UNIVERSIDAD DE SONORA
Departamento de Matemáticas

Teoremas Fundamentales del Cálculo Diferencial

TESIS

Que para obtener el Título de

LICENCIADO EN MATEMATICAS

Presenta

José de Jesús Ayala

Hermosillo, Sonora Diciembre de 1990

A MI MADRE

Doy mis mas sinceros agraecimientos al Dr. Ruben Flores Espinoza quien ha sido el autor de la idea y colaboraor de este material. Asi mismo agradezco a los profesores: Eduardo Tellechea armenta, Jacobo Nuñez Urias y Pedro Flores Perez de quienes obtuve valiosas observaciones y sugerencias para este trabajo..

CONTENIDO

INTRODUCCION.

Este trabajo tiene, como propósito principal, dar una presentación rigurosa y unificada de los teoremas considerados como más relevantes del cálculo diferencial en espacios finito dimensionales. Estos son, *el teorema de la función inversa, el teorema de la función implícita, el teorema del rango* y el no menos importante *teorema de la variedad invariante*. Tales resultados son pilares en el análisis no lineal, cuya escencia radica en que describen el comportamiento local de las *funciones continuamente diferenciables*, (funciones de clase $\mathcal{C}^1$), sustentado en el hecho de que en vecindades pequeñas de cada punto, admiten una muy buena aproximación lineal determinada por su diferencial en el punto. Mas aún se puede decir en términos generales, que las funciones diferenciables heredan las principales propiedades cualitativas de la transformación lineal que las aproxima alrededor del punto. Por esta razón algunos autores expresan esto diciendo que *"las funciones continuamente diferenciables se comportan localmente como sus derivadas"*.

En este trabajo se presentan, demostraciones completas y rigurosas tanto del teorema de la función inversa que nos permite asegurar invertibilidad de la función alrededor del punto a partir de la invertibiliad de su diferencial en el punto; del teorema de la función implícita que nos permite caracterizar localmente al conjunto de solúciones de una ecuación no lineal como la gráfica de una función cuando el sistema lineal asociado por la diferencial de n ecuaciones con n+m variables es tal que su rango es igual al número de ecuaciones. Del teorema del rango que nos permite representar localmente a una función con el menor número posible de variables a partir del conocimiento el rango de la diferencial y finalmente del teorema de la variedad invariante que asegura la existencia de una variedad invariantes bajo la función a partir de la existencia de un subespacio invariante para la diferencial en un punto

además, en términos generales, las demostraciones que se presentan, están fuertemente basadas en las técnicas y métodos iterativos que para encontrar puntos fijos se utilizan en el análisis funcional a la manera del teorema de contracción de Banach. En este último sentido, este trabajo dá una presentación unificada de los teoremas antes señalados que para los cuales, los esquemas de aproximación son, como mencionamos, los que se utilizan en el teorema de contracción de Banach y mas aún, esto nos permitirá, bajo un cambio apropiado, generalizar específicamente el teorema de la función inversa y el teorema de la función implícita a espacios infinito dimensionales o bien a espacios de Banach arbitrarios.

En realidad, el estudio de funciones no linales mediante transformaciones lineales resulta ser todo un éxito del cálculo diferencial. Es en este sentido por el cuál hemos intentado, como otras de nuestras metas, marcar el puente de lo lineal a lo no lineal motivando para cada uno de los teoremas antes señalados con un breve repaso de la teoría básica de los *sistemas de ecuaciones lineales*, *cambios de coordenadas* y *subespacios invariantes* de una transformación lineal respectivamente.

El trabajo se ha dividido en siete capítulos. en el **capitulo cero** (conceptos y resultados básicos) se definen los espacios de Banach, medio en el cual, los dos primeros capítulos se desarrolla, se ilustran algunos ejemplos que serán importantes para nuestro estudio como lo es el espacio de las funciones continuas $\mathscr{C}(D;\mathbb{R}^n)$ con D compacto y el espacio de las transformaciones lineales y continuas $\mathscr{L}(\mathbb{X};\mathbb{Y})$ donde $\mathbb{X}$, $\mathbb{Y}$ también son espacios de banach.

La teoría del cálculo diferencial en los espacios de banach representa la generalización natural de la ya conocida en los $\mathbb{R}^n$ y debido a la analogía con esta, solamente nos ocuparemos, en este mismo capítulo, en dar una breve exposición de las cuestiones mas importantes que utilizaremos como el concepto de *diferenciabilidad diferenciabilidad parcial*, y de algunos resultados conectados con estos así como la versión generalizada del *teorema del valor medio*. En el **capitulo uno** aparecen el *teorema de contracción de Banach*, válido en cualquier espacio métrico completo, y *el teorema de dependencia continua y diferenciable de puntos fijos*. como hicimos notar, el hecho de que aquí se hallan incluídos es que en base a ellos, daremos una prueba, no clásica, del teorema de la función inversa que a diferencia de las pruebas tradicionales, la diferenciabilidad de la función inversa se sigue directamente del teorema de dependencia diferenciable de puntos fijos sin recurrir formalmente a la definición. En tanto el teorema de la función implícita será dado como una aplicación al de la inversa y ambos aparecen en el **capitulo dos**.

En el **capitulo tres** encontramos uno de los teoremas fundamentales en el estudio de las llamadas variedades, como lo es *el teorema del rango*. El enfoque de este resultado se ha dado en términos del concepto de equivalencia entre funciones de clase $\mathscr{C}^1$, $\mathscr{C}^1(\mathbb{R}^n,\mathbb{R}^m)$, mediante el cuál queda dividido este espacio en clases de equivalencia. Como es sabido, esto nos permite hablar, no de una función en sí, sino de funciones representativas de toda una clase, las cuales, quedan localmente caracterizadas, bajo este teorema, por expresiones particularmente simples. Además la versión presentada aquí, tendrá su importancia en el **capitulo cuatro** en el cual se mostrará un importante resultado de equivalencia, es decir, se satisface que:

TEOREMA DE LA FUNCION INVERSA $\Longleftrightarrow$ *TEOREMA DEL RANGO* $\Longleftrightarrow$ *TEOREMA DE LA FUNCION IMPLICITA*.

Entre las múltiples aplicaciones de estos últimos, introduciremos el concepto de *dependencia funcional* como contraparte de la dependencia lineal que tiene sentido en espacios vectoriales y que aparece en el **capitulo cinco**. El problema a tratar aquí es el de determinar cuando en un cierto conjunto de funcionales es posible expresar al menos una de ellas en términos de las demás restantes.

Finalmente en el **capitulo seis** se verá un resultado un poco más complicado debido al tipo de demostración, pero interesante desde el punto de vista matemático, se llama el *teorema de la variedad invariante* y su utilidad alcanza niveles importantes dentro de las ecuaciones diferenciales avanzadas. Sin precisar, este teorema establece, bajo ciertas condiciones, que si un cierto mapeo T de clase $\mathscr{C}^1$, definido en un mismo espacio, es tal que su diferencial DT_o parte dicho espacio en dos subespacios invariantes entonces existen variedades de clase $\mathscr{C}^1$, representadas por la gráfica de alguna función, las cuales quedan localmente invariantes bajo T en el sentido de que manda puntos de la variedad en puntos de la misma, es decir, se preserva bajo pequeñas perturbaciónes la existencia de conjuntos invariantes.

CAPITULO O

CONCEPTOS Y RESULTADOS BASICOS

Este capítulo, que consta de tres secciones, lo hemos dedicado a dar los conceptos y resultados básicos con los cuales se sustentará
este trabajo. En la sección uno se definen los espacios de Banach y los espacios producto entre ellos; se dá por ejemplo el espacio de las funciones continuas, $\mathscr{C}(D\subset\mathbb{R}^n,\mathbb{R}^m)$, así como también una consecuencia importante de los espacios finito dimensionales. Se enunciará también el *teorema de Ascoli-Arzelá*, un resultado fuerte del análisis que será utilizado hasta el capitulo seis. En la sección dos se prersentan las transformaciones lineales y la norma suprema definida para ellas que hace del espacio de todos las transformaciones lineales y continuas, $\mathscr{L}(\mathscr{X},\mathscr{Y})$,un espacio de Banach. Por último en la sección tres encontramos la herramienta necesaria del cálculo que ocuparemos de los espacios de Banach que, como se verá, es una versión generalizada del cálculo estudiado en los euclidianos $\mathbb{R}^n$.

OBSERVACION. Los teoremas que aquí hemos incluído se darán sin demostración, por lo que se les adjuntará, en letra pequeña, la referencia contenida en la bibliografía misma que aparece hasta el final.

1. ESPACIOS DE BANACH

DEF.0.1.1 Un espacio vectorial $\mathscr{X}$, sobre un campo K, ($\mathbb{R}$ ó $\mathbb{C}$), es un espacio vectorial normado, (e.v.n), si a cada $x \in \mathscr{X}$, le corresponde a un número real, $\|x\|_X$, llamada la norma de x, el cual satiface las siguientes propiedades:

i) $\|x\|_X > 0$ si $x \neq 0$ y $\|0\|_X = 0$

ii) $\|x+y\|_X \leq \|x\|_X + \|y\|_X$

iii) $\|ax\|_X = |a|\,\|x\|_X$ $\forall$ $a \in K$, $x \in \mathscr{X}$.

NOTA. Cuando no exista peligro de confusión se escribirá $\|\ \|$ en vez de $\|\ \|_X$.

OBS.1 Todo espacio vectorial normado $\mathscr{X}$, es un espacio métrico, donde la métrica para este espacio es inducida por la norma y que viene dada por

$$d(x,y) = \|x - y\|,\ x,\ y \in \mathscr{X}$$

DEF.0.1.2 Sea $\mathscr{X}$ un e.v.n. y sea $\{x_n\}$ una sucesión en este espacio. Decimos que $\{x_n\}$ converge a $x \in \mathscr{X}$ si $\lim_{n\to\infty} \|x_n - x\| = 0$ y escribimos $\lim_{n\to\infty} x_n = x$.

OBS.2. Alternativamente, lo anterir significa que para cada $\varepsilon > 0$, $\exists$ $N(\varepsilon)$ tal que $\|x_n - x\| < \varepsilon$ si $n > N$.

DEF.0.1.3. Sea $\mathscr{X}$ un e.v.n.,una sucesión $\{x_n\}$ en $\mathscr{X}$ se llama *sucesión de cauchy* si para cada $\varepsilon > 0$, $\exists$ $N(\varepsilon) > 0$ tal que

$$\|x_n - x_m\| < \varepsilon \text{ si } n,\ m \geq N(\varepsilon).$$

Sabemos que toda sucesión convergente es una sucesión de cauchy. Si la recíproca tambien es cierta entonces:

DEF.0.1.4. Sea $\mathcal{X}$ un e.v.n.,se dice que $\mathcal{X}$ es completo, con respecto a la métrica inducida por la norma, si toda sucesión de cauchy en $\mathcal{X}$ converge a un elemento del mismo espacio.

De esta última se desprende la siguiente

DEF.0.1.5. Sea $\mathcal{X}$ un e.v.n. y completo, entonces a $\mathcal{X}$ se le llama *espacio de Banach*.

OBS.3.Todo espacio de Banach, como espacio métrico, posee una estructura topológica la cual queda determinada precisamente por métrica inducida por la norma.

EJEMPLOS

- El espacio $\mathbb{R}^n$ con cualquiera de las siguientes normas es un espacio de Banach.

$$\|x\| = \left(\sum_{i=1}^{n} | x_i |^2 \right)^{1/2}, \quad \|x\| = \sum_{i=1}^{n} | x_i |, \quad \|x\| = \sup_{i=1,\ldots,n} | x_i |.$$

Sea D un subconjunto compacto de $\mathbb{R}^n$. El espacio vectorial de las funciones continuas que van de D a $\mathbb{R}^m$, $\mathcal{C}(D, \mathbb{R}^m)$, es un espacio, de banach con la siguiente norma:

$$\|\Phi\| = \sup_{x\in D} \|\Phi(x)\|.$$

DEF.0.1.6. Sea $\mathcal{X}$ un e.v.n. y $\|\circ\|$, $\|\circ\|_1$ dos normas para $\mathcal{X}$. Se dice que $\|\circ\|$ y $\|\circ\|_1$ son equivalentes si existen numeros m > 0, M > 0 tales se satisface la desigualdad:

$$m\|\circ\|_1 \le \|\circ\| \le M\|\circ\|_1$$

OBS.4. Cuando dos normas son equivalentes, en el sentido de la definición anterior, la convergencia de una sucesión en una norma implica la convergencia de la suceción en la otra.

El siguiente teorema caracteriza a las normas en los espacios finito dimensionales.

<u>TEOREMA</u> 0.1.1. Sea $\mathcal{X}$ un e.v.n. finito dimensional y $\|\circ\|$, $\|\circ\|_1$ dos normas en $\mathcal{X}$. Entonces $\|\circ\|$ y $\|\circ\|_1$ son equivalentes.(véase,sección 2.4 del libro 9 de la bibligrafía)

OBS.5. Cuando dos normas son equivalentes, estas generan la misma topología, por lo cual, es usual trabajar los espacios finito dimensionales, independientemente de la norma que se trate.

Como tendremos la oportunidad de considerar funciones definidas en espacios producto, por ejemplo en el teorema de la función implícita, es conveniente hacer la siguiente definición.

DEF.0.1.7. Sean $\mathcal{X}$, $\mathcal{Y}$ espacios de banach. Definimos el espacio producto como:

$\mathcal{X} \times \mathcal{Y} = \{ (x_1, x_2) / \ x_1 \in \mathcal{X}, x_2 \in \mathcal{Y} \}$

así, $\mathcal{X} \times \mathcal{Y}$ está dotado de una estructura de espacio vectorial con las operaciones dadas por las siguientes fórmulas:

$\forall \ (x_1,x_2), \ (y_1,y_2) \in \mathcal{X} \times \mathcal{Y}$, λ escalar, $\mathbb{R}$ ó $\mathbb{C}$,

i) $(x_1,x_2) + (y_1,y_2) = (x_1 + y_1, \ x_2 + y_2)$

ii) $\lambda(x_1,x_2) = (\lambda x_1,\lambda x_2)$

en particular $(x_1, \ x_2) = (x_1, \ o) + (o, \ x_2)$.

Ahora si sobre el e.v $\mathcal{X} \times \mathcal{Y}$ definimos $\|(x_1,x_2)\| = \|x_1\| + \|x_2\|$ se puede comprobar que efectivamente es una norma y que según ella $\mathcal{X} \times \mathcal{Y}$ es completo pues $\mathcal{X}$, $\mathcal{Y}$ lo son.

DEF.0.1.8. Sean $\mathcal{X}$, $\mathcal{Y}$ espacios de Banach y consideremos una aplicación $f:\mathcal{X} \to \mathcal{Y}$. Se dice que f es continua en $x_o \in \mathcal{X}$ si para h en $\mathcal{X}$ se cumple que $\lim\limits_{h\to o} \|f(x_o + h) - f(x_o)\| = 0$; esto de igual manera significa que dado $\varepsilon > 0$, $\exists\ \delta(\varepsilon, x_o)$ tal que se satisface que $\|f(x_o + h) - f(x_o)\|_{\mathcal{Y}} < \varepsilon$ si $\|h\|_{\mathcal{X}} < \delta$.

DEF.0.1.9. Se dice que una función $f \in \mathcal{C}(D;\mathbb{R}^m)$ es uniformemente continua sobre D si dado $\varepsilon > 0$ existe $\delta(\varepsilon) > 0$ tal que

$$\|f(x + h) - f(x)\| < \varepsilon \quad \text{si } \|h\| < \delta,\ \forall\ x,\ x + h \in D.$$

DEF.0.1.10. Se dice que una sucesión $\{f_n\} \in \mathcal{C}(D;\mathbb{R}^m)$ es equicontinua sobre D, si dado $\varepsilon > 0$ $\exists\ \delta(\varepsilon)$ tal que

$\|f_n(x + h) - f_n(x)\| < \varepsilon$ si $\|h\| < \delta$, $\forall n$ y $\forall\ x,\ x + h \in D$.

DEF.0.1.11. Una sucesión $\{f_n\}$ en $\mathcal{C}(D;\mathbb{R}^m)$ se dice que es uniformemente convergente sobre D, si existe una función $f{:}D \to \mathbb{R}^m$ tal que dado $\varepsilon > 0$, $\exists\ N(\varepsilon) > 0$ de manera que

$$\|f_n(x) - f(x)\| < \varepsilon\ \forall\ n \text{ y } \forall\ x \in D.$$

DEF.0.1.12. Una sucesión $\{f_n\}$ en $\mathcal{C}(D;\mathbb{R}^m)$ se dice uniformemente acotada si existe $M > 0$ tal que $\|f_n(x)\| \le M\ \forall\ n$ y $\forall\ x \in D$.

Las tres últimas definiciones vienen conectadas fuertemente por el siguiente teorema

<u>TEOREMA</u> 0.1.2. (ASCOLI - ARZELA)

Toda sucesión úniformemente acotada y equicontinua en $\mathcal{C}(D;\mathbb{R}^m)$, contiene una subsucesión uniformemente Convergente sobre D.(véase el capítulo 7 del libro 12)

2. TRANSFORMACIONES LINEALES

En esta sección $\mathscr{X}$, $\mathscr{Y}$ son espacios vectoriales sobre un mismo campo K.

*DEF.*0.2.1. Sean $\mathscr{X}$, $\mathscr{Y}$ espacios de Banach. Una transformación $T:\mathscr{X} \to \mathscr{Y}$ con la propiedad de que $T(ax + by) = aTx + bTy \quad \forall\ x, y \in \mathscr{X}$ y escalares a, b, es llamada una transformación lineal.

DEF. 0.2.2. Una transformacion lineal $T:\mathscr{X} \to \mathscr{Y}$ se dice acotada, si existe una constante λ tal que $\|Tx\|_{\mathscr{Y}} \leq \lambda\|x\|\ \forall\ x \in \mathscr{X}$.

Las transformaciones lineales tienen una interesante propiedad en términos de su continuidad y acotamiento dada por lo siguiente.

<u>TEOREMA</u> 0.2.1. Sean $\mathscr{X}$, $\mathscr{Y}$ espacios de Banach. una transformación lineal $T:\mathscr{X} \to \mathscr{Y}$ es acotada si y solo si es continua.(véase sección 2.7 del libro 9).

OBS.6. En los espacios finito dimensionales las transformaciones lineales son automáticamente contínuas. Además cada transformación lineal $T:\mathbb{R}^n \to \mathbb{R}^m$, puede ser representado por una matriz A de m×n.

*DEF.*0.2.3. Sean $\mathscr{X}$, $\mathscr{Y}$ espacios de banach y $T:\mathscr{X} \to \mathscr{Y}$ una transformación lineal continua. Definimos la norma $\|T\|$ como:

$$\|T\| = \operatorname{Sup}\left\{\|Tx\|_{\mathscr{Y}}:\ \|x\|_{\mathscr{X}} \leq 1\right\}$$

OBS.7. De aquí se desprende la siguiente desigualdad

$$\|Tx\|_{\mathscr{Y}} \leq \|T\|\,\|x\|_{\mathscr{X}}\ .$$

TEOREMA 0.2.2. Sean $\mathscr{X}$, $\mathscr{Y}$ espacios de Banach y sea $\mathscr{L}(\mathscr{X},\mathscr{Y})$ el espacio de todas las transformaciones lineales y continuas que que van de $\mathscr{X}$ en $\mathscr{Y}$. Entonces $\mathscr{L}(\mathscr{X},\mathscr{Y})$ es un espacio de Banach con respecto a la norma definida anteriormente.(véase sección 2 del libro 9).

La estructura algebraica y topológica que poseen los espacios de Banach, son su principal objeto de estudio y es en base a estos con los cuales se formula y demuestra el teorema de la función inversa e implícita en el caso particular de los $\mathbb{R}^n$. De esta forma la generalización de estos resultados a tales espacios queda automáticamente resuelta mediante la identificación entre ellos por vía del concepto de *isomorfismo*. Es de esta manera por el cual podemos actuar en un espacio y otro tratándolos como si fueran el mismo sin recurrir a la dimensión de estos o a la naturaleza de los elementos que los constituyen. De ahí la importancia de la siguiente definición.

DEF.0.2.4. Sean $\mathscr{X}$, $\mathscr{Y}$ e.v.n. Se dice que $\mathscr{X}$, $\mathscr{Y}$ son isomorfos si existe una transformación lineal $T:\mathscr{X} \to \mathscr{Y}$ tal que:

i) T es continua

ii) T es invertible con inversa lineal continua.

Cuando esto es así, a T se le llama *isomorfismo* de $\mathscr{X}$ en $\mathscr{Y}$.

OBS.8. En otras palabras para que $\mathscr{X}$ y $\mathscr{Y}$ sean isomorfos es necesario y suficiente que exista un *homeomorfismo* lineal $T:\mathscr{X} \to \mathscr{Y}$.

Como un ejemplo sencillo, tenemos que los espacios n-dimensionales son isomorfos a $\mathbb{R}^n$, donde el isomorfismo natural esta dado por x -----→ vector coordenado correspondiente a una base elegida.

3. APLICACIONES DIFERENCIABLES

DEF.0.3.1. Sean $\mathscr{X}$, $\mathscr{Y}$ espacios de Banach y D un abierto en $\mathscr{X}$. Una función $f:D \to \mathscr{Y}$ se dice diferenciable en un punto $x \in D$, si existe una transformación lineal contínua $T:\mathscr{X} \to \mathscr{Y}$ tal que $\forall$ $h \in \mathscr{X}$ con $x + h \in D$ se cumple que:

$$(*) \quad \lim_{h \to 0} \frac{\|f(x + h) - f(x) - Th\|_{\mathscr{Y}}}{\|h\|_{\mathscr{X}}} = 0$$

Esta relación puede escribirse en la siguiente forma. Si definimos $f(x + h) - f(x) - Th = O(h)$, entonces $O(h)$ representa el error en la aproximación de $f(x + h) - f(x)$ por medio de la transformación lineal Th y claramente se tiene que

$$(**) \quad \lim_{h \to 0} \frac{\|O(h)\|}{\|h\|} = 0$$

así equivalentemente se dice que f es diferenciable en un punto $x \in D$, si existe una transformación lineal continua $T:\mathscr{X} \to \mathscr{Y}$ tal que se satisfaga la igualdad:

$$(***) \quad f(x + h) - f(x) = Th + O(h)$$

en donde $O(h)$ debe ser de orden mas pequeño que h, es decir, que satisfaga (**).

Nótese que (***) lo podemos asociar con la fórmula de Taylor de primer orden que nos proporcina la aproximación de la diferencia $f(x + h) - f(x)$ por medio de la transformación lineal Th y en donde el error en dicha aproximación esta representado por $O(h)$. En este caso a T se le llama la diferencial de f en x y se denota como $T = Df(x) = f'(x)$

OBS.8. Si f es diferenciable en todo punto $x \in D$, se dice que f es diferenciable en D.

Algunos resultados básicos de calculo diferencial como la regla de la cadena, entre otros, son tambien generalizables, de forma análoga que en los $\mathbb{R}^n$, a espacios de Banach arbitrarios, por lo cual, nos permitimos omitirlos en esta sección.

Ahora, con la definición de diferenciabilidad que se ha dado, pasaremos a definir el concepto de diferenciabilidad parcial considerando el caso en que D es conjuntoun abierto de un producto de espacios de Banach.

DEF.0.3.2. Supongamos que $\mathscr{X} = \mathscr{X}_1 \times \mathscr{X}_2$ donde $\mathscr{X}_1$, $\mathscr{X}_2$ son espacios de Banach y sea D un abierto en $\mathscr{X}$. consideremos una función $f: D \to \mathscr{Y}$ y elijamos un punto (x_o, y_o) en D con $x_o \in \mathscr{X}_1$, $y_o \in \mathscr{X}_2$. Ahora sea el mapeo $P: \mathscr{X}_1 \to \mathscr{Y}$ dado por $P(x) = f(x_o + x, y_o)$. Si P es diferenciable en el punto $x = 0$ la derivada $DP(0)$ es llamada la derivada parcial de f con respecto a x en el punto (x_o, y_o) y se denota como

$$DP(0) = \frac{\partial f}{\partial x}(x_o, y_o).$$

Por definición $DP(0)$ es una transformación lineal continua de $\mathscr{X}_1$ en $\mathscr{Y}$ que satisface que

$\displaystyle \lim_{h \to o} \frac{\|P(0 + h) - P(0) - DP(0)h\|}{\|h\|} = 0$. Esto es equivalente a que

$$\lim_{h \to o} \frac{\left\| f(x_o + h, y_o) - f(x_o, y_o) - \frac{\partial f}{\partial x}(x_o, y_o)h \right\|}{\|h\|} = 0.$$

De manera semejante se define $\frac{\partial f}{\partial y}(x_o, y_o)$ a la transformación lineal continua de $\mathscr{X}_2$ en $\mathscr{Y}$ que satisface que $\rightarrow$

$$\lim_{k \to 0} \frac{\left\| f(x_o, y_o + k) - f(x_o, y_o) - \frac{\partial f}{\partial y}(x_o, y_o)k \right\|}{\|k\|} = 0$$

Informalmente hablando las derivadas parciales son transformaciones lineales que aproximan las variaciones de f en una dirección del espacio.

OBS.9. de manera analoga el concepto de derivada parcial se extiende al caso en que $\mathbb{X} = \mathbb{X}_1 \times \mathbb{X}_2 \times \ldots \times \mathbb{X}_n$.

La diferenciabilidad y la diferenciablidad parcial estan conectadas por el siguiente teorema:

<u>TEOREMA</u> 0.3.1. Sean $\mathbb{X}_1$, $\mathbb{X}_2, \ldots,$ $\mathbb{X}_n$, $\mathbb{Y}$ espacios de Banach y pongamos $\mathbb{X} = \mathbb{X}_1 \times \mathbb{X}_2 \times \ldots \times \mathbb{X}_n$. Sea D un abierto en $\mathbb{X}$. Si $f: D \to \mathbb{Y}$ es diferenciable en un punto $a = (a_1, a_2, \ldots, a_n) \in D$ entonces las derivadas parciales $\frac{\partial f}{\partial x_i}(a_1, a_2, \ldots, a_n)$, $i=1,2,\ldots n$ existen y

$f'(a) \circ (h_1, h_2, \ldots, h_n) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a) h_i$ donde $h_i \in \mathbb{X}$. (véase sección 2.6 del libro 5).

OBS.10. En nuestra notación el símbolo $\mathbb{R}^k$ significa que $\mathbb{R}^k = \mathbb{R} \times \mathbb{R} \times \ldots \times \mathbb{R}$ k-veces. Si además $f: \mathbb{R}^n \to \mathbb{R}^m$ es diferenciable en un punto $a = (a_1, a_2, \ldots, a_n)$ entonces la derivada $f'(a)$ esta representada por una matriz de m×n dada por:

$$\left(\frac{\partial f_i}{\partial x_j}(a) \right) \quad i=1,2,\ldots,m \quad j=1,2,\ldots n$$

la cual, es llamada la matriz jacobiana de f con respecto al vector x en elpunto a.

DEF.0.3.3. Se dice que una función $f:D \to \mathcal{Y}$ es continuamente diferenciable ó que también es de clase $\mathcal{C}^1(D)$ si cumple con lo siguiente:

i) f es diferenciable en D

ii) f' es continua en todo punto $x \in D$.

OBS.11. Para cada $x \in D$, $f'(x)$ es un elemento de $\mathcal{L}(\mathcal{X},\mathcal{Y})$ y ii) significa que la aplicación $f':D \to \mathcal{L}(\mathcal{X},\mathcal{Y})$ es continua.

TEOREMA 0.3.2. Sean $\mathcal{X}_1$, $\mathcal{X}_2,\ldots,\mathcal{X}_n$, $\mathcal{Y}$ espacios de Banach. Pongamos $\mathcal{X} = \mathcal{X}_1 \times \mathcal{X}_2 \times \ldots \times \mathcal{X}_n$ y sea D un abierto en $\mathcal{X}$. Entonces una función $f:D \to \mathcal{Y}$ es de clase $\mathcal{C}^1(D)$ si y sólo si las derivadas parciales $\frac{\partial f}{\partial x_i}$ $i = 1,2,\ldots n$, existen y son continuas en D. (secc. 2.5, libro 5)

Para terminar con este capítulo, enunciaremos un teorema que se verifica en los espacio normados y que para efectos prácticos es tan útil como el teorema del valor medio. Dicho resultado es llamado con el mismo nombre y está dado por el siguiente.

TEOREMA 0.3.3. Sean $\mathcal{X}$, $\mathcal{Y}$ espacios de Banach, D un abierto en $\mathcal{X}$. Si $f:D\to \mathcal{Y}$ es diferenciable en D y si el segmento que une a los puntos x_o, $x_o+ h$ definido por $[x_o,x_o+ h] = \{c \in \mathcal{X} \ / \ c = x_o+th,\ 0 \le t \le 1\}$ esta en D, entonces:

$$\|f(x_o + h) - f(x_o)\|_{\mathcal{Y}} \le \|h\|_{\mathcal{X}} \sup_{c\in[x_o,x_o+h]} \|f'(c)\|_{\mathcal{L}(\mathcal{X},\mathcal{Y})}$$

(véase sección 3.1 del libro 5)

CAPITULO 1

DEPENDENCIA CONTINUA Y DIFERENCIABLE DE PUNTOS FIJOS

El objetivo de este capitulo es dar básicamente los resultados claves con los cuales nos apoyaremos para dar una demostración rigurosa del Teorema de la función inversa. El primero de ellos determina su existencia y es un teorema de punto fijo llamado, el Teorema de Contracción de Banach cuya validez se extiende a los espacios métricos completos como son, en particular los espacios de Banach, medio en el cual aquí nos encontraremos. Los siguientes resultados ,que también son sobre puntos fijos, se presentan cuando se tiene una familia de contracciones uniforme definida en un mismo espacio completo. informamente hablando estos resultados expresan que si las variaciones entre los elementos de dicha familia se comportan continua y diferenciablemente, entonces los puntos fijos correspondientes se comportan de igual forma; respectivamente estos son el teorema de dependencia continua y el teorema de dependencia diferenciable de puntos fijos. Este último nos llevará a concluír que la función inversa es localmente de clase $\mathscr{C}^1$.

1. EL TEOREMA DE CONTRACCION DE BANACH

*DEF.*1.1.1. Sea $\mathcal{X}$ un conjunto cualquiera y $T:\mathcal{X} \to \mathcal{X}$ una aplicación arbitraria. Se dice que $x \in \mathcal{X}$ es un punto fijo para T si $Tx = x$.

*DEF.*1.1.2. Sean $\mathcal{X}$, $\mathcal{Y}$ espacios de Banach y F un subconjunto de $\mathcal{X}$. Se dice que una aplicación $T:F \to \mathcal{Y}$ es una contracción sobre F si existe una constante λ, con $0 \leq \lambda < 1$, tal que:

$$\|Tx - Ty\| \leq \lambda \|x - y\| \quad \forall\ x,y \in F.$$

*OBS.*1. El número λ es llamado constante de contracción para T sobre F.

Las definiciones anteriores, juntas, dan por resultado el teorema central de esta sección que a continuación se enuncia.

<u>TEOREMA</u> 1.1.1. (DE CONTRACCION DE BANACH). Sea $\mathcal{X}$ un espacio de Banach y F un subconjunto cerrado del mismo. Si $T:F \to F$ es una contracción, entonces T tiene un único punto fijo en F.

*OBS.*2. Equivalentemente el teorema afirma que la ecuación $Tx = x$ posee solución única.

DEMOSTRACION: Como tradicionalmente se hace, la prueba se efectuará por aproximaciones sucesivas construyendo una sucesión $\{x_n\}$ en F y mostrando que tal sucesión es de cauchy. Luego con F completo, se deducirá que tal sucesión tiene como límite un punto en F, el cual será precisamente el punto fijo para T. Finalmente se verá que T no tiene otros puntos fijos.

Se elige un punto x_0 arbitrario y se define la sucesión dada por

$$x_1 = Tx_0, \quad x_2 = Tx_1, \ldots, x_n = Tx_{n-1}.$$

Para probar que esta sucesión es de cauchy se hace lo siguiente:

Sea m, n con $m \geq n$, entonces se tiene que

$\|x_n - x_m\| = \|Tx_{n-1} - Tx_{m-1}\| \leq \lambda\|x_{n-1} - x_{m-1}\| \leq \lambda\|Tx_{n-2} - Tx_{m-2}\| \leq$

$\lambda^2\|x_{n-2} - x_{m-2}\| \leq \lambda^2\|Tx_{n-3} - Tx_{m-3}\| \leq \ldots \leq \lambda^n\|x_o - x_{m-n}\|$.

nótese que aquí hemos usado repetidamente el hecho de que T es contracción. De esta manera obtenemos la desigualdad

$$(*) \qquad \|x_n - x_m\| \leq \lambda^n\|x_o - x_{m-n}\|.$$

Si en el lado derecho de esta última se aplica la propiedad triangular para los puntos $x_o, x_1, \ldots x_{m-n}$ se obtiene

$\lambda^n\|x_o - x_{m-n}\| \leq \lambda^n\left(\|x_o - x_1\| + \|x_1 - x_2\| + \ldots + \|x_{m-n-1} - x_{m-n}\|\right)$ y

por (*) tenemos

$\lambda^n\left(\|x_o - x_1\| + \lambda\|x_o - x_1\| + \lambda^2\|x_o - x_1\| + \ldots + \lambda^{m-n-1}\|x_o - x_1\|\right) =$

$\lambda^n\|x_o - x_1\|\left(1 + \lambda + \lambda^2 + \ldots + \lambda^{m-n}\right) = \lambda^n\|x_o - x_1\|\left(\frac{1 - \lambda^{m-n}}{1 - \lambda}\right)$.Como

$0 \leq \lambda < 1$ se tiene que $1 - \lambda^{m-n} < 1$ y

$\lambda^n\|x_o - x_{m-n}\| \leq \lambda^n\|x_o - x_1\|\left(\frac{1}{1 - \lambda}\right)$. Por tanto se llega a que

$$(**) \qquad \|x_n - x_m\| \leq \frac{\lambda^n}{1 - \lambda}\|x_o - x_1\| \qquad m \geq n.$$

Ahora como λ es positivo y menor que uno, entonces el término del lado derecho se puede hacer tan pequeño como se quiera eligiendo n suficientemente grande. Mas precisamente esto significa que $\forall\ \varepsilon > 0$, $\exists$ N tal que

$\|x_n - x_m\| \leq \frac{\lambda^n}{1 - \lambda}\|x_o - x_1\| \leq \varepsilon$ si $n \geq N$ y puesto que $m \geq n \geq N$ entonces $\|x_n - x_m\| \leq \varepsilon$ si $m, n \geq N$. Esto último significa que la sucesión $\langle x_n \rangle$ es de cauchy. Siendo $\mathcal{X}$ completo y $\langle x_n \rangle$ una sucesión de cauchy en $F \subset \mathcal{X}$, entonces converge a un punto en $\mathcal{X}$. Pero como F es cerrado, contiene todos sus puntos límites y por tanto dicha convergencia esta en F. Ahora, sea $x \in F$ tal que $\lim_{n\to\infty} x_n = x$. Como T es contracción en F es fácil ver que T es uniformemente continua por lo que

$Tx = T(\lim_{n\to\infty} x_n) = \lim_{n\to\infty} Tx_n = \lim_{n\to\infty} x_{n+1} = x$. Esto demuestra la exdstencia del punto fijo.

La unicidad es consecuencia también del hecho de que T es contracción en F, para ver esto supongamos que $x,y \in F$ y que se tiene que $Tx = x$, $Ty = y$. Entonces $\|x - y\| = \|Tx - Ty\| \le \lambda\|x - y\|$ y por tanto la desigualdad $\|x - y\| \le \lambda\|x - y\|$ junto con el hecho de que $0 \le \lambda < 1$ implican que $\|x - y\| = 0$ por lo tanto $x = y$. ■

OBS.3. Si se toma el límite cuendo $m \to \infty$ en la desigualdad (**) se tiene por la continuidad de la norma que se verifica la desigualdad $\|x_n - x\| \le \frac{\lambda^n}{1 - \lambda}\|x_0 - x_1\|$ lo cual nos dice que $\langle x_n \rangle$ converge a x con la rapidez conque la sucesión de lado derecho converge a cero.

El siguiente resultado es una aplicación de teorema anterior el cual será utilizado en la sección posterior.

<u>TEOREMA</u> 1.1.2. Sea $\mathscr{X}$ un espacio de Banach, T una transformación lineal continua tal que $T: \mathscr{X} \to \mathscr{X}$ y $\|T\| < 1$. Entonces el operador $I - T$,donde I es la transformación identidad, es invertible y su inversa es continua.

DEMOSTRACION: Para que $I - T$ sea invertible es necesario y suficiente que la ecuación $(I - T)x = y$ tenga solución única para cada $y \in \mathscr{X}$. Escribamos $(I - T)x \doteq x - Tx$, así la ecuación $x - Tx = y$ es equivalente a $x = y + Tx$ y definiendo el operador $Ax = y + Tx$

el problema se vuelve en probar la existencia del punto fijo, para cada $y \in \mathscr{X}$, del operador A.

Primero se observa que $A:\mathscr{X} \to \mathscr{X}$. Sean $x_1, x_2 \in \mathscr{X}$, entonces se tiene que $\|Ax_1 - Ax_2\| = \|y + Tx_1 - (y + Tx_2)\| = \|Tx_1 - Tx_2\| = \|T(x_1 - x_2)\| \leq \|T\|\|x_1 - x_2\|$ y puesto que $\|T\| < 1$, por hipótesis, se tiene que A es una contracción sobre $\mathscr{X}$. Por el teorema 1.1.1, con $F = \mathscr{X}$, se deduce que A tiene un único punto fijo para cada $y \in \mathscr{X}$. De esta manera se satisface que, para cada $y \in \mathscr{X}$, $Ax = x$ para una única x y esto a su vez implica que la ecuación dada por $(I - T)x = y$ posee solución única por lo que $I - T$ es invertible.

Para probar la continuidad de la inversa definamos como $P(y)$,para cada $y \in \mathscr{X}$, el punto fijo correspondiente. Así P es la inversa de $I - T$ y es tal que $y + TP(y) = AP(y) = P(y)$ ó también $P(y) = y + TP(y)$ por lo que necesitamos verificar que $\forall\ h \in \mathscr{X}$ se cumple que $\lim_{h\to 0}\|P(y + h) - P(y)\| = 0$. Esto se hace como sigue.

$P(y) = y + TP(y) \quad \Rightarrow$

$P(y + h) - P(y) = y + h + TP(y + h) - y - TP(y)$. Esto implica que

$\|P(y + h) - P(y)\| \leq \|TP(y + h) - TP(y)\| + \|h\| = \|T[P(y + h) - P(y)]\| + \|h\| \Rightarrow \|P(y + h) - P(y)\| \leq \|T\|\|P(y + h) - P(y)\| + \|h\| \Rightarrow$

$(1 - \|T\|)\|P(y + h) - P(y)\| \leq \|h\| \Rightarrow \|P(y + h) - P(y)\| \leq \dfrac{\|h\|}{1 - \|T\|}$.

Nótese que $1 - \|T\| \neq 0$ pues $\|T\| < 1$ y tomando límites en ambos lados se concluye que $\lim_{h\to 0}\|P(y + h) - P(y)\| = 0$. ∎

2. DEPENDENCIA CONTINUA Y DIFERENCIABLE DE PUNTOS FIJOS

DEF.1.2.1. Sean $\mathcal{X}$, $\mathcal{Y}$ espacios de Banach, $F \subseteq \mathcal{X}$, $G \subseteq \mathcal{Y}$ con F cerrado. Sea también una familia de operadores $\{T_y : y \in G\}$ tal que $T_y : F \to \mathcal{X}$ para cada $y \in G$. Se dice que T_y es de contracción uniforme sobre F, si existe un número λ, con $0 \leq \lambda < 1$, tal que satisface que:

$$\|T_y x_1 - T_y x_2\| \leq \lambda \|x_1 - x_2\| \quad \forall\ y \in G, \quad \forall\ x_1, x_2 \in F.$$

El siguiente teorema utiliza fuertemente el concepto de contracción uniforme.

TEOREMA 1.2.1. (DEPENDENCIA CONTINUA DE PUNTOS FIJOS)

Sean $\mathcal{X}$, $\mathcal{Y}$ espacios de Banach, $F \subseteq \mathcal{X}$, $G \subseteq \mathcal{Y}$ con F cerrado. Supongamos que $T : F \to F$, $y \in G$, es de contracción uniforme y que el mapeo $y \to T_y x$, $y \in G$, es continuo para cada x fija en F. Entonces el punto fijo correspondiente a T_y, $y \in G$, es continuo en y.

DEMOSTRACIÓN: Sea $g(y)$ el punto fijo correspondiente a T_y para cada $y \in G$, que existe pues T_y es contracción sobre F para cada $y \in G$. Así $g(y) = T_y g(y)$ y la continuidad se tiene si probamos que $\lim_{h \to 0} \|g(y + h) - g(y)\| = 0$. Para esto hacemos lo siguiente.

$$g(y) = T_y g(y) \quad \Rightarrow \quad g(y + h) - g(y) = T_{y+h} g(y + h) - T_y g(y) =$$

$$T_{y+h} g(y + h) - \underline{T_{y+h} g(y) + T_{y+h} g(y)} - T_y g(y) \quad \Rightarrow$$

$$\|g(y + h) - g(y)\| \leq \|T_{y+h} g(y + h) - T_{y+h} g(y)\| + \|T_{y+h} g(y) - T_y g(y)\|$$

como T_y es contración uniforme entonces

$$\|T_{y+h} g(y + h) - T_{y+h} g(y)\| \leq \lambda \|g(y + h) - g(y)\| \quad , \ 0 \leq \lambda < 1, \quad \Rightarrow$$

$$\|g(y + h) - g(y)\| \leq \lambda \|g(y + h) - g(y)\| + \|T_{y+h} g(y) - T_y g(y)\| \quad \Rightarrow$$

$(1 - \lambda)\|g(y + h) - g(y)\| \leq \|T_{y+h}g(y) - T_y g(y)\|$ y así finalmente $\|g(y + h) - g(y)\| \leq (1 - \lambda)^{-1}\|T_{y+h}g(y) - T_{y+h}g(y)\|$. Puesto que por hipótesis el mapeo $y \to T_y x$ es continuo para x fija en F tenemos que $\lim_{h\to 0}\|T_{y+h}g(y) - T_y g(y)\| = 0$ y por tanto se cumple que

$$\lim_{h\to 0}\|g(y + h) - g(y)\| = 0. \blacksquare$$

El que sigue es un una aplicación conjunta de los teoremas, 1.1.1, 1.1.2, 1.2.1 y que será importante en el resultado posterior.

<u>TEOREMA</u> 1.2.2. Sean $\mathcal{X}$, $\mathcal{Y}$ espacios de Banach, $G \subseteq \mathcal{Y}$ y $T_y:\mathcal{X} \to \mathcal{X}$ transformaciones lineales continuas para cada $y \in G$ con $\|T_y\| \leq \lambda < 1$ $\forall\ y \in G$, λ fijo. Supongamos que el mapeo $y \to T_y x$ es continuo para cada x fija en $\mathcal{X}$. Entonces:

i) El operador $I - T_y$ tiene inversa continua para cada $y \in G$.

ii) El mapeo $y \to [I - Ty]$ es continuo.

DEMOSTRACION: Nuevamente el hecho de que $I - T_y$ sea invertible, para cada $y \in G$, es equivalente a que la ecuación $(I - T_y)x = z$ tenga solución única para cada $z \in \mathcal{X}$ y esto a su vez es equivalente a probar la existencia de los puntos fijos para el operador $A_{y,z}x = T_y x + z$, $x \in \mathcal{X}$. Para ver esto se tiene, para cada $y \in G$, $z \in \mathcal{X}$ que $A_{y,z}:\mathcal{X} \to \mathcal{X}$ y además $\forall\ s,t \in \mathcal{X}$ se verifica lo siguiente

$$\|A_{y,z}s - A_{y,z}t\| = \|T_y s + z - (T_y t + z)\| = \|T_y s - T_y t\| = \|T_y(s - t)\|$$

$$\leq \|T_y\|\|s - t\| \leq \lambda\|s - t\| \quad \forall\ y \in G, \forall\ z \in \mathcal{X}.$$

Así $A_{y,z}$ es de contración uniforme sobre $\mathcal{X}$ y por el teorema 1.1.1 existe, para cada $y \in G$, un único punto fijo correspondiente a cada $z \in \mathcal{X}$. si tal punto fijo lo denotamos como $g(y, z)$ entonces $g(y, z) = [I - T_y]^{-1}z$, para cada $y \in G$, y es continuo en z por el teorema 1.1.2. Por lo tanto se tiene i). Para ver ii) debemos probar que $\lim_{h\to 0}\|[I - T_{y+h}]^{-1} - [I - T_y]^{-1}\| = 0$; Nótese que la norma aquí es tomada en $\mathcal{L}(\mathcal{X},\mathcal{X})$. Primeramente, para cada $x \in \mathcal{X}$ tenemos que $A_{y+h,z}x - A_{y,z}x = T_{y+h}x + z - T_y x - z = T_{y+h}x - T_y x$, por lo tanto

$\lim_{h\to 0}\|A_{y+h,z}x - A_{y,z}x\| = \lim_{h\to 0}\|T_{y+h}x - T_y x\| = 0$ pues el mapeo $y \to T_y x$ es continuo, por hipótesis. Luego también el mapeo $y \to A_{y,z}x$ es continuo para cada x fija en $\mathcal{X}$ y es uniforme respecto a z. Por el teorema 1.2.1 se cumple que:

$$\|[I - T_{y+h}]^{-1}z - [I - T_y]^{-1}z\| = \|g(y + h, z) - g(y, z)\| \leq$$

$$(1 - \lambda)^{-1}\|A_{y+h,z}g(y, z) - A_{y,z}g(y, z)\| \qquad \forall\, z \in \mathcal{X}.$$

Como esta desigualdad es válida $\forall\, z \in \mathcal{X}$, en particular también será válida si tomamos el supremo sobre la bola unitaria en el lado izquierdo. Así se sigue satisfaciendo que:

$\|[I - T_{y+h}]^{-1} - [I - T_y]^{-1}\| \leq (1 - \lambda)^{-1}\|A_{y+h,z}g(y, z) - A_{y,z}g(y,z)\|$ y

Puesto que $\lim_{h\to 0}\|A_{y+h,z}g(y, z) - A_{y,z}g(y, z)\| = 0$ se concluye finalmente que $\lim_{h\to 0}\|[I - T_{y+h}]^{-1} - [I - T_y]^{-1}\| = 0$ ∎

Ahora estamos dispuestos a dar otro de los teoremas centrales de este capitulo.

<u>TEOREMA</u> 1.2.3. **(DEPENDENCIA DIFERENCIABLE DE PUNTOS FIJOS)**

Sean $\mathcal{X}$, $\mathcal{Y}$ espacios de Banach, $F \subseteq \mathcal{X}$, $G \subseteq \mathcal{Y}$ con F cerrado. Sea también $T_y:F \to F$, $y \in G$, una contracción uniforme sobre F continua en y, para cada x fija en F. Si F, G son las clausuras de los conjuntos abiertos F°, G° y si $T_y x$ tiene primeras derivadas parciales continuas $A(x,y)$, $B(x,y)$ respectivamente, entonces el punto fijo correspondiente $g(y)$ es continuamente diferenciable en G°, esto es, $g \in \mathcal{C}^1(G^\circ)$.

DEMOSTRACION: Se supondrá por un momento que $g(y)$ es diferenciable y en donde la diferencial esta dada por la transformación lineal, en h, $Z(y,h) = S(y)h$. partiendo de esta hipótesis se probará que tal transformación lineal existe, es única y que además $S(y)$ es la derivada continua con respecto a $y \in G^\circ$.

Para principiar pongamos $T_y x = T(x,y)$ de manera que:

$T:F^\circ \times G^\circ \to F^\circ$ y $A(x,y) = \frac{\partial T}{\partial x}(x,y)$, $B(x,y) = \frac{\partial T}{\partial x}(x,y)$. Aceptando que la regla de la cadena es válida podemos aplicarla a la ecuación $g(y) = T(g(y), y)$ y representando con Z a la diferencial de g tenemos la expresión:

$Z = \frac{\partial T}{\partial x}(g(y),y)Z + \frac{\partial T}{\partial y}(g(y),y)h \quad h \in \mathcal{Y}$. De Aquí que también tenemos $[I - \frac{\partial T}{\partial x}(g(y),y)]Z = \frac{\partial T}{\partial x}(g(y),y)h$ por lo tanto para poder despejar Z debemos invertir el término dado por $I - \frac{\partial T}{\partial x}(g(y),y)$. Esto será la aplicación del teorema 1.2.2.

Puesto que las parciales son continuas tenemos que $\frac{\partial T}{\partial x}(x,y)$ es continua en $y \in G^\circ$. Ahora se probará que $\|\frac{\partial T}{\partial x}(x, y)\| < \sigma < 1$ para $x \in F^\circ$, $y \in G^\circ$. Por definición, para $x \in F^\circ$, $y \in G^\circ$, $\frac{\partial T}{\partial x}(x, y)$ es el mapeo lineal de $\mathscr{X}$ en $\mathscr{X}$ que satisface que

$\lim\limits_{k\to 0} \dfrac{\left\|T(x + k,y) - T(x,y) - \frac{\partial T}{\partial x}(x,y)k\right\|}{\|k\|} = 0$. Esto a su vez significa que dado $\varepsilon > 0$ existe $\delta > 0$ de tal manera que se cumple que:

$$\left\|T(x + k,y) - T(x,y) - \frac{\partial T}{\partial x}(x,y)k\right\| < \varepsilon\|k\| \qquad \text{si } \|k\| < \delta \Rightarrow$$

$$\left\|\frac{\partial T}{\partial x}(x, y)k\right\| - \|T(x + k, y) - T(x, y)\| < \varepsilon\|k\| \qquad \text{si } \|k\| < \delta \Rightarrow$$

$$\left\|\frac{\partial T}{\partial x}(x, y)k\right\| < \varepsilon\|k\| + \|T(x + k, y) - T(x, y)\| \qquad \text{si } \|k\| < \delta$$

debido a que T es de contracción uniforme sobre F entonces se tiene que $\|T(x + k,y) - T(x,y)\| \leq \lambda\|k\|$ $\forall\ y \in G$, x, $x + k \in F^\circ$. Así $\|\frac{\partial T}{\partial x}(x,y)k\| < \varepsilon\|k\| + \lambda\|k\| = (\varepsilon + \lambda)\|k\|$ si $\|k\| < \delta$ para $x \in F^\circ$, $y \in G^\circ$. Como $0\leq \lambda < 1$ entonces siempre es posible elegir un número ε de forma que $\sigma = \varepsilon + \lambda < 1$, por tanto tenemos $\|\frac{\partial T}{\partial x}(x,y)k\| < \sigma\|k\|$ si $\|k\| < \delta$. De aquí se deduce que $\|\frac{\partial T}{\partial x}(x, y)\| < \sigma$ $x \in F^\circ$, $y \in G^\circ$. Por el teorema 1.2.2, para cada $y \in G^\circ$, el operador $I - \frac{\partial T}{\partial x}(g(y),y)$ tiene inversa continua en $y \in G^\circ$, por lo que podemos encontrar una única solución, para cada $y \in G^\circ$, de forma que para $h \in \mathscr{Y}$ tengamos que $Z(y, h) = \left[I - \frac{\partial T}{\partial x}(g(y),y)\right]^{-1}\frac{\partial T}{\partial y}(g(y),y)h$, la cual además es continua en y, h respectivamente. Si $Z(y, h) = S(y)h$ entonces S es el mapeo lineal de $\mathscr{Y}$ en $\mathscr{X}$ que esta dado por la fórmula:

(*) $\quad S(y) = \left[I - \frac{\partial T}{\partial x}(g(y), y)\right]^{-1}\frac{\partial T}{\partial y}(g(y),y)$ el cual además es continuo en h y además es continuo en y en G°. Para ver que $S(y)$ es la derivada de $g(y)$ tenemos que probar que:

$$\lim_{h\to 0} \frac{\|g(y + h) - g(y) - S(y)h\|}{\|h\|} = 0.$$

las Siguientes relaciones nos serán útiles para este fin. Sea h pequeño de tal forma que si $y \in G^{\circ}$, $y + h \in G^{\circ}$. Entonces tenemos

$T(g(y + h),y + h) - T(g(y + h),y) = \frac{\partial T}{\partial y}(g(y + h),y)h + O(h)$ y

$T(g(y + h), y) - T(g(y),y) = \frac{\partial T}{\partial x}(g(y),y)[g(y + h) - g(y)] + O[g(y + h) - g(y)]$. De estas dos igualdades obtenemos que

$g(y + h) - g(y) = T(g(y + h),y + h) - T(g(y),y) =$

$T(g(y + h),y + h) - T(g(y + h),y) + T(g(y + h),y) - T(g(y),y) =$

$\frac{\partial T}{\partial y}(g(y + h),y)h + O(h) + \frac{\partial T}{\partial x}(g(y),y)[g(y + h) - g(y)] +$

$O[g(y + h) - g(y)]$. Reordenando estos últimos términos tenemos que

$g(y + h) - g(y) = \frac{\partial T}{\partial x}(g(y),y)[g(y + h - g(y)] + \frac{\partial T}{\partial y}(g(y + h,y)h +$

$O(h) + O[g(y + h) - g(y)]$.Si en el lado derecho de ésta sumamos y restamos la expresión dada por $\frac{\partial T}{\partial x}(g(y),y)S(y)h$ y $S(y)$ consecutivamente, obtenemos que

$$g(y + h) - g(y) = \frac{\partial T}{\partial x}(g(y),y)[g(y + h - g(y) - S(y)h] + \frac{\partial T}{\partial x}(g(y),y)S(y)h + \frac{\partial T}{\partial y}(g(y + h), y)h + O(h) + O[g(y + h) - g(y)] \Rightarrow$$

$$g(y + h) - g(y) - S(y)h = \frac{\partial T}{\partial x}(g(y), y)[g(y + h - g(y) - S(y)h] + \frac{\partial T}{\partial x}(g(y),y)S(y)h + \frac{\partial T}{\partial y}(g(y + h),y)h + O(h) + O[g(y + h) - g(y)] - S(y)h.$$

Ahora, si agrupamos los términos obtenemos la igualdad

$$\left[I - \frac{\partial T}{\partial x}(g(y),y)\right][g(y + h) - g(y) - S(y)h] = -\left[I - \frac{\partial T}{\partial x}(g(y),y)\right]S(y)h + \frac{\partial T}{\partial y}(g(y + h), y)h + O(h) + O[g(y + h) - g(y)].$$

Debido a que por (*) se cumple que $\left[I - \frac{\partial T}{\partial x}(g(y),y)\right]S(y)h = \frac{\partial T}{\partial y}(g(y),y)h$ entonces

$$g(y + h) - g(y) - S(y)h = \left[I - \frac{\partial T}{\partial x}(g(y),y)\right]^{-1}\left[- \frac{\partial T}{\partial y}(g(y),y)h + \frac{\partial T}{\partial y}(g(y + h),y)h + O(h) + O[g(y + h) - g(y)]\right] \Rightarrow$$

$$\|g(y + h) - g(y) - S(y)h\| \le \left\|\left[I - \frac{\partial T}{\partial x}(g(y),y)\right]^{-1}\right\|\left(\left\|\frac{\partial T}{\partial y}(g(y + h,y)h - \frac{\partial T}{\partial y}(g(y),y)h\right\| + \|O(h)\| + \|O[g(y + h) - g(y)]\|\right) \Rightarrow$$

$$\frac{\|g(y + h) - g(y) - S(y)h\|}{\|h\|} \le$$

$$\left\|\left[I - \frac{\partial T}{\partial x}(g(y),y)\right]^{-1}\right\|\left(\frac{\left\|\frac{\partial T}{\partial y}(g(y + h),y) - \frac{\partial T}{\partial y}(g(y),y))\right\|\|h\|}{\|h\|} + \frac{\|O(h)\|}{\|h\|} + \frac{\|O[g(y + h) - g(y)]\|}{\|h\|}\right).$$

Ahora al tomar

límites se observa que:

i) $\lim_{h\to 0} \left\| \frac{\partial T}{\partial y}(g(y + h),y) - \frac{\partial T}{\partial y}(g(y),y) \right\| = 0$. Por la continuidad de la parcial.

ii) $\lim_{h\to 0} \frac{\|O(h)\|}{\|h\|} = 0$. Por definición.

De aquí, que es claro que solo necesitamos probar que también se tiene que

iii) $\lim_{h\to 0} \frac{\|O[g(y + h) - g(y)]\|}{\|h\|} = 0$. Para esto, pongamos la igualdad

$$\frac{\|O[g(y + h) - g(y)]\|}{\|h\|} = \frac{\|O[g(y + h) - g(y)]\|}{\|g(y + h) - g(y)\|} \cdot \frac{\|g(y + h) - g(y)\|}{\|h\|}$$

y en vista de que $\lim_{h\to 0} \frac{\|O[g(y + h) - g(y)]\|}{\|g(y + h) - g(y)\|} = 0$ entonces

$\lim_{h\to 0} \frac{\|O[g(y + h) - g(y)]\|}{\|h\|} = 0$ si $\frac{\|g(y + h) - g(y)\|}{\|h\|}$ está acotado.

Para esto último tenemos que por el toerema 1.2.1 se satisface la desigualdad:

$$\|g(y + h) - g(y)\| \leq (1 - \lambda)^{-1} \|T(g(y),y + h) - T(g(y),y)\|$$

además, por el torema 0.3.2, T es diferenciable con continuidad pues sus derivadas parciales existen y son continuas. Ahora aplicando el teorema 0.3.3 se tiene la desigualdad

$$\|g(y + h) - g(y)\| \leq (1 - \lambda)^{-1} \sup_{\xi \in [y, y+h]} \|DT(\xi)\| \, \|h\| \quad \text{o también}$$

$$\frac{\|g(y + h) - g(y)\|}{\|h\|} \le (1 - \lambda)^{-1} \sup_{\xi \in (y, y+h)} \|DT(\xi)\|$$

puesto que DT es continua, es acotada en el compacto [y, y + h] por lo cual el sup existe y es finito.Por lo tanto se tiene que

$$\frac{\|g(y + h) - g(y)\|}{\|h\|}$$ esta acotado y finalmente los limites i), ii) y iii) implican que:

$$\lim_{h \to 0} \frac{\|g(y + h) - g(y) - S(y)h\|}{\|h\|} = 0. \blacksquare$$

CAPITULO 2

TEOREMA DE LA FUNCION INVERSA Y TEOREMA DE LA FUNCION IMPLICITA

El trabajo aquí es hacer un estudio detallado del teorema de la función inversa y de teorema de la función implícita. Ambos resultados pudieramos decir son los ejemplos mas plausibles en los que se refleja el hecho de que la técnica del cálculo diferencial, el de *aproximación por tangentes o linealización*, es la base fuerte con la cual se investiga el comportamiento local de las funciones continuamente diferenciables. Sin precisar el primero establece, entre otras cosas, que si la linealización de alguna función $y = f(x)$ en un punto x_o es invertible, esto es $Df(x_o)$ invertible, también lo es f cerca de x_o. Así mismo el segundo ataca el problema de determinar cuando una relación entre dos variables x, y, dada por la ecuación $F(x,y) = 0$, define a una como función de la otra, es decir, si existe, por ejemplo, una función de la forma $x = g(y)$ de forma que resuelva la ecuación anterior en el sentido de que $F(g(y),y) = 0$. Además la demostración del teorema de la función inversa se hace primeramente planteándolo como un problema de punto fijo y resolviendo esto último mediante la aplicación del *teorema de contracción de Banach* y del *teorema de dependencia continua y diferenciable de puntos fijos* ya expuestos en el capítulo anterior. Por último es interesante observar que en el teorema de la función inversa y en el teorema de la función implícita, el problema escencial redunda en la posibilidad de poder "*despejar*" una variable en términos de la otra, por lo cual, es una de las cuestiones mas importantes en las matemáticas.

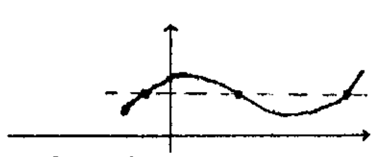
1. EL TEOREMA DE LA FUNCION INVERSA

Para empezar sabemos que si A, B son conjuntos arbitrarios y $f: A \to B$ una función que es uno a uno y sobre, entonces se dice que f es invertible y comunmente suele denotarse como f^{-1} a la inversa de f y en donde f^{-1} es la regla que asocia cada elemento de B un único elemento de A y que satisface las siguientes fórmulas:

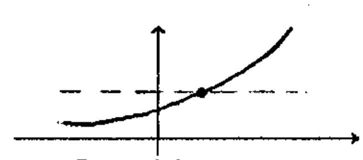
$$f[f^{-1}(y)] = y \quad y \in B \quad \text{y} \quad f^{-1}[f(x)] = x \quad x \in A.$$

De entre las funciones reales de variable real mas sencillas de analizar son aquellas que describen rectas, no verticales, en en el plano y éstas están dadas en la forma $y=f(x)=mx + b$. Es claro que toda recta no horizontal determina una función uno a uno y sobre de $\mathbb{R}$ en $\mathbb{R}$. Esto ocurre cuando el valor de la pendiente, m, es distinto de cero, por lo cual, en este caso es fácil calcular la inversa despejando directamente la variable x, esto es, $x=f^{-1}(y)=\frac{1}{m}y + \frac{1}{m}b$. Nótese que aquí la pendiente de la recta $x=f^{-1}(y)$ es precisamente la recíproca de la pendiente de la recta $y=f(x)$. Esta situación es mas complicada cuando consideramos otro tipo de funciones, y mas aún, cuando las funciones están definidas en espacios mas generales como los $\mathbb{R}^n$. En el caso real de variable real contamos con un criterio conocido para determinar cuando una función es uno a uno y es geométrico:

"si toda recta horizontal corta a la gráfica de una función f en a lo mas un punto, entonces f es uno a uno".



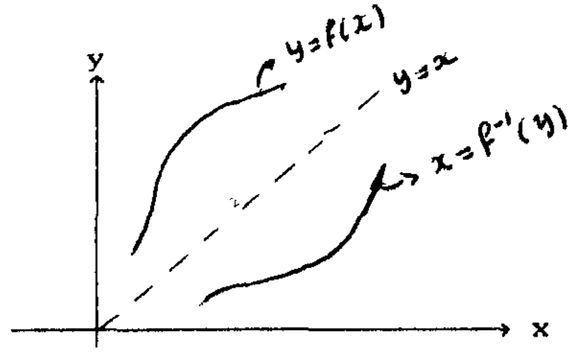
función no uno a uno



función uno a uno

Otro criterio utilizado es meramente algebraćico y es el siguiente; "*si* $f(x_1)=f(x_2) \Rightarrow x_1=x_2$, *entonces* f *es uno a uno*".

Cuando una función es invertible, es usual representar la gráfica de f^{-1} tomando la reflexión de la gráfica de f sobre la recta y=x, como se ilustra a continuación

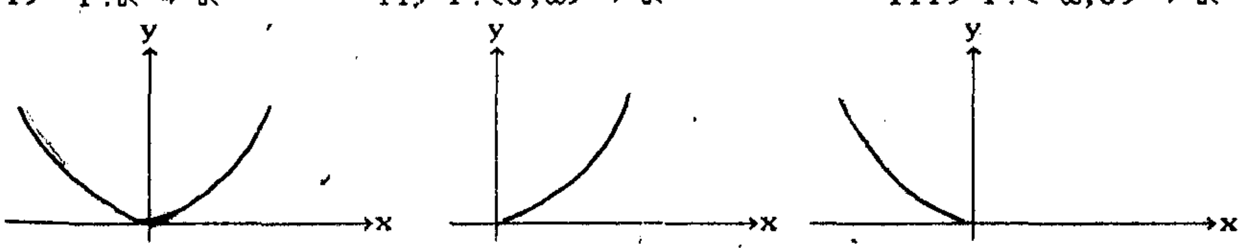


Esto se hace simplemente intercambiando los papeles de las variabiables x, y respectivamente.

Si queremos ser mas estrictos con la cuestión de las inversas el siguiente ejemplo nos muestra, entre otras cosas, la importancia de considerar el dominio de definición de f.

Sea la función dada por $f(x)=x^2$ y consideremos los siguientes casos:

i) $f:\mathbb{R} \to \mathbb{R}$ ii) $f:(0,\infty) \to \mathbb{R}^+$ iii) $f:(-\infty,0) \to \mathbb{R}^+$



f no es uno a uno f es uno uno f es uno a uno

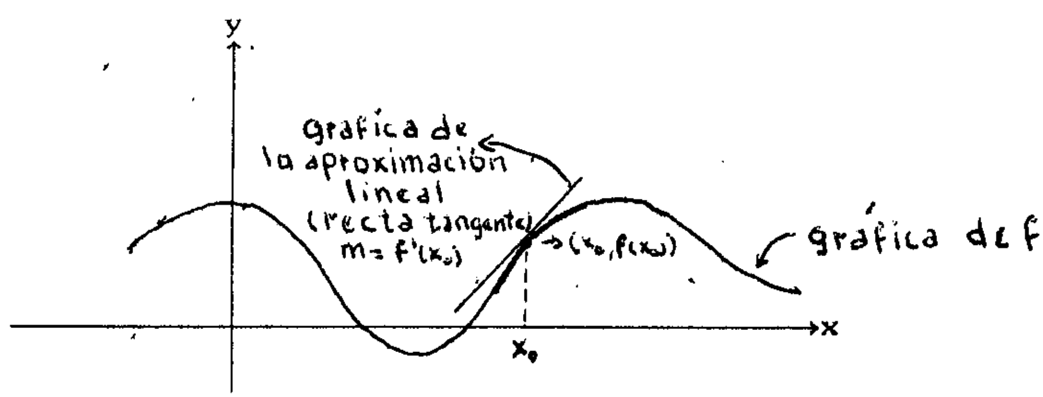
Observemos que las ramas ii) y iii) definen la misma función pero en dominios distintos, en ambqs casos f es invertible mas no con la misma inversa, es decir, podemos encontrar dominios donde una misma función posea inversa, pero no necesariamente ésta tiene que ser la misma. Sin embargo, donde existen son únicas.

En este ejemplo las inversas vienen dadas por:

$$\text{ii)}\ x = f^{-1}(y) = \sqrt{y} \qquad \text{iii)}\ x = f^{-1}(y) = -\sqrt{y}.$$

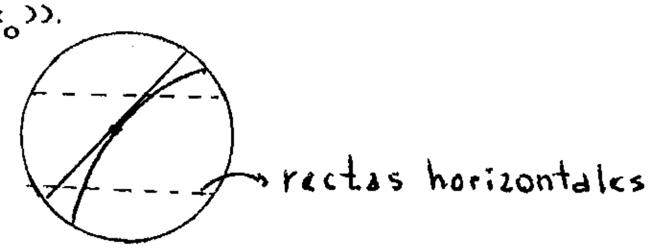
Ahora estudiaremos la invertivilidad de una función f, desde el punto de vista de cálculo. La idea central consiste en lo siguiente: *Si una función f es diferenciable, y más aun, con continuidad en algún conjunto abierto D, la derivada $f'(x_0)$, con $x_0 \in D$, nos dá una muy buena aproximación lineal en las cercanías del punto x_0. Dicho de otro modo, f es localmente muy parecida a tal aproximación cuando nos limitamos a vecindades pequeñas de x_0, de ahí que esperamos tener que el comportamiento local de f sea como el de la transformación lineal $f'(x_0)$.*

Discutamos lo dicho anteriormente volviendo al caso de funciones reales de variable real. Iniciemos suponiendo que f es diferenciable y que para algun punto x_0, $f'(x_0)$ es invertible , es decir,(*)$f'(x_0) \neq 0$. Representemos esto con la figura siguiente

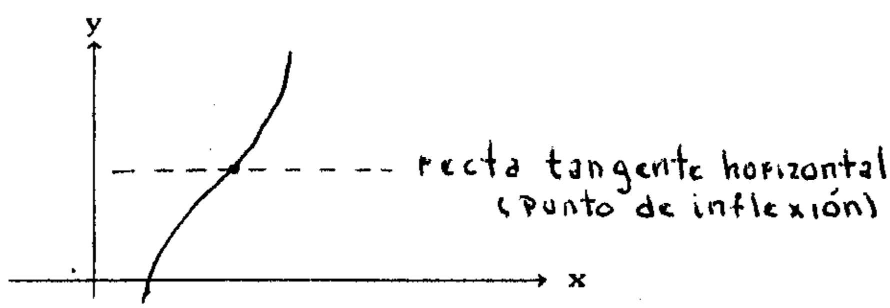


(*) En el caso general en que $f: D \subseteq \mathbb{R}^n \to \mathbb{R}^n$, la derivada $f'(x)$ es precisamente la matriz jacobiana de nxn y pedir su invertibilidad es pedir que su determinante,(jacobiano), sea distito de cero.

si $m=f'(x_o)\neq 0$, la recta tangente a la curva $y=f(x)$ en el punto $(x_o, f(x_o))$ no es horizontal y por tanto cada recta horizontal la corta una sola vez. Asi, es de esperarse, que esta propiedad, uno a uno, que es la condición para que una función se pueda invertir, se transfiera a la curva $y=f(x)$ por estar tan próxima a la recta tangente cerca del punto $(x_o, f(x_o))$.



Donde pudiera darse complicaciones, es en los puntos donde la tangente es horizontal, ahi la derivada es nula y puede resultar una concavidad. Se descarta entonces esta posibilidad aun cuando se tengan funciones uno a uno en prescencia de tangentes horizontales.



Hasta aquí, de alguna manera, hemos argumentado de que la hipótesis de invertivilidad de la derivada en un punto conduce a concluir que f es localmente invertible.Sin embargo, esto no basta en todos los casos para asegurar dicha afirmación pues, como veremos, la función definida por;

$$f(x)=\begin{cases} x + x^2 \operatorname{sen} 1/x, & x\neq 0 \\ 0 & x=0 \end{cases}$$

tiene la propiedad de que $f'(0)\neq 0$ sin ser uno a uno en ninguna vecindad de cero. Aquí se necesita una hipótesis adicional mas fuerte que permita asegurar loa invertivilidad de una función f alrededor de un punto y esta es que además f' sea continua en el punto en cuestión. Veamos como es esto.

(**)"Si $f'(x_o)\neq 0$ y f'es continua en x_o,entonces, debido al principio de conservación de signo de los mapeos continuos, existe un abierto U de x_o en donde $f'(x)\neq 0$ y teniendo el mismo signo que $f'(x_o)$". Podemos ver que f, bajo estas condiciones, en realidad es uno a uno en U.

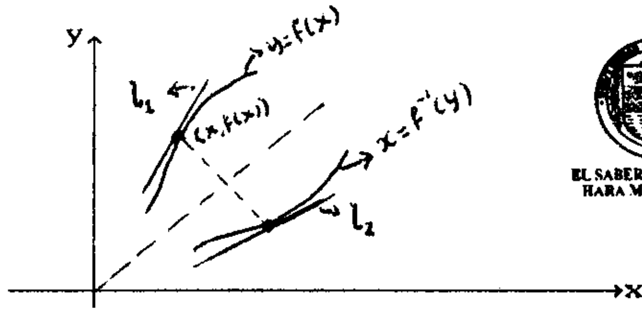
i) Geométricamente esto nos dice que las tangentes sobre la gráfica de f en los puntos $(x,f(x))$, $x \in U$, no son horizontales evitando así la posibilidad de que se presente alguna concavidad.

ii) Analíticamente se ve de la siguiente manera, si aplicamos el teorema del valor medio tenemos que $\forall$ x, y $\in$ U se verifica la igualdad $|f(x) - f(y)|=|f'(\xi)||x - y|$ donde ξ esta entre x y y; siendo $f'(x)\neq 0$ en U entonces $|f'(\xi)|\neq 0$ y si $f(x)=f(y)$ tenemos que $|f(x) - f(y)|=0 \Rightarrow |x - y|=0$ y por tanto x=y.De esto último se deduce que f es uno a uno en U.

Por lo pronto, ya se ha visto las condiciones a imponer a una función f que garanticen su invertibilidad cerca de un punto x_o. Sin embargo, éstas, nada nos dicen respecto a como se puede encontrar la inversa f^{-1}; pues, solo nos hablan de su existencia local. Pero algo muy importante acerca de las inversas es que

(**)En el caso general en que $f:D\subseteq\mathbb{R}^n \rightarrow \mathbb{R}^n$, el hecho de que la matriz jacobiana $f'(x_o)$ sea invertible y continua en x_o, significa que a variaciones pequeñas de x_o, se tienen variaciones pequeñas en las componentes de dicha matriz de tal forma que se mantiene, cerca de x_o, un jacobiano no nulo o en otras palabras que la matriz jacobiana $f'(x_o)$ se conserva invertible cerca de x_o.

cuando existen, heredan muchas de las propiedades de las funciones de las cuales provienen, como por ejemplo, la continuidad y la diferenciabilidad. De hecho existe una relación entre la derivada de f y la corespondiente a su inversa, cuando esta también es derivable. Ilustremos esto con lo que sigue a continuación.



Si l_1 es la tangente gráfica de f en el punto $(x,f(x))$, su pendiente esta dada por $f'(x)$. Además si l_2 es la tangente a la gráfica de f^{-1} en el punto simétrico $(y,f^{-1}(y))$ entonces l_2 es la inversa de l_1 y por tanto su pendiente debe ser la reciproca $\frac{1}{f'(x)}$ si $f'(x)\neq 0$. Admitiendo ahora que f^{-1} es diferenciable entonces tendremos que $[f^{-1}(y)]'=$ pendiente de l_2 en $(y,f^{-1}(y)) = \frac{1}{f'(x)} = \frac{1}{f'(f^{-1}(y))} = [f'(f^{-1}(y)]^{-1}$. Así la relación esta dada por:

$$[f^{-1}(y)]' = [f'(f^{-1}(y)]^{-1}.$$

Esto último es una consecuencia inmediata de la regla de la cadena aplicada a la ecuación $f(f^{-1}(y))=y$.

Estamos yá en condiciones de formular y demostrar lo que sería el teorema de la función inversa, pero antes de hacerlo, nos devolveremos para introducir el caso general estudiando el problema relacionado con las soluciones de sistemas de ecuaciones no lineales. Veamos esto, iniciando con el caso mas sencillo. Se sabe que un sistema lineal de n ecuaciones con n, incógnitas la forma:

$$
(*)\qquad
\begin{array}{l}
a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = y_1 \\
a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = y_2 \\
\dots\dots\dots\dots\dots\dots\dots\dots \\
a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n = y_n
\end{array}
$$

donde las a's y las y's son dados. Decimos que este sistema tiene solución única si el determinante de la matriz asociada es distinto de cero. Esto lo representamos como:

$$
\begin{vmatrix}
a_{11} & a_{12} & \dots & a_{1n} \\
a_{21} & a_{22} & \dots & a_{2n} \\
\dots & \dots & \dots & \dots \\
a_{n1} & a_{n2} & \dots & a_{nn}
\end{vmatrix} \neq 0.
$$
Si ponemos que

$$
A = \begin{pmatrix}
a_{11} & a_{12} & \dots & a_{1n} \\
a_{21} & a_{22} & \dots & a_{2n} \\
\dots & \dots & \dots & \dots \\
a_{n1} & a_{n2} & \dots & a_{nn}
\end{pmatrix}, \quad
x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad
y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}
\quad \text{y} \quad Tx = Ax \ ,
$$

entonces el sistema anterior es abreviado por la ecuación **Tx = y**. Así el determinar si (*) tiene solución única es equivalente a probar que Tx = y se satisface para una única x. Obsérvese que T es una transformación lineal de $\mathbb{R}^n$ en $\mathbb{R}^n$ y por tanto una única solución existe, si y solo si, la matriz que lo representa, A, es invertible. Nótese también que en este caso la regla que asocia, para cada $y \in \mathbb{R}^n$, un único vector $x \in \mathbb{R}^n$ es conocida como **la regla de cramer.**

Una generalización del sistema (*) esta dado por:

$$(**)\qquad \begin{aligned} f_1(x_1,x_2,\dots,x_n) &= y_1\\ f_2(x_1,x_2,\dots,x_n) &= y_2\\ &\dots\dots\dots\\ f_n(x_1,x_2,\dots,x_n) &= y_n \end{aligned}$$

donde $f_1, f_2, \dots f_n : D \subseteq \mathbb{R}^n \to \mathbb{R}$, $f_i \in \mathscr{C}^1(D)$, con D abierto.

De nueva cuenta deseamos dar un criterio que nos permita determinar la existencia da soluciones únicas para este sistema no lineal. Nótese que (*), es justamente el caso especial de (**) cuando los f_i son lineales y en el que dicho criterio se reduce a invertir una matriz.

Como antes si $x = (x_1,x_2,\dots,x_n)$, $y = (y_1,y_2,\dots,y_n)$ y la función $f(x_1,x_2,\dots,x_n) = (f_1(x),f_2(x),\dots,f_n(x))$ entonces el sistema (**) se representa por $f(x) = y$. Así nuestro problema se traduce en resolver para las x's en función de las y's. Asi mismo, esto equivale a buscar la existencia de la función inversa f^{-1} que expresa a x en términos de y y esto es precisamente a lo que responde nuestro primer resultado principal.

<u>TEOREMA</u> 2.1.1 **(DE LA FUNCION INVERSA)**

Sea D un subconjunto abierto de $\mathbb{R}^n$ y $f:D \to \mathbb{R}^n$ con $f \in \mathscr{C}^1(D)$. supongamos que para $a \in D$, $f(a) = b$ y que $f'(a)$ es invertible. Entonces:

i) Existen conjuntos abiertos U, V de $\mathbb{R}^n$ tales que $a \in U$, $b \in V$, f es uno a uno en U y $f(U) = V$.

ii) Sea g la inversa de f que existe por i). Entonces $g \in \mathscr{C}^1(V)$ y $g'(y) = [f'(g(y))]^{-1}$.

DEMOSTRACION: La prueba, como sugiere el teorema, será dado en dos partes. La parte i), que es de existencia, se planteará como un problema de punto fijo introduciendo el operador dado por $T(x,y) = x + [f'(a)]^{-1}(y - f(x))$ para cada $y \in \mathbb{R}^n$. En este proceso se ajustará el abierto V y se mostrará, por el teorema de contracción de Banach, que T tiene un único punto fijo, el cual denotaremos por $g(y)$, para cada $y \in V$. Este resultará ser precisamente la inversa de f. La parte ii) será una consecuencia inmediata del teorema de dependencia diferenciable de puntos fijos.

Parte i). Primero, como $[f'(a)]^{-1}$ es un operador lineal de $\mathbb{R}^n$ en $\mathbb{R}^n$, es continuo y por tanto acotado. Así podemos elegir un número k que sea tal que $\|[f'(a)]^{-1}\| = \frac{1}{2k}$. Como f' es continua en a entonces para $\varepsilon = k$ existe $\delta > 0$ tal que se satisface que $\|f'(x) - f'(a)\| < k$ si $\|x - a\| < \delta$. Además, siempre nos es posible elegir un $\delta_o < \delta$ de tal manera que también se tenga que $\|f'(x) - f'(a)\| < k$ si $\|x - a\| \leq \delta_o$. Definamos el conjunto $W = \left\{x \in \mathbb{R}^n : \|x - a\| \leq \delta_o\right\}$; así W es cerrado y se tiene en este caso que $\|f'(x) - f'(a)\| < k$ si $x \in W$. Ahora pongamos para $x \in D$, $y \in \mathbb{R}^n$ el operador $T: D \times \mathbb{R}^n \to \mathbb{R}^n$ definido por

$$T(x,y) = x + [f'(a)]^{-1}(y - f(x))$$

demostraremos primero que dicho operador, es de contracción uniforme sobre W. Para hacer esto tenemos, calculando la derivada parcial con respecto a x, que

$$\frac{\partial T}{\partial x}(x,y) = I - [f'(a)]^{-1}f'(x) = [f'(a)]^{-1}(f'(a) - f'(x)].$$

Ahora, al tomar las normas, obtenemos la desigualdad:

$$\left\|\frac{\partial T}{\partial x}(x,y)\right\| = \|[f'(a)]^{-1}(f'(a) - f'(x)\| \leq \|[f'(a)]^{-1}\|\,\|(f'(a) - f'(x)\|$$

$< \frac{1}{2k}k = \frac{1}{2}$. $\quad x \in W, \quad y \in \mathbb{R}^n$. De esta manera se tiene que

$$\left\|\frac{\partial T}{\partial x}(x,y)\right\| < \frac{1}{2} \qquad x \in W,\ y \in \mathbb{R}^n.$$

Ahora para $x_1,\ x_2 \in W$ se deduce del teorema 0.3.3 la desigualdad

$$\|T(x_1,y) - T(x_2)\| \leq \sup_{\xi \in [x_1, x_2]} \left\|\frac{\partial T}{\partial x}(\xi,y)\right\| \|x_1 - x_2\| \qquad \Rightarrow$$

$$\|T(x\ ,y) - T(x\ ,y)\| < \frac{1}{2}\|x_1 - x_2\| \qquad \forall\ y \in \mathbb{R}^n.$$

Esto último nos dice que el operador T, es de contracción uniforme en W. Ahora probaremos eligiendo adecuadamente las y's que T manda puntos de W en puntos de W. Fijemos $x = a$ y usando la definición de T se cumple que $\|T(a,y) - a\| = \|a + [f'(a)]^{-1}(y - f(a)) - a\| =$ $\|[f'(a)]^{-1}(y - b)\| \leq \|[f'(a)]^{-1}\|\,\|y - b\|$. De esta manera tenemos la desigualdad $\|T(a,y) - a\| \leq \|[f'(a)]^{-1}\|\,\|y - b\|$. Pongamos el conjunto $V = \left\{y \in \mathbb{R}^n : \|y - b\| < k\delta_o\right\}$; entonces V es abierto y tenemos que $\|T(a,y) - a\| \leq \frac{1}{2k}\|y - b\| < \frac{1}{2k}k\delta_o = \frac{\delta_o}{2}$ si $y \in V$. Así tenemos que $T(a,y) \in W$ si $y \in V$.

Tomemos Ahora x arbitrario en W. En este caso obtenemos que $\|T(x,y) - a\| = \|T(x,y) - T(a,y) + T(a,y) - a\| \leq \|T(x,y) - T(a,y)\|$ $+ \|T(a,y) - a\| < \frac{1}{2}\|x - a\| + \|T(a,y) - a\| < \frac{1}{2}\delta_o + \frac{1}{2}\delta_o = \delta_o$ si $x \in W$, $y \in V$. De aquí se deduce que para cada $y \in V$ y $x \in W$ se tiene que $T(x,y) \in W$. Siendo T de contracción en el cerrado W, se dá, por el teorema de contracción de Banach la existencia de un único punto fijo para cada $y \in V$.

Designemos, como antes, a $g(y)$ como el punto fijo correspondiente a $T(\circ,y)$ para cada $y \in V$, esto es, se satisface que $T(g(y),y)=g(y)$

si $y \in V$. De aquí que:

$T(g(y),y) = g(y) + [f'(a)]^{-1}[y - f(g(y))] = g(y)$. Esto implica que $f(g(y)) = y$ y por tanto g es la inversa del mapeo f y como f es continua $f^{-1}(V)$ es abierto. Finalmente se pone $U = f^{-1}(V)$ y es claro que f es uno a uno y sobre de U en V con $x = g(y) = f^{-1}(y)$ como la inversa de f.

Parte ii). Pongamos $T:U \times V \to U$ y, como antes, con T dado por $T(x,y) = x + [f'(a)]^{-1}(y - f(x))$. Sabemos que para cada $y \in V$, $g(y) \in U$ y $T(g(y),y) = g(y)$. También se ve que T es de clase $\mathscr{C}^1$, puesto que:

$\frac{\partial T}{\partial x}(x,y) = I - [f'(a)]^{-1}f'(x)$ y $\frac{\partial T}{\partial y}(x,y) = [f'(a)]^{-1}$ son continuos. Ahora por el teorema 1.2.3 , de dependencia diferenciable, se deduce que g es continuamente diferenciable sobre V, esto es, $g \in \mathscr{C}^1(V)$. Además, como $f(g(y)) = y$, $y \in V$, entonces la regla de la cadena nos dice que:

$f'(g(y))g'(y) = I$ y por tanto se concluye la afirmación de que

$$g'(y) = [f'(g(y)]^{-1}. \blacksquare$$

OBSERVACION: El hecho de que $g'(y) = [f'(g(y)]^{-1}$ también es una consecuencia del teorema 1.2.3. Para verlo dicho teorema nos dice

que $g'(y) = \left[I - \frac{\partial T}{\partial x}(g(y),y)\right]^{-1}\frac{\partial T}{\partial y}(g(y),y)$ $\quad y \in V. \quad \Rightarrow$

$$g'(y) = \left[I - \left(I - [f'(a)]^{-1}f'(g(y))\right)\right]^{-1}[f'(a)]^{-1}$$

$$= \left([f'(a)]^{-1}f'(g(y))\right)^{-1}[f'(a)]^{-1}$$

$$= [f'(g(y))]^{-1}([f'(a)]^{-1})^{-1}[f'(a)]^{-1} = [f'(g(y)]^{-1}.$$

Antes de dar ejemplos ilustrativos haremos ver que el teorema de la función inversa, bajo un cambio apropiado, sigue siendo válido en los espacios de Banach. Para esto obsérvese que la demostración presentada aquí es claramente independiente de la dimensión excepto, en la parte inicial donde se hace uso del hecho de que todas las transformaciones lineales en espacios finito dimensionales son continuas y que por lo tanto son acotadas. El resto esta basado en que podemos sumar y comparar elementos en los $\mathbb{R}^n$, pues estos, Poseen la estructura de un espacio métrico vectorial. Dicha estructura es una caracteristica propia de los espacios de Banach en general por lo que es de esperarse que el cambio apropiado tenga que ver con la transformación lineal **$f'(a)$**. Esta es que $f'(a)$ sea un isomorfismo, es decir, que también tenga inversa continua. En estos términos el teorema 2.1.1 se reformula como sigue:

Sean $\mathscr{X}$, $\mathscr{Y}$ espacios de Banach y $\mathbf{D}$ un abierto en $\mathscr{X}$. Consideremos $f:\mathbf{D} \to \mathscr{Y}$ con $f \in \mathscr{C}^1(\mathbf{D})$ y supongamos que para $\mathbf{a} \in \mathbf{D}$, con $\mathbf{b} = f(a)$, se tenga que $f'(a)$ es un isomorfismo de $\mathscr{X}$ en $\mathscr{Y}$. Entonces se cumple lo siguiente:

i) Existen conjuntos abiertos $\mathbf{U} \subset \mathbf{D}$, $\mathbf{V} \subset \mathscr{Y}$ de manera que $\mathbf{a} \in \mathbf{U}$, $\mathbf{b} \in \mathbf{V}$ y f es uno a uno y sobre de $\mathbf{U}$ en $\mathbf{V}$.

ii) Si g es la inversa de f entonces $g \in \mathscr{C}^1(\mathbf{V})$ y $g'(y) = [f'(g(y))]^{-1}$.

OBSERVACION: Dado que se supone que $f'(a)$ es un isomorfismo de $\mathscr{X}$ en $\mathscr{Y}$ entonces obviamente $\mathscr{X}$, $\mathscr{Y}$ son isomorfos y por tanto esto implica que deben de ser de la misma dimensión. En particular tenemos $\mathscr{X} = \mathbb{R}^n$, $\mathscr{Y} = \mathbb{R}^n$ el cual acabamos de ver.

2.1. EJEMPLOS

Este primer ejemplo ilustra como la invertibilidad de f' no implica necesariamente que f sea globalmente uno a uno. Aquí se establece pues el caracter local de teorema.

2.1.1. Sea $f:\mathbb{R}^2 \to \mathbb{R}^2$ dado por $f(x,y) = (e^x\cos y,\ e^x\operatorname{sen} y)$. Así

$$f'(x,y) = \begin{bmatrix} e^x\cos y & -e^x\operatorname{sen} y \\ e^x\operatorname{sen} y & e^x\cos y \end{bmatrix}$$ por lo que se ve que en este caso

$\operatorname{Det}\ [f'(x,y)] = e^{2x}\cos^2 y + e^{2x}\operatorname{sen}^2 y = e^{2x}(\cos^2 y + \operatorname{sen}^2 y) = e^{2x}$. Por lo tanto $\operatorname{Det}\ [f'(x,y)] \neq 0$ en todo punto de $\mathbb{R}^2$ y el teorema de la función inversa dice que en dichos puntos f admite una inversa local.

Ahora $\forall\ (x,y) \in \mathbb{R}^2$ de tiene que

$f(x,y + 2\pi) = (e^x\cos(y + 2\pi),\ e^x\operatorname{sen}(y + 2\pi)) = (e^x\cos y,\ e^x\operatorname{sen} y)$
$= f(x,y)$. Esto implica que f no es uno a uno en todo punto de $\mathbb{R}^2$.

El siguiente ejemplo nos muestra que puede existir la inversa de f aún cuando f'(a) no sea invertible y que pero, en estos casos, g no puede ser diferenciable en $b = f(a)$.

2.1.2. Sea $f:\mathbb{R} \to \mathbb{R}$ con $y = f(x) = x^3$. Observemos que $f(0)=0$ y que $f'(0)=0$. Sin embargo la inversa g existe y en este caso esta dada por $x = g(y) = y^{1/3}$. También se ve que g no es diferenciable en $f(0)=0$ pues , $g'(y) = \frac{1}{3\,y^{2/3}}$.

De un modo general, si f'(a) no es invertible y si g es la inversa de f entonces, g no puede ser diferenciable en $b = f(a)$ pues si esto ocurriese la regla de la cadena nos diría que

$\underbrace{f'(g(b))g'(b)}_{\text{producto matricial}} = I$, pero esto implica que $\operatorname{Det} f'(g(b)) \operatorname{Det} g'(b) = 1$

lo cual es un absurdo pues dado que $f'(a)$ no es invertible se tiene que $\operatorname{Det} f'(g(b)) = \operatorname{Det} f'(a) = 0$.

En el ejemplo siguiente se mostrará que la hipótesis de continuidad de la derivada es imprescindible en el teorema de la función inversa.

2.1.3. Sea $f:\mathbb{R} \to \mathbb{R}$ definida por $f(x) = \begin{cases} x + x^2 \operatorname{sen}\frac{1}{x}, & x \neq 0 \\ 0, & x = 0 \end{cases}$.

Entonces se cumplen las siguientes tres afirmaciones:

i) $f'(0)$ es invertible

ii) f' no es continua en $x=0$

iii) f no tiene inversa alrededor de cero.

i) $$f'(0) = \lim_{x\to 0} \frac{f(x) - f(0)}{x - 0} = \lim_{x\to 0} \frac{x - x^2 \operatorname{sen} {}^1\!/x}{x}$$

$$= \lim_{x\to 0} (1 + x\operatorname{sen} {}^1\!/x) = 1 + \lim_{x\to 0} x\operatorname{sen} {}^1\!/x = 1 + 0 = 0.$$ así

$f'(0) = 1$ y por tanto es invertible.

ii) Si $x \neq 0$ entonces $f'(x) = 1 + 2\operatorname{sen} {}^1\!/x + x^2(-{}^1\!/x^2 \cos {}^1\!/x)$

$= 1 + 2\operatorname{sen} {}^1\!/x - \cos {}^1\!/x$ .De esta manera f' esta definida para toda x y esta dada por:

$$f'(x) = \begin{cases} 1 + 2x\operatorname{sen}\frac{1}{x} - \cos\frac{1}{x}, & x \neq 0 \\ 1, & x = 0 \end{cases}$$

si f' fuera continua en $x=0$, se tendría que $\lim_{x\to 0} f'(x) = f'(0) = 1$,

pero $\lim_{x\to 0} f'(x) = \lim_{x\to 0}(1 + 2x\operatorname{sen} {}^1\!/x - \cos {}^1\!/x) = 1 + 0 - \lim_{x\to 0} \cos {}^1\!/x$

el cual no existe pues $\lim_{x\to 0} \cos {}^{1}/x$ no esta definido. por lo tantof'(x) no es continua en x=0.

iii) Para ver que alrededor de cero f no es invertible probaremos que en cualquier vecindad V de él, existen puntos $x\neq 0$ para los cuales $f'(x)=0$ y $f''(x) < 0$. Ambas cosas señalan que en tales puntos f toma un máximo local ó que cerca de ellos la gráfica de f es cóncava hacia abajo y que por tanto obliga a que f no sea uno a uno en V.

Primero tenemos que para $x\neq 0$ la derivada de f es:

$$f'(x) = 1+2x\operatorname{sen}\frac{1}{x} - \cos\frac{1}{x}$$

Sean los puntos de la forma $x = \frac{1}{2n\pi}$, n = 1, 2, ..., de esta forma $f'\left(\frac{1}{2n\pi}\right) = 1 + \frac{1}{n\pi}\operatorname{sen}2n\pi - \cos 2n\pi = 1 + 0 - 1 = 0$. Si V es cualquier vecindad de cero entonces podemos elegir una N tal que $x = \frac{1}{2n\pi} \in V$ si $n > N$. Ahora la segunda derivada de f es:

$f''(x) = 2\operatorname{sen}\frac{1}{x} - \frac{2}{x}\cos\frac{1}{x} + \frac{1}{x^2}\operatorname{sen}\frac{1}{x}$ y concluimos verificando que :

$$f''\left(\frac{1}{2n\pi}\right) = 2\operatorname{sen}2n\pi - 4n\pi\cos 2n\pi + 4n^2\pi^2\operatorname{sen}2n\pi = 0 - 4n\pi + 0$$

$$= -4n\pi < 0 \qquad n = 1, 2, \ldots$$

2.1.4. Para finalizar aquí, damos como ejemplo una aplicación sencilla a sistemas de no lineales. Para esto consideremos el siguiente:

$$x_1^3 + x_2^3 = y_1$$
$$x_1^2 + x_1^2 = y_2$$

Supongamos que se desea investigar si en tal sistema es posible resolver para x_1, x_2 en términos de y_1, y_2 alrededor del punto $(x_1,x_2)=(1,2)$. Para esto definamos $f:\mathbb{R}^2 \to \mathbb{R}^2$ dado por el mapeo $f(x_1,x_2)=(x_1^3 + x_2^3,\ x_1^2 + x_2^2)$, así, $f(1,2)=(9,5)$ y:

$$f'(1,2) = \begin{bmatrix} 3 & 12 \\ 2 & 4 \end{bmatrix}.$$

Puesto que $\mathrm{Det}[f'(1,2)] = 12 - 24 = 12 \neq 0$ entonces $f'(1,2)$ es invertible por lo que el teorema de la función inversa garantiza que en vecindades pequeñas del punto $(1,2)$, x_1, x_2 pueden ser resueltas en términos de y_1, y_2 siempre y cuando (y_1, y_2) esté próximo al punto $(9,5)$.

2. EL TEOREMA DE LA FUNCION IMPLICITA

En muchas ocasiones, una ecuación entre dos variables determina a una como función de la otra e inclusive es posible darla en forma explícita, por ejemplo, $x - 2y = 0$ define a y como función de x, esta es, $y = x/2$. También a x como función de y dado por $x=2y$. Otras son mucho menos evidentes como las relaciones:

$$xe^{y} + x^{2}y + 4 = 0 \ , \ x^{2}y + xy^{2} - 1 = 0$$

En general si pensamos que una ecuación esta dada por $f(x,y) = 0$ nos podemos preguntar:

Es posible definir una función $y = h(x)$ de modo que $f(x,h(x))=0$? ó de manera analoga; una función $x = g(y)$ tal que $f(g(y),y) = 0$? Esto es, en realidad, lo que se hace cuando uno intenta "*despejar*" en la ecuación $f(x,y)=0$ para una variable en términos de la otra.

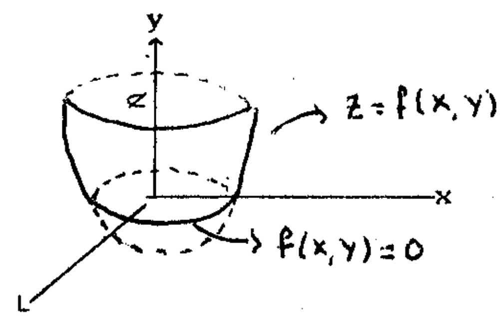
En este caso a las funciones $y = h(x)$, $x = g(y)$ cuando existen, se dicen que están definidas *implicitamente*, ó que son *funciones implicitas*, en las ecuaciones $f(x,h(x))$ y $f(g(y),y)$ resp.

Deseamos ahora, dar un criterio sencillo que nos permita asegurar la existencia de funciones implícitas. Para llevar esto acabo, nuestro estudio se centrará, particularmente, en determinar cuando la ecuación $f(x,y)=0$ puede ser resuelta para x en términos de y. El otro caso se seguirá analogamente.

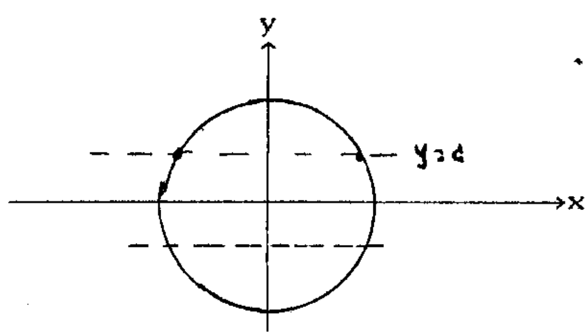
Para simplificar nuestro análisis, supondremos que x, y y $f(x,y)$ son reales de modo que $f(x,y)=0$ representa la relación dada por el corte de la superficie $z=f(x,y)$ sobre el plano xy. Iniciemos esto mediante el caso mas sencillo, es decir, Consideremos el plano dado por $f(x,y) = ax + by + c$ y estudiemos la relación dada por la ecuación $ax + by + c = 0$. Esta representa una recta en el plano xy y podemos despejar x en términos de y siempre y cuando $a \neq 0$. Explícitamente tenemos $x = g(y) = -a^{-1}by + a^{-1}c$. Nótese que la pendiente m de dicha recta es $m = a^{-1}b$.

OBSERVACION: Si $a=0$, $(b \neq 0)$, el plano $f(x,y) = ax + by + c$ resulta ser paralelo a xz y en este caso es imposible de tener a x como función de y en el plano xy.

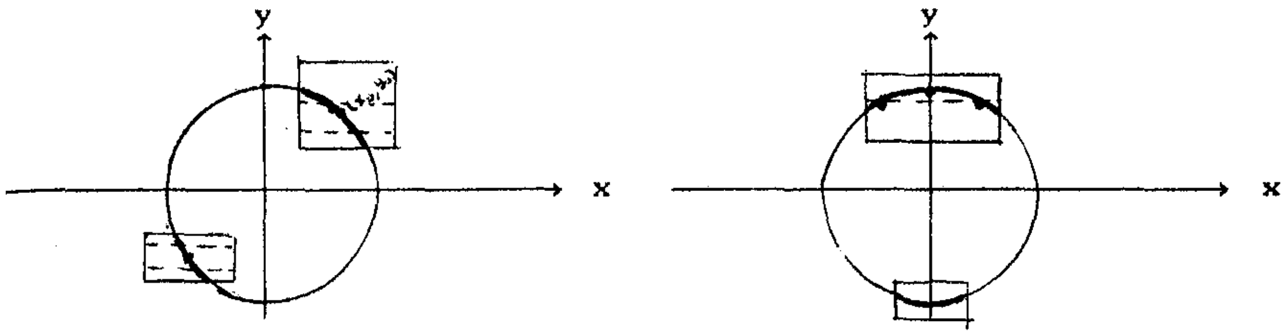
Un ejemplo menos simple, pero que es típico en esta clase de problemas, es la circunferencia unitaria dada por la ecuación $z = f(x,y) = x^2 + y^2 - 1 = 0$.



Ahora bien, para que el conjunto $f(x,y)=0$ represente a x como función de y debe de pasar que tada recta $y=c$, en el plano xy, atraviese a este conjunto una sola vez. Evidentemente no ocurre así en este caso el cual se ve de la siguiente manera.



Sin embargo, si se ha elegido un punto (x_o, y_o) tal que $f(x_o, y_o)=0$, es factible pensar que en una pequeña vecindad de dicho punto la condición de que toda recta $y=c$ atraviese a la gráfica una sola vez se satisfaga excepto para los puntos (0,+1) y (0,-1) como se muestra gráficamente a continuación.



i) *ii)*

Las porciones que quedan dentro de las vecindades que se indican en *i)*, representan ejemplos de funciones implícitas para la relación $x^2 + y^2 - 1 = 0$. En la parte superior, $x > 0$, tenemos que $x = g_1(y) = \sqrt{1 - y^2}$ mientras que para la parte inferior , $x < 0$, es $x = g_2(y) = -\sqrt{1 - y^2}$. Nótese que g_1 y g_2 son diferenciables.

Todo lo contrario sucede en la figura *ii)* donde es imposible tener a x como función de y alrededor de los puntos (0,+1) y (0,-1) respectivamente. Sin embargo, estos son precisamente en los cuales $\frac{\partial f}{\partial x} = 0$ por lo que podemos sospechar que una de las condiciones a imponer, es que se satisfaga que $\frac{\partial f}{\partial x}(x_o,y_o) \neq 0$. De hecho, adicionalmente la hipótesis de continuidad de dicha parcial en (x_o,y_o), garantiza la existencia única y local de funciones implícitas para para la relación $f(x,y)=0$. Lo anteriormente dicho se puede justificar en la siguiente forma. Si $z=f(x,y)$ es diferenciable entonces podemos aproximar linalmente a la superficie alrededor de un punto (x_o,y_o) mediante el plano tangente dado por:

$$P_t(x,y) = \frac{\partial f}{\partial x}(x_o,y_o)(x - x_o) + \frac{\partial f}{\partial y}(x_o,y_o)(y - y_o) + f(x_o,y_o)$$

y la ecuación $P_t(x,y)=0$ puede resolverse para x en términos de y simpre y cuando $\frac{\partial f}{\partial x}(x_o,y_o) \neq 0$. Ahora, si $f(x_o,y_o)=0$ entonces puesto que $z = f(x,y)$ se comporta como P_t cerca de dicho punto es de esperarse que, al menos localmente, la ecuación $f(x,y)=0$ también pueda "resolverse" para x en términos de y ó, en otras palabras, que el conjunto $f(x,y)=0$ representa, cerca de (x_o,y_o), la gráfica de alguna función de la forma $x=g(y)$. Pero si además, imponemos que la parcial, $\frac{\partial f}{\partial x}$, sea continua en el punto (x_o,y_o) entonces existirá alguna vecindad V del mismo en el cual se tenga que $\frac{\partial f}{\partial x} \neq 0$. Esto nos dice que los planos tangentes a la superficie $z=f(x,y)$ no son paralelos a xz, lo que asegura que localmente la gráfica $f(x,y)=0$ no alcanza a dar "vuelta" como sucede alrededor de los puntos $(0,+1)$ y $(0,-1)$ de ejemplo que acabamos de ver.

Ya hemos visto las condiciones a imponer para determinar la existencia de funciones implícitas en una relación de la forma $f(x,y)=0$, en donde x, y y $f(x,y)$ son reales. Sin embargo, podemos sacar mas información respecto a tales funciones si admitimos en este momento que las condiciones antes mencionadas garantizan su existencia local. En este sentido digamos que ya se tiene definida una función $x=g(y)$ por medio de la relación implícita $f(g(y),y)=0$; ahora, la ecuación $P_t(x,y)=0$ representa la tangente a la gráfica del conjunto $f(x,y)=0$ y cuya pendiente esta dada por la fórmula

$$m = \left[\frac{\partial f}{\partial x}(x,y)\right]^{-1} \frac{\partial f}{\partial x}(x,y)$$

puesto que $f(x,y)=0 \Rightarrow x=g(y)$ y si suponemos también que g es diferenciable entonces $g'(y)$ es la pendiente de la tangente en el punto $(g(y),y)$ sobre la gráfica de g y por lo cual se deberá tener que $g'(y) = \left[\frac{\partial f}{\partial x}(g(y),y)\right]^{-1} \frac{\partial f}{\partial y}(g(y),y)$.

Introduciremos ahora el problema a mas dimensiones mediante la discusión que sigue.

Ya hemos visto, en particular, que la ecuación lineal en dos incógnitas $ax + by = 0$ se resuelve para x en términos de y siempre y cuando a sea invertible, es decir, $a \neq 0$. En general el sistema lineal de n + m incógnitas con n ecuaciones dado por

$$
(*) \quad \begin{array}{l}
a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n + b_{11}y_1 + b_{12}y_2 + \dots + b_{1m}y_m = 0 \\
a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n + b_{21}y_1 + b_{22}y_2 + \dots + b_{2m}y_m = 0 \\
\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\
a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n + b_{n1}y_1 + b_{n2}y_2 + \dots + b_{nm}y_m = 0
\end{array}
$$

permite resolver para las x_i's en términos de las y_j's solo si la la matriz

$$
A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}
$$
es invertible, esto es , que $\operatorname{Det}(A) \neq 0$.

Nótese que este sistema, corresponde a una transformación lineal $T:\mathbb{R}^{n+m} \to \mathbb{R}^n$ y que a su vez, es posible descomponerla en dos transformaciones lineales $T_1:\mathbb{R}^n \to \mathbb{R}^n$, $T_2:\mathbb{R}^m \to \mathbb{R}^n$ con $T_1x = Ax$, $Ty = By$ donde:

$$
x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad y \quad B = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1m} \\ b_{21} & b_{22} & \dots & b_{2m} \\ \dots & \dots & \dots & \dots \\ b_{n1} & b_{n2} & \dots & b_{nm} \end{bmatrix}.
$$
Así, (*) puede ser escrita como $T_1x + T_2y = 0$ y si A es invertible, T_1^{-1} existe y por tanto $x = -T_1^{-1} \circ T_2y$. En términos de las coordenadas esto esto nos dice que las primeras n variables $x_1, x_2, \dots, x_n$ son funciones de $y_1, y_2, \dots, y_m$. Esta es precisamente la versión

lineal del teorema de la función implícita.

Una generalización del sistema (*) esta dado por:

$$
(**)\qquad
\begin{array}{l}
f_1(x_1,x_2,\dots,x_n,\, y_1,y_2,\dots,y_m) = 0\\
f_2(x_1,x_2,\dots,x_n,\, y_1,y_2,\dots,y_m) = 0\\
\dots\dots\dots\dots\dots\dots\dots\dots\dots\\
f_n(x_1,x_2,\dots,x_n,\, y_1,y_2,\dots,y_m) = 0
\end{array}
$$

donde $f_i:D \subseteq \mathbb{R}^{n+m} \to \mathbb{R}$, $f_i \in \mathscr{C}^1(D)$, D abierto, $i=1,2,\dots n$. Cuando las f_i son lineales, el criterio que nos permite pasar de unas variables a otras radica en la posibilidad de invertir una matriz. En este caso ¿ cual es el criterio que garantice que las variables x_1, $x_2,\dots,$ x_n puedan ser resueltas en función de y_1, $y_2,\dots,$ y_m?. Este problema se puede ver de la siguiente forma;

si definimos $f:\mathbb{R}^{n+m} \to \mathbb{R}^n$ cuya i-ésima componente es f_i entonces el sistema (**) puede escribirse como:

$$f(x,y)=0 \quad \text{donde} \quad x = (x_1,x_2,\dots,x_n)\ ,\ y = (y_1,y_2,\dots,y_m)$$

Además, si $x=g(y)$ es tal que $f(g(y),y)=0$ se ve que en términos de las coordenadas, esto implica las siguientes igualdades:

$x_1=g_1(y_1,y_2,\dots,y_m)$, $\qquad x_2=g_2(y_1,y_2,\dots,y_m),\dots,\qquad x_n=g_n(y_1,y_2,\dots,y_m)$.

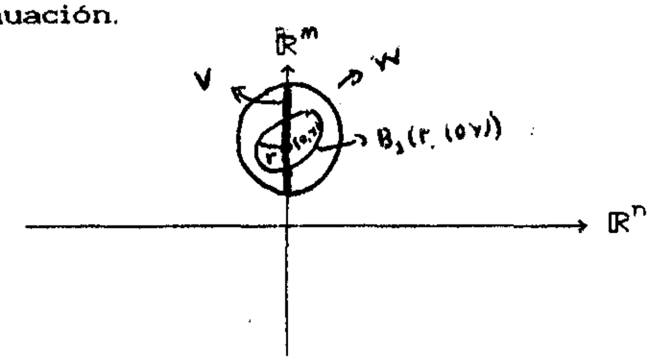
De aquí que la respuesta a la pregunta planteada equivale a dar condiciones que permitan asegurar la existencia de la función implícita en la ecuación $f(x,y)=0$. Esto es precisamente lo que asegura el teorema de la función implícita en su forma mas general y para lo cual exige que la función sea de clase $\mathscr{C}^1$ alrededor de un punto y que la matriz $\frac{\partial f}{\partial x}$ sea invertible de una manera análoga a como se ha observado en el caso lineal.

Antes de establecer el teorema daremos dos resultados que serán utilizados en la demostración.

TEOREMA 2.2.1. Considérese el espacio euclidiano $\mathbb{R}^{n+m} = \mathbb{R}^n \times \mathbb{R}^m$ con la topologia usual. Si W es un abierto en $\mathbb{R}^{n+m}$ entonces el conjunto $V = \left\{y \in \mathbb{R}^m : (0,y) \in W\right\}$ es un abierto en $\mathbb{R}^m$.

DEMOSTRACION: Para ver esto debemos probar que para cada $y \in V$ existe una bola abierta en $\mathbb{R}^m$ de radio k y centro en y, $B(k,y)$, tal que $B(k,y) \subset V$.

Si $V = \emptyset$, V es abierto por definición. Supongamos que $V \neq \emptyset$ y sea $y \in V$. Entonces $(0,y) \in W$ y siendo W abierto, existe una bola abierta $B_1(r,(0,y))$ completamente contenida en W, es decir, que $B_1(r,(0,y)) = \left\{(s,t) \in \mathbb{R}^{n+m} : d((s,t),(0,y)) < r\right\} \subset W$. Esto se muestra a continuación.



Ahora sea $\left\{t \in \mathbb{R}^m : (0,t) \in B_1(r,(0,y))\right\} \subset V$. Para esas t's se sigue cumpliendo que $d((0,t),(0,y)) < r$ y además $d(t,y) = d((0,t),(0,y)) < r$ por lo que eligiendo $k=r$ tenemos la conclusión de que $B(k,y) = \{t \in \mathbb{R}^m : d(t,y) < k\} \subset V$. ∎

El segundo resultado es sobre transformaciones lineales y es:

TEOREMA 2.2.2. Sean $\mathcal{X}$, $\mathcal{Y}$ espacios de Banach y consideremos el producto $\mathcal{Z} = \mathcal{X} \times \mathcal{Y}$. Sea $T_1 \in \mathcal{L}(\mathcal{X},\mathcal{X})$, $T_2 \in \mathcal{L}(\mathcal{Y},\mathcal{X})$ y definamos el mapeo $T:\mathcal{Z} \to \mathcal{Z}$ dado por $T(x,y) = (T_1x + T_2y,\ y)$. Si T_1 es invertible, T es invertible. El recíproco también es cierto.

DEMOSTRACION: Supongamos que T_1 es invertible, debemos mostrar que la ecuación $T(x,y) = (h,k)$ posee solución única. Para esto tenemos que la ecuación anterior es equivalente al sistema:

$$\begin{aligned} T_1x + T_2y &= h \\ y &= k \end{aligned} \qquad \text{o} \qquad \begin{aligned} T_1x + T_2k &= h \\ y &= k \end{aligned}$$

como T_1 es invertible se tiene que para cada (h,k) existe una única solución dada por $y = k$ y $x = T^{-1}(h - T_2k)$. De aquí que T sea uno a uno y sobre, es decir, invertible y cuya inversa es:

$$T^{-1}(h,k) = (x,y) = (T_1^{-1}(h - T_2k), k).$$

Finalmente tenemos:

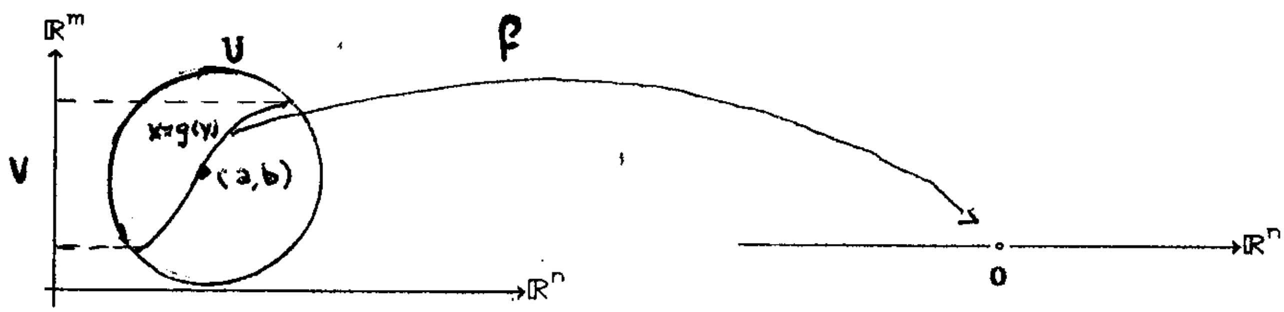
<u>TEOREMA</u> 2.2.3. **(DE LA FUNCION IMPLICITA)**

Sea $\mathbf{D}$ un subconjunto abierto de $\mathbb{R}^n \times \mathbb{R}^m$, $\mathbf{f}:\mathbf{D} \to \mathbb{R}^n$ con $\mathbf{f} \in \mathscr{C}^1(\mathbf{D})$; consideremos la ecuación $\mathbf{f}(x,y)=\mathbf{0}$ y sea $(a,b) \in \mathbf{D}$ de modo que $\mathbf{f}(a,b)=\mathbf{0}$. Supongamos ahora que $\frac{\partial \mathbf{f}}{\partial x}(a,b)$ es invertible. Entonces:

i) Existen conjuntos abiertos $\mathbf{U} \subset \mathbf{D}$, $\mathbf{V} \subset \mathbb{R}^m$ con $(a,b) \in \mathbf{U}$, $b \in \mathbf{V}$ y una única función $g:\mathbf{V} \to \mathbb{R}^n$, con $g \in \mathscr{C}^1(\mathbf{V})$, tal que satisface que para cada $y \in \mathbf{V}$, $(g(y),y) \in \mathbf{U}$, $g(b)=a$ y $\mathbf{f}(g(y),y)=\mathbf{0}$. Mas aún

ii) $g'(y) = \left[\frac{\partial \mathbf{f}}{\partial x}(g(y),y)\right]^{-1}\frac{\partial \mathbf{f}}{\partial y}(g(y),y)$ para cada $y \in \mathbf{V}$.

OBSERVACION: En otras palabras la parte *i)* nos dice que todas las soluciones de la ecuación $\mathbf{f}(x,y)=\mathbf{0}$ en $\mathbf{U}$ están dadas por la gráfica de g en $\mathbf{U}$ como se ilustra a continuación.



DEMOSTRACION: La conclusión *i)* será extraída mediante la aplicación del teorema de la función inversa a la función auxiliar $F(x,y)=(f(x,y),y)$. La parte *ii)* será una consecuencia inmediata de la regla de la cadena aplicada a la ecuación $f(g(y),y)=0$.

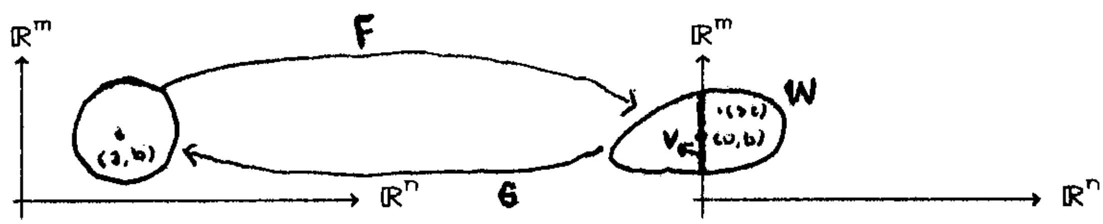
Para empezar tenemos que claramente F es continuamente diferenciable pues sus componentes lo son. Además en el punto (a,b), $F(a,b)=(f(a,b),b)=(0,b)$ y ahora probaremos que $F'(a,b)$ es invertible. Para esto sea (h,k) de modo $(a + h, b + k) \in D$; ahora $F(a + h, b + k) - F(a,b) = (f(a + h, b + k), b + k) - (0,b)$ $= (f(a + h, b + k), k)$. Como f es diferenciable en (a,b) entonces, por definición, $f'(a,b)$ es el mapeo lineal que manda $(h,k) \to f'(a,b)(h,k)$ y que satiface que $f(a + h, b + k)=f(a + h, b + k) - f(a,b)=f'(a,b)(h,k) + O(h,k)$. De esto resulta que:

$F(a + h, b + k) - F(a,b) = (f'(a,b)(h,k) + O(h,k), k)$ $= (f'(a,b)(h,k), k) + (O(h,k), 0)$. De esta manera se ve que $F'(a,b)$ es el mapeo lineal que manda $(h,k) \to (f'(a,b)(h,k), k)$, es decir, $F'(a,b) = (f'(a,b)(h,k), k)$. Por el teorema 0.3.1 tenemos que $f'(a,b)(h,k) = \frac{\partial f}{\partial x}(a,b)h + \frac{\partial f}{\partial y}(a,b)k$ por lo que $F'(a,b)(h,k) = \left(\frac{\partial f}{\partial x}(a,b)h + \frac{\partial f}{\partial y}(a,b)k \;,\; k\right)$ y puesto que la parcial $\frac{\partial f}{\partial x}(a,b)$ es invertible, se deduce del teorema 2.2.2 que $F'(a,b)$ es invertible. Ahora del teorema de la función inversa se deduce que existen conjuntos abiertos $U \subset D$, $W \subset \mathbb{R}^n \times \mathbb{R}^m$ de (a,b) y $F(a,b)=(0,b)$ respectivamente en donde $F:U \to W$ es invertible.

Además si G es la inversa definida por $G(s,t)=(g_1(s,t), g_2(s,t)$, $(s,t) \in W$, $G \in \mathscr{C}^1(W)$ y $F \circ G(s,t)=(s,t) \in W$. Desarrollando la composición tenemos que:

$F \circ G(s,t) = F\left[g_1(s,t),\ g_2(s,t)\right] = \left[f(g_1(s,t),\ g_2(s,t)),\ g_2(s,t)\right] = (s,t) \in W$. Ahora definamos el conjunto $V = \left\{y \in \mathbb{R}^m :\ (0,y) \in W\right\}$; entonces por el teorema 2.2.1 V es abierto pues W lo es.

Para $y \in V$, $(0,y) \in W$ y $\left[f(g_1(0,y),g_2(0,y)),\ g_2(0,y)\right] = (0,y)$. De aquí que $g_2(0,y) = y$ y poniendo $g_1(0,y) = g(y)$ tenemos que $(f(g(y),y),\ y) = (0,y)$.



Dado que G es uno a uno de W sobre U, para cada $y \in V$ existe un único $g(y)$ para el cual $(g(y),y) \in U$ y $f(g(y),y)=0$. Por último, siendo $G \in \mathscr{C}^1(W)$, la ecuación $G(0,y)=G\left[f(g(y),y),\ y)\right]=G\left[F(g(y),y)\right]$ $=(g(y),y)$ que $g \in \mathscr{C}^1(V)$. Esto prueba la parte *i*).

Ahora para ver *ii*) apliquemos la regla de la cadena a la ecuación $f(g(y),y)=0$. De esto resulta que:

$f'(g(y),y) = \frac{\partial f}{\partial x}(g(y),y)\frac{\partial x}{\partial y} + \frac{\partial f}{\partial y}(g(y),y)\frac{\partial y}{\partial y}$ y como $x=g(y)$, tenemos, $\frac{\partial x}{\partial y} = g'(y)$, $\frac{\partial y}{\partial y} = I$. Así:

$$\frac{\partial f}{\partial x}(g(y),y)g'(y) + \frac{\partial f}{\partial y}(g(y),y) = 0$$

y

$$g'(y) = -\left[\frac{\partial f}{\partial x}(g(y),y)\right]^{-1}\frac{\partial f}{\partial y}(g(y),y) \qquad y \in V. \blacksquare$$

Prácticamente el teorema de la función implícita aquí expuesto, ha sido dado como una aplicación del teorema de la función inversa el cual, bajo isomorfismo lineal, es extendido a los espacios de banach en general. Así mismo, dicho teorema sigue siendo válido en tales espacios si pedimos en este caso que $\frac{\partial f}{\partial x}(a,b)$ sea precisamente el isomorfismo lineal. De esta forma, todos los detalles de la demostración se mantienen intactos cuando en vez de $\mathbb{R}^n$, $\mathbb{R}^m$ se ponen, en general, los espacios de Banach $\mathbb{X}$, $\mathbb{Y}$ respectivamente.

2.2. EJEMPLOS

Los ejemplos que a continuación se presentan, ilustran la conclusión básica del teorema de la función implícita que nos permite poder "resolver" para una variable en términos de la otra en una ecuación de la forma $f(x,y) = 0$.

2.2.1. Sea el sistema:

$$x^2 + \frac{1}{2} y^2 + z^3 - z^2 - 3/2 = 0$$
$$x^3 + y^3 - 3y + z + 3 = 0$$

Puede ser resuelto para y, z en términos de x en una vecindad del punto $(-1,1,0)$?

solución. Sea $f:\mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}^2$ definido por: $f(x,y,z)=(x^2 + \frac{1}{2}y^2 + z^3 - z^2 - 3/2,\ x^3 + y^3 - 3y + z + 3)$. Así, $f(-1,1,0)=0$ y podemos ver que $\frac{\partial f}{\partial (y,z)} = \begin{pmatrix} 2y & 3z^2-2z \\ 3y^2-3 & 1 \end{pmatrix}$ la cual es claramente continua.

entonces $\dfrac{\partial f}{\partial(y,z)}(-1,1,0) = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$ que es claramente invertible y el teorema de la función implícita garantiza que es posible resolver para y,z en términos de x en alguna vecindad del punto (-1,1,0).

2.2.2. Sea $f:\mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^2$ dado porla fórmula:

$$f(x_1,x_2,y_1,y_2)=(x_1y_1 + x_2y_1 - 1,\ x_1x_2 - y_1y_2).$$

Sea también los puntos a=(1,0), b=(0,1) de modo que f(a,b)=0. Pongamos ahora $y=(y_1,y_2)$, $x=(x_1,x_2)$; mostraremos que y puede ser resuelta para x cerca de a y denotando $y=\phi(x)$ evaluaremos $\phi'(a)$.

solucion. $\dfrac{\partial f}{\partial y}(a,b) = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ que es invertible y por lo cual podemos expresar y en términos de x cerca del punto a=(1,0). Así tenemos la existencia local de una función ϕ tal que $f(x,\phi(x))=0$ con $\phi(a)=b$. Además sabemos que ϕ' esta dada por la expresión $\phi'(x)=-\left[\dfrac{\partial f}{\partial y}(\phi(x),x)\right]^{-1}\dfrac{\partial f}{\partial x}(\phi(x),x)$. Además podemos ver que:

$$\frac{\partial f}{\partial x}(a,b) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \text{ y } \left[\frac{\partial f}{\partial y}(a,b)\right]^{-1} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}. \text{ Por lo tanto}$$

$$\phi'(a) = \phi'(1.0) = -\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

En terminos de las componentes tenemos las igualdades:

$$\frac{\partial \phi_1}{\partial x_1}(1,0) = 0,\ \frac{\partial \phi_1}{\partial x_2}(1,0) = 1,\ \frac{\partial \phi_2}{\partial x_1}(1,0) = -1,\ \frac{\partial \phi_2}{\partial x_2}(1,0) = 0.$$

2.2.3. La ecuación $xy - z\ln y + e^{xz} - 1 = 0$, puede resolverse para z en términos de x,y alrededor de (0,1,1)?

solución. Pongamos $f(x,y,z) = xy - z\ln y + e^{xz} - 1$. Así, $f(0,1,1)=0$ y con $\frac{\partial f}{\partial z}(x,y,z) = -\ln y + xe^{xz}$,que claramente es continua, tenemos que $\frac{\partial f}{\partial z}(0,1,1) = 0$ por lo que el teorema de la función implícita no garantiza nada.

Podemos hacerlo para y en términos de x,z?

solución. $\frac{\partial f}{\partial y}(x,y,z) = x - \frac{z}{y}$ y $\frac{\partial f}{\partial y}(0,1,1) = -1 \neq 0$ por lo que en este caso si es posible resolver para y en función de x, z alrededor del punto (0,1,1).

CAPITULO 3

EL TEOREMA DEL RANGO

En el algebra lineal se introduce en los espacios finitos dimensionales, el concepto de cambio de coordenadas para un vector que ha sido expresado en dos bases distintas. Esto permite también para las transformaciones lineales de un cierto rango el buscar bases apropiadas en las cuales su matriz adquiera una expresión lo mas sencilla posible. De hecho es cierto que si $T:\mathbb{R}^n \to \mathbb{R}^m$ es una transformación lineal de rango=p y A una representación matricial de T entonces existen matrices invertibles Q, P tal que $QAP^{-1} = \left(\begin{array}{c|c} I_p & 0 \\ \hline 0 & 0 \end{array}\right)$ en donde I_p es la matriz identidad pxp. Las matrices Q, P son llamadas los cambios de coordenadas que me llevan de A a la forma $\left(\begin{array}{c|c} I_p & 0 \\ \hline 0 & 0 \end{array}\right)$. Evidentemente las matrices A y QAP^{-1} se pueden decir que son equivalentes en cuanto son expresiones de una misma transformación lineal. Mas aun entre ellas se define una relación de equivalencia cuyas clases de equivalencia quedan caracterizadas por el rango de la transformación.

Una teoría similar, como se verá en la sección 2, resulta cuando se consideran funciones $f:D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ de clase $\mathcal{C}^1(D)$. Aquí el concepto de equivalencia, definición 3.2.1, también conduce a una relación de equivalencia en el espacio de tales funciones como en el caso lineal, siendo el llamado TEOREMA DEL RANGO el que justifica lo antes afirmado. El teorema del rango establece que la transformación lineal Df es de rango=p en **D**, entonces **f** es localmente equivalente a la lineal:

$$I_p(x_1,x_2,\ldots,x_n) = (x_1,x_2,\ldots,x_p,0,\ldots 0)$$

en el sentido ahora de que existen difeomorfismos **h**, **g** tales que es válida la ecuación $\mathbf{h}\circ\mathbf{f}\circ\mathbf{g}^{-1}(\mathbf{x}) = I_p(\mathbf{x})$.

Este es pues, un resultado mas que refleja el principio básico general de que las funciones de clase $\mathscr{C}^1$ actúan localmente como lo hacen sus derivadas.

1. MATRICES Y TRANSFORMACIONES LINEALES

Sea $B_1 = \{\mathbf{u}_1,\mathbf{u}_2,\ldots,\mathbf{u}_n\}$ una base para $\mathbb{R}^n$ y sea $\mathbf{x} \in \mathbb{R}^n$. Entonces $\mathbf{x} = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \ldots + \alpha_n\mathbf{u}_n$ y al vector dado por $[\mathbf{x}]_{B_1} = (\alpha_1,\alpha_2,\ldots,\alpha_n)$ se le llama el vector coordenada de **x** respecto a la base B_1. Este vector, pudiera decirse, es la etiqueta ó nombre que recibe **x** de parte de B_1. De la misma manera si B_2 es una nueva base para $\mathbb{R}^n$, entonces en esta base el nombre para **x** es el vector coordenada $[\mathbf{x}]_{B_2}$. Geométricamente **x** representa el mismo vector en ambas bases, pero, con nombres distintos.

Cuando expresamos las coordenadas de un vector de una base en términos de la otra decimos que se ha efectuado un cambio de coordenadas. En este caso dicho cambio es de la base B_1 a la base B_2. En realidad siempre es posible encontrar una matriz **P** de n×n que realice dicha operación, es decir, una matriz **P** que satisfaga la ecuación:

$$[\mathbf{x}]_{B_2} = \mathbf{P}[\mathbf{x}]_{B_1}$$

en donde a **P** se le llama la matriz de transición de la base B_1 a la base B_2 ó simplemente la matriz del cambio dé coordenadas. Además ésta siempre posee la inversa que es la matriz recíproca de coordenadas, esto es:

$$P^{-1}[x]_{B_2} = [x]_{B_1} \ .$$

Ahora sea una trasformación lineal $T:\mathbb{R}^n \to \mathbb{R}^m$ y consideremos las bases B_1 , B_2 de $\mathbb{R}^n$ y $\mathbb{R}^m$ respectivamente. Podemos asociar a **T** una matriz **B** de m×n de forma que **B** sea la matriz de la transformación **T** con respecto a B_1 , como base para el dominio y B_2 como base del contradominio. Si por otra parte, **A** es la representación matricial de **T** en las bases canónicas, de $\mathbb{R}^n$ y $\mathbb{R}^m$ respectivamente entonces **A** y **B** están relacionadas por el siguiente teorema:

<u>TEOREMA</u> 3.1.1. Sea **P** la matriz del cambio de coordenadas de la base canónica en $\mathbb{R}^n$ a B_1 y **Q** la matriz del cambio de coordenadas de la base canónica en $\mathbb{R}^m$ a B_2. Si **A** es, como anteriormente. la matriz de una transformación **T** entonces se cumple que:

$$QAP^{-1} = B$$

donde **B** es la representación matricial de **T** en las bases B_1 y B_2.
(ver libro 11 cap. 7)

Lo anterior da lugar la siguiente definición.

DEF.3.1.1. Se dice que una matriz **A** de m×n es equivalente a una matriz **B** del mismo tamaño, escribimos $A \approx B$, si existe una matriz cuadrada **Q** invertible de m×m y una matriz cuadrada **P** invertible de n×n tal que $QAP^{-1} = B$.

Es fácil deducir que si $A \approx B$ entonces $B \approx A$ por lo que suele decirse simplementes matrices equivalentes. Además, en vista del teorema 3.1.1, es inmediato que si **A** y **B** son representaciones matriciales de una misma transformación lineal $T:\mathbb{R}^n \to \mathbb{R}^m$ entonces

A y **B** son equivalentes. También el inverso es cierto, es decir, si **A** y **B** son matrices equivalentes, éstas representan a una misma transformación lineal en dos bases distintas. Algo que para nosotros será importante es que éste concepto de equivalencia define una *relación de equivalencia* sobre el espacio vectorial de todas las matrices reales de mxn y, por isomorfismo, sobre el espacio vectorial $\mathcal{L}(\mathbb{R}^n,\mathbb{R}^m)$ por lo cual, queda dividido este espacio en clases disjuntas por pares y en donde cada clase esta formada por el conjunto de todas las transformaciones lineales equivalentes a algún elemento típico ó representante de dicha clase. Simbólicamente si para algún conjunto de índices I, el elemento $\mathbf{T}_i$, $i \in$ I, es el representante de alguna clase, entonces la clase de $\mathbf{T}_i$ es denotada por $[\mathbf{T}_i]$ y que esta dada por $[\mathbf{T}_i] = \{\mathbf{T} \in \mathcal{L}(\mathbb{R}^n,\mathbb{R}^m): \mathbf{T} \approx \mathbf{T}_i\}$ en donde $\bigcup_{i\in I}[\mathbf{T}_i] = \mathcal{L}(\mathbb{R}^n,\mathbb{R}^m)$. En realidad, a la hora de trabajar con clases es conveniente tener a un representante en cada una de ellas y como cada clase es independiente del elemento que lo represente es obvio buscar aquel que por sus características sea mas sencillo de tratar.

Ahora estudiaremos como el rango de las transformaciones lineales determinan una partición en n+1 clases de equivalencia en el espacio $\mathcal{L}(\mathbb{R}^n,\mathbb{R}^m)$ y en donde además los representantes se pueden escoger particulartmente simples.

*DEF.*3.1.2. Sea **A** una matriz real de mxn y sea p el número máximo de columnas, (o renglones), linelmente independientes de **A**. LLamamos a p el rango de esta matriz y escribimos ran(A)=p.

La definición anterior puede ser formulada como sigue: el rango de una matriz real **A** es p, si existe una submatriz cuadrada de tamaño $p \times p$ con determinante distinto de cero, esto es, invertible y que además cualquier submatriz de tamaño mayor a $p \times p$ tiene determinante igual a cero.

DEF.3.1.3. Sea $T:\mathbb{R}^n \to \mathbb{R}^m$ una transformación lineal y sea **A** la representación matricial de **T**; definimos el rango de **T** como ran(T)=dim(imagen de T)=ran(**A**).

Estamos ahora en condiciones de enunciar lo que sería la versión lineal del teorema del rango.

TEOREMA 3.1.2. Sea $T:\mathbb{R}^n \to \mathbb{R}^m$ y supongamos que el rango de **T** es p.- Entonces existen bases $B_1 = \{u_1,u_2,\ldots,u_n\}$ y $B_2 = \{v_1,v_2,\ldots,v_m\}$ respecto a las cuales **B**, la matriz de dicha transformación, es de la forma:

$$\mathbf{B} = \left[\begin{array}{c|c} I_p & 0 \\ \hline 0 & 0 \end{array}\right]$$, donde I_p es la matriz identidad de tamaño pxp.

(ver libro 11 capitulo 7).

De aquí se deduce que todas las representaciones matriciales, **A**, de una transformacion lineal $T:\mathbb{R}^n \to \mathbb{R}^m$ cuyo rango es p, son equivalentes a la matriz **B** del teorema anterior. De esta manera queda pues descompuesto el espacio $\mathcal{L}(\mathbb{R}^n,\mathbb{R}^m)$ en n+1 clases de equivalencia a saber; las de rango 0, 1, 2,..., n respectivamente y en donde los representantes para cada clase pueden tomarse mediante las matrices mas simples asociadas a estas transformaciones y que tienen la forma:

$$\left[\begin{array}{c|c} I_p & 0 \\ \hline 0 & 0 \end{array}\right] \quad p = 0, 1, 2, \ldots, n \text{ respectivamente.}$$

2. FUNCIONES EQUIVALENTES

En esta sección generalizamos el estudio del caso lineal precedente considerando ahora funciones de clase $\mathscr{C}^1$.

DEF.3.2.1. Sean ϕ, $\psi:\mathbb{R}^n \to \mathbb{R}^m$ funciones de clase $\mathscr{C}^1$. Diremos que ϕ es equivalente a ψ , $\phi \approx \psi$, si existen (*)difeomorfismos $\mathbf{g}:\mathbb{R}^n \to \mathbb{R}^n$ y $\mathbf{h}:\mathbb{R}^m \to \mathbb{R}^m$ de clase $\mathscr{C}^1$ tales que:

$$\mathbf{h}\circ\phi\circ\mathbf{g}^{-1}(x) = \psi(x)$$

El concepto de equivalencia en este caso también define una *relación de equivalencia* sobre el espacio vectorial $\mathscr{C}^1(\mathbb{R}^n,\mathbb{R}^m)$ y cuyas consecuencias ya hemos visto con anterioridad. Aquí nos detendremos un momento para verificar que efectivamente $\phi \approx \psi$ es una relación de equivalencia. Para esto tenemos que hacer ver que se satisfacen las siguientes tres propiedades:

i) $\phi \approx \phi$ "reflexibilidad"

ii) $\phi \approx \psi \Rightarrow \psi \approx \phi$ "simetría"

iii) $\phi \approx \psi$ y $\psi \approx \gamma \Rightarrow \phi \approx \gamma$ "transitividad"

i) si tomamos $\mathbf{h}=\mathbf{I}_m$, $\mathbf{g}=\mathbf{I}_n$ entonces $\mathbf{g}^{-1}=\mathbf{I}_n$ y $\mathbf{h}\circ\phi\circ\mathbf{g}^{-1}=\mathbf{I}_m\circ\phi\circ\mathbf{I}_n=\phi$ por lo que $\phi \approx \phi$.

ii) puesto que $\phi \approx \psi$, existen difeomorfismos $\mathbf{h}$, $\mathbf{g}$ de clase $\mathscr{C}^1$ tales que $\mathbf{h}\circ\phi\circ\mathbf{g}^{-1}=\psi$. Pero esto implica que $\mathbf{h}^{-1}\circ\psi\circ\mathbf{g}(x)=\phi(x)$ y poniendo $\mathbf{u}=\mathbf{h}^{-1}$, $\mathbf{v}=\mathbf{g}^{-1}$ tenemos que $\mathbf{v}^{-1}=\mathbf{g}$ por lo que podemos escribir que $\mathbf{u}\circ\psi\circ\mathbf{v}^{-1}(x)=\phi(x)$. Por lo tanto $\psi \approx \phi$.

iii) de igual forma $\phi \approx \psi \Rightarrow \mathbf{h}\circ\phi\circ\mathbf{g}^{-1}=\psi$ y $\psi \approx \gamma \Rightarrow \mathbf{p}\circ\psi\circ\mathbf{q}^{-1}=\gamma$. sustituyendo ψ de la primera implicación a la segunda tenemos que $\mathbf{p}\circ(\mathbf{h}\circ\phi\circ\mathbf{g}^{-1})\circ\mathbf{q}^{-1}=\gamma \Rightarrow (\mathbf{p}\circ\mathbf{h})\circ\phi\circ(\mathbf{g}^{-1}\circ\mathbf{q}^{-1})=\gamma$. Si definimos ahora $\mathbf{u}=\mathbf{p}\circ\mathbf{h}$,

(*)funciones diferenciables invertibles con inversa diferenciable.

$v=q\circ g$ tenemos , puesto que $g^{-1}\circ q^{-1}=(q\circ g)^{-1}$, la condlusión de que $u\circ\phi\circ v^{-1}(x)=\gamma(x)$. por lo tanto $\phi \approx \gamma$.

Concluiremos este capítulo presentando el resultado que es la generalización del teorema 3.1.2. Dicho resultado expresa que las funciones $f:D \subseteq \mathbb{R}^n \to \mathbb{R}^m$, con $f \in \mathscr{C}^1(D)$ y cuya diferencial Df es de rango constante p en D, se comportan localmente equivalentes a la función lineal I_p que representa la proyección en las primeras p coordenadas en el espacio $\mathbb{R}^m$. En otras palabras, que existen difeomorfismos locales h,g de clase $\mathscr{C}^1$ tales que:

$$h\circ f\circ g^{-1}(x)=I_p(x)=(x_1,x_2,\ldots,x_p,0,\ldots,0) \in \mathbb{R}^m.$$

Obsérvese que cuando f es lineal la matriz $A=Df$ es precisamente la matriz asociada a f en las bases canónicas de $\mathbb{R}^n$, $\mathbb{R}^m$ respectivamente y en donde aquí los difeomorfismos g, h corresponden a las transformaciones P, Q del teorema 3.1.2.

En otras palabras, el teorema del rango asegura que la dimension de la imagen de un abierto bajo una función continuamente diferenciable es igual al rango de la diferencial cuando este es constante en el mismo abierto.

Precisando lo anterior Tenemos:

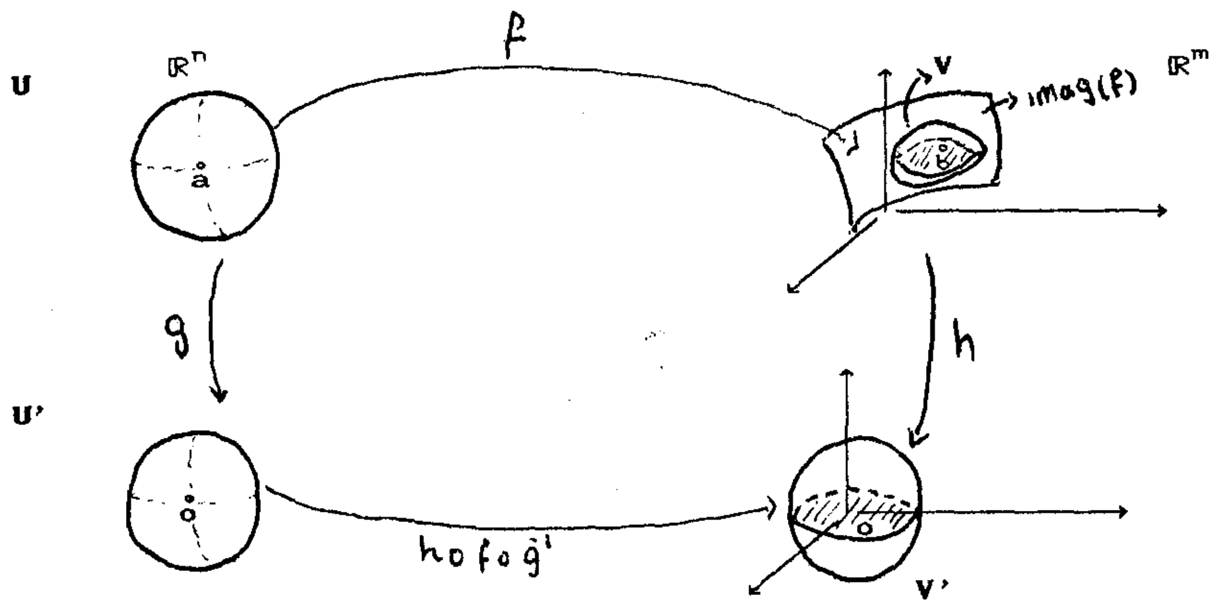
<u>TEOREMA</u> 3.2.1 **(DEL RANGO)**. Sea $D \subseteq \mathbb{R}^n$ abierto y $f:D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ con $f \in \mathscr{C}^1(D)$. Supongamos que Df tiene rango p $\forall\ x \in D$ y para $a \in D$ pongamos $b = f(a) \in \mathbb{R}^m$. Entonces:

i) Existen conjuntos abiertos U de a, U' de $0 \in \mathbb{R}^n$ y un difeomorfismo $g:U \to U'$ de clase $\mathscr{C}^1$.

ii) Existen conjuntos abiertos V de b, V' de $0 \in \mathbb{R}^m$ y un difeomorfismo $h:V \to V'$ de clase $\mathscr{C}^1$.

iii) $h\circ f\circ g^{-1}(x_1,x_2,\ldots,x_n)=(x_1,x_2,\ldots,x_p,0,0,\ldots,0)\in \mathbb{R}^m$.

Geométricamente, esto se ilustra como sigue:



$$(x_1, x_2, \ldots, x_n) \longrightarrow (x_1, x_2, \ldots, x_p, 0, 0, \ldots, 0)$$

DEMOSTRACION: Primero se supondrá que $a=\mathbf{0}$ y $\mathbf{f}(a)=b=\mathbf{0}$,pues, el caso general se obtendrá poniendo $F(x)=f(x + a) - b$. Se construye $\mathbf{g}$ usando las p componentes de $\mathbf{f}$ que hagan que $Df(\mathbf{0})$ sea de rango p y se le aplica a $\mathbf{g}$ el teorema de la función inversa. Se hace $\mathbf{f}\circ\mathbf{g}^{-1}$ y bajo ciertas consideraciones, se construye $\mathbf{h}$ de forma que $\mathbf{h}\circ\mathbf{f}\circ\mathbf{g}^{-1}$ tenga la forma requerida. Finalmente se le aplica a $\mathbf{h}$ el teorema de la función inversa.

Debido a que $Df(x)$ tiene rango p $\forall\ x \in \mathbf{D}$ entonces la matriz

$$Df(\mathbf{0}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}$$

tiene rango p. Por lo tanto sabemos que existe una submatriz de

tamaño pxp invertible y que esto no se cumple para submatrices de mayor tamaño. Además se puede suponer que tal submatriz esta colocada en la esquina de la
parte superior de $Df(0)$. De no ser así, podemos permutar los renglones y columnas para que esto se pueda tener como lo muestra la matriz:

$$\left[\begin{array}{ccc|c} \frac{\partial f_1}{\partial x_1} & \cdots\cdots & \frac{\partial f_1}{\partial x_p} & \\ \cdots & \cdots & \cdots & \text{———} \\ \frac{\partial f_p}{\partial x_1} & \cdots\cdots & \frac{\partial f_p}{\partial x_p} & \\ \hline & \text{———} & & \text{———} \end{array}\right]$$

Después de esto definamos $g:D\subseteq\mathbb{R}^n \to \mathbb{R}^n$ dado por la fórmula $g(x_1,x_2,\ldots,x_n)=(f_1(x),f_2(x),\ldots,f_p(x),x_{p+1},\ldots,x_n)$, en donde x es el vector $x=(x_1,x_2,\ldots x_n)$. Es claro que $g \in \mathscr{C}^1(D)$ y que $g(0)=0$; entonces si calculamos la diferencial de g podemos ver que es

$$Dg(x) = \left[\begin{array}{ccc|cccc} \frac{\partial f_1}{\partial x_1} & \cdots\cdots & \frac{\partial f_1}{\partial x_p} & & & & \\ \cdots & \cdots & \cdots & & \text{———} & & \\ \frac{\partial f_p}{\partial x_1} & \cdots\cdots & \frac{\partial f_p}{\partial x_p} & & & & \\ \hline & 0 & & 1 & & & \\ & & & & 1 & & \\ & & & & & 1 & \\ & & & & & & \ddots \ 1 \end{array}\right]$$

la cual es invertile en $x=0$ y por el teorema de la función inversa existen abiertos U de 0, U' de $g(0)=0$ y un difeomorfismo $g:U \to U'$ de clase $\mathscr{C}^1$ en donde además $g(U)=U'$. Esto prueba la parte *i*).

Pongamos Ahora en las coordenadas de g que:

$(y_1,y_2,\ldots,y_n)=g(x_1,x_2,\ldots,x_n)=(f_1(x),f_2(x),\ldots,f_p(x),x_{p+1},\ldots,x_n)$

y observemos que g^{-1} funciona como la inversa de f en las primeras p coordenadas, pues, g, f son iguales en esas componentes. De aquí que tengamos la siguiente igualdad:

$f\circ g^{-1}(y_1,y_2,\dots,y_n)=(y_1,y_2,\dots,y_p,f_{p+1}\circ g^{-1}(y),\dots,f_m\circ g^{-1}(y))$

$=(y_1,y_2,\dots,y_p,\psi_{p+1}(y),\dots,\psi_m(y))$; en donde $y=(y_1,y_2,\dots,y_n)\in U'$. Si calculamos la derivada de esta composición se puede ver que:

$$D[f\circ g^{-1}(y)] = \left[\begin{array}{c:c} I_p & 0 \\ \hdashline & \begin{matrix} \dfrac{\partial\psi_{p+1}}{\partial y_{p+1}} & \cdots\cdots & \dfrac{\partial\psi_{p+1}}{\partial y_n} \\ \cdots & \cdots & \cdots \\ \dfrac{\partial\psi_m}{\partial y_{p+1}} & \cdots\cdots & \dfrac{\partial\psi_m}{\partial y_n} \end{matrix} \\ \text{———} & \end{array}\right].$$

Ahora, el rango de esta matriz debe de ser p lo cual implica que todas las parciales deben de ser cero ya que si para algún $y\in U'$ al menos una de ellas es distinta de cero se tendría una submatriz de tamaño $(p+1)\times(p+1)$ lo cual sería contrario a nuestra hipótesis. Por lo tanto las funciones $\psi_{p+1},\dots,\psi_m$ no dependen de las variables $y_{p+1},\dots,y_m$ y en base a esto definamos $h:\mathbb{R}^m \to \mathbb{R}^m$ dado por:

$h(z_1,z_2,\dots,z_m)=(z_1,z_2,\dots,z_{p+1} - \psi_{p+1}(z),\dots,z_m - \psi_m(z))$ donde $z=(z_1,z_2,\dots,z_m)$. Así la diferencial de h es,

$$Dh(z) = \left[\begin{array}{c:c} \begin{matrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & \ddots \\ & & & & 1 \end{matrix} & 0 \\ \hdashline \text{———} & \begin{matrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & \ddots \\ & & & & 1 \end{matrix} \end{array}\right]$$

la cual es obviamente invertible y nuevamente por el teorema de la función inversa existen conjuntos abiertos **V**, **V'** de **0** y **h(0)=0** en $\mathbb{R}^m$, respectivamente, tales que la función $\mathbf{h}:\mathbf{V} \to \mathbf{V'}$ es un difeomorfismo de clase $\mathscr{C}^1$, en donde además **h(V)=V'**. Esto prueba la parte *ii*).

Por último , haciendo la composición de $\mathbf{f}\circ\mathbf{g}^{-1}$ con **h** obtenemos

$\mathbf{h}\circ\mathbf{f}\circ\mathbf{g}^{-1}(y_1,y_2,\dots,y_n)=(y_1,y_2,\dots,y_p,\psi_{p+1} - \psi_{p+1},\dots,\psi_m - \psi_m)$

$= (y_1,y_2,\dots,y_p,0,0,\dots,0) \in \mathbb{R}^m,\ \forall\ y \in \mathbf{U'}.$ ∎

CAPITULO 4

TEOREMAS EQUIVALENTES

En este breve capítulo daremos a conocer que los teoremas 2.1.1, 2.2.3, 3.2.1, son equivalentes entre sí, es decir, como un hecho interesante mostraremos que se cumple:

TEOREMA DE LA FUNCION INVERSA $\Longleftrightarrow$ TEOREMA DEL RANGO $\Longleftrightarrow$ TEOREMA DE LA FUNCION IMPLICITA.

El ciclo que obviamente determina esta equivalencia esta dado en el siguiente orden;

I.- **TEOREMA DE LA FUNCION INVERSA $\Rightarrow$ TEOREMA DEL RANGO**
II.- **TEOREMA DEL RANGO $\Rightarrow$ TEOREMA DE LA FUNCION IMPLICITA**
III.- **TEOREMA DE LA FUNCION IMPLICITA $\Rightarrow$ TEOREMA DE LA FUNCION INVERSA.**

DEMOSTRACION I. La prueba de esta implicación esta dada ya en el **teorema 3.2.1.**

DEMOSTRACION II. Las hipótesis del teorema de la función implícita están dadas por las siguientes.

i) $\mathbf{f}:\mathbf{D} \subseteq \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ con $\mathbf{f} \in \mathcal{C}^1(\mathbf{D})$

ii) $(\mathbf{a},\mathbf{b}) \in \mathbf{D}$, $\mathbf{f}(\mathbf{a},\mathbf{b})=\mathbf{0}$ y $\dfrac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{a},\mathbf{b})$ invertible

La conclusión es que existen conjuntos $\mathbf{U}$ de $(\mathbf{a},\mathbf{b})$, $\mathbf{V}$ de $\mathbf{b}$ y una función $\mathbf{g}:\mathbf{V} \to \mathbb{R}^n$ con $\mathbf{g} \in \mathcal{C}^1(\mathbf{V})$ tal que para cada $\mathbf{y} \in \mathbf{V}$, el punto $(\mathbf{g}(\mathbf{y}),\mathbf{y}) \in \mathbf{U}$ y $\mathbf{f}(\mathbf{g}(\mathbf{y}),\mathbf{y})=\mathbf{0}$.

Para ver que lo anterior es consecuencia del teorema del rango, supondremos primeramente que $(a,b)=(0,0)$, el caso general se obtendrá haciendo $F(x,y) = f[(x,y) - (a,b)]$. Como $f \in \mathcal{C}^1(D)$ tenemos que $\frac{\partial f}{\partial x}$ es continua en D y dado que $\frac{\partial f}{\partial x}(0,0)$ es invertible entonces $\frac{\partial f}{\partial x}$ es invertible en alguna vecindad W de $(0,0)$. De esta manera la matriz Df tiene rango máximo n $\forall\ x \in W$ y por el teorema del rango, tenemos la existencia de los difeomorfismos locales de clase $\mathcal{C}^1$ $\psi:U_{(0,0)} \to U'_{(0,0)}$, $\phi:Z_{f(0,0)=0} \to Z'_0$ que están dados en este caso por $\psi(x,y)=(f(x,y),y)$, $\phi(u)=u \in \mathbb{R}^n$ y satisfaciendo que:

$\phi \circ f \circ \psi^{-1}(s,t)=s \in \mathbb{R}^n \quad \forall(s,t) \in U'_{(0,0)}$.

Siendo ϕ la identidad y dado que ψ deja fija la segunda coordenada, ψ^{-1} también lo hace y poniendo $\psi^{-1}(s,t)=(\psi^*(s,t),t)$ tenemos $f \circ \psi^{-1}(s,t)=f(\psi^*(s,t),t)=s \in \mathbb{R}^n \quad \forall(s,t) \in U'_{(0,0)}$. Sea el conjunto $V = \{y \in \mathbb{R}^m: (0,y) \in W\}$ que es abierto pues W lo es; para esas y's tenemos que $f(\psi^*(0,y),y)=0$ en donde $(\psi^*(0,y),y) \in U_{(0,0)}$.

Finalmente se pone $g(y)=\psi^*(0,y)$ y es claro que $g \in \mathcal{C}^1(V)$ por lo que se tiene el resultado deseado. ∎

DEMOSTRACION III. En este caso las hipótesis de teorema de la función inversa són las siguientes.

i) $f:D \subseteq \mathbb{R}^n \to \mathbb{R}^n$ con $f \in \mathcal{C}^1(D)$

ii) $a \in D$, $b=f(a)$ y $f'(a)$ invertible.

La conclusión es que f es un difeomorfismo local de clase $\mathcal{C}^1$. Mas precisamente, existen abiertos U de a, V de b y una función $g:V \to U$ que es la inversa se f y con $g \in \mathcal{C}^1(V)$.

Primero definamos $E = D \times \mathbb{R}^n$ y la función $F : E \to \mathbb{R}^n$ dada por $F(x,y) = f(x) - y$. Así $F \in \mathscr{C}^1(E)$ y $F(a,b) = f(a) - b = 0$; además tenemos que $\frac{\partial F}{\partial x}(a,b) = f'(a)$ y por lo tanto es invertible por hipótesis. De aquí que por el teorema de la función implícita existe un abierto V de b y una función $g : V \to \mathbb{R}^n$, $g \in \mathscr{C}^1(V)$, tal que $F(g(y),y) = 0$ para cada $y \in V$. Pero $F(g(y),y) = f(g(y)) - y = 0$, por lo que resulta que $f(g(y)) = y$. De esta manera tenemos que g es la inversa de f y poniendo $U = f^{-1}(V)$ se tiene el resultado. ∎

CAPITULO 5

APLICACION: DEPENDENCIA FUNCIONAL

De manera análoga al concepto de dependencia lineal que se tiene en espacios vectoriales finito dimensionales y que nos permite expresar un vector como combinacion lineal de otros, en este capítulo daremos un criterio que nos permita saber si en un cierto conjunto de m funciones reales de n variables, $\psi_1, \psi_2, \ldots, \psi_m$ con $m \leq n$, existe una dependencia entre ellas en el sentido de poder expresar funciones de ese conjunto en términos de las restantes. Por ejemplo, en el sistema dado por las funciones $\psi_1(x,y,z) = x^2 + y^2 + z^2$, $\psi_2(x,y,z) = xy + yz + zx$, y $\psi_3(x,y,z) = x + y + z$ tenemos que ψ_1 depende de ψ_2 y de ψ_3 pues:

$$\psi_1(x,y,z) = [\psi_3(x,y,z)]^2 - 2\psi_2(x,y,z).$$

Tales sistemas se dicen dependientes y es el problema que nos ocupará como una aplicación conjunta del teorema del rango e implícita respectivamente.

1. DEPENDENCIA E INDEPENDENCIA

Recordemos que si $\mathscr{X}$ es un espacio vectorial sobre un campo de escalares K, entonces un conjunto finito de vectores $x_1, x_2, \ldots, x_n$ se dicen linealmente dependientes si para $k_1, k_2, \ldots, k_n \in K$, no todos iguales a cero, se cumple que $k_1x_1 + k_2x_2 + \ldots + k_nx_n = 0$. Si esta ecuación es válida únicamente para $k_1 = k_2 = \ldots = k_n = 0$ se dicen en este caso linealmente independientes.

Observemos que si por ejemplo $k_n \neq 0$ entonces podemos poner que

$x_n = -\frac{k_1}{k_n} x_1 - \frac{k_2}{k_n} x_2 - \dots - \frac{k_{n-1}}{k_n} x_{n-1}$ por lo que de manera equivalentemente se dice que el conjunto $x_1, x_2, \dots, x_n$ es linealmente dependiente si al menos uno de ellos es expresable en términos de los n-1 restantes.

Discutiremos brevemente ahora el concepto de dependencia sobre funcionales lineales para después pasar con el caso general en la siguiente sección.

Supongamos que se ha dado el siguiente sistema de funcionales lineales,

$$(*)\qquad \begin{array}{l} f_1 = a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ f_2 = a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ f_m = a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \end{array} \qquad m \leq n$$

es claro que cada f_i, $i=1,2,\dots,m$, queda completamente determinada por el vector $(a_{i1}, a_{i2}, \dots, a_{in})$ por lo que nos es permitido hablar del concepto de dependedencia lineal, con todas sus propiedades, para el caso de funcionales lineales. La matriz de los coeficientes de este sistema es:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ a_{m1} & a_{m1} & \dots & a_{mn} \end{pmatrix}$$

por lo que si el rango de esta matriz es m, los vectores fila son linealmente independientes y por tanto $f_1, f_2, \dots, f_m$ es un conjunto independiente. En tanto, si el rango es inferior a m, hay una dependencia lineal entre las filas que forman a A y por tanto una dependencia entre $f_1, f_2, \dots, f_m$.

OBSERVACION. Se omite el caso $m > n$ pues obviamente se tiene la dependencia automática entre los elementos del sistema (*).

2. DEPENDENCIA FUNCIONAL

Generalizaremos las ideas anteriores para el caso en que las funcionales lineales f_i son remplazadas por funciones ψ_i que son continuamente diferenciables. Para esto sea **D** un abierto en $\mathbb{R}^n$ y consideremos el sistema de m funciones dado por:

$$(**)\qquad \begin{aligned} y_1 &= \psi_1(x)\\ y_2 &= \psi_2(x)\\ &\cdots\cdots\cdots\\ y_m &= \psi_m(x) \end{aligned}\qquad m \le n$$

en donde $x \in \mathbb{R}^n$, $\psi_i:\mathbf{D} \to \mathbb{R}$ y $\psi_i \in \mathscr{C}^1(\mathbf{D})$. Definimos aquí el concepto de dependencia funcional de la siguiente manera.

DEF.5.2.1. Sea **V** un conjunto abierto de $\mathbb{R}^{m-1}$ y sea una función $\Phi:\mathbf{V} \to \mathbb{R}$ con $\Phi \in \mathscr{C}^1(\mathbf{V})$ tal que $\forall\ x \in \mathbf{D}$ se cumplan las condiciones:

i) $(\psi_1(x),\psi_2(x),\ldots,\psi_{m-1}(x)) \in \mathbf{V}$

ii) $\Phi(\psi_1(x),\psi_2(x),\ldots,\psi_{m-1}(x)) = \psi_m(x)$

entonces la función ψ_m se dice dependiente en **D** de las funciones $\psi_1(x)$, $\psi_2(x)$,..., $\psi_{m-1}(x)$ y al sistema (**) se le llama un sistema dependiente.

En particular si para algún $x_0 \in \mathbf{D}$, existe un abierto **U** de x_0 y un abierto **V** del punto $(\psi_1(x_0),\psi_2(x_0),\ldots,\psi_{m-1}(x_0))$ tal que $\forall\ x \in \mathbf{U}$ se cumple *i)* y *ii)* entonces ψ_m se dice localmente dependiente de las funciones ψ_1, $\psi_2,\ldots,\psi_{m-1}$ ó que (**) es un sistema localmente dependiente.

DEF.5.2.2. Si ninguna de las funciones anteriores dependen de las funciones restantes, en el sentido que acabamos de describir, entonces el conjunto de funciones $\psi_1(x)$, $\psi_2(x),\ldots,\psi_{m-1}(x)$, $\psi_m(x)$ se dicen independientes en el conjunto D y obviamente a (**) sistema independiente.

Supongamos, como antes, que se tiene el sistema

$$(**)\qquad \begin{array}{c} y_1 = \psi_1(x) \\ y_2 = \psi_2(x) \\ \cdots\cdots\cdots \\ y_m = \psi_m(x) \end{array} \qquad m \leq n$$

y asociemos en él la matriz jacobiana dada por

$$\begin{bmatrix} \dfrac{\partial\psi_1}{\partial x_1} & \dfrac{\partial\psi_1}{\partial x_2} & \cdots & \dfrac{\partial\psi_1}{\partial x_n} \\ \dfrac{\partial\psi_2}{\partial x_1} & \dfrac{\partial\psi_2}{\partial x_2} & \cdots & \dfrac{\partial\psi_2}{\partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \dfrac{\partial\psi_m}{\partial x_1} & \dfrac{\partial\psi_m}{\partial x_2} & \cdots & \dfrac{\partial\psi_m}{\partial x_n} \end{bmatrix}$$

nótese que ésta es precisamente la matriz A cuando las funciones ψ_i son lineales, y en dónde es el rango de A quien determina si hay ó no dependencia.

El siguiente teorema dá condiciones necesarias para la dependencia de funciones.

<u>**TEOREMA**</u> 5.2.1. Si las funciones en el sistema (**) son dependientes en D, en todo punto de este conjunto el rango de la matriz jacobiana es inferior a m.

OBSERVACION. Como se ve, las filas de la matriz jacobiana son los gradientes respectivos $\nabla\psi_i$, $i=1,2,...,m$, por lo que 'este teorema puede expresarse diciendo que si el sistema de funciones (**) es dependiente en **D**, entonces los gradientes $\nabla\psi_i$ de estas funciones son linealmente dependientes en todo punto $x \in$ **D**.

DEMOSTRACION. Por hipótesis, el conjunto ψ_1, ψ_2,..., ψ_m es dependiente en **D**, por lo tanto al menos una de esas funciones es expresable en términos de las restantes. Para ser mas explícitos, supongamos que ψ_m depende de las funciones ψ_1, ψ_2,..., ψ_{m-1}; entonces se debe tener que para alguna función Φ se tenga que $\psi_m(x) = \Phi(\psi_1(x),\psi_2(x),...,\psi_{m-1}(x))$ $x \in$ **D**. Aplicando a esta ecuacion la regla de la cadena obtenemos el siguiente sistema

$$\frac{\partial\psi_m}{\partial x_1} = \frac{\partial\Phi}{\partial y_1}\frac{\partial\psi_1}{\partial x_1} + \frac{\partial\Phi}{\partial y_2}\frac{\partial\psi_2}{\partial x_1} + \dots + \frac{\partial\Phi}{\partial y_{m-1}}\frac{\partial\psi_{m-1}}{\partial x_1}$$

$$\frac{\partial\psi_m}{\partial x_2} = \frac{\partial\Phi}{\partial y_1}\frac{\partial\psi_1}{\partial x_2} + \frac{\partial\Phi}{\partial y_2}\frac{\partial\psi_2}{\partial x_2} + \dots + \frac{\partial\Phi}{\partial y_{m-1}}\frac{\partial\psi_{m-1}}{\partial x_2}$$

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

$$\frac{\partial\psi_m}{\partial x_n} = \frac{\partial\Phi}{\partial y_1}\frac{\partial\psi_1}{\partial x_n} + \frac{\partial\Phi}{\partial y_2}\frac{\partial\psi_2}{\partial x_n} + \dots + \frac{\partial\Phi}{\partial y_{m-1}}\frac{\partial\psi_{m-1}}{\partial x_n}$$

Esto demuestra que la e-mésima fila de la matriz jacobiana, en todo punto $x \in$ **D**, es una combinación lineal de las filas restantes, vistas como columnas, por lo que el rango de esta matriz es inferior a m. ∎

De este teorema se desprenden los siguientes dos corolarios, el segundo de los cuales será utizado en el resultado posterior y dá condiciones suficientes para la independencia de las funciones.

COROLARIO #1. Supongamos que m=n y que el sistema es dependiente en **D**. En este caso tenemos que:

$$\mathrm{DET}\begin{pmatrix} \frac{\partial\psi_1}{\partial x_1} & \frac{\partial\psi_1}{\partial x_2} & \cdots\cdots & \frac{\partial\psi_1}{\partial x_n} \\ \frac{\partial\psi_2}{\partial x_1} & \frac{\partial\psi_2}{\partial x_2} & \cdots\cdots & \frac{\partial\psi_2}{\partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial\psi_n}{\partial x_1} & \frac{\partial\psi_n}{\partial x_2} & \cdots\cdots & \frac{\partial\psi_n}{\partial x_n} \end{pmatrix} = 0 \qquad \forall\ x \in \mathbf{D}.$$

COROLARIO #2. Supongamos que el rango de la matriz jacobiana por lo menos en un punto del abierto **D** es igual a m. En este caso el sistema de funciones es independiente en el conjunrto **D**.

DEMOSTRACION. (POR REDUCCION A LO ABSURDO)

Supongamos que el rango de la matriz jacobiana en al menos un punto de **D** es m y que las funciones $y_1=\psi_1$, $y_2=\psi_2,\dots,y_m=\psi_m$ son dependientes en **D**, entonces por el teorema 5.2.1 el rango de esta matriz debe de ser estrictamente menor que m $\forall\ x \in \mathbf{D}$.∇

$\therefore$ $y_1=\psi_1$, $y_2=\psi_2,\dots,y_m=\psi_m$ esun conjunto independiente en **D**. ∎

Para terminar, presentamos el teorema que da las condiciones suficientes para la dependencia local de las funciones.

TEOREMA 5.2.2. Nuevamente consideremos el sistema

$$(**)\qquad \begin{aligned} y_1 &= \psi_1(x) \\ y_2 &= \psi_2(x) \\ &\cdots\cdots\cdots \\ y_m &= \psi_m(x) \end{aligned} \qquad m \le n$$

y supongamos que en todos los puntos de un conjunto abierto $W \subset D$ se tenga que los gradientes $\nabla\psi_1$, $\nabla\psi_2, \ldots, \nabla\psi_{m-1}$ son linealmente independientes y junto con $\nabla\psi_m$ son linealmente dependientes. En otras palabras que el rango de la matriz jacobiana del sistema (**) es m-1 en todo punto del conjunto **W**. Sea $x_o \in W$; entonces las funciones ψ_1, $\psi_2, \ldots, \psi_{m-1}$ son independientes en **D** y mas aún, existe un abierto $U \subset W$ de x_o en el cual, la función ψ_m restante depende de las funciones ψ_1, $\psi_2, \ldots, \psi_{m-1}$.

DEMOSTRACION. Es claro, por el corolario 5.2.2, que las funciones ψ_1, $\psi_2, \ldots, \psi_{m-1}$ son independientes en **D**. Ahora definamos la función $f:W \to \mathbb{R}^m$ dada por $f(x_1,x_2,\ldots,x_n)=(\psi_1(x),\psi_2(x),\ldots,\psi_m(x))$ en donde $x=(x_1,x_2,\ldots,x_n)$. Así $f \in \mathscr{C}^1(W)$ y por hipótesis Df tiene rango constante m-1 en **W**. Por el teorema del rango existen difeomorfismos locales **h**, **g** tales que

$h\circ f\circ g^{-1}(z_1,z_2,\ldots,z_n) = (z_1,z_2,\ldots,z_{m-1},0) \in \mathbb{R}^m$ y en términos de las componentes de **h** podemos escribir que

$$\left(h_1(f[g^{-1}(z)]),h_2(f[g^{-1}(z)]),\ldots,h_m(f[g^{-1}(z)]\right) = (z_1,z_2,\ldots,z_{m-1},0)$$

donde $z=(z_1,z_2,\ldots,z_n)$. Sea **F** la última componente de **h**, es decir, $F(f[g^{-1}(z)])=h_m(f[g^{-1}(z)])$ y pongamos $x=g^{-1}(z)$ de manera que tengamos $F(f(x))=(\psi_1(x),\psi_2(x),\ldots,\psi_m(x))$; es claro que **F** es de clase $\mathscr{C}^1$ pues **h** lo es y como **h** es un difeomorfismo su inversa también es diferenciable, luego entonces **Dh** debe de ser invertible en particular para el punto $f(x_o)=(\psi_1(x_o),\psi_2(x_o),\ldots,\psi_m(x_o))$ lo cual implica que forzozamente para este punto se deberá tener que $DF = \left(\frac{\partial F}{\partial y_1}, \frac{\partial F}{\partial y_2},\ldots, \frac{\partial F}{\partial y_m}\right) \neq (0,0,..0)$ pues DF es la e-mésima fila de la matriz Dh que debe de ser invertible. Ahora para el punto x_o, pongamos $(\psi_1(x_o),\psi_2(x_o),\dots,\psi_m(x_o))=(y_1^\circ,y_2^\circ,\dots,y_m^\circ)$ y con $y_i=\psi_i$ tenemos que $F(y_1,y_2,\dots,y_m)=0$ y $F(y_1^\circ,y_2^\circ,\dots,y_m^\circ)=0$. La conclusión del teorema ahora es mas evidente si mostramos que, en este caso se satisface que $\frac{\partial F}{\partial y_m} \neq 0$. Para empezar si aplicamos la regla de la cadena a la ecuación $F(y_1,y_2,\dots,y_m)=0$, obtenemos el siguiente sistema:

$$\frac{\partial F}{\partial x_1} = \frac{\partial F}{\partial y_1}\frac{\partial \psi_1}{\partial x_1} + \frac{\partial F}{\partial y_2}\frac{\partial \psi_2}{\partial x_1} + \dots + \frac{\partial F}{\partial y_{m-1}}\frac{\partial \psi_{m-1}}{\partial x_1} + \frac{\partial F}{\partial y_m}\frac{\partial \psi_m}{\partial x_1} = 0$$

$$\frac{\partial F}{\partial x_2} = \frac{\partial F}{\partial y_1}\frac{\partial \psi_1}{\partial x_2} + \frac{\partial F}{\partial y_2}\frac{\partial \psi_2}{\partial x_2} + \dots + \frac{\partial F}{\partial y_{m-1}}\frac{\partial \psi_{m-1}}{\partial x_2} + \frac{\partial F}{\partial y_m}\frac{\partial \psi_m}{\partial x_2} = 0$$

$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots$$

$$\frac{\partial F}{\partial x_n} = \frac{\partial F}{\partial y_1}\frac{\partial \psi_1}{\partial x_n} + \frac{\partial F}{\partial y_2}\frac{\partial \psi_2}{\partial x_n} + \dots + \frac{\partial F}{\partial y_{m-1}}\frac{\partial \psi_{m-1}}{\partial x_n} + \frac{\partial F}{\partial y_m}\frac{\partial \psi_m}{\partial x_n} = 0$$

en términos de los gradientes y para los puntos x_o, $(y_1^\circ,y_2^\circ,\dots,y_m^\circ)$ el sistema anterior toma la forma siguiente

$\frac{\partial F}{\partial y_1}(y_1^\circ,y_2^\circ,\dots,y_m^\circ)\nabla\psi_1(x_o) + \frac{\partial F}{\partial y_2}(y_1^\circ,y_2^\circ,\dots,y_m^\circ)\nabla\psi_2(x_o) + \dots +$ $\frac{\partial F}{\partial y_{m-1}}(y_1^\circ,y_2^\circ,\dots,y_m^\circ)\nabla\psi_{m-1}(x_o) + \frac{\partial F}{\partial y_m}(y_1^\circ,y_2^\circ,\dots,y_m^\circ)\nabla\psi_m(x_o) = \mathbf{0}$. Si pasara que $\frac{\partial F}{\partial y_m}(y_1^\circ,y_2^\circ,\dots,y_m^\circ) = 0$, entonces tendríamos también que $\frac{\partial F}{\partial y_i}(y_1^\circ,y_2^\circ,\dots,y_m^\circ) = 0$ para $i=1,2,\dots,m-1$ pues, los gradientes $\nabla\psi_1$, $\nabla\psi_2,\dots,$ $\nabla\psi_{m-1}$ son linealmente independientes en x_o. Pero esto implicaría que $DF = \left(\frac{\partial F}{\partial y_1}, \frac{\partial F}{\partial y_2},\dots, \frac{\partial F}{\partial y_m}\right) = (0,0,..0)$ en el punto $(y_1^\circ,y_2^\circ,\dots,y_m^\circ)$ lo cual, es una contradicción. Por lo tanto es cierto que $\frac{\partial F}{\partial y_m}(y_1^\circ, y_2^\circ, \ldots, y_m^\circ) \neq 0$ y por el teorema de la función implícita en algún abierto **V** del punto $(y_1^\circ, y_2^\circ, \ldots, y_{m-1}^\circ)$ existe una función $\Phi: V \to \mathbb{R}$, con $\Phi \in \mathscr{C}^1(V)$, de modo que para los puntos $(y_1, y_2, \ldots, y_{m-1}) \in V$ se cumple que $y_m = \Phi(y_1, y_2, \ldots, y_{m-1})$. Sea el conjunto $U = \{x \in \mathbb{R}^n : (\psi_1(x), \psi_2(x), \ldots, \psi_{m-1}(x)) \in V\}$, así **U** es un abierto de x_o pues el mapeo $x \to (\psi_1(x), \psi_2(x), \ldots, \psi_{m-1}(x))$ es continuo. Por lo tanto, se ha probado la existencia del abierto **V** y de una función $\Phi: V \to \mathbb{R}$ tal que en el abierto **U** de x_o el punto $(\psi_1(x), \psi_2(x), \ldots, \psi_{m-1}(x)) \in V$ y $\psi_m(x) = \Phi(\psi_1(x), \psi_2(x), \ldots, \psi_{m-1}(x))$.∎

5.2. EJEMPLOS

5.2.1. consideremos el conjunto de funciones dado por

$$\psi_1(x,y) = \text{sen}(x + y),\ \psi_2(x,y) = \cos(x + y)$$

calculando el jacobiano tenemos:

$$\begin{vmatrix} \cos(x + y) & \cos(x + y) \\ -\text{sen}(x + y) & -\text{sen}(x + y) \end{vmatrix} = 0$$

$\forall\ (x,y) \in \mathbb{R}^2$ y como fácilmente se ve el rango es uno en todos los puntos del plano. Por el teorema 5.2.2 el conjunto de funciones es dependiente en un entorno de cada punto de $\mathbb{R}^2$. En este caso la dependencia se puede dar en forma explícita,por ejemplo, en el conjunto abierto de los puntos (x,y) para los cuales $\cos(x + y) > 0$ tal dependencia ersta dada por la fórmula:

$$\psi_2(x,y) = \sqrt{1 - [\psi_1(x,y)]^2}$$

5.2.2. Sea el conjunto $\psi_1(x,y) = e^{x^2-y^2}$, $\psi_2(x,y) = x^2+y^2$, y $\psi_3(x,y) = x^2$. Aquí en términos de los gradientes tenemos que $\nabla\psi_1 = (2xe^{x^2-y^2}, -2xe^{x^2-y^2})$, $\nabla\psi_2 = (2x,2y)$, $\nabla\psi_3 = (2x,0)$ por lo tanto son dependientes en todos puntos del plano pues Dim $\mathbb{R}^2 = 2$. En este caso ψ_1, ψ_2, ψ_3 es un conjunto dependiente y la ecuación que expresa esta dependencia es:

$$\psi_1(x,y) = e^{2\psi_3(x,y)-\psi_2(x,y)}.$$

5.2.3. Sea el sistema dado por las funciones

$\psi_1(x,y) = \ln x - \ln y$, $\psi_2(x,y) = \dfrac{x^2 + 3y^2}{2xy}$. Calculando el jacobiano

tenemos que $\begin{vmatrix} 1/x & -1/y \\ 1/2y - 3y/2x^2 & -x/2y^2 + 3/2x \end{vmatrix} = 0$. Por lo tanto las

funciones ψ_1, ψ_2 son dependientes. Para encontrar la ecuación que expresa la dependencia entre estas funciones observemos que $x/2y = \frac{1}{2}\ e^{\ln x-\ln y} = \frac{1}{2}\ e^{\psi_1(x,y)}$ y $3y/2x = \frac{3}{2}\ e^{-(\ln x-\ln y)} = \frac{3}{2}\ e^{-\psi_1(x,y)}$. De aquí obtenemos que:

$$\psi_2(x,y) = x/2y + 3y/2x = \frac{1}{2}\ e^{\psi_1(x,y)} + \frac{3}{2}\ e^{-\psi_1(x,y)}.$$

CAPITULO 6

EL TEOREMA DE LA VARIEDAD INVARIANTE

Ya hemos analizado que para una función $f \in \mathscr{C}^1$, las propiedades de invertibilidad y de rango de la diferencial Df, son transferidas localmente a la misma función. Ahora la cuestión es ¿Que podemos decir de un cierto mapeo T de case $\mathscr{C}^1$, definido en un mismo espacio, cuando la diferencial DT parte dicho espacio en dos (*)subespacios invariantes?

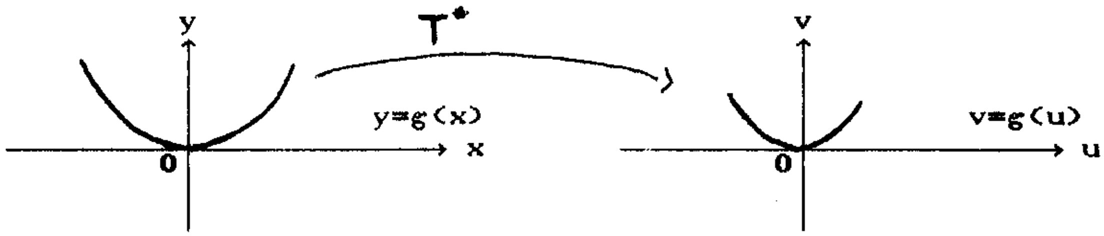
Por ejemplo, en el caso lineal, la transformación $T:\mathbb{R}^2 \to \mathbb{R}^2$ dada por $T(x,y) = (ax,by)$ manda cada uno de los ejes coordenados en sí mismo, es decir, son subespacios invariantes respecto a la transformación T. En el caso general, el problema se modifica en la siguiente forma: Si adicionalmente añadimos pequeñas perturbaciones en ambas componentes por medio de los términos f, h satisfaciendo que $f, h \in \mathscr{C}^1$ y junto con sus derivadas se anulan en el origen, obtenemos el mapeo resultante

$$T^*(x,y) = (ax + f(x,y),\ by + h(x,y))$$

Nos preguntamos ahora si la propiedad, antes mencionada, es de algun modo preservada por el mapeo T^*. Concretamente el teorema de la variedad invariante dá una respuesta a esta cuestión. Sin precisar, este teorema garantiza, bajo ciertas condiciones, la existencia de dos variedades las cuales, cada una de ellas, vienen representadas por la gráfica de una función de clase $\mathscr{C}^1$, con derivada nula en el origen, y de tal forma que se mantienen localmente invariante respecto a T^*. En particular, esto se

(*) Se dice que un subespacio $W \subset \mathbb{R}^s$ es invariante respecto a una transformación lineal $T:\mathbb{R}^s \to \mathbb{R}^s$ si $Tx \in W$, $\forall\ x \in W$ o que $T(W) \subset W$.

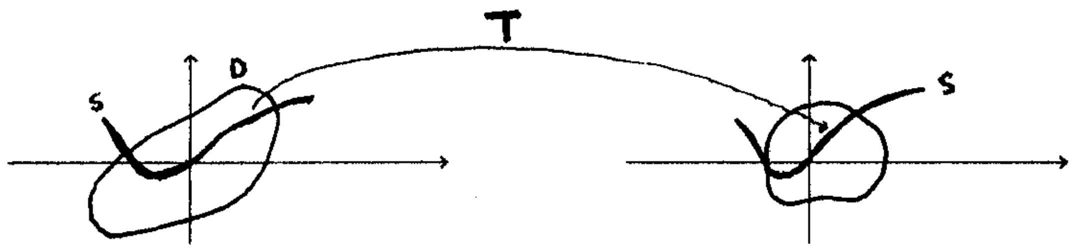
ilustra a continuación.



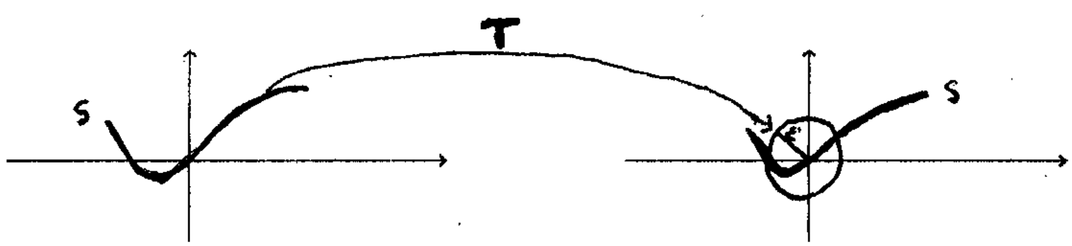
La demostración de este teorema, se obtiene mediante un método de aproximaciones sucesivas para algún operador definido en un espacio de funciones. La convergencia de sus iteradas conduce en el límite a una variedad invariante buscada. Dicha convergencia se establece mediante la aplicación de algunos resultados sobre convergencia uniforme así como también del teorema de Ascoli-Arzelá.De esta forma se presenta una modalidad del método de aproximaciones sucesivas análogo al que aparece en el teorema de la función inversa, sólo que en este último, se hace en un espacio de puntos.

1. CONJUNTOS INVARIANTES

DEF.6.1.1. Sea $\mathbf{T}:\mathbb{R}^k \to \mathbb{R}^k$ un mapeo continuo en una vecindad D de $\mathbf{x=0}$ al interior de una vecindad de $\mathbf{x=0}$ del mismo espacio, con $\mathbf{T(0)=0}$. Un conjunto S es llamado invariante con respecto a $\mathbf{T}$ si $\mathbf{T}(D \cap S) \subset S$.



DEF. 6.1.2. Un conjunto S es llamado localmente invariante con respecto a $\mathbf{T}$ si existe un $\varepsilon > 0$ tal que $x \in S$ implica que $\mathbf{T(x)} \in S$ y $\|\mathbf{T(x)}\| < \varepsilon$.



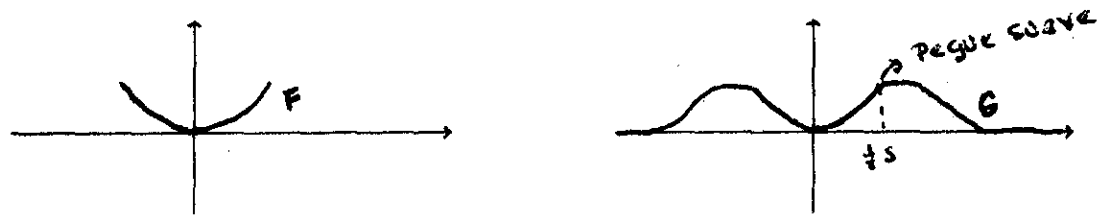
Podemos observar que si S es un conjunto invariante, entonces la intersección de S con la esfera $\|\mathbf{x}\| < \varepsilon$ es localmente invariante. Inversamente, si S es un conjunto localmente invariante entonces $S^\circ = \mathbf{T}(S \cap D)$ es un conjunto invariante. De esta manera la investigación de conjuntos invariantes puede ser reducida al estudio de conjuntos localmente invariantes y visceversa. Esto es conveniente por la siguiente razón: Si por ejemplo, en el mapeo $\mathbf{T}^*(ax + f(x,y),\ by + h(x,y))'$ las perturbaciones f, h se han modificado fuera de la esfera $\|(x,y)\| < \varepsilon$, y un conjunto invariante S° es determinado por el nuevo mapeo, entonces la intersección de S° y la esfera $\|(x,y)\| < \varepsilon$, es un conjunto localmente invariante por el mapeo original T^{*}.

Pensando en lo anterior, damos sin demostración el siguiente teorema , el cual será aplicado en el resultado que dá título a este capítulo.

TEOREMA 6.1.1. Sea $F:\mathbb{R}^n \to \mathbb{R}^m$ de clase $\mathscr{C}^1$ para $\|x\|$ pequeño, satisfaciendo que $F(0) = 0$ y $F'(0) = 0$. Sea $\theta > 0$ arbitrario. Entonces, existe un número $s = s(\theta) > 0$, ($s \to 0$ si $\theta \to 0$), y una función $G(x)$ de clase $\mathscr{C}^1$, definida para toda x tal que

$$G(x) = \begin{cases} F(x) & \text{si } \|x\| \leq \frac{1}{2} s \\ 0 & \text{si } \|x\| \geq s \end{cases}$$

y $\|G(x)\| \leq \theta \quad \forall\, x$. (véase, Philip Hartman capítulo 9 sección 1.).



Finalmente tenemos lo siguiente

TEOREMA 6.1.2 **(DE LA VARIEDAD INVARIANTE)**

Sea **A** una matriz de n×n, **B** una matriz *no singular* de m×m tales que

i) $\|A\| = a$, $\|B^{-1}\| = \frac{1}{b}$

ii) $a < b \qquad a < 1$

Consideremos el mapeo $T:\mathbb{R}^n \times \mathbb{R}^m$ definido por:

iii) $T(x,y) = (u,v) = (Ax + f(x,y), By + h(x,y))$

donde $f:\mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$, $h:\mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$ con $f, h \in \mathscr{C}^1$ para $\|x\|$, $\|y\|$ pequeños y satisfaciendo que

iv) $f, h, \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial h}{\partial x}, \frac{\partial h}{\partial y}$ se anulan en (0,0).

La conclusión es que existe una función $y = g(x) \in \mathscr{C}^1$ para $\|x\|$ pequeño que satisface que:

v) $g(0) = 0$, $g'(0) = 0$ y los mapeos

vi) $R(x,y)=(x^*,y^*)=(x,y - g(x))$, $R^{-1}(x^*,y^*)=(x,y)=(x^*,y^* - g(x^*))$

de tal modo que transforman T en la forma:

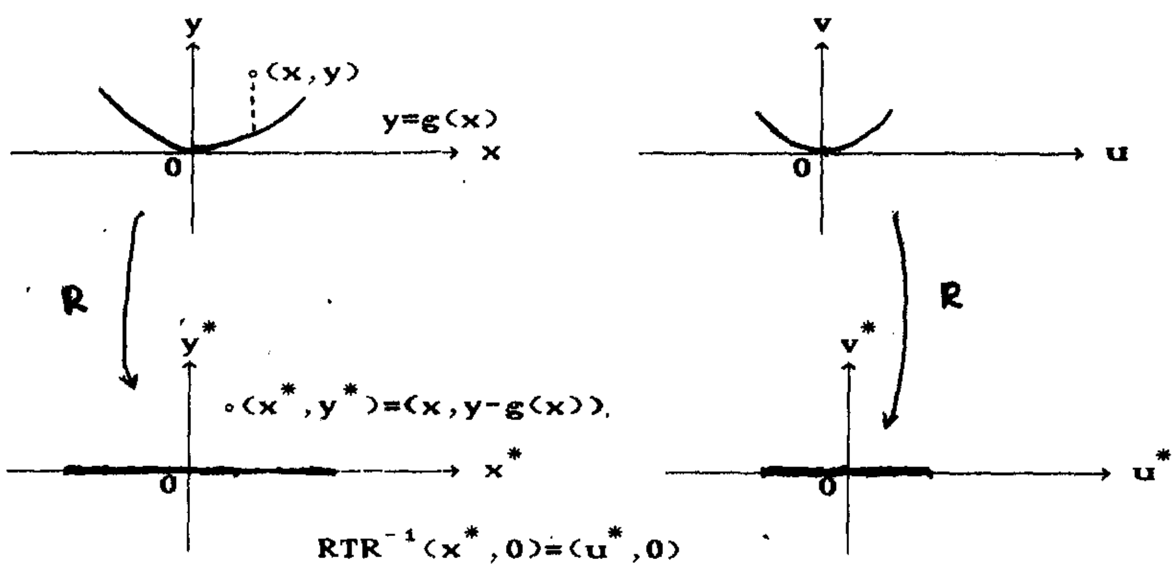
vii) $RTR^{-1}(x^*,y^*) = (u^*,v^*) = (Ax^* + U(x^*,y^*), By^* + V(x^*,y^*))$

donde

viii) $U, V, \frac{\partial U}{\partial x^*}, \frac{\partial U}{\partial y^*}, \frac{\partial V}{\partial x^*}, \frac{\partial V}{\partial y^*}$ se anulan en $(0,0)$ y

ix) $V(x^*,0) = 0$

La condición *ix)* muestra que el conjunto de puntos (x^*,y^*) cerca del origen y sobre la "recta" $y^* = 0$ es invariante por el mapeo dado en *vii)*; es decir, la variedad $y=g(x)$ queda localmente invariente por el mapeo T. Esto se ilustra a continuación.



DEMOSTRACION. Escencialmente la prueba se efectuará por **aproximaciones** sucesivas. En este proceso, utilizaremos algunos resultados fuertes de la convergencia uniforme que nos llevarán a probar la existencia del mapeo g. Además, nos apoyaremos en el teorema de Ascoli-Arzelá para mostrar que efectivamente g es de clase $\mathscr{C}^1$.

Supongamos por un momento que g(x) es conocido. Entonces si desarollamos la composición RTR^{-1} obtenemos lo siguiente

$$RTR^{-1}(x^*,y^*) = RT(x^*, y^* + g(x^*))$$

$$= R[Ax^* + f(x^*, y^* + g(x^*)), B(y^* + g(x^*)) + h(x^*, y^* + g(x^*))]$$

$$= \Big\{Ax^* + f(x^*, y^* + g(x^*)), By^* + Bg(x^*) + h(x^*, y^* + g(x^*)) -$$

$$g[Ax^* + f(x^*, y^* + g(x^*))]\Big\}.$$ De aquí se observa que

$U(x^*,y^*) = f(x^*, y^* + g(x^*))$ y además

$V(x^*,y^*) = Bg(x^*) + h(x^*, y^* + g(x^*)) - g[Ax^* + f(x^*, y^* + g(x^*))]$

en donde claramente U, V satisfacen *viii)*.

Dado que buscamos que $V(x^*,0) = 0$ y puesto que B es no singular entonces g debe satisfacer la ecuación funcional:

$$x)\quad g(x^*) = B^{-1}\Big\{g[Ax^* + f(x^*,g(x^*))] - h(x^*,g(x^*))\Big\}$$

Por lo tanto, el problema queda resuelto si probamos la existencia de una función $g \in \mathscr{C}^1$ que satisfaga *x)* y la condición *v)*. es para este fin que utilizaremos las aproximaciones sucesivas.

"Nótese que, equivalentemente, el problema es, determinar la existencia del punto fijo g para el operador $\mathcal{T}$, definido en un espacio de funciones, y que esta dado por:

$$\mathcal{T}g(x^*) = \bar{B}^1\Big\{g[Ax^* + f(x^*,g(x^*))] - h(x^*,g(x^*))\Big\}".$$

Tomemos el mapeo $g_0(x) = 0 \quad \forall\, x$ y definamos la suseción cuyo e-nésimo término esta dado por:

$xi)\quad g_n(x) = B^{-1}\left\{g_{n-1}[Ax + f(x,g_{n-1}(x))] - h(x,g_{n-1}(x))\right\}$

De esta manera, por ejemplo, tenemos que

$$g_1(x) = B^{-1}\left\{-h(x,0)\right\}$$

$$g_2(x) = B^{-1}\left\{g_1[Ax + f(x,g_1(x))] - h(x,g_1(x))\right\} \quad \text{etc.}$$

Ahora, el teorema 6.2.1 nos dice que podemos suponer, sin que se pierda la naturaleza local del teorema, que para algún número θ se tenga que

$xii)\quad f,\ h \in \mathscr{C}^1 \quad \forall\ (x,y)$

$f = 0\ ,\ h = 0$ si $\|(x,y)\| \geq s$ en donde s depende de θ satisfaciendo que $s(\theta) \to 0$ si $\theta \to 0$ y en donde además

$$\left\|\frac{\partial f}{\partial x}\right\|,\ \left\|\frac{\partial f}{\partial y}\right\|,\ \left\|\frac{\partial h}{\partial x}\right\|,\ \left\|\frac{\partial h}{\partial y}\right\| \leq \theta \qquad \forall\ (x,y).$$

Así, los mapeos g_1 , g_2 , ... son de clase $\mathscr{C}^1$ $\forall$ x. Ahora calculando la derivada en la ecuación $xi)$ obtenemos, aplicando la regla de la cadena, que:

$$xiii)\quad g_n'(x) = B^{-1}\left\{g_{n-1}'[Ax + f(x,g_{n-1}(x))]\left[A + \frac{\partial f}{\partial x}(x,g_{n-1}(x)) + \frac{\partial f}{\partial y}(x,g_{n-1}(x))g_{n-1}'(x)\right] - \left[\frac{\partial h}{\partial x}(x,g_{n-1}(x)) + \frac{\partial h}{\partial y}(x,g_{n-1}(x))g_{n-1}'(x)\right]\right\}.$$

Ahora, elejamos θ de modo que satisfaga la desigualdad

$xiv)\quad 0 < \theta < \min.\left\{\dfrac{b - a}{4}\ ,\ \dfrac{1 - a}{2}\right\}$ y definamos el número σ por

$xv)\quad \sigma = \dfrac{\theta}{b - a - 3\theta}$ de manera que $0 < \sigma < 1.$

Estamos en condiciones de probar los siguientes hechos.

I.- La sucesión $\{g_n'(x)\}$ está uniformemente acotada por σ. Esto es, se satisface que $\|g_n'(x)\| \leq \sigma \quad \forall$ x.

II.- La sucesión $\{g_n'(x)\}$ es equicontinua.

III.- La sucesión $\{g_n(x)\}$ es uniformemente convergente sobre todo conjunto acotado.

Estas tres afirmaciones serán probadas usando el método de inducción.

I.- $\|g_n'(x)\| \leq \sigma \quad \forall\ x.$

Es claro que se tiene para n=0. Supongamos que se vale para n - 1, esto es, $\|g_{n-1}'(x)\| \leq \sigma \quad \forall\ x$. Ahora, por *xiii)* tenemos que

$$\|g_n'(x)\| \leq \|B^{-1}\|\Big\{\|g_{n-1}'[Ax + f(x,g_{n-1}(x))]\|\Big(\|A\| + \Big\|\frac{\partial f}{\partial x}(x,g_{n-1}(x))\Big\| + \Big\|\frac{\partial f}{\partial y}(x,g_{n-1}(x))\Big\|\,\|g_{n-1}'(x)\|\Big) + \Big\|\frac{\partial h}{\partial x}(x,g_{n-1}(x))\Big\| + \Big\|\frac{\partial h}{\partial y}(x,g_{n-1}(x))\Big\|\,\|g_{n-1}'(x)\|\Big\}$$

aplicando *i)*, *xii)* y la hipótesis de inducción resulta que

$$\|g_n'(x)\| \leq b^{-1}\Big\{\sigma(a + \theta + \theta\sigma) + \theta + \theta\sigma\Big\} = b^{-1}\Big\{\sigma(a + 3\theta) + \theta\Big\}$$

La igualdad es porque también $0 < \sigma < 1$. Ahora por *xv)* tenemos que $\theta = \sigma(b - a - 3\theta) = \sigma b - \sigma(a + 3\theta)$ y por lo tanto

$$\|g_n'(x)\| \leq b^{-1}\Big\{\sigma(a + 3\theta) + \theta\Big\} = b^{-1}\Big\{\sigma(a + 3\theta) + \sigma b - \sigma(a + 3\theta)\Big\}$$

$= b^{-1}(\sigma b) = \sigma$. Esto completa la inducción.

II.- La sucesión $\{g_n'(x)\}$ es equicontinua.

Para esto sea $\Delta\frac{\partial f}{\partial x} = \frac{\partial f}{\partial x}(x + \Delta x,\ y + \Delta y) - \frac{\partial f}{\partial x}(x,y)$ y de la misma forma para $\Delta\frac{\partial f}{\partial y}$, $\Delta\frac{\partial h}{\partial x}$, $\Delta\frac{\partial h}{\partial y}$. Definamos aquí la siguiente función

$$xvi)\quad \gamma(\delta) = \operatorname{Sup}\Big\{ \Big\|\Delta\frac{\partial f}{\partial x}\Big\|,\ \Big\|\Delta\frac{\partial f}{\partial y}\Big\|,\ \Big\|\Delta\frac{\partial h}{\partial x}\Big\|,\ \Big\|\Delta\frac{\partial h}{\partial y}\Big\| \Big\}$$

Para $\|\Delta x\|$, $\|\Delta y\| \leq \delta$. $(\gamma(\delta) \to 0$ si $\delta \to 0)$

Pongamos ahora

$$xvii)\quad \Gamma(\delta) = \frac{4\gamma(\delta)}{b - a - 4\theta}$$

Obsérvese que $\|\Delta g_{n-1}\| = \|g_{n-1}(x + \Delta x) - g_{n-1}(x)\| \le \operatorname{Sup}\|g'_{n-1}(\xi)\|\,\|\Delta x\|$

$\le \sigma\|\Delta x\| \le \|\Delta x\|$. Así, por *xvi)* tenemos que

$$\left\|\Delta\frac{\partial f}{\partial x}(x,g_{n-1}(x))\right\| = \left\|\frac{\partial f}{\partial x}(x + \Delta x,g_{n-1}(x+\Delta x)) - \frac{\partial f}{\partial x}(x,g_{n-1}(x))\right\| \le \gamma(\|\Delta x\|)$$

puesto que $\|\Delta x\|$, $\|\Delta g_{n-1}\| \le \|\Delta x\|$. De aquí que

xviii) $\left\|\Delta\frac{\partial f}{\partial x}(x,g_{n-1}(x))\right\| \le \gamma(\|\Delta x\|) \le \gamma(\delta)$ para $\|\Delta x\| \le \delta$ y lo mismo pasa para las restantes.

Además, aplicando el teorema 0.3.3 se tiene la desigualdad

$$\|\Delta(Ax + f[x,g_{n-1}(x)])\| \le \left\|A + \frac{\partial f}{\partial x}(\xi,g_{n-1}(\xi)) + \frac{\partial f}{\partial y}(\xi,g_{n-1}(\xi))g'_{n-1}(\xi)\right\|\,\|\Delta x\|$$

$$\le \left(\|A\| + \left\|\frac{\partial f}{\partial x}(\xi,g_{n-1}(\xi))\right\| + \left\|\frac{\partial f}{\partial y}(\xi,g_{n-1}(\xi))\right\|\,\|g'_{n-1}(\xi)\|\right)\|\Delta x\|$$

$$\le (a + \theta + \theta\sigma)\|\Delta x\| \le (a + 2\theta)\|\Delta x\| \le \|\Delta x\|. \; (a + 2\theta < 1)$$

Por lo que

xix) $\|\Delta(Ax + f[x,g_{n-1}(x)])\| \le \|\Delta x\|$

Probaremos, para la equicontinuidad, que para $\Gamma(\delta)$ definido en *xviii)* se satisface que

$$\|\Delta g'_n(x)\| = \|g'_n(x + \Delta x) - g'_n(x)\| \le \Gamma(\delta) \quad \text{si} \quad \|\Delta x\| \le \delta.$$

Es claro que esto es cierto para n = 0. Supongamos que es válido para n - 1. Usando el hecho de que $\Delta fg = \Delta f\circ g(x + \Delta x) - f\circ\Delta g$ tenemos que por *xiii)*

$$\Delta g'_n(x) = B^{-1}\Big\{\Delta g'_{n-1}[Ax + f(x,g_{n-1}(x))]\Big(A + \frac{\partial f}{\partial x}(x+\Delta x,g_{n-1}(x+\Delta x)) +$$

$$\frac{\partial f}{\partial y}(x+\Delta x,g_{n-1}(x+\Delta x))g'_{n-1}(x+\Delta x)\Big) + g'_{n-1}[Ax+f(x,g_{n-1}(x))]\Big(\Delta\frac{\partial f}{\partial x}(x,g_{n-1}(x)) +$$

$$\Delta\frac{\partial f}{\partial y}(x,g_{n-1}(x))\circ g'_{n-1}(x+\Delta x) + \frac{\partial f}{\partial y}(x,g_{n-1}(x))\Delta g'_{n-1}(x)\Big) - \Big(\Delta\frac{\partial h}{\partial x}(x,g_{n-1}(x)) +$$

$$\Delta\frac{\partial h}{\partial y}(x,g_{n-1}(x))\circ g'_{n-1}(x+\Delta x) + \frac{\partial h}{\partial y}(x,g_{n-1}(x))\Delta g'_{n-1}(x)\Big)\Big\}.$$ Al tomar normas

se observa que en el lado derecho se tienen las siguientes desigualdades:

1) $\|\Delta g'_{n-1}[Ax + f(x,g_{n-1}(x))]\| \leq \Gamma(\delta)$ dado *xix)* y la hipótesis de inducción.

2) $\|\Delta g'_{n-1}(x)\| \leq \Gamma(\delta)$ hipótesis de inducción.

3) $\left\|\Delta\frac{\partial f}{\partial x}\right\|$, $\left\|\Delta\frac{\partial f}{\partial y}\right\|$, $\left\|\Delta\frac{\partial h}{\partial x}\right\|$, $\left\|\Delta\frac{\partial h}{\partial y}\right\| \leq \gamma(\delta)$ por *xvi)*.

4) $\|g'_{n-1}\| \leq \sigma$ por la parte *I*.

5) $\left\|\frac{\partial f}{\partial x}\right\|$, $\left\|\frac{\partial f}{\partial y}\right\|$, $\left\|\frac{\partial h}{\partial x}\right\|$, $\left\|\frac{\partial h}{\partial y}\right\| \leq \theta$ por el teorema 6.2.1.

De aquí que obtengamos que

$$\|\Delta g'_n(x)\| \leq b^{-1}\Big\{\Gamma(\delta)(a + \theta + \sigma\theta) + \sigma[\gamma(\delta) + \gamma(\delta)\sigma + \theta\Gamma(\delta)] + \gamma(\delta) + \gamma(\delta)\sigma + \theta\Gamma(\delta)\Big\}.$$

$$\leq b^{-1}\Big\{\Gamma(\delta)(a + 2\theta) + \gamma(\delta) + \gamma(\delta) + \gamma(\delta) + \gamma(\delta) + 2\theta\Gamma(\delta)\Big\}$$

$$= b^{-1}\Big\{\Gamma(\delta)(a + 4\theta) + 4\gamma(\delta)\Big\}.$$ Finalmente, puesto que por *xvii)* $4\gamma(\delta) = \Gamma(\delta)b - \Gamma(\delta)(a + 4\theta)$, se tiene entonces que

$\|\Delta g'_n(x)\| \leq \Gamma(\delta)$ si $\|\Delta x\| \leq \delta$. Esto completa la inducción.

III.- La sucesión g_0, g_1,..., es uniformemente convergente sobre todo conjunto acotado.

Esto se satisface si existen constantes M, r con 0< r <1 tales que

$$\|g_n(x) - g_{n-1}(x)\| \leq Mr^n\|x\|$$

Para n = 1 tenemos que

$$\|g_1(x) - g_0(x)\| = \|g_1(x)\| = \|g_1(x) - g_1(0)\| \leq \mathrm{Sup}\|g_1'(\xi)\|\,\|x\| \leq \sigma\|x\|$$

así, primeramente si M, r son tales que $Mr = \sigma$ entonces resulta → $\|g_1(x) - g_0(x)\| \leq Mr\|x\|$. Esto es, se satisface para n = 1. Ahora supongamos que también es es válido para n - 1, es decir, que se tenga la desigualdad

$$\|g_{n-1}(x) - g_{n-2}(x)\| \leq Mr^{n-1}\|x\|$$

Debido a *xi)* tenemos

$$\|g_n(x) - g_{n-1}(x)\| \leq b^{-1}\Big\{\Big\|g_{n-1}[Ax + f(x,g_{n-1}(x))] - h(x,g_{n-1}(x)) -$$

$$\Big(g_{n-2}[Ax + f(x,g_{n-2}(x))] - h(x,g_{n-2}(x))\Big)\Big\|\Big\} \Rightarrow$$

$$b\|g_n(x)-g_{n-1}(x)\| \leq \Big\|g_{n-1}[Ax+f(x,g_{n-1}(x))]-\underline{g_{n-2}[Ax+f(x,g_{n-1}(x))]} +$$

$$\underline{g_{n-2}[Ax+f(x,g_{n-1}(x))]} - g_{n-2}[Ax+f(x,g_{n-2}(x))]\Big\| +$$

$$\Big\|h(x,g_{n-1}(x)) - h(x,g_{n-2}(x))\Big\| \Rightarrow$$

$$b\|g_n(x)-g_{n-1}(x)\| \leq \Big\|g_{n-1}[Ax+f(x,g_{n-1}(x))]-g_{n-2}[Ax+f(x,g_{n-1}(x))]\Big\| +$$

$$\Big\|g_{n-2}[Ax+f(x,g_{n-1}(x))]-g_{n-2}[Ax+f(x,g_{n-2}(x))]\Big\| +$$

$$\Big\|h(x,g_{n-1}(x)) - h(x,g_{n-2}(x))\Big\|.$$

Ahora, reduciremos los tres términos del lado derecho de ésta última desigualdad como sigue

1') $$\Big\|g_{n-1}[Ax + f(x,g_{n-1}(x))] - g_{n-2}[Ax + f(x,g_{n-1}(x))]\Big\| \leq$$

$$\underbrace{M\Big\|Ax+f(x,g_{n-1}(x))\Big\|r^{n-1}}_{\text{por hip. de inducción}} \leq \underbrace{M(a + 2\theta)r^{n-1}}_{\text{por } xix)}$$

2') $$\Big\|g_{n-2}[Ax + f(x,g_{n-1}(x))] - g_{n-2}[Ax + f(x,g_{n-2}(x))]\Big\| \leq$$

$$\underbrace{\mathrm{Sup}\Big\|g'_{n-2}(\xi_1)\Big\|\Big\|f(x,g_{n-1}(x)) - f(x,g_{n-2}(x))\Big\|}_{\text{por el teorema 0.3.3}} \leq$$

$$\mathrm{Sup}\ \Big\|g'_{n-2}(\xi_1)\Big\| \circ \underbrace{\mathrm{Sup}\Big\|\frac{\partial f}{\partial y}(\xi_2)\Big\|\Big\|g_{n-1}(x) - g_{n-2}(x)\Big\|}_{\text{teorema 0.3.3}} \leq \sigma\theta Mr^{n-1}\|x\|$$

La última desigualdad es debido a la parte *I*, el teorema *xii)* y la hipótesis de inducción.

3') $\left\|h(x,g_{n-1}(x)) - h(x,g_{n-2}(x))\right\| \le \underbrace{\mathrm{Sup}\left\|\frac{\partial h}{\partial y}(\xi_3)\right\|\left\|g_{n-1}(x) - g_{n-2}(x)\right\|}_{\text{teorema 0.3.3}}$

$$\le \theta M r^{n-1}\|x\|.$$

Juntando 1', 2', 3' obtenemos que

$$b\|g_n(x)-g_{n-1}(x)\| \le M(a + 2\theta)r^{n-1} + \sigma\theta M r^{n-1}\|x\| + \theta M r^{n-1}\|x\|$$

$$= Mr^{n-1}(a + 2\theta + \sigma\theta + \theta)\|x\| \le Mr^{n-1}(a + 4\theta)\|x\|$$

De aquí que tengamos que

$$\|g_n(x)-g_{n-1}(x)\| \le Mr^{n-1}\frac{(a + 4\theta)}{b}\|x\|$$

Tomando $r = \dfrac{(a + 4\theta)}{b}$ se puede ver que $0 < r < 1$ y con $M = \sigma/r$ se verifica que en todo conjunto acotado

$$\|g_n(x)-g_{n-1}(x)\| \le Mr^n\|x\|.$$

Esto completa la inducción.

Ahora, daremos la última etapa de la demostración del teorema a continuación.

Debido a la parte *I*, tenemos el siguiente diagrama:

$$\begin{array}{ccc} g_n(x) = B^{-1}\Big\{g_{n-1}[Ax & + f(x,g_{n-1}(x))] & - h(x,g_{n-1}(x))\Big\} \\ \Big\downarrow u & \Big\downarrow u & \\ g(x) = B^{-1}\Big\{g[Ax & + f(x,g(x))] & - h(x,g(x))\Big\} \end{array}$$

Es decir, se ha probado la existencia de una función $y = g(x)$ satisfaciendo *x)* y como $g_n(0) = 0$, también $g(0) = 0$.

Probemos finalmente que $g \in \mathscr{C}^1$ con $g'(0) = 0$.

Como la sucesión de matrices jacobianas $(g_n'(x))$ es unformemente acotada,(por *I*), y equicontinua,(por *II*), se deduce del teorema **0.1.2,(ASCOLI-ARZELA)**, que existe una subsucesión $\mathbf{g}'_{n_k}$ tal que

$$\mathbf{g}'_{n_k} \xrightarrow{\ u\ } \psi$$

para alguna ψ y en donde ψ es continua pues $\mathbf{g}'_{n_k}$ lo son $\forall$ k. Para esos mismos subíndices tenemos que

$$\mathbf{g}_{n_k}(\mathbf{0}) \longrightarrow \mathbf{g}(\mathbf{0})$$

y como consecuencia de la convergencia uniforme de $\mathbf{g}'_{n_k}$ se deduce que $\psi = \mathbf{g}'$. Así $g \in \mathcal{C}^1$ y como $\mathbf{g}'_n(\mathbf{0}) = \mathbf{0}$, también $\mathbf{g}'(\mathbf{0}) = \mathbf{0}$.∎

CONCLUSIONES Y SUGERENCIAS

La concordancia existente entre las propiedades cualitativas de las transformaciones lineales y las propiedades locales de las funciones diferenciables, que está reflejada en los cuatros teoremas fundamentales de este trabajo, es un hecho satisfactorio que permite establecer la estrecha relación entre el cálculo diferencial y el algebra lineal. En realidad, el cálculo diferencial a traves del algebra lineal, constituye una alternativa poderosa con la cual se pueden atacar una gran variedad de problemas no lineales que se presentan en matemáticas. De hecho esto nos ha permitido, en términos generales, hacer un análisis completo de las funciones continuamente diferenciables a traves de su diferencial que la aproxima alrededor de un punto.

presentamos ahora el siguiente diagrama que nos muestra el caso lineal y su correspondiente generalización al caso no lineal.

CASO LINEAL		GENERALIZACION (CASO NO LINEAL)
Sistemas de ecuaciones lineales (regla de cráner)	→	teorema de la función inversa teorema de la función implícita
Trasnsformaciones lineales de un cierto rango. (cambios de coordenadas)	→	teorema del rango.
Descomposición en subespacios invariantes por una transformación lineal.	→	teorema de la variedad invariante

SUGERENCIAS: Existen muchas vertientes mediante las cuales, esta tesis puede ser extendida ya sea en el terreno del análisis, o dentro del campo de las aplicaciones que en este caso no se han incluído. Por ejemplo, se puede generalizar nuestro estudio para funciones de clase $\mathscr{C}^m$ o buscar las condiciones para que se tenga invertibilidad global en el teorema de la función inversa y de una gran variedad de problemas aplicados que justifican la enorme potenciabilidad de los resultados ya expuestos. Esto sugiere el camino que permite darle un mas completo acabado a este trabajo a manera de una invitación a futuras investigaciones.

BIBLOGRAFIA

1.- APOSTOL TOM. L. ANALISIS MATEMATICO. EDITORIAL REVERTE, BARCELONA 1981.

2.- ABRAHAM\MARSDEN\RATIU. MANIFOLDS TENSOR ANALISIS and APLICATIONS, ADIDISON-WESLEY, PUBLISHING COMPANY, INC. MASSACHUSETS 1983.

3.- BARTLE ROBERT. INTRODUCCION AL ANALISIS MATEMATICO, VERSION ESPAÑOLA, LIMUSA, MEXICO 1982.

4.- GITLER S.\ANTONIANO E. TOPOLOGIA DIFERENCIAL, I.P.N. OAXTEPEC MORELOS 1979.

5.- CARTAN HENRI. CALCULO DIFERENCIAL, OMEGA, BARCELONA 1972.

6.- HARTMAN PHILIP. ORDINARY DIFFERENTIAL ECUATIONS, WILEY, NEW YORQ 1964.

7.- STEPHEN H. FRIEDBERG\ARNOLD J INSEL\LAWRENCE. E. SPENCE. ALGEBRA LINEAL, ILLINOIS STATE UNIVERISITE 1982.

8.- HALE J.K. ORDINARY DIFFERENTIAL ECUATIONS, WILEY INTERSCIENCIE, NEW YORQ 1969.

9.- KREYSING ERWIN. INTRODUCTORY FUNCIONAL ANALISIS WHIT APLICATIONS, JHON WILEY & SONS INC. E.E.U.U.

10.- KUDRIASTEV L. D. CURSO DE ANALISIS MATEMATICO, TOMO II, MIR, MOSCU 1984.

11.- LIPSCHUTS SEYMOUR. ALGEBRA LINEAL, MCGRAUW-HILL, MEXICO 1979.

12.- RUDIN WALTER. PRINCIPLES OF MATHEMATICAL ANALISIS, GRAW-HILL, BOOK COMPANY 1976.

13.- TAYLOR\MAN. ADVACED CALCULUS, SEGUNDA EDICION, XEROX CORPORACION 1972.