UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

FACULTAD DE CIENCIAS

"CONSIDERACIONES DIDACTICAS SOBRE EL CALCULO DIFERENCIAL VECTORIAL"

T E S I S

QUE PARA OBTENER EL TITULO DE

M A T E M A T I C O

PRESENTA

ANTONIO ARGUELLES OROZCO

MEXICO, D. F. 1988

INDICE

INTRODUCCION

> "Platón, en la puerta de la Academia donde enseñaba a conocer las ideas eternas, había puesto un rótulo que rezaba así: "Que no entre aquí quien no sepa geometría". En nuestra época, las Facultades donde se enseña a conocer profunda y eficazmente a la naturaleza material que nos rodea prescriben, con menos solemnidad quizá, pero con igual exigencia, una regla de conducta que podría resumirse en estos términos: "que no entre aquí quien no sepa cálculo infinitesimal". (Roberto Saumells)."

La idea fundamental del Cálculo diferencial vectorial es la de aproximación local de funciones ($f: D \subseteq \mathbb{R}^n \to \mathbb{R}^m$) por medio de funciones lineales. Es la enseñanza clásica del Cálculo elemental, la derivada en un punto - para una función de $\mathbb{R}$ en $\mathbb{R}$ - se define como un número y no se hace referencia alguna a funciones lineales por existir una correspondencia biunívoca entre transformaciones lineales de $\mathbb{R}$ en $\mathbb{R}$ y números. A esto se debe que la idea fundamental quede prontamente obscurecida (cfr. Dieudonné J. Fundamentos de Análisis Moderno. Reverté. Barcelona. 1979. página 145). Más adelante, al tocar el Cálculo de funciones de varias variables, los cursos se centran en enseñar algoritmos: como calcular derivadas parciales, direccionales, vectores gradientes; cómo aplicar todos estos conceptos para diferentes propósitos, etc.. Pero con mucha frecuencia se omite la explicación de esa idea fundamental: pensar en la diferenciabilidad como la posibilidad de aproximar linealmente de manera local.

Ocurre entonces que cuando los estudiantes quieren abordar el estudio de algunas ramas de las Matemáticas superiores vinculadas al Cálculo vectorial, como pueden ser Geometría diferencial, Topología diferencial, Teoría de operadores, ecuaciones diferenciales, etc., encuentran una solución de continuidad entre el contenido de los cursos que han recibido y esta perspectiva del Cálculo que es la requerida en el estudio de esas áreas.

El presente trabajo constituye una propuesta dadáctica para contribuir a resolver esa situación. Se trata de unas consideraciones dirigidas al profesor, a fin de orientarle sobre el modo de presentar este enfoque del Cálculo diferencial vectorial desde que se introduce el Cálculo de varias variables, lo cual suele hacerse en los primeros semestres de las carreras en donde se estudia esa asignatura. De esta forma se evita el tener que aguardar hasta cursos más avanzados de Análisis matemático para cubrir ese material. Resulta interesante esto, pues en los planes de estudio de varias disciplinas científicas: Ingeniería, Economía, Física, etc. no están previstos tales cursos y sí en cambio se requiere a veces de esa perspectiva un tanto más abstracta del Cálculo.

Las páginas de este trabajo no pretenden ser unas notas para el curso de Cálculo diferencial vectorial, pues además de considerarlo una pretensión demasiado ambiciosa, resultaría inoperante ya que un curso así varía bastante dependiendo del contexto en el que se estudie. (Como se mencionó anteriormente, la idea es que el material que se propone pueda servir para distintas especialidades). Se incluyen, para ser presentados a los estudiantes, textos con el contenido matemático necesario para conseguir el objetivo propuesto. Esas partes van siempre entre comillas para distinguirlas del texto principal que está dirigido al profesor. En él se hacen comentarios y observaciones acerca del contenido y la forma de presentar el material: aspectos que conviene resaltar, recursos que pueden facilitar el aprendizaje, etc..

Cabe señalar que los temas aquí presentados deben complementarse con los que constituyen el temario tradicional de un curso de Cálculo diferencial de varias variables. Por otro lado, no se trata de convertir el curso de Cálculo en uno de Análisis matemático. Por tanto, si bien se busca reforzar la parte conceptual, el énfasis debe enfocarse - según la opinión del autor - en aprender a <u>usar</u> las técnicas del Cálculo para las distintas aplicaciones de éste. Sin embargo, el tener cierta comprensión de esa idea central del Cálculo diferencial redundará en un mejor manejo de los algoritmos. Además se estará en condiciones de más fácilmente proseguir el estudio de otras ramas de las Matemáticas.

Cada capítulo comienza con una introducción al mismo (aunque no en todos los casos dicha introducción se señale de manera explícita) en la que se explica el contenido y el orden de los temas tratados. No obstante, a manera de presentación, es oportuno bosquejar brevemente el esquema general de la obra:

Se comienza revisando la noción de derivabilidad para las funciones de $\mathbb{R}$ en $\mathbb{R}$, aprovechando esta ocasión para hacer ver que la existencia de la derivada de una función en un punto implica la existencia de una función lineal que aproxima a la función dada alrededor del punto en cuestión.

En el capítulo 2 (funciones de $\mathbb{R}$ en $\mathbb{R}^n$) y la primera parte del capítulo 3 (funciones de $\mathbb{R}^n$ en $\mathbb{R}$) se recoge un material usual en los cursos de Cálculo: la manera de introducir a las funciones en cuestión, los recursos para visualizar su comportamiento y los conceptos de límite, continuidad y derivada, pues son temas preliminares necesarios para el estudio de la diferenciabilidad.

Lo que sirve para motivar el estudio de la diferenciabilidad desde otra perspectiva: no como la existencia de un punto (en $\mathbb{R}$ ó $\mathbb{R}^n$) (que resulta de un cierto límite) sino como aproximación lineal local, es precisamente el advertir - en el caso de funciones reales de varias variables - que las derivadas (parciales y direccionales) no resultan una generalización completamente adecuada del concepto de derivada y sus implicaciones para las funciones de $\mathbb{R}$ en $\mathbb{R}$. Lo anterior supone un conocimiento de las transformaciones lineales al que está dedicado el capítulo 4.

Ya con los elementos necesarios sobre transformaciones lineales, se analiza con ese enfoque la diferenciabilidad de las funciones en los distintos casos que se han trabajado: funciones de $\mathbb{R}$ en $\mathbb{R}$, de $\mathbb{R}$ en $\mathbb{R}^n$, de $\mathbb{R}^n$ en $\mathbb{R}$.

(capítulo 5). De esta forma se puede emprender el estudio de la diferenciabilidad para el caso general, i.e. el de las funciones $\mathbb{R}^n$ en $\mathbb{R}^m$ (capítulo 6).

En el último capítulo se presenta la Regla de la Cadena. Este tema manifiesta claramente la ventaja de haber hecho énfasis en la diferenciabilidad como aproximación lineal. Ya que las fórmulas clásicas que dan las derivadas parciales de una función compuesta pueden parecer bastante artificiosas, en cambio la afirmación del teorema resulta clara si se razona a base de aproximaciones lineales: la diferencial de una composición de funciones es la composición de las diferenciales.

Las características didácticas con las que se busca facilitar el aprendizaje y que están presentes a lo largo de todo el trabajo son las siguientes:

- Al introducir algún objeto matemático o concepto nuevo, tratar de relacionarlo con otros ya conocidos.

- El recurso a la geometría: si hay imágenes que puedan ser asociadas a los conceptos, éstos se asimilan más fácilmente.

- Todo el desarrollo de la diferenciabilidad se hace a través de dos lineas paralelas: la perspectiva analítica y la geometría del tema.

- Remitirse frecuentemente a ejemplos específicos que permitan ver al estudiante cómo se llevan a la práctica y se trabaja operativamente con las ideas que se han estado manejando.

- Hacer las formulaciones para los casos generales después de haber analizado casos concretos.

Con la presentación de la diferenciabilidad en los términos que se proponen, se conseguirá mayor unidad en los temas que tradicionalmente integran el curso de Cálculo diferencial vectorial (derivadas, diferencial, diferencial total, matriz jacobiana, vector gradiente, obtención de planos tangentes a superficies, etc.). Esta unificación del material facilita al estudiante estructurar este nuevo conocimiento, lo cual repercute en una mayor claridad en los conceptos.

Es interesante resaltar que la concepción misma del Cálculo diferencial como la teoría de aproximaciones lineales a funciones, corresponde a una idea muy refinada y relativamente novedosa del Cálculo y dista bastante del concepto original de esta rama de las Matemáticas. Ya lo había dicho David Hilbert:

> "Cuán propio de la naturaleza de la ciencia matemática es que cada auténtico avance vaya unido a la invención de herramientas más potentes y de métodos más sencillos, herramientas y métodos que ayuden a comprender teorías previas y a desechar otros desarrollos más complicados."

CAPITULO 1

LA DERIVADA DE FUNCIONES DE $\mathbb{R}$ EN $\mathbb{R}$.

Cuando se emprende un conocimiento nuevo, es preferible, porque lo facilita, apoyarse en conocimientos que ya se poseen y que se manejan con familiaridad y soltura.

Esta es la razón por la que se sugiere empezar el estudio de la diferenciabilidad de las funciones de varias variables, revisando la idea de derivada para funciones real valuadas. Además de servir de repaso, se puede aprovechar la oportunidad para reformular la definición de diferenciabilidad que los estudiantes manejan, a fin de presentarla en términos que permitan, luego, relacionarla fácilmente con la correspondiente para funciones de varias variables.

La definición del Cálculo elemental que los estudiantes conocen es la siguiente:

"Definición. La función $f: D \subseteq \mathbb{R} \to \mathbb{R}$ es derivable en x_0 ($x_0 \in D$) si existe

$$\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}$$

(i.e. $$\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h} = L \quad \text{con } L \in \mathbb{R})$$

o equivalentemente si existe $$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

Este límite se llama la derivada de f en X_0 y se denota tradicionalmente por $f'(X_0)$ ó

$$\left.\frac{df}{dx}\right|_{x = x_0}"$$

Asociada a esta definición está la correspondiente interpretación geométrica que les resulta familiar:

"En la gráfica de la función f, la expresión:

$$\frac{f(x) - f(x_0)}{x - x_0}$$

es la pendiente (la tangente del ángulo de inclinación) de la recta que pasa por los puntos $(x_0, f(x_0))$ y $(x, f(x))$. (ver Figura 1.1)

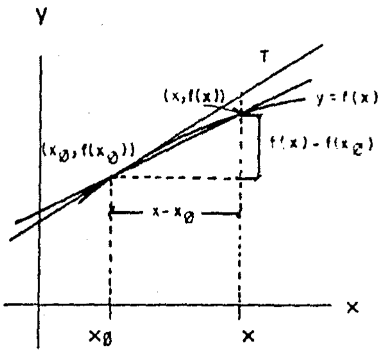


Figura 1.1

A medida que x se aproxima a x_0 esta recta se acerca cada vez más a la recta T que es tangente a la gráfica de la función en el punto $(x_0, f(x_0))$. Por lo tanto $f'(x_0)$ que es el valor límite de las pendientes de las rectas que pasan por $(x_0, f(x_0))$ y $(x, f(x))$ es igual al valor de la pendiente de la recta tangente a la curva en el punto $(x_0, f(x_0))$."

Aprovechando que los estudiantes ya han asimilado estas ideas, conviene hacerles notar que, si bien se tiene una idea intuitiva de lo que es una recta tangente a una curva en un punto, definir este concepto en términos geométricos precisos resulta complicado. Es por ello que la definición más económica que puede darse de recta tangente a una curva en un punto es precisamente la que se da en base al concepto de derivada, a saber: "la recta tangente a la gráfica de la función $y = f(x)$ en el punto $(x_0, f(x_0))$ es aquella recta que pasa por el punto $(x_0, f(x_0))$ y tiene pendiente igual a $f'(x_0)$ donde

$$f'(x_0) = \lim_{h \to 0} \frac{f(x_0+h) - f(x_0)}{h} = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0};$$

siempre y cuando este límite esté bien definido, i.e. la recta tangente no siempre existe."

Como los estudiantes han manejado el criterio de que si está bien definida la recta tangente a la gráfica de la función en un punto (propiedad geométrica), entonces la función es derivable en el punto (condición análitica); la observación anterior contribuye a hacerles ver que pueden intercambiarse los papeles, y que sean las condiciones analíticas (si existe el límite

$$\lim_{h\to 0} \frac{f(x_0 + h) - f(x_0)}{h}$$)

las que definan las propiedades geométricas (entonces se define la recta tangente como la recta con pendiente $f'(x_0)$ que pasa por el punto $(x_0, f(x_0))$ de la gráfica de f). Esto permitirá luego, por ejemplo, definir el espacio tangente a una función real de varias variables si ésta es diferenciable.

Nuestro objetivo final será hacer ver que la existencia de dicho límite es equivalente a la existencia de una transformación lineal de $\mathbb{R}$ en $\mathbb{R}$ (*) que satisface una cierta condición de aproximación local a la función, pues éste es el concepto de diferenciabilidad para funciones vectoriales en general. Sin embargo, para no introducir demasiado material nuevo en esta primera parte, que sólo pretende ser una revisión de diferenciabilidad para el caso unidimensional, será suficiente por el momento con ver que la definición de derivabilidad de una función f en el punto x_0, puede reformularse introduciendo a la función cuya gráfica es la recta tangente a la gráfica de f en el punto $(x_0, f(x_0))$. Para ello se puede seguir el siguiente razonamiento:

(*) A este nivel hablaremos de funciones lineales de $\mathbb{R}$ en $\mathbb{R}$ en el contexto de la Geometría analítica elemental, que los estudiantes conocen. No se hará en el contexto del Algebra lineal, pues se propone hacer un estudio detallado de las transformaciones lineales más adelante, ya que éstas no son, por ahora, necesarias: se puede adelantar bastante en el curso sin recurrir a ellas.

"Si la derivada de una función $y = f(x)$ en el punto x_0, se definió como:

$$f'(x_0) = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h} = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

$$\Rightarrow \lim_{x \to x_0} \left[\frac{f(x) - f(x_0)}{x - x_0} - f'(x_0) \right] = 0 \ldots\ldots\ (1)$$

Ahora, $f(x)$ es la ordenada del punto $(x, f(x))$ en la gráfica de f, mientras que la recta tangente a la gráfica de f en el punto $(x_0, f(x_0))$ tiene por ecuación:
$y - f(x_0) = f'(x_0)(x-x_0) \Rightarrow$ la función F cuya gráfica es la recta tangente a la curva en $(x_0, f(x_0))$ está dada por:

$$F(x) = f(x_0) + f'(x_0)(x-x_0)$$

Observamos que es una función de la forma $F = mx + b$ con m y b constantes. En general a una función de esa forma se le llama una función afín, en particular si $b = 0$, es decir, F es de la forma $F(x) = mx$, se le llama una función lineal. La gráfica de una función afín es siempre una línea recta, y si $b = 0$ dicha recta pasa por el origen.

Podemos considerar la diferencia dada por:
$R(x, x_0) = f(x) - F(x) = f(x) - f'(x_0)(x-x_0) - f(x_0)$."

Hay que hacer la observación de que $R(x, x_0)$ es el error que cometemos si en vez de tomar el valor $f(x)$ como la imagen de x bajo f, tomamos $F(x)$. "La diferencia $R(x, x_0) = f(x) - F(x)$ depende de x y de x_0. Tenemos entonces que

$$\frac{R(x,x_0)}{x-x_0} = \frac{f(x) - F(x)}{x-x_0} = \frac{f(x) - f(x_0)}{x-x_0} - \frac{f'(x_0)(x-x_0)}{x-x_0}$$

$$\Rightarrow \lim_{x \to x_0} \frac{R(x,x_0)}{x-x_0} = \lim_{x \to x_0} \left[\frac{f(x) - f(x_0)}{x-x_0} - f'(x_0) \right] = 0$$

(en virtud de (1)).

Lo que se interpreta diciendo que cuando $x \to x_0$, la diferencia $R(x,x_0) = f(x) - F(x)$ se va a cero más rápido que $x-x_0$. En términos más precisos:

$$\forall \varepsilon > 0 \quad \exists \delta > 0 \quad \text{tal que} \quad \left| \frac{R(x,x_0)}{x-x_0} \right| < \varepsilon \text{ si } |x-x_0| < \delta$$

$$\Rightarrow |R(x,x_0)| = |f(x) - F(x)| < \varepsilon\, |x-x_0| \quad \ldots\ldots\ldots\ldots (2)$$

Lo que nos dice que podemos hacer que la diferencia entre los valores de $f(x)$ y de $F(x)$ sea tan pequeña como queramos siempre y cuando x esté suficientemente cercano a x_0.

Se dice entonces que F es la aproximación lineal de f en x_0, en el sentido de que la función cuya gráfica es la recta tangente da una buena aproximación de la función cerca del punto de tangencia. Donde por buena aproximación, entendemos que para cada x que esté cerca de x_0, la diferencia $|f(x) - F(x)|$ es pequeña, más pequeña en general que la diferencia $|x-x_0|$.

Es claro de (2) que entre más nos acercamos a x_0, la aproximación entre F y la función f es mejor.

Por ejemplo en la figura 1.2, en x_1 estamos demasiado lejos."

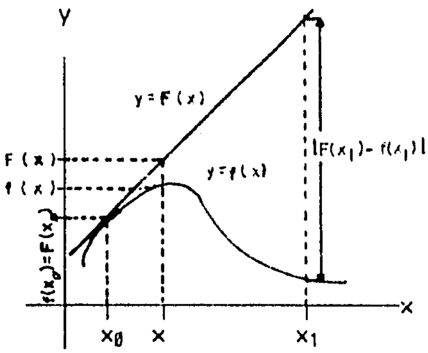


Figura 1.2

Se podría hacer a los estudiantes el comentario de que en sentido estricto la expresión aproximación lineal constituye un abuso de lenguaje porque en general F es una función afín, pero la función afín puede pensarse como una función lineal trasladada.

Como es muy factible que la idea de pensar en una función como "aproximación" a otra función sea novedosa para los estudiantes, conviene presentar algunos ejemplos específicos que les permitan ver en términos prácticos en qué consiste esa idea. Pueden ser ejemplos como los propuestos a continuación:

"Dadas las funciones siguientes, encontraremos la función afín F que aproxima a cada una alrededor del punto especificado. Verifiquemos con algunos valores, que para puntos cercanos al especificado las imágenes de dichos puntos bajo la función f y bajo la función afín F están cercanas.

(a) $f(x) = 1/x \quad x_0 = 1$

Hemos visto que $F(x) = f'(x_0)(x-x_0) + f(x_0)$ nos da la función afín que aproxima bien a f.

En nuestro caso $f'(x) = -1/x^2 \Rightarrow f'(1) = -1/(1)^2 = -1$

$$f(x_0) = f(1) = 1$$

$\therefore F(x) = (-1)(x-1) + 1 = -x + 1 + 1 = -x + 2$

con lo que cerca de $x_0 = 1$ esperamos $1/x \sim -x + 2$

$f(x) = 1/x$	$F(x) = -x + 2$	$\lvert R(x,x_0)\rvert = \lvert f(x)-F(x)\rvert$
$f(.5)=2$	$F(.5)=1.5$	$\lvert R(.5)\rvert = 0.5$
$f(.8)=1.25$	$F(.8)=1.2$	$\lvert R(.8)\rvert = 0.05$
$f(.9)=1.111111$	$F(.9)=1.1$	$\lvert R(.9)\rvert = 0.011111$
$f(.999)=1.001001$	$F(.999)=1.001$	$\lvert R(.999)\rvert = 0.000001$
$f(1)=1$	$F(1)=1$	$\lvert R(1)\rvert = 0$

(b) $f(x) = \ln x \quad x_0 = 1$

$F(x) = f(x_0) + f'(x_0)(x-x_0)$

$f'(x) = 1/x \Rightarrow f'(1) = 1$

$f(x_0) = f(1) = 0$

$\therefore F(x) = 0 + 1(x-1) = x-1$

$\therefore$ cerca de $x_0 = 1 \quad \ln x \sim x-1$

$f(x) = \ln x$	$F(x) = x - 1$	$\lvert R(x,x_0)\rvert = \lvert f(x)-F(x)\rvert$
$f(2)=.6931$	$F(2)=1$	$\lvert f(2)-F(2)\rvert = .3069$
$f(1.5)=.4054$	$F(1.5)=.5$	$\lvert f(1.5)-F(1.5)\rvert = .0945$
$f(1.01)=.0099$	$F(1.01)=.01$	$\lvert f(1.01)-F(1.01)\rvert = .0001$
$f(1.001)=.000999$	$F(1.001)=.001$	$\lvert f(1.001)-F(1.001)\rvert = .000001$
$f(1)=0$	$F(1)=0$	$\lvert f(1)-F(1)\rvert = 0$

La situación geométrica de ambos casos se ilustra en la figura 1.3."

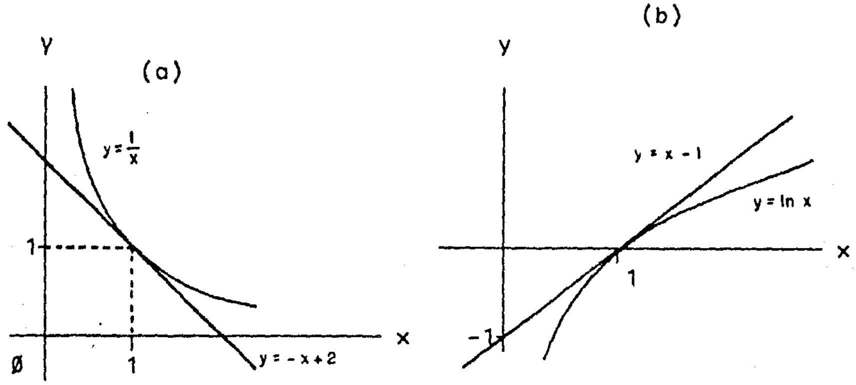


Figura 1.3

El énfasis en estos ejemplos debe ponerse en que si una función f es derivable en $x_\emptyset$, disponemos de un algoritmo para encontrar una función afín F (que es una función especialmente simple) que tiene la propiedad de que si estamos próximos al punto $x_\emptyset$, el error que cometemos al aproximar los valores de f con los de F es pequeño, más pequeño mientras más cerca estemos del punto $x_\emptyset$. De hecho podemos elegir x para que $R(x,x_\emptyset)$ sea tan pequeño como deseemos.

Es frecuente que en los cursos, al hablar de estas ideas, se propongan como aplicación de lo que se ha visto problemas del tipo: "encontrar una aproximación para $(65)^{1/2}$" que han de resolverse considerando que la función $f(x)=(x)^{1/2}$ es derivable es $x_\emptyset=64$, en donde además conocemos su valor $f(64)=8$; se encuentra la aproximación lineal F para esa f y una aproximación para f(65) estará dada por F(65). En este tipo de ejemplos se debe reforzar la parte conceptual: se dispone de un algoritmo para aproximar con una función afín a una función alrededor de un punto y no tanto en la materialidad del resultado obtenido; pues con los recursos con los que se dispone actualmente, estas aplicaciones numéricas pueden parecer innecesarias.

El profesor debe hacer notar que como conclusión del trabajo anterior se tiene que:

"Si una función f es derivable en x_0, es decir, existe el límite $\lim\limits_{h\to 0} \dfrac{f(x_0+h) - f(x_0)}{h} = f'(x_0)$, entonces se puede definir la recta tangente a la curva $y=f(x)$ en el punto $(x_0, f(x_0))$ y dicha recta es la gráfica de la función afín dada por $F(x)=f'(x_0)\ (x-x_0) + f(x_0)$ que nos da una buena aproximación para los valores de la función f alrededor (o cerca) del punto x_0. Bueno en el sentido de que la diferencia entre los valores de $f(x)$ y $F(x)$ para x's cercanas a x_0 se va a cero más rápido que $x-x_0$, i.e. si $R(x,x_0) = f(x) - F(x)$:

$$\lim_{x\to x_0} \frac{R(x,x_0)}{x-x_0} = \lim_{x\to x_0} \frac{f(x)-f(x_0)-f'(x_0)(x-x_0)}{x - x_0} = 0 \quad \ldots\ldots\ldots(*)"$$

Como en su momento nos interesará hablar de transformaciones lineales más que de funciones afines, podemos reescribir de una vez $(*)$ en términos de funciones lineales, haciendo las siguientes observaciones:

"Si hacemos $h=x-x_0$, tenemos que $(*)$ es equivalente a $\lim\limits_{h\to 0} \dfrac{f(x_0+h) - f(x_0) - f'(x_0)h}{h} = 0$. Con lo que, si definimos una función lineal $L:\mathbb{R} \to \mathbb{R}$ por medio de $L(h) = f'(x_0)\ h$ ($f'(x_0)$ es un número real), L es una función lineal, pues es del tipo $f(x)=mx$ con m constante. En otras palabras, hemos probado que si $f:\ D \subseteq \mathbb{R} \to \mathbb{R}$ es derivable en x_0, entonces existe una función lineal dada por $L(h) = f'(x_0)h$ tal que :

$$\lim_{h\to 0} \frac{f(x_0+h) - f(x_0) - L(h)}{h} = 0 \quad \ldots\ldots\ldots\ldots(**)"$$

Resulta muy importante destacar los papeles distintos que juegan cada una de las variables que intervienen en $(**)$: x_0 es un punto del dominio de la función f en el que pedimos que f sea derivable, una vez escogido, queda "fijo" para el resto de la exposición. Mientras que h es la variable que evoca movimiento, es la que propiamente "varía" (al grado que la estamos considerando como tendiendo a 0); denota al incremento a partir de x_0 (la antigua x es ahora $x = x_0 + h$). Para calcular L, h tiene que satisfacer $x_0 + h \in D$, el dominio de f (ver expresión (**)). No obstante, la función $L(h)=f'(x_0)h$ está definida $\forall\ h \in \mathbb{R}$.

Una vez hecha la distinción entre el carácter "estático" de la variable x_0 y el carácter "dinámico" de la variable h, se puede hacer ver más claramente cuál es la interpretación de (**):

"En estos términos (**) nos dice que si consideramos las funciones de h dadas por

$$\tilde{f}(h) = f(x_0+h) - f(x_0)$$

$$L(h) = f'(x_0)h$$

su diferencia $R(h) = \tilde{f}(h) - L(h) = (f(x_0+h)-f(x_0))-(f'(x_0)h)$

satisface que $\lim\limits_{h\to 0} \dfrac{R(h)}{h} = 0$. Es decir que cerca de $h=0$

L aproxima a $\tilde{f}$. Esta situación se ilustra en la figura 1.4."

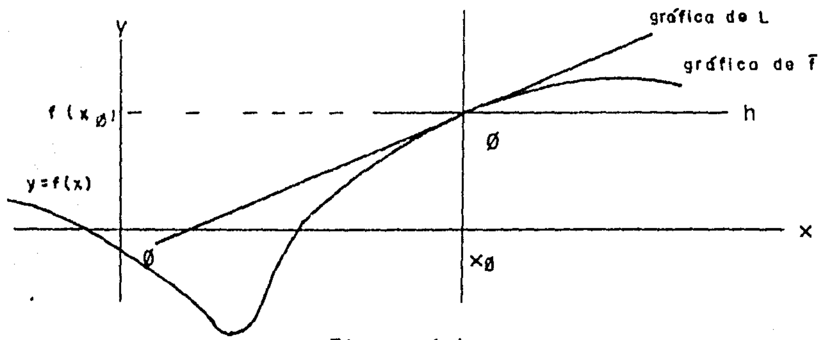


Figura 1.4

De esta forma, se han proporcionado todos los elementos para reescribir la conclusión en una notación muy conveniente, porque será la que se usará en dimensiones superiores:

"Si una función $f:D \subseteq \mathbb{R} \to \mathbb{R}$ es derivable en x_0, entonces existe una función lineal L (de $\mathbb{R}$ en $\mathbb{R}$) definida por $L(h) = f'(x_0)h$ que satisface:

si (i) $R(h) = f(x_0+h) - f(x_0) - L(h)$; entonces

(ii) $\lim_{h \to 0} \frac{R(h)}{h} = 0$ "

Se puede comentar que con frecuencia se unen (i) y (ii) en una sola condición, pidiendo que la función lineal L satisfaga:

$$\lim_{h \to 0} \frac{f(x_0+h) - f(x_0) - L(h)}{h} = 0$$

Es interesante contraponer el concepto de diferenciabilidad al de continuidad:

"Si una función $f: D \subseteq \mathbb{R} \to \mathbb{R}$ es derivable entonces es continua. Sin embargo sabemos que el recíproco no es cierto. Se tiene un ejemplo clásico en la función $f: \mathbb{R} \to \mathbb{R}$ dada por $f(x) = |x|$ que es continua en $\mathbb{R}$ y no es derivable en 0. Si bien los ejemplos más comunes son de funciones continuas que no son derivables en algún o en algunos puntos, se pueden dar ejemplos de funciones que no son derivables en un número infinito de puntos (figura 1.5)

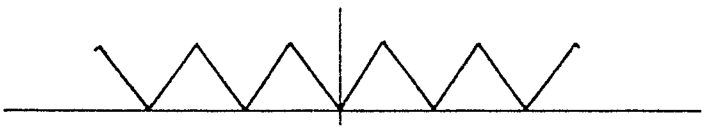


Figura 1.5

Incluso la situación puede ser más radical. Pues existen funciones que son continuas en todo $\mathbb{R}$ y en ningún punto de $\mathbb{R}$ son derivables."

Para un ejemplo de una función con las características anteriores se puede consultar: Spivak Michael, Calculus (volumen II), Reverté; Barcelona, 1981, páginas 624-627.

Sin embargo estudiar con detalle ese ejemplo supone el conocimiento de la convergencia uniforme de series infinitas, tema que los estudiantes desconocen a este nivel.

CAPITULO 2

FUNCIONES VECTORIALES DE VARIABLE REAL

Una vez estudiada la diferenciabilidad de funciones de $\mathbb{R}$ en $\mathbb{R}$, la pregunta es ¿con qué conviene seguir?, se sugiere emprender entonces el estudio de funciones de variable real cuyos valores están en $\mathbb{R}^n$. La razón de esta propuesta es que, al seguir siendo funciones de una variable son por tanto del tipo de funciones con las que ya se ha trabajado, además estas funciones las podemos pensar como colecciones de funciones reales de una variable y por tanto pueden reducirse al caso anterior.

Por otro lado, como sus contradominios son subconjuntos de $\mathbb{R}^n$; son una oportunidad de ir adquiriendo familiaridad con el ambiente natural del Cálculo vectorial. En síntesis, este tema constituye una buena etapa de transición para el estudio de las funciones vectoriales de varias variables.

Al trabajar con estas funciones se recomienda que, aunque los resultados se enuncien en su forma general, se haga una referencia constante a los casos $n = 2$, $n = 3$. Pues por una parte son los que en la práctica aparecen con más frecuencia. Y por otra, porque los elementos de $\mathbb{R}^n$ para esos casos, pueden interpretarse geométricamente como vectores en el sentido de esta palabra que los estudiantes conocen y manejan, por ejemplo, desde sus cursos de Física. Y es un principio útil en la enseñanza de las Matemáticas que el aprendizaje se facilita si es posible ver a un mismo objeto desde distintos puntos de vista. Además de que el recurso que proporcionan los dibujos, que para estas dimensiones es posible hacer, es también de mucha ayuda.

Se supone que los estudiantes poseen un conocimiento previo de la estructura de $\mathbb{R}^n$ así como del tema de vectores.

Una manera tradicional y al parecer eficiente de introducir a las funciones vectoriales de manera natural es considerar el movimiento de una partícula en el plano o en el espacio. Un prototipo de esta forma de razonamiento sería el siguiente:

"Si pensamos en un punto que se mueve en el plano a lo largo de una curva, para poder describir su posición necesitamos dos coordenadas: x,y. (Ver figura 2.1)

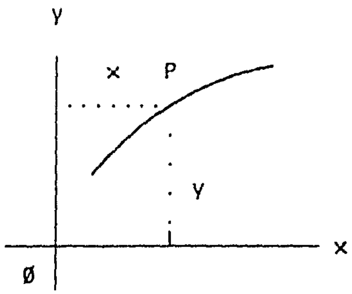


Figura 2.1

Más aún x, y varían con el tiempo t; por lo tanto necesitamos dos funciones: una que relacione x y t, a saber que para cada tiempo t nos dé la abscisa del punto P al tiempo t; y otra que relacione y con t, que nos dé la ordenada del punto P al tiempo t, i.e.:

$x: \mathbb{R} \to \mathbb{R} \quad x = x(t)$

$y: \mathbb{R} \to \mathbb{R} \quad y = y(t)$

Por tanto, podemos formar una función $f: \mathbb{R} \to \mathbb{R}^2$ que para cada tiempo t nos dé: $f(t) = (x(t), y(t)) \in \mathbb{R}^2$ que es la posición en el plano del punto P. (Figura 2.2)

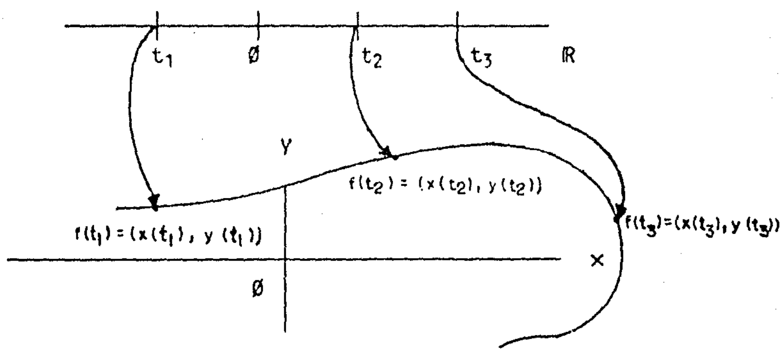


Figura 2.2

Ahora bien, podemos generalizar esta idea pensando en funciones que, en general a cada número real (sin necesidad de interpretarlo como tiempo) de un cierto subconjunto $D \subseteq \mathbb{R}$ le asocien un vector de $\mathbb{R}^n$. A una función de este tipo $f: D \subseteq \mathbb{R} \to \mathbb{R}^n$ se le llama una función vectorial de variable real.

Como para cada $x \in D$, $f(x) \in \mathbb{R}^n \Rightarrow$ $f(x) = (y_1, y_2, \ldots, y_n)$. Cada uno de los números y_i que constituyen las componentes del vector $f(x)$ depende del número x, por tanto, cada una de las componentes y_i define una función real valuada cuyo dominio es D.

$$y_i = f_i(x) \qquad f_i: D \subseteq \mathbb{R} \to \mathbb{R}$$

con lo que $f: D \xrightarrow{f} \mathbb{R}^n$ está definido por:

$$f(x) = (f_1(x), f_2(x), \ldots, f_n(x))$$

Las funciones f_i se llaman las funciones coordenadas de f."

Un comentario casi obligado de estas funciones es los recursos con los que contamos para visualizarlas, si bien, al ser funciones cabría preguntarse por su gráfica, la manera más adecuada de darse una idea geométrica acerca de ellas es por medio de su imagen, que corresponde generalmente, a la idea intuitiva de curva que es familiar para los estudiantes. Sobre esto se podría mencionar que:

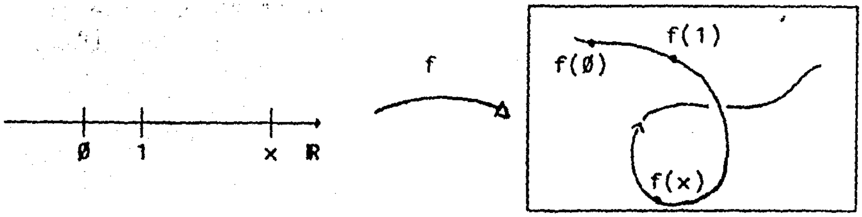
"La imagen de una función $f: D \subseteq \mathbb{R} \to \mathbb{R}^n$ es el subconjunto de $\mathbb{R}^n$ dado por $\{ f(x) \in \mathbb{R}^n \mid x \in D \}$. Como $f(x) \in \mathbb{R}^n$ $\Rightarrow f(x) = (y_1, y_2, \ldots, y_n)$

$$(*) \begin{cases} y_1 = f_1(x) \\ y_2 = f_2(x) & x \in D \\ \ldots\ldots\ldots\ldots \\ y_n = f_n(x) \end{cases}$$

y para $i = 1, 2, \ldots, n$ $\qquad f_i: D \subseteq \mathbb{R} \to \mathbb{R}$

A las expresiones (*) se les llama las ecuaciones paramétricas de la imagen de f con parámetro x. Aunque se les llama ecuaciones son propiamente funciones reales de una variable.

Podemos pensar que, generalmente, una función vectorial lo que hace es "meter" una copia de la recta real (como si fuera un alambre muy dúctil) en el espacio $\mathbb{R}^n$ deformándolo, (como "enrollándolo" y torciéndolo) según el patrón prescrito por las ecuaciones paramétricas. (Ver Figura 2.3)"



$\mathbb{R}^n$

Figura 2.3

Cabe mencionar que hay funciones vectoriales cuyo comportamiento no corresponde a la imagen intuitiva referida anteriormente. En el caso de las curvas de Peano, que son curvas que cubren una superficie. Un ejemplo específico es el siguiente:

Sea $f: [0, 1] \to [0, 1] \times [0, 1] \subseteq \mathbb{R}^2$ definida de la siguiente forma:

Cada $t_0 \in [0,1]$ puede expresarse de manera única como un decimal infinito:

$$t_0 = .a_1a_2a_3 \ldots = \frac{a_1}{10} + \frac{a_2}{10^2} + \frac{a_3}{10^3} + \ldots = \sum_{i=1}^{\infty} \frac{a_i}{10^i} \quad \text{con } 0 \leq a_i \leq 9$$

(la unicidad se obtiene eliminando las expansiones que terminan con puros 9's, i.e.

$$\frac{1}{10} + \frac{1}{10^2} + \frac{9}{10^3} + \frac{9}{10^4} + \ldots\ldots = \frac{1}{10} + \frac{2}{10^2} \quad)$$

De esta forma f se define como:

$$.a_1a_2a_3 \ldots\ldots a_n \ldots \xrightarrow{f} (.a_1a_3a_5 \ldots\ldots ,\ .a_2a_4a_6 \ldots\ldots) \text{ i.e.}$$

$$\sum_{i=1}^{\infty} \frac{a_i}{10^i} \xrightarrow{f} \left(\sum_{i=0}^{\infty} \frac{a_{2i+1}}{10^{i+1}} ,\ \sum_{i=1}^{\infty} \frac{a_{2i}}{10^i} \right)$$

f es suprayectiva pues dado $p = (.a_1a_2a_3\ldots\ldots, .b_1b_2b_3\ldots\ldots) =$

$$= \left(\sum_{i=1}^{\infty} \frac{a_i}{10^i} ,\ \sum_{i=1}^{\infty} \frac{b_i}{10^i} \right) \in [0,1] \times [0,1]$$

escogemos $t_0 = .a_1b_1a_2b_2a_3b_3\ldots\ldots \in [0,1]$ y $f(t_0) = p$

Para probar que f es continua en t_0, $\forall t_0 \in [0,1]$, debemos probar que dada $\varepsilon > 0$, existe $\delta > 0$ tal que si $|t| < \delta$ entonces $\| f(t_0 + t) - f(t_0) \| < \varepsilon$. Sea $t_0 \in [0,1]$ entonces $t_0 = .a_1a_2a_3a_4....$ Escogemos $N \in \mathbb{N}$ tal que

$$2/10^N < \varepsilon \quad \text{y sea} \quad \delta < 1/10^{2N}$$

$f(t_0) = (.a_1a_3a_5... a_{2N-1}\, a_{2N+1}...,\ .a_2a_4a_6... a_{2N}\, a_{2N+2}...)$

Si $|t| < \delta$ entonces

$t_0 + t = .a_1a_2a_3... a_N... a_{2N}\, \alpha_{2N+1}\, \alpha_{2N+2}\, \alpha_{2N+3}\, ...$

$$\Rightarrow f(t_0+t) = (.a_1a_3a_5...a_{2N-1}\, 0\, 0\, 0...,.a_2a_4a_6...a_{2N}\, 0\, 0\, 0...) + (.0\, 0\, 0\, ...0\, \alpha_{2N+1}\, \alpha_{2N+3}...,.0\, 0\, 0...0\, \alpha_{2N+2}\, \alpha_{2N+4}...)$$

$$\therefore \| f(t_0+t) - f(t_0) \| = \| (.00...0\, b_{N+1}\, b_{N+3}\, b_{N+5}....,.00...0\, b_{N+2}\, b_{N+4}\, b_{N+6}...) \|$$

donde $b_{N+j} = \alpha_{2N+j} - a_{2N+j}$

Si $\bar{\beta} = (.b_{N+1}\, b_{N+3}\, b_{N+5}...,\ .b_{N+2}\, b_{N+4}\, b_{N+6}...)$

$\| \bar{\beta} \| \leq 2$

$$\text{y } \| f(t_0+t)-f(t_0) \| = \| \bar{\beta} \| \frac{1}{10^{N+1}} < \frac{2}{10^N} < \varepsilon \text{ l.c.q.d.}$$

i.e. f es continua y suprayectiva. Sin embargo si las funciones vectoriales son diferenciables, un comportamiento de esa naturaleza queda excluido, y el comportamiento de la función sí se adecúa a la imagen que se propuso.

Habitualmente al hablar de las funciones vectoriales, los textos ponen el énfasis en la parte operativa y de aplicaciones, en términos de las observaciones que se hicieron al principio del capítulo, parece un tratamiento adecuado.

Sólo se harán consideraciones acerca de algunas cuantas cuestiones sobre el tema, de carácter más bien teórico. Cuestiones que van directamente encaminadas a preparar el estudio de la diferenciabilidad de las funciones vectoriales.

Sobre la definición de límite para estas funciones, es frecuente que en la bibliografía se den formulaciones

como la siguiente:
"Si $r: D \subseteq \mathbb{R} \to \mathbb{R}^3$ $r(t) = (f_1(t), f_2(t), f_3(t))$

entonces $\lim_{t\to a} r(t) = (\lim_{t\to a} f_1(t), \lim_{t\to a} f_2(t), \lim_{t\to a} f_3(t))$ "

Esta formulación tiene la ventaja de ser operativamente muy útil, además reduce inmediatamente el problema de límites a límites de funciones reales, donde se cuenta con experiencia y múltiples recursos para calcular límites. Sin embargo, se propone como una mejor opción, dar primeramente la definición de límite en términos de epsilons y deltas, que una vez interpretada geométricamente permite entender mejor lo que está pasando. Además, el enfrentarse con la formulación de $\varepsilon - \delta$ desde ahora, facilita que el estudiante adquiera familiaridad con esos argumentos lo cual resulta de mucha utilidad, pues si bien para las funciones vectoriales de variable real se pudo evitar esa formulación, al llegar a las funciones de varias variables, se hace ineludible recurrir a ella y será mejor para el estudiante si ya ha podido ejercitarse. Además esta formulación tiene la ventaja conceptual de tratar a las funciones vectoriales como "unidades", como entidades propias y no sólo a través de sus funciones coordenadas.

Como segundo paso, para tener también las ventajas operativas de la formulación en términos de funciones coordenadas, se puede probar la equivalencia de ambas formas.

Desde esta perspectiva tendríamos entonces que:
"Si $f: D \subseteq \mathbb{R} \to \mathbb{R}^n$, decimos que el límite de f conforme t se aproxima (o tiende) a $a (t \in D)$ es b (b vector de $\mathbb{R}^n$) denotado esto por $\lim_{t\to a} f(t) = b$;
si dado $\varepsilon > 0$ es posible encontrar $\delta > 0$ tal que
$\| f(t) - b \| < \varepsilon$ siempre que $0 < |t - a| < \delta$

Intuitivamente podemos pensar en una situación como la que se sugiere en la figura 2.4, donde si $\overrightarrow{OP}$ y $\overrightarrow{OB}$ son los vectores de posición correspondientes a f(t) y b respectivamente, entonces conforme t esté más cerca de a, el vector $\overrightarrow{OP}$ se aproxima al vector $\overrightarrow{OB}$ en el sentido de que P se acerca cada vez más a B.

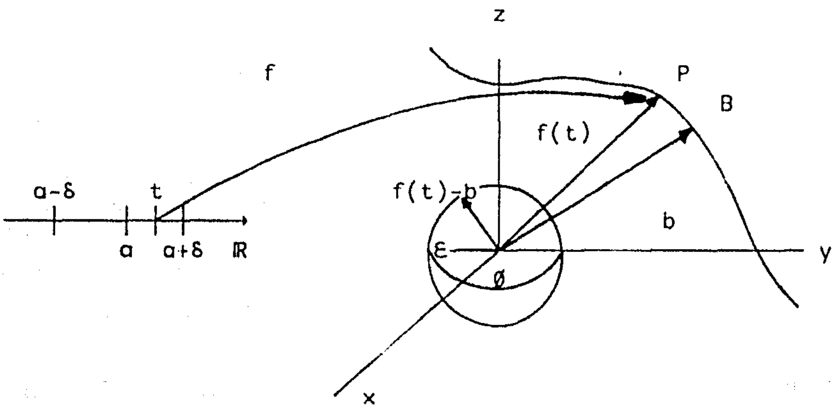


Figura 2.4

Tenemos el siguiente :

Teorema. La función $f: D \subseteq \mathbb{R} \to \mathbb{R}^n$ definida por $f(t) = (f_1(t), f_2(t), \ldots, f_n(t))$ tiene límite conforme $t \to a$ si y sólo si f_i tiene límite conforme $t \to a$; $f_i: D \subseteq \mathbb{R} \to \mathbb{R}$ para $i = 1, 2, \ldots, n$. En cuyo caso:

$$\lim_{t\to a} f(t) = \left(\lim_{t\to a} f_1(t), \lim_{t\to a} f_2(t), \ldots, \lim_{t\to a} f_n(t) \right)$$

Demostración.

Si f tiene límite conforme $t\to a \Rightarrow \exists\, b = (b_1, b_2, \ldots, b_n) \in \mathbb{R}^n$ tal que $\lim_{t\to a} f(t) = b$ i.e. Dado $\varepsilon > \emptyset$, $\exists\, \delta > \emptyset$ tal que

$$\| f(t) - b \| < \varepsilon \quad \text{siempre que } \emptyset < | t - a | < \delta$$

$$\|f(t)-b\| = ((f_1(t)-b_1)^2 + (f_2(t)-b_2)^2 + \ldots + (f_n(t)-b_n)^2)^{1/2} < \varepsilon$$

$\Rightarrow$ para cualquier $i \in \{1, 2, \ldots n\}$

$$|f_i(t) - b_i| \leq \|f(t) - b\| < \varepsilon \quad \text{si } \emptyset < |t - a| < \delta$$

$\Rightarrow$ para $i = 1, 2, \ldots n$ $\quad \lim_{t\to a} f_i(t) = b_i$

Inversamente si $\lim_{t\to a} f_i(t)$ existe para $i = 1, 2, \ldots n$, y denotamos dicho límite por b_i, i.e. $\lim_{t\to a} f_i(t) = b_i$ para $i = 1, 2, \ldots n$. Por definición de límite para el caso real

tenemos que dado $\varepsilon/(n)^{1/2}$, $\exists\ \delta_i$ para cada $i \in \{1, 2, \dots n\}$ tal que
$|f_i(t) - b_i| < \varepsilon/(n)^{1/2}$ siempre que $\emptyset < |t - a| < \delta_i$.

Sea $\delta = \min \{\delta_1, \delta_2, \dots \delta_n\}$, $\delta > \emptyset$

Si $\emptyset < |t - a| < \delta$ se tiene que para cada i, $i \in \{1, 2, \dots n\}$
$|f_i(t) - b_i| < \varepsilon/(n)^{1/2} \Rightarrow |f_i(t) - b_i|^2 < \varepsilon^2/n$.

Sumando las expresiones correspondientes para $i = 1, 2, \dots n$ tenemos que si $\emptyset < |t - a| < \delta$
$|f_1(t)-b_1|^2 + |f_2(t)-b_2|^2 + \dots + |f_n(t)-b_n|^2 < n(\varepsilon^2/n) = \varepsilon^2$

$\Rightarrow ((f_1(t) - b_1)^2 + \dots\dots + (f_n(t) - b_n)^2)^{1/2} < \varepsilon$

$\Rightarrow \|f(t) - b\| < \varepsilon$ donde $b = (b_1, b_2, \dots b_n) =$
$= (\lim_{t\to a} f_1(t), \lim_{t\to a} f_2(t), \dots\dots, \lim_{t\to a} f_n(t))$
i.e. $\lim_{t\to a} f(t) = b$. "

Se puede señalar a los estudiantes que este teorema justifica que en ocasiones se defina el límite de una función vectorial conforme t tiende a a ($t \in D$) como el vector cuyas coordenadas son respectivamente los límites de las funciones coordenadas conforme $t \to a$, en el caso de que dichos límites existan.

Habiendo trabajado con detalle la definición de límite; la definición de continuidad no ofrece ningún problema. Se puede presentar en términos de ε - δ repitiendo la exposición anterior simplemente sustituyendo al vector b por f(a) (el mismo cambio hace que la interpretación geométrica dada para el límite sirva para la continuidad). De manera más escueta se puede definir la continuidad de una función vectorial en a si se satisface que $\lim_{t\to a} f(t) = f(a)$.
Al igual que en el caso real hay que hacer notar que la expresión anterior supone que, de hecho, se están dando tres condiciones:

(i) f está definida en a

(ii) $\lim_{t \to a} f(t)$ existe

(iii) $\lim_{t \to a} f(t) = f(a)$

El mismo teorema citado anteriormente (he aquí otra razón para presentarlo) permite concluir que si $f(t) = (f_1(t), f_2(t), \ldots, f_n(t))$, entonces f es continua en a $\Leftrightarrow$ $f_1, \ldots, f_n$ son continuas en a, lo que nos proporciona una manera sencilla de estudiar la continuidad de una función vectorial de variable real: estudiando la continuidad de las funciones coordenadas, que al ser funciones reales de una variable, los estudiantes pueden en muchos casos por simple inspección saber dónde son continuas.

Para demostrar que la continuidad de una función vectorial es equivalente a la continuidad de sus funciones coordenadas, se puede calcar la demostración del teorema anterior sustituyendo b_i por $f_i(a)$ para $i = 1, 2, \ldots, n$.

Se ha de mencionar que la continuidad de f en un subconjunto de $\mathbb{R}$ se define de la manera usual:
"$f: D \subseteq \mathbb{R} \to \mathbb{R}^n$ es continua; si es continua en t, $\forall\, t \in D$."

Acerca de la definición de diferenciabilidad:

Nuestra idea es, en breve, hablar de diferenciabilidad de una función vectorial en un punto como la existencia de una transformación lineal que aproxima localmente a la función. Sin embargo aún no ha sido necesario hablar de transformaciones lineales y en este momento no hay suficiente motivación para hacerlo, pues para estas funciones, al igual que en el caso anterior, la existencia de dicha aproximación lineal es equivalente a la existencia del familiar cociente diferencial, es decir

$$\lim_{h \to 0} \frac{f(t_0 + h) - f(t_0)}{h},$$

donde hemos convenido en que

$$\frac{f(t_0+h) - f(t_0)}{h} = \frac{1}{h}\,[f(t_0+h) - f(t_0)] \;;$$

de esa forma el concepto de diferenciabilidad es formalmente el mismo que teníamos para funciones reales de una variable. Por ello se puede proponer, a este nivel, como definición de diferenciabilidad, la derivabilidad de la función:

"$f: D \subseteq \mathbb{R} \to \mathbb{R}^n$ es derivable en t_0 ($t_0 \in D$) si

$$\lim_{h \to 0} \frac{1}{h} [f(t_0+h) - f(t_0)]$$

existe en $\mathbb{R}^n$, en cuyo caso lo denotamos por $f'(t_0)$ y el vector $f'(t_0)$ se llama la derivada de f en t_0. (Se consideran aquellas h's para las cuales $t_0+h \in D$ y $h \neq 0$).

Si convenimos en que $\dfrac{f(t_0+h) - f(t_0)}{h} = \dfrac{1}{h} [f(t_0+h) - f(t_0)]$, la definición de derivada para este caso es formalmente la misma que para el caso real:

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

aunque no hay que olvidar que para el caso vectorial, la expresión anterior es una igualdad entre vectores de $\mathbb{R}^n$."

El teorema citado en el capítulo nos permite nuevamente dar la "versión operativa" de la definición de diferenciabilidad para funciones vectoriales, puesto que como consecuencia del mismo obtenemos que una condición necesaria y suficiente para la derivabilidad de una función en un punto, es la derivabilidad de las funciones coordenadas en el punto. Con lo cual, para calcular la derivada de una función vectorial en un punto, basta con formar el vector cuyas coordenadas son las derivadas de las funciones coordenadas en el punto ya que: $\dfrac{1}{h} [f(t_0+h) - f(t_0)] =$

$$= \frac{1}{h} [(f_1(t_0+h)-f_1(t_0), f_2(t_0+h)-f_2(t_0), \ldots, f_n(t_0+h)-f_n(t_0))]$$

$$= \left(\frac{f_1(t_0+h)-f_1(t_0)}{h}, \frac{f_2(t_0+h)-f_2(t_0)}{h}, \ldots, \frac{f_n(t_0+h)-f_n(t_0)}{h} \right). (*)$$

Con lo que si tomamos el límite cuando $h \to 0$ en ambos miembros de (*) el límite del lado izquierdo que es $f'(t_0)$, existirá si y sólo si cada uno de los límites

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

existe, para $i = 1, 2, 3, \ldots n$.
En cuyo caso tendríamos que:

$$\lim_{h\to 0}\frac{f(t_0+h)-f(t_0)}{h} = \left(\lim_{h\to 0}\frac{f_1(t_0+h)-f_1(t_0)}{h}, \ldots, \lim_{h\to 0}\frac{f_n(t_0+h)-f_n(t_0)}{h}\right)$$

i.e. $f'(t_0) = (f_1'(t_0), f_2'(t_0), \ldots\ldots, f_n'(t_0))$.

Que es la definición de la derivada de la función vectorial f en el punto t_0 que dan varios textos; pero que se propone presentar más bien como una forma práctica de calcular la derivada.

Un comentario que cabría hacer sobre la notación, motivado por la confusión que a veces se da en el caso real, es que se ha de precisar su sentido a fin de distinguir cuándo se habla de la derivada de una función vectorial en un punto, que es un vector de $\mathbb{R}^n$, y cuándo de la función derivada denotada por f', que es una función vectorial que a cada t le asocia el vector $f'(t)$ en caso de que éste exista. En términos de las funciones coordenadas es la función definida por:

$$f'(t) = (f_1'(t), f_2'(t), \ldots\ldots, f_n'(t))$$

para aquellos t's para los cuales $\lim_{h\to 0} \frac{f(t+h)-f(t)}{h}$ existe.

Como nos será de utilidad al hablar de la diferenciabilidad de una función vectorial en términos de aproximaciones lineales, es bueno que aquí se mencione la interpretación de la derivada de una función vectorial como un vector tangente. Se puede hacer con planteamientos como:

"La definición de la derivada de una función vectorial admite una interpretación geométrica interesante en $\mathbb{R}^2$ y $\mathbb{R}^3$. Por ejemplo sea $f(t) = (f_1(t), f_2(t), f_3(t))$ y f_i para $i=1,2,3$ es una función derivable. Denotamos por C la curva con ecuaciones paramétricas

$$x = f_1(t)$$
$$y = f_2(t)$$
$$z = f_3(t)$$

como ilustrado en (i) de figura 2.5.

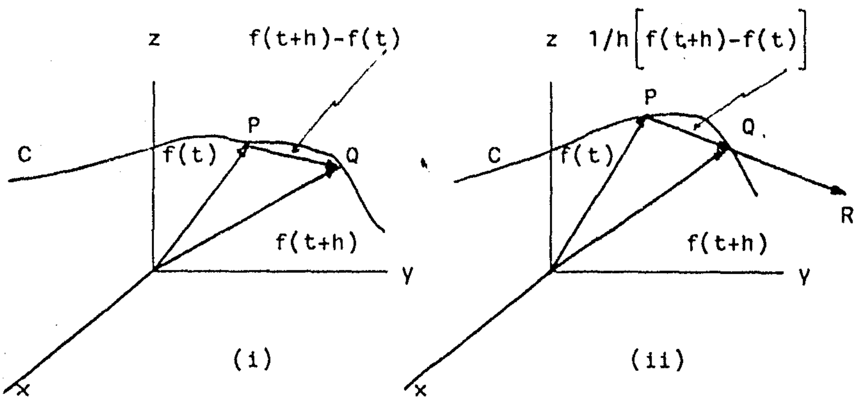


Figura 2.5

Si $\overrightarrow{OP}$ y $\overrightarrow{OQ}$ son los vectores de posición correspondientes a $f(t)$ y $f(t+h)$ respectivamente, entonces $\overrightarrow{PQ} = \overrightarrow{OQ}-\overrightarrow{OP}$, es una representación geométrica del vector $f(t+h) - f(t)$. Se sigue que $1/h\ \overrightarrow{PQ}$ es una representación geométrica de $1/h\ (f(t+h) - f(t))$. Si $h>0$ entonces $1/h\ \overrightarrow{PQ}$ es un vector $\overrightarrow{PR}$ que tiene la misma dirección que $\overrightarrow{PQ}$. Más aún si $0<h<1$ entonces $1/h > 1$ y $|\overrightarrow{PR}| > |\overrightarrow{PQ}|$, como se ilustra en (ii) de la Figura 2.5. Si $h \to 0$ entonces Q se acerca a P y el vector $\overrightarrow{PR}$ se aproxima a lo que intuitivamente llamamos el vector tangente a C en P. La recta tangente se define como el conjunto $\{f(t)+\lambda f'(t)/\lambda\in\mathbb{R}\} \subseteq \mathbb{R}^n$, es pues la recta en $\mathbb{R}^n$ que pasa por $f(t)$ en la dirección del vector $f'(t)$".

Un comentario que contribuye a sacar mayor provecho a esta interpretación geométrica es el mencionar que el vector $f'(t)$ apunta en la dirección en la que la curva está siendo "trazada" por $f(t)$ conforme t aumenta, i.e. $f'(t)$ señala la dirección en la que la curva está siendo recorrida (conforme t crece)(Figura 2.6)

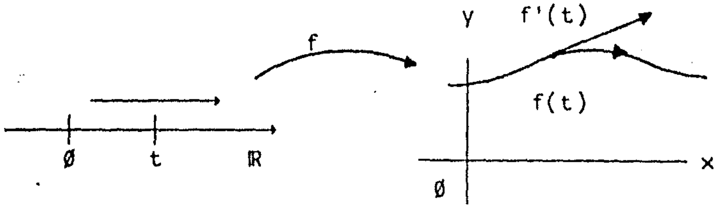


Figura 2.6

Finalmente, interesa dejar claro que la derivabilidad para funciones vectoriales implica continuidad (al igual que en el caso real). Esto sirve para, posteriormente, darse cuenta del contraste al llegar a funciones de varias variables, en donde la derivabilidad entendida como existencia de derivadas parciales o direccionales, no implica la continuidad de la función en el punto. Eso contribuirá, a su vez a motivar el estudio de la diferenciabilidad para esas funciones en términos de aproximaciones lineales. Conviene entonces presentar el siguiente resultado:

"Si $f:D\subseteq\mathbb{R}\to\mathbb{R}^n$ es derivable en $t_0 \in D$, entonces f es continua en t_0.

Necesitamos probar que $\lim\limits_{h\to 0} f(t_0+h)=f(t_0)$.

Consideremos $\lim\limits_{h\to 0} (f(t_0+h) - f(t_0)) =$

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

$$f'(t_0) \lim_{h\to 0} h = 0.$$

$$\therefore \lim_{h\to 0} f(t_0+h) = f(t_0) \Rightarrow f \text{ es continua en } t_0.$$

Obsérvese que, como un caso particular, se tiene el resultado anterior para las funciones reales de una variable."

CAPITULO 3

FUNCIONES REALES DE VARIAS VARIABLES

3.Ø INTRODUCCION

Nuestro objetivo final es estudiar la diferenciabilidad de las funciones de $\mathbb{R}^n$ a $\mathbb{R}^m$. Estas funciones o transformaciones se descomponen en funciones coordenadas de la manera usual: $f: D \subseteq \mathbb{R}^n \to \mathbb{R}^m$

$$f(x) = (f_1(x), f_2(x), \ldots\ldots, f_m(x))$$

para $i = 1, 2, \ldots .m \quad f_i: D \subseteq \mathbb{R}^n \to \mathbb{R}$

Dado que operativamente resulta muy útil trabajar con las funciones coordenadas de una transformación, tiene especial interés estudiar las funciones reales de varias variables.

Los comentarios y observaciones que sobre las funciones reales de varias variables se hacen en el presente capítulo, se centran en aquellos temas directamente relacionados con la idea de diferenciabilidad para dichas funciones. El esquema que se sigue es: posibles modos de introducir a las funciones de varias variables, los recursos que se utilizan para visualizarlas, las definiciones de límite y continuidad para dichas funciones; para finalmente abordar la noción de diferenciabilidad, tema que presenta sus dificultades desde el punto de vista didáctico. Pues para las funciones de una variable, la diferenciabilidad suponía simplemente la existencia de un límite, para las funciones de varias variables en cambio, la diferencibilidad supone la existencia de una cierta transformación lineal.

Este paso conceptual ha de hacerse gradualmente. Es por ello que para estudiar la diferencibilidad de las funciones de varias variables se propone hacerlo en etapas sucesivas: intentando generalizar la idea de derivada que se tiene para el caso anterior se hablaría primeramente de derivadas parciales, luego de derivadas direccionales, ya que en su definición y en su manipulación sólo se necesita del concepto de derivada que ya se posee.

Posteriormente se propone hacer ver que estas dos extensiones de la idea de derivada, no son suficientemente adecuados en el sentido de que por sí solas no aseguran que se preserven dos propiedades que nos interesaría que la definición de diferenciabilidad para funciones de varias variables lo hiciera - como sucedía en el caso de una variable - : la continuidad y consecuentemente la existencia de espacios tangentes.

3.1 FUNCIONES REALES DE VARIAS VARIABLES

Nos encontramos propiamente ante el primer tema que resulta totalmente novedoso para los estudiantes, pues si bien en el capítulo anterior se habló de funciones cuyos valores eran vectores; éstas no constituían esencialmente objetos nuevos, ya que de ordinario en los primeros cursos de Cálculo se estudia el tema de las ecuaciones paramétricas al menos para representar curvas en el plano y una adecuada comprensión de ese caso proporciona los elementos necesarios para entender lo que sucede en el caso general. Además, en el estudio de dichas funciones aún estamos dentro de un mismo ámbito, de un mismo nivel conceptual: el cálculo de funciones de una sola variable (real).

Sin embargo con el estudio de las funciones real valuadas de varias variables nos aventuramos en un tema en que los estudiantes no tienen ninguna experiencia previa y por otra parte, supone dar un paso conceptual: pasar de una a más variables.

Es recomendable que a lo largo de la exposición de este tema; sin dejar de comentar que los resultados y afirmaciones pueden ser generalizados, el profesor se refiera continuamente a los casos de dos y tres variables. La sugerencia anterior obedece a varias razones: el hombre en el proceso del conocer, va de lo particular a lo general, de lo concreto a lo abstracto; es por ello que si de primera instancia los resultados se presentan de manera abstracta y para casos generales, el estudiante seguramente tendrá más dificultades para entender el material que si, por el contrario, va teniendo una cierta experiencia de lo que sucede en casos particulares y sencillos como en el plano y en el espacio. Por otra parte, estos casos son susceptibles de ser visualizados geométricamente y esto, sobre todo estando en una etapa nueva de aprendizaje en Matemáticas, proporciona una gran ayuda. Finalmente, porque en la vida profesional de muchos de los estudiantes que cursan el Cálculo vectorial, estos casos serán suficientes para cubrir las necesidades reales que se presentan.

Por ser un tema nuevo, es interesante motivar su estudio. Una manera de lograrlo es hacer ver a los estudiantes que situaciones en Matemáticas con las que ya se ha experimentado previamente pueden interpretarse desde esta perspectiva: relaciones cuantitativas que se han manejado, al traducirse al lenguaje formal de las Matemáticas resultan ser funciones de este tipo. Por ejemplo el volumen V de un cilindro está dado en términos de su radio r y de su altura h por la fórmula $V = \pi r^2 h$. En este hecho está "latente" una función de dos variables, r y h definida por $f(r,h) = \pi r^2 h$, que estará definida para $r > \emptyset$ $h > \emptyset$, formalmente $f: D \subseteq \mathbb{R}^2 \to \mathbb{R}$ donde $D = \{(r,h) \in \mathbb{R}^2 / r > \emptyset, h > \emptyset\}$. Análogamente el volumen de un sólido rectangular cuyas dimensiones son a, b, c es función de tres variables (pues $V = abc$) lo que se denotaría por $f: D \subseteq \mathbb{R}^3 \to \mathbb{R}$ con
$D = \{(x,y,z) \in \mathbb{R}^3 / x > \emptyset, y > \emptyset, z > \emptyset\}$ y $f(a,b,c) = abc$.

También se puede hacer ver que dichas funciones aparecen de manera natural en muchos contextos al ser la manera más conveniente de describir matemáticamente algún problema. Por ejemplo:

"En el estudio de las condiciones climáticas, la temperatura se mide en varios puntos del espacio (de la Tierra) y en diferentes tiempos. Escogiendo marcos de referencia para el tiempo y el espacio; una manera adecuada de describir la temperatura es mediante una "función de temperatura", una función $\theta: \mathbb{R}^4 \to \mathbb{R}$ tal que $\theta(x,y,z,t)$ sea el valor de la temperatura en grados centígrados en el punto del espacio con vector de posición (x,y,z) para un tiempo de "t" unidades."

Cabe hacer la siguiente observación sobre la terminología.

"Dada una función de este tipo $f: A \subseteq \mathbb{R}^n \to \mathbb{R}$, i.e. $\forall\ a \in A$, $f(a) \in \mathbb{R}$ (A es el dominio de la función y es un cierto subconjunto de $\mathbb{R}^n$). Como cada punto p de $\mathbb{R}^n$ es una n-ada de números reales $p = (x_1, x_2, \ldots x_n)$ podemos escribir $f(p) = f(x_1, \ldots, x_n)$ con $x_i \in \mathbb{R}$ $i = 1, 2, \ldots n$. Esto hace que a dichas funciones se les llame indistintamente funciones reales de variable vectorial o funciones reales de varias variables lo que enfatiza el hecho de que consideramos a las coordenadas de un punto $p \in \mathbb{R}^n$
$p = (x_1, x_2, \ldots, x_n)$ como varias variables reales y $f(p) = f(x_1, \ldots, x_n) \in \mathbb{R}$ depende de esas variables."

Sobre los recursos para visualizar geométricamente a estas funciones puede mencionarse que si bien para representar pictóricamente a las funciones $f: \mathbb{R} \to \mathbb{R}^n$ ($n = 2$ ó 3) dibujamos sus imágenes - como subconjuntos de $\mathbb{R}^n$ -, para estas funciones no es útil ese recurso pues como sus imágenes están contenidas en $\mathbb{R}$, los detalles finos sobre su comportamiento se pierden; es por eso que, - como era lo habitual para el caso real - solemos representar gráficamente a dichas funciones (para el caso $n = 2$) por medio de su gráfica:

"La gráfica de $f: D \subseteq \mathbb{R}^n \to \mathbb{R}$ es el subconjunto de $\mathbb{R}^{n+1}$ dado por $\{(x_1, x_2, \ldots, x_n, x_{n+1}) \in \mathbb{R}^{n+1} / x_{n+1} = f(x_1, \ldots, x_n)\}$ $= \{(p, f(p)) \in \mathbb{R}^{n+1} / p \in D\}$.

A la expresión $x_{n+1} = f(x_1, x_2, \ldots, x_n)$ ($z = f(x, y)$ para el caso de 2 variables) se le llama la ecuación de la gráfica. Para el caso de dos variables por lo común la gráfica adopta la forma de una superficie en el espacio, siendo $f(x, y)$ la distancia orientada desde $(x, y, 0)$ hasta $(x, y, f(x, y))$ (ver figura 3.1).

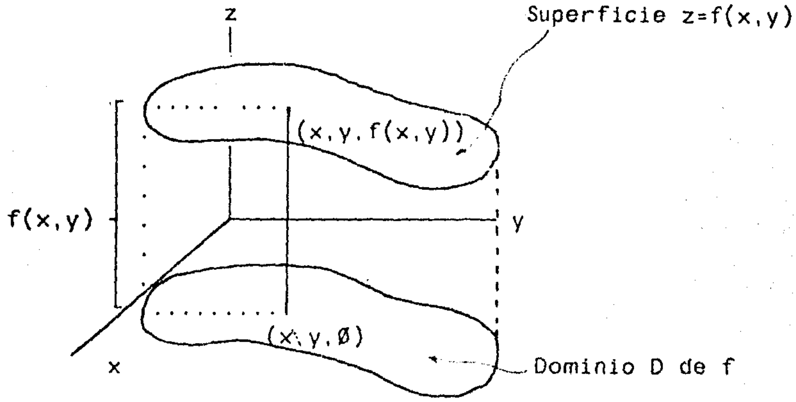


Figura 3.1

Para n=3 la gráfica está en $\mathbb{R}^4$ y ya no la podemos dibujar, sí se pueden dibujar, en cambio, los conjuntos de nivel :

$f:D\subseteq\mathbb{R}^n\to\mathbb{R}$ y $c \in \mathbb{R}$ el conjunto de puntos en D para los cuales la función toma el valor c, se llama el conjunto de nivel correspondiente a c, i.e. es el conjunto $\{p\in D/f(p)=c\}$.

Estos conjuntos de nivel también pueden dibujarse para el caso n=2 resultando generalmente curvas en el plano. Los conjuntos de nivel proporcionan un recurso útil también para darse una idea geométrica del comportamiento de la función."

Es conveniente que se trabajaran con detalle algunos ejemplos sobre cómo construir las gráficas de funciones de dos variables, determinando dominios e imágenes de las funciones correspondientes, puesto que esto ayuda a los estudiantes a adquirir práctica en visualizar objetos en $\mathbb{R}^3$ y si alcanzan a "ver" y "dibujar" lo que pasa ahí, podrán luego "imaginar" o "intuir" lo que sucede en más dimensiones.

Después de haberse familiarizado con el concepto y las maneras de visualizar a estas funciones se suele emprender el estudio del análisis de las mismas: límites, continuidad, etc..

En el caso de las funciones de una variable, incluso cuando no se señalaba de manera explícita, los estudiantes inferían que los dominios eran o toda la recta o intervalos o uniones (de un número en general pequeño) de intervalos. Dicha suposición obviaba el tener que hacer mayores comentarios al respecto. Sin embargo ahora los dominios son subconjuntos de $\mathbb{R}^n$ $(n \geq 2)$ y por tanto es fácil que los estudiantes no tengan "a priori" una idea sobre cómo han de ser dichos conjuntos. El profesor, en cambio, sabe que existe una "forma general" o al menos una "forma particularmente apropiada" para los conjuntos que constituyen los dominios de definición de las funciones con las que se va a trabajar (pensando en que finalmente llegaremos a estudiar diferenciabilidad): que dichos dominios sean conjuntos abiertos (abierto con la topología de $\mathbb{R}^n$ inducida por la norma (usualmente se considera la norma euclidiana)). Esta situación podría motivar al profesor a comenzar a dar las definiciones topológicas correspondientes: punto interior, punto exterior, punto frontera, conjunto abierto, etc..

Sobre esto, una advertencia: el curso de Cálculo vectorial es un curso muy extenso y con la consiguiente limitación de tiempo para ver el material previsto, esto fuerza al profesor a prescindir frecuentemente de lo que no sea estrictamente necesario para la continuidad del curso. Esta referencia a conceptos topológicos puede evitarse, por un lado, por no ser estrictamente necesaria y además porque supone sobrecargar de material nuevo al estudiante.

Por tanto, por simplicidad y a efectos de ahorrar tiempo, se recomienda limitarse a dominios que sean conjuntos abiertos, de esta forma el único concepto que se necesita introducir es el de ε - vecindad y se hace a los estudiantes la observación de que:

"Así como en el Cálculo de una variable, al considerar un punto en el que la función estuviese definida, pensábamos en que de hecho, lo estaba en todo un intervalo que contuviera al punto, ahora al considerar un punto que esté en el dominio de la función, supondremos que la función de hecho, está definida en todos los puntos "que rodean" al punto en cuestión si éstos están suficientemente cercanos. Para precisar esta idea introducimos la idea de ε - vecindad de un punto p de $\mathbb{R}^n$ como el conjunto $\{x \in \mathbb{R}^n \ / \ \|x - p\| < \varepsilon\}$

que se denota $B_\varepsilon(p)$. Veamos qué forma tienen las vecindades para los casos $n = 1, 2, 3$.

(i) En $\mathbb{R}$ con $p = 2$, una ε - vecindad de p es un intervalo abierto $\{t \in \mathbb{R} \;/\; |t - 2| < \varepsilon\} = (2 - \varepsilon, 2 + \varepsilon)$ (Figura 3.2 (i)).

(ii) En $\mathbb{R}^2$ con $p = (-1, 1)$ una ε - vecindad de p es el interior de un círculo (o disco) de radio ε con centro en p: $\{(x_1, x_2) \in \mathbb{R}^2 \;/\; ((x_1 + 1)^2 + (x_2 - 1)^2)^{1/2} < \varepsilon\}$ (Figura 3.2 (ii)).

(iii) En $\mathbb{R}^3$ con $p = (-1, 1, 2)$ una ε - vecindad de p, es el interior de una bola de radio ε con centro en p: $\{(x_1,x_2,x_3) \in \mathbb{R}^3 \;/\; ((x_1+1)^2 + (x_2-1)^2 + (x_3-2)^2)^{1/2} < \varepsilon\}$. (Figura 3.2 (iii)).

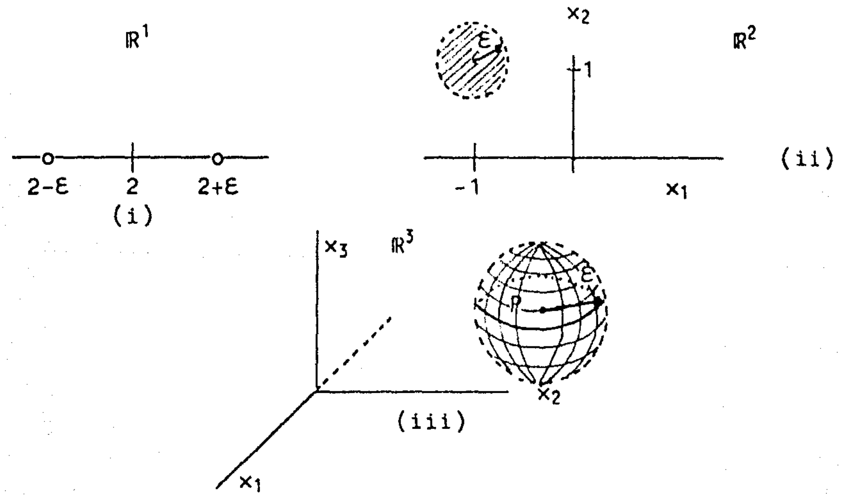


Figura 3.2

Por ello para designar a una ε - vecindad para los casos $n = 1, 2, 3$ también se utilizan los nombres más familiares de intervalo, disco y bola.

Con esta terminología al trabajar con alguna función y considerar algún punto p de su dominio, supondremos que existe un número real $\varepsilon > 0$ tal que f está definida para $B_\varepsilon(p)$."

Con ese breve antecedente se facilita estudiar límites, continuidad y diferenciabilidad para estas funciones.

Para la definición de límite se propone presentar la formulación clásica en términos de epsilons y deltas, con la correspondiente interpretación geométrica:

"Definición.

Dada una función $f\colon D \subseteq \mathbb{R}^n \to \mathbb{R}$ y $p \in D$ decimos que $f(x)$ tiende a L ($L \in \mathbb{R}$) conforme x tiende a p denotado $\lim_{x\to p} f(x) = L$ si para cualquier $\varepsilon > 0$ se puede encontrar un número $\delta > 0$ tal que $|f(x) - L| < \varepsilon$ siempre que $0 < ||x - p|| < \delta$ ($x \in D$).

Para $f\colon D \subseteq \mathbb{R} \to \mathbb{R}$ teníamos que la definición $\lim_{x\to p} f(x) = L$ quería decir, geométricamente que para cualquier intervalo de radio ε alrededor de L, podíamos encontrar un intervalo de radio δ alrededor de p tal que la imagen bajo f de cualquier punto de este segundo intervalo - a excepción quizá de p mismo - caía en el intervalo alrededor de L. (Ver Figura 3.3)

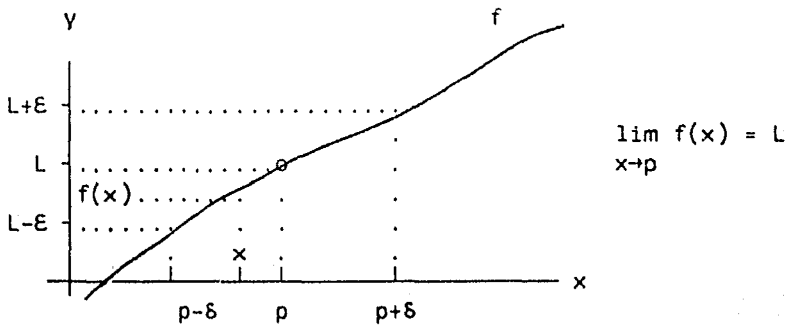


Figura 3.3

Para el caso de dos dimensiones $f\colon D \subseteq \mathbb{R}^2 \to \mathbb{R}$ tenemos una interpretación análoga sólo que en lugar de considerar intervalos de radio δ con centro en p, serían discos de radio δ con centro en p. Concretamente si $f\colon D \subseteq \mathbb{R}^2 \to \mathbb{R}$, $p = (a,b) \in \mathbb{R}^2$ y

$\lim\limits_{(x,y)\to(a,b)} f(x,y) = L$, tenemos que para cualquier intervalo de radio ε alrededor de L: $(L - \varepsilon, L + \varepsilon)$; es posible encontrar una $\delta > 0$ tal que para todo punto (x,y) dentro del círculo de radio δ con centro en (a,b) - excepto, posiblemente el punto (a,b) mismo - el número $f(x,y)$ se encuentra en el intervalo $(L - \varepsilon, L + \varepsilon)$ (i.e. si $0<((x-a)^2 + (y-b)^2)^{1/2} < \delta$ entonces $|f(x,y) - L| < \varepsilon$) como se ilustra en la Figura 3.4.

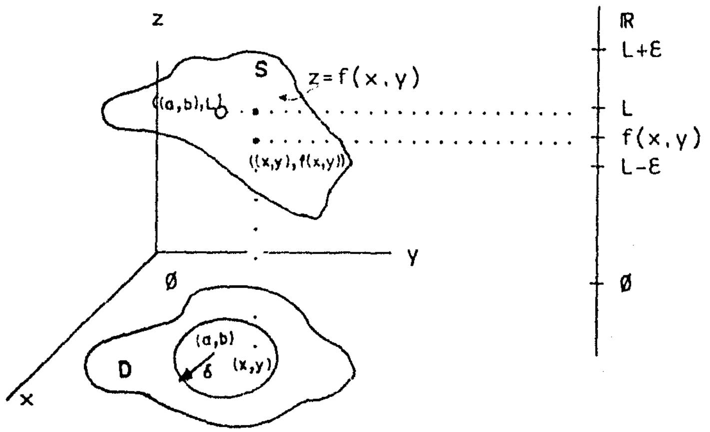


Figura 3.4

Si pensamos en términos de la gráfica, como la gráfica de una función $f: D \subseteq \mathbb{R}^2 \to \mathbb{R}$ es una superficie S en $\mathbb{R}^3$ cuya ecuación es $z = f(x,y)$. Intuitivamente $\lim\limits_{(x,y)\to(a,b)} f(x,y) = L$ nos dice que cuando el punto $(x,y,0)$ se acerca al $(a,b,0)$ sobre el plano xy; el punto correspondiente $(x,y, f(x,y))$ sobre S se acerca al (a,b,L) (que puede estar o no estar sobre S).

A fin de cuentas, la definición de límite; al igual que en $\mathbb{R}$, sólo precisa matemáticamente la idea de que "f(x) está tan cerca de L como queramos, con tal que x esté suficientemente cerca de p"."

La experiencia ha demostrado que el tema de límites en general resulta ser un tema difícil para los estudiantes, más ahora, si están tratando con un objeto nuevo para ellos que son las funciones de varias variables. Es por tanto conveniente dotar a los estudiantes de recursos concretos que les permitan enfrentar los problemas. En el siguiente texto se propone un esquema que los estudiantes pueden seguir: sustitución directa, probar que el límite no existe (viendo que la función tiene límites distintos conforme nos acercamos al mismo punto de maneras diferentes), o aplicar otros recursos como cambio a otro sistema de coordenadas o intentar aplicar directamente la definición.

"En la práctica, para calcular $\lim_{x \to p} f(x)$ se puede intentar primeramente sustituir directamente como se hacía en el caso de funciones de una variable, si la sustitución directa, da lugar a una forma indeterminada, se habrá de investigar el límite siguiendo diversas trayectorias, lo que puede mostrar que el límite no existe; o bien, puede resultar conveniente usar otro sistema de coordenadas para calcular el límite. Se proporcionan tres ejemplos:

Ejemplo 1. Calcular $\lim_{(x,y)\to(0,0)} \frac{x^2y}{x^4+y^2}$ si este límite existe. Como la sustitución directa nos da una forma indeterminada, intentamos acercarnos al $(0,0)$ por algunas trayectorias. Intentemos acercarnos primeramente por los ejes: por el eje $x \Rightarrow y = 0$

$$\lim_{(x,y)\to(0,0)} f(x,y) = \lim_{(x,y)\to(0,0)} f(x,0) = \lim_{(x,y)\to(0,0)} \frac{0}{x^4} = 0$$

Por el eje $y \Rightarrow x = 0$

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

Intentemos acercarnos por cualquier recta no horizontal que pase por el origen, $y = mx \quad m \neq 0$.

$$\lim_{(x,y)\to(0,0)} \frac{x^2y}{x^4+y^2} = \lim_{(x,y)\to(0,0)} \frac{x^2(mx)}{x^4+(mx)^2} =$$

$$\lim_{(x,y)\to(0,0)} \frac{mx^3}{x^4+m^2x^2} = \lim_{(x,y)\to(0,0)} \frac{mx}{x^2+m^2} = \frac{0}{0+m^2} = 0.$$

De acuerdo con esto; podríamos pensar que $f(x,y)$ tiene el límite 0 cuando (x,y) tiende a $(0,0)$. Sin embargo: si (x,y) se aproxima a $(0,0)$ a lo largo de la parábola $y=x^2$ entonces:

$$\lim_{(x,y)\to(0,0)} \frac{x^2y}{x^4+y^2} = \lim_{(x,y)\to(0,0)} \frac{x^2(x^2)}{x^4+(x^2)^2} =$$

$$\lim_{(x,y)\to(0,0)} \frac{x^4}{2x^4} = \lim_{(x,y)\to(0,0)} \frac{1}{2} = \frac{1}{2}.$$

Hemos visto que no toda trayectoria que pasa por $(0,0)$ lleva al mismo valor límite y por lo tanto el límite no existe.

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

Cambiando a coordenadas polares tenemos:

$$\lim_{(x,y)\to(0,0)} \frac{5x^2y}{x^2+y^2} = \lim_{r\to 0} \frac{5(r\cos\theta)^2\,(r \operatorname{sen}\theta)}{r^2}$$

(pues si $(x,y)\to(0,0)$ entonces $r\to 0$)

$$= \lim_{r\to 0} 5\,r \operatorname{sen}\theta\cos^2\theta = (5\operatorname{sen}\theta\cos^2\theta)\lim_{r\to 0} r = 0.$$

Ejemplo 3. Probar que $\displaystyle\lim_{(x,y)\to(0,0)} \frac{x^2y^2}{x^2+y^2} = 0$

Sea $\varepsilon>0$ dada. Queremos ver que existe δ tal que $|f(x,y)-L| = |f(x,y)| < \varepsilon$ si $\|(x,y)-(0,0)\|=\|(x,y)\|=(x^2+y^2)^{1/2} < \delta$

como $x^2 \le x^2 + y^2$
$y^2 \le x^2 + y^2$

tenemos que $\displaystyle |f(x,y)| = \frac{x^2y^2}{x^2+y^2} \le \frac{(x^2+y^2)^2}{x^2+y^2} = x^2+y^2$

con lo que si $\|(x,y)\| = (x^2+y^2)^{1/2} < (\varepsilon)^{1/2} \Rightarrow x^2+y^2 < \varepsilon$ y tendríamos que $|f(x,y)| \le x^2+y^2 < \varepsilon$ $\therefore$ si $\delta = (\varepsilon)^{1/2}$ $\Rightarrow |f(x,y)| < \varepsilon$ siempre que $\|(x,y)\| < \delta$

i.e. por definición $\lim\limits_{(x,y)\to(0,0)} \dfrac{x^2y^2}{x^2+y^2} = 0$."

La idea de los ejemplos anteriores es, como se mencionó, mostrar al estudiante distintos recursos con los que puede contar para determinar límites de funciones.

Después de estudiar la idea de límite para estas funciones, la continuación habitual y lógica es hablar de la continuidad. Sobre esto, el único comentario que se hará es señalar las pautas mínimas que habría que cubrir:

"Para definir la continuidad de una función $f: D \subseteq \mathbb{R}^n \to \mathbb{R}$, en p ($p \in D$) necesitamos al igual que en el caso real, las tres condiciones:

(i) f esté definida en p

(ii) $\lim\limits_{x\to p} f(x)$ existe

(iii) $\lim\limits_{x\to p} f(x) = f(p)$

Tenemos la siguiente:

Definición. Sea $f: D \subseteq \mathbb{R}^n \to \mathbb{R}$, $p \in D$. Decimos que f es continua en p si: $\lim\limits_{x\to p} f(x) = f(p)$

En general, nos interesará la continuidad de una función en su dominio. Decimos que $f: D \subseteq \mathbb{R}^n \to \mathbb{R}$ es continua si f es continua en p, $\forall\ p \in D$.

Vemos nuevamente que es, escribir en términos precisos, la idea de que f(x) esté muy próxima a f(p) si x está suficientemente cerca de p.

Para $f: D \subseteq \mathbb{R}^2 \to \mathbb{R}$ la continuidad de la función significa geométricamente que su gráfica (que es una superficie en $\mathbb{R}^3$) no tiene rupturas. De hecho podemos pensar que para una función $f: D \subseteq \mathbb{R}^n \to \mathbb{R}$ la continuidad de la función significa que su gráfica, que la podemos imaginar como una especie de "superficie" en $\mathbb{R}^{n+1}$ no tiene rupturas."

Sobre el tema de continuidad, una observación que a veces se omite y sí conviene hacerla pues hace ver a los estudiantes el tipo de dificultades que aparecen al trabajar con funciones de varias variables. Específicamente les hace advertir que para entender el comportamiento de una función no basta con analizar su comportamiento en cada variable por separado. La observación es que si una función de varias variables es continua, entonces también lo es en cada variable (i.e. si mantenemos fijas las demás variables). Mientras que una función puede ser continua en cada variable por separado y la función misma no ser continua. Como ejemplo de la situación anterior se puede proponer:

$$f(x,y) \begin{cases} \dfrac{x\,y}{x^2+y^2} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{si } (x,y) = (0,0) \end{cases}$$

$f(x,0) \equiv 0 \equiv f(0,y) \therefore f(x,0)$ como función de x es continua en $x = 0$ análogamente con y. Y sin embargo f no es continua en $(0,0)$ como función de dos variables ya que $\lim\limits_{(x,y)\to(0,0)} f(x,y)$ no existe. (Esto puede verse fácilmente por ejemplo si utilizamos coordenadas polares).

Al encarar el tema de la diferenciabilidad para las funciones reales de varias variables, lo hacemos guiados por un objetivo claro: generalizar para estas funciones el concepto de derivada. La primera idea que se ocurre es seguir un recurso de las Matemáticas que consiste en reducir de dimensión un problema. En términos de enseñanza y aprendizaje esto lleva a reducir el problema nuevo a uno cuya resolución es conocida.

En el contexto en el que estamos, esto se traduce en buscar reducir a las funciones de varias variables a funciones de una variable y aplicarles las técnicas de diferenciación que ya se conocen. Esto se logra a través de los conceptos de derivada parcial y de derivada direccional.

Es por ello que la mayoría de los textos de la materia, después de haber hablado de los límites y la continuidad para funciones de varias variables, abordan alguno de los dos temas mencionados.

Sobre esta posibilidad para el orden del curso se sugiere optar por el estudio de las derivadas parciales por las siguientes razones:

i) Es lo coherente con la línea que se ha mencionado de ir de lo particular a lo general: las derivadas direccionales son una generalización de las derivadas parciales.

ii) Si bien en ambos conceptos se reduce la función a una función de una variable, en el caso de las derivadas parciales esta reducción es inmediata: simplemente considerar todas las variables menos una como constantes. Mientras que para las derivadas direccionales la reducción es a través de un razonamiento indirecto: hay que parametrizar una recta que pasa por un punto y con una dirección determinados y pensar en la función restringida a esta recta, que de esta forma puede pensarse como función de una variable, a saber el parámetro (de la recta).

iii) El proceso mismo de derivación parcial es una experiencia más familiar en el sentido de que cuando se ha de trabajar con un fenómeno en el que intervienen varias variables, la manera natural de investigar las relaciones cuantitativas entre las variables es dejar fijas todas menos una y analizar el cambio que se obtiene al modificarla, lo que expresa Larson, en términos más formales, en la introducción al tema de derivadas parciales de su libro:

"Al considerar aplicaciones de las funciones de varias variables, surge a menudo la cuestión: ¿cómo se verá afectada la función si se varían, una, varias o todas sus variables independientes? Podemos contestar dicha cuestión, al menos en parte, si consideramos las variables independientes una a una. Por ejemplo, el producto nacional bruto es función de muchas variables, como pueden ser los tipos de impuestos, la tasa de desempleo o las guerras. Un economista que intente determinar el efecto que produciría un aumento de los impuestos mantendría todas las demás variables constantes, mientras se suben o bajan los impuestos. Asimismo, para determinar el efecto de un agente catalítico en un experimento, un químico repetiría el experimento un cierto número de veces, usando diversas cantidades de agente catalítico, mientras mantiene constantes otras variables como pueden ser la temperatura y la presión.

En matemáticas se sigue un procedimiento similar para determinar la razón de cambio de una función f con respecto a una o a varias de sus variables independientes. Dicho procedimiento consiste en obtener la derivada de f con respecto a cada variable independiente mientras se mantienen constantes los valores de las demás variables. Este proceso recibe el nombre de **derivación parcial** y el resultado se denomina **derivada parcial** de f con respecto a la variable independiente elegida."

(Larson R.- Hostetler R.; <u>Cálculo y Geometría Analítica</u>. México, McGraw-Hill, 1986, pág. 689).

Es por ello que la mayoría de los textos de la materia, después de haber hablado de los límites y la continuidad para funciones de varias variables, abordan alguno de los dos temas mencionados.

Sobre esta posibilidad para el orden del curso se sugiere optar por el estudio de las derivadas parciales por las siguientes razonés:

i) Es lo coherente con la línea que se ha mencionado de ir de lo particular a lo general: las derivadas direccionales son una generalización de las derivadas parciales.

ii) Si bien en ambos conceptos se reduce la función a una función de una variable, en el caso de las derivadas parciales esta reducción es inmediata: simplemente considerar todas las variables menos una como constantes. Mientras que para las derivadas direccionales la reducción es a través de un razonamiento indirecto: hay que parametrizar una recta que pasa por un punto y con una dirección determinados y pensar en la función restringida a esta recta, que de esta forma puede pensarse como función de una variable, a saber el parámetro (de la recta).

iii) El proceso mismo de derivación parcial es una experiencia más familiar en el sentido de que cuando se ha de trabajar con un fenómeno en el que intervienen varias variables, la manera natural de investigar las relaciones cuantitativas entre las variables es dejar fijas todas menos una y analizar el cambio que se obtiene al modificarla, lo que expresa Larson, en términos más formales, en la introducción al tema de derivadas parciales de su libro:

"Al considerar aplicaciones de las funciones de varias variables, surge a menudo la cuestión: ¿cómo se verá afectada la función si se varían, una, varias o todas sus variables independientes? Podemos contestar dicha cuestión, al menos en parte, si consideramos las variables independientes una a una. Por ejemplo, el producto nacional bruto es función de muchas variables, como pueden ser los tipos de impuestos, la tasa de desempleo o las guerras. Un economista que intente determinar el efecto que produciría un aumento de los impuestos mantendría todas las demás variables constantes, mientras se suben o bajan los por:

$$\frac{\partial f}{\partial x}(a,b);\ f_x(a,b);\ f_1(a,b) \text{ ó } \frac{\partial f}{\partial x}(P_\emptyset);\ f_x(P_\emptyset);\ f_1(P_\emptyset).$$

Sobre el conjunto de puntos P donde $f_1(P)$ existe, estos valores definen una función que será llamada la derivada parcial de f con respecto a x y que se denota por $\frac{\partial f}{\partial x}$; f_x ó f_1, es decir

$$(a,b) \xrightarrow{\frac{\partial f}{\partial x}} \frac{\partial f}{\partial x}(a,b).$$

De manera similar, si $\lim\limits_{h\to\emptyset} \frac{f(a,b+h) - f(a,b)}{h}$ existe, se denota por $\frac{\partial f}{\partial y}(a,b)$; $f_y(a,b)$ ó $f_2(P_\emptyset)$, ...

La función definida para aquellos puntos P tales que $\frac{\partial f}{\partial y}(P)$ existe, se llama la derivada parcial de f con respecto a y que denotamos por $\frac{\partial f}{\partial y}$, f_y ó f_2.

Es decir

$$P \xrightarrow{\frac{\partial f}{\partial y}} \frac{\partial f}{\partial y}(P)$$

Las derivadas parciales heredan su interpretación geométrica de la correspondiente para derivadas ordinarias. La derivada de una función f de una variable en el punto a, es la pendiente de la recta tangente a la curva descrita por y = f(x) en x = a. Por lo tanto $f_x(a,b)$ es la pendiente de la recta tangente a la curva descrita por z = f(x,b) en el

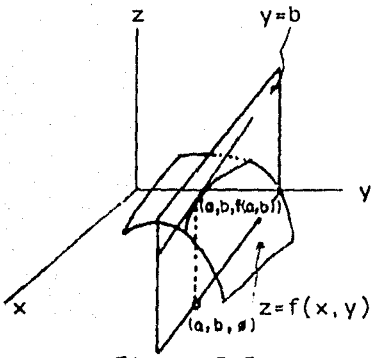


Figura 3.5

punto (a,b). Curva que está en el plano y = b, i.e. la curva que resulta de la intersección de la superficie z = f(x,y) y el plano y = b (Figura 3.5). Análoga interpretación se tiene para $f_y(a,b)$ - i.e. : es la pendiente de la recta tangente a la curva descrita por z = f(a,y) en el punto (a,b). Dicha curva está en el plano x = a.

Se puede hacer también la siguiente interpretación de las derivadas parciales:

Si f: $D \subseteq \mathbb{R}^2 \to \mathbb{R}$, entonces podemos pensar en que f es una función tal que a cada punto (x,y) de una cierta región D del plano le asocia su temperatura f(x,y) y ésta se registra en un termómetro representado por un eje w. (Figura 3.6).

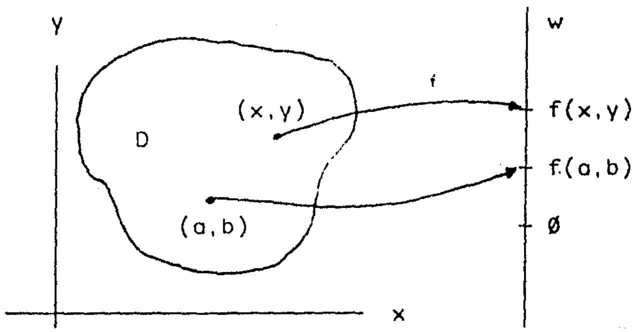


Figura 3.6

Observamos que los puntos (a,b) y (a+h,b) se encuentran sobre la misma recta horizontal, como se indica en la Figura 3.7

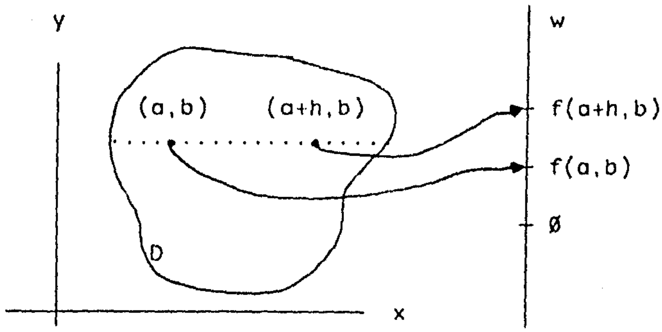


Figura 3.7

La diferencia f(a+h,b) - f(a,b) es el cambio que sufriría la temperatura cuando el punto se mueve de (a,b) a (a+h,b) y el cociente f(a+h,b) - f(a,b) / h es el cambio medio de temperatura. Por ejemplo si el cambio en la temperatura es de dos grados y h = 4 entonces esta razón es 1/2, es decir, en promedio la temperatura cambia (sobre la recta y = b entre (a,b) y (a+h,b)) a razón de 1/2 grado por unidad de distancia. Si ahora tomamos el límite cuando $h \to 0$, i.e. $\lim_{h \to 0} \frac{f(a+h,b) - f(a,b)}{h} = \frac{\partial f}{\partial x}(a,b)$ tenemos la razón de cambio instantánea de la temperatura con respecto a la distancia en el punto (a,b) a lo largo de la dirección horizontal. Análogamente $\frac{\partial f}{\partial y}(a,b)$ mide la razón de cambio instantánea de f en (a,b) a lo largo de la dirección vertical. (cfr. Swokowski Earl, Calculus with Analytic Geometry. Boston, Mass.- Prindle, Weber & Schmidt, 2nd. Ed., 1979, 768 - 771).

Como las derivadas parciales de una función de dos variables vuelven a ser funciones de dos variables, la operación puede repetirse. Las segundas derivadas se definen como derivadas de primeras derivadas, las terceras derivadas como derivadas de segundas derivadas; de esta forma tenemos definidas derivadas de órdenes superiores. El orden de una derivada es el número total de derivaciones que se han realizado para calcular la derivada en cuestión. Las notaciones de las derivadas de órdenes superiores son usualmente:

Segundas derivadas

$$\frac{\partial}{\partial x}\left[\frac{\partial f}{\partial x}\right] = \frac{\partial^2 f}{\partial x^2} = f_{xx} = f_{11}$$

$$\frac{\partial}{\partial x}\left[\frac{\partial f}{\partial y}\right] = \frac{\partial^2 f}{\partial x \partial y} = f_{yx} = f_{21}$$

$$\frac{\partial}{\partial y}\left[\frac{\partial f}{\partial x}\right] = \frac{\partial^2 f}{\partial y \partial x} = f_{xy} = f_{12}$$

$$\frac{\partial}{\partial y}\left[\frac{\partial f}{\partial y}\right] = \frac{\partial^2 f}{\partial y^2} = f_{yy} = f_{22}$$

Terceras derivadas

$$\frac{\partial}{\partial x}\left[\frac{\partial^2 f}{\partial x^2}\right] = \frac{\partial^3 f}{\partial x^3} = f_{xxx} = f_{111}$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

Definición. Sea $f\colon D \subseteq \mathbb{R}^n \to \mathbb{R}$, continua. Se dice que f es de clase C^K ($K \in \mathbb{N}$) en D si todas las parciales de orden $\leq K$ existen y son funciones continuas en D.

Estas ideas se pueden generalizar a $\mathbb{R}^n$ para $n > 2$, definiendo la derivada parcial para funciones de n variables.

Definición. Sea $f\colon D \subseteq \mathbb{R}^n \to \mathbb{R}$. Entonces las derivadas parciales

$$\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n}$$

de f con respecto a la primera, segunda y n-ésima variable también son funciones de n variables con valores reales definidas en el punto $x = (x_1, x_2, \ldots, x_n) \in D$ como:

$$\frac{\partial f}{\partial x_j}(x)=\frac{\partial f}{\partial x_j}(x_1,\ldots,x_n)=\lim_{h\to 0}\frac{f(x_1,x_2,\ldots x_j+h,\ldots x_n)-f(x_1,x_2,\ldots x_n)}{h}$$

$$=\lim_{h\to 0}\frac{f(x+he_j)-f(x)}{h} \quad \text{si el límite existe}$$

donde $1 \leq j \leq n$ y e_j es el vector cuyas coordenadas son todas cero, salvo un 1 en la j-ésima coordenada.

Es decir $\frac{\partial f}{\partial x_j}$ es sólo la derivada de f con respecto a la variable x_j manteniendo las otras variables fijas."

Después de haber visto lo concerniente a derivadas parciales, como la siguiente idea en la línea de generalizar para las funciones de varias variables el concepto de derivada que tenemos para las funciones de una sola variable, conviene introducir el tema de las derivadas direccionales. Pues pueden verse como una continuación lógica de las derivadas parciales a la luz del siguiente cuestionamiento:

Si las derivadas parciales nos miden la razón de cambio de la función (con respecto a la distancia) a lo largo de las direcciones horizontal y vertical (para el caso de dos variables) ¿no será posible definir la razón de cambio de una función en un punto a lo largo de cualquier dirección dada?

En primer lugar habría que precisarles a los estudiantes qué se entiende por "dirección" y por "moverse a lo largo de una dirección". Esto se puede lograr comentando lo siguiente:

"En $\mathbb{R}^1$ tenemos sólo dos direcciones: izquierda y derecha. En $\mathbb{R}^2$ podemos pensar en ángulos como la manera adecuada de describir direcciones (cada dirección puede escribirse como $(\cos\theta, \operatorname{sen}\theta)$ con $0 \leq \theta < 2\pi$). En $\mathbb{R}^3$ y en general en $\mathbb{R}^n$ es más fácil decir que una dirección es simplemente un punto u en $\mathbb{R}^n$ tal que $\|u\| = 1$.

Tal punto está sobre la esfera unitaria, intuitivamente, pensamos en dicho punto como el vector unitario cuyo extremo inicial es el origen y cuyo extremo final corresponde al punto u. Por ejemplo, si deseamos alejarnos del punto x_0 en la dirección u, entendemos que hemos de movernos desde x_0 a lo largo del segmento de recta hacia el punto $x_0 + u$. (Ver Figura 3.8). Sabemos que en general, la recta con dirección u y que pasa por x_0 es el conjunto $\{x \in \mathbb{R}^n \;/\; x = x_0 + tu\}_{t\in\mathbb{R}}$."

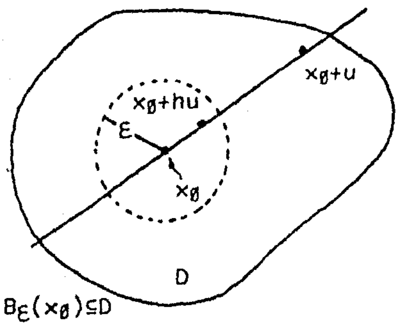


Figura 3.8

Con estos elementos se puede definir entonces para una función $f: D \subseteq \mathbb{R}^n \to \mathbb{R}$ la razón instantánea de cambio de f en x_0 en la dirección u precisamente como la derivada direccional de f en x_0 en la dirección u: $(D_u f)(x_0)$, i.e.

$$"(D_u f)\,(x_0) = \lim_{h\to 0} \frac{f(x_0+hu)-f(x_0)}{h}$$ si dicho límite existe."

Al tratar este tema habría que hacer ver que este concepto es efectivamente , una generalización del concepto de derivada que teníamos para funciones de una variable, i.e., si $f: D\subseteq\mathbb{R}\to\mathbb{R}$ y $u=1$ entonces $(D_u f)\,(x_0)=f'(x_0)$.

(N.B. En $\mathbb{R}$ hay dos direcciones posibles e_1 y $-e_1$ que se identifican con los escalares 1 y -1 respectivamente).

Más aún, convendría comentar que para funciones de varias variables la definición de derivada direccional incluye como casos particulares a las derivadas parciales, pues éstas serían simplemente -en esta nueva terminología- las derivadas direccionales en las direcciones e_1, e_2, ..., e_n

"para $i=1,2,\ldots,n$ $(De_if)(x_\emptyset)=\lim\limits_{h\to\emptyset} \dfrac{f(x_\emptyset+he_i) - f(x_\emptyset)}{h}$

$$= \lim_{h\to\emptyset} \frac{f(x_1,\ldots,x_i+h,\ldots,x_n) - f(x_1,\ldots,x_n)}{h} = \frac{\partial f}{\partial x_i}(x_\emptyset)$$

(Suponiendo que dicho límite exista)"

$$\text{i.e. } (De_if)\,(x_\emptyset) = \frac{\partial f}{\partial x_i}(x_\emptyset)$$

Al tratar con derivadas direccionales se puede señalar que, sustancialmente estamos reduciendo a las funciones de varias variables a funciones de una variable y aplicando la definición de diferenciabilidad que tenemos para ese caso. Pues la línea que pasa por $x_\emptyset$ con dirección u es el conjunto $\{x_\emptyset+hu \ / \ h \in \mathbb{R}\}$ y por tanto los puntos sobre esta línea están determinados por el valor del parámetro h. Con lo que si definimos a la función $F:\mathbb{R}\to\mathbb{R}$ por la regla:

$$F(h) = f(x_\emptyset + hu)$$

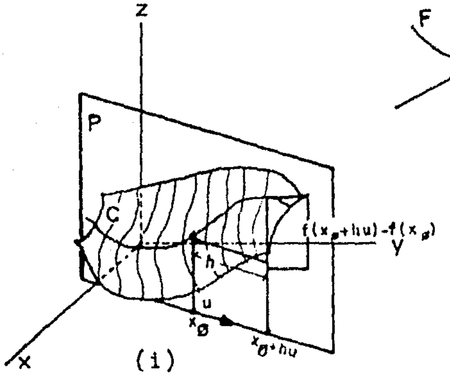
entonces $(D_uf)\,(x_\emptyset) = \left.\dfrac{dF}{dh}\right|_{h=\emptyset}$

Además este proceso permite presentar la interpretación geométrica usual de la derivada direccional para el caso de funciones $f:D\subseteq\mathbb{R}^2\to\mathbb{R}$:

"Si definimos la función F por la regla: $F(h) = f(x_\emptyset + hu)$ entonces F es una función de $\mathbb{R}$ en $\mathbb{R}$ y su gráfica corresponde a la intersección de la gráfica de f con el plano P que es perpendicular al plano xy y que contiene a la recta paralela a u que pasa por $x_\emptyset$. El punto $(\emptyset, F(\emptyset))$ corresponde al punto $(x_\emptyset, f(x_\emptyset))$ (ver figura 3.9). Como

$$D_uf(x_\emptyset) = \lim_{h\to\emptyset} \frac{f(x_\emptyset+hu)-f(x_\emptyset)}{h} = \lim_{h\to\emptyset} \frac{F(h)-F(\emptyset)}{h} = F'(\emptyset)$$

decimos entonces que $(D_u f)(x_\emptyset)$ es la pendiente en $(x_\emptyset, f(x_\emptyset))$ de la curva C formada por la intersección de la gráfica de f y el plano P".



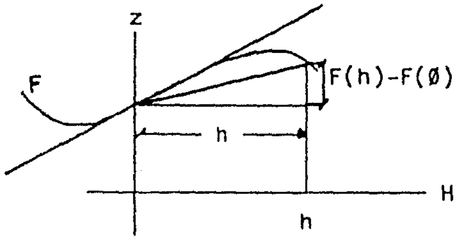


Figura 3.9

Es instructivo ver, en la práctica por medio de ejemplos específicos cómo evaluar las derivadas direccionales a través de considerar a la función restringida a los puntos de la recta, como función de una variable y derivando aplicando las reglas de derivación conocidas:

"Como ilustración consideramos el siguiente ejemplo:
$f(x,y) = x^2+3xy \qquad x_\emptyset = (2,\emptyset) \qquad u = (1/(2)^{1/2}, -1/(2)^{1/2})$
(obsérvese que u especifica la dirección - 45°).

Nos interesa calcular $(D_u f)(x_\emptyset)$

Solución 1. Como $x_\emptyset+hu = (2+(2)^{-1/2}h, -(2)^{-1/2}h)$
tenemos $f(x_\emptyset+hu) =$

$$= (2+(2)^{-1/2}h)^2 + 3(2+(2)^{-1/2}h)(-(2)^{-1/2}h) = 4 - 2^{1/2}h-h^2$$

$$\therefore (D_u f)(x_\emptyset) = \lim_{h\to\emptyset} \frac{(4 - (2)^{1/2}h - h^2) - 4}{h} = -(2)^{1/2}$$

Solución 2.

$$F(h) = f(x_\emptyset+hu) = 4 - \frac{2}{(2)^{1/2}} h - h^2 \Rightarrow \frac{dF}{dh} = - \frac{2}{(2)^{1/2}} -2h$$

$\therefore (D_u f)(x_\emptyset) = F'(\emptyset) = -(2)^{1/2}$."

El paso conceptual siguiente consistiría en hablar de la diferencial de una función de varias variables en un punto. Sin embargo resulta conveniente motivar la introducción de ese tema; ya que los estudiantes podrían pensar que con las derivadas parciales y las derivadas direccionales ya se cuenta con "la" generalización del concepto de derivada para estas funciones, pues después de todo están familiarizados con la idea de que para funciones de una variable, la derivada nos proporciona la razón instantánea de cambio de la función en el punto. Y las derivadas direccionales nos proporciona la razón instantánea de cambio de una función en un punto, a lo largo de cualquier dirección.

Para ver que las derivadas parciales y direccionales son una extensión todavía "algo insatisfactoria" del concepto de derivada en una variable se podrían hacer las siguientes consideraciones:

"En el Cálculo de una variable se probó (cfr. Capítulo 2) que si una función $f:\ D \subseteq \mathbb{R} \to \mathbb{R}^n$ $(n \geq 1)$ es derivable en $x_\emptyset$, entonces f resulta ser continua en $x_\emptyset$.

Analicemos por contraste, los siguientes ejemplos para funciones de dos variables:

Ejemplo a):

$$\text{Sea } f(x,y) = \begin{cases} \dfrac{xy}{x^2+y^2} & (x,y) \neq (\emptyset,\emptyset) \\ \emptyset & \text{si } (x,y) = (\emptyset,\emptyset) \end{cases}$$

f no tiene límite en $(\emptyset,\emptyset)$; pues si nos aproximamos a $(\emptyset,\emptyset)$ por rectas de la forma $y = mx$ tendríamos que:

$$\lim_{(x,y)\to(\emptyset,\emptyset)} \frac{xy}{x^2+y^2} = \lim_{(x,y)\to(\emptyset,\emptyset)} \frac{x(mx)}{x^2+m^2x^2} =$$

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

que nos da diferentes valores dependiendo de la elección de m, $\therefore$ el límite no existe => f no es continua en (0,0). Y sin embargo:

$$f_x(0,0) = \lim_{h\to 0} \frac{f(h,0) - f(0,0)}{h} = \lim_{h\to 0} \frac{0 - 0}{h} = \lim_{h\to 0} 0 = 0$$

$$f_y(0,0) = \lim_{h\to 0} \frac{f(0,h) - f(0,0)}{h} = \lim_{h\to 0} \frac{0 - 0}{h} = \lim_{h\to 0} 0 = 0$$

Se observa entonces que $f_x(0,0)$; $f_y(0,0)$ existen y sin embargo f no es ni siquiera continua en (0,0).

La razón es la siguiente: la existencia de la parcial f_x en un punto p tiene que ver con el comportamiento de la función f sólamente a lo largo de la dirección horizontal, análogamente la existencia de $f_y(p)$ nos dice que f debe comportarse "bien" a lo largo de la dirección vertical (es decir a lo largo de la línea paralela al eje y que pasa por p).

Sin embargo entre estas líneas, la función puede comportarse "bastante mal". Con lo que nos gustaría que el concepto que demos de diferenciabilidad de una función f en un punto - de alguna manera - nos asegure que la función f se comporte bien a lo largo de cualquier dirección.

En el ejemplo anterior, se puede verificar que las derivadas direccionales de f en (0,0) no existen para ninguna dirección salvo las de los ejes con lo que se podría pensar entonces que si todas las derivadas direccionales existieran en el punto, entonces al menos sí tendríamos asegurada la continuidad de la función en el punto. Consideremos para ello el ejemplo (b):

Ejemplo b):

$$\text{Sea } f(x,y) = \begin{cases} \dfrac{xy^2}{x^2+y^4} & \text{si } x \neq 0 \\ 0 & \text{si } x = 0 \end{cases}$$

Sea $u = (a_1, a_2)$ cualquier vector unitario en $\mathbb{R}^2$, tenemos entonces que

$$\frac{f(hu) - f(\emptyset,\emptyset)}{h} = \frac{f(ha_1, ha_2) - f(\emptyset,\emptyset)}{h} = \frac{\dfrac{h^3 a_1 a_2^2}{h^2(a_1^2 + h^2 a_2^4)}}{h}$$

$$= \frac{a_1 a_2^2}{a_1^2 + h^2 a_2^4} \qquad \therefore \text{ si } a_1 \neq \emptyset \qquad (D_u f)\ (\emptyset,\emptyset) = \frac{a_2^2}{a_1}$$

si $a_1 = \emptyset$ entonces $u = (\emptyset, a_2)$ $a_2 = 1$ ó $a_2 = -1$

$$\therefore \frac{f(hu) - f(\emptyset,\emptyset)}{h} = \frac{f(\emptyset, ha_2) - f(\emptyset,\emptyset)}{h} = \frac{\emptyset - \emptyset}{h} = \emptyset$$

$\therefore$ $(D_u f)\ (\emptyset,\emptyset) = \emptyset$

i.e. $(D_u f)\ (\emptyset,\emptyset)$ existe para todas las direcciones u. Por otro lado la función f toma el valor 1/2 en cada punto (x,y) de la parábola $x = y^2$ pues en esos puntos:

$$f(x,y) = f(y^2,y) = \frac{y^2 y^2}{y^4 + y^4} = \frac{y^4}{2y^4} = \frac{1}{2}$$

y $f(\emptyset,\emptyset) = \emptyset$ $\therefore$ f no es continua en el $(\emptyset,\emptyset)$."

La presentación de estos ejemplos debe acompañarse siempre del oportuno comentario del profesor. Ya que la conclusión precipitada que de ellos se puede obtener es que las derivadas parciales no son de utilidad, cuando en la práctica la mayor parte de las funciones con las que se trabaja son C^1; si no en todo su dominio, en un cierto subconjunto de él y son por tanto continuas en dichos puntos (son de hecho diferenciables como repasaremos en el Capítulo 5), y usualmente nos interesa lo que sucede con la función en el conjunto en el que es C^1.

Se puede hacer ver que en ambos ejemplos las derivadas parciales no son funciones continuas en el $(\emptyset,\emptyset)$ y ya desde ahora se puede señalar que si las parciales fueran continuas, no se daría ese comportamiento. Hay que dejar claro que los ejemplos presentados son casos atípicos, que sin embargo, se le presentan al estudiante para convencerle de que la sola existencia de las derivadas (parciales o direccionales) no constituye una extensión completamente adecuada del concepto de derivada; pues para el caso de una variable, la sola existencia de la derivada de la función en un punto sí garantiza la continuidad de la función en el punto.

Una idea que se les podría transmitir a los estudiantes con el ánimo de reforzar la importancia de que la extensión del concepto de derivada implique la continuidad, podría ser la siguiente, que al ser de carácter geométrico y no analítico les suele resultar más asequible:

Los estudiantes han aprendido que el que una función de $\mathbb{R}$ en $\mathbb{R}$ sea diferenciable en un punto es equivalente a que la recta tangente esté bien definida en el punto correspondiente de la gráfica de la función. Si se piensa en funciones de dos variables, por analogía con la situación anterior se esperaría que una generalización del concepto de diferenciabilidad (que se tenía para el caso real) para estas funciones fuese tal que asegurara la existencia de un "plano tangente" bien definido en el punto correspondiente de la gráfica de la función. Es claro, geométricamente hablando, que para que esto suceda, la gráfica no debe tener "rupturas", es decir, la función debe ser continua, sin embargo, no es suficiente esta condición, pues para que el plano tangente esté bien definido, la gráfica no debe tener dobleces cortantes, esquinas o picos (¿cómo se podría definir de manera única el plano tangente en el vértice de un cono?). Por tanto, si no hay continuidad de la función en el punto, no se puede definir un plano tangente en el correspondiente punto de la gráfica.

La misma idea geométrica de definir la diferenciabilidad para las funciones de varias variables como la existencia del plano tangente puede ser más esclarecedora para los estudiantes.

Otra observación que podría hacerse en torno a las dificultades que aparecen al tratar de definir la diferenciabilidad para las funciones de varias variables generalizando la definición de $f'(x_0)$ para una función de una variable es la siguiente:

"Se ha visto que la condición de diferenciabilidad para una función $f: D \subseteq \mathbb{R} \to \mathbb{R}$ en un punto p estaba dada por la existencia de $\lim_{h \to 0} \dfrac{f(p + h) - f(p)}{h}$.

Si $f: D \subseteq \mathbb{R} \to \mathbb{R}^n$, la condición de diferenciabilidad para f en p está dada por la existencia de

$$\lim_{h \to 0} \frac{f(p + h) - f(p)}{h} .$$

Claramente la expresión anterior - que formalmente es la misma para ambos casos: el de funciones de $\mathbb{R}$ en $\mathbb{R}$ y de funciones de $\mathbb{R}$ en $\mathbb{R}^n$ - no puede trasladarse sin más a este nuevo objeto de las funciones de varias variables $f: D \subseteq \mathbb{R}^n \to \mathbb{R}$ pues en este caso h es un vector de $\mathbb{R}^n$ y no tiene sentido dividir entre vectores de $\mathbb{R}^n$ ($n \geq 2$). Sin embargo, si tratamos de conservar la forma de la definición podríamos pensar en definir la derivada de una función $f: D \subseteq \mathbb{R}^n \to \mathbb{R}$ en un punto $p \in D$ como $\lim_{h \to 0} \dfrac{f(p + h) - f(p)}{||h||}$.

Para ver que este intento no resulta, consideremos el siguiente ejemplo.

Sea $f: \mathbb{R}^2 \to \mathbb{R}$ definida por $f(x, y) = x + y$. Es una función muy simple, su gráfica $z = x + y$ es un plano que pasa por el origen. Si se entiende diferenciabilidad como la existencia del plano tangente, f debería ser una función diferenciable en todo $\mathbb{R}^2$, y sin embargo si tratamos de aplicar a f la "definición" de diferenciabilidad propuesta anteriormente tenemos que:

$$\lim_{h \to 0} \frac{f(p + h) - f(p)}{||h||} = \lim_{(h_1, h_2) \to (0, 0)} \frac{f((x, y)+(h_1, h_2))-f(x, y)}{(h_1^2+h_2^2)^{1/2}}$$

$$= \lim_{(h_1, h_2) \to (0, 0)} \frac{f(x+h_1, y+h_2)-f(x, y)}{(h_1^2+h_2^2)^{1/2}}$$

$$= \lim_{(h_1,h_2)\to(\emptyset,\emptyset)} \frac{x+h_1+y+h_2-x-y}{(h_1^2+h_2^2)^{1/2}} =$$

$$= \lim_{(h_1,h_2)\to(\emptyset,\emptyset)} \frac{h_1 + h_2}{(h_1^2+h_2^2)^{1/2}}$$

Para ver que dicho límite no existe, podemos cambiar a coordenadas polares:

$$\lim_{(h_1,h_2)\to(\emptyset,\emptyset)} \frac{h_1 + h_2}{(h_1^2+h_2^2)^{1/2}} = \lim_{r\to\emptyset} \frac{r\ (\cos\theta+\operatorname{sen}\theta)}{r} =$$

$$\lim_{r\to\emptyset} (\cos\theta+\operatorname{sen}\theta) = \cos\theta+\operatorname{sen}\theta$$

que da valores diferentes para diferentes maneras de acercarse al $(\emptyset,\emptyset)$ $\therefore$ para cualquier $p = (x,y) \in \mathbb{R}^2$

$$\lim_{h\to\emptyset} \frac{f(p + h) - f(p)}{||h||} \quad \text{no existe.}$$

Es decir, la función $f(x,y) = x + y$ según nuestra "definición" no sería diferenciable."

Las consideraciones anteriores tienen por objeto haber motivado suficientemente la presentación a los estudiantes del tema de la diferencial de una función en un punto como la generalización adecuada para las funciones de varias variables del concepto de derivada de una función de variable real (en un punto). Generalización adecuada en el sentido de que su existencia nos asegurará la continuidad de la función y la existencia del respectivo espacio tangente (en el punto).

Una buena concepción del concepto de la diferencial de una función supone un cierto conocimiento y manejo de las transformaciones lineales. Es por ello, que se propone hacer en este momento del curso, una digresión para hablar de las transformaciones lineales. Un propósito importante del presente trabajo es convencer al profesor de la conveniencia de esta interrupción en el curso de la asignatura. Que si bien, de momento le supone al profesor recortar el tiempo, de por sí restringido, del que dispone para cubrir el material, esta inversión se verá luego recuperada cuando a la luz de las transformaciones lineales, a los estudiantes se les facilite aprender con mayor unidad, temas que tradicionalmente constituyen escollos dentro de la enseñanza del Cálculo de varias variables como es la Regla de la Cadena.

Otra razón que podría reforzar el convencimiento sobre las ventajas de introducir en este momento un estudio somero acerca de las transformaciones lineales es el hecho de que en buena parte de la literatura existente sobre el tema , se omite este paso para hablar directamente de la diferencial total de una función -donde el mismo nombre sugiere que en dicho concepto se toma en consideración, de alguna manera, el comportamiento de todas las variables de las que depende la función - con un tratamiento que resultará mucho más claro para los estudiantes si previamente se han introducido los elementos necesarios sobre transformaciones lineales.

A manera de ejemplo analicemos con cierto detalle la manera en que uno de esos libros desarrolla estas ideas. El contenido y la notación son bastante parecidos a los que se encuentran en varios textos.

cfr. Kaplan, Wilfred. Cálculo Avanzado. México, CECSA. 1a. edición, Junio'83. cfr. páginas 173 -180.

Se comienza hablando de las derivadas parciales para funciones de dos variables de manera muy parecida a la que se propuso en el presente capítulo, si acaso se prefiere la siguiente notación:

$$\frac{\partial f}{\partial x}(x_0, y_0) = \lim_{\Delta x \to 0} \frac{f(x_0 + \Delta, y_0) - f(x_0, y_0)}{\Delta x}$$

$$\frac{\partial f}{\partial y}(x_0, y_0) = \lim_{\Delta y \to 0} \frac{f(x_0, y_0 + \Delta y) - f(x_0, y_0)}{\Delta y}$$

Se define como Δz el incremento dado por: $\Delta z = f(x + \Delta x, y + \Delta y) - f(x, y)$, considerando x,y como constantes, esto define a Δz como una función de Δx, Δy con la propiedad de que $\Delta z = 0$ cuando $\Delta x = 0$, $\Delta y = 0$.

A continuación se suele analizar algunos ejemplos concretos para ver que para determinadas funciones $z=f(x,y)$; Δz puede escribirse de la forma: $\Delta z = a\Delta x+b\Delta y$ + términos de orden más alto en Δx y Δy, donde $a,b \in \mathbb{R}$.

"En general se dice entonces que la función $z=f(x,y)$ tiene una diferencial total en el punto (x,y) si en ese punto $\Delta z = a\Delta x + b\Delta y + \varepsilon_1\Delta x + \varepsilon_2\Delta y$ donde a,b son independientes de Δx, Δy y ε_1, ε_2 son funciones de Δx, Δy tales que: $\lim_{(\Delta x,\Delta y)\to(\emptyset,\emptyset)} \varepsilon_1 = \emptyset$, $\lim_{(\Delta x,\Delta y)\to(\emptyset,\emptyset)} \varepsilon_2 = \emptyset$. "

Una observación que se puede hacer aquí es que, como puede verse del texto anterior, no queda clara la naturaleza de las funciones ε_1, ε_2. Se menciona que son funciones de Δx, Δy pero no se menciona si son funciones de otras variables además de Δx, Δy. Sólo se especifica que se van a cero cuando Δx, Δy se van a cero, y cabría entonces preguntarse que tan rápido lo hacen. Prosiguiendo con el texto: "A la función de Δx y de Δy dada por $a\Delta x + b\Delta y$ se le denomina entonces la diferencial total de z en el punto (x,y) y se denota con dz; i.e. $dz = a\Delta x + b\Delta y$.

Teorema: Si $z = f(x,y)$ tiene una diferencial total $dz = a\Delta x + b\Delta y$ en el punto (x,y) entonces

$$a = \frac{\partial z}{\partial x} \qquad b = \frac{\partial z}{\partial y},$$

i.e. las dos derivadas parciales existen en (x,y) y tienen los valores dados.

Dem. Si hacemos $\Delta y = \emptyset$, entonces como $\Delta z = a\Delta x + b\Delta y + \varepsilon_1\Delta x + \varepsilon_2\Delta y$, tenemos que:

$$\frac{\partial z}{\partial x} = \lim_{\Delta x\to\emptyset} \frac{\Delta z}{\Delta x} = \lim_{\Delta x\to\emptyset} \frac{\Delta x(a + \varepsilon_1)}{\Delta x} = \lim_{\Delta x\to\emptyset} (a + \varepsilon_1) = a$$

Análogamente se demuestra que $\frac{\partial z}{\partial y} = b$

por lo tanto si $z = f(x,y)$ tiene una diferencial total en (x,y) entonces

$$dz = \frac{\partial z}{\partial x}\Delta x + \frac{\partial z}{\partial y}\Delta y$$

Por razones que más tarde se explican tanto Δx como Δy se pueden reemplazar por dx, dy en la expresión para la diferencial total. Así se tiene:

$$dz = \frac{\partial z}{\partial x}\,dx + \frac{\partial z}{\partial y}\,dy$$

la cual es la forma acostumbrada de escribir la diferencial. "

Se observa que la diferencial total así definida es una función lineal en los argumentos dx, dy. Sin embargo si los estudiantes desconocen la definición y las propiedades de las transformaciones lineales, el concepto anterior puede quedar reducido para ellos a una mera expresión formal.

Contribuye a este peligro el hecho de que la notación que se usa con frecuencia puede resultar un tanto confusa para los estudiantes. No es una notación que sugiera que dz es, de hecho, una función, que además es lineal en los argumentos dx, dy. De hecho, el uso mismo de notaciones como dx, dy para designar a las variables independientes, debería acompañarse de algún señalamiento para los estudiantes pues no están familiarizados con denotar de esta forma a las variables independientes. Como dato significativo, no aparece en el capítulo de donde está tomado el texto anterior la explicación a la que hace referencia el autor sobre el cambio de notación de Δx, Δy por dx, dy.

Sin embargo, salvada esta confusión en la notación, el desarrollo anterior es útil como un razonamiento heurístico para hacer ver porqué es de esperar que la extensión, para las funciones de varias variables, del concepto de derivada de una función de $\mathbb{R}$ en $\mathbb{R}$ en un punto, es la existencia de una función lineal que aproxima a la función en una vecindad del punto en cuestión.

CAPITULO 4

ALGUNOS ELEMENTOS DE TRANSFORMACIONES LINEALES

4.0 INTRODUCCION

El porqué de presentar el material de este capítulo obedece a una razón muy sencilla: para tener una adecuada comprensión de la diferencial de una función $F : D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ se ha de manejar el concepto de transformación lineal. Sin embargo, es común que los estudiantes a este nivel no hayan llevado un curso de Algebra lineal. En ocaciones llevan un curso de Algebra lineal simultáneo al de Cálculo vectorial y suele suceder que cuando se requiere en Cálculo de los conceptos del Algebra lineal; éstos aún no se han visto o se han estudiado de una manera un tanto abstracta, de tal modo que se le dificulta al estudiante aplicar esas ideas a las necesidades específicas del Cálculo. En otros muchos casos los estudiantes no llevan ningún curso de Algebra lineal.

Con la idea de solventar cualquiera de esas dificultades, se propone incorporar al curso, el tema de las transformaciones lineales. Es importante que ese estudio no se haga de una manera abstracta sino remitiéndose a ejemplos específicos de $\mathbb{R}^n$ a $\mathbb{R}^m$, a fin de facilitar la aplicación de esas ideas al Cálculo.

Con ese objetivo en mente; qué temas hay que tocar y qué aspectos habría que resaltar es lo que constituye el contenido del presente capítulo. En la primera sección se habla sobre los preliminares de Algebra lineal que son necesarios para entender las transformaciones lineales: independencia lineal, bases, coordenadas, subespacios vectoriales. Para luego abordar el estudio de las funciones lineales.

Una idea importante es no pretender dar un curso completo de la materia, sino limitarse a hacer una síntesis de los resultados que sean de utilidad para trabajar con transformaciones lineales. El profesor es quien debe graduar en base al tiempo de que disponga, al interés y la preparación previa de los estudiantes, etc., con qué tanto detalle explicará los aspectos concretos del Algebra lineal, teniendo presente que a efectos del Cálculo lo que interesa es que los estudiantes sepan cómo operar con las transformaciones lineales y sean capaces de advertir la conveniencia de que una función arbitraria pueda ser aproximada localmente por una transformación lineal.

Como prerequisito para el material de Algebra lineal que se propone presentar; se presupone que el estudiante maneja con soltura el álgebra de matrices, incluyendo el producto de matrices y la aplicación de las matrices para resolver sistemas de ecuaciones de 1er grado, lo cual no es una condición difícil de cubrir, pues prácticamente la totalidad de los estudiantes que han de llevar Cálculo vectorial, han pasado ya por un curso de Algebra superior, donde se revisa ese material.

4.1 ESPACIOS VECTORIALES

Los estudiantes están familiarizados ya con las operaciones de suma de vectores y de multiplicación de un vector por un escalar; y las correspondientes interpretaciones geométricas para los casos en 2 y 3 dimensiones.

Asimismo, les resultan muy claras, al ser muy fáciles de demostrar, las propiedades que dichas operaciones satisfacen:

"Si $u=(u_1,u_2,\dots,u_n)$, $v=(v_1,v_2,\dots,v_n)$, $w=(w_1,w_2,\dots,w_n) \in \mathbb{R}^n$ k, $s \in \mathbb{R}$ entonces

$$(*)\left\{\begin{array}{ll}1) & u + v = v + u \\ 2) & u + (v + w) = (u + v) + w \\ 3) & u + \emptyset = \emptyset + u = u \\ 4) & u + (-u) = \emptyset \quad (\text{i.e. } u - u = \emptyset) \\ 5) & k(su) = (ks)u \\ 6) & k(u + v) = ku + kv \\ 7) & (k + s)u = ku + su \\ 8) & 1u = u \qquad (1 \in \mathbb{R})\end{array}\right.$$ "

En esta sección se pretende presentar a los estudiantes los conceptos del Algebra lineal que son de utilidad para conseguir una mejor comprensión de las transformaciones lineales, como es la independencia lineal:

"Se dice que un conjunto de vectores $\{v_1,\dots,v_k\}$ de

$\mathbb{R}^n$ es linealmente independiente si la ecuación:

(1)... $c_1v_1 + c_2v_2 +\ldots+ c_kv_k = \emptyset \qquad c_i \in \mathbb{R} \quad i = 1,2,\ldots,k$

se satisface si y sólo si $c_1 = c_2 =\ldots= c_k = \emptyset$

Si la expresión (1) se cumple no con todos los c_i's necesariamente iguales a $\emptyset$, entonces se dice que los vectores son linealmente dependientes.

Notamos que para toda matriz A de $n \times k$ cuyos vectores columna son $v_1,\ldots,v_k$ con $v_j = (a_{1j}, a_{2j},\ldots,a_{nj})$ $1 \leq j \leq k$

$$A\begin{bmatrix} c_1 \\ \cdot \\ \cdot \\ \cdot \\ c_k \end{bmatrix} = \begin{bmatrix} | & & | \\ v_1 & \ldots & v_k \\ | & & | \end{bmatrix}\begin{bmatrix} c_1 \\ \cdot \\ \cdot \\ \cdot \\ c_k \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \ldots & a_{1k} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ a_{n1} & a_{n2} \ldots & a_{nk} \end{bmatrix}\begin{bmatrix} c_1 \\ \cdot \\ \cdot \\ \cdot \\ c_k \end{bmatrix} =$$

$$= \begin{bmatrix} a_{11}c_1 + a_{12}c_2 +\ldots+ a_{1k}c_k \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ a_{n1}c_1 + a_{n2}c_2 +\ldots+ a_{nk}c_k \end{bmatrix} = c_1\begin{bmatrix} a_{11} \\ \cdot \\ \cdot \\ \cdot \\ a_{n1} \end{bmatrix} + c_2\begin{bmatrix} a_{12} \\ \cdot \\ \cdot \\ \cdot \\ a_{n2} \end{bmatrix} +\ldots+ c_k\begin{bmatrix} a_{1k} \\ \cdot \\ \cdot \\ \cdot \\ a_{nk} \end{bmatrix} =$$

$$= c_1v_1 + c_2v_2 +\ldots+ c_kv_k$$

Por lo que si consideramos a los v_i's como vectores columna, sus combinaciones lineales se pueden expresar como Ac donde $c = (c_1, c_2, \ldots, c_k)$. Por tanto, decir que $v_1, \ldots, v_k$ son linealmente independientes es lo mismo que decir que $Ac = \emptyset$, se satisface sólo para $c = \emptyset$.

En particular si $k = n$, A es una matriz cuadrada y $Ac = \emptyset$ es equivalente a n ecuaciones lineales homogéneas en n indeterminadas. Los vectores columna de A: $v_1,\ldots, v_n$ serán entonces linealmente independientes, precisamente cuando estas ecuaciones sólo tienen la solución trivial $c_1=c_2\ldots= c_n=\emptyset$; es decir, cuando $\det A \neq \emptyset$ (la matriz A es no singular).

Conclusión: n vectores $v_1, v_2, \ldots v_n$ de $\mathbb{R}^n$ son linealmente independientes si y sólo si la matriz A cuyos vectores columna son $v_1, \ldots, v_n$ es no singular."

También necesitamos el concepto de base (de $\mathbb{R}^n$):

"Se dice que k vectores $v_1, \ldots, v_k$ de $\mathbb{R}^n$ son una base de $\mathbb{R}^n$ si todo vector $v \in \mathbb{R}^n$ se puede expresar de manera única como una combinación lineal de $v_1, \ldots, v_k$, i.e. si $\forall\ v \in \mathbb{R}^n$, $v = c_1v_1 + c_2v_2 + \ldots + c_kv_k$; para elecciones únicas de los números $c_1, c_2, \ldots, c_k$. A esos números se les llama las componentes o las coordenadas del vector v con respecto a la base $\{v_1, v_2, \ldots, v_k\}$.

Ahora podemos enunciar algunas reglas concernientes a independencia lineal y bases.

(1) Los vectores $\{v_1, \ldots, v_k\}$ son linealmente dependientes si y sólo si uno de estos vectores se puede expresar como una combinación lineal de los otros.

(2) Si uno de los vectores $\{v_1, \ldots, v_k\}$ es el vector cero, entonces $\{v_1, \ldots, v_k\}$ son linealmente dependientes.

(3) Si $v_1, \ldots, v_k$ son linealmente independientes, pero $v_1, \ldots, v_k, v_{k+1}$ son linealmente dependientes, entonces v_{k+1} se expresa como una combinación lineal de $v_1, \ldots, v_k$.

(4) Si $v_1, \ldots, v_k$ son linealmente independientes y $h < k$ entonces $v_1, \ldots, v_h$ son linealmente independientes.

(5) Si $v_1, \ldots, v_k$ son linealmente independientes y $a_1v_1 + a_2v_2 + \ldots + a_kv_k = b_1v_1 + b_2v_2 + \ldots + b_kv_k$, entonces $a_1 = b_1$, $a_2 = b_2$, $\ldots$, $a_k = b_k$.

(6) Existen n vectores linealmente independientes en $\mathbb{R}^n$, por ejemplo los vectores:
$e_1=(1,\emptyset,\ldots,\emptyset)$ $e_2=(\emptyset,1,\emptyset,\ldots,\emptyset)$, $\ldots$, $e_n=(\emptyset,\emptyset,\ldots,1)$.

(7) No existen $n + 1$ vectores linealmente independientes en $\mathbb{R}^n$.

(8) Si $v_1, \ldots, v_n$ son vectores linealmente independientes en $\mathbb{R}^n$, entonces forman una base para $\mathbb{R}^n$, en particular $e_1, e_2, \ldots, e_n$ forman una base de $\mathbb{R}^n$ llamada la base canónica de $\mathbb{R}^n$.

(9) Toda base de $\mathbb{R}^n$ consiste de n vectores lineal-

mente independientes.

(1∅) Si $k < n$ y $v_1, v_2, \ldots, v_k$ son linealmente independientes, entonces existen vectores de $\mathbb{R}^n$: $v_{k+1}, \ldots, v_n$ tales que $\{v_1, v_2, \ldots, v_k, v_{k+1}, \ldots, v_n\}$ constituyen una base para $\mathbb{R}^n$."

La lista anterior proporciona un resumen de los resultados usados con más frecuencia acerca de bases e independencia lineal, para nuestros fines, dicha enumeración es más que suficiente.

Para que los alumnos se familiaricen con estos conceptos, convendría que demostraran algunos; las demostraciones en su mayoría no son difíciles y además constituyen un magnífico ejercicio para aprender a demostrar cosas.

A manera de ejemplo se muestra (7): "No existen n+1 vectores linealmente independientes en $\mathbb{R}^n$. Usaremos el hecho de que dados $b \in \mathbb{R}^n$ y A una matriz de $n \times n$, el sistema $Ax=b$ tiene solución y esta es única si y sólo si A es no singular. Supongamos que $v_1, \ldots, v_{n+1}$ son vectores linealmente independientes en $\mathbb{R}^n$. Aceptando la validez de (4); $v_1, \ldots v_n$ también son linealmente independientes. En consecuencia, la matriz A cuyas columnas son $v_1, \ldots, v_n$ es no singular. Por consiguiente la ecuación $Ac = v_{n+1}$ tiene una solución única; digamos $c^{\emptyset} \neq \emptyset$ esto es $v_{n+1} = c_1^{\emptyset} v_1 + c_2^{\emptyset} v_2 + \ldots + c_n^{\emptyset} v_n$ con $c^{\emptyset} = (c_1^{\emptyset}, c_2^{\emptyset}, \ldots, c_n^{\emptyset})$.

Pero por (1); $v_1, v_2, \ldots, v_{n+1}$ serían entonces linealmente dependientes, contrario a la suposición. Por tanto, no puede haber n+1 vectores linealmente independientes en $\mathbb{R}^n$."

Como es la base que aparecerá con más frecuencia, vale la pena destacar la conveniencia de trabajar con la base canónica de $\mathbb{R}^n$:

"La base $e_1, e_2 \ldots e_n$ mencionada en (6) se conoce como la base canónica de $\mathbb{R}^n$ (para n=3 en ocasiones se usa la notación i, j, k en vez de e_1, e_2, e_3). Es una base con la que resulta especialmente fácil de trabajar pues $\forall\ v \in \mathbb{R}^n$ tenemos que: $v=(v_1, v_2, \ldots, v_n) = v_1e_1+v_2e_2+\ldots+v_ne_n$."

Es interesante considerar el concepto de subespacio vectorial de $\mathbb{R}^n$, puesto que en su momento se hablará de espacios tangentes:

"Si F es un conjunto (no vacío) tal que $F \subseteq \mathbb{R}^n$, F se llama un subespacio vectorial si se satisface que para todos $u, v \in F$; $\lambda \in \mathbb{R}$ se tiene

$$u + v \in F$$
$$\lambda u \in F \text{ "}$$

Se puede hacer la observación de que el nombre de subespacio vectorial se sigue del hecho de que con las operaciones definidas para $\mathbb{R}^n$, F tiene estructura de espacio vectorial. Es decir, la misma estructura algebraica de $\mathbb{R}^n$ que está caracterizada por el hecho de que los elementos de F sastisfacen (*). El concepto se aclara con algunos ejemplos como los siguientes:

"Ejemplo 1. Sea F cualquier plano en $\mathbb{R}^3$ que pasa por el origen. Los puntos $(x, y, z) \in F$ satisfacen una ecuación del tipo: $ax+by+cz = 0$ con $a, b, c \in \mathbb{R}$.

Sean $u = (u_1, u_2, u_3)$, $v = (v_1, v_2, v_3) \in F$ entonces

$$au_1 + bu_2 + cu_3 = 0$$
$$av_1 + bv_2 + cv_3 = 0$$

Sumando estas ecuaciones obtenemos que:

$a(u_1+v_1) + b(u_2+v_2) + c(u_3+v_3) = 0 \quad \therefore u+v \in F$.

Sea $u = (u_1, u_2, u_3) \in F$, entonces $au_1+bu_2+cu_3 = 0$ multiplicando por $\lambda \in \mathbb{R}$ (arbitrario):

$0 = \lambda\ (au_1+bu_2+cu_3) = a\lambda u_1+b\lambda u_2+c\lambda u_3 \quad \therefore \lambda u \in F \Rightarrow F$ es un subespacio vectorial de $\mathbb{R}^3$.

Ejemplo 2. Si $F \subseteq \mathbb{R}^n$ y F es cualquier recta que pasa por el origen, entonces F es un subespacio vectorial pues los puntos $x = (x_1, x_2, \ldots, x_n)$ de F satisfacen:

$x = ta \quad a \in \mathbb{R}^n$ i.e.

$$\left\{\begin{array}{l} x_1 = ta_1 \\ x_2 = ta_2 \qquad t \in \mathbb{R} \\ \cdots\cdots \\ x_n = ta_n \end{array}\right.$$

Si $x, y \in F$ $x+y = (x_1+y_1, x_2+y_2, \ldots, x_n+y_n)$ como para $i = 1, \ldots, n$ $\quad x_i = ta_i \quad y_i = t'a_i \quad t, t' \in \mathbb{R}$

$\Rightarrow x_i+y_i = (t+t')a_i = t''a_i \quad t'' \in \mathbb{R}. \quad i = 1, 2, \ldots\ldots, n$
$\therefore x + y \in F.$

Si $x \in F$ entonces $\lambda x = (\lambda x_1, \lambda x_2, \ldots\ldots, \lambda x_n)$ como $x_i = ta_i \quad i = 1, 2, \ldots, n \Rightarrow \lambda x_i = \lambda t a_i = \tau a_i \ \tau \in \mathbb{R}$ $i = 1, 2, \ldots, n \quad \therefore \lambda x \in F$ i.e. F es un subespacio vectorial.

Un ejemplo algebraico familiar es el siguiente:

Ejemplo 3. Si tenemos un sistema de m ecuaciones con n incógnitas:

$$\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \ldots\ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots\ldots + a_{2n}x_n = b_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots\ldots + a_{mn}x_n = b_m \end{array}$$

Que se escribe en forma matricial como $Ax=b$. Se dice que un vector $s = (s_1, s_2, \ldots, s_n)$ es un vector solución del sistema si

$$\left.\begin{array}{l} x_1 = s_1 \\ x_2 = s_2 \\ \ldots\ldots \\ x_n = s_n \end{array}\right\} \text{ es una solución del sistema.}$$

Si $b = \emptyset$, el sistema se llama homogéneo. Probaremos que el conjunto de vectores solución de un sistema homogéneo es un subespacio vectorial de $\mathbb{R}^n$.

Sea F el conjunto de vectores solución del sistema $Ax = \emptyset$. Sean $v_1, v_2 \in F$ entonces $Av_1 = \emptyset$ y $Av_2 = \emptyset$ $A(v_1+v_2) = Av_1+Av_2 = \emptyset + \emptyset = \emptyset \quad \therefore \quad v_1+v_2 \in F$. Como $Av_1 = \emptyset$ si $\lambda \in \mathbb{R} \Rightarrow \lambda Av_1 = \emptyset \Rightarrow A(\lambda v_1) = \emptyset \therefore \lambda v_1 \in F$.

A F se le llama el espacio solución del sistema $Ax = \emptyset$.

Ejemplo 4. Sea $S = \{v_1, \ldots\ldots, v_k\}$ una colección de vectores de $\mathbb{R}^n$ y consideremos todas sus combinaciones lineales, i.e. el conjunto $F = \{t_1v_1 + t_2v_2 + \ldots + t_kv_k / t_i \in \mathbb{R}$ $i = 1, 2, \ldots k\}$. Claramente F es un subespacio vectorial de $\mathbb{R}^n$ llamado el subespacio vectorial generado por $\{v_1, \ldots, v_k\}$, también se dice que el conjunto de vectores $S = \{v_1, \ldots\ldots, v_k\}$ genera a F."

Se puede hacer el señalamiento de que al tener estructura de espacio vectorial, los conceptos de bases, coordenas e independencia lineal se definen de igual modo para subespacios vectoriales de $\mathbb{R}^n$. En particular, en virtud del último ejemplo, cabe la observación de que, alternativamente, una base de un subespacio vectorial $F \subseteq \mathbb{R}^n$ puede definirse como un conjunto de vectores de F que sean linealmente independientes y que generan a F.

4.2 TRANSFORMACIONES LINEALES

Con el material de la sección anterior, se está en condiciones de estudiar con cierto detalle a las transformaciones lineales.

"Definición. Una función o transformación T de $\mathbb{R}^n$ a $\mathbb{R}^m$ se dice que es lineal si tiene las siguientes propiedades: $\forall u,v \in \mathbb{R}^n$; $\lambda \in \mathbb{R}$

$$T(u + v) = T(u) + T(v)$$
$$T(\lambda u) = \lambda T(u).$$

Como una primera observación se tiene que si T es lineal, entonces T manda al vector cero de $\mathbb{R}^n$ en el vector cero de $\mathbb{R}^m$. Pues dado cualquier $v \in \mathbb{R}^n$ $\emptyset = \emptyset v$ (el de la izquierda es el cero de $\mathbb{R}^n$ y el de la derecha el cero de $\mathbb{R}$).

$$\therefore \quad T(\emptyset) = T(\emptyset v) = \emptyset T(v) = \emptyset \qquad \emptyset \in \mathbb{R}^m."$$

Se puede sugerir a los estudiantes un recurso gráfico que les ayude a asociar una imagen al concepto de transformación lineal:

"Es útil pensar en una transformación lineal $T:\mathbb{R}^n \to \mathbb{R}^m$ como se sugiere en la Figura 4.1, en la cual los vectores están representados por segmentos dirigidos desde el origen, en este caso n = 2, m = 3."

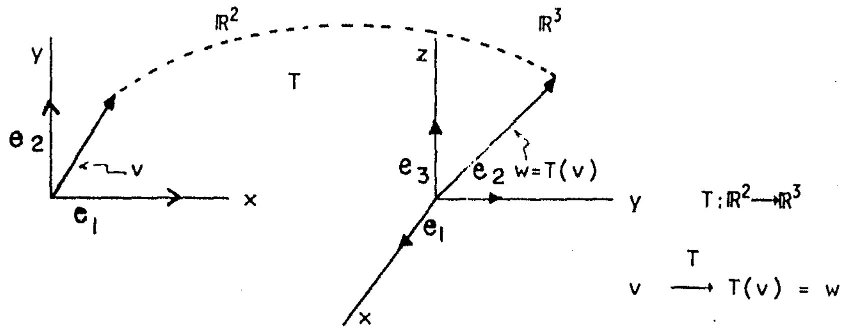


Figura 4.1 (Transformación lineal de $\mathbb{R}^2$ a $\mathbb{R}^3$)

Después de la definición, hay que presentar una selección adecuada de ejemplos que incluyan el caso en el que las funciones coordenadas están dadas de manera explícita, lo cual sirve para que los estudiantes puedan advertir que éstas tienen una forma particularmente simple. (Sobre este punto se ahonda al final del capítulo).

"Ejemplo 1. Sea $F: \mathbb{R}^2 \to \mathbb{R}^3$ dada por: $F(u) = F(x,y) = (x, x+y, x-y)$ Si $u = (1,1)$ $F(u) = (1,2,\emptyset)$
Si $u = (x_1, y_1)$ $v = (x_2, y_2)$ entonces $u+v = (x_1+x_2, y_1+y_2)$
$\Rightarrow F(u + v) = F(x_1 + x_2, y_1 + y_2) =$
$(x_1 + x_2, (x_1 + x_2) + (y_1 + y_2), (x_1 + x_2) - (y_1 + y_2)) =$
$= (x_1, x_1+y_1, x_1-y_1)+(x_2, x_2+y_2, x_2-y_2) = F(u) + F(v)$.

Si $\lambda \in \mathbb{R}$ $\lambda u = (\lambda x, \lambda y)$ por lo que:
$F(\lambda u) = F(\lambda x, \lambda y) = (\lambda x, \lambda x+\lambda y, \lambda x-\lambda y) = \lambda(x, x+y, x-y) = \lambda F(u)$
$\therefore$ F es una función lineal."

Un ejemplo particularmente importente es el de las transformaciones matriciales, pues todas las transformaciones lineales de $\mathbb{R}^n$ a $\mathbb{R}^m$ pueden reducirse a este caso.

"Ejemplo 2. Sea A una matriz de m x n y consideremos la función $T: \mathbb{R}^n \to \mathbb{R}^m$ definida mediante la ecuación:

$$T(x) = Ax \quad x \in \mathbb{R}^n.$$

Si x es una matriz de nx1 (un vector de $\mathbb{R}^n$ puesto como matriz columna); el producto Ax es una matriz de m x 1 y es por tanto un elemento de $\mathbb{R}^m$. Además T es lineal pues si u, v son matrices de n x 1; $\lambda \in \mathbb{R}$, por las propiedades de la multiplicación de matrices se tiene que:

$$T(u + v) = A(u + v) = Au + Av = T(u) + T(v)$$
$$T(\lambda u) = A(\lambda u) = \lambda Au = \lambda T(u).$$

A la transformación lineal de este ejemplo se le denomina multiplicación por A. A las transformaciones lineales de este tipo se les llama transformaciones matriciales."

Además del caso general, es bueno mostrar ejemplos específicos que ayuden a ver cómo se trabaja con las transformaciones matriciales:

"Ejemplo 3. Sea T: $\mathbb{R}^2 \to \mathbb{R}^2$ dada por $T(x) = Ax$

donde $A = \begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix}$ Calcular $T(x)$ para $x_1 = (1, 0)$; $x_2 = (0,1)$; $x_3 = (2,-1)$ y dibujar las imágenes siguiendo la idea de la Figura 4.1

$$T(x_1) = T(e_1) = \begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$$T(x_2) = T(e_2) = \begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

$$T(x_3) = T(e_3) = \begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Otra forma de calcular $T(x_3)$ hubiera sido observando que $x_3 = 2e_1 - e_2$ $\therefore T(x_3) = T(2e_1 - e_2) = 2T(e_1) - T(e_2) = 2(2,1) - (3,2) = (1,0)$. La situación se describe gráficamente en la Figura 4.2.

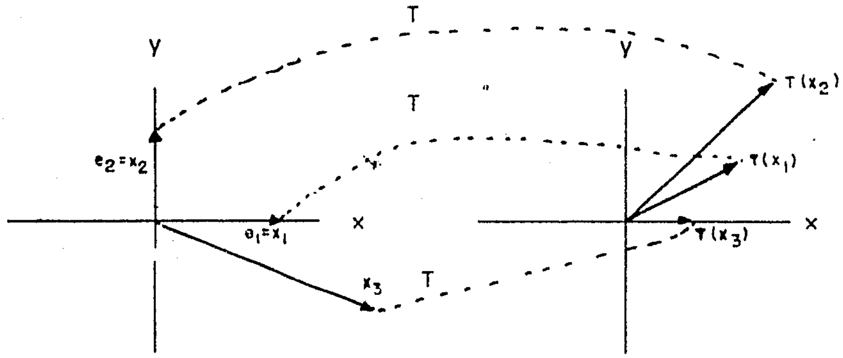


Figura 4.2

Obsérvese que en este ejemplo para $i = 1,2$; $T(e_i)$ es la i-ésima columna de la matriz A."

Se pueden presentar homotecias y rotaciones como ejemplos a través de los cuales se puede entender más acerca del contenido geométrico tanto de las transformaciones lineales como de la estructura misma de $\mathbb{R}^n$.

"Ejemplo 4. Sea $k \in \mathbb{R}$ fijo. La función $T: \mathbb{R}^n \to \mathbb{R}^n$ definida mediante $T(v) = kv \quad v \in \mathbb{R}^n$ es lineal:

$$T(u + v) = k(u + v) = ku + kv = T(u) + T(v)$$
$$T(\lambda u) = k(\lambda u) = \lambda(ku) = \lambda T(u)$$

Si $k > 1$, T es una dilatación de $\mathbb{R}^n$ y si $\emptyset < k < 1$ entonces T se denomina una contracción de $\mathbb{R}^n$. En términos geométricos una dilatación "alarga" cada vector de $\mathbb{R}^n$ por un factor k y una contracción "comprime" cada vector por un vector k (Ver Figura 4.3).

(i) Contracción de $\mathbb{R}^n$ (ii) Dilatación de $\mathbb{R}^n$

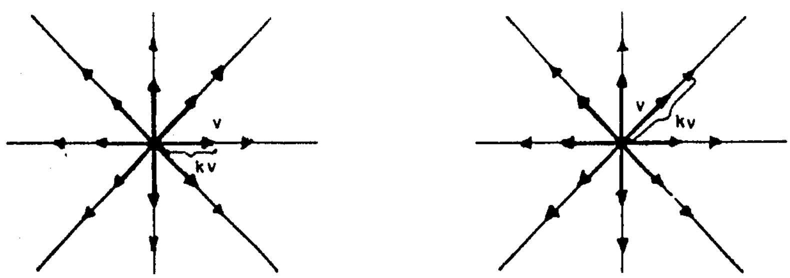


Figura 4.3

Las rotaciones tienen la ventaja adicional de que pueden verse directamente como casos particulares de las transformaciones matriciales:

"Ejemplo 5. Sea θ un ángulo fijo y sea $T: \mathbb{R}^2 \to \mathbb{R}^2$ la multiplicación por la matriz

$$A = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}$$

Si v es el vector (x,y) tenemos que:

$$T(v) = Av = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \begin{bmatrix} x\cos\theta - y\operatorname{sen}\theta \\ x\operatorname{sen}\theta + y\cos\theta \end{bmatrix}$$

Interpretando geométricamente, $T(v)$ es el vector que se obtiene al girar v un ángulo θ en sentido positivo. Para ver esta afirmación, sea φ el ángulo entre v y la dirección positiva del eje x y sea $v' = (x', y')$ el vector que resulta de girar v un ángulo θ (ver Figura 4.4). Tenemos que demostrar $T(v) = v'$.

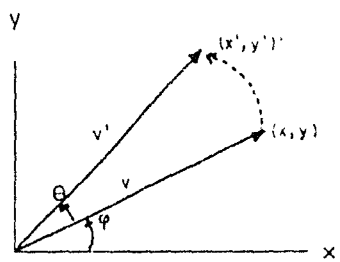


Figura 4.4

Si r denota la longitud de v entonces $x = r\cos\varphi$
$y = r\operatorname{sen}\varphi$

De igual manera, ya que v' tiene la misma longitud que v, se tiene:
$x' = r\cos(\theta + \varphi)$
$y' = r\operatorname{sen}(\theta + \varphi)$.

Por tanto:

$$v' = \begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} r\cos(\theta+\varphi) \\ r\operatorname{sen}(\theta+\varphi) \end{bmatrix} = \begin{bmatrix} r(\cos\theta\cos\varphi - \operatorname{sen}\theta\operatorname{sen}\varphi) \\ r(\operatorname{sen}\theta\cos\varphi + \cos\theta\operatorname{sen}\varphi) \end{bmatrix} = \begin{bmatrix} x\cos\theta - y\operatorname{sen}\theta \\ x\operatorname{sen}\theta + y\cos\theta \end{bmatrix}$$

$$= \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = Av = T(v)$$

A esta transformación se le llama rotación de $\mathbb{R}^2$ de un ángulo θ."

Dos ejemplos muy familiares para los estudiantes son la función cero y la identidad: funciones especialmente sencillas que son lineales.

"Ejemplo 6. Sea $T: \mathbb{R}^n \to \mathbb{R}^m$ definida como $T(x) = \emptyset$ $\forall\ x \in \mathbb{R}^n$. Se suele llamar a T la transformación cero; es lineal ya que: $T(u + v) = \emptyset = \emptyset + \emptyset = T(u) + T(v)$.
$T(ku) = \emptyset = k\emptyset = kT(u)$.

Ejemplo 7. La función $T: \mathbb{R}^n \to \mathbb{R}^n$ definida como $T(v) = v\ \ \forall\ v \in \mathbb{R}^n$ conocida como la función o transformación identidad en $\mathbb{R}^n$. Es lineal pues $\forall\ x,y \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$: $T(x + y) = x+y = T(x) + T(y)$; $T(\lambda x) = \lambda x = \lambda T(x)$. "

Sin embargo, conviene mencionar, por contraposición al ejemplo 6 que las funciones constantes - salvo la función cero - no son funciones lineales:

"Ejemplo 8. Sea $F: \mathbb{R}^n \to \mathbb{R}^m$ definida como $F(x) = a$ $\forall\ x \in \mathbb{R}^n$, $a \in \mathbb{R}^m - \{\emptyset\}$. $F(\lambda x) = a$ mientras que $\lambda F(x) = \lambda a$. Si $\lambda \neq 1 \Rightarrow F(\lambda x) \neq \lambda F(x)$ y $F(x+y) = a \neq 2a = a + a = F(x) + F(y)$. $\therefore$ F <u>no</u> es lineal.

Es decir, las funciones constantes - para toda constante diferente de cero - no son funciones lineales."

Como es una situación frecuente en la práctica, cabe mencionar - explicándolo convenientemente - que el dominio de una transformación lineal puede no ser todo $\mathbb{R}^n$ sino un subespacio vectorial de éste, y a su vez, el contradominio ser un subespacio vectorial de $\mathbb{R}^m$. Al tener estos conjuntos estructura de espacio vectorial está definida una suma y un producto por escalares, que son los elementos necesarios para dar sentido a la definición de transformación lineal.

4.3 PROPIEDADES DE LAS TRANSFORMACIONES LINEALES

Después de haber definido y dado ejemplos de transformaciones lineales, se pueden estudiar algunas propiedades específicas de dichas funciones, que proporcionen a los estudiantes los elementos necesarios para operar adecuadamente con ellos, como son los conceptos de núcleo e imagen:

"Si $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ es una transformación lineal, entonces el núcleo de T, denotado KerT, es el conjunto que consta de todos los vectores en $\mathbb{R}^n$ tales que T les asocia el $\emptyset$ de $\mathbb{R}^m$, es decir, $KerT = \{v \in \mathbb{R}^n: T(v) = \emptyset\}$. Por la observación de que si T es lineal $T(\emptyset) = \emptyset$, el $\emptyset$ de $\mathbb{R}^n$ está en KerT.

La imagen de T, que se denota por ImT, es el conjunto de todos los vectores en $\mathbb{R}^m$ que son imagen bajo T de al menos un vector en $\mathbb{R}^n$, i.e. $ImT = \{y \in \mathbb{R}^m$; para los que existe $x \in \mathbb{R}^n$ tal que $T(x) = y\}$, equivalentemente ImT es el conjunto de todas las imágenes bajo T de los puntos de $\mathbb{R}^n$, i.e. $ImT = \{T(x): x \in \mathbb{R}^n\} = T(\mathbb{R}^n)$.

Dichos conceptos de KerT e ImT se definen de la misma forma, mutatis - mutandis si en lugar de $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ tuviéramos: $T: E \subseteq \mathbb{R}^n \rightarrow F \subseteq \mathbb{R}^m$ con E, F subespacios vectoriales.

Ejemplo 1. Sea $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ la transformación cero. Como T mapea a todos los vectores en $\emptyset$ => $KerT = \mathbb{R}^n$. Como $\emptyset$ es la única imagen posible bajo T => $ImT = \{\emptyset\}$.

Ejemplo 2. Sea $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ la multiplicación por:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots\cdots & a_{1n} \\ a_{21} & a_{22} & \cdots\cdots & a_{2n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ a_{m1} & a_{m2} & \cdots\cdots & a_{mn} \end{bmatrix}$$

El núcleo de T consta de todos los vectores $x = (x_1, x_2, \ldots, x_n)$ que son vectores solución del sistema homogéneo:

$$A \begin{bmatrix} x_1 \\ \cdot \\ \cdot \\ x_n \end{bmatrix} = \begin{bmatrix} \emptyset \\ \cdot \\ \cdot \\ \emptyset \end{bmatrix}$$

$\therefore$ KerT es el espacio solución del sistema $Ax = \emptyset$ (cfr. Ejemplo 3 de la sección 4.1).

La imagen de T consta de todos los vectores $b = (b_1, b_2, \ldots, b_m)$ tales que el sistema $Ax = b$ tiene al menos una solución.

Dada $T: \mathbb{R}^n \to \mathbb{R}^m$ una transformación lineal, entonces ,
(i) $\text{KerT} \subseteq \mathbb{R}^n$ es un subespacio vectorial de $\mathbb{R}^n$.
(ii) $\text{ImT} \subseteq \mathbb{R}^m$ es un subespacio vectorial de $\mathbb{R}^m$.

Demostración

(i) Sean $v_1, v_2 \in \text{KerT}$; $\lambda \in \mathbb{R}$
Por demostrar
$$\begin{cases} v_1 + v_2 \in \text{KerT} \\ \lambda v_1 \in \text{KerT} \end{cases}$$

$$\left.\begin{array}{l} v_1 \in \text{KerT} \Rightarrow T(v_1) = \emptyset \\ v_2 \in \text{KerT} \Rightarrow T(v_2) = \emptyset \end{array}\right\}$$

Ahora bien $T(v_1+v_2) = T(v_1) + T(v_2) = \emptyset + \emptyset = \emptyset$
$\therefore$ $v_1 + v_2 \in \text{KerT}$.
$T(\lambda v_1) = \lambda T(v_1) = \lambda\emptyset = \emptyset \Rightarrow \lambda v_1 \in \text{KerT}$.

(ii) Sean $w_1, w_2 \in \text{ImT}$; $\lambda \in \mathbb{R}$
Por demostrar
$$\begin{cases} w_1 + w_2 \in \text{ImT} \\ \lambda w_1 \in \text{ImT} \end{cases}$$

Como $w_1 \in \text{ImT} \Rightarrow \exists\, v_1 \in \mathbb{R}^n \ni T(v_1) = w_1$
Como $w_2 \in \text{ImT} \Rightarrow \exists\, v_2 \in \mathbb{R}^n \ni T(v_2) = w_2$

Sean $\begin{cases} v = v_1 + v_2 \\ u = \lambda v_1 \end{cases}$

Entonces

$T(v) = T(v_1+v_2) = T(v_1)+T(v_2) = w_1+w_2 \Rightarrow w_1+w_2 \in \text{ImT}$

$T(u) = T(\lambda v_1) = \lambda T(v_1) = \lambda w_1 \Rightarrow \lambda w_1 \in ImT$."

Conviene destacar, que con los conceptos de núcleo e imagen se facilita estudiar las propiedades de inyectividad y suprayectividad para las funciones lineales:

"Dada $T: \mathbb{R}^n \to \mathbb{R}^m$ una transformación lineal, tiene interés preguntarse si T es inyectiva o suprayectiva. Recordando que T es inyectiva ó uno - a - uno si $x_1 \neq x_2$, $x_1, x_2 \in \mathbb{R}^n$ entonces $T(x_1) \neq T(x_2)$ y T es suprayectiva o sobre si $ImT = \mathbb{R}^m$.

De nuestra experiencia previa en el Cálculo de una variable, recordamos que en ocasiones no es fácil saber si dada una cierta función $f: D \subseteq \mathbb{R} \to \mathbb{R}$ ésta es inyectiva o suprayectiva. En el caso de las transformaciones lineales este problema se simplifica bastante, por ejemplo una transformación lineal $T: \mathbb{R}^n \to \mathbb{R}^m$ es inyectiva si y sólo si $KerT = \{\emptyset\}$ $(\emptyset \in \mathbb{R}^n)$. Pues supongamos que T es inyectiva y $x \in KerT$ entonces $T(x) = \emptyset = T(\emptyset)$ (la última igualdad se da por ser T lineal) y por ser T inyectiva $x = \emptyset \therefore KerT = \{\emptyset\}$. Inversamente, si $KerT=\{\emptyset\}$ supongamos que $Tx = Ty \Rightarrow Tx - Ty = \emptyset \Rightarrow T(x - y) = \emptyset \therefore x - y \in KerT \Rightarrow x-y = \emptyset \Rightarrow x = y$.

Ejemplo 3. Consideremos la siguiente transformación lineal $T: \mathbb{R}^3 \to \mathbb{R}^2$:

$$T(x_1, x_2, x_3) = \begin{bmatrix} 1 & -1 & \emptyset \\ \emptyset & 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 - x_2 \\ x_2 - x_3 \end{bmatrix}$$

Para hallar el núcleo debemos resolver
$T(x_1,x_2,x_3) = (\emptyset,\emptyset) \Leftrightarrow (x_1-x_2, x_2-x_3) = (\emptyset,\emptyset)$

$$\begin{cases} x_1 - x_2 = \emptyset \Leftrightarrow x_1 = x_2 \\ x_2 - x_3 = \emptyset \Leftrightarrow x_2 = x_3 \end{cases}$$

$\therefore$ las soluciones están dadas por: $\{(x_1,x_2,x_3)\in\mathbb{R}^3/x_1=x_2=x_3\} = \{t(1,1,1)/t\in\mathbb{R}\}$ son pues los puntos de la recta en $\mathbb{R}^3$ que pasa por el origen, en la dirección del vector (1,1,1). Como ya vimos, este lugar geométrico es un subespacio vectorial de $\mathbb{R}^3$.

Como $\text{KerT} \neq \{\emptyset\} \Rightarrow$ T no es inyectiva. Sin embargo $\text{ImT} = \mathbb{R}^2$ ya que las ecuaciones:

$$x_1 - x_2 = y_1$$
$$x_2 - x_3 = y_2$$

siempre se pueden resolver para cualquier elección de $(y_1, y_2) \in \mathbb{R}^2$, por ejemplo si hacemos $x_2 = \emptyset$ entonces $x_1 = y_1$, $x_3 = -y_2$ $\therefore$ T es suprayectiva.

Ejemplo 4. Sea $T: \mathbb{R}^n \to \mathbb{R}^n$ dada por $T(x) = x$; la transformación identidad. T es inyectiva, pues si $T(x)=T(y)$ por un lado $T(x-y) = T(x)-T(y) = \emptyset$ y por otro $T(x-y) = x-v$ $\therefore$ $x-y = \emptyset \Rightarrow x = y$. T es suprayectiva, pues dado $y \in \mathbb{R}^n$ $T(y) = y$. T es uno a uno y sobre es decir T es biyectiva.

En Algebra lineal, cuando una transformación lineal T es biyectiva se dice que es un isomorfismo."

Dos resultados ponen de relieve la facilidad operativa de las transformaciones lineales: 1) el hecho de que basta conocer los valores de estas funciones en los elementos de una base de $\mathbb{R}^n$, para saber el valor de la función en cualquier elemento de $\mathbb{R}^n$. 2) Que siempre es posible obtener una transformación lineal que mande a los elementos de una base de $\mathbb{R}^n$ en vectores de $\mathbb{R}^m$ preasignados:

"Si dos funciones lineales de $\mathbb{R}^n$ en $\mathbb{R}^m$ coinciden en n vectores linealmente independientes, entonces son iguales.

Dem. Sean: $T: \mathbb{R}^n \to \mathbb{R}^m$, $F: \mathbb{R}^n \to \mathbb{R}^m$ transformaciones lineales. Sean $\{v_1, \ldots, v_n\} \subseteq \mathbb{R}^n$ n vectores linealmente independientes $\Rightarrow$ constituyen una base de $\mathbb{R}^n$.

Si $T(v_1) = F(v_1), \ldots\ldots, T(v_n) = F(v_n)$, tenemos que probar que $\forall\ v \in \mathbb{R}^n$: $T(v) = F(v)$. Sea $v \in \mathbb{R}^n$, v puede expresarse entonces como: $v = c_1v_1+c_2v_2+\ldots+c_nv_n$ para escalares c_i únicos $i = 1,2,\ldots,n$ $\therefore$ $T(v) = T(c_1v_1+c_2v_2+\ldots+c_nv_n) = c_1T(v_1)+c_2T(v_2)+\ldots+c_nT(v_n) = c_1F(v_1)+c_2F(v_2)+\ldots+c_nF(v_n) = F(c_1v_1+c_2v_2+\ldots+c_nv_n) = F(v)$.

Por tanto, con conocer los valores que una transformación lineal de $\mathbb{R}^n$ en $\mathbb{R}^m$ toma en n puntos convenientemente elegidos (formen una base, es decir, sean linealmente independientes), ésta queda completamente definida para todo punto de $\mathbb{R}^n$.

Cabría la siguiente pregunta ¿es posible encontrar una transformación lineal que mande a los vectores de una base de $\mathbb{R}^n$ en puntos (o vectores) de $\mathbb{R}^m$ previamente determinados? La respuesta es afirmativa y por el resultado anterior dicha transformación es única.

Teorema. Si $\{v_1, \ldots, v_n\}$ es una base de $\mathbb{R}^n$ y si $\{w_1, \ldots, w_n\}$ son puntos arbitrarios (no necesariamente distintos) de $\mathbb{R}^m$, entonces es posible obtener una transformación lineal $T: \mathbb{R}^n \to \mathbb{R}^m$ tal que: $T(v_1) = w_1$, $T(v_2) = w_2, \ldots\ldots, T(v_n) = w_n$.

Dem. Sea $x \in \mathbb{R}^n \Rightarrow x = c_1v_1+c_2v_2+\ldots\ldots+c_nv_n$ donde los escalares c_i son únicos. Definimos $T: \mathbb{R}^n \to \mathbb{R}^m$ como $T(x) = c_1w_1+c_2w_2+\ldots\ldots+c_nw_n$ ....(a).

Que T es lineal es inmediato pues dados $x, y \in \mathbb{R}^n$ se tiene $x = c_1v_1+c_2v_2+\ldots\ldots+c_nv_n$ $y = d_1v_1+d_2v_2+\ldots\ldots+d_nv_n$ donde los escalares c_i y los d_i son únicos

$x + y = (c_1+d_1)v_1 + (c_2+d_2)v_2 + \ldots\ldots\ldots\ldots + (c_n+d_n)v_n$

$\therefore T(x+y) = (c_1+d_1)w_1 + (c_2+d_2)w_2 + \ldots\ldots\ldots\ldots + (c_n+d_n)w_n =$
$c_1w_1+c_2w_2+\ldots\ldots+c_nw_n+d_1w_1+d_2w_2+\ldots\ldots d_nw_n = T(x) + T(y)$ y si

$\lambda \in \mathbb{R}$ $\lambda x = \lambda c_1v_1 + \lambda c_2v_2 + \ldots\ldots + \lambda c_nv_n$

$T(\lambda x) = \lambda c_1w_1+\lambda c_2w_2+\ldots\ldots+\lambda c_nw_n = \lambda(T(x))$.

Además si $x = v_1 \Rightarrow c_1 = 1$ $c_2=c_3=\ldots\ldots=c_n=0$ $\therefore T(v_1) = w_1$.

De igual manera si $x=v_i \Rightarrow c_i=1$ $c_j=0$ $j=1\ldots\ldots n$, $j \neq i$ $\therefore T(v_i) = w_i$ $i = 1, \ldots\ldots, n$.

Ejemplo 5. Consideremos la base $S = \{v_1, v_2, v_3\}$ de $\mathbb{R}^3$ donde $v_1 = (1,1,1)$, $v_2 = (1,1,0)$, $v_3 = (1,0,0)$. Deseamos encontrar una transformación lineal $T: \mathbb{R}^3 \to \mathbb{R}^2$ tal que: $T(v_1) = (1,0)$, $T(v_2) = (2,-1)$, $T(v_3) = (4,3)$ y calcular $T(u)$ para $u = (2, -3, 5)$.

Sea $v = (x,y,z) \in \mathbb{R}^3$, entonces la expresión de v como combinación lineal de los vectores de S es: $v=(x,y,z)= zv_1 + (y-z)v_2 + (x-y)v_3$. Si $w_1 = T(v_1) = (1,0)$; $w_2 = T(v_2) = (2,-1)$; $w_3 = T(v_3) = (4,3)$ resulta de (a) que:

$T(v) = zw_1+(y-z)w_2+(x-y)w_3 = z(1,0)+(y-z)(2,-1)+(x-y)(4,3)$
$= (z+2(y-z)+4(x-y), z-y+3(x-y)) = (4x-2y-z, 3x-4y+z)$

En particular si $u = (2,-3,5)$ $T(2,-3,5) =$
$= (4(2)-2(-3)-5,\ 3(2)-4(-3)+5) = (8+6-5, 6+12+5) = (9,23)$."

Sin demostrarlo, porque implicaría desviarse un tanto de los objetivos centrales de un curso de Cálculo vectorial, se podría presentar también el Teorema de la Dimensión, como un resultado útil para trabajar con funciones lineales, pues establece una relación entre el núcleo y la imagen de una función lineal. Por ser especialmente importante, señalando el caso cuando el dominio y el contradominio son espacios de la misma dimensión.

"Teorema. Sea $T: E \subseteq \mathbb{R}^n \to \mathbb{R}^m$ una transformación lineal. E es un subespacio vectorial de $\mathbb{R}^n$, entonces

$$\dim(\mathrm{Im}T) + \dim(\mathrm{Ker}T) = \dim E.$$

En particular si $T: \mathbb{R}^n \to \mathbb{R}^n$, entonces por el teorema anterior las siguientes afirmaciones son equivalentes:

(a) $\mathrm{Ker}T = \{\emptyset\}$
(b) $\mathrm{Im}T = \mathbb{R}^n$
(c) T es un isomorfismo.

Dem. a) => b) si $\mathrm{Ker}T = \{\emptyset\}$ entonces $\dim(\mathrm{Ker}T) = \emptyset$ y en el teorema tenemos que $\dim(\mathrm{Im}T) = \dim \mathbb{R}^n = n \Rightarrow \mathrm{Im}T = \mathbb{R}^n$.
b) => c) si $\mathrm{Im}T = \mathbb{R}^n$; es decir T es sobre => (del teorema) $\dim(\mathrm{Im}T) = n$ y $\therefore n + \dim(\mathrm{Ker}T) = n \Rightarrow \dim(\mathrm{Ker}T) = \emptyset$ $\therefore \mathrm{Ker}T = \{\emptyset\}$ => T es inyectiva y por hipótesis era suprayectiva, por tanto es biyectiva, i.e. T es un isomorfismo.
c) => a) si T es isomorfismo en particular es inyectiva => $\mathrm{Ker}T = \{\emptyset\}$.

En otras palabras si $T: \mathbb{R}^n \to \mathbb{R}^n$ es una transformación lineal inyectiva, necesariamente es suprayectiva y vice-versa."

4.4 FORMA DE LAS TRANSFORMACIONES LINEALES DE $\mathbb{R}^n$ A $\mathbb{R}^m$.

Con el material de las secciones anteriores, el estudiante ha podido darse cuenta de que las transformaciones lineales son funciones con las qae resulta relativamente fácil trabajar en virtud de sus propiedades: basta conocer los valores en los vectores de una base y de esta manera quedan definidas en todo su dominio. Siempre se puede construir una transformación lineal que mande a los elementos de una base cualquiera de $\mathbb{R}^n$ en los puntos de $\mathbb{R}^m$ que se desee. Hay una manera sencilla de saber si son inyectivas (viendo que el núcleo esté compuesto sólo del vector $\emptyset$). Y en el caso particular de que la transformación lineal vaya de $\mathbb{R}^n$ en $\mathbb{R}^n$ es inyectiva si y sólo si es suprayectiva.

Lo anterior resulta interesante porque además de que el estudiante va aprendiendo a trabajar con las transformaciones lineales, está en condiciones de valorar más la conveniencia de tener una función lineal que aproxime localmente a una función dada. De igual manera podrá asociar a un nuevo concepto - la diferenciabilidad - un objeto que para él ya resultará familiar -las transformaciones lineales -. De esta forma el estudiante se sentirá en confianza para abordar el estudio de la diferenciabilidad, lo que facilitará el aprendizaje.

Con el fin de reforzar el convencimiento en los alumnos de que las funciones lineales son de una forma algebraica simple, y en general, funciones con las que resulta conveniente trabajar, además de proporcionar una manera operativa muy eficiente de trabajar con las transformaciones lineales, se propone presentar el resultado de que toda transformación lineal puede verse como una transformación matricial y determinar la matriz asociada a una transformación lineal. Para conseguir este objetivo se podrían utilizar razonamientos como el siguiente:

"En los ejemplos que hemos analizado, dada una transformación lineal de $\mathbb{R}^n$ en $\mathbb{R}^m$ sus funciones coordenadas eran de una forma especialmente simple: eran siempre polimonios de 1er. grado en las variables $x_1, x_2, \ldots, x_n$ (las coordenadas del punto $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$); cabe la pregunta ¿siempre es así? la respuesta es sí, como demuestra el siguiente teorema:

Una transformación lineal L de $\mathbb{R}^n$ a $\mathbb{R}^m$ tiene la siguiente forma coordenada:

$$
(1)\quad \begin{cases} y_1 = f_1(x_1,\ldots,x_n) = a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n \\ y_2 = f_2(x_1,\ldots,x_n) = a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ y_m = f_m(x_1,\ldots,x_n) = a_{m1}x_1+a_{m2}x_2\ \ldots+a_{mn}x_n \end{cases}
$$

donde los coeficiente $a_{ij} \in \mathbb{R}$."

Nota: En esta sección se usa la L para denotar a las transformaciones lineales en lugar de la T, por ser una notación muy frecuente con la que también debe familiarizarse el estudiante.

"Dem. Sea $\{e_1,\ldots,e_n\}$ la base canónica de $\mathbb{R}^n$, por lo tanto $x=(x_1,x_2,\ldots,x_n) \in \mathbb{R}^n$ se escribe como $x=x_1e_1+x_2e_2+\ldots+x_ne_n \therefore L(x) = L(x_1e_1+\ldots+x_ne_n) = x_1L(e_1)+x_2L(e_2)+\ldots+x_nL(e_n)$. Como $L(e_k) \in \mathbb{R}^m$ para $k=1,2,\ldots n$ $\Rightarrow L(e_k) = (a_{1k}, a_{2k},\ldots,a_{mk})$.

Sea $y = L(x)$ con $y = (y_1,y_2,\ldots,y_m)$ entonces $y = (y_1,y_2,\ldots,y_m) = L(x) = x_1L(e_1)+x_2L(e_2)+\ldots+x_nL(e_n) =$
$= x_1(a_{11},a_{21},\ldots,a_{m1})+x_2(a_{12},a_{22},\ldots,a_{m2})+\ldots\ldots$
$+x_n(a_{1n},a_{2n},\ldots,a_{mn}) = (a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n,\ a_{21}x_1+a_{22}x_2+\ldots$
$\ldots+a_{2n}x_n,\ldots,a_{m1}x_1+a_{m2}x_2+\ldots+a_{mn}x_n)$ que es efectivamente la expresión (1).

A su vez, otra manera de expresar (1) es:

$$
\begin{bmatrix} y_1 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ y_m \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13}\ldots & a_{1n} \\ a_{21} & a_{22} & a_{23}\ldots & a_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ a_{m1} & a_{m2} & a_{m3}\ldots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix}
$$

Que se escribe en forma más compacta como

$$y = Ax, \quad \text{donde } A = (a_{ij}).$$

Por lo que dada la matriz $A = (a_{ij})$ de los coeficientes, la transformación lineal L descrita en forma coordenada por (1), queda completamente determinada.

Vemos que las columnas de la matriz A son los vectores de $\mathbb{R}^m$: $L(e_k)$ $k = 1,2,\ldots,n$.

Hemos probado entonces que toda transformación lineal L de $\mathbb{R}^n$ en $\mathbb{R}^m$ puede verse como una transformación matricial: la multiplicación por la matriz A de mxn cuyas columnas son los vectores de $\mathbb{R}^m$ dados por $L(e_1)$, $L(e_2)$, ..., $L(e_n)$, es decir: $\forall\ x \in \mathbb{R}^n \quad L(x) = Ax$.

Decimos que A es la matriz que representa a la transformación lineal L o que A es la matriz de L, con respecto a la bases canónicas en $\mathbb{R}^n$ y $\mathbb{R}^m$. Se puede repetir el razonamiento si hubiéramos elegido otras bases para el dominio y el contradominio, obteniendo igualmente que toda función lineal de $\mathbb{R}^n$ a $\mathbb{R}^m$ tiene asociada una matriz A' que depende de la elección de las bases."

Se puede mencionar que, en el Cálculo, habitualmente se trabaja con la base canónica en el dominio y en el contradominio y por ello se habla simplemente de la matriz asociada a una transformación lineal.

"Ejemplo 1. Encontrar la matriz que representa a la transformación lineal $L: \mathbb{R}^3 \to \mathbb{R}^4$ definida por $L(x_1,x_2,x_3)= (x_1+x_2,x_1-x_2,x_3,x_1)$, o en forma de vectores columna

$$L\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1+x_2 \\ x_1-x_2 \\ x_3 \\ x_1 \end{bmatrix}$$

$$\left.\begin{array}{l} L(e_1) = L(1,\emptyset,\emptyset) = (1,1,\emptyset,1) \\ L(e_2) = L(\emptyset,1,\emptyset) = (1,-1,\emptyset,\emptyset) \\ L(e_3) = L(\emptyset,\emptyset,1) = (\emptyset,\emptyset,1,\emptyset) \end{array}\right\}$$

=> Si escribimos la matriz A cuyos vectores columna son $L(e_1)$, $L(e_2)$, $L(e_3)$ resulta:

$$A = \begin{bmatrix} | & | & | \\ L(e_1) & L(e_2) & L(e_3) \\ | & | & | \end{bmatrix} = \begin{bmatrix} 1 & 1 & \emptyset \\ 1 & -1 & \emptyset \\ \emptyset & \emptyset & 1 \\ 1 & \emptyset & \emptyset \end{bmatrix}$$

Como comprobación veamos que:

$$Ax = \begin{bmatrix} 1 & 1 & \emptyset \\ 1 & -1 & \emptyset \\ \emptyset & \emptyset & 1 \\ 1 & \emptyset & \emptyset \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1+x_2 \\ x_1-x_2 \\ x_3 \\ x_1 \end{bmatrix} = L\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = L(x)."$$

Como nos interesará en su momento estudiar la diferencial de funciones de $\mathbb{R}$ a $\mathbb{R}^n$ y de $\mathbb{R}^n$ a $\mathbb{R}$, es ahora un buen momento para ver qué forma tienen los transformaciones lineales para esos casos:

"Estudiemos cómo son las transformaciones lineales de $\mathbb{R}^n$ a $\mathbb{R}$, es decir funciones real valuadas de varias variables que además sean lineales. Sea $L: \mathbb{R}^n \to \mathbb{R}$ lineal.

Dado $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$
$L(x) = L(x_1e_1+x_2e_2+\ldots+x_ne_n) = x_1L(e_1)+x_2L(e_2)+\ldots+x_nL(e_n)$
como para $i = 1,2,\ldots,n$ $L(e_i) \in \mathbb{R}$ podemos denotar a $L(e_i) = a_i$ y obtenemos:

$$(2)\ldots.\quad L(x) = a_1x_1+a_2x_2+\ldots+a_nx_n$$

y la matriz A que representa a L es una matriz de un renglón por n columnas, a saber: $A = (L(e_1)\ L(e_2)\ldots L(e_n)) = (a_1\ \ a_2\ldots a_n)$.

Podemos reescribir (2) como $L(x) = a \cdot x$ (donde . denota el producto interior usual de $\mathbb{R}^n$ y $a = (a_1, a_2, \ldots, a_n) \in \mathbb{R}^n$ es el vector que se obtiene al considerar a la matriz A de $1 \times n$ como vector de $\mathbb{R}^n$) i.e. $L(x) = Ax = a \cdot x$."

Este último hecho interesa destacarlo, pues el vector que se llamó a, cuando dicha transformación lineal sea la diferencial en un punto de una función f de $\mathbb{R}^n$ a $\mathbb{R}$ será justamente el vector gradiente de f en el punto.

"Se tiene entonces que toda función lineal L de $\mathbb{R}^n$ en $\mathbb{R}$ es de la forma $L(x) = a \cdot x$ donde $a = (a_1, a_2, \ldots, a_n)$ es un vector asociado a L. Dada L, dicho vector a queda completamente determinado (para $i = 1,2,\ldots,n$ $a_i = L(e_i)$). Inversamente todo vector v de $\mathbb{R}^n$ determina una transformación lineal L de $\mathbb{R}^n$ en $\mathbb{R}$ definida como $L(x) = v \cdot x$."

Vale la pena reparar especialmente en los casos que pueden ser dibujados:

"Si nos fijamos en los casos $n = 1$, $n = 2$ observamos lo siguiente: Sea $L: \mathbb{R} \to \mathbb{R}$ lineal. Como $x \in \mathbb{R}$ puede escribirse como $x = x \cdot 1$ $1 \in \mathbb{R}$ – aquí $\{1\}$ es base para $\mathbb{R}$ – $\Rightarrow L(x) = xL(1) = ax$ si $a = L(1)$.

La gráfica de L que está en $\mathbb{R}^2$ es una recta que pasa por el origen. (cfr. capítulo 1). (Obsérvese que dicho lugar geométrico es un subespacio vectorial de $\mathbb{R}^2$).

Si $L: \mathbb{R}^2 \to \mathbb{R}$ es lineal, entonces existe $v = (a,b) \in \mathbb{R}^2$ tal que para todo $p = (x,y) \in \mathbb{R}^2$ $L(p) = v \cdot p$
$L(p) = (a,b) \cdot (x,y) = ax + by$.

La gráfica (en $\mathbb{R}^3$) de dicha función $L(x,y)=ax+by$ tiene por ecuación $z = ax+by$ y por tanto corresponde a un plano que pasa por el origen. (Obsérvese que la gráfica es un subespacio vectorial de $\mathbb{R}^3$)."

De ser posible, porque el tiempo lo permita, porque los estudiantes tengan una cierta capacidad para la abstracción, se podrían generalizar las ideas anteriores y hablar de hiperplanos:

"En general a las soluciones de una ecuación $a \cdot x = c$ $a, x \in \mathbb{R}^n$, $c \in \mathbb{R}$ que corresponde al conjunto $H_c = \{x \in \mathbb{R}^n \,/\, a \cdot x = c,\ a \neq \emptyset\}$ se le llama un hiperplano (en $\mathbb{R}^n$).

En $n = 1$ $H_c = \{x \in \mathbb{R} / ax = c \ a \neq \emptyset\}$ es un punto ($x = c/a$). En $n = 2$ $H_c = \{x \in \mathbb{R}^2 / a \cdot x = c\} = \{x \in \mathbb{R}^2 / ax+by = c\}$ corresponde geométricamente a una recta en $\mathbb{R}^2$. En $n = 3$ $H_c = \{x \in \mathbb{R}^3 / a \cdot x = c\} = \{x \in \mathbb{R}^3 / ax+by+dz = c\}$ que corresponde a un plano en $\mathbb{R}^3$.

Si $c = \emptyset$, entonces un hiperplano $H_\emptyset$ en $\mathbb{R}^n$ es un subespacio vectorial de $\mathbb{R}^n$ de dimensión $n - 1$.

Dem. $H_\emptyset = \{x \in \mathbb{R}^n \,/\, a \cdot x = \emptyset \quad a \in \mathbb{R}^n\}$; $H_\emptyset \subseteq \mathbb{R}^n$

Por demostrar: $\begin{cases} x, y \in H_\emptyset \Rightarrow x+y \in H_\emptyset \\ x \in H_\emptyset,\ \lambda \in \mathbb{R} \Rightarrow \lambda x \in H_\emptyset \end{cases}$

Sean $x, y \in H_\emptyset \Rightarrow a \cdot x = \emptyset,\ a \cdot y = \emptyset$. Entonces $a \cdot (x+y) = a \cdot x + a \cdot y = \emptyset + \emptyset = \emptyset \quad \therefore\ x+y \in H_\emptyset$
$a \cdot (\lambda x) = \lambda(a \cdot x) = \lambda\emptyset = \emptyset \quad \therefore\ \lambda x \in H_\emptyset$.

Para poder probar que $\dim H_\emptyset = n - 1$. Podemos pensar en la transformación lineal $L: \mathbb{R}^n \to \mathbb{R}$ dada por $L(x) = a \cdot x$. En este caso $H_\emptyset = \operatorname{Ker}L \ \therefore \dim H_\emptyset = \dim(\operatorname{Ker}L)$ y sabemos

$\dim(\mathrm{Im}L) + \dim(\mathrm{Ker}L) = \dim(\mathbb{R}^n) = n$. Lo que nos gustaría probar es que $\dim(\mathrm{Im}L) = 1$, i.e. $\mathrm{Im}L = \mathbb{R}$ (L es suprayectiva). Para eso debemos probar que $\forall\ c \in \mathbb{R}$, $\exists\ x = x(c) \in \mathbb{R}^n \ni L(x) = a \cdot x = c$. Como $a \neq \emptyset$ $\Rightarrow \exists\ i \in \{1,\dots,n\} \ni a_i \neq \emptyset$ $(a = (a_1,a_2,\dots,a_n))$.

Dado $c \in \mathbb{R}$ escogemos $x = (\emptyset,\emptyset,\dots,c/ai,\dots,\emptyset)$ con c/ai en el i-ésimo lugar. Para dicho $x \in \mathbb{R}^n$ tenemos que $L(x) = (a_1,a_2,\dots,a_i,\dots,a_n) \cdot (\emptyset,\emptyset,\dots,c/a_i,\dots,\emptyset) = a_i(c/a_i) = c$. i.e. Hemos sido capaces de probar que existe $x \in \mathbb{R}^n$ tal que $L(x) = c$ $c \in \mathbb{R}$ arbitrario $\therefore$ L es sobre. Y por tanto $\dim H_\emptyset = \dim(\ker L) = n - \dim(\mathrm{Im}L) = n - 1$.

Vemos que en el caso $n = 2$; $H_\emptyset$ corresponde a una recta que pasa por el origen: la gráfica de una función lineal de $\mathbb{R}^1$ a $\mathbb{R}$.

Para $n = 3$, $H_\emptyset$ corresponde a un plano que pasa por el origen: la gráfica de una función lineal de $\mathbb{R}^2$ en $\mathbb{R}$.

En general $H_\emptyset$ es un subespacio vectorial de $\mathbb{R}^n$ de dimensión $n - 1$ y corresponde geométricamente a la gráfica de una función lineal de $\mathbb{R}^{n-1}$ a $\mathbb{R}$, pues una función lineal de $\mathbb{R}^{n-1}$ en $\mathbb{R}$ debe ser de la forma: $L: \mathbb{R}^{n-1} \to \mathbb{R}$ $L(x_1,\dots,x_{n-1}) = a_1x_1+a_2x_2+\dots+a_{n-1}x_{n-1}$ para algunos $a_i \in \mathbb{R}$ $i = 1,2,\dots,n-1$. La gráfica de dicha función está dada por: $GrL = \{(x_1,x_2,\dots,x_{n-1},x_n) \in \mathbb{R}^n \ /\ x_n = L(x_1,x_2,\dots,x_{n-1}) = a_1x_1+\dots+a_{n-1}x_{n-1}\} = \{(x_1,x_2,\dots,x_n) \in \mathbb{R}^n \ /\ a \cdot x = \emptyset$ con $a = (a_1,a_2,\dots,a_{n-1},-1)\}$ que es efectivamente un hiperplano en $\mathbb{R}^n$."

"Veamos qué forma tienen las transformaciones lineales $L: \mathbb{R} \to \mathbb{R}^n$. Como $t = t \cdot 1$, $1 \in \mathbb{R}$ y $\{v_1\}$ con $v_1=1$, es base para $\mathbb{R}$, se tiene que $L(t) = L(t \cdot 1) = tL(1)$. En otras palabras para encontrar $L(t)$ sólo necesitamos conocer $L(1) \in \mathbb{R}^n$ y multiplicar escalarmente dicho vector de $\mathbb{R}^n$ por t.

Se tienen entonces dos posibilidades: $L(1) = \emptyset \in \mathbb{R}^n$ en cuyo caso $L(t) = \emptyset$ $\forall\ t \in \mathbb{R}$ (i.e. L es la transformación cero de $\mathbb{R}$ en $\mathbb{R}^n$) ó $L(1) \neq \emptyset$ en cuyo caso la imagen de L es la recta en $\mathbb{R}^n$, que pasa por el origen y por el punto $L(1)$.

Se puede ver que hay una correspondencia biunívoca entre transformaciones lineales $L: \mathbb{R} \to \mathbb{R}^n$ y elementos de $\mathbb{R}^n$, pues dado $a = (a_1, a_2, \ldots, a_n) \in \mathbb{R}^n$ dicho vector determina una transformación lineal de $\mathbb{R}$ en $\mathbb{R}^n$, a saber: $L(t) = ta$. Inversamente dada una transformación lineal $L: \mathbb{R} \to \mathbb{R}^n$ ésta es de la forma $L(t) = tL(1)$ donde $L(1) = (a_1, a_2, \ldots, a_n) \in \mathbb{R}^n$."

Como es un hecho que se usa al hablar de la diferencial de una suma de funciones diferenciables, cabe hablar sobre la suma de funciones lineales:

"La suma L_1+L_2 de dos transformaciones lineales $L_1, L_2: \mathbb{R}^n \to \mathbb{R}^m$ es una nueva transformación definida de la manera ususal, es decir $(L_1+L_2)(x) = L_1(x)+L_2(x)$.

Además si L_1, L_2 son lineales entonces L_1+L_2 es lineal, pues $\forall\ x, y \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$: $(L_1+L_2)(x+\lambda y) = L_1(x+\lambda y)+L_2(x+\lambda y)$ $= L_1(x)+\lambda L_1(y)+L_2(x)+\lambda L_2(y) = L_1(x)+L_2(x)+\lambda(L_1(y)+L_2(y)) =$ $(L_1+L_2)(x) + \lambda(L_1+L_2)(y)$.

Más aún, si A_1, A_2 son las matrices que representan a L_1 y L_2 respectivamente, entonces: $L_1(x) = A_1x$ $L_2(x) = A_2x$ $\therefore$ $(L_1+L_2)(x) = L_1(x)+L_2(x) = A_1x+A_2x = (A_1+A_2)x \Rightarrow A_1 + A_2$ es la matriz que representa a $L_1 + L_2$."

En el curso de Cálculo vectorial diferencial un tema importante es la composición de funciones diferenciables, lo que presupone el manejo de la composición de funciones lineales. Es este momento del curso, un lugar adecuado para introducir el tema:

"Si $f: \mathbb{R} \to \mathbb{R}$, $g: \mathbb{R} \to \mathbb{R}$. Podemos formar la función compuesta $g \circ f: \mathbb{R} \to \mathbb{R}$ definida por $(g \circ f)(x) = g(f(x))$.

Si ahora consideramos transformaciones lineales: $T: \mathbb{R}^n \to \mathbb{R}^m$, $S: \mathbb{R}^m \to \mathbb{R}^p$, podemos igualmente formar la función compuesta de $\mathbb{R}^n$ en $\mathbb{R}^p$ denotada por $S \circ T$ y definida por $(S \circ T)(x) = S(T(x))$

$$\mathbb{R}^n \xrightarrow{T} \mathbb{R}^m \xrightarrow{S} \mathbb{R}^p$$

$$S \circ T: \mathbb{R}^n \longrightarrow \mathbb{R}^p \,."$$

Los objetivos mínimos a cubrir son:
- La composición de funciones lineales es lineal:

"La transformación $S \circ T: \mathbb{R}^n \to \mathbb{R}^p$ es lineal si $T: \mathbb{R}^n \to \mathbb{R}^m$ y $S: \mathbb{R}^m \to \mathbb{R}^p$ lo son.

Dem. Sean $x, y \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$

$$(S \circ T)(x+\lambda y) = S(T(x+\lambda y)) = S(T(x)+\lambda T(y)) =$$
$$= S(T(x))+\lambda S(T(y)) = (S \circ T)(x)+\lambda(S \circ T)(y)."$$

- La matriz de la composición es el producto de las matrices:

"Para encontrar la matriz que representa a la transformación $S \circ T$, sean A y B las matrices que representan a T y S respectivamente,

$$(S \circ T)(x) = S(T(x)) = S(Ax) = B(Ax) = BAx.$$

Por lo tanto BA es la matriz que representa a la transformación lineal $S \circ T$.

Observamos que las restricciones en los tamaños de las matrices para poder efectuar el producto son satisfechas pues B es de $p{\times}m$, A de $m{\times}n$ $\therefore$ el producto BA se puede realizar y es una matriz de $p{\times}n$ que representa a una transformación lineal de $\mathbb{R}^n$ a $\mathbb{R}^p$.

Como una ilustración consideremos las siguientes transformaciones lineales del plano ($\mathbb{R}^2$) en él mismo.

$S: \mathbb{R}^2 \to \mathbb{R}^2$, $S(x,y) = (2x-3y, x+y)$ i.e. $B = \begin{bmatrix} 2 & -3 \\ 1 & 1 \end{bmatrix}$

$T: \mathbb{R}^2 \to \mathbb{R}^2$, $T(x,y) = (x+y, 3x+y)$ i.e. $A = \begin{bmatrix} 1 & 1 \\ 3 & 1 \end{bmatrix}$

$$BA = \begin{bmatrix} 2 & -3 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} -7 & -1 \\ 4 & 2 \end{bmatrix}$$

$\therefore$ la transformación S o T está representada por la matriz
$$\begin{bmatrix} -7 & -1 \\ 4 & 2 \end{bmatrix}$$
En funciones coordenadas:

$$(S \circ T)(x,y) = \begin{bmatrix} -7 & -1 \\ 4 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -7x-y \\ 4x+2y \end{bmatrix} \Rightarrow \left\{ \begin{array}{l} f_1(x,y) = -7x-y \\ f_2(x,y) = 4x+2y \end{array} \right.$$

Como comprobación:
$$(S \circ T)(x,y) = S(T(x,y)) = S(x+y,3x+y) =$$
$$= (2(x+y)-3(3x+y),\ (x+y)+(3x+y)) = (-7x-y,4x+2y)$$
$$\therefore\ f_1(x,y) = -7x-y$$
$$f_2(x,y) = 4x+2y."$$

- Lo concerniente a la inversa de una transformación lineal:

Se supone que los alumnos manejan la definición de función inversa, y el hecho de que en general para una función $f: A \to B$ (A,B conjuntos) f tiene inversa $f^{-1}: B \to A$ si y sólo si f es inyectiva y suprayectiva:

"Sea $L: \mathbb{R}^n \to \mathbb{R}^n$ una transformación lineal. Decimos que L es invertible si existe una transformación denotada $L^{-1}: \mathbb{R}^n \to \mathbb{R}^n$ tal que $(L \circ L^{-1}): \mathbb{R}^n \to \mathbb{R}^n$ y $(L^{-1} \circ L): \mathbb{R}^n \to \mathbb{R}^n$ resultan ser la función identidad en $\mathbb{R}^n$.

Si $L: \mathbb{R}^n \to \mathbb{R}^n$ es lineal y tiene inversa L^{-1}, entonces L^{-1} también es lineal.

Dem. Sean $y_1, y_2 \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$

Por demostrar $L^{-1}(y_1+\lambda y_2) = L^{-1}(y_1)+\lambda L^{-1}(y_2)$ como L es biyectiva (pues existe la inversa L^{-1})

$\exists\ x_1 \in \mathbb{R}^n$ único $\ni L(x_1) = y_1$ (i.e. $x_1 = L^{-1}(y_1)$)

$\exists\ x_2 \in \mathbb{R}^n$ único $\ni L(x_2) = y_2$ (i.e. $x_2 = L^{-1}(y_2)$)

$\therefore\ L^{-1}(y_1+\lambda y_2) = L^{-1}(L(x_1)+\lambda L(x_2)) = L^{-1}(L(x_1+\lambda x_2))$ (Por la linealidad de L) $= x_1+\lambda x_2$ (por definición de inversa) $= L^{-1}(y_1)+\lambda L^{-1}(y_2)$ l.c.q.d.

Sea $L: \mathbb{R}^n \to \mathbb{R}^n$ una transformación lineal y sea A la matriz que la representa $\therefore$ $L(x) = Ax$.

Supongamos que L^{-1} existe y sea B la matriz que la representa, entonces $(L^{-1} \circ L)(x) = BAx$. Pues la matriz de una composición es el producto de las matrices. Por otro lado $(L^{-1} \circ L)(x) = L^{-1}(L(x)) = x \Rightarrow \forall x \in \mathbb{R}^n \; BAx = x \Rightarrow BAx - Ix = \emptyset$ (donde I es la matriz identidad de nxn) $\therefore \; (BA-I)x = \emptyset \quad \forall x \in \mathbb{R}^n$.

En particular $(BA-I)e_k = \emptyset$; $k = 1,2,\ldots,n$ (como en general Ce_k nos da la k-ésima columna de la matriz C)

$$\Rightarrow BA-I = \emptyset$$

$$BA = I$$

Análogamente también debemos tener $ABx = x \quad \forall x \in \mathbb{R}^n$ de donde $AB = I \qquad \therefore \; B = A^{-1}$

En otras palabras si A es la matriz que representa a una transformación lineal invertible $L: \mathbb{R}^n \to \mathbb{R}^n$, entonces A^{-1} es la matriz que representa a L^{-1}.

Por lo cual, dada una transformación lineal de $\mathbb{R}^n$ en $\mathbb{R}^n$, para saber si es invertible (es decir L es un isomorfismo) basta ver que el determinante de su matriz asociada es $\neq \emptyset$, pues en ese caso A^{-1} existe, y por tanto L^{-1} existe, ya que $L^{-1}: \mathbb{R}^n \to \mathbb{R}^n$ es la transformación lineal definida mediante la ecuación: $L^{-1}(x) = A^{-1}x \qquad x \in \mathbb{R}^n$."

CAPITULO 5

DE REGRESO AL CALCULO

5.Ø INTRODUCCION

Después del estudio de las transformaciones lineales, retomamos el Cálculo diferencial. Al disponer ya de los antecedentes necesarios, se puede presentar la definición de diferenciabilidad de una función en un punto como una aproximación lineal local a la función, lo que constituye la generalización adecuada, para funciones de varias variables, del concepto de derivada que se manejaba para funciones reales de una variable y cuya búsqueda quedó incoada en el capítulo 3.

Para que este concepto sea mejor comprendido cuando se plantée en el caso general de funciones $f:D \subseteq \mathbb{R}^n \to \mathbb{R}^m$, se propone irlo presentando progresivamente. Con tal objeto, el presente capítulo se aboca a revisar con cierto detalle en qué consiste analítica y geométricamente la diferenciabilidad para los distintos casos que ya se han estudiado: las funciones de $\mathbb{R}$ en $\mathbb{R}$, de $\mathbb{R}$ en $\mathbb{R}^n$ y de $\mathbb{R}^n$ en $\mathbb{R}$.

En los dos primeros, se muestra que la derivabilidad, -aquello en lo cual han trabajado los estudiantes- implica la existencia de una aproximación lineal local (diferenciabilidad). Esto permite introducir un nuevo concepto, aprovechando su relación con otro que ya es conocido. Una vez familiarizados con el nuevo concepto, se estudia la relación de éste con el anterior, para concluir que ambos son equivalentes, lo cual sirve para reforzar el conocimiento de esta nueva propiedad que es la diferenciabilidad.

Para el caso de funciones de $\mathbb{R}^n$ en $\mathbb{R}$, el esquema es diferente. Basados en que ya se introdujo y se trabajó -en los casos anteriores- la definición de diferenciabilidad, ésta se presenta como una propiedad de algunas funciones. Una propiedad que tiene un contenido propio, que resulta más inteligible para los estudiantes después de haberla analizado en los casos más familiares de funciones de una variable. Como el propósito de presentar la diferencial es mostrar que ésta es la generalización del concepto de derivada para las funciones de $\mathbb{R}^n$ en $\mathbb{R}$, se analizan las implicaciones de la diferenciabilidad, demostrando que con la diferenciabilidad se satisfacen las condiciones deseables para una tal generalización, tanto analíticas (la continuidad), como geométricas (existencia del plano tangente). Al trabajar esta línea se obtienen los resultados que relacionan este nuevo concepto de diferenciabilidad con la derivación manejada anteriormente, destaca la utilidad que reportan las derivadas para trabajar con la diferenciabilidad y, de hecho, constituyen una parte importante del curso de Cálculo vectorial, la obtención de la matriz Jacobiana, del vector gradiente, el uso del vector gradiente para distintas aplicaciones, el inferir la diferenciabilidad a partir de la continuidad de las parciales, etc.

5.1 UN RETORNO A LAS FUNCIONES DE $\mathbb{R}$ EN $\mathbb{R}$

Muchos libros de texto sobre el Cálculo de una variable, tienen algún capítulo dedicado al estudio del tema que denominan "Diferenciales". Ocurre que cuando los alumnos en cursos de Cálculo vectorial aprenden que la diferencial de una función (función vectorial de variable vectorial) en un punto, es una transformación lineal que aproxima bien a la función dada en una vecindad del punto en cuestión (donde la frase "aproxima bien" tiene un sentido matemático preciso) y tratan de relacionar estas ideas con lo que aprendieron en el Cálculo de una variable con el nombre de "Diferenciales", les parece que, fuera del nombre, tienen poco en común.

Citemos algún texto que nos sirva de muestra para analizar como es tratado el tema en buena parte de la bibliografía existente. (N.B. Esto sucede especialmente en libros escritos hace ya tiempo, una bibliografía más reciente presenta este material de una manera similar a la que se sugiere en el presente trabajo).

"Diferenciales.

Definiciones: Si $f'(x)$ es la derivada de $f(x)$ para un valor particular de x y Δx es un incremento de x, arbitrariamente elegido, la diferencial de $f(x)$ que se representa por el símbolo df(x) se define por la igualdad:

$$\text{(A)}\quad df(x) = f'(x)\,\Delta x = \frac{dy}{dx}\,\Delta x$$

Si $f(x) = x$, entonces $f'(x) = 1$ y (A) se reduce a $dx = \Delta x$. Así, cuando x es la variable independiente, la diferencial de x (=dx) es idéntica a Δx. Por tanto si $y = f(x)$, (A) puede, en general escribirse en la forma:

$$\text{(B)}\quad dy = f'(x)dx = \frac{dy}{dx}\,dx$$

La diferencial de una función es igual al producto de su derivada por la diferencial de la variable independiente.

Veamos una interpretación geométrica de lo que esto significa. Construyamos la curva $y = f(x)$ (Figura 5.1). Sea $f'(x)$ el valor de la derivada en P. Tomemos $dx = PQ$. Entonces,

$$dy = f'(x)dx = \tan \tau \; PQ$$

$$= \frac{QT}{PQ}\,PQ = QT$$

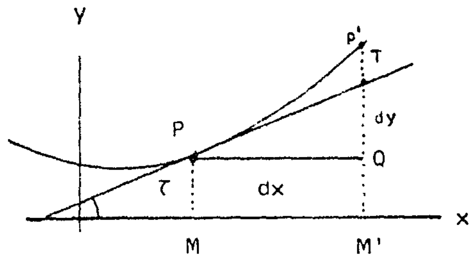


Figura 5.1

Luego dy, o sea df(x), es el incremento (= QT) de la ordenada de la tangente correspondiente a dx.

Esto da la siguiente interpretación de la derivada como fracción:

Si se representa por dx un incremento arbitrariamente elegido de la variable independiente x para un punto P(x,y) en la curva y=f(x), entonces en la derivada

$$\frac{dy}{dx} = f'(x) = \tan \tau$$, dy representa el incremento correspondiente de la ordenada de la tangente en P.

El lector debe advertir especialmente que la diferencial (=dy) y el incremento (=Δy) no son, en general iguales. En efecto, en la figura 5.1 dy = QT pero Δy = QP'.

Es claro que Δy(=QP' en la figura) y dy(=QT) son aproximadamente iguales cuando dx(=PQ) es pequeño. Cuando solamente se desea un valor aproximado del incremento de una función, es más fácil, la mayor parte de las veces, calcular el valor de la diferencial correspondiente y emplear este valor." (Granville, Smith, Longley, Cálculo Diferencial e Integral, México, D.F. - UTEHA; Edición revisada, reimpresión 1962, págs. 164 - 166).

Después de una exposición como la anterior, los libros suelen presentar ejemplos y ejercicios de aproximación de funciones y de estimación de errores. Sobre este tipo de aplicaciones ya se comentó en el capítulo 1.

En la exposición anterior del tema queda muy diluida la noción de diferenciabilidad entendida como la existencia de una función lineal que aproxima localmente a la función dada.

Por ello, para evitar esa falta de continuidad, habría que tratar el tema de diferenciales para funciones reales de una variable con el enfoque con que se hará para funciones vectoriales.

Pueden servir textos como el siguiente, que dejan clara esa idea: si la función es derivable, existe una función lineal que aproxima localmente a la función.

"Para una función $f:D \subseteq \mathbb{R} \to \mathbb{R}$, la derivada de f en x_0, denotada $f'(x_0)$, se define como el límite:

$$f'(x_0) = \lim_{h \to 0} \frac{f(x_0+h) - f(x_0)}{h}$$ en caso de que el límite exista.

Si la derivada existe para todo x en algún subconjunto $D' \subseteq D$, entonces podemos definir una nueva función llamada la derivada de f que denotamos f'; $f':D' \subseteq \mathbb{R} \to \mathbb{R}$

de la siguiente forma: $x \xrightarrow{f'} f'(x) \quad x \in D'$.

Para cada punto $x_0 \in D'$ (i.e. $f'(x_0)$ existe) definimos la función $df(x_0)$; $df(x_0): \mathbb{R} \to \mathbb{R}$ por:

$$df(x_0)(h) = f'(x_0)h \qquad \forall\ h \in \mathbb{R}$$

Nótese que aunque f' esté definida sólo en un cierto subconjunto $D' \subseteq \mathbb{R}$, para cada $x_0 \in D'$, $df(x_0)$ es una función definida sobre todo $\mathbb{R}$.

A la función $df(x_0)$ la llamamos la diferencial de f en x_0. Se observa que $f'(x_0)$ es un número por lo que la función $df(x_0)$ es una función lineal en la variable h.

Esta función lineal aproxima localmente a la función en el sentido de que si consideramos la diferencia:

$$R(x_0,h) = f(x_0+h) - f(x_0) - df(x_0)(h)$$

observamos que: $\displaystyle\lim_{h\to 0} \frac{R(x_0,h)}{h} \quad \lim_{h\to 0} \frac{f(x_0+h) - f(x_0) - f'(x_0)h}{h}$

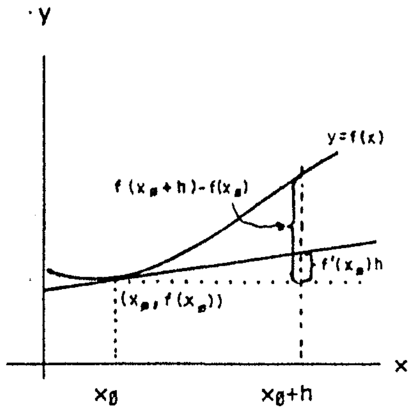
$$= \lim_{h\to 0} \left[\frac{f(x_0+h) - f(x_0)}{h} - f'(x_0) \right] = f'(x_0) - f'(x_0) = 0$$

$$\text{i.e.} \quad \lim_{h\to 0} \frac{R(x_0,h)}{h} = 0 \qquad \text{"}$$

Si el profesor al hablar de la diferencial para funciones reales, decidiera utilizar la notación:

$dy = f'(x)\,dx$ (cfr. la expresión (B) del texto citado)

por ser ésta de uso muy frecuente en la bibliografía; el uso de esta notación, debe acompañarse de la explicación detallada de la misma, algo que con frecuencia se omite. Concretamente, convendría señalar que dy, dx se emplean en este contexto para designar variables en exactamente la misma forma que x, h, t, etc.



Si se sigue esta línea habría que dar, adecuada a este lenguaje, la interpretación geométrica en términos que fueran muy explícitos. De tal manera que quede clara la idea: La gráfica de la diferencial de f en x_0 es la recta tangente a $y = f(x)$ en $(x_0, f(x_0))$, en un sistema de ejes coordenados con origen en $(x_0, f(x_0))$. (ver Figura 5.2).

Figura 5.2

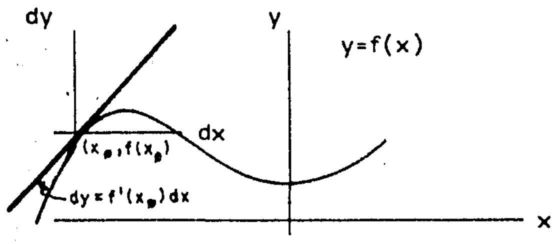
"Si marcamos los ejes coordenados de un plano cartesiano por medio de las variables dx, dy (en la misma forma que se usan x,y para los ejes horizontal y vertical habitualmente); la recta por el origen y con pendiente m sería la gráfica de la función lineal: $dx \longrightarrow mdx$. Con lo que $dy = mdx$ sería la ecuación en dicho plano de la gráfica de la función.

Supongamos que una función f es derivable en x_0. Superponemos ahora el plano dx, dy sobre el plano de tal manera que:

(1) El origen del plano dx, dy coincida con el punto $(x_0, f(x_0))$.

(2) Los ejes dx, dy sean paralelos a los ejes x, y respectivamente y apunten en las mismas direcciones.

(3) Los ejes dx, dy tengan la misma escala que los ejes x y y respectivamente. (Figura 5.3)



La recta T es la tangente a la gráfica de f en el punto $(x_0, f(x_0))$ y tiene pendiente $m = f'(x_0)$.

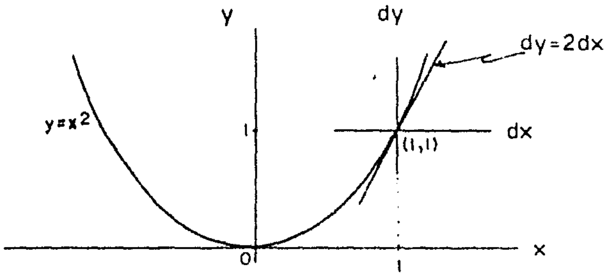
Figura 5.3

Consideremos ahora la función lineal dada por $dy = f'(x_0)dx$. La gráfica de esta función en el plano dx, dy coincide con la recta tangente a la gráfica de f en $(x_0, f(x_0))$ en el plano x, y. Esta tangente tiene por ecuación $y - f(x_0) = f'(x_0)(x - x_0)$. Si hacemos la traslación de ejes (en eso consiste el proceso geométrico descrito en los 3 puntos anteriores) dada por:

$$\begin{cases} dy = y - f(x_0) \\ dx = x - x_0 \end{cases}$$

La ecuación de la tangente en el plano x, y se convierte en la ec. $dy = f'(x_0)dx$ en el plano dx, dy.

Ejemplo 1. Si $y = x^2$ entonces $dy = 2xdx$. En particular para $x = 1$ obtenemos la función lineal $dy = 2dx$. (Figura 5.4)



La ec. de la tangente a la curva $y=x^2$ en el plano dx, dy está dada por $dy = 2dx$.

Figura 5.4

Como en la notación $dy = f'(x)dx$ queda un poco obscurecido el hecho de que dy es una función lineal en la variable dx, se recomienda denotar la diferencial de la función $y = f(x)$ en el punto x_0 como se hizo anteriormente:

$$df(x_0)(h) = f'(x_0)h \qquad df(x_0): \mathbb{R} \to \mathbb{R}."$$

Otra notación, que se utilizará sistemáticamente en todo lo que resta del presente trabajo es la siguiente:

$$L_{f,x_0}: \mathbb{R} \to \mathbb{R}\ ,\ L_{f,x_0}(h) = f'(x_0)h.$$

Se recomienda esta notación porque pone de relieve por un lado la naturaleza de función lineal que tiene la diferencial de la función en el punto, al mismo tiempo se hace mención explícita de la función y del punto en cuestión.

Un objetivo didáctico deseable es presentar la diferenciabilidad de las funciones reales de una variable, de forma tal que coincida con la definición que posteriormente se dará para el caso general. ($f: D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ es diferenciable en $x_\emptyset \in D$ si existe una transformación lineal $L: \mathbb{R}^n \to \mathbb{R}^m$ tal que

$$\lim_{h\to\emptyset} \frac{||f(x_\emptyset+h)-f(x_\emptyset)-L(h)||}{||h||} = \emptyset \quad).$$

Hasta ahora, el estado de conocimientos de los estudiantes es el siguiente:

Una función $f: D \subseteq \mathbb{R} \to \mathbb{R}$ es derivable en $x_\emptyset \in D$ si el límite

$$\lim_{h\to\emptyset} \frac{f(x_\emptyset+h)-f(x_\emptyset)}{h} = f'(x_\emptyset)$$

existe. Si la función es derivable en $x_\emptyset$, entonces existe una función lineal llamada la diferencial de f en $x_\emptyset$, denotada $L_{f,x_\emptyset}$ definida por:

$$L_{f,x_\emptyset}(h) = f'(x_\emptyset)h \qquad L_{f,x_\emptyset}: \mathbb{R} \to \mathbb{R}$$

que satisface:

$$\lim_{h\to\emptyset} \frac{f(x_\emptyset+h)-f(x_\emptyset)-L_{f,x_\emptyset}(h)}{h} = \emptyset$$

Cuya gráfica en un sistema de ejes coordenados con el origen en $(x_\emptyset, f(x_\emptyset))$ es la recta tangente a la curva $y = f(x)$ en el punto $(x_\emptyset, f(x_\emptyset))$.

Con estos elementos se puede entonces, centrando la atención en la función lineal así obtenida, definir la diferenciabilidad de una función $f: D \subseteq \mathbb{R} \to \mathbb{R}$ en un punto $x_\emptyset$.

"Definición. Decimos que $f: D \subseteq \mathbb{R} \to \mathbb{R}$ es diferenciable en $x_\emptyset$ si existe una función lineal $L: \mathbb{R} \to \mathbb{R}$ tal que:

$$\lim_{h\to\emptyset} \frac{f(x_\emptyset+h)-f(x_\emptyset)-L(h)}{h} = \emptyset."$$

Para que los estudiantes tengan completo el cuadro de la relación entre el familiar concepto de derivabilidad (como la existencia de un número) y esta nueva idea de diferenciabilidad (como aproximación lineal local a la función) habría que hacerles notar, que de la discusión anterior se infiere que si una función es derivable en un punto, entonces es diferenciable en ese punto; pero como para este caso ambos conceptos son equivalentes, la afirmación recíproca es válida:

"Dada $f: D \subseteq \mathbb{R} \to \mathbb{R}$ $x_0 \in D$.
Si existe una transformación lineal $L: \mathbb{R} \to \mathbb{R}$ tal que

$$\lim_{h\to 0} \frac{f(x_0+h)-f(x_0)-L(h)}{h} = 0$$

i.e., f es diferenciable en x_0, entonces f es derivable en x_0, pues si $L: \mathbb{R} \to \mathbb{R}$ es lineal $\Rightarrow \exists\, A \in \mathbb{R} \ni L(h) = Ah$

$$\therefore 0 = \lim_{h\to 0} \frac{f(x_0+h)-f(x_0)-L(h)}{h} = \lim_{h\to 0} \frac{f(x_0+h)-f(x_0)-Ah}{h} =$$

$$= \lim_{h\to 0} \frac{f(x_0+h)-f(x_0)}{h} - A.$$

$$\Rightarrow \lim_{h\to 0} \frac{f(x_0+h)-f(x_0)}{h} = A$$

i.e. $A = f'(x_0)$ "
(f es derivable en x_0).

Como en este caso ambos conceptos son equivalentes, cabe comentar que en ocasiones se usan indistintamente las expresiones de función derivable o función diferenciable. Sin embargo, como esta situación no se da para las funciones de varias variables, ambos conceptos son equivalentes en cuanto que uno implica al otro, pero cada uno tiene una razón formal diferente. La derivabilidad (de una función f en el punto x_0) hace referencia a la existencia de un número real

(denotado $f'(x_0)$ con $f'(x_0) = \lim_{h\to 0} \frac{f(x_0+h)-f(x_0)}{h}$).

Mientras que la diferenciabilidad (de f en x_0) connota la existencia de una función lineal que aproxima localmente a la función dada (específicamente, si la función lineal se denota $L: \mathbb{R} \to \mathbb{R}$ ésta satisface:

$$\lim_{h\to 0} \frac{f(x_0+h)-f(x_0)-L(h)}{h} = 0).$$

5.2 UN RETORNO A LAS FUNCIONES DE $\mathbb{R}$ EN $\mathbb{R}^n$.

La situación con estas funciones es análoga a la de las funciones de $\mathbb{R}$ en $\mathbb{R}$, en el sentido de que lo que es conocido por los estudiantes con respecto a estas funciones es la derivabilidad de una función $f: D \subseteq \mathbb{R} \to \mathbb{R}^n$ en un punto $x_0 \in D$ (cfr. capítulo 2) entendida como la existencia del límite

$$\lim_{h \to 0} \frac{f(x_0+h)-f(x_0)}{h} = f'(x_0)$$

siendo el único cambio, que $f'(x_0)$ es un vector de $\mathbb{R}^n$.

El objetivo de esta sección es estudiar la diferenciabilidad de una función en un punto, como la existencia de una transformación lineal que aproxima localmente a la función.

Aprovecharemos el hecho, de que para este caso ambos conceptos son equivalentes para introducir la idea de aproximación lineal local a una función a partir del concepto - ya conocido por los estudiantes - de derivada. Con la ventaja de que los estudiantes a través de este caso se irán familiarizando con la idea misma de aproximar localmente con transformaciones lineales y la correspondiente interpretación geométrica.

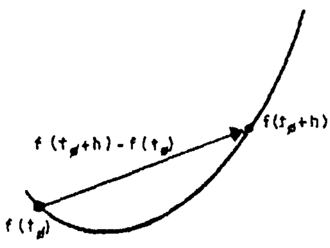


Figura 5.5

Para ello, se podría seguir el siguiente desarrollo:

"Dada una función $f: D \subseteq \mathbb{R} \to \mathbb{R}^n$, estamos interesados en la manera en la que los valores de la función f cambian conforme nos movemos de un punto $t_0 \in D$.

Para ello consideramos la diferencia $f(t_0+h)-f(t_0)$ para aquellos $h \in \mathbb{R}$ para los cuales $t_0 + h \in D$ (Figura 5.5)

Ejemplo 1. Sea $D = (0,3)$ y $f: D \subseteq \mathbb{R} \to \mathbb{R}$ dada por $f(x) = 1/x(x-3)$. Si $t_0 = 1$, tenemos entonces que $f(1+h) - f(1) = [1/(h+1)(h-2)] + 1/2$. Dicha diferencia está definida para $h \in (-1,2)$.

Supongamos que la función $f: D \subseteq \mathbb{R} \to \mathbb{R}^n$ es derivable en $t_0 \in D$. Entonces

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

Se sigue que para cada $h \ni t_0 + h \in D$ $(h \neq 0)$ existe un vector $R(t_0,h) \in \mathbb{R}^n$ tal que:

(1) ..... $f(t_0+h) - f(t_0) = hf'(t_0) + R(t_0,h)$ en donde

$$\lim_{h \to 0} \frac{R(t_0,h)}{h} = \lim_{h \to 0} \left[\frac{f(t_0+h) - f(t_0)}{h} - f'(t_0) \right] = 0.$$

La expresión (1) indica que $f(t_0+h) - f(t_0)$ se descompone como la suma de dos vectores, el primero de ellos $hf'(t_0)$ es un vector tangente a f en t_0; el segundo vector $R(t_0,h)$ es esencialmente el error que se comete al moverse a lo largo de la tangente en lugar de hacerlo a lo largo de la curva. La descomposición se ilustra en la Figura 5.6

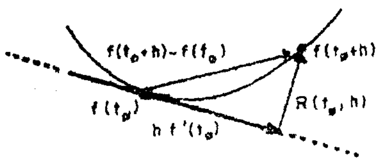


Figura 5.6

Ejemplo 2. Definimos $f: \mathbb{R} \to \mathbb{R}^2$ por $f(t) = (t^3+1, t)$. Para $t_0 \in \mathbb{R}$ tenemos $f'(t_0) = (3t_0^2, 1)$

$(\forall\ h \in \mathbb{R})$ $f(t_0+h) - f(t_0) = ((t_0+h)^3+1, (t_0+h)) - (t_0^3+1, t_0)$
$= (3t_0^2h + 3t_0h^2 + h^3, h) = h(3t_0^2, 1) + (3t_0h^2 + h^3, 0)$

$\therefore\ R(t_0,h) = (3t_0h^2 + h^3, 0)\quad h \in \mathbb{R}$

$$\lim_{h \to 0} \frac{R(t_0,h)}{h} = \lim_{h \to 0} \frac{1}{h} h^2(3t_0+h, 0) = \lim_{h \to 0} h(3t_0+h, 0) = (0,0)$$

Ejemplo 3. Definamos $f: \mathbb{R} \to \mathbb{R}^2$ como $f(t) = (1,2) + t(1,1)$. La imagen de f es una recta que pasa por $(1,2)$. En $t_0 \in \mathbb{R}$ tenemos: $f(t_0+h) - f(t_0) = (1,2) + (t_0+h)(1,1) - (1,2) - t_0(1,1) = h(1,1) + (0,0)$. Se observa que $f'(t_0) = (1,1)$

$\therefore\ f(t_0+h) - f(t_0) = hf'(t_0) \Rightarrow R(t_0,h) = (0,0)\quad \forall\ h$.

Claramente $\lim_{h \to \emptyset} \dfrac{R(t_\emptyset, h)}{h} = (\emptyset, \emptyset)$

Podemos reescribir la expresión (1) si definimos una función lineal $L_{f,t_0}: \mathbb{R} \to \mathbb{R}^n$ por: $L_{f,t_\emptyset}(t) = tf'(t_\emptyset)$ $t \in \mathbb{R}$.

(Por nuestro estudio del capítulo anterior, sabemos que toda transformación lineal de $\mathbb{R}$ en $\mathbb{R}^n$ es de la forma $L(t) = ta$ para algún $a \in \mathbb{R}^n$ (constante)).

De (1) obtenemos entonces que:

(2)..... $f(t_\emptyset+h) - f(t_\emptyset) = L_{f,t_\emptyset}(h) + R(t_\emptyset, h)$

(en toda nuestra discusión se sobreentiende que las $h \in \mathbb{R}$ permitidas son aquellas para las que $t_\emptyset + h \in D$). $L_{f,t_\emptyset}$ se llama la aproximación lineal a f en $t_\emptyset$ o la diferencial de f en $t_\emptyset$.

Ejemplo 4. Consideremos nuevamente la función $f: \mathbb{R} \to \mathbb{R}^2$ definida como $f(t) = (t^3+1, t)$ (Ver Figura 5.7), entonces $f(\emptyset) = (1, \emptyset)$; $f(-1) = (\emptyset, -1)$ y tenemos: $f(\emptyset+h) - f(\emptyset) =$

$$(h^3+1, h)-(1, \emptyset) = (h^3, h) = h(\emptyset, 1) + (h^3, \emptyset)$$

$$L_{f,\emptyset}(h) = hf'(\emptyset) = h(\emptyset, 1) \quad \therefore \quad R(\emptyset, h) = (h^3, \emptyset).$$

$$f(-1+h)-f(-1) = ((h-1)^3+1, h-1)-(\emptyset, -1) =$$
$$= (h^3-3h^2+3h-1+1, h-1)-(\emptyset, -1) = (h^3-3h^2+3h, h) =$$
$$= h(3, 1)+(h^3-3h^2, \emptyset) \qquad L_{f,-1}(h) = hf'(-1) = h(3, 1)$$
$$\therefore \quad R(-1, h) = (h^3-3h^2, \emptyset) \text{ ."}$$

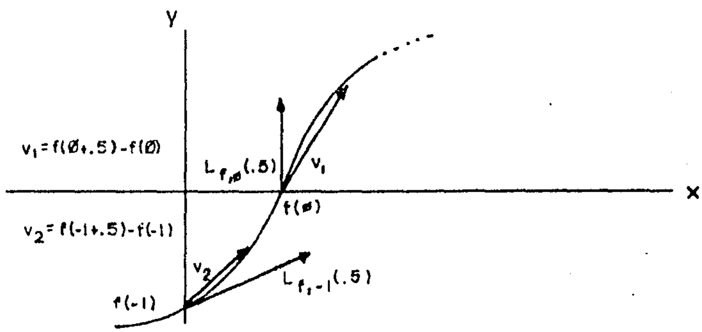


Figura 5.7

De esta forma ya tenemos los elementos para definir la diferenciabilidad de una función $f: D \subseteq \mathbb{R} \to \mathbb{R}^n$:

"f: $D \subseteq \mathbb{R} \to \mathbb{R}^n$ se dice que es diferenciable en $t_0 \in D$, si existe una transformación lineal $L: \mathbb{R} \to \mathbb{R}^n$ tal que:

$$\lim_{h\to 0} \frac{f(t_0+h)-f(t_0)-L(h)}{h} = 0 \text{ ."}$$

De la discusión anterior los estudiantes pueden notar que por lo pronto derivabilidad implica diferenciabilidad pues si la función es derivable, una transformación lineal que satisface la condición pedida está dada por

$$L_{f,t_0}(h) = hf'(t_0) \qquad h \in \mathbb{R}$$

(a la que habíamos llamado la diferencial de f en t_0).

Ahora habría que probar que la afirmación recíproca también es válida, i.e. que si una función es diferenciable entonces es derivable (en el punto) y que de hecho sólo puede haber una transformación lineal que satisfaga la definición de diferenciabilidad, lo que justificaría el nombre de la diferencial de la función en el punto.

"Supongamos que $f: D \subseteq \mathbb{R} \to \mathbb{R}^n$ es diferenciable en $t_0 \in D$ $\Rightarrow$ existe una transformación lineal $L: \mathbb{R} \to \mathbb{R}^n$ tal que:

$$\lim_{h\to 0} \frac{f(t_0+h)-f(t_0)-L(h)}{h} = 0$$

Como L es lineal; para $h \in \mathbb{R}$ $L(h) = hL(1)$ $1 \in \mathbb{R}$.

$$\therefore \lim_{h\to 0} \frac{f(t_0+h)-f(t_0)-hL(1)}{h} = 0$$

$$\Rightarrow \lim_{h\to 0} \left[\frac{f(t_0+h)-f(t_0)}{h} - L(1) \right] = 0$$

$$\therefore \lim_{h\to 0} \frac{f(t_0+h)-f(t_0)}{h} = L(1)$$

Con lo que, si f es diferenciable en t_0, entonces f es derivable en t_0 y $L(1) = f'(t_0)$.

Para probar la unicidad de la aproximación lineal, cuya existencia está postulada por la definición de diferenciabilidad, basta recordar que una transformación lineal de $\mathbb{R}$ en $\mathbb{R}^n$ es la forma $L(h) = hL(1)$ $h \in \mathbb{R}$, $L(1) \in \mathbb{R}^n$. $\therefore$ si f es diferenciable por lo que acabamos de probar, la aproximación lineal L estará definida por $L(h) = hL(1) = hf'(t_0)$ $h \in \mathbb{R}$ $\therefore L = L_{f,t_0}$. "

Cabe hacer entonces, la misma observación que se hizo para el caso de funciones reales de una variable, en el sentido de que por ser la derivabilidad y la diferenciabilidad equivalentes para las funciones $f: D \subseteq \mathbb{R} \to \mathbb{R}^n$, en ocasiones se utilizan de manera indistinta los términos derivable y diferenciable para estas funciones. Aunque nuevamente sería bueno insistir en el sentido de cada una: f es derivable en t_0 si existe un vector en $\mathbb{R}^n$, denotado $f'(t_0)$ tal que: $\lim_{h \to 0} (1/h) [f(t_0+h) - f(t_0)] = f'(t_0)$,

y donde dicho vector lo podemos pensar como el vector tangente a la imagen de f en $f(t_0)$. Mientras que f es diferenciable en t_0, si existe una transformación lineal que denotamos L_{f,t_0} llamada la diferencial de f en t_0: $L_{f,t_0}: \mathbb{R} \to \mathbb{R}^n$ tal que:

$$\lim_{h \to 0} \frac{f(t_0+h) - f(t_0) - L_{f,t_0}(h)}{h} = 0.$$

De hecho por las consideraciones anteriores esa función lineal está definida por $L_{f,t_0}(h) = hf'(t_0)$. Y la correspondiente interpretación geométrica de la diferenciabilidad es que si L_{f,t_0} no es la función cero de $\mathbb{R}$ a $\mathbb{R}^n$, entonces la imagen de $\mathbb{R}$ bajo L_{f,t_0} es una recta - la recta en $\mathbb{R}^n$ que pasa por el origen y por $L_{f,t_0}(1) = f'(t_0)$ - que trasladada a $f(t_0)$ es tangente a la imagen de f en $f(t_0)$. (Ver Figura 5.8)

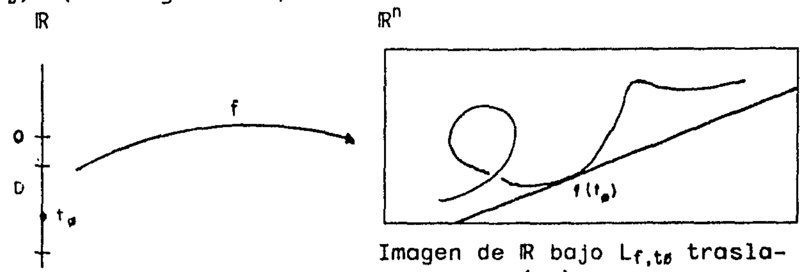


Imagen de $\mathbb{R}$ bajo L_{f,t_0} trasladada por $f(t_0)$

Figura 5.8

La función cuya imagen es esta recta aproxima a la función f alrededor de t_0, siendo esta aproximación mejor, mientras más cerca estemos de t_0; es decir, mientras h está más próxima a 0.

La intención de las consideraciones anteriores ha sido presentar a los estudiantes la noción de diferenciabilidad para funciones de $\mathbb{R}$ a $\mathbb{R}^n$ a partir de la noción de derivada que ya se manejaba.

Con el propósito de que, una vez familiarizados con la idea de aproximar linealmente de forma local a las funciones que geométricamente equivale a aproximar a la imagen de la función en un punto por medio de la recta tangente; puedan luego entender la definición de diferenciabilidad que se dará para funciones vectoriales de varias variables.

Para reforzar aún más el nexo entre la definición de diferenciabilidad que se ha dado para las funciones de $\mathbb{R}$ a $\mathbb{R}^n$ y la formulación de esta definición para el caso general, se puede hacer un pequeño ajuste en la definición de diferenciabilidad:

"Definición. Una función $f: D \subseteq \mathbb{R} \to \mathbb{R}^n$ es diferenciable en $t_0 \in D$ si existe una función lineal $L_{f,t_0}: \mathbb{R} \to \mathbb{R}^n$ llamada la diferencial de f en t_0 tal que:

si (i) $f(t_0+h)-f(t_0)-L_{f,t_0}(h) = R(t_0,h)$ (para $t_0+h \in D$)

entonces (ii) $$\lim_{h\to 0} \frac{||R(t_0,h)||}{|h|} = 0$$

Ambas condiciones pueden expresarse en una sola, pidiendo que L_{f,t_0} satisfaga:

$$\lim_{h\to 0} \frac{||f(t_0+h)-f(t_0)-L_{f,t_0}(h)||}{|h|} = 0$$

(El cambio puede hacerse ya que para $h \in \mathbb{R}$, $u \in \mathbb{R}^n$; $h \to 0 \iff |h| \to 0$ y $u \to 0 \iff ||u|| \to 0$."

Después de haber estudiado la diferencial de una función en un punto como una transformación lineal que aproxima localmente a la función, se puede introducir de manera natural, la matriz Jacobiana de la diferencial como la matriz que representa a la transformación lineal con respecto a las bases canónicas.

Esta manera de proceder, de presentar la diferencial de una función como una transformación lineal y no simplemente como una matriz, como suele hacerse en algunos textos se justificará en el siguiente capítulo.

"Si la función $f\colon D \subseteq \mathbb{R} \to \mathbb{R}^n$ es diferenciable en $t_0 \in D$, entonces existe L_{f,t_0}. Al ser una transformación lineal de $\mathbb{R}$ en $\mathbb{R}^n$, nos interesará, especialmente para efectos de operar con dicha función lineal, la matriz que la representa. Tenemos entonces la siguiente:

Definición. Si la función $f\colon D \subseteq \mathbb{R} \to \mathbb{R}^n$ es diferenciable en $t_0 \in D$, entonces la matriz de nx1 que denotaremos por J_{f,t_0} que representa a la transformación lineal $L_{f,t_0}\colon \mathbb{R} \to \mathbb{R}^n$ con respecto a las bases canónicas se llama la matriz Jacobiana de f en t_0. Como $L_{f,t_0}(t)=tL_{f,t_0}(1)=tf'(t_0)$
y $f'(t_0) = (f_1'(t_0),\ f_2'(t_0),\ \ldots,\ f_n'(t_0)) =$
$= f_1'(t_0)e_1 + f_2'(t_0)e_2 + \ldots + f_n'(t_0)e_n$.

Al ser la matriz de la transformación lineal en este caso la matriz cuyo (único) vector columna es $L_{f,t_0}(e_1) =$
$= L_{f,t_0}(1) \in \mathbb{R}^n$ y $L_{f,t_0}(1) = f_1'(t_0)e_1+f_2'(t_0)e_2+\ldots+f_n'(t_0)e_n$
$\Rightarrow$ la matriz Jacobiana de f en t_0 es:

$$J_{f,t_0} = \begin{bmatrix} f_1'(t_0) \\ f_2'(t_0) \\ \cdots\cdots \\ f_n'(t_0) \end{bmatrix}$$

Es decir la matriz Jacobiana de f en t_0, es la matriz que se forma transponiendo al vector tangente de f en t_0 y considerando dicho vector columna como matriz columna."

La notación J_{f,t_0} para denotar a la matriz Jacobiana de f en t_0, al hacer mención explícita de la función y del punto en donde está considerada la diferencial, resultará una notación conveniente para trabajar con la Regla de la Cadena.

5.3 UN RETORNO A LAS FUNCIONES DE $\mathbb{R}^n$ EN $\mathbb{R}$.

El hecho de que si una función $f: D \subseteq \mathbb{R} \to \mathbb{R}^n$ es derivable en $t_0 \in D$, entonces existe una transformación lineal L_{f,t_0} tal que la norma del vector diferencia o el error entre esta función lineal L_{f,t_0} y $f(t_0+h)-f(t_0)$ se va a cero más rápido que h, cuando h se va a cero, permitió introducir la definición de diferenciabilidad de una función en un punto, como la existencia de dicha aproximación lineal local (única).

Aprovechando esta experiencia con la definición de diferenciabilidad se puede presentar ahora la diferenciabilidad como un concepto propio, sin hacer referencia a otra propiedad de la función como la derivabilidad.

Se recomienda que, por simplificación, en este nivel de un primer curso de Cálculo vectorial, al hablar de la diferenciabilidad de una función en un punto, se considere siempre, que éste, es un punto interior del dominio de la función. (Dado $p \in D$, $\exists\, \varepsilon > 0 \ni B_\varepsilon(p) \subseteq D$).

"Consideremos la diferencia $f(p+h)-f(p)$ donde $p \in D \subseteq \mathbb{R}^n$ y $h \in \mathbb{R}^n \ni p+h \in D$. Dicha diferencia expresa el cambio en $f(x)$ conforme x se aleja de p. (Ver Figura 5.9)

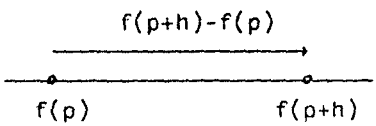


Figura 5.9

Definición 1. Una función $f: D \subseteq \mathbb{R}^n \to \mathbb{R}$ es diferenciable en $p \in D$ si existe una transformación lineal $L_{f,p}: \mathbb{R}^n \to \mathbb{R}$ tal que:

si (i) $R(p,h) = f(p+h)-f(p)-L_{f,p}(h)$

entonces (ii) $\displaystyle\lim_{h\to 0} \frac{R(p,h)}{||h||} = 0$

(se consideran aquellas h tales que $p + h \in D$).

o bien (reuniendo (i) y (ii)):

$$\lim_{h\to 0} \frac{f(p+h)-f(p)-L_{f,p}(h)}{||h||} = 0$$

Ejemplo 1. Sea $f(x_1,x_2) = x_1^2+x_2^2$ para cada $p \in \mathbb{R}^2$

$f(p+h)-f(p)=2p_1h_1+2p_2h_2+h_1^2+h_2^2 \quad p=(p_1,p_2) \quad h=(h_1,h_2)$

Si $L_{f,p}(h)=2p_1h_1+2p_2h_2$ tenemos que $f(p+h)-f(p)-L_{f,p}(h) =$

$h_1^2+h_2^2 = ||h||^2$

$$\therefore \lim_{h\to\emptyset} \frac{f(p+h)-f(p)-L_{f,p}(h)}{||h||} = \lim_{h\to\emptyset} \frac{||h||^2}{||h||} = \lim_{h\to\emptyset} ||h|| = \emptyset$$

$\therefore$ La función f es diferenciable en $p = (p_1,p_2)(\forall\ p \in \mathbb{R}^2)$.

Ejemplo 2. Sea $f(x,y) = 2x-3y$
$p = (p_1,p_2) \in \mathbb{R}^2$ arbitrario $f(p+h) - f(p)=$
$= 2(p_1+h_1) - 3(p_2+h_2) - (2p_1-3p_2) = 2h_1-3h_2$
si $L_{f,p}(h) = L_{f,p}(h_1,h_2) = 2h_1 - 3h_2$
tenemos que $f(p+h)-f(p)-L_{f,p}(h)=2h_1-3h_2-(2h_1-3h_2)=\emptyset$

$$\therefore \lim_{h\to\emptyset} \frac{f(p+h)-f(p)-L_{f,p}(h)}{||h||} = \lim_{h\to\emptyset} \frac{\emptyset}{||h||} = \emptyset.$$

$\therefore$ la función f es diferenciable en $p = (p_1,p_2)(\forall\ p \in \mathbb{R}^2)$.
Obsérvese que $L_{f,p} = f$ para cualquier $p = (p_1,p_2)$.

De hecho el ejemplo anterior es un caso particular de una afirmación más general:

Ejemplo 3. Una función lineal $L:\mathbb{R}^n\to\mathbb{R}$ es diferenciable en todo punto. Pues para cada $p \in \mathbb{R}^n$, $L_{L,p} = L$ es lineal y satisface

$$\lim_{h\to\emptyset} \frac{L(p+h)-L(p)-L_{L,p}(h)}{||h||} = \lim_{h\to\emptyset} \frac{L(p+h)-L(p)-L(h)}{||h||}$$

$$= \lim_{h\to\emptyset} \frac{L(p)+L(h)-L(p)-L(h)}{||h||} = \emptyset \text{ ."}$$

Recordemos que la introducción de la diferenciabilidad está motivada, según la idea del presente trabajo, por el hecho de que, cuando se abordó el estudio de las funciones reales de varias variables (cfr. capítulo 3) con la idea de tener para estas funciones una generalización del concepto de derivada que teníamos para funciones de $\mathbb{R}$ en $\mathbb{R}$; los candidatos para esta generalización que se manejaron entonces fueron desechados pues no implicaban la continuidad de la función en el punto, propiedad deseable en una generalización del concepto de derivada.

Como nuestra intención es hacer ver a los estudiantes que la generalización buscada es justamente la existencia de la diferencial de la función en el punto, habría que ir apuntando hacia esa dirección.

Para ello, como primer paso, hay que ver que la diferenciabilidad de una función en el punto sí implica la continuidad:

"Teorema 1. Una función $f:D\subseteq\mathbb{R}^n\to\mathbb{R}$ diferenciable en $p \in D$ es continua en p.

Por demostrar $\lim\limits_{h\to 0} f(p+h) = f(p)$

Dem. Al ser f diferenciable en p, existe $L_{f,p}:\mathbb{R}\to\mathbb{R}$ lineal tal que

$$\lim_{h\to 0} \frac{f(p+h) - f(p) - L_{f,p}(h)}{||h||} = 0 \overset{(*)}{\Rightarrow} \quad \text{en particular}$$

$$\lim_{h\to 0} (f(p+h)-f(p)-L_{f,p}(h)) = 0$$

Como las funciones lineales son continuas $\Rightarrow \lim\limits_{h\to 0} L_{f,p}(h) =$ $L_{f,p}(0) = 0$ (por ser $L_{f,p}$ lineal)

$$\therefore \lim_{h\to 0} (f(p+h)-f(p)) - \lim_{h\to 0} L_{f,p}(h) = 0$$

$$\Rightarrow \lim_{h\to 0} (f(p+h)-f(p)) = 0 \quad \therefore \quad \lim_{h\to 0} f(p+h) = f(p) \text{ l.c.q.d."}$$

La implicación (*) tiene relevancia pues de alguna forma se hace referencia a ella con frecuencia. Por eso hay que justificarla con detalle:

"Sea $f: \mathbb{R}^n \to \mathbb{R}$ si $\lim_{x\to 0} \frac{f(x)}{||x||} = 0$ entonces $\lim_{x\to 0} f(x)=0$

Tenemos que probar que:

$\forall\ \varepsilon > 0,\ \exists\ \delta > 0$ tal que si $||x|| < \delta$ entonces $|f(x)| < \varepsilon$.

Dem. Dada $\varepsilon > 0$, como $\lim_{x\to 0} \frac{f(x)}{||x||} = 0$, es posible encontrar $\delta_\varepsilon > 0$ tal que

(1)......... si $||x|| \leq \delta_\varepsilon \Rightarrow \frac{|f(x)|}{||x||} < \varepsilon \Rightarrow |f(x)| < \varepsilon\ ||x||$

Sea $\delta = \min\{\delta_\varepsilon, 1\}$. Si $||x|| < \delta \Rightarrow ||x|| \leq \delta_\varepsilon$ y $||x|| \leq 1$ (Por (1)) como $||x|| \leq \delta_\varepsilon$ tenemos $|f(x)| < \varepsilon\ ||x|| \leq \varepsilon(1)=\varepsilon$ l.c.q.d."

Una de las utilidades del Teorema que afirma que una función lineal L que satisfaga

$$\lim_{h\to 0} \frac{f(p+h)-f(p)-L_{f,p}(h)}{||h||} = 0$$

está determinada de manera única, es proporcionar la igualdad

$L_{f,p}(e_j) = \frac{\partial f}{\partial x_j}(p)$ que servirá, por una parte, para trabajar con la diferencial de f en el punto p, y por otra, para entender - en este caso de las funciones de varias variables - la relación entre la nueva noción de diferenciabilidad y los conceptos de derivabilidad que se conocen para estas funciones: existencia de derivadas parciales y direccionales.

"Teorema 2. Sea $f: D \subseteq \mathbb{R}^n \to \mathbb{R}$ diferenciable en $p \in D$. Entonces sólo existe una función lineal $L: \mathbb{R}^n \to \mathbb{R}$ tal que:

$$(*)\quad \lim_{h\to 0} \frac{f(p+h)-f(p)-L(h)}{||h||} = 0$$

Dem. Sea $\{e_1, e_2, \ldots, e_n\}$ la base canónica de $\mathbb{R}^n$. Observamos primero que si $v \in \mathbb{R}^n$ es una dirección i.e. $||v|| = 1$, $||tv|| = |t|$ $t \in \mathbb{R}$ $\therefore$ si t es suficientemente pequeña $p+tv \in D$. Por ello en la definición de diferenciabilidad podemos tomar $h = te_j$ y pedir que $h \to 0$ es equivalente a pedir $t \to 0$.

Como f es diferenciable en p, tenemos para cada j: $1 \leq j \leq n$:

$$0 = \lim_{t \to 0} \frac{f(p+te_j)-f(p)-L_{f,p}(te_j)}{|t|} = \lim_{t \to 0} \frac{f(p+te_j)-f(p)-tL_{f,p}(e_j)}{t}$$

$$\therefore \quad \lim_{t \to 0} \left[\frac{f(p+te_j)-f(p)}{t} - \frac{tL_{f,p}(e_j)}{t} \right] = 0$$

$$\therefore \quad \text{Para } 1 \leq j \leq n \qquad \lim_{t \to 0} \frac{f(p+te_j)-f(p)}{t} = L_{f,p}(e_j) \quad \ldots .(1)$$

Ahora si $L^*: \mathbb{R}^n \to \mathbb{R}$ fuese otra función lineal que satisfaciera (*), repitiendo el argumento anterior probaríamos que

$$\lim_{t \to 0} \frac{f(p+te_j)-f(p)}{t} = L^*(e_j)$$

$$\therefore \quad L_{f,p}(e_j) = L^*(e_j) \quad \text{para } j \in \{1,2,\ldots\ldots,n\}.$$

Pero una función lineal en $\mathbb{R}^n$ está determinada de manera única por su efecto en cualquier base de $\mathbb{R}^n$, en particular en la canónica $\{e_1, e_2, \ldots, e_n\}$. Por lo tanto hemos probado que $L_{f,p} = L^*$.

Si una función $f: D \subseteq \mathbb{R}^n \to \mathbb{R}$ es diferenciable en p, la única función lineal que satisface la condición (*) que hemos denotado por $L_{f,p}$ se llama la diferencial de f en p."

Una vez dada la definición de diferenciabilidad, se puede recorrer el camino inverso que se siguió para presentar dicha definición para las funciones de $\mathbb{R}$ a $\mathbb{R}^n$. En aquella ocasión a partir de la derivabilidad, llegamos a la diferenciabilidad (para luego regresar a la derivabilidad y ver que en ese caso ambos conceptos eran equivalentes). Ahora, a partir de la diferenciabilidad podemos llegar a la derivabilidad; lo cual resulta muy útil pues además de que las derivadas constituyen objetos familiares para los estudiantes, éstas nos proporcionan la manera operativa de trabajar con la diferencial de una función. Y luego en términos de las derivadas reformular la diferenciabilidad que es la manera usual de trabajar: dada una función, inspeccionar sus derivadas parciales y donde éstas sean continuas, sabemos que la función es diferenciable. A su vez, para trabajar con la diferencial lo hacemos a través de la matriz que representa a dicha transformación lineal (la matriz Jacobiana) que involucra precisamente a las derivadas parciales.

En el siguiente texto se les presentarían a los estudiantes los resultados que permiten relacionar la diferenciabilidad con las derivadas, incluyendo el uso de éstas para trabajar con la diferencial:

"La diferenciabilidad de $f:D \subseteq \mathbb{R}^n \to \mathbb{R}$ en $p \in D$, implica que existe una función lineal $L_{f,p} : \mathbb{R}^n \to \mathbb{R}$ tal que $\forall h$ con $p+h \in D$ tenemos que:

$$\lim_{h \to \emptyset} \frac{f(p+h)-f(p)-L_{f,p}(h_1e_1+h_2e_2+....+h_ne_n)}{||h||} = \emptyset$$

$\Rightarrow$

$$\lim_{h \to \emptyset} \frac{f(p+h)-f(p)-(h_1L_{f,p}(e_1)+h_2L_{f,p}(e_2)+....+h_nL_{f,p}(e_n))}{||h||} = \emptyset$$

$\Rightarrow$

$$\lim_{h \to \emptyset} \frac{f(p+h)-f(p)-(h_1,h_2,...,h_n) \, . \, (L_{f,p}(e_1),...,L_{f,p}(e_n))}{||h||} = \emptyset$$

Recordamos que toda trasformación lineal $L \; \mathbb{R}^n \to \mathbb{R}$ es de la forma $L(x) = a \, . \, x \quad x \in \mathbb{R}^n$ para algún $a \in \mathbb{R}^n$, que estaba dado por $a = (a_1,a_2,...,a_n)$ donde $a_i = L(e_i)$ $i = 1, 2, ...n$.

En nuestro caso, la diferencial de f en p: $L_{f,p}$ está dada por $L_{f,p}(x) = (L_{f,p}(e_1), L_{f,p}(e_2),......, L_{f,p}(e_n)) \, . \, x$.

Y la clave para determinar $L_{f,p}(e_j)$ para cada $j=1..n$ nos la da la expresión (1).

$$L_{f,p}(e_j) = \lim_{t \to \emptyset} \frac{f(p+te_j) - f(p)}{t}$$

Reconocemos el lado derecho precisamente como la derivada parcial de f con respecto a x_j en p, i.e. $\frac{\partial f}{\partial x_j}(p)$ es decir, tenemos que $L_{f,p}(e_j) = \frac{\partial f}{\partial x_j}(p)$

Recordamos que las derivadas parciales son derivadas direccionales en las direcciones determinadas por los vectores de la base canónica. Con un argumento análogo al utilizado en el teorema 2, podemos probar el siguiente:

Teorema 3. Si f es diferenciable en p, entonces:
(i) $(D_u f)(p)$ existe para cualquier dirección u.
(ii) $(D_u f)(p) = L_{f,p}(u)$

Dem. Como f es diferenciable en p $\Rightarrow \exists\, L_{f,p}: \mathbb{R}^n \to \mathbb{R}$ una función lineal tal que:

$$\lim_{h\to\emptyset} \frac{f(p+h) - f(p) - L_{f,p}(h)}{||h||} = \emptyset$$

sea $h=tu$ con $||u||=1$ como $||h||=|t|$ $\Rightarrow h \to \emptyset \Leftrightarrow t \to \emptyset$

$$\therefore \lim_{t\to\emptyset} \frac{f(p+tu) - f(p) - L_{f,p}(tu)}{t} = \emptyset$$

$$\Rightarrow \lim_{t\to\emptyset} \left[\frac{f(p+tu) - f(p)}{t} - \frac{tL_{f,p}(u)}{t} \right] = \emptyset$$

$$\Rightarrow \lim_{t\to\emptyset} \left[\frac{f(p+tu) - f(p)}{t} - L_{f,p}(u) \right] = \emptyset$$

$\Rightarrow$ el límite $\lim_{t\to\emptyset} \frac{f(p+tu) - f(p)}{t}$ que es precisamente $(D_u f)(p)$ existe y es $L_{f,p}(u)$.

En particular si $f: D \subseteq \mathbb{R}^n \to \mathbb{R}$ es diferenciable en $p \in D$, entonces las derivadas parciales $\frac{\partial f}{\partial x_j}(p)$ existen para cada $1 \leq j \leq n$ y $\frac{\partial f}{\partial x_j}(p) = L_{f,p}(e_j)$

(basta tomar $u=e_j$ en (ii) del Teorema 3).

La matriz de 1xn que representa a la transformación lineal $L_{f,p}: \mathbb{R}^n \to \mathbb{R}$ (la diferencial de f en p) con respecto a las bases canónicas, que denotamos por $J_{f,p}$ está dada por:

$$J_{f,p} = \begin{bmatrix} L_{f,p}(e_1) & L_{f,p}(e_2)\ldots\ldots L_{f,p}(e_n) \end{bmatrix}$$

$$= \begin{bmatrix} \frac{\partial f}{\partial x_1}(p) & \frac{\partial f}{\partial x_2}(p) & \ldots\ldots\ldots\ldots & \frac{\partial f}{\partial x_n}(p) \end{bmatrix}$$

y se llama la matriz Jacobiana de f en p.

Podemos también formular la definición de diferenciabilidad en términos de las derivadas parciales.

Teorema 4. La función $f: D \subseteq \mathbb{R}^n \to \mathbb{R}$ es diferenciable en $p \in D$ si y sólo si:

(i) Todas las derivadas parciales de f en p existen y

(ii) $$\lim_{h\to\emptyset} \frac{f(p+h)-f(p)-\left(\frac{\partial f}{\partial x_1}(p),\frac{\partial f}{\partial x_2}(p),\ldots,\frac{\partial f}{\partial x_n}(p)\right).(h_1,h_2,\ldots,h_n)}{||h||} = \emptyset$$

Dem. Si la función es diferenciable todas las derivadas parciales de f en p existen, y además:

$$\emptyset = \lim_{h\to\emptyset} \frac{f(p+h) - f(p) - L_{f,p}(h)}{||h||} =$$

$$= \lim_{h\to\emptyset} \frac{f(p+h)-f(p)-\left(\frac{\partial f}{\partial x_1}(p),\ldots,\frac{\partial f}{\partial x_n}(p)\right).(h_1,\ldots,h_n)}{||h||}$$

Inversamente si todas las derivadas parciales de f en p:

$\dfrac{\partial f}{\partial x_j}(p)$ $1 \leq j \leq n$ existen y se satisface (ii); nos fijamos en que la función definida por:

$$L(h) = \left(\frac{\partial f}{\partial x_1}(p), \ldots, \frac{\partial f}{\partial x_n}(p)\right) . (h_1, \ldots, h_n) \ \forall h=(h_1, \ldots, h_n) \in \mathbb{R}^n$$

resulta ser una función lineal que satisface las hipótesis del Teorema 2, y por ese Teorema, esta aproximación lineal debe ser $L_{f,p}$ y $\therefore$ f es diferenciable en p.

Ejemplo 4. Consideremos la función f: $\mathbb{R}^2 \to \mathbb{R}$ dada por $f(x_1,x_2) = x_1^2 + x_2^2$. Sean $x = (x_1,x_2)$ $h = (h_1,h_2)$.
$f(x+h)-f(x) = 2x_1h_1+2x_2h_2+h_1^2+h_2^2 = 2x_1h_1+2x_2h_2+||h||^2$.

Como $\dfrac{\partial f}{\partial x_1}(x) = 2x_1$ y $\dfrac{\partial f}{\partial x_2}(x) = 2x_2$, la matriz Jacobiana de f en x es: $J_{f,x} = [2x_1 \ 2x_2]$

$\therefore L_{f,x}(h) = (2x_1, 2x_2) . (h_1,h_2) = 2x_1h_1 + 2x_2h_2$

$\therefore f(x+h) - f(x) - L_{f,x}(h) = ||h||^2$

$$\Rightarrow \lim_{h\to 0} \frac{f(x+h)-f(x)-L_{f,x}(h)}{||h||} = \lim_{h\to 0} \frac{||h||^2}{||h||} = 0.$$

El Teorema 4 nos dice que si alguna de las derivadas parciales de f en p, no existe, entonces f no es diferenciable en p. ($f: D \subseteq \mathbb{R}^n \to \mathbb{R}$; $p \in D$).

Ejemplo 5. Sea f: $\mathbb{R}^n \to \mathbb{R}$ definida por $f(x) = ||x||$. La derivada parcial $\dfrac{\partial f}{\partial x_1}(0)$ en caso de existir, se calcula considerando la expresión:

$$\frac{f(0 + te_1) - f(0)}{t} = \frac{|t|}{t}$$

pero $\lim\limits_{t\to 0} \dfrac{|t|}{t}$ no existe y por ello f no es diferenciable en 0.

Es interesante recordar que el recíproco del Teorema 4 (i) no es cierto: la existencia de las derivadas parciales de una función $f: D \subseteq \mathbb{R}^n \to \mathbb{R}$ en $p \in D$, no es suficiente para asegurar que f sea diferenciable en p.

Recuérdese el ejemplo analizado en el capítulo 3.

$$f(x,y) = \begin{cases} \dfrac{x\,y}{x^2+y^2} & (x,y) \neq (0,0) \\ 0 & \text{si } (x,y) = (0,0) \end{cases}$$

existían las parciales de f en $(0,0)$ y sin embargo la función no era continua y por lo tanto no es diferenciable en $(0,0)$."

En las funciones de $\mathbb{R}$ a $\mathbb{R}^n$, derivabilidad y diferenciabilidad resultaban ser conceptos equivalentes, en este caso en sentido estricto no es así, como se les recuerda a los estudiantes al final del texto anterior. Sin embargo, en la práctica las funciones con las que habitualmente se trabaja son C^1, al menos en algún subconjunto del dominio, y son por tanto diferenciables en dicho subconjunto.

Por ello interesaría presentar este resultado, es decir, bajo qué condiciones derivabilidad implica diferenciabilidad. Además de mencionar también el hecho de que en la práctica las funciones que se manejan cumplen esa condición, i.e., no sólo se da la mera existencia de las derivadas parciales, sino que dichas derivadas son continuas en todo un subconjunto del dominio de la función.

Por otra parte, con ese resultado, se introduce; pues se requiere como lema para la demostración, una primera versión del Teorema del Valor Medio para funciones de varias variables, que tiene un interés y utilidad propios.

"Consideremos el siguiente ejemplo:

Ejemplo 6. Sea $f: \mathbb{R}^3 \to \mathbb{R}$ definida por: $f(x,y,z) = x^2y + 2z$. En este ejemplo usaremos la notación (x,y,z) en vez de (x_1,x_2,x_3).

Las derivadas parciales de f están dadas por:

$$\frac{\partial f}{\partial x}(p) = 2xy \qquad \frac{\partial f}{\partial y}(p) = x^2 \qquad \frac{\partial f}{\partial z}(p) = 2 \qquad \text{donde } p = (x,y,z)$$

Para determinar si f es diferenciable en $p \in \mathbb{R}^3$, debemos preguntarnos si la función lineal $L_{f,p}: \mathbb{R}^3 \to \mathbb{R}$ dada por

$$L_{f,p}(h)=\left(\frac{\partial f}{\partial x}(p),\frac{\partial f}{\partial y}(p),\frac{\partial f}{\partial z}(p)\right) \cdot (h_1,h_2,h_3) \qquad h=(h_1,h_2,h_3)\in\mathbb{R}^3$$

i.e. $$L_{f,p}(h) = \frac{\partial f}{\partial x}(p)h_1 + \frac{\partial f}{\partial y}(p)h_2 + \frac{\partial f}{\partial z}(p)h_3 =$$

$= 2xyh_1 + x^2h_2 + 2h_3$ satisface que:

$$\lim_{h\to\emptyset} \frac{f(p+h) - f(p) - L_{f,p}(h)}{||h||} = \emptyset$$

Ahora $f(p+h)-f(p)=2xyh_1+x^2h_2+2h_3+2xh_1h_2+h_1^2y+h_1^2h_2$

$\therefore \quad f(p+h) - f(p) - L_{f,p}(h) = 2xh_1h_2 + h_1^2y + h_1^2h_2$

i.e. $R_p(h) = 2xh_1h_2 + h_1^2y + h_1^2h_2$

Para concluir la diferenciabilidad de f en p necesitaríamos ver que:

$$\emptyset = \lim_{h\to\emptyset} \frac{R_p(h)}{||h||} = \lim_{h\to\emptyset} \frac{2xh_1h_2 + h_1^2y + h_1^2h_2}{||h||}$$

como $|hi| \leq ||h||$ $i = 1,2,3$ (pues para $i = 1,2,3$ $|hi| = (h_i^2)^{1/2} \leq (h_1^2+h_2^2+h_3^2)^{1/2} = ||h||$) tenemos que:

$$\emptyset \leq \frac{|R_p(h)|}{||h||} = \frac{|2xh_1h_2 + h_1^2y + h_1^2h_2|}{||h||} \leq$$

$$\frac{1}{||h||}\left[2|x| \ |h_1| \ |h_2| + |y| \ |h_1|^2 + |h_1|^2 \ |h_2| \right]$$

$$\leq \frac{1}{||h||}\left[2|x| \ ||h||^2 + |y| \ ||h||^2 + ||h||^3 \right] =$$

$$||h|| \ (2|x| + |y| + ||h|| \ )$$

$$\text{i.e.}\quad \emptyset \le \frac{|R_p(h)|}{||h||} \le ||h|| \ (2|x| + |y| + ||h|| \ )$$

$$\therefore \quad \emptyset \le \lim_{h\to\emptyset} \frac{R_p(h)}{||h||} \le \lim_{h\to\emptyset} ||h|| \ (2|x| + |y| + ||h|| \ ) = \emptyset$$

$$\text{i.e.}\quad \lim_{h\to\emptyset} \frac{R_p(h)}{||h||} = \emptyset = \lim_{h\to\emptyset} \frac{f(p+h)-f(p)-L_{f,p}(h)}{||h||}$$

$\therefore$ f es diferenciable en p.

Obsérvese que diferentes elecciones del punto $p = (x,y,z) \in \mathbb{R}^3$ dan lugar a diferentes expresiones para $R_p(h)$, pero en cada caso, la conclusión es válida.

En el ejemplo 6, las derivadas parciales

$$\frac{\partial f}{\partial x} : \mathbb{R}^3 \to \mathbb{R} \quad \text{definida como:} \quad (x,y,z) \xrightarrow{\ \frac{\partial f}{\partial x}\ } 2xy$$

$$\frac{\partial f}{\partial y} : \mathbb{R}^3 \to \mathbb{R} \quad \text{definida como:} \quad (x,y,z) \xrightarrow{\ \frac{\partial f}{\partial y}\ } x^2$$

$$\frac{\partial f}{\partial z} : \mathbb{R}^3 \to \mathbb{R} \quad \text{definida como:} \quad (x,y,z) \xrightarrow{\ \frac{\partial f}{\partial z}\ } 2$$

son funciones continuas en $\mathbb{R}^3$ y la función f resultó ser diferenciable para todo punto de $\mathbb{R}^3$. Esta es una propiedad relevante que nos proporcionará un recíproco parcial del Teorema 4 (i).

Necesitaremos un lema previo, que constituye una primera versión del Teorema del Valor Medio para funciones de varias variables.

Lema 1

Sea $f: D \subseteq \mathbb{R}^n \to \mathbb{R}$ donde $D = B_\varepsilon(p)$ para algún

$p \in \mathbb{R}^n$ y $\varepsilon > 0$ ($\varepsilon \in \mathbb{R}$). Supongamos que las derivadas parciales de f (de primer orden) existen y son continuas en D. Sea $h = (h_1, h_2, \ldots, h_n) \in \mathbb{R}^n$ tal que $p+h \in D$, i.e. $||h|| < \varepsilon$ (Ver Figura 5.10).

Entonces existen puntos $p_1, p_2, \ldots, p_n \in D$ tales que

$$(*) \quad f(p+h)-f(p) = \frac{\partial f}{\partial x_1}(p_1)h_1 + \frac{\partial f}{\partial x_2}(p_2)h_2 + \ldots\ldots + \frac{\partial f}{\partial x_n}(p_n)h_n$$

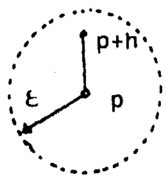


Figura 5.10
$||h|| < \varepsilon$

N.B. En esta versión cada una de las derivadas parciales $\frac{\partial f}{\partial x_i}$ está evaluada en un punto diferente $p_i \in D$. Se puede mejorar este resultado probando que los puntos p_i pueden ser todos reeplazados por un único punto p^* convenientemente elegido que está en el segmento de recta que une p y p+h.

Dem. La prueba del lema se hace para el caso de dos dimensiones, pues esto ilustra el método general. La notación para este caso será $h = (h_1, h_2)$ y $p = (x, y)$.

Sea q el punto $p+(h_1, 0) = (x+h_1, y)$ y escribamos:

$$f(p+h)-f(p) = (f(p+h)-f(q)) + (f(q)-f(p)) =$$
$$= (f(x+h_1, y+h_2) - f(x+h_1, y)) + (f(x+h_1, y) - f(x, y)).$$

Se observa que en cada término entre paréntesis, sólo una variable ha sido alterada, aplicamos a cada término el Teorema del Valor Medio para funciones de una variable, posible pues cada derivada parcial existe. Podemos entonces asegurar la existencia de una x^* entre x y $x+h_1$ y de una y^* entre y e $y+h_2$, obteniendo puntos p' y p'' tales que:

$$f(q)-f(p)=f(x+h_1,y)-f(x,y) = h_1 \frac{\partial f}{\partial x}(x^*,y) = \frac{\partial f}{\partial x}(p')\, h_1 \ldots\ldots (a)$$

$$f(p+h)-f(q)=f(x+h_1,y+h_2)-f(x+h_1,y) =$$

$$= h_2 \frac{\partial f}{\partial y}(x+h_1,y^*) = \frac{\partial f}{\partial y}(p'')\, h_2 \ldots\ldots (b)$$

(Ver Figura 5.11)

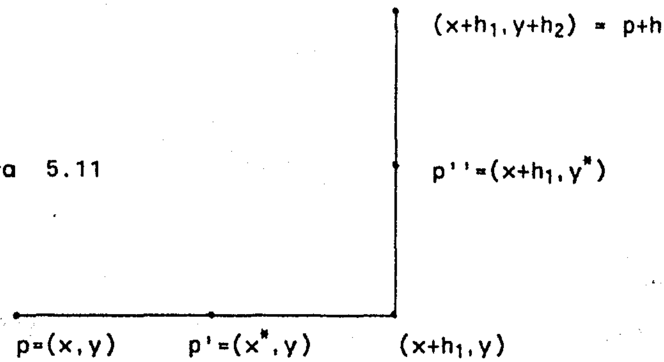


Figura 5.11

Sumando (a) y (b) tenemos que:

$$f(p+h) - f(p) = \frac{\partial f}{\partial x}(p')h_1 + \frac{\partial f}{\partial y}(p'')h_2$$

Que es la expresión para dos variables de (*).

Con este lema, estamos en condiciones de probar un recíproco parcial del Teorema 4 (i).

Teorema 5. Sea $f\colon D \subseteq \mathbb{R}^n \to \mathbb{R}$ una función cuyas derivadas parciales existen en $B_\varepsilon(p) \subseteq D$ $(p \in D)$ si

$\dfrac{\partial f}{\partial x_i} : B_\varepsilon(p) \subseteq \mathbb{R}^n \to \mathbb{R}$ es continua en p, para i = 1,2,3 entonces f es diferenciable en p.

Dem. Debemos probar que existe una transformación lineal $L_{f,p}$ tal que:

$$\lim_{h\to 0} \frac{f(p+h) - f(p) - L_{f,p}(h)}{||h||} = \lim_{h\to 0} \frac{R_p(h)}{||h||} = 0$$

Como las parciales de f en p existen podemos formar el vector

$$\left(\frac{\partial f}{\partial x_1}(p), \frac{\partial f}{\partial x_2}(p), \ldots\ldots, \frac{\partial f}{\partial x_n}(p) \right)$$

y considerar la función lineal definida por

$$L_{f,p}(h) = \left(\frac{\partial f}{\partial x_1}(p), \frac{\partial f}{\partial x_2}(p), \ldots\ldots, \frac{\partial f}{\partial x_n}(p) \right) . (h_1, h_2, \ldots\ldots, h_n)$$

con $h = (h_1, h_2, \ldots\ldots, h_n)$.

Para el caso de tres variables $p = (x,y,z)$; $h = (h_1,h_2,h_3)$ lo que necesitamos ver es que:

si (1).....$R_p(h) = f(p+h)-f(p)- \frac{\partial f}{\partial x}(p)h_1 - \frac{\partial f}{\partial y}(p)h_2 - \frac{\partial f}{\partial z}(p)h_3$

entonces $\lim\limits_{h \to \emptyset} \frac{R_p(h)}{||h||} = \emptyset$

Probaremos el Teorema 5 para este caso, pues el argumento fácilmente se generaliza a n variables.

Por el lema 1 tenemos que

$$f(p+h) - f(p) = \frac{\partial f}{\partial x}(p')h_1 + \frac{\partial f}{\partial y}(p'')h_2 + \frac{\partial f}{\partial z}(p''')h_3$$

Donde p', p'', p''' son puntos en cada uno de los tres segmentos que forman una poligonal de p a $p+h$. Si $p+h \in B_\varepsilon(p)$ esta poligonal está contenida en $B_\varepsilon(p)$.

Regresando a (1) tenemos entonces que:

$$R_p(h) = \left[\frac{\partial f}{\partial x}(p') - \frac{\partial f}{\partial x}(p) \right] h_1 + \left[\frac{\partial f}{\partial y}(p'') - \frac{\partial f}{\partial y}(p) \right] h_2 +$$

$$+ \left[\frac{\partial f}{\partial z}(p''') - \frac{\partial f}{\partial z}(p) \right] h_3$$

como $|hi| \leq ||h||$, $i = 1,2,3$

$$\emptyset \leq \frac{|R_p(f)|}{||h||} \leq \frac{1}{||h||} \left[\left| \frac{\partial f}{\partial x}(p') - \frac{\partial f}{\partial x}(p) \right| |h_1| + \right.$$

$$+\left|\frac{\partial f}{\partial y}(p'') - \frac{\partial f}{\partial y}(p)\right| |h_2| + \left|\frac{\partial f}{\partial z}(p''') - \frac{\partial f}{\partial z}(p)\right| |h_3| \Bigg]$$

$$\Rightarrow (2)\ldots\ldots \emptyset \leq \frac{|R_p(h)|}{||h||} \leq \left|\frac{\partial f}{\partial x}(p') - \frac{\partial f}{\partial x}(p)\right| + \left|\frac{\partial f}{\partial y}(p'') - \frac{\partial f}{\partial y}(p)\right| +$$

$$+ \left|\frac{\partial f}{\partial z}(p''') - \frac{\partial f}{\partial z}(p)\right|$$

si $h \to \emptyset \Rightarrow p' \to p,\ p'' \to p,\ p''' \to p$

y como $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}$ son continuas en p

$$\lim_{h \to \emptyset} \left|\frac{\partial f}{\partial x}(p') - \frac{\partial f}{\partial x}(p)\right| = \lim_{p' \to p} \left|\frac{\partial f}{\partial x}(p') - \frac{\partial f}{\partial x}(p)\right|$$

$$= \left|\lim_{p' \to p} \frac{\partial f}{\partial x}(p') - \frac{\partial f}{\partial x}(p)\right| = \left|\frac{\partial f}{\partial x}(p) - \frac{\partial f}{\partial x}(p)\right| = \emptyset$$

Análogamente con los otros dos términos del lado derecho de(2)

$$\therefore \lim_{h \to \emptyset} \frac{R_p(h)}{||h||} = \emptyset \Rightarrow \text{f es diferenciable en p.}$$

Si recordamos la definición de función de clase C^k en D, el Teorema 5 nos dice que cualquier función de clase C^1 en D es diferenciable en D. Donde por diferenciable en D, entendemos que para cada punto $p \in D$ existe una función lineal $L_{f,p} : \mathbb{R}^n \to \mathbb{R}$ tal que:

$$\lim_{h \to \emptyset} \frac{f(p+h)-f(p)-L_{f,p}(h)}{||h||} = \emptyset \text{ "}$$

A fin de facilitarle al estudiante el tener clara la relación entre los conceptos de derivabilidad y diferenciabilidad se le puede presentar, a modo de resumen el esquema siguiente. Se incluyen también, para facilitar la consulta, una colección de contraejemplos, para que le sirvan de referencia sobre la invalidez de las inversas de algunas implicaciones.

"Es importante tener clara la relación entre los diferentes conceptos que hemos estudiado, estas relaciones quedan resumidas en el siguiente esquema:

	Teorema 6		Teorema 3		Definición de derivada direccional	
Parciales Continuas (de f en p)	=>	Diferenciabilidad (de f en p)	=>	Existencia de las derivadas direccionales (de f en p)	=>	Existencia de las derivadas parciales (de f en p)

Cada afirmación recíproca obtenida al invertir una flecha no es válida, i.e.

Existencia de las derivadas parciales de f en p $\not\Rightarrow$ Existencia de las derivadas direccionales de f en p.

Ejemplo 7.

$$\text{sea } f(x,y) = \begin{cases} \dfrac{xy}{x^2+y^2} & (x,y) \neq (\emptyset,\emptyset) \\ \emptyset & \text{si } (x,y) = (\emptyset,\emptyset) \end{cases}$$

Se probó que

$$\frac{\partial f}{\partial x}(\emptyset,\emptyset) \text{ y } \frac{\partial f}{\partial y}(\emptyset,\emptyset)$$

existían (ambas eran iguales a $\emptyset$). (cfr. capítulo 3). Veamos ahora que salvo para las direcciones de los ejes, no existen las derivadas direccionales de f en $(\emptyset,\emptyset)$.

Debemos considerar $\dfrac{f(tv) - f(\emptyset,\emptyset)}{t}$

con $v = (\cos\theta, \operatorname{sen}\theta)$

$$\therefore \quad \frac{f(tv) - f(\emptyset,\emptyset)}{t} = \frac{1}{t}\left[\frac{t^2 \cos\theta \operatorname{sen}\theta}{t^2} - \emptyset\right] = \frac{\cos\theta \operatorname{sen}\theta}{t}$$

si v determina una dirección diferente de cualquiera de los dos ejes $\Rightarrow \cos\theta \operatorname{sen}\theta \neq \emptyset$

$$\therefore (D_v f)(\emptyset,\emptyset) = \lim_{t\to\emptyset} \frac{f(tv) - f(\emptyset,\emptyset)}{t} = \lim_{t\to\emptyset} \frac{\cos\theta \operatorname{sen}\theta}{t}$$

pero el límite $\displaystyle\lim_{t\to\emptyset} \frac{\cos\theta \operatorname{sen}\theta}{t}$ no existe

$\therefore$ si $v \neq \pm(1,\emptyset)$ ó $v \neq \pm(\emptyset,1)$ $(D_v f)(\emptyset,\emptyset)$ no existe.

Existencia de las derivadas direccionales (de f en p) $\not\Rightarrow$ Diferenciabilidad (de f en p)

Ejemplo 8.

$$\text{sea } f(x,y) = \begin{cases} \dfrac{xy^2}{x^2+y^4} & \text{si } x \neq \emptyset \\ \emptyset & \text{si } x = \emptyset \end{cases}$$

Este ejemplo se trabajó en el capítulo 3, vimos que $(D_u f)(\emptyset,\emptyset)$ existía para todas las direcciones u, y sin embargo f no era continua en $(\emptyset,\emptyset)$. Por el Teorema 1, f no es diferenciable en $(\emptyset,\emptyset)$ y $(D_u f)(\emptyset,\emptyset)$ existe para toda $u \in \mathbb{R}^2 \ni ||u|| = 1$

Diferenciadilidad de f en p. $\not\Rightarrow$ Las parciales de f sean continuas en p.

Ejemplo 9.

Consideremos la función $g:\mathbb{R}\to\mathbb{R}$ definida por

$$g(x) = \begin{cases} x^2 \operatorname{sen}(1/x) & \text{si } x \neq 0 \\ 0 & \text{si } x = 0 \end{cases}$$

Para encontrar g', si $x \neq 0$ $g(x) = x^2 \operatorname{sen}(1/x)$ y aplicamos las reglas de derivación usuales para obtener

$$g'(x) = 2x \operatorname{sen}(1/x) - \cos(1/x) \qquad (\text{si } x \neq 0)$$

Para calcular $g'(0)$ usamos la definición

$$g'(0) = \lim_{h\to 0} \frac{g(h) - g(0)}{h}$$

Para ello:

$$\frac{g(h) - g(0)}{h} = \frac{h^2 \operatorname{sen}(1/h)}{h} = h \operatorname{sen}(1/h)$$

$$\therefore \quad g'(0) = \lim_{h\to 0} \frac{g(h) - g(0)}{h} = \lim_{h\to 0} h \operatorname{sen}(1/h) = 0$$

$$\left(0 \le \left|h \operatorname{sen}\frac{1}{h}\right| \le |h| \Rightarrow \lim_{h\to 0} \left|h \operatorname{sen}\frac{1}{h}\right| = 0 \Rightarrow \lim_{h\to 0} h \operatorname{sen}\frac{1}{h} = 0 \right)$$

$$\therefore \quad g'(x) = \begin{cases} 2x \operatorname{sen}(1/x) - \cos(1/x) & \text{si } x \neq 0 \\ 0 & \text{si } x = 0 \end{cases}$$

Queremos ver que g' no es continua en 0; para que lo fuese necesitaríamos que $\lim_{x\to 0} g'(x) = g'(0) = 0$

pero $\lim_{x\to 0} g'(x) = \lim_{x\to 0} (2x \operatorname{sen}(1/x) - \cos(1/x))$

y como $\lim_{x\to 0} \cos(1/x)$ no existe $\Rightarrow$ $\lim_{x\to 0} g'(x)$ no existe $\Rightarrow$ g' no es continua en 0.

En conclusión g es diferenciable en $\emptyset$ (La diferenciabilidad de g en $\emptyset$ consiste en la existencia del límite $\lim_{h\to\emptyset} \frac{g(h) - g(\emptyset)}{h}$ pues para las funciones de $\mathbb{R}$ en $\mathbb{R}$, la existencia de dicho límite implica la existencia de la aproximación lineal) pero la derivada parcial de g (que en este caso sólo hay una, a saber $g' = \frac{dg}{dx}$) no es continua en $\emptyset$."

A fin de cuentas, lo que resulta necesario, es que el alumno tenga claro el "mecanismo operativo" a seguir para saber en qué puntos una función $f:D \subseteq \mathbb{R}^n \to \mathbb{R}$ es diferenciable y cómo calcular la diferencial en esos puntos.

Para ello se les puede recomendar la metodología habitual. Por otra parte, este modo de proceder proporciona una manera más nítida de presentar la idea de la diferencial total que se mencionó en el capítulo 3.

"En la práctica se sugiere proceder de la siguiente forma. Dada $f:D \subseteq \mathbb{R}^n \to \mathbb{R}$ se obtienen las derivadas parciales; se ve para qué puntos de D las derivadas parciales resultan ser funciones continuas. Sabemos entonces que la función dada será diferenciable en esos puntos. (Obsérvese que f puede ser diferenciable en otros puntos; pero en la práctica esos casos son muy poco frecuentes). Para los puntos p en los que f es diferenciable va a existir la transformación lineal $L_{f,p}:\mathbb{R}^n \to \mathbb{R}$ que está representada por la matriz $J_{f,p}$ (la matriz Jacobiana de f en p), donde la matriz $J_{f,p}$ está dada por:

$$J_{f,p} = \left[\frac{\partial f}{\partial x_1}(p) \quad \frac{\partial f}{\partial x_2}(p) \ldots\ldots \frac{\partial f}{\partial x_n}(p) \right]$$

Por tanto, la diferencial de f en p, $L_{f,p}:\mathbb{R}^n \to \mathbb{R}$ está definida por:

$$L_{f,p}(h) = h_1 \frac{\partial f}{\partial x_1}(p) + h_2 \frac{\partial f}{\partial x_2}(p) + \ldots + h_n \frac{\partial f}{\partial x_n}(p)$$

donde $h = (h_1, h_2, \ldots, h_n)$
(o bien como multiplicación de matrices:

$$L_{f,p}(h) = \left[\frac{\partial f}{\partial x_1}(p) \ldots \frac{\partial f}{\partial x_n}(p) \right] \begin{bmatrix} h_1 \\ \cdot \\ \cdot \\ \cdot \\ h_n \end{bmatrix}$$

Se ilustra el procedimiento anterior con dos ejemplos específicos:

Ejemplo 1Ø.

Sea $f: \mathbb{R}^3 \to \mathbb{R}$ dada por $f(x,y,z) = x^2y + xe^z$

$$\frac{\partial f}{\partial x}(x,y,z) = 2xy + e^z, \quad \frac{\partial f}{\partial y}(x,y,z) = x^2, \quad \frac{\partial f}{\partial z}(x,y,z) = xe^z$$

Las derivadas parciales son funciones continuas en todo punto $p = (x,y,z)$ de $\mathbb{R}^3 \Rightarrow f$ es diferenciable en $\mathbb{R}^3$. Si queremos calcular la diferencial de f, digamos en los siguientes puntos: (i) (Ø,Ø,Ø) (ii) (1,Ø,2) (iii) (a,b,c)

(i)
$$\left.\begin{array}{l} \dfrac{\partial f}{\partial x}(\emptyset,\emptyset,\emptyset) = 1 \\[2ex] \dfrac{\partial f}{\partial y}(\emptyset,\emptyset,\emptyset) = \emptyset \\[2ex] \dfrac{\partial f}{\partial z}(\emptyset,\emptyset,\emptyset) = \emptyset \end{array}\right\} \Rightarrow$$

La matriz que representa a $L_{f,(\emptyset,\emptyset,\emptyset)}$ es

$$J_{f,(\emptyset,\emptyset,\emptyset)} = [\, 1 \ \emptyset \ \emptyset \,]$$

$$\therefore L_{f,(\emptyset,\emptyset,\emptyset)}(h_1,h_2,h_3) = [\, 1 \ \emptyset \ \emptyset \,] \begin{bmatrix} h_1 \\ h_2 \\ h_3 \end{bmatrix} = h_1$$

Obsérvese que podemos cambiar de notación para denotar a las variables de la función lineal $L_{f,p}$ de una manera más usual, concretamente:

$$L_{f,(\emptyset,\emptyset,\emptyset)}:\mathbb{R}^3\to\mathbb{R}$$

$$L_{f,(\emptyset,\emptyset,\emptyset)}(x,y,z) = [\,1\ \emptyset\ \emptyset\,]\begin{bmatrix} x \\ y \\ z \end{bmatrix} = x \qquad (x,y,z)\in\mathbb{R}^3$$

Por ejemplo:

$$L_{f,(\emptyset,\emptyset,\emptyset)}(3,1,-2) = [\,1\ \emptyset\ \emptyset\,]\begin{bmatrix} 3 \\ 1 \\ -2 \end{bmatrix} = 3(1)+1(\emptyset)+(-2)(\emptyset)=3$$

(ii)
$$\left.\begin{aligned} \frac{\partial f}{\partial x}(1,\emptyset,2) &= e^2 \\ \frac{\partial f}{\partial y}(1,\emptyset,2) &= 1 \\ \frac{\partial f}{\partial z}(1,\emptyset,2) &= e^2 \end{aligned}\right\}$$
=> La matriz que representa a $L_{f,(1,\emptyset,2)}$ es

$$J_{f,(1,\emptyset,2)} = [e^2\ 1\ e^2]$$

$$\Rightarrow L_{f,(1,\emptyset,2)}(x,y,z) = [e^2\ 1\ e^2]\begin{bmatrix} x \\ y \\ z \end{bmatrix} = e^2x+y+e^2z$$

Si queremos calcular $L_{f,(1,\emptyset,2)}(3,-4,5)$ tendríamos que:
$L_{f,(1,\emptyset,2)}(3,-4,5) = 3e^2 + (-4) + e^2 5 = 8e^2 - 4$

(iii)
$$\left.\begin{aligned} \frac{\partial f}{\partial x}(a,b,c) &= 2ab+e^c \\ \frac{\partial f}{\partial y}(a,b,c) &= a^2 \\ \frac{\partial f}{\partial z}(a,b,c) &= ae^c \end{aligned}\right\} \Rightarrow$$

$$\Rightarrow L_{f,(a,b,c)}(x,y,z) = [2ab+e^c \;\; a^2 \;\; ae^c]\begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

$$\therefore L_{f,(a,b,c)}(x,y,z) = (2ab+e^c)x + a^2y + ae^cz$$

Donde (a,b,c) es cualquier punto en donde f sea diferenciable (en este caso, puede ser cualquier punto de $\mathbb{R}^3$).

Como vemos: $L_{f,(a,b,c)}(x,y,z) = (2ab+e^c)x + a^2y + ae^cz$

El lado derecho en sentido estricto es función de 6 variables independientes reales a saber $((a,b,c,),(x,y,z))$; (a,b,c) corresponde a un punto en donde la diferencial de la función existe y por lo tanto al sustituir dichas variables por las coordenadas de un punto específico p en donde la función sea diferenciable, obtendríamos una expresión de la forma:

$L_{f,p}(x,y,z) = Ax + By + Cz$ con A, B, C constantes que identificamos como una función lineal en los argumentos x, y, z.

Ejemplo 11: Sea $f(x,y) = 2x - 3y$

$$\left.\begin{array}{l} \dfrac{\partial f}{\partial x}(x,y) = 2 \\[2ex] \dfrac{\partial f}{\partial y}(x,y) = -3 \end{array}\right]$$

Las parciales son funciones constantes, en particular son continuas en todo punto de $\mathbb{R}^2$ $\therefore$ f es diferenciable para todo punto de $\mathbb{R}^2$

Para calcular la diferencial de f en un punto $(a,b) \in \mathbb{R}^2$ tenemos que:

$$L_{f,(a,b)}(x,y) = \begin{bmatrix} \dfrac{\partial f}{\partial x}(a,b) & \dfrac{\partial f}{\partial y}(a,b) \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = [2 \;\; -3]\begin{bmatrix} x \\ y \end{bmatrix} = 2x-3y = $$
$$= f(x,y)$$

i.e. $\forall (a,b) \in \mathbb{R}^2 \quad L_{f,(a,b)} = f$ "

Como la notación para la diferencial de una función que se utiliza en muchos textos es distinta de la que se ha usado sistemáticamente en este capítulo, es oportuno hacer algún comentario al respecto. Este mismo comentario se puede aprovechar para aclarar algún posible vestigio de confusión que todavía pudieran tener los estudiantes con relación a la notación diferencial de Leibniz.

"En el caso de dos variables, si $f: D \subseteq \mathbb{R}^2 \to \mathbb{R}$ y f es diferenciable en $p \in D$, entonces $L_{f,p}: \mathbb{R}^2 \to \mathbb{R}$ está definida por:

$$L_{f,p}(x,y) = \begin{bmatrix} \frac{\partial f}{\partial x}(p) & \frac{\partial f}{\partial y}(p) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

$$\therefore (1)\ldots\ldots L_{f,p}(x,y) = \frac{\partial f}{\partial x}(p)x + \frac{\partial f}{\partial y}(p)y$$

A la expresión 1 (que no es otra cosa que la expresión algebraica de la diferencial de f en p) es a lo que muchos autores denominan la diferencial total de f en p.

Algunos textos en lugar de la notación (x,y) para denotar las coordenadas del punto en donde se evalúa la transformación lineal utilizan (dx,dy) y en vez de $L_{f,(a,b)}$ escriben $df_{(a,b)}$ obteniendo expresiones como:

$$df_{(a,b)} = \frac{\partial f}{\partial x}(a,b)dx + \frac{\partial f}{\partial y}(a,b)dy$$

Donde (dx,dy) denotarían sencillamente las coordenadas del punto en donde se evalúa la transformación lineal $df_{(a,b)} : \mathbb{R}^2 \to \mathbb{R}$ y no tienen nada que ver con la idea vaga de "cantidades infinitamente pequeñas"."

5.3.1 EL GRADIENTE

Después de este estudio más detallado de la diferenciabilidad para funciones de varias variables, el estudiante podrá tener un mejor entendimiento del papel que juega el vector gradiente. Pues con frecuencia, el estudiante aprende en los cursos de Cálculo vectorial, a calcular el vector gradiente y quizás a utilizarlo haciendo uso de su interpretación geométrica. Pero su conocimiento está a un nivel meramente algorítmico, cuando interesaría que supiese porqué resulta especialmente útil considerar el vector formado con las derivadas parciales.

Como una posible introducción del vector gradiente en este contexto de aproximaciones lineales se propone la siguiente:

"Hemos visto que $f: D \subseteq \mathbb{R}^n \to \mathbb{R}$ es diferenciable en $p \in D$, si existe una transformación lineal llamada la diferencial de f en p denotada por $L_{f,p}$, que satisface:

$$\lim_{h \to \emptyset} \frac{f(p+h) - f(p) - L_{f,p}(h)}{||h||} = \emptyset$$

$L_{f,p}: \mathbb{R}^n \to \mathbb{R}$ está representada por la matriz Jocobiana de f en p, i.e. $L_{f,p}(x) = J_{f,p}x \quad x \in \mathbb{R}^n$ (puesto como matriz columna).

$$\text{Donde } J_{f,p} = \begin{bmatrix} L_{f,p}(e_1) & L_{f,p}(e_2) \ldots . L_{f,p}(e_n) \end{bmatrix} =$$

$$\begin{bmatrix} \frac{\partial f}{\partial x_1}(p) & \frac{\partial f}{\partial x_2}(p) \ldots & \frac{\partial f}{\partial x_n}(p) \end{bmatrix} \Rightarrow (1) \ldots . \; L_{f,p}(x_1, x_2, \ldots, x_n) =$$

$$\begin{bmatrix} \frac{\partial f}{\partial x_1}(p) \ldots & \frac{\partial f}{\partial x_n}(p) \end{bmatrix} \begin{bmatrix} x_1 \\ . \\ . \\ . \\ x_n \end{bmatrix} = \frac{\partial f}{\partial x_1}(p)x_1 + \frac{\partial f}{\partial x_2}(p)x_2 + \ldots + \frac{\partial f}{\partial x_n}(p)x_n$$

La expresión (1) puede reescribirse utilizando la noción de producto punto de 2 vectores en $\mathbb{R}^n$ como:

$$L_{f,p}(x_1,x_2,\ldots,x_n) = \left(\frac{\partial f}{\partial x_1}(p),\frac{\partial f}{\partial x_2}(p),\ldots,\frac{\partial f}{\partial x_n}(p) \right).(x_1,x_2,\ldots,x_n)$$

Al vector de $\mathbb{R}^n$ dado por $\left(\frac{\partial f}{\partial x_1}(p),\frac{\partial f}{\partial x_2}(p),\ldots,\frac{\partial f}{\partial x_n}(p) \right)$ i.e. al vector cuyas coordenadas son las entradas de la matriz Jacobiana en el mismo orden; se le denomina el vector gradiente de f en p y se denota (gradf)(p) ó $\nabla f(p)$. Con lo que podemos reescribir (1) como:

(2)... $L_{f,p}(x) = (\mathrm{grad}f)(p).x \quad x \in \mathbb{R}^n$ "

Aquí cabe el peligro de una formulación que puede resultar demasiado abstracta para los estudiantes, sobre la que es oportuno advertir al profesor.

Si f es diferenciable en D, entonces (gradf) (p) = $\left(\frac{\partial f}{\partial x_1}(p),\ldots\ldots,\frac{\partial f}{\partial x_n}(p) \right)$ define una función vectorial gradf: $D \subseteq \mathbb{R}^n \to \mathbb{R}^n$ mediante:

$$p \in D \subseteq \mathbb{R}^n \qquad p \xrightarrow{(\mathrm{grad}f)(p)} \left(\frac{\partial f}{\partial x_1}(p),\ldots,\frac{\partial f}{\partial x_n}(p)\right)$$

Algunos autores hablan entonces de la función gradf: $D \subseteq \mathbb{R}^n \to \mathbb{R}^n$ como "la derivada total de f" o como la "diferencial de f" (en contraposición a la diferencial de f en algún punto) denotándola en ocasiones por Df.

Por esta expresión habría entonces que entender - en el lenguaje de las transformaciones lineales - que Df sería la función (asociación) que a cada $p \in D$ le asigna, no el vector de $\mathbb{R}^n$ dado por (grad f) (p) -que es sencillamente un elemento de $\mathbb{R}^n$ - sino la transformación lineal L de $\mathbb{R}^n$ en $\mathbb{R}$ determinada por dicho vector; a saber:

$L(x) = (\text{grad } f)(p) \cdot x \quad x \in \mathbb{R}^n$ i.e. $Df: D \to \text{Lin}(\mathbb{R}^n, \mathbb{R})$ definida por:

$$p \xrightarrow{Df} L_{f,p} : \mathbb{R}^n \to \mathbb{R} \text{ con } L_{f,p}(x) = (\text{grad } f)(p).x \quad x \in \mathbb{R}^n.$$

donde $\text{Lin}(\mathbb{R}^n, \mathbb{R}) = \{F : \mathbb{R}^n \to \mathbb{R} / F \text{ es lineal}\}$

Entender esta terminología supone una cierta madurez matemática y por ello no se recomendaría introducirla.

Desde luego que al tratar el tema del vector gradiente hay que presentar también su uso en aplicaciones. Como algunas de las más relevantes están:

- Su uso para calcular derivadas direccionales:

"Si $f: D \subseteq \mathbb{R}^n \to \mathbb{R}$ es diferenciable en $p \in D$, $(D_u f)(p)$ existe para todo vector unitario u y además $(D_u f)(p) = L_{f,p}(u)$. En virtud de (2) tenemos que:

(3)................$(D_u f)(p) = (\text{grad } f)(p).u$

Lo cual proporciona una manera especialmente sencilla de calcular las derivadas direccionales; resolvamos nuevamente el problema de calcular $(D_u f)(p)$ donde $f: \mathbb{R}^2 \to \mathbb{R}$ está definida por $f(x,y) = x^2 + 3yx \quad p = (2,\emptyset) \quad u = (1/(2)^{1/2}, -1/(2)^{1/2})$ (cfr. capítulo 3) pero ahora aplicando (3).

$$\frac{\partial f}{\partial x}(x,y) = 2x + 3y, \quad \frac{\partial f}{\partial y}(x,y) = 3x \Rightarrow$$

$$\frac{\partial f}{\partial x}(2,\emptyset) = 4 \quad \frac{\partial f}{\partial y}(2,\emptyset) = 6 \Rightarrow (\text{grad } f)(2,\emptyset) = (4,6)$$

$\therefore (D_u f)(p) = (\text{grad } f)(p) \cdot u = (4,6).(1/(2)^{1/2}, -1/(2)^{1/2}) = 4/(2)^{1/2} - 6/(2)^{1/2} = -2/(2)^{1/2}$."

- Como consecuencia de la anterior, su uso para determinar direcciones de máximo y mínimo cambio de la función:

"Como corolario de (3) tenemos que el gradiente de f(en un punto) es un vector (en caso de que (grad f)(p) $\neq \emptyset$) que apunta en la dirección de máxima razón de cambio y cuya magnitud es el valor máximo posible de las derivadas direccionales en p. Pues la razón de cambio de f en p en una dirección u ($||u|| = 1$) está dada por: $(D_u f)(p) = (\text{grad } f)(p).u = ||(\text{grad } f)(p)|| \cos\theta$ donde θ es el ángulo que forman los vectores u y (grad f)(p). Por lo tanto, el valor de $(D_u f)(p)$ será máximo cuando $\cos\theta = 1 \Rightarrow \theta = \emptyset \Rightarrow$ u sea un vector colineal y apuntando en la misma dirección que (grad f)(p), y será mínimo cuando $\cos\theta = -1 \Rightarrow \theta = \pi \Rightarrow$ u sea un vector colineal que apunte en la dirección contraria a (grad f)(p).

En cualquiera de los dos casos el vector (grad f)(p) apunta en la dirección de máxima razón de cambio instántanea de la función en p. La longitud del vector gradiente es la medida de la razón máxima de crecimiento, pues hemos visto que la razón de cambio es máxima cuando u es un vector unitario en la dirección del vector gradiente

$$(D_u f)(p) \text{ es máxima si } u = \frac{(\text{grad } f)(p)}{||(\text{grad } f)(p)||}$$

en cuyo caso

$$(D_u f)(p) = (\text{grad } f)(p).u = (\text{grad } f)(p).\frac{(\text{grad } f)(p)}{||(\text{grad } f)(p)||}$$

$$= \frac{||(\text{grad } f)(p)||^2}{||(\text{grad } f)(p)||} = ||(\text{grad } f)(p)||.$$

Otro corolario que se desprende de $(D_u f)(p) = (\text{grad} f)(p).u$ es que si queremos encontrar las direcciones a lo largo de las cuales la función permanece constante (i.e. f no experimenta ninguna razón de cambio) basta ver que $\emptyset = (D_u f)(p) = (\text{grad } f)(p) . u$ para concluir que si (grad f)(p) $\neq \emptyset$ las direcciones en cuestión serán las determinadas por los vectores unitarios ortogonales a (grad f)(p)."

- Como ilustración de la aplicación anterior, se presentan algunos ejemplos tomados de la literatura disponible sobre el tema. Este tipo de ejemplos se proponen porque facilitan al estudiante relacionar el concepto matemático con su correspondiente aplicación a situaciones físicas muy concretas:

"Como ejemplo de la propiedad de que el gradiente tiene la dirección de cambio máximo de f, pensemos en un esquiador que desciende por una ladera. Si $f(x,y)$ denota la altitud del esquiador en el punto (x,y), entonces $-\nabla f(x,y)$ indica la dirección geográfica que el esquiador debe seguir para descender por la trayectoria de más pendiente. ($\nabla f(x,y)$ indica la dirección de ascenso máximo).

Nótese que el vector $\nabla f(x,y) = (\frac{\partial f}{\partial x}(x,y), \frac{\partial f}{\partial y}(x,y))$ indica una dirección en el plano horizontal (norte, sur, este, etc.) y no apunta hacia arriba o hacia abajo.

Como ejemplo adicional de gradiente, considérese la temperatura $T(x,y)$ en un punto cualquiera (x,y) de una placa caliente. En este caso $\nabla T(x,y)$ nos da la dirección de máximo incremento de temperatura en un punto (x,y), y $-\nabla T(x,y)$ señala la dirección hacia la que fluiría el calor (hacia el punto más frio).

Ejemplo. Si la distribución de temperatura en una placa de metal viene dada por la función $T(x,y)=20-4x^2-9y^2$.

(i)¿En qué dirección con origen en $(2,-3)$ aumenta la temperatura más rápidamente? ¿Con qué proporción aumenta en dicha dirección?

$\nabla T(x,y) = (-8x,-18y)$. En $(2,-3)$ la temperatura aumenta más rápidamente en la dirección dada por $\nabla T(2,-3) = (-16,54)$ y lo hace a un ritmo dado por $||\nabla T(2,-3)||=(16^2+54^2)^{1/2}$

(ii)¿En qué dirección con origen en $(2,-3)$ decrece la temperatura más rápidamente?

La temperatura disminuye en mayor medida en la dirección $-\nabla T(2,-3) = (16,-54)$." (cfr.Larson R., Hostetler R.; Cálculo y Geometría Analítica. México, McGraw-Hill. 1986, págs. 720-721).

- Como otra aplicación se puede mencionar que el gradiente proporciona el recurso, análogo a la derivada para funciones de una variable, para encontrar extremos relativos de una función:

"Recordemos que si una función f de $\mathbb{R}$ en $\mathbb{R}$ es derivable; en los puntos x en los que f alcanza algún extremo relativo (máximo o mínimo) se tiene que $f'(x) = 0$.

Para el caso de funciones $f: D \subseteq \mathbb{R}^n \to \mathbb{R}$ si f es C^1 en D y f tiene un máximo o mínimo local en un punto $p_0 \in D$, entonces $(\text{grad } f)(p_0) = 0$, i.e., todas las derivadas parciales se hacen cero en p_0.

Dem. Supongamos que p_0 es un máximo local para f. Entonces existe una ε-vecindad de $p_0: B_\varepsilon(p_0)$ tal que $f(p_0+h) \leq f(p_0)$ si $p_0+h \in B_\varepsilon(p_0)$ (i.e. $\|h\| < \varepsilon$).

$$\therefore \quad f(p_0+h) - f(p_0) \leq 0$$

Con lo que para cualquier dirección u si t es suficientemente pequeña, ($|t| < \varepsilon$), $p_0+tu \in B_\varepsilon(p_0)$

$$\therefore \lim_{t \to 0} \frac{f(p_0+tu)-f(p_0)}{t} \leq 0 \Rightarrow (D_u f)(p_0) \leq 0$$

(Pues por ser f diferenciable en p_0 el límite $(D_u f)(p_0)$ existe independientemente de cómo nos aproximemos a 0, con lo que, si lo hacemos con $t > 0$, $\dfrac{f(p_0+tu)-f(p_0)}{t} \leq 0$)

$$\therefore \quad (D_u f)(p_0) = (\text{grad } f)(p_0).u \leq 0$$

si reemplazamos u por -u tenemos igualmente

$$(\text{grad } f)(p_0) \, . \, (-u) \leq 0$$
$$\Rightarrow (\text{grad } f)(p_0) \, . \, (u) \geq 0$$
$$\Rightarrow (\text{grad } f)(p_0) \, . \, u = 0$$

para cualquier dirección u de donde se sigue que $(\text{grad } f)(p_0) = 0$.

Cualquier punto $p \in D$ en donde $(\text{grad } f)(p) = 0$ se llama punto crítico de f. Si p es un extremo local de f, entonces es un punto crítico de f. (Al igual que en el caso real el recíproco de la afirmación anterior no es cierto, no todo punto crítico es un extremo local de la función)."

"Recordemos que si una función f de $\mathbb{R}$ en $\mathbb{R}$ es derivable; en los puntos x en los que f alcanza algún extremo relativo (máximo o mínimo) se tiene que $f'(x) = \emptyset$.

Para el caso de funciones $f: D \subseteq \mathbb{R}^n \to \mathbb{R}$ si f es C^1 en D y f tiene un máximo o mínimo local en un punto $p_\emptyset \in D$, entonces $(\text{grad } f)(p_\emptyset) = \emptyset$, i.e., todas las derivadas parciales se hacen cero en $p_\emptyset$.

Dem. Supongamos que $p_\emptyset$ es un máximo local para f. Entonces existe una ε-vecindad de $p_\emptyset : B_\varepsilon(p_\emptyset)$ tal que $f(p_\emptyset+h) \leq f(p_\emptyset)$ si $p_\emptyset+h \in B_\varepsilon(p_\emptyset)$ (i.e. $||h|| < \varepsilon$).

$$\therefore \quad f(p_\emptyset+h) - f(p_\emptyset) \leq \emptyset$$

Con lo que para cualquier dirección u si t es suficientemente pequeña, ($|t| < \varepsilon$), $p_\emptyset+tu \in B_\varepsilon(p_\emptyset)$

$$\therefore \lim_{t\to\emptyset} \frac{f(p_\emptyset+tu)-f(p_\emptyset)}{t} \leq \emptyset \Rightarrow (D_uf)(p_\emptyset) \leq \emptyset$$

(Pues por ser f diferenciable en $p_\emptyset$ el límite $(D_uf)(p_\emptyset)$ existe independientemente de cómo nos aproximemos a $\emptyset$, con lo que, si lo hacemos con $t > \emptyset$, $\dfrac{f(p_\emptyset+tu)-f(p_\emptyset)}{t} \leq \emptyset$)

$$\therefore \quad (D_uf)(p_\emptyset) = (\text{grad } f)(p_\emptyset).u \leq \emptyset$$

si reemplazamos u por -u tenemos igualmente

$$(\text{grad } f)(p_\emptyset) \,.\, (-u) \leq \emptyset$$
$$\Rightarrow (\text{grad } f)(p_\emptyset) \,.\, (u) \geq \emptyset$$
$$\Rightarrow (\text{grad } f)(p_\emptyset) \,.\, u = \emptyset$$

para cualquier dirección u de donde se sigue que $(\text{grad } f)(p_\emptyset) = \emptyset$.

Cualquier punto $p \in D$ en donde $(\text{grad } f)(p) = \emptyset$ se llama punto crítico de f. Si p es un extremo local de f, entonces es un punto crítico de f. (Al igual que en el caso real el recíproco de la afirmación anterior no es cierto, no todo punto crítico es un extremo local de la función)."

5.3.2 DIFERENCIABILIDAD Y PLANOS TANGENTES.

A fin de que al estudiante le quedara claro que la diferenciabilidad de una función real de varias variables en un punto es la generalización más adecuada del concepto de derivabilidad para una función de una variable, en un punto, faltaría hablar del aspecto geométrico de la cuestión.

Hasta ahora ya se ha analizado que, desde la perspectiva analítica la diferenciabilidad cumple con las condiciones deseadas: implica la continuidad, supone la existencia de una transformación lineal que aproxima localmente bien, en el mismo sentido que "mutatis mutandis" se tenía para las funciones de una variable (según se revisó en el capítulo 1.)

Para completar el cuadro faltaría hablar de la perspectiva geomértrica: La diferenciabilidad de una función $\varphi:\mathbb{R}\to\mathbb{R}$ era equivalente a la existencia de la recta tangente a la gráfica de φ en el punto en cuestión. Habría entonces que mostrar que la diferenciabilidad para una función $f:D\subseteq\mathbb{R}^2\to\mathbb{R}$ es equivalente a la existencia del plano tangente en el punto, que sería la propiedad análoga que se esperaría de la diferenciabilidad para una función de varias variables.

"Supongamos que $f: D\subseteq\mathbb{R}^2 \to \mathbb{R}$ es diferenciable en $p\in D$ con diferencial $L_{f,p}$; entonces:

(1)...... $f(p+h)-f(p)-L_{f,p}(h)=R_p(h)$ donde $\lim\limits_{h\to\emptyset} \dfrac{R_p(h)}{||h||} = \emptyset$

Haciendo $x = p+h$, podemos reescribir (1) como:

(2)...... $f(x)=f(p)+L_{f,p}(x-p)+R_p(x-p)$ donde $\lim\limits_{x\to p} \dfrac{R_p(x-p)}{||x-p||} = \emptyset$

con $x \in D$.

La gráfica de f es el conjunto G de $\mathbb{R}^3$ con ecuación (3)....... $z = f(x) = f(x_1,x_2)$.

Consideremos el conjunto T de $\mathbb{R}^3$ con ecuación:

(3)'........ $z = f(p) + L_{f,p}(x-p)$.

Reescribiendo (3)' tenemos que si $p = (p_1, p_2)$ $x = (x_1, x_2)$:

$$L_{f,p}(x-p) = \begin{bmatrix} \dfrac{\partial f}{\partial x_1}(p) & \dfrac{\partial f}{\partial x_2}(p) \end{bmatrix} \begin{bmatrix} x_1 - p_1 \\ x_2 - p_2 \end{bmatrix}$$

$$= \frac{\partial f}{\partial x_1}(p)(x_1-p_1) + \frac{\partial f}{\partial x_2}(p)(x_2-p_2)$$

$$\Rightarrow z = f(p_1, p_2) + \frac{\partial f}{\partial x_1}(p)(x_1-p_1) + \frac{\partial f}{\partial x_2}(p)(x_2-p_2) \ldots\ldots (4)$$

$\therefore$ T es un plano en $\mathbb{R}^3$ que pasa por $(p, f(p))$ y es paralelo a la gráfica S de $L_{f,p}$ ya que la gráfica de $L_{f,p}$ está definida por la ecuación:

$$z = L_{f,p}(x_1, x_2) = \begin{bmatrix} \dfrac{\partial f}{\partial x_1}(p) & \dfrac{\partial f}{\partial x_2}(p) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \frac{\partial f}{\partial x_1}(p)x_1 + \frac{\partial f}{\partial x_2}(p)x_2 \ \ldots (4)'$$

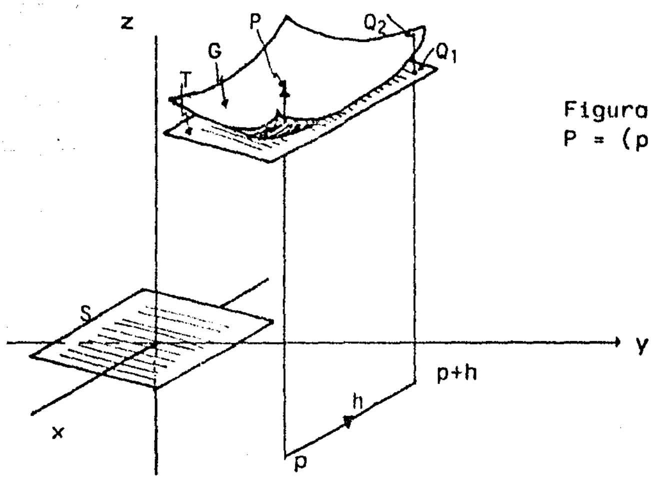


Figura 5.12
$P = (p, f(p))$

Si nos movemos del punto p al punto p+h en $\mathbb{R}^2$, entonces el correspondiente punto Q_1 en el plano T tiene tercera coordenada $z = f(p) + L_{f,p}(h)$.

Por otra parte, el correspondiente punto Q_2 en la gráfica G tiene tercera coordenada $z = f(p + h) = f(p) + L_{f,p}(h) + R_p(h)$.

La magnitud del vector $\overrightarrow{Q_1Q_2}$ indica el error que hacemos al suponer que el cambio en el valor de f es lineal conforme nos movemos de p. La diferenciabilidad de f en p nos dice que el error es "pequeño" en el sentido de que

$$\overrightarrow{Q_1Q_2} = (\emptyset, \emptyset, R_p(h))$$

$$\begin{aligned}(\text{pues } \overrightarrow{Q_1Q_2} &= (p+h, f(p+h)) - (p+h, f(p)+L_{f,p}(h)) \\ &= (\emptyset, f(p+h) - f(p) - L_{f,p}(h)) \\ &= (\emptyset, \emptyset, R_p(h))\end{aligned}$$

$$\text{y} \therefore \lim_{h\to\emptyset} \frac{\|\overrightarrow{Q_1Q_2}\|}{\|h\|} = \lim_{h\to\emptyset} \frac{|R_p(h)|}{\|h\|} = \emptyset \quad \ldots\ldots\ldots\ldots (5)$$

Así como en el caso de una variable, si bien se tenía una idea intuitiva de la recta tangente ésta se definía de una manera que resultaba operativa en términos de la derivada. Aquí también, la definición operativa del plano tangente es la que se da precisamente en términos de la diferencial:

Definición 1. Sea $G \subseteq \mathbb{R}^3$ la gráfica de una función diferenciable $f: D \subseteq \mathbb{R}^2 \to \mathbb{R}$. Para cualquier punto $p \in D$ el plano en $\mathbb{R}^3$ con ecuación $z = f(p)+L_{f,p}(x-p)$ $(x=(x_1,x_2) \in \mathbb{R}^2)$ se llama el plano tangente a G en $(p, f(p))$.

El plano tangente a G en $(p, f(p))$ es el plano que aproxima mejor a G cerca de $(p, f(p))$ en el sentido que se satisfaga (5).

Inversamente, si para una función arbitraria f, existe un plano tal que los puntos correspondientes Q_1 y Q_2 como en la Figura 5.12 satisfacen la condición (5), entonces f es diferenciable en p. La aproximación lineal $L_{f,p}$ puede obtenerse de la ecuación del plano remitiéndose a (3)'.

Podemos reescribir la ecuación del plano tangente a la gráfica de f en $(a,b,f(a,b))$ como

$$z = f(a,b)+L_{f,p}((x_1,x_2)-(a,b)) = f(a,b)+L_{f,p}(x_1-a, x_2-b)$$

$$= f(a,b) + \left[\frac{\partial f}{\partial x_1}(a,b) \quad \frac{\partial f}{\partial x_2}(a,b)\right] \begin{bmatrix} x_1-a \\ x_2-b \end{bmatrix}$$

$$\Rightarrow \quad z = f(a,b) + (x_1-a)\frac{\partial f}{\partial x_1}(a,b) + (x_2-b)\frac{\partial f}{\partial x_2}(a,b) \quad \ldots\ldots(6)$$

Ejemplo 1. La función $f(x,y) = x^2+y^2$ tiene gráfica con ecuación $z = x^2+y^2$. Tenemos $\frac{\partial f}{\partial x}(x,y)=2x$ $\frac{\partial f}{\partial y}(x,y)=2y$.

La ecuación del plano tangente a la gráfica de f en $(a, b, a^2 + b^2)$ está dada en virtud de (6) por:
$z = (a^2 + b^2) + (x - a)2a + (y - b)2b$
i.e. $z = 2ax + 2by - (a^2 + b^2)$.

En particular el plano tangente en $(\emptyset,\emptyset,\emptyset)$ es $z = \emptyset$ (el plano xy) como esperaríamos al ver el dibujo de la gráfica de f. (Ver Figura 5.13)."

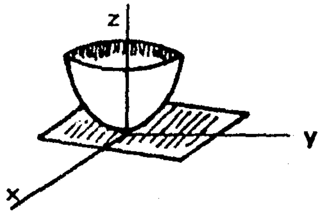


Figura 5.13

Si se deseara dar la generalización para funciones definidas sobre dominios de dimensión más alta y hablar por tanto de espacios tangentes sería conveniente tomar en consideración lo siguiente:

Si definimos el espacio tangente a la gráfica G en $(p,f(p))$ de una función diferenciable $f: D \subseteq \mathbb{R}^{n-1} \to \mathbb{R}$ $(G \subseteq \mathbb{R}^n)$ en $p \in D$, como el conjunto $T \subseteq \mathbb{R}^n$ con ecuación:
$x_n = f(p) + L_{f,p}(x - p)$; $x = (x_1,x_2,\ldots,x_{n-1}) \in \mathbb{R}^{n-1}$.

La definición les resultaría a los estudiantes más natural si manejasen el hecho de que la gráfica de la función lineal: $L_{f,p}: \mathbb{R}^{n-1} \to \mathbb{R}$ es un subespacio vectorial de $\mathbb{R}^n$ de dimensión n-1 y por tanto el espacio tangente T es la traslación de dicho subespacio de modo que pase por $(p,f(p))$. Sin embargo como es muy probable que desconozcan esos resultados, se les podría ayudar a que efectivamente vean a un espacio tangente como el análogo en dimensión superior de un plano tangente fijándose en la forma que tiene la ecuación del espacio tangente:

"El espacio tangente tiene ecuación:
si $p = (p_1,p_2,\ldots,p_{n-1})$; $x = (x_1,x_2,\ldots,x_{n-1})$

$$x_n = f(p)+L_{f,p}(x-p)=f(p)+ \begin{bmatrix} \dfrac{\partial f}{\partial x_1}(p) & \dfrac{\partial f}{\partial x_2}(p) \ldots \dfrac{\partial f}{\partial x_{n-1}}(p) \end{bmatrix} \begin{bmatrix} x_1-p_1 \\ x_2-p_2 \\ \ldots\ldots \\ x_{n-1}-p_{n-1} \end{bmatrix}$$

$$x_n = f(p)+(x_1-p_1)\frac{\partial f}{\partial x_1}(p)+(x_2-p_2)\frac{\partial f}{\partial x_2}(p)+\ldots+(x_{n-1}-p_{n-1})\frac{\partial f}{\partial x_{n-1}}(p)$$

Una recta en $\mathbb{R}^2$ no paralela a la dirección $(\emptyset,1)$ tiene ecuación $x_2 = ax_1 + b$. Un plano en $\mathbb{R}^3$ no paralelo a la dirección $(\emptyset,\emptyset,1)$ tiene ecuación $x_3 = a_1x_1 + a_2x_2 + b$.

La forma de la ecuación del espacio tangente es: $x_n = a_1x_1 + a_2x_2 + \ldots\ldots + a_{n-1}x_{n-1} + b$.

Con lo que podemos pensar en el espacio tangente como una cierta "copia" de $\mathbb{R}^{n-1}$ metida en $\mathbb{R}^n$."

CAPITULO 6

LA DIFERENCIAL DE FUNCIONES VECTORIALES

6.0 INTRODUCCION

Al hablar de funciones vectoriales, nos encontramos de alguna manera en "Terra Ignota". Es frecuente que los libros de texto de Cálculo a nivel básico, aun cuando hablen del Cálculo de varias variables no mencionen a las funciones vectoriales de varias variables. Lo que constituye una laguna, pues por una parte, su estudio tiene un interés propio ya que dichas funciones aparecen con frecuencia en las Matemáticas, y por otra, los mismos textos cuando - en su momento - se refieren a la fórmula del cambio de variables para integrales múltiples, implícitamente están hablando de una de estas funciones; o bien, algunos libros hablan de la divergencia y el rotacional de un campo vectorial, poniendo especial énfasis en la parte algorítmica del tema, sin reparar suficientemente en el hecho de que un campo es precisamente una función vectorial; y al quedar un tanto velada la naturaleza de función que tiene un campo no se investigan cuestiones como su dominio, su imagen, su diferenciabilidad, etc., que sirven para entender mejor el concepto. En síntesis, otra razón para introducir el estudio de las funciones vectoriales es que tarde o temprano, aparecerán en el desarrollo de los cursos de Cálculo de varias variables y se tendrá un conocimiento más profundo de aquellos temas en las que aparecen si se las ha estudiado con cierto detalle.

Una vez vista la conveniencia de presentar el material, la pregunta obligada es ¿con qué tanto detalle? Teniendo en mente que nuestro objetivo es que los estudiantes entiendan y se familiaricen con el concepto de diferencial de una función vectorial, la respuesta a esta pregunta pretende encontrarse en el presente capítulo.

Los elementos básicos necesarios son:

1) Qué es una función vectorial o transformación - nombre, éste último, que enfatiza el carácter geométrico de estas funciones como transformaciones o deformaciones de conjuntos -.

2) Cómo se opera con ellas. Esto supone explicar la

descomposición de una función vectorial en sus funciones coordenadas. Lo cual, además de proporcionar una manera práctica de trabajar con las transformaciones; desde el punto de vista didáctico es recomendable, pues se reduce un objeto nuevo a uno que ya es familiar para los estudiantes (las funciones real valuadas).

3) Con qué recursos se dispone para visualizar las funciones vectoriales.

4) Límites y continuidad para estas funciones, que pueden reducirse a los conceptos correspondientes para las funciones coordenadas.

Los temas anteriores se estudian en las primeras dos secciones. Con las definiciones de límite y continuidad para las funciones vectoriales se puede hablar de la diferenciabilidad de estas funciones como la existencia de (a lo más) una función lineal que aproxima a la función en una vecindad del punto.

El uso de funciones coordenadas permite reducir el estudio de la diferenciabilidad de las funciones vectoriales al de las funciones real valuadas en donde ya se cuenta con una serie de resultados importantes. Esto se trata con detalle en la sección 6.3.

6.1 FUNCIONES VECTORIALES DE VARIABLE VECTORIAL

En esta sección pretendemos abordar el estudio de las funciones vectoriales de variable vectorial, es decir aquellas que son de la forma $f\colon D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ con $n, m \in \mathbb{N}$. Donde $D \subseteq \mathbb{R}^n$ es un subconjunto abierto de $\mathbb{R}^n$. Ya se han trabajado los casos particulares: $n = 1$ y $m = 1$. Lo que pretendemos en este estudio es generalizar muchas de las ideas que se manejaron entonces.

El primer objetivo sería hacer ver que dada una función vectorial, ésta puede descomponerse en sus funciones coordenadas, lo cual resulta de una gran utilidad, pues éstas son funciones real valuadas y por tanto del tipo de las que los estudiantes ya han trabajado. Además de que operativamente, el uso de las funciones coordenadas simplifica el trabajar con las funciones vectoriales.

Por ello se sugiere la siguiente presentación:

"Si $A \subseteq \mathbb{R}^n$ y f es una función definida sobre A con rango B contenido en $\mathbb{R}^m$, f es una regla que permite asociar a cada x en $A(x \in \mathbb{R}^n)$ un elemento y de B, la imagen de x bajo f, hecho que denotamos como $y = f(x)$. Como y es un elemento de $\mathbb{R}^m \Rightarrow y = (y_1, y_2, \ldots, y_m)$. Cada uno de los números reales y_i depende de x, lo cual define m funciones real valuadas cuyo dominio es $D \subseteq \mathbb{R}^n$, es decir para cada $i = 1, 2, \ldots m$

$$y_i = f_i(x) \qquad f_i: D \subseteq \mathbb{R}^n \to \mathbb{R}$$

En conclusión, una función $f: D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ la podemos descomponer en m funciones real valuadas:
$f(x) = y = (y_1, y_2, \ldots, y_m) = (f_1(x), f_2(x), \ldots, f_m(x))$
con $f_i: D \subseteq \mathbb{R}^n \to \mathbb{R}$ $\qquad i = 1, 2, \ldots, m$.

A las funciones $f_i: D \subseteq \mathbb{R}^n \to \mathbb{R}$ se les denomina las funciones coordenadas de f.

Esta descomposición de una función cuyo rango está en $\mathbb{R}^m$ $(m \geq 2)$ en funciones coordenadas es muy útil, pues podemos conocer propiedades de la función a través del estudio de sus funciones coordenadas - que al ser funciones real valuadas - nos resultan más familiares.

Podemos mediante combinaciones de estas funciones originar otras de la manera usual:

Sean $f: D \subseteq \mathbb{R}^n \to \mathbb{R}^m$, $g: D \subseteq \mathbb{R}^n \to \mathbb{R}^m$, $\varphi: D \subseteq \mathbb{R}^n \to \mathbb{R}$ definimos

(i) La función suma $f + g: D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ como $(f + g)(x) = f(x) + g(x)$, $x \in D$.

(i)' La función diferencia $f - g: D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ como $(f - g)(x) = f(x) - g(x)$, $x \in D$.

(ii) El producto $\varphi f: D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ como
$(\varphi f)(x) = \varphi(x) f(x) = (\varphi(x) f_1(x), \ldots\ldots, \varphi(x) f_m(x))$, $x \in D$.

(iii) El producto punto $f \cdot g: D \subseteq \mathbb{R}^n \to \mathbb{R}$ como
$(f \cdot g)(x) = f_1(x) g_1(x) + \ldots + f_m(x) g_m(x)$, $x \in D$.
donde $f_i: D \subseteq \mathbb{R}^n \to \mathbb{R}$ $\quad g_i: D \subseteq \mathbb{R}^n \to \mathbb{R}$ $\quad i = 1, 2, \ldots m$
son las funciones coordenadas de f y g."

A este nivel, en el que los estudiantes entran en contacto por primera vez con las funciones vectoriales es especialmente interesante enfatizar el carácter geométrico de estas funciones como transformaciones de conjuntos en sus respectivas imágenes. Para ilustrar la naturaleza geométrica de estas funciones es de gran ayuda, presentar ejemplos como los que se proponen, trabajados con cierto detalle a fin de que los estudiantes adquieran el hábito de pensar en estas funciones justamente como transformaciones

"Ejemplo 1. Sea $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ dada por $f(x,y) = (x^2+y^2, x+y)$. Tenemos entonces que: $f(x,y) = (u,v)$

$$u = f_1(x,y) = x^2+y^2 \qquad f_1\colon \mathbb{R}^2 \to \mathbb{R}$$
$$v = f_2(x,y) = x+y \qquad f_2\colon \mathbb{R}^2 \to \mathbb{R}$$

Bajo f el punto (x,y) es mandado al (u,v). De este modo la imagen de $(1,2)$ es $(5,3)$; la imagen de $(\emptyset,2)$ es $(4,2)$ y la de $(2,\emptyset)$ también es $(4,2)$ (Ver Figura 6.1)

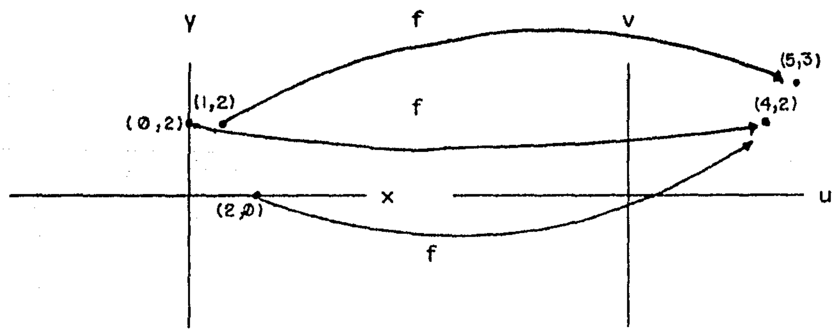


Figura 6.1

Cualquier conjunto de puntos en el plano XY es "mandado" o transformado por la función en un conjunto de puntos en el plano UV.

Podemos describir la imagen bajo f de algunas curvas y regiones. Para determinar la imagen de una línea horizontal $y = c$, hacemos la sustitución correspondiente en las ecuaciones que definen a f obteniendo

$$u = x^2 + c^2, \quad v = x + c$$

Que pueden considerarse como ecuaciones paramétricas para la curva imagen en el plano UV. Para visualizar la imagen podemos eliminar a x (que es el parámetro) obteniendo $u = (v - c)^2 + c^2$. Que corresponden (para diferentes elecciones de c) a parábolas que abren en sentido positivo y con eje de simetría paralelo al eje u. Algunas de estas curvas se muestran en la Figura 6.2

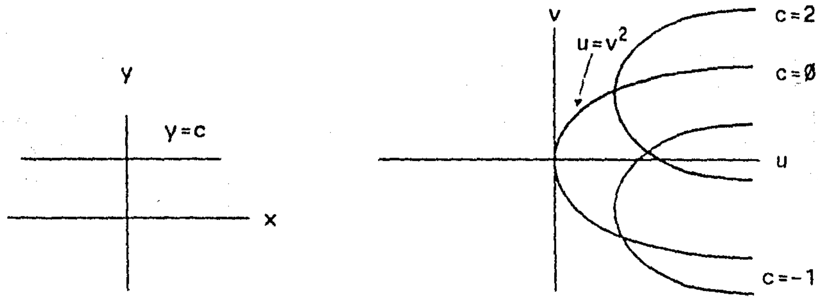


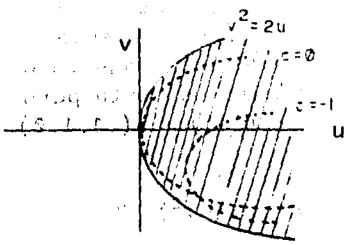
Figura 6.2

Con lo que respecta a las regiones, observamos primero que el punto (a,b) y el punto (b,a) tienen la misma imagen. Por lo tanto, la línea $y = x$ divide al plano XY en dos semiplanos que bajo f tienen por imagen el mismo conjunto del plano UV. Para determinar este conjunto, primero determinamos la imagen de la recta $x = y$ substituyendo en las ecuaciones que definen a f obtenemos:

$$\left.\begin{aligned} u &= 2x^2 \\ v &= 2x \end{aligned}\right]$$

que son las ecuaciones paramétricas para la parábola $v^2 = 2u$.

La imagen de cualquier punto (x,y) está en una de las dos regiones del plano UV en las que la parábola $v^2 = 2u$ divide al plano (Ver Figura 6.3)



pues se tiene:

$$\begin{aligned} 2u-v^2 &= 2(x^2+y^2)-(x+y)^2 \\ &= x^2 + y^2 - 2xy \\ &= (x - y)^2 \geq \emptyset \end{aligned}$$

$$\therefore \quad 2u \geq v^2$$

Figura 6.3

Inversamente se puede verificar que todo punto $(u_\emptyset, v_\emptyset)$ tal que $v_\emptyset^2 \leq 2u_\emptyset$ es la imagen bajo f de algún punto $(x,y) \in \mathbb{R}^2$. De hecho si $(u_\emptyset, v_\emptyset)$ está en el plano UV y satisface $v_\emptyset^2 \leq 2u_\emptyset$. Sean

$x_1 = v_\emptyset/2 + (2u_\emptyset - v_\emptyset^2)^{1/2}/2 \qquad y_1 = v_\emptyset/2 - (2u_\emptyset - v_\emptyset^2)^{1/2}/2$
$x_2 = v_\emptyset/2 - (2u_\emptyset - v_\emptyset^2)^{1/2}/2 \qquad y_2 = v_\emptyset/2 + (2u_\emptyset - v_\emptyset^2)^{1/2}/2$

se verifica que $f(x_1, y_1) = (u_\emptyset, v_\emptyset) = f(x_2, y_2)$.

De esta forma uno puede imaginarse el efecto de f aproximadamente como sigue: primero doblar el plano XY a lo largo de la línea x = y; colocar el "pliegue" a lo largo de la parábola $v^2 = 2u$ y luego aplanar suavemente el resto para cubrir el interior de la parábola (para permitir la necesaria deformación podemos pensar en el plano XY como si fuese una hoja hecha de goma). (Ver Figura 6.4)

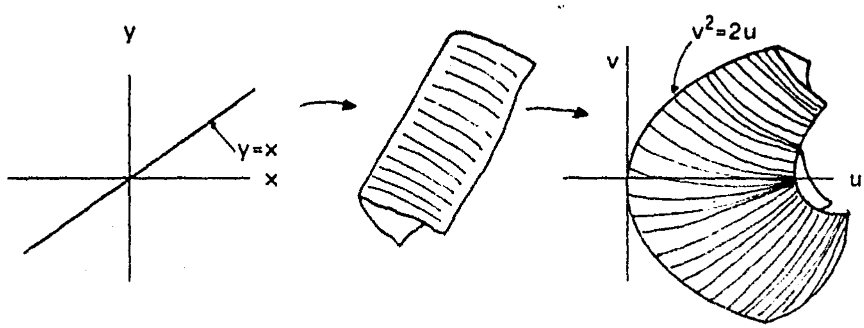


Figura 6.4

Como otra ilustración, considermos la siguiente transformación del plano en $\mathbb{R}^3$, dada por $S(x,y)=(u,v,w)$ $S:\mathbb{R}^2 \to \mathbb{R}^3$ $u(x,y) = x+y$, $v(x,y) = x-y$, $w(x,y) = x^2$.

$S(1,2) = (3,-1,1)$. La imagen de la línea $y = x$ es la curva con ecuaciones paramétricas dadas por: $u = 2x$, $v = \emptyset$, $w = x^2$. Que es una parábola en el plano UW. La imagen del plano XY bajo f es un cilindro parabólico tangente al plano UV. (Ver Figura 6.5). Pues un cilindro así está parametrizado por $(u,v,w) \in \mathbb{R}^3$ } $(u,v,w) = (2t,\emptyset,t^2) + \lambda(-1,1,\emptyset) = (2t-\lambda, \lambda, t^2)$ donde $t, \lambda \in \mathbb{R}$.

Si $t = x$, $\lambda = x-y$ $\qquad (u,v,w) = (x+y, x-y, x^2)$."

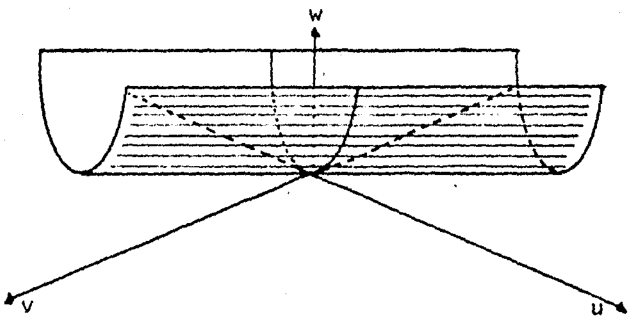


Figura 6.5

El hecho de que la geometría siempre resulta un gran apoyo para el aprendizaje de las Matemáticas, hace que convenga mencionar algunos otros recursos que pueden ser de utilidad para visualizar, i.e. para tener una imagen geométrica de las funciones vectoriales:

- La Gráfica:

"Como las transformaciones son funciones, también podríamos considerar sus gráficas. La gráfica de $f: D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ es el conjunto definido por $\{(p, f(p))/p \in D\} \subseteq \mathbb{R}^n \times \mathbb{R}^m$. Sólo podemos dibujar la gráfica de f en casos muy simples (necesitamos $m+n \leq 3$). Por ejemplo, la gráfica de la función f descrita en el primer ejemplo es el conjunto de puntos $(x, y, u, v) \in \mathbb{R}^4$ para los cuales $u = x^2+y^2$; $v = x+y$, i.e. el conjunto de puntos $\{(x, y, u, v) \in \mathbb{R}^4 /\ u=x^2+y^2\ v=x+y\ (x, y) \in \mathbb{R}^2\}$"

- Conjuntos de nivel:

El hecho de que la gráfica sólo puede dibujarse para casos muy sencillos, motiva el adoptar otros recursos para visualizar las transformaciones; como los conjuntos de nivel:

"El conjunto de nivel de $f: D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ correspondiente a $c \in \mathbb{R}^m$ es el conjunto $\{x \in D/f(x) = c\}$. Podemos dibujar conjuntos de nivel de f para m arbitraria siempre y cuando $n \leq 3$. Veamos dos ejemplos de este recurso.

Ejemplo (i) Definamos $f: \mathbb{R}^3 \to \mathbb{R}^2$ por $f(x) = (x_1^2+x_2^2+x_3^2, x_1+x_2+x_3) \qquad x = (x_1, x_2, x_3)$.

Las funciones coordenadas de f son f_i: $i = 1,2$ definidas como $f_1(x) = x_1^2+x_2^2+x_3^2$; $f_2(x) = x_1+x_2+x_3$. El conjunto de nivel correspondiente a $(-1,1)$ es vacío. El conjunto de nivel correspondiente a $(1,1)$ es $\{(x_1,x_2,x_3) \in \mathbb{R}^3 \ / \ x_1^2+x_2^2+x_3^2 = 1 \ \text{ y } \ x_1+x_2+x_3 = 1\}$.

El conjunto de nivel correspondiente a $(1,1)$ de f es la intersección de una esfera y un plano en $\mathbb{R}^3$ -un círculo-.

Ejemplo (ii) Asociada a una función diferenciable $f: D \subseteq \mathbb{R}^n \to \mathbb{R}$ existe una función vectorial grad f: $D \subseteq \mathbb{R}^n \to \mathbb{R}^n$. Las funciones coordenadas de (grad f) son las funciones definidas por las derivadas parciales de f:

$\dfrac{\partial f}{\partial x_j} : D \subseteq \mathbb{R}^n \to \mathbb{R} \qquad j = 1, \dots, n$. El conjunto de nivel de (grad f) correspondiente a $\emptyset$ es el conjunto de puntos críticos de f."

- Un recurso geométrico del que siempre se puede echar mano es pensar en las funciones vectoriales como transformaciones:

"Sin embargo lo más práctico al trabajar con funciones vectoriales $f: D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ es pensar en que f transforma a conjuntos de un espacio $\mathbb{R}^n$ en conjuntos de $\mathbb{R}^m$; "deformando" a cada conjunto $A \subseteq D$ en su imagen $f(A) \subseteq \mathbb{R}^m$ que podemos imaginarla como una especie de "superficie" en $\mathbb{R}^m$.

Por esta razón a estas funciones también se les denomina transformaciones, lo que enfatiza este aspecto geométrico. (Figura 6.6)."

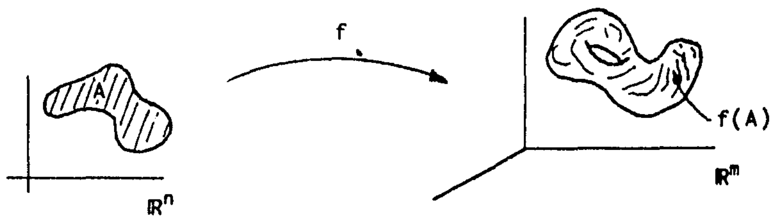


Figura 6.6

6.2 LIMITES Y CONTINUIDAD.

Después de haber considerado la geometría de las funciones vectoriales, se puede tocar el tema del análisis para dichas funciones. Si nuestro objetivo es llegar a estudiar la diferenciabilidad, es preciso reparar primero en las definiciones de límite y continuidad. Para ello es necesario ver cómo se traducen en este contexto ambas definiciones. Si bien formalmente sólo hay que sustituir el valor absoluto en las definiciones correspondientes para el caso de funciones reales de variable real, por la norma euclidiana; sí resulta conveniente esclarecer el sentido de las definiciones a base de explicar la geometría asociada a ellas.

"Para el caso de una variable tenemos que la expresión $\lim_{x \to p} f(x) = L$ quiere decir que dada $\varepsilon > 0$ es posible encontrar $\delta > 0$ tal que $|f(x)-L| < \varepsilon$ siempre que $0 < |x-p| < \delta$.

La correspondiente expresión $\lim_{x \to p} f(x) = L$ para el caso $f: D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ significa que dado $\varepsilon > 0$ es posible encontrar $\delta > 0$ tal que $||f(x)-L|| < \varepsilon$ siempre que $0 < ||x-p|| < \delta$.

Lo que supone que dada una ε-vecindad en $\mathbb{R}^m$ con centro en el punto L, podemos encontrar una δ-vecindad en $\mathbb{R}^n$ con centro en el punto p de forma que la imagen de cualquier punto de $\mathbb{R}^n$ que esté en el interior de dicha δ-vecindad (a excepción, quizá, del punto p) estará dentro de la ε-vecindad de $\mathbb{R}^m$ con centro en L. (Figura 6.7)

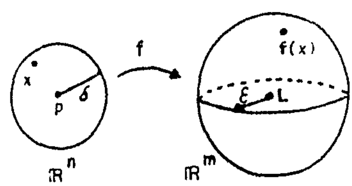


Figura 6.7

En términos de vectores la definición nos dice que podemos hacer la norma (en $\mathbb{R}^m$) del vector $||f(x)-L||$ tan pequeña como queramos siempre y cuando estemos suficientemente cerca del punto p; i.e. la norma del vector x-p en $\mathbb{R}^n$ sea suficientemente pequeña (aunque diferente de cero). (Ver Figura 6.8)"

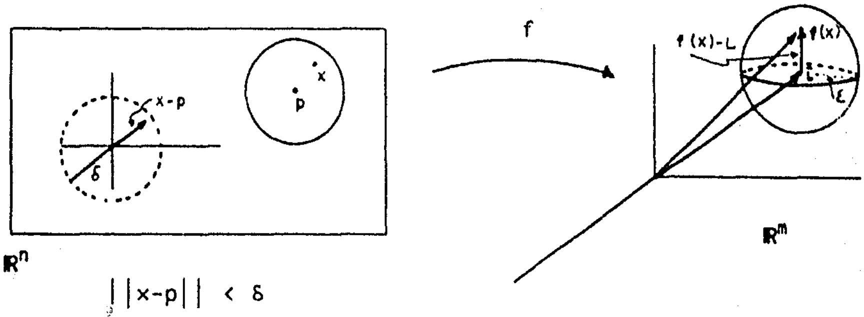


Figura 6.8

Algo que facilita el operar con las transformaciones es presentar a los estudiantes la demostración de que la existencia del límite en un punto para una función vectorial es equivalente a la existencia del límite en el punto de las funciones coordenadas ya que por un lado, esta demostración es en sí misma, un buen ejercicio en el manejo de argumentos del tipo: ε - δ, lo que sirve a los estudiantes para adquirir madurez matemática. Por otro lado, este teorema justifica el remitirse a las funciones coordenadas de una función vectorial para estudiar propiedades como la existencia de límites, continuidad de la función, etc.. Lo cual operativamente es recomendable pues para estudiar límites y continuidad de funciones real valuadas, los estudiantes ya disponen de varios recursos (cfr. capítulos 3 y 5).

"Teorema 1: Si $f: D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ y descomponemos a f en sus funciones coordenados $f_i: D \subseteq \mathbb{R}^n \to \mathbb{R}$ $i = 1,2,\ldots,m$ $f(x)=(f_1(x),f_2(x),\ldots\ldots,f_m(x))$ entonces si $L=(l_1,l_2,\ldots,l_m)$ se tiene que $\lim_{x\to p} f(x)=L$ si y sólo si $\lim_{x\to p} f_i(x)=l_i$ $i=1,2,\ldots,m$

Dem.

Supongamos que $\lim_{x\to p} f(x)=L$, entonces dada $\varepsilon > 0$, existe $\delta > 0$ tal que si $0<||x-p|| < \delta$ se tiene que $||f(x)-L|| < \varepsilon$

$$\Rightarrow ||(f_1(x)-l_1,\ f_2(x)-l_2,\ldots\ldots,f_m(x)-l_m)|| < \varepsilon$$

$$\Rightarrow (f_1(x)-l_1)^2 + (f_2(x)-l_2)^2 +\ldots\ldots+(f_m(x)-l_m)^2 < \varepsilon^2$$

lo que implica que para cada $i=1,2,\ldots,m$ $|f_i(x)-l_i| < \varepsilon$

$\therefore$ Dada $\varepsilon > 0$, hemos encontrado $\delta > 0$ tal que si $0 < ||x-p|| < \delta$ entonces $|f_i(x)-l_i| < \varepsilon \Rightarrow \lim_{x \to p} f_i(x) = l_i$ para $i = 1,2,\ldots m$

Inversamente, supongamos que $\lim_{x \to p} f_i(x) = l_i$, $i = 1,2,\ldots m$

Entonces dada $\varepsilon > 0$, exite $\delta_i > 0$ tal que si $||x-p|| < \delta_i$ se tiene que $|f_i(x)-l_i| < \varepsilon/(m)^{1/2}$ (para $i = 1,2\ldots m$).

Sea $\delta = \min\{\delta_1, \delta_2 \ldots \delta_m\}$ $\delta > 0$ si $||x-p|| < \delta$ entonces $|f_i(x)-l_i| < \varepsilon/(m)^{1/2}$ para $i = 1,2,\ldots m$ teniendo entonces que:

$(f_1(x)-l_1)^2 + (f_2(x)-l_2)^2 + \ldots + (f_m(x)-l_m)^2 < \varepsilon^2$

$\therefore$ $||f(x)-L||^2 < \varepsilon^2 \Rightarrow ||f(x)-L|| < \varepsilon$ $\therefore$ $\lim_{x \to p} f(x) = L$."

Entendida la idea de límite para estas funciones; la definición de continuidad de una función $f\colon D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ no ofrece ninguna dificultad:

$f\colon D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ es continua en $p \in D$ si $\lim_{x \to p} f(x) = f(p)$.

Lo que resulta conveniente es resaltar la naturaleza geométrica de la definición de continuidad con señalamientos como: "f es continua en p si puntos arbitrariamente cercanos a f(p), son imágenes de puntos suficientemente cercanos a p. (Ver Figura 6.9)."

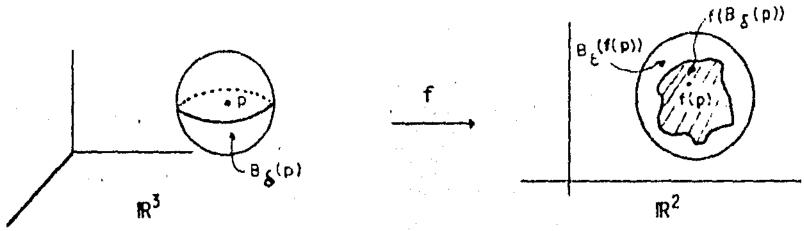


Figura 6.9 Geometría de la definición de continuidad en p, para una función $f\colon \mathbb{R}^3 \to \mathbb{R}^2$.

La definición de continuidad de una función en todo un subconjunto es la usual: "f es continua en D si f es continua en p, para todo $p \in D$". Conviene señalar que: "Intuitivamente, una función $f: D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ que sea continua en D es aquella que deforma a D, doblándolo y torciéndolo, pero sin romper (Ver Figura 6.1Ø)."

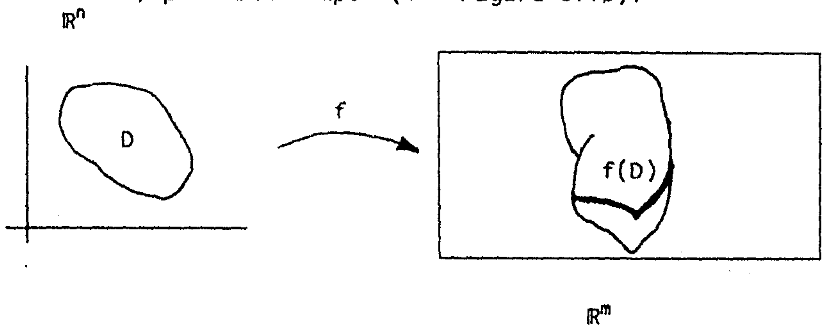


Figura 6.1Ø

Con el Teorema 1, es inmediato que la continuidad de una función $f: D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ en $p \in D$ es equivalente a la continuidad de cada una de las funciones coordenadas en p. (Se puede repetir la demostración del Teorema 1 simplemente sustituyendo $L = (l_1, l_2, \ldots, l_m)$ por $f(p) = (f_1(p), f_2(p), \ldots, f_m(p))$.)

Este hecho conviene señalarlo pues reduce el estudio de la continuidad para funciones vectoriales al estudio de la continuidad para funciones real valuadas, que es un tema ya familiar.

Si el profesor juzgara oportuno hablar del tema de la continuidad uniforme, es éste un buen momento para hacerlo, pues ya se han proporcionado los elementos necesarios. La continuidad uniforme es una propiedad cuya importancia se pone de manifiesto en situaciones de carácter teórico. Es un hecho experimentado por los estudiantes (al menos para el caso de funciones de $\mathbb{R}$ en $\mathbb{R}$) lo mucho que simplifica trabajar con funciones que tienen esa propiedad. En particular, si se introduce el tema, se puede probar que las funciones lineales son uniformemente continuas, lo cual resalta la conveniencia de contar con ellas como aproximaciones locales para funciones que pueden ser muy complicadas. Por otra parte, el resultado de la continuidad uniforme para las funciones lineales se usa en la demostración de la Regla de la Cadena (cfr. capítulo 7).

"La definición de continuidad uniforme también puede "calcarse" de la correspondiente para el caso real simplemente cambiando valores absolutos por las normas euclidianas de los espacios respectivos:

> $f: D \subseteq \mathbb{R} \to \mathbb{R}$ es uniformemente continua en D si dada $\varepsilon > 0$, existe $\delta > 0$ tal que $|x-y| < \delta$ implica $|f(x)-f(y)| < \varepsilon$, para cualquier par de puntos x, y en D.
>
> $f: D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ es uniformemente continua en D si dada $\varepsilon > 0$, existe $\delta > 0$ tal que $||x-y|| < \delta$ implica $||f(x)-f(y)|| < \varepsilon$, para cualquier par de puntos x, y en D.

Probemos el siguiente resultado:

Cualquier función lineal L $\mathbb{R}^n \to \mathbb{R}^m$ es uniformemente continua en $\mathbb{R}^n$. Pues dada $L: \mathbb{R}^n \to \mathbb{R}^m$ lineal, existe una matriz $A = (a_{ij})$ de mxn tal que $L(x) = Ax$ para $x \in \mathbb{R}^n$.

Sea $x = (x_1, \ldots, x_n)$ $y = L(x) = (y_1, \ldots, y_m)$

Tenemos $||x||^2 = \sum_{j=1}^{n} |x_j|^2$ $\qquad ||y||^2 = \sum_{i=1}^{m} |y_i|^2$

Como
$$\begin{bmatrix} y_1 \\ \cdot \\ y_i \\ \cdot \\ y_m \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \cdot & & \cdot \\ a_{i1} & \cdots & a_{in} \\ \cdot & & \cdot \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix} \Rightarrow y_i = \sum_{j=1}^{n} a_{ij} x_j$$

$$\Rightarrow |y_i|^2 \leq \left[\sum_{j=1}^{n} |a_{ij}| \; |x_j| \right]^2 \leq \left[\sum_{j=1}^{n} |a_{ij}|^2 \right] \left[\sum_{j=1}^{n} |x_j|^2 \right] =$$

$$||x||^2 \left[\sum_{j=1}^{n} |a_{ij}|^2 \right] \quad ....(1)$$

(por la desigualdad de Schwartz $|a \cdot b| \leq ||a|| \, ||b|| \; a, b \in \mathbb{R}^n$)

Sumando las expresiones (1) correspondientes a $i = 1,2,...m$ tenemos:

$$||y||^2 = \sum_{i=\emptyset}^{m} |y_i|^2 \leq ||x||^2 \left[\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2 \right]$$

$$\therefore \quad ||L(x)|| = ||y|| \leq M||x|| \quad ....(2)$$

donde $M = \left[\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2 \right]^{1/2}$

Con la expresión (2), la continuidad uniforme de L es fácil de verificar pues dada $\varepsilon > \emptyset$, escogemos $\delta = \varepsilon/M$ (suponiendo que la transformación no es la cero) entonces si $||u-v|| < \delta \quad u,v \in \mathbb{R}^n$

$$||L(u)-L(v)|| = ||L(u-v)|| \leq M||u-v|| < M\delta = M\,\varepsilon/M = \varepsilon \text{ ."}$$

6.3 DIFERENCIABILIDAD

En esta sección pretendemos enfrentar el problema de presentar la definición de diferenciabilidad de una manera didáctica eficiente. Es frecuente que los textos - quizá con el ánimo de facilitar las cosas - den la siguiente definición:

"Sea $f: D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$, decimos que f es diferenciable en $p \in D$, si existe una matriz T de mxn tal que:

$$\lim_{h \to \emptyset} \frac{||f(p+h) - f(p) - Th||}{||h||} = \emptyset \text{ ."}$$

Que puede no resultar una presentación adecuada, puesto que la diferencial de una función f en un punto p es propiamente una transformación lineal, una función. Y si bien es cierto que toda transformación lineal es una transformación matricial, matrices y funciones son objetos matemáticos distintos, que además tienen connotaciones claramente distintas para los alumnos. Las matrices son pensadas por los estudiantes como objetos de un carácter más bien estático: son arreglos de números. Mientras que las funciones son percibidas como un objeto dinámico: implican variabilidad, cambio, movimiento.

Por otra parte, si se ha seguido la línea de pensamiento del capítulo anterior y ya se ha estudiado el material sobre Algebra lineal presentado en el capítulo 4, a estas alturas, se tiene suficiente madurez matemática para poder entender la definición de diferencial de una función en un punto en términos de transformaciones lineales; ya que, a través de los casos de funciones de $\mathbb{R}$ a $\mathbb{R}$, de $\mathbb{R}$ a $\mathbb{R}^n$ y de $\mathbb{R}^n$ a $\mathbb{R}$, los estudiantes han tenido la oportunidad de familiarizarse con lo que supone, analítica y geométricamente la diferenciabilidad entendida como aproximación lineal local.

A continuación se proponen textos que buscan lograr una presentación adecuada al tema:

"Dada una función $f: D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ y un punto $p \in D$, podemos considerar la diferencia $f(p+h) - f(p)$ para todas aquellas h tales que $p+h \in D$. Obsérvese que para cada h, $f(p+h) - f(p)$ es un punto de $\mathbb{R}^m$.

Definición. Una función $f: D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ es diferenciable en $p \in D$, si existe una transformación lineal $L:\mathbb{R}^n \to \mathbb{R}^m$ tal que:

$$\text{si} \quad (i) \; f(p+h) - f(p) - L(h) = R(h)$$

$$\text{entonces} \; (ii) \lim_{h \to \emptyset} \frac{R(h)}{||h||} = \emptyset \text{ ."}$$

Se puede señalar que $R(h)$ es el vector diferencia o el error que se comete al aproximar la función f en una vecindad del punto p, por la función afín $h \longrightarrow f(p)+L(h)$.

"Ambas condiciones se pueden resumir en una pidiendo que

$$\lim_{h \to \emptyset} \frac{f(p+h) - f(p) - L(h)}{||h||} = \emptyset \text{ ."}$$

El siguiente teorema es muy importante pues nos permite utilizar el material acerca de la diferenciabilidad para las funciones real valuadas en el estudio de la diferenciabilidad de funciones vectoriales.

"Teorema 1. La función $f: D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ es diferenciable en $p \in D$ si y sólo si cada una de sus funciones coordenadas $f_i: D \subseteq \mathbb{R}^n \to \mathbb{R}$ $i = 1,2,\ldots m$ es diferenciable en $p \in D$.

Dem. Una función $L: \mathbb{R}^n \to \mathbb{R}^m$ es lineal si y sólo si cada una de sus funciones coordenadas es lineal. (cfr. capítulo 4). Supongamos que f es diferenciable en $p \in D$ $\Rightarrow \exists\, L: \mathbb{R}^n \to \mathbb{R}^m$ lineal tal que:

(i) $f(p+h) - f(p) - L(h) = R(h)$

$$\text{(ii)}\ \lim_{h \to \emptyset} \frac{R(h)}{||h||} = \emptyset .$$

Considerando la i-ésima coordenada de los vectores involucrados en (i) obtenemos las siguientes expresiones en términos de las funciones coordenadas para cada $i=1,2,\ldots m$:

$$f_i(p+h) - f_i(p) - L_i(h) = R_i(h)$$

y como $|R_i(h)| \leq ||R(h)||$, para cada $i \in \{1,2,\ldots .m\}$ tenemos que:

$$\emptyset \leq \lim_{h \to \emptyset} \frac{|R_i(h)|}{||h||} \leq \lim_{h \to \emptyset} \frac{||R(h)||}{||h||} = \emptyset$$

$\therefore$ f_i es diferenciable en p $i = 1,2,\ldots m$.

Inversamente si cada función coordenada de f es diferenciable en $p \in D$, entonces para cada $i = 1,2,\ldots m$ existe una función lineal $L_i: \mathbb{R}^n \to \mathbb{R}$ tal que

(i)' $f_i(p+h) - f_i(p) - L_i(h) = R_i(h)$

$$\text{(ii)'}\ \lim_{h \to \emptyset} \frac{R_i(h)}{||h||} = \emptyset .$$

Definamos una función lineal $L: \mathbb{R}^n \to \mathbb{R}^m$ por:
$L(h) = (L_1(h), L_2(h), \ldots\ldots, L_m(h))$ $h \in \mathbb{R}^n$
y consideremos $R(h) = (R_1(h), R_2(h), \ldots\ldots, R_m(h))$
(para aquellos $h \in \mathbb{R}^n$ tales que $p+h \in D$).

Ahora (i)' y la manera en que hemos definido $L:\mathbb{R}^n \to \mathbb{R}^m$ implican que: $f(p+h) - f(p) - L(h) = R(h)$

y como $\|R(h)\|^2 = |R_1(h)|^2 + |R_2(h)|^2 + \ldots + |R_m(h)|^2$

el hecho de que $\lim\limits_{h\to\emptyset} \dfrac{R_i(h)}{\|h\|} = \emptyset$ implica que $\lim\limits_{h\to\emptyset} \dfrac{\|R(h)\|}{\|h\|} = \emptyset$

por lo tanto f es diferenciable en $p \in D$.

Se sigue inmediatamente de la definición de diferenciabilidad que las funciones lineales son diferenciables en todo punto, pues dada $L: \mathbb{R}^n \to \mathbb{R}^m$, entonces $\forall\ p \in \mathbb{R}^n$

$$L(p+h) - L(p) - L(h) = L(p) + L(h) - L(p) - L(h) = \emptyset$$

i.e. $R(h) = \emptyset$ y por lo tanto $\lim\limits_{h\to\emptyset} \dfrac{R(h)}{\|h\|} = \emptyset$.

Es decir, que para cualquier punto de $\mathbb{R}^n$, la diferencial de L en p es la misma función L, lo cual es claro, pues para una función lineal su mejor aproximación lineal en una vecindad de cualquiera de sus puntos es ella misma."

Desde que se comenzó el estudio de la diferenciabilidad, siempre se tuvo la preocupación de que ésta implicara la continuidad. Con el trabajo del capítulo anterior, esta implicación para el caso de las funciones vectoriales es inmediata como corolario del teorema anterior.

"Si $f: D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ es diferenciable en $p \in D$, entonces f es continua en p.

Dem. Como f es diferenciable en p, sus funciones coordenadas son diferenciables en p, por resultados previos de funciones real valuadas, esto implica que son continuas en p, y por tanto la función f es continua en p."

Otro resultado que hay que presentar es la unicidad de la aproximación lineal a una función:

"Si $f: D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ es diferenciable en $p \in D$, entonces existe una única función lineal $L: \mathbb{R}^n \to \mathbb{R}^m$, tal que

$$\lim_{h \to \emptyset} \frac{||f(p+h) - f(p) - L(h)||}{||h||} = \emptyset.$$

Dicha función lineal $L: \mathbb{R}^n \to \mathbb{R}^m$ está dada por:
$L(h) = (L_{f_1,p}(h), L_{f_2,p}(h), \ldots\ldots\ldots, L_{f_m,p}(h)) \quad h \in \mathbb{R}^n$

Dem. Si $L: \mathbb{R}^n \to \mathbb{R}^m$ es lineal y satisface:

$$\lim_{h \to \emptyset} \frac{f(p+h) - f(p) - L(h)}{||h||} = \emptyset$$

Por la primera parte del Teorema 1, teníamos que para $i = 1,2,\ldots m$

$$f_i(p+h) - f_i(p) - L_i(h) = R_i(h)$$

$$\text{y} \quad \lim_{h \to \emptyset} \frac{R_i(h)}{||h||} = \emptyset .$$

y por el resultado de unicidad que ya tenemos para el caso de funciones real valuadas; las funciones $L_i: \mathbb{R}^n \to \mathbb{R}$ que son las funciones coordenadas de la función L, deben ser las diferenciales en p de las correspondientes funciones coordenadas de f, i.e. para $i=1,2,\ldots,m$ $L_i = L_{fi,p}$ de donde obtenemos la conclusión deseada. "

Con estos antecedentes se justifica la terminología usual:

"Si $f: D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ es diferenciable en $p \in D$. Entonces la única función lineal denotada por $L_{f,p}$, que satisface:

$$\lim_{h \to \emptyset} \frac{||f(p+h) - f(p) - L_{f,p}(h)||}{||h||} = \emptyset$$

se llama "la" diferencial de f en p."

Se destacan nuevamente las conveniencias de la notación que se ha usado y recomendado consistentemente a lo largo del trabajo para denotar a la diferencial de una función en un punto: $L_{f,p}$. Al aparecer tanto la función como el punto de manera explícita, la hacen una notación funcional para trabajar, por ejemplo, en las composiciones. A su vez, el énfasis está puesto en la letra L que pone de manifiesto la naturaleza de la diferencial como una función lineal.

Conviene que los estudiantes sepan cómo usar la notación, un error que puede presentarse, por ejemplo, es cuando se afirma que una función lineal L es diferenciable, y que su diferencial en cada punto $p \in \mathbb{R}^n$ es ella misma, algunos estudiantes podrían escribir lo anterior como $L_{f,p} = L(p) \quad \forall\ p \in \mathbb{R}^n$. Lo que no tiene sentido, pues el lado izquierdo es una función, mientras que el lado derecho es un vector de $\mathbb{R}^m$, lo correcto es entonces escribir:

$$L_{f,p} = L \quad \forall\ p \in \mathbb{R}^n.$$

Por ser la diferencial de una función f en un punto p; $L_{f,p}$: $\mathbb{R}^n \rightarrow \mathbb{R}^m$ una función lineal, los estudiantes ya están habituados a asociar a dicha función la matriz que la representa. A esto va dirigido el siguiente teorema:

Teorema 2. Si la función $f: D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ es diferenciable en $p \in D$, entonces todas las derivadas parciales $\left(\dfrac{\partial f_i}{\partial x_j}\right)(p)$ existen, y la matriz que representa a la diferencial $L_{f,p}$ (con respecto a las bases canónicas de $\mathbb{R}^n$ y $\mathbb{R}^m$) es la matriz de mxn dada por:

$$\begin{bmatrix} \dfrac{\partial f_1}{\partial x_1}(p) & \dfrac{\partial f_1}{\partial x_2}(p) & \cdots & \dfrac{\partial f_1}{\partial x_n}(p) \\ \dfrac{\partial f_2}{\partial x_1}(p) & \dfrac{\partial f_2}{\partial x_2}(p) & \cdots & \dfrac{\partial f_2}{\partial x_n}(p) \\ \vdots & \vdots & & \vdots \\ \dfrac{\partial f_m}{\partial x_1}(p) & \dfrac{\partial f_m}{\partial x_2}(p) & \cdots & \dfrac{\partial f_m}{\partial x_n}(p) \end{bmatrix} = \begin{bmatrix} \dfrac{\partial f_i}{\partial x_j}(p) \end{bmatrix}$$

Dem. Si f es diferenciable en p, entonces cada función coordenada f_i $1 \leq i \leq m$ también lo es, y por resultados previos, las derivadas parciales

$$\frac{\partial f_i}{\partial x_j}(p) \quad 1 \leq j \leq n \text{ existen para cada } i \in \{1,2,\ldots m\}.$$

Tenemos entonces que para cada $1 \leq j \leq n$:

$$L_{f,p}(e_j) = (L_{f_1,p}(e_j), L_{f_2,p}(e_j), \ldots\ldots, L_{f_m,p}(e_j))$$

$$= \left(\frac{\partial f_1}{\partial x_j}(p), \frac{\partial f_2}{\partial x_j}(p), \ldots\ldots\ldots, \frac{\partial f_m}{\partial x_j}(p)\right)$$

Pero las coordenadas de $L_{f,p}(e_j)$ con respecto a la base canónica de $\mathbb{R}^m$ forma la j-ésima columna de la matriz que representa a $L_{f,p}$ de donde se sigue la afirmación del teorema.

Si $f: D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ es diferenciable en $p \in D$, la matriz de $m \times n$ cuyas entradas son las derivadas parciales $\left[\frac{\partial f_i}{\partial x_j}(p) \right]$ se llama la matriz Jacobiana de f en p y la denotamos por $J_{f,p}$.

Se observa que la matriz Jacobiana de una función diferenciable $f: D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ es una matriz de $m \times n$, como corresponde a cualquier matriz, que represente a una transformación lineal de $\mathbb{R}^n$ a $\mathbb{R}^m$."

Los vectores $L_{f,p}(e_j) \in \mathbb{R}^m$ $1 \leq j \leq n$ cuyas coordenadas forman las columnas de la matriz $J_{f,p}$ son importantes en el estudio de una función $f: D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ que sea diferenciable en $p \in D$. Una notación especialmente conveniente es denotarlos por:

$$\frac{\partial f}{\partial x_j}(p) = \left(\frac{\partial f_1}{\partial x_j}(p), \frac{\partial f_2}{\partial x_j}(p), \ldots\ldots, \frac{\partial f_m}{\partial x_j}(p) \right) = L_{f,p}(e_j)$$

Es una notación conveniente porque sugiere que al vector $f(x) = (f_1(x), f_2(x), \ldots, f_m(x))$ $x = (x_1, \ldots, x_n)$ hay que derivarlo con respecto a la j-ésima variable mientras que las demás variables permanecen fijas, al modo que se tenía para las funciones vectoriales de una variable; en donde para derivar un vector se derivaba cada función coordenada:

$$\frac{\partial f}{\partial x_j} = \frac{\partial}{\partial x_j}(f_1, f_2, \ldots, f_m) = \left(\frac{\partial f_1}{\partial x_j}, \frac{\partial f_2}{\partial x_j}, \ldots\ldots, \frac{\partial f_m}{\partial x_j} \right)$$

Se puede incluso dar una interpretación geométrica útil para los estudiantes de los vectores $\frac{\partial f}{\partial x_j}(p)$, que se hará en el siguiente capítulo, como vectores tangentes a ciertas curvas.

Además la notación anterior, permite una formulación especialmente sencilla de algunos resultados; concretamente podemos reescribir a la diferencial $L_{f,p}$:

"Si $h \in \mathbb{R}^n$ $h = (h_1, h_2, \ldots, h_n)$

$$L_{f,p}(h) = L_{f,p}(h_1e_1+h_2e_2+\ldots+h_ne_n) = h_1 \frac{\partial f}{\partial x_1}(p)+\ldots+h_n \frac{\partial f}{\partial x_n}(p)$$

La condición que $f\colon D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ fuese diferenciable en $p \in D$ implicaba que: $f(p+h)-f(p) - L_{f,p}(h) = R(h)$

$$\text{i.e.}\quad f(p+h)-f(p) - \left[h_1\frac{\partial f}{\partial x_1}(p)+h_2\frac{\partial f}{\partial x_2}(p)+\ldots+h_n\frac{\partial f}{\partial x_n}(p)\right] = R(h)$$

$$\text{y} \lim_{h \to \emptyset} \frac{R(h)}{||h||} = \emptyset \text{ ."}$$

Aquí $R(h)$ expresa una diferencia de vectores en $\mathbb{R}^m$, y así escrita es fácil de recordar por la semejanza formal que guarda con la expresión que ya se conoce para el caso de funciones real valuadas.

"Como una consecuencia inmediata de lo que sucedía en el caso de funciones real valuadas, tenemos que la sola existencia de las derivadas parciales $\frac{\partial f_i}{\partial x_j}(p)$ $1 \leq i \leq m$ $1 \leq j \leq n$ no implica que la función f sea diferenciable en p; pero al igual que entonces (y por los resultados correspondientes) si las derivadas parciales existen en una vecindad del punto p y son continuas en p, entonces sí se satisface que f es diferenciable en p. Una función $f: D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ cuyas derivadas parciales existan y sean continuas en D, se dice que es continuamente diferenciable en D, o que es C^1 en D. Tenemos pues que cualquier función C^1 en D es diferenciable en D."

Sería útil señalar a los estudiantes que cuando se dice de una función $f: D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ que sus derivadas parciales existen en algún punto $p \in D$, hay que entender que si pensamos en las funciones coordenadas de f, $f = (f_1, f_2, \ldots, f_m)$ estamos pidiendo que para cada una de ellas, las derivadas parciales con respecto a $x_1, x_2, \ldots, x_n$ existan en p.

A manera de resumen, se puede proponer a los estudiantes un "esquema operativo" como se hizo al hablar de las funciones real valuadas de varias variables:

"Dada una función $f: D \subseteq \mathbb{R}^n \to \mathbb{R}^m$, calculamos todas sus derivadas parciales. En el conjunto $D' \subseteq D$ en donde todas las parciales sean continuas, la función es diferenciable.

Es decir, que $L_{f,p}$ existe para cada $p \in D'$ y para cada $p \in D'$, $L_{f,p}$ es la transformación lineal representada por la matriz Jacobiana de f en p."

El profesor tiene la experiencia de lo útil que resulta disponer de una interpretación geométrica de los conceptos. Por ejemplo en el caso de una función $f: D \subseteq \mathbb{R} \to \mathbb{R}$, si ésta es derivable en el punto $p \in D$, asociamos a este hecho la imagen de una curva en $\mathbb{R}^2$ (que corresponde a la gráfica de la función) y de una recta tangente a dicha curva en el punto $(p, f(p))$. Más aún, con el tratamiento que se hizo de ese caso, los estudiantes han aprendido que para puntos cercanos a p, la función cuya gráfica es la recta tangente les asigna valores aproximados a los que les asigna la función f.

"Para el caso de una función $f: D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ que sea diferenciable en algún punto $p \in D$, puede resultar conveniente pensar, no en la gráfica de f (que sería una especie de "superficie" en $\mathbb{R}^n \times \mathbb{R}^m$), sino en la imagen de f que consiste en una versión distorsionada de D "metida" en $\mathbb{R}^m$. Al ser $L_{f,p}$ $\mathbb{R}^n \rightarrow \mathbb{R}^m$ lineal, su imagen $L_{f,p}(\mathbb{R}^n)$ es un subespacio vectorial de $\mathbb{R}^m$ que trasladamos al punto $f(p) \in \mathbb{R}^m$, dicho subespacio vectorial trasladado se denomina espacio tangente a la imagen de f en el punto $f(p)$(Ver Figura 6.11).

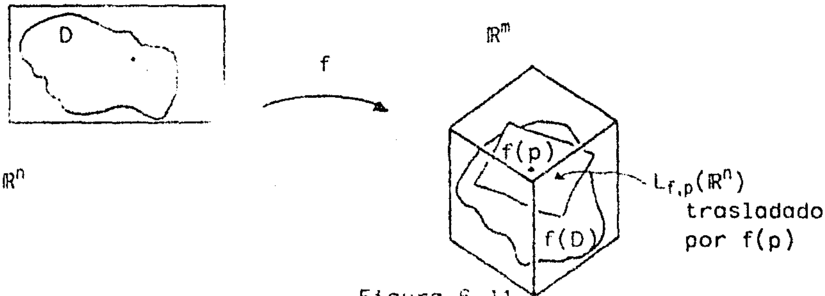


Figura 6.11

Análogamente para este caso, para puntos h cercanos a p, los valores que la función lineal $L_{f,p}$ asigna a dichos puntos estarán cercanos a las imágenes de esos puntos bajo $f(p+h) - f(p)$. Técnicamente, la función $h \longrightarrow f(p)+L_{f,p}(h)$ es una aproximación de primer orden a la función $h \longrightarrow f(p+h)$."

De esta forma quedan cubiertas tanto la perspectiva analítica como la geométrica de la diferenciabilidad para el caso de las funciones vectoriales.

Un elemento que facilita el aprendizaje de este tema es remitirse a ejemplos numéricos específicos, a manera de sugerencia se trabajan dos de ellos:

"Ejemplo 1. Sea $f:\mathbb{R}^3 \rightarrow \mathbb{R}^2$ $f(x,y,z) = (2x^2+y^2+z,\ x^2-y^2)$ vemos que las parciales de f son las funciones:

$$\frac{\partial f_1}{\partial x}(x,y,z) = 4x \qquad \frac{\partial f_1}{\partial y}(x,y,z) = 2y \qquad \frac{\partial f_1}{\partial z}(x,y,z) = 1$$

$$\frac{\partial f_2}{\partial x}(x,y,z) = 2x \qquad \frac{\partial f_2}{\partial y}(x,y,z) = -2y \qquad \frac{\partial f_2}{\partial z}(x,y,z) = 0$$

Como todas las parciales resultaron ser funciones continuas sobre $\mathbb{R}^3 \Rightarrow f$ es diferenciable para cualquier punto de $\mathbb{R}^3$, en particular si $p = (0,1,2)$ tenemos que $L_{f,(0,1,2)}$ es la transformación lineal de $\mathbb{R}^3$ a $\mathbb{R}^2$ representada por la matriz:

$$\begin{bmatrix} \dfrac{\partial f_1}{\partial x}(0,1,2) & \dfrac{\partial f_1}{\partial y}(0,1,2) & \dfrac{\partial f_1}{\partial z}(0,1,2) \\ \\ \dfrac{\partial f_2}{\partial x}(0,1,2) & \dfrac{\partial f_2}{\partial y}(0,1,2) & \dfrac{\partial f_2}{\partial z}(0,1,2) \end{bmatrix} = \begin{bmatrix} 0 & 2 & 1 \\ \\ 0 & -2 & 0 \end{bmatrix}$$

$$\therefore \quad L_{f,(0,1,2)}(h) = \begin{bmatrix} 0 & 2 & 1 \\ 0 & -2 & 0 \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ h_3 \end{bmatrix} = \begin{bmatrix} 2h_2 + h_3 \\ -2h_2 \end{bmatrix}$$

El hecho de que $L_{f,p}$ aproxima a f en una vecindad de p, quiere decir que si h es pequeño, el punto de $\mathbb{R}^2$ dado por $f(p) + L_{f,p}(h)$ está cercano al correspondiente dado por $f(p+h)$. Para fijar ideas, si $h = (h_1,h_2,h_3) = (0.01, 0.03, 0.02)$

$L_{f,p}(h) = (2(0.03)+0.02,\ -2(0.03)) = (0.08,\ -0.06)$

$f(p) = f(0,1,2) = (3,-1)$

$\therefore\ f(p) + L_{f,p}(h) = (3.08,\ -1.06)$

Mientras que $f(p+h) = f(0.01, 1.03, 2.02) = (3.0811,\ -1.0608)$

Ejemplo 2. Consideremos la función $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ dada por $f(x,y) = (|x-y|,\ x+y)$ las parciales de f son:

$$\frac{\partial f_1}{\partial x}(x,y) = \frac{x-y}{|x-y|} \qquad \frac{\partial f_1}{\partial y}(x,y) = -\frac{(x-y)}{|x-y|}$$

$$\frac{\partial f_2}{\partial x}(x,y) = 1 \qquad \frac{\partial f_2}{\partial y}(x,y) = 1.$$

Las parciales de f existen y son continuas en el complemento (en $\mathbb{R}^2$) de la recta $y = x$ $\Rightarrow$ para cualquier $(x,y) \in \mathbb{R}^2$ con $x \neq y$, f es diferenciable en (x,y). Si $(x,y) \in \mathbb{R}^2$ con $x \neq y$, la matriz Jacobiana de f en (x,y) estará dada entonces por

$$J_{f,(x,y)} = \begin{bmatrix} \dfrac{\partial f_1}{\partial x}(x,y) & \dfrac{\partial f_1}{\partial y}(x,y) \\ \\ \dfrac{\partial f_2}{\partial x}(x,y) & \dfrac{\partial f_2}{\partial y}(x,y) \end{bmatrix} = \begin{bmatrix} \dfrac{x-y}{|x-y|} & -\dfrac{(x-y)}{|x-y|} \\ \\ 1 & 1 \end{bmatrix}$$

$\Rightarrow$ si $p = (x,y) \in \mathbb{R}^2$ con $x \neq y$, la diferencial de f en p; $L_{f,p}: \mathbb{R}^2 \to \mathbb{R}^2$ está dada por:

$$L_{f,p}(h) = \left(\frac{(x-y)}{|x-y|} h_1 - \frac{(x-y)}{|x-y|} h_2,\ h_1 + h_2 \right)$$

si $h = (h_1,h_2) \in \mathbb{R}^2$. "

CAPITULO 7.

LA REGLA DE LA CADENA.

Una de las aplicaciones más útiles de la presentación de la diferenciabilidad en los términos en los que se ha hecho en los capítulos anteriores, es la simplicidad que toma la formulación para la Regla de la Cadena.

Es muy frecuente que en los textos de Cálculo al hablar de la manera de derivar parcialmente una composición de funciones de varias variables proporcionen una lista de posibles casos con una fórmula correspondiente para cada caso.

A manera de ejemplo, citemos un texto de la literatura tradicional sobre el tema:

"Supóngase ...que f(u,v) es una función diferenciable de u,v que a su vez son funciones diferenciables de una variable independiente x. Entonces..

$$\frac{df}{dx} = \frac{\partial f}{\partial u}\frac{du}{dx} + \frac{\partial f}{\partial v}\frac{dv}{dx}$$

Considérese ahora la función f(u,v) cuando u,v son funciones diferenciables de dos variables independientes x,y. Entonces.. $F(x,y) = f(u(x,y), v(x,y))$ es una función diferenciable de x,y cuyas derivadas parciales están dadas por:

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial x}$$

$$\frac{\partial f}{\partial y} = \frac{\partial f}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial y}$$

Cuando f(u,v,w) es diferenciable y u,v,w son funciones de x, tenemos..

$$\frac{df}{dx} = \frac{\partial f}{\partial u}\frac{du}{dx} + \frac{\partial f}{\partial v}\frac{dv}{dx} + \frac{\partial f}{\partial w}\frac{dw}{dx}$$

Cuando u, v, w son funciones... de tres variables x, y, z tenemos:

$$f_x = f_u u_x + f_v v_x + f_w w_x$$
$$f_y = f_u u_y + f_v v_y + f_w w_y$$
$$f_z = f_u u_z + f_v v_z + f_w w_z \text{ ."}$$

(Brand, Louis. Cálculo Avanzado. México, D.F.; CECSA; 1961; cfr. pág. 204 - 208).

Todos estos distintos casos quedan unificados al presentar la Regla de la Cadena desde la perspectiva de las transformaciones lineales.

Los ingredientes necesarios para esta presentación son el tener claro el concepto mismo de composición de funciones para funciones vectoriales de varias variables:

"Definición. Sean $g: E \subseteq \mathbb{R}^p \to \mathbb{R}^m$ y $f: D \subseteq \mathbb{R}^m \to \mathbb{R}^n$ tales que el dominio D de f contenga a la imagen $g(E)$ de g. La función composición $f \circ g: E \subseteq \mathbb{R}^p \to \mathbb{R}^n$ está definida por:

$(f \circ g)(x) = f(g(x)) \quad x \in E$. (Ver Figura 7.1)

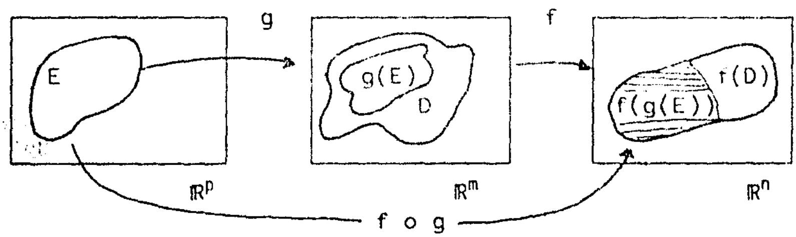


Figura 7.1

Ejemplo 1. Sean $g: \mathbb{R}^3 \to \mathbb{R}^2$, $f: \mathbb{R}^2 \to \mathbb{R}^2$ definidas como:

$g(t_1, t_2, t_3) = (t_1t_2, t_2t_3) \qquad (t_1, t_2, t_3) \in \mathbb{R}^3$
$f(x_1, x_2) = (\text{sen } x_1, x_1x_2) \qquad (x_1, x_2) \in \mathbb{R}^2$
Entonces
$(f \circ g)(t_1, t_2, t_3) = f(g(t_1, t_2, t_3)) = f(t_1t_2, t_2t_3) =$
$= (\text{sen}(t_1t_2), t_1t_2^2t_3)$.

Ejemplo 2. $T: \mathbb{R}^2 \to \mathbb{R}^3$ dada por $T(x,y) = (xy, 2x, -y)$

$S: \mathbb{R}^3 \to \mathbb{R}^2$ dada por $S(x,y,z) = (x-y, yz)$

$\Rightarrow$ $S \circ T: \mathbb{R}^2 \to \mathbb{R}^2$. Para calcular la composición, S o T, dadas las ecuaciones que describen a S y a T es conveniente usar las mismas variables en ambas transformaciones para designar a los puntos del espacio de "en medio". Por tanto si T manda (x,y) en (r,s,t) y S manda (r,s,t) en (u,v) como muestra la Figura 7.2, y si las ecuaciones para T y S son:

$$T: \begin{cases} r = xy \\ s = 2x \\ t = -y \end{cases} \qquad S: \begin{cases} u = r - s \\ v = st \end{cases}$$

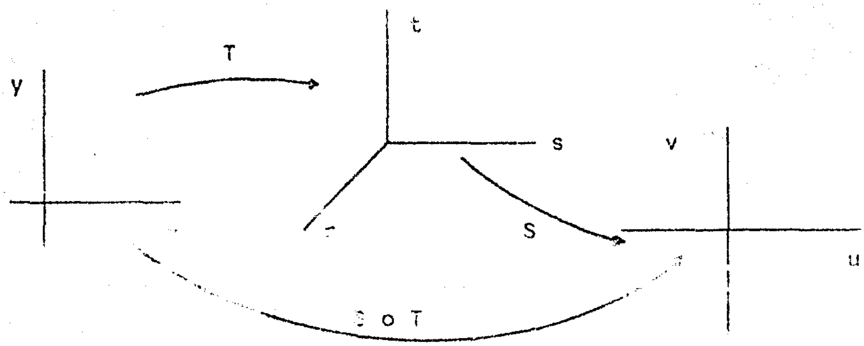


Figura 7.2

Entonces S o T manda (x,y) en (u,v) y las ecuaciones que describen a S o T están dadas por:

$$S \circ T: \begin{cases} u = xy - 2x \\ v = -2xy \end{cases}$$

En este ejemplo particular también podemos formar la composición $T \circ S: \mathbb{R}^3 \to \mathbb{R}^3$. Para dar las ecuaciones que describen a esta composición podemos reformular las ecuaciones de las transformaciones S y T en términos de variables nuevas y suponer que S manda a (x,y,z) en (u,v) y T manda a (u,v) en (r,s,t). De esta forma:

$$S: \begin{cases} u = x - y \\ v = yz \end{cases} \qquad T: \begin{cases} r = uv \\ s = 2u \\ t = -v \end{cases}$$

de donde

$$T \circ S = \begin{cases} r = (x - y)yz \\ s = 2(x - y) \\ t = -yz \end{cases}$$

Claramente, por otra parte, $S \circ T \neq T \circ S$."

La Regla de la Cadena se enuncia entonces como:

"Sea $g: E \subseteq \mathbb{R}^p \to \mathbb{R}^m$ ($E \subseteq \mathbb{R}^p$ un conjunto abierto), y sea $f: D \subseteq \mathbb{R}^m \to \mathbb{R}^n$ ($D \subseteq \mathbb{R}^m$ un conjunto abierto), y $g(E) \subseteq D$. Sea $F: E \subseteq \mathbb{R}^p \to \mathbb{R}^n$ la función composición de f y g, i.e. $F(t) = (f \circ g)(t) = f(g(t))$ $t \in E$.

Supongamos que g es diferenciable en $a \in E$ y f es diferenciable en $g(a) \in D$. Entonces F es diferenciable en a y su diferencial está dada por:

$$(1)\ldots\ldots \quad L_{F,a} = L_{f,g(a)} \circ L_{g,a}$$

Y la matriz Jacobiana de F en a, está dada por el siguiente producto de matrices:

$$(2)\ldots\ldots \quad J_{F,a} = J_{f,g(a)}\, J_{g,a}$$

Dem. Sea $k = g(a+h)-g(a) \Rightarrow g(a)+k = g(a+h)$ y $(f \circ g)(a+h) = f(g(a+h)) = f(g(a)+k)$.

Como f es diferenciable en g(a)
$\Rightarrow f(g(a)+k) - f(g(a)) - L_{f,g(a)}(k) = R_1(k)$

$$\text{donde} \quad \lim_{k \to 0} \frac{R_1(k)}{||k||} = 0$$

A su vez, g es diferenciable en a
$\therefore k = g(a+h) - g(a) = L_{g,a}(h) + R_2(h)$

$$\text{donde} \quad \lim_{h \to 0} \frac{R_2(h)}{||h||} = 0$$

$$(f \circ g)(a+h) = f(g(a+h)) = f(g(a)+k) = f(g(a)) + L_{f,g(a)}(k) + R_1(k)$$
$$= (f \circ g)(a) + L_{f,g(a)}(k) + R_1(k)$$
pero $k = L_{g,a}(h) + R_2(h)$
$$\therefore (f \circ g)(a+h) = (f \circ g)(a) + L_{f,g(a)}(L_{g,a}(h) + R_2(h)) + R_1(k)$$
$$(f \circ g)(a+h) = (f \circ g)(a) + (L_{f,g(a)} \circ L_{g,a})(h) + R(h)$$
donde (*)..... $R(h) = L_{f,g(a)}(R_2(h)) + R_1(k)$.

Para probar que $(f \circ g)$ es diferenciable en a, bastaría probar que
$$\lim_{h \to 0} \frac{||R(h)||}{||h||} = 0$$

Como f y g son diferenciables $\Rightarrow$
$||R_1(k)|| < \varepsilon ||k||$ (si $||k||$ es suficientemente pequeña)

$||R_2(h)|| < \varepsilon ||h||$ (si $||h||$ es suficientemente pequeña)

El hecho de que las transformaciones lineales son de Lipschitz (cfr. capítulo 6) nos permite saber que existe un número M tal que:
$$||L_{g,a}(h)|| \leq M||h||$$
$$||L_{f,g(a)}(R_2(h))|| \leq M||R_2(h)|| \quad \text{(para cualquier } h\text{)}.$$

Como $k = L_{g,a}(h) + R_2(h)$ tenemos que:
$$||k|| \leq M||h|| + ||R_2(h)|| \leq M||h|| + \varepsilon||h|| < (M+1)||h||.$$

Y usando este hecho en (*):
$$||R(h)|| \leq ||L_{f,g(a)}(R_2(h))|| + ||R_1(k)||$$
$$\leq M||R_2(h)|| + \varepsilon||k|| \leq M\varepsilon||h|| + \varepsilon(M+1)||h|| \leq (2M+1)\varepsilon||h||$$
$$\Rightarrow \quad ||R(h)|| \,/\, ||h|| \leq (2M+1)\,\varepsilon$$

Como $\varepsilon > 0$ es arbitrariamente pequeño
$$\Rightarrow \lim_{h \to 0} \frac{||R(h)||}{||h||} = 0$$
$\Rightarrow$ (por resultados del capítulo 6)
$F = (f \circ g)$ es diferenciable en a y su diferencial en a, $L_{F,a}$, es la función lineal $L_{F,a}: \mathbb{R}^p \to \mathbb{R}^n$ dada por:
$$L_{F,a} = L_{f,g(a)} \circ L_{g,a}$$

La expresión (2) se sigue de inmediato del hecho de que la matriz de una composición de funciones lineales es el producto de las matrices, con lo que queda demostrado el Teorema."

De la expresión $J_{F,a} = J_{f,g(a)} \; J_{g,a}$ se pueden obtener las formas de enunciar la Regla de la Cadena que aparecen en los textos. Ya que la entrada (i,j) de $J_{F,a}$ es el producto punto del i-ésimo renglón de $J_{f,g(a)}$ y la j-ésima columna de $J_{g,a}$. Si denotamos a las variables en E por t y a las variables en D por x, tenemos que para $a \in E$:

$$(3)\ldots \quad \frac{\partial F_i}{\partial t_j}(a) = \frac{\partial f_i}{\partial x_1}(g(a)) \frac{\partial g_1}{\partial t_j}(a) + \ldots + \frac{\partial f_i}{\partial x_m}(g(a)) \frac{\partial g_m}{\partial t_j}(a)$$

Es entonces en este momento cuando se puede mencionar a los estudiantes qua la expresión (3) es lo que corresponde a las fórmulas que presentan los libros al hablar de la Regla de la Cadena. Aunque suele escribirse de una forma simplificada. El razonamiento que en el fondo se hace y que es importante hacerlo explícito a los estudiantes es el siguiente:

Al considerar una función vectorial f de m variables $x_1,\ldots,x_m$; cada una de las cuales es función de las variables $t_1,\ldots,t_p$ la expresión:

$$(4)\ldots \quad f(x_1(t_1,t_2,\ldots,t_p), x_2(t_1,t_2,\ldots,t_p),\ldots,x_m(t_1,\ldots,t_p))$$

lleva a considerar a f como función de las variables $t_1,\ldots,t_p$. Si definimos una función $g\colon E \subseteq \mathbb{R}^p \to \mathbb{R}^m$ como:

$$g(t_1,\ldots,t_p) = (x_1,\ldots,x_m) = (x_1(t_1,\ldots,t_p),\ldots,x_m(t_1,\ldots,t_p))$$
$$(t_1,\ldots,t_p) \in E.$$

Entonces (4) da los valores de la función:

$$F\colon E \subseteq \mathbb{R}^p \to \mathbb{R}^n,\ F(t_1,\ldots,t_p) = (F_1(t_1,\ldots,t_p),\ldots,F_n(t_1,\ldots,t_p))$$
$$F(t_1,\ldots,t_p) = (f \circ g)(t_1,\ldots,t_p) \qquad (t_1,\ldots,t_p) \in E.$$

La Regla de la Cadena es simplificada a la identidad (compárese con (3)):

$$(5)\ldots\ldots \quad \frac{\partial f_i}{\partial t_j} = \frac{\partial f_i}{\partial x_1}\,\frac{\partial x_1}{\partial t_j} + \frac{\partial f_i}{\partial x_2}\,\frac{\partial x_2}{\partial t_j} + \ldots\ldots + \frac{\partial f_i}{\partial x_m}\,\frac{\partial x_m}{\partial t_j}$$

Donde $\dfrac{\partial f_i}{\partial t_j}$ es realmente $\dfrac{\partial F_i}{\partial t_j}$ y $\dfrac{\partial x_k}{\partial t_j}$ es realmente $\dfrac{\partial g_k}{\partial t_j}$

y si $\dfrac{\partial f_i}{\partial t_j}$ ha de evaluarse en $t \in \mathbb{R}^p$ usando (5)

entonces $\dfrac{\partial x_k}{\partial t_j}$ debe evaluarse en t y $\dfrac{\partial f_i}{\partial x_k}$ debe evaluarse en

$(x_1(t), x_2(t), \ldots , x_m(t))$.

Una observación muy importante para hacerla notar a los estudiantes es que esta manera de enunciar la Regla de la Cadena para el caso de funciones vectoriales proporciona una auténtica generalización de la Regla de la Cadena que se manejaba en el Cálculo elemental (de funciones reales de una variable) en la forma: $\dfrac{dy}{dt} = \left[\dfrac{dy}{dx}\right]\left[\dfrac{dx}{dt}\right]$:

"Sea $g\colon \mathbb{R} \to \mathbb{R}$ diferenciable en $a \in \mathbb{R}$. La diferencial $L_{g,a}$ está definida como $L_{g,a}(h) = (g'(a))h$ y la matriz Jacobiana de g en a es la matriz de 1x1: $J_{g,a} = (g'(a))$. Supongamos que $f\colon \mathbb{R} \to \mathbb{R}$ es diferenciable en $g(a)$ entonces la Regla de la Cadena implica que f o g es diferenciable en a. Más aún como

$$J_{f,g(a)} = (f'(g(a))) \text{ y } J_{g,a} = (g'(a))$$

la expresión (2) es una afirmación acerca de matrices de 1 x 1:

$$J_{f o g,a} = ((f \circ g)'(a)) = J_{f,g(a)}J_{g,a} = (f'(g(a)))\ (g'(a))$$

$$\text{i.e. } (f \circ g)'(a) = f'(g(a))g'(a).$$

que es la forma elemental de la Regla de la Cadena."

Con esta presentación se tiene la ventaja de que los estudiantes pueden aplicar expresiones del tipo (5), que operativamente resultan útiles, pero teniendo claro como interpretar adecuadamente ese formalismo. Pues la expresión (5) es una versión simplificada de (3) que permite recordarla con mayor facilidad, pero tal como aparece escrita, está en términos menos precisos que la (3). Es instructivo mostrar a los estudiantes el tipo de errores en los que se puede incurrir cuando se efectúan manipulaciones poco cuidadosas con expresiones del tipo (5):

"Considérese el siguiente "argumento":
Si $z=f(x,y)$, $y=g(x,t)$ donde f y g son funciones real valuadas. Entonces z está definida como una función de x,t.
Por la Regla de la Cadena:

$$(*)\quad \frac{\partial z}{\partial x} = \frac{\partial z}{\partial x}\,\frac{\partial x}{\partial x} + \frac{\partial z}{\partial y}\,\frac{\partial y}{\partial x} = \frac{\partial z}{\partial x} + \frac{\partial z}{\partial y}\,\frac{\partial y}{\partial x}$$

$$\text{entonces}\quad \frac{\partial z}{\partial y}\,\frac{\partial y}{\partial x} = \emptyset$$

En particular, si $z = x^2 + y$; $y = x + t$.

$$\text{Entonces}\quad \frac{\partial z}{\partial y} = 1 \qquad \frac{\partial y}{\partial x} = 1$$

$$\therefore\ 1 = \emptyset$$

El error del razonamiento anterior consiste en el abuso de notación que identifica $\frac{\partial z}{\partial x}$ en la izquierda de (*) con $\frac{\partial z}{\partial x}$ en la derecha, cuando en realidad son cantidades distintas. Si llamamos h al cambio de coordenadas: $(x,t) \xrightarrow{h} (x,y)$, entonces el primero es $\frac{\partial}{\partial x}(foh)(x,t)$ y el segundo es $\frac{\partial f}{\partial x}(x,y)$ $\left(= \frac{\partial f}{\partial x} oh(x,t) \right)$."

Resulta fundamental ilustrar el uso de la Regla de la Cadena a través de ejemplos concretos:

"Ejemplo 3. $g: \mathbb{R}^3 \to \mathbb{R}^2$
$f: \mathbb{R}^2 \to \mathbb{R}^2$

$g(t_1,t_2t_3) = (t_1t_2, t_2t_3) \qquad (t_1,t_2,t_3) \in \mathbb{R}^3$
$f(x_1,x_2) = (2x_1 + x_2^2,\ 3x_1^2 - x_2) \qquad (x_1,x_2) \in \mathbb{R}^2$

la función composición $F = f \circ g: \mathbb{R}^3 \to \mathbb{R}^2$ está dada por:
$F(t_1, t_2, t_3) = f(g(t_1, t_2, t_3)) = f(t_1 t_2, t_2 t_3)$
$= (2t_1 t_2 + t_2^2 t_3^2,\ 3t_1^2 t_2^2 - t_2 t_3).$

De esta manera se puede calcular la matriz Jacobiana de F en t o bien directamente o haciendo uso de la Regla de la Cadena: $J_{F,t} = J_{f,g(t)}\, J_{g,t}$.

Obsérvese que $g(t) = (x_1(t),\ x_2(t))$
con $x_1(t) = t_1 t_2$ $\quad x_2(t) = t_2 t_3$

$$J_{f,g(t)} = \begin{bmatrix} 2 & 2x_2(t) \\ 6x_1(t) & -1 \end{bmatrix} \qquad J_{g,t} = \begin{bmatrix} t_2 & t_1 & \emptyset \\ \emptyset & t_3 & t_2 \end{bmatrix}$$

$$\therefore\ J_{F,t} = \begin{bmatrix} 2 & 2t_2t_3 \\ 6t_1t_2 & -1 \end{bmatrix} \begin{bmatrix} t_2 & t_1 & \emptyset \\ \emptyset & t_3 & t_2 \end{bmatrix} = \begin{bmatrix} 2t_2 & 2t_1+2t_2t_3^2 & 2t_2^2t_3 \\ 6t_1t_2^2 & 6t_1^2-t_3 & -t_2 \end{bmatrix}"$$

Se puede hacer notar que, alternativamente, cada una de las entradas de $J_{F,t}$ se pueden calcular usando (5), por ejemplo:

$$\frac{\partial f_1}{\partial t_2} = \frac{\partial f_1}{\partial x_1}\frac{\partial x_1}{\partial t_2} + \frac{\partial f_1}{\partial x_2}\frac{\partial x_2}{\partial t_2} = 2t_1 + 2x_2 t_3 = 2t_1 + 2t_2 t_3^2.$$

Si se entiende la justificación de esta manera informal, por lo demás usual de trabajar, se puede hacer uso de ella, ya que, como se mencionó resulta útil operativamente.

"Ejemplo 4. Sean $g: E \subseteq \mathbb{R}^2 \to \mathbb{R}^2$
$f: \mathbb{R}^2 \to \mathbb{R}^2$
donde $E = \{(t_1, t_2) : t_1 > t_2\}$; definidas como

$g(t_1, t_2) = ((t_1 - t_2)^{1/2},\ t_1 + t_2) \quad (t_1, t_2) \in E$

$f(x_1, x_2) = ((x_2 + x_1^2)/2,\ (x_2 - x_1^2)/2) \quad (x_1, x_2) \in \mathbb{R}^2.$

La función compuesta $F = f \circ g: E \subseteq \mathbb{R}^2 \to \mathbb{R}^2$ es diferenciable en E pues tanto g como f son diferenciables sobre sus dominios. Podemos calcular la matriz Jacobiana $J_{F,t}$ escribiendo:

$x_1(t_1,t_2) = (t_1-t_2)^{1/2} \qquad x_2(t_1,t_2) = t_1 + t_2$

y encontramos que:

$$\frac{\partial F_1}{\partial t_1} = \frac{\partial f_1}{\partial x_1}\,\frac{\partial x_1}{\partial t_1} + \frac{\partial f_1}{\partial x_2}\,\frac{\partial x_2}{\partial t_1} = x_1 \left[\frac{1}{2(t_1-t_2)^{1/2}}\right] + \frac{1}{2}\cdot 1$$

$$= (t_1 - t_2)^{1/2} \left[\frac{1}{2(t_1-t_2)^{1/2}}\right] + \frac{1}{2} = \frac{1}{2} + \frac{1}{2} = 1.$$

$$\frac{\partial F_1}{\partial t_2} = \frac{\partial f_1}{\partial x_1}\,\frac{\partial x_1}{\partial t_2} + \frac{\partial f_1}{\partial x_2}\,\frac{\partial x_2}{\partial t_2} = x_1 \left[\frac{-1}{2(t_1-t_2)^{1/2}}\right] + \frac{1}{2} = 0$$

De la misma manera obtenemos que:

$$\frac{\partial F_2}{\partial t_1} = 0 \qquad \frac{\partial F_2}{\partial t_2} = 1$$

$\therefore$ La matriz Jacobiana $J_{F,t}$ es la matriz identidad I_2. $\therefore$ La diferencial de F en t, $L_{F,t}\colon \mathbb{R}^2 \to \mathbb{R}^2$ es la función identidad."

Es interesante resaltar que para derivar funciones compuestas, aun en casos más complicados, se puede hacer a partir del tratamiento propuesto para la Regla de la Cadena.

Considérese por ejemplo el siguiente conjunto de ecuaciones $u = f(x,y,z) \qquad z = g(x,y,t)$

$$y = h(x,t).$$

Estas expresan a u como función de x, t. Un diagrama de la relación entre las variables sería:

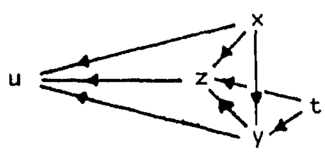


Figura 7.3

El sentido de la flecha indica que se es variable independiente de la variable que aparece en el lugar hacia donde se dirige la flecha.

Introducimos tres transformaciones R, S, T tales que:

$u = (S \circ T \circ R)(x,t)$ donde R, S, T están definidas como:
$R:\mathbb{R}^2 \to \mathbb{R}^3$ está dada por $T:\mathbb{R}^3 \to \mathbb{R}^3$ dada por $S:\mathbb{R}^3 \to \mathbb{R}$ dada por

$$R: \begin{cases} x = x \\ y = h(x,t) \\ t = t \end{cases} \qquad T: \begin{cases} x = x \\ y = y \\ z = g(x,y,t) \end{cases} \qquad u = f(x,y,z)$$

Esquemáticamente:

$$\begin{tikzcd} (x,t) \arrow[r, "R"] & (x,y,t) \arrow[r, "T"] & (x,y,z) \arrow[d, "S"] \\ & & u \end{tikzcd}$$

Para calcular las derivadas parciales
$\dfrac{\partial u}{\partial x}$, $\dfrac{\partial u}{\partial t}$; encontramos la matriz Jacobiana:

$$J_{SoToR,(x,t)} = \begin{bmatrix} \dfrac{\partial u}{\partial x} & \dfrac{\partial u}{\partial t} \end{bmatrix}$$

y por la Regla de la Cadena:
$J_{SoToR,(x,t)} = J_{S,T(R(x,t))} \; J_{T,R(x,t)} \; J_{R,(x,t)}$
de donde obtenemos, omitiendo los argumentos:

$$J_S = \begin{bmatrix} \dfrac{\partial f}{\partial x} & \dfrac{\partial f}{\partial y} & \dfrac{\partial f}{\partial z} \end{bmatrix} \qquad J_R = \begin{bmatrix} 1 & \emptyset \\ \dfrac{\partial h}{\partial x} & \dfrac{\partial h}{\partial t} \\ \emptyset & 1 \end{bmatrix}$$

$$J_T = \begin{bmatrix} 1 & \emptyset & \emptyset \\ \emptyset & 1 & \emptyset \\ \dfrac{\partial g}{\partial x} & \dfrac{\partial g}{\partial y} & \dfrac{\partial g}{\partial z} \end{bmatrix}$$

$\therefore$ El producto de $J_S J_T J_R$ da:

$$\begin{bmatrix} \dfrac{\partial f}{\partial x}+\dfrac{\partial f}{\partial z}\dfrac{\partial g}{\partial x}+\dfrac{\partial f}{\partial y}\dfrac{\partial h}{\partial x}+\dfrac{\partial f}{\partial z}\dfrac{\partial g}{\partial y}\dfrac{\partial h}{\partial x} & \dfrac{\partial f}{\partial y}\dfrac{\partial h}{\partial t}+\dfrac{\partial f}{\partial z}\dfrac{\partial g}{\partial y}\dfrac{\partial h}{\partial t}+\dfrac{\partial f}{\partial z}\dfrac{\partial g}{\partial t} \end{bmatrix}$$

que es lo que se obtendría para $\begin{bmatrix} \dfrac{\partial u}{\partial x} & \dfrac{\partial u}{\partial t} \end{bmatrix}$
si se utilizan las fórmulas correspondientes de la Regla de la Cadena para este caso: $u(x,t) = f(x,\ h(x,t),\ g(x,y,t))$.

El hecho de que si f y g están definidas de manera explícita; situación que se presenta con frecuencia en la práctica; se pueden calcular las derivadas parciales para $f \circ g$ sin necesidad de la Regla de la Cadena, simplemente obteniendo la expresión para $f \circ g$ en términos de las variables $t_1, \ldots, t_p$; hace que sean necesario motivar la presentación de la Regla de la Cadena. Esto se puede hacer señalando la importancia conceptual que tiene ese resultado para situaciones en donde esa forma directa no es posible.

Como ejemplo concreto de lo anterior, está la generalización del Teorema del Valor Medio:

"Teorema del Valor Medio. Sea $f: D \subseteq \mathbb{R}^m \to \mathbb{R}^n$ una función diferenciable. $D \subseteq \mathbb{R}^m$ un conjunto abierto que contiene a los puntos q y $q+h$ y al segmento que los une. Entonces, para cualquier vector $u \in \mathbb{R}^n$, existe $0 < \theta < 1$ tal que: $(f(q+h) - f(q)) \cdot u = L_{f,q+\theta h}(h) \cdot u$.

Dem. La función $\tilde f: D \subseteq \mathbb{R}^m \to \mathbb{R}$ definida como

$$\tilde f(x) = (f(x)-f(q)) \cdot u \qquad x \in D$$

es diferenciable. Sea $\{e_1, \ldots, e_m\}$ la base canónica de $\mathbb{R}^m$. Se tiene entonces para cada $k = 1, \ldots, m$

$$(*)\ldots\ldots \frac{\partial \tilde f}{\partial x_k}(x) = \frac{\partial f}{\partial x_k}(x) \cdot u = L_{f,x}(e_k) \cdot u \qquad x \in D$$

Ahora consideremos la función $x: E \subseteq \mathbb{R} \to \mathbb{R}^m$
$[0,1] \subseteq E$ y $x(E) \subseteq D$ dada por $x(t) = q + th$
y consideremos la composición: $F = \tilde f \circ x: E \subseteq \mathbb{R} \to \mathbb{R}$

$$\begin{tikzcd} t \arrow[r, "x"] \arrow[rr, bend right, "\tilde f \circ x"'] & x(t) = q + th \arrow[r, "\tilde f"] & \tilde f(q+th) = (f(q+th)-f(q)) \cdot u \end{tikzcd}$$

$F(t) = (\bar{f} \circ x)(t) = \bar{f}(x(t)) = \bar{f}(q+th) = (f(q+th)-f(q)) \cdot u$
x es diferenciable en E y $\bar{f}$ es diferenciable en $x(t)(t \in E)$
$\therefore$ F es diferenciable en E.

Y por la Regla de la Cadena

$$J_{F,t}=F'(t)=J_{\bar{f},x(t)}\ J_{x,t} = \left[\frac{\partial \bar{f}}{\partial x_1}(x(t)) \ldots \frac{\partial \bar{f}}{\partial x_m}(x(t)) \right] \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_m \end{bmatrix}$$

$$h = (h_1, h_2, \ldots, h_m) \in \mathbb{R}^m$$

$$\therefore \quad F'(t) = h_1 \frac{\partial \bar{f}}{\partial x_1}(x(t))+h_2 \frac{\partial \bar{f}}{\partial x_2}(x(t))+\ldots+ h_m \frac{\partial \bar{f}}{\partial x_m}(x(t))$$

$$\text{i.e.} \quad F'(t) = h_1 \frac{\partial \bar{f}}{\partial x_1}(q+th)+h_2 \frac{\partial \bar{f}}{\partial x_2}(q+th)+\ldots+ h_m \frac{\partial \bar{f}}{\partial x_m}(q+th)$$

Por (*)

$$F'(t) = (h_1 L_{f,q+th}(e_1)+h_2 L_{f,q+th}(e_2)+\ldots\ldots+h_m L_{f,q+th}(e_m)) \cdot u = L_{f,q+th}(h) \cdot u$$

El Teorema del Valor Medio aplicado a F implica que existe $0 < \theta < 1$ tal que $F(1) - F(0) = F'(\theta)$

$(f(q+h)-f(q)) \cdot u = F(1)-F(0) = F'(\theta) = L_{f,q+\theta h}(h) \cdot u$ L.c.q.d.

Un caso particular del teorema anterior que aparece con frecuencia es cuando $n = 1$ y $u = 1$. El teorema afirmaría entonces que para cualesquiera dos puntos $q, q+h$ en el dominio de una función diferenciable $f: D \subseteq \mathbb{R}^m \to \mathbb{R}$ existe $0 < \theta < 1$ tal que $f(q+h)-f(q) = L_{f,q+\theta h}(h)$

$$L_{f,q+\theta h}(h)= \left[\frac{\partial f}{\partial x_1}(q+\theta h) \quad \frac{\partial f}{\partial x_2}(q+\theta h) \ \ldots \ \frac{\partial f}{\partial x_m}(q+\theta h) \right] \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_m \end{bmatrix}$$

$$\therefore\ f(q+h)-f(q) = h_1 \frac{\partial f}{\partial x_1}(q+\theta h)+\ldots\ldots+ h_m \frac{\partial f}{\partial x_m}(q+\theta h)$$

De hecho el Teorema del Valor Medio del Cálculo elemental es a su vez un caso particular de este caso cuando $m = 1$. "

En la enseñanza de las Matemáticas un aprendizaje se facilita si es posible acudir a alguna interpretación geométrica de forma que los estudiantes tengan una imagen asociada al concepto que se pretende enseñar.

Es por esta razón que se propone introducir en este curso, la interpretación de la diferencial de una función vectorial de varias variables que suele hacerse en la Geometría diferencial. Ya que tiene la ventaja de ser de un carácter muy geométrico y por otra parte, a estas alturas del curso se dispone de todos los elementos necesarios para entenderla. (cfr. por ejemplo, Do Carmo Manfredo P., Differential Geometry of Curves and Surfaces; Prentice-Hall, 1976.)

"Dada una función $F:\ D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ diferenciable. Podemos interpretar $L_{F,p}(h)$ de la siguiente manera:

Si consideramos una curva diferenciable $\alpha:(-\varepsilon,\varepsilon) \to D$ tal que $\alpha(\emptyset)=p$ $\alpha'(\emptyset)=h$. (Lo anterior siempre es posible, pensemos por ejemplo en

$$t \xrightarrow{\alpha} p+th \qquad \alpha:(-\varepsilon,\varepsilon) \to D \subseteq \mathbb{R}^n)$$

Entonces por la Regla de la Cadena

la función $\beta = F \circ \alpha$

$\beta:\ (-\varepsilon,\varepsilon) \to \mathbb{R}^m$ es diferenciable

y $L_{F,p}(h) = \beta'(\emptyset)$.

Esta definición no depende de la elección de la curva α mientras ésta satisfaga que: $\alpha(\emptyset) = p$ $\alpha'(\emptyset) = h$.

Dem. Como $\alpha:(-\varepsilon,\varepsilon) \to \mathbb{R}^m$ es diferenciable sabemos que $L_{\alpha,t}(h) = J_{\alpha,t}h$ $h \in \mathbb{R}$ y $J_{\alpha,t} = (\alpha'(t))$ ($\alpha'(t)$ colocado como vector columna).

Como $F: D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ es diferenciable,
$\Rightarrow \beta = F \circ \alpha: (-\varepsilon, \varepsilon) \to \mathbb{R}^m$ es diferenciable y
$L_{\beta,t}(h) = J_{\beta,t}h \quad h \in \mathbb{R} \quad$ con $\quad J_{\beta,t} = (\beta'(t))$.

A su vez, por la Regla de la Cadena, si
$\beta = F \circ \alpha \Rightarrow J_{\beta,t} = J_{F,\alpha(t)} \; J_{\alpha,t}$
$J_{\beta,t} = J_{F,\alpha(t)} \, [\alpha'(t)]$
Evaluando en $t = \emptyset$
$J_{\beta,\emptyset} = [\beta'(\emptyset)] = J_{F,\alpha(\emptyset)}[\alpha'(\emptyset)] = J_{F,p}[h] = L_{F,p}(h)$.

Esto prueba la afirmación: $\beta'(\emptyset) = L_{F,p}(h)$ y el hecho de que la definición dada no depende de la elección de la curva α.

Es decir, que para calcular $L_{F,p}(h)$ pensamos en una curva que pase por el punto p, con vector de velocidad h al tiempo $t = \emptyset$ y restringimos F a los puntos sobre esa curva, de esta manera obtenemos una curva en $\mathbb{R}^m$ tal que al tiempo $t = \emptyset$ pasa por el punto F(p), se calcula su vector de velocidad al tiempo $t = \emptyset$ y dicho vector de $\mathbb{R}^m$ es precisamente $L_{F,p}(h)$. (Ver Figura 7.4)."

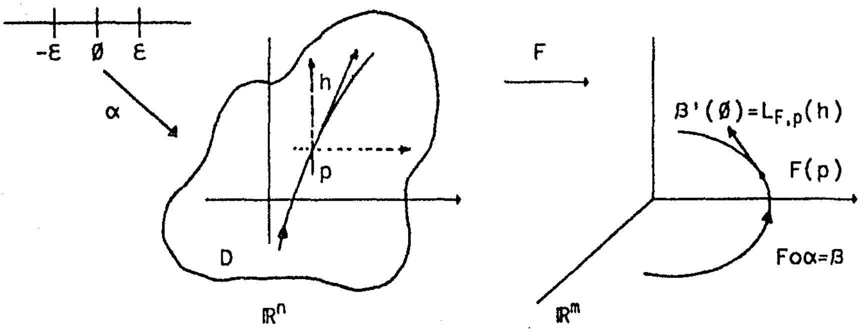


Figura 7.4

Lo interesante del resultado anterior, es que permite interpretar la diferencial de una función en un punto en términos de un objeto que resulta muy familiar para los estudiantes: el vector tangente a una curva.

BIBLIOGRAFIA

Abreu, J. L., et al. Cálculo Diferencial e Integral. (volúmenes 1, 3 y 5). Limusa. México. 1983.

Amazigo, J., Rubenfeld, L. Advanced Calculus. Wiley. New York. 1980.

Anton, H. Introducción al Algebra Lineal. Limusa. México. 1984.

Apostol, T. Mathematical Analysis. Addison - Wesley. Massachusetts. 1958.

Bartle, R. Introducción al Análisis Matemático. Limusa. México. 1982.

Baxandall, P.R., Liebeck, H. Differential Vector Calculus. Longman. New York. 1981.

Bers, L., Karal, F. Cálculo. Interamericana. 2a edición. México. 1978.

Brand, L. Cálculo Avanzado. CECSA. 2a edición. México. 1961.

Buck, R.C. Advanced Calculus. McGraw Hill. 3rd edition. New York. 1978.

Courant, R. Differential and Integral Calculus. (Volume I). Interscience. 2nd. edition. Glasgow. 1961.

Courant, R., John, F. Introduction to Calculus and Analysis. (2 volúmenes). Wiley. New York. 1974.

Craven, B.D. Functions of Several Variables. Chapman and Hall. New York. 1981.

Chillingworth, D.R. Differential Topology with a view to Applications. Pitman. London. 1977.

Dieudonné, J. Fundamentos del Análisis Moderno. Reverté. Barcelona. 1979.

Do Carmo, M. Differential Geometry of Curves and Surfaces. Prentice - Hall. New Jersey. 1976.

Fleming, W. Functions of Several Variables. Addison-Wesley. Reading, Mass. 1965.

Fraleigh, J. Cálculo con Geometría Analítica. FEI. México. 1984.

Fulks, W. Advanced Calculus. John Wiley & Sons. New York. 1961.

Granville, A., et al. Cálculo Diferencial e Integral. UTEHA. México. 1962.

Haaser, N., et al. A Course in Mathematical Analysis. Blaisdell Publishing Company. New York. 1964.

Hirsch, M., Smale, S. Differential Equations, Dynamical Systems and Linear Algebra. Academic Press. New York. 1974.

Hofmann, J. Historia de la Matemática. (Tomo II). UTEHA. México. 1961.

Kaplan, W. Cálculo Avanzado. CECSA. México. 1983.

Kline, M. Calculus: An Intuitive And Physical Approach. Wiley. 2nd edition. U.S.A. 1977.

Kolmogórov, A., Fomín, S. Elementos de la Teoría de Funciones y del Análisis Funcional. Mir. 2a edición. Moscú. 1975.

Lang, S. A Second Course in Calculus. Addison - Wesley. 2nd edition. Reading, Mass. 1964.

Lang, S. Calculus of Several Variables. Addison - Wesley. Reading, Mass. 1973.

Larson, R., Hostetler, R. Cálculo y Geometría Analítica. McGraw - Hill. 2a edición. México. 1986.

Lipschutz, M. Theory and Problems of Differential Geometry. McGraw - Hill. New York. 1969.

Loomis, L., Sternberg, S. Advanced Calculus. Addison-Wesley. Massachusetts. 1968.

Marsden, J., Tromba, A. Cálculo Vectorial. FEI. México. 1986.

McAloon, K., Tromba, A. Cálculo de una Variable. PCSA. México. 1975.

Milne, R.D. Applied Functional Analysis (An Introductory Treatment). Pitman. London. 1980.

Osserman, R. Two - Dimensional Calculus.

Phillips, H.B. Elementos de Cálculo Infinitesimal. UTEHA. México. 1947.

Saumells, R. Fundamentos de Matemática y de Física. Rialp. Madrid. 1961.

Sokolnikoff, I. Advanced Calculus. McGraw - Hill. New York. 1939.

Sokolnikoff, I., Redheffer, R. Mathematics of Physics and Modern Engineering. McGraw - Hill. New York. 1958.

Spivak, M. Cálculo Infinitesimal. (volúmenes 1 y 2). Reverté. Barcelona. 1981.

Spivak; M. Cálculo en Variedades. Reverté. Barcelona. 1982.

Swokowski, E. Calculus with Analytic Geometry. Prindle, Weber & Schmidt. 2nd edition. Boston, Mass. 1979.

Trench, W., Kolman, B. Multivariable Calculus with Linear Algebra and Series. Academic Press. New York. 1972.

Woods, F., Bailey, F. Geometría Analítica y Cálculo Infinitesimal. UTEHA. México. 1952.