UNIVERSIDAD CENTROCCIDENTAL “LISANDRO ALVARADO”
DECANATO DE CIENCIAS Y TECNOLOGÍA
LICENCIATURA EN CIENCIAS MATEMÁTICAS

Algoritmo subgradiente en Variedades Riemannianas

Autor: Br. Ismerai Daniela Morillo Cordero
Tutor: Dr. Eibar Hernández (UCLA)

TRABAJO ESPECIAL DE GRADO

Presentado ante la Ilustre
Universidad Centroccidental “Lisandro Alvarado”
como requisito final para optar al grado de
Licenciado en Ciencias Matemáticas.

BARQUISIMETO, VENEZUELA
Enero, 2012.

En honor al *Rey de Reyes* Jesús.

Agradecimientos

Al único y sabio *Dios*, porque cada día de mi vida he podido ver como me cuida, me protege, me sostiene y me mantiene, sin Él sencillamente no habria podido llegar hasta aqui, su amor ha sido mi fortaleza y esperanza en todo este camino. Gracias mi amado *Dios* tú eres merecedor de toda gloria, honra y alabanza, *No hay Nadie Como Tú.*

A mis *Padres*, por ser los mejores padres que cualquier hijo desearia tener, ustedes han sido un regalo de *Dios* a mi vida, son tan importantes para mi, siempre he podido contar con su apoyo incondicional y cuidado, lo único que pido a Dios es que me alcance la vida y pueda regresarles a ustedes aunque sea un poco de lo mucho que me dan.

A mis *Hermanos*, por todo su apoyo, y por ser tan especiales, *Dios* me bendijo al permitirme ser parte de ustedes, sus cuidados y sus juegos me han hecho muy feliz.

A mi iglesia *Pan de Vida*, por todo su apoyo y ayuda para que yo pudiera estar aqui, a mi equipo *Castillo del Rey* porque ustedes muchachos han sido inspiración a mi vida, ustedes son el sueño de todo lider.

A mi *tutor* Eibar Hernández, primero que nada por aceptarme como tesista y transmitirme los conocimientos necesarios para la elaboración de este trabajo, por toda su dedicación, y sobre todo porque más que un tutor se convirtió en un buen amigo.

A el profesor *Sergio Muñoz* por todo su apoyo en los incios de mi carrera, gracias profesor por creer en mi y ayudarme a superar todos aquellos miedos que no me dejaban avanzar.

A mis *Amigos* Naty, Yori, Laurita, Kim, Yohana, Rona, Eloimar, Jonathan, por todo su apoyo y compañia, ustedes más que amigos son mis hermanos, el saber que puedo contar siempre con ustedes llena de felicidad mi corazon.

Faltaría por mencionar una gran cantidad de personas que de una manera u otra me apoyaron durante todo este tiempo, a todos ustedes muchas gracias.

Br. Ismerai Morillo

Introducción

El método subgradiente es uno de los algoritmos clásicos de optimización no diferenciable que fue descubierto por Short, N. en los años sesenta, el cual siempre ha sido objeto de estudio, en este trabajo el próposito es estudiar y desarrollar el artículo [1] donde se analiza el algoritmo subgradiente para resolver el siguiente problema con restricciones

$$\min_{x \in M} f(x),$$

donde M es una variedad Riemanniana conexa y completa con curvatura seccional $K \geq 0$ y f una función convexa de valores reales definida en M no necesariamente diferenciable.

En el capitulo 1, se presentarán una serie de definiciones sobre funciones convexas en $\mathbb{R}^n$, curvas diferenciables, superficies, plano tangente, primera y segunda forma fundamental, derivada covariante, transporte paralelo, variedades diferenciables y Riemanniana, entre otras definiciones que serán necesario para la construcción de este algoritmo.

En el capítulo 2, se estudian algunas definiciones sobre funciones convexas en variedades riemannianas, se da la definición de subgradiente y se prueban algunos teoremas relacionados con el subgradiente en una variedad.

En el capítulo 3, se presenta el desarrollo donde se da la definición del problema y se presenta una prueba de convergencia, donde es de vital importancia la desigualdad

$$d^2(x_{k+1}, y) \leq d^2(x_k, y) + t_k^2 + 2[t_k \|s_k\|][f(y) - f(x_k)] \ \ \forall k \in \mathbb{N}.$$

Que se obtiene para variedades con curvatura seccional no-negativa.

Índice general

Índice de figuras

Capítulo 1

Preliminares de Geometría

1.1. Conceptos y teoremas básicos

Funciones convexas en $\mathbb{R}^n$:

Definición 1.1. *Un subconjunto C de $\mathbb{R}^n$ es llamado convexo si y solo si*

$$(1-\lambda)x + \lambda y$$

para cada $x, y \in C$, $\lambda \in (0,1)$. Es decir, un conjunto es convexo, cuando el segmento de recta que une cualquier par de puntos del conjunto esta totalmente contenido en éste.

Definición 1.2. *Una función $f : \mathbb{R} \to \mathbb{R}^n$ no identicamente igual a $+\infty$ se dice que es convexa cuando para todo para $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$ y para todo $\lambda \in (0,1)$ se cumple:*

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y).$$

Obsérvese que la desigualdad anterior es en $\mathbb{R} \cup +\infty$ y denotaremos a la clases de tales funciones por $Conv\mathbb{R}^n$.

Definición 1.3. *El dominio efectivo de* $f \in Conv\mathbb{R}^n$ *es el conjunto no-vacío*

$$dom f = \{x \in \mathbb{R}^n : f(x) < +\infty\}$$

Definición 1.4. *Dada* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *una funcion no identicamente igual a* $+\infty$*, el epigrafo de* f *es el conjunto no-vacío dado por:*

$$epi(f) := \{(x, r) \in \mathbb{R}^n \times \mathbb{R} : x \geq f(x)\}$$

el epigrafo es estricto si reemplazamos $\geq$ *por* $>$ *.*

Proposición 1.1. *Sea* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *una función no identicamente igual a* $+\infty$*. Las siguientes propiedades son equivalentes:*

- $f \in Conv\mathbb{R}^n$
- $epi(f)$ *es un conjunto convexo en* $\mathbb{R}^n \times \mathbb{R}$*.*
- $epi(f)$ *estricto es un conjunto convexo en* $\mathbb{R}^n \times \mathbb{R}$*.*

Proposición 1.2. *Sea* $f \in Conv\mathbb{R}$ *en todo* x_0 *en el interior de su dominio,* f *admite una derivada por la izquierda finita y una derivada por la derecha finita:*

$$D_- f(x_0) := \lim_{x \to x_0^-} \frac{f(x) - f(x_0)}{x - x_0} = \sup_{x < x_0} \frac{f(x) - f(x_0)}{x - x_0} \tag{1.1}$$

$$D_+ f(x_0) := \lim_{x \to x_0^+} \frac{f(x) - f(x_0)}{x - x_0} = \inf_{x > x_0} \frac{f(x) - f(x_0)}{x - x_0} \tag{1.2}$$

ellas satisfacen:

$$D_- f(x_0) \leq D_+ f(x_0)$$

Observación 1.1. *La derivada direccional de* f *en el punto* x_0 *en la dirección* d *es denotada por:*

$$f'(x_0, d) = \lim_{t \downarrow 0} \frac{f(x_0 + td) - f(x_0)}{t} = \inf_{t \downarrow 0} \frac{f(x_0 + td) - f(x_0)}{t}. \tag{1.3}$$

Si $f : \mathbb{R} \to \mathbb{R} \cup +\infty$*, se tiene* $d = 1$ *ó* $d = -1$*.*

Acontinuación se presentara una caracterización de las subderivadas.

Definición 1.5. *Sea $f \in Conv\mathbb{R}$. Decimos que $s \in \mathbb{R}$ es una subderivada de f en $x \in domf$ cuando*

$$D_{-}f(x_0) \leq s \leq D_{+}f(x_0).$$

Definición 1.6. *Sea $f \in Conv\mathbb{R}$ diremos que s es una subderivada f en x_0 si y sólo si*

$$f(x) \geq f(x_0 + s(x - x_0)) \;\; \forall x \in \mathbb{R}^n.$$

El subdiferencial $\partial f(x)$ es el conjunto de todas las subderivadas de f en x y se denota por:

$$\partial f(x) = \{s \in \mathbb{R}^n : \langle s, d\rangle \leq f'(x, d) \;\; \forall d \in \mathbb{R}^n\}.$$

El vector $s \in \partial f(x)$ es llamado subgradiente de f en x.

1.2. Método del Subgradiente en $\mathbb{R}^n$

Sea $f : \mathbb{R}^n \to \mathbb{R}$, una función convexa

$$x_{k+1} = x_k - \alpha_k g_k$$

donde g_k denota un subgradiente de f en x_k. Si f es diferenciable, entonces sólo es el vector gradiente ∇f en si mismo. Puede suceder que g_k no es una dirección de descenso para f en x_k. Por lo tanto, mantener una lista f_{best} que realiza un seguimiento del valor objetivo encontrado hasta el momento, es decir:

$$f_{best}^{k} = min\{f_{best}^{k-1}, f(x_k)\}$$

Reglas de tamaño de paso:

- Tamaño de paso constante $\alpha = \alpha_k$

- Longitud de paso constante, $\alpha_k = \lambda/\|g_k\|^2$ lo que da $\|x_{k+1} - x_k\|^2 = \lambda$
- $\alpha_k \geq 0$, $\sum_{k=1}^{\infty} \alpha_k^2 < \infty$, $\sum_{k=1}^{\infty} \alpha_k = \infty$
- $\alpha_k \geq 0$, $\lim_{k\to\infty} \alpha_k = 0$, $\sum_{k=1}^{\infty} \alpha_k = \infty$
- $\lambda \geq 0$,$\lim_{k\to\infty} \lambda_k = 0$,$\sum_{k=1}^{\infty} \lambda_k = \infty$

Resultado de Convergencia: el método subgradiente converge a una aproximación arbitrariamente cercana al valor mínimo, es decir:

$$\lim_{k\to\infty} f_{best}^k - f^* < \epsilon$$

1.3. Curvas

Definición 1.7. *Una curva en un conjunto* $\mathbb{X} \subset \mathbb{R}^n$ *es una funcion continua* $\alpha : I \to \mathbb{X}$*, donde* $I = (a, b)$

Definición 1.8. *Sean* $a, b \in \mathbb{X} \subset \mathbb{R}^n$ *diremos que* a *y* b *pueden ser conectado mediante una curva* α *cuando existe* $\alpha : I \to \mathbb{X}$ *curva, y existen* δ, β *con* $\delta < \beta$ *tal que* $\alpha(\delta) = a$ *y* $\alpha(\beta) = b$.

Proposición 1.3. *Sean* $\mathbb{X} \subset \mathbb{R}^n$ *y* $a, b \in \mathbb{X}$ *tales que* a *y* b *estan conectados por una curva* $\alpha : I \to \mathbb{X}$ *entonces existe una curva* $\varphi : [0, 1] \to \mathbb{X}$ *tales que* $\varphi(0) = a$ *y* $\varphi(1) = b$.

Definición 1.9. *Sea* $\alpha : I \to \mathbb{R}^n$*,una curva en* $\mathbb{R}^n$ *sea* $t_0 \in I$*, diremos que* α *es diferenciable en* t_0 *si y solo si el siguiente limite existe:*

$$\lim_{h\to 0} \frac{\alpha(t_{0+h}) - \alpha(t_0)}{h}$$

y el valor de tal limite lo denotaremos por $\alpha'(t_0)$.

Teorema 1.1 (Desigualdad de Valor medio). *Sea $\alpha : [a,b] \to \mathbb{R}^n$ una curva diferenciable en (a,b) con $\|\alpha'(t)\| \leq M$, para algún $M > 0$*

$$\|\alpha(b) - \alpha(a)\| \leq M(b-a).$$

Demostración

$$\begin{aligned}\|\alpha(b) - \alpha(a)\| &= \sqrt{\langle \alpha(b) - \alpha(a), \alpha(b) - \alpha(a)\rangle} \\ &= \sqrt{\langle \alpha(b), \alpha(b) - \alpha(a)\rangle - \langle \alpha(b), \alpha(b) - \alpha(a)\rangle}\end{aligned}$$

Definimos: $\varphi : [a,b] \to \mathbb{R}$ dada por:

$$\begin{aligned}\varphi(t) &= \langle \alpha(t), \alpha(b) - \alpha(a)\rangle \\ &= \sqrt{\varphi(b) - \varphi(a)}\end{aligned}$$

Así existe $c \in (a,b)$ tal que $\varphi(b) - \varphi(a) = \varphi'(yc)(b-a)$ Luego:

$$\begin{aligned}\|\alpha(b) - \alpha(a)\|^2 &= \varphi'(c)(b-a) \\ &= \langle \alpha'(c), \alpha(b) - \alpha(a)\rangle (b-a) \\ &\leq \|\alpha'(c)\| \|\alpha(b) - \alpha(a)\| (b-a)\end{aligned}$$

Asi:

$$\|\alpha(b) - \alpha(a)\| \leq M(b-a).$$

Proposición 1.4. *Si $\alpha : [a,b] \to \mathbb{R}^n$ es curva con $\alpha'(t) = 0$ para todo $t \in (a,b)$ entonces, α es constante.*

Demostración Sea $t \in [a,b]$ por Desigualdad de valor medio tenemos:

$$0 \leq \|\alpha(t) - \alpha(a)\| \leq 0(b-a) \Longrightarrow \alpha(b) = \alpha(a) \;\; \forall t \in [a,b].$$

Definición 1.10. *Una curva parametrizada definida en $[0,\ell]$, $\alpha : [0,\ell] \to S$ es la restricción a $[0,\ell]$ de una función diferenciable $\alpha : (0-\epsilon, \ell+\epsilon) \to S$, $\epsilon > 0$. Si $\alpha(0) = p$ y $\alpha(\ell) = q$ decimos que α conecta p con q y la distancia de p a q, se define como $d(p,q) = inf\{\ell(\alpha) : \alpha \text{ conecta a } p \text{ con } q.\}$*

Definición 1.11. *Sea $\alpha : I \to \mathbb{R}^n$ una curva diferenciable ,diremos que α es una curva regular si y sólo si $\alpha'(t) \neq 0_n$, para cualquier $t \in I$.*

Observación 1.2. *Una curva α es rectificable si para $t \in [a, b]$ finito, existe un recorrido $f(t)$ finito a lo largo de $f(a)$ hasta $f(b)$.*

Definición 1.12. *Dado $t \in I$, la longitud de arco de una curva regular desde t_0 hasta t, se define por:*

$$\alpha(t_o, t) = S(t) = \int_{t_o}^{t} \|\alpha'(t)\|.dt.$$

Ejemplo: Sea $\alpha(t) = (\cos t, \sin t)$, $t \in [0, 2\pi)$, ¿Cúal es la longitud de arco?
Solución:

$$\|\alpha'(t)\| = \sqrt{(-\sin t)^2 + (\cos t)^2} = \sqrt{\sin^2 t + \cos^2 t} = 1.$$

Luego,

$$S_\alpha(t) = \int_0^{2\pi} \|\alpha'(t)\|.dt = \int_0^{2\pi} 1.dt = 2\pi.$$

Observemos que la curva β dada por $\beta(t) = (\cos 2t, \sin 2t)$, $t \in [0, 2\pi)$, tiene la misma traza de la curva α, pero $\beta'(t) = (-2\sin 2t, 2\cos 2t)$, de esta manera $\|\beta'\| = \sqrt{4\sin^2 t + 4\cos^2 t} = 2$; así: $S_\beta(t) = \int_0^{2\pi} 2.dt = 4\pi = 2S_\alpha(t)$; es decir la longitud de arco de β es el doble de la longitud de arco de α a pesar de tener el mismo gráfico.

Observación 1.3. *Sea $\alpha : I \to \mathbb{R}^n$, $n = 2, 3$ una curva diferenciable parametrizada regular. Si $\|\alpha'(t)\| = 1$, entonces*

$$S(t) = \int_{t_o}^{t} 1.dt = t - t_o.$$

Cuando $\|\alpha'(t)\| = 1$, se dice que la curva α es parametrizada por longitud de arco.

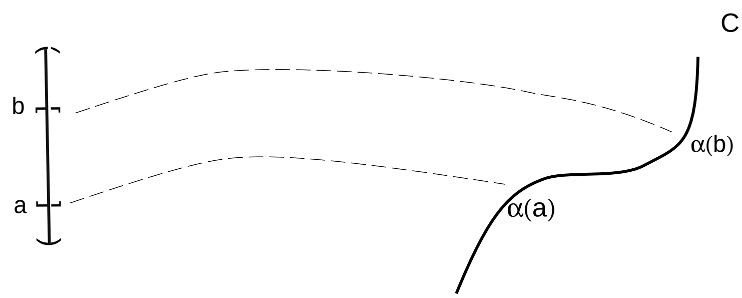

Figura 1.1: Longitud de arco

Definición 1.13. ***(Definición de diferenciabilidad uniforme)*** *Una curva $\alpha : I \to \mathbb{R}^n$ se dice uniformemente diferenciable cuando para cualquier $t \in I$, existe un vector $f(t) \in \mathbb{R}^n$ con la siguiente propiedad: Dado $\epsilon > 0$, podemos encontrar $\delta > 0$ tal que $0 < |h| < \delta$, $t + h \in I$ implica que $\|\alpha(t+h) - \alpha(t) - \alpha'(t).h\| < \epsilon.|h|$, para cualquier $t \in I$.*

Teorema 1.2. ***(Teorema fundamental del cálculo para curvas)*** *Sea $f : [a,b] \to \mathbb{R}^n$ una curva de clase C^1, entonces se cumple que:*

$$\int_a^b f'(t).dt = f(a) - f(b).$$

Definición 1.14. *Sea $\alpha : I \to \mathbb{R}^3$ una curva parametrizada por longitud de arco $s \in I$. Se le llama curvatura de α en s al número real $\|\alpha''(s)\| = K(s)$.*

Observación 1.4.

- *Si α es una linea recta, $\alpha(s) = us + v$, donde u, v son vectores constantes (además $\|u\| = 1$), entonces $K = \|\alpha''(s)\| = 0$.*
- *Si $K = \|\alpha''(s)\| = 0$, entonces por integración $\alpha(s) = us + v$, y la curva es una linea recta.*

Por tanto concluimos que $\alpha : I \to \mathbb{R}^3$ es una linea recta si y sólo si $K(s) = 0$

Observación 1.5. *En los puntos donde $K(s) \neq 0$, esto es $\alpha''(s) \neq 0$, se define $n(s) = \dfrac{\alpha''(s)}{\|\alpha''(s)\|}$, el cual es un vector unitario y además es paralelo a $\alpha''(s)$ y* *con el mismo sentido. Como α es parametrizada por longitud de arco s, se tiene que para cualquier $s \in I$, $\|\alpha'(s)\| = 1$.*

$$\begin{aligned}
\|\alpha'(s)\| = 1 &\Rightarrow \|\alpha'(s)\|^2 = 1 \\
&\Rightarrow \langle \alpha'(s), \alpha'(s) \rangle = 1 \\
&\Rightarrow \langle \alpha''(s), \alpha'(s) \rangle + \langle \alpha'(s), \alpha''(s) \rangle = 0 \\
&\Rightarrow 2\langle \alpha''(s), \alpha'(s) \rangle = 0 \\
&\Rightarrow \langle \alpha''(s), \alpha'(s) \rangle = 0
\end{aligned}$$

Así concluimos que $\alpha''(s)$ es perpendicular a $\alpha'(s)$ y como $\alpha''(s) = K(s)n(s)$, obtenemos tambien que $n(s)$ es perpendicular a $\alpha'(s)$, para cualquier $s \in I$. Como $n(s)$, $\alpha'(s)$ son linealmente independiente (por se ortogonales), ellos determinan un plano
$\Pi_s = \{x \in \mathbb{R}^3 : x = \alpha(s) + \alpha.\alpha'(s) + \beta.\alpha''(s),\ \alpha,\ \beta \in \mathbb{R}\}$, el cual se llamará plano osculador en s

1.4. Superficies en $\mathbb{R}^n$

Definición 1.15. *Sea $F : U \subset \mathbb{R}^n \to \mathbb{R}^m$ diferenciable, en cada $p \in U$ se le asocia una aplicación lineal $dF_p : \mathbb{R}^n \to \mathbb{R}^m$ que llamaremos la diferencial de F en p que se define así:*
Sea $w \in \mathbb{R}^n$ y sea $\alpha : (-\epsilon, \epsilon) \to U$ una curva diferenciable tal que: $\alpha(0) = p$; $\alpha'(0) = w$. Si consideramos $\beta = F \circ \alpha : (-\epsilon, \epsilon) \to \mathbb{R}^m$ se tiene así que β es diferenciable y se define: $dF_p(w) = \beta'(0) = (F \circ \alpha)'(0) = F'(\alpha(0)).\alpha'(0)$

Teorema 1.3. ***(Teorema de la Función Inversa)*** *Si $F : U \subset \mathbb{R}^n \to \mathbb{R}^n$ es una función diferenciable (C^1, C^∞) en $p \in U$, U abierto de $\mathbb{R}^n$ y $F'(p)$ es no singular para algún punto $p \in U$, entonces existen $V_p, W_{F(p)}$ vecindades de*

p y $F(p)$ respectivamente tales que: $F : V_p \to W_p$ es biyectiva y su inversa $G : W_p \to V_p$ es diferenciable (C^1, C^∞).

Teorema 1.4. *Sea $U \subset \mathbb{R}^{k+n}$ abierto, $F : U \subset \mathbb{R}^{k+n} \to \mathbb{R}^n$ de clase C^r en U. Escribir*
$F(x,y) = F(x_1, ..., x_k)$, para $y \in \mathbb{R}^n$, $x \in \mathbb{R}^k$, supongamos que $(a,b) \in U$ es tal que $F(a,b) = 0$
y $\dfrac{\partial F}{\partial y}(a,b)$ es no singular, entonces existe B vecindad de $a \in \mathbb{R}^k$ y una única función
$G : B \subset \mathbb{R}^k \to \mathbb{R}^n$ de clase C^r tal que $G(a) = b$ y $F(x, G(x)) = 0$, para cualquier $x \in B$. Además G es de clase C^r en B.

Demostración: Definir $f : U \subset \mathbb{R}^{k+n} \to \mathbb{R}^{k+n}$ por $f(x,y) = (x, F(x,y))$. Ver que f cumple las hipótesis del teorema de la función inversa.
f es de clase C^r en U, púes F es de clase C^r en U
además $f'(a,b) = \begin{bmatrix} I_{k\times k} & 0_{k\times k} \\ \dfrac{\partial F}{\partial x}(a,b) & \dfrac{\partial F}{\partial y}(a,b) \end{bmatrix}$ y como $det\dfrac{\partial F}{\partial y}(a,b) \neq 0$ por hipótesis, se tiene que $det f'(a,b) \neq 0$, luego usando el teorema de la función inversa se tiene que existen $C((a,b),r)$ y $W_{f(a,b)}$ tal que: $g = f^{-1} : C((a,b),r) \to W_{f(a,b)}$ es invertible. Observemos que $f(a,b) = (a, F(a,b)) = (a,0)$, (pues $F(a,b) = 0$ para todo $a, b \in U$), por otro lado, como $W_{(a,0)}$ es un abierto que contiene a $(a,0)$, entonces existe $C(a,0) \subset W_{(a,0)}$; llamemos $B_a = C(a,0) \cap \mathbb{R}^k \subset \mathbb{R}^k$.

Se sabe que $f(x,y) = (x, F(x,y))$, ¿cómo será $g = f^{-1}$?
$(x,y) = g(f(x,y)) = g(x, F(x,y))$, de esta manera $g(x,y) = (x, h(x,y))$ para alguna función $h(x,y)$, donde $h : \mathbb{R}^{k+n} \to \mathbb{R}^n$ y $g(x,0) = (x, h(x,0))$. Definir $G : B \subset: \mathbb{R}^k \to: \mathbb{R}^n$ por $G(x) = h(x,0)$ la cual es claramente C^r, observemos que:

$$\begin{aligned} f(a,b) = (a,0) &\Rightarrow (a,b) = g(a,0) \\ &\Rightarrow (a,b) = (a, h(a,0)) \\ &\Rightarrow a = a \ \wedge \ h(a,o) = b \end{aligned}$$

Además;

$$\begin{aligned} f(g(x,0)) = f(x,h(x,0)) = f(x,G(x)) &\Rightarrow (x,0) = (x, F(x, G(x,0))) \\ &\Rightarrow x = x \ \wedge \ F(x,G(x)) = 0 \end{aligned}$$

Por lo tanto $G(a) = h(a,0) = b$ y $F(x,G(x)) = 0$, para cualquier $x \in B$. Ahora se probará la unicidad de G. Supongamos que $G_1 : B \subset \mathbb{R}^k \to \mathbb{R}^n$, $G_2 : B \subset \mathbb{R}^k \to \mathbb{R}^n$, donde $G_1(x) \neq G_2(x)$ para cualquier $x \in B$,
$G_1(a) = b = G_2(a)$ y $F(x, G_1(x)) = 0 = F(x, G_2(x))$ para cualquier $x \in B$.

$$\begin{aligned} g(f(x,G_1(x))) = g(x, F(x,G_1(x))) &\Rightarrow (x, G_1(x)) = g(x, F(x, G_1(x))) \\ &\Rightarrow (x, G_1(x)) = g(x, F(x, G_1(x))) \\ &\Rightarrow (x, G_1(x)) = g(x, 0) \end{aligned}$$

También se tiene que:

$$\begin{aligned} g(f(x,G_2(x))) = g(x, F(x,G_2(x))) &\Rightarrow (x, G_2(x)) = g(x, F(x, G_2(x))) \\ &\Rightarrow (x, G_2(x)) = g(x, F(x, G_2(x))) \\ &\Rightarrow (x, G_2(x)) = g(x, 0) \end{aligned}$$

Por lo tanto; $(x, G_1(x)) = g(x,0) = (x, G_2(x))$, lo que implica que
$G_1(x) = G_2(x)$.

Definición 1.16. *Sea $F : U \subset \mathbb{R}^n \to \mathbb{R}^m$ $m \leq n$ una aplicación diferenciable, diremos que F es una inmersión si y solo si para todo $p \in U$ $dF_p : \mathbb{R}^n \to \mathbb{R}^m$ es inyectiva. Si $m = n$,entonces*
dF_p resulta isomorfismo y por tanto F es difeomorfismo local.

Ejemplo: $\alpha(t) = (t^3 - t, 2t)$, $\alpha'(t) = (3t^2 - 1, 2) \neq 0$ $\forall t$.
$\therefore$ α es curva diferenciable e inmersión.
sin embargo, α no es inyectiva, $\alpha(1) = (0,1) = \alpha(-1)$.

Definición 1.17. *Una parametrización de clase C^k y dimesión m de un conjunto $V \subset \mathbb{R}^n$ es una inmersión $\varphi : V_0 \to V$ de clase C^k que al mismo tiempo es homeomorfismo del abierto $V_0 \subset \mathbb{R}^m$ en V.*

Definición 1.18. *Un conjunto M es superficie de dimensión m y clase C^k, cuando para todo $p \in M$ existe $U_p \subset \mathbb{R}^n$ tal que $V_p = U \cap M$ es la imagen de alguna parametrización $\varphi_p : V_0 \subset \mathbb{R}^m \to V_p$ de dimensión m y clase C^k. El conjunto V_p es abierto en M y es llamado vecindad parametrizada de p y se dirá que $dimM = m$.*

Definición 1.19. *Si $M \subset \mathbb{R}^n$, M superficie de dimensión m con $m \leq n$, diremos que M tiene codimensión $n - m \geq 0$.*

Definición 1.20. *En general, llamaremos proyección $\pi : \mathbb{R}^n \to \mathbb{R}^m$ a cualquier aplicación $\pi(x_1, ..., x_n) = \pi(x_{i_1}, ..., x_{i_m})$ definida a partir de la escogencia de indices $i_1 < ... < i_m$; $i_j\{1, ..., n\}$.*

Teorema 1.5. *Sea $M \subset \mathbb{R}^n$ una superficie de dimensión m y clase C^k, en $\mathbb{R}^n$, sea $\varphi : V_0 \to V$ una parametrización en M, para cada $p \in V$ tal que $p \in \varphi(x_0)$, existe $\pi : \mathbb{R}^n \to \mathbb{R}^m$ existe $Z_0 \subset \mathbb{R}^m$ abierto tal que $x_0 \in Z_0 \subset V_0 \subset \mathbb{R}^m$ y existe $W_0 \subset \mathbb{R}^m$ tal que: $\pi \circ \varphi : Z_0 \to W_0$ es difeomorfismo.*

Demostración:

Sea $\varphi : V_0 \to V$ parametrización entonces $d\varphi_{x_0} : \mathbb{R}^m \to \mathbb{R}^n$ así:
$[d\varphi_{x_0}]_n \times m$ tiene m filas L.I., llamemos $i_1 < ... < i_m$ tales filas. Se define $\pi : \mathbb{R}^n \to \mathbb{R}^m$ por $\pi(x_1, ..., x_n) = \pi(x_{i_1}, ..., x_{i_m})$ entonces:
$\pi \circ \varphi : V_0 \subset \mathbb{R}^m \to \mathbb{R}^n$ es C^k. así:
$[d(\pi \circ \varphi)_{x_0}] = [d\pi_\varphi(x_0)][d\varphi_{x_0}] = \pi[d\varphi_{x_0}] = \left[\frac{\partial \varphi_i j}{\partial X_j}\right] = J$ es no- singular, luego por el teormea de la funcion inversa $\pi \circ \varphi : Z_0 \to W_0$ es difeomorfismo C^k.

Proposición 1.5. *Toda superficie C^k es localmente el gráfico de una aplicación C^k.*

Definición 1.21. *Sean $\varphi : V_0 \to W_0$ y $\psi : W_0 \to W$ dos parametrizaciones tales que: $V \cap W \neq \emptyset$, para $p \in V \cap W$ existen $x \in V_0$ y $y \in W_0$ donde: $\varphi(x) = p = \psi(y)$ entonces $(\varphi^{-}1 \circ \varphi)(x) = y$ a esta aplicación:*
$\psi^{-1} \circ \varphi : \varphi^{-1}(V \cap W) \to \varphi^{-1}(V \cap W)$ se le llama mudanza o cambio de parametrización

Proposición 1.6. *Tal aplicación $\psi^{-1} \circ \varphi$ es difeomorfismo C^k.*

Definición 1.22. *Al conjunto formado por todos los vectores tangentes a M en p, lo llamaremos plano tangente y lo denotaremos como sigue;*
$T_pM = \{\alpha'(0) = \boldsymbol{v} : \alpha : (-\epsilon, \epsilon) \to M$, *es una curva diferenciable con* $\alpha(0) = p\}$

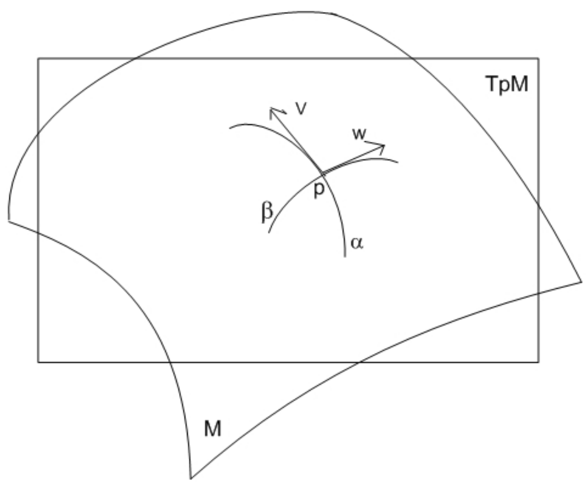

Figura 1.2: Plano Tangente

Definición 1.23. *Sea $f : U \subset \mathbb{R}^{m+n} \to \mathbb{R}^n$ diferenciable en U, diremos que c es un valor regular de f cuando $\forall x \in U : f(x) = c$ (es decir, $\forall x \in f^{-1}(c)$, esto es, $f'(x) : \mathbb{R}^{m+n} \to \mathbb{R}^n$ es sobreyectiva, $[f']$ tiene rango n)*

Teorema 1.6. *Sea $c \in \mathbb{R}^n$ un valor regular de $f : U \subset \mathbb{R}^{m+n} \to \mathbb{R}^n$ de clase C^k entonces $M = f^{-1}(c) = \{x \in U / f(x) = c\}$ es una superficie C^k con $dim M = m$; $M \subset \mathbb{R}^{m+n}$ y $T_pM = Ker f'(p) \forall p \in M$*

Demostración: Por teorema de la función inversa se tiene que $M = f^{-1}(c)$ es localmente el gráfico de una función g de clase C^k, entonces M es superficie parametrizada por $\varphi(x) = (x, g(x)),\ \ x \in \mathbb{R}^m$ asi $dimM = m$. Ahora probemos que $\forall p \in M,\ T_pM = Kerf'(p)$ Sea $p \in M$ y $v \in T_pM$, asi $v = \lambda'(0)$ para alguna $\lambda : (-\varepsilon, \varepsilon) \to M$ con $\lambda(p) = p, \lambda$ diferenciable. Así,

$$f'(p)(v) = df_p(v) = (f \circ \lambda)'(0) = 0.$$

$$\begin{aligned} \lambda(t) \in M \Rightarrow f(\lambda(t)) = c \Rightarrow (f \circ \lambda)(t) = c \\ \Rightarrow (f \circ \lambda)'(t) = 0, \forall t \in (-\varepsilon, \varepsilon) \end{aligned}$$

Luego,

$$f'(p)(v) = 0 \Rightarrow v \in Kerf'(p)$$

Por tanto $T_pM \subset Kerf'(p)$. Por otro lado, tenemos que:

$$dimKerf'(p) + dimImf'(p) = m + n \Rightarrow dimKerf'(p) = m.$$

pero $dimT_pM = m$ y como $T_pM \subset Kerf'(p)$ entonces podemos concluir que:

$$T_pM = Kerf'(p).$$

Observación 1.6.

- *Sea $M \subset \mathbb{R}^3$ una superficie en $\mathbb{R}^3$, donde tenemos un producto interno $\langle , \rangle$ definido como:*

$$\langle x, y \rangle = x_1y_1 + x_2y_2 + x_3y_3$$

- *Consideremos T_pM, si tenemos $w_1, w_2 \in T_pM$, definamos $\langle w_1, w_2 \rangle_p := \langle w_1, w_2 \rangle$*

- *Definamos la forma cuadratica $I_p : T_pM \to \mathbb{R}$ dada por:*

$$I_p(w) = \langle w, w \rangle_p = \langle w, w \rangle$$

A I_p le llamaremos 1^{ra} Forma fundamental.

- *Sea* $(v, +, .)$ *un espacio vectorial, una forma cuadratica sobre* v *es* $h : v \to \mathbb{R}$ *dada por:*

$$h(v) = \sum_{i,j=1}^{m} a_{i,j} v_i v_j$$

con $v = (v_1, v_2, ..., v_m)$ *y* $h_{i,j}$ *son escalares si* $H = (a_{i,j})$

1.5. La primera forma fundamental

Sea $M \subset \mathbb{R}^3$ una superficie, $p \in M$ y consideremos $T_pM \subset \mathbb{R}^3$. Como en $\mathbb{R}^3$ tenemos el producto interno $\langle , \rangle$ usual, entonces dado $v_p, w_p \in T_pM \subset \mathbb{R}^3$ podemos calcular

$$\langle v_p, w_p \rangle = \sum_{i=1}^{3} v_i w_i.$$

Así, al variar p en M podemos definir $\langle v_p, w_p \rangle_p = \langle v, w \rangle$.

Ahora bien, con $\langle , \rangle_p$ que es una forma bilineal, se define una forma cuadrática $I_p : T_pM \to \mathbb{R}$ dada por:

$$I_p(w) = \langle w_p, w_p \rangle_p = \|w\|^2 \geq 0.$$

A esta le llamaremos primera forma fundamental de la superficie $M \subset \mathbb{R}^3$ en el punto $p \in M$.

Observemos que al cambiar p, cambia I_p y expresemos la primera forma fundamental en términos de la base $\{\varphi_u, \varphi_v\}$ asociada a la parametrización $\varphi(u, v)$ en p. Como $w \in T_pM$, entonces $w = \alpha'(0)$ para alguna curva $\alpha : (-\epsilon, \epsilon) \to M$ con $\alpha(0) = p$; por otro lado, tenemos también que $\alpha(t) = \varphi(u_t, v_t)$. Así,

$p = \alpha(0) = \varphi(u(0), v(0)) = \varphi(u_0, v_0)$

$$\begin{aligned} I_p(w) &= I_p(\alpha'(0)) \\ &= \langle \alpha'(0), \alpha'(0)\rangle_p \\ &= \langle u'(0)\varphi_u(u_0,v_0) + v'(0)\varphi_v(u_0,v_0), u'(0)\varphi_u(u_0,v_0) + v'(0)\varphi_v(u_0,v_0)\rangle_p \\ &= E(u'(0))^2 + 2Fu'(0)v'(0) + G(v'(0))^2 \\ &= g_{11}(u'(0))^2 + 2g_{12}u'(0)v'(0) + g_{22}(v'(0))^2 \end{aligned}$$

Donde:

$$g_{11} = E = \langle \varphi_u(u_0,v_0), \varphi_u(u_0,v_0)\rangle_p,$$

$$g_{12} = g_{21} = F = \langle \varphi_u(u_0,v_0), \varphi_v(u_0,v_0)\rangle_p,$$

$$g_{22} = G = \langle \varphi_v(u_0,v_0), \varphi_v(u_0,v_0)\rangle_p$$

.

Más adelante la métrica Riemanniana será dada así: $\mathbf{G} = \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix}$

Observemos que si $p = \varphi(u, v)$ varia en una vecindad coordenada, entonces $E(u, v)$, $F(u, v)$, $G(u, v)$ serán funciones diferenciables en dicha vecindad. La importancia de la primera forma fundamental proviene del hecho de que al conocer I, podemos tratar cuestiones métricas sobre una superficie M, sin hacer referencia al espacio ambiente $\mathbb{R}^3$. Como ya sabemos la longitud de arco parametrizada por la curva $\alpha : \mathrm{I} \to M$ es dada por:

$$S(t) = \int_0^t \|\alpha'(t)\| dt = \int_0^t \sqrt{\langle \alpha'(t), \alpha'(t)\rangle} dt = \int_0^t \sqrt{I_p(\alpha'(t))} dt.$$

Pero si φ es una parametrización local de M, entonces $\alpha(t) = \varphi(u(t), v(t))$ e $I_P(\alpha'(t)) = E(u'(t))^2 + 2Fu'(t)v'(t) + G(v'(t))^2$; de esta manera obtenemos que:

$$S(t) = \int_0^t \sqrt{I_p(\alpha'(t))}dt = \int_0^t \sqrt{E(u'(t))^2 + 2Fu'(t)v'(t) + G(v'(t))^2}dt.$$

Luego, derivando y elevando al cuadrado tenemos que:

$$\left(\frac{dS(t)}{dt}\right)^2 = E\left(\frac{du(t)}{dt}\right)^2 + 2F\left(\frac{du(t)}{dt}\right)\left(\frac{dv(t)}{dt}\right) + G\left(\frac{dv(t)}{dt}\right)^2$$

y así: $dS^2 = E(du)^2 + 2Fdu.dv + G(dv)^2$. Esta es la expresión del teorema de pitágoras para un triángulo sobre la superficie M.

Definición 1.24. *Si $\varphi : U \subset \mathbb{R}^2 \to \mathbb{R}^3$ es una parametrización, se define $E = \langle \varphi_u, \varphi_u \rangle$, $F = \langle \varphi_u, \varphi_v \rangle$ y $G = \langle \varphi_v, \varphi_v \rangle$ y se le llama métrica Riemanniana a la primera forma fundamental $I_p : T_p\varphi(U) \to \mathbb{R}$ dada por:*

$$I_p = \langle w_p, w_p \rangle_p.$$

Luego, si llamamos $u = x_1$, $v = x_2$ podemos escribir que:

$$dS^2 = E(dx_1)^2 + 2Fdx_1dx_2 + G(dx_2)^2.$$

Más aún, si llamamos $g_{11} = E, \quad g_{12} = g_{21} = F \quad$ y $\quad g_{22} = G$ entonces:

$$dS^2 = g_{11}(du)^2 + 2g_{12}du.dv + g_{22}(dv)^2 = \sum_{i,j=1}^{2} g_{ij}dx_idx_j.$$

Así, podemos representar estos términos g_{ij} en una matriz y llamar

$$\mathbf{G} = (g_{ij}) = \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix} = \begin{pmatrix} E & F \\ F & G \end{pmatrix}.$$

Observemos que al variar el punto $p = \varphi(u, v)$ en la superficie, las entradas de la matriz g_{ij} son funciones que dependen de p; así, para dar una métrica Riemanniana sobre una superficie basta dar una matriz $\mathbf{G} = (g_{ij})$; donde:

$$g_{ij}(p) = \langle \frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j} \rangle_p = \langle \varphi_u, \varphi_v \rangle_p.$$

Si la superficie tiene dimensión 2, la matriz $\mathbf{G} = (g_{ij})$ será de orden 2×2; si la superficie tiene dimensión m, la matriz $\mathbf{G} = (g_{ij})$ será de orden $m \times m$.

Definición 1.25. *Sea $M \subset \mathbb{R}^3$ una superficie y $\varphi : U \subset \mathbb{R}^2 \to M$ una parametrización, entonces $g_{ij} : U \to \mathbb{R}$ son funciones a valores reales diferenciables en U para cualquier $i, j = 1, 2$; púes es claro que*
$g_{11} = E = \langle \varphi_u, \varphi_u \rangle = \|\varphi_u\|^2 > 0$, $g_{11} = G = \langle \varphi_v, \varphi_v \rangle = \|\varphi_u\|^2 > 0$ y para el caso donde $i \neq j$ será
$g_{12} = g_{21} = F = \langle \varphi_u, \varphi_v \rangle = \langle \varphi_v, \varphi_u \rangle = \|\varphi_u\| \|\varphi_v\|. \cos(\theta) = \sqrt{E.F}. \cos(\theta)$.

Observemos así, que a través $F = g_{ij}$ $(i \neq j)$ podemos encontrar el ángulo que forma φ_u con φ_v, por otro lado si $v = v_1\varphi_u + v_2\varphi_v$ y $w = w_1\varphi_u + w_2\varphi_v$, entonces:

$$\begin{aligned} \langle v, w \rangle &= \langle v_1\varphi_u + v_2\varphi_v, w_1\varphi_u + w_2\varphi_v \rangle \\ &= v_1w_1E + (v_1w_2 + v_2w_1)F + v_2w_2G. \end{aligned}$$

Es decir g_{ij} determina completamente el producto escalar de dos vectores. Consideremos ahora el producto vectorial de los vectores básicos.

$$\varphi_u \times \varphi_v = det \begin{pmatrix} i & j & k \\ \\ \dfrac{\partial x_1}{\partial u} & \dfrac{\partial x_2}{\partial u} & \dfrac{\partial x_3}{\partial u} \\ \\ \dfrac{\partial x_1}{\partial v} & \dfrac{\partial x_2}{\partial v} & \dfrac{\partial x_3}{\partial v} \end{pmatrix}.$$

Ahora bien, $\|\varphi_u \times \varphi_v\|^2 = \|\varphi_u\|^2\|\varphi_v\|^2 \sin^2(\theta)$, lo que implica

$$\begin{aligned}\|\varphi_u \times \varphi_v\|^2 + \langle \varphi_u, \varphi_v\rangle^2 &= \|\varphi_u\|^2\|\varphi_v\|^2 \sin^2(\theta) + \|\varphi_u\|^2\|\varphi_v\|^2 cos^2(\theta)\\ &= \|\varphi_u\|^2\|\varphi_v\|^2\end{aligned}$$

de acá obtenemos que $\|\varphi_u \times \varphi_v\| = \sqrt{EG - F^2}$.

Definición 1.26. *Un atlas de una superficie M, es un conjunto*

$A = \{\varphi : V_0 \to V/\ \varphi$ *es parametrización de V y la unión de tales V cubren M .*$\}$

Definición 1.27. *Dos parametrizaciones $\varphi : V_0 \to V$, $\psi : W_0 \to W$ se dicen compatibles cuando $V \cap W \neq 0$ y $\psi^{-1} \circ \varphi : \varphi^{-1}(V \cap W) \to \psi^{-1}(V \cap W)$ tiene $det J(\psi^{-1} \circ \varphi)(x) > 0$, para cualquier $x \in \varphi^{-1}(V \cap W)$.*

Definición 1.28. *Un atlas se dice coherente cuando todo par de parametrizaciones son compatibles.*

Definición 1.29. *Una superficie se dice orientable cuando admite un atlas coherente.*

Definición 1.30. *Dado una parametrización $\varphi : U \subset \mathbb{R}^2 \to M$ sobre una superficie regular M en el punto $p \in M$. Nosotros podemos escoger un vector normal unitario en cada punto de $\varphi(U)$, dado por la siguiente regla.*

$$N(q) = \frac{\varphi_u \times \varphi_v}{\|\varphi_u \times \varphi_v\|}(q), \ \ q \in \varphi(U).$$

Así, se tiene una aplicación $N : \varphi(U) \subset\!\to \mathbb{R}^3$ diferenciable, la cual asocia a cada $q \in \varphi(U)$ un vector normal unitario $N(q)$.

Definición 1.31. *Sea $M \subset \mathbb{R}^3$ una superficie con orientación N; la aplicación $N : M \to \mathbb{R}^3$ tiene sus valores en la esfera unitaria*

$$\boldsymbol{S}^2 = \{(x, y, z) \in \mathbb{R}^3 / x^2 + y^2 + z^2 = 1\}.$$

$N : M \to S^2 \subset \mathbb{R}^3$ se le llama función de Gauss de M, la cual es diferenciable y $dN_p : T_pM \to T_pM$ es lineal. Esta opera de la siguiente forma: Dado $\alpha(t) = \varphi\big(u(t), v(t)\big)$ una curva parametrizada en M con $\alpha(0) = p$, y consideremos $N \circ \alpha(t) = N(\alpha(t))$ en S^2; luego:

$$\begin{aligned} dN_p(\alpha'(0)) &= dN_p\Big(u'(0)\varphi_u + v'(0)\varphi_v\Big) \\ &= u'(0)dN_p(\varphi_u) + v'(0)dN_p(\varphi_v) \\ &= u'(0)N_u + v'(0)N_v \end{aligned}$$

Por otro lado tenemos que $\langle N, \varphi_u\rangle = 0$ y $\langle N, \varphi_v\rangle = 0$; así al derivar la primera respecto v y la segunda respecto u, obtenemos lo siguiente
$\langle N_v, \varphi_u\rangle = -\langle N, \varphi_{uv}\rangle$ y $\langle \varphi_v, N_u\rangle = -\langle \varphi_{vu}, N\rangle$, pero φ es C^2, entonces: $\varphi_{uv} = \varphi_{vu}$.

Por lo tanto, $\langle N_v, \varphi_u\rangle = \langle \varphi_v, N_v\rangle$ o lo que es equivalente,
$\langle dN_p(\varphi_v), \varphi_u\rangle = \langle \varphi_v, dN_p(\varphi_u)\rangle$.

De acá concluimos que $dN_p : T_pM \to T_{N(p)}S^2$ es autoadjunta, y así podemos asociarle una forma cuadrática $II_p : T_pM \to \mathbb{R}$ dada por:
$II_p(v) = \langle dN_p(\varphi_v), v\rangle$, la cual es llamada segunda forma fundamental.

Como T_pM y $T_{N(p)}M$ son paralelos podemos escribir $dN_p : T_pM \to T_pM$ en lugar de $dN_p : T_pM \to T_{N(p)}S^2$; así como $\{\varphi_u, \varphi_v\}$ es base de T_pM y $N_u, N_v \in T_pM$ tenemos que:

$N_u = a_{11}\varphi_u + a_{21}\varphi_v$.

$N_v = a_{12}\varphi_u + a_{22}\varphi_v$.

Y de esta manera:

$$\begin{aligned} dN_p(\alpha'(0)) &= u'(0)N_u + v'(0)N_v \\ &= u'(0)[a_{11}\varphi_u + a_{21}\varphi_v] + v'(0)[a_{12}\varphi_u + a_{22}\varphi_v] \\ &= [u'(0)a_{11} + v'(0)a_{12}]\varphi_u + [u'(0)a_{21} + v'(0)a_{22}]\varphi_v. \end{aligned}$$

Por lo tanto,

$$(dN_p).\begin{pmatrix} u'(0) \\ v'(0) \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.\begin{pmatrix} u'(0) \\ v'(0) \end{pmatrix}.$$

Luego la segunda forma fundamental $II_p(\alpha'(0))$ en la base $\{\varphi_u, \varphi_v\}$ será:

$$\begin{aligned} II_p(\alpha'(0)) &= -\langle dN_p(\alpha'(0)), \alpha'(0)\rangle \\ &= -\langle u'(0)N_u + v'(0)N_v, u'(0)\varphi_u + v'(0)\varphi_v\rangle \\ &= -(u'(0))^2\langle N_u, \varphi_u\rangle - u'(0)v'(0)\langle N_u, \varphi_v\rangle \\ &- u'(0)v'(0)\langle N_v, \varphi_u\rangle - (v'(0))^2\langle N_v, \varphi_v\rangle \\ &= e(u'(0))^2 + 2fu'(0)v'(0) + g(v'(0))^2 \end{aligned}$$

Donde $-\langle N_u, \varphi_u\rangle = e$, $\langle N_u, \varphi_v\rangle = \langle N_v, \varphi_u\rangle = f$ y $\langle N_v, \varphi_v\rangle = g$.

1.6. Transporte Paralelo y geodésica

Definición 1.32. *Sea $\alpha : I \to M$, una curva parametrizada en M. Un campo vectorial $\mathbb{W}$ a lo largo de α es una correspondencia que asigna a cada punto $t \in I$ un vector $\mathbb{W}(t) \in T_{\alpha(t)}M$. Diremos que $\mathbb{W}$ es un campo diferenciable en $t \in I$ si y sólo si existe $\varphi(u, v)$ parametrización en $\alpha(t_0)$ tal que $\mathbb{W}(t) = a\varphi_u + b\varphi_v$ en la base $\{\varphi_u, \varphi_v\}$, y las aplicaciones $a(t)$ y $b(t)$ son diferenciables.*

Definición 1.33. *Sea $\mathbb{W}$ un campo vectorial diferenciable sobre el conjunto abierto $U \subset M$ y $p \in M$. Sea $\mu \in T_pM$, consideremos una curva parametrizada $\alpha : (-\epsilon, \epsilon) \to U$, con $\alpha(0) = p$ y $\alpha'(0) = \mu$. El vector obtenido al proyectar $(\frac{d\mathbb{W}}{dt})(0)$ normalmente sobre el plano T_pS es llamado derivada covariante en p sobre el campo vectorial $\mathbb{W}$ relativo al vector μ.*
La derivada covariante es denotada por $(\frac{D\mathbb{W}}{dt})(0)$ o $(D_\mu\mathbb{W})(p)$.

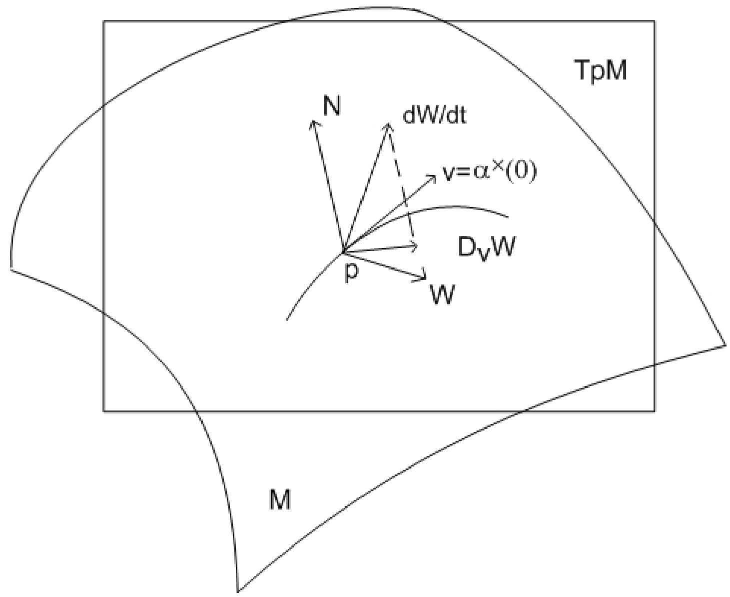

Figura 1.3: Derivada Covariante

Definición 1.34. *Sea* $\mathbb{W}$ *un campo vectorial diferenciable a lo largo de* $\alpha : I \to M$*, la expresión* $\frac{D\mathbb{W}}{dt}(t)$*,* $t \in I$ *está bien definida y es llamada derivada covariante de* $\mathbb{W}$ *en* t*.*

Definición 1.35. *Un campo vectorial* $\mathbb{W}$ *a lo largo de una curva parametrizada* $\alpha : I \to M$ *es paralelo si* $\frac{D\mathbb{W}}{dt}(t) = 0$ *para todo* $t \in I$*.*

Proposición 1.7. *Sea* v *y* $\mathbb{W}$ *campos vectoriales paralelos a lo largo de* $\alpha : I \to M$*, entonces* $\langle \mathbb{W}(t), v(t) \rangle$ *es constante. En particular,* $\|\mathbb{W}(t)\|$*=ctte,* $\|v(t)\|$*=ctte y* $\angle(\mathbb{W}(t), v(t))$ *es constante.*

Demostración: v y $\mathbb{W}$ son campos vectoriales paralelos a lo largo de una curva $\alpha : I \to M$, entonces: $\frac{D\mathbb{W}}{dt} = 0 = \frac{Dv}{dt}$ y $\mathbb{W}'(t) = \frac{d\mathbb{W}(t)}{dt} = \lambda_1 N$ y $v'(t) = \frac{dv(t)}{dt} = \lambda_2 N$.

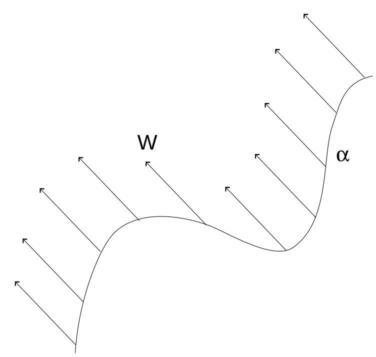

Figura 1.4: Campo Vectorial Paralelo

Así; $\langle \mathbb{W}'(t), v(t)\rangle = 0$ y $\langle \mathbb{W}(t), v'(t)\rangle = 0$
Derivemos $\langle \mathbb{W}(t), v(t)\rangle$.

$(\langle \mathbb{W}(t), v(t)\rangle)' = \langle \mathbb{W}'(t), v(t)\rangle + \langle \mathbb{W}(t), v'(t)\rangle = 0.$
Por lo tanto $\langle \mathbb{W}(t), v(t)\rangle = ctte$.

Proposición 1.8. *Sea $\alpha : I \to M$ una curva parametrizada en M y sea $_0 \in T_{\alpha(t_0)}M$, $t_0 \in I$ entonces existe un único campo vectorial $\mathbb{W}(t)$ a lo largo de $\alpha(t)$, con $\mathbb{W}(t_0) = \mathbb{W}_0$.*

Definición 1.36. *Sea $\alpha : I \to M$ una curva parametrizada y $\mathbb{W}_0 \in T_{\alpha(t_0)}M$, $t_0 \in I$. Sea $\mathbb{W}$ el campo vectorial paralelo a lo largo de α, con $\mathbb{W}(t_0) = \mathbb{W}_0$. El vector $\mathbb{W}_{t_1}$, con $t_1 \in I$ es llamado transporte paralelo de $\mathbb{W}_0$ a lo largo de α en t_1.*

Definición 1.37. *Fijemos $p, q \in M$ y $\alpha : I \to M$ con $\alpha(0) = p$ y $\alpha(1) = q$. Denotemos por $\tau_\alpha(v) : T_pM \to T_qM$ la función que asigna a cada $v \in T_pM$ un transporte paralelo a lo largo de α en q.*

Definición 1.38. *Una curva parametrizada, no constante $\gamma : I \to M$ se dice geodésica en $t \in I$ si el campo vectorial tangente $\gamma'(t)$ es paralelo a lo largo de γ en t; esto es que $\frac{D\gamma'(t)}{dt} = 0$ γ es geodésica si y sólo si γ es geodésica en todo $t \in I$.*

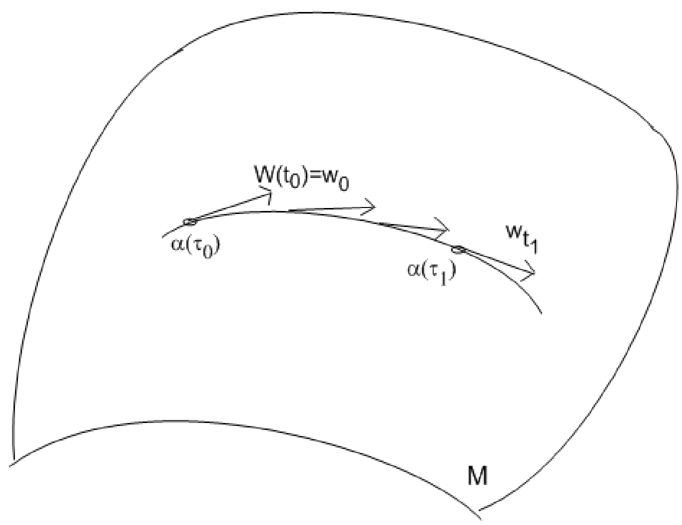

Figura 1.5: Transporte Paralelo.

Ejemplo 1.1. *Supongamos que M es una superficie y que $\alpha : I \to M$ es una curva diferenciable en M, donde $\alpha(0) = p$ y $\alpha'(0) = \lambda$. Consideremos el campo* $\mathbb{W}$ *formado por todos $\alpha'(0) = \lambda \in T_pM$ a lo largo de la curva; así: $\frac{d\mathbb{W}}{dt} = \alpha''(0)$. Si consideramos el caso particular en el cual $\alpha''(0) = \frac{d\mathbb{W}}{dt}$ es paralelo al normal N a lo largo de la curva α, entonces $\frac{D\mathbb{W}}{dt}(t) = 0$, para cualquier $t \in I$.*

1.7. La Función Exponencial.

Dado $p \in M$, M es una superficie y $v \in T_pM$, $v \neq \oslash$. Sabemos que existe una única geodésica $\gamma : (-\epsilon, \epsilon) \to S$ con $\gamma(0) = p$ y $\gamma'(0) = v$ Denotaremos $\gamma(t, v) = \gamma(t) = \gamma$ y $\gamma(t, v)$ indicará que $\gamma'(0) = v$.

Lema 1.1. *Si $\gamma(t, v)$ es geodésica, $t \in (-\epsilon, \epsilon)$ entonces la geodésica $\gamma(t, \gamma)$, $\gamma \in \mathbb{R}$, $\gamma \neq 0$, está definida para $t \in (\frac{-\epsilon}{\gamma}, \frac{\epsilon}{\gamma})$ y $\gamma(t, \lambda v) = \gamma(\lambda t, v)$.*

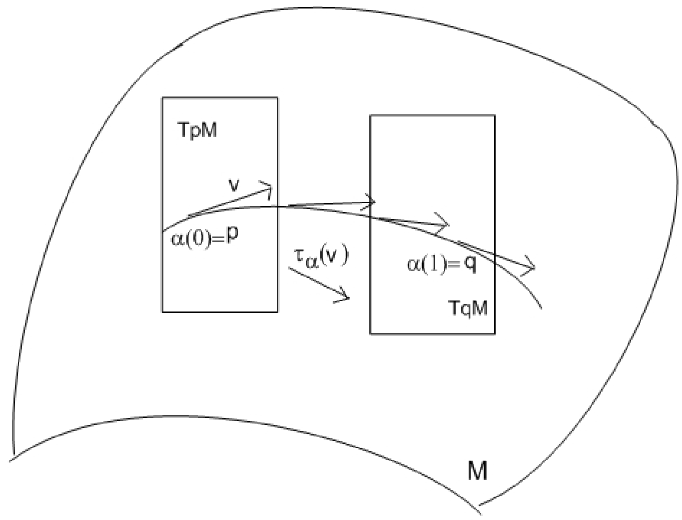

Figura 1.6: Función Transporte Paralelo.

Demostración: Sea $\alpha : (\frac{-\epsilon}{\lambda}, \frac{\epsilon}{\lambda}) \to M$ una curva parametrizada definida por $\alpha(t) = \gamma(\lambda t)$, entonces $\alpha'(0) = \lambda\gamma'(0) = \lambda v$. Por la linealidad de la derivada covariante se tiene que:
$D_{\alpha'(t)}\alpha'(t) = D_{\lambda\gamma'(t)}\lambda\gamma'(t) = \lambda^2 D_{\gamma'(t)}\gamma'(t)$, de acá se tiene que α es una geodésica con las siguientes condiciones iniciales $\gamma(0)$, $\lambda\gamma'(0)$.
Así por unicidad, $\alpha(t) = \gamma(t, \lambda v) = \gamma(\lambda t, v)$ Notación: si $v \in T_pM$ y $v \neq 0$ es tal que $\gamma(\|v\|, \frac{v}{\|v\|}) = \gamma(1, v)$, se define la exponencial dada como sigue: $exp_p(v) = \gamma(1, v)$ y $exp_p(0) = p$.

Proposición 1.9. *Dado $p \in M$, existe un $\epsilon > 0$ tal que $exp_p : B_\epsilon(0) \subset T_pM \to M$, es un difeomorfismo de $B_\epsilon(0)$ sobre un subconjunto abierto de M.*

Demostración Consideremos la curva $\alpha : (-\epsilon, \epsilon) \to M$ tal que $\alpha(t) = tv$, $v \in T_pM$, $\alpha(0) = 0$, $\alpha'(t) = v$, $\alpha'(0) = v$, $\forall t \in (-\epsilon, \epsilon)$

Luego:

$$(exp_p \circ \alpha)(t) = exp_p(\alpha(t)) = exp_p(tv).$$

Derivando: $\frac{d}{dt}(exp_p(tv))\Big|_{t=0} = \frac{d}{dt}(\lambda(t;v))\Big|_{t=0} = v$

De ahí, $d(exp_p)_0$ es la identidad de T_pM, y se sigue por teorema de la función inversa que exp_p es un difeomormisfo local en una vecindad de 0.

1.8. Superficies Abstractas.

Definición 1.39. *Una variedad diferenciable de dimension n es un conjunto M junto con una familia de funciones uno a uno, $X_\alpha : U_\alpha \to M$, donde $U_\alpha \subset (\mathbb{R})$ son abiertos en M tales que:*

1. $\bigcup_\alpha X_\alpha(U_\alpha) = M$
2. *$\forall \alpha, \beta$ con $X_\alpha(U_\alpha) \cap X_\beta(U_\beta) = M \neq \varnothing$ se tiene que: $X_\alpha^{-1} \circ X_\beta$ y $X_\beta^{-1} \circ X_\alpha$ son diferenciables.*
3. *La familia $\{U_\alpha, X_\alpha\}$ (estructura diferenciable para M) es maximal relativo a 1 y 2.*

Definición 1.40. *$f : M \to \mathbb{R}$ sera diferenciable si y solo si $\forall \alpha : f \circ X_\alpha : U_\alpha \subset \mathbb{R} \to \mathbb{R}$ es diferenciable.*

Definición 1.41. *Sea $D = \{f : M \to \mathbb{R} : f$ es diferenciable $\}$ y sea $\alpha : (-\epsilon, \epsilon) \to M$ curva diferenciable. Llamaremos vector tangente a α en $t = 0$ a la funcion $\alpha'(o) : D \to \mathbb{R}$, dada por: $\alpha'(o)(f) = \frac{d}{dt}(f \circ \alpha)\Big|_{t=0}$. Una base para $T_pM = \{\alpha'(0)/\alpha : (-\epsilon, \epsilon) \to M$ es curva diferenciable$\}$ sería:*

$$\left\{ \left(\frac{\partial}{\partial x_1}\right)_0, \left(\frac{\partial}{\partial x_2}\right)_0, ..., \left(\frac{\partial}{\partial x_n}\right)_0 \right\}.$$

Definición 1.42. *Una funcion diferenciable* $\alpha : (-\epsilon, \epsilon) \to M$ *es llamada curva sobre* M*. Supongamos que* $\alpha(0) = p$, $D = \{f : M \to \mathbb{R}/\ f$ *es diferenciable en* $p\}$ *El vector tangente a* α *en* $t = 0$ *es la funcion:*

$$\alpha'(0) : D \to \mathbb{R}$$

dada por:

$$\begin{aligned}
\alpha'(0)(f) &= \frac{d}{dt}(f \circ \alpha)\bigg|_{t=0}, f \in D \\
&= u'(0)\left(\frac{\partial f}{\partial u}\right)_0 + v'(0)\left(\frac{\partial f}{\partial v}\right)_0 \\
&= \left\{u'(0)\left(\frac{\partial}{\partial u}\right)_0 + v'(0)\left(\frac{\partial}{\partial v}\right)_0\right\} f.
\end{aligned}$$

Definición 1.43. *Una Variedad Riemanniana de dimension* α *es una superficie abstracta junto a la escogencia de un producto interno* $\langle , \rangle_p$ *en cada* $T_pM, p \in M$ *el cual varia diferencialmente con* p *en el siguiente sentido, para alguna (y por tanto para todas) parametrizacion* $X_\alpha : U_\alpha \to M$ *las funciones:*

- $E(u, v) = g_{11} = \langle \frac{\partial}{\partial u}, \frac{\partial}{\partial u} \rangle$
- $F(u, v) = g_{12} = g_{21} = \langle \frac{\partial}{\partial u}, \frac{\partial}{\partial v} \rangle$
- $G(u, v) = g_{22} = \langle \frac{\partial}{\partial v}, \frac{\partial}{\partial v} \rangle$

son funciones diferenciables en U_α. Al producto interno $\langle , \rangle$ se le llama metrica Riemanniana sobre M.

Definición 1.44. *Sean* M_1, M_2 *superficies abstractas y sea* $\varphi : M_1 \to M_2$ *diferenciable* $\forall p \in M_1, \forall w \in T_pM_1$ *consideremos* $\alpha : (-\epsilon, \epsilon) \to M_1$ *con* $\alpha(0) = p$ *y* $\alpha'(0) = w$*. Sea* $\beta = \varphi \circ \alpha$ *y la función:* $d\varphi_p : T_pM_1 \to T_{\varphi(p)}M_2$ *dada por* $d\varphi_p(w) = \beta'(0)$*, esta bien definida y es lineal, la cual llamaremos la diferencial de* φ *en* p

Definición 1.45. *Una funcion diferenciable* $\varphi_{:}M \to \mathbb{R}$ *de una superficie abstracta, es una inmersion si* $d\varphi_p : T_pM \to T_p\mathbb{R}$ *es inyectiva y si ademas* M *tiene metrica* $\langle , \rangle$ *y* $\langle d\varphi_p(v), d\varphi_p(w)\rangle_{d\varphi_p} = \langle v, w\rangle_p$, *se dice que* φ *es una inmersion isometrica.*

Teorema 1.7 (Hopf-Rinow)**.** *Sea* M *(conexa) una variedad Riemanniana. Las siguientes proposiciones son equivalentes:*

1. *Para cada* $x \in M$,exp_x, *es definida en todo* T_xM, *esto es* M *es completo.*
2. (M, d) *es un espacio métrico completo, es decir, cualquier sucesión de Cauchy de* M *es una sucesión convergente.*

   Además, cada una de las afirmciones anteriores implica que:
3. *Dos puntos cualesquiera* $x, x' \in M$, *se pueden unir por geodésicas de longitud* $l(\gamma) = d(x, x')$, *las geodésicas con esta propiedad se llama minimal.*

Demostración: *Ver referencia [2]*

Lema 1.2 (Gauss)**.** *Sea* $p \in M$ *y sea* $v \in T_pM$ *tal que* exp_pv *es definida.*
Sea $w \in T_pM \approx T_v(T_pM)$ *Entonces:*

$$\langle (dexp_p)_v(v), (dexp_p)_v(w)\rangle = \langle v, w\rangle. \tag{1.4}$$

Demostración: Consideremos $w = w_T + w_N$, donde w_T es paralelo a v y w_N es normal a v. Así:

$$\begin{aligned}\langle (dexp_p)_v(v), (dexp_p)_v(w)\rangle &= \langle (dexp_p)_v(v), (dexp_p)_v(w_T + w_N)\rangle \\ &= \langle (dexp_p)_v(v), (dexp_p)_v(w_T) + (dexp_p)_v(w_N)\rangle \\ &= \langle (dexp_p)_v(v), (dexp_p)_v(w_T)\rangle \\ &+ \langle (dexp_p)_v(v), (dexp_p)_v(w_N)\rangle\end{aligned}$$

Veamos que es suficiente probar 1.4 para $w = w_N$ donde se asume que $w_N \neq 0$ Como $exp_p v$ esta definida, entonces existe $\epsilon > 0$ tal que $exp_p u$ esta dada por:

$$u = tv(s), 0 \leq t \leq 1, -\epsilon < s < \epsilon.$$

donde $v(s)$ es una curva en T_pM con $v(0) = v$,$v'(0) = w_N$ y $|v(s) = const|$. Consideremos además la parametrización de la superficie: $f : A \to M$ donde $A = \{(t, s), 0 \leq t \leq 1, -\epsilon < s < \epsilon\}$ dada por $f(t, s) = exp_p tv(s)$ Observemos que las curvas $t \to f(t, s)$ son geodésicas Vemos que para $w = w_N$ se tiene que:

$$\langle \frac{\partial f}{\partial s}, \frac{\partial f}{\partial t} \rangle (1, 0) = \langle (dexp_p)_v(w_N), (dexp_p)_v(v) \rangle \tag{1.5}$$

Además, para todo (t, s), tenemos:

$$\frac{\partial}{\partial t} \langle \frac{\partial f}{\partial s}, \frac{\partial f}{\partial t} \rangle = \langle \frac{D}{\partial t} \frac{\partial f}{\partial s}, \frac{\partial f}{\partial t} \rangle + \langle \frac{\partial f}{\partial s}, \frac{D}{\partial t} \frac{\partial f}{\partial t} \rangle.$$

El último termino de la expresión es igual a cero, como $\frac{\partial f}{\partial s}$ es el vector tangente de una geodésica. Así el primer término de la suma es transformado en:

$$\langle \frac{D}{\partial t} \frac{\partial f}{\partial s}, \frac{\partial f}{\partial t} \rangle = \langle \frac{D}{\partial s} \frac{\partial f}{\partial t}, \frac{\partial f}{\partial t} \rangle = \frac{1}{2} \frac{\partial}{\partial s} \langle \frac{\partial f}{\partial t}, \frac{\partial f}{\partial t} \rangle = 0.$$

En lo que sigue $\langle \frac{\partial f}{\partial s}, \frac{\partial f}{\partial t} \rangle$ es independiente de t. Dado que:

$$\lim_{t \to 0} \frac{\partial \partial f}{\partial s}(t, 0) = \lim_{t \to 0} (dexp_p)_{tv} t w_N = 0.$$

Concluimos que $\langle \frac{\partial f}{\partial s}, \frac{\partial f}{\partial t} \rangle (1, 0) = 0$, uniendo con 1.5 nos queda:

$$\langle (dexp_p)_v(w_N), (dexp_p)_v(v) \rangle = 0.$$

Así:

$$\langle (dexp_p)_v(v), (dexp_p)_v(w) \rangle = \langle v, w \rangle.$$

Definición 1.46. *Una bisagra geodésica $(p; \gamma_1, \gamma_2)$ en M es una figura que consiste en un punto $p \in M$ llamado vertice y segmentos geodésicos minimales γ_1 y γ_2 emanados de p llamados lados. Nosotros denotaremos por α el ángulo entre los vectores tangentes a γ_1 y γ_2 en p, el cual es llamado el ángulo de* *la bisagra geodésica* $(p;\gamma_1,\gamma_2)$, *si* γ_1 *es minimal, pero* γ_2 *no necesariamente minimal, llamaremos a está bisagra geodésica generalizada. Si* $\ell_1=\ell(\gamma_1)$, $\ell_2=\ell(\gamma_2),\ell_3=\ell(\gamma_3)$ *y* $\alpha=\angle(\gamma_1'(0),\gamma_2'(0))$, *obtenemos el siguiente resultado:*

Teorema 1.8 (Toponogov)**.** *Sea* M *una variedad Riemanniana completa cuya curvatura seccional satisfaciendo que* $K\geq 0$. *Y consideremos* T_pM *variedad de curvatura constante* 0. *Sea* $(p;\gamma_1',\gamma_2')$ *una bisagra geodésica en* T_pM *tal que:* $\ell(\gamma_1')=\ell(\gamma_1),\ell(\gamma_2')=\ell(\gamma_2)$ *y su ángulo* α' *es igual al de* $(p;\gamma_1,\gamma_2)$. *Sean* q,s *los puntos finales de* γ_1 *y* γ_1' *respectivamente y sean* r,t *los puntos finales de* γ_2 *y* γ_2' *respectivamente, entonces,*

$$d(q,r)\leq d(s,t)$$

Demostración: Ver referencia [2]

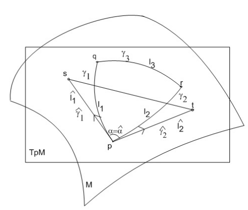

Figura 1.7: Teorema de Toponogov.

Corolario 1.1. *Sea* M *una variedad Riemanniana completa con curvatura seccional* $K\geq 0$. *Si* γ_{v_1} *y* γ_{v_2} *son geodésicas normalizadas tal que* $\gamma_{v_1}(0)=\gamma_{v_2}(0)$, *entonces:*

$$d(\gamma_{v_1}(t_1),\gamma_{v_2}(t_2))\leq\|t_2v_2-t_1v_1\|.$$

Definición 1.47. *Una sucesión* $\{y_k\}$ *es el espacio métrico compacto* (M,d) *es casi-Féjer convergente al conjunto* $W\subset M$ *si, para cada* $w\in W$, *existe una sucesión* $\{\epsilon_k\}\subset\mathbb{R}$ *tal que* $\epsilon_k\geq 0,\sum_{t=0}^{\infty}\epsilon_k<+\infty$ *y*

$$d^2(y_{k+1},w)\leq d^2(y_k,w)+\epsilon_k\ \ \forall k \tag{1.6}$$

Teorema 1.9. *Sea $\{y_k\}$ una sucesión es el espacio métrico (M, d). Si $\{y_k\}$ es casi-Féjer convergente a un conjunto no-vacío $W \subset M$, entonces $\{y_k\}$ es acotada. Si además y es un punto clausura de y_k perteneciente a W, entonces*

$$\lim_{k\to\infty} y_k = y.$$

Demostración: Sea $\{y_k\}$ una sucesión casi-Féjer convergente al conjunto $W \subset M$. Probemos que $\{y_k\}$ es acotada. Sea $w \in W$, así:

$$\begin{aligned} d^2(y_k, w) &\leq d^2(y_{k-1}, w) + \epsilon_{k+1} \\ &\leq d^2(y_{k-2}, w) + \epsilon_{k-2} + \epsilon_{k-1} \\ &. \\ &. \\ &. \\ &\leq d^2(y_{k-k}, w) + \epsilon_0 + \epsilon_1 + \epsilon_{k-2} + \epsilon_{k-1} \\ &= d(y_0, w) + \sum_{j=0}^{k-1} \epsilon_j \\ &\leq d(y_0, w) + \sum_{j=0}^{\infty} \epsilon_j \end{aligned}$$

Así, si fijamos w, tenemos que $\{y_k\}$ es acotada ya que esta contenida dentro de la bola con centro en W y radio $d(y_0, w) + \sum_{j=0}^{\infty} \epsilon_j$. Sea $y \in W$ un punto clausura de $\{y_k\}$ y $\delta > 0$ entonces existe y_{l_k} sub-sucesión de y_k tal que $y_{l_k} \to y$ como y_{l_k} es casi-Féjer, existe $\epsilon_k \subset \mathbb{R}$ y $\sum_{j=0}^{\infty} \epsilon_j < \infty$. Para $\delta > 0$ existe k_0 tal que $\sum_{j=0}^{\infty} \epsilon_j < \frac{\delta}{2}$ y existe $k_1 \geq k_0$ tal que:

$$k \geq k_1 \Rightarrow d^2(\{y_k\}, y) < d^2(y_{l_k}) + \sum_{j=0}^{k-1} \epsilon_j \leq \frac{\delta}{2} + \frac{\delta}{2} = \delta.$$

Así:

$$\lim_{k\to+\infty} \{y_k\} = y.$$

Capítulo 2

Preliminares de Optimización

2.1. Funciones Convexas en Variedades de Riemann

Definición 2.1. *Una función real f definida en una variedad Riemanniana completa, se dice que es una función convexa, si f es convexa cuando es restringida a cualquier geodesica de M, lo que significa:*

$$(f \circ \gamma)(ta + (1-t)b) \leq tf(\gamma(a)) + (1-t)f(\gamma(b)).$$

Se cumple para cualquier $a, b \in \mathbb{R}$ y $0 \leq t \leq 1$

Derivada direccional y subgradiente:

Definición 2.2. *Sea f una función de valores reales y $\gamma(t)$ con $0 < t < 1$ es la geodésica para la cual $\gamma(0) = x$, $\gamma'(0) = v_x$ entonces el límite:*

$$f'(x; v_x) = \lim_{t \downarrow 0} \frac{f(\gamma(t)) - f(x)}{t} \tag{2.1}$$

es llamada la derivada direccional de f en x con respecto a v_x. Dado

$$-f'(x; -v_x) = \lim_{t \uparrow 0} \frac{f(\gamma(t)) - f(x)}{t} \tag{2.2}$$

la derivada direccional $f'(x; v_x)$ es bilateral si y solo si $f'(x; -v_x)$ existe y $f'(x; -v_x) = -f'(x; v_x)$.

Definición 2.3. *Sea f una función convexa y $x \in M$. El vector $s \in T_xM$ es llamado subgradiente de f en x si por cualquier geodesica γ de M con $\gamma(0) = x$*

$$(f \circ \gamma)(t) \geq f(x) + t\langle s, \gamma'(0)\rangle, \tag{2.3}$$

para cualquier $t \geq 0$. El conjunto de todos los subgradientes denotado por $\partial f(x)$, es llamado subdiferencial de f en x.

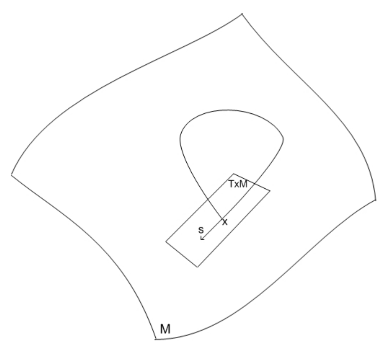

Figura 2.1: Definición de Subgradiente.

Definición 2.4. *Una función $f : M \to \mathbb{R}$ es llamada lipschitziana continua si existe un número real B tal que:*

$$|f(x) - f(y)| \leq Bd(x, y),$$

para todo $x, y \in M$, donde $d(x, y)$ es la distancia entre los puntos x, y. El número B es llamado constante de Lipschitz para f.

Lema 2.1. *Una función $f : M \to \mathbb{R}$ es continua lipschitz con constante lipschitz B si y solo si f es lipschitz continua con constante lipschitz B en una vecindad de cada punto de M, es decir, para cada $x \in M$, existe una vecindad U_x de x tal que:*

$$|f(x_1) - f(x_2)| \leq Bd(x, y)$$

para cada $x_1, x_2 \in U_x$.

Demostración: ($\Rightarrow$) La continuidad Lipschitz implica la condición local ya que se puede tomar $U_x = M$ para cada $x \in M$.
($\Leftarrow$) Sean $x_1, x_2 \in M$ consideremos la $d(x_1, x_2)$ la se sabe que esta definida por $d(x_1, x_2) = \inf L(\alpha)$, donde $\alpha : [0, 1] \to M$, es una curva regular que satisface $\alpha(0) = x_1$, $\alpha(1) = x_2$. Así, para establecer la continuidad Lipschitz de f con constante de Lipschitz B es suficiente probar que para cualquier curva α se tiene:

$$|f(x_1) - f(x_2)| \leq BL(\alpha).$$

Para esto se escoge una partición finita de $[0, 1]$ por puntos
$t_0 = 0 < t_1 < ... < t_n = 1$ tal que para todo $i = 0, 1, ..., n - 1$, la imagen $\alpha([t_i, t_{i+1}])$ esta contenida en una vecindad U_x sastifaciendo la condición del lema para algun $x \in M$. Ahora, escogiendo $\{U_x\}_x \in \alpha([0, 1])$ la cual es posible hacerla debido a que $\{U_x\}_{x \in M}$ forman un cubrimiento abierto de M.
Entonces:

$$\begin{aligned}
|f(x_1) - f(x_2)| &\leq |f(x_1) - f(\alpha(t_1))| + |f(\alpha(t_1)) - f(\alpha(t_2))| + ... + |f(\alpha(t_{n-1})) - f(x_2)| \\
&= \sum_{i=0}^{n-1} |f(\alpha(t_i)) - f(\alpha(t_{i+1}))| \\
&\leq \sum_{i=0}^{n-1} Bd(\alpha(t_i), \alpha(t_{i+1})) \\
&\leq B \sum_{i=0}^{n-1} d(\alpha(t_i), \alpha(t_{i+1})) \\
&= Bd(x_1, x_2).
\end{aligned}$$

Por tanto f es lipschitz.

Teorema 2.1. *Sea f una función convexa entonces f es localmente lipschitziana.*

Demostración: Sea f una funcion convexa, entonces por definición tenemos que:

$$(f \circ \gamma)(ta + (1-t)b) \leq tf(\gamma(a)) + (1-t)f(\gamma(b))$$

para cualquier $a, b \in \mathbb{R}$ y $t \in [0,1]$. Consideremos $x_1, x_0, x_2 \in M$ y $\gamma : [-1,1] \to M$ una geodésica en $\overline{B}(x_0, r)$ que conecta x_1 con x_2 tal que $\gamma(-1) = x_1$, $\gamma(0) = x_0$, $\gamma(1) = x_2$. Sabemos que $\overline{B}(x_0, r)$ es compacto y f por ser convexa es continua entonces existe c tal que $f(x) \leq c$.
Ahora, si consideramos $x \in B(x_0, r)$ denotamos $\alpha(t) = x$, donde $t = \frac{d(x_0,x)}{r} \in [0,1]$. Asi:

$$\begin{aligned} f(\gamma(t)) &\leq (1-t)f(x_0) + tf(x_2) \\ &\leq f(x_0) - tf(x_0) + tf(x_2) \\ &\leq f(x_0) - tf(x_0) + tc \end{aligned}$$

Entonces:

$$f(x) - f(x_0) \leq t(c - f(x_0)).$$

La restricción del arco de geodésica que conecta x_1 y x es la restricción $\gamma(u)$, con $u \in [-1, t]$. Haciendo $u = -1 + s(t+1)$, $s \in [0,1]$ tomando una reparametrización:

$$\alpha(s) = \gamma(-1 + s(t+1)),$$

$s \in [0,1]$ así; $\alpha(0) = \gamma(-1) = x_1$; $\alpha(\frac{1}{1+t}) = \gamma(0) = x_0$, $\alpha(1) = \gamma(t) = x$. Ahora, por la continuidad de f tenemos:

$$f(\alpha(s)) \leq (1-s)f(x_1) + sf(x) \leq (1-s)c + sf(x).$$

Se tiene que:

$$f(x_0) \leq \frac{t}{1+t}c + \frac{1}{1+t}f(x).$$

ó

$$-t(c - f(x_0)) \leq f(x) - f(x_0).$$

Así;

$$\begin{aligned}
|f(x) - f(x_0)| &\leq t(c - f(x_0)) \\
&= \frac{d(x_0, x)}{r}(c - f(x_0)) \\
&= \frac{c - f(x_0)}{r} d(x_0, x) \\
&= Bd(x_0, x)
\end{aligned}$$

Por tanto, toda función convexa es localmente lipschitziana.

Teorema 2.2. *Sea f una función convexa. Entonces, para cada $x \in M$, $\partial f(x)$ es no vacía, convexa y compacta.*

Demostración: Sea f una función convexa restringida a $\gamma : [0, 1] \to M$ geodésica en M tal que $\gamma(0) = x$, $\gamma(1) = y$ y $\gamma'(0) = v_x$. Así se tiene que:

$$\begin{aligned}
(f \circ \gamma)(t) &\leq tf(\gamma(1)) + (1 - t)f(\gamma(0)) \\
f(\gamma(t)) &\leq f(x) - tf(x) + tf(y) \\
f(\gamma(t)) - f(x) &\leq tf(y) - tf(x) \\
\frac{f(\gamma(t)) - f(x)}{t} &\leq f(x) - f(y)
\end{aligned}$$

Entonces:

$$f(x) - f(y) \geq \frac{f(\gamma(t)) - f(x)}{t}.$$

Pero:

$$\frac{f(\gamma(t)) - f(x)}{t} \geq Df(x; v_x)$$

donde $Df(x; v_x)$ es la derivada unilateral de f en x con respecto a $\gamma'(0) = v_x$, donde γ es cualquier geodésica de M. Sea $s \in T_xM$ tal que:

$$\langle s, \gamma'(0)\rangle \leq Df(x; v_x).$$

Así:

$$\frac{f(\gamma(t)) - f(x)}{t} \geq Df(x; v_x) \geq \langle s, \gamma'(0) \rangle.$$

Lo que implica que:

$$f(\gamma(t)) - f(x) \geq t\langle s, \gamma'(0) \rangle.$$

Entonces:

$$f(\gamma(t)) \geq f(x) + t\langle s, \gamma'(0) \rangle.$$

Es decir, $s \in \partial f(x)$, por tanto $\partial f(x)$ no es vacío. Sea $s_1, s_2 \in \partial f(x)$. Entonces:

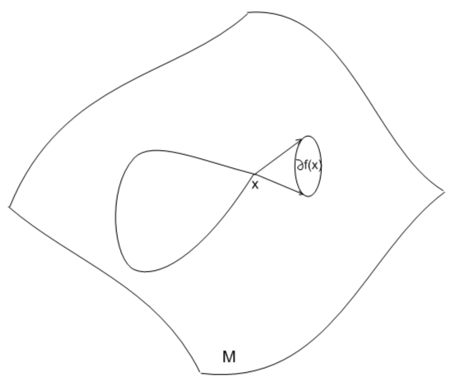

Figura 2.2: El Conjunto Subdiferencial.

$$(f \circ \gamma)(t) \geq f(x) + t\langle s_1, \gamma'(0) \rangle, \forall t \geq 0 \tag{2.4}$$

$$(f \circ \gamma)(t) \geq f(x) + t\langle s_2, \gamma'(0) \rangle, \forall t \geq 0 \tag{2.5}$$

multipliquemos 2.4 por $(1 - \alpha)$ y 2.5 por α donde $\alpha \in [0, 1]$ se tiene:

$$(1 - \alpha)(f \circ \gamma)(t) \geq (1 - \alpha)f(x) + (1 - \alpha)t\langle s_1, \gamma'(0) \rangle.$$

$$\alpha(f \circ \gamma)(t) \geq \alpha f(x) + \alpha t\langle s_2, \gamma'(0) \rangle.$$

Luego, sumando:

$$(1-\alpha)(f\circ\gamma)(t)+\alpha(f\circ\gamma)(t) \geq (1-\alpha)f(x)+(1-\alpha)t\langle s_1, \gamma'(0)\rangle+\alpha f(x)+\alpha t\langle s_2, \gamma'(0)\rangle.$$

Tenemos:

$$
\begin{aligned}
(f \circ \gamma)(t) &\geq f(x) + t\langle (1-\alpha)s_1, \gamma'(0)\rangle + t\langle \alpha s_2, \gamma'(0)\rangle \\
&\geq f(x) + t\langle (1-\alpha)s_1 + \alpha s_2, \gamma'(0)\rangle
\end{aligned}
$$

así:

$$(1-\alpha)s_1 + \alpha s_2 \in \partial f(x).$$

Por tanto $\partial f(x)$ es convexo.

Probemos ahora que $\partial f(x)$ es cerrado y acotado, es decir, compacto. Primero veamos que es cerrado: Consideremos $\{s_n\} \subset \partial f(x)$ una sucesión de subgradientes convergente a s_x, supongamos por absurdo que $s_x \notin \partial f(x)$, es decir, $\exists \epsilon > 0$ tal que para cualquier $t \geq 0$ se tiene que:

$$(f \circ \gamma)(t) + \epsilon = f(x) + t\langle s_x, \gamma'(0)\rangle \tag{2.6}$$

Por otro lado tenemos que:

$$(f \circ \gamma)(t) \geq f(x) + t\langle s_n, \gamma'(0)\rangle \tag{2.7}$$

Por sustracción de 2.6 y 2.7 nos queda:

$$(f \circ \gamma)(t) + \epsilon - (f \circ \gamma)(t) \leq f(x) + t\langle s_x, \gamma'(0)\rangle - f(x) - \langle s_n, \gamma'(0)\rangle.$$

Así:

$$
\begin{aligned}
&\epsilon \leq t\langle s_x, \gamma'(0)\rangle - t\langle s_n, \gamma'(0)\rangle \\
&\epsilon \leq t(\langle s_x, \gamma'(0)\rangle - \langle s_n, \gamma'(0)\rangle) \\
&\epsilon \leq t(\langle s_x - s_n, \gamma'(0)\rangle) \\
&\epsilon \leq t\|s_x - s_n\|\|\gamma'(0)\| \cos\angle(s_x - s_n, \gamma'(0))
\end{aligned}
$$

Lo que implica que:

$$0 < \frac{\epsilon}{t\|\gamma'(0)\|} \leq \|s_x - s_n\| \ \ \forall n \in \mathbb{N}.$$

Lo cual contradice el hecho de que $\{s_n\}$ converge a s_x, de ahi, $s_{x \in \partial f(x)}$. Por tanto $\partial f(x)$ es cerrado. Veamos ahora que es acotado: Sea $s_1 \in \partial f(x)$, es decir:

$$t\langle s_1, \gamma'(0)\rangle + f(x) \leq (f \circ \gamma)(t).$$

ó

$$t\|s_1\|\|\gamma'(0)\| \cos \angle(s_1, \gamma'(0)) + f(x) \leq (f \circ \gamma)(t)$$

$\forall t \geq 0$ y $\forall \gamma \in M$ con $\gamma(0) = x$ y como $\|s_1\|$ es finito por tanto $\partial f(x)$ es acotada. Vemos entonces que $\partial f(x)$ es cerrado y acotado es decir compacto.

Capítulo 3

Desarrollo

Observación 3.1. *Antes de plantear la definición del problema se presentan algunas observaciones que serán útiles para definiciones y teoremas que serán estudiados.*

- Cuando la referencia al punto x no es necesaria ni esta ímplicita la notación γ_v significa que $\gamma_v'(0) = v$
- Si $\|v\| = 1$, la geodésica γ_v se dice que esta parametrizada por longitud de arco o normalizada.
- Si $exp_x : V \to U$ donde $V \subset T_xM$ y $U \subset M$, es un difeomorfismo, entonces U es llamada vecindad normal de x
- Si $B_\epsilon(0) := \{v \in T_xM / \|v\| < \epsilon\}$ es tal que $\overline{B_\epsilon(0)} \subset V$, llamaremos $exp_x B_\epsilon(0) := B_\epsilon(x)$ la bola normal o geodésica con centro en x y radio ϵ.

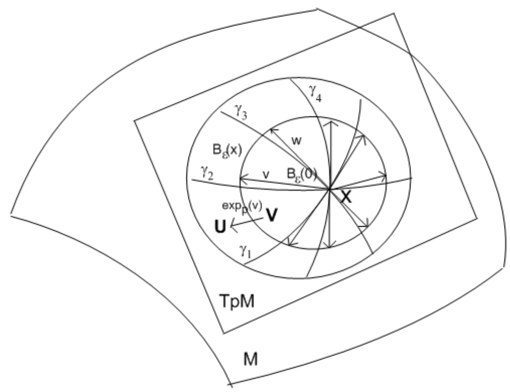

Figura 3.1: Difeomorfismo de la Exponencial..

3.1. Definición del problema y algoritmo:

Definición 3.1. *Sea M una variedad Riemanniana completa y conexa y sea f una función convexa no necesariamente diferenciable. El conjunto θ^* denota el conjunto de minimizadores de f y*

$$f^* = \inf_{x \in M} f(x).$$

denota el valor infimal. El problema es estimar f^* y tambien encontrar un punto de θ^*, si tal punto existe, esto es $\theta^* \neq \emptyset$

Algoritmo: Algoritmo Subgradiente. La sucesión $\{t_k\}$ es dada, con $t_k > 0$ para $k = 1, 2, \ldots$

- Paso 0: Inicializar. Se escoge $x_1 \in M$ y se obtiene $s_1 \in \partial f(x_1)$. Hacer $k = 1$

- Paso 1: Si $s_k = 0$, Paramos. Caso contrario, calculamos la geodésica γ_{v_k} con $\gamma_{v_k}(0) = x_k$, escoger $0 < t_k < [2/\|s_k\|][f(x_k) - f(x_*)]$ donde $\gamma'_{v_k}(0) = v_k$ y $v_k = -s_k/\|s_k\|$.

- Paso 2: Hacer $x_{k+1} = \gamma_{v_k}(t_k)$.
- Paso 3: Obtener $s_{k+1} \in \partial f(x_{k+1})$. Hacer $k = k + 1$ y volver al caso 1

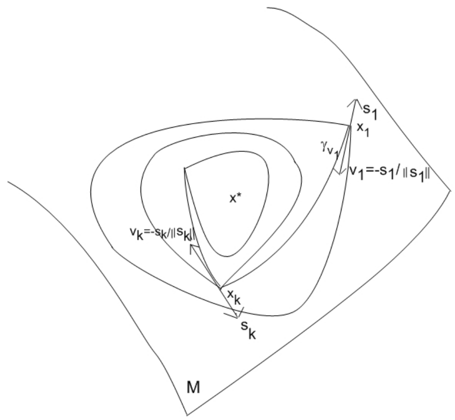

Figura 3.2: Iteración del Algoritmo Subgradiente.

3.2. Resultados Preliminares

Teorema 3.1. *Sea $x_* \in \theta^* \neq \emptyset$, y sea $B_\epsilon(x_*)$ una bola normal. Si $x_k \notin \theta^*$, $x_k \in B_\epsilon(x_*)$, entonces existe $\delta_k > 0$ tal que escogiendo $0 < t_k < \delta_k$ en el algoritmo.*

$$d(x_{k+1}, x_*) < d(x_k, x_*).$$

Demostración: Sea γ_v una geodésica parametrizada por longitud de arco, dada por $\gamma_v(0) = x_k$ y $\gamma_v(t_k) = x_*$ donde $d(x_k, x_*) = t^*$. Consideremos la esfera normal $S_{t^*}(x_*)$, la cual es límite de la bola normal B_{t^*}, con:

$$T_{x_k}S = \{u \in T_{x_k}M / \langle u, v\rangle = 0\}.$$

Sea $s_k \in \partial f(x_k)$. Así por definición tenemos que:

$$(f \circ \gamma_v)(t) \geq f(x_k) + t\langle s_k, v\rangle.$$

Ahora haciendo $t = t_*$ nos queda que:

$$\begin{aligned}(f \circ \gamma_v)(t_*) &\geq f(x_k) + t_* \langle s_k, v \rangle \\ f(\gamma_v(t_*)) &\geq f(x_k) + t_* \langle s_k, v \rangle \\ f(x_*) &\geq f(x_k) + t_* \langle s_k, v \rangle \\ f(x_*) - f(x_k) &\geq t_* \langle s_k, v \rangle\end{aligned}$$

Pero como $x_k \notin \theta^*$, tenemos que:

$$f(x_*) < f(x).$$

Lo que implica que $\langle s_k, v \rangle < 0$ ya que $t_* > 0$. Asi, existe $\delta_k > 0$ tal que $\gamma_{v_k}(t) \in B_{t^*}(x_*)$ para todo $0 < t < \delta_k$, donde $v_k = -s_k/\|s_k\|$. Por lo tanto, para todo $0 < t_k < \delta_k$

$$d(\gamma_{v_k}(t_k), x_*) < d(x_k, x_*)$$

Así:

$$d(x_{k+1}, x_*) < d(x_k, x_*).$$

Observación 3.2. *De aquí en adelante consideraremos la curvatura seccional de la superficie $K \geq 0$, debido a su influencia en el comportamiento de la sucesión $\{x_k\}$ definida en el algoritmo, ya que cuando $K \geq 0$ las geodésicas tiende a aproximarse unas a otras, lo que sugiere que podemos ir más lejos a lo largo de las geodésicas sin distanciarnos de x_*.*

Lema 3.1. *Sea $\{x_k\}$ una sucesión generada por algoritmo. Si M tiene curvatura seccional $K \geq 0$, entonces para todo $y \in M$.*

$$d^2(x_{k+1}, y) \leq d^2(x_k, y) + t_k^2 + 2[t_k\|s_k\|][f(y) - f(x_k)] \;\; \forall k \in \mathbb{N}$$

Demostración: Sea γ_{v_1} una geodésica parametrizada por longitud arco, esto es, para v_1 se tiene que $\|v_1\| = 1$, donde $\gamma_{v_1}(0) = x_k$, $\gamma_{v_1}(t_1) = y$, y $t_1 = d(x_k, y)$. Por el algoritmo podemos considerar otra geodésica γ_{v_2} tal que

$v_2 = v_k = -s_k/\|s_k\|$, $\gamma_{v_2}(0) = x_k$, $\gamma_{v_2}(t_k) = x_{k+1}$ con $t_k = t_2$. Así, por corolario 1,1 tenemos que:

$$\begin{aligned}
d_2(\gamma_{v_1}(t_1), \gamma_{v_k}(t_k)) &\leq \|t_k v_k - t_1 v_1\|^2 \\
&\leq \| - t_k[s_k/\|s_k\|] - t_1 v_1\|^2 \\
&= t_1^2\|v_1\| + 2[t_k/\|s_k\|]\langle s_k, t_1 v_1\rangle + t_k^2[s_k/\|s_k\|] \\
&= t_1^2 + t_k^2 + 2[t_k/\|s_k\|]\langle s_k, t_1 v_1\rangle \\
&\leq d^2(x_k, y) + t_k^2 + 2[t_k/\|s_k\|][f(y) - f(x_k)]
\end{aligned}$$

Entonces:

$$d^2(x_{k+1}, y) \leq d^2(x_k, y) + 2[t_k/\|s_k\|][f(y) - f(x_k)].$$

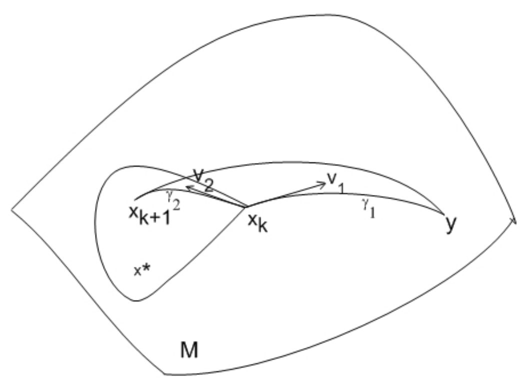

Figura 3.3: Lema 3.1

- Consideremos el conjunto:

$$\theta = \{z \in M / f(z) \leq \inf_k f(x_k)\}.$$

Corolario 3.1. *Sea $\{x_k\}$ la sucesión generada por algoritmo y sea $K \geq 0$ la curvatura seccional de M. Para todo $z \in \theta$, tenemos:*

$$d^2(x_{k+1}, z) \leq d^2(x_k, z) + t_k^2 \;\; \forall k \in \mathbb{N}.$$

Demostración: Sea γ_{v_1} una geodésica parametrizada por longitud de arco, donde $\gamma_{v_1}(0) = x_k$, $\gamma_{v_1}(t_1) = z$, $t_1 = d(x_k, z)$ y $z \in \theta$. Ahora, por el algoritmo podemos considerar otra geodésica γ_{v_2} tal que $v_2 = v_k = -s_k/\|s_k\|$, $\gamma_{v_2}(0) = x_k$ y $\gamma_{v_2}(t_k) = x_{k+1}$ y $t_k = t_2$.

Luego aplicando el corolario 1,1 tenemos que:

$$\begin{aligned} d_2(\gamma_{v_1}(t_1), \gamma_{v_k}(t_k)) &\leq \|t_k v_k - t_1 v_1\|^2 \\ &\leq \| - t_k[s_k/\|s_k\|] - t_1 v_1\| \\ &= t_1^2 + t_k^2 + 2[t_k/\|s_k\|]\langle s_k, t_1 v_1\rangle \\ &\leq d^2(x_k, y) + t_k^2 + 2[t_k/\|s_k\|][f(z) - f(x_k)] \end{aligned}$$

Pero dado que $z \in \theta$ se tiene entonces que $f(z) - f(x) \leq 0$ así:

$$d_2(\gamma_{v_1}(t_1), \gamma_{v_k}(t_k)) \leq d^2(x_k, z) + t_k^2.$$

Teorema 3.2. *Sea $x_* \in \theta^* \neq \emptyset$ y sea $\{x_k\}$ la sucesión generada por el algoritmo. Si M tiene curvatura $K \geq 0$ y $x_k \notin \theta^*$, entonces:*

$$d(x_{k+1}, x_*) < d(x_k, x_*).$$

Para todo

$$0 < t_k < [2/\|s_k\|][f(x_k) - f(x_*)].$$

Demostración: En el lema 3.1, hacemos $y = x_*$ entonces:

$$d^2(x_{k+1}, x_*) \leq d^2(x_k, x_*) + 2[t_k/\|s_k\|][f(x_*) - f(x_k)].$$

Como $x_* \neq x_k$, se sigue que, para todo

$$0 < t_k < [2/\|s_k\|][f(x_k) - f(x_*)].$$

Se tiene que:

$$\begin{aligned} t_k - [2/\|s_k\|][f(x_k) - f(x_*)] &< 0 \\ t_k + [2/\|s_k\|][f(x_*) - f(x_k)] &< 0 \\ t_k(t_k + [2/\|s_k\|][f(x_k) - f(x_*)]) &< 0 \end{aligned}$$

Lo que implica que:

$$t_k^2 + 2[t_k/\|s_k\|][f(x_*) - f(x_k)] < 0.$$

Lo cual prueba el teorema.

3.3. Convergencia del algoritmo

Teorema 3.3. *Sea $\{x_k\}$ la sucesión generada por algoritmo y sea $K \geq 0$ la curvatura seccional de M. Si la sucesión $\{t_k\}$ se elige para sastifacer:*

$$\sum_{k=0}^{\infty} t_k = +\infty \tag{3.1}$$

$$\sum_{k=0}^{\infty} t_k^2 < +\infty. \tag{3.2}$$

Entonces:

$$\liminf_{k} f(x_k) = f^* \tag{3.3}$$

Además de, si $\theta^* \neq \emptyset$, entonces la sucesión $\{x_k\}$ converge al punto $x_* \in \theta^*$.

Demostración:

Por contradicción, supongamos que:

$$\liminf_{k} f(x_k) > f^*.$$

Por un lado, esto implica que $\theta \neq \emptyset$, donde $\theta = \{z \in M / f(z) \leq \text{ínf}_k f(\{x_k\})\}$ así, por 3.2 y por el corolario 3.1 $\{x_k\}$ es casi-Féjer convergente a θ, donde $\epsilon_k = t_k^2$, por tanto, $\{x_k\}$ es acotada y consecuentemente $\{s_k\}$, donde $\{s_k\} \in \partial f(x_k)$ tambien es acotado.

Digamos $\|s_k\| < C_0$, $\forall k \in \mathbb{N}$, donde $C_0 > 0$.

Afirmación: Dado $z \in \theta$ existe $C_1 > 0$ y $k_0 \in \mathbb{N}$ tal que:

$$f(z) < f(x_k) - C_1 \;\; \forall k > k_0.$$

En efecto: Procedamos por reducción al absurdo: Así, dado $z \in \theta$, se tiene que para todo $C_1, k_0 > 0$, existe $k > k_0$, tal que:

$$f(z) \geq f(x_k) - C_1.$$

$$\text{Entonces } f(z) \geq f(x_k)$$

para algún $k > k_0$ lo cual contradice la definición de θ. Así:

$$f(z) < f(x_k) - C_1 \;\; \forall k > k_0$$

Para este z, tenemos por lema 3.1 que:

$$d^2(x_{k+1}, z) \leq d^2(x_k, z) + t_k^2 + 2[t_k / \|s_k\|][f(z) - f(x_k)] \tag{3.4}$$

Donde tenemos que:

$$f(z) < f(x_k) - C_1 \Rightarrow f(z) - f(x_k) < -C_1.$$

$$\|s_k\| < C_0 \Rightarrow \frac{1}{\|s_k\|} > \frac{1}{C_0} \Rightarrow \frac{-1}{\|s_k\|} < \frac{-1}{C_0}.$$

Lo que implica que:

$$\begin{aligned} 2\frac{t_k}{\|s_k\|}[f(z) - f(\{x_k\})] &< \frac{2t_k}{\|s_k\|}(-C_1) \\ &= \frac{-2t_k}{\|s_k\|}C_1 < -2\frac{C_1}{C_0}t_k \end{aligned}$$

Así, 3.4 se transforma en:

$$d^2(x_{k+1}, z) \leq d^2(x_k, z) - t_k \frac{C_1}{C_0} \ \ \forall k > k_0 \tag{3.5}$$

Si despejamos t_k nos queda:

$$t_k \leq (d^2(x_k, z) - d^2(x_{k+1}, z))(\frac{C_1}{C_0}).$$

Así, adicionando tenemos que:

$$\sum_{t_j}^{l+k} t_j \leq (\frac{C_1}{C_0})(d^2(x_{k_0}, z) - d^2(x_{l+k_0}, z)) \ \ \forall l.$$

Lo cual contradice la hipotesis 3.4 ya que $d(x_{l+k_0}, z)$ es acotada. Luego, se tiene de la primera parte que $\{f(x_k)\}$ posee una subsucesión monótona decreciente $\{f(x_{k_j})\}$ tal que:

$$\lim_{j\to\infty} f(x_{k_j}) = f^*.$$

Sin perdida de generalidad supongamos que la sucesión $\{f(x_k)\}$ es decreciente, monótona y converge a f^*. Ahora, siendo la sucesión $\{x_k\}$ limitada entonces posee una sub-sucesión convergente $\{x_{k_j}\}$. Digamos:

$$\lim_{j\to\infty} x_{k_j} = x^*.$$

Así, por la continuidad de f tenemos:

$$f(x^*) = \lim_{j\to\infty} f(x_{k_j}) = f^*.$$

Y por lo tanto $x_* \in \theta$. Así, $\{x_k\}$ tiene un punto de acumulación $x_* \in \theta$, y como $\{x_k\}$ es casi-Féjer convergente a θ, tenemos por el teorema 1.9 que la sucesión $\{x_k\}$ converge x^*.

Conclusión

En este trabajo se observa la necesidad de considerar la curvatura seccional de la variedad, debido al uso importante del corolario $1,1$ en las pruebas presentadas, el cual se usa solo en curvatura seccional $K \geq 0$. Además se considera el hecho importante de que las geodésicas en presencia de curvatura seccional positiva tienden a crecer mas cerca lo cual hace posible creer que se puede mejorar la estimación de $0 < t_k < [2/\|s_k\|][f(x_k) - f(x_*)]$. Este algoritmo resuelve el problema

$$\min_{x \in M} f(x),$$

donde M es una variedad Riemanniana conexa y completa con curvatura seccional $K \geq 0$ y f una función convexa de valores reales definida en M no necesariamente diferenciable. La prueba de la convergencia sin una hipótesis acerca de la curvatura se mantiene abierto.

Bibliografía

[1] O.P. Ferreira and P.R. Oliveira. Subgradient algorithm on riemannian manifolds. *Journal of Optimization Theory and Applications*, 97(1):93–104, April 1998.

[2] D. M.P. doCarmo. *Riemannian Geometry*. Birkhauser, Boston, Nashua, USA, 1992.