

**UNIVERSIDAD CENTROCCIDENTAL
"LISANDRO ALVARADO"**

Decanato de Ciencias y Tecnología
Licenciatura en Ciencias Matemáticas



**"ECUACIÓN DE EVOLUCIÓN DEL MÉTODO DE DESCENSO
ACELERADO SOBRE UNA VARIEDAD RIEMANNIANA; EL CASO
CONVEXO"**

TRABAJO ESPECIAL DE GRADO PRESENTADO POR

BR. ESTHER LORENA GONZÁLEZ HERNÁNDEZ

COMO REQUISITO FINAL

PARA OBTENER EL TÍTULO DE LICENCIADO

EN CIENCIAS MATEMÁTICAS

ÁREA DE CONOCIMIENTO: **OPTIMIZACIÓN Y GEOMETRÍA RIEMANNIANA**

TUTOR: DR. EIBAR R. HERNÁNDEZ R.

Universidad Centroccidental "Lisandro Alvarado"

**"ECUACIÓN DE EVOLUCIÓN DEL MÉTODO DE DESCENSO
ACELERADO SOBRE UNA VARIEDAD RIEMANNIANA; EL CASO
CONVEXO"**

Br. Esther Lorena González Hernández

Tutor: Dr. Eibar R. Hernández R.

RESUMEN

En el presente trabajo se desarrollará el artículo [1], en el cual se estudia la vecindad asintótica de las trayectorias de la ecuación de evolución del método de descenso acelerado en variedades Riemannianas, la cual viene dada por:

$$\dot{x}(t) + \text{grad}\phi(x(t)) = 0.$$

Se mostrará como la convexidad de la función objetivo ϕ , contribuye a establecer la convergencia de las trayectorias, cuando el tiempo tiende a infinito, hacia los puntos que minimizan la función objetivo. Además, el trabajo mostrará ilustraciones numéricas para la función de Rosenbrock.

A Dios todopoderoso, la Misión Adgentes, mis padres y hermanos.

Riemann reconoció, al igual que su auspiciador Gauss que ya fuera como un medio de expresión o como la adopción de una creencia en realidad falsa, los dogmas de corte aristotelico y euclideanos son obstaculos que aprisionan a la mente en un mundo falso limitandola a solo porder asomarse con impotencia a lo que la víctima creyera ser un mundo externo real.

AGRADECIMIENTOS

Agradezco primeramente a Dios por haber permitido la culminación de este trabajo.

A mis padres por ser mis mejores amigos y permitirme que les caliente la cabeza con todas mis dificultades, además gracias por ser tan creativos e inteligentes y por haberme aportado un poco de todas esas virtudes en mi mezcla genética.(Los quiero!!!).

A mis hermanos por animarme las veces que me deprimía y me hacían reír cuando eran necesario, especialmente a mi bella hermana Mariena y a su esposo Blas, por su total apoyo.

Muchas gracias a mi tutor el Dr. Eibar Hernández por su guía y apoyo continuos en mi carrera, y también por soportar mis ocurrencias siempre con una sonrisa .

A mi hija Karlita por hacerme compañía las veces que me sentía sola en la residencia.

A mi amiga Aleja por tenerme paciencia las veces que me estresaba.

A mi amiga Maria José por tener siempre unas palabras de apoyo cuando lo necesitaba.

A mi comunidad, por aceptarme como soy y quererme así, sin ustedes estar sola en Bqto no hubiese sido lo mismo. Tengo que decirles que sé que han puesto todo su amor y lo aprecio más de lo que parece.

Un montón de gracias a todos mis tíos, primos, amigos, compañeros y demás personas que en algún momento me apoyaron y animaron para culminar mi carrera.

ÍNDICE

Agradecimientos	i
Introducción	1
Capítulo 1. Preliminares	3
1.1. En El Espacio Euclídeo	3
1.1.1. Conceptos y Resultados	3
1.1.2. Método de Descenso Acelerado Euclídeo	6
1.1.3. Teoremas de Existencia, Unicidad y Extensión de Soluciones de E.D.O.	7
1.2. En Variedades de Riemann	9
1.2.1. Algunos resultados Preliminares	9
1.2.2. Método de Descenso Acelerado en Variedades de Riemann	18
1.2.3. Ecuación de Evolución en Variedades de Riemann	19
Capítulo 2. Desarrollo	27
2.1. Caso Convexo	31
Capítulo 3. Ejemplo: Función de Rosenbrock	39
Conclusiones	47
Referencias	48

INTRODUCCIÓN

En el estudio de la convexidad Riemanniana se tienen tres aspectos relevantes. En primer lugar, la convexidad Riemanniana depende fuertemente de la métrica asociada, tanto por su relación con las geodésicas, como por su relación con la conexión Riemanniana. Segundo, la convexidad Riemanniana es independiente del sistema de coordenadas (para el caso C^2), lo que hace que esta pueda ser conectada fácilmente con la computación simbólica. Tercero, la estructura Riemanniana está involucrada en la resolución de los problemas, por medio de las métricas, las conexiones, la geodésica y la curvatura. Estas razones justifican los esfuerzos por extender y generalizar la teoría de optimización sobre espacios Euclídeos a variedades de Riemann, sin embargo, últimamente se están planteando otros enfoques considerando la llamada ecuación de Evolución, obteniéndose por medio de ésta, avances notablemente importantes, que permiten obtener resultados donde otros métodos fallan.

La ecuación de Evolución ha sido estudiada por Udriste ver [5], bajo el nombre de línea de menos gradiente. Tal ecuación está generalizada a Variedades de Riemann con el nombre de ecuación de descenso gradiente y tal sistema está estrechamente relacionado con el problema de minimizar una función objetivo.

Sea $(M, \langle \cdot, \cdot \rangle)$ una variedad dotada de una métrica Riemanniana. La longitud de una curva es denotada por $\ell(\cdot)$ y la distancia por $\rho(\cdot, \cdot)$. Sea $\phi : M \rightarrow \mathbb{R}$ una función diferenciable, denótese por $grad\phi(x)$ el gradiente con respecto al producto interno Riemanniano $\langle \cdot, \cdot \rangle_x$, lo cual significa que para cada $u \in T_x M$, el espacio tangente a la Variedad en el punto x , se cumple que:

$$d\phi(x)u = \langle grad\phi(x), u \rangle_x.$$

Supongamos que (M, ρ) es completo, ϕ es C^1 con gradiente Lipschitz continuo y ϕ es acotada inferiormente. Considérese la siguiente ecuación de Evolución

$$\dot{x}(t) + grad\phi(x(t)) = 0$$

Es posible probar que para $x_0 \in M$ el problema de Cauchy generado tiene una única solución $x(\cdot)$ definida en $[0, +\infty)$, y además esta solución es continuamente diferenciable. También se prueba que

cualquier solución $x(\cdot)$ de la ecuación de Evolución cumple que $\phi(x(\cdot))$ decrece, converge y además $\|\dot{x}(t)\|_{x(t)} \in L^2(0, +\infty)$.

Si se supone además que ϕ es convexa en M , que el $\operatorname{argmin}\phi$ es no vacío, y tomando y en el conjunto $L = \{y \in M / \phi(y) \leq \phi_\infty\}$, donde $\phi_\infty = \lim_{t \rightarrow +\infty} \phi(x(t))$, entonces se prueba que la función $t \mapsto \rho(x(t), y)$ decrece y converge. Finalmente bajo todas las condiciones mencionadas anteriormente cada solución $x(\cdot)$ de la ecuación de evolución tiene un punto límite $\bar{x} \in \operatorname{argmin}\phi$. Como resultado adicional se estudiará la función de Rosenbrock, la cual no es convexa en el sentido Euclídeo, pero al dotarse \mathbb{R}^2 de una métrica de Riemann conveniente, hace convexa a tal función en el sentido Riemanniano, y utilizando un algoritmo desarrollado en base de los teoremas previos, se encuentra el mínimo para tal función.

CAPÍTULO 1

PRELIMINARES

§1.1. EN EL ESPACIO EUCLÍDEO

§1.1.1. Conceptos y Resultados

DEFINICIÓN 1.1.1. Un subconjunto C de \mathbb{R}^n es llamado convexo si y sólo si

$$(1 - \lambda)x + \lambda y \in C,$$

para cada $x, y \in C, \lambda \in (0, 1)$.

DEFINICIÓN 1.1.2. Sea $C \subset \mathbb{R}^n$ un conjunto convexo. Una función $f : C \rightarrow \mathbb{R}$ se llama convexa cuando, para cualquier $x, y \in C$ y $t \in [0, 1]$, se tiene

$$f((1 - t)x + ty) \leq (1 - t)f(x) + tf(y). \tag{1.1.1}$$

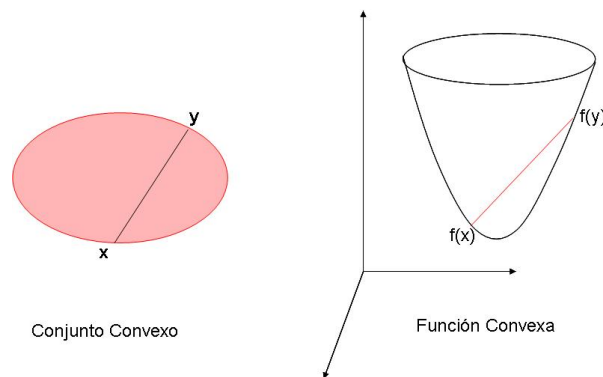


FIGURA 1.1.1: Izquierda: Un conjunto convexo; Derecha: Una función convexa

DEFINICIÓN 1.1.3. Sea $C \subset \mathbb{R}^n$ un conjunto convexo cerrado. La dirección $s \in \mathbb{R}^n$ es llamada normal a C en $x \in C$, cuando

$$\langle s, y - x \rangle \leq 0 \quad \forall y \in C. \tag{1.1.2}$$

El conjunto de todas estas direcciones es llamado cono normal a C en x .

DEFINICIÓN 1.1.4. Sea $E \subset \mathbb{C}$. Diremos que la función $f : E \rightarrow \mathbb{C}$ pertenece a los espacios de las funciones integrables a la p si $f \in L^p$ donde

$$L^p = \left\{ f : E \rightarrow \mathbb{C} / \int_E |f|^p < \infty \right\} \text{ donde } \|f\| = \left(\int_E |f|^p \right)^{\frac{1}{p}}. \quad (1.1.3)$$

PROPOSICIÓN 1.1.1. Sea $P_x = (x_1, x_2), P_u = (u_1, u_2)$ y $P'_x = (x'_1, x'_2)$ tres puntos en \mathbb{R}^2 , con $u \in [x, x']$. Entonces las tres propiedades siguientes son equivalentes:

- P_u esta por debajo de $P_x P'_x$
- Pendiente $(P_x P_u) \leq$ pendiente $(P_x P'_x)$
- Pendiente $(P_x P'_x) \leq$ pendiente $(P_u P'_x)$

Demostración: ver [3]

PROPOSICIÓN 1.1.2. (Criterio de pendiente creciente)

Sea I un intervalo no vacío de \mathbb{R} . Una función $f : I \rightarrow \mathbb{R}$ es convexa en I si y sólo si, para todo $x_0 \in I$, la función - pendiente

$$x \mapsto \frac{f(x) - f(x_0)}{x - x_0} := s(x), \quad (1.1.4)$$

es creciente en $I \setminus x_0$.

Demostración: ver [3]

TEOREMA 1.1.1. Sea p, q números reales mayores que 1 tales que $\frac{1}{p} + \frac{1}{q} = 1$.

Para cuales quiera números complejos x_i, y_i con $(i = 1, 2, \dots)$ se tiene que

$$\sum_{i=1}^n |x_i y_i| \leq \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}} \left(\sum_{i=1}^n |y_i|^q \right)^{\frac{1}{q}} \quad (1.1.5)$$

Además

$$\int_E |x_i y_i| \leq \left(\int_E |x_i|^p \right)^{\frac{1}{p}} \left(\int_E |y_i|^q \right)^{\frac{1}{q}} \quad (1.1.6)$$

Demostración: ver [6]

DEFINICIÓN 1.1.5. Sea $f : U \rightarrow \mathbb{R}$ una función definida en un abierto $U \subset \mathbb{R}^n$. Para cada $i = 1, 2, \dots, n$, la i -ésima derivada parcial de f en el punto $a = (a_1, a_2, \dots, a_n) \in U$ es el número

$$\frac{\partial f}{\partial x_i}(a) = \lim_{t \rightarrow 0} \frac{f(a + te_i) - f(a)}{t} = \lim_{t \rightarrow 0} \frac{f(a_1, \dots, a_i + t, \dots, a_n) - f(a)}{t}.$$

En caso que este límite exista.

DEFINICIÓN 1.1.6. Sea $f : U \rightarrow \mathbb{R}$ una función que posee las n derivadas parciales en todos los puntos del abierto $U \subset \mathbb{R}^n$. Quedan entonces definidas n funciones

$$\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} : U \rightarrow \mathbb{R}, \quad \text{donde} \quad \frac{\partial f}{\partial x_i} : x \mapsto \frac{\partial f}{\partial x_i}(x)$$

Si estas funciones son continuas en U , diremos que f es una función de clase C^1 .

Una aplicación $f : U \rightarrow \mathbb{R}^n$, definida en un abierto $U \subset \mathbb{R}^m$, se dice de clase C^1 , cuando cada una de sus funciones coordenadas $f_1, f_2, \dots, f_n : U \rightarrow \mathbb{R}$ es de clase C^1 .

DEFINICIÓN 1.1.7. Una función $f : U \rightarrow \mathbb{R}$, definida en un abierto $U \subset \mathbb{R}^n$ se dice diferenciable en un punto $a \in U$ cuando cumple las siguientes condiciones:

1. Existan las derivadas parciales $\frac{\partial f}{\partial x_1}(a), \dots, \frac{\partial f}{\partial x_n}(a)$.
2. Para todo $v = (\alpha_1, \dots, \alpha_n)$ tal que $a + v \in U$, se tiene

$$f(a + v) - f(a) = \sum_{i=1}^n \frac{\partial f}{\partial x_i} \cdot \alpha_i + r(v), \quad \text{donde} \quad \lim_{|v| \rightarrow 0} \frac{r(v)}{|v|} = 0. \quad (1.1.7)$$

DEFINICIÓN 1.1.8. Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$ una función diferenciable. Se dice que un punto $x_0 \in \mathbb{R}^n$, es un punto crítico de f , si se satisface:

$$\nabla f(x_0) = 0. \quad (1.1.8)$$

DEFINICIÓN 1.1.9. Sea $f : U \rightarrow \mathbb{R}$ una función que posee las derivadas parciales $\frac{\partial f}{\partial x_1}(x), \dots, \frac{\partial f}{\partial x_n}(x)$ en todo punto x de el abierto $U \subset \mathbb{R}^n$. La j -ésima derivada parcial de la función $\frac{\partial f}{\partial x_i} : U \rightarrow \mathbb{R}$ en un punto $x \in U$ sera indicado por

$$\frac{\partial^2 f}{\partial x_j \partial x_i}(x) = \frac{\partial}{\partial x_j} \left(\frac{\partial f}{\partial x_i} \right) (x), \quad i, j = 1, \dots, n$$

Si estas derivadas parciales de segundo orden existen en cada punto $x \in U$, tendremos $\frac{\partial^2 f}{\partial x_j \partial x_i} : U \rightarrow \mathbb{R}$.

Cuando tales funciones son continuas, diremos que f es de clase C^2 .

LEMA 1.1.1. Sea $r : B \rightarrow \mathbb{R}$ de clase C^2 en la bola abierta $B \subset \mathbb{R}^n$, de centro 0.

Si $r(0) = \frac{\partial r}{\partial x_i}(0) = \frac{\partial^2 r}{\partial x_i \partial x_j}(0) = 0$ para cualquier $i, j = 1, 2, \dots$, entonces

$$\lim_{v \rightarrow 0} \frac{r(v)}{|v|^2} = 0. \quad (1.1.9)$$

Demostración: ver [2]

§1.1.2. Método de Descenso Acelerado Euclídeo

A través de la historia de la optimización han existido distintos problemas para los cuales se ha querido hacer un análisis de convergencia, por esta razón se han creado distintos métodos para hallar el mínimo de una función establecida.

En esta sección se les hablará un poco de el método de descenso acelerado y se explicará las reglas de medida de paso. Primero habrá que considerar un problema de minimización irrestricta de una función de clase C^1 , $f : \mathbb{R}^n \rightarrow \mathbb{R}$, se iniciará con un punto inicial x_0 y una sucesión de vectores generada x_1, x_2, \dots , tal que, f decrecerá en cada iteración, es decir:

$$f(x_k) > f(x_{k+1}) \quad k = 0, 1, \dots \tag{1.1.10}$$

Nuestro objetivo será que en cada iteración se mejorará la solución estimada y así llegar al mínimo de dicho problema.

Ahora comenzaremos con el algoritmo del método anteriormente mencionado.

Fijemos un vector $x_k \in \mathbb{R}^n$ con $\nabla f(x_k) \neq 0$, para algún $k \in \{0, 1, \dots\}$, consideremos la semirecta de vectores

$$x_{k+1} = x_k - \alpha_k \nabla f(x_k) \quad \forall \alpha_k \geq 0 \tag{1.1.11}$$

Utilizando la expansión de primer orden de la serie de Taylor alrededor de x_k se obtendrá

$$\begin{aligned} f(x_{k+1}) &= f(x_k) + (\nabla f(x_k))^t (x_{k+1} - x_k) + o(\|x_{k+1} - x_k\|) \\ &= f(x_k) - \alpha_k \|\nabla f(x_k)\|^2 + o(\alpha_k \|\nabla f(x_k)\|) \\ &= f(x_k) - \alpha_k \|\nabla f(x_k)\|^2 + o(\alpha_k) \end{aligned}$$

donde $o(\alpha_k) = \zeta(\alpha_k) \Rightarrow \lim_{\alpha_k \rightarrow 0} \frac{\zeta(\alpha_k)}{\alpha_k} = 0$, luego para α_k positivo y suficientemente pequeño se obtendrá

$$f(x_{k+1}) = f(x_k) - \alpha_k \|\nabla f(x_k)\|^2$$

de lo anterior

$$f(x_{k+1}) < f(x_k) \tag{1.1.12}$$

Si $\nabla f(x_{k+1}) \neq 0$, entonces se precede de manera análoga a lo visto anteriormente.

Si $\nabla f(x_{k+1}) = 0$, entonces el algoritmo se detiene.

Generalizando más este algoritmo, consideraremos la semirecta del vector

$$x_{k+1} = x_k - \alpha_k d_k \quad \alpha_k \geq 0 \quad (1.1.13)$$

donde la dirección del vector $d_k \in \mathbb{R}^n$ se le hará un ángulo mayor de 90° con el $\nabla f(x_k)$, es decir

$$(\nabla f(x_k))^t \cdot d_k < 0, \quad (1.1.14)$$

por Taylor se obtendrá,

$$f(x_{k+1}) = f(x_k) - \alpha_k \cdot \nabla f(x_k)^t \cdot d_k + o(\alpha_k) \quad (1.1.15)$$

del mismo modo a lo anterior, cuando α es positiva y suficientemente pequeña

$$f(x_k + \alpha_k d_k) < f(x_k). \quad (1.1.16)$$

Estas observaciones construyen el importante algoritmo clásico

$$x_{k+1} = x_k - \alpha_k d_k \quad k = 0, 1, \dots \quad (1.1.17)$$

el cual cuando $\nabla f(x^k) \neq 0$ la dirección del vector d^k es escogida tal que $\nabla f(x^k)^T \cdot d^k < 0$.

El tamaño del paso α^k es construido positivo. En el caso en que $\nabla f(x^k) = 0$, el método se detendrá, es decir $x_{k+1} = x_k$.

En vista de la relación $(\nabla f(x_k))^t \cdot d_k < 0$, de la dirección d_k y el $\nabla f(x_k)$ es común llamar a estos algoritmos de la ecuación (1.1.17) métodos gradientes, en este trabajo se le llamará Método de Descenso Acelerado.

§1.1.3. Teoremas de Existencia, Unicidad y Extensión de Soluciones de E.D.O.

DEFINICIÓN 1.1.10. Dado un abierto U de \mathbb{R}^n . Interpretaremos cualquier aplicación $X : U \rightarrow \mathbb{R}^n$ como un campo de vectores en U , i.e. en cada punto $x \in U$, $X(x)$ se considera como el punto final del vector anclado en x .

Asociado a un campo vectorial $X : U \rightarrow \mathbb{R}^n$ se tiene la ecuación diferencial ordinaria

$$y' = X(y). \quad (1.1.18)$$

Una solución de (1.1.18) es cualquier función $\alpha : I \rightarrow U$, I es un intervalo en \mathbb{R} que satisface

- α es diferenciable en I .

- $\alpha'(t) = X(\alpha(t))$ para cada $t \in I$.

TEOREMA 1.1.2. (Peano)

Si el campo vectorial $X : U \rightarrow \mathbb{R}^n$ es continuo, entonces, para todo $t_0 \in \mathbb{R}$ y cada $y_0 \in U$, el problema de Cauchy

$$\begin{aligned} y' &= X(y) \\ y(t_0) &= y_0, \end{aligned}$$

admite al menos una solución α definida en $[t_0 - \delta, t_0 + \delta]$ para algún $\delta > 0$

Demostración ver: [7]

DEFINICIÓN 1.1.11. Una función $f : U \rightarrow \mathbb{R}^n$ es localmente Lipschitz, si para cada $x_0 \in U$ existe una constante $K > 0$ y una vecindad $V \subset U$ de x_0 tales que

$$\|f(x) - f(x_0)\| \leq K \|x - x_0\| \quad \forall x \in V$$

TEOREMA 1.1.3. (Cauchy-Picard)

Si el campo vectorial $X : U \rightarrow \mathbb{R}^n$ es localmente Lipschitz, entonces para cada t_0 y cada $y_0 \in U$ existe una constante $\delta > 0$ tal que

$$(\star) = \begin{cases} y' = X(y), \\ y(0) = y_0, \end{cases}$$

tiene una única solución definida en $[t_0 - \delta, t_0 + \delta]$.

Demostración: ver [7]

TEOREMA 1.1.4. Si I y J son intervalos conteniendo a 0 en su interior y si $\alpha : I \rightarrow U$ y $\beta : J \rightarrow U$ son soluciones de (??) entonces para cada $t \in I \cap J$ se cumple $\alpha(t) = \beta(t)$.

Demostración ver [7]

DEFINICIÓN 1.1.12. Se dice que un intervalo I es maximal para el sistema (\star) si dada cualquier solución $\alpha : J \rightarrow U$ de (\star) , se tiene que $J \subset I$. I es abierto y único en cada $x_0 \in U$, de allí denotaremos por $I(x_0) = (w_-(x_0), w_+(x_0))$ al intervalo maximal. Se llama solución maximal a la solución de (\star) definida en $I(x_0)$

TEOREMA 1.1.5. Sea $I(x_0) = (w_-(x_0), w_+(x_0))$ el intervalo máximo de definición de la solución α_{x_0} de (\star) . Si $w_+(x_0) < +\infty$ (resp. $w_-(x_0) < -\infty$), entonces para cualquier compacto $W \subset U$ existe $t^* \in I(x_0)$ próximo de $w_+(x_0)$ (resp. $w_-(x_0)$) tal que $\alpha_{x_0} \notin W$.

Demostración: ver [7]

§1.2. EN VARIEDADES DE RIEMANN

§1.2.1. Algunos resultados Preliminares

DEFINICIÓN 1.2.1. Una variedad diferenciable de dimensión n en un conjunto M es una familia de aplicaciones biunívocas $x_\alpha : U_\alpha \subset \mathbb{R}^n \rightarrow M$ de abiertos U_α de \mathbb{R}^n en M tal que:

1. $\bigcup_{\alpha \in \Lambda} x_\alpha(U_\alpha) = M$.
2. Para todo α, β , con $x_\alpha(U_\alpha) \cap x_\beta(U_\beta) = W \neq \emptyset$, los conjuntos x_α^{-1} y x_β^{-1} son abiertos en \mathbb{R}^n y las aplicaciones $x_\beta^{-1} \circ x_\alpha$ son diferenciables ver figura(1.2.1).
3. La familia $\{(U_\alpha; x_\alpha)\}$ es máxima que cumple con (1) y (2).

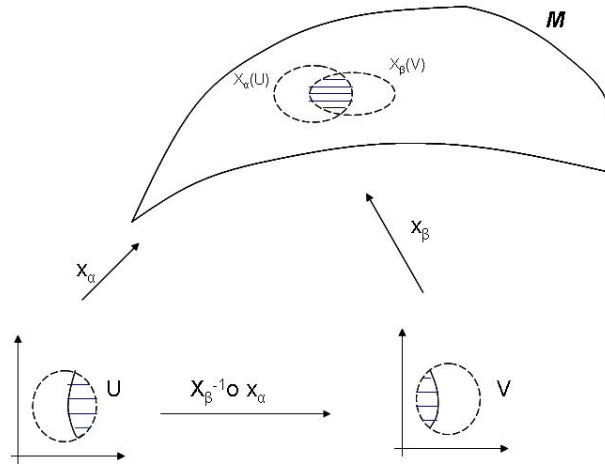


FIGURA 1.2.1

DEFINICIÓN 1.2.2. Sea $A \subset M$, A es abierta si y solo si $x_\alpha^{-1}(A \cap x_\alpha(U_\alpha))$ es abierto en $\mathbb{R}^n \forall \alpha$.

DEFINICIÓN 1.2.3. Sea M_1^n y M_2^m variedades diferenciables, una aplicación $\varphi : M_1 \rightarrow M_2$ es diferenciable en $p \in M_1$ sii $\forall y : V \subset \mathbb{R}^m \rightarrow M_2$ en $\varphi(p) \exists x : U \subset \mathbb{R}^n \rightarrow M_1$ en p tal que $\varphi(x(u)) \subset y(V)$ y

$$y^{-1} \circ \varphi \circ x : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m \text{ es diferenciable en } x^{-1}(p). \tag{1.2.1}$$

DEFINICIÓN 1.2.4. sea $\alpha : (-\varepsilon, \varepsilon) \rightarrow \mathbb{R}^n$ una curva diferenciable de \mathbb{R}^n con $\alpha(0) = p$. Escriba $\alpha(t) = (x_1(t), x_2(t), \dots, x_n(t))$ con $t \in (-\varepsilon, \varepsilon)$ y $(x_1, x_2, \dots, x_n) \in \mathbb{R}^n$, entonces $\alpha'(0) = (x'_1(0), x'_2(0), \dots, x'_n(0)) = v \in \mathbb{R}^n$.

Sea ahora la función diferenciable definida en una vecindad de p .

Podemos restringir f a una curva α y escribir la derivada direccional segunda en la dirección del vector $v \in \mathbb{R}^n$ como:

$$\left. \frac{d(f \circ \alpha)}{dt} \right|_{t=0} = \sum_{i=1}^n \left. \frac{\partial f}{\partial x_i} \right|_{t=0} \left. \frac{dx_i}{dt} \right|_{t=0} = \sum_{i=1}^n \left. \frac{\partial f}{\partial x_i} x'_i \right|_{t=0} = \left(\sum_{i=1}^n x'_i(0) \frac{\partial}{\partial x_i} \right) f \quad (1.2.2)$$

esto es:

$$\frac{d}{dt} (f \circ \alpha) (0) = \sum_{i=1}^n \frac{\partial f}{\partial x_i} (p) x'_i(0) = \sum_{i=1}^n x'_i(0) \frac{\partial f}{\partial x_i} (p) = \left[\sum_{i=1}^n \left(x'_i(0) \frac{\partial}{\partial x_i} \right) \right] f \quad (1.2.3)$$

DEFINICIÓN 1.2.5. Sea M una variedad diferenciable y $\alpha : (-\epsilon, \epsilon) \rightarrow M$ una curva diferenciable, suponga que $\alpha(0) = p$ y sea $D = \{f : M \rightarrow \mathbb{R}^n / f \text{ es diferenciable en } p\}$. El vector tangente a la curva α en $t = 0$ es la función $\alpha'(0) : D \rightarrow \mathbb{R}$ dada por:

$$\alpha'(0)(f) = \frac{d}{dt} (f \circ \alpha)(0) = \left(\sum_{i=1}^n \left(x'_i(0) \frac{\partial}{\partial x_i} \right) \right) f \quad (1.2.4)$$

donde $f \circ \alpha : (-\epsilon, \epsilon) \rightarrow \mathbb{R}^n$ es una curva diferenciable con $\alpha(0) = p$ y $\alpha(t) = (x_1(t), x_2(t), \dots, x_n(t))$, $(\frac{\partial}{\partial x_i})_0$ es el vector tangente en p sobre la curva coordenada $x_i \rightarrow x(0, \dots, 0, x_i, 0, \dots, 0)$

DEFINICIÓN 1.2.6. $T_p M = \{v : v \text{ es vector tangente a } M \text{ en } p\}$. Es conjunto de vectores tangentes a M en p . Este conjunto $T_p M$ con las operaciones usuales de funciones forman un espacio vectorial de dimensión n y, escogiendo una parametrización $X : U \rightarrow M$ se determina una base asociada a esta, la cual se denota por $\{(\frac{\partial}{\partial x_1})_0, \dots, (\frac{\partial}{\partial x_n})_0\}$ y es posible verificar que una estructura lineal en $T_p M$ así definida no depende de la parametrización. A $T_p M$ se le llama espacio tangente a M en p ver figura (1.2.2).

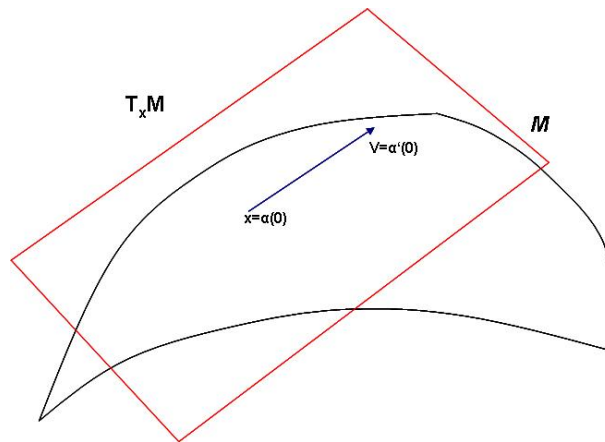


Ilustración de el plano tangente $T_x M$ en un punto x de M

FIGURA 1.2.2

DEFINICIÓN 1.2.7. sean M_1^n y M_2^m variedades diferenciables y sea $\varphi : M_1 \rightarrow M_2$ una aplicación diferenciable. Para cada $p \in M_1$ y cada $v \in T_p M_1$, escoja una curva diferenciable $\alpha : (-\varepsilon, \varepsilon) \rightarrow M_1$ con $\alpha(0) = p$ $\alpha'(0) = v$. Haga $\beta = \varphi \circ \alpha$. A la aplicación $d_{\varphi_p} : T_p M_1 \rightarrow T_{\varphi(p)} M_2$ dada por $d_{\varphi_p}(v) = \beta'(0)$ es una aplicación lineal que no depende de la escogencia de α .

Esta aplicación lineal $d\varphi_p$ se le llama diferencial de φ en p .

DEFINICIÓN 1.2.8. Sea M una variedad diferenciable y sea $TM = \{(p, v); p \in M, v \in T_p M\}$. Si dotamos a este conjunto de una estructura diferencial, este será una variedad de dimensión $2n$, dicho conjunto con tal estructura es llamado fibrado tangente; muy comunmente denotado por

$$TM = \bigcup_{p \in M} T_p M. \quad (1.2.5)$$

DEFINICIÓN 1.2.9. Sean M_1 y M_2 variedades diferenciables. Una aplicación

$$\varphi : M_1 \rightarrow M_2$$

es un difeomorfismo si es diferenciable y su inversa es diferenciable.

TEOREMA 1.2.1. sea $\varphi : M_1^n \rightarrow M_2^n$ una aplicación diferenciable y sea $p \in M_1$ tal que $d_{\varphi_p} : T_p M_1 \rightarrow T_{\varphi(p)} M_2$ es un isomorfismo, entonces φ es un difeomorfismo local en p .

Demostración: ver [4]

DEFINICIÓN 1.2.10. Un campo de vectores X en una variedad diferenciable M es una correspondencia que a cada punto $p \in M$ asocia un vector $X(p) \in T_p M$. En términos de aplicaciones, X es una aplicación de M en el fibrado tangente TM . Un campo es diferenciable si la aplicación $X : M \rightarrow TM$ es diferenciable.

Considerando una parametrización $U \subset \mathbb{R}^n \rightarrow M$ es posible escribir:

$$X(p) = \sum_{i=1}^n a_i(p) \frac{\partial}{\partial x_i}, \quad (1.2.6)$$

donde cada $a_i : U \rightarrow \mathbb{R}$ es una función en U y $\{\frac{\partial}{\partial x_i}\}_{i=1}^n$ es la base asociada al sistema de coordenada x , $i = 1, \dots, n$. Diremos que X es diferenciable si cada a_i lo es para alguna parametrización. A veces es conveniente utilizar la idea sugerida anteriormente para pensar un campo de vectores como una aplicación $X : D \rightarrow F$ donde D es el conjunto de funciones diferenciables en M y F el conjunto de funciones definidas en M , definida como sigue:

$$(Xf)(p) = \sum_{i=1}^n a_i(p) \frac{\partial f}{\partial x_i}(p), \quad (1.2.7)$$

donde f indica por un abuso de notación la expresión de f en la parametrización x . En verdad esta idea de vector como derivada direccional fue precisamente la manera como definimos la noción de vector tangente.

DEFINICIÓN 1.2.11. Una métrica Riemanianna (ó estructura Riemanianna) en una variedad diferencial M es una correspondencia que asocia a cada punto p de M un producto interno $\langle \cdot, \cdot \rangle_p$ (es decir una forma bilineal simétrica definida positiva) en el espacio tangente $T_p M$ que varía diferencialmente en el siguiente sentido sea $x : U \subset \mathbb{R}^n \rightarrow M$ un sistema de coordenadas locales en torno a p , con $x(x_1, x_2, \dots, x_n) = q \in x(U)$ y $\frac{\partial}{\partial x_i}(q) = dx(0, \dots, 1, \dots, 0)$ entonces, $\langle \frac{\partial}{\partial x_i}(q), \frac{\partial}{\partial x_j}(q) \rangle_q = g_{ij}(x_1, x_2, \dots, x_n)$ es una función diferenciable en U .

Una variedad diferenciable M con una métrica Riemanianna g se denomina Variedad Riemanianna y la denotamos por (M, G) , o solo por M si no surge confusión, donde G denota la matriz de representación de la métrica g .

DEFINICIÓN 1.2.12. El producto interno de dos vectores $u, v \in T_p M$ se escribe:

$$\langle u, v \rangle_p := g_p(u, v), \quad (1.2.8)$$

donde g_p es la métrica en el punto p .

DEFINICIÓN 1.2.13. Sea $v \in T_p M$ la norma de este vector esta definida como :

$$\|v\| := \sqrt{\langle v, v \rangle_p}. \quad (1.2.9)$$

DEFINICIÓN 1.2.14. Una aplicación diferenciable $c : I \rightarrow M$ de un intervalo abierto $I \subset \mathbb{R}$ en una variedad diferenciable M es llamada curva.

DEFINICIÓN 1.2.15. La longitud de una sección curva (trozo) de una curva plana $\alpha : [t_0, t_1] \rightarrow M$ que una $\alpha(t_0) = p'$ con $\alpha(t_1) = p$ es definida mediante la métrica como

$$L(\alpha) = \int_{t_0}^{t_1} \|\alpha'(t)\| dt. \quad (1.2.10)$$

DEFINICIÓN 1.2.16. La distancia Riemanianna entre dos puntos p' y p en M se define como:

$$d(p', p) = \inf\{L(\alpha) : \alpha \text{ curva suave que conecta a } p' \text{ con } p\}. \quad (1.2.11)$$

PROPOSICIÓN 1.2.1. Con la distancia d , M es un espacio métrico, esto es:

1. $d(p, r) \leq d(p, q) + d(q, r)$
2. $d(p, q) = d(q, p)$
3. $d(p, q) \geq 0$, y $d(p, q) = 0 \Leftrightarrow p = q$

Demostración: ver [4]

PROPOSICIÓN 1.2.2. La topología inducida por d en M coincide con la topología inicial de M .

Demostración : ver [4]

DEFINICIÓN 1.2.17. Una conexión afín ∇ en una variedad diferenciable M es una aplicación:

$$\nabla : X(M) \times X(M) \rightarrow X(M)$$

que se indica por $(X, Y) \xrightarrow{\nabla} \nabla_X Y$ 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 $X, Y, Z \in X(M)$ (conjunto de campos de vectores de clase C^∞ en M) y $f, g \in D(M)$ (conjunto de funciones de clase C^∞ definidas en M).

PROPOSICIÓN 1.2.3. Sea M una variedad diferenciable con una conexión afín ∇ . Entonces existe una única correspondencia que asocia a un campo vectorial X a lo largo de la curva diferencial $c : I \rightarrow \mathbb{R}$ otro campo vectorial $\frac{DX}{dt}$ a lo largo de c , denominado derivada covariante de X a lo largo de c tal que:

1. $\frac{D}{dt}(X + W) = \frac{DX}{dt} + \frac{DW}{dt}$.
2. $\frac{D}{dt}(fW) = \frac{df}{dt}W + f\frac{DW}{dt}$, donde W es un campo de vectores a lo largo de c y f es una función diferenciable en I .
3. Si W es inducido por el campo de vectores $Y \in X(M)$, es decir, $W(t) = Y(c(t))$, entonces $\frac{DW}{dt} = \nabla_{\frac{dc}{dt}} Y$.

Demostración: ver [4]

DEFINICIÓN 1.2.18. Sea M una variedad diferenciable con una conexión afín ∇ . Un campo vectorial X a lo largo de una curva $c : I \rightarrow M$ es llamado paralelo cuando $\frac{DX}{dt}(t) = 0$, para todo $t \in I$.

PROPOSICIÓN 1.2.4. Sea M una variedad diferenciable con una conexión afín ∇ . Sea $c : I \rightarrow M$ una curva diferenciable en M y V_0 un vector tangente a M en $c(t_0)$, $t_0 \in I$ (es decir $V_0 \in T_{c(t_0)}M$). Entonces existe un único campo de vectores paralelo X a lo largo de c tal que $X(t_0) = V_0$, ($X(t)$ es llamado transporte paralelo de V_0 a lo largo de c).

Demostración: ver [4]

DEFINICIÓN 1.2.19. Sea M una variedad diferencial con una conexión afín ∇ y una métrica Riemanniana \langle, \rangle . Una conexión se dice compatible con la métrica \langle, \rangle cuando para toda curva diferenciable c y cualquier par de campos de vectores paralelos X y X' a lo largo de c tenemos $\langle X, X' \rangle = \text{constante}$. En otras palabras decir que la conexión ∇ es compatible con la métrica \langle, \rangle , es decir que podemos diferenciar el producto interno con la regla del producto usual.

PROPOSICIÓN 1.2.5. Sea M una variedad Riemanniana. Una conexión ∇ es compatible con la métrica si y solo si para todo par X, X' de campo de vectores a lo largo de la curva diferencial $c : I \rightarrow M$ se cumple

$$\frac{d}{dt}\langle X, X' \rangle = \left\langle \frac{DX}{dt}, X' \right\rangle + \left\langle X, \frac{DX'}{dt} \right\rangle \quad t \in I. \quad (1.2.12)$$

Demostración: ver [4]

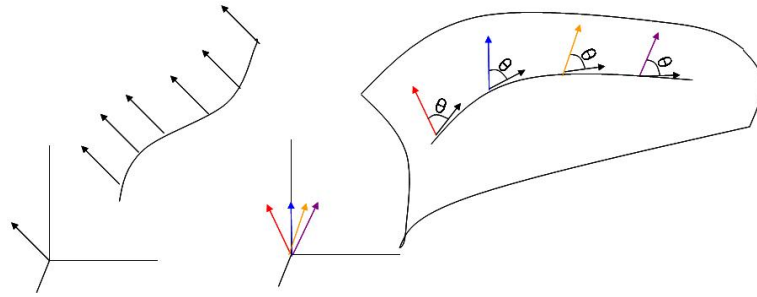


FIGURA 1.2.3: comparación de un transporte paralelo en el espacio Euclideo (Izquierda) y el transporte paralelo en el espacio Riemanniano (Derecha)

DEFINICIÓN 1.2.20. Un tensor T de orden r en una variedad Riemanniana es una aplicación multilinear

$$T : \underbrace{X(M) \times \dots \times X(M)}_{r \text{ factores}} \rightarrow D(M). \quad (1.2.13)$$

Esto quiere decir que, dados $Y_1, \dots, Y_r \in X(M)$, $T(Y_1, \dots, Y_r)$, es una función diferenciable en M donde T es lineal en cada argumento, esto es,

$$T(Y_1, \dots, fX + gY, \dots, Y_r) = fT(Y_1, \dots, X, \dots, Y_r) + gT(Y_1, \dots, Y, \dots, Y_r), \quad (1.2.14)$$

para todo $X, Y \in X(M)$ $f, g \in D(M)$.

COROLARIO 1.2.1. Para $p, q \in M$ y una curva parametrizada $I \rightarrow M$ con $\alpha(0) = p, \alpha(1) = q$. Denotando $\tau_\alpha : T_p(M) \rightarrow T_q(M)$ la aplicación que asigna a cada $v \in T_p(M)$ su transporte paralelo a lo largo de α por q se tiene que τ_α una isometría local.

Demostración: ver [4]

DEFINICIÓN 1.2.21. Una curva parametrizada $\gamma : I \rightarrow M$ es una geodésica en $t_0 \in I$ si $\frac{D}{dt}(\frac{d\gamma}{dt}) = 0$ en el punto t_0 ; si γ es geodésica en t , para cada $t \in I$, decimos que γ es una geodésica.

Si $[a, b] \subset I$ y $\gamma : I \rightarrow M$ es una geodésica, la restricción de γ a $[a, b]$ se llama el segmento geodésico uniendo $\gamma(a)$ con $\gamma(b)$.

DEFINICIÓN 1.2.22. Un segmento geodésico $\gamma : [a, b] \rightarrow M$ se llama minimizante si $\ell(\gamma) \leq \ell(c)$, donde $\ell(\cdot)$ indica la longitud de una curva y c es cualquier curva diferenciable por partes uniendo a $\gamma(a)$ con $\gamma(b)$.

DEFINICIÓN 1.2.23. Una curva integral de X en $x_0 \in M$ es una curva en x_0 tal que

$$\alpha'(t) = X(\alpha(t)) \quad \forall t \in I$$

donde $I \subset \mathbb{R}$ es un intervalo abierto, la imagen de la curva integral es llamado órbita o trayectoria.

La curva integral maximal $\alpha_x : I(x) \rightarrow M$ es definida en un intervalo abierto

$$I(x) = (w_-(x), w_+(x))$$

conteniendo $0 \in \mathbb{R}$.

Decimos además que α es una solución de la ecuación diferencial $\frac{d\alpha}{dt} = X(\alpha(t))$, con condición inicial $\alpha(0) = x_0$

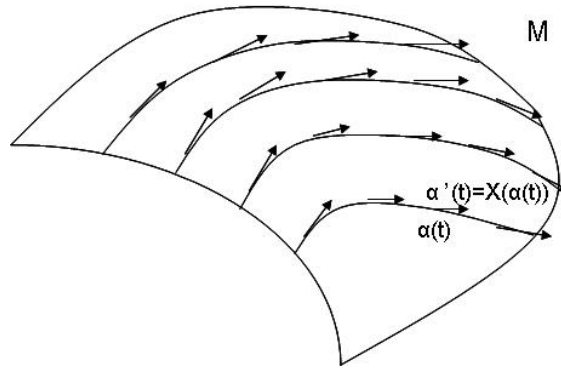


FIGURA 1.2.4: Curva integral del campo X

DEFINICIÓN 1.2.24. El flujo local generado por X en M es el mapeo C^∞ , $\varphi : D(X) \rightarrow M$,

$$\varphi_t(x) = \varphi(t, x) = \alpha_x(t)$$

,donde

$$D(X) = \{(t, x) \in \mathbb{R} \times M \mid w_-(x) < t < w_+(x)\}$$

- Si $D(X) = \mathbb{R} \times M$, el campo de vectores es llamado completo.
- Si $\varphi_t(x)$, $t \in I \subset \mathbb{R}$, $x \in D \subset M$, es el flujo generado por X, entonces la órbita de X en $x_0 \in M$ será el conjunto

$$\Theta(x_0) = \{\varphi_t(x_0) \mid t \in I\}$$

- Si x_0 es un cero de X, (es decir $X(x_0) = 0$), entonces $\Theta(x_0) = \{x_0\}$.
- Si $\Theta(x_0) \neq \{x_0\}$ el mapeo $\alpha : I \rightarrow M$, $\alpha(t) = \varphi_t(x_0)$ es una inmersión. Si α no es un mapeo inyectivo, entonces existe $t_1 > 0$ tal que $\alpha(t_1) = \alpha(0) = x_0$ y $\alpha(t) \neq x_0$ para $0 < t < t_1$, de aqui $\Theta(x_0)$ es difeomórfica al círculo S^1 y decimos que es una curva cerrada.

DEFINICIÓN 1.2.25. El conjunto Ω -límite de punto $x_0 \in M$, denotado por $\Omega(x_0)$, es el conjunto de aquellos puntos $x \in M$ para el cual existe una sucesión $t_n \rightarrow +\infty$ con

$$\lim_{n \rightarrow +\infty} \varphi_{t_n}(x_0) = x.$$

El conjunto α -límite de punto $x_0 \in M$, denotado por $\alpha(x_0)$, es el conjunto de aquellos puntos $x \in M$ para el cual existe una sucesión $t_n \rightarrow -\infty$ con

$$\lim_{n \rightarrow -\infty} \varphi_{t_n}(x_0) = x.$$

TEOREMA 1.2.2. Si X es un campo C^∞ en un abierto V en una variedad M y $p \in V$ entonces existe un abierto $V_0 \subset V$, $p \in V_0$, un número $\delta > 0$, y una aplicación C^∞ , $\varphi : (-\delta, \delta) \times V_0 \rightarrow V$ tal que la curva $t \rightarrow \varphi(t, q)$, $t \in (-\delta, \delta)$, es la única trayectoria de X que en el instante $t = 0$ pasa por el punto q , para cada $q \in V_0$.

PROPOSICIÓN 1.2.6. Dado $p \in M$, existe un abierto $V \subset M$, $p \in V$, números $\delta > 0$ y $\varepsilon_1 > 0$ y una aplicación C^∞

$$\gamma : (-\delta, \delta) \times U \rightarrow M, \quad U = (q, v); \quad q \in V, v \in T_p M, |v| < \varepsilon_1,$$

tales que la curva $t \rightarrow \gamma(t, q, v)$, $t \in (-\delta, \delta)$, es la única geodésica de M que en el instante $t = 0$ pasa por q con velocidad v , para cada $q \in V$ y cada $v \in T_q M$ con $|v| < \varepsilon_1$.

DEFINICIÓN 1.2.26. Sea $p \in M$ y $\mathcal{U} \subset TM$ un abierto tal que

$$\mathcal{U} = \{(q, v) \in TM; q \in V, v \in T_p M, |v| < \varepsilon\}$$

donde V es una vecindad de p . Entonces la aplicación $\exp : \mathcal{U} \rightarrow M$ dada por

$$\exp(q, v) = \gamma(1, q, v) = \gamma(|v|, q, \frac{v}{|v|}), \quad (q, v) \in \mathcal{U}, \tag{1.2.15}$$

es llamada aplicación exponencial en \mathcal{U} .

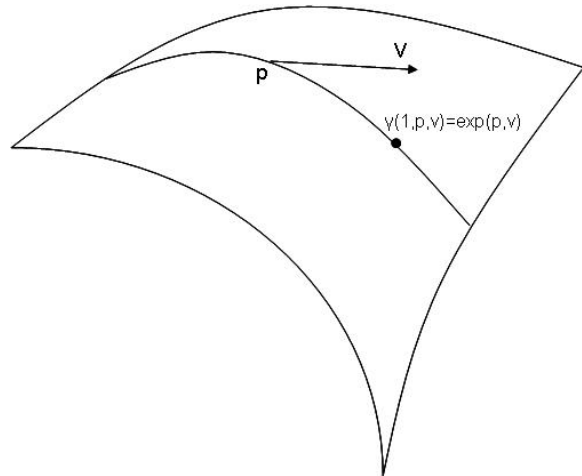


FIGURA 1.2.5: Función exponencial sobre M

PROPOSICIÓN 1.2.7. Sea $q \in M$, existe un $\epsilon > 0$ tal que $\exp_q : B_\epsilon(0) \subset T_q M \rightarrow M$ es un difeomorfismo de $B_\epsilon(0)$ sobre un abierto de \mathbb{R} .

Demostración: ver [4]

DEFINICIÓN 1.2.27. Un campo de vectores X es L -lipschitz continua en un subconjunto $U \subset M$ si para toda curva geodésica $\gamma : [0, 1] \rightarrow M$ con puntos finales en U se tiene

$$|P_{\gamma,1,0}X(\gamma(1)) - X(\gamma(0))|_{\gamma(0)} \leq L.l(\gamma) \quad (1.2.16)$$

Aquí $P_{\gamma,1,0}$ es el transporte paralelo de tiempo 1 a tiempo 0 a lo largo de γ .

DEFINICIÓN 1.2.28. Un subconjunto A es llamado totalmente convexo, si para toda curva geodésica γ , tenemos $(\gamma(0), \gamma(1)) \in A \times A \implies \forall t \in [0, 1], \gamma(t) \in A$.

DEFINICIÓN 1.2.29. $f : A \rightarrow \mathbb{R}$ se dice convexa en A , un conjunto totalmente convexo, si para toda curva geodésica γ con $(\gamma(0), \gamma(1)) \in A \times A$ y $\forall t \in [0, 1]$

$$f(\gamma(t)) \leq (1-t)f(\gamma(0)) + tf(\gamma(1)) \quad (1.2.17)$$

TEOREMA 1.2.3. Sea $A \subset M$ un conjunto abierto, totalmente convexo, y $f : A \rightarrow \mathbb{R}$ una función de clase C^2 . La función f es convexa si y solo si

$$\varphi_{xy} = f(\gamma_{xy}(t)), \quad t \in [0, 1]$$

satisface

$$\frac{d^2 \varphi_{xy}}{dt^2} \geq 0 \quad \forall t \in [0, 1] \quad \forall x, y \in A \quad \forall \gamma_{xy} \in \Gamma$$

Donde Γ es el conjunto de todas las geodésicas que unen a los puntos x y y .

Demostración: ver [3]

DEFINICIÓN 1.2.30. Sea $f : M \rightarrow \mathbb{R}$ una función de clase C^2 . El Hessiano de f es definido por

$$\text{Hess}f(X, Y) = \nabla_X(df)(Y), \quad \forall X, Y \in X(M) \quad (1.2.18)$$

TEOREMA 1.2.4. Sea $A \subset M$ un conjunto abierto, totalmente convexo, y $f : A \rightarrow \mathbb{R}$ una función de clase C^2 . La función f es convexa si y solo si $\text{Hess}f$ es semidefinido positivo en A .

Demostración: ver [3]

Comentario:

- La propiedad de convexidad es invariante bajo las transformaciones de coordenadas no lineales.
- La función f es llamada estrictamente convexa si $\text{Hess}f > 0$, para funciones de clase C^2

- la métrica Riemanniana es dada por:

$$ds^2 = \sum_{i,j} g_{ij} dx_i dx_j \quad (1.2.19)$$

TEOREMA 1.2.5. (*Hopf-Rinow*)

Sea M una variedad Riemanniana y sea $p \in M$, las siguientes afirmaciones son equivalentes:

- a) \exp_x está definida en todo el $T_p M$
- b) Los cerrados y acotados de M son compactos
- c) M es completa como espacio Métrico
- d) M es geodésicamente completa
- e) Existe una sucesión de compactos $K_n \subset M$, $K_n \subset \text{int} K_{n+1}$ y $\bigcup_n K_n = M$, tal que, si $q_n \notin K_n$ entonces $d(p, q_n) \rightarrow \infty$. (Aquí $\text{int} A$ indica el interior del conjunto A)
- f) Para todo $q \in M$ existe una geodésica γ uniendo p a q con $l(\gamma) = d(p, q)$.

Demostración: ver [4]

§1.2.2. Método de Descenso Acelerado en Variedades de Riemann

Sea (M, g) una variedad Riemanniana finito-dimensional, supongamos que se tiene una función real de clase C^2 , $f : M \rightarrow \mathbb{R}$, se desea encontrar el mínimo a esta función. Recordemos que $-\text{grad} f$ es un vector ortogonal a las superficies de nivel constantes de f . Se verá que en cada punto $x \in M$, la dirección y el sentido de el descenso acelerado, sugiere que la solución mas conveniente de la desigualdad

$$df(x_i)(X) \quad \text{es} \quad X_i = -\text{grad} f$$

con esta escogencia, el proceso iterativo resultante se le llamará el método de descenso acelerado ó Método Gradiente, recordando el bien conocido modelo Euclídeo, este método es descrito por el siguiente algoritmo.

1. Considere un punto inicial x_1 y calcule $\text{grad} f(x_1)$. Si $\text{grad} f(x_1) = 0$, entonces pare
2. Si $\text{grad} f(x_1) \neq 0$, entonces de el punto x_1 pase a algún punto

$$x_2 = \exp_{x_1}(-t_1 \text{grad} f(x_1)),$$

donde $t_1 > 0$ es escogido tal que $f(x_2) < f(x_1)$. Si $\text{grad} f(x_2) = 0$, entonces, pare

3. Si el $\text{grad} f(x_2) \neq 0$, entonces de el punto x_2 pase a otro punto

$$x_3 = \exp_{x_2}(-t_2 \text{grad} f(x_2)),$$

donde $t_2 > 0$ es escogido tal que $f(x_3) < f(x_2)$. Si $\text{grad}f(x_3) = 0$, entonces pare.

En la parada uno verifica si el punto establecido es un punto mínimo.

4. Generalmente, si $\text{grad}f(x_i) \neq 0$, entonces se pasa a

$$x_{i+1} = \exp_{x_i}(-t_i \text{grad}f(x_i)),$$

donde t_i , es determinado por la condición $f(x_{i+1}) < f(x_i)$.

§1.2.3. Ecuación de Evolución en Variedades de Riemann

Propiedades Generales del Flujo Gradiente

Sea (M, g) una variedad Riemanniana finito-dimensional, y sea $f : M \rightarrow \mathbb{R}$ una función de clase C^2 .

Veamos para la solución del problema de Cauchy

$$\frac{d\alpha}{dt} = \text{grad}f(\alpha(t)), \quad \alpha(0) = x$$

Esta solución es llamada línea gradiente de f a través de x . Este problema de Cauchy tiene una única solución maximal $\alpha_x : (w_-(x), w_+(x)) \rightarrow M$ por ser $\text{grad}f$ de clase C^1 y de allí localmente Lipschitz en M , donde $w_-(x) < 0 < w_+(x)$ y $\varphi_t(x) = \alpha_x(t)$ define el flujo gradiente.

TEOREMA 1.2.6. Sea $\alpha : I \rightarrow M$, $I = (w_-(x), w_+(x))$, la línea gradiente de f

1. α es una curva de crecimiento máximo local de f . Por consiguiente allí existen los límites

$$\lim_{t \searrow w_-(x)} f(\varphi_t(x)) \quad , \quad \lim_{t \nearrow w_+(x)} f(\varphi_t(x))$$

2. α es una curva cerrada sólo en el caso en el que esta reducido a un punto de equilibrio (punto crítico de f)

3. Si $w_+(x)$ es finita, entonces

$$\lim_{t \nearrow w_+(x)} f(\varphi_t(x)) = \infty$$

4. Si $w_+(x) = \infty$ y aquí existe $\lim_{t \rightarrow \infty} \alpha(t) = x_0$, entonces x_0 es un punto crítico de f

Demostración

1. Recordando la regla de la cadena, obtenemos el siguiente resultado

$$\begin{aligned} \frac{d}{dt} f(\alpha(t)) &= df(\alpha(t)) \left(\frac{d\alpha}{dt}(t) \right) \\ &= \|\text{grad}f(\alpha(t))\|^2 \\ &\geq 0 \quad \forall t \in I \end{aligned}$$

De aquí se tiene que $f(\alpha(t)) = f(x)$ para todo $t \in I(x)$, ó $f \circ \alpha$ es creciente y por consiguiente los límites en el teorema existen.

Mostraremos que para cada $t_0 \in I(x)$, la función f crece mas rápido a lo largo de α en $\alpha(t_0)$ que a lo largo de otra curva pasando a través de $\alpha(t_0)$ con la misma rapidez.

Verdaderamente, si $\alpha : (a, b) \rightarrow M$ es una curva integral de $\text{grad} f$ y $\beta : (c, d) \rightarrow M$ es cualquier otra curva de clase C^1 tal que $\beta(s_0) = \alpha(t_0)$, $s_0 \in (c, d)$, $t_0 \in (a, b)$ y $\|\beta'(s_0)\| = \|\alpha'(t_0)\|$, entonces

$$\begin{aligned} \frac{d(f \circ \beta)}{ds}(s_0) &= df(\beta(s_0))(\beta'(s_0)) \\ &\leq \|df(\beta(s_0))\| \cdot \|\beta'(s_0)\| \\ &= \|df(\alpha(t_0))\| \cdot \|\alpha'(t_0)\| \\ &= \|\text{grad} f(\alpha(t_0))\|^2 \\ &= \frac{d(f \circ \alpha)}{dt}(t_0). \end{aligned}$$

2. Sea $[a, b] \subset I$ y $\alpha : [a, b] \rightarrow M$ una curva integral cerrada, de $\text{grad} f$ (i.e. $\alpha(a) = \alpha(b)$).

Supongamos por absurdo que α no está reducida a un punto de equilibrio entonces

$$\begin{aligned} 0 = f(\alpha(b)) - f(\alpha(a)) &= \int_a^b df(\alpha(t))(\alpha'(t)) dt \\ &= \int_a^b \|\text{grad} f(\alpha(t))\|^2 dt > 0 \quad (\rightarrow \leftarrow) \end{aligned}$$

Fíjese que $\|\text{grad} f(\alpha(t))\|^2 \neq 0$, pues estamos suponiendo que α no se reduce a un punto de equilibrio, esto nos lleva a afirmar por la parte anterior que $\|\text{grad} f(\alpha(t))\|^2 > 0$. Así llegamos a la contradicción y por tanto α está reducida a un punto de equilibrio.

3. La relación

$$\frac{d}{dt} f(\alpha(t)) = \|\text{grad} f(\alpha(t))\|^2 \quad \forall t \in I$$

Implica que

$$f(\alpha(t)) = f(\alpha(s)) + \int_s^t \|\text{grad} f(\alpha(r))\|^2 dr \quad w_-(x) < r < w_+(x)$$

para $0 \leq s \leq t < w_+(x)$. La definición de la distancia d en M , la desigualdad de Cauchy Schwarz y la

relacion anterior implica

$$\begin{aligned}
 d(\alpha(t), \alpha(s)) &= \inf_{\beta \in \Omega} L(\beta) \leq L(\alpha) \\
 &= \int_s^t \|\alpha'(r)\| dr \\
 &\leq \int_s^t \|\text{grad}f(\alpha(r))\| dr \\
 &\leq (t-s)^{1/2} \left(\int_s^t \|\text{grad}f(\alpha(r))\|^2 \right)^{1/2} \\
 &= (t-s)^{1/2} (f(\alpha(t)) - f(\alpha(s)))^{1/2}
 \end{aligned}$$

Como $w_+(x) < \infty$, $\alpha(t)$ no converge cuando $t \rightarrow w_+(x)$, en efecto, supongamos que α converge a t , cuando t tiende a $w_+(x)$, entonces por el teorema de Picard la solución α puede ser extendida lo cual es una contradicción pues α es solución máxima de el problema de Cauchy con el que se está trabajando. Por tanto α no converge a t , cuando $t \rightarrow w_+(x)$ y como $f \circ \alpha$ es creciente, la inecuación de arriba implica que

$$\lim_{t \nearrow w_+(x)} f(\alpha(t)) = \infty.$$

4. Tomemos un $t_0 \in I(x)$ fijo y arbitrario.

Obviamente $\lim_{t \rightarrow \infty} f(\alpha(t)) = f(x_0)$, entonces

$$\begin{aligned}
 f(x_0) - f(\alpha(t_0)) &= \lim_{t \rightarrow \infty} (f(\alpha(t)) - f(\alpha(t_0))) \\
 &= \lim_{t \rightarrow \infty} \int_{t_0}^t df(\alpha(r))(\alpha'(r)) dr \\
 &= \lim_{t \rightarrow \infty} \int_{t_0}^t \|\text{grad}f(\alpha(r))\|^2 dr \\
 &= \int_{t_0}^{\infty} \|\text{grad}f(\alpha(r))\|^2
 \end{aligned}$$

La integral anterior converge, ahora por ser $t_0 \in I(x)$ arbitrario, se tiene

$$\int_t^{\infty} \|\text{grad}f(\alpha(r))\|^2 < \infty \quad \forall t \in I(x)$$

y esto implica que $\lim_{r \rightarrow \infty} \|\text{grad}f(\alpha(r))\|^2 = 0$, en efecto, supongamos que $\lim_{r \rightarrow \infty} \|\text{grad}f(\alpha(r))\|^2 \neq 0$.

Si $\lim_{r \rightarrow \infty} \|\text{grad}f(\alpha(r))\|^2 = c > 0$, entonces

$$+\infty = \int_t^{\infty} c.dr < \int_{t_0}^{\infty} \|\text{grad}f(\alpha(r))\|^2 < \infty \quad (\rightarrow \leftarrow)$$

Si $\lim_{r \rightarrow \infty} \|\text{grad}f(\alpha(r))\|^2 = +\infty$, entonces, existe un $c > 0$ tal que

$$+\infty = \int_t^{\infty} c.dr < \int_{t_0}^{\infty} \|\text{grad}f(\alpha(r))\|^2 < \infty \quad (\rightarrow \leftarrow)$$

Luego $\lim_{r \rightarrow \infty} \|\text{grad}f(\alpha(r))\|^2 = 0$, además,

$$\begin{aligned} \|\text{grad}f(x_0)\|^2 &= \left\| \lim_{t \rightarrow \infty} \text{grad}f(\alpha(t)) \right\|^2 \\ &= \lim_{t \rightarrow \infty} \|\text{grad}f(\alpha(t))\|^2 \\ &= 0 \end{aligned}$$

$\therefore x_0$ es punto crítico.

Ecuación de Evolución ó Flujo Menos Gradiente

Sea (M, g) una variedad Rimaniana completa finito-dimensional y sea $f : M \rightarrow \mathbb{R}$, una función real de clase C^2 . Consideremos el siguiente problema de Cauchy

$$\frac{d\alpha}{dt} = -\text{grad}f(\alpha(t)), \quad \alpha(0) = x.$$

Algunas veces a la solución α de este problema se le llama menos gradiente, pero en este trabajo se le llamará solución de la Ecuación de Evolución.

Por ser $\text{grad}f$ de clase C^1 y localmente lipschitz en M , el problema de Cauchy tiene una única solución máxima

$$\alpha_x : (w_-(x), w_+(x)) \rightarrow M \quad w_-(x) < 0 < w_+(x)$$

y $\varphi_t(x) = \alpha_x(t)$ define el flujo de $-\text{grad}f$. Los puntos críticos de f son puntos fijos de el flujo φ_t .

Primeramente recordemos el teorema anterior (ver teorema 2.0.9), con la diferencia de que

$$\alpha : I \rightarrow M, \quad I = (w_-(x), w_+(x))$$

es la línea menos gradiente de f , es decir.

1. α es una curva de máximo decrecimiento local de f . Por consiguiente allí existen los límites

$$\lim_{t \searrow w_-(x)} f(\varphi_t(x)) \quad , \quad \lim_{t \nearrow w_+(x)} f(\varphi_t(x))$$

2. α es una curva cerrada solo en el caso en el que esta reducido a un punto de equilibrio (punto crítico de f)
3. Si $w_+(x)$ es finita , entonces $\lim_{t \nearrow w_+(x)} f(\varphi_t(x)) = -\infty$
4. Si $w_+(x) = \infty$ y aquí existe $\lim_{t \rightarrow \infty} \alpha(t) = x_0$, entonces x_0 es un punto crítico de f

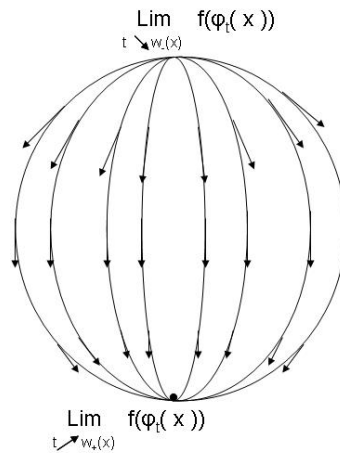


FIGURA 1.2.6: Ejemplo gráfico del teorema anterior sobre una superficie esférica, con la función altura.

Sea $\text{Crit}(f)$, el conjunto de todos los puntos críticos de f . Si $x \in M - \text{Crit}(f)$, entonces del teorema anterior (teorema 2.0.9) se tiene lo siguiente

$$\begin{aligned} f(\alpha(t)) &< f(x) && \text{para } t > 0 \\ f(\alpha(t)) &> f(x) && \text{para } t < 0 \end{aligned}$$

DEFINICIÓN 1.2.31. Sea $A \subset M$. Dado $\varepsilon > 0$, un subconjunto A_ε de A es llamado una ε -red de A , si

- A_ε es finito.
- Para cada $x \in A$ hay un $y \in A_\varepsilon$ tal que $d(x, y) < \varepsilon$.

El conjunto A es totalmente acotado si para cada $\varepsilon > 0$, A contiene una ε -red.

El conjunto A es llamado relativamente compacto si la clausura de A es secuencialmente compacto (toda sucesión tiene una subsucesión convergente) en (M, d) . Por ser M completo, acotamiento total y compacidad secuencial son equivalentes.

TEOREMA 1.2.7. Si $\alpha : (a, b) \rightarrow M$ es una curva de clase C^1 , con $L(\alpha) < \infty$, entonces la $\text{Im}(\alpha)$ es relativamente compacta en M .

Demostración:

Sea $\varepsilon > 0$ fijo y arbitrario, y consideremos la siguiente partición

$$a = t_0 < t_1 < \dots < t_n < t_{n+1} = b.$$

de $[a, b]$ tal que

$$\int_{t_i}^{t_{i+1}} \|\alpha'(t)\| dt < \varepsilon, \quad i = 1, 2, \dots, n + 1$$

esto si que el conjunto

$$A_\varepsilon = \{x_i = \alpha(t_i), \quad i = 1, 2, \dots, n\}$$

es una ε -red en $\text{im}(\alpha)$; por ser ε arbitrario, $\text{im}(\alpha)$ es totalmente acotada.

TEOREMA 1.2.8. sea $X \in \chi(M)$ y $\alpha : (w_-(x), w_+(x)) \rightarrow M$ es la solución máxima del problema de Cauchy

$$\frac{d\alpha}{dt} = X(\alpha(t)), \quad \alpha(0) = x \in M$$

Entonces lo siguiente es cierto

1. Si $w_+(x)$ es finito, entonces

$$\lim_{t \nearrow w_+(x)} \|X(\alpha(t))\| = \infty.$$

2. Si $w_-(x)$ es finito, entonces

$$\lim_{t \searrow w_-(x)} \|X(\alpha(t))\| = \infty.$$

Demostración:

1. Como $w_+(x) < \infty$ entonces $\lim_{t \nearrow w_+(x)} \alpha(t) = \infty$ ya que α es solución máxima, Así $\text{Im}(\alpha)$ no es relativamente compacto en M y por el teorema anterior (teorema 1.2.7), $L(\alpha) = \infty$.

Por consiguiente

$$\int_0^{w_+(x)} \|X(\alpha(t))\| dt = \int_0^{w_+(x)} \|\alpha'(t)\| dt = \infty$$

2. Análogo a la parte anterior.

Consideremos la función $f : M \rightarrow \mathbb{R}$ de clase C^2 que satisface las siguientes suposiciones:

[A] $\inf_M f = m > -\infty$;

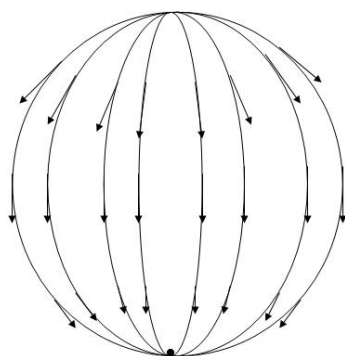
[B] (Condición de Palais-Smale). Cada sucesión $\{x_n\} \subset M$, tal que $\{f(x_n)\}$ es acotada y $\|\text{grad}f(x_n)\| \rightarrow 0$, contiene una subsucesión convergente.

TEOREMA 1.2.9. Asuma que f satisface las condiciones [A] y [B].

1. El flujo $\varphi_t(x)$ generado por $-\text{grad}f$ es definida en $M \times [0, \infty)$.
2. Para cada $x \in M$ se tiene la siguiente relación,

$$\overline{\{\varphi_t(x) : t > 0\}} \cap \text{Crit}(f) \neq \emptyset.$$

Demostración:



$$\overline{\{\varphi_t(x) / t > 0\}} \cap \text{Crit}(f) \neq \emptyset$$

FIGURA 1.2.7: Teorema anterior visto graficamente sobre una superficie esférica

1. Supongamos que $w_+(x) < \infty$. Sea $t_0 \in (0, w_+(x))$ y llamemos $h(t) = f(\varphi_t(x))$. Tenemos

$$\begin{aligned} m = \inf_m f \leq h(t_0) &= h(0) + \int_0^{t_0} h'(t) dt \\ &= h(0) - \int_0^{t_0} \|\text{grad}f(\varphi_t(x))\|^2 dt \\ &\leq h(0) - \int_0^{w_+(x)} \|\text{grad}f(\varphi_t(x))\|^2 dt. \end{aligned}$$

Por otro lado, por el teorema de Holder (ver teorema(1.1.1)) se tiene

$$\begin{aligned} \int_0^{w_+(x)} \|\text{grad}f(\varphi_t(x))\| dt &\leq (w_+(x))^{1/2} \left(\int_0^{w_+(x)} \|\text{grad}f(\varphi_t(x))\|^2 dt \right)^{1/2} \\ &\leq (w_+(x))^{1/2} (h(0) - m)^{1/2} < +\infty, \end{aligned}$$

pero esto contradice el teorema (1.2.8)

2. Probemos la existencia de una sucesión $\{t_n\}$ que satisfaga $\lim_{n \rightarrow \infty} t_n = \infty$ y el $\lim_{n \rightarrow \infty} \|\text{grad}f(\varphi_{t_n}(x))\| = 0$. De otro modo, existiría $C > 0$ tal que $\|\text{grad}f(\varphi_t(x))\| \geq C \quad \forall t \geq 0$, y de aquí $\int_0^{t_0} \|\text{grad}f(\varphi_t(x))\|^2 dt \geq C^2 t_0 \quad \forall t_0 \geq 0$. Considerando además que

$$m \leq h(0) - \int_0^{t_0} \|\text{grad}f(\varphi_t(x))\|^2 dt$$

Se tiene que

$$h(0) - m \geq C^2 t_0$$

lo cual es una contradicción para cuando $t_0 \rightarrow \infty$, de aquí $\lim_{n \rightarrow \infty} \|\text{grad}f(\varphi_{t_n}(x))\| = 0$, y por la condición de Palais-Smale, es asegurada la existencia de una subsucesión convergente $\varphi_{t_{n_k}}(x) \rightarrow x_0$, con $x_0 \in \text{Crit}(f)$

$$\therefore x_0 \in \overline{\{\varphi_t(x) : t \geq 0\}} \cap \text{Crit}(f).$$

COROLARIO 1.2.2. Si $x \in M - \text{Crit}(f)$, entonces existe un $x_0 \in \text{Crit}(f)$ tal que $f(x_0) < f(x)$

Demostración:

tomemos $x_0 \in \overline{\{\varphi_t(x) : t \geq 0\}} \cap \text{Crit}(f)$, encontramos que

$$f(x) = f(\varphi_0(x)) > f(\varphi_{t_n}(x)) \geq \lim_{n \rightarrow \infty} f(\alpha_{t_n}(x)) = f(x_0).$$

TEOREMA 1.2.10. Si f satisface las condiciones [A] y [B], entonces existe $x_0 \in \text{Crit}(f)$, tal que $f(x_0) = m$

Demostración:

Sea $m = \inf_M f$ y $\{x_n\} \subset M$ tal que $\lim_{n \rightarrow \infty} f(x_n) = m$, usando el corolario anterior, se puede suponer que $x_n \in \text{Crit}(f), \forall n \in \mathbb{N}$, así obtenemos lo siguiente

$$\|\text{grad}f(x_n)\| = 0 \longrightarrow 0,$$

de aquí $\{f(x_n)\}$ es acotada, y de la condición de Palais-Smale, la sucesión $\{x_n\}$ tiene una subsucesión convergente $x_{n_k} \longrightarrow x_0$, y $x_0 \in \text{Crit}(f)$

CAPÍTULO 2

DESARROLLO

Sea $(M, \langle \cdot, \cdot \rangle)$ una variedad dotada con una métrica Rimanmiana. Denotaremos la longitud de la curva como $\ell(\cdot)$ y la distancia $\rho(\cdot, \cdot)$.

Sea $\phi : M \rightarrow \mathbb{R}$ una función diferenciable, denotaremos $grad\phi(x)$ al gradiente con respecto a el producto interno Rimanmiano $\langle \cdot, \cdot \rangle_x$, Es decir para cada $u \in T_x M$ se sostiene lo siguiente:

$$d\phi(x).u = \langle grad\phi(x), u \rangle_x \quad (2.0.1)$$

Consideremos la siguiente Ecuación de Evolución

$$\dot{x}(t) + grad\phi(x(t)) = 0 \quad (2.0.2)$$

Desarrollaremos este trabajo suponiendo lo siguiente:

$$H_1 = \begin{cases} (M, \rho), & \text{es completo;} \\ \phi \in C^1, & \text{con gradiente localmente Lipschitz-continuo;} \\ \phi, & \text{es acotada por debajo.} \end{cases} \quad (2.0.3)$$

TEOREMA 2.0.11. *Asuma que se tiene (H_1) . Para todo punto inicial $x_0 \in M$, el problema de Cauchy*

$$\begin{cases} \dot{x}(t) + grad\phi(x(t)) = 0 \\ x(0) = x_0, \end{cases} \quad (2.0.4)$$

tiene una única solución $x(\cdot)$ definida en $[0, +\infty)$. Esta solución es continuamente diferenciable.

Demostración:

La existencia local y la unicidad de la solución de (2.0.2) es probada aplicando los resultados de teoremas de ecuaciones diferenciales. Estos proveen una solución máxima definida en un intervalo de tipo $[0, T)$.

Asumamos por contradicción que $T < +\infty$, como

$$\frac{d}{dt}\phi(x(t)) = \langle grad\phi(x(t)), \dot{x}(t) \rangle_{x(t)} = \langle -\dot{x}(t), \dot{x}(t) \rangle_{x(t)} = -\langle \dot{x}(t), \dot{x}(t) \rangle = -|\dot{x}(t)|_{x(t)}^2 \leq 0 \quad (2.0.5)$$

Luego la función $\phi \circ x$ es decreciente y como ϕ es acotada por debajo por H_1 entonces existe el límite cuando $t \rightarrow \infty$.

Por otro lado, observemos que:

$$\phi(x(T)) - \phi(x(0)) = - \int_0^T |\dot{x}(t)|_{x(t)}^2 dt \iff \phi(x(T)) = \phi(x_0) - \int_0^T |\dot{x}(t)|_{x(t)}^2 dt \quad (2.0.6)$$

Además sabemos por definición de ínfimo que

$$\inf_M \phi \leq \phi(x(t)) \quad \forall t, \quad (2.0.7)$$

en particular se cumple para $t = T$ esto es

$$\inf_M \phi \leq \phi(x(T)) \iff -\inf_M \phi \geq -\phi(x(T)) \iff -\phi(x(T)) \leq -\inf_M \phi \quad (2.0.8)$$

De 2.0.6 y 2.0.8 se tiene que

$$\int_0^T |\dot{x}(t)|_{x(t)}^2 dt = \phi(x_0) - \phi(x(T)) \leq \phi(x_0) - \inf_M \phi \quad (2.0.9)$$

Por ser $\int_0^T |\dot{x}(t)|_{x(t)}^2 dt \geq 0$ entonces

$$\left(\int_0^T |\dot{x}(t)|_{x(t)}^2 dt \right)^{\frac{1}{2}} \leq \left(\phi(x_0) - \inf_M \phi \right)^{\frac{1}{2}} \quad (2.0.10)$$

Luego por el Teorema de Holder (ver teorema 1.1.1)

$$\begin{aligned} \int_0^T |\dot{x}(t)|_{x(t)} dt &\leq \left(\int_0^T 1 dt \right)^{\frac{1}{2}} \left(\int_0^T |\dot{x}(t)|_{x(t)}^2 dt \right)^{\frac{1}{2}} \\ &= T^{\frac{1}{2}} \left(\int_0^T |\dot{x}(t)|_{x(t)}^2 dt \right)^{\frac{1}{2}} \\ &\leq T^{\frac{1}{2}} \left(\phi(x_0) - \inf_M \phi \right)^{\frac{1}{2}} \quad \text{por 2.0.10} \\ &= \sqrt{T \left(\phi(x_0) - \inf_M \phi \right)} \end{aligned}$$

Luego $\dot{x}(t)$ es integrable y $x(\cdot)$ tiene límite cuando $t \rightarrow T$. Por causa de la completitud de M . Así por resultados de Ecuaciones diferenciales ordinarias sobre variedades, se puede extender esta solución, con lo que contradice la maximalidad \diamond .

TEOREMA 2.0.12. Sea $x(\cdot)$ una solución de (2.0.2), bajo la suposición de (H_1) . Esto satisface lo siguiente.

i $\phi(x(\cdot))$ es decreciente y converge

ii $|\dot{x}(\cdot)|_{x(\cdot)} \in L^2(0, +\infty)$.

Demostración:

i Esto viene de (2.0.5) pues $\frac{d\phi}{dt}(x(t)) \leq 0$, luego $\phi(x(\cdot))$ es decreciente y por ser ϕ acotada por debajo, entonces $\phi(x(\cdot))$ converge. (denotaremos ϕ_∞ cuando $t \rightarrow +\infty$, es decir, $\phi_\infty = \lim_{t \rightarrow +\infty} \phi(x(t))$) \diamond .

ii Ahora como $\int_0^T |\dot{x}(t)|_{x(t)}^2 dt \leq \phi(x_0) - \inf_M \phi$ ver (2.0.9), por (2.0.10), la definición de los espacios de las funciones integrables a la p (ver Definición (1.1.4)) y por ser $\phi(x_0) - \inf_M \phi < \infty$ obtenemos que $|\dot{x}(t)|_{x(t)} \in L^2(0, +\infty) \diamond$.

TEOREMA 2.0.13. Asumiremos (H_1) . Sea $x(\cdot)$ una solución de (2.0.2) acotada, entonces se satisface que

$$\lim_{t \rightarrow \infty} |\text{grad}\phi(x(t))|_{x(t)} = 0 \quad (2.0.11)$$

demostración:

Sea L la constante de Lipschitz para el $\text{grad}\phi$ en el conjunto acotado $\{x(t), t \in [0, \infty]\}$. Denotemos para $s \leq t$, por $\gamma : [0, 1] \rightarrow M$ una geodésica minimal tal que $\gamma(0) = x(s)$ y $\gamma(1) = x(t)$, donde $P_{\gamma,1,0}$ es una isometría, tenemos lo siguiente:

$$\begin{aligned} \left| |\dot{x}(t)|_{x(t)} - |\dot{x}(s)|_{x(s)} \right| &\leq |\dot{x}(t) - \dot{x}(s)|_{x(s)} \\ &= |P_{\gamma,1,0}\dot{x}(t) - \dot{x}(s)|_{x(s)} \quad \text{ver proposición (1.2.4)} \\ &\leq L.l(\gamma) \quad \text{por definición (1.2.27)} \\ &\leq L. \int_s^t |\dot{x}(\tau)|_{x(\tau)} d\tau \\ &\leq L. \left(\int_s^t |\dot{x}(\tau)|_{x(\tau)}^2 d\tau \right)^{\frac{1}{2}} \left(\int_s^t 1^2 d\tau \right)^{\frac{1}{2}} \\ &\leq L. \left(\int_s^\infty |\dot{x}(\tau)|_{x(\tau)}^2 d\tau \right)^{\frac{1}{2}} \cdot \sqrt{t-s} \\ &\leq L.M.\sqrt{t-s} \quad \text{por ser } |\dot{x}(t)| \in L^2(0, \infty) \end{aligned}$$

Luego dado $\varepsilon > 0$, existe $\delta = \frac{\varepsilon^2}{L^2 M^2}$, tal que si

$$|t-s| < \delta \Rightarrow \left| |\dot{x}(t)|_{x(t)} - |\dot{x}(s)|_{x(s)} \right| \leq L.M.\sqrt{t-s} < L.M.\sqrt{\frac{\varepsilon^2}{L^2 M^2}}$$

De aquí $|\dot{x}(\cdot)|_{x(\cdot)}$ es uniformemente continua y su cuadrado es integrable, así necesariamente

$$\lim_{t \rightarrow +\infty} |\dot{x}(t)|_{x(t)} = 0,$$

en efecto, Supongamos por el absurdo que $\lim_{t \rightarrow +\infty} |\dot{x}(t)|_{x(t)} \neq 0$.

Por hipótesis $|\dot{x}(t)|_{x(t)}$ es cuadrado integrable es decir que

$$|\dot{x}(t)|_{x(t)} \in L^2(0, +\infty) \implies \int_0^{+\infty} |\dot{x}(t)|_{x(t)}^2 dt < \infty$$

Como $\lim_{t \rightarrow +\infty} |\dot{x}(t)|_{x(t)} \neq 0$, existen dos casos posibles:

Caso 1:

$$\lim_{t \rightarrow +\infty} |\dot{x}(t)|_{x(t)} = +\infty$$

Afirmación 1:

$$\lim_{t \rightarrow +\infty} |\dot{x}(t)|_{x(t)} = +\infty \Rightarrow \lim_{t \rightarrow +\infty} |\dot{x}(t)|_{x(t)}^2 = +\infty$$

Demostración:

$$\lim_{t \rightarrow +\infty} |\dot{x}(t)|_{x(t)}^2 = \left(\lim_{t \rightarrow +\infty} |\dot{x}(t)|_{x(t)} \right)^2 = +\infty \quad \diamond.$$

Afirmación 2:

$$\lim_{t \rightarrow +\infty} |\dot{x}(t)|_{x(t)}^2 = +\infty \Rightarrow \int_0^{+\infty} |\dot{x}(t)|_{x(t)}^2 dt = +\infty,$$

en efecto,

$$\begin{aligned} \forall M > 0, \exists N > 0 / t > N &\Rightarrow |\dot{x}(t)|_{x(t)}^2 > M \\ &\Rightarrow \int_N^{+\infty} |\dot{x}(t)|_{x(t)}^2 dt > \int_N^{+\infty} M dt = +\infty \\ &\Rightarrow \int_0^{+\infty} |\dot{x}(t)|_{x(t)}^2 dt = +\infty \quad \diamond. \end{aligned}$$

Pero esto es una contradicción pues $|\dot{x}(t)|_{x(t)}$ es cuadrado integrable

$$\therefore \lim_{t \rightarrow +\infty} |\dot{x}(t)|_{x(t)} \neq +\infty \quad \diamond.$$

Caso 2

$$\lim_{t \rightarrow +\infty} |\dot{x}(t)|_{x(t)} = \eta > 0.$$

Afirmación 3

$$\lim_{t \rightarrow +\infty} |\dot{x}(t)|_{x(t)} = \eta > 0 \Rightarrow \lim_{t \rightarrow +\infty} |\dot{x}(t)|_{x(t)}^2 = \eta^2 > 0$$

Demostración:

$$\lim_{t \rightarrow +\infty} |\dot{x}(t)|_{x(t)}^2 = \left(\lim_{t \rightarrow +\infty} |\dot{x}(t)|_{x(t)} \right)^2 = \eta^2 > 0 \quad \diamond.$$

Afirmación 4

$$\lim_{t \rightarrow +\infty} |\dot{x}(t)|_{x(t)}^2 = \beta > 0 \Rightarrow \int_0^{+\infty} |\dot{x}(t)|_{x(t)}^2 dt = +\infty$$

Demostración:

$$\begin{aligned} \forall \varepsilon > 0 \exists N > 0 / t > N &\Rightarrow \left| |\dot{x}(t)|_{x(t)}^2 - \beta \right| < \varepsilon \\ &\Rightarrow -\varepsilon < |\dot{x}(t)|_{x(t)}^2 - \beta < \varepsilon \\ &\Rightarrow \beta - \varepsilon < |\dot{x}(t)|_{x(t)}^2 < \beta + \varepsilon. \end{aligned}$$

Como se cumple $\forall \varepsilon > 0$ en particular para algún $\varepsilon_0 > 0$ tal que $\beta - \varepsilon_0 > 0$, luego

$$\begin{aligned} & \beta - \varepsilon_0 < |\dot{x}(t)|_{x(t)}^2 < \beta - \varepsilon_0 \\ \Rightarrow & \int_N^{+\infty} \beta - \varepsilon_0 < \int_N^{+\infty} |\dot{x}(t)|_{x(t)}^2 < \int_N^{+\infty} \beta - \varepsilon_0 \\ \Rightarrow & \int_N^{+\infty} |\dot{x}(t)|_{x(t)}^2 \rightarrow +\infty \quad (\rightarrow \leftarrow), \end{aligned}$$

puesto que, $|\dot{x}(t)|_{x(t)}$ es cuadrado integrable.

$\therefore \lim_{t \rightarrow +\infty} |\dot{x}(t)|_{x(t)} = 0$ que era lo que queriamos probar \diamond .

§2.1. CASO CONVEXO

Aparte de que se asume (H_1) , se asumirá además que

$$(H_2) \begin{cases} \phi, & \text{es convexa en } M \\ \arg \min \phi \neq \emptyset, \end{cases}$$

Para la definición de convexidad ver def (1.2.29)

Se denotará para algún y en M y para cada t , un vector $u(t) \in T_{x(t)}M$, tal que

$$\begin{aligned} \exp_{x(t)}(u(t)) &= y \\ \rho(x(t), y) &= |u(t)|_{x(t)} \end{aligned} \tag{2.1.1}$$

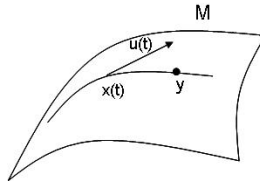


FIGURA 2.1.1: El punto $\exp_{x(t)}(u(t)) = y$ sobre M

La existencia del vector no los garantizará la completitud de M , lo mas seguro es que este no sea único, pero no es necesaria tal condición. La curva $[0, 1] \ni s \mapsto \exp_{x(t)}(su(t))$ será aquí una geodésica minimizante que une a $x(t)$ con y

PROPOSICIÓN 2.1.1. *Bajo la suposición (H_1) , si además ϕ es convexo en M y y pertenece al conjunto $L = \{y \in M / \phi(y) \leq \phi_\infty\}$, entonces la función $t \mapsto \rho(x(t), y)$ decrece y converge.*

Demostración:

Fijemos t y h .

En orden a simplificar las notaciones $x(t)$, $\text{grad}\phi(x(t))$, $u(t)$, $x(t+h)$ se denotará como x , g , u , x_h respectivamente. Considere la geodésica $\gamma(s) = \exp_{x(t)}(su)$ tal que $\gamma(0) = x$, $\dot{\gamma}(0) = u$ y

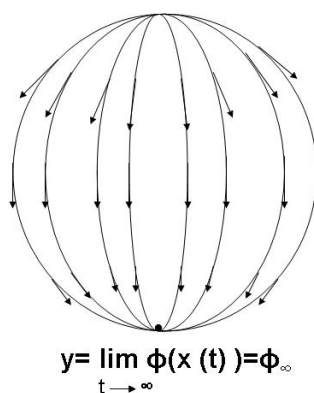


FIGURA 2.1.2: en esta superficie el conjunto L es unitario y el punto $y = \lim_{t \rightarrow +\infty} \phi(x(t)) = \phi_{\infty}$

$\gamma(1) = y$.

Como $\phi \circ \gamma$ es convexo entonces se obtiene que

$$\langle g, u \rangle_x \leq \phi(y) - \phi(x)$$

En efecto,

$$\begin{aligned} \phi(\gamma(t)) &\leq (1-t)\phi(\gamma(0)) + t\phi(\gamma(1)) \\ &= \phi(\gamma(0)) - t\phi(\gamma(0)) + t\phi(\gamma(1)) \\ &= \phi(\gamma(0)) - t\phi(x) + t\phi(y) \\ &= \phi(\gamma(0)) + t(\phi(y) - \phi(x)) \end{aligned}$$

Luego

$$\begin{aligned} \phi(\gamma(t)) - \phi(\gamma(0)) &\leq t(\phi(y) - \phi(x)) \\ \frac{\phi(\gamma(t)) - \phi(\gamma(0))}{t} &\leq \phi(y) - \phi(x) \end{aligned}$$

Por otro lado,

$$\begin{aligned} \langle g, u \rangle_x &= \langle \text{grad}\phi(x(t)), u(t) \rangle_x \\ &= \langle \text{grad}\phi(\gamma(0)), \dot{\gamma}(0) \rangle_x \\ &= \frac{d}{dt}\phi(\gamma(0)) \end{aligned}$$

Como ϕ es decreciente ver teorema (2.0.12) entonces

$$\langle g, u \rangle_x = \frac{d}{dt}\phi(\gamma(0)) \leq \frac{\phi(\gamma(t)) - \phi(\gamma(0))}{t} \leq \phi(y) - \phi(x),$$

así, $\langle g, u \rangle_x \leq \phi(y) - \phi(x)$.

Lo anterior en no positivo, pues ϕ es decreciente y $y \in L$, además es nulo si

$$\phi(x) = \phi(y) \leq \phi_\infty = \lim_{t \rightarrow \infty} \phi(x(t)) \leq \phi(x).$$

Así $\phi(x(t))$ será constante en $[0, \infty)$, el cual quiere decir con (2.0.2) y (2.0.5) que $g = 0$, en este caso la trayectoria es estacionaria y la proposición es facil ya que

$$\begin{aligned} \rho(x(t), y) &= |u(t)|_{x(t)} \\ &= |\dot{\gamma}(0)|_{x(t)} = 0, \end{aligned}$$

de aquí $\rho(x(\cdot), y)$ es constante y por tanto es decreciente y convergente.

Nosotros nos restringiremos al caso cuando $\langle g, u \rangle_x < 0$. Desde entonces el opuesto de u pertenece a el cono normal en el punto x de el conjunto $\{z \in M / \rho(y, z) \leq \rho(y, x)\}$, esto bastará para mostrar el decrecimiento deseado.

Daremos en lo siguiente una idea geométrica de este hecho ver Fig2.1.3

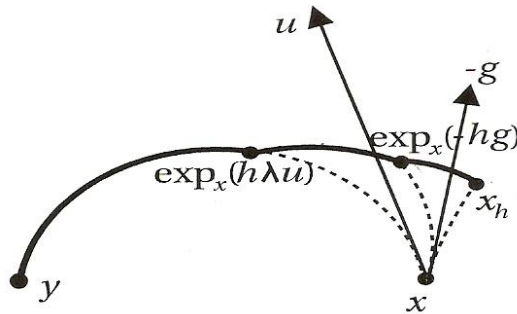


FIGURA 2.1.3

Como ambas trayectorias $h \rightarrow x_h$, $h \rightarrow \exp_x(-hg)$ son de clase C^1 y tiene la misma condición inicial de orden 0. Ver Fig2.1.4.

tenemos

$$\rho(x_h, \exp_x(-hg)) = o(h), \tag{2.1.2}$$

en efecto,

$$\rho(x_h, \exp_x(-hg)) = o(h) \Rightarrow \lim_{h \rightarrow 0^+} \frac{\rho(x_h, \exp_x(-hg))}{h} = 0$$

Definamos dos geodésicas $h \xrightarrow{\beta} x_h$ y $h \xrightarrow{\alpha} \exp_x(-hg)$, donde $\beta(0) = \alpha(0)$.

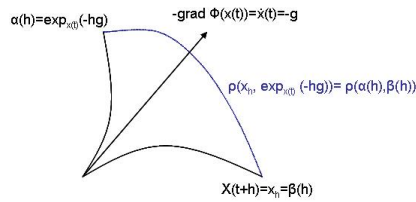


Ilustración de funciones de condición inicial de orden 0

FIGURA 2.1.4

LLamemos $\gamma_h(s)$ al segmento geodésico minimal que conecta al punto $\beta(h)$ con $\alpha(h)$ definamos la función $f(h) = \rho(x_h, \exp_x(-hg)) \geq 0$; $f : [0, 1] \rightarrow \mathbb{R}$ para algún $r > 0$, queremos probar que $\lim_{h \rightarrow 0^+} \frac{f(h)}{h} = 0$.

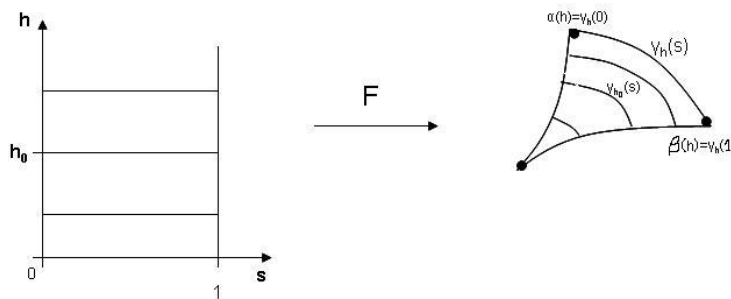
Definamos otra función $F : [0, 1] \times \mathbb{R}^+ \rightarrow M$ por $F(s, h) = \gamma_h(s)$ donde

$$F(0, h) = \gamma_h(0) = \alpha(h)$$

$$F(1, h) = \gamma_h(1) = \beta(h)$$

$$F(s, h_0) = \gamma_{h_0}(s)$$

$$F(s, 0) = \gamma_0(s) = x$$



Función $F : [0, 1] \times \mathbb{R} \rightarrow M$ por $F(s, h) = \gamma_h(s)$

FIGURA 2.1.5

De aquí $\gamma'_h(s) = \frac{\partial F}{\partial s}(s, h)$, supondremos que $\gamma'_h(s)$ es parametrizada por longitud de arco, $\|\gamma'_h\| = 1$ y así

$$L(\gamma_h) = \int_0^1 \|\gamma'_h(s)\| ds = \int_0^1 1 ds = 1$$

Pero $L(\gamma_h) = \rho(\alpha(h), \beta(h)) = \rho(x_h, \text{grad}(-hg)) = f(h) = \left\| \frac{\partial F}{\partial s}(s, h) \right\|$, Así

$$\lim_{h \rightarrow 0^+} \frac{\rho(x_h, \text{grad}(-hg))}{h} = \lim_{h \rightarrow 0^+} \frac{\left\| \frac{\partial F}{\partial s}(s, h) \right\|}{h}$$

Ahora definamos $H : [0, r] \rightarrow \mathbb{R}$ por

$$H(h) = \frac{1}{2} [f(h)]^2 = \frac{1}{2} \left\| \frac{\partial F}{\partial s}(s, h) \right\|^2 = \frac{1}{2} \left\langle \frac{\partial F}{\partial s}(s, h), \frac{\partial F}{\partial s}(s, h) \right\rangle$$

H es C^∞ y

$$H'(h) = \frac{1}{2} \left\langle \frac{D}{\partial h} \frac{\partial F}{\partial s}(s, h), \frac{\partial F}{\partial s}(s, h) \right\rangle + \frac{1}{2} \left\langle \frac{\partial F}{\partial s}(s, h), \frac{D}{\partial h} \frac{\partial F}{\partial s}(s, h) \right\rangle,$$

luego,

$$H'(h) = \left\langle \frac{D}{\partial h} \frac{\partial F}{\partial s}(s, h), \frac{\partial F}{\partial s}(s, h) \right\rangle \Rightarrow H'(0) = \left\langle \frac{D}{\partial h} \frac{\partial F}{\partial s}(s, 0), \frac{\partial F}{\partial s}(s, 0) \right\rangle = 0$$

Además

$$\begin{aligned} H''(h) &= \frac{D}{dh} H'(h) = \frac{D}{dh} \left\langle \frac{D}{\partial h} \frac{\partial F}{\partial s}(s, h), \frac{\partial F}{\partial s}(s, h) \right\rangle = \left\langle \frac{D^2}{\partial h^2} \frac{\partial F}{\partial s}(s, h), \frac{\partial F}{\partial s}(s, h) \right\rangle + \left\langle \frac{D}{\partial h} \frac{\partial F}{\partial s}(s, h), \frac{D}{\partial h} \frac{\partial F}{\partial s}(s, h) \right\rangle \\ &= \left\langle \frac{D^2}{\partial h^2} \frac{\partial F}{\partial s}(s, h), \frac{\partial F}{\partial s}(s, h) \right\rangle + \left\| \frac{D}{\partial h} \frac{\partial F}{\partial s}(s, h) \right\|^2 \end{aligned}$$

Así

$$H''(0) = \left\langle \frac{D^2}{\partial h^2} \frac{\partial F}{\partial s}(s, 0), \frac{\partial F}{\partial s}(s, 0) \right\rangle + \left\| \frac{D}{\partial h} \frac{\partial F}{\partial s}(s, 0) \right\|^2 = 0$$

De aquí $H(0) = H'(0) = H''(0) = 0$.

De lo anterior podemos probar que $\lim_{h \rightarrow 0^+} \frac{H(h)}{h} = 0$, en efecto,

$$0 = H'(0) = \lim_{h \rightarrow 0} \frac{H(0+h) - H(0)}{h} = \lim_{h \rightarrow 0} \frac{H(h)}{h},$$

además, por el lema (1.1.1) obtenemos que $\lim_{h \rightarrow 0^+} \frac{H(h)}{h^2} = 0$, Por tanto,

$$\lim_{h \rightarrow 0^+} \frac{f(h)}{h} = \lim_{h \rightarrow 0^+} \frac{\sqrt{2H(h)}}{h} = \lim_{h \rightarrow 0^+} \frac{\sqrt{2H(h)}}{\sqrt{h^2}} = \lim_{h \rightarrow 0^+} \sqrt{\frac{2H(h)}{h^2}} = 0$$

Un argumento del mismo tipo dará

$$\rho(\exp_x(-hg), \exp_x(h\lambda u)) = |-hg - h\lambda u|_x + o(h) \quad (2.1.3)$$

En efecto se prueba que

$$\rho(\exp_x(-hg), \exp_x(h\lambda u)) - |-hg - h\lambda u|_x = o(h),$$

lo que es,

$$\lim_{h \rightarrow 0} \frac{\rho(\exp_x(-hg), \exp_x(h\lambda u)) - |-hg - h\lambda u|_x}{h} = 0$$

$$\Rightarrow \lim_{h \rightarrow 0} \frac{\rho(\exp_x(-hg), \exp_x(h\lambda u))}{h} - \lim_{h \rightarrow 0} \frac{|-hg - h\lambda u|_x}{h} = 0$$

Con un argumento similar al anterior se prueba que

$$\lim_{h \rightarrow 0} \frac{\rho(\exp_x(-hg), \exp_x(h\lambda u))}{h} = 0$$

Esto se prueba definiendo dos funciones $h \xrightarrow{f} \exp_x(-hg)$, $h \xrightarrow{g} \exp_x(h\lambda u)$ donde $f(0) = g(0) = 0$.

Del mismo modo se hace con

$$\lim_{h \rightarrow 0} \frac{|-hg - h\lambda u|_x}{h} = 0,$$

definiendo $h \xrightarrow{\gamma_1} -hg$, $h \xrightarrow{\gamma_2} h\lambda u$, con $\gamma_1(0) = \gamma_2(0) = 0$

Por tanto quedaría probado que $\rho(\exp_x(-hg), \exp_x(h\lambda u)) = |-hg - h\lambda u|_x + o(h)$

Escojamos λ tal que $\langle -g - \lambda u, -g \rangle_x = 0$ (que es $\lambda = \frac{|g|_x^2}{-\langle u, g \rangle_x} > 0$).

Lo anterior es cierto ya que

$$\begin{aligned} (**) \quad \langle -g - \lambda u, -g \rangle_x = 0 &\Leftrightarrow \langle -g, -g \rangle_x + \langle -\lambda u, -g \rangle_x = 0 \\ &\Leftrightarrow \langle -g, -g \rangle_x = -\lambda \langle u, g \rangle_x \\ &\Leftrightarrow \lambda = \frac{|g|_x^2}{-\langle u, g \rangle_x} \quad (1) \end{aligned}$$

Además

$$\begin{aligned} \langle -g - \lambda u, -g \rangle_x = 0 &\Rightarrow |-g - \lambda u|_x |-g|_x \cos \theta = 0 \\ &\Rightarrow |-g - \lambda u|_x = 0 \vee \cos \theta = \frac{\pi}{2} \text{ ó } \frac{3\pi}{2} \\ &\Rightarrow |-g - \lambda u|_x = 0 \Rightarrow -g = \lambda u \quad (*) \end{aligned}$$

Sustituyendo (*) en (1) tenemos que

$$\lambda = \frac{|g|_x^2}{-\langle u, -\lambda u \rangle_x} = \frac{|g|_x^2}{\lambda \langle u, u \rangle_x} = \frac{|g|_x^2}{\lambda |u|_x^2}$$

$$\text{Luego } \lambda^2 = \frac{|g|_x^2}{|u|_x^2} > 0$$

$$\text{Esto dará } |-hg - h\lambda u|_x = h \left((\lambda |u|_x)^2 - |g|_x^2 \right)^{1/2}.$$

Veamos que se cumple

$$\begin{aligned}
 |-hg - h\lambda u|_x &= = (\langle -hg - h\lambda u, -hg - h\lambda u \rangle_x)^{1/2} \\
 &= \left(h^2 \langle -g - \lambda u, -g - \lambda u \rangle_x \right)^{1/2} \\
 &= h (\langle -g - \lambda u, -g \rangle_x + \langle -g - \lambda u, -\lambda u \rangle_x)^{1/2} \\
 &= h (\langle -g, -\lambda u \rangle_x + \langle -\lambda u, -\lambda u \rangle_x)^{1/2} \quad \text{por (**)} \\
 &= h \left(\langle -g, g \rangle_x + \lambda^2 \langle u, u \rangle_x \right)^{1/2} \quad \text{por (*)} \\
 &= h \left((\lambda |u|_x)^2 - |g|_x^2 \right)^{1/2}
 \end{aligned}$$

Finalmente

$$\rho(\exp_x(h\lambda u), y) = \rho(\exp_x(h\lambda u), \exp_x(u)) = |h\lambda u - u|_x = (1 - h\lambda) |u|_x \quad (2.1.4)$$

Combinando las ecuaciones (2.1.2), (2.1.3), (2.1.4), seremos capaces de construir una curva α continua y diferenciable a trozos que une a x_h con y , de longitud

$$\begin{aligned}
 l(\alpha) &= \rho(x_h, \exp_x(-hg)) + \rho(\exp_x(-hg), \exp_x(h\lambda u)) + \rho(\exp_x(h\lambda u), y) \\
 &= o(h) + h \left((\lambda |u|_x)^2 - |g|_x^2 \right)^{1/2} + (1 - h\lambda) |u|_x \\
 &= |u|_x - h\lambda |u|_x + h \left((\lambda |u|_x)^2 - |g|_x^2 \right)^{1/2} + o(h) \\
 &= |u|_x - h \left[\lambda |u|_x - \left((\lambda |u|_x)^2 - |g|_x^2 \right)^{1/2} \right] + o(h)
 \end{aligned}$$

Justamente el corchete de arriba es positivo, verdaderamente $(\lambda |u|_x)^2 > (\lambda |u|_x)^2 - |g|_x^2 \geq 0$
 Para h suficientemente pequeño tenemos que

$$\rho(x_h, y) \leq l(\alpha) \leq |u|_x = \rho(x, y) \quad \diamond.$$

Conforme a este resultado, podemos declarar el siguiente teorema

TEOREMA 2.1.1. *Bajo la completa suposición (H_1) - (H_2) . Toda solución $x(\cdot)$ de (2.0.2) tiene un límite \bar{x} que pertenece al $\arg \min \phi$*

Demostración:

Bajo las suposiciones (H_1) - (H_2) sabemos que el conjunto $L = \{y \in M \mid \phi(y) \leq \phi_\infty\}$ es no vacío, es cerrado por ser la imagen inversa de un conjunto cerrado mediante la función continua ϕ y Además es acotada por definición. Acordándonos del teorema de Hopf-Rinow ver (1.2.5), aquí existe una sucesión t_j tal que $x(t_j)$ converge a algún \bar{x} , pero \bar{x} pertenece a L , y entonces por teorema anterior $\rho(x(\cdot), \bar{x})$ converge y además el límite es cero, es decir, $x(t)$ converge a \bar{x} .

Como en el teorema anterior se demostró que la convexidad implica que

$$\begin{aligned}
0 \leq \phi(x(t)) - \phi(y) &\leq \langle -\text{grad}\phi(x(t)), u(t) \rangle_{x(t)} \\
&\leq |\text{grad}\phi(x(t))|_{x(t)} |u(t)|_{x(t)} \\
&= |\text{grad}\phi(x(t))|_{x(t)} \rho(x(t), y)
\end{aligned}$$

Una parte tiende a cero $\left(\lim_{t \rightarrow \infty} |\text{grad}\phi(x(t))|_{x(t)} = 0 \right)$ por teorema (2.0.13) y la otra es acotada

$$\begin{aligned}
0 \leq \lim_{t \rightarrow \infty} \phi(x(t)) - \lim_{t \rightarrow \infty} \phi(y) &\leq \lim_{t \rightarrow \infty} |\text{grad}\phi(x(t))|_{x(t)} \rho(x(t), y) \\
\therefore \phi(\bar{x}) = \lim_{t \rightarrow \infty} \phi(x(t)) &= \phi(y) = \inf_M \phi \quad \diamond.
\end{aligned}$$

CAPÍTULO 3

EJEMPLO: FUNCIÓN DE ROSENBROCK

Sea $(\mathbb{R}^2, g_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix})$, en el plano Euclidiano.

La función $F : \mathbb{R}^2 \rightarrow \mathbb{R}$, $F(y) = 100y_1^2 + y_2^2$, $y = (y_1, y_2)$ es convexa con respecto a g_0 .

Consideremos la transformación de coordenadas no lineales siguiente

$$\begin{cases} y_1 = 1 - x_1 \\ y_2 = x_2 - x_1^2 \end{cases}$$

Para esta transformación hallemos la métrica Riemanniana inducida por la Euclídea. Supongamos que

$$g(x_1, x_2) = \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix}$$

Por comentario hecho en los preliminares (ver 1.2.19) la métrica Riemanniana se puede calcular

$$ds^2 = \sum_{i,j} g_{ij} dx_i dx_j$$

esto es

$$ds^2 = g_{11} dx_1^2 + 2g_{12} dx_1 dx_2 + g_{22} dx_2^2 \quad (3.0.1)$$

Hallemos g_{11}, g_{12}, g_{22} , sabemos que la métrica Euclídea viene dada por $ds^2 = dy_1^2 + dy_2^2$ (c2),

$$\begin{aligned} dy_1 = -dx_1 &\Rightarrow dy_1^2 = dx_1^2 \\ dy_2 = dx_2 - 2x_1 dx_1 &\Rightarrow dy_2^2 = dx_2^2 - 4x_1 dx_1 dx_2 + 4x_1^2 dx_1^2 \end{aligned}$$

Luego igualando (c2) y (3.0.1) se obtiene

$$\begin{aligned} g_{11} dx_1^2 + 2g_{12} dx_1 dx_2 + g_{22} dx_2^2 &= dy_1^2 + dy_2^2 \\ &= (1 + 4x_1^2) dx_1^2 - 4x_1 dx_1 dx_2 + dx_2^2 \end{aligned}$$

lo que implica que $g_{11} = 1 + 4x_1^2$, $g_{12} = g_{21} = -2x_1$, $g_{22} = 1$.

Así la variedad Riemanniana

$$\left(\mathbb{R}^2, g_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right)$$

es cambiada por medio de esta transformación por $\left(\mathbb{R}^2, g(x) = \begin{pmatrix} 1 + 4x_1^2 & -2x_1 \\ -2x_1 & 1 \end{pmatrix} \right)$, $x = (x_1, x_2)$ y F es cambiada a la famosa función de Rosenbrock.

$$\phi : \mathbb{R}^2 \rightarrow \mathbb{R} \quad , \quad \phi(x_1, x_2) = 100(x_2 - x_1)^2 + (1 - x_1)^2 \quad (3.0.2)$$

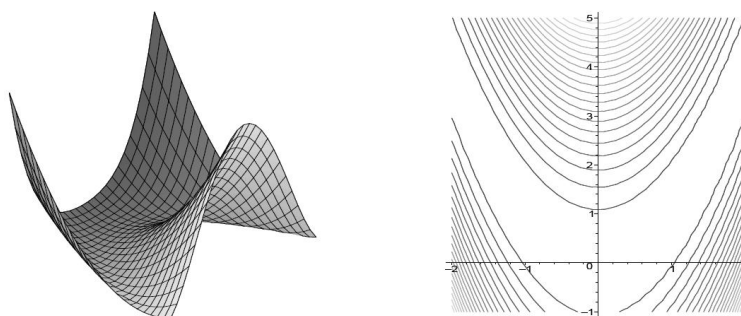


FIGURA 3.0.1: Gráfica de la función de Rosenbrock (Izquierda) y conjunto de nivel de dicha función en el espacio Euclídeo (Derecha)

Esta función es convexa con respecto a g , en efecto, encontremos primeramente $g^{-1}(x)$

$$\begin{pmatrix} u & v \\ z & w \end{pmatrix} \begin{pmatrix} 1 + 4x_1^2 & -2x_1 \\ -2x_1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\begin{cases} u(1 + 4x_1^2) - 2x_1v = 1 & \text{(I)} \\ u(-2x_1) + v = 0 & \Rightarrow v = 2ux_1 & \text{(II)} \\ z(1 + 4x_1^2) = 0 & \text{(III)} \\ -2x_1z + w = 1 & \Rightarrow w = 1 + 2x_1z & \text{(IV)} \end{cases}$$

Sustituyendo el valor de v y w

$$\begin{aligned} u + 4ux_1^2 - 2x_1(2ux_1) &= 1 \\ u + 4ux_1^2 - 4ux_1^2 &= 1 \quad \Rightarrow u = 1 \end{aligned}$$

$$\begin{aligned} z + 4zx_1^2 - 2x_1(1 + 2x_1z) &= 0 \\ z + 4zx_1^2 - 2x_1 - 4zx_1^2 &= 0 \quad \Rightarrow z = 2x_1 \end{aligned}$$

Sustituyendo $u = 1$ y $z = 2x_1$ en las ecuaciones (II) y (IV) se obtiene:

$$\begin{aligned} v &= 2(1)x_1 \Rightarrow v = 2x_1 \\ w &= 1 + 2x_1(2x_1) \Rightarrow w = 1 + 4x_1^2 \end{aligned}$$

$$\text{Así } g^{-1}(x) = \begin{pmatrix} 1 & 2x_1 \\ 2x_1 & 1 + 4x_1^2 \end{pmatrix}.$$

Denotando

$$g^{11} = 1, \quad g^{12} = 2x_1, \quad g^{21} = 2x_1, \quad g^{22} = 1 + 4x_1^2,$$

y sabiendo que

$$g_{11} = 1 + 4x_1^2, \quad g_{12} = g_{21} = -2x_1, \quad g_{22} = 1$$

, se pueden hallar los símbolos de Christofel:

$$\Gamma_{ij}^h = \frac{1}{2} \sum_{k=1}^2 g^{hk} \left(\frac{\partial g_{jk}}{\partial x_i} + \frac{\partial g_{ik}}{\partial x_j} - \frac{\partial g_{ij}}{\partial x_k} \right) \quad (*)$$

sabiendo que:

$$\begin{aligned} \frac{\partial g_{11}}{\partial x_1} &= 8x_1; & \frac{\partial g_{11}}{\partial x_2} &= 0 \\ \frac{\partial g_{12}}{\partial x_1} &= -2; & \frac{\partial g_{12}}{\partial x_2} &= 0 \\ \frac{\partial g_{21}}{\partial x_1} &= -2; & \frac{\partial g_{21}}{\partial x_2} &= 0 \\ \frac{\partial g_{22}}{\partial x_1} &= 0; & \frac{\partial g_{22}}{\partial x_2} &= 0 \end{aligned}$$

Desarrollemos (*)

$$\begin{aligned} \Gamma_{11}^1 &= \frac{1}{2} g^{11} \left(\frac{\partial g_{11}}{\partial x_1} + \frac{\partial g_{11}}{\partial x_1} - \frac{\partial g_{11}}{\partial x_1} \right) + \frac{1}{2} g^{12} \left(\frac{\partial g_{12}}{\partial x_1} + \frac{\partial g_{12}}{\partial x_1} - \frac{\partial g_{11}}{\partial x_2} \right) \\ &= \frac{1}{2} 8x_1 + \frac{1}{2} (-2 + (-2) - 0) \\ &= 4x_1 - 4x_1 \Rightarrow \Gamma_{11}^1 = 0 \end{aligned}$$

$$\begin{aligned} \Gamma_{12}^1 &= \frac{1}{2} g^{11} \left(\frac{\partial g_{21}}{\partial x_1} + \frac{\partial g_{11}}{\partial x_2} - \frac{\partial g_{12}}{\partial x_1} \right) + \frac{1}{2} g^{12} \left(\frac{\partial g_{22}}{\partial x_1} + \frac{\partial g_{12}}{\partial x_2} - \frac{\partial g_{12}}{\partial x_2} \right) \\ &= \frac{1}{2} (-2 + 0 - (-2)) + \frac{1}{2} 2x_1 (0) \Rightarrow \Gamma_{12}^1 = 0 \end{aligned}$$

$$\begin{aligned}\Gamma_{21}^1 &= \frac{1}{2}g^{11} \left(\frac{\partial g_{11}}{\partial x_2} + \frac{\partial g_{21}}{\partial x_1} - \frac{\partial g_{21}}{\partial x_1} \right) + \frac{1}{2}g^{12} \left(\frac{\partial g_{12}}{\partial x_2} + \frac{\partial g_{22}}{\partial x_1} - \frac{\partial g_{21}}{\partial x_2} \right) \\ &= \frac{1}{2}(-2 + 0 - (-2)) + \frac{1}{2}2x_1(0) \Rightarrow \Gamma_{21}^1 = 0\end{aligned}$$

$$\begin{aligned}\Gamma_{22}^1 &= \frac{1}{2}g^{11} \left(\frac{\partial g_{21}}{\partial x_2} + \frac{\partial g_{21}}{\partial x_2} - \frac{\partial g_{22}}{\partial x_1} \right) + \frac{1}{2}g^{12} \left(\frac{\partial g_{22}}{\partial x_2} + \frac{\partial g_{22}}{\partial x_2} - \frac{\partial g_{22}}{\partial x_2} \right) \\ &= 0\end{aligned}$$

$$\begin{aligned}\Gamma_{11}^2 &= \frac{1}{2}g^{21} \left(\frac{\partial g_{11}}{\partial x_1} + \frac{\partial g_{11}}{\partial x_1} - \frac{\partial g_{11}}{\partial x_1} \right) + \frac{1}{2}g^{22} \left(\frac{\partial g_{12}}{\partial x_1} + \frac{\partial g_{12}}{\partial x_1} - \frac{\partial g_{11}}{\partial x_2} \right) \\ &= \frac{1}{2}(2x_1)(8x_1 + 8x_1 - 8x_1) + \frac{1}{2}(1 + 4x_1^2)(-2 - 2) \\ &= 8x_1^2 - 2 - 8x_1^2 \Rightarrow \Gamma_{11}^2 = -2\end{aligned}$$

$$\begin{aligned}\Gamma_{12}^2 &= \frac{1}{2}g^{21} \left(\frac{\partial g_{21}}{\partial x_1} + \frac{\partial g_{11}}{\partial x_2} - \frac{\partial g_{12}}{\partial x_1} \right) + \frac{1}{2}g^{22} \left(\frac{\partial g_{22}}{\partial x_1} + \frac{\partial g_{12}}{\partial x_2} - \frac{\partial g_{12}}{\partial x_2} \right) \\ &= \frac{1}{2}(2x_1)(-2 - (-2)) + \frac{1}{2}(1 + 4x_1^2)(0) \\ &= 0\end{aligned}$$

$$\begin{aligned}\Gamma_{21}^2 &= \frac{1}{2}g^{21} \left(\frac{\partial g_{11}}{\partial x_2} + \frac{\partial g_{21}}{\partial x_1} - \frac{\partial g_{21}}{\partial x_1} \right) + \frac{1}{2}g^{22} \left(\frac{\partial g_{12}}{\partial x_2} + \frac{\partial g_{22}}{\partial x_1} - \frac{\partial g_{21}}{\partial x_2} \right) \\ &= \frac{1}{2}(2x_1)(-2 - (-2)) + \frac{1}{2}(1 + 4x_1^2)(0) \\ &= 0\end{aligned}$$

$$\begin{aligned}\Gamma_{22}^2 &= \frac{1}{2}g^{21} \left(\frac{\partial g_{21}}{\partial x_2} + \frac{\partial g_{21}}{\partial x_2} - \frac{\partial g_{22}}{\partial x_1} \right) + \frac{1}{2}g^{22} \left(\frac{\partial g_{22}}{\partial x_2} + \frac{\partial g_{22}}{\partial x_2} - \frac{\partial g_{22}}{\partial x_2} \right) \\ &= \frac{1}{2}(2x_1)(0) + \frac{1}{2}(1 + 4x_1^2)(0) \Rightarrow \Gamma_{22}^2 = 0\end{aligned}$$

Calculemos ahora las componentes del Hessiano $Hess\phi$.

$$\phi_{jk} = \frac{\partial^2 \phi}{\partial x_i \partial x_k} - \sum_{i=1}^2 \Gamma_{jk}^i \frac{\partial \phi}{\partial x_i}$$

Es decir:

$$\begin{aligned}\phi_{11} &= \frac{\partial^2 \phi}{\partial x_1^2} - \Gamma_{11}^2 \frac{\partial \phi}{\partial x_2} \\ \phi_{21} &= \frac{\partial^2 \phi}{\partial x_2 \partial x_1} \\ \phi_{12} &= \frac{\partial^1 \phi}{\partial x_i \partial x_2} \\ \phi_{22} &= \frac{\partial^2 \phi}{\partial x_2^2}\end{aligned}$$

Donde

$$\begin{aligned}\frac{\partial \phi}{\partial x_1} &= -400x_2x_1 + 400x_1^3 - 2 + 2x_1 \\ \frac{\partial^2 \phi}{\partial x_1^2} &= 1200x_1^2 + 2 - 400x_2 \\ \frac{\partial \phi}{\partial x_2} &= 200x_2 - 200x_1^2 \\ \frac{\partial^2 \phi}{\partial x_2^2} &= 200 \\ \frac{\partial^1 \phi}{\partial x_i \partial x_2} &= -400x_1\end{aligned}$$

Así sustituyendo los valores anteriores en las componenetes del Hessiano se tiene:

$$\begin{aligned}\phi_{11} &= 1200x_1^2 - 400x_2 + 2 - (-2)(200x_2 - 200x_1^2) \\ &= 1200x_1^2 - 400x_2 + 400x_2 - 400x_1^2 + 2 \\ &= 800x_1^2 + 2 \\ \phi_{21} &= -400x_1 \\ \phi_{12} &= -400x_1 \\ \phi_{22} &= 200\end{aligned}$$

Luego el Hessiano de ϕ es:

$$\text{Hess}\phi = \begin{pmatrix} 800x_1^2 + 2 & -400x_1 \\ -400x_1 & 200 \end{pmatrix}$$

Probemos que el Hessiano de la función ϕ es semidefinido positivo.

Sea un vector cualquiera $a = (a_1, a_2) \in M$. Si ϕ es convexa es por que

$$a(\text{hess}\phi)a^t \geq 0$$

Veamos que es cierto

$$\begin{aligned}
 & \begin{pmatrix} a_1 & a_2 \end{pmatrix} \begin{pmatrix} 800x_1^2 + 2 & -400x_1 \\ -400x_1 & 200 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \\
 &= \begin{pmatrix} a_1 800x_1^2 + 2a_1 - 400x_1 a_2 & -400x_1 a_1 + 200a_2 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \\
 &= 800x_1^2 a_1^2 + 2a_1^2 - 400x_1 a_2 a_1 - 400x_1 a_1 a_2 + 200a_2^2 \\
 &= 200 \left[4x_1^2 a_1^2 - 4x_1 a_1 a_2 + a_2^2 \right] + 2a_1^2 \\
 &= 200 [2x_1 a_1 - a_2]^2 + 2a_1^2 \\
 &\geq 0
 \end{aligned}$$

Por tanto ϕ es una función convexa.

Se presentará un algoritmo numérico hecho para este ejemplo. Contruyamos una sucesión de puntos obtenidos por una discretización de la ecuación (2.0.2)

$$x_{k-1} = x_k + \lambda_k v_k \tag{3.0.3}$$

El punto inicial x_0 es dado.

Se discutirá primero acerca de la escogencia del tamaño de paso. Por un lado, este se escogerá de la siguiente manera.

Algoritmo 1. Escogencia del tamaño del paso.

1. $\lambda = 1$, defina $y = x_k + \lambda v_k$ y $z = x_k + \frac{\lambda}{2} v_k$

2. mientras que $\phi(y) > \phi(z)$, hacer

$$\lambda = \frac{\lambda}{2},$$

$$y = z,$$

$$z = x_k + \frac{\lambda}{2} v_k.$$

3. Finalmente $\lambda_k = \lambda$ y $x_{k+1} = y$.

Este es un tipo de búsqueda lineal óptima reducida al intervalo $[0, 1]$. Ciertamente esto aproxima a la escogencia teórica

$$\lambda_k \in \operatorname{argmin} \{ \phi(x_k + \lambda v_k), \lambda \in [0, 1] \}$$

Por otro lado se calcula la sucesión obtenida por (3.0.3) con un tamaño de paso todavía dado por una búsqueda lineal óptima reducida al intervalo $[0;0,01]$. Esto es suficiente cambiar $\lambda = 1$ por $\lambda = 0,01$ en el algoritmo previo.

La dirección de descenso v_k es escogida primeramente como el opuesto del gradiente Riemanniano

$$v_k = -grad\phi(x_k) \tag{3.0.4}$$

que provee trayectorias lineales en la figura (3.0.2), y segundo como el opuesto de el gradiente Euclídeo

$$v_k = -\nabla\phi(x_k) \tag{3.0.5}$$

que provee las trayectorias puntuales.

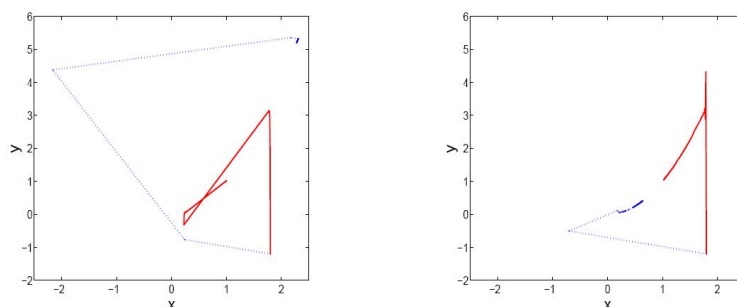


FIGURA 3.0.2: Comparación de la sucesión de puntos cuando $v_k = -\nabla\phi(x_k)$ (línea azul) y cuando $v_k = -grad\phi(x_k)$ (línea roja) con un tamaño de paso en el intervalo $[0, 1]$ (Izquierda),y Comparación de la sucesión de puntos cuando $v_k = -\nabla\phi(x_k)$ (línea azul) y cuando $v_k = -grad\phi(x_k)$ (línea roja) con un tamaño de paso en el intervalo $[0;0,1]$ (Derecha)

Como se esperó, los experimentos fueron hechos,escogiendo puntos al azar elegidos en $[-5,5] \times [-5,5]$, se muestra que ambas sucesiones parece que convergieran al mínimo $(1,1)$. Ciertamente (3.0.3) con (3.0.4) es una discretización de 2.0.2, cuyas trayectorias convergen al mínimo pues ϕ es convexa, se aplica entonces el teorema (2.1.1). En el otro lado (3.0.3) con (3.0.5) es una discretización de

$$\dot{x}(t) + \nabla\phi(x(t)) = 0 \tag{3.0.6}$$

Es conocido que el descenso acelerado (Euclideano) provee trayectorias convergentes bajo suposiciones de analiticidad de ϕ . La intención que se tiene es comparar los dos métodos.

Presentamos en la figura (3.0.3) la convergencia de las curvas obtenidas, relacionadas en el experimento presentado en la figura (3.0.2). El valor $\phi(x_i)$ es graficado en oposición a el número i de la iteración.

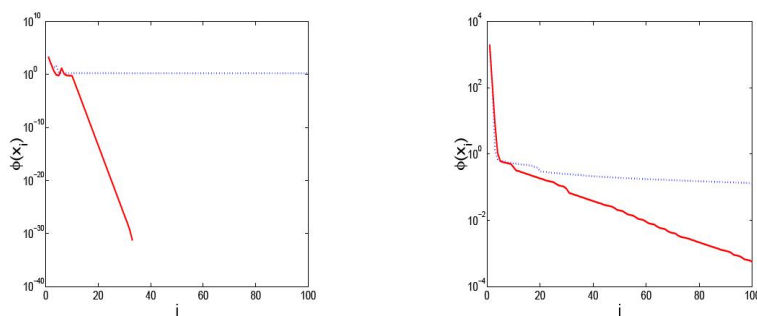


FIGURA 3.0.3: Comparación de los valores de la función ϕ con el método gradiente Euclídeo (línea azul) y con el método gradiente Riemanniano (Línea Roja), en oposición a el número i de la iteración.

Para ambas escogencias de tamaño de paso, el método de descenso Riemanniano (lineal) es más eficiente que el Euclidean (puntual).

La búsqueda lineal en $[0,0,1]$ parece estabilizar el método mientras los puntos no están en el fondo del valle "U", es decir, mientras el gradiente sea enorme. Mas aún podemos imaginar que esto reduce el número de cálculos computacionales hechos en el algoritmo 1, suministrando el tamaño de paso óptimo, pero esto ocurre debido que los puntos están en el fondo del valle, finalmente pasos de tamaño 1 o cuando el paso está entre 0,1 y 1 acelera hacia la convergencia.

Actualmente esto es cierto para la versión Riemanniana, pero esto puede producir inestabilidad para la versión Euclidean, ya que minimizarla en algunos casos, la sucesión parece ir de un lado de la "U" del valle hacia el otro lado.

CONCLUSIONES

Se ha estudiado las propiedades de la famosa Ecuación de Evolucion del método de descenso acelerado sobre variedades de Riemann, para resolver problemas de minimización de funciones.

Cuando (M, ρ) es completo, la función $\phi : M \rightarrow \mathbb{R}$ es C^1 y localmente lipschitz-continua, además de ser acotada inferiormente, se demostró que el problema de Cauchy:

$$\begin{cases} \dot{x}(t) + \text{grad}\phi(x(t)) = 0 \\ x(0) = x_0, \end{cases} \quad (3.0.7)$$

tiene única solución en $(0, +\infty)$, $|\dot{x}(t)|_{x(t)}$ es integrable a la 2 y además se obtiene que $\lim_{t \rightarrow \infty} |\text{grad}\phi(x)|_{x(t)} = 0$.

Sin embargo era necesario considerar que la función ϕ es convexa en M y que $\text{argmin} \neq \emptyset$ para poder conseguir que la distancia ρ de un punto inicial $x_0 \in M$ a un punto

$$y \in L = \{y \in M / \phi(y) \leq \phi_\infty\}$$

decreciera y convergiera, y así utilizar este resultado para finalmente garantizar la convergencia de la solución $x(\cdot)$ a un punto límite \bar{x} perteneciente a el $\text{argmin}\phi$.

Adicionalmente se estudió la famosa función de Rosenbrock, demostrando que con una métrica Riemanniana y por medio de un algoritmo basado en los teoremas previos existe un mínimo para esta función y se obtiene la convergencia.

REFERENCIAS

- [1] Julien Steepest descent Method on a Riemannian manifold: The convex case, *Balkan Journal of Geometry and Its Applications* Vol.12 No. 2 2007pp.98-106.
- [2] E. Lages, *Análise Real*, Vol.2, IMPA, R.J, Brazil, 2004, 1ra Edición.
- [3] Hiriart-Urruty and C. Lemarechal, "Convex Analysis and Minimization Algorithm I", Springer Verlag, New York, 1996.
- [4] M.P. Carmo, *Geometria Riemanniana*, Río de Janeiro, 2005.
- [5] C. Udriste, *Convex Function and Optimization Methods on Riemannian Manifolds*, Kluwer academic Publishers, 1994.
- [6] L. y George Bachman , *Functional Analysis*, Dover Publications , 2000, 1ra Edición.
- [7] Jorge, *Ecuacoes Diferenciais Ordinárias*, Río de Janeiro, Abril 1979.