

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



Minimización de una función de Morse sobre una variedad Riemanniana

AUTOR: BR. YORISBEL MARGARETTE PEÑA VELASQUEZ
TUTOR: DR. EIBAR HERNÁNDEZ (UCLA)

TRABAJO ESPECIAL DE GRADO

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

BARQUISIMETO, VENEZUELA

Junio, 2011.

“Todo lo puedo en Cristo que me Fortalece”

Dedicado a:

Dios todopoderoso

Mis padres...

Mi flaco y mis hermanos.

Agradecimientos

Primeramente a *Dios*, por estar conmigo en cada paso, por fortalecer mi corazón e iluminar mi mente, por darme la oportunidad de vivir, regalarme una familia maravillosa y por haber puesto en mi camino a aquellas personas que han sido mi soporte y compañía durante todo el periodo de estudio.

A *Mis Padres*, por darme la estabilidad emocional y económica para llegar a este logro y por enseñarme que todo esfuerzo al final es una recompensa.

A *Mi Flaco precioso*, por ese *Amor* que me brinda día a día, por su apoyo incondicional, y por su confianza en esos momentos tan difíciles cuando creí que no podía continuar.

A mi *tutor* Eibar Hernández, por su dedicación, por ayudarme en la realización de éste proyecto, por explicarme y transmitirme de manera paciente sus conocimientos y por ser más que mi tutor un buen amigo y consejero.

Al *profesor* Javier Hernández por explicarme latex y ayudarme con los gráficos de la tesis.

A mis *compañeros y amigos* Ninoska, Yohana, Hibelmar, Rona, Isme, Naty, M.Andreina, Laura, Mariú, Yasumir y M. Luisa, por todos esos momentos de estudios, risas, llantos, consejos, e innumerable momentos gratos que marcaron mi vida durante la carrera.

A todas aquellas personas que de una u otra forma colocaron un granito de arena para contribuir a que se culminara esta meta.

Gracias...

Resumen

En este trabajo se extiende el Método de Máximo Descenso Euclideo en variedades Riemanniana, donde se presenta el problema de minimizar una función con ciertas restricciones de igualdad. Estas restricciones se reescriben a conveniencia como una variedad y así se estudia sobre ella el concepto de curva Geodésica, la cual define la longitud minimal entre dos puntos de la variedad, de ésta manera se observa que dicha curva representa en la variedad el papel de las rectas en el espacio Euclideo, de allí surge la idea de crear el Método de Máximo Descenso a lo largo de Geodésicas, el cual es muy parecido al Método de Máximo Descenso Euclidiano, pero éste, en lugar de hacer la búsqueda por rectas la hace por Geodésica.

Introducción

El propósito de este trabajo es estudiar y desarrollar el artículo [2] en el cual se extienden los métodos conocidos, relacionados con el gradiente en la minimización sin restricciones de una función de \mathbb{R}^n a valores reales, al considerar el siguiente problema no lineal restringido:

$$\text{mín}\{f(x)/x \in \mathbb{R}^n : C_i(x) = 0; i = 1, 2, \dots, m\},$$

donde $m \leq n$, f y C_i son funciones de \mathbb{R}^n a valores reales, las cuales se suponen diferenciables \mathcal{C}^σ ($\sigma \geq 2$ a menos que se especifique lo contrario).

Se define $C : \mathbb{R}^n \rightarrow \mathbb{R}^m$ por $C(x) = (C_1(x), C_2(x), \dots, C_m(x))$ y se supone que cero es un valor regular de la función C , esto es que $C'(x)$ pertenece a $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m) = \{T : \mathbb{R}^n \rightarrow \mathbb{R}^m / T \text{ es lineal}\}$ y es de rango completo, para todo $x \in M = C^{-1}(0_m) = \{x \in \mathbb{R}^n / C(x) = 0_m\}$. Esta hipótesis de regularidad implica que existe una submatriz $m \times m$ de la matriz jacobiana C en $x \in M$, denotada por $[A_x]_{m \times n}$, la cual es no singular.

En el capítulo 1 se presentarán una serie de definiciones detalladas sobre curvas diferenciables, superficies, plano tangente, primera y segunda forma fundamental, campo vectorial, derivada covariante, transporte paralelo, variedades diferenciables y Riemanniana, entre otras definiciones que serán de gran ayuda para entender este trabajo.

En el capítulo 2 se prueba que $M = C^{-1}(0_m) = \{x \in \mathbb{R}^n / C(x) = 0_m\}$ es una subvariedad $(n - m)$ dimensional de \mathbb{R}^n , se da nuevamente el concepto de plano tangente pero relacionado con el núcleo de la matriz $[A_x]_{m \times n}$ y se dota $T_x M$ con una forma bilineal positiva $\gamma_x : T_x M \times T_x M \rightarrow \mathbb{R}$, la cual se llamará métrica Riemanniana. Esta forma bilineal γ_x es una función suave de x , y define una estructura Riemanniana γ en M ; luego se define el campo gradiente de f en M y se define el método gradiente reducido de minimización sin restricciones.

Este método consiste en la generación desde una aproximación x^k a un nuevo iterado x^{k+1} sobre la curva geodésica a partir de x^k y tangente a una dirección definida por el gradiente de f sobre M en x^k .

En el capítulo 3 se presentarán importantes resultados en la geometría de una variedad Riemanniana que se utilizan en este trabajo; también se establecerán las condiciones de optimalidad sobre la variedad M ; se definirá el método de máximo descenso a lo largo de geodésicas y serán analizadas sus propiedades de convergencia.

Índice general

| | |
|--|------------|
| Resumen | III |
| Introducción | IV |
| 1. Preliminares de Geometría | 1 |
| 1.1. Curvas Regulares y Longitud de arco | 1 |
| 1.2. La diferencial | 12 |
| 1.3. Superficie regular e imagen inversa de un valor regular | 17 |
| 1.4. La primera forma fundamental | 28 |
| 1.5. Transporte Paralelo y Geodésicas | 37 |
| 1.6. La Función Exponencial. | 41 |
| 1.7. Superficies Abstractas | 42 |
| 1.8. Función de Morse | 43 |

| | |
|--|-----------|
| 2. Preliminares de Optimización | 46 |
| 2.1. El conjunto de restricciones M visto como una variedad | 50 |
| 2.2. Geometría de la variedad M | 55 |
| 2.2.1. Estructura Riemanniana sobre M | 55 |
| 2.2.2. Diferenciación covariante y Geodésicas | 57 |
| 3. Métodos de descenso por Geodésicas | 63 |
| 3.1. Gradiente de una función sobre M | 67 |
| 3.2. Método de máximo descenso a lo largo de Geodésicas | 68 |
| 3.3. Método de Newton a lo largo de Geodésicas | 72 |

Índice de figuras

| | |
|--|----|
| 1.1. Hélice | 2 |
| 1.2. El Pez | 2 |
| 1.3. Longitud de una curva | 3 |
| 1.4. Poligonal | 4 |
| 1.5. Circunferencia | 5 |
| 1.6. Triedro de Frenet | 12 |
| 1.7. La diferencial | 13 |
| 1.8. Gráfico de una función | 18 |
| 1.9. Valor Regular | 22 |
| 1.10. Diferenciabilidad de f | 24 |
| 1.11. Definición de Plano Tangente | 25 |

| | |
|---|----|
| 1.12. Plano Tangente | 26 |
| 1.13. Primera Forma Fundamental | 29 |
| 1.14. Área de un paralelogramo | 32 |
| 1.15. Derivada covariante | 38 |
| 1.16. Campo Vectorial Paralelo | 39 |
| 1.17. Geodésica | 41 |
| 2.1. Construcción de Vecindades | 49 |

Capítulo 1

Preliminares de Geometría

1.1. Curvas Regulares y Longitud de arco

Definición 1.1. Una curva diferenciable parametrizada es una función diferenciable $\alpha : I \rightarrow \mathbb{R}^3$, donde $I = (a, b) \subset \mathbb{R}$.

Ejemplo: (ver Figura 1.1) $\alpha(t) = (a \cos t, a \sin t, bt)$ es una curva diferenciable parametrizada.

Ejemplo: (ver Figura 1.2) $\alpha : \mathbb{R} \rightarrow \mathbb{R}^2$, dada por $\alpha(t) = (t^3 - 4t, t^2 - 4)$ con $t \in \mathbb{R}$ es una curva diferenciable parametrizada.

Definición 1.2. Sea $\alpha : I \rightarrow \mathbb{R}^n$ una curva diferenciable parametrizada, llamaremos punto singular de α a todo punto $t \in I$, donde $\alpha'(t) = 0_n$.

Definición 1.3. Sea $\alpha : I \rightarrow \mathbb{R}^n$ una curva diferenciable parametrizada, diremos que α es una curva regular si y sólo si $\alpha'(t) \neq 0_n$, para cualquier $t \in I$; es decir que α no tiene puntos singulares. $C = \alpha(I)$ admite recta tangente en todo punto $\alpha(t)$, $t \in I$.

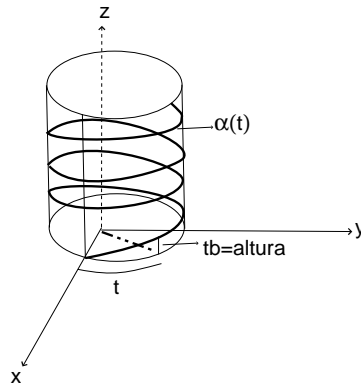


Figura 1.1: Hélice

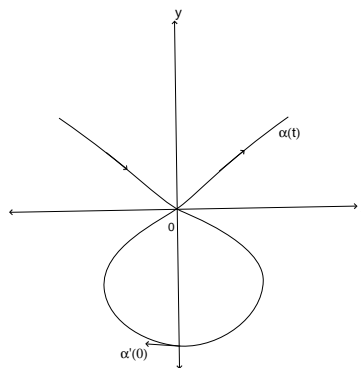


Figura 1.2: El Pez

Observación 1.1. Consideremos una curva diferenciable parametrizada $\alpha : I \rightarrow \mathbb{R}^n$ y observemos que $\alpha'(t)$ la podemos ver como la trayectoria de una partícula con rapidez $S(t)$, dada por $S(t) = \|\alpha'(t)\|$. Así, nos interesa saber ¿Qué longitud tiene el recorrido de la imagen, cuando t varía, por ejemplo desde a hasta b , con $a < b$? (ver Figura 1.3)

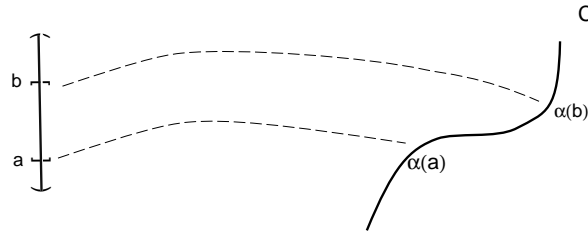


Figura 1.3: Longitud de una curva

Sea $\alpha : I \rightarrow \mathbb{R}^n$, $n = 2, 3$ una curva diferenciable parametrizada y sea $[a, b] \subset I$ un intervalo cerrado.

Sea P cualquier partición del intervalo $[a, b]$, $P: a = t_0 < t_1 < \dots < t_n = b$

$$[a, b] = \bigcup_{k=1}^n [t_{k-1}, t_k].$$

Designemos con $\Pi(P)$ la poligonal (o curva poligonal) cuyos vertices son $\alpha(t_0), \alpha(t_1), \dots, \alpha(t_n)$, respectivamente.

Por lo tanto;

$$\Pi(P) = \bigcup_{i=1}^n \overline{\alpha(t_i)\alpha(t_{i-1})}$$

donde $\overline{\alpha(t_i)\alpha(t_{i-1})}$, denota el segmento de recta que une $\alpha(t_k)$ y $\alpha(t_{k-1})$.
(ver Figura 1.3)

Los lados de la poligonal $\Pi(p)$ tienen longitud $\|\alpha(t_1) - \alpha(t_0)\|$, $\|\alpha(t_2) - \alpha(t_1)\|$, \dots , $\|\alpha(t_n) - \alpha(t_{n-1})\|$.

Así; la longitud $\Pi(p)$ está dada por

$$\ell(\Pi(p)) = \sum_{i=1}^{n-1} \|\alpha(t_i) - \alpha(t_{i-1})\| = \ell(\alpha, P).$$

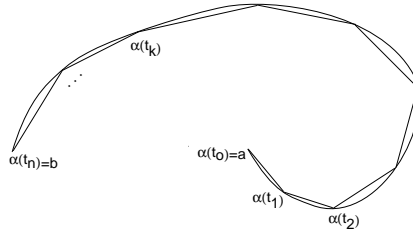


Figura 1.4: Poligonal

Sea $\mathcal{P}([a, b]) = \{ P : P \text{ es una partición de } [a, b] \}$ el conjunto de las particiones de $[a, b]$.

Definición 1.4. Sea $H = \{ \ell(\alpha, P) : P \in \mathcal{P}([a, b]) \}$. Si existe $M > 0$ tal que $\ell(\alpha, P) \leq M$, para cualquier $P \in \mathcal{P}([a, b])$, entonces diremos que la curva $C = \alpha([a, b])$ es rectificable y la longitud de arco α la indicaremos con $\Lambda(a, b)$ y esta dada por $\Lambda(a, b) = \text{Sup}(H)$. Si no existe $M > 0$, entonces diremos que la curva C no es rectificable.

Definición 1.5. Dado $t \in I$, la longitud de arco de una curva regular desde t_0 hasta t , se define

$$\Lambda(t_0, t) = S(t) = \int_{t_0}^t \|\alpha'(t)\| \cdot dt.$$

Ejemplo: (ver Figura 1.3) Sea $\alpha(t) = (\cos t, \sin t)$, $t \in [0, 2\pi)$, ¿Cuál es la longitud de arco?

Solución:

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

Luego,

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

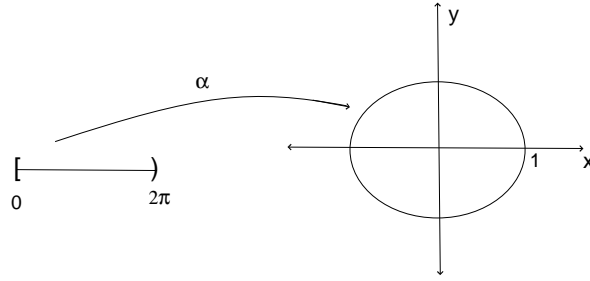


Figura 1.5: Circunferencia

Observemos que la curva β dada por $\beta(t) = (\cos 2t, \sin 2t)$, $t \in [0, 2\pi)$, tiene la misma traza de la curva α (ver Figura 1.3), pero $\beta'(t) = (-2 \sin 2t, 2 \cos 2t)$, de esta manera $\|\beta'\| = \sqrt{4 \sin^2 t + 4 \cos^2 t} = 2$; y se tiene que: $S_\beta(t) = \int_0^{2\pi} 2 \cdot dt = 4\pi = 2S_\alpha(t)$; es decir la longitud de arco de β es el doble de la longitud de arco de α a pesar de tener el mismo gráfico.

Observación 1.2. Sea $\alpha : I \rightarrow \mathbb{R}^u$, $u = 2, 3$ una curva diferenciable parametrizada regular. Si $\|\alpha'(t)\| = 1$, entonces

$$S(t) = \int_{t_0}^t 1 \cdot du = t - t_0.$$

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

Definición 1.6. (Definición de diferenciabilidad uniforme) Una curva $\alpha : I \rightarrow \mathbb{R}^n$ se dice uniformemente diferenciable cuando para cualquier $t \in I$, existe un vector $f(t) \in \mathbb{R}^n$ con la siguiente propiedad: Dado $\epsilon > 0$, se puede encontrar $\delta > 0$ de tal manera que si $0 < |h| < \delta$, $t + h \in I$ entonces $\|\alpha(t + h) - \alpha(t) - \alpha'(t) \cdot h\| < \epsilon \cdot |h|$, para cualquier $t \in I$.

Teorema 1.1. (Teorema fundamental del cálculo para curvas) Sea $f : [a, b] \rightarrow \mathbb{R}^n$ una curva de clase \mathcal{C}^1 , entonces se cumple que:

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

Demostración: [ver [3]]

Teorema 1.2. (Teorema de la diferenciabilidad uniforme) Toda curva $\alpha : [a, b] \rightarrow \mathbb{R}^n$, de clase \mathcal{C}^1 en un intervalo compacto es uniformemente diferenciable.

Demostración: Supongamos que $\alpha : [a, b] \rightarrow \mathbb{R}^n$, de clase \mathcal{C}^1 en $[a, b]$; así, $\alpha' : [a, b] \rightarrow \mathbb{R}^n$ es continua en $[a, b]$ y como $[a, b]$ es compacto, entonces α' es uniformemente continua, de acá se tiene que: Dado $\epsilon > 0$, existe $\delta > 0$ tal que $t, t+h \in [a, b]$ y $|h| < \delta$ implica que $\|\alpha'(t+h) - \alpha'(t)\| < \epsilon$, para cualquier $t \in [a, b]$. Observemos que para $t \in [a, b]$ fijo; vale que:

$$\int_t^{t+h} \alpha'(s) \cdot ds = \alpha'(t)h$$

Estudiamos lo siguiente:

$$\begin{aligned} \|\alpha(t+h) - \alpha(t) - \alpha'(t)h\| &= \left\| \int_t^{t+h} \alpha'(s) ds - \int_t^{t+h} \alpha'(t) ds \right\| \\ &= \left\| \int_t^{t+h} (\alpha'(s) - \alpha'(t)) ds \right\| \\ &\leq \int_t^{t+h} \|\alpha'(s) - \alpha'(t)\| ds \\ &< \epsilon \int_t^{t+h} ds \\ &= \epsilon |h| \\ &\leq \epsilon |h| \end{aligned}$$

De esta manera se tiene que dado $\epsilon > 0$, se pudo encontrar un $\delta > 0$ tal que $0 < |h| < \delta$, $t+h \in I$, entonces se cumple que $\|\alpha(t+h) - \alpha(t) - \alpha'(t)h\| < \epsilon|h|$, para cualquier $t \in I$.

Así se concluye que α es uniformemente diferenciable.

Teorema 1.3. Sea $\alpha : I \rightarrow \mathbb{R}^3$ una curva diferenciable y sea $[a, b] \subset I$, un intervalo cerrado. Para cada partición $P = \{a = t_0 < t_1 < \dots < t_n = b\}$, considérese $\ell(\alpha, P) = \sum \|\alpha(t_i) - \alpha(t_{i-1})\|$, donde P es cualquier partición dada. La norma $\|P\|$ de una partición P es definida como $\|P\| = \max(t_i - t_{i-1})$; $i = 1, 2, \dots, n$. Demostrar que: dado $\epsilon > 0$, existe $\delta > 0$ tal que si $\|P\| < \delta$, se tiene que:

$$\left| \int_a^b \|\alpha'(t)\| dt - \ell(\alpha, P) \right| < \epsilon.$$

Demostración: Sea $P \in \mathcal{P}[a, b]$, $P = \{a = t_0 < t_1 < \dots < t_n = b\}$.

Se define $\sum(P) = \sum \|\alpha'(t_i)\|(t_i - t_{i-1})$.

Afirmación: Dado $\epsilon > 0$, existe $\delta > 0$ tal que $\|P\| < \delta$ implica que $\|\alpha(t_i) - \alpha(t_{i-1})\| < (\alpha'(t_{i-1}) + \frac{\epsilon}{2(b-a)})(t_i - t_{i-1})$

En efecto; Por la continuidad uniforme de α' en $[a, b]$, se tiene que: dado $\epsilon > 0$, existe $\delta_1 > 0$ tal que $|h| < \delta$ y $t+h \in [a, b]$ implica que:

$$\|\alpha'(t+h) - \alpha'(t)\| < \frac{\epsilon}{2(b-a)}, \text{ para cualquier } t \in [a, b]$$

Luego por el teorema de la diferenciabilidad uniforme tenemos que:

$$|h| < \delta \text{ y } t+h \in [a, b] \text{ implica } \|\alpha'(t+h) - \alpha'(t) - \alpha'(t)h\| \leq \frac{\epsilon|h|}{2(b-a)} \quad (*)$$

Tomando el mismo $\delta_1 > 0$ de la continuidad uniforme de α' y $\|P\| < \delta$; esto es que: $\max\{t_j - t_{j-1} : j = 1, 2, \dots, k\} < \delta$

Sustituyendo $t = t_{i-1}$ y $h = t_i - t_{i-1}$ en $(*)$ se tiene:

$0 < |t_i - t_{i-1}| < \max\{t_j - t_{j-1} : j = 1, 2, \dots, k\} < \delta$, implica que

$$\|\alpha(t_i) - \alpha(t_{i-1}) - \alpha'(t_{i-1})(t_i - t_{i-1})\| < \frac{\epsilon |t_i - t_{i-1}|}{2(b-a)}$$

Luego; $\|\alpha(t_i) - \alpha(t_{i-1})\| - \|\alpha'(t_{i-1})(t_i - t_{i-1})\| < \frac{\epsilon(t_i - t_{i-1})}{2(b-a)}$

Así: $\|\alpha(t_i) - \alpha(t_{i-1})\| < (\|\alpha'(t_{i-1})\| + \frac{\epsilon}{2(b-a)})(t_i - t_{i-1})$

Por otro lado estudiemos lo siguiente:

$$\begin{aligned} \ell(\alpha, P) &= \sum_{i=1}^n \|\alpha(t_i) - \alpha(t_{i-1})\| \\ &< \sum_{i=1}^n (\|\alpha'(t_{i-1})\| + \frac{\epsilon}{2(b-a)})(t_i - t_{i-1}) \\ &= \sum_{i=1}^n \|\alpha'(t_{i-1})\|(t_i - t_{i-1}) + \frac{\epsilon}{2(b-a)} \sum_{i=1}^n (t_i - t_{i-1}) \\ &= \sum_{i=1}^n \|\alpha'(t_{i-1})\|(t_i - t_{i-1}) + \frac{\epsilon}{2(b-a)}(t_n - t_0) \\ &= \sum_{i=1}^n \|\alpha'(t_{i-1})\|(t_i - t_{i-1}) + \frac{\epsilon}{2} \end{aligned}$$

De esta manera obtenemos: $\ell(\alpha, P) < \sum_{i=1}^n \|\alpha'(t_{i-1})\|(t_i - t_{i-1}) + \frac{\epsilon}{2}$

Ahora bien; observemos que:

$$\ell(\alpha, P) - \sum(P) < \frac{\epsilon}{2} \quad \text{y} \quad \sum(P) - \ell(\alpha, P) > \frac{-\epsilon}{2}$$

Así tenemos que dado $\epsilon > 0$, existe $\delta_1 > 0$ tal que $\|P\| < \delta_1$, entonces $\|\ell(\alpha, P) - \sum(P)\| < \frac{\epsilon}{2}$

Por otro lado, como α es diferenciable en $[a, b]$ (C^∞), entonces $\alpha'(\cdot)$ es continua y como también se sabe que $\|\cdot\|$ es continua, y la compuesta de continuas es continua, entonces $\|\alpha'(t)\|$ es continua en $[a, b]$; así $\int \|\alpha'(t)\| dt$ existe.

Ahora bien; dado el mismo $\epsilon > 0$ de la continuidad uniforme y escogiendo $\xi_i = t_{i-1} \in [t_{i-1}, t_i]$ se tiene que:

$$\int_a^b \|\alpha'(t)\| dt = \lim_{\|P\| \rightarrow 0} \sum_{i=1}^n \|\alpha'(\xi_i)\| (t_i - t_{i-1}) = \lim_{\|P\| \rightarrow 0} \sum(P)$$

Esto es equivalente a decir que:

Dado $\epsilon > 0$, existe $\delta_2 > 0$ tal que $0 < \|P\| < \delta_2$, implica que $\|\int_a^b \|\alpha'(t)\| dt - \sum(P)\| < \frac{\epsilon}{2}$

Tomando $\delta = \min\{\delta_1, \delta_2\}$, se tiene que $\|P\| < \delta$

De esta manera:

$$\begin{aligned} \left| \int_a^b \|\alpha'(t)\| dt - \ell(\alpha, P) \right| &= \left| \int_a^b \|\alpha'(t)\| dt - \sum(P) + \sum(P) - \ell(\alpha, P) \right| \\ &\leq \left| \int_a^b \|\alpha'(t)\| dt - \sum(P) \right| + \left| \sum(P) - \ell(\alpha, P) \right| \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \end{aligned}$$

Así, hemos probado que dado $\epsilon > 0$, podemos encontrar un $\delta > 0$ tal que si $\|P\| < \delta$, entonces

$$\left| \int_a^b \|\alpha'(t)\| dt - \ell(\alpha, P) \right| < \epsilon$$

Definición 1.7. Sea $\alpha : I \rightarrow \mathbb{R}^3$ una curva parametrizada por longitud de arco $s \in I$, se le llama curvatura de α en s al número real $\|\alpha''(s)\| = K(s)$.

Observación 1.3. $\alpha : I \rightarrow \mathbb{R}^3$ es una línea recta si y sólo si $K(s) = 0$

En efecto:

Si α es una línea recta, $\alpha(s) = us + v$, donde u, v son vectores constantes, entonces $K = \|\alpha''(s)\| = 0$.

Si $K = \|\alpha''(s)\| = 0$, entonces por integración $\alpha(s) = us + v$, y la curva es una línea recta.

Observación 1.4. En los puntos donde $K(s) \neq 0$, esto es $\alpha''(s) \neq 0$, se define $n(s) = \frac{\alpha''(s)}{\|\alpha''(s)\|}$, el cual es un vector unitario y además es paralelo a $\alpha''(s)$ y con el mismo sentido. Por otro lado, como α es parametrizada por longitud de arco s , se tiene que para cualquier $s \in I$, $\|\alpha'(s)\| = 1$.

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

Así concluimos que $\alpha''(s)$ es perpendicular a $\alpha'(s)$ y como $\alpha''(s) = K(s)n(s)$, obtenemos también que $n(s)$ es perpendicular a $\alpha'(s)$, para cualquier $s \in I$.

Como $n(s)$, $\alpha'(s)$ son linealmente independiente (por se ortogonales), ellos determinan un plano $\Pi_s = \{x \in \mathbb{R}^3 : x = \alpha(s) + \lambda \cdot \alpha'(s) + \beta \cdot \alpha''(s), \lambda, \beta \in \mathbb{R}\}$, el cual se llamará plano osculador en s .

Definición 1.8. Dada $\alpha : I \rightarrow \mathbb{R}^3$ una curva diferenciable parametrizada por longitud de arco s . Dado $s \in I$ diremos que s es un punto singular de orden uno, si y sólo si $\alpha''(s) = 0$

En lo que sigue consideraremos curvas diferenciables parametrizadas por longitud de arco que no tengan puntos singulares de orden uno; esto es que $\alpha''(s) \neq 0$, para cualquier $t \in I$.

Se denotará a $\tau(s) = \alpha'(s)$, el vector tangente unitario; así $\tau'(s) = \alpha''(s) = K(s)n(s)$.

Ahora bien, para cualquier $s \in I$ definimos $b(s) = \tau(s) \times n(s)$, el cual es normal al plano osculador y es llamado **vector binormal de s** , además $\|b(s)\| = 1$. Por lo tanto, $b(s)$ es un vector unitario y normal al plano osculador el cual esta determinado por $\tau(s)$ y $n(s)$.

Para cualquier $s \in I$, $b(s) = \tau(s) \times n(s)$ es una función diferenciable en I .

Por lo tanto;

$$\begin{aligned} b'(s) &= \frac{d}{dt}(\tau(s) \times n(s)) \\ &= \tau'(s) \times n(s) + \tau(s) \times n'(s) \\ &= K(s)n(s) \times n(s) + \tau(s) \times n'(s) \\ &= K(s)(n(s) \times n(s)) + \tau(s) \times n'(s) \\ &= \tau(s) \times n'(s) \end{aligned}$$

De acá se sigue que $b'(s)$ es perpendicular a $\tau(s)$ y que $b'(s)$ es paralelo a $n(s)$, por tanto; existe $\tau(s) \in \mathbb{R}$ tal que $b'(s) = \tau(s)n(s)$.

Definición 1.9. Sea $\alpha : I \rightarrow \mathbb{R}^3$ una curva diferenciable parametrizada por longitud de arco s , tal que $\alpha''(s) \neq 0$ para cada $s \in I$. Al número $\tau(s)$ que cumple que $b'(s) = \tau(s)n(s)$ es llamado la torsión de α en s .

En resumen; a cada vector $s \in I$ le asociamos tres vectores unitarios $\tau(s)$, $n(s)$, $b(s)$ el cual recibe el nombre del triedro de Frenet en s , (ver Figura 1.6) además $\mathcal{B} = \{\tau(s), n(s), b(s)\}$ determina una base de \mathbb{R}^3 .

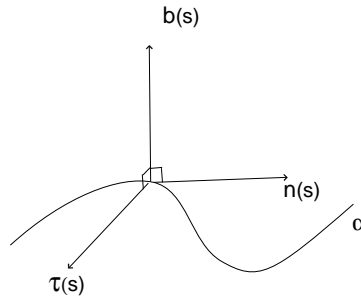


Figura 1.6: Triedro de Frenet

1.2. La diferencial

Definición 1.10. Sea $F : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ una función diferenciable, para cualquier $p \in U$ asociamos una aplicación lineal $dF_p : \mathbb{R}^n \rightarrow \mathbb{R}^m$ la cual llamaremos diferencial de F en p , y es definida como sigue: Sea $w \in \mathbb{R}^n$ y sea $\alpha : (-\epsilon, \epsilon) \rightarrow U$ una curva diferenciable tal que $\alpha(0) = p$ y $\alpha'(0) = w$; por la regla de la cadena se tiene que la curva $\beta = F \circ \alpha : (-\epsilon, \epsilon) \rightarrow \mathbb{R}^m$ es también diferenciable. (ver Figura 1.7)

$$\begin{aligned}
 dF_p(w) &= \beta'(0) \\
 &= (F \circ \alpha)'(0) \\
 &= [F'(\alpha(0))][\alpha'(0)] \\
 &= [F'(p)][w]
 \end{aligned}$$

Proposición 1.1. La definición de dF_p no depende de la elección de la curva la cual pasa por p con vector tangente w , además dF_p es en realidad un mapeo lineal.

Demostración:

Nota: Trabajaremos con el caso de $F : U \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$

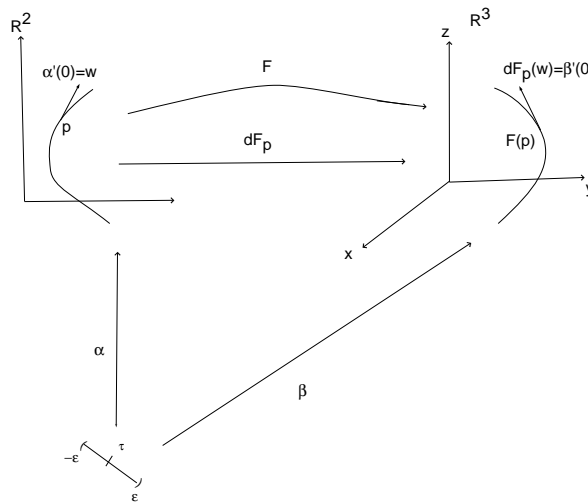


Figura 1.7: La diferencial

Sea $(u, v) \in \mathbb{R}^2$ y $(x, y, z) \in \mathbb{R}^3$.

Sea $e_1 = (1, 0)$, $e_2 = (0, 1)$ la base canónica de \mathbb{R}^2 y $f_1 = (1, 0, 0)$, $f_2 = (0, 1, 0)$, $f_3 = (0, 0, 1)$ la base canónica de \mathbb{R}^3 .

Sea $\alpha(t) = (u(t), v(t))$; $t \in (-\epsilon, \epsilon)$ y $F(u, v) = (X(u, v), Y(u, v), Z(u, v))$

Luego:

$$\begin{aligned} \alpha'(0) &= (u'(0), v'(0)) \\ &= u'(0)e_1 + v'(0)e_2 \\ &= w \end{aligned}$$

$$\begin{aligned} \beta(t) &= (F \circ \alpha)(t) \\ &= F(u(t), v(t)) \\ &= \left(X(u(t), v(t)), Y(u(t), v(t)), Z(u(t), v(t)) \right) \end{aligned}$$

Así, usando la regla de la cadena y tomando la derivada en $t = 0$, obtenemos:

$$\beta'(0) = \left(\frac{\partial X}{\partial u} \frac{du}{dt} + \frac{\partial X}{\partial v} \frac{dv}{dt}, \frac{\partial Y}{\partial u} \frac{du}{dt} + \frac{\partial Y}{\partial v} \frac{dv}{dt}, \frac{\partial Z}{\partial u} \frac{du}{dt} + \frac{\partial Z}{\partial v} \frac{dv}{dt} \right).$$

Luego, podemos escribir a $\beta'(0)$ como sigue:

$$\beta'(0) = \left(\frac{\partial X}{\partial u} \frac{du}{dt} + \frac{\partial X}{\partial v} \frac{dv}{dt} \right) f_1 + \left(\frac{\partial Y}{\partial u} \frac{du}{dt} + \frac{\partial Y}{\partial v} \frac{dv}{dt} \right) f_2 + \left(\frac{\partial Z}{\partial u} \frac{du}{dt} + \frac{\partial Z}{\partial v} \frac{dv}{dt} \right) f_3$$

$$\beta'(0) = \begin{pmatrix} \frac{\partial X}{\partial u} & \frac{\partial X}{\partial v} \\ \frac{\partial Y}{\partial u} & \frac{\partial Y}{\partial v} \\ \frac{\partial Z}{\partial u} & \frac{\partial Z}{\partial v} \end{pmatrix} \cdot \begin{pmatrix} \frac{du}{dt} \\ \frac{dv}{dt} \end{pmatrix} = dF_p(w)$$

Esto demuestra que dF_p es representada en la base canónica de \mathbb{R}^2 en \mathbb{R}^3 por una matriz, la cual depende sólo de las derivadas parciales de p sobre las funciones componentes de X, Y, Z sobre F . Así, dF_p es un mapeo lineal y claramente $dF_p(w)$ no depende de la elección de la curva α .

Teorema 1.4. (Teorema de la Función Inversa) Si $F : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ es una función diferenciable ($\mathcal{C}^1, \mathcal{C}^\infty$) en $p \in U$, U abierto de \mathbb{R}^n y $F'(p)$ es no singular para algún punto $p \in U$, entonces existen $V_p, W_{F(p)}$ vecindades de p y $F(p)$ respectivamente tales que: $F : V_p \rightarrow W_p$ es biyectiva y su inversa $G : W_p \rightarrow V_p$ es diferenciable ($\mathcal{C}^1, \mathcal{C}^\infty$).

Demostración: [ver [4]]

Teorema 1.5. Sea $U \subset \mathbb{R}^{k+n}$ abierto, $F : U \subset \mathbb{R}^{k+n} \rightarrow \mathbb{R}^n$ de clase \mathcal{C}^r en U . Escribir $F(x, y) = F(x_1, \dots, x_k)$, para $y \in \mathbb{R}^n$, $x \in \mathbb{R}^k$, suponga que $(a, b) \in U$ es tal que $F(a, b) = 0$ y $\frac{\partial F}{\partial y}(a, b)$ es no singular, entonces existe B vecindad de $a \in \mathbb{R}^k$ y una única función $G : B \subset \mathbb{R}^k \rightarrow \mathbb{R}^n$ de clase \mathcal{C}^r tal que $G(a) = b$ y $F(x, G(x)) = 0$, para cualquier $x \in B$. Además G es de clase \mathcal{C}^r en B .

Demostración: Definir $f : U \subset \mathbb{R}^{k+n} \rightarrow \mathbb{R}^{k+n}$ por $f(x, y) = (x, F(x, y))$. Ver que f cumple las hipótesis del teorema de la función inversa.

f es de clase \mathcal{C}^r en U , pues F es de clase \mathcal{C}^r en U , además

$$f'(a, b) = \begin{bmatrix} I_{k \times k} & 0_{k \times k} \\ \frac{\partial F}{\partial x}(a, b) & \frac{\partial F}{\partial y}(a, b) \end{bmatrix} \text{ y como } \det \frac{\partial F}{\partial y}(a, b) \neq 0 \text{ por hipótesis, se}$$

tiene que $\det f'(a, b) \neq 0$, luego usando el teorema de la función inversa se tiene

que existen vecindades $C((a, b), r)$ y $W_{f(a, b)}$ tal que la función

$g = f^{-1} : C((a, b), r) \rightarrow W_{f(a, b)}$ es invertible.

Observemos que $f(a, b) = (a, F(a, b)) = (a, 0)$, (pues $F(a, b) = 0$ para todo $a, b \in U$), por otro lado, como $W_{(a, 0)}$ es un abierto que contiene a $(a, 0)$, entonces existe $C(a, 0) \subset W_{(a, 0)}$; llamemos $B_a = C(a, 0) \cap \mathbb{R}^k \subset \mathbb{R}^k$.

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

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

Además;

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

Por lo tanto $G(a) = h(a, 0) = b$ y $F(x, G(x)) = 0$, para cualquier $x \in B$.

Ahora se probará la unicidad de G

Supongamos que $G_1 : B \subset \mathbb{R}^k \rightarrow \mathbb{R}^n$, $G_2 : B \subset \mathbb{R}^k \rightarrow \mathbb{R}^n$, donde $G_1(x) \neq G_2(x)$ para cualquier $x \in B$, $G_1(a) = b = G_2(a)$ y $F(x, G_1(x)) = 0 = F(x, G_2(x))$ para cualquier $x \in B$.

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

También se tiene que:

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

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

Definición 1.11. Sea $f : U \subset \mathbb{R}^m \rightarrow \mathbb{R}^m$, ($m \leq n$) una aplicación diferenciable, diremos que f es una **inmersión** sí y sólo si, para cualquier $p \in U$, $df_p : \mathbb{R}^m \rightarrow \mathbb{R}^m$ es inyectiva; es decir que el rango de $[df_p] = m$. Si $m = n$, df_p resulta isomorfismo y por tanto f es un difeomorfismo local.

Ejemplo: Si $\alpha : I \subset \mathbb{R} \rightarrow \mathbb{R}^n$ es una curva diferenciable regular, entonces es una inmersión.

En efecto: Como α es regular, entonces $\alpha'(t) = (\alpha_1(t), \dots, \alpha_n(t)) \neq (0, \dots, 0)$; es decir que el rango de $[\alpha'] = 1$

1.3. Superficie regular e imagen inversa de un valor regular

Definición 1.12. Un conjunto $S \subset \mathbb{R}^n$ es una superficie de dimensión m y de clase \mathcal{C}^k , cuando para cualquier $p \in S$ existe $U_p \subset \mathbb{R}^n$ tal que $V_p = U_p \cap S$ es la imagen de alguna parametrización $X_p : V_0 \subset \mathbb{R}^m \rightarrow V_p$ de dimensión m y de clase \mathcal{C}^k .

El conjunto V_p es abierto en S y es llamado vecindad parametrizada de p .

Definición 1.13. Sea $S \subset \mathbb{R}^3$, diremos que S es una superficie regular si, para cada $p \in S$, existe una vecindad V en \mathbb{R}^3 y una aplicación $X : U \rightarrow V \cap S$ en un conjunto abierto $U \subset \mathbb{R}^2$ sobre $V \cap S \subset \mathbb{R}^3$ tal que:

1. X es diferenciable.
2. X es un homeomorfismo.
3. (condición de regularidad) Para cada $q \in S$, la diferencial $dX_q : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ es inyectiva.

Proposición 1.2. Sea $U \subset \mathbb{R}^2$ abierto y $f : U \rightarrow \mathbb{R}^3$ una función diferenciable; entonces el gráfico de f que es un subconjunto de \mathbb{R}^3 dado por $(x, y, f(x, y))$ para $(x, y) \in U$, es una superficie regular. (ver Figura 1.8)

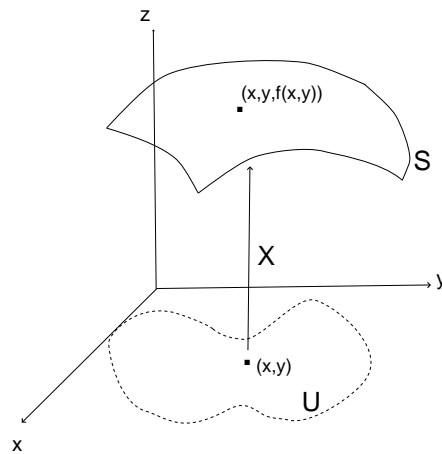


Figura 1.8: Gráfico de una función

Demostración:

Sea $X : U \rightarrow S \cap \mathbb{R}^3$, dada por:

$$X(u, v) = (u, v, f(u, v)), \text{ donde } (u, v) \in U \subset \mathbb{R}^2 \text{ y}$$

$$S = \{(u, v, f(u, v)) : (u, v) \in U\}$$

1. Claramente X es una parametrización del gráfico de f , cuya vecindad $X(U)$ cubre todos los puntos $p \in S$ y además $x(u, v) = u$, $y(u, v) = v$, $z(u, v) = z = f(u, v)$ las cuales son las funciones componentes de X , y admiten derivadas parciales continuas de todos los ordenes en U .

Por lo tanto; X es diferenciable en U .

2. X es inyectiva en U

En efecto:

Sea $(u_1, v_1), (u_2, v_2) \in U$

$$\begin{aligned}
X(u_1, v_1) = X(u_2, v_2) &\Rightarrow (u_1, v_1, f(u_1, v_1)) = (u_2, v_2, f(u_2, v_2)) \\
&\Rightarrow \begin{cases} u_1 = u_2 \\ v_1 = v_2 \end{cases} \\
&\Rightarrow (u_1, v_1) = (u_2, v_2)
\end{aligned}$$

Así concluimos que X es inyectiva, más aun $X : U \rightarrow X(U) \cap S$ es biyectiva, pues para cada $p \in S$; existe un único $(u, v) \in U$; tal que $X(u, v) = (x, y, z)$; pues con $u = x$, $v = y$ se tiene que $X(u, v) = (u, v, f(u, v)) = (x, y, z)$. Por lo tanto $X(U) = S$ y como X es diferenciable en U , entonces X es continua en U ; además existe $X^{-1} : S \rightarrow U$, dada por:

$$X^{-1}(u, v, f(u, v)) = (u, v) = \pi_{xy}(u, v, f(u, v))$$

donde $\pi_{xy} : \mathbb{R}^3 \rightarrow \mathbb{R}^2$, dado por $\pi_{xy}(x, y, f(x, y)) = (x, y)$ es continua en \mathbb{R}^2 .

Así, X^{-1} será continua en $S \subset \mathbb{R}^3$ vista como la restricción de π_{xy} en S .

Por lo tanto X es un homeomorfismo.

3. Probemos la condición de regularidad

Sea $q = (u, v) \in U$ y consideremos la diferencial $dX_q : \mathbb{R}^2 \rightarrow \mathbb{R}^3$, la cual es una aplicación lineal dada por:

$$dX_q = \begin{pmatrix} \frac{\partial u}{\partial u} & \frac{\partial u}{\partial v} \\ \frac{\partial v}{\partial u} & \frac{\partial v}{\partial v} \\ \frac{\partial f(u, v)}{\partial u} & \frac{\partial f(u, v)}{\partial v} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ \frac{\partial f(q)}{\partial u} & \frac{\partial f(q)}{\partial v} \end{pmatrix}$$

Observemos que el determinante jacobiano

$$\frac{\partial(x, y)}{\partial(u, v)} = \det \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1$$

Por lo tanto, dX_q es inyectiva; y como ya hemos probado que se cumplen 1, 2 y 3; concluimos que $S = \{(u, v, f(u, v)) : (u, v) \in U\}$ es una superficie regular.

Definición 1.14. Sea $U \subset \mathbb{R}^n$ abierto y $F : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ una función diferenciable. Diremos que $p \in U$ es un punto crítico de F cuando la diferencial $dF_p : \mathbb{R}^n \rightarrow \mathbb{R}^m$ no sea sobreyectiva, es decir; que el rango de $[dF_p]$ es menor que m .

Ejemplo: Sea $F : \mathbb{R}^n \rightarrow \mathbb{R}$, entonces :

$$\begin{aligned} dF_p(v) &= [DF(p)]_{1 \times n} [v]_{n \times 1} \\ &= \left[\frac{\partial F(p)}{\partial x_1} \quad \frac{\partial F(p)}{\partial x_2} \quad \dots \quad \frac{\partial F(p)}{\partial x_n} \right] \cdot \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \\ &= \langle \text{grad}F(p), v \rangle \end{aligned}$$

Queremos que $DF(p) = \text{grad}F(p)$ tenga rango menor que 1; pero $DF(p) = \text{grad}F(p)$ es una matriz unifila. ¿Cuándo una matriz unifila tiene rango menor que 1?, cuando cada uno de sus elementos son ceros. Por tanto, cuando $m = 1$ pediremos que $DF(p) = \text{grad}F(p) = [0 \dots 0]_{1 \times n}$

Así se concluye que p es un punto crítico de F si el $\text{grad}F(p) = 0$

Observación 1.5. Si p es un punto crítico de F , entonces a $F(p)$ lo llamaremos valor crítico.

Definición 1.15. Sea $F : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$, diremos que $a \in F(U)$ es un valor regular si no es un valor crítico.

Proposición 1.3. Sea $f : U \subset \mathbb{R}^3 \rightarrow \mathbb{R}$ diferenciable y $a \in f(U)$ un valor regular de f , entonces $f^{-1}(a)$ es una superficie regular.

Demostración:

Sea $p = (x, y, z) \in f^{-1}(a) = \{(x, y, z) \in \mathbb{R}^3 / f(x, y, z) = a\}$, donde $a \in f(U)$ es un valor regular. Como $a \in f(U)$ es un valor regular; entonces df_p es sobreyectiva, es decir el $grad(f(p))$ es no nulo.

$$grad(f(p)) = \left(\frac{\partial f(p)}{\partial x}, \frac{\partial f(p)}{\partial y}, \frac{\partial f(p)}{\partial z} \right) \neq 0,$$

Sin perder generalidad, suponga que $\frac{\partial f(p)}{\partial z} \neq 0$ y defina $F : \mathbb{R}_3 \rightarrow \mathbb{R}_3$, dada por: $F(x, y, z) = (x, y, f(x, y, z))$, donde para cada $(x, y, z) \in U$ su imagen $F(x, y, z) = (u, v, t)$.

Es claro que F es diferenciable; pues cada una de sus funciones componentes lo es; luego la diferencial de F en p viene dada por:

$$dF_p = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \frac{\partial F(P)}{\partial x} & \frac{\partial F(P)}{\partial y} & \frac{\partial F(P)}{\partial z} \end{pmatrix}$$

y el $det(dF_p) = \frac{\partial F(P)}{\partial z} \neq 0$, de esta manera podemos usar el teorema de la función inversa, el cual nos garantiza que existe una vecindad V_p de p y una vecindad $W_{F(p)}$ tal que $F : V_p \rightarrow W_{F(p)}$ es invertible y $F^{-1} : W_{F(p)} \rightarrow V_p$ es diferenciable C^∞ .

Por lo tanto, si $(x, y, z) \in V_p$, entonces

$$F(x, y, z) = (x, y, f(x, y, z)) = (u, v, t) \in W_{F(p)} \text{ y}$$

$$F^{-1}(u, v, t) = F^{-1}(x, y, f(x, y, z)) = (x, y, z)$$

Luego; si $F^{-1}(u, v, t) = (h(u, v, t), S(u, v, t), g(u, v, t))$, entonces

$h(u, v, t) = x = u$, $S(u, v, t) = y = v$, $g(u, v, t) = z$, son diferenciables con $(u, v, t) \in W_{F(p)}$ (ver Figura 1.9)

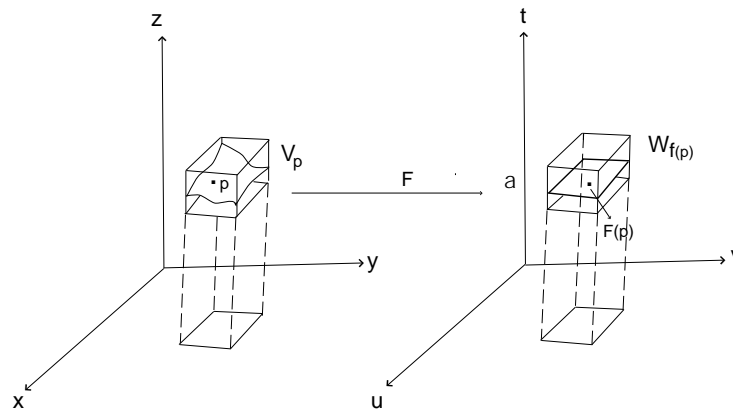


Figura 1.9: Valor Regular

Considérese la función proyección $\Pi : \mathbb{R}^3 \rightarrow \mathbb{R}^2$, dada por

$$\Pi(x, y, z) = (x, y), \text{ la cual es una función continua.}$$

Considérese también $H : \Pi(V_p) \rightarrow V_p$, dada por $H(x, y) = g(u, v, a) = z$ la cual es diferenciable en $\Pi(V_p)$, el cual es el conjunto abierto en \mathbb{R}^2 .

Ahora bien, probemos que: $F(f^{-1}(a) \cap v_p) = W_{f(p)} \cap \{(u, v, t) \in \mathbb{R}^3 / t = a\}$

En efecto:

- $\blacksquare F(f^{-1}(a) \cap V_p) \subset W_{f(p)} \cap \{(u, v, t) \in \mathbb{R}^3\}$

$$\begin{aligned} \xi \in F(f^{-1}(a) \cap V_p) &\Rightarrow F^{-1}(\xi) \in f^{-1}(a) \cap V_p \\ &\Rightarrow F^{-1}(\xi) \in f^{-1}(a) \wedge F^{-1}(\xi) \in V_p \\ &\Rightarrow f(F^{-1}(\xi)) = a \wedge \xi \in F(V_p) = W_{F(p)} \\ &\Rightarrow \xi \in \{(u, v, t) \in \mathbb{R}^3 / t = a\} \wedge \xi \in W_{F(p)} \\ &\Rightarrow \xi \in W_{f(p)} \cap \{(u, v, t) \in \mathbb{R}^3\} \end{aligned}$$

Justifiquemos porque $f(F^{-1}(\xi)) = a \Rightarrow \xi \in \{(u, v, t) \in \mathbb{R}^3\}$

Sea $\xi = (u, v, t)$, entonces

$$F^{-1}(\xi) = F^{-1}(u, v, t) = (u, v, g(u, v, t)) = (x, y, g(u, v, t))$$

$$\text{Luego; } f(F^{-1}(\xi)) = f(x, y, g(u, v, t)) = a$$

$$\text{Así, } F(F^{-1}(\xi)) = F(x, y, g(u, v, t)) = (x, y, f(u, v, g(u, v, t))) = (u, v, a)$$

y por tanto, $\xi \in \{(u, v, t) \in \mathbb{R}^3\}$

De acá, $F(f^{-1}(a) \cap V_p) \subset W_{f(p)} \cap \{(u, v, t) \in \mathbb{R}^3\}$.

- $\blacksquare W_{f(p)} \cap \{(u, v, t) \in \mathbb{R}^3\} \subset F(f^{-1}(a) \cap V_p)$

$$\begin{aligned} \xi \in W_{f(p)} \cap \{(u, v, t) \in \mathbb{R}^3\} &\Rightarrow \xi \in W_{F(p)} \wedge \xi \in \{(u, v, t) \in \mathbb{R}^3\} \\ &\Rightarrow F^{-1}(\xi) \in F^{-1}(W_{f(p)}) = V_p \wedge F^{-1}(\xi) \in f^{-1}(a) \\ &\Rightarrow F^{-1}(\xi) \in V_p \cap f^{-1}(a) \\ &\Rightarrow \xi \in F(V_p \cap f^{-1}(a)) \end{aligned}$$

Por lo tanto; $W_{f(p)} \cap \{(u, v, t) \in \mathbb{R}^3\} \subset F(f^{-1}(a) \cap V_p)$

De acá concluimos que $W_{f(p)} \cap \{(u, v, t) \in \mathbb{R}^3\} = F(f^{-1}(a) \cap V_p)$

Esto significa que el gráfico de H es:

$$\begin{aligned} f^{-1}(a) \cap V_p &= \{(x, y, z) \in \mathbb{R}^3 / z = H(x, y), (x, y) \in \Pi(V_p)\} \\ &= \{(x, y, H(x, y)) \in \mathbb{R}^3 / (x, y) \in \Pi(V_p)\} \\ &= \text{graf}(H) \end{aligned}$$

Luego, como el $\text{graf}(H) = f^{-1}(a) \cap V_p$, entonces por la **proposición 1.2**, $S_p = f^{-1}(a) \cap V_p$ es una superficie regular, más aun para una vecindad coordenada de p ($X : \Pi(V_p) \rightarrow S_p$), dada por $X(x, y) = (x, y, H(x, y))$ cumple las condiciones 2 y 3 de la definición de superficie regular.

Así, todo punto $p \in S = f^{-1}(a)$ puede ser cubierto por una vecindad coordenada. Por lo tanto, $S = f^{-1}(a)$ es una superficie regular.

Definición 1.16. Sea S una superficie regular, $V \subset S$ abierto y $f : V \subset S \rightarrow \mathbb{R}$. Diremos que f es diferenciable en $p \in V$ si y sólo si existe una parametrización $X : U \subset \mathbb{R}^2 \rightarrow S$ con $p \in X(U) \subset V$ y $f \circ X : U \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ diferenciable en $X^{-1}(p)$. (ver Figura 1.10)

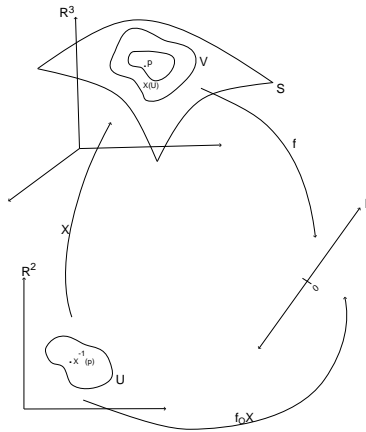


Figura 1.10: Diferenciabilidad de f

Definición 1.17. Diremos que f es diferenciable en V si ésta es diferenciable para cualquier punto en V .

Definición 1.18. Sea S una superficie, $S \subset \mathbb{R}^3$ y $p \in S$. Se llamará vector tangente a S en p al vector $\alpha'(0)$ donde $\alpha : (-\epsilon, \epsilon) \rightarrow S$ es una curva diferenciable con $\alpha(0) = p$.

Definición 1.19. Al conjunto formado por todos los vectores tangentes a S en p , lo llamaremos plano tangente y lo denotaremos como sigue;

$T_p S = \{\alpha'(0) = v : \alpha : (-\epsilon, \epsilon) \rightarrow S, \text{ es una curva diferenciable con } \alpha(0) = p\}$
 (ver Figura 1.11)

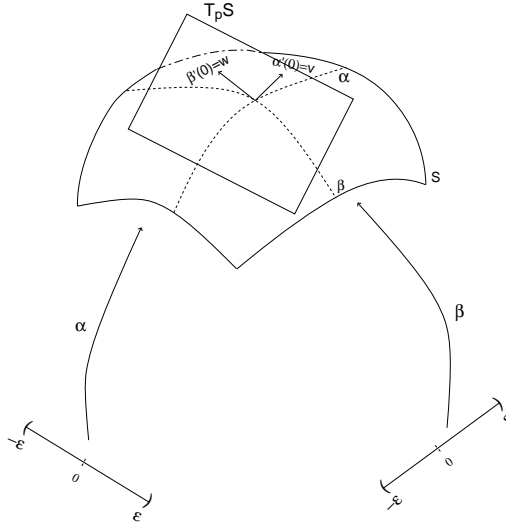


Figura 1.11: Definición de Plano Tangente

Proposición 1.4. Sea $X : U \subset \mathbb{R}^2 \rightarrow S$ una parametrización de la superficie regular S , y sea $q \in U$, entonces se tiene que el subespacio vectorial de dimensión 2; $dX_q(\mathbb{R}^2) = T_p S$.

Demostración:(ver Figura 1.12) Probemos que $dX_q(\mathbb{R}^2) \subset T_p S$.

Sea $w \in X_q(\mathbb{R}^2)$, entonces existe $v \in \mathbb{R}^2$ tal que $dX_q(v) = w$.

Ahora bien, consideremos $r : (-\epsilon, \epsilon) \rightarrow U$ dada por $r(t) = q + tv$, y observemos lo siguiente:

$r(0) = q$, $\alpha = X \circ r$. De esta manera tenemos:

- $\alpha(0) = (X \circ r)(0) = X(r(0)) = X(q) = p$

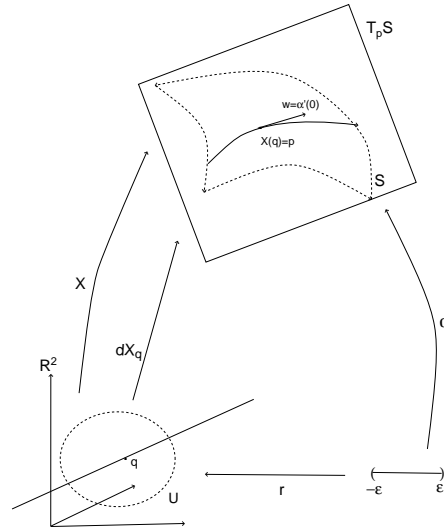


Figura 1.12: Plano Tangente

$$\blacksquare \alpha'(0) = (X \circ r)'(0) = X'(r(0))(r'(0)) = X'(q)v = dX_q(v) = w$$

Así, tenemos que $w \in T_p S$, y por tanto $dX_q(\mathbb{R}^2) \subset T_p S$.

Probemos ahora que $T_p S \subset dX_q(\mathbb{R}^2)$.

Sea $w \in T_p S$, entonces $w = \alpha'(0)$ para alguna curva $\alpha : (-\epsilon, \epsilon) \rightarrow S$ diferenciable con $\alpha(0) = X(q)$.

Observemos que $\beta = X^{-1} \circ \alpha : (-\epsilon, \epsilon) \rightarrow U$ es diferenciable y que $\alpha = X \circ \beta$; luego

$$\begin{aligned} \alpha'(0) &= (X \circ \beta)'(0) \\ &= X'(\beta(0))\beta'(0) \\ &= X'(q)\beta'(0) \\ &= dX_q(\beta'(0)) \\ &= w \end{aligned}$$

De acá $w \in dX_q(\mathbb{R}^2)$ y por tanto $T_pS \subset dX_q(\mathbb{R}^2)$.

Así, concluimos que $T_pS = dX_q(\mathbb{R}^2)$.

Definición 1.20. Sea \mathbb{V} un espacio vectorial, $\dim(\mathbb{V}) = 2$, diremos que la aplicación lineal $A : \mathbb{V} \rightarrow \mathbb{V}$ es autoadjunta si y sólo si:

$$\langle Av, w \rangle = \langle v, Aw \rangle, \text{ para cualquier } v, w \in \mathbb{V}$$

Ejemplo: Sea $A = a_{ij}$, $i, j = 1, 2$ y $\{e_1, e_2\}$ la base ortonormal de \mathbb{V} , entonces:

$$\begin{aligned} \blacksquare \langle Ae_1, e_2 \rangle &= \left\langle \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\rangle = a_{21} \\ \blacksquare \langle e_1, Ae_2 \rangle &= \left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\rangle = a_{12} \end{aligned}$$

Ahora bien, por la simetría del producto interno tenemos que:

$$\langle e_1, Ae_2 \rangle = \langle Ae_2, e_1 \rangle \text{ y } \langle e_2, Ae_1 \rangle = \langle Ae_1, e_2 \rangle.$$

Luego; si A es autoadjunta, entonces $a_{21} = \langle Ae_1, e_2 \rangle = \langle e_1, Ae_2 \rangle = a_{12}$.

De acá tenemos que si A es autoadjunta, entonces A es simétrica, esto es $a_{ij} = a_{ji}$.

Por otro lado, definiendo $B : \mathbb{V} \times \mathbb{V} \rightarrow \mathbb{R}$ por, $B(v, w) = \langle Av, w \rangle$, se tiene una aplicación bilineal y simétrica; es decir que se cumple:

1. $B(\alpha v, w) = \alpha B(v, w)$
2. $B(v_1 + v_2, w) = B(v_1, w) + B(v_2, w)$

3. $B(v, \alpha w) = \alpha B(v, w)$
4. $B(v, w_1 + w_2) = B(v, w_1) + B(v, w_2)$
5. $B(v, w) = B(w, v)$

Recíprocamente se tiene que, si $B : \mathbb{V} \times \mathbb{V} \rightarrow \mathbb{R}$ es bilineal simétrica, entonces $A : \mathbb{V} \rightarrow \mathbb{V}$, definida por $B(v, w) = (Av, w)$ es autoadjunta, más aún si $B : \mathbb{V} \times \mathbb{V} \rightarrow \mathbb{R}$ es bilineal simétrica y se define $Q(v) = B(v, v)$, $v \in \mathbb{V}$ se tiene una forma cuadrática. Además si se conoce la forma cuadrática Q que proviene de $B \frac{1}{2}[Q(u+v) - Q(u) - Q(v)]$ podemos obtener la forma bilineal simétrica de la siguiente manera:

$$\begin{aligned} \frac{1}{2}[Q(u+v) - Q(u) - Q(v)] &= \frac{1}{2}[B(u+v, u+v) - B(u, u) - B(v, v)] \\ &= \frac{1}{2}[B(u, u) + 2B(u, v) + B(v, v) - B(u, u) - B(v, v)] \\ &= B(u, v) \end{aligned}$$

Así, observamos que se puede establecer una correspondencia inyectiva entre la forma cuadrática en \mathbb{V} y la aplicación lineal autoadjunta de \mathbb{V} .

1.4. La primera forma fundamental

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

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

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

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

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

A ésta le llamaremos primera forma fundamental de la superficie regular $S \subset \mathbb{R}^3$ en el punto $p \in S$. (ver Figura 1.13)

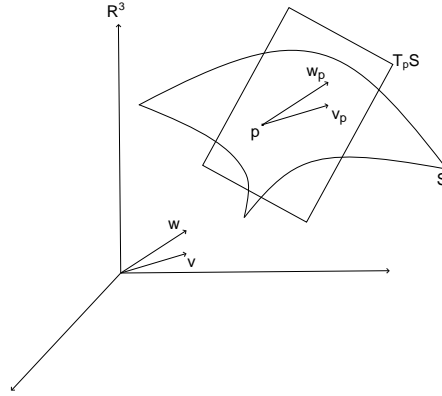


Figura 1.13: Primera Forma Fundamental

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

$$\text{Así, } p = \alpha(0) = X(u(0), v(0)) = X(u_0, v_0)$$

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

Donde,

$$g_{11} = E = \langle X_u(u_0, v_0), X_u(u_0, v_0) \rangle_p, \quad g_{12} = g_{21} = F = \langle X_u(u_0, v_0), X_v(u_0, v_0) \rangle_p$$

y $g_{22} = G = \langle X_v(u_0, v_0), X_v(u_0, v_0) \rangle_p$.

Más adelante la métrica Riemanniana será dada así:

$$\mathbf{G} = \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix}$$

Obsérvese que si $p = X(u, v)$ varia en una vecindad coordenada, entonces $E(u, v)$, $F(u, v)$, $G(u, v)$ serán funciones diferenciables en dicha vecindad.

La importancia de la primera forma fundamental proviene del hecho de que al conocer I , podemos tratar cuestiones métricas sobre una superficie regular S , sin hacer referencia al espacio ambiente \mathbb{R}^3 .

Como ya sabemos la longitud de arco parametrizada por la curva $\alpha : I \rightarrow S$ es dada por:

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

Pero si X es una parametrización local de S , entonces $\alpha(t) = X(u(t), v(t))$ e $I_p(\alpha'(t)) = E(u'(t))^2 + 2Fu'(t)v'(t) + G(v'(t))^2$; de esta manera se obtiene:

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

Luego, derivando y elevando al cuadrado se tiene que:

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

y así: $dS^2 = E(du)^2 + 2Fdu.dv + G(dv)^2$

Esta es la expresión del teorema de pitágoras para un triángulo sobre la superficie S .

Definición 1.21. Si $X : U \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$ es una parametrización, se define $E = \langle X_u, X_u \rangle$, $F = \langle X_u, X_v \rangle$ y $G = \langle X_v, X_v \rangle$ y se le llama métrica Riemanniana a la primera forma fundamental $I_p : T_p X(U) \rightarrow \mathbb{R}$ dada por:

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

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

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

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

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

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

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

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

$$g_{ij}(p) = \left\langle \frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j} \right\rangle_p = \langle X_u, X_v \rangle_p$$

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

Definición 1.22. Sea $S \subset \mathbb{R}^3$ una superficie y $X : U \subset \mathbb{R}^2 \rightarrow S$ una carta local, entonces $g_{ij} : U \rightarrow \mathbb{R}$ son funciones a valores reales diferenciables en U para cualquier $i, j = 1, 2$; pues es claro que $g_{11} = E = \langle X_u, X_u \rangle = \|X_u\|^2 > 0$, $g_{11} = G = \langle X_v, X_v \rangle = \|X_v\|^2 > 0$ y para el caso donde $i \neq j$ será $g_{12} = g_{21} = F = \langle X_u, X_v \rangle = \langle X_v, X_u \rangle = \|X_u\| \|X_v\| \cdot \cos(\theta) = \sqrt{E \cdot F} \cdot \cos(\theta)$.

Así se observa que a través de $F = g_{ij}$ ($i \neq j$) se puede encontrar el ángulo que forma X_u con X_v , por otro lado si $v = v_1X_u + v_2X_v$ y $w = w_1X_u + w_2X_v$, entonces:

$$\begin{aligned}\langle v, w \rangle &= \langle v_1X_u + v_2X_v, w_1X_u + w_2X_v \rangle \\ &= v_1w_1E + (v_1w_2 + v_2w_1)F + v_2w_2G\end{aligned}$$

Es decir g_{ij} determina completamente el producto escalar de dos vectores.

Considérese ahora el producto vectorial de los vectores básicos.

$$X_u \times X_v = \det \begin{pmatrix} i & j & k \\ \frac{\partial x_1}{\partial u} & \frac{\partial x_2}{\partial u} & \frac{\partial x_3}{\partial u} \\ \frac{\partial x_1}{\partial v} & \frac{\partial x_2}{\partial v} & \frac{\partial x_3}{\partial v} \end{pmatrix}$$

Y obsérvese que $\|X_u \times X_v\| = \|X_u\| \cdot \|X_v\| \cdot \sin(\theta)$ mide el área del paralelogramo y que dicha área es igual al área del rectángulo coloreado en azul (ver Figura 1.14).

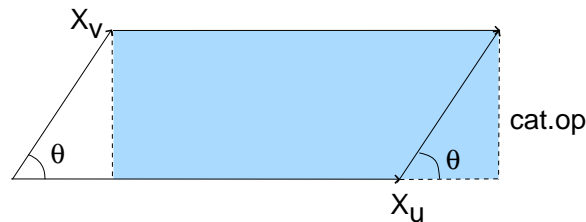


Figura 1.14: Área de un paralelogramo

$\sin(\theta) = \frac{\text{cat.op}}{h} = \frac{\text{cat.op}}{\|X_v\|}$, esto implica que $\text{cat.op} = \sin(\theta)\|X_v\|$; por otro lado denotemos por A_{\square} el área del rectángulo; así

$$A_{\square} = \|X_u\|\|X_v\|\sin(\theta) = \|X_u \times X_v\| = \text{área del paralelogramo}$$

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

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

de acá se obtiene que $\|X_u \times X_v\| = \sqrt{EG - F^2}$.

Definición 1.23. *Un atlas de una superficie S , es un conjunto*

$$A = \{\varphi : V_0 \rightarrow V / \varphi \text{ es parametrización de } V \text{ y la unión de tales } V \text{ cubren } S\}.$$

Definición 1.24. *Dos parametrizaciones $\varphi : V_0 \rightarrow V$, $\psi : W_0 \rightarrow W$ se dicen compatibles cuando $V \cap W \neq \emptyset$ y $\psi^{-1} \circ \varphi : \varphi^{-1}(V \cap W) \rightarrow \psi^{-1}(V \cap W)$ tiene $\det J(\psi^{-1} \circ \varphi)(x) > 0$, para cualquier $x \in \varphi^{-1}(V \cap W)$.*

Definición 1.25. *Un atlas se dice coherente cuando todo par de parametrizaciones son compatibles.*

Definición 1.26. *Una superficie se dice orientable cuando admite un atlas coherente.*

Definición 1.27. *Dado una parametrización $X : U \subset \mathbb{R}^2 \rightarrow S$ sobre una superficie regular S en el punto $p \in S$. Nosotros podemos escoger un vector normal unitario en cada punto de $X(U)$, dado por la siguiente regla.*

$$N(q) = \frac{X_u \times X_v}{\|X_u \times X_v\|}(q), \quad q \in X(U).$$

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

Definición 1.28. *Sea $S \subset \mathbb{R}^3$ una superficie con orientación N ; la aplicación $N : S \rightarrow \mathbb{R}^3$ tiene sus valores en la esfera unitaria*

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

$N : S \rightarrow S^2 \subset \mathbb{R}^3$ se le llama función de Gauss de S , la cual es diferenciable y $dN_p : T_p S \rightarrow T_p S$ es lineal. Esta opera de la siguiente forma:

Dado $\alpha(t) = X(u(t), v(t))$ una curva parametrizada en S con $\alpha(0) = p$, y consideremos $N \circ \alpha(t) = N(t)$ en S^2 ; luego

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

Por otro lado se tiene que $\langle N, X_u \rangle = 0$ y $\langle N, X_v \rangle = 0$; así al derivar la primera respecto v y la segunda respecto u , se obtiene lo siguiente $\langle N_v, X_u \rangle = -\langle N, X_{uv} \rangle$ y $\langle X_v, N_u \rangle = -\langle X_{vu}, N \rangle$, pero X es \mathcal{C}^2 , entonces $X_{uv} = X_{vu}$.

Por lo tanto, $\langle N_v, X_u \rangle = \langle X_v, N_u \rangle$ o lo que es equivalente,
 $\langle dN_p(X_v), X_u \rangle = \langle X_v, dN_p(X_u) \rangle$

De acá se concluye que $dN_p : T_p S \rightarrow T_{N(p)}S^2$ es autoadjunta, y así se le puede asociar una forma cuadrática $II_p : T_p S \rightarrow \mathbb{R}$ dada por $II_p(v) = \langle dN_p(X_v), v \rangle$, la cual es llamada segunda forma fundamental.

Como $T_p S$ y $T_{N(p)}S$ son paralelos podemos escribir $dN_p : T_p S \rightarrow T_p S$ en lugar de $dN_p : T_p S \rightarrow T_{N(p)}S^2$; así como $\{X_u, X_v\}$ es base de $T_p S$ y $N_u, N_v \in T_p S$ se tiene que:

$$N_u = a_{11}X_u + a_{21}X_v$$

$$N_v = a_{12}X_u + a_{22}X_v$$

Y de esta manera

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

Por lo tanto,

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

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

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

Donde $-\langle N_u, X_u \rangle = e$, $\langle N_u, X_v \rangle = \langle N_v, X_u \rangle = f$ y $\langle N_v, X_v \rangle = g$

Definición 1.29. Sea \mathbf{C} una curva regular en S que pasa por el punto $p \in S$ con curvatura K en p ; y el $\cos(\theta) = \langle n, N \rangle$, donde n es el vector normal a \mathbf{C} y N es el normal a S en p . El número $K_n = K \cos(\theta)$ es llamado curvatura normal de $\mathbf{C} \subset S$ en p .

Interpretaremos la segunda forma fundamental, Considérese una curva $\mathbf{C} \subset S$ parametrizada por $\alpha(s)$, donde s es la longitud en \mathbf{C} y $\alpha'(s) = p$.

Se denotará por $N(s)$ la restricción del vector normal N a lo largo de la curva $\alpha(s)$.

Como ya se sabe, $\langle N(s), \alpha'(s) \rangle = 0$, si deriva se obtiene que

$\langle N(s), \alpha''(s) \rangle = -\langle N'(s), \alpha'(s) \rangle$, por lo tanto.

$$\begin{aligned} II_p(\alpha'(0)) &= -\langle dN_p(\alpha'(0)), \alpha'(0) \rangle \\ &= -\langle N'(0), \alpha'(0) \rangle \\ &= -\langle N(0), \alpha''(0) \rangle \\ &= -\langle N, K_n \rangle \\ &= K_n(p) \end{aligned}$$

Definición 1.30. *Un campo en un conjunto abierto $U \subset \mathbb{R}^2$ es una aplicación que asigna a cada $q \in U$ un vector $w(q) \in \mathbb{R}^2$. se dirá que \mathbf{W} es un campo vectorial diferenciable si para cada $q = (x, y) \in U$ se puede escribir $W(q) = (a(x, y), b(x, y))$ y las funciones a y b son diferenciables en U .*

Definición 1.31. *Sea $X : U \subset \mathbb{R}^2 \rightarrow S$ una parametrización en la orientación de S . Es posible asignar a cada punto de $X(U)$ un triedro natural dado por los vectores X_u , X_v y N .*

Expresaremos la derivada de los vectores X_u , X_v y N en la base $\{X_u, X_v, N\}$.

$$\begin{aligned} X_{uu} &= \Gamma_{11}^1 X_u + \Gamma_{11}^2 X_v + L_1 N \\ X_{uv} &= \Gamma_{12}^1 X_u + \Gamma_{12}^2 X_v + L_2 N \\ X_{vu} &= \Gamma_{21}^1 X_u + \Gamma_{21}^2 X_v + \bar{L}_2 N \\ X_{vv} &= \Gamma_{22}^1 X_u + \Gamma_{22}^2 X_v + L_3 N \\ N_u &= a_{11} X_u + a_{21} X_v \\ N_v &= a_{12} X_u + a_{22} X_v \end{aligned}$$

Los coeficientes Γ_{ij}^k se les llaman símbolos de Cristoffel de S en la parametrización X . Como $X_{uv} = X_{vu}$, se concluye que $\Gamma_{12}^1 = \Gamma_{21}^1$ y $\Gamma_{12}^2 = \Gamma_{21}^2$; por otro lado, si se realiza el producto interno entre N y X_{uu} se obtiene $\langle N, X_{uu} \rangle = L_1 \langle N, N \rangle$.

$$\text{Así; } L_1 = \frac{\langle N, X_{uu} \rangle}{\langle N, N \rangle} = e$$

Haciendo el producto interno con las primeras cuatro ecuaciones anteriores y N , se obtiene que $L_1 = e$, $L_2 = \bar{L}_2 = f$ y $L_3 = G$, donde e , f y g son los coeficientes de la segunda forma fundamental.

1.5. Transporte Paralelo y Geodésicas

Definición 1.32. *Un campo vectorial (tangente) en un abierto $U \subset S$, donde S es una superficie regular; es una correspondencia \mathbf{W} que asigna a cada punto $p \in U$ un vector $\mathbf{W}(p) \in T_p S$.*

Se dirá que $\mathbf{W} : U \rightarrow TU$ es un campo diferenciable en p si y sólo si existe $X(u, v)$ parametrización en p tal que si $\mathbf{W}(p) = a(p)X_u + b(p)X_v$ en la base $\{X_u, X_v\}$, entonces las aplicaciones a y b son diferenciables.

\mathbf{W} es diferenciable en U si y sólo si \mathbf{W} es diferenciable para cualquier $p \in U$.

Definición 1.33. *Sea \mathbf{W} un campo vectorial diferenciable sobre el conjunto abierto $U \subset S$ y $p \in S$. Sea $v \in T_p S$, considérese una curva parametrizada $\alpha : (-\epsilon, \epsilon) \rightarrow U$, con $\alpha(0) = p$ y $\alpha'(0) = v$. El vector obtenido al proyectar $(\frac{d\mathbf{W}}{dt})(0)$ normalmente sobre el plano $T_p S$ es llamado derivada covariante en p sobre el campo vectorial \mathbf{W} relativo al vector v (ver Figura 2.23)*

La derivada covariante es denotada por $(\frac{D\mathbf{W}}{dt})(0)$ o $(D_v \mathbf{W})(p)$.

Sea $X(u, v)$ parametrización de S en p , $X(u(t), v(t))$ y $\mathbf{W}(t) = a(u(t), v(t)).X_u + b(u(t), v(t)).X_v$, entonces

$$\mathbf{W}(t) = a(t).X_u + b(t).X_v$$

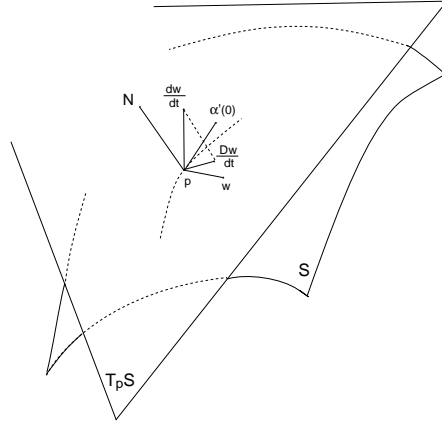


Figura 1.15: Derivada covariante

Por otro lado,

$$\begin{aligned} \frac{d\mathbf{W}}{dt} &= a(t).(u'X_{uu} + v'X_{uv}) + a'(t)X_u + b(t).(u'X_{vu} + v'X_{vv}) + b'(t)X_v \\ &= a(t)[u'(\Gamma_{11}^1 X_u + \Gamma_{11}^2 X_v + L_1 N) + v'(\Gamma_{12}^1 X_u + \Gamma_{12}^2 X_v + L_2 N)] + a'(t)X_u + \\ &\quad b(t)[u'(\Gamma_{21}^1 X_u + \Gamma_{21}^2 X_v + \bar{L}_2 N) + v'(\Gamma_{22}^1 X_u + \Gamma_{22}^2 X_v + L_3 N)] + b'(t)X_v. \end{aligned}$$

La derivada covariante, no es más $\frac{d\mathbf{W}}{dt}$ pero haciendo cero su parte normal.

$$\begin{aligned} \frac{D\mathbf{W}}{dt} &= a(t)[u'(\Gamma_{11}^1 X_u + \Gamma_{11}^2 X_v) + v'(\Gamma_{12}^1 X_u + \Gamma_{12}^2 X_v)] + a'(t)X_u + \\ &\quad b(t)[u'(\Gamma_{21}^1 X_u + \Gamma_{21}^2 X_v) + v'(\Gamma_{22}^1 X_u + \Gamma_{22}^2 X_v)] + b'(t)X_v. \end{aligned}$$

Definición 1.34. Una curva parametrizada definida en $[0, \ell]$, $\alpha : [0, \ell] \rightarrow S$ es la restricción a $[0, \ell]$ de una función diferenciable $\alpha : (0 - \epsilon, \ell + \epsilon) \rightarrow S$, $\epsilon > 0$. Sea $\alpha(0) = p$ y $\alpha(\ell) = q$, se dirá que α conecta p con q y la distancia de p a q , se define como $d(p, q) = \inf\{\ell(\gamma) : \gamma \text{ conecta a } p \text{ con } q\}$

Definición 1.35. Sea $\alpha : I \rightarrow S$ una curva parametrizada en S , un campo vectorial a lo largo de α es una correspondencia que asigna a cada $t \in I$ un vector $\mathbf{W}(t) \in T_{\alpha(t)}S$.

Se dirá que \mathbf{W} es un campo diferenciable en t_0 si y sólo si existe una parametrización $X(u, v)$ en $\alpha(t_0)$ tal que las componentes $a(t)$ y $b(t)$ de $\mathbf{W}(t) = a(t)X_u + b(t)X_v$ son diferenciables en t_0 .

\mathbf{W} es diferenciable en I si y sólo si \mathbf{W} es diferenciable para cualquier $t \in I$.

Definición 1.36. Sea \mathbf{W} un campo vectorial diferenciable a lo largo de $\alpha : I \rightarrow S$, la expresión $\frac{D\mathbf{W}}{dt}$, $t \in I$ está bien definida y es llamada derivada covariante de \mathbf{W} en t .

Definición 1.37. Un campo vectorial \mathbf{W} a lo largo de una curva parametrizada $\alpha : I \rightarrow S$ es paralelo si $\frac{D\mathbf{W}}{dt}(t) = 0$ para todo $t \in I$. (ver Figura 1.16)

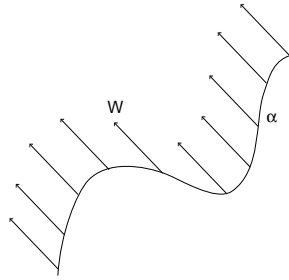


Figura 1.16: Campo Vectorial Paralelo

Proposición 1.5. Sea \mathbf{V} y \mathbf{W} campos vectoriales paralelos a lo largo de $\alpha : I \rightarrow S$, entonces $\langle \mathbf{W}(t), \mathbf{V}(t) \rangle$ es constante.

Demostración: \mathbf{V} y \mathbf{W} son campos vectoriales paralelos a lo largo de $\alpha : I \rightarrow S$, esto implica que: $\frac{D\mathbf{W}}{dt} = 0 = \frac{D\mathbf{V}}{dt}$, $\mathbf{W}'(t) = \frac{d\mathbf{W}(t)}{dt} = \lambda_1 N$ y $\mathbf{V}'(t) = \frac{d\mathbf{V}(t)}{dt} = \lambda_2 N$.

$$\text{Así; } \langle \mathbf{W}'(t), \mathbf{V}(t) \rangle = 0 \quad \text{y} \quad \langle \mathbf{W}(t), \mathbf{V}'(t) \rangle = 0$$

Derivemos $\langle \mathbf{W}(t), \mathbf{V}(t) \rangle$

$$\langle \mathbf{W}(t), \mathbf{V}(t) \rangle' = \langle \mathbf{W}'(t), \mathbf{V}(t) \rangle + \langle \mathbf{W}(t), \mathbf{V}'(t) \rangle = 0$$

Por lo tanto $\langle \mathbf{W}(t), \mathbf{V}(t) \rangle = cte$

Proposición 1.6. Sea $\alpha : I \rightarrow S$ una curva parametrizada en S y sea $\mathbf{w}_0 \in T_{\alpha(t_0)}S$, $t_0 \in I$ entonces existe un único campo vectorial $\mathbf{W}(t)$ a lo largo de $\alpha(t)$, con $\mathbf{W}(t_0) = \mathbf{W}_0$.

Demostración: [ver [1]]

Definición 1.38. Sea $\alpha : I \rightarrow S$ una curva parametrizada y $\mathbf{W}_0 \in T_{\alpha(t_0)}S$, $t_0 \in I$. Sea \mathbf{W} el campo vectorial paralelo a lo largo de α , con $\mathbf{W}(t_0) = \mathbf{W}_0$. El vector $\mathbf{W}_{t_1}S$, con $t_1 \in I$ es llamado transporte paralelo de \mathbf{W}_0 a lo largo de α en t_1 .

Definición 1.39. Una curva parametrizada, no constante $\gamma : I \rightarrow S$ se dice geodésica en $t \in I$ si el campo vectorial tangente $\gamma'(t)$ es paralelo a lo largo de γ en t ; esto es que $\frac{D\gamma'(t)}{dt} = 0$.

γ es geodésica si y sólo si γ es geodésica en todo $t \in I$.

Ejemplo: Suponga que S es una superficie regular y que $\alpha : I \rightarrow S$ es una curva diferenciable en S , donde $\alpha(0) = p$ y $\alpha'(0) = \gamma$.

Considérese el campo \mathbf{W} formado por todos $\alpha'(0) = \gamma \in T_pS$ a lo largo de la curva; así $\frac{d\mathbf{W}}{dt} = \alpha''(0)$.

Si se considera el caso particular en el cual $\alpha''(0) = \frac{d\mathbf{W}}{dt}$ es paralelo al normal N a lo largo de la curva α , entonces $\frac{D\mathbf{W}}{dt}(t) = 0$, para cualquier $t \in I$. (ver Figura 1.17)

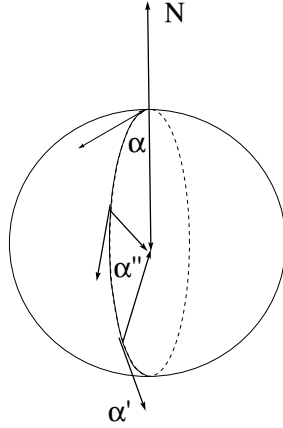


Figura 1.17: Geodésica

1.6. La Función Exponencial.

Dado $p \in S$, S es una superficie regular y $\mathbf{v} \in T_p S$, $\mathbf{v} \neq 0$. Se sabe que existe una única geodésica $\gamma : (-\epsilon, \epsilon) \rightarrow S$ con $\gamma(0) = p$ y $\gamma'(0) = \mathbf{v}$.

Se denotará $\gamma(t, \mathbf{v}) = \gamma(t) = \gamma$ y $\gamma(t, \mathbf{v})$ indicará que $\gamma'(0) = \mathbf{v}$.

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

Demostración: Sea $\alpha : (\frac{-\epsilon}{\lambda}, \frac{\epsilon}{\lambda}) \rightarrow S$ una curva parametrizada definida por $\alpha(t) = \gamma(\lambda t)$, entonces $\alpha'(0) = \lambda \gamma'(0) = \lambda \mathbf{v}$.

Por la linealidad de la derivada covariante se tiene que:

$D_{\alpha'(t)} \alpha'(t) = D_{\lambda \gamma'(t)} \lambda \gamma'(t) = \lambda^2 D_{\gamma'(t)} \gamma'(t)$, de acá α es una geodésica con las siguientes condiciones iniciales $\gamma(0)$, $\lambda \gamma'(0)$.

Por unicidad, $\alpha(t) = \gamma(t, \lambda \mathbf{v}) = \gamma(\lambda t, \mathbf{v})$

Notación: si $\mathbf{v} \in T_p S$ y $\mathbf{v} \neq 0$ es tal que $\gamma(\|\mathbf{v}\|, \frac{\mathbf{v}}{\|\mathbf{v}\|}) = \gamma(1, \mathbf{v})$, se define la exponencial dada como sigue: $exp_p(\mathbf{v}) = \gamma(1, \mathbf{v})$ y $exp_p(0) = p$.

1.7. Superficies Abstractas

Definición 1.40. *S es una superficie abstracta si S junto con una familia de parametrizaciones $X_\alpha : U_\alpha \rightarrow S$, con $U \subset \mathbb{R}^2$ abierto, X_α inyectiva se cumple que: (1) $\bigcup_\alpha X_\alpha(U_\alpha) = S$*

(2) $\forall \alpha, \beta$ con $X_\alpha(U_\alpha) \cap X_\beta(U_\beta) \neq \emptyset$; $X_\alpha \circ X_\beta^{-1}$ y $X_\beta \circ X_\alpha^{-1}$ son diferenciables.

Al par (U_α, X_α) se le llama sistema de coordenadas locales y $\{(U_\alpha, X_\alpha)\}_\alpha$ se le llama Estructura Diferenciable.

Definición 1.41. *Sean S_1 y S_2 superficies abstractas y sea $\varphi : S_1 \rightarrow S_2$ diferenciable, $\forall p \in S_1$ y $\forall w \in T_p(S_1)$. consideremos $\alpha : (-\epsilon, \epsilon) \rightarrow S_1$ en $\alpha(0) = p$ y $\alpha'(0) = w$, sea $\beta = \varphi \circ \alpha$, la función $d\varphi : T_p(S_1) \rightarrow T_{\varphi(p)}(S_2)$ dado por $d\varphi_p(w) = \beta'(0)$ es llamada la diferencial de φ en p .*

Definición 1.42. *Una superficie geométrica o Variedad Riemanniana de dimensión α , es una superficie abstracta junto a la escogencia de un producto interno $\langle \cdot, \cdot \rangle_p$ en cada $T_p S$, $p \in S$ el cual varía diferencialmente con p en el siguiente sentido, para alguna y (por tanto para toda) parametrización $X_\alpha : U_\alpha \rightarrow S$, las funciones:*

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

$$\blacksquare G(u, v) = g_{22}(u, v) = \left\langle \frac{\partial}{\partial v}, \frac{\partial}{\partial v} \right\rangle,$$

son funciones diferenciables en U_α .

Al producto interno \langle, \rangle se le llama *métrica Riemanniana sobre S* .

Definición 1.43. Una variedad Riemanniana es una n -dimensional variedad diferenciable M junto a una escogencia, para cualquier $p \in M$ de un producto interno \langle, \rangle_p en cada $T_p M$ que varía diferencialmente con p en el siguiente sentido, para alguna y (por tanto para toda) parametrización $X_\alpha : U_\alpha \rightarrow M$, con $p \in X_\alpha \subset M$ las funciones:

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

son diferenciables en $X_{\alpha'}(p)$ y x_1, \dots, x_n son coordenadas de $U_\alpha \subset \mathbb{R}^n$.

La familia $\{\langle, \rangle_p : p \in M\}$ la llamaremos *Estructura Riemanniana o Métrica Riemanniana para M* .

1.8. Función de Morse

Definición 1.44. Sea S una superficie y $f : S \rightarrow \mathbb{R}$ una función diferenciable, diremos que $p_0 \in S$ es un punto crítico de f si $f'(p_0) = 0$.

Ejemplo: Cuando $S \subset \mathbb{R}^3$ y $f : S \rightarrow \mathbb{R}$ es dad por $f(x, y, z) = z$ (la función altura), entonces los puntos críticos de f son aquellos donde el plano tangente es perpendicular al eje z .

Sea entonces $p_0 \in S$ un punto crítico de f y $\varphi : U \rightarrow S$, dado por:
 $\varphi(x, y) = (\varphi_1(x, y), \varphi_2(x, y), \varphi_3(x, y)) = \varphi_3(x, y)$, $(x, y) \in U$; por tanto una condición de p_0 para ser punto crítico de f implica que:

$$f'(\varphi(x_0, y_0)) = 0 \Rightarrow \frac{\partial \varphi_3(x_0, y_0)}{\partial x} = 0 = \frac{\partial \varphi_3(x_0, y_0)}{\partial y}$$

Así;

$$T_{p_0}S = \varphi'(x_0, y_0)(\mathbb{R}^2) = \left\{ \left(\begin{array}{cc} \frac{\partial \varphi_1(x_0, y_0)}{\partial x} & \frac{\partial \varphi_1(x_0, y_0)}{\partial y} \\ \frac{\partial \varphi_2(x_0, y_0)}{\partial x} & \frac{\partial \varphi_2(x_0, y_0)}{\partial y} \\ 0 \begin{pmatrix} \lambda \\ \mu \end{pmatrix} & 0 \end{array} \right), \lambda, \mu \in \mathbb{R} \right\}$$

$$= \{(\bar{\lambda}, \bar{\mu}, 0) : \bar{\lambda}, \bar{\mu} \in \mathbb{R}\}$$

que es lo que se quería demostrar.

Definición 1.45. Sea $f : S \rightarrow \mathbb{R}$ una función diferenciable y $p_0 \in S$ un punto crítico de f . Sea $\varphi : U \subset \mathbb{R}^2 \rightarrow V \subset S$ una parametrización con $\varphi(x_0, y_0) = p_0$.

El Hessiano de f en p_0 es una forma bilineal $f''(p_0) : T_{p_0}S \times T_{p_0}S \rightarrow \mathbb{R}$ definida por:

$$f''(p_0)(\varphi'(x_0, y_0)u, \varphi'(x_0, y_0)v) = (f \circ \varphi)''(x_0, y_0)(u, v)$$

En términos de la parametrización φ , el Hessiano de f está representado por la matriz

$$Hf(p_0) = \begin{pmatrix} \frac{\partial f \circ \varphi}{\partial x^2} & \frac{\partial f \circ \varphi}{\partial x \partial y} \\ \frac{\partial f \circ \varphi}{\partial y \partial x} & \frac{\partial f \circ \varphi}{\partial y^2} \end{pmatrix}$$

Definición 1.46. Un punto crítico $p_0 \in S$ de una función diferenciable $f : S \rightarrow \mathbb{R}$ es no degenerado si $\det H(f)(p_0) \neq 0$.

Proposición 1.7. Si S es una superficie compacta y todos los puntos críticos de la función $f : S \rightarrow \mathbb{R}$ son no degenerado, entonces tales puntos son un número finito.

Demostración: Como S es compacta, basta mostrar que los puntos críticos de f son aislados.

Sea $p_0 \in S$ un punto crítico de f y $\varphi : U \rightarrow V$, dada por $\varphi(x_0, y_0) = p_0$ una parametrización, $(f \circ \varphi)' : U \rightarrow \mathcal{L}(\mathbb{R}^2, \mathbb{R})$, donde $\mathcal{L}(\mathbb{R}^2, \mathbb{R}) = (\mathbb{R}^2)^*$ es el conjunto de las aplicaciones lineales de \mathbb{R}^2 en \mathbb{R} .

$(f \circ \varphi)' : \mathbb{R}^2 \rightarrow (\mathbb{R}^2)^*$ es un isomorfismo, pues p_0 es no degenerado, luego por el teorema de la función inversa se tiene que existen abiertos $U_1 \subset U$, en torno de (x_0, y_0) y $V_1 \subset (\mathbb{R}^2)^*$, con $(f \circ \varphi)'(x_0, y_0) = 0$, tal que $(f \circ \varphi)' : U_1 \rightarrow V_1$ es un difeomorfismo y por tanto (x_0, y_0) es el único punto de U_1 donde $(f \circ \varphi)'(x_0, y_0) = 0$.

De acá se sigue que $p_0 \in S$ es un punto crítico aislado de f .

Capítulo 2

Preliminares de Optimización

Sea f una función de \mathbb{R}^n a valores reales y considérese el siguiente problema no lineal con restricciones:

$$\text{mín}\{f(x)/x \in \mathbb{R}^n : C_i(x) = 0; i = 1, 2, \dots, m\}, \quad (2.1)$$

donde $m \leq n$, f y C_i son funciones de \mathbb{R}^n a valores reales, las cuales se suponen diferenciables \mathcal{C}^σ ($\sigma \geq 2$ a menos que se especifique lo contrario).

Se define $C : \mathbb{R}^n \rightarrow \mathbb{R}^m$ por $C(x) = (C_1(x), C_2(x), \dots, C_m(x))$ y se supone que cero es un valor regular de la función C , esto es que $C'(x)$ pertenece a $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m) = \{T : \mathbb{R}^n \rightarrow \mathbb{R}^m / T \text{ es lineal}\}$ y es de rango completo m , para todo $x \in M = C^{-1}(0_m) = \{x \in \mathbb{R}^n / C(x) = 0_m\}$. Esta hipótesis de regularidad implica que existe una submatriz $m \times m$ de la matriz jacobiana de C en $x \in M$, denotada por $[A_x]_{m \times n}$, la cual es no singular.

Denotar $B = \left(\frac{\partial C_i}{\partial x_j}\right)$, donde $i, j \in \{1, 2, \dots, m\}$ y sea $N = \{1, 2, \dots, m, m+1, \dots, n\}$, luego denotar $J = N - I = \{m+1, \dots, n\}$

Por otro lado, escribir \mathbb{R}^n como $\mathbb{R}^m \times \mathbb{R}^{n-m}$ y sea $[A_x] = [B \mid D]$, Así se utiliza el teorema de la función implícita, el cual garantiza que existe una vecindad $U_x = W \times V$ de $x = (x_I, x_J) \in \mathbb{R}^m \times \mathbb{R}^{n-m}$, $\psi_x : V \rightarrow W$ una función \mathcal{C}^σ tal que $\psi_x(y_J) = y_I$ si y sólo si $C(y_I, y_J) = 0_m$.

Por tanto, la restricción de f en $U_x \cap M$ puede ser representada por la función $\varphi_x : V \subset \mathbb{R}^{n-m} \rightarrow \mathbb{R}$, la cual es \mathcal{C}^σ y definida por:

$$\varphi_x(z) = f(\psi_x(z), z) \quad (2.2)$$

Suponiendo ahora que $x^* \in M$ es un minimizador local de f sobre M en la vecindad $U_x \cap M$ de x ; se tiene que $x^* = (x_I^*, x_J^*)$ con $\psi_x(y_J^*) = y_I^*$, y x_J^* es un minimizador local del problema de minimización sin restricciones en el conjunto abierto V .

$$\min_{z \in V} \varphi_x(z) \quad (2.3)$$

Aplicando el método tipo gradiente al problema sin restricciones (2.3), se construye una sucesión de aproximaciones $\{z^k\}$ que converge a x_J^* y se define iterativamente como sigue:

$$z^{k+1} = z^k + t_k P^k \quad (2.4)$$

donde P^k es una dirección de descenso basada en g^k , el gradiente de φ_x en z^k , y t_k es el tamaño de paso seleccionado por una búsqueda lineal de la función $j(t) = \varphi_x(z^k + t_k \cdot p^k)$ para $t > 0$. En el espacio original \mathbb{R}^n la iteración (2.4) corresponde al algoritmo:

$$x_J^{k+1} = x_J^k + t_k P^k \quad (2.5)$$

$$x_I^{k+1} = \psi_x(x_J^{k+1}) \quad (2.6)$$

El gradiente de φ_x puede ser expresado en términos de los datos originales del problema, ¿Cómo?.

Se define, $H : V \subset \mathbb{R}^{n-m} \rightarrow \mathbb{R}^m \times \mathbb{R}^{n-m} = \mathbb{R}^n$, dada por

$H(z) = (\psi_x(z), z)$, así:

$$\varphi_x(z) = f(\psi_x(z), z) = f \circ H(z)$$

Luego;

$$\begin{aligned} D\varphi_x(z) &= D(f \circ H)(z) \\ &= [Df(H(z))]_{1 \times n} [DH(z)]_{n \times (n-m)} \\ &= \begin{bmatrix} \frac{\partial f}{\partial x_I} & \frac{\partial f}{\partial x_J} \end{bmatrix}_{1 \times n} \begin{bmatrix} D\psi_x(z) \\ I_{n-m} \end{bmatrix}_{n \times (n-m)} \\ &= \frac{\partial f}{\partial x_I} D\psi_x(z) + \frac{\partial f}{\partial x_J} I_{n-m} \end{aligned} \quad (*)$$

Ahora se estudiara quien es $D\psi_x(z)$,

Se deriva $C(\psi_x(z), z) = C \circ H(z) = 0_m$

$$\begin{aligned} D(C \circ H)(z) &= [DC(H(z))]_{m \times n} [DH(z)]_{n \times (n-m)} \\ &\Rightarrow \begin{bmatrix} B & D \end{bmatrix}_{m \times n} \begin{bmatrix} D\psi_x(z) \\ I_{n-m} \end{bmatrix}_{n \times (n-m)} = [0]_{m \times (n-m)} \\ &\Rightarrow B \cdot D\psi_x(z) + D = [0] \\ &\Rightarrow D\psi_x(z) = -B^{-1}D \end{aligned} \quad (**)$$

Sustituyendo (*) en (**), se obtiene que:

$$\begin{aligned} D\varphi_x(z) &= \frac{\partial f}{\partial x_I} B^{-1}D + \frac{\partial f}{\partial x_J}, \text{ donde} \\ g^k &= \nabla \varphi_x(z) = [D\varphi_x(z)]^T \end{aligned}$$

De acá:

$$g^k = \nabla_J f(x^k) - D^T \cdot (B^{-1})^T \cdot \nabla_I f(x^k) \quad (2.7)$$

Este es llamado el **gradiente reducido** para la partición $N = I \oplus J$, el tamaño de paso t_k debe ser elegido a partir de x^k por una búsqueda a lo largo de la curva dada por:

$$X(t) = \left(\psi(x_J^k + t.p^k), x_J^k + t.p^k \right) \quad (2.8)$$

Así se produce un suficiente decrecimiento de la función objetivo, de esta manera Gabay y Luenberger propusieron ideas basadas en el **método gradiente reducido**, generalizando el método de máximo descenso y analizando sus propiedades de convergencia.

Notar que si x^* no pertenece a $U_{x_k} \cap M$, también es posible encontrar una aproximación x^{k+1} de acuerdo con (2.6) y (2.5) al restringir la búsqueda a lo largo de (2.8) para un intervalo $[0, \tilde{t}]$ tal que $x^{k+1} \in U_{x_k} \cap M$; para cada x^{k+1} se puede construir una nueva partición $N = I \oplus J$, y la iteración x^k es eventualmente alcanzada tal que $x^* \in U_{x_k} \cap M$.(ver Figura 2.1)

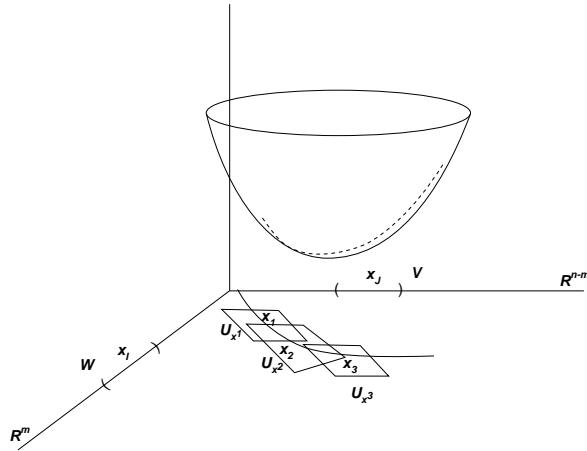


Figura 2.1: Construcción de Vecindades

2.1. El conjunto de restricciones M visto como una variedad

Se estudiarán las propiedades del conjunto de restricciones $M = C^{-1}(0_m)$, el cual es definido por la aplicación $C : \mathbb{R}^n \rightarrow \mathbb{R}^m$, ($m \leq n$) que se supone diferenciable ($\sigma \geq 2$).

Recordar que 0_m es un valor regular de C si y sólo si $C'(x)$ está dentro del conjunto $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m) = \{T : \mathbb{R}^n \rightarrow \mathbb{R}^m / T \text{ es lineal}\}$ y es de rango completo m para cualquier $x \in M$. Esto es, $C'(x)$ es sobreyectiva para todo $x \in M$.

Teorema 2.1. *Sea $C : \mathbb{R}^n \rightarrow \mathbb{R}^m$ una aplicación \mathcal{C}^σ , tal que 0_m es un valor regular de C , entonces el conjunto $M = C^{-1}(0_m)$ es una subvariedad $(n - m)$ dimensional en \mathbb{R}^n de clase $\mathcal{C}^{\sigma-1}$.*

Demostración: Se fija un punto $x \in M \subset \mathbb{R}^n$, así $C(x) = 0_m$. Como 0_m es un valor regular de C , entonces la derivada $C'(x)$ asigna una aplicación de \mathbb{R}^n sobre \mathbb{R}^m , la cual es sobreyectiva para todo $x \in M$.

Luego, por el teorema de nulidad y rango se tiene que:

$$\dim(N(C'(x))) + \dim(Im(C'(x))) = n.$$

Como $C'(x)$ es sobreyectiva, entonces $\dim(Im(C'(x))) = m$ y por tanto la $\dim(N(C'(x))) = n - m$. Ahora, se escoge una aplicación lineal $Z_x \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^{n-m})$, tal que $Z_x : \mathbb{R}^n \rightarrow \mathbb{R}^{n-m}$ es no singular en este subespacio $N(C'(x)) \subset \mathbb{R}^n$; lo que queremos es escoger $Z_x : \mathbb{R}^n \rightarrow \mathbb{R}^{n-m}$ de manera tal que $Z_x/N(C'(x))$ sea inyectiva; es decir: $N(Z_x/N(C'(x))) = \{0_{n-m}\}$, o lo que es lo mismo, para cualquier $v \in N(C'(x))$ se cumpla que $Z_x \cdot (v) = 0_{n-m}$ si y sólo si $v = 0_n$.

De esta manera se obtiene:

$$N(Z_x) \cap N(C'(x)) = \{0_n\} \tag{2.9}$$

Se define ahora una aplicación $s_x : \mathbb{R}^n \rightarrow \mathbb{R}^m \times \mathbb{R}^{n-m}$, dada por:

$$s_x(y) = \begin{pmatrix} C(y) \\ Z_x(y-x) \end{pmatrix}_{n \times 1}, \forall y \in \mathbb{R}^n \quad (2.10)$$

nótese que $s_x(x) = 0_n$ y que $s_x \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$, luego:

$$[s'_x](y) = \begin{bmatrix} C'(x) \\ Z_x(x-x) \end{bmatrix} (y) = \begin{bmatrix} C'(x) \\ Z_x \end{bmatrix} (y)$$

de esta forma se obtiene que $s'_x : \mathbb{R}^n \rightarrow \mathbb{R}^m \times \mathbb{R}^{n-m}$ viene dada por:

$$[s'_x](y) = \begin{bmatrix} C'(x)(y) \\ Z_x(y) \end{bmatrix}, \forall y \in \mathbb{R}^n \quad (2.11)$$

Afirmación: s'_x es no singular

En efecto; sea $v \in \mathbb{R}^n$

$$\begin{aligned} s_x(x)(v) = [0]_{n \times n} &\Leftrightarrow \begin{bmatrix} C'(x)(v) \\ Z_x(v) \end{bmatrix}_{n \times n} = [0]_{n \times n} \\ &\Leftrightarrow \begin{cases} C'(x)(v) = [0]_{m \times n} \\ Z_x(v) = [0]_{(n-m) \times n} \end{cases} \\ &\Leftrightarrow v \in N(C'(x)) \wedge v \in N(Z_x) \\ &\Leftrightarrow v \in N(Z_x) \cap N(C'(x)) = \{0_n\} \\ &\Leftrightarrow v = 0_n \end{aligned}$$

luego el teorema de la función inversa implica que para la aplicación $s_x : \mathbb{R}^n \rightarrow \mathbb{R}^n = \mathbb{R}^m \times \mathbb{R}^{n-m}$ existe alguna vecindad U_x de $x \in \mathbb{R}^n$, la cual es difeomorfa a la vecindad $W \times V$ de cero en $\mathbb{R}^m \times \mathbb{R}^{n-m}$.

En el cambio de coordenadas s_x , el conjunto M se transforma localmente en un subespacio $(n-m)$ dimensional de $\{0\} \times \mathbb{R}_{n-m}$, ya que s_x mapea $U_x \cap M$ difeomorficamente sobre $\{0\} \times \mathbb{R}_{n-m}$.

La aplicación Z_x define por restricción un sistema coordinado de M alrededor de x sobre el dominio coordinado $U_x \cap M$ y las componentes $(z_1, z_2, \dots, z_{n-m})$ de la imagen $z = Z_x(x - y)$, $y \in U_x \cap M$ son llamadas **coordenadas locales de y** .

El difeomorfismo inverso s_x^{-1} permite definir una **parametrización local no lineal** $\theta_x : V \rightarrow U_x \cap M$ de M alrededor de x , dada por:

$$y = \theta_x(z) = s_x^{-1}(0_m, z) \quad (2.12)$$

Probando que el mapeo Z_x es escogido como una función diferenciable $\mathcal{C}^{\sigma-1}$ de x , entonces el cambio de coordenadas locales alrededor de x al cambio de coordenadas x' tal que $(U_x \cap U_{x'} \cap M \neq \emptyset)$ definido por:

$$Z_{x'} \circ \theta_x : Z_x(U_x \cap U_{x'} \cap M) \longrightarrow Z_{x'}(U_x \cap U_{x'} \cap M)$$

es diferenciable $\mathcal{C}^{\sigma-1}$.

Se denotará por $[A_x]_{m \times n}$ la matriz Jacobiana de C en x , y se mantendrá $[Z_x]_{(n-m) \times n}$ para denotar la matriz que representa la aplicación lineal Z_x .

La matriz Jacobiana de s_x se puede ver como la matriz $[S_x]_{n \times n}$, dada por:

$$S_x = \begin{bmatrix} A_x \\ Z_x \end{bmatrix}_{n \times n}, \quad (2.13)$$

Nótese que la hipótesis de regularidad implica que: $\text{rang}(A_x) = m$

Definición 2.1. Sea $R \in M_{l \times n}$, la cual es de rango completo $l \leq n$, se llamará inversa por la derecha de R a la matriz $R^- \in M_{n \times l}$ la cual es de rango completo l y cumple que: $[R]_{l \times n}[R^-]_{n \times l} = [Id]_{l \times l}$, tal inversa existe, pero no es única.

El siguiente resultado se refiere a la escogencia de $[Z_x]$ y $[Z_x^-]$, en particular a la escogencia de una matriz inversa por la derecha de $[A_x]$, digamos $[A_x^-]$, y así expresar el inverso $[S_x^{-1}]$ de una forma más sencilla.

Proposición 2.1. Sea $[A_x^-]_{n \times m}$ una inversa por la derecha de $[A_x]_{m \times n}$, entonces existe una matriz $[Z_x]_{(n-m) \times n}$ de rango completo $(n-m)$ y una matriz inversa por la derecha $[Z_x^-]_{n \times (n-m)}$ que satisface:

$$[Z_x]_{(n-m) \times n} \cdot [A_x^-]_{n \times m} = [0]_{(n-m) \times m}, \quad [A_x]_{m \times n} \cdot [Z_x^-]_{n \times (n-m)} = [0]_{m \times (n-m)} \quad (2.14)$$

tal que la matriz $[S_x]_{n \times n}$ es no singular y su inversa viene dada por:

$$S_x^{-1} = \left[\begin{array}{cc} A_x^- & Z_x^- \\ \hline & \end{array} \right]_{n \times n} \quad (2.15)$$

Demostración: Una condición necesaria y suficiente para la no singularidad de $[S_x]_{n \times n}$ es:

$$N(Z_x) \cap N(A_x) = \{0_n\} \quad (2.16)$$

Dado que $[A_x]$ es de rango completo m ; es decir que $\text{rang}(A_x) = m = \dim I(A_x)$, se tiene por el teorema de nulidad y rango que

$$\dim N(A_x) + \dim I(A_x) = n$$

De esta manera podemos decir que $N(A_x)$ es un subespacio en \mathbb{R}^n de dimensión $(n-m)$, por otro lado recordemos que $Z_x : \mathbb{R}^n \rightarrow \mathbb{R}^{n-m}$ es inyectiva en el subespacio $N(A_x) \subset \mathbb{R}^n$, esto es $Z_x/N(A_x) : N(A_x) \rightarrow \mathbb{R}^{n-m}$ es sobreyectiva, de esta manera la $\dim I(Z_x/N(A_x)) = n-m$; más aún $Z_x : \mathbb{R}^n \rightarrow \mathbb{R}^{n-m}$ es sobreyectiva y $\dim I(Z_x) = n-m$

Así, nuevamente utilizando el teorema de nulidad y rango obtenemos que $\dim N(Z_x) = m$, de acá se tiene que $N(Z_x)$ es un subespacio en \mathbb{R}^n de dimensión m .

Ahora bien, $N(A_x)$ y $N(Z_x)$ son subespacios en \mathbb{R}^n de dimensión $(n-m)$ y m respectivamente, y además $N(Z_x) \cap N(A_x) = \{0_n\}$.

Por otro lado, la inversa de $[A_x^-]$ debe ser de rango m , es decir $[A_x^-]$ tiene m columnas linealmente independiente, es decir $\dim(Sg(col(A_x^-))) = m$, y sus columnas generan un subespacio $\mathcal{R}(A_x^-)$ en \mathbb{R}^n , la cual tiene dimensión m y cumple que:

$$N(A_x) \cap \mathcal{R}(A_x^-) = \{0_n\}$$

Por otro lado, se tiene que $N(A_x)$ y $\mathcal{R}(A_x^-)$ son subespacios ortogonales y como $N(Z_x) \cap N(A_x) = \{0_n\}$, indica que $N(Z_x)$ es complemento de $N(A_x)$ en \mathbb{R}^n ; y además $\dim N(Z_x) = \dim \mathcal{R}(A_x^-)$, por lo tanto: $\mathcal{R}(A_x^-) = N(Z_x)$ y se puede decir que:

$$[Z_x]_{(n-m) \times n} \cdot [A_x^-]_{n \times m} = [0]_{(n-m) \times m}$$

Un argumento similar demuestra que $\mathcal{R}(Z_x^-) = N(A_x)$, es decir que:

$$[A_x]_{m \times n} \cdot [Z_x^-]_{n \times (n-m)} = [0]_{m \times (n-m)}$$

Obsérvese que $y = A_x^- \cdot a + Z_x^- \cdot z$, donde $a \in \mathbb{R}^m$ y $z \in \mathbb{R}^{n-m}$ es solución de la ecuación (2.15)

En efecto,

$$\begin{aligned} S_x \cdot y &= \begin{bmatrix} A_x \\ Z_x \end{bmatrix}_{n \times n} \cdot [y]_{n \times 1} = \begin{bmatrix} A_x \\ Z_x \end{bmatrix}_{n \times n} \cdot [A_x^- \cdot a + Z_x^- \cdot z]_{n \times 1} \\ &= \begin{bmatrix} A_x \\ Z_x \end{bmatrix}_{n \times n} \cdot [A_x^- \cdot a]_{n \times 1} + \begin{bmatrix} A_x \\ Z_x \end{bmatrix}_{n \times n} \cdot [Z_x^- \cdot z]_{n \times 1} \\ &= \begin{bmatrix} a \\ 0 \end{bmatrix}_{n \times 1} + \begin{bmatrix} 0 \\ z \end{bmatrix}_{n \times 1} \\ &= \begin{bmatrix} a \\ z \end{bmatrix}_{n \times 1} \end{aligned}$$

Esta solución es única, pues S_x es no singular; por otro lado tenemos que:

$$S_x[y] = [S_x] \cdot [A_x^- \quad Z_x^-] \cdot \begin{bmatrix} a \\ z \end{bmatrix} = \begin{bmatrix} a \\ z \end{bmatrix}$$

Por lo tanto, $S_x^{-1} = [A_x^- \quad Z_x^-]$

Se define la matriz $[P_x]_{n \times n}$ por:

$$P_x = [Id] - [A_x^-][A_x]$$

Esta es la matriz proyección sobre $N(A_x)$, de esta manera $[A_x][P_x] = 0_{m \times n}$ y las matrices Z_x y Z_x^- definidas en la proposición 2.1 satisfacen:

$$[Z_x^-][Z_x] = P_x$$

2.2. Geometría de la variedad M

2.2.1. Estructura Riemanniana sobre M

Dado $x \in M$, se define el espacio tangente $T_x M$ a la variedad M en x , como el subespacio $(n - m)$ dimensional de \mathbb{R}^n

$$T_x M = N(A_x) = \{v \in \mathbb{R}^n / [A_x](v) = [0]_{n \times 1}\} \quad (2.17)$$

Al elegir el sistema de coordenadas local definido en la vecindad $U_x \cap M$ de x en M por:

$$Z_y = Z_x(y - x), \forall y \in U_x \cap M, \quad (2.18)$$

se tiene como se mostró en la proposición 2.1 que $\mathcal{R}(Z_x^-) = N(A_x)$; de esta manera podemos escribir el plano tangente como sigue:

$$T_x M = \mathcal{R}(Z_x^-) = \{Z_x^- \cdot p / p \in \mathbb{R}^{n-m}\}.$$

Es conveniente dotar a $T_x M$ con una forma bilineal positiva $\gamma_x : T_x M \times T_x M \rightarrow \mathbb{R}$, llamada **métrica Riemanniana**. La forma bilineal γ_x es una función suave de x y define una estructura Riemanniana γ en M .

Una elección natural consiste en tomar:

$$\gamma_x^E(v, w) = \langle v, w \rangle_n, \quad \forall v, w \in T_x M \subset \mathbb{R}^n, \quad (2.19)$$

donde \langle, \rangle_n denota el producto escalar ordinario en \mathbb{R}^n , γ_x^E es llamada la métrica Riemanniana en M inducida por la estructura Euclídea en \mathbb{R}^n .

En el sistema local (2.18) la métrica inducida puede ser expresada como:

$$\gamma_x^E(Z_x^- \cdot p, Z_x^- \cdot q) = \langle Z_x^{-T} \cdot Z_x^- \cdot p, Z_x^{-T} \cdot Z_x^- \cdot q \rangle_{n-m},$$

para cualquier $p, q \in \mathbb{R}^{n-m}$, el cual coincide con el producto escalar ordinario en \mathbb{R}^{n-m} si y sólo si $Z_x^{-T} \cdot Z_x^- = Id_{n-m}$

Puede ser conveniente definir directamente la formula en el sistema de coordenadas local por:

$$\gamma_x^Z(Z_x^- \cdot p, Z_x^- \cdot q) = \langle p, q \rangle_{n-m}, \quad \forall p, q \in \mathbb{R}^{n-m}. \quad (2.20)$$

Observemos que γ_x^Z se puede ver como la métrica sobre M inducida por un producto escalar en \mathbb{R}^n , definido alrededor de x por:

$$\langle v, w \rangle_{S_x} = \langle S_x \cdot v, S_x \cdot w \rangle_n, \quad \forall v, w \in \mathbb{R}^n,$$

donde S_x es la matriz definida en (2.13)

Una estructura Riemanniana general en M se puede definir en un sistema de coordenadas local alrededor de x por:

$$\gamma_x^G((Z_x^- \cdot p, Z_x^- \cdot q) = \langle p, G_x q \rangle_{n-m} \quad (2.21)$$

donde $G_x = (g_{ij})$ es una matriz simétrica definida positiva de dimensión $(n - m)$, de la función suave de x .

2.2.2. Diferenciación covariante y Geodésicas

Definición 2.2. *Un campo vectorial \mathbf{V} sobre la subvariedad M en \mathbb{R}^n , es una aplicación $\mathbf{V} : M \rightarrow T_x M \subset \mathbb{R}^n$, tal que $\mathbf{V}(x) \in T_x M$, para cualquier $x \in M$.*

Definición 2.3. *Sea $x \in M$, dado un vector $\mathbf{v} \in T_x M$ y el campo vectorial \mathbf{W} en M , se define un nuevo vector $D_{\mathbf{v}}(\mathbf{W}) \in T_x M$, llamado la derivada covariante de \mathbf{W} a lo largo de \mathbf{v} , la aplicación:*

$$\tau_x(\mathbf{W})(\mathbf{v}) = D_{\mathbf{v}}(\mathbf{W}), \quad (2.22)$$

la cual satisface:

$$\tau_x(\mathbf{W})(\alpha_1.v_1 + \alpha_2.v_2) = D_{(\alpha_1.v_1 + \alpha_2.v_2)}(\mathbf{W}) \quad (2.23a)$$

$$= \alpha_1.D_{v_1}(\mathbf{W}) + \alpha_2.D_{v_2}(\mathbf{W}), \quad (2.23b)$$

$$\tau_x(\mathbf{V} + \mathbf{W})(\mathbf{v}) = D_{\mathbf{v}}(\mathbf{V} + \mathbf{W}) = D_{\mathbf{v}}(\mathbf{V}) + D_{\mathbf{v}}(\mathbf{W}), \quad (2.23c)$$

$$\tau_x(\mathbf{f}.\mathbf{W})(\mathbf{v}) = f(x).D_{\mathbf{v}}(\mathbf{W}) + f'(x)(v).\mathbf{W}(\mathbf{x}), \quad (2.23d)$$

donde f es cualquier función suave sobre M de valor real que especifica una conexión afín en M sobre x .

Sean \mathbf{V} y \mathbf{W} campos vectoriales en M , se define el campo $D_{\mathbf{v}}(\mathbf{W})$, la derivada covariante de \mathbf{W} con respecto a \mathbf{V} en M , por estos valores:

$$D_{\mathbf{V}}(\mathbf{W})(x) = D_{\mathbf{v}}(\mathbf{W}), \quad \text{donde } v = \mathbf{V}(\mathbf{x}) \in T_x M, \quad (2.24)$$

la derivada covariante es así especificada globalmente en M .

Dado el sistema local de coordenadas (2.18) alrededor de x , los vectores columnas e_i , ($i = 1, 2, \dots, n - m$) de la matriz $[Z_x^-] = [e_1 e_2 \dots e_{n-m}]$ forman la base de $T_x M$ en \mathbb{R}^n .

La aplicación $E_i : M \rightarrow \mathbb{R}^n \mathcal{C}^{\sigma-1}$ es tal que $E_i(x) = e_i$ son campos de vectores y se dice que forman la base de los campos asociados en todo x .

Las propiedades (2.23) muestran que la derivada covariante en M está determinada por los campos $D_{E_i}E_j$

Es de costumbre expresar estos campos de vectores en la base de los campos

$$D_{E_i}E_j = \sum_{k=1}^{n-m} \Gamma_{ij}^k \cdot E_k \quad (2.25)$$

La función suave Γ_{ij}^k determina la derivada covariante alrededor de x y se les llaman coeficientes de la conexión.

Definición 2.4. Una curva parametrizada en M , es una aplicación $X(\cdot) : \mathbb{R} \rightarrow M \subset \mathbb{R}^n$ la cual es diferenciable; el campo $\dot{X}(\cdot)$ es definido por:

$$\dot{X}(t) = \frac{dX(t)}{dt} \in T_{X(t)}M, \quad \forall t \in \mathbb{R}. \quad (2.26)$$

El campo vectorial \mathbf{V} en M define por restricción un campo vectorial $\mathbf{V}(\cdot)$ a lo largo de la curva $X(\cdot)$, la cual asigna a cada $t \in \mathbb{R}$ un vector tangente $\mathbf{V}(\mathbf{t}) = \mathbf{V}(X(\mathbf{t})) \in T_{X(t)}M$.

Definición 2.5. $\mathbf{V}(\cdot)$ será llamado **Campo vectorial paralelo a lo largo de una curva** $X(\cdot)$, si su derivada covariante denotada por $\frac{DV}{dt}$, es idénticamente cero; es decir:

$$\frac{DV}{dt} = D_{\dot{X}}V = 0 \quad (2.27)$$

Usando el sistema de coordenadas local (2.18) alrededor del punto x sobre la curva y los campos asociados a la base E_i , el campo vectorial $\mathbf{V}(\cdot)$ puede ser expresado únicamente de la forma

$$\mathbf{V} = \sum_{i=1}^{n-m} v_i \cdot E_i,$$

donde v_i son funciones diferenciables de valor real, mientras que el campo velocidad es:

$$\dot{X} = \sum_{i=1}^n \frac{dZ_i}{dt} \cdot E_i, \text{ donde } Z_i(t) = Z[X(t)]$$

Al suponer que $\mathbf{V}(\cdot)$ es un campo vectorial paralelo a lo largo de la curva X , es decir que $\frac{dV}{dt} = D_{\dot{X}}V = 0$, se cumple que:

$$\frac{dV_k}{dt} + \sum_{i,j=1}^{n-m} \frac{dZ_i}{dt} \Gamma_{ij} v_j = 0, \quad k = 1, 2, \dots, n - m \quad (2.28)$$

Así se tiene el siguiente resultado de existencia y unicidad.

Proposición 2.2. *Sea $X(\cdot)$ una curva parametrizada sobre M definida en $[0, T]$. Para cada vector $v \in T_{X(0)}M$, hay un único campo paralelo $\mathbf{V}(\mathbf{0})$ a lo largo de $X(\cdot)$ tal que $\mathbf{V}(\mathbf{0}) = v$.*

Demostración: Observemos que (2.28) es un sistema de ecuaciones diferenciable de primer orden, donde nos dan una condición inicial $\mathbf{V}(\mathbf{0}) = v$; así el sistema de ecuación tiene una única solución. De esta manera concluimos que hay un único campo paralelo $\mathbf{V}(\mathbf{0})$ a lo largo de $X(\cdot)$ tal que $\mathbf{V}(\mathbf{0}) = v$.

Definición 2.6. *Se llamará transporte paralelo a lo largo de $X(\cdot)$, desde $X(0)$ hasta $X(t)$ a la aplicación $\pi_0^t : T_{X(0)}M \rightarrow T_{X(t)}M$, la cual se define como sigue $\pi_0^t(v(0)) = v(t)$ y es un isomorfismo lineal.*

Proposición 2.3. *Existe una única conexión (simétrica) sobre una variedad Riemanniana tal que el transporte paralelo preserva la métrica; es decir:*

$$D_{E_i}^{E_j} = D_{E_j}^{E_i}, \quad i, j = 1, 2, \dots, n - m$$

$$\frac{\partial}{\partial Z_k} \langle E_i, E_j \rangle = \langle D_{E_k}^{E_i}, E_j \rangle + \langle E_i, D_{E_k}^{E_j} \rangle$$

La conexión es simétrica, esto es $\Gamma_{ij}^k = \Gamma_{ji}^k, \forall i, j$, además dado el sistema de coordenadas locales $\{Z_i\}$ la función coeficiente Γ_{ij} define la conexión y está determinada únicamente por la matriz simétrica de la función G definida como la métrica Riemanniana en (2.21)

De la proposición 2.3 se tiene que:

$$\frac{\partial}{\partial Z_i} \langle E_j, E_k \rangle + \frac{\partial}{\partial Z_j} \langle E_k, E_i \rangle - \frac{\partial}{\partial Z_k} \langle E_i, E_j \rangle = 2 \langle D_{E_i}^{E_j}, E_k \rangle$$

luego;

$$\frac{\partial}{\partial Z_i} g_{jk} + \frac{\partial}{\partial Z_j} g_{ki} - \frac{\partial}{\partial Z_k} g_{ij} = 2 \sum_{l=1}^{n-m} \Gamma_{ij}^l \cdot g_{lk}$$

así,

$$\sum_{l=1}^{n-m} \Gamma_{ij}^l \cdot g_{lk} = \frac{1}{2} \left\{ \frac{\partial}{\partial Z_i} g_{jk} + \frac{\partial}{\partial Z_j} g_{ki} - \frac{\partial}{\partial Z_k} g_{ij} \right\}$$

Ya que el $\det(G) \neq 0$, entonces existe la matriz inversa de $G = (g_{lk})$ la cual denotaremos por $G^{-1} = (g^{lk})$, de esta manera podemos resolver el sistema y obtener Γ_{ij} como sigue:

$$\Gamma_{ij}^l = \frac{1}{2} \sum_{k=1}^{n-m} g^{lk} \cdot \left(\frac{\partial}{\partial Z_i} g_{jk} + \frac{\partial}{\partial Z_j} g_{ki} - \frac{\partial}{\partial Z_k} g_{ij} \right), \quad i, j, k = 1, 2, \dots, n-m. \quad (2.29)$$

A ésta se le llama segunda identidad de Christoffel.

Definición 2.7. A la curva $X(\cdot) : \mathbb{R} \rightarrow M$ se le llamará *geodésica* si el campo velocidad \dot{X} es paralelo a lo largo de x , es decir:

$$\frac{D\dot{X}}{dt} = D_{\dot{X}}\dot{X} = 0$$

En términos del sistema local de coordenadas (2.18) las funciones de coordenadas locales $Z = Z_k, k = 1, 2, \dots, n-m$ cumple con el siguiente sistema de ecuaciones diferenciables de segundo orden.

$$\frac{d^2 Z_k}{dt^2} + \sum \Gamma_{ij}^k \cdot \frac{dZ_i}{dt} \cdot \frac{dZ_j}{dt} = 0, \quad \text{para } k = 1, 2, \dots, n-m. \quad (2.30)$$

el cual se obtiene de (2.28) haciendo $V_k = \frac{dZ_k}{dt}$.

A partir de ahora se considerara M dotada con la métrica Riemanniana y se trabajara con la única conexión compatible con él, esto es que los coeficientes de la conexión cumplan con la segunda identidad de Christoffel.

El siguiente resultado sobre la solución de (2.30) establece la equivalencia de nuestra definición de geodésica con la elemental (Curva de longitud minimal)

Proposición 2.4. *Sea W un subconjunto conexo de M en \mathbb{R}^n . Dado $x \in W$ y $\mathbf{p} \in T_x M$, existe una única curva geodésica $X(\cdot)$ tal que:*

$$X(0) = x \quad , \quad \dot{X}(0) = \mathbf{p}$$

La aplicación X es definida para todo $t \in \mathbb{R}$ y toma estos valores en W , o se define en el intervalo $[-T, T]$ y $X(T)$ o $X(-T)$ pertenecen a la frontera de W . Cualquiera de los dos puntos de W pueden ser conectados por la única geodésica la cual minimiza la longitud de arco entre los puntos.

La demostración es una consecuencia inmediata del teorema de unicidad de ecuaciones diferenciables.

La curva geodésica se denota por:

$$X(t) = \exp_x(t \cdot \mathbf{p}), \quad (2.31)$$

y el sistema de coordenadas local $\exp_x^{-1} : U_x \cap M \longrightarrow V \subset T_x M \approx \mathbb{R}^{n-m}$ es llamado el sistema de coordenadas normal alrededor de x .

Observe que en tales coordenadas, la geodésica es localmente parametrizada en un intervalo lineal

$$\{t \cdot \mathbf{p} / t \in (-\epsilon, \epsilon)\}$$

Finalmente nótese que si en el sistema de coordenadas local la métrica Riemanniana es definida por la métrica constante G , la función de coordenadas local $Z = Z_k$, $K = 1, 2, \dots, n - m$ define la geodésica alrededor de x , satisfaciendo la ecuación diferencial de segundo orden

$$\frac{d^2 Z_K}{dt} = 0, \quad K = 1, 2, \dots, n - m$$

Pues, $\Gamma_{ij}^k = 0$; de acá $Z(t) = t \cdot \mathbf{p}$ y el sistema de coordenadas es común.

Capítulo 3

Métodos de descenso por Geodésicas

Se pasará a estudiar la solución del problema de programación no lineal (2.1), que ahora se escribirá de la siguiente manera

$$\min\{f(x) \mid x \in M\}, \quad (3.1)$$

es decir, el problema de minimizar la función diferenciable de valor real f en la variedad diferenciable M dotada de una estructura Riemanniana suave γ .

Definición 3.1. Dado $x \in M$, se define la derivada de f en x sobre M , como la forma lineal en $T_x M$; $D_M f(x) : T_x M \subset \mathbb{R}^n \rightarrow \mathbb{R}$ tal que:

$$D_M f(x)(v) = f'(x).v, \quad v \in T_x M, \quad (3.2)$$

donde $f'(x)$ es la derivada ordinaria de f en x considerada como una función en \mathbb{R}^n .

Ahora se dará una condición de optimalidad de primer orden necesaria para el problema (3.1)

Teorema 3.1. Si $x^* \in M$ es un mínimo local de f , entonces

$$D_M f(x^*) = 0. \quad (3.3)$$

Demostración: Sea $X : (-\epsilon, \epsilon) \rightarrow M$ una curva, tal que $X(0) = x^*$ y $X'(0) = v \in T_{x^*}M$, sea $f : M \rightarrow \mathbb{R}$; así podemos estudiar $f \circ X : (-\epsilon, \epsilon) \rightarrow \mathbb{R}$ y como x^* es un mínimo local de f en M , entonces se cumple que:

$$f(x) \geq f(x^*), \quad \forall x \in U_{x^*} \cap M. \quad (3.4)$$

donde U_{x^*} es una vecindad en \mathbb{R}^n , y es equivalente a

$$f(X(t)) \geq f(X(0)) = f(x^*). \quad (3.5)$$

Obsérvese lo siguiente:

$$\lim_{x \rightarrow 0^+} \frac{f(X(t)) - f(X(0))}{t} = \frac{+}{+} \geq 0.$$

$$\lim_{x \rightarrow 0^-} \frac{f(X(t)) - f(X(0))}{t} = \frac{+}{-} \leq 0.$$

Como $f \circ X$ es diferenciable, entonces

$$0 \leq \lim_{x \rightarrow 0^+} \frac{f(X(t)) - f(X(0))}{t} = \lim_{x \rightarrow 0} \frac{f(X(t)) - f(X(0))}{t} = \lim_{x \rightarrow 0^-} \frac{f(X(t)) - f(X(0))}{t} \leq 0$$

por lo tanto; $\lim_{x \rightarrow 0} \frac{f(X(t)) - f(X(0))}{t} = 0$; es decir:

$$0 = (f \circ X)'(0) = f'(X(0)).X'(0) = f'(x^*).v = D_M f(x^*).v,$$

así se cumple que $D_M f(x^*).v = 0$, $\forall v \in T_{x^*}M$.

Por lo tanto $D_M f(x^*) = 0$.

Definición 3.2. Cada punto $x^* \in M$ tal que $D_M f(x^*) = 0$ es llamado punto crítico de f .

Definición 3.3. Dado $v \in T_x M$, sea \mathbf{V} un campo vectorial diferenciable sobre M tal que $\mathbf{V}(x) = v$. Se define la función $\mathcal{C}^{\sigma-1}$ diferenciable

$Vf : M \rightarrow \mathbb{R}$ por

$$Vf(x) = D_M f(x).v, \quad (3.6)$$

puesto que $\sigma \geq 2$, esta función tiene una derivada sobre M en x , y $D_M Vf(x) \in \mathcal{L}(T_x M, \mathbb{R})$

Definición 3.4. Sea x^* un punto crítico, se define el **Hessiano de f en M** como la forma bilineal $Hf(x^*) : T_x M \times T_x M \rightarrow \mathbb{R}$ dado por:

$$Hf(x^*)(v, w) = D_M Vf(x^*)(w), \quad \forall v, w \in T_x M, \quad (3.7)$$

esta forma está bien definida; es decir que es independiente de la escogencia del campo vectorial \mathbf{V} y simétrica.

La siguiente definición es válida sólo en un punto crítico.

Definición 3.5. Un punto x^* se dice que es no degenerado si y sólo si $Hf(x^*)$ es no degenerado; es decir que:

$$Hf(x^*)(v, v) = D_M Vf(x^*)(v) = 0 \text{ implica } v = 0$$

En lo que sigue, se asumirá que todos los puntos críticos de f en M son no degenerados.

Definición 3.6. Una función diferenciable $f : M \rightarrow \mathbb{R}$ es una **función de Morse** si todos sus puntos críticos son no degenerados.

Teorema 3.2. Un punto $x^* \in M$ es un mínimo local estricto de una función de Morse f en la variedad Riemanniana M , si y sólo si x^* es un punto crítico de f y la forma de acción $Hf(x^*)$ es definida positiva.

Demostración:

(\Rightarrow) Sea x^* un mínimo local estricto de la función de Morse f en la variedad Riemanniana M .

Sea $X(\cdot)$ la curva geodésica a partir de x^* con $v \in T_{x^*}M$, y sea \mathbf{V} el único campo vectorial paralelo a lo largo de $X(\cdot)$, de manera que $\mathbf{V}(x^*) = v$

Aplicando la fórmula de Taylor a la función $f \circ X : \mathbb{R} \rightarrow \mathbb{R}$ la cual es \mathcal{C}^σ diferenciable se obtiene que:

$$f(X(t)) - f(X(x^*)) = D_M f(x^*)(v) \cdot t + \frac{1}{2} D_M \left(D_M f(x^*) \right) (v) t^2 \quad (3.8a)$$

$$= D_M f(x^*)(v) \cdot t + \frac{1}{2} D_M V f(x^*)(v) t^2 \quad (3.8b)$$

$$= D_M f(x^*)(v) \cdot t + \frac{1}{2} H f(x^*)(v, v) t^2 \quad (3.8c)$$

Por el teorema 3.1, se tiene que el primer miembro en el lado derecho se anula, mientras que el segundo miembro debe ser no negativo (por continuidad); así:

$$H f(x^*)(v, v) \geq 0, \quad \forall v \in T_{x^*}M,$$

La desigualdad estricta es válida para todo $v \neq 0$ (pues f es función de Morse) de esta manera se ha probado que $H f(x^*)$ es definido positivo y que x^* es un punto crítico.

(\Leftarrow) Sea x^* un punto crítico tal que $H f(x^*)$ es definido positivo, el lado izquierdo de (3.8) sigue siendo estrictamente positivo para todo $t \neq 0$ suficientemente pequeño y todo $v \neq 0 \in T_{x^*}M$.

Por lo tanto,

$$f(X(t)) \geq f(x^*), \quad \forall t \neq 0 \text{ y } \forall v \neq 0 \in T_{x^*}M,$$

Lo que garantiza que x^* es un minimizador estricto.

Observación 3.1. *Las pruebas del teorema 3.1 y 3.2 utilizan explícitamente la estructura Riemanniana de M , ya que ellas involucran geodésicas. Ellas*

generalizan resultados similares para el caso sin restricciones el cual hace uso de la estructura Euclidiana de \mathbb{R}^n .

Estos resultados se cumplen, sin embargo sobre una variedad Riemanniana y pueden ser creadas usando un sistema de coordenada local y demostrando que son independiente de su escogencia.

3.1. Gradiente de una función sobre M

Definición 3.7. Sea γ la estructura Riemanniana en M , diremos que el gradiente de f en x sobre M , es el vector $\nabla_M^\gamma f(x) \in T_x M$ tal que:

$$\gamma_x(\nabla_M^\gamma f(x), v) = D_M f(x) \cdot v \quad (3.9)$$

Claramente el vector $\nabla_M^\gamma f(x)$ depende de la métrica Riemanniana. Nótese que $\nabla_M^\gamma f(x)$ es un campo vectorial sobre M , llamado **campo gradiente**.

Si usamos la métrica γ_x^z definida en (2.20) como el producto escalar en \mathbb{R}^{n-m} , se obtiene:

$$\nabla_M^z f(x) = Z_x^- \cdot g_x^z, \text{ con } g_x^z = Z_x^{-T} \cdot \nabla f(x) \in \mathbb{R}^{n-m}. \quad (3.10)$$

Mientras que si usamos la métrica γ_x^E definido en (2.19) como el producto escalar en \mathbb{R}^n se obtiene:

$$\nabla_M^E f(x) = Z_x^- \cdot g_x^E, \text{ con } g_x^E = (Z_x^{-T} \cdot Z_x^-)^{-1} \cdot g_x^z. \quad (3.11)$$

A g_x^Z y g_x^E le llamaremos gradiente reducido y gradiente Euclidiano respectivamente, los cuales coinciden obviamente si y sólo si $Z_x^{-T} \cdot Z_x^- = Id$

3.2. Método de máximo descenso a lo largo de Geodésicas

Recordemos que para encontrar un mínimo local en \mathbb{R}^n de una función continuamente diferenciable f , el método de máximo descenso genera a partir de una estimación inicial x_0 , aproximaciones sucesivas de acuerdo a la iteración

$$x^{k+1} = x^k + t_k \cdot P^k, \quad k = 0, 1, \dots \quad (3.12)$$

donde P^k es una dirección que cumple $\frac{\langle \nabla f(x^k), P \rangle}{\|P\|_n} < 0$, y es dada por:

$$P^k = -\nabla f(x^k), \quad (3.13)$$

es llamada dirección de máximo descenso, y el tamaño de paso t_k es un escalar positivo, seleccionado como el primer mínimo local en \mathbb{R}^+ de la función $j(t) = f(x^k + t_k \cdot P^k)$. Diremos que t_k es determinado por una búsqueda lineal exacta y la solución se denota por:

$$t_k = \text{Argmin}\{f(x^k + t_k \cdot P^k) / t \geq 0\}. \quad (3.14)$$

En este trabajo se generaliza este método; y así se puede encontrar un mínimo local de f en una variedad Riemanniana M ; en este caso, la dirección de máximo descenso para f en $x^k \in M$ es dada por:

$$P^k = -\nabla_M^\gamma f(x^k) \quad (3.15)$$

y cumple que $\frac{\gamma_{x^k}(\nabla_M^\gamma f(x^k), P)}{\|P\|_\gamma} < 0$, para cualquier $P \in T_{x^k}M$, además se

tiene que la norma Riemanniana es dada por $\left(\|P\|_\gamma = (\gamma_{x^k}(P, P))^{\frac{1}{2}}\right)$.

Una observación muy importante es que las geodésicas en M desempeñan el papel de las rectas en \mathbb{R}^n , por esta razón se define el método de máximo descenso a lo largo de geodésicas como la iteración.

$$x^{k+1} = \text{exp}_{x^k}(t_k \cdot P_k), \quad (3.16)$$

donde P^k es definido por (3.15) y el tamaño de paso t_k es determinado por una búsqueda de geodésica exacta

$$t_k = \text{Argmin}\{f(\exp_{x^k}(t_k \cdot P^k)) / t \in \mathbb{R}^+\} \quad (3.17)$$

Por otro lado, se define W_k como la componente conexa que contiene x^k al conjunto de nivel $\{x \in M / f(x) \leq f(x^k)\}$ y se enuncia el siguiente teorema.

Teorema 3.3. *Supongamos que f es continuamente diferenciable y que W_0 es compacto, entonces la sucesión $\{x^k\}$ construida por el método de máximo descenso a lo largo de geodésica (3.15), (3.16) y (3.17) está bien definido, o bien es finita y termina en un punto crítico, o es infinita y cada punto de acumulación es un punto crítico. Si los valores críticos de f son distintos, toda la sucesión $\{x^k\}$ converge a un punto crítico.*

Demostración: Si x^k es un punto crítico, entonces $P^k = 0$ y el algoritmo no genera nuevas aproximaciones.

Si x^k no es un punto crítico; entonces $\nabla_M^\gamma f(x^k) \neq 0$; de acá:

$$\gamma_{x^k}(\nabla_M^\gamma f(x^k), P^k) = -\gamma_{x^k}(\nabla_M^\gamma f(x^k), \nabla_M^\gamma f(x^k)) < 0 \quad (3.18)$$

Se denota la curva geodésica a partir de x^k con P^k tangente, por:

$$X(t) = \exp_{x^k}(t \cdot P^k) \quad (3.19)$$

Se define la función $j : \mathbb{R}^+ \rightarrow \mathbb{R}$, dada por:

$$j(t) = f(X(t)) \quad (3.20)$$

Sea $\bar{t} = \limsup(J)$, donde J es definido por $J = \{t > 0 / X(t) \text{ es definida y } j(t) < j(0) = f(x^k)\}$

Por (3.18) el conjunto J es no vacío, por otro lado, la componente conexa de $\{x \in M / f(x) \leq f(x^k)\}$, denotada para nosotros como W_k es un conjunto cerrado, ya que la condición que lo define es cerrado, además esta contenido en W_0 , pues $f(x_0) \leq f(x^1) \leq \dots \leq f(x^k)$.

Como W_0 es compacto, se tiene que W_k también es compacto, pues es un cerrado dentro de un compacto. Así la proposición 2.4 garantiza que si $\bar{t} = +\infty$, $X(t) \in W_k$ para cualquier $t \in [0, \infty)$ o si \bar{t} es finito, $f(X(\bar{t})) = f(x^k)$ y $X(t) \in W_k$ para cualquier $t \in [0, \bar{t}]$.

En ambos casos el tamaño de paso de la regla (3.17) se define así;

$$t_k \in (0, \bar{t})$$

Obsérvese que t_k satisface:

$$j'(t_k) = \gamma_{x^{k+1}} \left(\nabla_M^\gamma f(x^{k+1}), \pi_0^{t_k}(P^k) \right) \quad (3.21a)$$

$$= \gamma_{x^k} \left(\pi_{t_k}^0(\nabla_M^\gamma f(x^{k+1})), P^k \right) = 0 \quad (3.21b)$$

Donde $\pi_0^{t_k}(P^k)$ es el vector en $T_x^{k+1}M$ obtenido por la traslación paralela a lo largo de la curva $X(\cdot)$, desde x^k hasta x^{k+1} ; y $\pi_{t_k}^0(\nabla_M^\gamma f(x^{k+1}))$ es el vector en $T_x^k M$ obtenido por la traslación paralela de la curva $X^{-1}(\cdot)$ desde x^{k+1} hasta x^k .

Se define la aplicación $g : [0, t_k] \rightarrow \mathbb{R}$ por:

$$g(t) = \gamma_{x^k} \left(\Pi_t^0(\nabla_M^\gamma f(X(t))), P^k \right) - \alpha \gamma_{x^k} \left(\nabla_M^\gamma f(x^k), P^k \right) \quad (3.22)$$

para $\alpha \in (0, \frac{1}{2})$.

Se quiere ver que existe $\hat{t} \in (0, t_k)$ de tal manera que $g(\hat{t}) = 0$; para ello se estudiara $g(0)$ y $g(t_k)$.

$$\begin{aligned} g(0) &= \gamma_{x^k} \left(\nabla_M^\gamma f(x^k), P^k \right) - \alpha \gamma_{x^k} \left(\nabla_M^\gamma f(x^k), P^k \right) \\ &= (1 - \alpha) \gamma_{x^k} \left(\nabla_M^\gamma f(x^k), P^k \right) < 0 \quad \text{por} \quad (3.18) \end{aligned}$$

$$\begin{aligned}
g(t_k) &= \gamma_{x^k} \left(\Pi_{t_k}^0 \left(\nabla_M^\gamma f(x^{k+1}), P^k \right) \right) - \alpha \gamma_{x^k} \left(\nabla_M^\gamma f(x^k), P^k \right) \\
&= -\alpha \gamma_{x^k} \left(\nabla_M^\gamma f(x^k), P^k \right) > 0 \quad \text{por (3.21)}
\end{aligned}$$

Así, existe un primer \hat{t} para el cual $g(\hat{t}) = 0$. De esta manera se tiene que para cualquier $t \in [0, \hat{t})$ se cumple que:

$$\gamma_{x^k} \left(\Pi_t^0 \left(\nabla_M^\gamma f(X(t)), P^k \right) \right) < \alpha \gamma_{x^k} \left(\nabla_M^\gamma f(x^k), P^k \right)$$

Luego; aplicando el teorema del valor medio se obtiene que:

$$\begin{aligned}
f(X(\hat{t})) - f(x^k) &= f(X(\hat{t})) - f(X(0)) \\
&= f'(X(\tilde{t}))X'(\tilde{t})(\hat{t} - 0) \\
&= \hat{t} j'(\tilde{t}) \\
&= \hat{t} \gamma_{x^k} \left(\pi_{t_k}^0 \left(\nabla_M^\gamma f(X(\tilde{t})), P^k \right) \right) \\
&< \hat{t} \alpha \gamma_{x^k} \left(\nabla_M^\gamma f(x^k), P^k \right) \\
&= -\hat{t} \alpha \|\nabla_M^\gamma f(x^k), P^k\|_{\gamma_{x^*}}^2 < 0
\end{aligned}$$

Por otro lado; como $f(x^{k+1}) < f(X(\hat{t}))$, se tiene que:

$$f(x^{k+1}) - f(x^k) < f(X(\hat{t})) - f(x^k) < 0$$

Así, $f(x^{k+1}) < f(x^k)$ y por tanto la sucesión $\{f(x^k)\}$ es monótona decreciente y esta acotada por debajo, ya que la función f alcanza este mínimo en el compacto.

De acá $\{f(x^k)\}$ converge a un limite; digamos $f(x^k) \rightarrow f(x^*)$, por otro lado como $\{x^k\}$ esta en el compacto W_0 se puede considerar una subsucesión $\{x^{k_i}\}$ de $\{x^k\}$ que converge a $x^* \in W_0$.

Suponer que x^* no es un punto crítico, es decir que $\|\nabla_M^\gamma f(x^k)\|_{\gamma_{x^*}} = \delta > 0$

La continuidad de la métrica Riemanniana γ y del campo gradiente $\nabla_M^\gamma f$ implica que:

$$\|\nabla_M^\gamma f(x^{k_i})\|_\gamma \geq \frac{\delta}{2}, \text{ para todo } i > I$$

De acá;

$$\begin{aligned} f(x^{k_{i+1}}) - f(x^{k_i}) &< -\alpha \hat{t} \|\nabla_M^\gamma f(x^k)\|_\gamma^2 \\ &\leq -\alpha \hat{t} \frac{\delta^2}{4} \end{aligned}$$

Tomando el limite en la desigualdad anterior, se obtiene $0 < -(-\alpha \hat{t} \frac{\delta^2}{4})$, lo cual es una contradicción.

Por lo tanto, x^* es un punto crítico.

Por último, se supone que x^* y x^{**} son dos puntos de acumulación de la sucesión $\{x^k\}$ en W_0 los cuales son distintos, y x^* y x^{**} son puntos críticos de la función f .

Como $\{f(x^k)\}$ converge, se debe tener que.

$$f(x^*) = f(x^{**})$$

Lo cual es imposible porque los valores críticos de f son distintos.

Por lo tanto, toda la sucesión $\{x^k\}$ converge a un punto crítico.

3.3. Método de Newton a lo largo de Geodésicas

Sea x^* un mínimo local sin restricciones de una función f diferenciable \mathcal{C}^σ , $\sigma \geq 3$, tal que la forma Hessiana $\nabla^2 f(x)$ es definida positivo y U_* es una vecindad de x^* tal que para todo $x \in U_*$, $\nabla^2 f(x)$ es definido positivo.

A partir de $x^0 \in U_*$, el método de Newton para la minimización sin restricciones genera una sucesión de aproximaciones de x^* de acuerdo a la iteración

$$x^{k+1} = x^k - (F_k)^{-1} \cdot \nabla f(x^k) \quad (3.23)$$

donde F_k denota la matriz simétrica no singular definida por el Hessiano $\nabla^2 f(x^k)$; U_* es siempre lo suficientemente pequeña para que las iteraciones permanezcan en U^* , y la sucesión $\{x^k\}$ converge a x^* con una velocidad de convergencia de segundo orden.

La extensión del método de Newton a la minimización de una variedad M , presenta mayor dificultad ya que no es posible definir la forma Hessiana de f sobre M fuera de un punto crítico.

Sin embargo, podemos definir en un punto no crítico, una forma cuadrática en el espacio tangente $T_x M$ mediante la aplicación de la estructura Riemanniana de la variedad M .

Sea $v \in T_x M$, y considérese la curva geodésica $X(\cdot)$ que parte desde x con vector tangente v dada por:

$$X(t) = \exp_x(t.v) \quad (3.24)$$

Se procederá como para la definición del Hessiano, pero ahora sea \mathbf{V} el único campo vectorial paralelo a lo largo de la curva $X(\cdot)$ tal que $\mathbf{V}(0) = v$

Se define la función $j : \mathbb{R}^+ \rightarrow \mathbb{R}$, diferenciable \mathcal{C}^σ por

$$j(t) = f(X(t)) \quad (3.25)$$

Por la definición del gradiente de f en M con respecto a la métrica Riemanniana γ , se tiene que:

$$j'(t) = \gamma_{X(t)}(\nabla_M f X(t), \mathbf{V}(t)) \quad y$$

$$\begin{aligned} j''(t) &= \gamma_{X(t)} \left(\frac{D\nabla_M f(X(t))}{dt}, \mathbf{V}(t) \right) + \gamma_{X(t)} \left(\nabla_M f(X(t)), \frac{D\mathbf{V}(t)}{dt} \right) \\ &= \gamma_{X(t)} \left(\frac{D\nabla_M f(X(t))}{dt}, \mathbf{V}(t) \right); \end{aligned}$$

ya que \mathbf{V} es un campo vectorial paralelo; luego se define la forma cuadrática $F(x) : T_x M \times T_x M \rightarrow \mathbb{R}$ por:

$$F(x)(v, v) = j''(0) = \gamma_x(D_v \nabla_M f, v), \quad \forall v \in T_x M \quad (3.26)$$

la regularidad de la solución (3.24) sobre la ecuación diferencial (2.29) y la definición de geodésica con respecto a las condiciones iniciales implican que F es $\mathcal{C}^{\sigma-2}$ diferenciable con respecto a x .

Nótese que, para un punto crítico x^* , $F(x^*)$ coincide con la forma Hessiana definida por (3.7)

$$F(x^*)(v, v) = Hf(x^*)(v, v), \quad \forall v \in T_{x^*} M \quad (3.27)$$

Se define la aplicación $F_x : T_x M \rightarrow T_x M$ por:

$$F(x)(v, v) = \gamma_x(F_x v, v), \quad \forall v \in T_{x^*} M \quad (3.28)$$

Note que F_x es simplemente la conexión Riemanniana del campo gradiente de x , dada por $F_x = \tau_x(\nabla_M f)$. Este se trata de un isomorfismo lineal del $T_x M$ y es autoadjunta (con respecto a la métrica Riemanniana γ_x .) Se denota por (F_x^{-1}) su inverso definido en $T_x M$.

Sea x^* un mínimo local no degenerado de f en M , por el teorema 3.2 el Hessiano $Hf(x^*)$ es definido positivo, además la identidad 3.27 junto con la continuidad de F , demuestra que existe una vecindad $U_* \cap M$ de x^* en M tal que $F(x)$ es definida para todo $x \in U_* \cap M$.

A partir de $x^0 \in U_* \cap M$, el método de Newton a lo largo de geodésicas genera la sucesión de aproximaciones sucesivas $\{x^k\}$ de acuerdo con la iteración

$$x^{k+1} = \exp_{x^k} p^k \quad (3.29)$$

donde la dirección de Newton $p^k \in T_{x^k}M$ es dada por:

$$p^k = -(F_k)^{-1} \nabla_M f(x^k) \quad (3.30)$$

El método está bien definido ya que $U_* \cap M$ es siempre escogida lo suficientemente pequeña como para que la sucesión $\{x^k\}$ pertenezca a la vecindad y converja a x^* cuadráticamente.

Teorema 3.4. *Supóngase que f es una función de Morse \mathcal{C}^σ sobre la variedad Riemanniana M , sea x^* un mínimo local de f en M , entonces existe una vecindad $U_* \cap M$ tal que, si $x^0 \in U_* \cap M$, el método de Newton a lo largo de geodésicas genera una sucesión $\{x^k\}$ en $U_* \cap M$ que converge a x^* , y existe una constante k tal que:*

$$\delta(x^{k+1}, x^k) \leq k (\delta(x^k, x^*))^2, \quad \forall k = 0, 1, \dots \quad (3.31)$$

donde $\delta(\cdot, \cdot)$ representa la distancia Riemanniana.

Demostración: Dado $x \in M$, y $v \in T_x M$, se define la función $\varphi = \mathbb{R}^+ \rightarrow \mathbb{R}$ diferenciable $\mathcal{C}^{\sigma-1}$ por:

$$\varphi(t) = \gamma_{X(t)}(\nabla_M f(X(t)), \mathbf{W}(t)), \quad (3.32)$$

donde $\mathbf{W}(t)$ es el campo vectorial paralelo a lo largo de la curva $X(\cdot)$ a dada por 3.24. Nótese que:

$$\varphi(t) = \gamma_{X(t)}(F_{X(t)} \mathbf{V}(t), \mathbf{W}(t)), \quad (3.33)$$

donde $\mathbf{V}(t)$ es el campo vectorial paralelo a lo largo de $X(\cdot)$ tal que $\mathbf{V}(0) = v$.

Considere la curva geodesia $X(\cdot)$ a partir de x^* y tangente a p^k dado por (3.29), ahora bien, si se aplica la fórmula de Taylor para la función $\varphi(t)$ se obtiene que:

$$\varphi(1) = \varphi(0) + \varphi'(0) + \int_0^1 \gamma_{x^{k+1}} (\varphi'(t) - \varphi'(0)) dt \quad (3.34)$$

De acá;

$$\gamma_{x^{k+1}} (\nabla_M f(x^{k+1}), \mathbf{W}(1)) = \gamma_{x^k} (\nabla_M f(x^k), \mathbf{W}(0)) + \int_0^1 \gamma_{X(t)} (F_{X(t)} \mathbf{V}(t), \mathbf{W}(t)) dt$$

Como F_x es $\mathcal{C}^{\sigma-2}$ diferenciable en $U_* \cap M$, se obtiene la siguiente estimación:

$$\begin{aligned} \|\nabla_M f(x^{k+1})\|_\gamma &\leq \frac{L}{2} \|p^k\|_\gamma^2 \\ &\leq \left(\frac{L}{2} m^2\right) \|\nabla_M f(x^*)\|_\gamma^2 \end{aligned} \quad (3.35)$$

si, $\|(F_x)^{-1}\| \leq \frac{1}{m}$, $\forall x \in U_* \cap M$

Considérese ahora, la curva geodésica que minimiza a x^{k+1} y x^k , la cual se sabe que existe, pues me lo garantiza la proposición 2.4, de acá se puede encontrar $q^{k+1} \in T_{k+1}M$ tal que la curva $X(t) = \exp_{x^{k+1}}(tp^{k+1})$ satisface:

$$X(0) = x^{k+1} \text{ y } X(1) = x^*$$

Para esta escogencia de $X(\cdot)$, la fórmula de Taylor para la función φ arroja:

$$\begin{aligned} \gamma_{x^*} (\nabla_M f(x^*), \mathbf{W}(1)) = 0 &= \gamma_{x^{k+1}} (\nabla_M f(x^{k+1}), \mathbf{W}(0)) + \\ &\int_0^1 \gamma_{X(t)} (F_{X(t)} \mathbf{V}(t), \mathbf{W}(t)) dt \end{aligned} \quad (3.36)$$

Asumiendo que:

$$m\|v\|^2 \gamma_x \leq \gamma_x (F_x v, v) \leq N\|v\|_{\gamma_x}^2, \text{ para todo } x \in U_* \cap M, \text{ } x \in T_x M$$

Se obtiene la siguiente estimación:

$$m\|q^{k+1}\|_\gamma \leq \|\nabla_M f(x^{k+1})\|_\gamma \leq N\|q^{k+1}\|_\gamma \quad (3.37)$$

Combinando (3.35) y (3.37) para k y $k - 1$, y notando que $\delta(x^{k+1}, x^*) = \|q^{k+1}\|_\gamma$, se obtiene:

$$\delta(x^0, x^*) \leq \left(\frac{LN}{2m^3} \right) (\delta(x^k, x^*))^2,$$

Lo que demuestra la convergencia cuadrática de $\{X^k\}$ a x^* , siempre que

$$\delta(x^0, x^*) \leq \frac{2m^3}{LN}$$

Bibliografía

- [1] M. Do Carmo. *Differential Geometry of Curves and Surfaces*. Prentice Hall, New Jersey, 1976.
- [2] D. Gabay. Minimizing a differentiable function over a differential manifold. *Journal of optimization theory and applications*, pages 1–25, 1982.
- [3] E.L. Lima. *Análise Real*, volume 2. IMPA, Rio de Janeiro, 2004.
- [4] J. Munkres. *Analysis on Manifolds*. Allan M. Wylde, 1930.