$B^* = \{x^* \in X^* \,/\, \| x^* \| \leq 1\}$$

Entonces, $f'(X)$ es denso en kB^*.

Demostración

Debemos demostrar que $\overline{f'(X)} = kB^*$. Es decir, dados $\epsilon > 0$ y $x^* \in kB^*$, debemos probar que existe $u_\epsilon \in X$ tal que

$$\| f'(u_\epsilon) - x^* \| < \epsilon.$$

En efecto, definamos la función

$$g : X \to \mathbb{R}$$
$$g(x) = f(x) - \langle x^*, x\rangle.$$

Se tiene que para cualquier $u \in X$

$$\begin{aligned}
\lim_{t\to 0^+} \frac{g(x+tu) - g(x)}{t} &= \lim_{t\to 0^+} \frac{f(x+tu) - \langle x^*, x+tu\rangle - [f(x) - \langle x^*, x\rangle]}{t} \\
&= \lim_{t\to 0^+} \frac{f(x+tu) - f(x) - \langle x^*, tu\rangle}{t} \\
&= \lim_{t\to 0^+} \frac{f(x+tu) - f(x)}{t} - \langle x^*, u\rangle \\
&= \langle f'(x), u\rangle - \langle x^*, u\rangle
\end{aligned}$$

lo cual demuestra que g es Gâteaux diferenciable y

$$g'(x) = f'(x) - x^*.$$

Claramente g es acotada inferiormente en $\{x \in X \,/\, \| x \| < r\}$ y por (2.18) g es acotada inferiormente en $\{x \in X \,/\, \| x \| \geq r\}$. Luego es acotado inferiormente sobre X.

Así, aplicando el Teorema 2.7, existe $u_\epsilon \in X$ tal que

$$\begin{aligned}
\| g^*(u_\epsilon) \|^* &\leq \epsilon, \quad \text{de aquí} \\
\| f'(u_\epsilon) - x^* \| &\leq \epsilon.
\end{aligned}$$

Concluimos que $f'(X)$ es denso en kB^*. ■

Teorema 2.10 *([5], [8]) Sean X un espacio de Banach, $f : X \to \mathbb{R}$ una función Gâteaux diferenciable sobre X y semicontinua inferiormente. $\phi : [0, +\infty\rangle \to \mathbb{R}$ una función continua, tal que*

$$\forall x \in X : f(x) \geq \phi(\| x \|), \quad y$$
$$\frac{\phi(t)}{t} \to \infty \quad cuando \quad t \to \infty.$$

Entonces $f'(X)$ es denso en X^*.

Demostración:

Sea $k > 0$ cualquiera. Como $\dfrac{\phi(t)}{t} \to \infty$ cuando $t \to \infty$, existe $N > 0$ tal que $\phi(t) \geq kt$ siempre que $t \geq N$. Tomando $t = \| x \|$, se tiene que $\phi(\| x \|) \geq k \| x \|$, siempre que $\| x \| \geq N$. De aquí

$$f(x) \geq \phi(\| x \|) \geq k \| x \|, \quad \| x \| \geq N$$

Aplicando el Teorema 2.9, se tiene que $f'(X)$ es denso en $kB^* \subset X^*$, y como $k > 0$ fue tomado arbitrariamente, concluimos que $f'(X)$ es denso en X^*. ■

Capítulo 3

APLICACIONES DE LA FORMA FUERTE DEL PRINCIPIO VARIACIONAL DE EKELAND

3.1. Programación Matemática

Proposición 3.1 *([8]) Sea X un espacio de Banach, $f : X \to \mathbb{R}$ una función semicontinua inferiormente sobre X. Supongamos que para cada $\epsilon > 0$, existe una función f_ϵ Gâteaux diferenciable tal que*

$$f_\epsilon \leq f, \tag{3.1}$$

$$\inf_X f_\epsilon \geq \inf_X f - \epsilon, \tag{3.2}$$

$$f'_\epsilon(v) \to \phi(u), \quad \textit{cuando} \quad \epsilon \to 0 \quad y \quad v \to u. \tag{3.3}$$

Entonces, cualquier punto $u \in X$ el cual minimiza f también satisface la ecuación

$$\phi(u) = 0 \quad \textit{en} \quad X^*. \tag{3.4}$$

Demostración:

Sea u el punto que minimiza f sobre X. Luego, de (3.1) y (3.2) se tiene que

$$f_\epsilon(u) \leq f(u) = \inf_X f$$

y como $\inf_X f \leq \inf_X f_\epsilon + \epsilon$ se concluye que

$$f_\epsilon(u) \leq \inf_X f_\epsilon + \epsilon$$

De aquí y del Corolario 2.2 se sigue que existe $v_\epsilon \in X$ tal que $\| u - v_\epsilon \| \leq \sqrt{\epsilon}$ y que $f_\epsilon(w) > f_\epsilon(v_\epsilon) - \sqrt{\epsilon} \| w - v_\epsilon \|$, para todo $w \neq v_\epsilon$. De la última desigualdad tenemos que v_ϵ minimiza a la función $g_\epsilon(w) = f_\epsilon(w) + \sqrt{\epsilon} \| w - v_\epsilon \|$ sobre X. Ahora la función f_ϵ es claramente diferenciable en v_ϵ y aunque la función $w \to \| w - v_\epsilon \|$ no es diferenciable en v_ϵ, tiene derivada direccional +1 en cada dirección. Se sigue entonces que

$$0 \in f'_\epsilon(v_\epsilon) + \sqrt{\epsilon} B^* \tag{3.5}$$

donde $B^* = \{x^* \in X^* / \| x^* \| \leq 1\}$. Luego, en (3.5) hacemos que ϵ tienda a cero, así v_ϵ converge a u y de (3.3), $f'_\epsilon(v_\epsilon)$ converge a $\phi(u)$. Se sigue que el conjunto en (3.5) converge a $\{\phi(u)\}$ y como 0 pertenece a dicho conjunto, concluimos que

$$\phi(u) = 0.$$

Consideremos ahora el problema de minimizar $f(x)$ sujeta a las siguientes restricciones de igualdad y desigualdad

$$\begin{aligned} g_i(x) &= 0, \quad 1 \leq i \leq r \\ h_j(x) &\leq 0, \quad 1 \leq j \leq k \end{aligned}$$

donde f, los g_i y los h_j son localmente Lipschitzianos sobre $\mathbb{R}^n$.
Establecemos el siguiente teorema.

Teorema 3.2 *([8]) Si $\overline{x} \in \mathbb{R}^n$ minimiza a la función f sujeta a las restricciones precedentes, entonces existen números reales $\lambda, \alpha_i, \beta_j$, no todos nulos, con $\beta_j \geq 0$ para todo j, y $\lambda \geq 0$ tales que*

$$\lambda \partial f(\overline{x}) + \sum_{i=1}^{r} \alpha_i \partial g_i(\overline{x}) + \sum_{j=1}^{k} \beta_j \partial h_j(\overline{x}) \ni 0.$$

Si todas las funciones son continuamente diferenciables, esta relación se puede escribir

$$\lambda f'(\overline{x}) + \sum_{i=1}^{r} \alpha_i g'_i(\overline{x}) + \sum_{j=1}^{k} \beta_j h'_j(\overline{x}) = 0\,.$$

Si además otras condiciones de regularidad se cumplen, por ejemplo que los $g'_i(\overline{x})$ y los $h'_j(\overline{x})$ para $h_j(\overline{x}) = 0$ sean linealmente independientes, entonces $\lambda \neq 0$; luego dividiendo por λ obtenemos la usual Regla de Multiplicadores de Lagrange.

Demostración:

Definamos la función $F : \mathbb{R}^n \to \mathbb{R}$, mediante

$$F(x) = \max_{i,j}\{f(x) - f(\overline{x}), |g_i(x)|, h_j(x)\}.$$

Notemos que $\overline{x}$ minimiza dicha función, que además es localmente Lipschitziana (pero no es diferenciable, aún si las funciones envueltas en la maximización son continuamente diferenciables). Usando cálculo generalizado, escribimos que 0 debería pertenecer al gradiente generalizado de F en $\overline{x}$, es decir

$$0 \in \partial F(\overline{x}).$$

Dicho gradiente puede ser visto además como la envoltura convexa de $\partial f(\overline{x})$, $\partial|g_i(\overline{x})|$ y $\partial h_j(\overline{x})$ para $h_j(\overline{x}) = 0$. Esto significa que se pueden hallar $\lambda, \alpha_i, \beta_j$ números no negativos tales que

$$\lambda\partial f(\overline{x}) + \sum_{i=1}^{r} \alpha_i \partial|g_i(\overline{x})| + \sum_{j=1}^{k} \beta_j \partial h_j(\overline{x}) = 0,$$

donde $\lambda + \sum_{i=1}^{r} \alpha_i + \sum_{j=1}^{k} \beta_j = 1.$

Pero, en la igualdad anterior $\partial|g_i(\overline{x})|$ está expresada como $\bigcup_{|t|\leq 1} t\partial g_i(\overline{x})$, de manera que dicha relación se puede expresar así

$$\lambda\partial f(\overline{x}) + \sum_{i=1}^{r} \alpha_i t_i \partial g_i(\overline{x}) + \sum_{j=1}^{k} \beta_j \partial h_j(\overline{x}) = 0,$$

donde $\quad \lambda + \sum_{i=1}^{r} \alpha_i + \sum_{j=1}^{k} \beta_j = 1 \quad \text{y} \quad -1 \leq t_i \leq 1.$

Luego, la solución trivial aparece tomando $\lambda = 0$, $\beta_j = 0$, $\alpha_i \neq 0$, pero $t_i = 0$; lo cual no se puede evitar.

Ahora, dado $\epsilon > 0$ cualquiera, consideremos la siguiente función

$$F_\epsilon(x) = \max_{i,j}\{f(x) - f(\overline{x}) + \epsilon, |g_i(x)|, h_j(x)\}.$$

Notemos que no necesariamente $\overline{x}$ minimiza dicha función, pero si se cumple claramente que

$$F_\epsilon(\overline{x}) = \epsilon \leq \inf_X F_\epsilon(x) + \epsilon,$$

luego, aplicando el Corolario 2.2, existe $x_\epsilon \in \mathbb{R}^n$ tal que

$$\| \overline{x} - x_\epsilon \| \leq \sqrt{\epsilon}$$

$$F_\epsilon(x_\epsilon) \leq F_\epsilon(x) + \sqrt{\epsilon} \mid\mid x - x_\epsilon \mid\mid, \quad \forall x \in \mathbb{R}^n.$$

Lo último significa que x_ϵ minimiza a la función $G_\epsilon(x) = F_\epsilon(x) + \sqrt{\epsilon} \mid\mid x - x_\epsilon \mid\mid$ sobre $\mathbb{R}^n$, esto implica que

$$0 \in \partial G_\epsilon(x_\epsilon) = \partial F_\epsilon(x_\epsilon) + \sqrt{\epsilon} B$$

donde B es la bola unitaria de $\mathbb{R}^n$. Luego, vemos en la expresión anterior que $\partial F_\epsilon(x_\epsilon)$ ha sido calculada para ser la envoltura convexa de las gradientes generalizadas de alguna de las funciones $f, |g_i(x)|, h_j$; cuyos valores en x_ϵ son exactamente $F_\epsilon(x_\epsilon)$. Pero debe ser que $F_\epsilon(x_\epsilon) > 0$, de otro modo x_ϵ satisfacería todas las restricciones y $f(x_\epsilon)$ sería estrictamente menor que $f(\overline{x})$, lo cual es una contradicción pues $\overline{x}$ es el punto que minimiza f. Se sigue entonces que tendremos que considerar $\partial|g_i(x_\epsilon)|$ solo para aquellos i donde se satisface $g_i(x_\epsilon) \neq 0$.

Pero notemos que la aplicación $t \to |t|$ es continuamente diferenciable en $\mathbb{R}\backslash\{0\}$, cuya derivada es la función $sgn(t)$, de manera que, aplicando la regla de la cadena se tiene que

$$\partial|g_i(x_\epsilon)| = [sgn\, g_i(x_\epsilon)]\partial g_i(x_\epsilon).$$

Finalmente, tendremos que

$$\lambda_\epsilon \partial f(x_\epsilon) + \sum_{i=1}^{r} \alpha_{\epsilon,i}[sgn g_i(x_\epsilon)]\partial g_i(x_\epsilon) + \sum_{j=1}^{k} \beta_{\epsilon,j} \partial h_j(x_\epsilon) \ni 0.$$

donde $\lambda_\epsilon \geq 0, \alpha_{\epsilon,i} \geq 0, \beta_{\epsilon,j} \geq 0$ y $\lambda_\epsilon + \sum_{i=1}^{r} \alpha_{\epsilon,i} + \sum_{j=1}^{k} \beta_{\epsilon,j} = 1$.

Ahora, hacemos $\epsilon \to 0$ y por compacidad, el vector $(k+r)-$dimensional $(\lambda_\epsilon, \alpha_{\epsilon,i}, \beta_{\epsilon,j})$ tiene un punto límite, afirmamos que sea $(\lambda, \alpha_i, \beta_j)$.

Sean ahora $u_\epsilon \in \partial f(x_\epsilon), v_\epsilon \in \partial g_i(x_\epsilon), w_\epsilon \in \partial h_j(x_\epsilon)$ tales que

$$\lambda_\epsilon u_\epsilon + \sum_{i=1}^{r} \alpha_{\epsilon,i} v_{\epsilon,i} + \sum_{j=1}^{k} \beta_{\epsilon,j} w_{\epsilon,j} = 0.$$

Luego, de propiedades conocidas del gradiente generalizado, se sigue que $(u_\epsilon, v_{\epsilon,i}, w_{\epsilon,j})$ tiene un punto límite, el cual es (u, v_i, w_j) donde $u \in \partial f(\overline{x}), v_i \in \partial g_i(\overline{x})$ y $w_j \in \partial h_j(\overline{x})$.

En el límite, tendremos que

$$\lambda u + \sum_{i=1}^{r} (\pm \alpha_i) v_i + \sum_{j=1}^{k} \beta_j w_j = 0$$

donde $\lambda \geq 0, \alpha_i \geq 0, \beta_j \geq 0$ y $\lambda + \sum_{i=1}^{r} \alpha_i + \sum_{j=1}^{k} \beta_j = 1$.

Con lo cual concluimos que existen números no negativos $\lambda, \alpha_i, \beta_j$ no todos nulos tales que

$$0 \in \lambda\partial f(\overline{x}) + \sum_{i=1}^{r} \alpha_i \partial g_i(\overline{x}) + \sum_{j=1}^{k} \beta_j \partial h_j(\overline{x}),$$

que es el resultado del teorema. ■

3.2. Cálculo

En esta sección, a no ser que se indique lo contrario, se establece lo siguiente: X un espacio de Banach, $f : X \to \mathbb{R}$ una función localmente Lipschitziana, $\partial f(u) \subset X^*$ el gradiente generalizado de f en u, $S \subset X$ un subconjunto cerrado, $N_S(u)$ el cono normal a S en u y finalmente $d_S(u)$ la distancia de u a S. Además se cumple la siguiente relación

$$N_S(u) = \bigcup_{\lambda > 0} \lambda\, \partial d_S(u), \quad \forall\, u \in S.$$

Teorema 3.3 *([8]) Sean $E = \{w \in S \,/\, f(w) = 0\}$ un subconjunto cerrado de S. Entonces, para cada $u \in S \setminus E$ y $a \in \langle 0, 1]$, existe algún $v \in S$ tal que*

$$\| v - u \| \leq a\, d_E(u)$$

$$\exists v_1^* \in \partial |f(v)| \quad y \quad v_2^* \in N_S(v) \quad con \quad \| v_1^* + v_2^* \|^* \leq \frac{|f(u)|}{a\, d_E(u)}$$

Demostración:

Sean $u \in S \setminus E$ y $a \in \langle 0, 1]$. Definamos la función:

$$F : X \to \mathbb{R} \cup \{+\infty\}$$

$$F(x) = \begin{cases} |f(x)|, & x \in S \\ +\infty, & x \notin S \end{cases}$$

Notemos que F es semicontinua inferiormente y acotada inferiormente debido a que f es continua (lipschitziana) y S es cerrado. Luego tomemos $\epsilon = |f(u)|$ y claramente se tiene que

$$F(u) \leq \inf_X F + |f(u)|,$$

luego, por el Teorema 2.1, tomando $k = \dfrac{1}{a\, d_E(u)}$, existe $v \in X$, tal que

$$\| v - u \| \leq \frac{1}{k} = a\, d_E(u)$$

$$\forall\, w \in X \,:\, F(w) \geq F(v) - \frac{\| w - v \| \, |f(u)|}{a\, d_E(u)} \tag{3.6}$$

Si tomamos a $w \in S$, se tiene que $F(w) = |f(w)|$ y entonces por la última desigualdad $F(v)$ sería finito, luego

$$v \in S \quad \text{y} \quad F(v) = |f(v)|$$

Así, (3.6) se puede escribir

$$\forall w \in S \,:\, |f(w)| + \frac{\| w - v \| \, |f(u)|}{a\, d_E(u)} \geq |f(v)|,$$

y si consideramos la función $g(w) = |f(w)| + \dfrac{\| w - v \| \, |f(u)|}{a\, d_E(u)}$, vemos que g alcanza su mínimo sobre S en $w = v$. En este sentido, la regla de Multiplicadores de Lagrange nos da que

$$0 \in \partial g(v) + N_S(v)$$

lo cual significa que

$$0 \in \partial |f(v)| + N_S(v) + B^* \frac{|f(u)|}{a\, d_E(u)}$$

donde $B^* = \{x^* \in X^* \,/\, \| x^* \| \leq 1\}$. De aquí se obtiene la conclusión del teorema. ■

Se puede entender el teorema anterior como una versión $n-$dimensional del Teorema del Valor Medio. Para ver esto, establezcamos el siguiente corolario.

Corolario 3.4 *Sea $g : \mathbb{R}^n \to \mathbb{R}$ una función de clase C^1. Para cada $u \in \mathbb{R}^n$, denotemos por $E[u]$ al siguiente conjunto de nivel*

$$E[u] = \{w \,/\, g(w) = g(u)\}.$$

Entonces, para cualesquiera dos puntos u_1 y u_2, existe un punto $v \in \mathbb{R}^n$ tal que

$$\| u_1 - v \| \leq d_{E[u_2]}(u_1)$$

$$\| g'(v) \| \leq \frac{|g(u_1) - g(u_2)|}{d_{E[u_2]}(u_1)}$$

Demostración:

Sean $u_1, u_2 \in \mathbb{R}^n$ cualesquiera. Utilizaremos el Teorema 3.3, para esto tomemos $S = \mathbb{R}^n$ y definamos la función

$$f(u) = g(u) - g(u_2),$$

con la cual, consideraremos el conjunto de nivel

$$E[u_2] = \{u \in \mathbb{R}^n \,/\, f(u) = 0\} \quad \text{cerrado en} \quad \mathbb{R}^n.$$

Entonces para $u_1 \in \mathbb{R}^n \backslash E[u_2]$ y $a \in \langle 0, 1]$, existe algún $v_a \in \mathbb{R}^n$ tal que

$$\| u_1 - v_a \| \leq a\, d_{E[u_2]}(u_1)$$

Existe $v_a^* \in \partial|f(v_a)|$ con $\| v_a^* \| \leq \dfrac{|g(u_1) - g(u_2)|}{a\, d_{E[u_2]}(u_1)}$.

Si $a \in \langle 0, 1\rangle$, de la primera desigualdad notamos que $v_a \notin E[u_2]$, de manera que $f(v_a) \neq 0$ y entonces $\partial|f(v_a)| = \pm f'(v_a) = \pm g'(v_a)$, de donde se verifica el resultado.

Cuando a tiende a 1, la sucesión v_a tendrá un punto límite $v \in \mathbb{R}^n$ pues este punto permanece en la bola de centro u_1 y radio $d_{E[u_2]}(u_1)$. Como $g'(v_a)$ tendrá como punto límite a $g'(v)$, de aquí se deduce el resultado. ■

El Teorema 3.3 se puede expresar en un sentido distinto, dicho resultado se muestra en el corolario siguiente, que tiene importantes aplicaciones en Teoría de Optimización.

Corolario 3.5 *Sean $E = \{w \in S \,/\, f(w) = 0\}$ un subconjunto cerrado de S y $w \in E$. Supongamos que existen ϵ y $c > 0$ tales que, para todo $v \in S \setminus E$ con $\| v - w \| \leq \epsilon$ se tiene que*

$$\text{Si} \quad v_1^* \in \partial|f(v)| \quad \text{y} \quad v_2^* \in N_S(v), \quad \text{entonces} \quad \| v_1^* + v_2^* \| \geq c. \tag{3.7}$$

Entonces, para todo $u \in S$ con $\| u - w \| \leq \dfrac{\epsilon}{2}$ se cumple que

$$d_E(u) \leq \frac{|f(u)|}{c}$$

Demostración:

Sea $u \in S$ con $\| u - w \| \leq \dfrac{\epsilon}{2}$. Luego, por el Teorema 3.3, para cada $a \in \langle 0, 1\rangle$, existe un punto $v_a \in S$ tal que

$$\| v_a - u \| \leq a\, d_E(u) \tag{3.8}$$

Existen $v_1^* \in \partial|f(v_a)|$ y $v_2^* \in N_S(v_a)$ con

$$\| v_1^* + v_2^* \| \leq \frac{|f(u)|}{a\, d_E(u)} \tag{3.9}$$

De la desigualdad (3.8), notemos que $v_a \in S \setminus E$ puesto que al tomar $a < 1$, tenemos que $\| v_a - u \| < d_E(u)$. Luego, de (3.8) también se deduce que

$$\begin{aligned}
\| v_a - w \| &\leq \| v_a - u \| + \| u - w \| \\
&\leq a\, d_E(u) + \| u - w \| \\
&\leq a \,\| u - w \| + \| u - w \| \\
&= (a+1) \,\| u - w \| \\
&\leq \frac{(a+1)\epsilon}{2} \\
&\leq \epsilon.
\end{aligned}$$

Además, de (3.9) también se tiene que

$$d_E(u) \leq \frac{|f(u)|}{a \parallel v_1^* - v_2^* \parallel}.$$

Y por la condición en (3.7) obtenemos que

$$d_E(u) \leq \frac{f(u)}{a\, c}.$$

Haciendo $a \to 1$, obtenemos el resultado deseado.

Conclusiones

En este trabajo se han abordado algunas aplicaciones del principio variacional de Ekeland en sus formas débil y fuerte. El punto esencial de dicho principio según se muestran en las aplicaciones se ubica en el Teorema 2.1, condición (2.3). Tal condición asegura, en particular, la existencia de un punto u tal que suponiendo X espacio de Banach y f Gâteaux diferenciable, $|| \, f'(u) \, ||$ puede ser tan pequeño como uno desee. Esto induce, en el caso cuando X es de Banach, a introducir un concepto de solución aproximada a un problema de minimización, como se estudió en las secciones posteriores.

Ahora, desde otro punto de vista, podemos estar interesados en construir a partir del punto inicial $\overline{u}$ dado (en el Teorema 2.1) para el cual el ϵ que satisface la ecuación en dicho teorema es conocido, otro punto tal que, además de satisfacer (2.3), $f(u)$ se aproxima al $\inf_X f$ con un error $\eta < \epsilon$. El principio no excluye la posibilidad de tener $u = \overline{u}$ si tal punto ya verifica (2.3). Observando (2.2) uno podría esperar $f(u) < f(\overline{u})$ si k es suficientemente pequeño. Si este fuere el caso, el principio variacional tal como fue establecido no nos da una nueva estima del error y por lo tanto no sabemos cuanto hemos mejorado el valor $f(\overline{u})$. Ciertamente, esto no fue el objetivo original del principio.

Bibliografía

[1] A. L. Brown and A. Page, Elements of Functional Analysis, Van Nostrand Reinhold Company (1970)

[2] Alexey Izmailov and Mikhail Solodov, Otimização, Volume 1, IMPA (2005)

[3] Erwin Kreyszig, Introductory Functional Analysis with applications, Jhon Wiley and Sons (1978)

[4] Fabián Florez Bazán, Optimización y cálculo de variaciones sin convexidad:Una introducción, Monografías del IMCA, N°15.

[5] Félix Villanueva Santos, El Principio Variacional de Ekeland y algunas de sus aplicaciones, Tesis de Licenciatura (1992)

[6] Frank H. Clarke, Optimization and Nonsmooth Analysis, Wiley Interscience Publication (1983).

[7] Haim Brezis, Análisis Funcional y Aplicaciones, Alianza Editorial Masson, París (1983).

[8] Ivar Ekeland, Nonconvex Minimization Problems, Bulletin of the American Mathematical Society 1, pp 443 - 474 (1979).

[9] Yboon García Ramos, Un estudio sobre funciones convexas, Tesis de Maestría (2003)