UNIVERSIDAD MAYOR DE SAN ANDRES
FACULTAD DE CIENCIAS PURAS Y NATURALES
CARRERA DE MATEMÁTICA

EL GRADIENTE GENERALIZADO DEL ANÁLISIS NO DIFERENCIABLE

Postulante: *Oscar Bautista Callizaya*
Tutor: *Willy Condori Equice*

La Paz, Agosto del 2013

Este trabajo está dedicado a mi
Tatiana, con amor y gratitud.
Por hacer mi realidad
mejor que los sueños.

Agradecimientos

Agradezco en primer lugar a mi madre Irma Virginia Callisaya Velasco, por estar siempre ahí apoyandome en todo proceso de mi vida y por tener fe en mí, a mis hermanas Patricia Beba Bautista Callisaya y Laura Cruz Bautista Callisaya, por soportarme, por esperar en vano que les escuche cuando estoy 'colgado' en mi mundo matemático y por apoyarme en todo.

A mi tutor del trabajo de licenciatura Willy Condori Equice, por todo el acompañamiento que me dio en el transcurso de la licenciatura, además fue quien aceptó la no muy grata tarea de corregir mis innumerables errores, convirtiendo la lectura de este texto más agradable.

A mi tribunal de Tesis, en particular a Javier Guachalla, por todas las recomendaciones y sugerencias.

Y en especial a mi Tatiana, una persona a la que tuve la suerte de encontrar cuando empezaba el camino matemático y la que ha conseguido, entre otras cosas, que éste sea mi mundo y que lo abrace con pasión e ilusión, por eso esta memoria le está dedicada. Gracias por tu apoyo cuando las cosas del todo no van bien.

Índice general

Resumen

Se pretende generalizar la idea de derivada para funciones no diferenciables, definiremos el Gradiente Generalizado como un subconjunto del dual de un espacio de Banch. Además, se desarrollan en detalle tres resultados esenciales de esta teoría:

i) El Teorema de el Valor Medio;

ii) La Regla de la Cadena;

iii) y una condicion necesaria para que una funcion tenga un minimo local.

Introducción

Análisis no diferencial se refiere al estudio de las funciones que no son diferenciables en el sentido usual, es decir, en ausencia de diferenciabilidad, en el sentido usual. Esto puede ser considerado como un subcampo de lo que se conoce como el análisis no lineal.

Consideremos un espacio de Banach X y una función f de X a R, si f fuese diferenciable, entonces a cada x en X le corresponde una única trasformación lineal de X a R. Nosotros veremos el caso para cuando una tal función va a los reales extendidos y no es necesariamente diferenciable, para dar una noción de diferenciabilidad a cada punto en x, haremos corresponder un conjunto de trasformaciones lineales, para así obtener información de f en dicho punto.

Para este fin primeramente consideraremos el caso en el que f es una función localmente Lipschitz, definiremos la *Derivada Generalizada* como una función de X a R y tomaremos el conjunto de trasformaciones lineales que son mayorizados por la Derivada Generalizada, a dicho conjunto lo llamaremos el *Gradiente Generalizado* de f.

Demostraremos una generalizacion del teorema del valor medio y regla de la cadena, claramente se notara que para funciones C^1 en espacios de dimension finita, los teoremas clasicos seran casos particulares de los teoremas ya mensionados.

Demostraremos una propiedad del Gradiente Generalizado, que nos ayudara a extender la idea de Gradiente Generalizado para funciones no Lipschitz y que van para los reales extendidos.

Capítulo 1

Gradientes Generalizados

Nosotros trabajaremos en un espacio de Banach X cuya norma denotaremos con $\|\cdot\|$, y a la bola unitaria como B.

Definición 1. *Sea Y un subconjunto de X. Una función $f : Y \to R$ es Lipschitz (en Y) si satisface que, para algún K, real no negativo, uno tiene*

$$|f(y) - f(y')| \leq K\|y - y'\|$$

para todo y, y' en Y; en este caso se dice que f es Lipschitz de grado K. Decimos que f es Lipschitz de grado K cerca a x si, para algún $\varepsilon > 0$, f satisface la condición de Lipschitz (de grado K) en $B(x, \varepsilon)$.

Como ya es sabido una función diferenciable es continua y podemos hallar una variedad de funciones localmente Lipschitz, basta que la derivada sea acotada, pero no toda función Lipschitz es diferenciable, ahora introduciremos una noción de derivada para funciones Lipschitz.

1.1. La Derivada Direccional Generalizada

Definición 2. *Sea f Lipschitz cerca a x, y sea v otro vector en X. La derivada direccional generalizada de f en x en la dirección v, denotado $f^{\circ}(x; v)$, is definido como sigue:*

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

donde y es un vector en X y t es un real positivo.

Notemos que esta definición no presupone la existencia del limite, esta es una definición de derivada que difiere de la derivada tradicional en que vemos la variación con respecto a x. Ahora damos algunas propiedades básicas de f°.

Proposición 1. *Sea f Lipschitz de grado K cerca a x. Entonces*
(a) La función $v \mapsto f^\circ(x;v)$ es finita, positivamente homogénea, subaditiva en X y satisface

$$|f^\circ(x;v)| \leq K\|v\|.$$

(b) $(x,v) \mapsto f^\circ(x;v)$ es semicontinua superior y $v \mapsto f^\circ(x;v)$ es Lipschitz de grado K en X.
(c) $f^\circ(x;-v) = (-f)^\circ(x;v)$.

Demostración. (a) Primero veamos que $f^\circ(x;v)$ es finito, por ser f lipschitz tenemos que

$$|\tfrac{f(y+tv)-f(y)}{t}| \leq K\|v\|,$$

equivalentemente a

$$-K\|v\| \leq \tfrac{f(y+tv)-f(y)}{t} \leq K\|v\|,$$

aplicando el limite superior miembro a miembro,

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

esto nos dice que $|f^\circ(x;v)| \leq K\|v\|$ y que $f^\circ(x;v)$ es finito. Por otro lado sea $\lambda > 0$

$$f^\circ(x;\lambda v) = \limsup_{\substack{y\to x\\ t\downarrow 0}} \tfrac{f(y+t\lambda v)-f(y)}{t}$$

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

con lo que demuestra que es positivamente homogénea. Para subaditividad, recordemos que el limite superior de una suma de funciones es menor o igual a la suma de los limites superiores de cada función([2]), sea v, w en X, entonces tenemos que

$$\tfrac{f(y+t(v+w))-f(y)}{t} = \tfrac{f((y+tw)+tv)-f(y+tw)}{t} + \tfrac{f(y+tw)-f(y)}{t}$$

cuanto $y \to x$ y $t \downarrow 0$, es claro que $(y+tw) \to x$, entonces aplicando el limite superior a la igualdad, este se convierte en desigualdad, por lo dicho anteriormente, así nos queda

$$f^{\circ}(x;v+w) \leq f^{\circ}(x;v) + f^{\circ}(x;w),$$

como se quería.

(b) Existen muchas equivalencias de la semicontinuidad superior, esta es una de ellas, basta probar que para $\{x_i\}$ y $\{v_i\}$ sucesiones arbitrarias en X convergiendo a x y v, respectivamente, se tenga que

$$\limsup_{i\to\infty} f^{\circ}(x_i;v_i) \leq f^{\circ}(x;v),$$

veamos que es así. Por la definición de limite superior para $\frac{1}{i} > 0$, existe $y_i \in B(x_i, \frac{1}{i})$ y $t_i \in (0, \frac{1}{i})$ tal que

$$\begin{aligned} f^{\circ}(x_i;v_i) - \tfrac{1}{i} &\leq \tfrac{f(y_i+t_iv_i)-f(y_i)}{t_i} \\ &= \tfrac{f(y_i+t_iv)-f(y_i)}{t_i} + \tfrac{f(y_i+t_iv_i)-f(y_i+t_iv)}{t_i} \\ &\leq \tfrac{f(y_i+t_iv)-f(y_i)}{t_i} + K\|v_i - v\|, \end{aligned}$$

la ultima parte es por se f Lipschitz, describiendo la desigualdad tenemos

$$f^{\circ}(x_i;v_i) \leq \tfrac{f(y_i+t_iv)-f(y_i)}{t_i} + K\|v_i - v\| + \tfrac{1}{i},$$

aplicando el limite superior (de sucesiones), que se comporta como el limite superior de funciones en las desigualdades([2]), nos queda

$$\limsup_{i\to\infty} f^\circ(x_i; v_i) \leq \limsup_{i\to\infty} \tfrac{f(y_i+t_i v)-f(y_i)}{t_i},$$

la parte derecha, de la desigualdad, es acotado por $f^\circ(x;v)$ ([2]), de aquí

$$\limsup_{i\to\infty} f^\circ(x_i; v_i) \leq f^\circ(x; v),$$

con esto es semicontinua superior. Finalmente para v y w en X. Nosotros tenemos por ser f Lipschitz

$$f(y+tv) - f(y+tw) \leq K\|v-w\|t,$$

haciendo operaciones elementales

$$f(y+tv) - f(y) \leq f(y+tw) - f(y) + K\|v-w\|t,$$

dividiendo por t y aplicando el limite superior con $y \to x$, $t \downarrow 0$, tenemos

$$f^\circ(x;v) \leq f^\circ(x;w) + K\|v-w\|.$$

Cambiando v por w y viceversa, tenemos lo requerido

$$|f^\circ(x;v) - f^\circ(x;w)| \leq K\|v-w\|.$$

(c) Haremos cálculos directos

$$\begin{aligned} f^\circ(x;-v) &= \limsup_{\substack{x'\to x\\ t\downarrow 0}} \tfrac{f(x'-tv)-f(x')}{t} \\ &= \limsup_{\substack{y\to x\\ t\downarrow 0}} \tfrac{(-f)(y-tv)-(-f)(y)}{t} \\ &= (-f)^\circ(x;v), \end{aligned}$$

donde $y = x' - tv$ y claramente $y \to x$, cuando $x' \to x$ y $t \downarrow 0$ (en la segunda igualdad). $\square$

1.2. El Gradiente Generalizado

Tomemos un v en X fijo, por el algebra lineal para $f^\circ(x;v)$ y v, existe ζ' en X^*, tal que $\zeta'(v) = f^\circ(x;v)$.
Afirmación

$$\zeta'(u) \leq f^\circ(x;u),$$

para todo u en $span\{v\}$. En efecto, denotemos $span\{v\} = V$, debemos mostrar que $\zeta'(\lambda v) \leq f^\circ(x;\lambda v)$, para $\lambda \in R$.

-Sea $\lambda \geq 0$, por ser $f^\circ(x;\cdot)$ positivamente homogénea, tenemos

$$\zeta'(\lambda v) = \lambda\zeta'(v) = \lambda f^\circ(x;v) = f^\circ(x;\lambda v).$$

Ahora probemos para $\lambda = -1$,

$$\zeta'(-v) = -\zeta'(v) = -f^\circ(x;v) \leq f^\circ(x;-v),$$

la ultima igualdad es por la subaditividad, es decir, es verdad que $0 = f^\circ(x;v-v) \leq f^\circ(x;v) + f^\circ(x;-v)$.
-Sea $\lambda < 0$, entonces

$$\zeta'(\lambda v) = -\lambda\zeta'(-v) \leq -\lambda f^\circ(x;-v) = f^\circ(x;\lambda v).$$

Con todo esto encontramos un funcional lineal $\zeta' : V \to R$ que es acotado por $f^\circ(x;\cdot)$ en V. Por el Teorema de Hahn-Banach ([3]),por ser $f^\circ(x;\cdot)$ positivamente homogénea y subaditiva, podemos extender ζ' a ζ, tal que

$$\zeta(u) \leq f^\circ(x;u),$$

para todo u en X y $\zeta\mid_V = \zeta'$. Con esto en manos y usando la siguiente notación de ahora en adelante $\zeta(v) = \langle\zeta, v\rangle = \langle v, \zeta\rangle$,definiremos el gradiente generalizado.

Definición 3. *El Gradiente Generalizado de f en x, denotado por $\partial f(x)$, es el subconjunto de X^* dado por*

$$\partial f(x) = \{\zeta \in X^* : \langle\zeta, v\rangle \leq f^\circ(x;v) \text{ para todo } v \text{ en } X\}.$$

Nosotros denotamos por $\|\zeta\|_*$ la norma en X^*:

$$\|\zeta\|_* = \sup\{\langle \zeta, v\rangle : v \in X, \|v\| \leq 1\}.$$

Ahora vamos con algunas propiedades básicas del gradiente generalizado.

Proposición 2. *Sea f Lipschitz de grado K cerca a x. Entonces*

(a) $\partial f(x)$ es no vació, convexo, débilmente-compacto en X^* y $\|\zeta\|_* \leq K$ para cada ζ en $\partial f(x)$.*

(b) Para cada v en X, uno tiene

$$f^\circ(x; v) = \text{máx}\{\langle \zeta, v\rangle : \zeta \in \partial f(x)\}$$

.

Demostración. (a) $\partial f(x)$ es no vació por la construcción que hicimos usando el teorema de Hahn-Banach. Para la convexidad sea ζ, ζ' en $\partial f(x)$ y $0 \leq t \leq 1$, entonces

$$\langle (1-t)\zeta + t\zeta', v\rangle = (1-t)\langle \zeta, v\rangle + t\langle \zeta', v\rangle \leq (1-t)f^\circ(x; v) + tf^\circ(x; v) = f^\circ(x; v),$$

con esto $(1-t)\zeta + t\zeta'$ esta en $\partial f(x)$. Ahora veamos que $\|\zeta\|_* \leq K$, en efecto pues para cada v en X

$$\langle \zeta, v\rangle \leq f^\circ(x; v) \leq K\|v\|,$$

en particular para $\|v\| \leq 1$, se tiene que

$$\langle \zeta, v\rangle \leq K,$$

por la definición de supremo $\|\zeta\|_* \leq K$. Podemos decir entonces que ζ esta en $B(0, K)$ que es compacto por el Teorema de Alaoglu ([3]), basta mostrar que $\partial f(x)$ es cerrado, ya que todo subconjunto cerrado de un compacto es compacto, para esto sea (ζ_i) una sucesión tal que converge a ζ, con ζ_i en $\partial f(x)$,entonces para todo v en X se tiene que

$$\langle \zeta_i, v\rangle \leq f^\circ(x; v)$$

por propiedades de convergencia débil* hacemos $i \to \infty$ y tenemos

$$\langle \zeta, v\rangle \leq f^\circ(x; v).$$

Ahora solo basta mostrar que ζ es continuo, en efecto por propiedades de convergencia débil*, se tiene que

$$\|\zeta\| \leq \liminf \|\zeta_i\|,$$

como $\|\zeta_i\| \leq K$, entonces $\|\zeta\| \leq K$ lo que nos dice que ζ es acotado y por tanto continuo.

(b) Para empezar notemos que tiene sentido tomar el máximo, pues como ya vimos $\partial f(x)$ es compacto y por la continuidad de φ_v, donde $\varphi(v) = \varphi_v$ (φ es la inclusion canonica), $\varphi_v(\partial f(x)) = \{\langle \zeta, v\rangle : \zeta \in \partial f(x)\}$ es compacto en R y así podemos tomar el máximo de la imagen. Es claro, por definición que

$$f^{\circ}(x; v) \geq \max\{\langle \zeta, v\rangle : \zeta \in \partial f(x)\},$$

sea $\langle \zeta, v\rangle$ el máximo, es decir $f^{\circ}(x; v) \geq \langle \zeta, v\rangle$, por contradicción supongamos que $f^{\circ}(x; v) > \langle \zeta, v\rangle$, exactamente como antes por el Teorema de Hahn-Banach, podemos encontrar ζ' en X^* tal que $f^{\circ}(x; v) = \langle \zeta', v\rangle$ y $\zeta' \in \partial f(x)$, pero $f^{\circ}(x; v) = \langle \zeta', v\rangle > \langle \zeta, v\rangle$ lo cual es una contradicción. $\square$

Ejemplo. 1. *Sea $f : R \to R$ definido como $f(x) = |x|$, calculemos $\partial f(x)$.*

Para $x > 0$

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

es claro que si $\zeta \cdot v \leq v$, para todo v en R, entonces $\zeta = 1$, así $\partial f(x) = \{1\}$. Análogamente para $x < 0$ se puede probar que $\partial f(x) = \{-1\}$. Para $x = 0$ haciendo cálculos se puede probar que $f^{\circ}(0; v) = |v|$, ahora si $\zeta \cdot v \leq |v|$, para todo v en R, entonces $-1 \leq \zeta \leq 1$, así $\partial f(0) = [-1, 1]$.

1.3. Funciones Soporte

Definición 4. *Llamamos función de soporte de un subconjunto no vació C de X a la función $\sigma_C : X^* \to R \cup \{\infty\}$ definido por*

$$\sigma_C(\zeta) = \sup\{\langle \zeta, x\rangle : x \in C\}.$$

Si Σ es un subconjunto de X^*, su función de soporte es $\sigma_\Sigma : X \to R \cup \{\infty\}$ definido por

$$\sigma_\Sigma(x) = \sup\{\langle \zeta, x \rangle : \zeta \in \Sigma\}.$$

Proposición 3. *Sea C, D subconjuntos de X convexos, cerrados y no vacíos, y sea Σ, Δ subconjuntos de X^* convexos, débilmente*-cerrados y no vacíos. Entonces*

(a) $C \subset D$ sii $\sigma_C(\zeta) \leq \sigma_D(\zeta)$; para todo $\zeta \in X^*$.

(b) $\Sigma \subset \Delta$ sii $\sigma_\Sigma(x) \leq \sigma_\Delta(x)$; para todo $x \in X$.

Demostración. (a) Si $C \subset D$, entonces $\langle \zeta, c \rangle \leq \sigma_D(\zeta)$ para c en C y ζ en X^* fijo, por la definición de supremo tenemos que $\sigma_C(\zeta) \leq \sigma_D(\zeta)$, para todo ζ en X^*.

Por otro lado si existe c en $C \backslash D$, entonces por el Teorema de separación de Hahn-Banach para $\{c\}$ y D, existen ζ en X^* y α, β en R, tal que

$$\langle \zeta, d \rangle \leq \alpha < \beta \leq \langle \zeta, c \rangle,$$

para todo d en D, por la definición de supremos $\sigma_D(\zeta) \leq \langle \zeta, c \rangle$, lo cual implica $\sigma_D(\zeta) \leq \sigma_C(\zeta)$, que es una contradicción. (b) Del mismo modo que (a), usando el Teorema de separación de Hahn-Banach en X^*.

□

Una $multifuncion$ $\Gamma : X \to Y$ es una correspondencia de X a los subconjuntos de Y.

Definimos *semicontinua superior* de Γ en x si, cumple que para todo $\varepsilon > 0$, existe $\delta > 0$ tal que si , x' esta en $x + \delta B$, entonces $\Gamma(x')$ es subconjunto de $\Gamma(x) + \varepsilon B$, donde $B = B(0,1)$.

Nosotros continuamos asumiendo $f : X \to R$, Lipschitz cerca de x.

Proposición 4. (a) *Sea x_n y ζ_n sucesiones en X tales que $\zeta_n \in \partial f(x_n)$. Supongamos que x_n converge a x, y que ζ es un punto clausura de ζ_n. Entonces $\zeta \in \partial f(x)$.*

(b) $\partial f(x) = \bigcap_{\delta>0} \bigcup_{y \in B(x,\delta)} \partial f(y)$.

(c) Si X es de dimension finita,entonces ∂f es semicontinua superior en x.

Demostración. (a) Ya demostramos que, para todo v en X, $f^{\circ}(\cdot; v)$ es semicontinua superior en x, es decir, para cada $\varepsilon > 0$, existe $\delta > 0$ tal que si, $\| x - x' \| < \delta$, entonces $f^{\circ}(x'; v) < \varepsilon + f^{\circ}(x; v)$. Ahora para un n suficientemente grande $\| x - x_n \| < \delta$ y $f^{\circ}(x_n; v) < \varepsilon + f^{\circ}(x; v)$, como $\langle \zeta_n, v \rangle \leq f^{\circ}(x_n; v)$, se sigue que $\langle \zeta_n, v \rangle < \varepsilon + f^{\circ}(x; v)$, ahora por hipotesis existe $\zeta_{n_j} \to \zeta$, ahora en particular para ζ_{n_j}, $\langle \zeta_{n_j}, v \rangle \leq \varepsilon + f^{\circ}(x; v)$, haciendo $j \to \infty$, tenemos que $\langle \zeta, v \rangle \leq \varepsilon + f^{\circ}(x; v)$, como ε es arbitrario, concluimos que $\langle \zeta, v \rangle \leq f^{\circ}(x; v)$, que es lo mismo que $\zeta \in \partial f(x)$.

(b) Sea $y' \in \partial f(x)$, es claro que $y' \in \bigcup\limits_{y \in B(x,\delta)} \partial f(y)$, para todo $\delta > 0$, entonces $y' \in \bigcap\limits_{\delta > 0} \bigcup\limits_{y \in B(x,\delta)} \partial f(y)$ y $\partial f(x) \subset \bigcap\limits_{\delta > 0} \bigcup\limits_{y \in B(x,\delta)} \partial f(y)$. Por otro lado sea $y' \in \bigcap\limits_{\delta > 0} \bigcup\limits_{y \in B(x,\delta)} \partial f(y)$, tomando $\delta = \frac{1}{n}$, existe $x_n \in B(x; \frac{1}{n})$ tal que $y' \in \partial f(x_n)$ y esta claro que x_n converge a x, por (a), se tiene que $y' \in \partial f(x)$. Así $\bigcap\limits_{\delta > 0} \bigcup\limits_{y \in B(x,\delta)} \partial f(y) \subset \partial f(x)$.

(c) Si suponemos que ∂f no es semicontinua superior en x de R^n, existe $\varepsilon > 0$, para todo $\frac{1}{m} > 0$, existe $x_m \in B(x, \frac{1}{m})$ y existe $y_m \in \partial f(x_m)$, para todo $z \in \partial f(x)$ tal que, $y_n \notin B(z, \varepsilon)$. Por otro lado ya demostramos, por ser f Lipschitz cerca de x, $|\langle y_m, e_i \rangle| \leq |f^{\circ}(x_m; e_i)| \leq K \| e_i \| = K$, para $i = 1, ...n$, entonces $\langle y_m, e_i \rangle$ posee una subsucesión convergente, tomemos $\langle y_{m_j}, e_i \rangle \to \alpha_i$, ahora tenemos un conjunto finito de sucesiones convergentes y podemos tomar $\{y_k\} \subset \{y_m\}$, tal que $\langle y_k, e_i \rangle \to \alpha_i$, para todo $i = 1, ..., n$, si tomamos $y = \sum\limits_{i=1}^{n} \alpha_i e_i$, esta claro que $y_k \to y$, ahora $\| y_k - z \| \geq \varepsilon$, $y \notin \partial f(x)$, lo cual es una contradicción a (b), pues $x_m \to x$ y y es un punto clausura de y_m. Así ∂f es semicontinua superior en x.

$\square$

Capítulo 2

Derivadas Clásicas

Sea $F : X \to Y$, nosotros decimos que F admite una *Derivada de Gateaux* en x, si existe $DF(x)$ en $\mathcal{L}(X, Y)$ tal que, para cada v en X, uno tiene

$$\lim_{t \downarrow 0} \frac{F(x+tv)-F(x)}{t} = \langle DF(x), v \rangle,$$

esta convergencia es puntual. Si la convergencia es uniforme para v en conjuntos compactos, la derivada es conocida como *Hadamard*.

2.1. Estrictamente Diferenciable

Definición 5. *Nosotros decimos que F admite deriva estricta en x, si existe un elemento en $\mathcal{L}(X, Y)$ denotado por $D_s F(x)$ que cumple para cada v en X*

$$\lim_{\substack{x' \to x \\ t \downarrow 0}} \frac{F(x'+tv)-F(x')}{t} = \langle D_s F(x), v \rangle$$

y proporciona la convergencia uniforme para v en conjuntos compactos.

Proposición 5. *Sea F una función de una vecindad de x de X a Y, y sea ζ un elemento de $\mathcal{L}(X, Y)$. Las siguientes afirmaciones son equivalentes:*

(a) F es estrictamente diferenciable en x y $D_s F(x) = \zeta$.

(b) F es Lipschitz cerca de x, y para cada v en X uno tiene

$$\lim_{\substack{x'\to x\\ t\downarrow 0}} \tfrac{F(x+tv)-F(x)}{t} = \langle \zeta, v\rangle.$$

Demostración. Si asumimos (a). La igualdad en (b) vale por la suposición, entonces para la prueba de (b) necesitamos solamente mostrar que F es Lipschitz cera de x. Si este no es le caso, existen sucesiones $\{x_i\}$ y $\{y_i\}$ tales que x_i, y_i viven en $B(x, \frac{1}{i})$ y

$$\|F(x_i) - F(y_i)\| > i\|x_i + y_i\|.$$

Podemos encontrar v_i en X y t_i en R^+ tal que $x_i = y_i + t_i v_i$ y $\|v_i\| = \frac{1}{\sqrt{i}}$. Se sigue que $t_i \to 0$. Sea $V = \{v_i\} \cup \{0\}$. Notemos que V es compacto, entonces por la definición de $D_x F(x)$ para todo $\varepsilon > O$ existe i_0 tal que, para cada $i > i_0$ y para todo v en V, uno tiene

$$\| \tfrac{F(y_i+t_i v)-F(y_i)}{t_i} - \langle D_s F(x), v\rangle \| < \varepsilon.$$

Pero esto es imposible, pues cuando $v = v_i$ tenemos que

$$\varepsilon > \| \tfrac{F(y_i+t_i v_i)-F(y_i)}{t_i} \| - \| \langle D_s F(x), v_i\rangle \| \geq \sqrt{i} - \| \langle D_s F(x), v_i\rangle \|$$

y el termino de la derecha crece, cuando $i \to \infty$. Así (b) vale.

Ahora Supongamos (b). Sea V cualquier compacto de X y ε cualquier numero positivo. Por la igualdad de (b), existe para cada v de V un numero $\delta'_v > 0$ tal que para todo $x' \in B(x, \delta'_v)$ y $t \in B(0, \delta'_v)$, se tiene

$$\| \tfrac{F(x'+tv)-F(x')}{t} - \langle \zeta, v\rangle \| < \tfrac{\varepsilon}{3}$$

Tomemos $\delta_v = \min\{\frac{\varepsilon}{3K}, \frac{\varepsilon}{3C}, \delta'_v\}$, donde $K > 0$ es la constante de Lipschitz para F y C cumple que $\| \zeta, u\rangle\| \leq C\|u\|$ para todo u en X. Ahora es claro que $V \subset \bigcup_{v\in V}\{B(v, \delta_v)$, como V es compacto podemos hallar un numero finito de $v_1, ..., v_k$ tales que $V \subset \bigcup_{i=1}^{k}\{B(v_i, \delta_{v_i})\}$.

Si nosotros tomamos $\delta = \min_{1\leq i\leq k} \delta_{v_i}$, así para todo $x' \in B(x, \delta)$, $t \in B(0, \delta)$ y v' en V, se tiene $v' \in B(v_i, \delta_{v_i})$ para algún $i \in [1, ...k]$ y $x' \in B(x, \delta_{v_i})$, $t \in B(0, \delta_{v_i})$, entonces

$\| \frac{F(x'+tv')-F(x')}{t} - \langle \zeta, v' \rangle \| \leq \| \frac{F(x'+tv')-F(x')}{t} - \frac{F(x'+tv_i)-F(x')}{t} \| + \| \frac{F(x'+tv_i)-F(x')}{t} - \langle \zeta, v_i \rangle \|$

$$+ \| \langle \zeta, v_i \rangle - \langle \zeta, v' \rangle \|$$

$$\leq K \| v_i - v' \| + \tfrac{\varepsilon}{3} + C \| v_i - v' \| \leq \tfrac{\varepsilon}{3} + \tfrac{\varepsilon}{3} + \tfrac{\varepsilon}{3} = \varepsilon.$$

□

Proposición 6. *Sea f Lipschitz cerca de x y admite derivada de Gateaux (Hadamard o estricta) $Df(x)$. Entonces $Df(x) \in \partial f(x)$.*

Demostración. Por definición, $f'(x;v)$ existe para cada v y es igual a $\langle Df(x), v \rangle$. Claramente uno tiene $f' \leq f^\circ$, entonces uno tiene $\langle Df(x), v \rangle \leq f^\circ(x;v)$ para todo v en X, lo que nos dice $Df(x) \in \partial f(x)$.

□

Ejemplo. 2. *$\partial f(x)$ puede tener otros puntos además de $Df(x)$, un ejemplo de esto es una función en R: $f(x) = x^2 \operatorname{sen} \frac{1}{x}$. esta función es Lipschitz cerca de 0, y es facil mostrar que $f^\circ(0;v) = |v|$. Se sigue que $\partial f(0) = [-1, 1]$, un conjunto el cual contiene la derivada (no estricta) $Df(0) = 0$.*

Definición 6. *f es continuamente diferenciable en x si, prueba que en una vecindad de x la derivada de Gateaux Df existe y es continua como una función de X a $\mathcal{L}(X, Y)$.*

Proposición 7. *Sea $f : X \to R$. Si f es estrictamente diferenciable en x, entonces f es Lipschitz cerca de x y $\partial f(x) = \{D_s f(x)\}$. Recíprocamente, si f es Lipschitz cerca de x y $\partial f(x) = \{\zeta\}$, entonces f es estrictamente diferenciable en x y $D_s f(x) = \zeta$.*

Demostración. Supongamos primero que $D_s f(x)$ existe, como ya vimos f es Lipschitz cerca de x. Entonces, por la definición de f°, uno tiene $f^\circ(x;v) = \langle D_s f(x), v \rangle$ para todo v, como no existe una función lineal que sea menor a otra función lineal, se tiene que $\partial f(x)$ se reduce a $\{D_s f(x)\}$. Para la recíproca solo basta que el limite existe y es igual a ζ, para empezar $f^\circ(x;v) = \langle \zeta, v \rangle$, para cada v en X, en efecto si no fuese así existiría v_0 en X tal que $f^\circ(x;v) > \langle \zeta, v_0 \rangle$, ahora podemos encontrar, por el teorema de Hahn-Banach, ζ' lineal tal que $\zeta' \in \partial f(x)$ y $f^\circ(x;v_0) = \langle \zeta', v_0 \rangle$, ahora $f^\circ(x;v_0) = \langle \zeta', v_0 \rangle > \langle \zeta, v_0 \rangle$, entonces $\zeta \neq \zeta'$, que es una contradicción, pues $\partial f(x)$ es unitario. Así $f^\circ(x;v) = \langle \zeta, v \rangle$, para cada v en X.

Nosotros ahora calculamos:

$$\liminf_{\substack{x'\to x\\ t\downarrow 0}}\tfrac{f(x'+tv)-f(x')}{t} = -\limsup_{\substack{x'\to x\\ t\downarrow 0}}\tfrac{f(x')-f(x'+tv)}{t} = -\limsup_{\substack{x'\to x\\ t\downarrow 0}}\tfrac{f(x'+tv-tv)-f(x'+tv)}{t} = -f^{\circ}(x;-v)$$

$$= -\langle \zeta, -v\rangle = \langle \zeta, v\rangle = f^{\circ}(x;v) = \limsup_{\substack{x'\to x\\ t\downarrow 0}}\tfrac{f(x'+tv)-f(x')}{t}.$$

Esto nos dice $\lim\limits_{\substack{x'\to x\\ t\downarrow 0}}\frac{f(x'+tv)-f(x')}{t}$ existe y es igual a $\langle \zeta, v\rangle$, como f es Lipschitz cerca a x, por la proposición 5, concluimos que es estrictamente diferenciable en x.

$\square$

Corolario

Si f es Lipschitz cerca a x en X, entonces $\partial f(x')$ se reduce a un solo punto para cada x' en $B(x,\varepsilon)$ si, y solo si, f es continuamente diferenciable en $B(x,\varepsilon)$.

Demostración. Veamos que $D_s f : X \to \mathcal{L}(X,Y)$ es continua en x' de $B(x,\varepsilon)$, por la hipótesis y la proposición anterior, f es G-diferenciable en x', ahora sabemos que ∂f es semicontinua superior (como multifunction) es decir, para cada $\varepsilon > 0$, existe $\delta > 0$ tal que para $y \in x' + \delta B$, se tiene que $\partial f(y) \in \partial f(x') + \varepsilon B$, por hipótesis $\partial f(y)$ y $\partial f(x')$ se reducen aun punto que son $D_s f(y)$ y $D_s f(x')$ respectivamente, lo ultimo se reduce a $D_s f(y) \in D_s f(x') + \varepsilon B$, lo que prueba la continuidad. Por otro lado si suponemos f continuamente diferenciable en x' de $B(x,\varepsilon)$, f es G-Diferenciable en x', es decir existe $Df(x')$, probemos que $Df(x') = D_s f(x')$. Primero para $\frac{\varepsilon}{\|v\|} > 0$ $(v \in X)$, existe $\delta'_v > 0$ tal que para $y \in x' + \delta'_v B$, se tiene que

$$\| Df(y) - Df(x') \| < \tfrac{\varepsilon}{2\|v\|}$$

Por otra parte para todo $\varepsilon > 0$, existe $\delta'' > 0$ tal que para $0 < t < \delta''$, se tiene que

$$\| \tfrac{f(y+tv)+f(y)}{t} - \langle Df(y), v\rangle \| < \varepsilon/2$$

Ahora para todo $\varepsilon > 0$, existe $\delta_v = min\{\delta'_v, \delta''\}$ tal que para $\| y - x' \| < \delta_v$ y $0 < t < \delta_v$ se tiene que

$$\| \tfrac{f(y+tv)+f(y)}{t} - \langle Df(x'), v\rangle \leq \| \tfrac{f(y+tv)+f(y)}{t} - \langle Df(y), v\rangle \| + \| \langle Df(y), v\rangle - \langle Df(x'), v\rangle \|$$

$$< \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

así

$$\lim_{\substack{x' \to x \\ t \downarrow 0}} \frac{f(y+tv)-f(y)}{t} = \langle Df(x'), v \rangle$$

como f es lipschitz cerca a x', $\partial f(x')$ se reduce a un solo punto.

$\square$

2.2. Funciones Convexas

Definición 7. *Sea U subconjunto de X abierto convexo. Decimos que una función $f : U \to R$ es convexa si satisface la propiedad de que, para todo u, u' en U y λ en $[0,1]$, uno tiene*

$$f(\lambda u + (1-\lambda)u') \leq \lambda f(u) + (1-\lambda)f(u').$$

Nosotros ahora mostraremos que, casi toda función convexa es Lipschitz.

Proposición 8. *Sea f acotado superiormente en algún abierto de algún punto de U. entonces, para cualquier x en U, f es Lipschitz cerca a x.*

Demostración. Primero probaremos f es acotado en algún abierto de x. Sin perdida de generalidad, supongamos que f es acotado superiormente por M en $B(0,\varepsilon) \subset U$. Tomemos $\rho > 1$ tal que $y = \rho x$ esta en U. Si $\lambda = 1/\rho$, entonces el conjunto

$$V = \{v : v = (1-\lambda)x' + \lambda y, x' \in B(0,\varepsilon)\} = B(x, (1-\lambda)\varepsilon).$$

Para todo v en V,por ser f convexo, uno tiene que

$$f(v) \leq (1-\lambda)f(x') + \lambda f(y) \leq M + \lambda f(y),$$

así f es acotado superiormente en V. Por otro lado tomemos z en $V = B(x,(1-\lambda)\varepsilon)$, tomemos $z' = 2x - z$ en $B(x,(1-\lambda)\varepsilon)$, así $x = (z+z')/2$, entonces

$$f(x) \leq \tfrac{1}{2}f(z) + \tfrac{1}{2}f(z').$$

Ahora despejando $f(z)$ y por lo demostrado

$$f(z) \geq 2f(x) - f(z') \geq 2f(x) - M - \lambda f(y),$$

con esto f es acotada inferiormente en V. Por tanto es acotada en V.

Sea N una cota de $|f|$ en $B(x, 2\delta)$, donde $\delta > 0$. Sea x_1, x_2 en $B(x, \delta)$, sea $x_3 = x_2 + (\delta/\alpha)(x_2 - x_1)$ donde $\alpha = \|x_2 - x_1\|$, notemos que x_3 esta en $B(x, 2\delta)$. Despejando x_2 tenemos

$$x_2 = \tfrac{\delta}{\alpha+\delta} x_1 + \tfrac{\alpha}{\alpha+\delta} x_3,$$

por la convexidad de f

$$f(x_2) \leq \tfrac{\delta}{\alpha+\delta} f(x_1) + \tfrac{\alpha}{\alpha+\delta} f(x_3).$$

Entonces restando $f(x_1)$ a la desigualdad y haciendo operaciones elementales

$$f(x_2) - f(x_1) \leq \tfrac{\alpha}{\alpha+\delta}[f(x_3) - f(x_1)] \leq \tfrac{\alpha}{\delta}|F(x_3) - f(x_1)|,$$

como $|f| \leq N$ y $\alpha = \|x_2 - x_1\|$ concluimos que

$$f(x_2) - f(x_1) \leq \tfrac{2N}{\delta}\|x_2 - x_1\|.$$

Como x_2 y x_1 son arbitrarios invirtiéndolos tenemos

$$|f(x_2) - f(x_1)| \leq \tfrac{2N}{\delta}\|x_2 - x_1\|,$$

con esto f es Lipschitz cerca a x.

$\square$

Ahora haremos una pequeña relación entre el gradiente generalizado y el subdiferencial, del análisis convexo.

Definición 8. *El $subdiferencial$ de una función convexa f en x es definido como el conjunto de los ζ en X^* que satisfacen*

$$\langle \zeta, x' - x \rangle \leq f(x') - f(x),$$

para todo x' en U.

Como nosotros denotamos al gradiente generalizado como ∂f denotaremos como $\partial_s f$ al Subdiferencial.

Nota

En el análisis convexo, al subdiferencial se lo denota como ∂f y al gradiente generalizado como $\overline{\partial f}$, pero aquí no usaremos esa notación.

Proposición 9. *Cuando f es convexo en U y Lipschitz cerca a x, entonces $\partial f(x) = \partial_s f(x)$, y $f^\circ(x; v)$ coincide la derivada direccional $f'(x; v)$, para cada v en X.*

Demostración. Por el análisis convexo, $f'(x; v)$ existe y es la función de soporte de $\partial_s f(x)$. Solo basta mostrar que $f^\circ(x; v) = f'(x; v)$. Por la definición de la derivada generalizada se tiene $f'(x; v) \leq f^\circ(x; v)$.

Para la otra desigualdad notemos que

$$f^\circ(x; v) = \lim_{\varepsilon \downarrow 0} \sup_{\|x-x'\|\varepsilon\delta} \sup_{0<t<\varepsilon} \tfrac{f(x'+tv)-f(x')}{t},$$

donde δ es cualquier numero positivo. Como f es convexo la función

$$t \mapsto \tfrac{f(x'+tv)-f(x')}{t}$$

es no decreciente, en efecto, sea $0 < t < t'$, entonces $0 < t/t' < 1$

$$\begin{aligned} f(x'+tv) &= f((1-\tfrac{t}{t'})x' + \tfrac{t}{t'}(x'+t'v)) \\ &\leq (1-\tfrac{t}{t'})f(x') + \tfrac{t}{t'}f(x'+t'v) \\ &= f(x') - \tfrac{t}{t'}f(x') + \tfrac{t}{t'}f(x'+t'v) \end{aligned}$$

haciendo operaciones elementales se tiene lo deseado

$$\tfrac{f(x'+tv)-f(x')}{t} \leq \tfrac{f(x'+t'v)-f(x')}{t'}.$$

Con esto la formula se puede escribir como

$$f^\circ(x; v) = \lim_{\varepsilon \downarrow 0} \sup_{\|x-x'\|\varepsilon\delta} \tfrac{f(x'+\varepsilon v)-f(x')}{\varepsilon}.$$

Ahora por la condición de Lipschitz, para $x' \in B(x, \varepsilon\delta)$, uno tiene

$$
\begin{aligned}
\left|\tfrac{f(x'+\varepsilon v)-f(x')}{\varepsilon}-\tfrac{f(x+\varepsilon v)-f(x)}{\varepsilon}\right| &\leq \tfrac{|f(x'+\varepsilon v)-f(x+\varepsilon v)|}{\varepsilon}+\tfrac{|f(x')-f(x)|}{\varepsilon}\\
&\leq \tfrac{K\|x'-x\|}{\varepsilon}+\tfrac{K\|x'-x\|}{\varepsilon}\\
&\leq \tfrac{2K\varepsilon\delta}{\varepsilon}=2K\delta.
\end{aligned}
$$

Esto implica que

$$\tfrac{f(x'+\varepsilon v)-f(x')}{\varepsilon}\leq \tfrac{f(x+\varepsilon v)-f(x)}{\varepsilon}+2\delta K,$$

como esto sucede para todo x' en $B(x,\varepsilon\delta)$, por la definición de supremo, tenemos que

$$\sup_{\|x-x'\|<\varepsilon\delta}\tfrac{f(x'+\varepsilon v)-f(x')}{\varepsilon}\leq \tfrac{f(x+\varepsilon v)-f(x)}{\varepsilon}+2\delta K,$$

haciendo $\varepsilon\downarrow 0$, se tiene

$$f^{\circ}(x;v)\leq f'(x;v)+2K\delta.$$

Como δ es arbitrario, concluimos que $f^{\circ}(x;v)\leq f'(x;v)$. Con esto la proposición esta completa.

$\square$

Ejemplo. 3. *Sea $f:R^n\to R$ definido como, $f(x_1,...,x_2)=\max\limits_{1\leq i\leq n}\{x_i\}$.*

Esta función es claramente convexa en R^n y como es acotada en $B(0,1)$ es localmente Lipschitz. Hallemos el gradiente generalizado de f; calculemos $f'(x;v)$, primero definamos

$$I(x)=\{i:x_i=f(x)\},$$

observe que este conjunto puede tener mas de un elemento. Sabiendo esto tenemos

$$
\begin{aligned}
f'(x;v) &= \lim_{t\downarrow 0}\tfrac{f(x+tv)-f(x)}{t}\\
&= \lim_{t\downarrow 0}\tfrac{\max\limits_{1\leq i\leq n}\{x_i+tv_i\}-f(x)}{t}\\
&= \lim_{t\downarrow 0}\tfrac{\max\limits_{i\in I(x)}\{x_i+tv_i\}-f(x)}{t}\\
&= \lim_{t\downarrow 0}\tfrac{\max\limits_{i\in I(x)}\{x_i+tv_i\}-\{x_i\}_{i\in I(x)}}{t}
\end{aligned}
$$

$$= \lim_{t\downarrow 0} \frac{\max\limits_{i\in I(x)} \{x_i + tv_i - x_i\}}{t}$$
$$= \max_{i\in I(x)} \{v_i\}$$

Como $f^\circ = f'$, $\zeta \in \partial f(x)$ *ssi*, $\langle \zeta, v\rangle \leq \max\limits_{i\in I(x)} \{v_i\}$.

Fácilmente se ve que $\zeta \in \partial f(x)$ *si, y solo si*, $\zeta_i \geq 0$, $\sum\limits_{I=1}^{n} \zeta_i = 1$ *y* $\zeta_i = 0$ *si*, $i \in I(x)$.

2.3. Cálculos Basicos

Ahora nosotros procederemos a demostrar algunas formulas que faciliten los cálculos de ∂f.

Nosotros seguimos asumiendo que f es Lipschitz cerca de x.

Proposición 10. *Sea cualquier* s *en* R, *uno tiene*

$$\partial(sf)(x) = s\partial f(x).$$

Demostración. Notemos que sf es Lipschitz cerca a x. Veamos cuando s es no negativo, $(sf)^\circ(x;\cdot) = sf^\circ(x;\cdot)$, y se sigue fácilmente que $\partial(sf)(x) = s\partial f(x)$. Cuando $s = -1$ tenemos para cada v en X

$$\begin{aligned}
\zeta \in \partial(-f)(x) &\Leftrightarrow \langle \zeta, v\rangle \leq (-f)^\circ(x; v)\\
&\Leftrightarrow \langle \zeta, v\rangle \leq f^\circ(x; -v)\\
&\Leftrightarrow \langle -\zeta, -v\rangle \leq f^\circ(x; -v)\\
&\Leftrightarrow -\zeta \in \partial f(x)\\
&\Leftrightarrow \zeta \in -\partial f(x),
\end{aligned}$$

con esto $\partial(-f)(x) = -\partial f(x)$. Para cuando $s < 0$, $\partial(sf)(x) = -\partial(-sf)(x) = s\partial f(x)$. Así es para todo s en R.

$\square$

Teorema 1. *Si f alcanza un mínimo o máximo local en x, entonces $0 \in \partial f(x)$.*

Demostración. Sea x un mínimo local de f, entonces para v fijo en X y para un $\delta > 0$ suficientemente pequeño, se tiene que para $t < \delta$

$$\tfrac{f(x+tv)-f(x)}{t} \geq 0,$$

y por la definición de la derivada generalizada tenemos

$$f^{\circ}(x;v) \geq \tfrac{f(x+tv)-f(x)}{t} \geq 0,$$

concluimos que

$$f^{\circ}(x;v) \geq 0 = \langle 0, v\rangle,$$

como v es arbitrario, entonces $0 \in \partial f(x)$.

Si x es un máximo local de f, x es un mínimo local de $-f$, entonces por lo demostrado $0 \in \partial(-f)(x)$, ahora por la proposición 10 $0 \in -\partial f(x)$, lo que implica que $0 \in \partial f(x)$.

$\square$

Si f_i $(i = 1, ..., n)$ es una familia finita de funciones Lipschitz cerca a x, es fácil ver que su suma $f = \sum f_i$ es Lipschitz cerca a x.

Proposición 11.

$$\partial(\sum f_i)(x) \subset \sum \partial f_i(x).$$

Demostración. Fácilmente se ve que, la función de soporte de $\partial f_1(x) + \partial f_2(x)$ es $f_1^{\circ}(x;v) + f_2^{\circ}(x;v)$. Entonces como

$$(f_1 + f_2)^{\circ}(x;v) \leq f_1^{\circ}(x;v) + f_2^{\circ}(x;v),$$

esto nos dice que $\partial(f_1 + f_2)(x) \subset \partial f_1(x) + \partial f_2(x)$. Si es verdad para $n-1$, tendremos

$$\begin{aligned}\partial(\sum_{i=1}^{n} f_i)(x) &= \partial(\sum_{i=1}^{n-1} f_i + f_n)(x)\\ &\subset \partial(\sum_{i=1}^{n-1} f_i)(x) + \partial f_n(x)\end{aligned}$$

$$\begin{aligned} &\subset \sum_{i=1}^{n-1} \partial f_i(x) + \partial f_n(x) \\ &= \sum_{i=1}^{n} \partial f_i(x). \end{aligned}$$

□

Corolario 1

$$\partial(\sum f_i)(x) = \sum \partial f_i(x),$$

si todas las funciones, o a lo mas uno no, son estrictamente diferenciable.

Demostración. Es fácil ver que si calculamos el limite superior de una suma de dos funciones, donde uno si tiene limite, es igual al limite de uno mas el limite superior del otro. Sabiendo esto supongamos que f_1 no estrictamente diferenciable y f_2 si lo es,

$$(f_1 + f_2)^\circ(x; v) = (f_1)^\circ(x; v) + D_s f(x; v) = f_1^\circ(x; v) + f_2^\circ(x; v),$$

así $\partial(f_1 + f_2)(x) = \partial f_1(x) + \partial f_2(x)$, procediendo como antes, se obtiene lo requerido.

□

Corolario 2

Para cualquier escalar s_i, uno tiene

$$\partial(\sum_{i=1}^{n} s_i f_i)(x) \subset \sum_{i=1}^{n-1} s_i \partial f_i(x),$$

la igualdad se cumple si todas las funciones, o a lo mas uno no, son estrictamente diferenciable.

2.4. Regularidad

Como vimos las formulas del gradiente generalizado, implica inclusion. la condición puede ser suavizada para tener la igualdad en algunos casos, pues lo requerido para esto es la diferencial estricta, ahora definiremos el concepto de función regular.

Definición 9. *f se dice regular en x, si cumplen lo siguiente*

(i) Para todo v, la derivada direccional $f'(x; v)$ existe.

(ii) Para todo v, $f'(x; v) = f^{\circ}(x; v)$

Corolario

Si f_i es regular para cada cada $i = 1, ..., n$, entonces

$$\partial(\sum f_i)(x) = \sum \partial f_i(x).$$

También si $s_i \geq 0$ para cada $i = 1, ..., n$, entonces

$$\partial(\sum s_i f_i)(x) = \sum s_i \partial f_i(x).$$

Demostración. En la siguiente proposición veremos que la suma de funciones regulares, es regular. Como antes es suficiente mostrar para $n = 2$. Probaremos la igualdad de sus funciones soporte. En la siguiente igualdad esta la función soporte de $\partial(f_q + f_2)(x)$ a la izquierda, $(f_1 + f_2)^{\circ}(x; \cdot) = (f_1 + f_2)'(x; \cdot) = f_1'(x; \cdot) + f_2'(x; \cdot) = f_1^{\circ}(x; \cdot) + f_2^{\circ}(x; \cdot)$, y a la derecha se encuentra la función soporte de $\partial f_1(x) + \partial f_2(x)$. □

Observación

Para que la igualdad se cumpla para todo s_i debe cumplirse que $-f_i$ sea regular, pero esto no se cumple en general, pues para $f(x) = |x|$ es regular, pero $-f$ no lo es.

Aquí vamos con la primera observación sobre las funciones regulares.

Proposición 12. *Sea f Lipschitz cerca a x.*

(a) Si f es estrictamente diferenciable en x, entonces f es regular en x.
(b) Si f es convexo, entonces f es regular en x.

(c) Una combinación lineal finita (para no negativos escalares) de funciones regulares en x, es regular en x.
(d) Si f admite la Derivada de Gateaux $Df(x)$ y es regular en x, entonces $\partial f(x) = \{Df(x)\}$.

Demostración. (a) Estrictamente diferenciable, implica que existe la derivada direccional y además $f'(x;v) = D_s f(x;v) = f^\circ(x;v)$.

(b) Por el análisis convexo $f'(x;v)$ existe y como f es lipschitz $f'(x;v) = f^\circ(x;v)$.

(c) Probaremos para $n = 2$, el resultado se deduce por inducción, sea $s \geq 0$ y f_1 y f_2 regulares en x, entonces $(sf_1)'(x;v) = sf_1'(x;v) = sf^\circ(x;v) = (sf)^\circ(x;v)$, con esto sf es regular en x, por otro lado $(f_1 + f_2)' = f_1' + f_2' = f_1^\circ + f_2^\circ \geq (f_1 + f_2)^\circ$ y como siempre vale $(f_1 + f_2)' \leq (f_1 + f_2)^\circ$, concluimos que $(f_1 + f_2)' = (f_1 + f_2)^\circ$.
(d) $f^\circ(x;\cdot) = \langle Df(x), \cdot\rangle$ tiene un comportamiento lineal y como vimos cuando sucede esto el gradiente generalizado es único y $\partial f(x) = \{Df(x)\}$.

□

2.5. El Teorema del Valor Medio

Teorema 2. *Sea x, y en X, supongamos que f el Lipschitz en un abierto conteniendo el segmento $[x, y]$. Entonces existe un punto u en (x, y) tal que*

$$f(y) - f(x) \in \langle \partial f(u), y - x\rangle.$$

Para el teorema del valor medio del gradiente generalizado necesitamos un lema donde denotaremos a $x + t(y - x)$ como x_t.

Lema 1. *La función $g : [0, 1] \to R$ definido por $g(t) = f(x_t)$ es Lipschitz en (0,1), y además*

$$\partial g(t) \subset \langle \partial f(x_t), y - x\rangle.$$

Demostración. Notemos que $\partial g(t)$ y$\langle \partial f(x_t), y - x\rangle$ son intervalos en R. Se debe mostrar que

$$\begin{gathered} \text{máx}\, \partial g(t) \leq \text{máx}\langle \partial f(x_t), y - x\rangle, \\ \text{mín}\, \partial g(t) \geq \text{mín}\langle \partial f(x_t), y - x\rangle \\ \text{ó máx}\, \partial g(t)(-1) \geq \text{máx}\langle \partial f(x_t), (y - x)(-1)\rangle, \end{gathered}$$

mostraremos algo mas general, que para un fijo v en R, se tiene que

$$\text{máx}\,\partial\{g(t)v\} \leq \text{máx}\langle \partial f(x_t), (y-x)v)\rangle$$

o equivalentemente, por propiedades del gradiente que

$$g^{\circ}(t;v) \leq f^{\circ}(x_t; v(y-x).$$

En efecto

$$\begin{aligned}
g^{\circ}(t;v) &= \limsup_{\substack{s\to t\\ \lambda\downarrow 0}} \tfrac{g(s+\lambda v)-g(s)}{\lambda}\\
&= \limsup_{\substack{s\to t\\ \lambda\downarrow 0}} \tfrac{f(x+(s+\lambda v)(y-x))-f(x_s)}{\lambda}\\
&= \limsup_{\substack{s\to t\\ \lambda\downarrow 0}} \tfrac{f(x_s+\lambda v(y-x))-f(x_s)}{\lambda};\ \text{cuando } s\to t,\ x_s\to x_t\\
&\leq \limsup_{\substack{s\to t\\ \lambda\downarrow 0}} \tfrac{f(y+\lambda v(y-x))-f(y)}{\lambda}\\
&= f^{\circ}(x_t; v(y-x)).
\end{aligned}$$

□

Demostracion del Teorema

Sea $\theta : [0,1] \to R$ definido como $\theta(t) = f(x_t) + (f(x) - f(y))t$, claramente $\theta(0) = \theta(1) = f(x)$. Como θ es continua existe t_0 en $(0,1)$ tal que $\theta(t_0)$ es máximo o mínimo, entonces $0 \in \partial\theta(t_0)$, por otro lado

$$\begin{aligned}
\partial\theta(t) &\subset \partial g(t) + (f(x) - f(y))\partial I(t)\\
&\subset \langle \partial f(x_t), y-x\rangle + (f(x) - f(y)),
\end{aligned}$$

para $t = t_0$, se tiene

$$0 \in (f(x) - f(y)) + \langle \partial f(x_{t_0}), y - x\rangle,$$

tomando $u = x_{t_0}$, tenemos que

$$(f(x) - f(y)) \in \langle \partial f(u), y - x\rangle. \ \square$$

2.6. La Regla de la Cadena

Teorema 3. *Sea $F : X \to R^n$, $g : R^n \to R$, supongamos que F es estrictamente diferenciable en x y que g es Lipschitz cerca a $F(x)$. Entonces $f = g \circ F$ es Lipschitz cerca a x, y uno tiene*

$$\partial f(x) \subset \partial(F(x)) \circ D_s F(x).$$

Demostración. Como se sabe la composición de funciones Lipschitz, es Lipschitz. Notemos que la inclusion arriba es una inclusion entre convexos y débilmente*-compactos. A la función soporte de $\partial(F(x)) \circ D_s F(x)$ denotemosle como $\sigma(\cdot)$, ahora debemos mostrar para todo v en X que

$$f^{\circ}(x; v) \leq \sigma(v),$$

sea un v fijo, entonces por definición

$$\sigma(v) = \text{máx}\{\langle \zeta, D_s F(x) v \rangle : \zeta \in \partial g(F(x))\},$$

ahora por definición para todo n en N, existe $x_n \in B(x, \frac{1}{n})$ y $t_n < \frac{1}{n}$, tal que

$$f^{\circ}(x; v) - \tfrac{1}{n} \leq \tfrac{f(x_n + t_n v) - f(x_n)}{t_n} = \tfrac{g(F(x_n + t_n v)) - g(F(x_n))}{t_n}$$

por el teorema del valor medio, existe u_n en $[F(x_n + t_n v), F(x_n)]$ tal que

$$f^{\circ}(x; v) - \tfrac{1}{n} \leq \langle \zeta_n, \tfrac{(F(x_n + t_n v) - F(x_n))}{t_n} \rangle,$$

para algún $\zeta_n \in \partial g(u_n)$, claramente $u_n \to F(x)$, por la semicontinuidad de ∂g.existe ζ en $\partial g(F(x))$ tal que $\zeta_n \to \zeta$, así haciendo $n \to \infty$, tenemos

$$f^{\circ}(x; v) \leq \langle \zeta, D_s F(x) \cdot v \rangle,$$

así, entonces $f^{\circ}(x; v) \leq \sigma(v)$ como queríamos.

□

Ejemplo. 4. *Supongamos que el siguiente conjunto es dado* $\{(x_0, y_0), (x_1, y_1), ..., (x_N, y_N)\} \subset R^2$ *considerando el problema de determinar la linea recta en* R^2*, tal que este recta sea lo mas cerca posible a cada punto (supondremos que los puntos no son colineales, si fueran colineales el problema ya estaría resuelto). Para la linea* $y = mx+b$ *el error de proximidad para* (x_i, y_i) *sera* $e_i = |mx_i + b - y_i|$*. El problema podemos verlo como, el problema de minimizar* $\sum\limits_{i=0}^{N} e_i$*, variando el ángulo* m *y la ordenada* b*, consideraremos la función* $f : R^2 \to R$ *definida como*

$$f(\alpha, \beta) = \sum_{i=0}^{N} |\alpha x_i + \beta - y_i|.$$

Específicamente, para N+1 puntos dados $(0, 0), (1, 1), ..., (N-1, N-1)$ *junto con* $(N, 0)$*, obtenemos la funciona a minimizar dada por*

$$f(\alpha, \beta) = |\alpha N + \beta| + \sum_{i=0}^{N-1} |\alpha i + \beta - i|.$$

Notemos primero que la función $f_{c,k}$ *definido por*

$$f_{c,k}(\alpha, \beta) = |\alpha c + \beta - k|$$

es la composición de g *y* f*, donde*

$$g(y) = |y|,\ F(\alpha, \beta) = \alpha c + \beta - k.$$

Notemos que F *es estrictamente diferenciable, con* $D_s F(\alpha, \beta) = (c, 1)$*. Además* g *es regular, se sigue de la regla de la cadena que* $\partial f_{c,k}(\alpha, \beta) = \partial g(F(x)) \cdot D_s F(x)$*, así uno tiene que*

-Si $\alpha c + \beta - k > 0$*, entonces* $\partial f_{c,k}(\alpha, \beta) = \{(c, 1)\}$*,*

-Si $\alpha c + \beta - k < 0$*, entonces* $\partial f_{c,k}(\alpha, \beta) = \{(-c, -1)\}$ *y*

-Si $\alpha c + \beta - k = 0$*, entonces* $\partial f_{c,k}(\alpha, \beta) = \{\lambda(-c, -1) : |\lambda| \leq 1\}$*.*

Existe un punto (α, β) *que es mínimo de* f*, además sabemos que* $0 \in \partial f(\alpha, \beta)$*, ahora si llamamos* $f_0(\alpha, \beta) = |\alpha N + \beta|$ *y* $f_i(\alpha, \beta) = |\alpha i + \beta - i|$*, entonces*

$$f(\alpha, \beta) = \textstyle\sum_{i=0}^{N-1} f_i(\alpha, \beta),$$

y $\partial f(\alpha,\beta) = \sum_{i=0}^{N-1} \partial f_i(\alpha,\beta)$, *asi* $0 \in \sum_{i=0}^{N-1} \partial f_i(\alpha,\beta)$, *esto es*

$$0 = \lambda_N(N,1) + \sum_{i=0}^{N-1} \lambda_i(i,1)$$

Si esta condición necesaria vale para $y = x$ *(es decir para* $\alpha = 1, \beta = 0$*), entonces* $\partial f_{N,0}(1,0) = \{(N,0)\}$*, así por la anterior ecuación, tenemos que* $\lambda_N = 1$ *y* $|\lambda_i| \leq 1$ *para* $i = 0, ..., n-1$*. Fácilmente se ve que si* $N \geq 3$*, entonces existen* λ_i'*s que satisfacen nuestra ecuación. Como* f *es convexo, la condición* $0 \in \partial f(1,0)$ *es también suficiente, lo cual implica que* $y = x$ *es una solución del problema.*

Capítulo 3

Una Definición Extendida de ∂f

3.1. Función Distancia

Definición 10. *Sea C un subconjunto de X no vació, consideremos la funcion distancia ; $d_C : X \to R$, definido por*

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

La función distancia es ciertamente no diferenciable en el sentido clásico, pero es globalmente Lipschitz de grado 1, como probaremos. Daremos una definición de normales, tangentes, etc. Finalmente usando las ideas geométricas para extender la definición de gradiente generalizado ∂f de una función f el cual no es necesariamente localmente Lipschitz, y posiblemente extender valores.

Proposición 13. *La función distancia d_C es localmente Lipschitz en X, es decir, para todo x,y en X se tiene que*

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

Demostración. Para $\varepsilon > 0$, existe $c \in C$ tal que

$$\|y - c\| \leq d_C(x) + \varepsilon,$$

además

$$d_C(x) \leq \|x - c\| \leq \|x - y\| + \|y - c\| \leq \|x - y\| + d_C(x) + \varepsilon,$$
$$d_C(x) - d_C(x) \leq \|x - y\| + \varepsilon,$$

como ε es arbitrario

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

del mismo modo intercambiando x por y y viceversa tenemos que

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

$\square$

3.2. Tangentes

Definición 11. *Supongamos ahora que x es un punto de C. Un vector v en V es tangente a C en x, si $d_C^\circ(x; v) = 0$. El conjunto de los tangentes a C en x lo denotaremos con $T_C(x)$, es decir*

$$T_C(x) = \{v \in X : d_C^\circ(x; v) = 0\}.$$

$T_C(x)$ es un cono convexo y cerrado en X (en particular, $T_C(x)$ siempre contiene al 0).En efecto Sea v en $T_C(x)$ y $t \geq 0$, entonces $d_C^\circ(x; tv) = t d_C^\circ(x; v) = 0$, con esto $tv \in T_C(x)$ y $T_C(x)$ es un cono. Sea (v_i) una sucesión en $T_C(x)$ convergiendo a v, como $d_C^\circ(x; \cdot)$ es continua, tenemos que

$$d_C^\circ(x; v) = \lim_{i \to \infty} d_C^\circ(x; v_i) = 0,$$

así $v \in T_C(x)$ y $T_C(x)$ es cerrado en X. Sea v, w en $T_C(x)$, solo basta mostrar que $d_C^\circ(x; v + w) = 0$ para la convexidad ([4]),

$$d_C^\circ(x; v + w) \leq d_C^\circ(x; v) + d_C^\circ(x; w) = 0,$$

para el otro sentido, por definición

$$\begin{aligned} d_C^\circ(x; v + w) &= \limsup_{\substack{y \to x \\ t \downarrow 0}} \tfrac{d_C(y + tv + tw) - d_C(y)}{t} \\ &\geq \limsup_{t \to 0} \tfrac{d_C(x + tv + tw) - d_C(x)}{t} \geq 0. \end{aligned}$$

3.3. Normales

Definición 12. *Nosotros definimos el Cono Normal a C en x, como*

$$N_C(x) = \{\zeta \in X^* : \langle \zeta, v\rangle \leq 0 \text{ para todo } v \text{ en } T_c(x)\}.$$

Proposición 14. *a) $N_C(x)$ es un cono convexo y débilmente*-cerrado,*
b) $N_C(x) = cl^*\{\bigcup_{\lambda\geq 0} \lambda\partial d_C(x)\}$ y
c) $T_C(x)$ es el cono polar de $N_C(x)$, esto es

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

donde cl^* denota la clausura en la topología débil.*

Demostración. *a*) Facilmente se ve esto.
b) Sea ζ en $\partial d_C(x)$ y v en $T_C(x)$, entonces $\langle \zeta, v\rangle \leq d_C^\circ(x; v) = 0$, con esto ζ esta en $N_C(x)$ y fácilmente se ve que $\lambda\zeta$ esta en $N_C(x)$, para $\lambda \geq 0$, por tanto $\bigcup_{\lambda\geq 0} \lambda\partial d_C(x) \subset N_C(x)$, como $N_C(x)$ es cerrado, $cl\{\bigcup_{\lambda\geq 0} \lambda\partial d_C(x)\} \subset N_C(x)$. Por otro llamemos $H = cl\{\bigcup_{\lambda\geq 0} \lambda\partial d_C(x)\}$, sea $\zeta \notin H$, entonces por los teoremas de separación, existe v en X tal que

$$\langle \zeta, v\rangle > \langle \zeta', v\rangle,$$

para todo ζ' en H, además $\langle \zeta', v\rangle \leq 0$, pues si existiera ζ' en H con $\langle \zeta', v\rangle > 0$, entonces como H es un cono $\langle \zeta, v\rangle > \langle t\zeta', v\rangle > 0$ y cuando $t \to \infty$, entonces $\langle \zeta, v\rangle = \infty$ lo cual es una contradicción,con esto podemos decir que

$$d_C^\circ(x; v) = \{\langle \zeta'', v\rangle >: \zeta'' \in \partial d_C(x)\} \leq 0,$$

pero como siempre $d_C^\circ(x; v) \geq 0$, entonces $d_C^\circ(x; v) = 0$ y $v \in T_C(x)$. Por otro lado como $0 \in H$, entonces $\langle \zeta, v\rangle > 0$ y con esto $\zeta \notin N_C(x)$, lo que demuestra el teorema.
c) Sea v en $T_C(x)$, entonces $d_C^\circ(x; v) = 0$ y por la definición del gradiente generalizado tenemos que

$$\langle \zeta, v\rangle \leq d_C^\circ(x; v) = 0,$$

para todo ζ en $\partial d_C(x)$, entonces

$$\langle \zeta, v\rangle \leq 0,$$

para todo ζ en $N_C(x)$, lo que implica que v pertenece a $N_C(x)^\circ$. Recíprocamente sea v en $N_C(x)^\circ$, entonces $\langle \zeta, v\rangle \leq 0$, para todo ζ en $\partial d_C(x)$, (por b)), ahora por la definición de gradiente generalizado $d_C^\circ(x; v) \leq 0$ y como $d_C^\circ(x; v) \geq 0$, entonces $d_C^\circ(x; v) = 0$, así v esta en $T_C(x)$.

□

Proposición 15. *Sea f Lipschitz de grado K en un conjunto S. Sea x en $C \subset S$ y supongamos que f contiene un mínimo sobre C en x. Entonces para cualquier $K' \geq K$, la función $g(y) = f(y) + K'd_C(y)$ contiene un mínimo sobre S en x. Si $K' > K$ y C es cerrado, entonces cualquier otro punto minimizando g sobre S debe estar en C.*

Demostración. Si suponemos que existe $K' \geq K$ y y en S tal que

$$\begin{gathered} g_{K'}(x) > g_{K'}(y) \\ f(x) > f(y) + K'd_C(y), \end{gathered}$$

además existe $\varepsilon > 0$ tal que

$$f(x) - K'\varepsilon > f(y) + K'd_C(y),$$

para este ε, existe c en C tal que $\|y - c\| \leq d_C(y) + \varepsilon$, como f Lipschitz entonces

$$f(c) \leq f(y) + K'\|y - c\| \leq f(y) + (d_C(y) + \varepsilon) < f(x),$$

lo cual es una contradicción a que x es un mínimo sobre C.
Sea $K' > K$ y y en S un mínimo de $g_{K'}$ sobre S, como $\frac{K'+K}{2} > K$, por lo demostrado

$$\begin{gathered} g_{\frac{K'+K}{2}}(x) \leq g_{\frac{K'+K}{2}}(y), \\ f(x) \leq f(y) + \tfrac{K'+K}{2} d_C(y), \end{gathered}$$

como $g_{K'}(y) \leq g_{K'}(x)$, entonces

$$\begin{gathered} f(y) + K'd_C(y) \leq f(x) \leq f(y) + \tfrac{K'+K}{2} d_C(y), \\ K'd_C(y) \leq \tfrac{K'+K}{2} d_C(y), \end{gathered}$$

pero esta ultima desigualdad se cumple, si $d_C(y) = 0$ lo cual nos dice que y esta en C, por ser C cerrado.

□

Corolario

Sea f Lipschitz cerca a x y conteniendo un mínimo sobre C en x. Entonces $0 \in \partial f(x) + N_C(x)$.

Demostración. Supongamos que f es Lipschitz en $S \ni x$ de grado K, $(N_{C\cap S} = N_C)$ supondremos que $C \subset S$, por la proposición x minimiza $f + Kd_C$. Así

$$0 \in \partial(f + Kd_C)(x) \subset \partial f(x) + K\partial d_C(x) \subset \partial f(x) + N_C(x).$$

□

Cuando C es convexo, hay un conocido concepto de vector normal: $\zeta \in X^*$ es normal a C en x (en el caso del análisis conexo) si

$$\langle \zeta, x - c\rangle \geq 0 \text{ ; para todo } c \text{ en } C,$$

aquí denotaremos como $N'_C(x)$ al cono de normales a C en x.

Lema 2. *Sea $C \subset X$ convexo, entonces $d_C(\cdot)$ es convexo.*

Demostración. Sea x, y en X y λ en $(0, 1)$. Para todo $\varepsilon > 0$, existe c_x, c_y en C tal que

$$\|c_x - x\| \leq d_C(x) + \varepsilon \text{ , } \|c_y - y\| \leq d_C(y) + \varepsilon,$$

sea $c = \lambda c_x + (1 - \lambda)c_y$ en C, pues C es convexo, por la definición de d_C

$$\begin{aligned} d_C(\lambda x + (1-\lambda)y) &\leq \|c - \lambda x - (1-\lambda)y\| \\ &= \|\lambda c_x + (1-\lambda)c_y - \lambda x - (1-\lambda)y\| \\ &\leq \lambda\|c_x - x\| + (1-\lambda)\|c_y - y\| \\ &\leq \lambda d_C(x) + \lambda\varepsilon + (1-\lambda)d_C(y) + (1-\lambda)\varepsilon \\ &= \lambda d_C(x) + (1-\lambda)d_C(y) + \varepsilon, \end{aligned}$$

como ε es arbitrario, se tiene lo deseado.

$\square$

Proposición 16. *Si C es convexo, $N_C(x) = N'_C(x)$.*

Demostración. Sea ζ en $N'_C(x)$ y $f(y) = \langle \zeta, x - y \rangle$, es claro que x minimiza a f sobre C, por el corolario anterior

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

f es estrictamente diferenciable y además $\langle D_s f(x), v \rangle = \langle -\zeta, v \rangle$, como ya vimos $\partial f(x)$ es unitario y $\partial f(x) = \{-\zeta\}$, y luego

$$0 \in -\zeta + N_C(x),$$

entonces $\zeta \in N_C(x)$, con esto $N'_C(x) \subset N_C(x)$. Por otro lado sea $\zeta \in \partial d_C(x)$, como d_C es convexo, el subdiferencial y el gradiente generalizado coinciden, y $\zeta \in \partial_s d_C(x)$, entonces por definición del subdiferencial

$$\langle \zeta, y - x \rangle \leq d_C(y) - d_C(x) \text{ ; para todo } y \in X,$$

como x esta en C

$$\langle \zeta, y - x \rangle \leq d_C(y),$$

lo anterior para todo y en C, tenemos que

$$\langle \zeta, y - x \rangle \leq 0,$$

lo que nos dice que ζ esta en $N'_C(x)$, y $\partial d_C(x) \subset N'_C(x)$, como $N'_C(x)$ es un cono convexo y cerrado

$$N_C(x) = cl\{\bigcup_{\lambda \geq 0} \lambda \partial d_C(x)\} \subset N'_C(x).$$

$\square$

Corolario

Si C es convexo, entonces v $T_C(x)$ ssi $d^{\circ}_C(x; v) = d'_C(x; v) = 0$.

3.4. Una Definición Intrínseca de $T_C(x)$

Nosotros ahora mostraremos que el concepto de tangente definido anteriormente es independiente de la norma (y así de la función distancia) usado en X. Sabiendo esto cambiaremos, en particulares circunstancias, la forma de calcular tangentes y normales.

Teorema 4. *Un elemento v de X es tangente a C en x sii, para cada sucesion x_i en C convergiendo a x y cada sucesion t_i en $(0, \infty)$ decresiendo a 0, existe una sucesion v_i en X convergiendo a v tal que $x_i + t_i v_i \in C$ para todo i.*

Demostración. Sea $v \in T_C(x)$, la primera desigualdad se sigue por propiedades del limite superior ([2])

$$0 = d_C^\circ(x; v) \geq \lim_{i\to\infty} \tfrac{d_C(x_i+t_iv)-d_C(x_i)}{t_i} = \lim_{i\to\infty} \tfrac{d_C(x_i+t_iv)}{t_i} \geq 0,$$

entonces

$$\lim_{i\to\infty} \tfrac{d_C(x_i+t_iv)}{t_i} \geq 0,$$

para $\frac{t_i}{i} > 0$, existe c_i en C tal que

$$\|x_i + t_iv - c_i\| \leq d_C(x_i + t_iv) + \tfrac{t_i}{i},$$

tomemos

$$v_i = \tfrac{c_i - x_i}{t_i},$$

$\|v - v_i\| \to 0$, en efecto

$$\|v - \tfrac{c_i-x_i}{t_i}\| \leq \tfrac{d_C(x_i+t_iv)}{t_i} + \tfrac{1}{i},$$

cuando $i \to \infty$, $v_i \to v$, además $c_i = x_i + t_iv_i \in C$ como queríamos.
Por otro lado, por propiedades del limite superior ([2]), existe y_i en X y $t_i \downarrow 0$ tal que

$$\lim_{i\to\infty} \tfrac{d_C(y_i+t_iv)-d_C(y_i)}{t_i} = d_C^\circ(x) \geq 0,$$

necesitamos probar que no es positivo y así v estará en $T_C(x)$. Para $\frac{t_i}{i} > 0$, existe c_i en C tal que

$$\|c_i - y_i\| \leq d_C(y_i) + \tfrac{t_i}{i},$$

utilizando la diferencia triangular claramente cuando $i \to \infty$, $c_i \to x$, por la hipótesis existe $v_i \to v$ tal que $c_i + t_i v_i \in C$, y como d_C el Lipschitz de grado 1

$$\begin{aligned} d_C(y_i + t_i v) - d_C(y_i + t_i v_i) &\leq \|y_i - c_i\| + t_i\|v - v_i\| \\ d_C(y_i + t_i v) &\leq d_C(y_i) + \tfrac{t_i}{i} + t_i\|v - v_i\| \\ \tfrac{d_C(y_i + t_i v) - d_C(y_i)}{t_i} &\leq \|v - v_i\| + \tfrac{1}{i}, \end{aligned}$$

cuando $i \to \infty$, $d_C^\circ(x; v) \leq 0$.

$\square$

3.5. Regularidad de Conjuntos

Definición 13. *Definimos el Cono Contingente, denotado por $K_C(x)$, al conjunto de los vectores tangentes a C en x. Un vector v en X pertenece a $K_C(x)$ sii, para todo $\varepsilon > 0$, existe t en $(0, \varepsilon)$ y un punto w en $B(v, \varepsilon)$ tal que $x + tw \in C$ (necesariamente x esta en cl$C$). Por el anterior teorema, siempre se tiene que $T_C(x) \subset K_C(x)$.*

Definición 14. *El conjunto C es Regular en x si, $T_C(x) = K_C(x)$.*

Observación

Cualquier conjunto convexo es regular en cada punto. En efecto, solo basta mostrar que $K_C(x) \subset T_C(x)$, para esto sea v en $K_C(x)$, entonces para $\frac{1}{n} > 0$, existe $t_n \in (0, \frac{1}{n})$ y $w_n \in B(v, \frac{1}{n})$ tal que $x + t_n w_n \in C$, calculemos $d'_C(x, v)$

$$\begin{aligned} d'_C(x, v) &= \lim_{t \downarrow 0} \tfrac{d_C(x+tv) - d_C(x)}{t} \\ &= \lim_{t \downarrow 0} \tfrac{d_C(x+tv)}{t} \\ &= \lim_{t \downarrow 0} \tfrac{d_C(x+t_n w_n)}{t_n} = 0, \end{aligned}$$

así v esta en $T_C(x)$.

Teorema 5. *Sea f Lipschitz cerca a x, y suponga que $0 \notin \partial f(x)$. Si definimos $C = \{y \in X \ : \ f(y) \leq f(x)\}$, entonces uno tiene que*

$$\{v \in X \ : \ f^\circ(x;v) \leq 0\} \subset T_C(x).$$

Si f es regular en x, entonces sea da igualdad y C es regular en x.

Demostración. Primero observemos que existe v' en X tal que $f^\circ(x;v') < 0$, pues $0 \notin \partial f(x)$. Si v es tal que $f^\circ(x;v) \leq 0$, entonces para todo $\varepsilon > 0$,

$$f^\circ(x; v + \varepsilon v') \leq f^\circ(x;v) + \varepsilon f^\circ(x;v') < 0,$$

en consecuencia es suficiente mostrar que para cualquier v, para el cual $f^\circ(x;v') < 0$, esta en $T_C(x)$ y es así como procederemos.
Sea $f^\circ(x;v') < 0$, por la definición de f°, existe $\varepsilon, \delta > 0$ tal que para $y \in B(x,\varepsilon)$ y $t \in (0,\varepsilon)$, se tiene que

$$f(y + tv) - f(y) \leq -\delta t.$$

Sea x_i cualquier sucesión en C convergiendo a x, y t_i cualquier sucesión decreciendo a 0. Por la definición de C, nosotros tenemos $f(x_i) \leq f(x)$, para un i suficientemente grade

$$f(x_i + t_i v) - f(x_i) \leq -\delta t_i,$$
$$f(x_i + t_i v) \leq f(x_i) - \delta t_i \leq f(x) - \delta t_i < f(x).$$

Se sigue que $x_i + t_i v$ pertenece a C (para i suficientemente grande), y lo que establece que $v \in T_C(x)$.
Ahora nosotros suponemos que f es regular en x. Probemos que si $v \in K_C(x)$ entonces $f^\circ(x;v') \leq 0$, así conseguiremos la igualdad y la regularidad del conjunto en x. Empezamos si $v \in K_C(x)$, para $\frac{1}{n} > 0$, existe $t_n \in (0, \frac{1}{n})$ y $v_n \in B(v, \frac{1}{n})$ tal que $x + t_n v_n \in C$, con esto tenemos que

$$d_C(x + t_n v) \leq d_C(x + t_n v_n) + \|x + t_n v_n - x - t_n v\| = t_n \|v_n - v\| < t_n \varepsilon < 2t_n \varepsilon,$$

por al definición de ínfimo, existe x_i en C tal que

$$\|x + t_n v - x_n\| < 2t_n \varepsilon.$$

ahora por ser f Lipschitz

$$f(x + t_n v) - f(x_n) \leq K\|x + t_n v - x_n\| < 2K t_n \varepsilon,$$

entonces haciendo operaciones elementales

$$\frac{f(x+t_nv)-f(x)}{t_n} \leq \frac{f(x_n)+2Kt_n\varepsilon - f(x)}{t_n} \leq 2K\varepsilon,$$

cuando $n \to \infty$ el lado izquierdo se transforma en $f'(x;v)$, entonces $f'(x;v) \leq 2K\varepsilon$, como ε es arbitrario, tenemos que $f'(x;v) \leq 0$, como queríamos.

$\square$

3.6. Epígrafos y Funciones no Lipschitz

Definición 15. *Definimos el epígrafo de f (denotado como $epif$) como sigue*

$$epif = \{(x,r) \in X \times R : f(x) \leq r\}.$$

Proposición 17. *Sea f Lipschitz cerca a x. Entonces*

$$T_{epi\ f}(x, f(x)) = epif^{\circ}(x;\cdot).$$

Demostración. Sea (v,r) en $T_{epif}(x,f(x))$, existen $y_i \to x$ y $t_i \to 0$ tal que

$$\lim_{i\to\infty} \frac{f(y_i+t_iv)-f(y_i)}{t_i} = f^{\circ}(x;v),$$

ademas para $(y_i, f(y_i)) \to (x, f(x))$ y $t_i \to 0$, por el teorema 4 existe $(v_i, r_i) \to (v, r)$, con $(y_i, f(y_i)) + t_i(v_i, r_i)$ en $epif$, entonces

$$f(y_i+t_iv) \leq f(y_i) + t_ir_i \Rightarrow \frac{f(y_i+t_iv)-f(y_i)}{t_i} \leq r_i,$$

ahora haciendo $i \to \infty$ tenemos

$$f^{\circ}(x;v) \leq r,$$

así (v,r)esta en $epif^{\circ}(x;\cdot)$.

Para el reciproco mostraremos que para cualquier v, y para cualquier $\delta > 0$, el punto $(v, f^{\circ}(x;v)+\delta)$ esta en $T_{epif}(x,f(x))$. Para esto usaremos el recíproco del teorema 4, sea (x_i, r_i) cualquier secuencia en $epif$ con $(x_i,r_i) \to (x, f(x))$ y $t_i \downarrow 0$. Nosotros ahora debemos mostrar que existe $(v_i, s_i) \to (v, f^{\circ}(x;v)+\delta)$ y $(x_i,r_i)+t_i(v_i,s_i)$ esta en $epif$ para cada i.

Para este fin definamos

$$v_i = v \text{ y } s_i = \text{máx}\{f^\circ(x;v) + \delta, \tfrac{f(x_i+t_iv)-f(x_i)}{t_i}\},$$

es claro que $(v_i, s_i) \to (v, f^\circ(x;v) + \delta)$. Por otro lado nosotros tenemos

$$r_i + t_i s_i \geq r_i + [f(x_i + t_i v) - f(x_i)] \geq f(x_i + t_i v),$$

con esto $(x_i + t_i v, r_i + t_i s_i) = (x_i, r_i) + t_i(v_i, s_i)$ esta en $epif$ y con esto terminamos.

$\square$

Corolario

$$\zeta \in \partial f(x) \Leftrightarrow (\zeta, -1) \in N_{epif}(x, f(x)).$$

Demostración. Sea ζ en $\partial f(x)$ ssi, para todo v en X $\langle \zeta, v \rangle \leq f^\circ(x;v)$ sii, para todo (v,r) en $epif^\circ(x;\cdot) = T_{epif}(x, f(x))$, nosotros tenemos

$$\langle (\zeta, -1), (v, r) \rangle \leq 0.$$

Asi $(\zeta, -1)$ vive en $N_{epif}(x, f(x))$.

$\square$

3.7. Definición Extendida de ∂f

El corolario arriba, el cual es analogo al hecho que $(f'(x), -1)$ es normal al grafo de la funcion suave $f : R \to R$, gracias al corolario de la proposicion 18, extenderemos la definicion del gradiente generalizado, el cual sera consistente con nuestra definicion.

Definición 16. *Sea $f : X \to R \cup \{\pm\infty\}$ finito en el punto x. Nosotros definimos $\partial f(x)$ como el conjunto de todos los ζ en X^* tal que $(\zeta, -1)$ esta en $N_{epif}(x, f(x))$.*

Se sigue que $\partial f(x)$ es un subconjunto debilmente*-cerrado de X^*, el cual no podria ser compacto y podria ser vacio, como muestra el ejemplo adelante.

Ejemplo. 5. *Sea $\sigma : R \to R$ la funcion signo, entonces $\partial\sigma(0) = \phi$*

Proposición 18. *Sea $f : X \to R \cup \{\pm\infty\}$ finito en el punto x, y supongamos que x es un minimo local de f. Entonces $0 \in \partial f(x)$.*

Demostración. Debemos mostrar que $(0,-1)$ esta en $N_{epif}(x, f(x))$, o equivalentemente, que $\langle (0,-1), (v,r) \rangle \leq 0$ para todo (v,r) en $T_{epif}(x, f(x))$, o lo que es lo mismo que $r \geq 0$ para todo (v,r) en $T_{epif}(x, f(x))$. Sea (v,r) en $T_{epif}(x, f(x))$ por el teorema 4, para $(x, f(x))$ en $epif$ y $t_i \downarrow 0$, existe $(v_i, r_i) \to (v,r)$ tal que $(x, f(x)) + t_i(v_i, r_i) \in epif$ esto es

$$f(x) + t_i r_i \geq f(x + t_i v_i),$$

como x en un minimo local, para un i suficientemente grande se tiene

$$f(x) + t_i r_i \geq f(x + t_i v_i) \geq f(x),$$

entonces $t_i r_i \geq 0$, lo que implica que $r_i \geq 0$. Se sigue que $\lim r_i = r \geq 0$. Como queriamos.

□

Terminaremos dando un ejemplo de que la derivada generalizada de la suma, no es igual a la suma de las derivadas generalizadas.

Ejemplo. 6. *Sea $f : R \to R$ definido como $f(x) = |x|$, como f es convexo, entonces*

$$f^{\circ}(0; v) = \lim_{t \downarrow 0} \frac{f(tv) - f(0)}{t} = |v|,$$

ademas por la proposicion 1, $(-f)^{\circ}(0; v) = f^{\circ}(0; -v) = |-v| = |v|$, asi para $v \neq 0$, tenemos

$$(f + (-f))^{\circ}(0; v) = 0 < 2|v| = f^{\circ}(0; v) + (-f)^{\circ}(0; v)$$

Apéndice

El Limite Superior

Definición

Sea $f: X \to R$ acotada, donde X es un espacio de Banach y $B_\delta = B^\circ(x,\delta)$ subconjunto de X, tomemos $A = \{supf(B_\delta) : \delta > 0\}$, entonces el *limite superior* de f en x es el infimo de A, es decir

$$\lim_{y\to x}\sup f(y) = inf A.$$

Notemos que $\delta \mapsto supf(B_\delta)$ es creciente y con esto tenemos que $\lim_{y\to x}\sup f(y) = \lim_{\delta\to 0}\sup f(B_\delta)$.

Propiedades

Sea $f, g : X \to R$ y $\lambda > 0$, entonces

$i)$ $\lim_{y\to x}\sup(f+g)(y) \leq \lim_{y\to x}\sup f(y) + \lim_{y\to x}\sup f(y)$,

$ii)$ $\lim_{y\to x}\sup(\lambda f)(y) = \lambda \lim_{y\to x}\sup f(y)$ y

$iii)$ Si $|f(y)| \leq K$, para todo y en X, entonces $|\lim_{y\to x}\sup f(y)| \leq K$.

Demostración. $i)$ Fácilmente se ve que

$$\sup(f+g)(B_\delta) \leq \sup f(B_\delta) + \sup f(B_\delta),$$

aplicando limites tenemos

$$\lim_{\delta\to 0}\sup(f+g)(B_\delta) \leq \lim_{\delta\to 0}\sup f(B_\delta) + \lim_{\delta\to 0}\sup f(B_\delta),$$

lo que por definición es

$$\limsup_{y\to x}(f+g)(y) \leq \limsup_{y\to x} f(y) + \limsup_{y\to x} f(y)$$

ii) Como $\lambda > 0$, tenemos que

$$\sup(\lambda f)(B_\delta) = \lambda \sup f(B_\delta),$$

aplicando limites tenemos

$$\lim_{\delta\to 0}\sup(\lambda f)(B_\delta) = \lim_{\delta\to 0}\lambda \sup f(B_\delta),$$

por definición

$$\limsup_{y\to x}(\lambda f)(y) = \lambda \limsup_{y\to x} f(y)$$

iii) Como $|f(y)| \leq K$, para todo y en X, entonces $-K \leq \sup f(B_\delta) \leq K$, aplicando limites, tenemos $-K \leq \lim_{\delta\to 0}\sup f(B_\delta) \leq K$, que equivale a $|\limsup_{y\to x} f(y)| \leq K$.

$\square$

Limite Superior de una Sucesión

Definición

Sea (x_n) una sucesión acotada en R y sea $X_n = \{x_n, x_{n+1}, x_{n+2}, ...\}$, llamemos $a_n = \sup X_n$, entonces *ellimite superior* de (x_n) es el limite de (a_n), es decir

$$\limsup_{n\to\infty} x_n = \lim_{n\to\infty} a_n$$

Propiedades

Sea (x_n) y (y_n) sucesiones en R, entonces

i) $\limsup_{n\to\infty}(x_n + y_n) \leq \limsup_{n\to\infty}(x_n) + \limsup_{n\to\infty}(y_n)$

ii) Si $x_n \leq y_n$, para todo n en N, entonces $\limsup_{n\to\infty}(x_n) \leq \limsup_{n\to\infty}(y_n)$.

Demostración. *i*) Sea $(X+Y)_n = \{x_n + y_n, x_{n+1} + y_{n+1}, x_{n+2} + y_{n+2}, ...\}$ y $Y_n = \{y_n, y_{n+1}, y_{n+2}, ...\}$, es claro que

$$\sup(X+Y)_n \leq \sup X_n + \sup Y_n,$$

aplicando limites

$$\lim_{n\to\infty} \sup(X+Y)_n \leq \lim_{n\to\infty} \sup X_n + \lim_{n\to\infty} \sup Y_n,$$

lo que equivale a

$$\limsup_{n\to\infty}(x_n + y_n) \leq \limsup_{n\to\infty}(x_n) + \limsup_{n\to\infty}(y_n)$$

ii) Como $x_n \leq y_n$, para todo n en N, entonces

$$\sup X_n \leq \sup Y_n,$$

similarmente a *i*), aplicando limites concluimos que

$$\limsup_{n\to\infty}(x_n) \leq \limsup_{n\to\infty}(y_n).$$

□

Afirmación

Sea $f : X \to R$ acotada y (x_n) en X, si $x_n \to a$, entonces $\limsup_{n\to\infty} f(x_n) \leq \limsup_{y\to a} f(y)$.

Demostración. Para cada $\delta > 0$, existe n_0 en N, tal que para todo $n > n_0$, se tiene que $|x_n - a| < \delta$ o lo que es lo mismo que $x_i \in B_\delta$, para todo $i \geq n$, entonces $f(x_i) \in f(B_\delta)$ y $f(x_i) \leq \sup f(B_\delta)$, si tomamos $X_i = \{f(x_i).f(x_{i+1}), ...\}$, tenemos que $\sup X_i \leq \sup f(B_\delta)$, como $\limsup_{n\to\infty} f(x_n)$ es el ínfimo de $\{\sup X_i\}$, se tiene que

$$\limsup_{n\to\infty} f(x_n) \leq \sup X_i \leq \sup f(B_\delta),$$

haciendo $\delta \to 0$, tenemos

$$\limsup_{n\to\infty} f(x_n) \leq \limsup_{y\to a} f(y).$$

□

Función Semicontinua Superior

Definición

Sea $f : X \to R$, f es semicontinua superior en a si, para todo $\varepsilon > 0$, existe $\delta > 0$ tal que para todo x en X, si $\| x - a \| < \delta$, implica $f(x) - f(a) < \varepsilon$.

f es semicontinua superior en $Y \subset X$ si, para todo a en Y, f es semicontinua superior en a.

Afirmación

Sea $f : X \to R$ acotada, si para todo (x_n) en X_n con $x_n \to a$, implica $\limsup_{n\to\infty} f(x_n) \leq f(a)$, entonces f es semicontinua superior en a.

Demostración. Supongamos por contradicción que f no es semicontinua superior en a, esto implica que existe un $\varepsilon_0 > 0$, talque para todo n en N, existe x_n en X con $\| x_n - a \| < \frac{1}{n}$ y $f(x_n) - f(a) \geq \varepsilon_0$, tomando el limite superior tenemos

$$\limsup_{n\to\infty} f(x_n) \geq f(a) + \varepsilon_0 > f(a),$$

lo cual es una contradicción, lo que prueba la afirmación.

□

Teorema de Hahn-Banach

Sean $Y \subset X$ un subespacio vectorial, $p : Y \to R$ lineal y $f : X \to R$ tal que

i) $f(x + y) \leq f(x) + f(y)$, para todo x, y en X,
ii) $f(\lambda x) = \lambda f(x)$, para todo x en X y $\lambda \geq 0$ y
iii) $p(y) \leq f(y)$, para todo y en Y,

entonces existe $P : X \to R$ tal que

$$P(y) = p(y) \text{ para todo } y \text{ en } Y \text{ y } P(x) \leq f(x).$$

Topología Débil*

Sea $\varphi : X \to X^{**}$ definido como, $\varphi(v) = \varphi_v : X^* \to R$ tal que $\varphi_v(\zeta) = \langle \zeta, v \rangle$, φ es llamado la *inclusion natural* de X a X^{**}

$$B = \{\varphi_v^{-1}(A) \subset X^* : v \in X \ y \ A \ es \ abierto \ de \ R\},$$

B es una base topológica, a la topología generada por B se lo conoce como *topologia debil**

Teorema (Alaoglu)

$S^* = \{\zeta \in X^* : \| \zeta \| \leq 1\}$ es compacto en la topología débil*.

Bibliografía

[1] FRANK H. CLARKE, *Optimization and Nonsmooth Analysis*, Montreal, 1990.

[2] H. L. ROYDEN, *Functional Analysis*, Stanford University, 1968.

[3] WALTER RUDIN, *Functional Analysis*, Wisconsin, 1973.

[4] ELON L. LIMA, *Algebra Lineal*, IMPA, 1998.

[5] FRANK H. CLARKE, *Nonsmooth Analysis and Control Theory*, New York, 1998 .