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

# Tópicos de Optimización sobre Variedades de Riemann.

Br. Waiberh José Acosta Aldana.
Tutor: Dr. Rómulo Castillo.
Area del 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 titulo de Licenciado en Ciencias Matematicas

Br. Waiberh José Acosta Aldana.
Tutor: Dr. Rómulo Castillo.

Barquisimeto, Marzo 2007

# Agradecimiento.

Primeramente a Dios Todopoderoso, que me acompaña en todo momento y me da vida y salud para culminar las cosas que me propongo. A mis padres, Rita E Aldana y Alfredo J Acosta por darme la vida y el apoyo necesario para culminar mis estudios y cada meta que me trazo. A mis hermanos, por siempre creer en mi. A la UCLA por darme la oportunidad de estudiar. A Dr. Rómulo Castillo por ser amigo, por la valiosa asesoría y su apoyo para el desarrollo de este trabajo. A todos aquellos profesores que de una u otra manera colaboraron en forma directa a mi formación profesional. A Dianmary Rojas por entenderme en todo momento y siempre apoyarme en mis estudios. A la mejor compañera de estudio que he tenido Virginia Morales gracias por ayudarme en todo momento de manera incondicional. A todos los compañeros de la Asociación de Estudiantes de Matemáticas.

# Dedicatoria.

A Dios todopoderoso, por haberme dado la vida, y hacer posible que termine este trabajo. A mis padres que siempre confiaron en mi, mi tia Maribel y en especial a mi abuela María Vargas.

# Resumen.

En el presente trabajo se hará un estudio introductorio de las principales definiciones y teoremas sobre algunos tópicos de optimización en variedades Riemannianas, pasando por una revisión previa de conceptos precedentes de la geometría diferencial. Todo ello con el fin de abordar algunos tópicos que incluye el método de Cauchy sobre la esfera unitaria y el cálculo de geodésicas.

Esto servirá como base de futuras investigaciones en el área en donde se aplican otras metodologías.

# Índice

# Capítulo 1

# Introducción.

En la programación matemática se tiene inicialmente como principal objetivo el problema de minimizar un funcional sobre un conjunto. En primer plano se aborda el problema en $R^n$, sin embargo, es usual extender el campo de acción a espacios mas generales, como por ejemplo, en nuestro caso, se considerará una aplicación sobre variedades Riemannianas.

En el presente trabajo hacemos un estudio introductorio de las principales definiciones y teoremas sobre algunos tópicos de optimización en variedades Riemannianas, pasando por una revisión previa de conceptos procedentes de la geometría diferencial y que pueden encontrarse en las referencias [2] y [4]. Por otra parte, se estudian y desarrollan extensiones de estos conceptos en variedades Riemannianas (ver [4]) y se espera analizar y comprender aspectos relacionados con el cálculo de geodésicas, establecer la ecuación diferencial que las rigen, finalmente estudiamos teoremas de convergencia de el método de Cauchy sobre variedades Riemannianas y lo ejemplificamos sobre la esfera unitaria.

Como antecedentes generales a este problema existen diversos trabajos que contemplan estudios sobre cálculo de geodésicas y teoría de minimización en variedades Riemannianas para los cuales la referencia [1] representa un compendio de trabajos diversos sobre el tema estudiado y constituye la fuente básica del desarrollo final del presente trabajo.

Esto servirá como base de futuras investigaciones en el área en donde se aplican otras metodologías ( ver [5], [6]).
Se espera escribir una monografía contentiva de algunos temas de la matemática en donde se detallan los principales resultados.

En los capítulos 2 y 3 se analizan y desarrollan algunos tópicos de la geometría Euclidiana y geometría diferencial como basamento teórico para comprender dichos conceptos en espacios mas generales como las variedades Rienmanianas. Luego en los capítulos 4 y 5 sabiendo que las geodésicas son las curvas con aceleración nula, vamos a deducir el sistema de ecuaciones diferenciales de segundo grado

$$\frac{d^2x^i}{dt^2} + \sum_{j,k=1}^{n} \Gamma^i_{jk} \frac{dx^j}{dt}\frac{dx^k}{dt} = 0, \quad i = 1, ..., n,$$

Donde $\Gamma^i_{jk}$ son los símbolos de Christoffel y cuyas soluciones serán las geodésicas de cada variedad, también definiremos el mapeo exponencial como un mapeo que va desde el espacio tangente $T_pM$ a la variedad $M$

$$X \in T_pM \rightarrow \gamma_X(1) \in M.$$

Teniendo en cuenta que frecuentemente en las matemáticas se trata de extender el campo de trabajo vamos a ver como aplicamos el método de cauchy de $\mathbb{R}^n$ o del gradiente, cuya dirección de descenso esta dado por $d = -\nabla f(x)$, y cuyo proceso iterativo esta dado por

$$x_{i+1} = x_i + \lambda d,$$

para minimizar una función sobre una variedad Riemanniana, específicamente sobre la esfera, usando la dirección de descenso dada por $-t_i gradf(x_i)$ para obtener el proceso iterativo

$$x_{i+1} = exp_{x_i}(-t_i gradf(x_i)).$$

# Capítulo 2

# Preliminares

En este capitulo se muestran algunos conceptos básicos de la geometría Euclídea que abarca desde la definición de vector tangente , derivadas, superficie regulares, métrica y algunas otras consideraciones, todo sobre $\mathbb{R}^n$.

El objetivo del mismo es construir una plataforma básica de conceptos que permitan una mejor comprensión de los contenidos de los capítulos siguientes en los cuales se discutirán muchos de estos conceptos aquí mostrados pero en un ambiente mas general como son las variedades diferenciales y posteriormente las variedades de Riemann.

Tambien es importante destacar que las Definiciones, Teoremas, Lemas y Corolarios de este capitulo fueron obtenidos de [2]

## 2.1. Aplicaciones Tangentes y Derivadas.

**Definición 2.1.1.** *Un vector tangente $v_p$ en el espacio euclídeo $\mathbb{R}^n$ consiste en un par de puntos $v, p \in \mathbb{R}^n$; $v$ se denomina **la parte vectorial** y $p$ el **punto de aplicación** de $v_p$.*

**Definición 2.1.2.** *Sea $p \in \mathbb{R}^n$. El **espacio tangente** a $\mathbb{R}^n$ en $p$ es el conjunto*

$$\mathbb{R}^n_p = \{v_p / v \in \mathbb{R}^n\}.$$

*El espacio tangente $\mathbb{R}^n_p$ es una copia del espacio vectorial $\mathbb{R}^n$; de hecho, existe una biyección canónica entre $\mathbb{R}^n_p$ y $\mathbb{R}^n$ dada por*

$$v_p \to v.$$

En la figura 2.1 se observa que un vector tangente $v_p$ a $\mathbb{R}^n$ en $p$ no es mas que el mismo vector $v$ trasladado a el punto $p$.

Además 2 vectores tangentes $v_p$ y $w_q$ son iguales si y sólo si sus partes vectoriales y sus puntos de aplicación coinciden, es decir si $v = w$ y $p = q$.

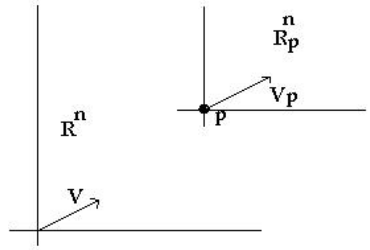



Figura 2.1: Vector Tangente

Para $v_p, w_p \in \mathbb{R}^n_p$.

Se define sobre $\mathbb{R}^n_p$ el producto escalar como:

$$v_p \cdot w_p = v \cdot w.$$

**Definición 2.1.3.** *Sea $f : U \subset \mathbb{R}^n \to \mathbb{R}$ una función diferenciable, y sea $v_p$ un vector tangente a $\mathbb{R}^n$ en $p \in U$.*

*Definimos*

$$v_p[f] = \frac{d}{dt}(f(p + tv))|_{t=0}. \tag{2.1}$$

*La ecuación 2.1 se conoce con el nombre de **Derivada Direccional de $f$ en la dirección de** $v_p$.*

*Otra forma de escribir 2.1 es*

$$v_p[f] = \lim_{t \to 0} \frac{f(p + tv) - f(p)}{t}.$$

**Lema 2.1.1.** *Sea $v_p = (v_1, ..., v_n)$ un vector tangente en un punto $p \in U \subset \mathbb{R}^n$,y sea $f : U \subset \mathbb{R}^n \to \mathbb{R}$ una función diferenciable. Entonces*

$$v_p[f] = \sum_{j=1}^{n} v_j \frac{\partial f}{\partial u_j}(p). \tag{2.2}$$

**Lema 2.1.2.** *Sean $f, g : U \subset \mathbb{R}^n \to \mathbb{R}$ funciones diferenciables, $v_p, w_p$ vectores tangentes a $\mathbb{R}^n$ en $p \in \mathbb{R}^n$, y $a, b$ números reales. Entonces*

$$(av_p + bw_p)[f] = av_p[f] + bw_p[f], \tag{2.3}$$

$$v_p[af + bg] = av_p[f] + bv_p[g], \tag{2.4}$$

$$v_p[fg] = f(p)v_p[g] + g(p)v_p[f]. \tag{2.5}$$

**Lema 2.1.3.** ***Regla de la Cadena.*** *Sean $g_1, ..., g_k : \mathbb{R}^n \to \mathbb{R}$ y $F : \mathbb{R}^k \to \mathbb{R}$ funciones diferenciables y sea $v_p \in \mathbb{R}^n_p$ con $p \in \mathbb{R}^n$. Escribamos $f = F(g_1, ..., g_k)$, Entonces*

$$v_p[f] = \sum_{j=1}^{k} \frac{\partial F}{\partial u_j}(g_1(p), ..., g_k(p))v_p[g_j].$$

Obsérvese que cualquier aplicación $F : \mathbb{R}^n \to \mathbb{R}^m$ puede ser escrita como

$$F(p) = (f_1(p), ..., f_m(p)), \tag{2.6}$$

donde $f_j : \mathbb{R}^n \to \mathbb{R}$ para cada $j = 1, ..., m$. Por conveniencia, utilizaremos la notación

$$F = (f_1, ..., f_m),$$

como una forma abreviada de 2.6.Ademas $U \subset \mathbb{R}^n$ se dice **abierto** si para $p \in U$ existe $\epsilon > 0$ tal que la bola abierta $\{q \in \mathbb{R}^m / \|q-p\| < \epsilon\}$ esta totalmente contenida en $U$.

**Definición 2.1.4.** *Sea $U \subset \mathbb{R}^n$ un subconjunto abierto. Una aplicación $F : U \to \mathbb{R}^m$ se dice **diferenciable** si cada $f_j$ es diferenciable. En general, Si $A$ es un subconjunto de $\mathbb{R}^n$, se dice que $F$ es diferenciable en $A$ si existe un conjunto abierto $U$ tal que $A \subset U$ y una aplicación diferenciable $\widetilde{F} : U \to \mathbb{R}^m$ tal que la restricción de $\widetilde{F}$ a $A$ es $F$.*

La Definicion anterior muestra como definir la diferenciabilidad sobre cualquier conjunto sin importar si este es abierto o cerrado.

**Definición 2.1.5.** *Sean $U, V \subset \mathbb{R}^n$, $F : U \to V$ es un **difeomorfismo** si posee inversa diferenciable $F^{-1} : V \to U$.*

**Lema 2.1.4.** *Sean $U \subset \mathbb{R}^1, V \subset \mathbb{R}^m$ y sean $F : U \to V$, $G : V \to \mathbb{R}^n$ aplicaciones diferenciables. Entonces la composición $G \circ F : U \to \mathbb{R}^n$ es diferenciable.*

**Definición 2.1.6.** *Sea $F : U \to \mathbb{R}^m$ una aplicación diferenciable, siendo $U$ un subconjunto abierto de $\mathbb{R}^n$, y sea $p \in U$. Para cada vector $v_p$ tangente a $\mathbb{R}^n$, sea $F_*(v_p)$ el vector velocidad inicial de la curva $t \mapsto F(p + tv)$, es decir,*

$$F_*(v_p) = F(p + tv)'(0).$$

*La aplicación $F_* : \mathbb{R}^n_p \to \mathbb{R}^m_{F(p)}$ así definida se denomina **aplicación tangente** de la aplicación diferenciable $F$ en el punto $p$.*

$p + tv$ es una recta en $\mathbb{R}^n$ y su imagen mediante $F$ es una curva en $\mathbb{R}^m$, ahora bien la derivada $F'$ a lo largo de la curva $p + tv$ es el vector velocidad de la curva $F(p + tv)$, ese vector velocidad lo denotaremos por $F_*(v_p)$ y es lo que llamaremos aplicacion tangente (ver figura 2.2).

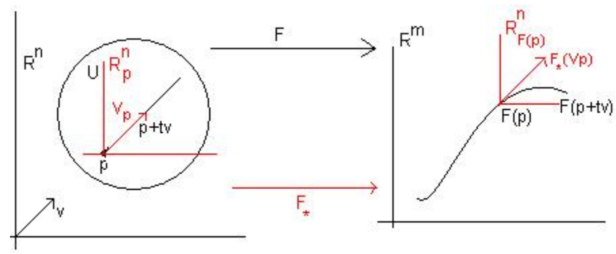



Figura 2.2: Aplicación Tangente

**Lema 2.1.5.** *Sea $F : U \to \mathbb{R}^m$ una aplicación diferenciable, siendo $U$ un subconjunto abierto de $\mathbb{R}^n$, y escribamos $F = (f_1, ..., f_m)$. Si $p \in U$ y $v_p$ tangente a $\mathbb{R}^n$ en $p$, entonces*

$$F_*(v_p) = (v_p[f_1], ..., v_p[f_m])_{F(p)}. \tag{2.7}$$

**Corolario 2.1.1.** *Si $F : U \to \mathbb{R}^m$ es una aplicación diferenciable, siendo $U$ un subconjunto abierto de $\mathbb{R}^n$, entonces en cada punto $p \in U$ la aplicación tangente*

$$F_* : \mathbb{R}^n_p \to \mathbb{R}^m_{F(p)},$$

*es una aplicación lineal.*

**Lema 2.1.6.** *Sea $\alpha : (a, b) \to \mathbb{R}^n$ una curva. Entonces, para $p \in (a, b)$ se verifica*

$$\alpha_*(1_p) = \alpha'(p).$$

**Lema 2.1.7.** *Sea $F : U \to \mathbb{R}^m$ una aplicación diferenciable, siendo $U$ un subconjunto abierto de $\mathbb{R}^n$, y sea $\alpha : (a, b) \to U$ una curva. Entonces*

$$(F \circ \alpha)'(t) = F_*(\alpha'(t)),$$

*para $a < t < b$.*

**Corolario 2.1.2.** *Sea $F : U \to \mathbb{R}^m$ una aplicación diferenciable, siendo $U$ un subconjunto abierto de $\mathbb{R}^n$, y sea $p \in U$. Sean $v_p \in \mathbb{R}^n_p$ y $\alpha : (a, b) \to U$ una curva tal que $\alpha'(0) = v_p$. Entonces*

$$F_*(v_p) = (F \circ \alpha)'(0).$$

**Corolario 2.1.3.** *Sean $F : U \subset \mathbb{R}^1 \to V \subset \mathbb{R}^m$ y $G : V \to \mathbb{R}^n$ aplicaciones diferenciales. Entonces*

$$(G \circ F)_* = G_* \circ F_*.$$

**Definición 2.1.7.** *Sea $F : U \to \mathbb{R}^m$ una aplicación diferenciable, siendo $U \subset \mathbb{R}^n$ abierto.Escribamos $F = (f_1, ..., f_m)$. La* ***Matriz Jacobiana*** *de $F$ es la función con valores matriciales $\jmath(F)$ dada por*

$$\jmath(F)(p_1, ..., p_n) = \begin{pmatrix} \frac{\partial f_1}{\partial u_1}(p_1, ..., p_n) & ... & \frac{\partial f_m}{\partial u_1}(p_1, ..., p_n) \\ : & & : \\ \frac{\partial f_1}{\partial u_n}(p_1, ..., p_n) & ... & \frac{\partial f_m}{\partial u_n}(p_1, ..., p_n) \end{pmatrix},$$

*para cada $(p_1, ..., p_n) \in U$.*

**Definición 2.1.8.** *Un* ***campo de vectores*** *$V$ sobre un subconjunto $U \in \mathbb{R}^n$ es una aplicación que a cada $p \in U$ le asigna un vector tangente $V(p) \in \mathbb{R}^n_p$. Si $f : U \to \mathbb{R}$ es una función diferenciable, hacemos actuar $V$ sobre $f$ mediante la fórmula*

$$V[f](p) = V(p)[f].$$

*El campo de vectores $V$ se dice* ***diferenciable*** *si la función $V[f] : U \to \mathbb{R}$ es diferenciable.*

**Definición 2.1.9.** *Las* ***funciones coordenadas naturales*** *de $\mathbb{R}^n$ son las funciones $u_1, ..., u_n$ dadas por*

$$u_i(p) = p_i,$$

*para cada $p = (p_1, ..., p_n)$.*

**Definición 2.1.10.** *El campo de vectores $U_j$ sobre $\mathbb{R}^n$ esta dado por*

$$U_j(p) = (0, .., 1, .., 0)_p,$$

*donde $1$ esta en la posición $j$. Llamaremos a $U_1, ..., U_n$ la* ***referencia natural de campos de vectores*** *sobre $\mathbb{R}^n$.*

Claramente los vectores $U_i(p)$ forman una base para el espacio tangente $\mathbb{R}^n_p$. Asi cualquier vector en $\mathbb{R}^n$ puede ser escrito como combinacion lineal de los vetores $U_i$ y cualquier vector en $\mathbb{R}^n_p$ puede ser escrito com combinacion lineal de los vectores $U_i(p)$.

**Lema 2.1.8.** *Sean $X$ e $Y$ campos de vectores sobre un subconjunto abierto $U \in \mathbb{R}^n$, y sea $a \in \mathbb{R}$ y $g : \mathbb{R}^n \to \mathbb{R}$ una función diferenciable. Definamos $X+Y$,$aX$,$gX$ y $X \cdot Y$ por:*

$$(X+Y)[f] = X[f] + Y[f], \qquad (aX)[f] = aX[f],$$
$$(X \cdot Y)(p) = X(p) \cdot Y(p), \qquad (gX)[f] = gX[f],$$

*donde $f : U \to \mathbb{R}$ es diferenciable y $p \in U$.*

*Así, $X+Y$,$aX$,$gX$ son campos de vectores $U$ y $X \cdot Y$ es una función.*

**Lema 2.1.9.** *Si $V$ es un campo de vectores arbitrarios sobre un subconjunto abierto $U \in \mathbb{R}^n$, entonces existen funciones $v_j : U \to \mathbb{R}$ con $j = 1...n$ tales que*

$$V = \sum_{j=1}^{n} v_j U_j, \tag{2.8}$$

*además diremos que $V$ es diferenciable si y sólo si las funciones $v_1, ..., v_n$ asociadas a $V$ dadas por 2.8 son diferenciables.*

**Definición 2.1.11.** *Sea $g : \mathbb{R}^n \to \mathbb{R}$ una función diferenciable. El **gradiente** de $g$ es el campo de vectores $\boldsymbol{grad}g$ dado por*

$$\boldsymbol{grad}g = \sum_{j=1}^{n} \frac{\partial g}{\partial u_j} U_j.$$

De esta manera el $\mathbf{grad}g = (\frac{\partial g}{\partial u_i}, ..., \frac{\partial g}{\partial u_n})$.

**Lema 2.1.10.** *El gradiente $\boldsymbol{grad}g$ de una función $g$ es el único campo de vectores diferenciable sobre $\mathbb{R}^n$ tal que*

$$\boldsymbol{grad}g \cdot V = V[g], \tag{2.9}$$

*para cualquier campo de vectores diferenciable $V$ sobre $\mathbb{R}^n$.*

**Definición 2.1.12.** *Sea $W$ un campo de vectores diferenciable sobre un subconjunto abierto $U \subset \mathbb{R}^n$, y sea $v_p$ un vector tangente a $\mathbb{R}^n$ en $p \in U$. Entonces la derivada de $W$ con respecto a $v_p$ es el vector tangente $D_v W \in \mathbb{R}^n_p$ dado por:*

$$D_v W = \widetilde{W}(p + tv)'(0)_p = \lim_{t \to 0} \frac{\widetilde{W}(p + tv) - \widetilde{W}(p)}{t}|_p,$$

*Donde $\widetilde{W}$ es el campo de vectores sobre la curva $p + tv$*

**Lema 2.1.11.** *Sea $W$ un campo de vectores diferenciable sobre un subconjunto abierto $U \subset \mathbb{R}^n$, y pongamos*

$$W = \sum_{i=1}^{n} w_i U_i.$$

*Entonces, para $p \in U$ y $v_p \in \mathbb{R}^n_p$, se verifica*

$$D_v W = \sum_{i=1}^{n} v_p[w_i] U_i(p) = (v_p[w_1], ..., v_p[w_n]). \tag{2.10}$$

**Corolario 2.1.4.** *Para cualquier vector $v_p$ tangente a $\mathbb{R}^n$ se verifica*

$$D_v U_j = 0, \tag{2.11}$$

*para cada $j = 1...n$.*

**Lema 2.1.12.** *Sean $X, Y$ campos de vectores diferenciables sobre un subconjunto abierto $U$ de $\mathbb{R}^n$, y sea $v_p, w_p$ vectores tangentes a $\mathbb{R}^n$ en $p \in U$. Entonces*

$$D_{av+bw} Y = aD_v Y + bD_w Y, \tag{2.12}$$

$$D_v(aY + bX) = aD_v Y + bD_v X, \tag{2.13}$$

$$D_v(fY) = v_p[f]Y(p) + f(p)D_v Y, \tag{2.14}$$

$$v_p[Y \cdot X] = D_v Y \cdot X(p) + Y(p) \cdot D_v X, \tag{2.15}$$

*siendo $a, b \in \mathbb{R}$ y $f$ diferenciable.*

**Lema 2.1.13.** *Sean $V$ y $W$ campos de vectores definidos sobre un subconjunto abierto $U \subset \mathbb{R}^n$, y supongamos que $W$ es diferenciable. Definamos $D_V W$ por*

$$(D_V W)(p) = D_{V(p)} W,$$

*para cada $p \in U$. Entonces $D_V W$ es un campo de vectores sobre $U$ que es diferenciable si $V$ y $W$ son ambos diferenciables.*
*Además, si*

$$W = \sum_{i=1}^{n} w_i U_i,$$

*entonces*

$$D_V W = \sum_{i=1}^{n} V[w_1] U_i. \tag{2.16}$$

La demostacion del lema anterior puede obtenerse usando el lema 2,1,12.

**Definición 2.1.13.** *Sea $F : U \to V$ un difeomorfismo entre 2 subconjuntos abiertos $U$ y $V$ de $\mathbb{R}^n$, y sea $V$ un campo de vectores diferenciables sobre $U$. Escribamos $F = (f_1, ..., f_n)$. Entonces $F_*(V)$ es el campo de vectores sobre $V$ dado por*

$$F_*(V) = \sum_{i=1}^{n} (V[f_i] \circ F^{-1}) U_i. \tag{2.17}$$

**Lema 2.1.14.** *Sea $F : U \to V$ un difeomorfismo entre dos subconjuntos abiertos $U$ y $V$ de $\mathbb{R}^n$, y sean $Y, Z$ campos de vectores diferenciables sobre $U$. Entonces*

$$F_*(aY + bZ) = aF_*(Y) + bF_*(Z), \tag{2.18}$$

$$F_*(fY) = (f \circ F^{-1})F_*(Y), \tag{2.19}$$

$$F_*(Y)[f \circ F^{-1}] = Y[f] \circ F^{-1}, \tag{2.20}$$

*siendo $a, b \in \mathbb{R}$ y $f : U \to \mathbb{R}$ una función diferenciable.*

**Lema 2.1.15.** *Sea $F : \mathbb{R}^n \to \mathbb{R}^n$ una transformación afín. Entonces*

$$D_{F_*(V)}(F_*(W)) = F_*(D_V W). \tag{2.21}$$

**Definición 2.1.14.** *El **campo de vectores velocidad** de una curva $\alpha$ es la aplicación*

$$t \to \alpha'(t)_{\alpha(t)},$$

*y la **aceleración** de $\alpha$ es la aplicación*

$$t \to \alpha''(t)_{\alpha(t)},$$

*puesto que $\alpha'(t)_{\alpha(t)}$ es el vector tangente en $\mathbb{R}^n$, es posible aplicarlo a una función diferenciable arbitraria $f : \mathbb{R}^n \to \mathbb{R}$.*

**Lema 2.1.16.** *Sean $\alpha : (a, b) \to \mathbb{R}^n$ una curva y $f : \mathbb{R}^n \to \mathbb{R}$ una función diferenciable. Entonces*

$$\alpha'(t)_{\alpha(t)}[f] = (f \circ \alpha)'(t).$$

**Lema 2.1.17.** *Sean $\alpha : (a, b) \to \mathbb{R}^n$ una curva y $W$ un campo de vectores diferenciable sobre $\mathbb{R}^n$. Entonces*

$$D_{\alpha'(t)} W = (W \circ \alpha)'(t)_{\alpha(t)}.$$

## 2.2. Superficies Parametrizadas en $\mathbb{R}^n$ y Métricas

**Definición 2.2.1.** *Una **superficie parametrizada o parametrización local** (carta local) es una aplicación diferenciable*

$$x : U \to \mathbb{R}^n,$$

*siendo $U \subset \mathbb{R}^2$ (abierto). En general, si $A$ es un subconjunto arbitrario de $\mathbb{R}^2$, se dice que una aplicación $x : A \to \mathbb{R}^n$ es una superficie parametrizada si $x$ puede ser extendida a una aplicación diferenciable de $U$ en $\mathbb{R}^n$, siendo $U$ un subconjunto abierto que contiene a $A$. La **traza** de $x$ es el conjunto imagen $x(U)$.*

Puesto que una superficie parametrizada se puede escribir como una $n - upla$ de funciones, en la forma

$$x(u,v) = (x_1(u,v), ..., x_n(u,v)), \tag{2.22}$$

es posible definir la derivada parcial $x_u$ de $x$ con respecto a $u$

$$x_u(u,v) = (\frac{\partial x_1}{\partial u}(u,v), ..., \frac{\partial x_n}{\partial u}(u,v)), \tag{2.23}$$

y la derivada parcial de $x$ con respecto a $v$

$$x_v(u,v) = (\frac{\partial x_1}{\partial v}(u,v), ..., \frac{\partial x_n}{\partial v}(u,v)), \tag{2.24}$$

de igual manera se define las derivadas $x_{uv}, x_{uu}, etc.$

**Lema 2.2.1.** *Sea $x : U \to \mathbb{R}^n$ una superficie parametrizada, y sea $q \in U$. Entonces*

$$x_*(\frac{\partial}{\partial u}\mid_q) = x_u(q) \quad y \quad x_*(\frac{\partial}{\partial v}\mid_q) = x_v(q),$$

*donde $x_*$ representa la aplicación tangente de $x$.*

**Definición 2.2.2.** *La **matriz Jacobiana** de una superficie parametrizada $x : U \to \mathbb{R}^n$ es la función con valores matriciales $\jmath(x)$ dada por*

$$\jmath(x)(u,v) = \begin{pmatrix} \frac{\partial x_1}{\partial u}(u,v) & ... & \frac{\partial x_n}{\partial u}(u,v) \\ \frac{\partial x_1}{\partial v}(u,v) & ... & \frac{\partial x_n}{\partial v}(u,v) \end{pmatrix} = \begin{pmatrix} x_u(u,v) \\ x_v(u,v) \end{pmatrix},$$

*para cada $(u,v) \in U$.*

**Corolario 2.2.1.** *Sea $x : U \to \mathbb{R}^n$ una superficie parametrizada. Entonces las condiciones siguientes son equivalentes*

1. *$x_u(u_0,v_0)$ y $x_v(u_0,v_0)$ son linealmente dependientes;*
2. *det $\begin{pmatrix} x_u \cdot x_u & x_u \cdot x_v \\ x_v \cdot x_u & x_v \cdot x_v \end{pmatrix}$ se anula en $(u_0,v_0)$;*
3. *la matriz jacobiana $\jmath(x)$ tiene rango menor que 2 en $(u_0,v_0)$;*

**Definición 2.2.3.** *Una **superficie parametrizada regular** es una superficie parametrizada $x : U \to \mathbb{R}^n$ cuya matriz Jacobiana $\jmath(x)(u,v)$ tiene rango 2 en todo punto $(u,v) \in U$. Una **Superficie parametrizada inyectiva** es una superficie parametrizada tal que $x(u_1,v_1) = x(u_2,v_2)$ implica que $u_1 = u_2$ y $v_1 = v_2$.*

**Lema 2.2.2.** *Sea $x : U \to \mathbb{R}^n$ una superficie parametrizada regular y sea $q \in U$.*

*Existe entonces un entorno $U_q$ de $q$ tal que $x$ aplica $U_q$ difeomorficamente sobre $x(U_q)$.*

**Definición 2.2.4.** *Sea $x : U \to \mathbb{R}^n$ una superficie parametrizada inyectiva, y sea $p \in x(U)$. Un* ***vector tangente*** *a $x$ en el punto $p$ es un vector tangente $v_p \in \mathbb{R}^n_p$ para el cual existe una curva $\alpha : (a, b) \to \mathbb{R}^n$, que se puede representar por*

$$\alpha(t) = x(u(t), v(t)), \qquad (a < t < b) \tag{2.25}$$

*tal que $\alpha(0) = p$ y $\alpha'(0) = v_p$. El conjunto de vectores tangentes a $x$ en el punto $p$ lo representamos por $x(U)_p$.*

**Definición 2.2.5.** *Sea $x : U \to \mathbb{R}^n$ una superficie parametrizada inyectiva, y sea $z_p \in \mathbb{R}^n_p$ con $p \in x(U)_p$.*

*Diremos que $z_p$ es* ***normal*** *o* ***perpendicular*** *a $x$ en $p$ si $z_p \cdot v_p = 0$ para todo vector $v_p$ tangentes a $x$ en el punto $p$.*

Si $x(U)^{\perp}_p$ denota el espacio de los vectores en $\mathbb{R}^n_p$ perpendiculares a $x(U)_p$, entonces

$$\mathbb{R}^n_p = x(U)_p \oplus x(U)^{\perp}_p.$$

**Definición 2.2.6.** *Sea $x : U \to \mathbb{R}^n$ una superficie parametrizada. Un* ***campo de vectores*** *$V$ sobre $x$ es una aplicación que asigna a cada punto $q \in U$ un vector tangente $V(q) \in \mathbb{R}^n_p$, siendo $p = x(q)$. Se dice que el campo de vectores $V$ es* ***tangente*** *a $x$ si $V(q) \in x(U)_{x(q)}$ para todo $q \in U$, y se dice que $V$ es* ***normal o perpendicular*** *a $x$ si $V(q) \cdot v_p = 0$ para todo $v_p \in x(U)_{x(q)}$ y todo $q \in U$.*

**Definición 2.2.7.** *Un subconjunto $M \subset \mathbb{R}^n$ es una* ***superficie regular*** *si para cada punto $p \in M$ existe un entorno abierto $V$ de $p$ en $\mathbb{R}^n$ y una aplicación $x : U \to \mathbb{R}^n$ de un subconjunto abierto $U \subset \mathbb{R}^2$ sobre $V \bigcap M$ tal que*

1. *$x$ es diferenciable.*
2. *$x : U \to V \bigcap M$ es un homeomorfismo. Esto significa que $x$ posee una inversa continua $x^{-1} : V \bigcap M \to U$ tal que $x^{-1}$ es la restricción a $V \bigcap M$ de una aplicación continua $F : W \to \mathbb{R}^2$, siendo $W$ un subconjunto abierto de $\mathbb{R}^n$ que contiene a $V \bigcap M$.*
3. *Cada aplicación $x : U \to M$ es una superficie parametrizada regular.*

**Lema 2.2.3.** *Sea $W \subset \mathbb{R}^n$ abierto, y sea $G : W \to \mathbb{R}^n$ una aplicación tal que $G : W \to G(W)$ es un difeomorfismo. Si $M \subset W$ es una superficie regular, entonces $G(M)$ es también una superficie regular.*

**Teorema 2.2.1.** *Sea $M \subset \mathbb{R}^n$una superficie regular, y sea $V \subset \mathbb{R}^n$ abierto, supongamos que $F : V \to \mathbb{R}^n$ una aplicación diferenciable tal que $F(V) \subseteq M$, y que $x : U \to M$ es una carta local de $M$ tal que $F(V) \subseteq x(U)$. Entonces $x^{-1} \circ F : V \to U$ es diferenciable.*

**Corolario 2.2.2.** *Sea $M$ una superficie regular, y sean $x : U \to M$ y $y : V \to M$ dos sistemas de coordenadas locales tales que $x(U) \bigcap y(V) = W$ es no vacía.*
*Entonces el* ***cambio de coordenadas***

$$x^{-1} \circ y : y^{-1}(W) \to x^{-1}(w), \tag{2.26}$$

*es un difeomorfismo entre subconjuntos abiertos de $\mathbb{R}^2$.*

**Definición 2.2.8.** *Sea $f : \mathcal{W} \subset \mathcal{M} \to \mathbb{R}$ una función definida sobre un subconjunto abierto $\mathcal{W}$ de una superficie regular $\mathcal{M}$. Se dice que $f$ es* ***diferenciable*** *en un punto $p \in \mathcal{W}$ si, para alguna carta local $x : U \subset \mathbb{R}^2 \to$ con $p \in x(U) \subset \mathcal{W}$, la composición*

$$f \circ x : U \subset \mathbb{R}^2 \to \mathbb{R},$$

*es diferenciable en $x^{-1}(p)$. Si $f$ es diferenciable en todos los puntos de $\mathcal{W}$ se dice que $f$ es* ***diferenciable*** *sobre $\mathcal{W}$.*

**Definición 2.2.9.** *Sea $\mathcal{M}$ una superficie regular y sea $v \in \mathbb{R}^n$.La* ***función altura*** *y la* ***función cuadrado de la función distancia*** *de $\mathcal{M}$ relativas a $v$ son las funciones $h, f : \mathcal{M} \to \mathbb{R}$ definidas, respectivamente por*

$$h(p) = p \cdot v \qquad y \qquad f(p) = \|p - v\|^2.$$

Esta definicion nos sera util cuando vamos a ejemplificar el algoritmo del metodo de mayor pendiente o metodo del gradiente.

**Definición 2.2.10.** *Sea $\mathcal{M}$ una superficie regular en $\mathbb{R}^n$ y sea $p \in \mathcal{M}$. Diremos que $v_p \in \mathbb{R}^n_p$ es* ***tangente*** *a $\mathcal{M}$ en $p$ si existe una curva $\alpha : (a, b) \to \mathbb{R}^n$ tal que $\alpha(0) = p$,$\alpha'(0) = v_p$ y $\alpha(t)$ esta en $\mathcal{M}$ para $a < t < b$. El* ***espacio tangente*** *a $\mathcal{M}$ en $p$ es el conjunto*

$$\mathcal{M}p = \{v_p \in \mathbb{R}^n_p / v_p \quad \text{es tangente a } M \text{ en } p\}.$$

**Lema 2.2.4.** *Sea $p$ un punto sobre una superficie regular $\mathcal{M} \subset \mathbb{R}^n$. El espacio $\mathcal{M}_p$ tangente a $\mathcal{M}$ en $p$ es un subespacio vectorial de dimensión 2 de $\mathbb{R}^n_p$. Si $x : U \to \mathcal{M}$ es una carta local cualquiera de $\mathcal{M}$ con $p = x(q)$, entonces*

$$x_*(\mathbb{R}^2_q) = \mathcal{M}_p.$$

**Definición 2.2.11.** *Sea $\mathcal{M}$ una superficie regular en $\mathbb{R}^n$ y sea $z_p \in \mathbb{R}^n_p$ con $p \in \mathcal{M}$. Se dice que $z_p$ es* ***normal o perpendicular*** *a $\mathcal{M}$ en $p$, si $z_p \cdot v_p = 0$ para todos los vectores tangentes $v_p \in \mathcal{M}_p$. Representaremos por $\mathcal{M}_p^{\perp}$ el conjunto de todos los vectores normales a $\mathcal{M}$ en $p$. Así $\mathbb{R}^n_p = \mathcal{M}_p \oplus \mathcal{M}_p^{\perp}$.*

**Definición 2.2.12.** *Un* ***campo de vectores*** *$V$ sobre una superficie regular $\mathcal{M}$ es una aplicación que asigna a cada punto $p \in \mathcal{M}$ un vector tangente $V(p) \in \mathbb{R}^n_p$.*

*Se dice que el campo de vectores $V$ es tangente a $\mathcal{M}$ si $V(p) \in \mathcal{M}_p$ para todo $p \in \mathcal{M}$ y se dice que $V$ es* ***normal o perpendicular*** *a $\mathcal{M}$ si $V(p) \in \mathcal{M}_p^{\perp}$ para todo $p \in \mathcal{M}$.*

**Definición 2.2.13.** *Una aplicación $F : \mathcal{M} \to \mathcal{N}$ de una superficie regular $\mathcal{M}$ en otra $\mathcal{N}$ se dice* ***diferenciable*** *si para dos cartas locales cualesquiera, $x$ de $\mathcal{M}$, $y$ de $\mathcal{N}$, la composición $y^{-1} \circ F \circ x$ es diferenciable. Cuando esto suceda diremos simplemente que $F$ es una* ***aplicación entre superficies.***

En la figura 2.3 se observa que la idea es llevar la diferenciabilidad entre superficies, a la diferenciabilidad que tenemos ya definida en espacios Euclídeos. Por ello definimos la diferenciabilidad entre superficies como la diferenciabilidad de la función compuesta $y^{-1} \circ F \circ x$ ya que las aplicaciones $x$,$y$ son diferenciables por su definición.

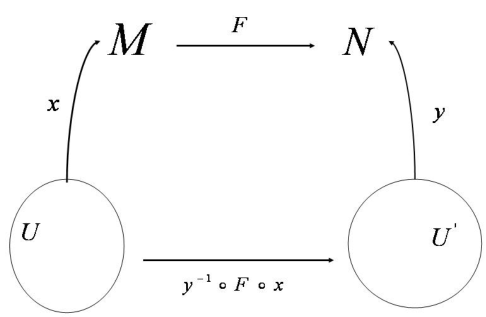



Figura 2.3: Aplicación entre Superficies

**Definición 2.2.14.** *Un* ***difeomorfismo*** *entre dos superficies regulares $\mathcal{M}$ y $\mathcal{N}$ es una aplicación diferenciable $F : \mathcal{M} \to \mathcal{N}$ que posea inversa diferenciable, es decir, existe una aplicación entre superficies $G : \mathcal{N} \to \mathcal{M}$ tal que*

$$G \circ F = 1_{\mathcal{M}} \quad y \quad F \circ G = 1_{\mathcal{N}},$$

*donde $1_{\mathcal{M}}$ y $1_{\mathcal{N}}$ denotan la aplicación identidad de $\mathcal{M}$ y $\mathcal{N}$ respectivamente.*

**Definición 2.2.15.** *Sean $\mathcal{M}$ y $\mathcal{N}$ superficies regulares en $\mathbb{R}^n$ y $p \in \mathcal{M}$. Sea $F : \mathcal{M} \to \mathcal{N}$ una aplicación entre superficies. La **aplicación tangente** de $F$ en el punto $p$ es la aplicación*

$$F_* : \mathcal{M}_p \to \mathcal{N}_{F(p)},$$

*definida como sigue, dado $v_p \in \mathcal{M}_p$ se elige una curva $\alpha : (a,b) \to \mathcal{M}$ tal que $\alpha(0) = p$, $\alpha'(o) = v_p$, se define entonces $F_*(v_p)$ como el vector velocidad inicial de la curva imagen*

$$F \circ \alpha : (a,b) \to \mathcal{N},$$

*es decir,*

$$F_*(v_p) = (F \circ \alpha)'(0).$$

Para describir como se mide una distancia sobre una superficie necesitaremos utilizar una noción , matemáticamente imprecisa, de infinitésimo. La versión infinitesimal (para la distancia en $\mathbb{R}^n$) en este caso n = 2 es

$$ds^2 = dx^2 + dy^2.$$

La ecuación correspondiente para una superficie (o más precisamente para una superficie parametrizada) es

$$ds^2 = Edu^2 + 2Fdudv + Gdv^2.$$

Donde $E, F, G$ son los coeficientes de la primera formula fundamental las cuales definiremos mas adelante y puede verse tambien en [2].

Esta es la versión clásica para representar la **métrica de una superficie**. Podemos considerar esta última ecuación como una versión infinitesimal "deformada"del teorema de Pitágoras.

**Definición 2.2.16.** *Sea $x : U \to \mathbb{R}^n$ una superficie parametrizada.*

*Definimos las funciones $E, F, G : U \to \mathbb{R}$, por*

$$E = \|x_u\|^2, \quad F = x_u \cdot x_v, \quad G = \|x_v\|^2.$$

*Entonces $ds_x^2 = Edu^2 + 2Fdudv + Gdv^2$ es la **métrica Riemanniana** o **primera formula fundamental** de la superficie parametrizada $x$ inducida a partir de $\mathbb{R}^n$.*

**Lema 2.2.5.** *Sea $\alpha : (a,b) \to \mathbb{R}^n$ una curva contenida en una superficie regular parametrizada e inyectiva $x : U \to \mathbb{R}^n$. Entonces la función longitud de arco $s$ de*

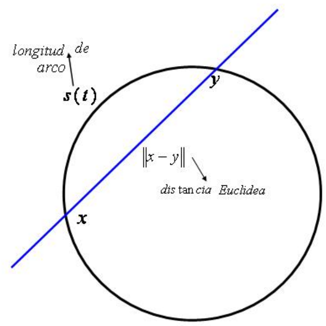



Figura 2.4: La longitud de arco $s(t) \geq \| x - y \|$

*la curva $\alpha$ comenzando en $\alpha(c)$, esta dada por*

$$s(t) = \int_{c}^{t} \sqrt{E(\frac{du}{dt})^2 + 2F\frac{du}{dt}\frac{dv}{dt} + G(\frac{dv}{dt})^2}dt. \tag{2.27}$$

Esta es la manera de medir la distancia entre dos puntos sobre una curva $\alpha$. Se puede probar que la distancia Euclidea siempre es menor o igual que la longitud de arco, es decir,

$$s(t) \geq \| x - y \|,$$

donde $\alpha(c) = x$ y $\alpha(t) = y$ (ver figura 2.4).

**Lema 2.2.6.** *La métrica $ds^2$ sobre $\mathbb{R}^2$, dada en coordenadas cartesianas por*

$$ds^2 = dx^2 + dy^2,$$

*y se expresa en coordenadas polares por*

$$ds^2 = dr^2 + r^2 d\theta^2.$$

**Definición 2.2.17.** *Sea $\mathcal{M} \subset \mathbb{R}^n$ una superficie regular, y sean $p, q \in \mathcal{M}$ la* ***distancia intrínseca*** *$\rho(p, q)$ es la mayor cota inferior de las longitudes de las curvas $\alpha$ contenidas en $M$ que conectan a $p$ y $q$, en general la distancia intrínseca $\rho(p, q)$ será mayor que las distancia dada por $\|p - q\|$.*

Es decir, para dos puntos $p, q$ sobre $\mathcal{M}$ pueden existir infinidades de curvas que pasen por $p, q$ pero de todas estas curvas se toma la que tenga menor longitud, a la longitud de esta curva se le llama longitud intrinseca entre $p, q$ y se denota por $\rho(p, q)$.

**Definición 2.2.18.** *Sea $x : U \to \mathbb{R}^n$ una superficie parametrizada regular. Los símbolos de Christoffel $\Gamma^i_{jk}$, $(i, j, k = 1, 2)$ correspondientes a $x$ se definen por*

$$
\begin{aligned}
\Gamma^1_{11} &= \frac{GE_u - 2FF_u + FE_v}{2(EG - F^2)}, & \Gamma^2_{11} &= \frac{2EF_u - EE_v - FE_u}{2(EG - F^2)}, \\
\Gamma^1_{12} &= \frac{GE_v - FG_u}{2(EG - F^2)}, & \Gamma^2_{12} &= \frac{EG_u - FE_v}{2(EG - F^2)}, \\
\Gamma^1_{22} &= \frac{2GF_v - GG_u - FG_v}{2(EG - F^2)}, & \Gamma^2_{22} &= \frac{EG_v - 2FF_v + FG_u}{2(EG - F^2)}, \\
\Gamma^1_{21} &= \Gamma^1_{12}, & \Gamma^2_{21} &= \Gamma^2_{12}.
\end{aligned}
$$

**Lema 2.2.7.** *Las formulas que determinan los símbolos de Christoffel de una superficie parametrizada regular $x$ para la cual $F = 0$ se simplifican a :*

$$
\begin{aligned}
\Gamma^1_{11} &= \frac{E_u}{2E}, & \Gamma^2_{11} &= \frac{-E_v}{2G}, \\
\Gamma^1_{12} &= \frac{E_v}{2E}, & \Gamma^2_{12} &= \frac{G_u}{2G}, \\
\Gamma^1_{22} &= \frac{-G_u}{2E}, & \Gamma^2_{22} &= \frac{G_v}{2G}.
\end{aligned}
$$

**Teorema 2.2.2.** *(Ecuaciones de Gauss) Sea $x$ una superficie parametrizada regular en $\mathbb{R}^3$ con norma unitaria $U$. Entonces*

$$
\left\{
\begin{aligned}
x_{uu} &= \Gamma^1_{11}x_u + \Gamma^2_{11}x_v + eU, \\
x_{uv} &= \Gamma^1_{12}x_u + \Gamma^2_{12}x_v + fU, \\
x_{vv} &= \Gamma^1_{22}x_u + \Gamma^2_{22}x_v + gU.
\end{aligned}
\right. \tag{2.28}
$$

**Lema 2.2.8.** *Sea $x$ una parametrización local regular. Entre los símbolos de Christoffel y $E, FyG$, se verifican las siguientes relaciones:*

$$
\left\{
\begin{array}{ll}
\Gamma^1_{11}E + \Gamma^2_{11}F = \frac{E_u}{2} & \Gamma^1_{11}F + \Gamma^2_{11}G = F_u - \frac{E_v}{2}, \\
\Gamma^1_{12}E + \Gamma^2_{12}F = \frac{E_v}{2} & \Gamma^1_{12}F + \Gamma^2_{12}G = \frac{G_u}{2}, \\
\Gamma^1_{22}E + \Gamma^2_{22}F = F_v - \frac{G_u}{2} & \Gamma^1_{22}F + \Gamma^2_{22}G = \frac{G_v}{2},
\end{array}
\right. \tag{2.29}
$$

*y*

$$
\left\{
\begin{aligned}
\Gamma^1_{11} + \Gamma^2_{12} &= (log\sqrt{EG - F^2})_u, \\
\Gamma^1_{12} + \Gamma^2_{22} &= (log\sqrt{EG - F^2})_v.
\end{aligned}
\right. \tag{2.30}
$$

# Capítulo 3

# Variedades Diferenciables y Variedades de Riemann.

Una noción de Variedad Diferenciable es necesaria para extender los métodos del calculo diferencial a espacios mas generales que el $R^n$. Un primer ejemplo de variedad accesible a nuestra experiencia es una superficie regular en $R^3$. Además se define el espacio (vectorial) tangente a una variedad diferenciable $\mathcal{M}$ en un punto $p \in \mathcal{M}$. Este espacio se puede considerar la mejor aproximación lineal a $\mathcal{M}$ en $p$. Como ya sabemos el concepto de vector tangente a una curva o a una superficie en $\mathbb{R}^n$ es claro intuitivamente; deseamos introducir ahora un concepto similar para variedades diferenciales arbitraria. Notese que no podemos generalizar el concepto de vector tangente a una variedad pues en el concepto que conocemos el vector tangente también es un vector tangente a $\mathbb{R}^n$, pero en general una variedad arbitraria no esta a priori contenida en un espacio Euclídeo, y por tanto, es necesario establecer una definición de vector tangente que no dependa de tal condición.

Definimos y estudiamos finalmente las variedades de Riemann y algunos conceptos en dicho espacio pues constituyen el ambiente de trabajo sobre el cual desarrollaremos los capítulos siguientes.

La gran parte de este capitulo se obtuvo de las referencias [2] y [4].

## 3.1. Definición de Variedad Diferenciable.

**Definición 3.1.1.** *Una Carta local de un espacio topológico $M$ es un par $(x, U)$ formado por un subconjunto abierto $U$ de $\mathbb{R}^n$ y una aplicación $x : U \to x(U) \subset M$ que es un homeomorfismo entre $U$ y el subconjunto abierto $x(U)$ de $M$. La aplicación $x$ se denomina homeomorfismo local de la carta local, y $x(U)$ su entorno coordenado.*

**Definición 3.1.2.** *Un Atlas $\mathfrak{U}$ sobre un espacio topológico $M$ es una colección de cartas locales $M$ tales que todos los homeomorfismos locales estén definidos en* *subconjuntos abiertos del mismo espacio Euclídeo* $\mathbb{R}^n$ *y además* $M$ *es la unión de todos los conjuntos abiertos* $x(U)$ *con* $(x, U) \in \mathfrak{U}$*. Un espacio topológico* $M$ *sobre el cual este definido un atlas se denomina variedad Topológica.*

**Definición 3.1.3.** *Una variedad diferenciable es un espacio topológico* $M$*, para compacto y Hausdorff, junto con un atlas* $\mathfrak{U}$ *tal que, para todo par de cartas locales* $(x, U), (y, V) \in \mathfrak{U}$ *con* $x(U) \cap y(V) = W \neq \phi$ *se verifica que el cambio de coordenadas* $(x^{-1} \circ y) : y^{-1}(W) \to x^{-1}(W)$ *es diferenciable, es decir, de clase* $C^\infty$ *en el sentido Euclídeo. Salvo que se especifique lo contrario, supondremos que todas las variedades diferenciables son conexas.*

Un par $(x, U)$ (o una aplicación $x$) con $p \in x(U)$ es llamado una parametrización o sistema de coordenadas de $M$ en $p$ y $x(U)$ es llamado una vecindad coordenada en $p$.

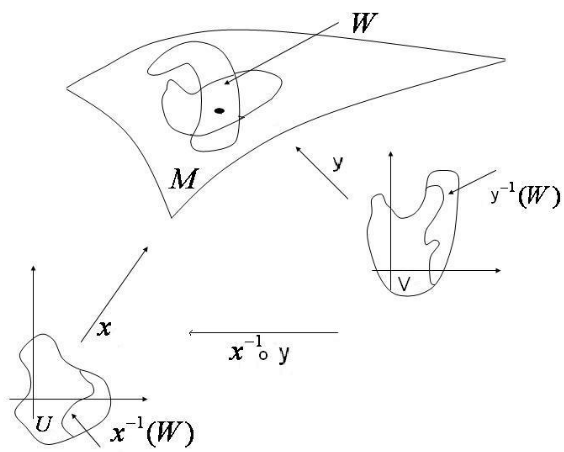



Figura 3.1: Variedad Diferenciable

Notese en la figura 3.1 que siempre se trata es de bajar los cálculos de diferenciabilidad a los espacios Euclídeos donde tenemos bien definidos estos conceptos, así mismo puede notarse que la definición de variedad diferenciable se basa en verificar la diferenciabilidad de la composición $(x^{-1} \circ y) : y^{-1}(W) \to x^{-1}(W)$ donde $W \neq \emptyset$ es la intersección de las cartas $(x, U); (y, V)$.

**Definición 3.1.4.** *La complección* $\overline{U}$ *de un atlas* $U$ *es la colección* $\overline{U} = \{(x, U)/x^{-1} \circ y \quad e \quad y^{-1} \circ x \quad son\ diferenciables \quad \forall (y, v) \in U\}$.

*Se dice que un atlas* $\mathfrak{U}$ *es completo si coincide con su compleción.*

## 3.2. Funciones diferenciables sobre variedades diferenciables.

**Definición 3.2.1.** *Sea $f : W \subset M \to \mathbb{R}$ una función definida en un subconjunto abierto $W$ de una variedad diferenciable $M$. Se dice que $f$ es diferenciable en un punto $p \in W$ si para alguna carta local $x : U \subset \mathbb{R}^n \to M$, con $p \in x(U) \subset W$, la composición $f \circ x : U \subset \mathbb{R}^n \to \mathbb{R}$ es diferenciable en un punto $x^{-1}(p)$ en el sentido Euclídeo. Si $f$ es diferenciable en todos los puntos de $W$, se dice que $f$ es diferenciable en $W$.*

**Definición 3.2.2.** *Sean $M$ y $N$ variedades diferenciables y sea $\psi : M \to N$ una aplicación. Se dice que $\psi$ es diferenciable si la función $y^{-1} \circ \psi \circ x$ es diferenciable, para cada carta local $(x, U)$ en el atlas de $M$ y cada carta local $(y, V)$ en el atlas de $N$ donde la correspondiente composición este definida. Un difeomorfismo entre 2 variedades $M$ y $N$ es una aplicación diferenciable $\phi : M \to N$ que posee una inversa $\phi^{-1} : N \to M$ que es diferenciable. Una aplicación $\psi : M \to N$ se denomina un difeomorfismo local si para cada punto $p \in M$ existe un entorno abierto $U$ tal que $\psi \mid_U : U \to \psi(U)$ es un difeomorfismo.*

**Lema 3.2.1.** *Sea $M$ una variedad diferenciable de dimensión $n$, y sea $x : U \to M$ una carta local. Escribamos $x^{-1} = (x_1, ..., x_n)$. Entonces:*
*i) $x$ es un aplicación diferenciable entre las variedades $U$ y $M$.*
*ii) $x^{-1} : x(U) \to \mathbb{R}^n$ es diferenciable.*
*iii) Para cada $i = 1, ..., n$ la función coordenada $x_i : x(u) \to \mathbb{R}$ es diferenciable.*

## 3.3. Vectores Tangentes a Variedades Diferenciables.

Observación: $\mathcal{F}(\mathcal{M}) = \{f : M \to \mathbb{R} \quad \text{diferenciables}\}$ es un Álgebra de funciones.

**Definición 3.3.1.** *Sea $p$ un punto de una variedad diferenciable $\mathcal{M}$. Un vector tangente a $\mathcal{M}$ en $p$ es una función $v_p : \mathcal{F}(\mathcal{M}) \to \mathbb{R}$, que verifica*

$$v_p[af + bg] = av_p[f] + bv_p[g], \qquad (prop \quad de \quad LINEALIDAD)$$
$$v_p[fg] = f(p)v_p[g] + g(p)v_p[f], \qquad (regla \quad de \quad LEIBNIZ)$$

*para cada $a, b \in \mathbb{R}$ y para cada $f, g \in \mathcal{F}(\mathcal{M})$*

**Definición 3.3.2.** *Sea $x : U \to M$ una carta local de una variedad diferenciable $M$, y escribamos $x^{-1} = (x_1, ..., x_n)$. Para $\mathcal{F}(\mathcal{M})$ y $p \in x(U))$, con $p = x(q)$ se define*

$$\frac{\partial f}{\partial x_i}(p) = \frac{\partial f}{\partial x_i} \mid_{x(q)} = \frac{\partial (f \circ x)}{\partial u_i} \mid_q, \tag{3.1}$$

*para cada $i = 1, .., n$. Aquí, las funciones $u_i$ son las funciones coordenadas naturales de $\mathbb{R}^n$, y la derivada parcial Euclídea ordinaria aparece en el termino de la derecha.*

**Definición 3.3.3.** *Sea $\mathcal{M}$ una variedad diferenciable, y $p \in \mathcal{M}$. El espacio tangente a $\mathcal{M}$ en $p$ es el conjunto $\mathcal{M}_p$ de todos los vectores tangentes a $\mathcal{M}$ en $p$, es decir,*

$$\mathcal{M}_p = \{v_p : \mathcal{F}(\mathcal{M}) \to \mathbb{R} \mid v_p[af+bg] = av_p[f]+bv_p[g], \quad v_p[fg] = f(p)v_p[g]+g(p)v_p[f]\},$$

*para cada $a, b \in \mathbb{R}$ y $f, g \in \mathcal{F}(\mathcal{M})$.*

**Corolario 3.3.1.** *Sea $x : U \to M$ una carta local de una variedad diferenciable $M$, y sea $p \in x(U)$; escribamos $x^{-1} = (x_1 ... x_n)$. Entonces, los vectores*

$$\frac{\partial}{\partial x_1}|_p, ..., \frac{\partial}{\partial x_n}|_p,$$

*forma una base del espacio tangente $\mathcal{M}_p$.Por tanto, la dimensión de cada espacio tangente $\mathcal{M}_p$, como espacio vectorial, es la misma que la dimensión de $\mathcal{M}$ como variedad.*

## 3.4. Campo de vectores sobre variedades diferenciables.

**Definición 3.4.1.** *Sea $A$ un álgebra (no necesariamente conmutativa ni asociativa) sobre un cuerpo $F$. Una **derivación de** $A$ es una aplicación $D : A \to A$ tal que*

$$D(af + bg) = aDf + bDg,$$
$$D(fg) = fD(g) + (Df)g,$$

*cualesquiera que sean $a, b \in F$ y $f, g \in A$.*

**Definición 3.4.2.** *Sean $D_1$ y $D_2$ derivaciones de un álgebra. Entonces, el **corchete** $[D_1, D_2]$ de $D_1$ y $D_2$ esta dado por*

$$[D_1, D_2] = D_1 D_2 - D_2 D_1.$$

**Definición 3.4.3.** *Sea $F$ un cuerpo. Un **álgebra de lie** sobre $F$ es un espacio vectorial $V$ sobre $F$ con un corchete (o producto corchete) $[., .] : V \times V \to V$ tal que*

$$\begin{aligned}[X, X] &= 0,\\ [aX + bY, Z] &= a[X, Z] + b[Y, Z],\\ [Z, aX + bY] &= a[Z, X] + b[Z, Y],\\ G_{xyz}[X, [Y, Z]] &= 0.\end{aligned}$$

*Cualesquiera que sean $X, Y, Z \in V$ y $a, b \in F$*

**Nota**: $G$ representa la suma cíclica.

**Definición 3.4.4.** *Un campo de vectores $X$ sobre un variedad diferenciable $\mathcal{M}$ es una derivación del álgebra $\mathcal{F}(\mathcal{M})$ de las funciones con valores reales y diferenciables sobre $\mathcal{M}$.*

*Así un campo de vectores $X$ sobre $\mathcal{M}$ es una aplicación $X : \mathcal{F}(\mathcal{M}) \to \mathcal{F}(\mathcal{M})$ que verifica*

$$\begin{aligned} X[af + bg] &= aX[f] + bX[g], && \textit{linealidad}\\ X[fg] &= fX[g] + gX[f], && \boldsymbol{\textit{regla de leibniz}}\end{aligned}$$

*para cada $a, b \in \mathbb{R}$ y $f, g \in \mathcal{F}(\mathcal{M})$.*

*El corchete $[X, Y]$ de 2 campos de vectores $X$ e $Y$ (también llamado corchete de Lie) esta definido por*

$$[X, Y] = XY - YX.$$

Denotemos por $\chi(\mathcal{M}) = \{X/X \quad \text{es un campo de vectores sobre M}\}$.

**Definición 3.4.5.** *El fibrado tangente de una variedad diferenciable $\mathcal{M}$ es el conjunto*

$$T(\mathcal{M}) = \{(p, v_p)/v_p \in \mathcal{M}_p, p \in \mathcal{M}\}.$$

## 3.5. Derivadas Covariantes.

**Definición 3.5.1.** *Una conexión o derivada covariante sobre una variedad diferenciable $\mathcal{M}$ es una aplicación*

$$D : \chi(\mathcal{M}) \times \chi(\mathcal{M}) \to \chi(\mathcal{M}),$$

*tal que, para cada $X \epsilon \chi(\mathcal{M})$, induce una aplicación*

$$D_X : \chi(\mathcal{M}) \to \chi(\mathcal{M}),$$

*que verifica las siguientes propiedades:*

$$D_{fX+gY} = fD_X + gD_Y, \tag{3.2}$$

$$D_X(Y+Z) = D_XY + D_XZ, \tag{3.3}$$

$$D_X(fY) = X[f]Y + fD_XY, \tag{3.4}$$

*cualesquiera que sean* $X, Y, Z \epsilon \chi(\mathcal{M})$ *y* $f, g \epsilon \mathcal{F}(\mathcal{M})$.

**Definición 3.5.2.** *Sea* $v_p$ *un vector tangente a* $\mathcal{M}$ *en* $p \epsilon \mathcal{M}$ *y sea* $V \in \chi(\mathcal{M})$ *tal que* $V_p = v_p$. *Entonces, para cada* $W \in \chi(\mathcal{M})$, *se define el vector* $D_VW$ *en* $\mathcal{M}_p$ *por*

$$D_VW = (D_VW)(p).$$

**Lema 3.5.1.** *Sea* $W \in \chi(\mathcal{M})$. *La definición de* $D_vW$ *es independiente de la elección de* $V \in \chi(\mathcal{M})$ *tal que* $V_p = v_p$.

**Lema 3.5.2.** *Sea* $W \in \chi(\mathcal{M})$, *y sea* $\alpha : (a,b) \to \mathcal{M}$ *una curva tal que* $\alpha'(t) = X_{\alpha(t)}$ *es decir,* $\alpha$ *es una curva integral de* $X$. *Sean* $Y, \widetilde{Y} \in \chi(\mathcal{M})$ *tales que*

$$Y_{\alpha(t)} = \widetilde{Y}_{\alpha(t)},$$

*para* $a < t < b$. *Entonces,*

$$(D_XY)_{\alpha(t)} = (D_X\widetilde{Y})_{\alpha(t)},$$

*para* $a < t < b$.

**Definición 3.5.3.** *Sea* $\alpha : (a,b) \to \mathcal{M}$ *una curva sobre una variedad diferenciable* $\mathcal{M}$. *Un campo de vectores a lo largo de* $\alpha$ *es una aplicación* $Y$ *que asigna a cada* $t$, *con* $a < t < b$, *un vector tangente* $Y(t) \epsilon \mathcal{M}_{\alpha(t)}$.

**Definición 3.5.4.** *Sea* $\alpha : (a,b) \to \mathcal{M}$ *una curva regular inyectiva sobre una variedad diferenciable* $\mathcal{M}$, *y sea* $Y$ *un campo de vectores a lo largo de* $\alpha$. *Entonces, la derivada covariante del campo de vectores* $Y$ *a lo largo de* $\alpha$ *es*

$$Y\prime(t) = D_{\alpha'(t)}\widetilde{Y},$$

*donde* $\widetilde{Y} \epsilon \chi(\mathcal{M})$ *es cualquier campo de vectores tal que* $\widetilde{Y}_{\alpha(t)} = Y(t)$, *para todo* $a < t < b$.

**Lema 3.5.3.** *Sean* $Y$ *y* $Z$ *campos de vectores a lo largo de una curva* $\alpha : (a,b) \to \mathcal{M}$ *sobre una variedad diferenciable* $\mathcal{M}$, *y sea* $f : (a,b) \to \mathbb{R}$ *una función diferenciable. Se verifican:*

$$(Y+Z)' = Y' + Z', \tag{3.5}$$

$$(fY)' = f'Y + fY'. \tag{3.6}$$

**Lema 3.5.4.** *Sea $Y$ un campo de vectores a lo largo de una curva $\alpha : (a,b) \to \mathcal{M}$ sobre una variedad diferenciable $\mathcal{M}$, y sea $h : (c,d) \to (a,b)$ una función diferenciable. Entonces*

$$(Y \circ h)'(t) = h'(t)Y'(h(t)),$$

*para $c < t < d$.*

**Definición 3.5.5.** *Sea $\alpha : (a,b) \to \mathcal{M}$ una curva inyectiva sobre una variedad diferenciable $\mathcal{M}$, y sea $D$ una conexión sobre $\mathcal{M}$. La aceleración de $\alpha$ con respecto a $D$ es el campo de vectores*

$$t \longmapsto D_{\alpha'(t)}\alpha'(t),$$

*que representaremos por $D_{\alpha'}\alpha'$. Se dice que una curva $\alpha : (a,b) \to \mathcal{M}$ es una geodésica con respecto a $D$ si*

$$D_{\alpha'(t)}\alpha'(t) = 0,$$

*para todo $a < t < b$.*

**Lema 3.5.5.** *Sea $\mathcal{M}$ una variedad diferenciable, y sea $D$ una conexión sobre $\mathcal{M}$. Sean $\alpha : (a,b) \to \mathcal{M}$ una curva sobre $\mathcal{M}$ y $\beta = \alpha \circ h$ una reparametrización de $\alpha$, siendo $h : (c,d) \to (a.b)$ una función diferenciable. Se verifica entonces*

$$D_{\beta'}\beta' = h'^2(D_{\alpha'}\alpha') \circ h + h''(\alpha' \circ h). \tag{3.7}$$

## 3.6. Métricas de Riemann.

**Definición 3.6.1.** *Una métrica Riemanniana (o estructura riemanniana) en una variedad diferenciable $M$ es una correspondencia que asocia a cada $p \in M$ un producto interno $<,>_p$ (esto es, una forma bilineal simétrica, definida positiva) en el espacio tangente $T_pM$ que varia diferencialmente en el siguiente sentido: si $x : U \subset \mathbb{R}^n \to M$ es un sistema de coordenadas locales en torno de $p$, con $x(x_1, ..., x_n) = q \in x(U)$ y $\frac{\partial}{\partial x_i}(q) = dx(0,.,1,.,0)$ entonces*

$$< \frac{\partial}{\partial x_i}(q), \frac{\partial}{\partial x_j}(q) >_p = g_{ij}(x_1, ..., x_n),$$

*es una función diferenciable en $U$. De ser así, $M$ se llamara* ***variedad Riemanniana.***

Asi tenemos que una Variedad Riemanniana es una Variedad Diferenciable en la que en cada $p \in M$ se puede definir un producto interno en el espacio tangente $T_pM$, esto es se puede definir una metrica de Riemann.

**Lema 3.6.1.** *Sea $V$ un espacio vectorial con un producto interno $\langle , \rangle$. Cada $x \in V$ esta completamente determinado por todos los productos interiores $\langle x, y \rangle$, para cada $y \epsilon V$.*

**Definición 3.6.2.** *Una conexión afín $\nabla$ en una variedad diferenciable $M$ es una aplicación*

$$\nabla : X(M) \times X(M) \to X(M), \quad (X,Y) \to \nabla_X Y,$$

*que satisface las siguientes propiedades :*

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

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

Si $\frac{d}{dt} < V, W > = < \frac{DV}{dt}, W > + < V, \frac{DW}{dt} >$ para cualquier par $V, W$ campo de vectores a lo largo de una curva $c$, entonces la conexión $\nabla$ es compatible con la métrica de Riemann.

**Proposición 3.6.1.** *Sea $M$ una variedad diferenciable con una conexión afín $\nabla$.*
*Entonces existe una única ley que asocia a un campo vectorial $V$ a lo largo de una curva diferenciable $c : I \to M$ con otro campo vectorial $\frac{DV}{dt}$ alo largo de c, denominado derivada covariante de $V$ a lo largo de c, tal que:*

1. $\frac{D}{dt}(V + W) = \frac{DV}{dt} + \frac{DW}{dt}$,
2. $\frac{D}{dt}(fV) = \frac{df}{dt} V + f \frac{DV}{dt}$, *donde $W$ es un campo de vectores a lo largo de c y f diferenciable sobre $M$.*
3. *Si $V(t) = Y(c(t))$, entonces $\frac{DV}{dt} = \nabla_{\frac{dc}{dt}} Y$.*

Demostracion: ver [4]
Esta proposicion forma parte importante del trabajo, pues se utilizara para demostrar que al calculo de geodesicas se reduce a las soluciones de un sistema de ecuaciones diferenciales de segundo grado.

**Teorema 3.6.1.** ***Levi-Civita*** *Dada una variedad Riemanniana $M$, existe una única conexión afín $\nabla$ en $M$ satisfaciendo:*

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

Toda conexión que cumpla con el teorema anterior es denominada conexión de Levi-Civita o Riemanniana de $M$.

**Definición 3.6.3.** *Sea $\mathcal{M}$ una variedad de Riemann. La conexión de Riemann $\nabla$ de $\mathcal{M}$ esta dado por*

$$2\langle \nabla_X Y, Z\rangle = X\langle Y, Z\rangle + Y\langle X, Z\rangle - Z\langle X, Y\rangle - \langle X, [Y, Z]\rangle - \langle Y, [X, Z]\rangle + \langle Z, [X, Y]\rangle, \tag{3.8}$$

*cualesquiera que sean $X, Y, Z \in \chi(\mathcal{M})$.*

**Lema 3.6.2.** *La conexión de Riemann $\nabla$ verifica las propiedades:*

$$\nabla_{fX+gY} = f\nabla_X + g\nabla_Y, \tag{3.9}$$

$$\nabla_X(aY + bZ) = a\nabla_X Y + b\nabla_X Z, \tag{3.10}$$

$$\nabla_X fY = X[f]Y + f\nabla_X Y, \tag{3.11}$$

$$\nabla_X Y - \nabla_Y X = [X, Y], \tag{3.12}$$

$$X\langle Y, Z\rangle = \langle \nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle, \tag{3.13}$$

*cualesquiera que sean $X, Y, Z \epsilon \chi(\mathcal{M})$, $f, g \epsilon \mathcal{F}(\mathcal{M})$ y $a, b \epsilon \mathbb{R}$.*
*Recíprocamente, si*

$$\nabla : \chi(\mathcal{M}) \times \chi(\mathcal{M}) \to \chi(\mathcal{M}),$$

*verifica 3.9-3.13, entonces $\nabla$ esta dada por 3.8*

**Lema 3.6.3.** *Sea $\langle , \rangle$ una métrica pseudo-Riemanniana sobre una variedad $\mathcal{M}$. Entonces, existe una única conexión $\nabla$ sobre $\mathcal{M}$ que verifica 3.12 y 3.13.*

## 3.7. Tratamiento clásico de las métricas.

**Lema 3.7.1.** *Una métrica $ds^2 = \langle \quad , \quad \rangle$ esta dada, en términos de las $g_{ij}$, por*

$$ds^2 = \sum_{ij=1}^{n} g_{ij} dx_i dx_j. \tag{3.14}$$

**Definición 3.7.1.** *Sean $g_{ij}$ $(1 \leq i, j \leq n)$ las componentes de una métrica con respecto a un sistema de coordenadas locales $(x_1, ..., x_n)$; entonces $(g^{ij})$ representara la matriz inversa de $(g_{ij})$. Escribamos*

$$g = det(g_{ij}).$$

**Lema 3.7.2.** *Para cualquier sistema de coordenadas locales* $(x_1, ..., x_n)$ *de una variedad de Riemann* $\mathcal{M}$*, el dual de la* $1-forma$ *diferencial* $dx_i$ *es el campo de vectores*

$$\sum_{h=1}^{n} g^{ih}\frac{\partial}{\partial x_h},$$

*para cada* $i = 1, ..., n$*, es decir,*

$$dx_i^* = \sum_{h=1}^{n} g^{ih}\frac{\partial}{\partial x_h}.$$

**Definición 3.7.2.** *Sea* $\langle,\rangle$ *una métrica sobre una variedad diferenciable* $\mathcal{M}$ *de dimensión* $n$*, y sea* $\nabla$ *su conexión de Riemann. Si* $(x_1, ..., x_n)$ *es un sistema de coordenadas locales de* $\mathcal{M}$*, los símbolos de Christoffel* $\Gamma_{ij}^k$ *de* $\nabla$ *relativos a* $(x_1, ..., x_n)$ *son las funciones (locales) dadas por*

$$\Gamma_{ij}^k = dx_k(\nabla_{\frac{\partial}{\partial x_i}}\frac{\partial}{\partial x_j}), \tag{3.15}$$

*para* $1 \leq i, j \leq n$*.*

**Lema 3.7.3.** *Sea* $(x_1, ..., x_n)$ *un sistema de coordenadas locales de una variedad* $\mathcal{M}$ *con una métrica, y sea* $\nabla$ *su conexión de Riemann. Entonces,*

$$\nabla_{\frac{\partial}{\partial x_i}}\frac{\partial}{\partial x_j} = \sum_{k=1}^{n}\Gamma_{ij}^k\frac{\partial}{\partial x_k}, \tag{3.16}$$

*para* $1 \leq i, j \leq n$*.*

El lema anterior muestra que los simbolos de Christoffel son las componentes de la conexion de Riemann.

**Lema 3.7.4.** *Los símbolos de Christoffel de una conexión de Riemann verifican la siguiente relación de simetría*

$$\Gamma_{ij}^k = \Gamma_{ji}^k, \tag{3.17}$$

*para cada* $i, j, k = 1, ..., n$*.*

**Lema 3.7.5.** *Se verifica*

$$\Gamma_{jk}^i = \sum_{h=1}^{n}\frac{g^{ih}}{2}(\frac{\partial g_{hj}}{\partial x_k} + \frac{\partial g_{hk}}{\partial x_j} - \frac{\partial g_{jk}}{\partial x_h}). \tag{3.18}$$

**Lema 3.7.6.** *Sean $X, Y \epsilon \chi(\mathcal{M})$, y sea $(x_1, ..., x_n)$ un sistema de coordenadas locales de una variedad de Riemann $\mathcal{M}$. Se verifica entonces*

$$\nabla_X Y = \sum_{k=1}^{n} XY[x_k] + \sum_{ij=1}^{n} X[x_i]Y[x_j]\Gamma_{ij}^{k}\frac{\partial}{\partial x_k}. \tag{3.19}$$

Sea $(M, g)$ una variedad Riemanniana, $\mathfrak{F}(M)$ el álgebra de todas las funciones reales $C^\infty$ sobre $M$ y $X(M)$ el álgebra de lie de todos los campos de vectores $C^\infty$ sobre $M$.

Definimos el campo de vectores $gradf$ por

$$g(X, gradf) = X(f) = df(X), \quad \forall X \in X(M), \quad f \in \mathfrak{F}(M)$$

y algunas veces es denotado por $\nabla f$.
Cuando $df(X) < 0$ es porque $X$ es el campo de vectores que forma angulo mayor (o menor) que $\frac{\pi}{2}$ $(-\frac{\pi}{2})$ con $gradf$, esta observacion sera util cuando veamos el algoritmo, pues esto es lo que nos da la forma de conseguir direcciones de mayor pendiente.

Si denotamos las componentes locales de la diferencial $df$ por $f_i = \frac{\partial f}{\partial x^i}$, entonces las componentes local de $gradf$ son $f^i = g^{ij}f_j$.

De igual manera definimos el Hessiano de $f$ $Hessf$ por

$$Hessf = \nabla(\nabla f) = \nabla df.$$

También se puede ver como

$$Hessf(X, Y) = g(\nabla_X gradf, Y) = \nabla_X(df)(Y) = X(Y(f)) - (\nabla_X Y)(f), \quad \forall X, Y \in X(M)$$

así las componentes locales de $Hessf$ están denotadas por

$$f_{jk} = \frac{\partial^2 f}{\partial x^j \partial x^k} - \Gamma_{jk}^{i}\frac{\partial f}{\partial x^i},$$

donde $\Gamma_{jk}^{i}$ son las componentes locales de la conexión de Riemann sobre $M$.

# Capítulo 4

# Geodésicas y Mapeo Exponencial.

En este capitulo estudiaremos las curvas sobre variedades que tienen menor longitud entre dos puntos y las llamaremos geodésicas, también calcularemos el sistema de ecuaciones diferenciales de segundo orden cuyas soluciones serán las geodésicas y estudiaremos el mapeo exponencial como un mapeo que va desde el espacio tangente hasta la variedad misma. Para finalizar veremos dos ejemplos de mapeo exponencial.

Sea $(M, g)$ una variedad Rienmanniana de dimensión $n$ y sea $I$ un intervalo de $\mathbb{R}$.

**Definición 4.1.** *Una curva $\gamma : I \to M$ $C^\infty$ cuyo campo de vectores aceleración $\frac{\nabla}{dt}\frac{d\gamma}{dt}$ es identicamente nula es llamada "GEODÉSICA."*

Obsérvese que $\gamma$ es geodésica si y solo si el campo de vectores velocidad es paralelo a lo largo de $\gamma$. También, si $\gamma$ es una geodésica de $(M, g)$ se tiene que,

$$0 = \frac{d}{dt}(g(\frac{d\gamma}{dt}, \frac{d\gamma}{dt})) = 2g(\frac{\nabla}{dt}\frac{d\nabla}{dt}, \frac{d\gamma}{dt}),$$

así la velocidad $\|\frac{d\gamma}{dt}\| = a$ es una constante a lo largo de $\gamma$ y con la longitud de arco

$$s(t) = \int \|\frac{d\gamma}{dt}\| dt,$$

obtenemos que $s(t) = at + b, \quad t \in I$.Asi podemos decir que para una geodésica la longitud de arco crece linealmente.

Sea $\gamma : I \to M$ una geodésica, sea $J$ un subintervalo de $I$ y $U$ una vecindad que tiene sistema coordenados $(x^1, ..., x^n)$, si $\gamma(J) \subset U$,
entonces $\gamma(t) = (x^1(t), ..., x^n(t))$ y la definición de geodésica toma la forma de un sistema de ecuaciones diferenciables de orden 2

$$\frac{d^2x^i}{dt^2} + \sum_{j,k=1}^{n} \Gamma^i_{jk} \frac{dx^j}{dt}\frac{dx^k}{dt} = 0, \quad i = 1, ..., n,$$

en efecto, sea $V=\sum\limits_{j=1}^{n} v^j X_j$ un campo de vectores con $X_1,...,X_n$ base del espacio vectorial.

luego,

$$\begin{aligned}
\tfrac{D}{dt}(V) &= \tfrac{D}{dt}(\sum_{j=1}^{n} v^j X_j),\\
&= \sum_{j=1}^{n} \tfrac{D}{dt}(v^j X_j),\\
&= \sum_{j=1}^{n} (\tfrac{dv^j}{dt} X_j + v^j \tfrac{D}{dt} X_j),\\
&= \sum_{j=1}^{n} (\tfrac{dv^j}{dt} X_j + v^j \nabla_{\frac{d\gamma}{dt}} X_j),\\
&= \sum_{j=1}^{n} (\tfrac{dv^j}{dt} X_j + v^j \nabla_{\sum_{i=1}^{n} \frac{dx^i}{dt} X_i} X_j),\\
&= \sum_{j=1}^{n} (\tfrac{dv^j}{dt} X_j + v^j \sum_{i=1}^{n} \tfrac{dx^i}{dt} \nabla_{X_i} X_j),\\
&= \sum_{j=1}^{n} \tfrac{dv^j}{dt} X_j + \sum_{j,i=1}^{n} v^j \tfrac{dx^i}{dt} \sum_{k=1}^{n} \Gamma_{ij}^k X_k,\\
&= \sum_{k=1}^{n} \tfrac{dv^k}{dt} X_k + \sum_{k,j,i=1}^{n} v^j \tfrac{dx^i}{dt} \Gamma_{ij}^k X_k,\\
&= \sum_{k=1}^{n} [\tfrac{dv^k}{dt} + \sum_{j,i=1}^{n} v^j \tfrac{dx^i}{dt} \Gamma_{ij}^k] X_k,
\end{aligned}$$

así obtenemos que $\frac{D}{dt}(V)=\sum\limits_{k=1}^{n}[\frac{dv^k}{dt}+\sum\limits_{i,j=1}^{n} v^j \frac{dx^i}{dt}\Gamma_{ij}^k]X_k$ y haciendo $V=\frac{d\gamma}{dt}$ se tiene que $v^j=\frac{dx^j}{dt}$ y $\frac{dv^j}{dt}=\frac{d^2x^j}{dt^2}$ y además $X_j=\frac{\partial}{\partial x_j}$ de esta manera se obtiene que:

$$\frac{D}{dt}(\frac{d\gamma}{dt}) = \sum_{k=1}^{n}(\ddot{x}^k + \sum_{i,j=1}^{n} \dot{x}^j \dot{x}^i \Gamma_{i.j}^k)\frac{\partial}{\partial x_k},$$

y si $\gamma$ es una geodésica se obtiene que

$$0 = \frac{D}{dt}(\frac{d\gamma}{dt}),$$

lo que implica que:

$$\ddot{x}^k + \sum_{i,j=1}^{n} \dot{x}^j \dot{x}^i \Gamma_{i.j}^k = 0.$$

Así la existencia de geodésicas se reduce a la existencia de solución del sistema de ecuaciones obtenido arriba.

**Lema 4.1.** *Para cada punto $x_0 \in M$ existe una vecindad $U_{x_0}$ y $\epsilon > 0$ tal que para cada $x \in U_{x_0}$ y algún $X \in M_x$ con $\|X\| < \epsilon$ existe una única geodésica*

$$\gamma_x : (-\epsilon, \epsilon) \to M,$$

*la cual satisface las condiciones iniciales*

$$\gamma_x(0) = x \qquad \frac{d\gamma}{dt} = X.$$

Demostraciones: ver [1]

En el algoritmo del metodo de mayor pendiente se utilizara la existencia unica de una geodesica, es decir, para $x \in M$ y $X_x \in T_xM$ existe unica geodesica de tal manera que pasa por $x$ y que el vector velocidad de la geodesica es precisamente $X_x$ y todo esto lo garantiza el lema anterior.

**Lema 4.2.** *Sea $\gamma_X : (-\epsilon, \epsilon) \to M$ geodésica determinada por las condiciones iniciales*

$$\gamma_X(0) = x, \qquad \frac{d\gamma_X}{dt}(0) = X,$$

*entonces $\gamma_{\lambda X} : (-\frac{\epsilon}{\lambda}, \frac{\epsilon}{\lambda}) \to M, \quad \lambda \in -\{0\}$ es una geodésica con la propiedad*

$$\gamma_{\lambda X}(t) = \gamma_X(\lambda t).$$

**Definición 4.2.** *Sea $X \in T_xM$ y suponga que existe una geodésica $\gamma_x : [0,1] \to M$ la cual satisface las condiciones iniciales*

$$\gamma_x(0) = x \qquad \frac{d\gamma}{dt}(0) = X,$$

*se define la exponencial de $X$ denotado por $exp_xX$ a el punto $\gamma_x(1)$, (ver figura 4.1)*

**Definición 4.3.** *Una variedad $M$ es llamada geodesicamente completa si $exp_xX$ es definida para cada $X \in T_xM$.*

También se puede decir que una variedad M es geodesicamente completa si un segmento geodésico $\gamma : [a,b] \to M$ puede ser extendido a una geodésica del tipo $\gamma : \mathbb{R} \to M$.

Sea $\mathfrak{M} = \cup_{x \in M} M_x$ el fibrado tangente de $M$. Las coordenadas locales en $\mathfrak{M}$ son de la forma $(x, y)$ con $x \in M$ y $y \in T_xM$.

**Definición 4.4.** *La función $(x, y) \to exp_xy$ es llamada el mapeo exponencial y es denotado por $exp$.*

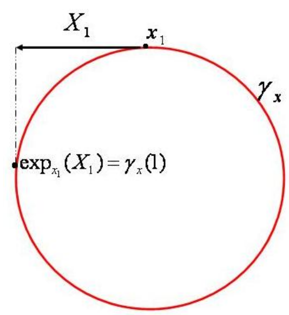



Figura 4.1: Mapeo Exponencial

1) **El mapeo exponencial para $\mathbb{R}^2$.**

Sea $x \in \mathbb{R}^2$ y $X \in \mathbb{R}^2_x$, la geodésica $\gamma$ de $\mathbb{R}^2$ que pasa por $x$ a través del tangente $X$ es la recta de ecuación

$$\gamma_x(t) = x + tX, \quad t \in \mathbb{R}$$

así

$$exp_x X = \gamma_x(1) = x + X$$

Como se muestra en la figura 4.2, en $\mathbb{R}^2$ las geodésicas son rectas de la forma $x + tX$, y notese que $\mathbb{R}^2_x$ es una copia de $\mathbb{R}^2$ pero trasladada al punto $x$. Así que

$$exp_x X = \gamma(1) = x + X.$$

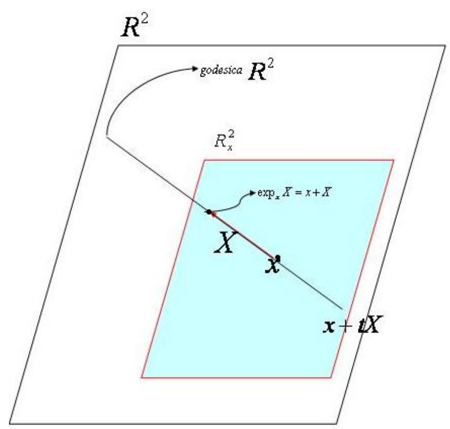



Figura 4.2: Geodésicas en $R^2$

(2) **El mapeo exponencial para el cilindro $M : x^2 + y^2 = 1$ en $\mathbb{R}^3$.**
Las geodésicas de $M$ están dadas por

$$\gamma(t) = (cos(at + b), sen(at + b), at + b), \quad t \in \mathbb{R}$$

claramente $\gamma(t)$ esta en M, ahora veamos que $\gamma(t)$ son geodésicas,
en efecto,

$$\gamma'(t) = (-asen(at + b), acos(at + b), a)$$

y así se obtiene

$$\|\gamma'(t)\| = 2a^2 = const.$$

Luego $\gamma(t)$ dada así, son geodésicas.

La geodésica que pasa por el punto $P = (cos(\varphi), sen(\varphi), z)$ siguiendo el tangente $v = (sen(\varphi), -cos(\varphi), v_3)$ en $t = 0$ es

$$\gamma_v(t) = (cos(-t + \varphi), sen(-t + \varphi), v_3 t + z), \quad t \in [0, +\infty)$$

esto se comprueba simplemente evaluando $\gamma_v$ en $p$ y observando que el vector velocidad de la curva $\gamma_v$ es precisamente el vector $v$ antes dado, así,

$$exp_p v = \gamma_v(1) = (cos(-1 + \varphi), sen(-1 + \varphi), v_3 + z).$$

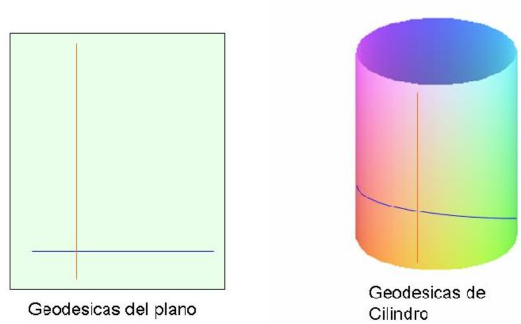



Figura 4.3: Geodésicas en el Cilindro

# Capítulo 5

# Método de Gradiente sobre Variedades Riemanniana.

## 5.1. Método de gradiente o de mayor descenso sobre $\mathbb{R}^n$.

El método de mayor descenso o del gradiente es uno de los mas fundamental para minimizar funciones diferenciables de varias variables.

**Definición 5.1.1.** *Un vector $d \in \mathbb{R}^n$ es una dirección de descenso de $f$ en $x$ si existe un $\delta > 0$ tal que*

$$f(x + \lambda d) < f(x), \quad \forall \lambda \in (0, \delta).$$

También se puede ver, si

$$\lim_{\lambda \to 0^+} \frac{f(x + \lambda d) - f(x)}{\lambda} < 0,$$

entonces $d$ es una dirección de descenso.
Entiendase $E_n$ como $\mathbb{R}^n$ con norma Euclídea.

**Lema 5.1.1.** *Suponga que $f : E_n \to E_1$ es diferenciable en $x$ y supongamos que $\nabla f(x) \neq 0$. Entonces la solución óptima a el problema a minimizar $f'(x, d)$ sujeto a $\| d \| \leq 1$ esta dado por*

$$\overline{d} = \frac{-\nabla f(x)}{\| \nabla f(x) \|},$$

*esto es, $\overline{d} = \frac{-\nabla f(x)}{\|\nabla f(x)\|}$ es la dirección de mayor descenso de $f$ en $x$.*

**Demostracion**: [1]

## 5.2. Algoritmo para el método de mayor pendiente o del gradiente.

Dado un punto $x_0$, el algoritmo de mayor pendiente procede haciendo una búsqueda lineal $\lambda$ a lo largo de la dirección

$$\frac{-\nabla f(x_0)}{\| \nabla f(x_0) \|},$$

o equivalente a lo largo de la dirección $-\nabla f(x_0)$.
Dado un punto inicial $x_k$.
Sea $\epsilon > 0$, si $x_k$ es tal que $\| \nabla f(x_k) \| < \epsilon$ no hay nada que hacer,pues ya $x_k$ sera el punto que estamos buscando, en otro caso.

Sea $d_k = -\nabla f(x_k)$ y sea $\lambda_k$ solución óptima del problema

$$min \quad f(x_k + \lambda d_k)$$

$$sujeto \quad \lambda \geq 0.$$

Y hacemos $x_{k+1} = x_k + \lambda_k d_k$.

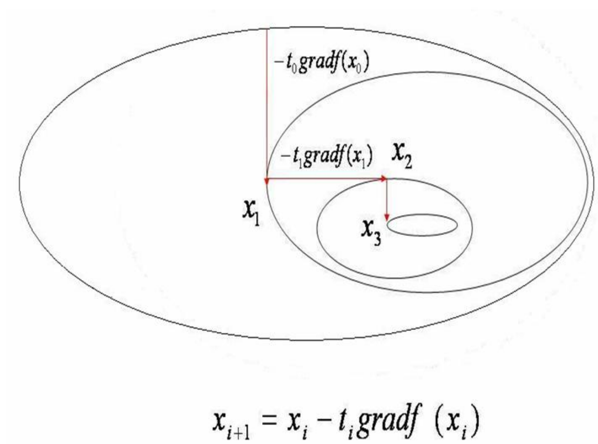



Figura 5.1: La dirección de descenso es $-gradf(x)$

Como se puede observar en la figura 5.1 el *gradf* siempre señala hacia el sentido de mayor crecimiento y el método del gradiente utiliza el $-gradf$ como el sentido de mayor descenso, en este sentido en la figura se muestra como en el método se van obteniendo puntos mas próximos al punto mínimo.

## 5.3. Método de mayor pendiente o Método de Gradiente.

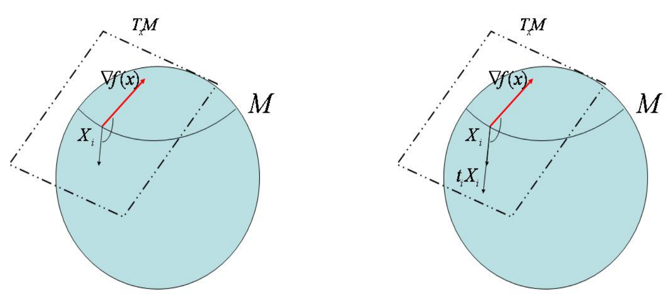
Sea $(M, g)$ una variedad Riemanniana completa de dimensión finita.
Suponga que tenemos una función $f : M \to \mathbb{R}$ diferenciable y deseamos calcular uno de sus mínimos.

El algoritmo general de descenso sobre una variedad de Riemann esta dado por el siguiente proceso:

1. sea $x_1$ un punto inicial.
2. Sea $i = 1$,
3. si $x_i$ cumple criterios de parada entones $x_i$ es minimo, si no se sigue.
4. Calculamos $X_i$ tal que $df(X_i) < 0$,
5. Luego se calcula $t_i$ tal que $f(x_i) > f(exp_{x_i}(t_i X_i))$,
6. Calculamos el punto $x_{i+1}$ como $x_{i+1} = exp_{x_i}(t_i X_i)$, y hacemos 2

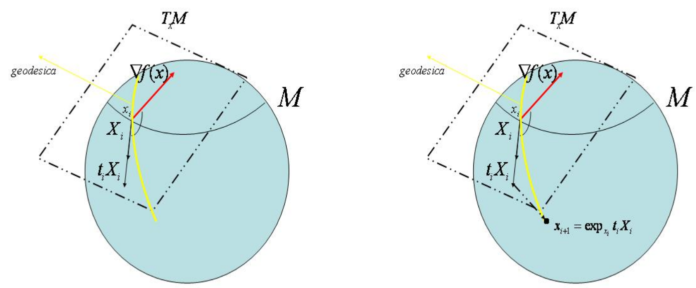
Como criterio de parada podemos utilizar:

- $\| \, gradf(x_i) \, \| < \epsilon, \quad \forall \epsilon > 0$
- $| \, f(x_i) - f(x_{i+1}) \, | < \eta, \forall \eta > 0$.



En las figuras anteriores se puede ver graficamente como el algoritmo cálcula los vectores $X_i$ y el $t_i$, para la función altura definida en los primeros capitulos, para una variedad como la esfera unitaria.

En las gráficas siguentes se observa como se cálcula la geodésica y como se aplica el mapeo exponencial para conseguir el punto $x_{i+1}$.



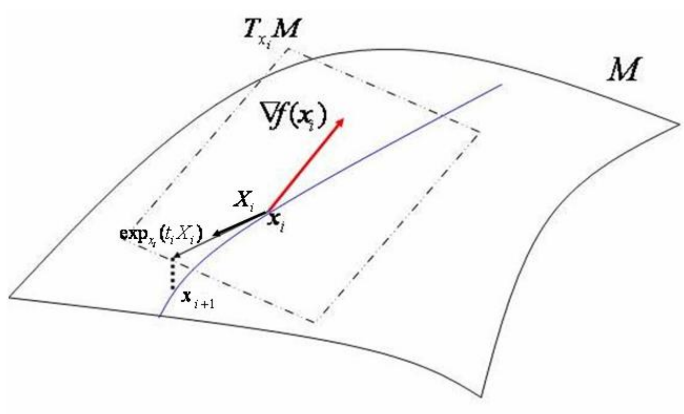




Figura 5.2: Ejemplo Gráfico

En la figura 5.2 observamos como el vector $X_i$ es exactamente una dirección de descenso, y como el mapeo exponencial nos proporciona de un punto en la variedad donde se cumple $f(x_i) > f(exp_{x_i}(t_iX_i))$.

Recordemos que $-gradf$ es el campo de vectores ortogonal a la hipersuperficie de nivel constante de $f$, el cual muestra a cada punto $x \in M$ la dirección y el sentido de mayor descenso, sugiriendo que la solución mas conveniente de la desigualdad

$$(df(x_1)(X)) < 0 \quad \text{es} \quad X = -gradf(x_1),$$

pues si $X = -gradf(x_1)$ obtenemos el valor mas negativo de $df(x_1)$.
Esto resulta en un proceso iterativo el cual llamaremos el método de mayor pendiente o el método del gradiente.

El método del gradiente es descrito por el siguiente algoritmo:

1. Consideramos un punto inicial $x_1$ y calculamos $gradf(x_1)$, si $gradf(x_1) = 0$ paramos,

2. Si $gradf(x_1) \neq 0$, estonces del punto $x_1$ pasamos al otro punto

$$x_2 = exp_{x_1}(-t_1 gradf(x_1)),$$

donde $t_1 > 0$ es calculado tal que $f(x_2) < f(x_1)$),

3. Si $gradf(x_2) = 0$ paramos, si $gradf(x_2) \neq 0$ calculamos

$$x_3 = exp_{x_2}(-t_2 gradf(x_2)),$$

donde $t_2$ es calculado de tal manera que $f(x_3) < f(x_2)$,

4. De manera general, si $gradf(x_i) \neq 0$ calculamos

$$x_{i+1} = exp_{x_i}(-t_i gradf(x_i)),$$

donde $t_i > 0$ es calculado de tal manera que $f(x_{i+1}) < f(x_1)$.

Recordar que:

1. Un problema de minimizar puede ser reducido a un problema de maximizar.

2. En el problema numérico de maximizar podemos usar $gradf$ el cual nos da la dirección y el sentido de mayor crecimiento y podemos utilizar $gradf(x_i)$

$$x_{i+1} = exp_{x_i}(t_i gradf(x_i)),$$

con $t_i > 0$ tal que $f(x_{i+1}) > f(x_i)$.

Ahora definamos los pasos para determinar los $t_i$ en el proceso iterativo

1. $x = exp_{x_i}(-tgradf(x_i))$ con $t > 0$ arbitrariamente establecido y el mismo para todas las iteraciones.

2. Calculamos $f(x) = f(exp_{x_i}(-tgradf(x_i)))$.

3. Verificamos la desigualdad $f(x) - f(x_i) \leq \epsilon tdf(X_i), \quad X_i = -gradf(x_i)$ donde $\epsilon \in (0,1)$ arbitrario pero fijo independiente de $i$.

4. Si la desigualdad es satisfecha entonces podemos cambiar $t_i = t$, si no se cumple la desigualdad entonces reemplazamos $t$ por $\lambda t$, $\lambda \in (0,1)$ fijo tal que es satisfecha la desigualdad.

## 5.4. Convergencia de $\{gradf\}$ a cero.

**Teorema 5.4.1.** *Sea $f : M \to \mathbb{R}$ una función $C^2$ acotada inferiormente. Denotemos por $X_x, X_y$ los vectores tangentes a $x, y$ respectivamente. Si para cualquier $x, y \in M$ la condición de Lipschitz*

$$|df(X_y) - df(X_x)| \leq rd^2(x,y), \quad r > 0$$

*es satisfecha y si el cambio del número $t_i$ es hecho como fue descrito, entonces en el proceso iterativo tenemos que*

$$\lim_{i\to\infty} gradf(x_i) = 0,$$

*para un punto inicial $x_1$ dado.*

**Demostracion**:
Sea $\gamma_{xy} : [0,1] \to M$ una geodésica que pasa por $x$ y $y$ tal que

$$\gamma_{xy}(0) = x \quad \text{y} \quad \gamma_{xy}(1) = y.$$

Dado que $f$ es de clase $C^1$, tenemos que

$$f(y) - f(x) = \int_0^1 \frac{d}{du}(f(\gamma_{xy}(u)))du = \int_0^1 df((\dot{\gamma_{xy}}(u)))du = df(\dot{\gamma_{xy}}(u_0))du \quad u_0 \in [0,1].$$

Si denotamos a $z = \gamma_{xy}(u_0)$ y usando la hipótesis tenemos que

$$\begin{array}{rrl} & |df(\dot{\gamma_{xy}}(u_0)) - df(X_k)| & \leq rd^2(x,z) \\ \Rightarrow & df(\dot{\gamma_{xy}}(u_0)) - df(X_k) & \leq rd^2(x,z) \\ \Rightarrow & df(\dot{\gamma_{xy}}(u_0)) & \leq df(X_k) + rd^2(x,z) \end{array}$$

pero $\parallel \dot{\gamma_{xy}}(u) \parallel = \parallel \dot{\gamma_{xy}}(0) \parallel = ctte$ así se tiene que

$$d^2(x,y) \leq (\int_0^{u_0} \parallel \dot{\gamma_{xy}}(u) \parallel du)^2 = \parallel \dot{\gamma_{xy}}(0) \parallel^2 u_0^2 \leq \parallel \dot{\gamma_{xy}}(0) \parallel^2,$$

poniendo $X_k = \dot{\gamma_{xy}}(0) = -tgradf(x), \quad t > 0$ se sigue que

$$\begin{aligned} f(y) - f(x) \quad &= df(\dot{\gamma_{xy}}(u_0)), \\ &\leq df(X_k) + rd^2(x,z), \\ &= df(-tgradf(x)) + rt^2 \parallel gradf(x) \parallel^2, \\ &= df(-tgradf(x)) + rt^2 \parallel gradf(x) \parallel^2, \\ &= -t \parallel gradf(x) \parallel^2 + rt^2 \parallel gradf(x) \parallel^2, \\ &= t \parallel gradf(x) \parallel^2 (-1 + rt), \end{aligned}$$

esto muestra que existe $t > 0$ tal que la desigualdad del proceso iterativo es satisfecha $(-1 + rt \leq -\epsilon)$. Como $r$ es estrictamente positivo y finito y $0 < \epsilon < 1$ la desigualdad es factible y da $t < \frac{1-\epsilon}{r}$, por consiguiente escogiendo $t_i$ como en las reglas del procedimiento, obtenemos:

$$f(x_{i+1}) - f(x_i) \leq -\epsilon t \parallel gradf(x_i) \parallel^2,$$

luego si $\parallel gradf(x_i) \parallel > 0$ entonces para cada $i \in N^*$ tenemos que

$$f(x_{i+1}) - f(x_i) < 0,$$

así $\{f(x_i)\}$ es decreciente.

Por otra parte tenemos que $f$ es acotado inferiormente así $\lim\limits_{i\to\infty} f(x_i)$ existe y por tanto

$$\lim_{i\to\infty} f(x_{i+1}) - f(x_i) = 0,$$

esto implica que

$$\lim_{i\to\infty} \parallel gradf(x_i) \parallel = 0.$$

## 5.5. Convergencia de $\{x_i\}$ a un punto crítico.

Una función es cual satisface las condiciones del teorema anterior puede o no tener puntos mínimos o un punto silla. El teorema muestra que el método del gradiente da la convergencia o a la cota inferior $inf\{f(x)\}$ o a un valor crítico de $f$. La convergencia de la sucesión $\{x_i\}$ a un punto crítico $x_*$ puede también ocurrir, sin embargo es difícil calcular las reglas de convergencia solo bajo las condiciones del teorema anterior.

**Teorema 5.5.1.** *Sea $f : M \to \mathbb{R}$ una función $C^2$, cuyo Hessiano satisface las condiciones*

$$a \parallel X \parallel^2 \leq Hessf(X,X) \leq b \parallel X \parallel^2; \quad b \geq a \geq 0, \forall X \in \mathfrak{X}(M).$$

1. *La función $f$ tiene un único punto mínimo $x_*$.*

2. *Si en el proceso iterativo el numero $t_i$ son establecidos como antes, entonces tenemos*
$$\lim_{i\to\infty} f(x_i) = f(x_*), \quad \lim_{i\to\infty} x_i = x_*,$$
*para un punto inicial estimado $x_1$ de $x_*$.*

3. *Las estimaciones siguientes de la regla de convergencia cumple:*
$$f(x_i) - f(x_*) \leq q^{i-1}(f(x_1 - f(x_*))),$$
$$d(x, x_*) \leq Cq^{\frac{i-1}{2}},$$
*con $C < \infty, \quad 0 < q < 1$.*

**Teorema 5.5.2.** *Una función $C^2$ $f : M \to \mathbb{R}$ es estrictamente convexa si y sólo si existe una constante $m$ positiva tal que*

$$Hessf(X,X) \geq m \parallel X \parallel^2, \forall X \in \mathbb{X}(M).$$

Demostración: ver [1]

La condición impuesta sobre el Hessiano de $f$ (implica que $Hessf$ es una métrica de Riemann sobre $M$ el cual induce sobre cada espacio tangente una norma la cual es equivalente a la norma inducida por g), así tendríamos que f es estrictamente convexa y de allí la existencia de un único punto mínimo $x_*$. Como $f$ es continua, es suficiente probar que $\lim_{i\to\infty} x_i = x_*$.

En efecto; por Taylor tenemos que

$$f(x_*) = f(x) + df(\dot{\gamma}_{xx_*}(0)) + \frac{1}{2}Hessf(\dot{\gamma}_{xx_*}(t_0), \dot{\gamma}_{xx_*}(t_0)), \quad t_0 \in [0,1]$$

así,

$$f(x) - f(x_*) = -df(\dot{\gamma}_{xx_*}(0)) - \frac{1}{2}Hessf(\dot{\gamma}_{xx_*}(t_0), \dot{\gamma}_{xx_*}(t_0)).$$

Utilizando la hipótesis se observa que:

$$f(x) - f(x_*) \leq \parallel gradf(x) \parallel d(x, x_*) - \frac{1}{2}d^2(x, x_*).$$

Aquí $\gamma_{xx_*}(t), \quad t \in [0, 1]$ es una geodésica minimal la cual pasa por los puntos $x$ y $x_*$.

Además nótese que si $\beta_{x_*x}(t) = \gamma_{xx_*}(1 - t), \quad t \in [0, 1]$ y $gradf(x_*) = 0$ implica que:

$$f(x) - f(x_*) = \frac{1}{2}Hessf(\dot{\beta}_{x_*x}(t_1), \dot{\beta}_{x_*x}(t_1)).$$

Así tenemos que

$$\frac{a}{2}d^2(x, x_*) \leq f(x) - f(x_*) \leq \frac{b}{2}d^2(x, x_*),$$

así

$$\frac{a}{2}d^2(x, x_*) \leq \parallel gradf(x) \parallel d(x, x_*) - \frac{a}{2}d^2(x, x_*),$$

o

$$d(x, x_*) \leq \frac{\parallel gradf(x) \parallel}{a},$$

también tenemos

$$d^2(x, x_*) \geq \frac{2}{b}(f(x) - f(x*)),$$

luego, usando todo esto podemos poner

$$\parallel gradf(x) \parallel^2 \geq a(1 + \frac{a}{b})(f(x) - f(x_*)),$$

y del teorema anterior sabemos que

$$f(x_{i+1}) - f(x_i) \leq -\epsilon t_i \parallel gradf(x_i) \parallel^2,$$

así encontramos que

$$f(x_{i+1}) - f(x_i) \leq -\epsilon t_i a(1 + \frac{a}{b})(f(x_i) - f(x_*)), \tag{5.1}$$

ahora bien, si consideramos las condiciones que le estamos poniendo a el $Hessf$ tenemos que:

$$\begin{aligned} f(x) - f(x_i) &= df(\dot{\gamma}_{x_ix}(0)) + \frac{1}{2}Hessf(\dot{\gamma}_{x_ix}(0), \dot{\gamma}_{x_ix}(0)), \\ &\leq -t \parallel gradf(x_i) \parallel^2 + \frac{t^2b}{2} \parallel gradf(x_i) \parallel^2, \\ &= -t(1 - \frac{tb}{2}) \parallel gradf(x_i) \parallel^2 . \end{aligned}$$

Así, dado $\epsilon > 0$ y tomando $t < \bar{t} = \frac{2(1-\epsilon)}{b}$ la condición antes escrita satisface

$$f(x) - f(x_i) \leq \epsilon t df(X_i), \quad X_i = -gradf(x_i).$$

Si en 5.1 sumamos y restamos en el lado izquierdo $f(x_*)$ y despejando $f(x_{i+1}) - f(x_*)$ tenemos lo siguiente

$$f(x_{i+1}) - f(x_*) \leq [1 - \epsilon t_i a(1 + \frac{a}{b})](f(x_i) - f(x_*)) \leq q(f(x_i) - f(x_*)),$$

donde $q = 1 - \epsilon \bar{t} a(1 + \frac{a}{b}) < 1$ entonces de forma iterativa se obtiene

$$f(x_i) - f(x_*) \leq q^{i-1}(f(x_1) - f(x_*)),$$

reemplazando $\bar{t} = \frac{2(1-\epsilon)}{b}$ podemos deducir que el valor mínimo de $q$ se obtiene en $\epsilon = \frac{1}{2}$.

Por otro lado, sabemos que

$$d(x_i, x_*) \leq (\frac{2}{a})^{\frac{1}{2}}(fx_i - f(x_*))^{\frac{1}{2}} \leq (\frac{2}{a})^{\frac{1}{2}}(fx_1 - f(x_*))^{\frac{1}{2}} q^{\frac{i-1}{2}} = cq^{\frac{i-1}{2}},$$

con $q < 1$.

Además, nótese que:

$$\begin{aligned} d^2(x_i, x_{i+1}) &\leq t_i^2 \parallel gradf(x_i) \parallel^2 \leq \frac{t_i}{\epsilon}(f(x_i - f(x_{i+1}))), \\ &\leq \frac{t_i}{\epsilon}(f(x_i - f(x_*)) \leq \frac{t_{max}}{\epsilon} q^{i-1}(f(x_i) - f(x_*)) \leq c_1 q^{i-1}, \end{aligned}$$

donde $t_{max}$ es el máximo valor admitido de $t$.

Para $j > i$ se tiene:

$$d(x_i; x_j) \leq \sum_{s=1}^{j-1} d(x_s, x_{s+1}) \leq C_1^{\frac{1}{2}} \sum_{s=1}^{j-1} q^{\frac{s-1}{2}} \leq C_1^{\frac{1}{2}} \frac{q^{\frac{i-1}{2}}}{1 - q^{\frac{1}{2}}}.$$

Luego;

$$\lim_{i \to \infty} d(x_i; x_j) = 0,$$

es decir, la sucesión $x_i$ converge a $x_*$.

## 5.6. Algoritmo del Método del Gradiente para la esfera.

Las geodésicas de $S$ son grandes circunferencias $\gamma : \mathbb{R} \to S$ dadas por

$$\gamma \begin{cases} x = acos(t) + lsen(t), \\ y = bcos(t) + msen(t)), \\ z = ccos(t) + nsen(t), \end{cases}$$

donde $\overrightarrow{u} = a\overrightarrow{i} + b\overrightarrow{j} + c\overrightarrow{k}$, $\overrightarrow{v} = l\overrightarrow{i} + m\overrightarrow{j} + \overrightarrow{k}$ y además $\| \overrightarrow{u} \| = 1$, $\| \overrightarrow{v} \| = 1$, $< \overrightarrow{u}, \overrightarrow{v} > 0$.

Las geodésicas $\gamma$ son tal que $\gamma(0) = u$ y $\dot{\gamma}(0) = v$.

Sea $f : \mathbb{R}^3 \to \mathbb{R}$ una función $C^2$ y sea $f_S$ la restricción de $f$ a la esfera $S$. Además nótese que las grandes circunferencias $\gamma(\mathbb{R})$ son un conjunto compacto así tenemos que la función $\varphi(t) = f(\gamma(t))$ tiene mínimo y máximo global, así el algoritmo para la esfera es el siguiente:

1. Para $i = 1$ fijamos el punto inicial $M_1 = (x_1, y_1, z_1) \in S$.
2. Determinamos el vector $X_1 = (l_1, m_1, n_1) \in T_{M_1}M$ tal que
$$df_S(M_1)(X_1) < 0.$$
3. Sea $\varphi : [0, +\infty] \to \mathbb{R}$ dada por
$$\varphi(t) = f(x_1cos(t) + l_1sen(t), y_1cos(t) + m_1sen(t), z_1cos(t) + n_1sen(t)),$$
así determinamos el valor de $t_1$ como la solución mas pequeña estrictamente positiva de la ecuación $\varphi'(t) = 0$ para el cual $\varphi(t_1)$ es un mínimo.
4. Calculamos $M_{i+1} = (x_{i+1}, y_{i+1}, z_{i+1})$ donde
$$x_{i+1} = x_icos(t_i) + l_isen(t_i),$$
$$y_{i+1} = y_icos(t_i) + m_isen(t_i),$$
$$z_{i+1} = z_icos(t_i) + n_isen(t_i).$$
5. Si $M_{i+1}$ satisface el criterio de convergencia dado, paramos. Si no, $i = i + 1$ y hacemos 2.

# Capítulo 6

# Conclusiones

Se han presentado una serie de teoremas y conceptos sobre algunos tópicos específicos de la geometría diferencial que sirvieron como base para analizar prácticamente sus adaptaciones en variedades Riemannianas, en donde los conceptos y teoría desarrollada no deben depender de la estructura que conocemos en $R^n$. Esta fue la motivación para hacer primeramente una revisión de conceptos en $R^n$ en capítulos iniciales y luego pasar a un espacio más general de tal forma que quede mejor entendible la secuencia mostrada. En esta fase se uso principalmente las referencias [2] y [4].

Posteriormente se hizo una revisión del concepto de geodésica y estudiamos la ecuación diferencial que las define, lo cual nos permite comprender la dificultad computacional que puede encontrarse en algunos métodos de optimización que se adapten a variedades, pues en muchos casos se necesitaría resolver esta ecuación diferencial en cada iteración principal del algoritmo usado.

Finalmente presentamos y detallamos dos teoremas de convergencia sobre el método de mayor descenso o método de cauchy en variedades Riemannianas que aparecen en [1], de lo cual observamos que el método presenta la misma forma como es usado en $R^n$ con la variante de necesitar la definición del mapeo exponencial y una búsqueda específica.

Cabe destacar que el trabajo presentado constituye una herramienta indispensable para poder abarcar otras metodologías en futuros trabajos, como por ejemplo, los citados en las referencias [5] y [6] en las cuales se abordan otras metodologías que ameritan una comprensión inicial de los aspectos aquí abordados.

# Bibliografía


[1] C. Udriste. Convex Functions and Optimization Methods on Riemannian Manifolds. Springer N.Y.1998.

[2] Cordero, Fernández y Gray. Geometrial Diferencial de Curvas y Supeficies. 1995

[3] R.T. Rockafellar. Convex Analysis. Princeton University. 1970.

[4] Manfredo P. Do Carmo.Geometria Riemanniana. 2005

[5] O.P. Ferreira and P.R. Oliveira. Proximal POint methods on Riemannian Manifolds. IME. Brasil. 2000.

[6] O.P. Ferreira and B.F. Svaiter. Kantorovich´s Theorem on Newton´s Method in Riemann Manifolds. Journal on Complexity. Vol. 18. nro.1. pag. 304-329. 2002.