

Universidad Centroccidental "Lisandro Alvarado"  
Decanato de Ciencias y Tecnología  
Programa de Licenciatura en Ciencias Matemáticas

ALGORITMO DE PUNTO PROXIMAL EN OPTIMIZACION SOBRE  
VARIETADES DE RIEMANN.

Br. María Virginia Morales Mendoza.  
Tutor: Dr. Eibar R. Hernández R.  
Area de conocimiento: Optimización y Geometría Diferencial.

Barquisimeto, Marzo 2007.

Universidad Centroccidental "Lisandro Alvarado"  
Decanato de Ciencias y Tecnología  
Programa de Licenciatura en Ciencias Matemáticas

Trabajo Especial de Grado presentado  
para optar al título de  
Licenciado en Ciencias Matemáticas.

Br. María Virginia Morales Mendoza.  
Tutor: Dr. Eibar R. Hernández R.

Barquisimeto, Marzo 2007.

## Agradecimiento.

A Dios Todopoderoso, que nos acompaña en todo momento y camina siempre a nuestro lado. A mis padres, Alirio J. Morales y Gisela Mendoza de Morales por darme la vida, el amor y el cariño, y sobretodo por apoyarme siempre en mis estudios. A mis hermanos, Alirio A. Morales y María Carolina Morales por animarme a terminar mi carrera con sus palabras de aliento. A mi tutor Dr. Eibar R. Hernández R, por su valiosa asesoría, su apoyo, sus enseñanzas, su colaboración para el desarrollo de esta labor y por sus grandes aportes para esta modesta obra. A mi Profesor Dr. Rómulo Castillo, por sus consejos, atención y ayuda. A mi amigo Waiberh Acosta, por su grata compañía a lo largo de mi carrera, en la que siempre conté con su apoyo y a Isaac Mendoza por ofrecerme su ayuda incondicional en todo momento. A mi amiga Raisbel Pérez por su apoyo, su confianza en mí y por sus valiosos consejos. Y a mis amigos, Albannis Ramos, Pavel Santeliz, María Rodriguez, María Carruido, Nancys Noguera, Elyseé Sánchez y Ruben Ortiz por su amena compañía, cariños y los inolvidables momentos entorno de la Universidad Centroccidental "Lisandro Alvarado".

## Dedicatoria.

A Dios nuestro Señor, primeramente por haberme dado el don de la vida, y hacer posible este logro que es para El, por enseñarme que a pesar de las dificultades no es vana la ofrenda de un nuevo día. Le agradezco haberme llenado de entendimiento y sabiduría, guiarme en el camino e iluminar mis pensamientos, enseñándome a encontrar las respuestas a mi vida en El, que todo lo puede, porque en la fuerza de Dios no existen los imposibles. A mis padres y hermanos por haberme apoyado siempre, e impulsado a salir adelante, por sobre todas las cosas, y haberme dado todo su amor y confiar en mí.

ALGORITMO DE PUNTO PROXIMAL EN OPTIMIZACION SOBRE VARIETADES DE RIEMANN.

## Resumen.

El objetivo de esta monografía es presentar la extensión de algunos conceptos y técnicas de la optimización en el espacio Euclideo  $\mathbb{R}^n$  a las variedades de Riemann, lo cual es bastante natural bajo ciertas condiciones. Consideraremos el problema de minimización con restricciones y probaremos que si el conjunto de restricciones es una variedad de Riemann de curvatura seccional no positiva y la función objetivo es convexa en esta variedad, el método de punto proximal es naturalmente extendido, probando que la sucesión generada esta bien definida y converge a un minimizador.

# Índice

<b>1. Introducción.</b>	<b>1</b>
<b>2. Conceptos Básicos.</b>	<b>3</b>
<b>3. Tópicos sobre Variedades de Riemann.</b>	<b>9</b>
<b>4. Tópicos sobre Análisis Convexo.</b>	<b>14</b>
4.1. Conjuntos Convexos. . . . .	14
4.2. Funciones Convexas. . . . .	23
<b>5. Regularización.</b>	<b>30</b>
<b>6. Algoritmo de Punto Proximal sobre Variedades de Riemann.</b>	<b>33</b>
<b>7. Convergencia del Algoritmo de Punto Proximal.</b>	<b>35</b>
<b>Bibliografía.</b>	<b>41</b>

# Capítulo 1

## Introducción.

Dado un problema de la forma  $(L + \lambda\mathcal{M})(f) = 0$ , donde  $f$  es un elemento de un conjunto  $X$  (usualmente un espacio de funciones) y  $L + \lambda\mathcal{M} : X \rightarrow X$  es un operador (usualmente diferenciable) con solución única  $f_\lambda$  para cada  $\lambda > 0$ . Se espera que cuando  $\lambda$  tienda a 0,  $f_\lambda$  tienda a una solución del problema  $L(f) = 0$  (usualmente esto no se comporta tan bien como se espera). Esta idea se conoce con el nombre de Regularización.

Si  $X = \mathbb{R}^n$  y  $L = \nabla f$  donde  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  es una función convexa, el problema  $\nabla f(x) = 0$  o equivalentemente  $\min\{f(x) : x \in \mathbb{R}^n\}$  puede ser regularizado suponiendo que  $f$  es acotada inferiormente y tomando  $g : \mathbb{R}^n \rightarrow \mathbb{R}$  estrictamente convexa y coerciva (es decir  $\lim_{\|x\| \rightarrow \infty} g(x) = +\infty$ ) [1] al considerar el problema

$$\min\{f(x) + \lambda g(x) : x \in \mathbb{R}^n\}$$

que tiene una única solución para cada  $\lambda > 0$  cuando se reduce el estudio a conjuntos compactos.

El algoritmo de punto proximal Euclideo está enmarcado en esta situación, este genera una sucesión  $\{x^k\} \subset \mathbb{R}^n$  en la siguiente forma:

$$x^0 \in \mathbb{R}^n$$
$$x^{k+1} = \arg \min_{x \in \mathbb{R}^n} \{f(x) + \lambda_k \|x - x^k\|\},$$

donde  $\lambda_k$  es un número real que cumple  $0 < \lambda_k \leq \bar{\lambda}$ , para algún  $\bar{\lambda} > 0$  y  $\|\cdot\|$  es la norma Euclidea. Este algoritmo converge bajo ciertas condiciones, como lo son: que la sucesión  $\{x^k\}$  sea Fejér convergente, que la función  $f$  sea convexa y continuamente diferenciable y el conjunto  $U$  de minimizadores de  $f$  sea no vacío.

La extensión de algunos conceptos básicos y técnicas de la optimización en el espacio Euclideo  $\mathbb{R}^n$  a las variedades de Riemann, es bastante natural bajo ciertas condiciones. En el artículo [7], que esta monografía desarrolla, se considera el problema de minimización con restricciones y se prueba que si el conjunto de restricciones es una variedad de Riemann de curvatura seccional no positiva y la función objetivo es convexa en esta variedad, el método de punto proximal es naturalmente extendido, probando que la sucesión generada está bien definida y converge a un minimizador.

Se desarrollarán y estudiarán los conceptos y teoremas básicos correspondientes a variedades de Riemann y optimización sobre los cuales se sustenta dicha metodología a través de la lectura de bibliografía especializada y análisis de artículos científicos. Es claro que tal generalización del método de punto proximal en espacios Euclideo, a variedades de Riemann ya es conocida y el objetivo en esta monografía es el de estudiar en detalle como fue extendido dicho método.

Sin embargo, se encuentran en [4] suficientes ejemplos que muestran que no es posible establecer la convergencia de algunos métodos bajo ciertas condiciones en el caso Euclideo y en su lugar, al considerar una métrica de Riemann conveniente en sustitución de la Euclídea, la convergencia se garantiza, por ello se hace evidente la necesidad de generalizar tales métodos a variedades de Riemann.

El método de punto proximal para variedades de Riemann viene dado de la siguiente forma [7]:

Sea  $\mathcal{M}$  una variedad Riemanniana completa y sea  $f : \mathcal{M} \rightarrow \mathbb{R}$  una función convexa.

Consideremos el problema de Optimización

$$\min_{x \in \mathcal{M}} \{f(x)\}.$$

El algoritmo de punto proximal en variedades Riemannianas genera, para un punto fijo  $p_0 \in \mathcal{M}$ , una sucesión  $\{x_k\} \subset \mathcal{M}$  por la iteración

$$x_{k+1} = \arg \min_{y \in \mathcal{M}} \{f(y) + \lambda_k \rho_{x_k}(y)\},$$

donde  $\rho_x : \mathcal{M} \rightarrow \mathbb{R}$  es definida por:

$$\rho_x(y) = \frac{1}{2} d^2(x, y),$$

donde  $d$  es la distancia Riemanniana y  $\{\lambda_k\}$  es una sucesión de números positivos.

# Capítulo 2

## Conceptos Básicos.

En este capítulo se desarrollarán algunos conceptos y teoremas sobre geometría Riemanniana que se requieren para comprender el ambiente sobre el cual se realizará la extensión del método de punto proximal.

**Definición 2.1.** Una variedad diferenciable de dimensión  $n$  es un conjunto  $\mathcal{M}$  y una familia de aplicaciones biunívocas  $x_\alpha : U_\alpha \subset \mathbb{R}^n \rightarrow \mathcal{M}$ , donde cada  $U_\alpha$  es un abierto de  $\mathbb{R}^n$  en  $\mathcal{M}$  tal que:

$$\cup_\alpha x_\alpha(U_\alpha) = \mathcal{M}.$$

Para todo par  $\alpha, \beta$ , con  $x_\alpha(U_\alpha) \cap x_\beta(U_\beta) = W \neq \emptyset$ , los conjuntos  $x_\alpha^{-1}(W)$  y  $x_\beta^{-1}(W)$  son abiertos en  $\mathbb{R}^n$  y las aplicaciones  $x_\beta^{-1} \circ x_\alpha$  son diferenciables.

**Definición 2.2.** Cada par  $(U_\alpha, x_\alpha)$  con  $p \in x_\alpha(U_\alpha)$  es llamado una parametrización de  $\mathcal{M}$  en  $p$ .

**Definición 2.3.** Sean  $\mathcal{M}_1^n$  y  $\mathcal{M}_2^m$  variedades diferenciables. Una aplicación  $\varphi : \mathcal{M}_1 \rightarrow \mathcal{M}_2$  es diferenciable en  $p \in \mathcal{M}_1$  si dada una parametrización

$$y : V \subset \mathbb{R}^m \rightarrow \mathcal{M}_2$$

en  $\varphi(p)$  existe una parametrización

$$x : U \subset \mathbb{R}^n \rightarrow \mathcal{M}_1$$

en  $p$  tal que  $\varphi(x(U)) \subset y(V)$  es una aplicación

$$y^{-1} \circ \varphi \circ x : x^{-1}(U) \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$$

diferenciable en  $x^{-1}(p)$ .

$\varphi$  es diferenciable en un abierto de  $\mathcal{M}_1$  si es diferenciable en todos los puntos de este abierto.

**Definición 2.4.** Sea  $\mathcal{M}$  una variedad diferenciable. Una aplicación diferenciable  $\alpha : (-\varepsilon, \varepsilon) \rightarrow \mathcal{M}$  es llamada una curva diferenciable en  $\mathcal{M}$ . Suponga que  $\alpha(0) = p \in \mathcal{M}$ , y sea  $\mathcal{D}$  un conjunto de funciones de  $\mathcal{M}$  diferenciable en  $p$ . Un vector tangente a una curva  $\alpha$  en  $t = 0$  es una función  $\alpha'(0) : \mathcal{D} \rightarrow \mathbb{R}$  dada por:

$$\alpha'(0)f = \left. \frac{d(f \circ \alpha)}{dt} \right|_{t=0}, \quad f \in \mathcal{D}.$$

**Definición 2.5.** Un vector tangente en  $p$  es el vector tangente en  $t = 0$  de alguna curva  $\alpha : (-\varepsilon, \varepsilon) \rightarrow \mathcal{M}$  con  $\alpha(0) = p$ . El conjunto de vectores tangentes a  $\mathcal{M}$  en  $p$  será indicado por  $T_p(\mathcal{M})$ .

**Proposición 2.1.** Sean  $\mathcal{M}_1^n$  y  $\mathcal{M}_2^m$  variedades diferenciables y sea  $\varphi : \mathcal{M}_1 \rightarrow \mathcal{M}_2$  una aplicación diferenciable. Para cada  $p \in \mathcal{M}_1$  y cada  $v \in T_p(\mathcal{M}_1)$ , escoja una curva diferenciable  $\alpha : (-\varepsilon, \varepsilon) \rightarrow \mathcal{M}_1$  con  $\alpha(0) = p$ ,  $\alpha'(0) = v$ . Para  $\beta = \varphi \circ \alpha$ . La aplicación  $d\varphi_p : T_p(\mathcal{M}_1) \rightarrow T_{\varphi(p)}(\mathcal{M}_2)$  dada por  $d\varphi_p(v) = \beta'(0)$  es una aplicación lineal que no depende de la escogencia de  $\alpha$ .

**Demostración:** Ver [5].

**Definición 2.6.** La aplicación lineal  $d\varphi_p$  dada por la definición anterior es llamada diferencial de  $\varphi$  en  $p$ .

**Definición 2.7.** Sean  $\mathcal{M}_1$  y  $\mathcal{M}_2$  variedades diferenciables. Una aplicación  $\varphi : \mathcal{M}_1 \rightarrow \mathcal{M}_2$  es un difeomorfismo si y solo si ella es diferenciable, biunívoca, sobreyectiva y su inversa  $\varphi^{-1}$  es diferenciable.  $\varphi$  es un difeomorfismo local en  $p \in \mathcal{M}$  si existen vecindades  $U$  de  $p$  y  $V$  de  $\varphi(p)$  tales que  $\varphi : U \rightarrow V$  es un difeomorfismo.

**Definición 2.8.** El fibrado tangente de  $\mathcal{M}$ , se define como:

$$T\mathcal{M} = \bigcup_{x \in \mathcal{M}} T_x\mathcal{M}.$$

**Definición 2.9.** Un campo de vectores  $X$  en una variedad diferenciable  $\mathcal{M}$  es una correspondencia que a cada punto  $p \in \mathcal{M}$  asocia un vector  $X(p) \in T_p(\mathcal{M})$ .

En término de aplicaciones,  $X$  es una aplicación de  $\mathcal{M}$  en el fibrado tangente  $T\mathcal{M}$ . Un campo es diferenciable si la aplicación  $X : \mathcal{M} \rightarrow T\mathcal{M}$  es diferenciable. Consideremos una parametrización  $x : U \subset \mathbb{R}^n \rightarrow \mathcal{M}^n$ , es posible escribir

$$X(p) = \sum_{i=1}^n a_i(p) \frac{\partial}{\partial x_i},$$

donde cada  $a_i : U \rightarrow \mathbb{R}$  es una función en  $U$  y  $\{\frac{\partial}{\partial x_i} : i = 1, \dots, n\}$  es la base asociada a  $x$ . Es claro que  $X$  es diferenciable si y sólo si las funciones  $a_i$  son diferenciable para alguna (y, por tanto, todas) parametrización.

**Definición 2.10.** Una métrica Riemanniana (o estructura Riemanniana) en una variedad diferenciable  $\mathcal{M}$  es una ley que hace corresponder a cada punto  $p$  de  $\mathcal{M}$  un producto interno (es decir, una forma bilineal simétrica, definida positiva)  $\langle \cdot, \cdot \rangle_p$  en el espacio tangente  $T_p\mathcal{M}$ , que varía diferenciablemente en el sentido siguiente: sea  $x : U \subset \mathbb{R}^n \rightarrow \mathcal{M}$  un sistema de coordenadas locales en  $p$ , con  $x(x_1, x_2, \dots, x_n) = q \in x(U)$  y  $\frac{\partial}{\partial x_i}(q) = dx(0, \dots, 1, \dots, 0)$ , entonces

$$\left\langle \frac{\partial}{\partial x_i}(q), \frac{\partial}{\partial x_j}(q) \right\rangle_q = g_{ij}(x_1, \dots, x_n)$$

es una función diferenciable en  $U$ .

**Definición 2.11.** Si  $c : [a, b] \rightarrow \mathcal{M}$  es una curva continuamente diferenciable en la variedad de Riemann  $\mathcal{M}$ , definimos la longitud  $L(c)$  como:

$$L(c)_a^b = \int_a^b \left\langle \frac{dc}{dt}, \frac{dc}{dt} \right\rangle^{\frac{1}{2}} dt.$$

**Definición 2.12.** La distancia  $d(x', x)$  entre dos puntos  $x'$  y  $x$  en  $\mathcal{M}$  se define como:

$$d(x', x) = \inf\{L(c)\},$$

donde  $c$  es una curva suave a trozos que conecta a  $x'$  y  $x$ .

**Definición 2.13.** Una conexión afín  $\nabla$  en una variedad diferenciable  $\mathcal{M}$  es una aplicación

$$\nabla : X(\mathcal{M}) \times X(\mathcal{M}) \rightarrow X(\mathcal{M}),$$

que se indica por  $(X, Y) \rightarrow \nabla_X Y$ , donde  $X(\mathcal{M})$  es el conjunto de todos los campos vectoriales en  $\mathcal{M}$  y que satisface las siguientes propiedades:

- i)  $\nabla_{fX+gY} Z = f\nabla_X Z + g\nabla_Y Z$
- ii)  $\nabla_X(Y + Z) = \nabla_X Y + \nabla_X Z$
- iii)  $\nabla_X(fY) = f\nabla_X Y + X(f)Y$ ,

donde  $X, Y, Z \in X(\mathcal{M})$  y  $f, g \in \mathcal{D}(\mathcal{M})$ .

**Proposición 2.2.** Sea  $\mathcal{M}$  una variedad diferenciable con una conexión afín  $\nabla$ .

Entonces existe una única ley que asocia a un campo vectorial  $V$  a lo largo de la curva diferenciable  $c : I \rightarrow \mathcal{M}$  otro campo vectorial  $\frac{DV}{dt}$  a lo largo de  $c$ , denominado derivada covariante de  $V$  a lo largo de  $c$ , tal que:

- a)  $\frac{D}{dt}(V + W) = \frac{DV}{dt} + \frac{DW}{dt}$ ,
- b)  $\frac{D}{dt}(fV) = \frac{df}{dt}V + f\frac{DV}{dt}$ ,

donde  $W$  es un campo vectorial a lo largo de  $c$  y  $f \in \mathcal{D}(\mathcal{M})$ .

c) Si  $V$  es inducido por un campo de vectores  $Y \in X(\mathcal{M})$ , es decir,

$$V(t) = Y(c(t)),$$

entonces  $\frac{DV}{dt} = \nabla_{\frac{dc}{dt}} Y$ .

**Demostración:** Ver [5].

**Definición 2.14.** Sea  $\mathcal{M}$  una variedad diferenciable con una conexión afín  $\nabla$ .

Un campo vectorial  $V$  a lo largo de una curva  $c : I \rightarrow \mathcal{M}$  es llamado paralelo cuando  $\frac{DV}{dt} = 0$ , para todo  $t \in I$ .

**Teorema 2.1. (LEVI-CIVITA)**

Sea  $\nabla$  la conexión "Levi-Civita" asociada a  $(\mathcal{M}, \langle \cdot, \cdot \rangle)$ .

Dada una variedad Riemanniana  $\mathcal{M}$ , existe una única conexión afín  $\nabla$  en  $\mathcal{M}$  satisfaciendo las condiciones:

- $\nabla$  es simétrica.
- $\nabla$  es compatible con la métrica Riemanniana.

**Definición 2.15.** Una curva parametrizada  $\gamma : I \rightarrow \mathcal{M}$  es una geodésica en  $t_0 \in I$  si  $\frac{D}{dt}(\frac{d\gamma}{dt}) = 0$  en el punto  $t_0$ ; si  $\gamma$  es geodésica en  $t$ , para cada  $t \in I$ , decimos que  $\gamma$  es una geodésica. Si  $[a, b] \subset I$  y  $\gamma : I \rightarrow \mathcal{M}$  es una geodésica, la restricción de  $\gamma$  a  $[a, b]$  se llama el segmento geodésico uniendo  $\gamma(a)$  con  $\gamma(b)$ .

**Definición 2.16.** Un segmento geodésico  $\gamma : [a, b] \rightarrow \mathcal{M}$  se llama minimizante si  $l(\gamma) \leq l(c)$ , donde  $l(\cdot)$  indica la longitud de una curva y  $c$  es cualquier curva diferenciable por partes uniendo  $\gamma(a)$  con  $\gamma(b)$ .

**Observación 2.1.** La ecuación geodésica  $\nabla_{\gamma'} \gamma' = 0$  es una ecuación diferencial ordinaria no lineal de segundo orden, entonces  $\gamma = \gamma_v(\cdot, x)$  es determinada por la posición  $x$  y la velocidad  $v$  en  $x$ .

**Observación 2.2.** Sea  $\gamma : I \rightarrow \mathcal{M}$  una geodésica, entonces

$$\frac{d}{dt} \left\langle \frac{d\gamma}{dt}, \frac{d\gamma}{dt} \right\rangle = 2 \left\langle \frac{D}{dt} \frac{d\gamma}{dt}, \frac{d\gamma}{dt} \right\rangle = 0,$$

es decir, la longitud del vector tangente  $\frac{d\gamma}{dt}$  es constante.

**Definición 2.17.** La longitud de arco  $s$  de  $\gamma$ , a partir de un origen fijo, digamos  $t = t_0$ , esta dado por:

$$s(t) = \int_{t_0}^t \left| \frac{d\gamma}{dt} \right| dt = c(t - t_0).$$

Por lo tanto, el parámetro de una geodésica es proporcional a la longitud de arco. Cuando el parámetro es la propia longitud de arco, es decir,  $c = 1$ , diremos que  $\gamma$  está normalizada.

**Definición 2.18.** Tomemos  $x \in \mathcal{M}$ , el mapeo exponencial  $exp_x : T_x\mathcal{M} \rightarrow \mathcal{M}$  es definido por  $exp_x v = \gamma_v(1, x)$ .

**Definición 2.19.** Una variedad Riemanniana  $\mathcal{M}$  es geodesicamente completa si para cada  $p \in \mathcal{M}$ , la aplicación exponencial,  $exp_p$ , se define para cada  $v \in T_p\mathcal{M}$ , es decir, si las geodésicas  $\gamma(t)$  que parten de  $p$  están definidas para todos los valores del parámetro  $t \in \mathbb{R}$ .

**Teorema 2.2.** Teorema de HOPF-RINOW'S.

Sea  $(\mathcal{M}, g)$  una variedad Riemanniana conexa. Entonces, son equivalentes:

- a)  $(\mathcal{M}, g)$  es completa.
- b)  $\forall p \in \mathcal{M}$ ,  $exp_p$  está definida en todo  $T_p\mathcal{M}$ .
- c)  $\exists p \in \mathcal{M}$  tal que  $exp_p$  está definida en todo  $T_p\mathcal{M}$ .
- d) La familia de compactos de  $\mathcal{M}$  coincide con la familia de cerrados y acotados.

Además, cada una de las afirmaciones anteriores implican que:

- e) Para todo  $p, q \in \mathcal{M}$  existe una geodésica  $\gamma$  que une  $p$  con  $q$ , con  $l(\gamma) = d(p, q)$ .

**Demostración:** Ver [5].

**Observación 2.3.** El Teorema de HOPF-RINOW'S afirma que si este es el caso entonces para algún par de puntos, digamos  $x'$  y  $x$ , en  $\mathcal{M}$  puede ser conectado por un (no necesariamente único) segmento minimal geodésico.

Más aún,  $(\mathcal{M}, d)$  es un espacio métrico completo y subconjuntos compactos son cerrados y acotados.

**Definición 2.20.** La curvatura  $R$  de una variedad Riemanniana  $\mathcal{M}$  es una ley que asocia cada par  $X, Y \in X(\mathcal{M})$  una aplicación  $R(X, Y) : X(\mathcal{M}) \rightarrow X(\mathcal{M})$  dada por:

$$R(X, Y)Z = \nabla_Y \nabla_X Z - \nabla_X \nabla_Y Z + \nabla_{[Y, X]} Z,$$

$Z \in X(\mathcal{M})$  dónde  $\nabla$  es la conexión Riemanniana de  $\mathcal{M}$ .

**Definición 2.21.** El tensor curvatura  $R$  esta definido por:

$$R(X, Y) = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[Y, X]} Z,$$

dónde  $X, Y$  y  $Z$  son los campo del vector de  $\mathcal{M}$  y  $[X, Y] = YX - XY$ .

---

**Definición 2.22.** Dado un punto  $p \in \mathcal{M}$  y un subespacio bidimensional  $\sigma \subset T_p\mathcal{M}$ , el número real  $K(X, Y) = K(\sigma)$ , donde  $\{X, Y\}$  es cualquier base de  $\sigma$ , es llamada la curvatura seccional de  $\sigma$  en  $p$ .

**Observación 2.4.** Entonces la curvatura seccional con respecto a  $X$  y  $Y$  está dado por:

$$K(X, Y) = \frac{\langle R(X, Y)Y, X \rangle}{(\|X\|^2\|Y\|^2 - \langle X, Y \rangle^2)},$$

donde  $\|X\|^2 = \langle X, X \rangle$ .

**Definición 2.23.** Si  $K(X, Y) \leq 0$  para todo  $X$  y  $Y$ , entonces  $\mathcal{M}$  es llamado una variedad Riemanniana de curvatura no positiva y usamos la notación  $K \leq 0$ .

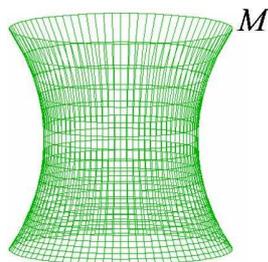
## Capítulo 3

# Tópicos sobre Variedades de Riemann.

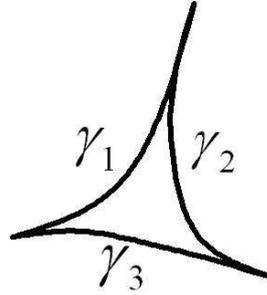
**Teorema 3.1.** *Sea  $\mathcal{M}$  una variedad Riemanniana completa simplemente conexa de curvatura seccional no positiva. Entonces  $\mathcal{M}$  es difeomórfica a el espacio Euclideo  $\mathbb{R}^n$ ,  $n = \dim \mathcal{M}$ . Más precisamente, para algún punto  $x \in \mathcal{M}$ , el mapeo exponencial  $\exp_x : T_x \mathcal{M} \rightarrow \mathcal{M}$  es un difeomorfismo.*

**Demostración:** Ver [5].

**Definición 3.1.** *Una variedad Riemanniana completa simplemente conexa de curvatura seccional no positiva es llamada una VARIEDAD DE HADAMARD.*



**Definición 3.2.** *Un TRIÁNGULO GEODÉSICO  $\Delta(x_1 x_2 x_3)$  en una variedad Riemanniana es el conjunto que consiste de tres puntos distintos  $x_1, x_2, x_3$ , los cuales son llamados los vértices y tres segmentos de geodésicas minimales  $\gamma_{i+1}$  que une  $x_{i+1}$  a  $x_{i+2}$  los cuales son llamados los lados, donde  $i = 1, 2, 3 \pmod{3}$ .*



**Proposición 3.1.** Sean  $\mathcal{M}_1$  y  $\mathcal{M}_2$  variedades Riemannianas y supongamos que para todo  $p_1 \in \mathcal{M}_1$ ,  $p_2 \in \mathcal{M}_2$ ,  $\sigma_1 \subset T_{p_1}(\mathcal{M}_1)$ ,  $\sigma_2 \subset T_{p_2}(\mathcal{M}_2)$ , se tiene que  $K_2(p_2, \sigma_2) \geq K_1(p_1, \sigma_1)$ . Sean  $p_1 \in \mathcal{M}_1$ ,  $p_2 \in \mathcal{M}_2$  y sea  $i : T_{p_1}(\mathcal{M}_1) \rightarrow T_{p_2}(\mathcal{M}_2)$  una función lineal inyectiva que preserva productos internos. Sea  $r > 0$  tal que  $\exp_{p_1}|_{B_r(0)}$  es un difeomorfismo y  $\exp_{p_2}|_{B_r(0)}$  es no singular. Entonces, para cualquier curva  $C : [0, l] \rightarrow \exp_{p_1}(B_r(0)) \subset \mathcal{M}_1$  una curva diferenciable y defina una curva  $C_1 : [0, l] \rightarrow \exp_{p_2}(B_r(0)) \subset \mathcal{M}_2$  por:

$$C_2(S) = \exp_{p_2} \circ i \circ \exp_{p_1}(C(S)),$$

$S \in [0, l]$ . Entonces,  $L(C_1) \geq L(C_2)$ .

**Demostración:** Ver [5].

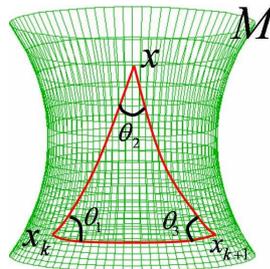
**Teorema 3.2.** Sea  $\mathcal{M}$  una variedad de Hadamard y  $\Delta(x_1x_2x_3)$  un triángulo geodésico. Denote por  $\gamma_{i+1} : [0, l_{i+1}] \rightarrow \mathcal{M}$  un segmento geodésico que va de  $x_{i+1}$  a  $x_{i+2}$  y el conjunto  $l_{i+1} := L(\gamma_{i+1})$ ,  $\theta_{i+1} = \angle(\gamma'_{i+1}(0), -\gamma'_i(l_i))$ , donde  $i = 1, 2, 3$ , (mod  $\cdot 3$ ).

Entonces,

$$\theta_1 + \theta_2 + \theta_3 \leq \Pi \quad (3.1)$$

$$l_{i+1}^2 + l_{i+2}^2 - 2l_{i+1}l_{i+2}\cos\theta_{i+2} \leq l_i^2 \quad (3.2)$$

$$l_{i+1}\cos\theta_{i+2} + l_i\cos\theta_i \geq l_{i+2}. \quad (3.3)$$



**Demostración.**

Veamos que se satisface la ecuación (3 · 1).

Si  $\gamma_{i+1}$ ,  $\gamma_{i+2}$  y  $\gamma_i$  son segmentos geodésicos,

$$L(\gamma_{i+1}) = l_{i+1}$$

$$L(\gamma_{i+2}) = l_{i+2}$$

$$L(\gamma_i) = l_i$$

que forman los lados de  $\Delta(x_1x_2x_3)$ .

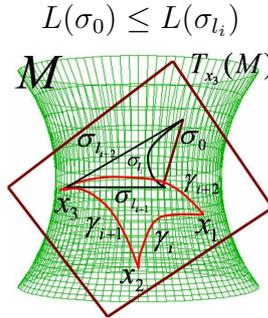
Sean  $\sigma_{l_{i+1}} = \exp_{x_3}^{-1}(\gamma_{i+1})$ ,  $\sigma_{l_{i+2}} = \exp_{x_3}^{-1}(\gamma_{i+2})$  y  $\sigma_{l_i} = \exp_{x_3}^{-1}(\gamma_i)$ , curvas en  $T_{x_3}(\mathcal{M})$ .

Como  $\gamma_{i+1}$  y  $\gamma_{i+2}$  son geodésicas, que surgen como radios de  $x_3$ , tenemos que:

$$l_{i+1} = L(\gamma_{i+1}) = L(\sigma_{l_{i+1}}),$$

$$l_{i+2} = L(\gamma_{i+2}) = L(\sigma_{l_{i+2}}).$$

Luego, indicando por  $\sigma_0$  un segmento de recta en  $T_{x_3}(\mathcal{M})$  que une las extremidades de  $\sigma_{l_i}$ , tenemos que:



y

$$L(\sigma_0)^2 = l_{i+1}^2 + l_{i+2}^2 - 2l_{i+1}l_{i+2}\cos\theta_{i+2},$$

como  $K \leq 0$  y  $T_{x_3}(\mathcal{M})$  tienen curvatura nula, podemos aplicar la Proposición 3 · 1 y obtenemos que:

$$L(\sigma_{l_i}) \leq L(\gamma_i), \quad (<, \quad \text{si } K < 0).$$

Así, concluimos que:

$$\begin{aligned} l_{i+1}^2 + l_{i+2}^2 - 2l_{i+1}l_{i+2}\cos\theta_{i+2} &\leq L(\sigma_{l_i})^2 \\ &\leq L(\gamma_i)^2 \\ &\leq l_i^2, \end{aligned}$$

( $<$ , si  $K < 0$ .)

Para obtener la ecuación (3 · 2) observemos que:

$$\begin{aligned} l_i &= d(x_1, x_2), \\ l_{i+2} &= d(x_1, x_3), \\ l_{i+1} &= d(x_2, x_3) \end{aligned}$$

y, por lo tanto, la longitud de cada segmento  $l_{i+1}$ ,  $l_{i+2}$  ó  $l_i$  es menor que la suma de los otros dos.

Podemos entonces encontrar en el espacio Euclideo  $T_{x_3}(\mathcal{M})$  un triángulo cuyos lados tenga segmentos  $l_{i+1}$ ,  $l_{i+2}$  y  $l_i$ . Indicando los ángulos opuestos en el triángulo por:  $\theta'_1$ ,  $\theta'_2$  y  $\theta'_3$ , respectivamente, obtenemos de la ecuación (3 · 1) que:

$$\theta_1 \leq \theta'_1, \quad \theta_2 \leq \theta'_2 \quad y \quad \theta_3 \leq \theta'_3, \quad (<, \text{ si } K < 0).$$

Como  $\theta'_1 + \theta'_2 + \theta'_3 = \pi$ , entonces

$$\begin{aligned} \theta_1 + \theta_2 + \theta_3 &\leq \theta'_1 + \theta'_2 + \theta'_3 = \pi \\ \Rightarrow \theta_1 + \theta_2 + \theta_3 &\leq \pi. \end{aligned}$$

La ecuación (3 · 3) es consecuencia inmediata de (3 · 2).

En efecto,

$$l_{i+1}^2 + l_{i+2}^2 - 2l_{i+1}l_{i+2}\cos\theta_{i+2} \leq l_i^2,$$

para  $i = 1$ , tenemos:

$$\begin{aligned} l_{1+1}^2 + l_{1+2}^2 - 2l_{1+1}l_{1+2}\cos\theta_{1+2} &\leq l_1^2, \\ \Rightarrow l_2^2 + l_3^2 - 2l_2l_3\cos\theta_3 &\leq l_1^2, \end{aligned} \tag{3.4}$$

para  $i = 2$ , se tiene:

$$\begin{aligned} l_{2+1}^2 + l_{2+2}^2 - 2l_{2+1}l_{2+2}\cos\theta_{2+2} &\leq l_2^2, \\ \Rightarrow l_3^2 + l_1^2 - 2l_3l_1\cos\theta_1 &\leq l_2^2, \end{aligned} \tag{3.5}$$

sumando (3 · 4) y (3 · 5) se tiene:

$$\begin{aligned} l_2^2 + l_3^2 - l_1^2 + l_3^2 + l_1^2 - l_2^2 &\leq 2l_2l_3\cos\theta_3 + 2l_3l_1\cos\theta_1 \Rightarrow 2l_3^2 \leq 2l_2l_3\cos\theta_3 + 2l_3l_1\cos\theta_1 \\ &\Rightarrow 2l_3^2 \leq 2l_3(l_2\cos\theta_3 + l_1\cos\theta_1) \\ &\Rightarrow l_3 \leq l_2\cos\theta_3 + l_1\cos\theta_1, \end{aligned}$$

Así,

$$l_{i+1}\cos\theta_{i+2} + l_i\cos\theta_i \geq l_{i+2}.$$

**Definición 3.3.** Sea  $\gamma : [a, b] \rightarrow \mathcal{M}$  un segmento geodésico normalizado.  
Una variación diferencial de  $\gamma$  esta definida por el mapeo diferencial

$$\alpha : [a, b] \times (-\varepsilon, \varepsilon) \rightarrow \mathcal{M},$$

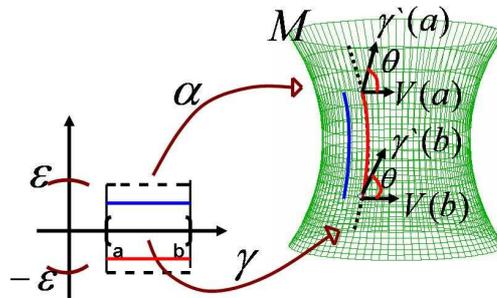
satisfaciendo  $\alpha(t, 0) = \gamma(t)$ .

El campo vectorial a lo largo de  $\gamma$  definido por  $V(t) = (\frac{\partial \alpha}{\partial s})(t, 0)$  es llamado el campo vectorial variacional de  $\alpha$ .

La primera forma variacional de el arco de longitud en  $\alpha$  esta dado como sigue:

$$L'(\gamma) := \frac{d}{ds} L(c_s)|_{s=0} = \langle V, \gamma' \rangle|_a^b, \quad (3.6)$$

donde  $c_s(t) = \alpha(t, s)$  con  $s \in (-\varepsilon, \varepsilon)$ .

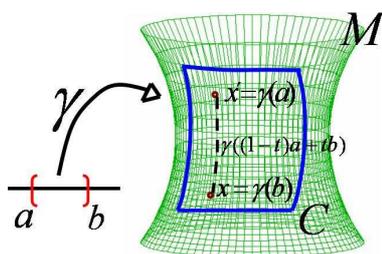


# Capítulo 4

## Tópicos sobre Análisis Convexo.

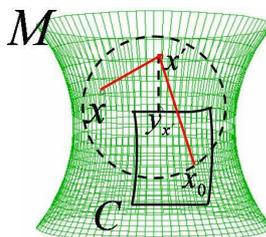
### 4.1. Conjuntos Convexos.

**Definición 4.1.1.** Sea  $\mathcal{C}$  un subconjunto de una variedad de Hadamard  $\mathcal{M}$ . El subconjunto  $\mathcal{C}$  de  $\mathcal{M}$  es llamado convexo si para puntos arbitrarios  $x'$  y  $x$  en  $\mathcal{C}$ , el segmento geodésico  $\gamma$  que une  $x'$  con  $x$  está contenido en  $\mathcal{C}$ , es decir, si  $x', x \in \mathcal{C}$  y  $\gamma$  es la geodésica tal que  $x' = \gamma(a)$  y  $x = \gamma(b)$ , entonces  $\gamma((1-t)a + tb) \in \mathcal{C}$ , para todo  $t \in [0, 1]$ .



Ahora, en lo que sigue, supongamos que  $\mathcal{C}$  es un subconjunto convexo cerrado de  $\mathcal{M}$ .

**Proposición 4.1.1.** Sea  $x' \in \mathcal{M}$  fijo, entonces existe un punto más cercano a  $x'$  en  $\mathcal{C}$ .



**Demostración.**

En efecto, para  $x' \in \mathcal{M}$  fijo, fijamos  $x_0 \in \mathcal{C}$  y definimos el conjunto,

$$A_{x_0} = \{x \in \mathcal{M} / d(x', x) \leq d(x', x_0)\}.$$

El mapeo  $x \mapsto d(x', x)$  es continuo.

Y además el conjunto  $\mathcal{C} \cap A_{x_0}$  es compacto, ya que  $\mathcal{C}$  es un subconjunto compacto,  $A_{x_0}$  es un conjunto cerrado y acotado, por lo tanto,  $\mathcal{C} \cap A_{x_0}$  es cerrado y acotado.

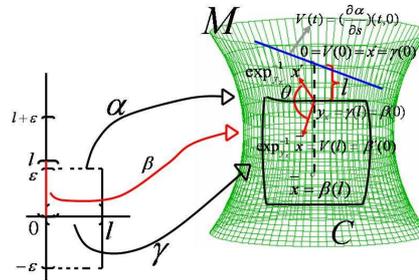
Así,  $\mathcal{C} \cap A_{x_0}$  es compacto.

Por tanto, existe un  $y_{x'} \in \mathcal{C}$  tal que  $d(x', y_{x'}) \leq d(x', x)$ , para todo  $x \in \mathcal{C}$ .

**Proposición 4.1.2.** *Sea  $x'$  en  $\mathcal{M}$ . Si  $y_{x'}$  en  $\mathcal{C}$  es tal que  $d(x', y_{x'}) \leq d(x', x)$  entonces,*

$$\langle \exp_{y_{x'}}^{-1} x', \exp_{y_{x'}}^{-1} x \rangle \leq 0, \tag{4.1}$$

para todo  $x \in \mathcal{C}$ .



**Demostración.**

Si  $x' \in \mathcal{C}$ , el resultado es trivial.

Supongamos que  $x' \notin \mathcal{C}$ . Fijemos  $l = d(x', y_{x'})$ , y sea  $\gamma : [0, l] \rightarrow \mathcal{M}$  la geodésica tal que  $\gamma(0) = x'$  y  $\gamma(l) = y_{x'}$ .

Recordemos que:

La distancia se define como,

$$d(x', x) = \inf\{L(c)\},$$

donde  $c$  es una curva suave a trozos que conecta a  $x'$  con  $x$ , y por definición de segmento geodésico minimal se tiene:

$$d(x, x') = L(\gamma).$$

Por otra parte, por definición de mapeo exponencial tenemos:

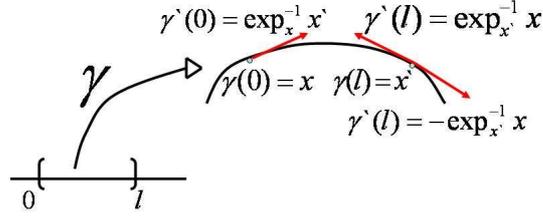
$$\exp_x : T_x \mathcal{M} \rightarrow \mathcal{M},$$

y esta dado por  $\exp_x v = \gamma_v(1, x)$ .

Y por definición de mapeo exponencial inverso:

$$\exp_x^{-1} : \mathcal{M} \rightarrow T_x \mathcal{M},$$

y esta dado por  $\exp_x^{-1} x' = \gamma'(0)$  y  $\exp_x^{-1} x = \gamma'(l)$ .



Así,

$$\gamma'(l) = -\frac{(\exp_{y_{x'}}^{-1} x')}{\|\exp_{y_{x'}}^{-1} x'\|},$$

esta normalizado.

Ahora supongamos lo contrario, que existe un punto  $\bar{x} \in \mathcal{C}$  tal que:

$$\langle \exp_{y_{x'}}^{-1} x', \exp_{y_{x'}}^{-1} \bar{x} \rangle > 0.$$

Considere la variación  $\alpha : [0, l] \times (-\varepsilon, \varepsilon) \longrightarrow \mathcal{M}$  de la geodésica  $\gamma$ , definida por:

$$\alpha(t, s) = \exp_{x'} t (\exp_{x'}^{-1} \beta(s)),$$

donde,

$$\beta : (-\varepsilon, l + \varepsilon) \longrightarrow \mathcal{M},$$

es la geodésica tal que  $\beta(0) = y_{x'}$  y  $\beta(l) = \bar{x}$ .

De esta manera el campo variacional del vector  $V(t) = (\frac{\partial \alpha}{\partial s})(t, 0)$  de  $\alpha$  satisface:

$$V(0) = 0 \quad y \quad V(l) = \beta'(0) = \exp_{y_{x'}}^{-1} \bar{x}.$$

Por otra parte de (3 · 6) se tiene que:

$$L'(\gamma) := \frac{d}{ds} L(\alpha(t, s))|_{s=0} = \langle v, \gamma' \rangle|_a^b,$$

donde  $c_s(\cdot) = \alpha(\cdot, s)$  y  $s \in (-\varepsilon, l\varepsilon)$ .

Así,

$$L'(\gamma) = \frac{d}{ds} L(c_s)|_{s=0} = \left\langle \frac{-\exp_{y_{x'}}^{-1} x'}{\|\exp_{y_{x'}}^{-1} x'\|}, \exp_{y_{x'}}^{-1} \bar{x} \right\rangle \leq 0,$$

ya que:

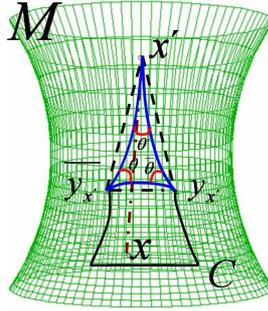
$$\begin{aligned} \left\langle \frac{-\exp_{y_{x'}}^{-1} x'}{\|\exp_{y_{x'}}^{-1} x'\|}, \exp_{y_{x'}}^{-1} \bar{x} \right\rangle &= \frac{-1}{\|\exp_{y_{x'}}^{-1} x'\|} \langle \exp_{y_{x'}}^{-1} x', \exp_{y_{x'}}^{-1} \bar{x} \rangle \\ &= \frac{-1}{\|\exp_{y_{x'}}^{-1} x'\|} \|\exp_{y_{x'}}^{-1} x'\|^2 \\ &\leq 0, \end{aligned}$$

como  $L'(\gamma) \leq 0$ , entonces  $L(\gamma)$  es decreciente.

Por lo tanto, de la desigualdad anterior existe  $\delta > 0$  tal que  $L(c_s) < L(\gamma)$  para cualquier  $0 < s < \delta$ , entonces  $d(x', \beta(s)) \leq d(x', y_{x'})$  para cualquier  $0 < s < \delta$ .

Además como  $\mathcal{C}$  es convexa,  $\beta(s) \in \mathcal{C}$  para cualquier  $0 < s < \delta$ . Entonces la desigualdad  $d(x', \beta(s)) \leq d(x', y_{x'})$  para cualquier  $0 < s < \delta$ , contradice la Proposición 4.1.1 de  $y_{x'}$ .

**Proposición 4.1.3.** *Sea  $x'$  en  $\mathcal{M}$ . Existe solamente un  $y_{x'}$  en  $\mathcal{C}$  tal que:  $d(x', y_{x'}) \leq d(x', x)$ , para todo  $x \in \mathcal{C}$ .*



### Demostración.

Si  $x' \in \mathcal{C}$ , el resultado es trivial.

Supongamos que  $x' \notin \mathcal{C}$ . Asumamos lo contrario, que existe  $y_{x'}$  y  $\bar{y}_{x'}$  en  $\mathcal{C}$ , con  $y_{x'} \neq \bar{y}_{x'}$  y tal que:

$$d(x', y_{x'}) = d(x', \bar{y}_{x'}) \leq d(x', x),$$

para todo  $x \in \mathcal{M}$ .

Ahora considere el triángulo geodésico  $\Delta(x', y_{x'}, \bar{y}_{x'})$  y fije  $\theta, \bar{\theta}$  y  $\theta'$  sus ángulos internos.

Nótese que:

$$\theta = \angle(\exp_{y_{x'}}^{-1} x', \exp_{y_{x'}}^{-1} \bar{y}_{x'}),$$

$$\bar{\theta} = \angle(\exp_{\bar{y}_{x'}}^{-1} x', \exp_{\bar{y}_{x'}}^{-1} y_{x'}),$$

y

$$\theta' = \angle(\exp_{x'}^{-1} y_{x'}, \exp_{x'}^{-1} \bar{y}_{x'}).$$

Además, como  $y_{x'} \neq \bar{y}_{x'}$  por la Proposición 4.1.2 se tiene que:

$$\theta \geq \frac{\Pi}{2}, \quad \bar{\theta} \geq \frac{\Pi}{2}.$$

Como  $d(x', y_{x'}) = d(x', \bar{y}_{x'})$  y  $\mathcal{C}$  es un conjunto convexo se observa que  $\theta' > 0$ .

Así,

$$\theta + \bar{\theta} + \theta' > \Pi,$$

y esto contradice el Teorema 3.2.

**Definición 4.1.2.** El punto único dado por la Proposición 4 · 1 · 3, se llama proyección de  $x'$  sobre el convexo  $\mathcal{C}$  y es denotado por  $P_C(x')$ .

**Corolario 4.1.1.** Sea  $x'$  en  $\mathcal{M}$ . Existe una única proyección  $P_C(x')$ .  
Además, de la siguiente desigualdad

$$\langle \exp_{P_C(x')}^{-1} x', \exp_{P_C(x')}^{-1} x \rangle \leq 0, \quad (4.2)$$

para todo  $x \in \mathcal{C}$ .

**Demostración.**

Si  $x' \in \mathcal{C}$ , el resultado es trivial.

Supongamos que  $x' \notin \mathcal{C}$ . Fijemos  $l = d(x', P_C(x'))$  y sea  $\gamma : [0, l] \rightarrow \mathcal{M}$  la geodésica tal que  $\gamma(0) = x'$  y  $\gamma(l) = P_C(x')$ .

Recordemos que:

Por definición de distancia:

$$d(x', x) = \inf\{L(c)\},$$

donde  $c$  es una curva suave a trozos que conecta a  $x'$  con  $x$ , y por definición de segmento geodésico minimal se tiene:

$$d(x, x') = L(\gamma).$$

Por definición de mapeo exponencial tenemos:

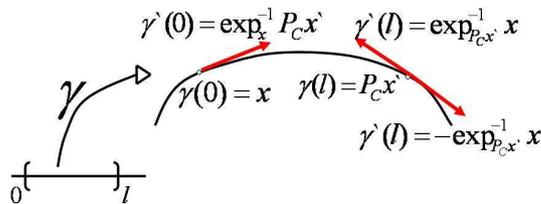
$$\exp_x : T_x \mathcal{M} \rightarrow \mathcal{M},$$

y esta dado por  $\exp_x v = \gamma_v(1, x)$ .

Y por definición de mapeo exponencial inverso:

$$\exp_x^{-1} : \mathcal{M} \rightarrow T_x \mathcal{M},$$

y esta dado por:  $\exp_x^{-1} P_C(x') = \gamma'(0)$  y  $\exp_{P_C(x')}^{-1} x = \gamma'(l)$ .



Así,

$$\gamma'(l) = \frac{-(\exp_{P_C(x')}^{-1}x')}{\|\exp_{P_C(x')}^{-1}x'\|},$$

esta normalizado.

Ahora supongamos lo contrario, que existe un punto  $\bar{x} \in \mathcal{C}$  tal que:

$$\langle \exp_{P_C(x')}^{-1}x', \exp_{P_C(x')}^{-1}\bar{x} \rangle > 0.$$

Considere la variación  $\alpha : [0, l] \times (-\varepsilon, \varepsilon) \longrightarrow \mathcal{M}$  de la geodésica  $\gamma$ , definida por:

$$\alpha(t, s) = \exp_{x'}t(\exp_{x'}^{-1}\beta(s)),$$

donde  $\beta : (-\varepsilon, l + \varepsilon) \longrightarrow \mathcal{M}$  es la geodésica tal que  $\beta(0) = P_C(x')$  y  $\beta(l) = \bar{x}$ .

De esta manera el campo variacional del vector  $V(t) = (\frac{\partial \alpha}{\partial s})(t, 0)$  de  $\alpha$  satisface  $V(0) = 0$  y  $V(l) = \beta'(0) = \exp_{P_C(x')}^{-1}\bar{x}$ .

Entonces de (3 · 6) se tiene que:

$$L'(\gamma) := \frac{d}{ds}L(C_s)|_{s=0} = \langle v, \gamma' \rangle|_a^b,$$

donde  $c_s(\cdot) = \alpha(\cdot, s)$  y  $s \in (-\varepsilon, l\varepsilon)$ .

Así,

$$L'(\gamma) = \frac{d}{ds}L(c_s)|_{s=0} = \left\langle \frac{-\exp_{P_C(x')}^{-1}x'}{\|\exp_{P_C(x')}^{-1}x'\|}, \exp_{P_C(x')}^{-1}\bar{x} \right\rangle \leq 0,$$

ya que:

$$\begin{aligned} \left\langle \frac{-\exp_{P_C(x')}^{-1}x'}{\|\exp_{P_C(x')}^{-1}x'\|}, \exp_{P_C(x')}^{-1}\bar{x} \right\rangle &= \frac{-1}{\|\exp_{P_C(x')}^{-1}x'\|} \langle \exp_{P_C(x')}^{-1}x', \exp_{P_C(x')}^{-1}\bar{x} \rangle \\ &= \frac{-1}{\|\exp_{P_C(x')}^{-1}x'\|} \|\exp_{P_C(x')}^{-1}x'\|^2 \\ &\leq 0, \end{aligned}$$

como  $L'(\gamma) \leq 0$ , entonces  $L(\gamma)$  es decreciente.

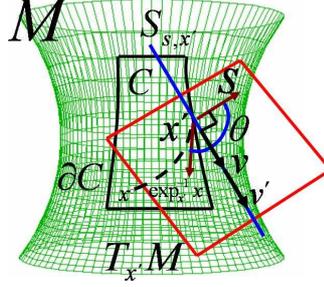
Por lo tanto, de la desigualdad anterior existe  $\delta > 0$  tal que  $L(c_s) < L(\gamma)$  para cualquier  $0 < s < \delta$ , entonces  $d(x', \beta(s)) \leq d(x', P_C(x'))$  para cualquier  $0 < s < \delta$ .

Además como  $\mathcal{C}$  es convexa,  $\beta(s) \in \mathcal{C}$  para cualquier  $0 < s < \delta$ . Entonces la desigualdad  $d(x', \beta(s)) \leq d(x', P_C(x'))$  para cualquier  $0 < s < \delta$ , contradice la Proposición 4 · 1 · 1 de  $P_C(x')$ .

**Definición 4.1.3.** Suponga que, la frontera  $\partial\mathcal{C}$  de  $\mathcal{C}$  es no vacío,  $x' \in \partial\mathcal{C}$ ,  $s \in T_{x'}\mathcal{M}$  y  $s \neq 0$ . El subespacio de  $T_{x'}\mathcal{M}$  definida por:

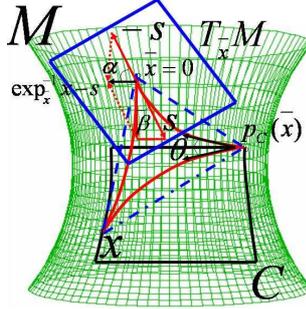
$$S_{s,x'} := \{v \in T_{x'}\mathcal{M} / \langle s, v \rangle = 0\}, \quad (4.3)$$

es llamado el soporte de el conjunto  $\mathcal{C}$  en  $x'$  si  $\langle s, \exp_{x'}^{-1}x \rangle \leq 0$ , para todo  $x \in \mathcal{C}$ .



**Lema 4.1.1.** Sea  $\bar{x}$  en  $\mathcal{M}$  tal que  $\bar{x} \in \mathcal{C}$ . Entonces existe  $s \in T_{\bar{x}}\mathcal{M}$  tal que:

$$\sup\{\langle -s, \exp_{\bar{x}}^{-1}x \rangle : x \in \mathcal{C}\} < 0. \quad (4.4)$$



**Demostración.**

Tomenos  $x \in \mathcal{C}$ .

Considere el triángulo geodésico  $\Delta(\bar{x}, x, p_C(\bar{x}))$ .

Fijemos,  $\beta = \angle(\exp_{\bar{x}}^{-1}p_C(\bar{x}), \exp_{\bar{x}}^{-1}x)$  y  $\theta = \angle(\exp_{p_C(\bar{x})}^{-1}\bar{x}, \exp_{p_C(\bar{x})}^{-1}x)$ .

Por (3 · 3) se tiene que:

$$d(\bar{x}, x)\cos\beta + d(x, p_C(\bar{x}))\cos\theta \geq d(\bar{x}, p_C(\bar{x})). \quad (4.5)$$

Ahora consideremos en  $T_{\bar{x}}\mathcal{M}$  el triángulo geodésico  $\Delta(0, \exp_{\bar{x}}^{-1}x, \exp_{\bar{x}}^{-1}p_C(\bar{x}))$ .

Sea  $s := \exp_{\bar{x}}^{-1}p_C(\bar{x})$  y  $\alpha = \angle(-s, \exp_{\bar{x}}^{-1}x - s)$ .

Observando la desigualdad (3 · 3), en  $T_{\bar{x}}\mathcal{M}$  se tiene la igualdad siguiente:

$$\|\exp_{\bar{x}}^{-1}x\|\cos\beta + \|\exp_{\bar{x}}^{-1}x - s\|\cos\alpha = \|s\|, \quad (4.6)$$

además,  $d(\bar{x}, p_C(\bar{x})) = \|\exp_{p_C(\bar{x})}^{-1}\bar{x}\|$  y claramente  $\|s\| = \|\exp_{\bar{x}}^{-1}p_C(\bar{x})\|$ .

De (4 · 5) se tiene:

$$d(\bar{x}, x)\cos\beta + d(x, p_C(\bar{x}))\cos\theta \geq d(\bar{x}, p_C(\bar{x})),$$

como  $\|\exp_{\bar{x}}^{-1}x\| = d(\bar{x}, x)$  y  $\|s\| = \|\exp_{\bar{x}}^{-1}P_C(\bar{x})\| = d(\bar{x}, P_C(\bar{x}))$ , entonces,

$$\|\exp_{\bar{x}}^{-1}x\|\cos\beta + d(x, P_C(\bar{x}))\cos\theta \geq \|s\|, \quad (4.7)$$

luego, de (4 · 6) y (4 · 7) se tiene que:

$$\begin{aligned} \|\exp_{\bar{x}}^{-1}x\|\cos\beta + \|\exp_{\bar{x}}^{-1}x - s\|\cos\alpha &= \|s\| \\ \|\exp_{\bar{x}}^{-1}x\|\cos\beta + d(x, P_C(\bar{x}))\cos\theta &\geq \|s\|, \end{aligned}$$

entonces,

$$\|\exp_{\bar{x}}^{-1}x\|\cos\beta - \|\exp_{\bar{x}}^{-1}x\|\cos\beta + d(x, P_C(\bar{x}))\cos\theta - \|\exp_{\bar{x}}^{-1}x - s\|\cos\alpha \geq \|s\| - \|s\|$$

$$\begin{aligned} \Rightarrow d(x, P_C(\bar{x}))\cos\theta - \|\exp_{\bar{x}}^{-1}x - s\|\cos\alpha &\geq 0 \\ \Rightarrow d(x, P_C(\bar{x}))\cos\theta &\geq \|\exp_{\bar{x}}^{-1}x - s\|\cos\alpha. \end{aligned}$$

Nótese que:

$$d(x, p_C(\bar{x})) = \|\exp_{p_C(\bar{x})}^{-1}x\|,$$

entonces esto implica que:

$$\langle -s, \exp_{\bar{x}}^{-1}x - s \rangle \leq \langle \exp_{p_C(\bar{x})}^{-1}\bar{x}, \exp_{p_C(\bar{x})}^{-1}x \rangle,$$

por el Corolario 4 · 1 · 1,

$$\langle \exp_{p_C(\bar{x})}^{-1}\bar{x}, \exp_{p_C(\bar{x})}^{-1}x \rangle \leq 0,$$

entonces por esto, y la desigualdad anterior, se tiene que:

$$\langle -s, \exp_{\bar{x}}^{-1}x - s \rangle \leq \langle \exp_{p_C(\bar{x})}^{-1}\bar{x}, \exp_{p_C(\bar{x})}^{-1}x \rangle \leq 0,$$

la cual implica que:

$$\langle -s, \exp_{\bar{x}}^{-1}x - s \rangle \leq 0, \quad (4.8)$$

luego, por propiedades se tiene:

$$\begin{aligned} \langle -s, \exp_{\bar{x}}^{-1}x - s \rangle &= \langle -s, \exp_{\bar{x}}^{-1}x + (-s) \rangle \\ &= \langle -s, \exp_{\bar{x}}^{-1}x \rangle + \langle -s, -s \rangle, \end{aligned}$$

así,

$$\langle -s, \exp_x^{-1}x - s \rangle = \langle -s, \exp_x^{-1}x + (-s) \rangle = \langle -s, \exp_x^{-1}x \rangle + \langle -s, -s \rangle, \quad (4.9)$$

luego de (4 · 8) y (4 · 9) se obtiene:

$$0 \geq \langle -s, \exp_x^{-1}x \rangle + \langle -s, -s \rangle \Rightarrow \langle -s, \exp_x^{-1}x \rangle \leq \langle -s, -s \rangle,$$

pero, como:

$$\langle -s, -s \rangle = \| -s \|^2 = \|s\|^2 = s \cdot s \cdot \cos \alpha' = s \cdot s \cdot \cos \alpha'$$

y

$$\|s\|^2 = (\sqrt{\langle s, s \rangle})^2 = |\langle s, s \rangle| = s \cdot s \cdot \cos \alpha',$$

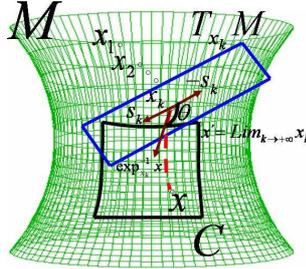
se tiene que:

$$\langle -s, -s \rangle = \|s\|^2,$$

por lo tanto,

$$\langle -s, \exp_x^{-1}x \rangle \leq \langle -s, -s \rangle = -\|s\|^2 \Rightarrow \langle -s, \exp_x^{-1}x \rangle \leq -\|s\|^2.$$

**Teorema 4.1.1.** *Sea  $x' \in \partial \mathcal{C}$ . Existe un subespacio soporte de  $\mathcal{C}$  en  $x'$ .*



**Demostración.**

Considere el único fibrado tangente  $UM := \bigcup_{x \in \mathcal{M}} U_x \mathcal{M}$ , donde,

$$U_x \mathcal{M} = \{u \in T_x \mathcal{M}; \|u\| = 1\}.$$

Tomemos la sucesión  $\{x_k\}$  en  $\mathcal{M}$  tal que  $x_k \in \mathcal{C}$ , para  $k = 1, 2, \dots$  y

$$\lim_{k \rightarrow +\infty} x_k = x'.$$

Para cada  $x_k$  por el Lema 4 · 1 · 1 existe  $s_k \in T_{x_k} \mathcal{M}$ , tal que:

$$\langle -s_k, \exp_{x_k}^{-1}x \rangle < 0, \quad (4.10)$$

para todo  $x \in \mathcal{C}$ .

Sin pérdida de generalidad, podemos asumir que:

$$\|s_k\| = 1,$$

es decir,

$$UM := \bigcup_{x \in \mathcal{M}} U_x \mathcal{M},$$

donde,

$$U_x \mathcal{M} = \{s_k \in T_{x_k} \mathcal{M}; \|s_k\| = 1\},$$

así, la sucesión  $\{s_k\}$  esta en  $UM$ .

Además como  $\lim_{k \rightarrow +\infty} x_k = x'$ , se tiene que el subconjunto

$$B := \bigcup_{i=1}^n U_{x_k} \mathcal{M}$$

de  $UM$  es compacto.

Nótese que  $\{s_k\} \in B$ , esto es, existe una subsucesión  $\{s_{k_j}\}$  de  $\{s_k\}$  y  $\bar{s} \in T_{x'} \mathcal{M}$  tal que:

$$\lim_{k \rightarrow +\infty} s_{k_j} = \bar{s}.$$

Por (4 · 10) se tiene:

$$\langle -s_k, \exp_{x_k}^{-1} x \rangle < 0,$$

entonces,

$$\langle -s_{k_j}, \exp_{x_{k_j}}^{-1} x \rangle < 0, \tag{4.11}$$

para todo  $x \in \mathcal{C}$ .

Ahora tomando límite en la desigualdad (4 · 11) obtenemos:

$$\langle -\bar{s}, \exp_{x'}^{-1} x \rangle \leq 0,$$

para todo  $x \in \mathcal{C}$ , pues  $\lim_{k \rightarrow +\infty} x_{k_j} = \bar{s}$  y  $\lim_{k \rightarrow +\infty} x_k = x'$ .

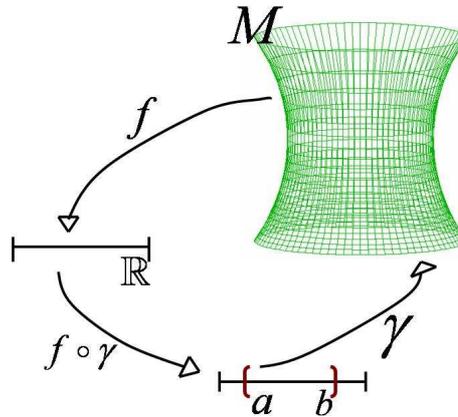
Por lo tanto,  $S_{s,x'}$  esta definido por (4 · 3):

$$S_{s,x'} = \{v \in T_{x'} \mathcal{M} / \langle s, v \rangle = 0\},$$

donde  $s = -\bar{s}$ , es el subespacio soporte a  $\mathcal{C}$  en  $x'$ .

## 4.2. Funciones Convexas.

**Definición 4.2.1.** *Sea  $\mathcal{M}$  una variedad de Hadamard. Una función  $f : \mathcal{M} \rightarrow \mathbb{R}$  es llamada convexa (respectivamente, estrictamente convexa) si la composición  $f \circ \gamma : \mathbb{R} \rightarrow \mathbb{R}$  es convexa (respectivamente, estrictamente convexa) para cualquier geodésica  $\gamma$  de  $\mathcal{M}$ .*

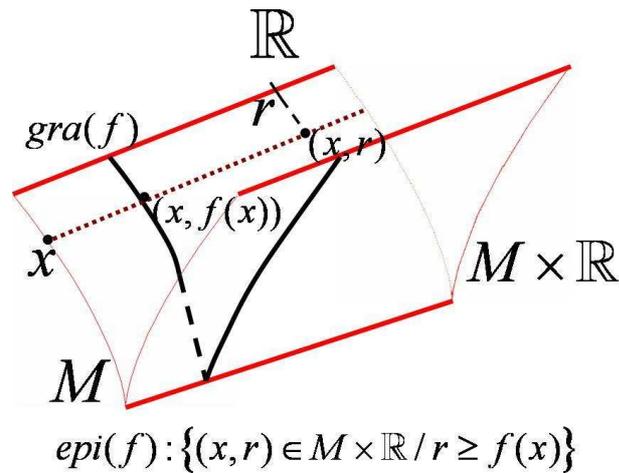


**Proposición 4.2.1.** *Todas las funciones convexas son continuas.*

**Demostración:** Ver [4].

**Definición 4.2.2.** *El epigrafo de una función  $f : \mathcal{M} \rightarrow \mathbb{R}$  se define como:*

$$\text{epi}(f) = \{(x, r) \in \mathcal{M} \times \mathbb{R} / r \geq f(x)\}.$$



**Definición 4.2.3.** *Sea  $\mathcal{M}$  una variedad de Hadamard y  $f : \mathcal{M} \rightarrow \mathbb{R}$  una función.*

a) *Si  $f(\gamma_{xy}(t)) \leq (1-t)f(x) + tf(y)$ ,  $\forall x, y \in \mathcal{M}$ ,  $\forall \gamma_{xy}$ ,  $\forall t \in [0, 1]$  entonces la función  $f$  es llamada convexa.*

**Teorema 4.2.1.** *Sea  $\mathcal{M}$  una variedad de Hadamard. Una función  $f : \mathcal{M} \rightarrow \mathbb{R}$  es convexa si y solo si su epigrafo*

$$\text{epi}(f) = \{(x, r) \in \mathcal{M} \times \mathbb{R} / r \geq f(x)\} \quad (4.12)$$

*es un subconjunto convexo en  $\mathcal{M} \times \mathbb{R}$ .*

**Demostración:**

( $\Rightarrow$ ) El  $\text{epi}(f)$  es el subconjunto de  $\mathcal{M} \times \mathbb{R}$ .

$$\text{epi}(f) = \{(x, r) \in \mathcal{M} \times \mathbb{R} / r \geq f(x)\}$$

Supongamos que  $f$  es convexa.

Tenemos que probar que el  $\text{epi}(f)$  es un subconjunto convexo, es decir,  $\forall (x_1, r_1), (x_2, r_2) \in \text{epi}(f)$  se verifica:

$$\lambda(x_1, r_1) + (1 - \lambda)(x_2, r_2) \in \text{epi}(f), \quad \forall \lambda \in [0, 1]$$

como  $(x_1, r_1), (x_2, r_2) \in \text{epi}(f)$  entonces,

$$(a) \quad r_1 \geq f(x_1)$$

y

$$(b) \quad r_2 \geq f(x_2),$$

con  $x_1, x_2 \in \mathcal{M}$ ,  $r_1, r_2 \in \mathbb{R}$ .

Al ser  $\mathcal{M}$  de Hadamard, se tiene que:

$$\gamma_{x_1 x_2} \in \mathcal{M},$$

luego,

$$(1 - \lambda)r_1 + \lambda r_2 \geq (1 - \lambda)f(x_1) + \lambda f(x_2), \text{ por } (a) \text{ y } (b),$$

así,

$$(1 - \lambda)r_1 + \lambda r_2 \geq (1 - \lambda)f(x_1) + \lambda f(x_2) \geq f \circ \gamma_{xy}(t), \quad \text{por definición.}$$

Por lo tanto,

$$(\gamma_{xy}(t), (1 - \lambda)r_1 + \lambda r_2) \in \text{epi}(f).$$

( $\Leftarrow$ ) Supongamos que el  $\text{epi}(f)$  es un subconjunto convexo.

Probemos que  $f$  es una función convexa.

Dados  $x_1, x_2 \in \mathcal{M}$ , los puntos  $(x_1, f(x_1)), (x_2, f(x_2))$  pertenecen al  $\text{epi}(f)$  pues:

$$f(x_1) \geq f(x_1)$$

y

$$f(x_2) \geq f(x_2).$$

Por ser convexo se verifica:

$$(\gamma_{xy}(t), (1 - \lambda)f(x) + \lambda f(y)) \in \text{epi}(f),$$

así,

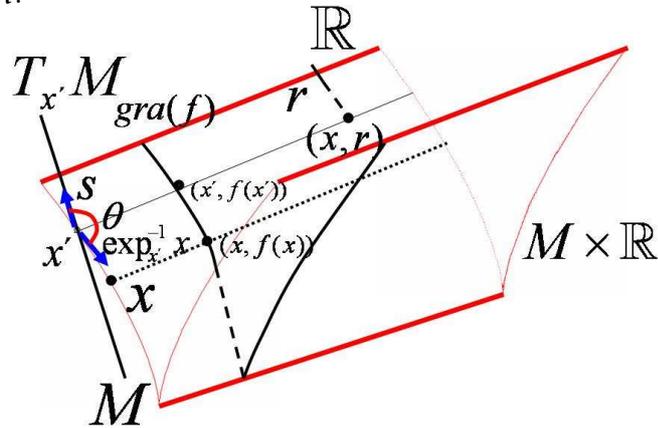
$$f(\gamma_{xy}(t)) \leq (1 - \lambda)f(x) + \lambda f(y).$$

Por lo tanto,  $f$  es convexa.

**Definición 4.2.4.** Sea  $\mathcal{M}$  una variedad de Hadamard, sea  $f : \mathcal{M} \rightarrow \mathbb{R}$  una función convexa. Tomemos  $x' \in \mathcal{M}$ , el vector  $s \in T_{x'}\mathcal{M}$  es llamado el subgradiente de  $f$  a  $x' \in \mathcal{M}$  si

$$f(x) \geq f(x') + \langle s, \exp_{x'}^{-1}x \rangle, \quad (4.13)$$

para todo  $x \in \mathcal{M}$ .



**Definición 4.2.5.** El conjunto de todos los subgradientes de  $f$  a  $x'$  es llamado el subdiferencial de  $f$  a  $x'$  y es denotado por  $\partial f(x')$ .

**Teorema 4.2.2.** Sea  $\mathcal{M}$  una variedad de Hadamard y sea  $f : \mathcal{M} \rightarrow \mathbb{R}$  convexa. Entonces, para cualquier  $x' \in \mathcal{M}$ , existe  $s \in T_{x'}\mathcal{M}$  tal que

$$f(x) \geq f(x') + \langle s, \exp_{x'}^{-1}x \rangle, \quad (4.14)$$

para todo  $x \in \mathcal{M}$ . En otras palabras, el subdiferencial  $\partial f(x')$  de  $f$  a  $x' \in \mathcal{M}$  es no vacío.

**Demostración.**

Como  $f$  es convexa entonces por definición  $f$  es continua, y por el Teorema 4.2.1 el  $\text{epi}(f)$  de  $f$  es un subconjunto convexo cerrado de  $\mathcal{M} \times \mathbb{R}$ .

Más aún, la frontera de el  $\text{epi}(f)$  de  $f$  es,

$$\partial \text{epi}(f) = \{(x, f(x)) : x \in \mathcal{M}\}.$$



Tomamos  $x = \bar{x}$  y  $r = \bar{r}$  en (4 · 15) obtenemos:

$$\begin{aligned} \langle s, \exp_{x'}^{-1}x \rangle + \alpha(r - f(x')) &= \langle s, \exp_{x'}^{-1}\bar{x} \rangle + \alpha(\bar{r} - f(x')) \\ &= \langle s, \exp_{x'}^{-1}\exp_{x'}s \rangle + \alpha(\bar{r} - f(x')) \\ &= \langle s, s \rangle + \alpha(\bar{r} - f(x')), \end{aligned}$$

como  $\|s\| = \sqrt{\langle s, s \rangle}$ , entonces  $\|s\|^2 = \langle s, s \rangle$ .

Así,

$$\langle s, \exp_{x'}^{-1}x \rangle + \alpha(r - f(x')) = \langle s, s \rangle + \alpha(\bar{r} - f(x')) = \|s\|^2 + \alpha(\bar{r} - f(x')), \quad (4.16)$$

luego, como por (4 · 15):

$$\langle s, \exp_{x'}^{-1}x \rangle + \alpha(r - f(x')) = \langle (s, \alpha), (\exp_{x'}^{-1}x, r - f(x')) \rangle \leq 0,$$

se tiene de (4 · 16) que:

$$\|s\|^2 + \alpha(\bar{r} - f(x')) \leq 0 \Rightarrow \|s\|^2 \leq \alpha(f(x') - \bar{r}). \quad (4.17)$$

Así se tiene que  $\alpha \neq 0$ , (ya que si  $\alpha = 0$ , se tendría de (4 · 17) que  $s = 0$  y eso no es posible ya que por definición de subespacio soporte  $s \neq 0$ ) pues el subespacio soporte es no trivial.

Sin pérdida de generalidad, tomemos  $\alpha = -1$ , entonces tomando  $r = f(x)$  y  $\alpha = -1$  en (4 · 15) se tiene que:

$$\begin{aligned} \langle s, \exp_{x'}^{-1}x \rangle + \alpha(r - f(x')) \leq 0 &\Rightarrow \langle s, \exp_{x'}^{-1}x \rangle + (-1)(r - f(x')) \leq 0 \\ &\Rightarrow \langle s, \exp_{x'}^{-1}x \rangle + (f(x') - f(x)) \leq 0. \end{aligned}$$

Por lo tanto,

$$\langle s, \exp_{x'}^{-1}x \rangle + f(x') \leq f(x) \Rightarrow f(x) \geq f(x') + \langle s, \exp_{x'}^{-1}x \rangle,$$

para todo  $x \in \mathcal{M}$ .

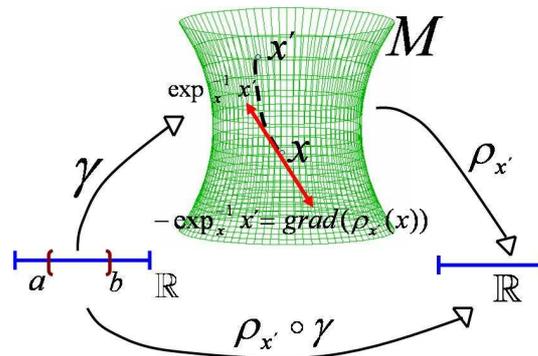
**Definición 4.2.6.** Sea  $\mathcal{M}$  una variedad de Hadamard. Para cualquier  $x' \in \mathcal{M}$  podemos definir el mapeo exponencial inverso  $\exp_{x'}^{-1} : \mathcal{M} \rightarrow T_{x'}\mathcal{M}$  el cual es  $C^\infty$ ; como  $d(x', x) = \|\exp_{x'}^{-1}x\|$ , entonces el mapeo  $\rho_{x'} : \mathcal{M} \rightarrow \mathbb{R}$  es definido por:

$$\rho_{x'}(x) = \frac{1}{2}d^2(x, x') \quad (4.18)$$

es  $C^\infty$  también.

**Proposición 4.2.2.** Sea  $\mathcal{M}$  una variedad de Hadamard. Sea  $x' \in \mathcal{M}$ , entonces el mapeo  $\rho_{x'}$ , definido en (4 · 18), es estrictamente convexo y su gradiente a  $x$  es

$$\text{grad}\rho_{x'}(x) = -\exp_x^{-1}x'.$$



**Demostración:** Ver [1].

**Proposición 4.2.3.**  $x_0$  es un minimizador de  $f$  si y solo si  $0 \in \partial f(x_0)$ .

**Demostración:**

$$\begin{aligned}
 x_0 \text{ es minimizador de } f &\Leftrightarrow \exists V_{x_0} \text{ tal que } f(x) \geq f(x_0), \quad \forall x \in V_{x_0} \\
 &\Leftrightarrow f(x) \geq f(x_0) + \langle 0, \exp_{x_0} x \rangle \\
 &\Leftrightarrow 0 \in \partial f(x_0).
 \end{aligned}$$

# Capítulo 5

## Regularización.

**Definición 5.1.** Sea  $\mathcal{M}$  una variedad de Hadamard y sea  $f : \mathcal{M} \rightarrow \mathbb{R}$  es una función convexa. Para  $\lambda > 0$ , la regularización MOREAU-YOSIDA  $f_\lambda : \mathcal{M} \rightarrow \mathbb{R}$  de  $f$  es definida por

$$f_\lambda(x) = \inf_{y \in \mathcal{M}} \{f(y) + \lambda \rho_x(y)\}, \quad (5.1)$$

donde  $\rho_x$  esta dado en (4.18).

**Definición 5.2.** Una función  $h : \mathcal{M} \rightarrow \mathbb{R}$  es llamado 1-coerciva en  $x \in \mathcal{M}$  si

$$\lim_{d(x,y) \rightarrow +\infty} \frac{h(y)}{d(x,y)} = +\infty. \quad (5.2)$$

**Definición 5.3.** Si  $h : \mathcal{M} \rightarrow \mathbb{R}$  es una función 1-coerciva en  $x \in \mathcal{M}$ , entonces el conjunto minimizante de  $h$  es no vacio.

**Lema 5.1.** Si  $f : \mathcal{M} \rightarrow \mathbb{R}$  es convexo y  $\lambda > 0$ , entonces la función  $(f + \lambda \rho_x) : \mathcal{M} \rightarrow \mathbb{R}$  es 1-coerciva en  $x \in \mathcal{M}$ .

**Demostración.**

Como  $f$  es convexo, por el Teorema 4.2.2, se tiene que existe  $s \in T_x \mathcal{M}$  tal que:

$$\begin{aligned} \frac{(f + \lambda \rho_x)(y)}{d(x,y)} &= \frac{f(y)}{d(x,y)} + \lambda \frac{\rho_x(y)}{d(x,y)} \\ &\geq \frac{f(x)}{d(x,y)} + \left\langle s, \frac{\exp_x^{-1} y}{d(x,y)} \right\rangle + \lambda \frac{\rho_x(y)}{d(x,y)} \\ &= \frac{f(x)}{d(x,y)} + \left\langle s, \frac{\exp_x^{-1} y}{d(x,y)} \right\rangle + \frac{\lambda}{2} d(x,y), \end{aligned}$$

ya que  $\rho_x(y) = \frac{1}{2} d^2(x,y)$ .

Aplicando limite en ambos lados, se obtiene que:

$$\lim_{d(x,y) \rightarrow +\infty} \frac{(f + \lambda\rho_x)(y)}{d(x,y)} \geq \lim_{d(x,y) \rightarrow +\infty} \frac{f(x)}{d(x,y)} + \langle s, \frac{\exp_x^{-1}y}{d(x,y)} \rangle + \frac{\lambda}{2}d(x,y),$$

entonces,

$$\lim_{d(x,y) \rightarrow +\infty} \frac{(f + \lambda\rho_x)(y)}{d(x,y)} = +\infty.$$

**Lema 5.2.** Si  $f : \mathcal{M} \rightarrow \mathbb{R}$  es convexo entonces, para cualquier  $x \in \mathcal{M}$  y  $\lambda > 0$  existe un único punto, denotado por  $P_\lambda(x)$ , tal que

$$f(P_\lambda(x)) + \lambda\rho_x(P_\lambda(x)) = f_\lambda(x). \quad (5.3)$$

caracterizada por

$$\lambda(\exp_{p_\lambda(x)}^{-1}x) \in \partial f(P_\lambda(x)). \quad (5.4)$$

### Demostración.

Por la Proposición 4·2·1, tenemos que  $\rho_x$  es estrictamente convexa, y así  $f + \lambda\rho_x$  es estrictamente convexa.

Por lo tanto, tiene a lo más un único punto mínimo (por ser  $f$  convexa y  $f + \lambda\rho_x$  es estrictamente convexa).

Por otra parte, por el Lema 5·1,  $f + \lambda\rho_x$  es 1-coerciva, entonces tiene por lo menos un mínimo.

Así,  $f + \lambda\rho_x$  tiene un único mínimo, y de la desigualdad

$$f_\lambda(x) = \inf_{y \in \mathcal{M}} \{f(y) + \lambda\rho_x(y)\},$$

se tiene que como  $f + \lambda\rho_x$  tiene un único punto minimo, digamos  $P_\lambda(x)$ , entonces:

$$f_\lambda(x) = f(P_\lambda(x)) + \lambda\rho_x(P_\lambda(x)),$$

esto implica que:

$$0 \in \partial(f + \lambda\rho_x)(P_\lambda(x)),$$

Por otra parte,

$$\partial(f + \lambda\rho_x)(y) = \partial f(y) + \lambda\partial(\rho_x)(y)$$

y como,  $\rho_x$  es  $C^\infty$  y el subdiferencial es el conjunto de todos los subgradiente, y

el subgradiente es unitario, por ser  $\rho_x$  diferenciable, se tiene entonces, que por la Proposición 4 · 2 · 2,  $\text{grad}\rho_{x'}(x) = -\text{exp}_x^{-1}x'$ .

Así,

$$\begin{aligned}\partial(f + \lambda\rho_x)(y) &= \partial f(y) + \lambda\partial(\rho_x)(y) \\ &= \partial f(y) + \{-\lambda\text{exp}_y^{-1}x\}\end{aligned}$$

es decir, (5 · 4) es una caracterización de  $P_\lambda(x)$ .

Luego, como  $0 \in \partial(f + \lambda\rho_x)(P_\lambda(x))$ , se tiene que:

$$\partial f(P_\lambda(x)) + \{-\lambda\text{exp}_{P_\lambda(x)}^{-1}x\} = 0,$$

entonces,

$$\partial f(P_\lambda(x)) = \lambda(\text{exp}_{P_\lambda(x)}^{-1}x).$$

Así,

$$\lambda(\text{exp}_{P_\lambda(x)}^{-1}x) \in \partial f(P_\lambda(x)).$$

## Capítulo 6

# Algoritmo de Punto Proximal sobre Variedades de Riemann.

Sea  $\mathcal{M}$  una variedad de Hadamard. Sea  $f : \mathcal{M} \rightarrow \mathbb{R}$  una función convexa.

El algoritmo de punto proximal genera, para un punto de partida  $p_0 \in \mathcal{M}$ , una sucesión  $\{x_k\} \subset \mathcal{M}$  por la iteración

$$x_{k+1} = P_{\lambda_k}(x_k) = \arg \min_{y \in \mathcal{M}} \{f(y) + \lambda_k \rho_{x_k}(y)\}, \quad (6.1)$$

donde  $\rho_{x_k}$  es definido en (4 · 18) y  $\{\lambda_k\}$  es una sucesión de números positivos.

En el caso particular en que  $\mathcal{M} = \mathbb{R}^n$  se tiene que  $\rho_{x_k}(y) = \|x_k - y\|$ . En este caso, la iteración (6 · 1) se reduce a

$$x_{k+1} = \arg \min_{y \in \mathbb{R}^n} \{f(y) + \lambda_k \|x_k - y\|^2\}.$$

**Teorema 6.1.** *Sea  $f : \mathcal{M} \rightarrow \mathbb{R}$  una función convexa, donde  $\mathcal{M}$  es una variedad de Hadamard. La sucesión  $\{x_k\}$  generada por (6 · 1) está bien definida, y caracterizada por*

$$\lambda_k (\exp_{x_{k+1}}^{-1} x_k) \in \partial f(x_{k+1}). \quad (6.2)$$

### **Demostración.**

Por la Proposición 4 · 2 · 1, tenemos que  $\rho_{x_k}$  es estrictamente convexa, y así  $f + \lambda_k \rho_{x_k}$  es estrictamente convexa. Esto es, tiene a lo más un único punto mínimo.

Por otra parte, por el Lema 5 · 1,  $f + \lambda_k \rho_{x_k}$  es 1-coerciva entonces tiene por lo menos un mínimo.

Así,  $f + \lambda \rho_x$  tiene un único mínimo, y de la desigualdad

$$f_\lambda(x) = \inf_{y \in \mathcal{M}} \{f(y) + \lambda \rho_x(y)\},$$

se tiene que como  $f + \lambda\rho_{x_k}$  tiene un único punto mínimo, digamos  $x_{k+1}$  se obtiene:

$$f_\lambda(x) = f(x_{k+1}) + \lambda\rho_{x_k}(x_{k+1}),$$

esto implica que:

$$0 \in \partial(f + \lambda\rho_{x_k})(x_{k+1}).$$

Por otra parte,

$$\partial(f + \lambda_k\rho_{x_k})(x_{k+1}) = \partial(f)(x_{k+1}) + \lambda_k\partial(\rho_{x_k})(x_{k+1})$$

y como,  $\rho_x$  es  $C^\infty$  y el subdiferencial es el conjunto de todos los subgradiente, y el subgradiente es unitario, por ser  $\rho_x$  diferenciable, se tiene entonces, que por la Proposicion 4 · 2 · 2,  $grad\rho_{x'}(x) = -exp_x^{-1}x'$ .

Así,

$$\begin{aligned} \partial(f + \lambda_k\rho_{x_k})(x_{k+1}) &= \partial(f)(x_{k+1}) + \lambda_k\partial(\rho_{x_k})(x_{k+1}) \\ &= \partial f(x_{k+1}) + \{-\lambda_k exp_{x_{k+1}}^{-1}x_k\}(x_{k+1}), \end{aligned}$$

es decir, (6 · 2) es una caracterización de  $x_{k+1}$ .

Entonces, como  $0 \in \partial(f + \lambda\rho_{x_k})(x_{k+1})$ , se tiene que:

$$\partial f(x_{k+1}) + \{-\lambda_k exp_{x_{k+1}}^{-1}x_k\}(x_{k+1}) = 0,$$

así,

$$\lambda_k(exp_{x_{k+1}}^{-1}x_k) \in \partial f(x_{k+1}).$$

Luego,

$$\lambda_k(exp_{x_{k+1}}^{-1}x_k) \in \partial f(x_{k+1}).$$

## Capítulo 7

# Convergencia del Algoritmo de Punto Proximal.

Probaremos que la secuencia dada por (6 · 1) converge, si

$$\sum_{k=0}^{\infty} \left(\frac{1}{\lambda_k}\right) = +\infty,$$

y el conjunto minimizador

$$0^* \neq \emptyset.$$

**Definición 7.1.** Sea  $(\mathcal{M}, d)$  un espacio métrico completo. La sucesión  $\{x_k\} \subset \mathcal{M}$  es llamada Fejér convergente al conjunto no vacío  $U \subset \mathcal{M}$  cuando

$$d(x_{k+1}, y) \leq d(x_k, y), \tag{7.1}$$

para todo  $y \in U$  y  $k \geq 0$ .

**Lema 7.1.** Sea  $(\mathcal{M}, d)$  un espacio métrico completo. Si  $\{x_k\} \subset \mathcal{M}$  es Fejér convergente a un conjunto no vacío  $U \subset \mathcal{M}$ , entonces  $\{x_k\}$  es acotada. Más aún, si un punto clausura  $x$  de  $\{x_k\}$  pertenece a  $U$ , entonces  $\lim_{k \rightarrow +\infty} x_k = x$ .

**Demostración.**

Tomemos  $u \in U$ . Nótese que de (7 · 1) se tiene:

$$d(x_{k+1}, u) \leq d(x_k, u)$$

Luego,

$$\begin{array}{ll}
\text{para } k = 0, & d(x_1, u) \leq d(x_0, u) \\
\text{para } k = 1, & d(x_2, u) \leq d(x_1, u) \\
\text{para } k = 2, & d(x_3, u) \leq d(x_2, u) \\
\vdots & \vdots \\
\text{para } k = n - 1, & d(x_n, u) \leq d(x_{n-1}, u)
\end{array}$$

así,

$$d(x_k, u) \leq d(x_0, u), \quad \forall k.$$

Por lo tanto, la sucesión  $\{x_k\}$  esta contenida en una vecindad de centro  $u$  y radio  $d(x_0, u)$ , así, esta es acotada.

Ahora sea  $\{x_{k_j}\}$  una subsucesión de  $\{x_k\}$  tal que:

$$\lim_{k \rightarrow +\infty} x_{k_j} = x, \quad \text{por teorema.}$$

Como  $x \in U$ , por (7 · 1), la sucesión de números positivos  $\{d(x_k, x)\}$  es decreciente, y como  $\{d(x_k, x)\}$  es acotada entonces tiene una subsucesión  $\{d(x_{k_j}, x)\}$ , convergente.

Además como  $\{d(x_k, x)\}$  es decreciente, entonces la subsucesión  $\{d(x_{k_j}, x)\}$  converge a 0.

Esto es, que si  $\{d(x_k, x)\}$  es convergente entonces,

$$\lim_{k \rightarrow +\infty} d(x_k, x) = 0,$$

y esto implica que:

$$\lim_{k \rightarrow +\infty} x_k = x.$$

**Lema 7.2.** *Sea  $\mathcal{M}$  una variedad de Hadamard y sea  $f : \mathcal{M} \rightarrow \mathbb{R}$  una función convexa. Si la sucesión  $\{x_k\}$  es generada por (6 · 1), entonces la siguiente desigualdad*

$$d^2(x_{k+1}, x) \leq d^2(x_k, x) - d^2(x_k, x_{k+1}) + \frac{2}{\lambda_k}(f(x) - f(x_{k+1})), \quad (7.2)$$

para todo  $x \in \mathcal{M}$ .

### **Demostración.**

Tomemos  $x \in \mathcal{M}$ . Consideremos el triángulo geodésico  $\Delta(x_k x_{k+1} x)$ .

Sea  $\theta = \angle(\exp_{x_{k+1}}^{-1} x_k, \exp_{x_{k+1}}^{-1} x)$ , por el Teorema 3 · 2 tenemos:

$$d^2(x_k, x_{k+1}) + d^2(x_{k+1}, x) - 2d(x_k, x_{k+1})d(x_{k+1}, x)\cos\theta \leq d^2(x_k, x),$$

además,

$$\langle \exp_{x_{k+1}}^{-1} x_k, \exp_{x_{k+1}}^{-1} x \rangle = \|\exp_{x_{k+1}}^{-1} x_k\| \|\exp_{x_{k+1}}^{-1} x\| \cos \theta,$$

y como,  $d(x_k, x_{k+1}) = \|\exp_{x_{k+1}}^{-1} x_k\|$  y  $d(x_{k+1}, x) = \|\exp_{x_{k+1}}^{-1} x\|$ , se tiene:

$$\langle \exp_{x_{k+1}}^{-1} x_k, \exp_{x_{k+1}}^{-1} x \rangle = d(x_k, x_{k+1}) d(x_{k+1}, x) \cos \theta,$$

de modo que:

$$\begin{aligned} d^2(x_k, x_{k+1}) + d^2(x_{k+1}, x) - 2d(x_k, x_{k+1})d(x_{k+1}, x)\cos\theta &\leq d^2(x_k, x) \\ \Rightarrow d^2(x_k, x_{k+1}) + d^2(x_{k+1}, x) - 2\langle \exp_{x_{k+1}}^{-1} x_k, \exp_{x_{k+1}}^{-1} x \rangle &\leq d^2(x_k, x), \end{aligned}$$

de esta desigualdad, despejamos  $d^2(x_{k+1}, x)$  :

$$d^2(x_{k+1}, x) \leq d^2(x_k, x) - d^2(x_k, x_{k+1}) + 2\langle \exp_{x_{k+1}}^{-1} x_k, \exp_{x_{k+1}}^{-1} x \rangle,$$

y por el teorema 6 · 1, se tiene:

$$\lambda_k(\exp_{x_{k+1}}^{-1} x_k) \in \partial f(x_{k+1}),$$

esto es,

$$\lambda_k(\exp_{x_{k+1}}^{-1} x_k) = \partial f(x_{k+1}) \Rightarrow \exp_{x_{k+1}}^{-1} x_k = \frac{\partial f(x_{k+1})}{\lambda_k},$$

luego,

$$\begin{aligned} d^2(x_{k+1}, x) &\leq d^2(x_k, x) - d^2(x_k, x_{k+1}) + 2\langle \exp_{x_{k+1}}^{-1} x_k, \exp_{x_{k+1}}^{-1} x \rangle \\ \Rightarrow d^2(x_{k+1}, x) &\leq d^2(x_k, x) - d^2(x_k, x_{k+1}) + 2\left\langle \frac{\partial f(x)}{\lambda_k}, \frac{\partial f(x_{k+1})}{\lambda_k} \right\rangle \\ \Rightarrow d^2(x_{k+1}, x) &\leq d^2(x_k, x) - d^2(x_k, x_{k+1}) + \frac{2}{\lambda_k} \langle \partial f(x), \partial f(x_{k+1}) \rangle, \end{aligned}$$

y de la definición de  $\partial f(x_{k+1})$  se tiene que:

$$d^2(x_{k+1}, x) \leq d^2(x_k, x) - d^2(x_k, x_{k+1}) + \frac{2}{\lambda_k} (f(x) - f(x_{k+1})).$$

**Teorema 7.1.** *Sea  $\{x_k\}$  una sucesión generada por (6 · 1). Si la sucesión  $\{\lambda_k\}$  es tal que  $\sum_{k=0}^{\infty} (\frac{1}{\lambda_k}) = +\infty$ , entonces  $\lim_{k \rightarrow +\infty} f(x_k) = f^*$ , donde  $f^* = \inf_{x \in \mathcal{M}} f(x)$ . Si, además, el conjunto minimizador  $U^*$  es no vacío, entonces*

$$\lim_{k \rightarrow +\infty} x_k = x_*$$

y  $x_* \in U^*$ .

**Demostración.**

Si  $x_k \notin U^*$ , sustituyendo  $x$  por  $x_k$  en el Lema 7 · 2 tenemos que:

$$\begin{aligned}
d^2(x_{k+1}, x) &\leq d^2(x_k, x) - d^2(x_k, x_{k+1}) + \frac{2}{\lambda_k}(f(x) - f(x_{k+1})) \\
\Rightarrow d^2(x_{k+1}, x_k) &\leq d^2(x_k, x_k) - d^2(x_k, x_{k+1}) + \frac{2}{\lambda_k}(f(x_k) - f(x_{k+1})) \\
\Rightarrow 0 &\leq d^2(x_k, x_k) + \frac{2}{\lambda_k}(f(x_k) - f(x_{k+1})) \\
\Rightarrow 0 &\leq \frac{2}{\lambda_k}(f(x_k) - f(x_{k+1})) \\
\Rightarrow 0 &\leq f(x_k) - f(x_{k+1}) \\
\Rightarrow f(x_{k+1}) &< f(x_k),
\end{aligned}$$

entonces la sucesión  $\{f(x_k)\}$  es estrictamente decreciente.

Ahora probemos que:

$$\lim_{k \rightarrow +\infty} f(x_k) = f^*.$$

Asumamos lo contrario, que:

$$\lim_{k \rightarrow +\infty} f(x_k) > f^*,$$

entonces existe un  $x \in \mathcal{M}$  y  $\delta > 0$  tal que:

$$f(x) < f(x_k) - \delta, \quad (7.3)$$

para todo  $k$ .

Por el Lema 7 · 2 se tiene que:

$$d^2(x_{k+1}, x) \leq d^2(x_k, x) - d^2(x_k, x_{k+1}) + \frac{2}{\lambda_k}(f(x) - f(x_{k+1})),$$

y de la desigualdad (7 · 3):

$$d^2(x_{k+1}, x) \leq d^2(x_k, x) - d^2(x_k, x_{k+1}) + \frac{2}{\lambda_k}(f(x_k) - \delta - f(x_{k+1})),$$

entonces,

$$\begin{aligned}
d^2(x_{k+1}, x) &< d^2(x_k, x) - d^2(x_k, x_{k+1}) - \frac{2\delta}{\lambda_k} + \frac{2}{\lambda_k}(f(x_k) - f(x_{k+1})) \\
\Rightarrow d^2(x_{k+1}, x) &< d^2(x_k, x) - \frac{2\delta}{\lambda_k}, \quad \forall k
\end{aligned}$$

y esta desigualdad implica que:

$$\begin{aligned}
d^2(x_{k+1}, x) \leq d^2(x_k, x) - \frac{2\delta}{\lambda_k} &\Rightarrow d^2(x_{k+1}, x) - d^2(x_k, x) \leq -\frac{2\delta}{\lambda_k} \\
&\Rightarrow d^2(x_k, x) - d^2(x_{k+1}, x) \geq \frac{2\delta}{\lambda_k} \\
&\Rightarrow \frac{1}{2\delta}(d^2(x_k, x) - d^2(x_{k+1}, x)) \geq \frac{1}{\lambda_k} \\
&\Rightarrow \frac{1}{\lambda_k} \leq \frac{1}{2\delta}(d^2(x_k, x) - d^2(x_{k+1}, x)),
\end{aligned}$$

para todo  $k$ .

Aplicando sumatoria en ambos lados de la desigualdad anterior, se tiene:

$$\sum_{k=0}^j \frac{1}{\lambda_k} \leq \sum_{k=0}^j \frac{1}{2\delta}(d^2(x_k, x) - d^2(x_{k+1}, x)),$$

así,

$$\sum_{k=0}^j \frac{1}{\lambda_k} \leq \frac{1}{2\delta}(d^2(x_0, x) - d^2(x_0, x_{j+1})) < \frac{1}{2\delta}d^2(x_0, x),$$

para todo  $j$ , lo cual contradice la ecuación,

$$\sum_{k=0}^{\infty} \frac{1}{\lambda_k} = +\infty.$$

Si el conjunto  $U^*$  es no vacío, dado  $\bar{x} \in U^*$ , entonces nuevamente por el Lema 7 · 2 se tiene que:

$$d^2(x_{k+1}, x) \leq d^2(x_k, x) - d^2(x_k, x_{k+1}) + \frac{2}{\lambda_k}(f(x) - f(x_{k+1})),$$

sustituyendo  $x$  por  $x_k$  y  $x_{k+1}$  por  $\bar{x}$ , se tiene:

$$\begin{aligned}
d^2(x_{k+1}, x) &\leq d^2(x_k, x) - d^2(x_k, x_{k+1}) + \frac{2}{\lambda_k}(f(x) - f(x_{k+1})) \\
\Rightarrow d^2(x_{k+1}, x_k) &\leq d^2(x_k, x_k) - d^2(x_k, x_{k+1}) + \frac{2}{\lambda_k}(f(x_k) - f(x_{k+1})) \\
\Rightarrow d^2(\bar{x}, x_k) &\leq d^2(x_k, x_k) - d^2(x_k, \bar{x}) + \frac{2}{\lambda_k}(f(x_k) - f(\bar{x})) \\
\Rightarrow 0 &\leq d^2(x_k, x_k) + \frac{2}{\lambda_k}(f(x_k) - f(\bar{x})) \\
\Rightarrow 0 &\leq \frac{2}{\lambda_k}(f(x_k) - f(\bar{x})) \\
\Rightarrow 0 &\leq (f(x_k) - f(\bar{x})) \\
\Rightarrow f(\bar{x}) &\leq f(x_k),
\end{aligned}$$

Así,

$$f(\bar{x}) \leq f(x_k),$$

para todo  $k$ .

Por esta desigualdad y sustituyendo  $x$  por  $\bar{x}$  en (7 · 2) obtenemos:

$$\begin{aligned} d^2(x_k, x_{k+1}) + d^2(x_{k+1}, \bar{x}) - 2d(x_k, x_{k+1})d(x_{k+1}, \bar{x})\cos\theta &\leq d^2(x_k, \bar{x}) \\ \Rightarrow d^2(x_k, x_{k+1}) + d^2(x_{k+1}, \bar{x}) - 2d(x_k, x_{k+1})d(x_{k+1}, \bar{x})\cos\theta &\leq d^2(x_k, \bar{x}) \\ \Rightarrow d^2(x_{k+1}, \bar{x}) &< d^2(x_k, \bar{x}), \end{aligned}$$

así, la sucesión  $\{x_k\}$  es Fejér convergente en  $U^*$ .

Por el Lema 7 · 1 tenemos que la sucesión  $\{x_k\}$  es acotada, y como  $\mathcal{M}$  es una variedad completa, esta tiene una subsucesión convergente.

Sea  $\{x_{k_j}\}$  una subsucesión de  $\{x_k\}$  tal que:

$$\lim_{k \rightarrow +\infty} x_{k_j} = x,$$

además  $f$  es continua y

$$\lim_{k \rightarrow +\infty} f(x_k) = f^*,$$

de esto se sigue:

$$\lim_{k \rightarrow +\infty} f(x_{k_j}) = f^*,$$

esto implica que  $x_* \in U^*$ .

Así, el punto clausura  $x_*$  de  $\{x_k\}$  pertenece a  $U^*$ .

Y nuevamente por el Lema 7 · 1 se tiene:

$$\lim_{k \rightarrow +\infty} x_k = x_*.$$

## Bibliografía.

- [1] Shiga, K. Hadamard manifolds, advanced studies in pure mathematics 3, Geometry of Geodesics and Related Topic, pp.239 – 281. 1984.
- [2] Iusem, A.N. Métodos de Punto Proximal en Optimización. 20<sup>a</sup> Colóquio Brasileiro de Matemática, IMPA. 1995.
- [3] Rockafellar, R. T, Monotone operators and the Proximal Point Algorithm. SIAM J. Control Optim. 14,877 – 898,1976
- [4] Udriste, C, Convex functions and optimization methods on Riemannian manifolds, Mathematics and its applications. Kluwer academy publishers. 1994.
- [5] Do Carmo, M.P., Riemannian manifolds, Journal of Optimization Theory. Boston, Birkhauser. 1992.
- [6] Hiriart-Hurruty, J,B. and Lemarechal, C. Convex analysis and minimization algorithms *I* and *II*. Springer-Verlag. 1993.
- [7] O.P Ferreira and P.R.Oliveira. Proximal Point algorithm on Riemannian manifolds. IME. Brazil. 2000.