Notas de Cálculo Diferencial en varias variables.

Luis Manuel Reyes de la Luz

Agosto 2024

Índice general

Parte I

Cálculo Diferencial en varias variables.

Capítulo 1

Espacio Euclideo n dimensional

En esta sección, denotamos $\mathbb{R}$ como el campo de los números naturales. Así como en Cálculo en una variable el campo $\mathbb{R}$ es fundamental, en varias variables seguiremos usando dicho campo y sus propiedades fundamentales, las cuales serán mencionados aquí como axiomas para estas notas.

Axioma 1.0.1 — Del supremo. Sea $X \subseteq \mathbb{R}$ no vacío, tal que existe una $M \in \mathbb{R}$ con la propiedad de que

$$x \leq M, \ \forall x \in X$$

entonces existe el supremo de X.

Axioma 1.0.2 — De densidad. Las siguientes afirmaciones son válidas:

1. Sean $r,t \in \mathbb{Q}$ dos números racionales tales que $r < t$ entonces existe un irracional $x \in \mathbb{R} - \mathbb{Q}$ tal que $r < x < t$.
2. Sean $r,t \in \mathbb{R} - \mathbb{Q}$ dos números irracionales tales que $r < t$ entonces existe un racional $x \in \mathbb{Q}$ tal que $r < x < t$.

Axioma 1.0.3 — Propiedad Arquimediana. El conjunto $\mathbb{N} \subseteq \mathbb{R}$ no es acotado, es decir no existe $r \in \mathbb{R}$ tal que $r > n$ para toda $n \in \mathbb{N}$.

Una consecuencia de este axioma es el siguiente resultado que se puede ver como la versión práctica de la propiedad arquimediana.

Proposición 1.0.4 Para todo par de números reales x,y con $y > 0$, existe un entero positivo n tal que $nx > y$

1.1. El espacio vectorial euclídeo n-dimensional.

Definición 1.1.1 — El espacio euclídeo n-dimensional estándar. Definimos el espacio euclídeo n-dimensional estándar como el conjunto

$$\mathbb{R}^n := \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mid x_1, \cdots, x_n \in \mathbb{R} \right\}$$

dotado de las siguientes operaciones;

1. Suma vectorial: $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} := \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}$.

2. Multiplicación escalar: $r \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} rx_1 \\ \vdots \\ rx_n \end{pmatrix}$.

Un resultado conocido del algebra es el siguiente:

Proposición 1.1.1 El espacio euclídeo n-dimensional $\mathbb{R}^n$ es un $\mathbb{R}$-espacio vectorial de dimensión n.

Hasta el momento se ha presentado al espacio euclídeo n-dimensional con su forma algebraica, sin embargo la razón de la palabra "euclídeo" tiene una razón geométrica que se explicará a continuación.

El concepto de vector. Tradicionalmente los elementos de $\mathbb{R}^n$ se les denominan vectores, pero originalmente, un vector es un concepto geométrico/físico que representa una magnitud física que contiene la distancia y su dirección. Geométricamente para realizar esto, un vector necesita una noción de "inicio" y de "fin", de manera formal se hace lo siguiente

R Dado $\vec{p}_i, \vec{p}_f$ dos puntos en $\mathbb{R}^n$, si yo quiero representar un vector que inicia en $\vec{p}_i$ y termina en $\vec{p}_f$ dibujare el segmento dirigido $\overline{p_i p_f}$. El conjunto de vectores que inician en $\vec{p}_i$ se va a denotar como $T_{p_i}\mathbb{R}^n$, formalmente:

1. $T_{p_i}\mathbb{R}^n$ es el conjunto $\mathbb{R}^n \times \{p_i\}$ y por abuso de notación cada elemento $(a, p_i) \in \mathbb{R}^n \times \{p_i\}$ se escribe simplemente como a.
2. Un vector que inicia en p_i y termina en p_f es el elemento $a \in T_{p_i}\mathbb{R}^n$ definido como $a := p_f - p_i$. Conversamente, cada elemento $a \in T_{p_i}\mathbb{R}^n$ se puede interpretar como un vector que inicia en p_i y termina en $p_i + a$.

Esta manera de ver, respeta un principio fundamental del espacio euclídeo, es **homogéneo**, esto significa coloquialmente que cada punto del espacio puede ser vista como el centro de observación de todo, formalmente se resumen en la siguiente afirmación sin demsotración.

Proposición 1.1.2 Para cada $p \in \mathbb{R}^n$ definimos $T_p \colon \mathbb{R}^n \to \mathbb{R}^n$ como $T_p(x) = x - p$, y denotemos $G := \{T_p \mid p \in \mathbb{R}^n\}$ entonces:

1. G es un grupo abeliano.
2. $T_0 = \mathrm{Id}_{\mathbb{R}^n}$.
3. $T_{p+q} = T_p \circ T_q$.
4. T_p es biyectiva y $T_p(p) = 0$.

Con la afirmación anterior, se tiene que T_p se le conoce como **la traslación centrada en** p. Así, cada $T_p\mathbb{R}^n$, conjuntistamente es como $\mathbb{R}^n$, algebraicamente es como $\mathbb{R}^n$, pero visualmente $T_p\mathbb{R}^n$ es formalmente tu centro de observación en el punto p y T_p es el cambio de origen. En particular $\mathbb{R}^n$, con esta perspectiva(denominada afín), se puede pensar como $T_0\mathbb{R}^n$, es decir todo punto $p \in \mathbb{R}^n$ se representa como un vector con origen 0. De esta manera, muchos objetos geométricos se pueden construir pensando primero en el origen y luego el caso general.

Rectas Desde la perspectiva vectorial (Algebra lineal + Geometría afín) una recta en el origen con dirección $u \in \mathbb{R}^n$ es el conjunto:

$$\mathscr{L}_u := \{x \in \mathbb{R}^n \mid x = \lambda u,\ \lambda \in \mathbb{R}\}$$

¿Y como se representa una recta en general? Sea $\mathscr{L}$ una recta que pasa por p y q, si aplicamos la traslación $T_p \colon \mathbb{R}^n \to \mathbb{R}^n$ entonces su imagen $\mathscr{S} := T_p(\mathscr{L})$ es una recta en el origen con dirección $q - p$. Así todo punto $s \in \mathscr{S}$ por un lado es de la forma

$$s = \lambda(p - q)$$

y por otro lado es de la forma

$$s = T_p(x) = x - p, \text{ para algún único } x \in \mathscr{L}$$

es decir

$$x = p + \lambda(p - q)$$

obteniendo la ecuación vectorial de una recta $\mathscr{L}$ que pasa por p y q en $\mathbb{R}^n$. Con el mismo principio que se aplicó en las rectas, podemos definir el concepto de plano, hiperplano, etc.

Definición 1.1.2 — Hiperplano. Sea $m < n$ natural. Un subconjunto $H \subseteq \mathbb{R}^n$ es denominado hiperplano afín de dimensión m si existen $u_1, \cdots, u_m \in \mathbb{R}^n$ linealmente independientes (denominado los vectores directrices de H) tal que

$$H = \{x \in \mathbb{R}^n \mid x = p + \lambda_1 u_1 + \cdots + \lambda_m u_m, \ \lambda_1, \cdots, \lambda_m \in \mathbb{R}\}$$

Como convención, cuando $m = 1$ entonces se denomina recta, cuando $m = 2$ entonces se denomina plano y cuando $m = n - 1$ se denomina simplemente hiperplano. Describir a los hiperplanos como un sistema de ecuaciones, es una tarea para álgebra lineal o geometría analítica.

1.2. Geometría del producto escalar.

Para hablar de más objetos geométricos desde la perspectiva de la geometría euclidiana, necesitamos una herramienta más para el espacio vectorial euclídeo n-dimensional, y es denominado como producto interior o producto escalar euclidiano.

Definición 1.2.1 — Producto interior. Definimos el producto interior o producto escalar euclidiano de $\mathbb{R}^n$ como una función $\langle , \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ como

$$\left\langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right\rangle := x_1 y_1 + x_2 y_2 + \cdots + x_n y_n.$$

Algunas propiedades importantes son

Proposición 1.2.1 Las siguientes afirmaciones son válidas para el producto interior:

1. Es bilineal, es decir, para toda $p, q, a, b \in \mathbb{R}^n$ y toda $\lambda, \xi \in \mathbb{R}$ se tiene que

$$\langle p + \lambda q, a + \xi b \rangle = \langle p, a \rangle + \lambda \langle q, a \rangle + \xi \langle p, b \rangle + \lambda \xi \langle q, b \rangle$$

2. Es no degenerada, es decir si $\langle x, y \rangle = 0$ para toda $y \in \mathbb{R}^n$ entonces $x = 0$.
3. Es positiva definida, es decir para toda $x \in \mathbb{R}^n$ entonces $\langle x, x \rangle \geq 0$.
4. Satisface la desigualdad de Cauchy-Schwarz, para toda $x, y \in \mathbb{R}^n$ se tiene que

$$|\langle x, y \rangle|^2 \leq \langle x, x \rangle \langle y, y \rangle$$

Denotamos $\|x\| := \sqrt{\langle x, x \rangle}$ y lo denominamos como **norma euclidiana.**

Definición 1.2.2 — Angulo entre vectores. Dado $x, y \in \mathbb{R}^n$ el ángulo entre ellos es un valor $\theta \in [0, \pi]$ tal que

$$\cos\theta = \frac{\langle x, y \rangle}{\|x\| \, \|y\|}$$

Además decimos que los vectores x, y son ortogonales si $\theta = \pi/2$.

Proposición 1.2.2 Son equivalentes:

1. $\langle x,y\rangle = 0$.
2. Los vectores x,y son ortogonales.

Definición 1.2.3 — La n-esfera. Definimos la n-esfera centrada en $p \in \mathbb{R}^n$ y con radio $r \in \mathbb{R}^+$ como el conjunto

$$S_r(p) := \{x \in \mathbb{R}^n \mid \|x-p\| = r\}$$

Además:

1. Cuando $p = 0$ lo denotamos simplemente como S_r y se conoce como n-esfera de radio r.
2. Cuando $p = 0$ y $r = 1$ lo denotamos como $\mathbb{S}^n$ y se conoce como n-esfera unitaria.

1.3. Normas y Distancias.

Las normas y distancias son el primer punto de estudio para poder comprender el concepto de funciones continuas y las diferentes de nociones de derivadas en varias variables.

Definición 1.3.1 — Espacio normado. Una norma en $\mathbb{R}^n$ es una función $\|\|: \mathbb{R}^n \to \mathbb{R}$ tal que

1. $\|x\| = 0$ si y sólo si $x = 0$.
2. $\|rx\| = |r|\,\|x\|$, con $r \in \mathbb{R}$ y $x \in \mathbb{R}^n$.
3. $\|x+y\| \leq \|x\| + \|y\|$.

■ **Ejemplo 1.1** Sea $\langle,\rangle$ como en 1.2.1, entonces $\|x\| := \sqrt{\langle x,x\rangle}$ es una norma, denominada como **norma euclidiana**. ■

■ **Ejemplo 1.2** Sea $\beta = \{e_1,\cdots,e_n\}$ una base de $\mathbb{R}^n$, definimos

$$\|x\|_{max} := \max\{|a_i|\}_{i=1}^n, \text{ donde } x = \sum_{i=1}^n a_i e_i$$

entonces $\|x\|_{max}$ es una norma. ■

■ **Ejemplo 1.3** Sea $\beta = \{e_1,\cdots,e_n\}$ una base de $\mathbb{R}^n$, definimos

$$\|x\|_1 := \sum_{i=1}^n |a_i|, \text{ donde } x = \sum_{i=1}^n a_i e_i$$

■

■ **Ejemplo 1.4** Sea $p \in [1,\infty)$ y $\beta = \{e_1,\cdots,e_n\}$ una base de $\mathbb{R}^n$, definimos

$$\|x\|_p := \left(\sum_{i=1}^n |a_i|^p\right)^{1/p}$$

■

■ **Ejemplo 1.5** Sea $A \in \mathbb{M}_{n\times n}(\mathbb{R})$ una matriz invertible, entonces la función

$$\|x\|_A := \|Ax\|$$

es una norma. ■

■ **Ejemplo 1.6** Sea $\{\|\|_i : \mathbb{R}^{n_i} \to \mathbb{R}\}_{i=1}^s$ una familia de normas, sea $n = \sum_{i=1}^s n_i$, entonces para cada $x \in \mathbb{R}^n$, definimos

$$\|x\| := \sum_{i=1}^n \|x_i\|_i$$

entonces es una norma. ■

Ejercicio 1.1 Sea $(\mathbb{R}^n, \|\|)$ un espacio normado, prueba que para toda $x,y \in \mathbb{R}^n$ se tiene

$$|\|x\| - \|y\|| \leq \|x-y\|$$

■

Definición 1.3.2 — Métrica. Sea X un conjunto, una métrica es una función $d \times X : X \to \mathbb{R}$ tal que

1. $d(x,y) \geq 0$.
2. $d(x,y) = 0$ si y sólo si $x = y$.
3. $d(x,y) = d(y,x)$.
4. $d(x,y) \leq d(x,z) + d(z,y)$.

A la pareja (X,d) se le conoce como espacio métrico.

Proposición 1.3.1 Sea $(\mathbb{R}^n, \|\|)$ un espacio normado, si definimos $d(x,y) := \|x-y\|$ entonces $(\mathbb{R}^n, d)$ es un espacio métrico.

■ **Ejemplo 1.7** Sea X un conjunto no vacio, definimos

$$d_{dis}(x,y) := \begin{cases} 0 & x = y \\ 1 & x \neq y \end{cases}$$

entonces (X, d_{disc}) es un espacio métrico. ■

A continuación mostraremos una construcción de espacio métrico que sera muy útil en posteriores temas.

Definición 1.3.3 — Funciones acotadas. Sea A un conjunto no vacio, una función $f : A \to \mathbb{R}^n$ es acotada si existe un $M \in \mathbb{R}^+$ tal que

$$\|f(x)\| \leq M, \ \forall x \in A$$

Proposición 1.3.2 Sea A un conjunto no vacio, definimos $V := B(A, \mathbb{R}^n)$ el espacio de las funciones acotadas y sea $\|\|_0$ una norma de $\mathbb{R}^n$ entonces:

1. V es un espacio vectorial real.
2. Si definimos $\|f\|_\infty := \sup\{\|f(x)\| \, |x \in A\}$ entonces $\|\|_\infty$ es una norma.

Demostración. Claramente $0 \in V$, ahora veamos que V es cerrado bajo sumas y producto escalar, tomemos $f, g \in V$ entonces existen $M_f, M_g \in \mathbb{R}^+$ tal que

$$\|f(x)\| \leq M_f, \ \|g(x)\| \leq M_g, \ \forall x \in A$$

y si tomamos $\alpha, \beta \in \mathbb{R}$ entonces

$$\|\alpha f(x) + \beta g(x)\| \leq |\alpha| M_f + |\beta| M_g, \ \forall x \in A$$

entonces $\alpha f + \beta g \in V$, por lo tanto V es espacio vectorial. Para el segundo punto, supongamos que $\|f\|_\infty = 0$ esto implica que para toda $x \in A$, $\|f(x)\| \leq 0$, entonces $\|f(x)\| = 0$ obteniendo $f(x) = 0$ por tanto $f = 0$. Conversamente, es claro que $\|0\|_\infty = 0$. Ahora dado $\alpha \in \mathbb{R}$ y $f \in V$, notemos que para toda $x \in V$ se tiene

$$\|\alpha f(x)\| \leq |\alpha| \, \|f\|_\infty$$

Ahora si $M \in \mathbb{R}^+$ tal que $\|\alpha f(x)\| \leq M$ entonces $\|f(x)\| \leq M/|\alpha|$, usando la definición de $\|f\|_\infty$ entonces

$$\|f\|_\infty \leq M/|\alpha|$$

por tanto

$$|\alpha| \, \|f\|_\infty \leq M$$

concluyendo, por la definición de supremo, obtenemos

$$\|\alpha f\|_\infty = |\alpha| \, \|f\|_\infty$$

Para la desigualdad triangular, notemos que para cada $x \in A$ entonces

$$\|f(x)+g(x)\| \leq \|f(x)\| + \|g(x)\| \leq \|f\|_\infty + \|g\|_\infty$$

Usando la definición del supremo $\|f+g\|_\infty$ obtenemos

$$\|f+g\|_\infty \leq \|f\|_\infty + \|g\|_\infty$$

■

Ahora estudiaremos algunas relaciones interesantes entre las normas de $\mathbb{R}^n$.

Definición 1.3.4 — Equivalencia de normas. Decimos que dos normas $\|\|_1, \|\|_2$ son equivalentes si existen $c_1, c_2 \in \mathbb{R}$ tal que

$$c_1 \|x\|_2 \leq \|x\|_1 \leq c_2 \|x\|_2$$

Proposición 1.3.3 Toda norma en $\mathbb{R}^n$ es equivalente.

Observemos que, de la definición de producto cartesiano (conjuntista) $\mathbb{R}^n = Fun(n,\mathbb{R}) = B(n,\mathbb{R})$ y la norma de (1.3.2) es descrito como

$$\|(x_1, \cdots, x_n)\|_\infty = \text{máx}\{|x_i| \mid i = 1, \cdots, n\}$$

Demostración. Vamos a probar que toda norma $\|\|$ es equivalente a la norma supremo $\|\|_\infty$. Tomemos $x = (x_1, \cdots, x_n) \in \mathbb{R}^n$ entonces por la desigualdad triangular

$$\|x\| = \left\| \sum_{k=1}^{n} x_i e_i \right\| \leq \sum_{k=1}^{n} |x_i| \, \|e_i\| \leq \left(\sum_{k=1}^{n} \|e_i\| \right) \|x\|_\infty$$

Por otro lado, notemos que ■

1.4. Introducción a la topología de los números reales.

Para profundizar un poco más el significado de normas equivalentes, exploraremos la noción de topología. Esta noción sirve para comprender los fenómenos de "proximidad", "vecindad" que se presentan naturalmente en cálculo y no depende explícitamente de la noción de métrica, para ello primero generalizaremos el concepto de intervalo abierto y cerrado. En esta sección $\|\|$ denota una norma fija arbitraria en $\mathbb{R}^n$.

Definición 1.4.1 — Disco abierto y cerrado. Dado $x \in \mathbb{R}^n$, $r > 0$, definimos:

1. El disco abierto centrado en x con radio r como

$$B_r(x) := \{y \in \mathbb{R}^n \mid \|x-y\| < r\}$$

2. El disco cerrado centrado en x con radio r como

$$\overline{B}_r(x) := \{y \in \mathbb{R}^n \mid \|x-y\| \leq r\}$$

Ejercicio 1.2 ¿Será cierto que $B_r(x)$ y $\overline{B}_r(x)$ son conjuntos numerables? ■

Ejercicio 1.3 Muestra las siguientes propiedades:

1. Para cada $x \in \mathbb{R}^n$,
$$B_r(x) \cap \overline{B}_r(x) = S_r(x)$$
2. Si $B_{r_1}(x_1) \cap B_{r_2}(x_2) \neq \emptyset$ entonces para todo $x \in \mathbb{R}^n$ existe $r > 0$ tal que
$$B_r(x) \subseteq B_{r_1}(x_1) \cap B_{r_2}(x_2)$$

■

Definición 1.4.2 — Conjuntos abiertos y cerrados. Un subconjunto $S \subseteq \mathbb{R}^n$ es abierto si para todo $x \in S$ existe una $r > 0$ tal que $B_r(x) \subseteq S$. Un subconjunto $F \subseteq \mathbb{R}^n$ es cerrado si $\mathbb{R}^n - F$ es abierto. Además, definimos $\emptyset$ como un conjunto abierto.

■ **Ejemplo 1.8** Definimos $R = [-1,1]^n$ entonces este conjunto es cerrado. Para ver esto, consideremos $X = \mathbb{R}^n - R$, entonces cada punto $(x_1, \cdots, x_n) \in X$ cumple la propiedad de que existe $i = 1, \cdots, n$ tal que
$$|x_i| > 1$$
Con esto en mente, fijemos $x = (x_1, \cdots, x_n) \in X$ con $|x_i| > 1$, definimos $r := |x_i - 1|$, usando 1.1 tenemos
$$r \geq ||x_i| - 1| > |1 - 1| = 0$$
entonces $B_{r/2}(x) \subseteq X$, mostrando que X es abierto y, por tanto, R es cerrado. ■

■ **Ejemplo 1.9** Todo disco abierto es un conjunto abierto. Supongamos que $D = B_r(x)$, y sea $y \in D$, definimos $s := d(x,y)$, por definición $s < r$, y sea $m := \text{mín}\{s, r-s\}$ entonces $B_{m/2}(y) \subseteq D$, obteniendo lo deseado. ■

■ **Ejemplo 1.10** Todo disco cerrado es cerrado. Supongamos que $F := \overline{B}_r(x)$ y sea $X := \mathbb{R}^n - F$, fijemos $y \in X$ entonces
$$\|x - y\| > r$$
sea $s := \|x - y\| - r$, entonces $B_{s/2}(y) \subseteq X$. ■

■ **Ejemplo 1.11** Si $\mathscr{F} := \{B_r(x)\}$ es una familia de discos abiertos, entonces $\bigcup \mathscr{F}$ es un abierto. ■

■ **Ejemplo 1.12** Sea $S := [0,1)^n$ entonces este conjunto no es abierto ni cerrado. Primero veamos que S no es abierto, tomemos $0 \in S$, entonces para cada $r > 0$ se tiene que $B_r(0) \cap S \neq \emptyset$, esto implica que $B_r(0) \not\subset S$. Ahora veamos que S no es cerrado, para ello veamos que $X = \mathbb{R}^n - S$ no es abierto, notemos que $x_1 := (1, \cdots, 1) \in X$, entonces para cada $r > 0$, se tiene que $B_r(x_1) \cap S \neq \emptyset$, esto implica que $B_r(x_1) \not\subset X$, por tanto, X no es cerrado. ■

Proposición 1.4.1 Las siguientes afirmaciones son válidas:

1. Los conjuntos $\emptyset, \mathbb{R}^n$ son abiertos.
2. Si $\mathscr{F}$ es una familia de conjuntos abiertos entonces $\bigcup \mathscr{F}$ es un abierto.
3. Si $X_1, \cdots, X_n$ son abiertos entonces $X_1 \cap \cdots \cap X_n$ tambien lo es.

Demostración. La primera afirmación es trivial. Supongamos que $\mathscr{F}$ es una familia de conjuntos abiertos, sea $x \in \bigcup \mathscr{F}$ entonces por definición de unión, existe $A \in \bigcup \mathscr{F}$ tal que $x \in A$ y A es abierto, entonces por definición de conjunto abierto, existe $r > 0$ tal que $B_r(x) \subseteq A$, notemos
$$B_r(x) \subseteq A \subseteq \bigcup \mathscr{F}$$

por tanto $\bigcup \mathscr{F}$ es abierto. Para la última afirmación, sea $x \in X_1 \cap \cdots \cap X_n$ entonces existen $r_1, \cdots, r_n > 0$ tales que

$$B_{r_i}(x) \subseteq X_i,\ i = 1, \cdots, n$$

Sea $r = \text{mín}\{r_1, \cdots, r_n\} > 0$ entonces

$$B_r(x) \subseteq B_{r_i}(x) \subseteq X_i \Rightarrow B_r(x) \subseteq X_1 \cap \cdots \cap X_n$$

mostrando que $X_1 \cap \cdots \cap X_n$ es abierto. ■

Como se ve en los ejemplos, existen conjuntos que no pueden ser abiertos ni cerrados, para estudiar su topología en este tipo de conjuntos, tenemos lo siguiente.

Definición 1.4.3 — Puntos importantes. Sea $S \subseteq \mathbb{R}^n$ y $s \in \mathbb{R}^n$ decimos que s es

1. Punto interior, si existe una $r > 0$ tal que $B_r(s) \subseteq S$. Al conjunto de puntos interiores lo denotamos como $int(S)$.
2. Punto de adherencia, punto de contacto, punto de cerradura, si para toda $r > 0$ se tiene $B_r(s) \cap S \neq \emptyset$. Al conjunto de puntos de adherencia lo denotamos como $cl(S)$.
3. Punto de acumulación, punto límite, si para toda $r > 0$ existe $y \neq s$ tal que $y \in B_r(s) \cap S$.
4. Punto aislado, si existe $r > 0$ tal que $S \cap (B_r(x) - \{x\}) \neq \emptyset$.

Proposición 1.4.2 — Propiedades generales de los puntos generales. Las siguientes afirmaciones son válidas:

1. Todo punto de acumulación es un punto de adherencia.
2. Un punto es aislado si y solo si el punto no es punto límite.
3. El conjunto $int(S)$ es abierto, tal que $int(S) \subseteq S$.
4. Para cada conjunto abierto H tal que $H \subseteq S$ entonces $H \subseteq int(S)$.
5. El conjunto $cl(S)$ es cerrado, tal que $S \subseteq cl(S)$.
6. Para cada conjunto cerrado F tal que $S \subseteq F$ entonces $cl(S) \subseteq F$.
7. Un punto $s \in S$ es punto límite de S si y solo si $s \in cl(S - \{x\})$.
8. El conjunto $cl(S)$ es la unión disjunta del conjunto de puntos límites y el conjunto de puntos aislados.

Corolario 1.4.3 Las siguientes afirmaciones son válidas

1. Un subconjunto S es abierto si y sólo si $S = int(S)$.
2. Un subconjunto S es cerrado si y sólo si $S = cl(S)$.

■ **Ejemplo 1.13** Para el conjunto $\{0\} \cup [1, 4]$, tenemos

1. Los puntos aislados son 0.
2. Es un cerrado.
3. El conjunto de puntos interiores es $(1, 4)$.
4. El conjunto de puntos límite es $[1, 4]$.

■

■ **Ejemplo 1.14** Para el conjunto $\mathbb{Z}$ tenemos

1. El conjunto de los puntos aislados es $\mathbb{Z}$.
2. El conjunto de los puntos límite es $\emptyset$.
3. El conjunto de los puntos interiores es $\emptyset$.
4. Es un conjunto cerrado.

■

■ **Ejemplo 1.15** Para el conjunto $\mathbb{Q}$ tenemos

1. El conjunto de los puntos aislados es $\emptyset$, esto se prueba usando 1.0.2.
2. El conjunto de puntos límite es $\mathbb{R}$.

3. El conjunto de los puntos interiores es $\emptyset$.
4. El conjunto de los puntos cerrados es $\mathbb{R}$.

■

■ **Ejemplo 1.16** Sea $p(x) \in \mathbb{R}[x]$, un polinomio y definimos $Z_p := \{a \in \mathbb{R}|\ p(a) = 0\}$, por el teorema del factor (algebra), tenemos que Z_p es finito. Entonces:

1. El conjunto de puntos aislados es Z_p.
2. El conjunto de puntos límite es $\emptyset$.
3. El conjunto de los puntos interiores es $\emptyset$.
4. El conjunto de los puntos cerrados es Z_p.

■

Algunos conceptos interesantes que se pueden definir usando estos puntos son los siguientes

Definición 1.4.4 — Conjunto denso. Un subconjunto $S \subseteq \mathbb{R}^n$ es denso si $cl(S) = \mathbb{R}^n$.

Definición 1.4.5 — Frontera topológica. Sea S un subconjunto de $\mathbb{R}^n$ definimos su frontera topológica como alguno de estos conjuntos

$$\partial S := cl(S) - int(S) = cl(S) \cap cl(\mathbb{R}^n - S)$$

A cada punto de ∂S se le conoce como punto frontera de S.

Ejercicio 1.4 Prueba la siguiente igualdad de conjuntos:

$$cl(S) - int(S) = cl(S) \cap cl(\mathbb{R}^n - S)$$

■

Ejercicio 1.5 Prueba que son equivalentes para S:

1. Un punto p es frontera de S.
2. Un punto p satisface que para cada $r > 0$ entonces $S \cap B_r(p) \neq \emptyset$ y $(\mathbb{R}^n - S) \cap B_r(p) \neq \emptyset$.

■

■ **Ejemplo 1.17** Tenemos que $\partial B_r(x) = S_r(x)$. ■

Ejercicio 1.6 Prueba que

$$\partial S = \partial(\mathbb{R}^n - S)$$

■

Ejercicio 1.7 Prueba que un subconjunto abierto $U \subseteq \mathbb{R}^n$ es denso si y sólo si $\partial U = \mathbb{R}^n - U$. ■

Definición 1.4.6 — Conjunto denso en ninguna parte. Un subconjunto S es denso en ninguna parte si $int(cl(S)) = \emptyset$.

■ **Ejemplo 1.18** Toda recta en $\mathbb{R}^2$ es denso en ninguna parte. ■

■ **Ejemplo 1.19** El conjunto $\mathbb{Z}$ es denso en ninguna parte en $\mathbb{R}$. ■

Ejercicio 1.8 Prueba que un conjunto es la frontera de un abierto si y sólo si este es cerrado y denso en ninguna parte. ■

■ **Ejemplo 1.20** Sea $H := \{(x,y) \in \mathbb{R}^2 | \ y > 0\}$ este conjunto es abierto y su frontera es el conjunto cerrado $L := \{(x,y) \in \mathbb{R}^2 | y = 0\}$ y este conjunto es denso en ninguna parte. ■

Ahora definiremos el concepto de límite de una sucesión en un espacio métrico.

Definición 1.4.7 — Limite de una sucesión. Sea (X,d) un espacio métrico, una sucesión $\{x_n\}$ y $x \in X$, decimos que la sucesión es:

1. de Cauchy si para todo $\varepsilon > 0$ existe una $N \in \mathbb{N}$ tal que para toda $n,m \geq N$ se tiene $d(x_n, x_m) < \varepsilon$.
2. convergente con límite x, si para todo $\varepsilon > 0$ existe una $N \in \mathbb{N}$ tal que para toda $n \geq N$ se tiene que $d(x_n, x) < \varepsilon$. En esta situación denotamos $\lim_{n \to \infty} x_n = x$

Vamos a probar que todo límite convergente tiene límite único, requerimos el siguiente lema técnico.

Lema 1.4.4 Sean x,y puntos distintos, entonces existen abiertos U,V disjuntos tal que $x \in U$ y $y \in V$.

Demostración. Sea $r = d(x,y)$, como $x \neq y$ entonces $r > 0$, definamos $U = B_{r/3}(x)$ y $V = B_{r/3}(y)$, entonces es fácil ver que dichos conjuntos son los que satisfacen el lema. ■

Proposición 1.4.5 Sea $\{x_n\}_{n \in \mathbb{N}}$ una sucesión que converge a x y a y entonces $x = y$. Es decir los límites son únicos.

Demostración. Supongamos por contradicción que $x \neq y$, por el lema anterior (1.4.4), existen abiertos disjuntos U,V tales que $x \in U$ y $y \in V$, entonces sean $\varepsilon_1, \varepsilon_2 > 0$ tales que

$$B_{\varepsilon_1}(x) \subseteq U, \ B_{\varepsilon_2}(y) \subseteq V$$

usando la definición de límite, existen $N_1, N_2 > 0$ tales que

$$\{x_n\}_{n \geq N_1} \subseteq B_{\varepsilon_1}(x), \ \{x_n\}_{n \geq N_2} \subseteq B_{\varepsilon_2}(y)$$

Este proceso se puede refinar a una $\varepsilon \leq \min\{\varepsilon_1, \varepsilon_2\}$ y sea $N \geq \max\{N_1, N_2\}$, entonces

$$\emptyset \neq \{x_n\}_{n \geq N} \subseteq B_\varepsilon(x) \cap B_\varepsilon(y) \subseteq U \cap V$$

obteniendo una contradicción, por lo tanto $x = y$. ■

Ejercicio 1.9 — Diferencia entre límites y puntos límites.

1. Sea $S = \{x_n\}$ una sucesión con límite x, prueba que x es un punto cerrado de S.
2. Considera la sucesión constante $x_n = x_0$, prueba que x_0 es el límite de la sucesión pero x_0 no es punto límite de S.

■

Proposición 1.4.6 Sea S un subconjunto de un espacio métrico, entonces S es cerrado si y sólo si para toda sucesión $\{x_n\} \subseteq S$ con límite x entonces $x \in S$.

Demostración. Si S es cerrado, entonces $S = cl(S)$. entonces si $\{x_n\} \subseteq S$ es una sucesión que converge a x entonces $x \in cl(S) = S$, obteniendo $x \in S$. Conversamente, sabemos que $S \subseteq cl(S)$, entonces dado $x \in cl(S)$, para cada $n \in \mathbb{N}$ tomemos $x_n \in B_{1/n}(x) \cap S$, entonces $\{x_n\} \subseteq S$ es una sucesión que converge a x, obteniendo que $x \in S$, por lo tanto $S = cl(S)$. ■

Otra propiedad importante del concepto de las sucesiones es la noción de espacio completo.

■ **Ejemplo 1.21** Sea $S = \{x\}$, entonces S es cerrado pues toda sucesión x_n contenido en S es una sucesión constante, entonces es una sucesión convergente con límite $x \in S$, entonces por 1,4,6 tenemos que S es cerrado. ■

Proposición 1.4.7 Toda sucesión convergente es de Cauchy.

Demostración. Sea $\varepsilon > 0$, entonces existe $N \in \mathbb{N}$ tal que

$$d(x_n, x) < \varepsilon/2, \ \forall n \geq N$$

Entonces para toda $n, m \geq N$ se tiene por desigualdad triangular

$$d(x_n, x_m) \leq d(x_n, x) + d(x, x_m) < \varepsilon/2 + \varepsilon/2 = \varepsilon$$

por tanto es una sucesión de Cauchy. ■

Proposición 1.4.8 En $\mathbb{R}^n$ con la topología euclidiana, toda sucesión de Cauchy es convergente.

Demostración. Se sigue de que las proyecciones canónicas $\pi_n : \mathbb{R}^n \to \mathbb{R}$ son continuas y en $\mathbb{R}$ toda sucesión de Cauchy es convergente.(Ver capítulo 3) ■

Definición 1.4.8 — Espacios métricos completos. Un espacio métrico en donde toda sucesión de Cauchy es convergente, se le denomina completo.

■ **Ejemplo 1.22** $\mathbb{R}^n$ con la métrica euclidiana es completo. ■

Proposición 1.4.9 Si (X, d) es un espacio completo y $S \subseteq X$ un subespacio cerrado entonces S con la métrica inducida es completo.

Demostración. Sea $\{x_n\} \subseteq S$ una sucesión de Cauchy, entonces $\{x_n\}$ es convergente a un elemento $x \in X$, pero por (1.4.6), como S es cerrado $x \in S$, por tanto S es completo. ■

■ **Ejemplo 1.23** El conjunto $[0, 1]$ es completo, pero $(0, 1)$ no es completo, pues la sucesión $1/n$ es una sucesión convergente con límite 0, por 1.4.7, es una sucesión de Cauchy en $(0, 1)$, pero como 0 no está en $(0, 1)$ y por 1.4.5, tenemos que no existe un límite de la sucesión en $(0, 1)$, por tanto $(0, 1)$ no es completo. ■

Por ultimo mencionaremos el concepto de topología, como una generalización de los espacios métricos, donde su enfoque es la importancia de la forma del conjunto como su cerradura, densidad, frontera, entre otros conceptos.

Definición 1.4.9 — Espacio topológico. Dado un conjunto X, una familia de conjuntos τ es llamado topología de X si cumple las siguientes propiedades:

1. Los conjuntos $\emptyset, X$ están en τ.
2. Si $\mathscr{F}$ es una familia de conjuntos en τ entonces $\bigcup \mathscr{F} \in \tau$.
3. Si $X_1, \cdots, X_n$ son conjuntos en τ entonces $X_1 \cap \cdots \cap X_n \in \tau$.

■ **Ejemplo 1.24** Dado $\|\|$ una norma de $\mathbb{R}^n$, definimos $\tau_{\|\|}$ como el conjunto de todos los subconjuntos abiertos según la definición 1.4.1, entonces por la proposición 1.4.1 muestra que $\tau_{\|\|}$ es una topología de $\mathbb{R}^n$. ■

Con esta construcción dado en el ejemplo anterior, podemos volver a comprender el concepto de 1.3.4.

Proposición 1.4.10 Sean $\|\|_1, \|\|_2$ dos normas, si son normas equivalentes entonces definen la misma topología, es decir $\tau_{\|\|_1} = \tau_{\|\|_2}$.

Demostración. Si $\|\|_1$ y $\|\|_2$ son equivalentes, entonces

$$c_1 \|\|_2 \leq \|\|_1 \leq c_2 \|\|_2 \tag{1.1}$$

Tomemos $U \in \tau_{\|\|_2}$ un abierto, entonces para cada $x \in U$ existe $r > 0$ tal que $B_r(x) \subseteq U$, usando (1.1) tenemos

$$B_r(x) = \{y| \ \|x-y\|_2 < r\} \subseteq \{y| \ \|x-y\|_1 < c_2 r\} = B'_{c_2 r}(x)$$

$$B'_{c_2 r}(x) \subseteq B_r(x)$$

entonces $B'_{c_2 r}(x) \subseteq U$, obtenemos que $U \in \tau_{\|\|_1}$, entonces $\tau_{\|\|_2} \subseteq \tau_{\|\|_1}$. De manera similar podemos ver $\tau_{\|\|_1} \subseteq \tau_{\|\|_2}$. ■

Corolario 1.4.11 La topología inducida por una norma en $\mathbb{R}^n$ es la topología inducida por la norma euclidiana.

1.5. Compacidad y conexidad.

De aqui en adelante se asume a $\mathbb{R}^n$ con la norma euclidiana y su topología inducida segun la definición (1.4.2).

Definición 1.5.1 — Conjunto compacto. Un subconjunto $S \subseteq \mathbb{R}^n$ es compacto si para cada cubierta abierta, es decir, una familia de abiertos $\mathscr{F} = \{U_i\}_{i \in I}$ tal que $S \subseteq \bigcup_{i \in I} U_i$, existe una subcubierta finita, es decir, existen $U_{i_1}, \cdots, U_{i_n} \in \mathscr{F}$ tal que $S \subseteq \bigcup_{k=1}^{n} U_{i_k}$

■ **Ejemplo 1.25** Un conjunto finito $\{x_1, \cdots, x_n\} \subseteq \mathbb{R}^n$ es compacto. ■

■ **Ejemplo 1.26** $\mathbb{N} \subseteq \mathbb{R}$ no es compacto, y esto se sigue observar que la siguiente cubierta abierta $U_n = (n - 1/5, n + 1/5)$ indexada por $\mathbb{N}$ no tiene subcubierta finita. ■

■ **Ejemplo 1.27** El conjunto $S = \{0\} \cup \left\{\frac{1}{n} \mid n \in \mathbb{N}_{>0}\right\}$ es numerable y compacto, pues si tomamos una cubierta abierta, es decir, una familia de abiertos $\mathscr{F} = \{U_i\}_{i \in I}$ tal que $S \subseteq \bigcup_{i \in I} U_i$, entonces existe $i_0 \in I$ tal que $0 \in U_{i_0}$, y por definición de abierto existe $0 \in (-\varepsilon, \varepsilon) \in U_{i_0}$, por la propiedad arquimediana 1.0.4, existe $n_0 \in \mathbb{N}_{>0}$ tal que $1/n_0 < \varepsilon$, esto significa que el subconjunto $\left\{\frac{1}{n} \mid n \geq n_0\right\} \subseteq (-\varepsilon, \varepsilon) \subseteq U_{i_0}$ para cada $k = 1, \cdots, n_0 - 1$ tomamos un abierto $\frac{1}{k} \in U_{i_k}$ obteniendo una subcubierta finita $U_{i_0}, U_{i_1}, \cdots, U_{i_{n_0 - 1}}$, por lo tanto S es compacto. ■

En los dos ejemplos presentados, se observa que la definición de un conjunto compacto está intrínsecamente ligada a la topología, que intuitivamente se relaciona con la forma del conjunto. Dado que la topología se deriva de la norma euclidiana, existen ciertos conceptos del cálculo que pueden tener un impacto significativo en si un conjunto puede ser o no compacto.

Ejercicio 1.10
1. Sea x_n una sucesión de números reales con límite x. Prueba que el conjunto $S := \{x\} \cup \{x_n \mid n \in \mathbb{N}\}$ es compacto.
2. Considera la sucesión $x_n = (-1)^n$, prueba que este conjunto no tiene límite pero $S = \{x_n\}$ es compacto. ¿Esto contradice el inciso anterior?

■

Proposición 1.5.1 Sea S un conjunto compacto y $F \subseteq S$ un cerrado, entonces F es compacto.

Demostración. Sea $\mathscr{F}$ una cubierta abierta de F, como $S - F$ es abierto, entonces $\mathscr{F} \cup \{S - F\}$ es una cubierta abierta de S, como S es compacto, existen $A_1, \cdots, A_n \in \mathscr{F}$ tal que $\{A_1, \cdots, A_n, S - F\}$ es una cubierta abierta de S, entonces $\{A_1, \cdots, A_n\}$ es cubierta abierta de F, por tanto F es compacto. ■

1.5.1. Teorema de Heine-Borel.

Proposición 1.5.2 Sea $E \subseteq K \subseteq \mathbb{R}^m$ donde E es un conjunto infinito, K compacto entonces E tiene un punto de acumulación en K

Demostración. Supongamos que E no tiene puntos de acumulación en K, asi que dado $x \in K$ existe $r_x > 0$ tal que $(B_{r_x}(x) - \{x\}) \cap E = \emptyset$, asi que la familia $\mathscr{F} := \{B_{r_x}(x) | \; x \in K\}$ es una cubierta abierta de K, como K es compacto, existen $x_1, \cdots, x_n$ tales que $B_{r_{x_1}}(x_1), \cdots, B_{r_{x_n}}(x_n)$ cubren a K, pero esto implica que $E = \{x_1, \cdots, x_n\}$ y esto contradice el hecho de que E es infinito. ■

Proposición 1.5.3 El conjunto $R := [a_1, b_1] \times \cdots \times [a_n, b_n]$ es compacto.

Vamos a dar una caracterización necesaria y suficiente para describir los conjuntos compactos en $\mathbb{R}^n$.

Teorema 1.5.4 — Caracterización de compactos en $\mathbb{R}^n$. Dado un subconjunto $S \subseteq \mathbb{R}^n$, son equivalentes

1. S es compacto.
2. (**Heine-Borel**) S es cerrado y acotado.
3. (**Bolzano-Weierstrass**) Todo subconjunto infinito de S tiene un punto de acumulación en S.

Demostración. Supongamos que S es cerrado y acotado, entonces $S \subseteq R$ para algún rectángulo $R = [a_1, b_1] \times \cdots \times [a_n, b_n]$, como R es compacto y S es cerrado entonces S es compacto, mostrando (2) implica (1). (1) implica (3) se sigue de (1.5.2). Ahora para probar (3) implica (2), supongamos que S no es acotado, entonces tomemos una sucesión $\{x_n\}$ tal que $\|x_n\| > n$, obteniendo una sucesión infinita de S sin punto de acumulación, lo cual es una contradicción. Tambien si S no es cerrado, sea $x_0 \in cl(S) - S$, entonces existen $x_n \in S$ tal que

$$\|x_n - x_0\| < \frac{1}{n}$$

entonces $E := \{x_n\}$ es un conjunto infinito con un punto de acumulación que no esta en K, contradiciendo la hipótesis (3). Por lo tanto (3) implica (2). ■

Corolario 1.5.5 Todo subconjunto $S \subseteq \mathbb{R}^n$ compacto es completo.

La formulación clásica del teorema de Bolzano-Weierstrass es la siguiente

Corolario 1.5.6 Una sucesión acotada en $\mathbb{R}^n$ tiene una subsucesión convergente.

1.5.2. Conexidad

Definición 1.5.2 — Abierto relativo. Sea $S \subseteq \mathbb{R}^n$ decimos que un subconjunto $A \subseteq S$ es un abierto relativo a S si A es de la forma:

$$A = U \subseteq S$$

donde $U \subseteq \mathbb{R}^n$ es un abierto.

Definición 1.5.3 — Conjunto conexo. Un subconjunto $S \subseteq \mathbb{R}^n$ es conexo si cada ves que tenemos una descomposición en abiertos disjuntos, relativos a S

$$S = A_1 \cup A_2$$

Entonces $A_1 = \emptyset$ o $A_2 = \emptyset$. En otro caso, decimos que S es disconexo.

■ **Ejemplo 1.28** Consideremos $[a,b] \subseteq \mathbb{R}$, entonces este conjunto es conexo. ■

■ **Ejemplo 1.29** El n-rectángulo $R = \prod_{i=1}^{n} [a_i, b_i]$ es conexo. ■

■ **Ejemplo 1.30** Un conjunto convexo es conexo. Decimos que un conjunto $S \subseteq \mathbb{R}$ es convexo si para todo $a, b \in S$ el conjunto $\{a + t(b-a) | \ t \in [0,1]\}$ está contenido en S. ■

■ **Ejemplo 1.31** Si tenemos una cadena $A_1 \subseteq A_2 \subseteq \cdots$ tal que cada A_i es conexo entonces $\bigcup_{i=1}^{\infty} A_i$ es conexo. ■

■ **Ejemplo 1.32** $\mathbb{R}^n$ es conexo. ■

■ **Ejemplo 1.33** El subconjunto $[0,1) \subseteq (1,2]$ es disconexo. ■

Proposición 1.5.7 Si $S \subseteq \mathbb{R}^n$ es conexo entonces $cl(S)$ es conexo.

Demostración. Usaremos (2.2.2), sea $f \colon cl(S) \to \{0,1\} \subseteq \mathbb{R}$ continua, usando que $i \colon S \to cl(S)$ es continua, entonces $f \circ i \colon S \to \{0,1\}$ es continua, pero como S es conexo, entonces $f \circ i$ es constante, obteniendo que f es constante, por (2.2.2) entonces $cl(S)$ es conexo. ■

1.6. Funciones en varias variables.

Tipicamente, cuando se piensa una función de la forma $f \colon A \to B$ se visualiza como una gráfica de la forma

$$Gra(f) := \{(x,y) \in A \times B | f(x) = y\}$$

es decir como un subconjunto de $A \times B$, de esta manera toda función de una variable $f \colon A \subseteq \mathbb{R} \to \mathbb{R}$ se puede visualizar como un subconjunto de $\mathbb{R}^2$. Sin embargo, cuando la dimensión de las variable suben, nuestra limitación de espacio tridimensional nos limita generalizar este proceso a varias variables, por ejemplo si seguimos el mismo punto de vista para visualizar una grafica de una función $f \colon \mathbb{R}^2 \to \mathbb{R}^2$ necesitariamos dibujarlo fielmente en un subconjunto de $\mathbb{R}^4$. Actualmente con la variedad de problemas que resuelve el cálculo diferencial, no todas las funciones se necesitan visualizar como una gráfica, existen diferentes formas de visualizar una función dependiendo el uso y su interpretación que se dé. A continuación mencionaremos 3 tipos de funciones y sus visualizaciones con su interpretación más común.

1.6.1. Funciones escalares.

Se entiene como función escalar a una función de la forma $f \colon \mathbb{R}^n \to \mathbb{R}$ con $n \geq 2$. Este tipo de funciones se les considera escalares porque asigna a cada vector de $\mathbb{R}^n$ un valor escalar de $\mathbb{R}$, cuyas interpretaciones varia de su modelo, por ejemplo, puede usarse para medir la altura de una región del espacio, para asignar el costo del terreno en el plano, para asignar una temperatura, etcétera. Cualitativamente se puede visualizar, como una gráfica como se muestra acontinuación:

Donde tipicamente en $\mathbb{R}^3$ el eje Z, se le considera los valores escalares de la función escalar definido por los vectores del plano XY.

Como conjuntos de nivel. Cabe mencionar que otra manera de visualizar es por medio de los conocidos conjuntos de nivel, formalmente, dado $f \colon \mathbb{R}^n \to \mathbb{R}$ una función escalar y $r \in \mathbb{R}$ definimos el conjunto de nivel

$$L_f := f^{-1}(r) = \{x \in \mathbb{R}^n | \ f(x) = r\}$$

entonces una forma de visualizar f es atravez del trazado de sus conjuntos de nivel, como se muestra en el ejemplo anterior:

Este tipo de situaciones se puede ver por ejemplo en mapas topográficos.

1.6.2. Parametrizaciones.

Se entiende como parametrización como una función (mayormente inyectiva y continua) de la forma $f\colon \mathbb{R}^n \to \mathbb{R}^m$ en donde $n < m$. Su principal uso, como dice su nombre, es poder parametrizar un objeto geométrico de dimensión n dibujado en $\mathbb{R}^m$ con las coordenadas que dicta f, como se muestra en el siguiente ejemplo:

Es importante mencionar que aunque de manera intuitiva una parametrización es una función inyectiva y continua, la imagen de dicho objeto en $\mathbb{R}^m$ puede generar objetos patológicos que suelen salir fuera de la ïntuición"de un objeto geométrico suave, por ejemplo

Entonces, es necesario que en un futuro se aclare como parametrizar objetos llamados típicamente "geometricos suaves"para poder estudiarlo sin muchas restricciones causadas por considerar una situación más general (encontrar los casos patológicos que el cálculo diferencial no le conviene).

1.6.3. Campos Vectoriales.

Un campo vectorial, se puede pensar, de manera introductoria, como un modelo geométrico para representar dentro de un subconjunto abierto de $\mathbb{R}^n$, a cada punto de $\mathbb{R}^n$ asignandole un vector de $\mathbb{R}^n$. Este tipo de gráficas se utilizan para representar por ejemplo como funciona la dinámica de un cuerpo o sistema moviendose en un ambiente. Este tipo de modelos puede requerir muchos requisitos, dependiendo la complejidad del modelo. Sin embargo para tener un primer acercamiento adecuado, en Cálculo Diferencial, se piensa un campo vectorial como una función de la forma $f\colon \mathbb{R}^n \to \mathbb{R}^n$ y su visualización es el subconjunto $Domf \subseteq \mathbb{R}^n$ junto con la propiedad de que a cada punto $p \in \mathbb{R}^n$ se trace un vector que inice en p y termine en $f(p) - p$, como se muestra en el siguiente ejemplo

Capítulo 2

Límites y Funciones Continuas.

Consideremos la definición de función continua en los reales de una variable, una función $f\colon (a,b) \to \mathbb{R}$ es continua en un punto $t \in (a,b)$ si para cada $\varepsilon > 0$ existe $\delta > 0$ tal que si

$$|t-s| < \delta \Rightarrow |f(t)-f(s)| < \varepsilon$$

Como $B_r(x) = (x-r, x+r)$ entonces tenemos la siguiente reformulación a nivel topológico.

Proposición 2.0.1 Sea $f\colon \mathbb{R} \to \mathbb{R}$ una función y $t \in \mathbb{R}$ entonces son equivalentes:

1. f es continua en t.
2. Para todo $\varepsilon > 0$ existe $\delta > 0$ tal que $f(B_\delta(t)) \subseteq B_\varepsilon(f(t))$.

Ejercicio 2.1 Prueba (2.0.1). ■

Definición 2.0.1 — Función continua. Sean $(X,d_1),(Y,d_2)$ dos espacios métricos, una función $f\colon X \to Y$ es continua en $t \in X$ si para todo $\varepsilon > 0$ existe $\delta > 0$ tal que $f(B_\delta(t)) \subseteq B_\varepsilon(f(t))$. Ademas decimos que f es continua en X si f es continua en cada $t \in X$.

■ **Ejemplo 2.1** Las funciones proyección $\pi_i\colon \mathbb{R}^n \to \mathbb{R}$ son continuas. Para ver esto, fijemos $x = (x_1,\cdots,x_n) \in \mathbb{R}^n$ y tomemos $\varepsilon > 0$, notemos que $A := f^{-1}(B_\varepsilon(\pi_i(x))) = \mathbb{R} \times \cdots \times (x_i - \varepsilon, x_i + \varepsilon) \times \cdots \times \mathbb{R}$, entonces si proponemos $\delta := \varepsilon/2$ entonces $B_\delta(x) \subseteq A$ obteniendo que $f(B_\delta(x)) \subseteq B_\varepsilon(\pi_i(x))$. Por tanto π_i es continua en cada $x \in \mathbb{R}^n$. ■

Proposición 2.0.2 Son equivalentes para una función $f\colon X \to Y$:

1. f es continua en X.
2. Para todo abierto $U \subseteq Y$ implica que $f^{-1}(U) \subseteq X$ es abierto.

Demostración. Si f es continua en X, sea $U \subseteq Y$ abierto, denotemos $V := f^{-1}(U)$, y tomemos $a \in V$ entonces $f(a) \in U$, pero por definición de abierto (1.4.2) existe una $\varepsilon > 0$ tal que $B_\varepsilon(f(a)) \subseteq U$, como f es continua en a, por (2.0.1), obtenemos que existe una $\delta > 0$ tal que $f(B_\delta(a)) \subseteq B_\varepsilon(f(a))$ esto implica que $f(B_\delta(a)) \subseteq U$, obteniendo que $B_\delta(a) \subseteq V$, por tanto V es abierto. Conversamente, supongamos (2) cierto, y sea $b \in X$ un punto arbitrario fijo, entonces $U := B_\varepsilon(f(b))$ es abierto en Y para cualquier $\varepsilon > 0$ fija arbitraria, por hipotesis $f^{-1}(U)$ es abierto en X y además $b \in f^{-1}(U)$, por definición (1.4.2), existe $\delta > 0$ tal que $B_\delta(b) \subseteq f^{-1}(U)$, esto implica que $f(B_\delta(b)) \subseteq B_\varepsilon(f(b))$, por tanto f es continua en b, obteniendo la continuidad en X. ■

Corolario 2.0.3 Las siguientes afirmaciones son válidas:

1. La función Id_X es continua.
2. Si $f\colon X\to Y$ y $g\colon Y\to Z$ son continuas entonces $g\circ f$ es continua.

Demostración. Usando (2.0.2), si $U\subseteq X$ es abierto entonces $(\mathrm{Id}_X)^{-1}(U)=U$ el cual es un abierto, por tanto Id_X es continua. Si $W\subseteq Z$ es abierto, como g es continua entonces $g^{-1}(W)$ es abierto en Y, entonces $f^{-1}(g^{-1}(W))$ es abierto en X, además $(g\circ f)^{-1}(W)=f^{-1}(g^{-1}(W))$, obteniendo por (2.0.2) que $g\circ f$ es continua. ■

Lema 2.0.4 La función $\pi_i\colon\mathbb{R}^n\to\mathbb{R}$ envía abiertos en abiertos, es decir, si $U\subseteq\mathbb{R}^n$ entonces $\pi_i(U)\subseteq\mathbb{R}$ es abierto.

Demostración. Sea $U\subseteq$ un abierto y tomemos $t\in\pi_i(U)$, entonces existe $x=(x_1,\cdots,x_n)\in U$ tal que $x_i=t$, como U es abierto, entonces existe $\varepsilon>0$ tal que $B_\varepsilon(x)\subseteq U$, aplicando π_i obtenemos $\pi_i(B_\varepsilon(x))\subseteq\pi_i(U)$, es decir $B_\varepsilon(t)\subseteq\pi_i(U)$, por tanto $\pi_i(U)$ es abierto. ■

■ **Ejemplo 2.2 — Continuidad en varias variables.** Sea $f\colon X\to\mathbb{R}^n$ una función desde un espacio métrico, definimos $f_i:=\pi_i\circ f\colon X\to\mathbb{R}$, entonces $f(x)=(f_1(x),f_2(x),\cdots,f_n(x))$. Afirmamos que f es continua si y sólo si todas las f_i son continuas. Para ver esto, si f es continua, como π es continua por (2.1) entonces por (2.0.3) tenemos que f_i es continua para cada $i=1,\cdots,n$. Conversamente si las f_i son continuas, sea $U\subseteq\mathbb{R}^n$ abierto de la forma $U=B_\varepsilon(x)$ con $x=(x_1,\cdots,x_n)\in\mathbb{R}^n$. Para cada $0<\delta<\varepsilon$ entonces $\pi_i(B_\delta(x))=B_\delta(x_i)\subseteq\mathbb{R}$ es abierto, pero f_i es continua, entonces $V_i:=f_i^{-1}(B_\delta(x_i)\subseteq X$ es abierto, sea $V_\delta:=\bigcap_{i=1}^n V_i$, por (1.4.1) es un abierto , además $V_\delta\subseteq V_i=f^{-1}(\pi_i^{-1}(B_\delta(x_i)))$ implica $f(V_\delta)\subseteq\pi_i^{-1}(B_\delta(x_i))$, entonces escogemos $\delta>0$ tal que $\bigcap_{i=1}^n\pi_i^{-1}(B_\delta(x_i))\subseteq B_\varepsilon(x)$. Como $x\in V_\delta$ y es un abierto, entonces existe $\delta_1>0$ tal que $B_{\delta_1}(x)\subseteq V_\delta$, aplicando f obtenemos $f(B_{\delta_1}(x))\subseteq B_\varepsilon(x)$, entonces f es continua en x, por tanto f es continua en X. ■

Ejercicio 2.2 Prueba que para cada $r>0$ existe $\delta>0$ tal que $\bigcap_{i=1}^n\pi_i^{-1}(B_\delta(x_i))\subseteq B_r(x)$ ■

■ **Ejemplo 2.3 — La norma es continua.** Sea $(\mathbb{R}^n,\|\|)$ un espacio normado, entonces la función $\|\|\colon\mathbb{R}^n\to\mathbb{R}$ es continua, pues si fijamos $x\in\mathbb{R}^n$, entonces para cada $\varepsilon>0$ sea $\delta=\varepsilon$, asi, cuando

$$\|x-y\|<\delta$$

entonces por (1.1)

$$|\|x\|-\|y\||\leq\|x-y\|<\delta=\varepsilon$$

por lo tanto $\|\|$ es continua en x, obteniendo la continuidad en $\mathbb{R}^n$. ■

■ **Ejemplo 2.4 — La distancia a un punto es continua.** El ejemplo anterior se puede extender a espacios métricos, pues si fijamos $a\in X$ con (X,d) un espacio métrico, definimos $f_a\colon X\to\mathbb{R}$ como $f_a(x):=d(x,a)$ entonces, usando la desigualdad

$$|d(x,a)-d(y,a)|\leq d(x,y)$$

que se puede obtener de la desigualdad triangular, obtenemos que f_a es continuo en X. Cuando $X=\mathbb{R}^n$ y la métrica es inducida por una norma, entonces $\|\|=f_0$. ■

Proposición 2.0.5 Sea $f\colon X\to Y$ continua en $t\in X$ y sea $\{x_n\}$ una sucesión convergente a t entonces $\{f(x_n)\}$ es una sucesión convergente a $f(t)$, es decir

$$f(\lim_{n\to\infty}x_n)=\lim_{n\to\infty}f(x_n)$$

Demostración. Sea $\varepsilon > 0$, por continuidad (2.0.1) existe $\delta > 0$ tal que

$$f(B_\delta(t)) \subseteq B_\varepsilon(f(t)) \tag{2.1}$$

Por definición de límite de sucesión (1.4.7) existe una $N \in \mathbb{N}$ tal que para toda $n \geq N$ se tiene $d(x_n, t) < \delta$, es decir

$$\{x_n\}_{n\geq N} \subseteq B_\delta(t)$$

aplicando imagen directa de f obtenemos $\{f(x_n)\} \subseteq f(B_\delta(t))$ y por (2.1) obtenemos $\{f(x_n)\}_{n\geq N} \subseteq B_\varepsilon(f(t))$ es decir $d(f(x_n), f(t)) < \varepsilon$ para toda $n \geq N$, por tanto $f(x_n)$ converge a $f(t)$. ■

Nuestro principal interés como se mencionó en el capitulo anterior son el estudio de las funciones de la forma

$$f\colon D \subseteq \mathbb{R}^n \to \mathbb{R}^m$$

ahora, si usamos el enfoque del ejemplo (2.2), tenemos que f se puede describir como

$$f(x_1, \cdots, x_n) = (f_1(x_1, \cdots, x_n), \cdots, f_m(x_1, \cdots, x_m))$$

donde las $f_i\colon D \subseteq \mathbb{R}^n \to \mathbb{R}$ son funciones escalares. Entonces podemos empezar a estudiar sobre las funciones escalares para despues extenderlo sobre funciones de $\mathbb{R}^n \to \mathbb{R}^m$.

2.1. Límites de Funciones.

Definición 2.1.1 — Límite de una función. Sea $f\colon D \subseteq \mathbb{R}^n \to \mathbb{R}^m$, $t \in D$ y $L \in \mathbb{R}^m$, decimos que f tiende a L cuando x tiende a t si para toda $\varepsilon > 0$ existe una $\delta > 0$ tal que

$$0 < \|x - t\| < \delta \Rightarrow \|f(x) - L\| < \varepsilon$$

En esta situación se escribe como

$$\lim_{x\to t} f(x) = L$$

Ejercicio 2.3 — Algebra de límites. Prueba las siguientes identidades:

1. $\lim_{x\to a}(f(x) + g(x)) = \lim_{x\to a} f(x) + \lim_{x\to a} g(x)$.
2. Si f, g son escalares entonces
$$\lim_{x\to a}(f(x)g(x)) = (\lim_{x\to a} f(x))(\lim_{x\to a} g(x))$$
3. Si f, g son escalares, $g \neq 0$ en una vecindad de a y $\lim_{x\to a} g(x) \neq 0$ entonces
$$\lim_{x\to a} \frac{f(x)}{g(x)} = \frac{\lim_{x\to a} f(x)}{\lim_{x\to a} g(x)}$$

■

Proposición 2.1.1 Sea $f\colon D \subseteq \mathbb{R}^n \to \mathbb{R}^m$, $g\colon S \subseteq \mathbb{R}^m \to \mathbb{R}^s$, $t \in D$, si f tiene límite en t y g es continua en el límite, entonces

$$\lim_{x\to t} g \circ f(x) = g\left(\lim_{x\to t} f(x)\right)$$

Demostración. Sea L el límite de f cuando $x \to t$. Por hipótesis, g es continua en L, entonces para cada $\varepsilon > 0$ existe $\varepsilon_1 > 0$ tal que

$$\|y - L\| \leq \varepsilon_1 \Rightarrow \|g(y) - L\| \leq \varepsilon \tag{2.2}$$

Para ε_1, por definición del límite L, existe $\delta > 0$ tal que si

$$0 < \|x-t\| < \delta \Rightarrow \ \|f(x) - L\| \leq \varepsilon_1$$

Combinando con (2.2) cuando $y = f(x)$ tenemos la igualdad deseada. ■

Proposición 2.1.2 Sea $f\colon D \subseteq \mathbb{R}^n \to \mathbb{R}^m$, $t \in D$ y $L \in \mathbb{R}^m$, a cada $i = 1, \cdots, n$ denotamos $f_i = \pi_i \circ f\colon D \to \mathbb{R}$. Si f tiene límite en t entonces

$$\lim_{x\to t}(f_i(x)) = (\lim_{x_i \to t_i} f_i(x))$$

Demostración. Se sigue del anterior, usando que $\pi_i\colon \mathbb{R}^m \to \mathbb{R}$ es continua, entonces de

$$\lim_{x\to t}(f_i(x)) = (a_i)$$

implica que al componer con π_i

$$\lim_{x\to t} f_i(x) = a_i$$

■

■ **Ejemplo 2.5** Vamos a calcular el límite de

$$\lim_{(x,y)\to(0,0)} (x^2+y^2)\sin(\frac{1}{x+y})$$

Usando el hecho de que

$$\left|(x^2+y^2)\sin(\frac{1}{x+y})\right| \leq |x^2+y^2|$$

obtenemos

$$\lim_{(x,y)\to(0,0)} (x^2+y^2)\sin(\frac{1}{x+y}) = 0$$

■

Como en el caso del cálculo diferencial en una variable, no hay una única manera de determinar si existe o no un límite, y en la práctica, suele ser más sencillo determinar si un límite no existe. Para varias variables, existe una método útil para empezar a determinar si un límite no existe.

Método de las líneas. Supongamos que deseamos determinar si existe o no el límite

$$\lim_{(x,y)\to(a,b)} f(x,y)$$

Consideremos $m \in \mathbb{R}$ y la recta con pendiente m que pase por (a,b), definido por la ecuación

$$y - b = m(x-a)$$

Entonces, usando la definición de límite, si existe el límite, digamos que tenga límite L, entonces al considerar el límite sobre la recta, es decir al realizar el caso particular de

$$x = t, y = b + m(t-a)$$

entonces tenemos que

$$lim_{(x,y)\to(a,b)} f(x,y) = \lim_{t\to a} f(t, b+m(t-a))$$

obteniendo el siguiente razonamiento por contraposición

R Si $\lim_{t\to a} f(t, b+m(t-a))$ tiene distintos límites al variar m entonces no existe $\lim_{(x,y)\to(a,b)} f(x,y)$.

■ **Ejemplo 2.6** Vamos a determinar si existe o no el siguiente límite

$$\lim_{(x,y)\to(0,0)} \frac{3xy}{x^2+y^2}$$

por el método de las líneas, consideremos las líneas $y = mx$ con $m \in \mathbb{R}$ fijo albitrario, entonces al calcular

$$\lim_{t\to 0} \frac{3t\cdot(mt)}{t^2+(mt)^2} = \lim_{t\to 0} \frac{t^2(3m)}{t^2(1+m^2)} = \lim_{t\to 0} \frac{3m}{1+m^2} = \frac{3m}{1+m^2}$$

entonces vemos que el límite varia cuando m cambia, entonces podemos asegurar que no existe el límite. ■

■ **Ejemplo 2.7** Vamos a estudiar el límite de

$$\lim_{(x,y)\to(0,0)} \frac{5x^2y^2}{x^2+y^2}$$

podemos empezar explorando el método de las líneas, consideremos las líneas $y = mx$ con $m \in \mathbb{R}$ fijo albitrario, entonces al calcular

$$\lim_{t\to 0} \frac{5t^2(mt)^2}{t^2+(mt)^2} = \lim_{t\to 0} \frac{t^4(5m)}{t^2(1+m^2)} = 0$$

vemos que el límite no depende de m, mostrando que hay una posibilidad de que exista dicho límite, ¿Cuál deberia ser el límite? debe ser 0 por el método de las líneas. Sea $\varepsilon > 0$ y consideremos $\delta < \sqrt{\varepsilon/5}$, notemos que se tiene la siguiente desigualdad

$$0 < \left|\frac{5y^2}{x^2+y^2}\right| < 5$$

Ahora, cada ves que

$$\|(x,y)\| < \delta$$

entonces

$$|f(x,y)-0| = \left|x^2\cdot\frac{5y^2}{x^2+y^2}\right| < \delta^2\cdot 5 < \varepsilon/5\cdot 5 = \varepsilon$$

concluyendo que

$$\lim_{(x,y)\to(0,0)} \frac{5x^2y^2}{x^2+y^2} = 0$$

■

2.2. Propiedades de la Continuidad.

En esta sección veremos propiedades que preserva la continuidad.

Proposición 2.2.1 Sea $f\colon D \to \mathbb{R}^m$ una función continua, entonces

1. Si $S \subseteq D$ es compacto, entonces $f(S)$ es compacto.
2. Si $S \subseteq D$ es conexo, entonces $f(S)$ es conexo.

Demostración. Para el primer punto, sea $\mathscr{F}$ una cubierta abierta de $f(S)$, como f es continua, por (2.0.2) tenemos que $\mathscr{G} = \{f^{-1}(U) | \; U \in \mathscr{F}\}$ es una cubierta abierta de S, pero S es compacto, entonces existen $U_1, \cdots, U_n$ tal que $f^{-1}(U_1), \subseteq, f^{-1}(U_n)$ es una subcubierta abierta finita de S, entonces $U_1, \cdots, U_n$ es una subcubierta abierta finita en $f(S)$. Para el segundo punto, supongamos que $f(S) = A \cup B$ es una descomposición disjunta de abiertos, entonces $S = f^{-1}(A) \cup f^{-1}(B)$ es una descomposición disjunta de abiertos, pero S es conexo entonces $f^{-1}(A)$ ó $f^{-1}(B)$ es vacio, entonces A o B es vacio, por tanto S es conexa. ∎

Tenemos una interesante caracterización de los conexos usando las funciones continuas.

Proposición 2.2.2 Sea $S \subseteq \mathbb{R}^n$ un subconjunto, entonces son equivalentes

1. S es conexo.
2. Las unicas funciones continuas $f\colon S \to \mathbb{R}$ con $imf = \{0,1\}$ son constantes.

Demostración. Si S es conexo, y sea $f\colon S \to \mathbb{R}$ continua con $imf = \{0,1\}$, entonces $imf = \{0\} \cup \{1\}$ es una descomposición disjunta de abiertos en imf, etnonces $S = f^{-1}(0) \cup f^{-1}(1)$ es una descomposición disjunta de abiertos, pero S es continua, entonces $f^{-1}(0)$ o $f^{-1}(1)$ es vacio, entonces f es constante con valor 0 ó 1. El regreso es directo. ∎

Proposición 2.2.3 Sea $A := A_1 \times \cdots \times A_n$ entonces A es conexo si y sólamente si A_i es conexo para cada i

Ejercicio 2.4 Considere $S = \{(x,y) \in [0,1] \times [0,1] | x < y\} \cup \{(x,y) \in [0,1] \times [0,1] | y < x\}$, prueba lo siguiente:

1. Muestra que S es disconexo.
2. Muestra que $\pi_1(S)$ y $\pi_2(S)$ son conexos.
3. ¿Contradice a la proposición (2.2.3)?

▪

2.2.1. Teorema del Valor intermedio.

Introduciremos una definión topológica importante para el siguiente teorema.

Definición 2.2.1 — Conjunto arco-conexo. Un conjunto $S \subseteq \mathbb{R}^n$ es arco-conexo si para todo par de puntos $a, b \in S$ existe una función continua $f\colon [0,1] \to S$ tal que:

1. $f(0) = a$ y $f(1) = b$.
2. $\gamma([0,1]) \subseteq R$.

▪ **Ejemplo 2.8** Considera $R := [a,b] \times [\hat{a}, \hat{b}]$, entonces este conjunto es arco-conexo, pues si tomamos $(x_1,y_1),(x_2,y_2) \in R$, definimos $\gamma\colon [0,1] \to R$ como sigue

$$\gamma(t) = \begin{pmatrix} x_1 \\ y_2 \end{pmatrix} + t \begin{pmatrix} x_2 - x_1 \\ y_2 - y_1 \end{pmatrix}, \; t \in [0,1]$$

claramente $\gamma([0,1]) \subseteq R$, $\gamma(0) = (x_1,y_1)$ y $\gamma(1) = (x_2,y_2)$. ▪

▪ **Ejemplo 2.9** Si S es arco-conexo entonces $int(S)$ también lo es. ▪

▪ **Ejemplo 2.10** Todo conjunto convexo es arco-conexo. ▪

▪ **Ejemplo 2.11** Sea $S_r \subseteq \mathbb{R}^2$ el arco superior de la circunferencia de radio r, entonces $S := \bigcup_{1 \le r \le 2} S_r$ es arco-conexo. ▪

▪ **Ejemplo 2.12** Sea $\{S_i\}_{i \in I}$ una familia de conjuntos arco-conexos tal que para toda $i, j \in I$ existen $i_0 = i, i_1, \cdots, i_{n-1}, i_n = j$ en I tal que $S_{i_k} \cap S_{i_{k+1}} \neq \emptyset$ entonces $S := \bigcup_{i \in I} S_i$ es arco-conexo. Esto se prueba como sigue, toma $a, b \in S$ entonces por definición de unión $\exists i, j \in I$ tal que $a \in S_i, b \in S_j$, luego por hipótesis, existen $i_0 = i, i_1, \cdots, i_{n-1}, i_n = j$ en I tal que $S_{i_k} \cap S_{i_{k+1}} \neq \emptyset$, así que toma $x_k \in S_{i_k} \cap S_{i_{k+1}}$ para cada $k = 1, \cdots, n-1$, define $x_0 = a$ y $x_n = b$ luego:

1. Construye $\gamma_k \colon [0,1] \to S_k$ que conecta x_k con x_{k+1}. Esto es posible porque S_k es arco-conexo.
2. Define $\gamma \colon [0,1] \to S$ como sigue
$$\frac{k}{n} \leq t \leq \frac{k+1}{n},\ \gamma(t) = \gamma_k(nt-k)$$
entonces es fácil ver que γ es continua.
3. Notemos que $\gamma(0) = \gamma_0(0) = x_0 = a$ y $\gamma(1) = \gamma_{n-1}(n(1)-(n-1)) = \gamma_{n-1}(1) = x_n = b$

Por tanto S es arco-conexo. ■

Proposición 2.2.4 Todo conjunto arco-conexo es conexo.

Demostración. Supongamos que S es arco-conexo pero no es conexo, por (2.2.2) existe una función $f \colon S \to \{0,1\} \subseteq \mathbb{R}$ que es continua y no es constante, toma $a \in f^{-1}(0)$ y $b \in f^{-1}(1)$, usando la definición de arco-conexo, entonces existe una función continua $\gamma \colon [0,1] \to S$ tal que $\gamma(0) = a$ y $\gamma(1) = b$, entonces $f \circ \gamma \colon [0,1] \to \{0,1\} \subseteq \mathbb{R}$ es continua y no constante por construcción, pero eso contradice (2.2.2) para el conexo $[0,1]$, por lo tanto S si es conexo. ■

Teorema 2.2.5 — Teorema del valor intermedio. Sea D un conjunto arco-conexo y $f \colon D \to \mathbb{R}$ continua, dados $a, b \in D$ tal que $f(a) \leq f(b)$ y $t \in (f(a), f(b))$ entonces existe $c \in D$ tal que
$$f(c) = t$$

Demostración. Por definición, existe una función continua $\gamma \colon [0,1] \to D$ tal que $f(0) = a$, $f(1) = b$ y $\gamma([0,1]) \subseteq D$ entonces $\gamma \circ f \colon [0,1] \to \mathbb{R}$ y aplicas el teorema del valor intermedio para una variable. ■

Ejercicio 2.5 Prueba usando el teorema del valor intermedio, que la ecuación
$$x^2 + 2y^2 = e^{(z-1/2)^2 \cos(e^{-\sin(y/(x+2))})}$$
tiene solución en $B_2(\vec{0})$. Hint: transforma el problema en una función $f(x,y,z) = 0$ y aplica el teorema del valor intermedio usando $(0,0,1/2), (0,1,1/2)$. ■

Ejercicio 2.6 Sea $S \subseteq \mathbb{R}^3$ la esfera unitaria, $f \colon S \to \mathbb{R}$, prueba que existe $x \in S$ tal que $f(x) = f(-x)$. Hint: aplica el valor intermedio para $g(x) = f(x) - f(-x)$ ■

2.3. Continuidad uniforme.

Definición 2.3.1 — Continuidad uniforme. Una función $f \subseteq D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ es uniformemente continua si para toda ε existe una δ tal que
$$\|x-y\| \leq \delta \Rightarrow \|f(x) - f(y)\| \leq \varepsilon$$

■ **Ejemplo 2.13** Consideremos una función $f \colon D \subseteq \mathbb{R}^n \to \mathbb{R}^m$, si existe una $k > 0$ tal que para todo $x, y \in D$ se cumple que
$$\|f(x) - f(y)\| \leq k\|x-y\|$$
entonces f es uniformemente continua. ■

■ **Ejemplo 2.14** Si f es uniformemente continua entonces f es continua. ■

Proposición 2.3.1 Si K es compacto y $f\colon K \to \mathbb{R}$ es continua entonces f es uniformemente continua.

Demostración. Para cada $x \in K$, como f es continua en x para cada $\varepsilon > 0$ entonces existe $\delta_x > 0$ tal que

$$f(B(x,\delta_x)) \subseteq B(f(x),\varepsilon/2)$$

Tenemos una cubierta abierta $\{B(x,\delta_x/2)\}_{x\in K}$ de K, como K es compacto, entonces existen $x_1,\cdots,x_n$ tal que $\{B(x_i,\delta_{x_i}/2)\}_{i=1}^{n}$ es cubierta abierta en K. Ahora sea $\delta := \text{mín}_{1\leq i\leq n}(\delta_{x_i}/2)$, y tomemos $x,y \in K$ tal que $d(x,y) < \delta$. Luego, existe $1 \leq i \leq n$ tal que

$$x \in B(x_i,\delta_{x_i}/2) \subseteq B(x_i,\delta_{x_i})$$

entonces

$$y \in B(x,\delta) \subseteq B(x_i,\delta_{x_i})$$

por lo tanto

$$d(f(x),f(y)) \leq d(f(x),f(x_i)) + d(f(x_i),f(y)) < \varepsilon/2 + \varepsilon/2 = \varepsilon$$

■

Capítulo 3

El concepto de Derivada

3.1. Derivada de funciones paramétricas.

En esta sección, se considera funciones de la forma $f\colon [a,b] \to \mathbb{R}^n$, sea $C = f([a,b])$ la imagen de la función a la que denominaremos por abuso de notación como una curva. Vamos a motivar la definición de la derivada usando la misma construcción de la derivada en una variable. Fijemos $t_0 \in [a,b]$ entonces para cada $h > 0$ tenemos que la recta

$$P_\lambda = f(t_0) + \lambda \frac{f(t+h) - f(t)}{h},\ \lambda \in \mathbb{R}$$

es una secante que pasa por $f(t_0) = P_0$ y $f(t+h) = P_h$. Geométricamente, si variamos h y lo acercamos a cero, la recta secante tiende a convertirse e una recta tangente, entonces la derivada $f'(t_0)$ debe ser un vector tal que

$$Q_\lambda = f(t_0) + \lambda f'(t_0),\ \lambda \in (a,b)$$

sea una recta tangente.

Definición 3.1.1 — Derivada 1. Sea $f\colon [a,b] \to \mathbb{R}$ y $t_0 \in (a,b)$ la derivada de f en t_0 es

$$f'(t_0) := \lim_{h \to 0} \frac{f(t+h) - f(t)}{h}$$

La definición anterior se puede ver equivalentemente como sigue, para cada ε existe una δ tal que cuando $\|h\| < \delta$ entonces

$$\left\| f(t+h) - f(t) - f'(t)h \right\| \leq \varepsilon \|h\| \tag{3.1}$$

Con esto, la derivada se puede reinterpretar como la mejor manera de aproximar linealmente f al rededor de $B_h(t_0)$ para algun h demasiado pequeño.

Ejercicio 3.1 Muestra que las únicas funciones lineales $L\colon R \to R^n$ son de la forma $L(h) = vh$, donde $v \in R^n$ es único determinado. ■

Proposición 3.1.1 Si la derivada de una función paramétrica existe entonces este es único.

Ejercicio 3.2 Probar la proposición (3.1.1) ■

Vamos a profundizar la interpretación de manera geométrica, sea $p := f(t_0)$ y usando la observación (1.1) tenemos que en $T_p\mathbb{R}^n$ hay una recta generada por el vector unitario $u := \frac{f'(t_0)}{\|f'(t_o)\|}$, denotemoslo como T_pC, entonces, este es un subespacio de dimensión 1 y si definimos $L\colon T_pC \to T_pC$ como $L(h) := f'(t)h$, esta es una función lineal, y al sustituirlo en (3.1), tenemos

$$\|\triangle_h f(t) - L(h)\| \leq \varepsilon \|h\|$$

Al vector u lo denominamos como el tangente unitario y $L(h)$ es la diferencial de f en t_0, y mide la dirección en la que aproxima la derivada $f'(t_0)$ con respecto a $f(t+h)$.

Proposición 3.1.2 Supongamos que $f\colon [a,b] \to \mathbb{R}^n$ es de la forma $f(t) = (f_1(t), \cdots, f_n(t))$ y $f'(t_0)$ existe, entonces $f_i'(t_0)$ existe y

$$f'(t_0) = (f_1'(t_0), \cdots, f_n'(t_0))$$

Demostración. Se sigue de (2.1.2) obteniendo

$$f'(t_0) = \lim_{h\to 0} \frac{f(t_0+h) - f(t_0)}{h} = \lim_{h\to 0} \left(\frac{f_i(t_0+h) - f_i(t_0)}{h} \right) = (f_i'(t_0))$$

■

■ **Ejemplo 3.1** Considera $f(t) = (\cos t, \sin t)$ entonces $f'(t) = (-\sin t, \cos t)$. ■

Proposición 3.1.3 Si f es derivable entonces es continua.

Demostración. Se sigue de la siguiente identidad y la definición de la derivada.

$$\|f(t) - f(t_0)\| = \left\| \frac{f(t) - f(t_0)}{t - t_0} \right\| |t - t_0|$$

■

Ejercicio 3.3 Prueba las siguientes identidades:

1. $(f+g)'(t) = f'(t) + g'(t)$.
2. $(\lambda f)'(t) = \lambda f'(t)$ para toda $\lambda \in \mathbb{R}$.
3. Si $g\colon [a,b] \to [c,d]$ entonces

$$(f\circ g)'(t) = f'(g(t))g'(t).$$

■

Teorema 3.1.4 — Teorema del valor medio vectorial. Para una función continua $f\colon [a,b] \to \mathbb{R}^n$ tal que es diferenciable en (a,b), entonces existe un número $c \in (a,b)$ tal que:

$$\|f(b) - f(a)\| \leq (b-a) \|f'(c)\|$$

Demostración. Sea $g\colon [a,b] \to \mathbb{R}$ definido como

$$g(t) = (f(b) - f(a)) \cdot f(t)$$

Notemos que

$$g'(t) = (f(b) - f(a)) \cdot f'(t)$$

como g es continua en $[a,b]$ y diferenciable en (a,b), por el teorema del valor medio de una variable, existe $c \in (a,b)$ tal que

$$g(b) - g(a) = g'(c)(b-a)$$

como $g(b) - g(a) = \|f(b) - f(a)\|^2$ tenemos

$$\|f(b) - f(a)\|^2 = (b-a)\|g'(c)\| \leq (b-a)\|f'(c)\| \|f(b) - f(a)\|$$

implicando el resultado deseado. ■

3.1.1. Introducción a las curvas diferenciables y su longitud.

Definición 3.1.2 — Curva Diferenciable. Una curva $\gamma\colon I\subseteq\mathbb{R}\to\mathbb{R}^n$ es diferenciable si existe $f'(t_0)$ para toda $t_0\in int(I)$ y f' es continua.

Se entiende como una reparametrización de una curva γ como una función $\gamma\circ\phi$ donde $\phi: I\subseteq\mathbb{R}\to J\subseteq\mathbb{R}$ es un homeomorfismo.

Ejercicio 3.4 Prueba que toda curva definido sobre un abierto (a,b) se puede reparametrizar a una curva de la forma $\gamma\colon (0,1)\to\mathbb{R}$. ¿Será que toda curva definida sobre $\mathbb{R}$ se puede reparametrizar a una curva de la forma $\gamma\colon [0,1]\to\mathbb{R}$? ■

Por simplicidad, podemos asumir que la curva es de la forma $\gamma\colon [0,1]\to\mathbb{R}$. Vamos a tomar algunas construcciones importantes para hablar de la geometría de una curva diferenciable.

Logitud de la curva. Sea $\mathscr{P}: t_0=0<t_1<t_2<\cdots<t_n=1$ una partición de $[0,1]$ entonces su longitud se obtiene como

$$l(\gamma,\mathscr{P})=\sum_{i=0}^{n}\|\gamma(t_{i+1})-\gamma(t_i)\| \tag{3.2}$$

Como γ es continuamente diferenciable, podemos aplicar el siguiente resultado:

Proposición 3.1.5 Si $\gamma\colon [a,b]\to\mathbb{R}^n$ es continuamente diferenciable entonces existe $c\in(a,b)$ tal que

$$\|\gamma(b)-\gamma(a)\|\le\left\|\gamma'(c)\right\||b-a|$$

Demostración. Definimos

$$\int_a^b\gamma(t)dt:=\left(\int_a^b\gamma_i(t)dt\right)$$

Entonces obtenemos del teorema fundamental del cálculo en una variable que:

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

Entonces

$$\|\gamma(b)-\gamma(a)\|=\left\|\int_a^b\gamma'(t)dt\right\|\le\int_a^b\left\|\gamma'(t)\right\|dt$$

Pero γ' es continua entonces $\|\gamma'\|$ es continua, y por el teorema del valor medio para integrales, existe $c\in(a,b)$ tal que

$$\int_a^b\left\|\gamma'(t)\right\|dt=\left\|\gamma'(c)\right\||b-a|$$

sustituyendo la igualdad anterior obtenemos el resultado deseado. ■

Aplicando el resultado anterior en (3.2) obtenemos

$$l(\gamma,\mathscr{P})\le\sum_{i=0}^{n}\left\|\gamma'(c_i)\right\||t_{i+1}-t_i|,\ c_i\in(t_{i+1}-t_i)$$

De esta manera si queremos definir la longitud de la curva como

$$l(\gamma)=\sup\{l(\gamma,\mathscr{P})|\mathscr{P}\text{ es una partición de }[0,1]\}$$

para curvas diferenciales obtenemos

$$l(\gamma)=\int_0^1\left\|\gamma'(t)\right\|dt$$

■ **Ejemplo 3.2** Consideremos la curva $\gamma(t) = (\cos(2\pi t), \sin(2\pi t))$ que dibuja una circunferencia en el sentido contrario a las manecillas del reloj, entonces

$$l(\gamma) = \int_0^1 \|(-2\pi\sin(2\pi t), 2\pi\cos(2\pi t))\|\, dt = \int_0^1 2\pi dt = 2\pi$$

■

Dado $t_0 \in [0,1]$ definimos la longitud del arco como

$$l_\gamma(t) := \int_0^t \|\gamma'(t)\|\, dt =: s(t)$$

Esta función mide la longitud de la curva en relación a como avanza por el intervalo $(0,1)$, si interpretamos un momento el intervalo $(0,1)$ como una unidad de tiempo, obtenemos las siguientes interpretaciones

1. $s(0) = 0$, significa que no hay distancia recorrida en el tiempo 0.
2. $s(1) = l(\gamma)$ significa que al final del tiempo $t = 1$, se ha recorrido toda la curva.

Uno de los objetivos con mayor sentido geométrico para una curva, es poder reparametrizar por longitud de arco, es decir, que el parametro del dominio está en función de la longitud, esto se hace bajo los siguientes principios; derivando, por el teorema fundamental del calculo obtenemos

$$\frac{ds}{dt} = \|\gamma'(t)\| > 0$$

por el teorema de la función inversa, es posible obtener t en función de s, obteniendo $t = t(s)$, permitiendo reparametrizar la curva como sigue $\gamma(t) = \gamma(t(s))$, en esta situación, notemos que, por regla de la cadena

$$\frac{d}{ds}\gamma(t(s)) = \frac{d\gamma}{dt}\frac{dt}{ds} = \frac{\frac{d\gamma}{dt}}{\frac{ds}{dt}} = \frac{\gamma'}{\|\gamma\|}$$

y entonces $\left\|\frac{d\gamma}{ds}\right\| = 1$ para toda $s \in dom(\gamma \circ t(s))$. A este procedimiento se le conoce como reparametrización por longitud de arco.

Definición 3.1.3 — Curva parametrizada por longitud de arco. Una curva diferenciable $\gamma(t)$ es parametrizada por longitud de arco si $\|\gamma'(t)\| = 1$ para toda t.

■ **Ejemplo 3.3** Sea $\gamma(t) = (3t-1, 4t+2)$, vamos a parametrizar la curva por longitud de arco, notemos

$$s(t) = \int_0^t \|(3,4)\|\, dt = 5t \Rightarrow\; t(s) = \frac{1}{5}s$$

Al reparametrizar

$$\gamma(t) = \gamma(t(s)) = \gamma_1(s) = (\frac{3}{5}s - 1, \frac{4}{5}s + 2)$$

Entonces $\gamma_1(0) = (-1,2)$ y $\gamma_1(2) = (1/5, 18/5)$, entonces $\gamma_1(2)$ tiene la propiedad de que $d(\gamma(2), \gamma(1)) = 2$ es decir ha recorrido 2 unidades, que son parte de la longitud de la curva. ■

3.2. Derivada direccional y sus propiedades.

Antes de extenderlo a funciones escalares, requerimos este concepto geométrico.

Definición 3.2.1 — Derivada direccional. Sea $n \in \mathbb{R}^n$ y $f\colon D \subseteq \mathbb{R}^n \to \mathbb{R}$ una función escalar, definimos la derivada direccional de f en $a \in D$ como

$$\partial_n f(a) := \lim_{h \to 0} \frac{f(a+hn) - f(a)}{h}$$

Ejercicio 3.5 Usando la definición, prueba que

$$\partial_{\lambda \cdot n+m} f(a) = \lambda \partial_n f(a) + \partial_m f(a)$$

■

■ **Ejemplo 3.4** Una manera practica de calcular $\partial_n f(a)$ es como sigue, definimos $g(t) = f(a+tv)$, entonces de la definición, obtenemos que $g'(0) = \partial_v f(a)$. Por ejemplo, sea $f(a) = \|a\|$, entonces al definir $g(t) = \|a+tv\|$, podemos elevar al cuadrado obteniendo

$$g^2(t) = \langle a+tv, a+tv \rangle = \|a\|^2 + 2t\langle a, v \rangle + t^2 \|v\|^2$$

Derivando con respecto a t, obtenemos

$$2g'(t)g(t) = 2\langle a, v \rangle + 2t\|v\|^2$$

Esto implica que

$$\partial_v f(a) = \langle \frac{a}{\|a\|}, v \rangle$$

cuando $a \neq 0$. ■

Fijemos $a \in D$, y consideremos $T_a\mathbb{R}^n$ como en (1.1) entonces tenemos que

$$\partial_{(\cdot)} f(a) \colon T_a\mathbb{R}^n \to \mathbb{R}$$

es una funcional lineal. En esta situación ya que $\dim T_a\mathbb{R}^n = n$ entonces al fijar una base $u_1, \cdots, u_n \in T_a\mathbb{R}^n$ obtenemos que para toda $w = \alpha_1 u_1 + \cdots + \alpha_n u_n \in T_a\mathbb{R}^n$ se cumple que

$$\partial_w f(a) = \sum_{i=1}^{n} \alpha_i \partial_{u_i} f(a)$$

Asi que esta propiedad indica que una condicion necesaria para la existencia de la derivada en cualquier dirección es si existe en una base de dimensión n, la suficiencia la probaremos cuando definamos la diferenciabilidad en la siguiente sección. Usando la derivada direccional, tenemos que mide la variación de f sobre la recta centrada en a y dirección n.

Proposición 3.2.1 Si $b \colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ es bilineal y $f, g \colon D \subseteq \mathbb{R}^n \to \mathbb{R}$ tienen derivada direccional en $a \in D$ entonces

$$\partial_n(b(f,g))(a) = b(\partial_n f(a), g(a)) + b(f(a), \partial_n g(a))$$

Demostración. Definamos $g(t) = b(f(a+tn), g(a+tn))$, entonces notemos que

$$\begin{aligned} g(t) - g(0) &= b(f(a+tn), g(a+tn)) - b(f(a+tn), g(a)) + b(f(a+tn), g(a)) - b(f(a), g(a)) \\ &= b(f(a+tn), (g(a+tn) - g(a))) + b(f(a+tn) - f(a), g(a)) \\ \frac{g(t) - g(0)}{t} &= b\left(f(a+tn), \frac{g(a+tn) - g(a)}{t}\right) + b\left(\frac{f(a+tn) - f(a)}{t}, g(a)\right) \end{aligned}$$

Aplicando el limite y usando la propiedad de que $g'(0) = \partial_n(b(f,g))(a)$, obtenemos el resultado deseado. ■

3.3. Gradiente de una función escalar.

Definición 3.3.1 — Derivada parcial. Sea $f \subseteq D \subseteq \mathbb{R}^n \to \mathbb{R}$, definimos la derivada parcial en x_i como

$$\partial_{x_i} f := \partial_{e_i} f$$

en donde e_i es el unico vector unitario tal que la proyección de $x = (x_1, \cdots, x_n)$ sobre e_i es x_i.

Definición 3.3.2 — Funcion diferenciable. Una función $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ es diferenciable en $a \in D$ si existen:

1. Una función lineal $L_a : \mathbb{R}^n \to \mathbb{R}$.
2. Una función escalar $\mathscr{E}(a,h) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, tal que

$$f(a+h) = f(a) + L_a(h) + \|h\| \mathscr{E}(a,h), \ \|h\| < r$$

De manera de que $\lim_{\|h\| \to 0} E(a,v) = 0$. A la función lineal L_a se le conoce como diferencial.

Algunas notaciones importantes:

1. A la diferencial L_a lo denotaremos como $d_a f$.
2. Si fijamos $\{e_i\} \subseteq \mathbb{R}^n$ la base estándar, entonces la matriz de representación de $d_a f$ es conocido como el gradiente de f y es denotado como $\nabla f(a)$.

Teorema 3.3.1 — Derivada total vs Derivada Parcial. Si f es diferenciable en a entonces para cada $n \in \mathbb{R}^n$, existe la derivada parcial $\partial_n f(a)$ y además

$$d_a f(n) = \partial_n f(a)$$

Demostración. El caso trivial es cuando $n = 0$. Supongamos $n \neq 0$ y usando la notación de la definición 3.4.1 tenemos

$$f(a+h) = f(a) + L_a(h) + \|h\| \mathscr{E}(a,h), \ \|h\| < r$$

Consideremos la recta $h = tn$, con $\|tn\| < r$, entonces tenemos:

$$\frac{f(a+tn) - f(a)}{t} = L_a(n) + \frac{|t| \, \|n\|}{t} \mathscr{E}(a,tn)$$

Cuando $t \to 0$ se tiene $\frac{|t|}{t} \to \pm 1$ y $\mathscr{E}(a,tn) \to 0$, entonces

$$\partial_n f(a) = d_a f(n)$$

■

Del teorema anterior obtenemos que

$$\nabla f(a) = (\partial_{x_1} f(a), \cdots, \partial_{x_n} f(a))$$

Puesto que la base estandar $\{e_i\}_{i=1}^n$ es una base ortonormal, por el teorema de representación de Riez (A.3.4), tenemos

Corolario 3.3.2 Para toda dirección $n = (a_1, \cdots, a_n)$ descrita en su base estándar, si f es diferenciable en a entonces

$$\partial_n f(a) = \langle \nabla f(a), n \rangle$$

Proposición 3.3.3 Si f es diferenciable en a, entonces f es continua en a.

■ **Ejemplo 3.5 — Derivable no implica diferenciable.** La función

$$f(x,y) = \frac{xy^2}{x^2+y^4} \text{ si } x \neq 0, f(0,y) = 0$$

existen $\partial_x f(0,0)$ y $\partial_y f(0,0)$, pero f no es continua en cero y por lo tanto f no es diferenciable en 0. ■

Proposición 3.3.4 Sea $f\colon D \subseteq \mathbb{R}^n \to \mathbb{R}$ una función, $a \in \partial D$, $\{e_i\}$ una base ortonormal de $\mathbb{R}^n$ tal que existe $\partial_{e_i} f(a)$ y son continuas en un abierto U, entonces f es diferenciable en a.

Demostración. Definamos $\mathscr{E}(a,h)$ como

$$f(a+h) - f(a) - \langle \nabla f(a), h \rangle = \|h\| \mathscr{E}(a,h)$$

veamos que $\mathscr{E}(a,h) \to 0$ cuando $\|h\| \to 0$. Sea $\lambda := \|h\|$ entonces $h = \lambda u$ donde $\|u\| = 1$, descomponiendo u en combinacion lineal de la base ortonormal

$$u = \omega_1 e_1 + \cdots + \omega_n e_n$$

Definamos $v_0 = 0$, $v_k = v_{k-1} + \omega_k e_k$ y $v_n = u$, entonces

$$f(a+h) - f(a) = g(a+\lambda u) - f(a) = \sum_{k=1}^{n} (f(a+\lambda v_k) - f(a+\lambda v_{k-1}))$$

Usando el teorema del valor medio

$$f(a+\lambda v_k) - f(a+\lambda v_{k-1}) = f(b_k + \lambda \omega_k e_k) - f(b_k) = \lambda \omega_K \partial_k f(c_k)$$

donde $b_k = a + \lambda v_{k-1}$ y $c_k \in [b_k, b_k + \lambda \omega_k e_k]$. Esto nos dice

$$f(a+h) - f(a) = \lambda \sum_{k=1}^{n} \partial_k f(c_k) \omega_k$$

y entonces al sustituir

$$\|h\| \mathscr{E}(a,h) = \sum_{k=1}^{n} (\partial_k f(c_k) - \partial_k f(a)) u_k$$

Como $c_k \to a$ cuando $\|h\| \to 0$ entonces $\mathscr{E}(a,h) \to 0$. ■

3.4. Derivada total y el Jacobinano.

Definición 3.4.1 — Funcion diferenciable. Una función $f\colon D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ es diferenciable en $a \in D$ si existen:

1. Una función lineal $L_a\colon \mathbb{R}^n \to \mathbb{R}^m$.
2. Una función $\mathscr{E}(a,h)\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^m$, tal que

$$f(a+h) = f(a) + L_a(h) + \|h\| \mathscr{E}(a,h), \ \|h\| < r$$

De manera de que $\lim_{\|h\| \to 0} E(a,v) = 0$. A la función lineal L_a se le conoce como diferencial.

De manera similar al caso escalar tenemos las convenciones

1. La diferencial L_a es denotado como $d_a f$.

2. Si se fija una base de $\mathbb{R}^n$ entonces la matriz de representación $d_a f$ se le conoce como Jacobinano y es denotado como

$$Df(a) = J(a) = \frac{\partial f}{\partial x}(a)$$

Si $f = (f_1, \cdots, f_m)$, usando que

$$\nabla f_j(a) = (\partial_{x_i} f_j(a))$$

y combinando con (A.2.1) obtenemos

Proposición 3.4.1 Si f es diferenciable en a y $f = (f_1, \cdots, f_n)$ en coordenadas cartesianas, entonces existen las $\partial_{x_i} f(a)$ y se tiene que

$$Df(a) = (\partial_{x_j} f_i(a))$$

Proposición 3.4.2 Si f es diferenciable en a entonces f es continuo en a.

Definición 3.4.2 — Función continuamente diferenciable. Decimos que una función $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}^m$ es continuamente diferenciable en el abierto Ω si para toda $a \in \Omega$, f es diferenciable en a y la función

$$\begin{array}{rccc} df: & \Omega & \to & L(\mathbb{R}^n, \mathbb{R}^m) \cong \mathbb{R}^{nm} \\ & a & \to & d_a f \end{array}$$

es contínua.

3.5. Regla de la cadena en varias variables.

Una consecuencia de A.1.8 aplicado a las diferenciales es el siguiente:

Teorema 3.5.1 — Regla de la cadena. Sean $f : D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ y $g : S \subseteq \mathbb{R}^n \to \mathbb{R}^s$, tomemos $a \in D$ tal que $g \circ f(a)$ existe. Si f es diferenciable en a y g es diferenciable en $f(a)$ entonces $g \circ f$ es diferenciable en a, además:

$$d_a(g \circ f) = d_{f(a)} g \circ d_a f$$

y

$$D(g \circ f)(a) = Dg(f(a)) \circ Df(a)$$

Cambio de coordenadas y ecuaciones diferenciales parciales. Definimos un cambio de coordenadas en $\mathbb{R}^n$ como una función biyectiva

$$\phi : D \subseteq \mathbb{R}^n \to T \subseteq \mathbb{R}^n$$

que es diferenciable en cada punto de D, en este tipo de situaciones, es común refererirse al cambio de coordendas como cambio de variables, digamos

$$u = \phi(x),\ u = (u_1, \cdots, u_n), x = (x_1, \cdots, x_n)$$

En esta situación, la regla de la cadena suele ser una importante herramienta para convertir los fenómenos descritos en la variable x a fenómenos descritos en la variable u. Si $f : T \subseteq \mathbb{R}^n \to \mathbb{R}$ es diferenciable, digamos que queremos calcular $a = \phi(b) \in T \subseteq \mathbb{R}^n$, entonces por regla de la cadena aplicado a $f \circ \phi$

$$\nabla_u f(a) = \nabla_x f(\phi(b)) D\phi(b)$$

Es decir, para cada $i = 1, \cdots, n$ se tiene

$$\partial_{u_i} f(a) = \sum_{k=1}^{n} \partial_{x_k} f(\phi(b)) \partial_{u_i} x_k(b)$$

■ **Ejemplo 3.6** Una ecuación diferencial parcial de primer orden lineal muy común es de la siguiente forma

$$\xi_1 \partial_x f + \xi_2 \partial_y f = 0$$

Sea $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ una matriz invertible. Vamos a encontrar un cambio de coordenadas de lineal tal que simplifique la ecuación parcial a una ecuación de la forma más simple. Para ello introducimos las coordenadas

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}$$

Entonces

$$\nabla_{(x,y)} f = A \cdot \nabla_{(u,v)} f$$

aplicando el producto escalar con $\xi = (\xi_1, \xi_2)$, entonces

$$\partial_\xi f = \langle \xi, \nabla_{(x,y)} f \rangle = \xi^T \cdot A \cdot \nabla_{(u,v)} f$$

entonces al sustituir

$$(\xi_1 a_{11} + \xi_2 a_{21}) \partial_u f + (\xi_1 a_{12} + \xi_2 a_{22}) \partial_v f = 0$$

proponemos que el primer coeficiente sea cero, obteniendo el sistema

$$\begin{cases} \xi_1 a_{11} + \xi_2 a_{21} = 0 \\ a_{11} a_{22} - a_{21} a_{12} = 1 \end{cases}$$

■

Ahora para el caso de ecuaciones de segundo grado, tenemos de la regla de la cadena

$$\partial_{u_j} f_i = \sum_{k=1}^{n} \partial_{x_k} f_i \partial_{u_j} x_k$$

al derivar con respecto a u_s obtenemos

$$\partial^2_{u_s u_j} f_i = \sum_{k=1}^{n} \left(\partial_{u_s} (\partial_{x_k} f_i) \partial_{u_j} x_k + \partial_{x_k} f_i \partial^2_{u_s u_j} x_k \right)$$

usando regla de la cadena nuevamente, tenemos $\partial_{u_s}(\partial_{x_k} f_i) = \sum_{h=1}^{n} \partial^2_{x_h x_k} f_i \partial_{u_s} x_h$ y al sustituir

$$\partial^2_{u_s u_j} f_i = \sum_{k,h=1}^{n} \partial^2_{x_h x_k} f_i \partial_{u_s} x_h \partial_{u_j} x_k + \sum_{k=1}^{n} \partial_{x_k} f_i \partial^2_{u_s u_j} x_k$$

A continuación presentaremos los cambios de coordenadas más comúnes y ejemplos de como la fórmula anterior se aplica para transformar ecuaciones diferenciales

■ **Ejemplo 3.7 — Coordenadas polares.** Sea $D = [0, 2\pi] \times [0, \infty)$ y definimos las coordenadas polares como el sistema

$$\begin{cases} x &= r\cos\theta \\ y &= r\sin\theta \end{cases}$$

este sistema induce una función diferenciable

$$\begin{array}{cccc} \phi: & D & \to & \mathbb{R}^2 \\ & (\theta, r) & \to & (x,y) \end{array}$$

Asi, se tiene la derivada total como

$$D\phi = \begin{pmatrix} \cos\theta & \sin\theta \\ -r\sin\theta & r\cos\theta \end{pmatrix}$$

Entonces la ecuación de Laplace

$$\partial^2_{xx} f + \partial^2_{yy} f = 0$$

en coordenadas polares es

$$\frac{1}{r}\partial_r\left(r\partial_r f\right) + \frac{1}{r^2}\partial^2_{\theta\theta} f = 0$$

■

■ **Ejemplo 3.8 — Coordenadas esfericas.** Sea $D = [0,\infty) \times [0,\pi] \times [0,2\pi]$, definimos las coordenadas esféricas por el sistema

$$\begin{cases} x &= r\sin\theta\cos\phi \\ y &= r\sin\theta\sin\phi \\ z &= r\cos\theta \end{cases}$$

este sistema induce una función diferenciable

$$\begin{array}{cccc} H: & D & \to & \mathbb{R}^3 \\ & (r,\theta,\phi) & \to & (x,y,z) \end{array}$$

Asi, se tiene la derivada total como

$$DH = \begin{pmatrix} \sin\theta\phi & \sin\theta\sin\phi & \cos\theta \\ r\cos\theta\cos\phi & r\cos\theta\sin\phi & -r\sin\theta \\ -r\sin\theta\sin\phi & r\sin\theta\cos\phi & 0 \end{pmatrix}$$

En esta situación la ecuación de Laplace

$$\partial^2_{xx} f + \partial^2_{yy} f + \partial^2_{zz} f = 0$$

se puede escribir en coordenadas polares como

$$\frac{1}{r}\partial^2_{rr}(rf) + \frac{1}{r^2\sin\theta}\partial_\theta\left(\sin\theta\,\partial_\theta f\right) + \frac{1}{r^2\sin^2\theta}\partial^2_{\phi\phi} f = 0$$

■

■ **Ejemplo 3.9 — Regla de Integración de Leibniz.** La regla de integración de Leibniz es una técnica utilizada para resolver una familia de integrales parametrizadas, permitiendo su aplicación en una integral específica. Esta regla establece que si tenemos una integral de la forma:

$$\phi(x) = \int_{a(x)}^{b(x)} f(x,t)\,dt$$

con $a(x) \le t \le b(x)$, f y f_x continuos en el plano t,x, entonces podemos determinar $D\phi(x)$ como sigue:

1. Define $F(a,b,x) := \int_a^b f(x,t)dt$.

2. Como f es continuo, por el teorema fundamental del cálculo, tenemos

$$\frac{\partial F}{\partial b} = f(x,b), \ \frac{\partial F}{\partial a} = -f(x,a)$$

3. Usando regla de la cadena tenemos que:

$$\phi'(x) = DF \circ (a(x), b(x), x) = f(x, b(x))b'(x) - f(x, a(x))a'(x) + \frac{\partial F}{\partial x} \tag{3.3}$$

■

Proposición 3.5.2 — Regla de integración de Leibniz. Con las hipótesis de (3.9), tenemos

$$\frac{d}{dx}\left(\int_{a(x)}^{b(x)} f(x,t)dt\right) = f(x,b(x)) \cdot b'(x) - f(x,a(x)) \cdot a'(x) + \int_{a(x)}^{b(x)} \frac{\partial f}{\partial x}(x,t)dt$$

Demostración. Por (3.3), basta verificar

$$\frac{d}{dx}\left(\int_a^b f(x,t)dt\right) = \int_a^b f_x(x,t)dt$$

Sea $\psi(x) := \int_a^b f(x,t)dt$. Para cada (x_0,t) tenemos, para toda $\varepsilon > 0$ existe $\delta > 0$ tal que si $|h| < \delta$ entonces

$$|f(x_0+h,t) - f(x_0,t) - f_x(x_0,t)h| \leq \varepsilon|h|/(b-a)$$

Usando lo anterior en lo siguiente, obtenemos

$$\begin{aligned} \left|\psi(x_0+h) - \psi(x_0) - \left(\int_a^b f_x(x_0,t)dt\right)h\right| &= \left|\int_a^b (f(x_0+h,t) - f(x_0,t) - f_x(x_0,t)h)dt\right| \\ &= \int_a^b |f(x_0+h,t) - f(x_0,t) - f_x(x_0,t)h|\, dt \\ &= \frac{\varepsilon|h|(b-a)}{b-a} = \varepsilon|h| \end{aligned}$$

obteniendo la derivada deseada. ■

El siguiente ejercicio es una aplicación muy útil en ecuaciones integro-diferenciales y precursor del cálculo fracional.

Ejercicio 3.6 — Fórmula de Cauchy para integraciones iteradas. Sea f continua, entonces para cada natural $n > 0$:

$$\int_a^x \int_a^{s_1} \cdots \int_a^{s_{n-1}} f(s_n)ds_n \cdots ds_2 ds_1 = \frac{1}{(n-1)!}\int_a^x (x-t)^{n-1} f(t)dt$$

■

Otros ejemplos útiles para mencionar, derivados de la fórmula de integración de Leibniz son los siguientes:

1. Para resolver $I(k) = \int_0^1 \frac{x^k - 1}{\ln x} dx$, derivamos con respecto a k obteniendo

$$\frac{dI}{dk} = \int_0^1 \frac{x^k \ln x}{\ln x} dx = \int_0^1 x^k dx = \frac{1}{k+1}$$

entonces, al integrar, por teorema fundamental del cálculo obtenemos

$$I(k) = I(0) + \ln(k+1) = \ln(k+1)$$

Un análogo al teorema de valor medio. En general no existe un teorema de valor medio para funciones de varias variables, sin embargo, es posible construir un análogo usando las siguientes observaciones.

Sea $f\colon \mathbb{R} \to \mathbb{R}$ continuamente diferenciable y fijemos $x_0 \in domf$, entonces por el teorema fundamental y por el cambio de variables $u = x_0 + th$ con $h > 0$ fijo albitrario, entonces:

$$f(x+h) - f(x) = \int_x^{x+h} f'(u)du = h\int_0^1 f'(x+th)dt$$

Proposición 3.5.3 — Teorema de valor medio para varias variables. Sea $f\colon U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ continuamente diferenciable, $x \in U$ y $h \in \mathbb{R}^n$ tal que $[x, x+h] \subseteq U$ entonces:

$$f(x+h) - f(x) = \left(\int_0^1 Df(x+th)dt\right) \cdot h$$

donde $\cdot$ es la multiplicación matricial.

Demostración. Sea $\gamma\colon [0,1] \to U$ definido como $\gamma(t) = x + th$ y sea $g := f \circ \gamma$ entonces

$$f(x+h) - f(x) = g(1) - g(0) = \int_0^1 g'(t)dt$$

pero si $f = (f_1, \cdots, f_m)$ entonces $g'(t) = (\frac{d}{dt}(f_i(x+th)))$, luego por regla de la cadena:

$$\frac{d}{dt}(f_i(x+th)) = \sum_{j=1}^{n} \frac{\partial f_i}{\partial x_j}(x+th)h_j$$

entonces

$$\int_0^1 g'(t)dt = \int_0^1 \left(\sum_{j=1}^{n} \frac{\partial f_i}{\partial x_j}(x+th)h_j\right) dt = \left(\sum_{j=1}^{n} \frac{\partial f_i}{\int}{}^{1}_{0} \partial x_j(x+th)dt \cdot h_j\right)$$

■

3.6. Geometría de Curvas Diferenciables

Aqui se considera curvas diferenciables de la forma $\gamma\colon [0,1] \to \mathbb{R}^n$ que sean parametrizadas por longitud de arco, es decir $\|\gamma'(s)\| = 1$. En esta situación, denotamos

$$T(s) := \gamma'(s)$$

y es conocido como el tangente unitario de la curva γ. En esta situación, tenemos

$$\langle T(s), T(s)\rangle = 1$$

Derivando con respecto a s, obtenemos

$$\langle T'(s), T(s)\rangle = 0$$

Mostrando que $T'(s)$ es ortogonal a $T(s)$ y es conocido como el primer marco de referencia de Frenet.

Definición 3.6.1 — Normal y plano osculador. Sea γ una curva parametrizada por longitud de arco, definimos:

1. el normal unitario (cuando $\gamma''(s) \neq 0$) como
$$N(s) := \frac{\gamma''(s)}{\|\gamma''(s)\|} = \frac{T'(s)}{\|T'(s)\|}$$
2. la curvatura como
$$\kappa(s) := \|\gamma''(s)\|$$
3. Si $\gamma''(s) \neq 0$, definimos el plano osculador $\Pi(s)$ a la curva en $\gamma(s)$ como el plano anclado en $\gamma(s)$ y generado por las direcciones $T(s), N(s)$

El plano osculador contiene la mejor circunferencia que se aproxima a la curva, en el punto $\gamma(s)$. Procedemos por casos.

Caso $\mathbb{R}^2$. Como la normal es unitaria, obtenemos
$$\langle N'(s), N(s)\rangle = 0$$
entonces $N'(s) = \alpha(s)T(s)$ para alguna función de una variable. Por otro lado, usando la identidad
$$\langle T(s), N(s)\rangle = 0$$
diferenciando, obtenemos
$$\kappa(s)\langle N(s), N(s)\rangle + \langle T(s), \alpha(s)T(s)\rangle = 0$$
obteniendo
$$N'(s) = -\kappa(s)T(s) \tag{3.4}$$

Ejercicio 3.7 Muestra que la curvatura cuando γ no está parametrizada por longitud de arco, se puede obtener como sigue
$$\kappa(t) = \frac{\det(\gamma'(t), \gamma''(t))}{\|\gamma'(t)\|^3}$$
Hint: Usa la composición $\gamma(t(s))$ y usa regla de la cadena. ■

Vamos a estudiar una interpretación de la curvatura en el plano. Consideremos la **curva paralela**:
$$\rho_d(s) = \gamma(s) + dN(s)$$
al derivar y usar (3.4) obtenemos
$$\rho_d'(s) = (1 - d\kappa(s))T(s)$$
Entonces
$$\rho_d'(s) = 0 \Leftrightarrow d = \frac{1}{\kappa(s)}$$
esto significa que cuando $d = \frac{1}{\kappa(s)}$, la curva paralela contiene todos los puntos singulares pues $\rho_d'(s)$ en relación con $T(s)$ es nula. Ahora lo que vamos a ver es que cada uno de esos puntos singulares son el centro de la mejor circunferenta que aproxima a la curva.

Fijemos $a \in \mathbb{R}^2$ y consideremos la función distancia $d_a : \mathbb{R}^2 \to \mathbb{R}$ como

$$d_a(p) = \|p-a\| = \langle p-a, p-a\rangle$$

tenemos las siguientes observaciones:

1. Un punto p está en la circunferencia $C(a,r)$ centrada en a con radio r si y sólamente si $d_a(p) = r^2$.
2. Si $g(s) = D_a(\gamma(s))$, entonces g mide los puntos de contacto de una circunferencia centrada en a en la curva γ. Es decir un punto $\gamma(s)$ esta en la circunferencia $C(a,r)$ si y sólamente si $g(s) = r^2$.
3. Si $a = \gamma(s) + \lambda N(s)$, toda circunferencia que está en contacto con γ tiene un contacto tangencial en γ, pues comparten en $\gamma(s)$ el mismo vector tangente.

Con lo anterior, vamos a determinar λ, es decir el radio de dicha circunferencia tangencial, de tal manera que tenga la mejor aproximación.

Definición 3.6.2 — Punto de contacto. Dada una función $f : \mathbb{R} \to \mathbb{R}$, un punto s tal que

$$f^{(i)}(s) = 0,\ i = 1, \cdots, k,\ f^{(k+1)}(s) \neq 0$$

se dice que tiene $k+1$ puntos de contacto.

Primero calculemos las derivadas de g en el caso general.

$$\begin{aligned}
g' &= 2(\langle T(s), \gamma(s)-a\rangle)\\
g'' &= 2\left(\kappa(s)\langle N(s), \gamma(s)-a\rangle + 1\right)\\
g''' &= 2(\kappa'(s)\langle N(s), \gamma(s)-a\rangle - \kappa^2\langle T(s), \gamma(s)-a\rangle)\\
g^{(iv)} &= 2(\kappa''-\kappa^3)\langle N, \gamma-a\rangle - 3\kappa\kappa'\langle T, \gamma-a\rangle - \kappa^2
\end{aligned}$$

Para el caso $a = \gamma(s) + \lambda N(s)$, entonces

$$\begin{aligned}
g' &= 0\\
g'' &= 2(-\kappa\lambda + 1)\\
g''' &= 2(-\lambda\kappa')\\
g^{(iv)} &= -2\lambda(\kappa''-\kappa^3) - \kappa^2
\end{aligned}$$

obteniendo

1. Todo punto en la normal tiene al menos 2 puntos de contacto.
2. Cuando $\lambda = \frac{1}{\kappa}$, entonces tiene al menos 3 puntos de contacto.
3. Cuando $\lambda = \frac{1}{\kappa}$, $\kappa'(s) = 0$ y $\kappa''(s) \neq 0$ entonces tiene al menos 4 puntos de contacto. Geométricamente se entiende que este centro es un vértice de la envoluta.

Mostrando que los puntos de la curva $\gamma(s) + \frac{1}{\kappa}N(s)$ contiene todos los centros de las circunferencias que hace contacto con la curva y mejor se aproxima.

■ **Ejemplo 3.10 — Evoluta en coordenadas cartesianas.** Supongamos que la curva tiene coordenadas cartesianas $\gamma(t) = (x(t), y(t))$, entonces

$$T(t) = \left(\frac{x'}{\sqrt{(x')^2+(y')^2}}, \frac{y'}{\sqrt{(x')^2+(y')^2}}\right)$$

obteniendo que $N(t)\left(\frac{y'}{\sqrt{(x')^2+(y')^2}}, \frac{-x'}{\sqrt{(x')^2+(y')^2}}\right)$, entonces la ecuación de la evoluta es

$$e(t) = \left(x - \frac{y'((x')^2+(y')^2)}{x'y''-x''y'}, y + \frac{x'((x')^2+(y')^2)}{x'y''-x''y'}\right)$$

■

■ **Ejemplo 3.11** Sea la curva $y = ax^2$ con $a > 0$, entonces se puede parametrizar como $\gamma(t) = (t, at^2)$, obteniendo

$$\begin{aligned} \gamma' &= (1, 2at) \\ \gamma'' &= (0, 2a) \\ \kappa &= \frac{2a}{(1+4a^2t^2)^{3/2}} \end{aligned}$$

Entonces la envoluta es

$$e(t) = (-4a^2t^3, \frac{1}{2a} + 3at^2))$$

■

Motivando la siguiente definición

Definición 3.6.3 — Radio de curvatura. Definimos el radio de curvatura de γ como

$$\rho(s) = \frac{1}{\kappa(s)}$$

El vector normal y la curvatura se pueden extender a mas dimensiones de manea natural

Proposición 3.6.1 — Primera relación de Frenet. Si $\gamma \colon I \to \mathbb{R}^n$ es parametrizada por longitud de arco y $\gamma'' \neq 0$ entonces

$$T'(s) = \kappa(s)N(s)$$

Caso $\mathbb{R}^3$. Ahora vamos a describir la relación sobre curvas en $\mathbb{R}^3$. Sea $B(s) := T(s) \times N(s)$.

Proposición 3.6.2 Sea $\gamma \colon I \to \mathbb{R}^3$ un curva parametrizada por longitud de arco tal que $B'(s)$ existe, entonces

$$B'(s) = -\tau(s)N(s)$$

Corolario 3.6.3 — Marco de Referencia de Frenet. Bajo las hipótesis anteriores tenemos:

$$\begin{pmatrix} T'(s) \\ N'(s) \\ B'(s) \end{pmatrix} = \begin{pmatrix} 0 & \kappa(s) & 0 \\ -\kappa(s) & 0 & \tau(s) \\ 0 & -\tau(s) & 0 \end{pmatrix} \begin{pmatrix} T(s) \\ N(s) \\ B(s) \end{pmatrix}$$

3.7. Derivadas de orden superior.

Si $f \colon U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ es diferenciable en U, esto nos permite crear una función llamado derivada

$$\begin{aligned} df \colon \quad U &\to L(\mathbb{R}^n, \mathbb{R}^m) \\ p &\to d_p f \end{aligned}$$

Si esta función es diferenciable, decimos que f es **doblemente diferenciable**. Para aclarar esto, notemos que el espacio vectorial $L(\mathbb{R}^n, \mathbb{R}^m)$ tiene una estructura de espacio normado, en donde para cada $T \in L(\mathbb{R}^n, \mathbb{R}^m)$, se le define su norma como

$$\|T\| := \sup_{0 < \|x\| \leq 1} \{\|T(x)\|\}$$

Obteniendo la desigualdad

$$\|T(x)\| \leq \|T\| \, \|x\|, \; \forall x \in \mathbb{R}^n$$

Con ella, si $\{e_i\}_{i=1}^n \subseteq \mathbb{R}^n$ es la base estándar, entonces $\pi_i = e_i^*\colon \mathbb{R}^n \to \mathbb{R}$ es su dual, si usamos las inclusiones canónicas $u_j\colon \mathbb{R} \to \mathbb{R}^m$, tenemos que las funciones $f_{ij} := u_j \circ \pi_i$ forman una base de $L(\mathbb{R}^n, \mathbb{R}^m)$ con $\|f_{ij}\| = 1$ y además, si $\{\hat{e}_s\}_{s=1}^{nm} \subseteq \mathbb{R}^{nm}$ es la base estandar, cualquier biyección $f_{ij} \leftrightarrow \hat{e}_s$ induce un isomorfismo de espacios normados

$$L(\mathbb{R}^n, \mathbb{R}^m) \cong \mathbb{R}^{nm}$$

Esto nos permite que desde la función $U \to \mathbb{R}^{nm}$ que envia $p \to Df(p)$, podamos definir la noción de diferenciabilidad.

Definición 3.7.1 — Segunda diferencial. Una función $f\colon U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ tiene segunda diferencial en $a \in U$ si df es diferenciable en a. Más precisamente, si existe una función lineal $d_a^2 f\colon \mathbb{R}^n \to L(\mathbb{R}^n, \mathbb{R}^m)$ tal que para toda $h \in \mathbb{R}^n$ con $\|h\| < r$ adecuado, se cumple

$$d_{a+h} f = d_a f + d_a^2 f(h) + \|h\|\, m(a,h)$$

donde $\lim_{\|h\| \to 0} m(a,h) = 0$.

Vamos a reinterpretar y describir $d_a^2 f$ en su forma matricial segun (A.4.3).

Caso escalar En esta situación, usando la función $U \to L(\mathbb{R}^n, \mathbb{R}) \cong \mathbb{R}^n$ que envia $p \to \nabla f(p)$, tenemos de la definición (3.7.1) que

$$(\partial_{x_i} f(a+h) - \partial_{x_i} f(a))_{i=1}^n = d_a^2 f(h) + \|h\|\, m(a,h)$$

Repitiendo la prueba de (3.3.1), si consideramos la recta $h = te_j$ con $|t| < r$, entonces

$$\left(\frac{\partial_{x_i} f(a+te_j) - \partial_{x_i} f(a)}{t} \right)_{i=1}^n = d_a^2 f(e_j) + \frac{|t|}{t} m(a, te_j)$$

obteniendo cuando $t \to 0$ que

$$(\partial_{x_j x_i}^2 f(a))_{i=1}^n = d_a^2 f(e_j)$$

Entonces la matriz de representación de $d_a^2 f$ es la matriz $Hf(a) \in \mathbb{M}_{n\times n}(\mathbb{R})$ definido como

$$Hf = D^2 f = \nabla^2 f := \left(\partial_{x_i x_j}^2 f \right)$$

conocido como la Hessiana de f.

Caso de campos vectoriales. De manera similar al caso escalar, si $f\colon \mathbb{R}^n \to \mathbb{R}^m$ es descrito en coordenadas cartesianas como $f = (f_1, \cdots, f_m)$ y es doblemente diferenciable, entonces:

$$Hf(a) = (\nabla^2 f_1(a), \cdots, \nabla^2 f_m(a)) \in \mathbb{M}_{n\times nm}(\mathbb{R})$$

Derivadas de Orden superior El teorema (B.0.1) muestra una forma alternativa de ver a $d_a^2 f$, pues le corresponde una única función bilineal (por abuso de notación) $d_a^2 f(-,-)\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^m$ y la definición (3.7.1) se puede reescribir para toda $(h,k) \in \mathbb{R}^n \times \mathbb{R}^n$ entonces

$$f(a+h+k) = f(a) + d_a f(h) + d_a f(k) + d_a^2 f(h,k) + (\|h\|^2 + \|k\|^2)\mathscr{E}(a,h,k)$$

Donde $\lim_{(h,k)\to 0} \mathscr{E}(a,h,k) = 0$. Entonces la función $f\colon \mathbb{R}^n \to \mathbb{R}^m$ es $s+1$-ésimo diferenciable si el mapeo de la s-derivada

$$f^{(s)}\colon U \to L(\mathbb{R}^n, \cdots, \mathbb{R}^n; \mathbb{R}^n)$$

es diferenciable, y entonces la $s+1$-diferencial, se entiende como la función $s+1$-lineal

$$d_a^{(s+1)}f\colon \mathbb{R}^n\times\cdots\times\mathbb{R}^n\to\mathbb{R}^m$$

Y como el caso bilineal, tiene una representación matricial, en donde sus coeficientes corresponden a derivadas parciales iteradas de la forma

$$\partial_{x_I}^{(s+1)}f,\ I=(i_1,\cdots,i_{s+1})\in\{1,\cdots,s+1\}^{s+1}$$

todas estas funciones en caso de existir son simétricas.

Proposición 3.7.1 Si f es doblemente diferenciable en a y continua, entonces la función bilineal d_a^2f es simétrica.

Demostración. Basta probarlo para el caso de funciones escalares. Sean $i\neq j$ entre $1,\cdots,n$, por simplicidad asumamos que $0\in dom f$ y el caso general se resuelve mediante regla de la cadena a una traslación de la forma $x\to tx$ con $t\in[0,1]$. Definamos $a_0:=\partial_{x_j}\partial_{x_i}f(0)$ y $a_1:=\partial_{x_i}\partial_{x_j}f(0)$, sea $\varepsilon>0$, por definición de continuidad en la segunda derivada, entonces existe $\delta>0$ tal que, cuando $\|x\|\leq 2\delta$ entonces

$$\left|\partial_{x_j}\partial_{x_i}f(x)-a_0\right|\leq\varepsilon,\ \left|\partial_{x_i}\partial_{x_j}f(x)-a_1\right|\leq\varepsilon \tag{3.5}$$

Definamos

$$X(\delta):=f(\delta e_i+\delta e_j)-f(\delta e_i)-f(\delta e_j)-f(0)$$

usando el teorema fundamental del calculo en la variable de i-ésima coordenada, tenemos

$$f(\delta e_i+\delta e_j)-f(\delta e_j)=\int_0^\delta\partial_{x_i}(x_ie_i+\delta e_j)dx_i$$

y

$$f(\delta e_i)-f(0)=\int_0^\delta\partial_{x_i}f(x_ie_i)dx_i$$

sustituyendo en la definición de $X(\delta)$ obtenemos

$$X(\delta)=\int_0^\delta\partial_{x_i}f(x_ie_i+\delta e_j)-\partial_{x_i}f(x_ie_i)dx_i$$

aplicando el teorema de valor medio existe una $0\leq h_j\leq\delta$ tal que

$$\partial_{x_i}f(x_ie_i+\delta e_j)-\partial_{x_i}f(x_ie_i)=\delta\partial_{x_j}\partial_{x_i}f(x_ie_i+h_je_j)$$

como para $x:=x_ie_i+h_je_j$ satisface $\|x\|\leq 2\delta$, usando (3.5) obtenemos

$$\left|\partial_{x_i}f(x_ie_i+\delta e_j)-\partial_{x_i}f(x_ie_i)-\delta a_0\right|\leq\varepsilon\delta$$

Integrando lo anterior desde 0 hasta δ obtenemos

$$\left|X(\delta)-\delta^2a_0\right|\leq\varepsilon\delta^2$$

De manera análoga, tenemos

$$\left|X(\delta)-\delta^2a_1\right|\leq\varepsilon\delta^2$$

Por desigualdad triangular, esto implica

$$\left|\delta^2a_0-\delta^2a_1\right|\leq 2\varepsilon\delta^2\Rightarrow\ |a_0-a_1|\leq 2\varepsilon$$

esta ultima desigualdad, es válida para todo $\varepsilon>0$, entonces $a_0=a_1$. ■

Usando inducción matemática, obtenemos que

Proposición 3.7.2 Si f es s-diferenciable en a y continua, entonces el mapeo s-diferencial d_a^sf es simétrico.

Ejercicio 3.8 Sea $f\colon \mathbb{R}^2 \to \mathbb{R}$ definido como

$$f(x,y) = \begin{cases} \frac{xy^3}{x^2+y^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0) \end{cases}$$

Prueba que f es doblemente diferenciable en $(0,0)$ pero las segundas derivadas cruzadas, no coinciden. ■

3.7.1. Expansiones de Taylor.

Proposición 3.7.3 Sea $f\colon U \subseteq \mathbb{R}^n \to \mathbb{R}$ una función con derivadas parciales continuas de orden a lo más r. Sea $P \in U$ y $H \in \mathbb{R}^n$(Interpretado como un vector en $Tp\mathbb{R}^n$). Definimos $g(t) = f(P+tH)$, entonces

$$g^{(r)}(t) = ((H\cdot\nabla)^r f)(P+tH)$$

para todos los valores en t tal que $P+tH \in U$

Corolario 3.7.4 — Formula de Taylor. Bajo las hipótesis anteriores, existe una $\tau \in (0,1)$ tal que

$$f(P+H) = f(P) + \frac{1}{1!}(H\cdot\nabla)f(P) + \cdots + \frac{1}{(r-1)!}(H\cdot\nabla)^{r-1}f(P) + \frac{1}{r!}(H\cdot\nabla)^r f(P+\tau H)$$

Capítulo 4

Aplicaciones y temas selectos de la Derivada.

4.1. Teorema de la función implícita.

A continuación presentaremos los dos principales teoremas de esta sección.

Teorema 4.1.1 — Teorema de la función implicita. Sea $f\colon A \subseteq \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^n$ una función $f = (f_1, \cdots, f_n)$ continuamente diferenciable, $(a,b) \in \mathbb{R}^m \times \mathbb{R}^n$ tal que $f(a,b) = 0$. Denotemos las variables independientes $(x,y) := (x_1, \cdots, x_m, y_1, \cdots, y_n) \in \mathbb{R}^m \times \mathbb{R}^n$ y $Df(a,b) := [D_x f(a,b), D_y f(a,b)]$. Si $D_y f(a,b)$ es una matriz invertible entonces existen:

1. $V \subseteq \mathbb{R}^{m+n}$ abierto, $(a,b) \in V$.
2. $W \subseteq \mathbb{R}^m$ con $a \in W$ descrito como

$$W = \{x \mid \exists y \in \mathbb{R}^n, f(x,y) = 0,\ (x,y) \in V\}.$$

3. Existe una función continuamente diferenciable $g\colon W \subseteq \mathbb{R}^m \to \mathbb{R}^n$.

Tales que

1. $f(x, g(x)) = 0$ para toda $x \in W$.
2. $Dg(x) = -(D_y(x, g(x))^{-1} D_x f(x, g(x))$ para toda $x \in W$.

El siguiente teorema garantiza la existencia de funciones invertibles y diferenciables, para ello requerimos la siguiente definición.

Definición 4.1.1 — Difeomorfismo. Una función $f\colon U \to V$ diferenciable es un difeomorfismo si existe una función diferenciable $g\colon V \to U$ tales que

$$f \circ g = Id_V,\ g \circ f = Id_U.$$

Notemos que todo difeomorfismo es homeomorfismo y por tanto una función biyectiva. Sin embargo, lo contrario no siempre es verdadero.

■ **Ejemplo 4.1** Considere $f\colon \mathbb{R} \to \mathbb{R}$ definido como $f(x) = x^3$ y $g\colon \mathbb{R} \to \mathbb{R}$ definido como $g(x) = \sqrt[3]{x}$, claramente ambas funciones son continuas y además

$$f \circ g = \mathrm{Id}_{\mathbb{R}}, g \circ f = \mathrm{Id}_{\mathbb{R}}$$

Pero g no es diferenciable en 0, por tanto f no es difeomorfo en cero. ■

Teorema 4.1.2 — Teorema de la función inversa. Sea $f\colon \Omega \subseteq \mathbb{R}^n \to \mathbb{R}^n$ una función continuamente diferenciable en Ω y $a \in Int\Omega$. Si $Df(a)$ es invertible entonces existen abiertos $U, V \subseteq \mathbb{R}^n$ con $a \in U \subseteq \Omega$ tales que:

1. $f|_U : U \to V$ es un difeomorfismo.
2. Sea $g\colon V \to U$ su inversa y $y = f(x)$ entonces

$$Dg(y) = (Df(g(y)))^{-1}$$

Lema 4.1.3 Supongamos que $f\colon U \to V$ es continuamente diferenciable con $Df(u)$ inyectiva para toda $u \in U$, entonces $f(U)$ es abierto.

Primero, mostraremos su equivalencia.

Teorema 4.1.4 — Equivalencia función implícita-función inversa. Son equivalentes:

1. El teorema de la función implícita (4.1.1).
2. El teorema de la función inversa. (4.1.2).

Demostración. Supongamos válido (1), Sea f una función que satisface las hipótesis del teorema (4.1.2), definimos $F : A \subseteq \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ como

$$F(x,y) := y - f(x), \ A := \Omega \times \operatorname{Im} f$$

Tomemos $a \in \Omega$, sea $b := f(a)$, entonces $(a,b) \in F^{-1}(0)$ y además

$$DF(a,b) := [\mathrm{Id}_n \mid -Df(a)]$$

como $Df(a)$ es invertible, por (4.1.1) existe

1. $(a,b) \in V \subseteq \mathbb{R}^n \times \mathbb{R}^n$.
2. $g\colon W \subseteq \mathbb{R}^n \to \mathbb{R}^n$ continuamente diferenciable tal que

$$Dg(b) = [Df(g(b))]^{-1}$$

Por ultimo como Dg es invertible, entonces $U := g(W)$ es abierto y $a \in V$. Es fácil ver que $f\colon U \to W$ es un difeomorfismo, entonces es válido (2). Ahora supongamos válido (2), sea $f\colon \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^m$ tal que satisface las hipótesis de (4.1.1), definimos

$$F(x,y) := (x, f(x,y))\colon \mathbb{R}^{n+m} \to \mathbb{R}^{n+m}$$

entonces, es fácil ver que F satisface las hipótesis de (4.1.2), obteniendo de que existe $G(u,v)$ tal que $F \circ G(u,v) = (u,v)$, digamos $G(u,v) = (G_1(u,v), f(G_1,G_2))$, entonces la función $g(x) := G_2(x,0)$ satisface el teorema de la función implicita. ■

Ahora probaremos el teorema de la función implicita, utilizando un teorema muy importante dentro del analisis y ecuaciones diferenciales:

Teorema 4.1.5 — Teorema del Punto fijo de Banach. Sea (X,d) un espacio métrico completo. Sea $f\colon X \to X$ tal que existe $L \in (0,1)$ tal que

$$\|f(x) - f(y)\| \leq L\|x - y\|$$

Entonces existe un único $\hat{x} \in X$ tal que $f(\hat{x}) = \hat{x}$.

Lema 4.1.6 Sea $B_r(0) \subseteq \mathbb{R}^n$ y $g\colon B_r(0) \to \mathbb{R}^m$ una función tal que existe $c \in (0,1)$ con

$$\|g(y) - g(x)\| \leq c\,\|y - x\|$$

Define $f(x) = x + g(x)$ en $B_r(0)$ entonces

$$(1-c)\,\|x - y\| \leq \|f(x) - f(y)\|$$

En particular f es inyectiva. Además, si $g(0) = 0$ entonces

$$B_{(1-c)r}(0) \subset f(B_r(0)) \subset B_{(1+c)r}(0)$$

Demostración. Las desigualdad queda como ejercicio al lector. Veamos que f es inyectiva, supongamos que $f(x) = f(y)$ entonces $g(x) - g(y) = y - x$, entonces $\|g(x) - g(y)\| \leq c\,\|x - y\|$, es decir $(1-c)\,\|x - y\| \leq 0$, pero $0 < 1 - c$ entonces $\|x - y\| = 0$, es decir $x = y$. Ahora solo resta ver que $B_{(1-c)r}(0) \subset f(B_r(0))$. Sea $y \in B_{(1-c)r}(0)$, define $F(x) = y - g(x)$ como una función $F\colon B_r(0) \to B_r(0)$, notemos que para cada $x \in B_r(0)$, se tiene

$$\|F(x)\| \leq \|y\| + \|g(x)\| \leq (1-c)r + \|g(x) - g(0)\| < (1-c)r + rc = r$$

Entonces F está bien definido, ademas notemos que $\|F(x) - F(x')\| = \|g(x') - g(x)\| \leq c\,\|x' - x\|$, obteniendo una contracción y por el teorema del punto fijo (4.1.5) existe x tal que $x = y - g(x)$, es decir $f(x) = y$, completando la prueba. ∎

4.1.1. Prueba del teorema de la función inversa.

Demostración. Basta probar el caso $a = 0$, $b = f(a) = 0$ y $f'(0) = I$. Definamos $g(x) = f(x) - I(x)$, usando el teorema del valor medio aplicado a la función $\xi(t) = g(x + t(y - x))$ dice que

$$\|g(y) - g(x)\| \leq \|y - x\| \sup_{0<t<1} \left\|\xi'(t)\right\|$$

como $g'(0) = 0$ y g' es continuo, entonces existe una r tal que

$$\|g(y) - g(x)\| \leq 2^{-1}\,\|y - x\|\,,\ \forall x, y \in B_0(r)$$

Por (4.1.6) se tiene que f es inyectiva en $B_0(r)$ y que $B_0(R/2) \subset f(B_0(r))$, entonces, si $U := B_0(r) \cap f^{-1}(B_0(r/2))$ y $v := B_0(r/2)$ entonces la función restringida

$$f\colon U \to V$$

es una biyección y por tanto tiene una inversa $f^{-1}\colon V \to U$. Basta ver que f^{-1} es de clase C^1. Sea $h := f^{-1}$, $A := Df(x)$ la matriz de representación de la derivada. Para $x = g(y)$, podemos escribir

$$h(y + k) = x + s \Leftrightarrow f(x + s) = y + k$$

entonces

$$\|s - k\| = \|f(x + s - f(x) - s)\| \leq \|s\|/2$$

y entonces $\|s\|/2 \leq \|k\|$. Entonces

$$\left\|h(y + k) - h(y) - A^{-1}k\right\| = \left\|h - A^{-1}(f(x + s) - f(x))\right\| \leq \left\|A^{-1}\right\| \|As - f(x + s) + f(x)\|$$

Cuando $k \to 0$, tenemos que $h \to 0$ y $\|h\| / \|k\|$ está acotado. Por lo tanto g es diferenciable en y y derivada continua. ∎

4.1.2. Aplicaciones.

Empecemos con una pequeña aplicación.

■ **Ejemplo 4.2** Sea $F: A \subseteq \mathbb{R}^2 \to \mathbb{R}$ y $s \in ImF$, definimos $L_s \subseteq \mathbb{R}^2$ la curva de nivel en s, $(x_0, y_0) \in A$ tal que $\partial_y F(x_0, y_0) \neq 0$. Nuestro objetivo es mostrar que en efecto una curva diferenciable. Por el teorema de la función implicita (4.1.1) existe:

1. Un abierto $x_0 \in U \subseteq \mathbb{R}$.
2. $f: U \to \mathbb{R}$ continuamente diferenciable tal que
   a) $f(x_0) = y_0$.
   b) $F(x, f(x)) = s$ para toda $x \in U$.
   c) $\frac{df}{dx}(x_0) = -\frac{\partial_x F(x_0,y_0)}{\partial_y F(x_0,y_0)}$

Permitiendo construir $\gamma_{(x_0,y_0)}: U \to L_a$ como sigue $\gamma_{(x_0,y_0)}(t) = (t, f(t))$, que es diferenciable y $\gamma'_{(x_0,y_0)}(t) = (1, f'(t))$.

Por otro lado, sea $H: A \to \mathbb{R} \times \mathbb{R}$ definido como $H(x,y) = (f(x,y), x)$, entonces

$$DH(x_0, y_0) = \begin{pmatrix} \partial_x f(x_0, y_0) & \partial_y f(x_0.y_0) \\ 1 & 0 \end{pmatrix}$$

es una matriz de rango 2, obteniendo que existe una inversa $G: V \times W \subseteq \mathbb{R}^2 \to U' \subseteq A$ dado $G = (G_1, G_2)$ tal que

$$\begin{array}{rcl} G \circ H(x,y) = (x,y) & \Rightarrow & (G_1(F(x,y),y), G_2(F(x,y),y))) = (x,y) \\ H \circ G(a,b) = (a,b) & \Rightarrow & (f(G_1(a,b), G_2(a,b)), G_2(a,b)) = (a,b) \end{array}$$

De allí G_2 es la proyección e la segunda entrada, obteniendo

$$\begin{array}{r} (G_1(F(x,y),y),y) = (x,y) \\ (f(G_1(a,b),b),b) = (a,b) \end{array} \Rightarrow x = G_1(F(x,y),y)$$

obteniendo que

$$\begin{array}{rccc} g: & W & \to & L_s \cap U' \\ & t & \to & G_1(s,t) \end{array}$$

es una función continuamente diferenciable con la propiedad

$$g(y) = x \Leftrightarrow f(x) = y$$

obteniendo un homeomorfismo $W \to L_s \cap U' \subseteq L_s$ ■

Las técnicas discutidas en el ejemplo anterior se pueden formalizar como sigue.

Definición 4.1.2 — Subvariedad k-dimensional. Un conjunto $M \subseteq \mathbb{R}^n$ es llamado subvariedad k-dimensional de $\mathbb{R}^n$ si para todo $x_0 \in M$ existe una vecindad abierta Ω de $x_0 \in \mathbb{R}^n$ y una función continuamente diferenciable $f: \Omega \subseteq \mathbb{R}^n \to \mathbb{R}^{n-k}$ tal que:

$$M \cap \Omega = f^{-1}(\{0\}),\ rankDf(x) = n-k,\ \forall x \in M \cap \Omega$$

Definición 4.1.3 — Gráfica de una función. Sea $f: A \to B$ una función, denotamos $G_f := \{(x, f(x)) \in A \times B |\ x \in A\}$ a dicho conjunto lo denotamos como la gráfica de la función f.

El siguiente resultado muestra que las subvariedades k-dimensionales son localmente gráficas de funciones locales.

Proposición 4.1.7 Son equivalentes:

1. M es una subvariedad k-dimensional de $\mathbb{R}^n$.

2. Para todo $x_0 \in M$, escribimos $x_0 = (y_0, z_0)$ con $y_0 \in \mathbb{R}^k$ y $z_0 \in \mathbb{R}^{n-k}$ entonces existen:
   a) Abiertos $y_0 \in U \subseteq \mathbb{R}^k$. $z_0 \in V \subseteq \mathbb{R}^{n-k}$
   b) Una función continuamente diferenciable $g\colon U \to V$.

   Tales que $g(y_0) = z_0$ y $M \cap (U \times V) = G_g$.

Ejercicio 4.1 Sea $M_t := \{(a,b,c) | a^4 + b^4 + c^4 = t\}$, muestra que para toda $t > 0$ los conjuntos M_t son subvariedades 2-dimensionales, para $t = 0$ es una subvariedad 0-dimensional y $t < 0$ son conjuntos vacios. ■

Este resultado es intrínseco, es decir, depende del espacio ambiente $\mathbb{R}^n$. Existe una manera para definir variedades diferenciables, pero la definición abstracta queda fuera de estas notas. Sin embargo presentaremos una idea de como construir dichas variedades con los siguientes ejemplos.

■ **Ejemplo 4.3** Sea V un espacio vectorial real de dimensión n. Fija una base β, este induce un isomorfismo $\mathbb{R}$-lineal $T\colon V \to \mathbb{R}^n$ que depende de la base. De esta manera podemos definir una topología en V como sigue:

- $U \subseteq V$ es abierto si y sólamente si $T(U) \subseteq \mathbb{R}^n$ es un abierto con la topología usual.

De esta manera decimos que V tiene una estructura de variedad diferenciable, pues para cada punto $v \in V$ y cada abierto $v \in U$ la función $T|_U\colon U \to T(U)$ es un homeomorfismo tal que si $v \in U'$ es otro abierto, entonces la función:

$$T|_U(U \cap U') \xrightarrow{(T|_U)^{-1}} U \cap U' \xrightarrow{T|_{U'}} T_{U'}(U \cap U')$$

es la restricción de la función identidad y por tanto **diferenciable**. Estableciendo una estructura a V como una variedad diferenciable de dimensión n. ■

■ **Ejemplo 4.4** Siguiendo el ejemplo anterior, tenemos que $V = \mathbb{M}_{n\times n}(\mathbb{R})$ es una variedad diferenciable de dimensión n^2. En algebra lineal se demuestra que existe un invariante importante denominada invariante, donde asigna a cada matriz A un número real $\det A \in \mathbb{R}$, entonces se puede ver como una función

$$\det\colon V \to \mathbb{R}$$

Por simplicidad en las cuentas, se escoge β la base estándar ordenada de las matrices, construyendo el isomorfismo $\mathbb{R}$-lineal $T\colon V \to \mathbb{R}^{n^2}$, de esta manera, por la fórmula de Leibniz para el determinante, se tiene que la función $f = \det \circ T^{-1}\colon \mathbb{R}^{n^2} \to \mathbb{R}$ es una función polinomial irreducible, por lo tanto una función continuamente diferenciable. Entonces:

1. $M = f^{-1}(\mathbb{R} - \{0\})$ es un abierto disconexo en $\mathbb{R}^n$ y por (4.1.7) M es una subvariedad de dimensión n^2, induciendo a $GL(n) \cong M$ una estructura de variedad diferenciable de dimensión n^2.
2. $M = f^{-1}(\{\pm 1\})$ es una subvariedad cerrada diferenciable de dimensión $n(n-1)/2$, por tanto $SO(n)$ es una variedad diferenciable de dimensión $n(n-1)/2$.

■

4.2. Multiplicadores de Lagrange.

En esta sección estudiaremos problemas de optimizacion para funciones de la forma $f\colon D \subseteq \mathbb{R}^n \to \mathbb{R}$, por (3.7.4) obtenemos que existe una $\tau \in (0,1)$ tal que para toda $H \in \mathbb{R}^n$ tal que $P + \tau H \in D$ entonces

$$f(P+H) = f(P) + \nabla f(P) \cdot H + \frac{1}{2} H^T H f(P) H + (H \cdot \nabla)^3 f(P + \tau H) \tag{4.1}$$

Donde ∇f es la gradiente y Hf la hessiana de f. Deseamos estudiar la optimización utilizando el espacio tangente $T_{f(P)}S$.

Proposición 4.2.1 Sea $f\colon \mathbb{R}^n \to \mathbb{R}^m$ continuamente diferenciable en a y $Df(x) = 0$ en una vecindad abierta U de a entonces $f|_U(x)$ es constante.

Demostración. Por el teorema de valor medio para varias variables (3.5.3 tenemos

$$f(a+h) - f(a) = \left(\int_0^1 Df(x+th)dt \right) \cdot h$$

entonces en la vecindad abierta U de a tenemos $f(a+h) = f(a)$. ■

Definición 4.2.1 — Punto Extremo. Sea $f\colon D \subseteq \mathbb{R}^n \to \mathbb{R}$ diferenciable, decimos que a es extremo si $Df(a) = 0$.

Una función puete tener un conjunto infinito de puntos extremos.

■ **Ejemplo 4.5** Sea $f(x,y) = \left(\sqrt{x^2+y^2} - a\right)^2$ con $a > 0$, entonces f es diferenciable en $\mathbb{R}^2 - \{0\}$, pues en esa situación

$$\nabla f(x,y) = 2\left(1 - \frac{a}{\sqrt{x^2+y^2}}\right)(x,y)$$

entonces el conjunto de ceros es $x^2 + y^2 = a^2$ en $\mathbb{R}^2$. ■

Como el ejemplo anterior, el conjunto de ceros puede ser grande que inclusive no se requieran todos para resolver un cierto problema. Muchos problemas de optimización de funciones diferenciables tienen además una restricción.

Caso cuadrático. Sea $X = (x_1, \cdots, x_n)$, definimos

$$q(x) = \sum_{1 \le i,j \le n} a_{i,j} x_i x_j$$

Definimos $B = (\frac{1}{2}(a_{ij} + a_{ji}))$, entonces se tiene

1. $\nabla q(x) = \left(2a_{ii}x_i + \sum_{j=1}^n (a_{ji} + a_{ij})x_j\right)$, entonces el único punto extremo es $x = 0$.
2. $Hq(x) := B$, como B es simétrico, entonces existe $D = diag(\lambda_1, \cdots, \lambda_n)$ tal que $B = PDP^T$ con P ortogonal. Definamos $y = P^T x$, obteniendo $x^T Bx = x^T (P^T)^T DP^T x = y^T Dy$.

Definición 4.2.2 Dado $f\colon \mathbb{R}^n \to \mathbb{R}$ y $x \in Dom(f)$, decimos que x es un punto:
1. máximo local si $f(x) \ge f(y)$ para toda y en alguna vecindad abierta de x.
2. mínimo local si $f(x) \le f(y)$ para toda y en alguna vecidnad abierta de x.
3. silla si existen dos direcciones $u, v \in T_pG_f$ tal que $f(p+tu)$ tiene un máximo local y $f(p+tv)$ tiene un mínimo local.

Proposición 4.2.2 Sea q una forma cuadrática con matriz de definición B, entonces el punto $x = 0$ es:

1. Un mínimo si B es definida positiva y $\lambda_{ii} > 0$ para toda i.
2. Un máximo si B es definida positiva y $\lambda_{jj} < 0$ para alguna j.
3. Un silla si B es definida negativa.

Demostración. Ejercicio. ■

En el caso de $n = 2$, esta caracterización se puede simplificar de las siguientes observaciones. Dado $H = \begin{pmatrix} a & c \\ c & b \end{pmatrix}$ entonces su polinomio característico es:

$$p(t) = t^2 - (a+b)t + ab - c^2$$

Sea Δ la discriminante del polinómio, entonces $\Delta = (a-b)^2 + c^2 \geq 0$, lo que significa que siempre tiene solución, es decir valores propios, denotemos

$$\lambda_1 = \frac{1}{2}(a+b+\sqrt{(a-b)^2+c^2}), \lambda_2 = \frac{1}{2}(a+b-\sqrt{(a-b)^2+c^2})$$

Proposición 4.2.3 Sea $q(x,y) = ax^2 + 2cxy + by^2$, entonces el punto $(x,y) = (0,0)$ es:

1. mínimo si $ab - c^2 > 0$ y $a > 0$.
2. máximo si $ab - c^2 > 0$ y $a < 0$.
3. silla si $ab - c^2 \leq 0$.

Demostración. Vamos a probar el punto 1, el resto se sigue de manera similar. Supongamos que $(x,y) = (0,0)$ es mínimo, entonces H es definida positiva y $\lambda_i > 0$, en particular implica que

$$a+b \geq \sqrt{(a-b)^2+c^2} \geq 0 \Rightarrow \; a+b \geq 0$$

por otro lado, de las hipótesis de H implica que

$$\det H = \lambda_1 \lambda_2 > 0 \Rightarrow ab - c^2 > 0$$

obteniendo que $ab > c^2 \geq 0$, entonces

$$\begin{cases} ab \geq 0 \\ a+b \geq 0 \end{cases}$$

Supongamos que $a \leq 0$ entonces de $b \geq -a$ al multiplicar por a obtenemos $ab \leq -a^2 \leq 0$ pero esto es una contradicción, por lo tanto $a > 0$. ∎

4.2.1. Criterios de máximos y mínimos para varias variables.

Sea f de clase C^2, entonces obtenemos para cada a en el dominio y una $x \in B_\varepsilon(a)$ que:

$$f(a+x) = f(a) + d_a f(x) + \frac{1}{2} d_a^2 f(x,x) + 2\|x\| \mathscr{E}(a,x,x)$$

Teorema 4.2.4 Sea $f\colon \mathbb{R}^2 \to \mathbb{R}$ de clase C^2 y a un punto crítico, entonces el punto a es:

1. máximo si $\frac{\partial^2 f}{\partial x \partial x}(a) \frac{\partial^2 f}{\partial y \partial y}(a) - \left(\frac{\partial^2 f}{\partial x \partial y}(a)\right)^2 > 0$ y $\frac{\partial^2 f}{\partial x \partial x}(a) < 0$.
2. mínimo si $\frac{\partial^2 f}{\partial x \partial x}(a) \frac{\partial^2 f}{\partial y \partial y}(a) - \left(\frac{\partial^2 f}{\partial x \partial y}(a)\right)^2 > 0$ y $\frac{\partial^2 f}{\partial x \partial x}(a) > 0$.
3. silla si $\frac{\partial^2 f}{\partial x \partial x}(a) \frac{\partial^2 f}{\partial y \partial y}(a) - \left(\frac{\partial^2 f}{\partial x \partial y}(a)\right)^2 < 0$.

Demostración. Si a es un punto crítico, entonces $d_a f = 0$, obteniendo

$$f(a+x) = f(a) + \frac{1}{2} d_a^2 f(x,x) + 2\|x\| \mathscr{E}(a,x,x)$$

luego, sea $B = (\frac{\partial^2 f}{\partial x_i \partial x_j}(a))$ la Hessiana en a, denominamos $q(x-a) := d_a^2 f(x,x)$ la forma cuadrática asociada. Notemos que a es un máximo para f si y sólo si a es un máximo para q, de manera similar para el mínimo, por lo que usando (4.2.3) obtenemos el resultado. ∎

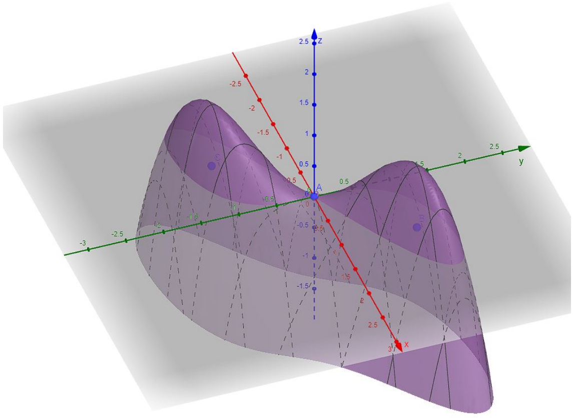

Figura 4.1: Gráfica de $f(x,y) = 4xy - 2x^2 - y^4$

■ **Ejemplo 4.6** Consideremos $f(x,y) = 4xy - 2x^2 - y^4$. Entonces su gradiente es

$$\nabla f(x,y) = (4y - 4x, 4x - 4y^3)$$

Encontrando los puntos críticos:

$$\nabla f(x,y) = (0,0) \Leftrightarrow (x,y) = (0,0), (1,1), (-1,-1)$$

Por otro lado tenemos que su Hessiana es

$$Hf(x,y) = \begin{pmatrix} -4 & 4 \\ 4 & -12y^2 \end{pmatrix}$$

notemos que

$$\det Hf(x,y) = 4^2(3y^2 - 1)$$

Con esto, aplicando (4.2.4), obtenemos que:

1. Los máximos son los puntos $(1,1)$ y $(-1,-1)$.
2. El punto $(0,0)$ es un punto silla.

■

4.2.2. Criterio de los multiplicadores de Lagrange

Empezamos con un lema importante acerca de los conjuntos de nivel

Proposición 4.2.5 Sea $f\colon \mathbb{R}^n \to \mathbb{R}$ de clase C^1 y L_a un conjunto de nivel a, entonces si $\nabla f \neq 0$, este es ortogonal al conjunto L_a.

Demostración. Sea $x_0 \in L_a$ tal que $\nabla f(x_0) \neq 0$, entonces, este es localmente una grafica (4.1.1), obteniendo una parametrización $\phi\colon U \subseteq \mathbb{R}^{n-1} \to L_a$ con $b \in U$ que satisface $\phi(b) = x$. Por regla de la cadena

$$\langle \nabla f(x), D\phi(b) \rangle = 0$$

obteniendo lo deseado. ■

El problema donde los multiplicadores de Lagrange toman lugar es el siguiente

R Estudiar los puntos críticos de $f(x)$ sobre el conjunto determinado por los conjuntos de nivel (denominados **restricciones**) $\{g_k(x)\}_{k=1}^{s}$.

4.3. Optimización por mínimos cuadrados.

Consideremos el siguiente problema

R Dado una colección finita $\beta := \{p_i\}_{i=1}^m \subseteq \mathbb{R}^{n+1}$ y $\{f_i(x)\}_{i=1}^s \subseteq Fun(\mathbb{R}^n, \mathbb{R})$ linealmente independientes, denotemos $V := \langle \{f_i(x)\}_{i=1}^s \rangle$ al espacio vectorial generado por dichas funciones. **Encontrar** $f(x) \in V$ tal que la curva $y = f(x)$ aproxime lo mejor posible a los datos β

Este problema es una generalización a encontrar una recta $y = a + bx$ tal que aproximen mejor los datos, aquí $V = \langle \{1, x\} \rangle$, y es conocido como **interpolación lineal**. Existen diversos métodos de aproximación, los cuales se pueden clasificar según el tipo de error de aproximación que generan. Para este problema usaremos el concepto de **error cuadrático**.

Error cuadrático. Dado $f \in V$, para cada $i = 1, \cdots, m$ denotemos $p_i := (x_i, y_i) \in \mathbb{R}^n \times \mathbb{R}$, entonces tenemos

$$y_i = f(x_i) + \varepsilon_i, \ \forall i$$

en notación vectorial tenemos

$$\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} f(x_1) \\ \vdots \\ f(x_m) \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

Definición 4.3.1 — Error cuadrático. Usando la notación anterior, definimos el error cuadrático de f para los datos β como

$$\mathscr{E} := d\left(\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}, \begin{pmatrix} f(x_1) \\ \vdots \\ f(x_m) \end{pmatrix} \right)^2$$

donde d es la distancia euclidiana de $\mathbb{R}^m$.

En esta situación el error cuadrático depende de $f \in V$, pero $f(x) = \sum_{i=1}^s c_i f_i(x)$ es una combinación linealmente independiente (base para V), entonces el error se puede pensar como una función $\delta \colon \mathbb{R}^s \to \mathbb{R}$ definido como

$$\delta(c_1, \cdots, c_s) := \sum_{i=1}^m \left(y_i - \sum_{j=1}^s c_j f_j(x_i) \right)^2$$

de esta manera la interpretación de aproximación equivale a minimizar δ. Así que el primer paso es calcular $\nabla\delta$

$$\frac{\partial \delta}{\partial c_k} = 2 \sum_{i=1}^m -f_k(x_i) \left(y_i - \sum_{j=1}^s c_j f_j(x_i) \right)$$

de esta manera, el sistema de ecuaciones $\frac{\partial \delta}{\partial c_k}$, $\forall k = 1, \cdots s$ se transforma en el sistema

$$\sum_{i=1}^m y_i f_k(x_i) = \sum_{i=1}^m \sum_{j=1}^s c_j f_j(x_i) f_k(x_i), \ \forall k = 1, \cdots s$$

Podemos reformular el sistema con la siguiente notación, dados $f, g \colon \mathbb{R}^m \to \mathbb{R}$ definimos

$$\langle f, g \rangle_\beta := \sum_{i=1}^m f(x_i) g(x_i)$$

de esta manera el problema anterior, tiene la siguiente forma matricial

$$\nabla\delta(c_1,\cdots,c_s)=0 \Leftrightarrow \begin{pmatrix} \langle f_1,f_1\rangle_\beta & \cdots & \langle f_1,f_s\rangle_\beta \\ \vdots & & \vdots \\ \langle f_s,f_1\rangle_\beta & \cdots & \langle f_s,f_s\rangle_\beta \end{pmatrix}\begin{pmatrix} c_1 \\ \vdots \\ c_s \end{pmatrix} = \begin{pmatrix} \langle f_1,y\rangle_\beta \\ \vdots \\ \langle f_s,y\rangle_\beta \end{pmatrix}$$

donde $y(x_i)=y_i$. Una ves determinado los puntos extremos, debemos usar (4.2.4) para el caso general. Primero, notemos

$$\frac{\partial^2\delta}{\partial c_k\partial c_t}=2\sum_{i=1}^{m} f_k(x_i)f_t(x_i)=2\langle f_k,f_t\rangle_\beta,\ \forall k,t=1,\cdots,s$$

Cuando $s=2$ tenemos la hessiana en cualquier (c_1,c_2) es

$$H\delta:=\begin{pmatrix} \langle f_1,f_1\rangle_\beta & \langle f_1,f_2\rangle_\beta \\ \langle f_1,f_2\rangle_\beta & \langle f_2,f_2\rangle_\beta \end{pmatrix}$$

cuyo determinante es

$$\det H\delta=\langle f_1,f_1\rangle_\beta\langle f_2,f_2\rangle_\beta-\langle f_1,f_2\rangle_\beta^2$$

Luego, uno puede usar el argumento de Cauchy-Swartz (1.2.1), ya que $\langle f,g\rangle_\beta$ es un producto interior en $Fun(\mathbb{R}^n,\mathbb{R})$, obteniendo

$$\det H\delta\geq 0$$

donde la igualdad ocurre si $f_1=f_2$. Por otro lado $\langle f_1,f_1\rangle_\beta\geq 0$, entonces por (4.2.4) tenemos

Proposición 4.3.1 Dado una colección finita $\beta:=\{p_i\}_{i=1}^m\subseteq\mathbb{R}^{n+1}$ y $\{f_i(x)\}_{i=1}^s\subseteq Fun(\mathbb{R}^n,\mathbb{R})$, entonces existen $c_1,\cdots,c_s\in\mathbb{R}$ tal que la función

$$f(x):=\sum_{i=1}^{s}c_if_i(x)$$

se aproxima mejor, por error cuadrático. Más aún,las c_i satisfacen la ecuación lineal

$$\begin{pmatrix} \langle f_1,f_1\rangle_\beta & \cdots & \langle f_1,f_s\rangle_\beta \\ \vdots & & \vdots \\ \langle f_s,f_1\rangle_\beta & \cdots & \langle f_s,f_s\rangle_\beta \end{pmatrix}\begin{pmatrix} c_1 \\ \vdots \\ c_s \end{pmatrix} = \begin{pmatrix} \langle f_1,y\rangle_\beta \\ \vdots \\ \langle f_s,y\rangle_\beta \end{pmatrix}$$

4.4. Ecuaciones diferenciales ordinarias exactas.

En esta sección trataremos un problema muy particular pero muy útil en la solución de ecuaciones diferenciales ordinarias. Para motivar el concepto de exactitud, Sea $u\colon D\subseteq\mathbb{R}^2\to\mathbb{R}$ diferenciable, tal que en el conjunto de nivel

$$u(x,y)=0$$

Entonces por el Teo. de la Función Implicita (4.1.1), u es localmente la gráfica de una función. Esto motiva a pensar si existe una parametrización a dicha curva de nivel $\gamma(t)=(x(t),y(t))$. En dicha situación

$$u(x(t),y(t))=0$$

Derivando

$$\partial_x u\cdot x'(t)+\partial_y u\cdot y'(t)=0$$

A nivel de 1-formas diferenciales

$$\partial_x u dx+\partial_y v dy=0$$

Si además, u es C^2, tenemos

$$\partial_y(\partial_x u)=\partial^2_{yx}u=\partial^2_{xy}u=\partial_x(\partial_y u)$$

Definición 4.4.1 — Ecuaciones Diferenciales Exactas. Una ecuación diferencial exacta es una ecuación de la forma

$$M(x,y)dx + N(x,y)dy = 0$$

tal que

$$\partial_y M(x,y) = \partial_x N(x,y)$$

Resolviendo ecuaciones exactas: Dado una ecuación exacta

$$M(x,y)dx + N(x,y)dy = 0$$

Los pasos a resolver son:

1. Supones una $u\colon D \subseteq \mathbb{R}^2 \to \mathbb{R}$ tal que

$$\begin{cases} \partial_x u = M \\ \partial_y u = N \end{cases}$$

2. Integras la primera ecuación con respecto a x, obteniendo

$$u = \int M dx + f(y)$$

3. Derivas parcialmente con respecto a y y lo comparas con la segunda ecuación

$$N = \partial_y u = \partial_y \int M dx + f'(y)$$

   Resuelve para encontrar $f(y)$.
4. La solución (implícita) está en la familia de curvas de nivel

$$u(x,y) = C,\ C \in \mathbb{R}$$

Ejercicio 4.2 Resolver

1. $(y + 2xe^y)dx + x(1 + xe^y)dy = 0$.
2. $(y - y^2)dx + xdy = 0$.
3. $(x^2y + y + 1)dx + x(1 + x^2)dy = 0$.
4. $(3x^2y + 8xy^2)dx + (x^3 + 8x^2y + 12y^2)dy = 0$.

Vamos a aplicar esta técnica para resolver un tipo de ecuaciónes diferenciales denominadas homogéneas.

Definición 4.4.2 — Funciones Homogéneas. Una función $f(x,y)$ se llama homogénea de grado d si

$$f(tx,ty) = t^d f(x,y)$$

Considera la ecuación (no necesariamente exacta), con M,N homogénea del mismo grado n,

$$M(x,y)dx + N(x,y)dy = 0$$

Entonces

$$x^n M(1,y/x)dx + x^n N(1,y/x)dy = 0$$

Sea $y = xv$, aplicando cambio de variable

$$x^n(M(1,v) + vN(1,v))dx + x^{n+1}N(1,v)dv = 0$$

Reescribiendolo como

$$\frac{1}{x}dx + \frac{N(1,v)}{M(1,v)+vN(1,v)}dv = 0$$

y esta ecuación es exacta, por lo cual podemos resolver la ecuación y obtener la solución implicita.

■ **Ejemplo 4.7** Al resolver

$$y' = \frac{x^2+xy+y^2}{x^2},\ x \neq 0$$

La transformación $y = xv$ simplifica en

$$xv' = 1 + v^2$$

Obteniendo la solución implicita.

$$\tan^{-1}(y/x) = \ln(cx),\ \forall c \in \mathbb{R}$$

■

Parte II

Cálculo Integral en varias variables

Capítulo 5

Integración en varias variables

5.1. Integral de Riemann en Rectángulos.

Vamos a empezar con notación importante:

1. Un rectángulo en $\mathbb{R}^n$ es un conjunto de la forma $R := [a_1, b_1] \times [a_2, b_2] \times \cdots [a_n, b_n]$.
2. Dado R un rectángulo en $\mathbb{R}^n$ definimos su volumen como

$$v(R) := \prod_{i=1}^{n} (b_i - a_i)$$

3. Una partición de R consiste en una familia de subconjuntos de R construidos de la siguiente forma:
   a) Para cada $i = 1, \cdots, n$ tomemos una partición del intervalo $x_0^{(i)} := a_i < x_1^{(i)} < \cdots < x_{s_i}^{(i)} = b_i$. Al conjunto lo denotamos $\mathscr{P}_i := \mathscr{P}([a_i, b_i])$.
   b) Construimos $\mathscr{P} := \mathscr{P}_1 \times \mathscr{P}_2 \times \cdots \times \mathscr{P}_n$ denominado **la partición** de R.
   c) Definimos un rectángulo R_s tal que sus vértices son elementos de $\mathscr{P}$ y $|R_s \cap \mathscr{P}| = 4$. A la colección de dichos rectángulos asociado a la partición de R lo denotaremos como $\mathscr{F}_{\mathscr{P}}$.

Notemos que R es un punto si y sólo si $v(R) = 0$.

Definición 5.1.1 — Refinamiento de particiones. Sean $\mathscr{P}$ y $\mathscr{Q}$ dos particiones de R, decimos que $\mathscr{Q}$ refina a $\mathscr{P}$ si para cada $i = 1, \cdots, n$ se tiene $\mathscr{P}_i \subseteq \mathscr{Q}_i$. En esta situación lo denotamos como $\mathscr{P} \leq \mathscr{Q}$.

Tenemos el siguiente resultado fundamental que describe el comportamiento de las particiones y su refinamiento.

Teorema 5.1.1 — . Fijemos $R \subseteq \mathbb{R}^n$ un rectángulo de volumen no cero y sea $\mathbf{Part}(R)$ el conjunto de todas las particiones de R, entonces:

1. $\mathbf{Part}(R)$ junto al órden definido en (5.1.1) es un conjunto parcialmente ordenado con elemento mínimo $\{a_1, b_1\} \times \cdots \times \{a_n, b_n\}$.
2. $\mathbf{Part}(R)$ es una retícula, donde para cada $\mathscr{P}, \mathscr{Q} \in \mathbf{Part}(R)$
   a) $\sup(\mathscr{P}, \mathscr{Q})$ es la partición $\mathscr{S} = \mathscr{S}_1 \times \cdots \times \mathscr{S}_n$, tal que $\mathscr{S}_i = \mathscr{P}_i \cup \mathscr{Q}_i$.
   b) $\inf(\mathscr{P}, \mathscr{Q})$ es la partición $\mathscr{S} = \mathscr{S}_1 \times \cdots \times \mathscr{S}_n$, tal que $\mathscr{S}_i = \mathscr{P}_i \cap \mathscr{Q}_i$

Nuestro objetivo es generalizar la construcción de Integral de Riemann en varias variables empezando por aproximar la integral de funciones $f\colon R \subseteq \mathbb{R}^n \to \mathbb{R}$.

R Sea R un rectángulo y $\mathscr{P} \in \mathbf{Part}(R)$, entonces

$$v(R) = \sum_{S \in \mathscr{F}_{\mathscr{P}}} v(S)$$

Definición 5.1.2 — Suma de Riemann. Sea $f\colon R \subseteq \mathbb{R}^n \to \mathbb{R}$ una función y $\mathscr{P} \in \mathbf{Part}(R)$, definimos

1. La suma de Riemann superior como

$$\overline{S}(f, \mathscr{P}) := \sum_{S \in \mathscr{F}_{\mathscr{P}}} M_S v(S)$$

donde $M_S := \sup\{f(x) | \, x \in S\}$.

2. La suma de Riemann inferior como

$$\underline{S}(f, \mathscr{P}) := \sum_{S \in \mathscr{F}_{\mathscr{P}}} m_S v(S)$$

donde $m_S := \text{ínf}\{f(x) | \, x \in S\}$.

Proposición 5.1.2 — Propiedades de las sumas de Riemann. Sea $f\colon R \subseteq \mathbb{R}^n \to \mathbb{R}$ una función, las siguientes afirmaciones son válidas:

1. Para cada partición $\mathscr{P}$, se tiene $\underline{S}(f, \mathscr{P}) \leq \overline{S}(f, \mathscr{P})$.
2. Si $\mathscr{P} \leq \mathscr{Q}$ entonces:
   a) $\underline{S}(f, \mathscr{P}) \leq \underline{S}(f, \mathscr{Q})$.
   b) $\overline{S}(f, \mathscr{Q}) \leq \overline{S}(f, \mathscr{P})$.

Definición 5.1.3 — Integral de Riemann superior e inferior. Sea $f\colon R \subseteq \mathbb{R}^n \to \mathbb{R}$ una función, definimos

1. La integral de Riemann inferior como

$$\text{ínf} \int_R f := \sup\{\underline{S}(f, \mathscr{P}) | \, \mathscr{P} \in \mathbf{Part}(R)\}.$$

2. La integral de Riemann superior como

$$\sup \int_R f := \text{ínf}\{\overline{S}(f, \mathscr{P}) | \, \mathscr{P} \in \mathbf{Part}(R)\}.$$

Notemos que gracias a (5.1.2) se tiene

$$\text{ínf} \int_R f \leq \sup \text{ínf}_R f$$

■ **Ejemplo 5.1** Sea $A = [0,1] \times [0,1] \cap \mathbb{Q}$ y definimos $f\colon [0,1] \times [0,1] \to \mathbb{R}$ como $f(x) = \chi_A(x)$, entonces $\text{ínf} \int_R f = 0 < \sup \text{ínf}_R f = 1$. ■

Notemos que los valores supremos e infimos existen cuando f es acotada.

Definición 5.1.4 — Funciones Riemann integrables. Sea $f\colon R \to \mathbb{R}$ una función acotada, decimos que f es Riemann integrable sobre R si

$$\text{ínf} \int_R f = \sup \int_R f$$

a dicho valor se le denotará como $\int_R f$.

Teorema 5.1.3 — Primera caracterización de funciones Riemann integrables. Sea $f\colon R \subseteq \mathbb{R}^n \to \mathbb{R}$ acotada. Entonces, f es Riemann integrable sobre R si y sólo si para toda $\varepsilon > 0$ existe una partición $\mathscr{P}$ de R tal que

$$\overline{S}(f,\mathscr{P}) - \underline{S}(f,\mathscr{P}) < \varepsilon.$$

Proposición 5.1.4 Si $f\colon R \to \mathbb{R}$ es continua entonces f es Riemann integrable.

Teorema 5.1.5 — Propiedad topológica de las funciones integrables. Sea $f\colon R \to \mathbb{R}$ Riemann integrable, entonces existe $x_0 \in int(R)$ tal que f es continua en x_0. Más aún e conjunto de discontinuidades tiene interior vació.

Proposición 5.1.6 — Propiedades de funciones integrables. Sean $f,g\colon R \to \mathbb{R}$ funciones integrables, son válidas:

1. $f+g$ es integrable y $\int_R (f+g) = \int_R f + \int_R g$.
2. Para cada $c \in \mathbb{R}$, cf es integrable y $\int_R (cf) = c\int_R f$.
3. fg es integrable sobre R.
4. Si $f \geq 0$ sobre R, entonces $\int_R f \geq 0$.
5. Si $f \leq g$ sobre R, entonces $\int_R f \leq \int_R g$.
6. $\|f\|$ es integrable sobre R y $\|\int_R f\| \leq \int_R \|f\|$.

Proposición 5.1.7 Sea $f\colon R \to \mathbb{R}$ integrable sobre R, las siguientes afirmaciones son válidas:

1. Si para toda $x \in R$, $f(x) \geq 0$, existe $x_0 \in R$ tal que f es continua con $f(x_0) > 0$ entonces $\int_R f > 0$.
2. Si para toda $x \in R$, $f(x) > 0$, entonces $\int_R f > 0$.

Proposición 5.1.8 Sea $f\colon R \subseteq \mathbb{R}^n \to [a,b] \subseteq \mathbb{R}$ integrable y $h\colon [a,b] \to \mathbb{R}$ continua, entonces $h \circ f\colon R \to \mathbb{R}$ es integrable.

Teorema 5.1.9 — Teorema del valor medio para integrales sobre rectangulos. Sean $f,g\colon R \to \mathbb{R}$ funciones con f continua y g integrable sobre R. Las siguientes afirmaciones son válidas:

1. Existe $h \in R$ tal que $\int_R f = f(h)v(R)$.
2. Si $g \geq 0$ en R entonces existe $h \in R$ tal que $\int_R fg = f(h)\int_R g$.

5.2. Conjuntos Jordan medibles.

Sea X un conjunto, tenemos una biyección

$$\begin{array}{ccc} \mathscr{P}(X) & \to & \mathbf{Fun}(X,\{0,1\}) \\ A & \to & \chi_A \end{array}$$

donde $\chi_A(x) = 1$ si $x \in A$ y $\chi_A(x) = 0$ si $x \notin A$, a esta función se le denomina indicador del subconjunto A.

Proposición 5.2.1 — Propiedades de la función indicador. Para cada conjunto X, pensando $\{0,1\} \subseteq \mathbb{R}$, tenemos las siguientes afirmaciones:

1. $\chi_A = 0$ si y sólo si $A = \emptyset$.
2. Si $A,B \subseteq X$ entonces son válidos:
   a) Si $A \subseteq B$ entonces $\chi_A \leq \chi_B$ como funciones.
   b) Si $A \cap B =$ entonces $\chi_{A\cup B} = \chi_A + \chi_B$.
   c) $\chi_{A\cup B} = \chi_A + \chi_B - \chi_{A\cap B}$.
   d) $\chi_{A\cap B} = \chi_A \chi_B$.
   e) $\chi_{X-A} = 1 - \chi_A$

3. Si $f\colon X\to Y$ es una función, entonces el siguiente diagrama de funciones es comutativo

$$\begin{array}{ccc} \mathscr{P}(Y) & \xrightarrow{\cong} & \mathbf{Fun}(Y,\{0,1\}) \\ {\scriptstyle f^*}\downarrow & & \downarrow{\scriptstyle -\circ f} \\ \mathscr{P}(X) & \xrightarrow[\cong]{} & \mathbf{Fun}(X,\{0,1\}) \end{array}$$

es decir, para cada subconjunto $S\subseteq Y$ entonces

$$\chi_S\circ f=\chi_{f^{-1}(S)}.$$

Definición 5.2.1 — Conjunto Jordan medible. Sea $A\subseteq\mathbb{R}^n$ decimos que A es Jordan medible si existe un rectángulo $R\subseteq\mathbb{R}^n$ tal que

1. $A\subseteq R$.
2. $\chi_A\colon R\to\mathbb{R}$ es Riemann integrable

Si A es Jordan medible, definimos su medida como

$$v(A):=\int_R\chi_A.$$

Proposición 5.2.2 Si $A\subseteq\mathbb{R}^n$ es Jordan medible, entonces la medida $v(A)$ no depende del rectángulo escogido. Es decir, si R,R' son rectángulos tales que $A\subseteq R$ y $A\subseteq R'$ entonces:

1. $\chi_A\colon R\to\mathbb{R}$ es integrable si y sólo si $\chi_A\colon R'\to\mathbb{R}$ es integrable.
2. $\int_R\chi_A=\int_{R'}A$.

Proposición 5.2.3 — Ejemplos de conjuntos Jordan-medibles. Las siguientes afirmaciones son válidas:

1. Todos los rectángulos son Jordan-medibles.
2. Para todo conjunto A que es Jordan-medible $v(A)>0$.
3. $v(\emptyset)=0$.
4. Si $A_1,\cdots,A_n$ son conjuntos Jordan medibles disjuntos entonces $\bigcup_{i=1}^n A_n$ es Jordan medible y

$$v\left(\bigcup_{i=1}^n A_n\right)=\sum_{i=1}^n v(A_n).$$

5. Si A,B son Jordan medibles entonces $A\cap B$ es Jordan medible.
6. Si A,B son Jordan medibles entonces $A\cup B$ lo es y

$$v(A\cup B)=v(A)+v(B)-v(A\cap B).$$

7. Si A es Jordan medible, para cada rectángulo R tal que $A\subseteq R$ entonces $R-A$ es Jordan medible.

Teorema 5.2.4 — Caracterización de conjuntos Jordan-medibles. Sea $A\subseteq\mathbb{R}^n$ acotado. Las siguientes afirmaciones son equivalentes

1. A es Jordan-medible.
2. Para toda $\varepsilon>0$ existen $R_1,\cdots,R_k$ rectángulos tales que:
   a) $Fr(A)\subset R_1\cup\cdots\cup R_k$.
   b) $\sum_{i=1}^k m(R_i)<\varepsilon$.
3. El conjunto $Fr(A)$ es Jordan-medible y $v(Fr(A))=0$.

Teorema 5.2.5 — Conjuntos de medida cero. Sea $A \subseteq \mathbb{R}^n$ acotado. Las siguientes afirmaciones son equivalentes

1. A es Jordan-medible y $v(A) = 0$.
2. Para toda $\varepsilon > 0$ existen $R_1, \cdots, R_k$ rectángulos tales que:
   a) $A \subset R_1 \cup \cdots \cup R_k$.
   b) $\sum_{i=1}^{k} m(R_i) < \varepsilon$.

Proposición 5.2.6 Sea $f : R \subseteq \mathbb{R}^n \to \mathbb{R}$ una función acotada. Si el conjunto de discontinuidades de f en R es Jordan-medible y de medida cero entonces f es integrable.

Definición 5.2.2 — Integral sobre conjuntos acotados.. Sea $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ una función acotada sobre un conjunto acotado A. Decimos que f es integrable sobre A si existe un rectángulo $A \subseteq R$ tal que $f\chi_A : R \to \mathbb{R}$ es integrable sobre R, además

$$\int_A f := \int_R f\chi_A$$

Aquí cabe mencionar que si $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ es una función acotada, lo extendemos naturalmente a $f : \mathbb{R}^n \to \mathbb{R}$ como sigue $f(x) = x$ si $x \in \mathbb{R}^n - A$.

Proposición 5.2.7 La integral de la definición (5.2.2) no depende de la elección del rectángulo.

Proposición 5.2.8 Sea $f : A \to \mathbb{R}^n$ acotada sobre un conjunto Jordan medible A. Si el conjunto de discontinuidades de f sobre A es un conjunto Jordan medible y de medida cero entonces f es integrable sobre A.

Proposición 5.2.9 Si $f : A \cup B \to \mathbb{R}$ es integrable sobre A y sobre B entonces f es integrable sobre $A \cap B$ y $A \cup B$, además

$$\int_{A\cup B} f = \int_A f + \int_B f - \int_{A\cap B} f.$$

Proposición 5.2.10 — Propiedades de integrales sobre conjuntos acotados. Sean $f, g : A \to \mathbb{R}$ funciones acotadas, entonces son válidos:

1. Si f es continua entonces f es integrable.
2. $f + g$ es integrable y $\int_A (f + g) = \int_A f + \int_A g$.
3. Para cada $c \in \mathbb{R}$, cf es integrable y $\int_A (cf) = c\int_A f$.
4. fg es integrable sobre A.
5. Si $f \geq 0$ sobre A, entonces $\int_A f \geq 0$.
6. Si $f \leq g$ sobre A, entonces $\int_A f \leq \int_A g$.
7. $\|f\|$ es integrable sobre A y $\|\int_A f\| \leq \int_A \|f\|$.

Teorema 5.2.11 — Teorema del valor medio para integrales. Sean $f, g : A \to \mathbb{R}$ funciones con f continua y g integrable sobre A. Las siguientes afirmaciones son válidas:

1. Existe $h \in A$ tal que $\int_A f = f(h)v(A)$.
2. Si $g \geq 0$ en R entonces existe $h \in A$ tal que $\int_A fg = f(h)\int_A g$.

5.3. Teoremas importantes de integración en varias variables.

5.3.1. Teorema de Fubini.

El siguiente teorema permite calcular ciertas integrales de varias variables como integrales iteradas. Sea $R := [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n] \subseteq \mathbb{R}$ un rectangulo y $f : R \to \mathbb{R}$ acotado. Para cada $i = 1, \cdots, n$ fijo aleatorio, tenemos la biyección

$$\begin{array}{ccc} \mathbf{Fun}(R, \mathbb{R}) & \cong & \mathbf{Fun}([a_1, b_1] \times [a_{i-1}, b_{i-1}] \times [a_{i+1}, b_{i+1}] \times \cdots [a_n, b_n], \mathbf{Fun}([a_i, b_i], \mathbb{R})) \\ f & \leftrightarrow & \hat{y} \to (f_{\hat{y}} \colon [a_i, b_i] \to \mathbb{R}) \end{array}$$

donde a cada $\hat{y} = (y_1, y_2, \cdots, y_{i-1}, y_i, \cdots, y_n)$, la función $f_{\hat{y}} \colon [a_i, b_i] \to \mathbb{R}$ es definido como

$$f_{\hat{y}}(x) := f(y_1, \cdots, y_{i-1}, x, y_{i+1}, \cdots, y_n)$$

Por simplicidad, denotamos $R_i := [a_1, b_1] \times [a_{i-1}, b_{i-1}] \times [a_{i+1}, b_{i+1}] \times \cdots [a_n, b_n]$

Teorema 5.3.1 — Teorema de Fubini. Usando la notación anterior, si f es integrable sobre R entonces las funciones de $R_i \times \cdots [a_n, b_n] \to \mathbb{R}$ definidos

$$\begin{array}{ccc} \phi_{\sup}(\hat{y}) & := & \sup \int_{[a_i, b_i]} f_{\hat{y}}(x)dx \\ \phi_{\text{ínf}}(\hat{y}) & := & \text{ínf} \int_{[a_i, b_i]} f_{\hat{y}}(x)dx \end{array}$$

son integrables sobre R_i y además

$$\int_R f = \int_{R_i} \phi_{\sup} = \int_{R_i} \phi_{\text{ínf}}$$

En particular obtenemos

Corolario 5.3.2 Sea $f : [a_1, b_1] \times [a_2, b_2] \to \mathbb{R}$ continua, entonces

$$\int_{[a_1, b_1] \times [a_2, b_2]} f = \int_{a_1}^{b_1} \int_{a_2}^{b_2} f(x, y)dydx = \int_{a_2}^{b_2} \int_{a_1}^{b_2} f(x, y)dxdy$$

Esto se puede generalizar sobre varias variables mayores a 2.

5.3.2. Teorema de Cambio de Variable.

Teorema 5.3.3 — Teorema de Cambio de Variable. Sea $A \subseteq \mathbb{R}^n$, $f \colon A \to \mathbb{R}$ continua en A, $g \colon \Omega \subseteq \mathbb{R}^n \to \mathbb{R}^n$ de clase C^1 en Ω y $B \subseteq \Omega$ Jordan-medible tal que $cl(B) \subseteq \Omega$. Si

1. g es inyectiva en casi B.
2. $\det(Dg(x)) \neq 0$ para casi toda $x \in B$.
3. $g(B) = A$.

Entonces A es Jordan medible y

$$\int_A f = \int_B (f \circ g) \left\| \det Dg \right\|$$

5.4. Algunas aplicaciones.

5.4.1. Centro de masa.

5.4.2. Integral de Riemann–Stieltjes y Probabilidad en variables aleatorias continuas.

Capítulo 6

Cálculo vectorial en variedades.

6.1. Interpretaciones de Campos Vectoriales.

6.2. Espacio tangente y conceptos del algebra multilineal.

6.3. Formas diferenciales en dimensiones bajas.

6.3.1. Gradiente de un campo vectorial

6.3.2. Operador estrella de Hodge y Campos vectoriales.

6.3.3. Volumen.

6.3.4. Rotacional.

6.3.5. Divergencia.

6.4. Integral sobre curvas y superficies.

6.4.1. Integración escalar.

6.4.2. Integración sobre campos vectoriales y sus diferentes presentaciones.

6.4.3. Integración y frontera geométrica.

Capítulo 7

Teoremas de Integración y aplicaciones

7.1. Teorema fundamental para curvas.

7.2. Caso Plano: Teorema de Green.

7.3. Caso Tridimensional: Teorema de Stokes.

7.4. Teorema de la divergencia y sus interpretaciones.

7.5. Aplicación 1: Electromagnetismo.

7.6. Aplicación 2: Mecánica de Fluidos, Ecuacion de Navier-Stokes.

7.7. Aplicacion 3: Generalizaciones de la Regla de Integral de Leibniz

Capítulo 8

Transformadas de Fourier.

8.1. Convergencia de funciones.

8.2. Convoluciones y sus interpretaciones.

8.3. La transformada de Fourier y su convergencia.

8.4. Aplicaciones de la transformada de Fourier.

8.5. La trasnformada de Laplace.

8.6. Aplicaciones de la transformada de Laplace.

Apéndice A

Notas sobre Algebra Lineal

Estas notas se especializarán para espacios vectoriales sobre $\mathbb{R}$, sin embargo mucho de lo discutido aquí, se puede extender a campos arbitrarios o campos totalmente ordenados (en el caso del producto interior), con sus debidos cuidados.

Definición A.0.1 — Espacio vectorial. Un espacio vectorial sobre $\mathbb{R}$ consiste en un conjunto V, dotado de una operación interna llamado suma $+: V \times V \to V$ y una operación externa llamado multiplicación escalar $\cdot: \mathbb{R} \times V \to V$ tal que satisfacen las siguientes propiedades, para toda $v, w, z \in V$ y toda $\lambda, \kappa \in \mathbb{R}$:

1. $v + (w + z) = (v + w) + z$.
2. Existe un elemento $0 \in V$ tal que $v + 0 = v = 0 + v$.
3. Para cada $v \in V$ existe $-v \in V$ tal que $v + (-v) = 0 = (-v) + v$.
4. $v + w = w + v$.
5. $\lambda * (\kappa * v) = (\lambda \kappa) v$.
6. $\lambda(v + w) = \lambda v + \lambda w$.
7. $(\kappa + \lambda) v = \kappa v + \lambda v$.
8. $1 * v = v$.

■ **Ejemplo A.1** $\mathbb{R}^n$ con las operaciones usuales. ■

■ **Ejemplo A.2** Si X es un conjunto, $V = Fun(X, \mathbb{R})$ junto a las operaciones usuales. ■

■ **Ejemplo A.3** $\mathbb{M}_{n \times m}(\mathbb{R})$. ■

■ **Ejemplo A.4** $\mathbb{R}[x] := \{a_0 + a_1 x + \cdots + a_n x^n \mid a_0, \cdots, a_n \in \mathbb{R}\}$ ■

Definición A.0.2 — Funciones lineales. Sean V, W dos espacios vectoriales, una función $f: V \to W$ es lineal si:

1. Para toda $v, w \in V$, se cumple $f(v + w) = f(v) + f(w)$.
2. Para toda $\lambda \in \mathbb{R}$ y $v \in V$ se cumple $f(\lambda v) = \lambda f(v)$

Ejercicio A.1 Sean V, W dos espacios vectoriales y una función $f: V \to W$. Muestra que son equivalentes:

1. f es lineal.
2. f satisface para toda $v, w \in V$ y toda $\lambda \in \mathbb{R}$, $f(v + \lambda w) = f(v) + \lambda f(w)$.

■

Ejercicio A.2 Si V, W son espacios vectoriales, entonces el conjunto de las funciones lineales $\mathrm{Hom}_{\mathbb{R}}(V, W)$ es un espacio vectorial. ■

Definición A.0.3 — Subespacio. Sea $W \subseteq V$ y V un espacio vectorial, decimos que W es un subespacio vectorial si:

1. $0 \in W$.
2. Para toda $\lambda \in \mathbb{R}$ y $v, w \in W$ entonces $v + \lambda w \in W$

■ **Ejemplo A.5** Si V es un espacio vectorial, entonces $\mathrm{Hom}_{\mathbb{R}}(V, \mathbb{R}) \subseteq Fun(V, \mathbb{R})$ es un subespacio vectorial. ■

A.1. Bases y Funciones lineales.

Definición A.1.1 — Subespacio generador. Sea $S \subseteq V$, entonces, definimos el subespacio generado por S

$$\langle S \rangle = \{\sum_{k=1}^{n} \lambda_k v_k |\ \lambda_k \in K,\ v_k \in S,\ n \in \mathbb{N}\}$$

En otras palabras, el subespacio generador es simplemente todas las combinaciones lineales posibles con los elementos de S.

■ **Ejemplo A.6** Si $V = \mathbb{R}^2$ entonces:

1. $\langle \{(1,0),(0,1)\} \rangle = \mathbb{R}^2$.
2. $\langle \{(2,3),(-1,1)\} \rangle = \mathbb{R}^2$.
3. $\langle \{(1,1),(2,2)\} \rangle = \{(x.x) |\ x \in \mathbb{R}\}$.
4. $\langle \{(1,1),(2,3),(6,-1) \rangle = \mathbb{R}^2$,
5. $\langle \{(2,4),(-1,-2),(10,20)\} \rangle = \{(t,2t) |\ t \in \mathbb{R}\}$

■

Notemos que es posible encontrar conjuntos distintos que generan el mismo espacio vectorial. Una pregunta natural ¿Que propiedades debe tener un conjunto de tal manera de que cada elemento sea importante para generar dicho conjunto?.

Proposición A.1.1 Sea $S \subseteq V$ entonces son equivalentes:

1. $\langle S \rangle = \langle S - \{x_0\} \rangle$, para algún $x_0 \in S$.
2. Existen $v_1, \cdots, v_n \in S$ tales que existen $\lambda_1, \cdots, \lambda_n$ no todos cero con la propiedad:

$$\sum_{k=1}^{n} \lambda_k v_k = 0$$

Demostración. Si para algún $x_0 \in S$ se cumple que $\langle S \rangle = \langle S - \{x_0\} \rangle$, entonces existen $x_1, \cdots, x_n \in S - \{x_0\}$ tales que:

$$x_0 = \sum_{i=1}^{n} \alpha_i x_i,\ \alpha_i \in K$$

con α_i no todos cero. Conversamente, tenemos que $\langle S - \{x_0\} \} \subseteq \langle S \rangle$ para todo $x_0 \in S$, por hipótesis, existen $v_1, \cdots, v_n \in S$ tales que existen $\lambda_1, \cdots, \lambda_n$ no todos cero con la propiedad:

$$\sum_{k=1}^{n} \lambda_k v_k = 0$$

supongamos sin pérdida de generalidad que $\lambda_1 \neq 0$ entonces:

$$v_1 = \sum (-\frac{\lambda_k}{\lambda_1}) v_k$$

esto implica que para cada $v \in \langle S \rangle$ tal que:

$$v = \alpha_1 v_1 + \sum_{v \neq v_0} \lambda_i v_i$$

entonces:

$$v = \sum_{v \in S - \{v_1\}} \delta_v v$$

por lo tanto $\langle S \rangle = \langle S - \{v_1\} \rangle$. ■

Definición A.1.2 — Dependencia e Independencia Lineal. Decimos que un subconjunto $S \subseteq V$ es linealmente dependendiente si satisface alguna de las condiciones equivalentes de (A.1.1). En caso contrario diremos que S es linealmente independiente.

Definición A.1.3 — Base. Un subconjunto $S \subseteq V$ se llama base de V si satisface las siguientes propiedades:

1. S es linealmente independiete.
2. $V = \langle S \rangle$.

En cursos de algebra lineal, se muestra las siguientes observaciones importantes:

1. Si $\beta \subseteq V$ es una base, entonces cada elemento $x \in V$ tiene únicos coeficientes $\lambda_b \in \mathbb{R}$ con una única combinación lineal
$$x = \sum_{b \in \beta} a_b b, \ b \text{ casi todos cero, excepto un número finito de ellos.}$$
2. Todo espacio vectorial V tiene una base(No necesariamente única).
3. Todas las bases son equipotentes, es decir, tienen la misma cardinalidad. A esa cardinalidad se le conoce como la dimensión de v.
4. $\dim_{\mathbb{R}} \mathbb{R}^n = n$ y la base $\{e_i\}$ donde $e_i = (0, \cdots, 0, 1, 0, \cdots, 0)$ es conocida como la base estándar.

Proposición A.1.2 — Propiedad universal de las bases. Sea V un espacio vectorial y $\beta \subseteq V$ una base, para toda función $f \colon \beta \to W$, existe una única función lineal $T \colon V \to W$ tal que $T|_\beta = f$.

Demostración. Mostremos la unicidad, supongamos que $T, H \colon V \to W$ son dos funciones lineales tales que $T|_\beta = f = H|_\beta$. Tomemos $x \in V$ entonces tiene una única combinación lineal

$$x = \sum_{i=1}^{n} \alpha_i e_i$$

Aplicando T y la propiedad, obtenemos

$$\begin{aligned}
T(x) &= T\left(\sum_{i=1}^{n} \alpha_i e_i\right) \\
&= \sum_{i=1}^{n} \alpha_i T(e_i) = \sum_{i=1}^{n} \alpha_i f(e_i) \\
&= \sum_{i=1}^{n} \alpha_i H(e_i) \\
&= H\left(\sum_{i=1}^{n} \alpha_i e_i\right) = H(x)
\end{aligned}$$

Obteniendo $H = T$, concluyendo la unicidad. Para la existencia, defina $T\colon V \to W$ como sigue:

$$T(x) := \sum_{i=1}^{n} \alpha_i f(e_i)$$

donde $x = \sum_{i=1}^{n} \alpha_i e_i$ es la única combinación lineal en términos de la base. Es fácil ver que T está bien definida, es lineal y satisface las hipótesis de la proposición. ■

Algunas consecuencias interesantes de (A.1.2) son las siguientes.

Corolario A.1.3 Se tiene un isomorfismo canónico

$$Fun(\beta, \mathbb{R}) \cong \mathrm{Hom}_k(V, K)$$

más aún, el espacio dual $V^* := \mathrm{Hom}_K(V, K)$ tiene la misma dimensión que V. Más especificamente, del isomorfismo canónico, si $e_1, \cdots, e_n$ es una base de V, sea $e_i^*\colon V \to K$ la única función K-lineal (definido por (A.1.2)) tal que $e_i^*(e_j) = \delta_{ij}$, entonces $e_1^*, \cdots, e_n^*$ es la base de V^*.

Corolario A.1.4 Sean V, W espacios vectoriales con bases β_V, β_W respectivamente. Consideremos $f\colon \beta_V \to \beta_W$ una función y $T_f\colon V \to W$ su función lineal asociada (A.1.2). Entonces:

1. T_f es isomorfismo si y sólo si f es biyectiva.
2. T_f es inyectiva si y sólo si f es inyectiva.

Ahora, sea V un espacio vectorial de dimensión finita (digamos n), tenemos que el isomorfismo (pasando por el isomorfismo de A.1.4):

$$\gamma_S\colon V \to K^n$$

Depende de la elección de la base. Más especificamente, si $S = \{v_1, \cdots, v_n\}$ es una base ordenada de V, el isomorfismo está dado como:

$$\gamma_S\left(\sum_{i=1}^{n} \lambda_i v_i\right) = (\lambda_1, \cdots, \lambda_n)$$

denominamos al valor $\gamma_S(v)$ como la **matriz (o el vector) de representación de v en la base S**. Ahora si $P = \{w_i\}_{i=1}^{n}$ es otra base de V, usando la biyección $S \cong P$ entonces existe un isomorfismo lineal $v_S^P\colon K^n \to K^n$ construido como sigue:

1. Si $\sum_{i=1}^{n} \lambda_i v_i = v = \sum_{i=1}^{n} \theta_i w_i$ son sus únicas descripciones de v en cada base. En particular para cada $j = 1, \cdots, n$ existen únicos escalares $\theta_i^{(j)}$ tales que:

$$v_j = \sum_{i=1}^{n} \theta_i^{(j)} w_i$$

2. Entonces en general se tiene:

$$\sum_{i=1}^{n} \theta_i w_i = v = \sum_{j=1}^{n} \lambda_j v_j = \sum_{j=1}^{n} \lambda_j \left(\sum_{i=1}^{n} \theta_i^{(j)} w_i\right)$$

3. Esto anterior usando sus matrices de representación, definamos la matriz cuadrada $A = (a_{ij})$ como $a_{ij} := \theta_i^{(j)}$ y entonces se tiene:

$$\begin{pmatrix} \theta_1 \\ \vdots \\ \theta_n \end{pmatrix} = \begin{pmatrix} \theta_1^{(1)} & \cdots & \theta_1^{(n)} \\ \vdots & & \vdots \\ \theta_n^{(1)} & \cdots & \theta_n^{(n)} \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$$

Es decir $v_S^P(\gamma_S(v)) = \gamma_P(v)$.
Basta probar que es un isomorfismo.

Proposición A.1.5 Para todo par de bases finitas S, P en V se cumple:

$$v_S^P \circ v_P^S = \mathrm{Id}_{K^n} = v_P^S \circ v_S^P$$

Para acompletar, tenemos el siguiente lema técnico.

Proposición A.1.6 Sea K^n con la base canónica, entonces para toda función lineal $f\colon K^n \to K^m$ le corresponde una única matriz $A \in \mathbf{M}_{n\times m}(K)$ tal que para toda $v \in K^n$ se cumple:

$$f(v) = A \cdot v$$

Esta manera de pensar en representaciones matriciales se puede extender para todo morfismo lineal, como sigue.

Teorema A.1.7 — Representación matricial. Sean V, W dos espacios de dimensión finita, fijemos las bases $S_V \subseteq V$ y $S_W \subseteq W$ y sea $f\colon V \to W$ una función K-lineal, entonces existe una única matriz T_f cuya única función lineal asociada $\hat{f}(v) = T_f \cdot v$ satisface el siguiente diagrama conmutativo de funciones K-lineales:

$$\begin{tikzcd} V \arrow[r, "\gamma_{S_V}"] \arrow[d, "f"'] & K^n \arrow[d, "T_f \cdot v"] \\ W \arrow[r, "\gamma_{S_W}"'] & K^m \end{tikzcd}$$

en donde $n = \dim_K V$ y $m = \dim_K W$.

Definición A.1.4 — Representación matricial. Sean V, W dos espacios de dimensión finita, fijemos las bases $S_V \subseteq V$ y $S_W \subseteq W$ y sea $f\colon V \to W$ una función K-lineal, definimos la matriz de representación de f con respecto a las bases como la única matriz A obtenida de (A.1.6) de la función lineal T_f asociada a f (A.1.7). Usualmente es denotado como

$$[f]_{S_V}^{S_W}$$

Proposición A.1.8 — Propiedades de una representación matricial. Sean V_1, V_2, V_3 espacios vectoriales, con bases $\beta_1, \beta_2, \beta_3$ respectivamente, entonces:

1. $[Id_{V_1}]_{\beta_1}^{\beta_1} = \mathrm{Id}_{\dim V}$.
2. Si $V_1 = V_2$ entonces $[Id_{V_1}]_{\beta_1}^{\beta_2}$ es la matriz de cambio de base de β_1 a β_2.
3. $[f_2]_{\beta_2}^{\beta_3}[f_1]_{\beta_1}^{\beta_2} = [f_2 \circ f_1]_{\beta_1}^{\beta_3}$.
4. Si f es invertible, entonces ${[f]_{\beta_1}^{\beta_2}}^{-1} = [f^{-1}]_{\beta_2}^{\beta_1}$.

A.2. Sobre productos de espacios vectoriales

Ahora hablaremos unas técnicas particulares del álgebra lineal que serán útiles para entender las demostraciones de la derivada.

Ejercicio A.3 — Producto de espacios vectoriales. Prueba lo siguiente; sean $V_1, \cdots, V_n$ espacios vectoriales sobre $\mathbb{R}$, entonces el conjunto $V_1 \times \cdots \times V_n$ dota de una estructura de espacio vectorial en donde la suma y multiplicación escalar son definidos como sigue:

1. $(v_1, \cdots, v_w) + (w_1, \cdots, w_n) = (v_1 + w_1, \cdots, v_n + w_n)$.
2. $\lambda(v_1, \cdots, v_n) = (\lambda v_1, \cdots, \lambda v_n)$.

De aquí en adelante, fijemos $V_1, \cdots, V_n$ los espacios vectoriales sobre $\mathbb{R}$ y sea $P := V_1 \times \cdots \times V_n$ el producto de espacios vectoriales según el ejercicio anterior, definimos:

$$\begin{array}{cccc} \pi_i & P & \rightarrow & V_i \\ & (v_1, \cdots, v_n) & \rightarrow & v_i \end{array}$$

esta función es denominada i-ésima proyección coordenada. También definimos

$$\begin{array}{cccc} u_i & V_i & \rightarrow & P \\ & v_i & \rightarrow & (0, \cdots, v_i, \cdots, 0) \end{array}$$

esta función es denominada i-ésima inclusión coordenada.

Ejercicio A.4 — **Propiedades de las proyecciones e inclusiones.** Usando la notación anterior prueba lo siguiente:

1. Cada π_i y u_i son funciones $\mathbb{R}$-lineales.
2. $\pi_i u_j = \delta_{ij}$.
3. $\mathrm{Id}_P = \sum_{i=1}^n u_j \pi_j$.
4. Si $\mathbb{R} = V_1 = \cdots = V_n$ y sea $\{e_i\}_{i=1}^n$ la base estándar, entonces $e_i^* = \pi_i$.

Técnica de las funcionales Gracias a las propiedades 2 y 3 del ejercicio (A.4), se puede caracterizar el estudio de funciones lineales sobre el producto de espacios vectoriales. Fijemos $V, W_1, W_2, \cdots, W_n$ espacios vectoriales y sea $F \colon V \rightarrow P$ lineal, entonces

$$\begin{aligned} F &= \mathrm{Id}_P \circ F \\ &= \left(\sum_{i=1}^n u_j \pi_j \right) \circ F \\ &= \sum_{i=1}^n u_j (\pi_j \circ F) \end{aligned}$$

Denotemos $f_i := \pi_i \circ F \colon V \rightarrow W_i$ y las denominaremos como componentes funcionales de F. Esto permite construir una función

$$\begin{array}{cccc} \alpha \colon & \mathrm{Hom}_K(V, P) & \rightarrow & \prod_{i=1}^n \mathrm{Hom}_K(V, W_i) \\ & F & \rightarrow & (f_1, \cdots, f_n) \end{array}$$

que es R-lineal y tiene una inversa

$$\begin{array}{cccc} \beta \colon & \prod_{i=1}^n \mathrm{Hom}_K(V, W_i) & \rightarrow & \mathrm{Hom}_K(V, P) \\ & (g_1, \cdots, g_n) & \rightarrow & \sum_{i=1}^n u_i g_i \end{array}$$

Obteniendo un isomorfismo

$$\prod_{i=1}^n \mathrm{Hom}_K(V, W_i) \cong \mathrm{Hom}_K(V, P)$$

En particular nos dice que toda función lineal F está completamente caracterizada por sus funcionales f_i.

Proposición A.2.1 — Tecnica de las funcionales. Sean (V,β_V), $(W_1,\gamma_1),\cdots,(W_n,\gamma_n)$, parejas de espacios vectoriales de dimensión finita con su respectiva base. Entonces:

1. $\gamma := \bigcup_{i=1}^{n} u_i(\gamma_i)$ es una base del producto P.
2. Si $F: V \to P$ es lineal, entonces

$$[F]_{\beta_V}^{\gamma} = \left([f_1]_{\beta_V}^{\gamma_1}|\cdots|[f_n]_{\beta_V}^{\gamma_n}\right)$$

A.3. Producto interior y teorema de Representación de Riez.

Definición A.3.1 — Producto Interior. Un producto interior de V es una función $b\colon V \times V \to \mathbb{R}$ tal que:

1. b es bilineal, es decir, satisface
$$\begin{aligned} b(a_1v_1+v_2,w) &= a_1b(v_1,w)+b(v_2,w) \\ b(v,a_1w_1+w_2) &= a_1b(v,w_1)+b(v,w_2). \end{aligned}$$
2. Si $b(v,v)=0$ entonces $v=0$.
3. b es simétrica, es decir $b(v,w)=b(w,v)$.
4. b es definida positivamente, es decir para toda $x \in V$ entonces $b(x,x) \geq 0$.

■ **Ejemplo A.7** Definimos en $\mathbb{R}^n$ con una base $\{e_1,\cdots,e_n\}$ la siguiente función bilineal

$$\langle \sum_{k=1}^{n} \alpha_k e_k, \sum_{k=1}^{n} \beta_k e_k \rangle = \sum_{k=1}^{n} \alpha_k \beta_k$$

entonces este es un producto interior. ■

Tenemos los isomorfismos naturales:

$$\begin{aligned} \mathscr{L} &: L_K(V,W;K) \to \mathrm{Hom}_K(V,\mathrm{Hom}_K(W,K)), \\ \mathscr{R} &: L_K(V,W;K) \to \mathrm{Hom}_K(W,\mathrm{Hom}_K(W,K)) \end{aligned}$$

Definición A.3.2 — Tipo de forma bilineal. Sea $b\colon V \times W \to K$ una forma bilineal decimos que b es:

1. No degenerado en V si $L_b = 0$.
2. No degenerado en W si $R_b = 0$.
3. Un pareo perfecto si b_V y b_W son isomorfismos.

Observemos que:

1. Decir que b es no degenerado en V es equivalente a pedir la siguiente propiedad: Si $v \in V$ es tal que $b(v,w)=0$ para toda $w \in W$, entonces $v=0$.
2. Decir que b_V es sobreyectiva, es equivalente a pedir que para toda $f\colon W \to K$ K-lineal, existe un vector $v \in V$ tal que
$$f(w) = b(v,w),\ \forall w \in W$$

Lema A.3.1 Si $f\colon V \to W$ es una función lineal inyectiva con $\dim V = \dim W$ entonces f es isomorfismo.

Demostración. Sea $\beta \subseteq V$ una base, entonces por inyectividad, se tiene que $\gamma := f(\beta) \subseteq W$ es un conjunto linealmente independiente de la misma cardinalidad que β, pero $\dim W = \dim V$, entonces γ es una base de W, obteniendo que $f|_\beta : \beta \to \gamma$ es una biyección entre bases y por (A.1.4), entonces f es isomorfismo. ■

Corolario A.3.2 Si $b\colon V\times V\to\mathbb{R}$ es una forma bilineal simetrica no degenerada, con V de dimensión finita, entonces b es un pareo perfecto.

Teorema A.3.3 Si $b\colon V\times V\to\mathbb{R}$ es un producto interior (A.3.1) y $\dim V<\infty$ entonces b es un pareo perfecto.

Demostración. Sea $v\in\ker b_V$ entonces $b(v,w)=0$ para toda $w\in V$, en particular $b(v,v)=0$, entonces por definición (A.3.1) obtenemos que $v=0$, por lo tanto b es simétrico no degenerado y como $\dim V<\infty$ entonces b es pareo perfecto. ∎

Una consecuencia directa es el teorema de Riez

Proposición A.3.4 — Teorema de Riez de dimensión finita.. Si b es un producto interior, entonces para toda función lineal $f\colon V\to\mathbb{R}$ existe un único $v_f\in V$ tal que:

$$f(x)=b(v_f,x)$$

Una manera práctica para encontrar v_f, seria:

1. Construye una base ortonormal con respecto a b, digamos $\{e_i\}_{i=1}^n$.
2. Define $v_f:=\sum_{i=1}^n f(e_i)e_i$.

A.4. Diagonalización y formas bilineales

Definición A.4.1 — Valor propio y vector propio. Sea $T\colon V\to V$ una función K-lineal, si existe $\lambda\in K$ y $v\in V-\{0\}$ tal que

$$T(v)=\lambda v$$

entonces decimos que:

1. λ es valor propio.
2. v es vector propio asociado a λ.

Tenemos que son equivalentes:

1. $T(v)=\lambda v$.
2. $(T-\lambda\,\mathrm{Id}_V)(v)=0$.
3. $\ker(T-\lambda\,\mathrm{Id}_V)\neq 0$.

Y cuando $\dim_K V$ es finita, todo lo anterior es equivalente a decir

$$\det(T-\lambda\,\mathrm{Id}_V)=0$$

Definición A.4.2 — Polinomio característico. Sea $T\colon V\to V$ una función K-lineal, entonces definimos el polinomio caracteristico como:

$$\chi_T(t):=\det|[T]-t\,\mathrm{Id}_V|$$

Esta definición no depende de la base en V, puse si $T=A=PBP^{-1}$, entonces de las siguientes igualdades

$$\det|A-t\,\mathrm{Id}_V|=\det|PBP^{-1}-tPP^{-1}|=\det|P(B-t\,\mathrm{Id}_V)P^{-1}|=\det|P|\det|B-t\,\mathrm{Id}_V|(\det|P|)^{-1}=\det|B-t\,\mathrm{Id}_V|$$

se sigue que $\chi(t)$ no depende de la matriz $[T]$ obtenida al escoger una base. El polinomio característico tiene como principal objetivo, encontrar todos los valores propios de T, calculando sus ceros. Es decir:

- $t = \lambda$ es valor propio de T si y sólo si $\chi_T(\lambda) = 0$.

■ **Ejemplo A.8** Considera $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix}$, entonces $\chi_T(t) = (t-1)^2(t-4)$ por tanto $4, 1$ son los únicos valores propios. ■

Una función $T: V \to V$ es diagonalizable si existe una base $\beta \subseteq V$ tal que $[T]_\beta^\beta$ es una matriz diagonal. En algebra lineal, uno de los temas principales es estudiar formas canonicas, es decir factorizaciones de V en función de T, por ejemplo si una matriz es diagonalizable se puede ver que la base β consiste en vectores propios asociados a cada valor propio, por tanto puedes descomponer V en términos de sus vectores propios (A este resultado suele ser un caso particular del Teorema espectral).

Definición A.4.3 — Matriz asociada a una forma bilineal. Sea $b: V \times V \to K$ una forma bilineal, y $\beta \subseteq V$ una base, definimos la matriz asociada a la forma bilineal (con respecto a la base β) como la matriz

$$B := (b(e_i, e_j))_{e_i, e_j \in \beta}$$

De esta manera para toda $x, y \in V$ escrito en base β, se satisface

$$b(x,y) = x^T B y$$

■ **Ejemplo A.9** Consideremos $b: K^3 \times K^3 \to K$ definido como

$$b\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}, \begin{pmatrix} a \\ b \\ c \end{pmatrix}\right) = xa + xc + zc$$

entonces su matriz asociada con respecto a su base estándar es

$$B = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

■

Teorema A.4.1 Toda matriz B asociada a una forma bilineal b simétrica es diagonalizable

Ahora vamos a discutir acerca de todas las formas bilineales utilizados en la geometria.

Definición A.4.4 — Formas bilineales positivas y negativas.. Sea $B \in \mathbf{M}_{n \times n}(\mathbb{R})$ y definimos $b(x,y) = x^T B y: \mathbb{R}^{2n} \to \mathbb{R}$ su forma bilineal asociada a la matriz B, decimos que b es:

1. semidefinida positivamente, si $b(x,x) \geq 0$.
2. semidefinida negativamente, si $b(x,x) \leq 0$.
3. definida positivamente, si $b(x,x) > 0$ para toda $x \neq 0$.
4. definida negativamente, si $b(x,x) < 0$ para toda $x \geq 0$.

Cuando B es una matriz simétrica, por (A.4.1), existe un cambio de base $B = P^{-1}DP$ donde

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

más aún la base de vectores en $P = (v_1 | v_2 | \cdots | v_n)$ correspondiente a los valores propios, son ortogonales con respecto al producto estandar. Para ver esto, toma λ_i, λ_j valores propios distintos con respecto a sus vectores propios v_i, v_J. Como $B = B^T$, por ser simétrica, tenemos que

$$\langle v_i, Bv_j\rangle = v_i^T Bv_j = (Bv_i)^T v_j = \langle Bv_i, v_j\rangle$$

usando que son vectores propios, entonces

$$\langle v_i, \lambda_j v_j\rangle = \langle \lambda_i v_i, v_j\rangle$$

es decir

$$(\lambda_i - \lambda_j)\langle v_i, v_j\rangle = 0, \Rightarrow \langle v_i, v_j\rangle = 0$$

Permitiendo encontrar la base de vectores propios ortogonales P, esto implica que $P^T = P^{-1}$ y entonces

$$B = P^T DP$$

asi, sea $\beta = \{v_1, \cdots, v_n\}$ entonces al hacer el cambio de base $u := Px$, $v := Py$ tenemos

$$b(x,y) = x^T By = (Px)^T D(Py) = u^T Dv$$

con esta nueva descripción podemos extraer información acerca de b, para ello, definamos $q(x) := b(x,x)$ denominado su forma cuadrática asociada, q satisface $q(\lambda x) = \lambda q(x)$

R

1. Si todos los valores propios son estrictamente positivos, entonces, $q(x) > 0$ para toda $x \neq 0$. En esta situación tenemos que los únicos puntos donde $q(x) = 0$ son $x = 0$.
2. Si todos los valores propios son estrictamente positivos, entonces $q(x) < 0$ para toda $x \neq 0$. En esta situación tenemos que los únicos puntos donde $q(x) = 0$ son $x = 0$.
3. Si el 0 es un valor propio, supongamos sin pérdida de generalidad que $\lambda_1 = \cdots = \lambda_r$ con $r < n$ y el resto son estrictamente positivos(o negativos) entonces el conjunto de los puntos donde $q(x) = 0$ es el subespacio $W = \langle v_1, \cdots, v_r\rangle$.

Apéndice B

Espacios Normados

B.0.1. Sobre funciones multilineales en espacios normados.

Teorema B.0.1 — Bilinealidad. Existe un único isomorfismo canónico entre los siguientes espacios normados

$$L(V, L(W, Z)) \cong B(V \times W, Z)$$

donde V, W, Z son espacios normados, $L(V, L(W, Z))$ el espacio de las funciones lineales continuas y $B(V \times W, Z)$ el espacio de las funciones bilineales continuas.