UNIVERSIDAD DE EL SALVADOR
FACULTAD DE INGENIERIA Y ARQUITECTURA

CONJUNTOS Y FUNCIONES CONVEXAS
EN ESPACIOS VECTORIALES NORMADOS

TRABAJO DE GRADUACION
PREVIO A LA OPCION
DEL TITULO DE
LICENCIADO EN MATEMATICA.

PRESENTADO POR:

MANUEL DE JESUS AREVALO BONILLA.
JUAN JUVENCIO CASTILLO MEZQUITA.

JULIO DE 1979

SAN SALVADOR, EL SALVADOR, CENTRO AMERICA.

UNIVERSIDAD DE EL SALVADOR

RECTOR a. i. : Lic. JOSE LUIS ARGUETA ANTILLON

SECRETARIO a. i. : Lic. OSCAR ARMANDO ACEVEDO VELASQUEZ

FACULTAD DE INGENIERIA Y ARQUITECTURA

DECANO a. i.: Ing. JOSE FRANCISCO AGUIRRE TOVAR

SECRETARIO a. i.: Ing. MANUEL ANTONIO CAÑAS LAZO

DEPARTAMENTO DE MATEMATICA

JEFE DEL DEPARTAMENTO a. i.: Lic. FRANCISCO MAURICIO FIGEAC

TRABAJO DE GRADUACION

ASESOR: Ing. CARLOS MAURICIO CANJURA LINARES.

INTRODUCCION

En el presente trabajo hemos recopilado alguna información sobre los conjuntos convexos y las funciones convexas y decimos alguna porque, sobre tales temas, el trabajo realizado por los Matemáticos Europeos y Norteamericanos es relativamente abundante.

Nuestro propósito al realizar esta recopilación comprende los siguientes aspectos:

1. Conocer de tales temas ya que, la teoría sobre los conjuntos y funciones convexos, no es muy conocida en nuestro medio.

2. Establecer un punto inicial en el estudio de los fundamentos de una de las ramas de la Matemática Aplicada: La Optimización.

Durante nuestra labor de recopilación hemos podido darnos cuenta de que para realizar una investigación seria sobre estos temas es necesario poseer bases teóricas sólidas sobre Topología, Geometría y Análisis Funcional. Con esta aclaración debe quedar entendido que nuestro trabajo es incipiente.

El Capítulo I está dedicado a recordar conceptos - topológicos y de análisis matemático básicos que son empleados explícita o implícitamente en los dos capítulos posteriores. Otros conceptos tales como las propiedades de las funciones definidas sobre espacios vectoriales y sus derivadas nos eran completamente desconocidos. Ahora, sabemos que se puede estudiar este campo con mayor amplitud.

Los Capítulos II y III tratan, respectivamente, sobre los conjuntos convexos y las funciones convexas. - El material que hemos reunido en esos capítulos es poco, puede investigarse más sobre estos temas. Existe una - íntima relación entre las funciones convexas y los conjuntos convexos ya que, éstos son el dominio natural de aquellas y las propiedades de esos conjuntos determinan las propiedades de las funciones convexas.

Finalmente, queremos dejar constancia de nuestro - agradecimiento a las personas que nos ayudaron en la elaboración de este trabajo: Ing. Carlos Mauricio Canjura, nuestro Asesor; Sra. Nohemy de Rovelo, quien realizó la mecanografía y Sr. Mauricio García, quien hizo la impresión.

I N D I C E

PAGINA

CAPITULO II

CONJUNTOS CONVEXOS.

CAPITULO III

FUNCIONES CONVEXAS.

CAPITULO I
ESPACIOS VECTORIALES Y ESPACIOS METRICOS

La teoría y métodos de los espacios vectoriales y topológicos se han convertido en instrumentos indispensables para el estudio de la optimización. En este capítulo revisamos propiedades fundamentales de la teoría de los espacios vectoriales, de las funciones definidas sobre ellos y de las derivadas de estas funciones.

1. Espacios Vectoriales.

Sea V un conjunto no vacío con elementos llamados puntos o vectores y $(K, +, \cdot)$, un campo conmutativo, a quien representaremos únicamente con el símbolo K y llamaremos, simplemente, campo. La estructura algebráica formada por el conjunto V; la operación binaria interna $+: V \times V \to V$, suma de vectores; la operación binaria externa $\circ: K \times V \to V$, multiplicación por escalar; que satisface los siguientes axiomas es llamada el espacio vectorial V sobre el campo K.

S1) $\forall x, y, z \in V, (x+y) + z = x + (y+z)$.

S2) $\forall x, y \in V, x + y = y + x$.

S3) Existe $\bar{o} \in V$, llamado el vector nulo tal que $\forall x \in V, x + \bar{o} = x$.

S4) Para cada $x \in V$, existe $(-x) \in V$, llamado el inverso aditivo de x, tal que $x + (-x) = \bar{o}$.

P1) $\forall \alpha, \beta \in K; \forall x \in V, (\alpha\beta) x = \alpha(\beta x)$.

P2) $\forall \alpha \in K; \forall x, y \in V, \alpha(x+y) = \alpha x + \alpha y$.

P3) $\forall \alpha, \beta \in K; \forall x \in V, (\alpha+\beta) x = \alpha x + \beta x$.

P4) $\forall x \in V, 1_{\circ}x = x$, con 1 representando la identidad multiplicativa.

Por comodidad, representaremos los espacios vectoriales con el mismo símbolo que empleamos para representar el correspondiente conjunto de vectores. Por otra parte, en la mayoría de veces, omitiremos la escritura delsímbolo $\circ$ de la multiplicación por escalar.

Algunas de las propiedades de los espacios vectoriales que se deducen, inmediatamente de los axiomas, aparecen en la proposición 1.2.

1.2 <u>Proposición.</u>

Sea V un espacio vectorial sobre el campo K. Entonces,

1. El vector nulo $\bar{o}$ es único.
2. El inverso aditivo de cada vector es único.
3. $\forall \alpha \in K, \alpha\bar{o} = \bar{o}$.
4. $\forall x \in V, o \cdot x = \bar{o}$, con 0 la identidad para la suma de escalares.
5. $\forall \alpha \in K; \forall x \in V, (-\alpha) x = \alpha(-x) = -(\alpha x)$.
6. $\alpha x = \bar{o}$, con $\alpha \in K$ y $x \in V$, implica $\alpha = 0$ ó $x = \bar{o}$.
7. $\forall \alpha \in K; \forall x, y \in V, \alpha(x-y) = \alpha x - \alpha y$.
8. $\alpha x = \alpha y$, con $\alpha \neq 0$, implica $x = y$.
9. $\alpha x = \beta x$, con $x \neq \bar{o}$, implica $\alpha = \beta$.

1.3 Ejemplos de Espacios Vectoriales.

1. Sea K un campo y $K^n = \{(\alpha_1, \ldots, \alpha_n) / \alpha_i \in K, i=1,\ldots,n\}$. Entonces, K^n con la adición de vectores y la multiplicación vector por escalar definidas así:

   $(\alpha_1,\ldots,\alpha_n)+(\beta_1,\ldots,\beta_n) = (\alpha_1+\beta_1,\ldots,\alpha_n+\beta_n)$,

   $k(\alpha_1,\ldots,\alpha_n) = (k\alpha_1,\ldots,k\alpha_n)$,

   con α_i, β_i, $k \in K$, es un espacio vectorial. Se representa con K^n. Un caso particular de esta clase de espacios vectoriales es $\mathbb{R}^n$; con $\mathbb{R}$, el campo de los números reales.

2. El conjunto de todos los polinomios en una indeterminada x con coeficientes a_i, $i = 1,\ldots,n$, en un campo K; con la suma y multiplicación por un escalar usuales.

3. El conjunto de todas las matrices de orden mxn cuyas componentes son elementos de un campo arbitrario K, con la adición y la multiplicación por un escalar usuales.

4. Sea K el campo de los números reales y F, el conjunto de todas las funciones reales contínuas definidas sobre el intervalo $[a,b] \subseteq \mathbb{R}$. F, con la suma de funciones reales y la multiplicación de un real por una función forman un espacio vectorial sobre $\mathbb{R}$.

La comprobación de que los ejemplos citados son espacios vectoriales, se omite.

2. Subespacios Vectoriales.

2.1 Definición.

Sea V un espacio vectorial sobre el campo K. Si $W \subseteq V$, con las operaciones adición de vectores y multiplicación vector por escalar, definidas en V, forman un espacio vectorial; entonces, W es llamado un subespacio de V.

2.2 Ejemplos de Subespacios Vectoriales.

1. Sea V un espacio vectorial sobre el campo K. V y $\{\bar{o}\}$ son los subespacios triviales de V. Cualquier subespacio de V que no sea V es llamado un subespacio propio de V.

2. El conjunto M de todas las matrices simétricas $A = [a_{ij}]_{nxn}$, con $a_{ij} = a_{ji}$, la suma de matrices y el producto de un escalar por una matriz, restringidas a M, es un subespacio del espacio vectorial de las matrices cuadradas de orden n.

3. El conjunto de los polinomios de grado menor o igual que n, con la suma de polinomios y el producto de un escalar por un polinomio es un subespacio del espacio vectorial de los polinomios.

2.3 Proposición.

Sea V un espacio vectorial sobre el campo K. $W \subseteq V$ es un subespacio de V si y sólo si

1. $W \neq \emptyset$.
2. $\forall x, y \in W$, $(x + y) \in W$.
3. $\forall \alpha \in K$, $\forall x \in W$, $\alpha x \in W$.

Demostración.

Si W es un subespacio de V, W es un espacio vectorial y satisface las condiciones 1, 2 y 3.

Ahora, supongamos que se satisfacen estas condiciones. - Los axiomas de conmutatividad y asociatividad para la adición de vectores y los axiomas de la multiplicación vector por escalar son válidas para los elementos de W, puesto que son válidas para todos los ejementos de V y $W \subseteq V$. Por lo tanto, para mostrar esta segunda parte del teorema, sólo - es necesario comprobar la existencia, en W, del vector nulo y de un inverso aditivo para cada $x \in W$. Sea $x \in W$ un elemento cualquiera , por hipótesis, $\alpha x \in W$, $\forall \alpha \in K$, $\forall x \in W$ luego, $0x = \bar{0}$ y $(-1)x = -x$ están en W.

2.4 Proposición.

Sea V un espacio vectorial sobre el campo K. $W \subseteq V$ es un subespacio de V si y sólo si $W \neq \emptyset$ y

$\forall\alpha,\beta\in K$; $\forall x, y\in W$, $(\alpha x + \beta y) \in W$.

Demostración.

Si W es un subespacio de V; entonces, de acuerdo con la prop. 2.3, pág. 5 , $W \neq 0$; αx, $\beta y \in W$ y $(\alpha x + \beta y)\in W$; $\forall\alpha, \beta\in K$; $\forall x,y\in W$.
Si $W \neq \emptyset$, podemos seleccionar en W, dos elementos cuales - quiera x,y. Luego $x + y = 1_\circ x+1_\circ y$, y $\alpha x+0_\circ y = \alpha x$ es - tán, por hipótesis, en W y de nuevo, por la prop. 2.3, pág. 5 , W es un subespacio de V.

2.4.1 Corolario.

Sea V un espacio vectorial sobre el campo K. $W \subseteq V$ es un subespacio de V si y sólo si $W \neq \emptyset$ y $\forall\alpha\in K$, $\forall x, y\in W$, $(\alpha x + y) \in W$.

Demostración.

Este corolario es una consecuencia inmediata de la - prop. 2.4, pág. 5, si consideramos $\beta = 1$.

2.5 Proposición.

Sea V un espacio vectorial sobre el campo K y C, una colección no vacía de subespacios de V; entonces, $\cap C_j$, con $C_j\in C$, es un subespacio de V.

Demostración.

$\cap C_j \neq \emptyset$ ya que $\bar{o} \in C_j$, $\forall_j$, por ser C_j un subespacio.
Si $x, y \in \cap C_j$; entonces, $x, y \in C_j$, $\forall_j$. Luego,
$(\alpha x + \beta y) \in C_j$, $\forall_j$; $\forall \alpha$, $\beta \in K$ y, en consecuencia, $(\alpha x + \beta y) \in \cap C_j$.
El teorema está probado.

3. Generadores Lineales y Combinaciones Lineales.

Sea X un subconjunto de un espacio vectorial sobre un campo K y F, la colección de todos los subespacios de V - que contienen a X. Como $X \subseteq V$ y V es un subespacio de sí mismo, $F \neq \emptyset$. La intersección de los elementos de F: $\cap F_j$, $F_j \in F$ es, de acuerdo con la proposición 2.5, un subespacio de V. Las propiedades características de este subespacio se muestran en la siguiente proposición:

3.1 Proposición.

El subespacio $\cap F_j$, $F_j \in F$, con F la colección de todos los subespacios del espacio vectorial V sobre el campo K - que contienen a $X \subseteq V$, tiene las siguientes propiedades:

1. $X \subseteq \cap F_j$.
2. Si un subespacio vectorial U de V contiene a X, $\cap F_j \subseteq U$.
3. $\cap F_j$ es único.

Demostración.

Para 1. Cada elemento de X es un elemento de cada subespacio F_j y, por lo tanto, un elemento de $\cap F_j$. Luego, $X \subseteq F_j$.

Para 2. Si U contiene a X, $U \in F$ y si $x \in \cap F_j$, x pertenece a cada elemento de F; en particular a U. Así que, $\cap F_j \subseteq U$.

Para 3. Sea W un subespacio que satisface las condiciones 1 y 2. Entonces, como $\cap F_j$ satisface 2, $W \subseteq \cap F$ y, como también W satisface 2, $\cap F_j \subseteq W$. Por lo tanto, $W = \cap F_j$. Este resultado nos permite afirmar que $\cap F_j$ es el mínimo subespacio vectorial de V que contiene a X.

3.2 <u>Definición.</u>

Sea V un espacio vectorial sobre el campo K, $X \subseteq V$ y F, la colección de subespacios de V que contienen a X. Al mínimo subespacio vectorial de V que contiene a X, $\cap F_j, F_j \in F$, con j recorriendo un conjunto de índices, le llamaremos el subespacio vectorial generado por X y lo denotaremos con - Gen(X). Recíprocamente, X es el generador de $\cap F_j$.

Definimos $Gen(\emptyset) = \{\bar{o}\}$.

Si S es una familia no vacía de subespacios del espacio vectorial V; entonces, US_j, $S_j \in S$ es un subconjunto de V y, de acuerdo con la prop. 3.1, pág. 7, $Gen(US_j)$ es el mínimo subespacio de V que contiene a US_j.

3.3 Definición.

Sea S_1 y S_2 dos subespacios del espacio vectorial V sobre el campo K. La suma de estos subespacios se denota y se define así:

$S_1 + S_2 = \{x \in V / x = s_1 + s_2,\ s_1 \in S_1,\ s_2 \in S_2\}$.

3.4 Proposición.

Si S_1 y S_2 son dos subespacios del espacio vectorial V sobre el campo K; entonces, la suma $S_1 + S_2$ es el menor subespacio de V que contiene a S_1 y a S_2 y, por lo tanto, a $S_1 \cup S_2$; es decir, $S_1 + S_2 = \mathrm{Gen}(S_1 \cup S_2)$.

Demostración.

Como $\bar{o} \in S_1, S_2$ y $\bar{o} = \bar{o} + \bar{o}$, $\bar{o} \in S_1 + S_2$. Luego, $S_1 + S_2 \neq \emptyset$.

Supongamos que $s_1 + s_2$, $s_3 + s_4$ están en $S_1 + S_2$ y k, en K. Entonces, $(s_1+s_2)+(s_3+s_4) = (s_1+s_3)+(s_2+s_4)$ y $k\,(s_1+s_2)=ks_1+ks_2$ están en S_1+S_2 ya que, por ser S_1 y S_2 subespacios, s_1+s_3, ks_1 están en S_1 y $s_2 + s_4$, ks_2, en S_2. Luego, $S_1 + S_2$ es un subespacio de V.

Puesto que $S_2=\{\bar{o}\}+S_2=\{x \in V/x=\bar{o}+s_2=s_2, \bar{o} \in S_1, s_2 \in S_2\}$ se deduce, inmediatamente, que $S_2 \subseteq S_1 + S_2$. De igual manera concluimos que $S_1 \subseteq S_1 + S_2$. Por lo tanto, S_1 y S_2 son subespa -

cios de $S_1 + S_2$ y, $S_1 \cup S_2 \subseteq S_1 + S_2$.

Ahora, sea Z un subespacio de V que contiene a S_1 y a S_2. Entonces, precisamente por ser subespacio debe contener todas las sumas $s_1 + s_2$, con $s_1 \in S_1$ y $s_2 \in S_2$. Es decir, $S_1 + S_2 \subseteq Z$. Así, $S_1 + S_2$ es el mínimo subespacio de V que contiene a $S_1 \cup S_2$; en otras palabras, $S_1+S_2=\text{Gen}(S_1 \cup S_2)$.

3.5 Definición.

Sea $S_1,\ldots,S_n$ subespacios del espacio vectorial V sobre el campo K. Entonces, $S_1+\ldots+S_n=\{x\in V/x=s_1+\ldots+s_n, s_i\in S_i, i=1,\ldots,n\}$.

3.6 Proposición.

Si $S_1,\ldots,S_n$ son subespacios del espacio vectorial V sobre el campo K; entonces, la suma $S_1+\ldots+S_n$ es el menor subespacio de V que contiene a S_i, $i=1,\ldots,n$ y, por lo tanto, a $\cup S_i$, $i=1,\ldots,n$; es decir, $S_1+S_2+\ldots+S_n=\text{Gen}(S_i), i=1,\ldots,n$.

Demostración.

Apoyándose en la prop. 3.4, pág. 9, la demostración se hace por inducción sobre n, el número de subespacios de V.

3.7 Proposición.

Sea V un espacio vectorial sobre el campo K; $X, Y \subseteq V$; U, un subespacio de V. Entonces,

1. Si $X \subseteq Y$, $Gen(X) \subseteq Gen(Y)$.
2. $Gen(U) = U$.
3. $Gen(X \cup Y) = Gen(X) + Gen(Y)$.

Demostración.

Para 1. $X \subseteq Y$, $Y \subseteq Gen(Y)$ implican $X \subseteq Gen(Y)$ y, por definición de generado por X, tenemos que $Gen(X) \subseteq Gen(Y)$.

Para 2. Por definición de generado por U, $U \subseteq Gen(U)$. - Por esa misma definición y por ser U un subespacio de V, $Gen(U) \subseteq U$. En consecuencia, $Gen(U) = U$.

Para 3. $X \subseteq Gen(X)$ y $Y \subseteq Gen(Y)$ implican $X \cup Y \subseteq Gen(X) \cup Gen(Y)$. Además, $Gen(X) \cup Gen(Y) \subseteq Gen[Gen(X) \cup Gen(Y)] = Gen(X) + Gen(Y)$. Por lo tanto, $X \cup Y \subseteq Gen(X) + Gen(Y)$ y como $Gen(X \cup Y)$ es - el mínimo subespacio vectorial que contiene a $X \cup Y$,
$Gen(X \cup Y) \subseteq Gen(X) + Gen(Y)$. (1)
Por otra parte, $X \subseteq X \cup Y$ y $Y \subseteq X \cup Y$ implican $Gen(X) \subseteq Gen(X \cup Y)$ y $Gen(Y) \subseteq Gen(X \cup Y)$; luego, $Gen(X) \cup Gen(Y) \subseteq Gen(X \cup Y)$. Como $G(X) + G(Y)$ es el mínimo subespacio vectorial de V que contiene a $Gen(X) \cup Gen(Y)$ tenemos que $Gen(X) + Gen(Y) \subseteq Gen(X \cup Y)$. (2)
Combinando (1) y (2) tenemos que $Gen(X \cup Y) = Gen(X) + Gen(Y)$.

3.8 Definición.

Sea X un subconjunto del espacio vectorial V sobre el campo K. Si $X = \emptyset$, se define la combinación lineal de sus

elementos como $\bar{o}$. Si $X \neq \emptyset, x_1,\ldots,x_n \in X$ y $\alpha_1,\ldots,\alpha_n \in K$; entonces, $x = \alpha_1 x_1+\ldots+\alpha_n x_n = \sum_1^n \alpha_i x_i$ es llamada una combinación lineal de $x_1,\ldots,x_n$. $\alpha_1 x_1,\ldots,\alpha_n x_n$ son los términos de la combinación lineal. Cada término no nulo es llamado no trivial. x es el valor de la combinación.
Dos combinaciones lineales de vectores de X son idénticas, si consisten de los mismos términos no triviales e iguales, si tienen el mismo valor.
Una combinación lineal con, al menos, un término no trivial es llamada no trivial.

3.9 Proposición.

Sea X un subconjunto del espacio vectorial V sobre el campo K. El conjunto de todas las combinaciones lineales de vectores de X es, precisamente, Gen(X).

Demostración.

Sea L(X) el conjunto de todas las combinaciones lineales de vectores de X, para $X \neq \emptyset$. $\forall x \in X$, $1_o x = x$ es una combinación lineal de vectores de X; luego, $X \subseteq L(X)$.
Ahora, mostraremos que L(X) es un subespacio de V. Suponiendo que v_1 y v_2 son combinaciones lineales de los mismos vectores $x_1,\ldots,x_n$ podemos escribir

$v_1 = \alpha_1 x_1+\ldots+ \alpha_n x_n$, con $\alpha_i \in K$.
$v_2 = \beta_1 x_1+\ldots+ \beta_n x_n$, con $\beta_i \in K$.

Luego, $v_1+v_2 = (\alpha_1 + \beta_1)x_1+\ldots+(\alpha_n + \beta_n)\, x_n$ y

$$kv_1 = (k\alpha_1)x_1+\ldots+(k\alpha_n)x_n, \text{ para cualquier } k \in K,$$

están en $L(X)$. Así, $L(X)$ es un subespacio de V y, por definición de generado por X, $Gen(X) \subseteq L(X)$.

Pos otra parte, como $X \subseteq Gen(X)$ y ésta es un subespacio de V, toda combinación lineal de elementos de X está en $Gen(X)$. Así, $L(X) \subseteq Gen(X)$. En conclusión, $L(X) = Gen(X)$.

Si $X = \emptyset$, definimos $L(X) = \{\bar{o}\}$.

3.10 Definición.

Sea V un espacio vectorial sobre el campo K. Un conjunto de vectores $\{x_i\} \subseteq V$, con i recorriendo un conjunto de índices, es llamado un conjunto de generadores si, para cada $x \in V$, existe un número finito de vectores $x_1,\ldots,x_n$ en $\{x_i\}$ tales que x es una combinación lineal de $x_1,\ldots,x_n$.

3.11 Ejemplos de Generadores.

1. Para el espacio vectorial de todos los polinomios en la indeterminada x, $\{1,x,x^2,\ldots,x^n,\ldots\}$ es un conjunto de generadores. Asimismo, $\{1,x,x^2,\ldots,x^n\}$ es un conjunto de generadores para el espacio vectorial de los polinomios de grado menor o igual que n.

2. $\{e_1,e_2,\ldots,e_n\}$ con

$$\begin{aligned} e_1 &= (1,0,0,\ldots,0) \\ e_2 &= (0,1,0,\ldots,0) \\ &\vdots \\ e_n &= (0,0,0,\ldots,1) \end{aligned}$$

es un conjunto generador del espacio vectorial $\mathbb{R}^n$.

4. Dependencia e Independencia Lineales.

4.1 Definición.

Sea V un espacio vectorial sobre el campo K. $X \subseteq V$ es linealmente dependiente si y sólo si existe $Y \subseteq X$ tal que $Gen(Y) = Gen(X)$. X es linealmente independiente si y sólo si no es linealmente dependiente.

Como $\emptyset$ no tiene subconjuntos propios es linealmente independiente.

Sea $x \in V$, $x \neq \bar{o}$. Entonces, $\{x\}$ es linealmente independiente pues su único subconjunto propio es $\emptyset$ y $Gen(\emptyset)=\{\bar{o}\}\neq Gen(\{x\})$.
Sea $X \subseteq V$, con $\bar{o} \in X$. Entonces, X es linealmente dependiente ya que $Y = X-\{\bar{o}\} \subset X$, $Y\cup\{\bar{o}\} = X$ y

$Gen(X) = Gen(Y\cup\{\bar{o}\}) = Gen(Y) + Gen(\{\bar{o}\})$; por 3, prop. 3.7, pág. 10

$\qquad = Gen(Y) + \{\bar{o}\}$; por 2, prop. 3.7, pág. 10.

$\qquad = Gen(Y)$.

4.2 Proposición.

Sea V un espacio vectorial sobre el campo K y $X \subseteq V$. Entonces, las siguientes proposiciones son equivalentes:

1. X es linealmente dependiente.
2. Algún vector $x \in V$ es el valor de una combinación lineal de vectores en $X - \{x\}$.

3. Algún vector de Gen(X) es el valor de más de una combinación lineal de vectores de X.
4. Hay una combinación lineal no trivial de vectores de X de valor $\bar{o}$.

Demostración.

1 $\Rightarrow$ 2. Sea $Y \subset X$ tal que Gen(Y) = Gen(X). Como Y es un subconjunto propio de X, existe $x \in X$ y $x \notin Y$. Así, $Y \subseteq X - \{x\} \subset X$. De estas inclusiones se deduce que $Gen(X) = Gen(Y) \subseteq Gen\ (X-\{x\}) \subset Gen(X)$ y que $Gen(X)=Gen(X-\{x\})$. Como $x \in X \subseteq Gen(X)$, $x \in Gen(X-\{x\})$ y se concluye 2.

2 $\Rightarrow$ 3. Sea x una combinación lineal de vectores de $X - \{x\}$. En esta combinación no aparece x. Como $x=1_\circ x$, x es el valor de, al menos, dos combinaciones lineales distintas de vectores de X, y siendo $x \in X \subseteq Gen(X)$, x satisface 3.

3 $\Rightarrow$ 4. Sea $x \in Gen(X)$ con $x=\alpha_1 x_1+\ldots+\alpha_n x_n=\beta_1 x_1+\ldots+\beta_n x_n$ tal que los vectores $x_1,\ldots,x_n$ sean todos distintos entre sí. Sin pérdida de generalidad podemos considerar que solamente $\alpha_1 \neq \beta_1$. Entonces, $\bar{o} = (\alpha_1 - \beta_1)x_1+\ldots+(\alpha_n - \beta_n)x_n$ y, puesto que $\alpha_1-\beta_1\neq 0$, $(\alpha_1 - \beta_1)x_1+\ldots+(\alpha_n - \beta_n)x_n$ es una combinación lineal no trivial de vectores de x con valor $\bar{o}$ y se cumple 4.

4 $\Rightarrow$ 1. Sea $\bar{o} = \alpha_1 x_1+\ldots+\alpha_n x_n$ una combinación lineal

de vectores de X con $\alpha_1 \neq 0$. Entonces,

$$x_1 = \frac{-\alpha_2}{\alpha_1} x_2 + \ldots + \frac{-\alpha_n}{\alpha_1} x_n \qquad (5)$$

Sea $Y = X - \{x_1\}$. Entonces, $Gen(Y) \subseteq Gen(X)$. Como $x_2,\ldots,x_n \in Y$, $x_1 \in Gen(X)$ de acuerdo con (5) y como $Y \subseteq Gen(Y)$ se sigue que $Y \cup \{x_1\} = X \subseteq Gen(Y)$ y X es linealmente dependiente.

4.3 Proposición.

Sea V un espacio vectorial sobre el campo K y $X \subseteq V$. Las siguientes proposiciones son equivalentes:

1. X es linealmente independiente.
2. Ningún vector $x \in X$ es el valor de una combinación lineal de vectores de $X - \{x\}$; es decir, para cada $x \in X$, $x \notin Gen\,(X-\{x\})$.
3. Cada vector de Gen(X) es el valor de una sola combinación lineal de vectores de X.
4. Toda combinación lineal de vectores de X de valor $\bar{o}$ es trivial.

Demostración.

La proposición es una consecuencia directa de la proposición 4.2.

5. Bases y Dimensión.

5.1 Definición.

Sea V un espacio vectorial sobre el campo K. Todo conjunto generador de V linealmente independiente es llamado una base de V.

La base del subespacio $\{\bar{o}\}$ es $\emptyset$ ya que Gen($\emptyset$)= $\{\bar{o}\}$ y $\emptyset$ es linealmente independiente.

Sea K el cuerpo de los números reales. Entonces, el conjunto generador de K^n, $\{e_1,\ldots,e_n\}$ el cual es linealmente independiente, es una base para K^n. Es la llamada base canónica.

$\{x^\circ, x^1, x^2,\ldots\}$ es una base del espacio vectorial P de todos los polinomios sobre $\mathbb{R}$ en la indeterminada x y $\{x^\circ, x^1,\ldots,x^n\}$, una base del subespacio de P que consiste de todos los polinomios de grado menor o igual que n.

5.2 Proposición.

Sea V un espacio vectorial sobre el campo K y $A=\{v_1,\ldots,v_n\}$, un conjunto de vectores de V. Entonces, A es un conjunto linealmente independiente o algún vector v_i de A puede ser escrito como una combinación lineal de los vectores de menor subíndice.

Demostración.

Si A es linealmente independiente ninguno de sus vectores puede ser escrito como una combinación lineal de los otros. Supongamos que A es linealmente dependiente; entonces, de acuerdo con la prop. 4.2, pág. 14 , $\alpha_1 v_1 + \ldots + \alpha_n v_n = \bar{0}$. Sea k el mayor entero para el cual $\alpha_k \neq 0$. Como $\alpha_i = 0$, para $i > k$, $\alpha_1 v_1 + \ldots + \alpha_k v_k = 0$ y

$$v_k = \alpha_k^{-1} (-\alpha_1 v_1 - \alpha_2 v_2 - \ldots - \alpha_{k-1} v_{k-1}) = (-\alpha_k^{-1} \alpha_1) v_1 + \ldots + (-\alpha_k^{-1} \alpha_{k-1}) v_{k-1}.$$

El teorema queda probado.

5.3 Proposición.

Sea V un espacio vectorial sobre el campo K, $A = \{v_1, \ldots, v_n\}$, un conjunto generador de W, un subespacio de V y $B = \{v_1, \ldots, v_k\}$, con $k \leq n$, un conjunto linealmente independiente. Entonces, podemos encontrar un subconjunto de A, $C = \{v_1, \ldots, v_k, v_{i_1}, \ldots, v_{i_r}\}$, linealmente independiente, que también genera a W.

Demostración.

Si A es linealmente independiente el teorema está probado. Si no lo es, eliminemos de A el primer vector v_j que sea una combinación lineal de los vectores de menor índice.

Como B es linealmente independiente, $j > k$. El subconjunto así determinado $\{v_1,\ldots,v_k,\ldots v_{j-1},v_{j+1},\ldots,v_n\}$ tiene $n - 1$ elementos y es claro que el subespacio que generan está contenido en W. Además, ese subespacio es igual a W ya que, cada $w \in W$, puede ser escrito como una combinación lineal de $v_1,\ldots,v_n$, combinación en la que podemos reemplazar v_j por una combinación lineal de $v_1,\ldots v_{j-1}$. Por lo tanto, w puede ser escrito como una combinación lineal de $v_1,\ldots,v_{j-1}, v_{j+1},\ldots,v_n$. Continuando con este proceso obtenemos un conjunto $\{v_1,\ldots,v_k,v_{i_1},\ldots,v_{i_r}\}$, que aún genera a W. Como en este conjunto no hay elementos que sean combinaciones lineales de otros de menor índice, de acuerdo con la prop. 5.2, pág. 17, debe ser un conjunto linealmente independiente.

5.3.1 <u>Corolario.</u>

Sea V un espacio vectorial de dimensión finita sobre el campo K. Entonces, V contiene un conjunto finito $A = \{v_1,\ldots,v_n\}$, linealmente independiente, que genera a V.

<u>Demostración.</u>

Como V tiene dimensión finita es generado por un número finito de elementos, digamos, $u_1,\ldots,u_m$. De acuerdo con la prop. 5.3, pág.18 podemos encontrar un número finito de -

tales vectores, los cuales denotamos con $v_1,\dots,v_n$, linealmente independientes que generan a V.

El enunciado de este corolario también puede ser expresado de la manera siguiente: Si V es un espacio vectorial de dimensión finita sobre el campo K y si $\{u_1,\dots,u_m\}$ genera a V; entonces, existe un subconjunto de ese generador, $\{v_1,\dots,v_n\}$ que constituye una base de V.

5.4 Proposición.

Sea V un espacio vectorial sobre el campo K. Si $\{v_1,\dots,v_n\}$ es una base de V y si $\{w_1,\dots,w_m\} \subseteq V$ es un conjunto linealmente independiente; entonces, $m \leq n$.

Demostración.

Cada vector de V y, en particular w_m, puede ser expresado como una combinación lineal de los elementos de $A = \{v_1,\dots,v_n\}$. Luego, $B = \{w_m, v_1,\dots,v_n\}$ es un conjunto linealmente dependiente que genera a V ya que A lo genera, por hipótesis. Por lo tanto, existe un subconjunto de B, $C = \left\{w_m, v_{i_1},\dots,v_{i_k}\right\}$, con $K \leq n - 1$, que es una base para V, por 5.3.1, pág. 19. Para formar esta base hemos cambiado, al menos, un vector $v \in A$ por w_m. Formemos el conjunto $D = \left\{w_{m-1}, w_m, v_{i_1},\dots,v_{i_k}\right\}$. D es un generador de V y, de acuerdo con el mismo corolario 5.3.1, de este conjunto linealmente dependiente podemos extraer una base $\left\{w_{m-1}, w_m, v_{j_1}, \ldots, v_{j_s}\right\}$, con $s \leq n - 2$.

Repitiendo este proceso llegamos a tener una base de la forma $E = \{w_2, \ldots, w_{m-1}, w_m, v_\alpha, v_\beta, \ldots\}$. Ya que w_1 no puede ser expresado como una combinación lineal de $w_2, \ldots, w_{m-1}$ y E es una base, E debe tener como elemento algún $v \in A$. Para tener la base E hemos introducido $m - 1$ elementos w y, por cada uno de ellos se ha eliminado, al menos, un elemento $v \in A$; no obstante, todavía hay en E algún elemento de A. Por lo tanto, $m-1 \leq n-1$ y $m \leq n$.

5.4.1 Corolario.

Sea V un espacio vectorial de dimensión finita sobre el campo K. Entonces, dos bases cualesquiera de V tienen el mismo número de elementos.

Demostración.

Sea $A = \{v_1, \ldots, v_n\}$ y $B = \{w_1, \ldots, w_m\}$ dos bases de V. Entonces, como B es un conjunto linealmente independiente, de acuerdo con la prop. 5.4, pág. 20, $m \leq n$. Con el mismo argumento concluimos que $n \leq m$. Luego, $m = n$.

5.5 Definición.

Sea V un espacio vectorial sobre el campo K. Definimos la dimensión de V de la manera siguiente:

dim V = 0, si V = $\{\bar{o}\}$.

= n, si V tiene una base con n vectores.

= ∞, si no ocurre alguno de los casos anteriores.

Si dim V = 0 ó dim V = n, decimos que V tiene dimensión finita. En el último caso, la dimensión es infinita.

Sea U = Gen ($\{x\}$), x $\in$ V. Como $\{x\}$ es una base de U, dim U = 1.

Sea P el espacio vectorial de los polinomios sobre $\mathbb{R}$ en una indeterminada x y $P_n \subseteq P$, el espacio vectorial sobre $\mathbb{R}$ de los polinomios de grado menor o igual que n.

Como $\{x^\circ, x^1, x^2, \ldots\}$ y $\{x^\circ, x^1, \ldots, x^n\}$ son, respectivamente, bases de P y P_n, dim P = ∞ y dim P_n = n + 1.

5.6 Proposición.

Sea V un espacio vectorial sobre el campo K, con $V \neq \{\bar{o}\}$, y $A=\{v_1,\ldots,v_n\}$ una base de V. Entonces, cualquier vector x $\in$ V tiene una representación única como combinación lineal de los vectores de A.

Demostración.

Supongamos que x $\in$ V tiene las dos representaciones

$x = \sum_{1}^{n} \alpha_i v_i$, $x = \sum_{1}^{n} \beta_i v_i$, con α_i, $\beta_i \in K$. Entonces,

$\bar{0} = \sum_{1}^{n} (\alpha_i - \beta_i) v_i$, $\alpha_i = \beta_i$ y, la representación de x es única.

6. Isomorfismos de Espacios Vectoriales.

6.1 Definición.

Sea U y V dos espacios vectoriales sobre el campo K. La función $h: U \to V$ es llamada un homomorfismo entre espacios vectoriales si y sólo si, $\forall\, x_1, x_2 \in U$ y $\forall\, k \in K$,

1. $h(x_1 + x_2) = h(x_1) + h(x_2)$ y
2. $h(kx_1) = k\, h(x_1)$.

Cuando h es biyectiva se dice que h es un isomorfismo de espacios vectoriales y U, V son llamados espacios vectoriales isomorfos.

6.2 Proposición.

Sea V un espacio n-dimensional sobre el campo K y K^n, el espacio vectorial de todas las n-uplas formadas con elementos de K. Entonces, V es isomorfo con K^n.

Demostración.

Sea $B = \{v_1, \ldots, v_n\}$ una base de V. Como cualquier -

vector $x \in V$ tiene una representación única $x = \sum_{1}^{n} \alpha_i v_i$, con $\alpha_i \in K$, para cada $x \in V$, dado B, hay sólo una n-upla $(\alpha_1,\ldots,\alpha_n) \in K^n$ asociada con x; es decir, existe una función biyectiva $f: V \to K^n$, con $f(x) = (\alpha_1,\ldots,\alpha_n)$.

Mostremos que f es un homomorfismo entre V y K^n. Consideremos los dos vectores $x=\sum_{1}^{n} \alpha_i v_i$ e $y=\sum_{1}^{n} \beta_i v_i$ de V. Entonces, $x + y = \sum_{1}^{n} (\alpha_i + \beta_i)v_i$. Para esta suma tenemos que

$f(x + y) = (\alpha_1 + \beta_1,\ldots,\alpha_n + \beta_n)=(\alpha_1,\ldots,\alpha_n)+(\beta_1,\ldots,\beta_n)=f(x)+f(y)$ y,

para $kx = \sum_{1}^{n} (k\alpha_i)v_i$, con $k \in K$, que

$f(kx) = (kv_1,\ldots,kv_n) = k(v_1,\ldots,v_n) = kf(x)$. El teorema está probado.

Como una consecuencia inmediata de prop. 6.2, pág. 23 tenemos

6.2.1 Corolario.

Cada espacio vectorial sobre el campo K y de dimensión finita n es isomorfo con el espacio vectorial $\mathbb{R}^n$.

El espacio vectorial de todos los polinomios en una indeterminada de grado menor o igual que n sobre el campo real $\mathbb{R}$ es isomorfo con K^{n+1}.

7. Espacios Vectoriales con Producto Interno.

La definición de espacio vectorial incluye un campo conmutativo K. Cuando K es el campo de los números reales el espacio es llamado espacio vectorial real. Es un espacio de esta clase que se definen los conceptos de norma, distancia y ortogonalidad.

7.1 Definición.

Sea V un espacio vectorial real. Diremos que V es un espacio real con producto interno si existe una función $g: V \times V \to \mathbb{R}$, con $g(x,y) = \langle x,y\rangle$, que satisface, $\forall x, y \in V$ y $\forall \alpha \in \mathbb{R}$, los siguientes axiomas:

H1. $\langle x,x\rangle \geq 0$, con $\langle x,x\rangle = 0$ si y sólo si $x = \bar{o}$.

H2. $\langle \alpha x,y\rangle = \alpha \langle x,y\rangle$.

H3. $\langle x,y\rangle = \langle y,x\rangle$.

H4. $\langle x+y,z\rangle = \langle x,z\rangle + \langle y,z\rangle$.

Sea el espacio vectorial $\mathbb{R}^n$ y $g:\mathbb{R}^n x \mathbb{R}^n \to \mathbb{R}$, con $g(x,y) = \sum_1^n x_i y_i$, $x=(x_1,\ldots,x_n)$ e $y=(y_1,\ldots,y_n)$. Entonces, $\mathbb{R}^n$ y la función g forman un espacio vectorial con producto interno llamado el n-espacio Euclideano.

Sea V el espacio vectorial de las funciones reales contínuas definidas en el intervalo $[a,b] \subseteq \mathbb{R}$ y $\langle , \rangle : V x V \to \mathbb{R}$, con $\langle f,g\rangle = \int_a^b f(t)\, g(t)\, dt$. V y la función $\langle , \rangle$ también for -

man un espacio con producto interno.

7.2 Definición.

Sea V un espacio vectorial real con producto interno y $x \in V$. Definimos la longitud o norma de x, la cual representamos con $\| x \|$, como la función

$\| \ \| : V \to \mathbb{R}$, con $\| x \| = \sqrt{<x\ ,\ x>}$.

Un espacio vectorial sobre el cual se ha definido una norma es llamado espacio vectorial normado.

$x \in V$ es un vector unitario si y sólo si $\| x \| = 1$. Si x no es un vector unitario podemos transformarlo en unitario haciendo $u = \frac{x}{\| x \|}$. Al hacer esto decimos que x ha sido normalizado.

7.3 Proposición. Desigualdad C.B.S (Cauchy-Bunyakovskii-Schwarz).

Sea V un espacio vectorial con producto interno. Entonces, $\forall x, y \in V$, $|<x,y>| \leq \| x \| \ \| y \|$.

Demostración.

Si $x = \bar{0}$ ó $y = \bar{0}$ los dos miembros de $|<x,y>| \leq \| x \| \ \| y \|$ son iguales a cero y la proposición es cierta. Supongamos, pues, que ni x ni y son el vector nulo.

Si $u = \frac{x}{\|x\|}$ y $v = \frac{y}{\|y\|}$; entonces, $\|u\| = \|v\| = 1$ y

$0 \leq \|u-v\|^2 = \langle u-v, u-v\rangle = \langle u,u\rangle + \langle u,-u\rangle + \langle -v,u\rangle + \langle -v,-v\rangle$,

$0 \leq \|u\|^2 + \|v\|^2 - 2\langle u,v\rangle = 2-2\langle u,v\rangle = 2(1-\langle u,v\rangle)$. Luego,

$0 \leq 1-\langle u,v\rangle$, $\langle u,v\rangle \leq 1$, $\langle \frac{x}{\|x\|}, \frac{y}{\|y\|}\rangle \leq 1$ y $\langle x,y\rangle \leq \|x\|\ \|y\|$.

Ahora, reemplazando x por -x tenemos que $\langle -x,y\rangle \leq \|-x\|\ \|y\|$, $-\langle x,y\rangle \leq \|x\|\ \|y\|$ y $\langle x,y\rangle \geq -\|x\|\ \|y\|$. Por lo tanto, $|\langle x,y\rangle| \leq \|x\|\ \|y\|$.

7.4 <u>Proposición.</u>

Sea V un espacio real normado. Entonces, $\forall x,y \in V$ y $\forall \alpha \in \mathbb{R}$,

N1. $\|x\| \geq 0$, con $\|x\| = 0$ si y sólo si $x = \bar{o}$.

N2. $\|\alpha x\| = |\alpha|\ \|x\|$.

N3. $\|x+y\| \leq \|x\| + \|y\|$. (Desigualdad triangular).

N4. $|\,\|x\| - \|y\|\,| \leq \|x-y\|$.

<u>Demostración.</u>

N1 es una consecuencia inmediata de H1.

Para N2. $\|\alpha x\| = \sqrt{\langle \alpha x, \alpha x\rangle} = \sqrt{\alpha^2 \langle x,x\rangle} = |\alpha|\ \|x\|$.

Para N3. $\|x+y\| = \sqrt{\langle x+y, x+y\rangle}$ implica

$$\|x+y\|^2 = \langle x,x\rangle + \langle y,y\rangle + 2\langle x,y\rangle$$

$$= \|x\|^2 + \|y\|^2 + 2\langle x,y\rangle$$

Ahora, por la desigualdad C.B.S tenemos que

$\|x+y\|^2 \leq \|x\|^2 + \|y\|^2 + 2\|x\| \; \|y\| = (\|x\| + \|y\|)^2$. Luego,
$\|x+y\| \leq \|x\| + \|y\|$.

Para N4. $\|x\| = \|(x-y)+y\| \leq \|x-y\| + \|y\|$ implica $\|x\| - \|y\| \leq \|x-y\|$. Similarmente, tenemos que $\|y\| - \|x\| \leq \|x-y\|$. Luego, $\|x\| - \|y\| \geq -\|x-y\|$. Así, $|\, \|x\| - \|y\| \,| \leq \|x-y\|$.

Sobre un espacio vectorial es posible definir más de una norma. Por ejemplo, sobre el espacio vectorial $\mathbb{R}^n$ - se definen las siguientes normas:

$\|x\|_1 = |x_1| + \ldots + |x_n|$

$\|x\|_2 = \text{máx}\ (|x_1|, \ldots, |x_n|)$. (la norma sup.)

$\|x\|_3 = (x_1^2 + \ldots + x_n^2)^{1/2}$. (la norma Euclídea)

para $x = (x_1, \ldots, x_n)$.

Para $C'[0,1]$, el espacio vectorial de todas las funciones de valor real continuas definidas sobre $[0,1]$ las dos funciones siguientes son normas:

$\| \; \|_1 : C[0,1] \to \mathbb{R}$, con $\|f\|_1 = \text{máx}\ |f(x)|$, $x \in [0,1]$.

$\| \; \|_2 : C[0,1] \to \mathbb{R}$, con $\|f\|_2 = \left[\int_0^1 f^2(x)dx\right]^{1/2}$, $x \in [0,1]$.

7.5 Definición.

Sea V un espacio vectorial real sobre el cual se ha definido un producto interno. Diremos que $x, y \in V$ son - ortogonales si $<x,y> = 0$. Es obvio que el vector nulo $\bar{o}$, es ortogonal con x, $\forall\, x \in V$.

7.6 Definición.

Sea V un espacio vectorial real sobre el cual se ha - definido un producto interno y $A = \{x_1,\ldots,x_n\} \subseteq V$. A es llamado un conjunto ortonormal si

1. $\|x_i\| = 1$, $\forall i$, $i = 1,2,\ldots,n$.
2. $<x_i,x_j> = 0$, $\forall i \neq j$.

En un conjunto ortonormal cada vector es unitario y, además, ortogonal con cada uno de los otros vectores.

7.7 Definición.

Sea V un espacio vectorial real sobre el cual se ha definido un producto interno. Una base es llamada ortogo nal si sus elementos son, mutuamente, ortogonales. Si, a demás, cada vector es unitario se dice que la base es orto normal.

8. Espacios Métricos.

8.1 Definición.

Sea L un conjunto no vacío. La función $d: L \times L \to \mathbb{R}$

es llamada una métrica o función distancia si, $\forall x,y,z \in L$, satisface los siguientes axiomas:

D1. Si $x \neq y$; entonces, $d(x,y) > 0$.

D2. $d(x,x) = 0$.

D3. $d(x,y) = d(y,x)$.

D4. $d(x,z) \leq d(x,y) + d(y,z)$.

$d(x,y)$ es la distancia entre los puntos x,y.

8.2 <u>Proposición.</u>

Si $d: L \times L \to \mathbb{R}$ es una función distancia; entonces, $|d(x,z) - d(y,z)| \leq d(x,y)$.

<u>Demostración.</u>

$d(x,z) \leq d(x,y) + d(y,z)$ implica que $d(x,z)-d(y,z) \leq d(x,y)$. Por otra parte, $d(y,z) \leq d(y,x) + d(x,z) = d(x,y)+d(x,z)$ implica que $d(y,z) - d(x,z) \leq d(x,y)$; luego, $d(x,z) - d(y,z) \geq -d(x,y)$. Así, tenemos que $|d(x,z) - d(y,z)| \leq d(x,y)$.

8.3 <u>Definición.</u>

Un espacio métrico es un par (L,d) formado por un conjunto no vacío, L y por una métrica, d, definida sobre L. Para $L = \mathbb{R}^n$, $d: \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, con $d(x,y) = \left\{ \sum_1^n (x_i - y_i)^2 \right\}^{1/2}$ y

$$d: \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}, \text{ con } d(x,y) = \sum_1^n |x_i - y_i|$$

son dos métricas definidas sobre $\mathbb{R}^n$; luego, para ambos casos, el par $(\mathbb{R}^n, d)$ es un espacio métrico.

Para $L=C[0,1]$, $d:L\times L\to\mathbb{R}$, con $d(f,g)=\int_0^1 |f(x)-g(x)|dx$ es una métrica.

En consecuencia, $(C[0,1], d)$ es también un espacio métrico.

8.4 Proposición.

Sea V un espacio vectorial real normado. Si definimos $d:V\times V\to\mathbb{R}$, con $d(x,y) = \|x - y\|$; entonces, (V,d) es un espacio métrico.

Demostración.

La función d es una métrica. En efecto, si $x,y\in V$, con $x\neq y$, $(x-y)\in V$ y $\|x-y\| > 0$; si $x=y$, $\|x-y\| = \|\bar{o}\| = 0$. Así, d satisface las condiciones D1 y D2.

Como $\|x-y\| = |-1|\ \|x-y\| = \|(-1)(x-y)\| = \|y-x\|$, $d(x,y) = d(y,x)$.
Finalmente, tenemos que
$d(x,z) = \|x-y + y-z\| \leq \|x-y\| + \|y-z\| = d(x,y) + d(y,z)$.
El teorema está probado.

d es llamada la métrica inducida por la norma. Cualquier espacio vectorial normado puede ser convertido en métrico por medio de la métrica inducida.

Algunas veces, representaremos un espacio métrico con el símbolo del conjunto que forma dicho espacio y consideraremos que d es la métrica asociada.

8.5 Definiciones.

Ahora, recordamos muchos conceptos fundamentales de los espacios métricos los cuales, citaremos con frecuencia. Sea, entonces, (L,d) un espacio métrico. Para $x \in L$ y $r > 0$, los conjuntos $B(x,r)=\{y \in L/d(y,x)<r\}$, $\bar{B}(x,r)=\{y \in L/d(y,x) \leq r\}$ son las bolas abierta y cerrada de centro x y radio r. $V \subseteq L$ es un vecindario de $x \in L$ si V contiene una bola cerrada de centro x y radio r. De acuerdo con esta definición $\bar{B}(x,r)$ y $B(x,r)$ son vecindarios de x.

El punto $x \in U \subseteq L$ es un punto interior de U si existe un vecindario de x completamente contenido en U. $x \in U$ es un punto exterior de U si existe un vecindario de x disjunto con U. $x \in U$ es un punto frontera de U si cada vecindario de x contiene, al menos un punto que pertenece a U y, al menos, un punto que no pertenece a U. La colección de todos los puntos frontera de U es la frontera de U.

Un conjunto es abierto si cada uno de sus elementos es un punto interior. El conjunto de todos los puntos interiores de U, es llamado el interior de U; lo representaremos con U°. Un conjunto es cerrado si su complemento es abierto.

$x \in U \subseteq L$ es un punto de acumulación de U si cada -

conjunto abierto que contiene a x contiene, al menos, un punto de U diferente de x. El conjunto de todos los puntos de acumulación de U es llamado el conjunto derivado de U. La clausura del conjunto U es la unión de U con su conjunto derivado. Lo representaremos con $\bar{U}$. En topología se muestra que un conjunto es cerrado si y sólo si contiene todos sus puntos de acumulación. Mostraremos esta proposición en $\mathbb{R}^n$.

Una colección de conjuntos abiertos O_i, i recorriendo un conjunto de índices, es un cubrimiento abierto de U si $U \subseteq \bigcup_i O_i$. El conjunto U es compacto si cada cubrimiento abierto de U contiene una subcolección finita que también es un cubrimiento de U.

Una sucesión de puntos de L, $\{x_n\}$ converge a $z \in L$, si para cada $\varepsilon > 0$ existe $N \in \mathbb{N}$ tal que si $n \in \mathbb{N}$ y $n > N$; entonces, $d(x_n, z) < \varepsilon$. Cuando esto ocurre escribimos $\lim_{n\to\infty} x_n = z$ y llamamos a z un punto límite. Una sucesión es convergente si existe, al menos, un punto al cual converge.

Se dice que el conjunto U es acotado si existe un número real M tal, que $\forall\ x \in U$, $\|x\| \leq M$.

Si (V,d) y (V_1,d_1) son dos espacios métricos; entonces, la biyección $f: V \to V_1$ es llamada una isometría si $d_1(f(x), f(y)) = d(x,y)$; $\forall x,y \in V$. Se dice que una isometría preserva las distancias.

Dos espacios métricos son isométricos si existe una isometría definida de uno sobre el otro.

8.6 Proposición.

Si (V,d) es un espacio métrico y si $A \subseteq V$, con $A \neq \emptyset$; entonces, (A,d'), con d' definida por $d'(x,y)=d(x,y), \forall x,y \in A$, es un espacio métrico. Este espacio es llamado un subespacio métrico de (V,d).

Demostración.

Puesto que $A \neq \emptyset$ para probar que (A,d') es un espacio métrico sólo es necesario mostrar que d' es una métrica. Esta demostración es inmediata ya que, d' es una restricción de d, sobre $A \subseteq V$.

8.7 Proposición.

Sea (V,d) un espacio métrico y $A \subseteq V$ un conjunto compacto. Entonces, si $F \subseteq A$ es cerrado en V, F es compacto.

Demostración.

Sea O_i, con i recorriendo un conjunto de índices, un cubrimiento abierto de F. Como F es cerrado, su complemento en V, CF, es abierto. Luego, si a la colección de los O_i agregamos CF obtendremos un recubrimiento abierto de A. Por ser A compacto existe una subcolección finita R de este recubrimiento que aún cubre a A y, por consiguiente, a F. Si CF no está en R hemos encontrado una subcolección finita de los O_i que recubre a F. Si CF está en R lo podemos descartar y tener, todavía, un cubrimiento finito de F. Por lo tanto, F es compacto.

8.8 Proposición.

Sea V un espacio métrico y $A \subseteq V$ un conjunto compacto. Entonces, todo subconjunto infinito de A tiene un punto de acumulación en A.

Demostración.

Sea $X \subseteq A$ un conjunto infinito. Supongamos que ningún punto de A es punto de acumulación de X; entonces, $\forall\, x \in A$ existe una bola abierta $B(x,r_x)$ tal que $B(x,r_x) \cap X = \emptyset$, si $x \notin X$ ó $B(x,r_x) \cap X = \{x\}$, si $x \in X$.

La colección de bolas $B(x,r_x)$, con $x \in A$, es un cubrimiento abierto de A y, como A es compacto, un número finito de ellas $B(x_1,r_{x_1}),\ldots, B(x_n,r_{x_n})$, lo recubren.

Puesto que $X \subseteq A$ este número finito de bolas, también recubre a X; pero esto es una contradicción ya que X es infinito y cada una de tales bolas contiene, a lo sumo, un punto de X.

8.9 Proposición.

Sea el espacio métrico $\mathbb{R}^n$. Entonces, $A \subseteq \mathbb{R}^n$ es cerrado si y sólo si contiene todos sus puntos de acumulación.

Demostración.

Supongamos que a es un punto de acumulación del conjunto cerrado A. Puesto que cada bola abierta $B(a,r)$ contiene puntos de A y CA es abierto, a no puede ser un punto de CA. Luego, $a \in A$.

Recíprocamente, supongamos que A contiene todos sus puntos de acumulación. Si $b \in C_A$, b no es punto de acumulación de A de modo que, existe una bola abierta $B(b,r)$ que no contiene puntos de A. Como b es un punto cualquiera de CA, éste conjunto es abierto y, en consecuencia, A es cerrado.

8.10 Proposición.

Sea V un espacio métrico y $\{x_n\}$ una sucesión de puntos de V. Si V es compacto, $\{x_n\}$ tiene una subsucesión convergente en V.

Demostración.

Sea S el conjunto de imágenes de la sucesión $\{x_n\}$. Si S es finito podemos definir una subsucesión constante de $\{x_n\}$ la cual, obviamente, será convergente.

Supongamos S infinito. Puesto que, por hipótesis, V es compacto, S tiene un punto de acumulación $x \in V$. (prop. 8.8, página 35 . Las bolas $B(x,1), B(x,1/2), \ldots, B(x,1/k), \ldots$ contienen, respectivamente, los puntos $x_{n_1}, x_{n_2}, \ldots, x_{n_k}, \ldots$ de S, con $n_2 > n_1, \ldots, n_k > n_{k-1}, \ldots$ Luego, podemos encontrar $x_{n_k} \in [B(x,1/k) \cap S]$, $\forall k \in \mathbb{N}$. De aquí se deduce que

$\lim\limits_{k\to\infty} x_{n_k} = x$; pues, $d(x_{n_k}, x) < \frac{1}{k}$, $\forall k \in \mathbb{N}$.

8.11 Proposición.

Si $A \subseteq \mathbb{R}^m$ y $B \subseteq \mathbb{R}^n$ son conjuntos compactos; entonces, $A \times B$ es un subconjunto compacto de $\mathbb{R}^{m+n}$.

Demostración.

Sea $\{(a_n, b_n)\}$ una sucesión de puntos de $A \times B$. Puesto que A es un conjunto compacto, la sucesión $\{a_n\}$ tiene una subsucesión $\{a_{n_i}\}$ que converge a un punto $a \in A$. Similarmente, $\{b_n\}$ tiene una subsucesión $\{b_{n_i}\}$ que converge a un punto $b \in B$. Así, $\{(a_{n_i}, b_{n_i})\}$ es una subsucesión de $\{(a_n, b_n)\}$ que converge al punto $(a,b) \in A \times B$.

8.11.1 Corolario.

Sea $I_1 = [a_1, b_1], \ldots, I_n = [a_n, b_n]$, con $I_i \subseteq \mathbb{R}$, $i = 1, \ldots, n$. Entonces, $I_1 x \ldots x I_n$ es un subconjunto compacto de $\mathbb{R}^n$.

8.12 Definición.

Una sucesión $\{x_n\}$ de puntos de un espacio métrico V es llamada acotada si el conjunto de imágenes de la sucesión es acotado; es decir, si el conjunto de imágenes está contenido en una bola $\bar{B}(x,r)$, para algún $x \in V$.

8.13 Proposición.

Toda sucesión acotada en $\mathbb{R}^n$, tiene una subsucesión convergente.

Demostración.

Si $\{x_n\}$ es una sucesión acotada, el conjunto de imágenes está contenido en el compacto
$I \subseteq \mathbb{R}^n$, con $I = I_1 x \ldots x I_n$ e $I_i \subseteq \mathbb{R}$, $i = 1, \ldots, n$, un intervalo -

cerrado y acotado. Luego, de acuerdo con la prop. 8.10, pág. 36, $\{x_n\}$ tiene una subsucesión convergente.

8.14 Teorema de Heine - Borel.

Sea el espacio métrico $\mathbb{R}^n$. Entonces, $A \subseteq \mathbb{R}^n$ es compacto si y sólo si A es cerrado y acotado.

Demostración.

Supongamos que A es compacto. Entonces, si a es un punto de acumulación de A para cada entero positivo n, existe un punto a_n tal que $d(a_n,a) < 1/n$. Luego, a es el único punto límite de la sucesión $\{a_n\}$. Como A es compacto $\{a_n\}$ debe tener un punto límite en A; por lo tanto, $a \in A$ y A es cerrado.

Ahora, supongamos que A es cerrado y acotado. Seleccionemos $r > 0$ de modo que $A \subseteq B(0,r)$. Si $I= [-r,r]$; entonces, A es un subconjunto cerrado del compacto IxIxI...xI (n veces) y, de acuerdo con la prop. 8.7, pág. 34, A es compacto.

8.15 Proposición.

Sea el espacio métrico $\mathbb{R}^n$. Entonces, A es compacto si y sólo si todo subconjunto infinito de A tiene un punto de acumulación en A.

Demostración.

Supongamos A compacto. Entonces, de acuerdo con la prop. 8.8,pág. 35, todo subconjunto infinito de A tiene un punto de acumulación en A.

Ahora, debemos mostrar que A es compacto; es decir, cerrado y acotado. Haremos la prueba por contradicción.

Supongamos que A no es acotado. Entonces, $\forall n \in \mathbb{N}$ existe $x_n \in A$ tal que $\|x_n\| > n$. Luego, el conjunto $S = \{x_n\}$, $n \in \mathbb{N}$ es infinito y, por hipótesis, existe un punto $y \in A$ que es punto de acumulación de S. Como existe $m \in \mathbb{N}$ tal que $m > 1 + \|y\|$; tenemos que, $-\|y\| > 1-m$, $\|x_m-y\| \geq \|x_m\| - \|y\| > \|x_m\|+1-m > m+1-m = 1$. De este resultado deducimos que la bola $B(y,1)$ sólo puede contener un número finito de puntos de S; pero, esto es una - contradicción al hecho de que si p es un punto de acumulación de un conjunto E, todo vecindario de p contiene - infinitos puntos de A (Principios de Análisis Matemático, Rudin, segunda edición, pág. 40, teorema 2.22). Por lo - tanto, A debe ser acotado.

Mostremos ahora que A es cerrado. Sea x un punto de acumulación de A. Las bolas abiertas $B(x,1/k)$, $K \in \mathbb{N}$ contienen puntos distintos $x_k \in A$. Sea $S = \{x_k\}$, $k \in \mathbb{N}$. Es claro que x es un punto de acumulación de S y; por ser éste un conjunto infinito, por hipótesis, S tiene un punto de acumulación en A. Mostremos que x es el único punto de acumulación de S. Supongamos que $y \neq x$ es otro punto de acumulación de S; entonces, para $x_k \in S$, tenemos que $\|y-x\| \leq \|y-x_k\| + \|x_k-x\| < \|y-x_k\| + 1/k$. Así, si escogemos $k_\circ \in \mathbb{N}$ tal que $k \geq k_\circ$ implique $\frac{1}{2}\|y-x\| > \frac{1}{k}$ tendremos que $\frac{1}{2}\|y-x\| < \|y-x_k\|$, lo cual indica que $x_k \notin B(y,r)$, con $r = \frac{1}{2}\|y-x\|$, si $k \geq k_\circ$. Luego, y no puede ser un punto de acumulación de S. Como x se tomó arbitrariamente y no puede haber para S un punto de acumulación que no esté en A, A contiene todos sus puntos de acumulación y por ello, es cerrado.

8.15.1 Corolario. Teorema de Weierstrass.

Todo subconjunto infinito y acotado del espacio métrico $\mathbb{R}^n$ tiene un punto de acumulación en $\mathbb{R}^n$.

8.16 Definición.

Sea (V,d) un espacio métrico. Una sucesión de puntos de V, $\{x_i\}$, es una sucesión de Cauchy si, para todo $\varepsilon > 0$ existe $N \in \mathbb{N}$ tal que si $m,n \in \mathbb{N}$ y $m,n \geq N$; entonces, $d(X_m,X_n) < \varepsilon$.

Cuando V es un espacio vectorial normado $\{x_i\}$ es una sucesión de Cauchy si para todo $\varepsilon > 0$ existe $N \in \mathbb{N}$ tal que si $m,n \in \mathbb{N}$ y $m,n \geq N$; entonces, $\|X_m - X_n\| < \varepsilon$.

8.17 Proposición.

En todo espacio métrico toda sucesión convergente es una sucesión de Cauchy.

Demostración.

Sea $\{x_n\}$ una sucesión convergente con $\lim x_n = a$. Entonces, para cada $\varepsilon > 0$ existe $N \in \mathbb{N}$ tal que $n \in \mathbb{N}$ y $n > N$ implica $d(x_n,a) < \varepsilon/2$.

Ahora, si $p, q \in \mathbb{N}$ y si $p,q \geq N$ tendremos que $d(x_p,a) < \varepsilon/2$ y $d(x_q,a) < \varepsilon/2$. Luego, $d(x_p,x_q) \leq d(x_p,a) + d(x_q,a) < \varepsilon/2 + \varepsilon/2 = \varepsilon$ y la sucesión es de Cauchy.

8.18 Definición.

Un espacio métrico V es llamado completo o de Banach si cada sucesión de Cauchy $\{x_n\}$ de puntos de V converge a un punto $x \in V$.

8.19 Proposición.

El espacio métrico $\mathbb{R}^n$ es completo.

Demostración.

Sea $\{x_n\}$ una sucesión de Cauchy en $\mathbb{R}^n$. Mostraremos que la sucesión es convergente.

Si el conjunto I de imágenes en la sucesión es un conjunto finito, a partir de un determinado subíndice las imágenes serán constantes y la sucesión, convergente.

Sea I un conjunto infinito. Mostremos que I es acotado. Puesto que la sucesión es de Cauchy existe $N \in \mathbb{N}$ tal que si $n \geq N$, $d(x_n, X_N) < 1$. Así, si

$M = \text{máx}\ \{\|x_1\ \ , \ldots, \|X_N\|\}$; entonces, $\|x_n\| < M + 1$, $\forall n \in \mathbb{N}$;

es decir, $I \subseteq B\ (0, M+1)$. Luego, I es un conjunto infinito y acotado en $\mathbb{R}^n$ y por eso, tiene un punto de acumulación $x \in \mathbb{R}$. Veamos ahora, que $\lim\limits_{n\to\infty} x_n = x$. Para cualquier $\varepsilon > 0$ existe $N \in \mathbb{N}$ tal que si $m,n \geq N$; entonces, $d(x_m, x_n) < {}^{\varepsilon}/2$. Además, la bola abierta $B(x, {}^{\varepsilon}/2)$ contiene un punto x_m de I, con $m \geq N$. Luego, si $n \geq N$, $d(x_n, x) \leq d(x_n, x_m) + d(x_m, x) < {}^{\varepsilon}/2 + {}^{\varepsilon}/2 = \varepsilon$. Así, $\lim\limits_{n\to\infty} x_n = X$.

9. Funciones Contínuas en Espacios Métricos.

9.1 Definición.

Sea (V_1, d_1) y (V_2, d_2) dos espacios métricos,

$A \subseteq V_1$, $x_\circ \in A$ y $f\colon A \to V_2$, una función. Decimos que f es contínua en $x_\circ$ si, para cada $\varepsilon > 0$ existe $\delta > 0$ tal que si $x \in A$ y si $d_1(x,x_\circ) < \delta$; entonces, $d_2(f(x), f(x_\circ)) < \varepsilon$.

Esta definición también puede ser expresada de la manera siguiente: f es contínua en $x_\circ \in A$ si, para cada bola abierta $B(f(x_\circ),\varepsilon)$ existe una bola abierta $B(x_\circ,\delta)$ tal que si $x \in (B(x_\circ,\delta) \cap A)$; entonces, $f(x) \in B(f(x_\circ),\varepsilon)$.

Diremos que f es contínua si es contínua en cada punto de su dominio A.

9.2 Proposición.

Sea V_1, V_2, dos espacios métricos y $f\colon V_1 \to V_2$, una función. Entonces, las siguientes afirmaciones son equivalentes:

1. f es contínua en V_1.
2. $f^{-1}(O)$ es un conjunto abierto, para todo conjunto abierto $O \subseteq V_2$.
3. $f^{-1}(C)$ es un conjunto cerrado, para todo conjunto cerrado $C \subseteq V_2$.

Demostración.

$1 \Rightarrow 2$. Sea O un conjunto abierto de V_2 y $x_\circ \in f^{-1}(O)$. Entonces, $f(x_\circ) \in O$ y, por ser este un conjunto abierto, existe una bola abierta $B(f(x_\circ),\varepsilon) \subset O$. Ahora, por ser f contínua en V_1 lo es en $x_\circ$ y; por lo tanto, existe $\delta > 0$ tal que $x \in B(x_\circ,\delta)$ implica $f(x) \in B(f(x_\circ),\varepsilon)$. Así, $f(B(x_\circ,\delta)) \subseteq B(f(x_\circ),\varepsilon)$. Luego,

$B(x_\circ,\delta) \subset f^{-1}[f(B(x_\circ,\delta)] \subseteq f^{-1}(B(f(x_\circ,\varepsilon)) \subseteq f^{-1}(O)$ y $x_\circ$, un punto arbitrario de $f^{-1}(O)$ es un punto interior; por eso, $f^{-1}(O)$ es un conjunto abierto.

$2 \Rightarrow 1$. Sea $\varepsilon > 0$ y $x_\circ \in V_1$, un punto cualquiera. Puesto que la bola $B(f(x_\circ),\varepsilon)$ es un conjunto abierto de V_2, $f^{-1}(B(f(x_\circ),\varepsilon)$ es un conjunto abierto de V_1. Ahora, como

$x_\circ \in f^{-1}(B(f(x_\circ),\varepsilon))$, existe $\delta > 0$ tal que $B(x_\circ,\delta) \subset f^{-1}(B(f(x_\circ),\varepsilon)$ y, en consecuencia, $f(B(x_\circ,\delta)) \subseteq B(f(x_\circ),\varepsilon)$

Luego, f es contínua en $x_\circ$ y, por la arbitrariedad con que fue seleccionado $x_\circ$, f es contínua en V_1.

$2 \Rightarrow 3$. Si O es un conjunto abierto, el complemento de O, CO, es cerrado. Como por hipótesis $f^{-1}(O)$ es abierto, $Cf^{-1}(O)$ es cerrado. De la igualdad $f^{-1}(CO) = Cf^{-1}(O)$ deducimos 3. Con un razonamiento similar se muestra que $3 \Rightarrow 2$.

Hemos mostrado que $1 \Longleftrightarrow 2$ y $2 \Longleftrightarrow 3$; luego, $1 \Longleftrightarrow 3$.

9.3 <u>Proposición.</u>

Sea V_1, V_2, espacios métricos, $A \subseteq V_1$ y $f:A \to V_2$, una función. Si A es un conjunto compacto y si f es contínua en A; entonces, f(A) es un subconjunto compacto de V_2.

Demostración.

Sea O_i, con i recorriendo un conjunto de índices I, un cubrimiento abierto de f(A). Como f es contínua en A, los conjuntos $f^{-1}(O_i)$, $i \in I$, son abiertos en A, cuando A es considerado un espacio métrico. (proposición 9.2, pág. 44). Además, resulta que $A \subseteq f^{-1}(\cup O_i) = \cup f^{-1}(O_i)$, $i \in I$; es decir, tenemos un cubrimiento abierto de A del cual, por ser A compacto, podemos extraer un cubrimiento finito, $f^{-1}(O_{i_1})$, ..., $f^{-1}(O_{i_n})$ que también cubre a A. Así, $A \subseteq \bigcup_{k=1}^{n} f^{-1}(O_{i_k})$ y ,

$$f(A) \subseteq f\left[\bigcup_{k=1}^{n} f^{-1}(O_{i_k})\right] = \bigcup_{k=1}^{n} f\left[f^{-1}(O_{i_k})\right] \subseteq \bigcup_{k=1}^{n} O_{i_k}.$$

Resulta, pues, que f(A) es compacto como se afirmaba.

9.3.1 Corolario.

Si en la prop. 9.2 pág. 44, $V_2 = \mathbb{R}^n$, entonces, de acuerdo con la prop. 8.14, pág. 39 , f(A) es cerrado y acotado.

9.4 Definición.

Sea V un espacio vectorial normado cuya norma es $\| \ \|$.

Entonces, decimos que la función $f\colon V \to \mathbb{R}^n$ es acotada, si existe $M > 0$ tal que $\|f(x)\| < M$, $\forall\, x \in V$.

9.5 <u>Definición.</u>

Sea (V_1,d_1), (V_2,d_2), espacios métricos y $f\colon V_1 \to V_2$, una función.

Decimos que f es uniformemente contínua en V_1 si, para cada $\varepsilon > 0$ existe $\delta > 0$ tal que si $x, y \in V_1$ y $d_1(x,y)<\delta$; entonces, $d(f(x), f(y)) < \varepsilon$.

Se deduce, inmediatamente, que toda función uniformemente contínua definida sobre un espacio métrico es contínua en tal espacio. Sólo es necesario considerar en la definición 9.5 que el punto y, después de haber sido seleccionado arbitrariamente, es fijo.

9.6 <u>Proposición.</u>

Sea (V_1,d_1), (V_2,d_2), espacios métricos; $f\colon V_1 \to V_2$, una función y $A \subseteq V_1$. Si A es compacto y f contínua en A; entonces, f es uniformemente contínua en A.

<u>Demostración.</u>

Sea $\varepsilon > 0$. Como f es contínua en cada punto $y \in A$, existe $\delta_y > 0$ tal que si $x \in [B(y,\delta_y) \cap A]$; entonces,

$d_2(f(x),\ f(y)) < {}^{\varepsilon}/2$.

La colección formada por los conjuntos $B(y, \frac{1}{2}\delta_y)$, - con $y \in A$, resulta ser un cubrimiento abierto de A y, como por hipótesis, A es compacto podemos encontrar puntos $y_1, \ldots, y_n \in A$ de modo que sea $A \subseteq \bigcup_1^n B(y_i, \frac{1}{2}\delta_{y_i})$. Ahora, - sea $\delta = \text{mín}\ \{\frac{1}{2}\delta_1, \ldots, \frac{1}{2}\delta_n\}$, x, y $\in$ A tales que $d_1(x,y) < \delta$. Es evidente que existe $y_{i_o} \in A$ tal que $x \in B(y_{i_o}, \frac{1}{2}\delta y_{i_o}) \subseteq B(y_{i_o}, \delta y_{i_o})$; por lo que, $d_2(f(x), f(y_{i_o})) < {}^{\varepsilon}/2$. Por otra parte,

$d_1(y, y_{i_o}) \leq d_1(y,x) + d_1(x, y_{i_o}) < \delta + \frac{1}{2}\delta_{y_{i_o}} \leq \delta_{y_{i_o}}$, puesto que $\delta = \text{mín}\ \{\frac{1}{2}\delta_{y_1}, \ldots, \frac{1}{2}\delta y_n\}$. Luego, $y \in [B(y_{i_o}, \delta y_{i_o}) \cap A]$, por ello tenemos que, $d_2(f(y),\ f(y_{i_o})) < {}^{\varepsilon}/2$. Así, resulta que $d_2(f(x), f(y)) \leq d_2(f(x), f(y_{i_o})) + d_2(f(y_{i_o}),\ f(y)) < {}^{\varepsilon}/2 + {}^{\varepsilon}/2 = \varepsilon$ y el teorema está probado.

9.7 <u>Proposición.</u>

Sea V un espacio métrico, $A \subseteq V$ y $f: A \to \mathbb{R}$, una función. Entonces, si A es un conjunto compacto y f, continua en A, f alcanza sus valores extremos superior e inferior.

Demostración.

El cor. 9.3.1. pág. 46 nos permite afirmar que $f(A)$ es acotado. Por el axioma del extremo superior sabemos que existen $M = \sup f(A)$ y $m = \inf f(A)$. Como m y M son puntos de acumulación de $f(A)$ y éste, es un conjunto cerrado, $m, M \in f(A)$; es decir, existen $x_1, x_2 \in A$ tales que $M = f(x_1)$ y $m = f(x_2)$.

9.8 Definición.

Sea V un espacio métrico. El diámetro de $U \subseteq V, d(U)$, con $U \neq \emptyset$, se define como la mínima cota superior de las distancias entre los puntos de U; es decir,
$d(U) = \sup \{d(x,y) \ / \ x, y \in U\}$.

Si $d(U)$ es finito; es decir, si $d(U) < \infty$ decimos que U es acotado y si $d(U) = \infty$, que U no es acotado.

9.9 Definición.

Sea (V,d) un espacio métrico. $U \subseteq V$ es un conjunto totalmente acotado si para cada $\varepsilon > 0$, existe un número finito de conjuntos
$A_1, \ldots, A_n$ tales que $U = A_1 \cup \ldots \cup A_n$, con $d(A_i) \leq \varepsilon$, $i = 1, \ldots, n$.

9.10 Proposición.

Sea (V,d) un espacio métrico. Si $U \subseteq V$ es un conjunto totalmente acotado; entonces, U es acotado.

Demostración.

Sea $\varepsilon = 1$. Por hipótesis, $U=A_1 \cup \ldots \cup A_n$, con $d(A_i)\leq 1, i=1,\ldots,n$. Seleccionemos en cada A_i un punto x_i y tomemos $\delta=\text{máx}\ d(x_j,x_i)$, $i,j = 1,\ldots,n$. Para x, $y \in A$ tenemos que $x \in A_i$ e $y \in A_j$, $i, j = 1,\ldots,n$. Luego,

$d(x,y) \leq d(x,x_j) + d(x_j,x_i) + d(x_i,y) \leq 1 + \delta + 1 = 2 + \delta$ y $d(U) < \infty$.

9.11 Proposición.

Sea (V,d) un espacio métrico. Entonces, toda sucesión de Cauchy $\{x_n\}$ de puntos de V es totalmente acotada y, en consecuencia, acotada.

Demostración.

Sea $\varepsilon > 0$. Mostraremos que existe una descomposición de $\{x_n\}$ en un número finito de conjuntos cuyos diámetros son menores que ε. Puesto que $\{x_n\}$ es una sucesión de Cauchy, existe $N \in \mathbb{N}$ tal que si $m,n \in \mathbb{N}$ y $m,n > N$; entonces, $d(x_m,x_n) < \varepsilon$. Luego, $B = \{x_{N+1}, x_{N+2},\ldots\}$ tiene un diámetro menor que ε. Así, $\{x_1\},\ldots, \{x_N\}$, B es una descomposición finita de $\{x_n\}$ en conjuntos cuyos diámetros son menores que ε. Por lo tanto, $\{x_n\}$ es totalmente acotada y, por ello, acotada.

9.12 Definición.

Sea V un espacio métrico. Una sucesión de funciones $\{f_n\}$ definida en V converge uniformemente en V hacia una función f si y sólo si para cada $\varepsilon > 0$ hay un número natural N tal que $n \geq N$ implica $d(f_n(x), f(x)) \leq \varepsilon$, $\forall\, x \in V$.

9.13 Proposición. Criterio de Cauchy sobre convergencia uniforme de funciones.

Sea V un espacio métrico. La sucesión de funciones $\{f_n\}$ definida en V converge, uniformemente en V si y sólo si para cada $\varepsilon > 0$ existe un número natural N tal que $m,n \geq N$ y $x \in V$ implica $d(f_n(x), f_m(x)) \leq \varepsilon$.

Demostración.

Supongamos que $\{f_n\}$ converge uniformemente en V hacia la función límite f. Entonces, existe un natural N tal - que $m,n \geq N$ y $x \in V$ implica que
$d(f_m(x),f(x)) \leq {}^{\varepsilon}/2$ y $d(f_n(x),f(x)) \leq {}^{\varepsilon}/2$. Luego,
$d(f_m(x),f_n(x)) \leq d(f_m(x),f(x)) + d(f_n(x),f(x)) \leq \varepsilon$.

Recíprocamente, supongamos que se cumple la condición de Cauchy. Entonces, la sucesión $\{f_n(x)\}$ converge, para -

todo x, hacia un límite que podemos llamar $f(x)$. Luego, la sucesión $\{f_n\}$ converge en V hacia f. Debemos mostrar que la convergencia es uniforme. Sea $\varepsilon > 0$ y elijamos un natural N de tal modo que, para cada $\varepsilon > 0$; $m,n \geq N$ y $x \in V$ implique $d(f_n(x), f_m(x)) \leq \varepsilon$. Ahora, fijemos n y hagamos $m \to \infty$; entonces,

$$\lim_{m\to\infty} d(f_n(x), f_m(x)) \leq \lim_{m\to\infty} \varepsilon$$

$$d(f_n(x), \lim_{m\to\infty} f_m(x)) \leq \varepsilon.$$

$$d(f_n(x), f(x)) \leq \varepsilon.$$

y f es uniformemente contínua.

10. Distancia entre conjuntos.

10.1 Definición.

Sea V un espacio métrico, $F \subseteq V$ un conjunto cerrado y $b \in V$, un punto cualquiera. Definimos la distancia de b a F, la cual denotaremos con $d(b,F)$, como la menor de las distancias de b a cada uno de los puntos de F; es decir, $d(b,F) = \inf \{d(b,x) \ / \ x \in F\}$.

Se acepta, convencionalmente, que $d(b,\emptyset) = \infty$.

10.2 Proposición.

Sea V un espacio métrico, $F \subseteq V$ un conjunto cerrado y

$b \in V$. Entonces, $d(b,F) \geq 0$.

Demostración.

Puesto que $d(x,y) \geq 0$, $\forall x, y \in V$; se sigue, inmediatamente, de la def. 10.1,pág.52,que $d(b,F) = 0$, si $b \in F$.

Mostremos que debe ser $d(b,F) > 0$, si $b \notin F$. Supongamos, junto con esta última condición, que $d(b,F) \leq 0$. Como, por definición, $d(b,F)$ no puede ser negativa, tiene que ser $d(b,F) = 0$. Ahora, por ser d un ínfimo, tenemos que $\forall\ \varepsilon > 0$ existe $y_o \in F$ tal que $d(b,y_o) < d(b,F) + \varepsilon = 0 + \varepsilon$. Así, resulta que la bola $B(b,\varepsilon)$ contiene, al menos, un punto de F distinto de b; y por eso, b es un punto de acumulación de F. Como F es un conjunto cerrado, $b \in F$; pero, esto contradice nuestra hipótesis de que $b \notin F$. Por lo tanto, si $b \notin F$, $d(b,F) > 0$.

10.3 Proposición.

Sea V un espacio métrico y $F \subseteq V$, un conjunto cerrado. Entonces, la función $f: V \to \mathbb{R}$, con $f(x) = d(x,F)$ es continua.

Demostración.

Sea a,b dos puntos de V. Para cualquiera que sea $\delta > 0$, existe un punto $x \in F$ tal que $d(a,x) \leq d(a,F) + \delta$ ya que,

$d(a,F)$ es un ínfimo y F es cerrado. Ahora, por la desigualdad triangular tenemos que $d(b,x) \leq d(a,x) + d(a,b)$ y luego, que $d(b,x) \leq d(a,F) + d(a,b)$. Por otra parte, $d(b,F) \leq d(b,x)$, $\forall\ x \in F$ ya que, $d(b,F) = \text{ínf } \{d(x,F)/x \in F\}$. Por último tenemos que $d(b,F) \leq d(a,F) + d(a,b)$ pues, $\delta > 0$ es un número arbitrario. Así, $d(b,F) - d(a,F) \leq d(a,b)$.

con un razonamiento similar; pero, intercambiando los papeles de a y b deducimos que $d(a,F) - d(b,F) \leq d(a,b)$ y luego, que $d(b,F) - d(a,F) \geq -d(a,b)$. Así, $|d(a,F) - d(b,F)| \leq d(a,b)$.

Ahora, si hacemos $d(a,b) = \varepsilon$ resulta que $|f(a)-f(b)| < \varepsilon$. De esta expresión se deduce que f es uniformemente contínua; luego, contínua.

10.4 Proposición.

Si en un espacio métrico V todas las bolas cerradas son conjuntos compactos; entonces, para $F \subseteq V$ cerrado, con $F \neq \emptyset$, y para todo $a \in V$, existe un punto $\alpha \in F$ tal que $d(a,\alpha)=d(a,F)$.

Demostración.

Hagamos $r = d(a,F)$ y consideremos la bola $\bar{B}(a,r+1)$. De acuerdo con la proposición 8.14, $B \cap F$ es un conjunto compacto. Ahora, definimos la función $f: B\cap F \to \mathbb{R}$, con $f(x)=d(a,x)$. Con un argumento similar al de la prop. 10.3, pág. 53, se deduce que f es contínua. Entonces, $f(B \cap F)$ es compacto como imagen directa de un compacto bajo una función contínua. Luego, f alcanza su valor mínimo en $B \cap F$, sea m - ese valor. Así, pues, existe $\alpha \in B \cap F$ tal que

$m = f(\alpha) = d(a,\alpha) = \text{ínf } \{d(a,x)/x \in B \cap F\}$. Ahora, podemos afirmar que

$d(a,\alpha) \leq d(a,x) \leq r + 1, \forall\, x \in B \cap F,$

$d(a,\alpha) \leq r+1 \leq d(a,x)$, para $x \in F$, $x \notin B \cap F$; y de aquí, que

$d(a,\alpha) \leq d(a,x)$, $\forall\, x \in F$. Luego, $d(a,\alpha) = d(a,F)$.

10.5 Definición.

Sea V un espacio métrico y F_1, F_2 dos subconjuntos cerrados de V. Llamaremos distancia de F_1 a F_2 al

$\text{ínf } \{d(x,y)/\ x \in F_1,\ y \in F_2\}$ y la denotaremos con $d(F_1,F_2)$.

Convencionalmente, se acepta que $d(A,\emptyset) = d(\emptyset,A) = \infty$.

10.6 Proposición.

Sea V un espacio métrico y F_1, F_2 subconjuntos de V, con F_1 compacto y F_2, cerrado. Si $F_1 \cap F_2 = \emptyset$; entonces, $d = d(F_1,F_2) > 0$. Además, si todas las bolas cerradas de V son conjuntos compactos, d es un mínimo.

Demostración.

Definamos la función $g: F_1 \to \mathbb{R}$, con $g(x_1)=d(x_1,F_2)$.

Razonando como en la prop. 10.3, pág. 53, podemos afirmar que g es contínua. Ahora, $g(F_1) \subseteq \mathbb{R}$ es compacto. Luego, g alcanza su valor mínimo, mayor que cero, sobre F_1 porque F_1 es compacto y $F_1 \cap F_2 = \emptyset$; es decir, existe $\alpha_1 \in F_1$ tal que $g(\alpha_1) = d(\alpha_1, F_2) > 0$. Así, $d = d(F_1, F_2) > 0$.

Ahora, si todas las bolas cerradas de V son conjuntos compactos, sabemos que existe $\alpha_2 \in F_2$ tal que

$d(\alpha_1, \alpha_2) = d(\alpha_1, F_2)$, $\alpha_1 \in F_1$.

Similarmente, podemos afirmar que existe $\alpha_1 \in F_1$ tal que $d(\alpha_2, \alpha_1) = d(\alpha_2, F_1), \alpha_2 \in F_2$. Luego, $d(\alpha_1, F_2) = d(\alpha_1, \alpha_2) = d(\alpha_2, F_1) = d(F_1, F_2) > 0$.

11. Isomorfismo Topológico en Espacios Vectoriales Normados.

11.1 Definición.

Sea V un espacio vectorial normado y $\| \ \|_1$, $\| \ \|_2$ dos normas definidas sobre V. Decimos que $\| \ \|_1$ y $\| \ \|_2$ son topológicamente equivalentes si existen números reales positivos m_1 y m_2 tales que $\| x \|_2 \leq m_1 \| x \|_1$ y $\| x \|_1 \leq m_2 \| x \|_2$, $\forall \; x \in L$.

Ejemplo.

En el espacio vectorial normado $\mathbb{R}^n$, las normas

$\| x \|_1 = \text{máx } |x_i|$, $i = 1,\ldots,n$;

$\| x \|_2 = |x_1|+\ldots+|x_n|$, con $x = (x_1,\ldots,x_n)$,

$\| x \|_3 = (x_1^2 +\ldots+ x_n^2)^{1/2}$

Son topológicamente equivalentes ya que, como puede fácilmente, mostrarse $\| x \|_1 \leq \| x \|_3 \leq \| x \|_2 \leq n \| x \|_1$

11.2 Proposición.

Sea V un espacio vectorial de dimensión finita. Entonces, todas las normas definidas sobre V son topológicamente equivalentes.

Demostración.

Sea $\{u_1,\ldots,u_n\}$ una base para V. Entonces, cualquier $x \in V$ puede ser representado en la forma $x=x_1u_1+\ldots+x_nu_n$. Es fácil mostrar que $\| x \| = (x_1^2+\ldots+x_n^2)^{1/2}$ es una norma en V y que la equivalencia topológica es una relación transitiva. De acuerdo con esta última afirmación, para probar la proposición, será suficiente mostrar que cualquier otra norma $\| \;\; \|_1$, definida sobre V, es topológicamente equivalente con $\| \;\; \|$.

Sea $M = \text{máx} \| u_i \|$, $i = 1,\ldots,n$; entonces,

$\|x\|_1 = \| x_1u_1+\ldots+x_nu_n\|_1 \leq |x_1| \ \|u_1\|_1+\ldots+|x_n| \ \|u_n\|_1 \leq$

$M(|x_1| +\ldots+|x_n|) \leq Mn \text{ máx } |x_i| \leq Mn\|x\|$.

Ahora, supongamos que no existe un número real m tal que $\|x\| \leq m \|x\|_1$. Entonces, correspondiente a cada entero positivo k, existe un vector y_k tal que $\|y_k\| > k \|y_k\|_1$.

Hagamos $z_k = \dfrac{y_k}{\|y_k\|}$; entonces, $\|z_k\| = 1$ y

$\|z_k\|_1 = \dfrac{\|y_k\|_1}{\|y_k\|} < \dfrac{1}{k}$. Escribamos $z_k=z_{1k}u_1+\ldots+z_{nk}u_n$. -

Puesto que $\|z_k\| = 1$, la sucesión $\{(z_{1k},\ldots,z_{nk})\}$ es acotada en el espacio Euclídeo $\mathbb{R}^n$ y, de acuerdo con la prop. - 8.13, pág. 38, dicha sucesión contiene una subsucesión que converge, digamos, a $(z_1,\ldots,z_n)$. Para simplificar la notación supongamos que tal subsucesión es la sucesión misma. Luego, si $z = z_1u_1 +\ldots+ z_nu_n$ tenemos que

$\lim\limits_{k\to\infty} \|z_k-z\| = \lim\limits_{k\to\infty} \left[(z_{1k}-z_1)^2+\ldots+(z_{nk}-z_n)^2\right]^{1/2} = 0$. Ahora, por N4 de la prop. 7.4, pág. 38, podemos escribir $\|z_k\| -\|z-z_k\|\leq\|z\|\leq\|z-z_k\|+\|z_k\|$. De aquí se sigue que $\lim\limits_{k\to\infty} \|z_k\|\leq\|z\|\leq\lim\limits_{k\to\infty} \|z_k\|$ y, por lo tanto, que

$\lim\limits_{k\to\infty} \|z_k\| = \|z\| = 1$.

Por otra parte, $\|z\|_1 \leq \|z-z_k\|_1 + \|z_k\|_1 \leq Mn\|z-z_k\|+1/k$ y, puesto que estas desigualdades son ciertas para cualquier k arbitrariamente grande, resulta que $\|z\|_1 = 0$; es decir, $z = \bar{0}$; pero, esto es una contradicción al hecho de que - $\|z\| = 1$. Así, existe, un número real positivo m tal que $\|x\| \leq m \|x_i\|$ y el teorema está probado.

Dentro del marco de esta demostración se ha estableci do una relación entre normas definidas sobre espacios vectoriales diferentes: $\| \ \|_1$ sobre V y $\| \ \|$ sobre $\mathbb{R}^n$. Sabemos, por la prop. 62, pág. 23, que $\Psi : V \to \mathbb{R}^n$, con $\Psi(x) = (x_1,\ldots,x_n)$ es un isomorfismo. Al considerar k=Mn hemos mostrado que $\|x\|_1 \leq k \|\Psi(x)\|$ y que $\|\Psi(x)\| \leq m \|x\|_1$.

11.3 Definición.

Sea U y V dos espacios vectoriales normados cuyas normas son, respectivamente, $\| \ \|$ y $\| \ \|_1$. Entonces, si $f: U \to V$ es un isomorfismo de espacios vectoriales y si existen constantes reales positivas, m y k, tales que $\|x\| \leq k \|f(x)\|_1$, $\|f(x)\|_1 \leq m \|x\|$, decimos que los espacios son topológicamente equivalentes. f preserva la estructura vectorial entre los dos espacios y también, la estructura topológica; es decir, en este último caso, preser va cualquier propiedad que dependa del concepto de conjunto abierto.

Lo que ha quedado establecido lo enunciamos como

11.4 Proposición.

Cada espacio vectorial normado de dimensión finita n es topológicamente isomorfo con $\mathbb{R}^n$.

De acuerdo con esta proposición toda propiedad que dependa del concepto conjunto abierto y que sea probada en $\mathbb{R}^n$ es, automáticamente, válida en cualquier espacio vectorial normado de dimensión finita n. Así,

11.5 Proposición.

Si V es un espacio vectorial normado de dimensión finita n; entonces: cada sucesión de Cauchy converge a un punto de V, cada subconjunto infinito acotado de V tiene un punto límite y, $A \subseteq V$ es compacto si y sólo si es cerrado y acotado.

Demostración.

Cada enunciado de esta proposición es una consecuencia inmediata de la combinación de las proposiciones 8.14, 8.15. (pág. 39), 8.19 (pág. 42) con la prop. 11.4 (pag. 60).

12. Funciones en Espacios Vectoriales Normados.

Ahora estudiamos propiedades de las funciones lineales y afines que tienen dominio y rango en espacios vectoriales normados. Estas funciones son llamadas transformaciones.

12.1 Definición.

Sea L y M dos espacios vectoriales normados. Decimos que

1. $T: U \subseteq L \to M$ es una transformación lineal si
   $\forall\ x, y \in U$ y $\alpha \in \mathbb{R}$,
   $T(x + y) = T(x) + T(y)$
   $T(\alpha x) = \alpha T(x)$.

2. $A: U \subseteq L \to M$ es una transformación afín si $\forall x \in L$,
   $A(x) = T(x) + b$, con T una transformación lineal y $b \in M$, una constante.

Cuando $M = \mathbb{R}$ las transformaciones son llamadas funcionales y, cuando $L = M = \mathbb{R}$ las funciones lineales y afines son descritas, respectivamente, con las ecuaciones $T(x) = mx$ y $A(x) = mx + b$. Observemos que si T es lineal, $T(\bar{o}) = T(0.x) = 0.T(x) = \bar{o}$ y si A es afín, $A(\bar{o}) = b$.

12.2 Proposición.

Sea $T: U \subseteq L \to M$ una función definida entre espacios vectoriales normados. Entonces, T es lineal si y sólo si $T(\alpha x + \beta y) = \alpha T(x) + \beta T(y)$; $\forall\, x, y \in U$ y para cualquier α, β en $\mathbb{R}$.

Demostración.

Si T es lineal, $T(\alpha x+\beta y) = T(\alpha x)+T(\beta y)=\alpha T(x)+\beta T(y)$, con $\alpha,\beta \in \mathbb{R}$. Si $T(\alpha x+\beta y) = \alpha T(x) + \beta T(y)$; entonces, al hacer sucesivamente, $\alpha=\beta=1$ y $\beta=0$ resulta que $T(x+y) = T(x) + T(y)$ y $T(\alpha x) = \alpha T(x)$.

12.3 Ejemplos de Funciones Lineales.

1. $I: L \to L$, con $I(x)=x$, $\forall\, x \in L$ y L, un espacio vectorial.
2. $\delta: D \to D$, con $\delta(f)$ la derivada de $f \in D$, con D el espacio vectorial de todas las funciones infinitamente diferenciables.
3. $\mathbb{I}: C \to \mathbb{R}$, con $\mathbb{I}(f) = \int_g^b f$ y C, el espacio vectorial de todas las funciones contínuas definidas sobre el intervalo $[a,b] \subseteq \mathbb{R}$.

12.4 Ejemplos de Funciones Afines.

1. $A: L \to L$, con $A(x) = x + y$, y un vector fijo del espacio vectorial L. La transformación lineal asociada con A es la identidad. A es llamada una traslación del vector y. La traslación del vector nulo; es decir, la transformación identidad es afín.
2. $A: L \to L$, con $A(x) = y + r\,(x - y)$, $y \in L$ es un punto -

fijo y $r \in \mathbb{R}$, un número fijo. A es llamada homotecia de centro y y radio r. La transformación lineal asociada con A es la homotecia vectorial $T: L \to L$, con $T(x) = rx$.

En general, con dos puntos cualesquiera $y, z \in L$ y una transformación lineal $T: L \to L$, podemos construir la transformación afín $A: L \to L$, con $A(x) = z+T(x-y)$.

12.5 Proposición.

Sea $T: U \subseteq L \to M$ una transformación lineal y $x = \sum_1^n \alpha_i x_i$, con x_i, $i = 1,\dots,n$, vectores de L y α_i, números reales. Entonces, $T(x) = \sum_1^n \alpha_i\, T(x_i)$.

Demostración.

La prueba es por inducción sobre n. Para n = 2 la demostración aparece en la proposición 12.2, pág. 61.

Si $T\left(\sum_1^k \alpha_i x_i\right) = \sum_1^k \alpha_i\, T(x_i)$; entonces, $T\left(\sum_1^{k+1} \alpha_i x_i\right) =$

$$T\left(\sum_1^k \alpha_i x_i + \alpha_{k+1}\, x_{k+1}\right) = \sum_1^k \alpha_i\, T(x_i) + \alpha_{k+1}\, T(x_{k+1}) = \sum_1^{k+1} \alpha_i\, T(x_i).$$

12.6 Proposición.

$A: L \to M$ es una transformación afín si y sólo si

$A\left(\sum_{1}^{n} \alpha_i x_i\right) = \sum_{1}^{n} \alpha_i A(x_i)$, (1) para cualquier selección de vectores $x_i \in L$ y para cualquier selección de números reales α_i, con $\sum_{1}^{n} \alpha_i = 1$.

Demostración.

Si A es afín, $A\left(\sum_{1}^{n} \alpha_i x_i\right) = T\left(\sum_{1}^{n} \alpha_i x_i\right) + b = \sum_{1}^{n} \alpha_i T(x_i) + b\sum_{1}^{n} \alpha_i$

$$= \sum_{1}^{n} \alpha_i [T(x_i)+b] = \sum_{1}^{n} \alpha_i A(x_i)$$

Supongamos que la función A satisface la condición (1) y hagamos $T(x) = A(x) - A(\bar{o})$. Para cualquier número real α,

$A(\alpha x) = A[\alpha x + (1-\alpha)\bar{o}] = \alpha A(x) + (1-\alpha) A(\bar{o})$. Luego,

$T(\alpha x) = A(\alpha x) - A(\bar{o}) = \alpha A(x) + (1-\alpha) A(\bar{o}) - A(\bar{o}) = \alpha[A(x) - A(\bar{o})] = \alpha T(x)$.

Además, $T(x_1+x_2) = T\left[2(\frac{1}{2} x_1 + \frac{1}{2} x_2)\right] = 2\, T(\frac{1}{2} x_1 + \frac{1}{2} x_2)$

$$= 2\left[A(\tfrac{1}{2} x_1 + \tfrac{1}{2} x_2) - A(\bar{o})\right] = 2\{\tfrac{1}{2}[A(x_1) - A(\bar{o})] + \tfrac{1}{2}[A(x_2) - A(\bar{o})]\}$$

$$= [A(x_1) - A(\bar{o})] + [A(x_2) - A(\bar{o})] = T(x_1) + T(x_2).$$

Así, T es una función lineal y $A(x) = T(x) + A(\bar{o})$, afín.

12.7 Definición.

Sea $f: U \subseteq L \to M$ una función entre espacios vectoriales normados. Decimos que f es contínua en $x \in U$ si para cada $\varepsilon > 0$ existe $\delta > 0$ tal que si $y \in U$ y $\|y-x\| \leq \delta$; entonces, $\|f(y)-f(x)\| < \varepsilon$.

Como hemos visto, en la sección precedente, las normas asociadas con los espacios vectoriales L y M pueden ser diferentes; pero, para representarlas usaremos el mismo símbolo $\| \ \|$, excepto en aquellos casos en los que, distinguirlas sea indispensable.

12.8 <u>Proposición.</u>

Sea L, M los espacios vectoriales normados y $T:L \to M$, una transformación lineal. Entonces, las tres propiedades que siguen son equivalentes:

1. Existe un número real $k>0$ tal que $\|t(x)\| \leq k\|x\|$, $\forall x \in L$.
2. T es contínua en x, $\forall\, x \in L$.
3. T es contínua en el origen.

<u>Demostración.</u>

$1 \Rightarrow 2$. Sea $x_o \in L$ y $\varepsilon > 0$. Si $\|x-x_o\| < \varepsilon/k$; entonces,

$$\|T(x)-T(x_o)\| = \|T(x-x_o)\| \leq k\|x-x_o\| < k\,\frac{\varepsilon}{k} = \varepsilon .$$

Así, T es contínua en x_o y, como x_o es un punto arbitrario de L, T es contínua en L.

$2 \Rightarrow 3$. Si T es contínua en x, $\forall\, x \in L$, es contínua en el origen.

$3 \Rightarrow 4$. Siendo T contínua en el origen existe $\delta > 0$ tal que si $\|x\| \leq \delta$; entonces, $\|T(x)\| < 1$. Luego, para cualquiera $x \in L$, con $x \neq \bar{o}$, tenemos que

$\|T(x)\| = \left\| T\left(\frac{\|x\|}{\delta} \cdot \frac{\delta}{\|x\|} x \right) \right\| = \frac{\|x\|}{\delta} \left\| T\left(\frac{\delta}{\|x\|} x \right) \right\|$. Ahora, como

$\left\| \frac{\delta}{\|x\|} x \right\| = \delta$, deducimos que, $\|T(x)\| < \frac{\|x\|}{\delta} .1 = k\|x\|$, con $k = \frac{1}{\delta}$.

12.9 <u>Definición.</u>

Sea $T: L \to M$ una transformación lineal entre espacios vectoriales normados y $B = \{\|T(x)\| / x \in L,\ \|x\| = 1\}$. Si B tiene una cota superior mínima diremos que la transformación lineal T es acotada.

12.10 <u>Corolario.</u>

La transformación lineal $T: L \to M$, con L y M dos espacios vectoriales normados, es contínua si y sólo si es acotada.

<u>Demostración.</u>

Si T es contínua existe $k > 0$ tal que $\|T(x)\| \leq k\|x\|$, $\forall\, x \in L$. Al considerar los vectores x de la bola unitaria con centro en el origen tenemos que, $\|T(x)\| \leq K$ y T es acotada.

Mostremos que si T es acotada; entonces, es contí - nua. Supongamos que T no es contínua; entonces, no lo - será en el origen. Luego, podemos encontrar una sucesión $\{x_i\}$, con $\|x_i\| < 1/i$ de modo que $\|T(x_i)\| > \varepsilon$, para al gún $\varepsilon > 0$. Sigue, entonces, que para todo i

$$\left\| T\left(\frac{x_i}{\|x_i\|}\right)\right\| = \frac{1}{\|x_i\|}\|T(x_i)\| > i\,\varepsilon$$

lo cual contradice el acotamiento de T sobre la bola unitaria.

12.11 <u>Proposición.</u>

Sea $T: L \to M$ una transformación lineal entre espa - cios vectoriales normados. Si L es de dimensión finita; entonces, T es contínua.

<u>Demostración.</u>

Porque todo espacio vectorial de dimensión finita es topológicamente isomorfo con $\mathbb{R}^n$ sólo es necesario probar el teorema para cuando $L = \mathbb{R}^n$. Usando la base canónica podemos representar cualquier vector unitario $x \in L$ de la siguiente manera:

$x = \sum_{i}^{n} \alpha_i e_i$, con $\alpha_i \in \mathbb{R}$. Ahora, si $m = \text{máx}\|T(e_i)\|$, $1 \leq i \leq n$, -

tendremos que

$\| T(x)\| = \| T(\sum_{1}^{n} \alpha_i e_i)\| \leq \sum_{1}^{n} |\alpha_i| \, \| T(e_i)\| \leq m\sum_{1}^{n} |\alpha_i|$ y, como $\alpha_i = \langle x, e_i \rangle$,

$|\alpha_i| = |\langle x, e_i\rangle| \leq \| x\| \, \| e_i\| = 1$, concluímos que $\| T(x)\| \leq mn$. Así,

T es acotada y, por ello contínua.

12.12 Proposición.

Sea $T: L \to \mathbb{R}$ una funcional lineal. Entonces,

$N = \{x \in L / T(x) = 0\}$ es un subespacio de L. Este subespacio es llamado el espacio nulo de T.

Demostración.

$T(\bar{o}) = T(0.x) = 0T(x) = \bar{o}$ y $\bar{o} \in N$. Por lo tanto, $N \neq \emptyset$.

Si $x, y \in N$ y $\alpha, \beta \in \mathbb{R}$; entonces, $T(\alpha x + \beta y) = \alpha T(x) + \beta T(y) = \alpha .0 + \beta .0 = 0$ y, $\alpha x + \beta y$ pertenece a N.

El teorema está probado.

12.13 Definición.

Sea L un espacio vectorial y N, un subespacio de L. N es un subespacio propio maximal de L si, para cualquier otro subespacio K de L, se tiene que si $N \subseteq K \subseteq L$; entonces, N = k ó k = L.

12.14 Proposición.

N es un subespacio propio maximal del espacio vecto -

rial normado L si y sólo si N es el espacio nulo de una - funcional lineal no idénticamente nula. El subespacio - propio maximal es cerrado si y sólo si la funcional lineal es contínua.

Demostración.

Sea N un subespacio propio maximal de L. Entonces, - existe un vector $y \in L$ que no pertenece a N. Con este vector formemos $K = \{\alpha y+x / x \in N \text{ y } \alpha \in \mathbb{R}\}$ y mostremos que este - conjunto con las operaciones suma de vectores y producto por escalar constituyen un subespacio de L.

Si $\alpha=-1$ y $x=y$, $\alpha y+x = \bar{o} \in k$. Así, $k \neq \emptyset$

Si $u,v \in k$; entonces $u+v= \alpha_1 y+x_1+\alpha_2 y+x_2=(\alpha_1+\alpha_2)y+(x_1+x_2)$, con $x_1,x_2 \in N$ y $\alpha_1,\alpha_2 \in \mathbb{R}$. Como $(\alpha_1+\alpha_2) \in \mathbb{R}$ y $x_1,x_2 \in N$ implica $(x_1+x_2) \in N$, por ser N un subespacio, deducimos - que $(u+v) \in K$.

Si $\lambda \in \mathbb{R}$ y $u \in K$; entonces, $u = \alpha y+x$, con $x \in N$, $\alpha \in \mathbb{R}$ y $\lambda u = (\lambda\alpha)y+\lambda x$. Como $\lambda\alpha \in \mathbb{R}$ y $\lambda x \in N$, por ser N un subespacio, $\lambda u \in k$.

Haciendo $\alpha = 0$ podemos ver que N está contenido propiamente en k; luego, $N \subseteq K \subseteq L$, $K = L$ y los elementos de L pueden ser escritos en la forma $\alpha y+x$.

Definamos $T: L \rightarrow \mathbb{R}$, con $T(\alpha y+x)= \alpha$, mostremos que T es una funcional lineal y N su espacio nulo.

$T[(\alpha_1 y+x_1) + (\alpha_2 y+x_2)] = T[(\alpha_1 +\alpha_2) y + (x_1+x_2)] = \alpha_1+\alpha_2 = T(\alpha_1 y+x_1) + T(\alpha_2 y+x_2)$.

$T[\lambda(\alpha_1 y+x_1)] = T(\lambda\alpha_1 y+\lambda x_1) = \lambda\alpha_1 = \lambda T(\alpha_1 y+x_1)$.

Ya vimos que N se obtiene a partir de K haciendo $\alpha = 0$. Luego, $T(0.y+x) = T(x) = 0$, con $x \in N$, y N es el espacio nulo de T.

Ahora, supongamos que T es una funcional lineal no - idénticamente nula con espacio nulo N. Sea, además, K un subespacio que contiene propiamente a N; es decir, sea $N \subseteq K \subseteq L$. Entonces, existe $y \in K$ tal que $T(y) = \alpha \neq 0$. Si hacemos $y_\circ = \frac{y}{\alpha}$, será $T(y_\circ) = \frac{1}{\alpha} T(y) = 1$.

Como K es un subespacio, K contiene todos los elementos de la forma $\beta y_\circ + x$, con $x \in N$ y $y_\circ \in K$. Seleccionemos cualquier vector $z \in L$ y hagamos $x = z - T(z)y_\circ$; entonces, - $T(x) = T(z) - T(z)T(y_\circ) = T(z)[1-T(y_\circ)] = T(z).\ 0 = 0$.

Luego, $x \in N$. Esto significa que cualquier $z \in L$ puede - ser escrito en la forma $z = T(z)y_\circ + x$ y que; por lo tanto, siendo $z \in K$, $L \subseteq K$.

Así, $K = L$ y N un subespacio propio maximal de L.

Hemos probado la primera parte del teorema. Probemos la segunda parte.

Si la funcional lineal es contínua, para cualquier sucesión $\{x_n\}$ de puntos de N que converja a x, 0 =

$$\lim_{n\to\infty} T(x_n) = T\left(\lim_{n\to\infty} x_n\right) = T(x).$$ Así, $x \in N$; es decir, N contiene sus puntos de acumulación y por eso, es cerrado.

Supongamos que el espacio nulo N de una funcional - lineal T es cerrado. Ya se ha mostrado que todo $z \in L$ puede ser escrito en la forma $z = \alpha y_0 + x$, con $y_0 \notin N, x\in N$ y $\alpha = T(z)$. Puesto que N es cerrado podemos encontrar un vecindario de y_0, $W = V(y_0, \varepsilon)$, con $W \cap N = \emptyset$. Entonces, $\forall\ x \in N$, $\| x - y_0 \| > \varepsilon$ y puesto que $-x \in N$, es $\| -x-y_0 \| = \| x+y_0 \| > \varepsilon$. Para cualquier $z \in L$.

$$\| z \| = \| \alpha y_0 + x \| = |\alpha| \, \| y_0 + \tfrac{1}{\alpha} x \| \geq |\alpha| \varepsilon = \varepsilon |T(z)| \, y, \, |T(z)| \leq \frac{\| z \|}{\varepsilon}$$

Esto muestra que T es una funcional lineal acotada; y por eso, contínua.

13. Transformaciones Lineales y Matrices.

13.1 La transformación lineal asociada con una matriz.

Sea $A = [a_{ij}]_{mxn}$ una matriz y L^n, L^m dos espacios vectoriales normados de dimensiones n y m, respectivamente.

Entonces, asociamos con la matriz A una función $T: L^n \to L^m$, con $T(x) = Ax$, para cada vector columna $x\in L^n$ y Ax, un producto de matrices.

Mostraremos que T es una transformación lineal - usando las propiedades del producto de matrices:

$T(x+y) = A(x+y) = Ax + Ay = T(x) + T(y)$,

$T(\alpha x) = A(\alpha x) = \alpha(Ax) = \alpha T(x)$; $\forall x, y \in L^n$ y $\forall \alpha \in \mathbb{R}$.

13.2 La matriz asociada con una transformación lineal.

Sea L^n y L^m dos espacios vectoriales normados con dimensiones n,m, respectivamente y $T: L^n \to L^m$, una transformación lineal. Probaremos que existe una matriz $A = [a_{ij}]_{m\times n}$ tal que $\forall\ x \in L$, $T(x) = Ax$.

Supongamos que $\{I_1,\dots,I_n\}$ y $\{J_1,\dots,J_m\}$ son bases con vectores unitarios de L^n y L^m, respectivamente. - Cualquier vector $x \in L^n$ puede ser expresado como una combinación lineal de los vectores $I_1,\dots,I_n$, así:

$x = x_1I_1+\dots+x_nI_n$, con $x_1,\dots,x_n$, las componentes de x.

Ahora, por linealidad tenemos que

$T(x) = x_1\ T(I_1)+\dots+x_n\ T(I_n)$.

Podemos escribir $T(I_1),\dots,T(I_n) \in L^m$ en términos de $J_1,\dots,J_m$; pues, existen escalares a_{ij} tales que

$$\begin{aligned}
T(I_1) &= a_{11}J_1 +\dots+ a_{m1}\ J_m \\
T(I_2) &= a_{12}J_1 +\dots+ a_{m2}J_m \qquad\qquad (*) \\
&\ \vdots \\
T(I_n) &= a_{1n}J_1 +\dots+ a_{mn}\ J_m \qquad\qquad ó
\end{aligned}$$

$$T(I_1) = \begin{bmatrix} a_{11} \\ a_{12} \\ \cdot \\ \cdot \\ \cdot \\ a_{m1} \end{bmatrix}, \quad T(I_2) = \begin{bmatrix} a_{12} \\ a_{22} \\ \cdot \\ \cdot \\ \cdot \\ a_{m2} \end{bmatrix}, \ldots, \quad T(I_n) = \begin{bmatrix} a_{1n} \\ a_{2n} \\ \cdot \\ \cdot \\ \cdot \\ a_{mn} \end{bmatrix},$$

usando vectores coordenados. Así,

$$T(x) = x_1 \begin{bmatrix} a_{11} \\ a_{12} \\ \cdot \\ \cdot \\ \cdot \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \cdot \\ \cdot \\ \cdot \\ a_{m2} \end{bmatrix} + \ldots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \cdot \\ \cdot \\ \cdot \\ a_{mn} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}x_1+\ldots+a_{1n}\,x_n \\ \cdot \\ \cdot \\ \cdot \\ a_{m1}x_1+\ldots+a_{mn}x_n \end{bmatrix} = \begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ a_{m1} & \ldots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix} = Ax$$

A, la matriz buscada, es la transpuesta de la matriz de coeficientes del sistema de ecuaciones (*).

13.3 Ejemplo.

Sea $T: \mathbb{R}^2 \to \mathbb{R}^3$ una transformación lineal con ecuaciones $T(I_1) = (1,1,1)$ y $T(I_2) = (1, -1,1)$; con $I_1=(1,0)$ e $I_2=(0,1)$. Determinemos la matriz asociada a T con respecto a las bases canónicas $\{I_1,I_2\}$ y $\{J_1,J_2,J_3\}$ de $\mathbb{R}^2$ y $\mathbb{R}^3$.

Solución.

Como $T(I_1) = (1,1,1) = J_1+J_2+J_3$ y $T(I_2)=J_1-J_2+J_3$, la matriz asociada con T es $A = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix}$.

Si $x=x_1I_1+x_2I_2$ y $T(x) = y_1J_1+y_2J_2+y_3J_3$; entonces,

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Recordemos que la matriz A, asociada con la transformación T, depende de las bases seleccionadas para los conjuntos de partida y de llegada en la definición de T. $\{k_1,k_2\}$, con $k_1=(1,1)$ y $k_2=(1,-1)$, es una base para $\mathbb{R}^2$; mientras que, $\{L_1,L_2,L_3\}$, con $L_1=(1,1,1)$, $L_2=(0,1,1)$ y $L_3=(0,0,1)$, lo es para $\mathbb{R}^3$. Para determinar la nueva matriz asociada con T debemos, otra vez, encontrar los elementos a_{ij} para los cuales, $T(k_i) = \sum_1^3 a_{ij} L_i$, $i = 1,2$.

$T(k_1)=T\,[(1,0)+(0,1)] = T(1,0)+T(0,1) = (1,1,1)+(1,-1,1) = (2,0,2)$

$=2\,(1,0,1) = 2(L_1-L_2+L_3) = 2L_1-2L_2+2L_3$.

$T(k_2)=T\,[(1,0)-(0,1)] = T(1,0)-T(0,1)=(1,1,1)-(1,-1,1)=0L_1+2L_2+0L_3$.

Por lo tanto,

$$A = \begin{bmatrix} 2 & 0 \\ -2 & 2 \\ 2 & 0 \end{bmatrix}$$

y si $x = z_1k_1 + z_2k_2$, $T(x) = w_1L_1 + w_2L_2 + w_3L_3$; entonces,

$$\begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ -2 & 2 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$$

Con el símbolo $[T]$ representaremos a la matriz asociada con una transformación lineal T para indicar que se ha - seleccionado una base para cada espacio que aparece en la definición de T.

Para que la imagen de $x \in L$, $T(x)$ pueda ser obtenida como un producto de matrices consideraremos a x como un vector columna, siempre que x aparezca en una ecuación matricial.

Cuando el producto interno de dos vectores $x, y \in \mathbb{R}^n$ deba considerarse como un producto de matrices, escribiremos $\langle x,y\rangle = x^t y$. Para ser consistentes con la convención de que x representa un vector columna deberíamos escribir $x^t = (x_1,\dots,x_n)$; pero, para evitar el uso de índices t escribiremos $x = (x_1,\dots,x_n)$ cuando no haya confusión.

14. La clase de las Transformaciones Lineales Contínuas.

14.1 Definición.

Sea L y M dos espacios vectoriales normados. Representaremos con L(L,M) la clase de todas las transformaciones lineales contínuas $T: L \to M$. Definimos, para esta clase, la suma y la multiplicación por escalar de la siguiente manera:

$+ : L(L,M) x L(L,M) \to T(L,M)$, con $+(S,T)(x)=(S+T)(x)= S(x) + T(x)$.

$\circ : K x L(L,M) \to T(L,M)$, con $\circ(\alpha,T)(x) = (\alpha T)(x) = \alpha T(x)$, con α un escalar del campo K.

T(L,M) representa la clase de todas las transformaciones lineales.

14.2 Proposición.

Si $S:L \to M$ y $T:L \to M$ son dos transformaciones lineales contínuas entre espacios vectoriales; entonces, la suma y el producto escalar definidos en 14.1 son transformaciones lineales contínuas.

Demostración.

La prueba es inmediata a partir de las definiciones de 14.1 y de la continuidad de S y T.

14.3 Proposición.

La clase L(L,M) y las operaciones suma y producto escalar definidas en 14.1 constituyen un espacio vectorial.

Demostración.

Es inmediata y la omitimos.

14.4 Proposición.

El conjunto $L(L,M)$, cuyos elementos son transformaciones lineales contínuas, puede ser descrito como la clase de las transformaciones lineales acotadas (cor. 12.10, pág. 66). Es decir, si T es una transformación lineal contínua, $B = \{\|T(x)\| / x \in L,\ \|x\| = 1\}$ está acotado superiormente por $k = 1/\delta$ (prop. 12.8, pág. 65) y, por lo tanto, tiene una cota inferior mínima. Basándonos en esta observación definimos la función

$\|\ \ \| : L(L,M) \to \mathbb{R}$, con $\|T\| = \sup \|T(x)\|$, $\|x\| = 1$. (1).

Puesto que si $T \in L(L,M)$, $\|T(x)\| \leq k\|x\|$. $\forall\, x \in L$ y siendo, por definición, $\|T\| = \sup \|T(x)\|$ el menor valor que puede tomar k es $\|T\|$; luego, podemos decir que

$\|T(x)\| \leq \|T\|\,\|x\|$, $\forall\, x \in L$.

14.5 Proposición.

La función definida en 14.4 es una norma sobre $L(L,M)$.

Demostración.

- Como $\|T(x)\| \geq 0$, $\forall\, x \in L$, es $\sup \|T(x)\| = \|T\| \geq 0$.

- Mostremos que $\|T_\circ\| = 0$ si y sólo si $T_\circ$ es la transformación nula.

  Sea $T = T_\circ$. Entonces, $T(x) = \bar{o}$, $\forall x \in L$. Luego, $\|T\| = \sup \|T(x)\| = 0$, con $x \in \bar{B}(0,1)$.

  Ahora, Sea $\|T\| = 0$. Entonces, $\sup \|T(x)\| = 0$, $\forall x \in \bar{B}(0,1)$.

  Para $x \in V$ arbitrario, $x \neq \bar{o}$, tenemos que $x = \|x\| \dfrac{x}{\|x\|}$; luego,

  $\|T(x)\| = \|x\| \left\| T\left(\dfrac{x}{\|x\|}\right)\right\| = \|x\| . \; 0 = \bar{0}$. Así,

  $T(x) = \bar{o}$, $\forall\, x \in L$ y $T = T_\circ$.

- $\|\alpha T\| = \sup \|(\alpha T)(x)\| = \sup (|\alpha| \; \|T(x)\|) = |\alpha| \sup \|T(x)\| = |\alpha| \; \|T\|$.

- $\sup \|T_1(x) + T_2(x)\| \leq \sup(\|T_1(x)\| + \|T_2(x)\|) =$

  $\sup\|T_1(x)\| + \sup \|T_2(x)\|$; luego, $\|T_1 + T_2\| \leq \|T_1\| + \|T_2\|$.

14.6 <u>Proposición.</u>

Si en la clase de las transformaciones lineales acotadas $L(L,M)$, M es un espacio de Banach; entonces, $L(L,M)$ es un espacio de Banach.

<u>Demostración.</u>

Supongamos que $\{T_n\}$ es una sucesión de Cauchy en $L(L,M)$. Para $x = \bar{o}$ se deduce, inmediatamente, que la sucesión $\{T_n(x)\}$ es de Cauchy. Sea, pues, $x \in L$, $x \neq \bar{o}$ y $\varepsilon > 0$, entonces,

$0 \leq \|T_m(x) - T_n(x)\| = \|(T_m - T_n)(x)\| \leq \|T_m - T_n\| \; \|x\|$.

Por ser $\{T_n\}$ una sucesión de Cauchy existe $N \in \mathbb{N}$ - tal que para $m,n>N$, $\|T_m-T_n\| < \frac{\varepsilon}{\|x\|}$ y, en consecuencia, $\|T_n(x)-T_m(x)\| < \varepsilon$ siempre que $m,n>N$. Así, $\{T_n(x)\}$ es una sucesión de Cauchy en M. Como por hipótesis, M es - completo, $\{T_n(x)\}$ converge a algún elemento en M; lo cual, nos permite definir la función $T: L \to M$, con $T(x) = \lim\limits_{n\to\infty} T_n(x)$.

Mostremos que $T \in L(L,M)$ y que $\{T_n\}$ converge a T.

Para $x, y \in L$ y $\alpha, \beta \in \mathbb{R}$

$T(\alpha x+\beta y) = \lim\limits_{n\to\infty} T_n(\alpha x+\beta y) = \lim\limits_{n\to\infty}[\alpha T_n(x)+\beta T_n(y)] = \alpha T(x)+\beta T(y)$ y T es lineal.

Sea k una cota para T_n (todas las sucesiones de Cau - chy son acotadas). Por la continuidad de la función norma tenemos que, para cualquier

$x \in L$, $\|T(x)\| = \left\|\lim\limits_{n\to\infty} T_n(x)\right\| = \lim\limits_{n\to\infty} \|T_n(x)\| \leq \lim\limits_{n\to\infty}\|T_n\| \|x\| \leq k\|x\|$.

Luego, T es acotada sobre la bola unitaria y por eso, continua.

Puesto que $\|T_n-T\| = \sup\|(T_n-T)(x)\|$, sobre la bola unitaria; existe para cualquier $\varepsilon > 0$ un vector unitario x tal que

$\|T_n-T\| \leq \|(T_n-T)(x)\| + \frac{\varepsilon}{2} = \|T_n(x)-T(x)\| + \frac{\varepsilon}{2}$ y como $\lim\limits_{n\to\infty} T_n(x)=T(x)$,

$\|T_n(x)-T(x)\|$ puede hacerse menor que $\varepsilon/2$ tomando n suficien temente grande. Para este n, $\|T_n - T\| < \varepsilon$ y T_n converge a T. Así, L(L,M) es un espacio de Banach.

15. El espacio dual de un espacio vectorial normado.

El conjunto de todas las funciones lineales contínuas $L(L, \mathbb{R})$ es un espacio vectorial normado llamado el espacio dual de L y es representado, usualmente, con L^*. Es decir, $L^* = \{f: L \to \mathbb{R} / f \text{ es lineal}\}$. De acuerdo con nuestra última proposición L^* es un espacio de Banach.

Sea L un espacio vectorial con producto interno; entonces, para cualquier $a \in L$, definimos $T: L \to \mathbb{R}$, con $T(x) = \langle a,x \rangle$.

15.1 Proposición.

$T: L \to \mathbb{R}$, con $T(x) = \langle a,x \rangle$ y $a \in L$, es un elemento del espacio L^*.

Demostración.

Sea $x, y \in L$ y $\alpha \in \mathbb{R}$. Entonces,

$T(x+y) = \langle a,x+y \rangle = \langle a,x \rangle + \langle a,y \rangle = T(x) + T(y)$.

$T(\alpha x) = \langle a,\alpha x \rangle = \langle \alpha x,a \rangle = \alpha \langle x,a \rangle$.

y T es una transformación lineal.

Ahora, usando la desigualdad C.B.S tenemos que $\|T(x)\| = |\langle a,x \rangle| \leq \|a\| \ \|x\|$; es decir, T es acotada en la bola unitaria; luego, T es contínua.

15.2 <u>Proposición.</u>

Sea $T: L \to M$ una transformación lineal biyectiva entre espacios vectoriales, con $T^{-1}: M \to L$, su transformación inversa. Entonces, T^{-1} es una transformación lineal.

<u>Demostración.</u>

Sea $y_1, y_2 \in M$ y $\alpha \in \mathbb{R}$. Supongamos que $x_1 = T^{-1}(y_1)$ y $x_2 = T^{-1}(y_2)$. Entonces, $y_1 = T(x_1)$ y $y_2 = T(x_2)$. Así que,

$T^{-1}(y_1+y_2) = T^{-1}[T(x_1)+T(x_2)] = T^{-1}[T(x_1+x_2)] = x_1+x_2 = T^{-1}(y_1)+T^{-1}(y_2)$,

$T^{-1}(\alpha y_1) = T^{-1}[\alpha T(x_1)] = T^{-1}[T(\alpha x_1)] = \alpha x_1 = \alpha T^{-1}(y_1)$, y T^{-1} es lineal.

15.3 <u>Proposición.</u>

Si $T: L \to M$ es una transformación lineal entre espacios vectoriales sobreyectiva y si su núcleo es $\{\bar{o}\}$; entonces, T tiene una aplicación lineal inversa.

<u>Demostración.</u>

Mostraremos que T es inyectiva. Sea $\{\bar{o}\}$ el núcleo de T y $T(x) = T(y)$, con $x, y \in L$. Entonces,

$T(x) - T(y) = T(x-y) = \bar{o}$ y $(x-y) \in \{\bar{o}\}$; luego,

$x-y = \bar{o}$, $x = y$. Así, T es biyectiva y tiene una transformación inversa.

15.4 Proposición.

Sea L un espacio vectorial normado de dimensión finita n. Entonces, el espacio dual L* es también de dimensión finita y dim L* = n.

Demostración.

Supongamos que $\{v_1,\ldots,v_n\}$ es una base para L. Determinaremos una base para L*. Sea $f_1,\ldots,f_n \in L^*$ las funcionales lineales definidas por

$$f_i(v_j) = \begin{cases} 1, & \text{si } i = j \\ 0, & \text{si } i \neq j \end{cases}$$

Mostraremos, primero, que $\{f_1,\ldots,f_n\}$ genera a L*. Sea f un elemento arbitrario de L* con $f(v_1)=k_1,\ldots,f(v_n) = k_n$. Si tomamos $g = k_1f_1+k_2f_2+\ldots+k_n f_n$; entonces,

$$\begin{aligned} g(v_i) &= (k_1f_1+k_2f_2 +\ldots+k_nf_n)(v_i), \quad i = 1,2,\ldots,n \\ &= k_1f_1(v_i)+k_2f_2(v_i)+\ldots+k_if_i(v_i)+\ldots+k_nf_n(v_i) \\ &= k_1.0+k_2.0 +\ldots+ k_i.1 +\ldots+ k_n.0 = k_i \end{aligned}$$

Así, $f(v_i) = g(v_i)$, $i = 1,2,\ldots,n$; es decir, f y g tienen los mismos valores para cada uno de los elementos de la base $\{v_1,v_2,\ldots,v_n\}$ y, en consecuencia, tienen los mismos valores sobre cada combinación lineal de los elementos de esta base, por eso, f y g son iguales sobre L. Es decir,

$f = k_1f_1 + k_2f_2 + \ldots + k_nf_n$ y $\{f_1,\ldots,f_n\}$ genera a L^*.

Ahora, mostraremos que $\{f_1,\ldots,f_n\}$ es un conjunto linealmente independiente. Supongamos que

$a_1f_1(v_1)+a_2f_2(v_2)+\ldots+a_if_i(v_i) +\ldots+ a_nf_n(v_i) = \bar{o} \circ v_i = 0$; entonces,

$a_1 .0 + a_2 .0 +\ldots+ a_i .1 +\ldots+ a_n .0 = 0.$

$a_i = 0, \ i=1,\ldots,n.$

Como $\{f_1,\ldots,f_n\}$ es una base para L^*, $\dim L^* = n$ y el teorema está probado.

$\{f_1,\ldots,f_n\}$ es llamada la base dual de $\{v_1,\ldots,v_n\}$.

15.5 <u>Definición.</u>

Recordemos que un producto escalar es llamado no degenerado si, cuando $x \in L$, con L un espacio vectorial, y $\langle x,y\rangle = 0$, $\forall y \in L$; entonces, $x = \bar{o}$.

15.6 <u>Proposición.</u>

Sea L un espacio vectorial normado de dimensión finita sobre el campo K con un producto escalar no degenerado. Entonces, $T:L\to L^*$, con $T(v) = L_v$ y $L_v: L \to \mathbb{R}$, con $L_v(x)=\langle v,x\rangle$ es una transformación lineal. Además, T es un isomorfismo y una isometría entre espacios vectoriales.

Demostración.

Sea u, v, $\in$ L y $\alpha \in$ K. Entonces,

$T(u+v)(x)=L_{u+v}=<u+v,x>=<u,x>+<v,x>=L_u(x)+L_v(x)=T(u)(x)+T(v)(x)$,

$T(\alpha u)(x) = L_{\alpha u}(x) = <\alpha u,x> = \alpha<u,x> = \alpha L_u(x) = \alpha[T(u)(x)]$, y T es lineal.

Ahora, como el producto escalar es no degenerado, el núcleo de T es $\{\bar{o}\}$ y T es inyectiva y, como dím L=dím L*, T es sobreyectiva. Así, T es biyectiva y tiene una transformación inversa $T^{-1}: L^* \to L$ y es, por eso, que T es un iso morfismo.

Para probar que T es una isometría mostremos, primero, que $\|T(v)\| = \|v\|$.

Por la desigualdad C.B.S. tenemos que

$\|T(v)(x)\| = |<v,x>| \leq \|v\| \, \|x\|$; luego,

$\sup_{x\in B(0,1)} \|T(v)(x)\| \leq \sup_{x\in B(0,1)} \|v\| \, \|x\| = \|v\|$. Así, $\|T(v)\| \leq \|v\|$.

Ahora, $\left\|T_v\left(\frac{v}{\|v\|}\right)\right\| = \left|< v, \frac{v}{\|v\|} >\right| = \frac{1}{\|v\|} \left|<v,v>\right| = \frac{1}{\|v\|} \cdot \|v\|^2 = \|v\|$.

Así, $\|T_v\| = \|v\|$.

Cuando ocurren las condiciones de la hipótesis de la proposición 15.6 cada transformación lineal de L^* puede - ser definida con un producto interno. Si $T: \mathbb{R}^n \to \mathbb{R}$ satis face dichas condiciones podemos escribir

$T(x) = \langle a,x\rangle = a^t x = [T]_x$, con $[T] := [a_1 \ldots a_n]$.

16. Derivadas en Espacios Vectoriales Normados.

Para funciones que tienen derivadas direccionales en todas las direcciones, la existencia de estas derivadas no es suficiente para conocer el comportamiento brusco o moderado de estas funciones. Con el fin de conocer este comportamiento se introduce la derivada de Fréchet, la cual, definimos a continuación.

16.1 Definición.

Sea L y M dos espacios vectoriales normados, con $\| \ \|$, la norma de L, y $U \subseteq L$, un conjunto abierto. La función $f: U \to M$ es diferenciable en $x_0 \in U$ si existe una transformación lineal $T: L \to M$ tal que, para $h \in L$ suficientemente pequeño, es

$f(x_0 + h) = f(x_0) + T(h) + \|h\| \varepsilon(x_0,h)$,

con $\varepsilon(x_0,h) \in M$ aproximándose a $\bar{0}$ cuando $\|h\| \to 0$. La transformación lineal T es llamada la derivada de f y denotada con $f'(x_0)$.

Consideremos esta definición aplicada a funciones $f: \mathbb{R}^n \to \mathbb{R}^m$, con

$$f: \begin{cases} y_1 = f_1\ (x_1,\dots,x_n) \\ y_2 = f_2\ (x_1,\dots,x_n) \\ \vdots \\ y_m = f_m\ (x_1,\dots,x_n) \end{cases}$$

y determinemos $[f'(x_\circ)]$, la matriz asociada con la transformación lineal $f'(x_\circ)$.

16.2 Proposición.

Si $f: \mathbb{R}^n \to \mathbb{R}^m$ es diferenciable en x; entonces, todas las derivadas parciales de las funciones coordenadas de f existen y,

$$[f'(x)] = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(x) & \dots & \frac{\partial f_1}{\partial x_n}(x) \\ \vdots & & \\ \frac{\partial f_m}{\partial x_1}(x) & \dots & \frac{\partial f_m}{\partial x_n}(x) \end{bmatrix}$$

Demostración.

De acuerdo con la definición de la derivada de Fréchet para f, existe una transformación lineal T: L → M, cuya matriz asociada es $[a_{ij}]_{m\times n}$, tal que para $h=te_j$(e_j un vector de la base canónica de $\mathbb{R}^n$),

$$\begin{bmatrix} f_1(x+te_j) \\ f_2(x+te_j) \\ \vdots \\ f_m(x+te_j) \end{bmatrix} - \begin{bmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_m(x) \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ t \\ \vdots \\ 0 \end{bmatrix} + |t| \begin{bmatrix} \varepsilon_1(x,te_j) \\ \varepsilon_2(x,te_j) \\ \vdots \\ \varepsilon_m(x,te_j) \end{bmatrix}$$

Entonces, para cualquier $i = 1,2,\dots,m$

$$f_i(x+te_j) - f_i(x) = ta_{ij} + |t|\varepsilon_i(x,te_j) \text{ y, para } x = (x_1,\dots,x_n) ,$$

$$\lim_{t\to 0} \frac{f_i(x_1,\dots,x_j+t,\dots,x_n) - f_i(x_1,\dots,x_n)}{t} = \frac{\partial f_i}{\partial x_j} = a_{ij},\ j=1,\dots,n.$$

16.3 <u>Ejemplos.</u>

1. Sea $f\colon \mathbb{R}^2 \to \mathbb{R}$, con $f(x,y) = (x-y)\,\mathrm{sen}(3x + 2y)$. Entonces, como

$$\frac{\partial f}{\partial x} = 3(x-y)\cos(3x+2y) + \mathrm{sen}(3x+2y) ,\quad \frac{\partial f}{\partial x}(0, \pi/3) = \frac{1}{2}(\pi + \sqrt{3}) \text{ y}$$

$$\frac{\partial f}{\partial y} = 2(x-y)\cos(3x+2y) - \mathrm{sen}(3x+2y) ,\quad \frac{\partial f}{\partial x}(0, \pi/3) = \frac{1}{6}(2\pi - 3\sqrt{3})$$

resulta que $[f'(0,\pi/3)] = \begin{bmatrix} \frac{1}{2}(\pi + \sqrt{3}) & \frac{1}{6}(2\pi - 3\sqrt{3}) \end{bmatrix}$.

2. Sea $f\colon \mathbb{R}^2 \to \mathbb{R}^2$, con $f(x,y) = (x^2-y^2, 2xy)$. Si hacemos $f_1(x,y)=x^2-y^2$ y $f_2(x,y) = 2\,x\,y$ tendremos que

$$\frac{\partial f_1}{\partial x} = 2x, \frac{\partial f_1}{\partial y} = -2y, \frac{\partial f_2}{\partial x} = 2y, \frac{\partial f_2}{\partial y} = 2x \text{ y, para } (x,y) = (2,1),$$

$$\frac{\partial f_1}{\partial x}(2,1) = 4,\ \frac{\partial f_1}{\partial y}(2,1) = -2,\ \frac{\partial f_2}{\partial x}(2,1) = 2,\ \frac{\partial f_2}{\partial y}(2,1) = 4; \text{ por consi-}$$

guiente $[f'(2,1)] = \begin{bmatrix} 4 & -2 \\ 2 & 4 \end{bmatrix}$.

16.4 <u>Proposición.</u>

Sea $U \subseteq \mathbb{R}^n$ un conjunto abierto y $f\colon U \to \mathbb{R}^m$ una función definida por funciones coordenadas que tienen derivadas parciales en cada punto de U. Entonces, $f'(x)$ existe para cada $x \in U$.

Excepto porque se trata de la derivada de Fréchet este teorema es similar al teorema de la existencia de f' en el cálculo diferencial de varias variables y puede, éste, ser consultado en los libros: Advanced Calculus -- [Creighton Buck, 1965, pág. 264] , Análisis Matematics - [T.M. Apostol, 1957, pág. 117].

16.5 Proposición.

Sea L, M dos espacios vectoriales normados y $U \subseteq L$, un conjunto abierto. Si la función $f: U \to M$ es diferenciable en $x_0 \in U$; entonces, $f - f'(x_0)$ es contínua en x_0.

Demostración.

$$\|(f-f'(x_0))(x) - (f-f'(x_0))(x_0)\| = \frac{\|(f-f'(x_0))(x)-(f-f'(x_0))(x_0)\|}{\|x - x_0\|} \|x-x_0\|$$

Luego, cuando $x \to x_0$, $\|(f-f'(x_0))(x) - (f-f'(x))(x_0)\| \to 0$ y; por eso, $f-f'(x_0)$ es contínua en x_0.

16.5.1 Corolario.

La transformación lineal $f'(x_0)$ es contínua en x_0 si y solamente si f es contínua en x_0.

Demostración.

Escribamos $f'(x_0)=f-(f-f'(x_0))$ y $f=f'(x_0)-(f'(x_0)-f)$. Entonces, si f es contínua en x_o, $f'(x_o)$ es contínua por ser, $f'(x_o)$ una diferencia de funciones contínuas. Con el mismo razonamiento se concluye que si $f'(x_o)$ es contínua, f lo será en x_o.

Si $f'(x_o)$ es contínua; entonces, $f'(x_o)$ es una transformación lineal acotada; luego, es un elemento de $L(L,M)$ y para cualquier $h \in L$ podemos escribir que
$\| f'(x_o)(h)\| \leq \|f'(x_o)\| \; \|h\|$.

16.6 Proposición. Regla de Cadena.

Sea L, M, N tres espacios vectoriales normados y $U \subseteq L$, $V \subseteq M$ dos conjuntos abiertos. Si $f: U \to M$ y $g: V \to N$, con $f(U) \subseteq V$, son funciones contínuas, si $y_o = f(x_o)$ y si $f'(x_o)$, $g'(y_o)$ existen; entonces, $H=g \circ f$ es diferenciable en x_o y $H'(x_o) = g'(y_o) \circ f'(x_o)$.

Demostración.

Las condiciones supuestas de diferenciabilidad nos permiten escribir, para h y k suficientemente pequeños, que

$f(x_o+h) = f(x_o) + f'(x_o)(h) + \|h\|\varepsilon_1(h)$,

$g(y_o+k) = g(y_o) + g'(y_o)(k) + \|k\|\varepsilon_2(k)$.

Luego, $H(x_o+h) = g[f(x_o+h)] = g\;[f(x_o)+f'(x_o)(h)+\| h \|\varepsilon_1(h)]$.

Ahora, al hacer $k = f'(x_o)(h)+\| h \|\varepsilon_1(h)$ tendremos que

$\|k\| \leq \|h\| \; [\|f'(x_\circ)\| + \|\varepsilon_1(h)\|$; de modo que, cuando $h \to \bar{o}$, $k \to \bar{o}$ y $H(x_\circ) = g[f(x_\circ)] = g(y_\circ)$. Además, por la linealidad de $g'(y_\circ)$, tenemos que

$$H(x_\circ+h) = g(f(x_\circ)+k) = g(y_\circ+k) = g(y_\circ)+g'(y_\circ)(k)+\|k\|\varepsilon_2(K)$$

$$= g(y_\circ)+g'(y_\circ)\circ f'(x_\circ)(h)+ [g'(y_\circ)(\|h\|\varepsilon_1(h) + \|k\|\varepsilon_2(k)]$$

Puesto que

$$\|g'(y_\circ)(\|h\|\varepsilon_1(h))+\|k\|\varepsilon_2(k)\| \leq \|g'(y_\circ)(\|h\|\varepsilon_1(h))\|+\|h\| \; \| [f'(x_\circ)\| + \|\varepsilon_1(h)\| \;]\varepsilon_2(k)\| \leq \|h\| \; \|g'(y_\circ)\varepsilon_1(h)+\| \; [f'(x_\circ)\|+\|\varepsilon_1(h)\| \;]\varepsilon_2(k)\|$$

cuando $h \to \bar{o}$, el miembro de la derecha tiende a cero.

Ahora, $H(x_\circ+h)-H(x_\circ)=g'(y_\circ)\circ f'(x_\circ)(h)+[g'(y_\circ)(\|h\|\varepsilon_1(h)+\|k\|\varepsilon_2(k)]$ y, como el segundo término del miembro derecho tiende a cero con $h \to \bar{o}$, se tiene que $H'(x_\circ) = g'(y_\circ)\circ f'(x_\circ)$, por definición de derivada de H en $x_\circ$.

Cuando L, M y N son espacios vectoriales de dimensión finita de modo que $g'(y_\circ)$ y $f'(x_\circ)$ pueden ser representadas con matrices, la regla de la cadena indica que la representación matricial de $H(x_\circ)$ es el producto de matrices $[H'(x_\circ)] = [g'(y_\circ)] \; [f'(x_\circ)]$.

16.7 Definición.

Sea $f: U \subseteq L \to M$ una función definida entre espacios vectoriales, contínua en U, derivable en todo U con la derivada una función contínua; es decir, con $f': U \to L(L,M)$ contínua. Entonces, f es llamada una función continuamente diferenciable.

Si además, f' es diferenciable en x_o, $f''(x)$ será una transformación lineal contínua definida de L hacia $L(L,M)$. Entonces, $f''(x)(h) \in L(L,M)$ y $[f''(x)(h)](k)$ queda definida para $k \in L$. $[f''(x)(h)]$ (k) es lineal en h y k. Es decir, $f''(x): LxL \to M$, denotada con $f''(x)(h,k)$, es una transformación bilineal.

17. Formas Bilineales.

17.1 Definición.

Sea V un espacio vectorial de dimensión finita sobre el campo K. Una forma bilineal sobre V es una aplicación $f: VxV \to K$ que satisface

1. $f(\alpha u_1 + \beta u_2, v) = \alpha f(u_1,v) + \beta f(u_2,v)$,
2. $f(u,\alpha v_1+\beta v_2)=\alpha f(u,v_1)+\beta f(u,v_2)$; $\forall \alpha,\beta \in K$, $\forall u_i, v_i \in V$, $i = 1,2$.

17.2 Ejemplos.

Sea $\emptyset$ y Ψ dos funcionales lineales definidas sobre V y $f: VxV \to \mathbb{R}$, con $f(u,v)=\emptyset(u)\Psi(v)$. Entonces, f es bilineal porque $\emptyset$ y Ψ son lineales. Una forma bilineal de esta clase es llamada producto tensorial de $\emptyset$ y Ψ.

Sea f el producto escalar sobre $\mathbb{R}^n$; es decir, sea $f(u,v) = \langle u,v \rangle$. f es una forma bilineal sobre $\mathbb{R}^n$.

Sea $A = [a_{ij}]$ nxn una matriz definida sobre el campo K. A es considerada como una forma bilineal f sobre K^n definida por

$$f(x,y) = X^t AY = (x_1, x_2, \ldots, x_n) \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \cdot \\ y_n \end{bmatrix}$$

$\forall a, b \in K$, $\forall X_i, Y_i \in K^n$.

En efecto,

$f(aX_1+bX_2,Y) = (aX_1+bX_2)^t AY = (aX_1^t+bX_2^t)AY = aX_1^t AY+bX_2^t AY = af(X_1,Y)+bf(X_2,Y)$,

y f es lineal en la primera variable.

$f(X,aY_1+bY_2) = X^t A(aY_1+bY_2) = aX^t AY_1+bX^t AY_2 = af(X,Y_1)+bf(X,Y_2)$, y f es lineal en la segunda variable. Luego, f es una forma bilineal sobre $\mathbb{R}^n$.

17.3 Formas Bilineales y Matrices.

Sea f una forma bilineal sobre el espacio vectorial V y $\{e_1, \ldots, e_n\}$, una base de V. Si u, v $\in$ V, con $u = a_1 e_1 + \ldots + a_n e_n$ y $v = b_1 e_1 + \ldots + b_n e_n$; entonces,

$$f(u,v) = f(a_1e_1 +\ldots+ a_ne_n,\ b_1e_1 +\ldots+ b_ne_n)$$

$$= a_1\ f(e_1,b_1e_1+\ldots+b_ne_n)+\ldots+a_n\ f(e_n,b_1e_1+\ldots+b_ne_n)$$

$$= a_1b_1\ f(e_1,e_1)+a_1b_2f(e_1,e_2)+\ldots+a_1b_nf(e_1,e_n)+\ldots+\ldots$$

$$+\ a_nb_1f(e_n,e_1)+a_nb_2f(e_n,e_2)+\ldots+a_nb_nf(e_n,e_n).$$

$$= \sum_{i,j=1}^{n} a_ib_j\ f(e_i,e_j)$$

Así, f está completamente determinada por los n^2 valores $f(e_i,e_j)$.

La matriz $A= [a_{ij}]_{n\times n}$ con $a_{ij}=f(e_i,e_j)$ es llamada la representación matricial de f relativa a la base $\{e_1,\ldots,e_n\}$.

Esta matriz representa a f en el sentido de que

$$f(u,v) = \sum_{i,j=1}^{n} a_ib_j\ f(e_i,e_j)=(a_1,\ldots,a_n)A\begin{bmatrix} b_1\\ b_2\\ \cdot\\ \cdot\\ \cdot\\ b_n\end{bmatrix} =[u]_e^t\ A[v]_e,\ \forall u_1v\epsilon V.$$

$[u]_e$ denota el vector columna coordenado de $u \in V$ en la base $\{e_1,\ldots,e_n\}$.

17.4 Definición.

Una forma bilineal f definida sobre el espacio vectorial V es llamada simétrica si $f(u,v) = f(v,u)$, $\forall u,v\epsilon V$.

17.5 Proposición.

Una forma bilineal f definida sobre el espacio vectorial V es simétrica si y sólo si su representación matricial es simétrica.

Demostración.

Si A es una representación matricial para f, podemos escribir $f(X,Y)=X^tAY=(X^tAY)^t=Y^tA^tX$, pues X^tAY es un escalar. Luego, si f es simétrica $Y^tA^tX=f(X,Y)=f(X,Y)=Y^tAX$. Como esta igualdad es cierta para todos los vectores x, y $\in$ V, tenemos que $A = A^t$, es decir, A es simétrica.

Recíprocamente, si A es simétrica,
$f(X,Y) = X^tAY = X^tA^tY = (X^tA^tY)^t = Y^tAX = f(Y,X)$.

17.6 Definición.

Sea V un espacio vectorial definido sobre $\mathbb{R}$. Una forma bilineal simétrica real es llamada semidefinida no negativa si $f(v,v) \geq 0$, $\forall$ v $\in$ V y, definida positiva si $f(v,v) > 0$, $\forall$ $v \in V$, $v \neq \bar{o}$.

17.7 Ejemplo.

1. Sea $f: \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, con $f(x,y) = \langle x,y\rangle$. f es simétrica ya que $f(x,y) = \langle x,y\rangle = \langle y,x\rangle = f(y,x)$. Además, f es definida positiva, pues, $f(x,x) = x_1^2+\ldots+x_n^2 > 0$, cuando $x = (x_1,\ldots,x_n) \neq \bar{o}$.

Ahora, utilizando la regla de cadena mostraremos la siguiente proposición, la cual; necesitaremos en el Capítulo III.

17.8 Proposición.

Sea V un espacio vectorial normado, $U \subseteq V$ un conjunto abierto y $f: U \to \mathbb{R}$ una función continuamente diferenciable sobre U. Además, supongamos que $f''(x)$ existe $\forall x \in U$. Entonces, para cualesquiera $x, x_o \in U$ existe $s \in (0,1)$ tal que $f(x) = f(x_o)+f'(x_o)(h)+ \frac{1}{2} f''(x_o+sh)(h,h)$, con $h = x - x_o$.

Demostración.

Dados x, $x_o \in U$, definimos $\emptyset:(a,b) \to \mathbb{R}$, con $\emptyset(t) = f(x_o + th)$ y $[0,1] \subseteq (a,b)$. Entonces, aplicando sucesivamente la regla de la cadena a $\emptyset(t)$ y a $f'(x_o+th)(h)$ resulta que

$\emptyset'(t) = f'(x_o + th)(h)$

$\emptyset''(t) = f''(x_o + th)(h,h)$.

Por otra parte, de acuerdo con el teorema de Taylor, sabemos que para $t > 0$ existe $s \in (0,t)$ tal que

$\emptyset(t) = \emptyset(o) + \emptyset'(o)t + \frac{1}{2} \emptyset''(s)t^2$.

Como $\emptyset(o)=f(x_o)$, $\emptyset'(o)=f'(x_o)(h)$ y $\emptyset''(s)=f''(x_o+sh)(h,h)$ tenemos que

$\emptyset(t) = f(x_o+th) = f(x_o)+f'(x_o)(th) + \frac{1}{2}f''(x_o+sh)(t^2h,t^2h)$

Ahora, haciendo $t = 1$ resulta

$f(x_o+h) = f(x_o) + f'(x_o)(h) + \frac{1}{2} f''(x_o+sh)(h,h)$.

CAPITULO II
CONJUNTOS CONVEXOS

La convexidad aparece como un concepto fundamental en diversos contextos teóricos: Algebra Lineal, Programación Lineal, Análisis Funcional, Geometría Analítica Local, Teoría de Juegos, Teoría Económica, etc. Su estudio, como puede apreciarse, no carece de importancia.

En este capítulo hemos reunido definiciones y proposiciones de carácter primario por lo cual, puede ser considerado como una introducción al estudio de la convexidad en espacios vectoriales.

Para evitar constantes repeticiones advertimos que todos los espacios vectoriales mencionados en las definiciones y proposiciones están asociados al campo de los números reales $\mathbb{R}$.

1. Conjuntos Convexos y Conjuntos Afines.

1.1 Definición.

Sea L un espacio vectorial. $U \subseteq L$ es llamado un conjunto

i) convexo, si $\forall$ x, y $\in$ U y $\forall \alpha \in [0,1]$, $z=\alpha x+(1-\alpha)$ y está en U.

ii) afín, si $\forall$ x, y $\in$ U y $\forall \alpha \in \mathbb{R}$, $z = \alpha x + (1-\alpha)$ y está en U.

1.2 Definición.

Sea L un espacio vectorial. $x = \sum_{1}^{n} \alpha_i x_i$, con $x_i \in L$ y $\sum_{1}^{n} \alpha_i = 1$ es llamada una combinación afín de los vectores $x_1,\ldots,x_n$. Si además, $\alpha_i \geq 0$, $i = 1,\ldots,n$, se dice que x es una combinación convexa de $x_1,\ldots,x_n$.

Para conjuntos convexos y afines $z = \alpha x + (1-\alpha) y$ es, respectivamente, una combinación lineal convexa, una combinación lineal afín. Las definiciones de conjuntos convexo y afín presuponen la formación de estas combinaciones por lo que, dichos conceptos son concebidos en espacios que admiten una estructura lineal. Luego, podemos hablar de rectas y segmentos de rectas.

Geométricamente, un conjunto es convexo si contiene al segmento de recta que une dos cualesquiera de sus puntos y afín, si contiene la recta que pasa por dos cualesquiera de sus puntos.

1.3 Ejemplos de Conjuntos Convexos.

- Los espacios y subespacios vectoriales son convexos ya que contienen toda combinación lineal de cualesquiera dos de sus puntos y, por consiguiente, la combinación convexa de ellos.

- Los conjuntos unitarios y el conjunto vacío.
- Los subespacios afines y los conjuntos afines.
- El conjunto de soluciones factibles para un problema de programación lineal.
- Cada bola cerrada (abierta) en un espacio vectorial normado. En efecto, si x_1, x_2 son elementos de $B=\{x \in L / d(x,y) \leq r\}$, una bola de centro y, radio r, con $d(x,y) = \|x-y\| \leq r$; entonces, $z = \alpha x_1 + (1-\alpha) x_2$, con $\alpha \in [0,1]$, está en B ya que
$\| z-y \| = \| \alpha(x_1-y)+(1-\alpha)(x_2-y) \| \leq \alpha \| x_1-y \| +(1-\alpha) \| x_2-y \| \leq \alpha r+(1-\alpha)r=r.$
Al considerar la desigualdad estricta queda mostrado que una bola abierta es un conjunto convexo.
- Los triángulos, paralelogramos y círculos en $\mathbb{R}^2$; las esferas, los cubos y paralelepípedos en $\mathbb{R}^3$, etc.

1.4 Ejemplos de Conjuntos Afines.

- Cada espacio vectorial, los conjuntos unitarios y el conjunto vacío.
- Las rectas en $\mathbb{R}^2$ y las rectas y planos en $\mathbb{R}^3$.

1.5 Proposición.

Sea L un espacio vectorial. $U \subseteq L$ es un conjunto afín si y sólo si es una traslación de un subespacio de L.

Demostración.

Sea $U = x_o + W = \{x_o + w / x_o \in L, w \in W\}$ una traslación del subespacio W de L. Si $x_1 = x_o + w_1$ y $x_2 = x_o + w_2$ pertenecen a U y si $\alpha \in \mathbb{R}$; entonces,

$z = \alpha x_1 + (1-\alpha)x_2 = \alpha(x_o + w_1) + (1-\alpha)(x_o + w_2) = x_o + [\alpha w_1 + (1-\alpha)w_2] = x_o + w_3$. Como W es un subespacio de L, $w_3 \in W$, $z \in U$ y éste es un conjunto afín.

Supongamos que U es un conjunto afín. Con x_o, un elemento cualquiera de U formemos el conjunto

$W = -x_o + U$. Si $w_1 = -x_o + x_1$ y $w_2 = -x_o + x_2$ pertenecen a W; entonces, para $\alpha \in \mathbb{R}$

$$\begin{aligned} w_1 + \alpha w_2 &= -x_o + \alpha x_1 - \alpha x_1 - \alpha x_o + x_1 + \alpha x_2 \\ &= -x_o + \alpha\,[(x_1 + x_2) - x_o] + (1-\alpha)x_1 \\ &= -x_o + \alpha\left[2\left(\tfrac{1}{2}x_1 + \tfrac{1}{2}x_2\right) - x_o\right] + (1-\alpha)x_1 \\ &= -x_o + z. \end{aligned}$$

Como U es afín, $y = \frac{1}{2}x_1 + \frac{1}{2}x_2$, $v = 2y + (-1)x_o$, $w = \alpha v + (1-\alpha)x_1$ son elementos que le pertenecen; luego, $-x_o + z$ está en W, W es un subespacio y $U = x_o + W$ es una traslación de W.

1.6 Proposición.

Sea L un espacio vectorial. $U \subseteq L$ es un conjunto -

convexo (afín) si y sólo si le pertenece toda combinación convexa (afín) de sus puntos.

Demostración.

Si cada combinación convexa (afín) de puntos de U - es un elemento de U, por definición, U es convexo (afín).

Supongamos que U es convexo y mostremos, por inducción sobre el número de puntos de U que aparecen en una - combinación convexa, que ella es un elemento de U. Para $n = 2$ la conclusión del teorema es consecuencia de la definición. Ahora, supongamos que la conclusión es cierta para cualquier combinación con n o menos puntos y consideremos una con $n + 1$ puntos, $x = \sum_{1}^{n+1} \alpha_i x_i$. Entonces,

$$x = \sum_{1}^{n} \alpha_i x_i + \alpha_{n+1} x_{n+1} = (1-\alpha_{n+1}) \sum_{1}^{n} \frac{\alpha_i}{1-\alpha_{n+1}} x_i + \alpha_{n+1} x_{n+1}.$$

Como $\alpha_i \geq 0$, $i=1,2,\ldots,n$ y como $\sum_{1}^{n+1} \alpha_i = 1$ tenemos que

$\frac{\alpha_i}{1-\alpha_{n+1}} \geq 0$, $1-\alpha_{n+1} = \sum_{1}^{n} \alpha_i$, $\sum_{1}^{n} \frac{\alpha_i}{1-\alpha_{n+1}} = 1$. Luego, de acuerdo con nuestra hipótesis inductiva $y = \sum_{1}^{n} \frac{\alpha_i}{1-\alpha_{n+1}} x_i$ pertenece a U y $x = (1-\alpha_{n+1})y + \alpha_{n+1} x_{n+1}$ es una combinación convexa de dos puntos de U que también pertenece a U.

La prueba, en el caso de un conjunto afín, es similar.

2. Formación de Conjuntos Convexos y Conjuntos Afines.

2.1 Proposición.

Si $\{U_i\}$, $i \in I$, es una familia de conjuntos convexos (afines); entonces, $M = \cap U_i$, $i \in I$, es convexo (afín).

Demostración.

Si $x, y \in M$; $x, y \in U_i$, $\forall i \in I$. Luego, $\alpha x+(1-\alpha)y$ está en U_i, $\forall i \in I$, con $\alpha \in [0,1]$ porque U_i es convexo. En consecuencia, $\alpha x + (1-\alpha)y$ está en $\cap U_i$, $i \in I$ y M es convexo.

El caso afín se trata de manera similar.

2.2 Proposición.

Sea $U_1, U_2, \ldots, U_m$, respectivamente, subconjuntos convexos (afines) de los Espacios Euclídeos $\mathbb{R}^{n_1}, \mathbb{R}^{n_2}, \ldots, \mathbb{R}^{n_m}$. Entonces, $\prod_1^m U_i$ es un subconjunto convexo (afín) de $\mathbb{R}^{\Sigma n_i}$.

Demostración.

Sea $(x_1^r, x_2^r, \ldots, x_m^r)$ y $(x_1^s, x_2^s, \ldots, x_m^s)$ dos puntos de $\prod_1^m U_i$, con $x_1^r, x_1^s \in U_1, \ldots, x_m^r, x_m^s \in U_m$. Para $\alpha \in [0,1]$ tenemos que

$\alpha(x_1^r, x_2^r, \ldots, x_m^r) + (1-\alpha)\ (x_1^s, x_2^s, \ldots, x_m^s) = (\alpha x_1^r + (1-\alpha)x_1^s, \ldots,$ $\alpha x_m^r + (1-\alpha)\ x_m^s)$. Ahora, este nuevo punto, dado que los conjuntos U_i son convexos para $i=1,2,\ldots,m$, tiene sus componentes

$$\alpha x_1^r + (1-\alpha)x_1^s \quad \text{en } U_1$$

$$\alpha x_2^r + (1-\alpha)x_2^s \quad \text{en } U_2$$

$$\vdots$$

$$\alpha x_m^r + (1-\alpha)\ x_m^s \text{ en } U_m,$$

de lo cual resulta que pertenece a $\prod_1^m U_i$. En otras palabras $\prod_1^m U_i$ es convexo.

El caso afín es similar.

2.3 Definición.

Sea $U_1, \ldots, U_r$, r subconjuntos del Espacio Euclídeo $\mathbb{R}^n$. Definimos la suma directa de estos subconjuntos $\sum_1^r U_i$, como el conjunto de todos los puntos $y \in \mathbb{R}^n$ tales que $y = \sum_1^r x_i$, con $x_i \in U_i$; $i = 1,2,\ldots,r$.

2.4 Proposición.

Sea $U_1, \ldots, U_r$, r subconjuntos convexos (afines) del

Espacio Euclídeo $\mathbb{R}^n$. Entonces, $\sum_1^r U_i$ es convexo (afín).

Demostración.

Hagamos $Y = \sum_1^r U_i$ y tomemos dos puntos $x, y \in Y$.

Es claro, que existen $x_i \in U_i$, $y_i \in U_i$, $i=1,\ldots,r$ tales que $x = \sum_1^r x_i$, $y = \sum_1^r y_i$. Tomemos $\alpha \in [0,1]$; entonces,

$\alpha x+(1-\alpha)\, y=\alpha \sum_1^r x_i+(1-\alpha)\sum_1^r y_i = \sum_1^r [\alpha x_i+(1-)y_i]$. Ahora, para $i=1,2,\ldots,r$ tenemos que $\alpha x_i + (1-\alpha)\, y_i$ está en U_i, pues cada U_i es convexo. Por lo tanto, $\alpha x + (1-\alpha)y$ está en Y y Y es convexo.

Seguimos el mismo patrón de prueba en el caso afín.

2.5 Definición.

Sea $U \subseteq L$ un conjunto del espacio vectorial L. La envolvente convexa de U, representada con $H(U)$, es definida como la intersección de todos los conjuntos convexos que contienen a U. Así, si $U \subseteq A$ y A es un conjunto convexo cualquiera, $H(U) \subseteq A$.

La envolvente afín de U, $A(U)$, es la intersección de todos los conjuntos afines que contienen a U. Si $U \subseteq A$ y A es un conjunto afín cualquiera, $A(U) \subseteq A$.

De acuerdo con la prop. 2.1, pág. 101, toda envolvente convexa es convexa y toda envolvente afín es afín.

2.6 Proposición.

Sea L un espacio vectorial. Para cualquier subconjunto U de L, la envolvente convexa (afín) de U es igual al conjunto de todas las combinaciones convexas (afines) de elementos de U.

Demostración.

Representemos con H(U) la envolvente convexa de U y con K(U), el conjunto de las combinaciones convexas de elementos de U. Mostraremos, por inclusión recíproca, que H(U) = K(U).

Por definición $U \subseteq H(U)$. Como H(U) es convexo, toda combinación convexa de sus puntos le pertenece por lo que, le pertenece toda combinación convexa de puntos de U. Así, $K(U) \subseteq H(U)$.

K(U) contiene a U ya que, $\forall u \in U$, $u = 1.u + 0.u$ está en K(U).

Si $x = \sum_1^n \alpha_i x_i$, $y = \sum_1^m \beta_j y_j$ son dos elementos de K(U),

$z = \lambda x + (1-\lambda)\, y = \sum_1^n \lambda\, \alpha_i x_i + \sum_1^m (1-\lambda)\, \beta_j y_j$ es otro elemento de K(U)

ya que, $\sum_1^n \lambda\, \alpha_i + \sum_i^m (1-\lambda)\beta_j = 1$. Luego, K(U) es un conjunto convexo que contiene a U; por eso, $H(U) \subseteq K(U)$.

La prueba para el caso afín es similar.

Para construir la envolvente convexa de $U \subseteq L$, L un espacio vectorial, es necesario considerar todas las posibles combinaciones convexas de puntos de U; pero, si L tiene dimensión finita n, sólo es necesario formar las combinaciones convexas de, a lo sumo, n + 1 puntos arbitrarios de U. Esta afirmación está expresada en el teorema de Carathéodory, al cual, hemos llamado proposición 2.8.

2.7 Definición.

Sea L un espacio vectorial y $U \subseteq L$. La dimensión de un conjunto afín U es la dimensión del subespacio del cual, U es una traslación. La dimensión de un conjunto convexo U es la dimensión de su envolvente afín A(U).

Antes de mostrar la validez del teorema de Carathéodory hagamos las siguientes observaciones:

(a) Si $X \subseteq Y$; entonces, $A(X) \subseteq A(Y)$ y $\dim A(X) \leq \dim A(Y)$.

Si $z \in A(X)$, $z = \sum_{1}^{r} \alpha_i x_i$, $x_i \in X$, $\sum_{1}^{r} \alpha_i = 1$. Como $X \subseteq Y$, $x_i \in Y, \forall_i$; por eso, $z \in A(Y)$. Luego, $A(X) \subseteq A(Y)$ lo cual significa que, el subespacio del cual A(X) es una traslación es un subconjunto del subespacio del cual A(Y) es una tras lación; por eso, dim $A(X) \leq$ dim $A(Y)$.

(b) $A(X) = A[H(X)]$

$X \subseteq H(X)$ implica que $A(X) \subseteq A[H(X)]$ de acuerdo con (a).

Si $z \in A[H(X)]$, $z = \sum_1^r \alpha_i x_i$, $x_i \in H(X)$, $\sum_1^r \alpha_i = 1$. Luego,

$z = \sum_1^r \alpha_i \left(\sum_1^s \beta_{ij} \; y_{ij}\right)$, con $\sum_1^s \beta_{ij} = 1$, $y_{ij} \in X$. Como

$\sum_1^r \alpha_i \left(\sum_1^s \beta_{ij}\right) = 1$, $z \in A(X)$ y $A[H(X)] \subseteq A(X)$.

(c) Sea M un subconjunto de un espacio vectorial L. Si M es afín y es la traslación del subespacio W, éste es único. Además, si $M = A(\{x_o, x_1, \ldots, x_n\})$, la envolvente afín de un conjunto finito; entonces, M es la traslación del subespacio $N = A(\{\bar{o}, x_1 - x_o, \ldots, x_n - x_o\})$.

En este caso, dim $M = n$ si y sólo si $\{x_1 - x_o, \ldots, x_n - x_o\}$ es un conjunto linealmente independiente.

. Supongamos que M es una traslación de los subespacios W y W_1 de L.

Mostraremos que $W = W_1$.

Si $v \in M$; entonces, $v = x_o + w$, con $w \in W$ y $v = y_o + w_1$, con $w_1 \in W_1$. Como $x_o \in M$ podemos escribir $x_o = y_o + w_2$, con $w_2 \in W_1$. Luego, $v = y_o + w_2 + w = y_o + w_1$, $w_1 = w_2 + w$, $w = w_1 + (-w_2)$. Siendo W_1 un subespacio

$- w_2 \in W_1$, $w \in W_1$ y $W \subseteq W_1$. Con un razonamiento análogo aplicado a $y_\circ$ se muestra que $W_1 \subseteq W$.

Puesto que M es afín, M es la traslación del subespacio $-x_\circ + M$ (prop. 1.5, pág.98). Ahora, sabemos que este subespacio es único. Debemos, por lo tanto, mostrar que $-x_\circ + M = N$. Si $x \in (-x_\circ + M)$,

$x = - x_\circ + m = - x_\circ + \sum_1^n \alpha_i x_i$, con $x_i \in M$ y $\sum_0^n \alpha_i = 1$.

Luego, $\alpha_\circ + \sum_1^n \alpha_i = 1$ implica $\alpha_\circ - 1 = - \sum_1^n \alpha_i$ y tenemos que

$$x = - x_\circ + \alpha_\circ x_\circ + \sum_1^n \alpha_i x_i = (\alpha_\circ - 1)\, x_\circ + \sum_1^n \alpha_i x_i$$

$$= - x_\circ \sum_1^n \alpha_i + \sum_1^n \alpha_i x_i = \sum_1^n \alpha_i\, (x_i - x_\circ) = \sum_1^n \alpha_i\, (x_i - x_\circ)$$ es un elemento de N. Así, $- x_\circ + M \subseteq N$.

Si suponemos que $x \in N$, un razonamiento inverso al que hemos seguido nos lleva a la conclusión de que $N \subseteq (-x_\circ + M)$.

Sea $B = \{x_1 - x_\circ, \ldots, x_n - x_\circ\}$ y Gen (B) el subespacio generado por B.

Si B es un conjunto linealmente independiente,

dim Gen(B) = dim Gen(B $\cup \{\bar{o}\}$) = n.

[Vea generadores, bases y dimensión en el capítulo I].

Como N es un subespacio generado por $\{\bar{o}, x_1 - x_\circ, \ldots, x_n - x_\circ\}$, dim N = n y, siendo M una traslación de N, es dim M = n.

Ahora, supongamos que dim M = n. Puesto que dim M = dim N = dim Gen(B) = n y N es un subespacio generado por $\{\bar{o}, x_1 - x_o, \ldots, x_n - x_o\}$ debe ser B una base para N y por consiguiente, ser un conjunto linealmente independiente.

2.8 <u>Proposición: Teorema de Carathéodory.</u>

Sea U un subconjunto del espacio vectorial L. Si la dimensión de la envolvente convexa de U, H(U) es m; entonces, para cada $z \in H(U)$, existen m + 1 puntos $x_o, \ldots, x_m$ en U tales que z es una combinación convexa de ellos.

<u>Demostración.</u>

Si $z \in H(U)$, $z = \sum_{o}^{n} \alpha_i x_i$, con $x_i \in U, \alpha_i > 0$, $\sum_{o}^{n} \alpha_i = 1$. Supongamos que n + 1, el número de términos en la combinación convexa z, es mayor que m + 1 y formemos el conjunto $B = \{x_o, \ldots, x_n\}$. De acuerdo con las observaciones que se hicieron previamente podemos afirmar que $\dim A(B) \leq \dim A(U) = \dim A(H(U)) = \dim H(U) = m \leq n - 1$ y, por consiguiente, que $\{x_1 - x_o, \ldots, x_n - x_o\}$ es un conjunto linealmente dependiente. Por lo tanto, existen constantes $\beta_1, \ldots, \beta_n$, no todas cero, tales que $\sum_{1}^{n} \beta_i (x_i - x_o) = \bar{o}$.

Sea $\beta_o = -\sum_{1}^{n} \beta_i$. Como $\sum_{1}^{n} \beta_i x_i = \sum_{1}^{n} \beta_i x_o$, al sumar $\beta_o x_o$ en ambos miembros de esta igualdad, resulta que $\sum_{0}^{n} \beta_i x_i = \beta_0 x_0 + \sum_{1}^{n} \beta_i x_0 = \bar{0}$ y $\sum_{0}^{n} \beta_i = 0$. Puesto que los escalares α_i son positivos podemos encontrar un número positivo t tal que $\gamma_i = \alpha_i - t\beta_i \geq 0$, $i = 0,1,\ldots,n$ y $\gamma_k = 0$, para algún k. tse obtiene tomando el mínimo de los cocientes $\frac{\alpha_i}{\beta_i}$, con $\beta_i > 0$. Ahora, $\alpha_i = \gamma_i + t\beta_i$ y $z = \sum_{0}^{n} \alpha_i x_i = \sum (\gamma_i + t\,\beta_i) x_i$

$$= \sum_{i \neq K} \gamma_i x_i + t \sum_{0}^{n} \beta_i x_i = \sum_{i \neq K} \gamma_i x_i$$

con $\sum_{i \neq K} \gamma_i = \sum_{0}^{n} (\alpha_i - t\beta_i) = \sum_{0}^{n} \alpha_i - t \sum_{0}^{n} \beta_i = \sum_{0}^{n} \alpha_i - 0 = \sum_{0}^{n} \alpha_i = 1.$

Así, hemos representado a z como una combinación convexa de los puntos $x_0, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n$. Si n, el número de estos puntos es igual a m + 1, el teorema está probado; si no es así, podemos repetir el proceso las veces que sea necesario hasta tener n = m + 1. Eventualmente, z puede ser representado con m + 1 puntos de U y, tal vez, con menos.

2.9 Poliedros Convexos.

Los conjuntos convexos más simples son las envolventes convexas de conjuntos finitos de puntos; es decir, conjuntos de la forma $U = H(\{x_0, \ldots, x_n\})$. Estos conjuntos son llamados poliedros convexos o politopos.

Si $x_1 - x_o, \ldots, x_n - x_o$ son linealmente independientes; entonces, $U = H(\{x_o, \ldots, x_n\})$ es llamado un n-simple con vértices $x_o, \ldots, x_n$. Un o-simple es un punto; un 1-simple, un segmento de recta; un 2-simple, un triángulo; un 3-simple, un tetraedro; etc. Un punto x de un n-simple tiene una representación única, $x = \sum_{o}^{n} \alpha_i x_i$, como combinación convexa de los vértices. Los escalares α_i son llamadas las coordenadas baricéntricas de x. El número de vértices menos uno es la dimensión del n-simple.

3. Propiedades Topológicas de los Conjuntos Convexos.

Todos los resultados presentados y las definiciones dadas en las secciones anteriores están basadas en las propiedades lineales de un espacio vectorial. Ahora, consideraremos algunas propiedades de los conjuntos convexos en un espacio vectorial normado L.

3.1 Proposición.

Si $U \subseteq L$ es un conjunto convexo; entonces, su interior $\overset{\circ}{U}$ y su cierre $\bar{U}$, son convexos.

Demostración.

Sea $x, y \in \overset{\circ}{U}$, $z = \alpha x + (1-\alpha)y$, con $\alpha \in (0,1)$. $z \in U$ porque $\overset{\circ}{U} \subseteq U$ y U es convexo. Como $x \in \overset{\circ}{U}$, existe una -

bola abierta $B(x,r) \subseteq U$; luego, si u es un vector tal que $\|u/\alpha\| \leq r$, $(x + u/\alpha) \in U$ y la combinación de vectores $z+u = \alpha x + (1-\alpha)y+u = [\alpha(x+u/\alpha) + (1-\alpha)y]$ está en U. Es decir, existe una bola con centro en z completamente contenida en U; por lo tanto, $z \in \overset{\circ}{U}$ y $\overset{\circ}{U}$ es convexo.

Sea $x, y \in \bar{U}$. Como $\bar{U}$ es cerrado existen sucesiones $\{x_j\}$ y $\{y_j\}$ de puntos de U que convergen a x e y, respectivamente. Si $\alpha \in (0,1)$, $z_j = \alpha x_j + (1-\alpha)y_j$ está en U y $z = \lim_{j\to\infty} z_j = \alpha \lim_{j\to\infty} x_j + (1-\alpha) \lim_{j\to\infty} y_j$ está en U. Es decir, $z = \alpha x + (1-\alpha)y$ es una combinación convexa de $x, y \in U$ que está en $\bar{U}$; por lo tanto, $\bar{U}$ es convexo.

3.2 Proposición.

Si U es un conjunto abierto en el espacio vectorial L, su envolvente convexa $H(U)$ es un conjunto abierto. Si U es compacto y L, de dimensión finita; entonces, $H(U)$ es compacto.

Demostración.

Sea U un conjunto abierto y $x = \sum_{1}^{n} \alpha_i x_i$, $x_i \in U$, un elemento cualquiera de $H(U)$. Entonces, $x + u = \sum_{1}^{n} \alpha_i(x_i+u)$ y como U es abierto, $(x_i + u) \in U$ cuando $\|u\|$ es suficientemente pequeña. Luego, $x + u$ está en $H(U)$ como combinación convexa de elementos de U y por eso, existe una bola abierta de centro x completamente contenida en $H(U)$. Así, x es un punto interior y $H(U)$, un conjunto abierto.

Ahora, supongamos que U es un conjunto compacto. Entonces, existe M tal que $\|x\| \leq M$, $\forall\, x \in U$. Puesto que L es de dimensión finita, digamos n, la dimensión de $H(U)$, llamémosla m, es menor o igual que n; luego, finita. Si $y \in H(U)$, por el teorema de Carathéodory sabemos que existen $m + 1$ puntos de U tales que

$y = \sum_{0}^{m} \alpha_i x_i$, con $m = \dim H(U)$. Entonces, $\|y\| \leq \sum_{0}^{m} \alpha_i \|x_i\| \leq \sum_{0}^{m} \alpha_i M = M$.

Así, $H(U)$ también está acotado por M. Ahora probaremos que toda sucesión de $H(U)$ converge a un punto de $H(U)$. Sea $\{x_j\}$ una sucesión de puntos de $H(U)$ que converge a x. Por el teorema de Carathéodory podemos escribir $x_j = \sum_{0}^{m} \alpha_{ij} x_{ij}$, con $x_{ij} \in U$. Para cada i, las sucesiones $\{\alpha_{ij}\}$ y $\{x_{ij}\}$ están acotadas por 1 y M, respectivamente; por lo que, contienen subsucesiones convergentes. Puesto que, la cantidad de subíndices i es finita podemos seleccionar una subsucesión $\{j_k\}$ de enteros positivos tales que $\{\alpha_{ij_k}\}$ y $\{x_{ij_k}\}$ converjan, para cada i, digamos a α_i y z_i, respectivamente. Así, tenemos que

$$x = \lim_{k\to\infty} x_{j_k} = \lim_{k\to\infty} \sum_{o}^{m} \alpha_{ij_k} x_{ij_k} = \sum_{o}^{m} \lim_{k\to\infty} (\alpha_{ij_k} x_{ij_k}) = \sum_{o}^{m} \alpha_i z_i .$$

Los números $\alpha_{i_{j_k}} \geq 0$ ya que, son elementos del conjunto de números $\alpha_{ij} \geq 0$. Además, cada x_{j_k} es un elemento del conjunto formado por los x_j por lo que, $\sum_{o}^{m}\alpha_{ij_k} = 1$.

Ahora, $\alpha_i = \lim_{k\to\infty} \alpha_{ij_k}$ y $\lim_{k\to\infty} \sum_{o}^{m} \alpha_{ij_k} = \sum_{o}^{m} \alpha_i = 1$. Como U es cerrado cada z_i está en U; así que, x es una combinación convexa de elementos de U y $x \in H(U)$.

4. <u>Hiperplanos.</u>

El concepto de hiperplano, en un espacio vectorial L, es una generalización de las nociones de recta en $\mathbb{R}^2$ y de plano en $\mathbb{R}^3$.

4.1 <u>Definición.</u>

Todo subconjunto afín maximal propio de un espacio vectorial L ó, equivalentemente, toda traslación de un subespacio propio maximal de L es llamado un hiperplano.

4.2 <u>Proposición.</u>

H es un hiperplano del espacio vectorial L si y sólo si $H = \{z \in L / f(z) = \alpha$, con f una funcional lineal no nula$\}$.

Demostración.

Sabemos que cada subespacio propio maximal de L es el espacio nulo N de una funcional lineal no nula, $f:L\to\mathbb{R}$. Como H es una traslación de N, cada $z \in H$ puede ser escrito en la forma $z = z_0 + x$, con $x \in N$ y, puesto que f es nula sobre N, $f(z) = f(z_0) + f(x) = \alpha$.

Luego, $H = \{z \in L/f(z) = \alpha$, con f una funcional no nula$\}$. (1)

Supongamos que H está definido en la forma (1) y seleccionemos cualquier $z_0 \in H$ para formar el conjunto $N = \{x \in L/x = z - z_0,\ z \in H\}$.

Como $f(x)=f(z-z_0)=f(z)-f(z_0) = 0$, N es un subespacio propio maximal de L y, puesto que H es el conjunto de todos los vectores $z = x + z_0$, con $x \in N$, H es una traslación de N y, por consiguiente, un hiperplano.

4.3 Definición.

Sea L un espacio vectorial y $H \subseteq L$, un hiperplano. Entonces, los conjuntos $H^+ = \{z\in L/f(z)>\alpha\}$ y $H^- = \{z\in L/f(z)<\alpha\}$ son llamados los semi-espacios abiertos determinados por H.
Las cerraduras de estos conjuntos $\bar{H}^+$ y $\bar{H}^-$ son los semi-espacios cerrados determinados por H.

4.4 Definición.

Sean U y V dos subconjuntos del espacio vectorial L. El hiperplano H separa a U y V si U y V están contenidos, respectivamente, en cada semi-espacio cerrado determinado por H. H separa fuertemente a U y V si H está, "estrictamente, situado" entre dos traslaciones de H que separan a U y V. La separación fuerte requiere que los conjuntos - U y V sean disjuntos; mientras que la separación simple, no.

4.5 Hiperplanos del Espacio Euclídeo $\mathbb{R}^n$.

Ahora, estudiamos los hiperplanos del Espacio Euclídeo $\mathbb{R}^n$. Puesto que $\mathbb{R}^n$ es un espacio vectorial sobre $\mathbb{R}$ con producto escalar no degenerado, $<,>: \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, entonces $f: \mathbb{R}^n \to \mathbb{R}$ definido por $f(x) = <c,x>$, con $c \in \mathbb{R}^n$ fijo, es una funcional lineal y por lo tanto para $\alpha \in \mathbb{R}$

$$H = \{x \in \mathbb{R}^n / <c,x> = \alpha,\ c \neq \bar{0}\}$$

es un hiperplano.

Si $c = (c_1,\ldots,c_n)$ y $x = (x_1,\ldots,x_n)$ la ecuación desarrollada del hiperplano H es $<c,x> = c_1x_1+\ldots+c_n\ x_n=\alpha$. (2)

Cualquier $x \in \mathbb{R}^n$ cuyas coordenadas satisfagan (2) - pertenece al hiperplano H.

Puesto que $c \neq \bar{o}$, un hiperplano H pasa por el origen si y sólo si $<c,x> = 0$. En este caso, c es ortogonal con cada vector de H y decimos que c es normal al hiperplano. Cuando $\alpha = 0$ y $x, x_1 \in H$, $<c, x-x_1> = 0$ y c es ortogonal con $x - x_1$. Como este vector es "paralelo" con H, c es normal al hiperplano. En todo caso, c es normal con H.

Si multiplicamos $<c,x> = \alpha$ por el escalar $\lambda \neq 0$ tenemos $<\lambda c, x> = \lambda\alpha$; es decir, si un hiperplano está definido por c y α, también lo está por λc y $\lambda\alpha$, $\lambda \neq 0$. Como c es normal al hiperplano, λc también lo es. Ahora, si $\lambda = \|c\|^{-1}$; entonces, $<\|c\|^{-1} c, x> = \|c\|^{-1}\alpha = <n,x> = b$ representa un hiperplano H y n es llamado un vector unitario normal con H. También, $-n$ es un vector unitario normal con H.

4.6 <u>Proposición.</u>

Los hiperplanos de $\mathbb{R}^n$ son conjuntos convexos cerrados.

<u>Demostración.</u>

Sea H un hiperplano con ecuación $<c,x> = \alpha$. Entonces, $C_{\mathbb{R}^n} H = \{x \in \mathbb{R}^n / f(x) \neq \alpha\}$, con $f(x) = <c,x>$.

Como f es una función lineal definida de $\mathbb{R}^n$ en $\mathbb{R}$, f es continua. Además, $C_{\mathbb{R}^n} H = f^{-1}[(-\infty,\alpha) \cup (\alpha,\infty)]$.

Por otra parte, como $(-\infty,\alpha)$, (α,∞) son conjuntos abiertos y la unión de conjuntos abiertos es un abierto,

$f^{-1}\left[(-\infty,\alpha) \cup (\alpha,\infty)\right]$ es un abierto en $\mathbb{R}$, ya que f es continua. Luego, H es un conjunto cerrado.

Mostremos que H es un conjunto convexo. Sea $x_1,x_2 \in H$; entonces, $<c,x_1>=<c,x_2>=\alpha$ y $x=\lambda x_2+(1-\lambda)x_1, \lambda\in[0,1]$ está en H ya que, $<c,x>=<c,\lambda x_2+(1-\lambda)x_1>=\lambda<c,x_2>+(1-\lambda)<c,x_1>=\lambda\alpha+(1-\lambda)\alpha=\alpha$.

Los teoremas que siguen tienen gran aplicación en un amplio rango de problemas en programación lineal, en teoría de la decisión y en teoría de juegos.

5. Teoremas de Separación de Conjuntos Convexos.

Aunque la separación fuerte de dos conjuntos convexos U y V requiere que sean disjuntos, esa exigencia no la garantiza; para asegurarla, uno de los conjuntos debe ser compacto y el otro cerrado. Es lo que se muestra en

5.1 Proposición.

Sea U y V dos subconjuntos de $\mathbb{R}^n$ disjuntos, convexos, no vacíos con U compacto y V cerrado. Entonces, existe un hiperplano que separa fuertemente a U y V.

Demostración.

Para la demostración nos apoyamos en la figura 2.1 y en la proposición 10.6, página 55.

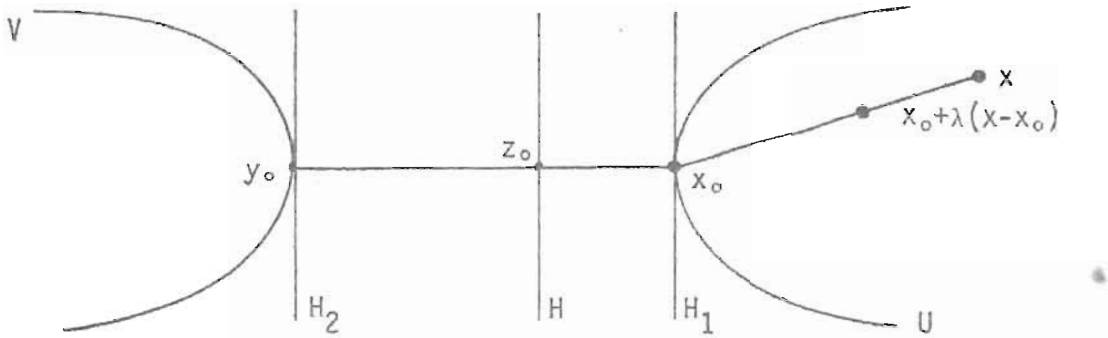


fig. 2.1

Por ser U compacto, V cerrado y $U \cap V = \emptyset$ existen $x_o \in U$, $y_o \in V$ tales que $d(U,V) = \inf\limits_{x \in U, y \in V} \|x-y\| = \|x_o - y_o\| > 0$. Mostraremos que un hiperplano H que pasa por el punto $z_o \in$ seg $[x_o, y_o]$ y que, además, es ortogonal con este segmento separa a U y V, fuertemente.

El hiperplano H_1 ortogonal con el segmento $[x_o, y_o]$ y que pasa por x_o tiene, por ecuación, $\langle y_o - x_o, z - z_o\rangle = 0$.

Con $x \in U$ definimos $\Psi : \mathbb{R} \to \mathbb{R}$, con

$\Psi(\lambda) = \| y_o - [x_o + \lambda(x - x_o)]\|^2$. Entonces,

$$\Psi(\lambda) = \langle (y_o - x_o) - \lambda(x-x_o), (y_o - x_o) - \lambda(x - x_o)\rangle$$

$$= \langle y_o - x_o, y_o - x_o\rangle - 2\lambda\langle y_o - x_o, x - x_o\rangle + \lambda^2 \langle x - x_o, x - x_o\rangle .$$

Como Ψ es diferenciable, por ser una función polino -

mial, tenemos que $\Psi'(\lambda) = -2<y_o-x_o, x-x_o>+2\lambda<x-x_o, x-x_o>$.
Siendo $x_o \in U$ el punto más cercano a y_o se tiene:

$0 < \| y_o-x_o \| \leq \| y_o-z \|$; $\forall z \in U$, con $z = x_o+\lambda(x-x_o)$.

Luego, $\|y_o-x_o\|^2 \leq \|y_o-z\|^2$ y $\Psi(0) \leq \Psi(\lambda)$, con $\lambda \in [0,1]$.

Así, $\Psi'_+(0) = \lim_{\lambda \to 0+} \frac{\Psi(\lambda)-\Psi(0)}{\lambda} \geq 0$; es decir, $\Psi'_+(0)=-2<y_o-x_o, x-x_o> \geq 0$.

Luego, $<y_o-x_o, x-x_o> \leq 0$, $\forall x \in U$.

Por medio de un argumento similar aplicado al hiperplano H_2 que pasa por y_o y que es ortogonal con el segmento $[x_o, y_o]$ encontramos que, $<x_o-y_o, y-y_o> \leq 0$, $\forall y \in V$. Así, si $y \in V$, $<y_o-x_o, y-y_o> \geq 0$. Como $<y_o-x_o, y_o-x_o> \geq 0$ tenemos que

$<y_o-x_o, y-y_o> + <y_o-x_o, y_o-x_o> = <y_o-x_o, y-x_o> \geq 0$. Ahora, combinando este resultado con el anterior $<y_o-x_o, x-x_o> \leq 0$, $\forall x \in U$, hemos probado que H_1 separa a U y V. Un resultado similar se obtiene para H_2. Ahora, como H "está entre H_1 y H_2", éstos son traslaciones de aquel y H separa fuertemente a U y V.

5.2 Proposición.

Sea U y V dos conjuntos convexos de $\mathbb{R}^n$ con $U^\circ \neq \emptyset$ y $U^\circ \cap V = \emptyset$. Entonces, existe un hiperplano que separa a U y V.

Demostración.

La verdad del teorema para $\bar{U}$ y $\bar{V}$ implica la verdad del mismo para U y V de modo que, no hay pérdida de generalidad si consideramos a U y V, cerrados.

Para $x_o \in U^o$ definimos

$B_n = \{x \in \mathbb{R}^n / \|x-x_o\| \leq n\}$ y $D_n = \{x_o + (1-\frac{1}{n})(x-x_o) / x \in U\}$. Ahora, formemos el conjunto $U_n = B_n \cap D_n$. [fig. 2.2].

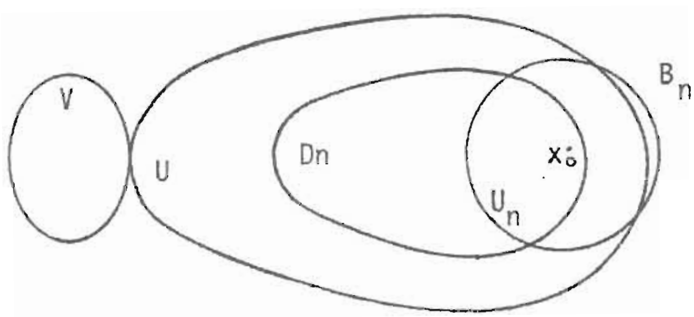


fig. 2.2

D_n es convexo.

Sea $y_1, y_2 \in D_n$ y $\alpha \in [0,1]$. Entonces,

$z = \alpha y_1 + (1-\alpha) y_2 \; \alpha[x_o + (1-\frac{1}{n})(x_1 - x_o)] + (1-\alpha)[x_o + (1-\frac{1}{n})(x_2 - x_1)]$; con $x_1, x_2 \in U$.

$= x_o + (1-\frac{1}{n})[\alpha(x_1-x_o) + (1-\alpha)(x_2 - x_o)]$

$= x_o + (1-\frac{1}{n})[\alpha x_1 + (1-\alpha)x_2 - x_o]$

$= x_o + (1-\frac{1}{n})(x-x_o)$ ya que, $x = \alpha x_1 + (1-\alpha) x_2$ está en U por ser éste un conjunto convexo.

D_n es cerrado.

Sea $\{Y_m\}$ una sucesión de puntos de D_n. Entonces,

$$\lim_{m\to\infty} Y_m = \lim_{m\to\infty} [x_\circ + (1-\tfrac{1}{n})(x_m - x_\circ)] \ , \ x_m \in U.$$

$$= x_\circ + (1-\tfrac{1}{n})(x - x_\circ) \text{ con } \lim_{m\to\infty} x_m = x \text{ en } U, \text{ puesto}$$

que U es cerrado. Así, hemos mostrado que D_n es convexo y cerrado.

Como toda bola cerrada en $\mathbb{R}^n$ es convexa, B_n es un conjunto convexo cerrado y; por consiguiente, también lo es U_n, como intersección de dos conjuntos convexos cerrados. Además, U_n es acotado puesto que B_n es acotada y $U_n \subseteq B_n$. En conclusión, U_n es convexo y compacto. Probemos que $U_n \cap V = \emptyset$. Como $U^\circ \cap V = \emptyset$, por hipótesis, y $U_n \subseteq D_n$ bastará mostrar que $D_n \subseteq U^\circ$. Sea $y \in D_n$; entonces, $y = \frac{1}{n} x_\circ + (1- \frac{1}{n})x$, con $x_\circ \in U^\circ$ y $x \in U$. Si $x \in U^\circ$, $y \in U^\circ$ porque U° es convexo. Luego, $U_n \subseteq U^\circ$ y $U_n \cap V = \emptyset$. Si $x \in F_r(U)$, $y \in \text{seg}\ [x_\circ, x[\ \subseteq U^\circ$ ya que y no puede ser igual a x; pues, si lo fuera tendríamos que

$x = \frac{1}{n} x_\circ + x - \frac{1}{n} x$, $\bar{0} = x_\circ - x$, $x_\circ = x$ y x sería un punto de U°, lo cual contradice nuestra hipótesis. Así, tam bién en este caso $U_n \cap V = \emptyset$.

Ahora, apoyándonos en la prop. 5.1,pág. 117 podemos afirmar que existe un hiperplano H_n, con ecuación $<u_n,x> = \alpha_n$ que separa fuertemente a U_n y V. Así, para $x \in U_n, <u_n,x> \leq \alpha_n$; mientras que, para $y \in V$, $<u_n,y> \geq \alpha_n$. Podemos asumir, al escribir estas relaciones, que los vectores u_n han sido normalizados de modo que, $\|u_n\| = 1$. Entonces, las sucesiones $\{u_n\}$ y $\{\alpha_n\}$ son acotadas, la última porque $<u_n,x_o> \leq \alpha_n \leq <u_n, y_o>$, con x_o, y_o dos puntos fijos de U y V, respectivamente. Así, podemos seleccionar subsucesiones $\{U_{n_k}\}$ y $\{\alpha_{n_k}\}$, las cuales son convergentes digamos, a u y a α. Para cualquier $x \in U^\circ$, x estará en U_{n_k}, para k suficientemente grande.

Para este k, $<u_{n_k},x> \leq \alpha_{n_k}$ y, consecuentemente, $<u,x> \leq \alpha$, $\forall\, x \in U^\circ$ ya que,

$$\lim_{k\to\infty} < u_{n_k},x> = \left\langle \lim_{k\to\infty} u_{n_k}, x\right\rangle \leq \lim_{k\to\infty} \alpha_{n_k} = \alpha \text{ ; y, } <u,x> \leq \alpha \text{ . } (3)$$

Sea $x \in F_r(U)$. Por ser U cerrado existe una sucesión de puntos de U, $\{x_n\}$ que convergen a x y que satisfacen la desigualdad (3).

Luego, $\lim_{n\to\infty} <u,x_n> = \left\langle u, \lim_{n\to\infty} x_n\right\rangle = <u,x> \leq \alpha$. Esta última -

desigualdad es cierta para todo $x \in U$.

Similarmente, se deduce que $\alpha \leq \langle u,y\rangle$, $\forall y \in V$. El hiperplano que separa a U y V es $H = \{z \in L / \langle u,z\rangle = \alpha\}$.

5.3 Definición.

Sea U un conjunto convexo y $x_0 \in F_r(U)$. Decimos que el hiperplano H, con ecuación $\langle u,z\rangle = \alpha$, es un soporte para U en x_0 si $\langle u,x_0\rangle = \alpha$ y si U es un subconjunto de uno de los semi-espacios cerrados determinados por H.

5.4 Proposición.

Si $x_0 \in U$ es un punto frontera del conjunto convexo $U \subseteq \mathbb{R}^n$, con $U^\circ \neq \emptyset$; entonces, existe un hiperplano soporte para U en x_0.

Demostración.

Hagamos $V = \{x_0\}$. Entonces, V es un conjunto convexo. Además, U, por hipótesis, es convexo con $U^\circ \neq \emptyset$. Siendo $x_0 \in F_r(U)$ tenemos que $V \cap U^\circ = \emptyset$. Luego, utilizando la prop. 5.2, pág. 119, podemos afirmar que existe un hiperplano H que separa a U y V pasando por x_0. Ahora, apoyándonos en la def. 5.3, pág. 123, afirmamos que H es un hiperplano soporte para U en x_0.

5.5 Definición.

Un punto x_0 de un conjunto convexo U es llamado un -

punto extremo si no es punto interior de algún segmento de recta contenido en U; es decir, si no existen puntos $x_1, x_2 \in U$ y $\alpha \in (0,1)$ de modo que $x_\circ = \alpha x_1 + (1-\alpha)\, x_2$.

Los puntos extremos de una bola cerrada son sus puntos - frontera; los de un cubo, en $\mathbb{R}^3$ sus ocho vértices; los de los polígonos convexos en $\mathbb{R}^2$, sus vértices, etc. Un semi-espacio, aunque sea cerrado, no tiene puntos extremos.

5.6 Proposición.

Sea U un subconjunto convexo cerrado de $\mathbb{R}^n$ y H, un hiperplano soporte para U en $x_\circ$. Si x es un punto extremo - de $H \cap U$; entonces, x es un punto extremo de U.

Demostración.

Supongamos que $x \in H \cap U$ no es un punto extremo de U. Entonces, existen $x_1, x_2 \in U$ tales que $x = \lambda x_1 + (1-\lambda)x_2$, con $\lambda \in (0,1)$. Mostraremos que, en estas condiciones, x_1 y x_2 deben ser elementos de $H \cap U$. Como $x_1, x_2 \in U$ sólo falta mostrar que están en H. Si $x_1, x_2 \notin H$, están en el semi-espacio abierto E determinado por H y que contiene a U. - Luego, el segmento $[x_1, x_2] \subseteq E$ ya que, E es un conjunto - convexo, y $x \notin H \cap U$. Así, tenemos que existen $x_1, x_2 \in H \cap U$ tales que $x = \lambda x_1 + (1-\lambda)x_2$, con $\lambda \in (0,1)$;

es decir, x no es un punto extremo de $H \cap U$; esta conclusión es una contradicción y x debe ser un punto extremo de U.

5.7 Proposición.

Sea $U \subseteq \mathbb{R}^n$ un conjunto convexo, compacto no vacío. Entonces, U es la envolvente convexa de sus puntos extremos.

Demostración.

En la demostración usaremos inducción sobre la dimensión m del conjunto U.

Si $m = 0$, $\dim A(U) = \dim W = 0$, con W un subespacio del cual $A(U)$ es una traslación. Entonces, $W = \{\bar{o}\}$ y

$A(U) = \{x_o + \bar{o}/x_o \in \mathbb{R}^n\} = \{x_o\}$. Como $U \neq \emptyset$ y $U \subseteq \{x_o\}$, debe ser $U = \{x_o\}$. Por otra parte,

$H(U) = \{x \in \mathbb{R}^n/x = \alpha x_o + (1-\alpha)x_o,\ x_o \in U,\ \alpha \in [0,1]\} = \{x_o\}$. Así, $U=H(U)$.

Si $m = 1$, $\dim A(U) = \dim W = 1$, con $A(U)$ una traslación del subespacio W. Entonces, $W=\{\lambda x/x \in \mathbb{R}^n, \lambda \in \mathbb{R}\}$, $A(U)=\{x_o+\lambda x/\lambda x \in W\}$ es una recta y $U = [a,b]$. Por otra parte,

$H(U) = \{x \in \mathbb{R}^n/x = \alpha a+(1-\alpha)b,\ \alpha \in [0,1]\}= [a,b]$. Así, también en este caso, es $U = H(U)$.

Ahora, supongamos cierto el resultado para cualquier conjunto compacto de dimensión, a lo sumo, m, con $m \leq n - 1$.

Sea $m + 1$ la dimensión de U y consideremos a U incluido en $\mathbb{R}^{m+1}$. Si $x_\circ$ es un punto frontera de U, por la prop. 5.4, pág. 123, sabemos que hay un hiperplano soporte H (un conjunto afín m - dimensional) para U que pasa por $x_\circ$. El conjunto $H \cap U$ es convexo y cerrado, por ser la intersección de dos conjuntos convexos cerrados. Además, es acotado ya que, - $U \cap H \subseteq U$ y U es acotado. Así que, $H \cap U$ es convexo, compacto y su dimensión no excede a m. Entonces, de acuerdo con nuestra hipótesis inductiva $x_\circ$ es una combinación convexa de los puntos extremos de $H \cap U$ y, por consiguiente, una - combinación convexa de los puntos extremos de U (prop. 5.6, pág. 124). Consideremos, por último, el caso en que $x_\circ$ es un punto interior de U. En este caso, cualquier recta que pasa por $x_\circ$ tiene como intersección con U un segmento con puntos extremos x_1, x_2 los cuales, son puntos frontera de U. Como x_1 y x_2 son combinaciones convexas de puntos extremos y $x_\circ$ lo es de x_1 y x_2, $x_\circ$ es una combinación convexa de puntos extremos.

Los teoremas relacionados con la separación, soporte - y puntos extremos han sido probados en $\mathbb{R}^n$; luego, por el isomorfismo vectorial y topológico que existe entre este espacio y los de dimensión finita, dichos teoremas están probados para cualquier espacio vectorial normado de dimensión finita.

6. Conos Convexos en $\mathbb{R}^n$.

La última clase de conjuntos convexos que estudiamos son los conos convexos. Estos conjuntos tienen un papel importante en el análisis económico, especialmente, en la teoría de la producción.

6.1 Definición.

Un cono $C \subseteq \mathbb{R}^n$ es un conjunto de puntos que presenta la siguiente propiedad: si $x \in C$; entonces, $\lambda x \in C$, $\forall\, \lambda > 0$.

El cono generado por un conjunto $U \subset \mathbb{R}^n$ es el conjunto $C = \{y \in \mathbb{R}^n / y = \lambda x,\ \forall\ x \in U,\ \forall\ \lambda \geq 0\}$.

El punto $\bar{0}$ es un elemento de cualquier cono y es llamado el vértice del cono.

El negativo de un cono C es el conjunto $C^- = \{-y / y \in C\}$. Es claro que, C^- es un cono si y sólo si C lo es.

6.2 Definición.

La suma de dos conos $C_1 \subseteq \mathbb{R}^n$, $C_2 \subseteq \mathbb{R}^n$ denotado con $C_1 + C_2$ se define como $\{u+v \in \mathbb{R}^n / u \in C_1,\ v \in C_2\}$.

6. 3 Definición.

Un cono $C \subseteq \mathbb{R}^n$ es un cono convexo si C es un conjunto convexo.

6.4 Proposición.

Un conjunto $C \subseteq \mathbb{R}^n$ es un cono convexo si y sólo si

(1) $x \in C$; entonces, $\lambda x \in C$, $\forall \lambda \geq 0$.

(2) $x_1, x_2 \in C$; entonces, $(x_1 + x_2) \in C$.

Demostración.

Sea C un cono convexo. Entonces, si $x \in C$, $\lambda x \in C$, $\forall \lambda \geq 0$, por definición de cono. Ahora, veamos que se cumple (2): Dado que C es un cono podemos escribir

$x_1 = \lambda v_1$, $0 \leq \lambda \leq 1$, $v_1 \in C$; $x_2 = (1-\lambda)\, v_2$, $v_2 \in C$

y afirmar que $x_1, x_2 \in C$. Luego, como C es convexo, $\lambda v_1 + (1-\lambda)\, v_2$ está en C; es decir, $x_1 + x_2$ está en C.

Supongamos que se cumplen las condiciones (1) y (2) y mostremos que C es un cono convexo. La primera condición indica, por definición, que C es un cono. El cono C será convexo si $\lambda v_1 + (1-\lambda)v_2, 0 \leq \lambda \geq 1$, está en C cuando v_1 y v_2 lo están. Si $v_1, v_2 \in C$, por la primera condición, podemos afirmar que $\lambda v_1, (1-\lambda)v_2 \in C$ y, por la segunda condición, que $\lambda v_1 + (1-\lambda)v_2$ está en C. Luego, C es convexo.

6.5 Ejemplos de Conos Convexos.

- La semi-recta $C = \{y \in \mathbb{R}^n / y = \lambda x, \lambda \geq 0\}$, con x un vector cualquiera de $\mathbb{R}^n$ diferente del vector nulo.

- Los semi-espacios, $\bar{H}^- = \{x \in \mathbb{R}^n / \langle u,x\rangle \leq 0\}$ y $\bar{H}^+ = \{x \in \mathbb{R}^n / \langle u,x\rangle \geq 0\}$.

- El ortante no negativo de $\mathbb{R}^n$, $\mathbb{R}^n_+ = \{x \in \mathbb{R}^n / x \geq \bar{o}\}$, con $x = (x_1,\ldots,x_n)$, $x_i \geq 0$, $i = 1,\ldots,n$.

- Cualquier subespacio vectorial de $\mathbb{R}^n$.

6.6 Proposición.

La suma de dos conos convexos C_1 y C_2 de $\mathbb{R}^n$ es un cono convexo.

Demostración.

Puesto que, $\lambda(u + v) = \lambda u + \lambda v$; $\lambda u \in C_1$, si $u \in C_1$ y $\lambda v \in C_2$, si $v \in C_2$, $\forall \lambda \geq 0$; $\lambda(u + v) \in C_1 + C_2$, $\forall \lambda \geq 0$ y $C_1 + C_2$ es un cono.

Sea x_1 y x_2 dos puntos cualesquiera de la suma C_1+C_2. Por la definición de suma es

$x_1 = u_1 + v_1$, $x_2 = u_2 + v_2$, con $u_1, u_2 \in C_1$ y $v_1, v_2 \in C_2$.

Luego, para cualquier punto x que sea una combinación convexa de x_1 y x_2 tenemos que

$x = \lambda x_1 + (1-\lambda)x_2 = \lambda u_1 + (1-\lambda)u_2 + \lambda v_1 + (1-\lambda)v_2$.

Como $\lambda u_1 + (1-\lambda)u_2$ está en C_1 y $\lambda v_1 + (1-\lambda)v_2$ está en C_2, $x \in C_1 + C_2$ y $C_1 + C_2$ es un cono convexo.

6.7 Proposición.

Sea $\{C_i\}$, $i = 1,\ldots,m$, una familia finita de conos de $\mathbb{R}^n$. Entonces, $C = \sum_1^m C_i$ es un cono. Además, si cada C_i, $i = 1,\ldots,m$, es un cono convexo, C es convexo.

Demostración.

La demostración es inmediata si aplicamos inducción sobre m, el número de conos, y los resultados de la proposición 6.6 , página 129.

6.8 Proposición.

Si $\{C_i\}$, $i = 1,\ldots,m$, es una familia de conos convexos de $\mathbb{R}^n$; entonces, la intersección $\bigcap_1^m C_i$ es un cono convexo.

Demostración.

Que $\bigcap_1^m C_i$ es un conjunto convexo se sigue de la proposición 2.1 , página 101.

Sea $C = \bigcap_1^m C_i$. Si $x \in C$, $x \in C_i$, $\forall_i$, $i = 1,\ldots,m$. Como C_i es un cono; $\forall_i$, $i = 1,\ldots,m$; $\lambda x \in C_i$, $\forall \lambda \geq 0$, $\forall_i$, $i=1,\ldots,m$; luego, $\lambda x \in C$ y C es un cono.

6.9 Proposición.

El cono C generado por un conjunto convexo es un cono convexo.

Demostración.

Dado el conjunto convexo U deseamos probar que

$C = \{y \in \mathbb{R}^n / y = \lambda x, \forall \lambda \geq 0, \forall x \in U\}$ es un cono convexo.

Es claro, por la definición, que C es un cono.

Sea $y_1, y_2 \in C$. Entonces, $y_1 = \lambda_1 x_1$, $y_2 = \lambda_2 x_2$, con $x_1, x_2 \in U$, $\lambda_1 \geq 0$ y $\lambda_2 \geq 0$. Vamos a mostrar que $y = \alpha y_1 + (1-\alpha) y_2$, con $0 \leq \alpha \leq 1$, está en C. Ahora bien,

$$y = \alpha\lambda_1 x_1 + (1-\alpha)\lambda_2 x_2 = \left[\alpha\lambda_1 + (1-\alpha)\lambda_2\right]\left[\frac{\alpha\lambda_1}{\alpha\lambda_1+(1-\alpha)\lambda_2}\, x_1 + \frac{(1-\alpha)\,\lambda_2}{\alpha\lambda_1+(1-\alpha)\lambda_2}\, x_2\right]$$

$$= \beta\left[\gamma\, x_1 + \frac{\alpha\lambda_1 + (1-\alpha)\lambda_2 - \alpha\lambda_1}{\beta}\, x_2\right] = \beta\left[\gamma x_1 + (1-\gamma) x_2\right],$$

con $\beta = \alpha\lambda_1 + (1-\alpha)\lambda_2 \neq 0$, $\gamma = \dfrac{\alpha\lambda_1}{\beta}$ y $0 \leq \gamma \leq 1$.

Como U es convexo $x = \gamma x_1 + (1-\gamma) x_2$ está en U e $y = \beta x$, en C.

Si $\lambda = 0$, $y = \bar{0}$, y $\bar{0} \in C$. El teorema está probado.

6.10 Definición.

Sea C un cono convexo en $\mathbb{R}^n$. Entonces, el dual o -

polar de C, denotado con C^* es definido como $\{y \in \mathbb{R}^n / \langle x,y\rangle \leq 0,\ \forall\ x \in C\}$.

6.11 Proposición.

Si C es un cono convexo en $\mathbb{R}^n$; entonces, su dual C^* es un cono convexo.

Demostración.

Puesto que C^* es el dual de C, $\langle x,y\rangle \leq 0,\ \forall x \in C$. - Luego, $\langle x,\lambda y\rangle \leq 0,\ \forall\ \lambda \geq 0,\ \forall\ x \in C$, $\lambda y \in C^*$ y C^* es un cono.

Sea $\lambda \in [0,1]$ y $u,v \in C^*$; entonces, $\langle x,u\rangle \leq 0$, $\langle x,v\rangle \leq 0$, $\langle x,\lambda u\rangle \leq 0$, $\langle x,(1-\lambda)v\rangle \leq 0$, $\langle x,\lambda u\rangle + \langle y,(1-\lambda)v\rangle = \langle x+y,\ \lambda u+(1-\lambda)v\rangle \leq 0$ y $\lambda u+(1-\lambda)v$ está en C^*.

Por lo tanto, C^* es un conjunto convexo. El teo - rema está probado.

CAPITULO III

FUNCIONES CONVEXAS

1. Funciones Convexas en $\mathbb{R}$.

1. Definición.

Sea $I \subseteq \mathbb{R}$ un intervalo. La función $f: I \to \mathbb{R}$ es convexa si $f[\alpha x + (1-\alpha)y] \leq \alpha f(x) + (1-\alpha)f(y)$; $\forall x,y \in I$; $\alpha \in (0,1)$. (1)

Alternativamente α podría tomarse del intervalo $[0,1]$.

f es cóncava si $f[\alpha x + (1-\alpha)y] \geq \alpha f(x)+(1-\alpha)f(y)$; $\forall x,y \in I$; $\alpha \in (0,1)$. (2)

f es estrictamente convexa o estrictamente cóncava cuando las desigualdades (1) y (2) son estrictas para $x \neq y$.

La concavidad y la concavidad estricta corresponden, respectivamente, a la convexidad y a la convexidad estricta; por consiguiente, el estudio sobre las funciones convexas puede adaptarse a las funciones cóncavas.

f es cóncava si y sólo si $-f$ es convexa ya que
$f[\alpha x + (1-\alpha)y] \geq \alpha f(x) + (1-\alpha)f(y)$ si y sólo si
$-f[\alpha x + (1-\alpha)y] \leq \alpha[(-f)(x)] + (1-\alpha)[(-f)(y)]$.

2. Proposición.

$f: I \to \mathbb{R}$ es una función convexa si y sólo si, para $x,z,y \in I$, con $x<z<y$, $\dfrac{f(z) - f(x)}{z-x} \leq \dfrac{f(y)-f(x)}{y-x} \leq \dfrac{f(y)-f(z)}{y-z}$ (3)

Demostración.

$x < z < y$ implica $z = \alpha x + (1-\alpha)y$, $\alpha\in(0,1)$, de lo cual se deduce que $\alpha = \frac{z-y}{x-y} = \frac{y-z}{y-x}$; $1-\alpha = \frac{x-z}{x-y} = \frac{z-x}{y-x}$.

Supongamos f convexa. Entonces,

$f[\alpha x+(1-\alpha)y] \leq \alpha f(x) + (1-\alpha) f(y)$; $\forall x, y\in I$; $\alpha\in(0,1)$ y,

$f(z) \leq \alpha [f(x) - f(y)] + f(y) = \frac{y-z}{y-x} [f(x) - f(y)] + f(y)$.

Luego,

$$\frac{f(z) - f(y)}{y - z} \leq \frac{f(x) - f(y)}{y - x} \; ; \; \frac{f(y) - f(x)}{y - x} \leq \frac{f(y)-f(z)}{y - z} \quad (4)$$

También, $f(z) \leq \alpha f(x)+f(x)-f(x)+(1-\alpha)f(y) = f(x)+(1-\alpha)[f(y)-f(x)]$

Así que, $f(z)-f(x) \leq \frac{z-x}{y-x} [f(y)-f(x)]$;

$$\frac{f(z) - f(x)}{z - x} \leq \frac{f(y) - f(x)}{y - x} . \quad (5)$$

Por transitividad, aplicada a (4) y (5), obtenemos las desigualdades del teorema.

Ahora, supongamos que las desigualdades en (3) son ciertas. Entonces,

$f(z) \leq f(x) + [f(y) - f(x)] \frac{z - x}{y - x} =$

$f(x) + (1-\alpha) [f(y) - f(x)] = \alpha f(x) + (1-\alpha) f(y)$,

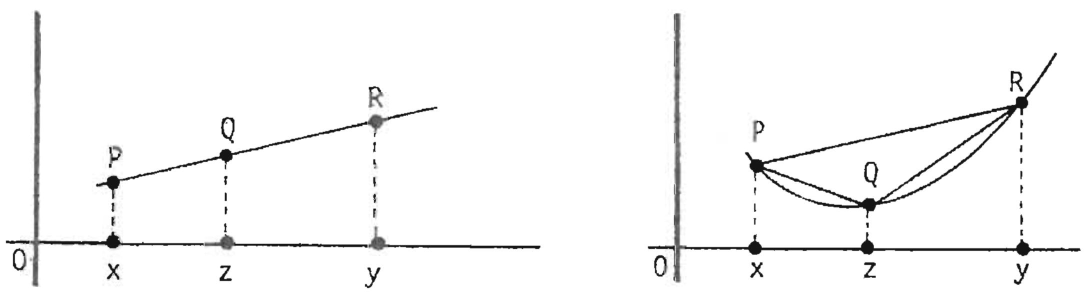
$-f(z) \geq [f(y) - f(x)] \frac{y-z}{y-x} - f(y) =$

$[f(y) - f(x)]\alpha - f(y) = -\alpha f(x) + (\alpha - 1) f(y)$ y

$f(z) \leq \alpha f(x) + (1 - \alpha) f(y)$.

En todo caso resulta que f es convexa.

Como x, z, y $\in$ I, los puntos P = (x,f(x)) , Q = (z,f(z)) y R = (y,f(y)) están en el gráfico de f. Las expresiones en las desigualdades de (3) representan respectivamente, de izquierda a derecha, las pendientes de los segmentos PQ, PR y QR. Así que, geométricamente, con la proposición 2 se afirma que si f es convexa el punto Q está en la cuerda PR o bajo ella, lo cual significa que el gráfico de f coincide con dicha cuerda o está bajo ella en el intervalo $[x,y]$ para cualesquiera x,y$\in$I. (fig. 3.1).



$z = \alpha x + (1-\alpha)y;\ \alpha \in (0,1).$

fig. 3.1

3. <u>Proposición.</u>

Sea f: I $\to \mathbb{R}$ una función y $S = \{(x,y) \in \mathbb{R}^2 / y \geq f(x), \forall x \in I\}$. Entonces, f es convexa si y sólo si S es un conjunto convexo.

<u>Demostración.</u>

Si f es convexa, si $(x_1, y_1),(x_2, y_2)$ son dos puntos

cualesquiera de S y si $\alpha \in [0,1]$; entonces, $f(z) =$
$f[\alpha x_1 + (1-\alpha)\, x_2] \leq \alpha\, f(x_1) + (1-\alpha)\, f(x_2) \leq \alpha y_1 + (1-\alpha)\, y_2$ y
$P = [\alpha x_1 + (1-\alpha)x_2,\ \alpha y_1 + (1-\alpha)y_2]$ está en S ya que, $z \in I$.
Además,
$P = \alpha(x_1,\ y_1) + (1-\alpha)\ (x_2, y_2)$; luego, S es convexo.

Si S es convexo, $(x_1,\ y_1)$, $(x_2,\ y_2)$ son dos puntos cualesquiera de S y f, con $y_1 = f(x_1)$, $y_2 = f(x_2)$ y si $\alpha \in [0,1]$; entonces,
$\alpha(x_1,y_1) + (1-\alpha)\ (x_2,\ y_2) = [\alpha x_1+(1-\alpha)x_2, \alpha y_1+(1-\alpha)y_2]$ está en S; luego, $f[\alpha x_1 + (1-\alpha)x_2] \leq \alpha\, f(x_1) + (1-\alpha)\ f(x_2)$ y
f es convexa.

Con un procedimiento similar se prueba que:

4. <u>Proposición.</u>

Si $f: I \to \mathbb{R}$ es una función y $S = \{(x,y) \in \mathbb{R}^2 / y \leq f(x),\ \forall\, x \in I\}$; entonces, f es cóncava si y sólo si S es convexa.

5. <u>Ejemplos de Funciones Convexas.</u>

- Toda función afín $f: I \subseteq \mathbb{R} \to \mathbb{R}$, con $f(x) = mx+b$.
- $g: \mathbb{R} \to \mathbb{R}$, con $g(x) = x^2$.
- $h: \mathbb{R} \to \mathbb{R}$, con $h(x) = e^x$.
- $i: \mathbb{R}^+ \to \mathbb{R}$, con $i(x) = x^p$, $p > 1$.
- $j: \mathbb{R}^+ \to \mathbb{R}$, con $j(x) = x \log x$.

1.1 Continuidad.

Si una función convexa no es contínua en todo su dominio, las discontinuidades sólo pueden ocurrir en los puntos frontera de su dominio, es lo que se deduce de la proposición 2 que demostraremos a continuación.

1. Proposición.

Sea $f: [a,b] \to \mathbb{R}$ una función convexa acotada. Entonces, f está acotada superiormente por $M = \text{máx}\{f(a), f(b)\}$ e, inferiormente, por $m = 2f\left(\frac{a+b}{2}\right) - M$.

Demostración.

Si $z = \alpha a + (1-\alpha)b$, con $\alpha \in [0,1]$, es un elemento cualquiera de $[a,b]$; entonces, $f[\alpha a+(1-\alpha)b] \leq \alpha f(a)+(1-\alpha)f(b) \leq \alpha M + (1-\alpha) M = M$.

Supongamos que $z = \frac{a+b}{2} + t$ es un elemento cualquiera de $[a,b]$.

Entonces,

$$f\left(\frac{a+b}{2}\right) = f\left[\frac{1}{2}\left(\frac{a+b}{2} + t\right) + \frac{1}{2}\left(\frac{a+b}{2} - t\right)\right] \leq$$

$$\frac{1}{2} f\left(\frac{a+b}{2} + t\right) + \frac{1}{2} f\left(\frac{a+b}{2} - t\right) \text{ y,}$$

$$f\left(\frac{a+b}{2} + t\right) \geq 2 f\left(\frac{a+b}{2}\right) - f\left(\frac{a+b}{2} - t\right) \geq$$

$$2 f\left(\frac{a+b}{2}\right) - M = m \text{ ya que, } -f\left(\frac{a+b}{2} - t\right) \geq -M.$$

Así, para cualesquiera $x,y \in [a,b]$, $f(x) - f(y) \leq M-m$.

2. Proposición.

Si $f: I \to \mathbb{R}$ es una función convexa, f satisface una condición de Lipschitz sobre cualquier intervalo cerrado $[a,b] \subseteq I°$, con $I°$ el interior de I. En consecuencia, f es absolutamente contínua en $[a,b]$ y contínua en $I°$.

Demostración.

Seleccionemos $\varepsilon > 0$ de modo que $a - \varepsilon$ y $b + \varepsilon$ pertenezcan a I. Sea M y m, respectivamente, el supremo y el ínfimo de f en $[a - \varepsilon, b + \varepsilon]$.

Con x e y dos puntos cualesquiera de $[a,b]$ (fig. 3.2) hagamos

$z = y + \frac{\varepsilon}{|y-x|}(y-x)$ y $\alpha = \frac{|y-x|}{\varepsilon+|y-x|}$. Entonces,

$z \in [a-\varepsilon, b+\varepsilon]$, $1-\alpha = \frac{\varepsilon}{\varepsilon+|y-x|}$ e $y = \frac{|y-x|}{\varepsilon+|y-x|} z + \frac{\varepsilon}{\varepsilon+|y-x|} x = \alpha z + (1-\alpha) x$.

Luego, tenemos que,

$f(y) \leq \alpha f(z) + (1-\alpha) f(x) = \alpha [f(z) - f(x)] + f(x)$ y

$f(y) - f(x) \leq \alpha(M-m) < \frac{|y-x|}{\varepsilon}(M-m) = K|y-x|$, con $K = \frac{M-m}{\varepsilon}$.

Como $f(y) - f(x) \leq K|y-x|$ es cierto para cualesquiera $x, y \in [a,b]$; es cierto, entonces que $f(x)-f(y) \leq K|y-x|$ y por ello, que $|f(y)-f(x)| \leq K|y-x|$.

Así, f es Lipschitz.

Sea $\{(a_i,b_i)\}$, $i=1,\dots,n$, una colección de subintervalos abiertos disjuntos de $[a,b]$. Por ser f Lipschitz, $|f(b_i) - f(a_i)| \leq K\,(b_i - a_i)$. Luego, $\sum_1^n |f(b_i)-f(a_i)| \leq K\sum_1^n (b_i-a_i)$. Si hacemos $K\sum_1^n(b_i-a_i) < \varepsilon$ tendremos que $\sum_1^n (b_i-a_i) < \frac{\varepsilon}{K}$ implica $\sum_1^n |f(b_i)-f(a_i)| < \varepsilon$ y f es absolutamente contínua en $[a,b]$; de aquí, uniformemente contínua en $[a,b]$, contínua en $[a,b]$ y en I°.

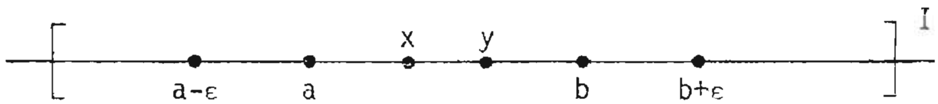


fig. 3.2

1.2 Diferenciabilidad.

Una función convexa no es necesariamente diferenciable, aún en puntos interiores de su dominio; pero, la diferenciabilidad lateral es una consecuencia inmediata de la convexidad.

1. Proposición.

Sea $f: I \to \mathbb{R}$ una función convexa. Entonces,

1) Si x es un punto interior de I, las derivadas laterales derecha $f'_+(x)$ e izquierda $f'_-(x)$ existen y $f'_-(x) \leq f'_+(x)$.
2) Si f es diferenciable en $y \in I$; entonces, para $x \in I$ se tiene $f'(y)\,(x-y) \leq f(x) - f(y)$.

3) Si x e y son dos puntos interiores de I con $x < y$; entonces, $f'_-(x) \leq f'_+(x) \leq f'_-(y) \leq f'_+(y)$.

Demostración.

Para (1). Sea $r,s,t, \in I$, con $r < s < t$. Por ser f una función convexa sabemos que

$$\frac{f(s) - f(r)}{s - r} \leq \frac{f(t) - f(r)}{t - r} \leq \frac{f(t) - f(s)}{t - s} . \qquad (6)$$

Como $x \in I°$, existe $\varepsilon > 0$ tal que $(x - \varepsilon, x + \varepsilon) \subseteq I$. Trabajaremos en este intervalo. Si aplicamos la desigualdad de la izquierda de (6) a $r = x$, $s = x + k_2$, $t = x + k_1$, con $0 < k_2 < k_1 < \varepsilon$, y a $r=x$, $s=x + h_1$, $t = x + h_2$, con $- \varepsilon < h_1 < h_2 < 0$ tendremos, respectivamente, que

$$\text{(a)} \quad \frac{f(x + k_2) - f(x)}{k_2} \leq \frac{f(x + k_1) - f(x)}{k_1}$$

$$\text{(b)} \quad \frac{f(x + h_1) - f(x)}{h_1} \leq \frac{f(x + h_2) - f(x)}{h_2}$$

De (6) se deduce que $\frac{f(r)-f(s)}{r - s} \leq \frac{f(t)-f(s)}{t - s}$. Si aplicamos esta desigualdad a $r = x + h_2$, $s = x$, $t = x + k_2$, con $- \epsilon < h_2 < 0 < k_2 < \epsilon$, tendremos

$$\text{(c)} \quad \frac{f(x + h_2) - f(x)}{h_2} \leq \frac{f(x + k_2) - f(x)}{k_2}$$

Combinando (a), (b) y (c) resulta

$$\frac{f(x+h_1)-f(x)}{h_1} \leq \frac{f(x+ h_2)-f(x)}{h_2} \leq \frac{f(x+k_2)-f(x)}{k_2} \leq \frac{f(x+k_1)-f(x)}{k_1} \qquad (7)$$

con $- \varepsilon < h_1 < h_2 < k_2 < k_1 < \varepsilon$.

De estas desigualdades se deduce, inmediatamente, que

i) $\dfrac{f(x+h)-f(x)}{h}$ no decrece y está acotada superiormente cuando $h<0$ tiende, monótonamente, a 0.

ii) $\dfrac{f(x+k)-f(x)}{k}$ no crece y está acotada inferiormente cuando $h>0$ tiende, monótonamente, a 0.

Luego,

$$f'_-(x)=\lim_{h\to o^-}\frac{f(x+h)-f(x)}{h}\ ;\ f'_+(x)=\lim_{k\to o^+}\frac{f(x+k)-f(x)}{k},\ \text{existen.}$$

Observando (7) se deduce, inmediatamente, que $f'_-(x)\leq f'_+(x)$.

Para (2).

(a) Sea $x>y$, $r=y$ e $t=x$ en (6). Entonces,

$$\frac{f(s)-f(y)}{s-y}\leq\frac{f(x)-f(y)}{x-y}\ \text{implica}\ \lim_{s\to y}\frac{f(s)-f(y)}{s-y}\leq\frac{f(x)-f(y)}{x-y};$$

es decir, $f'(y)\leq\dfrac{f(x)-f(y)}{x-y}$. Luego, $f'(y)(x-y)\leq f(x)-f(y)$.

(b) Sea $y>x$, $r=y$, $t=y$ en (6). Entonces,

$$\frac{f(y)-f(x)}{y-x}\leq\frac{f(y)-f(s)}{y-s}\ \text{implica}\ \frac{f(y)-f(x)}{y-x}\leq\lim_{s\to y}\frac{f(y)-f(s)}{y-s};$$

es decir,

$$\frac{f(y)-f(x)}{y-x}\leq f'(y),\ f'(y)(y-x)\geq f(y)-f(x)\ ;\ f'(y)(x-y)\leq f(x)-f(y).$$

Para (3).

De (6) deducimos que $\dfrac{f(s)-f(r)}{s-r}\leq\dfrac{f(s)-f(t)}{s-t}$. Si aplicamos esta desigualdad a $r=x$, $t=y$, s, con $x<s<y$, y si ha-

cemos $k = s - x > 0$, $h = s - y < 0$ tendremos que

$$\frac{f(x + k) - f(x)}{k} \leq \frac{f(y + h) - f(y)}{h}.$$

Luego, de acuerdo con la primera parte (a), tenemos que

$f'_+(x) \leq \frac{f(x+k)-f(x)}{k} \leq \frac{f(y+h)-f(y)}{h} \leq f'_-(y)$ y, además, que

$f'_-(x) \leq f'_+(x) \leq f'_-(y) \leq f'_+(y)$.

Esto significa que f'_- y f'_+ son crecientes en el interior de I.

Si f fuese estrictamente convexa, f'_- y f'_+ serían estrictamente crecientes en I°.

1.3 Caracterización.

Caracterizaremos las funciones convexas con propiedades de la primera y segunda derivadas y con la existencia de rectas soporte.

1. Proposición.

Sea $f:(a,b) \subseteq \mathbb{R} \to \mathbb{R}$ una función diferenciable en (a,b). Entonces, f es convexa [estrictamente convexa] en (a,b) si y sólo si f' es creciente [estrictamente creciente] en ese intervalo.

Demostración.

Supongamos f convexa [estrictamente convexa]. Enton ces, f'_+ y f'_- son crecientes [estrictamente crecientes] - en (a,b). Como f es diferenciable en (a,b), $f'=f'_-=f'_+$ y f' es creciente [estrictamente creciente] en dicho intervalo.

Ahora, supongamos que f' es creciente en (a,b). Para probar que f es convexa mostraremos que para cualesquiera $x,y\epsilon(a,b)$, $\frac{f(t)-f(x)}{t-x} \leq \frac{f(y)-f(x)}{y-x} \leq \frac{f(y)-f(t)}{y-t}$, con $x < t < y$.

Puesto que f es contínua en [x,y] y diferenciable en (x,y) el teorema del valor medio nos asegura que existe $z\epsilon(x,y)$ tal, que $f'(z) \leq \frac{f(y)-f(x)}{y-x}$ y como f' es creciente en (x,y) tenemos que

i) $f'(t) \leq \frac{f(y)-f(x)}{y-x}$, si $x \leq t \leq z$

ii) $f'(t) \geq \frac{f(y)-f(x)}{y-x}$ si $z \leq t \leq y$.

Por otra parte,

a) $f'(t)(x-t)\leq f(x)-f(t)$ implica $f'(t)\geq\frac{f(t)-f(x)}{t-x}$, si $x<t$. (8)

b) $f'(t)(y-t)\leq f(y)-f(t)$ implica $f'(t)\leq\frac{f(y)-f(t)}{y-t}$, si $t<y$. (9)

[prop. 1, página 139].

Al combinar, respectivamente, (8) con (i) y (9) con (ii) obtenemos

$$\frac{f(y)-f(x)}{y-x} \geq \frac{f(t)-f(x)}{t-x}, \text{ si } x<t \qquad (10)$$

$$\frac{f(y)-f(t)}{y-t} \geq \frac{f(y)-f(x)}{y-x}, \text{ si } t<y \qquad (11)$$

y, por transitividad aplicada a (10) y (11), resulta que

$$\frac{f(t) - f(x)}{t - x} \leq \frac{f(y) - f(x)}{y - x} \leq \frac{f(y) - f(t)}{y - t}, \text{ si } x < t < y$$

Para f estrictamente creciente se sigue el mismo procedimiento con la desigualdad estricta.

2. Proposición.

Sea $f:(a,b) \subseteq \mathbb{R} \to \mathbb{R}$ una función dos veces diferenciable en (a,b). f es convexa si y sólo si $f''(x) \geq 0$, $\forall x \in (a,b)$. Si $f''(x) > 0$ sobre (a,b), f es estrictamente convexa en ese intervalo.

Demostración.

Si f es convexa, f' es creciente y $f''(x) \geq 0$ sobre (a,b). Si $f''(x) \geq 0$, $\forall x \in (a,b)$, f' es creciente y f es convexa sobre (a,b). Cuando sea $f''(x) > 0$, $\forall x \in (a,b)$, f' será estrictamente creciente y f, estrictamente convexa sobre (a,b).

Ahora, resulta fácil comprobar que las siguientes funciones son convexas:

- $f: \mathbb{R} \to \mathbb{R}$, $f(x) = e^x$.
- $g: \mathbb{R}^+ \to \mathbb{R}$, $g(x) = -\log x$.
- $h: [0,\infty) \to \mathbb{R}$, $h(x) = -x^p$, $0 < p < 1$.
- $j: \mathbb{R}^+ \to \mathbb{R}$, $j(x) = x \log x$
- $k: \mathbb{R} \to \mathbb{R}$, $k(x) = (a + bx^2)^{1/2}$; $a \geq 0$, $b \geq 0$.

3. Definición.

Sea $f: I \subseteq \mathbb{R} \to \mathbb{R}$ una función y $x_o \in I$. f tiene soporte en x_o si y sólo si existe una función afín $A(x) = f(x_o) + m(x - x_o)$ tal, que $A(x) \leq f(x)$, $\forall x \in I$. El gráfico de la función soporte A es la recta de soporte para f en x_o.

4. Proposición.

$f:(a,b) \subseteq \mathbb{R} \to \mathbb{R}$ es una función convexa si y sólo si existe, al menos, una recta de soporte para f en cada $x_o \in (a,b)$.

Demostración.

Supongamos f convexa, $x_o \in (a,b)$ y $m \in [f'_-(x_o), f'_+(x_o)]$. Entonces, $\dfrac{f(x)-f(x_o)}{x - x_o} \geq m$, si $x > x_o$; $\dfrac{f(x)-f(x_o)}{x - x_o} \leq m$, si $x < x_o$.

En cualquiera de los dos casos resulta que $f(x) \geq f(x_o) + m\,(x-x_o)$; luego, hay al menos, una línea de soporte para f en cada punto de (a,b) ya que x_o es un elemento cualquiera de (a,b).

Antes de mostrar la segunda parte del teorema observemos que si $A:(a,b) \subseteq \mathbb{R} \to \mathbb{R}$ es una función afín; entonces, $A[\alpha x+(1-\alpha)y] = \alpha A(x) + (1-\alpha)\, A(y)$.

En efecto, por ser A afín, su ecuación es de la forma $A(x) = mx + b$. Luego,

$$A[\alpha x+(1-\alpha)y] = m[\alpha x+(1-\alpha)y] + b = m\alpha x + \alpha b + m(1-\alpha)y + b - \alpha b$$
$$= \alpha(mx+b) + (1-\alpha)(my+b) = \alpha A(x)+(1-\alpha)A(y)$$

Ahora, supongamos que f tiene una línea soporte en cada punto de (a,b).

Sea $x,y\in(a,b)$, $x_0=\alpha x+(1-\alpha)y$, $\alpha\in[0,1]$ y $A(x)=f(x_0)+m(x-x_0)$, la función soporte de f en x_0. Entonces, tenemos que $f(x_0)=f[\alpha x+(1-\alpha)y] = A(x_0)=A[\alpha x+(1-\alpha)y]=\alpha A(x)+(1-\alpha)A(y)\leq \alpha f(x) + (1-\alpha) f(y)$ y, f es convexa.

1.4 Operaciones Cerradas.

Ahora estudiamos las operaciones con funciones convexas que dan como resultado otra función convexa.

1. Proposición.

Si $f: I\subseteq\mathbb{R}\to\mathbb{R}$, $g: I\subseteq\mathbb{R}\to\mathbb{R}$ son funciones convexas y si $\lambda\geq 0$; entonces, $f+g$ y λf son funciones convexas sobre I.

Demostración.

Como f y g son convexas, tenemos que $\forall x, y\in I$ y para $\alpha\in(0,1)$

$$f[\alpha x + (1-\alpha)y] \leq \alpha f(x) + (1-\alpha)f(y),$$
$$g[\alpha x + (1-\alpha)y] \leq \alpha g(x) + (1-\alpha) g(y).$$

Luego,

$(f+g)[\alpha x+(1-\alpha)y] \leq \alpha[(f+g)(x)]+(1-\alpha)[(f+g)(y)]$ y $f+g$ es convexa.

De inmediato se deduce que λf es convexa pues,

$$\lambda f[\alpha x+(1-\alpha)y] = \lambda \{f[\alpha x+(1-\alpha)y]\} \leq \lambda [\alpha f(x)+(1-\alpha)f(y)]$$
$$\leq \alpha [(\lambda f)(x)] + (1-\alpha) [(\lambda f)(y)].$$

2. <u>Proposición.</u>

Sea $f: I \subseteq \mathbb{R} \to \mathbb{R}$ y $g: J \subseteq \mathbb{R} \to \mathbb{R}$ dos funciones tales que el rango de f es un subconjunto de J. Si f y g son convexas con g creciente; entonces, la función compuesta $g\circ f: I \subseteq \mathbb{R} \to \mathbb{R}$ es convexa.

<u>Demostración.</u>

Sea $x, y \in I$ y $\alpha \in (0,1)$.

$(g\circ f)[\alpha x+(1-\alpha)y] = g\{f[\alpha x+(1-\alpha)y]\} \leq g[\alpha f(x)+(1-\alpha)f(y)]$ ya que f es convexa y g, creciente. Luego,

$(g\circ f)[\alpha x+(1-\alpha)y] \leq g[\alpha f(x)]+g[(1-\alpha)f(y)] \leq \alpha[(g\circ f)(x)]+ (1-\alpha)[(g\circ f)(y)]$ porque g es convexa. El teorema está probado.

3. <u>Proposición.</u>

Si $f: I \subseteq \mathbb{R} \to \mathbb{R}$ y $g: I \subseteq \mathbb{R} \to \mathbb{R}$ son dos funciones no negativas, decrecientes [crecientes] y convexas; entonces; $h(x) = f(x)g(x)$ también es no negativa, decreciente [creciente] y convexa.

Demostración.

Es claro que el producto fg es no negativo.

Sea $x,y\in I$ con $x<y$. Entonces, $f(x)\geq f(y)$ y $g(x)\geq g(y)$ implica $(fg)(x)\geq(fg)(y)$. Es decir, fg es decreciente.

Es obvio el resultado $(fg)(x)\leq(fg)(y)$ cuando f y g son crecientes y $x < y$.

Ahora, mostramos que h es convexa.

Si $x < y$; entonces,

$$[f(x)-f(y)]\,[g(y)-g(x)] \leq 0 \text{ y}$$

$$f(x)g(y)+f(y)g(x) \leq f(x)g(x) + f(y)g(y) \tag{12}$$

Para $\alpha > 0$, $\beta > 0$ y $\alpha+\beta = 1$ tenemos

$$\begin{aligned} f(\alpha x+\beta y)g(\alpha x+\beta y) &\leq [\alpha f(x)+ \beta f(y)]\,[\alpha g(x)+\beta g(y)] \\ &\leq \alpha^2 f(x)g(x)+\alpha\beta[f(x)g(y)+f(y)g(x)]+\beta^2 f(y)g(y) \\ &\leq \alpha^2 f(x)g(x)+\alpha\beta[f(x)g(x)+f(y)g(y)]+\beta^2 f(y)g(y), \text{ por (12)} \\ &\leq \alpha(\alpha+\beta)f(x)g(x)+\beta(\alpha+\beta)f(y)g(y) \\ &\leq \alpha f(x)g(x)+\beta f(y)g(y) = \alpha[(fg)(x)]+\beta[(fg)(y)] \end{aligned}$$

La demostración de la convexidad de fg en el caso en que f y g son crecientes es similar.

4. Proposición.

Sea $(f_i)_{i\in A}$ una familia arbitraria de funciones convexas con $f_i\colon I \subseteq \mathbb{R} \to \mathbb{R}$. Si $f(x)=\sup_{i\in A} f_i(x)$ y $J=\{x\in I/f(x)<\infty\}$ es no vacío; entonces, J es un intervalo y f es convexa sobre J.

Demostración.

Sea $x, y \in J$ y $\alpha \in (0,1)$. Entonces,

$$f(z) = f[\alpha x + (1-\alpha)y] = \sup_{i\in A} f_i [\alpha x + (1-\alpha)y]$$

$$\leq \sup_{i\in A} [\alpha f_i(x)+(1-\alpha)f_i(y)] \text{ , por ser } f_i \text{ convexa, } i\in A.$$

$$\leq \alpha \sup_{i\in A} f_i(x)+(1-\alpha) \sup_{i\in A} f_i(y), \text{ por propiedad del sup.}$$

$$\leq \alpha f(x) + (1-\alpha) f(y).$$

Como $f(x)<\infty$, $f(y)<\infty$, $\alpha f(x)+(1-\alpha)f(y)<\infty$ y $z\in J$. Así, J es un intervalo (contiene todo punto entre dos cualesquiera de sus puntos) y f es convexa sobre él.

5. Proposición.

Si f_n: $I \subseteq \mathbb{R} \to \mathbb{R}$ es una sucesión de funciones convexas que convergen a una función límite acotada f: $I \subseteq \mathbb{R} \to \mathbb{R}$; entonces, f es convexa. Aún más, la convergencia es uniforme sobre cualquier subintervalo cerrado de I°, el interior de I.

Demostración.

Sea $x,y\in I$ y $\alpha\in(0,1)$. Entonces,

$$f[\alpha x+(1-\alpha)y] = \lim_{n\to\infty} f_n [\alpha x+(1-\alpha)y]$$

$$\leq \lim_{n\to\infty} [\alpha f_n(x) + (1-\alpha)f_n(y)]$$

$$\leq \alpha \lim_{n\to\infty} f_n(x) + (1-\alpha) \lim_{n\to\infty} f_n(y)$$

$$\leq \alpha f(x) + (1-\alpha)f(y) \text{ y f es convexa.}$$

Sea $a < c < b$ tres puntos cualesquiera de I, $\alpha = \sup f_n(a)$, $\gamma = \inf f_n(c)$, $\beta = \sup f_n(b)$. Sea, además, L_1, L_2 y L_3 las tres funciones afines que satisfacen $L_1(a) = \alpha$, $L_1(b) = \beta$, $L_2(c) = \gamma$, $L_2(b) = \beta$, $L_3(a) = \alpha$ y $L_3(c) = \gamma$. [fig. 3.3]

Mostraremos que $\{f_n\}$ está uniformemente acotada por estas funciones afines.

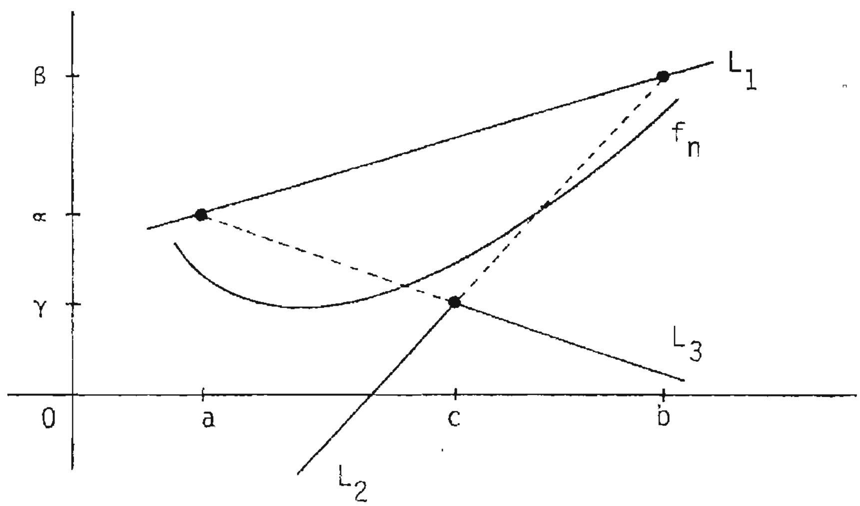


fig. 3.3

Si $x = \lambda a + (1-\lambda)b$ es cualquier punto de $[a,b]$; entonces, para un n arbitrario, $f_n(x) \leq \lambda f_n(a)+(1-\lambda)f_n(b) \leq \lambda L_1(a) + (1-\lambda)L_1(b) = L_1[\lambda a+(1-\lambda)b] = L_1(x)$, puesto que L_1 es afín.

Si x es cualquier punto de $[a,c]$ podemos escribir

$c = \lambda x + (1-\lambda)b$, con $\lambda\in(0,1)$ y $x = \frac{1}{\lambda} c + \frac{\lambda-1}{\lambda} b$. Luego,

$L_2(c) \leq f_n(c) \leq \lambda f_n(x) + (1-\lambda) f_n(b) \leq \lambda f_n(x) + (1-\lambda) L_2(b)$ y

$f_n(x) \geq \frac{1}{\lambda} L_2(c) + \frac{\lambda-1}{\lambda} L_2(b) = L_2\left(\frac{1}{\lambda} c + \frac{\lambda-1}{\lambda} b\right) = L_2(x)$, puesto que L_2 es afín.

Similarmente, si $x\in [c,b]$ escribimos $c = \lambda x + (1-\lambda)a, \lambda\in(0,1]$. Luego, $x = \frac{1}{\lambda} c + \frac{\lambda-1}{\lambda} a$,

$L_3(c) \leq f_n(c) = \lambda f_n(x) + (1-\lambda) f_n(a) \leq \lambda f_n(x) + (1-\lambda) L_3(a)$ y

$f_n(x) \geq \frac{1}{\lambda} L_3(c) + \frac{\lambda-1}{\lambda} L_3(a) = L_3\left(\frac{1}{\lambda} c + \frac{\lambda-1}{\lambda} a\right) = L_3(x)$, ya que también L_3 es afín.

Así, $\{f_n\}$ está acotada superiormente por L_1 sobre $[a,b]$, inferiormente por L_2 sobre $[a,c]$ y por L_3 sobre $[c,b]$.

Por otra parte, sabemos que una función convexa satisface una condición de Lipschitz sobre cualquier compacto contenido en el interior de su dominio I. Luego, existe K, independiente del número natural n, tal que
$|f_n(y) - f_n(x)| \leq K|y-x|$; $\forall n\in\mathbb{N}$, $\forall x, y\in I$.

Ahora, sea $\varepsilon > 0$ y seleccionemos un conjunto finito E de puntos de $[a,b]$ de modo que cada punto de este intervalo esté a una distancia menor o igual que $\frac{\varepsilon}{3K}$ de, al menos, un punto de E. Puesto que E es finito existe $N \in \mathbb{N}$ para el -

cual $m,n \geq N$ implica $|f_n(z) - f_m(z)| \leq \frac{\varepsilon}{3}$, $\forall z \in E$, ya que - la sucesión es convergente. Luego, si

$x \in [a,b]$, $z \in E$, $|z - x| < \frac{\varepsilon}{3K}$ y $m,n \geq N$ tendremos que

$$|f_n(x)-f_m(x)| \leq |f_n(x)-f_n(z)|+|f_n(z)-f_m(z)|+|f_m(z)-f_m(x)|$$
$$\leq K|x-z| + \frac{\varepsilon}{3} + K|x-z| \leq \varepsilon.$$

lo cual es, precisamente, la condición de Cauchy para la convergen - cia uniforme sobre $[a,b]$.

1.5 Funciones Logarítmicamente Convexas.

1. Definición.

La función $f: I \subseteq \mathbb{R} \to \mathbb{R}$ es logarítmicamente con - vexa si f es positiva y si la función log f es convexa sobre I.

2. Proposición.

$f: I \subseteq \mathbb{R} \to \mathbb{R}$ es logarítmicamente convexa si y sólo si f es positiva y $f(\alpha x + \beta y) \leq f^{\alpha}(x)\ f^{\beta}(y)$, con $x,y \in I$, $\alpha > 0$, $\beta > 0$, $\alpha + \beta = 1$.

Demostración.

Supongamos f logarítmicamente convexa. Entonces, f es positiva y log f, convexa sobre I; luego, para $\alpha > 0$, $\beta > 0$ y $\alpha + \beta = 1$ tenemos que

$$\begin{aligned}\log f(\alpha x + \beta y) &\leq \alpha \log f(x) + \beta \log f(y)\\ &\leq \log f^{\alpha}(x) + \log f^{\beta}(y)\\ &\leq \log f^{\alpha}(x)\ f^{\beta}(y)\end{aligned}$$

De esta última desigualdad se deduce que $f(\alpha x+\beta y) \leq f^{\alpha}(x)\, f^{\beta}(y)$.

El recíproco se muestra siguiendo un proceso inverso.

Puesto que $f(x) = \exp[\log f(x)]$ es una composición de dos funciones convexas con exp una función creciente, se sigue que una función logarítmicamente convexa es convexa.

3. Proposición.

La suma y el producto de dos funciones logarítmicamente convexas definidas sobre un intervalo I son funciones logarítmicamente convexas. También el límite de una sucesión convergente de funciones logarítmicamente convexas es logarítmicamente convexa, si tal función existe y es positiva.

Demostración.

Sea $a,b,c,d,\alpha,\beta\in \mathbb{R}^{+}$ con $\alpha+\beta = 1$. Entonces,

$a^{\alpha}b^{\beta}= \exp[\log a^{\alpha}b^{\beta}]=\exp[\alpha\log a+\beta\log b]\leq\alpha[\exp(\log a)]+\beta[\exp(\log b)]$ y

$a^{\alpha}b^{\beta}\leq\alpha a + \beta b$.

Luego,

$$\frac{a^{\alpha}b^{\beta} + c^{\alpha}d^{\beta}}{(a+c)^{\alpha}(b+d)^{\beta}} = \left(\frac{a}{a+c}\right)^{\alpha}\left(\frac{b}{b+d}\right)^{\beta} + \left(\frac{c}{a+c}\right)^{\alpha}\left(\frac{d}{b+d}\right)^{\beta}$$

$$\leq \alpha\frac{a}{a+c} + \beta\frac{b}{b+d} + \alpha\frac{c}{a+c} + \beta\frac{d}{b+d} = \alpha + \beta = 1 \text{ y},$$

$$a^{\alpha}b^{\beta}+ c^{\alpha}d^{\beta}\leq(a+c)^{\alpha}(b+d)^{\beta}. \qquad (13)$$

Como f y g son logarítmicamente convexas; para $x, y \in I$ tenemos que

$f(\alpha x+\beta y)+g(\alpha x+\beta y) \leq f^{\alpha}(x)\, f^{\beta}(y) + g^{\alpha}(x)\, g^{\beta}(y) \leq [(f+g)^{\alpha}(x)]\, [(f+g)^{\beta}(y)]$, de acuerdo con la desigualdad (13). Lo que muestra la afirmación para la suma.

Si $x,y \in I$, $\alpha>0, \beta>0$ y $\alpha+\beta=1$; entonces,

$$\begin{aligned}
\log [(fg)(\alpha x+\beta y)] &= \log f(\alpha x+\beta y) + \log g(\alpha x+\beta y)\\
&\leq \alpha \log f(x) + \beta \log f(y) + \alpha \log g(x) + \beta \log g(y)\\
&\leq \alpha \log [(fg)(x)] + \beta \log [(fg)(y)]\\
&\leq \log [(fg)(x)]^{\alpha} [(fg)(y)]^{\beta}\ \text{ y,}
\end{aligned}$$

$(fg)(\alpha x+\beta y) \leq [(fg)(x)]^{\alpha} [(fg)(y)]^{\beta}$, y queda mostrada la afirmación para el producto.

Sea $\{f_n\}$, con $f_n: I \subseteq \mathbb{R} \to \mathbb{R}$, una sucesión convergente de funciones logarítmicamente convexas cuyo límite es $f: I \subseteq \mathbb{R} \to \mathbb{R}$; $\alpha>0$, $\beta>0$ y $\alpha+\beta=1$. Entonces,

$$\begin{aligned}
\log \left[\lim_{n\to\infty} f_n(\alpha x+\beta y)\right] &= \lim_{n\to\infty} [\log f_n(\alpha x+\beta y)] \text{ (Por continuidad de Log).}\\
&\leq \lim_{n\to\infty} [\alpha \log f_n(x) + \beta \log f_n(y)]\\
&\leq \alpha \left[\log \lim_{n\to\infty} f_n(x)\right] + \beta \left[\log \lim_{n\to\infty} f_n(y)\right]; \text{ es decir,}
\end{aligned}$$

$$\begin{aligned}
\log f(\alpha x+\beta y) &\leq \alpha \log f(x) + \beta \log f(y)\\
&\leq \log f^{\alpha}(x)\, f^{\beta}(y).
\end{aligned}$$

De esta última desigualdad resulta que $f(\alpha x+\beta y) \leq f^{\alpha}(x) f^{\beta}(y)$.

2. Funciones Convexas en Espacios Vectoriales Normados.

En esta sección generalizamos el concepto de función convexa con valores reales tomando, como conjunto de partida, un espacio vectorial normado cualquiera L. Luego, estudiamos la continuidad de estas funciones y algunas caracterizaciones de las funciones convexas, por medio de la diferenciabilidad.

1. Definición.

Sea L un espacio vectorial normado y $U \subseteq L$ un conjunto convexo. Entonces; $f:U \to \mathbb{R}$ es una función convexa sobre U si

$f[\alpha x+(1-\alpha)y] \leq \alpha f(x) + (1-\alpha)f(y)$; $\forall x, y \in U$; $\alpha \in (0,1)$

Las demostraciones de las proposiciones que se refieren al comportamiento de una función convexa $f: U \subseteq L \to \mathbb{R}$, en un punto particular $x_0 \in U$, se simplifican utilizando la función

$g:V \subseteq L \to \mathbb{R}$, con $g(x) = f(x+x_0) - f(x_0)$ y $V = \{ x \in L/(x+x_0) \in U \}$ ya que, el estudio de g es equivalente al de f, como lo veremos más adelante. Observemos que si $y = x + x_0$, $x = y - x_0$ y cuando $y = x_0$, $x = \bar{o}$; es decir, el origen está en el dominio de g. Además, $g(\bar{o}) = 0$.

2. Proposición.

Sea $f: U \subseteq L \to \mathbb{R}$ una función convexa y $g: V \subseteq L \to \mathbb{R}$, la función definida por $g(x) = f(x+x_o) - f(x_o)$ con $V = \{x \in L/(x + x_o) \in U\}$. Entonces, V es un conjunto convexo y g, una función convexa. Además, V es un conjunto abierto (cerrado) si y sólo si U es un conjunto abierto (cerrado).

<u>Demostración.</u>

Sea $x_1, x_2 \in V$ y $\alpha \in (0,1)$. Entonces, $(x_1 + x_o)$, $(x_2 + x_o)$ pertenecen a U e $y = \alpha(x_1+x_o)+(1-\alpha)(x_2+x_o) = \alpha x_1+(1-\alpha)x_2+x_o$ está en U, por ser U convexo; luego, $\alpha x_1+(1-\alpha)x_2$ está en V y V es convexo.

V es abierto (cerrado) si y sólo si U es abierto (cerrado) ya que toda traslación es una función contínua y V se obtiene de U por traslación.

Por último, mostramos que g es convexa.

$$
\begin{aligned}
g[\alpha x_1+(1-\alpha)x_2] &= f[\alpha x_1+(1-\alpha)x_2+x_o] - f(x_o)\\
&= f[\alpha(x_1+x_o)+(1-\alpha)(x_2+x_o)] - f(x_o)\\
&\leq \alpha f(x_1+x_o)+(1-\alpha)f(x_2+x_o) - f(x_o)\\
&\leq \alpha[f(x_1+x_o)-f(x_o)]+(1-\alpha)[f(x_2+x_o)-f(x_o)]\\
&\leq \alpha\, g(x_1) + (1-\alpha)\, g(x_2)
\end{aligned}
$$

3. <u>Definición.</u>

Sea U un subconjunto del espacio vectorial normado

L y $f: U \to \mathbb{R}$, una función. Se dice que f está localmente acotada en U, si para cada $x \in U$ existe un vecindario sobre el cual f está acotada.

4. Proposición.

Sea $U \subseteq L$ un conjunto abierto y convexo del espacio vectorial normado L y $f: U \to \mathbb{R}$, una función convexa. f está localmente acotada en x_o si y sólo si la función g, de la proposición 2, está localmente acotada en $\bar{o}$.

Demostración.

Si f está localmente acotada en x_o existe un vecindario $V(x_o,\varepsilon) \subseteq U$ tal que $|f(x)| \leq M$, $\forall\ x \in V(x_o,\varepsilon)$. Sea $(x + x_o) \in V\ (x_o,\varepsilon)$; entonces, $(x + x_o) \in U$ y $x \in V$. Además, como $||x+x_o-x_o|| = ||x|| < \varepsilon$, $x\in V\ (\bar{o},\varepsilon)$; luego, $g(x) = f(x+x_o)-f(x_o) \leq M- f(x_o)$ y g está local y superiormente acotada en el origen. Mostraremos que g también está local e inferiormente acotada en $\bar{o}$. Puesto que, $||-x||=||x|| < \varepsilon$, $-x\in V(\bar{o},\varepsilon)$ y como $\bar{o} = \frac{1}{2} x + \frac{1}{2} (-x)$ y g es convexa, $g(\bar{o}) \leq \frac{1}{2} g(x) + \frac{1}{2} g(-x)$. Luego, $g(x)\geq 2g(\bar{o})-g(-x) =-g(-x) \geq f(x_o)-M$ ya que, $g(\bar{o}) = 0$ y $g(-x) \leq M-f(x_o)$. Así, g está localmente acotada en el origen.

Recíprocamente, supongamos que g está localmente acotada en el origen. Entonces, existe un vecindario $V(\bar{o},\varepsilon) \subseteq V$ tal que

$g(x) \leq M$, $\forall x \in V(\bar{o},\varepsilon)$. Si x está en ese vecindario, $\|x\| < \varepsilon$, $x \in V$ y si $x_o \in U$, $(x+x_o) \in U$. Además, $(x+x_o) \in V(x_o,\varepsilon)$ ya que, $\|x+x_o - x_o\| = \|x\| < \varepsilon$. Luego, $f(x+x_o) - f(x_o) = g(x) \leq M$ implica que $f(x+x_o) \leq M+f(x_o)$.

Así, f está local y superiormente acotada en x_o puesto que $x + x_o$ es un punto cualquiera de $V(x_o,\varepsilon)$. Sólo falta mostrar que f está local e inferiormente acotada en x_o.

Como $\|x_o-x-x_o\| = \|-x\| < \varepsilon$, $(x_o - x) \in V(x_o, \varepsilon)$, f convexa y $x_o = \frac{1}{2}(x+x_o)+\frac{1}{2}(x_o-x)$, $f(x_o) \leq \frac{1}{2}f(x+x_o) + \frac{1}{2} f(x_o-x)$. Así, tenemos que $f(x + x_o) \geq 2 f(x_o) - f(x_o-x) \geq f(x_o) - M$ pues, $f(x_o-x) \leq M + f(x_o)$ implica $- f(x_o-x) \geq -M-f(x_o)$

5. <u>Proposición.</u>

Sea L un espacio normado y d, la distancia inducida por la norma de L. Entonces, para cada $A \subseteq E$ no vacío y para cada $r \in \mathbb{R}^+$, $V(A,r)=\{x \in L / d(x,A) < r\}$ es un vecindario abierto de A.

<u>Demostración.</u>

Si $x \in V(A,r)$, $d(x,A) < r$ y $r-d(x,A) > 0$. Consideremos la bola abierta $B(x,r-d(x,A))$ y mostremos que está incluida en $V(A,r)$. Si y pertenece a esta bola, $d(y,x) < r-d(x,A)$. Como $d(y,A) - d(x,A) < d(y,x)$ tenemos que $d(y,A) < r$. Así, $y \in V(A,r)$ y para cada $x \in V(A,r)$ podemos encontrar una bola $B(x, r-d(x,A))$

completamente contenida en $V(A,r)$. Además $A \subseteq V(A,r)$ ya que $d(x,A) = 0 < r$, si $x \in A$. Así, $V(A,r)$ es un vecindario abierto de A.

Cuando $A = \{a\}$, $V(A,r) = V(\{a\},r) = \{x \in L / d(x,a) < r\} = B(a,r)$ y se dice que esta bola abierta es un vecindario de a.

6. Proposición.

Sea $U \subseteq L$ un conjunto abierto y convexo del espacio vectorial normado L y $f: U \to \mathbb{R}$, una función convexa. Si f está acotada superiormente en un vecindario de un punto $x_0 \in U$; entonces, f está localmente acotada.

Demostración.

Supongamos, por conveniencia, que f está acotada superiormente por B en el vecindario $V(\bar{o}, \varepsilon)$. Entonces, de acuerdo con la primera parte de la demostración de la prop. 4, pág. 157, f está inferiormente acotada en $V(\bar{o}, \varepsilon)$.

Considerando como cierto que f está acotada superiormente por B en $V(\bar{o}, \varepsilon)$ mostraremos que f está acotada en un vecindario de $y \in U$, con $y \neq \bar{o}$.

Seleccionemos $r > 1$ de modo que $z = ry$ estén en U y hagamos $\lambda = 1/r$; entonces, de acuerdo con la prop. 5, pág. 158, $M = \{v \in L / v = (1-\lambda)\, x + \lambda z,\ x \in V(\bar{o},\varepsilon)\}$ es un vecindario de $y = \lambda z$ con radio $(1-\lambda)\varepsilon$ ya que,

$\|(1-\lambda)x + \lambda z - \lambda z\| = (1-\lambda)\|x\| < (1-\lambda)\varepsilon$.

Luego, $f(v) \leq (1-\lambda) f(x) + \lambda f(z)$.

Ahora, puesto que $0 < \lambda < 1$, $\lambda f(z) \leq f(z)$, si $f(z) \geq 0$.

Además, $(1-\lambda) f(x) \leq B$. Así que, en este caso, $f(v) \leq B + f(z)$.

Por otra parte, si $f(z) \leq 0$, $\lambda f(z) \leq 0$. Además, como $x \in V(\bar{o},\varepsilon)$, $f(x) \geq f(\bar{o}) - B$, $-\lambda f(x) \leq \lambda [B - f(\bar{o})]$ y $\lambda [f(z) - f(x)] \leq \lambda [B - f(\bar{o})]$. Así, de $f(v) \leq f(x) + \lambda[f(z) - f(x)]$ se deduce que $f(v) \leq B + \lambda [B - f(\bar{o})]$.

En ambos casos, f está local y superiormente acotada en un vecindario de y; luego, de acuerdo con la segunda parte de la prop. 4, pág.157, f está localmente acotada en el punto arbitrario $y \in U$ y por eso, localmente acotada en U.

7. Proposición.

Sea f una función convexa definida sobre el conjunto abierto y convexo U del espacio vectorial normado L y $x_o \in U$, un punto arbitrario. Entonces, $g: (a,b) \subseteq \mathbb{R} \to \mathbb{R}$, con $g(t) = f(x_o + ty)$, con $(x_o + ty) \in U$; (a,b) un intervalo que contiene al origen e $y \in L$, un punto cualquiera; es una función convexa.

Demostración.

(a,b) es un conjunto convexo abierto por ser un intervalo abierto. Ahora, si $t_1, t_2 \in (a,b)$; entonces, para $\alpha \in (0,1)$

$$g\,[\alpha t_1 + (1-\alpha)t_2] = f\,\{x_o + [\alpha t_1 + (1-\alpha)\,t_2]\,y\}$$

$$= f\,[\alpha(x_o + t_1 y) + (1-\alpha)\,(x_o + t_2\,y)]$$

$$\leq \alpha f(x_o + t_1 y) + (1-\alpha)\,f\,(x_o + t_2 y)$$

$$\leq \alpha g(t_1) + (1-\alpha)\,g\,(t_2).$$

8. Definición.

Sea $U \subseteq L$ un conjunto abierto en un espacio vectorial normado L y $f: U \to \mathbb{R}$, una función. Se dice que f es localmente - Lipschitz si para cada $x \in U$ existen, un vecindario $V(x,\varepsilon)$ y una constante $k(x)$ tales que, si y, $z \in V\,(x,\varepsilon)$; entonces, $|f(y) - f(z)| \leq k\,\|y-z\|$. Cuando esta desigualdad es cierta en $W \subseteq U$, con k independiente de x, se dice que f es - Lipschitz sobre W.

9. Proposición.

Sea f y g las funciones definidas en la prop. 2,pág.155. Entonces, f es Lipschitz sobre U si y sólo si g es Lipschitz sobre V.

Demostración.

Si f es Lipschitz en U, $|f(x)-f(y)| \leq k\|x-y\|$,$\forall x, y \in U$. Sea x_1, $x_2 \in V$; entonces, $g(x_1) = f(x_1 + x_o) - f(x_o)$,$g(x_2)=f(x_2+x_o)-f(x_o)$,

$$|g(x_1)-g(x_2)| = |f(x_1+x_o)-f(x_2+x_o)| \leq k\|x_1-x_2\|.$$

Si g es Lipschitz en V, $|g(x)-g(y)| \leq k\|x-y\|$, $\forall x$, $y \in V$. Enton ces, ya que $x_1 + x_\circ \in U$ y $x_2 + x_\circ \in U$,

$f(x_1 + x_\circ) - f(x_2 + x_\circ) = f(x_1 + x_\circ) - f(x_\circ) - [f(x_2 + x_\circ) - f(x_\circ)]$

y, $|f(x)-f(y)| = |g(x_1)-g(x_2)| \leq k\|x_1 - x_2\| = k\|x_1 + x_\circ - (x_2 + x_\circ)\| = k\|x - y\|$.

10. Proposición.

Sea $U \subseteq L$ un conjunto abierto y convexo del espacio vectorial normado L y $f\colon U \to \mathbb{R}$ una función convexa acotada superiormente en un vecindario de un punto de U; entonces, f es localmente Lipschitz en U y Lipschitz sobre cualquier subconjunto compacto de U.

Demostración.

De acuerdo con la prop. 6, pág. 159, f está localmente acotada y por eso, para $x_\circ \in U$ podemos encontrar un vecindario $V(x_\circ, 2\varepsilon) \subseteq U$ sobre el cual f está acotada, digamos, por M. Supongamos que f no satisface la condición de Lipschitz sobre $V(x_\circ, \varepsilon)$; entonces, podemos seleccionar x_1, $x_2 \in V(x_\circ, \varepsilon)$ tales que $\dfrac{f(x_2) - f(x_1)}{\|x_2 - x_2\|} > \dfrac{2M}{\varepsilon}$ (14).

Si $x_3 = x_2 + \alpha(x_2 - x_1)$ está en $V(x_\circ, 2\varepsilon)$ y $\|x_3 - x_2\| = \varepsilon$, $\alpha > 0$. En efecto, si $\|x_3 - x_2\| = \varepsilon$,

tenemos $\alpha = \dfrac{\varepsilon}{\|x_2 - x_1\|} > 0$ y si $\|x_3 - x_2\| < 2\varepsilon$ resulta que

$$\alpha \, \| x_2 - x_1 \| - \| x_2 - x_\circ \| < \| \alpha (x_2 - x_1) + (x_2 - x_\circ) \| < 2 \, \varepsilon,$$

$$\alpha \| x_2 - x_1 \| < 2\varepsilon + \| x_2 - x_\circ \| < 3\varepsilon \, , \; \alpha < 3 \frac{\varepsilon}{\| x_2 - x_1 \|} = 3\alpha \, , \; y \; \alpha > 0 .$$

De $x_3 = x_2 + \alpha (x_2 - x_1)$ se deduce que $x_2 = \frac{x_3}{1+\alpha} + \frac{\alpha x_1}{1+\alpha}$ es una combinación convexa de x_1 y x_3. Luego, por ser f convexa sobre el segmento que contiene a x_1, x_2 y x_3 podemos es - cribir que

$$\frac{f(x_3) - f(x_2)}{\| x_3 - x_2 \|} \geq \frac{f(x_2) - f(x_1)}{\| x_2 - x_1 \|} \; . \qquad (15)$$

Aplicando transitividad en (14) y (15) tenemos

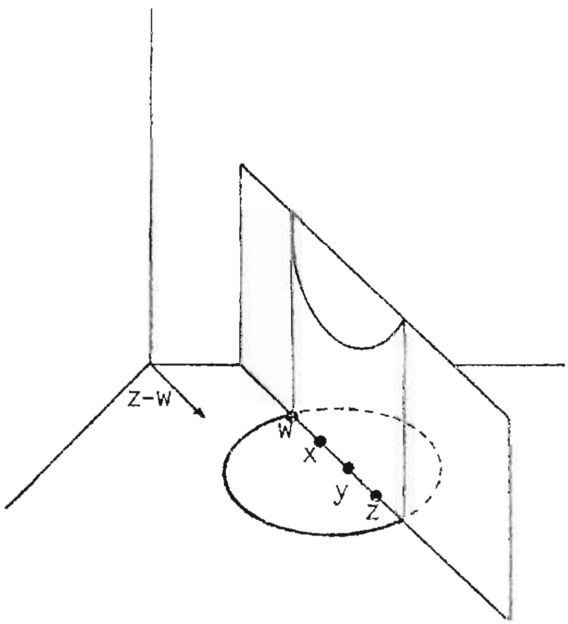
$\frac{f(x_3) - f(x_2)}{\| x_3 - x_2 \|} > \frac{2M}{\varepsilon}$, $f(x_3) - f(x_2) > 2M$ y $f(x_3) > M + f(x_2)$, contradiciendo el hecho de que f está acotada por M en $V(x_\circ, 2\varepsilon)$.

Ahora, mostremos que f es Lipschitz sobre cualquier - subconjunto compacto V de U. Sea $\{N_i\}$ un cubrimiento abierto de V. Siendo V compacto podemos extraer de $\{N_i\}$ un cubrimiento finito $C = \{N_1, \ldots, N_r\}$ tal que $V \subseteq \bigcup N_i$, $i=1,\ldots,r$.

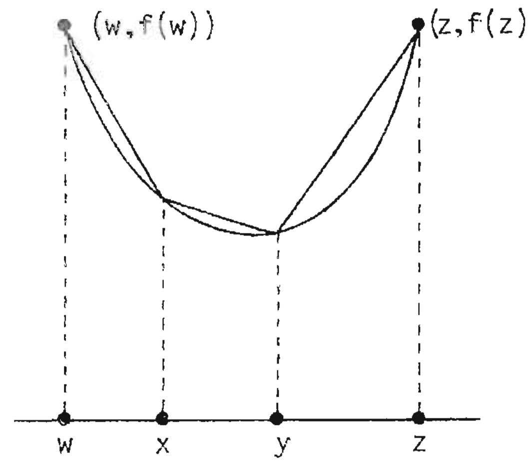
Tomemos $x_\circ \in N_s \in C$, $x_\circ \in U$; entonces, N_s es un vecindario abierto de $x_\circ$ y existen u, v $\in N_s$, con u,v $\in$ U para los cuales, por ser f localmente Lipschitz, $|f(u) - f(v)| \leq k \| u - v \|$. Se - leccionemos x, y $\in$ V; entonces, x $\in N_i$, y $\in N_j$ y existen - w, z $\in$ U con w $\in N_i$, z $\in N_j$ tales que x, y $\in$ Seg $[w,z]$.

Luego, definimos

$g : (a,b) \subseteq \mathbb{R} \to \mathbb{R}$, con $g(t) = f[w + t(z-w)]$; $a,b \in \mathbb{R}$. (fig. 3.4).



(a)



(b)

fig. 3.4

En el gráfico (b) de la figura 3.4 puede verse que la relación (16) $\frac{f(x) - f(w)}{x - w} \leq \frac{f(y) - f(x)}{y - x} \leq \frac{f(z) - f(y)}{z - y}$ es cierta porque g es convexa.

Para w, x $\in N_i$; y, z $\in N_j$ tenemos, respectivamente, que $-k_1 \leq \frac{f(x) - f(w)}{\|x - w\|} \leq k_1$, $-k_2 \leq \frac{f(z) - f(y)}{\|z - y\|} \leq k_2$, por ser f localmente Lipschitz.

Si tomamos $k = \text{máx}\ \{k_1, k_2\}$ podemos escribir, de acuerdo con la relación (16) que $-k \leq \frac{f(y) - f(x)}{\|y - x\|} \leq k$ y luego, que $|f(y) - f(x)| \leq k\|y-x\|$.

10.1 Corolario.

g, la función definida en la prop. 2, pág. 155 es localmente Lipschitz en V y Lipschitz sobre cualquier subconjunto compacto de V.

Este corolario se deduce inmediatamente de las proposiciones 9 y 10, páginas 161 y 162.

2.1 Continuidad.

1. Proposición.

Sea f y g las funciones definidas en la prop. 2, pág. 155. Entonces, f es continua en $x_\circ \in U$ si y sólo si g es continua en $\bar{o}$.

Demostración.

f continua en $x_\circ$ implica que $\forall\ \varepsilon > 0\ \exists\ \delta > 0$ tal que si $x \in U$ y $\|x - x_\circ\| < \delta$; entonces, $|f(x) - f(x_\circ)| < \varepsilon$. Como $x \in U$ podemos escribir $x = x_1 + x_\circ$, con $x_1 \in V$. Luego, $\forall\ \varepsilon > 0\ \exists\ \delta > 0$ tal que $\|x_1 + x_\circ - x_\circ\| = \|x_1\| < \delta$ implica $|f(x_1 + x_\circ) - f(x_\circ)| = |g(x_1)| < \varepsilon$. Así, g es continua en $\bar{o}$.

Con un razonamiento similar iniciado con g se deduce que f es continua en x_0.

2. Proposición.

Sea L un espacio vectorial normado y $f: U \subseteq L \to \mathbb{R}$ una función convexa definida sobre el conjunto abierto y convexo U. Si f está acotada superiormente en un vecindario de un punto de U; entonces, f es continua en U.

Demostración.

De acuerdo con la hipótesis podemos afirmar que f es localmente Lipschitz. Luego, para cada $x \in U$ existe un vecindario $V(x,\varepsilon)$ y una constante $K(x)$ tales que si $y, z \in V(x,\varepsilon)$, $|f(y) - f(z)| \leq K \|y-z\|$. Así que, para $z = x$ tenemos $|f(y) - f(x)| \leq K \|y - x\|$ y si hacemos $\delta = \varepsilon/K$ tendremos $K \|y - x\| < \varepsilon$ y $|f(y) - f(x)| < \varepsilon$. Por lo tanto, f es continua en x, $\forall\, x \in U$; es decir, continua en U.

3. Proposición.

Sea $U \subseteq \mathbb{R}^n$ un conjunto abierto y convexo, $f: U \to \mathbb{R}$ una función convexa. Entonces, f es Lipschitz sobre cualquier subconjunto compacto de U y continua en U.

Demostración.

Haremos la demostración considerando, de nuevo, la función g de la prop. 2, pág. 155. Seleccionemos $\alpha > 0$ suficientemente pequeño de modo que la envolvente convexa - $V = H(\bar{o}, \alpha e_1, \ldots, \alpha e_n)$ sea un subconjunto de U. Observemos que V tiene un interior $\overset{\circ}{V} \neq \emptyset$. Luego, cualquier $x \in V$ tiene la representación $x = \lambda_\circ \bar{o} + \sum_1^n \lambda_i(\alpha e_i)$, con $\lambda_i \geq 0$, $\sum_o^n \lambda_i = 1$ y $g(x) \leq \lambda_\circ g(\bar{o}) + \sum_1^n \lambda_i\ g(\alpha e_i)$.

Sea $M = \text{máx}\ \{g(\bar{o}), \ldots, g\ (\alpha e_n)\}$. Entonces,

$$g(x) \leq \lambda_\circ\ g(\bar{o}) + \sum_{i=1}^n \lambda_i\ g(e_i) \leq M \left\{ \sum_{i \geq o} \lambda_i \right\}.$$ Así, g

está acotada superiormente sobre el conjunto abierto no vacío V; por lo tanto, g es Lipschitz sobre cualquier subconjunto compacto de V, prop.10, pág. 162 y continua sobre V, prop. 2, pág. 166. Ahora, el resultado de nuestro teorema se deduce de la prop. 1, página 165.

2.2 Funciones Convexas Diferenciales.

1. Proposición.

Sea f una función definida sobre un conjunto convexo abierto U de un espacio vectorial normado L. Si f es convexa sobre U y diferenciable en $x_\circ$; entonces, para $x \in U$,

$$f(x) - f(x_\circ) \geq f'(x_\circ)\ (x - x_\circ). \qquad (17)$$

Si f es diferenciable en U; entonces, f es convexa si y sólo si (17) es cierta para todo $x, x_\circ \in U$. Aún más, f es estric tamente convexa si y sólo si la desigualdad (17) es estricta.

Demostración.

Si f es convexa; entonces, para $t \in (0,1)$

$f[x_0 + t(x-x_0)] = f[(1-t)\, x_0 + tx] \leq (1-t)\, f(x_0)+t\, f(x)$.

Haciendo $h = x - x_0$ obtenemos

$f(x_0+ th) \leq f(x_0) + t[f(x) - f(x_0)]$ y $f(x_0+th) - f(x_0) \leq$

$$t\, [f(x_0 + h) - f(x_0)] \qquad (18)$$

Restando $f'(x_0)(th)$ de ambos miembros de (18) y dividiendo por t resulta $\dfrac{f(x_0+th)-f(x_0)-f'(x_0)(th)}{t} \leq$

$$f(x_0 + h) - f(x_0) - f'(x_0)\,(h) \qquad (19)$$

Ahora, cuando $t \to 0$, el miembros de la izquierda se aproxima a cero; mientras que, el miembro de la derecha siendo independiente de t, permanece constante. Luego,

$0 \leq f(x) - f(x_0) - f'(x_0)\,(x-x_0)$; $f(x) - f(x_0) \geq f'(x_0)\,(x - x_0)$.

Si f es estrictamente convexa, (18) será una desigualdad estricta la cual, al combinarse con (17), da como resultado

$t[f(x_0+h) - f(x_0)] > f(x_0+ th) - f(x_0) \geq f'(x_0)(x_0 + th - x_0)=f'(x_0)(th)$

Ahora, $t[f(x_0+h) - f(x_0)] > t\, f'(x_0)\,(h)$ implica

$f(x_0 + h) - f(x_0) > f'(x_0)\,(h)$; es decir, $f(x)-f(x_0)>f'(x_0)(x-x_0)$.

Supongamos que f es diferenciable y que satisface (17) en todo U. Para $x_1, x_2 \in U$, $t \in (0,1)$ hacemos $x_0=tx_1+(1-t)x_2$.

Entonces, $\bar{o} = tx_1 - tx_o + (1-t)x_2 + tx_o - x_o = t(x_1 - x_o) + (1-t)x_2 - (1-t)x_o$

$$= t(x_1 - x_o) + (1-t)\,(x_2 - x_o)$$

Luego, $0 = f'(x_o)\,[t(x_1 - x_o) + (1-t)(x_2 - x_o)]$, porque $f'(x_o)$ es lineal. $f(x_o) = f(x_o) + f'(x_o)[t(x_1 - x_o) + (1-t)\,(x_2 - x_o)]$

Usando la propiedad de linealidad de $f'(x_o)$ podemos escribir esta última igualdad de las maneras siguientes:

$$f(x_o) = f(x_o) + t\,f'(x_o)\,(x_1 - x_o) + (1-t)\,f'(x_o)(x_2 - x_o) + tf(x_o) - tf(x_o)$$

$$f(x_o) = t[f(x_o) + f'(x_o)(x_1 - x_o)] + (1-t)f'(x_o)(x_2 - x_o) + (1-t)f(x_o)$$

$$f(x_o) = t[f(x_o) + f'(x_o)(x_1 - x_o)] + (1-t)\,[f(x_o) + f'(x_o)(x_2 - x_o)].$$

Puesto que (17) es cierta para $x = x_1$ y $x = x_2$ tenemos que

$$f(x_o) \leq t[f(x_o) + f(x_1) - f(x_o)] + (1-t)[f(x_o) + f(x_2) - f(x_o)]$$

$$f(x_o) \leq t\,f(x_1) + (1-t)\,f(x_2) \qquad (20)$$

Así, queda probado que f es convexa.

Por último, si (17) es una desigualdad estricta, (20) también lo será.

2. Definición.

Sea $f: U \subseteq L \to \mathbb{R}$ una función y L un espacio vectorial normado. Decimos que f' es monótona creciente si para

$$x, y \in U,\ [f'(x) - f'(y)]\ [x-y] \geq 0 \qquad (21)$$

f' es estrictamente monótona creciente si la desigualdad (21) es estricta para $x \neq y$.

3. Proposición.

Sea $f: U \subseteq L \to \mathbb{R}$ una función continua y diferenciable sobre el conjunto convexo abierto U. f es convexa (estrictamente convexa) si y sólo si f' es monótona (estrictamente monótona) creciente sobre U.

Demostración.

Para una función convexa diferenciable sobre U, la prop. 1, página 167, nos permite afirmar que

$$f(x) - f(y) \geq f'(y)\ (x-y)$$

$$f(y) - f(x) \geq f'(x)\ (y-x)$$

Luego, $0 \geq f'(y)(x-y) + f'(x)(y-x)$

$$0 \leq f'(x)(x-y) - f'(y)(x-y)$$

$$0 \leq [f'(x) - f'(y)]\,[x-y]$$

y hemos probado la primera parte del teorema. Las desigualdades estrictas ocurren cuando f es estrictamente convexa.

Ahora, supongamos que f' es monótona creciente y definamos la función $\emptyset : [0,1] \to \mathbb{R}$, con $\emptyset(\lambda) = f[\lambda x + (1-\lambda)y]$. Mostraremos que $\emptyset$ es convexa.

Para $0 \leq \lambda_1 < \lambda_2 \leq 1$, sea $u_1 = \lambda_1 x + (1-\lambda_1)y$, y

$u_2 = \lambda_2 x + (1-\lambda_2)y$. Entonces,

$u_2-u_1 = \lambda_2 x+y-\lambda_2 y-\lambda_1 x - y + \lambda_1 y = \lambda_2(x-y) - \lambda_1(x-y) = (\lambda_2-\lambda_1)(x-y)$ y

$0 \leq [f'(u_2)-f'(u_1)]\,[u_2-u_1] = (\lambda_2-\lambda_1)[f'(u_2)-f'(u_1)](x-y)$ pues f' es monótona creciente y u_1, u_2 están en U. Luego,

$f'(u_1)\ (x - y) \leq f'(u_2)\ (x - y)$.

Por otra parte, $\emptyset'(\lambda_1)=f'[\lambda_1 x + (1-\lambda_1)y](x+0-y) = f'(u_1)(x-y)$

$\emptyset'(\lambda_2)=f'[\lambda_2 x + (1-\lambda_2)y](x+0-y) = f'(u_2)(x-y)$.

Por tanto, $\emptyset'(\lambda_1) \leq \emptyset'(\lambda_2)$; es decir, $\emptyset'$ es creciente.

Luego, $\emptyset$ es convexa. (estrictamente convexa si las desigualdades son estrictas).

Finalmente, tenemos $f[\lambda x + (1-\lambda)y] = \emptyset(\lambda) = \emptyset[\lambda(1) + (1-\lambda)\ 0]$

$\leq \lambda\ \emptyset(1) + (1-\lambda)\ \emptyset\ (0)$

$\leq \lambda\ f(x) + (1-\lambda)f(y)$; es decir, f es convexa.

4. Proposición.

Sea $U \subseteq L$ un conjunto convexo abierto del espacio vectorial normado L y $f\colon U \to \mathbb{R}$, una función continuamente diferenciable cuya segunda derivada existe en todo U. Entonces, f es convexa sobre U si y sólo si $f''(x)$ es "definida no negativa" para cada $x \in U$. Si $f''(x)$ es "definida positiva" sobre U; entonces, f es estrictamente convexa.

Demostración.

De acuerdo con la prop. 17.8, pág. 95 del capítulo I podemos afirmar que, para cualesquiera x, $x_0 \in U$

$f(x) = f(x_0) + f'(x_0)(h) + \frac{1}{2} f''(x_0 + sh)\ (h,h)$, con $s \in (0,1)$ y $h = x - x_0$.

Supongamos que $f''(x)$ es definida no negativa. Entonces, es inmediato que $f(x) \geq f(x_o) + f'(x_o)(x - x_o)$

$$f(x) - f(x_o) \geq f'(x_o)(x - x_o)$$

y f es convexa.

Recíprocamente, supongamos f convexa y definamos, para $x \in U$ y $h \in L$, la función $g: (a,b) \subseteq \mathbb{R} \to \mathbb{R}$, con $g(t) = f(x + th)$. Entonces, de acuerdo con la prop. 7, página 160, g es convexa en un vecindario del origen y

$$g'(t) = f'(x + th)(h)$$

$$g''(t) = f''(x + th)(h,h).$$

La convexidad de g implica que para cada $t \in (a,b)$, $g''(t) \geq 0$, en particular, que $g''(0) \geq 0$; así que, $f''(h)(h,h) \geq 0$. Puesto que $h \in L$ se tomó arbitrariamente, $f''(x)$ es definida no negativa.

BIBLIOGRAFIA

1. CONVEX FUNCTIONS.
   A. Wayne Roberts - Dale E. Varberg.

2. ECONOMIC THEORY AND MATHEMATICAL ECONOMICS.
   Erwin Klein.

3. PRINCIPIOS DE ANALISIS MATEMATICO.
   Walter Rudin.

4. NOTAS DE ANALISIS I
   Mauricio Marroquín Escoto.

5. ALGEBRA MODERNA.
   I. N. Herstein.

6. ADVANCED CALCULUS OF SEVERAL VARIABLES.
   C. H. Edwards, Jr.

7. ALGEBRA LINEAL.
   G. Hadley.

8. FUNDAMENTAL STRUCTURES OF ALGEBRA.
   G. D. Mostaw - J. H. Sampson. - J. P. Meyer.